Optical Engineering

OpticalEngineering.SPIEDigitalLibrary.org

Cross-ratio–based line scan camera calibration using a planar pattern

Dongdong Li
Gongjian Wen
Shaohua Qiu

SPIE.

Cross-ratio–based line scan camera calibration using a planar pattern

Dongdong Li,* Gongjian Wen, and Shaohua Qiu
National University of Defense Technology, School of Electronic Science and Engineering, Yanwachi 137, Kaifu District, Changsha, Hunan 410073, China

Abstract. A flexible new technique is proposed to calibrate the geometric model of line scan cameras. In this technique, the line scan camera is rigidly coupled to a calibrated frame camera to establish a pair of stereo cameras. The linear displacements and rotation angles between the two cameras are fixed but unknown. This technique only requires the pair of stereo cameras to observe a specially designed planar pattern shown at a few (at least two) different orientations. At each orientation, a stereo pair is obtained including a linear array image and a frame image. Radial distortion of the line scan camera is modeled. The calibration scheme includes two stages. First, point correspondences are established from the pattern geometry and the projective invariance of cross-ratio. Second, with a two-step calibration procedure, the intrinsic parameters of the line scan camera are recovered from several stereo pairs together with the rigid transform parameters between the pair of stereo cameras. Both computer simulation and real data experiments are conducted to test the precision and robustness of the calibration algorithm, and very good calibration results have been obtained. Compared with classical techniques which use three-dimensional calibration objects or controllable moving platforms, our technique is affordable and flexible in close-range photogrammetric applications. [DOI: 10.1117/1.OE.55.1.014104]

Keywords: line scan camera; frame camera; cross-ratio; planar calibration pattern.

Paper 151204 received Sep. 1, 2015; accepted for publication Dec. 11, 2015; published online Jan. 20, 2016.

1 Introduction

A line scan camera or pushbroom camera is a nonconventional one-dimensional (1-D) imaging device widely used in areas such as industrial inspection[1] and satellite imagery.[2] Compared with frame cameras, line scan cameras provide higher sample rate and space resolution. They are preferred over frame cameras in many close-range photogrammetric applications such as vehicle-borne three-dimensional (3-D) scene reconstruction[3] and attitude measurement of high-speed targets.[4] In such metrology applications, line scan camera calibration is a necessary step in order to extract accurate metric information from linear array images.

In the last few decades, many mature frame camera calibration methods have been introduced, such as calibration methods of Tsai[5] and Zhang.[6] These methods recover intrinsic parameters from correspondences between 3-D feature points on a calibration object and their corresponding projections on the image. However, in the case of line scan cameras, only those points are imaged that lie on the view plane defined by the optical center and the linear array sensor, typically a CCD array. Therefore, without professional instruments such as a programmable stage, none of these frame camera calibration methods tackle line scan camera calibration due to the specific imaging model of line scan cameras.

In the scope of this paper, we aim at a calibration technique determining the line scan camera imaging model from a few linear array images of a planar calibration pattern. The inherent difficulty of this calibration technique is to establish correspondences between image points on the linear array images and their corresponding space points on the calibration pattern. Compared with frame camera calibration, only several papers focus on the subject of line scan camera calibration. We can classify the existent techniques roughly into two categories: scan-based calibration and line-based calibration.

1. Scan-based calibration: in scan-based calibration techniques,[7–9] the line scan camera is mounted on a moving platform (generally a programmable stage). The line scan camera moves along the moving platform at a known velocity and scans a calibration object over the image acquisition duration. A sequence of linear array images is obtained and rearranged into a two-dimensional synthetical frame image. In this way, corresponding projections of 3-D feature points on the calibration object are easily recognized and located in the synthetical frame image. Here, the considered camera imaging model is the translational pushbroom camera imaging model.[7] Generally, scan-based calibration techniques are accurate and robust. However, the moving platform which is expensive and requires an elaborate setup makes scanning-based calibration techniques not flexible and affordable for ordinary camera users in close-range photogrammetric applications.
2. Line-based calibration: line-based techniques do not use any moving platforms. Instead, they calibrate the line scan camera imaging model from a single linear array image of a 3-D calibration pattern.[10–12] Two groups of parallel feature lines are inscribed on the pattern. The equations of these feature lines are known in the pattern coordinate system with good precision.

*Address all correspondence to: Dongdong Li, Email: moqimubai@sina.cn

When the line scan camera "looks" at the pattern, each intersection point of a feature line and the view plane is projected onto the linear array sensor as an image point. Coordinates of intersection points can be calculated from image points based on the pattern geometry. After that, correspondences between space points and image points are established to recover the camera imaging model. Line-based calibration techniques avoid the use of moving platforms and are thus more flexible compared with scan-based calibration techniques. However, two factors may affect accuracy and robustness of line-based calibration techniques. First, the geometry of the 3-D calibration object has to be known with very good precision which requires an elaborate setup. Second, a single linear array image provided too limited projective information because only a limited number of feature lines are inscribed on the pattern.

Our current research is focused on outdoor survey work. Therefore, flexibility, robustness, and low cost are primarily considered. Unfortunately, all the calibration techniques mentioned in the above two categories are designed for laboratory environments and can hardly serve fieldwork due to extra requirements such as the use of moving platforms and 3-D calibration objects which require an elaborate setup and are inconvenient to carry.

In this paper, we present a flexible plane-based line scan camera calibration technique. This technique gets inspiration from Zhang's plane-based frame camera calibration method[6] and combines the advantages of both scan-based and line-based calibration techniques. Unlike scan-based calibration techniques that use a moving platform to guarantee the calibration pattern a constant velocity and orientation, our proposal couples the line scan camera to a calibrated frame camera that determines the position and pose of the planar calibration pattern. Unlike line-based calibration techniques, which calibrate the line scan camera imaging model from a single linear array image, our proposal requires the line scan camera and frame camera to capture several stereo pairs of the planar calibration pattern shown at a few (at least two) different orientations. At each orientation, a stereo pair is obtained including a linear array image and a frame image. The linear displacements and rotation angles between the line scan camera and frame camera are fixed, but unknown over the image acquisition duration. Both computer simulation and real data experiments show that the proposed calibration algorithm is accurate and robust. Compared with existent techniques which use 3-D calibration object or controllable moving platforms, our technique is easy to use and flexible in close-range photogrammetric applications.

The remainder of this paper is organized as follows. Section 2 derives the formulations of the line scan camera imaging model. In Sec. 3, a general scheme of our proposed calibration technique is explained. The calibration scheme mainly includes three steps: (1) establishing point correspondences based on the projective invariance of cross-ratio, (2) calculating a closed-form solution through direct linear transformation (DLT), and (3) nonlinear optimization. Section 4 details the experiments and results. Both computer simulation and real data are used to test the accuracy and robustness of this calibration technique. Finally, in Sec. 5, conclusions of this paper are exposed.

