

## Refractive geometry for underwater domes

Mengkun She<sup>\*</sup>, David Nakath, Yifan Song, Kevin Köser

Oceanic Machine Vision, GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany

### ARTICLE INFO

#### Keywords:

Underwater camera  
Dome port  
Refraction  
Spherical refractive geometry  
Calibration  
Decentering

### ABSTRACT

Underwater cameras are typically placed behind glass windows to protect them from the water. Spherical glass, a dome port, is well suited for high water pressures at great depth, allows for a large field of view, and avoids refraction if a pinhole camera is positioned exactly at the sphere's center. Adjusting a real lens perfectly to the dome center is a challenging task, both in terms of how to actually guide the centering process (e.g. visual servoing) and how to measure the alignment quality, but also, how to mechanically perform the alignment. Consequently, such systems are prone to being decentered by some offset, leading to challenging refraction patterns at the sphere that invalidate the pinhole camera model. We show that the overall camera system becomes an axial camera, even for thick domes as used for deep sea exploration and provide a non-iterative way to compute the center of refraction without requiring knowledge of exact air, glass or water properties. We also analyze the refractive geometry at the sphere, looking at effects such as forward- vs. backward decentering, iso-refraction curves and obtain a 6th-degree polynomial equation for forward projection of 3D points in thin domes. We then propose a pure underwater calibration procedure to estimate the decentering from multiple images. This estimate can either be used during adjustment to guide the mechanical position of the lens, or can be considered in photogrammetric underwater applications.

### 1. Introduction

More than two-thirds of Earth's surface is covered by water – or more specifically – by the oceans. Underwater imaging, vision, photogrammetry and robotic applications include recovering sunken cultural heritage, offshore installations, habitat mapping, resource estimation, deposited munition monitoring, optical quantification of processes in the ocean, human impact on ecosystems and last but not least exploration of the last uncharted terrain on Earth: the deep sea.

Underwater cameras are usually placed inside pressure housings and observe the outer world through some kind of *window*. Incident light rays travel through water, glass, and air before being sensed by the camera and each time they traverse media with different optical densities their direction might be changed. In particular for flat glass interfaces this requires more complex geometric reasoning (Treibitz et al., 2012; Agrawal et al., 2012; Maas, 2015; Jordt et al., 2016), such that alternative configurations are being explored. In oceanographic applications, but also for professional photographers, so-called *dome port systems* have become popular that rely on a spherical window to avoid view limitations. They are also mechanically more stable and relatively thin spherical glass can resist extremely high pressure at several

kilometers of depth. In principle, a well centered lens behind the dome port is able to avoid refraction, but it still remains challenging as both optical centers are physically invisible, especially for large lenses and rugged equipment as shown for example in Fig. 1. If the lens is not well centered, again refraction will occur as with flat ports, but now the decentered dome geometry produces even more complex refraction effects in the 2D image, a threat that might discourage people from using dome systems. Current practical solutions adopt standard pinhole calibration parameters to compensate the remaining refraction for an approximately centered lens, which allows to achieve high accuracy and has been widely applied in shallow water survey tasks (Drap, 2012b; Menna et al., 2013; Figueira et al., 2015; Nocerino et al., 2020). However, this solution is usually performed at an ideal working distance with a well designed control network, which is difficult to achieve in less controllable scenarios such as robotic mapping applications in the deep ocean.

In this contribution, we geometrically analyze refraction at the sphere for the common case that the lens is not exactly centered with the dome and show that the system is actually an axial camera. We derive a chessboard-based direct solver for the refraction axis, distinguish positive and negative decentering, and propose a complementary

<sup>\*</sup> Corresponding author.

E-mail address: [msh@geomar.de](mailto:msh@geomar.de) (M. She).

<https://doi.org/10.1016/j.isprsjprs.2021.11.006>

Received 16 July 2021; Received in revised form 31 October 2021; Accepted 17 November 2021

Available online 9 December 2021

0924-2716/© 2021 The Authors. Published by Elsevier B.V. on behalf of International Society for Photogrammetry and Remote Sensing, Inc. (ISPRS). This is an

open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

decentering calibration procedure that can support mechanical adjustment of the lens to avoid refraction in the first place. In case entirely avoiding refraction is not possible, the analyses and the methods for estimating the decentering presented in this paper are intended to facilitate further research, whether and when it could be beneficial to explicitly consider such physical parameters in photogrammetric surveys.

## 2. Related Work

It is well known in photogrammetry that refraction occurs when photographing through different media and in particular in underwater scenarios (see e.g. Shmutter, 1967; Moore, 1976; Kotowski, 1988; Fryer and Fraser, 1986; Maas, 1995a; Harvey and Shortis, 1998; Jaffe et al., 2001; Kunz and Singh, 2008; Drap, 2012a). Early photogrammetric methods have analytically analyzed the refractive scenario (Kotowski, 1988) and suggested linearized correction terms in multi-media photogrammetry at flat interfaces (Maas, 1995b), whereas other approaches suggested to absorb refractive effects into 2D lens distortion. Many practical underwater photogrammetry applications have shown that approximating the dome port camera system as a perspective camera model with standard distortion parameters is sufficient and convincing 3D survey results can be obtained.

Glaeser and Schröcker (2000) found the analytical forward projection for a single refraction. Treibitz et al. (2012) have shown that overall systems of cameras behind flat glass interfaces in the water can be considered axial cameras and Chari et al. (2009) have inspected related multi-view relations. Agrawal et al. (2012) have generalized analytical forward projection to multiple layers of flat interfaces.

For spherical glass domes, on the contrary, much less work exists, although they can sustain more pressure - and are thus better suited for deep sea applications - and they do not limit the field of view (see Fig. 1 for examples). Nocerino et al. (2016) compare how aerial standard calibration parameters change when submerging dome port cameras underwater, but disregard explicit consideration of refraction. Kunz and Singh (2008) discuss that domes avoid refraction when a pinhole camera is exactly centered with the dome and suggest that decentering could be determined using optimization. They analyze the 2D pixel error when approximating a misaligned dome port camera as a perspective camera and state that the error will increase as the camera moves closer or further to the observed scene since the refraction distortion is depth-dependent. Therefore, centering the camera with the dome port is critical to underwater vision applications. Here, Menna et al. (2016) practically measure a particular dome port geometry and discuss properties of domes as compared to flat port cameras. They also suggest how to align the nodal point of a lens with the dome's center by individually measuring both points and then mounting without a feedback loop. Recently, in Menna et al. (2020), the authors investigate the depth-dependent systematic errors introduced by refraction effects through dome ports in a statistical way and apply iterative look-up table corrections to reduce systematic residual patterns (Nocerino et al., 2021).

To support such a theory, and also to motivate this work, we conduct a numerical experiment which compares the physically-based decentered dome port model and the underwater perspective approximation model. We particularly look at the case of a camera-equipped robot that

sometimes comes close to underwater structures, and sometimes observes them from a distance, which means that calibration is required to hold for a range of different distances. We first model a ground-truth dome port camera system synthetically, in which the camera projection center is decentered in forward-backward and left-right direction with different magnitudes. Next, for each decentering, we simulate 30 virtual planar chessboard calibration images with 0.2 pixels Gaussian noise as if the target was photographed underwater. Then, the underwater dome port calibration and the perspective camera model calibration are both performed on the same set of images. After calibration, we randomly generate a ground-truth 3D point cloud that is in the camera's viewing frustum to evaluate the calibration quality by projecting the 3D points onto the image space and measure the mean absolute reprojection error. The point cloud consists of points with distances varying from 1m–10m, like in the robotic mapping scenario. As shown in Fig. 2, the underwater perspective approximation model works well for a mild decentering in optical axis direction (see top row), which is in agreement with the previous literature, but it is noticeable that the standard distortion parameters have more difficulty to absorb the refraction effect with larger decentering or when there is even only a slight sideward decentering (see bottom row). In these cases explicitly considering refraction does not suffer from such errors and this motivates the work in this contribution.

Already earlier, She et al. (2019) have proposed to mechanically adjust the dome and lens using through-the-lens feedback until no refraction effects are observable. They build their feedback loop on a human operator to judge the continuity of lines in a setup where the camera is looking parallel to the water surface in a special tank. Later they determine the remaining offset by an image pair of a chessboard in air and in water, both taken at exactly the same pose, which makes the approach delicate and complex. However, there are cases where the camera cannot be centered with the dome port. For instance, Bosch et al. (2015) have designed a special underwater camera housing for an omnidirectional camera system which consists of a cylinder for the lateral cameras and a hemispherical dome for the top-looking camera, then a calibration scheme to estimate the extrinsic parameters of each camera and the housing parameters is proposed. Refraction effects are considered using the ray-tracing based camera model. However, their work reports a relatively high residual after the parameter optimization in the real-world evaluation. Iscar and Johnson-Roberson (2020) have developed an underwater stereo camera system where two cameras share a single hemispherical glass port. Thus, the offset between the camera's optical center and the dome center is extreme in this scenario. They propose ideas to recover the camera's position inside the dome by measuring the point spread function of the overall camera system, which requires a complex dataset collection procedure. The authors report that their calibration approach is very time-consuming and practically limited. Besides underwater imaging, exact modeling of refraction in the image formation process can also improve in-air applications such as cameras behind the windshield of cars (Verbiest et al., 2020) or behind an optical cover for coal mining vehicles (Yang et al., 2021). Note however, that when imaging in air through a relatively thin glass pane, rays essentially only undergo a slight lateral shift depending on the thickness of the glass, whereas in our scenario the refraction effects are dominated by the other two media (air and water) with significantly

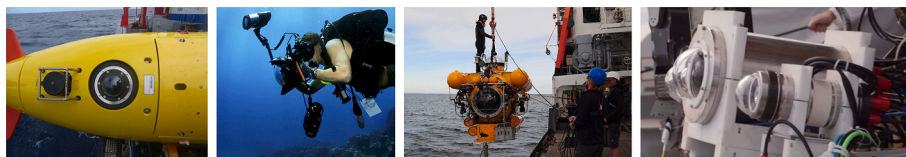
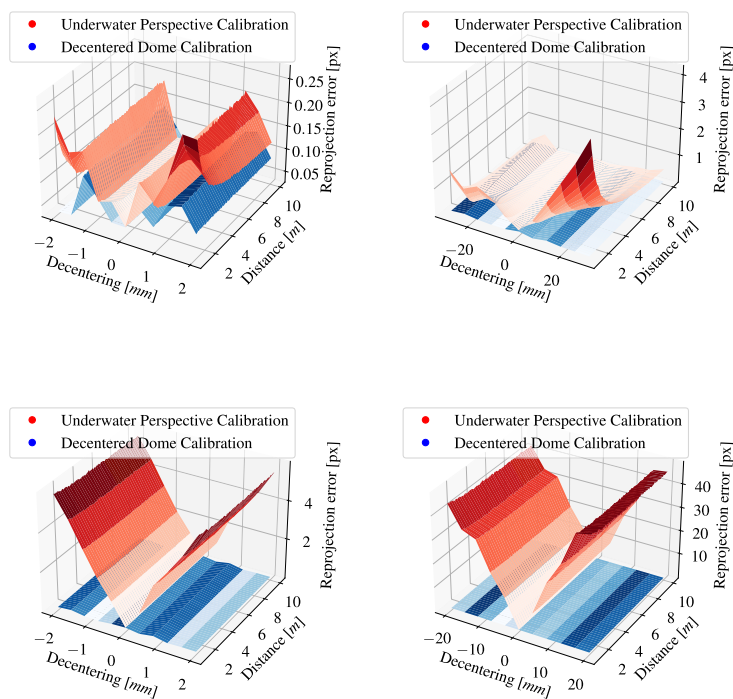


Fig. 1. From left to right: Autonomous underwater vehicle with dome port camera, an underwater photographer with dome for DSLR, submersible with huge spherical window, multi dome port camera system mounted on remotely operated deep sea robot. These domes can avoid refraction effects when all principal rays of the lenses pass the glass spheres in direction of the local normal. Perfectly centering the

lenses in the dome is however a hard task, since it is neither obvious where the center of the dome is, nor where the nodal point of the camera is, nor how good an alignment has been achieved.



