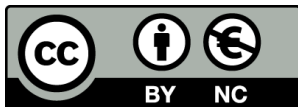


UNDERWATER 3D SENSING USING
STRUCTURED LIGHT: DEVELOPMENT OF AN
UNDERWATER LASER SCANNER AND A NON-
RIGID POINT CLOUD REGISTRATION METHOD

Miguel Castellón Sánchez



<http://creativecommons.org/licenses/by-nc/4.0/deed.ca>

Aquesta obra està subjecta a una llicència Creative Commons Reconeixement-NoComercial

Esta obra está bajo una licencia Creative Commons Reconocimiento-NoComercial

This work is licensed under a Creative Commons Attribution-NonCommercial licence



Doctoral Thesis

**Underwater 3D Sensing using Structured
Light: Development of an Underwater
Laser Scanner and a Non-Rigid Point
Cloud Registration Method**

MIGUEL CASTILLÓN SÁNCHEZ

2022



Doctoral Thesis

**Underwater 3D Sensing using Structured
Light: Development of an Underwater
Laser Scanner and a Non-Rigid Point
Cloud Registration Method**

MIGUEL CASTILLÓN SÁNCHEZ

2022

Doctoral Program in Technology

Supervised by:

PERE RIDAO RODRÍGUEZ
JOSEP FOREST COLLADO

Thesis submitted to the University of Girona
in fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

CERTIFICATE OF THESIS DIRECTION

Dr. Pere Ridao Rodríguez, director of the Institut de Recerca en Visió per Computador i Robòtica (ViCOROB) and member of the *Departament d'Arquitectura i Tecnologia de Computadors* of *Universitat de Girona*, and Dr. Josep Forest Collado, member of the *Departament d'Arquitectura i Tecnologia de Computadors* of *Universitat de Girona*,

DECLARE:

That the work entitled *Underwater 3D Sensing using Structured Light: Development of an Underwater Laser Scanner and a Non-Rigid Point Cloud Registration Method* presented by Miguel Castellón Sánchez to obtain the degree in Doctor of Philosophy has been developed under our supervision and fulfills the requirements to obtain the International Mention.

Therefore, in order to certify the aforesaid statement, we sign this document.

Girona, November 2022

Dr. Pere Ridao

Dr. Josep Forest

*A todos aquellos
que me quieren
y que quiero*

AGRADECIMIENTOS

Esta tesis no es más que una sucesión de palabras y ecuaciones que, entrelazadas con cariño, aspiran a resumir y justificar las horas que en los últimos años he pasado frente a una pantalla, sobre unos papeles o detrás de un láser. Pero tampoco es nada menos que eso: palabras y ecuaciones me han bastado, como a tantos antes que a mí, para expresar todas las ideas aquí presentadas.

Si bien es cierto que la he escrito yo, esta tesis no existiría sin Pere. Desde que contacté con él por primera vez, todo lo que he recibido de él ha sido positivo, tanto personal como académicamente: ideas, facilidades, correcciones, charlas, flexibilidad, apoyo. Después de estos años, estoy convencido de que el factor más importante a la hora de elegir qué tesis doctoral hacer no es el tema, sino el supervisor, y no creo haber encontrado nunca a nadie que reúna tantas cualidades buenas para serlo como Pere.

En este mismo sentido, estoy muy agradecido a César por haberme abierto las puertas de ASL y haberme acompañado durante esos seis intensos meses, que han supuesto un empujón que no podía haber predicho. Medio año después de volver, sigo procesando todo lo que pude aprender de él.

También tengo muchísimo que agradecer a Pep y a Albert. Al fin y al cabo, esta tesis no ha sido más que un pequeño paso en la dirección que ellos iniciaron. Su experiencia fue una inmejorable guía y me ahorró muchas horas de trabajo en el diseño y construcción del escáner. Un escáner que, por supuesto, le debe gran parte de su existencia a Lluís: suyos son el diseño y fabricación de los componentes electrónicos y mecánicos, con todos los quebraderos de cabeza que implicaron.

Me gustaría ahora hacer una mención especial para Roger. En este tiempo ha dedicado de forma altruista e incansable grandes esfuerzos a ayudar a sus compañeros, a menudo en detrimento del tiempo que podía dedicarle a su tesis. La integración del escáner en el robot y las interminables sesiones de experimentos, tanto en la piscina como en el mar, fueron durante meses un trabajo codo a codo con él. Sólo puedo estarle agradecido.

Finalmente, quiero mostrar mi gratitud a toda la gente con la que he coincidido en el CIRS por haberme ayudado en tantos aspectos: a Patryk, Eduardo, Joan, Pau, Hayat, Marta y Valerio; a Narcís y Nuno; a Rafa y Marc; a Khadidja, Klemen, Dina y Guillem; a Joseta, Mireia, Anna y Noèlia; a Tali, David, Lluís, Álex, Carles, Eduard y Jep. Igualmente, a todos con los que coincidí en ASL: Nick, Jen Jen, Paula, Francesco, Julian...

Para acabar, mi agradecimiento más especial es para mi familia y la gente que quiero, que tanto me han acompañado durante este viaje, y cuyo apoyo tanto bien me ha hecho. Y a todos los que, en algún momento, directa o indirectamente, conscientes de ello o no, me han enseñado a apreciar las palabras y las ecuaciones.

AGRAÏMENTS

Aquesta tesi no és res més que una successió de paraules i equacions que, entrelaçades amb afecte, aspiren a resumir i justificar les hores que els darrers anys he passat davant d'una pantalla, sobre uns papers o darrere d'un làser. Però tampoc no és menys que això: paraules i equacions m'han sigut prou, com a tants abans que a mi, per expressar totes les idees aquí presentades.

Si bé és cert que l'he escrita jo, aquesta tesi no existiria sense en Pere. Des que vaig contactar amb ell per primera vegada, tot el que n'he rebut ha estat positiu, tant personal com acadèmicament: idees, facilitats, correccions, converses, flexibilitat, suport. Després d'aquests anys, estic convençut que el factor més important a l'hora d'escollir quina tesi doctoral fer no és el tema, sinó el supervisor, i no crec que hagi trobat mai ningú que reuneixi tantes qualitats bones per ser-ho com en Pere.

En aquest mateix sentit, estic molt agraït a César per haver-me obert les portes d'ASL i haver-me acompanyat durant aquells sis intensos mesos, que han suposat una empenta que no podia haver predit. Mig any després de tornar, continuo processant tot el que en vaig poder aprendre.

També tinc moltíssim a agrair a en Pep i a l'Albert. Al cap i a la fi, aquesta tesi no ha estat més que un petit pas a la direcció que ells van iniciar. La seva experiència va ser una guia immillorable i em va estalviar moltes hores de treball en el disseny i construcció de l'escàner. Un escàner que, per descomptat, deu gran part de la seva existència a en Lluís: seus són el disseny i fabricació dels components electrònics i mecànics, amb tots els maldecaps que van implicar.

M'agradaria ara fer una menció especial per a en Roger. En aquest temps ha dedicat de forma altruista i incansable grans esforços a ajudar els seus companys, sovint en detriment del temps que podia dedicar-li a la seva tesi. La integració de l'escàner al robot i les interminables sessions d'experiments, tant a la piscina com al mar, van ser durant mesos una feina de costat amb ell. Només li puc estar agraït.

Finalment, vull mostrar la meva gratitud a tota la gent amb qui he coincidit al CIRS per haver-me ajudat en tants aspectes: a Patryk, Eduardo, Joan, Pau, Hayat, Marta i Valerio; a Narcís i Nuno; a Rafa i Marc; a Khadidja, Klemen, Dina i Guillem; a Joseta, Mireia, Anna i Noèlia; a Tali, David, Lluís, Àlex, Carles, Eduard i Jep. Igualment, a tots amb els que vaig coincidir a ASL: Nick, Jen Jen, Paula, Francesco, Julian...

Per acabar, el meu agraïment més especial és per a la meva família i la gent que estimo, que tant m'han acompanyat durant aquest viatge, i el suport dels quals tant m'ha ajudat. I a tots els qui, en algun moment, directament o indirectament, conscients o no d'això, m'han ensenyat a apreciar les paraules i les equacions.

ACKNOWLEDGMENTS

This thesis is nothing more than a series of words and equations that, intertwined with affection, aim at summarizing and justifying the hours that in the last years I have spent in front of a screen, over some papers or behind a laser. But it is also nothing less than that: words and equations have been enough for me, like for so many before me, to express all the ideas presented here.

Although it is me who has written it, this thesis would not exist without Pere. Since I first contacted him, everything I have received from him has been positive, both personally and academically: ideas, chats, corrections, flexibility, support. After these years, I am convinced that the most important factor when choosing which doctoral thesis to do is not the subject but the supervisor, and I have never found anyone who has as many good qualities for it as Pere.

Likewise, I am very grateful to César for opening the doors of ASL to me and accompanying me during those six intense months, which have given me a boost that I could not have predicted. Half a year after coming back, I'm still processing everything I could learn from him.

I also have a lot to thank Pep and Albert. After all, this thesis has been nothing more than a small step in the direction that they started. Their experience was an excellent guide and saved me many hours of work in the design and construction of the scanner. A scanner that, of course, owes a large part of its existence to Lluís: he did the design and manufacture of the electronic and mechanical components, with all the headaches that they implied.

I would now like to make a special mention of Roger. During this time he has altruistically and tirelessly devoted great efforts to helping his colleagues, often at the expense of the time he could devote to his thesis. The integration of the scanner in the robot and the endless experiment sessions, both in the pool and in the sea, were for months a hand-in-hand work with him. I can only be grateful to him.

Finally, I want to express my gratitude to all the people I have met at CIRS, for having helped me in so many ways: to Patryk, Eduardo, Joan, Pau, Hayat, Marta and Valerio; to Narcís and Nuno; to Rafa and Marc; to Khadidja, Klemen, Dina and Guillem; to Joseta, Mireia, Anna and Noèlia; to Tali, David, Lluís, Álex, Carles, Eduard and Jep. Likewise, to everyone I shared time with in ASL: Nick, Jen Jen, Paula, Francesco, Julian...

To finish, my most special thanks go to my family and the people I love, who have accompanied me so much during this journey, and whose support has done me so much good. And to all those who, at some point, directly or indirectly, aware of it or not, have taught me to appreciate words and equations.

LIST OF PUBLICATIONS

Publications in the compendium

The presented thesis is a compendium of the following research articles:

- **Miguel Castellón**, Albert Palomer, Josep Forest, and Pere Ridao. “State of the Art of Underwater Active Optical 3D Scanners”. In: *Sensors* 19.23 (2019), page 5161. DOI: [10.3390/s19235161](https://doi.org/10.3390/s19235161)
Quality index: JCR2019 Instruments & Instrumentation, Impact Factor: 3.275, Q1 (15/64).
Contributions: This publication reviews the state-of-the-art approaches present in the literature for underwater 3D laser scanning. The goal of this publication is to provide an exhaustive yet comprehensible guide of the challenges of underwater light-based 3D scanning, as well as of the proposed technological solutions and their limitations.
Author contributions: The author of this thesis was responsible for surveying the underwater 3D scanners in the literature and their components, establishing common metrics to compare the existing academic and commercial devices, and structuring and writing the survey.
- **Miguel Castellón**, Albert Palomer, Josep Forest, and Pere Ridao. “Underwater 3D scanner model using a biaxial MEMS mirror”. In: *IEEE Access* 9 (2021), pages 50231–50243. DOI: [10.1109/ACCESS.2021.3069189](https://doi.org/10.1109/ACCESS.2021.3069189)
Quality index: JCR2021 Engineering, Electrical & Electronic, Impact Factor: 3.476, Q2 (105/278).
Contributions: This article presented a ray-tracing model of an underwater laser scanner. The main novelty of this model is that it uses a two-axis mirror to steer the laser beam. The second rotational degree of freedom of the mirror can be used to project optimally-curved light shapes so that the refraction process transforms them into planes. Being able to model the light surfaces as planes rather than cones can significantly reduce the computational complexity of the 3D reconstruction.
Author contributions: The author of this thesis designed this ray-tracing model and validated it using synthetic data. He also studied the influence of the miscalibration of each of the model parameters on the 3D reconstruction error. Moreover, he found the optimal range of values in which each parameter should be in order to optimize scanning performance criteria such as field of view (FoV) or maximum

distortion of the projected light pattern at different distances. Finally, he structured and wrote the article.

- **Miguel Castellón**, Josep Forest, and Pere Ridao. “Underwater 3D scanner to counteract refraction: calibration and experimental results”. In: *IEEE/ASME Transactions on Mechatronics* (2022). DOI: [10.1109/TMECH.2022.3170504](https://doi.org/10.1109/TMECH.2022.3170504)

Quality index: JCR2021 Engineering, Electrical & Electronic, Impact Factor: 5.867, Q1 (47/278).

Contributions: This work presents an underwater 3D scanner prototype based on the model proposed in the previous article. To the best of the authors’ knowledge, the prototype presented in this paper is the first laser line scanner (LLS) that actively counteracts the refraction of the projected light in the context of underwater robotics. This prototype can achieve millimeter accuracy and scanning speeds of between 50 and 100 scanned lines per second, which is in the order of other state-of-the-art underwater laser scanners. Our approach allows for higher flexibility in the scanning pattern, potentially opening future lines of research. Its calibration routine based on numeric projection functions is much simpler than other scanners, reducing the time required for internal calibration.

Author contributions: The author of this thesis is the main developer of the prototype. He was responsible for designing the physical configuration of the hardware components, programming the microcontroller that governs the scanner, designing a simple but accurate calibration procedure, carrying out its underwater performance characterization, and writing the article.

- **Miguel Castellón**, Pere Ridao, Roland Siegwart, and César Cadena. “Linewise Non-Rigid Point Cloud Registration”. In: *IEEE Robotics and Automation Letters* 7.3 (2022), pages 7044–7051. DOI: [10.1109/LRA.2022.3180038](https://doi.org/10.1109/LRA.2022.3180038)

Quality index: JCR2021 Robotics, Impact Factor: 4.321, Q2 (11/30).

Contributions: This publication presented a novel non-rigid point cloud registration method. The main contribution of the article was the development of a probabilistic algorithm that fully exploits the characteristics of LLSs to achieve higher accuracy with a much lower computational complexity than state-of-the-art non-rigid registration methods. The approach was validated on synthetic and experimental data acquired with our underwater 3D scanner.

Author contributions: The author of this thesis is the main developer of the registration method. Among his tasks, he reviewed the state of the art concerning non-rigid registration methods, developed the mathematical formulation of the algorithm, evaluated it on synthetic and experimental data, and wrote the article.

Conference presentation: The contents of this paper were also selected for presentation at the *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems* (IROS 2022), held in Kyoto (Japan) in October 2022.

Publications derived from this thesis

The work developed in this thesis also led to the following publication:

- **Miguel Castellón**, Roger Pi, Narcís Palomeras, and Pere Ridao. “Extrinsic visual-inertial calibration for motion distortion correction of underwater 3D scans”. In: *IEEE Access* 9 (2021), pages 93384–93398. DOI: [10.1109/ACCESS.2021.3092180](https://doi.org/10.1109/ACCESS.2021.3092180)

Patents derived from this thesis

The work developed in this thesis also led to the following patent:

- **Miguel Castellón**, Pere Ridao, Albert Palomer, y Josep Forest. “Procedimiento de proyección de un haz de luz puntual sobre una superficie final de proyección y procedimiento de escaneo correspondiente”. ES202031301. Diciembre 2020.

Other publications

Parallel work at the time of this thesis led to the following publication:

- Jeygopi Panisilvam, **Miguel Castellón**, Nicholas Lawrance, and Roland Siegwart. “Conditioned deep feature consistent variational autoencoder for simulating realistic sonar images”. In: *OCEANS 2022 - Hampton Roads*. 2022

Finally, I also had the pleasure to co-supervise a master’s thesis:

- Dominic Marti. “Underwater Volumetric Occupancy Mapping with Imaging Sonar”. Master’s thesis. ETH Zurich, 2022. DOI: [20.500.11850/551563](https://doi.org/20.500.11850/551563)

ACRONYMS

AUV	autonomous underwater vehicle
CPD	Coherent Point Drift
DoF	degree of freedom
DVL	Doppler velocity log
FoV	field of view
ICP	Iterative Closest Point
INS	inertial navigation system
lidar	light detection and ranging
LLS	laser line scanner
PCL	The Point Cloud Library
ROV	remotely operated vehicle
SfM	structure from motion
SLAM	simultaneous localization and mapping
sonar	sound navigation and ranging
ToF	time of flight
UUV	unmanned underwater vehicle

Institutions

ASL	Autonomous Systems Lab
CIRS	Centre d'Investigació en Robòtica Submarina
ETH	Eidgenössische Technische Hochschule Zürich
UdG	Universitat de Girona
ViCOROB	Institut de Recerca en Visió per Computador i Robòtica

CONTENTS

Abstract	1
Resum	3
Resumen	5
1 Introduction	7
1.1 Motivation	8
1.2 Objectives	11
1.3 Context	11
1.4 Document structure	13
2 State of the Art of Underwater Active Optical 3D Scanners	15
3 Underwater 3D Scanner Model Using a Biaxial MEMS Mirror	51
4 Underwater 3D Scanner to Counteract Refraction	65
5 Linewise Non-Rigid Point Cloud Registration	75
6 Results and Discussion	85
6.1 Summary of the completed work	86
6.2 Experiments	86
6.2.1 Extrinsic calibration of the scanner	86
6.2.2 Sea experiments	88
7 Conclusions and Future Work	93
7.1 Contributions of this thesis	94
7.2 Future work	94

LIST OF FIGURES

1.1	Two different frames from the inspection of the Boreas shipwreck using the omnidirectional RGB camera developed by Bosch <i>et al.</i> [8]. The full 360° video is available on https://www.youtube.com/watch?v=ptkhwTPFvjk	8
1.2	Underwater laser scanner.	10
1.3	autonomous underwater vehicles (AUVs) developed at Centre d’Investigació en Robòtica Submarina (CIRS).	12
1.4	Infrastructure for experiments.	12
6.1	Girona 1000 autonomous underwater vehicle (AUV) equipped with our 3D laser scanner inspecting a mock-up of an off-shore structure at the Centre d’Investigació en Robòtica Submarina (CIRS) water tank. The red box delimits the scanned area considered in Fig. 6.2.	87
6.2	Combined partial dynamic scans into a single point cloud of the target underwater structure using the calibrated extrinsic parameters of the scanner. Point clouds (a) to (f) show the details A to F in (g). Figure (g) is made up of 25 scans in total. Each one of them is painted with an arbitrary colour for visualization. Please note that the scans were combined based on robot odometry only without using any registration method.	87
6.3	Experimental data gathering at sea at the harbour of Sant Feliu de Guíxols.	88
6.4	Two views of the same scanned rocks.	89
6.5	Incremental creation of the 3D reconstruction of some rocks by aggregating individual scans.	90
6.6	Top view of the autonomous seafloor survey. The total covered area was 30m × 8m at a depth of around 10 m.	90
6.7	Comparison on the degradation of the point clouds due to spurious artifacts with and without using post-processing outlier removal.	91
6.8	Comparison of how the drift accumulated by the inertial sensors of the robot affects the quality of the final 3D reconstruction.	92

ABSTRACT

ACCURATE underwater 3D perception is essential to advance towards the automation of expensive, dangerous and/or time-consuming tasks, such as the inspection, maintenance and repair of off-shore industrial sites. Accurate underwater 3D sensors can potentially have a large positive impact on the progress of tasks like object detection and semantic mapping, which are key to the development of robotic platforms capable of a higher level of abstraction. Moreover, these advances would decidedly contribute to the transition from remotely operated vehicles (ROVs) towards autonomous underwater vehicles (AUVs) in industrial operations. However, accurate underwater 3D perception is very hard to achieve because of the many physical particularities of light propagation in water, including refraction: the direction of light rays changes due to the different refraction indices of the media it travels through.

This thesis focuses on the development of a novel underwater 3D scanner and a non-rigid point cloud registration method aimed at enabling underwater 3D reconstructions with accuracies in the order of millimeters both in static and dynamic missions.

The thesis is structured according to these two main contributions, which resulted in five journal articles. The first main contribution of this thesis is designing and building an underwater 3D scanner using a 2-axis mirror. The second axis of the rotating mirror allows us to project optimally-curved scanning patterns designed to counteract refraction, so that they transform into straight lines when entering the water. This results in a decrease in computational complexity of the 3D reconstruction while maintaining millimeter accuracy. Minor contributions of this part of the thesis are the design of a ray-tracing model to study the effect of each optical component on the quality of the 3D reconstruction and the development of a simplified calibration algorithm based on numeric projection functions. The second main contribution of this thesis is the development of a non-rigid point cloud registration method that can successfully minimize the motion distortion that appears when the scanner is mounted on a moving robot.

Finally, this thesis also includes unpublished 3D reconstructions performed during missions both in the water tank at the Centre d'Investigació en Robòtica Submarina (CIRS) and at sea.

RESUM

UNA percepció submarina 3D d'alta precisió és essencial per avançar cap a l'automatització de tasques com la inspecció, el manteniment i la reparació d'estructures industrials submarines, las quals actualment comporten un elevat preu, risc i/o durada. Disposar de sensors 3D subaquàtics d'alta precisió podria accelerar el progrés de tasques com la detecció d'objectes i el mapejat semàntic, claus per al desenvolupament de plataformes robòtiques capaces d'un nivell d'abstracció més alt. A més, aquest avenç contribuiria de forma decisiva a la transició a nivell industrial des de vehicles submarins operats remotament (ROVs) cap a vehicles submarins autònoms (AUVs). No obstant això, aconseguir percepció 3D d'alta precisió sota l'aigua és difícil per les moltes particularitats físiques de la propagació de la llum a l'aigua, incloent-hi la refracció: la llum canvia de direcció a causa dels diferents índexs de refracció dels mitjans pels quals viatja.

Aquesta tesi se centra en el desenvolupament d'un nou escàner 3D submarí i un mètode de registre no-rígid de núvols de punts amb l'objectiu de crear reconstruccions submarines en 3D amb precisions a l'ordre de mil·límetres tant en missions estàtiques com dinàmiques.

La tesi s'estructura segons aquestes dues contribucions principals, las quals ens han permès publicar cinc articles en revistes científiques. La primera contribució principal és el disseny i la construcció d'un escàner 3D submarí usant un mirall de dos eixos de rotació. El segon eix del mirall ens permet projectar patrons d'escaneig corbats de manera òptima per contrarestar la refracció, de manera que es transformin en línies rectes en entrar a l'aigua. Així, podem disminuir la complexitat computacional de la reconstrucció 3D mentre mantenim una precisió mil·limètrica. Altres contribucions de menor rang d'aquesta part de la tesi són el disseny d'un model de la direcció del feix làser per estudiar l'efecte de cada component òptic en la qualitat de la reconstrucció 3D, i el desenvolupament d'un algoritme de calibració simplificat basat en funcions numèriques de projecció. La segona contribució principal és un mètode de registre no-rígid de núvols de punts que minimitza la distorsió present als escanejats quan el sensor està muntat en un robot en moviment.

Finalment, aquesta tesi també inclou reconstruccions 3D no publicades fins ara i que van ser realitzades en missions tant a la piscina del Centre d'Investigació en Robòtica Submarina (CIRS) com al mar.

RESUMEN

UNA percepción submarina 3D de alta precisión es esencial para avanzar hacia la automatización de tareas como la inspección, mantenimiento y reparación de estructuras industriales submarinas, que actualmente entrañan un elevado precio, riesgo y/o duración. Disponer de sensores 3D subacuáticos de alta precisión podría acelerar el progreso de tareas como la detección de objetos y el mapeado semántico, que son claves para el desarrollo de plataformas robóticas capaces de un mayor nivel de abstracción. Además, tal avance contribuiría de forma decisiva a la transición a nivel industrial desde vehículos submarinos operados remotamente (ROVs) hacia vehículos submarinos autónomos (AUVs). Sin embargo, conseguir percepción 3D de alta precisión bajo el agua es difícil por las muchas particularidades físicas de la propagación de la luz en el agua, incluyendo la refracción: la luz cambia de dirección debido a los diferentes índices de refracción de los medios por los que viaja.

Esta tesis se centra en el desarrollo de un novedoso escáner 3D submarino y un método de registro no-rígido de nubes de puntos con el objetivo de crear reconstrucciones submarinas en 3D con precisiones en el orden de milímetros tanto en misiones estáticas como dinámicas.

La tesis se estructura según estas dos contribuciones principales, que han resultado en cinco artículos publicados en revistas científicas. La primera contribución principal es el diseño y construcción de un escáner 3D submarino usando un espejo de dos ejes de rotación. El segundo eje del espejo nos permite proyectar patrones de escaneo curvados de forma óptima para contrarrestar la refracción, de forma que se transformen en líneas rectas al entrar en el agua. Así, podemos disminuir la complejidad computacional de la reconstrucción 3D mientras mantenemos una precisión milimétrica. Otras contribuciones de menor rango de esta parte de la tesis son el diseño de un modelo de la dirección del haz láser para estudiar el efecto de cada componente óptico en la calidad de la reconstrucción 3D, y el desarrollo de un algoritmo de calibración simplificado basado en funciones numéricas de proyección. La segunda contribución principal es un método de registro no-rígido de nubes de puntos que minimiza la distorsión presente en los escaneos cuando el sensor está montado en un robot en movimiento.

Finalmente, esta tesis también incluye reconstrucciones 3D no publicadas hasta ahora y que fueron realizadas en misiones tanto en la piscina del Centre d'Investigació en Robòtica Submarina (CIRS) como en el mar.

1

INTRODUCTION

THIS chapter summarizes the motivation behind the development of this PhD thesis. First, Section 1.1 presents a brief overview of the challenges of accurate 3D perception in the underwater environment and introduces some of the design decisions taken during the development of our laser scanner. Next, Section 1.2 states the objectives of the thesis and Section 1.3 describes the context in which this work has been developed. Finally, Section 1.4 concludes with a summary of the organization of this document.

1.1 Motivation

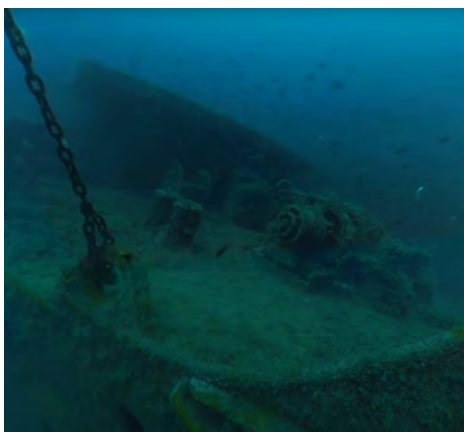
High-accuracy underwater 3D perception is very hard to achieve but it is essential in order to advance towards the automation of currently expensive, dangerous and/or time-consuming tasks, such as the inspection, maintenance and repair of off-shore industrial sites. In order to understand why underwater 3D perception is difficult, we first need to analyze the particularities of light propagation in water. We can identify two main factors by observing real underwater camera footage recorded in almost optimal visibility conditions (see Fig. 1.1):

- Both fresh and sea water carry **floating particles**, which increase water turbidity. Higher turbidity contributes to blurrier features in the image and accentuates *backscatter*: the floating particles reflect the projected light back to the camera and dazzle it, shortening the visibility range.
- Light has a **high attenuation rate** in water, which means that it loses most of its power at distances in the order of meters or tens of meters. Since the attenuation rate is wavelength-dependent, blue and green can travel further and are therefore much more predominant than red and violet (see Figure 2 of Chapter 2).

Given these special conditions, underwater 3D perception is typically done using one or more of three big sensor families¹: **sonars**, passive light-based, or active light-based.

- **Sonar** is probably the most used type of vision sensor in many underwater applications due to its long range and robustness. Its working principle is based on measuring the travelling time of ultrasounds. Since sound is a mechanical wave, it is not affected by the two challenges mentioned above and can reach hundreds of meters. Among its main drawbacks, however, its measurements have a low signal-to-noise ratio and it can only achieve resolutions in the order of tens of centimeters.

¹A more in-depth analysis of the motivation and challenges of underwater 3D scanning are presented in Sections 1 and 2 of Chapter 2, including an extensive list of relevant bibliographic references.



(a)



(b)

Figure 1.1: Two different frames from the inspection of the Boreas shipwreck using the omnidirectional RGB camera developed by Bosch *et al.* [8]. The full 360° video is available on <https://www.youtube.com/watch?v=ptkhwTPFvjk>.

- **Passive light-based** sensors are photogrammetric systems of one or more cameras. They are called passive methods because the only artificial light that they use is diffuse light to increase visibility in dark areas. They typically compute the 3D shape of the target by stereo triangulation or structure from motion (SfM). They are widely used by the underwater community because they are inexpensive and relatively easy to purchase, install and use. However, they can only measure in ranges of a few meters and they can only produce dense reconstructions in feature-rich environments. Moreover, the accuracy that they can typically achieve is in the order of centimeters.
- **Active light-based** sensors reconstruct the 3D geometry of the target by projecting a light source in known directions. Lasers are typically used because their high optical density minimizes backscatter and allows a more effective propagation than diffuse light. These sensors are usually called either **lidars** or laser line scanners (LLSs), depending on whether they compute the 3D reconstruction using time of flight (ToF) or triangulation. Like passive sensors, they usually achieve ranges of only a few meters. However, they can measure with an accuracy in the order of millimeters and can provide dense 3D point clouds in featureless scenes.

Despite their potential advantage of providing accurate and dense 3D point clouds, underwater laser scanners are still not broadly used by the underwater community because they have much shorter ranges than **sonars** and are much more expensive than cameras. In fact, the number of underwater laser scanners developed for academic or commercial purposes is very limited, as thoroughly reviewed in Chapter 2. The main contribution of this thesis is designing and creating a novel laser scanner that can potentially be useful for a wide range of tasks in the future, including inspection and object recognition and manipulation. One of the first design decisions was to use triangulation because it provides higher resolution at short ranges than **ToF**. The rest of the decisions will be explained in detail throughout this thesis.

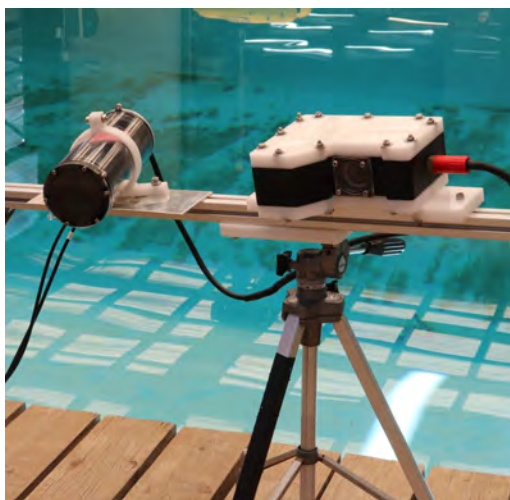
Underwater laser scanning: the effect of refraction Apart from the two challenges explained at the beginning of this chapter, a factor that affects the accuracy of underwater laser scanners is refraction. The optical components of the scanner (typically, laser module and rotating mirror) are encapsulated along with the electronics in a water-proof casing equipped with a transparent viewport, usually made of plastic or glass. As a consequence, the light projected onto the scene has to go through a double refraction process (air \rightarrow viewport material \rightarrow water), which modifies the direction of the projected light. This change of direction has a negative effect on the accuracy of the 3D reconstruction if not properly taken care of. With this in mind, we can briefly review the two main types of underwater triangulation laser scanners in the literature:

- **Profilers** directly project a fixed laser line onto the scene. Since the light goes through the viewport at a perpendicular angle, its direction is not modified. This approach is very robust and the only parameter to calibrate is the laser plane equation as seen from the camera. However, in order to scan a target, the whole scanner needs to move. They are typically used for short-range high-accuracy bathymetries. They are installed looking down at the bottom of an unmanned underwater vehicle (UUV), which covers the area of interest following a given pattern.

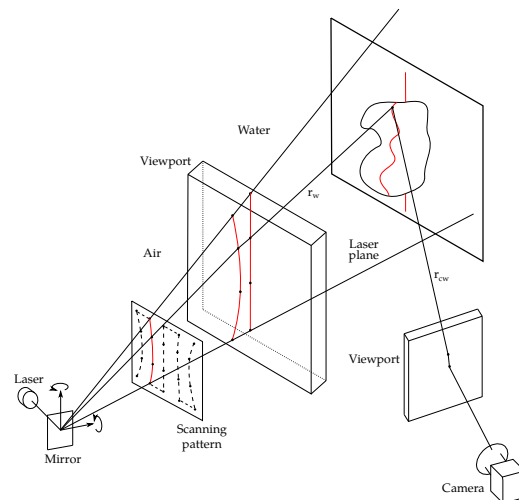
- **Scanners**, on the other hand, sweep the laser line across the scene by first projecting it onto a rotating mirror. As a consequence, the laser plane does not always go through the viewport at a perpendicular angle and therefore suffers refraction. Some authors [9, 10] try to model the change in the direction of the laser plane given the refraction index of water. However, Palomer *et al.* [11] proved that the laser plane does not only change direction but also gets distorted into an elliptic cone. Therefore, in order to maintain high accuracy in the 3D reconstruction, the calibration and the reconstruction must identify and use the equations of cones and not planes. This makes the calibration more cumbersome and the triangulation more computationally expensive.

In this thesis, we decided to develop a new scanner because it can cover a much larger area than a profiler, making it much more interesting for tasks like object recognition and manipulation. The main novelty of our scanner is that it explicitly models refraction-related distortions and counteracts them **actively** using a mirror rotating around 2 axes. The extra DoF introduced by the second rotation axis allows us to project optimally-curved lines that get transformed into planes when they enter the water. This way, the complexity introduced by refraction-related distortions is transferred from the triangulation step (which happens thousands of times per scan) to the pattern design step (which only needs to be computed once). The scanner prototype and its working scheme are shown in Fig. 1.2.

Dynamic underwater scanning Once the scanner was designed and built, and its performance was assessed, we would like to mount it on an autonomous underwater vehicle (AUV) and use it to scan while the robot is in motion. However, this is far from trivial. Unlike global-shutter cameras, which acquire all the scene information at once, LLSs need some time to sweep the laser scene across their whole field of view (FoV). If the displacement of the robot during a sweep is non-negligible, then that scan will present motion distortion. In this thesis, we approach this problem using two complementary approaches. First, we calibrate the extrinsic parameters of the scanner with respect to



(a) Prototype.



(b) Model.

Figure 1.2: Underwater laser scanner.

the robot by building and solving a landmark-based pose graph **SLAM** [5], and use that transformation along with the inertial measurements of the robot to project each line in the scan onto the world reference frame. This reduces the distortions when compared to using manually calibrated parameters. However, the accuracy of inertial systems is at least one order of magnitude worse than the scanning accuracy². This noise from the inertial sensors causes that the projected scans cannot be treated as rigid when matching. Therefore, we also developed a non-rigid point cloud registration method (Chapter 5) that achieves much better results than Iterative Closest Point (**ICP**) [12] and Coherent Point Drift (**CPD**) [13], a state-of-the-art non-rigid registration algorithm.

1.2 Objectives

With the motivations of this thesis described, we can now state that the main goal is:

To enhance the 3D sensing capabilities of unmanned underwater vehicles (UUVs) by enabling the dynamic acquisition of high-accuracy, distortion-free 3D point clouds.

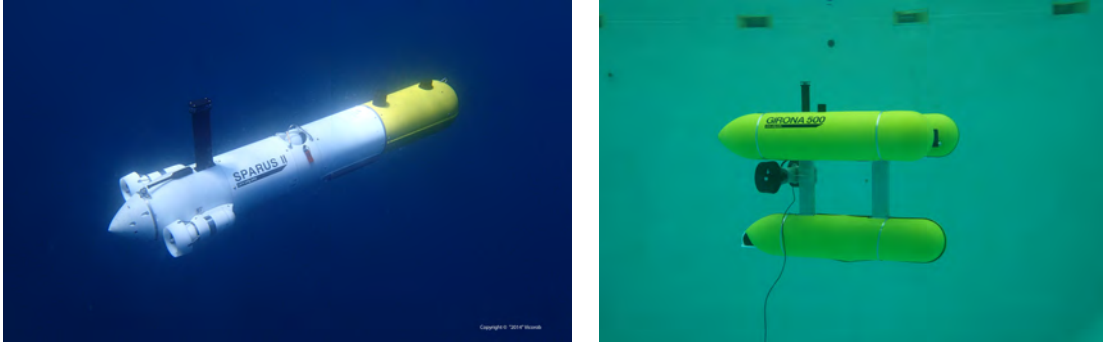
This general aim can be divided into the following objectives:

1. **Static high accuracy:** Designing and developing an underwater 3D scanning system that can achieve high accuracy in static conditions. This process was made up of the following sub-steps:
 - 1.1. **State of the art:** Reviewing the existing literature on underwater light-based 3D scanning and identifying current challenges and potentially interesting new technologies.
 - 1.2. **Model:** Designing the theoretical ray-based model for a new underwater 3D scanner that can address an open issue in the literature: pattern distortion due to refraction. Using this model to study the limitations of the chosen approach.
 - 1.3. **Prototype:** Building a prototype based on the previous model using novel light-projection technologies. Designing a simple **calibration** algorithm that accurately identifies the intrinsic parameters of the scanner. Characterizing its performance experimentally and validating the feasibility of the refraction-counteraction approach in practice.
2. **Dynamic high accuracy:** Designing and implementing an extrinsic calibration algorithm that can accurately estimate the 6-DoF pose of the scanner in the robot reference frame, enabling dynamic scanning.
3. **Minimizing distortion:** Formulating a non-rigid point cloud registration algorithm that can minimize motion distortion.

1.3 Context

The work presented in this thesis has been developed at Centre d'Investigació en Robòtica Submarina (**CIRS**), which is part of the Institut de Recerca en Visió per Computador i

²For instance, see the datasheet of iXblue's compact series of inertial navigation systems (**INSs**): <https://www.ixblue.com/wp-content/uploads/2022/01/Phins%20Compact%20Series%20-%20Datasheet.pdf>



(a) Sparus II in open sea.

(b) Girona 500 in the water tank at CIRS.

Figure 1.3: AUVs developed at CIRS.

(a) CIRS water tank.

(b) *Sextant* boat carrying the Girona 1000 AUV.**Figure 1.4:** Infrastructure for experiments.

Robòtica (ViCOROB) institute of the Universitat de Girona (UdG). Formed in 1992, this research group has become a leading team in the underwater robotics and computer vision community. One of the key elements that have allowed CIRS to have such a performance is the in-house availability of extraordinary infrastructure. On the one hand, there are two AUVs available as research platforms: the Sparus II [14] and the Girona 500 [15] (see Fig. 1.3). Sparus II can efficiently cover long distances, which makes it ideal for missions such as photogrammetric sea-bottom surveying, whereas Girona 500 can carry a heavier payload, so it is typically used for manipulation tasks. On the other hand, the lab counts with a fresh water tank of dimensions $16\text{m} \times 8\text{m} \times 5\text{m}$ (length \times width \times depth) and a crane-equipped boat named *Sextant* at St. Feliu harbor (see Fig. 1.4).

The robustness and ease to use of the AUVs along with the infrastructure allow for a relatively easy experimental data collection. In the context of this thesis, the AUV employed for experimental data collection both in the water tank and at sea was the Girona 1000, which is an evolution of the Girona 500. Among its improvements, it is rated to reach depths of up to 1000 m (instead of 500 m).

This thesis was mostly financed by the doctoral grant of the UdG IFUdG2019. It also benefited from a mobility grant IFMobUdG2019 by the UdG for a 6-month research stay at the Autonomous Systems Lab (ASL) at ETH Zurich (Switzerland). The experiments,

equipment and infrastructure resources used in this thesis have been partially funded by the following projects:

- Project GIRONA1000 (ref. DPI2017-86372-C3-2-R), funded by the Spanish Ministry of Economy, Industry and Competitiveness.
- Project EUMR (ref. H2020-INFRAIA-2017-1-twostage-731103), funded by the European Commission.
- Project ATLANTIS (ref. H2020-ICT-2019-2-871571), funded by the European Commission.
- Project PER2IAUV (ref. PID2020-115332RB-C32), funded by the Spanish Ministry of Science and Innovation.
- Project OPTIHROV (ref. PDC2021-120791-C21), funded by the Spanish Ministry of Science and Innovation.

1.4 Document structure

This document is structured into the following chapters:

- **Chapter 2:** In this chapter, we present our work *State of the Art of Underwater Active Optical 3D Scanners*. First, we make a general introduction to the context and the challenges of underwater 3D light-based perception. Then, we explain the two main reconstruction principles used by such scanners, namely triangulation and **ToF**. Later, we review the main technologies for active light projection used by underwater scanners. Finally, we make a quantitative comparative analysis of academic and commercial sensors.
- **Chapter 3:** In this chapter, we present the ray-tracing model of an underwater 3D scanner that uses a 2-axis mirror to counteract refraction, published under the title *Underwater 3D Scanner Model Using a Biaxial MEMS Mirror*. We also study the influence that the miscalibration of each parameter value would have on the accuracy of the light projection. Finally, we analyze how we should design the scanner using optimal parameter values to maximize the **FoV** and minimize the scanning pattern distortion.
- **Chapter 4:** In this chapter, we use our work *Underwater 3D Scanner to Counteract Refraction: Calibration and Experimental Results* to present the underwater 3D scanner prototype that we built based on the ray-tracing model introduced in Chapter 3. We discuss the *black-box* projection model and explain in detail the characteristics of the hardware components and their connections. This simplified projection model relates the output direction of the projected laser beam with the input mirror voltages only using polynomials. This way, the calibration process is much simpler than using the whole projection model of Chapter 3. Finally, we perform the experimental performance characterization of the prototype in the water tank at the **CIRS**.

- **Chapter 5:** In the work introduced in this chapter, *Linewise Non-Rigid Point Cloud Registration*, we present a novel non-rigid point cloud registration method that can successfully correct the distortion present in some experimental scans. It exploits the in-line rigidity of the point clouds acquired by LLSs to reduce the computational complexity while improving accuracy when compared to state-of-the-art non-rigid registration algorithms. We validated our method on synthetic data and on experimental scans gathered by our prototype LLS (see Chapter 4) mounted on the Girona1000 AUV [15]. This publication was one of the principal outcomes of the author’s research stay at ASL, ETH Zurich (Switzerland). The contents of this paper were also selected for presentation at the *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2022)*, held in Kyoto (Japan) in October 2022.
- **Chapter 6:** This chapter contains a summary of the results obtained in the context of this thesis.
- **Chapter 7:** Finally, the last chapter presents the conclusions and some guidelines for future work.

2

STATE OF THE ART OF UNDERWATER ACTIVE OPTICAL 3D SCANNERS

IN this chapter we review the state of the art of underwater 3D laser scanners. First, we make a general introduction to the context and the challenges of underwater 3D light-based perception. Then, we explain the two main reconstruction principles used by such scanners, namely triangulation and ToF. Later, we review the main technologies for active light projection used by underwater scanners. Finally, we make a quantitative comparative analysis of academic and commercial approaches.

Title: State of the Art of Underwater Active Optical 3D Scanners
Authors: **Miguel Castellón**, Albert Palomer, Josep Forest, and Pere Ridao
Journal: Sensors
Volume: 19, Number: 23, Pages: 5161, Published: 2019
DOI: [10.3390/s19235161](https://doi.org/10.3390/s19235161)
Quality index: JCR2019 Instruments & Instrumentation, Impact Factor: 3.275, Q1 (15/64)

*Review*

State of the Art of Underwater Active Optical 3D Scanners

Miguel Castellón *, Albert Palomer , Josep Forest and Pere Ridao

Computer Vision and Robotics Research Institute (VICOROB), University of Girona, 17003 Girona, Spain; albert.palo@gmail.com (A.P.); josep.forest@gmail.com (J.F.); pere.ridao.rodriguez@gmail.com (P.R.)

* Correspondence: miguel.castillon@udg.edu

Received: 28 October 2019; Accepted: 21 November 2019; Published: 25 November 2019



Abstract: Underwater inspection, maintenance and repair (IMR) operations are being increasingly robotized in order to reduce safety issues and costs. These robotic systems rely on vision sensors to perform fundamental tasks, such as navigation and object recognition and manipulation. Especially, active optical 3D scanners are commonly used due to the domain-specific challenges of underwater imaging. This paper presents an exhaustive survey on the state of the art of optical 3D underwater scanners. A literature review on light projection and light-sensing technologies is presented. Moreover, quantitative performance comparisons of underwater 3D scanners present in the literature and commercial products are carried out.

Keywords: underwater 3D laser scanners; 3D reconstruction; active 3D techniques; underwater imaging; underwater robotics

1. Introduction

Oceans cover 71% of the Earth's surface, of which 95% is still unexplored [1]. Oceanic exploration is a growing field but diving deeper than 50 m poses a huge challenge to human divers. Even though divers can use human-operated vehicles (i.e., submarines) to go deeper, there are still safety issues involved. On the other hand, robotic systems called unmanned underwater vehicles (UUVs) have been experimentally tested at full ocean depth (around 11,000 m) [2]. Their main advantage is that they can perform missions in a safer way and at a lower cost [3].

UUVs can be classified into remotely operated vehicles (ROVs) and autonomous underwater vehicles (AUVs). ROVs are connected to a vessel through a tether that transmits power and control signals for navigation or other tasks. AUVs do not need any pilot nor tether, so their operation vessels can be smaller and less costly. Their diving time is only restricted by the autonomy of the on-board batteries and the required power, so their missions can usually last for several hours [4]. There is a great abundance of marine applications currently carried out by robots, such as underwater archaeology [5,6], ocean monitoring [7,8], marine biology [9] and geology [10,11], damage assessment [12,13] and inspection, maintenance and repair (IMR) applications [14,15], to name a few.

Underwater robots require a large number of modules, each of them aimed at fulfilling a different requirement: IMUs for navigation, motors for thrust, acoustical or optical modems for communication, among others. Apart from all these, a key ability of an autonomous robotic system is sensing its environment. For UUVs it is especially important to acquire 3D data of its surroundings in order to perform tasks such as object recognition [16], inspection [17], manipulation [18] or navigation [19].

