Journal of Electronic Imaging

SPIEDigitalLibrary.org/jei

Geometric calibration of line-scan camera using a planar pattern

Ming Yao
Zuyun Zhao
Bugao Xu

SPIE

IS&T
imaging.org

Journal of Electronic Imaging 23(1), 013028 (Jan–Feb 2014)

Geometric calibration of line-scan camera using a planar pattern

Ming Yao, Zuyun Zhao, and Bugao Xu*
University of Texas at Austin, School of Human Ecology, 200 West 24th Street, Austin, Texas 78712

Abstract. We present an innovative calibration method for line-scan cameras to estimate the intrinsic parameters. The calibration involves using a stationary planar pattern that consists of repeated vertical and slanted lines, and constructing a two-dimensional (2-D) calibration framework with one-dimensional (1-D) data. A feature point reconstruction method is applied to transform the 1-D camera calibration problem into the 2-D scope. Camera parameters are then solved by using a 2-D camera model with constraints unique to 1-D geometry. In our tests over 12 calibrations with images of 2048×2048 pixels, the average of the reprojection errors is 0.46 pixels. As opposed to other line-scan camera calibration techniques, this method does not require the camera to progressively scan a pattern, thus eliminating the need for additional mechanical devices to assist the calibration. This method does not need a three-dimensional pattern as a calibration target, either. The stationary planar target makes the calibration more suitable for an application that has to be done in a nonlaboratory setting, such as highway pavement inspection. *© 2014 SPIE and IS&T* [DOI: 10.1117/1.JEI.23.1.013028]

Keywords: camera calibration; line-scan camera; homography; projective mapping.

Paper 13703 received Dec. 17, 2013; revised manuscript received Jan. 22, 2014; accepted for publication Jan. 29, 2014; published online Feb. 21, 2014.

1 Introduction

A line-scan camera is an imaging device with only one row of photonic sensors to capture a two-dimensional (2-D) image by accumulating scanned lines over time. With this feature, line-scan cameras have been useful for a wide range of machine vision applications, such as pavement inspection,[1–4] position and motion estimations,[5,6] target detection,[7,8] etc. For the applications which require quantitative measurements, camera parameters must be known through geometric camera calibration, which has been rigorously studied.[9–11] However, most of these studies have been performed only for 2-D cameras.

In the 2-D camera calibration, the mapping of a feature point from its pattern coordinate space to the image coordinate space is related by the homography matrix, and the pair of the coordinates in both spaces is referred to as a correspondence. The numerical solution of a homography is achieved by applying numbers of feature point correspondences. Among almost all of the calibration methods, the estimation of the homography is the first step in solving the intrinsic and extrinsic camera parameters. The line-scan camera calibration problem can be approached by following similar steps. However, there are two major challenges. The first one is the difficulty in determining the location of a feature point in the pattern coordinate space, and the second is the scarcity of mathematical models suitable for solving one-dimensional (1-D) perspective projection geometry in a static setup. When a static scene is captured by a line-scan camera, contrasts exist only in the x-direction. With 1-D information, the corresponding feature points on a 2-D pattern plane are hard to be determined unless the exact location of the line-of-view can be found on the pattern plane.

To solve this problem, Horaud et al.[12] proposed a pattern of four coplanar straight lines, three of which were parallel and the fourth intercepted the three with an arbitrary angle. The line-scan camera saw four points from this pattern. By looking at the ratio of the distances between each pair of adjacent points, the location of each point could be estimated. A linear stage was used to drive the pattern in both y- and z-axes, so that sufficient independent feature points can be found. Because the movements were in known directions with predetermined steps, the shifts of the camera coordinate system can be translated into the offsets of coordinates of feature point in the pattern coordinate space, leaving the homography unchanged. Later, Luna et al.[13] extended this technique by utilizing two sets of parallel lines (vertical lines and slanted lines) interconnected in a zig-zag configuration. To avoid the use of a mechanical device, a second tier of the same line sets were added to the pattern such that the pattern provided feature points in two-different elevations. Sufficient independent feature points can be obtained in one shot. Although these patterns seem to be viable in line-scan camera calibration given that no additional device is needed to assist feature points estimation, they both rely on the assumption that camera lens distortion is small enough to be ignored, or constant within each repeated parallel line.

Besides these static setups, calibration based on the dynamic-imaging model with 2-D chessboard pattern were reported.[14,15] 2-D images were captured by scanning the pattern with a line-scan camera attached to a linear stage. The feature points (chessboard corners) in the image coordinate space were detected from the scanned image. In this type of calibration, the scanning rate of the camera must be synchronized with the speed of the stage. A virtual three-dimensional (3-D) homography between the pattern coordinate space, and the 2-D image coordinate space was adopted. This virtual homography was the result of the linear motion of the camera's 1-D sensor array due to the scan.

*Address all correspondence to: Bugao Xu, E-mail: bxu@utexas.edu

In this research, we propose a new calibration approach that only requires a line-scan camera to view a stationary planar pattern from two orientations. The multiple observations can be done by adjusting either the camera or the pattern. The relative positions between the camera and the pattern do not have to be known before the calibration. The homography matrix for each pattern observation is first estimated from feature point correspondences reconstructed from orthogonal projection in both the image plane and pattern plane, and then the initial camera parameters are obtained by solving the equations derived from the homography matrix under the orthonormal constraints, among each column. Finally, these parameters are optimized in a nonlinear fashion, which minimizes the reprojection error. The calibration pattern in our proposed method is essentially similar to the ones reported in Refs. 12 and 13. However, our method differs from Ref. 12 in that repeated line pattern are used so that mechanical movement is not necessary to capture enough feature points for homography estimation. Our method also differs from Ref. 13 in that a 2-D calibration pattern instead of a 3-D pattern is sufficient for the task. Compared with the approaches presented in Refs. 14 and 15, this technique is considerably easier to be implemented because there is no high-precision motion or synchronization involved.