**Fig. 2.** Motivation experiment: the physically-based decentered dome port camera model and the underwater perspective approximation model are compared numerically. 30 virtual planar chessboard images with additive Gaussian noise of 0.2 pixels are simulated to perform calibration on both models, and a ground-truth 3D point cloud with randomly generated points (distances varying from 1m–10m) is used for evaluating the calibration quality by measuring the mean absolute reprojection error of the point cloud. Top row, decentering in forward–backward direction. Bottom row, decentering in left–right direction.

different optical densities that cause large direction changes of rays.

In summary, in practice dome port cameras are often successfully modeled by the pinhole camera system and very special setups of domes (e.g. stereo camera in a single dome) needed very special treatment that does not directly apply to the case of mildly decentered dome systems. In this contribution we therefore want to characterize the geometrical properties of decentered dome systems such as important axes, directions, symmetries, what kind of refraction patterns are to be expected and how refraction effects can be exploited to infer the physical parameters of the camera system.

Therefore, our contribution in this paper is twofold: In the next section, we derive the geometrical properties for decentered dome port systems and show that even thick domes are actually also axial cameras, similar to the findings of Treibitz et al. (2012), Agrawal et al. (2012) for flat interfaces. In contrast to flat interfaces, spheres have two intersections with the refraction axis, and we discuss their different properties. Second, we derive how to estimate the refraction axis from a single picture of a chessboard drawing on prior insights into radial distortion center estimation (Hartley and Kang, 2007). We also show that the apparent chessboard curvature (barrel vs. pincushion) is related to backward and forward decentering respectively and propose how to distinguish these cases in a 2D image. Afterwards, we analyze the forward projection of a 3D scene point onto the image plane given a calibrated decentered dome port camera system. We show that a 6th degree polynomial equation for thin domes can be derived by leveraging the property of the axial camera (Agrawal et al., 2010; Agrawal et al., 2012; Glaeser and Schröcker, 2000). Next, we derive an efficient optimization scheme to find the exact decentering from multiple underwater chessboard pictures, which provides a practical calibration procedure that does not require cumbersome in-air/underwater pair of the same chessboard. Finally, we discuss the limitation and the calibration accuracy when the decentering is not significant and propose an image pre-selection scheme to maximize the observable refraction effect to achieve a high calibration accuracy.

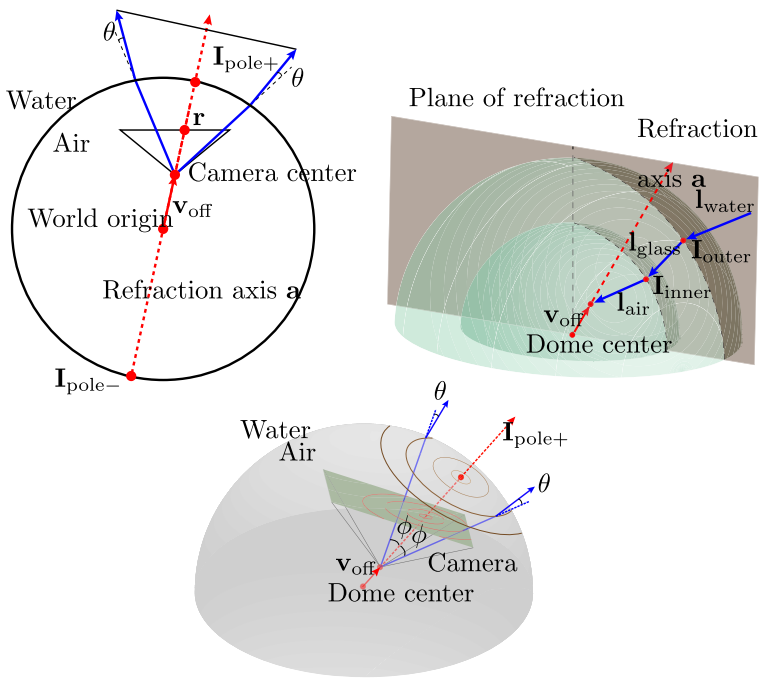
The remainder of this paper is structured as follows: In the following Section 3, we will discuss and derive the refractive dome geometry based upon which we propose the different novel steps of our calibration algorithm in Section 4. The single components and an overall calibration

approach are evaluated in Section 5 using synthetic and real datasets. Then, we discuss key findings and limitations in Section 6 before the conclusion in the final Section 7.

### 3. Decentered Dome Geometry

The setting is displayed in Fig. 3: We assume that a camera is positioned inside a sphere. The medium inside the sphere (e.g. air) has a different optical density  $\mu_{\text{air}}$  as compared to the medium outside the sphere (e.g. water, with density  $\mu_{\text{water}}$ ). The separating layer (e.g. glass) can either be considered of almost zero thickness (thin dome model), or, in particular for deep sea housings sustaining several hundred bars of pressure, can consist of several millimeters glass (thick dome model) with optical density  $\mu_{\text{glass}}$ . The exact numbers for the  $\mu$  parameters depend on the composition of the water, the exact “glass” material, but in the remainder of this paper we will assume  $\mu_{\text{air}} \leq \mu_{\text{water}} \leq \mu_{\text{glass}}$  to reason about inward- or outward bending. In case a pinhole camera<sup>1</sup> is positioned exactly in the center, all viewing rays will pass all dome layers in direction of the local normal and no refraction will occur. In practice, aligning the centers is very challenging, and therefore some decentering is likely to remain when assembling without visual feedback. The vector from the dome center to the camera center is the decentering offset vector  $\mathbf{v}_{\text{off}}$ , or short *decentering*. The line through the dome center and the camera center is the refraction axis with direction  $\mathbf{a}$ . It has two intersections with the (thin) dome surface which will be called the refraction poles, where we distinguish the pole closer to the camera center (positive refraction pole  $\mathbf{I}_{\text{pole}+}$ ) and the pole further away (negative refraction pole  $\mathbf{I}_{\text{pole}-}$ ). The refraction axis also intersects the

<sup>1</sup> In this contribution we consider the mathematical pinhole camera model, cf. to Hartley and Zisserman (2004). Real cameras use lenses to collect light over a larger area and need to be focused to the object under consideration. When used behind a dome port, the focusing behaviour of the lens depends on the medium outside the dome, and is different in air or in water (cf. to Menna et al., 2016). While focusing strategies are an interesting problem in their own right, in this contribution we consider the principal rays of the lens, and use the pinhole model to describe the mapping from world coordinates to image coordinates.



**Fig. 3.** Sketch of the dome refraction setup. Up left: The camera center is not located at the center of a thin dome (the origin). Still, the ray from the dome center through the camera center will not be refracted at the sphere. Another ray, going from the camera more to the left, will not continue straight (dashed line) but rather be refracted. The dotted line, the refracted line and the line from the origin to the dome center (the refraction axis) lie all in one plane. The upper right image shows the intersection of this plane with the sphere and for each viewing ray, a plane of refraction exists, even in case we use a thick dome. In the lower sketch, the dashed light paths form the same angle  $\phi$  with the refraction axis at the camera center. Consequently, they will be refracted by the same angle  $\theta$  at the interface. The cone of all those rays intersects the sphere in a 3D circle, and intersects the image plane in an ellipse.

image plane at the refraction center, which has the position  $\mathbf{r}$  in image coordinates.

We set the origin of the world coordinate system to the center of the dome. Further, we assume that the camera is calibrated, and omit the camera intrinsics for the sake of readability, so it can be described by the projection matrix

$$\mathbf{P} = (\mathbf{R} \mid -\mathbf{R}\mathbf{v}_{\text{off}}). \quad (1)$$

Because of refraction effects according to Snell’s law, light rays will change their direction at the interface between different media, unless they hit the interface at an angle coinciding with the normal of the intersection point.

Along the light path from an object in the water, through the glass dome and into the camera we consider 3 segments here, the water segment with light ray direction  $\mathbf{l}_{\text{water}}$ , the glass segment with direction  $\mathbf{l}_{\text{glass}}$  and the air segment with  $\mathbf{l}_{\text{air}}$  (see Fig. 3, Center). Note that incoming rays that travel along the refraction axis will not be refracted, as they hit the outer interface at  $90^\circ$  and then also the inner interface at  $90^\circ$  before they move towards the camera. We will now trace back a different light ray  $\mathbf{l}$  from the pinhole to its intersection  $\mathbf{I}_{\text{inner}}$  with the inner interface.

**Lemma 1.** *The surface normal  $\mathbf{n}_{\text{inner}}$  of the inner interface at  $\mathbf{I}_{\text{inner}}$  is a linear combination of  $\mathbf{v}_{\text{off}}$  and  $\mathbf{l}_{\text{air}}$ .*

**Proof.** The intersection point  $\mathbf{I}_{\text{inner}}$  must be somewhere on the ray starting from the camera center  $\mathbf{v}_{\text{off}}$  in direction of  $\mathbf{l}_{\text{air}}$ , consequently  $\mathbf{I}_{\text{inner}} = \mathbf{v}_{\text{off}} + \lambda\mathbf{l}_{\text{air}}$ . Since the surface normal at some point  $\mathbf{X}$  on an origin-centered sphere is simply  $\mathbf{X}/\|\mathbf{X}\|$ , we obtain  $\mathbf{n}_{\text{inner}} = \alpha_1\mathbf{v}_{\text{off}} + \alpha_2\mathbf{l}_{\text{air}}$ .  $\square$

**Lemma 2.** *All segments  $\mathbf{l}_s$  of a light path towards the pinhole, the normals at the inner and outer dome intersection point and the refraction axis all lie in one plane.*

**Proof.** From Lemma 1, it can be seen that  $\mathbf{I}_{\text{inner}}$  and the normal at this position are both linear combinations of  $\mathbf{v}_{\text{off}}$  and  $\mathbf{l}_{\text{air}}$ , consequently, they all lie in the plane spanned by  $\mathbf{v}_{\text{off}}$  and  $\mathbf{l}_{\text{air}}$ . According to Snell’s law, the glass segment direction  $\mathbf{l}_{\text{glass}}$  can be computed as a linear combination of the incoming direction  $\mathbf{l}_{\text{air}}$  and the local surface normal  $\mathbf{n}_{\text{inner}}$  (Glassner, 1989):

$$\mathbf{l}_{\text{glass}} = r\mathbf{l}_{\text{air}} + c\mathbf{n}_{\text{inner}} \quad (2)$$

where  $r = \mu_{\text{air}}/\mu_{\text{glass}}$ , and

$$c = \frac{-\mu_{\text{air}}\mathbf{l}_{\text{air}}^T\mathbf{n}_{\text{inner}}}{\mu_{\text{glass}}(\mathbf{n}_{\text{inner}}^T\mathbf{n}_{\text{inner}})} - \frac{\sqrt{\mu_{\text{air}}^2(\mathbf{l}_{\text{air}}^T\mathbf{n}_{\text{inner}})^2 - (\mu_{\text{air}}^2 - \mu_{\text{glass}}^2)(\mathbf{l}_{\text{air}}^T\mathbf{l}_{\text{air}})(\mathbf{n}_{\text{inner}}^T\mathbf{n}_{\text{inner}})}}{\mu_{\text{glass}}(\mathbf{n}_{\text{inner}}^T\mathbf{n}_{\text{inner}})}. \quad (3)$$

Consequently, following the same reasoning, also the outer interface point, its normal and the water segment are linear combinations of  $\mathbf{v}_{\text{off}}$  and  $\mathbf{l}_{\text{air}}$  and all lie in the same plane.  $\square$  In the flat refractive case (Agrawal et al., 2012), this plane is called the plane of refraction. We will now investigate by what angle a light ray changes its direction at the inner interface. We conceptually group light rays that start from  $\mathbf{v}_{\text{off}}$  and change their direction at the inner interface by some angle  $\theta$  into the set  $\mathcal{R}_\theta$  (see Fig. 3, Right). We call the corresponding set of image coordinates *iso-refraction-angle curves*.

**Theorem 1.** *Iso-refraction-angle curves are conic sections in the image.*

**Proof.** For symmetry reasons it can be seen that all rays starting at  $\mathbf{v}_{\text{off}}$  that form some angle  $\phi$  with the refraction axis (the  $\phi$ -cone) intersect the inner sphere in a 3D circle (see Fig. 3). They all form the same angle with the local normal at the sphere, and due to Snell’s law, they will thus all be refracted by the same amount. The intersection of the  $\phi$ -cone with the image plane forms a conic section.  $\square$

Finally, we will inspect the displacement induced by the different optical densities of the media. In case all media (air, glass, water) had the same refraction index, we could compute the image position  $\mathbf{x}_a$  of an observed 3D point  $\mathbf{X}$  simply by using the projection matrix:  $\mathbf{x}_a \simeq \mathbf{P}\mathbf{X}$ . This is the position where we would expect to see the point without a dome, just “in air”. In case the three media have different optical densities the light will undergo refraction and we will observe the same 3D point at another position  $\mathbf{x}_r$ . 3D points that lie on the refraction axis will not be refracted and rather be projected onto the refraction center  $\mathbf{r}$  in the image.