2 Line Scan Camera Imaging Model

In order to describe the line scan camera imaging model, we establish the line scan camera coordinate system $O_l - X_l Y_l Z_l$, the image coordinate system $o - uv$ and the world coordinate system $O - XYZ$ as shown in Fig. 1. $O_l - X_l Y_l Z_l$ is defined by the linear array sensor and the perspective center O_l. The Y_l-axis lies in the view plane parallel with the linear array sensor. The Z_l-axis lies in the view plane perpendicular to the Y_l-axis and directed so that the space points have positive Z_l coordinate. The X_l-axis is perpendicular to both the Y_l-axis and Z_l-axis, so that $O_l - X_l Y_l Z_l$ is a right-handed coordinate system. The image coordinate system $o - uv$ is defined on the linear array sensor. The origin o is foot of the perpendicular from perspective center O_l to the linear array sensor. The v-axis lies along the linear array sensor. The u-axis is parallel with the X_l-axis. Without loss of generality, we assume the view plane is on $X_l = 0$ of $O_l - X_l Y_l Z_l$ and the linear array sensor is on $u = 0$ of $o - uv$.

2.1 *Imaging Formulation of Perspective Projection*

In this subsection, perspective projection is adopted as a good approximation of the real camera imaging model. In this way, line scan cameras can be comprehended as a special case of frame cameras because perspective projection is satisfied only in the view plane.

Let us consider a space point $P(X, Y, Z)$ on the view plane. It is projected onto the linear array sensor at image point $p(0, v)$. Therefore, without consideration of lens distortion, the line scan camera imaging model can be represented as follows:

$$\lambda \begin{bmatrix} 0 \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & & \\ & f & v_0 \\ & & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}. \tag{1}$$

In Eq. (1), λ is an arbitrary scale factor. f is the focal length (in pixels) and v_0 is the coordinate of the principal point in image v-axis. f and v_0 are called intrinsic parameters which are fixed as long as the line scan camera is configured. r_{ij} and t_i $(i, j = 1, 2, 3)$ are, respectively, elements of rotation matrix $\mathbf{R}$ and translation vector $\mathbf{T}$. Rotation matrix $\mathbf{R}$ is parameterized by a vector of three parameters, denoted by $\mathbf{r} = [r_1 r_2 r_3]^T$ according to the Rodrigues formula.[13] r_1, r_2,

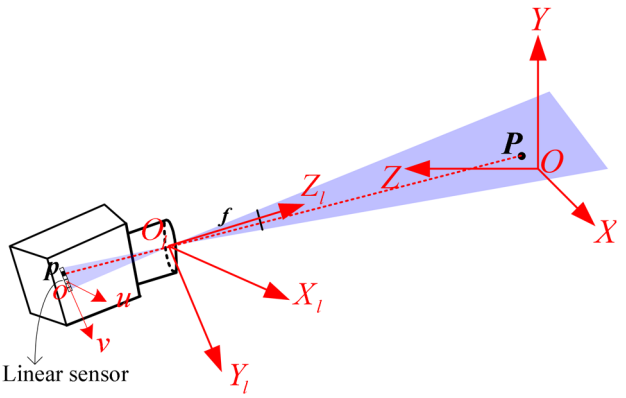

Fig. 1 A general view of the line scan camera imaging model.

r_3, t_1, t_2, t_3 are called extrinsic parameters which relate the world coordinate system $O-XYZ$ to the line scan camera coordinate system $O_l-X_lY_lZ_l$.

Consequently, the line scan camera imaging model including both intrinsic and extrinsic parameters can be rephrased as follows:

$$\begin{cases} 0 = r_{11}X + r_{12}Y + r_{13}Z + t_1 \\ v = f\frac{r_{21}X+r_{22}Y+r_{23}Z+t_2}{r_{31}X+r_{32}Y+r_{33}Z+t_3} + v_0 \end{cases}. \tag{2}$$

Actually in Eq. (2), the first equation is the function of the view plane in the world coordinate system while the second equation describes perspective projection in the view plane.

2.2 *Lens Distortion for Line Scan Cameras*

Up to now, we have not considered lens distortion of line scan cameras. In terms of high-precision photogrammetric applications, it is essential to rectify the lens distortion. In the case of industrial line scan cameras, lens distortion is small and mainly dominated by radial components especially for cameras with a long focal length.[14] Therefore, in this subsection, only the first term of radial distortion is taken into our consideration.

According to Brown's theory,[14] if we mark s as in Eq. (3), a rigorous line scan camera imaging model considering radial distortion is represented as follows

$$s = \frac{r_{21}X + r_{22}Y + r_{23}Z + t_2}{r_{31}X + r_{32}Y + r_{33}Z + t_3}, \tag{3}$$

$$\begin{cases} 0 = r_{11}X + r_{12}Y + r_{13}Z + t_1 \\ v = fs(1 + ks^2) + v_0 \end{cases}, \tag{4}$$

where k is the first-order coefficient of radial distortion.

In this line scan camera imaging model, a total of nine independent parameters f, v_0, k, $r_1r_2r_3$, t_1, t_2, t_3 are unknown and to be determined.

3 Our Calibration Scheme

3.1 *Calibration Pattern*

In this paper, a special planar calibration pattern is designed to establish correspondences between space points on the pattern and their projections in a linear array image. Different from traditional chessboard calibration pattern, our pattern comprises three right triangles and six cross corner markers as shown in Fig. 2. The equations of feature lines and the coordinates of feature points are known in the pattern coordinate system $O-XYZ$. L_1, L_3, L_5 are parallel to the X-axis while L_2, L_4, L_6 are parallel to each other. Without loss of generality, we assume the pattern is on $Z = 0$ of the pattern coordinate system $O-XYZ$.

3.2 *Establishing Point Correspondences*

In our proposal, the line scan camera is rigidly coupled to an auxiliary frame camera (see Fig. 3). The frame camera is accurately calibrated beforehand with Zhang's method.[6] Our calibration algorithm only requires the stereo cameras to observe the planar calibration pattern shown at a few different orientations as shown in Fig. 3. At each orientation, a stereo pair is obtained including a frame image and a linear

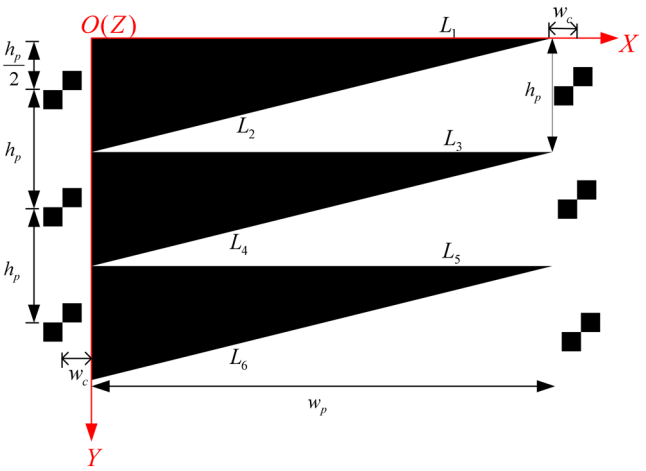

Fig. 2 Calibration pattern.