2 Camera Model and Imaging Geometry

We start with the review of the classic geometry in 2-D camera model, followed by the discussion of unique and important constraints in 1-D imaging.

2.1 *Preliminary*

Without lens distortion, a 3-D projection modeled by a pinhole camera is described by

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{M}[\,\mathbf{R} \quad \mathbf{t}\,]\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}, \tag{1}$$

where $[X, Y, Z]^T$ are the coordinates of a 3-D point in pattern coordinate space, $[u, v]^T$ are the coordinates of the projection on the image in pixels, and s is an arbitrary scale factor. $[\,\mathbf{R} \quad \mathbf{t}\,]$ is a 3×4 matrix of extrinsic parameters, in which $\mathbf{R}$ is the rotation matrix defined on Euler angles (α, β, γ) and constructed with Rodrigues's rotation formula, $\mathbf{t}$ is the translation vector between the pattern coordinate space and the image coordinate space. $\mathbf{M}$ is called the camera intrinsic matrix and is defined by

$$\mathbf{M} = \begin{bmatrix} f_x & k & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{2}$$

in which f_x and f_y are the horizontal and vertical focal lengths in pixels, k is the skewness of the axes in the image plane, and (u_0, v_0) are the coordinates of the principal point, or optical center, on the image.

If a planar pattern is used in the calibration, then all the feature points can be defined on a single plane. Assume the z-coordinate of each of these feature points to be 0, and let $\mathbf{r}_i$ be the i'th column of rotation matrix $\mathbf{R}$, Eq. (1) can be rewritten as

$$s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \mathbf{M}[\,\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}\,]\begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}. \tag{3}$$

Equation (3) defines a homogeneous transformation of a point from the pattern coordinate space to the image coordinate space. Therefore, the points are related by a homography $\mathbf{H}$:

$$\mathbf{H} = \mathbf{M}[\,\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}\,]. \tag{4}$$

2.2 *Estimation of Homography for 1-D Line-Scan Camera*

A complete homography to map a world homogeneous coordinate $[X, Y, Z, 1]^T$ to an image homogeneous coordinate $[u, v, 1]^T$ is a 3×4 matrix with 11 degrees of freedom. These degrees of freedom, or parameters, can be broken down into two categories: five intrinsic and six extrinsic parameters. The six extrinsic parameters relate the camera orientation to a world coordinate space (or the pattern coordinate space) and consist of three rotations and three translations. In order to estimate the homography with 11 degrees of freedom, theoretically six pairs of independent correspondences are needed, because each pair generates two formulas according to Eq. (1), and a total of 12 equations are sufficient to solve 11 parameters. In the case of 1-D line-scan camera, one can assume $f_y = 0$, $v_0 = 0$, and $k = 0$, then there are eight parameters left to be solved, thus four pairs of independent point correspondences are needed to estimate the homography of a line-scan camera. However, since these four imaged feature points will be collinear on the image plane due to the 1-D geometry of the sensor array, their corresponded feature points in the world coordinate space must not be collinear to avoid redundancy. Otherwise, the correspondences defined by the third and fourth point is over constrained because their coordinates in pattern coordinate space can be estimated by back projecting rays from the focal point, passing through their image points, and intercepting the line-of-view, as is shown in Fig. 1.

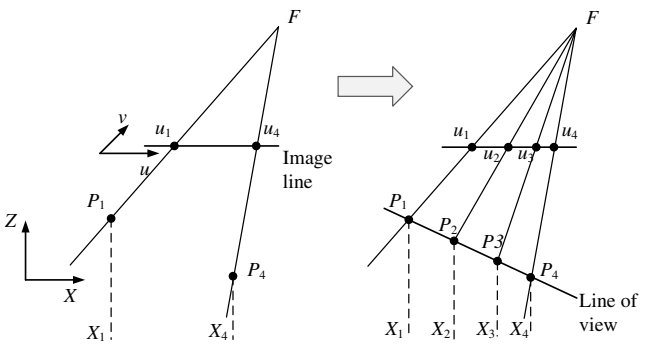

Fig. 1 More than two feature points that are colinear in the pattern coordinate space do not add new constraints in solving the homography. P_2 and P_3 can be estimated by extending the ray from the focal point F to the line-of-view.

3 Calibration Pattern and Feature Point Correspondences

3.1 *Calibration Pattern Design*

The calibration pattern we used is a planar pattern of two sets of mutually parallel lines. The first set consists of 58 evenly spaced vertical parallel lines $\{P_i\} = \{P_1, P_2, \cdots, P_N\}$ $(N = 58)$, and the second set consists of 57 slanted lines $\{D_i\} = \{D_1, D_2, \cdots, D_{N-1}\}$, which are placed between the vertical lines. Figure 2 displays two of the repeated units to show how these lines are constructed.

All these lines are defined on $Z = 0$ plane. The origin of the pattern coordinate space is placed at the intersection of lines P_1 and D_1. The direction of the x-axis is perpendicular to the vertical line set $\{P_i\}$, and the direction of y-axis is parallel to $\{P_i\}$. The pattern coordinate space is defined in such a way that the line sets $\{P_i\}$ and $\{D_i\}$ can be expressed in two simple formulas:

$$P_i\colon X = ih, \tag{5}$$

and

$$D_i\colon Y = \frac{l}{h}X - (i-1)l, \tag{6}$$

in which l is the length of each vertical line in set $\{P_i\}$, and h is the distance between each adjacent vertical line, i denotes the i'th line in the set.

3.2 *Feature Points Estimation*

Figure 2 illustrates the geometrical configuration of the calibration pattern being viewed by a line-scan camera. In this figure, F is the focal point of the camera, line Q_2 is the image line. The plane determined by F and Q_2 is the view plane, which intersects the pattern at line Q_1. Q_1 indicates the line-of-view of the camera. When the camera looks at this pattern, it sees a series of dots, which are the intersection points of the view line Q_1 and the pattern lines. The task of constructing point correspondences is to determine the coordinates of each feature point in both pattern coordinate space and image coordinate space.

