Check for updates

A Camera Model for Line-Scan Cameras with Telecentric Lenses

Carsten Steger[1] 🄳 · **Markus Ulrich[1]** 🄳

Abstract

We propose a camera model for line-scan cameras with telecentric lenses. The camera model assumes a linear relative motion with constant velocity between the camera and the object. It allows to model lens distortions, while supporting arbitrary positions of the line sensor with respect to the optical axis. We comprehensively examine the degeneracies of the camera model and propose methods to handle them. Furthermore, we examine the relation of the proposed camera model to affine cameras. In addition, we propose an algorithm to calibrate telecentric line-scan cameras using a planar calibration object. We perform an extensive evaluation of the proposed camera model that establishes the validity and accuracy of the proposed model. We also show that even for lenses with very small lens distortions, the distortions are statistically highly significant. Therefore, they cannot be omitted in real-world applications.

Keywords Line-scan cameras · Telecentric lenses · Camera models · Camera model degeneracies · Camera calibration

1 Introduction

Line-scan cameras play an important role in machine vision applications because they offer a better resolution per price in comparison to area-scan cameras. Today, line-scan cameras with lines of up to 16,384 pixels are available (Steger et al. 2018, Chapter 2.3.4). The height of the resulting image is essentially unlimited because it corresponds to the number of 1D images acquired over time, as described in more detail below. Hence, several hundred megapixels per image can be achieved easily.

In contrast to area-scan cameras, the sensor of a line-scan camera consists of a single line of photosensitive elements. Consequently, the image that is obtained from a line-scan camera would be one pixel high. To obtain a 2D image that can be processed in computer vision applications, multiple 1D images are stacked over time while moving the sensor with respect to the object that is to be imaged. In machine vision applications, the relative motion is realized either by

mounting the camera above the moving object or by moving the camera across the stationary object (Steger et al. 2018, Chapter 2.3.1.1). The motion can be effected, for example, by a conveyor belt, a linear motion slide, or other linear actuators. For practical applications, it is therefore not sufficient to calibrate the single sensor line only. Instead, also the process of creating the 2D image must be included in the calibration.

Obviously, the resulting image strongly depends on the relative motion of the camera with respect to the object. In almost all machine vision applications, a linear motion is applied. This requires the camera to move with constant velocity along a straight line relative to the object while the orientation of the camera is constant with respect to the object. Furthermore, the motion must be equal for all images (Gupta and Hartley 1997). In other application domains, e.g., in remote sensing, more general motion models are applied. For example, the motion of an airborne or spaceborne camera can be modeled by discrete positions and orientations (Haala et al. 1998) or approximated by polynomial functions (Lee et al. 2000; Poli 2007). Since our focus is on machine vision applications, we will assume a linear motion in this paper. In practice, a linear motion can be realized by using appropriate encoders that ensure a constant speed (Steger et al. 2018, Chapter 2.3.1.1; Beyerer et al. 2016, Chapter 6.8).

Because typical readout rates of line-scan cameras are in the range of 10–200 kHz (Steger et al. 2018, Chapter 2.3.1.1), in some applications the exposure time of each line needs to

---

Communicated by Jun Sato.

✉ Carsten Steger
   steger@mvtec.com

   Markus Ulrich
   ulrich@mvtec.com

[1] MVTec Software GmbH, Arnulfstraße 205, 80634 Munich, Germany

🄳 Springer

be very short. Therefore, line-scan applications often require a very bright illumination. Fortunately, the illumination only needs to cover the narrow footprint of the sensor line, and hence often line-shaped LED or halogen light sources are used. Another consequence of the short exposure time is that often an open diaphragm, i.e., a large aperture, is used to maximize the amount of light that passes through the lens. This must be taken into account when choosing an appropriate hardware setup because a large aperture severely limits the depth of field.

For line-scan cameras with linear motion and a conventional entocentric (i.e., perspective) lens, an appropriate camera model and calibration procedure have been described in Steger et al. (2018, Chapter 3.9.3). Because of the perspective projection of an entocentric lens, objects that are closer to the lens produce a larger image. Therefore, this kind of line-scan camera performs a perspective projection in the direction of the line sensor and a parallel projection perpendicular to the line sensor.

In contrast to entocentric lenses, telecentric lenses perform a parallel projection of the world into the image (Steger et al. 2018, Chapter 2.2.4). In many machine vision applications, and especially in measurement applications, telecentric lenses are preferred over entocentric lenses because they eliminate perspective distortions, which is especially important for gauging applications when non-flat objects must be inspected. Also, self-occlusions of objects that would occur because of perspective distortions are avoided (Luster and Batchelor 2012, Chapter 6.1).

Steger (2017, Section 4) and Steger et al. (2018, Chapter 2.2.4) show that the placement of the aperture stop is crucial for the projection properties of a lens. For entocentric lenses, the aperture stop is placed between the two focal points of the lens system. For telecentric lenses that perform a parallel projection in object space, the aperture stop is placed at the image-side focal point (Steger et al. 2018, Chapter 2.2.4; Lenhardt 2017, Chapter 4.2.14; Beyerer et al. 2016, Chapter 3.4.5). This effectively filters out all light rays that are not parallel to the optical axis in object space. Owing to the parallel projection in object space, the lens must be chosen at least as large as the object to be imaged.

Because of the advantages of line-scan cameras and telecentric lenses, many machine vision applications would benefit from a camera model for line-scan cameras with telecentric lenses. Obviously, a camera model and an appropriate calibration procedure are essential for accurate 2D measurements in world units. The calibration result facilitates the removal of image distortions and the rectification from the image to a world plane.

In addition to accurate measurement tasks, the calibration of line-scan cameras is important for many other applications. For example, if the motion direction is not perfectly perpendicular to the sensor line, i.e., has a non-zero motion component in the direction of the sensor line, skew pixels will be obtained (Steger et al. 2018, Chapter 3.9.3.4). Furthermore, if the speed of the motion is not perfectly adjusted, rectangular instead of square pixels will be obtained. Both effects cause many image processing operations that (often implicitly) assume square pixels to return false or unusable results. Examples are segmentation and feature extraction (e.g., the computation of moments or shape features from segmented regions), 2D template matching approaches, and stereo matching. Unfortunately, a perfect alignment of the camera is very cumbersome to realize in practice. For example, a sensor line with 16,384 pixels would have to be mounted with an accuracy of $1/140°$ in order to keep the total skew of the image below one pixel. Camera calibration allows us to rectify the images in order to eliminate lens distortions or skew, for example, and to ensure square pixels, hence making an exact alignment of the sensor line unnecessary.

In this paper, we introduce a versatile camera model for line-scan cameras with telecentric lenses. We first discuss work relating to camera models for entocentric line-scan cameras in Sect. 2. We then discuss the camera models for area-scan and entocentric line-scan cameras on which our model is based in Sect. 3. In Sect. 4, we describe our camera model for telecentric line-scan cameras, the calibration of its parameters, its relation to affine cameras, as well as the camera model's degeneracies and how to handle them. Various experiments that establish the validity and accuracy of the model are described in Sect. 5. Finally, Sect. 6 concludes the paper.

Our main contributions are the following:

– We propose a comprehensive and versatile camera model for line-scan cameras with telecentric lenses. The camera model allows to model a very large class of lens distortions. It does not assume that the sensor line is aligned with the optical axis of the lens. To the best of our knowledge, no camera model for line-scan cameras with telecentric lenses has ever been proposed in the scientific literature.
– The parameterization of the camera model is very intuitive for machine vision users. All parameters have a physical meaning that is easy to understand.
– We prove that for the division model of lens distortions, the projection of a 3D point to the image can be computed analytically.
– We establish that images of telecentric line-scan cameras with lens distortions and a potential skew can be rectified to have no lens distortions and no skew without knowing the 3D geometry of the scene in the image. This is in contrast to line-scan cameras with entocentric lenses, where this is impossible in general.

– We propose a camera calibration algorithm that determines the camera parameters using images of a planar calibration object.
– We examine how telecentric line-scan cameras without lens distortions are related to affine cameras and prove that every affine camera be regarded as a telecentric line-scan camera with appropriately chosen interior orientation parameters. We also show that every telecentric line-scan camera without lens distortions has an equivalent area-scan camera with a bilateral telecentric tilt lens.
– We comprehensively examine the degeneracies of the camera model and propose methods to handle them.
– We perform an extensive evaluation that establishes the validity and versatility of our camera model.
– We show that even for lenses with very small lens distortions, the distortions are statistically highly significant and therefore cannot be omitted in real-world applications.

2 Related Work

We have been unable to find any research relating to camera models for line-scan cameras with telecentric lenses. The research closest to the camera model we will propose is research on camera models for line-scan cameras with entocentric lenses. Therefore, we will discuss these approaches in this section. In our experience, it is very important for machine vision users to have a camera model with explicit parameters that are easy to understand and have a physical meaning. Hence, in the taxonomy of Sturm et al. (2010, Section 3), we require a global and not a local or discrete camera model. Therefore, although in principle any camera could be modeled by a generalized camera model, such as those described by Sturm et al. (2010, Sections 3.2 and 3.3) or Ramalingam and Sturm (2017), we do not consider and discuss these approaches.

Camera models for line-scan cameras with entocentric lenses can be grouped into two categories: static and dynamic camera models. In static camera models, the motion of the camera is not taken into account in the camera model and therefore is not calibrated. In contrast, dynamic camera models take the camera motion into account and calibrate it. In dynamic models, a linear camera motion is typically assumed. As discussed in Sect. 1, the ability to model the linear camera motion is essential in machine vision applications. Therefore, static camera models are less relevant there.

Another distinction is whether a camera model is able to model lens distortions. Machine vision applications often have very high accuracy requirements that can only be achieved if the camera model is able to model lens distortions accurately. Furthermore, the large size of the sensors also makes the ability to model lens distortions essential. For example, a maximum lens distortion of 0.1% (a figure often specified in data sheets of telecentric lenses) will cause a distortion of more than 8 pixels at the left and right edges of the image for a line-scan sensor with 16,384 pixels.

Even if a camera model supports lens distortions, often the model assumes that the sensor line is mounted exactly behind the principal point, i.e., that the optical axis intersects the sensor line. In a real camera, this assumption may not be fulfilled. In fact, in photogrammetry and remote sensing, multiple sensor lines are often mounted behind the lens at large distances to the principal point (Chen et al. 2003). Therefore, to provide general applicability, the camera model should be able to represent cameras in which the line sensor is not mounted directly behind the principal point.

A final distinction is whether a planar (2D) or a 3D calibration object is used to calibrate the camera. In machine vision applications, there is often limited space in the machine where the camera must be calibrated. Here, 3D calibration objects may be too cumbersome to handle or may not even fit into the available space. In contrast, planar calibration targets typically are much easier to handle for the users. Furthermore, a backlight illumination is used frequently. For these applications, 2D calibration targets are much easier to manufacture. Finally, 2D calibration targets can be produced more cheaply and accurately than 3D calibration objects.

As a result of the above discussion, we require that the camera model is dynamic, supports lens distortions, allows the line sensor to be mounted anywhere with respect to the principal point or optical axis, and uses a planar calibration object. In the following, we will not discuss every camera model in detail. Instead, we will only mention the requirements that the respective model does not fulfill.

Static entocentric line-scan camera models are described by Horaud et al. (1993), Luna et al. (2010), Lilienblum et al. (2013), Yao et al. (2014), Sun et al. (2016a, b, 2017), Niu et al. (2018) and Song et al. (2018). The camera models by Horaud et al. (1993) and Luna et al. (2010) do not model lens distortions. The camera models by Lilienblum et al. (2013), Yao et al. (2014), Niu et al. (2018), Sun et al. (2016a, b, 2017) and Song et al. (2018) model lens distortions, but assume that the sensor is mounted exactly behind the principal point. 3D calibration objects are used by Luna et al. (2010), Lilienblum et al. (2013), Niu et al. (2018) and Song et al. (2018). Finally, Sun et al. (2016b) use an additional area-scan camera to calibrate the line-scan camera, which is undesirable since it increases the cost of the camera setup.

Dynamic entocentric line-scan camera models are proposed by Gupta and Hartley (1997), MVTec Software GmbH (2005a, b), Steger et al. (2008, 2018), Draréni et al. (2011), Hui et al. (2012a, b, 2013) Donné et al. (2017) and Zhang et al. (2018). Lens distortions are not modeled in the camera models by Gupta and Hartley (1997), Draréni et al. (2011), Hui et al. (2012b), Donné et al. (2017) and Zhang et al.

(2018). The camera models by Hui et al. (2012a, 2013) take into account lens distortions, but assume that the sensor is mounted exactly behind the principal point. 3D calibration objects are used by Hui et al. (2012a, b). Furthermore, Hui et al. (2013) use an additional area-scan camera to calibrate the line-scan camera, which is undesirable for the reasons that were mentioned previously. A camera model that fulfills all of the above requirements (i.e., a dynamic camera model that supports lens distortions, allows an arbitrary line sensor position with respect to the principal point, and uses a planar calibration object) is described by MVTec Software GmbH (2005a, b) and Steger et al. (2008, 2018) . We will describe it in more detail in Sect. 3.2 and use it to develop our proposed model for telecentric line-scan cameras.