Most of the 3D sensing systems in the literature are either based on acoustic (SONAR) or light signals (LiDAR). Sonars can work at a much longer range (of up to some thousands of meters) and they are not affected by water turbidity. On the other hand, optical sensors provide a much higher

lateral resolution and refresh rate [20]. Their short-range (typically a few meters) does not limit the UUV's performance for intervention tasks since the robot needs to get close to the target.

Optical 3D sensors can be categorized as active or passive. According to Bruno et al. [21], an underwater sensor is said to use active (or structured) light when it projects light patterns onto the scene in a controlled way. These patterns can be a point, a line or more complex shapes. In active techniques, the information given by the structure of the pattern is key to reconstruct the scene in 3D. It is worth noticing that the pattern's structure is not limited to the spatial domain, but it can also be temporal (which is the base for time of flight (ToF) sensors). Active techniques determine the 3D position of the points in the environment either by ToF or by triangulation principles (see Section 3). On the other hand, passive lighting relies solely in ambient light to illuminate the scene, although artificial diffuse light may be used in dark environments [21] (This definition of active and passive lighting is characteristic of underwater sensors [21–23]. For in-air sensors, any projection of artificial light onto the scene is considered an active technique [24]). Passive techniques in underwater environments typically use stereo vision or structure from motion (SfM).

The main advantages of passive sensors such as passive stereo systems [22,25] and photometric stereo [26,27] are their low price and their theoretical high lateral resolution, which is mainly limited by forward-scattered light [28] (see Figure 1). Their main drawbacks are that they are computationally demanding and rely heavily on the target's texture to extract features from the image. Nevertheless, they are widely used by the marine research community [6,29–32]. On the other hand, active 3D sensors typically achieve a much higher point-cloud density, especially in low-contrast scenarios [5]. One of their main drawbacks is that their performance decreases under bright sunlight. Nonetheless, since sunlight is quickly attenuated by ocean water, this is usually not a major problem in relatively deep underwater environments.

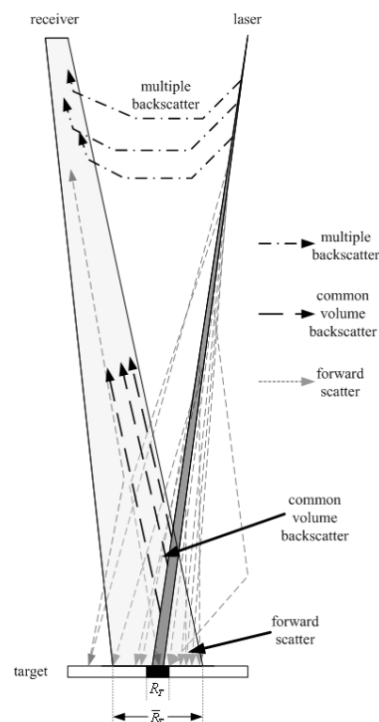


Figure 1. Possible trajectories that laser light can follow while scanning a target [28]. Forwardscatter reduces the lateral resolution. Backscatter leads time of flight (ToF) sensors to range errors.

Several surveys on underwater imaging systems have been published in the last two decades. In 2001, Jaffe et al. [33] summarized the history of underwater optical imaging and its relationship to other fields of ocean optics, focusing on technological advances in the last decade of the twentieth century. Kocak and Caimi [34] reviewed the historical progress of underwater imaging, with a special focus on the period from 2000 to 2005. Caimi et al. [35] made a survey on underwater optics in 2008, where they covered the advances in image formation and image processing methods, extended range imaging techniques, spatial coherency, and multi-dimensional imaging. Bianco et al. [22] compared the performance of two 3D underwater sensors (based on structured light and passive stereo, respectively) in 2013. In 2015, Massot-Campos and Oliver-Codina [20] presented a very complete review on underwater optical 3D reconstruction, including a quantitative comparison of performance criteria. The present survey reviews the state of the art of active underwater 3D optical sensors, focusing especially on the technologies for light projection and light sensing. Their working principles, as well as their practical limitations, are explained. Moreover, quantitative performance comparisons of underwater 3D scanners present in the literature and of commercial products are carried out.

This paper is structured as follows: the main challenges that underwater 3D sensors have to face are summarized in Section 2. Methods to reconstruct 3D scenes are explained in Section 3. Current technologies used for projecting light are gathered in Section 4. Next, the existing underwater 3D active optical sensors that can be found in the literature after 2015 are compared quantitatively in Section 5. Finally, the conclusions drawn by the authors concerning subsea 3D imaging are collected in Section 6.

2. Challenges of Underwater Imaging

One of the main challenges of underwater imaging is that light is strongly attenuated by water. This process is wavelength-dependent (see Figure 2). The visible spectrum can travel up to some hundreds of meters before being completely absorbed by water. Infrared (IR) wavelengths, on the other hand, do not propagate further than 30 cm [36,37]. Images taken in shallow waters (with a depth of less than 10 m) are less affected by water attenuation. However, there are other phenomena that degrade those images, such as flickering [4,38] and higher backscatter due to the presence of suspended particles: just like fog does above water, the floating particles and organisms randomly distributed in the water reflect the projected light back to the sensor and dazzle it [39] (see Figure 1).

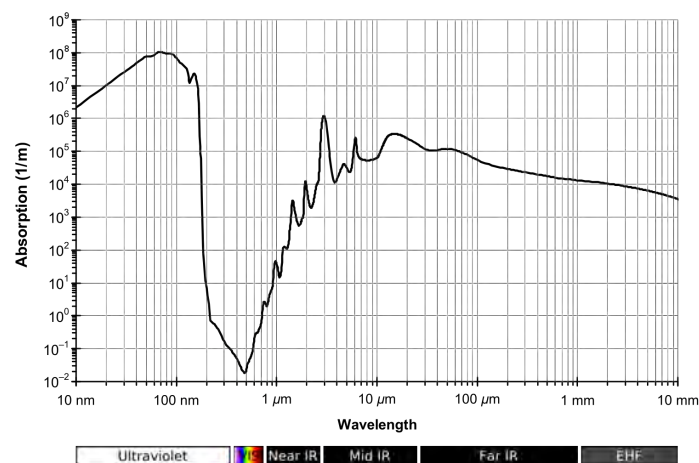


Figure 2. Water absorption spectrum of light [40].

Furthermore, vision systems are usually enclosed inside a sealed casing with a transparent viewport. This entails that light suffers a refraction process twice before arriving at the camera from

the scene, according to Snell's law [41]. This complicates further the computation of the 3D position of the observed object (see Section 2.1).

Two concepts are commonly used to characterize the underwater environment when testing an underwater sensor: turbidity and attenuation length. Turbidity is the cloudiness of a fluid caused by its suspended particles, and it is measured in NTU [42]. The attenuation length of a beam of particles (in this case, light) is defined as the distance where the intensity of the beam has dropped to $1/e$ ($\approx 37\%$) of its initial intensity [43].

All these optical differences between air and water entail that 3D sensing technology developed for in-air applications cannot be directly submerged and used for underwater tasks. The design of underwater scanners usually includes some of the following approaches to tackle the medium-specific challenges:

- The amount of light scattered back from suspended particles to the vision system can be reduced by increasing the baseline, which is the separation distance between the light source and the sensor. However, there is a limit to this increment defined by the maximum sensor size that the AUV can carry [44].
- A range-gated receiver synchronized with the laser system can also help differentiate between the backscattered noise and the light reflected by the target [45] (see Figure 3).
- Acquiring a pair of images using a polarizer at different orientations enhances the image contrast [46].
- Light wavelengths with low absorption under water can propagate longer distances. These wavelengths correspond to green or blue, but green laser sources are usually preferred because they are cheaper and more energy-efficient [47].
- Lasers sources permit a more efficient propagation when compared to diffuse light because they are highly collimated and have a high optical density [48].

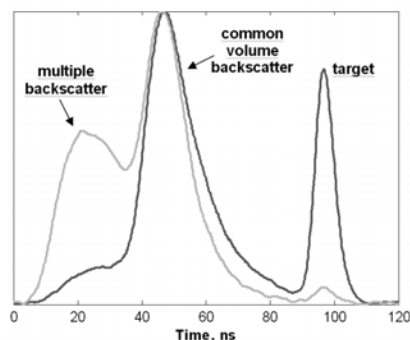


Figure 3. Typical measurement of a ToF sensor in a scattering medium, like water, for two different turbidity levels [28]. The vertical axis is the magnitude of the output signal of the ToF sensor, which relates to the intensity of the received light at each time instant. Higher turbidity (lighter curve) implies a higher peak from multiple backscatter and a very low target peak, which affects the accuracy of the measurement. If a range-gated sensor is used, its optimal opening time to accurately detect the target is around 90 ns.

2.1. Calibration

Calibration is a fundamental step in any vision system aimed at acquiring undistorted, accurate and reliable data and it usually comprises two steps. First, the intrinsic parameters of the camera (including the lens) must be computed [49]. Second, the position and orientation (extrinsic parameters) of the camera with respect to the laser projector (in case of a laser triangulation system) or with respect to the other camera (in case of stereo vision) must also be determined.

Underwater camera calibration has been widely studied in the literature. For instance, Shortis [50] presented a very complete survey on calibration techniques and considerations for underwater

photogrammetric systems, and Sedlazeck and Koch [51] compared perspective and non-perspective camera models in underwater imaging. The calibration parameters of a vision system change depending on external conditions: depth, temperature, and salinity change the refractive index of water [52–54]. On top of that, the shape of the camera housing is prone to deformations at increasing pressure levels [30].

Refraction provokes a pin-cushion distortion, which makes that the largest reconstruction errors appear at the edges of the target [55]. Due to the symmetric nature of this effect, it can be absorbed by the radial lens distortion component of the calibration parameters [50]. A practical method for calibrating a camera for underwater laser scanning is presented in [56]. However, the refraction effect entails systematic errors, since the assumption of a single projection centre for the camera (single view-point (SVP) camera model) does not hold (see Figure 4) [51]. A more complicated approach that can be followed in order to solve this issue is tracing the light rays through the refractive interfaces, such as in [57].

Underwater cameras mainly use two types of ports: flat or dome-shaped. At the expense of a more costly and difficult process of manufacture and assembly, dome ports can in principle reduce the refractive effect because there is a theoretical alignment between the interface normal and the incoming rays. However, due to small misalignments, this reduction is not usually total [51]. Performance comparisons of camera models and types of ports are done in [58,59]. Similarly, projected light also suffers this refraction process. For instance, Palomer et al. [57] demonstrated that an elliptic cone is a better geometry to describe the deformation of a laser line through a flat port in water than a plane, especially when the incidence angle between the laser and the port increases. Using an elliptic cone rather than a plane, however, makes the 3D reconstruction process more computationally demanding.

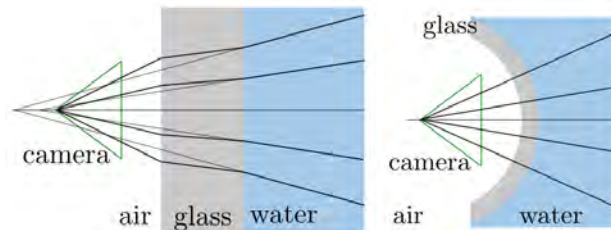


Figure 4. Refraction of light rays at flat and dome ports, respectively [51].

2.2. Open Issues

To the present date, underwater active optical 3D scanners in the literature lack two important abilities for UUVs' tasks:

- First, the data refresh rate of these sensors is too low for real-time applications in which highly dense point clouds are required. Acquisition time is important because it limits the accuracy of the 3D sensor. The relative motion during that period entails reconstruction errors. Consequently, a longer time means a larger error. One solution to mitigate this problem consists of using a very accurate and fast-refreshing navigation system, such as an inertial navigation system (INS). However, these devices have the disadvantage of being very expensive. Another approach is allowing an increase of the scanner's frequency by reducing either its field of view (FoV) or its lateral resolution. Other sensors use one-shot reconstruction so that the whole scene is captured at once, but backscatter effects and processing limitations bound the maximum lateral resolution [60]. While these approaches may be valid for certain conditions, a faster refresh rate is key to enable scanners to be mounted on realistic moving platforms.
- Second, these devices are generally not able to sense the color of the surrounding objects. Obtaining characteristics of the environment aside from its geometric description, such as the texture of each point, can be relevant in applications dealing with autonomous manipulation.

Bodenmann et al. [61] developed a laser system that enables the simultaneous capture of both structure and color from the images of a single camera and tested it for a mapping application. Nonetheless, it does not seem directly suitable for autonomous object manipulation, since the position of the laser plane with respect to the camera is fixed. Performing laser beam steering would reduce the scanning time significantly. Another existing method was presented by Yang et al. [62]. They used three lasers (RGB) to retrieve both color and 3D position of the point cloud. However, it cannot produce accurate color information as it returns three thin spectral peaks of light as opposed to a broad spectrum. As commercial products, Kraken Robotics [63] claims to have developed a working system similar to [62], which can be mounted on an UUV. It is important to note that, in general, the perceived color of an underwater scene or object is not the same as outside the water since the water absorption index of light depends heavily on its wavelength. Therefore, a color restoration process is usually needed [64–67].

There are other aspects that can be potentially improved, such as laser peak detection [68]. Several approaches have been proposed that use filters to deal with undesired lighting peaks, which are typical in underwater imaging [69,70]. Further refining the accuracy of the laser peak detection means improving the accuracy of the 3D reconstruction, which is especially relevant in media with high noise level like turbid water.

3. 3D Reconstruction Methods for Active Optical Sensors

Most of the devices at the receiving end of underwater active optical 3D sensors in the literature are based either on ToF or on triangulation methods. The general principles of both approaches are explained in Sections 3.1 and 3.2, respectively. In Section 3.3, their main characteristics are compared and the challenges of their underwater implementations are described.

3.1. Time of Flight

ToF sensors compute the depth d of a point by measuring the time Δt from the emission of a light ray until its reception, according to:

$$d = \frac{1}{2} \Delta t c_m. \quad (1)$$

The factor of $1/2$ is due to the fact that light travels a distance $2d$ until it arrives back to the sensor. Here, c_m is the speed of light in the medium where the sensor is immersed and it is given by the absolute refraction index of the medium η_m and the speed of light in vacuum $c \approx 2.9979 \times 10^8$ m/s [41]:

$$c_m = \frac{c}{\eta_m}. \quad (2)$$

Apart from positioning a point in the depth direction, its position in the other two axes needs to be determined in order to locate it in the 3D space. That spatial information can be retrieved using three different configurations:

- A punctual light source steered in 2D, along with a single detector.
- A linear light source swept in 1D, along with a 1D array of detectors.
- Diffuse light that illuminates the whole scene at once, along with a 2D array of detectors.

Most of the ToF sensors in the literature share a similar basic structure [71] (see Figure 5): a light source (transmitter), a matching detector (receiver), suitable circuitry to provide the needed power supply and control signals to transmitter and receiver, readout electronics and an analog–digital converter (ADC).

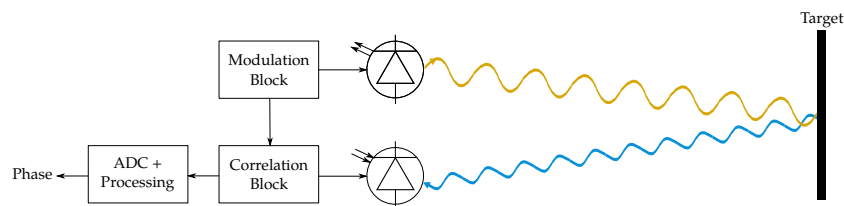


Figure 5. Scheme of a ToF 3D sensor using diffuse illumination and a 2D receiver.

The elapsed time Δt can be measured directly. However, due to the high speed of light, very accurate timers are needed: for a depth resolution of 1 mm, the accuracy of the time measurement mechanism needs to be better than 7 ps. There are some alternatives to try to obtain more accurate measurements:

- One of them is using a continuous wave (CW)-modulated light, so that the phase difference between the sent and received signals can be measured. As the modulation frequency is known, this measured phase difference corresponds to the time of flight [72].
- Another approach consists in using pulsed light. Pulsed light has a high signal to noise ratio, which makes the system more robust to background illumination. light emitting diodes (LEDs) and laser diodes are commonly used to generate pulses with repetition rates on the order of tens of kHz.

The reader is referred to [71,73,74] for more detailed explanations on ToF sensors, including data processing and calibration.

Receiver-End Technologies

Different technologies are used at the receiving end of ToF sensors. Range-gated cameras have the advantage of being more robust to backscatter effects [75]. More specifically, gated intensifiers coupled to charge-coupled device (CCD) image sensors allow for precise range gating. However, complementary metal-oxide-semiconductor (CMOS) chips achieve high frame rates and high spatial and depth resolution with a lower system complexity [76].

Apart from cameras, there are other sensing ToF technologies. Four of them are now briefly reviewed, namely PIN photodiodes, photomultiplier tubes (PMTs), avalanche photodiodes (APDs) and silicon photomultipliers (SiPMs).

- A PIN photodiode is a diode with an intrinsic semiconductor in the middle of a PN union that is sensitive to the incidence of light [77]. Its usage is rather limited due to its unity gain: only one electron is generated for each detected photon, which bounds its signal-to-noise ratio (SNR). Since conventional PIN photodiodes are much easier and cheaper to fabricate than other technologies and highly reliable all the time [78], they are used in very price-sensitive applications where gain is not a critical factor, such as timers in pulsed LiDAR [79]. Its bandwidth is up to 10 GHz [78].
- APDs can detect smaller quantities of light than PIN photodiodes, since their gain is around 10–100, which also improves their SNR. Their bandwidth is also high, of around 40 GHz [78]. They have been rarely used for underwater 3D measurements [80].
- PMTs present much larger gains of around 10^6 – 10^8 , while maintaining a high bandwidth. Their main disadvantage is their fragility and extremely high sensitivity to ambient light, which can damage the device irreversibly. They have been used in underwater robotics [81,82].
- SiPMs are composed of multiple single-photon avalanche photodiodes (SPADs), which are APDs in Geiger mode aimed at detecting single photons [83]. They are commercialized by Hamamatsu under the name multipixel photon counter (MPPC) [84]. They have a large gain of around 10^6 , although their bandwidth is lower [85]. Despite being used for in-air LiDAR sensors [86], they have not been mounted on underwater 3D scanners.

3.2. Triangulation

Active light triangulation sensors find the 3D position of a point in the scene by combining geometrical information from the light emitter and the receiver. This way, the position in space of the scanned point coincides with the intersection of two light rays: the one sent by the projector and the one going from the camera focal point through the illuminated pixel, assuming the camera has been simplified by a pin-hole camera model. Since all the needed information is known, finding the 3D position of a point comes down to solving a geometry problem. The derivations for a point laser scanner and a line laser scanner are done in Sections 3.2.1 and 3.2.2, respectively.

Regarding the used technologies, most underwater triangulation-based scanners use regular cameras protected inside a sealed housing. The difficulties that this fact entails have been explained in Section 2.1. Nevertheless, a new type of sensors called event cameras are being gradually used in in-air 3D robotic vision systems for scene reconstruction and tracking [87,88]. Instead of storing a full image frame at defined intervals, event cameras record an asynchronous sequence of per-pixel intensity changes, each with a precise time stamp [89]. Their low power consumption, high frame rate and absence of motion blur [90] postulate them as appropriate sensors for very agile robots [91].

3.2.1. Point Triangulation Scanners

The principle of calculating the depth of a point is based on the simplified 2D triangulation scheme depicted in Figure 6, where the following relationships hold for an arbitrary point P :

$$\tan \theta_C = \frac{x}{z} \quad (3)$$

$$\tan \theta_L = \frac{b-x}{z} \quad (4)$$

Substituting x from the first equation into the second one and rearranging:

$$z = \frac{b}{\tan \theta_L + \tan \theta_C}. \quad (5)$$

Therefore, calculating the z coordinate of a point requires knowing the baseline of the sensor, the angle at which the laser beam is steered and the angle that the reflected light ray makes with respect to the camera (which is given by the pixel position).

Designers of triangulation systems have to decide the geometrical configuration of its elements (regardless of them being a camera and a laser, a camera and a projector or two cameras) to comply with the requirements of depth resolution and maximum scanning range. In order to explain the concept of depth resolution geometrically, consider now that both the camera and the laser in Figure 6 are non-ideal and have finite angular resolution ($\Delta\theta_C$ and $\Delta\theta_L$, respectively). This entails that the observable FoV is discretized and measurements can only happen in the limited intersections of light rays coming out of the laser and sensed by the camera. Of all those intersections, the closest to P is P' . Their z coordinates differ by Δz , which is the depth resolution of the device at point P . The relationship of depth resolution with scan distance, baseline, and angular resolutions is given by the following equation [92]:

$$\Delta z = z - \frac{b}{\tan(\theta_L + \Delta\theta_L) + \tan(\theta_C + \Delta\theta_C)} \quad (6)$$

In order to express this relationship graphically, the x coordinate of point P is assumed to be halfway between the camera and the laser, which means that $x = b/2$ and $\theta_L = \theta_C$. Moreover, the values of the angular relationships are assumed to be $\Delta\theta_L = \Delta\theta_C = 0.1$ mrad. The resulting plot is depicted in Figure 7. The main conclusion is that both scanning at larger distances and using a shorter baseline influence negatively in the depth resolution.

Therefore, considering the mechanical design of the sensor, an approach to obtain measurements with finer depth resolution consists of increasing the baseline, which will especially benefit longer distance scans. However, there are two main limitations to this increment: the maximum size of the sensor that can be mounted on an UUV and the fact that a larger baseline entails a higher probability of occlusions in the short-range [93].

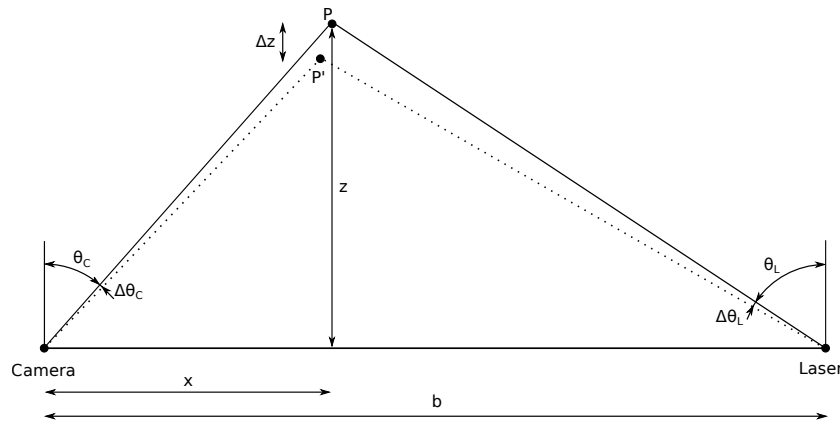


Figure 6. Scheme of the geometrical relationships used to compute the distance by point triangulation scanners.

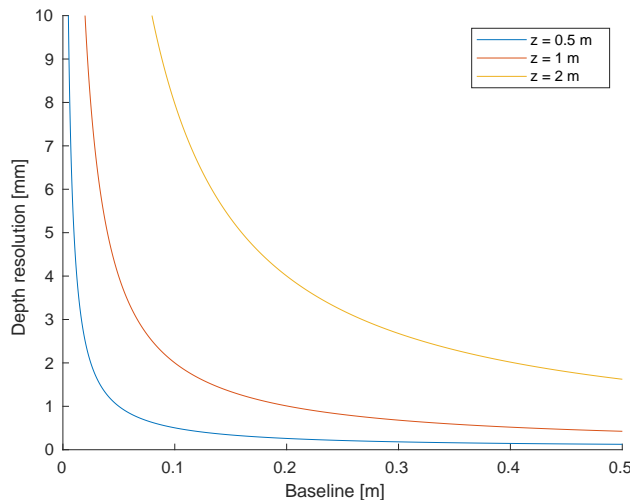


Figure 7. Depth resolution versus baseline for different scan distances.

3.2.2. Line Triangulation Scanners

The principle of laser line scanners (LLSs) is similar to point scanners, but they project a plane onto the scene. The intersection of this plane with an object creates an illuminated line that is registered by the camera. The 3D position in space of each line with respect to the camera coordinate frame can be calculated by combining information of the laser plane with the camera pixels [20,94,95].

The equation of the laser plane is assumed to be known with respect to the camera coordinate frame and can be written as $\pi \equiv Ax + By + Cz + D = 0$. The parametric equation of any point p of the line (see Figure 8) in the camera coordinate frame is

$$r(t) = \begin{bmatrix} \frac{u-c_x}{f_x} t \\ \frac{v-c_y}{f_y} t \\ t \end{bmatrix}, \quad (7)$$

where (f_x, f_y) is the camera focal length in the x and y axes, (c_x, c_y) is the position of the central pixel in the image and (u, v) is the position of one of the detected pixels in the image.

In order to find the intersection of π with $r(t)$, both equations are combined. Noting that $z = t$, the depth of p from the camera is:

$$z = \frac{-D}{A \frac{u-c_x}{f_x} + B \frac{v-c_y}{f_y} + C}. \quad (8)$$

The rest of the 3D components of p are directly obtained from Equation (7).

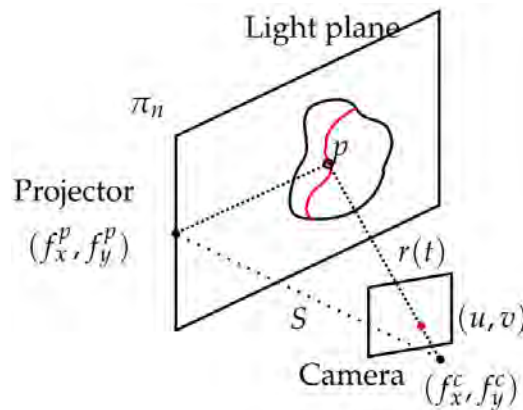


Figure 8. Line triangulation [20].

3.3. Conclusions

The main difference in performance between ToF and triangulation sensors concerns scan range and depth resolution. The depth resolution of a ToF scanner depends on the resolution of the time or phase measurement but not on the scan distance, unlike for triangulation scanners. This fact was studied by McLeod et al. [96], who identified a threshold of ranges (<1 m) where triangulation-based sensors provided higher depth accuracy. Above 2.5 m, ToF sensors were generally more accurate.

Moreover, implementing any of these measuring methods in an underwater sensor entails a number of challenges. As explained in Section 2.1, the main difficulty to accurately reconstruct the 3D position of a point has to do with the fact that every light ray travels through three different media, namely air, the viewport material, and water. This affects each method in a different way:

- In underwater triangulation sensors using flat viewports, the direction of light rays changes twice due to double refraction (see Figure 9), which can affect the accuracy of the reconstruction. At increasing incidence angles of the laser in the viewport, the laser plane transforms into an elliptic cone (see Figure 10), which makes the 3D reconstruction more computationally demanding [57].
- Along with the previous effect, underwater ToF sensors also suffer from the fact that the different propagation speeds of light (Equation (2)) affect the computation of the total travelled distance (Equation (1)). A detailed implementation of these geometrical calculations can be found in [97].

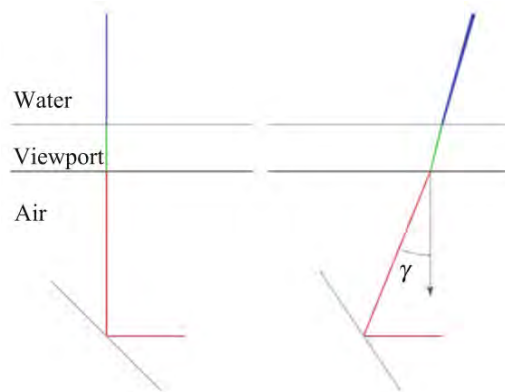


Figure 9. Change in direction of a laser line entering water due to double refraction [57].

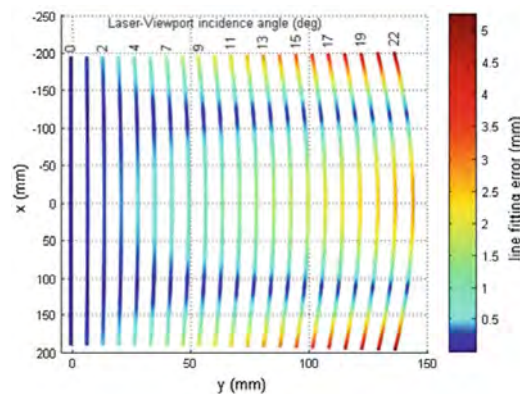


Figure 10. Projected laser lines are increasingly distorted for higher incidence angles of the laser in the viewport [57].

4. Active Light Projection Technologies

This section describes the different types of technologies for active light projection. Active optical 3D scanners project light in a known direction and/or at a known instant in time, which gives essential information for the 3D scene reconstruction. Their main advantage in underwater applications is that they help provide a more homogeneous, denser point cloud, less dependent on the scene's texture than passive methods such as stereo vision.

A big-picture classification of beam steering technologies is presented in Figure 11. First, active techniques that do not perform beam steering are explained in Section 4.1. Then, laser steering technologies are classified according to whether they involve mechanical elements or not. Both groups are explained in Sections 4.2 and 4.3, respectively. Finally, a comparison is carried out in Section 4.4.

Scanning technologies can also be classified in raster and random-access scanning. Raster scanners (such as polygon mirrors or MEMS in resonant mode) need to steer the flying spot or line through the entire FoV before they can start to scan again. On the other hand, random-access scanners can dynamically modify the scanned FoV in order to increase spatial resolution or decrease acquisition time, which makes them more flexible.

Usually, the main performance criteria of laser scanners are scanning speed, FoV, resolution, and accuracy. There are other important factors, such as optical efficiency or size.

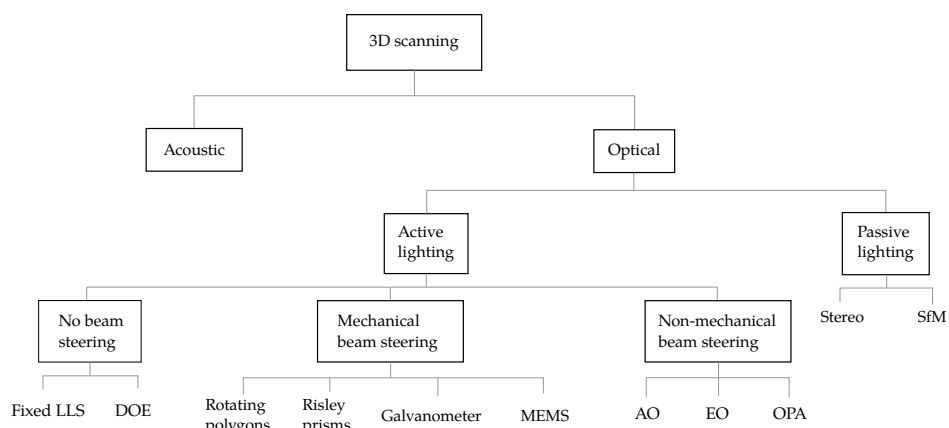


Figure 11. Classification of 3D scanning techniques.

4.1. No Beam Steering

This section gathers two light projection techniques that do not use any type of beam-steering mechanism. On the one hand, the whole scene can be illuminated at once using a homogeneizer diffractive optical element (DOE) (see Figure 12). This technique needs a 2D ToF sensor in order to resolve the 3D position of the scanned scene [75,76]. This is considered an active method because, even though the light direction is not actively controlled, its time structure provides information for the 3D reconstruction. A variation was introduced by Massot-Campos and Oliver-Codina [60], who used a special DOE to project a pattern of lines and resolved their 3D position by triangulation. Acquiring the whole scene at once has the advantage of being robust against the high dynamics of the scanned scene and against the sensor movement. However, it usually comes at the cost of reducing lateral resolution.

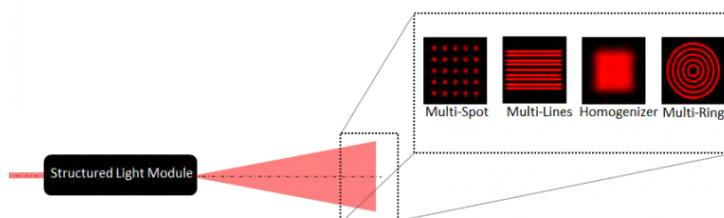


Figure 12. Different types of diffractive optical element (DOEs) [98].

On the other hand, a fixed laser line scanner (also called profiler) can be used. This setup usually consists of a laser line module and a camera. The relative position of the projected laser plane with respect to the camera is always the same. Even though the laser light is not swept across the scene, it is considered an active technique because the plane equation of the light is essential to reconstruct the 3D points. This 3D reconstruction is usually done by triangulation. This configuration makes these scanners simpler to build and calibrate, but they always need to be attached to a moving platform, usually either an UAV or a rotating tripod. Therefore, the accuracy of the final reconstruction depends greatly on the accuracy of the pose of the platform.

4.2. Mechanical Beam Steering

Mechanical beam steering is achieved by moving an object, usually a mirror, in a controlled way. A laser beam can also be steered by moving the whole sensor (for instance, a hand-held scanner). However, the smaller the inertia of the moving part, the faster the scanning can take place. Also, if the moving element is surrounded by air, the friction forces are smaller than in water. Hence,

the moving part of an underwater scanner is usually placed inside a sealed housing. In all mirror systems, the reflection angle is twice the mirror tilting angle. Consequently, mechanical scanners can achieve high deflection angles more easily. Moreover, mechanical systems are usually suitable for a wider range of laser wavelengths than solid-state beam steerers, since the mirror's surface generally has very broadband reflectance.

4.2.1. Rotating Polygon Mirror

Rotating polygon mirrors are the most simple and compact systems. They usually consist of a laser line source targeted at a 3D rotating polygon, whose faces are very reflective. Another possible configuration is using a laser point source and a rotating mirror that can also be tilted on another axis (see Figure 13). This way, other light patterns other than straight lines can be projected.

In any case, the FoV is inversely proportional to the number of faces of the polygon: the more faces, the smaller the scan angle. The deflection angle can be of up to 120° . Because of this multi-mirror characteristic, the polygon does not need to return to its initial position in order to start a new scan, it simply keeps rotating and the following face is used. Therefore, the polygon does not need to change the direction of rotation, which allows for very high rotating speeds and very short idle times. However, they are rather bulky and only allow raster scanning. Polygon mirrors have been rarely used in underwater 3D scanners [82].

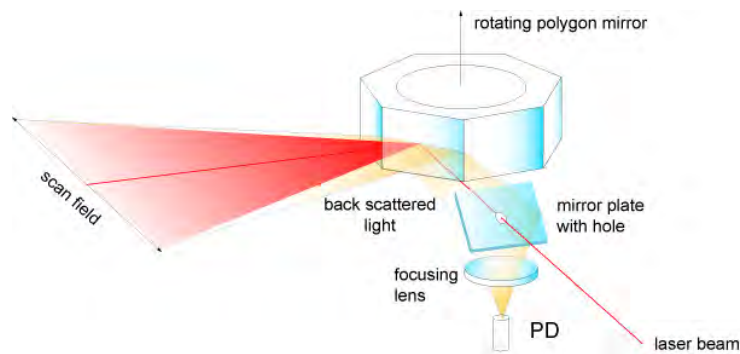


Figure 13. Rotating polygon mirror scanner [99].

4.2.2. Risleys or Wedge Prisms

Risley prisms deflect the light beam by rotating one wedge prism with respect to the other, as shown in Figure 14. Light passes through both prisms, experimenting with a double refraction process. When the prisms are rotated such that the normal of their refractive faces are parallel, the direction of the outgoing light beam is the same as the incoming (Figure 14a). However, when one of them is rotated and their normals are not parallel anymore, the outgoing light beam is deflected (Figure 14b).

This mechanism results in a very compact and robust scanners and achieves deflection angle of up to 120° . However, their main disadvantage is that they need large rotations to obtain a significant beam deflection, especially when scanning objects close to the optical axis [100].

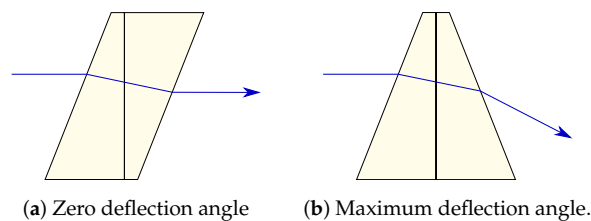


Figure 14. Risley prisms.

4.2.3. Galvanometer

Galvanometer scanners use a small electromechanical element attached to a flat mirror that can quickly turn it in small-angle increments [101]. There is an inherent trade-off between the mirror size and the maximum angular velocity, which means that there is a trade-off between the laser beam size and the scanning speed [79]. Usually, they achieve deflection angles of around 80° , and their scanning frequency is below 50 Hz. Galvanometers have been used in underwater 3D scanners [48,92,102,103].

4.2.4. MEMS Micromirrors

microelectromechanical systems (MEMS) micromirrors are very similar to galvanometers, but at a smaller scale. They can work on two regimes: linear and resonant. In linear operation, they typically achieve frequencies of around 1 kHz and deflection angles of up to 30° . In resonant mode, however, they can work at higher frequencies and deflection angles by exciting one of the mirror's axis at its resonance frequency. Nonetheless, this resonant mode is raster scanning, which means that the scene is scanned from beginning to end and arbitrary directions cannot be projected. Consequently, they cannot dynamically modify the scanned FoV in order to increase spatial resolution or decrease acquisition time. Due to their small inertia, MEMS consume less power and perform better at high resonant frequency than polygon or galvanometric scanners [104]. For more exhaustive studies on MEMS micromirrors, the reader is referred to [104–106]. MEMS micromirrors can be assembled in a number of configurations:

- Single biaxial MEMS scanner (also called 2D or flying spot). It consists of a single mirror that can be tilted around two axes (see Figure 15). The eigenfrequencies of the two axes are different so that they can perform resonant raster scanning at one of the natural frequencies.
- 1-dimensional array of MEMS micromirrors. It consists of several uniaxial or biaxial MEMS micromirrors, such as the one developed by Preciseley [107]. Another type of 1D array is the grating light valve (GLV)TM of Silicon Light MachinesTM [108]. They act as spatial light modulators (SLMs), controlling the amount of light projected at each location of a light line. They are mostly used for displays and projectors [109].
- 2-dimensional matrix of MEMS micromirrors. They are called digital micro-mirror devices (DMDs) and are normally used as SLMs in projectors. The resolution of their projection is equal to the number of micromirrors. Each of the mirrors is bistable, so they are always either *on* or *off*. However, they can achieve shades of gray by being *on* only a fraction of the total projection time of each frame. The best known commercial product is Texas Instruments' digital light processor (DLP) [110]. There are underwater 3D scanners that use DMDs to project patterns which are more complex than a line [111–113].

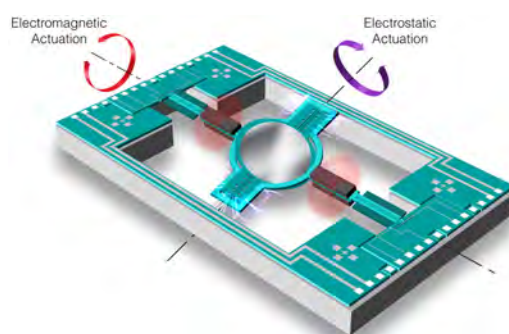


Figure 15. Maradin's MEMS 2D laser scanning mirror [114].

4.3. Non-Mechanical Beam Steering (Solid-State)

Mechanical beam scanners are robust and reliable. However, non-mechanical beam deflectors, also called solid-state deflectors, are being increasingly used in 3D scanners because of a number of reasons. Mainly, their mass-free nature allows them to achieve much higher deflection velocities. Moreover, they do not experience mechanical issues such as wear and drift. Although their maximum deflection angles are limited to a few degrees, they have high angular accuracy. Another positive characteristic is that they allow random-access scanning. For a more detailed study on electro-optic (EO) and acousto-optic (AO) laser beam scanners, the reader is referred to [115].

4.3.1. Electro-Optic Reflector (EOD)

electro-optic deflectors (EODs) perform beam steering by changing the refractive index of an optically transparent material as a result of an electric field [116] (see Figure 16). There are different types of EO technologies:

- Liquid crystal waveguides accomplish in-plane beam steering by changing the voltage on one or more prisms filled with liquid crystals. The in-plane deflection angle can be of 60° , while out-of-plane steering is of around 15° . Their response time is of less than $500 \mu\text{s}$. However, their main limitation is the size of the aperture of less than 1 cm [117].
- Electro-wetting-based systems use sealed cavities filled with two immiscible liquids, such as water and oil [118]. When a voltage difference is applied, the contact angle between the liquids is modified (see Figure 17), which deflects the laser beam. For large angles, light transmittance can drop to 30% [119]. Due to its high inertia, its maximum frequency of scene acquisition in a working scanner is around 2 Hz [120].
- potassium tantalate niobate (KTN) crystal has the maximum EO effects among existing materials. These devices are capable of very high-speed deflection (around 80 ns), but the maximum deflection angle is only of $\pm 7^\circ$ for IR wavelengths and only of $\pm 1^\circ$ for the visible spectrum [121]. Although only one-dimensional beam deflection has been achieved on a single KTN crystal, a 2D beam deflection can be obtained by lining up two deflectors appropriately. Nonetheless, this configurations is more complex and power-consuming.

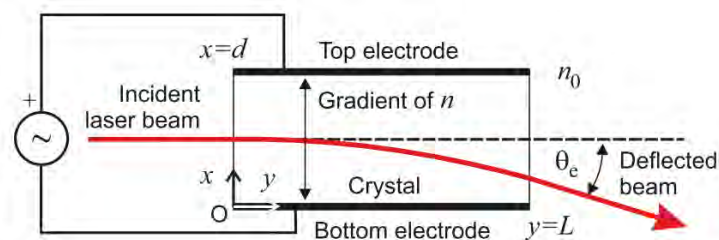


Figure 16. Electro-optic deflector (EOD) scheme [115].

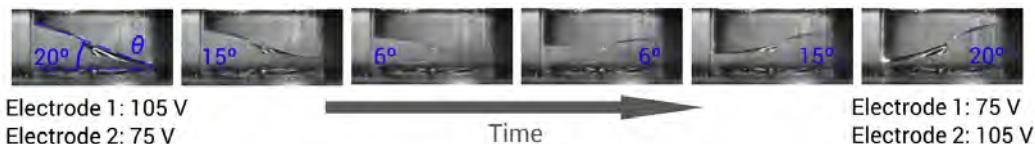


Figure 17. Electrowetting [120].

4.3.2. Acousto-Optic Deflector (AOD)

acousto-optic deflectors (AODs) are similar to EODs. However, they achieve the change of the refractive index of the material by propagating sound waves that induce a change of its density (see

Figure 18). Their maximum deflection angle is of approximately $\pm 2^\circ$ for the visible spectrum but they are slower than KTN deflectors (access time of around $10 \mu\text{s}$) [122]. These systems are commonly used in microscopy [123] and micromachining [124], but their small deflection angles reduce their applicability as 3D scanners for UAVs.

4.3.3. Optical Phased Array

An optical phased array (OPA) is an array of coherent optical sources, which means that they emit signals with a constant phase difference and the same frequency and waveform [41]. The deflection angle of the beam can be controlled by the phase and/or amplitude of the emitters [125] (see Figure 19). OPAs allow scanning frequencies of around 100 kHz. However, the insertion loss of the laser power is a drawback [126,127]. One of the few examples of OPA-based 3D scanner is [128], which aims at being implemented as LiDAR for autonomous cars. Nonetheless, these systems have not been used in real terrestrial nor aquatic vehicles to the best of the authors' knowledge.

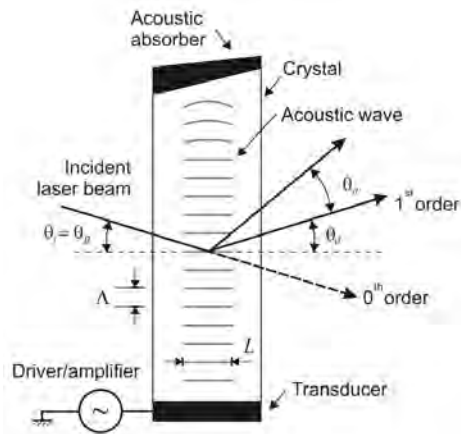


Figure 18. AOD scheme [115].

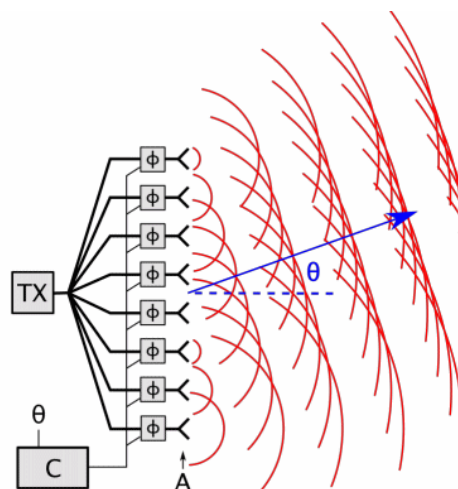


Figure 19. OPA scheme [129].

4.4. Conclusions

Performance comparisons of mechanical and non-mechanical beam steerers have been done by Bechtold et al. [130] and Römer and Bechtold [115]. The spatial resolutions (number of resolvable spots) versus the scanning speeds (rate of resolvable spots) of different devices are plotted in Figure 20. There is a clear division between mechanical and non-mechanical (or solid-state) laser scanners. In general, mechanical laser scanners provide higher angular resolution and maximum deflection angle, while non-mechanical devices allow for much higher speeds. In this respect, Ref. [115] proposed a combination of mechanical and non-mechanical elements could be highly useful in applications where high spatial resolution and high scanning speed are needed.

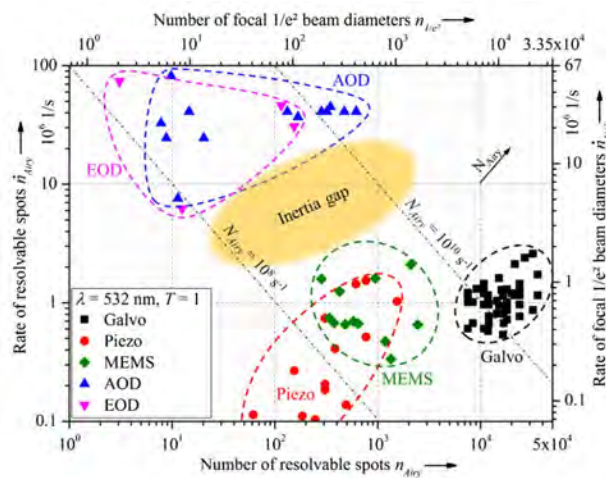


Figure 20. Number and rate of resolvable spots [130].

