PROCEEDINGS OF SPIE

SPIEDigitalLibrary.org/conference-proceedings-of-spie

Structured light sensor with telecentric stereo camera pair for measurements through vacuum windows

Beermann, Rüdiger, Quentin, Lorenz, Kästner, Markus, Reithmeier, Eduard

Rüdiger Beermann, Lorenz Quentin, Markus Kästner, Eduard Reithmeier, "Structured light sensor with telecentric stereo camera pair for measurements through vacuum windows," Proc. SPIE 11056, Optical Measurement Systems for Industrial Inspection XI, 1105614 (21 June 2019); doi: 10.1117/12.2526049

SPIE.

Event: SPIE Optical Metrology, 2019, Munich, Germany

Structured light sensor with telecentric stereo camera pair for measurements through vacuum windows

Rüdiger Beermann[a], Lorenz Quentin[a], Markus Kästner[a], and Eduard Reithmeier[a]

[a]Institute of Measurement and Automatic Control, Leibniz Universität Hannover, Nienburger Str. 17, 30167 Hannover, Germany

ABSTRACT

Within the Collaborative Research Centre 1153 *Tailored Forming* a process chain is being developed to manufacture hybrid high performance components made from different materials. The optical geometry characterization of red-hot workpieces directly after the forming process yields diverse advantages, e.g., the documentation of workpiece distortion effects during cooling or the rejection of deficient components in an early manufacturing state.

Challenges arise due to the high components temperature directly after forming (approximately 1000 °C): The applied structured light method is based on the triangulation principle, which requires homogeneous measurement conditions and a rectilinear expansion of light. This essential precondition is violated when measuring hot objects, as the heat input into the surrounding air leads to an inhomogeneous refractive index field. The authors identified low pressure environments as a promising approach to reduce the magnitude and expansion of the heat induced optical inhomogeneity. To this end, a vacuum chamber has been developed at the *Institute of Measurement and Automatic Control*.

One drawback of a measurement chamber is, that the geometry characterization has to be conducted through a chamber window. The sensors light path is therefore again affected - in this case by the window's discrete increase of refractive index, and also due to the different air density states at sensor location (density at ambient pressure conditions) and measurement object location (density at low pressure conditions). Unlike the heat induced deflection effect, the light path manipulation by the window and the manipulated air density state in the chamber are non-dynamic and constant over time.

The reconstruction of 3-D geometry points based on a structured light sensor measurement directly depends on the mathematical model of detection and illumination unit. The calibration routine yields the necessary sensor model parameters. The window light refraction complicates this calibration procedure, as the standard pinhole camera model used for entocentric lenses does not comprise enough degrees of freedom to adequately parametrize the pixel-dependent light ray shift induced by thick vacuum windows. Telecentric lenses only map parallel light onto a sensor, therefore the window induced ray shift is constant for all sensor pixels and can be directly reproduced by the so-called affine camera model.

In this paper, we present an experimental calibration method, and corresponding calibration data and measurement results for a structured light sensor with and without measurement window. The sensor comprises a telecentric stereo camera pair and an entocentric projector. The calibration of the telecentric cameras is conducted according to the well-known affine camera model. The projector is used as *feature generator* to solve the correspondence problem between the two cameras. The calibration data illustrates that the window refraction effect is fully reproduced by the affine camera model, allowing a precise geometry characterization of objects recorded through windows. The presented approach is meant to be used with the aforementioned vacuum chamber to enable a geometry characterization of hot objects at low pressure levels.

Keywords: 3-D measurement, triangulation, fringe projection, structured light, refractive index field, light deflection, optical inhomogeneity, measurement window

Further author information: (Send correspondence to Rüdiger Beermann)
Rüdiger Beermann: E-mail: ruediger.beermann@imr.uni-hannover.de, Telephone: +49 511 762 5816

Optical Measurement Systems for Industrial Inspection XI, edited by Peter Lehmann, Wolfgang Osten, Armando Albertazzi Gonçalves Jr., Proc. of SPIE Vol. 11056, 1105614 · © 2019
SPIE · CCC code: 0277-786X/19/$21 · doi: 10.1117/12.2526049

1. INTRODUCTION

Subproject C5 *Multiscale Geometry Inspection of Joining Zones* is part of the Collaborative Research Centre (CRC) 1153 *Process Chain for Manufacturing Hybrid High Performance Components by Tailored Forming*. Within the scope of the CRC new approaches and techniques are being analyzed and developed in order to manufacture hybrid workpieces made from different materials with locally adapted properties.

The optical geometry characterization of the wrought-hot, hybrid workpieces' directly after forming can yield divers advantages, e.g., the documentation of workpiece distortion effects during cooling or the rejection of deficient components in an early manufacturing state.

The high component temperature directly after forming of approximately 1000 °C influences the measurement accuracy of optical geometry sensors: Triangulation based, optical measurement techniques require homogeneous conditions and a rectilinear expansion of light. This precondition is violated, as the measurement object's heat input into the surrounding air leads to an inhomogeneous refractive index field. The authors identified low pressure measurements in air as a promising approach to reduce the magnitude and expansion of the heat induced optical inhomogeneity.[1]

New challenges arise when wrought-hot measurement objects are meant to be characterized geometrically under low pressure conditions: The object has to be positioned in a suitable low pressure chamber rapidly via a feeding system, and the measurement has to be conducted through a chamber window.

The sensors light path is therefore again affected: The emitted projector light traverses air with a density value at ambient pressure (from projector to chamber window), is then refracted at the boundary layer from air to window (due to the discrete increase in refractive index), and again refracted at the transition from window to chamber air (at low pressure). Unlike the heat induced deflection effect, the light path manipulation by the window and the manipulated air density state in the chamber are non-dynamic and constant over time. It is therefore possible to reproduce it's effect on the light path via a suitable sensor calibration model.

In this paper, the calibration model and procedure for a structured light sensor with telecentric stereo camera pair and entocentric projector is presented, considering the light refraction in a window installed in front of the sensor. The standard pinhole camera model for entocentric lenses cannot be used to model the projector's light path, as it does not comprise enough degrees of freedom to adequately parametrize the pixel-dependent light ray shift induced by thick windows. To simplify the calibration of the projector, the window is modeled as a plane-parallel plate of constant thickness, oriented in parallel to the projector sensor. Moreover, the refractive index of air is supposed to be constant – independently from the pressure state.