For each repeated unit, let A, B, C, and D denote the four feature points seen by the camera, i.e., the intersections of

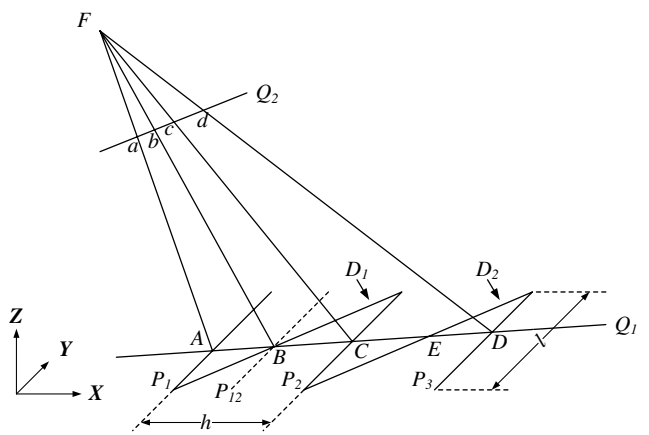

Fig. 2 The calibration pattern being viewed by a line-scan camera. Only two-repeated line structures are shown.

the view line with $\{P_i\}$ and $\{D_i\}$, and a, b, c, and d are their images, respectively. Since points A, B, C, and D are collinear and the rays FA, FB, FC, and FD are concurrent, the cross-ratio[16] of these points is thus defined as

$$r_{ABCD} = (\overline{CA}/\overline{CB})/(\overline{DA}/\overline{DB}). \tag{7}$$

Because the cross-ratio of collinear points does not change under perspective projection, we have

$$r_{abcd} = (\overline{ca}/\overline{cb})/(\overline{da}/\overline{db}) = r_{ABCD}. \tag{8}$$

Since the coordinates of a, b, c, and d can be readily detected from a line-scanned image, one can compute the cross-ratio $r = r_{abcd}$ from these coordinates, and further determine the coordinates of A, B, C, and D as follows. Assume η is the ratio of $\overline{BA}$ to $\overline{CA}$. With r being available and $\overline{DA} = 2 \times \overline{CA}$, the relationship between η and r can be solved

$$\eta = \frac{2r-2}{2r-1}. \tag{9}$$

Therefore, the x-coordinate of point B can be calculated as $x_B = \eta h$. Since point B is also on the line D_1, its y-coordinate can also be calculated by Eq. (6). Similarly, the intersection point E of the view line to D_2 can also be calculated by introducing the fourth vertical line P_4. With the coordinates of both B and E being determined, the equation of the view line section between P_1 and P_3 can be formulated, and the coordinates of points A, C, and D can be determined as well.

3.3 *Feature Point Correspondences*

From Sec. 2.2, we have concluded that to solve an 8 degrees of freedom 3-D homography, at least four feature points that are not collinear in the pattern coordinate space must be available. However, a planar pattern can only provide two independent feature point correspondences. To solve this problem, we project the original feature points onto the orthogonal direction of the view line on the pattern plane.

To create the projected feature points within the image coordinate space, we rotate the originally detected 1-D feature points by 90 deg around an estimated optical center u_0, which is initialized as half the width of the image. By doing this, it is equivalent to convert the 1-D image into 2-D in which the contrast in the second dimension is exactly duplicated from the first dimension. Thus, our 1-D camera calibration problem becomes a 2-D problem, on which some constraints on the intrinsic parameters can be applied: (1) the focal length must be the same in both vertical and horizontal direction, i.e., $f_x = f_y$; (2) and the optical center in both vertical and horizontal direction must be the same, i.e., $u_0 = v_0$. Assume the estimated optical center is at (u_0, v_0), the i'th feature point in the line-scanned image will be mapped to a point on a 2-D image with coordinate (u_i, v_0), in which u_i is the measured feature location from the line-scanned image. Then its projected counterpart can be computed as

$$\begin{bmatrix} u_i' \\ v_i' \end{bmatrix} = \mathbf{R}_{\text{proj}}\left(\begin{bmatrix} u_i \\ v_0 \end{bmatrix} - \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}\right) + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}, \tag{10}$$

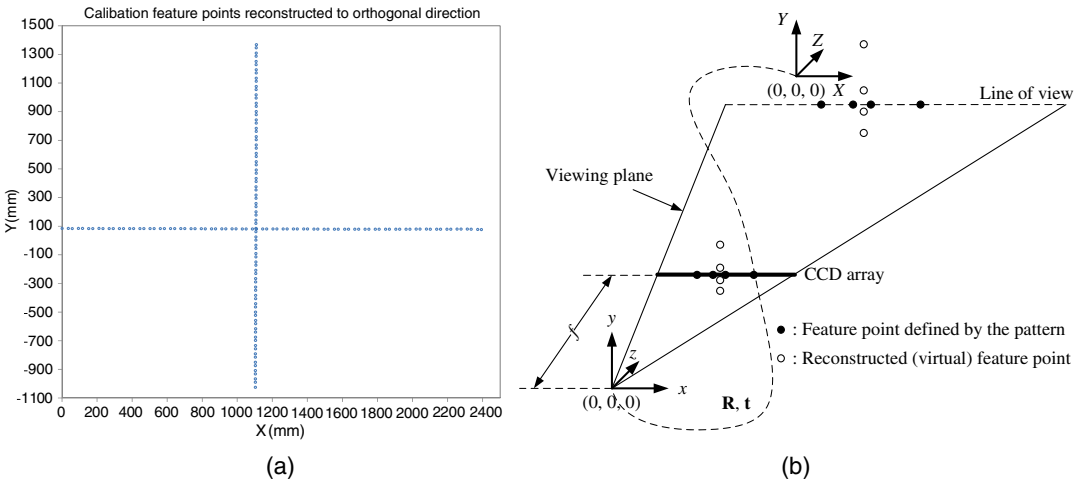

Fig. 3 Reconstruction of calibration feature points. (a) Feature points in the pattern coordinate space with reconstructed points in the orthogonal (vertical) direction. (b) Geometrical illustration of the features points viewed by a line-scan camera. Rotation and translation from the pattern coordinate space to the camera coordinate space are related by **R** and **t**.