Regarding optical efficiency, mirror-based scanners have very small losses for a wide range of light wavelengths. On the other hand, EODs show better performance for a narrower range of wavelengths: for instance, KTN achieves a much higher FoV for IR light than for visible light. Furthermore, the efficiency of AODs is typically limited to around 70% [115].

In mechanical systems, there seems to be a trade-off between FoV and size. Polygon mirrors achieve a FoV of up to 120°, but they are large and heavy. Galvanometers usually provide a FoV of 80° and have a smaller size. MEMS can have a diameter of less than 1 mm, but they can typically only deflect the laser beam 30°.

Concerning systems that do not perform beam steering and scan the whole scene at once, they have the advantage of being robust against high dynamics of the scanned scene and against the sensor movement, at the cost of reducing lateral resolution. Moreover, systems that use diffuse light are limited to shorter measuring ranges due to a higher attenuation of light and an increased effect of backscatter.

5. Quantitative Analysis of Current Technologies

This section collects and compares active optical underwater 3D sensors present in the literature. Consequently, neither acoustic nor passive light scanners are considered for this study. Moreover, only devices developed after 2015 are taken into account since others presented before that date have already been gathered in other surveys (see Section 1).

Some systems are left out of this analysis because they follow a hybrid approach of combining active and passive light techniques. For instance, some stereo cameras make use of active light projection [113,131,132] in order to ease the feature finding and matching processes. Duda et al. [133] use an iterative combination of active light projection with SfM.

In the first subsections, the different 3D scanners are initially grouped by the type of structured light used (regardless of whether they work on ToF or triangulation principles): one-shot illumination (Section 5.1), steered line (Section 5.2), non-steered line (Section 5.3), or steered point (Section 5.4). In Section 5.5, commercial depth cameras adapted for underwater scanning are gathered. Then, a discussion comparing all the analyzed sensors is made in Section 5.6. Finally, commercial underwater solutions are collected in Section 5.7.

5.1. One-Shot Illumination

One-shot illumination consists of illuminating and acquiring the whole scene at once. Risholm et al. [76,134] and Chua et al. [75] used a DOE to flash diffuse light in every direction with no spatial structure. Other authors employ spatially-structured light: Massot-Campos and Oliver-Codina [60] used a special DOE to project 25 lines onto the surface of the scanned object, whereas Sarafraz and Haus [135] and Risholm et al. [112,136] used commercial projectors to illuminate the scene with specifically-designed, more complex patterns. Bleier and Nüchter [137] simply used two laser lines forming a cross.

The devices presented by Risholm et al. [76], Mariani et al. [134] and Chua et al. [75] are range-gated ToF sensors. This means that their light receiver opens with a delay after the light has been sent out. This is done to make them more robust against backscatter (see Figures 1 and 3). Risholm et al. [76], Mariani et al. [134] developed a peak-finding algorithm that can deal with forward scatter and at the same time can perceive distant peaks that are barely higher than peaks caused by noise. This way, they can increase the theoretical depth resolution of 18.8 cm by a factor of 20. They measured the effect of scattering in a pool and the SNR versus depth resolution in sea. Chua et al. [75] introduced a new range estimation model to reduce the effects induced by distance, target reflection and range distortion based on time slicing reconstruction and bidirectional reflection distribution function (BRDF).

Maccarone et al. [138] developed a ToF scanner based on a SPADs detector array of 192×128 . Both stationary and moving targets were imaged under a variety of underwater scattering conditions of up to 6.7 attenuation lengths.

Massot-Campos and Oliver-Codina [60] presented a triangulation-based laser sensor for underwater close-range 3D reconstructions that projected 25 lines simultaneously. It was tested at high turbidity conditions. In [139], they compared it with stereo vision. They found out that a stereo-based reconstruction is best suited for long, high altitude surveys, granted that the scene has enough texture and light. On the other hand, their structured light reconstruction worked better at short distances where accurate dimensions of an object or structure where needed. For testing, both sensors were mounted, one at a time, on a Cartesian robot and performed a lawn-moving survey over a pool.

Sarafraz and Haus [135] developed a triangulation-based scanner to simultaneously estimate both the geometric shape of the water surface and the geometric shape of underwater objects from outside the water. They chose a complex pattern of red, green and blue dots using a liquid crystal display (LCD) commercial projector.

Risholm et al. [112] projected a gray code phase stepping (GCPS) pattern using a LED with a DLP projector. In a newer version [136], they developed a multi-frequency phase stepping (MFPS) pattern, which was reportedly more accurate in-depth and more robust to turbidity than GCPS.

Bleier and Nüchter [137] developed a self-calibrating hand-held scanner. They projected two crossed line lasers and exploited coplanarity constraints to perform the 3D reconstruction. Their sensors use triangulation principles with a baseline between 0.5 m and 1 m. The depth accuracy of the underwater 3D reconstruction was not reported.

5.2. Steered Line

The scanners in this section sweep rapidly a laser line across the scene using a laser line generator and a 1-DoF galvanometer:

Chi et al. [102] first calibrated the camera's internal parameters and the relative pose of the camera and the galvanometer in air. They also obtained several different laser plane equations to calculate the galvanometer rotating axis equation. The compensation of the double refraction was done by geometric relationships, assuming that the indices of refraction of all media were accurately known, as well as the relative distances of laser, camera, and viewport. The system was experimentally tested in a water tank by measuring spherical objects whose radii and distances had been accurately measured by a coordinate measurement machine (CMM), which represented the ground truth.

Palomer et al. [48] took into account the distortion introduced by the double refraction through flat viewports. Rather than a ray-based triangulation, they used elliptical cones to speed up the reconstruction while not increasing the error. This sensor was successfully used for manipulation [18] as well as for object recognition and SLAM tasks [16].

5.3. Non-Steered Line

Non-steered laser line scanners, also called profilers, always keep the same relative position of the laser plane with respect to the camera. Therefore, the whole system must be translated and/or rotated in order to scan a scene. This is usually done by attaching the sensor to a vehicle or to a robotic arm, or by moving it by hand.

Lopes et al. [140] developed a scanner with two-line lasers (red and green). The results using the red laser were reportedly more accurate than with the green one, although no explanation of the possible causes was given. Their goal was to build a general-purpose, low-cost prototype for underwater industries, but the final cost was not reported. The system was calibrated both in dry and underwater environments using two different methods, based on the cross-ratio invariance principle and on the robust fitting of the laser line projection, respectively. The second calibration method yielded more accurate results. During the tests, the scanner was not submerged, so the scan was performed from outside the water tank. In [141], the scanner was mounted on an AUV to explore underwater mines.

Constantinou et al. [142] developed a laser scanner to measure mesh-like objects, like fish nets. The system was calibrated automatically, with the help of a calibration box. Three lasers were used in order to cover a wider area. The system was mounted on a ROV and tested in a pool and in real environment at an offshore aquaculture installation, where the relative position of a fishnet with respect to the ROV was measured.

Matos et al. [143] assumed their laser to be perpendicular to the viewport so that the laser plane was not refracted. Moreover, the two refractions that occur between air, glass, and water for the camera were approximated as one refraction between air and water. The tests were carried out with the sensor mounted on a linear slider outside the water tank measuring two fixed spheres and a dummy head, which had been calibrated with a CMM.

Bodenmann et al. [61] developed a system to simultaneously capture both structure and color using a single camera, a line laser and diffuse LED light. The line laser was fixed on an underwater vehicle pointing vertically downwards and projected a line onto the seafloor that formed a 90° angle with the forward-moving direction. Half of the camera's FoV was devoted to laser detection and the other half to color imaging. The 3D color reconstructions of the seafloor were done with subcentimeter-order resolution at a depth of more than 2000 m. The color was added to the bathymetry by back-projecting each 3D point into the illuminated area of a photo to retrieve the RGB components for that point in space. Moreover, the texture map was corrected for the color-dependent attenuation of light in water to reproduce the colors as if the water were drained. The scanner was mainly used at a distance of 2 m, because it was found to be a good compromise between area covered and image quality: imaging at a lower altitude entailed smaller covered area while scanning at altitudes exceeding 3 m meant darker image with lower contrast. Laser detection was improved by using machine learning techniques. Then, a similar system was used to do SLAM [144].

5.4. Steered Point

Steered-point or flying-spot scanners are usually ToF scanners that sweep a point across the scene, building the 3D representation of the environment point by point. They typically steer the point in 2D, although they can also do it only in 1D if mounted on a moving platform.

Imaki et al. [82] chose a laser pointer and a 2-DoF polygon mirror. This approach resulted in a rather bulky system ($\varnothing 20 \text{ cm} \times 60 \text{ cm}$) but achieved a wider horizontal FoV (120°). They used a PMT as a ToF receiver. They carried out performance tests in a pool and 3D reconstruction in field experiments in the ocean. In its newer version [145], the authors used a laser line generator and a 1-DoF polygon mirror, which slightly reduced size and weight while providing the same FoV and resolution. It was mounted on an AUV to map the seafloor.

Maccarone et al. [146] developed a time-gated ToF sensor that consisted on two 1D galvanometers and a single-pixel SPAD. They used the time-correlated single-photon counting (TCSPC) technique. An object in a water tank was scanned from the outside at different water turbidity levels, and scans at distances of up to eight attenuation lengths were achieved. The spatial resolution was only slightly affected as the scattering level was increased, even at eight attenuation lengths, with an angular resolution of approximately $60 \mu\text{rad}$ in water samples with a high scattering level. Regarding acquisition times, the authors only reported that they needed to be long.

5.5. Off-the-Shelf IR Depth Cameras

This analysis also includes three papers that proposed to use different existing IR depth cameras and make them suitable for underwater applications. They all needed to post-process the resulting data in order to correct the effect of refraction, which involved a preliminary calibration step.

Digumarti et al. [97] used Intel RealSense to develop a cheap and compact solution that enabled handheld scanning of marine life for divers. They scanned the coral reef in the ocean at a depth of 20 m. Anwer et al. [147] used Kinect v2 from above the water to measure objects in a water tank. They found out that their measurements were very noisy, so they additionally applied a median filter that could partly deal with the noise. Chourasiya et al. [148] used a Kinect v1 from outside the water to measure objects in a water tank at different turbidity levels.

5.6. Discussion

The classification of the analyzed scanners can be visualized in Table 1. Moreover, Figures 21 and 22 show the absolute and relative depth error of each group of sensors. The reported performances of the scanners are gathered in Table 2.

Table 1. Classification of the analyzed systems.

	One shot	Steered line	Non-steered line	Steered point
ToF ○	Chua [76] ● Risholm [77] ● Maccarone [140] ●			Maccarone [147] ● Imaki [83] ●
Triangulation △	Massot [60,136] ▲ Sarafraz [137] ▲ Bleier [139] ▲ Risholm [113,138] ▲	Chi [103] ▲ Palomer [48] ▲	Lopes [141] ▲ Constantinou [143] ▲ Bondenmann [61] ▲ Matos [144] ▲	

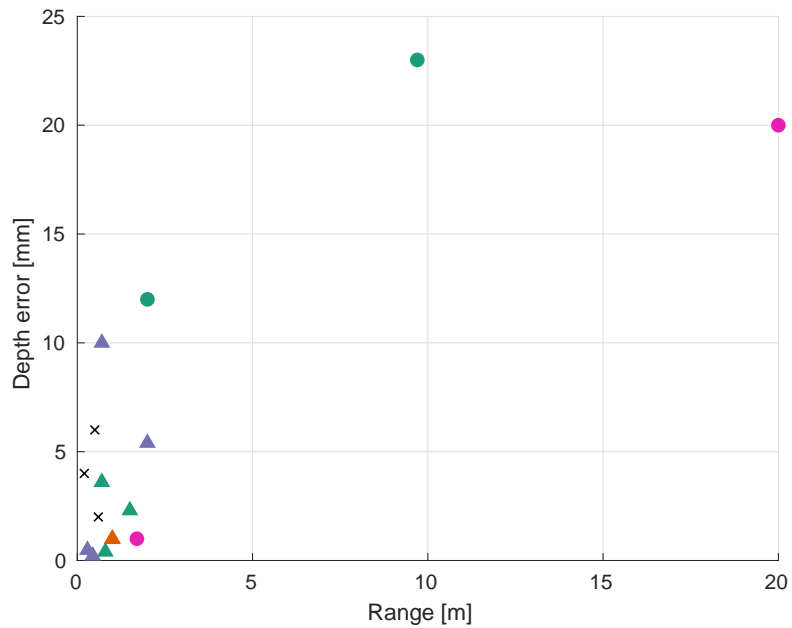


Figure 21. Reported absolute depth error of each system. See legend in Table 1. “X” represent off-the-shelf IR depth cameras [97,147,148].

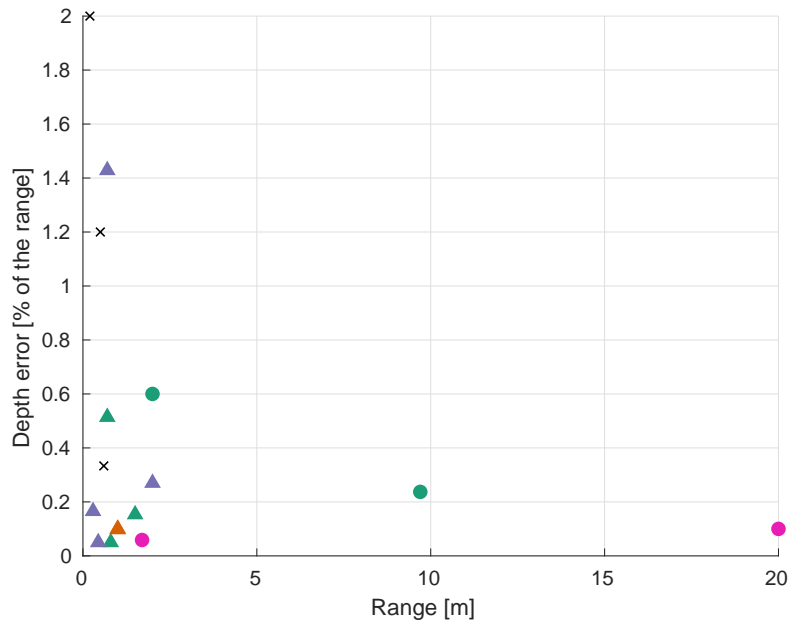


Figure 22. Reported relative depth error of each system. See legend in Table 1. “X” represent off-the-shelf IR depth cameras [97,147,148].

Table 2. Quantitative comparison of current underwater active optical 3D scanners. For an explanation of the symbol code, please see Table 1.

Reference	Type	Depth Accuracy [mm]	Range (min-max) [m]	Depth Accuracy [% of Range]	Frequency [Hz]	FoV [°]	Baseline (min-max) [mm]
Chua, 2017 [75]	●	12	2	0.60	-	-	-
Risholm, 2018 [76]	●	23	9.7	0.24	10	-	-
Maccarone, 2019 [138]	●	-	1.7	-	10	$\approx 1.7 \times 1.7$	-
Massot, 2015 [60,139]	▲	3.6	0.7	0.51	15	19	-
Sarafraz, 2016 [135]	▲	2.3	1.5	0.15	-	-	-
Bleier, 2017 [137]	▲	-	-	-	30/line	64	-
Risholm, 2018 [112,136]	▲	0.4	0.8 (0.5-2)	0.05	8	-	150
Chi, 2016 [102]	▲	1	1 (0.7-1)	0.10	-	-	-
Palmer, 2017 [48]	▲	0.98	1 (0.5 - 1.2)	0.098	0.1-6	-	300 (150-400)
Lopes, 2015 [140]	▲	0.22	0.44 (0.2-1.5)	0.05	-	-	100
Constantinou, 2016 [142]	▲	10	0.7 (0.1-1)	1.43	25/line	160	-
Bodenmann, 2017 [61]	▲	5.4	2 (0.8-3)	0.27	12/line	64	800 (800-1500)
Matos, 2019 [143]	▲	0.48	0.29 (0.15-0.29)	0.17	-	30	265
Maccarone, 2015 [146]	●	<1	1.7	0.059	-	1.6×1.6	-
Imaki, 2016 [82]	●	20	20	0.10	0.2	120×30	-
Digumarti, 2016 [97]	X	4	0.2	2.0	-	$\approx 80 \times 50$	-
Anwer, 2017 [147]	X	6	0.5 (0.35-0.65)	1.20	10	-	-
Chourasiya, 2017 [148]	X	2	0.6 (0.6-1)	0.33	30	-	75

There is a clear difference in performance between ToF and triangulation sensors. Most of the analyzed triangulation sensors achieve depth accuracies better than 5 mm, whereas the accuracies of ToF devices are in the range of a few centimeters (see Figure 21). In fact, the best absolute and relative depth accuracies are achieved by Lopes et al. [140] and Risholm et al. [136]. On the other hand, ToF sensors can work at longer ranges of up to 20 m, which makes their relative depth error at longer distances comparable to triangulation scanners working at short ranges (usually less than 2 m), as can be seen in Figure 22. Maccarone et al. [146] achieved the lowest depth error of all the ToF sensors (<1 mm), but they worked in ideal, dark conditions and with a very high acquisition time and a very small FoV. The fact that the depth accuracy of triangulation sensors decreases with longer measuring ranges, unlike ToF sensors, was expected given their different working principles (see Section 3.3).

Off-the-shelf IR depth cameras are represented with “X” marks in Figures 21 and 22. It can be seen that their high relative depth error and short-range make them not suitable for UUVs’ tasks such as mapping, inspection or manipulation. Nonetheless, Chourasiya et al. [148] could reportedly achieve 0.33% relative depth error under ideal conditions. In general, their low price and low development effort might make them useful in price-sensitive applications where the scanning range is a few tens of centimeters.

It must be noted that the authors of the presented papers tested depth accuracy in different ways. Some opted to directly scan a gauge object that was placed at a known distance from the sensor. Others placed the object on the floor, while the camera scanned at a known distance from the floor. Another used method consisted of scanning the same scene with another different sensor and then computing the average distance between point clouds, although in this case, the alternative method needs to have proved to be very accurate. Moreover, some authors report a relative depth accuracy of triangulation sensors with respect to the total object size, which is not as informative as with respect to the measuring range.

Regarding the used viewpoints, most of the analyzed sensors use flat windows. Nonetheless, some such as Imaki et al. [82] and Constantinou et al. [142] use hemispherical viewpoints. Flat viewpoints are easier to manufacture and mount, but they introduce distortions especially at high incidence angles of the laser beam in the viewport. Hemispherical or dome viewpoints do not in principle suffer from this distortion, but in practice, there are usually some misalignments (see Section 2.1).

The baseline of triangulation sensors is a very important design decision, as discussed in Section 3.2. The reported baselines are 100 mm (Lopes et al. [140]), 100 to 200 mm (Risholm et al. [112,136]), 265 mm (Matos et al. [143]) and 150 to 400 mm (Palomer et al. [48]). Bodenmann et al. [61] use a larger baseline between 800 and 1500 mm. The rest of authors did not inform on the baseline of their triangulation sensors.

There is a wide variety of applications to which each type of sensor can be applied. One-shot scanners acquire the whole scene at once, so there is no movement-related distortion. Therefore, they can provide undistorted data when they are mounted on moving vehicles or even moved manually. Consequently, applications include hand-held scanners for archaeological surveying or sea-life monitoring cameras. Non-steered line scanners are commonly used mounted on an UUV facing downwards to map the seafloor at a distance of a few meters. Steered line scanners are typically used to scan an object with a highly dense point cloud, which can be used in object recognition and manipulation tasks. In general, most scanners are quite versatile, so their applications are not strictly limited to the ones mentioned here.

5.6.1. Other Performance Criteria

Apart from depth accuracy, there are other important performance criteria. First, the number of complete scans per second measures the refresh rate of the sensor. Scanners that are able to provide more frames per second are more suitable for scenes with faster dynamics. The reported refresh rates are 0.1 Hz (Risholm et al. [136]), 0.1 Hz to 6 Hz, depending on lateral resolution (Palomer et al. [48]), 0.2 Hz (Imaki et al. [82]), 1 Hz (Digumarti et al. [97] with Intel RealSense) and 10 Hz (Risholm et al. [76,134] and Anwer et al. [147] with Kinect v2). Typical underwater manipulation tasks require a refresh rate higher than a few hertz. Others like sealife monitoring need faster frequencies. Seafloor mapping with non-steered laser line scanners usually have refresh rates of

a few tens of hertz for each line (12 Hz [61]–30 Hz [137]). Another very relevant criterion is the number of scanned points per second, which measures the trade-off between point cloud density and refresh rate. Unfortunately, many of the analyzed sensors do not report it.

Regarding the lateral resolution of ToF scanners, many authors simply report the pixel resolution of the detector. Even though it is an important parameter, the actual lateral resolution of ToF scanners is limited by the optical spreading of the beam, as depicted in Figure 1. Maccarone et al. compared a multi-pixel ToF sensor [138] with a previous single-pixel detector [146]. The multi-pixel ToF sensor provided a faster acquisition time but at the cost of a reduced spatial resolution. In the case of triangulation systems, however, subpixel laser detection algorithms can be used [149–151]. The lateral resolution of these devices employing subpixel methods is much more robust against the scattering effects of turbid water. In this sense, another relevant factor is the steerer's maximum number of resolvable spots. As explained in [115], the number of resolvable spots of a beam-steering mechanism is the factor by which the maximum deflection angle exceeds the beam divergence angle. It corresponds to the number of independent spots that can be addressed across the maximum deflection angle, so it constitutes an upper limit to the lateral resolution of the scanner. It is an objective parameter to allow a comparison of the maximum deflection angles of different technologies since it is invariant with respect to imaging optics.

Another relevant parameter is the scanner's FoV. Wide FoVs are usually preferred because they can in principle provide more information about the environment. However, a larger FoV entails a longer acquisition and processing time or a reduced lateral resolution. Moreover, it also involves higher incidence angles on the viewport, which introduces error if a flat viewport is used. The FoV of laser line modules depends on the fan angle induced by the optical elements, and it can range from 30° [143] to 160° [142].

5.6.2. Turbidity

Apart from the sensor's performance, water turbidity is one of the main challenges of underwater scanning (see Section 2), and it has various effects. On the one hand, Risholm et al. [112] report a worsening in-depth precision for higher turbidity levels. On the other hand, Massot-Campos and Oliver-Codina [60] states that the effect of higher turbidity on their sensor is a major decrease in the number of detected points due to a reduction of the maximum scanning range. Consequently, only a sparse reconstruction can be performed, but they claim that this sparse reconstruction does not lose accuracy. Moreover, the performance of ToF sensors at high turbidity levels depends strongly on the reflectivity of the scanned object. SNR is much lower for non-reflective, dark surfaces, which introduces errors in the 3D reconstruction [112].

The sensor presented by Risholm et al. [76] is capable of estimating depth at distances of over 4.5 attenuation lengths when imaging high albedo (highly reflective) targets at low attenuation lengths of less than 2 m. In their case, the attenuation length of the water is measured with a specifically designed sensor. Mariani et al. [134] states that water attenuation is not the limiting factor in very clear water with attenuation lengths larger than 4.5 m. They illuminate the whole scene at once with a flash laser, whose signal primarily drops with the inverse of the distance squared due to radial spreading. Apart from attenuation lengths, turbidity is also measured in other units, such as dB lost per meter [82] or NTU. Chourasiya et al. [148] correlate maximum scan distance and turbidity: they can scan up to 10 cm at 100 NTU with Kinect™ v1. The data represented in Figures 21 and 22 corresponds to the reported results at the lowest turbidity level of each paper (clear water).

5.7. Commercial Scanners

There are also commercial optical underwater 3D scanners. Their performances are compared quantitatively in Table 3. The only ToF scanners are the ones by 3DatDepth, the rest are triangulation-based. Regarding the light projection techniques, many of them (NewtonLabs, 3DatDepth, KrakenRobotics) steer the laser line. On the other hand, the scanners created by 2GRobotics and Savante perform fixed-line scanning.

Table 3. Quantitative comparison of commercial underwater active optical 3D scanners. ¹ Does not do beam steering. ² After CAD processing.

Company	Product	Range [m]	FoV [°]	Water Depth [m]	Depth Accuracy [mm]	Sample Rate [points/s]	Frequency [Hz]
2G Robotics	ULS-100 [152]	0.13–1	50 ¹	-	-	4750	-
	ULS-200 [153]	0.36–2.5	50 ¹	-	-	4750	-
	ULS-500 PRO [154]	1.5–20	50 ¹	4000	-	61,440	-
Newton Labs	M500UW [155]	0.15–0.9	≈25 × 25	500	0.15 at 0.9 m ²	-	-
	HRM1500UW [156]	0.15–0.9	≈25 × 25	1500	0.15 at 0.9 m ²	-	-
	HRM3200UW [157]	0.15–1	≈25 × 25	3200	0.15 at 0.9 m ²	-	-
	HRM4000UW [158]	0.15–1	≈25 × 25	4000	0.15 at 0.9 m ²	-	-
	M1500UW [159]	0.5–5	≈20 × 20	1500	1.41 at 5 m ²	-	-
	M3200UW [160]	0.5–5	≈20 × 20	3200	1.41 at 5 m ²	-	-
	M4000UW [161]	0.5–5	≈20 × 20	4000	1.41 at 5 m ²	-	-
	PL3200UW-LW [162]	1.5–10	≈35 × 35	3200	1.4 at 10 m ²	-	-
	PL4000UW [163]	1.5–10	≈35 × 35	4000	1.4 at 10 m ²	-	-
	M210UW [164]	0.15–0.9	≈25 × 25	100	0.7 at 0.9 m	-	-
	M310UW [165]	0.5–5	≈20 × 20	100	1.2 at 5 m ²	-	-
Savante	SLV-8000i [166]	0.75–15	60 ¹	4000	-	-	22
Kraken Robotics	Sea Vision [63]	0.5–8	65 × 50	-	3 at 2 m	300 k	0.1–1
3D at Depth	SL1 [167]	2–45	30 × 30	3000	6	-	-
	SL2 [168]	2–45	30 × 30	1500	6	-	-
	SL3 [169]	2–45	30 × 30	3000	6	-	-

Some of the performance results in Table 3 should be highlighted. Regarding range, they can all achieve longer distances than academic sensors. ToF sensors by 3DatDepth can measure at distances of up to 45 m, while 2GRobotics has developed the ULS-500 PRO [154], which allows for distances of up to 20 m. In order to achieve that, its baseline is very long (1.24 m). Most of the analyzed commercial scanners can work at water depths of a few thousand meters. Concerning depth accuracy, ToF scanners by 3DatDepth achieve errors of a few millimeters, which constitute relative errors of 0.013% at their maximum range. On the other hand, triangulation sensors, NewtonLabs' M210UW [164] and Kraken's SeaVision [63] achieve 0.08% and 0.15%, respectively.

6. Conclusions

The design of an active optical 3D sensor for underwater applications depends strongly on the characteristics of the task to be carried out. The decision of whether it should be a ToF or a triangulation scanner is mainly determined by the trade-off between depth accuracy and range: a ToF sensor is more suited for ranges up to a few tens of meters, whereas a triangulation sensor can be capable of submillimetric accuracy when working at shorter ranges.

Another design trade-off concerns lateral resolution, acquisition time and structural simplicity. One-shot systems acquire the whole scene at once—which makes them more suited for highly dynamic scenes—at the expense of reducing lateral resolution. Steered line scanners achieve very dense point clouds but if they are attached to an UUV, the movement of the vehicle can distort the resulting data. Non-steered line sensors are easier to build and calibrate, but they need to be attached to a moving platform, such as a vehicle or a rotating tripod. The accuracy of the measurements depends heavily on the accuracy of the data of the platform's position.

It is relevant to note that some of the reviewed papers lack a complete, systematic test on the performance of their presented sensors. Therefore, quantitative comparisons of the current technologies are cumbersome to carry out. It would be highly beneficial that researchers reported on basic performance criteria of underwater 3D scanners, such as depth resolution, measurement range, lateral resolution, number of scans per second and FoV.

As explained in Section 5, depth accuracy can be assessed by statistically analyzing the data resulting from measuring known objects at known distances. For triangulation sensors, the baseline and the distance to the object at which the depth measurements were made should always be reported, since their depth resolution depends on their baseline and range. A way to experimentally assess lateral resolution of ToF scanners is as performed in [138]. The authors scanned an object that was composed of bars whose widths decreased down to zero. Then, they estimated the lateral resolution of the sensor to be equal to the smallest bar width that the system is able to resolve. In order to assess 3D scanners in a complete, objective way, authors should ideally conform to metrological standards, such as VDI/VDE 2634 [170].

Furthermore, it would be interesting to systematically test the sensors with different targets and at varying turbidity conditions, such as [60,112,134], among others, because robustness against changing visibility situations is fundamental for autonomous robots to work in real environments.

Author Contributions: Conceptualization, M.C., A.P., J.F. and P.R.; methodology, M.C.; investigation, M.C.; writing, M.C.; reviewing, A.P., J.F. and P.R.; visualization, M.C.; supervision, A.P., J.F. and P.R.; project administration, P.R.; funding acquisition, P.R.

Funding: This research was funded by the GIRONA1000 project (grant number DPI2017-86372-C3-2-R, by the Spanish Ministry of Science) and EUMR project (grant number H2020-INFRAIA-2017-1-twostage-731103, by the European Commission).

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ADC	analog–digital converter
AO	acousto–optic
AOD	acousto–optic deflector
APD	avalanche photodiode
AUV	autonomous underwater vehicle
BRDF	bidirectional reflection distribution function
CAD	computer-aided design
CCD	charge-coupled device
CMM	coordinate measurement machine
CMOS	complementary metal-oxide-semiconductor
CW	continuous wave
DLP	digital light processor
DMD	digital micro-mirror device
DOE	diffractive optical element
DoF	degree of freedom
EO	electro–optic
EOD	electro–optic deflector
FoV	field of view
GCPS	gray code phase stepping
GLV	grating light valve
IMR	inspection, maintenance and repair
IMU	inertial measurement unit
INS	inertial navigation system
IR	infrared
KTN	potassium tantalate niobate
LCD	liquid crystal display
LED	light emitting diode
LiDAR	light detection and ranging
LLS	laser line scanner
MEMS	microelectromechanical systems
MFPS	multi-frequency phase stepping
MPPC	multipixel photon counter
NTU	nephelometric turbidity units
OPA	optical phased array
PMT	photomultiplier tube
ROV	remotely operated vehicle
SfM	structure from motion
SiPM	silicon photomultiplier
SLAM	simultaneous localization and mapping
SLM	spatial light modulator
SNR	signal-to-noise ratio
SONAR	sound navigation ranging
SPAD	single-photon avalanche photodiode
SVP	single view-point
TCSPC	time-correlated single-photon counting
ToF	time of flight
UUV	unmanned underwater vehicle

References

1. National Oceanic and Atmospheric Administration (NOAA), US Department of Commerce. Oceans & Coasts. Available online: <https://www.noaa.gov/oceans-coasts> (accessed on 12 April 2019).
2. Kyo, M.; Hiyazaki, E.; Tsukioka, S.; Ochi, H.; Amitani, Y.; Tsuchiya, T.; Aoki, T.; Takagawa, S. The sea trial of “KAIKO”, the full ocean depth research ROV. In Proceedings of the OCEANS ’95 MTS/IEEE ‘Challenges of Our Changing Global Environment’, San Diego, CA, USA, 9–12 December 1995; Volume 3, pp. 1991–1996. [CrossRef]
3. Foley, B.; Mindell, D. Precision Survey and Archaeological Methodology in Deep Water. *ENALIA J. Hell. Inst. Mar. Archaeol.* **2002**, *6*, 49–56.
4. García, R.; Gracias, N.; Nicosevici, T.; Prados, R.; Hurtós, N.; Campos, R.; Escartin, J.; Elibol, A.; Hegedus, R.; Neumann, L. Exploring the Seafloor with Underwater Robots. In *Computer Vision in Vehicle Technology*; John Wiley & Sons, Ltd.: Chichester, UK, 2017; pp. 75–99. [CrossRef]
5. Roman, C.; Inglis, G.; Rutter, J. Application of structured light imaging for high resolution mapping of underwater archaeological sites. In Proceedings of the OCEANS’10 IEEE Sydney, Sydney, NSW, Australia, 24–27 May 2010. [CrossRef]

6. Johnson-Roberson, M.; Bryson, M.; Friedman, A.; Pizarro, O.; Troni, G.; Ozog, P.; Henderson, J.C. High-Resolution Underwater Robotic Vision-Based Mapping and Three-Dimensional Reconstruction for Archaeology. *J. Field Robot.* **2017**, *34*, 625–643. [[CrossRef](#)]
7. Giguere, P.; Dudek, G.; Prahacs, C.; Plamondon, N.; Turgeon, K. Unsupervised learning of terrain appearance for automated coral reef exploration. In Proceedings of the 2009 Canadian Conference on Computer and Robot Vision (CRV), Kelowna, BC, Canada, 25–27 May 2009; pp. 268–275. [[CrossRef](#)]
8. Smith, R.N.; Schwager, M.; Smith, S.L.; Jones, B.H.; Rus, D.; Sukhatme, G.S. Persistent ocean monitoring with underwater gliders: Adapting sampling resolution. *J. Field Robot.* **2011**, *28*, 714–741. [[CrossRef](#)]
9. Pizarro, O.; Singh, H. Toward large-area mosaicing for underwater scientific applications. *IEEE J. Ocean. Eng.* **2003**, *28*, 651–672. [[CrossRef](#)]
10. Pascoal, A.; Oliveira, P.; Silvestre, C.; Sebastião, L.; Rufino, M.; Barroso, V.; Gomes, J.; Ayela, G.; Coince, P.; Cardew, M.; et al. Robotic ocean vehicles for marine science applications: The european ASIMOV project. *IEEE Ocean. Conf. Rec.* **2000**, *1*, 409–415. [[CrossRef](#)]
11. Yoerger, D.R.; Jakuba, M.; Bradley, A.M.; Bingham, B. Techniques for deep sea near bottom survey using an autonomous underwater vehicle. *Int. J. Robot. Res.* **2007**, *26*, 41–54. [[CrossRef](#)]
12. DeVault, J.E. Robotic system for underwater inspection of bridge piers. *IEEE Instrum. Meas. Mag.* **2000**, *3*, 32–37. [[CrossRef](#)]
13. Lirman, D.; Gracias, N.; Gintert, B.; Gleason, A.C.; Deangelo, G.; Dick, M.; Martinez, E.; Reid, R.P. Damage and recovery assessment of vessel grounding injuries on coral reef habitats by use of georeferenced landscape video mosaics. *Limnol. Oceanogr. Methods* **2010**, *8*, 88–97. [[CrossRef](#)]
14. Schjølberg, I.; Gjersvik, T.B.; Transeth, A.A.; Utne, I.B. Next Generation Subsea Inspection, Maintenance and Repair Operations. *IFAC-PapersOnLine* **2016**, *49*, 434–439. [[CrossRef](#)]
15. Liljebäck, P.; Mills, R. Eelume: A flexible and subsea resident IMR vehicle. In Proceedings of the IEEE OCEANS 2017—Aberdeen, Aberdeen, UK, 19–22 June 2017; pp. 1–4. [[CrossRef](#)]
16. Himri, K.; Ridao, P.; Gracias, N.; Palomer, A.; Palomeras, N.; Pi, R. Semantic SLAM for an AUV using object recognition from point clouds. *IFAC-PapersOnLine* **2018**, *51*, 360–365. [[CrossRef](#)]
17. Palomer, A.; Ridao, P.; Ribas, D. Inspection of an Underwater Structure using Point Cloud SLAM with an AUV and a Laser Scanner. *J. Field Robot.* **2019**, *36*, 1333–1344. [[CrossRef](#)]
18. Palomer, A.; Ridao, P.; Youakim, D.; Ribas, D.; Forest, J.; Petillot, Y.; Peñalver Monfort, A.; Sanz, P.J. 3D Laser Scanner for Underwater Manipulation. *Sensors* **2018**, *18*, 1086. [[CrossRef](#)] [[PubMed](#)]
19. Dalglish, F.R.; Tetlow, S.; Allwood, R.L. Experiments in laser-assisted visual sensing for AUV navigation. *Control Eng. Pract.* **2004**, *12*, 1561–1573. [[CrossRef](#)]
20. Massot-Campos, M.; Oliver-Codina, G. Optical sensors and methods for underwater 3D reconstruction. *Sensors* **2015**, *15*, 31525–31557. [[CrossRef](#)]
21. Bruno, F.; Bianco, G.; Muzzupappa, M.; Barone, S.; Rationale, A.V. Experimentation of structured light and stereo vision for underwater 3D reconstruction. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 508–518. [[CrossRef](#)]
22. Bianco, G.; Gallo, A.; Bruno, F.; Muzzupappa, M. A comparative analysis between active and passive techniques for underwater 3D reconstruction of close-range objects. *Sensors* **2013**, *13*, 11007–11031. [[CrossRef](#)]
23. Menna, F.; Agrafiotis, P.; Georgopoulos, A. State of the art and applications in archaeological underwater 3D recording and mapping. *J. Cult. Herit.* **2018**, *33*, 231–248. [[CrossRef](#)]
24. Seitz, S. *An Overview of Passive Vision Techniques*; Technical Report; The Robotics Institute, Carnegie Mellon University: Pittsburgh, PA, USA, 1999.
25. Sarafraz, A.; Negahdaripour, S.; Schechner, Y.Y. *Improving Stereo Correspondence in Scattering Media by Incorporating Backscatter Cue*; Technical Report; TECHNION—Israel Institute of Technology: Haifa, Israel, 2010.
26. Murez, Z.; Treibitz, T.; Ramamoorthi, R.; Kriegman, D.J. Photometric Stereo in a Scattering Medium. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *39*, 3415–3423. [[CrossRef](#)]
27. Fujimura, Y.; Iiyama, M.; Hashimoto, A.; Minoh, M. Photometric Stereo in Participating Media Considering Shape-Dependent Forward Scatter. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7445–7453.
28. Dalglish, F.R.; Caimi, F.M.; Britton, W.B.; Andren, C.F. Improved LLS imaging performance in scattering-dominant waters. *Ocean Sens. Monit.* **2009**, *7317*, 73170E. [[CrossRef](#)]

29. Johnson-Roberson, M.; Pizarro, O.; Williams, S.B.; Mahon, I. Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys. *J. Field Robot.* **2010**, *27*, 21–51. [CrossRef]
30. Menna, F.; Nocerino, E.; Troisi, S.; Remondino, F. A photogrammetric approach to survey floating and semi-submerged objects. *Videomet. Range Imaging Appl. XII Automat. Visual Inspection* **2013**, 8791, 87910H. [CrossRef]
31. Yamafune, K.; Torres, R.; Castro, F. Multi-Image Photogrammetry to Record and Reconstruct Underwater Shipwreck Sites. *J. Archaeol. Method Theory* **2017**, *24*, 703–725. [CrossRef]
32. Bosch, J.; Istenic, K.; Gracias, N.; Garcia, R.; Ridao, P. Omnidirectional Multicamera Video Stitching Using Depth Maps. *IEEE J. Ocean. Eng.* **2019**, 1–16. [CrossRef]
33. Jaffe, J.S.; Moore, K.D.; McLean, J.W.; Strand, M. Underwater Optical Imaging: Status and Prospects. *Oceanography* **2001**, *14*, 64–75. [CrossRef]
34. Kocak, D.M.; Caimi, F.M. The Current Art of Underwater Imaging- With a Glimpse of the Past and Vision of the Future. *Mar. Technol. Soc. J.* **2005**, *39*, 5–26. [CrossRef]
35. Caimi, F.M.; Kocak, D.M.; Dalglish, F.R.; Watson, J. Underwater imaging and optics: Recent advances. In Proceedings of the OCEANS 2008, Quebec City, QC, Canada, 15–18 September 2008; pp. 1–9. [CrossRef]
36. Hale, G.M.; Querry, M.R. Optical Constants of Water in the 200-nm to 200- μ m Wavelength Region. *Appl. Opt.* **1973**, *12*, 555. [CrossRef]
37. Smith, R.C.; Baker, K.S. Optical properties of the clearest natural waters (200–800 nm). *Appl. Opt.* **1981**, *20*, 177–184. [CrossRef]
38. Gracias, N.; Negahdaripour, S.; Neumann, L.; Prados, R.; Garcia, R. A motion compensated filtering approach to remove sunlight flicker in shallow water images. In Proceedings of the OCEANS 2008, Quebec City, QC, Canada, 15–18 September 2008; pp. 1–7. [CrossRef]
39. Cho, M.; Javidi, B. Three-Dimensional Visualization of Objects in Turbid Water Using Integral Imaging. *J. Disp. Technol.* **2010**, *6*, 544–547. [CrossRef]
40. Wikimedia Commons. File: Absorption Spectrum of Liquid Water.png—Wikimedia Commons, the Free Media Repository. 2016. Available online: https://commons.wikimedia.org/w/index.php?title=File:Absorption_spectrum_of_liquid_water.png&oldid=183806579 (accessed on 22 May 2019).
41. Hecht, E. *Optics*, 5th ed.; Pearson Education Limited: London, UK, 2017.
42. Wilde, F.; Radtke, D.; Gibbs, J.; Iwatsubo, R. *National Field Manual for the Collection of Water-Quality Data*, 9th ed.; US Geological Survey Techniques in Water-Resources Investigations: North Platte, NE, USA, 1998.
43. Tanabashi, M.; Hagiwara, K.; Hikasa, K. Review of Particle Physics. *Phys. Rev. D* **2018**, *98*, 030001. [CrossRef]
44. Kocak, D.M.; Dalglish, F.R.; Caimi, F.M.; Schechner, Y.Y. A Focus on Recent Developments and Trends in Underwater Imaging. *Mar. Technol. Soc. J.* **2008**, *42*, 52–67. [CrossRef]
45. Tan, C.S.; Seet, G.G.; Sluzek, A.; He, D.M. A novel application of range-gated underwater laser imaging system (ULIS) in near-target turbid medium. *Opt. Lasers Eng.* **2005**, *43*, 995–1009. [CrossRef]
46. Schechner, Y.Y.; Karpel, N. Clear underwater vision. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, 27 June–2 July 2004; Volume 1, pp. 536–543. [CrossRef]
47. Hildebrandt, M.; Kerdels, J.; Albiez, J.; Kirchner, F. A practical underwater 3D-Laserscanner. In Proceedings of the OCEANS, Quebec City, QC, Canada, 15–18 September 2008; pp. 1–5. [CrossRef]
48. Palomer, A.; Ridao, P.; Forest, J.; Ribas, D. Underwater Laser Scanner: Ray-based Model and Calibration. *IEEE/ASME Trans. Mechatronics* **2019**, *24*, 1986–1997. [CrossRef]
49. Brown, D.C. Close-range camera calibration. *Photogramm. Eng.* **1971**, *37*, 855–866. [CrossRef]
50. Shortis, M.R. Calibration techniques for accurate measurements by underwater camera systems. *Sensors* **2015**, *15*, 30810–30827. [CrossRef]
51. Sedlazeck, A.; Koch, R. *Perspective and Non-Perspective Camera Models in Underwater Imaging—Overview and Error Analysis*; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Cham, Switzerland, 2012, Volume 7474, pp. 212–242. [CrossRef]
52. Schiebener, P.; Straub, J.; Levelt Sengers, J.; Gallagher, J. Refractive index of water and steam as function of wavelength, temperature and density. *J. Phys. Chem. Ref. Data* **1990**, *19*, 677–717. [CrossRef]
53. Tengesdal, O.A. Measurement of Seawater Refractive Index and Salinity by Means of Optical Refraction. Ph.D. Thesis, University of Bergen, Bergen, Norway, 2012.