**Theorem 2.** *The “in-air” observation  $\mathbf{x}_a$  of a point  $\mathbf{X}$  and the underwater “refracted” position  $\mathbf{x}_r$  of that point  $\mathbf{X}$  form a line with the refraction center  $\mathbf{r}$ .*

**Proof.** From Lemma 2 it can be seen that all the segments of the light

path, including the 3D point X, are in a plane jointly with the refraction axis. The intersection of that plane with the image plane forms a line. Since the camera center is in that plane too, also the direct line from the camera center to X is in that plane, and thus the unrefracted "in-air" observation must be in the same intersection line of that plane with the image plane. □

This means that refraction happens actually along a line containing the unrefracted "in-air" observation and the refraction center (cf. also Fig. 5). Now we consider the overall dome camera system in the water as a special kind of camera. We simply extend the water segments that lead from 3D points to the outer sphere to (infinitely long) lines:

**Theorem 3.** *The decentered dome port camera system is an axial camera.*

**Proof.** From Lemma 2 it can be seen that each water segment of a light path is in a plane jointly with the refraction axis. The water segment is either parallel to the axis (intersection at an ideal point at infinity) or its extension will have a Euclidean intersection with the axis. Consequently, all water segments of paths reaching the pinhole, i.e. the viewing rays in water, intersect the axis and the overall system is an axial camera. □

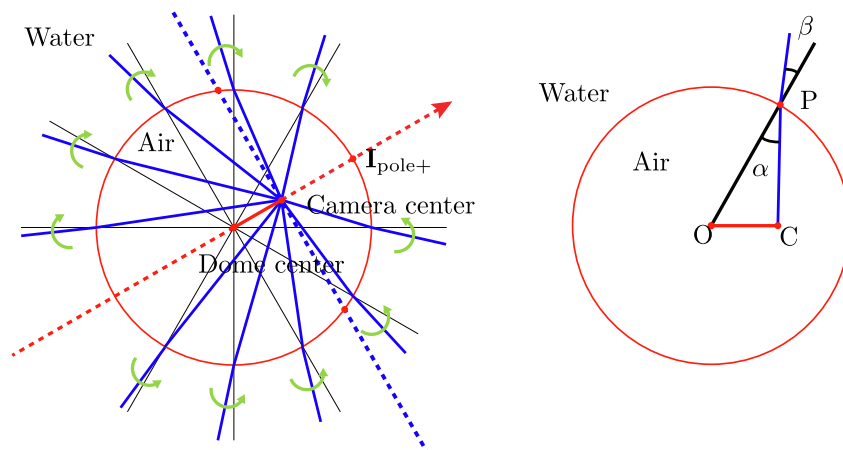
**Theorem 4.** *In the thin dome port camera system in water, all the rays from the camera are refracted towards the positive refraction pole.*

**Proof.** From Lemma 2 we know that all ray segments of a light path lie in a single plane with the axis, i.e. the incoming and outgoing ray at the refraction interface change direction inside this plane. According to Snell's law, when going from the lighter to the denser medium, rays change their direction towards the surface normal. Since all the surface normals cross the dome center, and since the camera is positioned closer to the positive refraction pole, the incident rays from the camera are thus always refracted towards the positive refraction pole (see also Fig. 4, Left). □

This means that in the image the 2D displacement direction with respect to the refraction center (inwards or outwards) depends on the sign of the 3D decentering vector and is the explanation of the barrel and pincushion distortion (depending on backward or forward decentering) known to underwater photographers and empirically reported e.g. in Menna et al. (2016).

**Theorem 5.** *In the thin dome port model the maximum change of direction (refraction at sphere) happens to the rays that approach the camera center inside the plane perpendicular to the refraction axis.*

**Proof.** Given a unit circle  $\odot O$ , a point C inside the circle has the distance  $k \in [0, 1]$  to the circle center O. A ray from C intersects  $\odot O$  at P with the incidence angle  $\angle\alpha$  and an outgoing angle  $\angle\beta$  (see Fig. 4, Right). The change of direction can be represented as  $\angle\text{diff} = \angle\alpha - \angle\beta$ .



**Fig. 4.** Left: Refraction of various blue rays from the camera center happens towards the positive refraction pole. Right: Maximum change of direction happens to the rays that enter the camera center from the plane perpendicular to the axis.

According to Snell's Law,  $n_{air}\sin\alpha = n_{water}\sin\beta$ , change of direction in range  $[0, \pi/2]$  can be rewritten to:

$$\angle\text{diff} = \alpha - \arcsin\left(\frac{n_{air}}{n_{water}}\sin\alpha\right), \quad \left(\alpha \in \left[0, \pi/2\right]\right) \tag{4}$$

Its first derivative is:

$$\frac{\partial\angle\text{diff}}{\partial\alpha} = 1 - \frac{1}{\sqrt{1 - \left(\frac{n_{air}}{n_{water}}\sin\alpha\right)^2}} \cdot \frac{n_{air}}{n_{water}}\cos\alpha = 1 - \frac{1}{\sqrt{\frac{n_{water}^2 - n_{air}^2}{n_{air}^2} \cdot \frac{1}{\cos^2\alpha} + 1}} > 0 \tag{5}$$

Since the derivative is strictly positive,  $\angle\text{diff}$  is monotonically increasing ( $\alpha \in [0, \pi/2]$ ) and does not have local maxima. Then the problem of finding a point P on the circle which has the largest changes of direction  $\angle\text{diff}$  is equivalent to find P which has the largest incident angle  $\angle\alpha$ .

Then, according to the Law of Sines,

$$\frac{OC}{\sin\alpha} = \frac{OP}{\sin\angle\text{POC}}, \tag{6}$$

since OC and OP are fixed, and  $\sin\alpha$  is monotonically increasing in the range  $[0, \pi/2]$ ,  $\angle\alpha$  must have its largest value when  $\sin\angle\text{POC}$  reaches its maximum 1. Since  $\angle\text{POC} \in [0, \pi]$ , it follows that  $\angle\text{POC} = \pi/2$ . Therefore, the maximum incident angle  $\angle\alpha$  on the circle happens when  $PC \perp OC$ .

Since, for the sphere all refractions happen in a plane of refraction, which always include the axis, we can subdivide the sphere surface into circles that include the poles, and in each of them consider the problem only as a 2F problem inside the specific plane of refraction. As shown above, in each of them the maximum change of direction happens perpendicular to the axis. □

This means that the ray that meets the camera center perpendicular to the refraction axis has suffered from the largest angular change (see Fig. 4, Left), whereas the ray on the axis is not at all refracted. This is an important finding for setting up experiments to observe or to calibrate the decentering. Note that the largest effect in the image also depends on the orientation of the camera, since the angular resolution of a pinhole camera increases towards the boundaries: Lateral (left–right, or up–down, in the camera coordinate system) decentering will provide a much clearer signal-to-noise ratio of refraction effects vs. corner detector uncertainty, as compared to forward–backward decenterings.

#### 4. Decentered Dome Calibration

In this section, we will derive a calibration procedure from the in-

sights of the previous section. First, we will present the geometrical considerations that allow directly inferring the refraction axis and a distinction between positive or negative direction decentering from one underwater photo showing refracted chessboard corners. The result can be used to initialize a  $v_{\text{off}}$ -optimization, using multiple images to actually measure the decentering. Throughout the calibration, we assume the camera intrinsics are known.

#### 4.1. Direct Estimation of the Refraction Center

We describe a corner's position on the original chessboard by  $\mathbf{x}_c$  (cf. to Fig. 5). When photographing a chessboard without refraction, the "as in air" image coordinates  $\mathbf{x}_a$  and the original chessboard pattern positions are related by a perspectivity (Hartley and Zisserman, 2004), a special kind of homography:  $\mathbf{x}_a \simeq \mathbf{H} \mathbf{x}_c$ .

Keeping the chessboard pose, now consider the camera being behind a dome port and that the entire system of camera, dome and chessboard is submerged in water (underwater in Fig. 5). Imagine a line  $\mathbf{q}$  through the unrefracted point in air  $\mathbf{x}_a$  and the refraction center:  $\mathbf{q} = [\mathbf{r}]_{\times} \mathbf{x}_a$ .

By Theorem 2, the refracted point  $\mathbf{x}_r$  must be somewhere on this line:  $\mathbf{x}_r^T \mathbf{q} = 0$ . If we now replace  $\mathbf{q}$  and  $\mathbf{x}_a$  we obtain a constraint that must hold between all points in the refracted image and their corresponding chessboard position:

$$\mathbf{x}_r^T \underbrace{[\mathbf{r}]_{\times} \mathbf{H}}_{\mathbf{F}} \mathbf{x}_c = 0 \tag{7}$$

This relation is reminiscent of epipolar geometry, where also all "displacement" (due to parallax) happens towards or away from the epipole. This principle has been exploited by Hartley and Kang (2007) for calibration of radial distortion. We use it in a similar way to find the refraction center, but working on 3D rays rather than 2D points, and we do not have radial symmetry with respect to displacement in the image. Note that (Agrawal et al., 2012) obtain an algebraically similar setting for refractive projection through flat interfaces. Essentially, as in epipolar geometry estimation, one can rearrange this equation using the Kronecker product and vectorization operator (Fusiello, 2007) to obtain constraints on the matrix  $\mathbf{F}$ :

$$\mathbf{x}_c^T \otimes \mathbf{x}_r^T \text{vec}(\mathbf{F}) = 0 \tag{8}$$

Basically, many of these equations can be stacked as for the eight-point algorithm for fundamental matrix estimation in order to estimate the vectorized matrix  $\mathbf{F}$ . After recomposition of  $\mathbf{F}$ , the refraction center  $\mathbf{r}$  can then be extracted as the left null vector of  $\mathbf{F}$ .

Since  $\mathbf{H}$  is actually a perspectivity, it would even be possible to use a 5-point algorithm for essential matrix estimation, but as we will use the refraction center only as an initial guess for subsequent optimization, and as many reliable correspondences are obtained using a chessboard detector, the normalized 8-point algorithm will be a good fit for our purposes (also avoiding the ambiguities of up to ten solutions). Multiple images of chessboards can be combined in the same way as described in Hartley and Kang (2007) for 2D radial distortion center estimation. Note however that according to Theorem 1 iso-refraction curves (with respect to refraction angle) are conic sections in the image and that the refraction effect in pixels is depth-dependent, so it cannot be described by radial distortion and underwater images cannot simply be "unrefracted" without 3D scene knowledge.

*Degenerate Cases* In case the camera is already perfectly centered, the second-smallest singular value of the resulting equation system will also be zero and no unique  $\mathbf{F}$  can be obtained. This corresponds to the case of no radial distortion in Hartley and Kang (2007) or fundamental matrix estimation in a planar scene and can easily be detected. Besides that the same algebraic conditions hold as for fundamental matrix estimation (e.g. number of correspondences, non-collinearity (Hartley and Zisserman, 2004)).

Now, using the camera orientation  $\mathbf{R}$ , the corresponding refraction

axis in world coordinates can be computed as:

$$\mathbf{a} = \frac{\mathbf{R}^T \mathbf{r}}{\|\mathbf{R}^T \mathbf{r}\|} \tag{9}$$

Note that so far, we have obtained the refraction axis, but the sign of the decentering along the axis (forward vs. backward) is still missing. When drawing a line through two refracted chessboard corner points  $\mathbf{x}_{r,1}, \mathbf{x}_{r,3}$  (using the cross product operator  $\times$ ), we can determine whether a third chessboard point  $\mathbf{x}_{r,2}$  in between them is projected onto the same side of their connection line as the refraction center by using the convexity test:

$$\text{convexity}(\mathbf{x}_{r,1}, \mathbf{x}_{r,2}, \mathbf{x}_{r,3}, \mathbf{r}) = \text{sign}\left(\left(\mathbf{x}_{r,1} \times \mathbf{x}_{r,3}\right)^T \mathbf{x}_{r,2} \cdot \left(\mathbf{x}_{r,1} \times \mathbf{x}_{r,3}\right)^T \mathbf{r}\right) \tag{10}$$

A positive sign in this equation relates to barrel distortion, whereas a negative sign determines pincushion distortion of the refracted chessboard corners. From Theorem 4 we can thus infer the sign of the decentering vector  $v_{\text{off}}$ . In most practically relevant cases one will have an estimate on the maximum expected decentering (e.g. 10% of the dome radius) allowing for a good initial guess for a later parameter optimization.