in which $u_i' \equiv v_0$, $\mathbf{R}_{\text{proj}}$ is the rotation matrix defined by Rodrigues equation. Since rotation only exists on the xy-plane, it can be simplified to be a 2×2 matrix. When a rotation angle of $\pi/2$ is applied,

$$\mathbf{R}_{\text{proj}} = \begin{bmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}. \tag{11}$$

The same technique is applied to the feature points in pattern coordinate space. Figure 3(a) shows an example of all the feature points plotted on the pattern coordinate space. In this figure, the horizontally arranged point array indicates the position of the view line, and the vertically arranged point array indicates the positions of the reconstructed feature points. The intersection of these two point arrays is determined by back-projecting the optical center (u_0, v_0) onto the pattern surface. The geometrical illustration of the feature points being viewed by a line-scan camera is shown in Fig. 3(b).

4 Solving the Camera Parameters

Due to the nonlinear mapping introduced by lens distortion in the camera model, a solution of all camera parameters cannot be estimated directly in a least-squares fashion. Generally, given a set of feature point correspondences, camera parameters are optimized iteratively.[9,10,14,15] However, the feature point estimation described in Sec. 3.3 requires the intrinsic parameters, i.e., the calibration output, to construct the point correspondences. This condition can be solved by nesting the iterative nonlinear optimization process within an iterative estimation of feature points. The combined calibration process thus consists of three steps: (1) initialize the optical center at the image center and estimate the camera parameters from a closed form solution; (2) apply nonlinear optimization of camera parameters based on the initial estimates from (1); and (3) refine the positions of feature points based on optimized camera parameters. Steps 2 and 3 are repeated until the convergence is reached.

4.1 *Closed-Form Solution*

In the closed-form solution, the homography is estimated first. There are various algorithms proposed for homography estimation. We looked at the direct linear transform,[17] which is simple yet effective when combined with geometric distance cost function. Once the homography is computed, the intrinsic and extrinsic parameters will be available by solving Eq. (4).

Denote $\mathbf{H} = [\,\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_3\,]$ for homography that is estimated for each observation of the calibration pattern, $\mathbf{h}_j$ is the j'th column of $\mathbf{H}$. Equation (4) can be rewritten as

$$[\,\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_3\,] = \mathbf{M}[\,\mathbf{r}_1 \quad \mathbf{r}_2 \quad \mathbf{t}\,]. \tag{12}$$

Because $\mathbf{r}_1$ and $\mathbf{r}_2$ are orthonormal, we have

$$\begin{cases} \mathbf{r}_1 \cdot \mathbf{r}_2 = 0 \\ \|\mathbf{r}_1\| = \|\mathbf{r}_2\| \end{cases} \Rightarrow \begin{cases} \mathbf{h}_1^T \mathbf{M}^{-T}\mathbf{M}^{-1}\mathbf{h}_2 = 0 \\ \mathbf{h}_1^T \mathbf{M}^{-T}\mathbf{M}^{-1}\mathbf{h}_1 = \mathbf{h}_2^T \mathbf{M}^{-T}\mathbf{M}^{-1}\mathbf{h}_2 \end{cases}. \tag{13}$$

Since the homography has 8 degrees of freedom and there are six extrinsic parameters, two constraints in Eq. (13) are left to be applied on intrinsic parameters. Let $\mathbf{B} = \mathbf{M}^{-T}\mathbf{M}^{-1}$, thus $\mathbf{B}^T = (\mathbf{M}^{-T}\mathbf{M}^{-1})^T \equiv \mathbf{B}$. So, $\mathbf{B}$ is symmetric and can be defined by a six-element vector $[\,B_{11} \quad B_{12} \quad B_{22} \quad B_{13} \quad B_{23} \quad B_{33}\,]^T$, which represents the upper diagonal of $\mathbf{B}$.[10] Furthermore, since part of our feature point correspondences is reconstructed through orthogonal projection, we have $f_x = f_y$, $u_0 = v_0$, and $k = 0$. Thus, we obtain $B_{11} = B_{22}$, $B_{12} = 0$, and $B_{13} = B_{23}$, only three numbers are required to uniquely define $\mathbf{B}$. Then, we have $\mathbf{b} = [\,B_{11} \quad B_{13} \quad B_{33}\,]^T$ and

$$\mathbf{h}_i^T \mathbf{B} \mathbf{h}_j = \mathbf{v}_{ij}^T \mathbf{b}, \tag{14}$$

with

$$\mathbf{v}_{ij} = \begin{bmatrix} h_{i1}h_{j1} + h_{i2}h_{j2} \\ h_{i3}(h_{j1} + h_{j2}) + (h_{i1} + h_{i2})h_{j3} \\ h_{i3}h_{j3} \end{bmatrix}. \tag{15}$$

Therefore, two constraints defined in Eq. (13) can be transformed into two linear equations of $\mathbf{b}$:

$$\begin{bmatrix} \mathbf{v}_{12}^T \\ (\mathbf{v}_{11} - \mathbf{v}_{22})^T \end{bmatrix} \mathbf{b} = \mathbf{0}. \tag{16}$$