54. Łuczynski, T.; Pflingstorn, M.; Birk, A. The Pinax-model for accurate and efficient refraction correction of underwater cameras in flat-pane housings. *Ocean Eng.* **2017**, *133*, 9–22. [[CrossRef](#)]
55. Kwon, Y.H. Object plane deformation due to refraction in two-dimensional underwater motion analysis. *J. Appl. Biomech.* **1999**, *15*, 396–403. [[CrossRef](#)]
56. Wang, C.C.; Cheng, M.S. Nonmetric camera calibration for underwater laser scanning system. *IEEE J. Ocean. Eng.* **2007**, *32*, 383–399. [[CrossRef](#)]
57. Palomer, A.; Ridaio, P.; Ribas, D.; Forest, J. *Underwater 3D Laser Scanners: The Deformation of the Plane*; Lecture Notes in Control and Information Sciences; Springer: Cham, Switzerland, 2017; Volume 474, pp. 73–88. [[CrossRef](#)]
58. Kunz, C.; Singh, H. Hemispherical refraction and camera calibration in underwater vision. In Proceedings of the OCEANS 2008, Quebec City, QC, Canada, 15–18 September 2008; pp. 1–7. [[CrossRef](#)]
59. Menna, F.; Nocerino, E.; Remondino, F. Flat versus hemispherical dome ports in underwater photogrammetry. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2017**, *42*, 481–487. [[CrossRef](#)]
60. Massot-Campos, M.; Oliver-Codina, G. Underwater laser-based structured light system for one-shot 3D reconstruction. In Proceedings of the IEEE Sensors, Valencia, Spain, 2–5 November 2014; pp. 1138–1141. [[CrossRef](#)]
61. Bodenmann, A.; Thornton, B.; Ura, T. Generation of High-resolution Three-dimensional Reconstructions of the Seafloor in Color using a Single Camera and Structured Light. *J. Field Robot.* **2017**, *34*, 833–851. [[CrossRef](#)]
62. Yang, Y.; Zheng, B.; Kan, L.Y.; Yu, J.; Wang, J.C. 3D color reconstruction based on underwater RGB laser line scanning system. *Optik* **2014**, *125*, 6074–6077. [[CrossRef](#)]
63. Kraken Robotics. Sea Vision®. Available online: <https://krakenrobotics.com/products/seavision/> (accessed on 28 March 2019).
64. Bryson, M.; Johnson-Roberson, M.; Pizarro, O.; Williams, S.B. True Color Correction of Autonomous Underwater Vehicle Imagery. *J. Field Robot.* **2016**, *33*, 853–874. [[CrossRef](#)]
65. Peng, Y.T.; Cosman, P.C. Underwater Image Restoration Based on Image Blurriness and Light Absorption. *IEEE Trans. Image Process.* **2017**, *26*, 1579–1594. [[CrossRef](#)] [[PubMed](#)]
66. Lu, H.; Li, Y.; Zhang, Y.; Kim, H.; Serikawa, S.; Chen, M. Underwater Optical Image Processing: A Comprehensive Review. *Mob. Netw. Appl.* **2017**, *22*, 1204–1211. [[CrossRef](#)]
67. Ancuti, C.O.; Ancuti, C.; De Vleeschouwer, C.; Neumann, L.; Garcia, R. Color transfer for underwater dehazing and depth estimation. In Proceedings of the International Conference on Image Processing, Beijing, China, 17–20 September 2018; pp. 695–699. [[CrossRef](#)]
68. Liu, J.J.; Jakas, A.; Al-Obaidi, A.; Liu, Y. Practical issues and development of underwater 3D laser scanners. In Proceedings of the 15th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA 2010), Bilbao, Spain, 13–16 September 2010; pp. 1–8. [[CrossRef](#)]
69. Forest, J.; Salvi, J.; Cabruja, E.; Pous, C. Laser stripe peak detector for 3D scanners. A FIR filter approach. In Proceedings of the International Conference on Pattern Recognition, Cambridge, UK, 26 August 2004; Volume 3, pp. 646–649. [[CrossRef](#)]
70. Duda, A.; Albiez, J. Back Projection Algorithm for Line Structured Light Extraction. In Proceedings of the 2013 OCEANS, San Diego, CA, USA, 23–27 September 2013; pp. 1–7.
71. Zanuttigh, P.; Mutto, C.D.; Minto, L.; Marin, G.; Dominio, F.; Cortelazzo, G.M. In *Time-of-Flight and Structured Light Depth Cameras: Technology and Applications*; Springer International Publishing: Cham, Switzerland, 2016; pp. 1–355. [[CrossRef](#)]
72. Lange, R. 3D Time-of-Flight Distance Measurement with Custom Solid-State Image Sensors in CMOS/CCD-Technology. Ph.D. Thesis, University of Siegen, Siegen, Germany, 2000.
73. Hansard, M.; Lee, S.; Choi, O.; Horaud, R. *Time of Flight Cameras: Principles, Methods, and Applications*; Springer Briefs in Computer Science; Springer: London, UK, 2012; p. 95. [[CrossRef](#)]
74. Giancola, S.; Valenti, M.; Sala, R. *A Survey on 3D Cameras: Metrological Comparison of Time-of-Flight, Structured-Light and Active Stereoscopy Technologies*; Springer Briefs in Computer Science; Springer International Publishing: Cham, Switzerland, 2018; pp. 89–90. [[CrossRef](#)]
75. Chua, S.Y.; Guo, N.; Tan, C.S.; Wang, X. Improved range estimation model for three-dimensional (3D) range gated reconstruction. *Sensors* **2017**, *17*, 2031. [[CrossRef](#)] [[PubMed](#)]

76. Risholm, P.; Thorstensen, J.; Thielemann, J.T.; Kaspersen, K.; Tschudi, J.; Yates, C.; Softley, C.; Abrosimov, I.; Alexander, J.; Haugholt, K.H. Real-time super-resolved 3D in turbid water using a fast range-gated CMOS camera. *Appl. Opt.* **2018**, *57*, 3927–3937. [[CrossRef](#)] [[PubMed](#)]
77. Li, K.K.; Law, H. Picosecond InGaAs PIN Photodiode For 0.95 μm –1.65 μm Operation. In *Ultrashort Pulse Spectroscopy and Applications*; Soileau, M.J., Ed.; International Society for Optics and Photonics: Lansdale, PA, USA, 2012; Volume 0533, p. 126. [[CrossRef](#)]
78. Kharraz, O.; Forsyth, D. Performance comparisons between PIN and APD photodetectors for use in optical communication systems. *Optik* **2013**, *124*, 1493–1498. [[CrossRef](#)]
79. Riu, J. Cámara LiDAR de Escaneo MEMS para Imagen 3D de Resolución Espacial Variable. Ph.D. Thesis, Universitat Politècnica de Catalunya, Barcelona, Spain, 2018.
80. Rumbaugh, L.K.; Bollt, E.M.; Jemison, W.D.; Li, Y. A 532 nm Chaotic Lidar Transmitter for High Resolution Underwater Ranging and Imaging. In Proceedings of the 2013 OCEANS, San Diego, CA, USA, 23–27 September 2013; pp. 1–6.
81. Dalglish, F.R.; Caimi, F.M.; Britton, W.B.; Andren, C.F. An AUV-deployable pulsed laser line scan (PLLS) imaging sensor. In Proceedings of the Oceans Conference Record, Vancouver, BC, Canada, 29 September–4 October 2007; pp. 1–5. [[CrossRef](#)]
82. Imaki, M.; Ochimizu, H.; Tsuji, H.; Kameyama, S.; Saito, T.; Ishibashi, S.; Yoshida, H. Underwater three-dimensional imaging laser sensor with 120-deg wide-scanning angle using the combination of a dome lens and coaxial optics. *Opt. Eng.* **2016**, *56*, 031212. [[CrossRef](#)]
83. Finkelstein, H.; Hsu, M.J.; Esener, S. An ultrafast Geiger-mode single-photon avalanche diode in 0.18- μm CMOS technology. *Adv. Photon Count. Tech.* **2006**, 6372, 63720W. [[CrossRef](#)]
84. Yamamoto, K.; Yamamura, K.; Sato, K.; Kamakura, S.; Ota, T.; Suzuki, H.; Ohsuka, S. Development of Multi-Pixel Photon Counter (MPPC). In Proceedings of the IEEE Nuclear Science Symposium Conference Record, Honolulu, HI, USA, 26 October–3 November 2007; Volume 2, pp. 1511–1515. [[CrossRef](#)]
85. Nassalski, A.; Moszyński, M.; Syntfeld-Kazuch, A.; Szcześniak, T.; Świdorski, L.; Wolski, D.; Batsch, T.; Baszak, J. Multi pixel photon counters (MPPC) as an alternative to APD in PET applications. *IEEE Trans. Nucl. Sci.* **2010**, *57*, 1008–1014. [[CrossRef](#)]
86. Agishev, R.; Comerón, A.; Bach, J.; Rodriguez, A.; Sicard, M.; Riu, J.; Royo, S. Lidar with SiPM: Some capabilities and limitations in real environment. *Opt. Laser Technol.* **2013**, *49*, 86–90. [[CrossRef](#)]
87. Kim, H.; Handa, A.; Benosman, R.; Ieng, S.H.; Davison, A. Simultaneous Mosaicing and Tracking with an Event Camera. In *Proceedings of the British Machine Vision Conference*; Valstar, M., French, A., Pridmore, T., Eds.; BMVA Press: Durham, UK, 2014. [[CrossRef](#)]
88. Bardow, P.; Davison, A.J.; Leutenegger, S. Simultaneous Optical Flow and Intensity Estimation from an Event Camera. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 884–892. [[CrossRef](#)]
89. Kim, H.; Leutenegger, S.; Davison, A.J. *Real-Time 3D Reconstruction and 6-DoF Tracking With an Event Camera*; Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Cham, Switzerland, 2016; Volume 9910, pp. 349–364. [[CrossRef](#)]
90. Gallego, G.; Lund, J.E.; Mueggler, E.; Rebecq, H.; Delbruck, T.; Scaramuzza, D. Event-Based, 6-DOF Camera Tracking from Photometric Depth Maps. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 2402–2412. [[CrossRef](#)]
91. Falanga, D.; Kim, S.; Scaramuzza, D. How Fast Is Too Fast? The Role of Perception Latency in High-Speed Sense and Avoid. *IEEE Robot. Autom. Lett.* **2019**, *4*, 1884–1891. [[CrossRef](#)]
92. Moore, K.D.; Jaffe, J.S.; Ochoa, B.L. Development of a new underwater bathymetric laser imaging system: L-Bath. *J. Atmos. Ocean. Technol.* **2000**, *17*, 1106–1117. [[CrossRef](#)]
93. Munaro, M.; So, E.W.Y.; Tonello, S.; Menegatti, E. Efficient completeness inspection using real-time 3D color reconstruction with a dual-laser triangulation system. In *Integrated Imaging and Vision Techniques for Industrial Inspection: Advances and Applications*; Springer: London, UK, 2015; pp. 201–225. [[CrossRef](#)]
94. Ji, Z.; Leu, M.C. Design of optical triangulation devices. *Opt. Laser Technol.* **1989**, *21*, 339–341. [[CrossRef](#)]
95. Narasimhan, S.G.; Nayar, S.K. Structured light methods for underwater imaging: Light stripe scanning and photometric stereo. In Proceedings of the MTS/IEEE OCEANS, Washington, DC, USA, 17–23 September 2005; Volume 2005, pp. 1–8. [[CrossRef](#)]

96. McLeod, D.; Jacobson, J.; Hardy, M.; Embry, C. Autonomous inspection using an underwater 3D LiDAR. In Proceedings of the 2013 OCEANS, San Diego, CA, USA, 23–27 September 2013; pp. 1–8. [CrossRef]
97. Digumarti, S.T.; Chaurasia, G.; Taneja, A.; Siegwart, R.; Thomas, A.; Beardsley, P. Underwater 3D capture using a low-cost commercial depth camera. In Proceedings of the 2016 IEEE Winter Conference on Applications of Computer Vision, WACV 2016, Lake Placid, NY, USA, 7–10 March 2016; pp. 1–9. [CrossRef]
98. HOLO/OR LTD. FOV Magnification Module. Available online: <https://www.holoor.co.il/products/structured-light-module/> (accessed on 25 September 2019).
99. Dr. Walter Luhs. © LEYBOLD/LD DIDACTIC GmbH, Hürth. Photonics: Rotating Polygon Mirror Scanner. Available online: <http://www.photonics.ld-didactic.com/Educational%20Kits/P5889.html> (accessed on 30 April 2019).
100. Schwarze, C. A new look at Risley prisms. *Photonics Spectra* **2006**, *40*, 67–71.
101. Montagu, J.I. Achieving optimal high resolution in galvanometric scanning systems. *Infrared Technol. Appl. Int. Soc. Opt. Photonics* **1986**, *590*, 47–53. [CrossRef]
102. Chi, S.; Xie, Z.; Chen, W. A Laser Line auto-scanning system for underwater 3D reconstruction. *Sensors* **2016**, *16*, 1534. [CrossRef]
103. Chantler, M.J. Calibration and operation of an underwater laser triangulation sensor: The varying baseline problem. *Opt. Eng.* **1997**, *36*, 2604. [CrossRef]
104. Holmström, S.T.; Baran, U.; Urey, H. MEMS laser scanners: A review. *J. Microelectromech. Syst.* **2014**, *23*, 259–275. [CrossRef]
105. Brown, M.; Urey, H. MEMS Microdisplays. In *Handbook of Visual Display Technology*; Springer: Berlin/Heidelberg, Germany, 2015; pp. 1–15. [CrossRef]
106. Song, Y.; Panas, R.M.; Hopkins, J.B. A review of micromirror arrays. *Precis. Eng.* **2018**, *51*, 729–761. [CrossRef]
107. Preciseley. MEMS Mirror Array. Available online: <https://www.preciseley.com/mems-mirror-array.html> (accessed on 26 April 2019).
108. Corrigan, R.; Cook, R.; Favotte, O. *Silicon Light Machines™-Grating Light Valve™ Technology Brief Breakthrough MEMS Component Technology for Optical Networks*; Technical Report; Silicon Light Machines: Sunnyvale, CA, USA, 2001.
109. Perry, T. Tomorrow's TV. *IEEE Spectr.* **2004**, *41*, 38–41. [CrossRef]
110. Hornbeck, L.J. The DMD™ Projection Display Chip: A MEMS-Based Technology. *MRS Bull.* **2001**, *26*, 325–327. [CrossRef]
111. Narasimhan, S.G.; Nayar, S.K.; Sun, B.; Koppal, S.J. Structured light in scattering media. In Proceedings of the IEEE International Conference on Computer Vision, Beijing, China, 17–21 October 2005; Volume I, pp. 420–427. [CrossRef]
112. Risholm, P.; Kirkhus, T.; Thielemann, J.T. High-resolution structured light 3D sensor for autonomous underwater inspection. In Proceedings of the OCEANS 2018 MTS/IEEE, Charleston, SA, USA, 22–25 October 2018; pp. 1–5. [CrossRef]
113. Detry, R.; Koch, J.; Pailevanian, T.; Garrett, M.; Levine, D.; Yahnker, C.; Gildner, M. Turbid-water subsea infrastructure 3D reconstruction with assisted stereo. In Proceedings of the 2018 OCEANS-MTS/IEEE Kobe Techno-Oceans 2018, Kobe, Japan, 28–31 May 2018. [CrossRef]
114. Maradin. MEMS 2D Laser Scanning Mirror. Available online: <http://www.maradin.co.il/products/mar1100-mems-2d-laser-scanning-mirror/> (accessed on 26 April 2019).
115. Römer, G.R.; Bechtold, P. Electro-optic and acousto-optic laser beam scanners. *Phys. Procedia* **2014**, *56*, 29–39. [CrossRef]
116. Maldonado, T.A. Electro-Optic modulators. *Handb. Opt.* **1995**, *2*. [CrossRef]
117. National Research Council. *Laser Radar: Progress and Opportunities in Active Electro-Optical Sensing*; National Academies Press: Washington, DC, USA, 2014. [CrossRef]
118. Hou, L.; Smith, N.R.; Heikenfeld, J. Electrowetting micro-prisms and micro-mirrors. In Proceedings of the Lasers and Electro-Optics Society Annual Meeting-LEOS, Lake Buena Vista, FL, USA, 21–25 October 2007; pp. 457–458. [CrossRef]
119. Han, W.; Haus, J.W.; McManamon, P.; Heikenfeld, J.; Smith, N.; Yang, J. Transmissive beam steering through electrowetting micropillar arrays. *Opt. Commun.* **2010**, *283*, 1174–1181. [CrossRef]
120. Zohrabi, M.; Cormack, R.H.; Supekar, O.D.; Lim, W.Y.; Gopinath, J.T.; Bright, V.M. Lidar system with nonmechanical electrowetting-based wide-angle beam steering. *Opt. Express* **2019**, *27*, 4404. [CrossRef]

121. Chao, J.H.; Zhu, W.; Chen, C.J.; Hoffman, R.C.; Campbell, A.L.; Henry, M.G.; Yin, S. High speed non-mechanical two-dimensional KTN beam deflector enabled by space charge and temperature gradient deflection. *Opt. Express* **2017**, *25*, 15481. [[CrossRef](#)]
122. ISOMET. *Acousto-Optic Scanning and Deflection*; Technical Report; ISOMET: City of Manassas, VA, USA, 2018.
123. Salomé, R.; Kremer, Y.; Dieudonné, S.; Léger, J.F.; Krichevsky, O.; Wyart, C.; Chatenay, D.; Bourdieu, L. Ultrafast random-access scanning in two-photon microscopy using acousto-optic deflectors. *J. Neurosci. Methods* **2006**, *154*, 161–174. [[CrossRef](#)]
124. Ngoi, B.A.; Venkatakrishnan, K.; Lim, L.; Tan, B. Angular dispersion compensation for acousto-optic devices used for ultrashort-pulsed laser micromachining. *Opt. Express* **2001**, *9*, 200. [[CrossRef](#)]
125. Heck, M.J. Highly integrated optical phased arrays: Photonic integrated circuits for optical beam shaping and beam steering. *Nanophotonics* **2017**, *6*, 93–107. [[CrossRef](#)]
126. Yaacobi, A.; Sun, J.; Moresco, M.; Leake, G.; Coolbaugh, D.; Watts, M.R. Integrated phased array for wide-angle beam steering. *Opt. Lett.* **2014**, *39*, 4575. [[CrossRef](#)] [[PubMed](#)]
127. Yoo, H.W.; Druml, N.; Brunner, D.; Schwarzl, C.; Thurner, T.; Hennecke, M.; Schitter, G. MEMS-based lidar for autonomous driving. *Elektrotechnik Und Informationstechnik* **2018**, *135*, 408–415. [[CrossRef](#)]
128. Poulton, C.V.; Russo, P.; Timurdogan, E.; Whitson, M.; Byrd, M.J.; Hosseini, E.; Moss, B.; Su, Z.; Vermeulen, D.; Watts, M.R. High-Performance Integrated Optical Phased Arrays for Chip-Scale Beam Steering and LiDAR. In Proceedings of the Conference on Lasers and Electro-Optics, San Jose, CA, USA, 13–18 May 2018; p. ATu3R.2. [[CrossRef](#)]
129. Wikimedia Commons. File: Phased Array Animation with Arrow 10frames 371x400px 100ms.gif—Wikimedia Commons, the Free Media Repository. 2019. Available online: https://en.wikipedia.org/wiki/File:Phased_array_animation_with_arrow_10frames_371x400px_100ms.gif (accessed on 22 May 2019).
130. Bechtold, P.; Hohenstein, R.; Schmidt, M. Evaluation of disparate laser beam deflection technologies by means of number and rate of resolvable spots. *Opt. Lett.* **2013**, *38*, 2934. [[CrossRef](#)]
131. Ekkel, T.; Schmik, J.; Luhmann, T.; Hastedt, H. Precise laser-based optical 3D measurement of welding seams under water. *Int. Arch. Photogram. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2015**, *40*, 117–122. [[CrossRef](#)]
132. Buschinelli, P.D.; Matos, G.; Pinto, T.; Albertazzi, A. Underwater 3D shape measurement using inverse triangulation through two flat refractive surfaces. In Proceedings of the OCEANS 2016 MTS/IEEE Monterey, Monterey, CA, USA, 19–23 September 2016; pp. 1–7. [[CrossRef](#)]
133. Duda, A.; Schwendner, J.; Gaudig, C. SRSL: Monocular self-referenced line structured light. In Proceedings of the IEEE International Conference on Intelligent Robots and Systems, Hamburg, Germany, 28 September–2 October 2015; pp. 717–722. [[CrossRef](#)]
134. Mariani, P.; Quincoces, I.; Haugholt, K.H.; Chardard, Y.; Visser, A.W.; Yates, C.; Piccinno, G.; Reali, G.; Risholm, P.; Thielemann, J.T. Range-Gated Imaging System for Underwater Monitoring in Ocean Environment. *Sustainability* **2018**, *11*, 162. [[CrossRef](#)]
135. Sarafraz, A.; Haus, B.K. A structured light method for underwater surface reconstruction. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 40–52. [[CrossRef](#)]
136. Risholm, P.; Kirkhus, T.; Thielemann, J.T.; Thorstensen, J. Adaptive Structured Light with Scatter Correction for High-Precision Underwater 3D Measurements. *Sensors* **2019**, *19*, 1043. [[CrossRef](#)]
137. Bleier, M.; Nüchter, A. Low-Cost 3D laser scanning in air or water using self-calibrating structured light. *Int. Arch. Photogram. Remote Sens. Spat. Inf. Sci. ISPRS Arch.* **2017**, *42*, 105–112. [[CrossRef](#)]
138. Maccarone, A.; Mattioli, F.; Rocca, D.; Mccarthy, A.; Henderson, R.; Buller, G.S. Three-dimensional imaging of stationary and moving targets in turbid underwater environments using a single-photon detector array. *Opt. Express* **2019**, *27*, 28437. [[CrossRef](#)]
139. Massot-Campos, M.; Oliver-Codina, G.; Kemal, H.; Petillot, Y.; Bonin-Font, F. Structured light and stereo vision for underwater 3D reconstruction. In Proceedings of the MTS/IEEE OCEANS 2015—Genova: Discovering Sustainable Ocean Energy for a New World, Genoa, Italy, 18–21 May 2015; pp. 1–6. [[CrossRef](#)]
140. Lopes, F.; Silva, H.; Almeida, J.M.; Martins, A.; Silva, E. Structured light system for underwater inspection operations. In Proceedings of the MTS/IEEE OCEANS 2015—Genova: Discovering Sustainable Ocean Energy for a New World, Genoa, Italy, 18–21 May 2015; pp. 1–6. [[CrossRef](#)]

141. Martins, A.; Almeida, J.; Almeida, C.; Dias, A.; Dias, N.; Aaltonen, J.; Heininen, A.; Koskinen, K.T.; Rossi, C.; Dominguez, S.; et al. UX 1 system design—A robotic system for underwater mining exploration. In Proceedings of the 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Madrid, Spain, 1–5 October 2019; pp. 1494–1500. [CrossRef]
142. Constantinou, C.C.; Loizou, S.G.; Georgiades, G.P. An underwater laser vision system for relative 3-D posture estimation to mesh-like targets. In Proceedings of the IEEE International Conference on Intelligent Robots and Systems, Daejeon, Korea, 9–14 October 2016; pp. 2036–2041. [CrossRef]
143. Matos, G.; Buschinelli, P.D.; Pinto, T. Underwater Laser Triangulation Sensor Model with Flat Refractive Interfaces. *IEEE J. Ocean. Eng.* **2019**, 1–9. [CrossRef]
144. Massot-Campos, M.; Oliver, G.; Bodenmann, A.; Thornton, B. Submap bathymetric SLAM using structured light in underwater environments. In Proceedings of the 2016 IEEE/OES Autonomous Underwater Vehicles 2016 (AUV 2016), Tokyo, Japan, 6–9 November 2016; pp. 181–188. [CrossRef]
145. Ishibashi, S.; Ohta, Y.; Sugawara, M.; Tanaka, K.; Yoshida, H.; Choi, S. Seabed 3D images created by an underwater laser scanner applied to an AUV. In Proceedings of the OCEANS 2017, Anchorage, AK, USA, 18–21 September 2017; pp. 1–5.
146. Maccarone, A.; McCarthy, A.; Ren, X.; Warburton, R.E.; Wallace, A.M.; Moffat, J.; Petillot, Y.; Buller, G.S. Underwater depth imaging using time-correlated single-photon counting. *Opt. Express* **2015**, *23*, 33911. [CrossRef] [PubMed]
147. Anwer, A.; Azhar Ali, S.S.; Khan, A.; Meriaudeau, F. Underwater 3-D Scene Reconstruction Using Kinect v2 Based on Physical Models for Refraction and Time of Flight Correction. *IEEE Access* **2017**, *5*, 15960–15970. [CrossRef]
148. Chourasiya, S.; Mohapatra, P.K.; Tripathi, S. Non-intrusive underwater measurement of mobile bottom surface. *Adv. Water Resour.* **2017**, *104*, 76–88. [CrossRef]
149. Izquierdo, M.A.; Sanchez, M.T.; Ibañez, A.; Ullate, L.G. Sub-pixel measurement of 3D surfaces by laser scanning. *Sens. Actuators A Phys.* **1999**, *76*, 1–8. [CrossRef]
150. Forest Collado, J. New Methods for Triangulation-Based Shape Acquisition Using Laser Scanners. Ph.D. Thesis, Universitat de Girona, Girona, Spain, 2005.
151. de Dominicis, L. Underwater 3D vision, ranging and range gating. In *Subsea Optics and Imaging*; Elsevier: Amsterdam, The Netherlands, 2013; pp. 379–408. [CrossRef]
152. 2G Robotics. ULS-100. Available online: <https://www.2grobotics.com/products/underwater-laser-scanner-uls-100/> (accessed on 4 April 2019).
153. 2G Robotics. ULS-200. Available online: <https://www.2grobotics.com/products/underwater-laser-scanner-uls-200/> (accessed on 4 April 2019).
154. 2G Robotics. ULS-500 PRO. Available online: <https://www.2grobotics.com/products/underwater-laser-scanner-uls-500/> (accessed on 4 April 2019).
155. Newton Labs. M500UW. Available online: http://www.newtonlabs.com/M500UW_landing.htm (accessed on 4 April 2019).
156. Newton Labs. HRM1500UW. Available online: http://www.newtonlabs.com/HRM1500UW_landing.htm (accessed on 4 April 2019).
157. Newton Labs. HRM3200UW. Available online: http://www.newtonlabs.com/HRM3200UW_landing.htm (accessed on 4 April 2019).
158. Newton Labs. HRM4000UW. Available online: http://www.newtonlabs.com/HRM4000UW_landing.htm (accessed on 4 April 2019).
159. Newton Labs. M1500UW. Available online: http://www.newtonlabs.com/M1500UW_landing.htm (accessed on 4 April 2019).
160. Newton Labs. M3200UW. Available online: http://www.newtonlabs.com/M3200UW_landing.htm (accessed on 4 April 2019).
161. Newton Labs. M4000UW. Available online: http://www.newtonlabs.com/M4000UW_landing.htm (accessed on 4 April 2019).
162. Newton Labs. PL3200UW-LW. Available online: http://www.newtonlabs.com/PL3200UW-LW_landing.htm (accessed on 4 April 2019).
163. Newton Labs. PL4000UW. Available online: http://www.newtonlabs.com/PL4000UW_landing.htm (accessed on 4 April 2019).

164. Newton Labs. M210UW. Available online: http://www.newtonlabs.com/scan_m200uw_sys_specs.html (accessed on 4 April 2019).
165. Newton Labs. M310UW. Available online: http://www.newtonlabs.com/scan_m300uw_sys_specs.html (accessed on 4 April 2019).
166. Savante. SLV-8000i. Available online: <https://www.savante.co.uk/slv80-long-range-subsea-laser-profiler> (accessed on 4 April 2019).
167. 3D at Depth. SL1 LiDAR Laser. Available online: <https://www.3datdepth.com/product/sl1-lidar-laser> (accessed on 14 May 2019).
168. 3D at Depth. SL2 LiDAR Laser. Available online: <https://www.3datdepth.com/product/sl2-lidar-laser> (accessed on 14 May 2019).
169. 3D at Depth. SL3 LiDAR Laser. Available online: <https://www.3datdepth.com/product/sl3-lidar-laser> (accessed on 14 May 2019).
170. Verein Deutscher Ingenieure. VDI/VDE 2634: Optical 3-D Measuring Systems. Available online: <https://standards.globalspec.com/std/9914533/vdi-vde-2634-blatt-2> (accessed on 25 October 2019).



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).

3

UNDERWATER 3D SCANNER MODEL USING A BIAXIAL MEMS MIRROR

IN this chapter, we present the ray-tracing model of an underwater 3D scanner that uses a 2-axis mirror to counteract refraction. We also study the influence that the miscalibration of each parameter value would have on the accuracy of the light projection. Finally, we analyze how we should design the scanner using optimal parameter values to maximize the FoV and minimize the scanning pattern distortion.

Title: Underwater 3D scanner model using a biaxial MEMS mirror
Authors: **Miguel Castellón**, Albert Palomer, Josep Forest, and Pere Ridao
Journal: IEEE Access
Volume: 9, Pages: 50231–50243, Published: 2021
DOI: [10.1109/ACCESS.2021.3069189](https://doi.org/10.1109/ACCESS.2021.3069189)
Quality index: JCR2021 Engineering, Electrical & Electronic, Impact Factor: 3.476, Q2 (105/278)

Received March 12, 2021, accepted March 20, 2021, date of publication March 29, 2021, date of current version April 7, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3069189

Underwater 3D Scanner Model Using a Biaxial MEMS Mirror

MIGUEL CASTILLÓN¹, ALBERT PALOMER², JOSEP FOREST¹,
AND PERE RIDAO¹, (Member, IEEE)

¹Computer Vision and Robotics Research Institute (VICOROB), University of Girona, 17003 Girona, Spain

²IQUA Robotics SL, 17003 Girona, Spain

Corresponding author: Miguel Castellón (miguel.castillon@udg.edu)

This work was supported in part by the Spanish Ministry of Science through the GIRONA1000 Project under Grant DPI2017-86372-C3-2-R, in part by the European Commission through the ATLANTIS Project under Grant H2020-ICT-2019-2-871571, and in part by the Doctoral Grant of the University of Girona under Grant IFUdG2019.


ABSTRACT Underwater 3D laser scanners are an essential type of sensor used by unmanned underwater vehicles (UUVs) for operations such as navigation, inspection and object recognition and manipulation. These sensors need to be able to provide highly accurate 3D data at fast refresh rates in order to accomplish these tasks. Usually, these scanners rely on a rotating mirror actuated by a galvanometer. However, the light planes steered by this type of mirrors are typically deformed into cones due to refraction. In order to produce accurate results, this distortion needs to be taken into account, which increases the computational cost of the 3D reconstruction. A novel approach consisting in using a biaxial MEMS mirror is proposed in this paper. The second rotational degree of freedom of the mirror can be used to project optimally curved light shapes, so that the refraction process transforms them into planes. Being able to model the light surfaces as planes rather than cones can significantly reduce the computation time of the 3D reconstruction. In order to do so, an exhaustive model of the complete light trajectories is presented. To the best of the authors' knowledge, this paper constitutes the first attempt to model and counteract the distortion in the scanning pattern introduced by a biaxial mirror and a double refraction process in the context of underwater robotics.

INDEX TERMS Mobile robotics, autonomous robots, underwater robotics, 3D laser scanner, 3D sensing.

I. INTRODUCTION

Unmanned underwater vehicles (UUVs) are being increasingly used in industry out of safety and cost reasons. Particularly, autonomous underwater vehicles (AUVs) are already performing tasks such as object recognition [1], inspection [2], manipulation [3] or navigation [4]. Sensing their surroundings is essential for them to achieve their tasks. Therefore, they are usually equipped with some type of 3D sensor, which are mainly based either on acoustic (SONAR) or light signals (LiDAR). The main advantage of optical sensors is that they can provide a much higher lateral resolution and refresh rate [5]. Their relatively short range is usually enough for intervention tasks, since the robot needs to get close to the target.

Underwater 3D sensing using optical sensors has two main challenges: the fast attenuation rate of light in water and the distortion introduced by refractive elements.

The associate editor coordinating the review of this manuscript and approving it for publication was Jing Yan .

Therefore, 3D sensing technology developed for in-air applications cannot be directly submerged and used for underwater tasks. Some authors have tried to adapt commercial in-air scanners to the underwater domain [6]–[8]. Even though their low price could make them useful for certain applications, their low range and accuracy would not suffice for real operations such as mapping or object manipulation. Consequently, custom-made sensors need to be developed for underwater tasks. Moreover, they should achieve high performance regarding two main aspects: high accuracy and fast refresh rate.

Underwater 3D scanners designed for inspection and manipulation typically use a rotating mirror actuated by a galvanometer [9], [10]. However, the flat refractive surfaces that seal the scanner deform the outgoing light planes into cones [11]. Modelling these cones in order to produce accurate results is possible but entails a higher computational cost. This problem can be tackled by approaching the refraction process the other way around: if the projected planes result in cones when entering the water, is there any surface that

gets transformed into a plane due to the refraction process? In order to find those surfaces, a novel approach consisting in using a biaxial microelectromechanical system (MEMS) mirror is proposed in this paper. MEMS mirrors are small-sized mirrors whose actuator transforms the applied voltage into rotation angle. Due to their small inertia, they can rotate faster than typical mirrors. They are further reviewed in section II-B. The second rotational degree of freedom of the mirror can be used to project optimally curved light shapes, so that they are refracted into planes. In order to do so, an exhaustive model of the complete light trajectories and its results using synthetic data are presented in the following sections.

Modelling the light surface as a plane rather than as an elliptic cone constitutes an advantage from the point of view of the computational cost. Triangulation-based 3D laser scanners reconstruct the shape of the 3D target by triangulating each illuminated pixel on the camera image with the corresponding laser surface. This means that the position of each 3D point in the point cloud is computed by intersecting a line and the modelled surface. Taking into account that a typical value of point cloud density in this type of scanners is 500k points per scan, reducing the computational complexity of the reconstruction of each point is of great importance. When using an elliptic cone, this triangulation consists in solving a quadratic equation and choosing the appropriate value from the two-valued solution [11]. On the other hand, triangulation using a plane boils down to simply solving a linear equation. This reduces the computational time of the 3D reconstruction of every point by a factor of 8 in our current implementation.

The main contribution of this paper is presenting a model to compensate for the distortion introduced by the flat refractive surface, so that the light surface entering in the water can be modelled as a plane rather than as an elliptic cone. This way, the computational time used for the 3D reconstruction of the scene can be significantly reduced without losing accuracy. This claim is further developed by means of exhaustive simulations, which also aim at modelling the non-ideal behaviour of the scanner. To the best of the authors' knowledge, this paper constitutes the first attempt to model and counteract the distortion in the scanning pattern introduced by a biaxial mirror and a double refraction process in the context of underwater robotics.

The paper is structured as follows. First, a literature review on different topics related to this research is summarized in section II. Then, the proposed model is presented in section III. The non-ideal behaviour of the mirror is simulated in section IV in order to quantify the magnitude of the errors related to it. Finally, the conclusions drawn are listed in section V. As support notes, basic geometrical concepts are refreshed in appendix A.

II. RELATED WORK

This section studies the state of the art regarding 3D laser scanners. First, literature regarding underwater scanners is summarized in section II-A, including descriptions and

examples of different types of sensors. Then, in-air solutions that may be proved useful if adapted to the underwater domain are studied in section II-B.

A. UNDERWATER 3D LASER SCANNERS

Optical 3D sensors are usually divided into active and passive. An underwater sensor is active if it projects light patterns onto the scene in a controlled way [12]. These patterns can be spatial (a point, a line or more complex shapes) and/or temporal. In active techniques, the information given by the structure of the pattern is key to reconstruct the scene in 3D. Laser scanners, which is the topic of this section, are therefore active sensors. A complete survey on underwater active optical sensors can be found in a previous work [13]. On the other hand, a passive sensor reconstructs its surrounding scene in 3D using information other than the structure of the light. This information comes from different simultaneous viewpoints or from the movement between consecutive image frames. Well-known passive techniques in underwater environments are stereo vision [12], [14] and structure from motion (SfM) [15], [16].

Active optical 3D sensors can be further classified depending on how they project light onto the scene. One of the main approaches can be categorized under the "one-shot" label [17]–[20]. These sensors illuminate the whole scene at once with a certain pattern. The acquisition time for the whole scene is therefore extremely short. For this reason, they are suitable to scan scenes in which high dynamics are present. However, they can only provide limited lateral resolution (perpendicularly to the depth direction).

Another popular type of sensors are laser line scanners (LLSs). These scanners generally use laser line modules, which project a light plane onto the target. One of their main advantages when compared to one-shot scanners is their high point cloud density. However, since they need a certain amount of time to sweep the laser along the scene of interest, their 3D reconstructions may be distorted when high dynamics are present. For this reason, fast scanning frequencies are usually preferred.

Some LLSs [21]–[24] do not steer the laser plane, so that its relative pose with respect to the camera is constant. This way, they achieve a robust performance, since the number of parameters to be calibrated is relatively low. Moreover, by making the laser plane enter the viewport perpendicularly, the effects of the double refraction are minimized. However, the speed at which they can scan a certain area is totally dependent on the speed of the platform used to steer the scanner. They are typically mounted on static rotational heads [25] or attached to the bottom of a vessel or an UUV facing downwards, projecting the laser plane in the direction perpendicular to the movement of the platform.

In order to increase the size of the area that can be scanned per second, some authors include some kind of laser steering mechanism in their LLSs. One of the most well-known approaches is reflecting the laser plane on a rotating mirror actuated by a galvanometer [9], [10]. The mirror is normally

placed inside the sealed case, so the friction caused by the surrounding air is negligible compared to water. This small friction along with the small inertia of the mirror allows it to achieve relatively high rotation speeds, which increases notably the area that can be scanned per second.

However, this approach presents some issues. First, deflecting the laser ray direction makes it enter the viewport at a non-perpendicular angle (assuming a flat viewport is used). Therefore, the double refraction process causes the laser ray travelling in water to have another different direction. Moreover, the refractive viewport deforms the original laser plane into a cone, as described in [11]. These two non-ideal behaviours need to be taken into account in order to produce accurate 3D reconstructions. Doing so, nonetheless, is computationally expensive. Using a dome viewport can in principle be a solution to these problems, since the laser ray would theoretically always intersect the viewport plane perpendicularly. In reality, however, aligning the center of curvature of the viewport with the center of rotation of the mirror is not straightforward, so distortions may not disappear completely. Second, introducing a new element makes the sensor model grow with more parameters, so its calibration becomes more difficult.

Another approach to designing an underwater scanner different from a LLS was followed in [26]. Instead of using a laser line module with a 1-DoF mirror, they used a laser pointer reflected on a 2-DoF polygon mirror.

B. BIAxIAL-MIRROR SCANNERS FOR IN-AIR APPLICATIONS

One of the main drives for the research of 3D scanners outside water is the automotive industry. The development of sensors aimed at improving autonomous capabilities of cars is a rapidly growing field. More particularly, LiDAR-based navigation is one of the current main applications of these sensors [27]–[32]. Most of the commercial LiDARs use some kind of rotating mirror. Currently, an increasingly popular element to produce a small-sized LiDAR are MEMS mirrors [33].

MEMS mirrors are small-sized mirrors whose actuator transforms the applied voltage into rotation angle. Due to their small inertia, they can rotate faster than typical mirrors. Moreover, they are usually biaxial: they can rotate around two perpendicular axes. They can work in two different regimes: linear and resonant. The resonant regime occurs around the first resonance frequency of one of the axes of the mirror (typically a few kHz) [34], [35]. In this operation mode, one of the axes vibrates fast at its resonance frequency sweeping the target while the other axis controls the inclination of the projected line. This mode is known as raster scanning, which means that the scene is scanned from beginning to end. On the other hand, linear operation on both axes can only provide a slower scanning speed. However, it has the advantage of being able to continuously project the laser ray at arbitrary directions. Consequently, this regime can project arbitrary patterns at different resolutions. MEMS mirrors have been

recently studied in [36]–[38]. It is very relevant to know that the rotations of biaxial mirrors introduce distortions in the light patterns, as studied in [39].

Recently, there have also been advances on so-called solid-state LiDARs, which deflect the laser ray without using any moving part. They typically use electro-optic deflectors (EODs) or acousto-optic deflectors (AODs), which can accurately deflect the ray at very high speeds. Despite their potential advantages, EODs and AODs currently present very strong limitations, such as extremely narrow field of view (FoV) and limited wavelengths [40], [41].

III. MODEL

This section develops the proposed geometrical model of the underwater laser scanner. It is structured as follows. First, the description of the model elements and parameters is done in section III-A. Then, the implementation of the model is described in detail in section III-B. Finally, the inverse use of the model to compute the required pairs of mirror angles is explained in section III-C.

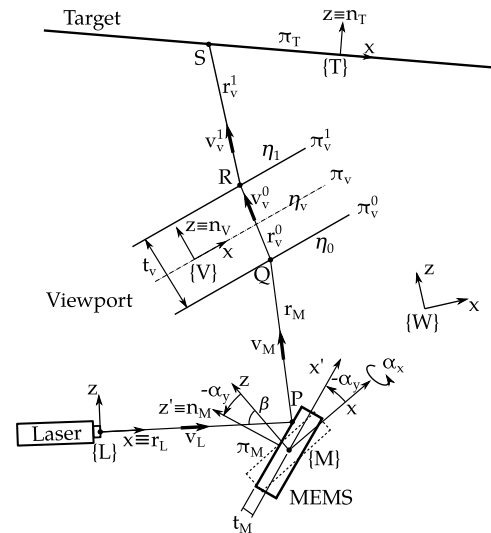


FIGURE 1. Underwater laser scanner model. It is drawn in 2D only for the sake of simplicity. In reality, the y directions of the different coordinate frames need not be parallel.

A. MODEL PARAMETERS

Figure 1 shows the model of the underwater laser scanner, which is composed of 4 elements: a laser point source, a biaxial mirror, a flat viewport used as transparent window separating two different media, and a target to be scanned. The following assumptions have been made when building the model:

1) LASER

The laser is considered a point light source. Its outgoing ray is modelled as a line, which represents the longitudinal

middle axis of the real ray. It is aligned with the x axis of its local reference frame $\{L\}$.

2) MEMS

The MEMS mirror is modelled as a rectangular prism that can be rotated around two of its local axes (x and y) by the angles α_x and α_y , respectively. The origin of its local reference frame $\{M\}$ is placed at the center of rotation. The reflective surface of the mirror is then represented by the plane π_M , which is at a distance t_M of the center of rotation, and whose normal is aligned with the local direction z' .

3) VIEWPORT

The flat viewport is also modelled as a rectangular prism. The middle XY plane π_V is the symmetry plane that goes through the origin of its local reference frame $\{V\}$. Parallel to it are the inner and the outer refraction planes (π_V^0 and π_V^1 , respectively) at a distance $t_V/2$ from it. Both of them deflect the light ray according to the refraction indices η_0 , η_V and η_1 . The normal vectors of the 3 planes are parallel to the local z direction.

4) TARGET

The target object to be scanned is represented by a plane, π_T . It goes through the origin of its local reference frame $\{T\}$ and its normal is aligned with the local z direction.

TABLE 1. Model parameters.

Element	Parameters
Laser	Pose ξ_L
Mirror	Pose ξ_M
	Distance to center of rotation t_M Angle pair $[\alpha_x \alpha_y]$
Viewport	Pose ξ_V
	Thickness t_V Indices of refraction η_0, η_V, η_1
Target	Pose ξ_T

The different elements of the model are parameterized according to table 1. Once the model has been built, the only parameters that can be actuated in order to perform the scan of a target are $[\alpha_x \alpha_y]$. Therefore, all the rest are assumed constant and defined as $\Gamma = [\xi_L \xi_M t_M \xi_V t_V \eta_0 \eta_V \eta_1 \xi_T]$.

As a side note, it should be highlighted that the presented model for underwater scanning can be immediately applied to in-air scanning without viewport by setting $\eta_0 = \eta_V = \eta_1$.

B. APPLYING THE MODEL

The first goal of the model is to express the position of the scanned point S with respect to the world reference frame as a function of all the model parameters. Formally, ${}^{\{W\}}S = f(\Gamma, [\alpha_x \alpha_y])$. Since all the parameters in Γ are assumed constant, for a given configuration of the scanner the final point only depends on the actuated angles of the mirror:

$${}^{\{W\}}S = f(\alpha_x, \alpha_y) \quad (1)$$

Algorithm 1 Computing S in $\{W\}$ as a Function of the Model Parameters

```

1: procedure PROJECT( $\alpha_x, \alpha_y$ )
2:    $\mathbf{O} \leftarrow [0 \ 0 \ 0]^T$ 
3:    $\mathbf{e}_x \leftarrow [1 \ 0 \ 0]^T$ 
4:    $\mathbf{e}_z \leftarrow [0 \ 0 \ 1]^T$ 
5:
6:   // Laser
7:    ${}^{\{L\}}\mathbf{r}_L \leftarrow (\mathbf{O}, \mathbf{e}_x)$ 
8:    $\mathbf{r}_L \leftarrow \xi_L \oplus {}^{\{L\}}\mathbf{r}_L$ 
9:
10:  // Mirror
11:   $\mathbf{z}' \leftarrow \mathbf{R}(\alpha_x, \alpha_y)\mathbf{z}$ 
12:   ${}^{\{M\}}\pi_M \leftarrow (\mathbf{z}', t_M)$ 
13:   $\pi_M \leftarrow \xi_M \oplus {}^{\{M\}}\pi_M$ 
14:   $\mathbf{n}_M \leftarrow \text{NORMAL}(\pi_M)$ 
15:   $\mathbf{P} \leftarrow \text{INTERSECT}(\mathbf{r}_L, \pi_M)$  ▷ eq. (20)
16:   $\mathbf{v}_M \leftarrow \text{REFLECT}(\mathbf{v}_L, \mathbf{n}_M)$  ▷ eq. (18)
17:   $\mathbf{r}_M \leftarrow (\mathbf{P}, \mathbf{v}_M)$ 
18:
19:  // Viewport
20:   ${}^{\{V\}}\pi_V^0 \leftarrow (\mathbf{z}, -t_V/2)$ 
21:   ${}^{\{V\}}\pi_V^1 \leftarrow (\mathbf{z}, t_V/2)$ 
22:   $\pi_V^0 \leftarrow \xi_V \oplus {}^{\{V\}}\pi_V^0$ 
23:   $\pi_V^1 \leftarrow \xi_V \oplus {}^{\{V\}}\pi_V^1$ 
24:   $\mathbf{n}_V \leftarrow \text{NORMAL}(\pi_V^0)$ 
25:   $\mathbf{Q} \leftarrow \text{INTERSECT}(\mathbf{r}_M, \pi_V^0)$  ▷ eq. (20)
26:   $\mathbf{v}_V^0 \leftarrow \text{REFRACT}(\mathbf{v}_M, \mathbf{n}_V, \eta_0, \eta_V)$  ▷ eq. (19)
27:   $\mathbf{r}_V^0 \leftarrow (\mathbf{Q}, \mathbf{v}_V^0)$ 
28:   $\mathbf{R} \leftarrow \text{INTERSECT}(\mathbf{r}_V^0, \pi_V^1)$  ▷ eq. (20)
29:   $\mathbf{v}_V^1 \leftarrow \text{REFRACT}(\mathbf{v}_V^0, \mathbf{n}_V, \eta_V, \eta_1)$  ▷ eq. (19)
30:   $\mathbf{r}_V^1 \leftarrow (\mathbf{R}, \mathbf{v}_V^1)$ 
31:
32:  // Target
33:   ${}^{\{T\}}\pi_T \leftarrow (\mathbf{O}, \mathbf{e}_z)$ 
34:   $\pi_T \leftarrow \xi_T \oplus {}^{\{T\}}\pi_T$ 
35:   $\mathbf{S} \leftarrow \text{INTERSECT}(\mathbf{r}_V^1, \pi_T)$  ▷ eq. (20)
36: end procedure

```

The implementation of function f is conceptually shown in algorithm 1. From now on, all the variables are assumed to be referred to the world coordinate frame $\{W\}$ unless otherwise explicitly stated. Lines are defined by a point and a vector and planes by a normal vector and the distance to the reference frame (see section A-3). Please, bear in mind that this implementation is conceptual. For instance, $\mathbf{r}_L \leftarrow \xi_L \oplus {}^{\{L\}}\mathbf{r}_L$ in line 8 is meant to signify that the laser ray \mathbf{r}_L in the world coordinate frame is obtained from the laser pose ξ_L and its equation relative to the local frame ${}^{\{L\}}\mathbf{r}_L$. In reality, this operation is implemented as explained in section A-4.

