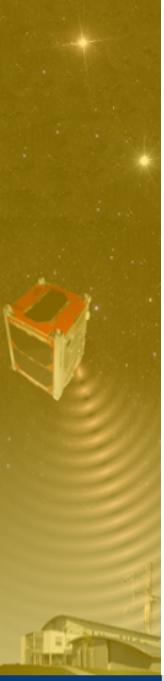


Institut für Informatik
Lehrstuhl für Robotik
Prof. Dr. A. Nüchter
Prof. Dr. K. Schilling

Forschungsberichte
in der Robotik

Research Notes
in Robotics

Julius-Maximilians-
UNIVERSITÄT
WÜRZBURG

Band 28

Dissertation an der Graduate School
of Science and Technology

Michael Bleier

Underwater Laser Scanning

Refractive Calibration, Self-calibration and Mapping for 3D Reconstruction

Julius-Maximilians-Universität Würzburg
Graduate School of Science and Technology

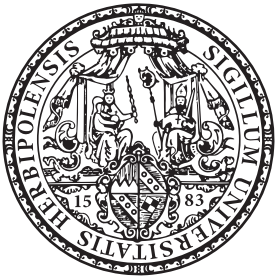
Doctoral thesis / ***Dissertation***
for the doctoral degree / *zur Erlangung des Doktorgrads*
Doctor rerum naturalium (Dr. rer. nat.)

Underwater Laser Scanning

Refractive Calibration, Self-calibration and Mapping for 3D Reconstruction

Laserscanning unter Wasser

Refraktive Kalibrierung, Selbstkalibrierung und Kartierung zur 3D Rekonstruktion



Submitted by / *Vorgelegt von*

Michael Bleier

from / *aus*
Bayreuth

Würzburg, 2023

Submitted on / *Eingereicht am:* 9 January 2023

Day of thesis defense / *Tag des Promotionskolloquiums:* 29 March 2023

Members of thesis committee / *Mitglieder des Promotionskomitees*

Chairperson / Vorsitz: Prof. Dr. Tobias Hertel

1. Reviewer and Examiner / 1. Gutachter und Prüfer: Prof. Dr. Andreas Nüchter

2. Examiner / 2. Prüfer: Prof. Dr. Sergio Montenegro

3. Reviewer and Examiner / 3. Gutachter und Prüfer: Prof. Dr. Stefan May

4. Reviewer and Examiner / 4. Gutachter und Prüfer: Prof. Dr. Thomas Luhmann

Additional Examiners / Weitere Prüfer: Prof. Dr. Mark Shortis

Abstract

There is great interest in affordable, precise and reliable metrology underwater: Archaeologists want to document artifacts in situ with high detail. In marine research, biologists require the tools to monitor coral growth, and geologists need recordings to model sediment transport. Furthermore, for offshore construction projects, maintenance and inspection millimeter-accurate measurements of defects and offshore structures are essential. While the process of digitizing individual objects and complete sites on land is well understood and standard methods, such as Structure from Motion or terrestrial laser scanning, are regularly applied, precise underwater surveying with high resolution is still a complex and difficult task. Applying optical scanning techniques in water is challenging due to reduced visibility caused by turbidity and light absorption. However, optical underwater scanners provide significant advantages in terms of achievable resolution and accuracy compared to acoustic systems.

This thesis proposes an underwater laser scanning system and the algorithms for creating dense and accurate 3D scans in water. It is based on laser triangulation, and the main optical components are an underwater camera and a cross-line laser projector. The prototype is configured with a motorized yaw axis for capturing scans from a tripod. Alternatively, it is mounted to a moving platform for mobile mapping. The main focus lies on the refractive calibration of the underwater camera and laser projector, the image processing and 3D reconstruction. For the highest accuracy, the refraction at the individual media interfaces must be taken into account. This is addressed by an optimization-based calibration framework using a physical-geometric camera model derived from an analytical formulation of a ray-tracing projection model. In addition to scanning underwater structures, this work presents the 3D acquisition of semi-submerged structures and the correction of refraction effects. As in-situ calibration in water is complex and time-consuming, the challenge of transferring an in-air scanner calibration to water without re-calibration is investigated, as well as self-calibration techniques for structured light.

The system was successfully deployed in various configurations for both static scanning and mobile mapping. An evaluation of the calibration and 3D reconstruction using reference objects and a comparison of free-form surfaces in clear water demonstrate the high accuracy potential in the range of one millimeter to less than one centimeter, depending on the measurement distance. Mobile underwater mapping and motion compensation based on visual-inertial odometry are demonstrated using a new optical underwater scanner based on fringe projection. Continuous registration of individual scans allows the acquisition of 3D models from an underwater vehicle. RGB images captured in parallel are used to create 3D point clouds of underwater scenes in full color. 3D maps are useful to the operator during the remote control of underwater vehicles and provide the building blocks to enable offshore inspection and surveying tasks. The advancing automation of the measurement technology will allow non-experts to use it, significantly reduce acquisition time and increase accuracy, making underwater metrology more cost-effective.

Zusammenfassung

Das Interesse an präziser, zuverlässiger und zugleich kostengünstiger Unterwassermesstechnik ist groß. Beispielsweise wollen Archäologen Artefakte in situ mit hoher Detailtreue dokumentieren und in der Meeresforschung benötigen Biologen Messwerkzeuge zur Beobachtung des Korallenwachstums. Auch Geologen sind auf Messdaten angewiesen, um Sedimenttransporte zu modellieren. Darüber hinaus ist für die Errichtung von Offshore-Bauwerken, sowie deren Wartung und Inspektion eine millimetergenaue Vermessung von vorhandenen Strukturen und Defekten unerlässlich. Während die Digitalisierung einzelner Objekte und ganzer Areale an Land gut erforscht ist und verschiedene Standardmethoden, wie zum Beispiel Structure from Motion oder terrestrisches Laserscanning, regelmäßig eingesetzt werden, ist die präzise und hochauflösende Unterwasservermessung nach wie vor eine komplexe und schwierige Aufgabe. Die Anwendung optischer Messtechnik im Wasser ist aufgrund der eingeschränkten Sichttiefe durch Trübung und Lichtabsorption eine Herausforderung. Optische Unterwasserscanner bieten jedoch Vorteile hinsichtlich der erreichbaren Auflösung und Genauigkeit gegenüber akustischen Systemen.

In dieser Arbeit werden ein Unterwasser-Laserscanning-System und die Algorithmen zur Erzeugung von 3D-Scans mit hoher Punktdichte im Wasser vorgestellt. Es basiert auf Lasertriangulation und die optischen Hauptkomponenten sind eine Unterwasserkamera und ein Kreuzlinienlaserprojektor. Das System ist mit einer motorisierten Drehachse ausgestattet, um Scans von einem Stativ aus aufzunehmen. Alternativ kann es von einer beweglichen Plattform aus für mobile Kartierung eingesetzt werden. Das Hauptaugenmerk liegt auf der refraktiven Kalibrierung der Unterwasserkamera und des Laserprojektors, der Bildverarbeitung und der 3D-Rekonstruktion. Um höchste Genauigkeit zu erreichen, muss die Brechung an den einzelnen Medienübergängen berücksichtigt werden. Dies wird durch ein physikalisch-geometrisches Kameramodell, das auf einer analytischen Beschreibung der Strahlenverfolgung basiert, und ein optimierungsbasiertes Kalibrierverfahren erreicht. Neben dem Scannen von Unterwasserstrukturen wird in dieser Arbeit auch die 3D-Erfassung von teilweise im Wasser befindlichen Strukturen und die Korrektur der dabei auftretenden Brechungseffekte vorgestellt. Da die Kalibrierung im Wasser komplex und zeitintensiv ist, wird die Übertragung einer Kalibrierung des Scanners in Luft auf die Bedingungen im Wasser ohne Neukalibrierung, sowie die Selbstkalibrierung für Lichtschnittverfahren untersucht.

Das System wurde in verschiedenen Konfigurationen sowohl für statisches Scannen als auch für die mobile Kartierung erfolgreich eingesetzt. Die Validierung der Kalibrierung und der 3D-Rekonstruktion anhand von Referenzobjekten und der Vergleich von Freiformflächen in klarem Wasser zeigen das hohe Genauigkeitspotenzial im Bereich von einem Millimeter bis weniger als einem Zentimeter in Abhängigkeit von der Messdistanz. Die mobile Unterwasserkartierung und Bewegungskompensation anhand visuell-inertialer Odometrie wird mit einem neuen optischen Unterwasserscanner auf Basis der Streifenprojektion demonstriert. Dabei ermöglicht die kontinuierliche Registrierung von Einzelscans die Erfassung von 3D-Modellen von einem Unterwasserfahrzeug aus. Mit Hilfe von parallel aufgenommenen RGB-Bildern werden dabei farbige 3D-Punktwolken der Unterwasserszenen erstellt. Diese 3D-Karten dienen beispielsweise dem Bediener bei der Fernsteuerung von Unterwasserfahrzeugen und bilden die Grundlage für Offshore-Inspektions- und Vermessungsaufgaben. Die fortschreitende Automatisierung der Messtechnik wird somit auch eine Verwendung durch Nichtfachleute ermöglichen und gleichzeitig die Erfas-

sungszeit erheblich verkürzen und die Genauigkeit verbessern, was die Vermessung im Wasser kostengünstiger und effizienter macht.

Danksagung

An dieser Stelle möchte ich die Gelegenheit nutzen, um mich bei allen zu bedanken, die mich während der Entstehung dieser Arbeit unterstützt und begleitet haben. Mein erster Dank gebührt meinem Doktorvater Prof. Dr. Andreas Nüchter für die hervorragende Betreuung und fortwährende Unterstützung. Über die Jahre durfte ich von seiner Erfahrung und Förderung Dinge weiter voranzutreiben profitieren. Mein ganz besonderer Dank gilt Prof. Dr.-Ing. Sergio Montenegro und Prof. Dr. Stefan May für die fachlichen Diskussionen, Unterstützung und Erstellung der Gutachten. Besonderer Dank gilt ebenfalls Prof. Dr.-Ing. habil. Dr. h.c. Thomas Luhmann für die schnelle Begutachtung meiner Dissertation, dem Vorsitzenden des Promotionskomitees Prof. Dr. Tobias Hertel, sowie Prof. Dr. Mark Shortis und Dr. Peter Janotta für ihr Mitwirken während der Disputation. Des Weiteren bedanke ich mich für die Unterstützung von Seiten der Graduate School Science and Technology, insbesondere durch Dr. Stephan Schröder-Köhne.

Außerdem möchte ich mich bei meinen vielen Kollegen an der Universität Würzburg, aber auch am Zentrum für Telematik e.V. bedanken. Besonderer Dank gilt meinem langjährigen Bürokollegen Joschka van der Lucht für alle Zusammenarbeit und gemeinsame Dienstreisen. Darüber hinaus geht großer Dank an Florian Leutert für die Zusammenarbeit, fachliche Diskussionen und Unterstützung bei Manipulatorexperimenten und an Sven Jörissen, Helge Andreas Lauterbach und Dorit Borrmann für das gemeinsame Meistern verschiedenster Herausforderungen in Projekten. Ebenfalls wesentlich danken möchte ich den ganzen wundervollen früheren und aktuellen Kolleginnen und Kollegen am Lehrstuhl für Robotik für den fachlichen Austausch, Zusammenarbeit, Unterstützung und gemeinsame verbrachte Zeit. An dieser Stelle also nochmal ein Dankeschön u.a. an Johannes Schauer Marin Rodrigues, Lakshminarasimhan Srinivasan, Sarah Menninger, Peter Janotta, Jasper Zevering, Yijun Yuan, Fabian Arzberger, Hermann Helgert, Christian Herrmann, Dieter Ziegler, Luca Anteunis, Maros Hladky und Heidi Frankenberger. Es war mir immer eine große Freude mit euch zusammenzuarbeiten.

Begonnen hat die Reise am Zentrum für Telematik e.V. bevor ich auf eine Stelle an der Universität Würzburg wechselte. Hier danke ich insbesondere Prof. Dr. Klaus Schilling, der mich nach Würzburg geholt hat und mir einen Einstieg in spannende Forschungsarbeiten ermöglichte. Des Weiteren danke ich der Mobile Systeme Gruppe und all den anderen herausragenden Kolleginnen und Kollegen. Stellvertretend stehen hier Robin Heß, Tobias Lindeholz, Florian Leutert, David Bohlig, Matthias Pröstler, Doris Aschenbrenner, Angel Martell, Bertram Koch und Philipp Wolf. Des Weiteren danke ich besonders Michael Albert und Kai Schwedhelm für die Versorgung mit Kaffee und Gesprächen, sowie Ulrike Göbel und Daniel Eck für die organisatorische Unterstützung am Zentrum für Telematik.

Über die Jahre ergab sich eine fruchtvolle Zusammenarbeit auch mit Partnern an anderen Universitäten, Instituten und Firmen. Ich danke im Speziellen allen Partnern der Projekte ¡VAMOS!, MUSCHEL-3D, TASTSINN-VR, LAUNCH-3D und UWSensor. Darüber hinaus danke ich Fraunhofer IOSB-AST in Ilmenau für die Unterstützung bei den Unterwassertests im UWSensor Projekt, sowie Fraunhofer IGP in Rostock und der IMAWIS GmbH aus dem Projekt „Offshore Wind Solution Mecklenburg-Vorpommern (2015 - 2018)“ für die Organisation und Unterstützung bei den Tests in der Schlepprinne am Lehrstuhl für Strömungsmechanik der Universität Rostock.

Abschließend möchte ich mich noch vor allem bei Jost Wittmann und Florian Leutert für ihre

wertvolle Unterstützung durch das Korrekturlesen der Arbeit bedanken. Des Weiteren danke ich Helge Andreas Lauterbach, Dorit Borrmann, Joschka van der Lucht, Julian Scharnagl und Sven Jörissen für ihre Anmerkungen.

Contents

Chapter 1

Introduction

While the process of digitizing individual objects and complete sites on land is well understood and standard methods, such as photogrammetry, Structure from Motion (SfM) or terrestrial laser scanning, are regularly applied, precise underwater surveying with high resolution is still a complex and difficult task.

On a global scale, we have only very low resolution maps of Earth's oceans and lakes. The ocean floor is less well mapped than the surfaces of celestial bodies, such as the Moon or Mars. Most of the Moon, as well as Mars, have been mapped with spatial resolution of much better than 1 km. The data of the Lunar Reconnaissance Orbiter has been processed to digital elevation maps of almost the complete lunar surface with a 100 m raster and a vertical accuracy of ca. 10 m [286]. On the other hand, available global bathymetric charts, such as the General Bathymetric Chart of the Oceans (GEBECO) [141], have a grid of 1 km and in most places the data is interpolated with the actual resolution being significantly lower [43]. The GEBECO is built from many different contributed archives of ship surveys and data from satellite missions. However, high-resolution bathymetry is only obtained from ship echosounding data. Therefore, more often traveled sea lanes are typically represented with higher accuracy. Only a small percentage of the oceans is covered by bathymetric surveys. Fig. 1.1 shows single beam and multibeam surveys in the archive of the National Oceanic and Atmospheric Administration (NOAA). Some smaller water bodies are mapped using bathymetric laser scanning and multibeam sonar with high resolution. For example, the 536 km^2 large Lake Constance in Germany has been mapped with point densities of up to 40 points per square meter during the project "Tiefenschärfe" [307].

Similarly, on a smaller scale creating 3D scans in water with high resolution and accuracy is challenging. Typically, sites in water are difficult to access, and deployment and recovery are time-consuming and expensive, especially for greater water depths. For many underwater mapping applications sonar technology is still the primary solution because of its large sensor range and its robustness to turbidity. However, specific measurement tasks require higher accuracies and resolutions. For example, archaeologists are interested in scanning and monitoring wooden structures, such as fragments of hull planks of ships, with millimeter resolution and accuracy to be able to investigate which parts fit together. Recovering these artifacts from the water and scanning them on land might lead to wrong conclusions because drying the wood causes deformations. Furthermore, there is also an interest in high-resolution underwater scanning for industrial

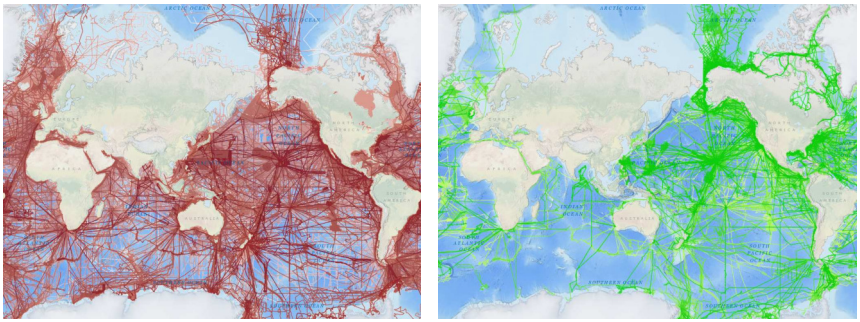


Fig. 1.1: Bathymetric surveys in the archive of the NOAA visualized on top of a rendering of the GEBECO (Source: [236]). Left: Single beam surveys displayed in red. Right: Multibeam surveys displayed in green.

applications, such as inspection of welding seams [114, 174] or metrology of pipelines [221] and other underwater structures.

1.1 Challenges of Optical Underwater 3D Sensing

High accuracy and resolution requirements make optical measurement technology interesting despite visibility limitations due to the high absorption of light in water and turbidity. However, sensors developed for applications in air are typically not directly applicable for underwater scanning. Due to the harsh environment, sensors must be appropriately protected from water ingress and pressure. Moreover, active optical sensors typically rely on lasers or Light-Emiting Diodes (LED) in the Infrared (IR) spectrum. IR lasers are desirable light sources because of cost-efficiency, high continuous power and pulse energy, low dispersion and invisibility to the human eye. However, application in water is mostly unfeasible due to the high absorption of the IR spectrum. Suitable underwater sensors also need to consider geometric effects, such as refraction at the water-glass-air interfaces and image degradation due to backscatter from particles in the water. This thesis concerns itself with the development of such a sensor system and associated image processing and precise calibration algorithms to increase robustness and accuracy.

One of the most reliable ways to calibrate optical sensors is using calibration artifacts with known geometry. Acquiring images of such a structure in water is time-consuming and tedious due to the high efforts of deploying and moving the sensor system or structure. If survey missions are carried out at significant water depths, for example, deeper than 1000 m, calibration stability is an issue. Under the extreme water pressure, the housings of the sensors, as well as the mounting structures, are elastically deformed. Underwater housings are not built completely rigid due to the rubber seals required between the aluminum or titanium housings and the glass window. Therefore, high pressure causes changes of the extrinsic and housing parameters. Moreover, the refractive index is affected by parameters, such as pressure, temperature and

soluted components. While changes in refractive index are typically small, they are relevant for applications that require high accuracy over large distances or hypersaline lakes. This makes self-calibration methods interesting that are applied in situ to compensate for errors. This thesis studies the application of geometric constraints for the estimation of calibration parameters of a camera and laser projector setup without the need for a special calibration structure or known scene geometry. Such methods are applicable for quality control and compensating errors due to parameter changes.

1.2 Industrial and Scientific Applications

The sea is becoming an increasingly important resource for the energy production industry and mining of raw materials. In order to ensure the safe and efficient operation of structures, such as offshore wind turbines, platforms for electrical substations or submarine pipelines, frequent inspections are necessary. High-resolution 3D scanners enable low-cost monitoring and dimensional control of such submerged structures. Additionally, every structure installed in the sea is manufactured on land. Thus, precise metrology is essential to ensure that pieces fit together, since adjustments are difficult to make once the structure is deployed in the water. For example, pipeline tie-in spools create a connection between subsea assets that allows for expansion due to the pressure and temperature of the transported fluid. To ensure correct alignment and a leak-tight connection, precise measurements of the relative positions and orientations of the connection points are necessary to fit a custom-built spool piece. Moreover, the spool is required to be built in a way that most of the weight lies on the seafloor, such that the mechanical stress on the interconnection points is minimized. The lifetime of the subsea structure is increased if the spool is properly anchored. Fig. 1.2 shows a subsea template, which is a large steel structure that carries various subsea structures, e.g., wellheads, subsea trees or manifolds. Here, photogrammetric targets are placed on the hub and scale bars on the structure for a metrology project to fit a spool between the hub of the template and a pipeline end termination. Accuracy requirements for spool metrology are very high. Typically, over a measurement distance of 100 m, the relative hub positions' accuracy must be within a few centimeters, and the angular errors are required to be less than half a degree.

Besides static subsea assets, there is also an interest in the inspection of moveable parts, such as floating platforms, cables or anchor chains. For the teleoperation of Remotely Operated Vehicles (ROV) and the automation of Autonomous Underwater Vehicles (AUV) 3D sensors play an important role in the perception of the environment and surveillance of operations [258, 261].

Moreover, there is a strong interest in underwater 3D data acquisition for scientific purposes, such as exploring ocean habitats. For example, marine biologists study the development of coral reefs as an important indicator of the well-being of the marine ecosystem. One aspect to characterize corals is structural complexity, which also varies between different species. Traditionally, this is measured by a diver using a 1 m long string. The string is wrapped over the corals and the 2D distance between the starting and the end point of the string is measured. The ratio between this shortened distance and the string length is used as an indicator of the structural complexity. More recently, marine biologists started to adopt SfM techniques to create 3D models of coral colonies [73]. This way, the marine habitat is digitized with high resolution for further

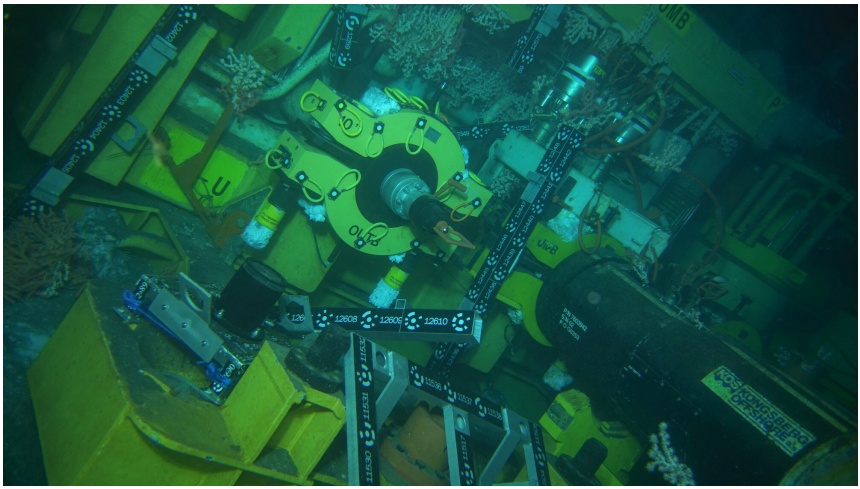


Fig. 1.2: Subsea template with photogrammetric targets and scale bars captured during a spool metrology project (Source: Equinor ASA and Jost Wittmann).

study, but the same indicators used before are also extracted from the 3D data. Accurately measuring corals' structural complexity gives scientists an indicator of the genetic diversity and spreading of diseases [74]. Besides corals, there are monitoring needs for other marine life, such as fish [272, 325]. Especially in aquaculture, the size of the fish, which indirectly provides the weight, and visual control of diseases, such as lice infestation, is important to dose food and medications correctly. This is challenging since it requires 3D metrology of a dynamic scene. Moreover, the aquaculture nets need to be inspected regularly for defects. Imaging sensors with automatic processing algorithms support doing this more efficiently.

Simple, low-cost 3D underwater sensors significantly impact our ability to study the oceans. This work looks at underwater mapping and 3D scanning in the context of 3D modeling of flooded mines and documentation of underwater assets.

1.2.1 Underwater Mining in the ¡VAMOS! Project

Part of the work in this thesis was carried out within the framework of the EU Horizon 2020 project Viable Alternative Mine Operating System (¡VAMOS!). The ¡VAMOS! project looked into alternative mining techniques to extract unexploited mineral resources in flooded inland mines that are not attractive from a conventional mining perspective in an environmentally and economically sustainable way. The approach has the potential to extend opencut mines with hydrological problems or allow to re-open mines that have been considered depleted in the past because with previous mining techniques it was not viable anymore to continue operations due to

Fig. 1.3: The ¡VAMOS! underwater mining system deployed during field trials at Lee Moor in Devon, United Kingdom.

technical, economic or environmental constraints. Today, with rising prices of certain rare ores, it might become interesting again to re-open abandoned mines in order to access deeper-seated minerals. However, conventional mining techniques require high treatment and dewatering costs. Moreover, from an environmental perspective, it is desirable that the water table of these flooded inland mines is not changed. The ¡VAMOS! project developed a prototype remotely controlled underwater mining machine and associated launch and recovery equipment and demonstrated the system in field trials in inactive mine sites. Fig. 1.3 shows the system deployed during field trials at the Lee Moor mine site in Devon, United Kingdom.

Excavation of raw materials in a water-filled open-pit mine requires a detailed mine model for teleoperation of the machine and planning of the mining operations. Therefore, part of the efforts in the project focused on positioning and navigation of the underwater assets and creating a detailed 3D terrain model. The combined data is presented to the operator in a virtual reality interface. The benefit of this approach in the ¡VAMOS! project is that the human operators gain a better situational overview and understanding of the mining operations, which assists remote control. Additionally, a full 3D model of the operations is valuable to monitor effectively what is happening below the water surface and communicate the status of mining operations. Moreover, it allows the use of a smaller and cheaper sensor kit since only the areas where change is expected need to be monitored regularly with surveying equipment. At the same time, the full context of the mine site is still visualized to the human operator.

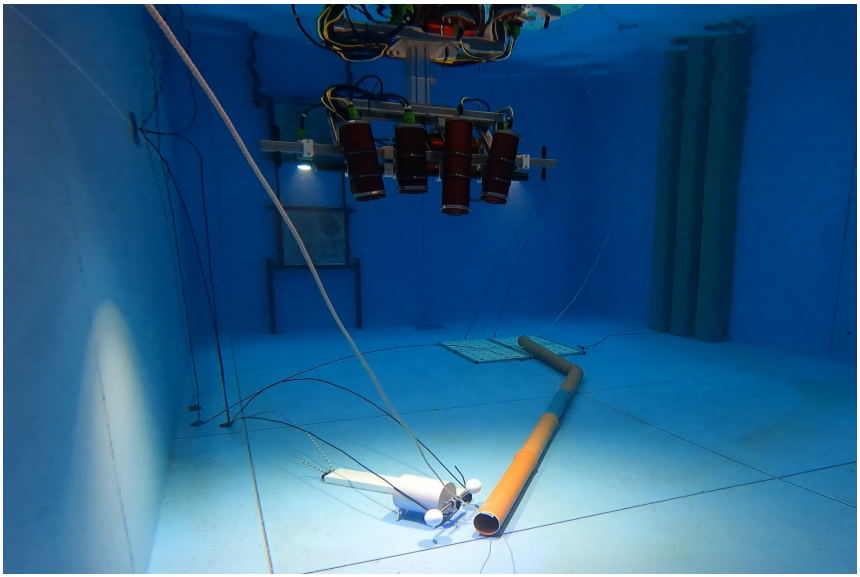


Fig. 1.4: UWSensor deployed in the test tank of Fraunhofer IOSB-AST in Ilmenau.

1.2.2 Offshore Metrology in the UWSensor Project

The UWSensor project developed a new 3D underwater sensor for mobile offshore surveying. The sensor is designed to be configured for operation by a ROV. The system is mounted rigidly on the frame of the ROV or to an actuator to control the viewing angle. The sensor operates while the ROV is in motion, and individual scans are fused together to expand the scan field continuously. The goal is a movement speed of the sensor of up to $1\,\frac{\mathrm{m}}{\mathrm{s}}$.

The 3D sensor system developed in the UWSensor research project is based on active stereo with pattern projection. In order to measure precisely during fast movements in the water and to register the individual 3D scans into maps, the movement must be compensated. This requires an accurate estimation of the sensor motion. There are several approaches to localization in an underwater environment. One of the difficult aspects of this effort is that approaches that work robustly on land cannot necessarily be implemented directly in underwater environments due to limited visibility conditions and media changes. Due to the nature of the sensor, camera-based approaches to underwater localization and methods based on the registration of 3D data are used individually or in combination. However, this task remains challenging in highly unstructured underwater environments.

Furthermore, the assumption of a static measurement environment is violated in the case of fast sensor movements since the employed 3D measurement method is based on multi-image

Fig. 1.5: Flooded mine shafts with historic timber installations.

evaluation, which is accompanied by a small temporal offset. These problems are addressed in this work by generating an initial trajectory based on visual odometry. Based on this trajectory, motion compensation for the 3D reconstruction is performed.

1.2.3 Documentation of Cultural Heritage

Cultural heritage sites are the legacy of the human past but are severely threatened by natural deterioration, man-made alteration, vandalism or robbery. In water, guarding and protecting archaeological structures and artifacts is challenging. Therefore, in addition to conservation and restoration efforts, many sites need precise digitization to preserve sensitive assets and their dissemination to the general public. Similarly to the rise of terrestrial laser scanning and structure from motion techniques for documentation on land, underwater archaeologists increasingly digitize and model sites in 3D. Smaller objects are often recovered and documented on land. However, this is not always feasible because bringing objects to the surface changes the finds or deteriorates the artifacts. Another aspect is that underwater archaeology is often time sensitive because of the high cost of vessels and equipment and restrictions on dive duration. Therefore, fast acquisition techniques are preferred in fieldwork. Hence, in the last ten years, methods for 3D reconstruction and recording are increasingly used in maritime archaeology [219]. Interesting archaeological structures, such as historic building remains, are also found in lakes. Fig. 1.5 shows flooded mine shafts with historic timbre installations in an inactive mine in the Harz in Germany, which was operated from the middle of the 16th century. The difficulty here is the access with scanning systems and the ability to carry heavy equipment to the site through narrow corridors of the mine.

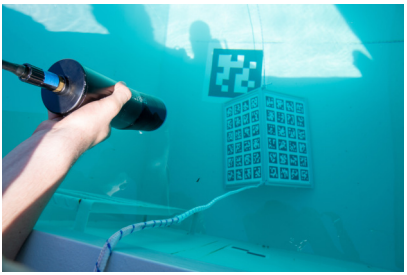
3D imaging approaches are not only applied for documentation purposes but also for monitoring sites, searching for archaeological traces and the study and analysis of the found evidence. Photogrammetric methods are regularly applied to underwater archaeological sites [222]. Larger areas are documented using sonar imaging. Unfortunately, the environmental conditions and requirements for each specific site vary. Therefore, no single sensor system or technique fits all sites and purposes.

1.3 Contributions

The accuracy of optical underwater sensors is negatively affected by imprecise physical-geometrical modeling of the light path and inaccurate estimates of the parameters of the refractive interfaces. Moreover, underwater imaging is heavily degraded by absorption, turbidity, dispersion and backscatter. Due to these factors, the achievable accuracy in water is reduced. In practice, reduced measurement quality, typically by a factor of two, is observed in the literature [205].

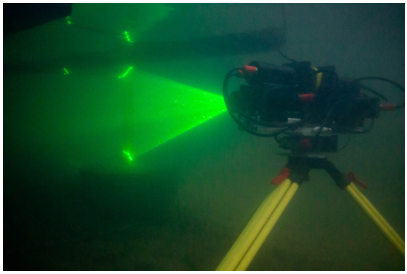
This thesis addresses this challenge by developing approaches for physical-geometric modeling of refraction and underwater calibration and techniques for self-calibrating structured light, which results in more robust and accurate results. Moreover, it describes the development of an underwater laser scanning system based on projecting a laser cross pattern for capturing dense 3D scans. Furthermore, mobile scanning using optical underwater sensors is demonstrated, and underwater mapping in the context of the teleoperation of an underwater mining vehicle is successfully applied. The contributions of this thesis are summarized in the following main topics:

Refractive camera calibration



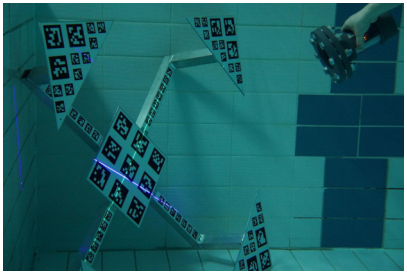
The challenge of modeling refraction for accurate 3D reconstruction in water is the estimation of the parameters and refractive surfaces. Available software packages for camera calibration and Bundle Adjustment (BA) do not provide multi-media projection models. Therefore, underwater camera models that precisely describe the light path for accurate 3D reconstruction in water, i.e., the estimation of the parameters and refractive surfaces, are still not widely available in the research community. Specific to the application for an underwater laser scanning system, a physical-geometric camera model using ray tracing is derived, and an optimization-based calibration framework is proposed (cf. Section 3.3, Section 3.4). Simulation shows that the Brown Model's approximation errors are large for underwater imaging, especially at close range. In simulation and experiments, the performance of the refractive calibration is validated, and it is shown that the proposed refractive calibration achieves lower residual errors compared to implicit modeling. Additionally, the effects of the test field and fiducial marker system on the calibration result are investigated. Based on the insights, a calibration approach using a 3D test field is developed to enable accurate calibration even if the images are degraded (cf. Section 3.5).

Underwater and multi-media laser scanning



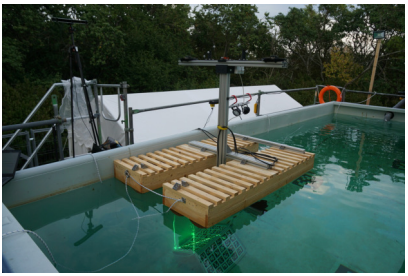
The development of an underwater laser scanning system and the algorithms for 3D reconstruction are described in [11]. The contributions include the hardware and data processing software development. A cross-line pattern is used to acquire redundant measurements and enable scanning while moving the scanner in different directions. Based on the refractive camera calibration and a model of the laser line distortion, improved accuracy is demonstrated (cf. Chapter 4). The system is evaluated and the accuracy is assessed in lab tests, in a towing tank, as well as in an underwater test environment [24]. The achieved accuracy is in the range of one millimeter up to below one centimeter, depending on the measurement range (cf. Section 4.6). In clear water, measurement distances of up to 8 m are demonstrated. A refractive correction approach for the system to enable scanning of semi-submerged objects is developed in [21–23]. The method's effectiveness in estimating the water plane and correcting the refraction effects is demonstrated in lab experiments (cf. Section 4.7).

Self-calibrating structured light



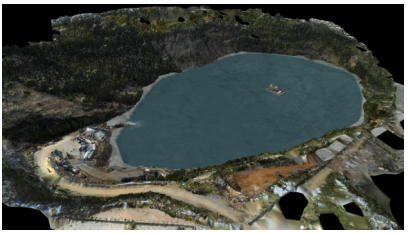
In-situ calibration of structured light scanners in underwater environments is time-consuming and complicated. Therefore, in [8] an algorithm for self-calibrating structured light using geometric constraints is developed. Coplanarities and orthogonality constraints are exploited to estimate the plane parameters of a cross-line laser projector (cf. Chapter 5). This is used to acquire 3D scans in air and in water using a single camera and a handheld laser projector. The approach is applicable without explicit calibration of the laser projector. Methods to estimate parameters in the presence of noise and making the approach more robust are presented in [9]. Some of the work on self-calibrating structured light was also transferred to medical applications [17].

Underwater mobile mapping



The developed underwater scanning system is extended for a mobile mapping application in [10, 12, 13]. Using Global Navigation Satellite System (GNSS) measurements combined with an Inertial Measurement Unit (IMU) based trajectory estimation system, 3D scans are acquired with the underwater scanner from a floating platform (cf. Section 6.1). The difficulty here is the co-calibration of the positioning system above-the-water and the underwater scanner, which is achieved by a joint optimization of the camera and GNSS trajectory. In [6, 7, 15, 16] a new optical underwater sensor system for mobile scanning is described. It has the advantage compared to the state-of-the-art in that it provides sub-millimeter accuracy and dense surface reconstruction at high data rates. Using Visual-Inertial Odometry (VIO) and continuous registration, mobile mapping of underwater structures is demonstrated (cf. Section 6.2). The proposed methods are transferable to other tasks, such as mobile laser scanning and personal mapping systems, which is demonstrated in [14, 18–20, 25].

3D underwater modelling for remote control



In [1, 3–5] applications of state-of-the-art techniques to 3D underwater modeling for remote control in the Horizon 2020 ¡VAMOS! project are demonstrated. From different acoustic sensors and a positioning system an underwater model of a flooded mine is built (cf. Section 7.2). In addition, an above-the-water model is built from terrestrial laser scanning. The different sensor measurements are fused in a consistent underwater terrain model using a Signed Distance Function (SDF) based map. In order to store large SDF voxel maps, a tree-based data structure is implemented. The data is presented to the operator of the mining vehicle using a Virtual Reality (VR) based user interface, which supports the remote control.

1.4 Outline

This thesis is structured as follows. Chapter 2 introduces the environmental and technical constraints for performing 3D data acquisition in the water and reviews the state-of-the-art of underwater acoustic and optical sensing modalities. Chapter 3 and Chapter 4 describe the development of an underwater laser scanning system as well as the needed algorithms for refractive calibration, image processing and 3D reconstruction. Moreover, methods for multi-media scanning of semi-submerged objects are introduced. Chapter 5 looks into self-calibrating structured light using geometric constraints. Approaches for mobile data acquisition using the proposed underwater scanning system are investigated in Chapter 6. Chapter 7 describes methods for building a 3D mine model using navigation data and sonar sensors for the support of the teleoperation of an underwater mining vehicle. Finally, Chapter 8 summarizes and discusses the main results and presents future work on open questions following this work.

Chapter 2

Underwater Metrology

Different sensing modalities and techniques are used to achieve underwater 3D acquisition and metrology tasks [79, 216, 222]. Depending on the environmental conditions and measurement goals, very different approaches have to be applied. Acoustic sensors are often used for imaging and bathymetric measurements because of the low absorption loss and large achievable measurement range. Acoustic signals are also used for underwater communication and distance measurements by measuring the travel time between acoustic transceivers. By building an array of acoustic transponders, which are, for example, statically installed using seabed tripods, positions are derived from the acoustic network solution based on multiple range measurements. Especially for long distances, the calibration and variation of the speed of sound in water affect the measurement accuracy.

High resolutions and accuracies are achieved using optical measurement and imaging techniques. Underwater laser scanners provide detailed scans of surfaces. Photogrammetric surveys using visual markers are one of the few approaches achieving millimeter-accurate metrology in water. However, visibility is a key limiting factor for cameras and optical sensors. High turbidity levels make optical imaging infeasible.

Similarly to tape measurements on land, taut wire systems are used to measure distances in water. The wire is rolled up on a winch, and it is measured how much wire is paid out. Relative positions are measured by recording the departure angle of the wire from the winch or through additional sensors, such as inclinometers or gyroscopes. Taut wires are also used for dynamic vessel positioning relative to an anchor. However, this metrology approach is only feasible for short distances since it is difficult to keep long wires under tension, and sagging of the wire due to the weight of the wire itself results in errors. Moreover, the measurements are affected by water currents, and it is difficult for the operator to control the measurement in low visibility. Elastic deformation of the wire, as well as temperature changes, also affect measurement accuracy.

While this works focuses on optical sensing in underwater environments, we also review acoustic sensors since acoustic sonar sensor are applied for the application of 3D modeling of submerged mines in Chapter 7. This chapter provides a brief overview of the relevant physical properties of the water for this work and introduces optical and acoustic underwater measurement techniques.

Fig. 2.1: Examples of underwater images. Left: Blue-green color cast visible in an uncorrected underwater image of a color checker photographed at a distance of approximately 1 m. Right: Scattered light of a green laser projector due to particles in the water.

2.1 Relevant Physical Properties of the Water

Working with optical sensors in underwater environments is constrained by the transmission properties of light in the water. Mainly, absorption and scattering limit the use of optical imaging sensors underwater. Absorption limits the range of sensors since light signals are increasingly attenuated with distance. Since attenuation is wavelength dependent, this also causes errors in color reproduction. For measurement purposes, the changes in direction of light rays due to reflection at a surface or refraction at a medium interface need to be considered.

2.1.1 Absorption and Scattering

Pure water is relatively transparent to visible light and near visible light, but it absorbs most of the ultraviolet, infrared and microwave spectrum. The attenuation of light is dependent on two processes: absorption and scattering. In the process of absorption, the light energy is transformed into different forms of internal energy of the medium. It occurs due to absorption by the water and other matter contained in the medium, such as dissolved salt in seawater, organic matter or other contaminants. Scattering is a divergence of the light path from a straight trajectory due to local non-uniformities, e.g., diffraction by small particles suspended in the water or refraction due to particulate matter with different refractive indices.

Typically, the total attenuation of wavelengths in the red spectrum is higher, which causes the blue-green color cast of underwater photos. The left image in Fig. 2.1 shows an underwater image of a color checker photographed at a distance of approximately 1 m in water. The image is brightened and scaled to correct for the camera sensor response, but no white balance or color correction was applied. The attenuation of the red color channel in water is visible in the image in the form of a blue-green color shift. The right image in Fig. 2.1 was captured in turbid water with a green laser line projector. Due to the scattering of the laser light in the water, the camera does not observe only the direct reflection of the laser projection at the surface. The whole image frame is illuminated by scattered laser light. Typically, two components are distinguished:

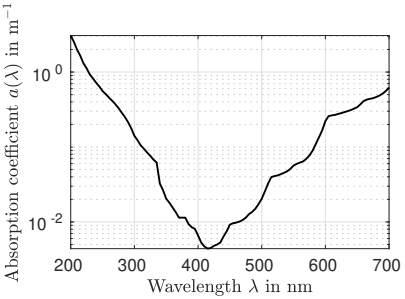


Fig. 2.2: Absorption coefficient $a(\lambda)$ of pure water based on the monograph [345].

forward scattering, denominating a light component with small angular changes of the direct light path, and backscatter, which is light that enters the camera without being reflected at the object but scattered within the volume of water.

Attenuation of Light

Using the total attenuation coefficient $c(\lambda)$, which describes the decay of light intensity per unit distance, the loss of irradiance is described by Lambert's law of absorption [169]:

$$E(r, \lambda) = E(o, \lambda) \mathrm{e}^{-c(\lambda) r} \quad , \tag{2.1}$$

where r is the distance, λ is the wavelength and E is the irradiance at positions o and r. The total attenuation coefficient is decomposed further as a sum of the absorption $a(\lambda)$ and scattering coefficients $b(\lambda)$:

$$c = a(\lambda) + b(\lambda) \quad , \tag{2.2}$$

where the total scattering coefficient $b(\lambda)$ is the integral of the volume scattering function $\beta(\lambda, \theta)$

$$b(\lambda) = 2\pi \int_0^{\pi} \beta(\lambda, \theta) \sin(\theta) \, \mathrm{d}\theta \quad . \tag{2.3}$$

The volume scattering function $\beta(\lambda, \theta)$ describes the dependence of the scattering effects on the angle θ relative to the light's direction of travel. Attenuation and scattering coefficients are not constant but change with the position in water due to local changes in the concentration of dissolved matter or density of suspended particles.

Fig. 2.2 shows the absorption coefficient for pure water depending on the wavelength. For typical wavelengths of blue, green and red lasers, we retrieve the absorption coefficients $a(480\mathrm{nm}) = 0.0127\mathrm{m}^{-1}$, $a(525\mathrm{nm}) = 0.0417\mathrm{m}^{-1}$ and $a(650\mathrm{nm}) = 0.34\mathrm{m}^{-1}$. This means that half of the initial irradiance is lost after 54.6 m for the blue laser, 16.6 m for the green laser and 2.0 m for the red laser. The absorption minimum is not necessarily in the blue wavelength range for natural water bodies. For example, high contents of dissolved organic matter, such as humus, peat

or decaying plant matter, result in an absorption minimum shifted more towards the yellow or orange wavelengths. Moreover, the scattering coefficient is higher for blue wavelengths.

Typically, blue or green laser sources are employed for underwater applications. This is not only a result of the absorption spectrum but also due to the availability of compact high-power laser sources with blue or green wavelengths.

Color Correction

The parameters that influence the color information measured by a camera are complex. The measured color depends on the illumination light source, the surface parameters of the object, the scene geometry and the optical transfer function of the camera [181]. In real-world scenes, this becomes more challenging because there are typically multiple light sources in a scene, e.g., sunlight and strobe light, and mixed illumination due to the interreflection of light between objects or shadows. Additionally, in underwater imaging, the medium plays an important role.

Color correction in air typically assumes some kind of constancy. For example, it is assumed that the scene features one dominant light source. The problem of color constancy is then expressed as the problem of estimating the light source's color. This is often solved by heuristics and constraints, physical models or machine learning [142]. Methods like Gray World or Gray Edge assume that the average surface reflectivity or edge response is constant to recover the scene illumination [71, 124, 323]. Moreover, not all theoretical possible light sources are observed in the real world. Therefore, constraints on the set of light sources and observed colors are sensible, e.g., as used in Gamut Mapping methods [122, 123, 129]. Physical methods for color correction, for example, exploit specularities in the image [315]. Machine learning methods often try to learn models for surface reflectivities and light sources [61, 134]. Therefore, sufficiently large training databases are necessary to generalize to various imaging conditions. Techniques for extracting and correcting for multiple light sources have been proposed in the literature [44, 143].

These methods are not directly applicable to underwater imaging because the underlying constancy assumptions are violated. Moreover, if part of the color spectrum is too strongly attenuated, it is challenging to reconstruct the complete color information. A method to describe the physics of light propagation in water is the Jaffe-McGlamery image model [169, 220]. It describes the image as a linear superposition of three components: a direct component, backscatter and forward scattering. However, this model cannot be directly inverted. Therefore, physical methods for underwater color correction rely on simplifying the Jaffe-McGlamery model for color restoration [346].

Recent methods for underwater color correction are based on machine learning algorithms using Convolutional Neural Networks (CNN). The problem is the lack of ground truth data for underwater color reproduction. Therefore, methods are often trained on artificially created databases. This is achieved, for example, by applying a physical underwater color model to images captured in air [333]. Machine learning is then used to learn a mapping, which inverts this process. Other approaches use machine learning to create artificial training data. These data sets are, for example, created by learning a mapping, which transforms a database of reference images in a way such that they are indistinguishable from a set of real underwater images. This idea of indirect training is used in Generative Adversarial Network (GAN) based methods for underwater color correction [117]. WaterGAN uses in-air RGB-D data sets as input to model

attenuation and backscattering more accurately based on the depth information [197].

Other methods rely on a set of filters to enhance colors of underwater images [27, 83, 164], such as contrast stretching, histogram modification or saturation and intensity equalization, or a combination of filtering and in air color correction methods [48]. Since the absorption of light is dependent on the distance of the object from the camera, depth maps from, e.g, SfM, are used in the literature to model color absorption more precisely [33, 327]. Often, color correction is combined with other image enhancements, such as increasing contrast, denoising and dehazing [328]. Overall, there is no single best approach for underwater color restoration because it is also dependent on the requirements of the application [211]. Some users want improved clarity in the image, others focus on visually pleasing results, while some applications require accurate color measurements.

2.1.2 Refraction

Typical optical sensors cannot be operated in water. Therefore, they need to be protected from water in an underwater housing with a view port. This means that light coming from the scene travels in water, enters the housing through a glass or acrylic window, and then travels in air before arriving at the actual sensor. At the water-glass, as well as at the glass-air interface, the light ray passes through a boundary between media with different optical densities. This causes a change in phase velocity, which results in a change of direction of the optical ray. The relationship between the angles of incidence and refraction is described by the law of refraction, also known as Snell's law.

Snell's Law

The law of refraction is derived, for example, from interference of all possible paths of light (Huygens–Fresnel principle) or by application of the general boundary conditions of the Maxwell equations for electromagnetic radiation. An intuitive derivation [171, pp. 40-44] stems from the application of Fermat's principle, which states that light will travel along the path which takes the least time [52, pp. 36-38]. The left illustration in Fig. 2.3 depicts the refraction of a light ray at a planar interface between two media of different refractive indices. The time t that is required for the light ray to travel from point $\mathbf{P}$ to point $\mathbf{Q}$ is computed by dividing the optical path length by the speed of light in the respective medium:

$$t = \frac{\sqrt{d^2 + x^2}}{v_1} + \frac{\sqrt{(l-d)^2 + (r-x)^2}}{v_2} \quad , \tag{2.4}$$

where v_i is the phase velocity in the respective medium and d,l,x,r denoted as in Fig. 2.3. Thus, if the travel time t is minimized, the distance x needs to satisfy [320]:

$$\frac{\mathrm{d}t}{\mathrm{d}x} = 0 = \frac{x}{v_1\sqrt{x^2 + d^2}} + \frac{-(r-x)}{v_2\sqrt{(r-x)^2 + (l-d)^2}} \quad . \tag{2.5}$$

Using the definition of the sine function, this results in

$$\frac{\mathrm{d}t}{\mathrm{d}x} = 0 = \frac{\sin\theta_1}{v_1} - \frac{\sin\theta_2}{v_2} \quad , \tag{2.6}$$

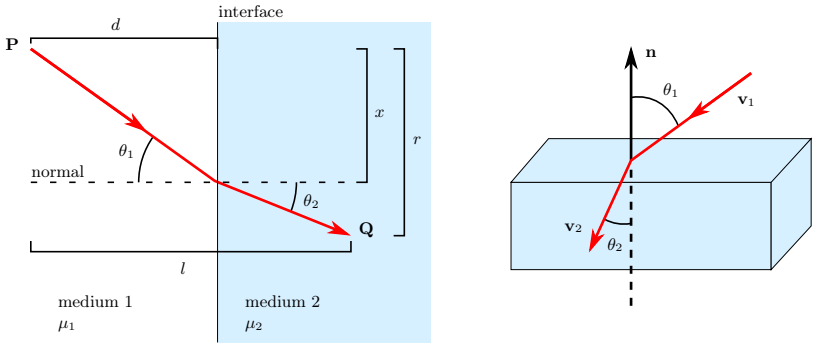


Fig. 2.3: Refraction of a light ray at a planar interface between two media of different refractive indices. Left: Planar case. Right: Refraction at an arbitrary interface plane with plane normal $\mathbf{n}$.

where θ_1 is the angle of incidence and θ_2 is the angle of refraction. By substituting the phase velocities in the two media with $v_i = \frac{c}{\mu_i}$, where μ_i is the refractive index of the respective medium and c is the speed of light in vacuum, we yield Snell's law:

$$\frac{\sin\theta_2}{\sin\theta_1} = \frac{v_2}{v_1} = \frac{\mu_1}{\mu_2} \quad . \tag{2.7}$$

If μ_1 is larger than μ_2, Eq. 2.7 may require for large angles of incidence that the sine of the angle of refraction is larger than 1. In this case, total internal reflection occurs. The largest angle where the incident ray is still refracted is computed by setting $\theta_2 = 90°$ in Eq. 2.7.

To apply the angular relationship of Snell's law in Eq. 2.7 to ray tracing, the incident and refracted ray need to be transformed to a common plane, which is perpendicular to the interface plane. Then Equation 2.7 is directly applicable. To simplify this computation for ray tracing a vector form of Snell's law is derived using the normal of the interface plane $\mathbf{n}$ and the normalized direction vector of the incident ray $\mathbf{v}_1$ and the refracted ray $\mathbf{v}_2$ as depicted in the right illustration of Fig. 2.3 [31, 320]. The normal vector $\mathbf{n}$ is chosen to point in the direction of the side where the incident ray is coming from, such that $\cos(\theta_1) = -\mathbf{n}^T\mathbf{v}_1$ is positive. Since the refracted ray $\mathbf{v}_2$ is coplanar with the incident ray $\mathbf{v}_1$ and the normal $\mathbf{n}$ of interface surface it is expressed as the linear combination

$$\mathbf{v}_2 = \alpha\mathbf{v}_1 + \beta\mathbf{n} \quad . \tag{2.8}$$

Using the physical law of Equation 2.7 as constraint α and β are computed as

$$\alpha = \frac{\mu_1}{\mu 2} \quad , \tag{2.9}$$

$$\beta = \frac{-\mu_1\mathbf{v}_1^T\mathbf{n} - \sqrt{\mu_1^2(\mathbf{v}_1^T\mathbf{n})^2 - (\mu_1^2 - \mu_2^2)(\mathbf{v}_1^T\mathbf{v}_1)}}{\mu_2} \quad . \tag{2.10}$$

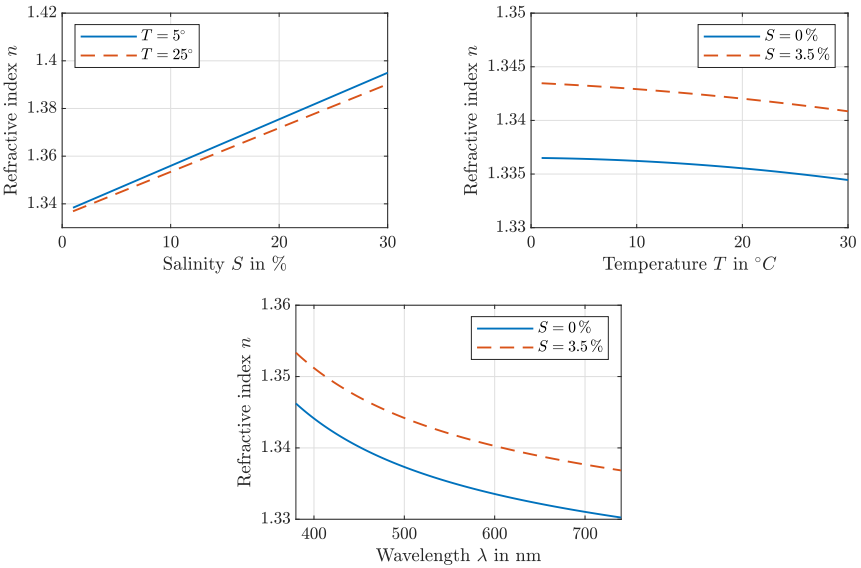


Fig. 2.4: Refractive index of seawater based on the empirical formulation in [265]. Top left: Refractive index for different salinity for a wavelength of 520 nm. Top right: Refractive index in relation to temperature for a wavelength of 520 nm. Bottom: Wavelength dependency of the refractive index at $T = 5°C$.

A derivation of the solution for α and β is found in [145, pp. 137-141]. Using the ratio of refractive indices $\mu_{1,2} = \frac{\mu_1}{\mu_2}$ and $c = \cos(\theta_1) = -\mathbf{n}^T\mathbf{v}_1$ this is expressed in a more compact form:

$$\mathbf{v}_2 = \text{RefractedRay}(\mathbf{v}_1, \mu_1, \mu_2) = \mu_{1,2}\mathbf{v}_1 + \left(\mu_{1,2}c - \sqrt{1 + \mu_{1,2}^2(c^2 - 1)}\right)\mathbf{n} \quad . \tag{2.11}$$

The refractive index of air is close to the refractive index of vacuum, which is 1 by definition. Commonly used silica glass or acrylic glass windows have a refractive index in the range of $1.4 - 1.5$. The refractive index of pure water is 1.336 at a temperature of $10°C$ and a wavelength of 520 nm, which is the primary laser wavelength used in this work. This changes to a refractive index of 1.343 for a salinity of 3.5%, which is the average salinity of seawater [265].

The main parameters for the change of refractive index in seawater are salinity, temperature, pressure and wavelength. The change due to pressure at deeper water depths is small. [265] reports a change in refractive index by $1.37 \cdot 10^{-6}$ per additional meter water depth in the top 100 m. Fig. 2.4 shows the refractive index change due to different parameters based on the empirical formulation in [265]. The top two figures show the dependence on salinity and temperature at a wavelength of 520 nm. The bottom figure shows the change in refractive index for the visible spectrum. The refractive index needs to be considered for the dominant wavelength

of the active illumination of the sensing system. High salinity strongly affects the refractive index. Hypersalient lakes, such as the Dead Sea, reach a salinity of more than 30%. Temperature effects on the refractive index are minor compared to the changes due to soluble contents.

Color Aberration

The refractive index of water varies with the wavelength of the light, see Fig. 2.4. Therefore, not all colors are projected on the same point on the image sensor, which is visible as color fringes, especially along edges. This effect is observed with glass lenses as well. Modern optics designs try to minimize the effect by using low-dispersion glass or adding correction elements to achieve lenses with approximately constant focal length for the entire visible wavelength range. However, the effect is very pronounced with underwater images, which might suggest that it is sensible to perform separate calibration for all color channels [226]. Strong chromatic aberration is observed for flat port cameras [158]. Color aberrations are reduced for housings with hemispherical dome ports [225].

2.1.3 Acoustic Properties

Acoustic waves are created by mechanical perturbation. An acoustic transducer creates local compressions and dilations. Due to the medium's elastic properties, these local perturbations propagate away from the source through the medium. Compared to air, water is an excellent conductor of sound. Moreover, acoustic signals propagate relatively favorably compared to optical signals in water. Acoustic signals are also affected by attenuation due to absorption of the acoustic waves. The direction of propagation is perturbed by inhomogeneities of the medium or reflections on the seafloor or sea surface. This creates additional echoes of a signal due to multi-path effects. If moving objects are measured, frequency changes due to the relative movement (Doppler effect) must be considered. Additionally, acoustic measurements are affected by ambient noise in the ocean, e.g., due to ship noise, seismic activity or rain. This background noise masks parts of the wanted acoustic signal.

Propagation Loss

The loss of intensity is described primarily by the geometric spreading losses and the absorption of the acoustic energy by the medium. Especially close to the water surfaces, air bubbles play an additional role. [203, pp. 18-26]

The energy of an acoustic wave created by a point source spreads in a homogeneous, infinite medium over a sphere with a growing radius. This results in a loss of the intensity I in $1/R^2$, where R is the radial distance from the source. Considering a two-way propagation, where the signal is transmitted from the sensor to the object and back, the loss becomes $I \sim 1/R^4$.

Additionally, the sound is attenuated due to dissipation through viscosity or chemical reaction. The attenuation of sound by water mainly depends on the sound frequency and density and viscosity of the water. Absorption for low frequencies is small. Therefore, frequencies below 30 kHz are typically used for long-range measurements of several kilometers. However, for shorter measurement ranges higher frequencies are desirable to achieve high resolution in compact sensor packages. The Francois-Garrison equation describes the sound absorption coefficient based on

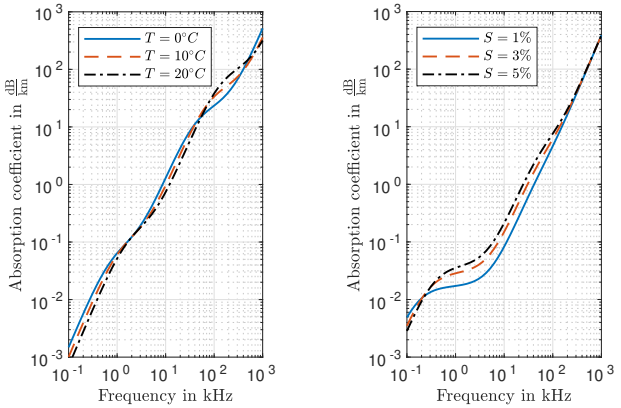


Fig. 2.5: Sound absorption coefficient depending on the frequency. Left: Absorption coefficient for different water temperatures. Right: Absorption coefficient for varying salinity. The plots are created for a water depth of zero meters and a pH of 8 based on the Francois-Garrison equation [130].

the temperature, salinity, pressure and sound frequency [130]. Fig. 2.5 shows the absorption coefficient depending on the sound frequency for varying temperature and salinity.

Speed of Sound

While the propagation rate of an acoustic wave through the air is about 343 $\frac{\text{m}}{\text{s}}$, the speed of sound in the oceans varies in the range of 1400 $\frac{\text{m}}{\text{s}}$ to 1600 $\frac{\text{m}}{\text{s}}$. It changes mainly with hydrostatic pressure, temperature and salinity. Especially near the water surface, larger changes in the speed of sound are observed due to the effects of temperature and salinity. Pressure becomes the main factor at greater depth since temperature and salinity vary less with increasing depth. The speed of sound changes by about 1.6 $\frac{\text{m}}{\text{s}}$ per 100 m water column and in the range of 2 $\frac{\text{m}}{\text{s}}$ to 4.5 $\frac{\text{m}}{\text{s}}$ per 1 °C temperature change depending on the temperature interval [97][p. 438]. Fig. 2.6 shows how temperature and pressure control the speed of sound based on the empirical formulation of [88]. The left graph shows the speed of sound for different temperatures at atmospheric pressure for pure and seawater. The right graph shows the speed of sound for seawater with a salinity of 3.5 % for different hydrostatic pressures.

Therefore, the calibration of the speed of sound is critical for acoustic ranging applications, such as sonar sensors or acoustic localization systems. For depth sounding, the change in speed of sound also needs to be considered for accurate water depth measurements.

Refraction and Scattering

Moreover, geometric effects at boundaries must be considered depending on the application. Similarly, Snell's law, as introduced in Section 2.1.2, applies to acoustic waves. Variations in the

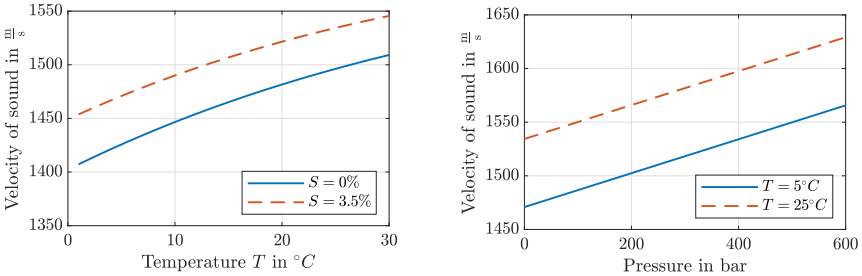


Fig. 2.6: Velocity of sound in water based on the empirical formulation of [88]. Left: Velocity of sound at atmospheric pressure in pure water and seawater for different temperatures. Right: Velocity of sound depended on hydrostatic pressure for seawater with a salinity of 3.5 %.

speed of sound cause refraction if the sound wave does not travel at a right angle to water layers with different temperatures or salinity. Scattering occurs at surfaces depending on the surface properties. An acoustic wave is scattered diffusely in random directions on rough surfaces. Volume scattering occurs from fluctuations of the medium and marine life or objects.

2.2 Optical Sensing

Passive optical methods, such as structure from motion or stereo vision, are regularly applied for 3D acquisition of underwater sites [222]. New commercial optical underwater scanning solutions are becoming available in the industry, as well as research projects demonstrating the use of optical scanners for tasks, such as underwater manipulation [257, 262] and inspection using underwater robots [277, 317]. A comprehensive overview of optical underwater sensing modalities is given in [216]. An overview of active optical 3D scanning systems is provided in [79].

2.2.1 Underwater Photogrammetry

Underwater photogrammetry using coded targets is often used for subsea metrology tasks that require high precision, such as measurement of welding seams or metrology of the relative poses of connection points for subsea installations. However, this approach requires the deployment of a large number of calibrated frames with optical targets and markers in the measurement volume using a remotely operated vehicle. Nevertheless, this method enables the relative pose measurements of subsea assets over large distances of more than 100 m with small distance and angular deviations.

The process for this subsea metrology technique is the following: First, rigid frames with coded circle targets are placed along a path between the two subsea assets using a ROV as depicted in the top row images of Fig. 2.7. This is an involved process since the frames have to be deployed to the seabed in a container, and the individual frames have to be placed using the remotely operated arm of the ROV. These frames have been calibrated before deployment

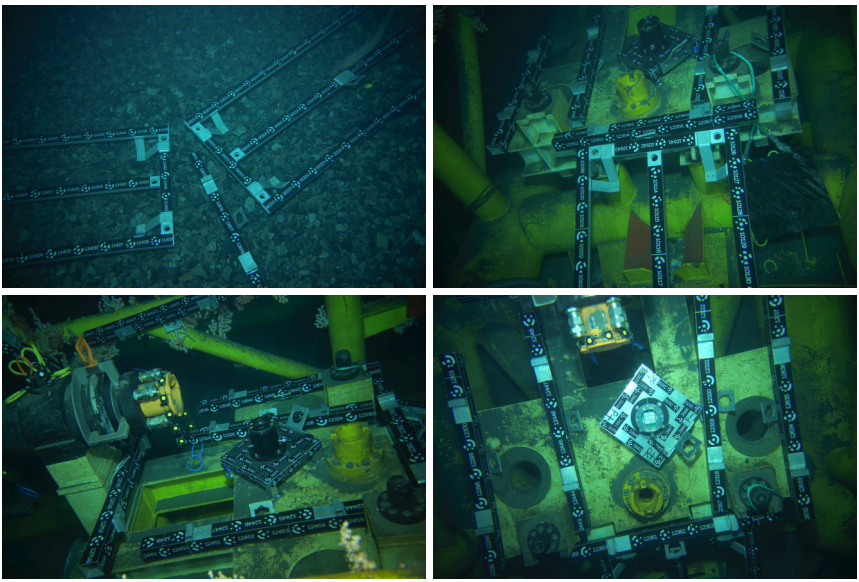


Fig. 2.7: Underwater photogrammetry using coded circle targets for the relative distance and orientation measurement of subsea assets (Source: Equinor ASA and Jost Wittmann).

in the air, such that the distances of all coded circle targets on each individual frame are known. If the pose of a hub or pipeline segment is required to be measured, additional circle targets are placed on the hub, as is visible in the bottom left image of Fig. 2.7. Then, many images with high overlap are captured of the placed targets along the path between the two connection points. Typically, the measurement camera is placed in the ROV arm for the image acquisition to capture multiple passes with different angles. This is necessary to achieve sufficient parallax between the camera poses for accurate depth triangulation. The point set is finally reconstructed offline. The image coordinates with associated Identification Number (ID) are extracted from the images mostly automatically using automatic marker detection algorithms. Point-matching between images is performed based on the ID. The multi-view reconstruction is then computed using BA software. The constraints of the known distances of the individually calibrated frames are used to scale the result accurately.

For leveling the created point set, inclinometers are used as depicted in the bottom right image of Fig. 2.7. The inclinometer is mounted and aligned to a planar base plate, which is marked with coded circle targets. This way, the plane of the inclinometer is computed from the reconstructed point set. Additionally, the inclinometer has a digital display, which shows the measured roll and pitch angle, which is manually read from the image and used to level the

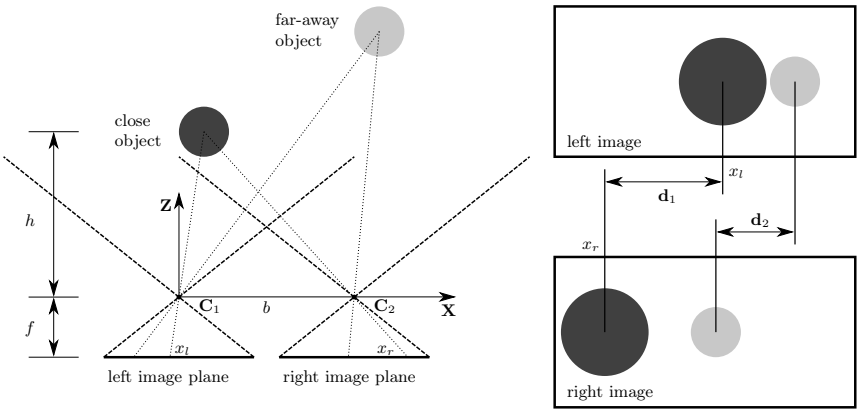


Fig. 2.8: Geometry of stereo vision. The observed disparity depends on the object's distance from the stereo sensor. A close object (dark grey) yields a larger disparity $\mathbf{d}_1$ compared to the disparity $\mathbf{d}_2$ of a far-away object (light gray).

result. Since the base plate is relatively small, these angle measurements are only applied locally. The hub pose or inclinometer surface are extracted by fitting cylinder or plane models to the reconstructed point set.

2.2.2 Multi-view Stereo and Structure from Motion

Passive stereo techniques [47] are regularly applied underwater for diverse tasks, such as fish size measurements [153, 297], monitoring of marine fauna [298] or interaction between divers and an underwater robot [146]. Depth from stereo is extracted by comparing two images captured from different viewpoints as visualized in Fig. 2.8. Objects observed at different distances are projected on different positions in the right and left camera images. The observed displacement of the object points in the two images, which is called disparity, depends on the object's distance from the stereo sensor. Looking at Fig. 2.8, we observe that a close object (dark grey) yields a larger disparity $\mathbf{d}_1$ compared to the disparity $\mathbf{d}_2$ of a far-away object (light gray). To compute 3D information from disparity, it is necessary to find the corresponding object points in the left and in the right camera image. Stereo matching between the two images is done using key points extracted using feature detectors or pixel-wise by comparison of the images using, for example, the Semi-Global Matching (SGM) method [160]. Objects without any texture, e.g., homogeneously colored surfaces without any visible surface structure, cannot be reconstructed from stereo matching. Moreover, low-light environments, changing illumination and harsh shadows are challenging for passive stereo techniques.

Based on triangulation the depth is reconstructed from two images of the same scene for each corresponding point. Looking at the dark gray circle in Fig. 2.8, the X- and Y- coordinates

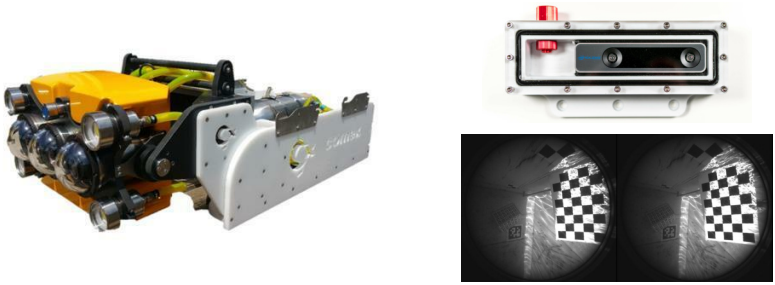


Fig. 2.9: Multi-camera sensors. Left: COMEX ORUS 3D system mounted to a ROV skid (Source: [229]). Right: Intel T265 tracking stereo camera in an underwater housing and fisheye images captured underwater by the T265.

parallel to the image plane are determined by

$$X = \frac{h}{f} x_l \quad , \qquad Y = \frac{h}{f} y_l \quad , \tag{2.12}$$

where x_l and y_l describe the position of the projection of the object point on the image plane, f is the focal length and h the distance in Z-direction of the object from the optical center $\mathbf{C}_1$. In the direction perpendicular to the image sensor, applies

$$\frac{h}{f} = \frac{b}{x_l - x_r} \quad , \tag{2.13}$$

which yields the depth

$$Z = h = \frac{bf}{x_l - x_r} = \frac{bf}{d_1} \quad , \tag{2.14}$$

where $d_1 = x_l - x_r$ is the measured disparity and b is the baseline distance between the two cameras. [202, pp. 341f]

Commercially, multi-camera systems find increasing adoption. For example, the COMEX ORUS 3D [229] underwater measurement system features three cameras with dome ports and lighting. The system is typically used from an ROV for subsea inspection. An acoustic distance sensor is installed on the sensor to support the ROV operator to keep optimal distance to the measurement object during data acquisition. The sensor unit is tilted in its mounting frame to achieve different camera angles. The data is typically post-processed but also allows real-time processing [224]. Fig. 2.9 shows on the left the COMEX ORUS 3D multi-camera system mounted to a ROV skid. The right image shows a low-cost Intel T265 stereo camera in a custom underwater housing and example images captured underwater for calibration [2].

Refraction typically has to be explicitly considered for underwater stereo geometry. The radial shift in the image caused by refraction at a flat port housing is distance-dependent. Therefore, a point in the left image corresponds to an epipolar line in the right image that is not straight anymore, as visualized in Fig. 2.10. This complicates dense image matching because either the

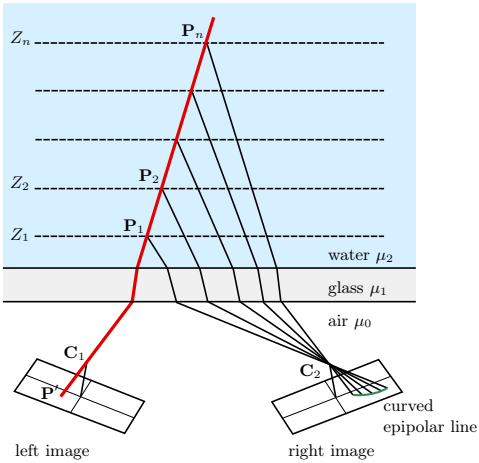


Fig. 2.10: Epipolar geometry in underwater imaging with flat port housings. Due to the distance-dependent radial shift caused by refraction, a point in the left image corresponds to a curve in the right image. Image reproduced based on [205].

images have to be normalized as a pre-processing step or the search must be performed along the curved epipolar line. If the camera centers are close to the window and the cameras are oriented perpendicular to the planar air-glass and glass-water interface, low errors are achieved using the Brown Model for calibration. However, for applications where the cameras are tilted relative to the medium interface plane, the refraction needs to be modeled explicitly for accurate measurements. [176, 205]

Structure from motion, is regularly applied for 3D acquisition of underwater sites [171, 219], e.g., for reconstruction and monitoring of coral reefs [73, 247, 248, 276] and underwater archeology [105, 106, 108]. Techniques to create metric point clouds from dense image matching are scaling with laser points as reference [166] or alignment with IMU measurements.

2.2.3 Pattern and Fringe Projection Systems

Different variants of structured light scanning have been successfully applied for underwater 3D scanning. Fringe projection has been applied successfully to acquire very detailed scans with high precision of small underwater objects [319]. Despite the limited illuminating power of standard digital projectors, working distances of more than 1 m have been reported in clear water [68]. The active stereo setup of [68] is depicted in the top left image of Fig. 2.11. [62] developed a diver-operated scanning system based on fringe projection, which is shown in the top right image of Fig. 2.11. Similarly to passive stereo vision, ray-based modeling and calibration are applied to compensate for refraction effects [63]. [271] proposes an adaptive phase-unwrapping

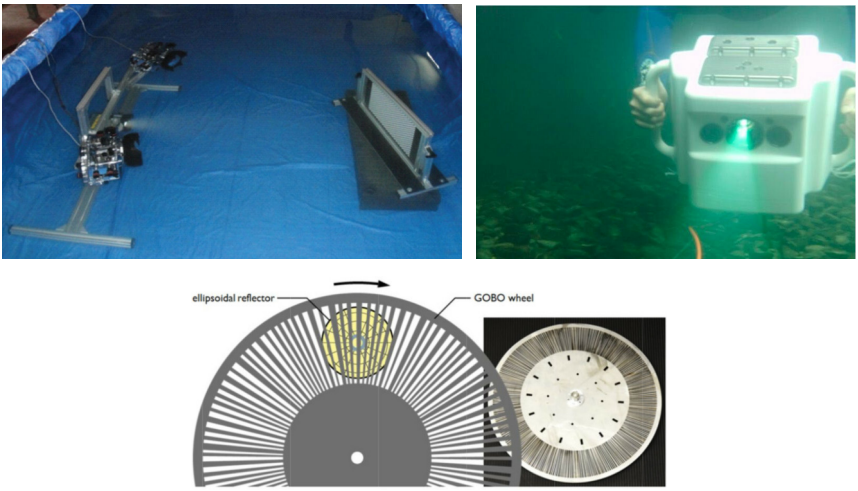


Fig. 2.11: Examples of pattern projections systems. Top left: Active stereo setup with a video projector (Source: [68]). Top right: Diver-operated scanner developed by Fraunhofer IOF (Source: [62]). Bottom: GOBO projector tested for underwater measurements (Source: [189]).

procedure and filters and models to make fringe projection more robust to scattering effects observed underwater.

The popularity of RGB-D cameras also inspired researchers to apply them to underwater imaging [40, 99, 182]. However, due to the infrared wavelength of the laser pattern projector, the achievable range is limited to less than 30 cm [94].