3 Fundamental Camera Models

The camera model for line-scan cameras with telecentric lenses that we will propose in Sect. 4 is based on camera models for area-scan cameras and on a camera model for line-scan cameras with entocentric lenses. Therefore, we will discuss these models first. Our presentation is based on the descriptions in Steger et al. (2018, Chapter 3.9) and Steger (2017, Section 6).

We start by discussing camera models for area-scan cameras since we will later model line-scan cameras conceptually as one particular line of an area-scan camera. This will enable us to model that the line sensor may not be perfectly aligned with the optical axis of the lens. This, in turn, will allow us to model a more general class of lens distortions for line-scan cameras.

Furthermore, we will describe some of the properties of the existing camera models. This will allow us to compare the properties of the camera model for telecentric line-scan cameras with those of the existing camera models in Sect. 4.

3.1 Camera Models for Area-Scan Cameras

The camera model for area-scan cameras is capable of modeling a multi-view setup with n_c cameras (Steger 2017, Section 6.1). In this paper, we will only consider single cameras. Therefore, we simplify the discussion to this case.

To calibrate the camera, n_o images of a calibration object in different poses are used. Each pose l ($l = 1, \ldots, n_o$) of the calibration object defines a transformation from the calibration object coordinate system to the camera coordinate system. The transformation of a point $\boldsymbol{p}_o = (x_o, y_o, z_o)^\top$ is given by

$$\boldsymbol{p}_l = \mathbf{R}_l \boldsymbol{p}_o + \boldsymbol{t}_l, \tag{1}$$

where $\boldsymbol{t}_l = (t_{l,x}, t_{l,y}, t_{l,z})^\top$ is a translation vector and $\mathbf{R}_l$ is a rotation matrix that is parameterized by Euler angles: $\mathbf{R}_l = \mathbf{R}_x(\alpha_l)\mathbf{R}_y(\beta_l)\mathbf{R}_z(\gamma_l)$. The transformation can also be written as a 4×4 homogeneous matrix:

$$\boldsymbol{p}_l = \mathbf{H}_l \boldsymbol{p}_o = \begin{pmatrix} \mathbf{R}_l & \boldsymbol{t}_l \\ \boldsymbol{0}^\top & 1 \end{pmatrix} \boldsymbol{p}_o, \tag{2}$$

where it is silently assumed that $\boldsymbol{p}_l$ and $\boldsymbol{p}_o$ have been extended with a fourth coordinate of 1.

As discussed by Steger (2017, Section 6.1) and Ulrich and Steger (2019), the origin of the camera coordinate system lies at the center of the entrance pupil of the lens.

Next, the point $\boldsymbol{p}_l = (x_l, y_l, z_l)^\top$ is projected into the image plane. For entocentric lenses, the projection is given by:

$$\begin{pmatrix} x_u \\ y_u \end{pmatrix} = \frac{c}{z_l} \begin{pmatrix} x_l \\ y_l \end{pmatrix}, \tag{3}$$

where c is the principal distance of the lens. For telecentric lenses, the projection is given by:

$$\begin{pmatrix} x_u \\ y_u \end{pmatrix} = m \begin{pmatrix} x_l \\ y_l \end{pmatrix}, \tag{4}$$

where m is the magnification of the lens.[1]

Subsequently, the undistorted point $(x_u, y_u)^\top$ is distorted to a point $(x_d, y_d)^\top$. We support two distortion models (Steger 2017, Section 6.1; Steger et al. 2018, Chapter 3.9.1.3): the division model (Lenz 1987, 1988; Lenz and Fritsch 1990; Lanser et al. 1995; Lanser 1997; Blahusch et al. 1999; Fitzgibbon 2001; Steger 2012) and the polynomial model (Brown 1966, 1971).

In the division model, the undistorted point $(x_u, y_u)^\top$ is computed from the distorted point by:

$$\begin{pmatrix} x_u \\ y_u \end{pmatrix} = \frac{1}{1 + \kappa r_d^2} \begin{pmatrix} x_d \\ y_d \end{pmatrix}, \tag{5}$$

where $r_d^2 = x_d^2 + y_d^2$. The division model can be inverted analytically:

$$\begin{pmatrix} x_d \\ y_d \end{pmatrix} = \frac{2}{1 + \sqrt{1 - 4\kappa r_u^2}} \begin{pmatrix} x_u \\ y_u \end{pmatrix}, \tag{6}$$

where $r_u^2 = x_u^2 + y_u^2$. The division model only supports radial distortion.

---

[1] The camera model for area-scan cameras also supports tilt lenses (Scheimpflug optics); see Steger (2017) for details. Furthermore, it supports hypercentric lenses (Ulrich and Steger 2019).

The polynomial model supports radial as well as decentering distortions. The undistorted point is computed by:

$$\begin{pmatrix} x_\mathrm{u} \\ y_\mathrm{u} \end{pmatrix} = \begin{pmatrix} x_\mathrm{d}\bigl(1 + K_1 r_\mathrm{d}^2 + K_2 r_\mathrm{d}^4 + K_3 r_\mathrm{d}^6\bigr) \\ \quad + \bigl(P_1\bigl(r_\mathrm{d}^2 + 2x_\mathrm{d}^2\bigr) + 2P_2 x_\mathrm{d} y_\mathrm{d}\bigr) \\ y_\mathrm{d}\bigl(1 + K_1 r_\mathrm{d}^2 + K_2 r_\mathrm{d}^4 + K_3 r_\mathrm{d}^6\bigr) \\ \quad + \bigl(2P_1 x_\mathrm{d} y_\mathrm{d} + P_2\bigl(r_\mathrm{d}^2 + 2y_\mathrm{d}^2\bigr)\bigr) \end{pmatrix}. \tag{7}$$

The polynomial model cannot be inverted analytically. The computation of the distorted point from the undistorted point must be performed by a numerical root finding algorithm.

Finally, the distorted point $(x_\mathrm{d}, y_\mathrm{d})^\top$ is transformed into the image coordinate system:

$$\begin{pmatrix} x_\mathrm{i} \\ y_\mathrm{i} \end{pmatrix} = \begin{pmatrix} x_\mathrm{d}/s_x + c_x \\ y_\mathrm{d}/s_y + c_y \end{pmatrix}. \tag{8}$$

Here, s_x and s_y denote the pixel pitches on the sensor and $(c_x, c_y)^\top$ is the principal point. Note that x refers to the horizontal axis of the image (increasing rightward) and y to the vertical axis (increasing downward).

The calibration of the camera model is described in detail in Steger (2017, Sections 9 and 10).

Remark 1 The above parameterization is very intuitive for machine vision users (Steger 2017, Section 6.1). All parameters have a physical meaning that is easy to understand. Approximate initial values for the interior orientation parameters simply can be read off the data sheets of the camera (s_x and s_y) and the lens (c or m) or can be obtained easily otherwise (the initial values for the principal point can be set to the center of the image and the distortion coefficients can typically be set to 0). Furthermore, the calibration results are easy to check for validity.

Remark 2 For telecentric lenses, $(c_x, c_y)^\top$ is solely defined by the lens distortions (Steger 2017, Remark 2). If there are no lens distortions, $(c_x, c_y)^\top$ and $(t_{l,x}, t_{l,y})^\top$ have the same effect. Therefore, in this case $(c_x, c_y)^\top$ should remain fixed at the initial value specified by the user (typically, the image center).

Remark 3 For telecentric lenses, the pose parameter $t_{l,z}$ obviously cannot be determined. We arbitrarily set it to 1 m (Steger 2017, Remark 4).

Remark 4 For telecentric cameras and planar calibration objects, the rotation part of the pose can only be determined up to a twofold ambiguity from a single camera (Steger 2017, Remark 5). This is a special case of a Necker reversal (Shapiro et al. 1995, Section 4.1) when the object is planar. The two sets of pose parameters $(\alpha_l, \beta_l, \gamma_l)$ and $(-\alpha_l, -\beta_l, \gamma_l)$ (with identical translation vectors) result in the same points in the image (Steger 2018, Section 2.4). If a correct exterior orientation of the calibration object is required in the application,

the user must resolve this ambiguity by selecting the correct pose based on prior knowledge.

Remark 5 An operation that we will use below is the calculation of the optical ray of an image point (also called camera ray or line of sight). We first invert (8):

$$\begin{pmatrix} x_\mathrm{d} \\ y_\mathrm{d} \end{pmatrix} = \begin{pmatrix} s_x(x_\mathrm{i} - c_x) \\ s_y(y_\mathrm{i} - c_y) \end{pmatrix}. \tag{9}$$

Then, we rectify the lens distortions by applying (5) or (7). Now, for entocentric lenses, the optical ray is given by:

$$(0, 0, 0)^\top + \lambda(x_\mathrm{u}, y_\mathrm{u}, c)^\top, \tag{10}$$

while for telecentric lenses, it is given by:

$$(x_\mathrm{u}/m, y_\mathrm{u}/m, 0)^\top + \lambda(0, 0, 1)^\top. \tag{11}$$

Remark 6 In machine vision applications, it is often desirable to remove the lens distortions from an image or from data, such as subpixel-precise contours, that were extracted from an image. For example, fitting a line to a contour only returns useful results if the lens distortions have been removed. Since (5) and (7) represent transformations that are performed purely within the image plane, this can easily be achieved. Furthermore, often it is also desirable to remove perspective distortions as well as lens distortions. This can be achieved from a single image if the scene exhibits a known geometry, e.g., a plane, by intersecting the optical ray with the plane (Steger et al. 2018, Section 3.9.5.4).

3.2 Camera Model for Line-Scan Cameras with Entocentric Lenses

The camera model for line-scan cameras was first described in MVTec Software GmbH (2005a, b) and Steger et al. (2008). Our discussion is based on Steger et al. (2018, Chapter 3.9.3).

As described in Sect. 1, we assume that the relative motion between the line-scan camera and the object is linear with constant velocity. Therefore, the camera motion can be described by the motion vector $\mathbf{v} = (v_x, v_y, v_z)^\top$. The vector $\mathbf{v}$ is described in units of meters per scan line in the camera coordinate system (i.e., the units are $\mathrm{m\,Pixel}^{-1}$).[2] This definition of $\mathbf{v}$ assumes a moving camera and a fixed object. If the camera is stationary and the object is moving, e.g., on a conveyor belt, we can simply use $-\mathbf{v}$ as the motion vector (see Fig. 1).

---

[2] If the speed of the motion vector $\mathbf{v}$ is specified in world units (i.e., $\mathrm{m\,s}^{-1}$), it must be divided by the readout rate of the camera (units: $\mathrm{Pixel\,s}^{-1}$, i.e., scan lines per second) to obtain the motion vector in meters per scan line.
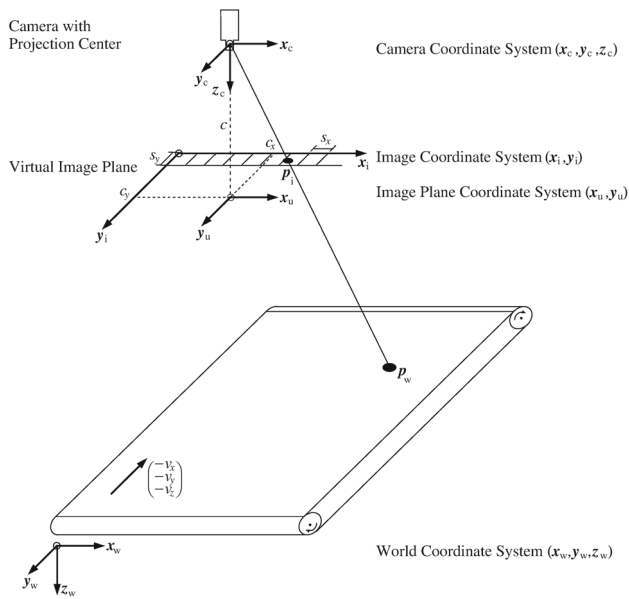
Fig. 1 Camera model for line-scan cameras with entocentric lenses

The camera model for line-scan cameras is displayed in Fig. 1. The origin of the camera coordinate system lies at the center of the entrance pupil of the lens. The z axis is identical to the optical axis and is oriented such that points in front of the camera have positive z coordinates. The $\boldsymbol{x}$ axis is parallel to the sensor line and perpendicular to the $\boldsymbol{z}$ axis. It points rightward in the image. The $\boldsymbol{y}$ axis is perpendicular to the sensor line and to the $\boldsymbol{z}$ axis such that a right handed coordinate system is obtained.

Like for area-scan cameras, the transformation from the calibration object coordinate system to the camera coordinate system is given by (1). In contrast to area-scan cameras, this exterior orientation refers only to the first line of the image. Since the camera moves relative to the object, the exterior orientation is different for each line. However, because we assume a linear motion, the motion vector $\boldsymbol{v}$ can be used to compute the exterior orientation of all lines. Therefore, the single exterior orientation is sufficient.

Since we want to be able to model line-scan cameras for which the sensor line is not perfectly aligned with the optical axis, we model the sensor line as one particular line of a virtual area-scan camera. We use the principal point $(c_x, c_y)^\top$ to model this misalignment (cf. Fig. 1). The semantics of c_y are slightly different than for area-scan cameras: $c_y = 0$ signifies that the sensor line is perfectly aligned with the optical axis in the $\boldsymbol{y}$ direction.

