ANALYTICAL FORWARD-PROJECTION AND SYSTEM CALIBRATION FOR IMAGING THROUGH REFRACTIVE WINDSHIELDS IN AUTOMOTIVE APPLICATIONS

Bashar Elnashef * Sagi Filin

Mapping and Geo-Information Engineering, Technion – Israel Institute of Technology, Haifa, Israel (basharnashef, filin)@technion.ac.il

Commission IIWG 1

KEY WORDS: Autonomous vehicles, Refraction, Pose estimation, System calibration

ABSTRACT:

We study in this paper models to mitigate effect of a windshield-related refraction on imaging systems. The literature shows that the distortions introduced by this curved surface are non-linear in their effect and reduce the performance of image-based analyses and depth estiamtion algorithms. We show that using geometric optics and local approximation of the windshield's surface to a spherical one, a direct analytical refractive forward projection (RFP) form of a 3-space point onto the image plane can be derived. Howevere, an exact form requires solving a 22-degree polynomial which may become numerically unstable. To stabilize the solution, we demonstrate how the introduction of valid local assumptions on the interface allows reducing the polynomial degree down to 8 and 4. Utilizing these forms we then show that the RFP can be used to jointly estimate the camera pose and the windshield's surface parameters through minimization of the reprojection error. The proposed models are tested on simulated data and validated on real-world observations. Results show stability and a sub-millimeter level of reconstruction accurecy, alluding to the validity and quality of our representations.

1. INTRODUCTION

The calibration of imaging systems has been the subject of extensive research as it sets the foundations of a variety of applications, including image-based scene analysis, multi-view 3-D reconstruction, structure-from-motion, visual odometry, and robotic navigation, to name only a few (Weilong and Lizuo, 2022). For autonomous driving applications, proper calibration of the imaging system is vital as it facilitates the localization, mapping, and estimation of distances and relative velocities of the surrounding traffic. In contrast with prototype versions of autonomous self-driving vehicles, where the cameras are mounted on the car roof, their placement in commercial vehicles has to be both functional and aesthetically pleasing, often being positioned behind the windshield (Hanel and Stilla, 2019). Handling refraction introduced by the windshield needs to address the non-linear ray trajectory and the surface model of the interface. The literature shows that while imaging through a flat refractive surface was researched extensively (Chadebecq et al., 2020), little can be found regarding the effect of curved surfaces, in general, and windshields, in particular (Hanel and Stilla, 2019). It has been demonstrated how distortions introduced by a windshield significantly affect the estimation of depth (Verbiest et al., 2020), and it has also been established that the refraction of the ray translates into nonlinear aberrations that are depth-dependent and are difficult to compensate for by direct image correction methods (Telem and Filin, 2010). In addition, Krebs et al. (2021), who analyzed the impact of these geometrical aberrations on classification tasks performed using state-of-the-art convolutional neural networks, demonstrated how the prediction accuracy was lowered by up to ten percent.

An explicit formulation of the refraction effect based on Snell's law was proposed by Chari and Sturm (2009), who analyzed

the multi-view relations when the scene contained a single refractive planar surface that separated two different media. Telem and Filin (2010) introduced an axial modification to the imaging system by developing a varifocal adaptation to handle the depth-dependent distortions introduced by refraction. This modified form was integrated into the collinearity equations and allowed the pose and system parameters to be estimated simultaneously. Agrawal et al. (2012) investigated the multi-layered flat-refractive geometry and demonstrated that such imaging systems correspond to an axial form. Elnashef and Filin (2022) developed a two-view flat-refractive model and proposed an orthogonality constraint that separates the rotational, and translational system-related components. In reference to non-planar interfaces, Kunz and Singh (2008) studied the effect of a misalignment between the camera and port centers when imaging through hemispherical housing ports in an underwater environment. They showed that the non-linearity of the ray's path requires an iterative nonlinear minimization to project a point from object-space onto the image. ? explicitly modeled the refraction through a curved surface. The authors approximated the refractive medium as a thick cone slice, where the inner and outer cones have the same aperture and the centers are such that the medium thickness is constant. To estimate the surface parameters, they used an elaborate minimization cost function, which was based on a backward-projection ray-casting model. In a later work, ? approximated the windshield surface by an ellipsoidal form. A similar minimization was used but the results were of numerical simulations only. Yoon et al. (2020) studied stereo-imaging through a refractive interface with an arbitrary surface form and derived a refraction-aware triangulation model for depth estimation. In order to estimate the surface parameters, targets were set on the inner refractive interface. Experimental results were compared to a baseline method that attempted to mitigate the refraction effect using lens distortions model. A clear improvement when using the refraction-aware triangulation model was reported. Verbiest et al. (2020) who

^{*}Corresponding author

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper. https://doi.org/10.5194/isprs-annals-V-2-2022-9-2022 | © Author(s) 2022. CC BY 4.0 License.

studied the effect of imaging through a windshield, demonstrated that the refraction effect leads to severe reconstruction errors when neglected. To reduce the complexity involved in modeling an arbitrary curved surface form, the authors used a local spherical surface approximation and applied a ray-tracing formulation to model rays trajectory through the system. They derived a polynomial solution as a function of the viewing angle, but no actual information on its degree, robustness, or performance was provided. An elaborate four-step procedure was proposed to estimate the system parameters, involving a non-linear minimization, which was followed by a refinement step using a local spline-based interpolation to accommodate for deviations from an ideal spherical shape. ? identified the axial nature of the imaging through hemispherical ports. An evaluation of the non-linear distortions was provided whose source was related to the camera displacement from the sphere center and the interface thickness. The authors developed a varifocal representation that followed the flat-refractive system modeling by Telem and Filin (2010). They then estimated the pose parameters and the decentering of their system.

