Article

Calibration of a Structured Light Imaging System in Two-Layer Flat Refractive Geometry for Underwater Imaging †

Domagoj Zoraja [1,*] , **Tomislav Petković** [1] , **Josep Forest** [2] and **Tomislav Pribanić** [1]

[1] University of Zagreb Faculty of Electrical Engineering and Computing, Unska 3, HR-10000 Zagreb, Croatia; tomislav.petkovic.jr@fer.hr (T.P.); tomislav.pribanic@fer.hr (T.P.)

[2] Computer Vision and Robotics Research Institute (VICOROB), University of Girona, 17003 Girona, Catalunya, Spain; forest@eia.udg.edu

* Correspondence: domagoj.zoraja@fer.hr

† This paper is an extended version of our conference paper published in Zoraja, D.; Petković, T.; Pribanić, T.; Forest, J. Projector Calibration in a Two-Layer Flat Refractive Geometry for Underwater Imaging. In Proceedings of the 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 23–27 May 2022; pp. 930–935.

Abstract: The development of a robust 3D imaging system for underwater applications is a crucial process in underwater imaging where the physical properties of the underwater environment make the implementation of such systems challenging. Calibration is an essential step in the application of such imaging systems and is performed to acquire the parameters of the image formation model and to enable 3D reconstruction. We present a novel calibration method for an underwater 3D imaging system comprising a pair of cameras, of a projector, and of a single glass interface that is shared between cameras and projector(s). The image formation model is based on the axial camera model. The proposed calibration uses a numerical optimization of a 3D cost function to determine all system parameters, thus avoiding the minimization of re-projection errors which require numerically solving a 12th order polynomial equation multiple times for each observed point. We also propose a novel stable approach to estimate the axis of the axial camera model. The proposed calibration was experimentally evaluated on four different glass interfaces, wherein several quantitative results were reported, including the re-projection error. The achieved mean angular error of the system's axis was under $6°$, and the mean absolute errors for the reconstruction of a flat surface were $1.38\,\text{mm}$ for normal glass interfaces and $2.82\,\text{mm}$ for the laminated glass interface, which is more than sufficient for application.

Keywords: calibration; structured light; underwater imaging; flat refractive geometry

1. Introduction

The applications of underwater 3D imaging systems span over various fields of research and have diverse applications, e.g., mapping of the underwater sites of archaeological value [1], monitoring of underwater wildlife in marine biology [2], and underwater inspection [3], to name a few. Recent developments also place such a system onto ROVs (remotely operated vehicles), UAVs (underwater autonomous vehicles), or similar underwater systems [4,5] thus making them perfect for underwater applications where human life might be endangered.

There are many challenges in the design and usage of an underwater 3D imaging system [6]. The difficulty of achieving the desired accuracy [7,8] of the imaging system is caused by image distortions due to the physical properties of light propagation in an underwater environment. To limit these, an underwater enclosure or a watertight housing must be designed accurately for each element of the imaging system, which include a camera and a projector. In the underwater environment, such enclosures introduce

refractions and reflections of light-rays on both the air-to-glass interface and the glass-to-water interface, while water itself introduces attenuation and dispersion of the transmitted and reflected rays. Together, these effects present significant challenges in the calibration of a 3D imaging system. To overcome these challenges, cameras and projectors are usually first calibrated individually to acquire the corresponding intrinsic parameters, followed by a calibration of the whole 3D imaging system to acquire and to refine all parameters of the image formation model [9].

Sensors of a 3D imaging system may be classified as passive or active. The key difference is that active 3D imaging introduces an artificial illumination in the observed scene. A typical example of passive 3D imaging is stereo vision [10], and an example of active 3D imaging is structured light or SL [11,12]. In an SL imaging system, a projector projects a code onto a surface of an observed object, which is then observed by one or more cameras to reconstruct the 3D shape. Therefore, an underwater 3D imaging system comprising one or more cameras and one or more projectors is an active SL 3D imaging device for which there are many reported uses in underwater imaging [1,3,6,13–15]. Next, underwater SL systems may be classified depending on the shape of the air-to-water interface, usually a protective glass, which may be flat or curved [9,16]. Although curved interfaces offer some advantages, including the compensation of unwanted refractions, they are difficult to design and to manufacture, thus making flat interfaces a more convenient and more flexible choice. Finally, when a camera or a projector is coupled with a flat refractive interface, the ubiquitous pinhole camera model [10,17,18] is no longer applicable and must be extended. One possible extension of the pinhole model for flat refractive interfaces is the axial camera model [19,20], which can account for the spatial spreading of the focal point of the pinhole camera model.

In this paper, we present a novel method for calibrating a structured light 3D imaging system, which comprises two cameras and of one projector and which uses a shared flat refractive interface (a single shared glass). The proposed method is a continuation of our previous work [9] on projector calibration for underwater imaging, which we here extend to the calibration of the whole SL imaging system. It is also based on the previous work by Agrawal et al. [19], which introduced flat refractive imaging geometry for a single camera, which we extended to a system of two cameras and a projector. The proposed calibration method uses a planar calibration board with white circles arranged in a hexagonal pattern as a calibration object [21]. Once the calibration object is imaged in several positions, the obtained data is combined with the in-the-air calibration data to obtain the axis of the axial camera first, which is then followed by a numerical optimization procedure using a 3D cost function to determine the final imaging parameters of the SL system. The proposed calibration method was thoroughly evaluated in a laboratory environment for four different glass interfaces. To summarize, compared to the previous works and to the state-of-the-art approaches, our contributions are the following:

1. The coplanarity constraint proposed by Agrawal et al. [19] and used in our previous work [9] was extended to the case of multiple cameras and projectors that share a single flat glass interface and was then applied to the estimation of the axial camera's axis. The proposed axis estimation was more stable than the method of [19] as demonstrated by the performed experiments.
2. The proposed optimization of the 3D cost function from our previous work [9] was extended to the whole SL imaging system, and we also introduced boundaries on allowed system parameters.
3. Calibration using the extended coplanarity constraint and the proposed 3D cost function were thoroughly evaluated on four different glass types, and the obtained results verify that the proposed method can cope with differing glass thicknesses.

Note that, to the best of our knowledge, the proposed method is the first method for calibration of a SL underwater imaging system that uses a single glass interface.

The rest of this paper is structured as follows: Section 2 describes state-of-the-art and related works. In Section 3, we present the proposed calibration method. Experimental results are presented in Section 4 and are discussed in Section 5. We conclude with Section 6.

2. Related Work

In this section, we present prior works considering the in-the-air geometric calibration of multiple cameras and projectors based on the work by Petković et al. [21], which follows with an individual camera (or projector) calibration for the underwater imaging, where a camera is represented using an axial parametric model with an additional calibration of the refractive parameters as proposed by Agrawal et al. [19]. This approach has been extended to projector calibration by Zoraja et al. [9]. In addition, we present prior works where a camera was modeled purely by the pinhole model [17], without any additional imaging parameters to model refractions, and by the Pinax model [22], which allows a fast refraction correction of the flat pane housing by utilizing a pre-computed lookup table.

The geometric calibration of multiple cameras and projectors for in-the-air imaging is an essential step in the everyday use of an SL imaging system [11]. A typical calibration procedure utilizes a planar calibration object that comprises bright circles on a dark background arranged in a regular lattice and follows a flexible method proposed by Zhang and Huang [23] for projector calibration. Such a calibration board makes the projected code easier to decode w.r.t. the standard checkerboard pattern. Petković et al. [21] summarized the whole procedure and also proposes to use a hexagonal calibration lattice, as it increases the number of calibration points, and a multiple phase shift (MPS) structured light code, as it inherently provides subpixel precise mapping w.r.t. the gray code, which requires interpolation for such a level of precision.

A theory of multi-layer flat refractive geometry proposed by Agrawal et al. [19] provides the basis for underwater camera calibration. Their calibration procedure is based on a pinhole camera model, which is extendend to an axial model with additional parameters to model refractions on flat layers. The camera is firstly calibrated in air using Zhang's [24] calibration procedure, followed by the underwater imaging of a planar calibration object to determine additional refractive parameters. Agrawal et al. introduced a multi-layer flat refractive geometry and a plane of refraction (POR), which have been used in our calibration method as well. They model the imaging system as an axial camera [20], since all refracted camera rays when extended to infinity intersect the axis at some point. Here, we stress the fact that the axis of the whole system is not the optical axis of a single camera, although, in a special case, they may coincide. The coplanarity constraint for a single camera yields a linear system, which is solved using singular value decomposition (SVD) to yield an essential matrix from which the axis of the axial system is determined. The pose parameters $\mathbf{R}$ and $\mathbf{t}$ are also computed from the essential matrix, and the layers' thicknesses d_i are computed using the flat refraction constraint (FRC), which states that the segment between the 3D point $\mathbf{P}$ and the point on the last refractive layer $\vec{q}_2$ should be parallel to the final refracted ray $\vec{v}_2$ (see Figure 1). The refinement of all calibration and pose parameters is achieved by minimizing the re-projection error, which requires solving the analytical forward projection (AFP) equation. The AFP describes a method for the analytical computation of the projection of a known 3D point and is computationally expensive, which has prompted further research into possible simplifications, as performed by Kawahara et al. [25]. Our proposed calibration method sidesteps this issue by using a 3D cost function instead of the AFP.

A projector calibration in a two-layer flat refractive geometry presented by Zoraja et al. [9] provides a projector calibration procedure where the projector is modeled as an inverse camera. The projector calibration proposed by Zoraja et al. [9] utilizes the geometric camera calibration procedure proposed by Petković et al. [21] and an additional calibration of the refractive parameters as proposed by Agrawal et al. [19]. Zoraja et al. [9] also discussed the noted instability in the estimation of the axis of an axial system as proposed by Agrawal et al. [19]; this instability has been addressed in our calibration method.
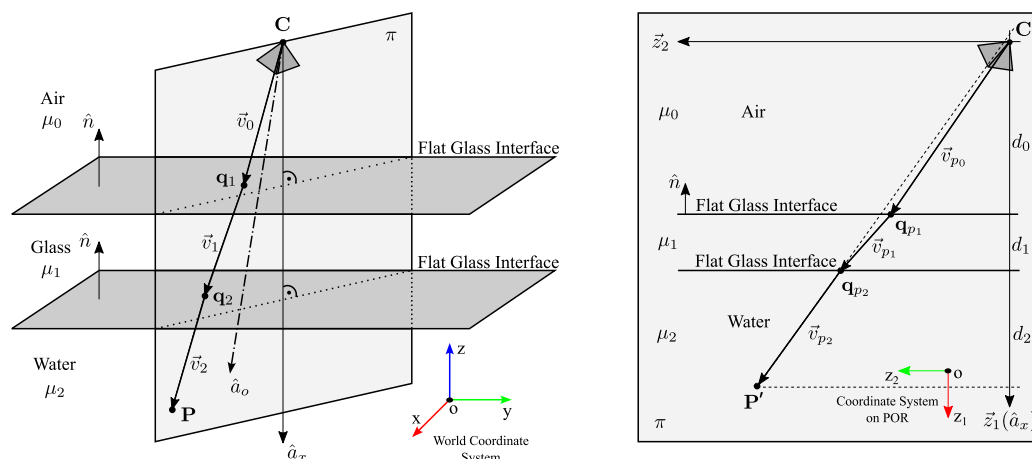
Figure 1. A two-layer flat refractive geometry and a plane of refraction π. **Left**: 3D view. **Right**: plane of refraction.