The remaining parameters of the model are identical to those of area-scan cameras (see Sect. 3.1): c is the principal distance, the lens distortions are described by (5) or (7), and s_x and s_y describe the pixel pitch on the sensor.

To compute the projection of a point $\boldsymbol{p}_c = (x_c, y_c, z_c)^\top$ that has been transformed into the camera coordinate sys-

tem,[3] we can use the fact that $\boldsymbol{p}_c$ moves along the straight line $\boldsymbol{p}_c - t\boldsymbol{v}$, where t denotes the number of scan lines that have been acquired since the first scan line. As the point moves, it must at some point intersect the optical ray of an image point if it projects to a point $\boldsymbol{p}_s = (x_s, 0)^\top$ on the sensor line.[4] This optical ray is given by (9) and (10).

Let us assume that we have transformed the point $\boldsymbol{p}_s$ to a distorted image point $\boldsymbol{p}_d$ by (9), i.e., $(x_d, y_d)^\top = (s_x(x_s - c_x), -s_y c_y)^\top$. Furthermore, let us call the undistortion function in (5) or (7) $\boldsymbol{u}(\boldsymbol{p}) = (u_x(x_d, y_d), u_y(x_d, y_d))^\top$. Then, the intersection of the moving point and the optical ray results in the following equation system:

$$\begin{aligned}
\lambda u_x(x_d, y_d) &= x_c - t v_x \\
\lambda u_y(x_d, y_d) &= y_c - t v_y \\
\lambda c &= z_c - t v_z.
\end{aligned} \tag{12}$$

The equation system (12), which for both distortion models is a polynomial equation system, must be solved for λ, t, and x_d. Once t and x_d have been determined, the point is transformed into the image coordinate system by:

$$\begin{pmatrix} x_i \\ y_i \end{pmatrix} = \begin{pmatrix} x_d/s_x + c_x \\ t \end{pmatrix}. \tag{13}$$

Thus, the interior orientation of line-scan cameras with entocentric lenses is given by: c; κ or K_1, K_2, K_3, P_1, P_2; s_x, s_y, c_x, c_y, v_x, v_y, and v_z.[5]

Remark 7 The model is overparameterized. The values of c and s_x cannot be determined simultaneously. This can be solved by fixing s_x at the initial value that was specified by the user. Furthermore, since there is only one image line, s_y is only used to specify the principal point in pixels. It has no physical meaning on the sensor and only occurs in the term $y_d = -s_y c_y$. Therefore, it cannot be determined and is kept fixed at the initial value specified by the user. In practice, $s_y = s_x$ is typically used.

Remark 8 The value of c_y is solely defined by the lens distortions. If there are no lens distortions, c_y should remain fixed at the initial value specified by the user (typically, $c_y = 0$ is used).

Remark 9 If $c_y = 0$, the effect of the lens distortions is purely along the sensor line, i.e., in the horizontal direction. If $c_y \neq$

---

[3] This corresponds to the point $\boldsymbol{p}_l$ in (1).

[4] The above definition of the principal point implies that points on the sensor line have a $\boldsymbol{y}$ coordinate of 0.

[5] One could also argue that the motion vector $\boldsymbol{v}$ logically belongs to the exterior orientation of the camera. However, as (12) shows, the motion vector affects the projection of a point given in the camera coordinate system. Therefore, we include it in the interior orientation.

0, lens distortions will also cause the sensor line to appear bent (Steger et al. 2018, Chapter 3.9.3.4).

Remark 10 The parameters P_1 and P_2 of the polynomial distortion model are highly correlated with other parameters of the entocentric line-scan camera model, especially if the radial distortions are small. Therefore, they typically cannot be determined reliably. Consequently, in practice, they should be set to 0 and should be excluded from the calibration.

Remark 11 In contrast to area-scan cameras, where the aspect ratio of the pixels is determined by s_x and s_y, the aspect ratio of the pixels is determined by s_x and v_y for line-scan cameras. Furthermore, in contrast to area-scan cameras, for which we have assumed orthogonal pixels, for line-scan cameras, a nonzero v_x will result in pixels that appear skewed (i.e., non-orthogonal). Therefore, to achieve square pixels, it is important that the sensor is aligned to be perpendicular to the motion vector and the motion speed or the line frequency of the camera must be selected appropriately (Steger et al. 2018, Chapter 3.9.3.4).

Remark 12 A pure removal of lens distortions is impossible for entocentric line-scan cameras if $c_y \neq 0$. To remove the lens distortions, we would have to compute the optical ray for an image point and then would have to reproject it into the rectified image. However, if there are lens distortions and if $c_y \neq 0$, the optical ray in general does not project to a single point in the rectified image. Therefore, a pure removal of lens distortions requires a 3D reconstruction to be available because we can then reconstruct a unique 3D point that we can project into the rectified image. What is possible, on the other hand, is to reproject the image onto a world plane (Steger et al. 2018, Chapter 3.9.5.4). This rectification also removes the lens distortions, of course.

4 Camera Model for Line-Scan Cameras with Telecentric Lenses

4.1 Camera Model

The first step of the camera model is identical to that in Sect. 3.2: We transform points from the calibration object coordinate system to the camera coordinate system by (1). Again, we will call the transformed point $\boldsymbol{p}_c$.

To project the point into the image, we use the same approach as in Sect. 3.2: We intersect the line on which the point moves with the optical ray of the point to which it projects. The equation of the optical ray is given by (11). This results in the following equation system:

$$u_x(x_d, y_d)/m = x_c - t v_x \tag{14}$$

$$u_y(x_d, y_d)/m = y_c - t v_y \tag{15}$$

$$\lambda = z_c - t v_z, \tag{16}$$

where $\boldsymbol{u}(\boldsymbol{p}) = (u_x(x_d, y_d), u_y(x_d, y_d))^\top$ is defined as in Sect. 3.2. It can be seen that λ does not occur in (14) and (15). Therefore, neither z_c nor v_z influence the projection and we can omit (16). Consequently, line-scan cameras with telecentric lenses perform an orthographic projection, similar to area-scan cameras with telecentric lenses. Thus, with respect to the taxonomy by Sturm et al. (2010), line-scan cameras with telecentric lenses are central cameras (Sturm et al. 2010, Section 3), unlike line-scan cameras with entocentric lenses, which are axial cameras (Ramalingam et al. 2006; Sturm et al. 2010, Section 3.1.4). Furthermore, with respect to the taxonomy by Ye and Yu (2014), line-scan cameras with telecentric lenses are orthographic cameras and not pushbroom cameras.

For the polynomial model, Eqs. (14) and (15) define a polynomial equation system of degree 7 in the unknowns x_d and t. Therefore, the equations cannot be solved analytically, i.e., a numerical root finding algorithm must be used to solve them. For the division model, however, an analytical solution is possible. Specializing (14) and (15) to the division model results in the following equation system:

$$u x_d/m = x_c - t v_x \tag{17}$$

$$u y_d/m = y_c - t v_y, \tag{18}$$

where $u = 1/(1+\kappa(x_d^2 + y_d^2))$. Since $y_d = -s_y c_y$ is constant, we solve (18) for t:

$$t = \frac{1}{v_y}\left(y_c - \frac{y_d}{m\left(1 + \kappa\left(x_d^2 + y_d^2\right)\right)}\right). \tag{19}$$

Substituting (19) into (17) results in:

$$\frac{x_d}{m\left(1 + \kappa\left(x_d^2 + y_d^2\right)\right)} = x_c - \frac{v_x}{v_y}\left(y_c - \frac{y_d}{m\left(1 + \kappa\left(x_d^2 + y_d^2\right)\right)}\right). \tag{20}$$

If we multiply both sides by $1+\kappa(x_d^2 + y_d^2)$, expand the terms, and sort them according to powers of x_d, we obtain:

$$\kappa\left(x_c - y_c \frac{v_x}{v_y}\right)x_d^2 - \frac{1}{m}x_d$$
$$+ \left(1 + \kappa y_d^2\right)\left(x_c - y_c \frac{v_x}{v_y}\right) + \frac{y_d}{m}\frac{v_x}{v_y} = 0. \tag{21}$$

The term $x_c - y_c v_x / v_y$ is the x coordinate of the point at which the line $(x_c, y_c)^\top - t(v_x, v_y)^\top$ intersects the x axis (or, in 3D, at which the line $(x_c, y_c, z_c)^\top - t(v_x, v_y, v_z)^\top$ intersects the xz plane). Let us call this term x_0. The term $1 + \kappa y_d^2$ represents the inverse of the undistortion factor u for $x_d = 0$. Let us call this term d_0. Then, we have:

$$\kappa x_0 x_d^2 - \frac{1}{m} x_d + x_0 d_0 + \frac{y_d}{m} \frac{v_x}{v_y} = 0. \tag{22}$$

Hence, for $\kappa \neq 0$ and $x_0 \neq 0$, we have:

$$x_d = \frac{\frac{1}{m} \pm \sqrt{\frac{1}{m^2} - 4\kappa x_0 \left(x_0 d_0 + \frac{y_d}{m} \frac{v_x}{v_y} \right)}}{2\kappa x_0}. \tag{23}$$

For $\kappa = 0$ or $x_0 = 0$, (22) reduces to a linear equation. We examine both cases in turn. For $x_0 = 0$, we have:

$$x_d = y_d \frac{v_x}{v_y}. \tag{24}$$

Inserting the value of x_d obtained from (23) or (24) into (19) returns the value of t in both cases. For $\kappa = 0$, we have:

$$x_d = m x_0 + y_d \frac{v_x}{v_y}. \tag{25}$$

In this case, Eq. (19) can be simplified to:

$$t = \frac{1}{v_y} \left(y_c - \frac{y_d}{m} \right). \tag{26}$$

Note that for $\kappa = 0$, y_d, i.e., c_y, is not meaningful (cf. Remarks 8 and 20). Therefore, if it is known a priori that $\kappa = 0$, c_y (and, therefore, y_d) should be set to 0, which simplifies the equations even further to:

$$x_d = m x_0 = m \left(x_c - y_c \frac{v_x}{v_y} \right) \tag{27}$$

$$t = \frac{y_c}{v_y}. \tag{28}$$

If there are lens distortions, we can see from (23) that there are two potential solutions for the projection into the image, whereas in the cases without distortion, (24) and (25), there is a unique solution. Intuitively, we expect that for the case with lens distortions, there also is a unique solution since there is only one particular instant of time when the point will appear in front of the sensor line.

Proposition 1 *In* (23)*, the correct solution is given by:*

$$x_d = \frac{\frac{1}{m} - \sqrt{\frac{1}{m^2} - 4\kappa x_0 \left(x_0 d_0 + \frac{y_d}{m} \frac{v_x}{v_y} \right)}}{2\kappa x_0}. \tag{29}$$

Proof To prove the assertion, we will examine the limit of (23) for $\kappa \to 0$. Obviously, this solution must converge to (25) for the correct solution because (17) and (18) are continuous around $\kappa = 0$. We first examine the solution in (29) and note that both the numerator and denominator converge to 0 for $\kappa \to 0$. Therefore, we use L'Hôpital's rule to compute the limit:

$$\lim_{\kappa \to 0} \frac{\frac{1}{m} - \sqrt{\frac{1}{m^2} - 4\kappa x_0 \left(x_0 d_0 + \frac{y_d}{m} \frac{v_x}{v_y} \right)}}{2\kappa x_0}$$

$$= \lim_{\kappa \to 0} \frac{\frac{d}{d\kappa} \left(\frac{1}{m} - \sqrt{\frac{1}{m^2} - 4\kappa x_0 \left(x_0 d_0 + \frac{y_d}{m} \frac{v_x}{v_y} \right)} \right)}{\frac{d}{d\kappa} (2\kappa x_0)}$$

$$= \lim_{\kappa \to 0} -\frac{-\left(x_0 d_0 + \frac{y_d}{m} \frac{v_x}{v_y} \right) - \kappa y_d^2 x_0}{\sqrt{\frac{1}{m^2} - 4\kappa x_0 \left(x_0 d_0 + \frac{y_d}{m} \frac{v_x}{v_y} \right)}}$$

$$= m x_0 + y_d \frac{v_x}{v_y}. \tag{30}$$

Hence, Eq. (29) converges to (25) for $\kappa \to 0$. We now examine the solution of (23) with the plus sign and note that the numerator converges to $2/m$ for $\kappa \to 0$, while the denominator converges to 0. Hence, the second solution converges to ∞ for $\kappa \to 0$. Therefore, Eq. (29) is the correct solution. $\qquad \square$

Remark 13 The optical ray of an image point $(x_i, y_i)^\top$ can be computed as follows. First, Eq. (13) is inverted:

$$\begin{pmatrix} x_d \\ t \end{pmatrix} = \begin{pmatrix} s_x(x_i - c_x) \\ y_i \end{pmatrix}. \tag{31}$$

Next, Eqs. (14) and (15) are solved for $(x_c, y_c)^\top$:

$$x_c = u_x(x_d, y_d)/m + t v_x \tag{32}$$

$$y_c = u_y(x_d, y_d)/m + t v_y, \tag{33}$$

where $y_d = -s_y c_y$. The optical ray is then given by:

$$(x_c, y_c, 0)^\top + \lambda(0, 0, 1)^\top. \tag{34}$$

Remark 14 In contrast to line-scan cameras with entocentric lenses (cf. Remark 12), a pure removal of lens distortions is possible for line-scan cameras with telecentric lenses because (14) and (15) do not depend on z_c. Given an image point $(x_i, y_i)^\top$, the corresponding point $(x_c, y_c)^\top$ in the camera coordinate system can be computed, as described in Remark 13. This point can then be projected into a rectified camera for which all distortion coefficients have been

set to 0. Moreover, any skew in the pixels can be removed by setting v_x to 0 in the rectified camera. Finally, square pixels can be enforced by setting s_x to $\min(s_x, mv_y)$ and then setting v_y to s_x/m. This approach ensures that no aliasing occurs when rectifying the image.