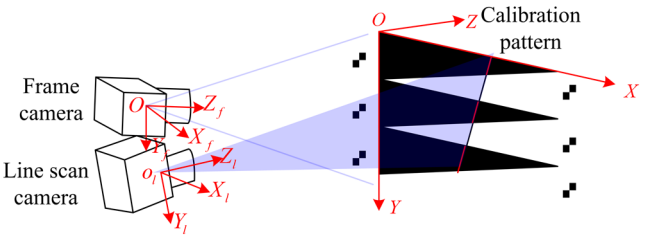

Fig. 3 A general view of the stereo cameras and the calibration pattern.

array image (see Fig. 4). In Fig. 4, $P_i(X_i, Y_i, 0)(i = 1,2 \cdots 6)$ is the intersection point of the view plane and feature line $L_i(i = 1,2 \cdots 6)$. $p_i(0, v_i)$ is the corresponding projection of $P_i(X_i, Y_i, 0)$ in the linear array image.

Cross-ratio is an important 1-D projective invariance in computer vision.[15] The cross-ratio of four arbitrary colinear points m_1, m_2, m_3, m_4 is represented as follows:

$$\text{Cross-ratio}_{m_1,m_2,m_3,m_4} = \frac{\vec{m_1m_3}}{\vec{m_2m_3}} : \frac{\vec{m_1m_4}}{\vec{m_2m_4}}. \tag{5}$$

We shall now show how the coordinates of P_3, P_5 are extracted from the pattern geometry based on the projective invariance of cross-ratio. According to the definition of cross-ratio, $\text{Cross-ratio}_{P_2,P_4,P_6,P_3}$ and $\text{Cross-ratio}_{P_4,P_6,P_2,P_5}$

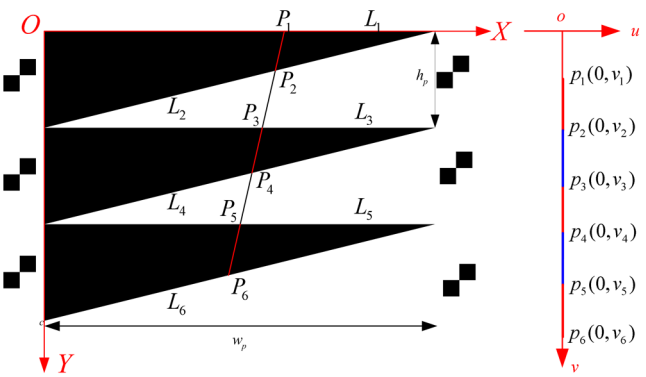

Fig. 4 The frame image and the linear array image.

can be represented with the directed line segments as follows:

$$\text{Cross-ratio}_{P_2,P_4,P_6,P_3} = \frac{P_2P_6}{P_4P_6} : \frac{P_2P_3}{-1 \cdot P_3P_4}, \tag{6}$$

$$\text{Cross-ratio}_{P_4,P_6,P_2,P_5} = \frac{-1 \cdot P_2P_4}{-1 \cdot P_2P_6} : \frac{P_4P_5}{-1 \cdot P_5P_6}. \tag{7}$$

From the parallelism of feature line L_2, L_4, L_6 in Fig. 4, the following equations can be derived:

$$P_2P_6 = 2 \cdot P_2P_4 = 2 \cdot P_4P_6, \tag{8}$$

$$\frac{P_2P_3}{P_3P_4} = \frac{X_3}{w_p - X_3}, \tag{9}$$

$$\frac{P_4P_5}{P_5P_6} = \frac{X_5}{w_p - X_5}. \tag{10}$$

Substituting Eqs. (8)–(10) into Eqs. (6) and (7), the following equations can be derived

$$X_3 = \frac{2w_p}{2 - \text{Cross-ratio}_{P_2,P_4,P_6,P_3}}, \tag{11}$$

$$X_5 = \frac{w_p}{1 - 2\text{Cross-ratio}_{P_4,P_6,P_2,P_5}}. \tag{12}$$

From the projective invariance of cross-ratio, the following equations are derived

$$\begin{aligned}\text{Cross-ratio}_{P_2,P_4,P_6,P_3} &= \text{Cross-ratio}_{p_2,p_4,p_6,p_3} \\ &= \frac{v_2 - v_6}{v_4 - v_6} : \frac{v_2 - v_3}{v_4 - v_3},\end{aligned} \tag{13}$$

$$\begin{aligned}\text{Cross-ratio}_{P_4,P_6,P_2,P_5} &= \text{Cross-ratio}_{p_4,p_6,p_2,p_5} \\ &= \frac{v_4 - v_2}{v_6 - v_2} : \frac{v_4 - v_5}{v_6 - v_5}.\end{aligned} \tag{14}$$

Substituting Eqs. (13) and (14) into Eqs. (11) and (12), the coordinates of P_3 and P_5 are worked out as in Table 1.

From the geometry of the pattern, the equations of feature line L_1, L_3, L_5 and L_2, L_4, L_6 are known as in Table 2.

P_3P_5 is the intersection line of the view plane and the planar pattern. It can be represented with the function $aX + bY + c = 0$, if we set $a = Y_3 - Y_5$, $b = X_5 - X_3$ and $c = X_3Y_5 - X_5Y_3$. Therefore, $P_i(i = 1, 2 \cdots 6)$ are calculated out in Table 3 as the intersections points of line P_3P_5 and feature line L_i in Table 2.

Table 1 Coordinates of P_3 and P_5.

	X	Y	Z
P_3	$\frac{w_p}{1-\frac{(v_2-v_6)(v_4-v_3)}{2(v_4-v_6)(v_2-v_3)}}$	h_p	0
P_5	$\frac{w_p}{1-\frac{2(v_4-v_2)(v_6-v_5)}{(v_6-v_2)(v_4-v_5)}}$	$2h_p$	0

Table 2 Functions of feature line $L_i(i = 1,2 \cdots 6)$.

$L_i(i = 1,3,5)$	$Y = \frac{i-1}{2} h_p$
$L_i(i = 2,4,6)$	$Y = -\frac{h_p}{w_p} X + \frac{i}{2} h_p$

Table 3 Coordinates of $P_i(i = 1,2 \cdots 6)$.