The presented approach is based on the calibration routine for a telecentric camera and an entocentric projector as developed by Li and Zhang,[2] who suggest a telecentric camera calibration with re-projected 3-D points in the projector coordinate frame. The entocentric projector model is extended with a window refraction model, similar to the model suggested by Bräuer et al.,[3] but with generally derived equations for light refraction. The necessity of an absolute pose reconstruction of window to projector[4] is not required – the window induced light ray shift does not depend on the distance from window to projector, but only on it's orientation.

An advantage results from choosing telecentric lenses for the cameras: As this lens type only maps parallel light onto a sensor, the window induced ray shift is constant for all camera pixels and can directly be reproduced by the so-called affine camera model – there is no need to introduce new parameters.

The presented measurement setup and calibration routine is suitable for all types of geometry measurements that need to be performed through windows – e.g., when it is necessary to protect the sensor from dirt and heat (in rough industrial environments), or in order to realize measurements in combinations with chambers.

2. STRUCTURED LIGHT SENSOR: HARDWARE AND THEORY

In this section, the structured light sensor is introduced and the basic theory on lens modeling and light refraction in a plane-parallel plate is given. The basic mathematical models for pinhole and orthographic projection are well-known principles, e.g., stated by Hartley et al.[5]

2.1 Hardware

The structured light sensor is composed of a telecentric stereo camera pair and an entocentric LED projector as depicted in Fig. 1. The monochromatic cameras are equipped with 1 in. CMOS sensors and provide a resolution

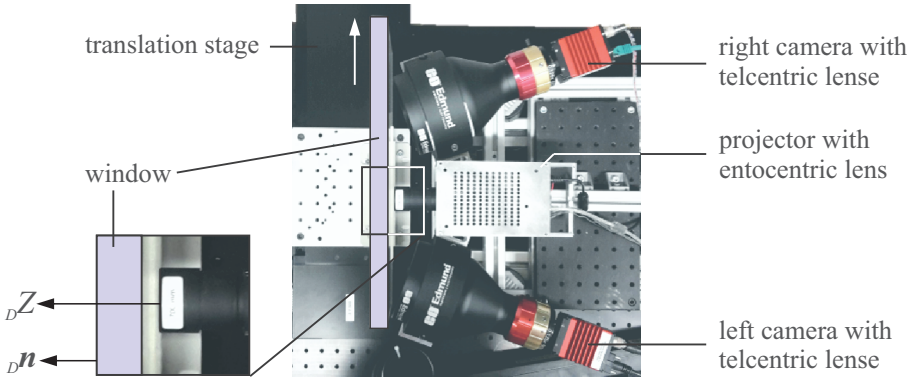


Figure 1: Structured light sensor comprising two cameras (stereo pair) with telecentric lenses, a projector with entocentric lens, and a measurement window on a translation stage in order to capture the same scenes (e.g., calibration plane poses) with and without window effect. The window has been arranged in parallel to the projector lenses' front, i.e., ideally the window's normal vector $_D\boldsymbol{n}$ is in parallel to the projector's optical axis $_DZ$.

of 2048 x 2048 pixels. The telecentric lenses are originally designed for 2/3 in. sensors, the usage of a bigger sensor leads to a vignetting effect and an actual circle field of view (FOV) of about 95 mm in diameter. The projector has a resolution of 1140 x 912 pixels. A virtual resolution of 3649 x 2281 pixels is hypothesized in the calibration process, due to the projector pixels' diamond shape. The sensor is calibrated for a working distance of approximately 280 mm, leading to a projector FOV of about 160 mm x 95 mm.

2.2 Mathematical Model of Entocentric Projector with Window Refraction

The projector is equipped with an entocentric lens and can therefore be described mathematically by the pinhole camera model in combination with a suitable distortion model to compensate for lens distortion effects.
The parametric representation of an entocentric projector's light ray $_Dg_{1,air}$ through a specific pixel point $_d\boldsymbol{u}$ (before being refracted) can be gained with help of the inverse projector camera matrix $(\boldsymbol{K}_{dD})^{-1}$ according to Eq.(1). $_D\boldsymbol{X}$ is defined in the coordinate system of the projector $_D\boldsymbol{COS}$ (indexes D and d stand for digital mirror device (DMD), matrix $\boldsymbol{K}_{dD}$ transforms a 3-D point from $_D\boldsymbol{COS}$ to the 2-D sensor system $_d\boldsymbol{COS}$), as well as $_D\boldsymbol{r}_{air}$, which defines the direction vector of the light ray in air.

$$
\begin{aligned}
{}_Dg_{1,air}: \quad & \underbrace{\begin{pmatrix} X \\ Y \\ Z \end{pmatrix}}_{{}_D\boldsymbol{X}} = \mu \underbrace{\begin{bmatrix} \frac{1}{f_x} & 0 & -\frac{c_x}{f_x} \\ 0 & \frac{1}{f_y} & -\frac{c_y}{f_y} \\ 0 & 0 & 1 \end{bmatrix}}_{(\boldsymbol{K}_{dD})^{-1}} \underbrace{\begin{pmatrix} u \\ v \\ 1 \end{pmatrix}}_{({}_d\boldsymbol{u},1)^T} = \mu \underbrace{\begin{pmatrix} \frac{u}{f_x} - \frac{c_x}{f_x} \\ \frac{v}{f_y} - \frac{c_y}{f_y} \\ 1 \end{pmatrix}}_{{}_D\boldsymbol{r}_{air}}, \\
{}_Dg_{1,air}: \quad & {}_D\boldsymbol{X} = \mu \quad (\boldsymbol{K}_{dD})^{-1} \quad ({}_d\boldsymbol{u},1)^T = \mu \quad {}_D\boldsymbol{r}_{air}.
\end{aligned}
\tag{1}
$$

The intrinsic parameters of the pinhole model are f_x, f_y, c_x and c_y. f_x and f_y are the focal lengths (in pixel) and c_x and c_y parametrize the location of the principle point (in pixel). μ is a scaling factor which depends on the $_DZ$-coordinate of the point, that has been projected onto the camera sensor. μ is measured in the metric domain of length, e.g., mm.
The projector lens distortion is modeled by a Zhang's[6] classical approach (e.g., as implemented in the *calibrate-camera*-function in the *OpenCV-library*[7,8]) with five parameters – three radial distortion coefficients for barrel or pincushion distortion (k_1, k_2 and k_3), and two parameters to model tangential distortion (p_1 and p_2). The distortion parameters are combined in vector $\boldsymbol{k}_D$, and $\tilde{\boldsymbol{k}}_D$ for the undistortion parameters respectively.
During the calibration process, it is necessary to transform 3-D points from the object coordinate frame $_O\boldsymbol{COS}$ of the calibration plane into the coordinate frame $_D\boldsymbol{COS}$ of the projector. This is achieved by a 4x4 rigid body transformation matrix $\boldsymbol{T}_{DO}$, combining rotation and translation in one matrix. For this operation, the 3-D points have to be homogeneous.