[189] employs the principle of GOBO projection for underwater 3D measurements. A rotating slide is placed in front of the optics of the illumination source to project aperiodic sinusoidal fringes as depicted in the bottom image of Fig. 2.11. [230] uses a Diffractive Optical Element (DOE) to project a static wave pattern for one shot 3D reconstruction.

2.2.4 Laser Line Scanning

Underwater laser scanning systems with larger measurement ranges often employ high-power line laser projectors. For example, commercial scanners from Voyis, formerly 2G Robotics, offer a range of up to 10 m depending on the water conditions [144, 329]. 3D scans are typically created by rotating the scanner and measuring the movement using rotational encoders or mounting the scanner to a moving platform. The top left image of Fig. 2.12 shows the Voyis Insight Micro system. It includes a RGB camera and high power LED flash to capture colored point clouds.

[86] and [258] use laser projectors based on galvanometer scanners. This allows to project sweeping laser lines on the scene, which enables 3D reconstruction for the entire field of view of

Fig. 2.12: Examples of different underwater laser line scanning systems and Time-of-Flight (ToF) laser scanners. Top left: Voyis Insight Micro / 2G Robotics ULS-500 Micro (Source: [329]). Top right: Newton Labs M4000UW underwater laser scanner (Source: [240]). Middle left: University of Girona underwater laser scanner with galvanometer projector (Source: [256]). Middle right: Kraken Robotics SeaVision Subsea 3D Laser imaging system with tri-color lasers [109, 186]. Bottom left: 3D at Depth SL1 subsea Light Detection and Ranging (LIDAR) (Source: [150]). Bottom right: Mitsubishi underwater laser scanner U4LE (Source: [163]).

the camera. An example of a galvanometer-based scanner [256] is depicted in the middle left image of Fig. 2.12. In this configuration, the laser hits the air-glass and glass-water interface surface at an angle. Hence, the laser projection cannot be described by a plane in water. This needs to be modeled explicitly using a physical refraction model. For example, Palomer et al. calibrate the parameters of a cone model to describe the laser projection surface [256]. [80] uses a biaxial Microelectromechanical System (MEMS) mirror to steer the laser beam.

Another approach is to mount line laser projectors inside a glass cylinder and rotate the line lasers with a motor. This approach is, for example, used in the UX-1 underwater mine exploration robotic system [213]. In this case, the line projectors are aligned manually, such that the projection is perpendicular to the air-glass and glass-water interface surfaces, and a straight line is projected. This way, the refraction effects are relatively small and are neglected depending on the accuracy requirements. Similarly, the Newton Labs M4000UW scanner, which is depicted in the top right image of Fig. 2.12, uses a rotating line laser projector. The SeaVision subsea 3D laser imaging system developed by Kraken Robotics [186] employs rotating red, green and blue line lasers as depicted in the middle right image of Fig. 2.12. This enables the system to produce colored scans by evaluating the intensity of the responses of the different laser projectors [109]. Alternatively, DOEs are used to project multiple lines for one shot 3D reconstruction [215]. Calibration of these systems is typically achieved using chessboard patterns or 3D calibration fixtures.

2.2.5 Time of Flight Sensors

Most commercially available underwater laser range sensors are based on laser stripe projection or other forms of structured light. More recently, companies started the development of ToF underwater laser scanners. For example, the company "3D at Depth" developed a commercial subsea LIDAR system, which is mounted on a pan-and-tilt unit to create 3D scans of underwater environments similar to terrestrial pulsed ToF laser scanning [26, 221]. The sensor is depicted in the bottom left image in Fig. 2.12. Combined with acoustic positioning, the subsea sensor has also been applied to dynamic scanning [90].

The Mitsubishi underwater laser scanner U4LE, depicted in the bottom right image of Fig. 2.12, uses a dome port with the scanner aligned in the optical center to achieve a wider field of view [163].

Stemmler et al. use an 8-sided-polygon for steering the laser beam to achieve high line scan rates of up to 800 Hz and a lateral resolution of 4 mm [310]. To increase dynamic range and suppress signal from backscatter at particles close to the scanner, they apply a time-dependent detector gain. Objects farther from the scanner experience a higher gain. In experiments, a standard deviation of 6 mm and a peak-to-peak deviation of less than 4 cm was measured by fitting a plane to a flat surface. [85] applies filtering to separate backscatter noise from the signal.

Range-gated imaging is used to increase visibility and minimize the effects of backscatter [190, 314]. The scene is illuminated with a laser pulse, illuminating objects in the water at different times. By capturing an image with a high-speed synchronized camera shutter, it is possible to expose the sensor only with the light returned from a certain depth range. For example, if the camera is closed for about 40 ns after the laser pulse is emitted, only objects farther than about

6 m are detected. Since most of the energy from backscatter comes from suspended particles close to the sensor, this is effectively suppressed and the Signal-to-noise Ratio (SNR) is improved. This way, high-contrast images are created that exclude most of the backscatter from particles in the water. By sweeping the range gate and capturing multiple images with different exposure windows, 3D information is reconstructed [155, 212]. For example, the UTOFIA camera [273] uses a CMOS ToF sensor that is able to capture images at a framerate of 1000 Hz and shift the opening of the camera shutter in increments of 1.67 ns in relation to the laser source. From an image set typically sampled for 25 different ranges, 3D data with a framerate of up to 10 Hz is reconstructed.

2.2.6 Airborne Laserscanning and Imaging

Airborne Laser Scanning (ALS) has been successfully applied to bathymetric mapping of flat water zones of lakes and archaeological documentation of underwater structures [101, 102, 209, 210, 288, 308]. Depending on the turbidity of the water body, penetration depths of multiple meters are achieved. The measured depths are corrected based on the information retrieved from the reflected signals. This is especially challenging for open water bodies with surface waves [339].

Other approaches estimate water depths from aerial images using multi-view reconstruction [131]. Dense image matching is applicable in favorable conditions if sufficient texture of the bottom of the water body is captured [208]. The water depth of through-water SfM models needs to be corrected based on an estimate of the water surface [98].

2.3 Acoustic Sensing

Acoustic signals are used for many different applications since they are transmitted with smaller losses compared to electromagnetic signals in water. This ranges from hydrophones, which are used to listen to marine life or noise produced by vessels, to Acoustic Doppler Current Profiler (ADCP) that measure water current velocities using the Doppler effect of sound waves scattered back from suspended particles. For underwater operated vehicles speed estimation using Doppler Velocity Log (DVL) is regularly used in the sensor suite for navigation systems. Moreover, acoustic signals are used for communication in acoustic modems.

Sonar sensors, named from sound navigation and ranging, are the main perception sensors for large distances and in turbid water conditions. Sonars are categorized in active and passive sonars. Passive sonars only listen to the sounds created by the environment. Using an array of sonar transducers the position of the sound source is triangulated and databases of known sounds, e.g., acoustic signatures of different ship motors or marine life, are used to identity the source. Active sonars emit acoustic signals themselves and analyze the echo of these signals. The emitted signals are typically acoustic pulses or chirp signals, which alternate in frequency to enable analyzing correlation between emitted and received signal. Acoustic transmitters and receivers are integrated into a single sensor unit or spread apart. In towed array sonar the hydrophones are towed behind the vessel on a cable with lengths of up to kilometers. This way the sonar sensors are far away from the noise generated by the vessel, which improves the SNR. Since low-frequency acoustic signal also penetrate the upper layers of the ocean bottom sonar

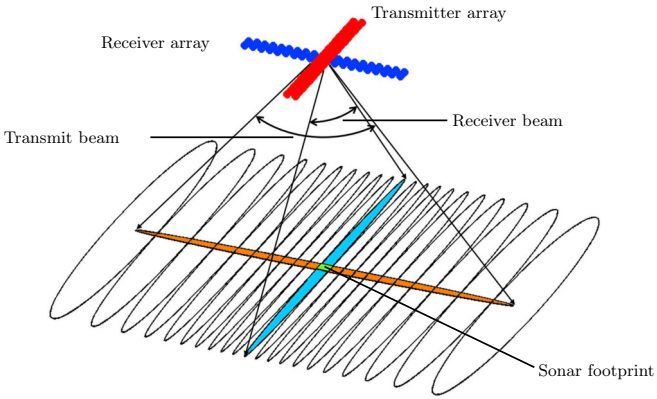


Fig. 2.13: The sonar footprint of a MBS is formed by the intersection of transmit and receiver beam (Source: [173]).

is also applicable for sub-bottom profiling. For imaging and 3D measurements typically active sonars are applied since knowing the emitted signal allows to directly apply ToF methods.

2.3.1 Active Sonar Sensors

Active sonars are often characterized by application, e.g., forward-looking sonar, Side-scan Sonar (SSS), the operation mode, such as imaging sonar or profiling sonar, and the number of acoustic beams, e.g., Single Beam Sonar (SBS) and Multibeam Sonar (MBS). SBS are regularly applied as altimeters or for obstacle detection. By placing the SBS on a pan-and-tilt head, 3D swaths are captured, which is also sometimes referred to as mechanically scanning sonar [351]. Multibeam echosounders often feature operation modes for application as imaging sonar as well as profiling sonar.

SSS or side imaging sonar captures a 2D intensity image of the seafloor. The sonar device illuminates the bottom of the sea typically with a fan-shaped pulse created using beamforming. The acoustic reflections are recorded based on the arrival time in a series of cross-track slices. This way, by moving the sonar, a 2D image of the seafloor along the direction of motion is created. While SSS does not directly deliver 3D information, there are techniques to infer geometry from imaging sonar similar to SfM techniques for optical cameras [89]. The main difference is the angular ambiguity of sonar images, which needs to be solved [340].

MBS similarly emit acoustic waves in a fan shape that is wide across-track and narrow along-track. This is achieved using beamforming by an acoustic transmitter array. The distance information is computed by measuring the ToF of the sound wave reflected of the seabed. Using an array of acoustic receivers, multiple receive beams are formed across-track. By evaluating the difference in time of arrival of the returned signals, it is possible to extract the direction

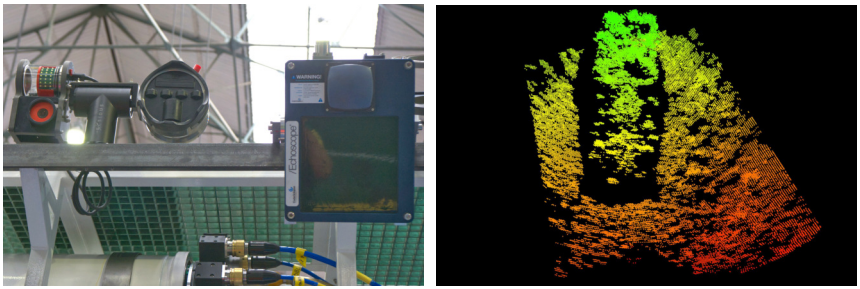


Fig. 2.14: Acoustic cameras capture dense 3D data. Left: Kongsberg M3 Multibeam sonar and 3D sonar Coda Octopus Echoscope mounted on the ¡VAMOS! mining vehicle. Right: Dense point cloud of a car on the lakebed captured by the 3D sonar at a distance of 5 m to 7 m.

information. Assuming the known speed of sound in the water column, a 2D profile, using the distance and direction of the return signals, is created. The sonar footprint results as the strip of the ocean floor, which is ensonified by the acoustic transmitter and the strip of the ocean observed by the receiver, see Fig. 2.13.

More recently, 3D sonars were developed that employ specialized acoustic transducers, beamforming and data processing to create dense point clouds of a volume in front of the sensor in real-time. For example, the Coda Octopus Echoscope is able to capture a Field of View (FoV) of $55° \times 55°$ at 20 Hz. Different FoV and update rates are configurable. Fig. 2.14 shows in the left image a Kongsberg M3 MBS and the 3D sonar Coda Octopus Echoscope mounted to the ¡VAMOS! mining vehicle. The MBS is placed with a camera on a pan-and-tilt to capture a large volume. In the right image, the point cloud of a car on the seafloor captured by the 3D sonar at a distance of approximately 5 m to 7 m is visualized.

Extracting accurate 2D profiles from MBS data is challenging due to multiple returns of acoustic signals. There is often ambiguity about which of the returns, e.g., first return, last return or strongest return, belongs to the surface that is measured. Hence, techniques for multibeam data processing have been developed, such as bottom following, which assumes that there are no large range discontinuities while profiling the seafloor. Moreover, registration of individual sonar scans is challenging due to high noise. Therefore, there is an interest in registration techniques that are more robust to noisy point clouds, e.g., spectral registration methods [72].

2.3.2 Acoustic Modems

Acoustic modems transmit data at low rates using acoustic signals. If the speed of sound is known, acoustic modems compute the distance between sender and receiver by measuring the time of flight of the acoustic signal. This is possible by using synchronized clocks between the two modems and measuring the time of transmission and time of arrival. Another possibility is to measure the round-trip time of one packet. If the receiver answers to an incoming packet within a predictable time, it is also possible to compute the time of flight from the round-trip

time of a data packet. By using an array of transducers and evaluating the difference in arrival time of an acoustic signal, the position of the sound source is triangulated.

2.4 Combination of Acoustic and Optical Modalities

Increasingly, different modalities, such as optical and sonar imaging, are used in combination to leverage the advantages of the individual methods. This enables creating more detailed and complete 3D underwater surveys [119]. A common application is the colorization of sonar bathymetry from camera images, which requires co-calibration of sonar and optical sensors [162]. Nevertheless, techniques have also been developed to directly fuse 3D optical mapping with sonar bathymetry [301]. This allows to exploit and combine the benefits of the large measurement range of sonar sensors with the detail and color information of optical mapping to create higher quality and more gap-free 3D models. Moreover, a combination of optical information and sonar information is used to improve registration [177].

Bathymetric mapping using multibeam sonar is often challenging in flat water zones with less than three meters of depth. Depending on the sonar sensor, the minimum measurement distance of a multibeam sonar sensor is typically in the range 0.5 m to 3 m. Moreover, sonar measurements below 1 m are challenging due to multipath effects and reverberation. Therefore, for mapping of lakes or river beds, there is a growing interest in combining multibeam sonar with airborne lidar bathymetry. For example, the most current bathymetry of the 536 km^2 large Lake Constance in Germany was conducted by a combination of multibeam sonar and airborne LIDAR [307]. All parts with a water depth larger than 5 m were mapped using multibeam sonar, while the seabed in the flat water zone was captured from the air with LIDAR.

Recently, Fitzpatrick et al. demonstrated in the lab a measurement technique for photoacoustic imaging [126] through the water surface. They fire high power 1070 nm laser pulses with peak power in the kilowatt range at the water surface. Due to thermoelastic expansion, an acoustic wave is generated, which travels in the water similar to sonar. The acoustic wave is then reflected off the targets and transmitted through the water-air interface. In the air, the acoustic wave is picked up by Capacitive Micromachined Ultrasonic Transducers (CMUT). From the captured signal, an image of the targets in the water is reconstructed. While this proof-of-concept device is not applicable in the field due to challenges with signal strength and water surface waves, it is a step towards future airborne acoustic imaging systems.

AUV often carry a wide range of different optical and acoustic sensor payloads [34, 213]. Sensor fusion leverages the combination of these sensors to estimate the vehicle trajectory more accurately and hence create more accurate maps [300]. Docking of AUVs often includes a multi-step process where every step relies on different sensors as the main source of information. For example, acoustic sensors are used for coarse localization to navigate the underwater vehicle close to the docking structure. Once the AUV is able to detect the structure in the camera images, docking control switches to a more precise localization based on relative pose estimation using the optical information [259]. Optical fiducials, such as coded tags, are often placed on the docking structure to assist pose estimation [84]. In close vicinity to the structure, electromagnetic sensors are applicable for navigation [59].

Chapter 3

Refractive Camera Calibration

The view port of the underwater housing acts like an optical element. Due to the change in refractive index between water, glass and air, refraction occurs, see Section 2.1.2. This changes the geometry of the light path, which results in a distorted image. Hence, for accurate measurements the physical-geometric effects of the housing need to be considered. Underwater cameras typically employ one of two types of cover glass in the housing: dome ports and flat ports.

Dome ports are built from a hemispherical glass shell and aim to minimize the geometric changes of the light path. If the camera center is aligned with the center of the dome port, the ray from the object point $\mathbf{P}$ through the center of projection $\mathbf{C}$ to the image point $\mathbf{P}'$ is perpendicular to the spherical media interfaces. Therefore, no refraction occurs and a near-perfect central projection is provided, which is visualized in Fig. 3.1.

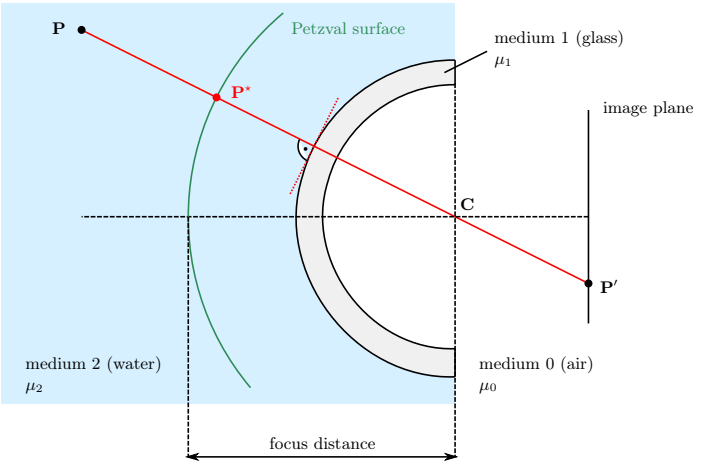


Fig. 3.1: A dome port preserves the field of view. Image reproduced based on [202, p. 419].

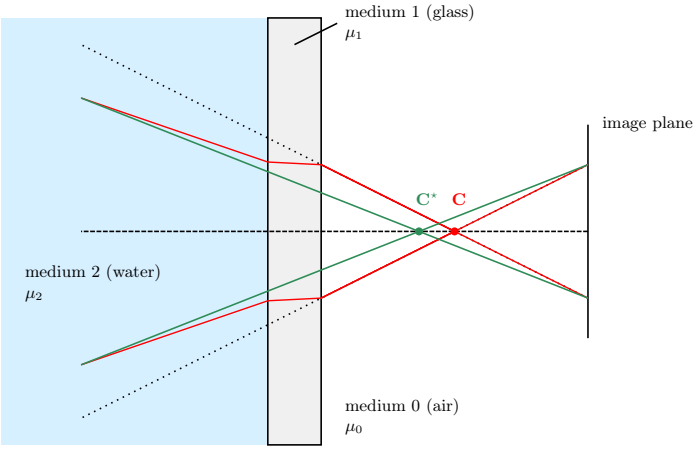


Fig. 3.2: A flat port with plane media interfaces causes a change in apparent field of view and focal length.

The advantage of the dome port is that the field of view and image distortions are unchanged if the camera is centered. However, the depth of field increases slightly and the focus distance is different compared to air. The spherical interface acts similarly to a negative lens element. Hence, the objects captured through the dome port appear to be closer. The object point $\mathbf{P}$ is projected onto a virtual image lying on a spherical surface called the Petzval surface. Therefore, the camera has to be focused on the virtual image $\mathbf{P}^\star$ [202, p. 419]. Wide angle lenses especially benefit from dome ports since the FoV is conserved and distortion is minimized. The main disadvantage of dome ports is that they are significantly more expensive to manufacture. Some residual errors and distortions will occur if the camera center is not perfectly aligned with the sphere center or due to manufacturing imperfections of the spherical dome. These errors need to be modeled and compensated with, for example, radial and tangential distortion models.

In contrast, a flat port causes a change in the apparent field of view and focal length. This is depicted in Fig. 3.2. The light rays visualized in red are refracted at the water-glass and glass-air interface. If the light path is approximated with a pinhole model, which is visualized in green, this leads to a reduction of field of view and a larger effective focal length. Moreover, the apparent camera center $\mathbf{C}^\star$ is shifted towards the media interface, which leads to a larger principal distance compared to the real camera center $\mathbf{C}$. In general, the image distortion is dependent on the object distance and cannot be compensated exactly by a power series distortion model. Engineered correction lens designs for flat ports have been developed that compensate for the refraction effects and reduce image distortions [167]. For flat port underwater cameras, objects appear closer to the camera. The focus distance is reduced by approximately the reciprocal of the refractive index of water. Color aberrations are stronger for flat ports compared to dome ports and astigmatism effects are observed [225].

Fig. 3.3: Dome port (top left) and flat port (top right) underwater camera. Below are two example images with the dome port (bottom left) and flat port (bottom right) partially submerged in water. The top half of the image is in air, while the bottom half is in water.

Examples of a dome port and flat port underwater housing are shown in Fig. 3.3. The top left image shows a dome port housing that is rated for a depth of 750 m. The flat port housing depicted in the top right images features a 19 mm thick glass window and a depth rating of 1750 m. In the two housings, two industrial machine vision cameras are integrated. Below the housings, example images of the two underwater cameras are shown. The images are captured with the dome port and flat port partially submerged in water.

The top half of the images is in air, while the lower half is in water. The dome port image shows little geometric distortion of the captured chessboard, which suggests that the camera is well-centered in the hemispherical glass window. The sharpness of the image is noticeably different for the part of the image in air and water since the camera needs to be focused at different distances. Here, the camera-lens system was focused for the in-water conditions. The bottom right image visualizes the different magnifications of the flat port interface in air and in water. The chessboard appears closer in the lower half of the image captured in water. The magnification and focal length of the combined camera and housing system are increased.

Both flat ports and dome ports have a role in underwater imaging depending on the application and the design of the imaging system. The main differences are summarized in Tab. 3.1. Flat port housings are typically more cost-effective to manufacture. Therefore, in this chapter

mainly plane media interfaces are examined. Calibration of the camera system is necessary to determine model parameters that precisely describe the internal geometric characteristics of the camera, lens optics and housing.

Table 3.1: Different properties of dome and flat port housings [225, 245] [202, p.419f].

	Dome Port	**Flat Port**
Description	Camera looking through a concentric, hemispherical glass lens	Camera looking through a plane parallel media interfaces
Field of view	Equal if camera is centered, focal length and magnification are preserved	Reduced in water, increased focal length and magnification, limited maximum field of view
Geometric lens distortion	Lens distortion is unchanged if the camera is centered • A shift of the perspective center perpendicular to the optical axis is modeled by tangential distortion • A shift along the optical axis is modeled by radial distortion	Pincushion distortion due to refraction • Significant distortion towards the edges of the field of view • Radial distortion due to refraction is distance dependent • Misalignment of media interface causes non-symmetric distortion
Focus distance	A smaller virtual image is formed and is projected onto a spherical surface	Objects appear closer by approximately the reciprocal of the refraction index of water
Other effects on the image	• Chromatic aberration is unchanged if the camera is centered • Increased Depth of field by the ratio of refractive indices • Spherical aberration may increase • Field curvature	• Strong chromatic aberration • Astigmatism
Practical aspects	• Prone to reflections • Easily scratched • Good mechanical properties for high pressure	• Prone to fracture due to uneven distribution of mechanical stress • Thick glass ports necessary for extreme pressure
System costs	More expensive to manufacture	Low-cost housings available

3.1 Modeling the Geometric Effects of Underwater Imaging

The refractive effects of water are reduced using specialized optics and underwater housing designs, such as the Rebikoff-Ivanoff correction lens [167,266]. There are also industrial underwater cameras, where the front element of the lens system is in direct contact with the seawater without an additional view port, e.g., ZEISS DUW DISTAGON. However, for high accuracy, calibration is still necessary. For accurate underwater metrology using cameras, the refractive effects need to be modeled implicitly or explicitly [296]. The approaches described in the literature are differentiated by the employed calibration approach, type of view port and camera model. Typically, only implicit models are applicable for both flat port and dome ports because physical camera models are specific to the housing. Based on how the camera models describe the refraction effects, the literature can be divided into the following categories:

1. *Implicit modeling* of refraction using a standard pinhole model with distortion. The camera is typically calibrated in situ, and the refraction effects are absorbed by the distortion parameters, such as Brown's distortion model [66, 67, 91].

2. *Explicit modeling using ray tracing.* These approaches try to find a physically correct model of the path of the light and model the refraction at each interface layer. This means that the parameters of the model have a physical-geometric interpretation, such as the refractive indices of the media or the geometry of the interfaces.

3. *Alternative camera models*, such as generic camera models that describe the ray direction for each pixel, camera models with non-single viewpoint or axial cameras. These methods neglect some of the physical-geometric terms and describe them by generic calibration parameters.

Tab. 3.2 lists a selection of literature for the different categories. Comprehensive literature reviews of different underwater camera models are found in [100, 171, 289, 295, 296].

Table 3.2: Selection of literature on underwater camera calibration.

Authors	View Port	Description
Implicit Modeling		
Letherdale and Turner (1983) [194]	dome	Calibration of 70 mm cameras in a test tank using a 3D structure built from a solid plate with variable length posts.
Fryer and Fraser (1986) [133]	flat	Underwater calibration using plumb-line test frame.
Harvey and Shortis (1995) [153]	flat	Calibration of stereo-rig using 3D calibration structure.
Drap et al. (2007) [108]	dome	Underwater self-calibration using markers on the seabed.

Menna et al. (2016, 2017) [223, 225, 226]	flat/dome	Characterization and calibration of dome port. Comparison of residual errors between flat and dome port using Brown parameters.
Menna et al. (2020) [228]	flat/dome	Improves results of BA using radial weighting of image observations and correction of systematic residuals using look-up tables.
Nocerino et al. (2021) [246]	flat	Introduction of an additional polynomial model in BA to compensate distance-dependent distortion caused by refraction.
Explicit Modeling		
Zaar (1948) [348], Rinner (1948) [270]	flat/dome	Early work on physical models for two-media photogrammetry.
Maas (1992) [204]	flat	Physical model for a stereo rig. Calibration using 3D structured with circle markers in different planes.
Li et al. (1997) [198]	dome	Geometric correction for concave lens covers. A two-step calibration process is applied. First, an air calibration is performed. Then, the housing parameters and refractive indices are calibrated in water.
Agrawal et al. [31]	flat	Solve multi-layer flat refractive geometry using polynomial approach.
Jordt and Koch (2012,2013) [172, 289]	flat/dome	Error comparison of different underwater camera models. Calibration using a synthesis of chessboard images.
Yau et al. (2013) [347]	flat	Considers dispersion for calibration of flat refractive layers.
Mulsow and Maas (2014) [231]	flat	Integration of ray tracing into BA.
Duda and Gaudig (2016) [111], Duda (2020) [109]	flat	Linearization of refractive forward projection using Taylor expansion for each observation.
Łuczyński et al. (2017) [354]	flat	In air calibration of camera and compensation using housing parameters in water. Refracted rays are saved in a map for computationally efficient look-up.
She et al. (2019, 2022) [292, 293]	dome	Alignment method for dome port. Calibration of a decentered dome port.
Rofallski et al. (2022) [274], Kahmen et al. (2019,2020) [175, 176]	flat	Flat port calibration using ray tracing model. The error function is minimized in object space to improve computation speed.

Rofallski et al. (2022) [275]	dome	Dome port calibration using ray tracing model with BA. Orthogonal line-point distance in object space is minimized to avoid solving forward projection of 3D points.
Alternative Models		
Telem and Filin (2010) [316]	flat	Modeling of planar housing using a perspective-center offset.
Treibitz et al. (2011) [320]	flat	Calibration of non-single viewpoint model using ray map to describe flat refractive geometry.
Agrawal et al. [31]	flat	Approximation of flat refractive geometry by an axial camera for calibration of refractive layer parameters.
Bräuer-Burchardt et al. (2015, 2020, 2022) [63, 64, 70]	flat	Calibration in air combined with underwater parameters. Extension of pinhole model for underwater applications using variable projection centers.
Chadebecq et al. (2017) [82]	flat	Generalized epipolar constraints for an axial camera for SfM.

3.1.1 Implicit Modeling

In the literature, explicit models for relative orientation for two-media photogrammetry have been developed [185, 291]. However, the currently available commercial and open source BA software packages do not implement physical multimedia models. Therefore, it is interesting to apply the standard physical parameter sets to underwater imaging. This way, no change to the calibration algorithm and method used in air is necessary, which allows the application of common workflows, such as commercial SfM software packages. The idea is that the pinhole, radial and tangential distortion parameters absorb the refraction effects. Depending on the underwater imaging system, very good results are achieved [295]. However, some residual errors remain that are not modeled accurately.

In [225] it is found that the residual systematic errors are smaller using implicit modeling for dome ports, even if decentered, compared to flat port cameras. For flat port cameras, the radial effects are theoretically compensated by the Brown model if the projection center lies on the planar air-water interface. In practice, the errors are small if the camera is aligned and the following conditions are met [176]: (1) The image plane is parallel to the air-glass and glass-water interface of the flat port. (2) The object point in the water is far away compared to the distance between the camera center and the planar view port.

For hemispherical ports, the error is reduced by careful alignment of the camera and dome center [174, 293]. Systematic errors are reduced by empirical weighting [228] or introducing distant-dependent distortion models [246]. An error analysis of implicit modeling is found in [176, 178, 289].

3.1.2 Explicit Modeling

Explicit modeling describes the geometry of the optical path as accurately as possible [172, 270, 289, 348]. This correction approach is more complex because it needs to model the refraction at each interface surface. Therefore, the solution is typically specific to the optical path and type of view port. The methods consider explicitly flat ports [31, 109, 111, 175, 176, 204, 231, 274, 347] or dome ports [198, 275, 292, 293]. Typically, ray tracing through the refractive interfaces is employed. An advantage of explicit modeling approaches is that the parameters have a physical interpretation, such as the refractive index of water or the thickness of the glass window.

In theory, this approach fully corrects the refraction effects. However, it is difficult to accurately estimate the physical parameters, e.g., the surface of the lens cover. Some parameters are also difficult to estimate stably from image observations since they are correlated with intrinsic camera parameters, such as focal length, or parameter variations have a small influence.

Explicit modeling requires the implementation of the physical-geometric model in the application. In order to include explicit modeling in standard SfM pipelines, some methods use an iterative refinement, e.g., by applying corrections to the input images [302, 304].

3.1.3 Alternative Models

Alternative models also aim to correct for the refraction effects fully. These models typically simplify the ray-based geometry or neglect some of the physical-geometric terms and describe them by generic calibration parameters [320]. For flat ports, the refraction effects are modeled by a shift of the perspective center [63, 70, 316]. Often, axial cameras are used as an approximation of the geometric effects [31, 82]. Instead of a central projection center, an axial camera defines a line, the camera axis, which cuts all of the projection rays.

Recently, learning-based correction models have been applied for aerial bathymetry [29, 30]. These methods try to correct the input images for refraction effects to enable processing in standard SfM pipelines.

3.2 Perspective Camera Model and Lens Distortions

Camera models describe the mapping of a 3D point in object space to a 2D point on the image plane. Here, finite projective camera models are considered, which are applicable to typical machine vision cameras. The basic model of a finite projective camera is the ideal perspective camera model or pinhole model [152, p. 153ff].

3.2.1 Pinhole Camera Model

The geometry of the pinhole model is depicted in Fig. 3.4. The center of projection $\mathbf{C}$ is also called the camera center or optical center and is the origin of the Euclidean camera coordinate system. Here, the image plane is defined by $Z = f$, where f is the focal length. The point $\mathbf{X} = (X, Y, Z)^T$ in world coordinates is projected on the image plane according to

$$(X, Y, Z)^T \longmapsto (fX/Z + p_x, fY/Z + p_y)^T = (x, y)^T \ , \tag{3.1}$$

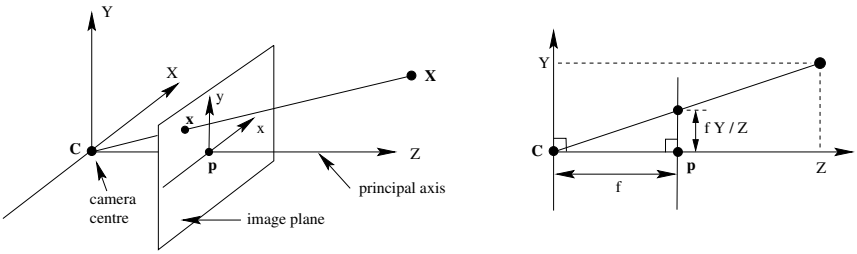


Fig. 3.4: Pinhole camera model (Source: [152, p. 154]).

where $\mathbf{x} = (x, y)^T$ are the image coordinates of the projection and $(p_x, p_y)^T$ are the coordinates of the principal point $\mathbf{p}$ in the image coordinate system.

Using homogeneous coordinates $\mathbf{X}_H = (X, Y, Z, 1)^T$ this is expressed as

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \rightarrow \begin{pmatrix} fX + Zp_x \\ fY + Zp_y \\ Z \end{pmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{I}_{3\times3} & \mathbf{0}_{3\times1} \end{bmatrix} \mathbf{X}_H \quad , \tag{3.2}$$

where

$$\mathbf{K} = \begin{bmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{bmatrix} \tag{3.3}$$

is called the camera (calibration) matrix, $\mathbf{I}_{3\times3}$ is the 3×3 identity matrix and $\mathbf{0}_{3\times1}$ is a matching zero column vector. In order to describe deviations in orthogonality and scale of the image's coordinate axis, the camera matrix is extended by an affinity and skew coefficient. In the computer vision literature, this is often parameterized by two different focal lengths f_x and f_y and the skew parameter s [152, p.185]:

$$\mathbf{K} = \begin{bmatrix} f_x & s & p_x \\ 0 & f_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \quad . \tag{3.4}$$

In the photogrammetry literature alternatively a scale difference parameter m and the principal distance f is regularly used [128, p.471]:

$$\mathbf{K} = \begin{bmatrix} 1 & s & p_x \\ 0 & (1+m) & p_y \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f & fs & p_x \\ 0 & f(1+m) & p_y \\ 0 & 0 & 1 \end{bmatrix} \quad . \tag{3.5}$$

The skew parameter s is zero for the typical machine vision cameras employed in this work.

In general, the object point $\mathbf{X}$ is not given in camera coordinates but with respect to a world coordinate system. Hence, the object point in the world frame $\mathbf{X}_{\text{world}}$ needs to be transformed to camera coordinates first:

$$\mathbf{X}_H = \begin{bmatrix} \mathbf{R} & -\mathbf{R}\mathbf{C}_{\text{world}} \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix} \begin{pmatrix} X_{\text{world}} \\ Y_{\text{world}} \\ Z_{\text{world}} \\ 1 \end{pmatrix} = \begin{bmatrix} \mathbf{R} & -\mathbf{R}\mathbf{C}_{\text{world}} \\ \mathbf{0}_{1\times 3} & 1 \end{bmatrix} \mathbf{X}_{\text{world},H} \quad , \tag{3.6}$$

where $\mathbf{C}_{\text{world}}$ is the camera center in world coordinates and $\mathbf{R}$ is the 3×3 rotation matrix representing the orientation of the camera coordinate frame in the world coordinate frame. Using Eq. 3.2 this yields for the projection of a point in world coordinates on the image plane in homogeneous coordinates:

$$\mathbf{x}_H = \mathbf{K}\mathbf{R} \begin{bmatrix} \mathbf{I}_{3\times 3} & -\mathbf{C}_{\text{world}} \end{bmatrix} \mathbf{X}_{\text{world},H} \quad . \tag{3.7}$$

This model is generalized through additional distortion models or physical parameters to describe the geometric effects of the lens optics or refraction.

3.2.2 Brown's Distortion Model

Real lens optics deviate from the ideal pinhole projection. Radial and tangential lens distortions are often compensated using polynomial models. An often employed model in the literature is the Brown or Brown-Conrady distortion model [66, 67, 91].

Radial Distortion

For an ideal pinhole camera, the angle of incidence of a light ray is equal to the angle of emergence. This is true since the model assumes a single entrance and exit pupil, which is infinitesimally small. This results in an undistorted, or orthographic, reproduction of the image. In compound lens designs, the entrance and exit pupil do not have the same position. Moreover, the position of the aperture influences how the light traverses the lens, as visualized in Fig. 3.5. If the aperture is moved towards the object, the light rays traverse the edges of the lens, which increases barrel distortion. Similarly, pincushion distortion is formed when the aperture is moved towards the image.

In underwater imaging, the refraction at flat ports causes a radial shift that is dependent on the angle of incidence. This shows up in the image as a pincushion distortion. Radial distortion is also observed with dome ports that are not perfectly centered. A shift of the optical center along the optical axis causes a radial distortion effect.

The relation of the actual projected point and the ideal point computed by the pinhole geometry is modeled by a radial displacement:

$$\mathbf{x}_d = \begin{pmatrix} x_d \\ y_d \end{pmatrix} = L(r) \begin{pmatrix} x - x_c \\ y - y_c \end{pmatrix} \quad , \tag{3.8}$$

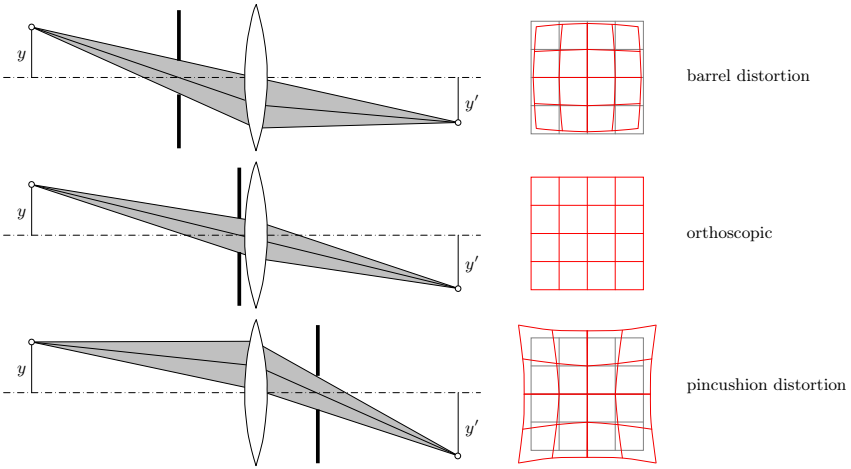


Fig. 3.5: Different forms of radial distortion in a spherical lens as a function of the aperture position (Source: [201, p.40]).

where $\mathbf{x}_d = (x_d, y_d)^T$ is the actual (distorted) image position, $\mathbf{x} = (x, y)^T$ is the ideal (undistorted) image position and $\mathbf{x}_c = (x_c, y_c)^T$ is the center of the radial distortion. $L(r)$ is a distortion factor, which is a function of the radius $r = \sqrt{(x - x_c)^2 + (y - y_c)^2}$ only. The center of radial distortion is typically chosen as the coordinates of the principal point in image space or as the center of the image. Different formulations for $L(r)$ are suggested in the literature. The function $L(r)$ is often approximated by a Taylor expansion $L(r) = 1 + k_1 r + k_2 r^2 + k_3 r^3 + \ldots$. In the computer vision literature [321, 350] often only the first few terms with even exponent, e.g., $L(r) = 1 + k_1 r^2 + k_2 r^4 + k_3 r^6$, are considered.

In the photogrammetry community, balanced distortion functions [202, p.172ff] with a zero crossing at a fixed image radius are suggested [188]. This has the advantage that the radial distortion parameters do not correlate as strongly with focal length, which results in more stable estimates of the camera parameters.

Tangential Distortion

Tangential distortion is caused, for example, by a misalignment of lens elements or a tilted image plane as visualized in Fig. 3.6. Since the housing of an underwater camera acts like an additional optical element, a misalignment causes tangential distortion. For example, a flat port that is tilted with respect to the image plane results in asymmetric distortion. Similarly, a misaligned dome port causes image distortions. A shift of the perspective center perpendicular to the optical axis causes a tangential distortion effect.

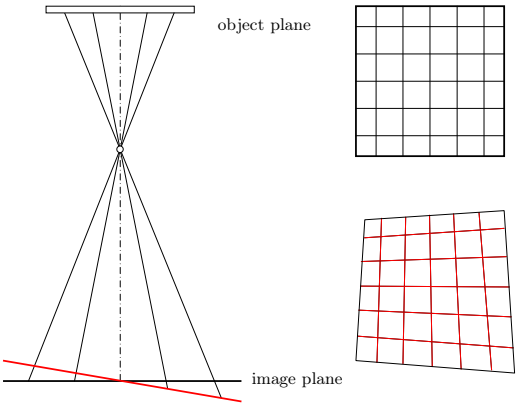


Fig. 3.6: Tangential distortion due to a tilted image plane (Source: [201, p.46]).

According to [67] or [91] and implemented, for example, in [60] or [56], the tangential and radial distortion is corrected in the following way:

Using the normalized pinhole projection

$$\mathbf{x}_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix} = \begin{pmatrix} X/Z \\ Y/Z \end{pmatrix} = \begin{pmatrix} \frac{x-p_x}{f_x} \\ \frac{y-p_y}{f_y} \end{pmatrix} \tag{3.9}$$

the complete model includes radial and tangential distortion defined as follows

$$\mathbf{x}_d = \mathbf{x}_n \left(1 + k_1 r^2 + k_2 r^4 + k_3 r^6\right) + \begin{pmatrix} 2p_1 x_n y_n + p_2(r^2 + 2x_n^2) \\ p_1(r^2 + 2y_n^2) + 2p_2 x_n y_n \end{pmatrix} , \tag{3.10}$$

where $\{k_1, k_2, k_3\}$ are the radial and $\{p_1, p_2\}$ are the tangential distortion parameters. Here, $\mathbf{x}_d = (x_d, y_d)^T$ are the measured (distorted) normalized point coordinates and $r^2 = x_n^2 + y_n^2$. This is extended in [125] by a division model with additional parameters to provide a more accurate approximation than the polynomial model. The parameters of the distortion model $\mathbf{q} = \{k_1, k_2, k_3, p_1, p_2\}$ are considered a part of the intrinsic parameters of the camera.

Here, the distortion function is defined as the function that transforms the ideal pinhole projection coordinates to the real, distorted coordinates. For example, in BA applications, the distortion function is sometimes defined for the back projection instead of the forward projection for optimization purposes [51]. This allows to project an image observation to a 3D vector in world space without inverting the distortion function. To compute the 3D camera ray using the formulation of Eq. 3.10, the distortion function needs to be inverted. This is done iteratively by computing the approximate solution

$$\mathbf{x}'_n = \frac{1}{1 + k_1 r^2 + k_2 r^4 + k_3 r^6} \left[\mathbf{x}_d - \begin{pmatrix} 2p_1 x_n y_n + p_2(r^2 + 2x_n^2) \\ p_1(r^2 + 2y_n^2) + 2p_2 x_n y_n \end{pmatrix} \right] , \tag{3.11}$$

where we set $\mathbf{x}_n = \mathbf{x}'_n$ and $r = \sqrt{{x'_n}^2 + {x'_n}^2}$ and iterate until the change of $\mathbf{x}'_n$ is negligible. In practice, this is typically achieved with less than ten iterations.

3.2.3 Absorption of Refraction Effects by the Distortion Model

Although in underwater conditions Brown's distortion model does not explicitly describe the physical properties of refraction, it may provide a sufficient approximation for some underwater imaging applications, and low errors are achievable [295]. In general, this model does not fully describe the multi-media imaging case of a flat port underwater camera, since the observed radial distortion is distance dependent.

To analyze the errors, the case of refraction at a single media interface parallel to the image plane is considered [202, p.415ff]. The geometry is depicted in Fig. 3.7. Here, the expression of Snell's law (cf. Section 2.1.2) as a function of the tangent is used [202, p.126]:

$$\frac{\tan\theta_1}{\tan\theta_2} = \sqrt{\mu^2 + (\mu^2 - 1)\tan^2\theta_1} \quad , \tag{3.12}$$

where θ_1 is the angle of incidence, θ_2 is the angle of refraction and $\mu = \frac{\mu_1}{\mu_2}$ is the ratio of the respective refractive indices with $\mu_2 > \mu_1$, the radial shift Δr due to refraction is given by:

$$\begin{aligned}
\Delta r = r_0 - r = & && (3.13)\\
= f\tan(\theta_1) - \frac{f}{Z}R = & && (3.14)\\
= f\tan(\theta_1) - \frac{f}{Z}\left(\tan(\theta_1)d + \tan\theta_2(Z-d)\right) = & && (3.15)\\
= f\tan(\theta_1)\frac{Z-d}{Z}\left(1 - \frac{\tan\theta_2}{\tan\theta_1}\right) \overset{\text{Eq. 3.12}}{=} & && (3.16)\\
= f\tan(\theta_1)\frac{Z-d}{Z}\left(1 - \frac{1}{\sqrt{\mu^2 + (\mu^2-1)\tan^2\theta_1}}\right) \quad . & && (3.17)
\end{aligned}$$

Here, f is the focal length, Z is the distance of the object point from the camera center along the optical axis, d is the distance of the plane of refraction from the camera center, θ_1 is the angle of the light ray with respect to the optical axis and μ is the ratio of the refractive indices. All other geometric variables are defined as shown in Fig. 3.7.

In Eq. 3.17 the radial shift is a function of the angle of the light ray θ_1 and the object point distance Z. The radial shift Δz is zero, when:

- $\theta_1 = 0$: A light ray perpendicular to the media interface is not refracted.
- $\mu = 1$: The angle of incidence is equal tp the angle of refraction if both media have equal refractive indices.
- $Z = d$: An object point that is co-planar with the media interface does not result in a radial shift.

The error of modeling the radial shift using a radial correction function as described in Section 3.2.2 is small, when:

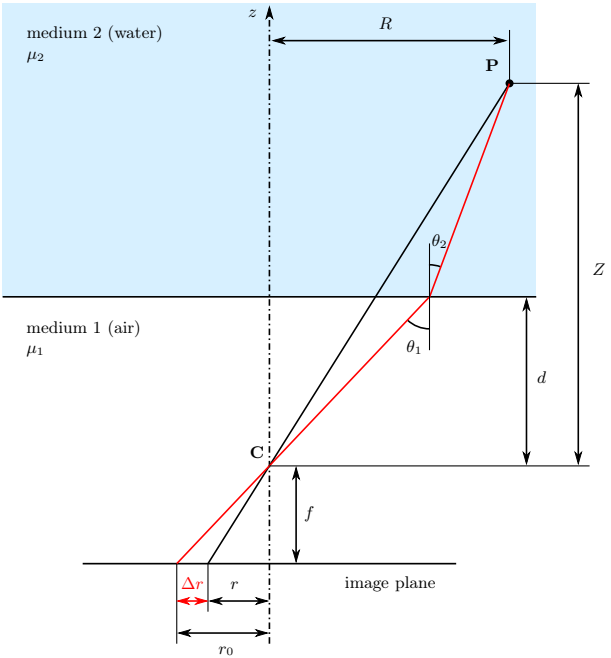


Fig. 3.7: Radial shift due to refraction at a single media interface parallel to the image plane. Image reproduced based on [202, p.416].

- $Z >> d$ and therefore $\frac{Z-d}{Z} \approx 1$: The distance dependence is small if the camera is very close to the interface and only objects far away from the camera are considered.
- $\tan\theta_1 \approx 0$: Objects close to the optical axis exhibit a smaller error since the effect of refraction is smaller because the light ray travels through the interface approximately perpendicular. Therefore, we observe the lowest deviations in the image center, and the error increases towards the image borders.
- $\frac{Z-d}{Z} \approx$ const.: If the radial correction function is tailored for a certain object distance, the errors are small for objects within this depth range.

Therefore, depending on the accuracy requirements good measurement results are obtained without an explicit physical-geometric camera model if the distance to the object surface is kept approximately constant, and object points farther away from the camera are considered.

However, due to reduced visibility for 3D reconstruction in water measurements at short distances are necessary. Moreover, it is difficult to perfectly align the image plane and refraction

plane, which introduces additional errors, and for large water depths thick glass windows are necessary, resulting in two plane parallel interfaces. This is not considered in Eq. 3.17.

3.3 Flat Refractive Geometry

The work at hand models an underwater camera with a flat port by refraction through parallel-sided media. This is valid for a homogeneous glass or acrylic window with constant thickness as the view port. Given a calibrated camera, the refraction effects are described by explicitly compensating for the refraction at each interface layer. Consequently, the housing parameters are the thickness of the glass window, the orientations of the interface planes and the corresponding distances from the camera center, and the refractive index of the glass.

Applying this model for calibration and 3D reconstruction requires deriving the forward and back projection models. While the back projection of a camera ray to the ray in water can be solved directly, this is not true for the forward projection. The forward projection of an object point on the camera plane requires finding the intersection points on the interfaces between the individual media that satisfy the refraction constraints.

3.3.1 Back Projection

The geometry of the back projection is visualized in Fig. 3.8. The camera ray $\mathbf{v}_0$ is refracted at the air-glass interface at point $\mathbf{q}_1$, resulting in the ray $\mathbf{v}_1$. Subsequently, the ray $\mathbf{v}_1$ is refracted at the glass-water interface at point $\mathbf{q}_2$. This results in the ray $\mathbf{v}_2$.

Using the vector form of Snell's law described in Eq. 2.11 $\mathbf{q}_2$ and $\mathbf{v}_2$ are computed analytically in the following way: First, the intersection of the camera ray with the air-glass interface is found. Assuming a normalized camera ray vector ($\|\mathbf{v}_0\| = 1$) the intersection point $\mathbf{q}_1$ is

$$\mathbf{q}_1 = -\frac{d_0}{\mathbf{n}^T\mathbf{v}_0}\mathbf{v}_0 \quad , \tag{3.18}$$

where $\mathbf{n}$ is the normal vector of the interface plane with unit length pointing towards the camera, and d_0 is the orthogonal distance of the camera center from the interface. The refracted ray $\mathbf{v}_1$ is computed using Eq. 2.11:

$$\mathbf{v}_1 = \text{RefractedRay}(\mathbf{v}_0, \mu_0, \mu_1) \quad , \tag{3.19}$$

where μ_0 and μ_1 are the refractive indices of air and the glass of the window. Finally, computing the intersection point with the glass-water interface $\mathbf{q}_2$ and the refracted ray $\mathbf{v}_2$ yields

$$\mathbf{q}_2 = \mathbf{q}_1 - \frac{d_1}{\mathbf{n}^T\mathbf{v}_1}\mathbf{v}_1 \tag{3.20}$$

$$\mathbf{v}_2 = \text{RefractedRay}(\mathbf{v}_1, \mu_1, \mu_2) \quad , \tag{3.21}$$

where μ_2 is the refractive index of water, and d_1 is the thickness of the glass window. The observed point $\mathbf{P}$ lies on the line

$$\mathbf{P} = \mathbf{q}_2 + \lambda\mathbf{v}_2 \tag{3.22}$$

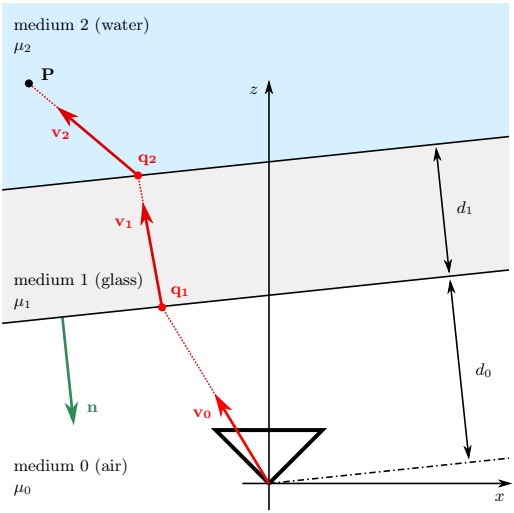


Fig. 3.8: Back projection of a camera ray $\mathbf{v}_0$ to the point $\mathbf{P} = \mathbf{q}_2 + \lambda \mathbf{v}_2$ in camera coordinates.

with $\lambda \geq 0$.

If the underwater camera is used in air, which means $\mu_0 = \mu_2$, then $\mathbf{v}_0 || \mathbf{v}_2$ applies. Therefore, the image of an object captured by the camera through the glass window appears laterally shifted compared to an image captured without the window.

3.3.2 Forward Projection

The forward projection of a 3D point on the camera sensor is more difficult to compute because it requires solving for the points of refraction on each interface surface, such that Snell's law holds true. This is numerically solvable by optimization using the back projection model. However, this requires running a non-linear optimizer for every single computed forward projection, which is inefficient. Therefore, this work follows the approach of Agrawal et al. [31] using an analytical formulation of the forward projection.

To simplify the derivation, we look only at the plane of refraction. This is visualized in the left part of Fig. 3.9. The light path lies in the plane spanned by the camera center, the 3D point $\mathbf{P}$ and the interface normal $\mathbf{n}$. An orthogonal coordinate system is constructed, which is defined by the unit vectors $\mathbf{z}_1$ and $\mathbf{z}_2$ lying on the plane of refraction and with $\mathbf{z}_1$ aligned with

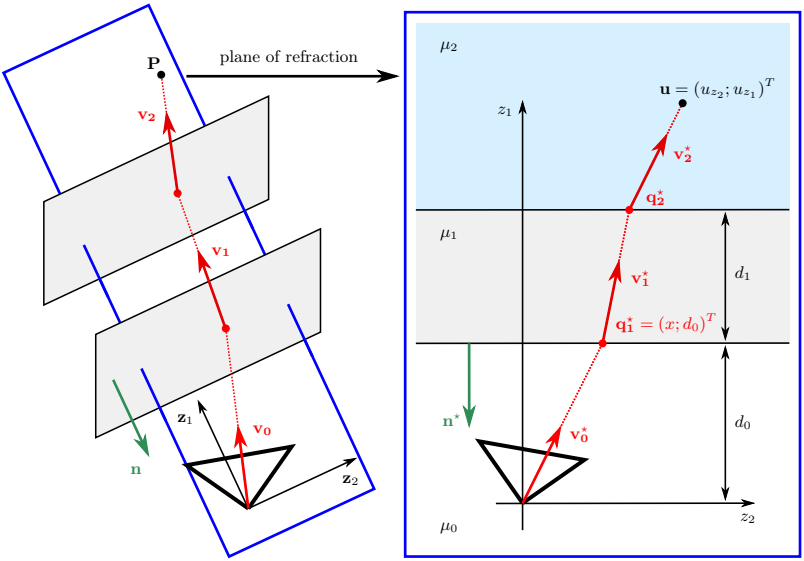


Fig. 3.9: Geometry of the forward projection. To solve for the intersection points, we consider the plane of refraction spanned by the camera center, the point $\mathbf{P}$ and the interface normal. Image reproduced based on [32].

the interface normal:

$$\mathbf{z}_1 = -\mathbf{n} \tag{3.23}$$

$$\mathbf{z}_2 = \frac{\mathbf{t} \times \mathbf{z}_1}{||\mathbf{t} \times \mathbf{z}_1||} \quad \text{with} \quad \mathbf{t} = \frac{\mathbf{n} \times \mathbf{P}}{||\mathbf{n} \times \mathbf{P}||} \quad , \tag{3.24}$$

where $\mathbf{t}$ is the normal of the plane of refraction and the camera center at the origin of the coordinate system.

By transforming $\mathbf{P}$ to this coordinate system, the forward projection is derived in 2D on the plane of refraction. The geometry is depicted in the right part of Fig. 3.9. The 3D point $\mathbf{P}$ is transformed to the point

$$\mathbf{u} = \begin{pmatrix} u_{z_2} \\ u_{z_1} \end{pmatrix} = \begin{pmatrix} \mathbf{z}_2^T \mathbf{P} \\ \mathbf{z}_1^T \mathbf{P} \end{pmatrix} \tag{3.25}$$

and the interface normal in the new coordinate system is $\mathbf{n}^\star = (0; -1)^T$. The vector $\mathbf{v}_i^\star$ denotes the projection of the ray $\mathbf{v}_i$ on the plane of refraction, and $\mathbf{q}_i^\star$ denotes the intersection points of the light path with the i-th medium boundary.

To solve the forward projection, we need to find the z_2-component, denoted by x, of the inner layer intersection point $\mathbf{q}_1^\star = (x; d_0)^T$ subject to the constraints that the rays $\mathbf{v}_i^\star$ adhere to Snell's law and that $\mathbf{v}_2^\star$ is parallel to $(\mathbf{u} - \mathbf{q}_2^\star)$. The constraint on the last refracted ray $\mathbf{v}_2^\star$ is given by

$$\mathbf{v}_2^\star \times (\mathbf{u} - \mathbf{q}_2^\star) = 0 \quad . \tag{3.26}$$

Using Eq. 2.8, the refracted ray is expressed as a linear combination of the incident ray and the interface normal

$$\begin{aligned} \mathbf{v}_2^\star &= \alpha_2 \mathbf{v}_1^\star + \beta_2 \mathbf{n}^\star = && (3.27)\\ &= \alpha_2 \left(\alpha_1 \mathbf{v}_0^\star + \beta_1 \mathbf{n}^\star\right) + \beta_2 \mathbf{n}^\star = && (3.28)\\ &= \alpha_2 \alpha_1 \mathbf{v}_0^\star + \left(\alpha_2 \beta_1 + \beta_2\right) \mathbf{n}^\star \quad . && (3.29) \end{aligned}$$

Without loss of generality $\mu_0 = 1$ is chosen since Snell's law depends only on the ratio of refractive indices. In order to simplify the equations, the parameterization $\mathbf{v}_0^\star = (x; d_0)^T$ is considered. Using Eq. 2.9 $\alpha_1 = \frac{1}{\mu_1}$ and $\alpha_2 = \frac{\mu_1}{\mu_2}$ are found. From Eq. 2.10 β_1 is given by

$$\begin{aligned} \beta_1 &= \frac{-\mathbf{v}_0^{\star T}\mathbf{n}^\star - \sqrt{(\mathbf{v}_0^{\star T}\mathbf{n}^\star)^2 - (1-\mu_1^2)(\mathbf{v}_0^{\star T}\mathbf{v}_0^\star)}}{\mu_1} = && (3.30)\\ &= \frac{d_0 - \sqrt{d_0^2 - (1-\mu_1^2)(x^2+d_0^2)}}{\mu_1} = \frac{d_0 - \sqrt{d_0^2\mu_1^2 + \mu_1^2 x^2 - x^2}}{\mu_1} && (3.31) \end{aligned}$$

and β_2 is given by

$$\begin{aligned} \beta_2 &= \frac{-\mu_1 \mathbf{v}_1^{\star T}\mathbf{n}^\star - \sqrt{\mu_1^2(\mathbf{v}_1^{\star T}\mathbf{n}^\star)^2 - (\mu_1^2-\mu_2^2)(\mathbf{v}_1^{\star T}\mathbf{v}_1^\star)}}{\mu_2} \quad \overset{\mathbf{v}_1^\star = \alpha_1 \mathbf{v}_0^\star + \beta_1 \mathbf{n}^\star}{=} && (3.32)\\ &= \frac{\sqrt{d_0^2\mu_1^2 + \mu_1^2 x^2 - x^2} - \sqrt{d_0^2\mu_2^2 + \mu_2^2 x^2 - x^2}}{\mu_2} \quad . && (3.33) \end{aligned}$$

Now using Eq. 3.29 and similar to Eq. 3.20 we find

$$\mathbf{v}_2^\star = \alpha_2\alpha_1\mathbf{v}_0^\star + (\alpha_2\beta_1 + \beta_2)\,\mathbf{n}^\star = \frac{1}{\mu_2}\begin{pmatrix} x \\ \sqrt{D_2} \end{pmatrix} \tag{3.34}$$

$$\mathbf{q}_2^\star = \mathbf{q}_1^\star - \frac{d_1}{\mathbf{n}^{\star T}\mathbf{v}_1^\star}\mathbf{v}_1^\star = \begin{pmatrix} \left(1 + \frac{d_1}{\sqrt{D_1}}\right) x \\ d_0 + d_1 \end{pmatrix} \quad , \tag{3.35}$$

where $D_1 = d_0^2\mu_1^2 + \mu_1^2 x^2 - x^2$ and $D_2 = d_0^2\mu_2^2 + \mu_2^2 x^2 - x^2$.

Substituting $\mathbf{v}_2^\star$ and $\mathbf{q}_2^\star$ in Eq. 3.26 and multiplying both sides of the equation with $\mu_2\sqrt{D_1}$ yields

$$k_1\sqrt{D_1} + k_2\sqrt{D_1 D_2} + k_3\sqrt{D_2} = 0 \quad , \tag{3.36}$$

where $k_1 = x(d_0 + d_1 - u_{z_1})$, $k_2 = u_{z_2} - x$ and $k_3 = -d_1 x$. By separating the terms and squaring twice the square roots in Eq. 3.36 are removed:

$$\left(k_1^2 D_1 + k_3^2 D_2 - k_2^2 D_1 D_2\right)^2 - 4k_1^2 k_3^2 D_1 D_2 = 0 \quad . \tag{3.37}$$

Expanding Eq. 3.37 yields a twelfth degree polynomial in x. By finding the roots of the polynomial, up to twelve unique solutions are created. The correct solution is retrieved by removing all solutions with an imaginary part and checking the validity of the remaining solution using the backprojection model described in Section 3.3.1. A valid solution for x is transformed to the 3D camera ray

$$\mathbf{v}_0 = \frac{x\mathbf{z}_2 + d_0\mathbf{z_1}}{||x\mathbf{z}_2 + d_0\mathbf{z_1}||} \quad . \tag{3.38}$$

Note this model can be extended to a higher number of plane parallel media interfaces. This yields higher-order polynomials, which need to be solved.

3.3.3 Implementation and Runtime

For the pinhole and distortion model as described in Section 3.2, the implementation provided in *OpenCV: Open Source Computer Vision Library* [60] is used. The refractive model described in Section 3.3 is an own implementation in C++ based on the *Eigen* library [148]. The polynom solver provided by the *GNU Scientific Library* [139] is used to solve Eq. 3.36.

The analytical forward projection, described in Section 3.3.2, is compared with a numerical solution by inverting the back projection model described in Section 3.3.1 using non-linear optimization. The *Ceres Solver* [28] is employed for the numerical solution of the forward projection. Considering the result, the two approaches are identical regarding numerical precision, but the analytical solution is more computationally efficient.

Fig. 3.10 shows the runtime for the projection of 1000 points using the analytical and numerical forward projection solution. As reference the runtime of the corresponding back projection is provided. The 1000 points are sampled randomly in image space. The experiment is conducted 4000 times, and the variation of the runtime is reported in Fig. 3.10 using a boxplot. The simulated flat port underwater camera is a distortion-free camera with a 12.5 mm lens and a 2 cm thick flat port 3 cm in front of the entrance pupil of the camera. The tilt of the glass window is varied relative to the image plane in the range 0 deg to 30 deg. The runtime numbers are produced on an Intel Core i7-8550U CPU with 4.0 GHz without parallelization. Therefore, the numbers reflect the single core execution time.

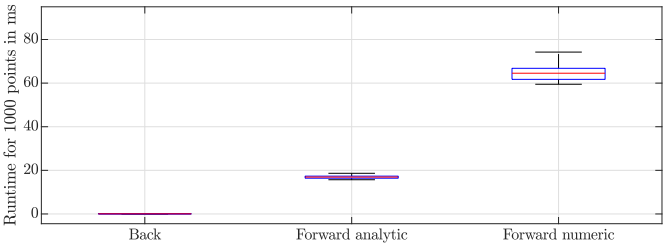


Fig. 3.10: Runtime comparison of back projection, analytical forward projection and numerical forward projection.

The back projection of 1000 points has a median runtime of 0.06 ms while the forward projection has a significantly longer runtime since it requires solving a polynomial or non-linear optimization. A median runtime of 16.9 ms for the analytical forward projection and a median runtime of 64.5 ms for the numeric solution using non-linear optimization is observed. Here, the analytical solution is approximately 3.8 times faster than the numerical solution. Moreover, the runtime of the analytical solution is more consistent. The non-linear optimization approach converges at different speeds depending on the quality of the initialization and model parameters.

3.4 Calibration of Housing and Camera Parameters

Camera Calibration means estimating the (intrinsic) parameters of the camera model. Additionally, in underwater imaging the housing parameters of the physical-geometric model need to be estimated. In total, we want to estimate the following 16 parameters:

- focal length f_x, f_y and image coordinates of the principal point p_x, p_y (see Section 3.2.1)
- parameters of the image errors: radial $\{k_1, k_2, k_3\}$ and tangential $\{p_1, p_2\}$ distortion parameters (see Section 3.2.2)
- housing parameters: interface normal $\mathbf{n} = (n_x, n_y, n_z)^T$, air-glass interface distance d_0, glass-water interface distance d_1 and refractive indices of the individual layers μ_1, μ_2 (see Section 3.3)

Camera calibration is typically performed using a test-field calibration or via self-calibration:

Test-field calibration is based on a suitable calibration fixture with precisely identifiable object points with optionally precisely known coordinates or relative distances [95, 156, 287, 321,334,337,350]. This setup allows to do calibration very efficiently. Planar grids of points or spatial test fields with points distributed in 3D space are regularly used. The parameters of the camera model are then calculated using the known correspondences of 3D calibration points and 2D image observations. However, there are also approaches that do not require the 3D coordinates to be fully known. Camera parameters are also calculated as part of a BA in which the exterior orientation of the camera, the unknown 3D calibration object points and the camera model parameters are estimated simultaneously [188]. This allows to introduce additional constraints, such as known relative distances between individual points or individual coordinates, to provide scale [202, p.683ff]. The disadvantage is an elaborate setup with calibration fixtures and additional calibration image acquisition.

Self-calibration techniques do not use a specific calibration object. Strictly speaking, the calibration object is replaced by the actual object of measurement, which is assumed to be static and rigid during image acquisition. The camera model is then estimated from the constraints of capturing a static scene by a moving camera using image information alone [218]. For a camera with fixed internal parameters correspondences between a minimum of three images are sufficient to recover the interior and exterior camera parameters and the 3D object points up to scale [151]. To define scale a reference length needs to be measured. The advantage of self-calibration is that it is more flexible and that calibration

is performed simultaneously with object measurement, which improves the accuracy of object reconstruction tasks. In order to have a defined reference for verifying calibration, test points are placed in the scene if needed. Self-calibration may not be applicable for situations that do not allow capturing of suitable images with configurations to obtain reliable results.

For the highest accuracy, the calibration is ideally performed under the same imaging conditions as the actual measurement task. This means the calibration for underwater imaging is conducted in situ with the same water temperature, salinity and pressure to keep the refraction effects constant. In more difficult imaging tasks, such as multi-media imaging, it is sensible to prefer test-field calibration because it provides more reliable calibration observations and has an in-built quality control due to the known structure of the calibration fixture.

3.4.1 Zhang's Method for Camera Calibration

A popular and robust method for camera calibration in the computer vision literature is Zhang's method [350]. The technique is implemented in open source libraries and toolboxes, such as OpenCV [60] or Bouget's Camera Calibration Toolbox for Matlab [56]. The method requires the camera to observe a planar pattern with known 2D coordinates at different orientations. The 2D metric information of the calibration pattern points and the corresponding image observations are used to estimate the camera's intrinsic parameters. At least two images with different orientations are necessary. However, in practice, a larger set of calibration images is required for stable results. Typically, at least 10 to 20 images are required to achieve low errors. The approach accurately determines the camera parameters and is easy to realize in practice because by using automatic grid corner extractors, such as chessboard detectors, the approach is highly automated. However, the results may be influenced by the quality of the captured calibration images. Some methods try to improve on this by guiding the user to capture good images of the calibration target using an information maximization scheme [269].

For typical machine vision cameras, distortion is expected to be small. Therefore, a reasonable initial guess of the intrinsic parameters of the pinhole model is found, ignoring the distortion. Zhang's method follows a three-step process:

1. Find pinhole model parameters and extrinsic parameters of the planar pattern using a closed-form solution

2. Estimate radial distortion based on a least-squares solution using the initial intrinsic parameters

3. Refine all camera model parameters using non-linear optimization

The first step of Zhang's method uses a simplification that all points of the calibration target lie

in the X/Y-plane. The assumption $Z = 0$ simplifies Eq. 3.7 to

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{bmatrix} f_x & s & p_x \\ 0 & f_y & p_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & t_1 \\ r_{21} & r_{22} & t_2 \\ r_{31} & r_{32} & t_3 \end{bmatrix} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} = \tag{3.39}$$

$$= \underbrace{\mathbf{K} \begin{bmatrix} \mathbf{r_1} & \mathbf{r_2} & \mathbf{t} \end{bmatrix}}_{\mathbf{H}} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} = \mathbf{H} \begin{pmatrix} X \\ Y \\ 1 \end{pmatrix} \quad , \tag{3.40}$$

where $\mathbf{H}$ is a 3x3 homography matrix. For each observed pattern point $(X_i, Y_i)^T$ in the same image and therefore with the same rotation and translation of the calibration board, we obtain

$$\begin{pmatrix} x_i \\ y_i \\ 1 \end{pmatrix} = \mathbf{H} \begin{pmatrix} X_i \\ Y_i \\ 1 \end{pmatrix} \quad i = 1, \ldots, I \quad . \tag{3.41}$$

From these equations a system of linear equations is built and solved, which leads to an estimate of $\mathbf{H}$ that is defined up to a scale factor. Since the homography $\mathbf{H}$ has eight Degrees of Freedom (DoF) at least four points in the image are necessary. Because the six extrinsic parameters for each image (three for rotation and three for translation) need to be found, a single image provides only two constraints on the camera matrix $\mathbf{K}$. From Eq. 3.40 follows

$$\mathbf{H} = \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3 \end{bmatrix} = \mathbf{K} \begin{bmatrix} \mathbf{r_1} & \mathbf{r_2} & \mathbf{t} \end{bmatrix} \tag{3.42}$$

$$\implies \begin{bmatrix} \mathbf{r_1} & \mathbf{r_2} & \mathbf{t} \end{bmatrix} = \mathbf{K}^{-1} \begin{bmatrix} \mathbf{h}_1 & \mathbf{h}_2 & \mathbf{h}_3 \end{bmatrix} \quad . \tag{3.43}$$

By exploiting that $\mathbf{r}_1 = \mathbf{K}^{-1}\mathbf{h}_1$ and $\mathbf{r}_2 = \mathbf{K}^{-1}\mathbf{h}_2$ are orthonormal basis vectors the constraints are rewritten as

$$\mathbf{r}_1^T \mathbf{r}_2 = 0 \implies \mathbf{h}_1^T \mathbf{K}^{-T} \mathbf{K}^{-1} \mathbf{h}_2 = 0 \tag{3.44}$$

$$\|\mathbf{r}_1\| = \|\mathbf{r}_2\| = 0 \implies \mathbf{h}_1^T \mathbf{K}^{-T} \mathbf{K}^{-1} \mathbf{h}_1 = \mathbf{h}_2^T \mathbf{K}^{-T} \mathbf{K}^{-1} \mathbf{h}_2 \tag{3.45}$$

Using homographies from multiple images, a system of linear equations is built from these constraints to solve for $\mathbf{B} = \mathbf{K}^{-T}\mathbf{K}^{-1}$. Using the constraints on the camera matrix $\mathbf{K}$, at least three different views are necessary to solve for $\mathbf{B}$. If less than three views are available, additional constraints are imposed, such as the skewless constraint $s = 0$ or fixing the image coordinates of the principal point $(p_x, p_y)^T$ to the image center. Finally, $\mathbf{K}$ is computed by matrix decomposition of $\mathbf{B}$.

This approach is only applicable to planar patterns, such as chessboards. For 3D test fields, this step is skipped and a reasonable initialization needs to be provided. This is done, for example, by using known parameters of the camera-lens system, using Zhang's method on a planar subset of the 3D test field or by applying another method for estimating the pinhole parameters.

In the second step, the distortion parameters are estimated given the estimates of the camera matrix $\mathbf{K}$ by minimizing the distance between the ideal pinhole projection $\mathbf{x_n}$ and the observed

distorted image coordinates $\mathbf{x}_d$ using Brown's distortion model as described in Section 3.2.2. Each 2D point in each image contributes to two equations by computing $\mathbf{x}_d - \mathbf{x_n}$ using the distortion model. The distortion parameters are then computed as the least-squares solution.

The full refined parameter set is calculated by minimizing the non-linear error function:

$$\min_{(\mathbf{K},\mathbf{q},\mathbf{R}_i,\mathbf{t}_i)} \sum_{i=1}^{n} \sum_{j=1}^{m} \left\| \mathbf{x}_{ij} - \breve{\mathbf{x}} \left(\mathbf{K}, \mathbf{q}, \mathbf{R}_i, \mathbf{t}_i, \mathbf{X}_j \right) \right\|^2 \quad , \tag{3.46}$$

where $\mathbf{x}_{ij}$ is the observation of the calibration point $\mathbf{X}_j$ in image i and $\breve{\mathbf{x}} \left(\mathbf{K}, \mathbf{q}, \mathbf{R}_i, \mathbf{t}_i, \mathbf{X}_j \right)$ is the projection of $\mathbf{X}_j$ according to the pinhole model with Brown's distortion model as described in Section 3.2. Here, the camera model is parameterized by the camera matrix $\mathbf{K}$ and the set of radial and tangential distortion parameters $\mathbf{q} = \{k_1, k_2, k_3, p_1, p_2\}$. The orientation of the calibration grid in image i is parameterized by the rotation $\mathbf{R}_i$ and translation $\mathbf{t}_i$. The results are the intrinsic camera parameters and the pose of the calibration target in each image. The minimization is performed using the Levenberg-Marquardt algorithm in [350].

3.4.2 Error Functions for Camera Calibration

Different error statistics are proposed in the literature to assess how well the model fits the observed projection of an actual camera [149, pp.27ff]. The reprojection error is the most common approach to describe the quality of the camera calibration. It measures the distance between the observed image points $\mathbf{x}_i$ and the back-projection of the corresponding 3D object point on the image plane $\breve{\mathbf{x}} \left(\mathbf{X}_i \right)$. This is a 2D error in image space and is measured in pixels. The error is also often reported in micrometers if the size of a single pixel on the image sensor is known. Typically, the Root Mean Squared Error (RMSE) of the complete calibration image set is reported:

$$RMSE = \sqrt{\frac{\sum_i \left\| \mathbf{x}_i - \breve{\mathbf{x}} \left(\mathbf{X}_i \right) \right\|^2}{n}} \quad , \tag{3.47}$$

where n is the number of point correspondences. Note that this error is computed during the calibration on the set of calibration images. Therefore, the reprojection error describes how well the camera model fits the calibration data. It does not provide information on how well the camera model generalizes to other data. Therefore, it is sensible to recompute the error on an additional set of test images to assess the calibration quality.

In the literature, projective error and re-projective error are often used interchangeably. Hanning [149, p.28] distinguishes between the projective error measured in the image plane and the re-projective error that is measured in the plane $\{z = 1\}$ in camera coordinates. This gives a normalization of the unit of the projective error. This work uses the term reprojection error to describe the error in the image plane.

In photogrammetric applications, the Euclidean error of the object coordinates is commonly used [149, pp.29ff]. This is useful for 3D reconstruction tasks since it provides an error measure in object space. However, compared to the reprojection error, this has the disadvantage that the Euclidean error is dependent on the distance to the calibration object.

3.4.3 Estimation of Housing Parameters

For calibration of the housing parameters, an initial estimate is required. A reasonable approach is to take the parameters from the Computer-Aided Design (CAD) model of the housing or by physically measuring the geometric dimensions. Refractive indices are provided by tables in the literature, are available from the data sheet of the glass manufacturer or are computed based on empirical functions [265].

If no prior knowledge about the underwater housing is available, Agrawal et al. suggest a method for estimating the axis and layer thickness for a system of multiple flat refractive layers [31]. This method corresponds the multi-layer flat refraction to an axial camera, assuming all interface layers are parallel. Here, the axis is the line through the camera center parallel to the plane normal of the refractive layers. The system is considered axial since all refracted rays intersect this axis. The approach finds the housing parameters in a two-step process:

1. Axis estimation based on coplanarity constraints

2. Find layer thickness based on the refractive constraint at each layer posed by Snell's law

Note that this approach requires the camera to be calibrated, i.e., the intrinsic parameters of the camera and subsequently the optical rays in the camera coordinate system are known.