#### 4.2. Analytical Forward Projection for Thin Domes

Now, we derive the analytical forward projection for decentered thin domes, where there is only one spherical layer of refraction. By Lemma 2, we know that all segments of a light path from the camera center and the refraction axis all lie in a single plane, therefore, the refraction on the spherical layer can be analyzed on the plane of refraction, which is similar to the derivation for multi-layer flat interfaces (Agrawal et al., 2012). But the difference is that the plane of refraction is constructed by the refraction axis and the 3D scene point. Let 3D vectors  $z_1$  and  $z_2$  be the vertical axis and the horizontal axis of the 2D local coordinate system on this plane, and the spherical center be at the origin. Let  $z_1$  align with the refraction axis and point from the spherical center to the camera's optical center, as illustrated in Fig. 6. Then, the normal of the plane can be found by taking the cross product between 3D point  $\mathbf{X}$  and the vertical axis  $z_1$ . Afterwards, we can find the horizontal axis  $z_2$  as the cross product between  $z_1$  and the normal of the plane. Therefore, a point on the plane of refraction has a 2D Euclidean coordinate of  $(z_2, z_1)^T$ .

Assume the camera is decentered by a distance of  $d$ , thus the camera center has a coordinate of  $(0, d)^T$ . The 2D coordinate of the 3D scene point  $\mathbf{X}$  on the plane is given by  $\mathbf{x} = (u_x, u_y)^T$ , where  $u_x = \mathbf{X}^T z_2$  and  $u_y = \mathbf{X}^T z_1$ . The 2D refraction point  $\mathbf{m} = (x, y)^T$  on the spherical interface satisfies the 2D circle equation  $x^2 + y^2 = r^2$ , where  $r$  represents the radius of the sphere. The 2D normal on the interface at the refraction point can also be simply computed as  $n = -(x, y)^T / r$ . Now, the 2D ray shot from the camera is  $\mathbf{l}_0 = (x, y - d)^T$ , and the refracted ray  $\mathbf{l}_1$  can be computed by Eq. 2. The line joining the refraction point  $\mathbf{m}$  and the scene point  $\mathbf{x}$  should be parallel to the refracted ray  $\mathbf{l}_1$ , which provides another constraint. By removing the square root term and substituting  $x^2$  by  $r^2 - y^2$ , we end up with a 6th degree polynomial equation of a single variable  $y$ . Solving the equation results in up to 6 solutions, the correct solution can be found by checking Snell's law at the spherical refraction layer.

#### 4.3. Iterative Forward Projection for Thick Domes

As outlined in the previous section, the thin flat port projection equation by Glaeser and Schröcker (2000) of degree 4 becomes degree 6 for the thin dome port, essentially because of the quadratic nature of the refraction surface. For thick flat ports, (Agrawal et al., 2012) had derived

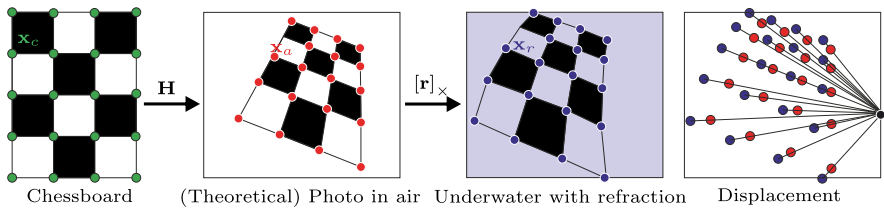


Fig. 5. Sketch of offset axis direction estimation principle: A chessboard (left sketch, corners  $x_c$  marked in green) is photographed from an oblique viewpoint (second image), the projected positions without refraction  $x_a$  marked in red. When submerging the camera and chessboard into the water behind a decentered dome, the light rays will be refracted (third image) and the projected positions  $x_r$  are marked in blue. The displacement vectors (right image) between corresponding  $x_a$  and  $x_r$  all intersect in a single point, which

corresponds to the refraction center  $r$ . This point can be estimated from correspondences of  $x_c$  and  $x_r$  (without knowledge of  $x_a$  !) similar to an epipole (where all epipolar lines meet).

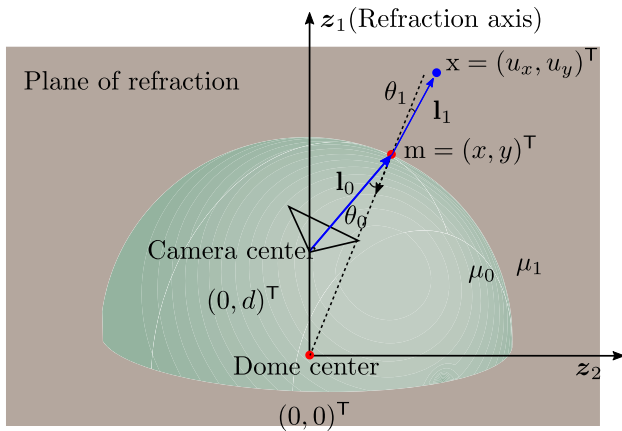


Fig. 6. Due to the property of the axial camera, the refraction on the spherical layer can be analyzed on the plane of refraction. In the local coordinate system of the plane of refraction,  $z_1$  aligns with the refraction axis and points from the spherical center to the camera center.

a 12th degree polynomial for the analytical forward projection. Both for this case, and also for the case of imaging through a solid glass sphere (degree 10, also by Agrawal et al. (2010)), only some extra constraints of the special setting helped to bring down the polynomial degree to 12 resp. 10. For the thick dome, we haven't found a similar extra constraint, but even if found, the quadratic nature of the refraction surface comes into play and it is likely that the degree of the polynomial will be significantly higher than 12. Solutions to high degree polynomials can become numerically unstable, and will also require iterative, numerical solvers. Consequently, for forward projection through thick domes we turn to the numerical approach as proposed in Kunz and Singh (2008) to find the projection by iteratively solving the inverse problem (back-projection), until the correct 2D point is found. Note that according to Theorem 2, the correct 2D point lies on the line joining the refraction center  $r$  and the "in-air" observation  $x_a$  of the 3D point. Rather than a 2D search, the search could now be restricted to the 1D line connecting  $x_a$  and  $r$ , which would simplify iterative forward projection.

#### 4.4. Decentering Estimation

Having obtained a good start value from the direct solver and convexity test, in this section we describe an optimization procedure to optimize the decentering vector using several images at the same time.

Different from the approach presented in She et al. (2019), this paper estimates the decentering vector  $v_{off} \in \mathcal{R}^3$  only from underwater imagery. Since there is no in-air photo (like in She et al. (2019)) to provide accurate pose estimation, the poses  $P_i$  of the chessboard images have to be estimated jointly with  $v_{off}$ . This results in  $6m+3$  parameters  $\Theta = (v_{off}, P_1, P_2, \dots, P_m)^T$ , where  $m$  defines the number of chessboard images and the estimation relies on  $2 \cdot m \cdot \#corners$  measurements. Assume Euclidean coordinates:  $X = (X, Y, 0)^T$ , and the measured  $i^{th}$  corner in  $j^{th}$

image is  $x_i^j$ . Then, we optimize the energy given by:

$$E(\Theta) = \sum_{i \in \Omega} \sum_{j \in \Phi} \|X_i^j - \hat{X}_i^j(x_i^j, v_{off}, P^i)\|^2 \quad (11)$$

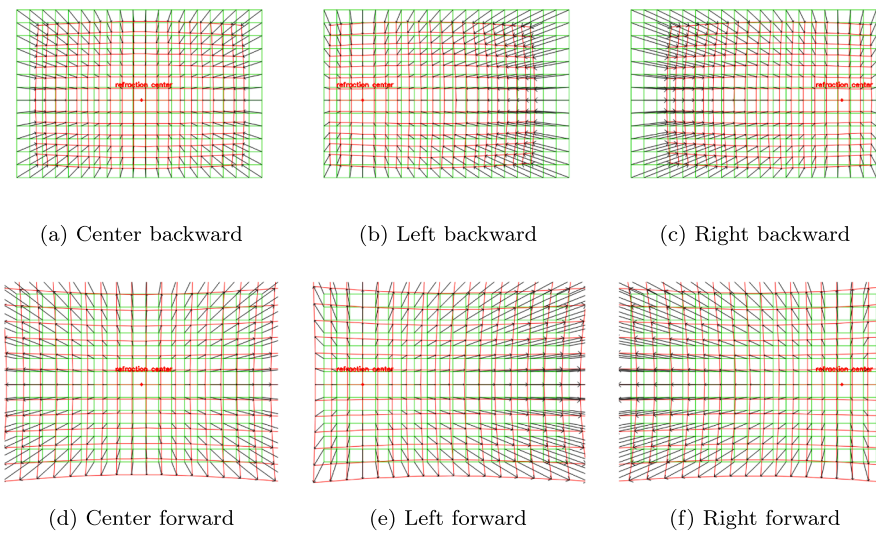
Herein,  $\hat{X}$  is the 3D point on the chessboard plane back-projected from  $x$ . Instead of minimizing the re-projection error, we back-project the corners detected in the images to the 3D chessboard, and the sum of squared differences is minimized inside the chessboard plane, since this is computationally much more efficient than doing the iterative forward projection within each optimization step.

#### 5. Evaluation

To validate the geometrical insights into the decentered dome and to evaluate the proposed decentering calibration approach, we have conducted three different types of experiments. The first type of experiment was performed using synthetic data, where we numerically simulated projections of 3D points by a decentered dome port camera system with perfectly known ground truth and noise models. This helps to understand performance and to look at the magnitude of effects. To take into account also effects as they occur in real images (corner detection problems, reflections, field of view issues and other physical limitations) we have also conducted real-world experiments with a deep sea dome port in a test tank. However, here evaluation becomes very indirect, as it is very hard to obtain ground truth information (in particular, real decentering), and experiments are sensitive to deformation of the tank due to the weight of the water, calibration uncertainties, inaccurate physical measurements of distances and many other effects. While all this will also occur in complex systems and real world applications, we think it is nevertheless important to isolate and understand the refraction effects. We have therefore put substantial effort into another step of evaluation that employs an open source ray-tracing toolbox (Community, 2018) to faithfully render images with ground truth settings for evaluation. In particular, we modeled a real deep sea dome port including all radiuses and materials in a virtual copy of our real water pool as realistic as possible and we have verified that we can accurately reproduce images taken by the real system. In this setup, we can control all physical parameters and understand effects in valid experiments.

##### 5.1. Synthetic experiments

First, to validate the proposed decentered dome geometry, we simulated the geometric projection of a dome port camera system with known decenterings. Here, the refraction displacement fields in the images were generated to visualize the theorems for the decentered dome geometry as shown in Fig. 7. The dome center in the local camera coordinate system was directly projected onto the image as the refraction center. In order to better visualize the displacement direction (inward vs. outward) with respect to the refraction center, the dome center was selected to be located in front of the camera center (Fig. 7, Top) and behind the camera center (Fig. 7, Bottom) respectively. The original patterns shown as green grids were created as if a chessboard was

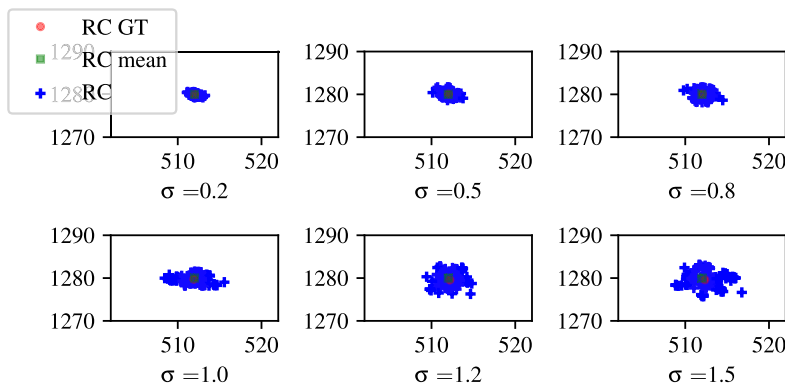


**Fig. 7.** Refraction displacement fields created by simulation. If a camera is decentered in a dome port, the red patterns show how the grids are refracted depending on the direction of the decentering. The black arrows point from the un-refracted grids towards the refracted grids. Top: the camera is decentered backward (along the viewing direction of the camera), the grids are refracted towards the refraction center. Bottom: the grids are refracted outwards from the refraction center because the camera is decentered forward.