An underwater 3D reconstruction using structured light was proposed by Bruno et al. [13], which provided experimental results for a highly turbid environment with a heavy presence of scattering and absorption. Their imaging system was experimentally assessed in different turbidity conditions. For the underwater calibration, Bruno et al. [13] used Bouguet's Camera Toolbox [26] and adopted the ubiquitous in-the-air calibration model; thus, they did not consider optical properties of different mediums, including air and water. Such a calibration procedure utilizing only the pinhole camera model with standard distortions, without any additional calibration of the refractive parameters, is simple and convenient to use, but it has a larger error than a proper model, which includes refractions [19]. Here, we note that the necessary requirements [22,27] for the pinhole model to be applicable are that the glass interface should be: (1) perpendicular to the camera's optical axis, (2) be as close to the lens as possible, and (3) be as thin as possible.

Finally, the Pinax model for accurate and efficient refraction correction, presented by Łuczyński et al. [22], allows for a pre-computation of a lookup-table for very fast refraction correction of the flat-pane with high accuracy. The model takes the refraction indices of water into account, especially with respect to salinity, and it is, therefore, sufficient to calibrate the underwater camera only once in-the-air. They require that the optical axis of the camera be perpendicular to the glass surface and that the distance d_0 (see Figure 1) between the glass and the center of projection be minimal; the posed requirements were not considered in our implementation i.e., the direction of camera's (projector's) optical axis and the distance d_0 was considered to be arbitrary.

3. Calibration Method

The proposed calibration method is applicable to a structured light imaging system comprising an arbitrary number of cameras and projectors that all share a single flat interface that separates them from the water. We first provide a description of the imaging geometry used in Section 3.1, which is followed by a description of the coplanarity constraint for an individual camera (or projector) in Section 3.2, which is then extended to an arbitrary number of cameras and projectors in Section 3.3). The in-the-air calibration of individual cameras and projectors and of a whole SL system are described in Section 3.4. Finally, the proposed numerical optimization using a 3D cost function is described in Section 3.5, and the final calibration procedure of a SL system is described in Section 3.6.

3.1. Imaging Geometry

The building blocks of an SL imaging system are cameras and projectors, and, for underwater imaging, we also require a protective flat glass. W.l.o.g., we limit the discussion to a system comprising two cameras and of one projector. The protective interface must be flat and must be shared between all cameras and projectors in the imaging system.

Then, the imaging geometry of the SL imaging system is based on the multi-layer flat refractive geometry introduced by Agrawal et al. [19]. For a submerged imaging system enclosed in a watertight housing using a single protective glass, there are two interfaces between mediums with different optical properties: an air-to-glass interface and a glass-to-water interface. This two-layer flat refractive geometry is presented in Figure 1. On the left, $\vec{v}_i$ represents the light path of a camera/projector ray in 3D (for a two-layer system $i = 0, 1, 2$): $\vec{v}_0$ is the initial camera ray, $\vec{v}_1$ and $\vec{v}_2$ are refracted rays, and q_{i+1} are points of refraction. A point P is the point in the 3D space, and C is the extrinsic center of a pinhole camera/projector model. The axis of the axial camera model $\hat{a}_x$ points in the opposite direction of the refractive interface's normal $\hat{n}$. Note that $\hat{a}_o$ denotes the optical axis of the camera, which is different than the axis $\hat{a}_x$.

A plane of refraction π introduced by [19] is the most important concept in the multi-layer flat refractive geometry, as it contains the entire light path of a camera ray for each pixel (see Figure 1, right). When considering the properties of the light propagation, both the incident ray and the surface normal lie on the unique plane π, and when considering Snell's law, the refracted ray must also lie on π; hence, by induction, the entire light path and the point P are also on π [19].

A 2D coordinate system on the POR is defined by a vector $\vec{z}_1$, which points in the direction of the axis $\hat{a}_x$ of the axial system, and a vector $\vec{z}_2$, which is defined for an unique camera ray $\vec{v}_0$ as $\vec{z}_2 = \vec{z}_1 \times (\vec{z}_1 \times \vec{v}_0)$. A typical POR π for an unique pixel is shown in Figure 1: on the right, it contains the axis $\hat{a}_x$, the light path $\vec{v}_{p_i}$ of a camera/projector ray in 2D, points of refraction $q_{p_{i+1}}$, and the point P' denoting the point P in the coordinate system on the POR. Note that $\vec{v}_{p_i}$ (see Figure 1, right) denotes the camera/projector rays $\vec{v}_i$ (see Figure 1, left) in the coordinate system on the POR. Vector d_i represents the layers' thickness; d_0 is the distance between camera's (projector's) extrinsic center to the boundary of the first refractive interface; d_1 is the thickness of the second interface (glass interface); and d_2 is the distance between the boundary of the second refractive interface and the point P'. Note that, for each pixel, there exists a unique POR and that all PORs for all corresponding pixels comprise a pencil of planes through the straight line, which is the axis of the axial camera $\hat{a}_x$ [9,20].

3.2. The Coplanarity Constraint for a Camera/Projector Using a Single Interface

A constraint that binds the point P and the corresponding initial camera ray $\vec{v}_0$ to the same plane (POR) is called the *coplanarity* constraint [19], and it may be expressed in a matrix form as

$$\left([\mathbf{a}]_\times \mathbf{v}_0\right)^T (\mathbf{Sp} + \mathbf{u}) = 0, \tag{1}$$

where $[\mathbf{a}]_\times$ is a skew-symmetric matrix representing $\hat{a}_x$ in the camera coordinate system, $\mathbf{v}_0$ is a 3×1 vector representing the initial camera ray $\vec{v}_0$ in the camera coordinate system, and $\mathbf{p}$ is a 3×1 vector representation of the illuminated point P in the calibration frame. As Equation (1) is expressed in the camera coordinate system, the parameters $\mathbf{S}$ (rotation) and $\mathbf{u}$ (translation) represent the relative pose of the world coordinate system to the camera coordinate system (see Figure 2). The coplanarity constraint that binds a relative pose of the world coordinate system to the camera coordinate system may be re-written as

$$\mathbf{v}_0^T [\mathbf{a}]_\times \mathbf{Sp} + \mathbf{v}_0^T [\mathbf{a}]_\times \mathbf{u} = 0. \tag{2}$$

The product of a skew-symmetric matrix $[\mathbf{a}]_\times$ and a rotation matrix $\mathbf{S}$ is a 3×3 a matrix $\mathbf{E}$, which has the properties of an essential matrix [10]; note that this is not an essential matrix of a stereo vision system. The product of a skew-symmetric matrix $[\mathbf{a}]_\times$ and a translation vector $\mathbf{u}$ is a 3×1 vector $\mathbf{h}$ representing a constraint on the translation vector $\mathbf{u}$. Considering $[\mathbf{a}]_\times \mathbf{S} = \mathbf{E}$ and $[\mathbf{a}]_\times \mathbf{u} = \mathbf{h}$, Equation (2) may be expressed as

$$\mathbf{v}_0^T \mathbf{Ep} + \mathbf{v}_0^T \mathbf{h} = 0, \tag{3}$$

and, by forming a Kronecker's product (denoted as $\otimes$) between points $\mathbf{p}(i)$ in 3D space and their corresponding initial camera rays $\mathbf{v}_0(i)$, we may express Equation (3) as a linear system by stacking equations for their points' correspondences

$$\underbrace{\begin{bmatrix} \mathbf{p}(1)^T \otimes \mathbf{v}_0(1)^T & \mathbf{v}_0(1)^T \\ \mathbf{p}(2)^T \otimes \mathbf{v}_0(2)^T & \mathbf{v}_0(2)^T \\ \vdots & \vdots \\ \mathbf{p}(N)^T \otimes \mathbf{v}_0(N)^T & \mathbf{v}_0(N)^T \end{bmatrix}}_{\mathbf{Q}} \begin{bmatrix} \mathbf{E}(:) \\ \mathbf{h} \end{bmatrix} = \mathbf{0}. \tag{4}$$

We denote a $N \times 9$ matrix of the linear system with $\mathbf{Q}$. N is the number of points' correspondences. Vector $\mathbf{E}(:)$ is a column vector formed by stacking columns of the essential matrix $\mathbf{E}$.
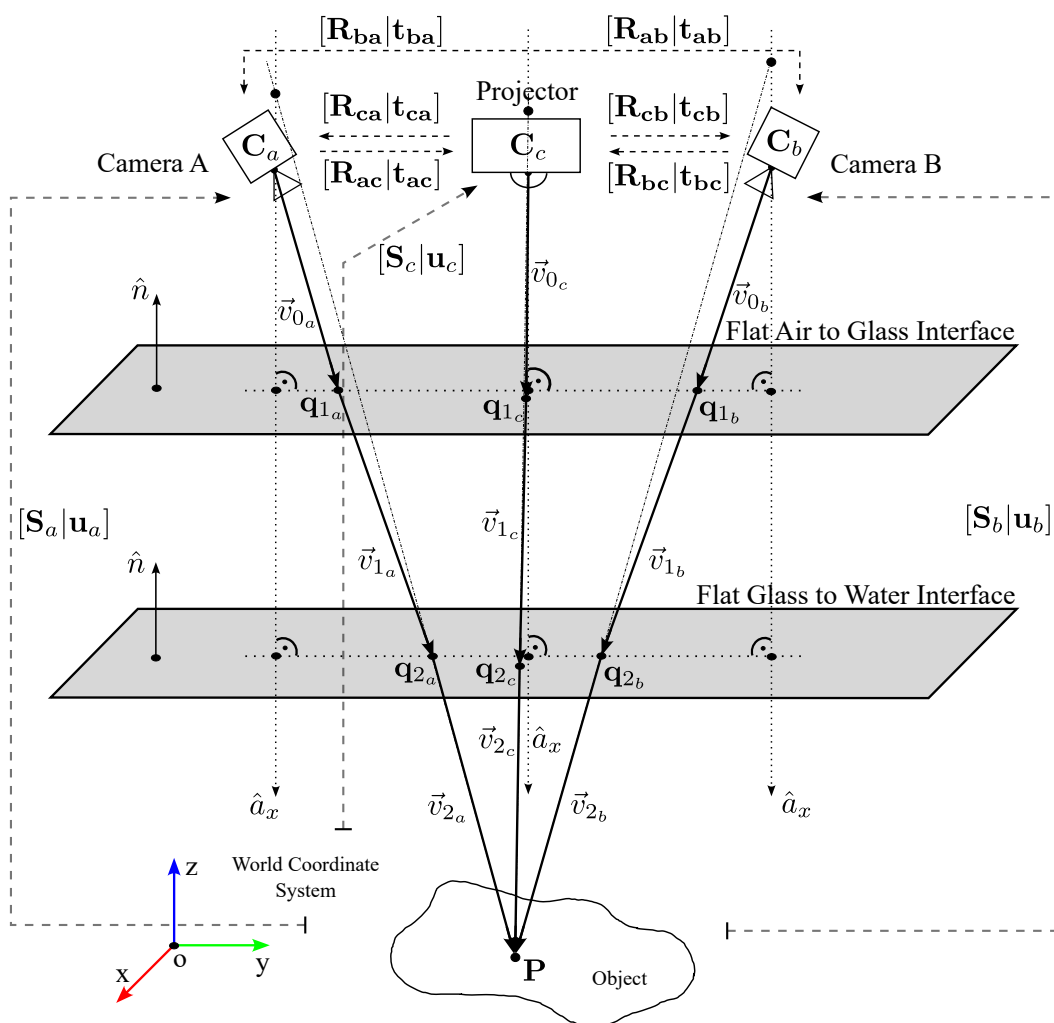

Figure 2. Relative poses of cameras and a projector w.r.t. the world coordinate system.