As the literature shows, imaging through refractive surfaces introduces severe nonlinear distortions, and as the ray's path through the system becomes non-linear it also necessitates an iterative minimization procedure to project a point from object space onto the image plane (Verbiest et al., 2020). Considering the non-parametric geometric form of the windshield, modeling its consequent effect on the ray path and the pose parameters becomes involved, even when simplifying its shape. To alleviate these modeling-related complexities, we propose an alternative formulation of the refraction-related effect on the system. We show that even when modeling the surface by a simplified spherical form, a direct ray tracing formulation would translate to a 22-degree polynomial in terms of the image coordinates. This complex form is prohibitive for solving and has led many researchers to adopt complex optimization schemes. Instead, we show that with local approximations that have little effect on the outcome, the non-linear imaging system can translate into 8- or 4-degree polynomials for the forward projection. Our modeling facilitates an efficient and direct system calibration solution where all parameters can be estimated with high accuracy. Evaluations of our model demonstrate high levels of accuracy in the estimation of both pose and system parameters, robustness to high levels of noise, and a sub-millimeter level of accuracy in 3-D reconstruction.

2. THEORY

We consider an imaging system consisting of a camera mounted behind a thick windshield. The imaging system is of a threelayered air-glass-air form, where μ is the index of refraction of the windshield material. Without loss of generality, we set the index of refraction of the air to 1. Following Verbiest et al. (2020), we locally approximate the windshield surface to a spherical form. We define the axis between the camera and the sphere centers as the unit vector $\mathbf{a} = [0, 0, -1]^T$ and the distance between them as d (Fig. 1). Thus, the sphere and camera centers are set to $\mathbf{o} = [0, 0, 0]^T$, and $\mathbf{c} = \mathbf{o} + d\mathbf{a}$, respectively. When d = 0 the camera projection center and sphere center coincide and no refraction occurs. Defining the sphere thickness by d_s , the inner and outer radii become r and $r + d_s$, respectively. The image plane coordinates $\mathbf{x} = [x, y]^T$, are given in a calibrated camera frame so that an image ray would have the form of $\mathbf{v} = [x, y, f]^T$, and be measured in reference to the principal point, also the following correction to the lens distortion

effects. The angular offset between the optical axis and system frame can be expressed by two rotations, off and about the optical axis $\mathbf{M} = \mathbf{M}_z(\tau)\mathbf{M}_y(\eta)$. However, the rotation about the z-axis is strongly correlated with the camera rotation parameters (Telem and Filin, 2010), reducing the transformation to a single rotation, $\mathbf{v}_0 = \mathbf{M}_y(\eta)\mathbf{v}$. Following this transformation, the image ray in the system frame is given by $\mathbf{v}_0 = \mathbf{M}\mathbf{v}$. The ray sequence $\mathbf{v}_0, \mathbf{v}_1, \mathbf{v}_2$ from the camera center to an object-space point describes the light traversal through the system (Fig. 1), where \mathbf{q}_1 and \mathbf{q}_2 are the points of refraction on the inner and outer spheres, respectively. For each refraction, the intersection on the interface can be determined by a ray-sphere intersection. We define $\mathbf{Q}_{SU} = diag\{1, 1, 1, -1\}$ as a unit sphere where all points, **q**, lying on it fulfill the quadratic form $\mathbf{q}^T \mathbf{Q}_{SU} \mathbf{q} = 0$. The inner and outer spheres \mathbf{Q}_{Si} and \mathbf{Q}_{So} , can be regarded as a transformed version of \mathbf{Q}_{SU} by:

$$\mathbf{Q}_{Si} = \mathbf{T}_i^{-T} \mathbf{Q}_{SU} \mathbf{T}_i^{-1} \tag{1}$$

$$\mathbf{Q}_{So} = \mathbf{T}_o^{-T} \mathbf{Q}_{SU} \mathbf{T}_o^{-1} \tag{2}$$

where

Т

$$_{i} = diag\left(r, r, r, -1\right) \tag{3}$$

$$\mathbf{\Gamma}_{o} = diag\left(r + d_{s}, r + d_{s}, r + d_{s}, -1\right) \tag{4}$$

and where both spheres are centered at the origin $\mathbf{o}_i = \mathbf{o}_o = [0, 0, 0]^T$. To characterize a ray passing through that system we first seek the position on \mathbf{q}_1 . Defining an image ray by $\mathbf{q}_1 = \mathbf{c} + \lambda \mathbf{v}_0$, we wish to find its intersection point, \mathbf{q}_1 , by solving the quadratic form:

$$\mathbf{q}_1^T \mathbf{Q}_{Si} \mathbf{q}_1 = 0 \quad \rightarrow \quad \lambda_{i+,i-} \tag{5}$$