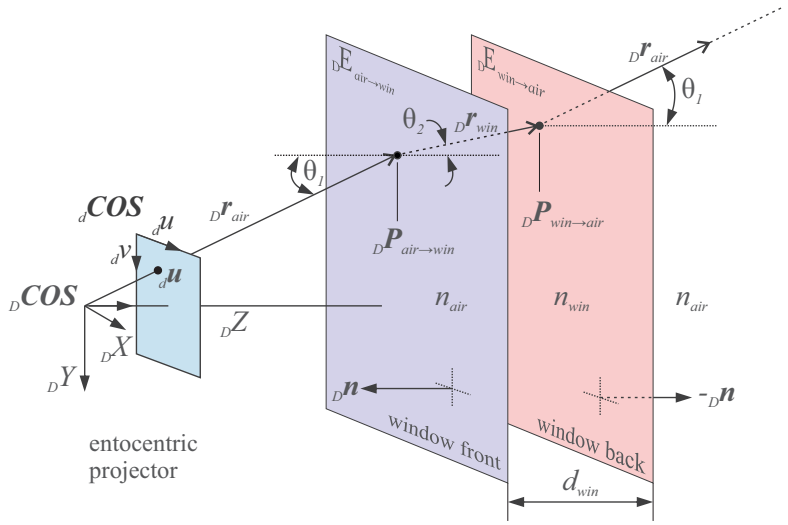


Figure 2: Principle outline of light refraction induced by measurement window for an exemplary projector light ray through pixel position ${}_d\boldsymbol{u} = ({}_du, {}_dv)$. The window is hypothesized to be a plane-parallel plate with constant thickness d_{win}. The projector's optical axis ${}_DZ$ and the window normal vector ${}_D\boldsymbol{n}$ are assumed to be parallel. As the window's refractive index n_{win} is larger than n_{air}, the light is deflected towards the optical axis – the projector's field of view (FOV) is reduced. After passing through the window, the original light path orientation ${}_D\boldsymbol{r}_{air}$ is restored, but the ray's correct mathematical description requires a new support vector ${}_D\boldsymbol{P}_{win\mapsto air}$ according to Eq. (5).

To correctly model the projector light path after window refraction, a general 3-D formulation of window refraction is needed. The following boundary conditions are assumed: The window is a plan-parallel plate with constant thickness d_{win} and constant refractive index n_{win}. Furthermore, the refractive index of air n_{air} is hypothesized to be approximately one – both at ambient pressure and at lower pressure conditions. These two assumptions allow a helpful simplification: A light ray is only displaced by the window, but it's vectorial orientation is not changing at all, meaning that ${}_D\boldsymbol{r}_{air}$ will not change for a specific projector light ray, but it's support vector will. In a first step, ${}_D\boldsymbol{r}_{air}$, ${}_D\boldsymbol{n}$, n_{air} and n_{win} are used to gain the direction of the refracted light ray ${}_D\boldsymbol{r}_{win}$ (refracted at first boundary layer between air and window). This can be achieved by Snell's law and an algebraic formulation of the refracted light ray[9] ${}_D\boldsymbol{r}_{win}$ according to Eqs. (2) and (3):

$$ {}_D\boldsymbol{r}_{win} = \frac{n_{air}}{n_{win}} \cdot \left(\frac{{}_D\boldsymbol{r}_{air}}{|{}_D\boldsymbol{r}_{air}|} + {}_D\,\boldsymbol{n} \cdot cos(\theta_1) \right) + (-{}_D\boldsymbol{n}) \cdot \sqrt{1 - \left(\frac{n_{air}}{n_{win}} \right)^2 \cdot (1 - cos(\theta_1)^2)}, \tag{2} $$

with

$$ cos(\theta_1) = -{}_D\boldsymbol{n} \cdot \frac{{}_D\boldsymbol{r}_{air}}{|{}_D\boldsymbol{r}_{air}|}. \tag{3} $$

The window orientation is formulated in the coordinate frame of the projector and given by normal vector ${}_D\boldsymbol{n} = (n_1, n_2, n_3)^T$, θ_1 is the ray's incidence angle (as depicted in Fig. 2).
The parametric representation of the plane equations for the window boundary layers (${}_DE_{air\mapsto win}$ and ${}_DE_{win\mapsto air}$, compare to Fig. 2) can be gained with the normal vector ${}_D\boldsymbol{n}$ and the window thickness d_{win}. The support vector of plane ${}_DE_{air\mapsto win}$ is set to ${}_D\boldsymbol{s}_{air\mapsto win} = (0, 0, 1)^T$. The ${}_DZ$-component of ${}_D\boldsymbol{s}_{air\mapsto win}$ is set to one – the amount of light ray shift does not depend on the distance between window and projector, as long as the window orientation is the same. As depicted in Fig. 1, the window has been arranged in parallel to the projector lens. Therefore the normal vector is set to ${}_D\boldsymbol{n} = (0, 0, -1)$. The support vector of plane ${}_DE_{win\mapsto air}$ can directly be determined from ${}_D\boldsymbol{s}_{air\mapsto win}$, window pose ${}_D\boldsymbol{n}$ and thickness d_{win}. The direction vectors for ${}_DE_{air\mapsto win}$ and ${}_DE_{win\mapsto air}$ are equal, as the window is hypothesized to be a plane-parallel plate. The basic procedure to gain the equation for the line-of-sight of a specific, refracted projector light ray ${}_Dg_{3,air}$ is (compare to Fig. 2):

(1) Determine line-of-sight ${}_D g_{1,air}$ of projector according to Eq. (1) without window, result is direction vector ${}_D\boldsymbol{r}_{air}$ for specific ray.

(2) Determine intersection between ${}_D g_{1,air}$ and ${}_D E_{air \mapsto win}$, result is 3-D point ${}_D\boldsymbol{P}_{air \mapsto win}$.