In the first step, the constraint from Snell's law is applied, which requires that the incident ray, normal and refracted ray lie in the same plane. If the correspondence of a 3D object point $\mathbf{P}$ and the camera ray $\mathbf{v}_0$ is known, a coplanarity constraint is formulated. The 3D object point transformed in the camera coordinate system $(\mathbf{RP}+\mathbf{t})^T$ lies in the plane spanned by the axis direction vector $\mathbf{A}$ and the camera ray $\mathbf{v}_0$ (cf. Fig. 3.9). This is true independent of the number of refractions at plane parallel refractive media interfaces. This is expressed as the coplanarity constraint:

$$(\mathbf{RP}+\mathbf{t})^T (\mathbf{A}\times\mathbf{v}_0) = 0 \quad , \tag{3.48}$$

where $(\mathbf{A}\times\mathbf{v}_0)$ is the normal to the plane of refraction. By reformulating this constraint and using multiple known 3D object points of a calibration structure, a system of linear equations is set up. Agrawal et al. [31] suggest two approaches: an 11-point algorithm based on solving the linear system using Singular Value Decomposition (SVD) and an 8-point algorithm inspired by [244]. This yields the axis $\mathbf{A}$, the rotation $\mathbf{R}$ and two components of $\mathbf{t}$. The component of $\mathbf{t}$ in the direction of the axis cannot be recovered from the coplanarity constraint alone. Agrawal et al. suggest computing the 8-point and 11-point algorithm in a Random Sample Consensus (RANSAC) method to find the best solution based on the coplanarity error.

In the second step, the refractive constraint from Eq. 3.26 is used. The refracted rays $\mathbf{v}_\mathbf{i}^\star$ are computed based on the axis $\mathbf{A}$. The projection of the 3D object point on the plane of refraction $\mathbf{u}$ is computed based on the estimated orientation from the first step. It is found up to an unknown offset α in the component in the direction of the axis. Using Eq. 3.26 this yields

$$\mathbf{v}_2^\star \times \left[\mathbf{u} + \begin{pmatrix}0\\ \alpha\end{pmatrix} - \mathbf{q}_2^\star\right] = 0 \quad , \tag{3.49}$$

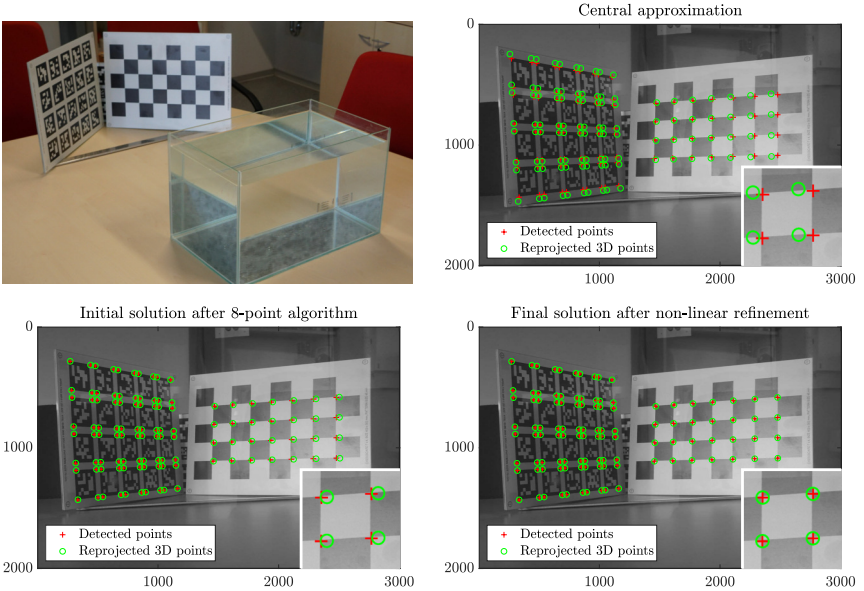


Fig. 3.11: Camera looking at a 3D test field through a water tank at an angle of approx. 45°. Top left: Experimental setup with water tank. Top right: Approximation using pinhole model. Bottom left: Initial solution using Agrawal's 8-point algorithm. Bottom right: Refined solution.

where α is an unknown offset in the direction of the axis. Substituting with Eq. 3.18 and Eq. 3.20 results in

$$\mathbf{v}_2^\star \times \begin{bmatrix} \frac{\mathbf{v}_0^\star}{\mathbf{n}^{\star T}\mathbf{v}_0^\star} & \frac{\mathbf{v}_1^\star}{\mathbf{n}^{\star T}\mathbf{v}_1^\star} & \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{bmatrix} \begin{bmatrix} d_0 \\ d_1 \\ \alpha \end{bmatrix} = -\mathbf{v}_2^\star \times \mathbf{u} \quad , \tag{3.50}$$

where d_i are the layer thicknesses. Hence, this yields a linear equation in d_0, d_1, α for each correspondence between $\mathbf{u}$ and $\mathbf{v}_0^\star$. Using multiple correspondences, a system of linear equations is built to obtain the layer thicknesses d_i and α. Note that if the first layer and the last layer have the same refractive index $\mu_0 = \mu_2$ the layer thickness d_0 cannot be recovered because in this case $\mathbf{v}_0^\star = \mathbf{v}_1^\star$ and the corresponding lines in the linear system become equal. Hence, for the underwater camera in air only the thickness of the glass d_1 is recovered but not the distance d_0.

Then, the result is further refined by non-linear optimization of the refractive forward projection. Fig. 3.11 shows a reproduction of the experiment of [31] for imaging through a water tank. The top left image shows the experimental setup of the tank and calibration fixture. A planar grid of AprilTag markers and a chessboard are captured through a water tank. The camera

is calibrated separately using Zhang's method (cf. Section 3.4.1). The relative transformation between the two calibration patterns is estimated based on images of the calibrated camera captured directly in air. Therefore, the 3D coordinates of all AprilTag and chessboard corners are known in the object coordinate system.

Fig. 3.11 depicts the results for capturing an image at an angle of approx. 45° through the water tank. Agrawal's approach is applied to the undistorted images based on the camera calibration, such that the camera is described by a pinhole model. Two planar refraction layers are modeled. Only an air-water and a water-air interface are estimated while the thin glass sheets of the water tank are neglected in the model. The refractive index of water is assumed to be $\mu_{\text{water}} = 1.33$.

The top left image shows the detected object points in red and the reprojected 3D points based on the pinhole model of the calibrated camera. In this case, only the pose of the calibration fixture is estimated. This results in a high RMS reprojection error of 15.391 pixels. A magnification of the top-right chessboard quad is shown in the bottom right corner of the graphics in order to show the errors between detected and reprojected 3D points.

The result of the 8-point algorithm with 100 RANSAC trials is visualized in the bottom left image. The RMS reprojection error is reduced to 5.772 pixels. The final result after non-linear refinement is shown in the bottom right image. This models the refraction well and yields an RMS reprojection error of 0.268 pixels.

The 8-point algorithm estimates a thickness of the water tank of 353 mm while the non-linear refinement yields 279 mm. The true thickness is 250 mm. Estimating the layer thicknesses based on a single image has a high uncertainty using the approach of Agrawal et al. [31]. This was also found in the literature [354].

Besides calibration using test points, there are approaches that do optimization using simulated images. Jordt and Koch [172] use a simulated view of a calibration target that is rendered based on the refractive imaging model and compared pixel-wise with the acquired image. The camera and housing parameters are then optimized based on the difference between the captured and simulated images.

Another approach by Gu et al. [147] achieves calibration of the housing parameters by mechanically fixing the camera relative to a planar calibration pattern and capturing an image in air, with the glass port and finally in water. By exploiting the different image locations for the three different scenarios, the normal of the glass port and interface distance is recovered. However, this involves disassembling the underwater camera, which does not allow fast re-calibration of the assembled camera system.

3.4.4 Underwater Camera Calibration Methodology

In order to achieve a high-accuracy calibration, an in situ calibration in water is performed. The non-linear optimization of the combined camera and refractive housing model is bootstrapped using an in-air calibration of the underwater camera and approximate housing parameter estimates. To summarize, for the calibration of the combined camera and housing system a three-step process is followed:

1. Calibrate camera-lens system in air with Zhang's method (cf. Section 3.4.1)

2. Find initial estimates of housing parameters from CAD or applying the method of Agrawal et al. (cf. Section 3.4.3)

3. Refinement of model parameters by non-linear optimization of the reprojection error of the combined camera and housing model

The first and second steps create an initial solution of the camera and housing parameters. In the third step, the complete parameter set of the camera model (cf. Section 3.2) and analytic forward projection based on raytracing (cf. Section 3.3.2) are optimized. Since the forward projection involves a polynomial solver, it is not possible to use analytical derivatives in the refinement step. Therefore, the reprojection error is minimized by applying numerical derivatives using the Ceres solver [28]. Basically, the intrinsic camera parameters found in air are also applicable to the underwater case. However, optimizing the full parameter set, including the distortion parameters of the camera, allows to absorb residual errors, which reduces the overall reprojection error. Moreover, it is usually disadvantageous to disassemble the underwater camera to find the exact parameters in air. Therefore, the camera calibration in air, strictly speaking, includes the refraction effects of the glass port in air, which causes, for example, a shift of the principal point. This requires a re-calibration of the intrinsic camera parameters in water.

In principle, it is possible to optimize all camera and housing parameters in the framework of the ray tracing based model of flat refraction. However, some of the housing parameters and the intrinsic parameters of the camera are highly correlated. Furthermore, some geometric parameters are also measured precisely by other means, e.g., the thickness of the view port window. Therefore, it is sensible to fix some of the model parameters. This work fixes the thickness of the glass window, which is measured accurately with calipers. Moreover, the refractive indices are taken from tables in the literature and datasheets and are not optimized in the refinement step.

3.5 Calibration Artifacts

The algorithms for camera calibration require as input precise correspondences between object points in 3D space and the projection of these points on the image plane. The quality of these correspondences significantly impacts the estimation of the camera model parameters. Therefore, target marks that are automatically detected in the image with high accuracy and precision even under difficult imaging conditions, such as water turbidity and blurred images, enable to create a low noise input set of correspondences. If a 3D calibration fixture is used, uniquely coded targets facilitate association between image and object points. This is especially true if only part of the calibration fixture is visible in the image frame or if some target marks are not detected due to challenging lighting conditions, occlusions, or dirt on the calibration target. In general, a larger calibration structure is beneficial, especially in water, since it is detected over larger distances, which allows one to take calibration data in the whole measurement range.

Fig. 3.12 shows different calibration artifacts used in this work. The top left image shows a 3D calibration structure built from rectangular aluminum rods and sheets. AprilTag markers [252] are printed on vinyl waterproof stickers and glued to the structure. The 2D barcode allows to assign every corner of the square tags a unique ID number. This has the advantage that

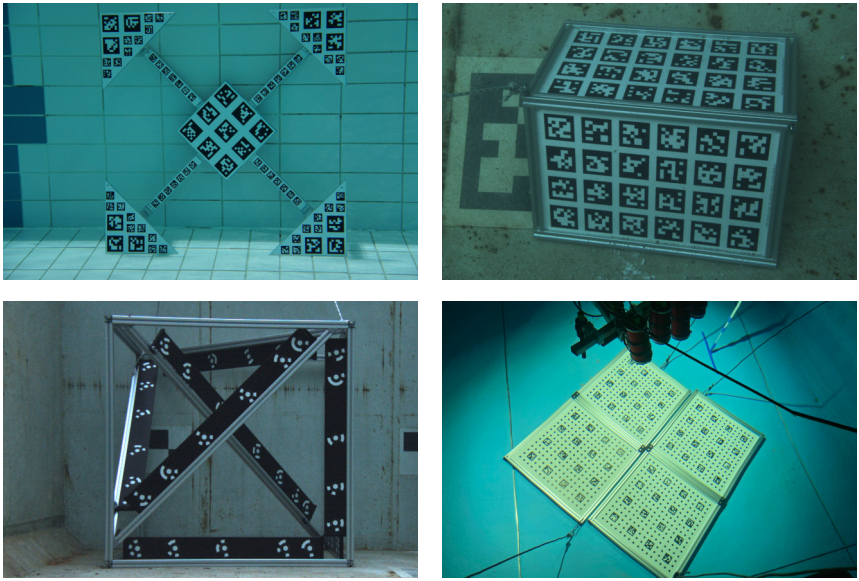


Fig. 3.12: Different underwater calibration fixtures. Top left: 3D calibration structure composed of AprilTags with different sizes. Top right: 3D calibration fixture composed of two planar AprilTag grids. Bottom left: Cuboidal calibration artifact with coded circle targets. Bottom right: Planar circle grids with ArUco marker.

calibration points are extracted automatically even if only part of the structure is visible in the image. AprilTags with different sizes are used to enable detection at large distances as well as a high number of calibration points at close distances. The top right image shows a calibration fixture composed of two planar AprilTag grids directly printed on glass sheets. The two planes are perpendicular to each other. According to the print service provider, the dimensional accuracy of the print is approximately one tenth of a millimeter. The relative orientation of the two AprilTag grids is computed in air using BA with the print coordinates as constraints. This problem is solved using the Damped Bundle Adjustment Toolbox (DBAT) [76]. The bottom left image shows a cuboidal calibration artifact with coded circle targets. The bottom right image shows a calibration pattern composed of four circle grids printed directly on glass sheets. Here, ArUco [279] markers are used to encode individual ID numbers for each circular target.

The dimensional accuracy and especially the dimensional stability of the calibration artifact directly influences the quality of the camera calibration. In particular, Zhang's calibration approach assumes planarity of the calibration object. Calibration patterns printed on acrylic or aluminum composite sheets often deform slightly. If the targets are printed on adhesive stickers

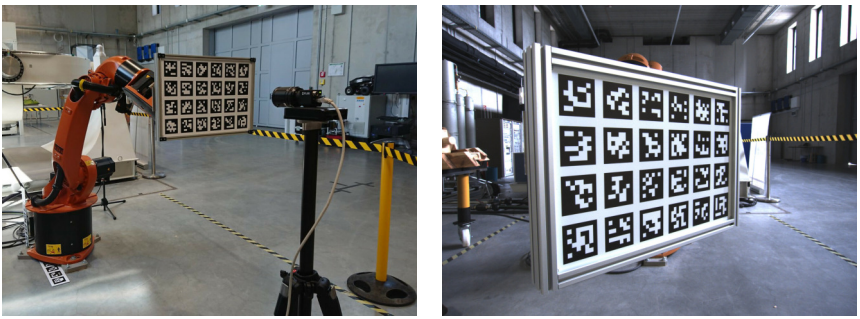


Fig. 3.13: Dataset of Leutert [195] for the evaluation of camera calibration. Left: Experimental setup using KUKA industrial manipulator (Source: [195]). Right: Example image of the dataset (Source: [195]).

and then transferred on the object, the print is slightly stretched due to the flexible material. On the other hand, direct print on glass sheets delivers a rigid calibration structure with good dimensional accuracy. Therefore, for underwater camera calibration and calibration of the laser projector, this work employs fixtures built from glass prints.

3.5.1 Calibration with Bundle Adjustment

If Zhang's method (cf. Section 3.4.1) is applied for camera calibration, the precisely known coordinates of the calibration points are required. Errors of the object points directly affect the quality of the calibration result.

Leutert [195, p.204ff] shows the effect of imprecise calibration targets in a systematic evaluation using images captured from an industrial manipulator for different calibration artifacts. Three planar grids with AprilTag markers and a size of 600 × 400 mm are compared. Two fixtures are built from a print on a vinyl adhesive label, which is glued on an aluminum composite panel or a glass sheet. The third fixture is manufactured using direct print on a glass sheet. Both patterns using adhesive labels show some scale error of the print coordinates because the soft vinyl material is stretched during print and application on the base plate. Moreover, aluminum composite material is not dimensionally stable and is less rigid than the glass sheets. This results in some deformation of the aluminum composite sheet with continued use.

The systematic evaluation in [195] was performed in a laboratory using an industrial robot (KUKA KR16 arc HW). Different planar calibration grids were clamped in a carrier frame and attached to the flange of the robot. The experimental setup and an example image captured by the measurement camera is depicted in Fig. 3.13. The camera was manually focused on the calibration target, and the aperture and focus were fixed. Subsequently, an automatic calibration program was started. This moves the calibration artifact systematically and repeatable to different poses in the camera's field of view. A total of 560 images for each calibration fixture were recorded. The images were captured with a FLIR Blackfly PGE-31S4C-C camera and a Lensagon CVM0411ND lens with a focal length of 4.4 mm to 11 mm. The sensor is a Sony

Table 3.3: Errors of camera calibration with Zhang's method compared with BA using different calibration targets.

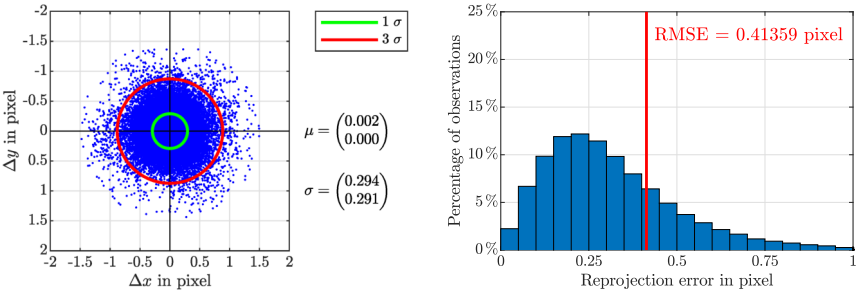
Calibration	**RMS reprojection error** Calibration	**RMS reprojection error** Validation on glass print
Vinyl foil on composite (Zhang)	0.929 pixels	0.414 pixels
Vinyl foil on glass (Zhang)	0.513 pixels	0.336 pixels
Vinyl foil on composite (BA)	0.423 pixels	0.330 pixels
Vinyl foil on glass (BA)	0.340 pixels	0.322 pixels
Direct printing on glass (Zhang)	0.318 pixels	

IMX265 with a resolution of 2048 × 1536 pixels and a sensor format of 1/1.8".

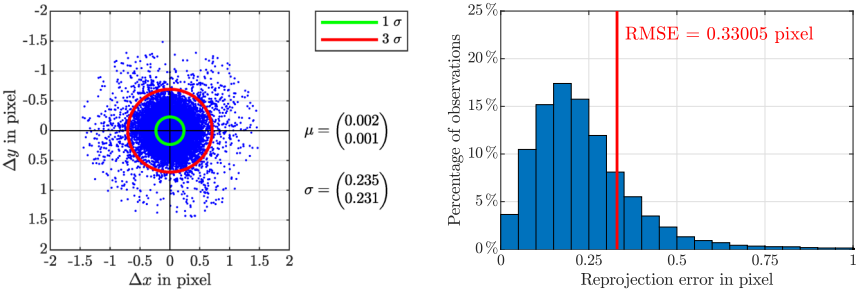
The findings of [195, p.207ff] show that valid calibration results are achieved with all three calibration targets. However, the errors are lowest for the direct printing on glass target due to the higher accuracy of the calibration coordinates.

In BA the object coordinates are estimated simultaneously. This compensates for inaccuracies of the calibration fixture as long as it is rigid and dimensionally stable during the image acquisition. Therefore, this work reprocesses the data set using BA to validate if simultaneous optimization of the object coordinates improves the camera calibration results. DBAT [51] is used for that. To add a scale constraint, the coordinates of a single AprilTag marker in the center of the calibration grid are fixed. The assumption is that the errors introduced by the deformation of the composite sheet and the stretched vinyl foil are small for a single AprilTag. The results of the BA show that the composite sheet has a planarity error of 1.1 mm and the glass sheet shows a planarity error of 0.3 mm. Here, the planarity error is computed as the maximum deviation of the estimated object points from the best-fit plane. The mean deviation of the resulting object points from the print coordinates is 0.71 mm for the vinyl foil on an aluminum composite panel and 0.55 mm for the vinyl foil on a glass sheet. Besides the deformation of the aluminum composite sheet, there is also some scale error due to the stretching of the adhesive stickers during application. This is consistent with the observations of [195].

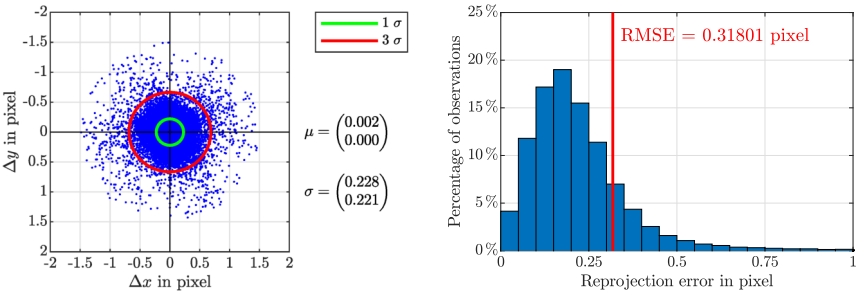
The results of BA are compared with a calibration using Zhang's method. The resulting RMS reprojection errors (cf. Eq. 3.47) are reported in Tab. 3.3. For the computation of the error measurements, any outliers with errors larger than six times the standard deviation were removed. It is assumed that these errors are caused by wrong or inaccurate detections. The middle column shows the RMS reprojection errors during calibration. Using Zhang's method on the vinyl foil on aluminum composite data set a RMS error of 0.929 pixels, and on the vinyl foil on glass data set a RMS error of 0.513 pixels is observed. BA yields lower RMS errors of 0.423 pixels and 0.340 pixels, respectively. While this shows an improvement in the calibration error, the values are not comparable since they are computed on different data sets. Therefore, all calibrations are evaluated using the different calibration targets and methods on the same 560 images captured of the target with direct printing on glass. The RMS reprojection errors are reported in the last column of Tab. 3.3. Here, the experiments show an improvement of the RMS reprojection error from 0.414 pixels (vinyl foil on composite) and 0.336 pixels (vinyl foil on



Calibration using the vinyl foil on aluminum composite target and Zhang's method.



Calibration using the vinyl foil on aluminum composite target and BA.



Calibration using direct printing on glass target and Zhang's method.

Fig. 3.14: Distribution and histogram of the reprojection errors computed on 560 images of the direct printing on glass target with an AprilTag grid.

glass) down to 0.330 pixels (vinyl foil on composite) and 0.322 pixels (vinyl foil on glass). This is comparable to the RMS error of 0.318 pixels produced by applying Zhang's method directly on the validation data set using the direct printing on glass targets.

To analyze the errors more closely, Fig. 3.14 shows the distribution of the residuals and the histogram of the residuals in image space. The residuals are computed on the glass print data set. The top row shows the result of the calibration using Zhang's method on the composite sheet data, the middle row shows the result using BA on the composite sheet data, and the bottom row shows the result of the calibration on the direct printing on glass data as a comparison. Using BA with adjustment of the object coordinates of the aluminum composite target, the magnitude of the residuals and the spread of the residuals is reduced. For comparison, the additional results for the vinyl foil on glass pattern are found in Appendix B.1.

The result shows that BA is able to improve imprecise calibration point coordinates and improve camera calibration results. However, for underwater applications, it is difficult to acquire unbiased, high-quality calibration image data sets. Hence, self-calibration is not always well-conditioned. Moreover, the refraction introduces additional unknowns in the calibration. Therefore, besides performing self-calibration of the camera using BA, it is sensible to work with calibration fields with known 3D coordinates and calibrate the structures in air using BA before deploying them in the water in order to provide an accurate reference.

3.5.2 Planar and 3D Calibration Structures

From the literature it is known that 3D structures provide more consistent and accurate calibration results [295]. Boutros et al. [58] compare a cuboidal calibration artifact with A3 and A4 planar chessboard patterns. They found that the cuboidal calibration artifact displayed improved accuracy and precision compared to the planar patterns. However, this evaluation does not only include the difference between planar and 3D calibration fixtures. The three calibration artifacts have different sizes and three different calibration image sets are compared. Hence, the coverage and distribution of the object points and the quality of the calibration image sets influence the end result as well.

Therefore, this work proposes an experiment where the calibration on a planar structure and a 3D structure is compared on the exact same image observations. To achieve this, the calibration pattern depicted in the top-left image of Fig. 3.12 is employed. It is constructed from two planar grids of AprilTag markers that are mounted approximately orthogonal to each other. This way, every image is evaluated as the observation of a single 3D structure or two separate planar calibration boards. For the evaluation, the exact same image coordinates are used. The only difference is that for the 3D calibration, the relative transform between the two calibration boards is applied as an additional constraint. The relative transform between the two calibration boards is calibrated in air using BA on a set of 629 images.

The experiments perform calibration on a set of 600 underwater images. First, a calibration using Zhang's method using the 1200 planar patterns observed in the images is performed. Then, a 3D calibration based on the 600 observations of the 3D calibration pattern is computed. The RMS reprojection error (cf. Eq. 3.47) computed for the two planar calibration boards and the full 3D structure is reported.

The reprojection error computed on the planar patterns is very similar for both calibrations.

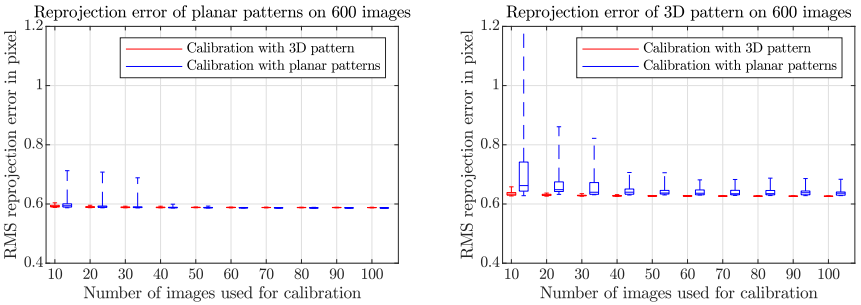


Fig. 3.15: Comparison of the reprojection error of a calibration using planar or 3D targets.

The 2D calibration yields a RMS reprojection error of 0.586 pixels, while the 3D calibration produces a RMS reprojection error of 0.587 pixels. The reprojection error computed on the 3D structure is slightly lower for the 3D calibration. The 2D calibration results in a RMS reprojection error of 0.637 pixels, while the 3D calibration yields a RMS reprojection error of 0.625 pixels. In summary, the achieved reprojection error is comparable for the 2D and 3D calibration given the complete set of 600 calibration images.

A second experiment examines how fast 2D and 3D calibration converges to a low error. To do this, a random subset of n images is drawn from the complete set of calibration images. Then, calibration is performed on this subset of images and the error of the resulting calibration is computed on the complete set of 600 calibration images. The 2D and 3D calibration is performed for the exact same set of images and this is repeated 20 times for each subset of n images. Fig. 3.15 shows a boxplot of the RMS reprojection errors for different numbers of images used for the calibration. The left image shows the RMS reprojection error computed on the planar pattern, and the right image shows the RMS reprojection error evaluated on the 3D pattern. The results of the 3D calibrations are visualized in red, while the results of the 2D calibration are depicted in blue. For small subsets of calibration images, the spread of the error measure is much higher for the 2D calibration than the 3D calibration. Moreover, the maximum observed RMS reprojection error is significantly higher for the 2D calibration if the number of calibration images is smaller than 30. Hence, given a small set of calibration images, the calibration performed using the 3D structure results in more reliable results and lower errors. In summary, convergence of the calibration is faster for the 3D calibration, i.e., fewer images are needed to achieve the same error.

This is also visible in the trend of the calibration parameters in relation to the number of images used for the calibration. Fig. 3.16 visualizes the calibration results in relation to the size of the calibration set. The top row shows boxplots of the focal length f_x and f_y, and the middle row shows boxplots of the coordinates of the principal point p_x and p_y. Here, we see again that the spread of the results is higher for the 2D calibration than the 3D calibration. The bottom row shows the trend of the different radial and tangential distortion parameters in relation to the number of calibration images. For visualization purposes, only the mean of the

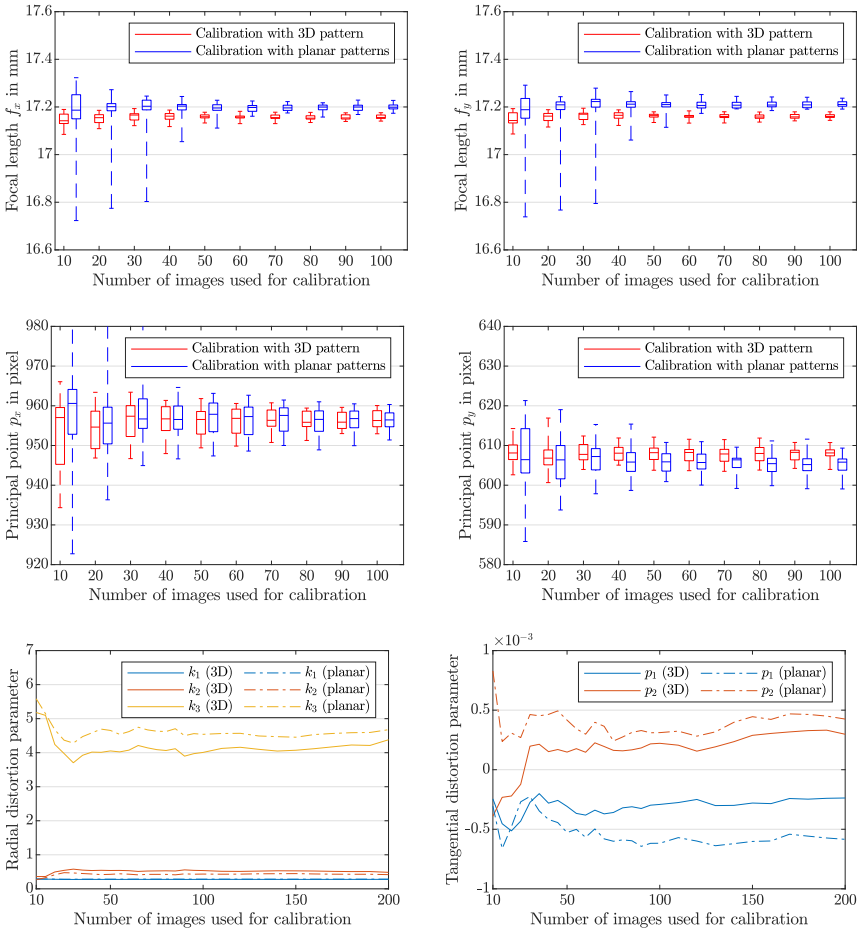


Fig. 3.16: Comparison of the variation of the model parameters of a calibration using planar or 3D targets.

different calibrations is reported instead of a boxplot. The change of the parameters is significant for small sets of calibration images. Only after 50 to 100 images the change of the parameters become less dominant. 2D and 3D calibration converge to slightly different parameters. However,

this does not tell us anything about the quality of the individual parameters since we always need to evaluate the complete camera parameter set.

For comparison purposes, the results of this experiment reproduced on a calibration set captured in air instead of water are found in Appendix B.2. From the results, the same conclusions are drawn: Given a sufficient set of calibration images, both 2D and 3D calibration yield low errors. However, the 3D calibration estimates a camera parameter set, which yields a low RMS reprojection error on the complete set of 600 images, more reliably if the number of calibration images is small, i.e., below 50 images.

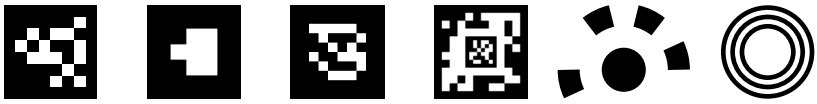
Note that in the presented experiment, the subset of images for calibration are randomly drawn. This does not compare to an expert calibration with a manual selection of good images. If the images of a planar target are carefully captured from different angles, it is possible to achieve good results with the 2D calibration on a small set of images. However, it is often difficult to capture high-quality calibration datasets in water with a large number of images and large camera pose variance. Therefore, employing 3D calibration patterns for underwater camera calibration provides more reliable and accurate results.

3.5.3 Fiducial Marker Detection

Calibration patterns are often constructed from circle grid, chessboard, or square grid patterns. The centroid of circular markers is extracted with subpixel accuracy using image moments, correlation, or by tracing the edge of the marker and estimating the parameters of the circle, or more general, the ellipse [202, pp. 482-494]. For high-accuracy applications the eccentricity due to the perspective distortion needs to be considered [202, pp. 260-262]. Chessboards are among the most popular patterns for camera calibration due to the support in readily available open-source software, such as Bouget's camera calibration toolbox for Matlab [56] and the camera calibration algorithms implemented in the OpenCV library [60]. Detected positions are refined using, for example, the subpixel corner localization approach proposed by [127]. A disadvantage of regular patterns, such as chessboard patterns, is that it is difficult to sample calibration points close to the image corners. Especially for the calibration of lenses with high distortion, it is difficult to capture images with the chessboard captured completely in the image with calibration points close to the image corners. This is not the case for coded targets like coded square tags [269] because the object points are still detected if part of the complete pattern is outside the field of view.

Additionally, there are more specialized target marks for optical metrology. For example, in combination with terrestrial laser scanning or in medical photogrammetry applications with infrared cameras, retro-reflective markers are used to achieve a higher contrast between the marker and the background. If the markers have to be detected over a wide range of view angles, ball markers with or without reflective coating are employed. A typical example is the spherical markers with reflective tape used for human skeleton and motion tracking [237]. For camera calibration, planar markers are typically employed since a restricted viewing angle in the order of $\pm 45^\circ$ is often sufficient.

Distinguishable IDs are often encoded using matrix barcodes, such as QR codes or the fiducial systems of AprilTag [252], ARToolkit [179] ARTag [121], and ArUco [140, 279]. In turbid water conditions, the AprilTag detector was found to perform best in experiments in the liter-



(a) (b) (c) (d) (e) (f)

Fig. 3.17: Examples for common fiducial systems that are also applied for camera calibration: (a) AprilTag [252], (b) ARToolkit [179], (c) ARTag [121], (d) ArUco fractal marker [279], (e) Ring code targets [285], (f) CCTag [77].

ature [104]. The detection speed of the AprilTag detector was significantly improved in recent versions [332]. 2D Barcodes are also added to circular targets for identification as well as circular ring codes [285]. Ring coded targets are supported by popular photogrammetry software packages and have been also succesfully applied for underwater photogrammetry [107]. Alternatively, identification numbers are also encoded using concentric circles [77]. Furthermore, combinations of different marker systems are used, such as coded tags combined with circle targets or chessboard corners [38]. Improvements to chessboard detection accuracy by applying the Radon transform and extracting the corners from the response map are proposed by [110]. To increase the information for centroid detection [287] uses star patterns, which the authors think of as a generalization of chessboard patterns, as feature points. Examples of these fiducial systems are depicted in Fig. 3.17.

The results in this work were created using AprilTag, chessboard patterns, and circle targets. In most of the experiments, AprilTag fiducial markers were employed due to the robustness of the detector, fast decoding of the tags, and availability of an excellent open-source implementation, which is easy to integrate in custom applications.

3.5.4 Subpixel Detection Accuracy

Fiducial systems should ideally provide methods to interpolate the target position in an image with subpixel accuracy. Here, circle targets, chessboard corners, AprilTag and ArUco markers are compared, which are all regularly applied for camera calibration. The library OpenCV 4.2.0 [60] is used for detecting circles, chessboard corners and ArUco markers and compared with version 3.1.4 of the AprilTag detector [187, 252, 332].

AprilTag achieves subpixel accuracy by fitting lines to the gradient edges. For this comparison, the option of AprilTag to refine corner positions is enabled, which is slower but more accurate. Circles are detected very accurately because the whole outline of the circle is traced, and the centroid is interpolated based on the complete circle. OpenCV provides different position refinement methods for ArUco based on corner subpixel refinement or contour tracing. Compared to single matrix tags, a chessboard has the advantage that the position of a single corner is encoded by black-to-white transitions in two directions, which is accurately localized with the method of [127].

The localization error is compared based on synthetically generated images with a single tag with four corner positions. For all different markers, the corners have a distance of 80 pixels. The ArUco and AprilTag detectors are tested on the same images using a common codebook

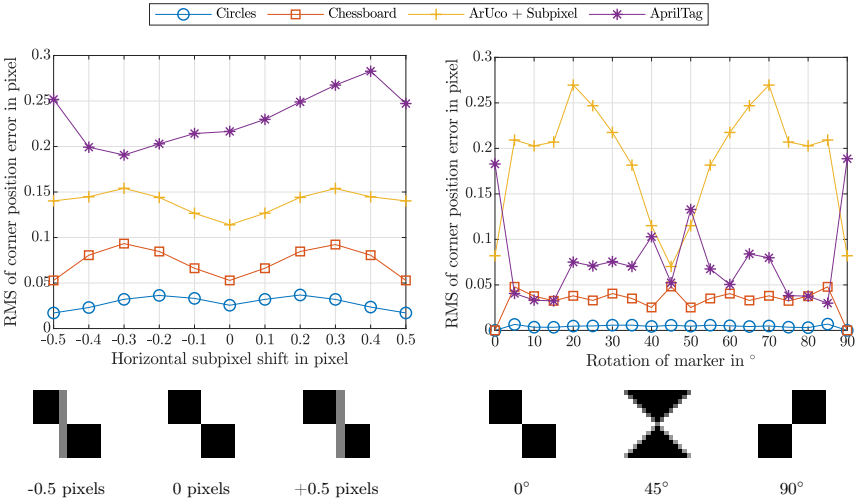


Fig. 3.18: Localization error of corner positions for circle targets, chessboard corners, ArUco and AprilTag markers. Left: RMS error in pixels depending on horizontal shift for images with subpixel shift in horizontal and vertical direction. Right: RMS error in pixels depending on rotation of the marker for images with horizontal/vertical subpixel shift and rotation.

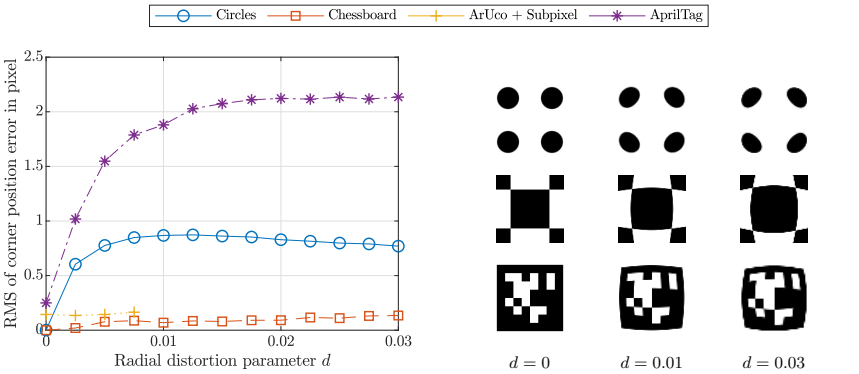


Fig. 3.19: Localization error of corner position for images with radial distortion. Left: RMS error in pixels depending on the radial distortion parameter. Right: Examples of distorted tags for different radial distortion parameters d.

with tags with 6x6 bits. The tags have a size of 80 pixels without the white border. Similarly, a chessboard pattern is generated with a square size of 80 pixels, and the experiments evaluate the four corner positions of one square. To test the circle detector, four circles are generated with a diameter of 40 pixels and a distance of the center positions of 80 pixels. The different generated markers are visualized in the right images of Fig. 3.19. From top to bottom: circle targets, chessboard pattern and 6x6 bit matrix tags. The test images are generated by transforming or filtering the marker images. For subpixel rendering a bilinear interpolation is applied.

First, this work looks at subpixel shifts and rotations of the images and examines how well the individual marker detectors are able to localize the corner positions. To do this, the images are shifted by ± 0.5 pixels in horizontal and vertical directions in increments of 0.1 pixels and subsequently are rotated by 0° to 90° in increments of 1°. Hence, a total of 11011 synthetic images are generated.

The RMS error for the four corner point positions in the generated 11011 images is 0.029 pixels for the circle markers, 0.075 pixels for the chessboard corners and 0.234 pixels for AprilTag markers. The result for ArUco depends on the selected subpixel refinement. Applying no subpixel refinement results in a RMS error of 0.673 pixels, contour-based refinement in a RMS error of 0.747 pixels and corner subpixel refinement in a RMS error of 0.140 pixels. Therefore, the following experiments look only at ArUco with the corner subpixel refinement method since it empirically produces better results than the other refinement approaches.

The left image in Fig. 3.18 shows a plot of the errors depending on the horizontal subpixel shift for the subset of images that only apply horizontal/vertical shift and no rotation. This represents the subpixel error for images with the edges of the tags/chessboard aligned with the axes of the image's pixel grid. A subpixel shift of 0 pixels denotes here that the corner position is exactly located at the edge between two pixels, while a subpixel shift of ± 0.5 pixels means the corner position lies in the center of a pixel. Here, circles and chessboard corners perform better. This is expected since the edge positions of the matrix coded tags are less well defined by only the black-to-white edge. Centroid estimation of the circles or interpolation over the four quadrants of the chessboard corner is expected to be more accurate. Lower errors are observed if the corner position is exactly in the center or at the edge of a pixel. The error produced by AprilTag seems to be non-symmetrical if the images are shifted by the same amount to the left or to the right. All different marker detectors yield errors of less than half a pixel on all images.

If rotations are added to the transformations, the errors of the AprilTag detector decrease while the errors for ArUco with corner subpixel refinement increase. The right image in Fig. 3.18 shows the RMS error of the corner positions for horizontally and vertically shifted images for different rotation angles. Again, circles and chessboard corners provide the best subpixel localization accuracy on the transformed images.

Moreover, image distortion strongly affects localization accuracy. To test this, images are generated with different quadratic radial distortion

$$r' = r + d \cdot r^2 \quad , \tag{3.51}$$

where d is the distortion parameter, r is the distance of the pixel from the image center and r' is the distorted radial distance. The image center is aligned with the center of the marker for images without subpixel shift or rotation. The right image of Fig. 3.19 visualizes the effect of different distortion parameters on the different markers. A total of 5250 test images with different subpixel

shifts, rotations and distortions were generated. The left image of Fig. 3.19 visualizes the RMS error of the corner localization depending on the radial distortion parameter. The chessboard pattern seems to be largely unaffected by the applied distortion. ArUco with corner subpixel refinement also produces good results for small distortion. However, for distortions of $d > 0.0075$ the ArUco detector fails to detect the marker. The localization error of the circles increases with higher distortion. This is explained by the inhomogeneous distortion of the circles by the radial distortion. Therefore, the center of the ellipse does not represent the correct position anymore and the localization error increases. Similarly, the line fitting based corner position refinement of AprilTag exhibits problems with accurately localizing the corners in images with high distortion. In principal, some of these problems are mitigated by correcting for distortion. However, this is only possible if an estimate for the distortion is available.

3.5.5 Effects of Image Degradation

Another important issue for target point detection is robustness to image degradation. Underwater it is often difficult to capture clean images of the calibration target. Achieving good lighting is challenging, which results in incorrectly exposed images, inhomogeneous lighting and reduced contrast. The sharpness of the images may be reduced due to the limited depth of field. Moreover, particles in the water or dirt on the calibration targets significantly degrade underwater images and influence detection accuracy.

To test for different effects of image degradation, contrast is reduced synthetically, the images are blurred, or noise is added. The experiment creates a test set of 250 images with different subpixel shifts and rotations, applies filters to these images, and examines the resulting localization error of the individual detectors. The results are visualized in Fig. 3.20.

The top left image shows the RMS error of the corner position depending on the image's contrast. A contrast of 100 % means that the full spread of the 8-bit image between the values 0 and 255 is used to encode the white background and the black marker pixels. A contrast of 50 % means that only half of the range is used to separate foreground and background. Circle and chessboard detectors are mostly unaffected by reduced contrast. However, the corner detection accuracy for the matrix code tags is reduced since contrast reduction affects the edge gradient.

The top right image shows the effect of image blur on the localization. Here, the images are smoothed with a Gaussian kernel with a standard deviation $\sigma_b \in [0; 2]$. AprilTag and ArUco fail to localize the blurred corners accurately, and the error increases with higher standard deviation. Circle and chessboard corner targets are still localized accurately. With increasing blur, some of the methods also fail to detect the markers. This is partially intended since some of the methods check for edge contrast to remove bad detection results.

In order to examine the effects of noise on the detection, Gaussian noise with zero mean and variance σ_n^2 or salt and pepper noise with a given noise density is added. The bottom left image shows the effect of Gaussian noise, and the bottom right image the effect of salt and pepper noise. In both cases, the circle detector is more robust against noise compared to the other detectors and provides the lowest corner position errors. However, detection run-time significantly increases with increasing noise. Gaussian noise above a variance of $\sigma_n^2 > 0.02$ causes the ArUco detector to fail on some of the test images. Salt and pepper noise heavily affects the chessboard detector and AprilTag. The AprilTag detector fails on some images for noise densities

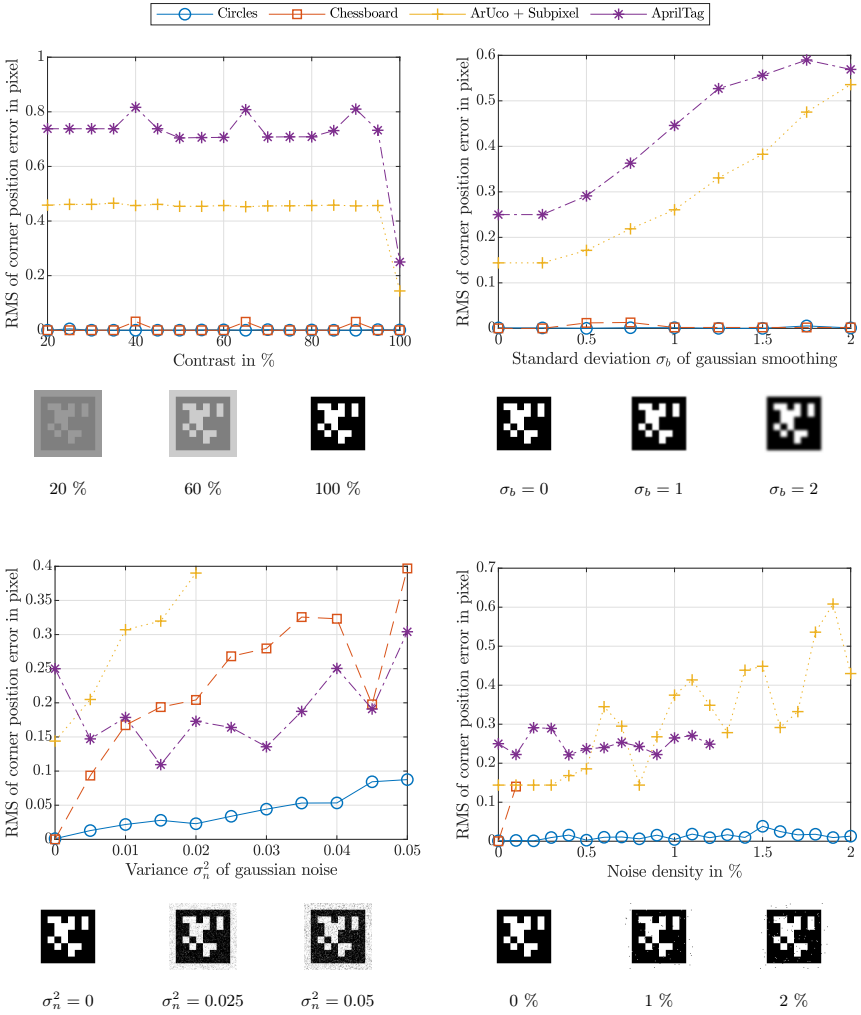


Fig. 3.20: Localization error for different types of image degradation. Top left: RMS corner position error depending on image contrast. Top right: RMS corner position error depending on Gaussian blur. Bottom left: RMS corner position error depending on additive Gaussian noise with zero mean. Bottom right: RMS corner position error depending on added salt and pepper noise.

larger than 1.25 %. The chessboard detector is unable to reliably detect all corners on all test images for noise densities above 0.1 %. This is explained by the fact that the subpixel method of [127] is not entirely robust to outliers. If noise pixels are present inside the window for the subpixel approximation, the extracted center is displaced. This means for very noisy images, it is beneficial to apply some denoising or smoothing of the images prior to running the detectors.

3.6 Analysis of Approximation Errors

First, simulation results are considered for the analysis of the approximation errors of implicit vs. explicit modeling of the refraction. Based on the ray tracing model of multi-layer refraction, it is investigated in simulation how well Brown's distortion model describes the ray geometry. In a second step, the results are compared on a large set of images of a calibration fixture captured in water. The errors between the standard camera parameter set using Brown's distortion model and the proposed physical-geometric model are inspected. Moreover, the results of the proposed calibration approach are compared with a calibration performed solely in air. The experiments demonstrate that an in-air calibration is successfully transferred to underwater imaging if the refraction at the housing is taken into account.

3.6.1 Simulation Result

For the analysis of the errors in simulation, a camera is considered with parameters similar to the underwater cameras employed in this work: An ideal distortion-free camera is simulated with a 12.5 mm lens, a sensor resolution of 1920 × 1200 pixels and a pixel size of 5.86 µm. In front of the camera, a glass window with a thickness of 2 cm is placed at a distance of 3 cm in front of the entrance pupil of the lens. In simulation, a refractive index of 1.33 for water and 1.5 for the glass window is applied. For the orientation of the glass window, the experiments consider the case that the optical axis crosses the media interface perpendicular as well as the effects of a rotation of the glass window around the yaw axis of the camera.

The simulation is performed in the following way: First, 9000 points are evenly sampled in image space (1 point per 16 pixels). These points are then projected into 3D space according to the pinhole camera model and the back projection model described in Section 3.3.1. A depth range up to three meters in front of the camera is sampled in 10 cm increments. Second, a Brown model is fitted to the projection of these 3D points on the image coordinates. This is done using non-linear optimization by minimizing the reprojection error. Note that the camera's position is not fixed for this optimization process. The virtual camera center is allowed to move relative to the simulated view port, such that the pinhole model with Brown parameters best fits the ray geometry. Finally, the projection according to Brown's distortion model is compared with the simulated projection based on ray tracing. The RMS reprojection error is computed on the complete point set since the simulation has perfect knowledge and no outliers.

The Brown parameters are optimized such that the reprojection error over the range of 0 - 3 m is minimized. As we have seen in theory in Section 3.2.3, the radial distortion is dependent on the measurement distance, and the largest errors are observed close to the camera. Fig. 3.21 shows a plot of the maximum residuals in image space of the simulated points. These are the residuals of the points directly in front of the view port window. The left image shows the residuals for

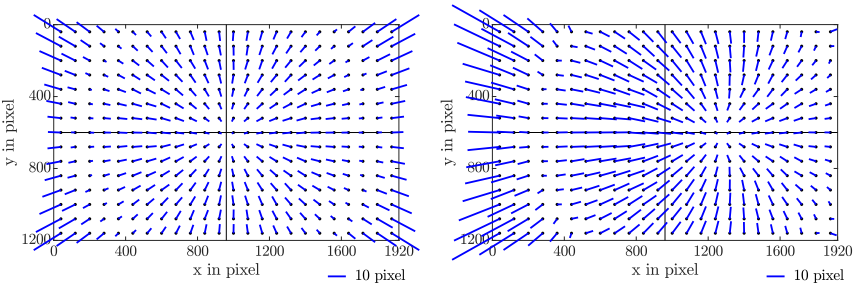


Fig. 3.21: Plot of sampled maximum residuals of the Brown model in image space. Left: Orthogonal view port. Right: View port rotated by 5°.

the view port orthogonal to the optical axis. Here, a systematic residual error is observed in the radial direction. The optimized distortion parameters do not fully describe the refractive effects. The RMS reprojection error over the range of 0 - 3 m is 0.977 pixels with a maximum residual of 18.00 pixels. The right image shows the image residuals for a view port rotated around the yaw axis by 5°. In this case, a systematic error is observed in the radial direction as well as a decentering effect. The residuals are not distributed evenly around the image center since the camera looks through the refractive layers at an angle. The RMS reprojection error over the range of 0 - 3 m is 1.552 pixels and the maximum image residual is 30.28 pixels. In summary, very large image residuals are observed in simulation. The Brown parameters do not fully absorb the refraction effects for the simulated flat port camera.

Fig. 3.22 visualizes the RMS reprojection error depending on the measurement distance. In the left image, the black solid line shows the error curve for a calibration that minimizes the error for the full range of 0 - 3 m. The RMS reprojection error over the range of 0 - 3 m is 0.977 pixels. The RMSE is 4.880 pixels directly in front of the view port. Then the error decreases to a minimum at approximately 0.6 m. For measurements at larger distances, the error increases. The RMS projective error of the Brown model approximation is 0.433 pixels at 3 m and 0.497 pixels at 6 m.

In practice, optical measurement systems are often optimized for a certain depth range. For example, by fitting a model to the simulated observations at a far distance, it is possible to reduce the modeling error for these measurements while sacrificing accuracy for close-range measurements. The red dot-dashed line in the left image of Fig. 3.22 shows the error curve for a Brown model fitted to the simulated observations at the measurement distance of 1.5 - 3 m. While this increases the errors for measurements below 0.8 m, this results in a good fit for measurements at a distance of $\geq$ 1.5 - 3 m. For example, at 2 m distance now an RMS reprojection error of 0.008 pixels is observed compared to 0.382 pixels for the Brown model optimized on the full range. The error of the Brown model approximation is only negligible for a certain depth and increases again with distance. For example, at 6 m distance, the RMS reprojection error increases again to 0.107 pixels for the calibration on the range 1.5 - 3 m.

The blue dotted curve shows the errors for a Brown model fitted to the observations closer

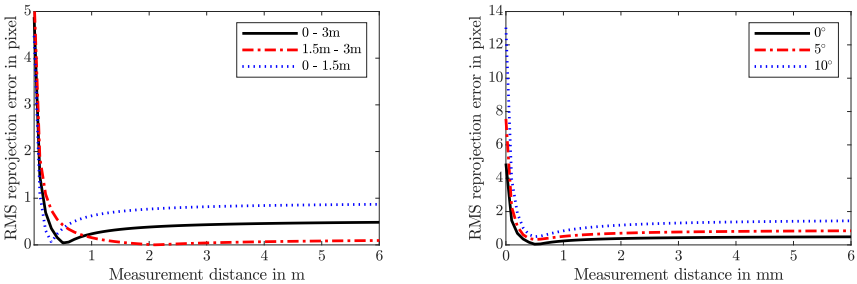


Fig. 3.22: RMS reprojection error depending on the measurement distance. Left: Difference between fitting the Brown model to the full range 0 - 3 m, to the far range 1.5 - 3 m or to the close range 0 - 1.5 m. Right: Error curve for different orientations of the view port (calibration distance 0 - 3 m).

than 1.5 m in front of the view port of the camera. This improves the error for close-range measurements but approximately doubles the error for measurements at a distance larger than 1.5 m.

The right image of Fig. 3.22 shows the RMS reprojection error depending on the measurement distance for different rotations of the view port window around its yaw axis. In this case, the Brown parameters are optimized to minimize the error for the full range of 0 - 3 m. Here, the experiment finds that the error significantly increases with a larger rotation of the window over the entire measurement range. The refraction effects of an angled view port are not absorbed well by the Brown model's radial and tangential distortion parameters.

In summary, if the measurement range is limited to points that are more than 1.5 m away from the camera, a Brown model is found in the simulation that describes the ray geometry well with low errors for an orthogonal flat port with the given parameters. The refraction effects that occur for measurements at close range, as well as imaging through an angled view port, is not described well by the Brown parameters.

3.6.2 Underwater Calibration Result

In this section, the proposed refractive model and calibration approach are applied to a set of underwater images. The images were captured with a custom-built underwater camera. The camera is a FLIR Grasshopper3 GS3-U3-23S6M-C industrial vision camera with a monochrome Sony IMX174 sensor. The sensor features a resolution of 1920 × 1200 pixels, a pixel size of 5.86 µm and a sensor format of 1/1.2". The camera is integrated into an off-the-shelf underwater housing with a 19 mm thick acrylic flat port.

Fig. 3.23 shows the handheld underwater camera during calibration in the left image and an example of the calibration data set of 640 images in the right image. First, the camera is calibrated in air on a set of 640 images using Brown parameters. This calibration is performed with the camera mounted in the underwater housing. Therefore, the double refraction at the view port in air is absorbed in this model. Then, the camera is calibrated in water using the

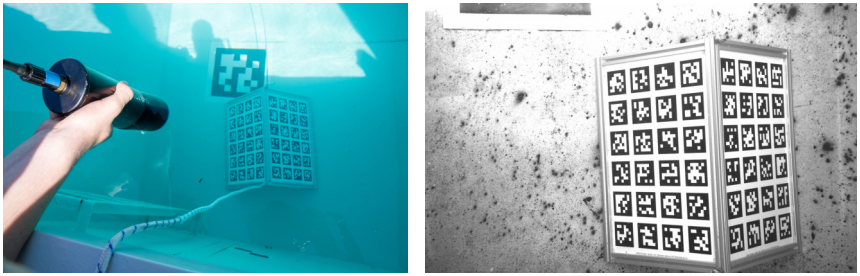


Fig. 3.23: Underwater camera calibration setup. Left: Handheld underwater camera with calibration fixture. Right: Example image of the calibration data set.

Table 3.4: Calibration result of Brown model in air, Brown model in water and refractive model in water.

Parameter	**Brown model** air	**Brown model** water	**Refractive model** water
Camera			
(f_x, f_y) in pixels	$(2140.01, 2138.66)$	$(2857.29, 2856.41)$	$(2147.63, 2146.96)$
(p_x, p_y) in pixels	$(938.31, 619.96)$	$(936.68, 614.76)$	$(934.92, 615.84)$
k_1	-0.11661	0.17705	-0.11068
k_2	0.26094	1.06306	0.30910
k_3	0.00293	0.00233	0.00192
p_1	-0.00207	-0.00352	-0.00264
p_2	-0.20044	-0.89965	-0.29905
Housing			
d_0 in mm	-	-	32.45
d_1 in mm	-	-	19.00
n_x	-	-	-0.00029
n_y	-	-	-0.00028
n_z	-	-	-0.99999
(μ_0, μ_1, μ_2)	-	-	$(1, 1.5, 1.33)$
Calibration error			
RMSE in pixels	0.287	0.263	0.261

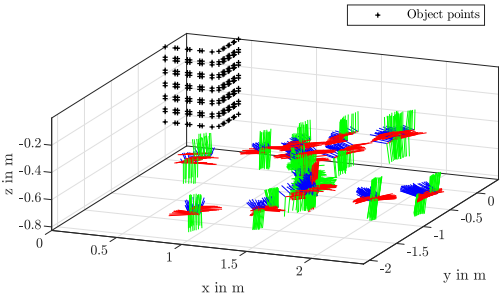


Fig. 3.24: Camera poses for a calibration set of 640 images.

Brown model and the proposed refractive model.

The resulting parameters are summarized in Tab. 3.4. The RMS reprojection error for the Brown model in water is 0.263 pixels, and 0.261 pixels for the refractive model.

Note that this calibration set is captured handheld from the surface of an approximately 2 m deep water tank. Therefore, the calibration images are captured at a distance of 1.5 - 2.5 m from the calibration object. The camera poses of the calibration set are visualized in Fig. 3.24.

In this case, the distance to the object is large compared to the distance between camera and view port. Hence, a small difference between the Brown model result and refractive calibration is expected. The observed RMS reprojection errors of the calibration results are similar, see Tab. 3.4. This is consistent with the simulation results in Section 3.6.1.

Fig. 3.25 shows a plot of the distortion model for the in-air calibration, the Brown model in water and the refractive model in water. The distortion parameters of the in-air and underwater calibration using the refractive model are similar. The underwater calibration with the Brown model deviates strongly, especially the radial distortion parameters k_1 and k_2. This is more clearly shown in Fig. 3.26, where the radial and tangential components of the distortion model are plotted separately for the three calibrations. The underwater calibration using the Brown model absorbs the refraction effects primarily in an increased focal length and the radial distortion model.

The camera and distortion parameters for the in-air calibration and refractive model are similar. This suggests that the refraction effects are described well by the physical-geometric model using the estimated housing parameters. Some variation is expected because the individual parameters are correlated, and the distortion parameters absorb some of the residual errors that are not described in the refractive model, e.g., inhomogeneities of the view port or inexact estimates of the housing parameters. Moreover, the camera was calibrated in-air mounted in the housing. Thus, this model contains the refraction effects of the view port in air. Consequently, it is expected that a small shift of the principal point and a change of the parameters between the in-air model and the refractive model occurs.

The resulting RMS reprojection error of the refractive calibration is 0.261 pixels. The 640 images show 118212 point observations. From these, a total number of 225 outliers were removed

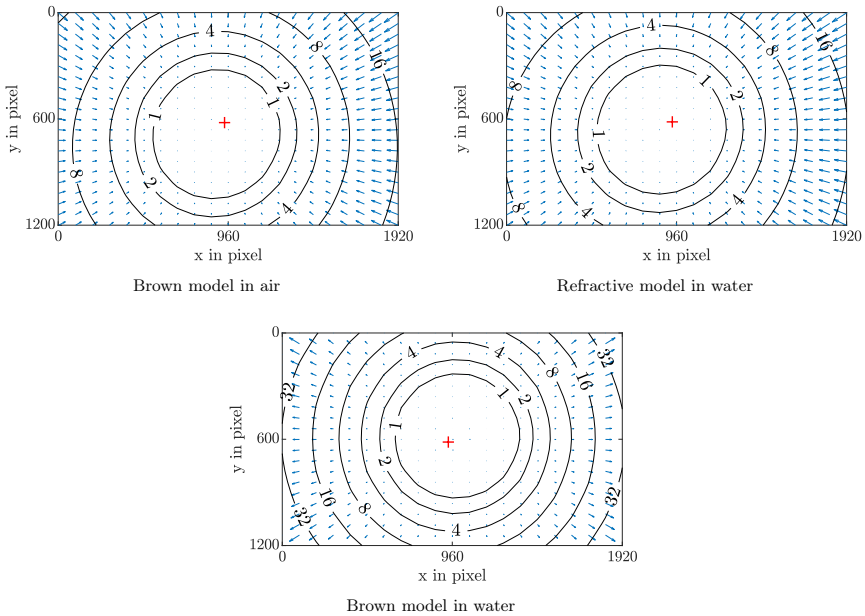


Fig. 3.25: Complete distortion model for three different calibrations: Calibration of the camera in air, underwater calibration with refractive model applied, and underwater calibration without refractive model. The red cross visualizes the principal point. The contour lines denominate the magnitude of the shift caused by the distortion model in pixels.

with residuals larger than six times the standard deviation. The RMS error is computed on the remaining 117987 point observations. In Fig. 3.27 the histogram of the reprojection errors in image space is depicted.

Fig. 3.28 shows the distribution of residuals in image space. The left image shows a sub-sampled plot of the largest residual vectors. Residuals with a norm of more than 1 pixel are plotted in red. The right image shows a plot of the mean residuals on a regular grid with a cell size of 80 × 80 pixels. In both of the plots, no systematic pattern of the residual is observed.

Fig. 3.29 depicts the distribution of the residuals. The left image is a scatter plot of the x/y-components of the residual vector. The 1σ and 3σ ellipses of the standard deviation σ of the residuals are plotted in green and red color. Here, the spread is similar in x- and y-direction. The right image shows a scatter plot of the residuals along the radial direction from the image center. The mean reprojection error is approximately constant and increases towards the image corners. However, this is also an effect of the small number of point observations close to the image corners.

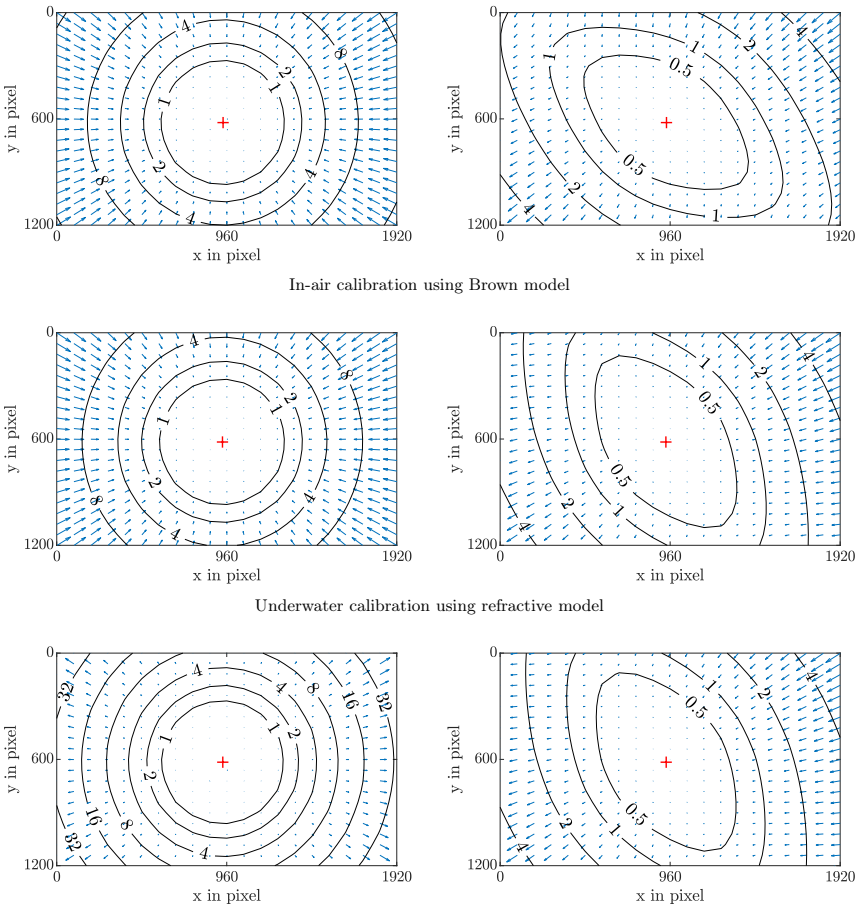


Fig. 3.26: Radial component (left) and tangential component (right) of Brown's distortion model for three different calibrations: Calibration of the camera in air, underwater calibration with refractive model applied, and underwater calibration without refractive model. The red cross visualizes the principal point. The contour lines denominate the magnitude of the shift caused by the distortion model in pixels.

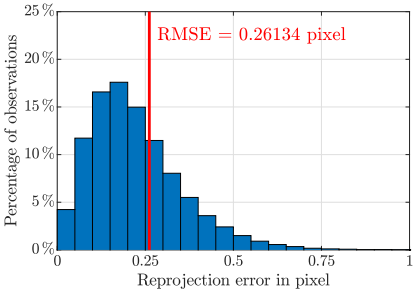


Fig. 3.27: Histogram of the reprojection errors computed for a calibration set of 640 images.

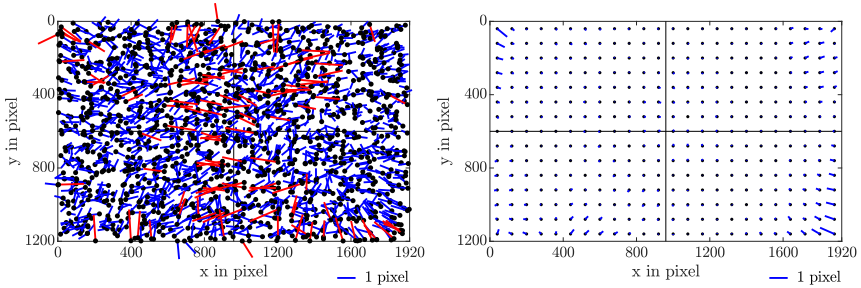


Fig. 3.28: Distribution of residuals in image space computed for a calibration set of 640 images. The lengths of the residual vectors are scaled by a factor of 100 for visualization purposes. Residuals with a norm of larger than 1 pixel are plotted in red. Left: Plot of sampled residuals in image space. For visualization purposes, only the four largest residuals in every 80×80 pixels cell are shown. Right: Mean residuals on a regular grid with a cell size of 80×80 pixels.

Overall, a valid underwater calibration is achieved using the refractive model. Small residuals comparable to an in-air calibration are observed.

3.6.3 Comparison of A priori and In situ Calibration

To compare the results, the calibration is validated on a second data set comprising 74 images of the calibration fixture. The structure is captured at distances of 0.75 to 6 m in front of the camera. Three different calibrations are compared:

- an a priori calibration performed solely in air,
- an underwater calibration using the refractive model, and
- an underwater calibration using the Brown model.

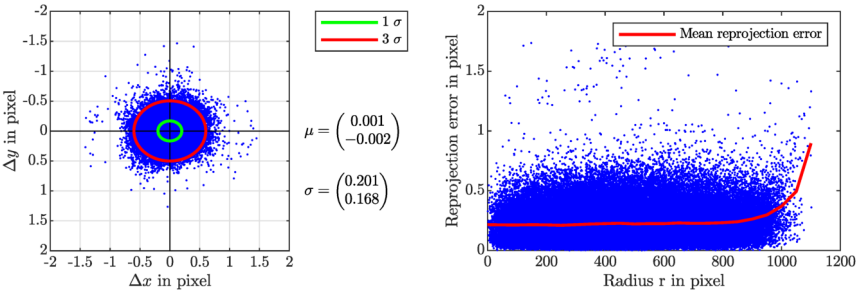


Fig. 3.29: Distribution of residuals computed for a calibration set of 640 images. Left: Scatter plot of the residuals with mean and standard deviation. Right: Scatter plot of the reprojection errors dependent on the radial distance from the image center.

For the a priori calibration, the camera parameters calibrated in air are employed, and the refractive model is applied with the estimated housing parameters. The underwater calibrations are optimized on the calibration data set of 640 underwater images.

The a priori calibration yields a RMS reprojection error of 0.344 pixels on the test data set. The refractive underwater calibration shows an RMSE of 0.320 pixels, and the Brown model has a RMSE of 0.341 pixels. Fig. 3.30 shows the histograms of the reprojection errors for the three different calibrations. Plots of the distributions of the image residuals are found in Appendix B.3.

Overall, all three calibration approaches lead to valid camera parameters. The RMSE on the test data set is comparable with the refractive model, resulting in the smallest error.

From the simulation results we know that the difference between the Brown model and refractive model is expected to be larger for close range measurements. Hence, the RMSE is computed for observations of the calibration pattern closer than 1 m. Here, a larger difference is observed: The Brown model yields an RMS reprojection error of 0.813 pixels, the a priori calibration a RMSE of 0.704 pixels and the refractive model a RMSE of 0.683 pixels. Based on the simulation results, larger discrepancies between implicit and explicit modeling are expected for measurements even closer to the camera. Besides inaccurate modeling of the refraction effects, the error is also higher for the close-range measurements since the camera is focused at a distance of approximately 1.5 m and therefore the images start to show a slight blur below 1 m measurement distance.

If the error is computed for all calibration patterns with a distance larger than 1 m, similar errors are observed. The Brown model has an RMS reprojection error of 0.307 pixels, the RMSE of the a priori calibration is 0.309 pixels, and the refractive calibration yields a RMSE of 0.285 pixels.

In summary, for the close-range measurement, a small improvement of the RMSE is achieved using the refractive model. For measurements distances larger than 1 m, the models perform similarly on the test data. Based on the estimated housing parameters, the in-air calibration is successfully transferred to underwater conditions by applying the refractive model.

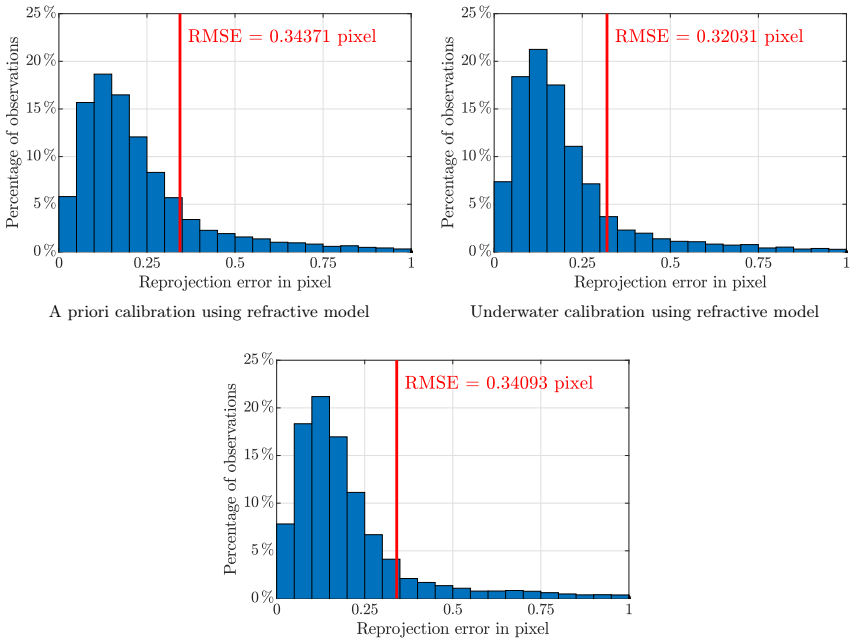


Fig. 3.30: Histograms of the reprojection errors of the validation of the calibration results on 74 images captured at a distance of 0.75 to 6 m in front of the camera.

3.7 Discussion

In this chapter, it is shown that implicit modeling of a flat port underwater camera using Brown parameters yields a high error, especially for close-range measurement. Additionally, the refractive effects introduced by a tilted view port are not absorbed well by the Brown model. Therefore, a physical-geometric underwater camera model and calibration approach is proposed, improving the accuracy. The refractive model allows to transfer in-air calibration results to underwater imaging as well as refinement of the parameters in situ. This yields an underwater camera model with low errors for precise measurements. These findings are substantiated by experiments for the calibration of an underwater camera using a 3D structure composed of two AprilTag grids in a water tank.

Considering the practical aspects of underwater camera calibration, the experiments showed that 3D calibration fixtures provide faster convergence. Moreover, the additional constraints of the 3D structure allow calibrating with reliable results on fewer images. If the relative position

of the calibration points is not precisely known, BA allows to estimate the structure of the object points simultaneously. Considering marker detection, there are different available fiducial systems that enable fast and accurate detection. If clean images with high sharpness are available, all fiducial systems provide good results. On the other hand, in low-contrast and noisy images, circle targets are detected with higher accuracy. Since image quality is often degraded in natural waters, circle targets are a good choice. However, circle grids introduce a bias due to nonlinear distortion [207], which needs to be considered. Detection of the underwater markers is further improved by applying pre-processing of the image, e.g., dehazing [353].

Chapter 4

Underwater and Multi-media Laser Scanning

In this chapter, the development of a self-built structured light underwater laser scanning system and associated data processing is described. On a hardware level, it consists of an underwater camera and a cross-line laser projector, and it is based on laser triangulation as the measurement principle. A cross pattern has the benefit of an internal overlap between the created point clouds of the individual line projections. This is especially advantageous for the application of the scanner for mobile mapping, as discussed in Chapter 6. The chapter at hand describes the image processing, calibration and 3D reconstruction methods used for creating accurate dense point clouds using the scanner. The system is applied for the 3D reconstruction in water as well as scanning of semi-submerged objects, for which a refractive correction approach is proposed.

Fig. 4.1 depicts the proposed scanner mounted on a tripod and deployed for testing in a natural lake. Here, the scanner was employed for 3D acquisition of the concrete base of a wooden pier. Green high-power lasers are employed to achieve sufficient contrast of projected pattern. Due to the turbidity, backscatter of the laser light is present. This causes the light sheet to be visible in the image.

