Underwater Laser Scanner: Ray-Based Model and Calibration

Albert Palomer , Pere Ridao, *Member, IEEE*, Josep Forest , and David Ribas

Abstract—Underwater 3-D perception is crucial for autonomous manipulation and mapping. Using a mirror-galvanometer system to steer a laser line projector and using triangulation with a camera it is possible to produce full 3-D perception, while avoiding moving the sensor with respect to the scene. This paper presents a ray-based model of a laser scanner and a calibration method for such a model. The model takes into account the distortion introduced to the system by the flat viewports used for both the laser and the camera. The paper also introduces the computation of elliptical cones as a way to synthesize the ray-based model and speed up the triangulation computation. The present study reports results obtained using the real sensor in a water tank to reconstruct objects and planes to assess the sensor's accuracy, as well as a comparison between the two triangulation methods (ray-based and elliptical cone).

Index Terms—Laser scanner, refraction, structured light, three-dimensional (3-D), triangulation, underwater.

I. INTRODUCTION

DEPTH and RGB-D cameras have accelerated the development of new computer vision and robotics techniques due to their low cost, good accuracy, and resolution. The work done in mapping using this type of sensors has been extensive [1]. In [2], graphics processing unit implemented algorithms which allowed tracking the position of a depth camera used to reconstruct indoor scenes. Moreover, this type of sensor has also been used for on-line three-dimensional (3-D) object identification [3] and semantic mapping [4]. These types of 3-D-perception cameras, besides other types of 3-D-perception sensors, such as laser scanners, have also helped to advance the field of path planning [5]. Several studies using real robots, as well as robotic manipulators, have been presented using *MoveIt!* and its integration with Octomap and open motion planning library.
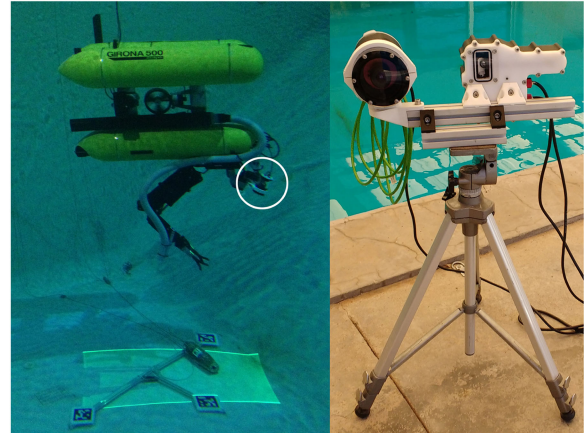
Fig. 1. Laser scanning system developed in this study mounted on the Girona 500 AUV [6] during experiments in the water tank (left, highlighted with a white circle) and mounted on a tripod (right). Please, note that the left-hand image presents an area illuminated by the laser instead of a laser line, because the camera used to take the picture is much slower than the mirror-galvanometer system used to steer the laser light.

Despite all these works, underwater 3-D-perception is still a developing field of study. Although some work has been done using scanning multibeam sonars producing low resolution reconstructions (see [7], [8]), computer vision techniques are a popular approach to the underwater 3-D-perception problem because of the quality of the produced models. These computer vision techniques can be divided in two groups depending on the source of light being used. There are passive techniques which use a light source that illuminates the whole scene as homogeneously as possible by using ambient light or artificial light such as lamps and spotlights. In contrast, active techniques use a pattern projected onto the scene to triangulate the different points producing reconstructions with better results than passive techniques, especially in the featureless areas where the 3-D of passive techniques is computed assuming continuity from areas with features.

Passive techniques are essentially based on multiple views of a scene that can be obtained using stereo vision or a single camera in different vantage positions [9]–[11]. Here, the triangulation is done by intersecting the two rays that pass through a pixel in each image identified with the same feature. This type of technique only works in environments where distinct features can be identified in the images and, as far as the present authors are aware, there is no work presenting underwater on-line 3-D

point clouds computation using such techniques, except in the case of stereo vision.

Active vision techniques can project different patterns: either t2-D (see [12]–[14]) or 1-D (see [15]–[17]). In both cases, the triangulation is computed by intersecting the ray that passes through a pixel lit up by the pattern, and the projected pattern itself. Depth cameras fall within the type of active sensor that uses 2-D patterns. In [14], a Kinect sensor was mounted in an underwater housing with a flat viewport. Although this study presents a model to correct for the flat viewport distortion, the attenuation of the infrared light which it employs made it ineffective at distances greater than 20 cm. Using visible light and multiple binary patterns, the authors of [12] present a 3-D reconstruction framework with submillimetric accuracy but with an acquisition time that does not allow for real-time operations. At the cost of decreasing the resolution, [13] present a study where the 3-D is gathered with a single shot. In their study, they use a lens to diffract a laser beam into a multiple line configuration. The advantage of using a laser in comparison to the noncoherent light used in [14] and [12] is that laser light can be highly collimated and has a higher optical density. These characteristics allow for a better propagation of the light in the water medium. The simplest example was presented in [16]. There, the authors used a single laser line projector and a camera to obtain profiles of the seabed. The sensor was mounted on a moving underwater vehicle so that profiles of different parts of the scene could be gathered for the overall 3-D reconstruction. In [17], a similar system was used for mapping small areas. In this case, the laser was mounted on an underwater robotic arm and the camera was mounted with a fixed baseline to the base of the arm. In this method, the motion of the laser line on the scene was produced by moving the arm. In [15], the light was projected onto different parts of the scene by steering the light using a mirror actuated with a galvanometer. This setup has several advantages when compared with vehicle- or robotic arm-mounted systems. The light can be projected to different parts of the scene much faster because the mirror-galvanometer system has a very low inertia and also, because it moves in air and therefore, suffers less friction. There are some companies, such as 2G Robotics [18], Newton Labs [19], or Savante [20], which have market-ready products with systems similar to these laser scanners.

This paper is structured as follows. First, the related work is outlined in Section II. Then, the sensor (see Fig. 1) is described in Section III; the mathematical definition of the sensor model is presented in Section IV; the triangulation process in Section V and the calibration procedure in Section VI. Finally, the results and the conclusions and future work are explained in Sections VII and VIII, respectively.

II. RELATED WORK

To the best of the authors' knowledge, only three works have been presented using an underwater triangulation-based laser scanner [15], [21], [22]. In [15], the authors used a planar viewport for sealing the laser scanner housing despite the well known refraction that light suffers when traveling through flat
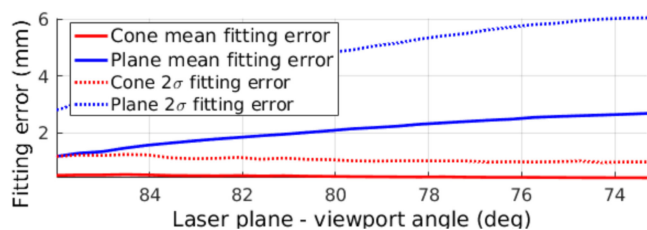

Fig. 2. Comparison between plane and cone fitting over laser 3-D points [23].

viewports [23]. The model for the light projection assumed that the refraction on the laser viewport only changed the plane equation but did not change the projected shape. However, because the rays of the laser are not parallel when they meet the viewport surface, the refraction of each ray is different and the plane is distorted.

In contrast to using a flat viewport, [21] and [22] attempted to minimize the deformation using a better hardware design. To avoid such deformation, it is necessary to mount the system in a way that the laser plane contains the normal to the viewport surface for all the intersection points between the laser plane and the viewport. The shapes meeting this condition are spheres and cylinders. The only layout in which the plane equation does not change is when it contains and rotates around an axis containing the center of the sphere. Moreover, if the focal point is not coincident with the center of the sphere the aperture of the laser will be affected. When using a cylinder, as was done in [22], the surface containing the light remains planar as long as it contains and rotates around the revolution axis of the cylinder. Nevertheless, there is no position for the focal point in which the cylindrical viewport does not affect the aperture of the laser. To achieve high accuracy, any deviation from such ideal conditions should be taken into account in the sensor model. In contrast, using a flat viewport makes the manufacturing and mounting process easier at the cost of using a suitable surface to represent the light projection.

The same light bending effect is observed in both airborne and terrestrial laser scanners [24]. However, as far as the authors are aware, these systems work with individual laser beams instead of laser fans.[1] For this reason, their models take into account such deformation for the individual beams. However, in the discussed underwater laser scanners, the equations of the individual laser beams that illuminated the specific pixels of the camera are not known. For this reason, to triangulate the 3-D points using a ray model requires an iterative process (see Section V-A). In [23], the authors presented theoretical and experimental results using an elliptic cone to model projected light after it had passed through a flat viewport (i.e. modeling a shape containing all the rays of the laser fan). There, the experimental results showed that the average fitting error for the elliptical cone is significantly reduced when compared to a plane (see Fig. 2). Moreover, the line-cone intersection can be solved analytically

---

[1]Some works use multiple laser beams which might, or might not, be configured into a fan. In any case, the equations of the individual laser beams are known.

(see Section V-B), drastically reducing the computation time when compared to the iterative ray model triangulation (see Section VII-A). Although [23] uses an elliptical cone, any quadratic surface could potentially be used with the closed-form solution triangulation advantage. However, studding different surfaces to represent the laser light underwater is beyond the scope of this article.