as a function of λ . We select λ_{i+} , the positive root, as the scene is defined in our system to lie in front of the camera. With the choice of λ_{i+} the first intersection point, \mathbf{q}_1 , is directly derived. In addition, as the center of the sphere coincides with the system origin, the normal to the sphere at \mathbf{q}_1 , is given by $\mathbf{n}_i = \mathbf{q}_1/|\mathbf{q}_1||$. As the ray propagates through the medium, it refracts. To compute the new direction, \mathbf{v}_1 , we use the vector form of Snell's law of refraction:

$$\mathbf{v}_1 = \xi \mathbf{v}_0 + \delta_1 \mathbf{n}_i \tag{6}$$

where $\xi = 1/\mu$, and:

$$\delta_1 = -\xi \mathbf{v}_0^T \mathbf{n}_i - \sqrt{\xi^2 \left(\mathbf{v}_0^T \mathbf{n}_i\right)^2 + (1 - \xi^2) \mathbf{v}_0^T \mathbf{v}_0} \qquad (7)$$

We compute \mathbf{q}_2 using the intersection of \mathbf{v}_1 and \mathbf{Q}_{So} , by first solving for λ_{o+} for $\mathbf{q}_2^T \mathbf{Q}_{So} \mathbf{q}_2 = 0$, where $\mathbf{q}_2 = \mathbf{q}_1 + \lambda_{o+} \mathbf{v}_1$. In addition, we have $\mathbf{n}_o = \mathbf{q}_2/r + d_s$. Finally, the direction of the object-space ray, \mathbf{v}_2 , is given by:

 $\mathbf{v}_2 = \mu \mathbf{v}_1 + \delta_2 \mathbf{n}_o$

where

$$\delta_2 = -\xi \mathbf{v}_1^T \mathbf{n}_o - \sqrt{\xi^2 \left(\mathbf{v}_1^T \mathbf{n}_o\right)^2 + \left(1 - \mu^2\right) \mathbf{v}_1^T \mathbf{v}_1} \qquad (9)$$

We have traced the ray trajectory from image-space to objectspace, and if the distance to the object-space point is known, its coordinates can be fully reconstructed.

(8)



Figure 1: Refraction through a spherical surface where all refractions take place in the plane of refraction.

2.1 Ray-tracing based system calibration by iterative forward projection (IFP)

The ray-tracing form sets the foundation for defining the cost function in reference to the camera pose and the system parameters. To begin, a set of N corresponding points $\{\mathbf{x}_i \rightarrow \mathbf{P}_i\}$, $\forall i = 1, \dots, N$ is given and our objective is to calibrate the system parameters and solve the camera pose, namely π = $\{\mathbf{R}, \mathbf{t}, d, r, d_s, \eta\}$ as the unknowns to solve for. The objective is to estimate them optimally by minimizing the reprojection error. A problem that arises in the way of realizing this is the unknown path of the forward projection from an object-space point to its image-space location. This is a well-known iterative procedure, in which the ray is first linearly traced to the image plane. The ray is then backward traced to object-space using our ray-tracing solution and by using the depth of the objectspace point. This iterative procedure continues until convergence is reached. Essentially minimizing

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\| \mathbf{P} - \hat{\mathbf{P}} \right\|^2 \tag{10}$$

where $\hat{\mathbf{P}}$ is the back-projected 3-D coordinates of the objectspace point, and $\hat{\mathbf{x}}$ is the corresponding image-space point. Next, given a set of N corresponding points, a calibration 10 step commences, as a function of the image-space reprojection 11 error:

$$\{\mathbf{R}, \mathbf{t}, d, r, d_s, \eta\}^* = \operatorname*{argmin}_{\{\mathbf{R}, \mathbf{t}, \mathbf{a}, d, r, d_s\}} \sum_{i=1}^N \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|^2 \quad (11) \quad \frac{13}{14}$$

In each iteration, all N 3-D points are first iteratively being projected to the image plane by using Eq. (10) and the estimated values of π . Then, Eq. (11) is used to estimate a correction to π . This nested minimization form has proven to limit the accuracy of the estimated parameters, introduce instabilities, and is also inefficient. It is common to find algorithms that fix some of the parameters as a means to stabilize the solution and secure 20convergence. Alternatively, we propose here an efficient solu-21 tion to project 3-D points directly onto the image plane, thus, eliminating the need for minimization in Eq. (10).

Forward projection through refractive quadratic sur-2.2 faces

Our objective is to alleviate the need for the iterative raytracing-based forward projection phase in the solution. To deAlgorithm 1: Polynomial forward-projection

Input: Object-space 3-D coordinate $\mathbf{P} = [X, Y, Z]^T$ **Output:** Forward-projected image-space point $\mathbf{x} = [x, y]^T$ **Data:** Pose (**R**, **t**) and system parameters $(\mathbf{a}, r, d, d_s, \mu)$ 1 $\mathbf{n}_P = \frac{\mathbf{a} \times \mathbf{P}}{\|\mathbf{a} \times \mathbf{P}\|}$ // Plane of refraction normal /* Define $[\hat{x},\hat{z}]\text{,}$ the coordinate system on the plane */

of refraction 2
$$\hat{\mathbf{z}}_2 = \mathcal{N}(\mathbf{a})$$

$$\hat{\mathbf{z}}_1 = \mathbf{n}_P \times \hat{\mathbf{z}}_2$$

4 $\mathbf{p} = \mathbf{RP} + \mathbf{t}$ // Transform an object-space point into the camera frame.