The underwater scanner hardware (cf. Section 4.1) and the experimental results presented in Section 4.5 were previously published in [10, 11]. The contributions of the author are the development of the scanner hardware and data processing software as well as the calibration approach and conduction of the experiments. The work at hand improves on the published methods with a calibration model that takes the distortion of the laser curve into account and an enhanced calibration of the rotation axis of the scanner (cf. 4.4). A thorough evaluation of the achieved measurement quality is presented in Section 4.6. The refractive correction method and experiments for scanning semi-submerged objects in Section 4.7 is the result of collaborative work and was previously published by van der Lucht et al. in [21–23]. The author's contributions are the 3D reconstruction methods, co-calibration and the details for the practical application of the refraction correction.

Fig. 4.1: Developed structured light underwater laser scanning system deployed for testing in a natural lake. Left: 3D underwater scanner with cross-line laser projector mounted on a tripod. Right: Underwater scanner with active laser projector.

4.1 Underwater Scanner Hardware

The developed structured light underwater laser scanner consists of two housings with flat port glass windows, one containing the camera and the other one the cross-line laser projector. The system is depicted in the left image in Fig. 4.2. The two housings are mounted on a 0.5 m long aluminum bar. Custom mounts for the housings were manufactured using 3D printing. The camera housing is mounted at an angle of 30° to the bar. On top, a larger housing is mounted, which contains the motor control electronics and a network switch that connects the underwater scanner to the surface via an underwater cable. The main components of the scanner, such as the underwater housings, cables and connectors, are rated for more than 1000 m water depth. However, the motor and 3D printed parts used in the prototype limit the whole system's depth rating to below 80 m.

The scanner is placed on a motorized joint with slewing ring bearings and a 1:50 worm gear, which is driven by a stepper motor. This allows rotating the scanner around the yaw axis to capture 360° scans. A magnetic encoder sealed in epoxide resin is used to measure the rotation angle of the scanner. All housings include embedded PCs with network interfaces that read the sensors and control the motor and laser drivers. For image processing, an embedded PC with an Intel Atom x5-Z8350 processor is included in the camera housing. The Robot Operating System (ROS) is used as a middleware for sensor interfaces, logging and data processing. All embedded PCs are time-synchronized using Network Time Protocol (NTP). For synchronization of the camera with the laser projector, a dedicated trigger pulse signal is used. A microcontroller realizes the projector control and timing of the trigger scheme.

The right image in Fig. 4.2 shows the electronics and optics components mounted inside the underwater housings. The camera assembly includes the lens with a focal length of 12.5 mm and a FLIR Blackfly 2.3 Megapixel color camera with a 1/1.2" Sony Pregius IMX249 CMOS sensor. The image resolution is 1920×1200 pixels with 5.86 µm pixel size and a maximum framerate of 41 fps.

The cross-line projector is constructed from Powell laser line optics and the laser diodes with

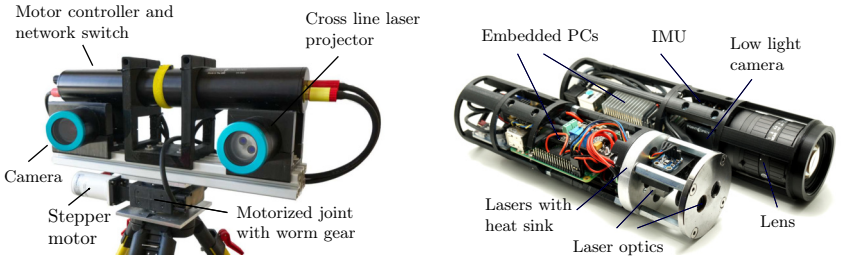


Fig. 4.2: Structured light underwater laser scanner. Left: Scanner with motorized yaw axis mounted on a tripod. Right: Detail view of the camera and laser projector assemblies mounted inside the underwater housings.

collimators. Powell lenses achieve a more even brightness distribution compared to cylindrical line lenses. The lasers are two 1 W green diode lasers with a wavelength of 525 nm, which are mounted to an aluminum heat sink. The laser output power is controlled by two laser diode drivers, which are adjusted via PWM signals generated by a microcontroller connected to an embedded PC. The two laser lines project a laser cross consisting of two approximately perpendicular lines in the scene. The fan angle of the laser lines is 45°, which is reduced in water to approximately 32°. The total field of view of the scanner is limited by the vertical field of view of the camera and is approximately 360°× 30°. The lasers are fired synchronized to the camera shutter using trigger pulse signals. An alternating firing order of the individual lasers is employed, such that each image captured by the camera includes only one of the two laser lines.

The two laser planes are projected at an angle of 45° with respect to the vertical camera axis. This way, both projected laser sheets have approximately the same point-plane distance to the camera center. However, compared to a laser sheet aligned with the vertical camera axis, the resulting triangulation baseline is reduced. For the proposed system, the baseline is approximately 30 cm.

4.2 Triangulation for Depth Estimation

The proposed laser line scanner is based on the measurement principle of triangulation. This means that the depth is measured by determining the angles in the triangle formed by the camera center $\mathbf{C}$, the world point $\mathbf{P}$ and the laser projection. The geometry is depicted in Fig. 4.3. The camera and the laser plane are separated by a baseline of length b. Here, it is assumed that the laser plane is perpendicular to the baseline. This means that the depth Z is determined by the angle of the camera ray θ:

$$\tan\theta = \frac{Z}{b} \quad . \tag{4.1}$$

An image sensor does not measure this angle directly, but the distance x' of the projection of the world point on the image plane from the optical axis. If the image plane is parallel to the baseline,

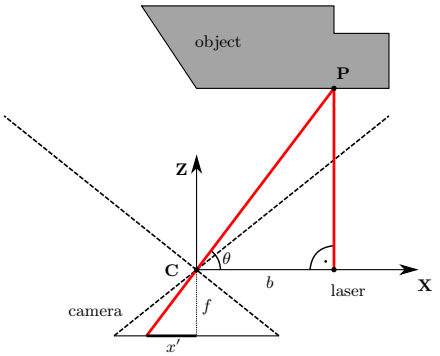


Fig. 4.3: Geometry of triangulation. The observed disparity x' depends on the distance Z of the object point $\mathbf{P} = (X, Y, Z)^T$ from the sensor.

then $\theta = \frac{f}{x}$, which leads to the basic equation of triangulation depth sensors [299, p.786]:

$$Z = \frac{fb}{x'} \quad , \tag{4.2}$$

where f is the focal length of the camera. Differentiation of Eq. 4.2 and application of error propagation yields the accuracy estimation in Z-direction:

$$s_Z = \frac{Z^2}{bf} s_{x'} \quad , \tag{4.3}$$

where $s_{x'}$ is the accuracy of the disparity measurement. Here, the focal length c and baseline b are assumed free of error. The equation shows that the triangulation accuracy is a function of the measurement distance Z, baseline b, focal length f and disparity measurement accuracy $s_{x'}$. The depth measurement accuracy falls off with the square of the distance to the object. Increasing the baseline or the focal length increases the accuracy linearly.

Similarly, the accuracy in X/Y-direction is estimated. Parallel to the image plane the X/Y-coordinates are determined as

$$X = \frac{Z}{f} x' \quad \text{and} \quad Y = \frac{Z}{f} y' \quad . \tag{4.4}$$

Analogous results the accuracy of the X and Y coordinates [202, p.344]:

$$s_X = \sqrt{(\frac{x'}{f} s_Z)^2 + (\frac{Z}{f} s_{x'})^2} \quad , \tag{4.5}$$

$$s_Y = \sqrt{(\frac{y'}{f} s_Z)^2 + (\frac{Z}{f} s_{y'})^2} \quad . \tag{4.6}$$

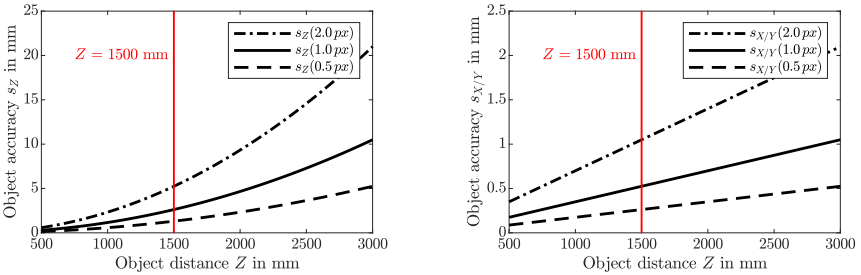


Fig. 4.4: Theoretical object accuracy s_Z and $s_{X/Y}$ depending on the measurement distance Z for a baseline length $b = 300\,\text{mm}$ and a focal length $f = 16.75\,\text{mm}$. Note that the scale is different for the y-axis of the two plots. Left: Object accuracy in Z-direction. Right: Object accuracy in X/Y-direction.

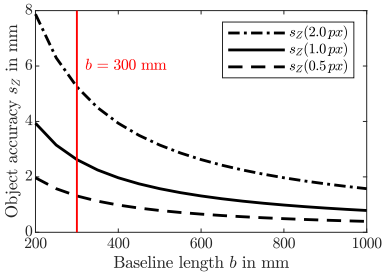


Fig. 4.5: Theoretical depth measurement accuracy s_Z depending on the baseline length b for a focal length $f = 16.75\,\text{mm}$ and an object distance of $Z = 1.5\,\text{m}$.

The dominant term is the second summand of the square root since $Z >> x'$. Therefore, the accuracy is estimated simplified as

$$s_{X/Y} \approx \frac{Z}{f} s_{x'/y'} \tag{4.7}$$

Hence, the accuracy in X/Y-direction linearly degrades with the distance.

Fig. 4.4 depicts plots of Eq. 4.3 and Eq. 4.7 depending on the measurement distance for the proposed underwater laser scanner. Note that the scale is different for the y-axis of the two plots. Here, the approximately equivalent focal length in water $f = 16.75\,\text{mm}$, a baseline length $b = 300\,\text{mm}$ and a pixel size of $5.86\,\mu\text{m}$ is used. The left image shows the object accuracy in Z-direction, and the right image shows the object accuracy in X/Y-direction for the assumption that the disparity is measured with an accuracy of 0.5, 1 and 2 pixels. At a nominal measurement distance of $Z = 1.5\,\text{m}$ and pixel accurate measurements the theoretical estimation yields $s_Z = 2.62\,\text{mm}$ and $s_{X/Y} = 0.52\,\text{mm}$.

Fig. 4.5 shows the effect of the baseline on the object accuracy in Z-direction for a nominal

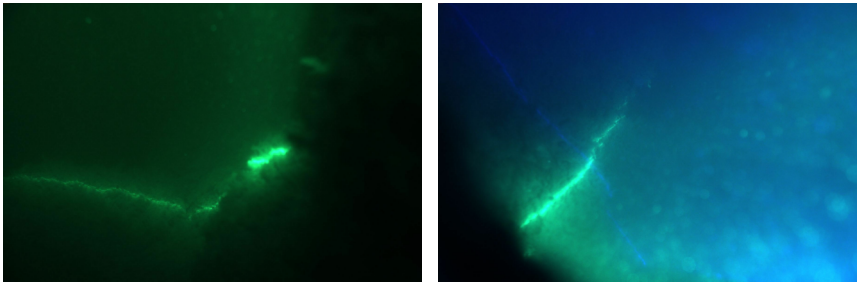


Fig. 4.6: Example images captured with green and blue laser line projectors in turbid water conditions with less than 1 m visibility.

measurement distance of 1.5 m. If one pixel accurate disparity measurements are assumed, an accuracy of $s_Z = 1\,\text{mm}$ at a distance of 1.5 m is achieved according to the theoretical estimates with a baseline length b of 787 mm. However, the actual achieved accuracy depends on additional factors, such as the accuracy of the calibration and subpixel localization of the image features.

4.3 Laser Line Extraction

A significant challenge during underwater 3D laser scanning is the image processing and extraction of the laser curves. Fig. 4.6 shows images captured by an underwater camera in turbid water conditions. Particle backscatter complicates the automatic extraction of the laser lines.

A simple approach to extracting laser lines from an image is to use maximum detection along horizontal or vertical scanlines in the image. However, this requires a high contrast between the bright pixels of the laser line and the background. Furthermore, this does not allow to extract multiple laser lines from a single image, which is required for the work on self-calibrating structured light described in Chapter 5.

Therefore, a ridge detector is employed for the extraction of the laser lines in the image. For this work, Steger's line algorithm [306] is applied since it is very robust and traces the center of the lines with subpixel accuracy. The idea behind this algorithm is to find curves in the image that have in the direction perpendicular to the line a characteristic 1D line profile, i.e., a vanishing gradient and high curvature. The line detector is applied to a grayscale image. If the input is a color image, it is converted accordingly.

The direction of the line in the two-dimensional image is estimated locally by computing the eigenvalues and eigenvectors of the Hessian matrix

$$H(x,y) = \begin{bmatrix} \frac{\partial^2 g_\sigma(x,y)}{\partial x^2} & \frac{\partial^2 g_\sigma(x,y)}{\partial x \partial y} \\ \frac{\partial^2 g_\sigma(x,y)}{\partial y \partial x} & \frac{\partial^2 g_\sigma(x,y)}{\partial y^2} \end{bmatrix} * I(x,y) = \begin{bmatrix} r_{\text{xx}} & r_{\text{xy}} \\ r_{\text{yx}} & r_{\text{yy}} \end{bmatrix} \quad , \tag{4.8}$$

where $g_\sigma(x,y)$ is the 2D gaussian kernel with standard deviation σ, $I(x,y)$ is the image and r_{xx}, r_{xy}, r_{yx}, r_{yy} are the partial derivatives. The direction perpendicular to the line is the eigenvector

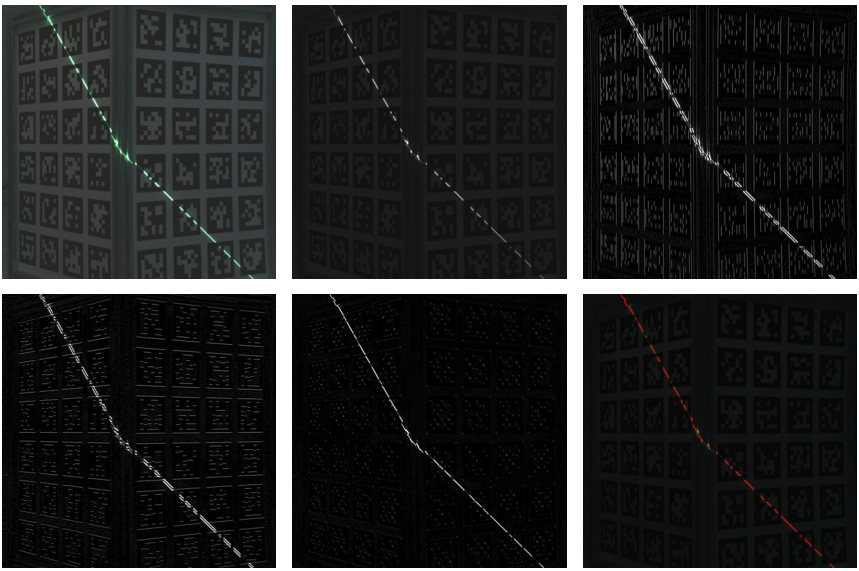


Fig. 4.7: Detection of the laser line. From top left to bottom right: Input image, grayscale image, r_{xx}, r_{yy}, r_{xy}, and detected points in red overlayed on the image.

$(n_x, n_y)^T$ with $\|(n_x, n_y)^T\|_2 = 1$ corresponding to the eigenvalue with the largest absolute value. For bright lines, the eigenvalue needs to be smaller than zero.

Fig. 4.7 shows a visualization of the filter responses for line detection. The figure shows the input image, the grayscale image, the partial derivatives r_{xx}, r_{yy}, r_{xy}, and the detected points in red overlayed on the image.

Instead of searching directly for the zero-crossing, a second-order Taylor expansion is employed to determine the location $(q_x, q_y)^T$ where the first derivative in the direction perpendicular to the line vanishes with subpixel accuracy:

$$(q_x, q_y)^T = t\,(n_x, n_y)^T \quad , \tag{4.9}$$

where

$$t = -\frac{r_x n_x + r_y n_y}{r_{xx} n_x^2 + 2 r_{xy} n_x n_y + r_{yy} n_y^2} \quad . \tag{4.10}$$

Here, $r_x = \frac{\partial g_\sigma(x,y)}{\partial x}$ and $r_y = \frac{\partial g_\sigma(x,y)}{\partial y}$ are the first partial derivatives.

For valid line points, the position must lie within the current pixel. Therefore, $(q_x, q_y) \in [-0.5, 0.5] \times [-0.5, 0.5]$ is required. Individual points are then linked together to line segments. This is done by choosing starting points with high responses and tracing along the detected

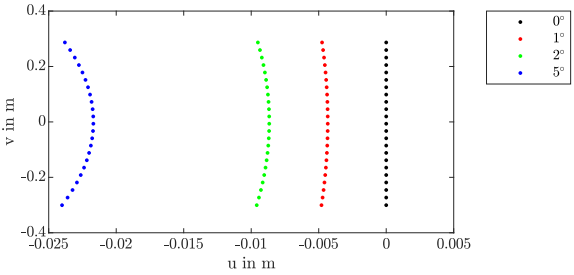


Fig. 4.8: Distortion effect of a tilted glass window on the laser line projection. The graph shows the projection on a plane at a distance of 1 m.

ridge points to form line segments until all detected ridge points have been processed. Double responses are explicitly detected and removed from the final output.

The response of the ridge detector given by the value of the maximum absolute eigenvalue is a good indicator of the saliency of the extracted line points. Only line points with a sufficiently high response are considered.

Since the line detector requires applying computationally expensive convolution filters, the image is first segment based on color cues and intensity thresholds, which is implemented using a look-up table. This way the line extraction algorithm needs to be computed only for parts of the image, which significantly reduces processing time.

4.4 Calibration of the Underwater Scanner System

The underwater camera is calibrated, which is described in Chapter 3. Additionally, it is necessary to calibrate the laser projector and the rotation axis to achieve 3D reconstruction. Due to misalignment of the optical components of the laser line projector as well as refraction at the glass window of the underwater housing, the projected laser line is not described accurately by a plane equation. To absorb these errors, the distortion of the laser curve is modeled as follows.

4.4.1 Laser Curve Parameters

The refraction of the laser line projection at the flat port of the underwater housing has two effects: First, assuming that the projector is perfectly orthogonal to the glass window, the fan angle of the laser line is reduced. Or, put another way, the focal length of the line projector is increased. This causes an inhomogeneous intensity distribution of the laser line. Towards the end of the line, the laser rays have a larger angle of incidence. Therefore, the refraction effect is stronger. This leads to a brighter laser projection at the ends of the line compared to the center.

Second, if the glass window is tilted with respect to the laser plane, the line is distorted. Fig. 4.8 shows a simulation of the effect. The simulation considers a single refraction interface

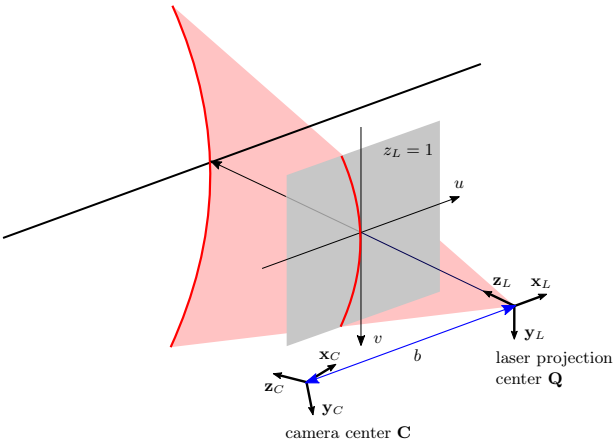


Fig. 4.9: Model of the laser curve by a tangential distortion defined in the plane $z_L = 1$.

directly in front of the laser projector based on Snell's law using Eq. 2.11. The refractive index of water is set to 1.33, and a laser fan angle of 45° is applied. The view port is tilted around the vertical axis with an angle of 1°, 2° and 5°. The image shows the projection of the simulated laser rays on a plane at a distance of 1 m. Note that the horizontal and vertical axes are scaled differently in this figure to emphasize the effect. The displacement is in the range of less than 5 mm for a window tilt of 1°.

In [281] the refracted rays are directly measured in water using a calibration fixture that splits the laser light sheet into multiple sub-beams. This way, a model based on interpolating the sub-beams is built. Other work models the deformation of the laser plane by the approximation of an elliptic cone [256]. In [109] the laser line projector is modeled as a line camera with the following parameters: focal length, principal point, tangential distortion, orientation of the laser plane and baseline.

An exact model based on ray tracing depends on a calibration of the origin and direction vector of the laser projection as well as the view port of the underwater housing. This requires purposely built calibration fixtures for accurate calibration. The self-built laser projector is fairly well aligned with a tilt in the range of 1°. However, it makes sense to compensate for this since even a small misalignment causes errors in the range of multiple millimeters at larger distances.

In this work, a simplified calibration based on implicit modeling of the laser line distortion is applied. This allows a workflow with the same calibration structure used for the underwater camera calibration. Similar to a line camera model, a parabolic correction function for the tangential distortion is applied. As visualized in Fig. 4.9, the proposed model defines the parameters using a fixed reference plane at a distance of $z_L = 1$ in front of the laser projector. Here, $\mathbf{y}_L$ and $\mathbf{z}_L$ span the plane of the laser sheet and $\mathbf{x}_L$ is orthogonal to the laser sheet. The $\mathbf{z}_L$ axis points in the direction of projection. Then, the tangential distortion of the laser sheet is defined based

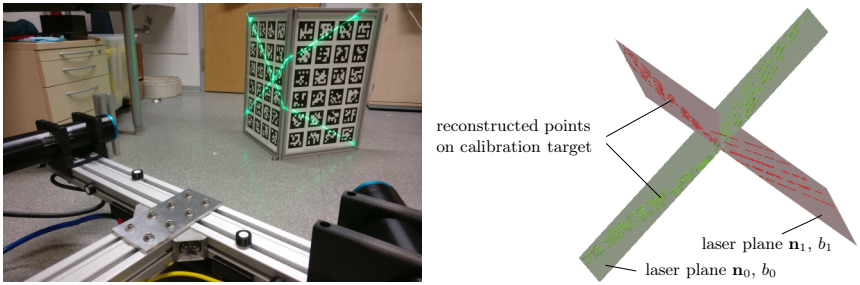


Fig. 4.10: Calibration of the laser planes in air using a calibration fixture. Left: Cross-line projection on the structure constructed from two planar targets. Right: Example of estimated planes based on the reconstructed points on the calibration target.

on the laser point $(u, v)^T$ in the plane $z_L = 1$. The distorted normalized laser point coordinate u_d is computed as

$$u_d = u + t_i \cdot r^2 \quad \text{with} \quad r^2 = u^2 + v^2, \tag{4.11}$$

where t_i is the distortion coefficient. This model does not use a principal point for the projection model since this is absorbed in the orientation of the laser coordinate system.

Hence, the complete parameter set of the laser curve i comprises the laser projection center $\mathbf{Q} = (q_x, q_y, q_z)^T$ in the camera coordinate system, the orientation (laser plane normal) $\mathbf{n}_i = (n_{1,x}, n_{1,y}, n_{1,z})^T$, the baseline of the projector b_i and the tangential distortion parameter t_i. For simplification, a single common laser projector center is applied although the two line projectors of the underwater scanner are approximately 18 mm apart.

The laser projection center and direction is estimated based on the intersection of the two laser planes and an external manual measurement of the approximate center position. Hence, the laser plane normal, baseline and distortion coefficient are calibrated in water.

The left image of Fig. 4.10 shows the setup for calibration of the laser plane parameters. The cross is projected on a calibration structure with known geometry. The underwater camera is pre-calibrated for this process. A calibration sequence of three images is captured: a bright image with the lasers switched off and two images with exposure suitable for the laser projection of each individual laser line. During the image sequence, the scanner and calibration fixture are static. First, the bright image is used for the localization of the target in the camera coordinate system based on the visual markers on the calibration structure. Second, the two images of the laser lines are used to reconstruct the laser points on the structure by intersection with the known geometry of the calibration structure. This is facilitated by using automatically created image masks based on the localization of the structure in the bright image. This way, only laser points projected on the two planar sides of the calibration structure are considered. The laser points of each line are then reconstructed by intersecting the camera rays with the respective plane of the calibration fixture.

In principal, a single calibration sequence at one position of the target is sufficient to estimate the plane equations. However, the end result is more robust and accurate if multiple positions

and orientations of the calibration fixture are captured. From the multiple observations, two point clouds of the two laser sheets are created. To extract the plane equations, a least square fit is applied to the point clouds, which is visualized in the right image of Fig. 4.10. In practice, it is important to capture the calibration fixture at different distances across the whole measurement range to achieve the most accurate plane parameter estimates. This yields the plane normal as well as the baseline.

The distortion coefficients are not estimated robustly from this data since the calibration fixture is too small. This way, typically only part of the laser lines is visible on the target and the observed distortion is small. To estimate the distortion coefficients, a planar target, e.g., a wall, is scanned at a distance of approximately 4 m. The distortion coefficient is then found by optimization, such that the scan of the planar target shows a minimal flatness error.

This calibration technique is suitable in air as well as in water. Calibrating the laser curve distortion in air is also advantageous to absorb errors, for example, due to misalignment of the laser optics. Finally, the 3D reconstruction is computed by intersecting the camera rays with the distorted laser curve.

4.4.2 Hand-Eye Calibration

In order to capture dense 3D scans, the underwater scanner needs to be moved. Therefore, it is necessary to measure or estimate the movement of the sensor to build the point cloud along the trajectory. One possibility to achieve this is by applying an external pose measurement, for example, by employing an optical tracking system for pose measurement.

Since the tracking system measures in a different coordinate system than the camera, it is necessary to align the coordinate systems. This is called hand-eye calibration in the literature [290]. To achieve this, two sets of 6-DoF poses with two unknowns need to be registered. The problem is visualized in Fig. 4.11. The tracking system and object are at fixed poses with respect to a global coordinate system and do not move. The body comprised of the camera and tracking target are moving. However, the transform between the target and the camera does not change.

Using the tracking system, the pose of the target $\mathbf{A}_i$ relative to the coordinate system of the tracking system is measured. In this work, the transform $\mathbf{A}_i$ is, for example, measured by an optical tracking system, satellite navigation or actuated axes with encoders. At the same time, the pose of the camera $\mathbf{B}_i$ is estimated with respect to the coordinate system of the object. This is typically done by estimating the position and orientation of the calibrated camera relative to known object points. Then, hand-eye calibration aims to find the two unknowns: (1) The rigid transformation $\mathbf{Y}$ between the tracking target and the camera coordinate system. (2) The relative position of the tracking system and the object coordinate system $\mathbf{X}$.

To find $\mathbf{X}$ and $\mathbf{Y}$, the camera body is moved to poses with different orientations and positions. For each camera body pose i the transforms $\mathbf{A}_i$ and $\mathbf{B}_i$ are recorded. The transformation from the target frame to the object coordinate system is determined by $\mathbf{A}_i\mathbf{X}$ or $\mathbf{Y}\mathbf{B}_i$. Therefore, every different pose of the camera yields the constraint

$$\mathbf{A}_i\mathbf{X} = \mathbf{Y}\mathbf{B}_i \quad \text{with} \quad i = 1 \dots n \quad , \tag{4.12}$$

where n is the number of different camera poses.

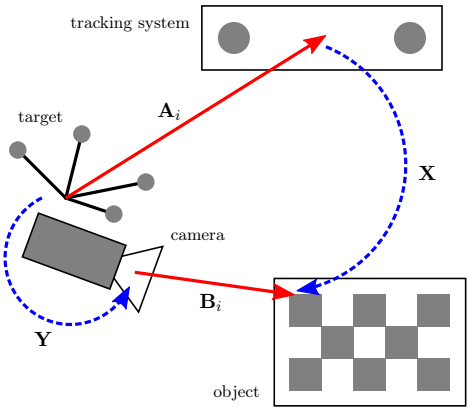


Fig. 4.11: Hand-eye calibration with two unknowns $\mathbf{X}$ and $\mathbf{Y}$.

To find the unknowns $\mathbf{X}$ and $\mathbf{Y}$ of hand-eye calibration, the problem of the form $\mathbf{AX} = \mathbf{YB}$ needs to be solved. Different approaches to do this are proposed in the literature [290]. Some approaches solve this with a separable solution by solving the rotation and translation components individually [103,352]. In this work, the method of [196] is applied, which solves the orientation and position simultaneously.

4.4.3 Calibration of the Rotation Axis

For 3D scanning while using the motorized yaw axis, a precise estimate of the axis of rotation is necessary. Calibration of the rotation axis parameters poses a problem similar to hand-eye calibration. In this work, the transform is described by the radius and the rotation of the camera, which is depicted in Fig. 4.12. The four parameters that need to be calibrated are the radius r and the rotations around the camera axes $\mathbf{x}_c$, $\mathbf{y}_c$ and $\mathbf{z}_c$.

This is achieved using the proposed hand-eye calibration approach. By placing a calibration fixture in front of the scanner and rotating the scanner, the rotation angle and the pose of the camera relative to the calibration object is recorded. From this data, it is possible to estimate the parameters. However, a single calibration structure is only observed for rotation angles in a range of 10° to 20°. This results in an unstable estimate of the radius using a least-squares fit since only a small part of the circular trajectory is observed.

One possibility to achieve reliable estimates is to build a larger calibration structure, which is visible in a larger range of rotation angles. In this work, the rotation parameters are estimated from external trajectory measurements using an optical tracking system. Additionally, a refinement approach is proposed based on optimizing the point cloud consistency using the captured 3D data of the individual laser lines.

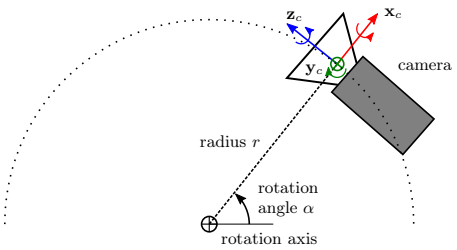


Fig. 4.12: Parameterization of the rotation axis: Radius r and the rotation angles around the camera axes $\mathbf{x}_c$, $\mathbf{y}_c$ and $\mathbf{z}_c$.

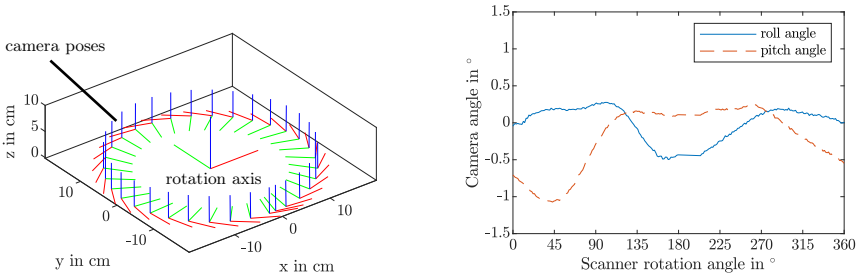


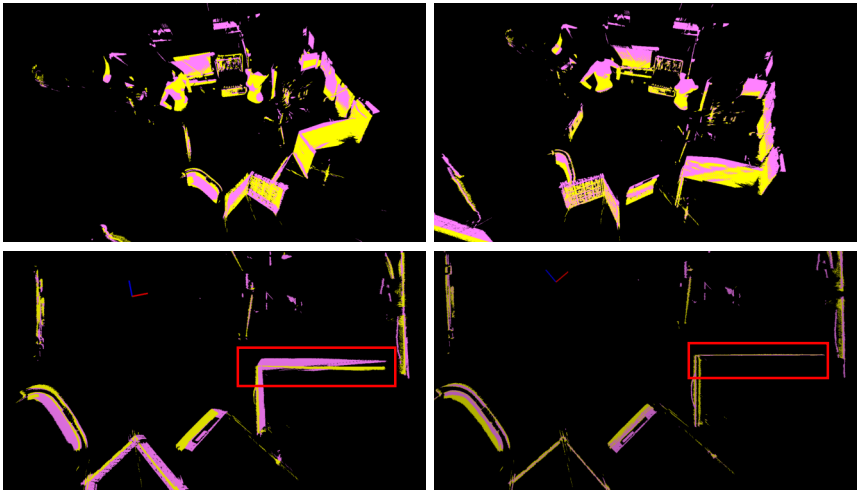
Fig. 4.13: Camera path measured by the optical tracking system. Left: Measured camera poses transformed in a coordinate system spanned by the rotation axis and plane of the trajectory. Right: Plot of camera roll and pitch angle depending on the scanner rotation angle shows a precession of the rotation axis.

External Measurement of the Rotation

The trajectory of the camera is measured using an OptiTrack V120:Trio tracking system [237]. Calibration of the transformation between the tracking body and the camera is achieved by means of hand-eye calibration (cf. Section 4.4.2).

The radius and plane of rotation are estimated by fitting a circle to the recorded camera positions. The left image of Fig. 4.13 shows a visualization of a subset of the camera poses transformed into a coordinate system spanned by the rotation axis and plane of the trajectory.

The right plot in Fig. 4.13 shows the camera roll and pitch angle depending on the scanner rotation angle. Since the rotation axis is only supported on one side, a precession of the rotation axis is observed. The precession is not a reproducible movement. The change in orientation of the axis depends on the orientation of the scanner due to play in the bearings of the rotary joint. Therefore, the average camera roll and pitch angles are applied. In this example the measurement yields a radius of 178.21 mm, a yaw angle of $-14.829°$, a pitch angle of $-0.614°$, and a roll angle of $-0.102°$.



Initial result. Optimized result.

Fig. 4.14: Initial and optimized point cloud result based on rotation axis calibration. Top: Point clouds created by the individual laser lines in yellow and pink. Bottom: Top view of the point cloud with prominent alignment errors highlighted by the red rectangle.

Calibration Based on Point Cloud Optimization

Another approach for the calibration of the rotation axis is based on optimization of the point cloud consistency. The idea is similar to the calibration approach for multiple LIDAR scanners proposed in [294] and the calibration approach described in [115]. This method has the advantage that no external infrastructure or particular calibration structure is necessary. It is directly applied to the scan data of a scene, which is required to provide sufficient depth features.

The idea is to exploit the multiple measurements of surfaces by the two laser lines. A 360° scan is performed with the underwater scanner. Then, the generated point cloud is divided into the scans acquired by the two individual laser lines. The deviations of the point measurements of these two scans are very small in static environments if the rotation parameters are correctly estimated. To find the parameters of the rotation axis, an error measurement similar to the nearest neighbor distance used in Iterative Closest Point (ICP) [45] is applied. The rotation axis parameter, which minimizes the point-to-point distance between the two scans of the individual lines, is found using Powell's method. This is a local optimization method; therefore, a good initial estimate is required.

Fig. 4.14 shows an example of the two point clouds using the initial parameters on the left and after optimization on the right. The point clouds of the individual laser lines are colored yellow and pink. In the bottom images, the improvement in alignment is highlighted by a red

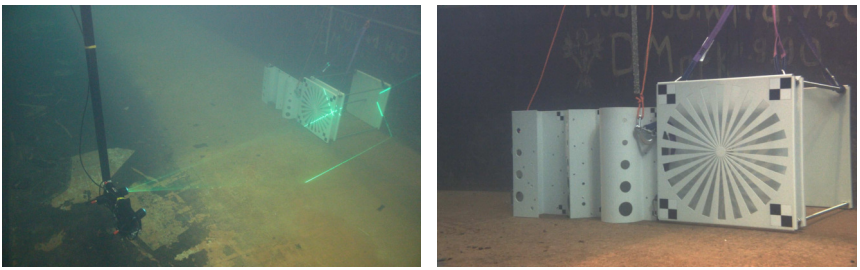


Fig. 4.15: Experiments conducted in the towing tank of the chair of fluid dynamics at the University of Rostock. Left: Structured light scanner deployed in the towing tank. Right: Image of the test objects captured with the camera of the scanner.

rectangle. While the point clouds deviate with high error in the initial result, they align well after the optimization.

However, using this approach, it is difficult to estimate the radius of the rotation trajectory precisely. Both point clouds of the individual laser lines are captured from a similar central position. This typically results in an underestimation of the radius by the optimization technique. This is potentially solved by using scans from different positions of the same scene. However, this requires simultaneous estimation of the rotation parameters as well as the relative transform between the individual scan positions. Moreover, matching explicit features in the scene, e.g., individual planes, is a direction to improve the results. This also yields a better interpretable error measurement, such as the alignment error of selected planes in the scene.

4.5 Underwater Scanning Results

This section shows examples of underwater scans created with the proposed scanning system. The results in this section are earlier work and were achieved without the proposed refractive camera calibration and the laser distortion model. The camera was calibrated with Brown's distortion model, and the laser sheet is assumed planar. Moreover, less precise calibration fixtures were employed. Therefore, the achieved accuracy is lower compared to the evaluation of the measurement quality that is described in the following section. However, the examples demonstrate the 3D reconstruction of larger objects and scenes in water.

4.5.1 Experiments in a Towing Tank

For testing, the underwater laser scanner was deployed in the towing tank at the chair of fluid dynamics at the University of Rostock. The deployed scanner is depicted in the left image of Fig. 4.15. The water tank is 5 m wide and provides a depth of up to 3 m. The scanner was deployed at about 2.5 m water depth using a vertical bar.

Static scans were acquired by rotating the system using the yaw motor of the scanner. For evaluation purposes, complex objects, which were built by the Fraunhofer Research Institution

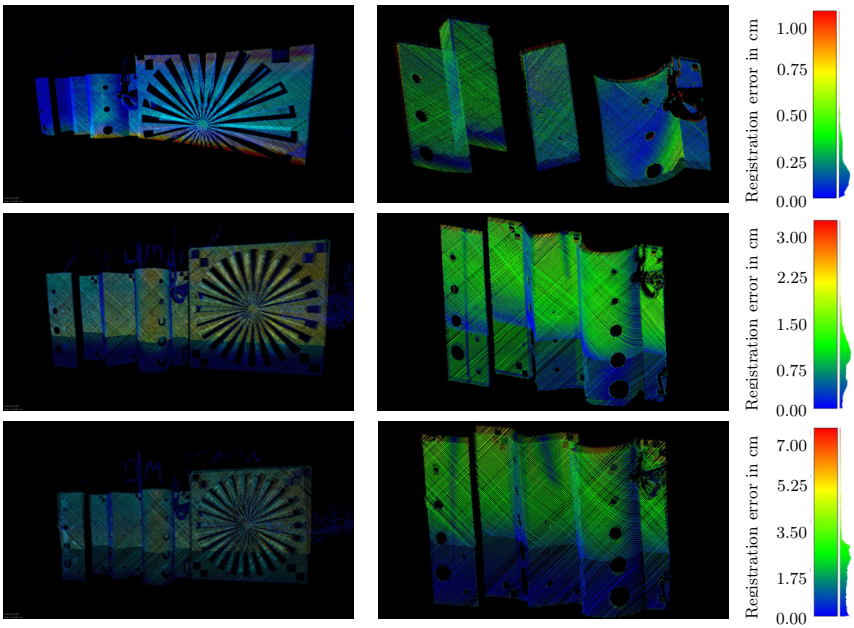


Fig. 4.16: Scans of the test objects captured at different distances of 1 m, 2 m and 3.5 m (from top to bottom). Left: point cloud colored using intensities, middle: registration of the scan with the reference model (white), right: distance between reference model and 3D scan.

for Large Structures in Production Engineering (IGP) in Rostock and the company IMAWIS GmbH, are placed in front of the scanner at different distances. The test objects shown in the right image of Fig. 4.15 were scanned in air with a high-precision structured light scanner GOM ATOS III to create a reference model of the geometry [242].

Fig. 4.16 shows in the left column point cloud results captured at varying distances. The point clouds are colored by intensity. The top scan was captured at about 1 m distance, the middle scan at 2 m distance, and the bottom scan at 3.5 m distance from the objects.

The point clouds are registered using the ICP algorithm with the reference models. The distance between the reference model and the captured point cloud is computed based on the nearest neighbors. The scan colored by the difference to the reference model is visualized in the right column of Fig. 4.16. Note that the color scaling is adjusted for each picture to the range of errors present in the particular scan to highlight the distribution of errors within the scan. While for the close range small errors in the range of few millimeters are observed, the error significantly increases with distance. For the top scan captured at 1 m distance, the errors are

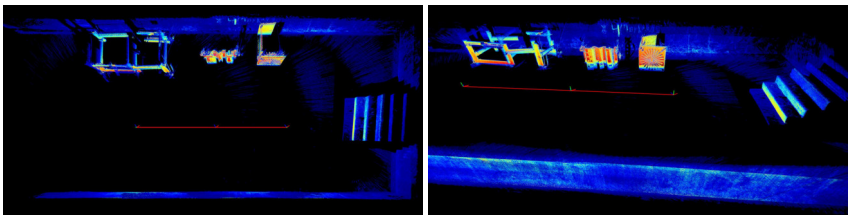


Fig. 4.17: Point cloud of the towing tank colored with laser intensities created by registering three underwater scans (the scan poses are marked along the red line).

below 1 cm. The middle scan captured at 2 m shows larger errors of up to 2 cm. The bottom scan captured at 3.5 m distance shows errors of up to 5 cm. At the larger scanning distances, some misalignment of the measurements of the two laser lines is visible, which is caused by calibration inaccuracies.

The towing tank was also scanned from multiple different positions. The top images in Fig. 4.17 show 360° scans of the towing tank. The point cloud was created by registering three scans using the ICP algorithm. The individual scanner poses are marked along the red line.

4.5.2 Experiments in a Water Tank

For the next experiments, an underwater test environment was set up, which is described in [24]. The setup is depicted in the left images of Fig. 4.18. It is based on a roll-off container with a size of 7 m × 2.4 m × 2.4 m. The total volume is 40 m^3. A test structure constructed from aluminum profiles and polypropylene pipes is placed in the water tank, which is shown in the right images of Fig. 4.18. The pipe object is placed at a distance of approximately 2 m in front of the scanner. Multiple scans from different view points are captured with the underwater scanner by rotating around the motorized yaw axis.

Fig. 4.19 depicts results of the static underwater scans from a tripod. In the top left image, the resulting point cloud colored by intensity is shown. The intensity image is computed from the brightness of the detected laser points in the camera image. The top right image depicts the top view and perspective view of a point cloud registered from four underwater scans. The point cloud is colored by height, and the poses of the individual poses are visualized by small coordinate systems drawn in the image.

In the bottom row, the comparison between the reference scan and the underwater point cloud is shown. The reference scan of the pipe structure was acquired in air with a terrestrial laser scanner. On the left, the reference point cloud is visualized in pink, and the underwater point cloud is shown in yellow. The right image shows the registration errors between the reference and underwater scans. For this object, which encompasses a depth range of approximately 2 m to 4 m, the achieved errors are in the centimeter range, as shown in the error bar on the right side of the image. In this particular case 90 % of the errors are below 2 cm.

Fig. 4.18: Scanning setup in the test tank. Left: $40\,\mathrm{m}^3$ roll-off container. Right: Underwater scanning of a test structure constructed from aluminum profiles and polypropylene pipes.

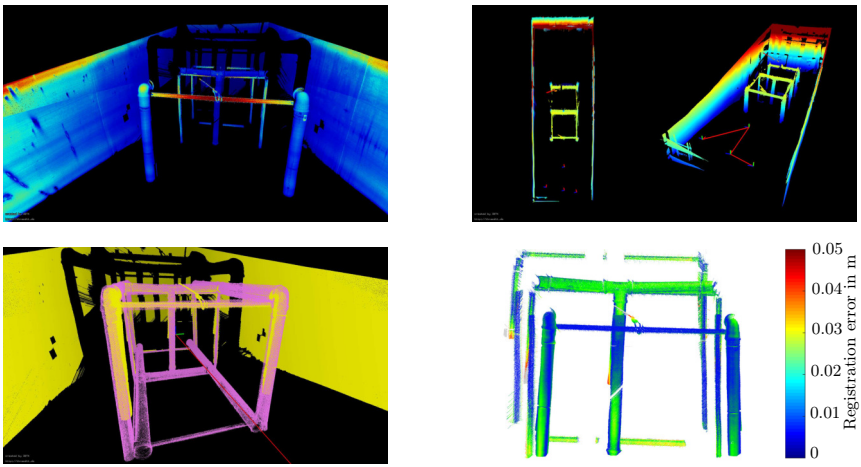


Fig. 4.19: Results of underwater scanning from a tripod. Top left: Point cloud of a scan of the pipe structure colored by intensity. Top right: Top view and perspective view of the registered point cloud from four scans. Bottom left: Underwater point cloud in yellow and reference scan in pink. Bottom right: Registration error between reference scan and underwater scan.

4.6 Evaluation of Measurement Quality

For the benchmark of the measurement quality of the underwater laser scanning system, geometric errors of the reconstructed point cloud are evaluated, such as shape and length measurement errors, as well as the achievable resolution of the scans. Specifically, the following error measurements are considered:

1. The *sphere probing error* characterizes surface measurement errors in a small volume [202, p. 680]. The probing error evaluates how well the point measurements fit the form of a sphere as visualized in Fig. 4.20. It is reported as the range e_f of the measured deviations from the best-fit sphere. Additionally, the size error e_s is reported as the difference between the diameter of the calibrated reference sphere and the best-fit sphere.

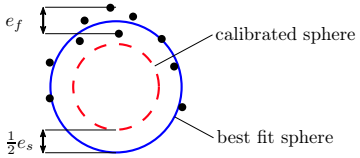


Fig. 4.20: Form error e_f and size error e_s of the probed sphere. The black dots visualize the point measurements of the scanning device.

2. The *sphere-spacing error* measures the deviation of a distance measurement between two spheres [202, p. 681]. This is done by fitting two spheres with fixed diameters to the point cloud as visualized in Fig. 4.21. The deviation of the distance between the fitted spheres and the calibrated distance of the reference spheres is reported as sphere-spacing error

$$e_l = l_m - l_c \quad . \tag{4.13}$$

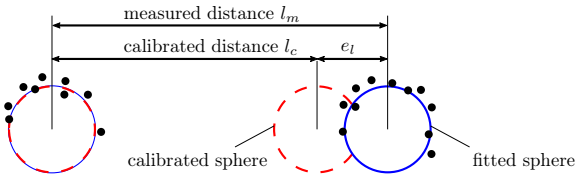


Fig. 4.21: Sphere spacing error e_l. The black dots visualize the point measurements of the scanning device.

3. The *length measurement error of scalebars* characterizes similarly distance error between specific points. The difference here is that planar circular markers on scalebars are measured, which are distributed in the measurement volume instead of spherical targets. Generally, this error is expected to be higher than the sphere-spacing error because the sphere center is retrieved more accurately by computing the fit over a larger number of points.

4. The *flatness or planarity error* evaluates the performance of the scanner to measure a planar surface [202, p. 680]. It also characterizes the measurement noise since it reports the spread of the point measurements from a best-fit plane. This is descriptive of the thickness of measured surfaces in the point cloud. The flatness error e_p is reported as the range of deviations of the points from a best-fit plane as visualized in Fig. 4.22

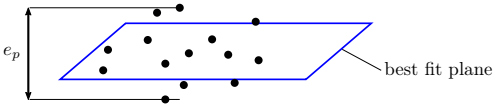


Fig. 4.22: Flatness error e_p. The black dots visualize the point measurements of the scanning device.

5. The *measurement error of free-form surfaces* captures how well objects of known geometry are reproduced by the scanning system. This is evaluated by scanning a geometrically stable object in air with a precise scanning system to create a reference scan. Then, the point clouds of the reference scan and underwater scans are compared with each other. This is done by aligning the two point clouds and computing the nearest neighbor distance, which is sometimes also referred to as cloud-to-cloud distance, for each point of the underwater scan.

6. The *resolution* indicates the minimum size increment of geometric features that are reproduced by the scanning system. The depth resolution is evaluated separately from the overall spatial resolution. The depth resolution mainly depends on baseline, focal length and the minimum resolvable feature distance in image space. The spatial resolution is typically lower since it is also affected by the movement increment of the scanner, thickness of the laser line and more strongly affected by blur and image degradation.

The described tests in this section for probing, sphere-spacing and flatness error are inspired by the recommendations of VDI/VDE 2634 Part 2 guidelines for testing area-scanning systems. However, the tests do not strictly follow the verification procedures of the norm. The employed test targets are smaller than $10 - 20\%$ of the diagonal of the measurement space, which is the minimum size recommended by VDI/VDE 2634. Moreover, the distance dependency of measurement errors is considered because for structured-light 3D scanning systems the measurement quality primarily decreases with the distance between the object and the scanner. Therefore, scans of the test objects were acquired mainly at different distances and are not evenly distributed over the complete measurement volume.

4.6.1 Reference Artifacts

The reference artifacts used for the evaluation are depicted in Fig. 4.23. The first image in the top row shows a target with four spheres, which is used for the evaluation of the sphere probing and sphere-spacing error. The second image shows scalebars mounted in a $1\,\mathrm{m}^3$ measurement volume for length measurement error determination. The third image shows gray-painted float glass sheets for characterizing the flatness error. In the bottom row, the first two images show

Fig. 4.23: Reference artifacts. From top left to bottom right: Sphere target, Scalebars in a $1\,m^3$ measurement volume, Grey painted float glass planes, resolution test targets, resolution test target mounted to calibration fixture, dwarf figure being scanned with the FARO ScanArm.

staircase patterns used for resolution evaluation. The bottom right image shows a dwarf figure being scanned with the FaroArm Quantum S and a FaroBlu Laser Line Probe to create a reference scan for the evaluation of free-form geometric objects.

Sphere Target

The sphere target is built from four plastic ball screw nuts, which are mounted on a 3D-printed base plate. The spheres are grey painted with 48% grey spray paint (RAL 7037). The Dimensions of the fixture are visualized in Fig. 4.24. All dimensions are checked using a three-camera tracking system OptiTrack V120:Trio. The accuracy of the tracking system is specified as submillimeter by the manufacturer [237]. The measured distances of the spheres differ from the

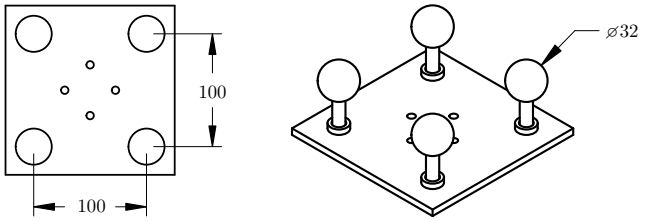


Fig. 4.24: Ball Target. Dimensions are in mm.

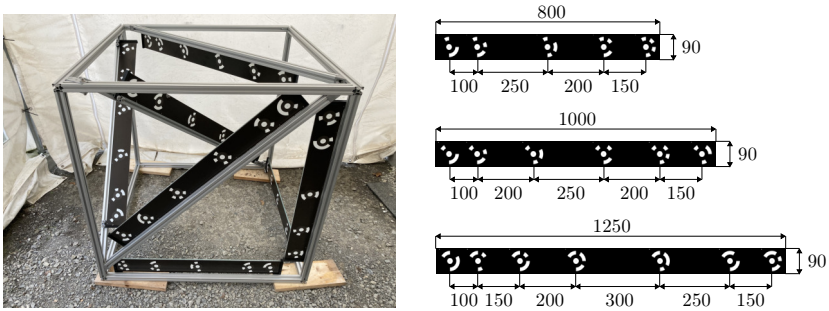


Fig. 4.25: Scalebars mounted in an aluminum frame with a size of 1 × 1 × 1 m. Dimensions of the scalebars are in mm.

design dimensions by less than 0.2 mm. The diameter measurements of the individual balls were additionally checked using a caliper. While commercially available sphere targets are calibrated more accurately using, e.g., coordinate-measuring machines, the test measurements indicate that the errors of the dimensions are in the sub-millimeter range. Therefore, given the higher magnitude of the evaluated errors, the accuracy of the sphere target is sufficient for the performed evaluations.

Scalebar Target

The scalebar target ist constructed from an aluminum frame with a size of 1 × 1 × 1 m. Each individual scale bar is manufactured using direct print on a float glass sheet. Due to the manufacturing process of curing molten glass on a bed of molten metal, very flat surfaces are achieved. The print accuracy, according to the manufacturer, is 0.1 mm. Three scalebars with a total length of 800 mm, three scalebars with a total length of 1000 mm and one scalebar with a total length of 1250 mm are mounted in the aluminum frame. Coded circle markers are printed in white color on a black background on the glass sheets. The central white circles of the markers have a diameter of 24 mm. The markers are distributed with varying distances over the length of the scalebars. A total of 40 points on 7 glass sheets are available. The dimensions of the scalebars are depicted in Fig. 4.25.

Plane Target

To evaluate the flatness error, often granite plates are used because they are manufactured with high overall surface flatness. Moreover, granite has low thermal expansion. For this work, a sheet of float glass is employed, which is expected to have a very flat surface due to the manufacturing process. The glass is coated with mate spray paint with 48% gray color (RAL 7037) to achieve a surface with good scanning conditions.

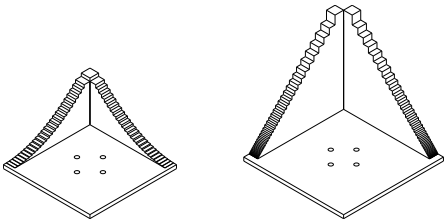


Fig. 4.26: Resolution test structures. Left: Test structure with stair pattern with decreasing height. Right: Stair pattern with equally decreasing height and tread depth.

Resolution Targets

The spatial resolution of terrestrial laser scanners is evaluated, for example, with the so-called Boehler Star. This is an adaption of the Siemens Star, an often employed image resolution test pattern, to three dimensions. However, for evaluating triangulation-based sensors, the Boehler Star has a disadvantage: For a structured light scanner, the light travels a different path from the illumination source to the object than from the object back to the receiver or camera. Therefore, also shadowing of the back plane of the Boehler Star occurs due to the different viewpoints. This is not the case for ToF laser scanners, where the beam of laser light travels along the same path from the scanner to the object and back.

Hence, in this work, resolution test targets similar to a staircase pattern are applied. The test targets are depicted in Fig. 4.26. Each pattern has two staircase patterns to characterize resolution in two dimensions, e.g., horizontally and vertically in the 3D image. The left test pattern in Fig. 4.38 has a constant run of 5 mm and a decreasing rise from 5 mm to 0.2 mm in increments of 0.2 mm. With rise, the height of a step is denoted, which is measured from the top of one tread to the top of the next tread. With run, the length of the tread is denoted, which is measured from the outer edge of the step to the vertical portion of the stair. This pattern is used to characterize only the depth resolution of the sensor since only the step's height decreases.

The second test pattern, which is depicted on the left in Fig. 4.38, features an equally decreasing rise and decreasing run. The step sizes are in the range of 0.2 mm to 15 mm. The largest step has a run and rise of 15 mm. From 10 mm to 5 mm the run and rise decrease in increments of 0.5 mm. Below 5 mm the increment size is 0.2 mm.

Free-form Object

For comparison with a free-form object, a reference point cloud of a dwarf figure was created. The dwarf figure has a size of approximately $16 \times 15 \times 23$ cm. The figure was scanned with the FARO Quantum S ScanArm. The object was placed on a rotating table and scanned from all sides with the line profiler of the ScanArm as depicted in the bottom right image in Fig. 4.23. The manufacturer specifies the overall accuracy of the system to be below 0.1 mm (depending on the configuration). The scan covers the complete figure and has only small holes in areas that are hard to scan, e.g., underneath the arm holding the soccer ball.

Table 4.1: Calibration result of *Brown Model* without explicit compensation of refraction effects.

Parameter	Value
Camera parameters	
Focal length (f_x, f_y)	$(2943.17, 2944.02)$ pixels
Principal point (c_x, c_y)	$(959.87, 611.99)$ pixels
Radial distortion (k_1, k_2, k_3)	$(0.27628, 0.40508, 3.60879)$
Tangential distortion (p_1, p_2)	$(0.00114, 0.00127)$
Laser plane parameters	
Laser 0 normal $(n_{0,x}, n_{0,y}, n_{0,z})$	$(-0.55111, -0.80701, 0.21218)$
Laser 0 baseline b_0	215.72 mm
Laser 1 normal $(n_{1,x}, n_{1,y}, n_{1,z})$	$(-0.75716, 0.59128, 0.27765)$
Laser 1 baseline b_1	301.57 mm

4.6.2 Calibration Result and Experimental Conditions

All tests for the evaluation are performed with the same calibration. The results of three different models are compared: First, the Brown model was calibrated on underwater images with the assumption of planarity of the laser planes. This model does not explicitly compensate for the refraction but absorbs the effects in radial and tangential distortion parameters. Results using this model are reported in the following experiments as *Brown Model*. Second, a refractive model with explicit modeling of the refraction at the air-glass and glass-water interface is employed. The laser planes are assumed planar. This model's results are reported in the following experiments as *Refractive Model*. Third, a model that additionally models the distortion of the laser line with a tangential distortion parameter is considered. Results using this model are reported in the following experiments as *Line Distortion*.

The parameters of the three models are estimated on the same input data and the same settings for the optimization algorithm of the model. The camera models are calibrated on a set of 248 underwater images of the 3D float glass structure with AprilTags described in Section 3.5. The laser line parameters are estimated on a set of 11 different poses of the 3D structure in water using the method described in Section 4.4.1. The distance between the scanner and the calibration pattern is in the range of 1 m to 4 m. All housing parameters are calibrated as described in Section 3.4.

The calibration results using the *Brown Model* without explicit compensation of refraction effects is reported in Tab. 4.1. The RMS projection error of the camera calibration is 0.46995 pixels. The estimated focal length in water is approximately 17.25 mm compared to 12.93 mm in air. This is the expected increase in focal length due to refraction. Additionally, the refraction effects are absorbed in the radial and tangential distortion parameters. Especially, the third radial distortion parameter k_3 is very high. This means high shifts in pixel position due to distortion near the image borders. The angle between the estimated laser plane normals is 90.06°, which means the laser planes are approximately perpendicular. The RMS plane fit error

Table 4.2: Calibration result of *Refractive Model* and laser line distortion parameters.

Parameter	Value
Camera parameters	
Focal length (f_x, f_y)	$(2211.85, 2212.55)$ pixels
Principal point (c_x, c_y)	$(957.51, 612.54)$ pixels
Radial distortion (k_1, k_2, k_3)	$(-0.05818, 0.14644, 0.18660)$
Tangential distortion (p_1, p_2)	$(0.00091, 0.00004)$
Housing parameters	
Air-glass interface distance d_0	33.14 mm
Glass-water interface distance d_1	52.14 mm
Interface normal (n_x, n_y, n_z)	$(0.00478, 0.00001, -0.99999)$
Laser plane parameters	
Laser 0 normal $(n_{0,x}, n_{0,y}, n_{0,z})$	$(-0.55133, -0.80690, 0.21201)$
Laser 0 baseline b_0	212.18 mm
Laser 1 normal $(n_{1,x}, n_{1,y}, n_{1,z})$	$(-0.75746, 0.59080, 0.27786)$
Laser 1 baseline b_1	296.84 mm
Laser line distortion	
Laser projection center (q_x, q_y, q_z)	(-325.03, -0.18, 164.69) mm
Laser fan angle α	33.8°
Laser 0 tangential distortion t_0	0.00250
Laser 1 tangential distortion t_1	0.00125

of the two laser planes is 0.487 mm and 0.282 mm, respectively.

The result of the calibration using the *Refractive Model* with housing parameters as well as the parameters for the laser line distortion, are reported in Tab. 4.2. The intrinsic camera parameters are close to the parameters of an in-air calibration. The focal length is approximately 12.95 mm, which is close to the focal length result of 12.93 mm of a camera calibration in air. The distortion parameters are close to the air calibration and mostly an order of magnitude smaller than the parameters of the *Brown Model* calibrated in water. This indicates that the housing parameters compensate well for the refraction effects. The refractive camera model has a RMS projection error of 0.46921 pixels, which is slightly smaller than the RMS projection error of the *Brown Model* without explicit compensation of the refraction effects. As already noted in Section 3.4.2, the RMS error is the training residual of the camera model, which on its own has limited significance for the expected error of the model. The estimated plane normals are similar to the plane normals of the *Brown Model* with an angle between the laser planes

of 90.01°. A bigger difference is visible for the estimated baselines of the laser planes. Scale is computed based on the extrinsic parameter estimation of the 3D float glass structure. Both models estimate slightly different extrinsic parameters, which results in different baselines. This means a difference in the scale of the resulting point clouds is expected depending on the employed model. The RMS plane fit error of the two laser planes is 0.545 mm and 0.271 mm, respectively.

The laser planes in air have a deviation of 1.4° and 0.8° to the normal of the air-glass interface. This results in a deformation of the laser planes, which is compensated by the *Line Distortion* parameters in Tab. 4.2. The parameters can be interpreted as a deviation of 2.5 mm and 1.25 mm from the laser plane for a laser curve projected in water of the length of 2 m. The reported laser projection center is relative to the camera center. The projection center is computed based on the intersection of the laser planes. The same projection center is used for both laser line projectors; also there is a physical offset of 18 mm between the two projectors.

The calibration of the rotation axis of the scanner as described in Section 4.4.3 yields a radius of 178.21 mm, a yaw angle of −14.829°, a pitch angle of −0.614°, and a roll angle of −0.102°. This means the horizontal and vertical axis of the camera is approximately aligned with the plane of rotation since the pitch and roll angle are both below 1°.

The evaluation was performed in clear mains water in the 40 m^3 water tank described in Section 4.5.2. The scanner was placed at a water depth of approximately 1.6 m. The focus distance of the camera was set at approximately 1.5 m in air, which translates to a distance with the best sharpness of the image at approximately 2 m in water. Below 1 m, there is a visible decrease of the sharpness of the image, resulting in slightly blurred images, which affects measurement quality.

4.6.3 Sphere Probing Error

The sphere probing error is evaluated using scans of the sphere target as depicted in the top left image of Fig. 4.27. The target from the perspective of the scanner camera is shown in the top right image. A total of 14 scans are captured at different measurement distances in the range of 0.75 m to 3.7 m. The results are compared for a calibration using the *Brown Model*, *Refractive Model* and additional compensation of the *Line Distortion*. Additionally, the scans of the two laser lines of the scanner are evaluated separately to investigate the influence of the different baselines of the individual line projectors.

For the computation of the sphere probing error, the individual spheres are cropped manually from the point cloud, which is depicted for a measurement distance of 0.75 m in the bottom left image of Fig. 4.27, using box filters. Then, a best-fit sphere is fitted to the individual point clouds of the spheres, which is visualized in the bottom right image of Fig. 4.27. Here, the best-fit spheres are rendered in green color. Following the recommendations of VDI/VDE 2634 a maximum of 0.3% of outliers are removed, and the sphere form error e_f and the sphere size error e_s are computed. No post-processing of the raw reconstruction data or filtering of the point clouds was applied except for removing the mentioned outliers.

The sphere form error depending on the measurement distance is plotted in the graphs in the top row of Fig. 4.28. These are the results for the point cloud reconstructed from the laser line with the larger baseline $b_1 \approx 0.3$ m. The left graph is a scatter plot of the form errors of the four individual spheres. The right graph shows the mean sphere form error for each scan. The

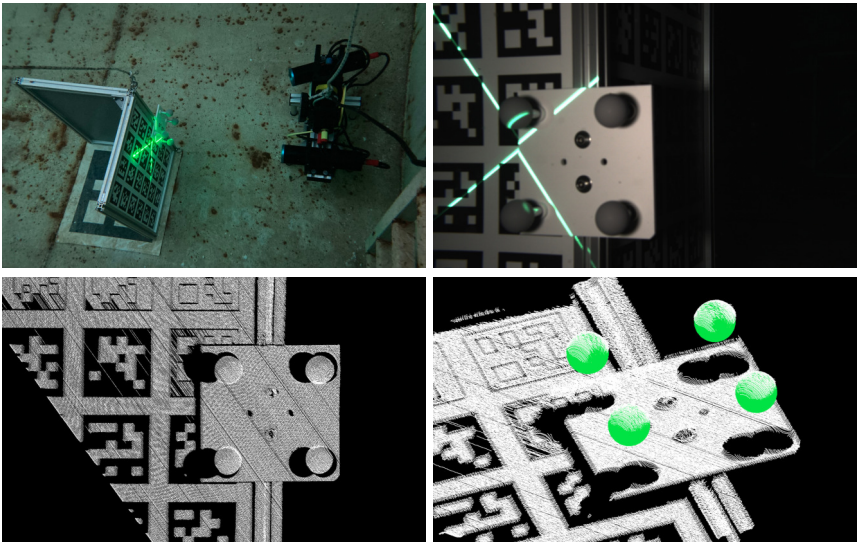


Fig. 4.27: Sphere target evaluation. Top left: Scanner deployed in the water tank during evaluation. Top right: Scanner image at close range with laser line projection. Bottom left: Point cloud acquired at a distance of ca. 0.75 m. Bottom right: Spheres fitted to the point cloud visualized in green color.

resulting sphere form errors are in the range of 1 mm to 3 mm in the measurement range up to 3 m and increases for larger distances due to higher noise of the point clouds. Moreover, with a larger distance, less points and a smaller part of the spheres are captured. No significant difference between the individual calibration models is observed. This is expected since the spheres with a diameter of 32 mm are too small for the distortion effects due to refraction to have a high impact.

The middle row in Fig. 4.28 shows the size error of the sphere diameters depending on the measurement range. Here, the results for the point cloud reconstructed from the laser line with the larger baseline $b_1 \approx 0.3$ m are shown. On the left, a scatter plot of the errors of the individual spheres, and on the right, the mean sphere size error per scan is depicted. The error follows a similar trend for all calibration models. The mean error over the whole measurement range for the *Refractive Model* with or without line distortion applied is slightly lower than the *Brown Model*. The reason for this is a higher scale error of the point clouds created using the *Brown Model* since the baselines of the two laser lines are estimated less accurately in the calibration process using the *Brown Model*. Overall, the mean error of the diameters of the spheres is in the range ± 1 mm, and the maximum deviation over the whole measurement range is below 2 mm.

The bottom row of Fig. 4.28 shows a comparison of the mean sphere form error and size error for the two laser lines. Here, only the result for the explicit refractive calibration with *Line Distortion* is reported. In close range, the errors for the two different baselines are comparable.

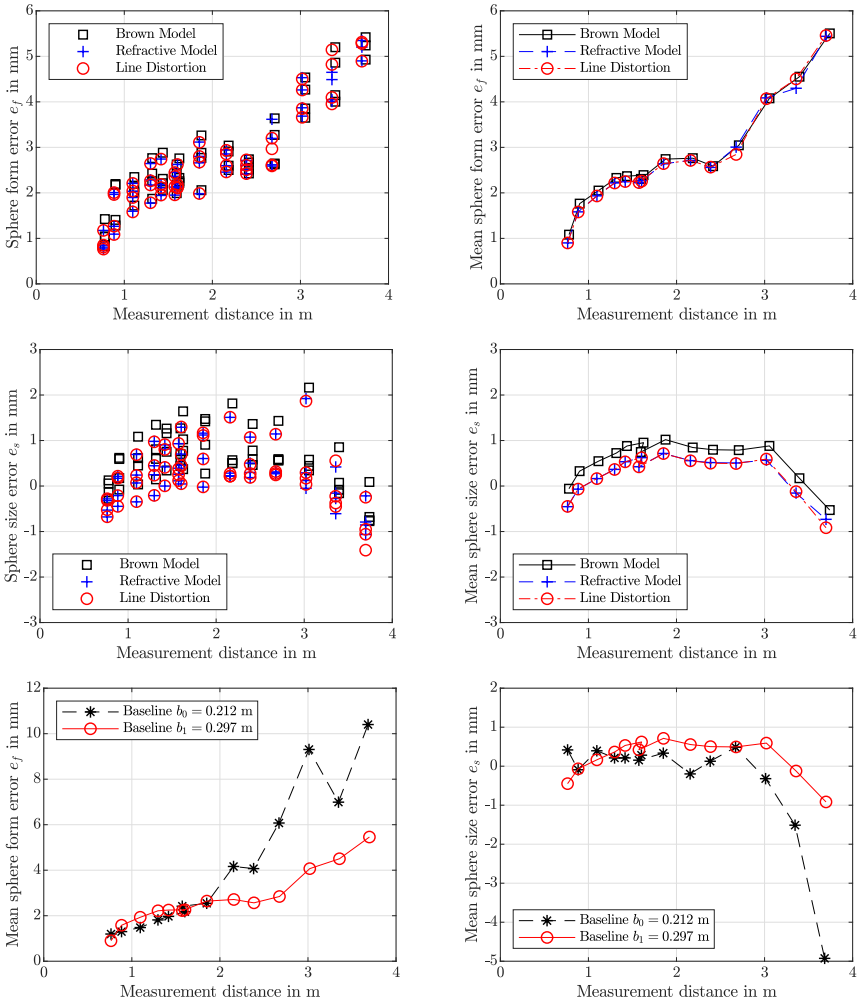


Fig. 4.28: Sphere probing error. Top row: Sphere form errors for the individual spheres and mean sphere form errors depending on the measurement distance. Middle row: Sphere size errors for the individual spheres and mean sphere size errors depending on the measurement distance. Bottom row: Mean sphere form and mean sphere size error for the two laser lines of the scanner with different baselines.

However, at distances larger than two meters, the errors for the laser line projection with the smaller baseline $b_p \approx 0.2\,\text{m}$ is significantly higher. This effect is partially due to the smaller baseline, which causes higher noise and measurement errors. A second effect here is the different geometry of the two laser line projections. This results in less of the surface of the spheres being captured in the point clouds of the laser line with a smaller baseline. As a result, the sphere fit is less stable and errors in diameter of the best-fit sphere of up to 8 mm and high form errors of up to 12 mm are observed.

4.6.4 Sphere Spacing Error

The sphere spacing error is computed on the exact same input data as the sphere probing error. However, in contrast to the sphere probing error, spheres with a fixed diameter of 32 mm are fitted to the point cloud. All possible distances between the four spheres are checked in the evaluation. This means the error is computed for the four distances of 10 cm between adjacent spheres as well as the two diagonal distances of 14.14 cm.

The results for the sphere spacing error e_l depending on measurement distance are reported in Fig. 4.29. The left graphs show a scatter plot of the six individual distances between the spheres. The right graphs show the mean sphere spacing error per scan.

In the top row, the errors are reported for the different calibration models and the laser line with a larger baseline. Here, the effect of the higher scale error of the calibration using the *Brown Model* becomes visible again. Compensation of line distortion does not yield different results since only distances in the range below 15 cm are evaluated, which means the errors due to line distortion are bounded. Due to the different scale error, the calibration using the *Brown Model* yields distance errors in the range of up to 7 mm, while all the observed errors of the *Refractive Model* are below 6 mm. The average errors using the *Brown Model* are in the range of 3 mm to 5 mm while the average errors using the *Refractive Model* are in the range of 1 mm to 3 mm.

The bottom row compares the sphere spacing errors for the two laser lines with different baselines using the *Line Distortion* model. Here, the trend of the mean sphere spacing error is similar for the two lines. Since spheres with fixed diameter are fit to the point cloud, the higher point errors of the reconstruction with a smaller baseline do not affect the error measurement as strongly as the sphere probing error. However, looking at the scatter plot of the individual errors, the spread of the errors is significantly higher for the line with smaller baseline at larger measurement distances.

4.6.5 Length Measurement Error of Scalebars

For the evaluation of length measurement errors of large distances throughout the measurement volume, a target with a size of $1 \times 1 \times 1\,\text{m}$ with scalebars was scanned. The cube was scanned at a distance of 3.2 m from the scanner. The setup is depicted in the top left image of Fig. 4.30, and the resulting point cloud colored by intensity is shown in the top right image.

This experiment is only performed with the refractive model with line distortion compensation. The two laser lines are not evaluated separately since, due to the measurement distance and geometry of the scanner, not all targets were visible in the point clouds created using only

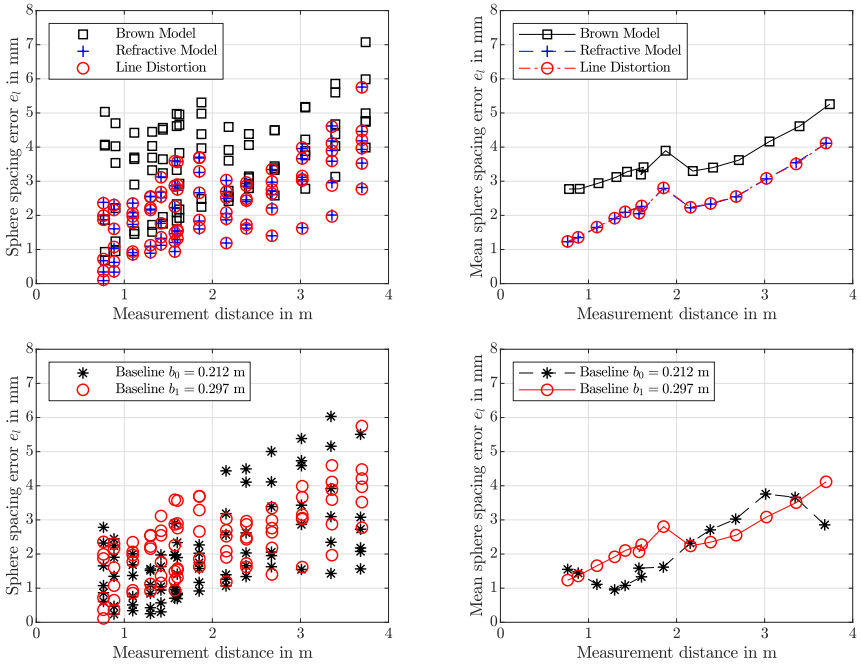


Fig. 4.29: Sphere spacing error. Top row: Sphere spacing errors for all combinations of distances between the four spheres and mean sphere spacing error depending on the measurement range. Bottom row: Sphere spacing errors for all combinations of distances between the four spheres and mean sphere spacing error depending on the measurement range for the two laser lines of the scanner with different baselines.

a single laser line. The image below shows the extraction steps for the point positions from the point cloud. The glass sheets are manually segmented. Then, a threshold filter by intensity is applied to the point cloud. The circles are manually separated using box filters, and finally the centroid is estimated for each coded target.

Of the seven scalebars mounted in the cubical frame, one scalebar is completely occluded. On the remaining six scalebars six coded targets are not evaluable due to occlusion by other scalebars or the frame of the target. This leads to a total of 29 points extracted from the scan. The distribution of the extracted points is visualized in the left image of Fig. 4.31. In the right image, the length measurement errors for all combinations of distances between the points of each individual scalebar are plotted. A total of 62 different distances are evaluable. The errors are in the range of −6.7 mm to 4.2 mm The RMS error of all length measurements is 2.8 mm.

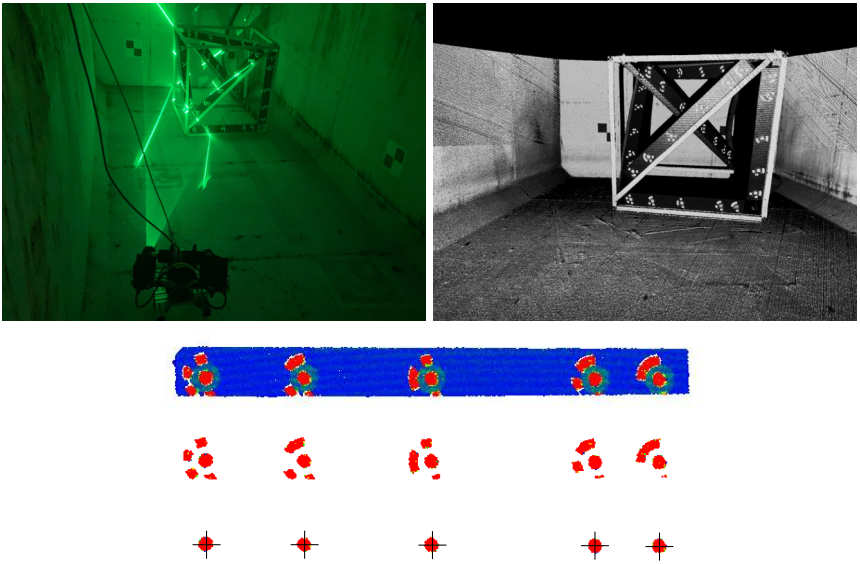


Fig. 4.30: Measurement of the scalebar target at a distance of 3.2 m from the scanner. Top left: Scanner and scalebar target deployed in the water tank. Top right: Resulting point cloud colored by intensity. Bottom: Extraction steps for the positions of the coded circles using filtering of the point cloud by intensity threshold.