The solution of the linear system, given by Equation (4), is acquired by performing a singular value decomposition (SVD) of the matrix $\mathbf{Q}$. A planar calibration object is commonly represented as the xy plane with the z coordinate set to zero [24], so $\mathbf{p} = [x_i, y_i, 0]^T$ and columns 7, 8, and 9 all reduce to zero. Therefore, we cannot directly estimate the full matrix $\mathbf{E}$; we may only estimate the first two columns of $\mathbf{E}$. The last column of $\mathbf{E}$ is estimated using Demazure's [28] constraints (also see Nister [29]). The axis $\mathbf{a}$ of the axial system, a 3×1 vector, may be extracted from the acquired $\mathbf{E}$ matrix by performing an SVD. Note that $\mathbf{h}^T \mathbf{a} = 0$; hence, the full translation cannot be extracted, as the component of $\mathbf{u}$ in

the direction of the axis vanishes in **h**. By using the acquired solution for **h**, we may only estimate a part of translation orthogonal to the axis $\mathbf{u}_{\text{orth}}$ as a cross product of **h** and **a**. Note that the full translation may be estimated from a flat refraction constraint (FRC) defined on the POR for each point correspondence; see Agrawal et al. [19]. Alternatively, a central approximation may be used to obtain an approximate estimate. In a central approximation, 3D points in the world frame are mapped to 2D points in the image frame by a central projection model in which a ray defined by a 3D point in space and a fixed point in space, the center of projection, intersects a specific plane in space that is chosen as the image plane (see Hartley and Zisserman [10]). In our calibration procedure, the rotation and translation estimates were computed using central approximation, which is simpler than solving the FRC and which yields appropriate initial points for the numerical optimization.

The coplanarity constraint may be derived separately for each pose of the camera/projector w.r.t. the calibration object. In the case of multiple cameras/projectors, we set the coordinate system of the camera **a** as the referent one; hence, the coplanarity constraint may be re-written w.r.t. to this referent frame as

$$\mathbf{v}_{0_a}^T \left[[\mathbf{a_1}]_{\times} \mathbf{S_a p_a} + [\mathbf{a_1}]_{\times} \mathbf{u_a} \right] = 0, \tag{5}$$

$$\mathbf{v}_{0_b}^T \left[[\mathbf{R_{ab} a_1}]_{\times} (\mathbf{R_{ab} S_a p_b} + \mathbf{R_{ab} u_a} + \mathbf{t_{ab}}) \right] = 0, \tag{6}$$

$$\mathbf{v}_{0_c}^T \left[[\mathbf{R_{ac} a_1}]_{\times} (\mathbf{R_{ac} S_a p_c} + \mathbf{R_{ac} u_a} + \mathbf{t_{ac}}) \right] = 0, \tag{7}$$

where Equation (5) represents the coplanarity constraint of the camera **a**, Equation (6) is the coplanarity constraint of the camera **b** in the frame **a**, and Equation (7) denotes the coplanarity constraint of the projector in the frame **a** (see Figure 2). Each constraint yields a certain pose of the camera (projector) w.r.t. the calibration object: the pose $[\mathbf{S_a}|\mathbf{u_a}]$ for the correspondence between the object's frame and the frame **a**, the pose $[\mathbf{S_b}|\mathbf{u_b}]$ for the correspondence between the object's frame and the frame **b**, and the pose $[\mathbf{S_c}|\mathbf{u_c}]$ for the correspondence between the object's frame and the frame **c**.

The rotation matrix $\mathbf{R_{ab}} = \mathbf{R_b R_a}^T$ is the rotation from the camera coordinate system **a** to the frame **b**, and $\mathbf{R_{ac}} = \mathbf{R_c R_a}^T$ is the rotation from the camera coordinate system **a** to the projector's frame (frame **c**). The corresponding translations are $\mathbf{t_{ab}} = -\mathbf{R_b C_b} + \mathbf{R_b C_a}$ and $\mathbf{t_{ac}} = -\mathbf{R_c C_c} + \mathbf{R_c C_a}$, where $\mathbf{C}_i$ denotes the extrinsic center of the corresponding camera (projector). These rotations and translations are determined during the in-the-air calibration. The axis $\mathbf{a_1}$ of an axial system is the system's axis expressed in the referent frame (frame **a**). In Equations (6) and (7), the axes $\mathbf{a_2}$ and $\mathbf{a_3}$ are rotated into the referent frame, since $\mathbf{a_2} = \mathbf{R_{ab} a_1}$ and $\mathbf{a_3} = \mathbf{R_{ac} a_1}$.

The rotation matrices $\mathbf{S_b}$ and $\mathbf{S_c}$ w.r.t. the frame **a** are $\mathbf{S_b} = \mathbf{R_{ab} S_a}$ and $\mathbf{S_c} = \mathbf{R_{ac} S_a}$. The corresponding translations w.r.t. the frame **a** are $\mathbf{u_b} = \mathbf{R_{ab} u_a} + \mathbf{t_{ab}}$ and $\mathbf{u_c} = \mathbf{R_{ac} u_a} + \mathbf{t_{ac}}$. The $3 \times N_i$ matrices of initial camera rays are denoted with $\mathbf{v}_{0_a}$, $\mathbf{v}_{0_b}$, and $\mathbf{v}_{0_c}$, and they are expressed in the corresponding camera/projector frame (see Figure 2). The corresponding $\mathbf{p_i}$ denotes $3 \times N_i$ matrices of the object's points in the world coordinate system (the object's frame). Note that the coplanarity constraint is performed for each pose of the calibration object. For a complete derivation of the coplanarity constraint, see Appendix A.1.

3.3. Unified Coplanarity Constraint for a System Using a Single Interface

Following up on Section 3.2, we proposed to calibrate the system by using an unified coplanarity constraint. We formed the coplanarity constraint equations separately for each pose of the camera/projector w.r.t. the calibration board (see Equations (8)–(10)), and we

combined them into a single linear equation for the whole system (see Equation (14)). Considering $[a]_\times \mathbf{R} = \mathbf{E}$ and $[a]_\times \mathbf{t} = \mathbf{h}$, Equations (5)–(7) may be simplified as

$$\mathbf{v}_{0_a}^T \mathbf{E_1 p_a} + \quad \mathbf{v}_{0_a}^T \mathbf{h_1} \qquad\qquad = 0 \qquad\qquad (8)$$

$$\mathbf{v}_{0_b}^T \mathbf{R_{ab}} \mathbf{E_1 p_b} + \mathbf{v}_{0_b}^T \mathbf{R_{ab}} \mathbf{h_1} + \mathbf{v}_{0_b}^T [\mathbf{R_{ab} a_1}]_\times \mathbf{t_{ab}} = 0 \qquad (9)$$

$$\mathbf{v}_{0_c}^T \mathbf{R_{ac}} \mathbf{E_1 p_c} + \mathbf{v}_{0_c}^T \mathbf{R_{ac}} \mathbf{h_1} + \mathbf{v}_{0_c}^T [\mathbf{R_{ac} a_1}]_\times \mathbf{t_{ac}} = 0, \qquad (10)$$

where the essential matrix $\mathbf{E_1}$, the parameter $\mathbf{h_1}$, and the axis $\mathbf{a_1}$ denote the unknown parameters in the frame $\mathbf{a}$ (the referent frame). Furthermore, we may refactor the term $\mathbf{v}_{0_b}^T [\mathbf{R_{ab} a_1}]_\times \mathbf{t_{ab}}$ in Equation (9) as $\left[\mathbf{R_{ab}}^T \left(\mathbf{t_{ab}} \times \mathbf{v}_{0_b}\right)\right]^T \mathbf{a_1}$ and the term $\mathbf{v}_{0_c}^T [\mathbf{R_{ac} a_1}]_\times \mathbf{t_{ac}}$ in Equation (10) as $\left[\mathbf{R_{ac}}^T \left(\mathbf{t_{ac}} \times \mathbf{v}_{0_c}\right)\right]^T \mathbf{a_1}$, thus forming a linear system

$$\mathbf{v}_{0_a}^T \mathbf{E_1 p_a} + \quad \mathbf{v}_{0_a}^T \mathbf{h_1} \qquad\qquad\qquad = 0 \qquad (11)$$

$$\mathbf{v}_{0_b}^T \mathbf{R_{ab}} \mathbf{E_1 p_b} + \mathbf{v}_{0_b}^T \mathbf{R_{ab}} \mathbf{h_1} + \left[\mathbf{R_{ab}}^T \left(\mathbf{t_{ab}} \times \mathbf{v}_{0_b}\right)\right]^T \mathbf{a_1} = 0 \qquad (12)$$

$$\mathbf{v}_{0_c}^T \mathbf{R_{ac}} \mathbf{E_1 p_c} + \mathbf{v}_{0_c}^T \mathbf{R_{ac}} \mathbf{h_1} + \left[\mathbf{R_{ac}}^T \left(\mathbf{t_{ac}} \times \mathbf{v}_{0_c}\right)\right]^T \mathbf{a_1} = 0, \qquad (13)$$

which may be refactored further w.r.t. Equation (3), by using Kronecker's product, and represented in the matrix form:

$$\begin{bmatrix} \mathbf{p_a(1)^T} \otimes \mathbf{v}_{0_a}(1)^T & \mathbf{v}_{0_a}(1)^T & \mathbf{0} \\ \mathbf{p_a(2)^T} \otimes \mathbf{v}_{0_a}(2)^T & \mathbf{v}_{0_a}(2)^T & \mathbf{0} \\ \vdots & \vdots & \vdots \\ \mathbf{p_b(1)^T} \otimes \left(\mathbf{v}_{0_b}(1)^T \mathbf{R_{ab}}\right) & \mathbf{v}_{0_b}(1)^T \mathbf{R_{ab}} & \left[\mathbf{R_{ab}}^T \left(\mathbf{t_{ab}} \times \mathbf{v}_{0_b}(1)\right)\right]^T \\ \mathbf{p_b(2)^T} \otimes \left(\mathbf{v}_{0_b}(2)^T \mathbf{R_{ab}}\right) & \mathbf{v}_{0_b}(2)^T \mathbf{R_{ab}} & \left[\mathbf{R_{ab}}^T \left(\mathbf{t_{ab}} \times \mathbf{v}_{0_b}(2)\right)\right]^T \\ \vdots & \vdots & \vdots \\ \mathbf{p_c(1)^T} \otimes \left(\mathbf{v}_{0_c}(1)^T \mathbf{R_{ac}}\right) & \mathbf{v}_{0_c}(1)^T \mathbf{R_{ac}} & \left[\mathbf{R_{ac}}^T \left(\mathbf{t_{ac}} \times \mathbf{v}_{0_c}(1)\right)\right]^T \\ \mathbf{p_c(2)^T} \otimes \left(\mathbf{v}_{0_c}(2)^T \mathbf{R_{ac}}\right) & \mathbf{v}_{0_c}(2)^T \mathbf{R_{ac}} & \left[\mathbf{R_{ac}}^T \left(\mathbf{t_{ac}} \times \mathbf{v}_{0_c}(2)\right)\right]^T \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \mathbf{E_1(:)} \\ \mathbf{h_1} \\ \mathbf{a_1} \end{bmatrix} = \mathbf{0} \qquad (14)$$

Equation (14) is the unified coplanarity constraint for a system using a single interface w.r.t. the coordinate system of the camera $\mathbf{a}$. Since we are using a planar calibration object, only the first two columns of the essential matrix $\mathbf{E_1}$ may be estimated using SVD; the third column is estimated using the Demazure's [28] constraints (also see Nister [29]). Note that the key feature in the proposed unified coplanarity constraint is the estimation of the system axis $\hat{a}_x$ (expressed in the referent frame as axis $\mathbf{a_1}$) derived directly from the posed linear system (Equation 14). The estimated parameter $\mathbf{h_1}$ is not used, since we compute the initial pose parameters directly from the central approximation. Note that the unified coplanarity constraint is performed for each pose of the calibration object, thereby implying that we acquire as many solutions as there are positions of the imaged calibration object. The system's axis is then acquired as a circular mean of the estimated axes. For a full derivation of the unified coplanarity constraint, see Appendix A.2.