/* Project p into the plane of refraction, (u,v). $\hat{\mathbf{z}}_{2}^{T}\mathbf{p}$

s
$$v = \mathbf{z}_2 \mathbf{p}$$

6 $u = \hat{\mathbf{z}}_1^T \mathbf{p}$

Compute the polynomial M coefficients, k_1, \ldots, k_{M+1} and the roots of the polynomial $\{x_1, \ldots, x_M\}$.

$$\{x_1, \ldots, x_M\} = roots (k_1, \ldots, k_{M+1})$$
/* From the acquired six solutions, choose the m
real ones $\{x_1, \ldots, x_m\}$. Loop over all the real
 x_i values for $i = \{1, \ldots, m\}$ and choose the
solution that abides to Snell's law of
refraction.
*

9 for *i*:=1, *m* do

12

15

16

17

18

19

22

23

$$\left|\begin{array}{c} z_i = \sqrt{r^2 - x_i^2} \\ \mathbf{q}_{1i} = [x_i, z_i]^T \\ \mathbf{v}_{0i} = \mathbf{q}_{1i} - \mathbf{c} \qquad // \ \mathbf{c} = [0, d]^T \text{defines the camera} \\ \text{center of projection.} \\ \mathbf{n}_i = -\mathcal{N}(\mathbf{q}_{1i}) \\ \mathbf{v}_{1i} = \xi \mathbf{v}_0 + \delta \mathbf{n}_i \qquad // \ \text{Compute the refracted ray} \\ \text{direction.} \\ \mathbf{v}_{1i}^T \mathbf{Q}_{So} \mathbf{v}_{i1} = 0 \rightarrow \lambda_{+o} \\ \mathbf{q}_{2i} = \mathbf{q}_{1i} + \lambda_{o+} \mathbf{v}_1 \\ \mathbf{n}_o = -\mathcal{N}(\mathbf{q}_{2i}) \\ \mathbf{v}_{2i} = \mu \mathbf{v}_1 + \delta_{2i} \mathbf{n}_o \qquad // \ \text{Compute the refracted ray} \\ \text{direction.} \\ \overline{\mathbf{v}}_{2i} = [u, v]^T - \mathbf{q}_{2i} \\ \varepsilon = \frac{\mathbf{v}_{2i} \times \overline{\mathbf{v}_{2i}}}{\|\mathbf{v}_{2i} \times \overline{\mathbf{v}_{2i}}\|} \\ /* \ \text{If } v_{2i} \ \text{and } \overline{v}_{2i} \ \text{are parallel then} \qquad */ \\ \text{if } \varepsilon < 1e - 6 \ \text{then} \\ \\ \mathbf{v}_0 = x_i \hat{\mathbf{z}}_1 + (z_i - d) \hat{\mathbf{z}}_2 \qquad // \ \text{Compute the image} \\ \text{ray, } \mathbf{v}_0 \ \text{ in 3-D sapce} \end{aligned}$$

rive a direct form, observe that the surface normal at the point of refraction, the incident ray, and the axis defined by the sphere and camera centers form a coplanar triad that defines the plane of refraction (POR). In addition, all planes of refraction form a pencil of planes about an axis that passes through the camera and sphere centers, **a**, and we use this axis to define the system frame (Fig. 1). For convenience, we develop our model within the POR and for that define a planar reference frame within it. Consider the camera center to be positioned at a distance d from the center of the sphere and let $\hat{\mathbf{z}}_2 = \mathcal{N}(\mathbf{a})$ define the system axis, where N is a normalization operator. The orthogonal direction to $\hat{\mathbf{z}}_2$ can be defined by $\hat{\mathbf{z}}_1 = \mathcal{N}(\hat{\mathbf{z}}_2 \times (\hat{\mathbf{z}}_2 \times \mathbf{v}_0))$. Within this plane, the origin is at the sphere center, $\mathbf{o} = [0, 0]^T$ suggesting that the camera center coordinates are $\mathbf{c} = \begin{bmatrix} 0, d \end{bmatrix}^T$. The rays, v_1 , and v_2 along the light path can be represented by 2-D vectors $\mathbf{vp}_i = [\hat{\mathbf{z}}_1^T \mathbf{v}_i, \hat{\mathbf{z}}_2^T \mathbf{v}_i]^T, i = \{1, 2\}$ in the POR. At the inner interface, we set $\mathbf{q}_1 = [x, z]^T$, as the point of intersection with the sphere. This allows us to express \mathbf{vp}_0 and **n** at \mathbf{q}_1 by $\mathbf{v}\mathbf{p}_0 = \mathbf{q}_1 - \mathbf{c} = [x, y - d]^T$ and $\mathbf{n}_i = \mathcal{N}(\mathbf{q}_1)$. To compute \mathbf{q}_2 and \mathbf{n}_o the intersection between the ray $\mathbf{v}\mathbf{p}_1$ and the outer sphere needs to be computed, from which the values of \mathbf{n}_{oi} and then \mathbf{vp}_2 follow. Defining $\mathbf{p} = \mathbf{RP} + \mathbf{t}$ as an objectspace point transformed to the camera system frame, where **R**, $\mathbf{t} = [t_x, t_y, t_z]^T$ as the rotation matrix and translation vector, and $\mathbf{\bar{p}} = [\mathbf{\hat{z}}_1^T \mathbf{p}, \mathbf{\hat{z}}_2^T \mathbf{p}]$ we can write the constraint