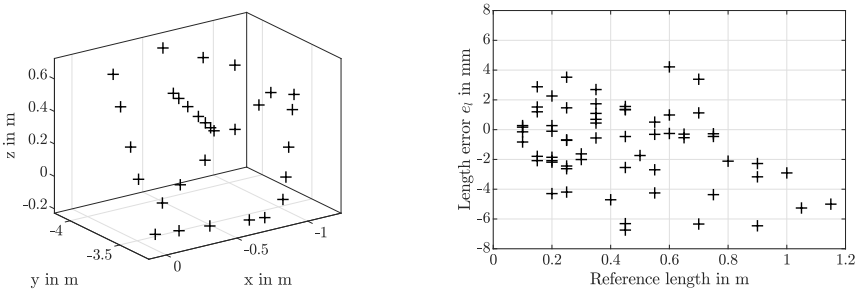


Fig. 4.31: Length measurement errors of the scalebar target. Left: Measured points extracted from the point cloud. Right: Errors of length measurements depending on the point distance.

4.6.6 Flatness Error

The flatness error is evaluated on 13 scans of the plane target at varying measurement distances. For all scans, the planar target with a size of roughly 60 cm by 40 cm is approximately parallel to the scanner. For close measurements, only part of the target is in the field of view, as is visible in the left image in Fig. 4.32. Therefore, for close measurements, the number of points does not increase quadratically. This is visible in the right image of Fig. 4.32, which shows the graph of the number of points reconstructed on the planar target depending on the measurement distance.

The plane is cut out of the point cloud for evaluation manually. A plane is fitted to the remaining points in a least squares optimization. No denoising of the scans or averaging over multiple scans was performed. Following the recommendations of VDI/VDE 2634 a maximum of 0.3% of outliers are removed, and the flatness error e_p is computed.

The result is reported in Fig. 4.33. Here, the errors for the points of the laser line with larger

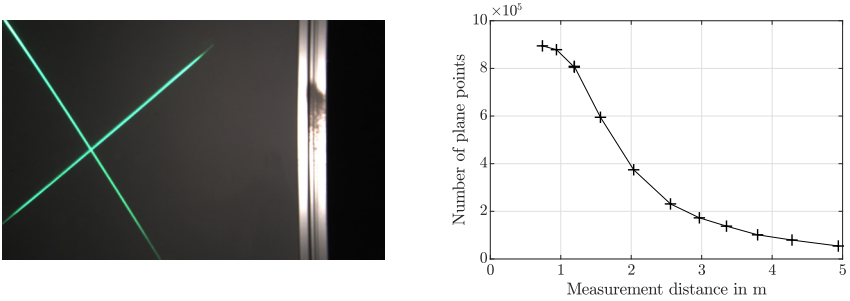


Fig. 4.32: Size of the plane in the point cloud. Left: Image of the scanner camera of the plane target at close distance. Right: Number of plane points used for the evaluation.

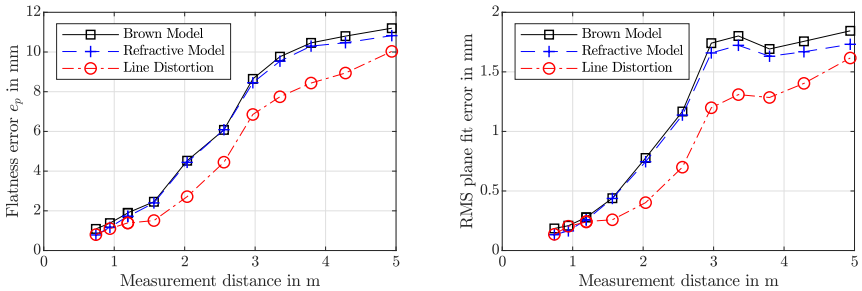


Fig. 4.33: Flatness error for different measurement distances. Left: Plot of flatness error e_p. Right: Plot of root mean squared error of the points to the best-fit plane.

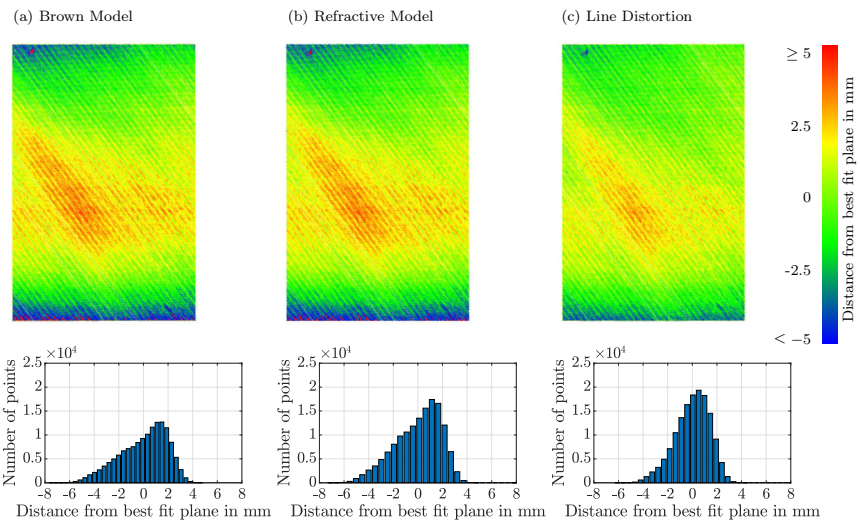


Fig. 4.34: Point clouds of the plane target at a measurement distance of 3.35 m colored by the distance from the best-fit plane and histogram of the distances. Left: Brown model, Middle: Refractive model, Right: Refractive model with line distortion.

baseline $b_1 \approx 0.3$ m are reported. The two laser lines cannot be compared directly on the data because, due to the geometry, the scan coverage is different for the two laser lines. However, it is expected that the laser line with a larger baseline performs better in this test. The flatness error is evaluated for the *Brown Model* without explicit refraction correction, the *Refractive Model* and the additional compensation of the *Line Distortion*. The left image shows the computed flatness error for different measurement distances. The right image shows the RMS error of the plane fit. Errors of the *Brown Model* and the *Refractive Model* are mostly similar. Only at close distance and at large measurement distances the *Refractive Model* performs slightly better. This is due to smaller errors close to the image borders if explicit refraction correction is used. Moreover, the distortion parameters are mainly optimized for a measurement distance of about 2 m since most of the camera calibration images are taken at this distance in the calibration data set. Therefore, higher errors at the edges of the measurement range are expected. Note that with larger measurement distances, the flatness error does not increase as rapidly as for measurements at close range. This is partially an effect of the evaluation methodology. With larger distances less points are sampled on the plane target, as is visible in the left image of Fig. 4.32, and the plane target does not cover the entire FoV anymore. A significant improvement is visible using the additional *Line Distortion* compensation. Without this compensation, the plane exhibits a larger deformation.

This is also visible in Fig. 4.34. Here, scans of the plane target at a measurement distance

of 3.35 m are visualized. The top row shows the scans colored by the signed distance from the best-fit plane for the point clouds using the three calibration results. In the bottom row, the corresponding error histograms are depicted. The point clouds without line distortion compensation are more strongly deformed, which is especially visible in the errors close to the outer edges of the plane target scans. Moreover, the error histograms of the *Brown Model* and *Refractive Model* are skewed. The error histogram of the *Line Distortion* case resembles more closely a Gaussian distribution.

4.6.7 Measurement Error of Free-form Surfaces

To evaluate the accuracy of scanning free-form surfaces, an underwater scan of a dwarf figure is compared with a reference scan. The reference scan was captured with a FARO ScanArm. The dwarf figure is captured in water at a distance of approximately 0.9 m. The measured point cloud of the dwarf with the underwater scanner has 0.35 million points, and the reference point cloud of the complete figure has a size of 2.70 million points.

Note that the reference model, which was scanned from all sides and integrated into a closed

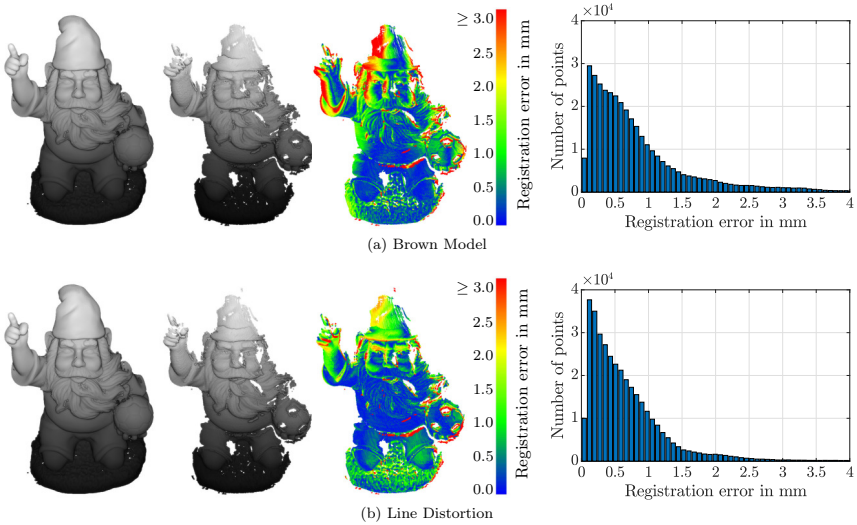
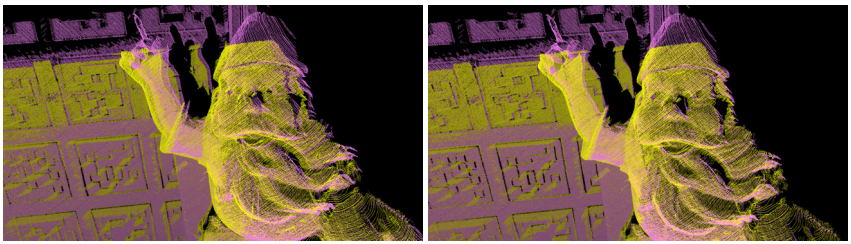


Fig. 4.35: Comparison of dwarf figure scanned with the FARO ScanArm and the underwater scanner. The top row shows the comparison for reconstruction using the *Brown Model* calibration. The bottom row shows the result for reconstruction using the refractive model with *Line Distortion* compensation. From left to right: Reference scan colored by height, underwater scan colored by height, underwater scan colored by registration error, histogram of registration errors. For visualization purposes, the point clouds are rendered with larger point sizes and shading.



(a) Brown Model (b) Line Distortion

Fig. 4.36: Point clouds colored separately by the scans created by the two individual laser lines in yellow and in pink. The reconstruction using the *Brown Model* on the left shows a higher misalignment between the two point clouds of the individual laser lines.

model, is compared with a single scan of the underwater scanner. The underwater scan was taken with the scanner placed lower than the dwarf during scanning. The top of the dwarf is close to the top image border of the scanner camera. Therefore, parts of the dwarf, especially around the nose and eyes, cannot be captured well due to the scanning angle and occlusion. Moreover, some parts of the dwarf that are colored with dark colors, such as the black spots on the soccer ball, are not captured completely.

The reference scan is aligned with the underwater scan using the ICP algorithm. Then, the registration error, or cloud-to-cloud distance, is computed for each point of the underwater scan.

Fig. 4.35 depicts the scan results and the comparison with the reference scan. The top row shows the result using the *Brown Model* while the bottom row shows results for the refractive model with *Line Distortion* compensation. The left dwarf in both rows is the point cloud of the reference scan colored by height. The middle dwarf image shows the underwater scan colored by height, and the right image shows the underwater scan colored by the registration error between the reference scan and the reconstruction. For visualization purposes, the point clouds are rendered with larger point sizes and shading. On the right, the histogram of the errors is depicted.

For the scan reconstructed with the *Brown Model* some misalignment is visible between the points reconstructed by the two laser lines. This is visible, especially in the raised hand/arm of the dwarf and the cap. The fingers of the dwarf are misaligned as well as the arm is reconstructed too wide. Fig. 4.36 shows the point clouds of the two laser lines rendered with different colors in yellow and pink. Here, the misalignment is especially visible in the left image created using the *Brown Model*. The arm of the dwarf and the AprilTags in the background show a stronger misalignment compared to the reconstruction with explicit compensation of refraction effects.

At close distances, the distortion parameters of the *Brown Model* do not accurately model the refraction effects close to the image border. Therefore, the parts of the dwarf that are scanned close to the image border are distorted. This is also visible by the higher errors of the top part of the dwarf figure created using the *Brown Model* visualized in Fig. 4.35.

In total, the scan of the dwarf figure computed using explicit modeling of refraction effects is more accurate. For the reconstruction created using the *Brown Model* 95 % of the points have a

registration error below 2.5 mm. In contrast, 95 % of the points have a registration error below 1.8 mm for the refractive model with *Line Distortion*.

4.6.8 Depth and Spatial Resolution

Besides accuracy, it is interesting what structures the scanning system is able to resolve. The depth resolution depends mainly on the baseline of the laser projection, pixel size and focal length. Horizontal and lateral resolution are additionally affected by the motor's rotation speed and the width of the laser line. Therefore, the achievable depth and spatial resolution are considered separately. Moreover, it is interesting what kind of subpixel accuracy is achieved for localizing the laser line in real-world conditions. While often in the literature, resolution is approximated using the triangulation equation and a theoretical assumption of a resolvable pixel distance, e.g., 1/10 of a pixel, it is not clear how detection resolution in image space translates to resolution in object space. This is approximated with tests using the described resolution targets with staircase patterns.

The two staircases of the targets are approximately aligned to follow the horizontal and vertical axis of the camera. A total of 11 scans were captured of the target with constant run size for evaluation of depth resolution at a distance of 0.6 m to 3 m. The target with decreasing run and rise for the evaluation of spatial resolution was captured with 13 scans at varying distances in the range of 0.6 m to 4 m.

The smallest step size that is resolved in the scan was extracted manually. Hence, there is some bias in the evaluation since a human is still able to determine structures despite higher numbers of outliers and noise, which an automatic algorithm might reject. Fig. 4.37 shows scans of the resolution target at different distances. From left to right the scans at approximately 0.85 m, 1.5 m, and 2.3 m are depicted. Moreover, a detailed view of the steps with a rise (step height) of less than 1.5 mm is shown. In the left point cloud, all steps are identifiable to the human eye. However, the smallest step is not considered resolvable due to the small difference to noise. In the point cloud in the middle, the sub-millimeter steps are not resolved. In the point

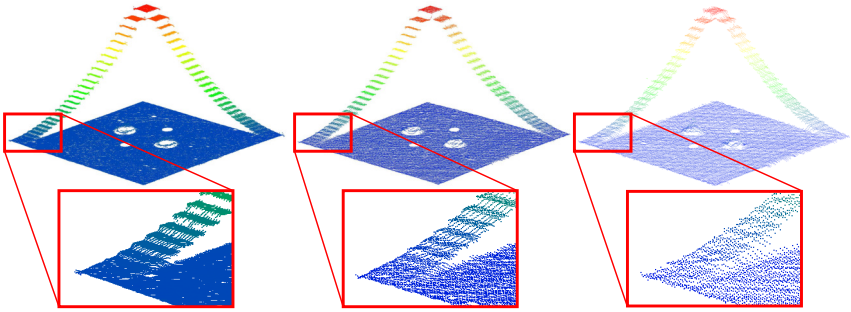


Fig. 4.37: Scans of resolution target and detail view at measurement distance 0.85 m, 1.5 m, and 2.3 m (from left to right).

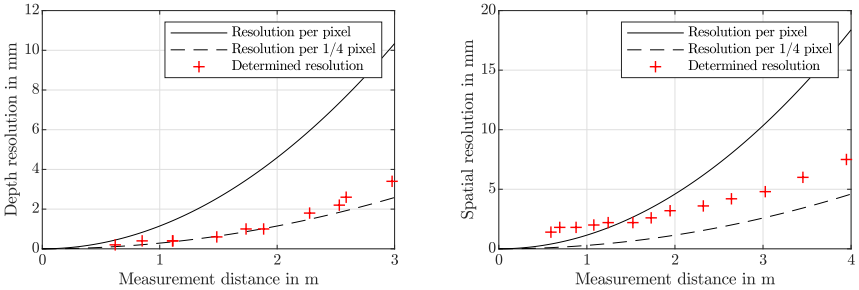


Fig. 4.38: Smallest resolvable staircase height in the point cloud for different measurement distances. Left: Plot of the depth resolution estimated from the stair case pattern with decreasing rise and constant run. Right: Plot of the spatial resolution from the staircase pattern with decreasing rise and run.

cloud on the right, none of the steps in the crop of the detail view are resolved.

The results of this evaluation are summarized in Fig. 4.38. In the graphs, the curves for the theoretical depth resolution are plotted based on Eq. 4.3 as explained in Section 4.2. The curves for the theoretical resolution for pixel-accurate extraction as well as 1/4 subpixel accurate laser curve extraction are shown. Here, a pixel size of 5.86 µm, the equivalent focal length in water of 17 mm and a baseline of 30 cm was assumed. This corresponds to the larger baseline of the two laser projectors, which is more representative due to the manual extraction of the values from the point cloud. Distances smaller than 0.5 m cannot be evaluated due to the geometry of the scanner.

The left image in Fig. 4.38 shows the experimentally determined depth resolution depending on measurement distance. The curve roughly follows the theoretical values for 1/4 pixel resolution in image space. The deviation of the theoretical curve is higher for larger measurement distances. In the right image, the experimentally determined spatial resolution is plotted. As expected, the determined spatial resolution is lower than the depth resolution. It corresponds roughly to the theoretical value for 1/2 pixel resolution in image space. However, note that below about 1.5 m, the spatial resolution is mostly constant and no longer improves with smaller measurement distances. The reason for this is the focus point of the laser projection and camera. Below 1.5 m, the sharpness of the camera image decreases and the laser line becomes thicker, consequently decreasing the minimum resolvable structure.

4.6.9 Discussion of the Results

The overall results of the scanner evaluation are summarized in Tab. 4.3. The average errors for all measurements below 2 m, above 2 m and the average of all measured errors are reported. Here, only the findings for the best performing calibration are reported, which is the refractive model with line distortion compensation. The average errors are in the millimeter range. However, in the evaluation, individual errors of up to about 1 cm were observed. The main limitation of the scanner is the small baseline of 0.3 m. Increasing the baseline improves the results. Moreover,

Table 4.3: Summary of the underwater scanner evaluation results with the best performing calibration model, which is the refractive model with line distortion compensation.

Quality parameter	**Close range** ($\leq$ 2 m)	**Far range** ($\geq$ 2 m)	**Full range**
Mean sphere form error e_f	2.0 mm	3.7 mm	2.7 mm
Mean sphere size error $\lvert e_s \rvert$	0.5 mm	0.6 mm	0.5 mm
Mean sphere spacing error $\lvert e_l \rvert$	1.9 mm	3.0 mm	2.4 mm
Mean flatness error e_p	1.2 mm	7.0 mm	4.6 mm
Mean depth resolution	0.6 mm	2.5 mm	1.3 mm
Mean spatial resolution	2.2 mm	5.2 mm	3.3 mm

high flatness errors are observed at larger measurement distances. That suggests that some residual model errors remain. The laser projection model can be improved by accurately determining the projection center and explicitly modeling the refraction effects. Another source of inaccuracies is the precession of the rotation axis of the scanner. However, accurately compensating this movement is difficult since the change in orientation of the axis is not completely repeatable and alternates based on the orientation of the scanner due to gravity.

Comparison with results for underwater scanning in the literature is difficult because most publications do not report errors at all, use different quality parameters or evaluation methods that are not comparable. Gu et al. [147] evaluate an underwater line laser scanner with a baseline of 0.8 m at measurement distances of 1 to 2.5 m. They compute the error of the 3D data separately for the different spatial axes using plane fits (Z-axis) or line fits (X/Y-axis). The reported mean error is 2.30 mm for the X-axis, 1.31 mm for the Y-axis and 0.03 mm for the Z-axis. Scanned reference planes show an RMSE of 0.09 mm and a maximum deviation of 0.39 mm. In experiments in clear water described by Niemeyer et al. [241], the 2G Robotics ULS-200 underwater laser scanner achieves a measurement noise of approximately 1.3 mm at a close range of just under 1 m. The noise increases to 3.6 mm at a distance of 2.3 m. At 3.7 m, which is outside of the specified measurement range, a measurement noise of 20.4 mm was observed. The Kraken SeaVision [186] 3D laser system achieves similar results of a measurement noise between 1.5 mm and 1.6 mm at a distance of 1.5 m. For offshore tests reported in [242] a measurement noise of 1 - 2 mm at 1.2 m distance and a measurement noise of 4 mm at 2.7 m distance is achieved using the SeaVision system. Bruno et al. [68] compare scans of reference objects that are acquired with an underwater stereo vision system with pattern projection. They compute the distance of the scan from the best fit of the reference geometry, which is a planar or cylindrical surface. A maximum deviation of 0.1 mm and a standard deviation of 0.05 mm is observed at a measurement distance of 0.8 m. At 1.2 m distance the maximum deviation is reported as 0.4 mm and the standard deviation as 0.2 mm. Buschinelli et al. [75] report sphere form errors below 0.18 mm and sphere spacing errors below 0.24 mm for close range measurements at distances below 25 cm with a stereo system.

In summary, the length and flatness errors reported in the literature for optical underwater scanning systems are in the range of 0.05 to 1 % of the measurement distance. The errors are

Table 4.4: Results of Kersten et al. [180] for different photogrammetric handheld 3D scanners for the quality parameters according to VDI/VDE 2634.

System	**Form error** e_f	**Size error** $\lvert e_s \rvert$	**Spacing error** $\lvert e_l \rvert$	**Flatness error** e_p
Mantis PocketScan 3D	4.56 mm	4.33 mm	2.89 mm	1.01 mm
Mantis F5-SR	2.71 mm	0.58 mm	1.38 mm	0.68 mm
Mantis F5-B	10.48 mm	0.45 mm	0.11 mm	0.35 mm
Mantis F6 SMART	5.02 mm	1.94 mm	0.25 mm	0.55 mm
Artec Spider	0.41 mm	0.12 mm	1.28 mm	2.01 mm
FreeScan X7	1.01 mm	0.10 mm	0.23 mm	0.12 mm
HandySCAN 700	0.72 mm	0.23 mm	0.10 mm	0.10 mm

typically in the range of a few millimeters. Some results show errors in the sub-millimeter range for close distances. However, measurement errors in the sub-millimeter range are typically only observed for stereo systems with pattern projection [47, 63, 68, 75]. For comparison, the achieved results of the proposed system of the work at hand show errors below 3 mm at a distance of 2 m, which is a percentage error of 0.15 %.

Tab. 4.4 summarizes the result of an evaluation of handheld 3D scanners in the price range of €2.000 to €49.000 based on VDI/VDE 2634 by Kersten et al. [180]. The evaluation of Kersten et al. is not directly comparable to the evaluation performed in this section. Kersten et al. use larger spheres and reference structures that are more accurately calibrated. Moreover, the handheld scanners are evaluated at close range, and the commercial scanning systems perform integration of multiple area scans and run denoising filters on the scans. Compared to that, the individual underwater scans are created by rotating the scanner. While it is prudent not to compare the values directly, it is visible that the observed errors in the evaluation of the underwater scanner are not magnitudes higher, but at least for the close-range measurements in the range of the errors observed by [180] for state-of-the-art handheld structured light systems in air.

4.7 Laser Scanning of Semi-submerged Objects

In [21–23] we look at the 3D acquisition of semi-submerged structures with a triangulation-based underwater laser scanning system. The motivation is that we want to simultaneously capture data above and below water to create a consistent model without any gaps. In order to reconstruct precise surface models of the object, it is necessary to model and correct for the refraction of the laser line and camera rays at the water-air boundary.

Since the employed structured light scanner captures only 2D profiles, the scanner is moved in lab experiments using an industrial manipulator to create a complete 3D scan of the scene. The experimental setup is depicted in Fig. 4.39. The 6-DoF movement of the scanner system is captured externally with a tracking system. This work assumes a planar air-water interface.

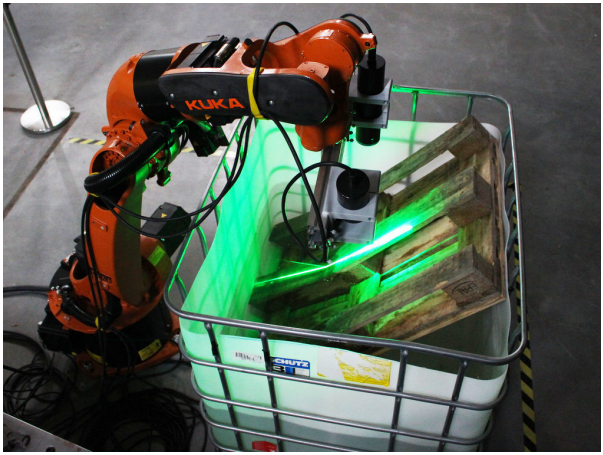


Fig. 4.39: Prototype underwater line laser scanner mounted on an industrial manipulator for scanning a semi-submerged scene.

From the sensor data the water surface is extracted and all 3D point measurements inside the water body are corrected using a ray-based approach.

Correction of optical measurements through the water surface has been studied for various applications of laser scanning and photogrammetry, e.g., airborne laser scanning of coast lines [159, 165, 283, 331] or sunken archaeological sites [101], photogrammetric measurements of river beds [338] or convection flow estimation in a glass vessel [205].

[235] investigated the application of light stripe projection and photogrammetric stereo on multimedia scanning. They employ a camera-projector setup in air that measures objects in a water tank. To correct for the refraction at the transition between air and water, the refraction of the camera ray as well as the projector ray need to be considered. For calibration, two planes at a known distance are placed in the water tank.

[182] apply a Microsoft Kinect RGB-D camera to capture bathymetry with the sensor placed above the water. Although the infrared pattern projector of the Kinect suffers from high absorption in water, measurements at depths of up to 40cm were achieved. The sensor was mounted parallel to the water surface, and the distance between the water plane and the sensor was measured. The point clouds captured by the Kinect were then corrected based on a ray-based refraction model.

In [227] a photogrammetric survey of a floating ship above-the-water and underwater is described. The images are captured separately in air and water and subsequently aligned using photogrammetric targets.

Some of the published literature also addresses inaccuracies introduced by waves, which are very difficult to model precisely [132, 251, 280].

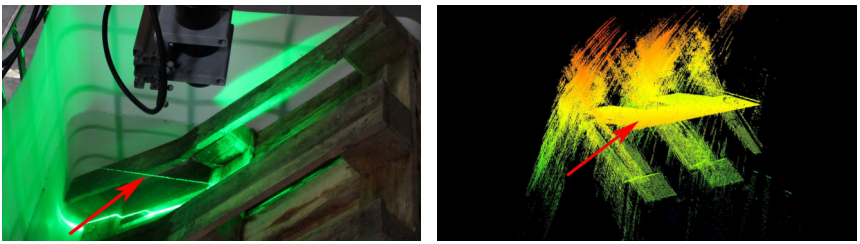


Fig. 4.40: Reflections of the laser line projector on the water surface are exploited to estimate the water plane. Left: Reflection of laser line. Right: Reconstructed point cloud including the water plane.

4.7.1 Refractive Correction Approach

In this work, a flat water surface is assumed, which is semi-automatically detected and estimated using only the measurements of the scanner. As depicted in the left image of Fig. 4.40 a reflection of the laser line on the water surface is visible. If a low threshold for the laser line extraction of the structured light scanner is set, these reflections are extracted as laser line points and reconstructed. The resulting point cloud is shown in the right image of Fig. 4.40. It is seen that the water plane is visible in the 3D point cloud. The water surface is estimated by a robust fit to the dominant plane of these measurements. The reflections are later removed to create the final 3D scan of the scene by re-running the laser line extraction with an adjusted threshold.

To create accurate scans of the semi-submerged structures, the refraction at the air-water interface needs to be considered. After performing 3D reconstruction based on light section and applying the 6-DoF scanner trajectory, the resulting point cloud is split into an above-the-water and an underwater point cloud based on the estimated water plane. Only the points that lie below the water plane are then corrected based on a ray-based approach similar to the work of [182]. Since, in this case, the laser ray and camera ray do not follow the same optical path, the refraction of the camera rays and the laser rays at the water surface need to be computed.

The principles of the ray-based correction approach are depicted in Fig. 4.41. Using the 6-DoF trajectory of the scanner and the estimated water surface, for each individual line scan of the structured light scanner the position of the camera center $\mathbf{C}$, the position of the laser projection center $\mathbf{L}$ and the water plane are established in a common world coordinate system. For each point $\mathbf{P}$ of the structured light scan located below the water surface, a ray-based approach is applied to find the corrected point position $\mathbf{P}'$.

First, the intersection points of the camera ray and laser ray with the water plane are computed. $\mathbf{S}_C$ is the intersection point of the line from the camera center $\mathbf{C}$ to the point $\mathbf{P}$, which is visualized in red in Fig. 4.41. $\mathbf{S}_L$ is the intersection point of the line from the laser projection center $\mathbf{L}$ to the point $\mathbf{P}$, which is visualized in green in Fig. 4.41.

Then, at the intersection points $\mathbf{S}_C$ and $\mathbf{S}_L$, the refracted camera and laser ray is computed based on Eq. 2.11.

In principle, the corrected point $\mathbf{P}'$ results from intersecting the refracted camera and laser ray. However, it is not directly known which laser ray belongs to the observed camera ray.

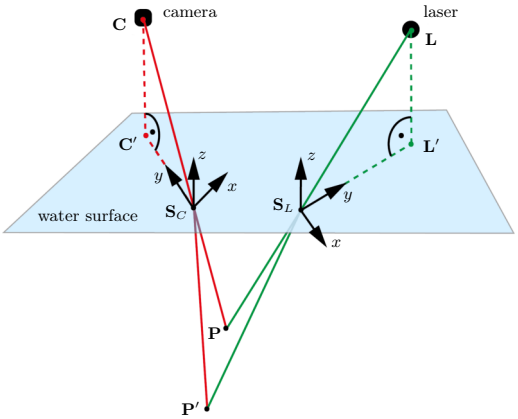


Fig. 4.41: Geometry of the ray-based correction approach. The point $\mathbf{P}$ is the initial reconstructed point using light section without accounting for refraction and the point $\mathbf{P}'$ is the point after refractive correction. $\mathbf{C}$ is the camera and $\mathbf{L}$ the laser projection center. $\mathbf{C}'$ and $\mathbf{L}'$ are the projections of the camera and laser center along the water plane normal on the water plane. $\mathbf{S}_C$ and $\mathbf{S}_L$ are the respective intersections between the camera and laser ray with the water plane.

Therefore, for an exact solution, it is necessary to solve for the laser ray that intersects with the camera ray and fulfills the constraint of Snell's law.

In [22] we compute an approximate solution using the laser ray derived based on the uncorrected point position $\mathbf{P}$. Instead of computing the intersection of the two refracted rays, the intersection between the refracted camera ray and the refracted laser plane is computed. The plane is constructed from the refracted laser ray and the intersection line between the laser and water plane. This constrains the solution to lie inside this plane, which limits the error due to the inaccurately estimated incident rays.

4.7.2 Scanning of Semi-submerged Objects

Experiments with the scanner mounted to a KUKA KR-16 industrial manipulator are performed to validate the model. This allows to move the scanner with different repeatable angles of incidence for the laser plane with respect to the water surface.

The setup is depicted in Fig. 4.42. Three different scenes are placed in a half-filled water tank with an edge length of 1 m. The first scanned object is a wooden Euro-pallet, as shown in Fig. 4.39. The second scene, depicted in Fig. 4.42, is more complex and contains a dwarf figure, a coffee pot, a plastic pipe and two chessboards. The last scene only contains a large chessboard. The objects are placed in front of the scanner at a distance of approximately 1 m.

First, the scanner is moved along a linear track with the laser projector pointing straight down, such that the laser plane is orthogonal to the water plane. Second, multiple scans are

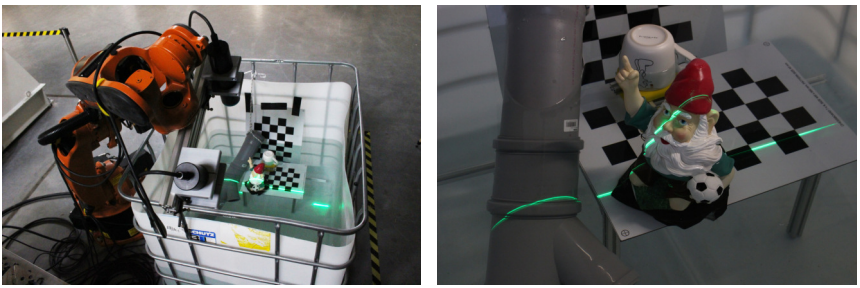


Fig. 4.42: Left: Scanning a semi-submerged scene with the scanner mounted to an industrial manipulator. Right: Detail view of the captured scene with a laser line projected by the scanner.

recorded with the same motion and increasing rotation of the scanner in 5° steps up to an angle of 20°. This way, the influence of different angles of incidence of the laser projector rays is observed. Angles larger than 20° are not considered for experimental reasons due to constraints on the possible scanner movements without collision with the water tank.

To record the trajectory of the scanner, an external OptiTrack V120:Trio 6-DoF tracking system is employed. Based on the scanner trajectory, measured by the tracking system, all line scans are combined to a complete 3D scan of the entire scene.

The resulting point clouds of two different scenes are visualized in Fig. 4.43. The top scan shows a scene with multiple objects, and in the bottom scan a semi-submerged chessboard is captured. The point clouds are colored by height. The images on the left show the uncorrected initial 3D reconstruction result, while the images on the right show the result after refractive correction. The point clouds are the combined point clouds of multiple scans, each consisting of around 800 scan lines, with varying scanner rotation between 0° and 20°. The influence of refraction and the different angles of incidence are most clearly visible in the top left image of Fig. 4.43. Especially the measurements of the bottom of the water tank do not line up between the individual scans with different scanner rotations. In the corrected scan result on the right, the point cloud measurements of the floor of the water tank are consistent. Please note the compressed chessboard in the bottom left image to compare the uncorrected and corrected scans.

To illustrate this more clearly, Fig. 4.44 shows cross-sections of uncorrected scans of a wooden Euro-pallet for different scanner angles and the corrected scan. In this visualization, the points above the water are colored in red and the points below the water surface are colored in blue. With increasing scanner angle, the underwater points become more compressed in the vertical direction, and the bend visible in the wooden plank becomes stronger. At an scanner angle of 20° there is a strong deviation between the surface of the plank in the underwater and the above-the-water scan. However, after correction, these errors are not visible anymore, and the points of the plank surface form a straight line. This is true for all scans independent of the scanner angle. The image in Fig. 4.44 bottom right shows a cross-section of the combined point cloud of all corrected scans.

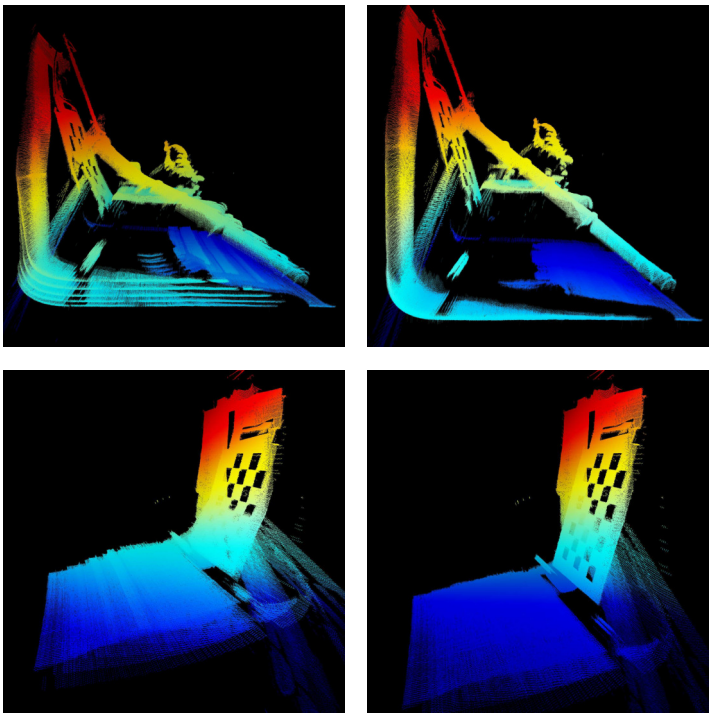


Fig. 4.43: Point clouds of the initial uncorrected reconstruction on the left side and the corrected result on the right side. The point clouds are colored by height.

To investigate this in more detail, measurements are performed using the scans of the chessboard, which are shown in the bottom images of Fig. 4.43. In this scene, the chessboard is positioned half below and half above the water table. Since the chessboard provides a planar reference surface, the plane normals fitted to the underwater and above-the-water parts are compared.

For each angle of incidence, the submerged part and the part above water are considered as separate planes. Subsequently, the angle between the normals of these two planes is calculated. For the uncorrected and corrected point clouds, the angular error results for the different scanner angles are listed in Tab. 4.5. The steeper the angle of incidence, the faster the angular error increases for the uncorrected scan.

After the correction of each scan, low angular errors are observed. For the vertical scanner alignment the error is reduced from 5.56° to 0.19°. The remaining error for a scanner alignment

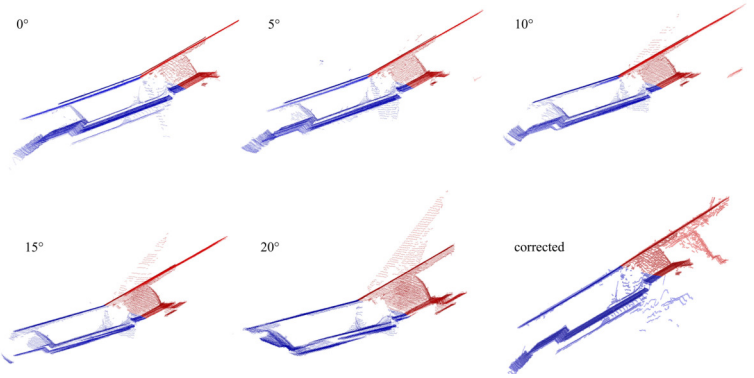


Fig. 4.44: Cross-sections through 3D scans of a wooden pallet. Points below the water are colored in blue and points above the water in red. Starting from the top left with the scanner vertically positioned (0°), scans with different rotation angles of the structured light scanner are depicted. The last figure, bottom right, shows the cross-section of the reconstruction result after applying the refractive correction method. In the corrected scan the surface of the wooden pallet lines up well.

Scanner angle	Angular error uncorrected	Angular error corrected
0°	5.56°	0.19°
5°	6.65°	0.34°
10°	8.12°	0.37°
15°	10.76°	0.35°
20°	14.53°	0.53°

Table 4.5: Angular errors between the plane normals of the chessboard above and below water for the initial uncorrected and the corrected reconstruction result.

of 5°, 10° and 15° is similar, ranging from around 0.34° up to 0.37°. The largest remaining error of 0.53° is observed for a scanner rotation of 20°. This shows that the angular error is significantly reduced after applying the refractive correction and supports the visual observation that the large errors introduced due to refraction are almost completely removed.

4.8 Discussion

This chapter describes the hardware development of an underwater laser line scanning system, the data processing and the necessary calibration approaches for 3D reconstruction. The measurement quality of the proposed approach is validated in experiments in a test tank using reference objects, and the different calibration processes are compared. In contrast to other approaches, such as [147], the proposed calibration technique requires only observations of a calibration fix-

ture. It does not require disassembling of the camera, e.g., to find the accurate camera model in air, for the underwater calibration.

As a benchmark, the work at hand investigates surfaces and distance measurement errors in a small volume using sphere targets, length measurement errors of scale bars and the flatness error of a reference plane. Depending on the measurement distance, errors in the range of one millimeter up to less than one centimeter in clear water conditions are observed. At a measurement distance of 2 m the length, shape and flatness errors are below 3 mm, which is a percentage error of 0.15 %. This is comparable to the better performing approaches demonstrated in the literature (cf. Section 4.6.9). Using a staircase pattern, the experiments find a measurement resolution in the millimeter range, which is consistent with the predicted theoretical resolution based on subpixel feature extraction. Besides scanning underwater structures, the 3D acquisition of semi-submerged structures using the system is investigated. In lab experiments, an effective approach for correcting the refraction at a planar water surface is demonstrated.

Chapter 5

Self-calibrating Structured Light

In situ calibration of structured light scanners in underwater environments is time-consuming and complicated. Therefore, this chapter looks at self-calibration techniques for handheld 3D line laser scanning.

In particular, laser scanning with a fixed calibrated camera and a handheld cross-line laser projector is investigated. The proposed system is depicted in Fig. 5.1. The left image shows the setup for scanning in air, and the right image shows the cross-line laser projector integrated in a watertight housing for underwater scanning. Two different laser line colors are employed to facilitate the separation of the two laser lines in a single camera image. A green and a blue line laser is employed since absorption in water is significantly lower for these wavelengths than, for example, a red laser. By employing an off-the-shelf underwater camera and a waterproof housing with high-power line lasers, an affordable 3D scanning solution is built. Experiments demonstrate that the scanning system is applicable for above-the-water as well as underwater scenes.

The developed approach is based on self-calibration techniques proposed by [137], which exploit coplanarity and orthogonality constraints between multiple laser planes to recover the depth information. It is applicable without any prior knowledge of the position and orientation of the laser projector. To capture the full 3D geometry multiple images are necessary, such that the whole scene in the field of view of the camera is illuminated by the lasers. The scene must remain static during capture, and the camera needs to be fixed since the reconstruction algorithm depends on the properties associated with a point in 3D space being illuminated from different positions.

Similar to SfM, the scale of the 3D point cloud is not recovered using only two laser lines due to the perspective projection of the camera. In this work, scale is typically inferred from a known distance of the scanned object or place scale references in the scene. Moreover, scale definition is achieved by co-registration of the scans with data captured using other sensor modalities, e.g., time-of-flight laser scanning.

In general, uncalibrated scanning with the projector moving without restrictions makes it more difficult to obtain accurate scans due to noisy estimates of the laser plane parameters. However, the accuracy of triangulation-based depth estimation is also dependent on the baseline. Uncalibrated Structured Light has the advantage that it is not limited to a fixed distance, and

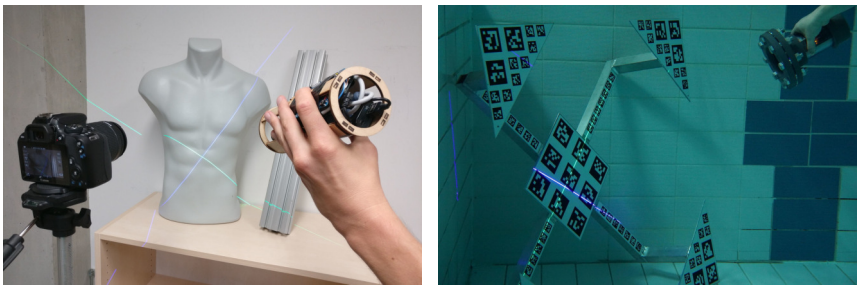


Fig. 5.1: Uncalibrated structured light. Left: Handheld scanning with a static camera and a cross-line laser projector. Right: Cross-line laser projector in a waterproof housing for underwater scanning.

scanning with very large baselines becomes possible. A suitable baseline, which depends on the depth range of the scene, is chosen by simply moving closer or further away from the camera. This allows to record details that would otherwise not appear in scans with a fixed small baseline.

The work described in this chapter was previously published in [8, 9]. The author's contributions include the development and implementation of the proposed self-calibrating structured light approach as well as the experiments and evaluation. In [17] the developed approach was applied for self-calibration of a camera and projector system for medical applications.

5.1 Self-calibration

Exploiting the projection of planar curves on surfaces for recovering 3D shape has been studied for diverse applications, such as automatic calibration of structured light scanners [138], single image 3D reconstruction [324] and shape estimation from cast shadows [57]. Here, the focus lies on self-calibrating line laser scanning, which recovers 3D information from the projection of planar curves. Previous work typically either employs a fixed camera and tries to estimate the plane parameters of the laser planes or works with a setup where the camera and laser are mounted rigidly relative to each other and automatically estimates the extrinsic parameters.

Some methods solve the online calibration problem by placing markers or known reference planes in the scene [343, 349]. For example, [343] proposed a method for a handheld laser line scanning system, which estimates the laser planes by placing the object in front of a corner with two known reference planes. By intersecting the image rays with the two reference planes, the 3D point positions of the laser projection on the reference target is reconstructed. Then, by fitting a plane to these points, the plane parameters of the line laser are computed. This approach became later popular as the *David Laser Scanner*.

The approach in [170] demonstrated calibration of a line laser scanning system where camera and line laser are mounted fixed relative to each other without requiring a special target. Given an initial estimate, the relative transformation parameters between the laser plane and the image plane are refined by matching multiple profiles taken from different viewpoints.

In contrast, Furukawa and Kawasaki proposed a self-calibration method for movable laser planes observed by a static camera that does not require placing any specific objects in the scene [136]. The approach exploits coplanarity constraints from intersection points and additional metric constraints, e.g., the angle between laser planes, to perform 3D reconstruction. It was demonstrated that 3D laser scanning is possible with a simple cross-line laser pattern. Later, Furukawa and Kawasaki extended the approach and showed how additional unknowns, such as the parameters of a pinhole model (without distortion), are estimated if a suitable initial guess is provided [137]. This work follows this direction and extends the method to underwater scanning and adds heuristics and pre-processing steps to improve the results and remove outliers.

The underlying plane parameter estimation problem leads to a linear system of coplanarity constraints for which a direct least-squares approach does not necessarily yield a unique solution. For example, projecting all points of all planes in the same common plane fulfills the coplanarity constraints. However, this does not describe the real scene geometry. In [113] additional constraints on the distance of the points from the best-fitting plane are incorporated in the optimization problem to avoid unmeaningful solutions.

More recently, Nagamatsu et al. showed how coplanarity and epipolar constraints are used for the online self-calibration of a system with multiple line lasers [233]. In this approach, camera and line laser are fixed relative to each other, and the whole system is applied for handheld scanning. This approach leverages Visual Simultaneous Localization and Mapping (SLAM) for mobile scanning. Using the results from Visual SLAM, no additional geometric constraints on the angular relationship of the individual plane are necessary.

5.2 Self-calibrating Cross Line Laser Scanning

The configuration for creating a 3D scan is visualized in Fig. 5.2. The scene is captured with a fixed calibrated camera by moving the handheld laser projector in order to project laser crosses in the scene from different positions. In each image of the video camera, two laser curves are observed, which are known to have plane normals that are perpendicular to each other due to the cross configuration of the employed laser line projector. The camera parameters are explicitly calibrated, because this allows to acquire robust estimates of the distortion parameters, which are especially necessary for underwater imaging and uncorrected optics.

By aggregating a sequence of images over time, many different laser curves are extracted on the image plane. This allows us to exploit the intersections between laser curves to estimate the laser planes.

Solving for the plane parameters is a two-step process: First, coplanarity constraints from intersections between laser curves are exploited. Since the camera is fixed and the scene is static, intersection points correspond to the same 3D point. By extracting many laser curves, many more intersections than the number of laser planes are obtained. This allows to build a linear system to solve for the plane parameters up to a scale and an offset. In the second step, the additional DoF of the parameters are solved by considering the orthogonality constraint between the laser planes in the cross configuration. The plane parameters are found up to a scale by solving a non-linear optimization problem.

Finally, the 3D point positions of each laser curve are computed by intersecting the camera

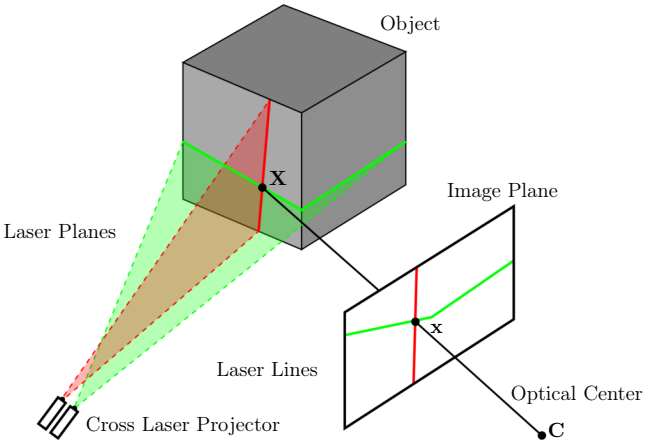


Fig. 5.2: Configuration of the cross line laser projector and a fixed calibrated camera.

rays with the laser planes. The following sections explain the individual steps and employed models in more detail.

Fig. 5.3 shows an image of the scene captured from the perspective of the fixed camera used for scanning, a subset of the extracted laser curves in white, the computed intersection points in red and the reconstructed 3D point cloud. This result was created from 3 min of video recorded at 30 fps. A total of 7,817 valid laser curves were extracted with 1,738,187 intersection points. The 3D reconstruction was computed using a subset of 400 laser planes. The final point cloud created from all valid laser curves has a size of 11,613,200 points. The individual steps and employed models are explained in more detail in the following sections.

5.2.1 Cross Line Detection

To apply the method, the two laser lines need to be extracted separately from the individual images. Since the camera is static, background subtraction is applicable to enhance the contrast of the laser lines. First, the laser line extraction approach as described in Section 4.3 is applied. To distinguish between the two laser lines, thresholds in the HSV color space are applied. This is implemented using look-up tables to speed up color segmentation. Only very low thresholds for saturation and brightness are applied for the detection of the laser line. Depending on the object's surface, the laser line is barely visible and appears desaturated in the image.

After performing laser line extraction, all image coordinates of the detected line points are undistorted based on the camera model. Therefore, the distortion parameters do not have to be considered during the 3D reconstruction step, simplifying the equations presented in the following sections.

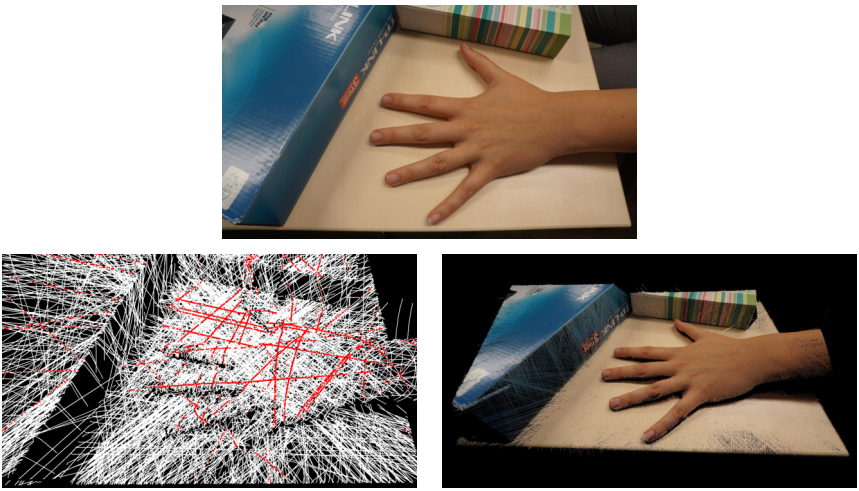


Fig. 5.3: Visualization of the scene, extracted laser curves, and 3D point cloud. Top: Image of the scene from the perspective of the fixed camera used for scanning. Bottom left: Visualization of a subset of the extracted laser curves in white and intersection points in red. Bottom right: Colored 3D point cloud.

5.2.2 3D Reconstruction Using Light Section

If the parameters of the laser plane are known, the 3D coordinates of the detected laser points are found by intersecting the image rays with the laser planes. A line laser can be considered a tool to extract points on the image plane that are projections of object points that lie on the same plane in 3D space. Here, the laser plane π_i is described using the general form

$$\pi_i : \quad a_i X + b_i Y + c_i Z = 1 \ , \tag{5.1}$$

where (a_i, b_i, c_i) are the plane parameters and $\mathbf{X} = (X, Y, Z)^T$ is a point in world coordinates. Using the perspective camera model described in Eq. 3.1, this is expressed as

$$\pi_i : \quad a_i \frac{x - p_x}{f_x} + b_i \frac{y - p_y}{f_y} + c_i = \frac{1}{Z} \ , \tag{5.2}$$

where $\mathbf{x} = (x, y)^T$ are the image coordinates of the projection of $\mathbf{X}$ on the image plane, $\mathbf{p} = (p_x, p_y)^T$ is the principal point and f_x, f_y are the respective focal lengths.

If the plane and camera parameters are known, the coordinates of a 3D object point $\mathbf{X} = (X, Y, Z)^T$ on the plane are computed from its projection on the image plane $\mathbf{x} = (x, y)^T$ by

intersecting the camera ray with the laser plane:

$$\begin{aligned} Z &= \frac{1}{a_i \frac{x-p_x}{f_x} + b_i \frac{y-p_y}{f_y} + c_i} \\ X &= Z \frac{x-p_x}{f_x} \\ Y &= Z \frac{y-p_y}{f_y} \ . \end{aligned} \tag{5.3}$$

5.2.3 Self-calibration Approach

Self-calibration in this work is considered as the problem of estimating the parameters of all observed laser planes. From the recorded sequence of images, the laser curves are extracted as polygonal chains. The points that exist on multiple laser planes are found by intersecting the polylines. This computation is accelerated by spatial sorting, such that only line segments that possibly intersect are tested for intersections. Moreover, the polylines are simplified to reduce the number of line segments. However, this needs to be done with a very low threshold (less than half a pixel) in order not to degrade the accuracy of the extracted intersection positions.

The plane parameters are estimated in a two-step process based on the approach described in [137]. First, by solving a linear system of coplanarity constraints the laser planes are reconstructed up to 4-DoF indeterminacy. Second, further indeterminacies are recovered from the orthogonality constraints between laser planes in the cross-configuration in a non-linear optimization.

First, using Eq. 5.2, the coplanarity constraint between two laser planes π_i and π_j are expressed in the perspective system of the camera for an intersection point $\mathbf{x}_{ij} = (x_{ij}, y_{ij})^T$ as

$$\begin{aligned} &\frac{1}{Z_i(x_{ij}, y_{ij})} - \frac{1}{Z_j(x_{ij}, y_{ij})} = \\ &(a_i - a_j)\frac{x_{ij} - p_x}{f_x} + (b_i - b_j)\frac{y_{ij} - p_y}{f_y} + (c_i - c_j) = 0 \ . \end{aligned} \tag{5.4}$$

These linear equations are combined in a homogeneous linear system:

$$\mathbf{A}\mathbf{v} = 0 \ , \tag{5.5}$$

where $\mathbf{v} = (a_1, b_1, c_1, \ldots, a_N, b_N, c_N)^T$ is the combined vector of the planes' parameters and $\mathbf{A}$ is a matrix whose rows contain $\pm(x_{ij} - p_x){f_x}^{-1}$, $\pm(y_{ij} - p_y){f_y}^{-1}$ and ± 1 at the appropriate columns to form the linear equations of Eq. 5.4.

This problem has a trivial solution for $\mathbf{v}$, which is the zero vector. Therefore, the system is solved under the constraint $\|\mathbf{v}\| = 1$ using SVD. If the system is solvable and it is not a degenerate condition, the perspective solution of the plane parameters (a_p, b_p, c_p) are obtained with 4-DoF indeterminacy. The solution is represented by an arbitrary offset $\mathbf{o}$ and an arbitrary scale s:

$$(a, b, c) = s(a_p, b_p, c_p) + \mathbf{o} \ . \tag{5.6}$$

A degenerate condition is caused, for example, by planes with only collinear intersection points. For example, consider the intersection points one, two and three (visualized by circles) in Fig. 5.4

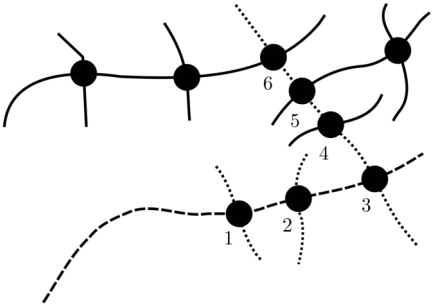


Fig. 5.4: Collinear intersections between laser curves.

that lie on the dashed curve. These intersections are collinear on the image plane, which means that they lie in the same plane in 3D space. In this case, the parameters of the dashed line cannot be recovered even if all other planes are determined. Therefore, it is necessary to remove these curves before trying to solve the linear system.

The described problem does not only occur with single laser curves but also with groups of curves. For example, consider the solid lines in Fig. 5.4 as one group and the dashed and dotted lines as a second group. These two groups are only tied together by the intersection points four, five and six, which are collinear. In this case, the group of dashed and dotted planes has indeterminacies even if all solid planes are determined.

The second step is finding all plane parameters up to scale by minimizing a non-linear optimization problem. With the cross-line laser configuration, an additional orthogonality constraints is obtained between each of the two cross laser planes which is used to recover the offset vector. The offset is computed, such that the error of the orthogonality constraints is minimized. The offset vector $\hat{\mathbf{o}}$ that minimizes the sum of the inner product between planes in the set $C = \{(i,j)|(\pi_i \perp \pi_j)\}$ of orthogonal laser planes is found by:

$$\hat{\mathbf{o}} = \arg\min_{\mathbf{o}} \sum_{(i,j)\in C} \mathbf{n}(a_i, b_i, c_i, \mathbf{o})^T \mathbf{n}(a_j, b_j, c_j, \mathbf{o}) \ , \tag{5.7}$$

where $\mathbf{n}$ is the normal of the plane computed from the plane parameters and offset vector. The scale cannot be recovered with only two laser planes and needs to be estimated from other measurements, such as a known distance in the scene.

Only a subset of the laser planes is used to solve for the plane parameters. The other planes are then reconstructed by fitting a plane to the intersection points with planes of the already solved subset of laser planes. Although it is possible to perform 3D reconstruction with fewer planes, empirically it is found that 100 - 200 planes are necessary to compute a robust solution.

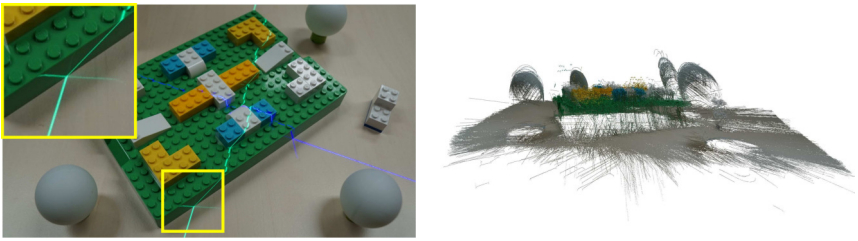


Fig. 5.5: Reflections of the laser line due to glossy surfaces and resulting distorted point cloud reconstruction.

5.2.4 Computing Robust Laser Plane Estimates

In order to choose a solvable subset of planes, all planes that have only collinear intersection points are removed, which is tested using principal component analysis (PCA). Moreover, a heuristics is applied to select planes that have distinct orientations and positions in the image. To do this, planes are picked in a way that they are spread apart in time. Consecutive image frames are very similar if the projector is moved slowly. Additionally, planes that have more than one intersection with each other are rejected. There are situations where it is valid that two planes have multiple intersections. In practice, this mostly happens for almost identical planes or due to erroneous or noisy intersection detections.

Problematic for the self-calibration are incorrect intersection constraints. This problem occurs quite often when scanning glossy surfaces. Depending on the incidence angle of the laser light, reflections are visible as depicted in the left image in Fig. 5.5. It is not always possible to reduce this effect by discarding detected laser lines with low brightness because this also discards lines on darker surfaces. This means that additional line segments are detected that form erroneous intersection with other laser curves. This significantly distorts the 3D reconstruction, as shown in the right image in Fig. 5.5.

This problem is addressed by explicitly detecting these outliers among the intersections and labeling them as invalid. To do this the intersection point in 3D space is computed using the plane parameters of both of the intersecting planes. Since the plane parameters are noisy, there is an error between the two computed point positions. For each laser curve, all intersections are labeled as invalid based on a threshold if they have a higher error than the median error of all intersections of that particular curve. Then, the self-calibration is recomputed without taking these invalid intersection constraints into account.

Moreover, the line points associated with the reflection are not reconstructed correctly because they do not lie in the original laser plane. These reflections are typically detected as distinct line segments by the laser line extraction algorithm. Removing line segments with only invalid intersection points is an effective technique for removing these erroneous points from the final point cloud result. This is visualized in Fig. 5.6. The top left image shows the input image, and the top right image shows the detected laser curves. In the bottom left image, all intersections with other laser curves are highlighted in green. The bottom right image shows the intersections

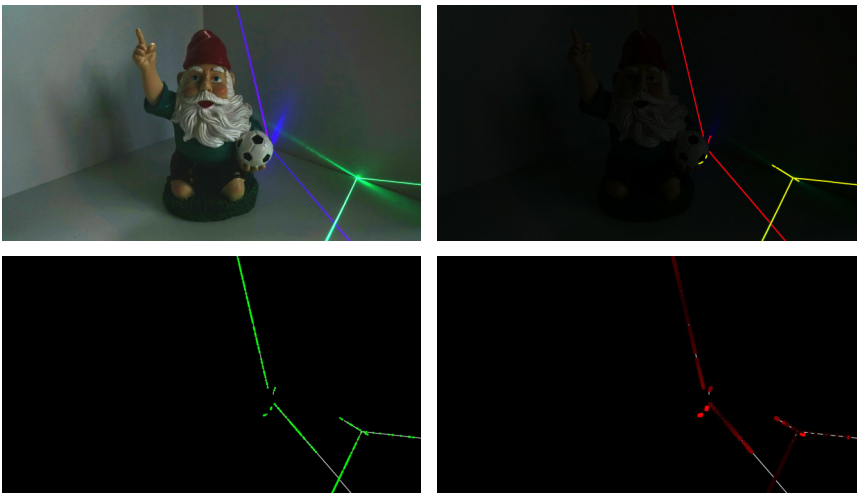


Fig. 5.6: Filtering of line segments based on reconstruction error. Top left: Input image. Top right: Detected laser curves. Bottom left: Intersections with other laser planes depicted in green. Bottom right: Intersections colored by reconstruction error. Bright red means high error.

with high errors in bright red. In this plot, the intersections are colored by the reconstruction error from black (low error) to red (high error).

It is generally difficult to verify that a valid 3D reconstruction is found. It cannot be discerned if the solution is good or bad by only looking at the residuals of Eq. 5.5 and Eq. 5.7 since the approach optimizes for these values, and they are expected to be small. Therefore, the errors of the planes that are not take into account in the plane parameter optimization step are considered. Specifically, the root-mean-square angular error for all orthogonal laser planes is computed.

A disadvantage of the presented method compared to other structured light approaches, e.g., gray code projector based systems, is that a high number of images is necessary since only two lines are reconstructed from a single image. However, in order to achieve real-time reconstruction, it is only feasible to compute self-calibration for a subset of the detected laser lines in the images. Moreover, not all laser lines are estimated directly using the proposed method, such as curves with collinear intersections. All other planes are only determined by the intersection points with the subset of solved laser planes.

5.3 Experiments

The proposed approach is demonstrated on different data sets acquired in air or water. For the experiments in air, a consumer camera Sony a6000 with an APS-C sized sensor was used

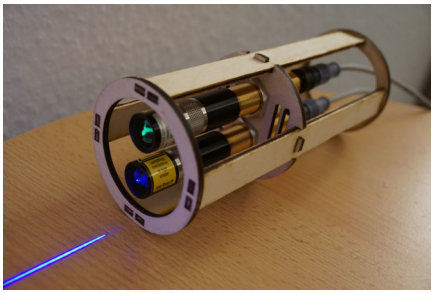


Fig. 5.7: Cross line laser projector in wooden frame employed for the experiments.

in video mode. The scenes were captured with a wide-angle lens with a focal length of 16 mm. The underwater scans are captured using a GoPro action camera. Video at 30 fps in Full-HD resolution (1920 × 1080 pixels) is recorded. A high shutter speed is beneficial since the laser is moved by hand. An exposure time in the range of 1 ms to 10 ms was used in the experiments in order to reduce motion blur.

The cross-line projector is built from a 450 nm and 520 nm line lasers with a fan angle of 90° and an adjustable output power of up to 40 mW. Note that the fan angle is reduced in water to approximately 64°. The lasers were mounted in a custom-built frame constructed from laser-cut wood as depicted in Fig. 5.7. It is also possible to use DOE to project a cross with a single laser. However, line lasers typically emit a significantly thinner line, which improves accuracy, and using two lasers with different colors simplifies the separation of the two laser curves in the image. The laser focus is set such that the line is as thin as possible over the whole depth range of the scene.

5.3.1 Comparison with Explicit Plane Parameter Estimation

In order to verify the plane parameter estimation, the proposed approach is compared with an explicit online calibration of the laser plane parameters using known reference planes in the scene. The setup is inspired by the David Laser Scanner [343], which is depicted in Fig. 5.8. This approach exploits a known background geometry. Two reference planes with a known angle are placed behind the scan object. These background planes are calibrated with respect to the camera. Then, the object is scanned with a handheld line laser.

If the image pixels that belong to projections of the laser line on the background are detected, these points are reconstructed by computing the intersections of the camera rays with the planes. These points are extracted from the image using the assumption that they belong to linear segments of the extracted laser line. These segments are found either by application of the RANSAC algorithm or a Hough line transform. If points on both planes are found, the laser plane is computed by fitting a plane. The scan object is finally reconstructed using the intersection of the camera projection rays and the online estimated laser plane.

Similar to this approach in this work, three planes are placed around the object as shown

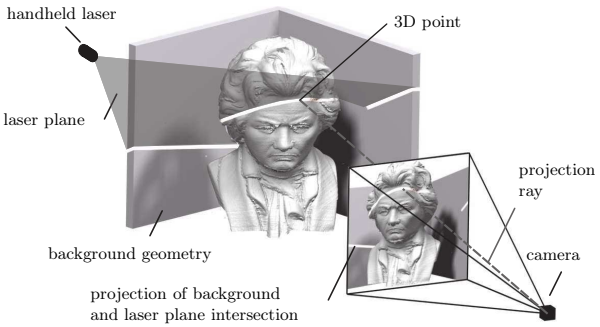


Fig. 5.8: Laser triangulation of the David Laser Scanner. Image adopted from [343].

in the top image of Fig. 5.9. The third plane is necessary, such that vertical laser lines are determined. The plane equations of the three reference planes are estimated from multiple images of chessboard patterns placed on the plane surfaces. The laser points that lie on the reference plane are found using manually created mask images. If the laser line is visible on two of the calibrated planes with a sufficient amount of points, the plane parameters are extracted by fitting a plane to the laser line points.

The same extracted line points and camera parameters are used as an input for both methods. Moreover, the reference solution using online calibration is also degraded by any inaccuracies of the performed camera calibration. Therefore, both methods are affected by the same input data errors. The only difference is the plane parameter estimation. For comparison purposes, the following two error metrics are employed: Firstly, the angular error between the estimated plane normal vector $\mathbf{n}_{\text{est}}$ using the proposed self-calibration technique and the reference plane normal vector $\mathbf{n}_{\text{ref}}$ estimated using the known reference planes is computed by

$$e_{\text{angular}} = \arccos(\mathbf{n}_{\text{est}}^T \mathbf{n}_{\text{ref}}) \ . \tag{5.8}$$

Secondly, the error of the distances of the planes from the origin is computed:

$$e_{\text{dist}} = |d_{\text{est}} - d_{\text{ref}}| \ , \tag{5.9}$$

where d_{est} is the distance of the estimated plane from the origin and d_{ref} is the distance of the reference plane from the origin. In this experiment, the origin was chosen as the projection center of the camera.

The RMSE is reported for a total of 890 extracted laser curves. The RMS of the angular error between the plane normals is 2.70° (mean error 0.34°) and the RMSE of the distances of the planes from the origin is 8.29 mm (mean error 1.59 mm).

The reconstruction result of the proposed method, see bottom left images of Fig. 5.9, compares very well to the result using the explicit online calibration result, see bottom right images of Fig. 5.9. However, the scans created by the proposed method are visibly more noisy.

Experimental setup.

Proposed method. Reference.

Fig. 5.9: Setup for comparison with explicit plane parameter estimation. Top: Object with three known reference planes in the background. Bottom left: Top and detail view of the 3D reconstruction using the proposed method. Bottom right: Top and detail view of the 3D reconstruction based on the calibrated reference planes.

5.3.2 Measurement Results

Fig. 5.10 shows examples of the achievable results. On the left, a subset of the extracted laser lines is depicted. On the right, the final reconstructed point cloud is shown with RGB color mapped to the points. The top scene showing the table tennis balls and Lego bricks was reconstructed from 4 minutes of video with 9,065 valid laser planes detected. The final point cloud created from all valid laser curves has a size of 5,522,983 points. The bottom scene showing the hand was reconstructed from 3.5 minutes of video with 7,817 valid laser planes detected. The final point cloud created from all valid laser curves has a size of 11,613,200 points. For both examples, the self-calibration was computed using a subset of 400 laser curves.

An example of an underwater scan created using the presented method is depicted in Fig. 5.11. The left image shows the pipe structure scene in a water tank. The right image shows the 3D reconstruction result. In this case, the point cloud is colored using the automatic white balance

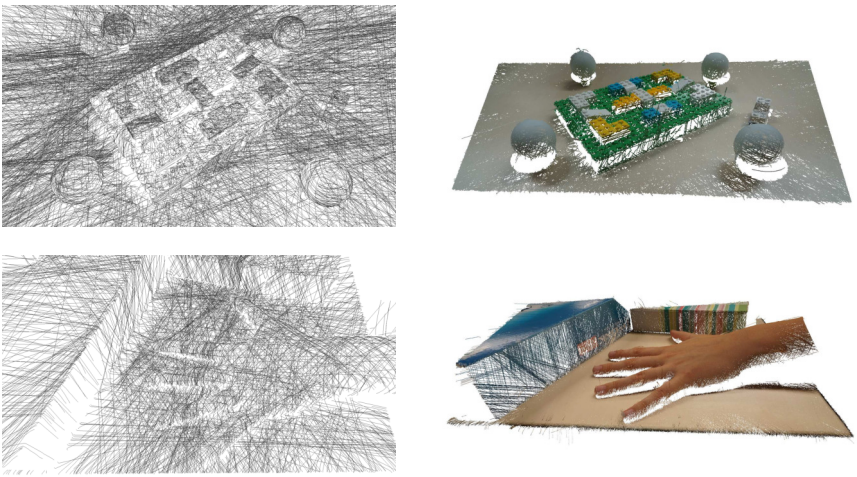


Fig. 5.10: Examples of the reconstruction result. Left: Subset of the extracted laser lines in the image. Right: Colored reconstructed point cloud.

of the GoPro camera. Therefore, in the underwater scan the color reproduction is not accurate. In this experiment, the projector was approximately 0.5 m to 1 m away from the camera.

The method applies with some restrictions also to larger outdoor scenes. Scans of a chapel and a comparison with LIDAR data created from multiple scans using a Riegl VZ-400 terrestrial laser scanner are depicted in Fig. 5.12. The scan in the top left image of Fig. 5.12 was captured at night to achieve high contrast of the laser lines. Flash photography was used to capture the color information. Measured surfaces in the scene are up to 20 m away from the camera. Hence, the cross-line laser projector is moved further away from the camera to scan with baselines in the range of 2 m to 3 m. In the top right image of Fig. 5.12, two scans created using the presented approach are visualized in pink and turquoise. The data was registered with the LIDAR reference data, colored in yellow, using the ICP algorithm. In this case, scale definition of the self-calibrating approach was achieved by registering the scan with scale adjustment.

The bottom left image of Fig. 5.12 shows the point cloud created using the proposed method colored with the point-to-point distance from the point cloud acquired using the terrestrial laser scanner. The reference data was acquired multiple months before scanning the chapel with the cross-line laser system. Therefore, note that the points of the tree in the right part of the image do not match the LIDAR data because the vegetation has changed in the time that passed between the scans. The error histogram is depicted in the bottom right image of Fig. 5.12. 78 % of the points in the two scans created using the proposed method differ less than 10 cm from the scans captured using the Riegl VZ-400.

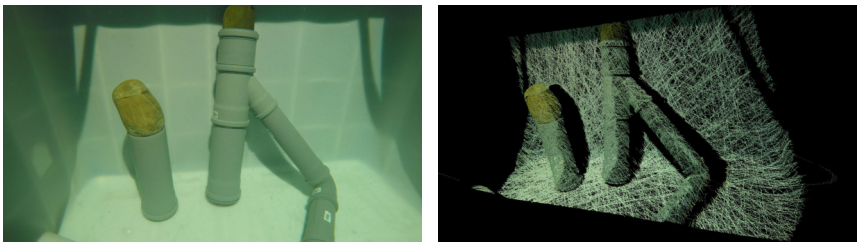


Fig. 5.11: Example of the proposed method on underwater data. Left: Test scene with pipe structures. Right: 3D point cloud created using the proposed method.

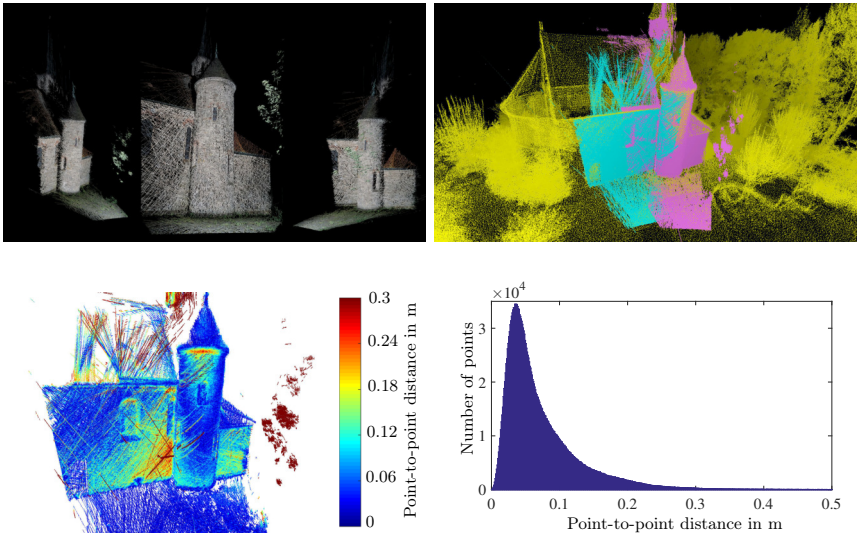


Fig. 5.12: Scans of a chapel and comparison with LIDAR data captured using a Riegl VZ-400. Top left: Three views of a scan created using the proposed method. Top right: LIDAR data in yellow and two scans using the proposed method in pink and turquoise. Bottom left: Visualization of the result using the proposed method colored with the distance from the reference point cloud. Bottom right: Histogram of the point-to-point errors.

5.4 Discussion

This chapter demonstrates how self-calibration techniques for line laser scanning, first introduced in [136, 137] are extended to apply to underwater imaging. The proposed approach is low-cost because the system only requires a single consumer video camera and two line laser projectors.

The detection accuracy of the intersection points strongly influences the self-calibration result. The solution is improved by explicitly detecting and removing outliers. This work provides the implementation details to recover robust 3D estimates. Moreover, it shows how the geometric constraints employed in the self-calibration process are used to remove noisy or erroneous detections due to reflections. In experiments, it was demonstrated that good quality scans are achieved, and a similar performance to 3D line laser scanning using online calibration based on known reference planes in the scene is possible.