Remark 15 As is the case for line-scan cameras with entocentric lenses (cf. Remark 12), a reprojection of the image onto a world plane is possible for line-scan cameras with telecentric lenses.

4.2 Calibration

Like the camera models that were described in Sect. 3, the camera is calibrated by using the planar calibration object introduced in Steger (2017, Section 9). The calibration object has a hexagonal layout of circular control points. It has been designed in such a way that it can cover the entire field of view. Further advantages of this kind of calibration object are discussed in Steger (2017, Section 9).

Let the known 3D coordinates of the centers of the control points of the calibration object be denoted by $\boldsymbol{p}_j$ ($j = 1, \ldots, n_m$, where n_m denotes the number of control points on the calibration object). The user acquires n_o images of the calibration object. Let us denote the exterior orientation parameters of the calibration object in image l by $\boldsymbol{e}_l$ ($l = 1, \ldots, n_o$), the interior orientation parameters of the camera by $\boldsymbol{i}$, and the projection of a point in the calibration object coordinate system to the image coordinate system by $\boldsymbol{\pi}$ (cf. Sect. 4.1). In addition, let v_{jl} denote a function that is 1 if the control point j of the observation l of the calibration object is visible with the camera, and 0 otherwise. Finally, let $\boldsymbol{p}_{jl}$ denote the position of control point j in image l. Then, the camera is calibrated by minimizing the following function:

$$\varepsilon^2 = \sum_{l=1}^{n_o} \sum_{j=1}^{n_m} v_{jl} \| \boldsymbol{p}_{jl} - \boldsymbol{\pi}(\boldsymbol{p}_j, \boldsymbol{e}_l, \boldsymbol{i}) \|_2^2. \tag{35}$$

The minimization is performed by a suitable version of the sparse Levenberg–Marquardt algorithms described in Hartley and Zisserman (2003, Appendix A6).

The points $\boldsymbol{p}_{jl}$ are extracted by fitting ellipses (Fitzgibbon et al. 1999) to edges extracted with a subpixel-accurate edge extractor (Steger 1998b, Chapter 3.3; Steger 2000). As discussed by Steger (2017, Section 5.2) and Mallon and Whelan (2007), this causes a bias in the point positions. Since telecentric line-scan cameras perform an orthographic projection, there is no perspective bias, i.e., the bias consists solely of distortion bias. The bias can be removed with the approach for entocentric line-scan cameras described by Steger (2017, Section 10).

The optimization of (35) requires initial values for the unknown parameters. Initial values for the interior orientation

parameters, except for the motion vector, can be obtained from the specification of the camera and the lens, as described in Remark 1. In contrast to area-scan cameras, $c_y = 0$ is typically used as the initial value. An approximate value for v_y usually will be known from the considerations that led to the line-scan camera setup. Finally, v_x typically can be set to 0. With known initial values for the interior orientation, the image points $\boldsymbol{p}_{jl}$ can be transformed into metric coordinates in the camera coordinate system using (31)–(33). This allows us to use the OnP algorithm described by Steger (2018) to obtain estimates for the exterior orientation of the calibration object.

Remark 16 The Levenberg–Marquardt algorithm requires the partial derivatives of x_i and y_i with respect to the interior and exterior orientation parameters of the camera model. This, in turn, requires the partial derivatives of x_d and t with respect to the interior orientation parameters. These can be computed analytically using the implicit function theorem (de Oliveira 2013).

4.3 Line-Scan Cameras with Telecentric Lenses in Projective Geometry

In this section, we will consider telecentric line-scan cameras without lens distortions. In this case, Eqs. (27), (28), and (13) can be written as the following calibration matrix:

$$\mathbf{K} = \begin{pmatrix} a & -a\,v_x/v_y & 0 \\ 0 & 1/v_y & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{36}$$

Since m and s_x cannot be determined simultaneously (cf. Remark 19), we have removed this overparameterization by using the parameter $a = m/s_x$. Furthermore, since the principal point is undefined if there are no distortions (see Remark 20), we have used $c_x = c_y = 0$.[6] The orthographic projection that the telecentric line-scan camera performs can be written as:

$$\mathbf{O} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{37}$$

Finally, the pose of the camera with respect to the world coordinate system can be written as a 4×4 homogeneous matrix $\mathbf{H}$ (see Sect. 3.1). Hence, if there are no distortions, line-scan cameras with telecentric lenses are affine cameras (Hartley

---

[6] A principal point $(c_x, c_y)^\top \neq (0, 0)^\top$ would lead to $\mathbf{K}_{1,3} = -c_y(s_y/s_x)(v_x/v_y) + c_x$ because of (25) and (13) and to $\mathbf{K}_{2,3} = c_y s_y/(mv_y)$ because of (26). We will refer to these values as c_x' and c_y' where necessary.

and Zisserman 2003, Chapter 6.3.4) with the following camera matrix:

$$\mathbf{P} = \mathbf{KOH}. \tag{38}$$

A general affine camera can be written as the following camera matrix:

$$\mathbf{M} = \begin{pmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{39}$$

where the top left 2×3 submatrix must have rank 2 (Hartley and Zisserman 2003, Chapter 6.3.4). $\mathbf{M}$ has eight degrees of freedom: its elements m_{ij} ($i = 1, 2$, $j = 1, \ldots, 4$) (Hartley and Zisserman 2003, Chapter 6.3.4). The camera matrix in (38) also has eight degrees of freedom: a, v_x, v_y, t_x, t_y, α, β, and γ. Therefore, it is natural to examine whether a general affine camera matrix $\mathbf{M}$ can be decomposed uniquely into the eight parameters of a telecentric line-scan camera without lens distortions.

Theorem 1 *Every affine camera matrix $\mathbf{M}$ can be decomposed into the eight parameters a, v_x, v_y, t_x, t_y, α, β, and γ of a telecentric line-scan camera without lens distortions. There is a twofold ambiguity in the decomposition: If a valid decomposition of $\mathbf{M}$ is given by $(a, v_x, v_y, t_x, t_y, \alpha, \beta, \gamma)$, another valid decomposition is given by $(a, v_x, -v_y, t_x, -t_y, \alpha + \pi, \beta, \gamma)$.*

Proof To prove Theorem 1, we will make use of the dual image of the absolute conic (DIAC) (Hartley and Zisserman 2003, Chapter 8.5), given by

$$\boldsymbol{\omega}^* = \mathbf{P}\mathbf{Q}_\infty^*\mathbf{P}^\top, \tag{40}$$

where $\mathbf{Q}_\infty^* = \mathrm{diag}(1, 1, 1, 0)$ is the canonical form of the absolute dual quadric (Hartley and Zisserman 2003, Chapter 3.7).[7] This will allow us to remove the exterior orientation from the equations to be solved. If we denote the entries of a camera matrix $\mathbf{M}$ by m_{ij}, the elements ω_{ij} of the DIAC $\boldsymbol{\omega}^*$ are given by

$$\omega_{ij} = \sum_{k=1}^{3} m_{ik} m_{jk}. \tag{41}$$

Note that $\boldsymbol{\omega}^*$ is a symmetric matrix.

Let us denote the DIAC of $\mathbf{M}$ by $\boldsymbol{\omega}_\mathbf{M}^*$ and the DIAC of the camera matrix (38) by $\boldsymbol{\omega}_\mathbf{P}^*$. We require that both DIACs are identical:

$$\boldsymbol{\omega}_\mathbf{P}^* = \boldsymbol{\omega}_\mathbf{M}^*. \tag{42}$$

---

[7] The function diag constructs a diagonal matrix with the specified elements.

The DIAC $\boldsymbol{\omega}_\mathbf{P}^*$ is given by:

$$\boldsymbol{\omega}_\mathbf{P}^* = \mathbf{KK}^\top = \begin{pmatrix} a^2\left(1 + \frac{v_x^2}{v_y^2}\right) & -a\frac{v_x}{v_y^2} & 0 \\ -a\frac{v_x}{v_y^2} & \frac{1}{v_y^2} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{43}$$

Hence, we have the following three equations to determine a, v_x, and v_y:

$$\omega_{11} = a^2\left(1 + \frac{v_x^2}{v_y^2}\right) \tag{44}$$

$$\omega_{12} = -a\frac{v_x}{v_y^2} \tag{45}$$

$$\omega_{22} = \frac{1}{v_y^2}. \tag{46}$$

We can solve (46) for v_y:

$$v_y = \pm\frac{1}{\sqrt{\omega_{22}}}. \tag{47}$$

Substituting v_y into (45) and solving for v_x results in:

$$v_x = -\frac{\omega_{12}}{a\omega_{22}}. \tag{48}$$

Substituting v_x and v_y into (44) and solving for a yields:

$$a = \pm\sqrt{\omega_{11} - \frac{\omega_{12}^2}{\omega_{22}}}. \tag{49}$$

We can assume that $a = m/s_x$ is positive. Hence, only the positive square root in (49) yields a valid result:

$$a = \sqrt{\omega_{11} - \frac{\omega_{12}^2}{\omega_{22}}}. \tag{50}$$

Substituting a into (48) results in:

$$v_x = -\frac{\omega_{12}}{\sqrt{\omega_{22}(\omega_{11}\omega_{22} - \omega_{12}^2)}}. \tag{51}$$

Note that $\omega_{11} > 0$ and $\omega_{22} > 0$ because $\mathbf{M}$ has rank 3 and that $\omega_{11}\omega_{22} - \omega_{12}^2 \geq 0$ (and, therefore, $\omega_{11} - \omega_{12}^2/\omega_{22} \geq 0$) because of the Cauchy-Schwarz inequality (Steger 2017, Appendix A.1). Consequently, all equations can always be solved.

The above derivation shows that that there are two solutions: (a, v_x, v_y) and $(a, v_x, -v_y)$, where we have selected $v_y = 1/\sqrt{\omega_{22}}$. These solutions allow us to compute the respective calibration matrix $\mathbf{K}$. If we compute $\mathbf{K}^{-1}\mathbf{M}$, we obtain the first two rows of $\mathbf{H}$ (Steger 2017, Appendix A.1).

The left 2×3 submatrix of $\mathbf{K}^{-1}\mathbf{M}$ contains the first two rows of the rotation matrix of the pose. The third row of the rotation matrix can be computed as the vector product of the first two rows. This rotation matrix can then be decomposed into the parameters α, β, and γ. The right 2×1 submatrix of $\mathbf{K}^{-1}\mathbf{M}$ contains t_x and t_y.

We now examine what effect the two different solutions for the interior orientation parameters have on the pose. The two different calibration matrices are given by:

$$\mathbf{K}_1 = \begin{pmatrix} a & -a\,v_x/v_y & 0 \\ 0 & 1/v_y & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{52}$$

$$\mathbf{K}_2 = \begin{pmatrix} a & a\,v_x/v_y & 0 \\ 0 & -1/v_y & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{53}$$

Their inverses are given by:

$$\mathbf{K}_1^{-1} = \begin{pmatrix} 1/a & v_x & 0 \\ 0 & v_y & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{54}$$

$$\mathbf{K}_2^{-1} = \begin{pmatrix} 1/a & v_x & 0 \\ 0 & -v_y & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{55}$$

Since the only difference between the two inverses is the sign of the element $(2, 2)$, this means that when $\mathbf{H}_{1,2}$ are computed as $\mathbf{K}_{1,2}^{-1}\mathbf{M}$, the two solutions $\mathbf{H}_1$ and $\mathbf{H}_2$ will have inverse signs in their second row. If the third row of the rotation matrix is computed by the vector product, the two rotation matrices $\mathbf{R}_1$ and $\mathbf{R}_2$ will differ by having inverse signs in their second and third rows. This corresponds to a multiplication by a matrix $\mathbf{Q} = \mathrm{diag}(1, -1, -1)$ on the left. Note that $\mathbf{Q}$ is a rotation by π around the $\mathbf{x}$ axis. Since in our Euler angle representation the rotation around the $\mathbf{x}$ axis is performed last (see Sect. 3.1), multiplying by $\mathbf{Q}$ on the left corresponds to adding π to α. This shows that for the first solution (a, v_x, v_y) of the interior orientation parameters, the solution for the pose parameters is given by $(t_x, t_y, \alpha, \beta, \gamma)$, while for the second solution $(a, v_x, -v_y)$ of the interior orientation parameters, the solution for the pose parameters is given by $(t_x, -t_y, \alpha + \pi, \beta, \gamma)$. □

Remark 17 A rotation by π around the $\mathbf{x}$ axis corresponds to looking at the front or at the back of an object. Therefore, if the camera is acquiring images of opaque objects, it is typically possible to select the correct solution. However, if images of transparent objects are acquired, which is the case in applications with backlight illumination (at least the calibration object must be transparent in this case), the ambiguity of Theorem 1 can occur in practice.