When n images of the calibration pattern are taken, by stacking n equations of Eq. (16) on top of each other, a total of $2n$ equations are available to solve $\mathbf{b}$ by applying the Levenberg–Marquardt algorithm.[18] Once $\mathbf{b}$ is estimated, we can compute the intrinsic parameters as

$$f_x = f_y = \sqrt{\frac{\lambda}{B_{11}}}, \tag{17}$$

$$u_0 = v_0 = -\frac{B_{13} f_x^2}{\lambda}, \tag{18}$$

$$k = 0, \tag{19}$$

with $\lambda = B_{33} - 2B_{13}^2/B_{11}$. Then, the camera matrix $\mathbf{M}$ can be constructed according to Eq. (2). Once $\mathbf{M}$ is known, the extrinsic parameters for each observation can be readily computed as

$$\mathbf{r}_1 = \lambda' \mathbf{M}^{-1} \mathbf{h}_1, \tag{20}$$

$$\mathbf{r}_2 = \lambda' \mathbf{M}^{-1} \mathbf{h}_2, \tag{21}$$

$$\mathbf{r}_3 = \mathbf{r}_1 \times \mathbf{r}_2, \tag{22}$$

$$\mathbf{t} = \lambda' \mathbf{M}^{-1} \mathbf{h}_3, \tag{23}$$

with $\lambda' = 1/\|\mathbf{M}^{-1}\mathbf{h}_1\| = 1/\|\mathbf{M}^{-1}\mathbf{h}_2\|$.

4.2 *Nonlinear Optimization*

After the initial estimation of the intrinsic and extrinsic parameters has been worked out through the closed-form solution, an optimization procedure can be applied to minimize the reprojection error in a maximum likelihood fashion.

To include the lens distortion, the projection of a feature point defined in the pattern coordinate space needs to be adjusted to reflect the geometrical distortion. Because the image captured by a line-scan camera only contains 1-D information, tangential distortion actually does not exist, so lens distortion is essentially radial distortion only. Based on literature,[9,19] although the distortion can spread over more than 20 orders of magnitude, it is dominated by the first order. The nonlinear optimization procedure follows Zhang's method[10] by applying Levenberg–Marquardt algorithm to minimize the reprojection error.

5 Experiments and Discussions

The proposed calibration method is implemented within MATLAB® and has been tested on both synthetic data and real images. The synthetic data are used to test the performance of calibration considering two factors: (1) the tolerance to the noise in points detected from images, and (2) the effect of pattern orientations on the calibrated results. The calibration's practicability is verified by real images.

5.1 *Synthetic Data*

The accuracy and robustness of the calibration method is tested on a set of synthetic feature point correspondences. The simulated intrinsic parameters are $f_x = f_y = 1260$ pixels, $k = 0$ and $u_0 = v_0 = 1020$ pixels. The image resolution is 2048×2048 pixels. The simulated pattern is an array of 41 points evenly distributed alone a straight line. The distance between each neighboring points is set to be 60 mm, thus the point array covers a width of 2.4 m. The orientation and displacement of the pattern with respect to the camera are represented by a 3-D rotation vector $\mathbf{r}$ and a translation vector $\mathbf{t}$.

5.1.1 *Noise test*

In this test, Gaussian noise with 0 mean and σ standard deviation is added to the image points. Two observations are taken for each calibration with $\mathbf{r} = [180 \text{ deg}, 0, 0]^T$, $\mathbf{t} = [0,0,1800]^T$, and $\mathbf{r} = [180 \text{ deg}, 5 \text{ deg}, 0]^T$, $\mathbf{t} = [0,0,2000]^T$, respectively. In the first observation, the pattern plane is parallel with the imaging plane with the surface normal pointing to the reverse direction. In the second observation, the pattern plane makes a 5-deg angle around the y-axis with the imaging plane. The estimated camera parameters are compared with the ground truth. The level of noise varies from 0.1 to 1 pixel in 0.1-pixel increments. For each noise level, we conduct 60 independent trials and take the average residual errors as the results. As is shown in Figs. 4(a) and 4(b), both the reprojection error and the errors of each intrinsic parameter increase when the noise level increases. Although the reprojection error increases almost linearly, the errors of intrinsic parameters show slight fluctuations. This is the result of global optimization, which may favor some parameters over the others at certain noise levels.

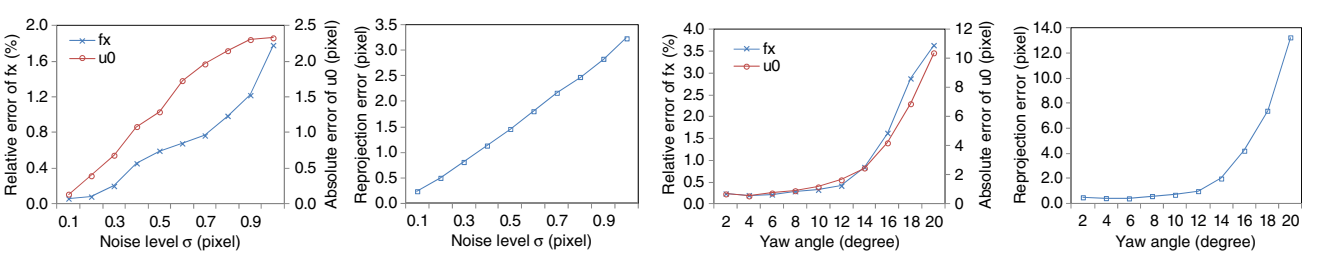

Fig. 4 Calibration errors from simulations. (a) Errors in intrinsic parameters with respect to noise; (b) Reprojection error with respect to noise; (c) Errors in intrinsic parameters with respect to rotation angle; (d) Reprojection error with respect to rotation angle.

5.1.2 *Orientation test*

In the perspective projection of a camera model, the location of an image feature point on the image plane is orientation-dependent. When generating the orthogonally reconstructed feature points with respect to the optical center, the distribution pattern of the original feature points is preserved in the reconstructed points. This simulation is to find out if a rotation in one dimension is significant to compromise the calibration results.