The rotation matrix $\mathbf{R}(\alpha_x, \alpha_y)$ in line 11 is built as a multiplication of two matrices: one expressing the rotation around the y axis of the mirror and another around x . The mirror is considered to first rotate around the y axis an angle α_y .

Then, it rotates around the unrotated x axis an angle α_x . Since both rotations occur around the original axes, the 3×3 rotation matrix on the y axis $\mathbf{R}_y(\alpha_y)$ is premultiplied by the 3×3 rotation matrix on the x axis $\mathbf{R}_x(\alpha_x)$:

$$\mathbf{R}(\alpha_x, \alpha_y) = \mathbf{R}_x(\alpha_x) \mathbf{R}_y(\alpha_y) \quad (2)$$

C. APPLYING THE REVERSE MODEL

The proposed model can also be used to solve the inverse problem: given a point to be scanned \mathbf{S} , which angles $[\alpha_x, \alpha_y]$ should the mirror be rotated? This is a highly relevant problem because solving it allows to project the laser ray in any arbitrary direction inside its FoV. This is directly applicable to 3D scanning: when scanning a target, the laser ray must be steered in order to project a given pattern onto the scene. This pattern is a set of known setpoints that the laser needs to illuminate in a particular order. By using a biaxial MEMS mirror and solving the aforementioned inverse problem, any arbitrary pattern can in principle be projected.

Solving the inverse problem means inverting the function f in equation (1). However, doing so analytically is a very cumbersome task because the variables $[\alpha_x, \alpha_y]$ appear multiple times with different exponents and inside different trigonometric functions. Therefore, a numerical approach is followed.

1) ONE SINGLE POINT

The pair of angles $[\alpha_x^*, \alpha_y^*]$ that achieve the projection of \mathbf{S} on the target must comply with the equation

$$[\alpha_x^* \alpha_y^*] = f^{-1}(\mathbf{S}). \quad (3)$$

They can be found iteratively by using the non-linear least-squares Ceres solver [42]. In order to do so, the residual error to be minimized is

$$r = \|\mathbf{S} - f(\alpha_x, \alpha_y)\|^2 \quad (4)$$

Like in any iterative solver, an initial guess is required. In this problem, starting with $\alpha_x = \alpha_y = 0^\circ$ is usually good enough for the solver to converge to the solution.

2) APPROXIMATED FUNCTION

Real 3D scanning patterns can be made up of hundreds or thousands of setpoints. A straightforward way to compute the required angle pairs would be computing each pair of angles corresponding to each setpoint, as seen before. This approach, however, would not be efficient. Finding a more efficient approach could in principle allow the projection pattern to be changed online. This could enable the sensor to increase its resolution in interesting areas of the FoV as it scans or to decrease it in order to speed up the 3D reconstruction.

An arbitrary pattern Ω consisting of n points can be expressed as:

$$\Omega = f(\boldsymbol{\alpha}), \quad (5)$$

where

$$\Omega = \begin{bmatrix} \mathbf{S}^1 \\ \mathbf{S}^2 \\ \vdots \\ \mathbf{S}^n \end{bmatrix} \text{ and } f(\boldsymbol{\alpha}) = \begin{bmatrix} f(\alpha_x^1, \alpha_y^1) \\ f(\alpha_x^2, \alpha_y^2) \\ \vdots \\ f(\alpha_x^n, \alpha_y^n) \end{bmatrix}. \quad (6)$$

Consequently, in order to project the pattern, the set of needed angle pairs can be computed as

$$\boldsymbol{\alpha} = f^{-1}(\Omega). \quad (7)$$

Like before, finding the analytical expression of the inverse function f^{-1} is cumbersome. Instead, it would be highly beneficial to find a simpler, approximated function g that yielded approximated values for the angle pair $[\tilde{\alpha}_x, \tilde{\alpha}_y]$ corresponding to an arbitrary projection point \mathbf{S} :

$$[\tilde{\alpha}_x, \tilde{\alpha}_y] = g(\mathbf{S}) \quad (8)$$

These values would be then used to project an approximated setpoint $\tilde{\mathbf{S}}$, which would be close enough to the original \mathbf{S} :

$$\tilde{\mathbf{S}} = f(\tilde{\alpha}_x, \tilde{\alpha}_y) \approx \mathbf{S} \quad (9)$$

This would be done for all the points in the pattern:

$$\tilde{\Omega} = f(g(\Omega)) \approx \Omega \quad (10)$$

More specifically, the function g is chosen to be a 5th polynomial on S_x and S_y :

$$\tilde{\alpha}_x = g(S_x, S_y, \boldsymbol{\sigma}_x) \quad (11)$$

$$\tilde{\alpha}_y = g(S_x, S_y, \boldsymbol{\sigma}_y) \quad (12)$$

$\boldsymbol{\sigma}_x$ and $\boldsymbol{\sigma}_y$ are the coefficient vectors of the polynomial that yield the corresponding approximated angles $\tilde{\alpha}_x$ and $\tilde{\alpha}_y$, respectively. The z component of \mathbf{S} is not used because it is a redundant parameter that is determined by the position of the target plane π_T .

The coefficient vectors $\boldsymbol{\sigma}_x$ and $\boldsymbol{\sigma}_y$ are found numerically using Ceres [42]. First, an arbitrary pattern Ω containing n setpoints is defined. Then, the residual passed to the solver is

$$r = \sum_{0 \leq i \leq n} \|\mathbf{S}^i - f(\tilde{\alpha}_x^i, \tilde{\alpha}_y^i)\|^2, \quad (13)$$

where $\tilde{\alpha}_x^i$ and $\tilde{\alpha}_y^i$ are computed using equations (11) and (12). The resulting joint formal expression for the optimized coefficient vectors $[\boldsymbol{\sigma}_x^* \boldsymbol{\sigma}_y^*]$ is then:

$$[\boldsymbol{\sigma}_x^* \boldsymbol{\sigma}_y^*] = \arg \min_{\boldsymbol{\sigma}_x, \boldsymbol{\sigma}_y} \sum_{0 \leq i \leq n} \|\mathbf{S}^i - f(g(S_x^i, S_y^i, \boldsymbol{\sigma}_x), g(S_x^i, S_y^i, \boldsymbol{\sigma}_y))\|^2 \quad (14)$$

In order to evaluate the performance of the proposed approximated function, a pattern consisting in 50×50 equidistant setpoints spread over the whole FoV is used. Please note that these are waypoints through which the scanning pattern will pass. The values of all the different sensor

parameters are gathered in table 2. Please note that the target is at a distance of 1 m from the MEMS.

Computing each angle pair corresponding to each of the setpoints takes around 7 s for such a pattern. If the pattern were required to change, new angles would need to be computed once again, unless some of the setpoints could be reused or interpolated. Using the approximated approach, the required time to compute all the angles is greatly reduced to 7 ms for the same pattern. Regarding accuracy, the maximum and average point-to-point distances between the desired and the approximated patterns are 0.013 mm and 0.002 mm, respectively. The resulting angles computed using the approximated function are also very similar to the ones obtained point by point: their maximum and average differences are 0.12° and 0.07° , respectively. These results confirm that the 5th-order polynomial is a suitable approximation function for f^{-1} , since it allows to reduce computation time while keeping a very high accuracy. Computing the coefficient vectors σ_x and σ_y takes around 1 minute, but it only needs to be done once because they would only vary if some of the sensor parameters changed.

The order of the polynomial g has been chosen to be 5 heuristically. Additional investigation of more potentially appropriate functions could be carried out. However, given that the evaluation times for different polynomial orders are very similar, their use would probably not increase the sensor performance noticeably.

TABLE 2. Nominal values for the model parameters. The units are millimeters for lengths and degrees for angles.

$\xi_L = [0 \ 0 \ 0 \ 0 \ 0]^T$	$\xi_M = [10 \ \delta \ 0 \ \theta_M \ 0]^T$
$\xi_V = [0 \ 0 \ l_V \ 0 \ \theta_V \ 0]^T$	$\xi_T = [0 \ 0 \ 1000 \ 0 \ \theta_T \ 0]^T$
$\delta = 0$	$\theta_M = -70$
$l_V = 100$	$\theta_V = 0$
$\theta_T = 0$	
$t_M = 0$	$t_V = 20$
$\eta_0 = 1.00$	$\eta_V = 1.51$
$\eta_1 = 1.33$	$\alpha_x, \alpha_y = [-5, 5]$

IV. STUDY OF PARAMETER INFLUENCE

This section is devoted to examining the influence that each one of the model parameters has on global performance. In order to study the individual effect of each parameter, they will be modified one at a time while keeping the rest constant. Let us consider a synthetic model with the nominal values of all the parameters defined in table 2. The poses ξ of all the elements are referenced to the world coordinate frame $\{W\}$. The rotation angles around the local y axes of the MEMS, viewport and target are represented with θ_M , θ_V and θ_T , respectively. δ is the displacement of the mirror reference frame $\{M\}$ in world y and z directions. l_V is the position of the reference frame of the viewport in the world z direction. A schematic top-down view of this configuration is shown

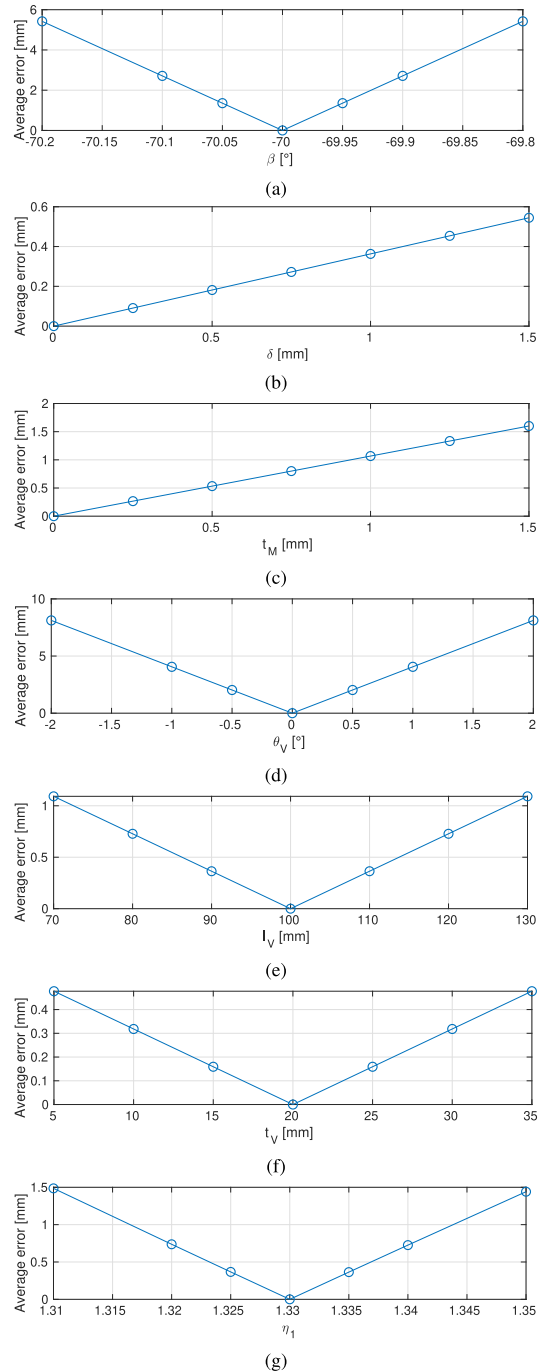


FIGURE 2. Influence of miscalibrating each parameter. Please note the different error scales in the vertical axes.

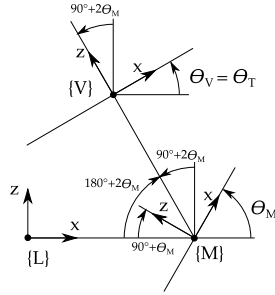


FIGURE 3. Viewport and target are kept perpendicular to the outgoing light ray for $\alpha_x = \alpha_y = 0^\circ$ throughout the simulations. Therefore, the y angles of the viewport and target depend directly on the y angle of the MEMS. In this configuration, $\theta_V = \theta_T = 90^\circ + 2\theta_M$. Please note that the values of angles θ_M , θ_V and θ_T as drawn in the figure would have negative sign.

in figure 3. Note that in this configuration:

$$\beta = 90^\circ + \theta_M, \quad (15)$$

where β is the angle defined by the laser ray and the normal of the mirror plane when $\alpha_x = \alpha_y = 0$ (see figure 1).

The different indices of refraction η are chosen from air, plexiglass and water, respectively [43]. The maximum mechanical angle for both axes of the MEMS is 5° , which makes both the horizontal and vertical mechanical FoVs equal to 10° . This is a typical value for MEMS mirrors [34].

The section is structured as follows: first, the magnitude of the error that results from miscalibrating each parameter individually is reported in section IV-A. Then, other sources of error in the final projection are described in section IV-B. Finally, the effect of the parameters in the FoV of the scanner is studied in section IV-C.

A. EFFECT OF INDIVIDUAL PARAMETER MISCALIBRATION

The model presented in this paper is aimed at evaluating and reducing the projection errors of the scanner. One of the main sources of projection errors is parameter miscalibration. A parameter is miscalibrated when its actual value differs from its nominal one. The projection error introduced by this miscalibration is defined as the distance between the nominal (expected) position of a point S and its deviated (actual) position S_m .

The projection errors due to miscalibrated parameters are shown in figure 2. The metric used in all their vertical axes is the average distance between the nominal projection and the miscalibrated one for each projected point (for n waypoints in total, with $n = 100$):

$$\text{avg error} = \frac{1}{n} \sum_{0 \leq i < n} \|S^i - S_m^i\| \quad (16)$$

There is a number of conclusions that can be drawn from these figures. First, it can be seen that θ_M (directly related to β , see equation (15)) is a very critical parameter: a slight

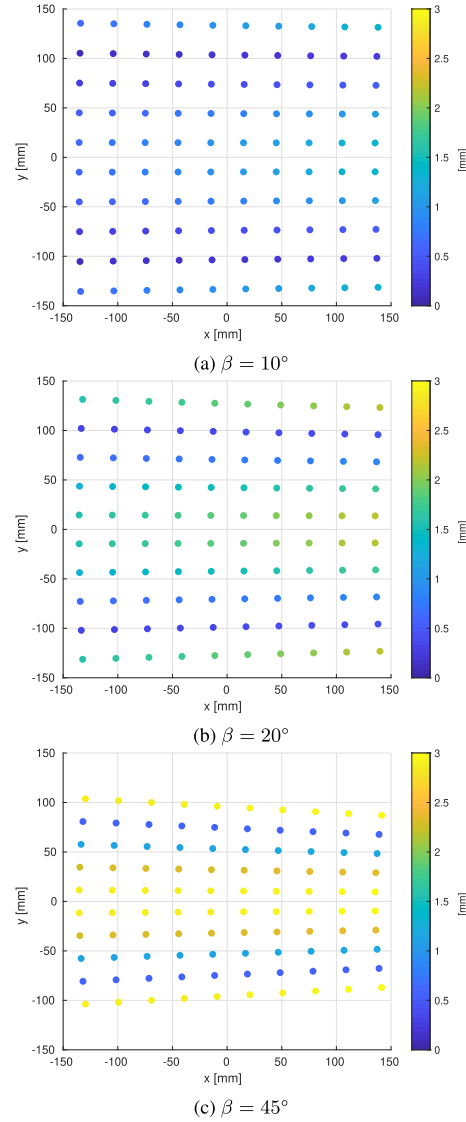


FIGURE 4. Distorted patterns for different values of β at 1 m distance. The color bar measures the distance of each point to the average x -coordinate of its corresponding vertical line.

deviation from its nominal value of only 0.2° makes the error increase rapidly (see figure 2a). Similarly, a deviation in the value of θ_V of only half a degree causes an average error well over 1 mm (see figure 2d). Regarding the index of refraction of water, it should be noted that it generally depends on different factors, such as light wavelength, salinity and temperature. For a green laser source of wavelength $\lambda = 532$ nm, the index of refraction of sea water η_1 is between 1.330 and 1.345 for a wide range of salinity values [44]. A discrepancy

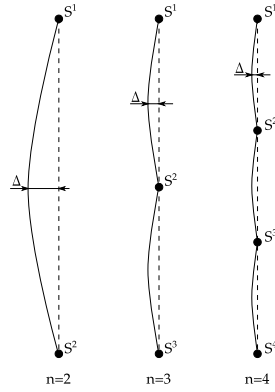


FIGURE 5. Maximum deviation Δ from the desired pattern of straight lines for an increasing number of waypoints. The curved lines show the actual light trajectories between waypoints.

of this order would entail an average error of around 1 mm (see figure 2g).

Moreover, the possible variations of δ and t_M considered are up to 1.5 mm, since that is in the order of a typical MEMS mirror diameter [34]. These values are enough to cause noticeable deviations, especially if the required precision is submillimetric (see figure 2b and 2c).

On the other hand, the influence of a slightly miscalibrated viewport thickness is not too high (1 cm of deviation entails an average error of around 0.3 mm, see figure 2f). Similarly, an error in the viewport distance of up to 1 cm would not make the average error increase further than 0.4 mm (see figure 2e). In these last two cases, standard measurement tools exist that can easily provide lengths measurements with a greater accuracy than 1 cm.

B. OTHER SOURCES OF ERROR

The presented model allows the investigation of different sources of errors other than parameter miscalibration. This section studies the magnitude of projection errors with perfectly calibrated parameters. The sources of error considered depend on the number of waypoints used per line (section IV-B1) and on the real distance to the target (section IV-B2). The aim of this section is helping to have a deeper insight on which nominal parameter values are optimal.

In the following results, both viewport and target planes are kept perpendicular to the outgoing reflected laser beam when $\alpha_x = \alpha_y = 0^\circ$ (see figure 3). Therefore:

$$\theta_V = \theta_T = 90^\circ + 2\theta_M \tag{17}$$

1) NUMBER OF WAYPOINTS PER LINE

The model presented in this paper can be used to compensate for the distortions introduced by the biaxial mirror and by

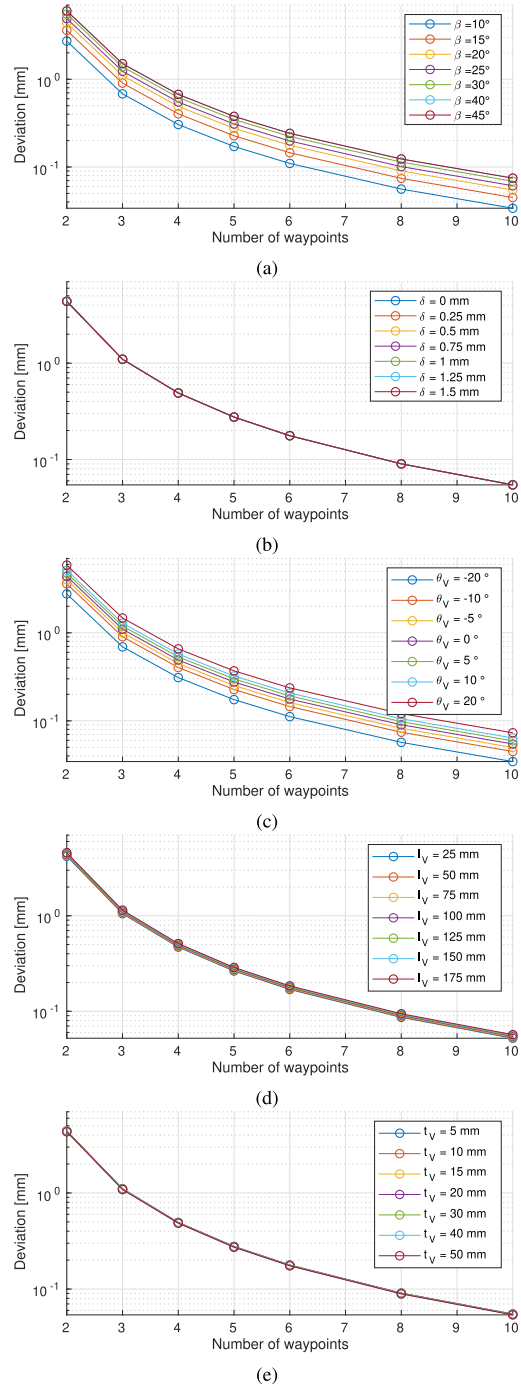


FIGURE 6. Influence of the number of waypoints per line in the maximum deviation from the straight light between waypoints. Please note that the vertical axes use a logarithmic scale.

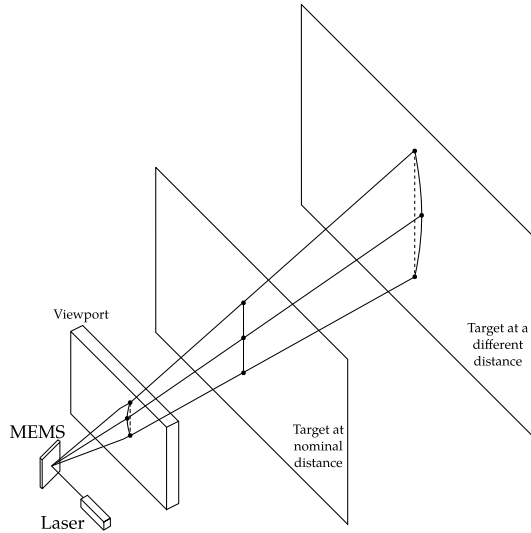


FIGURE 7. Scheme of the projection of straight lines on a target placed at a nominal distance. In order to achieve this, the points at the outside surface of the viewport do not lie exactly on top of each other. Therefore, the laser surface is actually a cone with very slight curvatures. As a consequence, when the target lies at a different distance, the waypoints will not be exactly on top of each other. Please note that the curvature of this surface is in reality much lower than drawn.

the double refraction, which mainly depend on β [39] (see figure 4). This way, it can compute the required angles to project a given number of waypoints at their desired positions. Nonetheless, the movement of the mirror between waypoints still follows the distorted trajectories shown in figure 4, introducing a distortion with respect to the corresponding straight line.

The deviation Δ of the actual light pattern from the desired one largely depends on the number of waypoints used per line, as shown in figure 5. Intuitively, the deviation Δ should decrease when more waypoints per line are used: the smaller the distance between two consecutive waypoints, also the smaller the deviation Δ . The actual evolution is depicted in figure 6.

Some conclusions can be drawn from these plots. In general, the higher the number of waypoints per line, the lower the deviation. However, this introduces a trade-off, since increasing the number of waypoints per line may in principle mean higher memory requirements and scan time.

The deviation can also be minimized by reducing β (see figure 6a). This is an expected result, since a smaller β naturally introduces less deviation, as seen in figure 4. However, this reduction is usually limited in practice by the laser module (or an alternative mirror combination), which would obstruct the reflected light for $\beta \rightarrow 0^\circ$. Another reason to keep β as small as possible is to increase the effective reflective area of the mirror (in reality, a laser ray has a certain beam diameter and is not a 1D line).

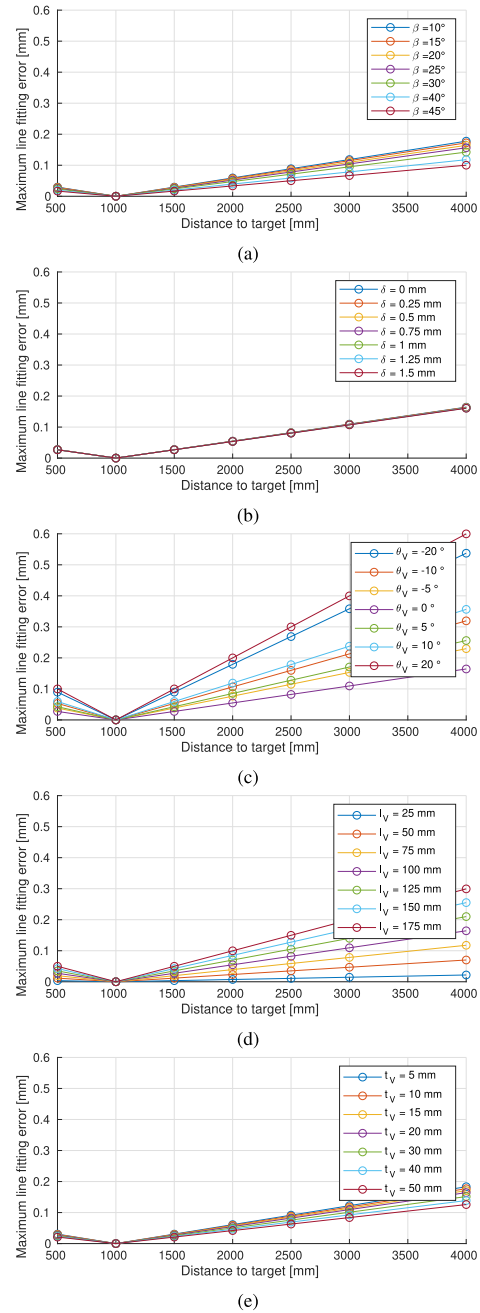


FIGURE 8. Maximum line fitting error when the actual scan plane does not necessarily lie at the nominal distance of 1000 mm.

In a similar fashion, increasing θ_V in the negative direction has a positive influence in the pattern deviation (see figure 6c). However, making $\theta_V \neq 0^\circ$ has other implications

(see section IV-B2). The rest of the parameters (δ , l_V and t_V) have a much smaller influence (see figure 6b, 6d and 6e).

2) DIFFERENT TARGET DISTANCE

The set of angles required to project straight lines depends on the distance to the target, so a certain value needs to be chosen as nominal. However, during the actual scanning, the target may lie at a different distance, which is not known a priori. This fact results in the projection points corresponding to the same line not lying exactly one on top of the other. The reason for it is that the double refraction causes points R^i in figure 1 not to form a perfectly straight line. Therefore, the outgoing light planes are in reality curved surfaces with very slight curvatures (see figure 7).

In order to assess the actual magnitude of the error, a straight line was fitted using least squares to each line of waypoints S^i . The metric used is the maximum distance from the points in the pattern to their corresponding fitted line. Results are shown in figure 8.

It can be seen that the most influential parameters are viewport distance and orientation. With a wise choice of a small viewport distance l_V and with $\theta_V = 0^\circ$, the maximum distance to the corresponding fitted line is well below 0.05 mm (see figures 8c and 8d), even if the scanned target lies at a distance of 4 m (whereas the angles have been calibrated for 1 m). The rest of the parameters (β , δ and t_V) have a much lower influence on the magnitude of this error (see figures 8a, 8b and 8e).

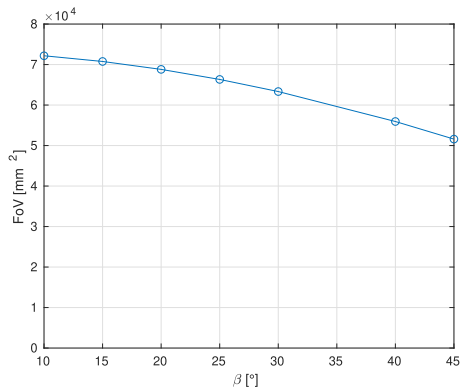


FIGURE 9. FoV at a distance of 1 m.

C. FIELD OF VIEW

A relevant performance characteristic of a scanner is its FoV, which should be as big as possible. In the presented model, the area covered by the FoV is mainly dependent on β , as plotted in figure 9. According to this graph, reducing β also contributes to increasing the scanner's FoV.

V. CONCLUSION AND FUTURE WORK

The model presented in this paper constitutes a tool to prove that using a biaxial mirror as deflecting element in an

underwater 3D scanner can minimize the distortions introduced by a flat refractive viewport. Moreover, the model is also useful to study the theoretical performance of the scanner for a set of variable parameters. One of the first conclusions that can be drawn is that calibrating properly all the parameters is essential for an accurate light projection, as has been proved with synthetic data.

Furthermore, the model can be used to have a better understanding with respect to the optimal nominal values of the different parameters. It has been proved that the angle β should be kept small, so a reasonable value for the MEMS orientation with respect to the laser could be $\theta_M \approx -80^\circ$. The flat viewport should be placed close to the MEMS (l_V as small as possible) and perpendicular to the reflected laser ray ($\theta_V \approx 0^\circ$). Viewport thickness does not influence much, so in practice its value will mainly depend on mechanical and production requirements. It is also important to keep $\delta \approx 0^\circ$, so laser and MEMS should be as well aligned as possible. With this configuration and using a high enough number of waypoints, the errors from the studied sources should not introduce noticeable deviations in the straightness of the laser lines, even when scanning at different distances.

Coming back to the contributions of this work introduced in section I, it has been proved that the presented approach can in principle reduce significantly the 3D reconstruction time while keeping a high accuracy. The authors in [11] claim that using a 1-DoF mirror, the maximum fitting error at a distance of 100 mm was 0.195 mm using elliptic cones and 4.05 mm using planes. In our approach with a 2-DoF mirror, and using the discussed nominal values for each parameter, the maximum fitting error using planes is less than 0.1 mm at a 4 m distance.

Finally, changes of the index of refraction of the water η_1 should be taken into account. By following these recommendations, submillimetric accuracy in the projection of straight lines could in principle be achieved for a wide range of distances to the target.

The next step in this line of research should be building a first prototype of the scanner using the knowledge presented in this paper. Once it has been built, its implementation and testing will likely yield hardware-related sources of error which have not been considered yet. Examples of these types of sources are limitations in the MEMS dynamic control and in the accuracy of the calibration result.

APPENDIX A GEOMETRICAL CONCEPTS

This appendix first gathers three well-known principles, namely reflection, refraction and line-plane intersection, upon which the mathematical model is built. Then, coordinate composition to transform points and vectors into different reference frames are reviewed.

1) REFLECTION (figure 10a)

The direction vector of the light ray resulting from the reflection of a ray with vector v_{in} at a surface with normal n

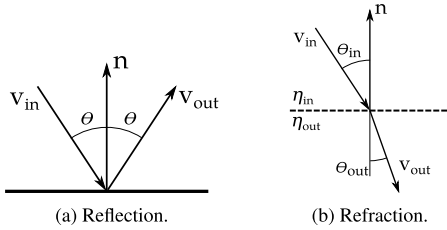


FIGURE 10. The direction of a light ray impacting on a surface changes due to either reflection or refraction, mainly depending on the properties of the surface material.

(given that \mathbf{v}_{in} and \mathbf{n} are unit vectors) is computed as:

$$\mathbf{v}_{out} = \mathbf{v}_{in} - 2(\mathbf{v}_{in} \cdot \mathbf{n}) \mathbf{n} \quad (18)$$

2) REFRACTION (figure 10b)

In case a light ray with vector \mathbf{v}_{in} encounters a refractive surface with normal \mathbf{n} that separates two media with different refraction coefficients (η_{in} and η_{out} , respectively), and assuming that \mathbf{v}_{in} and \mathbf{n} are unit vectors, the direction vector of the resulting refracted ray can be expressed as:

$$\mathbf{v}_{out} = \frac{\eta_{in}}{\eta_{out}} (\mathbf{n} \times (-\mathbf{n} \times \mathbf{v}_{in})) - \mathbf{n} \sqrt{1 - \left(\frac{\eta_{in}}{\eta_{out}}\right)^2 (\mathbf{n} \times \mathbf{v}_{in}) \cdot (\mathbf{n} \times \mathbf{v}_{in})} \quad (19)$$

3) LINE-PLANE INTERSECTION

The intersection point \mathbf{P} between a line $\mathbf{l} = \mathbf{l}_0 + \lambda \mathbf{v}$ and a plane $\pi \equiv \mathbf{n} \cdot \mathbf{x} = d$ is computed as:

$$\mathbf{P} = \mathbf{l} \cap \pi = \frac{d - \mathbf{n} \cdot \mathbf{l}_0}{\mathbf{n} \cdot \mathbf{v}} \mathbf{v} + \mathbf{l}_0 \quad (20)$$

As seen here, a line is defined by a point and a direction vector $\mathbf{l} = (\mathbf{l}_0, \mathbf{v})$, whereas a plane is defined by its normal vector and the distance to the reference frame $\pi = (\mathbf{n}, d)$.

4) 3D POSE COMPOSITION

The pose of a given coordinate frame A with respect to the world coordinate frame W is given by ${}^W\xi_{\{A\}}$ but, for the sake of readability, from now on it will be named as ξ_A . In 3D space, it is made up of 6 DoF:

$$\xi_A = [x \ y \ z \ \phi \ \theta \ \psi]^T \quad (21)$$

In order to change the reference frame in which the 3D coordinates of a point are expressed, the following equation is generally used:

$${}^W\mathbf{P} = {}^W\mathbf{t}_{\{A\}} + {}^W\mathbf{R}_{\{A\}} \{A\}\mathbf{P}, \quad (22)$$

where ${}^W\mathbf{t}_{\{A\}}$ is the relative translation of $\{A\}$ with respect to $\{W\}$ and ${}^W\mathbf{R}_{\{A\}}$ is the relative rotation of $\{A\}$ with respect to $\{W\}$. This operation is called composition and can be expressed in a more compact form:

$${}^W\mathbf{P} = \xi_A \oplus \{A\}\mathbf{P} \quad (23)$$

A more detailed explanation of compound geometrical relationships in 3D can be found in the appendix A of [45]. The coordinates of a vector can be expressed with respect to a different coordinate frame by simply rotating it:

$${}^W\mathbf{v} = {}^W\mathbf{R}_{\{A\}} \{A\}\mathbf{v}, \quad (24)$$

Lines and planes can be characterized by using one point and one vector. A line is defined by any arbitrary point on it and its direction vector, whereas a plane is represented by any arbitrary point on it and its normal vector. Therefore, expressing a line or a plane with respect to a different reference frame boils down to applying the corresponding transformation to its defining point and vector.

ABBREVIATIONS

AOD	acousto-optic deflector
AUV	autonomous underwater vehicle
DoF	degree of freedom
EOD	electro-optic deflector
FoV	field of view
LiDAR	light detection and ranging
LLS	laser line scanner
MEMS	microelectromechanical system
SfM	structure from motion
SONAR	sound navigation ranging
UUV	unmanned underwater vehicle

REFERENCES

- [1] K. Himri, P. Ridaou, N. Gracias, A. Palomer, N. Palomeras, and R. Pi, "Semantic SLAM for an AUV using object recognition from point clouds," *IFAC-PapersOnLine*, vol. 51, no. 29, pp. 360–365, 2018. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S2405896318321864>
- [2] A. Palomer, P. Ridaou, and D. Ribas, "Inspection of an underwater structure using point-cloud SLAM with an AUV and a laser scanner," *J. Field Robot.*, vol. 36, no. 8, pp. 1333–1344, Dec. 2019.
- [3] A. Palomer, P. Ridaou, D. Youakim, D. Ribas, J. Forest, Y. Pettillot, A. P. Monfort, and P. J. Sanz, "3D laser scanner for underwater manipulation," *Sensors*, vol. 18, no. 4, pp. 1–14, 2018.
- [4] F. R. Dalgleish, S. Tetlow, and R. L. Allwood, "Experiments in laser-assisted visual sensing for AUV navigation," *Control Eng. Pract.*, vol. 12, no. 12, pp. 1561–1573, Dec. 2004. [Online]. Available: https://ac.els-cdn.com/S0967066103002594/1-s2.0-S0967066103002594-main.pdf?_tid=7d62f36a-e123-4310-a955-cff1e6446ed1&acdnat=1548408210_461d9c24e11c75c9fccc5244ac192b6e6
- [5] M. Massot-Campos and G. Oliver-Codina, "Optical sensors and methods for underwater 3D reconstruction," *Sensors*, vol. 15, no. 12, pp. 31525–31557, Dec. 2015. [Online]. Available: <http://www.mdpi.com/1424-8220/15/12/29864>
- [6] S. T. Digumarti, G. Chaurasia, A. Taneja, R. Siegart, A. Thomas, and P. Beardsley, "Underwater 3D capture using a low-cost commercial depth camera," in *Proc. IEEE Winter Conf. Appl. Comput. Vis. (WACV)*, Mar. 2016, pp. 1–9. [Online]. Available: <http://ieeexplore.ieee.org/document/7477644/>
- [7] A. Anwer, S. S. Azhar Ali, A. Khan, and F. Meriaudeau, "Underwater 3-D scene reconstruction using kinect v2 based on physical models for refraction and time of flight correction," *IEEE Access*, vol. 5, pp. 15960–15970, 2017. [Online]. Available: <http://ieeexplore.ieee.org/document/8000305/>
- [8] S. Chourasiya, P. K. Mohapatra, and S. Tripathi, "Non-intrusive underwater measurement of mobile bottom surface," *Adv. Water Resour.*, vol. 104, pp. 76–88, Jun. 2017. [Online]. Available: <https://linkinghub.elsevier.com/retrieve/pii/S0309170817302385>
- [9] S. Chi, Z. Xie, and W. Chen, "A laser line auto-scanning system for underwater 3D reconstruction," *Sensors*, vol. 16, no. 9, p. 1534, Sep. 2016. [Online]. Available: <http://www.mdpi.com/1424-8220/16/9/1534>

- [10] A. Palomer, P. Ridaio, J. Forest, and D. Ribas, "Underwater laser scanner: Ray-based model and calibration," *IEEE/ASME Trans. Mechatronics*, vol. 24, no. 5, pp. 1986–1997, Oct. 2019.
- [11] A. Palomer, P. Ridaio, D. Ribas, and J. Forest, "Underwater 3D laser scanners: The deformation of the plane," in *Sensing and Control for Autonomous Vehicles* (Lecture Notes in Control and Information Sciences), vol. 474, Cham, Switzerland: Springer, 2017, pp. 73–88, doi: 10.1007/978-3-319-55372-6_4.
- [12] F. Bruno, G. Bianco, M. Muzzupappa, S. Barone, and A. V. Razonale, "Experimentation of structured light and stereo vision for underwater 3D reconstruction," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 4, pp. 508–518, Jul. 2011. [Online]. Available: www.elsevier.com/locate/isprsjprs
- [13] M. Castellón, A. Palomer, J. Forest, and P. Ridaio, "State of the art of underwater active optical 3D scanners," *Sensors*, vol. 19, no. 23, p. 5161, Nov. 2019. [Online]. Available: <https://www.mdpi.com/1424-8220/19/23/5161>
- [14] A. Sedlazeck and R. Koch, "Calibration of housing parameters for underwater stereo-camera rigs," in *Proc. Brit. Mach. Vis. Conf.*, 2011, pp. 1–11. [Online]. Available: <http://www.mip.informatik.uni-kiel.de>
- [15] A. Jordt-Sedlazeck and R. Koch, "Refractive structure-from-motion on underwater images," in *Proc. IEEE Int. Conf. Comput. Vis.*, Dec. 2013, pp. 57–64.
- [16] K. Istenič, N. Gracías, A. Arnaubec, J. Escartín, and R. García, "Automatic scale estimation of structure from motion based 3D models using laser scalars in underwater scenarios," *ISPRS J. Photogramm. Remote Sens.*, vol. 159, pp. 13–25, Jan. 2020.
- [17] 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," *Appl. Opt.*, vol. 57, no. 14, pp. 3927–3937, 2018. [Online]. Available: <http://ao.osa.org/abstract.cfm?URI=ao-57-14-3927>, doi: 10.1364/AO.57.003927.
- [18] M. Massot-Campos and G. Oliver-Codina, "Underwater laser-based structured light system for one-shot 3D reconstruction," in *Proc. IEEE Sensors*, Nov. 2014, pp. 1138–1141. [Online]. Available: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6985208>
- [19] M. Bleier and A. Nüchter, "LOW-cost 3D laser scanning in air or water using self-calibrating structured light," *ISPRS Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XLII-2/W3, pp. 105–112, Feb. 2017. [Online]. Available: <https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-2-W3/105/2017/isprs-archives-XLII-2-W3-105-2017.pdf>
- [20] P. Risholm, T. Kirkhus, and J. T. Thielemann, "High-resolution structured light sensor for autonomous underwater inspection," in *Proc. OCEANS MTS/IEEE Charleston*, Oct. 2018, pp. 1–5. [Online]. Available: <https://ieeexplore.ieee.org/document/8604930>
- [21] F. Lopes, H. Silva, J. M. Almeida, A. Martins, and E. Silva, "Structured light system for underwater inspection operations," in *Proc. OCEANS Genova*, May 2015, pp. 1–6. [Online]. Available: <http://ieeexplore.ieee.org/document/7271564/>
- [22] C. C. Constantinou, S. G. Loizou, and G. P. Georgiades, "An underwater laser vision system for relative 3-D posture estimation to mesh-like targets," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2016, pp. 2036–2041. [Online]. Available: <http://ieeexplore.ieee.org/document/7759320/>
- [23] A. Bodenmann, B. Thornton, and T. Ura, "Generation of high-resolution three-dimensional reconstructions of the seafloor in color using a single camera and structured light," *J. Field Robot.*, vol. 34, no. 5, pp. 833–851, Aug. 2017, doi: 10.1002/rob.21682.
- [24] G. Matos, P. D. Buschinelli, and T. Pinto, "Underwater laser triangulation sensor model with flat refractive interfaces," *IEEE J. Ocean. Eng.*, vol. 45, no. 3, pp. 937–945, Apr. Jul. 2020. [Online]. Available: <https://ieeexplore.ieee.org/document/8644043/>
- [25] *2G Robotics. Uls-200*. Accessed: Apr. 4, 2020. [Online]. Available: <https://www.2grobotics.com/products/underwater-laser-scanner-uls-200/>
- [26] 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," *Opt. Eng.*, vol. 56, no. 3, Oct. 2016, Art. no. 031212. [Online]. Available: <http://opticalengineering.spiedigitallibrary.org/article.aspx?doi=10.1117/1.OE.56.3.031212>
- [27] Luminar Technologies, Inc. *Lidar*. Accessed: Mar. 28, 2020. [Online]. Available: <https://www.luminartech.com/>
- [28] Velodyne Lidar, Inc. *Lidar*. Accessed: Mar. 28, 2020. [Online]. Available: <https://velodynelidar.com/>
- [29] Quanergy Systems, Inc. *Lidar*. Accessed: Mar. 28, 2020. [Online]. Available: <https://quanergy.com/technology-2/#LiDAR>
- [30] Ouster, Inc. *Lidar*. Accessed: Mar. 28, 2020. [Online]. Available: <https://ouster.com/>
- [31] Valeo. *Valeo Scala*. Accessed: Mar. 28, 2020. [Online]. Available: <https://www.valeo.com/en/valeo-scala/>
- [32] Ledartech. *Lidar*. Accessed: Mar. 28, 2020. [Online]. Available: <https://ledartech.com/>
- [33] Blickfeld GmbH. *MEMS-Based Lidar*. Accessed: Mar. 28, 2020. [Online]. Available: <https://www.blickfeld.com/technology/>
- [34] Mirrorcle Technologies, Inc. *MEMS Mirrors Catalogue*. Accessed: Mar. 28, 2020. [Online]. Available: <https://www.mirrorcletech.com/wp/products/mems-mirrors/>
- [35] K. K. H. Photonics. *MEMS Mirrors Catalogue*. Accessed: Mar. 28, 2020. [Online]. Available: <https://www.hamamatsu.com/jp/en/product/optical-components/mems-mirror/index.html>
- [36] S. T. S. Holmstrom, U. Baran, and H. Urey, "MEMS laser scanners: A review," *J. Microelectromech. Syst.*, vol. 23, no. 2, pp. 259–275, Apr. 2014. [Online]. Available: <http://ieeexplore.ieee.org/document/6714402/>
- [37] M. Brown and H. Urey, "MEMS Microdisplays," in *Handbook of Visual Display Technology*. Berlin, Germany: Springer, 2015, pp. 1–15. [Online]. Available: http://link.springer.com/10.1007/978-3-642-35947-7_128-2
- [38] Y. Song, R. M. Panas, and J. B. Hopkins, "A review of micromirror arrays," *Precis. Eng.*, vol. 51, pp. 729–761, Jan. 2018. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0141635917302210>
- [39] M. Hafez, T. Sidler, and R.-P. Salathé, "Study of the beam path distortion profiles generated by a two-axis tilt single-mirror laser scanner," *Opt. Eng.*, vol. 42, no. 4, pp. 1048–1057, Apr. 2003. [Online]. Available: <http://opticalengineering.spiedigitallibrary.org/article.aspx?doi=10.1117/1.1557694>
- [40] P. Bechtold, R. Hohenstein, and M. Schmidt, "Evaluation of disparate laser beam deflection technologies by means of number and rate of resolvable spots," *Opt. Lett.*, vol. 38, no. 16, p. 2934, Aug. 2013. [Online]. Available: <https://www.osapublishing.org/abstract.cfm?URI=ol-38-16-2934>
- [41] G. R. B. E. Römer and P. Bechtold, "Electro-optic and acousto-optic laser beam scanners," *Phys. Procedia*, vol. 56, pp. 29–39, Jan. 2014. [Online]. Available: www.sciencedirect.com
- [42] S. Agarwal and K. Mierle. *Ceres Solver*. Accessed: Mar. 28, 2020. [Online]. Available: <http://ceres-solver.org>
- [43] E. Hecht, *Optics*, 5th ed. London, U.K.: Pearson, 2017. [Online]. Available: www.pearsonglobaleditions.com
- [44] R. W. Austin and G. Halikas, "The index of refraction of seawater," Visibility Lab., Univ. California, Oakland, CA, USA, Tech. Rep. 76-1, 1976. [Online]. Available: http://misclab.umeoce.maine.edu/education/VisibilityLab/reports/SIO_76-1.pdf
- [45] R. Smith, M. Self, and P. Cheeseman, "Estimating uncertain spatial relationships in robotics," in *Autonomous Robot Vehicles*. New York, NY, USA: Springer, 1990, pp. 167–193. [Online]. Available: https://link.springer.com/chapter/10.1007/978-1-4613-8997-2_14



MIGUEL CASTILLÓN received the B.Sc. degree in industrial engineering from the University of Zaragoza, Spain, in 2015, and the M.Sc. degree in mechanical engineering from KU Leuven, Belgium, in 2018. He is currently pursuing the Ph.D. degree in robotics with the University of Girona, Spain. His research interests include computer vision applied to robotics and autonomous navigation.