$$\mathbf{v}\mathbf{p}_2 \times (\bar{\mathbf{p}} - \mathbf{q}_2) = \mathbf{0} \tag{12}$$

where \times denotes the cross product. As \mathbf{q}_1 is a point on the inner circle with radius r, we can write $z = \sqrt{r^2 - x^2}$. Substituting this form into Eq. (12) transforms the constraint into a function of only x. Assuming that the pose and system parameters are known, one can project a point **P** by solving a single variable constraint. However, because of the complex trajectory, this constraint contains a cascade of root terms, observable from the root terms in Snell's law of refraction, (Eqs. 7 & 9) and the two quadratic equations when computing the points \mathbf{q}_1 and \mathbf{q}_2 (Eq. 5). Agrawal et al. (2012) demonstrated that a solution for the flat-refractive system can be reached by removing the square-root terms through mathematical manipulation leading to root-free constraint with one unknown. The authors then demonstrated that for a 3-flat-layers system with two refractions, Eq. (12) correspondence to 12-degree polynomial. However, in our case and because of the higher dimensionality of the surface compared to the simple planar case, Eq. (12) yields a 22-degree polynomial with the monomials $[1, x, x^2, \dots, x^{22}]$. This form facilitates a forward projection of a point from objectspace onto the image directly for a noise-free system. However, simulations show that this projection is unstable when noise is introduced. To improve the stability of the polynomial, we propose two approximations to the ray-trajectory through the system, which we demonstrate are negligible in effect, and then validate their effect in our real-world experiments.

First, we observe that the change in the normal direction between \mathbf{n}_i and \mathbf{n}_o results in a change between the directions of the rays \mathbf{v}_0 and \mathbf{v}_2 which would be similar if $\mathbf{n}_i = \mathbf{n}_o$. Analytically, from Fig. (1), it can be shown that the angle between the two normals is given by:

$$\varphi = \frac{d_s}{r + d_s} \tan \theta_1 \tag{13}$$

where θ_1 is the refraction angle formed between the normal \mathbf{n}_i and \mathbf{v}_1 . For a windshield, the radius would be much larger than the glass thickness ($r >> d_s$). For example, a standard 1 : 100 ratio would amount to $\phi \approx 0.01 \tan \theta_1$ and for a camera with a 45° filed-of-view, the offset angle will not be greater than $\phi \approx [-0.01^\circ, 0.01^\circ]$. Hence, setting $\mathbf{n}_i = \mathbf{n}_o$ is a reasonable approximation with a byproduct of $\mathbf{v}_0 = \mathbf{v}_2$. The effect on Eq. (12) is vast, as the polynomial form of the forward projection reduces to the degree 8 with the monomials $[1, x, x^2, x^3, x^4, x^5, x^6, x^7, x^8]$. This form is stable and facilitates application for the system parameters estimation. Details on the polynomial coefficient can be found in Appendix (A.1).

Another viable simplification (see e.g., Yoon et al., 2020) is to consider the two interfaces to be locally parallel. Computing \mathbf{q}_2 entails solving a quadratic equation of the intersection between \mathbf{v}_1 and the outer sphere, \mathbf{Q}_{So} . For a windshield, these computations can be approximated by computing the distance to a parallel interface at a distance d_s as follows:

$$\lambda_{+o}^{approx} = \frac{ds}{\cos\theta_1} \tag{14}$$

Hence,

$$\mathbf{q}_{2}^{approx} = \mathbf{q}_{1} + \lambda_{+o}^{approx} \mathbf{v}_{1}$$
 (15)

With both approximations, Eq. (12) is further simplified, leading to a 4-degree polynomial, with the monomials $[1, x, x^2, x^3, x^4]$. This form is advantageous as it also allows computing the roots linearly in a closed form by applying Ferrari's solution (e.g., Kurosh, 1980). Details on the polynomial coefficient can be found in Appendix (A.2). Using both approximations translates to a local flat-refractive approximation with \mathbf{n}_i as the flat-surface normal. Furthermore, if the windshield surface is smooth and has sufficiently low curvature, the approximation errors are practically negligible, regardless of whether the medium is thick or thin. When the medium is sufficiently thin, the error caused by the parallel assumption is also negligible, also confirmed by our experiments. Given a known set of camera pose and system parameters, a point P can be forward-projected onto the image, \mathbf{x} , directly and linearly using either the 8- or 4-degree polynomials, respectively. The correct solution (root) is the one that best abides by Snell's law refraction. Algorithm (1) details the polynomial forward-projection procedure.

Calibrations Calibrating the system would require minimizing the function in Eq. (11). However, because of the large number of unknowns, the minimization may prove to be hard to solve. Instead of solving the entire set of parameters, which includes the camera pose and system parameters it was shown by Elnashef and Filin (2019) that the problem can be separated into two parts, refraction dependent and refraction independent parameters in a flat-refractive system. This model can be extended to the refraction through curved surfaces, as it is a coplanar constraint on rays within the POR. Harnessing this, we can solve \mathbf{R}, t_x, t_y linearly and independently from refraction and system parameters with as little as five coplanar points (e.g., checkerboard target). Thus, reducing the parameters to solve for to only six, including $t_z, r, d_s, d, \mu, \eta$.