Remark 18 Theorem 1 shows that every affine camera is equivalent to a telecentric line-scan camera with no distortions. On the other hand, Theorem 1 in Steger (2017) shows that every affine camera is equivalent to an area-scan camera with a bilateral telecentric tilt lens with no distortions. Therfore, telecentric line-scan cameras with no distortions are equivalent to telecentric area-scan cameras (with tilt lenses if the pixels are skewed). This means that we can replace telecentric line-scan cameras with no distortions by telecentric area-scan cameras with no distortions if this is convenient. In particular, this allows us to reuse existing algorithms for telecentric area-scan cameras for telecentric line-scan cameras.

We note that a telecentric line-scan camera that has been rectified with the approach described in Remark 14 fulfills the above criterion of having no distortions. Therefore, a rectified telecentric line-scan camera with parameters $m_1 = m$, $s_{x,1} = s_x$, $s_{y,1} = s_y$, $c_{x,1} = c_x$, $c_{y,1} = c_y$, $v_{x,1} = 0$, and $v_{y,1} = v_y = s_x/m$ can be represented by a telecentric area-scan camera (cf. Sect. 3.1) with $m_a = m$, $s_{x,a} = s_x$, $s_{y,a} = s_x$, $c_{x,a} = c_x$, and $c_{y,a} = c_y s_x/s_y$.

4.4 Model Degeneracies

Remark 19 The model for telecentric line-scan cameras is overparameterized. The values of m and s_x cannot be determined simultaneously. This can be solved by fixing s_x at the initial value that was specified by the user. Furthermore, like for entocentric line-scan cameras, s_y is only used to specify the principal point in pixels (cf. Remark 7) and is therefore kept fixed at the initial value specified by the user.

Remark 20 Like for telecentric area-scan cameras (cf. Remark 2), $(c_x, c_y)^\top$ is solely defined by the lens distortions for telecentric line-scan cameras. If there are no lens distortions, $(c_x, c_y)^\top$ and $(t_{l,x}, t_{l,y})^\top$ have the same effect. Therefore, in this case $(c_x, c_y)^\top$ should remain fixed at the initial value specified by the user (typically, c_x is set to the horizontal image center and c_y is set to 0).

Remark 21 Like for entocentric line-scan cameras (cf. Remark 10), the parameters P_1 and P_2 of the polynomial distortion model are highly correlated with other parameters of the telecentric line-scan camera model, especially if the radial distortions are small. Therefore, they typically cannot be determined reliably. Consequently, in practice, they should be set to 0 and should be excluded from the calibration.

Remark 22 Neither t_z nor v_z can be determined since they have no effect on the projection (cf. Sect. 4.1). We leave v_z at the initial value specified by the user and set t_z to 1 m (see also Remark 3).

Remark 23 As described in Theorem 1 and Remark 17, there is a sign ambiguity for v_y. Therefore, the user must specify the initial value of v_y with the correct sign to ensure the calibration converges to the correct solution.

Remark 24 Like for telecentric area-scan cameras (see Remark 4), the rotation of the pose of a planar calibration object can only be determined up to a twofold ambiguity. This is a special case of a Necker reversal (Shapiro et al. 1995, Section 4.1) when the object is planar. The two sets of pose parameters $(\alpha_l, \beta_l, \gamma_l)$ and $(-\alpha_l, -\beta_l, \gamma_l)$ (with identical translation vectors) result in the same points in the image. If a correct exterior orientation of the calibration object is required in the application, the user must resolve this ambiguity by selecting the correct pose based on prior knowledge.

Proposition 2 *For a planar calibration object, if $c_y = 0$, or if $c_y \neq 0$ and there are no distortions (i.e., $\kappa = 0$ for the division model), and if $\beta = 0$, then v_y, α, and t_y cannot be determined simultaneously.*

Proof Without loss of generality, we can assume that the planar calibration object lies in the plane $z = 0$ in the calibration object coordinate system. Let us first assume we have cameras with $c_y = 0$. Furthermore, let us assume that the first camera has $\alpha_1 = 0$, $t_{y,1} = t_y$, and $v_{y,1} = v_y$, where t_y and v_y are arbitrary but fixed. In addition, let us assume the remaining interior and exterior orientation parameters m, c_x, s_x, s_y, v_x, t_x, and γ are arbitrary and identical for both cameras. We now can select an arbitrary value α for the rotation around the **x** axis for the second camera, i.e., $\alpha_2 = \alpha$. Then, by setting $v_{y,2} = v_y \cos \alpha$ and $t_{y,2} = t_y \cos \alpha$, we obtain a camera with identical projection geometry for points in the plane $z = 0$. If $c_y \neq 0$ and if there are no distortions, we must also set $c_{y,1} = c_y$ and $c_{y,2} = c_y \cos \alpha$.

To prove that both cameras result in the same projection geometry, we can construct their camera matrices $\mathbf{P}_{1,2}$, as described in Sect. 4.3. Then, we can project an arbitrary point $\mathbf{p}_c = (x_c, y_c, 0)^\top$, i.e., we can compute $\mathbf{p}_{i,1,2} = \mathbf{P}_{1,2}\mathbf{p}_c$. Comparing the resulting expressions for $\mathbf{p}_{i,1,2}$, which we omit here, shows that they are identical. □

Proposition 3 *For a planar calibration object, if $c_y = 0$, or if $c_y \neq 0$ and there are no distortions (i.e., $\kappa = 0$ for the division model), and if $\alpha = 0$, then m, v_x, β, and t_x cannot be determined simultaneously.*

Proof Without loss of generality, we can assume that the planar calibration object lies in the plane $z = 0$ in the calibration object coordinate system. Let us first assume we have cameras with $c_y = 0$. Furthermore, let us assume that the first camera has $\beta_1 = 0$, $t_{x,1} = t_x$, $m_1 = m$, and $v_{x,1} = v_x$, where t_x, m, and v_x are arbitrary but fixed. In addition, let us assume the remaining interior and exterior orientation parameters c_x, s_x, s_y, v_y, t_y, and γ are arbitrary and identical for both cameras. We now can select an arbitrary value β for the rotation around the **y** axis for the second camera, i.e., $\beta_2 = \beta$. Then, by setting $m_2 = m/\cos \beta$, $v_{x,2} = v_x \cos \beta$, and $t_{x,2} = t_x \cos \beta$, we obtain a camera with identical projection geometry for points in the plane $z = 0$. If $c_y \neq 0$ and if there are no distortions, we must also set $c_{y,1} = c_y$ and $c_{y,2} = c_y/\cos \beta$.

To prove the assertion, we can proceed in the same manner as in the proof of Proposition 2. □

Remark 25 Proposition 2 shows that a rotation of a planar calibration object around the **x** axis can be exchanged with different values for the speed v_y and the translation t_y. Proposition 3 shows that a rotation of a planar calibration object around the **y** axis can be exchanged with different values for the magnification m, the speed v_x, and the translation t_x. Since the interior and exterior orientation parameters that are affected by these degeneracies are independent of each other, we conjecture that there is a universal degeneracy that implies that the interior and exterior orientation parameters cannot be determined from a single image of a planar calibration object, no matter how the calibration object is oriented in 3D. We prove that this is the case in the following theorem.

Theorem 2 *For a single image of a planar calibration object, m, v_x, v_y, α, β, t_x, and t_y cannot be determined simultaneously if $c_y = 0$, or if $c_y \neq 0$ and there are no distortions (i.e., $\kappa = 0$ for the division model).*

Proof Without loss of generality, we can assume that the planar calibration object lies in the plane $z = 0$. Furthermore, we can immediately see that a rotation around the **z** axis is immaterial since a rotation of a point in the plane $z = 0$ by an angle γ around the **z** axis merely corresponds to a different point in the plane $z = 0$. Consequently, we can use $\gamma = 0$ in the following.

We will now show that a camera with parameters m, s_x, s_y, $c_x = 0$, $c_y = 0$, v_x, v_y, t_x, t_y, α, and β leads to an identical projection as a camera with parameters $m/\cos \beta$, s_x, s_y, $c_x = 0$, $c_y = 0$, $v_x \cos \beta$, $v_y f$, $t_x \cos \beta$, $t_y f$, ϕ, and ψ for a suitably chosen factor f and suitably chosen angles ϕ and ψ. We will only examine the case of no lens distortions since for $c_y = 0$ the distortion is purely along the horizontal direction of the image and therefore can be rectified within each line independently.

To prove that the two camera parameter sets above result in the same projection, we will construct the affine transformation matrix $\mathbf{A}$ of points in the plane $z = 0$ in the calibration object coordinate system to points in the image plane. The affine transformation is given by multiplying the camera matrix in (38) by $\mathbf{O}^\top$ from the right, where $\mathbf{O}$ is given by (37). Hence, we have:

$$\mathbf{A} = \mathbf{P}\mathbf{O}^\top = \mathbf{K}\mathbf{O}\mathbf{H}\mathbf{O}^\top. \tag{56}$$

For the first set of camera parameters, this results in:

$$
\mathbf{A}_1 = \begin{pmatrix} a\cos\beta - \frac{av_x \sin\alpha \sin\beta}{v_y} & -\frac{av_x \cos\alpha}{v_y} & \frac{a(t_x v_y - t_y v_x)}{v_y} \\ \frac{\sin\alpha \sin\beta}{v_y} & \frac{\cos\alpha}{v_y} & \frac{t_y}{v_y} \\ 0 & 0 & 1 \end{pmatrix},
$$
(57)

where $a = m/s_x$. For the second set of camera parameters, we obtain:

$$
\mathbf{A}_2 = \begin{pmatrix} \frac{a\cos\psi}{\cos\beta} - \frac{av_x \sin\phi \sin\psi}{v_y f} & -\frac{av_x \cos\phi}{v_y f} & \frac{a(t_x v_y - t_y v_x)}{v_y} \\ \frac{\sin\phi \sin\psi}{v_y f} & \frac{\cos\phi}{v_y f} & \frac{t_y}{v_y} \\ 0 & 0 & 1 \end{pmatrix}.
$$
(58)

Hence, we can see that the translation part of $\mathbf{A}_1$ and $\mathbf{A}_2$ (their last columns) are identical. This leaves us with the following four equations for f, ϕ, and ψ:

$$
\frac{a\cos\psi}{\cos\beta} - \frac{av_x \sin\phi \sin\psi}{v_y f} = a\cos\beta - \frac{av_x \sin\alpha \sin\beta}{v_y}
$$
(59)

$$
\frac{av_x \cos\phi}{v_y f} = \frac{av_x \cos\alpha}{v_y}
$$
(60)

$$
\frac{\sin\phi \sin\psi}{v_y f} = \frac{\sin\alpha \sin\beta}{v_y}
$$
(61)

$$
\frac{\cos\phi}{v_y f} = \frac{\cos\alpha}{v_y}.
$$
(62)

We solve (62) for f and obtain:

$$
f = \frac{\cos\phi}{\cos\alpha}.
$$
(63)

By substituting (63) into (60), we obtain an equation that is fulfilled tautologically. Hence, we substitute (63) into (61) and simplify to obtain:

$$
\tan\phi \sin\psi = \tan\alpha \sin\beta,
$$
(64)

which we solve for $\tan\phi$:

$$
\tan\phi = \frac{\tan\alpha \sin\beta}{\sin\psi}.
$$
(65)

By substituting (63) into (59) and simplifying, we obtain:

$$
\frac{av_y \cos\psi}{\cos\beta} - av_x \cos\alpha \tan\phi \sin\psi
$$
$$
= av_y \cos\beta - av_x \sin\alpha \sin\beta.
$$
(66)

By substituting (65) into (66), we obtain:

$$
\frac{av_y \cos\psi}{\cos\beta} - av_x \sin\alpha \sin\beta
$$
$$
= av_y \cos\beta - av_x \sin\alpha \sin\beta.
$$
(67)

Hence, we have:

$$
\cos\psi = \cos^2\beta,
$$
(68)

where $\cos^2\beta$ is an abbreviation for $(\cos\beta)^2$. Thus:

$$
\psi = \arccos(\cos^2\beta).
$$
(69)

By substituting (68) into (65) and using the identity $\sin\theta = \sqrt{1 - \cos^2\theta}$, we obtain:

$$
\tan\phi = \frac{\tan\alpha \sin\beta}{\sqrt{1 - \cos^4\beta}}.
$$
(70)

Therefore, we have:

$$
\tan\phi = \arctan\left(\frac{\tan\alpha \sin\beta}{\sqrt{1 - \cos^4\beta}}\right).
$$
(71)

Finally, by substituting (71) into (63) and using the identity $\cos(\arctan\theta) = 1/\sqrt{\theta^2 + 1}$, we obtain:

$$
f = \left(\cos\alpha \sqrt{1 + \frac{\tan^2\alpha \sin^2\beta}{1 - \cos^4\beta}}\right)^{-1}.
$$
(72)

To extend the proof to the case, $(c_x, c_y)^\top \neq (0, 0)^\top$, we note (see Footnote 6) that $c'_{y,1} = c_{y,1}s_y/(mv_y)$ for the first camera and $c'_{y,2} = c_{y,2}s_y \cos\beta/(mv_y f)$ for the second camera, whence $c_{y,2} = c_{y,1}(f/\cos\beta)$. Furthermore, for the first camera, $c'_{x,1} = -c_{y,1}(s_y/s_x)(v_x/v_y) + c_{x,1}$, while for the second camera, $c'_{x,2} = -c_{y,2}(s_y/s_x)(v_x \cos\beta)/(v_y f) + c_{x,2}$. Substituting $c_{y,2} = c_{y,1}(f/\cos\beta)$ into the last equation shows that $c_{x,2} = c_{x,1}$. □

Remark 26 Table 1 displays an example for interior and exterior orientation parameters that result in identical projections for planar objects in the plane $z = 0$ that were obtained with solution in the proof of Theorem 2.