	X	Y	Z
$P_i(i = 1,3,5)$	$\frac{-\frac{i-1}{2}bh_p - c}{a}$	$\frac{i-1}{2} h_p$	0
$P_i(i = 2,4,6)$	$\frac{-\frac{i}{2}bw_ph_p - cw_p}{aw_p - bh_p}$	$\frac{\frac{i}{2}aw_ph_p}{aw_p - bh_p}$	0

In our proposal, several views of the calibration pattern are required to be obtained at different orientations. At each orientation, the coordinates of six intersection points $P_i(i = 1, 2 \cdots 6)$ are obtained in the pattern coordinate system as in Table 3. However, these coordinates are not represented in a common coordinate system. To transform these coordinates into a common coordinate system, the line scan camera is combined to a calibrated frame camera. The position and pose of the pattern at each orientation can be determined from the feature points on the pattern and their corresponding projections in the frame image, which is a classical P6P problem. Let us define the frame camera coordinate system $O_f - X_fY_fZ_f$ as the common coordinate system and assume $n(n \geq 2)$ frame images of the pattern are captured. $\mathbf{R_j}$ and $\mathbf{T_j}$ are, respectively, the rotation matrix and translation vector of the pattern at the $j_{th}(j = 1, 2 \cdots n)$ orientation from the pattern coordinate system $O - XYZ$ to the frame camera coordinate system $O_f - X_fY_fZ_f$. Then the coordinates of $P_i(i = 1,2 \cdots 6)$ can be transformed into $O_f - X_fY_fZ_f$ as follows

$$\begin{bmatrix} X_{ij} \\ Y_{ij} \\ Z_{ij} \end{bmatrix} = \mathbf{R_j} \begin{bmatrix} X_i \\ Y_i \\ 0 \end{bmatrix} + \mathbf{T_j}, \tag{15}$$

where $(X_{ij}, Y_{ij}, Z_{ij})^T$ is the coordinates of P_i represented in the frame camera coordinate system.

3.3 *Closed-Form Solution*

In this step, we adopt DLT to get a closed-form solution. Before applying DLT to the line scan camera imaging model, we extract X from the first formula of Eq. (2) and substitute it into the second formula to achieve mathematical simplification. Therefore without consideration of lens distortion, a simplified line scan camera imaging model is obtained as follows:

$$\begin{cases} 0 = r_{11}X + r_{12}Y + r_{13}Z + t_1 \\ v = \frac{(fr_{33} - v_0r_{23})Y + (-fr_{32} + v_0r_{22})Z + fr_{11}t_2 - fr_{21}t_1 + v_0r_{11}t_3 - v_0r_{31}t_1}{-r_{23}Y + r_{22}Z + r_{11}t_3 - r_{31}t_1} \end{cases}. \tag{16}$$

Given n views of the pattern, we can establish $6n$ correspondences between space points in the frame camera

coordinate system and image points in the linear array image. Substituting the $6n$ correspondences into the simplified line scan camera imaging model in Eq. (16), two homogeneous equation systems can be obtained as follows:

$$\begin{bmatrix} X_{11} & Y_{11} & Z_{11} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ X_{ij} & Y_{ij} & Z_{ij} & 1 \\ \vdots & \vdots & \vdots & \vdots \\ X_{6n} & Y_{6n} & Z_{6n} & 1 \end{bmatrix} \mathbf{J} = 0, \tag{17}$$

$$\begin{bmatrix} Y_{11} & Z_{11} & 1 & -v_{11}Y_{11} & -v_{11}Z_{11} & -v_{11} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Y_{ij} & Z_{ij} & 1 & -v_{ij}Y_{ij} & -v_{ij}Z_{ij} & -v_{ij} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Y_{6n} & Z_{6n} & 1 & -v_{6n}Y_{6n} & -v_{6n}Z_{6n} & -v_{6n} \end{bmatrix} \mathbf{K} = 0. \tag{18}$$

In Eqs. (17) and (18), (X_{ij}, Y_{ij}, Z_{ij}) is the coordinate of $P_i (i = 1, 2 \cdots 6)$ corresponding to $j_{th} (j = 1, 2 \cdots n)$ view represented in the frame camera coordinate system and v_{ij} is the corresponding image coordinate of P_i. The explicit forms of $\mathbf{J}$ and $\mathbf{K}$ are as follows:

$$\mathbf{J} = \begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \\ t_1 \end{bmatrix}, \tag{19}$$

$$\mathbf{K} = \begin{bmatrix} fr_{33} - v_0 r_{23} \\ -fr_{32} + v_0 r_{22} \\ fr_{11}t_2 - fr_{21}t_1 + v_0 r_{11}t_3 - v_0 r_{31}t_1 \\ -r_{23} \\ r_{22} \\ r_{11}t_3 - r_{31}t_1 \end{bmatrix}. \tag{20}$$

If a closed-form solution exists, $\mathbf{J}$ and $\mathbf{K}$ can be estimated up to a scale factor by solving the two homogeneous equation systems via singular value decomposition of the coefficient matrix. The pending parameters f, v_0, $r_1 r_2 r_3$, t_1, t_2, t_3 can be solved from the two column vectors $\mathbf{J}$ and $\mathbf{K}$.

If $J(i)(i = 1, 2, 3, 4)$ denotes the i'th element of $\mathbf{J}$, r_{11}, r_{12}, r_{13}, t_1 can be solved with the constraint $r_{11}^2 + r_{12}^2 + r_{13}^2 = 1$ as follows:

$$\begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \\ t_1 \end{bmatrix} = \frac{\pm 1}{\sqrt{J(1)^2 + J(2)^2 + J(3)^2}} \begin{bmatrix} J(1) \\ J(2) \\ J(3) \\ J(4) \end{bmatrix}. \tag{21}$$

Let us assume $\mathbf{K}$ is retrieved up to a scale factor γ, then the following equation system can be derived

$$\begin{cases} \gamma K(4) = -r_{23} \\ \gamma K(5) = r_{22} \\ \gamma K(1) = f_y r_{33} - v_0 r_{23} \\ \gamma K(2) = -f_y r_{32} + v_0 r_{22} \\ \gamma K(3) = f_y r_{11} t_2 - f_y r_{21} t_1 + v_0 r_{11} t_3 - v_0 r_{31} t_1 \\ \gamma K(6) = r_{11} t_3 - r_{31} t_1 \end{cases}. \tag{22}$$

According to the orthonormality of the rotation matrix, the following two equations can be derived

$$\begin{cases} r_{11}r_{21} + r_{12}r_{22} + r_{13}r_{23} = 0 \\ r_{21}^2 + r_{22}^2 + r_{23}^2 = 1 \end{cases}. \tag{23}$$

Substituting the first two equations of Eq. (22) into Eq. (23), the scale factor γ can be solved as follows:

$$\gamma = \pm \left\{ K(5)^2 + K(4)^2 + \left[\frac{K(4) \cdot r_{13} - K(5) \cdot r_{12}}{r_{11}} \right]^2 \right\}^{-1}. \tag{24}$$

Accordingly, r_{21}, r_{22}, r_{23} can be solved as follows:

$$\begin{cases} r_{22} = \gamma K(5) \\ r_{23} = -\gamma K(4) \\ r_{21} = -(r_{12}r_{22} + r_{13}r_{23})/r_{11} \end{cases}. \tag{25}$$

Therefore, r_{31}, r_{32}, r_{33} are obtained from the orthonormality of the rotation matrix as follows:

$$\begin{bmatrix} r_{31} \\ r_{32} \\ r_{33} \end{bmatrix} = \begin{bmatrix} r_{11} \\ r_{12} \\ r_{13} \end{bmatrix} \times \begin{bmatrix} r_{21} \\ r_{22} \\ r_{23} \end{bmatrix}. \tag{26}$$

The remaining pending parameters f, v_0, t_2, t_3 can be solved from the last four equations in Eq. (22).