3. EXPERIMENTS

To test the performance of our proposed calibration solution, experiments were carried out via simulated experiments and validated using real-world data. For the simulations, the gener-



Figure 2: 3-D model for or the windshield surface we used a publicly available open-source 3-D model of a Mazda MX-5 Miata, 2002 form https://grabcad.com/library/mazda-miatawindshield-2002.

ated data followed similar settings as the ones used in the realworld experiments. A Canon PowerShot G9 X camera model was used, with a 5472×3648 pixels frame, and a 2.41μ m pixel size, while the principal distance was set to 32 mm. The index of refraction was set to $\mu = 1.5$ for the glass-made windshield. All image points were perturbed by a random Gaussian noise with a standard deviation ranging from 0 to 2 pixels, and for each noise level, 100 trials were performed. As for the number of images used in each trial, though a single image suffices to estimate all of the system- and pose-parameters, for redundancy, ten images of a 2-D gridded target were used. The target comprised of 48 points forming a grid of 6×8 spaced by 2.5 cm with a total size of 17.5×22.5 cm. This was simulated while positioning at depths ranging between 0.5 and 5 m in front of the camera and rotating randomly by up to 20° in each test. We randomly placed the camera at a distance d_w from the windshield, at a predefined range [2, 30] cm along the **a**, while offaxis deviations were randomly set up to 5° off **a**. To generate simulated data, independent of our model, the ray-tracing model described in Kunz and Singh (2008) was used. For the windshield surface, we used a publicly available open-source 3-D model of the Mazda MX-5 Miata, 2002 (Fig. 2).¹

Validation metrics: We used the following metrics to evaluate our results: the *positional error* was defined as the relative difference between the known and estimated pose $\varepsilon_t = \|\mathbf{t}_{true} - \mathbf{t}_{estimated}\|/\|\mathbf{t}_{true}\|$, and the *rotational error* is measured by the minimal rotation angle required to go from the estimated rotation matrix to the actual one. For its computation, we note that for the 3-D rotation group $\mathbb{SO}(3)$ there is an intrinsic notion of distance, which we present here in a matrix form (Trefethen and Bau, 1997). Let **R** and $\hat{\mathbf{R}}$ be the true and estimated rotation matrices, respectively. The difference-rotation-matrix is defined by: $\delta \mathbf{R} \triangleq \mathbf{R} \hat{\mathbf{R}}^T$, and the angle of the difference rotation can be retrieved by the trace of $\delta \mathbf{R}$ such that: $\varepsilon_{\phi} = \cos^{-1} \left((1 - trace(\delta \mathbf{R}))/2 \right)$.

Pose estimation Plots of the estimation accuracy of the algebraic and reprojection-error minimizations are presented in Fig. (3). As shown, for a noise level ± 0.5 pixels and using a 2-D target, the positional error is as low as 0.075%, and the angular error is 43.2'' (arc seconds). Applying the optimal form improved the results to 0.018% and 10.8'' for the positional-and angular error, respectively. The quality of these results is comparable to the ones obtained in an in-air imaging setup. Additionally, the reprojection error exhibits a linear trend as a function of the noise level increase, an indication of the stability of both the linear and non-linear forms.

Calibration In the calibration process, the estimated pose parameters were fixed and we only solved the remaining system

parameters. To evaluate our proposed polynomial-based calibration, the 4- and 8-degree forward projection (FP) forms were compared to the iterative FP (IFP) counterpart (Sec. 2.1). Fig. (4) plots the results of estimating the system parameters as a function of Gaussian noise. Results show that using our solutions the errors of all the system parameter estimates are on a low fraction of a percent in reference to their actual values. This is while the IFP solution error increases sharply as more noise is added to the system, reaching more than 2% error in the estimation of the camera distance from the sphere center, the windshield thickness, and index of refraction, d, ds, and μ , respectively. For example, while using a 2-D target and a ± 0.5 pixel noise, an error of 0.016%, 0.009% and 0.625% for d were recorded using 4-, 8-degree FP, and the IFP solutions, respectively. Similarly, and for the same level of noise, an error of 0.015%, 0.021%, and 0.33% for d_s was recorded. In both d and d_s , an improvement by an order of magnitude or more was recorded when compared to the IFP solution. The refraction index, μ for the same noise level was estimated with an error of 0.021%, 0.022%, and 0.691%. When estimating the reprojection error, a linear trend was observed through the different levels of added noise, which indicates the stability of the estimation model (Fig. 5). Note that when comparing both 4and 8-degree FP forms comparable results were reached with the 4-degree FP being more efficient.