This study proposes a real-time underwater scanner. It uses a laser line steered with a mirror galvanometer and projected through a flat viewport, coupled to a real-time camera. This paper explains in detail the ray-based model of the light projection used by [23] to study the elliptical cone representation of the laser light. We also present the new light receiving model that includes the camera as well as the model of its flat viewport. The main contribution of this paper is the calibration procedure of the ray-based model of both light-projection and light-receiving parts of the sensor. Moreover, this paper presents two different triangulation techniques for the laser scanner: one based on the ray model and another based on the elliptical cone representation of the light.

III. SENSOR DESCRIPTION

The present study presents a hardware system similar to the one in [15]. On the light emission side, there are no remarkable differences between the two works. In both cases, a laser line is reflected by a mirror steered with a galvanometer and projected onto the observed scene through a flat viewport. However, the light model for the triangulation used in both works differs. While [15] uses a plane, here we use an elliptic cone surface or a ray-based model to represent the light in the underwater medium to better describe the shape of the projected light [23]. On the light receiving side, both sensors have similar hardware. However, the camera used in this study implements the laser peak detection from [25] in the camera hardware itself. Then, the camera sends the 2-D laser points to the computer for the triangulation. Another point of divergence is the model of the overall light receiving side. While the previous work uses a pin-hole camera model with a traditional radial and tangential distortion model such as the Open Source Computer Vision (OpenCV) one [26], here we model the refraction of the flat viewport interface by computing each ray's path from the camera, through the viewport and onto the scene. The calibration method proposed here estimates the position of the different elements of the sensor using just a subset of the possible mirror positions, although the results are applicable to the complete range of motion.

This laser scanner can work at a wide variety of angular resolutions for the galvanometer-mirror system (see Fig. 3). This allows the sensor to work at a very high resolution (0.008 degrees between projected lines) with a slow acquisition time, or at a lower resolution with a higher number of scans per second. This second mode of operation, and any improvement that increments the scanning frame rate, makes this sensor potentially useful for moving vehicles such as autonomous underwater vehicles (AUV) or remotely operated underwater vehicles, because the distortion introduced by the sensor motion while gathering the
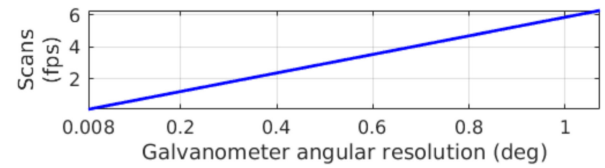

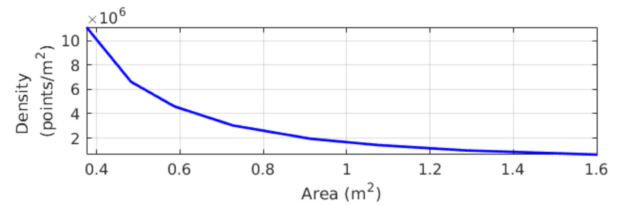Fig. 3. Galvanometer resolution and scanning rate of the laser scanner.

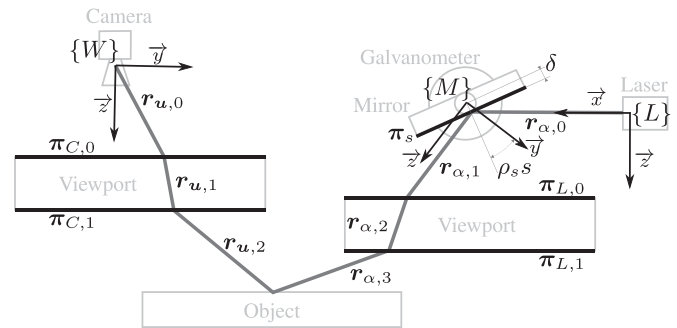Fig. 4. Point cloud density evolution with distance.

Fig. 5. Illustration with the different elements of the sensor (clear gray), the elements of the mathematical model (black) and the light path (dark gray).

scan can be reduced by increasing the frame rate of the sensor. In general, decreasing the motion distortion makes this sensor more useful for mapping and intervention applications. However, in this study, the sensor will be mounted on a tripod (i.e., static with respect to the scene) to gather the calibration data as well as the data used in the results (see Section VII) to achieve the best accuracy. It is important to note that although the sensor is calibrated using high resolution data, it will also be able to operate at lower resolutions without recalibration. The current configuration of the sensor mounts a 2048×1088 camera with 8-mm focal length optics and a 50-mW laser line with an aperture of $45°$ in air. With this configuration, and scanning at a resolution of $0.04°$, the point cloud density variation with distance to scanned area varies according to Fig. 4.

IV. SENSOR MODEL

This section presents the ray-based model of the sensor (see Fig. 5). The model has two main parts, one where the light is emitted and projected onto the scene (referred to as "laser model" and presented in Section IV-A); and another where the light is projected back onto the image plane (referred to as "camera model," and detailed in Section IV-B). The model is composed of the different models for the different parts: camera,

camera viewport, laser fan emitter, mirror-galvanometer system, and laser viewport. The model proposes a way to compute the different surfaces (mirror and viewports) so that the light path can be computed for each ray, either from the laser source or from the camera.

Only two phenomena are taken into account in the model: reflection and refraction. Light reflection is the process by which light changes its direction when it reaches a given surface without crossing it

$$r_l = i - 2 \left(i \cdot n \right) n. \qquad (1)$$

This process only depends on the angle ν between the incident ray i and the normal of the surface where it is reflected n.

Light refraction is the change in direction that a light beam suffers when crossing the surface separating two different media and follows Snell's law:

$$r_r = \frac{n_1}{n_2} \left(n \times (-n \times i) \right) - n \sqrt{1 - \left(\frac{n_1}{n_2} \right)^2 (n \times i) \cdot (n \times i)}. \qquad (2)$$

This law involves three elements: 1) the index of refraction of the first medium n_1, 2) the index of refraction of the second medium n_2, and 3) the angle ν between the incident light i and the normal of the surface separating the two media n.

In the ray-based model, the ray-to-plane intersection is a problem that needs to be solved several times. In this study, a ray $r = \left[v \ o \right]^T : p = \lambda v + o$, with free parameter λ, is defined by its direction vector $v = \left[v_x \ v_y \ v_z \right]^T$ and its origin point $o = \left[o_x \ o_y \ o_z \right]^T$. Given a ray r and a plane $\pi = \left[n \ d \right]^T : n \cdot p = d$, where n is the normal of the plane and d is the distance between the plane and the origin, there exists an intersection point as long as the ray direction vector and the normal of the plane are not perpendicular:

$$p = r \bigcap \pi = \frac{d - n \cdot o}{n \cdot v} v + o \qquad (3)$$

A. Laser Model

On the light emission side, the sensor consists of a laser fan emitter, a mirror where the laser light is redirected towards different parts of the scene and a sealing viewport through which the laser is transmitted to the water. Each element is modeled in the following subsections so the equation of the laser rays underwater can be computed.

1) Laser Emitter: The laser plane is produced using a diffractive lens and a laser beam [27]. One could model the laser plane by introducing the lens and the beam into the overall model. However, for the sake of simplicity, it is assumed that the laser plane is produced by a source of light expanding at a given aperture angle.

The light projection is modeled by emitting rays on the plane, perpendicular to the direction vector $\overrightarrow{z}$ of the laser frame $\{L\}$ with a certain angle $\alpha \in [\alpha_l, \alpha_e]$ with respect to the $\overrightarrow{x}$ direction. The laser ray with respect to $\{L\}$ generated at an angle α is computed as

$$^{\{L\}}r_\alpha : \ ^{\{L\}}p_\alpha = \lambda \left[\cos \alpha \ \sin \alpha \ 0 \right]^T + \left[0 \ 0 \ 0 \right]^T \qquad (4)$$

Given the transformation of the laser with respect to the world $^{\{W\}}t_{\{L\}} = \left[t_x \ t_y \ t_z \ t_\phi \ t_\theta \ t_\psi \right]^T$, the ray $^{\{L\}}r_\alpha$ can be transformed from $\{L\}$ to $\{W\}$ using (5) where $^{\{W\}}R_{\{L\}}$ is the RPY rotation matrix ($^{\{W\}}R_{\{L\}} = \mathbf{Rot}_z(t_\psi) \, \mathbf{Rot}_y(t_\theta) \, \mathbf{Rot}_x(t_\phi)$)

$$^{\{W\}}r_\alpha : \ ^{\{W\}}p = \ ^{\{W\}}t_{\{L\}} \oplus \ ^{\{L\}}p_\alpha$$
$$= \ ^{\{W\}}R_{\{L\}} \left(\lambda \ ^{\{L\}}v_\alpha \right.$$
$$\left. + \ ^{\{L\}}p_\alpha \right) + ^{\{W\}} \left[t_x \ t_y \ t_z \right]^T. \qquad (5)$$

2) Mirror-Galvanometer: In an ideal scenario, the mirror surface where the laser is reflected would rotate around an axis contained in the reflection plane. However, a real mirror is not a single plane but a volume where reflection happens only on the surface. The galvanometer-mirror system will be able to change its position faster if the inertia of the mirror is small, which happens when the rotation axis passes through the center of mass of the mirror. This forces the rotation axis to not be contained in the reflective plane. This misalignment is taken into account in the model.