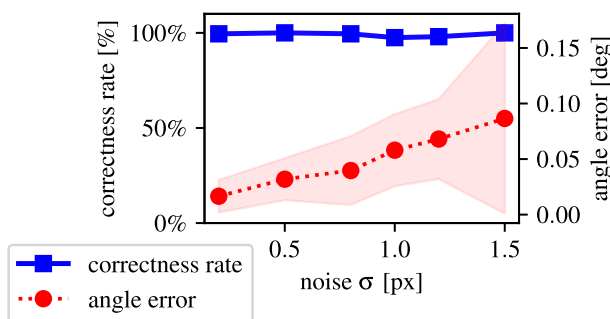
observed in air (not refracted). Then, the pattern was back-projected to the 3D space at 1m depth to obtain the pattern coordinates. Afterwards, we projected the pattern into the image considering the refraction effect, as if it was observed underwater, which is shown as red grids. As can be seen, the grids are refracted either towards or away from the refraction center depending on the direction of the decentering. In addition, the lines joining the un-refracted points and refracted points all intersect at the refraction center as exploited in the collinear constraint proposed in Theorem 2. Note that the refraction distortion pattern looks somewhat similar to the barrel/pincushion distortion, however, the refraction distortion center varies with the decentering direction, and the magnitude of displacement varies with the scene distance. Although it is not explicitly marked in the figure, it is also possible to identify the iso-

refraction-angle curves.

Next, to evaluate the effectiveness and stability of the proposed calibration approach, synthetic experiments for the refraction center estimation and determine the sign of the decentering from noisy input data. The calibration target was simulated as a  $7 \times 8$  planar chessboard with  $0.05m \times 0.05m$  square size. The virtual camera had a resolution of  $2048 \times 1536$  pixels and a field of view of  $90^\circ$ . The dome port was simulated with a radius of  $50mm$  and a thickness of  $7mm$  (which mimics real domes for 6 km ocean depth). Since the refraction center estimation and the decentering sign determination require a single image of the chessboard as input, we generated 10 random poses of the chessboard to be photographed, and it is guaranteed that the chessboard is at a pose



(a) Refraction center estimation



(b) Refraction axis direction estimation

**Fig. 8.** Synthetic evaluation of refraction center estimation and the distinction between positive and negative decentering along the axis using synthetic data. (a), each plot shows the distribution of the estimated refraction center from 200 trials in total for each noise level. (b), the red curve shows the average and standard deviation of the angle error between the estimated refraction axis direction and the ground truth direction; the blue curve shows the correctness rate of detecting the sign of the decentering.



that is inside the field of view of the camera. In order to evaluate the stability of the approach against small decentering offsets, which the refraction effect in the image is expected to be weaker in this scenario, we explicitly set the decentering vector as  $\mathbf{v}_{\text{off}} = (-0.001, 0.001, 0.002)^\top$  (all numbers in meter, if not stated otherwise) which is an extremely small decentering in the millimeter range. The 3D calibration targets are projected to the images using the iterative thick dome projection and zero-mean Gaussian noise with  $\sigma = 0.2, 0.5, 0.8, 1.0, 1.2, 1.5$  pixels is added to each projected corner. We performed 20 trials for each of the poses and recorded the estimated refraction center and the sign of the decentering, then we measured the angle error between the estimated direction and the ground truth direction and computed the correctness rate of the sign of the axis direction. The results are shown in Fig. 8, from which we can infer that the refraction center can be successfully estimated despite a high noise level. In addition, the sign of the decentering can also be correctly determined by the convexity test.

Then, we present the simulation experiments of the decentering vector calibration. Similar to the above experiment, the dome port system, the 3D chessboard and corner observations at different poses were simulated. The only difference is that the calibration approach is entirely based on multiple underwater images. We set the decentering vector as  $\mathbf{v}_{\text{off}1} = (0.0, -0.002807, -0.013)^\top m$  and  $\mathbf{v}_{\text{off}2} = (-0.001, 0.001, 0.002)^\top m$ . The first set aims to validate the effectiveness of the calibration approach and the second set aims to evaluate the stability in the case of small decentering offsets. The evaluation was repeated 20 times on each set, and 10 images were taken per trial. The experiment results can be found in Fig. 9, where the rotation angle error and the translation error show the differences between the estimated camera poses and the ground truth poses. As can be seen, the unknown parameters were computed with high accuracy even though a considerable amount of noise was added to the measurement. Note that the uncertainties of the estimated parameters in the second set are higher than the first set since the refraction effect is much weaker in the images, but the results can still be considered accurate.

## 5.2. Experiments on Rendered Images

To validate the imaging formation model of our dome port camera system and also to take imaging and corner detection effects into account in the evaluation, we employed the ray-tracing toolbox Blender<sup>2</sup> to render underwater imagery with physically plausible ground truth parameters. Blender is a physics-based ray-tracing software to create 3D worlds with realistic optical parameters for lighting and media, such as the index of refraction. It also allows setting virtual cameras to simulate capturing the scene into 2D images. To validate our camera model, we have modeled the water pool and the dome port into Blender mimicking the same setup as it exists in reality in our lab. This allowed us to identify discrepancies between real images and those from the simulation, and strongly supports that our simulated experiments are valid. The setup for rendering is visualized in Fig. 10. The dome port was modeled as concentric hemispheres with 7 mm thick interface material of borosilicate glass 3.3 (index of refraction = 1.473). Then the ray-tracer will track back light rays starting from the camera sensor and compute the light paths when traveling through the glass dome, thus, the refraction effect will be simulated separately for each pixel in the image. This is different from refraction simulation in rasterization-based renderers (e.g. Song et al., 2019) that apply refraction effects to the image after the actual rendering process. Instead, using ray-tracing, the refraction is handled naturally as can be seen in Fig. 11a, showing a chessboard half underwater and half in the air. When the camera is exactly centered in the dome (see Fig. 11a, Center), we can clearly see that there is no

refraction in the image. When pulling the camera away from the center, the lower part of the image is refracted inwards whereas it is refracted outwards when the camera is positioned in front of the dome center. Fig. 11b shows quantitatively the angular change of the viewing rays due to refraction at the dome sphere for each decentering case.

To validate the image formation model, we projected chessboard corners onto the images and compared their position against the rendered images by Blender and the real photos. The parameters for synthesizing were exactly the same as were used in rendering, among which, the camera intrinsics were obtained via calibrating the real camera in-air, and the poses of the chessboard where the real images were taken were computed via the decentering calibration. The resulting images are shown in Fig. 12, as can be seen, the synthesized chessboard corners, the rendered images and the real images match very well.

Next, we rendered 8 sets of images for different decentering situations and evaluate the refraction center estimation and the decentering vector calibration. The chessboard corners were automatically detected and the calibration results are shown in Table 1. Fig. 13 provides some of the resulting images, where the estimated refraction centers are marked as green spots, and the reprojections are marked as green crosses. The ATE (Average Transformation Error, ATE) in the table represents the average pose error over all used images in the translational components between the estimated camera poses and the ground truth poses, which is given as Sturm et al. (2012):

$$\text{ATE}_{\text{trans}} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left\| \text{trans} \left( \mathbf{P}_{\text{est},i}^{-1} \cdot \mathbf{P}_{\text{gt},i} \right) \right\|_2^2} \quad (12)$$

where  $\mathbf{P}_{\text{est},i}$  and  $\mathbf{P}_{\text{gt},i}$  are the estimated camera pose and the ground truth pose for the  $i^{\text{th}}$  image respectively.  $\text{trans}(\cdot)$  will extract the translational components from the transformation matrix. As can be seen, the system can determine the decentering with high accuracy. Note that in set +4, the camera is decentered horizontally, thus the refraction axis is parallel to the image plane and the refraction center is located at infinity.

## 5.3. Real-World Experiments

The proposed geometry insights and the calibration approach have been demonstrated on synthetic data and rendered images, but we finally want to validate them also in a real-world scenario. The experimental setup is shown in Fig. 10, top, the dome port with 50mm radius and 7mm thickness was attached to the sidewall of a test tank, then the tank was filled with water. The camera had a resolution of  $1280 \times 1024$  and a horizontal field of view of approximately  $73^\circ$ , then the intrinsic parameters of the camera were calibrated in air, also to obtain potential lens distortion. The camera was then attached to the dome port afterwards. Next, we have placed a planar checkerboard inside the pool, captured images and performed chessboard detection (OpenCV). The positions of the detected chessboard corners are then corrected for lens distortion, and are given, jointly with the known chessboard corner grid positions (cf. to Fig. 5)) to the eight-point algorithm for estimating the  $\mathbf{F}$  matrices. From each estimated  $\mathbf{F}$  the refraction center is extracted (for each of the images) without using any information about the index of refraction or thickness of the glass. Then, the sign of the decentering (along the refraction axis) was determined by the convexity test. Finally, non-linear optimization was performed to calibrate the decentering vector. The estimated decentering vector was  $\mathbf{v}_{\text{off}} = (-0.04, -0.02, -20.18)^\top mm$ , which is in agreement with our coarse manual measurements, the average refraction center was  $(641.56, 513.62)^\top$ , and the RMSE of the reprojection error was 0.32 pixels. Fig. 14a presents some exemplary images of reprojection results. Since measuring the actual offset between the camera center and the dome center is not trivial, we employed two strategies to validate the calibration results.

The first method is similar to the validation method suggested in

<sup>2</sup> <http://www.blender.org>

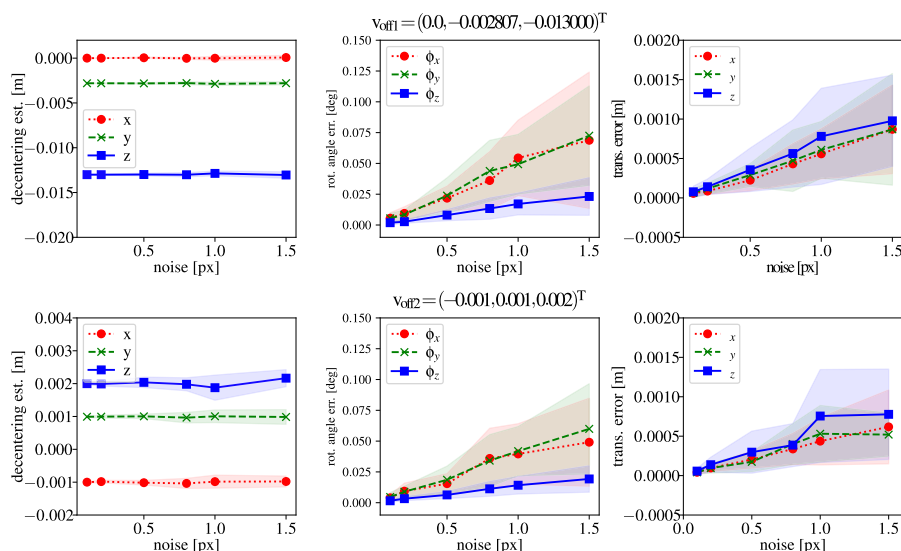
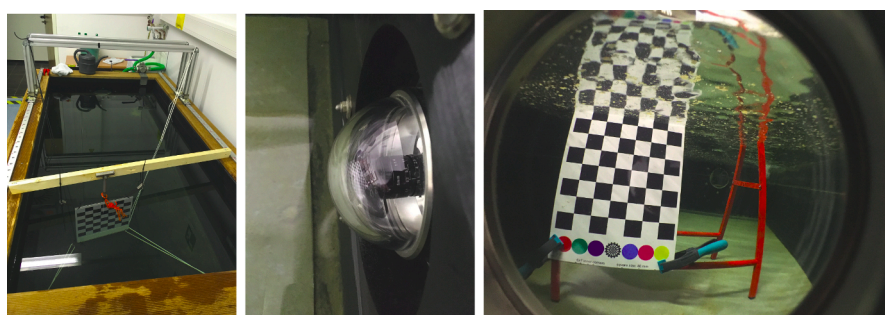
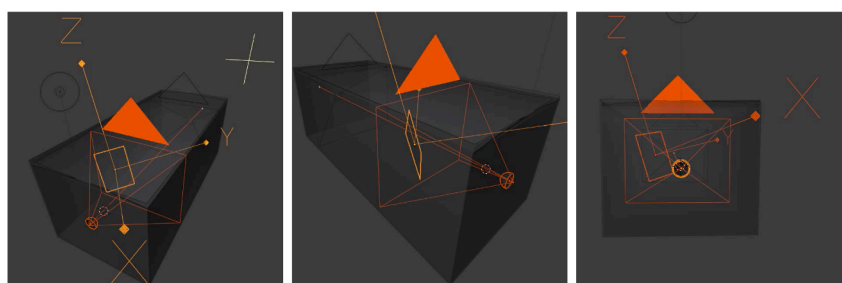


Fig. 9. Synthetic evaluation of decentering vector calibration in presence of a growing noise level. The plots from left to right: average and standard deviation of the estimated decentering vector, rotation angle error and translation error of the estimated camera poses.