Up to now, a closed-form solution of all the camera parameters has been determined except the radial distortion coefficient k. However, we should notice that mirror-pose solutions exist due to the sign ambiguity in Eqs. (21) and (24). Fortunately, given our choice of the coordinate systems, we can pick out the right solution that guarantees the planar pattern in front of the line scan camera.

3.4 *Nonlinear Optimization*

In Sec. 3.3, the closed-form solution is solved based on the assumption that perspective projection is the line scan camera imaging model. Under this assumption, space points are derived from the pattern geometry and cross-ratio. However, in the real line scan camera imaging model, the projective invariance of cross-ratio is not strictly satisfied due to image noise. As a result, the space points derived in Tables 1 and 3 are not strictly lying in the view plane theoretically. Instead, they are lying randomly on both sides of the view plane. Therefore, a nonlinear optimization is applied to solve this problem and promote the calibration precision further.

3.4.1 *Updating intersection points*

In this subsection, we mark all the pending camera parameters with a vector $\phi = [f, v_0, k, r_1, r_2, r_3, t_1, t_2, t_3]$. The

nonlinear optimization is initialized with the closed-form solution ϕ_0 in which k is initialized with zero because radial distortion is extremely small for industrial line scan cameras. Unlike Luna's calibration method[10] in which space points derived from cross-ratio remain unchanged during the whole calibration process, our method updates space points from the derived extrinsic camera parameters after each cycle of iteration. We shall now show how the coordinates of space points are derived from the extrinsic parameters.

Let us assume that in a certain cycle of iteration, the rotation matrix $\mathbf{R}$ and translation vector $\mathbf{T}$ from the frame camera coordinate system to the line scan camera coordinate system are derived from $r_1 r_2 r_3$, t_1, t_2, t_3. The rotation matrix and translation vector from the pattern coordinate system at the j'th orientation to the frame camera coordinate system is defined by $\mathbf{R_j}$, $\mathbf{T_j}$. Then, the rotation matrix $\mathbf{R'_j}$ and translation vector $\mathbf{T'_j}$ from the pattern coordinate system at the j'th orientation to the line scan camera coordinate system are derived as follows

$$\begin{cases} \mathbf{R'_j} = \mathbf{R}\mathbf{R_j} \\ \mathbf{T'_j} = \mathbf{R}\mathbf{T_j} + \mathbf{T} \end{cases}, \tag{27}$$

where $\mathbf{R'_j}$ and $\mathbf{T'_j}$ are denoted as follows:

$$\mathbf{R'_j} = \begin{bmatrix} r'_{11} & r'_{12} & r'_{13} \\ r'_{21} & r'_{22} & r'_{23} \\ r'_{31} & r'_{32} & r'_{33} \end{bmatrix}, \tag{28}$$

$$\mathbf{T'_j} = \begin{bmatrix} t'_1 \\ t'_2 \\ t'_3 \end{bmatrix}. \tag{29}$$

Therefore, the view plane is derived in the pattern coordinate system as follows:

$$0 = r'_{11}X + r'_{12}Y + r'_{13}Z + t'_1. \tag{30}$$

We assume the pattern is on $Z = 0$ of the pattern coordinate system $O - XYZ$. Then the intersection line of the view plane and the pattern is $0 = r'_{11}X + r'_{12}Y + t'_1$. Thus, the coordinates of $P_i (i = 1, 2 \cdots 6)$ are calculated as in Table 3, if we set $a = r'_{11}$, $b = r'_{12}$, and $c = t'_1$.

3.4.2 *Nonlinear optimization*

If n views of the pattern are captured, the nonlinear optimization procedure can be applied to reduce the reprojection error by minimizing the following cost function

$$\sum_{j=1}^{n} \sum_{i=1}^{6} [v_{ij} - \hat{v}_{ij}(f, v_0, k, r_1 r_2 r_3, t_1, t_2, t_3, P_i)]^2, \tag{31}$$

where P_i is the i'th$(i = 1, 2 \cdots 6)$ intersection points derived from the line scan camera extrinsic parameters as described in Sec. 3.4.1. $\hat{v}_{ij}(f, v_0, k, r_1 r_2 r_3, t_1, t_2, t_3, P_i)$ is the projection of $P_i(r_1 r_2 r_3, t_1, t_2, t_3)$ in the j'th linear array image according to the line scan camera imaging model described in Eq. (4). For n views of the pattern, we must optimize nine camera parameters $\phi = [f, v_0, k, r_1 r_2 r_3, t_1, t_2, t_3]$.

The explicit form of the cost function in Eq. (31) is highly nonlinear and complicated. Therefore, it is not wise to use gradient-based optimization methods to carry out the nonlinear iteration. In our implementation, we use the Nelder–Mead Simplex algorithm[16] to minimize the cost function.

The nonlinear optimization procedure is carried out as follows:

Step 1: set counter $k = 0$ and initialize the camera parameters ϕ with the close-from solution ϕ_0.
Step 2: in the k'th cycle of iteration, calculate the intersection points on the planar pattern at each orientation with extrinsic parameters in ϕ_k using the method described in Sec. 3.4.1.
Step 3: calculate the correction $\Delta\phi_k$ for ϕ_k with the Nelder–Mead Simplex algorithm.
Step 4: the simplex size E is used as a stopping criteria. If E is smaller than the given threshold, the nonlinear iteration stops; else set $\phi_{k+1} = \phi_k + \Delta\phi_k$ and $k = k + 1$, then turn to step 2.

When the nonlinear iteration stops, ϕ_k is the optimal solution for the line scan camera parameters ϕ.