3.4. In-the-Air Calibration

In-the-air geometric calibration is required in order to determine the intrinsic parameters of camera/projector and the relative poses between cameras and projectors [21]. This geometric calibration is a necessary prerequisite for the proposed calibration of the underwater imaging system and is performed in three steps:

1. Create an image of a planar calibration board in many positions;

2. Extract the image coordinates of the calibration points;
3. Optimize all the parameters by minimizing the re-projection error.

Regarding projector calibration, the projector is not able to acquire images of the observed illuminated scene. The necessary data is instead acquired by cameras in the imaging system [21,23,30]. Since the projector is modeled as an inverse camera, the standard pinhole camera model with radial distortion is used. The state-of-the-art projector calibration algorithms follow the procedure proposed by Zhang and Huang [23] using a structured light scanning technique to obtain calibration data [11,12]. The projector's row and column coordinates are embedded in the projected coded pattern, which yields direct correspondences between an uncalibrated camera and the projector [21].

Considering the calibration object, we used a planar calibration board with a regular hexagonal lattice of bright circles on a dark background, which is shown in Figure 3. The projected SL code is easier to decode for white circles when compared to the standard checkerboard pattern [30]. The calibration board should be imaged in many different positions w.r.t. the camera/projector position [21]. For each camera and projector, two different positions of the calibration board constitute the minimum [24]. We proposed to use MPS-coded patterns, which enable simultaneous acquisition in a system comprising of an arbitrary number of projectors [31,32]. Another key point of the proposed approach is that, if there are multiple cameras observing the same part of the calibration board illuminated by a projector, then the number of extracted points for calibrating that particular projector is increased as measurements are repeated up to the number of cameras observing the illuminated board [21].
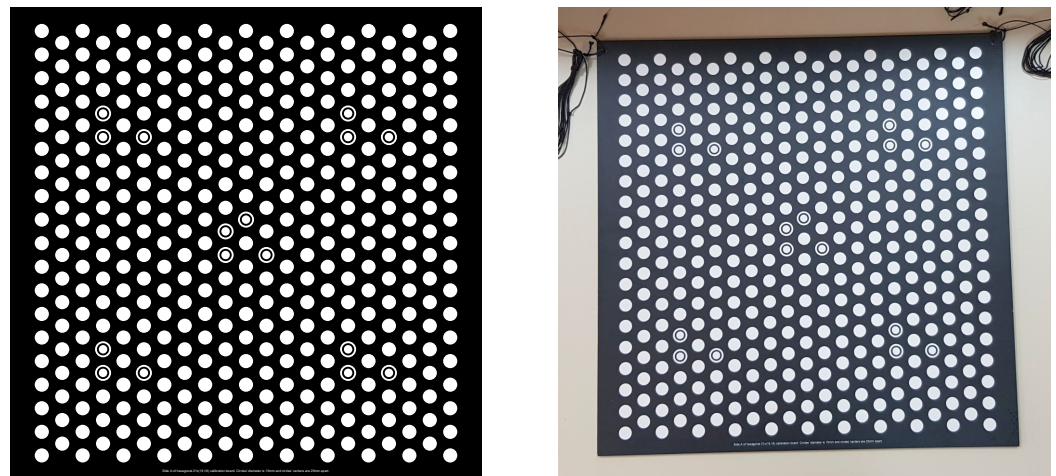
Figure 3. Planar calibration board (**left**: ideal calibration pattern; **right**: physical board used).

3.5. Numerical Optimization

As proposed in Section 3.3, we did not estimate the pose parameters $[\mathbf{S}|\mathbf{u}]$ from the coplanarity constraint (1); instead, we utilized a central approximation algorithm to acquire the initial pose of the camera(s) and projector(s). To refine the acquired pose parameters, we designed a numerical optimization algorithm, which follows our previous implementation [9], which was extended to include parameter boundaries. Instead of minimizing the re-projection error, we minimized the 3D cost function of [9], which we describe here for completeness. This avoided the minimization of the re-projection error, which introduces high computational requirements of an analytical forward projection algorithm for which a 12th degree polynomial must be solved for each calibration point and for each iteration of the minimization [19,25].

The total error in 3D comprises three components: (1) coplanarity error; (2) frustum error; and (3) backprojection error. Posed errors may be used to form an objective function in a nonlinear least-squares minimization algorithm.

3.5.1. Coplanarity Error

The coplanarity constraint (1) limits the position of a point $\vec{p}$ and its corresponding ray $\vec{v}_0$ to an unique plane (the POR, see Figure 1). The shortest vector $\vec{r}_{\text{CPL}}$ connecting the point $\vec{p}$ to the plane is

$$\vec{r}_{\text{CPL}} = (\delta - \hat{n} \cdot \vec{p}) \, \hat{n}, \tag{15}$$

where a plane in the Hesse normal form is described by its normal $\hat{n}$ and its signed distance to the origin δ. It is implied that a posed constraint must be zero for all points and their corresponding rays; hence, the sum of squares of two-norms is

$$\sum_i ||\vec{r}_{\text{CPL},i}||_2^2 \tag{16}$$

for all ray-point correspondences $(\vec{v}_{0,i}, \vec{p}_i)$ must be minimized in order to find the local minimum.

3.5.2. Frustum Error

Since all calibration points must be in a field of view (FoV), we proposed a frustum error [9] as the distance to the projector's clipping planes (see Figure 4). Frustum is defined by six clipping planes: front, back, top, bottom, left, and right. We may apply a constraint by considering the distance to a clipping plane, with the purpose to guide the minimization by constraining the pose parameters. The equations of all six clipping planes are easily expressed in the Hesse normal form from the known in-the-air calibration. If the normals of all the clipping planes are oriented to point inside the frustum, then the vector $\vec{r}_{\text{FRS}}$ connecting the point $\vec{p}$, which is outside of the frustum to the clipping plane, $(\hat{n}_j, \delta_j)$ is

$$\vec{r}_{\text{FRS}} = \begin{cases} (\delta_j - \hat{n}_j \cdot \vec{p})\hat{n}_j, & \hat{n}_j \cdot \vec{p} - \delta_j > 0 \\ 0, & \text{else.} \end{cases} \tag{17}$$

Equation (17) implies that, if the point is on the correct side of the clipping plane (inside the frustum), then we set $\vec{r}_{\text{FRS}}$ to zero. All calibration points that end up outside of or on the boundaries of the frustum are fed to the numerical optimization algorithm. The sum of squares of two-norms is

$$\sum_i ||\vec{r}_{\text{FRS},i}||_2^2, \tag{18}$$

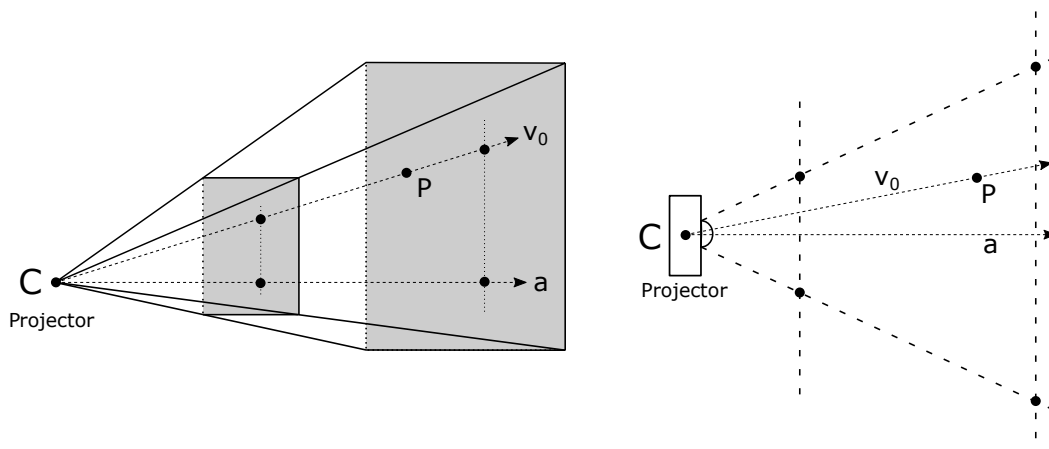which defines the objective function in a minimization scheme.

Figure 4. Projector's clipping planes (**left**: 3D view; **right**: top 2D view).

3.5.3. Backprojection Error

From the known imaging geometry, we may analytically derive the equations of refracted rays in the water [19]. We are only interested in the part of the last ray in the

water, which is inside the projector's frustum. This ray segment is best modeled via its two endpoints, which are denoted as $\vec{q}_2$ and $\vec{q}_3$ (see Figures 1 and 5):

$$\vec{q}_3 = \vec{q}_2 + \ell\hat{v}_2, \tag{19}$$

where ℓ represents the distance between $\vec{q}_2$ and $\vec{q}_3$, which is defined by the projector's frustum. The point $\vec{q}_2$ denotes the point of refraction on the second refractive layer, and $\vec{q}_3$ is the point on the camera ray in the 3rd medium (water). The distance of a point $\vec{p}_i$ to the line segment connecting $\vec{q}_2$ to $\vec{q}_3$ is defined as the backprojection error. Expressing the line in a parametric form as $\vec{q}_2 + t\vec{q}_3$ enables a simple computation of t, which defines the closest point as

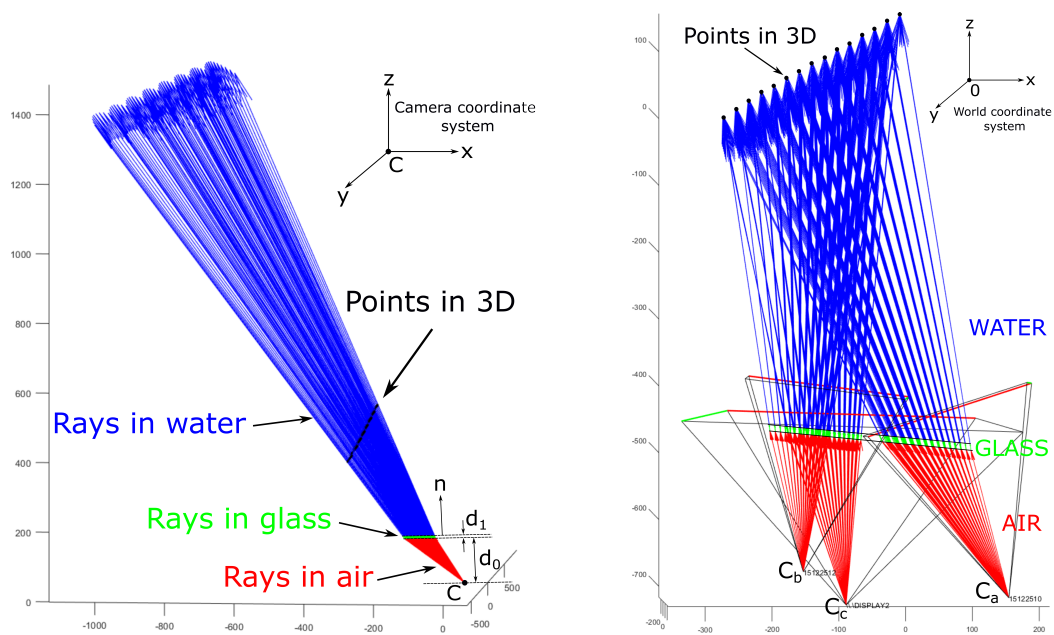$$t = \frac{(\vec{p}_i - \vec{q}_2) \cdot (\vec{q}_3 - \vec{q}_2)}{||\vec{q}_3||}. \tag{20}$$

The shortest vector $\vec{r}_{\text{BPR}}$ connecting a point $\vec{p}$ to the line segment is then

$$\vec{r}_{\text{BPR}} = \begin{cases} \vec{q}_2 - \vec{p}, & t \leq 0 \\ \vec{q}_2 + t\vec{q}_3 - \vec{p}_i, & 0 < t < 1 \\ \vec{q}_3 - \vec{p}, & 1 \leq t. \end{cases} \tag{21}$$

Hence, the sum of squares of two-norms

$$\sum_i ||\vec{r}_{\text{BPR},i}||_2^2 \tag{22}$$

for all ray-points correspondences $(\vec{v}_{0,i}, \vec{p}_i)$ must be minimized.