After an initial solution on a subset of planes is found, an approach to improve the geometric accuracy and consistency is to optimize all plane parameters based on the initial solution. However, this introduces additional noise due to varying localization accuracy of line segments and intersection points, which requires appropriate weighting of the individual observations. The choice of the optimal weighting function is unclear, requiring further investigation.

Choosing good parameters for the reconstruction step is challenging because the scale of the scene is unknown. Scaling the parameters, e.g., by the depth range of the reconstructed point cloud, is not always possible. Therefore, automatically determining good parameters remains to be investigated in future work.

Chapter 6

Underwater Mobile Mapping

Optical underwater scanners feature a limited measurement volume. Absorption in water and turbidity, usually limit the measuring range to a few meters. Therefore, to capture large objects or areas in the water, we need to move the scanner, e.g., by mounting the sensor to a ROV or surface vessel. Mobile mapping is the process of collecting spatial data from a mobile platform. For inspection tasks, scanning from a mobile vehicle is interesting to analyze elongated objects, such as pipelines or cables. However, fast movements cause distortion of the point clouds acquired using multi-shot structured light techniques. Hence, even for accurate single scans, motion compensation is necessary.

Mobile mapping with underwater laser scanners is typically achieved by using the navigation data of the vehicle. For example, GNSS data of a ship [193] or a combination of acoustic underwater positioning system information, DVL and inertial navigation [213] is used to measure the vehicle trajectory. Furthermore, SfM algorithms are applied to create a trajectory estimate of the scanning system [81,112]. With more recently developed underwater laser scanners with a high update rate and larger field of view, it is possible to incrementally register the point clouds using the ICP algorithm. For example, Palomer et al. use a laser scanner for mapping from an AUV. They find an initial coarse alignment using point features and refine the solution using ICP [255]. In [233, 234] an underwater laser scanning system using Visual SLAM is described.

This chapter describes the application of the underwater laser scanning system described in Chapter 4 for mobile 3D data acquisition from a floating platform. A GNSS based localization system is employed for trajectory estimation. The system consists of an under- and an above-the-water part and enables to scan the ground surface in shallow waters from a floating platform. Above water, two GNSS antennas with a baseline of 1 meter are used to calculate the Real-Time Kinematic (RTK)-position and the heading angle. The full 6-DoF position is estimated by fusing the GNSS navigation data with IMU measurements. The 3D data is captured in water using a structured light scanner consisting of a low-light underwater camera and a green cross-line laser projector.

Moreover, this chapter presents the development of an optical underwater scanning system for mobile mapping. The system realizes motion estimation and the fusion of the individual scans into one globally consistent 3D model. This is achieved through motion compensation based on visual odometry and scan matching. This work was carried out in the UWSensor project with

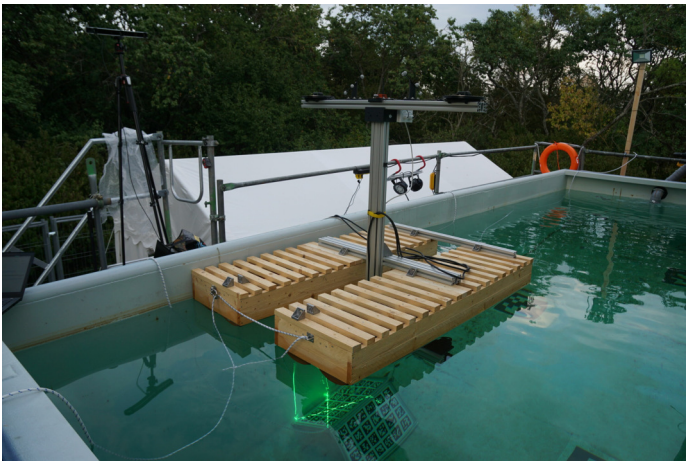


Fig. 6.1: Underwater scanning system mounted to a floating platform for evaluation in a water tank.

the goal to apply the system for offshore inspection tasks.

The application of the proposed underwater laser scanning for mobile mapping described in Section 6.1 was previously published in [10–13]. The contributions of the author include the development and implementation of the approaches for trajectory estimation, calibration and mapping and the experiments. Moreover, the contributions are the hardware and software design and the construction of the overall underwater laser scanning system.

The development of the novel optical underwater 3D scanner for mobile mapping described in Section 6.2 is joint work with the UWSensor project partners. The work was previously published in [6, 7, 15, 16]. The contributions of the author encompass the camera-IMU calibration, VIO for real-time motion compensation, post-processing and registration of the underwater scans and experiments. On the hardware side, the contributions are the electronics and mechanical design of the triggered high-power LED lights, IMU unit and calibration fixtures.

6.1 Mobile Underwater Laser Scanning

This section presents first results on applying the underwater structured light laser scanner described in Chapter 4 for mobile scanning in the water. The scanning system mounted to a floating platform, which is depicted in Fig. 6.1. For the experiments, the platform is moved manually to acquire 3D data in a test tank.

The system consists of an above-the-water and an underwater component. Above-the-water a GNSS RTK positioning system is used to derive position data of the scanning system. Below the water surface, a custom-built underwater laser scanner is used for 3D data acquisition. The

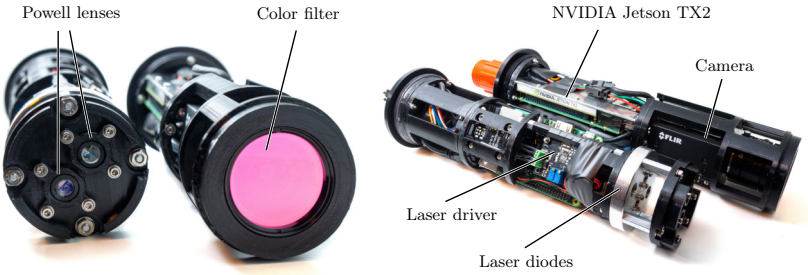


Fig. 6.2: Camera and laser projector assembly of the underwater scanner for mobile mapping.

structured light scanner is placed in the water pointing downwards. The 6-DoF trajectory is estimated by fusing satellite navigation data with a MEMS-based IMU. Both components are mounted rigidly to the same aluminum profile in order to keep the relative orientation between the trajectory estimation system and the scanner constant during scanning. Since the system uses GNSS for trajectory estimation, it is only applicable for mapping in shallow water up to approximately 5 m water depth depending on the turbidity.

6.1.1 Underwater Laser Scanning System

The data presented in this work was captured with a self-built structured light underwater laser scanning system. For the mobile mapping application, a second system, seen in Fig. 6.2, was set up with modifications compared to the underwater scanner described in Section 4.1. The color camera was replaced by a monochrome camera with a frame rate of 163 Hz and otherwise similar specifications. The industrial camera is a FLIR Grasshopper3 GS3-U3-23S6M-C with a Sony IMX174 sensor with a resolution of 1920 × 1200 pixels and a sensor format of 1/1.2".

When scanning in shallow water, the ambient light of the sun is still strong and interferes with the line pattern projection. Moreover, the high light absorption especially in turbid water conditions further degrades the image. Therefore, a green bandpass filter, seen in the left part of Fig. 6.2, was installed in front of the camera to block some of the ambient illumination. The laser projection needs to be bright enough, such that a sufficient contrast of the green laser line against the background is achieved. The employed high-power lasers and a high-sensitivity camera with a large dynamic range mitigate these problems to some degree.

Fig. 6.3 shows different scan patterns: single laser line, two parallel lines and cross-line configuration. Compared to a single laser line configuration, scanning using multiple laser lines has the advantage that an internal overlap between the created point clouds of the individual line projection is achieved. A single line usually only has overlap between the point clouds of individual laps of the vehicle that is carrying the scanner. Additionally, with a cross-line configuration, a large swath is also scanned for across-track movement. This enables scanning

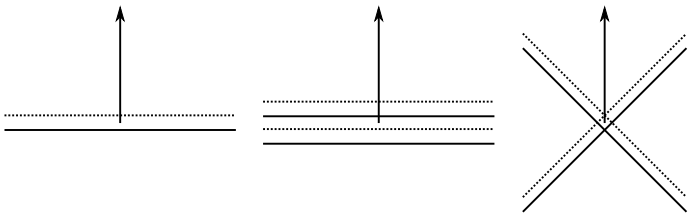


Fig. 6.3: Overlap of scan patterns for different laser line configurations. Left: single line, middle: two parallel lines, right: cross line. The arrow visualizes the direction of movement.

along a trajectory that is perpendicular to the direction of movement visualized by the arrow in Fig. 6.3. Therefore, a cross-line laser pattern is selected for the proposed system because this enables scanning with less restricted movement. Only the distance between the scanner and the object has to be kept in a specific range because of field-of-view and focus restrictions.

The two laser planes are projected at an angle of 45 deg with respect to the vertical camera axis. This way, both projected laser lines have similar baselines. However, the resulting baseline is reduced compared to the distance between the camera and laser projection centers.

6.1.2 Satellite and Inertial Navigation System

The employed satellite navigation system is based on the cost-efficient u-blox ZED-F9P receivers with multiband GNSS patch antennas. The receivers are able to concurrently receive GPS, GLONASS, Galileo and BeiDou satellite navigation data. Two antennas are employed to measure the heading of the vehicle. In addition, an Xsens MTi-300 MEMS-based Inertial Navigation System (INS) is integrated for measuring the system's orientation. The INS has a measurement rate of 100 Hz.

The GNSS antennas and IMU are mounted on the same aluminum profile, such that the relative orientations are rigid and do not change during the transport of the system. The hardware setup is shown in Fig. 6.4. Additionally, a radio transceiver is mounted on the aluminum bar for receiving real-time correction data from the GNSS base station. This comprises the surface component of the scanning system.

A RTK base station is set up, which sends correction data to the GNSS receivers via a long-range radio link. This way, real-time RTK positioning and heading data is computed.

For real-time processing, the integrated RTK solution of the u-blox receivers is employed, which provides update rates of up to 10 Hz. Post-processing is done with RTKLIB [313], which computes the positioning solution with the full 20 Hz update rate of the receivers. With this setup, centimeter precise positioning is achieved depending on the circumstances of the data acquisition, such as occlusion of satellite or multi-path effects. For trajectory estimation, the position and heading data from satellite navigation is fused with the filtered orientation computed based on the INS data to create a full 6-DoF trajectory.

For reference measurements, gray ball markers are attached to the system and the 6-DoF trajectory is measured using an optical tracking system. Fig. 6.5 shows a trajectory measured

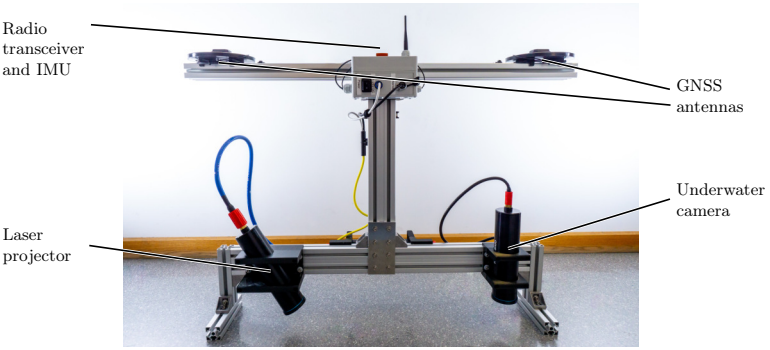


Fig. 6.4: Satellite navigation system with two multiband GNSS antennas and IMU. The correction data from the RTK base station is received via a long-range radio transceiver.

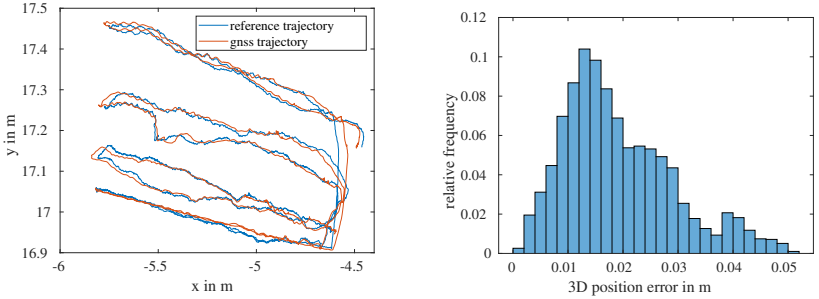


Fig. 6.5: Validation of the estimated GNSS trajectory by comparison with a reference trajectory acquired using an optical tracking system. Left: Plot of the trajectories. Right: Histogram of the 3D position error between the two trajectories.

using the proposed GNSS RTK system compared to a measurement using an optical tracking system as well as the error histogram. The depicted trajectory has a length of 14 m. The RMSE of the 3D position error between the RTK trajectory and the reference trajectory is 2.2 cm.

An important point is the co-calibration between the scanner and the coordinate frame of the GNSS trajectory. First, an initial estimate is found by manual measurements of the antenna positions and the CAD model. Then, a calibration structure is placed in the water, and the sensor system is moved linearly and rotated above the structure. Based on the calibration structure, a camera trajectory is estimated. This trajectory is recorded synchronously with the estimated GNSS-IMU trajectory. The extrinsic parameters are optimized based on an alignment of the camera and satellite navigation trajectory using the approach described in Section 4.4.2.

Fig. 6.6: Mobile data acquisition experiments conducted in the towing tank of the chair of fluid dynamics at the University of Rostock. Left: Towing tank at the chair of fluid dynamics in Rostock, Right: Underwater laser scanner deployed in the tank.

6.1.3 Experiments in a Towing Tank

Initial tests for mobile data acquisition with the underwater laser scanner were carried out in the towing tank at the chair of fluid dynamics at the University of Rostock. The towing carriage and tank is depicted in the left image of Fig. 6.6. The water tank is 5 m wide and provides a depth of up to 3 m. The scanner was deployed at about 2.5 m water depth using a vertical bar, which is shown in the right image of Fig. 6.6.

Mobile scans are acquired by moving the scanner in the water using the towing carriage. For this experiment, the scanner is moved with constant velocity at a slow speed of $0.1\,\frac{\mathrm{m}}{\mathrm{s}}$ through the water. It is assumed that the trajectory is approximately linear. However, the angular parameters of the relative pose of the scanner with respect to the towing carriage need to be estimated from the data.

The result for a trajectory with a length of 8 m is shown in Fig. 6.7. The red line visualizes the trajectory of the scanner. The top images show two renderings of the point cloud colored by intensities, while the bottom images are colored by height.

6.1.4 Mobile Mapping Results

For testing, the scanning system is deployed in a water test tank built from a container with $40\,\mathrm{m}^3$ of water. For mobile scanning experiments, the scanner was mounted to a raft, which is seen in the left image of Fig. 6.8. The floating platform was manually moved above the scan objects. A result of the point cloud created by projecting the individual scans into 3D space along the estimated trajectory from GNSS-INS is shown in the right image of Fig. 6.8. The scan was created of the scene depicted in the middle image of Fig. 6.8.

The conditions in the test tank are not ideal since the lid of the container occludes part of the sky, which is visible in the left image of Fig. 6.8. This means that the measurement results of the GNSS antenna close to the lid are degraded. Nevertheless, 3D scans were captured successfully.

In the tank a test object with a size of 2 m × 2 m × 1 m is placed. The test artifact is constructed from aluminum profiles and polypropylene pipes. A picture of this artifact is depicted

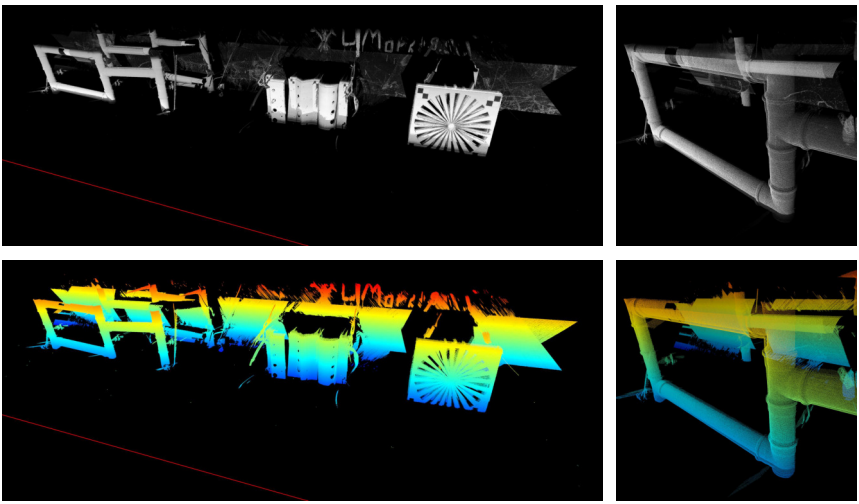


Fig. 6.7: Underwater scan created by moving the scanner in the water along a linear trajectory (red line). Top: Point cloud and detail view colored with intensities. Bottom: Point clouds colored by height.

in left image of Fig. 6.9. An example of the captured point cloud from the floating platform of the pipe object is shown in the right image of Fig. 6.9. While the structure of the scanned objects is visible, the created point clouds still exhibit coarse errors in the centimeter range. This requires future work to improve the estimated scanner trajectory as well as optimizing the misalignment errors of the system.

Fig. 6.10 shows the point cloud of a calibration fixture scanned in the water tank. The left image depicts the result using satellite navigation. The right image shows the point cloud using pose estimation of the scanner using the optical markers on the structure. While the GNSS results show coarse errors in the centimeter range, the point cloud based on marker tracking is

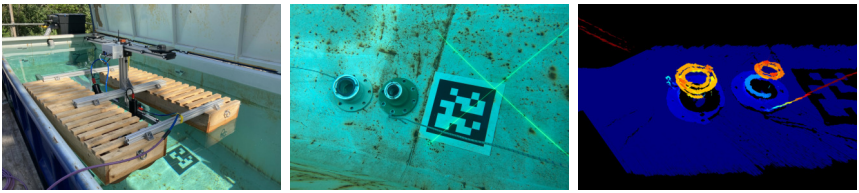


Fig. 6.8: Experimental setup and mobile scan result. Left: Underwater scanning system mounted on a floating platform. Middle: Picture of the scene. Right: Resulting point cloud colored by height.

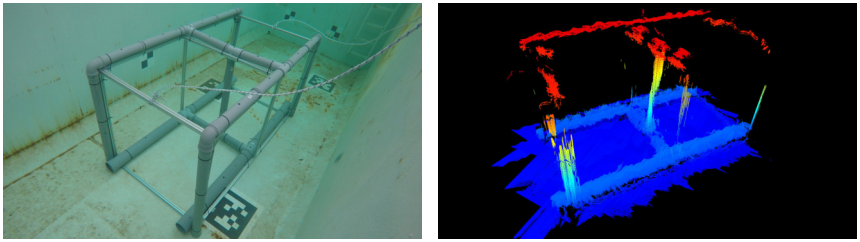


Fig. 6.9: Example of the captured point cloud from the floating platform using the GNSS-INS trajectory. Left: Pipe artifact in the water tank. Right: Point cloud of the pipe structure colored by height.

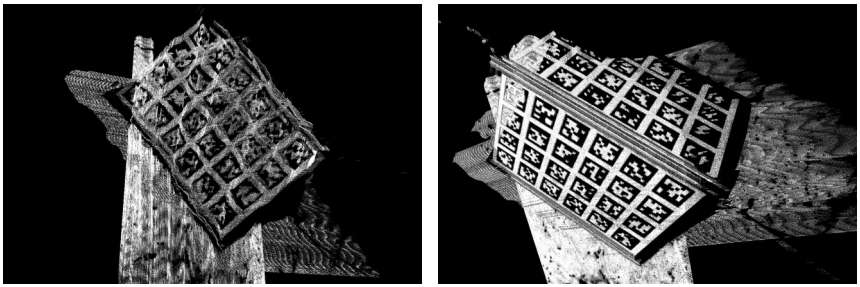


Fig. 6.10: Comparison of satellite navigation trajectory and pose estimation based on visual markers. Left: Point cloud of a calibration target colored by intensity based on GNSS-INS trajectory. Right: Point cloud colored by intensity based on pose estimation using the markers on the target object.

significantly more accurate. This suggests that trajectory estimation based on visual odometry improves the results, which is the direction pursued in the next section.

6.2 Mobile Mapping with an Optical Underwater Sensor System

In the UWSensor project, a novel optical 3D underwater sensor based on fringe projection was developed [15]. A stereo camera and a GOBO projector for 3D measurement, a color camera and an IMU are integrated into the underwater housings of the sensor. The sensor system enables dense 3D acquisition in a measurement volume of approximately 1 m × 1 m × 1 m, as well as simultaneous acquisition of color data. The pattern projection is adapted to the underwater context using high power blue LEDs. The advantages of the sensor system are a very high measurement accuracy and a high scan rate. Special attention is paid to the calibration procedures for the underwater stereo scanner, which takes the refraction at the media transitions into account and applies additional refinement procedures. This leads to very low systematic errors. The dense surface acquisition with simultaneous high update rate enables continuous

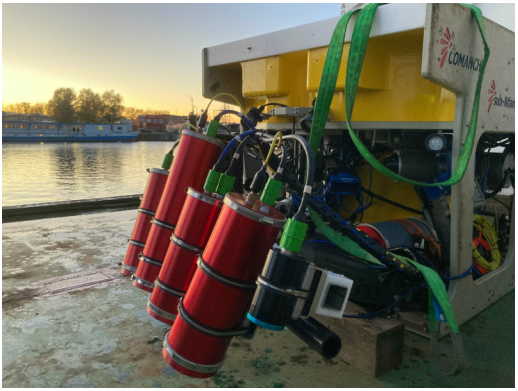


Fig. 6.11: UWSensor mounted in the manipulator of a light Work-Class ROV.

registration of individual scans. This enables the use for mobile mapping and 3D acquisition with an underwater vehicle. Fig. 6.11 shows the sensor mounted in the manipulator of a light Work-Class ROV during preparations of offshore field tests.

The system realizes a motion estimation as well as a fusion of the single scans into a globally consistent 3D model. Here, visual odometry is initially used in combination with a fiber optic gyroscope to estimate the sensor's motion. The resulting low-latency derived pose estimates are provided to the 3D reconstruction process to realize an online motion compensation of the 3D scans. A multi-stage registration strategy is then applied to further refine the trajectory estimation and create consistent 3D maps.

6.2.1 Underwater 3D Sensor

The underwater sensor uses a multi-image 3D reconstruction method based on active stereo with pattern projection. A GOBO projector with a high power LED is used as the projector's light source. The sensor is shown in Fig. 6.12 during calibration in a water tank. In the outer two housings, the stereo cameras for 3D measurement are integrated. In the middle, there are two housings: one for the GOBO-projector and the control unit of the sensor. A second housing contains an additional color camera and an IMU with a fiber optic gyroscope. In addition, two LED strobes for illuminating the color images are also mounted on the sensor system.

The control unit triggers cameras, projector and LED strobe lights in such a way that a measurement series of stereo images with pattern projection and color images with LED flash illumination are recorded in a configurable, alternating trigger scheme. The sensor system delivers 3D point clouds with a scan rate of up to 60 Hz. A single scan contains approximately one million points. At a measuring distance of 2 m, this results in a measuring volume of approximately 1 m $\times$ 1 m $\times$ 1 m. Synchronously to the point clouds, color images with 7 megapixels are recorded with a frame rate of 25 Hz. The color camera has a larger field of view than the stereo cameras.

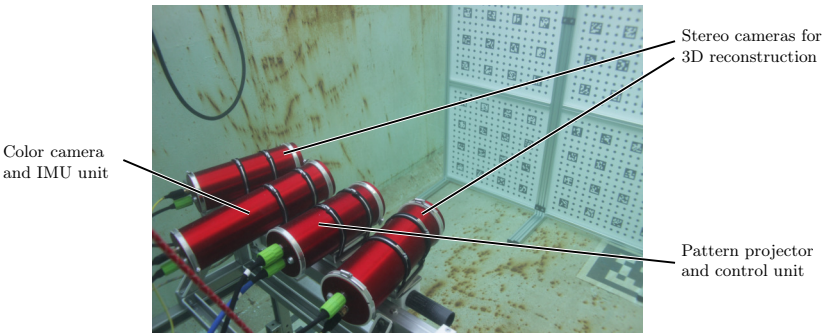


Fig. 6.12: Underwater 3D sensor system.

At a measuring distance of 2 m an area of approximately 3 m × 3 m is recorded. The IMU sensor data stream is acquired and recorded in parallel.

6.2.2 Mobile Data Acquisition

Due to the high scan rate and the dense surface reconstruction, the underwater 3D sensor is very well suited for mobile mapping. In contrast to line laser scanners, which only record individual 2D profiles, it is possible to register individual scans with each other directly. However, the limited measurement field poses the challenge that sometimes there is insufficient structure for precise registration of single scans. Therefore, a multi-stage continuous-time registration is performed.

The 3D reconstruction by active pattern projection [157] that is employed uses temporal correlation over a sequence of gray-scale images to accurately determine the correspondences between the left and right measurement cameras. This multi-shot structured light technique is typically computed on a set of 10 - 15 images. The mathematical model assumes that each pixel of this gray-scale image stack maps to the same object point. This static measurement configuration is violated in the planned mobile scanning approach using a ROV. The turbidity and light absorption in water require longer integration times for the cameras. Hence, high-speed projection and acquisition to allow for quasi-static scans is not applicable in challenging underwater conditions. This causes, if not compensated, measurement artifacts at higher measurement speeds. Fig. 6.13 shows an underwater 3D scan acquired at a speed of ca. $0.7\,\frac{\mathrm{m}}{\mathrm{s}}$ if the motion is not considered. In this example, the spheres of the test target are not fully reconstructed. Therefore, a low-latency sensor pose estimation is necessary to compensate for the motion within a measurement sequence.

6.2.3 Calibration

The stereo calibration of the sensor is performed using a circle pattern [65]. Here, additional ArUco markers are used to give each circle mark a unique ID. The calibration plates are built

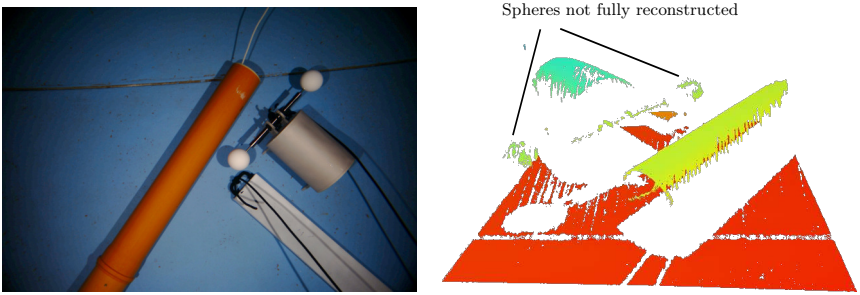


Fig. 6.13: 3D point cloud with scan artifacts due to fast sensor motion. The spheres are not fully reconstructed. Left: Color image acquired by the sensor. Right: Reconstructed 3D point cloud without motion compensation colored by height.

from direct printing on float glass and installed in aluminum frames as a mechanical protection. For calibration, BINGO is used as the bundle block adjustment software [188].

The setup for calibration is shown in Fig. 6.14. The calibration of the alignment between the IMU and the color camera is performed in air. The calibration of the cameras and stereo calibration is performed in water. The external orientation of the color camera relative to the measurement cameras is determined in the stereo calibration process.

The IMU unit is installed in an underwater housing together with the color camera. Hence, it is assumed that the color camera and IMU are rigidly mounted with respect to each other. In order to relate the IMU measurements to the observations of the cameras and the 3D measurements, the extrinsic orientation of the sensors must be determined. To do this, the rotation and translation between the color camera and IMU is calibrated. The translation and rotation between the color camera and IMU is determined via an optimization between IMU measurements and the camera trajectory. For this purpose, the intrinsically calibrated color camera is moved in front of a known static structure with optical marks. This allows the trajectory of the color camera to be determined with all six degrees of freedom.

From the trajectory, the expected IMU measurements are determined. The linear displacement and rotation of the IMU relative to the color camera are then determined, such that the actual measurements deviate as little as possible from the angular velocities and accelerations predicted from the camera trajectory. The method used is based on the work presented in [135] and was adapted accordingly to the system parameters and camera model of the underwater sensor.

6.2.4 Visual Odometry

For the visual approach, sufficient visibility in the water and adequate illumination are required. Since there is little ambient light from the sun at depth, LED strobes are employed. In addition, the sensor must be in an environment that contains sufficient optical features for tracking. For computing visual odometry, feature detection is applied to analyze the camera images. Features

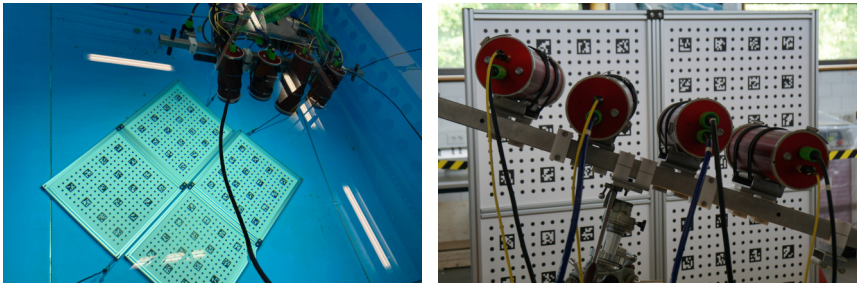


Fig. 6.14: Calibration of the sensor system in air and in water.

are local, distinct and recognizable points in the image. The visual approach here is based on [263, 264] and uses the Harris feature detector (also known as the Shi-Tomasi method) to extract the features from the image automatically. Fig. 6.15 depicts an example of the successfully tracked feature points overlayed in red on the image stream of the sensor.

At the beginning of the acquisition process, the pose estimation must be initialized. This means that the method cannot be used directly from a standstill. The sensor must first be moved to create enough parallax to allow a SfM reconstruction of the feature points without scale. This result is then scaled by an adjustment using the integrated IMU data. Once this initialization process is complete, the actual measurement acquisition begins.

The time offset between the camera and the IMU data stream is determined online. The relative motion is then determined by adjustment of the visual and IMU measurements in a sliding window approach. Optimization is performed locally using a continuous group of keyframes. Keyframes are selected based on a threshold on the mean parallax of each feature point. Note that no relocalization or loop closure based on visual features is used in this step. The goal is to obtain a continuous trajectory without jumps. Visual odometry is primarily used for local motion compensation. Any drift due to residual error is later compensated in the 3D registration process.

6.2.5 Registration of the 3D Data

The registration pipeline is depicted in Fig. 6.16. It is based on the processing pipeline developed in [192] and the algorithms described in [116] and [54].

Using the initial estimation of the movement trajectory by visual-inertial odometry, a registration strategy with multiple steps is used for further refinement of the trajectory. In preparation of the registration, the 3D data is filtered to reduce the size of the 3D point cloud and achieve an equalized spatial distribution of the scan points. In this step, a range filter and an Octree based sub-sampling are applied. In the experiments, a voxel size of 1 cm is used for the point reduction. Then, every 3D point cloud is registered sequentially against its predecessors using the ICP algorithm with the goal of local optimization of the trajectory and improvement of the resulting 3D map.

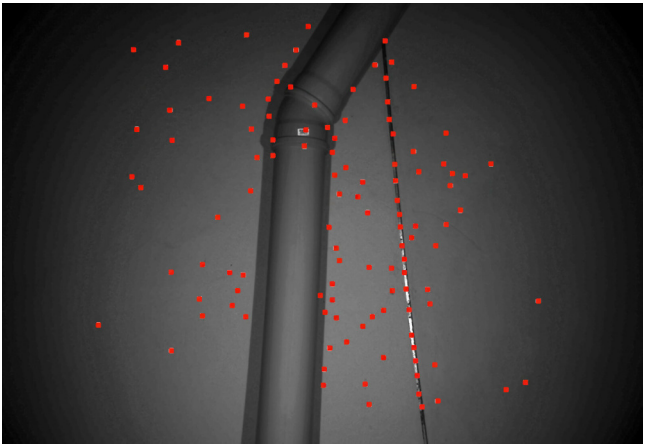


Fig. 6.15: Example of tracked feature points while scanning a pipe in the water.

For this, a metascan created from a sliding window of registered preceding scans is used to provide more structure during ICP registration. A metscan size that corresponds to the number of scans captured in the previous 3 s is selected. The search radius is set quite small (only a few centimeters) depending on the velocity of the sensor to reduce the risk of rough registration errors because of the limited field of vision and the expected weak geometric structure of the scene.

Since the remaining residual errors accumulate and the existing drift of visual-inertial odometry is not completely eliminated, a second registration with a continuous-time ICP method is performed in the next step. The basic idea is that the trajectory error in temporal proximity of a considered pose is negligible. The trajectory is then split into subsections, and several successive 3D scans are combined to form a submap. After that, the partial maps are again registered against their predecessors. The change in pose of a reference scan is then distributed to the poses between two reference scans to maintain the continuity of the trajectory. For small trajectory changes, a linear distribution (translation) or Spherical Linear Interpolation (SLERP) (rotation) is sufficient.

To correct the accumulated drift over longer trajectory sections, loops are detected and closed [305]. For this purpose, the poses of the aggregated submaps are optimized, and the changes are subsequently distributed to the individual poses analogously. Post-processing of the data to create the final 3D point cloud is performed with continuous-time SLAM [116]. In this step, the subsections of the trajectory are created in a way that they partially share scans with the previous and next submap in order to ensure sufficient overlap for registration.

The resulting map is provided for live visualization during the measurement process. This is done by aggregating single scans in a combined point cloud based on the optimized trajectory.

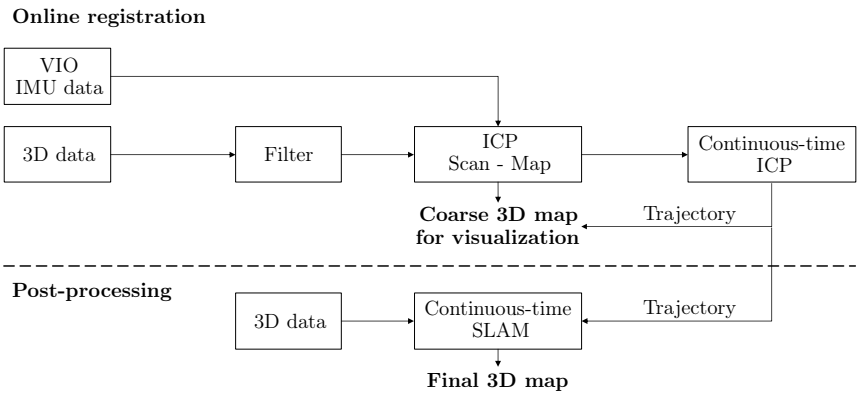


Fig. 6.16: 3D registration pipeline.

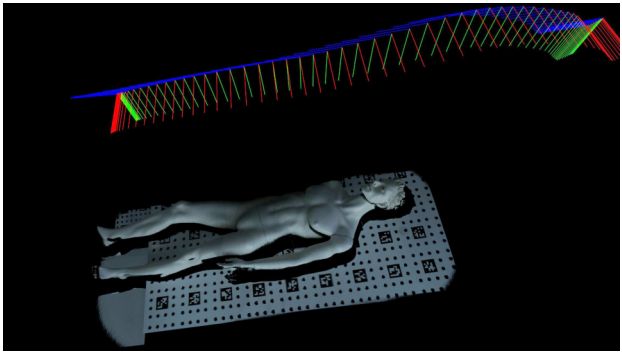
The point cloud is recomputed and sub-sampled in a parallel process. This way, approximately every 1 to 5 s an update of the complete map is created for visualization.

6.2.6 Motion Compensation

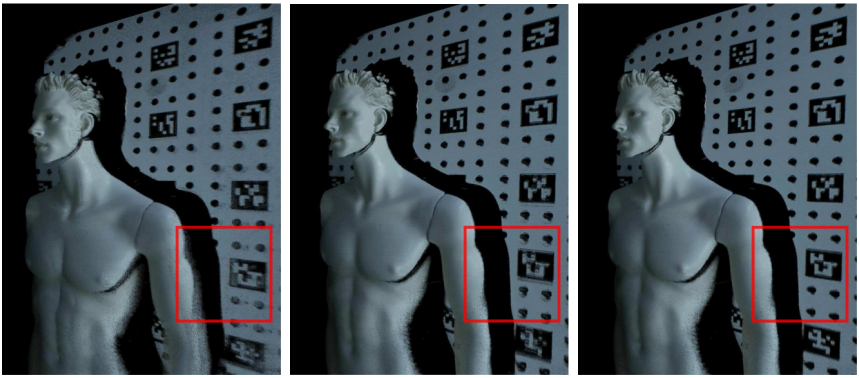
As discussed in Section 6.2.2, motion compensation is necessary for every 2D stereo image of the measurement sequence because otherwise the assumption of a static scene for the correlation-based correspondence search in the multi-shot structured light approach is violated. However, precise compensation is difficult because it requires knowledge of the object points to compute the transformation in image space. The object coordinates are not known before the 3D triangulation.

Hence, certain simplifications are considered: First, a linear uniform motion is assumed. This is a reasonable assumption for scan trajectories with limited rotation. While the overall trajectory is not linear, locally within the window of a single scan, a linear movement is assumed. Second, the shift in image space is computed based on an assumption of the typical measurement distance. This is done based on the interpolated pose estimates of the VIO. Specifically, the average shift in image space is computed for object points in a plane parallel to the sensor coordinate frame at the nominal measurement distance of 2 m. The shift is computed for the rectified 2D stereo images. This way, after applying the 2D pixel shift vector, the 3D reconstruction pipeline [157] is applied as usual.

Besides motion compensation of the 3D scans, the trajectory estimate is also necessary for accurate color mapping. Since the color images are captured interleaved between the stereo images, there is a time difference of up to 20 ms between the 3D scan and the next color image. Therefore, the movement of the sensor needs to be compensated to accurately map color information on the point cloud. This is done the following way: First, the VIO trajectory is interpolated to compute the camera's movement between the acquisition time of the 3D scan



Point cloud and trajectory.



Visual-inertial odometry. 3D registration. Improved color mapping.

Fig. 6.17: Point cloud of a mannequin acquired in the lab by moving the sensor. Top: Point cloud and visualization of the sensor poses. Bottom left: Initial result using VIO. Bottom middle: Result after 3D registration. Bottom right: Improved color mapping based on trajectory interpolation.

and the color image. This is done based on the timestamp of the sensor information. Then, the point cloud is transformed to the color camera frame based on the co-calibration of the camera and stereo sensor and the estimated camera movement. Each point is then colored by projecting the point in the color image based on the camera model.

Fig. 6.17 shows an example of a point cloud captured in a lab setting. A mannequin is scanned in air in front of a calibration board. The top image shows the colored point cloud combined with a visualization of the sensor poses. The bottom image show the effect of the different trajectories

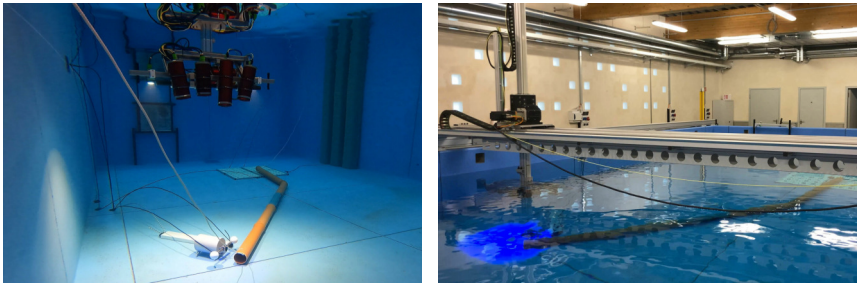


Fig. 6.18: Underwater test tank at Fraunhofer IOSB AST in Ilmenau. Left: Underwater scene with test objects and pipe structure. Right: Gantry system with three-axis.

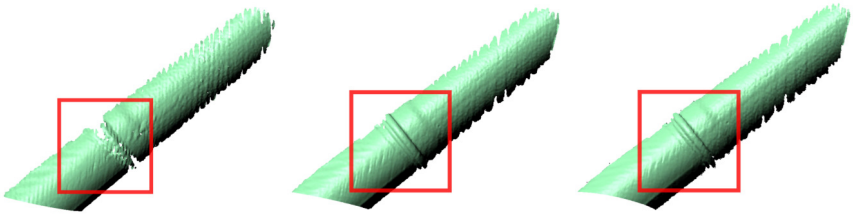
on the point cloud result. The bottom left images shows the result using the VIO trajectory estimate computed in real-time. In the highlighted section, some misalignment of the scans is observed. The middle image shows an improved result based on 3D registration. This results in a consistent point cloud. However, the color mapping is inaccurate since the color information is projected based on the next color image without considering the sensor motion. Errors in the color projection are especially visible at the edges. These errors are reduced in the final result shown in the bottom right image. Here, the sensor movement is taken into account for accurate color mapping.

6.2.7 Measurement Results

The underwater sensor was tested at Fraunhofer IOSB AST in Ilmenau in the underwater test facility. Here, dynamic recordings were carried out on a three-axis gantry, shown in the right image in Fig. 6.18. A pipe structure was placed in the water as well as a spherical dumbbell and a planar target. The underwater scene is shown in the left image of Fig. 6.18. Recordings were captured with two sensor velocities of about $0.1\,\frac{\mathrm{m}}{\mathrm{s}}$ and $0.7\,\frac{\mathrm{m}}{\mathrm{s}}$. The measurements were carried out in clear water.

Fig. 6.19 depicts an example of a single scan of the pipe structure to show the effectiveness of the motion compensation. Here, an area of approximately $0.9 \times 0.7\,\mathrm{m}$ of the pipe is captured at a sensor velocity of ca. $0.7\,\frac{\mathrm{m}}{\mathrm{s}}$. The left image shows the 3D reconstruction result without motion compensation. The middle image shows the result with a manually estimated pixel shift. The right image shows the automatic motion compensation result based on the VIO trajectory. The effect of motion compensation is most prominently visible in the part of the pipe around the socket, which is highlighted in red. The reconstruction quality and in particular the completeness of the point cloud is improved by applying the motion compensation. Here, little difference is observed between the manual and automatic motion compensation. Moreover, the experiments showed that the standard error of measurements of the spheres of a dumbbell target is reduced using automatic motion compensation [6].

Fig. 6.20 shows the generated point cloud, which was assembled from 680 individual scans



No motion compensation. Manual motion compensation. Automatic motion compensation.

Fig. 6.19: Single scan of the pipe structure with and without motion compensation. Left: Reconstruction result without motion compensation. Middle: Result with manually estimated pixel shift. Right: Automatic motion compensation result based on VIO. (Source: [6])

based on the trajectory determined with VIO. In the upper part of the image, the determined sensor poses are visualized. The top image shows the point cloud colored according to the height. The bottom image shows the point cloud textured based on the RGB images of the color camera.

The trajectory determined by visual odometry is verified by comparison with the movement recorded by the gantry system. From the gantry, the translation is precisely determined via encoders of the individual linear axes. In contrast, visual odometry provides both translation and rotation of the sensor. In this comparison, therefore, only the estimate of the translation is compared.

Here, no calibration was performed between the gantry system and the underwater sensor. The trajectories are compared by aligning the two trajectories against each other at the starting point. The reference trajectory of the gantry system is planar. The trajectory of the visual odometry is also approximately planar with a maximum deviation of about $\pm 1\,\mathrm{cm}$. Fig. 6.21 shows the top view of the trajectory and the deviation between gantry data and the visual odometry solution for two different motion speeds. The deviation here is calculated as the distance between the two 3D trajectories. The acquisition starts at the zero point.

The top images of Fig. 6.21 show the result for the slow sensor speed of about $0.1\,\frac{\mathrm{m}}{\mathrm{s}}$. In the left graph, it is seen that the two trajectories are well aligned. The right graph shows small deviations at the beginning of the linear trajectory, increasing significantly around the turning point. This is probably due to a residual error in the scaling of the visual odometry data. Moreover, the images in the middle part of the image show less structure since here only the tube is visible against a background of homogeneous color.

The result for the higher velocity of $0.7\,\frac{\mathrm{m}}{\mathrm{s}}$ is comparable, but the errors are generally higher, as is seen in the bottom images of Fig. 6.21. The trajectory here is not identical to the measurement at slow speed due to the different acceleration and deceleration behavior of the linear axes of the gantry system.

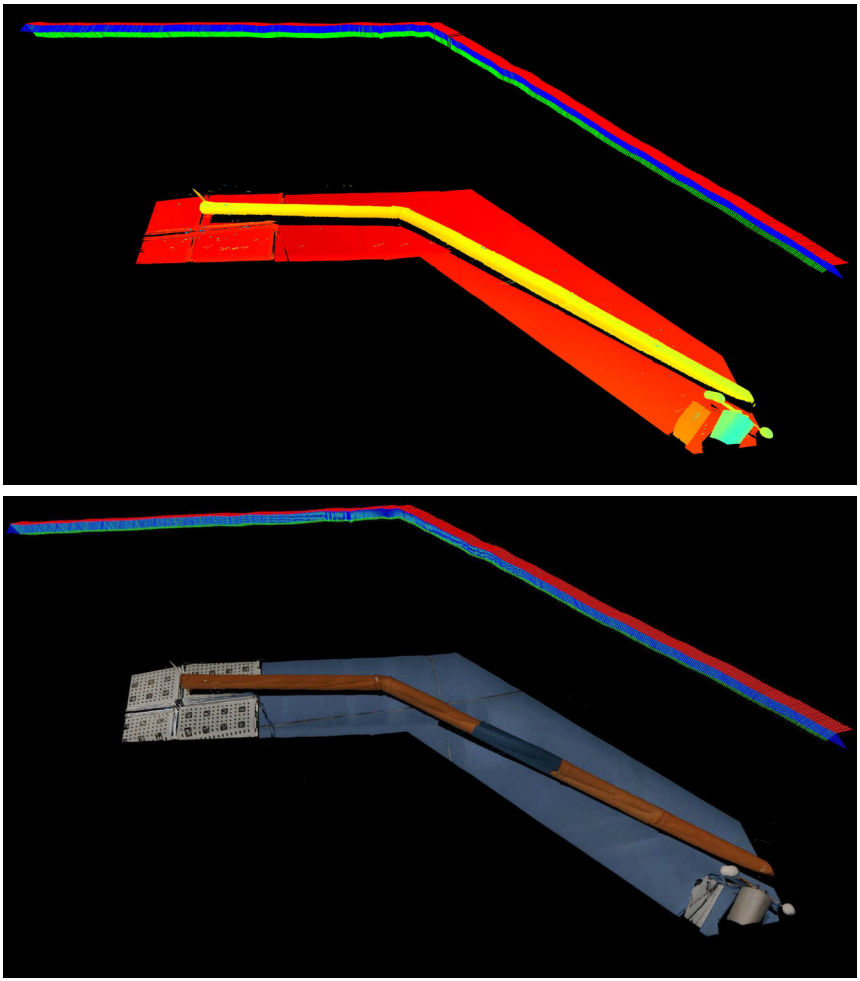


Fig. 6.20: Combined point cloud and visualization of the sensor trajectory. Top: Point cloud colored by height. Bottom: Point cloud colored using RGB images.

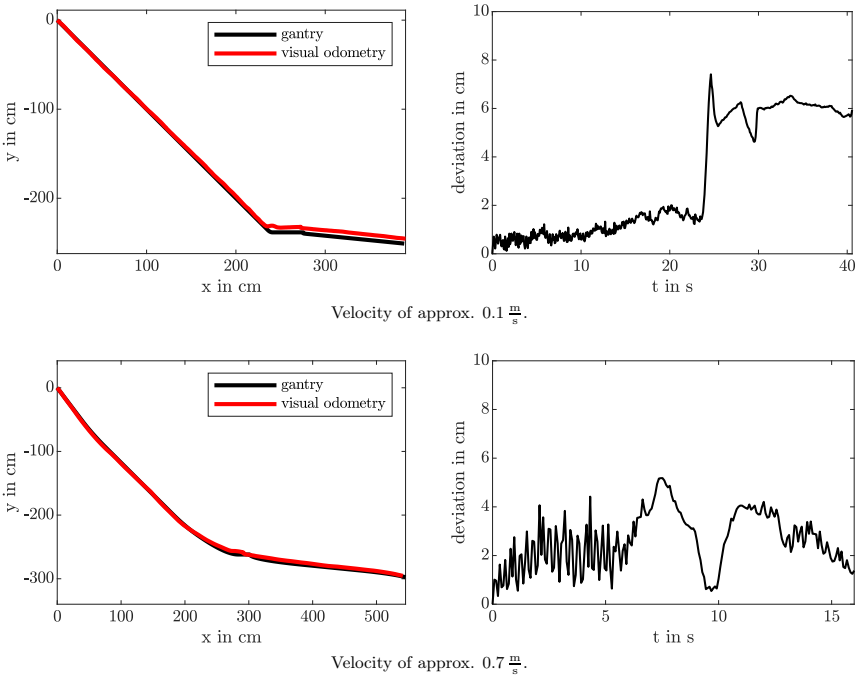


Fig. 6.21: Trajectory and deviation between gantry data and visual odometry for two different sensor velocities.

6.3 Discussion

In this chapter, the proposed underwater laser scanning system was extended for mobile mapping using satellite navigation. While it is demonstrated that this enables scanning from a floating platform, the accuracy of the estimated trajectory is too low, given the achievable 3D point measurement accuracy of the scanner.

Therefore, an approach based on visual odometry and SLAM is developed for an optical 3D underwater sensor in the UWSensor project. The 3D underwater scanner allows a dense reconstruction and high sampling rate and is consequently well suited for mobile 3D mapping in water. However, with fast sensor movement, the assumption of a static measurement scene for the multi-shot 3D reconstruction method is violated. With a linear sensor movement of 1 $\frac{\text{m}}{\text{s}}$, the sensor moves approx. 1-2 cm within a 3D image sequence, depending on the selected acquisition frequency and the number of images. This movement must be considered for precise measure-

ments at fast sensor speeds. The presented compensation via displacement in the image space already achieves a significant improvement in the envisioned application. The reconstruction accuracy and completeness of the scan are enhanced. This is demonstrated by experiments with a gantry system for different velocities of the data acquisition in water. Furthermore, the developed motion compensation approach improves the 3D output data, allowing a finer registration and the recording of more accurate 3D models. The UWSensor system was recently successfully tested in an offshore measurement campaign in the Baltic Sea. The evaluation of the acquired data is ongoing work [7, 16].

Chapter 7

3D Underwater Modeling for Remote Control

The Horizon 2020 ¡VAMOS! project developed a novel prototype mining system to extract raw materials from water-bearing open pit mines [41, 312, 344]. To achieve this, it uses a remotely operated mining vehicle, which is launched from a pontoon vessel. The left image in Fig. 7.1 shows the ¡VAMOS! mining machine hoisted on the crane during launch. The remotely operated miner is used for cutting the rock and for collecting the material. The challenge of operating such a large mining vehicle in water is that there is no intervisibility between the vehicle and the operator. Therefore, the machine's state and the surroundings is only communicated via a computer interface. This impedes precise control since the pilot has to rely only on the available data that is streamed from the vehicle's sensors. Additionally, optical cameras, which are often used for the teleoperation of robots and machines, provide very little information in this scenario due to the high turbidity of the water. Especially during cutting operations or if the tracks raise fine particles, the visibility is typically below one meter. The monitors depicted in the right image of Fig. 7.1 show examples of the camera feeds of the mining vehicle during a field test. Due to turbidity, even the cutter and the backhoe bucket of the machine itself are barely visible. This necessitates employing perception sensors that are less affected by turbidity in the water, such as imaging sonar or acoustic cameras. However, sonar images are more difficult to interpret visually by a remote operator.

Hence, one part of the efforts in the ¡VAMOS! project concentrated on enhancing the operator's situational awareness by creating a 3D model of the mine above-the-water and underwater, which captures the mining site as detailed as possible. It is well known that a map of the environment, in addition to the raw sensor data, is extremely helpful in supporting remote control and enhances spatial awareness [238, 282].

The work described in this chapter was previously published in [1,3–5]. This is joint work with the ¡VAMOS! project partners, particularly with INESC Technology and Science Porto, which developed the perception, positioning and navigation systems, and with BMT WBM, which developed the VR system. The author's contributions encompass the methods and approaches to create a 3D map of the mining environment from the sensor and positioning data for visualization of a consistent digital replica of the mining operations in VR.

Fig. 7.1: Left: ¡VAMOS! mining vehicle hoisted on the crane of the launch and recovery vessel (Source: ¡VAMOS! project [318]). Right: Camera feeds of the mining vehicle displayed in the control cabin.

7.1 The ¡VAMOS! Underwater Mining System

The main components of the ¡VAMOS! underwater mining system are depicted in Fig. 7.2. On the pit floor, the underwater mining vehicle is deployed. It is propelled with continuous tracks and receives power and control signals via an umbilical cable. As the mining vehicle cuts the rock face with a roadheader, small rock fragments are created, which are mixed with water and subsequently collected by a dredge suction mouth. The material is then transported as slurry to the surface with a built-in pump using a flexible riser hose. On the launch and recovery vessel, an additional support pump transports the material to the shore via a floating pipe. A dewatering plant separates the raw material from the slurry. Any excess water than flows back into the mining pit. This way, the water table is not significantly altered.

The launch and recovery vessel is used to deploy the mining vehicle on the mine pit and recovering it. The vessel is propelled using electric winches and four cables, which are connected to anchor points on the shore. A portal crane built on top of the vessel is used to hoist the mining machine. On the end of the winch wire, a bullet is mounted, which latches on to a mechanism on top of the mining vehicle. This way, the machine is deployed to the pit floor and detached from the winch wire. Later, the bullet is re-attached for pick-up and recovery. However, precise positioning is necessary to move the mining vehicle back below the vessel's crane with small errors.

In order to support the mining operations, an Hybrid Remotely Operated Vehicle (HROV) was developed in the project. It is a hybrid ROV in the sense that it is manually controlled via an acoustic link, but it also carries out pre-programmed tasks autonomously, such as performing automatic surveys. The HROV is used for performing multibeam sonar surveys to create updated bathymetry of the complete mine pit. Additionally, optional a 3D sonar is mounted, which enables to provide real-time imaging of the mining operations.

The whole system is controlled from a control cabin on shore. Part of the system, such as crane operations, is also controlled for safety reasons directly via control interfaces on board the vessel. Wired and wireless intercom systems are used for communication between personnel on board the vessel and the operators in the control cabin.

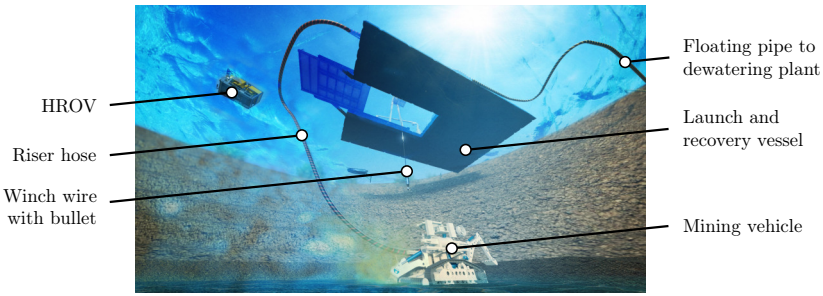


Fig. 7.2: Rendering visualizing the concept of the ¡VAMOS! underwater mining project (Source: ¡VAMOS! project [318], original rendering by DAMEN Dredging Equipment).

7.1.1 The Underwater Mining Vehicle

The main tool of the mining machine is a hydraulic roadheader cutter mounted on a boom on the front of the vehicle. The cutter head and boom are the red-painted part of the machine depicted in the top left image of Fig. 7.3. Bellow the cutter is a dredge suction mouth with a rotating auger for collecting the material. Tracks with different widths and materials are mounted to adapt the locomotion to soft and hard terrain. On the top of the vehicle, a cone with a diameter of 1 m is mounted. This is visible in the image taken from the crane tower in the top middle image of Fig. 7.3. The cone is used to allow for some error while catching the bullet attached to the winch wire. The bullet is centered by the cone and then fixed by a latching mechanism. During launch and recovery, a thruster is used to prevent the vehicle from spinning in the water column and maintains the orientation. On the back of the vehicle a hydraulic arm with an interchangeable tool connector is mounted. The top right image of Fig. 7.3 shows a backhoe shovel tool, used for excavation or collecting rock samples.

For maneuvering and supervising the operations, multiple perception sensors are installed on the mining vehicle [1]. The perception sensors installed on the sensor bar at the front of the mining vehicle are depicted in the bottom image of Fig. 7.3. A Kongsberg M3 multibeam sonar is installed on a pan-and-tilt unit with rotary encoder feedback, which is automatically panned/tilted for 3D scanning and mapping. The multibeam sonar has a profiling as well as in imaging mode. The sensor has a measurement range of 0.2 m to 120 m. In profiling mode, the horizontal field of view is 120 deg and the vertical beam width is 3 deg. A profile with 256 beams is measured with an update rate of up to 40 Hz, depending on the range limit. Typically, measurement rates in the range of 3 Hz up to 10 Hz are used during operations. The measurement plane of the sensor is aligned horizontally with the vehicle and the sensor is tilted up and down for 3D scanning. The vertical field of view of the scanning multibeam is 60 deg and a full 3D scan is created every 10 s. Alternatively, a Coda Octopus Echoscope 3D sonar provides real-time scans of the work surface with a high update rate of 20 Hz. In contrast to the multibeam sonar, this sensor captures dense data of a FoV of 40 $\times$ 40° with 128 $\times$ 128 beams. The measurement range is 80 m. Both sonar sensors provide centimeter range resolution.

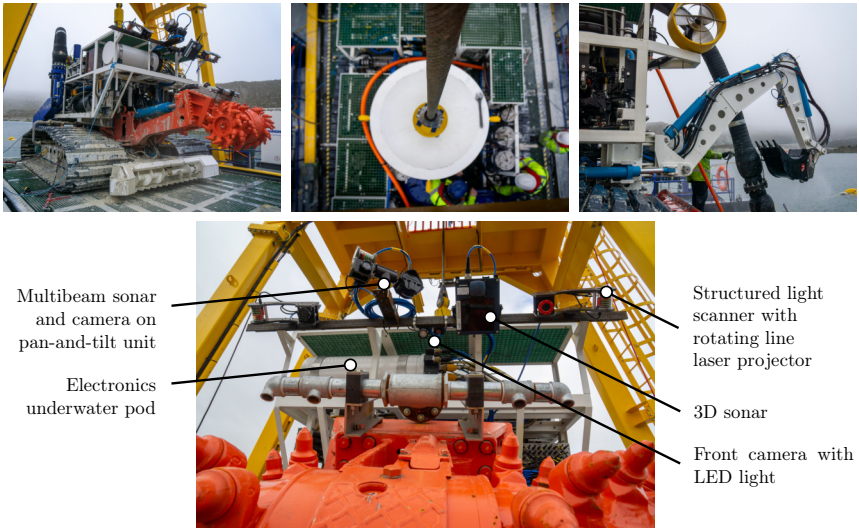


Fig. 7.3: Top left: ¡VAMOS! mining vehicle on the deck of the vessel. Top middle: Crane bullet latched in the cone-shaped mechanism on top of the mining vehicle. Top right: Backhoe shovel mounted on the interchangeable tool carrier on the back of the vehicle. Bottom: Perception sensors mounted on the front of the mining vehicle.

Cameras and LED lights are mounted on the pan-and-tilt unit, pointed at the cutter head, and pointing backward to the backhoe bucket. Furthermore, two custom-built structured light scanners with rotating line lasers and LED flashes are installed on each side of the sensor bar. However, due to the high turbidity of the water during cutting operations, optical sensors provide little information for the operator.

For positioning and navigation, an Evologics Ultra Short Baseline (USBL) transceiver combined with a KVH1750 Fiber Optic Gyro (FOG) based IMU is installed on top of the vehicle behind the sensor bar [36]. To measure the height above ground during the lowering of the vehicle an acoustic altimeter is employed. A pressure sensor additionally supports the dive depth estimates. Additional sensors measure the articulation of the cutter boom, suction mouth, backhoe shovel, and movement of the tracks.

7.1.2 The Launch and Recovery Vessel

The floating platform is built from modular pontoons. The assembled vessel is shown in the left image of Fig. 7.4. A portal crane is employed for launching and recovering the mining vehicle. Below the crane is a moveable platform. This way, for deployment, the mining vehicle is hoisted on the crane, the platform then moves to the side, and the vehicle is lowered directly

Fig. 7.4: Left: Launch and recovery vessel (Source: ¡VAMOS! project [318]). Middle: GNSS antennas mounted on top of the portal crane tower. Right: SBL transceivers mounted on the corners of the vessel (Source: [214]).

into the water. Furthermore, the vessel houses the generator for the operation of the crane and winches, part of the slurry transport circuit, such as a support pump, and the control systems and connections to the control cabin on shore.

The position and orientation of the Launch and Recovery Vessel (LARV) is estimated using GNSS RTK. Three GNSS antennas are mounted on top of the portal crane tower, which is shown in the middle image of Fig. 7.4. Additionally, multiple cameras are installed on the LARV to provide an overview of the LARV operations to the operators in the control cabin. Winch sensors provide the load and length of the unspooled winch wire. This is also used to provide a rough estimate of the position of the mining vehicle before the vehicle is deployed deep enough in the water for the acoustic positioning to work. On three edges of the LARV Evologics Short Baseline (SBL) transceivers are mounted, which are used in the underwater positioning network.

7.1.3 The Hybrid ROV

To support the mining operations, a HROV was developed in the project, depicted in Fig. 7.5. A detailed description of the HROV is found in [213]. For surveying the mine the same type of sonar sensors as installed on the mining vehicle are used. A 70 deg downwards tilted Kongsberg M3 MBS or alternatively a CodaOctopus Echoscope 3D sonar is mounted.

The HROV and the installed perception and positioning sensors are shown in Fig. 7.5. On the water surface the HROV directly infers precise positioning information using GNSS. This is achieved with two antennas so the heading is also directly computed from the RTK solution. Additionally, a fiber optic gyro based INS, a pressure sensor and a DVL are used for pose estimation. During dives of the HROV acoustic positioning using an Evologics USBL modem is used. Additionally, two custom-built Structured Light Scanner (SLS) are installed on the front of the vehicle.

The HROV is controlled via an acoustic link. On the surface, Wireless Local Area Network (WLAN) is used to stream live sensor data to the control cabin. During dives, only limited data is available due to the bandwidth restrictions of the acoustic link. In order to stream multibeam and 3D sonar scans to the control center in real-time during dives, a small support boat with a WLAN access point and a short umbilical to the HROV is optionally employed. All data is also logged on hard disks, so the complete data is available after a mission. For easier deployment

Fig. 7.5: Sensors installed on the HROV.

and transport on land four removable wheels are attached to the HROV.

7.1.4 The Control Cabin

The system is remotely controlled from a control cabin on the shore, depicted in Fig. 7.6. It features two operator stations, as is visible in the left image. One operator pilots the mining vehicle while the second operator supervises the LARV operations and the slurry circuit. Joystick and touchscreen controls are used for most of the machine and vessel systems. Additionally, keyboard and mouse interfaces are used to interface with the computer systems. In front of the operators, a configurable video wall is placed to display system information and real-time sensor data and visualizations. The right image in Fig. 7.6 shows the operator remotely controlling the underwater mining machine using the VR interface. The VR system is used during planning, e.g., for finding a suitable landing position on the bottom of the pit and during operations. Moreover, live data from some of the environmental sensors is fed into the system to visualize measurements, such as suspended sediments in the water, for real-time monitoring.

The sensor data of the mining vehicle and launch and recovery vessel is directly streamed to the control center via an umbilical and optical fiber. This provides low communication delays. The HROV streams high-bandwidth sensor data on the surface via a long-range WLAN connection to the control center. The antenna of the GNSS RTK base station is mounted on top of the control cabin on land.

7.1.5 Positioning and Navigation

The positioning and navigation system of ¡VAMOS! is described in [1,36,37]. It was designed with the goal that the system works without infrastructure installed in the mine. This way, the system is also applicable for off-shore use. Additionally, the position determination is computed on board the vehicles. The reason for this is that only on the vehicles the complete sensor information of all navigation and positioning sensors is directly available. This way, positioning is computed even if the acoustic link is unavailable or some sensors do not provide valid information. For georeferencing, the pose of all assets is computed relative to a fixed GNSS RTK base station, which is mounted on top of the control cabin. The base station antenna is chosen as the origin of the local coordinate system used for the positioning.

Fig. 7.6: The control cabin has two work places: one for the operator of the mining vehicle and one for the operator of the LARV and slurry circuit.

The pose of the LARV is determined using three multi-frequency GNSS antennas mounted on top of the crane tower. The position of each antenna is solved individually using GNSS RTK with the fixed base station. One of the antennas is chosen as the origin of the local vessel coordinate system, and the two vectors to the other antennas are used to compute the orientation.

Underwater positioning is achieved using an acoustic positioning network present on all vehicles. The network architecture is visualized in Fig. 7.7. It combines SBL positioning with inverted Ultra Short Baseline (iUSBL) positioning. A SBL network consisting of three Evologics Mini Modems is mounted to the corners of the vessel with fixed relative baselines of approximately 15 to 20 m. The position of these transceivers is georeferenced using the GNSS RTK determined pose and a calibration of the relative poses of the transceivers. Subsequently, all relative underwater position measurements are converted to global positions.

On the HROV and the mining vehicle an USBL transceivers are installed. Typically, USBL transceivers are installed on vessels looking downwards to locate mobile subsea transponders. Here, the USBL transceivers are used in an inverted (iUSBL) configuration. This means the underwater vehicles carry the USBL transceivers looking upwards to the vessel. USBL transceivers calculate both range and bearing. The range is computed from the travel time of the acoustic signal, and the angles are measured using an array of transducers. By looking at the phase difference within this transducer array, the direction is computed. The iUSBL devices of the vehicles interact cyclically with the SBL transceivers installed on the vessel. This way, every time the SBL network pings the vehicles, both vehicles determine an iUSBL position measurement at the same time. In order to determine the full 6-DoF poses, the vehicles carry additional sensors, such as INS integrating triaxial accelerometers, FOG and magnetometers. During launch and recovery, the winch length is used as additional pose information input. Furthermore, pressure sensors are used for depth measurement. Additionally, odometry computed using the movement of the tracks is used for pose estimation while driving on the pit floor. The individual sensor measurements are fused using Kalman Filtering [36].

While close to the water surface, the HROV directly infers precise positioning information using GNSS. This is achieved with two antennas so the heading is also directly computed from the RTK solution. Additionally, the HROV uses INS, a DVL and depth sensors for pose estimation.

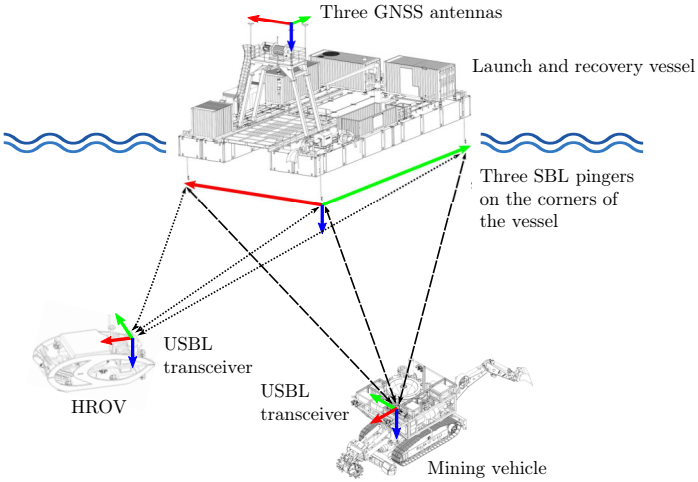


Fig. 7.7: Principle sketch of the acoustic underwater positioning network. Image reproduced based on [1].

7.1.6 Virtual Reality Human-Machine Interface

The raw sensor data, terrain maps, positioning and navigation data, and system information is transferred to the Human-Machine Interface (HMI) via Local Area Network (LAN). The HMI is based on a custom-built VR application built on top of the Unity gaming engine [322]. It provides a 3D VR model of the entire mining operations. This includes models of all vehicles and vessels and the riser system and relevant static structures. The model is dynamically adjusted such that it faithfully replicates the current state of the real operations. As output device, standard monitors are chosen and a computer mouse is used for interacting with the 3D scene in VR. This input scheme was selected because it integrates well with the monitor wall used in the control center for displaying camera streams and data visualizations from other ¡VAMOS! subsystems as well as the joystick and touchscreen-based control interfaces for the mining vehicle and vessel. Using the VR system with more immersive output devices, such as VR headsets like Oculus Rift or HTC Vive, is also possible. However, for prolonged operations, problems like simulator sickness and increased eye strain are an issue.

Fig. 7.8 shows the digital replica of the Magcobar mine site in Silvermines, Ireland, and the components of the ¡VAMOS! underwater mining system in VR. In the top image, the mesh model created from the terrestrial laser scans together with the flooded open-cut mine and the launch and recovery vessel is shown. The water level is dynamically adjusted using measurements from a tide gauge. The bottom images show detailed views of the vessel during the deployment of the mining vehicle together with the underwater terrain.

The VR system is used for control of the mining vehicle, the LARV and the HROV. It provides

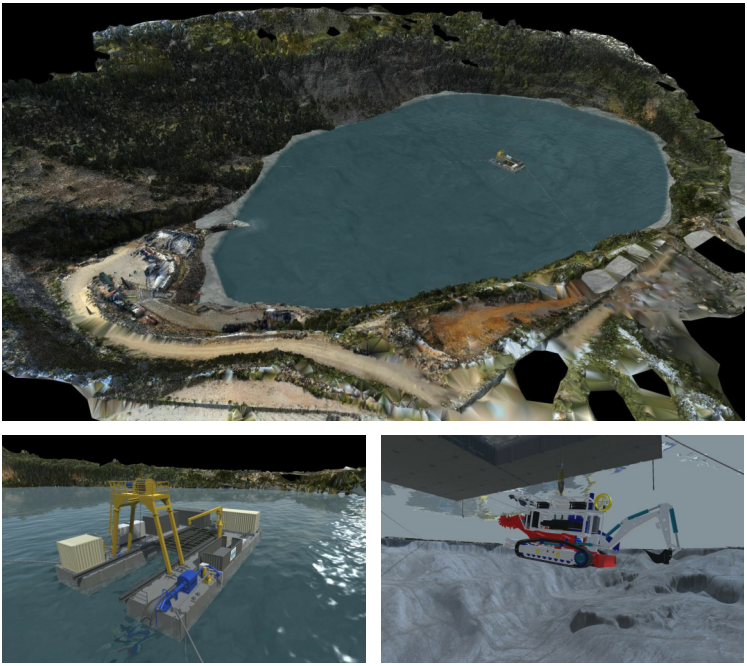


Fig. 7.8: Model of the mining site in VR. Top: model of the mine with the colored mesh of the above-the-water terrain. Bottom left: Detail view of the launch and recovery vessel. Bottom right: Visualization of the mining vehicle with underwater terrain model during deployment.

a range of functionalities to support teleoperation as well as assistance during the planning of operations. For example, it provides guidance in maneuvering the vessel to lower or pick up the mining vehicle at the chosen position. It enables path input for autonomous surveys using the HROV, and it provides awareness when maneuvering the mining vehicle on the pit floor and during cutting operations.

Fig. 7.9 shows different visualization options available in the VR systems. The top left image shows the above-the-water point cloud of the mine site combined with real-time sensor data from multibeam sonar colored by height during an HROV survey. The top right image shows the fully articulated model of the mining vehicle combined with real-time point cloud data from the mechanically panning multibeam sonar during cutting operations. The bottom left image shows the mining vehicle with terrain model updates of the work surface. The bottom right image shows split views for piloting the mining vehicle during cutting operations. The view point for each split view is individually adjustable.

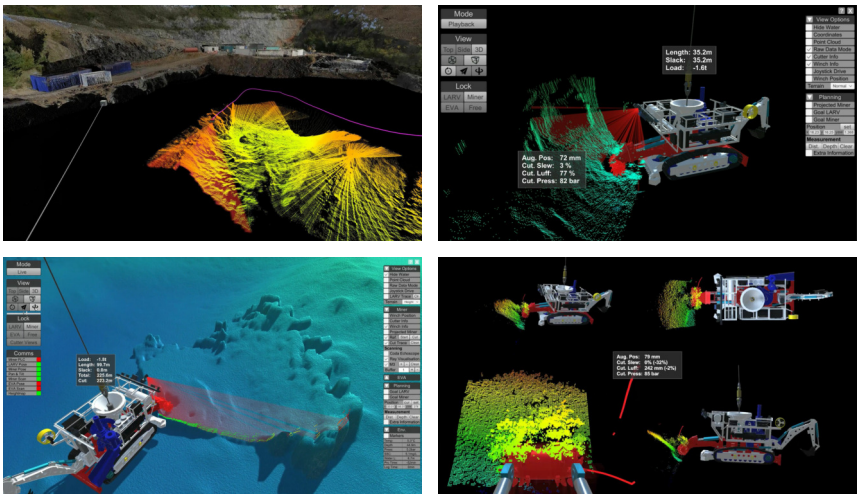


Fig. 7.9: Different views of the VR based HMI. Top left: Live multibeam data visualization during a survey using the EVA HROV. Top right: Mining vehicle with visualization of multibeam sonar scans. Bottom left: Mining vehicle with continuously updating terrain model using a mechanically panning multibeam sonar. Bottom right: Split-views for piloting the mining vehicle during cutting operations.

7.1.7 Test Sites

The complete ¡VAMOS! mining system was tested in two field trials at the Lee Moor mine site near Plymouth, United Kingdom and at the Magcobar mine in Silvermines, Ireland. Additional tests of the perception, positioning and navigation systems were carried out in Portugal at the Bejanca mine near Queirã and the Malaposta stone quarry near Santa Maria da Feira. Images of the test sites are shown in Fig. 7.10.

The Bejanca mine is a tin and tungsten deposit and was in operation from 1919 to 1944. During that time, about 30 tons per month of mineral of cassiterite and wolframite was produced. The pit is approximately 160 m long and 80 m wide. The maximum water depth is ca. 30 m. The site was not selected for field trials, but a test data set for underwater mapping and modeling of the mine was acquired at Bejanca.

The Whitehill Yeo Pit is located at the Lee Moor mine site north of Plymouth in Devon, UK. The open pit was mined hydraulically for China clay production and went non-operational in 2008. The pit is approximately 350 m long and 260 m wide. The maximum water depth was expected to be larger than 50 m. However, the depth observed during trials was less than 30 m with a thick silt layer on the bottom of the pit.

The Magcobar mine in Silvermines in Tipperary, Ireland, is a flooded opencast baryte mine. The mine was closed in 1993. The site was selected as the second trial site for the ¡VAMOS!

Bejanca mine

Whitehill Yeo Pit, Lee Moor

Magcobar mine, Silvermines

Fig. 7.10: Three of the ¡VAMOS! test sites: The Bejanca mine, the Whitehill Yeo Pit at Lee Moor and the Magcobar mine in Silvermines.

project. During the trials, the mining machine was deployed in depths of up to 57 m, cutting in different rock types under various operating conditions. The pit is approximately 410 m long and 260 m wide.