ALBERT PALOMER received the M.Sc. degree in industrial engineering, the M.Sc. degree in information technologies and automation, and the Ph.D. degree from the University of Girona, in 2012, 2013, and 2018, respectively. In 2012, he joined the Underwater Robotics Research Center (CIRS), University of Girona. He is currently a Software Engineer with IQUA Robotics. His research interests include AUVs, on underwater 3D perception, and underwater navigation.



JOSEP FOREST received the B.Sc. degree in industrial informatics from the University of Girona, in 1992, the M.Sc. degree in electronics engineering from the Autonomous University of Barcelona, in 1998, and the Ph.D. degree from the University of Girona, in 2004. His research interests include 3D-machine vision, including laser triangulation, calibration, detection, Point Cloud processing, usability of 3D applied to the industry for dimensional testing, and quality control applications.



PERE RIDAO (Member, IEEE) received the Ph.D. degree in industrial engineering from the University of Girona, Spain, in 2001. He is currently the Director of the Computer Vision and Robotics Research Institute (VICOROB), the Head of the Underwater Robotics Research Center (CIRS), and an Associate Professor with the Department of Computer Engineering, University of Girona. Since 1997, he has been participated in 24 research projects (15 European and nine National), he is an author of more than 100 publications, and he has directed nine Ph.D. thesis (four more under direction) and 14 M.S. thesis. He is also a coauthor of four licenses and one Spanish/European patent, being the Co-Founder of Iqua Robotics S.L. spin-off company. His research interests include designing and developing Autonomous Underwater Vehicles for 3D Mapping and Intervention. He has served as the Chair of the IFAC's Technical Committee on Marine Systems.

...

4

UNDERWATER 3D SCANNER TO COUNTERACT REFRACTION

IN this chapter we present the underwater 3D scanner prototype that we built based on the ray-tracing model introduced in Chapter 3. We discuss the *black-box* projection model and explain in detail the characteristics of the hardware components and their connections. This simplified projection model relates the output direction of the projected laser beam with the input mirror voltages only using polynomials. This way, the calibration process is much simpler than if the whole projection model of Chapter 3 was used. Finally, we perform the experimental performance characterization of the prototype in the water tank at the [CIRS](#).

Title: Underwater 3D scanner to counteract refraction: calibration and experimental results
Authors: **Miguel Castellón**, Josep Forest, and Pere Ridao
Journal: IEEE/ASME Transactions on Mechatronics
Volume: , Published: 2022
DOI: [10.1109/TMECH.2022.3170504](https://doi.org/10.1109/TMECH.2022.3170504)
Quality index: JCR2021 Engineering, Electrical & Electronic, Impact Factor: 5.867, Q1
(47/278)

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.



IEEE/ASME TRANSACTIONS ON MECHATRONICS

1

Underwater 3D Scanner to Counteract Refraction: Calibration and Experimental Results

Miguel Castellón ^{ORCID}, Josep Forest ^{ORCID}, and Pere Ridao ^{ORCID}, *Member, IEEE*

Abstract—Underwater 3-D laser scanners are an essential type of sensors used by unmanned underwater vehicle (UUVs) for operations such as inspection, navigation, and object recognition and manipulation. This article presents a novel 3-D laser scanner, which uses a 2-axis mirror to project straight lines into the water by compensating for refraction-related distortions. This is achieved by projecting optimally curved lines, so that the refraction when they enter the water transforms them into straight lines. The relevance of this approach lies in the fact that 3-D triangulation using planes is noticeably faster than using elliptic cones. The goal of this work is twofold: first, to prove that refraction-related distortions can in practice be compensated for by using a 2-axis mirror, and second, to present a simple calibration algorithm that only needs to compute the coefficients of polynomial functions. To the best of the authors' knowledge, the prototype presented in this article is the first laser line scanner that actively counteracts the refraction of the projected light in the context of underwater robotics.

Index Terms—3-D reconstruction, 3-D sensing, autonomous underwater vehicle (AUV), underwater 3-D scanner.

NOMENCLATURE

Abbreviations

AUV	Autonomous underwater vehicle.
CMOS	Complementary metal-oxide-semiconductor.
DAC	Digital-analog converter.
DoF	Degree of freedom.
FPGA	Field-programmable gate array.
FoV	Field of view.
fps	Frames per second.
LLS	Laser line scanner.
LPF	Low-pass filter.

Manuscript received January 18, 2022; accepted April 23, 2022. Recommended by Technical Editor J. Liu and Senior Editor K. J. Kyriakopoulos. This work was supported in part by the ATLANTIS project under Grant H2020-ICT-2019-2-871571, in part by the European Commission, the PER2IAUV project under Grant PID2020-115332RB-C32, in part by the Spanish Ministry of Science, and in part by the doctoral grant of the University of Girona under Grant IFUdG2019. (Corresponding author: Miguel Castellón.)

The authors are with the Computer Vision and Robotics Research Institute (VICOROB), University of Girona, 17003 Girona, Spain (e-mail: miguel.castillon@udg.edu; josep.forest@udg.edu; pere@eia.udg.edu).

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TMECH.2022.3170504>.

Digital Object Identifier 10.1109/TMECH.2022.3170504

This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License. For more information, see <https://creativecommons.org/licenses/by-nc-nd/4.0/>

MEMS	Microelectromechanical system.
SVD	Single value decomposition.
UUV	Unmanned underwater vehicle.

I. INTRODUCTION

UNDERWATER 3-D laser line scanners (LLSs) are currently being used by unmanned underwater vehicles (UUVs) in inspection [1], object recognition [2], manipulation [3], [4], and navigation [5] tasks. This type of sensors can provide the robot with a high point cloud density at relatively fast refresh rates.

A particular type of LLSs are steered-plane scanners. These scanners typically use a rotating mirror to steer a laser plane [6], [7], which allows them to acquire 3-D information of a relatively broad field of view (FoV). This characteristic makes them suitable for manipulation tasks, in which the UUV moves at slow speeds at close distances to the target structure (around 1 m). However, this approach presents two main challenges. First, their data suffers from motion distortion when they are used to scan dynamically. A method to compensate for these distortions was proposed in [8]. Second, steering the laser plane causes it to enter the refractive surface and then the water at a nonperpendicular angle. This double refraction process deforms the original light plane into an elliptic cone [9]. Performing 3-D triangulation using elliptic cones is noticeably more computationally expensive than using planes [9]. In order to be able to project light planes in the water with a steered-plane LLS, a theoretical projection model was introduced in a previous work [10]. This model proved that using a 2-axis mirror enabled the projection of optimally curved surfaces that become planes when entering water (or, more accurately speaking, cones with a negligible curvature).

This article presents the hardware realization of that theoretical model into a prototype (see Fig. 1), whose working principle is schematically shown in Fig. 2. Thanks to the calibration process, an optimally curved scanning pattern can be computed, which when going through the flat viewport is transformed into straight lines due to refraction. The laser module is a point source. However, the trigger of the camera is synchronized with the mirror so that it opens at the beginning of each line and closes at the end of it. This means that for the camera the scanning pattern is actually made up of lines.

This article has the following two main goals: 1) to prove that refraction-related distortions can in practice be compensated for by using a 2-axis mirror; and 2) to present a simple calibration

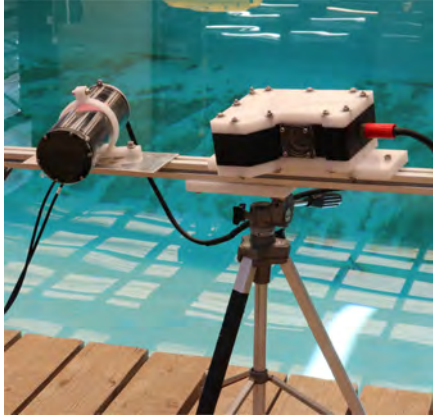


Fig. 1. Prototype of the underwater 3-D scanner. All its components are sealed together in the waterproof casing, except for the camera, which is placed in the waterproof cylinder.

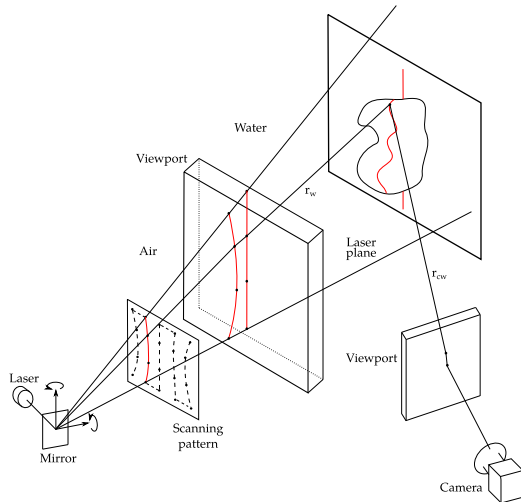


Fig. 2. Working principle of the scanner. The 2-axis mirror steers the laser beam following a set of pre-computed waypoints that form an optimally-curved set of lines, so that they become straight vertical lines when they enter the water.

algorithm that only needs to compute the coefficients of polynomial functions. These goals may be considered fulfilled if the results of 3-D triangulation using planes are accurate enough. In the context of 3-D sensing in autonomous underwater missions, a good enough accuracy should be in the order of millimeters.

A thorough review of the working principles and performance characteristics of state-of-the-art underwater 3-D laser scanners was presented in [11, Sec. 5]. However, comparing the performance of different underwater scanners in the literature is not a trivial task, since the authors report different performance

TABLE I
ACCURACY COMPARISON OF UNDERWATER TRIANGULATION SCANNERS IN THE LITERATURE

Authors	Accuracy [mm]	Range [m]
Chi [6]	1.0	1.0
Constantinou [12]	10	0.7
Lopes [13]	0.22	0.44
Massot [14]	3.6	0.7
Matos [15]	0.48	0.29
Palomer [7]	0.98	1.0
Risholm [16]	0.4	0.8
Sarafraz [17]	2.3	1.5

parameters measured in different ways. The results of this study have been summarized here in Table I for triangulation-based scanners. Broadly speaking, they all report accuracies in the order of millimeters at ranges of around 1 m. The scanning range is important, because the resolution of triangulation-based scanners worsens for increasing distances to the object [11]. As previously stated, the main goal of the presented scanner is to counteract refraction while keeping an accuracy of the 3-D reconstruction in the order of millimeters, like other systems in the literature.

The rest of this article is structured as follows. First, an overview of the system is done in Section II. Then, a new calibration algorithm based on simple projection functions is proposed in Section III. The experimental results of scanning under water are analyzed in Section IV. Finally, Section V concludes this article.

II. SENSOR DESCRIPTION

This section provides with an overview of the presented prototype. First, its theoretical projection model is explained in Section II-A. Then, a detailed explanation of its working principle is presented in Section II-B. Finally, the performance characteristics of the hardware components used in the prototype are gathered in Section II-C. Throughout the whole explanation, all the geometric variables are assumed referred to the camera reference frame $\{C\}$.

A. Sensor Model

The working principle of the proposed scanner is conceptually shown in Fig. 2 and it consists basically on two main steps: light projection and light sensing. Regarding the light projection part, the laser ray first comes out of the laser source and is reflected by the 2-axis mirror, which is driven to visit the sequence of precomputed waypoints in the scanning pattern. Then, the light passes through a wide-angle lens, which amplifies the incoming angle in order to increase the scanner FoV (not drawn in the figure for the sake of clarity). Later, the laser beam suffers a double refraction process due to the different refraction indices of the subsequent media (air, glass, and water), according to Snell's Law [18]. Finally, the laser beam hits the target to be scanned and bounces back toward the camera through the camera viewport, where the double refraction happens in reverse order (water to glass, and then glass to air). This last step constitutes the light sensing part.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

CASTILLÓN *et al.*: UNDERWATER 3D SCANNER TO COUNTERACT REFRACTION

3

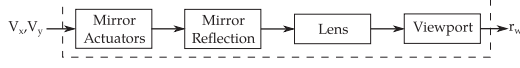


Fig. 3. Projection light procedure. The direction of the outgoing laser beam r_w is the result of four sequential processes. First, the input voltage pair $[V_x, V_y]$ tilts the mirror around two perpendicular axes. Then, the mirror reflects the light in a direction that depends on those two rotation angles. Later, the wide-angle lens opens the FoV of the beam. Finally, the ray goes through the viewport into the water, experiencing a double refraction process. In practice, only the global joint effect is considered.

Creating a complete light projection model that related the direction of the outgoing laser beam r_w with all the parameters of every component is in principle possible, as shown in [10]. This model included, among many other parameters, the thickness of the viewport and its distance to the mirror, and the offset angle of the mirror with respect to the laser source. However, this approach presents two main drawbacks when applied in practice. First, the mirror is a microelectromechanical system (MEMS) whose rotation angles show a strongly nonlinear response to the input voltages, especially for rotations close to the mechanical limits. Second, the exact optical model of the wide-angle lens is not known. These two challenges could in principle be overcome by finding suitable functions that successfully modeled these behaviors. Nonetheless, a more direct approach is followed instead: all the optical components of the scanner are treated as a black box, as shown in Fig. 3. This means that it is not needed to model the individual effect of each one of the components in the direction of the outgoing laser beam r_w . Instead, simple polynomials are found that express r_w as a function of the input voltages $[V_x, V_y]$, modeling the joint effect of all the optical components.

The outgoing laser beam r_w is modeled as a line, made up of a point on the scanning pattern \mathbf{p} and a unit direction vector \mathbf{v}_w , as shown in Fig. 4: $r_w \equiv \mathbf{p} + \lambda \mathbf{v}_w$, where $\lambda \in \mathbb{R}$. The indices (i, j) in Fig. 4 make reference to the horizontal and vertical position of the ray in the scanning pattern, respectively (see Section II-B).

Formally, the goal is finding the set of functions that relate each one of the free DoF of r_w with $[V_x, V_y]$

$$p_x = f_1(V_x, V_y) \quad (1)$$

$$p_y = f_2(V_x, V_y) \quad (2)$$

$$v_{wx} = f_3(V_x, V_y) \quad (3)$$

$$v_{wy} = f_4(V_x, V_y). \quad (4)$$

Please note that p_z is not needed to be computed because its value is forced by the fact that \mathbf{p} must lay on the projection plane π_P : $\mathbf{p} \in \pi_P$. The projection plane is an arbitrary plane chosen perpendicular to the direction of the outgoing laser beam for zero voltage (parallel to the viewport). Similarly, it is not needed to compute v_{wz} because the direction vector \mathbf{v}_w is forced to have unit norm.

As will be clear in Section II-B, the inverse problem is also extremely relevant, which pair of voltages are needed in order to project the laser beam in a certain direction? These values are

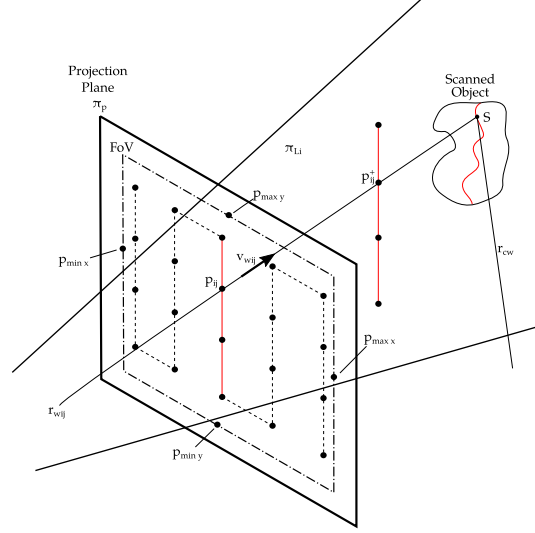


Fig. 4. Scheme of the principle behind the triangulation process followed to reconstruct the scanned 3-D points. Please note that all the elements in the scene are under water.

given by the following two functions:

$$V_x = f_5(p_x, p_y) \quad (5)$$

$$V_y = f_6(p_x, p_y). \quad (6)$$

As will be seen in Section III, all these functions turn out to be smooth enough to adjust polynomials to them.

B. 3-D Triangulation

This section explains the procedure by which the scanner reconstructs the 3-D shape of the target object. This triangulation-based approach is depicted in Fig. 4 and can be summarized in five steps.

- 1) First, compute the extreme horizontal and vertical points, which delimit the FoV of the projected laser beam on the projection plane using (1) and (2)

$$p_{\max x} = f_1(V_x = V_{\max x}, V_y = 0) \quad (7)$$

$$p_{\min x} = f_1(V_x = V_{\min x}, V_y = 0) \quad (8)$$

$$p_{\max y} = f_2(V_x = 0, V_y = V_{\max y}) \quad (9)$$

$$p_{\min y} = f_2(V_x = 0, V_y = V_{\min y}). \quad (10)$$

- 2) Design a scanning pattern. A practical choice is a set of equidistant points inside the limits of the FoV. This pattern is defined by the number of lines n and the number of waypoints per line m ($n \cdot m$ points in total). The point p_{ij} in such a pattern refers to the j th waypoint of line i . On the one hand, a higher number of waypoints per line contributes to a smaller planarity error of the outgoing light plane, as studied in [10]. This happens because the

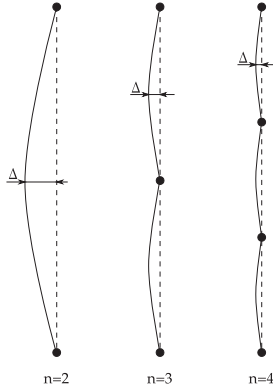


Fig. 5. Maximum deviation Δ of the actual light trajectory between waypoints with respect to the ideal one. An increasing number of waypoints per lines results in a smaller deviation.

discrete waypoints in the pattern are optimally placed to counteract refraction. Nonetheless, the mirror follows the shortest path between waypoints in angle space, which causes a deviation (see Fig. 5). On the other hand, a higher number of waypoints per line also results in a slower scanning speed (see Section IV-C).

- 3) Compute the voltage pair needed to project each point using (5) and (6)

$$\begin{aligned} V_{ijx} &= f_5(p_{ijx}, p_{ijy}) \\ V_{ijy} &= f_6(p_{ijx}, p_{ijy}). \end{aligned} \quad (11)$$

- 4) Fit a plane to the set of points corresponding to the same line. Since all the points corresponding to the same line are colinear, an intermediate step is required. For each waypoint \mathbf{p}_{ij} , another waypoint \mathbf{p}_{ij}^+ can be defined in the direction of its corresponding direction vector \mathbf{v}_{wij}

$$\mathbf{p}_{ij}^+ = \mathbf{p}_{ij} + \lambda \mathbf{v}_{wij} \quad (12)$$

where the unit direction vector is computed using (3) and (4)

$$\mathbf{v}_{wij} = \begin{bmatrix} f_3(V_{ijx}, V_{ijy}) \\ f_4(V_{ijx}, V_{ijy}) \\ \sqrt{1 - v_{ijx}^2 - v_{ijy}^2} \end{bmatrix}. \quad (13)$$

This way, the extended set of waypoints corresponding to the i th line \mathcal{P}_i can be defined as made up of all the corresponding \mathbf{p}_{ij} and \mathbf{p}_{ij}^+ . This set can now be used to compute the best-fitting plane

$$\pi_{Li} = \text{fit_plane}(\mathcal{P}_i) \quad (14)$$

where the `fit_plane` function is based on the single value decomposition (SVD) of the 3-D coordinates of the points in \mathcal{P}_i [19].

- 5) Finally, the 3-D position of each reconstructed scanned point is computed as the intersection of the camera ray

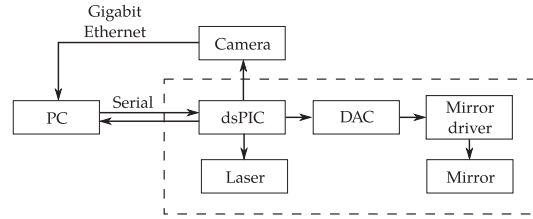


Fig. 6. Connections between the different hardware components of the proposed scanner.

with the corresponding light plane

$$S = \pi_{Li} \cap \mathbf{r}_{cw}. \quad (15)$$

The direction of the camera ray \mathbf{r}_{cw} can be computed in the reference frame of the camera $\{C\}$ as the line that passes through the camera focal point and the pixel (u, v) at which the laser line is detected:

$$\mathbf{r}_{cw} = \mathbf{0} + \lambda \begin{bmatrix} u - c_x & v - c_y \\ f_x & f_y \\ 1 \end{bmatrix}^T \quad (16)$$

where $[c_x, c_y]$ is the camera center and $[f_x, f_y]$ is the focal length.

C. Hardware

The scanner prototype is divided in two sealed water-proof boxes with transparent viewports, as shown in Fig. 1. The cylinder seals the camera, whereas the box encloses the laser module, the MEMS mirror, the wide-angle lens, and all the electronic components. The scheme of electronic connections between the different components is depicted in Fig. 6. Basically, the central computer communicates with all the elements by programming a dsPIC microcontroller, which synchronizes the rotations of the mirror with the trigger of the camera. In order to pass the required voltages to the mirror actuators, the signal coming out of the processing goes first through the digital-analog converter and then to the analog driver of the mirror. The camera transfers the information of the scanned laser line directly to the computer through a gigabit Ethernet connection. The PC may also be integrated as an on-board computer, so that the UUV can dive freely. The performance characteristics of the different components are now reviewed.

The camera used in the scanner is a CMOS sensor equipped with an FPGA for fast laser line detection. At full resolution (2048×1088 pixels) it can work in laser-detection mode at a rate of 339 fps. A lens with a focal length of 8 mm is mounted on the sensor.

The laser module projects a laser beam of 60 mW of output power at a wavelength of 520 nm. The beam diameter at the aperture is of 1 mm and the beam divergence is of 1.1 mrad.

The wide-angle lens used to increase the FoV of the scanner has an approximately linear behavior: it increases the angle of the incoming light ray by a factor of roughly 3.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

The mirror has a circular reflective surface with a diameter of 1.2 mm and maximum mechanical tilting angles of approximately $\pm 5^\circ$ in each axis. The output angle achieved by the mirror actuators is only linear with respect to the input voltage for small angles. As the input voltage gets closer to the maximum voltage, severe nonlinearities appear.

The dynamic behavior of each of both of its axes can in principle be approximated to a second-order system [20] with resonance frequency of 1.3 kHz and a damping ratio $\xi = 6.5 \times 10^{-3}$. This means that it is a highly resonant system and that even small changes in the input voltages cause large overshoots in its tilting angles. In order to limit these overshoots, an LPF is applied to the input voltage signal that avoids exciting the resonance frequency. The downside of this approach is that it limits the speed of the scanner, as reported in Section IV-C. Nonetheless, it is already enough for typical autonomous object manipulation missions. An option to increase its speed would be substituting the LPF by an inverse plant filter [21], which can reportedly drive the mirror 2 orders of magnitude faster without overshooting. Yet another option could be controlling the tilting angles in closed loop [22]. However, this last approach would need to have feedback of the actual tilting angles, which is unavailable in the current configuration.

III. CALIBRATION

This section explains in detail the procedure to calibrate the projection functions of the laser scanner in order to enable the 3-D reconstruction. In brief, this method follows three subsequent steps: first, the camera intrinsic parameters are calibrated in air according to the pinhole model, along with the distortion parameters (see Section III-A). Second, the parameters of the camera viewport are calibrated in water (see Section III-B). Finally, the coefficients of the numeric projection functions are calibrated (also in water, see Section III-C). The first two steps correspond to the application of already presented methods. The third one, however, has been designed for this prototype with the aim of reducing the complexity of the calibration algorithm. Consequently, the discussion in Section III-C is noticeably longer than Sections III-A and III-B.

A. Camera Intrinsic Parameters (in Air)

The first step consists in calibrating the intrinsic parameters of the camera: the camera center $[c_x, c_y]$ and the focal lengths $[f_x, f_y]$, as well as its radial and tangential distortion coefficients. This can be done following a standard, well-known calibration routine, such as the one implemented in the OpenCV library [23]. This step is necessary for the triangulation process [see (15) and (16)].

B. Camera Viewport (in Water)

In this step, the parameters to be calibrated are the 6-DoF pose of the camera viewport with respect to the camera, the thicknesses of the camera viewport and the refraction indices of the water and of the viewport material, respectively. The index of refraction of air is used as reference ($n_{\text{air}} = 1$).

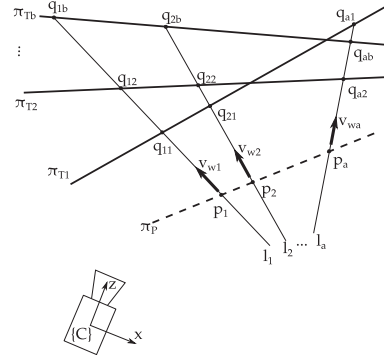


Fig. 7. In-air calibration data gathering.

The procedure basically consists in gathering an underwater dataset made up of images of a calibration pattern. Then, an optimization algorithm finds the viewport parameters by minimizing the reprojection error. The interested reader can find in [7, Sec. VI-C], a detailed explanation of the algorithm.

C. Projection Functions (in Water)

The last part of the calibration process is estimating the coefficients of the projection functions. The steps of this procedure are explained in detail in this section.

1) *Dataset Gathering*: First, a set of a voltage pairs is defined

$$\mathbf{V} = [\mathbf{V}_1 \dots \mathbf{V}_a]^T. \quad (17)$$

This set of voltages is usually chosen as an equidistant pattern covering the whole FoV of the scanner in both axes. The scanner is first placed looking at a calibration plane π_{T1} at an arbitrary position (see Fig. 7). This calibration plane has some fiducial markers attached to it. The ground-truth position of the plane with respect to the camera is computed by applying the Perspective-n-Point algorithm to those known fiducial markers. The camera grabs an image of the projected laser dot for each voltage pair in \mathbf{V} that is applied to the 2-axis mirror. This is repeated for b different positions of the calibration plane, resulting in a total of $a \times b$ images.

For each one of these images, the pixel position of the laser dot on the image plane (u, v) is found by means of the circle Hough transform [24]. The camera ray is computed by substituting the values of (u, v) in (16). The intersection of the camera ray corresponding to voltage \mathbf{V}_k with the calibration plane π_{T1} defines the 3-D position of the point \mathbf{q}_{kl} . This way, a dataset is built that contains $a \times b$ 3-D points.

2) *Voltage Line Fitting*: The next step is fitting a line to the 3-D points corresponding to the same voltage pair. For the voltage pair \mathbf{V}_k , the best-fitting line \mathbf{l}_k is defined as

$$\mathbf{l}_k \equiv \mathbf{p}_k + \lambda \mathbf{v}_{wk}. \quad (18)$$

This way, there is a set of a 3-D points \mathbf{p}_k with their corresponding unit direction vectors \mathbf{v}_{wk} . The line-fitting algorithm

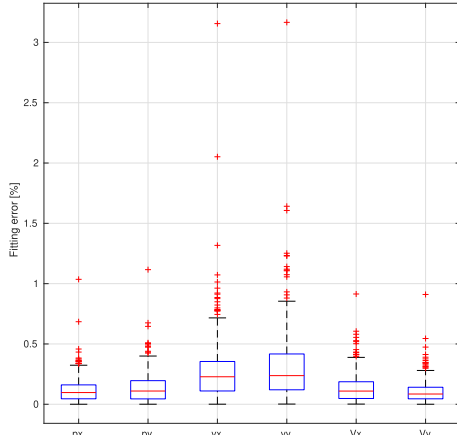


Fig. 8. Percent fitting error of the projection functions to real calibration data for functions f_1 to f_6 .

is based on SVD applied to the set of corresponding 3-D points [19].

3) *Function Coefficients*: The coefficients of all the functions in (1) to (6) are found by minimizing the following residuals:

$$\beta_1^* = \arg \min \sum_{k=1}^a \|p_{kx} - f_1(V_{kx}, V_{ky}, \beta_1)\|^2 \quad (19)$$

$$\beta_2^* = \arg \min \sum_{k=1}^a \|p_{ky} - f_2(V_{kx}, V_{ky}, \beta_2)\|^2 \quad (20)$$

$$\beta_3^* = \arg \min \sum_{k=1}^a \|w_{wkx} - f_3(V_{kx}, V_{ky}, \beta_3)\|^2 \quad (21)$$

$$\beta_4^* = \arg \min \sum_{k=1}^a \|w_{wky} - f_4(V_{kx}, V_{ky}, \beta_4)\|^2 \quad (22)$$

$$\beta_5^* = \arg \min \sum_{k=1}^a \|V_{kx} - f_5(p_{kx}, p_{ky}, \beta_5)\|^2 \quad (23)$$

$$\beta_6^* = \arg \min \sum_{k=1}^a \|V_{ky} - f_6(p_{kx}, p_{ky}, \beta_6)\|^2 \quad (24)$$

where β_i is the vector containing the coefficients of function f_i . It was experimentally chosen that all the functions f_1 to f_6 were fifth-order polynomials. In order to validate this choice, the fitting error to real calibration data is shown in Fig. 8. It can be seen in the figure that the vast majority of data points have a fitting error well below 1%. This proves that this choice of projection functions is suitable to actual data. All the minimization algorithms are implemented in C++ with the Ceres solver [25].

IV. EXPERIMENTAL RESULTS

This sections presents and discusses the numerical results of three sets of experiments. The first two of these sets are aimed at

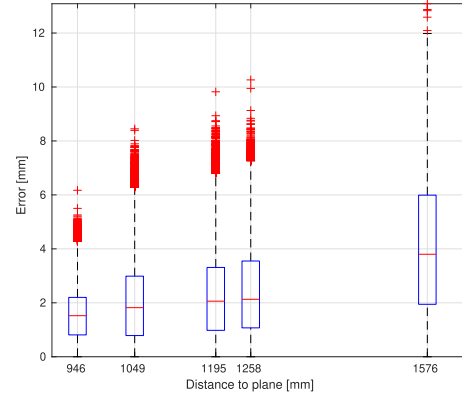


Fig. 9. Boxplot of plane fitting error.

studying the accuracy of the underwater 3-D laser scanner. On the one hand, a calibration plane with very high planarity was scanned to analyze the depth accuracy of the 3-D reconstruction at different scan ranges (see Section IV-A). On the other hand, the lateral accuracy of the scanner was assessed by scanning a calibration sphere of known radius (see Section IV-B). The goal of the third set of experiments was to study the effect of the number of waypoints per line in two important performance characteristics: accuracy and scanning speed (see Section IV-C).

All the data were gathered in the water tank of CIRS lab. The scanner was mounted on a tripod and submerged at a depth of around 1 m in clear water. In all the cases studied here, both the scanner and the scanned object (plane or sphere) were static. The standard scanning pattern used was made up of 500 lines with 50 waypoints per line (25 k waypoints in total). The FoV of the scanner is of approximately $35^\circ \times 35^\circ$. However, not the whole FoV was used in all the experiments.

With the current configuration of baseline of approximately 0.6 m between laser box and camera, the scanning range is roughly between 0.5 m and 3 m. However, most of the presented results were taken at a range between 1 m and 1.5 m, which is the most relevant for object manipulation. Measurements at shorter ranges are expected to have greater accuracy (see Fig. 13, with range of 0.6 m), whereas longer distances are typically out of the manipulation range of UUVs.

As a side note, it is worth mentioning that the apparently large number of outliers present in several of the boxplots is only due to the much larger amount of data points. The maximum outlier ratio is 0.7%.

A. Plane Fitting Error

In this set of experiments, a calibration plane was scanned at five different positions. The resulting data is shown in the boxplot of Fig. 9. The error measure is the distance of each point to the fitted plane. Additionally, the spatial distribution of the plane fitting error is plotted in Fig. 10.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

CASTILLÓN *et al.*: UNDERWATER 3D SCANNER TO COUNTERACT REFRACTION

7

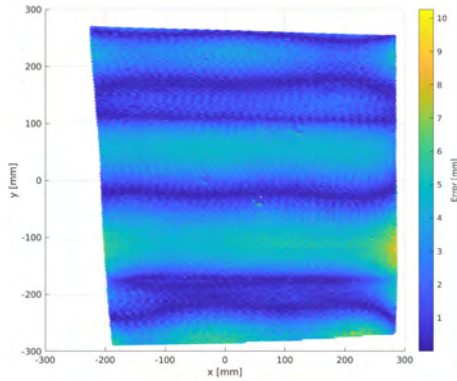


Fig. 10. Spatial distribution of the plane fitting error at an average distance of 1257.7 mm.

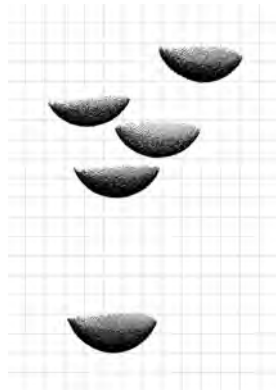


Fig. 11. Top view of the reconstructed point clouds of the calibration sphere at five different positions. The size of the blue grid is 20 mm × 20 mm.

Two main conclusions can be drawn from these charts. On the one hand, the resulting error is in the order of a few millimeters, which is a satisfactory for the typical object manipulation tasks performed by UUVs. On the other hand, it can be seen that the error increases for an increasing scanning range. This is an expected result of a triangulation-based 3-D scanner, as explained in [11]. The accuracy level is in the order of what has been reported for other systems in the literature (see Table I).

B. Sphere Fitting Error

In this set of experiments, the scanned object was a calibration sphere with a radius of 100 mm. The reconstructed point clouds are shown in Fig. 11. The sphere fitting error of those resulting point clouds along with the radius error of the least-squares fitted sphere are shown in Fig. 12. Similarly to the previous section, the magnitude of the errors are in the order of a few millimeters. In this case, the errors are distributed between the radius error and the fitting error. Consequently, establishing the relationship between error and scan range is not as simple as with the plane.

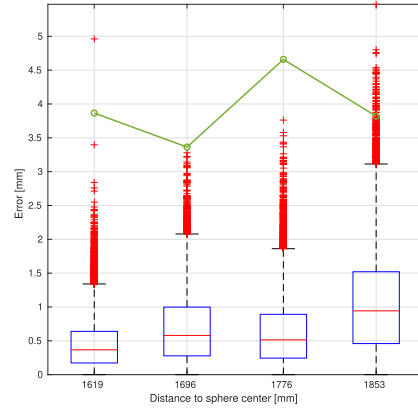


Fig. 12. Boxplot of sphere fitting error. The green line shows the radius error for a ground-truth radius of 100 mm.

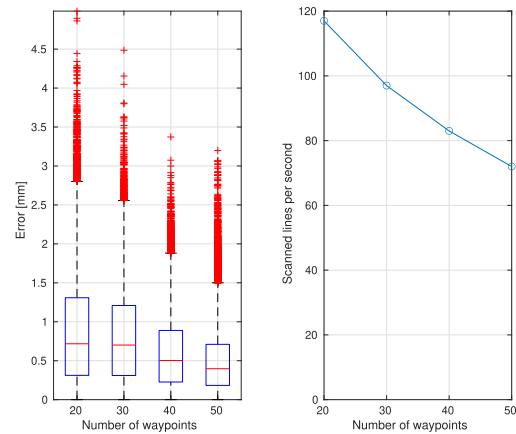


Fig. 13. Boxplot in the left shows the plane fitting residual for scan patterns with increasing number of waypoints per line. Similarly, the plot in the right shows the number of scanned lines per second for scan patterns with increasing number of waypoints per line.

C. Accuracy and Speed Versus Number of Waypoints Per Line

As mentioned in Section II-B, the choice on the number of waypoints per line in the pattern has an effect on the accuracy and speed of the 3-D reconstruction. In order to study this tradeoff in practice, scanning patterns with different number of waypoints per line were used to scan the calibration plane. The position of the plane was kept constant throughout the experiment at an approximate distance of 0.6 m. The results are shown in Fig. 13. As studied in [10], a higher number of waypoints per line means a higher accuracy. However, due to the dynamic behavior** of the mirror, it also results in a slower scanning speed. Despite having room for improvement, a scanning speed of roughly 70 lines per second is equivalent to 140 k points/s, which is already enough for a dense reconstruction of the object to be manipulated.

V. CONCLUSION

A novel 3-D underwater laser scanner has been presented that can actively counteract refraction-related distortions caused by a flat refractive surface in the projection of light. The low 3-D reconstruction errors reported in this work prove that the scanner can be used by UUVs to perform object manipulation tasks satisfactorily.

In reference to the goals of the work presented in the abstract, it has been experimentally proved that refraction-related distortion can be compensated for by using a biaxial 2-axis mirror. This has been reliably indicated by the low 3-D reconstruction errors of the different scanned objects, namely planes and spheres. Throughout the whole reconstruction process (including the calibration), it has always been assumed that the scanner can project planes in the water. Low 3-D reconstruction errors using planes consequently validate this assumption. For the same reason, the simple calibration based on numeric projection functions has been proved suitable. Therefore, and in spite of the practical limitations of the current system, we believe that the approach presented in this work can open the door to other interesting lines of research, such as optimal scanning patterns and refraction counteraction strategies.

The future steps in this line of research will be to use this 3-D underwater laser scanner mounted on a UUV in object manipulation tasks in realistic scenarios. Despite the fact that the current scanning speed is enough for typical manipulation tasks, it would be beneficial to investigate how to increase it. The most plausible strategy would be a more sophisticated dynamic control of the mirror actuators, such as the one presented in [21]. A different possible upgrade would be using a laser source with higher output power. This would probably enable 3-D reconstruction in more challenging light conditions. Finally, it will also be interesting to study the effect of varying water refraction index on the accuracy of the output point cloud. This step will be fundamental for operations in sea water.

REFERENCES

- [1] A. Palomer, P. Ridaó, and D. Ribas, "Inspection of an underwater structure using point cloud SLAM with an AUV and a laser scanner," *J. Field Robot.*, vol. 36, no. 8, pp. 1333–1344, 2019.
- [2] K. Himri, P. Ridaó, and N. Gracias, "Underwater object recognition using point-features, bayesian estimation and semantic information," *Sensors*, vol. 21, no. 5, pp. 1–27, Mar. 2021. [Online]. Available: <https://www.mdpi.com/1424-8220/21/5/1807>
- [3] A. Palomer, P. Ridaó, D. Youakim, D. Ribas, J. Forest, and Y. Pettillot, "3D laser scanner for underwater manipulation," *Sensors*, vol. 18, no. 4, pp. 1–14, 2018.
- [4] R. Pi, P. Cieslak, P. Ridaó, and P. J. Sanz, "TWINBOT: Autonomous underwater cooperative transportation," *IEEE Access*, vol. 9, pp. 37668–37684, 2021. [Online]. Available: <https://ieeexplore.ieee.org/document/9367135/>
- [5] F. R. Dalgleish, S. Tetlow, and R. L. Allwood, "Experiments in laser-assisted visual sensing for AUV navigation," *Control Eng. Pract.*, vol. 12, no. 12, pp. 1561–1573, 2004. [Online]. Available: https://ac.els-cdn.com/S0967066103002594/1-s2.0-S0967066103002594-main.pdf?_tid=7d62f36a-e123-4310-a955-cff1e6446ed1&acdnat=1548408210_461d9c24e11c759fce5244ac192b6e6
- [6] S. Chi, Z. Xie, and W. Chen, "A laser line auto-scanning system for underwater 3D reconstruction," *Sensors*, vol. 16, no. 9, Sep. 2016, Art. no. 1534. [Online]. Available: <http://www.mdpi.com/1424-8220/16/9/1534>
- [7] A. Palomer, P. Ridaó, J. Forest, and D. Ribas, "Underwater laser scanner: Ray-based model and calibration," *IEEE/ASME Trans. Mechatron.*, vol. 24, no. 5, pp. 1986–1997, Oct. 2019.
- [8] M. Castellón, R. Pi, N. Palomeras, and P. Ridaó, "Extrinsic visual-inertial calibration for motion distortion correction of underwater 3D scans," *IEEE Access*, vol. 9, pp. 93384–93398, 2021. [Online]. Available: <https://ieeexplore.ieee.org/document/9464334/>
- [9] A. Palomer, P. Ridaó, D. Ribas, and J. Forest, "Underwater 3D laser scanners: The deformation of the plane," in *Lecture Notes in Control and Inf. Sci.*, vol. 474. Cham, Switzerland: Springer, 2017, pp. 73–88. [Online]. Available: http://link.springer.com/10.1007/978-3-319-55372-6_4
- [10] M. Castellón, A. Palomer, J. Forest, and P. Ridaó, "Underwater 3D scanner model using a biaxial MEMS mirror," *IEEE Access*, vol. 9, pp. 50231–50243, 2021. [Online]. Available: <https://ieeexplore.ieee.org/document/9388657/>
- [11] M. Castellón, A. Palomer, J. Forest, and P. Ridaó, "State of the art of underwater active optical 3D scanners," *Sensors*, vol. 19, no. 23, 2019, Art. no. 5161. [Online]. Available: <https://www.mdpi.com/1424-8220/19/23/5161>
- [12] C. C. Constantinou, S. G. Loizou, and G. P. Georgiades, "An underwater laser vision system for relative 3-D posture estimation to mesh-like targets," in *Proc. IEEE Int. Conf. Intell. Robots Syst.*, 2016, pp. 2036–2041. [Online]. Available: <https://ieeexplore.ieee.org/document/7759320/>
- [13] F. Lopes, H. Silva, J. M. Almeida, A. Martins, and E. Silva, "Structured light system for underwater inspection operations," in *Proc. IEEE/MTS OCEANS Genova: Discovering Sustain. Ocean Energy for a New World*, 2015, pp. 1–6. [Online]. Available: <https://ieeexplore.ieee.org/document/7271564/>
- [14] M. Massot-Campos and G. Oliver-Codina, "Underwater laser-based structured light system for one-shot 3D reconstruction," in *Proc. IEEE Sensors*, 2014, pp. 1138–1141. [Online]. Available: <https://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6985208>
- [15] G. Matos, P. D. Buschinelli, and T. Pinto, "Underwater laser triangulation sensor model with flat refractive interfaces," *IEEE J. Ocean. Eng.*, vol. 45, no. 3, pp. 937–945, Jul. 2020. [Online]. Available: <https://ieeexplore.ieee.org/document/8644043/>
- [16] P. Risholm, T. Kirkhus, and J. T. Thielemann, "High-resolution structured light 3D sensor for autonomous underwater inspection," in *Proc. IEEE/MTS OCEANS Conf.*, Charleston, SC, USA, 2018, pp. 1–5. [Online]. Available: <https://ieeexplore.ieee.org/document/8604930/>
- [17] A. Sarafraz and B. K. Haus, "A structured light method for underwater surface reconstruction," *ISPRS J. Photogramm. Remote Sens.*, vol. 114, pp. 40–52, Apr. 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0924271616000290>
- [18] E. Hecht, Optics. 5th ed. London, U.K.: Pearson Education Limited, 2017. [Online]. Available: www.pearsoneditions.com
- [19] I. Söderkvist, "Using SVD for some fitting problems," no. 2, pp. 2–5, 2009. [Online]. Available: https://www.ltu.se/cms_fs/1.51590/svd-fitting.pdf
- [20] G. F. Franklin, J. D. Powell, M. L. Workman, and A. A. > Others, *Digital Control of Dynamic Systems*, vol. 3. Reading, MA, USA: Addison-Wesley, 1998.
- [21] V. Milanović, K. Castelino, and S. P. Ave, "Sub-100 μ s settling time and low voltage operation for gimbal-less two-axis scanners," in *IEEE/LEOS Opt. MEMS*, Takamatsu, Japan, vol. 15, no. 2, pp. 2003–2004, Aug. 2004. [Online]. Available: http://www.adriaticresearch.org/Research/pdf/VM_MOEMS04.pdf
- [22] V. Milanović, A. Kasturi, J. Yang, and F. Hu, "Closed-loop control of gimbal-less MEMS mirrors for increased bandwidth in LiDAR applications," in *Proc. Laser Radar Technol. Appl.*, 2017, Art. no. 101910N. [Online]. Available: www.mirrorleotech.com
- [23] OpenCV, "Tutorial: Camera calibration with OpenCV," Accessed: May 4, 2022. [Online]. Available: https://docs.opencv.org/4.x/d4/d94/tutorial_camera_calibration.html
- [24] J. Illingworth and J. Kittler, "The adaptive hough transform," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-9, no. 5, pp. 690–698, Sep. 1987.
- [25] S. Agarwal, K. Mierle, and The Ceres Solver Team, "Ceres solver," Mar. 2022. [Online]. Available: <https://github.com/ceres-solver/ceres-solver>

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

CASTILLÓN *et al.*: UNDERWATER 3D SCANNER TO COUNTERACT REFRACTION

9



Miguel Castellón received the B.Sc. degree in industrial engineering from the University of Zaragoza, Zaragoza, Spain, in 2015 and the M.Sc. degree in mechanical engineering from KU Leuven, Leuven, Belgium, in 2018. He is currently working toward the Ph.D. degree in robotics with the University of Girona, Girona, Spain.

His research interests include computer vision applied to robotics and autonomous navigation.



Josep Forest received the B.Sc. degree in industrial informatics from the University of Girona, Girona, Spain, in 1992, the M.Sc. degree in electronics engineering from the Autonomous University of Barcelona, Bellaterra, Spain, in 1998, and the Ph.D. degree in industrial engineering from the University of Girona, in 2004.

He is a former Co-Founder of AQSENSE (now part of COGNEX) and OPSIS Vision Technologies Spin-Off companies. His research interests are focused on 3D-machine vision, including laser triangulation, calibration, detection and point cloud processing. His research also includes the usability of 3-D applied to the industry for dimensional testing and quality control applications.



Pere Ridao (Member, IEEE) received the Ph.D. degree in industrial engineering from the University of Girona, Girona, Spain, in 2001.

He was the Chair of the IFAC's Technical Committee on Marine Systems. He is currently the Director of the Computer Vision and Robotics Research Institute (VICOROB) and the head of the Underwater Robotics Research Center (CIRS), and an Associate Professor with the Department of Computer Engineering, University of Girona. He is the Co-Author of four licenses and one Spanish/European patent, being co-founder of Iqua Robotics S.L. spin-off company. Since 1997, he has participated in 24 research projects (15 European and 9 National), he is also the author of more than 100 publications, and he has directed nine Ph.D. thesis (4 more under direction) and 14 M.S. thesis. His research activity focuses on designing and developing autonomous underwater vehicles for 3-D mapping and intervention.