Figure 5. Initial and refracted camera rays. **Left**: for a single camera in the corresponding camera coordinate system. **Right**: for a system in the world coordinate system. Initial camera rays in the air are displayed in red, rays refracted in the glass are displayed in green, and final refracted rays in the water are displayed in blue.

3.5.4. Total Error

If the cameras and the projector are calibrated in the air, implying that we know the corresponding intrinsic parameters for each camera (projector) expressed in the same physical units, then all three vectors given by Equations (15) (17) and (21) are measured using the same physical units, and they may be joined together in an objective function for

the optimization algorithm. Therefore, the total error, which defines the objective function to minimize, is

$$\sum_i ||\vec{r}_{\text{CPL},i}||_2^2 + ||\vec{r}_{\text{FRS},i}||_2^2 + ||\vec{r}_{\text{BPR},i}||_2^2. \tag{23}$$

The value of (23) is minimized over the space of parameters, which includes the system axis $\hat{a}_x$, the pose parameters (a 3×3 rotation matrix **S** and a 3×1 translation vector **u**), as well as distances d_0 and d_1. For a nonlinear least-squares minimization scheme, we utilized the Levenberg–Marquardt optimization algorithm (i.e., Matlab's `lsqnonlin` function [33]).

3.5.5. Boundary Constraints

We proposed additional tweaks on the minimization of the objective function, i.e., we imposed boundary constraints on the optimization algorithm. A nonlinear least-squares fitting problem optionally may be bounded by upper and/or lower bounds. The components of an initial value x_0 that violate the bounds

$$\text{LB} \leq x \leq \text{UB}, \tag{24}$$

where x is an output, LB is a lower bound, and UB is an upper bound, are reset to the interior of the box defined by the bounds. Components that satisfy the boundary constraint are not changed. We set boundaries for the distance d_0 between the camera's (projector's) extrinsic center and the glass interface, and, for thickness d_1 of the glass interface, the ranges of these two parameters are easily estimated (see Figure 1).

3.6. SL System Calibration

The purpose of geometric calibration is to determine all parameters that describe the image formation model of a particular SL system. Our proposed calibration method required two rounds of data acquisition: the first one was performed in-the-air without the shared protective glass, and the second one was performed underwater, with both using the same calibration object. The proposed geometric calibration of an underwater SL imaging system comprising an arbitrary number of cameras and projectors is performed as follows:

1. Calibrate all cameras and projectors in-the-air using a standard pinhole model with distortions as described in Section 3.4.
2. Acquire as many images of the calibration board in the water as is practical and process the data using the procedure of [21] to extract the calibration data.
3. Estimate the axis using Equation (14). In addition, for each position of the calibration board, estimate the initial pose of the calibration board w.r.t. the camera/projector frame using central approximation.
4. Use a numerical optimization with the objective function comprising the coplanarity error and of the frustum error (see Section 3.5) to estimate true relative poses and to refine the axis; this is performed separately for each position of the calibration board.
5. Use the numerical optimization with the complete objective function comprising the backprojection, the coplanarity, and the frustum errors to refine all parameters (see Section 3.5).

4. Evaluation

Evaluation of the proposed calibration method was achieved by performing several quantitative measurements. First, we computed the angular error of the estimated axis of the system; second, we computed the error in 3D (23); third, we computed the re-projection error; and, fourth, we computed the distance error for a plane fitted on the reconstructed calibration board. The errors are reported as a mean for each position of the calibration board w.r.t. the different glass interface. For the comparisons with the state-of-the-art techniques, we used the method proposed by Agrawal et al. [19] as the baseline.

4.1. Laboratory Setup

The proposed method was evaluated on a laboratory setup comprising a water tank with a 300 L capacity filled with fresh water, a projector, and two cameras mounted on a fixed rig (see Figure 6). The water tank comprised four glass interfaces, with each interface having different properties. We labeled the interfaces with indices from I to IV: the interface I was the **8** mm thick glass; the interface II was the **10** mm thick glass; the interface III was the **10** mm thick laminated glass; and the interface IV was the **12** mm thick glass. The two cameras used were PointGrey's Grasshopper 3 GS3-U3-23S6C-C cameras equipped with Fujinon HF12.5SA-1 lenses, and the projector used was an Acer S1383WHne.
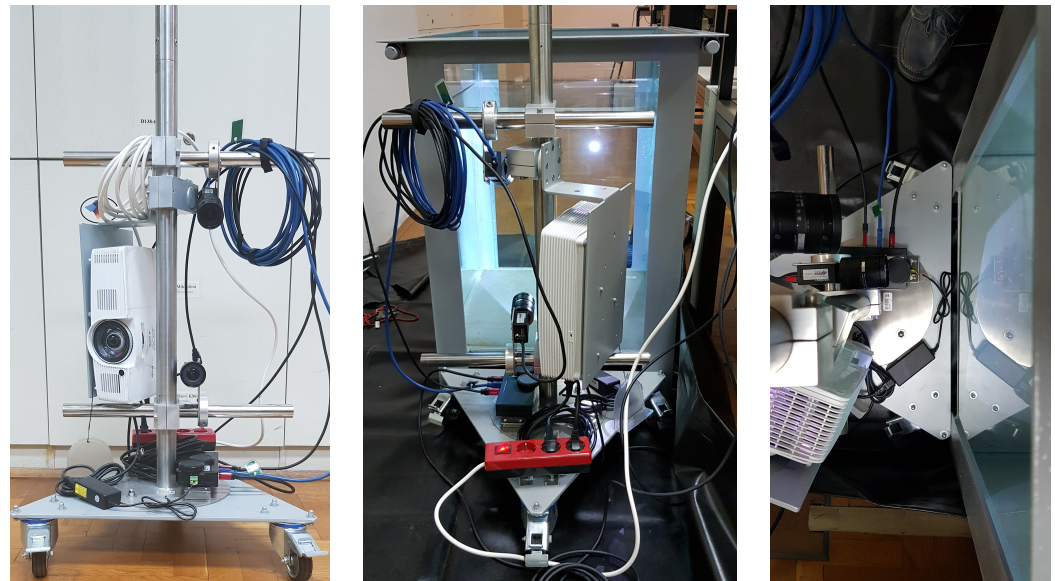
Figure 6. Laboratory setup. **Left**: SL scanner. **Middle**: side view. **Right**: top view.

4.2. Data Acquisition

Since an in-air calibration of both the camera and projector is a prerequisite, we first performed the geometrical calibration using the procedure of [21]. Then, we acquired the calibration data for four relative positions of the imaging system w.r.t. to the four glass interfaces; we performed this calibration once per each glass interface (see Section 4.1). For every position of the imaging system, the calibration board (Figure 3) was imaged underwater in at least five positions. The extraction of the calibration points (Figure 7, right) was performed using the procedure of [21], and the SL code (Figure 7, left) was MPS-coded using 20:21:25 ratios [32].
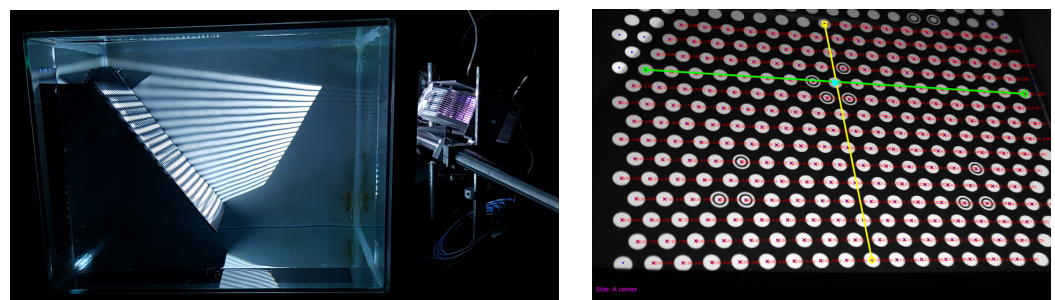
Figure 7. Underwater imaging (**left**: structured light projection; **right**: extraction of points' coordinates).

4.3. Axis Estimation

The axis of an axial system may be estimated from the proposed unified coplanarity constraint given by Equation (14) directly as $\mathbf{a}_1$ or indirectly by the decomposition of the

matrix $\mathbf{E}$ (see Section 3.3). We estimated the axis of the system directly from the coplanarity constraint. The mean angular error for the obtained axis of the system is presented in Table 1 for four different glass interfaces.

Table 1. Mean angular errors and standard deviation for the axis direction w.r.t. the true axis of the system (measured in air): column i lists sides of the tank; column d_1 lists corresponding glass thickness; column Ours lists error for the axis $\mathbf{A}$ estimated from the Equation (14) and for the axis $\mathbf{a}$ extracted by decomposition of the essential matrix $\mathbf{E_1}$; column Agrawal et al. lists errors by utilizing the algorithm proposed by Agrawal et al. [19]. Sides II and III of the tank are both 10 mm thick; the exponent (a) denotes a real glass, while the exponent (b) denotes a laminated glass.

i	d_1	Ours		Agrawal et al. [19]		
	[mm]	$\theta_\mathbf{A}$ [°]	$\theta_\mathbf{a}$ [°]	$\theta_{\mathbf{CAM}_a}$ [°]	$\theta_{\mathbf{CAM}_b}$ [°]	$\theta_{\mathbf{PRJ}_c}$ [°]
I	8	4.85 ± 0.73	4.83 ± 0.71	81.22 ± 18.36	78.44 ± 17.04	47.00 ± 38.30
II	$10^{(a)}$	5.34 ± 0.81	5.35 ± 0.80	81.81 ± 22.16	73.65 ± 26.76	46.96 ± 27.23
III	$10^{(b)}$	17.38 ± 16.72	17.11 ± 16.58	79.16 ± 32.19	79.04 ± 17.59	84.76 ± 6.22
IV	12	5.31 ± 0.83	5.32 ± 0.82	72.95 ± 31.19	64.61 ± 22.33	64.98 ± 17.31

4.4. Errors in 2D and 3D

Three components form the total error in 3D: the coplanarity error, the backprojection error, and the frustum error (23). The total error in 3D is presented in Table 2 as a mean backprojection and coplanarity error for four different glass interfaces. Note that the frustum error was expected to be zero, since all calibration points needed to be within the field of view (FOV).