(3) Determine refracted light ray ${}_D\boldsymbol{r}_{win}$ with Eqs. (2) and (3) and ${}_D\boldsymbol{r}_{air}$.

(4) Formulate linear equation for the light ray when passing through the window according to:

$$Dg_{2,win} :_D \boldsymbol{X} = \lambda_1 \cdot_D \boldsymbol{r}_{win} +_D \boldsymbol{P}_{air \mapsto win}. \tag{4}$$

(5) Determine intersection between ${}_D g_{2,win}$ and ${}_D E_{win \mapsto air}$, result is 3-D point ${}_D\boldsymbol{P}_{win \mapsto air}$.

(6) Formulate linear equation for the light ray after window refraction in air according to:

$${}_D g_{3,air} :_D \boldsymbol{X} = \lambda_2 \cdot_D \boldsymbol{r}_{air} +_D \boldsymbol{P}_{win \mapsto air}. \tag{5}$$

${}_D\boldsymbol{P}_{win \mapsto air}$ has to be determined individually for each projector light ray in order to describe the ray path after window refraction correctly, whereas ${}_D\boldsymbol{r}_{air}$ is not changing at all for a plane-parallel plate with constant refractive index.

2.3 Mathematical Model of Telecentric Cameras

The cameras are equipped with telecentric lenses. Therefore only parallel light is mapped onto the sensor (orthographic projection). The mathematical model according to Eq. (6) (compare to Chen et al.[10]) differs from the pinhole model, as the projection center of a so-called affine camera is located at infinity – there is no principle point.

$$\underbrace{\begin{pmatrix} u \\ v \\ 1 \end{pmatrix}}_{{}_c} = \underbrace{\begin{bmatrix} \frac{m}{s_x} & \gamma & o_x \\ 0 & \frac{m}{s_y} & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{K}_{cC}} \underbrace{\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\tilde{\boldsymbol{T}}_{CO}} \underbrace{\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}}_{{}_O} = \underbrace{\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{\boldsymbol{P}_{cO}} \underbrace{\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}}_{{}_O}, \tag{6}$$

$$({}_c\boldsymbol{u}, 1)^T = \qquad ({}_O\boldsymbol{X}, 1)^T = \qquad ({}_O\boldsymbol{X}, 1)^T.$$

An arbitrary homogeneous 3-D object point $({}_O\boldsymbol{X}, 1)$ is transformed by a truncated rigid body matrix $\tilde{\boldsymbol{T}}_{CO}$ into the coordinate frame of the camera ${}_C\boldsymbol{COS}$. The resulting homogeneous 2-D point $({}_C\boldsymbol{X}, 1)$ in the camera coordinate frame ${}_C\boldsymbol{COS}$ is multiplied with the affine camera matrix $\boldsymbol{K}_{cC}$, which projects the point onto the sensor in position ${}_c\boldsymbol{u}$ (in pixel). The intrinsics of the camera model are given by magnification factor m (no unit), and the pixel sizes s_x and s_y in x and y-direction (in metric domain length). The skewing factor γ is given in pixel per metric domain length.

To define a centre for a telecentric lens distortion model, the origin of the image coordinate system is fixed to the middle of the camera sensor ($o_x = w_{sensor}/2, o_y = h_{sensor}/2$), with sensor width w_{sensor} and height h_{sensor}. The distortion vector is defined as $\boldsymbol{k}_C = (k_1, k_2, p_1, p_2)$, the undistortion vector as $\tilde{\boldsymbol{k}}_C = \left(\tilde{k}_1, \tilde{k}_2, \tilde{p}_1, \tilde{p}_2\right)$.

When triangulating a 3-D point based on an affine camera model, the projection matrix $\boldsymbol{P}$ is needed. It can be gained by multiplication of $\boldsymbol{K}_{cC}$ and $\tilde{\boldsymbol{T}}$ (for a specific measurement coordinate frame).

An additional advantage is inherent in the orthographic projection model of a telecentric camera when measuring geometries through windows: Only parallel light is mapped onto the camera sensor. As the amount of light ray displacement in a plan-parallel plate (like a window) depends on it's incidence angle, and the incidence angle is constant for parallel light rays, the window refraction effect results in a virtual shift of the whole camera. Therefore, no additional parameter set-up is necessary to reproduce the window refraction effect.

3. CALIBRATION ROUTINE

The calibration routine of the structured light sensor is divided into three different steps: First of all, the projector is calibrated without window to guarantee a correct pinhole camera estimation. Secondly, the pinhole model of the projector is extended by a window refraction model, hypothesizing a parallel orientation of window to projector sensor. The estimated projector camera matrix from the first step (without window) are fixed and the distortion parameters of the projector are re-optimized, e.g., to account for additional window distortion effects. In the last step, the telecentric cameras are calibrated with re-projected 3-D points in the frame of the projector ${}_D\boldsymbol{COS}$. The routine is based on a calibration procedure developed by Li and Zhang,[2] but expanded by a mathematical description of window refraction.

3.1 Projector Calibration

The projector is calibrated as a pinhole camera with planar calibration patterns based on the concept of Zhang.[6] To this end, the approach of an image-capturing projector is applied, as introduced by Zhang and Huang.[11] The calibration planes are captured in different poses by the telecentric cameras. To gain the necessary correspondences between calibration plane points and projector sensor points, absolute phasemaps in both horizontal and vertical direction are recorded for each plane pose. The calibration plane markers are white circles on black background to guarantee phasemap information in each marker position.

In order to provide comparative reference data, the plane poses are recorded with and without window in front of the structured light sensor. This is achieved by moving the window via translation stage for each single plane pose. The main advantage of this approach is, that the correct orientation of plane i is given by rigid body transformation ${}^i\boldsymbol{T}_{DO}$. ${}^i\boldsymbol{T}_{DO}$ is known via the conventional calibration routine without window, and can be compared to the plane orientation estimated by the expanded calibration routine with window.