(a) Real experimental setup



(b) Blender experimental setup

Fig. 10. Evaluation setup in real world and in Blender. The scene in Blender mimics the setup in the real world. The black cube represents the water pool, and the orange plane with a rectangular shape represents the planar checkerboard. The dome is modeled as concentric hemispheres and it is attached to the sidewall of the pool.

Peng and Sturm (2019), which is faced with a similar challenge. We separated our dataset into a calibration set and a validation set. The calibration set was used to perform the evaluation as described above whereas the latter one was used for validating the estimated decentering vector. For validation, the 4 outermost 3D-2D pairs of chessboard corners were utilized to compute the relative transformation of the camera with respect to the chessboard while keeping other parameters constant, then the remaining chessboard corners were projected and the average reprojection error was computed. The reason we used the outermost 4 corners for pose estimation is that the camera poses should be measured independently. There are 24 images for validation, thus there are  $24 \times$

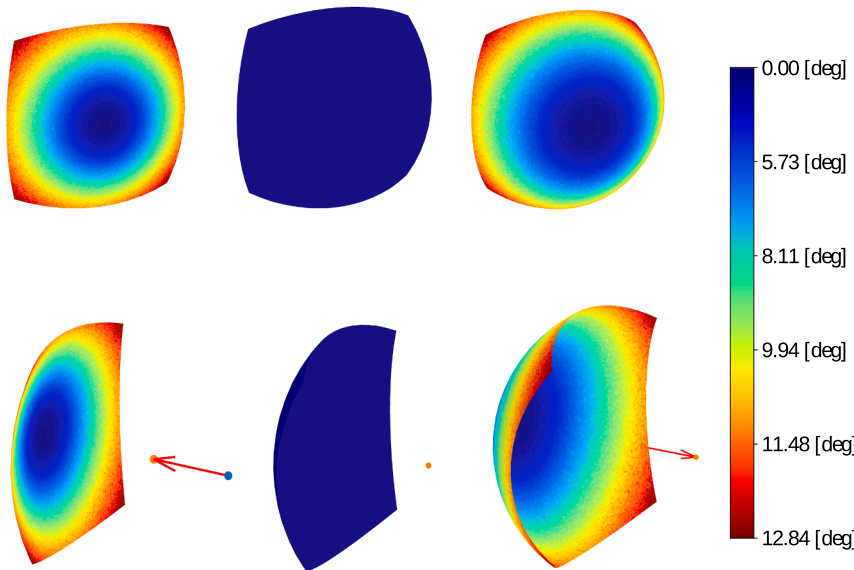
$52 = 1248$  points in total, and the average reprojection error was 0.611 pixels. Some example images are selected to be visualized in Fig. 14b. Since localization with only 4 pairs of correspondences is less accurate than using all pairs, thus, below 1 pixel validation error can already be considered as satisfactory, from which we conclude that the decentering was estimated accurately.

However, the relative transformation estimation could still partially remedy an estimation error in the decentering vector, therefore, we devised a second, complementary, validation experiment. We took a pair of photos of the chessboard underwater and in air, respectively, while keeping the chessboard position and orientation fixed. The

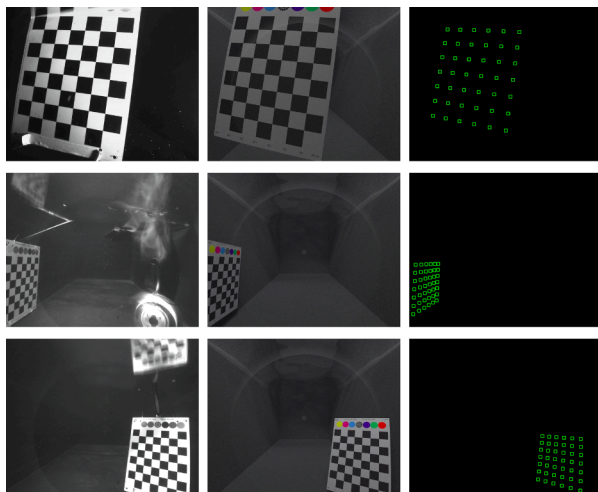


(a)

**Fig. 11.** (a) Half-air/half-underwater simulation, showing the different behaviours of the air-glass, glass-air interface as well as glass-water interface in different decentering settings. Images obtained with Blender (Community, 2018). (b) Refraction pattern shown as the angle between the incident light ray and the surface normal of the inner dome sphere plus the angle between the normal of the outer dome sphere and the outgoing ray. The lower row shows the same setting from another perspective, additionally indicating  $v_{off}$  as a red arrow. Images obtained with Mitsuba 2 (Nimier-David et al., 2019). Validation of the physically correct ray-tracing image renderer in different decentering settings. Left: Forward-decentering. Center: Perfectly centered. Right: Backward-decentering.



(b)



(a) Real (b) Blender (c) Ours

**Fig. 12.** Validation of the imaging formation model of the dome port camera system. From left to right: Real chessboard images, rendered chessboard images by Blender, synthetic chessboard corners by our thick dome port projection model.

experimental setup and the captured images are shown in Fig. 15. Given the calibrated dome port camera system, we first used the in-air photo for estimating the chessboard pose, which can be denoted as  $P_{air}$ . Next, we used the underwater photo to estimate the chessboard pose again, which we denote as  $P_{unw}$ . Theoretically, the two estimated poses should

**Table 1**

Evaluation results of decentering calibration on 8 sets of rendered images.

set	$r$ GT	$r$ Est.	–
•1	$(870.40, 921.60)^T$	$(884.85, 942.92)^T$	–
•2	$(1024, 768)^T$	$(1027.28, 774.88)^T$	–
•3	$(511.99, 1280.01)^T$	$(507.94, 1351.56)^T$	–
•4	$(1075.09, -1.2e + 08)^T$	$(1194.02, 286.07)^T$	–
•5	$(1024, 1342.88)^T$	$(1064.23, 1291.11)^T$	–
•6	$(1024, 989.10)^T$	$(1104.97, 906.42)^T$	–
•7	$(1183.69, 927.69)^T$	$(1187.7, 906.37)^T$	–
•8	$(652.80, 623.77)^T$	$(586.46, 565.22)^T$	–
set	$v_{off}$ GT [mm]	$v_{off}$ Est. [mm]	ATE [mm]
•1	$(-3.0, 3.0, 20.0)^T$	$(-3.06, 3.05, 20.39)^T$	0.61
•2	$(0.0, 0.0, 30.0)^T$	$(0.001, 0.09, 30.49)^T$	0.51
•3	$(-1.0, 1.0, 2.0)^T$	$(-1.09, 0.95, 1.66)^T$	0.78
•4	$(0.0, 2.81, 0.0)^T$	$(-0.08, 2.81, -0.26)^T$	0.48
•5	$(0.0, 2.81, 5.0)^T$	$(-0.10, 2.92, 4.77)^T$	0.69
•6	$(0.0, -2.81, -13.0)^T$	$(-0.16, -2.77, -13.48)^T$	0.90
•7	$(-2.81, -2.81, -18)^T$	$(-2.93, -2.75, -18.39)^T$	0.50
•8	$(0.42, 3.67, 28.39)^T$	$(0.46, 3.64, 28.43)^T$	0.50

be exactly the same as we assumed that the chessboard position and orientation had not changed while taking the image pair.

Consequently, we can measure the relative pose error between the two poses to validate the calibration results. The measured pose error in the translation component was 0.0107m. It should be mentioned that

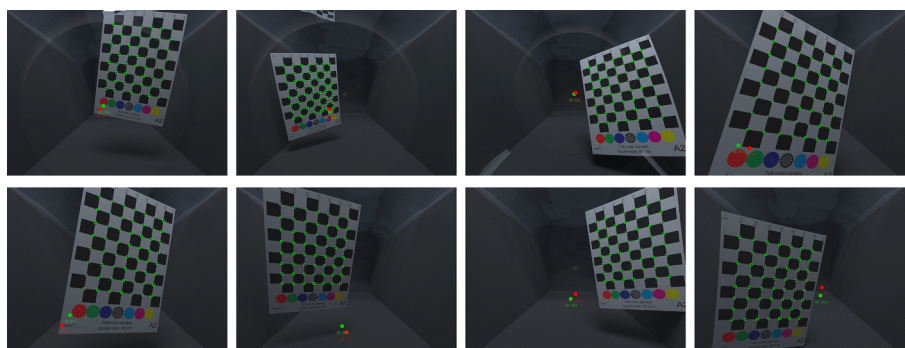
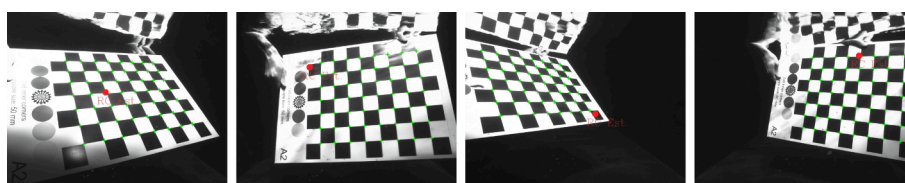
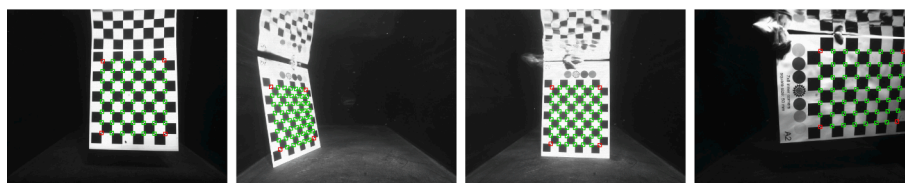


Fig. 13. Exemplary images showing the evaluation results of the experiments on the rendered images. The green squares and the green dots represent the reprojection results and the estimated refraction centers respectively. The ground truth refraction center is depicted as a red dot.

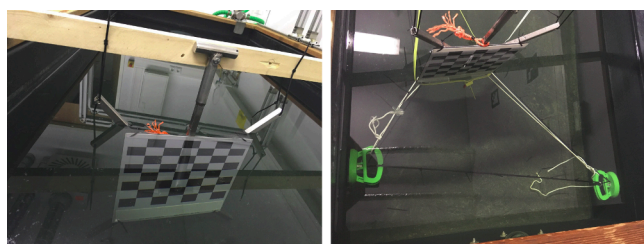


(a) Images used for evaluation

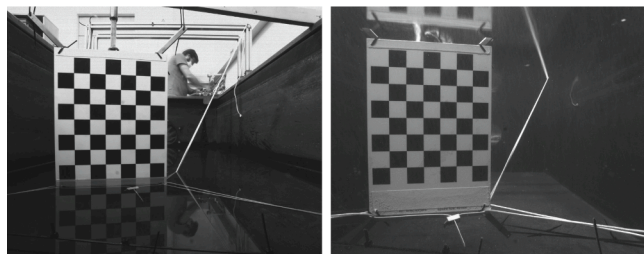


(b) Images used for validation

Fig. 14. Real-world evaluation results. (a), the results of the underwater decentering calibration in a water tank. (b), the validation results on the test set. The outermost 4 corners are marked as red squares, the inner reprojected corners are marked as green squares.