Remark 27 From the proof of Theorem 2, it might appear that there is only a twofold ambiguity. However, this is only caused by the fact that we have chosen the specific values of $m/\cos\beta$, $v_x \cos\beta$, and $t_x \cos\beta$ for the second camera to simplify the proof as much as possible. If other factors instead of $\cos\beta$ had been chosen, the values of f, ϕ, and ψ would change accordingly. Therefore, like in Propositions 2 and 3, the degeneracy is completely generic.

Table 1 An example for interior and exterior orientation parameters that result in identical projections for planar objects in the plane $z = 0$

	Camera 1	Camera 2
Interior orientation		
m (Scalar)	0.3	0.346410162
κ (m^{-2})	-2000	-2000
s_x (μm Pixel^{-1})	10	10
s_y (μm Pixel^{-1})	10	10
c_x (Pixel)	950	950
c_y (Pixel)	0	0
v_x (μm Pixel^{-1})	1.5	1.299038106
v_y (μm Pixel^{-1})	55	56.43276971
Exterior orientation		
t_x (m)	0.01	0.008660254
t_y (m)	0.1	0.102605036
α (°)	20	15.38347596
β (°)	30	41.40962211
γ (°)	50	50

Corollary 1 *The results above show that telecentric line-scan cameras cannot be calibrated from a single image of a planar calibration object. Consequently, multiple images of a planar calibration object with different exterior orientation must be used to calibrate the camera if all parameters are to be determined unambiguously.*

Remark 28 In machine vision applications, it sometimes is important to calibrate the camera from a single image. As the above discussion shows, this will lead to camera parameters that differ from their true values. However, if the residual error of the calibration is sufficiently small, a camera geometry that is consistent within the plane that is defined by the exterior orientation of the calibration object (and all planes parallel thereto) will be obtained. Therefore, an image or features extracted from an image can be rectified to this plane (cf. Remark 15). On the other hand, algorithms that solely rely on the interior orientation, e.g., the pure removal of radial distortions in Remark 14, are less useful because the ambiguities with respect to m and v_y imply that we cannot reliably undistort an image or features extracted from an image to have square pixels.

5 Calibration Experiments

5.1 Robustness of Principal Point Estimation

In our first experiment, we evaluate the importance of modeling c_y and lens distortions in general. We mounted an area-scan camera with a telecentric lens approximately 30 cm above a linear stage. The camera was oriented such that its viewing direction was vertically downwards onto the linear stage and its y axis was approximately parallel to the linear motion of the stage. An encoder that triggered the image acquisition was used to ensure a constant speed. We acquired an image at each trigger event and saved the obtained image array. We restricted the part that was read out from the sensor to the center image rows only: we selected the 90 sensor rows above and below the center image row, resulting in images of height 181. This setup enabled us to generate images of a virtual line-scan camera that consist of one of the 181 image rows. The line-scan image for one selected row was obtained by stacking the selected image row of all images in the array on top of each other. The frequency of the encoder was chosen such that the pixels in the generated line-scan images were approximately square.

For the tests, we chose two different hardware setups. In the first setup, we used an IDS GV-5280CP-C-HQ color camera (2/3 inch sensor size, CMOS, 3.45 μm pixel pitch, 2448 × 2048) with a telecentric Vicotar T201/0.19 lens (nominal magnification: 0.19). We set up the camera, which uses a color filter array to capture color information, to directly return gray-scale images. The generated line-scan images are of size 2448 × 3330.

In the second setup, we used an IDS UI-3080CP-M-GL monochrome camera (2/3 inch sensor size, CMOS, 3.45 μm pixel pitch, 2456 × 2054) with a telecentric V.S. Technologies L-VS-TC017 lens (nominal magnification: 0.17). Because the lens was designed for a maximum sensor size of 1/2 inch, we cropped the images to 70% of their width. The generated line-scan images are of size 1719 × 2954.

With both setups, we acquired 16 image arrays of a 4×3 cm^2 planar calibration object with a hexagonal layout of circular marks in different poses, as described in Steger et al. (2018, Chapter 3.9.4.1). For each of the 181 image rows, we generated 16 virtual line-scan images, one from each of the 16 image arrays. Consequently, for each of the 181 image rows, we obtained 16 calibration images that we used for calibration by minimizing (35). Lens distortions were taken into account by applying the division model. The variation of the resulting camera parameters depending on the selected sensor row is shown in Figs. 2 and 3. In addition, the variation of the root mean square (RMS) calibration error, i.e., $\sqrt{\varepsilon^2 / \sum_{l=1}^{n_\mathrm{o}} \sum_{j=1}^{n_\mathrm{m}} v_{jl}}$, is plotted.

From the plots, it can be seen that all parameters except for c_y do not change substantially when selecting different sensor rows (note the scaling of the vertical axes). This is the expected behavior because c_y measures the misalignment of the sensor line with respect to the optical axis in the y direction (see Sect. 3.2). Therefore, there is a linear relationship between the selected sensor row and c_y, which is evident from the plots. While the expected change of c_y over the sensor rows is 181, it is 182 for the color camera but only
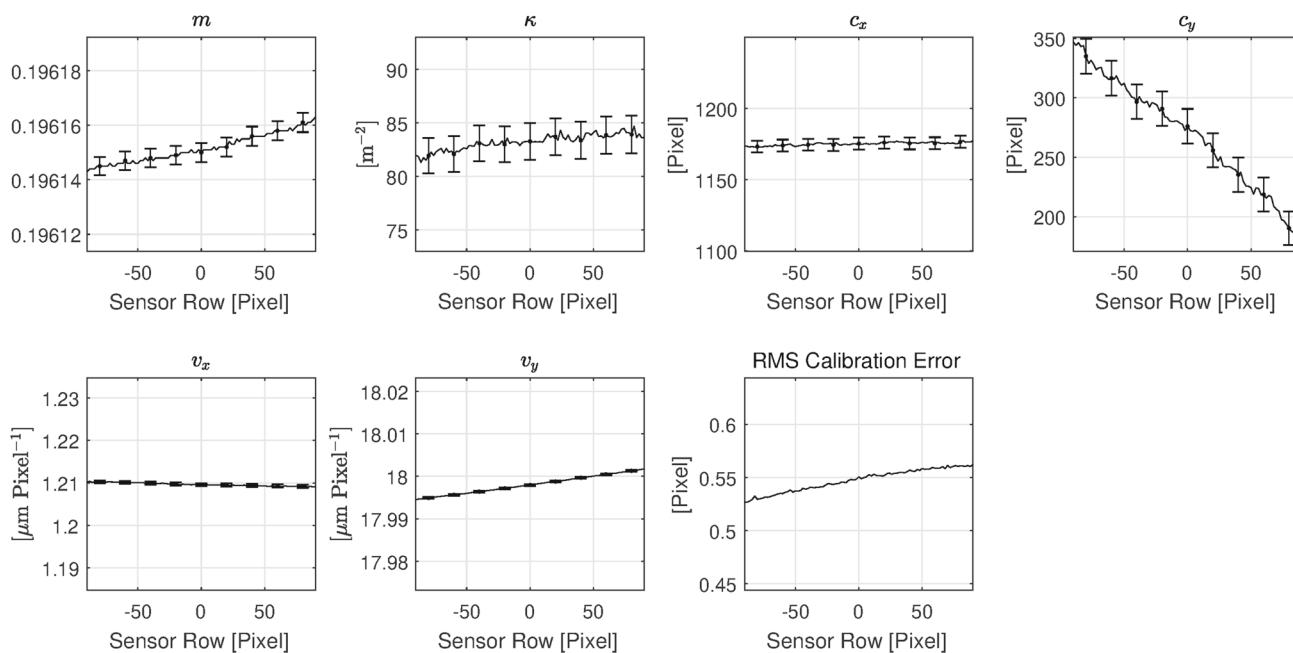
Fig. 2 Variation of the calibrated interior orientation of the *IDS GV-5280CP-C-HQ color camera* with the *Vicotar T201/0.19 lens* depending on the selected sensor row when applying the *division* distortion model. For each twentieth sensor row, the standard deviation is indicated by error bars. Additionally, the root mean square (RMS) calibration error is plotted
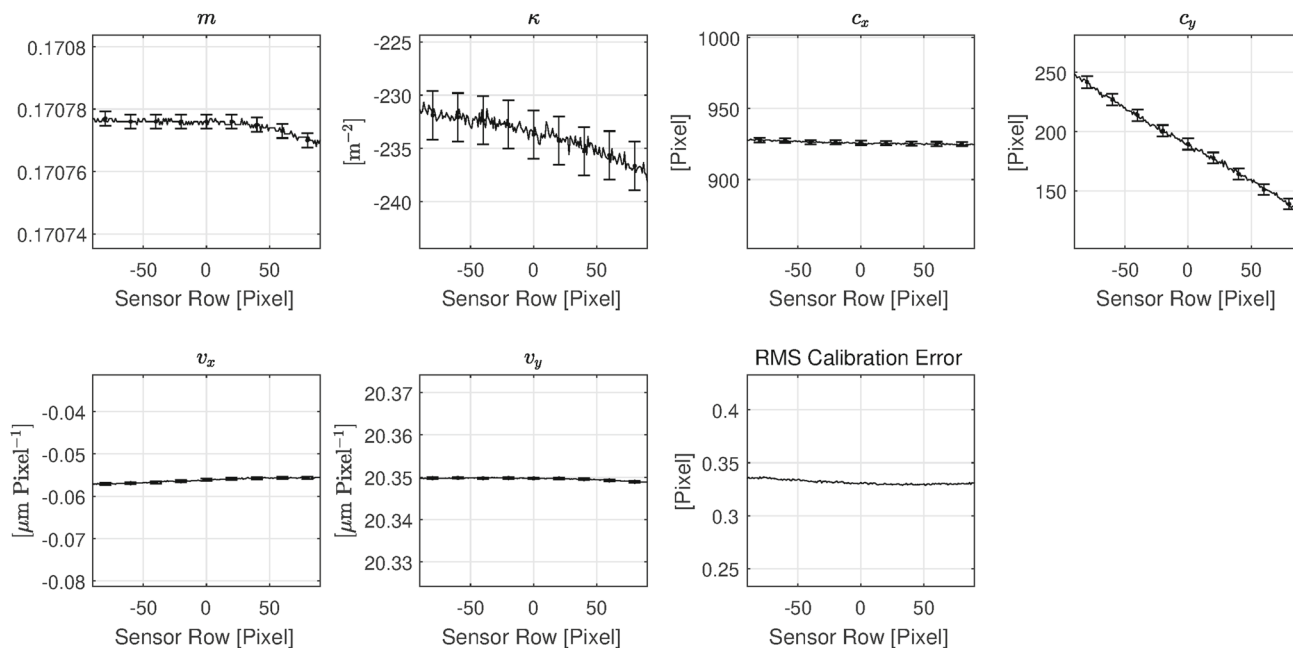

Fig. 3 Variation of the calibrated interior orientation of the *IDS UI-3080CP-M-GL monochrome camera* with the *V.S. Technologies L-VS-TC017 lens* depending on the selected sensor row when applying the *division* distortion model. For each twentieth sensor row, the standard deviation is indicated by error bars. Additionally, the root mean square (RMS) calibration error is plotted

131 for the monochrome camera. Although we are not perfectly sure, we assume that the reason for c_y changing only by 131 pixels is that the circular actuator of the linear stage is not perfectly centered, resulting in a non-uniform movement (see Sect. 5.2 for more details). Because of the different magnifications of the lenses, for the color camera, there seems to be an integer number of periods of the movement bias in the images, whereas for the monochrome camera there seems to be a non-integer number of periods. Since c_y is primarily determined by the bending of the image lines, an asymmetry in the movement bias in the images may prevent the reliable extraction of c_y for the monochrome camera.

For increasing sensor rows, the values for m and κ increase slightly for the color camera and decrease slightly for the monochrome camera. This is because a larger magnification can be at least partially compensated by a larger value of κ, which causes a non-uniform scaling in the image row, and vice versa.

Also note that the RMS error of the color camera is significantly larger than that of the monochrome camera, which probably is caused by image artifacts that are caused by the color filter array.

Despite the relatively small lens distortions of at most 1.8 pixels in the images, c_x and c_y are estimated consistently over different sensor rows. Nevertheless, the standard deviations of the estimated principal point are significantly larger for the color camera than for the monochrome camera because the magnitude of κ is larger by a factor of almost 3 for the monochrome camera. Consequently, the principal point of the monochrome camera is better defined (see Remark 20).