The mirror model $\Pi = \left[\rho_s \ s \ \delta \ ^{\{W\}}t_{\{M\}} \right]$ consists of the angular step ρ_s around the $\overrightarrow{x}$ axis of its reference frame $\{M\}$, the step number s, the distance to the rotation axis δ, and the transformation to the world reference frame $^{\{W\}}t_{\{M\}}$. The equation of the mirror plane $^{\{M\}}\pi_s$ can be computed using (6) and then transformed to $\{W\}$ as follows:

$$^{\{M\}}\pi_s = \left[0 \ \cos \left(\rho_s s \right) \ \sin \left(\rho_s s \right) \ \delta \right]^T \qquad (6)$$

$$^{\{W\}}\pi_s = \left[\begin{array}{c} ^{\{W\}}R_{\{M\}} \ ^{\{M\}}n_s \\ ^{\{W\}}n_s \cdot \left(^{\{W\}}t_{\{M\}} \oplus \frac{^{\{M\}}\delta}{\|^{\{M\}}n_s\|^2} \ ^{\{M\}}n_s \right) \end{array} \right]. \qquad (7)$$

3) Laser Viewport: The laser viewport $\Omega_L = \left[\pi_L \ t_L \right]$ is modeled by the central plane π_L with respect to $\{W\}$ and the thickness of the viewport. The two interfaces of the viewport are the two planes parallel to π_L at a distance $t_L / 2$.

4) Laser Ray: Since the mirror plane and the light ray are in the same reference frame $\{W\}$, the reflection $^{\{W\}}r_{\alpha,1}$ of the ray $^{\{W\}}r_{\alpha,0}$ on the mirror $^{\{W\}}\pi_s$ is computed using (1) and (3). After the light ray is reflected onto the mirror, it travels through the laser viewport Ω_L. In the first viewport surface $\pi_{L,0}$, the direction of $^{\{W\}}r_{\alpha,2}$ is computed using (2) with the incident ray $^{\{W\}}r_{\alpha,1}$ and the normal of the refraction surface n_L. The origin of the ray $^{\{W\}}r_{\alpha,2}$ can be calculated using (3) with $r_{\alpha,1}$ and $\pi_{L,0}$. The same process can be done to compute the light ray $^{\{W\}}r_{\alpha,3}$ in the water medium using $^{\{W\}}r_{\alpha,2}$ and $\pi_{L,1}$.

B. Camera Model

It is known that camera lenses introduce distortion to the pin-hole camera model. One of the effects of this distortion is that straight lines in the real world map onto curved lines on the image plane. In our case, it is extremely important that straight lines in the real world appear as straight lines in the image, otherwise, the reconstructed surface will appear bent. The OpenCV library [26] provides widely used methods for camera calibration. This report uses the implementation of [28]

in OpenCV for the in-air camera model. This includes the distortion model as well as the pin-hole camera model with focal distance f_x and f_y and camera principal point c_x and c_y.

During the image acquisition process, the light travels from the scene to the image. However, for triangulation purposes, the light ray is traced the other way around since it triangulates using an identified feature in the image, in this case a laser detection, with the projected light. The light ray that passes through a specific pixel $\boldsymbol{u}_u = [u_u \; v_u]$ in the undistorted image can be computed as follows:

$$r_{\boldsymbol{u},0} : \boldsymbol{p} = \lambda \left[(u_u - c_x)/fx \; (v_u - c_y)/fy \; 1 \right]^T + \boldsymbol{0}. \quad (8)$$

Given the camera viewport $\boldsymbol{\Omega}_C = [\boldsymbol{\pi}_C \; t_C]$, the light ray that traverses the viewport $\boldsymbol{r}_{\boldsymbol{u},1}$ and the one that travels through the water $\boldsymbol{r}_{\boldsymbol{u},2}$ can be calculated in the same way as those of the laser using equations (2) and (3) (see Section IV-A).

V. TRIANGULATION

In this study, two different triangulation techniques are presented for the same sensor: ray–ray and ray–cone. Ray–ray triangulation is more accurate than the ray–cone one. However, it is also slower because there is no closed form solution to find which angle α a laser ray $\boldsymbol{r}_{\alpha,0}$, producing $\boldsymbol{r}_{\alpha,3}$ in the water, has to follow to intersect a given camera ray $\boldsymbol{r}_{\boldsymbol{u},2}$. In contrast, the ray-cone method does have a closed form solution. However, to be able to use ray-cone intersection, an elliptic cone has to be fitted for each mirror position and this process is time-consuming. The calibration process presented in Section VI is iterative and, because of that, it is faster to use ray–ray intersection, thus avoiding the need to fit a cone for each mirror angle at each iteration of the calibration process.

A. Ray–Ray Triangulation

This triangulation process finds the angle α in which a light ray emitted from $\{L\}$ intersects a given camera ray. The ray to ray distance can be computed by finding, for each emitted ray, the point which is closest to the camera ray. Given two rays $\boldsymbol{r}_1(\lambda_1)$ and $\boldsymbol{r}_2(\lambda_2)$ the closest point can be calculated by solving

$$\begin{cases} (\lambda_2 \boldsymbol{v}_2 + \boldsymbol{p}_2 - (\lambda_1 \boldsymbol{v}_1 + \boldsymbol{p}_1)) \cdot \boldsymbol{v}_1 = 0 \\ (\lambda_2 \boldsymbol{v}_2 + \boldsymbol{p}_2 - (\lambda_1 \boldsymbol{v}_1 + \boldsymbol{p}_1)) \cdot \boldsymbol{v}_2 = 0 \end{cases}. \quad (9)$$

Defining $\mathrm{d}(\boldsymbol{r}_1, \boldsymbol{r}_2)$ as the distance function between the two rays (i.e., the distance between the two closest points of the two rays) it is possible to find the parameter α that minimizes this function between the laser ray $\boldsymbol{r}_{\alpha,3}$ and the camera ray $\boldsymbol{r}_{\boldsymbol{u},2}$ [see (10)]. From this, it is possible to find the intersection point of the camera ray with the laser model using again (9)

$$\alpha = \underset{\alpha}{\operatorname{argmin}} \, \mathrm{d}(\boldsymbol{r}_{\alpha,3}, \boldsymbol{r}_{\boldsymbol{u},2}). \quad (10)$$

B. Ray–Cone Triangulation

An elliptic cone is a two-dimensional surface $\boldsymbol{c}(h,\beta)$ in the 3-D-space with parametric equation (11) where a and b are its

aperture in the $\overrightarrow{x}$ and $\overrightarrow{y}$ directions, respectively

$$\boldsymbol{c}(h,\beta) = \begin{bmatrix} a \; h \; \cos(\beta) & b \; h \; \sin(\beta) & h \end{bmatrix}^T. \quad (11)$$

This equation represents an elliptic cone with respect to its reference frame $\{Q\}$ with its main axis aligned with the $\overrightarrow{z}$ axis and its vertex placed on the origin. Therefore, a transformation $^{\{W\}}\boldsymbol{t}_{\{Q\}}$ is necessary to represent the elliptic cone in any position in the 3D-space $\{W\}$

$$\boldsymbol{g}(h,\beta,\boldsymbol{t}) = {}^{\{W\}}\boldsymbol{t}_{\{Q\}} \oplus {}^{\{Q\}}\boldsymbol{c}(h,\beta). \quad (12)$$

Given a ray $^{\{W\}}\boldsymbol{r}(\lambda)$ in the world reference frame, it can be transformed to $\{Q\}$ using (13) to obtain $^{\{Q\}}\boldsymbol{v}$ and $^{\{Q\}}\boldsymbol{o}$

$$^{\{Q\}}\boldsymbol{r}(\lambda) : \lambda \, {}^{\{Q\}}\boldsymbol{v} + \, {}^{\{Q\}}\boldsymbol{o}$$

$$= {}^{\{W\}}\boldsymbol{R}_{\{Q\}}^{-1} \left(\lambda \, {}^{\{W\}}\boldsymbol{v} + \, {}^{\{W\}}\boldsymbol{o} - [t_x \; t_y \; t_z]^T \right). \quad (13)$$

Then, the intersection of such a ray and an elliptic cone can be computed using (14). It should be noted that the ray direction vector and origin point are referenced to $\{Q\}$

$$^{\{Q\}}\boldsymbol{r}(\lambda) = \boldsymbol{c}(h,\beta) \rightarrow \lambda \begin{bmatrix} v_x \\ v_y \\ v_z \end{bmatrix} + \begin{bmatrix} o_x \\ o_y \\ o_z \end{bmatrix} = \begin{bmatrix} a \; h \; \cos(\beta) \\ b \; h \; \sin(\beta) \\ h \end{bmatrix}. \quad (14)$$

By applying mathematical transformations to this system of equations, it is possible to obtain (15), which is a second degree polynomial. Depending on the number of real solutions, the ray and the cone can either intersect in two points, one or none. Solving the ambiguity of choosing between the two intersection points is later explained in Section VI-E

$$\lambda^2 \left(b^2 v_x^2 + a^2 v_y^2 - a^2 b^2 v_z^2 \right)$$
$$+ 2\lambda \left(b^2 v_x o_x + a^2 v_y o_y - a^2 b^2 v_z o_z \right)$$
$$+ \left(b^2 o_x^2 + a^2 o_y^2 - a^2 b^2 o_z^2 \right) = 0. \quad (15)$$

VI. CALIBRATION

This section describes the procedure for obtaining all the parameters of the sensor model presented in the previous section (see Section IV). The calibration process assumes that the relative position between laser emitter, rotation axis of the galvanometer, and camera focal point do not change regardless of the sensor being in air and without viewports, or in water with viewports. The process for obtaining the model parameters consists of four steps. First, the camera intrinsics are estimated (see Section VI-A) and then, the relative position of the laser and the mirror is computed (see Section VI-B). These two steps of the calibration process are done in air and without the viewports being mounted on the sensor. Then, both the camera and the laser viewports are mounted and their positions are estimated underwater (see Sections VI-C and VI-D). Finally, a cone is computed for each mirror angle (see Section VI-E).