The basic step-by-step routine for the projector calibration is given below: First, a conventional projector calibration is conducted without window, based on a Levenberg-Marcquardt algorithm by minimizing function $\mathfrak{e}_{O,nowin}$ given as

$$\mathfrak{e}_{O,nowin} = \sum_{i=0}^{m-1} \sum_{j=0}^{{}^in-1} ||{}^{ij}_{d,nowin}(u,v) - f_1({}^{ij}_{O,nowin}(X,Y), \boldsymbol{K}_{dD,nowin}, \boldsymbol{k}_{D,nowin}, {}^i\boldsymbol{T}_{DO,nowin})||^2. \tag{7}$$

$\mathfrak{e}_{O,nowin}$ is the sum of squared errors between the matched feature points ${}^{ij}_{d,nowin}(u,v)$ and the corresponding projected points ${}^{ij}_{d,nowin}(\hat{u},\hat{v})$ according to

$${}^{ij}_{d,nowin}(\hat{u},\hat{v}) = f_1({}^{ij}_{O,nowin}(X,Y), \boldsymbol{K}_{dD,nowin}, \boldsymbol{k}_{D,nowin}, {}^i\boldsymbol{T}_{DO,nowin}). \tag{8}$$

The projection error is given in pixel, as it is defined in the projector sensor coordinate frame ${}_d\boldsymbol{COS}$. i defines the calibration plane number (total amount of poses is m), j the number of point correspondences on a plane (total amount ${}^in$ is specific for plane pose i). $\boldsymbol{K}_{dD,nowin}$, $\boldsymbol{k}_{D,nowin}$ and ${}^i\boldsymbol{T}_{DO,nowin}$ are optimized.

In the next step, the projector's distortion coefficients are calibrated with window, again realized with a Levenberg-Marcquardt algorithm, but in this case the window refraction is modeled in the light path of the projector, and the optimization direction is inverse, i.e., from sensor to calibration plane.

$$\mathfrak{E}_{O,win} = \sum_{i=0}^{m-1} \sum_{j=0}^{{}^in-1} ||{}^{ij}_{O,win}(X,Y,0) - f_2({}^{ij}_{d,win}(u,v), (\boldsymbol{K}_{dD,nowin})^{-1}, \tilde{\boldsymbol{k}}_{D,win}, ({}^i\boldsymbol{T}_{DO,win})^{-1}, {}_D\boldsymbol{n}, n_{air}, n_{win}, d_{win})||^2. \tag{9}$$

$\mathfrak{E}_{O,win}$ is the sum of squared errors between the matched and projected feature points in the object coordinate frame of a specific calibration plane. It follows for a projected feature point

$${}^{ij}_{O,win}(\hat{X},\hat{Y},\hat{Z}) = f_2({}^{ij}_{d,win}(u,v), (\boldsymbol{K}_{dD,nowin})^{-1}, \tilde{\boldsymbol{k}}_{D,win}, ({}^i\boldsymbol{T}_{DO,win})^{-1}, {}_D\boldsymbol{n}, n_{air}, n_{win}, d_{win}). \tag{10}$$

Equation (9) not only depends on the camera intrinsics and the plane extrinsics, but also on the window refraction parameters ${}_D\boldsymbol{n}$, n_{air}, n_{win} and d_{win}.

$\boldsymbol{K}_{dD,nowin}$ is fixed to the values according to the calibration without window. Moreover, none of the refraction parameters are estimated in order to reduce the parameter space, only the extrinsics $({}^{i}\boldsymbol{T}_{DO,win})^{-1}$ and the undistortion parameters $\tilde{\boldsymbol{k}}_{D,win}$ are optimized. Assumptions are:

- The data sheet values for the measurement window are correct ($n_{win} = 1.47$ and d_{win}=21 mm).
- The refractive index of air is supposed to be constant with a value of approximately $n_{air} \approx 1$.
- The window pose is hypothesized to be ${}_{D}\boldsymbol{n} = (0, 0, -1)$ and will not be optimized. This assumption is considered to be valid, if the window is accurately positioned in front of the projector, and the window normal vector directs in the negative direction of the projector's optical axis.

Deviations from the assumptions above are meant to be reproduced by a re-estimation of the distortion coefficients, leading to distortion coefficients with window effect $\boldsymbol{k}_{D,win}$, and undistortion coefficients $\tilde{\boldsymbol{k}}_{D,win}$ respectively. The projection of a 2-D projector sensor point ${}^{ij}_{d}(u, v)$ onto a specific calibration plane (as required in Eq. (9)) has to be realized according to Subsec. 2.2, by formulating the linear light ray equation with a pixel-dependent support vector as given in Eq. (5).

3.2 Camera Calibration

Subsequently, the telecentric cameras are calibrated with help of the calibrated projector and the camera-corresponding calibration plane data sets from the previous step. To this end, the detected feature points ${}^{ij}_{d,win}(u, v)$ in the 2-D coordinate frame ${}_{d}\boldsymbol{COS}$ are projected back into the projector's 3-D coordinate frame ${}_{D}\boldsymbol{COS}$. The basic approach is simple: The extrinsic parameters ${}^{i}\boldsymbol{T}_{DO,win}$ of a specific calibration plane pose i can be used to formulate a plane equation in coordinate frame ${}_{D}\boldsymbol{COS}$, which again can be intersected with the corresponding line-of-sight of a detected projector feature point ${}^{ij}_{d,win}(u, v)$. Again, the light ray equation has to be formulated according to Eq. (5) in order to consider window refraction. The resulting 3-D points ${}^{ij}_{D,win}(X, Y, Z)$ are then used to calibrate the cameras.
The calibration of the telecentric cameras with projector 3-D points has a major advantage – the cameras are already calibrated in the correct coordinate frame ${}_{d}\boldsymbol{COS}$. Triangulated measurement results between camera one and projector, camera two and projector, or the stereo camera pair already appear in the same coordinate frame – there is no need for further rigid body transformations. As described in Subsec. 2.3, the window refraction results in a virtual shift of the telecentric cameras. The affine projection matrix $\boldsymbol{P}_{cD,win}$ (compare to Eq. (1)) from 3-D projector coordinate frame ${}_{D}\boldsymbol{COS}$ to 2-D camera coordinate frame ${}_{c}\boldsymbol{COS}$ fully reproduces this virtual shift, when optimized based on the re-projected point cloud ${}^{l}_{D,win}(X, Y, Z)$.
Again, a nonlinear Levenberg-Marquardt algorithm is used to minimize the sum of squared errors between matched feature points ${}^{l}_{c}(u, v)$ and projected feature points ${}^{l}_{c}(\hat{u}, \hat{v})$ – in this case in the coordinate frame of the corresponding camera ${}_{c}\boldsymbol{COS}$ in pixel.

$$\mathfrak{e}_{c,win} = \sum_{l=0}^{k-1} ||{}^{l}_{c,win}(u, v) - f_3({}^{l}_{D,win}(X, Y, Z), \boldsymbol{P}_{cD,win}, \boldsymbol{k}_{C,win})||^2, \tag{11}$$

$${}^{l}_{c,win}(\hat{u}, \hat{v}) = f_3({}^{l}_{D,win}(X, Y, Z), \boldsymbol{P}_{cD,win}, \boldsymbol{k}_{C,win}). \tag{12}$$

$\mathfrak{e}_{c,win}$ is the sum of squared errors between the matched feature points ${}^{l}_{c,win}(u, v)$ and the projected feature points ${}^{l}_{c,win}(\hat{u}, \hat{v})$. k parametrizes the total amount of point correspondences.