Two observations are taken for each calibration with $\mathbf{r} = [180\ \text{deg}, 0, 0]^T$, $\mathbf{t} = [0,0,1800]^T$, and $\mathbf{r} = [180\ \text{deg}, \theta\ \text{deg}, 0]^T$, $\mathbf{t} = [0,0,2000]^T$, respectively. The rotation angle θ varies from 2 to 20 deg at a 2-deg interval. Gaussian noise with 0 mean and 0.3 pixels standard deviation is added to image points. Sixty independent trials are conducted for each orientation angle. The results are shown in Figs. 4(c) and 4(d). It is discovered that at orientation angle about 5 deg the minimum error is observed. This can be explained that the algorithm needs the variation of depths in the feature points to accurately estimate the focal length. The greater the orientation angle, the larger the variations of feature point depth along the z-axis in the camera coordinate space. At around 5 deg the duplicated distribution of the reconstructed feature points are not significant enough to generate errors in calibration, while the depth variation in feature points are sufficient for the estimation of focal length. It can also be observed that when the rotation angle gets greater than 10 deg, the reporjection error increases almost exponentially. This indicates that rotations at this level or greater can make the calibration unstable. At 12-deg rotation, we started to observe calibration failures. At 20-deg rotation, about 65% of calibrations failed.

5.2 *Real Images*

To capture line-scanned images, a Dalsa (Waterloo, Ontario, Canada) Spyder III line-scan camera is used. The camera is mounted on a heavy-duty tripod so that the orientation of the camera can be adjusted. A Nikon (Tokyo, Japan) 20-mm-fixed focal length F-mount lens is attached to the camera through an F-mount adapter. The captured images are 2048-pixels wide. The calibration pattern is printed by a poster printer. Our pattern consists of 57 repeated units as described in Sec. 3.1, with $h = 60$ mm and $l = 220$ mm. The whole pattern covers a width of about 3.4 m. The printed pattern is then attached to two-laminated wood boxes that are connected together to create a uniformed flat surface. The camera and pattern are mounted at the same elevation, so that the camera has almost 0-pitch angle. A total of 12 calibrations are conducted. In each calibration, two pictures are taken at two-different poses. In the first pose, the camera looks at the pattern almost perpendicularly. In the second pose, a randomly selected +5- or −5-deg yaw angle is applied by either turning the camera to its right or to its left. The yaw angle does not need to be precisely controlled. The mean and standard deviations of the 12 calibrations are shown in Table 1.

In the captured picture of the calibration pattern, the image of each feature point is a vertical line, its horizontal position is calculated by averaging the left and right edges of the line. Edges are detected up to 0.1-pixel resolution using subpixel interpolation. Feature point correspondences are computed according to Sec. 3.3. The iterative process of refining the optical center takes 3 to 5 iterations to converge. In order to verify the correctness of the method, all feature points in the pattern are reprojected to the image coordinate space using the computed intrinsic and extrinsic parameters. The mean reprojection error of the 12 calibration is 0.46 pixels with standard deviation of 0.066.

Table 1 Calibration results for intrinsic parameters and reprojection errors.

	f_x	u_0	k_1	k_2	Reprojection errors
Mean	962.18	1021.46	−0.036	0.015	0.46
STD	1.913	0.494	5.41E − 04	3.25E − 04	0.066

Three-pose calibrations were also investigated in our experiments. Images used in these three-pose calibration are the perpendicular observation, +5- deg observation and −5- deg observation. Our existing 12 sets of calibration images contributed to 6 three-pose calibrations. The mean reprojection error was 0.81 pixels with standard deviation of 0.064, compared with 0.46 pixels in two-pose calibration. This suggested a degradation in calibration when including more angled observations. The source of the degraded performance can be traced to the error in 1-D to 2-D conversion in constructing feature point correspondences, in which angled observation introduces more error than perpendicular observation. Adding a new pose into the calibration with more errors in feature point correspondences ultimately decreases the accuracy of the results, thus two observations with sufficient feature points in each are preferred in our calibration method.

To find out if the lens distortion coefficients are effective in correcting the radial distortion in the images, we performed distortion correction with obtained calibration results. Figure 5 shows the results of the correction. The image was captured with the camera looking at our calibration pattern in a perpendicular orientation. Each vertical line in the image indicates a feature point on the pattern seen by the camera. If a perfect lens is considered and radial distortion is zero, the horizontal spacing between each neighboring line segments found in the middle of the image should be equal to the ones found on the sides of the image. The top part of Fig. 5 shows the original line-scanned image with radial distortion, from which the spaces between line segments on the sides are narrowed. The bottom part of Fig. 5 shows the result after distortion is removed. All the line segments can be perfectly aligned, regardless where in the image they are sampled.

To further verify the quality of distortion correction, we looked at the space within each repeated pattern unit, referred to as the unit length. Figure 6 shows the unit lengths measured before and after distortion correction. Figure 6(a) displays the correction results from a picture that was taken at a perpendicular camera orientation, whereas Fig. 6(b) displays the results from a picture taken with a yaw angle. It is clear that radial distortion exists in both pictures. The unit lengths in the middle of the pictures, where radial distortion is less, are quite consistent before and after correction. The effectiveness of the correction is evaluated by the linearity of the plots of unit lengths with respect to their locations on

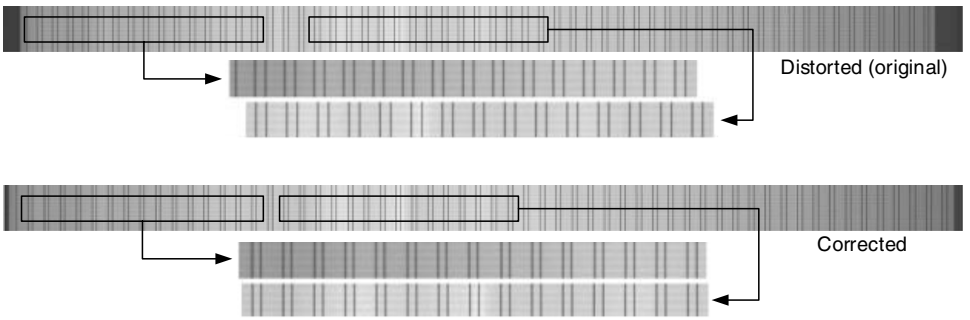

Fig. 5 Examples of distortion correction on pattern images. Top part: original image with lens distortion. Pattern lines are unevenly spaced; bottom part: distortion corrected. Pattern lines are evenly spaced and can be aligned.