5

LINEWISE NON-RIGID POINT CLOUD REGISTRATION

IN this chapter we present a novel non-rigid point cloud registration method that can successfully correct the distortion present in some experimental scans. It exploits the in-line rigidity of the point clouds acquired by LLSs to reduce the computational complexity while improving accuracy when compared to state-of-the-art non-rigid registration algorithms. We validated our method on synthetic data and on experimental scans gathered by our prototype LLS (see Chapter 4) mounted on the Girona1000 AUV [15]. This publication was one of the principal outcomes of the author’s research stay at ASL, ETH Zurich (Switzerland). The contents of this paper were also selected for presentation at the *2022 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2022)*, held in Kyoto (Japan) in October 2022.

Title: Linewise Non-Rigid Point Cloud Registration
Authors: **Miguel Castellón**, Pere Ridao, Roland Siegwart, and César Cadena
Journal: IEEE Robotics and Automation Letters
Volume: 7, Number: 3, Pages: 7044–7051, Published: 2022
DOI: [10.1109/LRA.2022.3180038](https://doi.org/10.1109/LRA.2022.3180038)
Quality index: JCR2021 Robotics, Impact Factor: 4.321, Q2 (11/30)

Linewise Non-Rigid Point Cloud Registration

Miguel Castellón ¹, Pere Ridao ¹, *Member, IEEE*, Roland Siegwart ², *Fellow, IEEE*,
and César Cadena ², *Member, IEEE*

Abstract—Robots are usually equipped with 3D range sensors such as laser line scanners (LLSs) or lidars. These sensors acquire a full 3D scan in a line by line manner while the robot is in motion. All the lines can be referred to a common coordinate frame using data from inertial sensors. However, errors from noisy inertial measurements and inaccuracies in the extrinsic parameters between the scanner and the robot frame are also projected onto the shared frame. This causes a deformation in the final scan containing all the lines, which is known as *motion distortion*. Rigid point cloud registration with methods like ICP is therefore not well suited for such distorted scans. In this paper we present a non-rigid registration method that finds the rigid transformation to be applied to *each line* in the scan in order to match an existing model. We fully leverage the continuous and relatively smooth robot motion with respect to the scanning time to formulate our method reducing the computational complexity while improving accuracy. We use synthetic and real data to benchmark our method against a state-of-the-art non-rigid registration method. Finally, the source code for the algorithm is made publicly available.¹

Index Terms—Autonomous vehicle navigation, non-rigid registration, point set registration, range sensing.

I. INTRODUCTION

POINT set registration is typically used to process range data coming from sensors mounted on robotic platforms, such as laser line scanners (LLSs) or lidars. These sensors do not acquire all the 3D points at once, but rather line by line. All the lines can be referred to a common coordinate frame using data from inertial sensors. However, errors from noisy inertial measurements and inaccuracies in the extrinsic calibration are also projected onto the shared frame. This causes a deformation in the final scan containing all the lines, which is known as *motion distortion*. This distortion varies smoothly through the scan lines: each scan line tends to have suffered a similar deformation to its neighbouring lines (see Fig. 1(c)). Motion distortion is

Manuscript received February 23, 2022; accepted May 27, 2022. Date of publication June 3, 2022; date of current version June 13, 2022. This letter was recommended for publication by Associate Editor A. Nuechter and Editor S. Behnke upon evaluation of the reviewers' comments. This work was supported in part by the European Commission through ATLANTIS Project under Grant H2020-ICT-2019-2-871571, in part by the Spanish Ministry of Science through Projects PER2IAUV and OPTIHTROV under Grants PID2020-115332RB-C32 and PDC2021-120791-C21, in part by the Doctoral Grant of the University of Girona under Grant IFUDG2019, and in part by the mobility Grant IF-MobUdG2019. (*Corresponding author: Miguel Castellón.*)

Miguel Castellón and Pere Ridao are with the Computer Vision and Robotics Research Institute (VICOROB), University of Girona, 17003 Girona, Spain (e-mail: miguel.castillon@udg.edu; pere@eia.udg.edu).

Roland Siegwart and César Cadena are with the Autonomous Systems Lab (ASL), ETH Zurich, 8092 Zurich, Switzerland (e-mail: rsiegwart@ethz.ch; cesarc@ethz.ch).

Digital Object Identifier 10.1109/LRA.2022.3180038

¹The source code of the algorithm can be found in <https://github.com/miguelcastillon/lnrr>

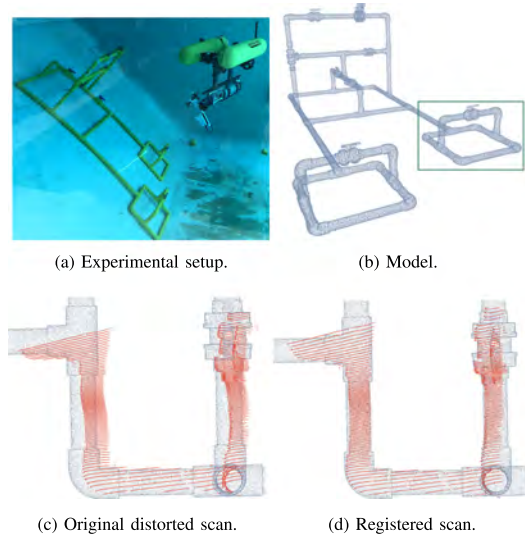


Fig. 1. Linewise non-rigid registration of real scans. The experiments (a) were carried out with our underwater LLS [1], [2] mounted on the AUV Girona1000 [3]. The 3D CAD model of the structure (b) was used as reference model. Errors in navigation data result in scans affected by motion distortion (c). Our method successfully finds the set of transformations to be applied to the scan in order to fit the model (d).

especially relevant for sensors with low refresh frequency or for robots with relatively high dynamics. In our case, we use an underwater LLS [1], [2] with a typical scan density of 50 lines per scan running at 1 Hz (50 lines per second). Since it only takes around 20 ms to acquire all the points in each line, we can safely assume in-line rigidity.

In order to undistort the scan, we can model the spatial transformation \mathcal{T} that needs to be applied to the scan point cloud \mathbf{Y} as a set of rigid transformations, one for each scan line: $\mathcal{T} = \{\mathcal{T}_1 \dots \mathcal{T}_L\}$ for a scan made up of L lines. Thus, a 3D point $\mathbf{y} \in \mathbb{R}^3$ in the l -th line of \mathbf{Y} will be transformed by the rigid transform $\mathcal{T}_l \in \mathbb{SE}(3)$, with rotation $\mathbf{R}_l \in \mathbb{SO}(3)$ and translation $\mathbf{t}_l \in \mathbb{R}^3$, as $\mathcal{T}_l(\mathbf{y}) = \mathbf{R}_l \mathbf{y} + \mathbf{t}_l$.

In this letter, we present a novel non-rigid registration algorithm that fully leverages a priori knowledge on range sensors to allow smooth inter-line deformation while keeping in-line rigidity. This notably limits the space of feasible solutions and results in improved accuracy and reduced computational complexity when compared to a state-of-the-art non-rigid registration method. Furthermore, we offer an open-source implementation of the method.

Contextualization: A great effort has been done in mobile robotics in the last decade to obtain accurate and consistent 3D data from range sensors mounted on moving platforms [4]–[6]. These works typically parameterize the robot trajectory using a continuous representation, such as B-splines, rather than discrete poses, which they aim to optimize by fitting observations from the range sensor. This continuous-time SLAM formulation presents the advantage of allowing an easy fusion of inertial and range sensors, even if they are not perfectly synchronized, while minimizing the motion distortion present in the scans.

However, these methods are based on the implicit assumption that there is enough overlap between consecutive scans. This assumption is met by terrestrial lidars but not by underwater LLSs. State-of-the-art underwater LLSs typically have a maximum range of around 3 to 5 meters and a field of view (FoV) of around $40^\circ \times 40^\circ$ in realistic visibility conditions, due to the high attenuation rate of visible light when travelling through water [7]. Furthermore, autonomous underwater vehicles (AUVs) need to keep a safe distance of at least 1 m during real operations in presence of water currents to avoid crashing into the scanned structure. As a consequence of these factors, it is very common that a relatively high percentage of scans fail at capturing enough informative points and can effectively be considered as empty point clouds. Nonetheless, in the context of missions performed by AUVs such as inspection of an industrial underwater structure, the CAD model of the target to be scanned is usually available. For all these reasons, our method considers the non-rigid registration of one scan at a time against the point cloud sampled from the model of the structure. The result of our method is an undistorted point cloud, which may later be used in downstream applications, such as structural damage assessment. Please note that this approach is motivated by our application but its applicability is not limited to the underwater inspection case, since very few assumptions have been made during the design of the algorithm.

II. RELATED WORK

Many point set registration approaches have been proposed in the literature. They can be classified according to how they i) find correspondences between point clouds and ii) model the transformation \mathcal{T} . This section provides a brief overview of point set registration algorithms that are closely related to ours: more precisely, only feature-less methods are considered.

Regarding the choice of correspondences, iterative closest point (ICP) [8], [9] follows the very simple yet effective approach of assigning the closest point at each iteration. This is known as *hard-assignment*. Thanks to its simplicity and low computational complexity, ICP is probably the most popular registration method. One of the factors that severely limits the performance of ICP is the existence of noise in the data points. Probabilistic methods overcome this limitation by *soft-assigning* a correspondence probability between each pair of points. The first method to use soft assignment in point set registration was robust point matching (RPM) [10]. The alternating soft-assignment of correspondences and transformation in RPM is an equivalent to the expectation maximization (EM) algorithm [11] for a Gaussian mixture model (GMM), if we consider one of the sets as GMM centroids and the other as data points [12]. As a matter of fact, point set registration is modelled as a GMM likelihood maximization problem in several methods [13]–[15].

These methods normally add an extra distribution to the GMM in order to account for outliers.

Regarding the choice of the transformation \mathcal{T} , registration methods are typically classified into rigid and non-rigid. Rigid methods (like ICP) assume that a rigid transformation is enough to relate both point sets, whereas non-rigid methods allow deformation. Non-rigid registration is naturally more challenging due to the increased number of degrees of freedoms (DoFs). Therefore, it is common for non-rigid techniques to include a regularization term in their cost functions to avoid overfitting. Non-rigid methods usually parameterize the transformation with local displacement fields, which typically use either thin plate spline (TPS) [16] or Gaussian kernels. An example of the former is [13], which combined them with RPM resulting in TPS-RPM. Another example is the correlation-based approach proposed in [17] and extended in [14], which models both point sets as GMMs and estimates the TPS parameters by minimizing the L_2 norm between the distributions. One of the arguably most popular non-rigid methods, coherent point drift (CPD) [18], uses a Gaussian kernel to define the radius of the local displacement field. CPD is related to motion coherence theory [19], [20], which imposes the assumption that neighbouring points tend to move similarly. Among its strengths, its Gaussian kernel provides a free parameter to control the locality of deformations. Moreover, the algorithm is designed to estimate the Gaussian width within the minimization framework instead of using deterministic annealing, which results in reportedly shorter runtimes and better performance [18]. When compared to the correlation-based method in [14], CPD effectively minimizes the KL divergence between two distributions rather than the L_2 norm, which yields better results because it weighs the error according to its probability [18]. Several variants of CPD have been presented which focus on improving different aspects of the algorithm. For instance, [21] improves registration quality and robustness in applications where correspondence priors are available by integrating them in closed form, [22] uses efficient Gaussian filtering methods to achieve substantially faster computational performance while maintaining robustness, and [23], [24] propose a Bayesian approach that accelerates the registration process and guarantees convergence.

Our method draws inspiration from CPD in modelling the problem as a GMM and in parameterizing the regularization term as a local displacement field ruled by a Gaussian kernel. However, it differs from it in the parameterization of the transformations \mathcal{T} : we exploit the a priori knowledge of the working principles of LLSs to reduce the dimensionality of the non-rigid displacement field by imposing in-line rigidity.

Within the family of non-rigid methods that characterize the deformation as a displacement field, CPD is the most common method to compare against. Like ICP for rigid methods, CPD is a simple and general method that can be set as a fair baseline for comparison. Moreover, CPD has an open-source implementation that can be readily used. For these reasons, we selected it as the non-rigid method against which we benchmark our algorithm in Section IV.

III. METHOD

The goal of our method is to register two 3D point sets. The model \mathbf{X} is composed of N points, $\mathbf{X} = [\mathbf{x}_1 \dots \mathbf{x}_n \dots \mathbf{x}_N]^T$, and the deformable point set \mathbf{Y} contains M points, $\mathbf{Y} = [\mathbf{y}_1 \dots \mathbf{y}_m \dots \mathbf{y}_M]^T$. In fact, \mathbf{Y} is assumed to be made up of L lines, where line l has P_l points, with $\sum_l P_l = M$. The variables

TABLE I
VARIABLES IN THE PROPOSED METHOD

Name	Meaning	Size
\mathbf{G}	Gaussian kernel	$L \times L$
\mathbf{g}_l	l -th row in \mathbf{G}	$1 \times L$
\mathbf{J}_R^T	Jacobian of the rotation matrix	$3L \times 9L$
L	Number of lines in the scan \mathbf{Y}	1×1
M	Number of points in the scan \mathbf{Y}	1×1
N	Number of points in the model \mathbf{X}	1×1
P_l	Number of points in line l of the scan \mathbf{Y}	1×1
p_{mn}	Correspondence probability between \mathbf{x}_n and \mathbf{y}_m	1×1
\mathbf{P}	Probability matrix made up of elements p_{mn}	$M \times N$
\mathbf{R}_l	Rotation applied to line l in the scan \mathbf{Y}	3×3
\mathbf{R}	Matrix of rotation matrices \mathbf{R}_l	$3 \times 3L$
\mathbf{t}_l	Translation applied to line l in the scan \mathbf{Y}	1×3
\mathbf{t}	Matrix of translation vectors \mathbf{t}_l	$L \times 3$
\mathbf{X}	Model point set	$N \times 3$
\mathbf{x}_n	n -th point (row) in the model \mathbf{X}	1×3
\mathbf{Y}	Scan point set	$M \times 3$
\mathbf{y}_m	m -th point (row) in the scan \mathbf{Y}	1×3
β	Gaussian width	1×1
λ	Regularization parameter	1×1
w	Weight of the outlier distribution	1×1

TABLE II
SIZES OF THE AUXILIARY MATRICES

Name	Size	Name	Size	Name	Size
\mathbf{A}	$L \times 3$	\mathbf{F}	$M \times L$	\mathbf{W}	$L \times 6$
\mathbf{B}	$L \times 3L$	\mathbf{S}	$3L \times 3L$	\mathbf{T}	$3 \times 3L$
\mathbf{C}	$M \times 3$	\mathbf{U}	$L \times 3$	\mathbf{Y}_D	$M \times 3L$
\mathbf{D}	$M \times 3L$	\mathbf{V}	$L \times 3$	\mathbf{Z}	$L \times L$

defined in this section are gathered in Tables I and II. Please note that $M \gg L$.

The points in \mathbf{Y} represent the centroids of a GMM, and the points \mathbf{X} represent data points generated by the GMM. In order to account for outliers, we add an extra uniform distribution to the model. The probability density function of the resulting mixture is:

$$p(\mathbf{x}_n) = w \frac{1}{N} + (1-w) \sum_{m=1}^M P(m) p(\mathbf{x}_n|m), \quad (1)$$

with w , $0 \leq w \leq 1$ weighting the outliers contribution. We assign equal weights to all the components in the GMM: $P(m) = \frac{1}{M}$. All the normal distributions in the GMM ($p(\mathbf{x}_n|m) = \mathcal{N}(\mathbf{x}_n|\mathcal{T}(\mathbf{y}_m, \xi), \Sigma_m)$) are assumed to have equal, isotropic variances σ^2 :

$$p(\mathbf{x}_n|m) = \frac{1}{(2\pi\sigma^2)^{3/2}} \exp\left(-\frac{\|\mathbf{x}_n - \mathcal{T}(\mathbf{y}_m, \xi)\|^2}{2\sigma^2}\right), \quad (2)$$

where ξ are the updated centroids of the GMM. Given all the observations \mathbf{X} , the combined negative log-likelihood to minimize is:

$$E(\xi, \sigma^2) = -\sum_{n=1}^N \log \sum_{m=1}^M P(m) p(\mathbf{x}_n|m) \quad (3)$$

Due to the difficulty in directly minimizing E , we use the expectation maximization (EM) technique instead. This method consists on iterating over two steps. First, in the E-step we fix the parameters (ξ, σ^2) and compute the probability of every possible correspondence. Then, in the M-step we fix these correspondence probabilities and minimize a cost function. In the E-step, we compute the probability that data point \mathbf{x}_n was

generated by the component m as its posterior probability:

$$p_{mn} = P(m|\mathbf{x}_n) = \frac{\exp\left(-\frac{\|\mathbf{x}_n - \mathcal{T}(\mathbf{y}_m, \xi)\|^2}{2\sigma^2}\right)}{\sum_{m'=1}^M \exp\left(-\frac{\|\mathbf{x}_n - \mathcal{T}(\mathbf{y}_{m'}, \xi)\|^2}{2\sigma^2}\right) + c}, \quad (4)$$

where $c = \frac{w}{1-w} \frac{M(2\pi\sigma^2)^{3/2}}{N}$. Then, for the M-step we define the expectation of the complete negative log-likelihood function E [18]:

$$Q(\xi, \sigma^2) = \frac{1}{2\sigma^2} \sum_n \sum_m p_{mn} \|\mathbf{x}_n - \mathcal{T}(\mathbf{y}_m, \xi)\|^2 + \frac{3}{2} N_P \log \sigma^2, \quad (5)$$

where $N_P = \sum_n \sum_m p_{mn}$. Jensen's inequality states that Q is an upper bound for E [25], so minimizing Q also minimizes E . Up to this point, the derivation is the same as in [18]. However, our method proposes to define \mathcal{T} as a set of L rigid transformations. We still need to introduce a regularization term in the objective function in order to apply the a priori knowledge that transformations of neighbouring lines tend to be similar. We can define this field as $\mathbf{p} : \mathbb{N} \mapsto \mathbb{SE}(3)$, which relates each line index with a 6-DoF rigid transformation.² In order to transform centroid \mathbf{y} , \mathbf{p} is first evaluated at the line number of \mathbf{y} , l :

$$\mathcal{T}(\mathbf{y}) = \mathbf{p}(l) \oplus \mathbf{y}, \quad (6)$$

where the operator \oplus denotes 3D composition [26]. Then, we can introduce the regularization term $\phi(\mathbf{p})$ weighted with a regularization parameter $\lambda > 0$:

$$Q(\xi, \sigma^2) = \frac{1}{2\sigma^2} \sum_n \sum_m p_{mn} \|\mathbf{x}_n - \mathbf{p}(l) \oplus \mathbf{y}\|^2 + \frac{3}{2} N_P \log \sigma^2 + \frac{\lambda}{2} \phi(\mathbf{p}) \quad (7)$$

The term $\phi(\mathbf{p})$ is introduced to ensure *smoothness* in the field of rigid transformations. In our case, smoothness refers to a measure of the oscillatory nature of the field \mathbf{p} . In the frequency domain, the field \mathbf{p} can be said smooth if it has most of its energy at low frequency (small bandwidth). Therefore, we can define $\phi(\mathbf{p})$ as a measure of the remaining energy in \mathbf{p} after applying a high-pass filter to it [27]:

$$\phi(\mathbf{p}) = \int \frac{|\tilde{\mathbf{p}}(\mathbf{s})|}{\tilde{G}(\mathbf{s})} d\mathbf{s}, \quad (8)$$

where the symbol $\tilde{\cdot}$ indicates the Fourier transform, and \tilde{G} is some positive function that tends to zero as $\|\mathbf{s}\| \rightarrow \infty$ (so that $\frac{1}{\tilde{G}}$ is a high-pass filter), where \mathbf{s} belongs to the complex frequency domain resulting from the Laplace transform. The field \mathbf{p} that minimizes this energy in (8) has the form [28]:

$$\mathbf{p}(l) = \sum_{k=1}^L \mathbf{w}_k G(l, k) + \psi(l), \quad (9)$$

where $\psi(l)$ is a term in the nullspace of ϕ . From the theory of reproducing kernels [29] it is known that if G is a positive definite function, then (8) is a norm in its corresponding reproducing kernel Hilbert space (RKHS). The key intuition behind this

²Please note that this field is different from the field $\mathbf{v} : \mathbb{R}^3 \mapsto \mathbb{R}^3$ proposed in [18].

definition is that this squared norm of function ϕ can be thought of as a generalization to functions of the quadratic form of an n -vector [30]. Therefore, its nullspace contains only the zero element: $\psi(l) = 0$. In order to meet this requirement, we choose the radially symmetric Gaussian function:

$$G(l, k) = \exp\left(-\frac{(l-k)^2}{2\beta^2}\right), \quad (10)$$

where the Gaussian width β controls the *locality* of the area in which smoothness is applied. As explained in [18], this choice of kernel agrees with motion coherence theory [19]. There, the authors chose it because it has second order derivatives and it generates analytic solutions.

If we stack w_k and $G(l, k)$ in (9) into matrices \mathbf{W} and \mathbf{G} , respectively, we obtain: $\mathbf{p}(l) = \mathbf{g}_l \mathbf{W}$ and $\phi(\mathbf{p}) = \|\mathbf{p}\|_{\mathbf{G}}^2 = \text{tr}(\mathbf{W}^T \mathbf{G} \mathbf{W})$, where \mathbf{g}_l is the l -th row of \mathbf{G} . Substituting in the cost function:

$$Q = \frac{1}{2\sigma^2} \sum_n \sum_m p_{mn} \|\mathbf{x}_n - \mathbf{g}_l \mathbf{W} \oplus \mathbf{y}_m\|^2 + \frac{3}{2} N_P \log \sigma^2 + \frac{\lambda}{2} \text{tr}(\mathbf{W}^T \mathbf{G} \mathbf{W}) \quad (11)$$

Please note that index l changes according to the line in the scan to which the sum index m belongs. Now, our goal is to find the weights \mathbf{W} that minimize Q . In order to ease our task, we separate the rotational from the translational components: \mathbf{U} and \mathbf{V} , respectively. \mathbf{U} are the first three columns of \mathbf{W} , and \mathbf{V} the last three. We can then rewrite the distance function in (11):

$$\mathbf{g}_l \mathbf{W} \oplus \mathbf{y}_m = \mathbf{y}_m \text{Rot}(\mathbf{g}_l \mathbf{U})^T + \mathbf{g}_l \mathbf{V} \quad (12)$$

Please note that the rotation matrix is transposed because \mathbf{y}_m is a row vector. Rotations are parameterized using the Euler angles roll around x , pitch around y , and yaw around z , in ZYX order. Moreover, we can also use the properties of the trace to rewrite the regularization term in (11), $\text{tr}(\mathbf{W}^T \mathbf{G} \mathbf{W}) = \text{tr}(\mathbf{U}^T \mathbf{G} \mathbf{U}) + \text{tr}(\mathbf{V}^T \mathbf{G} \mathbf{V})$, and we can define $\mathbf{t}_l = \mathbf{g}_l \mathbf{V}$ and $\mathbf{R}_l = \text{Rot}(\mathbf{g}_l \mathbf{U})$, so that the final cost function can be rewritten as:

$$Q = \frac{1}{2\sigma^2} \sum_n \sum_m p_{mn} \|\mathbf{x}_n - (\mathbf{y}_m \mathbf{R}_l^T + \mathbf{t}_l)\|^2 + \frac{3}{2} N_P \log \sigma^2 + \frac{\lambda}{2} \text{tr}(\mathbf{U}^T \mathbf{G} \mathbf{U}) + \frac{\lambda}{2} \text{tr}(\mathbf{V}^T \mathbf{G} \mathbf{V}) \quad (13)$$

subject to $\mathbf{R}_l^T \mathbf{R}_l = \mathbf{I}$ and $\det(\mathbf{R}_l) = 1$, $\forall l \leq L$. Please note that these two constraints are naturally observed in the construction of each rotation matrix \mathbf{R}_l out of Euler angles. In the remaining of the section we show how to compute the weights \mathbf{U} and \mathbf{V} that minimize the cost function in (13).

A. Solving for \mathbf{V}

First, we solve for \mathbf{V} by setting the corresponding partial derivative of the cost function in (13) to zero. The resulting expression is:

$$(\sigma^2 \lambda \mathbf{I}_{L \times L} + \mathbf{F}^T \text{diag}(\mathbf{P} \mathbf{1}_N) \mathbf{F} \mathbf{G}) \mathbf{V} = \mathbf{F}^T \mathbf{P} \mathbf{X} - \mathbf{F}^T \text{diag}(\mathbf{P} \mathbf{1}_N) \mathbf{Y}_D \mathbf{R}^T, \quad (14)$$

where \mathbf{P} is the probability matrix made up of elements p_{mn} , $\mathbf{I}_{L \times L}$ is the $L \times L$ identity matrix, $\mathbf{1}_N$ is an N -vector of all

ones, $\mathbf{R} = [\mathbf{R}_1 \cdots \mathbf{R}_l \cdots \mathbf{R}_L]$, and

$$\mathbf{F} = [\text{diag}(\mathbf{1}_{P_1}, \dots, \mathbf{1}_{P_L})], \quad (15)$$

$$\mathbf{Y}_D = \left[\text{diag} \left(\frac{\mathbf{y}_1}{\mathbf{y}_{P_1}}, \frac{\mathbf{y}_{P_1+1}}{\mathbf{y}_{P_1+P_2}}, \dots \right) \right]. \quad (16)$$

The optimal \mathbf{V} ((14)) can be rewritten as:

$$\mathbf{V} = \mathbf{A} - \mathbf{B} \mathbf{R}^T, \quad (17)$$

with $\mathbf{A} = \mathbf{Z} \mathbf{F}^T \mathbf{P} \mathbf{X}$, $\mathbf{B} = \mathbf{Z} \mathbf{F}^T \text{diag}(\mathbf{P} \mathbf{1}_N) \mathbf{Y}_D$ and $\mathbf{Z} = (\sigma^2 \lambda \mathbf{I}_{L \times L} + \mathbf{F}^T \text{diag}(\mathbf{P} \mathbf{1}_N) \mathbf{F} \mathbf{G})^{-1}$.

B. Solving for \mathbf{U}

The cost function ((13)) can be written in matrix form:

$$Q = \frac{1}{2\sigma^2} \left[\text{tr}(\mathbf{X}^T \text{diag}(\mathbf{P}^T \mathbf{1}_M) \mathbf{X}) - 2 \text{tr}(\mathbf{X}^T \mathbf{P}^T \mathbf{Y}_R) + \text{tr}((\mathbf{Y}_R)^T \text{diag}(\mathbf{P} \mathbf{1}_N) \mathbf{Y}_R) \right] + \frac{3}{2} N_P \log \sigma^2 + \frac{\lambda}{2} \text{tr}(\mathbf{U}^T \mathbf{G} \mathbf{U}) + \frac{\lambda}{2} \text{tr}(\mathbf{V}^T \mathbf{G} \mathbf{V}), \quad (18)$$

where \mathbf{Y}_R is the registered point set: $\mathbf{Y}_R = \mathcal{T}(\mathbf{Y}) = \mathbf{Y}_D \mathbf{R}^T + \mathbf{F} \mathbf{t}$, with $\mathbf{t} = [\mathbf{t}_1 \cdots \mathbf{t}_l \cdots \mathbf{t}_L]^T$. In fact, \mathbf{Y}_R can be expressed as a function of \mathbf{R} :

$$\mathbf{Y}_R = \mathbf{C} + \mathbf{D} \mathbf{R}^T, \quad (19)$$

where $\mathbf{C} = \mathbf{F} \mathbf{G} \mathbf{A}$ and $\mathbf{D} = \mathbf{Y}_D - \mathbf{F} \mathbf{G} \mathbf{B}$. Using (17) and (19) and dropping out the terms that do not depend on \mathbf{U} , (18) can be rewritten as:

$$Q = \frac{1}{2} \text{tr}(\mathbf{R} \mathbf{S} \mathbf{R}^T) + \text{tr}(\mathbf{T} \mathbf{R}^T) + \frac{\lambda}{2} \text{tr}(\mathbf{U}^T \mathbf{G} \mathbf{U}), \quad (20)$$

where $\mathbf{S} = \lambda \mathbf{B}^T \mathbf{G} \mathbf{B} + \frac{1}{\sigma^2} \mathbf{D}^T \text{diag}(\mathbf{P} \mathbf{1}_N) \mathbf{D}$ and $\mathbf{T} = -\lambda \mathbf{A}^T \mathbf{G} \mathbf{B} + \frac{1}{\sigma^2} (\text{diag}(\mathbf{P} \mathbf{1}_N) \mathbf{C} - \mathbf{P} \mathbf{X})^T \mathbf{D}$. Please note that \mathbf{S} is symmetric. The optimal \mathbf{U} that minimizes the cost function is found by setting the corresponding partial derivative to zero:

$$\frac{\partial Q}{\partial \mathbf{U}} = \mathbf{J}_R^T \text{vec}(\mathbf{R} \mathbf{S} + \mathbf{T}) + \text{vec}(\lambda \mathbf{G} \mathbf{U}) = \mathbf{0} \quad (21)$$

The operator $\text{vec}(\cdot)$ flattens the matrix in column-major order. The transpose of the jacobian \mathbf{J}_R is a sparse matrix computed as:

$$\mathbf{J}_R^T = \begin{bmatrix} \text{diag} \left(\text{vec} \left(\frac{\partial \mathbf{R}_1}{\partial \phi} \right), \dots, \text{vec} \left(\frac{\partial \mathbf{R}_L}{\partial \phi} \right) \right) \\ \text{diag} \left(\text{vec} \left(\frac{\partial \mathbf{R}_1}{\partial \theta} \right), \dots, \text{vec} \left(\frac{\partial \mathbf{R}_L}{\partial \theta} \right) \right) \\ \text{diag} \left(\text{vec} \left(\frac{\partial \mathbf{R}_1}{\partial \psi} \right), \dots, \text{vec} \left(\frac{\partial \mathbf{R}_L}{\partial \psi} \right) \right) \end{bmatrix} \quad (22)$$

The elements in \mathbf{J}_R are the jacobians of \mathbf{R} with respect to roll, pitch and yaw, respectively. Finally, (21) is solved numerically.

The optimal σ^2 is found by setting the corresponding partial derivative of (18) equal to 0:

$$\sigma^2 = \frac{1}{3N_P} \left[\text{tr}(\mathbf{X}^T \text{diag}(\mathbf{P}^T \mathbf{1}_M) \mathbf{X}) - 2 \text{tr}(\mathbf{X}^T \mathbf{P}^T \mathbf{Y}_R) + \text{tr}((\mathbf{Y}_R)^T \text{diag}(\mathbf{P} \mathbf{1}_N) \mathbf{Y}_R) \right] \quad (23)$$

Algorithm 1: Linewise Non-Rigid Registration.

-
- 1: Inputs: $\mathbf{X}, \mathbf{Y}, \lambda > 0, \beta > 0, 0 \leq w \leq 1$
 - 2: Initialize: $\mathbf{Y}_R = \mathbf{Y}, \sigma^2 = \frac{1}{3NM} \sum_n \sum_m \|\mathbf{x}_n - \mathbf{y}_m\|^2$
 - 3: Build $\mathbf{F}, \mathbf{Y}_D, \mathbf{G}$: $g_{ij} = \exp(-\frac{(i-j)^2}{2\beta^2})$
 - 4: **while** not converged **do**
 // E-Step:
 5: $p_{mn} = \frac{\exp(-\frac{\|\mathbf{x}_n - \mathbf{y}_m\|^2}{2\sigma^2})}{\frac{w}{1-w} \frac{M(2\pi\sigma^2)^{3/2}}{N} + \sum_{m'=1}^M \exp(-\frac{\|\mathbf{x}_n - \mathbf{y}_{m'}\|^2}{2\sigma^2})}$
 // M-Step:
 6: $\mathbf{A} = \mathbf{Z}\mathbf{F}^T\mathbf{P}\mathbf{X}, \mathbf{B} = \mathbf{Z}\mathbf{F}^T \text{diag}(\mathbf{P}\mathbf{1}_N)\mathbf{Y}_D$
 7: $\mathbf{C} = \mathbf{F}\mathbf{G}\mathbf{A}, \mathbf{D} = \mathbf{Y}_D - \mathbf{F}\mathbf{G}\mathbf{B}$
 8: $\mathbf{S} = \lambda\mathbf{B}^T\mathbf{G}\mathbf{B} + \frac{1}{\sigma^2}\mathbf{D}^T \text{diag}(\mathbf{P}\mathbf{1}_N)\mathbf{D}$
 9: $\mathbf{T} = -\lambda\mathbf{A}^T\mathbf{G}\mathbf{B} + \frac{1}{\sigma^2}(\text{diag}(\mathbf{P}\mathbf{1}_N)\mathbf{C} - \mathbf{P}\mathbf{X})^T\mathbf{D}$
 10: Solve \mathbf{U} from $\mathbf{J}_R^T \cdot \text{vec}(\mathbf{R}\mathbf{S} + \mathbf{T}) + \text{vec}(\lambda\mathbf{G}\mathbf{U}) = \mathbf{0}$
 11: Build \mathbf{R}
 12: $\mathbf{Y}_R = \mathbf{C} + \mathbf{D}\mathbf{R}^T$
 13: $\sigma^2 = \frac{1}{3NF} [\text{tr}(\mathbf{X}^T \text{diag}(\mathbf{P}^T\mathbf{1}_M)\mathbf{X}) - 2\text{tr}(\mathbf{X}^T\mathbf{P}^T\mathbf{Y}_R) + \text{tr}((\mathbf{Y}_R)^T \text{diag}(\mathbf{P}\mathbf{1}_N)\mathbf{Y}_R)]$
 14: **end while**
 15: Return: $\mathbf{Y}_R, \mathcal{T}$
-

The method presented in this section is summarized in Algorithm 1. In comparison with CPD, both methods share the same E-step (building the \mathbf{P} matrix). However, the different parameterization of the transformation \mathcal{T} applied to the moving point set \mathbf{Y} makes the M-step different. It is worth noting that our method does not optimize all the parameters ($\mathbf{U}, \mathbf{V}, \sigma^2$) at once but rather iteratively by means of partial derivatives. Such iterations decrease the Q function but not to an exact minimum, which is known as the generalized EM algorithm [11].

We followed the proposal in [18] to alleviate the computational burden of the E-step by using fast Gauss transform (FGT) [31]³ to evaluate the Gaussian kernel and to compute the matrix-vector products $\mathbf{P}\mathbf{X}$, $\mathbf{P}\mathbf{1}_N$, and $\mathbf{P}^T\mathbf{1}_M$.

IV. RESULTS

This section presents the registration results of applying our method to synthetic and real scans (Section IV-A and IV-B, respectively). The tests were conducted using an Intel i7-9750H CPU at 2.60 GHz with 16 GB of RAM. The C++ implementation exploited parallelization on the 12 cores whenever possible. Results of the proposed method were benchmarked against a rigid and a non-rigid method. The chosen rigid method was generalized ICP (G-ICP) [32]⁴. The chosen non-rigid method was CPD⁵. Different performance aspects were quantitatively evaluated, such as registration accuracy, robustness to noise, robustness to different types and amplitudes of deformations, and runtimes. Throughout the experiments, all methods were given identical, reasonably good initial guesses. In all the boxplots presented in this section, the length of the whiskers represent the 1.5 interquartile range of the data.

As already explained in Section I, our method has been developed in the context of mobile underwater scanning. AUVs

³Implementation from <https://github.com/gadomski/fgt>

⁴Implemented in PCL https://pointclouds.org/documentation/classpcl_1_1_generalized_iterative_closest_point.html

⁵Implementation from <https://github.com/gadomski/cpd>

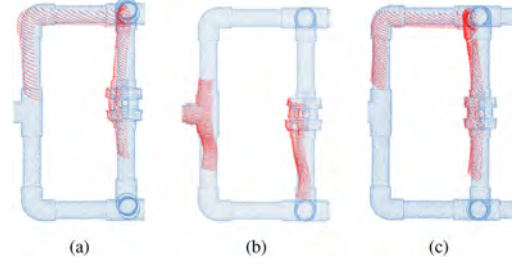


Fig. 2. Three different examples of real distorted scans (before registration). Each scan only covers a rather small area of the structure due to the limited FoV of the LLS. The distortions present a smooth distribution.

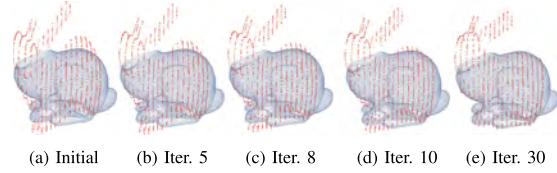


Fig. 3. Registration process of simulated vertical scan lines (red) with an incomplete version of the Stanford bunny model (blue) using the proposed method.

are usually equipped with modern inertial navigation systems (INSS), which are more accurate than inertial sensors mounted on terrestrial robots, especially for measuring orientation. For example, the INS mounted on Girona1000 has orientation errors smaller than 0.10° around all three axes⁶ thanks to its high-end fiber optic gyroscopes (FOGs). By examining the experimental scans (such as Fig. 2) and the navigation data, it can be concluded that the translational component of the distortion is predominant over the rotational one. The synthetic dataset used in Section IV-A was designed to be consistent with the experimental dataset used in Section IV-B but with higher deformations: we applied errors of tens of millimeters in translation and of up to 2° in rotation (which is 20x higher than expected in reality). The applicability of our method to realistic scenarios is proved by the low registration errors in Section IV-B (for a visual example, see the registration process from Fig. 1(c) to Fig. 1(d)). Please note that the fact that we test our method on datasets relevant to our use case does not shrink its applicability, since no assumption was made in the design of the algorithm other than in-line rigidity.

A. Synthetic Data

The proposed method was first validated using simulated data from two different datasets: the Stanford bunny and an underwater industrial site. The Stanford bunny was used to visually verify the performance of the proposed method on a well-known model. We simulated line scans and applied different, smoothly-changing rigid transformations to each line (see Fig. 3(a)), following a sinusoidal function. As shown through the iterations in the figure, the proposed method achieved good results (see Fig. 3(e)). Despite the upper part

⁶<https://www.ixblue.com/wp-content/uploads/2022/01/Phins%20Compact%20Series%20-%20Datasheet.pdf>

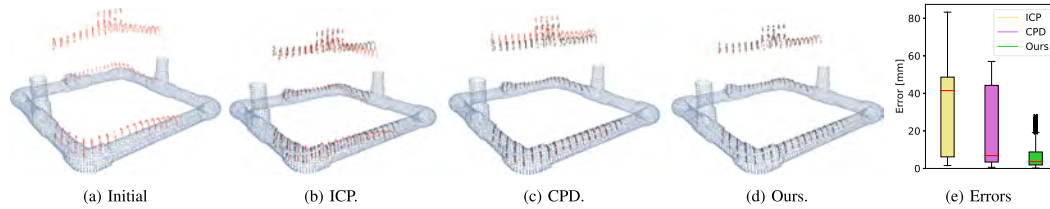


Fig. 4. Challenging example of a synthetically generated scan registered against an incomplete model of the underwater pipe structure. The upper pipe and valve are missing in the model but present in the scan. Figures (b)–(d) compare visually the registration results of ICP, CPD and our method. Black dots in these figures show the ground truth.

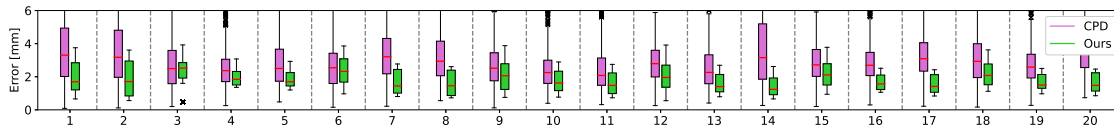


Fig. 5. Synthetic benchmark based on 20 different simulated scans of the underwater pipe structure.

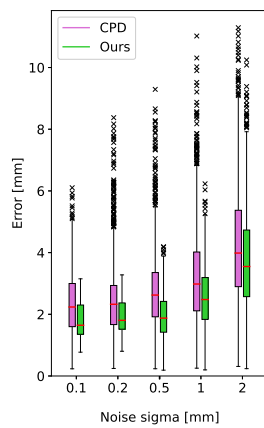


Fig. 6. Accuracy comparison for increasing levels of noise in the scan.

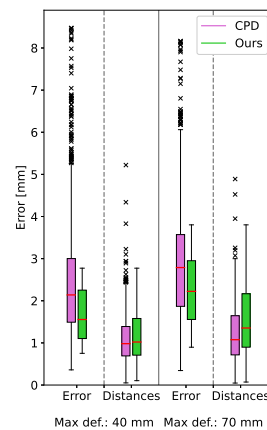


Fig. 7. Accuracy comparison for different deformation levels using two different metrics.

of the model being missing, the in-line rigidity enforced by our method pulls the points without correspondences towards their true position.

The presented method was also synthetically validated using a mock-up model of an underwater industrial site (see Fig. 1(b)). The dimensions of the whole structure are approximately $3500 \times 2100 \times 1300$ mm and it is made up of different types of pipes and valves. The diameter of the pipes is of 60 mm. The points for \mathbf{X} were sampled from the model using a voxel grid of 5 mm. This scenario was designed to be more realistic and therefore challenging than the Stanford bunny. First, as it can be seen in Fig. 4, the scan does not contain points that would be occluded from a given point of view of the scanner: in this case, the scanner is assumed to view the structure from above. Second, points in the scan and in the model are subsampled differently and therefore there are no direct correspondences. Unless otherwise stated, the metric used in all the boxplots of synthetic data is the 3D distance of each point to its ground truth after registration. We use this metric because the main goal of

our work is to minimize the distortion present in the scans. The red lines in the boxplots refer to the median.

A visual example of how the three different methods compare in this dataset can be found in Fig. 4. Note that in this case only an incomplete subset of the model was considered in order to test the robustness of the method: the upper pipe and its valve are present in the scan but missing in the model. A different rigid transformation was applied to each scan line. The distribution of these transformations followed a sinusoidal pattern, with a maximum distortions of 40 mm in translation and 2° in rotation. The boxplot in Fig. 4(e) shows how the proposed method achieves a more accurate registration than ICP and CPD. ICP naturally fails at registering the distorted scan because it incorrectly assumes that all the points in it can be treated as a rigid body (see Fig. 4(b)). CPD is able to accurately register the points that are close to the visible parts of the structure. However, it fails at registering the points corresponding to the missing upper pipe and valve (see Fig. 4(c)). The reason for it is that coherence in CPD applies to 3D distances between points,

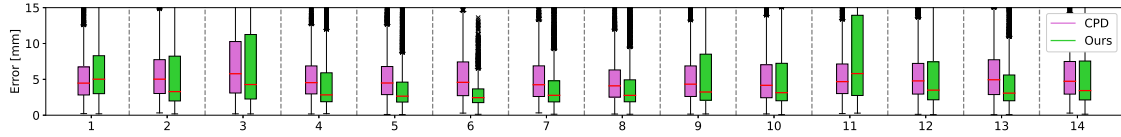


Fig. 8. Experimental benchmark based on 14 different scans of the underwater pipe structure.

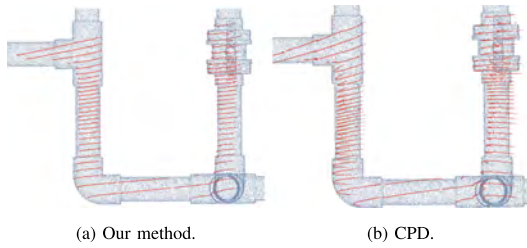


Fig. 9. Comparison of a registration result of a real scan using both methods (number 4 in Fig. 8). In this case, both methods achieve similar errors when using the distance-to-model. However, our method maintains the structure of the scan much better than CPD thanks to the in-line rigidity constraint.

irrespective of which line they correspond to. On the other hand, our method exploits the fact that all the points in each line share the same rigid transform, even if they are distant in 3D space. This makes that the final registration result is closer to the ground truth (see Fig. 4(e)).

1) *Benchmark Ideal Conditions*: Our method was benchmarked against CPD on 20 simulated scans of the underwater pipe structure (see Fig. 5). ICP was no longer considered because it consistently achieved much worse results than the non-rigid methods (see Fig. 4(e)). All the scans were noise-free and made up of 20 lines. The structure did not have any missing parts. The different scans featured different deformation shapes and directions (always smooth), and different number of points per line. The maximum deformation level was 40 mm in translation and 2° in rotation, following a sinusoidal pattern. It can be seen in Fig. 5 that the proposed method systematically outperforms CPD in this diverse set of scans. The parameter values chosen for our method were $(\beta, \lambda, w) = (60, 80, 0.1)$, and equivalent values for CPD. These values were chosen empirically on a subset of the simulated data and it was observed that slight variations of up to around 40% did not have a very noticeable impact on the results. However, an automatic procedure to estimate these parameters should be further studied, as proposed in Section V.

2) *Robustness to Noise*: The performance of our method in presence of noise was compared to CPD in Fig. 6. The noise was only applied to the 3D position of the points in scan and it followed a normal distribution centered around 0 with increasing values of standard deviation, as shown in the x axis. Our method consistently yielded lower errors than CPD for all levels of noise.

3) *Deformation Amplitude*: The performance of both methods on relatively large deformations was studied in Fig. 7. Considering for now only the columns with the label “Error” in the figure, it can be seen that our method can successfully register scans with deformations of 70 mm achieving errors in the order of a few mm (please note that the diameter of

the pipes is 60 mm). Our method once again outperforms CPD.

4) *Metric Comparison*: Until now we have only worked with synthetic data. The registration metric that we have used was the final error with respect to the ground truth. However, this ground truth is not available when working with experimental data. One of the most commonly used metrics in that case is the distance to the closest point in the model. This metric, however, tends to underestimate the registration error because it assigns low errors to points converging towards *any* point in the model, even if far from their ground truth. This effect is depicted in Fig. 7. For both levels of deformation, our method outperforms CPD when we consider the error with respect to the ground truth, as explained in Section IV-A3. If we use the distance-to-model metric instead, both methods are assigned lower errors, which can lead to false interpretations of the results.