4. RESULTS

In this section, the calibration results with and without window refraction are compared in terms of intrinsic and extrinsic projector and camera parameters.

4.1 Calibration

As the different calibration plane poses are recorded with and without window refraction by moving the window via translation stage for each single plane pose (compare to Fig. 1), the ground truth for the poses is known in terms of ${}^{i}\boldsymbol{T}_{DO,nowin}$.

Results for three scenarios are depicted: Scenario one corresponds to a conventional calibration using the *OpenCV-library* (no window refraction model implemented). As the calibration plane data was recorded without window refraction, the chosen model fits the application well. The mean projection error is given in pixel, as the marker pixel distance on the sensor is minimized.

The same calibration function is used for scenario two. However, in contrast to scenario one, the plane data was recorded through the window – therefore the model can only approximate the actual projector light path roughly, as window refraction is not considered. The focal lengths f_x and f_y of scenario two are estimated too large by approximately 20 pixels, due to the missing window model.

In scenario three, the projector matrix $\boldsymbol{K}_{dD,nowin}$ (from scenario one) is used, and the window refraction is additionally modeled according to Subsec. 3.1. $\tilde{\boldsymbol{k}}_{D,nowin}$ is used as a start value for the optimization of the undistortion coefficients $\tilde{\boldsymbol{k}}_{D,win}$ with window. The projection error is given in mm, as the marker distances are minimized in the object coordinate frame of the calibration planes, but can be transformed to pixel with the projector's FOV for the applied working distance. The FOV is approximately 160 mm x 95 mm, which leads to a projection error of approximately 0.8 pixels for the given error of 0.034 mm (when considering the virtual projector resolution). This means, that the projection errors are of same scale. It is therefore not sufficient to evaluate the calibration quality only based on the error values according to the sixth column in Tab. 1.

A closer analysis of the estimated extrinsics reveals the effect of the window refraction model. The last column in Tab. 1 depicts the mean distance between the estimated translation vector t_i as part of ${}^{i}\boldsymbol{T}_{DO,i}$ and the actual ground truth value $t_{i,nowin}$ as part of ${}^{i}\boldsymbol{T}_{DO,nowin}$, leading to a mean distance value of approximately 5.9 mm for scenario two, but only 71.5 µm for scenario three. A comparison of the optimized rotation matrices ${}^{i}\boldsymbol{R}_{DO,i}$ is not given in Tab. 1, but again scenario three yields better results than scenario two.

The projector intrinsics are used to re-project the points ${}^{ij}_{d}(u,v)$ back into the 3-D coordinate frame ${}_{D}\boldsymbol{COS}$ of the projector in order to calibrate the telecentric cameras. The results for the right camera (according to Fig. 1) are given in Tab. 2, the same scenario numbering is used as in the previous table. The affine camera matrix $\boldsymbol{K}_{cC,nowin}$ for scenario one is considered ground truth. Scenario one, two and three yield similar results.

The light ray shift induced by window refraction is illustrated when comparing matrix entry $\tilde{\boldsymbol{T}}_{CD,nowin}[1,4] = t_{x,CD,nowin}$ of scenario one and entry $\tilde{\boldsymbol{T}}_{CD,win}[1,4] = t_{x,CD,win}$ of scenario three. These values carry the information about the cameras x-coordinate with regard to the projector coordinate frame ${}_{D}\boldsymbol{COS}$. The difference between non-refracted and refracted x-values Δx_{calib} (estimated in the calibration) should therefore be equal to the theoretical light ray shift Δx_{ref}. The calibration based difference Δx_{calib} yields:

$$\Delta x_{calib} = |t_{x,CD,nowin} - t_{x,CD,win}| = |-109.9894 - (-112.2775)| = 2.2881\,\text{mm}. \tag{13}$$

The theoretical reference value is

$$\Delta x_{ref} = d_{win} \cdot sin(\alpha) \cdot \left[1 - \frac{cos(\alpha)}{\sqrt{\frac{n_{win}^2}{n_{air}^2} - sin(\alpha)^2}}\right] = 21\,\text{mm} \cdot sin(20°) \cdot \left[1 - \frac{cos(20°)}{\sqrt{\frac{1.47^2}{1^2} - sin(20°)^2}}\right] = 2.46\,\text{mm}, \tag{14}$$

with α as angle of incidence according to the hardware setup (equals the triangulation angle between projector and camera, approximately 20°). The values just differ slightly, which is possibly due to the roughly determined angle value. As mentioned before, the window refraction results in a virtual shift Δx_{calib} of the telecentric cameras, which is fully inherent in the truncated rigid body transformation $\tilde{\boldsymbol{T}}_{CD,win}$ of scenario three – but only when calibrated with re-projected points based on the projector model with window refraction. As the remaining matrix entries of $\tilde{\boldsymbol{T}}_{CD,nowin}$ for scenario one and $\tilde{\boldsymbol{T}}_{CD,win}$ for scenario three are very similar, the reproduction of the refraction effect in the calibration data is considered to be valid. For the sake of completeness, the projection errors are given as well – the difference of the values in Tab. 2 is not significant. The small advantage of the scenario two data is possibly a random effect, maybe due to outliers elimination in the calibration procedure.