We alternatively applied the polynomial distortion model to the calibration of the monochrome camera. As suggested in Remark 21, we set P_1 and P_2 to 0 and excluded them from the calibration. The results are shown in Fig. 4. The obtained values for the parameters are very similar to those obtained with the division model. The RMS error decreases only marginally from 0.3318 (division) to 0.3316 (polynomial), on average. As for the division model, a correlation between m and the radial lens distortion parameters (K_i) is observable. The experiment shows that the division model obviously represents the lens distortions sufficiently well and that both models return consistent results.

These experiments show that the proposed camera model is able to estimate the principal point of the camera accurately, and hence is able to handle lens distortions effectively, even for lenses with only small lens distortions.

5.2 Calibration of Line-Scan Cameras

In the second experiment, we calibrated two monochrome Basler raL2048-48gm line-scan cameras (14.3 mm sensor size, CMOS, 7.0 μm pixel pitch, 2048 × 1) with Opto Engineering telecentric lenses. On the first camera, we mounted

a TC2MHR048-F (nominal magnification: 0.268, working distance: 133 mm) and on the second camera, we mounted a TC2MHR058-F (nominal magnification: 0.228, working distance: 158 mm). We used the same setup as described in Sect. 5.1 and acquired 16 images of a $8 \times 6 \, cm^2$ planar calibration object in different poses. Each calibration was performed with the division and polynomial distortion models. The results are shown in Tables 2 and 3.

The low RMS errors indicate that the cameras and lenses can be represented very accurately by our proposed model. Both high-quality lenses have very small lens distortions. This also causes the principal points to be poorly defined, resulting in significantly different values for c_y for the division and polynomial distortion models in Table 2. When setting $c_x = 1024$, $c_y = 0$, all distortion parameters to 0, and excluding these parameters from the calibration, the RMS errors only marginally increase to 0.3545 (division model) and 0.3541 (polynomial model) for the TC2MHR048-F lens and to 0.2928 (division model) and 0.2927 (polynomial model) for the TC2MHR058-F lens. Nevertheless, the maximum absolute distortion in the images is approximately 1.2 pixels (TC2MHR048-F) and 1.4 pixels (TC2MHR058-F), which would decrease the accuracy of measurements when being ignored.

The small distortions raise the question whether any of the parameters in the model are redundant for these lenses, i.e., whether overfitting has occurred. Obviously, m, v_x, and v_y are significant geometrically and therefore cannot be omitted. Hence, the question is whether (c_x, c_y) or the distortion parameters κ or K_1, K_2, and K_3 are significant. We use the significance test proposed by Grün (1978) to test whether any useful combinations of these parameters are significant. Each test hypothesis was obtained by setting the respective distortion parameters to 0 and the principal point to the center of the sensor line. In Table 4, we display the results of this test for the TC2MHR058-F lens. The results for the TC2MHR048-F lens are omitted because they are similar. As can be seen from Table 4, all distortion-related parameters are highly significant. Therefore, no overfitting occurs, even for these small distortions.

Figure 5 shows one of the calibration images that were used to calibrate the camera with the TC2MHR048-F lens. In addition, the residuals (scaled by a factor of 130) are visualized for each circular calibration mark. The residuals are the differences between the extracted centers of the calibration marks in the image and the projections of the corresponding points on the calibration object into the image. The projection is performed by using the calibrated camera parameters of the interior and exterior orientation while applying polynomial distortion model. The mean and maximum length of the residuals was 0.284 pixels and 0.538 pixels, respectively. This corresponds to 7.21 μm and 13.67 μm in the world. It can be seen that the predominant part of the residuals is a sys-
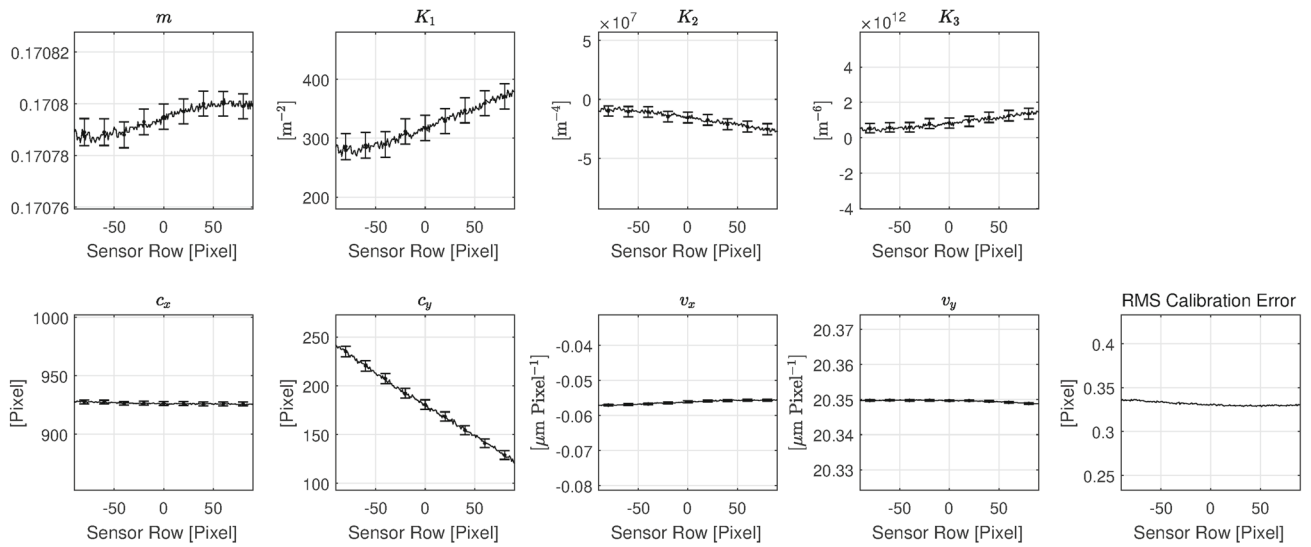
Fig. 4 Variation of the calibrated interior orientation of the *IDS UI-3080CP-M-GL monochrome camera* with the *V.S. Technologies L-VS-TC017 lens* depending on the selected sensor row when applying the *polynomial* distortion model. For each twentieth sensor row, the standard deviation is indicated by error bars. Additionally, the root mean square (RMS) calibration error is plotted

Table 2 Calibration results for a Basler raL2048-48gm line-scan camera with an Opto Engineering TC2MHR048-F telecentric lens for the division and polynomial lens distortion models

RMS eror (Pixel)	Division 0.3211	Polynomial 0.3210
m (Scalar)	0.26717	0.26716
s_x (μm Pixel^{-1})*	7.0	7.0
s_y (μm Pixel^{-1})*	7.0	7.0
c_x (Pixel)	854.725	855.278
c_y (Pixel)	− 16.030	− 157.289
v_x (μm Pixel^{-1})	− 0.0963	− 0.0963
v_y (μm Pixel^{-1})	25.2901	25.2906
κ (m^{-2})	14.0165	–
K_1 (m^{-2})	–	− 12.4707
K_2 (m^{-4})	–	− 110858.0
K_3 (m^{-6})	–	1409814283.2
P_1 (m^{-1})*	–	0.0
P_2 (m^{-1})*	–	0.0

Parameters indicated by * are excluded from the calibration

Table 3 Calibration results for a Basler raL2048-48gm line-scan camera with an Opto Engineering TC2MHR058-F telecentric lens for the division and polynomial lens distortion models

RMS eror (Pixel)	Division 0.2522	Polynomial 0.2521
m (Scalar)	0.22802	0.22803
s_x (μm Pixel^{-1})*	7.0	7.0
s_y (μm Pixel^{-1})*	7.0	7.0
c_x (Pixel)	1231.955	1237.057
c_y (Pixel)	− 130.684	− 131.088
v_x (μm Pixel^{-1})	− 1.0491	− 1.0491
v_y (μm Pixel^{-1})	30.0455	30.0455
κ (m^{-2})	− 13.7166	–
K_1 (m^{-2})	–	18.3485
K_2 (m^{-4})	–	− 196535.0
K_3 (m^{-6})	–	1983345655.0
P_1 (m^{-1})*	–	0.0
P_2 (m^{-1})*	–	0.0

Parameters indicated by * are excluded from the calibration

tematic periodic error in the direction of the movement, i.e., in vertical direction in the image. It should be noted that the used encoder reacts to the angle position of the electric motor of the linear stage. Therefore, we assume that the major part of the residuals is caused by the circular actuator that is not perfectly centered. Another indication of this assumption is the fact that the periodicity of the error corresponds to one full revolution of the actuator. In this case, the calibration error could be further reduced by using a higher-quality electric

actuator that better realizes a constant speed. In comparison, the residuals in horizontal direction are very small, which again shows that the proposed camera model represents the true projection very well.

5.3 Example Application: Image Rectification

To be able to precisely measure distances and angles in the image, the obtained line-scan images must be rectified to

Table 4 Results of the significance test for parameter combinations for a Basler raL2048-48gm line-scan camera with an Opto Engineering TC2MHR058-F telecentric lens for the division and polynomial lens distortion models

Parameters	T	m	n	α
Division model				
κ	3194.29	1	20,662	$< 1 \times 10^{-16}$
$\kappa, (c_x, c_y)$	6780.65	3	20,662	$< 1 \times 10^{-16}$
Polynomial model				
K_3	18.87	1	20,660	1.4028×10^{-5}
K_2, K_3	14.12	2	20,660	7.4479×10^{-7}
K_1, K_2, K_3	758.99	3	20,660	$< 1 \times 10^{-16}$
$K_3, (c_x, c_y)$	559.51	3	20,660	$< 1 \times 10^{-16}$
$K_2, K_3, (c_x, c_y)$	492.51	4	20,660	$< 1 \times 10^{-16}$
$K_1, K_2, K_3, (c_x, c_y)$	4801.52	5	20,660	$< 1 \times 10^{-16}$

The values of the significance level α were computed using the MAT-LAB function $1 - \mathrm{fcdf}(T, m, n)$. T is the test statistic proposed by Grün (1978), m is the number of parameters tested for significance, and n is the redundancy of the system ($2 \times$ number of calibration marks extracted $-$ number of camera parameters optimized $- 5n_o$). The entry $< 1 \times 10^{-16}$ indicates that $\mathrm{fcdf}(T, m, n) = 1$ to double precision accuracy. Note that all distortion-related parameters are highly significant

eliminate lens distortions and skew to ensure square pixels. While this is generally impossible for entocentric lenses (see Remark 12), we can perform such a rectification for telecentric lenses (see Remark 14).

For the example shown in Fig. 6, we acquired a line-scan image of a graph paper with the setup described in Sect. 5.2 with the TC2MHR058-F lens. The acquired image is shown in Fig. 6a. Because the motion direction was not perfectly perpendicular to the sensor line, i.e., has a significant non-zero motion component in the direction of the sensor line ($v_x = -1.0491\,\mu\mathrm{m}\,\mathrm{Pixel}^{-1}$, see Table 3), the squares of the graph paper appear skewed. Furthermore, because the speed of the motion was not perfectly adjusted, rectangular instead of square pixels are obtained, causing a non-uniform scaling of the squares on the graph paper in the image.

By setting all distortion coefficients to 0, v_x to 0, s_x to $\min(s_x, m v_y)$, and v_y to s_x/m (see Remark 14), we can generate an image mapping that rectifies the images acquired with the setup. After the rectification, the images have no lens distortions and square pixels. Fig. 6b shows the resulting rectified image. The squares on the graph paper are squares in the rectified image. Hence, in the rectified image, it is possible to measure angles and distances and areas in world units correctly.