A. Camera Calibration

The camera calibration process is based on multiple views of a given pattern. For each view of the pattern, pairs of image
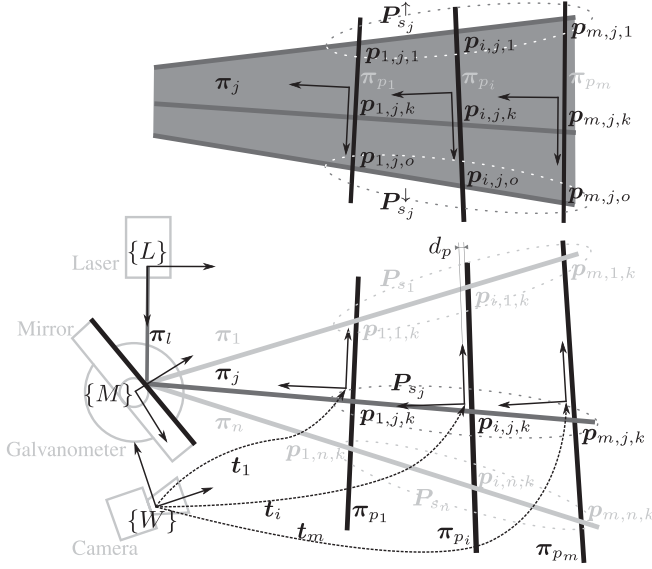
Fig. 6. Illustration with all the elements from the air laser calibration.

points (pixel detections of the pattern in the image) and object points (3-D points of the pattern) are gathered. With the set of image-object points, OpenCV is used to estimate the pin-hole camera parameters $[f_x \ f_y \ c_x \ c_y]$ as well as the pattern to camera transformation for each view.

B. Laser Calibration

The calibration of the laser model (i.e. computing the laser pose, the mirror-galvanometer pose, angular step, and distance to rotation axis), is the most complex step of the calibration process (see Fig. 6). This procedure consists of using a calibration data set (see Section VI-B1) to find a simplified model (see Section VI-B2), refining it using least squares (see Section VI-B3), and finally computing the laser pose (see Section VI-B4).

1) Laser Calibration Dataset: The calibration data used consist of several scans of a plane π_p where each scan observes the plane from a different position $\boldsymbol{T} = \{\boldsymbol{t}_1, \boldsymbol{t}_2, \dots, \boldsymbol{t}_m | \boldsymbol{t}_i = [t_{i,x} \ t_{i,y} \ t_{i,z} \ t_{i,\phi} \ t_{i,\theta} \ t_{i,\psi}]^T\}$. For each view of the plane, a pattern is placed on the projection plane, and an image is taken to estimate each $\boldsymbol{t}_i$ using OpenCV. Then, the laser stripe is projected onto the projection plane for certain mirror positions $\boldsymbol{s} = \{s_1, s_2, \dots, s_n\}$ producing a line in the image plane for each mirror position s_j and for each projection plane π_{p_i}. In each image, the sorted list of undistorted pixels $\boldsymbol{\lambda}_{i,j} = \{\boldsymbol{u}_{i,j,1}, \boldsymbol{u}_{i,j,2}, \dots, \boldsymbol{u}_{i,j,o} | \boldsymbol{u}_{i,j,k} = [u_{i,j,k} \ v_{i,j,k}] | u_{i,j,k} < u_{i,j,k+1}\}$ are identified as reflections of the laser on the scene. For each identified laser pixel on the image, its corresponding 3-D point can be computed by intersecting the ray associated to $\boldsymbol{u}_{i,j,k}$ [see (8)] with the projection plane using (3). Hence, the set of 3-D points $\boldsymbol{P}_{i,j} = \{\boldsymbol{p}_{i,j,1}, \boldsymbol{p}_{i,j,2}, \dots, \boldsymbol{p}_{i,j,o}\}$ related to $\boldsymbol{\lambda}_{i,j}$ can be computed. Given the thickness d_p of the pattern laid over the projection plane π_{p_i}, the equation of the projection plane in the pattern reference

frame $\boldsymbol{\pi}_p = [0 \ 0 \ -1 \ d_p]^T$ can be transformed to the world reference frame $\{W\}$ using $\boldsymbol{t}_i$ and (7) obtaining $\boldsymbol{\pi}_{p_i}$.

2) Simplified model: The first step of the calibration process consists in finding a simplified model where the laser source is modeled using a plane instead of a set of rays and where the reflective surface of the mirror contains the rotation axis of the galvanometer. To estimate this simplified model, first, a plane $\boldsymbol{\pi}_j$ is fitted amongst the 3-D points belonging to each angle along the m scans, in other words, $\boldsymbol{P}_{s_j} = \bigcup_{i=1}^{m} \boldsymbol{P}_{i,j}$ for each step angle s_j. Using the angle between the computed planes a first estimate of ρ_s is obtained, and using the intersection of the planes, the galvanometer rotation axis can be estimated while leaving $\delta = 0$ (this will later be refined in Section VI-B3). Then, the original laser plane can be estimated by reflecting back the individually fitted laser planes $\boldsymbol{\pi}_j$ onto their corresponding mirror-galvanometer positions.

The optical angle for each step can be computed by averaging the angular distance between each consecutive pair of planes $\boldsymbol{\pi}_j$ and $\boldsymbol{\pi}_{j+1}$

$$\rho_s = \frac{1}{2(n-1)} \sum_{j=1}^{n-1} \arccos\left(\frac{\boldsymbol{n}_j \cdot \boldsymbol{n}_{j+1}}{\|\boldsymbol{n}_j\|\|\boldsymbol{n}_{j+1}\|}\right). \quad (16)$$

Note that it is divided by two because the mechanical angle is half of the optical angle.

For each pair of planes, the intersection line is also computed and for each line, the point where $p_x = 0$ is chosen and the average amongst these points for all the lines is used as an initialization of the translation part of $^{\{W\}}\boldsymbol{t}_{\{M\}}$ which are given as

$$\boldsymbol{p}_j = \lambda\left(\boldsymbol{n}_j \times \boldsymbol{n}_{j+1}\right) + \frac{\left(d_j \boldsymbol{n}_{j+1} - d_{j+1} \boldsymbol{n}_j\right)\left(\boldsymbol{n}_j \times \boldsymbol{n}_{j+1}\right)}{\|\boldsymbol{n}_j \times \boldsymbol{n}_{j+1}\|^2}$$

$$(17)$$

$$\begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} = \frac{1}{n-1} \sum_{j=1}^{n-1} \boldsymbol{p}_j|_{p_x=0}. \quad (18)$$

Then, the $\overrightarrow{x}$ direction vector $\boldsymbol{x}$ of the mirror reference frame $\{M\}$ can be computed by averaging the direction vector of the intersecting lines [see (19)]. Note that $\boldsymbol{x}$ is referenced to $\{W\}$. Then, the $\overrightarrow{z}$ direction is defined by the vector that is perpendicular to $\boldsymbol{x}$ as well as to the normal of the plane of the first angle $\boldsymbol{n}_1$. The rotation matrix $^{\{W\}}\boldsymbol{R}_{\{M\}}$ can be generated using these two vectors and the vector perpendicular to both of them as shown in the following:

$$\boldsymbol{x} = (1)/(n-1) \sum_{j=1}^{n-1} \boldsymbol{n}_j \times \boldsymbol{n}_{j+1} \quad (19)$$

$$^{\{W\}}\boldsymbol{R}_{\{M\}} = [\boldsymbol{x} \ (\boldsymbol{x} \times \boldsymbol{n}_1) \times \boldsymbol{x} \ \boldsymbol{x} \times \boldsymbol{n}_1]. \quad (20)$$

Given the plane $\boldsymbol{\pi}_j$, its reflection $\boldsymbol{\pi}_j^-$ with the mirror $\boldsymbol{\pi}_{s_j}$ can be computed by choosing a point along the intersection line [see (17) using $\boldsymbol{\pi}_{s_j}$ instead of $\boldsymbol{\pi}_{j+1}$] and using (1) with the plane normal $\boldsymbol{n}_j$ as the input ray and $\boldsymbol{n}_{s_j}$ as the normal of the reflection surface. Note that in ideal conditions, $\boldsymbol{\pi}_j^-$ should be

the same for all planes $\boldsymbol{\pi}_j$ because each $\boldsymbol{\pi}_j$ plane has its own mirror position s_j. Therefore, $\boldsymbol{\pi}_l$, the laser source plane, can be computed by averaging all $\boldsymbol{\pi}_j^-$.