scenario	window	model	projector camera matrix $\boldsymbol{K}_{dD}$	distortion coefficients $\boldsymbol{k}_D$	mean proj. error / pixels or mm	mean distance $\frac{\sum_{i=0}^{m-1} \|t_{i,nowin} - t_i\|}{m}$ / mm
1	no	no	$\begin{bmatrix} 6938.1 & 0 & 1852.9 \\ 0 & 6939.6 & 1182.1 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} k_1 \\ k_2 \\ p_1 \\ p_2 \\ k_3 \end{bmatrix} = \begin{bmatrix} -3.9072 \cdot 10^{-2} \\ -5.9844 \cdot 10^{-1} \\ -6.1830 \cdot 10^{-5} \\ 1.9296 \cdot 10^{-3} \\ 6.3360 \end{bmatrix}$	0.774 pixels	0 (ground truth)
2	yes	no	$\begin{bmatrix} 6957.6 & 0 & 1845.6 \\ 0 & 6957.9 & 1183.2 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} k_1 \\ k_2 \\ p_1 \\ p_2 \\ k_3 \end{bmatrix} = \begin{bmatrix} -3.1109 \cdot 10^{-2} \\ -4.1705 \cdot 10^{-1} \\ 4.3338 \cdot 10^{-4} \\ 1.5615 \cdot 10^{-3} \\ 3.5144 \end{bmatrix}$	0.801 pixels	5.9008
3	yes	yes	$\begin{bmatrix} 6938.1 & 0 & 1852.9 \\ 0 & 6939.6 & 1182.1 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} k_1 \\ k_2 \\ p_1 \\ p_2 \\ k_3 \end{bmatrix} = \begin{bmatrix} -3.7383 \cdot 10^{-2} \\ -7.0535 \cdot 10^{-1} \\ 1.5284 \cdot 10^{-4} \\ 2.0169 \cdot 10^{-3} \\ 6.7085 \end{bmatrix}$	0.034 mm, (approx. 0.8 pixels)	0.0715

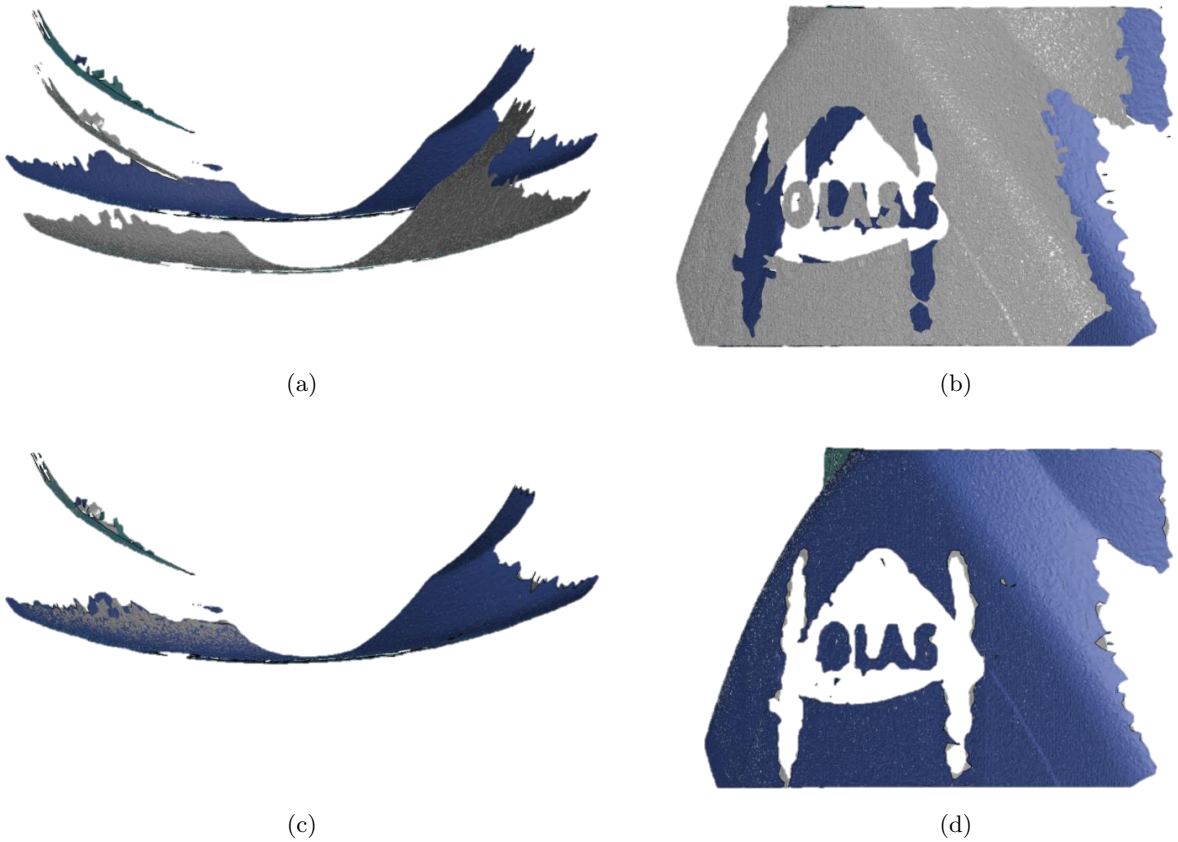
Table 1: Projector intrinsics based on different calibration scenarios.

scenario	window	model	camera matrix $\boldsymbol{K}_{cC}$	truncated rigid body matrix $\tilde{\boldsymbol{T}}_{CD}$	distortion coefficients $\boldsymbol{k}_C$	mean proj. error / pixels
1	no	no	$\begin{bmatrix} 23.937 & 0.0373 & 1024 \\ 0 & 23.935 & 1024 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0.9375 & -0.0009 & 0.3480 & -109.9894 \\ -0.0026 & 0.99995 & 0.0096 & -0.5080 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} k_1 \\ k_2 \\ p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 5.0630 \cdot 10^{-6} \\ -2.7627 \cdot 10^{-9} \\ 4.4397 \cdot 10^{-6} \\ -3.4561 \cdot 10^{-6} \end{bmatrix}$	0.768
2	yes	no	$\begin{bmatrix} 23.939 & 0.0218 & 1024 \\ 0 & 23.952 & 1024 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0.9384 & -0.0009 & 0.3456 & -110.0420 \\ -0.0026 & 0.99995 & 0.0098 & -0.5239 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} k_1 \\ k_2 \\ p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 3.8998 \cdot 10^{-6} \\ -2.3790 \cdot 10^{-9} \\ 4.7773 \cdot 10^{-6} \\ -5.5019 \cdot 10^{-6} \end{bmatrix}$	0.653
3	yes	yes	$\begin{bmatrix} 23.941 & 0.0295 & 1024 \\ 0 & 23.954 & 1024 \\ 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} 0.9378 & -0.0008 & 0.3472 & -112.2775 \\ -0.00265 & 0.99995 & 0.0096 & -0.5768 \\ 0 & 0 & 0 & 1 \end{bmatrix}$	$\begin{bmatrix} k_1 \\ k_2 \\ p_1 \\ p_2 \end{bmatrix} = \begin{bmatrix} 3.9164 \cdot 10^{-6} \\ -2.3926 \cdot 10^{-9} \\ 4.9568 \cdot 10^{-6} \\ -5.7750 \cdot 10^{-6} \end{bmatrix}$	0.703

Table 2: Camera intrinsics for right camera for different calibration scenarios.