3.5 *Complete Calibration Algorithm*

In this subsection we present the complete plane-based algorithm for line scan camera calibration. From n views of the pattern:

1. Locate the cross-corner markers in the frame images and determine orientation of the calibration pattern in each view.
2. Calculate six intersection points on the pattern at each orientation from the linear array image with cross-ratio and the pattern geometry (see Sec. 3.2), and transform these intersection points from the pattern coordinate system to the frame camera coordinate system.
3. Estimate the coefficient matrix $\mathbf{J}$, $\mathbf{K}$ using point correspondences. The estimation of $\mathbf{J}$, $\mathbf{K}$ is equivalent to the so-called DLT and can be done by solving a linear equation system.
4. Extract the intrinsic camera parameters f, v_0 and extrinsic camera parameters $r_i, t_i (i = 1, 2, 3)$ from $\mathbf{J}$, $\mathbf{K}$.
5. Optional but recommended: nonlinear optimization of the intrinsic and extrinsic parameters by minimizing the reprojection error.

The flowchart of the complete plane-based algorithm is shown in Fig. 5.

4 Experimental Results

The proposed algorithm has been tested on both computer-simulated data and real data. Both tests are detailed in Secs. 4.1 and 4.2.

4.1 *Computer Simulations*

Three groups of experiments are performed using computer-simulated data to evaluate the performances of our calibration algorithm with respect to the following three factors: (1) the noise level for locating image points, (2) the number

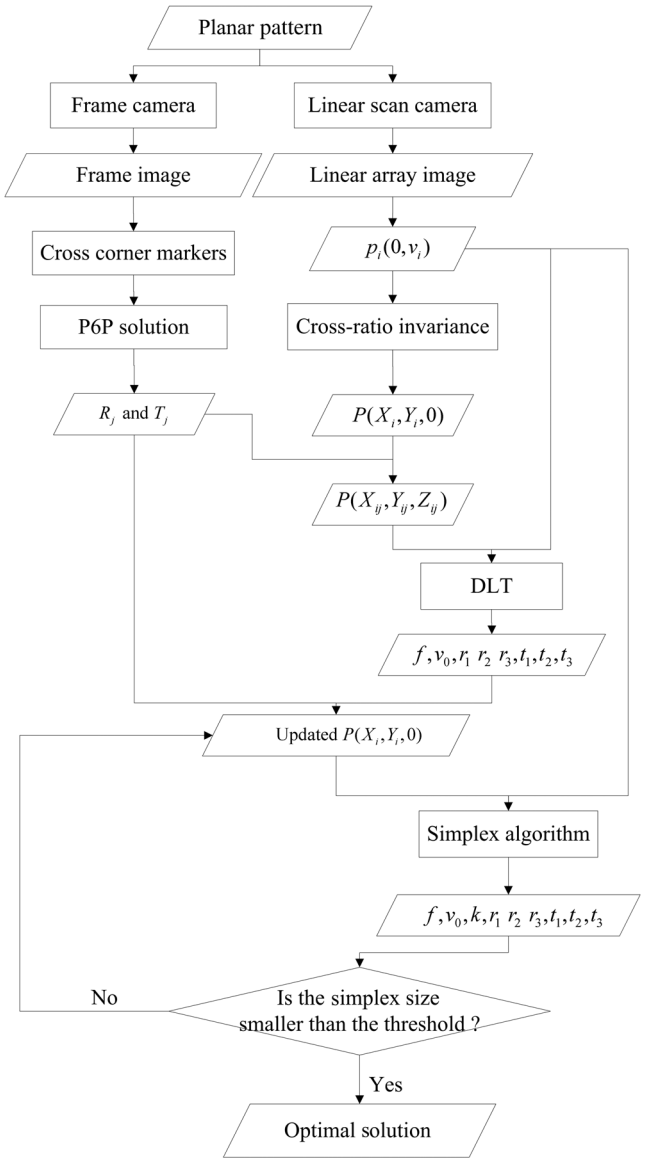

Fig. 5 The flowchart of the complete plane-based algorithm.

of pattern planes, and (3) the accuracy of frame camera intrinsic parameters.

In all the experiments, a simulated line scan camera, a simulated frame camera and a simulated planar pattern are used. The width and height of the simulated planar pattern are $w_p = 0.24$ m, $h_p = 0.06$ m. The sensor length of the simulated line scan camera is 2048 pixels; the sensor size of the simulated frame camera is 1600×1200 pixels. The intrinsic parameters of the two simulated cameras are listed in Table 4 along with the rigid transform parameters from the frame camera coordinate system to the line scan camera coordinate system. In this subsection, if there is no special announcement, the two simulated cameras are defined by parameters in Table 4.

In each group of experiments, three steps are taken to evaluate the performances of our calibration algorithm as follows:

1. The cross-corner markers on the pattern are projected on the simulated frame camera to produce the simulated frame image according to the frame camera imaging model. Noises are added to the coordinates of the projections of the feature points.
2. Intersection points of the view plane and the feature lines are derived from the rigid transform parameters in Table 4 using the method described in Sec. 3.4.1. The intersection points are then projected on the simulated line scan camera to produce the simulated linear array image according to the line scan camera imaging model. Noised are added to the coordinates of the projections of the intersection points.
3. We use the simulated frame images and the simulated linear array images as input data to recover the intrinsic parameters of the simulated line scan camera and the rigid transform parameters between the two simulated cameras through the calibration algorithm proposed in this paper.
4. The performances of our proposal are evaluated by comparing the recovered parameters with the ground truth listed in Table 4.

Table 4 Parameters of the line scan camera and frame camera.

Line scan camera	Frame camera	Rigid transform parameters
$f = 5000$ pixel	$f_x = f_y = 3000$ pixel	$r_1 = 0$ deg, $t_1 = 0$ m
$v_0 = 1024$ pixel	$u_0 = 800$ pixel	$r_2 = 0$ deg, $t_2 = -0.1$ m
$k = 0.001$	$v_0 = 600$ pixel	$r_3 = 0$ deg, $t_3 = 0$ m

An overview of the simulated experiments is shown in Fig. 6. The planar pattern is placed at several different positions with only a rotation around the X-axis of the pattern coordinate system. $\pi/18$ radian increment is added for each neighboring positions. Especially, for pattern M_1, there are only displacements between the pattern coordinate system $O - XYZ$ and the frame camera coordinate system $O_f - X_f Y_f Z_f$. In addition, the pattern center $(0.5w, 1.5h)$ of pattern M_1 is just 1 m away along the Z_f-axis.

4.1.1 *Performance with respect to the noise level for locating image points*

In this experiment, we use three patterns M_1, M_6, M_7. Gaussian noise with 0 mean and σ standard deviation is added to the projected image points in the frame image and linear array image. We vary σ from 0.1 pixels to 1.0 pixels in 0.1 pixels increment. As in Zhang,[6] we perform 100

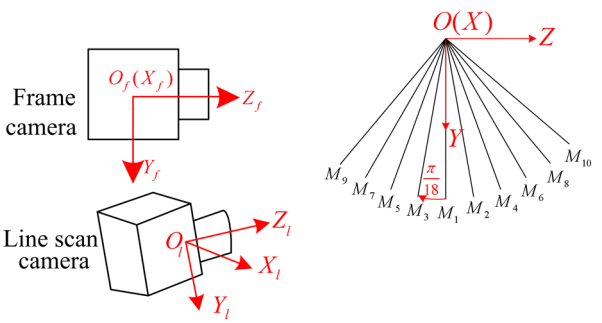

Fig. 6 Overview of the simulated experiments.