3.1 Real-world experiments

Real-world experiments were performed as a two-step calibration procedure. First, the camera was calibrated outside of the car and the intrinsic calibration parameters were estimated and corrected. Next, the camera was set in the car and in front of the windshield, and images of the checkerboard target were captured through the refractive surface (Fig. 6). To validate these results, a reconstruction experiment followed where our ability to restore intrinsic geometric properties within the scene was evaluated. For that, all 48 corners of the 2.5 cm-spaced, 17.5×22.5 cm, grid were measured (Fig. 6). Our intrinsic measures included the coplanarity of the points, collineation, parallelism, and orthogonality, where the measured grid lines facilitated our analyses. External metric measures of the grid dimensions provided a means to evaluate the metric accuracy. In this experiment, the images that were used for calibration were not used for the reconstruction. The working distance for the reconstruction was also different from that used for the calibration. The first test evaluated the coplanarity of the mapped points, and a plane that was fitted to them yielded a 0.08 mm deviation using our 4-degree FP model compared to 3.94 mm using the IFP solution. The collinearity test for points lying along straight lines, and as a consequence the removal of the refraction effect that bends them, yielded a mean deviation of 0.18 mm and a maximal deviation of 0.31 mm from collinearity compared to a mean deviation of 2.88 mm and a maximal deviation of 4.95 mm. Computation of the angles between parallel and orthogonal lines of our solution shows a mean absolute error of 12'' and 15'' in parallelism and orthogonality, respectively, whereas the application of IFP solution was of much lower quality, with 263" and 347". Finally, the dimensions of the target were computed by measuring the distance between edge points along the width and height of the target. The errors were 0.48 mm (0.27% of 175mm), and 0.79 mm (0.35% of 225 mm) with respect to the actual dimensions in our solution, again showing significant improvement compared to the IFPbased solution with 2.31 mm (1.32% of 175mm), and 3.29 mm (1.46% of 225 mm).

¹https://grabcad.com/library/mazda-miata-windshield-2002



Figure 3: Pose estimation error, using the refraction-invariant model by Elnashef and Filin (2019) with 2-D targets, as a function of increasing Gaussian noise. Normalized positional error (left); Angular difference between the estimated and true rotation matrix (center); Reprojection error (right).



Figure 4: Quality of estimation of system parameters as a function of increasing Gaussian noise with a 2-D calibration target. All evaluations are done for the proposed 4th and 8th degree polynomials and compared to the backward-projection model.



Figure 5: Evaluation of the calibration reprojection error estimation as a function of increasing Gaussian noise with a 2-D calibration target. All evaluations are done for the proposed 4- and 8-degree polynomials and compared to the iterative forward-projection solution.



Figure 6: Setups for the real-world experiments.

4. CONCLUSIONS

The literature has shown that refraction through a curved nonparametric interface surface introduces non-linear distortions. Attempts to model it has largely been based on the iterative forward projection ray-tracing model. This leads to an inaccurate system and pose parameter estimations, and even divergence in cases where all parameters are estimated simultaneously. Instead, we developed here a direct forward projection representation through a refractive spherical interface that yielded a 22degree polynomial form. However, with slight relaxation on the ray propagation model, this form was reduced to 4- and 8degree forms. Our experiments have demonstrated that these relaxations have little to no effect on the quality of the pose and system parameters. Also, the reconstruction accuracy was on a sub-millimeter level. This suggests that our simplified form offers a viable calibration and modeling framework for automotive applications.

REFERENCES

Agrawal, A., Ramalingam, S., Taguchi, Y., Chari, V., 2012. A theory of multi-layer flat refractive geometry. *IEEE Conference on Computer Vision and Pattern Recognition*, 3346–3353.

Chadebecq, F., Vasconcelos, F., Lacher, R., Maneas, E., Desjardins, A., Ourselin, S., Vercauteren, T., Stoyanov, D., 2020. Refractive two-view reconstruction for underwater 3d vision. *International Journal of Computer Vision*, 128(5), 1101–1117.

Chari, V., Sturm, P., 2009. Multiple-view geometry of the refractive plane. *BMVC 20th British Machine Vision Conference*, 1–11.

Elnashef, B., Filin, S., 2019. Direct linear and refractioninvariant pose estimation and calibration model for underwater imaging. *ISPRS Journal of Photogrammetry and Remote Sensing*, 154, 259–271.

Elnashef, B., Filin, S., 2022. Target-free calibration of flat refractive imaging systems using two-view geometry. *Optics and Lasers in Engineering*, 150, 106856.

Hanel, A., Stilla, U., 2019. Semantic road scene knowledge for robust self-calibration of environment-observing vehicle cameras. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLII-2/W16, 103–110. https://www.int-archphotogramm-remote-sens-spatial-inf-sci.net/XLII-2-W16/103/2019/.

Krebs, C., Müller, P., Braun, A., 2021. Impact of windshield optical aberrations on visual range camera based classification tasks performed by cnns. *London Imaging Meeting*, 2021number 1, Society for Imaging Science and Technology, 83–87.

Kunz, C., Singh, H., 2008. Hemispherical refraction and camera calibration in underwater vision. *OCEANS 2008*, IEEE, 1–7.

Kurosh, A. G., 1980. Higher algebra. MIR Publishers.

Telem, G., Filin, S., 2010. Photogrammetric modeling of underwater environments. *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(5), pp. 433–444.

Trefethen, L. N., Bau, D., 1997. *Numerical linear algebra*. 50, SIAM.

Verbiest, F., Proesmans, M., Van Gool, L., 2020. Modeling the effects of windshield refraction for camera calibration. *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part VI 16*, Springer, 397–412.

Weilong, H., Lizuo, J., 2022. On camera calibration and distortion correction based on bundle adjustment. *Advances in Guidance, Navigation and Control*, Springer, 429–441.