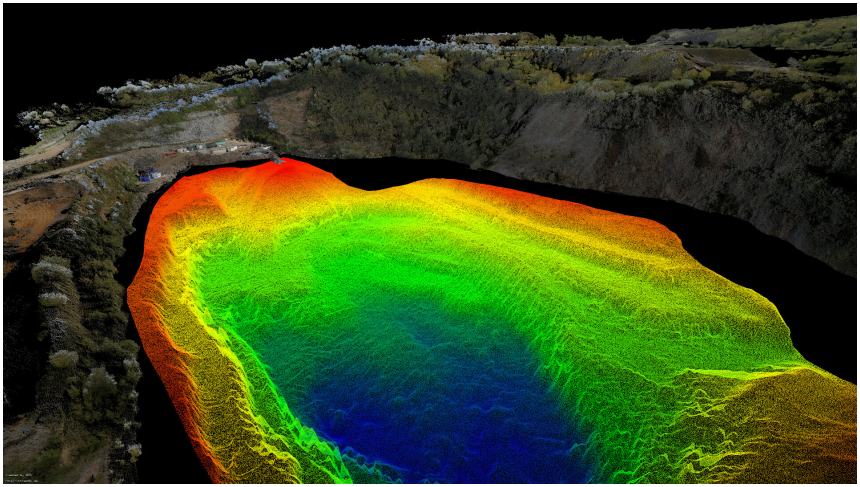


Fig. 7.11: Combined point cloud above-the-water and underwater of the Silvermines site. The above-the-water model is colored with RGB data. The underwater model is colored by water depth.

7.2 3D Mine Modelling

An above-the-water model is created from terrestrial laser scanning and camera images. The scans are registered into a consistent model, and a high-resolution point cloud as well as a lower-resolution mesh for faster rendering of the complete scene are created. The model of the underwater site is acquired using multibeam sonar data captured using an HROV. The models are visualized to the operator in a VR system. Fig. 7.11 shows an example of the combined point cloud of the Silvermines site above the water and underwater.

Since we know the location of all mining assets using the positioning information from GNSS and an underwater positioning system, the mining vehicle, launch and recovery vessel and HROV are rendered accurately in the 3D scene [1]. This way, a replica of the physical system and the environment is built which provides real-time and contextual information in a consistent and comprehensible way.

All navigation sensor data is combined and fused to provide real-time, accurate and precise information about the position and orientation of all ¡VAMOS! systems [36]. This navigation information and the mine perception data from multibeam and 3D sonar are used to update the 3D model with real-time data and with off-line survey data.

7.2.1 Time Synchronization and Calibration

Time synchronization is achieved using the NTP and Pulse-Per-Second (PPS) signals. All computer systems and sensors are synchronized this way to GNSS time. This allows to also correlate offline measurements to data logs in retrospect based on timestamps.

The sensors are individually calibrated depending on the sensing modality. Calibration of the sensor poses is challenging because many different sensors are employed, and it is costly to design calibration fixtures which are visible, for example in optical sensors as well as in sonar sensors. Moreover, considering the large size of the mining vehicle, very large calibration targets would be necessary for accurate calibration, such that they are visible in multiple sensors. Therefore, sensor mounting positions are estimated from a combination of manufacturing CAD models, laser scans taken of the vehicles after integration, GNSS positions, and tape measure measurements.

For example, initial estimates of the sensor poses were created by laser scanning of the mining vehicle and then fitting CAD models of the individual sensors to the point cloud. Similarly, the LARV was scanned to find the positions of the GNSS antennas and the relative poses of the mounting poles for the SBL transducers for the underwater acoustic network. The rotational offsets of the 3D perception sensors are further optimized based on an error measurement which determines point cloud quality similar to [115, 294].

7.2.2 Above-the-water and Underwater Mine Mapping

The above-the-water model is created by registering the laser scans using ICP scan matching combined with a global relaxation and loop closing [54,250]. Color information from RGB images is mapped on the point cloud based on a co-calibration of laser scanner and camera [53, p.155ff]. For the above-the-water survey, a Riegl VZ-400 terrestrial laser scanner combined with a Single-Lens Reflex camera (SLR) mounted on a custom-built camera mount is employed [55]. The complete point cloud is georeferenced using GNSS position measurements and is sub-sampled to provide a static high-resolution point cloud for visualization. Additionally, the point cloud is meshed and imported as a simplified triangular mesh with color since this allows faster rendering of the complete scene in VR.

The bathymetry is initially created from a multibeam survey using the HROV and later updated with additional surveys and perception data from the mining vehicle. The initial survey is processed offline using a continuous-time SLAM algorithm, which optimizes point cloud consistency globally, i.e., for all the sensor measurements of the complete map.

Sensor measurements from the perception sensor systems, such as multibeam sonar and 3D imaging sonar, are fused into a consistent 3D representation. For integrating measurements from multiple sensors SDF-based mapping is selected. SDF represent the surfaces implicitly by storing in each voxel cell the signed distance to the closest surface. Typically, the signed distance is only stored in a narrow band around the surfaces, which is referred to as a Truncated Signed Distance Function (TSDF). A SDF voxel map is a beneficial surface representation because noisy measurements are smoothed over multiple observations.

As the mine changes over time due to the mining operations themselves, the internal representation of the mining environment needs to be constantly updated based on new sensor observations. For real-time processing, the model is only updated in a small window of 3 - 5 m

around the initial survey. This way, only a small number of cells need to be updated, and coarse outliers or measurements of the riser etc. are not integrated into the terrain model.

Global optimization using continuous-time SLAM is not applied during online processing. Since a valid mine model from the initial survey already exists, this model is used to minimize misalignment for the real-time processing. This is done by registering new sensor data with the established mine model. Since this requires only finding an alignment between the sensor scans and the model, it is possible to compute this in real-time. This alignment compensates for certain accumulated errors during the mapping process.

The underwater model is created with a grid resolution of 10 cm. The terrain model is split into 50 m × 50 m tiles, which are individually sent to the VR-based human-machine interface. Therefore, updating the underwater terrain model in real time requires transferring only the modified part of the map over the network.

7.2.3 Continuous-time SLAM

Misalignments between multiple laps of the survey due to inaccuracies of the vehicle trajectory measurement or imprecise calibration result in errors of the SDF model. In this case, individual scan segments might not line up very well, which creates artifacts in the resulting surface model, such as additional surfaces or gaps in the 3D reconstruction.

Therefore, in post-processing a continuous-time SLAM technique is applied to compute an improved vehicle trajectory, which optimizes point cloud consistency. Similar to the mapping approach described in Section 6.2.5, the algorithm described in [116] is adopted, which was developed in the context of mobile mapping with a spinning laser scanner. The approach does not require explicit tie points or feature descriptors.

In the case of multibeam sonar mapping, a single scan slice does not provide enough structure, such that individual scans can be matched with each other. A 2D scan slice does not sufficiently constrain a six DoF rigid registration. Therefore, the trajectory is first split into sections. These sub-maps are created with overlp so there is always a match between consecutive sections. For optimizing the trajectory, the approach of [116] makes no rigidity assumptions, except for the computation of the point correspondences. The algorithm uses a starting solution based on a globally consistent scan matching of the sub-maps [54]. After that, the full trajectory is optimized based on a semi-rigid matching as described in [116], such that the pose of every individual scan slice is improved and drift within sub-maps is compensated. The method has the advantage that it does not require a motion or sensor model, i.e., it works solely on the trajectory and the point measurements themselves. This makes it flexible and allows application to different 3D sensor types. A challenge with multibeam data is that the resolution across track is typically much higher than along track. This often results in the problem that during scan matching individual scan slices are aligned instead of the geometric structure. Here, this problem is circumvented by applying an Octree-based sub-sampling to the sub-maps, such that the resolution is more homogeneous.

An important issue in continuous-time SLAM is the search for closest point pairs. In order to yield a more sparse graph, a time threshold for the point pairs is applied, i.e., scans are only matched if they were recorded at least a certain number of time steps away. This way, consecutive sub-scans are not matched with each other, which also improves run-time. In addition, a distance

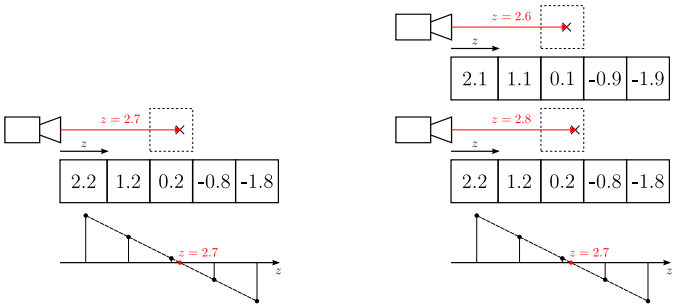


Fig. 7.12: Example of implicitly storing surfaces using the signed distance. Each edge of the voxel cells has unit length. The surface is recovered by searching the zero crossing along a ray.

threshold on the maximal allowed point-to-point distance is applied.

7.2.4 Signed Distance Function Based Mapping

All sonar scans are integrated into a SDF voxel model based on the optimized trajectory solution. An example of a SDF model is shown in the left image Fig. 7.12. The signed distance is positive outside the surface and negative inside the surface. The value is zero for points on the surface. The signed distances are stored with respect to the centers of the voxel cells. Although the voxel grid in the depicted example has only a low resolution (each voxel cell has a size of unit length), the exact position of the surface is still computed by searching for the zero crossing along the z-axis. Additionally, interpolation between neighboring cells is performed to get more accurate results.

The right image of Fig. 7.12 shows the integration of two different measurements of the same surface into the signed distance map. After the integration using averaging of the cell values, the surface position has changed to the average of the two measurements as well.

The signed distance measurement $d(\mathbf{v})$ for a voxel with center $\mathbf{v}$ is computed as follows

$$d(\mathbf{v}) = m - ||\mathbf{p} - \mathbf{v}|| \quad , \tag{7.1}$$

where $\mathbf{p}$ is the sensor position and m is the distance measurement of the sensor. Multiple measurements of the same voxel cell are integrated based on a weighting function f. This way, noise cancels out over multiple observations. Each voxel cell stores the signed distance $s(\mathbf{v})$ and the weight $w(\mathbf{v})$. To integrate a new measurement $d(\mathbf{v})$ at iteration $k+1$ the weighted average is computed:

$$s_{k+1}(\mathbf{v}) = \frac{w_k(\mathbf{v})s_k(\mathbf{v}) + f d_{k+1}(\mathbf{v})}{w_k(\mathbf{v}) + f} \quad , \tag{7.2}$$

where f is a weight assigned to the new measurement. The signed distance is truncated to the interval $[s_{\min}; s_{\max}]$. Since no accurate noise model of the sonar sensor is available, uniform

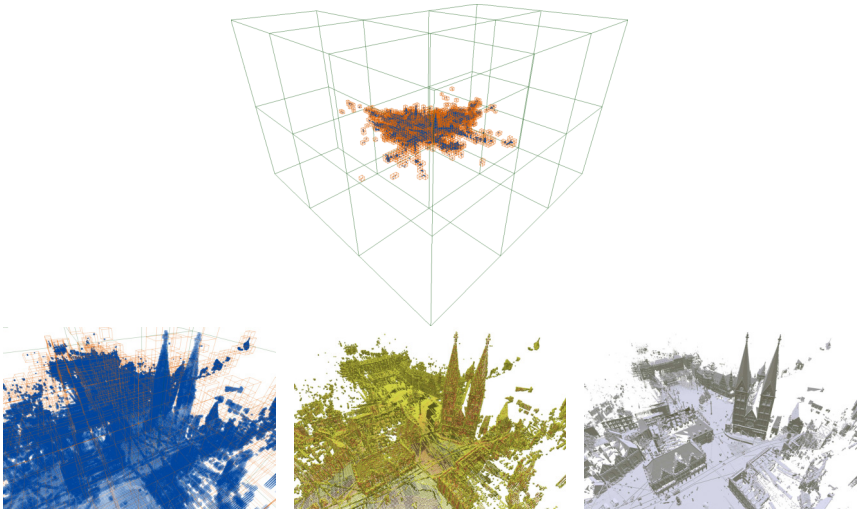


Fig. 7.13: Example of the OpenVDB tree using the Bremen City data set [53, p. 57]. Top: Visualization of the tree topology. Only populated cells are visualized. Bottom left: Detail view of the tree topology. Bottom middle: Visualization of surface points. Bottom right: Extracted mesh from the SDF.

weights ($f = 1$) are employed. The weight is updated by

$$w_{k+1}(\mathbf{v}) = \min(w_k(\mathbf{v}) + f, w_{\max}) \ , \tag{7.3}$$

where $w_{\max}$ is the maximum weight. Since the environment changes over time, the maximum weight needs to be limited. Otherwise, new measurements will not result in an update of the voxel cells and a changing environment cannot be modeled. For the experiments in this work $w_{\max} = 20$ is chosen.

SDF-based mapping is not entirely robust to coarse outliers. Noisy surfaces are only smoothed if the individual measurements lie within a certain band, which is determined by the penetration depths $D_{\min}$ and $D_{\max}$ of the TSDF. Underwater sonar sensors typically exhibit some amount of coarse outliers. Measurement points that lie outside the truncation thresholds are integrated as additional surfaces. To address this problem, a large truncation threshold of 1.5 m is selected. This limits the minimum thickness of objects that are represented by the SDF model. However, in the particular case of the submerged inland mine, this is not an issue because we only want to represent a single surface of the mine floor. To remove erroneous integrated surfaces, the SDF voxels are filtered based on the weight. This is based on the assumption that voxels representing real surfaces carry a higher weight, i.e., are observed more often than voxels filled from measurement outliers.

For modeling the mine a voxel resolution of 10 cm is chosen. This means the TSDF space of

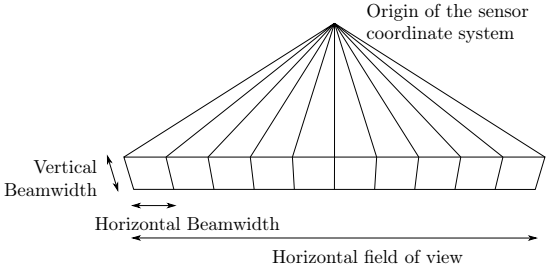


Fig. 7.14: Projection of multibeam sonar simplified described by a perspective line camera model.

the entire mine has a size in the order of a billion voxels. In order to store large maps with low memory consumption, the free space needs to be encoded efficiently. Methods for storing the entire TSDF are voxel hashing [243] or octree data structures [309]. For this work, the B-tree based data structure of OpenVDB [232] is used to store the complete sparse TSDF grid. The tree has constant depth, which allows constant time local and random traversals. A three-level tree with branching factors decreasing closer to the leaves is applied. Fig. 7.13 shows an example of the employed tree structure of OpenVDB.

To integrate the multibeam data in the TSDF, the work at hand follows the generalized sensor fusion approach proposed by [217]. The multibeam is modeled as a perspective line sensor with a vertical beam width of 1.5 deg, which is depicted in Fig. 7.14. Individual voxel cells within the measurement range are then updated based on back projection using this sensor model. Finally, from the scalar TSDF grid a surface mesh or 2.5D Digital Elevation Model (DEM) is extracted for visualization.

7.3 Results

In this section, the underwater mapping results of the Bejanca mine using the Autonomous Surface Vehicle (ASV) ROAZ [35, 120] are described. Moreover, we look at the model of the Silvermines site acquired with multibeam sonar using the HROV during the second ¡VAMOS! field trials. The results are mostly qualitative since no accurate ground truth is available.

7.3.1 Multibeam Sonar Survey of the Bejanca Mine

The robotic boat used for the mine mapping is the ASV ROAZ [120]. It is a 4m long twin hull robotic vehicle with electric propulsion and autonomous navigation and control, see Fig. 7.15. For bathymetric mapping, it is equipped with an Imagenex Delta T multibeam profiling sonar, which has a fan angle of 120 deg and a maximum range of 100 m. The experiments were conducted with a resolution of 480 beams and a beam width of 1.5 deg. Sonar data was recorded at 10 Hz.

For positioning and localization of the vehicle, a precision GNSS unit with RTK differential corrections and a fiber optic based INS were installed on the robotic boat. A high-precision

Fig. 7.15: ROAZ surface vehicle [120] at the Bejanca mine site (Source: [5]).

localization solution is later obtained by post-processing the raw INS data in combination with the raw GNSS data. The post-processing step is performed using the Inertial Explorer software [249], where all raw GNSS observations are processed in RTK and integrated with raw inertial measurements in a tightly coupled manner.

The multibeam sonar, GNSS antenna and INS were mounted rigidly to the same sensor bar. This was done to ensure that the relative positions and orientations stay consistent even during transport, which requires some disassembly. The sensor bar was mounted to the front of the vehicle with the multibeam sonar only a few centimeters below the water surface. The translation offsets between the individual sensors were measured manually. Rotational offsets between the INS and multibeam sonar are later estimated using a calibration routine. Even very small alignment errors introduce large inconsistencies in the resulting point cloud. To do this, a short trajectory segment, which is assumed to have minimal drift, is selected. Then, the rotational offset is optimized based on an error measurement, which determines point cloud quality similar to [294]. The error measurement is computed by splitting the trajectory into overlapping parts and calculating a point distance error based on the closest point correspondences. The rotational offset parameters are found by minimizing the error, and the result is verified on different trajectory segments.

All sensor measurements are recorded with GPS timestamps for correct data association. A trajectory was chosen, such that there is about 30 - 50% overlap between individual laps of the surface vehicle, and all parts of the mine are covered multiple times.

To demonstrate the continuous-time SLAM algorithm Fig. 7.16 shows results on a trajectory with significant drift of multiple meters. In this specific case, the GNSS signal was lost temporarily during data acquisition, which explains the large trajectory errors. The dataset consists of 7291 multibeam scans captured at 10 Hz. It was captured in 13 min, and the trajectory is approximately 757 m long. The plots on the left in Fig. 7.16 depict the initial GNSS/INS trajectory as a black dashed line and the optimized trajectory as a red continuous line. The x/y-plane is aligned

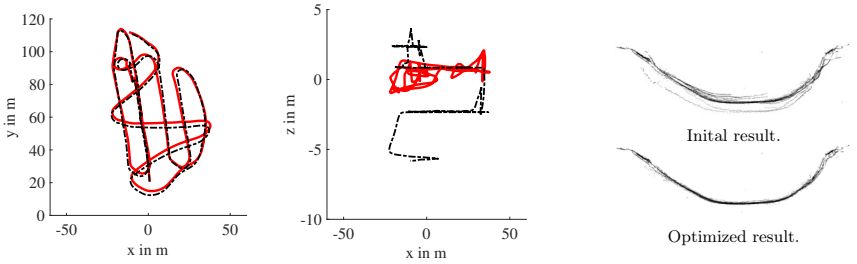


Fig. 7.16: Trajectory and cross-section of the resulting point cloud for a data set with significant drift. Left: Initial (black dashed line) and optimized (red line) trajectory of the sensor system. Right: Cross-section of the initial and optimized point cloud in the x/z-plane.

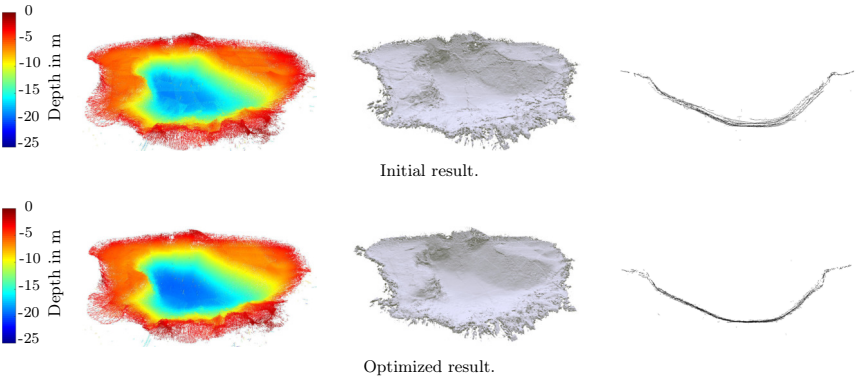


Fig. 7.17: Initial result using GNSS/INS trajectory (top) and optimized result (bottom). From left to right: 3D point cloud, surface mesh extracted from SDF model, and a cross-section of the point cloud.

parallel to the water surface. We observe that SLAM converges to a solution that puts the sensor poses closer to a planar motion as expected for a surface vehicle. Please note that the algorithm does not impose any movement constraints or rely on a vehicle motion model. Cross sections of the resulting point cloud are displayed in the right images of Fig.7.16. Misalignment between multiple passes of the multibeam sonar is visible in the initial result. Point measurements align well using the improved trajectory estimate based on continuous-time SLAM.

Moreover, data captured with a good GNSS/INS result is further improved using the proposed techniques as depicted in Fig. 7.17. The top row of Fig. 7.17 shows the resulting point cloud using the GNSS/INS trajectory, while the bottom row of Fig. 7.17 shows the result using the optimized trajectory from continuous-time SLAM. The left images show the point cloud colored

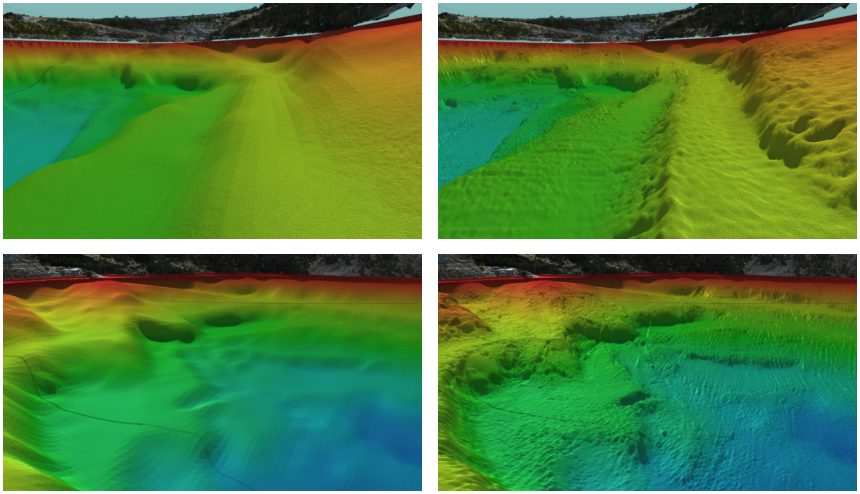


Fig. 7.18: Bathymetry of the Silvermines flooded opencast mine visualized in the VR system. Left: Pre-survey created by a surveyor. Right: Updated survey created using multibeam sonar.

by depth. This dataset comprises 12786 multibeam scans captured at 10 Hz. It was captured in 22 min, and the trajectory is approximately 1567 m long. Especially at the bottom of the mine, it is visible that the multibeam measurements are more consistent in the optimized results. This is more clearly visible in the cross sections of the point clouds presented in the right images of Fig. 7.17.

Consequently, the extracted mesh from the SDF representation using the optimized continuous-time SLAM solution exhibits smoother surfaces than the initial result. The mesh models are depicted in the middle images of Fig. 7.17. Despite the noise of the measurements, a smooth surface is extracted if a sufficient amount of repeated observations are available.

The borders of the mine show holes in the mesh. This is a result of the irregular and low point density of the sonar measurements due to limited coverage close to the borders of the mine. Since this is undesirable, the holes are later interpolated for display in the VR system.

7.3.2 Mine Model of the Silvermines Site

For the Silvermines site, the bathymetry model covers an area of approximately 450 m × 280 m. Fig. 7.18 shows results of the underwater bathymetry created using data from multibeam sonar surveys of the HROV compared to results from a pre-survey created by a surveyor using a single beam echosounder. We observe that the updated bathymetry includes more detail than the initial pre-survey of the site.

Fig. 7.19 shows the difference between a pre-survey created by a surveyor and the bathymetry

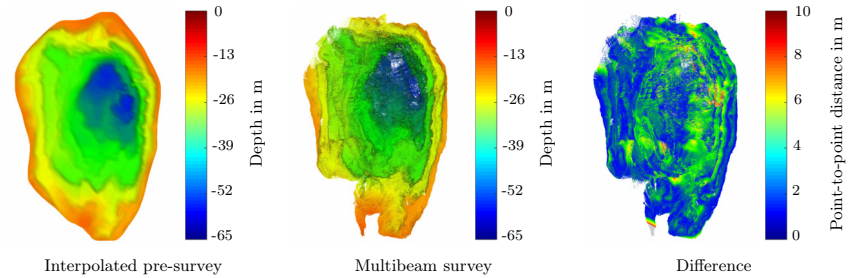


Fig. 7.19: Bathymetry of the Silvermines site. Left: Interpolated pre-survey colored by depth. Middle: Multibeam survey created in the ¡VAMOS! project colored by depth. Right: Difference between pre-survey and multibeam survey colored by point-to-point distance.

acquired during the field trials of the ¡VAMOS! project. The pre-survey was created using a single beam echosounder from a boat. It was interpolated to extract a dense point cloud of the mine. The multibeam survey was acquired on two different days using the EVA HROV.

The difference computed by the point-to-point distance between the two models is visualized in the right image of Fig. 7.19. The overall geometry agrees well between the pre-survey and the updated model. Especially at edges, higher deviations are observed, which is expected due to the higher resolution of the multibeam survey.

For the field trials, the system is sufficiently accurate to map newly extracted bathymetric surfaces with processing times that are adequate to keep up with mining progress. The overall positioning accuracy was sufficient for vessel maneuvering and driving of the mining vehicle.

7.4 Related Work

Bathymetric data gathering of larger water bodies is often performed using acoustic sensors, such as single beam or multibeam sonar [96, 311]. Typically, in industrial applications such as underwater mining operations, water turbidity is high and the visibility with optical sensors and cameras is very low. In these applications, sonar sensors have the advantage of providing a significantly higher measurement range in water, and measurements are also possible in very turbid environments.

Underwater Mapping For surveys from a surface vehicle, the motion is compensated using an attitude reference system and GNSS. For an underwater vehicle, it is more difficult since a global reference through a positioning system, such as USBL, is not always available. Different SLAM techniques have been proposed to improve underwater surveys.

In man-made structured environments, such as harbors, feature-based SLAM approaches have been proven to be effective. Typically employed features include planar patches [253, 260] and line features [268].

An example of a feature-less approach is the algorithm described by [42]. They use Particle Filter based SLAM to create a 2.5D point cloud of the seafloor. Individual particles are weighted based on how well the multibeam measurements agree with the global elevation map. Loop closures are detected using a Gaussian process regression of previous sonar beam observations. This allows to detect loop closures with minimal overlap and enforces consistency of neighboring map borders even if there is no overlap.

The approach of [278] divides the terrain map into smaller sub-maps that are assumed to be error-free. Overlapping sub-maps are first coarsely aligned using cross-correlation, and then the ICP algorithm is used for fine registration. The relative pose measurements are then used to constrain an Extended Kalman Filter (EKF) based mapping procedure further. In [49] this approach is extended with a Graph SLAM based framework to improve the full trajectory. In [254] a coarse-to-fine scan matching technique using ICP is proposed, which takes point measurement uncertainties into account during submap registration. Furthermore, spectral registration methods [72] for the registration of noisy sonar scans or probabilistic scan matching techniques [118] that weight point observations by the uncertainty based on a sonar sensor model are described in the literature.

Map Representations For many bathymetry applications 2.5D DEMs are created. More recently, creating dense surface models from sonar imaging has become of interest. However, the significant amount of noise in acoustic measurements makes it challenging to extract surface meshes directly from the point cloud. Therefore, for reconstructing meshes from noisy data, often the point measurements are integrated into an implicit surface description [161] or robust local surface descriptors are fitted to the 3D point cloud [78].

Our work follows this direction and uses an SDF voxel map [92] as an intermediate implicit surface model to create a more noise-free representation. This representation became popular in the robotic mapping community with KinectFusion [168,239], which demonstrated excellent real-time 3D reconstruction and tracking results. SDF have the advantage that the weighting scheme describes noise models during measurement update and different sensors are integrated based on a geometric projection model [217]. Techniques based on SDF were successfully applied for the reconstruction of underground mine shafts [183, 184] and aerial mapping using drones [267].

A disadvantage of the initial KinectFusion implementation is the large memory consumption due to the dense storage of the voxel map. Kintinuous [341] resolved this issue by keeping only a local volume in the memory. Parts of the map that drop out of this volume are saved memory efficiently by triangularization.

OpenVDB [232], which is used in this work to store the SDF map memory efficiently, has recently become more popular for robotic mapping [46, 326, 330]. Point density and resolution in object space strongly depend on the measurement distance for sonar sensors. Multiresolution SDF models might be a way to represent different levels of detail efficiently [284].

3D User Interfaces for Teleoperation The combination of underwater mapping technology, especially 3D reconstruction techniques using photogrammetr, and visualization in a VR environment has become popular in the field of cultural heritage conservation and underwater archaeology [108]. Underwater cultural heritage is typically difficult to access, and exploitation of

underwater archaeological sites for a large-scale public audience and tourism is not sustainable. This sparked an interest in creating immersive experiences and virtual tours using VR [154]. Increasingly, there is an interest not only to show the 3D underwater environment but to augment it with additional educational and archaeological information [69, 303].

Similarly, VR has been applied to assist teleoperation of mobile robots, remote control of industrial manipulators and the piloting of underwater robots [199]. The difference here is that for a teleoperation scenario, live sensor data needs to be fed back into the system to update the virtual environment in order to create a faithful representation of the current environment of the remote system. In combination with simulation of the environment, VR is also used in pilot training of ROV for underwater operations [200]. For dexterous tasks, such as controlling an underwater manipulator and effectors, haptic user interfaces support the teleoperation [50]. As an underlying technology, game engines, such as the Unity 3D cross-platform game engine [322], are increasingly used in research and industry for serious applications, e.g., simulation and teleoperation of robots using VR [87].

7.5 Discussion

This chapter described techniques for creating an above-the-water and underwater survey of a mine site. The above-the-water model is created from laser scanning and camera images. It is used as a static map, which is delivered either as high-resolution point cloud or textured mesh in order to achieve a better orientation in the 3D scene of the remote operator. The underwater model is initially created from a pre-survey and then regularly updated from multibeam sonar data. For the underwater model SDF grid maps are used to update the model incrementally with new data. The 3D maps are used in the ¡VAMOS! project to create a virtual reality 3D replica of the real environment to assist the teleoperation of an underwater mining vehicle. This model is used for immersive data visualization of the mining operations, for planning during development and operations during the testing phase.

A 3D model of the operations is valuable to effectively monitor the events and the mining process below the water surface. Moreover, it enables the use of a smaller and cheaper sensor kit since only the areas where change is expected need to be monitored and updated continuously with surveying equipment while the full context of the mine site is still visualized to the human operator.

Chapter 8

Conclusions and Future Work

In this last chapter, we review the main theoretical and experimental results achieved in this work and summarize the technological outcomes. Moreover, we wish to provide an outlook for possible directions to follow for further investigation and research.

8.1 Summary

Scientific and industrial operations, such as monitoring coral growth, pipeline surveying or defect assessment in offshore structures, require precise subsea metrology. While no single underwater measurement technology covers all use cases, optical underwater scanners provide advantages in terms of achievable resolution and accuracy.

This dissertation presents an underwater laser scanning system and the algorithms for creating dense 3D scans in water. The main focus lies on the calibration of the underwater camera using a physical-geometric model, the calibration of the laser projector, the image processing and 3D reconstruction. Since in-situ calibration in water is complex and time-consuming, the challenge of transferring an in-air calibration of the scanner to water without re-calibration, as well as self-calibration techniques, are investigated. The system was successfully applied in different configurations for static scanning and mobile mapping. 3D maps are useful to the operator during the remote control of underwater vehicles and enable offshore inspection and surveying tasks.

The first part of this thesis reviews optical and acoustic modalities for underwater sensing. The main challenges of applying optical measurement techniques are identified as the visibility constraints due to the attenuation of light in the water and the refraction effects present in multimedia imaging. This work shows that if the refraction effects of a flat port camera are modeled using polynomial distortion models, high errors are observed for close-range measurements. Since turbidity is an issue for optical metrology, 3D scanning is typically performed at short distances. Therefore, for the highest accuracy, it is necessary to model the refraction effects explicitly. While Snell's law and its application for ray tracing is well understood, the challenge of calibration lies in the precise estimation of the required camera and housing parameters. This work addresses this problem with an optimization-based calibration framework using a physical-geometric camera model derived from an analytical formulation of a ray tracing projection model. Moreover, the

results of the proposed calibration approach are compared with a calibration performed solely in air and a calibration based on the standard camera parameter set using Brown's distortion model. An in-air calibration is successfully transferred to underwater imaging if the refraction at the housing is compensated. However, some type of in-situ calibration and validation is recommended to achieve the highest accuracy. Additionally, the effects of the test field and fiducial marker system on the calibration result are analyzed. The insights recommend a calibration approach using a 3D test field with circle marks. Circle targets show better feature localization accuracy if the images are degraded by noise and blur, which is often the case in underwater applications.

Based on these findings, a triangulation-based underwater scanner using a cross-line laser projector and the associated data processing algorithms are developed. The projection pattern was chosen, such that overlapping scan lines are acquired during movement of the scanner. In order to scan larger volumes, the underwater scanner is configured with a motorized yaw axis. The contributions include approaches for calibrating the rotational axis and laser curve parameters. In order to achieve robust laser line extraction under difficult conditions, line constraints in the image processing pipeline are applied. The technology was successfully applied to collect data in various underwater scenarios. The measurement quality is validated in experiments in a test tank using reference objects, and the different calibration processes are compared. As a benchmark, the work at hand investigates surfaces and distance measurement errors in a small volume using sphere targets, length measurement errors of scale bars and the flatness error of a reference plane. Depending on the measurement distance, errors in the range of one millimeter up to less than one centimeter in clear water conditions are observed. Using a staircase pattern, the experiments find a measurement resolution in the millimeter range, which is consistent with the predicted theoretical resolution based on subpixel feature extraction.

Besides scanning underwater structures, the 3D acquisition of semi-submerged structures using the system is investigated. In the lab, scanning of structures partially covered by water is demonstrated with the line laser scanner moved by an industrial manipulator. Using a semi-automatic refractive correction approach, which extracts the parameters of the flat water surface from the imaging data, the errors in the scans through the water surface are effectively compensated.

The work at hand exploits coplanarity and orthogonality constraints to achieve self-calibrating structured light. The proposed method is based on the assumption of a static scene relative to the camera frame. A handheld cross-line laser projector with an unknown position and orientation is freely moved with respect to the camera and projects two laser curves on the scene. The intersections between consecutive projections of laser lines are exploited for solving the parameters of the projection planes. After that an euclidean solution is obtained through non-linear optimization by introducing orthogonality constraints on the two perpendicular laser planes of the cross-line laser projector. This work leverages heuristics and outlier rejection to improve the results and make the approach more robust and applicable to real-world data.

A limitation of optical scanning in water is the reduced range and small measurement volume compared to sonar. Consequently, underwater laser scanning is extended with simultaneous trajectory estimation to apply it for mobile mapping. Initially, this is addressed using Global Navigation Satellite System (GNSS) and Inertial Navigation System (INS), which is only applicable in surface water because the GNSS antennas need to be above the water. While this is a feasible direction, the accuracy of the GNSS-INS trajectory is low compared to the point

measurement accuracy of the scanner. Therefore, the results are improved using Visual-Inertial Odometry (VIO) for trajectory estimation. In the UWSensor project, a new sensor system based on active stereo using fringe projection was developed. This sensor delivers dense reconstruction at high update rates of approximately a 1 m×1 m×1 m measurement volume. This technology was successfully applied for mobile underwater mapping of a pipe structure using VIO for motion compensation. The results are further improved by registration of the point cloud data.

Finally, in the Viable Alternative Mine Operating System (¡VAMOS!) project 3D underwater mapping was applied to support a remotely controlled underwater mining system. Individual sonar scans are fused based on a volumetric map representation using signed distance functions. The insights during the field trials showed that a 3D model of the operations is valuable for effective monitoring and awareness of the situation below the water surface.

In summary, the experimental part of this thesis substantiates that the proposed underwater laser scanning system provides an accurate technique for acquiring detailed 3D scans in water. Overall, we believe that the insights of this work with respect to refractive calibration and self-calibration to mitigate residual measurement errors and the proposed mobile mapping approaches improve the applicability. Mobile mapping with optical scanners in water without additional infrastructure opens up the system for diverse applications in other fields, such as enabling cost-effective real-world inspection and metrology tasks in archaeology, the offshore industry and marine science.

8.2 Technological Outcomes

Throughout this thesis, several hardware and software systems were developed. This includes two underwater scanning systems with cross-line laser projectors that have been successfully tested in different scenarios (cf. Chapter 4 and Chapter 6). One system is configured with a motorized yaw axis to capture scans from a tripod, and a second system with a larger baseline for scanning from a moving platform. This includes the design of the laser projector, the engineering of the mechanical structures and integration in the underwater housings. Moreover, the necessary power electronics, Light-Emiting Diode (LED) flash and laser drivers, trigger circuits and control electronics are part of the outcomes of this work. For these scanners, the control software, microcontroller firmware and algorithms for image processing and 3D reconstruction were implemented. In low turbidity conditions, the system is able to create dense, high-resolution scans in water with errors in the millimeter to sub-1 cm range depending on the measurement distance (cf. Section 4.6). With the optical underwater scanner developed in the UWSensor project (cf. Sec 6.2), underwater mobile mapping was achieved. The results include the software and hardware design of the Inertial Measurement Unit (IMU) system and real-time motion compensation based on visual-inertial odometry.

An additional outcome of this thesis is the developed software framework for underwater scanning, which implements refractive calibration of underwater cameras and laser projectors, self-calibrating structured light, image processing, 3D reconstruction, mapping and motion compensation. For calibration and validation, different calibration fixtures (cf. Section 3.5) and test artifacts (cf. Section 4.6) were designed as well as infrastructure for an underwater test environment [24].

8.3 Outlook

The results presented in this thesis open several opportunities for further research. Future work will address technological and theoretical challenges to make optical underwater scanners more useful and applicable for real-world inspection and metrology tasks. Furthermore, we point out aspects for further investigation, which are considered most promising or interesting to the scientific community.

Range and Visibility Limitations By far, the biggest limitations of applying optical 3D imaging technologies in water are the range and visibility constraints. A lot of work remains to be done to improve the robustness to water turbidity. Increasing the output power of the projector for active imaging works only up to a certain point. Backscatter at particles in the water column in front of the sensor causes noise that masks the useful signal. Range gating is effectively used to reduce backscattering in active underwater imaging [39, 335]. More recently, affordable CMOS sensors with fast-shutter durations of down to ten microseconds became available for commercial applications. These sensors have been successfully customized for underwater applications in combination with pulse lasers to achieve gating with delay steps in the nanoseconds range [212, 273]. This allows to work at long ranges and directly provide Time-of-Flight (ToF) measurements. However, the range-gated signal delivers only a depth resolution in the centimeter range. Consequently, combining the increased range of gated imaging with triangulation-based depth estimation is interesting. Gated imaging has already been successfully applied to line laser scanning in water [93]. The newer fast CMOS sensors might enable active stereo sensors with pattern projection that capture dense 3D data with increased range and high precision.

Another interesting direction is the combination of different sensing modalities as discussed in Section 2.4. Multi-sensor setups combining sonar and camera sensors leverage the benefits of the long range and robustness to turbidity of acoustic ranging with the precision and resolution of optical technologies. Today, this is often applied using a pre-survey of the seabed using multibeam sonar. Then, selected regions of interest are captured in high detail using optical sensing. The data of the different sensors is then fused in post-processing. Volumetric representations, such as Signed Distance Function (SDF) have been successfully applied to multi-sensor fusion. While sensor models allow a weighting of the observations based on the expected noise of the sensors, it is challenging to apply these methods to sensing modalities with very different properties. Sonar data features strong outliers, high noise and sparse data compared to optical imaging. Multiresolution SDF models might be a way to represent different levels of detail efficiently [284]. Therefore, future research is required to adapt these methods to the characteristics of underwater sensors.

Scanning Through Complex Media Interfaces In Chapter 4, scanning through a planar air-water interface with a moving sensor is demonstrated. While this works well in a laboratory setting, in real-world applications the water surface is never perfectly still due to surface waves generated, for example, by wind. Hence, it is interesting to investigate how accurate measurements are achieved in more complex situations. Especially for aerial imaging or bathymetric Light Detection and Ranging (LIDAR), measurement results are affected by refraction. If we

have perfect knowledge of the water surface, it is possible to correct for the refraction. Therefore, the problem is reduced to the problem of estimating the surface of the media interface. For airborne LIDAR this is partially recovered by evaluating the signal returns from the water surface and water bottom in full-waveform analysis [206]. While modeling water waves is well understood for physics simulations, estimating the parameters of wave models from imaging data remains a challenging problem.

Motion Compensation of Multi-shot Structured Light A general limitation of multi-shot structured light based on fringe projection is the assumption of no motion between the individual recordings of a single sequence. This is violated if the sensor movement speed is too high compared to the duration of the image sequence. Additionally, in underwater imaging the required exposure times are typically higher compared to in air because of light absorption. Therefore, the motion needs to be compensated to enable scanning at larger sensor movement speeds. In the literature, approaches applying correlation-based image registration [336,342] are described. Most methods are based on assumptions like a linear motion along a single axis [191]. In [15] we demonstrated motion compensation of an underwater active stereo sensor by shifting the images based on a sensor motion estimate from VIO. Here, only a global linear translation is considered. Future work focuses on estimating a more accurate dense image warping function for compensating the sensor's motion by exploiting the pose estimates from VIO and overlapping previous 3D structured light scans.

Joint Alignment of 3D Scans and Color Images Optical underwater scanners provide a limited measurement range and field of view. Therefore, the 3D data does not always provide enough structure to constrain the registration fully. In the work at hand, motion is initially estimated using VIO, and a second step performs registration and loop closing based on the 3D data. While this exploits both color and 3D information, it does this in sequential steps and not in an integrated approach. Further research is necessary for achieving reliable registration in all situations, e.g., by joint alignment of 3D scans and color images.

Self-calibration and Quality Control Measurement accuracy for inspection tasks depends on precise calibration. Performing in-situ calibration is time-consuming and difficult. Some results of calibrating in air and transferring the calibration to imaging in water are presented in Chapter 3. However, for precise measurements in-situ calibration is preferred. Moreover, the problem remains that at high water pressures the underwater housings and structure deform, which affects some of the calibration parameters. Self-calibration using multi-view constraints or projection-based constraints, as discussed in Chapter 5, provide additional information. However, this typically does not allow to estimate precise scale. Therefore, further investigations are aspired to achieve in-situ calibration with minimal infrastructure. Moreover, approaches for quality control of the calibration parameters are necessary to achieve traceable measurement results.

Low-cost Technologies as a Driver for Data Set Acquisition A challenge in underwater imaging technologies is the limited availability of data sets to the scientific community. Especially

for machine learning based approaches, it is often difficult to obtain suitable training data. Large data sets are difficult and expensive to collect. Images acquired in subsea metrology campaigns or the data sets used for training sonar target detection in the industry are typically well-guarded secrets and are not available to the scientific community. The availability of a wide variety of low-cost waterproof action cameras, e.g., GoPro HERO, and affordable Remotely Operated Vehicle (ROV) platforms, such as BlueRobotic's BlueROV2, have democratized access to underwater imaging data. We hope the work at hand contributes to this trend towards affordable underwater sensor systems.

Appendix A

Acronyms

ADCP	Acoustic Doppler Current Profiler. 30
ALS	Airborne Laser Scanning. 30
ASV	Autonomous Surface Vehicle. 185
AUV	Autonomous Underwater Vehicle. 3, 33, 149
BA	Bundle Adjustment. 8, 23, 40, 41, 46, 54, 62, 64–66, 85, 203, 204
CAD	Computer-Aided Design. 58, 61, 153, 181
CMUT	Capacitive Micromachined Ultrasonic Transducer. 33
CNN	Convolutional Neural Networks. 16
DBAT	Damped Bundle Adjustment Toolbox. 62, 64
DEM	Digital Elevation Model. 185, 190
DOE	Diffractive Optical Element. 27, 29, 142
DoF	Degrees of Freedom. 56, 97, 125, 127, 129, 135, 138, 149, 151, 152, 175, 182
DVL	Doppler Velocity Log. 30, 149, 173, 175
EKF	Extended Kalman Filter. 190
FOG	Fiber Optic Gyro. 172, 175
FoV	Field of View. 32, 36, 119, 171
GAN	Generative Adversarial Network. 16
GEBECO	General Bathymetric Chart of the Oceans. 1, 2

GNSS	Global Navigation Satellite System. 10, 149–156, 173–175, 180, 181, 185–187, 189, 194
HMI	Human-Machine Interface. 176, 178
HROV	Hybrid Remotely Operated Vehicle. 170, 173–178, 180, 181, 185, 188, 189
ICP	Iterative Closest Point. 100, 102, 103, 121, 145, 149, 160, 161, 181, 190
ID	Identification Number. 23, 61, 62, 69, 158
IMU	Inertial Measurement Unit. 10, 26, 149–153, 156–160, 172, 195
INS	Inertial Navigation System. 152, 154, 156, 173, 175, 185–187, 194
IR	Infrared. 2
iUSBL	inverted Ultra Short Baseline. 175
LAN	Local Area Network. 176
LARV	Launch and Recovery Vessel. 173–176, 181
LED	Light-Emiting Diode. 2, 27, 150, 156, 157, 159, 172, 173, 195
LIDAR	Light Detection and Ranging. 28, 29, 33, 100, 145, 146, 196, 197
MBS	Multibeam Sonar. 31, 32, 173
MEMS	Microelectromechanical System. 29, 151, 152
NOAA	National Oceanic and Atmospheric Administration. 1, 2
NTP	Network Time Protocol. 88, 181
PPS	Pulse-Per-Second. 181
RANSAC	Random Sample Consensus. 58, 60, 142
RMSE	Root Mean Squared Error. 57, 76, 83, 124, 143, 153
ROS	Robot Operating System. 88
ROV	Remotely Operated Vehicle. 3, 6, 22, 23, 25, 149, 157, 158, 170, 191, 198
RTK	Real-Time Kinematic. 149, 150, 152, 153, 173–175, 185, 186
SBL	Short Baseline. 173, 175, 181
SBS	Single Beam Sonar. 31
SDF	Signed Distance Function. 10, 181–184, 187, 188, 190, 191, 196

SfM	Structure from Motion. 1, 3, 17, 30, 31, 41, 42, 133, 149, 160
SGM	Semi-Global Matching. 24
SLAM	Simultaneous Localization and Mapping. 135, 149, 161, 167, 181, 182, 186–190
SLERP	Spherical Linear Interpolation. 161
SLR	Single-Lens Reflex camera. 181
SLS	Structured Light Scanner. 173
SNR	Signal-to-noise Ratio. 30
SSS	Side-scan Sonar. 31
SVD	Singular Value Decomposition. 58, 138
ToF	Time-of-Flight. 28–31, 109, 196
TSDF	Truncated Signed Distance Function. 181, 184, 185
USBL	Ultra Short Baseline. 172, 173, 175, 189
VIO	Visual-Inertial Odometry. 10, 150, 162–165, 195, 197
VR	Virtual Reality. 10, 169, 174, 176–178, 180–182, 188, 190, 191
WLAN	Wireless Local Area Network. 173, 174
¡VAMOS!	Viable Alternative Mine Operating System. 4, 5, 10, 32, 169–174, 176, 178–180, 185, 189, 191, 195

Appendix B

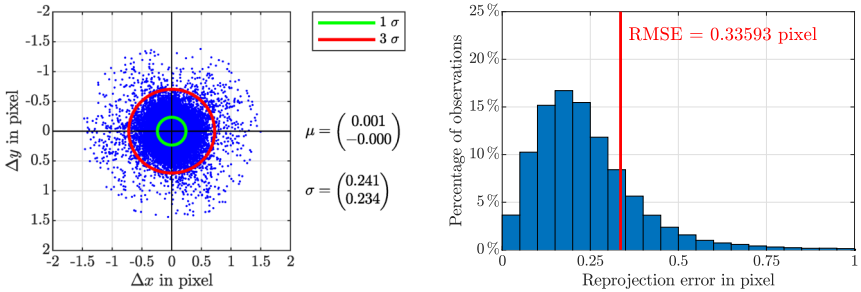
Additional Results

This appendix lists additional results and figures of the experiments described in this thesis. The additional material is mainly provided for completeness and comparison purposes.

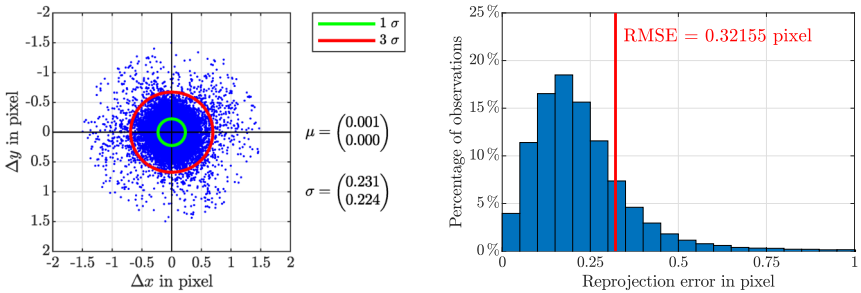
B.1 Comparison of Calibration Using Zhang's Method and Bundle Adjustment

This section provides additional results for the experiment described in Section 3.5.1. The results of a calibration using Zhang's method and Bundle Adjustment (BA) are compared on imprecise calibration targets. Each calibration is performed on a set of 560 images.

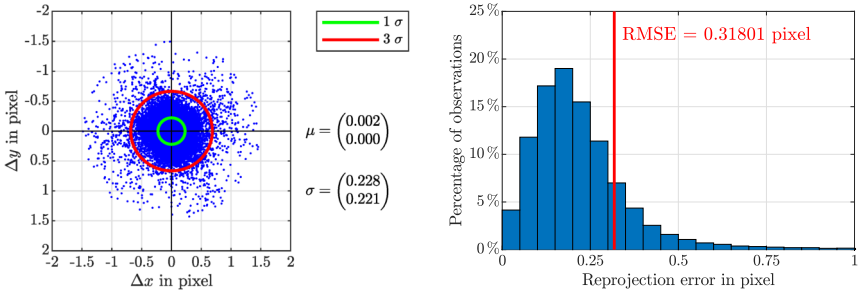
The difference of the RMS projective errors for Zhang's method compared to BA is small for the glass pattern with vinyl foil. Fig. B.1 shows the distribution of the residuals and the histogram of the residuals in image space. The residuals are computed on the same set of 560 images of the direct printing on glass target. The top row shows the result of the calibration using Zhang's method on the glass pattern with vinyl foil data, the middle row shows the result using BA on the glass pattern with vinyl foil data, and the bottom row shows the result of the calibration on the direct printing on glass data as a comparison. Here, the error using BA provides only marginally smaller residuals than the pattern manufactured using direct printing.



Calibration using the vinyl foil on glass sheet target and Zhang's method.



Calibration using the vinyl foil on glass sheet target and BA.



Calibration using direct printing on glass target and Zhang's method.

Fig. B.1: Distribution and histogram of the projective errors computed on 560 images of the vinyl foil target compared to the direct printing on glass target.

B.2 Comparison of Planar and 3D Structures for Calibration

This section provides the results of the comparison between a planar and 3D structure as described in Section 3.5.2 for the experiment carried out in air instead of water. The experiment was carried out with the same calibration structure as the experiment in water. The experimental procedure is the same as described for the evaluation conducted on the underwater data set. Similarly, the calibration is evaluated on a data set of 600 images. The employed camera is a Sony a6000 mirrorless camera with an APS-C size sensor and a 35mm lens. The images were downsampled to a resolution of 3000×2000 pixels.

The reprojection error computed on the planar patterns is very similar for both calibrations. The 2D calibration yields a RMS reprojection error of 0.389 pixels, while the 3D calibration produces a RMS reprojection error of 0.390 pixels. The reprojection error computed on the 3D structure is slightly lower for the 3D calibration. The 2D calibration results in a RMS reprojection error of 0.537 pixels, while the 3D calibration yields a RMS reprojection error of 0.523 pixels. In summary, the achieved reprojection error is comparable for the 2D and 3D calibration given the complete set of 600 calibration images.

Fig. B.2 shows a boxplot of the RMS reprojection errors for different numbers of images used for the calibration. The left image shows the RMS reprojection error computed on the planar pattern, and the right image shows the RMS reprojection error evaluated on the 3D pattern. Fig. B.3 visualizes the calibration results in relation to the size of the calibration set.

Similar to the results in water described in Section 3.5.2, faster convergence is observed for the calibration with the 3D structure. Moreover, fewer images are necessary for the 3D structure to achieve a low residual error on the complete data set.

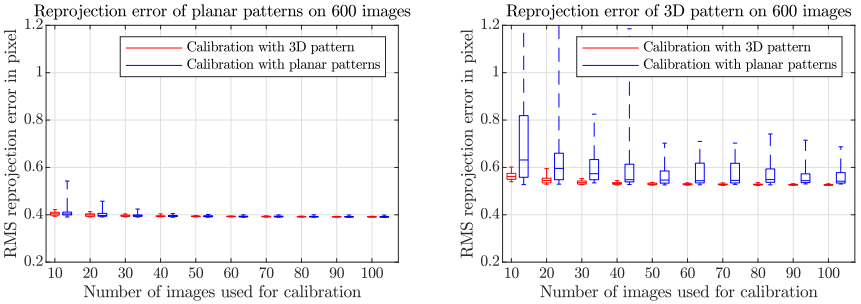


Fig. B.2: Comparison of the projective error of a calibration using planar or 3D targets.

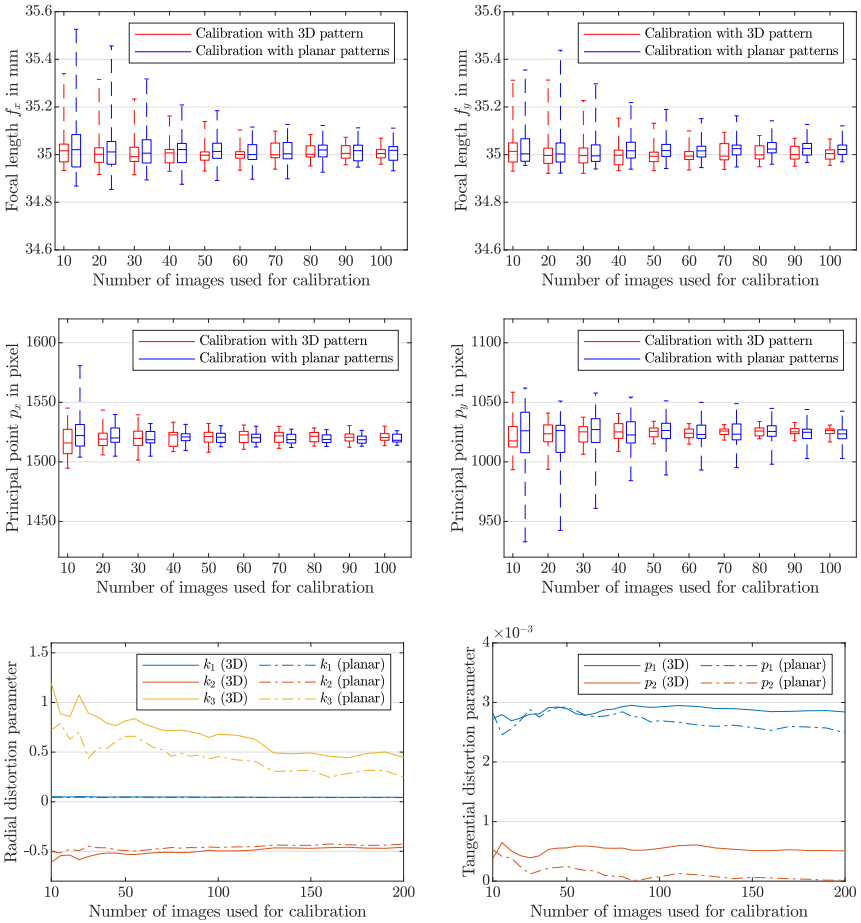


Fig. B.3: Comparison of the variation of the model parameters of a calibration using planar or 3D targets.

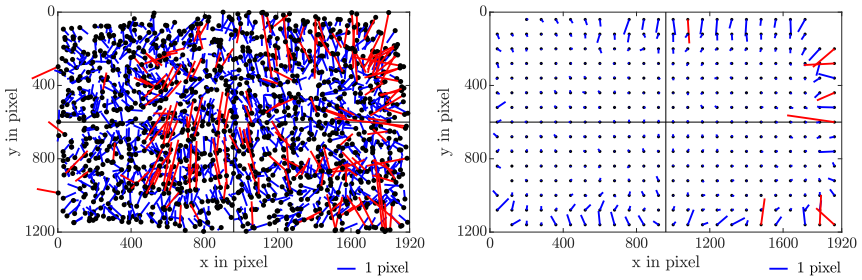
B.3 Comparison of A priori and Underwater Calibration

This section provides additional plots of the error distribution of the comparison between a priori calibration, refractive calibration and underwater calibration using the Brown model as described in Section 3.6.3. Three different calibrations are compared:

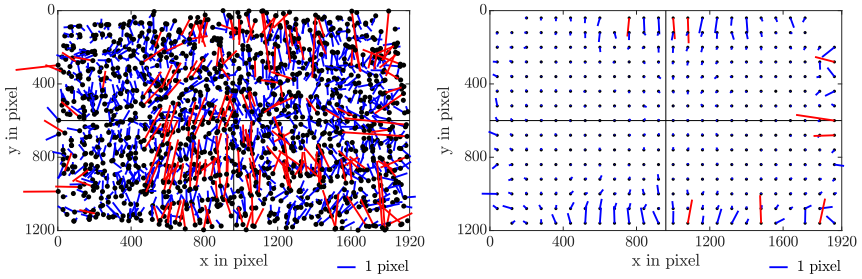
- an a priori calibration performed solely in air,
- an underwater calibration using the refractive model, and
- an underwater calibration using the Brown model.

Fig. B.4 shows the distribution of residuals in image space. The left image shows a sub-sampled plot of the largest residual vectors. Residuals with a norm of more than 1 pixel are plotted in red. The right image shows a plot of the mean residuals on a regular grid with a cell size of 80×80 pixels.

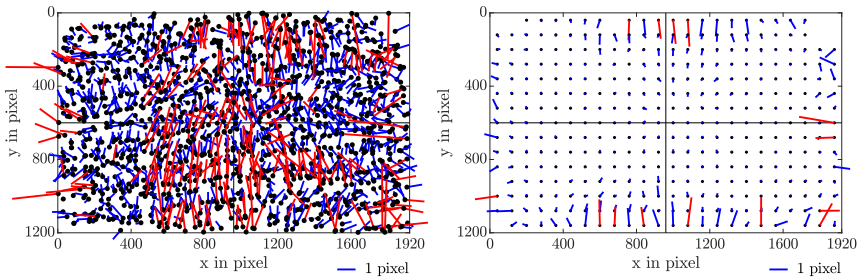
Fig. B.5 depicts the distribution of the residuals. The left image is a scatter plot of the x/y-components of the residual vector. The 1σ and 3σ ellipses of the standard deviation σ of the residuals is plotted in green and red color. The right image shows a scatter plot of the residuals along the radial direction from the image center.



A priori calibration using refractive model

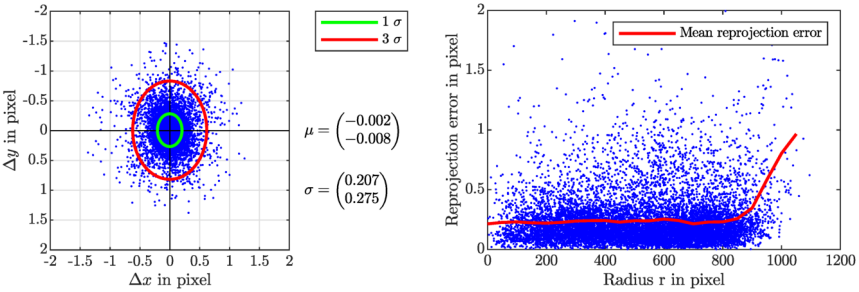


Underwater calibration using refractive model

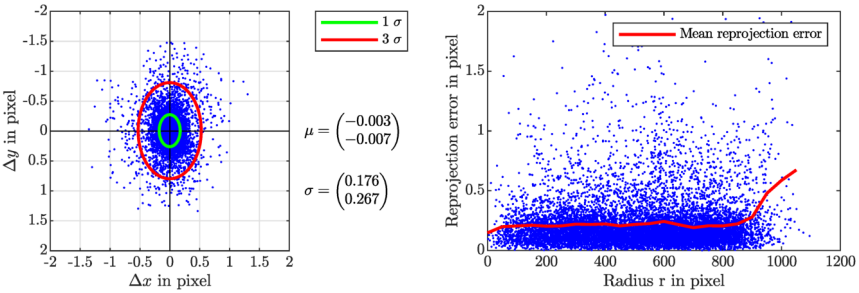


Underwater calibration using Brown model

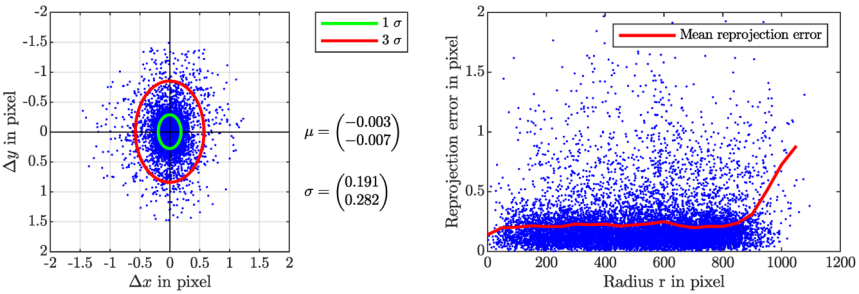
Fig. B.4: Distribution of residuals in image space computed for a validation set of 74 images. The lengths of the residual vectors are scaled by a factor of 100 for visualization purposes. Residuals with a norm of larger than 1 pixel are plotted in red. Left: Plot of sampled residuals in image space. For visualization purposes only the four largest residuals in every 80×80 pixels cell are shown. Right: Mean residuals on a regular grid with a cell size of 80×80 pixels.



A priori calibration using refractive model



Underwater calibration using refractive model



Underwater calibration using Brown model

Fig. B.5: Distribution of residuals computed for a validation set of 74 images. Left: Scatter plot of the residuals with mean and standard deviation. Right: Scatter plot of the reprojection errors dependent on the radial distance from the image center.

Bibliography

Own Work

[1] J. Almeida, A. Martins, C. Almeida, A. Dias, B. Matias, A. Ferreira, P. Jorge, R. Martins, M. Bleier, A. Nüchter, J. Pidgeon, S. Kapusniak, and E. Silva. Positioning. navigation and awareness of the ¡VAMOS! underwater robotic mining system. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1527–1533, Madrid, Spain, Oct. 2018. IEEE.

[2] T. Appelt, J. van der Lucht, M. Bleier, and A. Nüchter. Calibration and validation of the Intel T265 for visual localisation and tracking underwater. In *Proceedings of the XXIV ISPRS Congress (2021 edition)*, volume XLIII-B2-2021 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 635–641. Copernicus GmbH, June 2021.

[3] M. Bleier, C. Almeida, A. Ferreira, R. Pereira, B. Matias, J. Almeida, J. Pidgeon, J. van der Lucht, K. Schilling, A. Martins, E. Silva, and A. Nüchter. 3D underwater mine modelling in the ¡VAMOS! project. In *Proceedings of the 2nd ISPRS International Workshop Underwater 3D Recording & Modelling: A Tool for Modern Applications and CH Recording*, volume XLII-2/W10 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 39–44, Limassol, Cyprus, May 2019. Copernicus GmbH.

[4] M. Bleier, A. Dias, A. Ferreira, J. Pidgeon, J. Almeida, E. Silva, K. Schilling, and A. Nüchter. Real-time 3D mine modelling in the ¡VAMOS! project. In M. Buxton and J. Benndorf, editors, *Real Time Mining*, pages 91–102, Amsterdam, The Netherlands, Oct. 2017. Technische Universität Bergakademie Freiberg, Institut für Markscheidewesen und Geodäsie.

[5] M. Bleier, A. Dias, A. Ferreira, J. Pidgeon, J. M. Almeida, E. Silva, K. Schilling, and A. Nüchter. Signed distance function based surface reconstruction of a submerged inland mine using continuous-time SLAM. In *Proceedings of the 20th World Congress of the International Federation of Automatic Control (IFAC WC '17)*, Toulouse, France, July 2017.

[6] M. Bleier, C. Munkelt, M. Heinze, C. Bräuer-Burchardt, H. A. Lauterbach, J. van der Lucht, and A. Nüchter. Visuelle Odometrie und SLAM für die Bewegungskompensation und mobile Kartierung mit einem optischen 3D-Unterwassersensor. In T. Luhmann and C. Schumacher, editors, *Photogrammetrie Laserscanning Optische 3D-Messtechnik, Beiträge*

der Oldenburger 3D-Tage 2022, Jade Hochschule, pages 394–405. Wichmann Verlag, Feb. 2022.

[7] M. Bleier, C. Munkelt, M. Heinze, C. Bräuer-Burchardt, S. Marx, D. Fromme, C. Dettmer, H. A. Lauterbach, and A. Nüchter. UWSensor: Ein optisches Unterwasser-3D-Scannersystemfür mobile Kartierung. In Thomas P. Kersten and Nora Tilly, editor, *43. Wissenschaftlich-Technische Jahrestagung der DGPF in München – Publikationen der DGPF*, volume 31, pages 332–346, Munich, Germany, Mar. 2023. DGPF.

[8] M. Bleier and A. Nüchter. Low-cost 3D laser scanning in air or water using self-calibrating structured light. In *Proceedings of the 7th ISPRS International Workshop 3D-ARCH 2017: 3D Virtual Reconstruction and Visualization of Complex Architectures*, volume XLII-2/W3 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 105–112, Nafplio, Greece, Mar. 2017. Copernicus GmbH.

[9] M. Bleier and A. Nüchter. Towards robust self-calibration for handheld 3D line laser scanning. In *Proceedings of the 5th International Workshop LowCost 3D – Sensors, Algorithms, Applications*, volume XLII-2/W8 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 31–36, Hamburg, Germany, Nov. 2017. Copernicus GmbH.

[10] M. Bleier, J. van der Lucht, and A. Nüchter. SCOUT3D – an underwater laser scanning system for mobile mapping. In *Proceedings of Optical 3D Metrology 2019*, volume XLII-2/W18 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 13–18, Strasbourg, France, Dec. 2019. Copernicus GmbH.

[11] M. Bleier, J. van der Lucht, and A. Nüchter. Towards an underwater 3D laser scanning system for mobile mapping. In *Proceedings of the IEEE ICRA Workshop on Underwater Robotic Perception (ICRAURP '19)*, Montreal, Canada, May 2019.

[12] M. Bleier, J. van der Lucht, and A. Nüchter. SCOUT3D – ein Unterwasser-Laserscanner für mobile Kartierung. In *Photogrammetrie Laserscanning Optische 3D-Messtechnik, Beiträge der Oldenburger 3D-Tage 2020, Jade Hochschule*, pages 82–93, Feb. 2020.

[13] M. Bleier, J. van der Lucht, and A. Nüchter. SCOUT3D – ein Unterwasser-Laserscanner für mobile Kartierung. *Allgemeine Vermessungs-Nachrichten (AVN)*, 128(4):171–180, Aug. 2021.

[14] M. Bleier, J. van der Lucht, and A. Nüchter. Kinect backpack for rapid mobile indoor mapping. In *Proceedings of the XXIV ISPRS Congress (2022 edition)*, volume V-1-2022 of *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 121–127, Nice, France, June 2022. Copernicus GmbH.

[15] C. Bräuer-Burchardt, C. Munkelt, M. Bleier, M. Heinze, I. Gebhart, P. Kühmstedt, and G. Notni. A new sensor system for accurate 3D surface measurements and modeling of underwater objects. *Applied Sciences*, 12(9), Apr. 2022.

[16] C. Bräuer-Burchardt, C. Munkelt, M. Bleier, M. Heinze, I. Gebhart, P. Kühmstedt, and G. Notni. Underwater 3D scanning system for cultural heritage documentation. *Remote Sensing*, 15(7), 2023.

[17] S. Jörissen, M. Bleier, and A. Nüchter. Self-calibrated surface acquisition for integrated positioning verification in medical applications. In *Proceedings of the IS&T International Symposium on Electronic Imaging Science and Technology*, Electronic Imaging, pages 353–1–353–8, San Francisco, USA, Jan. 2019. Society for Imaging Science and Technology.

[18] S. Menninger, M. Bleier, and A. Nüchter. Entwicklung einer randomisierten Hough-Transformation zur Zylinderdetektion in Laserscans. In T. Luhmann and C. Schumacher, editors, *Photogrammetrie Laserscanning Optische 3D-Messtechnik, Beiträge der Oldenburger 3D-Tage 2022, Jade Hochschule*, pages 260–267. Wichmann Verlag, Feb. 2022.

[19] A. Nüchter, M. Bleier, J. Schauer, and P. Janotta. Improving Google's cartographer 3D mapping by continuous-time SLAM. In *Proceedings of the 7th ISPRS International Workshop 3D-ARCH 2017: 3D Virtual Reconstruction and Visualization of Complex Architectures*, volume XLII-2/W3 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 543–549, Nafplio, Greece, Mar. 2017. Copernicus GmbH.

[20] A. Nüchter, M. Bleier, J. Schauer, and P. Janotta. Continuous-time SLAM—improving Google's cartographer 3D mapping. In F. Remondino, A. Georgopoulos, D. Gonzalez-Aguilera, and P. Agrafiotis, editors, *Latest Developments in Reality-Based 3D Surveying and Modelling*, pages 53–73. MDPI, Basel, Switzerland, Jan. 2018.

[21] J. van der Lucht, M. Bleier, F. Leutert, K. Schilling, and A. Nüchter. Korrektur der Brechung an der Wasseroberfläche beim triangulationsbasierten 3D-Laserscannen. In *Photogrammetrie Laserscanning Optische 3D-Messtechnik, Beiträge der Oldenburger 3D-Tage 2018, Jade Hochschule*, pages 87–102, Feb. 2018.

[22] J. van der Lucht, M. Bleier, F. Leutert, K. Schilling, and A. Nüchter. Structured-light based 3D laser scanning of semi-submerged structures. In *Proceedings of the Technical Commision II Mid-term Symposium "Towards Photogrammetry 2020"*, volume IV-2 of *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 287–294, Riva del Garda, Italy, June 2018. Copernicus GmbH.

[23] J. van der Lucht, M. Bleier, F. Leutert, K. Schilling, and A. Nüchter. Korrektur der Brechung an der Wasseroberfläche beim triangulationsbasierten 3D-Laserscannen. *Allgemeine Vermessungs-Nachrichten (AVN)*, 126(3):43–52, Mar. 2019.

[24] J. van der Lucht, M. Bleier, and A. Nüchter. A low cost underwater test environment. In *Proceedings of the 6th International Workshop LowCost 3D – Sensors, Algorithms, Applications*, volume XLII-2/W17 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 399–404, Strasbourg, France, Dec. 2019. Copernicus GmbH.

[25] T. Wahl, D. Borrmann, M. Bleier, A. Nüchter, T. Wiemann, T. Hänel, and N. Aschenbruck. WIP: Real-world 3D models derived from mobile mapping for ray launching based propagation loss modeling. In *Proceedings of the 23rd IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM '22)*, pages 1–4, Belfast, UK, June 2022.

References

[26] 3D at Depth. SL3 Subsea LiDAR. `https://3datdepth.com/product/subsea-lidar-sl3`. Webpage, accessed December 22, 2022.

[27] A. S. Abdul Ghani and N. A. Mat Isa. Underwater image quality enhancement through integrated color model with rayleigh distribution. *Applied Soft Computing*, 27:219–230, Feb. 2015.

[28] S. Agarwal, K. Mierle, and The Ceres Solver Team. Ceres Solver. software release 2.1, Mar. 2022. Library available online `https://github.com/ceres-solver/ceres-solver`.

[29] P. Agrafiotis, K. Karantzalos, A. Georgopoulos, and D. Skarlatos. Correcting image refraction: Towards accurate aerial image-based bathymetry mapping in shallow waters. *Remote Sensing*, 12(2):322, Jan. 2020.

[30] P. Agrafiotis, K. Karantzalos, A. Georgopoulos, and D. Skarlatos. Learning from synthetic data: Enhancing refraction correction accuracy for airborne image-based bathymetric mapping of shallow coastal waters. *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 89:91–109, May 2021.

[31] A. Agrawal, S. Ramalingam, Y. Taguchi, and V. Chari. A theory of multi-layer flat refractive geometry. In *Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3346–3353, Providence, USA, June 2012. IEEE.

[32] A. Agrawal, S. Ramalingam, Y. Taguchi, and V. Chari. A theory of multi-layer flat refractive geometry supplementary materials. Technical report, Mitsubishi Electric Research Labs (MERL) and INRIA Rhône-Alpes, 2012. Available online `http://www.amitkagrawal.com/cvpr12/AgrawalCVPR12Supplementary.pdf`, accessed December 22, 2022.

[33] D. Akkaynak and T. Treibitz. Sea-thru: A method for removing water from underwater images. In *Proceedings of the 2019 IEEE/CVF conference on computer vision and pattern recognition (CVPR)*, pages 1682–1691, Long Beach, USA, June 2019. IEEE.

[34] J. Albiez, S. Joyeux, C. Gaudig, J. Hilljegerdes, S. Kroffke, C. Schoo, S. Arnold, G. Mimoso, P. Alcantara, R. Saback, J. Britto, D. Cesar, G. Neves, T. Watanabe, P. Merz Paranhos, M. Reis, and F. Kirchner. FlatFish - a compact subsea-resident inspection AUV. In *Proceedings of OCEANS 2015 - MTS/IEEE Washington*, Washington, USA, Oct. 2015. IEEE.

[35] J. Almeida, A. Ferreira, B. Matias, A. Dias, A. Martins, F. Silva, J. Oliveira, P. Sousa, M. Moreira, T. Miranda, C. Almeida, and E. Silva. Air and underwater survey of water enclosed spaces for ¡VAMOS! project. In *Proceedings of OCEANS 2016 MTS/IEEE Monterey*, Monterey, USA, Sept. 2016. IEEE.

[36] J. Almeida, A. Ferreira, B. Matias, C. Lomba, A. Martins, and E. Silva. ¡VAMOS! underwater mining machine navigation system. In *Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1520–1526, Madrid, Spain, Oct. 2018. IEEE.

[37] J. Almeida, B. Matias, A. Ferreira, C. Almeida, A. Martins, and E. Silva. Underwater localization system combining iUSBL with dynamic SBL in ¡VAMOS! trials. *Sensors*, 20(17):4710, Aug. 2020.

[38] G. H. An, S. Lee, M.-W. Seo, K. Yun, W.-S. Cheong, and S.-J. Kang. Charuco board-based omnidirectional camera calibration method. *Electronics*, 7(12):421, Dec. 2018.

[39] A. Andersson. Range gated viewing with underwater camera. Master's thesis, Linköping University, The Institute of Technology., 2005.

[40] A. Anwer, S. S. A. Ali, A. Khan, and F. Mériaudeau. Real-time underwater 3D scene reconstruction using commercial depth sensor. In *Proceedings of the 2016 IEEE International Conference on Underwater System Technology: Theory and Applications (USYS)*, pages 67–70, Penang, Malaysia, Dec. 2016. IEEE.

[41] E. Bakker, G. Žibret, and J. Rainbird. The ¡VAMOS! sustainable underwater mining solution. *Journal of the European Federation of Geologists - Geology and a sustainable future*, 44:58–62, Nov. 2017.

[42] S. Barkby, S. B. Williams, O. Pizarro, and M. V. Jakuba. Bathymetric particle filter SLAM using trajectory maps. *The International Journal of Robotics Research*, 31(12):1409–1430, Oct. 2012.