Table 2. Errors in 2D and 3D: column i lists sides of the tank, while column d_1 lists corresponding glass thickness; column Total Error (Mean) lists mean coplanarity and backprojection errors in 3D; column Re-Projection Error (Mean) lists mean errors in 2D for a corresponding camera (projector); and column Re-Projection Error (Median) lists median of errors in 2D. Sides II and III of the tank are both 10 mm thick; the exponent (a) denotes a real glass, while the exponent (b) denotes a laminated glass.

i	d_1	Total Error (Mean) [mm] (23)		Re-Projection Error (Mean) [px]			Re-Projection Error (Median) [px]		
	[mm]	$e_{\mathbf{BPR}}$	$e_{\mathbf{CPL}}$	e_a	e_b	e_c	e_a	e_b	e_c
I	8	0.63	0.17	5.11	2.38	1.91	5.24	1.99	1.32
II	$10^{(a)}$	0.47	0.13	4.22	2.01	1.80	4.47	1.72	1.23
III	$10^{(b)}$	2.72	0.78	16.33	6.88	5.20	18.54	6.00	2.64
IV	12	0.50	0.14	4.12	2.19	1.66	4.16	1.97	1.24

The re-projection error, i.e., the distance in pixels in the image frame between the observed 2D point and the re-projected 3D point, is computed using the AFP equation [19] for a two-layer flat refractive system. Note that the AFP equation was not used in the calibration procedure. Table 2 presents the mean and median values of the re-projection error for four different glass interfaces (tank sides I–IV).

To evaluate the 3D reconstruction quality, we also fit 3D planes onto the reconstructed 3D points for each position of the calibration board. Table 3 lists the mean absolute distances of the reconstructed 3D points for the ideal fitted plane.

Table 3. Fitting errors: column i lists sides of the tank; column d_1 lists corresponding glass thickness; column N_i lists number of positions in which the calibration board was imaged; column Fitting Error (Mean) lists mean errors computed as distances from the fitted plane in 3D; and column Fitting Error (Median) lists median errors computed as distances from the fitted plane in 3D. Sides II and III of the tank are both 10 mm thick; the exponent (a) denotes a real glass, while the exponent (b) denotes a laminated glass.

i	d_1 [mm]	N_i	Fitting Error (Mean) [mm] $[e_j], j = 1, \ldots, N_i$	Fitting Error (Median) [mm] $[e_j], j = 1, \ldots, N_i$
I	8	7	[1.83], [1.31], [0.60], [2.29], [1.08], [1.12], [1.69]	[1.34], [1.06], [0.46], [1.80], [0.90], [0.93], [1.39]
II	$10^{(a)}$	6	[1.18], [1.16], [0.84], [2.16], [1.43], [1.37]	[0.99], [0.95], [0.69], [1.68], [1.14], [1.07]
III	$10^{(b)}$	7	[2.70], [1.80], [2.42], [4.16], [1.76], [3.21], [3.68]	[2.36], [1.56], [2.13], [3.40], [1.55], [2.82], [3.07]
IV	12	6	[1.24], [1.09], [0.96], [2.12], [1.40], [1.30]	[1.04], [0.90], [0.80], [1.62], [1.12], [1.02]

5. Discussion

The calibrated imaging geometry is presented in Figure 5, where calibrated poses of the calibration board are shown. Backprojected rays were drawn for the position of the board w.r.t. the side I of the tank; rays in the air are displayed in red, green rays are the refracted rays in the glass interface, and blue rays are the refracted rays in the water.

Regarding the axis estimation, Zoraja et al. [9] conjectured that the coplanarity constraint given by Equation (1) exhibits a coupling between the axis $\hat{a}_x$ and the pose parameters $\mathbf{S}$ and $\mathbf{u}$. The coplanarity constraint of Equation (1) states that the points in 3D must lie on the POR, but, if we flatten the 3D space onto the POR, the constraint holds as well. Let $\mathbf{O} = \mathrm{Null}(\mathbf{a}^T)^T$ be a 3×2 matrix representing the orthographic projection of the 3D space onto the null space of $\hat{a}_x$. Then,

$$\left(\mathbf{O}([\mathbf{a}]_\times \mathbf{v}_0) \right)^T \left(\mathbf{O}(\mathbf{S}\mathbf{p} + \mathbf{u}) \right) = \mathbf{v}_0^T [\mathbf{a}]_\times \mathbf{O}^T \mathbf{O}(\mathbf{S}\mathbf{p} + \mathbf{u}) = 0 \tag{25}$$

holds, thus indicating a possible instability in the solution, as points $\mathbf{p}$ are planar and losing one degree of freedom, and, due to coupling the null-space of Equation (4), are susceptible to noise. Zoraja et al. [9] analyzed the data provided by Agrawal et al. [19] by minimizing Equation (1) to start with many random initial axis directions, and they found at least three local minima having an average CPL error less than 1 mm. Axis estimations using the procedure of Agrawal et al. [19] are shown in Figure 8 on the left: in red are estimated axes of the camera $\mathbf{a}$, in green are the estimated axes of the camera $\mathbf{b}$, and in blue are the estimated axes of the projector $\mathbf{c}$. The ground truth axis of the system is displayed in black. On the other hand, if the axis is estimated using the unified coplanarity constraint, (14) then the results are much more robust. The axis estimated in such way is a robust initial point for the proposed optimization algorithm, and this approach is presented in Figure 8 on the right: in red are the estimated axes of the system, and in black is the ground truth axis of the system.

Numerical values for mean angular errors (before optimization) w.r.t. different sides of the tank are listed in Table 1: the column Ours lists angles between the mean estimated axes of the system and the ground truth (the sub-column $\theta_\mathbf{A}$ for axes was extracted from Equation (14), and the sub-column $\theta_\mathbf{a}$ for axes was obtained by the decomposition of the essential matrix of the system), and the column Agrawal et al. [19] lists angles between the mean estimated axes of the single camera/projector obtained by the decomposition of the corresponding essential matrices [19]. The mean angular error of the estimated systems' axes for each position of the SL system w.r.t. the tank's side was 5.17 deg (without considering the side III), which is acceptable and is a great improvement considering the error we obtained from Agrawal et al. [19] (see also Zoraja et al. [9]). Furthermore, system's axis estimated by the unified coplanarity constraint was a good initial value for

the optimization algorithm. Note that the angular error obtained w.r.t. the side III was greater than others due to the optical properties of the laminated glass, which we modeled as a single glass medium (without considering different optical properties for each layer of the laminated glass).
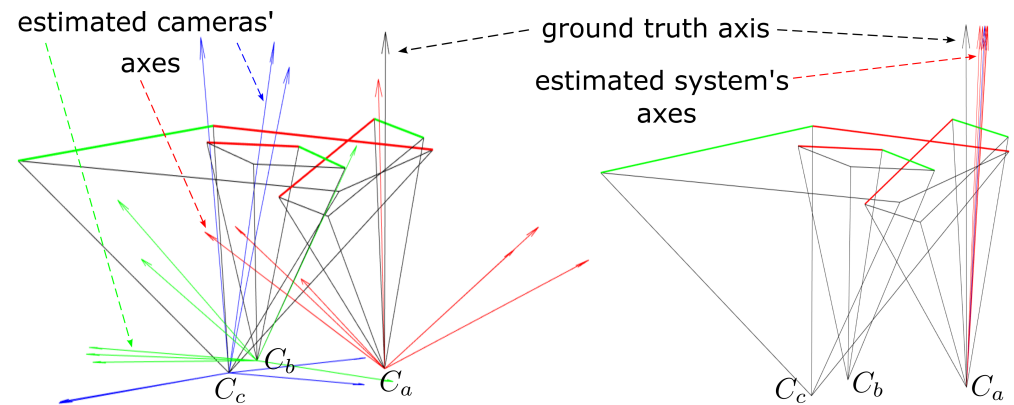
Figure 8. Estimated cameras' axes for different views of the calibration board (before optimization); the estimates were used as an initial point in the optimization. **Left**: using the coplanarity constraint for a single camera/projector where red, green, and blue vectors correspond to axes acquired for each device in the system; **Right**: using the unified coplanarity constraint for a system where the axis of the system is denoted in red. The ground truth axis is denoted in black for both the left and the right case.

The numerical results for the total error w.r.t. the different sides of the tank are listed in Table 2. Both the experimental backprojection and the coplanarity errors were less than one millimeter, which was sufficient for the 3D reconstruction using SL imaging when considering a working distance of up to 2 m. The numerical results for the re-projection error w.r.t. the different side of the tank are listed in Table 2; both the mean and median of the re-projection error are presented. For better visualization of the estimated re-projection error, we provide Figure 9 for a single position of the calibration board in water (imaged over the 8 mm thick glass interface). Notice that re-projection error was greater for the camera **a**, w.r.t. the camera **b** and the projector **c**, due to the greater angle between the optical axis and the axis of the system (see Figure 5). The re-projection error increased with the aforementioned angle.
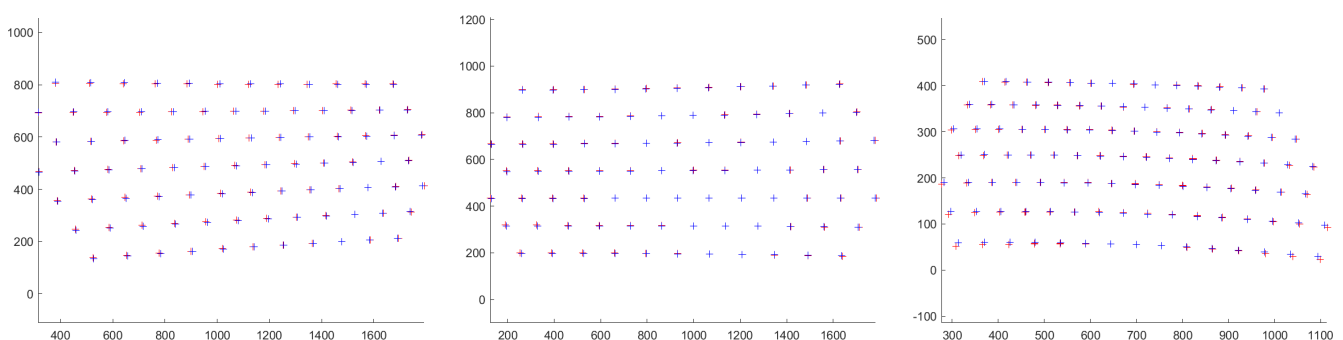
Figure 9. Extracted 2D points and corresponding re-projected 3D points for a single position of the calibration board in water (imaged over the 8mm thick glass interface): **left**—error for the camera **a**, **middle**—error for the camera **b**, and **right**—error for the projector. Red crosses represent re-projected 3D points in the camera (projector) frame, and blue crosses are the extracted 2D points.

The last quantitative measure is the distance error for a fitted plane. The numerical results for the proposed error w.r.t. the different sides of the tank are listed in Table 3. Considering the number of positions N_i in which the calibration board was imaged, for every

position (I–IV) of the imaging system, we reconstructed the calibration board and fitted the plane on the acquired dataset, thus acquiring the mean distance error. For positions where the angle δ between the axis $\hat{a}_x$ and the normal of the calibration board was minimal, the mean error was less than a millimeter. By increasing δ, the mean distance error increased.

Note that the error in 3D, the re-projection error, and the distance error for the fitted plane were greater for the pose where an imaging system observed the calibration board over the side III of the tank due to the optical properties of the laminated glass, which we modeled as a simple glass interface (disregarding reflections and refractions on multiple layers of the laminated glass).

6. Conclusions

We have presented a novel calibration procedure of a structured light underwater 3D imaging system using the theory of two-layer flat refractive geometry. The proposed calibration procedure is applicable to SL scanners with an arbitrary number of cameras and projectors, which share a common flat glass interface. It also uses a simple planar calibration board, which is readily available. The obtained experimental results demonstrated a low mean angular error for the estimation of the axis of the system using the proposed unified coplanarity constraint. This indicates that the proposed unified coplanarity constraint gives a robust initial estimate for the optimization algorithm. The experimentally assessed errors in 3D were less then one millimeter, which is sufficient for a 3D reconstruction and which indicates that the proposed calibration is applicable in the real world.

Future Work

Given that the proposed system was evaluated in a laboratory with a tank simulating an underwater environment, our future work will focus on the design and implementation of a prototype of an underwater SL scanner that will be tested in a real underwater environment. We will also investigate the possibility of using a 3D calibration object comprising two planar boards at a 90° angle. Finally, we will investigate the robustness of the MPS-structured light pattern under conditions of increased turbidity.