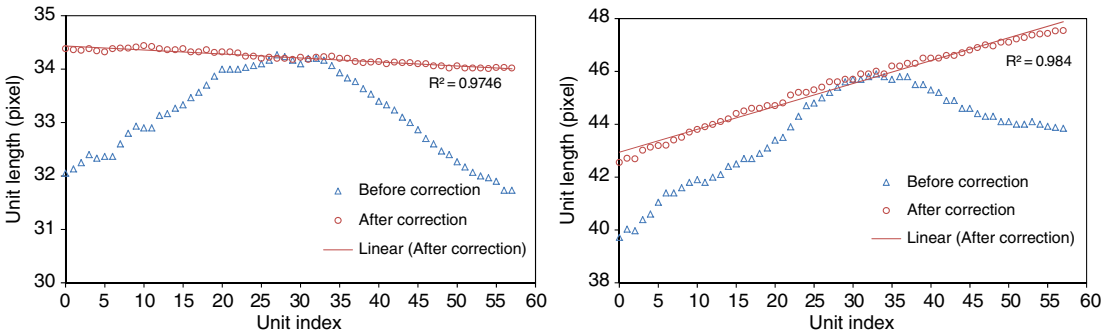

Fig. 6 Plots of unit lengths before and after radial distortion correction. (a) Measurements taken from the observation at a perpendicular pose; (b) measurements taken from the observation with yaw angle.

the image. We fitted a line to the graph after correction. The R^2 values of the two observations are found to be 0.975 and 0.985, respectively, which indicate good linearity for both cases.

This calibration technique was applied to one of our pavement imaging projects for dual-camera registration. Two line-scan cameras were used to capture pavement images simultaneously but configured in different exposure settings. The paired images captured by these cameras were fused to generate images free from shadows of roadside objects, such as trees and traffic signs. The line-scan cameras were mounted on one piece of metal plate and were next to each other. Therefore, the image planes of the two cameras were coplanar and the image registration only involved shifts and rotations once the lens distortion was corrected. Figure 7 shows the results of the image fusion with and without the lens distortion correction. The image fusion from uncorrected images shows misalignment, because the lens distortion is nonlinear and the images cannot be registered correctly. A perfect alignment was achieved when lens distortion was removed by applying the camera calibration method presented in this article. The distortion free image ensures correct measurement of pavement distress and features, such as the length of a crack, as well as accurate feature localization.

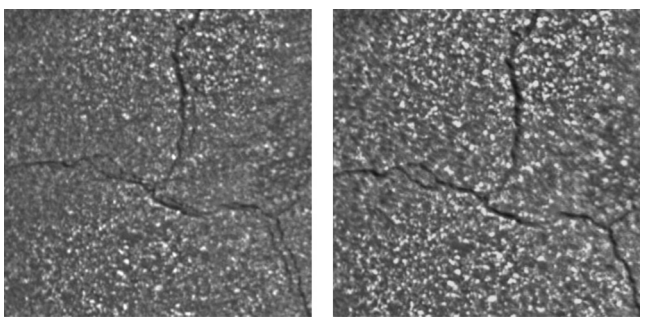

Fig. 7 Fusion results from images captured by two identical line-scan cameras with two different exposure settings. (a) Fused from images without lens distortion correction; (b) lens distortion corrected.

5.3 *Calibration Errors*

As observed in the calibrations with real images, the mean reprojection error we have reached is 0.46 pixels, which indicate that there are other sources that may affect the performance of the calibration method. Here, we briefly review some of the most significant error sources.

5.3.1 *1-D to 2-D conversion*

We superpose the 1-D calibration problem into 2-D scope and treat the line-scan camera as a 2-D matrix camera. The calibration problem is solved by an extensively studied camera model with added constraints derived from the unique geometry of a 1-D camera. The feature point correspondences are handled so that the coordinates on the second dimension are exactly duplicated from the first dimension. As a result, the calculated intrinsic parameters are symmetric, i.e., $f_x = f_y$ and $u_0 = v_0$. In reality, it is common to have the same focal lengths in both x- and y-directions for most of the lenses and cameras with square pixels. However, the optical center is usually not symmetric, because these values depend on the relative position between the photonic sensor and the

lens. Given only the 1-D locations of feature points in the image, it is technically difficult to estimate the optical center in 2-D. Since the magnitude of radial distortion is related to the distance from a current image point to the optical center, with one dimension missing, the amount of distortion cannot be accurately estimated. Although our experiments with real images show that the plot of unit lengths in distortion-corrected images is quite linear, this is the result of small vertical offset between the optical center and the image center. The radial distortion along the horizontal direction is largely determined by the horizontal location of each feature point.

5.3.2 *Radial distortion*

It is reasonable to accept that the camera sensor array is aligned in a straight line. However, one can argue that the view line is not necessarily a straight line when the optical center is not on the sensor array and lens distortion is present. In fact, the actual view line may be slightly curved. Since the distance between each of two parallel lines is small (60 mm in our experiments), we consider the view line segment within each repeated pattern unit to be a straight line, and estimate the feature point coordinates individually for each repeated pattern unit. Furthermore, this strategy is based on the assumption that the amount of distortion is uniform within each pattern unit. This could also introduce error in the calculation of cross ratio within each pattern unit, and ultimately error in the estimated feature point location.