(a) Photos of the setup



(b) Chessboard images captured in the air and underwater

Fig. 15. Experimental setup for in-air/underwater image pair validation.

when computing the in-air chessboard pose, we considered the refraction pattern when light rays travel through the air-glass-air interfaces. Afterwards, we did the reciprocal reprojection validation, where we projected the 3D chessboard corners onto the underwater image using the in-air estimated pose  $P_{air}$  and then onto the in-air image using the underwater estimated pose  $P_{unw}$ .

The average corner displacement induced by the refraction effect between the in-air and underwater image is 25.24 pixels. When considering refraction, as obtained from the disjoint set of calibration images, the measured distance is reduced to 0.59 and 0.62 pixels, respectively, which can be explained by unavoidable non-rigidity of the setup when filling the tank with water and some other potential sources of uncertainty as we will outline in the discussion. The results are displayed in Fig. 16. Overall, the reduction of error and the previous validation show that the methods proposed are valid and can be applied in practice.

### 5.3.1. Practical Calibration of an AUV Camera

Finally, and this was also the motivation of this work in the first place, the insights and the developed techniques have been applied to a newly developed AUV camera for ocean research. This machine vision camera has a wide angle lens inside of a 50 mm dome port and a pressure housing. We have first calibrated the camera in air, before we applied the mechanical adjustment method as proposed in She et al. (2019) until no refraction effects were observable. Afterwards, the entire system was submerged in a tank and chessboard images were recorded. The images were undistorted according to the in-air calibration, then we apply the steps as reported in the previous sections. We obtain a calibration

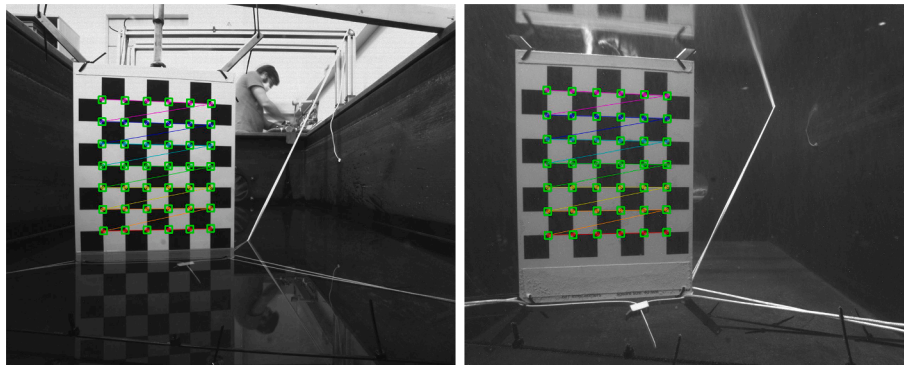


Fig. 16. Reciprocal reprojection validation results on the in-air/underwater image pair.

residual of 0.299 pixels for a decentering estimate of  $\mathbf{v}_{\text{off}} = (-0.36, 0.14, 0.27)mm$ , which means less than half a millimeter decentering in total. The camera system, sample calibration images with reprojected corners, as well as the AUV are displayed in Fig. 18.

6. Discussion

As the proposed geometrical insights and the decentering calibration approach rely on refraction effects caused by a decentered camera photographing through a spherical interface, in this section we would like to report some lessons learned.

*Homography Mapping Error.* In the classical approach of using a chessboard for calibration, which we also follow in this manuscript, we have to estimate a chessboard pose for each calibration image, on top of the actually desired refraction parameters. If we can directly compute a homography between an underwater image and the chessboard pattern, and the residual error for the homography is below the corner detector noise, this means that the center of refraction is effectively not observable in this calibration image, i.e. refraction effects are *drowned* in noise

(see Fig. 17b). Typically, this particular scenario happens when the decentering is very small or the chessboard corners in the image are located close to the refraction center (low refraction effects, see Section 3). Also, in case the refracted image point  $\mathbf{x}_r$  and the un-refracted image point  $\mathbf{x}_a$  are related by a 2D similarity transformation

$$\mathbf{x}_r \simeq \mathbf{H}_s \mathbf{x}_a, \tag{13}$$

then the overall relation between refracted image coordinates and 3D chessboard points becomes

$$\mathbf{x}_r \simeq \mathbf{H}_s \mathbf{H} \mathbf{x}_c \simeq \mathbf{H}' \mathbf{x}_c. \tag{14}$$

This means that the overall homography  $\mathbf{H}'$  can absorb all refraction effects and the equation system arising from Eq. 8 is underconstrained (ambiguous).

To detect and avoid these situations, we define the *Homography Mapping Error* as the residual of a homography mapping from the 3D points to the image points. We recommend users to monitor the Homography Mapping Error for each calibration image, and if it is low

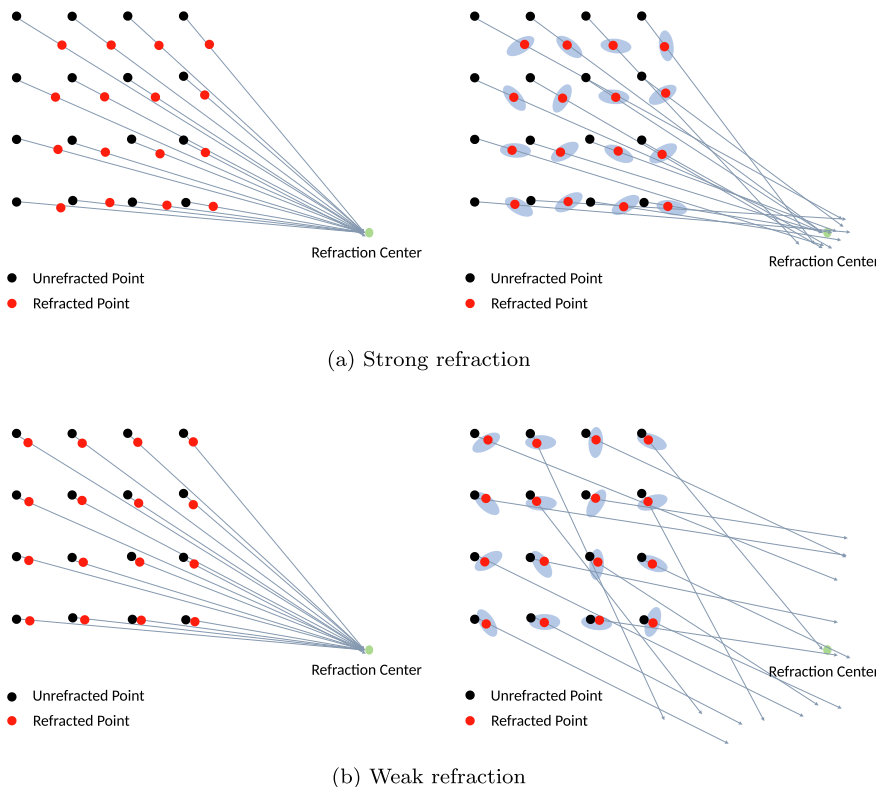


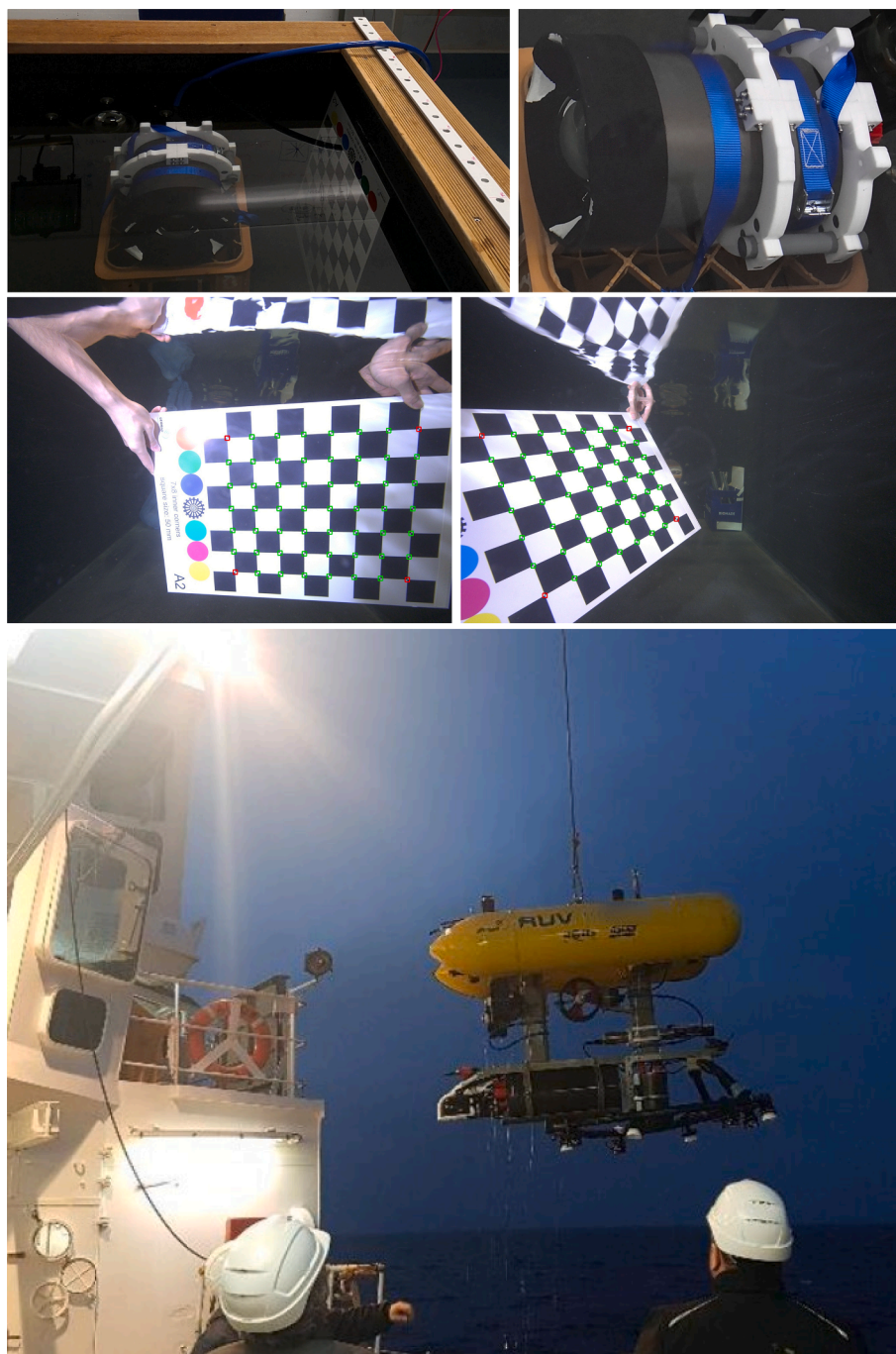
Fig. 17. (a), When there is strong refraction, noise does not break the refractive colinear constraint, and the Homography Mapping Error is high. (b), When the refraction is weak, noise can severely impair the refractive colinear constraint, which affects the calibration of the decentering, and the refraction center. In this case, the Homography Mapping Error is low. The gray ellipses represent the uncertainties of the chessboard corner detection. For illustration purposes, the uncertainties are exaggerated.

(compared to the noise level), to place the chessboard away from the refraction axis where we can see stronger effects (larger signal-to-noise-ratio). Additionally, the *Homography Mapping Error* can also be used to pre-select or to weight images for decentering calibration.

*Non-Single-View-Point Camera.* The geometry in this paper is analyzed based on the assumption that the camera has a single center of projection. For non-single-view-point lenses such as fisheye lenses, there is no perfect pinhole but a caustic where all light rays pass through. Centering such a lens with the dome port is in principle difficult as we can only bring one point to the dome center.

*Uncertainty of Intrinsic Calibration.* It is well known that the principal point of a camera is difficult to observe in air (Pollefeys et al., 1999; Hartley and Kaucic, 2002). Conversely, this also means of course that it

does not have a big impact on 3D mapping in air. Underwater however, the situation is different: Since the principal point is the intersection point of the optical axis with the image plane, the ray direction and refraction for every pixel will be impacted once the principal point changes, creating a different refraction pattern. Chessboard-based experiments to calibrate both the camera intrinsics and the decentering jointly failed, although we can achieve a lower calibration residual, the estimated parameters were off from the ground truth. Those parameters are correlated and probably more powerful (non-flat) calibration targets are required to better constrain the principal point. As in this contribution we discuss refraction effects, calibration of intrinsics is beyond the scope of this paper, but high-precision principal point calibration using underwater refraction could be an interesting option for future



**Fig. 18.** Top: Camera detached from AUV, calibrated in a tank. Center: Sample images from refraction calibration. Bottom: AUV with camera shown after successfully completing visual mapping mission.

research.