Author Contributions: Conceptualization, D.Z. and T.P. (Tomislav Petković); methodology, D.Z. and T.P. (Tomislav Petković); validation, T.P. (Tomislav Pribanić) and J.F.; investigation, D.Z.; writing—original draft preparation, D.Z.; writing—review and editing, T.P. (Tomislav Petković); visualization, D.Z.; supervision, T.P. (Tomislav Petković); funding acquisition, T.P. (Tomislav Petković). All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: On request.

Conflicts of Interest: The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

ROV	Remotely Operated Vehicle
UAV	Underwater Autonomous Vehicle
SL	Structured Light
FRC	Flat Refraction Constraint
POR	Plane of Refraction
MPS	Multiple Phase Shift

SVD   Single Value Decomposition
AFP   Analytical Forward Projection
CPL   Coplanarity
BPR   Backprojection
FRS   Frustum
FOV   Field of View

Appendix A

Here, we present the complete derivation of the coplanarity constraint individually for a single camera/projector and the derivation of the unified coplanarity constraint for an imaging system. The equation for the unified coplanarity constraint is derived w.r.t. the referent coordinate system, which we set to be the coordinate system of the camera **a**.

Appendix A.1. Coplanarity Constraint for Camera (Projector) Using a Single Interface

A coplanarity constraint is given as:

$$\left([\mathbf{a}]_\times \mathbf{v}_0\right)^T (\mathbf{S}\mathbf{p} + \mathbf{u}) = 0. \tag{A1}$$

Considering Equation (A1), the coplanarity constraint between the calibration object and the camera **a** may me derived as:

$$\left([\mathbf{a_1}]_\times \mathbf{v}_{0_a}\right)^T (\mathbf{S_a}\mathbf{p_a} + \mathbf{u_a}) = 0$$

$$-\mathbf{v}_{0_a}^T \left[[\mathbf{a_1}]_\times (\mathbf{S_a}\mathbf{p_a} + \mathbf{u_a})\right] = 0$$

$$\mathbf{v}_{0_a}^T \left[[\mathbf{a_1}]_\times \mathbf{S_a}\mathbf{p_a} + [a_1]_\times \mathbf{u_a}\right] = 0$$

$$\mathbf{v}_{0_a}^T [\mathbf{a_1}]_\times \mathbf{S_a}\mathbf{p_a} + \mathbf{v}_{0_1}^T [a_1]_\times \mathbf{u_a} = 0. \tag{A2}$$

By using the following identities

$$[\mathbf{a_1}]_\times \mathbf{S_a} = \mathbf{E_1} \tag{A3}$$

$$[\mathbf{a_1}]_\times \mathbf{u_a} = \mathbf{h_1} \tag{A4}$$

the Equation (A2) may be simplified as

$$\mathbf{v}_{0_a}^T \mathbf{E_1}\mathbf{p_a} + \mathbf{v}_{0_a}^T \mathbf{h_1} = 0. \tag{A5}$$

The resulting linear system from Equation (A5) may be expressed in the matrix form:

$$\begin{bmatrix} \mathbf{p_a}(\mathbf{1})^{\mathbf{T}} \otimes \mathbf{v}_{0_a}(1)^T & \mathbf{v}_{0_a}(1)^T \\ \mathbf{p_a}(\mathbf{2})^{\mathbf{T}} \otimes \mathbf{v}_{0_a}(2)^T & \mathbf{v}_{0_a}(2)^T \\ \vdots & \vdots \end{bmatrix} \begin{bmatrix} \mathbf{E_1}(:) \\ \mathbf{h_1} \end{bmatrix} = \mathbf{0}, \tag{A6}$$

where $\otimes$ denotes a Kronecker's product, and $\mathbf{E_1}(:)$ is the vector formed by stacking the columns of the essential matrix $\mathbf{E_1}$.

Considering Equation (A1), the coplanarity constraint between the calibration object and the camera **b** is

$$\left([\mathbf{a_2}]_\times \mathbf{v}_{0_b}\right)^T (\mathbf{S_b}\mathbf{p_b} + \mathbf{u_b}) = 0$$

$$-\mathbf{v}_{0_b}^T \left[[\mathbf{a_2}]_\times (\mathbf{S_b}\mathbf{p_b} + \mathbf{u_b})\right] = 0$$

$$\mathbf{v}_{0_b}^T \left[[\mathbf{R_{ab}}\mathbf{a_2}]_\times (\mathbf{R_{ab}}\mathbf{S_a}\mathbf{p_b} + \mathbf{R_{ab}}\mathbf{u_a} + \mathbf{t_{ab}})\right] = 0, \tag{A7}$$

and the coplanarity constraint between the calibration object and the projector **c** is

$$\left([\mathbf{a_3}]_\times \mathbf{v}_{0_c}\right)^T (\mathbf{S_c p_c} + \mathbf{u_c}) = 0$$

$$-\mathbf{v}_{0_c}^T \left[[\mathbf{a_3}]_\times (\mathbf{S_c p_c} + \mathbf{u_c})\right] = 0$$

$$\mathbf{v}_{0_c}^T \left[[\mathbf{R_{ac} a_3}]_\times (\mathbf{R_{ac} S_a p_c} + \mathbf{R_{ac} u_a} + \mathbf{t_{ac}})\right] = 0. \tag{A8}$$

By using the following identities depicting rotations and translations between the coordinate systems **a**, **b**, and **c**,

$$\mathbf{R_{ab}} = \mathbf{R_b R_a}^T \qquad\qquad \mathbf{t_{ab}} = -\mathbf{R_b C_b} + \mathbf{R_b C_a}$$
$$\mathbf{R_{ac}} = \mathbf{R_c R_a}^T \qquad\qquad \mathbf{t_{ac}} = -\mathbf{R_c C_c} + \mathbf{R_c C_a}$$

and the identities for the rotations and translations between the aforementioned coordinate systems and the coordinate system of the calibration object (world frame),

$$\mathbf{S_b} = \mathbf{R_{ab} S_a} \qquad\qquad \mathbf{u_b} = \mathbf{R_{ab} u_a} + \mathbf{t_{ab}}$$
$$\mathbf{S_c} = \mathbf{R_{ac} S_a} \qquad\qquad \mathbf{u_c} = \mathbf{R_{ac} u_a} + \mathbf{t_{ac}}$$

the coplanarity constraint for the camera **b** w.r.t. the referent frame (A7) may be re-written as

$$\mathbf{v}_{0_b}^T \left[[\mathbf{R_{ab} a_1}]_\times \mathbf{R_{ab} S_a p_b} + [\mathbf{R_{ab} a_1}]_\times \mathbf{R_{ab} u_a} + [\mathbf{R_{ab} a_1}]_\times \mathbf{t_{ab}}\right] = 0$$

$$\mathbf{v}_{0_b}^T \left[\mathbf{R_{ab}}\left([\mathbf{a_1}]_\times \mathbf{S_a p_b}\right) + \mathbf{R_{ab}}\left([\mathbf{a_1}]_\times \mathbf{u_a}\right) + [\mathbf{R_{ab} a_1}]_\times \mathbf{t_{ab}}\right] = 0$$

$$\mathbf{v}_{0_b}^T \mathbf{R_{ab}}\left([\mathbf{a_1}]_\times \mathbf{S_a p_b}\right) + \mathbf{v}_{0_b}^T \mathbf{R_{ab}}\left([\mathbf{a_1}]_\times \mathbf{u_a}\right) + \mathbf{v}_{0_b}^T [\mathbf{R_{ab} a_1}]_\times \mathbf{t_{ab}} = 0, \tag{A9}$$

and the coplanarity constraint for the projector **c** w.r.t. the referent frame (A8) may be re-written as

$$\mathbf{v}_{0_c}^T \left[[\mathbf{R_{ac} a_1}]_\times \mathbf{R_{ac} S_a p_c} + [\mathbf{R_{ac} a_1}]_\times \mathbf{R_{ac} u_a} + [\mathbf{R_{ac} a_1}]_\times \mathbf{t_{ac}}\right] = 0$$

$$\mathbf{v}_{0_c}^T \left[\mathbf{R_{ac}}\left([\mathbf{a_1}]_\times \mathbf{S_a p_c}\right) + \mathbf{R_{ac}}\left([\mathbf{a_1}]_\times \mathbf{u_a}\right) + [\mathbf{R_{ac} a_1}]_\times \mathbf{t_{ac}}\right] = 0$$

$$\mathbf{v}_{0_c}^T \mathbf{R_{ac}}\left([\mathbf{a_1}]_\times \mathbf{S_a p_c}\right) + \mathbf{v}_{0_c}^T \mathbf{R_{ac}}\left([\mathbf{a_1}]_\times \mathbf{u_a}\right) + \mathbf{v}_{0_c}^T [\mathbf{R_{ac} a_1}]_\times \mathbf{t_{ac}} = 0. \tag{A10}$$

By using the identities from Equations (A3) and (A4), we may simplify Equations (A9) and (A10) as

$$\mathbf{v}_{0_b}^T \mathbf{R_{ab} E_1 p_b} + \mathbf{v}_{0_b}^T \mathbf{R_{ab} h_1} + \mathbf{v}_{0_b}^T [\mathbf{R_{ab}} a_1]_\times \mathbf{t_{ab}} = 0 \tag{A11}$$

and

$$\mathbf{v}_{0_c}^T \mathbf{R_{ac} E_1 p_c} + \mathbf{v}_{0_c}^T \mathbf{R_{ac} h_1} + \mathbf{v}_{0_c}^T [\mathbf{R_{ac}} a_1]_\times \mathbf{t_{ac}} = 0, \tag{A12}$$

where Equation (A11) denotes the coplanarity constraint between the calibration object and the camera **b** w.r.t. the referent frame, and Equation (A12) denotes the coplanarity constraint between the calibration object and the projector **c** w.r.t. the referent frame.