Yoon, S., Choi, T., Sull, S., 2020. Depth estimation from stereo cameras through a curved transparent medium. *Pattern Recognition Letters*, 129, 101–107.

A. POLYNOMIALS

Manipulating Eq. (12) to eliminate the square-root terms, we arrive at a 22-degree polynomial with the monomials $[1, x, x^2, \ldots, x^{22}]$. Although the solution is direct simulations showed that its unstable when noise are introduced to the system. To improve the stability of the polynomial, we propose two approximations to the ray-trajectory through the system.

A.1 Eight degree polynomial

When the windshield is modeled as a sphere, the radius is much larger than the interface thickness (r >> ds). Setting $\mathbf{n}_i = \mathbf{n}_o$ is a reasonable approximation with the byproduct of similar direction of both \mathbf{v}_0 and \mathbf{v}_2 . Given this, Eq. (12) can be simplified to a 8-degree polynomial with the monomials $[1, x, x^2, x^3, x^4, x^5, x^6, x^7, x^8]$. The coefficients of the polynomial are given by,

$$\begin{split} k_0 &= 4r^2 A_4^2 - A_5^2 \\ k_1 &= 2A_5 A_6 - 8r^2 A_2 A_4 \\ k_2 &= 4r^2 (2A_3 A_4 + A_2^2) - 2A_1 A_5 - 4A_4^2 - A_6^2 \\ k_3 &= 2A_1 A_6 + 8A_2 A_4 + 2A_5 A_7 - 4r^2 (2A_2 A_3 + 4d^2 u v A_4) \\ k_4 &= 4r^2 (A_3^2 + 4d^2 u v A_2) - 8A_4 A_3 - 4A_2^2 - 2A_6 A_7 - A_1^2 \\ k_5 &= 8A_2 A_3 + 2A_7 A_1 + 16d^2 u v A_4 - 16d^2 r^2 u v A_3 \\ k_6 &= 16d^4 r^2 u^2 v^2 - 16A_7^2 - 16d^2 u v A_2 - 4A_3^2 \\ k_7 &= 16d^2 u v A_3 \\ k_8 &= -16d^4 u^2 v^2 \end{split}$$

where

$$\begin{split} A_{1} &= r^{2}(d^{2}u^{2} - d^{2}v^{2}) + u^{2}A_{8} + v^{2}A_{8} + A_{10}^{2} + A_{11}^{2} + 2duA_{9} \\ A_{2} &= A_{10}A_{11} - uvA_{8} + dvA_{9} \\ A_{3} &= dvA_{10} - du^{2}A_{13} + dv^{2}A_{13} + duA_{11} \\ A_{4} &= A_{9}A_{11} + dr^{2}u^{2}A_{13} \\ A_{5} &= A_{9}^{2} - r^{2}(u^{2}A_{8} + A_{11}) \\ A_{6} &= 2A_{9}A_{10} - r^{2}(2dvA_{11} - A_{12}) \\ A_{7} &= 2duA_{10} + 2dvA_{11} - A_{12} \\ A_{8} &= d^{2}mu^{2} + d^{2}r^{2} - d^{2} + mu^{2}r^{2} + r^{4} - r^{2} \\ A_{9} &= -dur^{2} + du \\ A_{10} &= vr^{2} + d - v \\ A_{11} &= -ur^{2} + u \\ A_{12} &= 4duvA_{13} \\ A_{13} &= \mu^{2} + r^{2} - 1 \end{split}$$

Solving the polynomial leads to eight solutions, four which are real. To choose the actual one, we examine which of the solutions is physically correct and abide to Snell's law of refraction, leading to the minimal cost function.

A.2 Four degree polynomial

Given a thin windshield the position of \mathbf{q}_2 can approximated by computing the distance to a parallel interface at a distance d_s (windshield thickness) as follows. In addition to the $\mathbf{n}_i = \mathbf{n}_o$ assumption, Eq. (12) is simplified further leading to a 4degree polynomial with the monomials $[1, x, x^2, x^3, x^4]$. The coefficients of the polynomial are given by,

$$k_{0} = \frac{d^{2}r^{3}u}{r}$$

$$k_{1} = d^{2}A_{5} + A_{2} + A_{3}$$

$$k_{3} = -\frac{d^{3}r^{3}u}{r} - \frac{d^{2}r^{5}u}{r} - A_{4} - A_{1}$$

$$k_{4} = -r^{2}(A_{2} + A_{3}) - A_{5}A_{6} - dd_{s}r^{3}$$

$$k_{5} = r^{2}(A_{1} + A_{4}) + \frac{dr^{3}uA_{6}}{r}$$

This contribution has been peer-reviewed. The double-blind peer-review was conducted on the basis of the full paper. https://doi.org/10.5194/isprs-annals-V-2-2022-9-2022 | © Author(s) 2022. CC BY 4.0 License. where

$$A_{1} = 2ud^{2}\mu^{2}r^{4}$$

$$A_{2} = d^{2}d_{s}r$$

$$A_{3} = 2d\mu^{2}r^{2}A_{5}$$

$$A_{4} = r^{2}uA_{6}$$

$$A_{5} = dr^{2} - vr^{2} + dd_{s}rA_{6} = d^{2}\mu^{2}r^{2} + \mu^{2}r^{4}$$