[43] J. J. Becker, D. T. Sandwell, W. H. F. Smith, J. Braud, B. Binder, J. Depner, D. Fabre, J. Factor, S. Ingalls, S.-H. Kim, R. Ladner, K. Marks, S. Nelson, A. Pharaoh, R. Trimmer, J. V. Rosenberg, G. Wallace, and P. Weatherall. Global bathymetry and elevation data at 30 arc seconds resolution: SRTM30_PLUS. *Marine Geodesy*, 32(4):355–371, Nov. 2009.

[44] S. Beigpour, C. Riess, J. Van De Weijer, and E. Angelopoulou. Multi-illuminant estimation with conditional random fields. *IEEE Transactions on Image Processing*, 23(1):83–96, Oct. 2013.

[45] P. Besl and N. McKay. A method for registration of 3–D shapes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2):239–256, Feb. 1992.

[46] M. G. Besselmann, L. Puck, L. Steffen, A. Roennau, and R. Dillmann. VDB-mapping: A high resolution and real-time capable 3D mapping framework for versatile mobile robots. In *Proceedings of the 2021 IEEE 17th International Conference on Automation Science and Engineering (CASE)*, pages 448–454, Lyon, France, Aug. 2021. IEEE.

[47] G. Bianco, A. Gallo, F. Bruno, and M. Muzzupappa. A comparative analysis between active and passive techniques for underwater 3D reconstruction of close-range objects. *Sensors*, 13(8):11007–11031, Aug. 2013.

[48] G. Bianco, M. Muzzupappa, F. Bruno, R. Garcia, and L. Neumann. A new color correction method for underwater imaging. In *Proceedings of the ISPRS/CIPA Underwater 3D Recording and Modeling Workshop*, volume XL-5/W5 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 25–32, Piano di Sorrento, Italy, Apr. 2015. Copernicus GmbH.

[49] V. Bichucher, J. M. Walls, P. Ozog, K. A. Skinner, and R. M. Eustice. Bathymetric factor graph SLAM with sparse point cloud alignment. In *Proceedings of OCEANS 2015 - MTS/IEEE Washington*, Washington, USA, Oct. 2015. IEEE.

[50] A. Birk, T. Doernbach, C. Mueller, T. Łuczynski, A. G. Chavez, D. Koehntopp, A. Kupcsik, S. Calinon, A. K. Tanwani, G. Antonelli, et al. Dexterous underwater manipulation from onshore locations: Streamlining efficiencies for remotely operated underwater vehicles. *IEEE Robotics & Automation Magazine*, 25(4):24–33, Dec. 2018.

[51] N. Börlin and P. Grussenmeyer. Camera calibration using the damped bundle adjustment toolbox. In *Proceedings of the ISPRS Technical Commission V Symposium*, volume II-5 of *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 89–96, Riva del Garda, Italy, June 2014. Copernicus GmbH.

[52] M. Born and E. Wolf. *Principles of Optics – Electromagnetic Theory of Propagation Interference and Diffraction of Light*. Pergamon Press, 4th edition, 1970.

[53] D. Borrmann. *Multi-modal 3D mapping – Combining 3D point clouds with thermal and color information*. PhD thesis, Julius Maximilian University of Würzburg, Faculty of Mathematics and Computer Science, 2018.

[54] D. Borrmann, J. Elseberg, K. Lingemann, A. Nüchter, and J. Hertzberg. Globally consistent 3D mapping with scan matching. *Robotics and Autonomous Systems*, 56(2):130–142, Feb. 2008.

[55] D. Borrmann, R. Hess, D. Eck, H. Houshiar, A. Nüchter, and K. Schilling. Evaluation of methods for robotic mapping of cultural heritage sites. In *Proceedings of the 2th IFAC conference on Embedded Systems, Computer Intelligence and Telematics (CESCIT)*, pages 105–110, Maribor, Slovenia, June 2015.

[56] J.-Y. Bouget. Camera Calibration Toolbox for MATLAB (1.0). Technical report, CaltechDATA, May 2022. Software available at: `https://doi.org/10.22002/D1.20164`.

[57] J.-Y. Bouguet, M. Weber, and P. Perona. What do planar shadows tell about scene geometry? In *Proceedings of the 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Fort Collins, USA, June 1999. IEEE.

[58] N. Boutros, M. R. Shortis, and E. S. Harvey. A comparison of calibration methods and system configurations of underwater stereo-video systems for applications in marine ecology. *Limnology and Oceanography: Methods*, 13(5):224–236, May 2015.

[59] F. Boyer, V. Lebastard, C. Chevallereau, S. Mintchev, and C. Stefanini. Underwater navigation based on passive electric sense: New perspectives for underwater docking. *The International Journal of Robotics Research*, 34(9):1228–1250, Aug. 2015.

[60] G. Bradski. The OpenCV library. *Dr. Dobb's Journal of Software Tools*, 2000. Library available online `https://github.com/opencv/opencv`.

[61] D. H. Brainard and W. T. Freeman. Bayesian color constancy. *Journal of the Optical Society of America A*, 14(7):1393–1411, July 1997.

[62] C. Bräuer-Burchardt, M. Heinze, I. Schmidt, P. Kühmstedt, and G. Notni. Underwater 3D surface measurement using fringe projection based scanning devices. *Sensors*, 16(1), Jan. 2016.

[63] C. Bräuer-Burchardt, C. Munkelt, I. Gebhart, M. Heinze, S. Heist, P. Kühmstedt, and G. Notni. A-priori calibration of a structured light underwater 3D sensor. *Journal of Marine Science and Engineering*, 8(9):635, Aug. 2020.

[64] C. Bräuer-Burchardt, C. Munkelt, M. Heinze, I. Gebhart, P. Kühmstedt, and G. Notni. Underwater 3D measurements with advanced camera modelling. *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 90:55–67, Feb. 2022.

[65] C. Bräuer-Burchardt, S. Ölsner, P. Kühmstedt, and G. Notni. Comparison of calibration strategies for optical 3D scanners based on structured light projection using a new evaluation methodology. In *Proceedings of Videometrics, Range Imaging, and Applications XIV, SPIE Optical Metrology*, volume 10332, pages 94–103, Munich, Germany, June 2017. SPIE.

[66] D. C. Brown. Decentering distortion of lenses. *Photogrammetric Engineering and Remote Sensing*, 32(3):444–462, 1966.

[67] D. C. Brown. Close-range camera calibration. *Photogrammetric Engineering and Remote Sensing*, 37(8):855–866, 1971.

[68] F. Bruno, G. Bianco, M. Muzzupappa, S. Barone, and A. V. Razionale. Experimentation of structured light and stereo vision for underwater 3D reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 66(4):508–518, July 2011.

[69] F. Bruno, A. Lagudi, L. Barbieri, M. Muzzupappa, G. Ritacco, A. Cozza, M. Cozza, R. Peluso, M. Lupia, and G. Cario. Virtual and augmented reality tools to improve the exploitation of underwater archaeological sites by diver and non-diver tourists. In *Proceedings of the Euro-Mediterranean Conference (EuroMed)*, volume 10058 of *Lecture Notes in Computer Science*, pages 269–280, Nicosia, Cyprus, Oct. 2016. Springer.

[70] C. Bräuer-Burchardt, P. Kühmstedt, and G. Notni. Combination of air- and water-calibration for a fringe projection based underwater 3D-scanner. In *Proceedings of the International Conference on Computer Analysis of Images and Patterns (CAIP)*, volume 9257 of *Lecture Notes in Computer Science*, pages 49–60, Valletta, Malta, Sept. 2015. Springer.

[71] G. Buchsbaum. A spatial processor model for object colour perception. *Journal of the Franklin institute*, 310(1), July 1980.

[72] H. Bülow and A. Birk. Spectral registration of noisy sonar data for underwater 3D mapping. *Autonomous Robots*, 30(3):307–331, Feb. 2011.

[73] J. Burns, D. Delparte, R. Gates, and M. Takabayashi. Integrating structure-from-motion photogrammetry with geospatial software as a novel technique for quantifying 3D ecological characteristics of coral reefs. *PeerJ*, 3:e1077, July 2015.

[74] J. Burns, A. Fukunaga, K. Pascoe, A. Runyan, B. Craig, J. Talbot, A. Pugh, and R. Kosaki. 3D habitat complexity of coral reefs in the northwestern hawaiian islands is driven by coral assemblage structure. In *Proceedings of the 2nd ISPRS International Workshop Underwater 3D Recording & Modelling: A Tool for Modern Applications and CH Recording*, volume XLII-2/W10 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 61–67, Limassol, Cyprus, May 2019. Copernicus GmbH.

[75] P. D. Buschinelli, G. Matos, T. Pinto, and A. Albertazzi. Underwater 3D shape measurement using inverse triangulation through two flat refractive surfaces. In *Proceedings of OCEANS 2016 MTS/IEEE Monterey*, Monterey, USA, Sept. 2016. IEEE.

[76] N. Börlin and P. Grussenmeyer. Bundle adjustment with and without damping. *The Photogrammetric Record*, 28(144):396–415, Oct. 2013.

[77] L. Calvet, P. Gurdjos, C. Griwodz, and S. Gasparini. Detection and accurate localization of circular fiducials under highly challenging conditions. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 562–570, Las Vegas, Nevada, June 2016. IEEE.

[78] R. Campos, R. Garcia, P. Alliez, and M. Yvinec. A surface reconstruction method for in-detail underwater 3D optical mapping. *The International Journal of Robotics Research*, 34(1):64–89, Jan. 2015.

[79] M. Castillón, A. Palomer, J. Forest, and P. Ridao. State of the art of underwater active optical 3D scanners. *Sensors*, 19(23):5161, Nov. 2019.

[80] M. Castillón, A. Palomer, J. Forest, and P. Ridao. Underwater 3D scanner model using a biaxial MEMS mirror. *IEEE Access*, 9:50231–50243, Mar. 2021.

[81] M. Castillón, R. Pi, N. Palomeras, and P. Ridao. Extrinsic visual–inertial calibration for motion distortion correction of underwater 3D scans. *IEEE Access*, 9:93384–93398, June 2021.

[82] F. Chadebecq, F. Vasconcelos, G. Dwyer, R. Lacher, S. Ourselin, T. Vercauteren, and D. Stoyanov. Refractive structure-from-motion through a flat refractive interface. In *Proceedings of the 2017 IEEE International Conference on Computer Vision (ICCV)*, pages 5315–5323, Venice, Italy, Oct. 2017. IEEE.

[83] M. Chambah, D. Semani, A. Renouf, P. Courtellemont, and A. Rizzi. Underwater color constancy: enhancement of automatic live fish recognition. In *Proceedings of Color Imaging IX: Processing, Hardcopy, and Applications, Electronic Imaging*, volume 5293, pages 157–168, San Jose, California, Dec. 2003. SPIE.

[84] A. G. Chavez, C. A. Mueller, T. Doernbach, and A. Birk. Underwater navigation using visual markers in the context of intervention missions. *International Journal of Advanced Robotic Systems*, 16(2):1729881419838967, Mar. 2019.

[85] Z. Cheng, K. Yang, J. Han, Y. Zhou, L. Sun, W. Li, and M. Xia. Improved time-of-flight range acquisition technique in underwater lidar experiments. *Applied optics*, 54(18):5715–5725, June 2015.

[86] S. Chi, Z. Xie, and W. Chen. A laser line auto-scanning system for underwater 3D reconstruction. *Sensors*, 16(9):1534, Sept. 2016.

[87] C. S. Chin, N. B. Kamsani, X. Zhong, R. Cui, and C. Yang. Unity3D serious game engine for high fidelity virtual reality training of remotely-operated vehicle pilot. In *Proceedings of the 2018 10th International Conference on Modelling, Identification and Control (ICMIC)*, Guiyang, China, July 2018. IEEE.

[88] C. S. Clay and H. Medwin. *Acoustical oceanography: principles and applications*. Wiley-Interscience, New York, USA, 1st edition, 1977.

[89] E. Coiras, Y. Petillot, and D. M. Lane. Multiresolution 3-D reconstruction from side-scan sonar images. *IEEE Transactions on Image Processing*, 16(2):382–390, Jan. 2007.

[90] S. Collings, T. J. Martin, E. Hernandez, S. Edwards, A. Filisetti, G. Catt, A. Marouchos, M. Boyd, and C. Embry. Findings from a combined subsea LiDAR and multibeam survey at kingston reef, western australia. *Remote Sensing*, 12(15):2443, July 2020.

[91] A. E. Conrady. Decentred lens-systems. *Monthly notices of the royal astronomical society*, 79(5):384–390, 1919.

[92] B. Curless and M. Levoy. A volumetric method for building complex models from range images. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, pages 303–312. ACM, Aug. 1996.

[93] F. R. Dalgleish, F. M. Caimi, W. B. Britton, and C. F. Andren. Improved LLS imaging performance in scattering-dominant waters. In *Proceedings of Ocean Sensing and Monitoring, SPIE Defense, Security, and Sensing*, volume 7317, pages 89–100, Orlando, USA, Apr. 2009. SPIE.

[94] A. Dancu, M. Fourgeaud, Z. Franjcic, and R. Avetisyan. Underwater reconstruction using depth sensors. In *Proccedings of SIGGRAPH Asia 2014 Technical Briefs*, Shenzhen, China, Dec. 2014.

[95] J. P. de Villiers, F. W. Leuschner, and R. Geldenhuys. Centi-pixel accurate real-time inverse distortion correction. In J. T. Wen, D. Hodko, Y. Otani, J. Kofman, and O. Kaynak, editors, *Proceedings of Optomechatronic Technologies 2008, International Symposium on Optomechatronic Technologies*, volume 7266, pages 320–327, San Diego, USA, Nov. 2008. SPIE.

[96] A. R. Denny, T. O. Sæbø, R. E. Hansen, and R. B. Pedersen. The use of synthetic aperture sonar to survey seafloor massive sulfide deposits. *The Journal of Ocean Technology*, 10(1):37–53, Jan. 2015.

[97] J. Dera. *Marine physics*, volume 53 of *Elsevier Oceanography Series*. Elsevier, 1992.

[98] J. T. Dietrich. Bathymetric structure-from-motion: extracting shallow stream bathymetry from multi-view stereo photogrammetry. *Earth Surface Processes and Landforms*, 42(2):355–364, Oct. 2017.

[99] S. T. Digumarti, G. Chaurasia, A. Taneja, R. Siegwart, A. Thomas, and P. Beardsley. Underwater 3D capture using a low-cost commercial depth camera. In *Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, Lake Placid, USA, Mar. 2016. IEEE.

[100] T. Dolereit. *A Virtual Object Point Model for the Calibration of Underwater Stereo Camerasto Recover Accurate 3D Information*. PhD thesis, University of Rostock, Faculty of Computer Science and Electrical Engineering, 2018.

[101] M. Doneus, N. Doneus, C. Briese, M. Pregesbauer, G. Mandlburger, and G. Verhoeven. Airborne laser bathymetry–detecting and recording submerged archaeological sites from the air. *Journal of Archaeological Science*, 40(4):2136–2151, Apr. 2013.

[102] M. Doneus, I. Miholjek, G. Mandlburger, N. Doneus, G. Verhoeven, C. Briese, and M. Pregesbauer. Airborne laser bathymetry for documentation of submerged archaeological sites in shallow water. In *Proceedings of the ISPRS/CIPA Underwater 3D Recording and Modeling Workshop*, volume XL-5/W5 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 99–107, Piano di Sorrento, Italy, Apr. 2015. Copernicus GmbH.

[103] F. Dornaika and R. Horaud. Simultaneous robot-world and hand-eye calibration. *IEEE Transactions on Robotics and Automation*, 14(4):617–622, Aug. 1998.

[104] D. B. dos Santos Cesar, C. Gaudig, M. Fritsche, M. A. dos Reis, and F. Kirchner. An evaluation of artificial fiducial markers in underwater environments. In *Proceedings of OCEANS 2015-Genova*, Genova, Italy, May 2015. IEEE.

[105] P. Drap. Underwater photogrammetry for archaeology. In D. C. D. Silva, editor, *Special Applications of Photogrammetry*, volume 114, chapter 6. IntechOpen, Rijeka, Croatia, 2012.

[106] P. Drap, D. Merad, B. Hijazi, L. Gaoua, M. M. Nawaf, M. Saccone, B. Chemisky, J. Seinturier, J.-C. Sourisseau, T. Gambin, et al. Underwater photogrammetry and object modeling: a case study of xlendi wreck in malta. *Sensors*, 15(12):30351–30384, Dec. 2015.

[107] P. Drap, D. Merad, A. Mahiddine, J. Seinturier, P. Gerenton, D. Peloso, J.-M. Boï, O. Bianchimani, and J. Garrabou. In situ underwater measurements of red coral: Non-intrusive approach based on coded targets and photogrammetry. *International Journal of Heritage in the Digital Era*, 3(1):123–139, Mar. 2014.

[108] P. Drap, J. Seinturier, D. Scaradozzi, P. Gambogi, L. Long, and F. Gauch. Photogrammetry for virtual exploration of underwater archeological sites. In *Proceedings of the XXI International CIPA Symposium*, Athens, Greece, Oct. 2007. CIPA.

[109] A. Duda. *Underwater Visual Multi-Modal 3D Sensing*. PhD thesis, University of Bremen, Faculty of Mathematics and Computer Science, 2020.

[110] A. Duda and U. Frese. Accurate detection and localization of checkerboard corners for calibration. In *Proceedings of the 29th British Machine Vision Conference (BMVC)*, Newcastle, UK, Sept. 2018.

[111] A. Duda and C. Gaudig. Refractive forward projection for underwater flat port cameras. In *Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2022–2027, Daejeon, South Korea, Oct. 2016. IEEE.

[112] A. Duda, J. Schwendner, and C. Gaudig. SRSL: Monocular self-referenced line structured light. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 717–722, Hamburg, Germany, Sept. 2015. IEEE.

[113] A. Ecker, K. N. Kutulakos, and A. D. Jepson. Shape from planar curves: A linear escape from flatland. In *Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Minneapolis, USA, June 2007. IEEE.

[114] T. Ekkel, J. Schmik, T. Luhmann, and H. Hastedt. Precise laser-based optical 3D measurement of welding seams under water. In *Proceedings of the ISPRS/CIPA Underwater 3D Recording and Modeling Workshop*, volume XL-5/W5 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 117–122, Piano di Sorrento, Italy, Apr. 2015. Copernicus GmbH.

[115] J. Elseberg, D. Borrmann, and A. Nüchter. Automatic and full calibration of mobile laser scanning systems. In *Proceedings of the 13th International Symposium on Experimental Robotics (ISER)*, Springer Tracts in Advanced Robotics, pages 907–917, Québec City, Canada, June 2012. Springer.

[116] J. Elseberg, D. Borrmann, and A. Nüchter. Algorithmic solutions for computing precise maximum likelihood 3D point clouds from mobile laser scanning platforms. *Remote Sensing*, 5(11):5871–5906, Nov. 2013.

[117] C. Fabbri, M. J. Islam, and J. Sattar. Enhancing underwater imagery using generative adversarial networks. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 7159–7165, Brisbane, Australia, May 2018. IEEE.

[118] A. Ferreira, J. Almeida, A. Martins, A. Matos, and E. Silva. 3DupIC: An underwater scan matching method for three-dimensional sonar registration. *Sensors*, 22(10):3631, May 2022.

[119] F. Ferreira, D. Machado, G. Ferri, S. Dugelay, and J. Potter. Underwater optical and acoustic imaging: A time for fusion? a brief overview of the state-of-the-art. In *Proceedings of OCEANS 2016 MTS/IEEE Monterey*, Monterey, USA, Sept. 2016. IEEE.

[120] H. Ferreira, C. Almeida, A. Martins, J. Almeida, N. Dias, A. Dias, and E. Silva. Autonomous bathymetry for risk assessment with ROAZ robotic surface vehicle. In *Proceedings of OCEANS 2009 - Europe*, Bremen, Germany, May 2009. IEEE.

[121] M. Fiala. ARTag, a fiducial marker system using digital techniques. In *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 590–596, San Diego, USA, June 2005. IEEE.

[122] G. D. Finlayson. Color in perspective. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(10):1034–1038, Oct. 1996.

[123] G. D. Finlayson, S. D. Hordley, and I. Tastl. Gamut constrained illuminant estimation. *International Journal of Computer Vision*, 67(1):93–109, Apr. 2006.

[124] G. D. Finlayson and E. Trezzi. Shades of gray and colour constancy. In *Proceedings of the IS&T/SID 12th Color Imaging Conference: Color Science and Engineering Systems, Technologies, Applications*, pages 37–41, Scottsdale, Arizona, Nov. 2004. Society for Imaging Science and Technology.

[125] A. W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Kauai, USA, Dec. 2001. IEEE.

[126] A. Fitzpatrick, A. Singhvi, and A. Arbabian. An airborne sonar system for underwater remote sensing and imaging. *IEEE Access*, 8:189945–189959, Oct. 2020.

[127] W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *Proceedings of the ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data*, pages 281–305, Interlaken, Switzerland, June 1987. Interlaken.

[128] W. Förstner and B. P. Wrobel. *Photogrammetric computer vision*, volume 11 of *Geometry and Computing*. Springer, 2016.

[129] D. A. Forsyth. A novel algorithm for color constancy. *International Journal of Computer Vision*, 5(1):5–35, Aug. 1990.

[130] R. Francois and G. Garrison. Sound absorption based on ocean measurements. part ii: Boric acid contribution and equation for total absorption. *Journal of the Acoustical Society of America*, 72:1879–1890, June 1982.

[131] J. Fryer. A simple system for photogrammetric mapping in shallow water. *The Photogrammetric Record*, 11(62):203–208, Oct. 1983.

[132] J. Fryer and H. Kniest. Errors in depth dertermination caused by waves in through-water photogrammetry. *The Photogrammetric Record*, 11(66):745–753, Oct. 1985.

[133] J. G. Fryer and C. Fraser. On the calibration of underwater cameras. *The Photogrammetric Record*, 12(67):73–85, Apr. 1986.

[134] B. Funt, V. Cardei, and K. Barnard. Learning color constancy. In *Proceedings of the 4th IS&T/SID Color Imaging Conference Color Science, Systems, andApplications*, pages 58–60, Scottsdale, USA, Nov. 1996. Society for Imaging Science and Technology, Society for Imaging Science & Technology.

[135] P. Furgale, J. Rehder, and R. Siegwart. Unified temporal and spatial calibration for multi-sensor systems. In *Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1280–1286, Tokyo, Japan, Nov. 2013. IEEE.

[136] R. Furukawa and H. Kawasaki. Self-calibration of multiple laser planes for 3D scene reconstruction. In *Proceedings of the 3rd International Symposium on 3D Data Processing, Visualization, and Transmission (3DPVT)*, pages 200–207, Chapel Hill, USA, June 2006. IEEE.

[137] R. Furukawa and H. Kawasaki. Laser range scanner based on self-calibration techniques using coplanarities and metric constraints. *Computer Vision and Image Understanding*, 113(11):1118–1129, Nov. 2009.

[138] R. Furukawa, H. Q. H. Viet, H. Kawasaki, R. Sagawa, and Y. Yagi. One-shot range scanner using coplanarity constraints. In *Proceedings of the 15th IEEE International Conference on Image Processing (ICIP)*, pages 1524–1527, San Diego, USA, Oct. 2008. IEEE.

[139] M. Galassi, J. Davies, J. Theiler, B. Gough, G. Jungman, P. Alken, M. Booth, F. Rossi, and R. Ulerich. GNU scientific library reference manual. software release 2.7, June 2021. Library available online `https://www.gnu.org/software/gsl`.

[140] S. Garrido-Jurado, R. Munoz-Salinas, F. J. Madrid-Cuevas, and R. Medina-Carnicer. Generation of fiducial marker dictionaries using mixed integer linear programming. *Pattern Recognition*, 51:481–491, Mar. 2016.

[141] GEBCO. The general bathymetric chart of the oceans. `https://www.gebco.net/`. Webpage, accessed December 22, 2022.

[142] A. Gijsenij, T. Gevers, and J. van de Weijer. Computational color constancy: Survey and experiments. *IEEE Transactions on Image Processing*, 20(9):2475–2489, Feb. 2011.

[143] A. Gijsenij, R. Lu, and T. Gevers. Color constancy for multiple light sources. *IEEE Transactions on Image Processing*, 21(2):697–707, Aug. 2011.

[144] J. Gillham. Development of a field deployable underwater laser scanning system. Master's thesis, University of Waterloo, Faculty of Engineering, 2011.

[145] A. S. Glassner. *An introduction to ray tracing*. Morgan Kaufmann Publishers, Inc., 1989.

[146] A. Gomez Chavez, A. Ranieri, D. Chiarella, E. Zereik, A. Babić, and A. Birk. CADDY underwater stereo-vision dataset for human–robot interaction (HRI) in the context of diver activities. *Journal of Marine Science and Engineering*, 7(1):16, Jan. 2019.

[147] C. Gu, Y. Cong, and G. Sun. Three birds, one stone: Unified laser-based 3-D reconstruction across different media. *IEEE Transactions on Instrumentation and Measurement*, 70, Nov. 2020.

[148] G. Guennebaud, B. Jacob, et al. Eigen v3. software release 3.4, Aug. 2021. Library available online `https://eigen.tuxfamily.org/`.

[149] T. Hanning. *High precision camera calibration*. Springer, 2011.

[150] M. Hardy. Subsea LiDAR metrology, June 2014. Presentation at SUT Subsea Metrology Seminar 2014, The Hydrographic Society.

[151] R. Hartley. An algorithm for self calibration from several views. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 94, pages 908–912, Seattle, USA, June 1994. IEEE.

[152] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd edition, 2000.

[153] E. Harvey and M. Shortis. A system for stereo-video measurement of sub-tidal organisms. *Marine Technology Society Journal*, 29(4):10–22, Jan. 1995.

[154] M. Haydar, D. Roussel, M. Maïdi, S. Otmane, and M. Mallem. Virtual and augmented reality for cultural computing and heritage: a case study of virtual exploration of underwater archaeological sites. *Virtual Reality*, 15(4):311–327, Oct. 2011.

[155] D.-M. He and G. G. Seet. Underwater lidar imaging in highly turbid waters. In *Proceedings of Ocean Optics: Remote Sensing and Underwater Imaging, International Symposium on Optical Science and Technology*, volume 4488, pages 71–81, San Diego, USA, Jan. 2002. SPIE.

[156] J. Heikkila and O. Silvén. A four-step camera calibration procedure with implicit image correction. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1106–1112, San Juan, USA, June 1997. IEEE.

[157] S. Heist, P. Dietrich, M. Landmann, P. Kühmstedt, and G. Notni. High-speed 3D shape measurement by GOBO projection of aperiodic sinusoidal fringes: a performance analysis. In *Proceedings of Dimensional Optical Metrology and Inspection for Practical Applications VII, SPIE Commercial + Scientific Sensing and Imaging*, volume 10667, pages 33–40, Orlando, USA, May 2018. SPIE.

[158] P. Helmholz and D. D. Lichti. Investigation of chromatic aberration and its influence on the processing of underwater imagery. *Remote Sensing*, 12(18):3002, Sept. 2020.

[159] R. C. Hilldale and D. Raff. Assessing the ability of airborne lidar to map river bathymetry. *Earth Surface Processes and Landforms*, 33(5):773–783, Sept. 2008.

[160] H. Hirschmuller. Accurate and efficient stereo processing by semi-global matching and mutual information. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 807–814, San Diego, USA, June 2005. IEEE.

[161] A. Hornung and L. Kobbelt. Robust reconstruction of watertight 3D models from non-uniformly sampled point clouds without normal information. In *Proceedings of the Fourth Eurographics Symposium on Geometry Processing*, pages 41–50, Cagliari, Italy, June 2006.

[162] N. Hurtós, X. Cufí, and J. Salvi. Calibration of optical camera coupled to acoustic multibeam for underwater 3D scene reconstruction. In *Proceedings of OCEANS'10 IEEE SYDNEY*, pages 1–7, Sydney, Australia, Oct. 2010. IEEE.

[163] M. Imaki, H. Ochimizu, H. Tsuji, S. Kameyama, T. Saito, S. Ishibashi, and H. Yoshida. Underwater three-dimensional imaging laser sensor with 120-deg wide-scanning angle using the combination of a dome lens and coaxial optics. *Optical Engineering*, 56(3), Oct. 2017.

[164] K. Iqbal, R. A. Salam, A. Osman, and A. Z. Talib. Underwater image enhancement using an integrated colour model. *IAENG International Journal of Computer Science*, 34(2), Jan. 2007.

[165] J. L. Irish and W. J. Lillycrop. Scanning laser mapping of the coastal zone: The SHOALS system. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54(2):123–129, July 1999.

[166] K. Istenič, N. Gracias, A. Arnaubec, J. Escartín, and R. Garcia. Automatic scale estimation of structure from motion based 3D models using laser scalers in underwater scenarios. *ISPRS Journal of Photogrammetry and Remote Sensing*, 159:13–25, Jan. 2020.

[167] A. Ivanoff and P. Cherney. Correcting lenses for underwater use. *Journal of the SMPTE*, 69(4):264–266, Apr. 1960.

[168] S. Izadi, D. Kim, O. Hilliges, D. Molyneaux, R. Newcombe, P. Kohli, J. Shotton, S. Hodges, D. Freeman, A. Davison, and A. Fitzgibbon. KinectFusion: Real-time 3D reconstruction and interaction using a moving depth camera. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology*, pages 559–568, New York, USA, Oct. 2011. ACM.

[169] J. S. Jaffe. Computer modeling and the design of optimal underwater imaging systems. *IEEE Journal of Oceanic Engineering*, 15(2):101–111, 1990.

[170] O. Jokinen. Self-calibration of a light striping system by matching multiple 3-D profile maps. In *Proceedings of the 1999 Second International Conference on 3-D Digital Imaging and Modeling*, pages 180–190, Ottawa, Canada, Oct. 1999. IEEE.

[171] A. Jordt. *Underwater 3D reconstruction based on physical models for refraction and underwater light propagation*. PhD thesis, Christian-Albrecht University of Kiel, Faculty of Engineering, 2013.

[172] A. Jordt-Sedlazeck and R. Koch. Refractive calibration of underwater cameras. In *Proceedings of the 12th European Conference on Computer Vision (ECCV)*, pages 846–859, Firenze, Italy, Oct. 2012. Springer.

[173] D. Jung, J. Kim, and G. Byun. Numerical modeling and simulation technique in time-domain for multibeam echo sounder. *International Journal of Naval Architecture and Ocean Engineering*, 10(2):225–234, Mar. 2018.

[174] O. Kahmen and T. Luhmann. Monocular photogrammetric system for 3D reconstruction of welds in turbid water. *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 90:19–35, Mar. 2022.

[175] O. Kahmen, R. Rofallski, N. Conen, and T. Luhmann. On scale definition within calibration of multi-camera systems in multimedia photogrammetry. In *Proceedings of the 2nd ISPRS International Workshop Underwater 3D Recording & Modelling: A Tool for Modern Applications and CH Recording*, volume XLII-2/W10 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 93–100, Limassol, Cyprus, May 2019. Copernicus GmbH.

[176] O. Kahmen, R. Rofallski, and T. Luhmann. Impact of stereo camera calibration to object accuracy in multimedia photogrammetry. *Remote Sensing*, 12(12):2057, June 2020.

[177] J. Kalwa, D. Tietjen, M. Carreiro-Silva, J. Fontes, L. Brignone, N. Gracias, P. Ridao, M. Pfingsthorn, A. Birk, T. Glotzbach, et al. The european project MORPH: Distributed UUV systems for multimodal, 3D underwater surveys. *Marine Technology Society Journal*, 50(4):26–41, July 2016.

[178] L. Kang, L. Wu, and Y.-H. Yang. Experimental study of the influence of refraction on underwater three-dimensional reconstruction using the SVP camera model. *Applied Optics*, 51(31):7591–7603, Oct. 2012.

[179] H. Kato and M. Billinghurst. Marker tracking and hmd calibration for a video-based augmented reality conferencing system. In *Proceedings of the 2nd IEEE and ACM International Workshop on Augmented Reality (IWAR)*, pages 85–94, San Francisco, USA, Oct. 1999. IEEE.

[180] T. Kersten, M. Lindstaedt, and D. Starosta. Comparative geometrical accuracy investigations of hand-held 3D scanning systems – an update. In *Proceedings of the Technical Commision II Mid-term Symposium "Towards Photogrammetry 2020"*, volume XLII-2 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Riva del Garda, Italy, June 2018. Copernicus GmbH.

[181] G. J. Klinker, S. A. Shafer, and T. Kanade. A physical approach to color image understanding. *International Journal of Computer Vision*, 4(1):7–38, Jan. 1990.

[182] F. Klopfer, M. Hämmerle, and B. Höfle. Assessing the potential of a low-cost 3-D sensor in shallow-water bathymetry. *IEEE Geoscience and Remote Sensing Letters*, 14(8):1388–1392, July 2017.

[183] P. Koch, H. Engelhardt, S. May, M. Schmidpeter, J. Ziegler, and A. Nüchter. Signed distance based reconstruction for exploration and change detection in underground mining disaster prevention. In *Proceedings of the IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, pages 360–365, Abu Dhabi, United Arab Emirates, Nov. 2020. IEEE.

[184] P. Koch, S. May, H. Engelhardt, J. Ziegler, and A. Nüchter. Signed distance based reconstruction for exploration and change detection in underground mining disaster prevention. In *Proceedings of the IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, Würzburg, Germany, Sept. 2019. IEEE.

[185] R. Kotowski. Phototriangulation in multi-media photogrammetry. In *Proceedings of the VI ISPRS Congress Technical Commission V: Other Non-Cartographic Applications of Photogrammetry and Remote Sensing*, volume XXVII-B5 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 324–334, Kyoto, Japan, July 1988.

[186] Kraken Robotics. Kraken SeaVision subsea 3D laser imaging system. `https://krakenrobotics.com/our-services/seavision/`. Webpage, accessed December 22, 2022.

[187] M. Krogius, A. Haggenmiller, and E. Olson. Flexible layouts for fiducial tags. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1898–1903, Macau, China, Nov. 2019. IEEE.

[188] E. Kruck. *Lösung großer Gleichungssysteme für photogrammetrische Blockausgleichungen mit erweitertem funktionalen Modell*. PhD thesis, Leibniz University Hannover, Faculty of Civil Engineering and Geodetic Science, 1983.

[189] P. Kühmstedt, C. Bräuer-Burchardt, S. Heist, I. Schmidt, and G. Notni. GOBO projection for underwater 3D measurement technique. In *Proceedings of Optical Measurement Systems for Industrial Inspection X, SPIE Optical Metrology*, volume 10329, page 103290N, Munich, Germany, June 2017. International Society for Optics and Photonics.

[190] Z. L, M. E, S. Gayen, M. Alrubaiee, and R. Alfano. Time-gated backscattered ballistic light imaging of objects in turbid water. *Applied Physics Letters*, 86(1):011115, Dec. 2005.

[191] T. F. Lam, H. Blum, R. Siegwart, and A. Gawel. SL Sensor: An open-source, real-time and robot operating system-based structured light sensor for high accuracy construction robotic applications. *Automation in Construction*, 142:104424, Oct. 2022.

[192] H. A. Lauterbach and A. Nüchter. Aerial 3D mapping with continuous time ICP for urban search and rescue. In J. Będkowski, editor, *Autonomous Mobile Mapping Robots*, chapter 6. IntechOpen, Rijeka, Croatia, Nov. 2022.

[193] S. Leather. High-quality surveys of man-made structures as an aid to improved decision making. The Teledyne Marine Channel. `https://www.video.teledynemarine.com/video/10368962/high-quality-surveys-of-man-made-structures-as-an`, Sept. 2014. Webpage, accessed December 22, 2022.

[194] J. Leatherdale and D. Turner. Underwater photogrammetry in the North Sea. *The Photogrammetric Record*, 11(62):151–167, Oct. 1983.

[195] F. Leutert. *Flexible Augmented Reality Systeme für robotergestützte Produktionsumgebungen*. PhD thesis, Julius Maximilian University of Würzburg, Faculty of Mathematics and Computer Science, 2021.

[196] A. Li, L. Wang, and D. Wu. Simultaneous robot-world and hand-eye calibration using dual-quaternions and kronecker product. *International Journal of Physical Sciences*, 5(10):1530–1536, Sept. 2010.

[197] J. Li, K. A. Skinner, R. Eustice, and M. Johnson-Roberson. WaterGAN: Unsupervised generative network to enable real-time color correction of monocular underwater images. *IEEE Robotics and Automation Letters (RA-L)*, 3(1):387–394, Jan. 2018.

[198] R. Li, H. Li, W. Zou, R. G. Smith, and T. A. Curran. Quantitative photogrammetric analysis of digital underwater video imagery. *IEEE Journal of Oceanic Engineering*, 22(2):364–375, Apr. 1997.

[199] Q. Lin and C. Kuo. Virtual tele-operation of underwater robots. In *Proceedings of the International Conference on Robotics and Automation (ICRA)*, pages 1022–1027, Albuquerque, USA, Apr. 1997. IEEE.

[200] Q. Lin and C. Kuo. On applying virtual reality to underwater robot tele-operation and pilot training. *International Journal of Virtual Reality (IJVR)*, 5(1):71–91, Jan. 2015.

[201] T. Luhman. *Erweiterte Verfahren zur geometrischen Kamerakalibrierung in der Nahbereichsphotogrammetrie*, volume 645 of *Reihe C: Dissertationen*. Deutsche Geodätische Kommission, Verlag der Bayerischen Akademie der Wissenschaften in Kommission beim Verlag C. H. Beck, 2010.

[202] T. Luhmann, S. Robson, S. Kyle, and J. Boehm. *Close-range photogrammetry and 3D imaging*. De Gruyter STEM. Walter de Gruyter, 3rd edition, 2020.

[203] X. Lurton. *An introduction to underwater acoustics: principles and applications*. Springer Praxis Books (PRAXIS), Geophysical Sciences (GEOPHYS). Springer, 2nd edition, 2002.

[204] H.-G. Maas. *Digitale Photogrammetrie in der dreidimensionalen Strömungsmesstechnik*. PhD thesis, ETH Zurich, Institute of Geodesy and Photogrammetry, 1992.

[205] H.-G. Maas. On the accuracy potential in underwater/multimedia photogrammetry. *Sensors*, 15(8):18140–18152, July 2015.

[206] D. Mader, K. Richter, P. Westfeld, and H.-G. Maas. Potential of a non-linear full-waveform stacking technique in airborne LiDAR bathymetry. *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 89:139–158, June 2021.

[207] J. Mallon and P. F. Whelan. Which pattern? biasing aspects of planar calibration patterns and detection methods. *Pattern Recognition Letters*, 28(8):921–930, June 2007.

[208] G. Mandlburger. Through-water dense image matching for shallow water bathymetry. *Photogrammetric Engineering & Remote Sensing*, 85(6):445–455, June 2019.

[209] G. Mandlburger, M. Pfennigbauer, F. Steinbacher, and N. Pfeifer. Airborne hydrographic LiDAR mapping–potential of a new technique for capturing shallow water bodies. In *Proceedings of the 19th International Congress on Modelling and Simulation*, pages 2416–2422, Perth, Australia, Dec. 2011.

[210] G. Mandlburger, M. Pfennigbauer, M. Wieser, U. Riegl, and N. Pfeifer. Evaluation of a novel UAV-borne topo-bathymetric laser profiler. In *Proceedings of the XXIII ISPRS Congress*, volume XLI-B1 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 933–939, Prague, Czech Republic, July 2016. Copernicus GmbH.

[211] M. Mangeruga, F. Bruno, M. Cozza, P. Agrafiotis, and D. Skarlatos. Guidelines for underwater image enhancement based on benchmarking of different methods. *Remote Sensing*, 10(10):1652, Oct. 2018.

[212] P. Mariani, I. Quincoces, K. H. Haugholt, Y. Chardard, A. W. Visser, C. Yates, G. Piccinno, G. Reali, P. Risholm, and J. T. Thielemann. Range-gated imaging system for underwater monitoring in ocean environment. *Sustainability*, 11(1):162, Dec. 2019.

[213] A. Martins, J. Almeida, C. Almeida, B. Matias, S. Kapusniak, and E. Silva. EVA a hybrid ROV/AUV for underwater mining operations support. In *Proceedings of 2018 OCEANS - MTS/IEEE Kobe Techno-Oceans (OTO)*, Kobe, Japan, May 2018. IEEE.

[214] A. Martins, A. Ferreira, A. Dias, and B. Matias. ¡VAMOS! project deliverable D4.4: Multi-sensor navigation system. `http://vamos-project.eu/wp-content/uploads/2018/07/D4.4-Multi-Sensor-Navigation-System.pdf`, Oct. 2016.

[215] M. Massot-Campos and G. Oliver-Codina. Underwater laser-based structured light system for one-shot 3D reconstruction. In *Proceedings of IEEE Sensors 2014*, pages 1138–1141, Valencia, Spain, Nov. 2014. IEEE.

[216] M. Massot-Campos and G. Oliver-Codina. Optical sensors and methods for underwater 3D reconstruction. *Sensors*, 15(12):31525–31557, Dec. 2015.

[217] S. May, P. Koch, R. Koch, C. Merkl, C. Pfitzner, and A. Nüchter. A generalized 2D and 3D multi-sensor data integration approach based on signed distance functions for multi-modal robotic mapping. In *Proceedings of the 19th International Workshop on Vision, Modeling and Visualization (VMV)*, pages 95–102, Darmstadt, Germany, Oct. 2014. The Eurographics Association.

[218] S. J. Maybank and O. D. Faugeras. A theory of self-calibration of a moving camera. *International Journal of Computer Vision*, 8(2):123–151, 1992.

[219] J. McCarthy, J. Benjamin, T. Winton, and W. van Duivenvoorde. The rise of 3D in maritime archaeology. In *3D Recording and Interpretation for Maritime Archaeology*, volume 31 of *Coastal Research Library*, chapter 1, pages 1–10. Springer, Mar. 2019.

[220] B. McGlamery. Computer analysis and simulation of underwater camera system performance. Technical Report SIO Ref. 75-2, Visibility Laboratory, University of California, San Diego, June 1975.

[221] D. McLeod, J. Jacobson, M. Hardy, and C. Embry. Autonomous inspection using an underwater 3D LiDAR. In *Proceedings of 2013 OCEANS - San Diego*, San Diego, USA, Sept. 2013. IEEE.

[222] F. Menna, P. Agrafiotis, and A. Georgopoulos. State of the art and applications in archaeological underwater 3D recording and mapping. *Journal of Cultural Heritage*, 33:231–248, Sept. 2018.

[223] F. Menna, E. Nocerino, F. Fassi, and F. Remondino. Geometric and optic characterization of a hemispherical dome port for underwater photogrammetry. *Sensors*, 16(1):48, Jan. 2016.

[224] F. Menna, E. Nocerino, M. M. Nawaf, J. Seinturier, A. Torresani, P. Drap, F. Remondino, and B. Chemisky. Towards real-time underwater photogrammetry for subsea metrology applications. In *Proceedings of OCEANS 2019 - Marseille*, Marseille, France, June 2019. IEEE.

[225] F. Menna, E. Nocerino, and F. Remondino. Flat versus hemispherical dome ports in underwater photogrammetry. In *Proceedings of the 7th ISPRS International Workshop 3D-ARCH 2017: 3D Virtual Reconstruction and Visualization of Complex Architectures*, volume XLII-2/W3 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 481–487, Nafplio, Greece, 2017. Copernicus GmbH.

[226] F. Menna, E. Nocerino, and F. Remondino. Optical aberrations in underwater photogrammetry with flat and hemispherical dome ports. In *Proceedings of Videometrics, Range Imaging, and Applications XIV, SPIE Optical Metrology*, volume 10332, pages 28–41, Munich, Germany, June 2017. SPIE.

[227] F. Menna, E. Nocerino, S. Troisi, and F. Remondino. A photogrammetric approach to survey floating and semi-submerged objects. In *Proceedings of Videometrics, Range Imaging,*

and Applications XII; and Automated Visual Inspection, SPIE Optical Metrology, volume 8791, Munich, Germany, 2013. SPIE.

[228] F. Menna, E. Nocerino, S. Ural, and A. Gruen. Mitigating image residuals systematic patterns in underwater photogrammetry. In *Proceedings of XXIV ISPRS Congress (2020 edition)*, volume XLIII-B2-2020 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 977–984. Copernicus GmbH, 2020.

[229] F. Menna, A. Torresani, E. Nocerino, M. M. Nawaf, J. Seinturier, F. Remondino, P. Drap, and B. Chemisky. Evaluation of vision-based localization and mapping techniques in a subsea metrology scenario. In *Proceedings of the 2nd ISPRS International Workshop Underwater 3D Recording & Modelling: A Tool for Modern Applications and CH Recording*, volume XLII-2/W10 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 127–134, Limassol, Cyprus, May 2019. Copernicus GmbH.

[230] H. Morinaga, H. Baba, M. Visentini-Scarzanella, H. Kawasaki, R. Furukawa, and R. Sagawa. Underwater active oneshot scan with static wave pattern and bundle adjustment. In T. Bräunl, B. McCane, M. Rivera, and X. Yu, editors, *Proceedings of PSIVT 2015: Image and Video Technology*, volume 9431 of *Lecture Notes in Computer Science*, pages 404–418. Springer, Feb. 2016.

[231] C. Mulsow and H.-G. Maas. A universal approach for geometric modelling in underwater stereo image processing. In *Proceedings of the 2014 ICPR Workshop on Computer Vision for Analysis of Underwater Imagery*, pages 49–56, Stockholm, Sweden, Aug. 2014. IEEE.

[232] K. Museth. VDB: High-resolution sparse volumes with dynamic topology. *ACM Transactions on Graphics*, 32(3):27, June 2013.

[233] G. Nagamatsu, T. Ikeda, T. Iwaguchi, D. Thomas, J. Takamatsu, and H. Kawasaki. Self-calibration of multiple-line-lasers based on coplanarity and epipolar constraints for wide area shape scan using moving camera. In *Proceedings of the 2022 26th International Conference on Pattern Recognition (ICPR)*, pages 3959–3965, Montreal, Canada, Aug. 2022. IEEE.

[234] G. Nagamatsu, J. Takamatsu, T. Iwaguchi, D. Thomas, and H. Kawasaki. Self-calibrated dense 3D sensor using multiple cross line-lasers based on light sectioning method and visual odometry. In *Proceedings of the 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 94–100, Prague, Czech Republic, Oct. 2021. IEEE.

[235] S. G. Narasimhan and S. K. Nayar. Structured light methods for underwater imaging: light stripe scanning and photometric stereo. In *Proceedings of OCEANS 2005 MTS/IEEE*, pages 2610–2617, Washington, USA, Sept. 2005. IEEE.

[236] National Oceanic and Atmospheric Administration. Bathymetric data viewer. `https://www.ncei.noaa.gov/maps/bathymetry/`. Webpage, accessed August 1, 2022.

[237] Natural Point Inc. OptiTrack V120:Trio. `https://www.optitrack.com/cameras/v120-trio/specs.html`. Webpage, accessed December 22, 2022.

[238] Y. Nevatia, T. Stoyanov, R. Rathnam, M. Pfingsthorn, S. Markov, R. Ambrus, and A. Birk. Augmented autonomy: Improving human-robot team performance in urban search and rescue. In *Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2103–2108, Nice, France, Sept. 2008. IEEE.

[239] R. A. Newcombe, A. J. Davison, S. Izadi, P. Kohli, O. Hilliges, J. Shotton, D. Molyneaux, S. Hodges, D. Kim, and A. Fitzgibbon. KinectFusion: Real-time dense surface mapping and tracking. In *Proceedings of the 10th IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*, pages 127–136, Basel, Switzerland, Oct. 2011. IEEE.

[240] Newton Labs. M4000UW underwater laser scanner. `https://www.newtonlabs.com/M4000UW_landing.htm`. Webpage, accessed December, 2022.

[241] F. Niemeyer, T. Dolereit, M. Neumann, J. Albiez, M. Vahl, and M. Geist. Untersuchungen von optischen Scansystemen zur geometrischen Erfassung von Unterwasserstrukturen. *Hydrographische Nachrichten*, 113:16–25, 2019.

[242] F. Niemeyer, M. Neumann, J. Albiez, A. Duda, and M. Geist. Untersuchungen zur Messgenauigkeit von Laserscannern unter Wasser am Beispiel des SeaVision 3D Laser System. In *Photogrammetrie Laserscanning Optische 3D-Messtechnik, Beiträge der Oldenburger 3D-Tage 2018, Jade Hochschule*, pages 16–25, Feb. 2018.

[243] M. Nießner, M. Zollhöfer, S. Izadi, and M. Stamminger. Real-time 3D reconstruction at scale using voxel hashing. *ACM Transactions on Graphics*, 32(6):169, Nov. 2013.

[244] D. Nistér. An efficient solution to the five-point relative pose problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6):756–770, June 2004.

[245] E. Nocerino, F. Menna, F. Fassi, F. Remondino, et al. Underwater calibration of dome port pressure housings. In *Proceedings of the 2016 European Calibration and Orientation Workshop (EuroCOW)*, volume XL-3/W4 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 127–134, Lausanne, Switzerland, Feb. 2016. Copernicus GmbH.

[246] E. Nocerino, F. Menna, and A. Gruen. Bundle adjustment with polynomial point-to-camera distance dependent corrections for underwater photogrammetry. In *Proceedings of the XXIV ISPRS Congress (2021 edition)*, volume XLIII-B2-2021 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 673–679. Copernicus GmbH, 2021.

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

[248] E. Nocerino, F. Neyer, A. Grün, M. Troyer, F. Menna, A. Brooks, A. Capra, C. Castagnetti, and P. Rossi. Comparison of diver-operated underwater photogrammetric systems for coral reef monitoring. In *Proceedings of the 2nd ISPRS International Workshop Underwater 3D Recording & Modelling: A Tool for Modern Applications and CH Recording*, volume XLII-2/W10 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 143–150, Limassol, Cyprus, May 2019. Copernicus GmbH.

[249] NovAtel. Inertial Explorer. `https://novatel.com/products/waypoint-post-processing-software/inertial-explorer`. Webpage, accessed December 22, 2022.

[250] A. e. a. Nüchter. 3DTK–The 3D Toolkit. `http://www.threedtk.de/`. Webpage, accessed December 22, 2022.

[251] A. Okamoto. Wave influences in two-media photogrammetry. *Photogrammetric Engineering and Remote Sensing*, 48(9):1487–1499, Sept. 1982.

[252] E. Olson. AprilTag: A robust and flexible visual fiducial system. In *Proceedings of the 2011 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3400–3407, Shanghai, China, May 2011. IEEE.

[253] P. Ozog and R. M. Eustice. Real-time SLAM with piecewise-planar surface models and sparse 3D point clouds. In *Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1042–1049, Tokyo, Japan, Nov. 2013. IEEE.

[254] A. Palomer, P. Ridao, and D. Ribas. Multibeam 3D underwater SLAM with probabilistic registration. *Sensors*, 16(4):560, Apr. 2016.

[255] A. Palomer, P. Ridao, and D. Ribas. Inspection of an underwater structure using point-cloud SLAM with an AUV and a laser scanner. *Journal of Field Robotics*, 36(8):1333–1344, Sept. 2019.

[256] A. Palomer, P. Ridao, D. Ribas, and J. Forest. Underwater 3D laser scanners: The deformation of the plane. In T. I. Fossen, K. Y. Pettersen, and H. Nijmeijer, editors, *Sensing and Control for Autonomous Vehicles*, volume 474 of *Lecture Notes in Control and Information Sciences*, pages 73–88. Springer, May 2017.

[257] A. Palomer Vila. *3D underwater SLAM using sonar and laser sensors*. PhD thesis, University of Girona, 2018.

[258] A. Palomer Vila, P. Ridao Rodríguez, D. Youakim, D. Ribas Romagós, J. Forest Collado, and Y. R. Petillot. 3D laser scanner for underwater manipulation. *Sensors*, 18(4):1086, Apr. 2018.

[259] N. Palomeras, P. Ridao, D. Ribas, and G. Vallicrosa. Autonomous I-AUV docking for fixed-base manipulation. In *Proceedings of the 19th World Congress of the International Federation of Automatic Control*, volume 47 (3) of *IFAC Proceedings Volumes*, pages 12160–12165, Cape Town, South Africa, Aug. 2014. Elsevier.

[260] K. Pathak, A. Birk, and N. Vaskevicius. Plane-based registration of sonar data for underwater 3D mapping. In *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4880–4885, Taipei, Taiwan, Oct. 2010. IEEE.

[261] L. Paull, S. Saeedi, M. Seto, and H. Li. AUV navigation and localization: A review. *IEEE Journal of Oceanic Engineering*, 39(1):131–149, Dec. 2013.

[262] M. Prats, J. J. Fernández, and P. J. Sanz. Combining template tracking and laser peak detection for 3D reconstruction and grasping in underwater environments. In *Proceedings of the 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 106–112, Vilamoura-Algarve, Portugal, Oct. 2012. IEEE.

[263] T. Qin, P. Li, and S. Shen. VINS-Mono: A robust and versatile monocular visual-inertial state estimator. *IEEE Transactions on Robotics*, 34(4):1004–1020, July 2018.

[264] T. Qin and S. Shen. Online temporal calibration for monocular visual-inertial systems. In *Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3662–3669, Madrid, Spain, Oct. 2018. IEEE.

[265] X. Quan and E. S. Fry. Empirical equation for the index of refraction of seawater. *Applied Optics*, 34(18):3477–3480, June 1995.

[266] D. Rebikoff. History of underwater photography. *Photogrammetric Engineering*, 33(8):897–904, 1967.

[267] V. Reijgwart, A. Millane, H. Oleynikova, R. Siegwart, C. Cadena, and J. Nieto. Voxgraph: Globally consistent, volumetric mapping using signed distance function submaps. *IEEE Robotics and Automation Letters*, 5(1):227–234, Nov. 2019.

[268] D. Ribas, P. Ridao, J. D. Tardós, and J. Neira. Underwater SLAM in man-made structured environments. *Journal of Field Robotics*, 25(11-12):898–921, July 2008.

[269] A. Richardson, J. Strom, and E. Olson. AprilCal: Assisted and repeatable camera calibration. In *Proceedings of the 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1814–1821, Tokyo, Japan, Nov. 2013. IEEE, IEEE.

[270] K. Rinner. Abbildungsgesetz und Orientierungsaufgaben in der Zweimedienphotogrammetrie. *Sonderheft 5 der Österreichischen Zeitschrift für Vermessungswesen*, 5:1–46, 1948.

[271] P. Risholm, T. Kirkhus, J. T. Thielemann, and J. Thorstensen. Adaptive structured light with scatter correction for high-precision underwater 3D measurements. *Sensors*, 19(5):1043, Mar. 2019.

[272] P. Risholm, A. Mohammed, T. Kirkhus, S. Clausen, L. Vasilyev, O. Folkedal, Øistein Johnsen, K. H. Haugholt, and J. Thielemann. Automatic length estimation of free-swimming fish using an underwater 3D range-gated camera. *Aquacultural Engineering*, 97:102227, Mar. 2022.

[273] P. Risholm, J. Thorstensen, J. T. Thielemann, K. Kaspersen, J. Tschudi, C. Yates, C. Softley, I. Abrosimov, J. Alexander, and K. H. Haugholt. Real-time super-resolved 3D in turbid water using a fast range-gated CMOS camera. *Applied Optics*, 57(14):3927–3937, May 2018.

[274] R. Rofallski and T. Luhmann. An efficient solution to ray tracing problems in multimedia photogrammetry for flat refractive interfaces. *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 90:37–54, Mar. 2022.

[275] R. Rofallski, F. Menna, E. Nocerino, and T. Luhmann. An efficient solution to ray tracing problems for hemispherical refractive interfaces. In *Proceedings of the XXIV ISPRS Congress (2022 edition)*, volume V-2-2022 of *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 333–342, Nice, France, June 2022. Copernicus GmbH.

[276] R. Rofallski, C. Tholen, P. Helmholz, I. Parnum, and T. Luhmann. Measuring artificial reefs using a multi-camera system for unmanned underwater vehicles. In *Proceedings of XXIV ISPRS Congress (2020 edition)*, volume XLIII-B2-2020 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 999–1008. Copernicus GmbH, 2020.

[277] C. Roman, G. Inglis, and J. Rutter. Application of structured light imaging for high resolution mapping of underwater archaeological sites. In *Proceedings of OCEANS'10 IEEE SYDNEY*, Sydney, Australia, 2010. IEEE.

[278] C. Roman and H. Singh. Improved vehicle based multibeam bathymetry using sub-maps and SLAM. In *Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 3662–3669, Edmonton, Canada, Aug. 2005. IEEE.

[279] F. Romero-Ramirez, R. Muñoz-Salinas, and R. Medina-Carnicer. Speeded up detection of squared fiducial markers. *Image and Vision Computing*, 76:38–47, Aug. 2018.

[280] A. Sarafraz and B. K. Haus. A structured light method for underwater surface reconstruction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114:40–52, Apr. 2016.

[281] H. Sardemann, C. Mulsow, and H.-G. Maas. Accuracy analysis of an oblique underwater laser lightsheet triangulation system. *PFG–Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 90:3–18, Feb. 2022.

[282] M. Sauer. *Mixed-Reality for Enhanced Robot Teleoperation*. PhD thesis, Julius Maximilian University of Würzburg, Faculty of Mathematics and Computer Science, 2011.

[283] K. Saylam, R. A. Brown, and J. R. Hupp. Assessment of depth and turbidity with airborne Lidar bathymetry and multiband satellite imagery in shallow water bodies of the Alaskan North Slope. *International Journal of Applied Earth Observation and Geoinformation*, 58:191–200, June 2017.

[284] L. Schmid, J. Delmerico, J. L. Schönberger, J. Nieto, M. Pollefeys, R. Siegwart, and C. Cadena. Panoptic Multi-TSDFs: a flexible representation for online multi-resolution volumetric mapping and long-term dynamic scene consistency. In *Proceedings of the 2022 International Conference on Robotics and Automation (ICRA)*, pages 8018–8024, Philadelphia, USA, May 2022. IEEE.

[285] C. Schneider. 3-D Vermessung von Oberflächen und Bauteilen durch Photogrammetrie und Bildverarbeitung. *Proceedings of IDENT/VISION*, pages 14–17, 1991.

[286] F. Scholten, J. Oberst, K.-D. Matz, T. Roatsch, M. Wählisch, E. J. Speyerer, and M. S. Robinson. GLD100: The near-global lunar 100 m raster DTM from LROC WAC stereo image data. *Journal of Geophysical Research: Planets*, 117(E12), Mar. 2012.

[287] T. Schops, V. Larsson, M. Pollefeys, and T. Sattler. Why having 10,000 parameters in your camera model is better than twelve. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2535–2544, Seattle, USA, June 2020. IEEE.

[288] R. Schwarz, G. Mandlburger, M. Pfennigbauer, and N. Pfeifer. Design and evaluation of a full-wave surface and bottom-detection algorithm for LiDAR bathymetry of very shallow waters. *ISPRS Journal of Photogrammetry and Remote Sensing*, 150:1–10, Apr. 2019.

[289] A. Sedlazeck and R. Koch. Perspective and non-perspective camera models in underwater imaging–overview and error analysis. In F. Dellaert, J. Frahm, M. Pollefeys, L. Leal-Taixé, and B. Rosenhahn, editors, *Outdoor and large-scale real-world scene analysis*, volume 7474 of *Lecture Notes in Computer Science*, pages 212–242. Springer, 2012.

[290] M. Shah, R. D. Eastman, and T. Hong. An overview of robot-sensor calibration methods for evaluation of perception systems. In *Proceedings of the Workshop on Performance Metrics for Intelligent Systems (PerMIS)*, pages 15–20, College Park, USA, Mar. 2012.

[291] J. Shan. Relative orientation for two-media photogrammetry. *The Photogrammetric Record*, 14(84):993–999, Oct. 1994.

[292] M. She, D. Nakath, Y. Song, and K. Köser. Refractive geometry for underwater domes. *ISPRS Journal of Photogrammetry and Remote Sensing*, 183:525–540, Jan. 2022.

[293] M. She, Y. Song, J. Mohrmann, and K. Köser. Adjustment and calibration of dome port camera systems for underwater vision. In G. Fink, S. Frintrop, and X. Jiang, editors, *Proceedings of the 41th German Conference on Pattern Recognition (DAGM GCPR)*, volume 11824 of *Lecture Notes in Computer Science*, pages 79–92. Springer, Springer, Sept. 2019.

[294] M. Sheehan, A. Harrison, and P. Newman. Self-calibration for a 3D laser. *The International Journal of Robotics Research*, 31(5):675–687, Apr. 2012.

[295] M. Shortis. Calibration techniques for accurate measurements by underwater camera systems. *Sensors*, 15(12):30810–30826, Dec. 2015.

[296] M. Shortis. Camera calibration techniques for accurate measurement underwater. In *3D Recording and Interpretation for Maritime Archaeology*, volume 31 of *Coastal Research Library*, chapter 2, pages 11–27. Springer, Cham, Mar. 2019.

[297] M. Shortis and E. H. D. Abdo. A review of underwater stereo-image measurement for marine biology and ecology applications. In *Oceanography and Marine Biology*, volume 47, chapter 6, pages 269–304. CRC Press, 1st edition, June 2016.

[298] M. R. Shortis and E. S. Harvey. Design and calibration of an underwater stereo-video system for the monitoring of marine fauna populations. In *Proceedings of the ISPRS Commission V Symposium: Real-Time Imaging and Dynamic Analysis*, volume XXXII Part 5 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 792–799, June 1998.

[299] B. Siciliano and O. Khatib. *Springer Handbook of Robotics*. Springer Handbooks. Springer, 2016.

[300] H. Singh, C. Roman, O. Pizarro, R. Eustice, and A. Can. Towards high-resolution imaging from underwater vehicles. *The International Journal of Robotics Research*, 26(1):55–74, Jan. 2007.

[301] H. Singh, C. Roman, L. Whitcomb, and D. Yoerger. Advances in fusion of high resolution underwater optical and acoustic data. In *Proceedings of the 2000 International Symposium on Underwater Technology*, pages 206–211, Tokyo, Japan, May 2000. IEEE.

[302] D. Skarlatos and P. Agrafiotis. A novel iterative water refraction correction algorithm for use in structure from motion photogrammetric pipeline. *Journal of Marine Science and Engineering*, 6(3):77, July 2018.

[303] D. Skarlatos, P. Agrafiotis, T. Balogh, F. Bruno, F. Castro, B. D. Petriaggi, S. Demesticha, A. Doulamis, P. Drap, A. Georgopoulos, et al. Project iMARECULTURE: Advanced VR, immersive serious games and augmented reality as tools to raise awareness and access to european underwater cultural heritage. In *Proceedings of the Euro-Mediterranean Conference (EuroMed)*, volume 10058 of *Lecture Notes in Computer Science*, pages 805–813, Nicosia, Cyprus, Oct. 2016. Springer.

[304] Y. Song, K. Köser, T. Kwasnitschka, and R. Koch. Iterative refinement for underwater 3D reconstruction: Application to disposed underwater munitions in the baltic sea. In *Proceedings of the 2nd ISPRS International Workshop Underwater 3D Recording & Modelling: A Tool for Modern Applications and CH Recording*, volume XLII-2/W10 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 181–187, Limassol, Cyprus, May 2019. Copernicus GmbH.

[305] J. Sprickerhof, A. Nüchter, K. Lingemann, and J. Hertzberg. A heuristic loop closing technique for large-scale 6D SLAM. *Automatika*, 52(3):199–222, 2011.

[306] C. Steger. An unbiased detector of curvilinear structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(2):113–125, Feb. 1998.

[307] F. Steinbacher, T. Bodmer, and R. Baran. Gewässervermessung aus der Luft – Tiefenschärfe am Bodensee und die neuen Möglichkeiten der ökologischen Bewertung von Gewässern. *Österreichische Wasser- und Abfallwirtschaft*, 67:441, Dec. 2015.

[308] F. Steinbacher, M. Pfennigbauer, M. Aufleger, and A. Ullrich. High resolution airborne shallow water mapping. In *Proceedings of the XXII ISPRS Congress*, volume XXXIX-B1 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 55–60, Melbourne, Australia, Sept. 2012. Copernicus GmbH.

[309] F. Steinbrücker, J. Sturm, and D. Cremers. Volumetric 3D mapping in real-time on a CPU. In *Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 2021–2028, Hong Kong, China, May 2014. IEEE.

[310] S. Stemmler, C. S. Werner, and A. Reiterer. Development of a time-of-flight laser scanning system for underwater applications. In C. R. B. Jr., X. Neyt, and F. Viallefont-Robinet, editors, *Proceedings of Remote Sensing of the Ocean, Sea Ice, Coastal Waters, and Large Water Regions 2019, SPIE Remote Sensing*, volume 11150, pages 168 – 173, Strasbourg, France, Oct. 2019. SPIE.

[311] M. S. J. Surtees. Bathymetric survey of flooded open cast mine workings. Bachelor's thesis, University of Southern Queensland, Faculty of Engineering and Surveying, 2009.

[312] C. M. Sword. Viable alternative mine operating system: A novel underwater robotic excavation system for flooded open-cut mines. *Energy Procedia*, 125:50–55, Sept. 2017.

[313] T. Takasu and A. Yasuda. Development of the low-cost RTK-GPS receiver with an open source program package RTKLIB. In *Proceedings of the International Symposium on GPS/GNSS*, Jeju, South Korea, Nov. 2009.

[314] C. Tan, G. Seet, A. Sluzek, X. Wang, C. T. Yuen, C. Y. Fam, and H. Y. Wong. Scattering noise estimation of range-gated imaging system in turbid condition. *Optics Express*, 18(20):21147–21154, Sept. 2010.

[315] R. T. Tan, K. Nishino, and K. Ikeuchi. Color constancy through inverse-intensity chromaticity space. *Journal of the Optical Society of America A*, 21(3):321–334, Mar. 2004.

[316] G. Telem and S. Filin. Photogrammetric modeling of underwater environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(5):433–444, Sept. 2010.

[317] S. Tetlow and J. Spours. Three-dimensional measurement of underwater work sites using structured laser light. *Measurement Science and Technology*, 10(12):1162, Dec. 1999.

[318] The ¡VAMOS! Project Consortium. Viable alternative mine operating system. `http://vamos-project.eu/`, 2019. Webpage, accessed December 22, 2022.

[319] N. Törnblom. Underwater 3D surface scanning using structured light. Master's thesis, Uppsala University, Faculty of Science and Technology, 2010.

[320] T. Treibitz, Y. Schechner, C. Kunz, and H. Singh. Flat refractive geometry. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(1):51–65, Jan. 2012.

[321] 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*, 3(4):323–344, Aug. 1987.

[322] Unity Technologies. Unity cross-platform game engine. `https://unity3d.com/`. Webpage, accessed December 22, 2022.

[323] J. van de Weijer, T. Gevers, and A. Gijsenij. Edge-based color constancy. *IEEE Transactions on Image Processing*, 16(9):2207–2214, Sept. 2007.

[324] F. A. Van den Heuvel. 3D reconstruction from a single image using geometric constraints. *ISPRS Journal of Photogrammetry and Remote Sensing*, 53(6):354–368, Dec. 1998.

[325] S. Villon, M. Chaumont, G. Subsol, S. Villéger, T. Claverie, and D. Mouillot. Coral reef fish detection and recognition in underwater videos by supervised machine learning: Comparison between deep learning and HOG+ SVM methods. In *Proceedings of the International Conference on Advanced Concepts for Intelligent Vision Systems (ACIVS)*, volume 10016 of *Lecture Notes in Computer Science*, pages 160–171, Lecce, Italy, Oct. 2016. Springer.

[326] I. Vizzo, T. Guadagnino, J. Behley, and C. Stachniss. VDBFusion: Flexible and efficient TSDF integration of range sensor data. *Sensors*, 22(3):1296, Feb. 2022.

[327] M. Vlachos, A. Calantropio, D. Skarlatos, and F. Chiabrando. Modelling colour absorption of underwater images using SFM-MVS generated depth maps. In *Proceedings of the XXIV ISPRS Congress (2022 edition)*, volume XLIII-B2-2022 of *ISPRS International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pages 959–966, Nice, France, June 2022. Copernicus GmbH.

[328] M. Vlachos and D. Skarlatos. An extensive literature review on underwater image colour correction. *Sensors*, 21(17):5690, Aug. 2021.

[329] Voyis. Insight Micro underwater laser scanner. `https://voyis.com/insight-micro/`. Webpage, accessed December 22, 2022.

[330] V. Walker, F. Vanegas, and F. Gonzalez. NanoMap: A GPU-accelerated OpenVDB-based mapping and simulation package for robotic agents. *Remote Sensing*, 14(21):5463, Oct. 2022.

[331] C.-K. Wang and W. D. Philpot. Using airborne bathymetric lidar to detect bottom type variation in shallow waters. *Remote Sensing of Environment*, 106(1):123–135, Jan. 2007.

[332] J. Wang and E. Olson. AprilTag 2: Efficient and robust fiducial detection. In *Proceedings of the 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4193–4198, Daejeon, South Korea, Oct. 2016. IEEE.

[333] Y. Wang, J. Zhang, Y. Cao, and Z. Wang. A deep CNN method for underwater image enhancement. In *Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP)*, pages 1382–1386, Beijing, China, Sept. 2017. IEEE.

[334] G.-Q. Wei and S. De Ma. Implicit and explicit camera calibration: Theory and experiments. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(5):469–480, May 1994.

[335] A. Weidemann, G. R. Fournier, L. Forand, and P. Mathieu. In harbor underwater threat detection/identification using active imaging. In *Proceedings of Photonics for Port and Harbor Security, Defense and Security*, volume 5780, pages 59–70, Orlando, USA, May 2005. SPIE.

[336] T. Weise, B. Leibe, and L. Van Gool. Fast 3D scanning with automatic motion compensation. In *Proceedings of the 2007 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Minneapolis, USA, June 2007. IEEE.

[337] J. Weng, P. Cohen, M. Herniou, et al. Camera calibration with distortion models and accuracy evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(10):965–980, Oct. 1992.

[338] R. Westaway, S. Lane, and D. Hicks. Remote survey of large-scale braided, gravel-bed rivers using digital photogrammetry and image analysis. *International Journal of Remote Sensing*, 24(4):795–815, 2003.

[339] P. Westfeld, H.-G. Maas, K. Richter, and R. Weiß. Analysis and correction of ocean wave pattern induced systematic coordinate errors in airborne LiDAR bathymetry. *ISPRS Journal of Photogrammetry and Remote Sensing*, 128:314–325, June 2017.

[340] E. Westman, I. Gkioulekas, and M. Kaess. A volumetric albedo framework for 3D imaging sonar reconstruction. In *Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9645–9651, Paris, France, May 2020. IEEE.

[341] T. Whelan, M. Kaess, M. Fallon, H. Johannsson, J. Leonard, and J. McDonald. Kintinuous: Spatially extended KinectFusion. Technical Report MIT-CSAIL-TR-2012-020, Massachusetts Institute of Technology, Computer Science and Artificial Intelligence Laboratory, July 2012.

[342] J. Wilm, O. V. Olesen, R. R. Paulsen, and R. Larsen. Correction of motion artifacts for real-time structured light. In *Proceedings of the Scandinavian Conference on Image Analysis (SCIA)*, volume 9127 of *Lecture Notes in Computer Science*, pages 142–151, Copenhagen, Denmark, June 2015. Springer.

[343] S. Winkelbach, S. Molkenstruck, and F. M. Wahl. Low-cost laser range scanner and fast surface registration approach. In *Proceedings of the Joint Pattern Recognition Symposium (DAGM)*, volume 4174 of *Lecture Notes in Computer Science*, pages 718–728, Berlin, Germany, Sept. 2006. Springer.

[344] M. Winkelman, D. Dijxhoorn, and O. Marcus. Novel solution for mining minerals: ¡VAMOS! In *Proceedings of the Dredging Summit and Expo 2018*, pages 25–28, Norfolk, USA, June 2018.

[345] B. Woźniak and J. Dera. Light absorption by suspended particulate matter (SPM) in sea water. In *Light Absorption in Sea Water*, volume 33 of *Atmospheric and Oceanographic Sciences Library*, chapter 5, pages 167–294. Springer, 2007.

[346] M. Yang, J. Hu, C. Li, G. Rohde, Y. Du, and K. Hu. An in-depth survey of underwater image enhancement and restoration. *IEEE Access*, 7:123638–123657, Aug. 2019.

[347] T. Yau, M. Gong, and Y.-H. Yang. Underwater camera calibration using wavelength triangulation. In *Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2499–2506, Portland, USA, June 2013. IEEE.

[348] K. Zaar. Zweimedienphotogrammetrie. *Sonderheft 4 der Österreichischen Zeitschrift für Vermessungswesen*, 4:1–36, 1948.

[349] L. Zagorchev and A. Goshtasby. A paintbrush laser range scanner. *Computer Vision and Image Understanding*, 101(2):65–86, Feb. 2006.

[350] Z. Zhang. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11):1330–1334, Nov. 2000.

[351] M. Zhou, R. Bachmayer, and B. deYoung. Towards autonomous underwater iceberg profiling using a mechanical scanning sonar on a underwater Slocum glider. In *Proceedings of 2016 IEEE/OES Autonomous Underwater Vehicles (AUV)*, pages 101–107, Tokyo, Japan, Nov. 2016. IEEE.

[352] H. Zhuang, Z. S. Roth, and R. Sudhakar. Simultaneous robot/world and tool/flange calibration by solving homogeneous transformation equations of the form AX=YB. *IEEE Transactions on Robotics and Automation*, 10(4):549–554, Aug. 1994.

[353] M. Žuži, J. Čejka, F. Bruno, D. Skarlatos, and F. Liarokapis. Impact of dehazing on underwater marker detection for augmented reality. *Frontiers in Robotics and AI*, 5:92, Aug. 2018.

[354] T. Łuczyński, M. Pfingsthorn, and A. Birk. The Pinax-model for accurate and efficient refraction correction of underwater cameras in flat-pane housings. *Ocean Engineering*, 133:9–22, Mar. 2017.

Die Schriftenreihe

wird vom Lehrstuhl für Informatik XVII: Robotik der Universität Würzburg herausgegeben und präsentiert innovative Forschung aus den Bereichen der Robotik und der Telematik.

Die Kombination fortgeschrittener Informationsverarbeitungsmethoden mit Verfahren der Regelungstechnik eröffnet hier interessante Forschungs- und Anwendungsperspektiven. Es werden dabei folgende interdisziplinäre Aufgabenschwerpunkte bearbeitet:

- Sensorik: Integration von Sensoren in robotische Systeme, Kalibrierung, Lokalisierung, Kartierung und Interpretation von Sensordaten in Echtzeit.

- Robotik und Mechatronik: Kombination von Informatik, Elektronik, Mechanik, Aktuatorik, Regelungs- und Steuerungstechnik, um Roboter adaptiv und flexibel ihrer Arbeitsumgebung anzupassen.

Anwendungsschwerpunkte sind u.a. mobile Roboter, Tele-Robotik, Raumfahrtsysteme und Medizin-Robotik.

Lehrstuhl Informatik XVII
Robotik
Am Hubland
D-97074 Wuerzburg

Tel.: +49 (0) 931 - 31 - 88790

andreas.nuechter@uni-wuerzburg.de
https://www.informatik.uni-wuerzburg.de/robotics

ISSN: 2940-6145 (online)
ISSN: 2940-6137 (print)
ISBN: 978-3-945459-45-4 (online)

Zitation dieser Publikation

BLEIER, M. (2023). Underwater Laser Scanning - Refractive Calibration, Self-calibration and Mapping for 3D Reconstruction. Schriftenreihe Forschungsberichte in der Robotik, Band 28. Würzburg: Universität Würzburg. DOI: 10.25972/OPUS-32269

Dissertation an der Universität Würzburg im Rahmen der Graduate School of Science and Technology