Appendix A.2. Unified Coplanarity Constraint for System Using a Single Interface

Equations (A5), (A11), and (A12) denote the coplanarity constraints of camera **a**, camera **b**, and projector **c** w.r.t. the referent frame, respectively. The third terms of Equations (A11) and (A12) may be reformed, which, for Equation (A11), results in

$$\mathbf{v}_{0_b}^T[\mathbf{R_{ab}a_1}]_\times \mathbf{t_{ab}} = \mathbf{v}_{0_b}^T[\mathbf{R_{ab}}a_1]_\times\left(\mathbf{R_{ab}R_{ab}}^T\mathbf{t_{ab}}\right)$$

$$= \mathbf{v}_{0_b}^T\mathbf{R_{ab}}\left[[\mathbf{a_1}]_\times\left(\mathbf{R_{ab}}^T\mathbf{t_{ab}}\right)\right]$$

$$= \mathbf{a}_1^T\left[\left(\mathbf{R_{ab}}^T\mathbf{t_{ab}}\right)\times\left(\mathbf{v}_{0_b}^T\mathbf{R_{ab}}\right)^T\right]$$

$$= \mathbf{a}_1^T\left[\left(\mathbf{R_{ab}}^T\mathbf{t_{ab}}\right)\times\left(\mathbf{R_{ab}}^T\mathbf{v}_{0_b}\right)\right]$$

$$= \mathbf{a}_1^T\left[\mathbf{R_{ab}}^T\left(\mathbf{t_{ab}}\times\mathbf{v}_{0_b}\right)\right]$$

$$= \left[\mathbf{R_{ab}}^T\left(\mathbf{t_{ab}}\times\mathbf{v}_{0_b}\right)\right]^T\mathbf{a_1}.$$

Similarly, the third term of Equation (A12) may be reformed as

$$\mathbf{v}_{0_c}^T[\mathbf{R_{ac}a_1}]_\times \mathbf{t_{ac}} = \mathbf{v}_{0_c}^T[\mathbf{R_{ac}}a_1]_\times\left(\mathbf{R_{ac}R_{ac}}^T\mathbf{t_{ac}}\right) = \cdots = \left[\mathbf{R_{ac}}^T\left(\mathbf{t_{ac}}\times\mathbf{v}_{0_c}\right)\right]^T\mathbf{a_1}.$$

Considering the reformation of the third term in both Equations (A11) and (A12), both linear systems may be re-written as

$$\mathbf{v}_{0_b}^T\mathbf{R_{ab}E_1p_b} + \mathbf{v}_{0_b}^T\mathbf{R_{ab}s_1} + \left[\mathbf{R_{ab}}^T\left(\mathbf{t_{ab}}\times\mathbf{v}_{0_b}\right)\right]^T\mathbf{a_1} = 0 \qquad (A13)$$

$$\mathbf{v}_{0_c}^T\mathbf{R_{ac}E_1p_c} + \mathbf{v}_{0_c}^T\mathbf{R_{ac}s_1} + \left[\mathbf{R_{ac}}^T\left(\mathbf{t_{ac}}\times\mathbf{v}_{0_c}\right)\right]^T\mathbf{a_1} = 0. \qquad (A14)$$

The matrix form of Equation (A11) is

$$\begin{bmatrix} \mathbf{p_b(1)^T}\otimes\left(\mathbf{v}_{0_b}(1)^T\mathbf{R_{ab}}\right) & \mathbf{v}_{0_b}(1)^T\mathbf{R_{ab}} & \left[\mathbf{R_{ab}}^T\left(\mathbf{t_{ab}}\times\mathbf{v}_{0_b}(1)\right)\right]^T \\ \mathbf{p_b(2)^T}\otimes\left(\mathbf{v}_{0_b}(2)^T\mathbf{R_{ab}}\right) & \mathbf{v}_{0_b}(2)^T\mathbf{R_{ab}} & \left[\mathbf{R_{ab}}^T\left(\mathbf{t_{ab}}\times\mathbf{v}_{0_b}(2)\right)\right]^T \\ \vdots & \vdots & \vdots \end{bmatrix}\begin{bmatrix}\mathbf{E_1}(:) \\ \mathbf{s_1} \\ \mathbf{a_1}\end{bmatrix} = \mathbf{0}, \qquad (A15)$$

and the matrix form of Equation (A12) is

$$\begin{bmatrix} \mathbf{p_c(1)^T}\otimes\left(\mathbf{v}_{0_c}(1)^T\mathbf{R_{ac}}\right) & \mathbf{v}_{0_c}(1)^T\mathbf{R_{ac}} & \left[\mathbf{R_{ac}}^T\left(\mathbf{t_{ac}}\times\mathbf{v}_{0_c}(1)\right)\right]^T \\ \mathbf{p_c(2)^T}\otimes\left(\mathbf{v}_{0_c}(2)^T\mathbf{R_{ac}}\right) & \mathbf{v}_{0_c}(2)^T\mathbf{R_{ac}} & \left[\mathbf{R_{ac}}^T\left(\mathbf{t_{ac}}\times\mathbf{v}_{0_c}(2)\right)\right]^T \\ \vdots & \vdots & \vdots \end{bmatrix}\begin{bmatrix}\mathbf{E_1}(:) \\ \mathbf{s_1} \\ \mathbf{a_1}\end{bmatrix} = 0. \qquad (A16)$$

Considering linear systems, posed by Equations (A6), (A15), and (A16), by denoting the coplanarity constraints between the calibration object and the corresponding camera/projector frame w.r.t. the referent frame, we obtain the unified coplanarity constraint of Equation (14) for the imaging system w.r.t. the referent frame.

References

1. 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, Australia, 4–27 May 2010; pp. 1–9. [CrossRef]
2. Shortis, M.; Ravanbakskh, M.; Shaifat, F.; Harvey, E.; Mian, A.; Seager, J.; Culverhouse, P.; Cline, D.; Edgington, D. A review of techniques for the identification and measurement of fish in underwater stereo-video image sequences. In *Videometrics, Range Imaging, and Applications XII; and Automated Visual Inspection*; SPIE: Bellingham, WA, USA, 2013; Volume 8791, p. 87910G. [CrossRef]
3. Lopes, F.; Silva, H.; Almeida, J.; Martins, A.; Silva, E. Structured light system for underwater inspection operations. In Proceedings of the OCEANS 2015–Genova, Genova, Italy, 18–21 May 2015; pp. 1–6. [CrossRef]
4. Castillón, M.; Forest, J.; Ridao, P. Underwater 3D Scanner to Counteract Refraction: Calibration and Experimental Results. *IEEE/ASME Trans. Mechatronics* **2022**, *27*, 1–9. [CrossRef]
5. Castillón, M.; Palomer, A.; Forest, J.; Ridao, P. Underwater 3D Scanner Model Using a Biaxial MEMS Mirror. *IEEE Access* **2021**, *9*, 50231–50243. [CrossRef]
6. Fan, J.; Wang, X.; Zhou, C.; Ou, Y.; Jing, F.; Hou, Z. Development, Calibration, and Image Processing of Underwater Structured Light Vision System: A Survey. *IEEE Trans. Instrum. Meas.* **2023**, *72*, 1–18. [CrossRef]
7. Shortis, M. Camera Calibration Techniques for Accurate Measurement Underwater. In *3D Recording and Interpretation for Maritime Archaeology*; McCarthy, J.K., Benjamin, J., Winton, T., van Duivenvoorde, W., Eds.; Springer International Publishing: Cham, Switzerland, 2019; pp. 11–27. [CrossRef]
8. Maas, H.G. On the Accuracy Potential in Underwater/Multimedia Photogrammetry. *Sensors* **2015**, *15*, 18140–18152. [CrossRef] [PubMed]
9. Zoraja, D.; Petković, T.; Pribanić, T.; Forest, J. Projector Calibration in a Two-Layer Flat Refractive Geometry for Underwater Imaging. In Proceedings of the 2022 45th Jubilee International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 23–27 May 2022; pp. 930–935. [CrossRef]
10. Hartley, R.I.; Zisserman, A. *Multiple View Geometry in Computer Vision*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2004; ISBN 0521540518.
11. Geng, J. Structured-light 3D surface imaging: A tutorial. *Adv. Opt. Photon.* **2011**, *3*, 128–160. [CrossRef]
12. Salvi, J.; Fernandez, S.; Pribanic, T.; Llado, X. A state of the art in structured light patterns for surface profilometry. *Pattern Recognit.* **2010**, *43*, 2666–2680. [CrossRef]
13. Bruno, F.; Bianco, G.; Muzzupappa, M.; Barone, S.; Razionale, A. Experimentation of structured light and stereo vision for underwater 3D reconstruction. *ISPRS J. Photogramm. Remote Sens.* **2011**, *66*, 508–518. [CrossRef]
14. Sarafraz, A.; Haus, B.K. A structured light method for underwater surface reconstruction. *ISPRS J. Photogramm. Remote Sens.* **2016**, *114*, 40–52. [CrossRef]
15. Narasimhan, S.; Nayar, S.; Sun, B.; Koppal, S. Structured light in scattering media. In Proceedings of the Tenth IEEE International Conference on Computer Vision (ICCV'05), Beijing, China, 17–21 October 2005; Volume 1, pp. 420–427. [CrossRef]
16. Bräuer-Burchardt, C.; Munkelt, C.; Heinze, M.; Gebhart, I.; Kühmstedt, P.; Notni, G. Underwater 3D measurements with advanced camera modelling. *PFG—J. Photogramm. Remote Sens. Geoinf. Sci.* **2022**, *90*, 55–67. [CrossRef]
17. Young, M. The pinhole camera: Imaging without lenses or mirrors. *Phys. Teach.* **1989**, *27*, 648–655. [CrossRef]
18. Sedlazeck, A.; Koch, R. Perspective and Non-perspective Camera Models in Underwater Imaging—Overview and Error Analysis. In *Outdoor and Large-Scale Real-World Scene Analysis, Proceedings of the 15th International Conference on Theoretical Foundations of Computer Vision: Outdoor and Large-Scale Real-World Scene Analysis, Dagstuhl Castle, Germany, 26 June–1 July 2011*; Dellaert, F., Frahm, J.M., Pollefeys, M., Leal-Taixé, L., Rosenhahn, B., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 212–242.
19. Agrawal, A.; Ramalingam, S.; Taguchi, Y.; Chari, V. A Theory of Multi-Layer Flat Refractive Geometry. Available online: http://www.amitkagrawal.com/cvpr12/FlatRefraction.html (accessed on 28 November 2022).
20. Ramalingam, S.; Sturm, P.; Lodha, S. Theory and Calibration for Axial Cameras. In Proceedings of the Computer Vision—ACCV 2006: 7th Asian Conference on Computer Vision, Hyderabad, India, 13–16 January 2006; Volume 3851. [CrossRef]
21. Petković, T.; Gasparini, S.; Pribanić, T. A Note on Geometric Calibration of Multiple Cameras and Projectors. In Proceedings of the 2020 43rd International Convention on Information, Communication and Electronic Technology (MIPRO), Opatija, Croatia, 28 September–2 October 2020; pp. 1157–1162. [CrossRef]
22. Łuczyński, T.; Pfingsthorn, 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]
23. Zhang, S.; Huang, P. Novel method for structured light system calibration. *Opt. Eng.* **2006**, *45*, 083601. [CrossRef]
24. Zhang, Z. A flexible new technique for camera calibration. *IEEE Trans. Pattern Anal. Mach. Intell.* **2000**, *22*, 1330–1334. [CrossRef]
25. Kawahara, R.; Nobuhara, S.; Matsuyama, T. A Pixel-Wise Varifocal Camera Model for Efficient Forward Projection and Linear Extrinsic Calibration of Underwater Cameras with Flat Housings. In Proceedings of the 2013 IEEE International Conference on Computer Vision Workshops, Sydney, Australia, 1–8 December 2013; pp. 819–824. [CrossRef]
26. Bouguet, J.Y. Camera Calibration Toolbox for Matlab. Available online: http://robots.stanford.edu/cs223b04/JeanYvesCalib/ (accessed on 29 November 2022).
27. Silvatti, A.P.; Salve Dias, F.A.; Cerveri, P.; Barros, R.M. Comparison of different camera calibration approaches for underwater applications. *J. Biomech.* **2012**, *45*, 1112–1116. [CrossRef]

28. Demazure, M. Sur Deux Problemes de Reconstruction. Ph.D. Thesis, National Institute for Research in Digital Science and Technology, Paris, France, 1988.
29. Nister, D. An efficient solution to the five-point relative pose problem. *IEEE Trans. Pattern Anal. Mach. Intell.* **2004**, *26*, 756–770. [CrossRef]
30. Li, B.; Karpinsky, N.; Zhang, S. Novel calibration method for structured-light system with an out-of-focus projector. *Appl. Opt.* **2014**, *53*, 3415–3426. [CrossRef]
31. Petković, T.; Pribanić, T. Multiprojector Multicamera Structured Light Surface Scanner. *IEEE Access* **2022**, *10*, 90321–90337. [CrossRef]
32. Petković, T.; Pribanić, T.; Đonlić, M. Temporal phase unwrapping using orthographic projection. *Opt. Lasers Eng.* **2017**, *90*, 34–47. [CrossRef]
33. Mathworks. Optimization Toolbox. Available online: https://www.mathworks.com/help/optim/ug/lsqnonlin.html (accessed on 2 December 2022).