**Advantage over In-Air/Underwater Pair Calibration.** In order to obtain independent measurements of the decentering we have performed the method proposed in She et al. (2019) that is based on a pair of images, one taken in water, the other one in air, but with the same pose. While this method works in principle, the accuracy is limited by physical constraints.

When trying to bring the air and water pair into as good as possible agreement, we realized that it was extremely difficult to keep the chessboard steady when capturing the in-air/underwater image pair during the experiment. When the water is injected into the pool, waves are generated at the water surface, making the checkerboard unstable. In addition, we also found that the water pool deforms slightly when carrying more water due to the increasing pressure. Consequently, we had to spend substantial effort to keep the chessboard steady. As can be seen from Fig. 15a, we first firmly attached the chessboard to a metal frame and then attached them together on a wooden bar on top of the pool. Next, nylon ropes were used to connect the chessboard with the sidewall of the pool. Nevertheless, we still believe that there is half a pixel error due to non-rigidity of the setup. We therefore conclude that the in-air/underwater pair method is limited in the achievable accuracy. The pure underwater calibration suggested in this paper is more practical, as it can be cumbersome in some cases to bring the chessboard and camera (e.g. attached on a robot) from water into air without changing relative pose, while in the proposed method an underwater camera just has to be submerged, as shown in Fig. 18 for a camera of an AUV. Other cameras as shown in Fig. 1 can also be calibrated using this method even in the ocean.

## 7. Conclusion

In this paper, we have presented new insights into refraction effects caused by a decentered camera behind a spherical window. Somewhat similar to the flat refractive geometry case, the overall system acts as an axial camera, but here the axis intersects the sphere in two different poles that determine barrel- or pincushion-like refraction effects. Refraction happens along a line that connects the refraction center with the unrefracted projection, reducing the 2D search for projection to a 1D search.

It was then shown how to directly estimate the refraction center from a single underwater chessboard image, and that it is possible to infer the decentering from underwater chessboard images only. We then presented a novel practical approach to underwater calibration of dome port camera systems.

The approach can be used without a special pool with windows at the water level and does not need an in-air/underwater image pair, facilitating calibration also for bulky systems like submersibles. The results obtained by our method are not only relevant for adjusting, but the remaining decentering offset can be considered when using the camera in practice, e.g. similar to refractive structure from motion with flat ports as proposed by Jordt et al. (2016).

Later, evaluations have demonstrated the presented geometry insights and the results have shown the effectiveness and accuracy of the approach. But the error analysis was mostly performed in 2D image space, therefore in future works, we strive to integrate decentered dome port cameras into the state-of-art structure from motion pipeline to enable 3D reconstruction using dome port cameras, to facilitate a more detailed 3D error analysis, and also compare uncertainties and performance of different camera models.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgment

The authors would like to thank Dr. Amit Agrawal for offering generous help when deriving the analytical forward projection for the thin dome model. The authors also would like to thank Mareike Kampmeier for providing the AUV photos. This publication has been funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) Projektnummer 396311425, through the Emmy Noether Programme. Last but not least, the authors are also grateful for support from the Chinese Scholarship Council (CSC) for Mengkun She and Yifan Song.

## References

- Agrawal, A., Taguchi, Y., Ramalingam, S., 2010. Analytical forward projection for axial non-central dioptric and catadioptric cameras. In: European Conference on Computer Vision. Springer, pp. 129–143.
- Agrawal, A., Ramalingam, S., Taguchi, Y., Chari, V., 2012. A theory of multi-layer flat refractive geometry. In: CVPR, 2012.
- Bosch, J., Gracias, N., Rida, P., Ribas, D., 2015. Omnidirectional underwater camera design and calibration. *Sensors* 15 (3), 6033–6065.
- Chari, V., Sturm, P., 2009. Multiple-View Geometry of the Refractive Plane. In: Cavallaro, A., Prince, S., Alexander, D.C., (Eds.), BMVC 2009 - 20th British Machine Vision Conference, The British Machine Vision Association (BMVA), London, United Kingdom, 2009, pp. 1–11. doi:10.5244/C.23.56. URL <https://hal.inria.fr/inria-00434342>.
- B.O. Community, 2018. Blender - a 3D modelling and rendering package, Blender Foundation, Stichting Blender Foundation, Amsterdam. URL <http://www.blender.org>.
- Drap, P., 2012. Underwater photogrammetry for archaeology. In: da Silva, D.C. (Ed.), Special Applications of Photogrammetry, IntechOpen, Rijeka, 2012, Ch. 6. doi: 10.5772/33999. doi: 10.5772/33999.
- Drap, P., 2012b. Underwater photogrammetry for archaeology. Special applications of photogrammetry 114.
- Figueira, W., Ferrari, R., Weatherby, E., Porter, A., Hawes, S., Byrne, M., 2015. Accuracy and precision of habitat structural complexity metrics derived from underwater photogrammetry. *Remote Sensing* 7 (12), 16883–16900.
- Fryer, J.G., Fraser, C.S., 1986. On the calibration of underwater cameras. *Photogram. Rec.* 12, 73–85.
- Fusiello, A., 2007. A matter of notation: Several uses of the kronecker product in 3d computer vision. *Pattern Recogn. Lett.* 28, 2127–2132.
- Glaeser, G., Schröcker, H.-P., 2000. Reflections on refractions. *Journal for Geometry and Graphics (JGG)* 4, 1–18.
- Glassner, A.S., 1989. An introduction to ray tracing. Elsevier.
- Hartley, R., Kang, S.B., 2007. Parameter-free radial distortion correction with center of distortion estimation. *IEEE Trans. Pattern Anal. Mach. Intell.* 29 (8), 1309–1321.
- Hartley, R.L., Kaucic, R., 2002. Sensitivity of calibration to principal point position. In: European Conference on Computer Vision. Springer, pp. 433–446.
- Hartley, R., Zisserman, A., 2004. Multiple View Geometry in Computer Vision (Second Edition), 2nd Edition. Cambridge University Press.
- Harvey, E.S., Shortis, M.R., 1998. Calibration stability of an underwater stereo-video system: Implications for measurement accuracy and precision. *Marine Technology Society journal* 32, 3–17.
- Iscar, E., Johnson-Roberson, M., 2020. Towards distortion based underwater domed viewport camera calibration. In: 2020 IEEE International Conference on Robotics and Automation (ICRA), IEEE, 2020, pp. 1896–1902.
- Jaffe, J.S., Moore, K.D., McLean, J., Strand, M.P., 2001. Underwater optical imaging: Status and prospects. *Oceanography* 14.
- Jordt, A., Köser, K., Koch, R., 2016. Refractive 3d reconstruction on underwater images. *Methods in Oceanography* 15–16, 90–113. <https://doi.org/10.1016/j.mio.2016.03.001>. URL <http://www.sciencedirect.com/science/article/pii/S221122015300086>.
- Kotowski, R., 1988. Phototriangulation in multi-media photogrammetry. *International Archives of Photogrammetry and Remote Sensing* 27 (B5), 324–334.
- Kunz, C., Singh, H., 2008. Hemispherical refraction and camera calibration in underwater vision. In: OCEANS 2008, IEEE, 2008, pp. 1–7.
- Maas, H.-G., 1995. New developments in multimedia photogrammetry, Optical 3-D Measurement Techniques III.
- Maas, H.-G., 1995b. New developments in multimedia photogrammetry. In: Optical 3-D Measurement Techniques III. Wichmann Verlag, Karlsruhe.
- Maas, H.-G., 2015. On the accuracy potential in underwater/multimedia photogrammetry. *Sensors* 15 (8), 18140–18152.
- Menna, F., Nocerino, E., Fassi, F., Remondino, F., 2016. Geometric and optic characterization of a hemispherical dome port for underwater photogrammetry. *Sensors* 16 (1). <https://doi.org/10.3390/s16010048>. URL <http://www.mdpi.com/1424-8220/16/1/48>.
- Menna, F., Nocerino, E., Ural, S., Gruen, A., 2020. Mitigating image residuals systematic patterns in underwater photogrammetry. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLIII-B2-2020 977–984. <https://doi.org/10.5194/isprs-archives-XLIII-B2-2020-977-2020>.
- Menna, F., Nocerino, E., Troisi, S., Remondino, F., 2013. A photogrammetric approach to survey floating and semi-submerged objects. In: Videometrics, Range Imaging, and

- Applications XII; and Automated Visual Inspection, Vol. 8791, International Society for Optics and Photonics, 2013, p. 87910H.
- Moore, E.J., 1976. Underwater photogrammetry. *Photogram. Rec.* 8 (48), 748–763. <https://doi.org/10.1111/j.1477-9730.1976.tb00852.x>. URL <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1477-9730.1976.tb00852.x>. arXiv:<https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1477-9730.1976.tb00852.x>.
- Nimier-David, M., Vicini, D., Zeltner, T., Jakob, W., 2019. Mitsuba 2: A retargetable forward and inverse renderer. *Transactions on Graphics (Proceedings of SIGGRAPH Asia)* 38 (6). <https://doi.org/10.1145/3355089.3356498>.
- Nocerino, E., Menna, F., Fassi, F., Remondino, F., et al., 2016. Underwater calibration of dome port pressure housings. *International Archives of the Photogrammetry, Remote Sensing and Spatial. Inf. Sci.* 40, 127–134.
- Nocerino, E., Menna, F., Gruen, A., Troyer, M., Capra, A., Castagnetti, C., Rossi, P., Brooks, A.J., Schmitt, R.J., Holbrook, S.J., 2020. Coral reef monitoring by scuba divers using underwater photogrammetry and geodetic surveying. *Remote Sensing* 12 (18), 3036.
- Nocerino, E., Menna, F., Gruen, A., 2021. Bundle adjustment with polynomial point-to-camera distance dependent corrections for underwater photogrammetry. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLIII-B2-2021* 673–679. <https://doi.org/10.5194/isprs-archives-XLIII-B2-2021-673-2021>.
- Peng, S., Sturm, P., 2019. Calibration wizard: A guidance system for camera calibration based on modelling geometric and corner uncertainty. In: *Proceedings of the IEEE International Conference on Computer Vision*, 2019, pp. 1497–1505.
- Pollefeys, M., Koch, R., Van Gool, L., 1999. Self-calibration and metric reconstruction inspite of varying and unknown intrinsic camera parameters. *Int. J. Comput. Vision* 32 (1), 7–25.
- She, M., Song, Y., Mohrmann, J., Köser, K., 2019. Adjustment and calibration of dome port camera systems for underwater vision. In: *German Conference on Pattern Recognition*. Springer, pp. 79–92.
- Shmutter, L.B.B., 1967. Orientation problems in two-media photogrammetry. *Photogrammetric Engineering* 1421–1428.
- Song, Y., Köser, K., Kwasnitschka, T., Koch, R., 2019. Iterative refinement for underwater 3d reconstruction: Application to disposed underwater munitions in the baltic sea. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences XLII-2/W10* 181–187. <https://doi.org/10.5194/isprs-archives-XLII-2-W10-181-2019>.
- Sturm, J., Engelhard, N., Endres, F., Burgard, W., Cremers, D., 2012. A benchmark for the evaluation of rgb-d slam systems. In: *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, IEEE, 2012, pp. 573–580.
- Treibitz, T., Schechner, Y., Kunz, C., Singh, H., 2012. Flat refractive geometry. *IEEE Trans. Pattern Anal. Mach. Intell.* 34 (1), 51–65. <https://doi.org/10.1109/TPAMI.2011.105>.
- Verbiest, F., Proesmans, M., Van Gool, L., 2020. Modeling the effects of windshield refraction for camera calibration. In: *Vedaldi, A., Bischof, H., Brox, T., Frahm, J.-M. (Eds.), Computer Vision – ECCV 2020*. Springer International Publishing, Cham, pp. 397–412.
- Yang, W., Zhang, X., Ma, H., Zhang, G., Yang, G., 2021. Non-central refractive camera calibration using co-planarity constraints for a photogrammetric system with an optical sphere cover. *Opt. Lasers Eng.* 139, 106487.