3) Simplified Model Refinement: Now, a simplified model that uses planes instead of rays can be computed. For each step s_j, the light plane $\boldsymbol{\pi}_j^+$ generated by the simplified model can be computed by reflecting $\boldsymbol{\pi}_l$ on the mirror at the position s_j. For each ray associated with $\boldsymbol{u}_{i,j,k}$ [see (8)], a 3-D point $\boldsymbol{p}_{i,j,k}^+$ can be computed by intersecting such ray with the laser plane $\boldsymbol{\pi}_j^+$. Then, an error can be defined as the distance between this point and the previously computed point $\boldsymbol{p}_{i,j,k}$. Finally, the total addition of these errors can be minimized to refine the estimation of the simplified model parameters as follows:

$$
\begin{aligned}
&\left[\boldsymbol{\pi}_l, \rho_s, \delta, {}^{\{W\}}\boldsymbol{t}_{\{M\}} \right] \\
&= \underset{\left[\boldsymbol{\pi}_l, \rho_s, \delta, {}^{\{W\}}\boldsymbol{t}_{\{M\}} \right]}{\operatorname{argmin}} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{o} \boldsymbol{p}_{i,j,k}^+ - \boldsymbol{p}_{i,j,k}.
\end{aligned}
\tag{21}
$$

4) Laser Pose: The last part of the in-air laser calibration is finding the transformation ${}^{\{W\}}\boldsymbol{t}_{\{L\}}$. This is done by finding the closest point on $\boldsymbol{\pi}_l$ to all the first and last laser rays for each angular position of the mirror-galvanometer.

The sensor is mounted in such a way that both ends of the projected lines are visible in the image (i.e. the laser has a smaller aperture than the camera). Therefore, the light ray that produced each end point of the laser line projection can be computed as follows. For all the s_j mirror positions, the two end points of each line ($\boldsymbol{p}_{i,j,1}$ and last $\boldsymbol{p}_{i,j,o}$) are classified in two different sets $\boldsymbol{P}_{s_j}^{\uparrow} = \bigcup_{i=1}^{m} \boldsymbol{p}_{i,j,1}$ and $\boldsymbol{P}_{s_j}^{\downarrow} = \bigcup_{i=1}^{m} \boldsymbol{p}_{i,j,o}$. For each mirror position s_j, the rays $\boldsymbol{r}_{s_j}^{\uparrow}$ and $\boldsymbol{r}_{s_j}^{\downarrow}$ are fitted using linear regression amongst the set of points $\boldsymbol{P}_{s_j}^{\uparrow}$ and $\boldsymbol{P}_{s_j}^{\downarrow}$ respectively. Then, each ray is reflected onto each corresponding mirror plane $\boldsymbol{\pi}_{s_j}$ to obtain $\boldsymbol{r}_{s_j}^{\uparrow -}$ and $\boldsymbol{r}_{s_j}^{\downarrow -}$. In ideal conditions, all these rays would intersect in a single point which would be the translational $[t_x \ t_y \ t_z]$ part of ${}^{\{W\}}\boldsymbol{t}_{\{L\}}$. However, this does not happen with real data. The plane $\boldsymbol{\pi}_l$ containing the laser focal point was previously found. Therefore, the translational part of ${}^{\{W\}}\boldsymbol{t}_{\{L\}}$ corresponds to the point on $\boldsymbol{\pi}_l$ that minimizes the total error between $[t_x \ t_y \ t_z]$ and each ray [see (22) and (23)]. Note that the minimization is formulated only with two of the three variables $[t_x \ t_y]$ because the third one can be extracted from the plane equation $t_z(\boldsymbol{\pi}_l)$. Finally, the rotation matrix ${}^{\{W\}}\boldsymbol{R}_{\{L\}}$ of the transformation ${}^{\{W\}}\boldsymbol{t}_{\{L\}}$ can be computed by aligning the $\overrightarrow{z}$ direction with the normal of the laser plane $\boldsymbol{n}_l$ and the $\overrightarrow{y}$ direction with the direction vector of the intersection line between the laser plane $\boldsymbol{\pi}_l$ and the mirror plane $\boldsymbol{\pi}_{s_0}$ [see (24)].

$$
\begin{aligned}
\left[t_x, t_y \right] = \underset{\left[t_x, t_y \right]}{\operatorname{argmin}} \sum_{i=1}^{m} \mathrm{d}\left(\left[t_x, t_y, t_z(\boldsymbol{\pi}_l) \right], \boldsymbol{r}_{s_j}^{\uparrow -} \right) \\
+ \sum_{i=1}^{m} \mathrm{d}\left(\left[t_x, t_y, t_z(\boldsymbol{\pi}_l) \right], \boldsymbol{r}_{s_j}^{\downarrow -} \right)
\end{aligned}
\tag{22}
$$

$$
\mathrm{d}(\boldsymbol{p}, \boldsymbol{r}) = |((\boldsymbol{o} - \boldsymbol{p}) \cdot \boldsymbol{v})\boldsymbol{v} + \boldsymbol{o} - \boldsymbol{p}|
\tag{23}
$$

$$
{}^{\{W\}}\boldsymbol{R}_{\{L\}} = \left[\boldsymbol{n}_l \times (\boldsymbol{n}_l \times \boldsymbol{n}_{s_0}) \ \boldsymbol{n}_l \times \boldsymbol{n}_{s_0} \ \boldsymbol{n}_l \right]
\tag{24}
$$

C. Camera Viewport Calibration

Given the in-air pin-hole camera calibration parameters $[f_x, f_y, c_x, c_y]$, the viewport $\boldsymbol{\Omega}_C$, and the transformation $\boldsymbol{t}$ between the reference frame $\{O\}$ and the camera frame $\{C\}$, there is no closed form solution to find the projection of the point $\boldsymbol{p}$ referenced in $\{O\}$ on to the undistorted image. The viewport camera calibration uses a set of different positions ${}^{\{C\}}\boldsymbol{t}_{\{O\}}$ of a pattern with respect to the camera $\boldsymbol{T} = \{{}^{\{C\}}\boldsymbol{t}_{\{O\},1}, {}^{\{C\}}\boldsymbol{t}_{\{O\},2}, \ldots, {}^{\{C\}}\boldsymbol{t}_{\{O\},m}\}$. The pattern consists of a set of object points $\boldsymbol{P} = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$ referenced in the object frame $\{O\}$. For each position ${}^{\{C\}}\boldsymbol{t}_{\{O\},i}$, each point $\boldsymbol{p}_j$ has an associated undistorted pixel $\boldsymbol{u}_{u,i,j}$. Using (23), it is possible to compute the distance between the point $\boldsymbol{p}_j$ and the camera ray $\boldsymbol{r}_{u,2,i,j}$ for a given ${}^{\{C\}}\boldsymbol{t}_{\{O\}}$ and $\boldsymbol{\Omega}_C$. Minimizing the total addition of these distances for all the points and positions, it is possible to estimate the camera viewport position as follows:

$$
\left[\boldsymbol{\Omega}_C, \boldsymbol{T} \right] = \underset{\left[\boldsymbol{\Omega}_C, \boldsymbol{T} \right]}{\operatorname{argmin}} \sum_{i=1}^{m} \sum_{j=1}^{n} \mathrm{d}\left({}^{\{C\}}\boldsymbol{t}_{\{O\},i} \oplus \boldsymbol{p}_j, \boldsymbol{r}_{u,2,i,j} \right).
\tag{25}
$$

D. Laser Viewport Calibration

The laser viewport calibration uses the same type of data as the in-air laser calibration. The laser is projected onto a planar surface at different positions $\boldsymbol{t}_i$ for certain mirror positions s_j producing a sorted list of undistorted pixels $\boldsymbol{\lambda}_{i,j} = \{\boldsymbol{u}_{i,j,1}, \boldsymbol{u}_{i,j,2}, \ldots, \boldsymbol{u}_{i,j,o}\}$. For each view of the pattern, the transformation between the pattern and the camera can be found using (25) forcing $\boldsymbol{\Omega}_C$ to be constant. Then, for each pixel $\boldsymbol{u}_{i,j,k}$, its 3-D point $\boldsymbol{p}_{i,j,k}$ can be computed by intersecting its associated ray $\boldsymbol{r}_{u,2}$, which has been refracted through the viewport, with the projection plane. On the other hand, given an estimate of the laser viewport, the 3-D point with the estimated model $\boldsymbol{p}_{i,j,k}^+$ can be computed using the ray–ray triangulation method (see Section V-A). Therefore, a total error function between all the 3-D points from the estimated model, and all the 3-D points obtained by intersecting camera rays with each respective projection plane can be minimized to estimate $\boldsymbol{\Omega}_L$ as follows:

$$
\boldsymbol{\Omega}_L = \underset{\boldsymbol{\Omega}_L}{\operatorname{argmin}} \sum_{i=1}^{m} \sum_{j=1}^{n} \sum_{k=1}^{o} \boldsymbol{p}_{i,j,k}^+ - \boldsymbol{p}_{i,j,k}.
\tag{26}
$$