For a quantitative comparison, we extracted subpixel-precise lines (Steger 1998a, b, 2013) in the original and in the rectified image, fitted straight lines, and computed their intersection angles. The mean over all angles was 87.971° in the original image and 90.004° in the rectified image. Further-
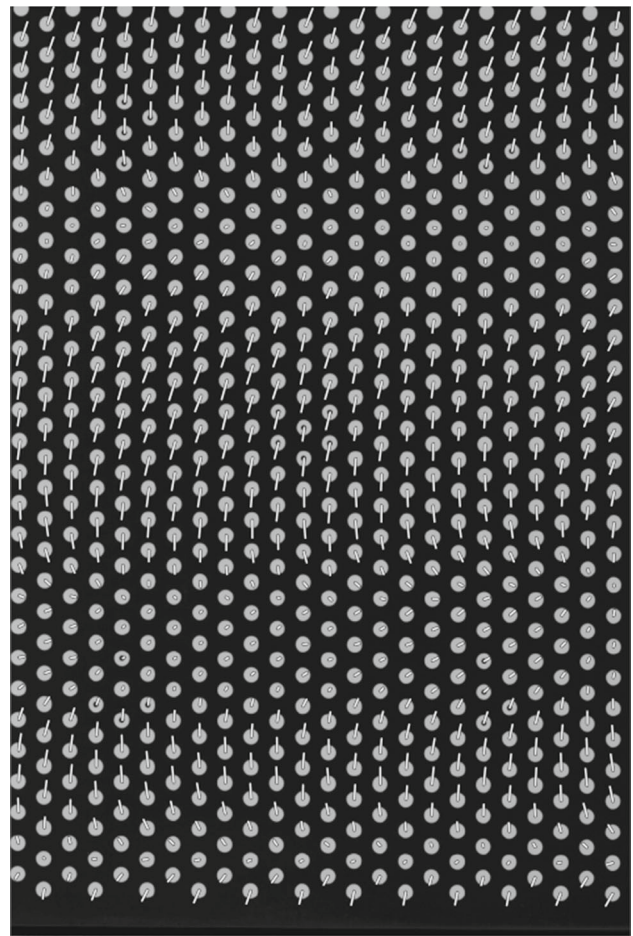

Fig. 5 One of the images of the planar calibration object that were used to calibrate the raL2048-48gm line-scan camera with the TC2MHR048-F lens. Residuals are overlaid for each circular calibration mark as white lines. The residuals were scaled by a factor of 130 for better visibility. The predominant part of the residuals is a systematic periodic error in the direction of the movement, i.e., in vertical direction in the image
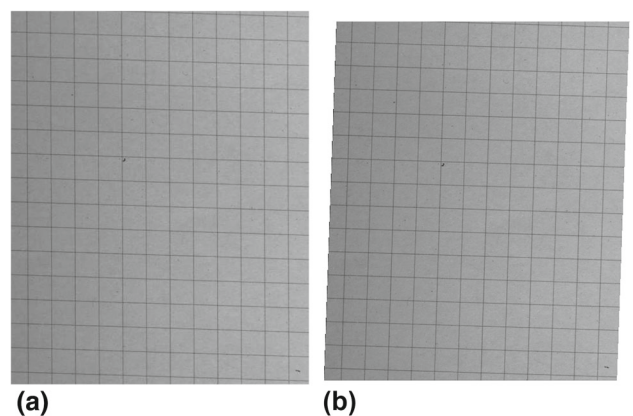

(a)　　　　　　　　　　**(b)**

Fig. 6 Example application: image rectification by eliminating lens distortions and ensuring square pixels. **a** Original line-scan image. The squares on the graph paper are skewed and have a non-uniform scaling in the image. **b** Rectified line-scan image. The squares on the graph paper are squares in the image

more, we computed the area of each square on the graph paper in both images and transformed it to metric units by multiplying it by $s_x s_y / m^2$. Here, s_x and s_y are the pixel pitches on the sensor and m is the lens magnification obtained from calibration. The mean area was $25.50\,mm^2$ in the original image and $24.98\,mm^2$ in the rectified image, while the actual size of the squares on the paper was $25\,mm^2$.

6 Conclusions

We have proposed a camera model for line-scan cameras with telecentric lenses. The model assumes a linear motion of the camera with constant velocity. It can model general lens distortions by allowing the sensor line to lie anywhere with respect to the optical axis. The model is parameterized by camera parameters that have a physical meaning and is therefore easy to understand by machine vision users. We have described an algorithm to calibrate the camera model using a planar calibration object. Furthermore, we have investigated the degeneracies of the model and have discussed how they can be handled in practice. In addition, we have shown that every affine camera can be interpreted as a telecentric line-scan camera and vice versa, provided the lens does not exhibit any lens distortions. Experiments with real setups have been used to establish the validity of the model. In particular, we have shown that even for lenses with very small lens distortions, the distortions statistically are highly significant and therefore cannot be omitted in real-world applications.

One direction for future research is to derive an explicit stereo or multi-view camera model for telecentric line-scan cameras.

Author contributions Carsten Steger designed, analyzed, and implemented the camera model and wrote Sects. 2–4 and 6. Markus Ulrich designed the experiments, acquired the images, performed the experiments, and wrote Sect. 5. Both authors wrote Sect. 1 and critically reviewed the paper.

References

Beyerer, J., Puente León, F., & Frese, C. (2016). *Machine vision: Automated visual inspection: Theory, practice and applications*. Berlin: Springer.

Blahusch, G., Eckstein, W., Steger, C., & Lanser, S. (1999). Algorithms and evaluation of a high precision tool measurement system. In *5th International conference on quality control by artificial vision* (pp. 31–36).

Brown, D. C. (1966). Decentering distortion of lenses. *Photogrammetric Engineering*, *32*(3), 444–462.

Brown, D. C. (1971). Close-range camera calibration. *Photogrammetric Engineering*, *37*(8), 855–866.

Chen, T., Shibasaki, R., & Shunji, M. (2003). Development and calibration of the airborne three-line scanner (TLS) imaging system. *Photogrammetric Engineering and Remote Sensing*, *69*(1), 71–78.

de Oliveira, O. (2013). The implicit and the inverse function theorems: Easy proofs. *Real Analysis Exchange*, *39*(1), 207–218.

Donné, S., Luong, H., Dhondt, S., Wuyts, N., Inzé, D., Goossens, B., & Philips, W. (2017). Robust plane-based calibration for linear cameras. In *International conference on image processing* (pp. 36–40).

Draréni, J., Roy, S., & Sturm, P. (2011). Plane-based calibration for linear cameras. *International Journal of Computer Vision*, *91*(2), 146–156.

Fitzgibbon, A., Pilu, M., & Fisher, R. B. (1999). Direct least square fitting of ellipses. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *21*(5), 476–480.

Fitzgibbon, A. W. (2001). Simultaneous linear estimation of multiple view geometry and lens distortion. In *IEEE conference on computer vision and pattern recognition* (Vol. I, pp. 125–132).

Grün, A. (1978). Accuracy, reliability and statistics in close-range photogrammetry. In *Inter-congress Symposium, Commission V, International Society for Photogrammetry*, Stockholm, unbound paper No. 9.

Gupta, R., & Hartley, R. I. (1997). Linear pushbroom cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *19*(9), 963–975.

Haala, N., Stallmann, D., & Cramer, M. (1998). Calibration of directly measured position and attitude by aerotriangulation of three-line airborne imagery. In T. Schenk & A. Habib (Eds.), *ISPRS Commission III symposium on object recognition and scene classification from multispectral and multisensor pixels* (pp. 23–30).

Hartley, R., & Zisserman, A. (2003). *Multiple view geometry in computer vision* (2nd ed.). Cambridge: Cambridge University Press.

Horaud, R., Mohr, R., & Lorecki, B. (1993). On single-scanline camera calibration. *IEEE Transactions on Robotics and Automation*, *9*(1), 71–75.

Hui, B., Wen, G., Zhang, P., & Li, D. (2013). A novel line scan camera calibration technique with an auxiliary frame camera. *IEEE Transactions on Instrumentation and Measurement*, *62*(9), 2567–2575.

Hui, B., Wen, G., Zhao, Z., & Li, D. (2012a). Line-scan camera calibration in close-range photogrammetry. *Optical Engineering*, *51*(5), 053602-1–053602-12.

Hui, B., Zhong, J., Wen, G., & Li, D. (2012b). Determination of line scan camera parameters via the direct linear transformation. *Optical Engineering*, *51*(11), 113201-1–113201-6.

Lanser, S. (1997). Modellbasierte Lokalisation gestützt auf monokulare Videobilder. Dissertation, Forschungs- und Lehreinheit Informatik IX, Technische Universität München.

Lanser, S., Zierl, C., & Beuthauser, R. (1995). Multibildkalibrierung einer CCD-Kamera. In G. Sagerer, S. Posch, & F. Kummert (Eds.), *Mustererkennung* (pp. 481–491). Informatik aktuell: Springer.

Lee, C., Theiss, H. J., Bethel, J. S., & Mikhail, E. M. (2000). Rigorous mathematical modeling of airborne pushbroom imaging systems.

Photogrammetric Engineering and Remote Sensing, *66*(4), 385–392.

Lenhardt, K. (2017). Optical systems in machine vision. In A. Hornberg (Ed.), *Handbook of machine and computer vision* (2nd ed., pp. 179–290). Weinheim: Wiley.

Lenz, R. (1987). Linsenfehlerkorrigierte Eichung von Halbleiterkameras mit Standardobjektiven für hochgenaue 3D-Messungen in Echtzeit. In E. Paulus (Ed.), *Mustererkennung* (Vol. 149, pp. 212–216). Berlin: Springer.

Lenz, R. (1988). *Viedeometrie mit CCD-Sensoren und ihre Anwendung in der Robotik*. Habilitationsschrift, Lehrstuhl für Nachrichtentechnik der Technischen Universität München.

Lenz, R., & Fritsch, D. (1990). Accuracy of videometry with CCD sensors. *ISPRS Journal of Photogrammetry and Remote Sensing*, *45*(2), 90–110.

Lilienblum, E., Al-Hamadi, A., & Michaelis, B. (2013). A coded 3D calibration method for line-scan cameras. In J. Weickert, M. Hein, & B. Schiele (Eds.), *Pattern recognition* (Vol. 8142, pp. 81–90). Lecture notes in computer science Berlin: Springer.

Luna, C. A., Mazo, M., Lázaro, J. L., & Vázquez, J. F. (2010). Calibration of line-scan cameras. *IEEE Transactions on Instrumentation and Measurement*, *59*(8), 2185–2190.

Luster, S. D., & Batchelor, B. G. (2012). Telecentric, fresnel and micro lenses. In B. G. Batchelor (Ed.), *Machine vision handbook* (pp. 259–281). London: Springer.

Mallon, J., & Whelan, P. F. (2007). Which pattern? Biasing aspects of planar calibration patterns and detection methods. *Pattern Recognition Letters*, *28*(9), 921–930.

MVTec Software GmbH. (2005a). HALCON application note: Machine vision in world coordinates. HALCON Version 7.1.

MVTec Software GmbH. (2005b). HALCON/HDevelop reference manual. HALCON Version 7.1.

Niu, M., Song, K., Wen, X., Zhang, D., & Yan, Y. (2018). The line scan camera calibration based on space rings group. *IEEE Access*, *6*, 23711–23721.

Poli, D. (2007). A rigorous model for spaceborne linear array sensors. *Photogrammetric Engineering and Remote Sensing*, *73*(2), 187–196.

Ramalingam, S., & Sturm, P. (2017). A unifying model for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *39*(7), 1309–1319.

Ramalingam, S., Sturm, P., & Lodha, S. K. (2006). Theory and calibration for axial cameras. In P. J. Narayanan, S. K. Nayar, & H. Y. Shum (Eds.), *7th Asian conference on computer vision* (Vol. 3851, pp. 704–713). Lecture notes in computer science Berlin: Springer.

Shapiro, L. S., Zisserman, A., & Brady, M. (1995). 3d motion recovery via affine epipolar geometry. *International Journal of Computer Vision*, *16*(2), 147–182.

Song, K., Hou, B., Niu, M., Wen, X., & Yan, Y. (2018). Flexible line-scan camera calibration method using a coded eight trigrams pattern. *Optics and Lasers in Engineering*, *110*, 296–307.

Steger, C. (1998a). An unbiased detector of curvilinear structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(2), 113–125.

Steger, C. (1998b). Unbiased extraction of curvilinear structures from 2D and 3D images. Dissertation, Fakultät für Informatik, Technische Universität München.

Steger, C. (2000). Subpixel-precise extraction of lines and edges. *International Archives of Photogrammetry and Remote Sensing*, *33*(B3), 141–156.

Steger, C. (2012). Estimating the fundamental matrix under pure translation and radial distortion. *ISPRS Journal of Photogrammetry and Remote Sensing*, *74*, 202–217.

Steger, C. (2013). Unbiased extraction of lines with parabolic and Gaussian profiles. *Computer Vision and Image Understanding*, *117*(2), 97–112.

Steger, C. (2017). A comprehensive and versatile camera model for cameras with tilt lenses. *International Journal of Computer Vision*, *123*(2), 121–159.

Steger, C. (2018). Algorithms for the orthographic-*n*-point problem. *Journal of Mathematical Imaging and Vision*, *60*(2), 246–266.

Steger, C., Ulrich, M., & Wiedemann, C. (2008). *Machine vision algorithms and applications*. Weinheim: Wiley.

Steger, C., Ulrich, M., & Wiedemann, C. (2018). *Machine vision algorithms and applications* (2nd ed.). Weinheim: Wiley.

Sturm, P., Ramalingam, S., Tardif, J. P., Gasparini, S., & Barreto, J. (2010). Camera models and fundamental concepts used in geometric computer vision. *Foundations and Trends in Computer Graphics and Vision*, *6*(1–2), 1–183.

Sun, B., Zhu, J., Yang, L., Guo, Y., & Lin, J. (2017). Stereo line-scan sensor calibration for 3D shape measurement. *Applied Optics*, *56*(28), 7905–7914.

Sun, B., Zhu, J., Yang, L., Yang, S., & Guo, Y. (2016a). Sensor for in-motion continuous 3D shape measurement based on dual line-scan cameras. *Sensors*, *16*(11), 1949-1–1949-15.

Sun, B., Zhu, J., Yang, L., Yang, S., & Niu, Z. (2016b). Calibration of line-scan cameras for precision measurement. *Applied Optics*, *55*(25), 6836–6843.

Ulrich, M., & Steger, C. (2019). A camera model for cameras with hypercentric lenses and some example applications. *Machine Vision and Applications*, *30*(6), 1013–1028.

Yao, M., Zhao, Z., & Xu, B. (2014). Geometric calibration of line-scan camera using a planar pattern. *Journal of Electronic Imaging*, *23*(1), 013028–1–013028–8.

Ye, J., & Yu, J. (2014). Ray geometry in non-pinhole cameras: a survey. *The Visual Computer*, *30*(1), 93–112.

Zhang, J., Zhang, Z., Shen, F., Zhang, F., & Su, H. (2018). A new calibration method of line scan camera for high-precision two-dimensional measurement. In *14th International conference on automation science and engineering* (pp 678–683).