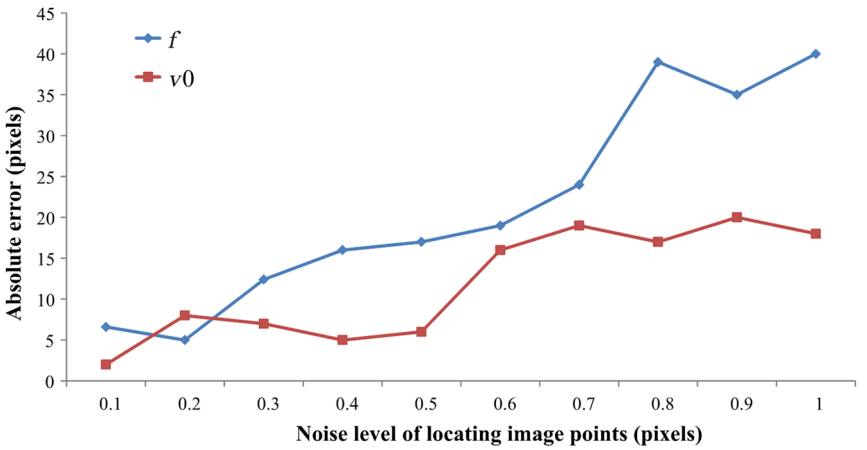

Fig. 7 Errors versus the noise level for locating the image points.

independent runs for each noise level and compute the average absolute errors for f and v_0. As we can see from Fig. 7, errors increase almost linearly with the noise level. For $\sigma = 0.5$ (which is larger than the normal noise in practical calibration), error in f is less than 20 pixels and error in v_0 is less than 10 pixels.

4.1.2 *Performance with respect to the number of pattern planes*

This experiment investigates the performance of our calibration algorithm with respect to the number of pattern planes (more precisely, the number of pairs of linear array images and frame images). At least two planes are required in our proposal. We vary the number of pattern planes from 2 to 10. For each number N, pattern $M_1, M_2 \cdots M_N$ are used and 100 independent runs are performed in which Gaussian noise with 0 mean and 0.5 pixels standard deviation is added. The average result is shown in Fig. 8. We notice that errors decrease as more pattern planes are used. From 2 to 5, the errors decrease significantly.

4.1.3 *Performance with respect to accuracy of frame camera intrinsic parameters*

The accuracy of frame camera intrinsic parameters determines orientations of the pattern planes and thus indirectly determines accuracy of line scan camera calibration. This experiment investigates the performance of our calibration algorithm with respect to accuracy of frame camera intrinsic parameters. Patterns $M_1, M_2 \cdots M_{10}$ are used. The linear array images and frame images are dirtied by Gaussian noise with 0 mean and 0.5 pixels standard deviation. Gaussian noise with 0 mean and σ standard deviation is added to the primary four frame camera intrinsic parameters f_x, f_y, u_0, v_0. We vary σ from 0 to 5 pixels in 0.5 pixels increment. For each noise level, 100 independent runs are performed. The average result is shown in Fig. 9. We notice that errors increase as accuracy of frame camera intrinsic parameters deteriorates.

4.2 *Real Data*

The line scan camera to be calibrated is a Balser sprint camera with Nikon 50 mm lens.[17] The basic characteristics of the line scan camera are listed in Table 5. The pixel size is 10 μm and thus the focal length is around 5000 pixels. The sensor length is 2048 pixels. The auxiliary frame camera is an ordinary off-the-shelf CCD camera whose sensor size is 1600×1200 pixels. The frame camera has been calibrated using the OpenCV library plane-based calibration routines and its intrinsic parameters are listed in Table 5. The line scan camera and frame camera are not required to be synchronous. The calibration pattern can be printed with a laser printer and attached to a planar object. The width

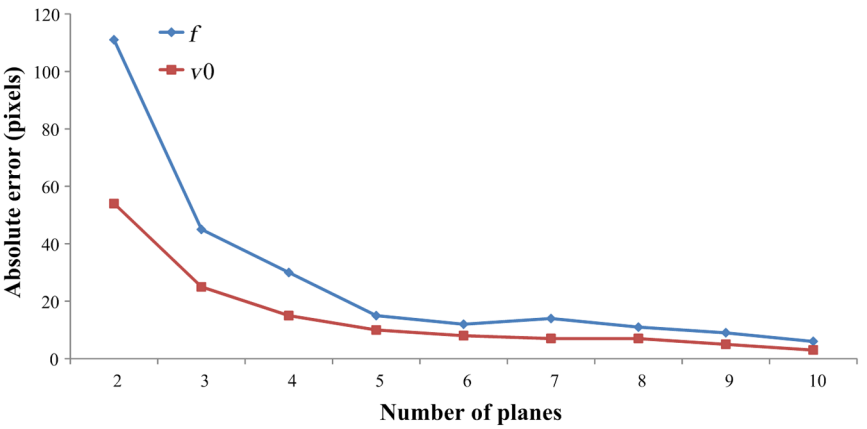

Fig. 8 Errors versus the number of the pattern planes.

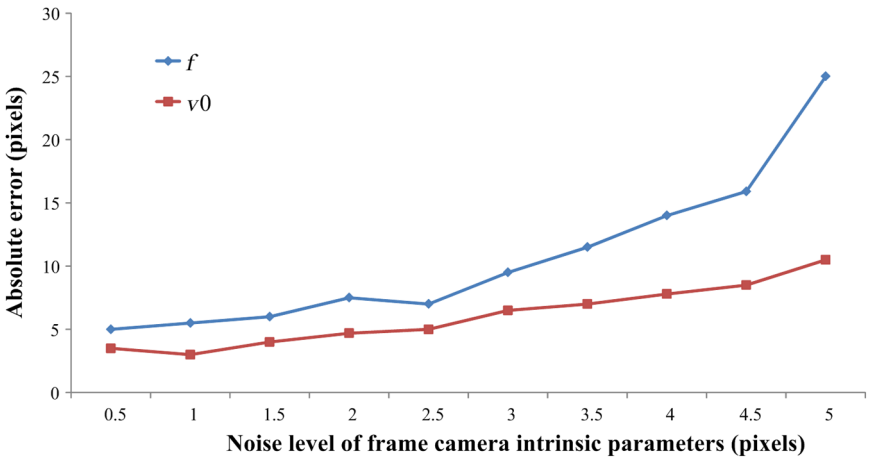

Fig. 9 Errors versus accuracy of frame camera intrinsic parameters.