5.3.3 *Calibration target*

It is essential that the coordinates of feature points on the pattern are known with good precision so that their errors are not observable from the image. If this condition is not met, the feature point correspondences are not accurate, and the estimated camera parameters become inaccurate. It is commonly accepted that the relative accuracy in the pattern coordinate space should be better than the accuracy in the image coordinate space.[11] For example, in order to achieve an accuracy of 0.1 pixels as we did in feature point detection from captured pictures, the locations of a feature point in pattern coordinate space should be known to a resolution better than 0.16 mm, given the size of the pattern being about 3.4 m in lateral direction. This is still challenging because the accuracy can be affected by many factors, such as the precision of the printer, the flatness of the pattern mounting surface, any distortion on the printed paper due to the stretch when attaching the paper to the mounting surface, etc.

6 Conclusion

In this article, we have presented a line-scan calibration method that is aimed to be applicable for implementation outside a laboratory setting. A planar pattern consisting of repeated parallel lines is used in the calibration. This pattern allows the estimation of feature points in the pattern coordinate space without mechanical scanning, thus image capturing can be done in a static configuration. A typical calibration can be performed by taking two pictures of the pattern, with one being a perpendicular observation of the pattern and the other at a small angle (5 deg). The calibration method is tested on both synthetic data and real images. The average reprojection error in calibrating real images is 0.46 pixels. Our results also show that the distribution of unit length in a distortion-corrected image is linear along the width of the image. This indicates the distortion coefficients together with the optical center are effective in correcting the radial distortion in line-scanned images.

References

1. K. H. McGhee, "Automated pavement distress collection techniques," Chapter 2 in *Transportation Research Board, National Research*, Vol. 334 (2004).
2. X. Yao, M. Yao, and B. Xu, "Automated measurements of road cracks using line-scan imaging," *J. Test. Eval.* **39**(4), 621–629 (2011).
3. K. C. P. Wang et al., "Automated imaging technique for runway condition survey," Technical Report (2012).
4. C. Wang, A. Sha, and Z. Sun, "Pavement core sample image acquisition system based on line-scan camera," in *Proc. 2010 IEEE 3rd Int. Congr. Image Signal Process. (CISP)*, Vol. 6, pp. 2614–2618, IEEE (2010).
5. R. S. Petty, M. Robinson, and J. P. O. Evans, "3D measurement using rotating line-scan sensors," *Meas. Sci. Technol.* **9**(3), 339–346 (1998).
6. Z. Zhao and G. Wen, "Model-based estimation of axisymmetric target's pose and speed using line-scan camera," in *International Conference on System Science, Engineering Design and Manufacturing Informatization (ICSEM), 2011*, Vol. 2, pp. 31–34, IEEE, Guiyang, China (2011).
7. A. Hirahara and K. Ikeuchi, "Detection of street-parking vehicles using line scan camera and scanning laser range sensor," in *Proc. 2004 IEEE Intel. Vechicles Symp.*, pp. 656–661, IEEE (2004).
8. M.-S. Lim and J. Lim, "Visual measurement of pile movements for the foundation work using a high-speed line-scan camera," *Pattern Recognit.* **41**(6), 2025–2033 (2008).
9. 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. Robot. Autom.* **3**(4), 323–344 (1987).
10. Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(11), 1330–1334 (2000).
11. J. Heikkila, "Geometric camera calibration using circular control points," *IEEE Trans. Pattern Anal. Mach. Intell.* **22**(10), 1066–1077 (2000).
12. R. Horaud, R. Mohr, and B. Lorecki, "On single-scanline camera calibration," *IEEE J. Robot. Autom.* **9**(1), 71–75 (1993).
13. C. A. Luna et al., "Calibration of line-scan camera," *IEEE Trans. Instrum. Meas.* **59**(8), 2185–2190 (2010).
14. B. Hui et al., "Line-scan camera calibration in close-range photogrammetry," *Opt. Eng.* **51**(5), 053602 (2012).
15. J. Drareni, S. Roy, and P. Sturm, "Plane-based calibration for linear cameras," *Int. J. Comput. Vision* **91**(2), 145–156 (2011).
16. J. N. Cederberg, *A Course in Modern Geometries*, 2nd ed., Springer, New York, NY (1995).
17. R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, 2nd ed., Cambridge University Press, Cambridge, UK (2004).
18. P. E. Gill, W. Murray, and M. H. Wright, "The Levenberg-Margqardt method," in *Pratical Optimization*, Emerald Group Publishing Limited, London (1981).
19. J. Weng, P. Cohen, and M. Herniou, "Camera calibration with distortion models and accuracy evaluation," *IEEE Trans. Pattern Anal. Mach. Intell.* **14**(10), 965–980 (1992).

Ming Yao is a PhD candidate at the University of Texas at Austin. He received his BS degree in communication engineering, and his MS degree in pattern recognition and intelligent systems from Donghua University, Shanghai, China, in 2003 and 2006, respectively. He was a research fellow in the School of Human Ecology, University of Texas at Austin, from 2006 to 2008. He was then enrolled in the PhD program in biomedical engineering at the University of Texas at Austin. His research interests include image processing, computer vision, and computer graphics.

Zuyun Zhao received her BS and MS degrees in communication engineering from Donghua University, Shanghai, China, in 2008 and 2010, respectively. Currently, she is a graduate student in the Department of Material Science and Engineering, University of Texas at Austin. Her research interests include digital imaging and image processing.

Bugao Xu received his PhD in 1992 from the University of Maryland at College Park and joined the University of Texas faculty in 1993. Now, he is a full professor in the School of Human Ecology at the College of Natural Science. He is also on the graduate faculty of Biomedical Engineering, and Material Science and Engineering at UT Austin. His investigative work composes multidisciplinary areas including textiles, biometrics, and transportation.