E. Elliptical Cones Calibration

Fitting a cone to a set of points requires defining the distance between each point and the cone. Finding this distance using the normal to the cone as a function of the two parameters h and β and the point to which the distance of the cone is computed requires solving a fourth degree polynomial equation. However, the problem can also be formulated as an optimization, minimizing the distance between a parametric point (h, β) on the cone and the given point $\boldsymbol{p}$ which is easier to solve using least squares

$$
\mathrm{d}(\boldsymbol{g}, \boldsymbol{p}) = \underset{h, \beta}{\operatorname{argmin}} \| \boldsymbol{g}(h, \beta, \boldsymbol{t}) - \boldsymbol{p} \|.
\tag{27}
$$

A total error function for a set of points $\boldsymbol{P} = \{\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_n\}$, can be computed, and the cone parameters a, b, and $\boldsymbol{t}$ can be estimated by minimizing the total error function as follows:

$$[a, b, \boldsymbol{t}] = \operatorname*{argmin}_{[a, b, \, \boldsymbol{t}]} \sum_{i=1}^{n} \mathrm{d} \left(\boldsymbol{g}(h, \beta, \boldsymbol{t}), \boldsymbol{p}_i \right). \qquad (28)$$

In the present study, a different elliptical cone is generated for each mirror position s at which the scanner will operate. For each mirror position s, a set of points $\boldsymbol{P}_\alpha$ is sampled along the ray $\boldsymbol{r}_{\alpha,3}$. Then, all the sets $\boldsymbol{P}_\alpha$ are merged into a bigger set $\boldsymbol{P}_s = \bigcup_\alpha \boldsymbol{P}_\alpha$ on which the cone for the mirror position s is fitted using (28).

In Section V-B, the ray–cone intersection was presented. There, it can be seen how this intersection can either be two, one, or zero points. If a ray does not intersect with the cone, it means that the laser light could not have illuminated the pixel through which that ray travels. If (14) has only one intersection, it either means that the ray is tangential to the cone (this includes the ray passing through the origin of the cone), or that one of the parameters of the cone a or b is 0. The first case is not possible with the sensor presented here, while the second will only happen if the laser plane traverses the laser viewport perfectly perpendicular. In any case, there would be no ambiguity. However, in the great majority of times that (14) is used for triangulation, such an equation will have two solutions. Due to the way that the light is refracted through the viewport, and the layout of the sensor, the cone will always be wider in the $\overrightarrow{x}$ direction of its local reference frame than in the $\overrightarrow{y}$ one. In other words, $a > b$. Moreover, the laser-projected line will always be more or less aligned with the world $\overrightarrow{x}$ axis due to the way the model is formulated. Therefore, it is possible to differentiate between the two solutions because during the cone fitting process, for each cone, the sign of the point along the $\overrightarrow{y}$ direction is saved as part of the calibration. However, it might happen that both solutions of the ray–cone intersection are on the same side of the $\overrightarrow{y}$ direction in the cone reference frame. In this case, the point that is valid for triangulation will be the one with $z > 0$ in the world frame $\{W\}$.

VII. Results

This section presents the experimental results. First, the different triangulation techniques from Section V are compared in Section VII-A. Then, the calibration of the laser scanner (see Section VI) is evaluated both in air and underwater in Section VII-B. Finally, the sensor is mounted onto Girona 500 AUV [6] and used to scan a realistic pipe structure emulating an installation typical of an underwater offshore industry in Section VII-C.

A. Ray–Cone Triangulation

With the sensor already calibrated, it is possible to use two different triangulation methods as explained in Section V. In Fig. 7, a cloud triangulated using the cone model, is coloured according to the distance between the point and the cloud gathered in the same scene but triangulated using the ray model.
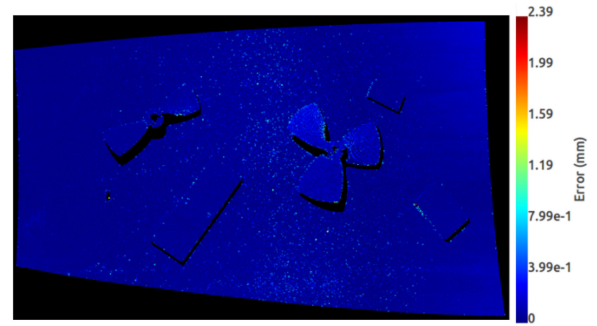

Fig. 7. Scene with propellers and calibration gauges triangulated using the cone model and coloured according to the distance to the cloud gathered on the same scene but triangulated using the ray model.
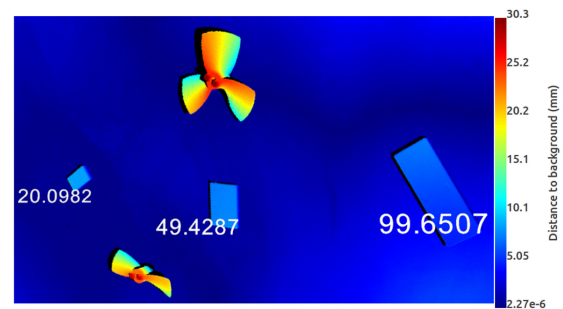

Fig. 8. Partial view of a reconstructed scene with the three ceramic gauges with their corresponding measurements in mm, and two propellers. Points are colored according to the distance between each point and the background.

In the figure, it is possible to observe how the differences are submillimetric, with a mean distance between points of 0.05 mm and a standard deviation of 0.062 mm. The total scanning time, including projecting the laser in all positions and triangulating all the points, was approximately 21 s when using ray triangulation, while the cone method allowed for a much faster scan gathered in approximately 5 s.

B. Three-Dimensional Reconstruction

The performance of the laser scanner is evaluated in three stages. First, the distance between points in the gathered clouds is measured in Section VII-B1. Then, the shape reconstruction of an object is evaluated in Section VII-B2. Finally, the reconstruction of a given plane for different distances is evaluated in Section VII-B3. For all these sections, the triangulation is computed using the cone model for the underwater scans and the standard plane triangulation in air.

1) Measurement Accuracy: Three calibrated ceramic gauges of different sizes (20, 50, and 100 mm) are scanned at different distances from the sensor and in different positions within the scan. Then, the size of the gauges is measured for all the different scans and positions by manually selecting points on the gathered scans.

Fig. 8 presents one of the evaluated scans gathered underwater. In the figure, coloured according to each point's distance
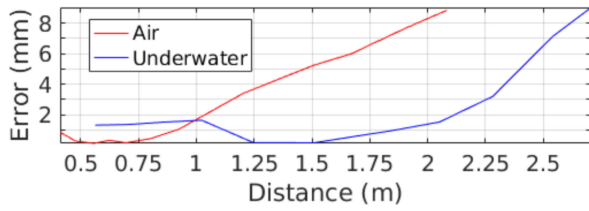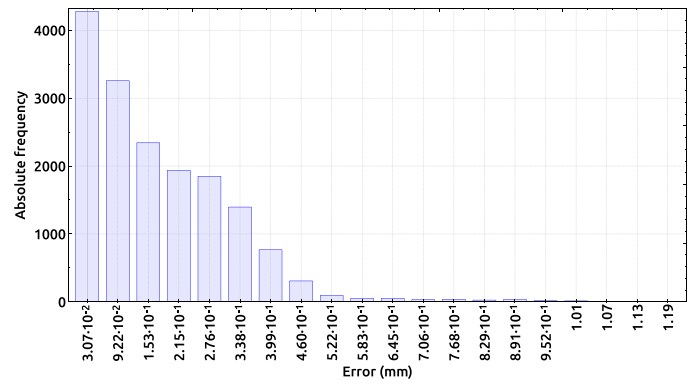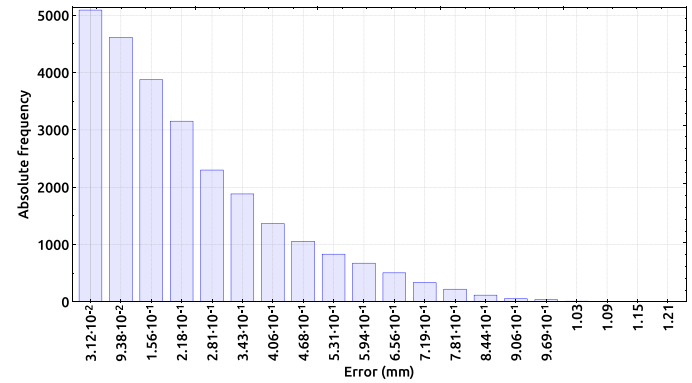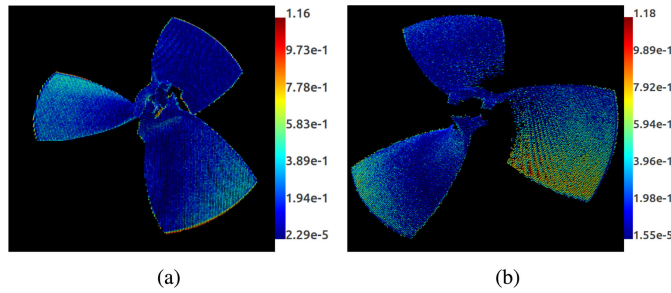
Fig. 9.   Distance–error evolution.

Fig. 10.   Reconstruction results of a propeller of 100 mm diameter in (a) air and (b) underwater. The point cloud coloured according to the distance in mm between each point and the mesh surface.