5) *Runtimes*: A comparison of runtimes for different sizes of the scan and of the model can be found in Table III. It can be seen how an increasing number of points per line affects CPD much more than it does our method: a 5x increment in the number of points per line (from 50 to 250) results in a 40x increment in CPD runtime but only 3.5x increment for our method. Increasing the number of points in the model has little effect in both of them, mainly thanks to the use of FGT in the E-step of both. Finally, increasing the number of lines in the scan affects both CPD and our method: doubling the number of lines (from 20 to 40) results in a similar increment in runtime for both methods.

B. Real Data

The proposed method was also validated on real data. The data was gathered in a set of experiments carried out in the water tank at the CIRS lab using our recently developed underwater scanner [2] mounted on Girona1000 [3] (see Fig. 1(a)). The goal of the experiments was to navigate around the structure in the figure while acquiring 3D data. Each scan line was projected to the world reference frame using data from the INS in the robot.

Out of these experiments, a set of 14 scans was established to benchmark our method against CPD (see Fig. 8). The metric

TABLE III
RUNTIME. FASTEST METHOD IS MARKED IN BOLD

L	P	N	CPD [s]	Ours [s]
20	50	5.6k	1.44	0.7
20	100	5.6k	4.02	1.2
20	150	5.6k	10.7	1.8
20	200	5.6k	30.7	2.0
20	250	5.6k	57.6	2.5
20	200	12.6k	29.6	2.3
30	200	12.6k	93.7	6.1
40	200	12.6k	232.4	20.6

used in this figure is the distance to the closest point in the model. The magnitude of the errors of both methods is larger than in the synthetic dataset, mainly due to experimental errors such as small geometric discrepancies between the model and the real structure. In any case, our method consistently outperforms CPD, achieving final reconstruction errors in the order of a few mm. The parameter values chosen for our method were $(\beta, \lambda, w) = (40, 100, 0.1)$, and equivalent values for CPD. These values were chosen empirically on a subset of the simulated data. A more robust method to determine these values according to the actual deformation of the scans should be further investigated (see experiments 1 and 11 in Fig. 8).

In order to better understand these results, we present a visual comparison of both registration methods in Fig. 9. We see that both methods can fit the model relatively well. However, our method successfully maintains the global topology of the scan by enforcing rigidity of points belonging to the same line. On the other hand, CPD achieves low distance to model but fails at maintaining straight scan lines (please bear in mind the effect of the distance metric, as explained in Section IV-A4).

V. CONCLUSION

In this work we propose a novel non-rigid registration algorithm that fully leverages on knowledge of the working principles of 3D scanners used by robots. We proved using synthetic and experimental data that our method is able to outperform state-of-the-art non-rigid registration methods, achieving higher accuracy with a lower computational complexity. Moreover, we have made the source code publicly available. Future work in this line of research may involve making our method more robust to relatively bad initial guesses, and including a way of automatically estimating the needed parameters. Finally, the results of applying this algorithm in real robotic scenarios can be used as observations in a SLAM framework or to refine the extrinsic calibration parameters of the scanner.

REFERENCES

- [1] M. Castillón, A. Palomer, J. Forest, and P. Ridao, "Underwater 3D scanner model using a biaxial MEMS mirror," *IEEE Access*, vol. 9, pp. 50231–50243, 2021.
- [2] M. Castillón, J. Forest, and P. Ridao, "Underwater 3D scanner to counteract refraction: Calibration and experimental results," *IEEE/ASME Trans. Mechatronics*, 2022, doi: [10.1109/TMECH.2022.3170504](https://doi.org/10.1109/TMECH.2022.3170504).
- [3] D. Ribas, N. Palomer, P. Ridao, M. Carreras, and A. Mallios, "Girona 500 AUV: From survey to intervention," *IEEE/ASME Trans. Mechatronics*, vol. 17, no. 1, pp. 46–53, Feb. 2012.
- [4] H. Alismail, L. D. Baker, and B. Browning, "Continuous trajectory estimation for 3D SLAM from actuated lidar," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2014, pp. 6096–6101.
- [5] D. Droschel and S. Behnke, "Efficient continuous-time SLAM for 3D lidar-based online mapping," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2018, pp. 5000–5007.
- [6] K. Daun, M. Schnaubelt, S. Kohlbrecher, and O. Von Stryk, "HectorGrapher: Continuous-time lidar SLAM with multi-resolution signed distance function registration for challenging terrain," in *Proc. IEEE Int. Symp. Safety, Secur., Rescue Robot.*, 2021, pp. 152–159.
- [7] M. Castillón, A. Palomer, J. Forest, and P. Ridao, "State of the art of underwater active optical 3D scanners," *Sensors*, vol. 19, no. 23, 2019, Art. no. 5161.
- [8] P. J. Besl and N. D. McKay, "Method for registration of 3-D shapes," in *Sensor Fusion IV: Control Paradigms and Data Structures*, vol. 1611. Bellingham, WA, USA: SPIE, Apr. 1992, pp. 586–606.
- [9] Z. Zhang, "Point matching for registration of free-form surfaces (Lecture Notes in Computer Science)," in *Comput. Anal. of Images and Patterns*, vol. 719. Berlin, Germany: Springer, 1993, no. 2, pp. 460–467.
- [10] S. Gold, A. Rangarajan, C. P. Lu, S. Pappu, and E. Mjolsness, "New algorithms for 2D and 3D point matching: Pose estimation and correspondence," *Pattern Recognit.*, vol. 31, no. 8, pp. 1019–1031, 1998.
- [11] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. Roy. Stat. Society: Ser. B*, vol. 39, no. 1, pp. 1–22, 1977.
- [12] H. Chui and A. Rangarajan, "Feature registration framework using mixture models," in *Proc. Workshop Math. Methods Biomed. Image Anal.*, 2000, pp. 190–197.
- [13] H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Comput. Vis. Image Understanding*, vol. 89, no. 2/3, pp. 114–141, 2003.
- [14] B. Jian and B. C. Vemuri, "Robust point set registration using Gaussian mixture models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 8, pp. 1633–1645, Aug. 2011.
- [15] J. Ma, J. Wu, J. Zhao, J. Jiang, H. Zhou, and Q. Z. Sheng, "Nonrigid point set registration with robust transformation learning under manifold regularization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 12, pp. 3584–3597, Dec. 2019.
- [16] F. L. Bookstein, "Principal warps: Thin-plate splines and the decomposition of deformations," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 11, no. 6, pp. 567–585, Jun. 1989.
- [17] Y. Tsai and T. Kanade, "A correlation-based approach to robust point set registration," in *Proc. Eur. Conf. Comput. Vis.*, vol. 3023, 2004, pp. 558–569.
- [18] A. Myronenko and X. Song, "Point set registration: Coherent point drifts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 12, pp. 2262–2275, Dec. 2010.
- [19] A. L. Yuille and N. M. Grzywacz, "Motion coherence theory," in *Proc. 2nd Int. Conf. Comput. Vis.*, 1988, pp. 344–353.
- [20] A. L. Yuille and N. M. Grzywacz, "A mathematical analysis of the motion coherence theory," *Int. J. Comput. Vis.*, vol. 3, no. 2, pp. 155–175, 1989.
- [21] V. Golyanik, B. Taetz, G. Reis, and D. Stricker, "Extended coherent point drift algorithm with correspondence priors and optimal subsampling," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, 2016, pp. 1–9.
- [22] W. Gao and R. Tedrake, "Filterreg: Robust and efficient probabilistic point-set registration using gaussian filter and twist parameterization," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2019, 2019, pp. 11087–11096.
- [23] O. Hirose, "A Bayesian formulation of coherent point drift," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 7, pp. 2269–2286, Jul. 2021.
- [24] O. Hirose, "Acceleration of non-rigid point set registration with downsampling and Gaussian process regression," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 43, no. 8, pp. 2858–2865, Aug. 2021.
- [25] R. Durrett, *Probability: Theory and Examples*. Cambridge, U.K.: Cambridge Univ. Press 2019.
- [26] J. Blanco, "A tutorial on SE (3) transformation parameterizations and on-manifold optimization," Univ. Malaga, Málaga, Spain, Tech. Rep. 3, 2013.
- [27] F. Girosi, M. Jones, and T. Poggio, "Regularization theory and neural networks architectures," *Neural Computation*, vol. 7, no. 2, pp. 219–269, Mar. 1995.
- [28] T. Poggio and F. Girosi, "Networks for approximation and learning," *Proc. IEEE*, vol. 78, no. 9, pp. 1481–1497, Sep. 1990.
- [29] N. Aronszajn, "Theory of reproducing kernels," *Trans. Amer. Math. Soc.*, vol. 68, no. 3, pp. 337–404, 1950.
- [30] C. E. Rasmussen and C. Williams, "Gaussian processes for machine learning," in *Adaptive Computation and Machine Learning*. Cambridge, MA, USA: MIT Press, 2006.
- [31] L. Greengard and J. Strain, "The fast Gauss transform," *SIAM J. Sci. Stat. Comput.*, vol. 12, no. 1, pp. 79–94, 1991.
- [32] A. V. Segal, D. Haehnel, and S. Thrun, "Generalized-ICP," in *Proc. Robot. Sci. Syst.*, vol. 2, 2009, Art. no. 435.

6

RESULTS AND DISCUSSION

IN this chapter we discuss and extend the results presented in this thesis. First, in Section 6.1 we briefly summarize the results reported in each chapter of the thesis. Second, we present unpublished experimental 3D data in Section 6.2. These point clouds were gathered with our underwater laser scanner mounted on Girona 1000 AUV in several mapping missions at the harbour of Sant Feliu de Guíxols (Girona, Spain).

6.1 Summary of the completed work

In this thesis, we have focused on developing systems and methods aimed at providing UUVs with accurate 3D underwater perception. In this section, we briefly summarize the main contributions presented in each chapter of this document.

In Chapter 2, we studied the characteristics and challenges of light-based underwater 3D perception. We also analyzed technologies currently used for underwater light projection, focusing on their features and their limitations. Moreover, we quantitatively compared recently developed underwater 3D scanners for academic and commercial purposes.

In Chapter 3, we presented the ray-tracing model of an underwater 3D scanner using a 2-axis mirror. This model allowed us to study theoretically the expected quality of the light projection according to possible miscalibrations of the different model parameters. Moreover, we used this model to establish the optimal characteristics that the scanner components should have in order to optimize performance criteria, such as FoV.

In Chapter 4, we explained the development and the hardware realization of a novel underwater 3D scanner based on the aforementioned ray-tracing model. Moreover, we introduced a calibration method based on numeric projection functions aimed at simplifying the calibration process. Finally, we characterized the scanner performance based on 3D reconstruction quality metrics with data gathered experimentally in our water tank. To the best of our knowledge, this is the first underwater 3D scanner that actively counteracts refraction when projecting light.

In Chapter 5, we presented a novel non-rigid point cloud registration method. The aim of the algorithm is to reduce the motion distortion typically present in scans gathered by LLSs. In comparison to state-of-the-art non-rigid registration methods, we use in-line rigidity constraints to reduce computational complexity and increase accuracy. We evaluated the algorithm on synthetic data and on experimental scans acquired using our underwater 3D scanner.

6.2 Experiments

Apart from the results already presented in previous chapters, the underwater 3D laser scanner developed during this thesis was also tested at sea mounted on Girona 1000 AUV. The process of integrating the scanner into the AUV is explained in Section 6.2.1. The scanning results at the harbour of Sant Feliu de Guíxols (Girona, Spain) are shown in Section 6.2.2. The point clouds shown in the figures of this section have been given arbitrary colours to make them easier to visualize.

6.2.1 Extrinsic calibration of the scanner

When mounting a sensor in a mobile robot, an essential step is calibrating its extrinsic parameters with respect to the robot, so that its pose in the world reference frame can be computed using robot odometry. In the case of a 3D scanner, achieving a high-quality calibration is paramount to obtaining high-quality point clouds. Since the range and FoV of the sensor are limited, many scans need to be combined in order to cover a sufficiently large part of the inspected structure. This composition is done using the most recent estimate of the robot pose in the world, and the scanner pose with respect to the robot.

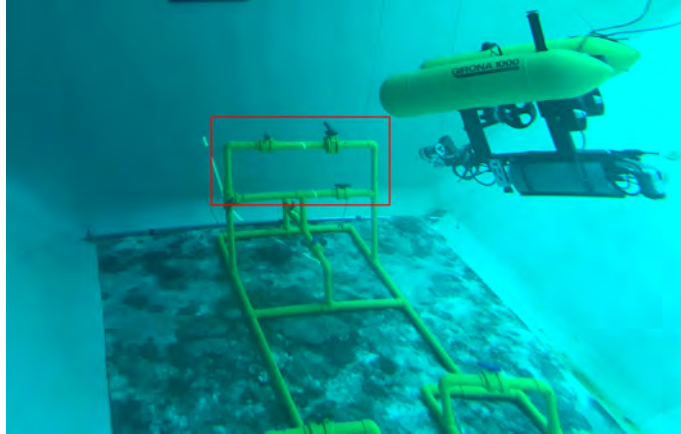


Figure 6.1: Girona 1000 AUV equipped with our 3D laser scanner inspecting a mock-up of an off-shore structure at the CIRS water tank. The red box delimits the scanned area considered in Fig. 6.2.

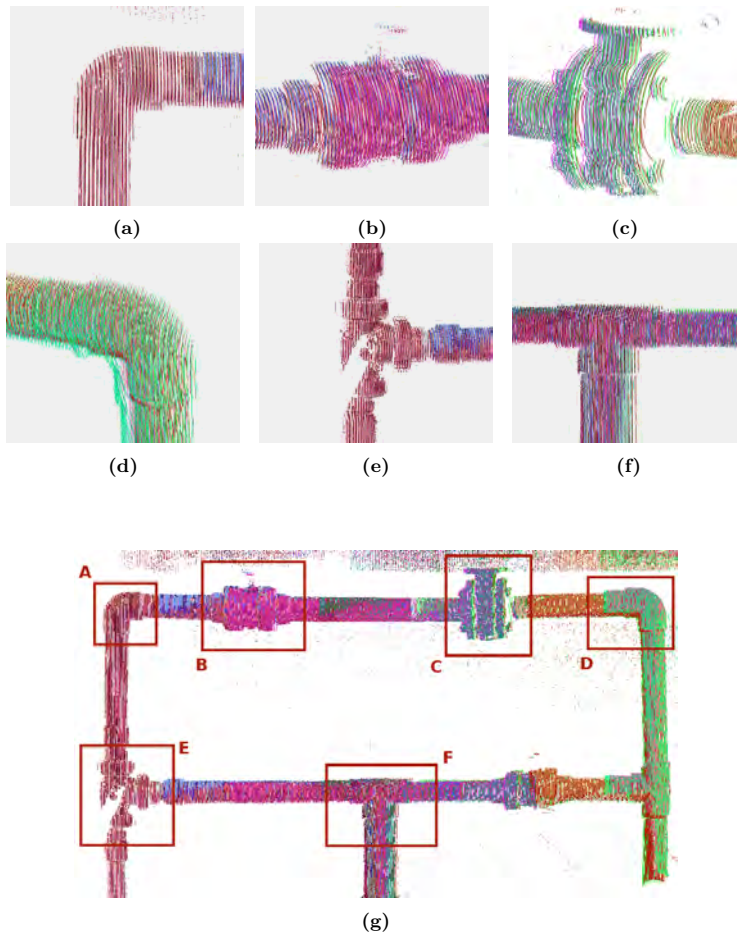


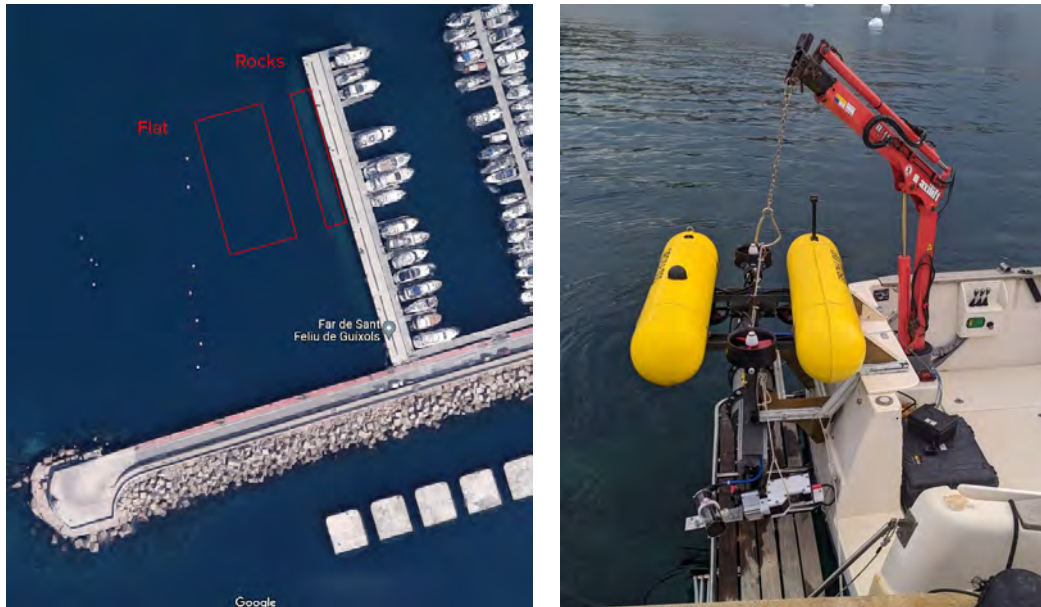
Figure 6.2: Combined partial dynamic scans into a single point cloud of the target underwater structure using the calibrated extrinsic parameters of the scanner. Point clouds (a) to (f) show the details A to F in (g). Figure (g) is made up of 25 scans in total. Each one of them is painted with an arbitrary colour for visualization. Please note that the scans were combined based on robot odometry only without using any registration method.

In our case, the extrinsic calibration of the 6D pose of the camera in the reference frame of the robot was performed using landmark graph SLAM, as explained in [5]. The results of the extrinsic calibration were validated in the water tank at CIRS (see Fig. 6.1). The resulting point clouds of such an inspection are shown in Fig. 6.2. We can see that using the high-end INSSs with which the Girona 1000 AUV is equipped allows obtaining a good geometric consistency of the combined map.

6.2.2 Sea experiments

A highly relevant contribution of this thesis is the experimental demonstration of accurate underwater 3D perception capabilities in a real environment. When compared to working under the controlled conditions of our water tank, carrying out experiments at sea presents significant challenges:

- First, the lighting conditions at sea during the day are potentially much worse. The 3D reconstruction method used by laser scanners assumes that the projected light can be clearly detected. However, strong sunlight increases the brightness of the overall image, making the laser harder to detect. This is especially noticeable at shallow depths. For this reason, we decided to carry out our experiments at night.
- Moreover, a change in water salinity modifies the refraction index of water. This may have a negative impact on the quality of the 3D reconstruction if the scanner is calibrated in the freshwater of the water tank and used in seawater. However, we decided not to recalibrate the scanner in seawater in order to assess the robustness against different salinity conditions.



(a) Satellite view of St. Feliu Harbour (source: Google Maps). The inspected areas of flat seafloor and rocks are marked in red.

(b) Girona 1000 AUV equipped with our prototype laser scanner before deployment.

Figure 6.3: Experimental data gathering at sea at the harbour of Sant Feliu de Guíxols.

- Finally, the large quantities of floating particles in seawater present a major challenge to light-based underwater sensors. As explained in Chapter 1, these particles reflect the projected light back to the sensor, producing backscatter. The implications of backscatter and the strategies used to tackle it will be further explained in Section 6.2.2.2.

For logistic convenience, the sea experiments covered in this section were carried out in two different areas of the St. Feliu harbour (see Fig. 6.3): the rocks close to the pier at a depth of 2 m to 5 m and the 10m-deep flat seafloor. This last part was done autonomously by the robot following a lawnmower trajectory.

6.2.2.1 Results

In Fig. 6.4 we see an example of a local map of the rock area. The area covered by the scans shown in this figure is around $3.5\text{m}\times 1.5\text{m}$. Despite not having a ground truth 3D model of the scanned rocks, we can qualitatively appreciate the geometric consistency of the resulting local map using only dead reckoning. Further tests are needed, as proposed in Chapter 7, but scans seem robust against different salinity conditions.

In Fig. 6.5 we can see in another part of the rock area how the global map is built incrementally with partial observations. Please note that just like in Fig. 6.1, the position of the scans in the common reference frame is computed using only robot odometry, without employing any registration method.

Finally, in Fig. 6.6 we see the top view of the autonomous mapping of the seafloor. In this case, the resulting 3D map was mostly flat and did not contain many objects with remarkable shapes apart from a few rocks. However, a potential application of such a mapping strategy could be finding lost objects that are believed to be constrained to an area of several hundred square meters.

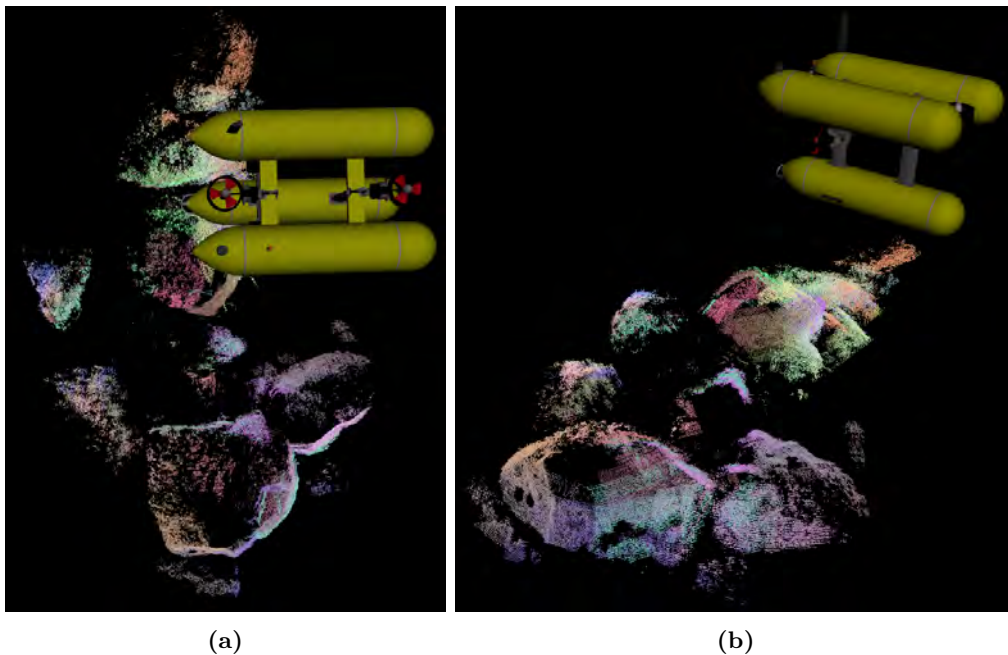


Figure 6.4: Two views of the same scanned rocks.

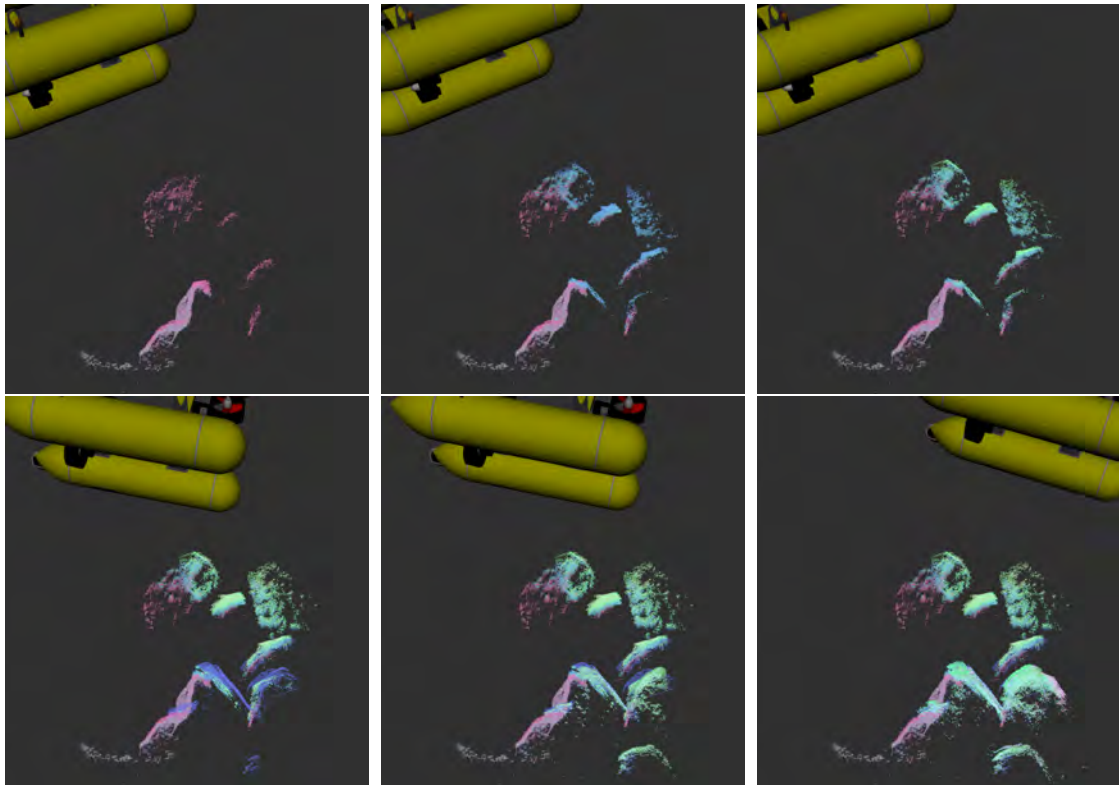


Figure 6.5: Incremental creation of the 3D reconstruction of some rocks by aggregating individual scans.

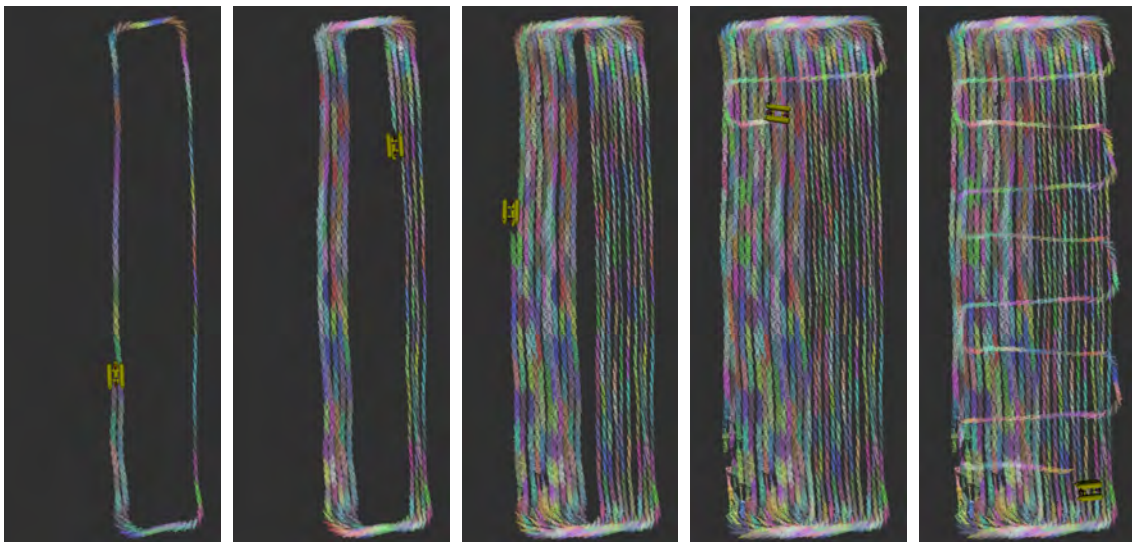


Figure 6.6: Top view of the autonomous seafloor survey. The total covered area was $30\text{m} \times 8\text{m}$ at a depth of around 10 m.

6.2.2.2 Backscatter

Particles floating in seawater cause backscatter: they reflect the projected light back to the sensor, which results in spurious points in the final scan, as shown in Fig. 6.7a. Our approach to minimize the number of these noisy detections was to post-process each scan using the statistical outlier removal algorithm implemented on PCL [16]. This method first computes the mean distance from each point in the scan to its closest N neighbours, and then removes from the dataset all the points whose mean distances are outside an interval defined by the global distances mean and standard deviation. The result of this backscatter removal is shown in Fig. 6.7b.

6.2.2.3 Odometry drift

The INS mounted on Girona 1000 is very reliable, particularly when the robot navigates over a flat or smooth sea bottom because the velocity readings of the Doppler velocity log (DVL) are very accurate. However, when the robot navigates above a rougher terrain (such as the rocks), the INS accumulates drift much faster. In Fig. 6.8 we can see how the drift accumulated by the INS affects the quality of the final 3D reconstruction. In Fig. 6.8a we see one of the few rocks on the planar seafloor. The aggregation of scans using only dead reckoning is very consistent. However, in Fig. 6.8b we see how the scans are composed incorrectly due to drift, which produces a geometrically inconsistent 3D reconstruction.

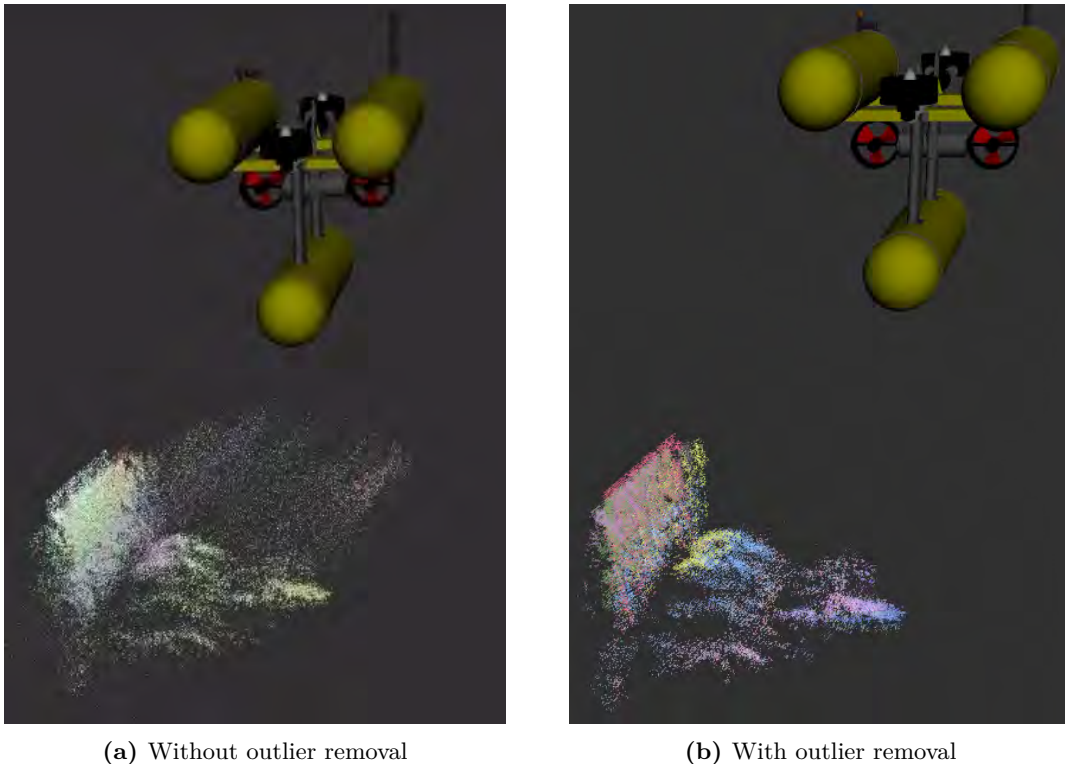
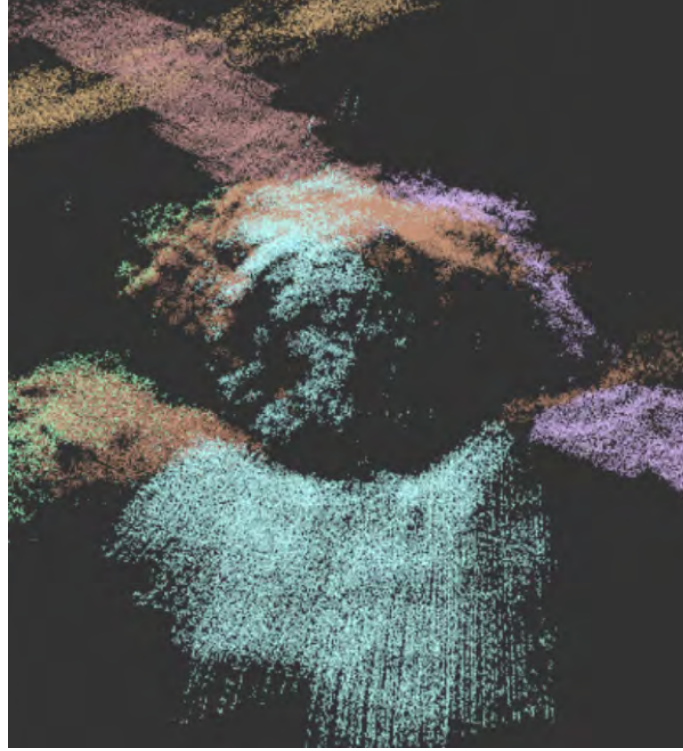
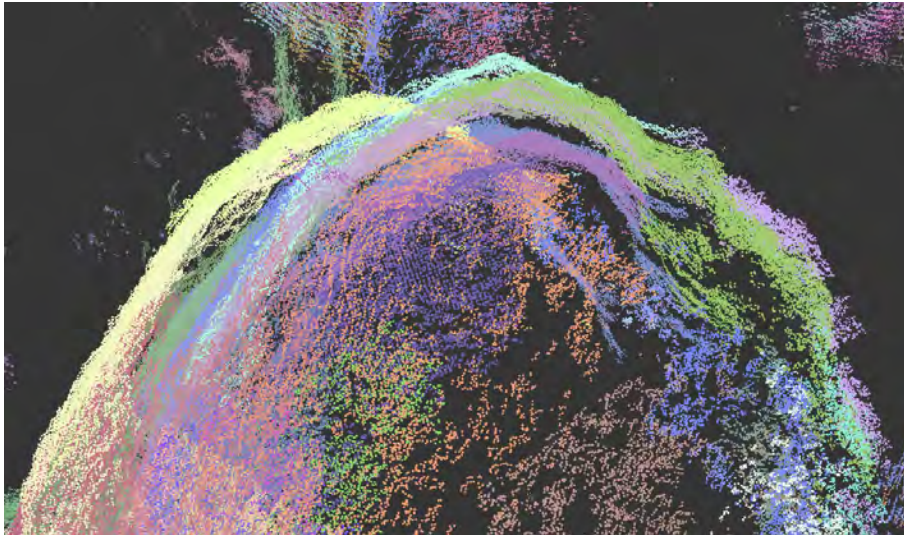


Figure 6.7: Comparison on the degradation of the point clouds due to spurious artifacts with and without using post-processing outlier removal.



(a) Aggregated scans at the flat area of the seafloor (low drift).



(b) Aggregated scans at the rocky area of the seafloor (high drift).

Figure 6.8: Comparison of how the drift accumulated by the inertial sensors of the robot affects the quality of the final 3D reconstruction.

7

CONCLUSIONS AND FUTURE WORK

THIS chapter closes this thesis in two last sections. First, the main conclusions of this work are summarized in Section 7.1. Finally, new research lines for future work are proposed in Section 7.2.

7.1 Contributions of this thesis

This thesis has contributed to advancing the state of the art in underwater 3D laser-based perception and in non-rigid 3D point cloud registration. We can break down these two general contributions into four more particular points:

Extensive and comprehensible literature survey The first contribution presented in this thesis was an extensive and comprehensible review of the principles and technologies of underwater 3D laser scanning. In the article presented in Chapter 2 we explained the challenges of underwater optical imaging and compiled light projection technologies along with their uses and limitations. Moreover, we compared quantitatively academic and commercial underwater laser scanners present in the literature.

Counteracting refraction at light projection In this thesis, we have also presented a ray-tracing model of a 2-axis underwater scanner. The extra rotation axis of the mirror allows us to control the direction of the laser beam projection. We applied this ability to counteract refraction with the aim of reducing the computational complexity of the 3D triangulation. Moreover, we created a prototype scanner to prove we could counteract refraction in practice. We proved the usability of the scanner in experiments carried out both in the water tank and at the sea.

Simple calibration based on projection functions During this thesis, we also developed a simple calibration procedure for the scanner based on numeric projection functions. This allows us to treat all the parameters related to each optical component in the scanner as a black box, so that we only estimate their joint effect on the laser beam direction without assuming any projection model *a priori*.

Non-rigid point cloud registration Finally, we also present a novel non-rigid 3D point cloud registration method able to reduce the motion distortion present in some dynamic scans acquired by LLSs. Our algorithm exploits in-line rigidity to limit the dimensionality of the non-rigid displacement field, which results in reduced computational complexity and increased accuracy. We proved this method both on synthetic and real scans.

7.2 Future work

Despite the efforts of the work presented in this thesis, accurate underwater laser-based 3D perception is far from solved. Consequently, we conclude this thesis by pointing out and discussing research lines that can potentially contribute to this end in the future.

Underwater 3D scanner The underwater 3D scanner presented in Chapter 4 can scan at high accuracy at a range of up to around 5 meters in the clear water of the water tank. However, achieving such a scanning range in the open sea is much more challenging due to floating particles and the strong exposition to sunlight. Possible approaches to tackle this and other challenges are now proposed.

Higher laser power and backscatter A direct approach to achieving longer scanning ranges would be installing a more powerful laser source. However, illuminating the underwater scene with a stronger light may have the counter-intuitive result of a shorter range due to backscatter. As shown in Fig. 6.7a, the particles floating in the water reflect the laser light back to the camera before it reaches the target object, creating many artifacts. Even though we can minimize the number of outliers using statistical removal methods, ideally the scanner should produce as few as possible. A straightforward approach to mitigate backscatter could be increasing the scanner baseline, but that would have side effects on the FoV. A more elaborated approach could be redesigning the sensor into a range gating scanner. Such scanners delay the opening of the camera exposure in order to capture light once it has already travelled a minimum distance.

Influence of the water refraction index The design of the scanning pattern is dependent on the refraction index of water. This might compromise the quality of the 3D data if the same configuration is used for missions in different types of water. It is well known that the water refraction index varies according to different factors, including water salinity, temperature, and pressure [17]. Despite the relative robustness of the scanner shown in Section 6.2.2, ideally, the refraction index should either be measured or estimated, so that the scanning pattern could be modified accordingly.

Scanning speed Currently, our scanner typically operates at scanning speeds of 50 to 100 scanned lines per second, which corresponds to up to 200k scanned points per second. Despite being enough for manipulation applications, it would be desirable to increase this speed for other operations such as inspection and mapping. A critical step in this direction would be to design a more robust dynamic control of the 2-axis mirror which could work at higher frequencies without exciting its resonance frequencies.

New scanning patterns We have developed a method to accurately project light underwater in the desired direction. We have only tested it to project planes because that is probably optimal for scanning. However, if the scope of applications of the projection system is broadened, other projection patterns could prove more suitable.

Automatic calibration Accurate underwater 3D scanners could become a crucial sensor for marine archaeologists and biologists. In order to expand its usability, a desirable property of the sensor would be automatic calibration. Therefore, it might be worth it to investigate possible automatic calibration procedures so that it could be easily operated and reconfigured by non-technical users.

Multimodal perception For certain applications, it could be desirable to capture the texture of the 3D data. This could be achieved by acquiring simultaneously the scene using a colour camera, which would result in a conceptually similar system to [18].

Non-rigid point cloud registration The non-rigid point cloud registration method proposed in Chapter 5 has been proved able to successfully remove the distortion from real underwater 3D scans. However, it could potentially benefit from the following improvements:

Higher robustness A very desirable property of the algorithm would be to show a more robust behaviour when matching point clouds deformed by large rotations. Likewise, increasing its robustness to worse initial guesses would broaden its applicability.

Automatic parameter estimation The results of the registration algorithm are relatively sensitive to the input parameters β and λ . Ideally, the value of these parameters could be inferred from navigation data, such as current linear and rotational velocities. This would contribute to making it easier to use by non-experts.

SLAM integration The next natural step in this research line would be to investigate how to use the outputs of this method as observations in a **SLAM** framework. Our registration algorithm provides the relative 6-DoF transformation that needs to be applied to each robot pose. These results can be used as constraints to the **SLAM** problem to obtain more accurate navigation and mapping results.

BIBLIOGRAPHY

- [1] **Miguel Castellón**, Albert Palomer, Josep Forest, and Pere Ridao. “State of the Art of Underwater Active Optical 3D Scanners”. In: *Sensors* 19.23 (2019), page 5161. DOI: [10.3390/s19235161](https://doi.org/10.3390/s19235161) (cited on pages [xiii](#), [15](#)).
- [2] **Miguel Castellón**, Albert Palomer, Josep Forest, and Pere Ridao. “Underwater 3D scanner model using a biaxial MEMS mirror”. In: *IEEE Access* 9 (2021), pages 50231–50243. DOI: [10.1109/ACCESS.2021.3069189](https://doi.org/10.1109/ACCESS.2021.3069189) (cited on pages [xiii](#), [51](#)).
- [3] **Miguel Castellón**, Josep Forest, and Pere Ridao. “Underwater 3D scanner to counteract refraction: calibration and experimental results”. In: *IEEE/ASME Transactions on Mechatronics* (2022). DOI: [10.1109/TMECH.2022.3170504](https://doi.org/10.1109/TMECH.2022.3170504) (cited on pages [xiv](#), [65](#)).
- [4] **Miguel Castellón**, Pere Ridao, Roland Siegwart, and César Cadena. “Linewise Non-Rigid Point Cloud Registration”. In: *IEEE Robotics and Automation Letters* 7.3 (2022), pages 7044–7051. DOI: [10.1109/LRA.2022.3180038](https://doi.org/10.1109/LRA.2022.3180038) (cited on pages [xiv](#), [75](#)).
- [5] **Miguel Castellón**, Roger Pi, Narcís Palomeras, and Pere Ridao. “Extrinsic visual–inertial calibration for motion distortion correction of underwater 3D scans”. In: *IEEE Access* 9 (2021), pages 93384–93398. DOI: [10.1109/ACCESS.2021.3092180](https://doi.org/10.1109/ACCESS.2021.3092180) (cited on pages [xiv](#), [11](#), [88](#)).
- [6] Jeygopi Panisilvam, **Miguel Castellón**, Nicholas Lawrance, and Roland Siegwart. “Conditioned deep feature consistent variational autoencoder for simulating realistic sonar images”. In: *OCEANS 2022 - Hampton Roads*. 2022 (cited on page [xv](#)).
- [7] Dominic Marti. “Underwater Volumetric Occupancy Mapping with Imaging Sonar”. Master’s thesis. ETH Zurich, 2022. DOI: [20.500.11850/551563](https://doi.org/20.500.11850/551563) (cited on page [xv](#)).
- [8] Josep Bosch, Klemen Istenic, Nuno Gracias, Rafael Garcia, and Pere Ridao. “Omni-directional Multicamera Video Stitching Using Depth Maps”. In: *IEEE Journal of Oceanic Engineering* (2019), pages 1–16. DOI: [10.1109/joe.2019.2924276](https://doi.org/10.1109/joe.2019.2924276) (cited on page [8](#)).
- [9] Shukai Chi, Zexiao Xie, and Wenzhu Chen. “A Laser Line auto-scanning system for underwater 3D reconstruction”. In: *Sensors* 16.9 (Sept. 2016), page 1534. DOI: [10.3390/s16091534](https://doi.org/10.3390/s16091534) (cited on page [10](#)).

- [10] Changjun Gu, Yang Cong, and Gan Sun. “Three Birds, One Stone: Unified Laser-Based 3-D Reconstruction across Different Media”. In: *IEEE Transactions on Instrumentation and Measurement* 70 (2021). DOI: 10.1109/TIM.2020.3039641 (cited on page 10).
- [11] Albert Palomer, Pere Ridao, Josep Forest, and David Ribas. “Underwater Laser Scanner: Ray-based Model and Calibration”. In: *IEEE/ASME Transactions on Mechatronics* 24.5 (2019), pages 1986–1997. DOI: 10.1109/TMECH.2019.2929652 (cited on page 10).
- [12] A. Segal, Dirk Haehnel, and Sebastian Thrun. “Generalized-ICP”. In: *Robotics: Science and Systems V*. Robotics: Science and Systems Foundation, June 2009. DOI: 10.15607/RSS.2009.V.021 (cited on page 11).
- [13] Andriy Myronenko and Xubo Song. “Point set registration: Coherent point drifts”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32.12 (2010), pages 2262–2275. DOI: 10.1109/TPAMI.2010.46 (cited on page 11).
- [14] Marc Carreras, Juan David Hernandez, Eduard Vidal, Narcis Palomeras, David Ribas, and Pere Ridao. “Sparus II AUV - A Hovering Vehicle for Seabed Inspection”. In: *IEEE Journal of Oceanic Engineering* 43.2 (Apr. 2018), pages 344–355. DOI: 10.1109/JOE.2018.2792278 (cited on page 12).
- [15] David Ribas, Narcís Palomeras, Pere Ridao, Marc Carreras, and Angelos Mallios. “Girona 500 AUV: From survey to intervention”. In: *IEEE/ASME Transactions on Mechatronics* 17.1 (Feb. 2012), pages 46–53. DOI: 10.1109/TMECH.2011.2174065 (cited on pages 12, 14, 75).
- [16] The Point Cloud Library (PCL). *Statistical outlier removal*. https://pcl.readthedocs.io/projects/tutorials/en/latest/statistical_outlier.html. [Online; accessed 2022-06-28] (cited on page 91).
- [17] P Schiebener, Johannes Straub, J. M. H. Levelt Sengers, and J. S. Gallagher. “Refractive index of water and steam as function of wavelength, temperature and density”. In: *Journal of Physical and Chemical Reference Data* 19.3 (May 1990), pages 677–717. DOI: 10.1063/1.555859 (cited on page 95).
- [18] Andry Maykol Pinto and Anibal C. Matos. “MARESyE: A hybrid imaging system for underwater robotic applications”. In: *Information Fusion* 55 (Mar. 2020), pages 16–29. DOI: 10.1016/j.inffus.2019.07.014 (cited on page 95).