4.2 Measurement

The performance of the window refraction model can be demonstrated by triangulating 3-D geometry points. In this case the surface of a detergent plastic bottle is used as measurement object. Within the scope of this work, the triangulation routine will not be explained. Just some basic remarks: The 3-D surface reconstruction is executed by triangulation between the two telecentric cameras (stereo camera pair). Although the projector calibration is necessary for the camera calibration, the projector model will not be used for the 3-D reconstruction of the surface points. The projector is simply used as *feature generator* to solve the correspondence problem between the cameras, based on vertical and horizontal absolute phasemap information in both camera views.
The triangulation is basically conducted as given by Chen et al.[10] The necessary projection matrix $\boldsymbol{P}_{cD}$ is gained from $\boldsymbol{K}_{cC}$ and $\tilde{\boldsymbol{T}}_{CD}$ for the different scenarios according to Tab. 2. Of course, the second camera's projection matrix needs to be determined as well.
To allow comparability, the measurement object's pose has not been changed between the single measurements. Moreover, it is important to notice that the meshed surface data according to Fig. 3 has not been moved and registered by an iterative closest point (ICP) algorithm – the data has just been imported to the same inspection software without additional rigid body transformations. The blue mesh is considered ground truth (calibrated and triangulated without window refraction). The gray meshes have been triangulated with window. Overexposed surface regions are excluded from the reconstruction algorithm, leaving areas with no surface information in the mesh.
Subfigures (a) and (b) in Fig. 3 depict a comparison between scenarios one and two (no window refraction model), Subfigs. (c) and (d) between scenarios one and three (additional window refraction model).



(a) (b)

(c) (d)

Figure 3: Comparison of 3-D surface data of a plastic bottle, gained on basis of different calibration routines. Blue: Ground truth data, calibrated and triangulated without window refraction (corresponding to calibration data according to scenario one in Tabs. 1 and 2.) (a,b) Triangulated data through window, without window model (scenario two). (c,d) Triangulated data through window, with window model (scenario three). Data evaluated with *GOM Inspect*.[12]

Without additional refraction model, the surface data is reconstructed in front of the actual reference surface and therefore too close to the projector center – the meshed data is uniformly shifted. This effect is not present in Subfigs. (c) and (d). Although not perfectly congruent with the reference geometry, the triangulated data with window model fits the original data quite well.
A surface data alignment between reference and window effected geometries via iterative closest point (ICP) algorithm permits a further analysis of the measurement object's possible shape manipulation due to the window effect. Again, the triangulated geometry with refraction model fits the reference surface well. But even the measurement without additional window model yields reasonable results. A further analysis of window induced shape manipulations is necessary. Measurement of objects with greater surface gradients need to be examined – the results might reveal the advantage of the window refraction model in a more distinct way.

5. CONCLUSION AND FUTURE WORK

In this paper, a calibration routine for a telecentric stereo camera pair with measurement window is presented. The entocentric projector is combined with a window refraction model and used to re-project 2D points into the 3-D coordinate frame of the projector in order to provide 3-D points for the calibration of the telecentric cameras. Calibration results both for projector, as well as for affine cameras are stated, illustrating that the window refraction effect is fully reproduced by the introduced model.
In forthcoming work, the window pose is meant to be optimized robustly in the calibration routine as well.

Promising research in this field has been presented by Agrawal et al.[13] in their approach on modeling multi-layer flat refractive geometries.

ACKNOWLEDMENTS

The authors would like to thank the German Research Foundation (DFG) for the financial and organisational support of this project, as the presented results were obtained within the Collaborative Research Centre 1153 *Process Chain for Manufacturing Hybrid High Performance Components by Tailored Forming*.

REFERENCES

[1] Beermann, R., Quentin, L., Reithmeier, E., and Kästner, M., "2d refractive index field measurements in air in different pressure scenarios," in [*Proceedings of SPIE/COS Photonics Asia, 2018, Beijing, China*], **10819**, SPIE (2018).

[2] Li, B. and Zhang, S., "Flexible calibration method for microscopic structured light system using telecentric lens," *Opt. Express* **23**, 25795–25803 (Oct 2015).

[3] Bräuer-Burchardt, C., Kühmstedt, P., and Notni, G., "Combination of air- and water-calibration for a fringe projection based underwater 3d-scanner," in [*Computer Analysis of Images and Patterns*], Azzopardi, G. and Petkov, N., eds., 49–60, Springer International Publishing (2015).

[4] Haner, S. and Åström, K., "Absolute pose for cameras under flat refractive interfaces," in [*2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*], 1428–1436 (June 2015).

[5] Hartley, R. and Zisserman, A., [*Multiple View Geometry in Computer Vision*], Cambridge University Press, 2 ed. (2004).

[6] Zhang, Z., "A flexible new technique for camera calibration," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(11), 1330–1334 (2000).

[7] Bradski, G., "OpenCV library," *Dr. Dobb's Journal of Software Tools* (2000).

[8] Bradski, G. and Kaehler, A., [*Learning OpenCV: Computer Vision with the OpenCV Library*], O'Reilly & Associates, 1 ed. (2008).

[9] Kidger, M. J., [*Fundamental Optical Design*], SPIE - The International Society for Optical Engineering (2001).

[10] Chen, Z., Liao, H., and Zhang, X., "Telecentric stereo micro-vision system: Calibration method and experiments," *Optics and Lasers in Engineering* **57**, 8292 (06 2014).

[11] Zhang, S. and Huang, P. S., "Novel method for structured light system calibration," *Optical Engineering* **45**, 45 – 45 – 8 (2006).

[12] GOM, "GOM Inspect." http://www.gom-inspect.com/de (2019). (Accessed: 08 April 2019).

[13] Agrawal, A., Ramalingam, S., Taguchi, Y., and Chari, V., "A theory of multi-layer flat refractive geometry," in [*2012 IEEE Conference on Computer Vision and Pattern Recognition*], 3346–3353 (June 2012).