(a)

(b)

Fig. 11.   Histograms of the errors from Fig. 10 in (a) air and (b) underwater.

to the background for easier display, the three ceramic blocks as well as two other reference objects (propellers) can be seen. There, we can observe that the error for this scan is below one millimeter for all the three gauges. This has been done for a total of 29 scans at distances ranging from 0.5 to 1.2 m both in air and underwater. For each measurement in the scan, the error is evaluated as the absolute value of the difference between the actual measurement and the gauge size. The average error for the 29 scans is 0.44 and 0.98 mm with a standard deviation of 0.35 and 0.72 mm in air and underwater, respectively.

To assess the evolution of the error with the distance to the scanned area, a calibration chessboard pattern of 50×50 mm squares with 9×11 inner corners is scanned at different distances. Since the pattern consists of black and white squares, the laser light is not reflected from the black squares with enough power for the camera to detect the return and compute the 3-D points. Fig. 9 presents the error measured with the maximum number of squares visible in each scan. There we can see that the in air (red) error line is shifted to the left (i.e. a smaller distance) when compared to the underwater one (blue). This is because the effect of the refraction makes the fields of view smaller in water than in air, displacing the measurements of a given resolution or error in air to a larger distance underwater.

2) Overall Shape Reconstruction: A 3-D-printed propeller of 100-mm diameter is scanned at an approximate distance of 65 cm. After manually selecting the points in the scan that belong to the propeller, they are registered using iterative closest point (ICP) against the original CAD model. The results are presented in Figs. 10 and 11 for air and underwater, respectively. Each one of these figures shows the gathered cloud where each point is coloured according to its distance with respect to the CAD model [see Fig. 10(a) and (b)] and the histogram of these distances [see Fig. 11(a) and (b)]. In both cases it can be seen that the vast

TABLE I
AIR STATISTICS FOR PLANARITY IN MM

Distance	Mean	Standard deviation	Maximum
492.4	0.46	0.31	1.56
562.17	0.45	0.33	1.93
660.76	1.12	0.58	2.81
737.76	2.53	0.84	5.79
877.71	1.29	1.02	5.92
1085.69	1.79	1.17	7.35

majority of the error is below one millimeter, the calibration in air (error mean of 0.17 mm and standard deviation of 0.15 mm), being a bit better than underwater (error mean of 0.23 mm and standard deviation of 0.19 mm). Moreover, the maximum error is also lower in air (1.17 mm) than underwater (1.19 mm).

3) Plane Surface Reconstruction: To better assess the performance of the calibration, a flat surface is scanned. The results of reconstructing this surface at different distances, which are computed from the sensor to the centroid of the points belonging to the plane, in air and underwater are presented in Tables I and II, respectively. The tables show how the error, which is computed using the distance between each point and the best fitting plane, of the set of points in air is below the underwater one, and in both cases the mean error grows from values below one millimeter when scanning at short range, to larger values such as 2.53 mm (air) or 4.39 mm (underwater) obtained at longer distances. It is worth noting that the decrement in the

TABLE II
UNDERWATER STATISTICS FOR PLANARITY IN MM

Distance	Mean	Standard deviation	Maximum
381.34	0.37	0.24	1.93
486.84	0.54	0.37	2.97
534.67	0.69	0.48	3.46
619.79	1.04	0.74	3.88
691.82	1.48	1.12	6.71
828.61	4.39	1.87	12.31
831.728	3.77	1.87	11.7
955.315	3.15	2.19	12.79

Fig. 14. Underwater plane fitting error for the plane at 619.79 mm distance.

Fig. 12. Air plane fitting error for the plane at 492.4 mm distance.

Fig. 15. Synthetically generated scan of a flat surface underwater with error on the estimation of different sensor parameters. Points are coloured with the distance to the best fitting plane. (a) Camera viewport normal misalignment. (b) Camera viewport distance error. (c) Laser viewport normal misalignment. (d) Laser viewport distance error.

Fig. 13. Synthetically generated scan of a flat surface in air with error on the estimation of different sensor parameters. Points are coloured with the distance to the best fitting plane. (a) Camera distortion estimation error. (b) Laser z position error.

C. Pipe Structure Scene

mean error that can be observed beyond the 737.76 mm distance in air and the 826.61 mm distance underwater is caused by the plane not occupying the whole scanned area.

In Fig. 12, the in-air plane reconstruction for the distance of 492.4 mm is coloured with each point's distance to the best fitting plane. In this case, it is possible to observe how the error is distributed unevenly across the cloud, indicating some poor estimation of the sensor's parameters. The error distribution does not follow the pattern that a bad estimate of a single parameter of the model would produce. Using the sensor model, it is possible to compute synthetically how different estimation errors on the parameters of the sensor model can lead to distortions on the reconstructed point clouds. The error that can be observed in Fig. 12 is potentially similar to the combination of a bad estimation of the camera distortion parameters [see Fig. 13(a)] with a bad estimate on the position of the laser [see Fig. 13(b)].

In Fig. 14, the underwater plane reconstruction for the distance of 619.79 mm is coloured according to each point's distance to the best fitting plane. In this case, the shape is similar to the one observed with a bad estimation of the camera viewport normal [see Fig. 15(a)] rather than any other errors on the viewports of the system [see Fig. 15(b)–(d)].

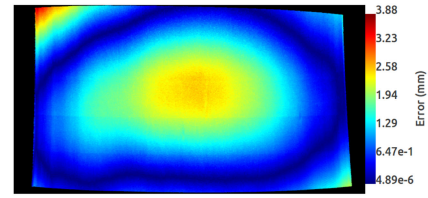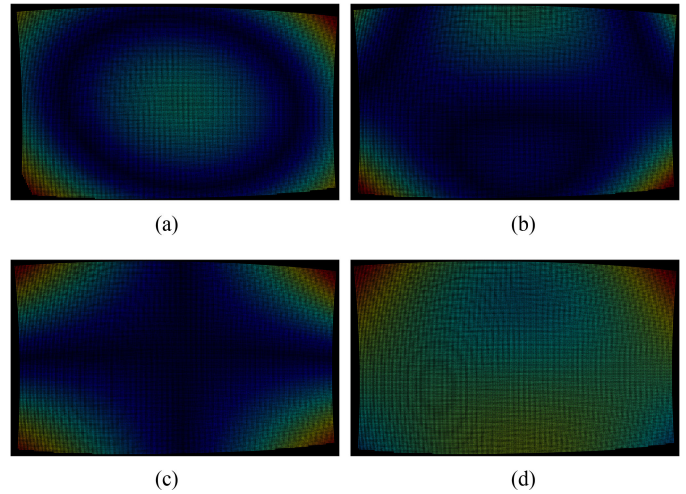To demonstrate the scanning capabilities of real scenes, the sensor was mounted onto Girona 500 AUV and used to scan a pipes-and-valves structure of approximately 145 × 145 × 140 cm, simulating an offshore underwater industry installation [see Fig. 16(a)] at an approximate distance of 1.6 m from the top bar. The point clouds presented in Fig. 16 are gathered with the robot moving from left to right [see Fig. 16(b)–(d)] and with the robot in keep position while free floating in front of the structure [see Fig. 16(e)]. The measured distance between the top left and top right valves in the real structure is 51.5 cm (measured using a measuring tape because this is a hand-made structure) and 51.7 cm in the point cloud gathered with the robot in keep position mode [see Fig. 16(e)]. However, the measured distance while the robot was moving was 43.2 cm [see Fig. 16(c)]. This difference in measurement is a consequence of the combination of scanning direction and robot moving direction. When they are in opposite directions, the object appears larger, while when directions are equal, objects appear smaller. This is confirmed as Fig. 16(c) was gathered with scanning in the same direction as the robot. Finally, when comparing point clouds in Fig. 16(b)–(d) it can be seen how the horizontal pipes appear bent in different directions depending on the point cloud.
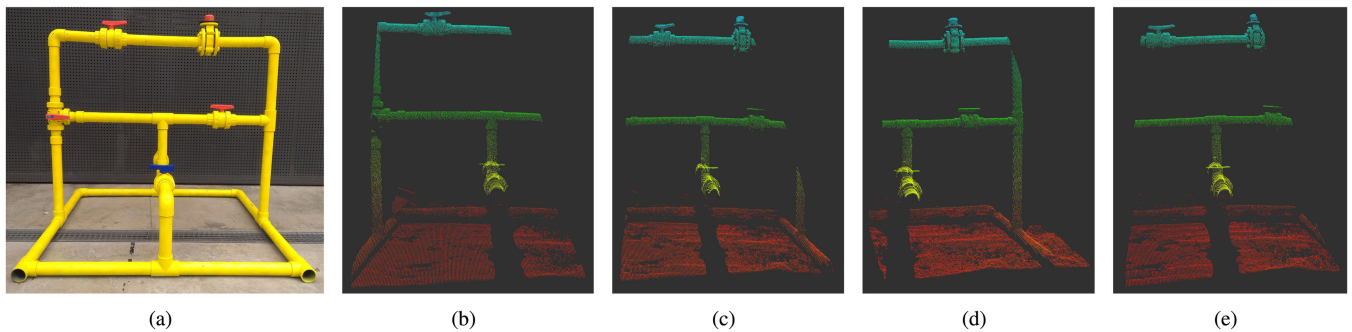
Fig. 16. Scanned pipe structure (a) and different scans coloured according to depth (b)–(e) gathered with the sensor mounted on Girona 500 autonomous underwater vehicles (AUV).
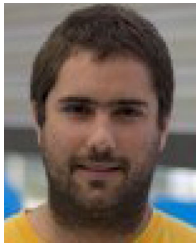
VIII. Conclusion

This study has presented the model and calibration process for an underwater laser scanner. First, a ray-based model for the sensor is presented. The ray-based model consists of two distinct components—light emission and light reception. On the emission side, the path of the light is tracked from the laser focal point to the mirror and through the viewport onto the scene projection. On the receiving side, the path of the ray passing through a specific pixel is traced through the camera viewport to the scene. Following this, the report presents a calibration method for this model in which all the parameters are estimated. The results for the calibration show that the error of the measurements on the point cloud is below one millimeter (measurements in the range 0.5 to 1.2 m and for objects of maximum length 100 mm) in both experiments (point-to-point measurement and point cloud-to-mesh fitting error after ICP alignment). However, the reconstruction presents a deformation when scanning flat surfaces. Finally, the current study also presents a comparison between a ray-based triangulation and an elliptical-cone triangulation. While the error between both gathered clouds is negligible, the speed performance is much better using the elliptical-cone triangulation.