Table 5 Characteristics of the line scan camera and frame camera.

Instruments	Characteristics
Line scan camera	Pixel size: 10 μm × 10 μm
	Sensor size: 2048 × 1 pixels
	Lens: 50 mm
Frame camera	Sensor size: 1600 × 1200 pixels
	$f_x = f_y = 2895$ pixels
	$u_0 = 795$ pixel, $v_0 = 608$ pixels

and height of the black triangles on the pattern are $w_p = 24$ cm and $h_p = 4$ cm.

Ten pairs of linear array images and frame images of the calibration pattern under different orientations are captured. We apply our calibration algorithm to all the images. The retrieved line scan camera intrinsic parameters and the rigid transform parameters between the two cameras are listed in Table 6. The former column is the linear solution

Table 6 Calibration result for line scan camera parameters.

Parameters	DLT	Optimal
f (pixels)	5198	5333
v_0 (pixels)	1090	1002
k	0.0032	0.0028
r_1 (deg)	0.0863	0.07185
r_2 (deg)	−0.3335	−0.3333
r_3 (deg)	0.0581	0.0614
t_1 (m)	0.2275	0.2278
t_2 (m)	0.0188	0.0200
t_3 (m)	−0.0534	−0.0310

Table 7 Reprojected errors for the real data.

	Max residual error	RMSE
v (pixel)	1.579	0.348

Table 8 Experimental results and comparisons.

Parameter	Hui's method	Our method	Error (%)
f (pixels)	5084	5155	1.40
v_0 (pixels)	1022	1004	1.76

solved through direct linear transform and the latter column is the optimal solution after nonlinear refinement.

To assess the quality of our calibration, all the intersection points on the calibration pattern are calculated and reprojected onto the linear array images using the calibration result in Table 6. The reprojected image points are compared with the real image points in the linear array images. The max residual error and RMSE for all the reprojected image points are shown in Table 7. The max residual error is less than 2 pixels and RMSE is less than 0.5 pixels which shows that our calibration algorithm performs well.

Since no ground truth was available to evaluate the calibration accuracy of the intrinsic parameters, we conducted another experiment and took the calibration method[8] proposed by Hui et al. as a reference. Experimental results and comparisons are reported in Table 8. We can see that the focal length and the principle point estimated by our method are very close to the estimation made by Hui's method (both parameters differ by less than 2%) which means our proposal works well.

5 Conclusion

In this paper, we proposed a flexible calibration technique for line scan cameras. The usage of an auxiliary frame camera and a planar pattern is a typical characteristic of our calibration technique. The technique only requires the stereo cameras to observe a planar pattern under (at least two) different

orientations. The calibration scheme includes three stages. In the first stage, points correspondences are established from the pattern geometry based on cross-ratio. The second stage consists of a close-form solution derived from DLT. In the third stage, a nonlinear optimization procedure is applied with the close-form solution as an initial guess.

Compared with classical techniques which use 3-D calibration object or controllable moving platforms, our technique is much more flexible and affordable for ordinary camera users. In addition, once the auxiliary frame camera is located and orientated in the world coordinate system, it can help determine the extrinsic parameters of the line scan camera in close-range photogrammetric applications.

Acknowledgments

The authors would like to thank Miss Yangliu Kuai for the language improvement, and the OE associate editor and anonymous reviewers for their constructive feedback on an earlier version of our paper.

References

1. K. Hirahara and K. Ikeuchi, "Detection of street-parking vehicles using line scan camera and scanning laser range sensor," in *Proc. IEEE Int. Vehicles Symp.*, pp. 656–661 (2004).
2. R. Gupta and R. I. Hartley, "Linear pushbroom cameras," *IEEE Trans. Pattern Anal. Mach. Int.* **19**(9), 963–975 (1997).
3. K. Kataoka et al., "3D building façade model reconstruction using parallel images acquired by line scan camera," in *Proc. IEEE Int. Conf. on Image Processing*, pp. 1009–1012 (2005).
4. Z. Zhao et al., "Model-based estimation for pose, velocity of projectile from stereo linear array image," *Meas. Sci. Rev.* **12**(3), 104–110 (2012).
5. R. Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses," *IEEE J. Rob. Autom. Mag.* **3**(4), 323–344 (1987).
6. Z. Zhang, "A flexible new techinique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(11), 1330–1334 (2000).
7. J. Drareni, S. Roy, and P. Sturm, "Plane-based calibration for linear cameras," *Int. J. Comput. Vision* **91**(2), 146–156 (2011).
8. B. Hui et al., "Line-scan camera calibration in close-range photogrammetry," *Opt. Eng.* **51**, 053602 (2012).
9. B. Hui et al., "A novel line scan camera calibration technique with an auxiliary frame camera," *IEEE Trans. Instrum. Meas.* **62**(9), (2013).
10. R. Horaud, R. Mohr, and B. Lorecki, "On single-scan line camera calibration," *IEEE Trans. Rob. Autom.* **9**(1), 71–75 (1993).
11. C. A. Luna et al., "Calibration of line-scan cameras," *IEEE Trans. Instrum. Meas.* **59**(8), 2185–2190 (2010).
12. D. Li et al., "Cross-ratio invariant based line scan camera geometric calibration with static linear data," *Opt. Lasers Eng.* **62**, 119–125 (2014).
13. O. Faugeras, *Three-Dimensional Computer Vision: A Geometric Viewpoint*, MIT Press (1993).
14. D. C. Brown, "Close-range camera calibration," *Photogramm. Eng.* **37**(8), 855–866 (1971).
15. R. Mohr and L. Morin, "Relative positioning from geometric invariants," in *Proc. Computer Vision Pattern Recognition Conf.*, Lahaina, Maui, Hawaii, pp. 139–144 (1991).
16. J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.* **7**, 308–315 (1965).
17. Balser Vision Technologies, http://www.balserweb.com/ (2009).

Dongdong Li received his BS degree from Wuhan University in 2012 and his MS degree from National University of Defense Technology in 2014. He has been pursuing his doctor's degree in information and communication engineering since 2015. His main research interests include camera calibration and efficient visual tracking.

Gongjian Wen received his BS, MS and PhD degrees from the National University of Defense Technology in 1994, 1997 and 2000, respectively. He completed two-year postdoctoral assignment at Wuhan University. Currently, he is the head of the fourth department of the Key Laboratory of ATR and is mainly interested in image understanding and photogrammetry and remote sensing.

Shaohua Qiu received his BS degree from Wuhan University in 2011 and his MS degree from National University of Defense Technology in 2013. He has been pursuing his doctor's degree in information and communication engineering since 2014. His main research interests include structure from motion and multi-view object recognition.