Future work will include testing the calibrated system at different depths, or using finite-elements analysis and sensor simulation, to assess the distortion introduced by the viewport deformation because of the water pressure. Future work will also include improving the sensor's planarity estimation which can be addressed either by changing the calibration procedure or improving the currently presented one. Finally, future work will also include studying the quadratic surfaces generically to represent the underwater laser light and determine whether or not it presents an advantage over the elliptical cone method.
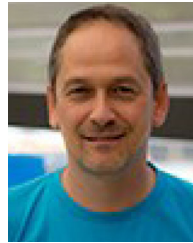
References

[1] E. E. Hitomi, J. V. L. Silva, and G. C. S. Ruppert, "3D scanning using RGBD imaging devices: A survey," in *Developments in Medical Image Processing and Computational Vision*, vol. 19. Cham, Switzerland: Springer, 2015, pp. 379–395.

[2] S. Izadi *et al.*, "KinectFusion: Real-time 3D reconstruction and interaction using a moving depth camera," in *Proc. 24th Annu. ACM User Interface Softw. Technol. Symp.*, 2011, pp. 559–568.

[3] S. Hinterstoisser *et al.*, "Gradient response maps for real-time detection of textureless objects," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 5, pp. 876–888, May 2012.

[4] R. F. Salas-Moreno, R. A. Newcombe, H. Strasdat, P. H. Kelly, and A. J. Davison, "SLAM++: Simultaneous localisation and mapping at the level of objects," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit.*, Portland, OR, USA, 2013, pp. 1352–1359.

[5] F. Flacco, T. Kröger, A. De Luca, and O. Khatib, "A depth space approach to human-robot collision avoidance," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2012, pp. 338–345.

[6] D. Ribas, N. Palomeras, P. Ridao, M. Carreras, and A. Mallios, "Girona 500 AUV: From survey to intervention," *IEEE/ASME Trans. Mechatron.*, vol. 17, no. 1, pp. 46–53, Feb. 2012.

[7] S. Barkby, S. B. Williams, O. Pizarro, and M. Jakuba, "A featureless approach to efficient bathymetric slam using distributed particle mapping," *J. Field Robot.*, vol. 28, no. 1, pp. 19–39, 2011.

[8] A. Palomer, P. Ridao, and D. Ribas, "Multibeam 3D underwater SLAM with probabilistic registration," *Sensors*, vol. 16, no. 560, pp. 1–23, 2016.

[9] S. Williams and I. Mahon, "Simultaneous localisation and mapping on the great barrier reef," *Proc. IEEE Int. Conf. Robot. Autom.*, vol. 2, pp. 1771–1776, 2004.

[10] R. Eustice, H. Singh, J. Leonard, M. Walter, and R. Ballard, "Visually navigating the RMS titanic with SLAM information filters," *Proc. Robot. Sci. Syst.*, 2005. [Online]. Available: http://www.roboticsproceedings.org/rss01/p08.html

[11] R. Campos, N. Gracias, A. Palomer, and P. Ridao, "Global alignment of a multiple-robot photomosaic using opto-acoustic constraints," *IFAC-PapersOnLine*, vol. 48, no. 2, pp. 20–25, 2015.

[12] F. Bruno, G. Bianco, M. Muzzupappa, S. Barone, and A. V. Razionale, "Experimentation of structured light and stereo vision for underwater 3D reconstruction," *ISPRS J. Photogrammetry Remote Sens.*, vol. 66, pp. 508–518, 2011.

[13] 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.

[14] S. T. Digumarti, G. Chaurasia, A. Taneja, R. Siegwart, A. Thomas, and P. Beardsley, "Underwater 3D capture using a low-cost commercial depth camera," in *Proc. IEEE Winter Conf. Appl. Comput. Vis.*, 2016, vol. 1, pp. 1–9.

[15] M. J. Chantler, J. Clark, and M. Umasuthan, "Calibration and operation of an underwater laser triangulation sensor: The varying baseline problem," *Opt. Eng.*, vol. 36, no. 9, pp. 2604–2611, 1997.

[16] G. Inglis, C. Smart, I. Vaughn, and C. Roman, "A pipeline for structured light bathymetric mapping," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst.*, 2012, pp. 4425–4432.

[17] M. Prats, J. J. Fernandez, and P. J. Sanz, "Combining template tracking and laser peak detection for 3D reconstruction and grasping in underwater environments," in *Proc. IEEE Intell. Robots Syst*, 2012, pp. 106–112.

[18] 2G Robotics, Waterloo, ON, Canada. Accessed: Mar. 19, 2018. [Online]. [Available] https://www.2grobotics.com/

[19] Newton Labs, Renton, WA, USA. Accessed: Mar. 19, 2018. [Online]. [Available] http://www.newtonlabs.com/

[20] Savante, Aberdeenshire, U.K. Accessed: Mar. 19, 2018. [Online]. [Available] https://www.savante.co.uk/

[21] D. Kocak, F. Caimi, P. Das, and J. Karson, "A 3-D laser line scanner for outcrop scale studies of seafloor features," in *Proc. IEEE Oceans Conf. Exhib.*, 1999, vol. 3, pp. 1105–1114.

[22] M. Hildebrandt, J. Kerdels, J. Albiez, and F. Kirchner, "A practical underwater 3D-Laserscanner," in *Proc. Oceans*, 2008, pp. 1–5.

[23] A. Palomer, P. Ridao, D. Ribas, and J. Forest, "Underwater 3D laser scanners: The deformation of the plane," in *Lecture Notes in Control and Information Sciences*, vol. 474, T. I. Fossen, K. Y. Pettersen, and H. Nijmeijer, Eds., Cham, Switzerland: Springer, 2017, pp. 73–88.

[24] G. Vosselman and H.-G. Maas, Eds., *Airborne and Terrestrial Laser Scanning*. Dunbeath, Scotland: Whittles Publishing, 2010.

[25] J. Forest Collado, "New methods for triangulation-based shape acquisition using laser scanners," Ph.D. dissertation, Dept. Elect. Inf. Autom., , Universitat de Girona, Girona, Spain, 2004. [Online]. Available: http://www.tdx.cat/handle/10803/7730

[26] G. Bradski, "The OpenCV Library," *Dr Dobbs J. Softw. Tools*, vol. 25, pp. 120–125, 2000.

[27] I. Powell, "Design of a laser beam line expander," *Appl. Opt.*, vol. 26, no. 17, pp. 3705–3709, 1987.

[28] J. Heikkila and O. Silven, "A four-step camera calibration procedure with implicit image correction," in. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 1997, pp. 1106–1112.

Pere Ridao (M'03) received the Ph.D. degree in computer engineering in 2001 from the University of Girona, Girona, Spain.

He is 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 of the University of Girona. His research interests include designing and developing autonomous underwater vehicles for 3-D mapping and intervention.
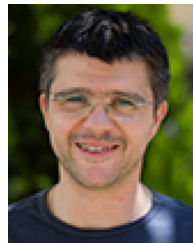
Josep Forest received the B. Sc in industrial informatics from the University of Girona, Girona, Spain in 1992, the M.Sc degree in electronics engineering from the Autonomous University of Barcelona, Barcelona, Spain, in 1998 and the Ph.D degree in 2004 from the University of Girona.

His research interests include 3-D-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.

Albert Palomer received the M.S. degree in industrial engineering from the University of Girona, Girona, Spain, in 2012. At that time, he joined the Underwater Robotics Research Center (CIRS) at the same university. In 2013, he received the M.S. degree in information technologies and automation and Ph.D. degree in 2018 from the University of Girona. He is currently a Researcher at CIRS.

David Ribas received the M.Sc. in industrial engineering in 2003 and obtained the Ph.D. in 2008, both at the University of Girona, Girona, Spain.

He is currently a Researcher with the Department of Computer Engineering at the University of Girona, and a Member of the Underwater Robotics Research Center (CIRS). His research interests include development of AUVs and more particularly the autonomous navigation problem using simultaneous localization and mapping (SLAM) techniques.