Stereo line-scan sensor calibration for 3D shape measurement

5 authors, including:

Bo Sun
Tianjin University
8 PUBLICATIONS **98** CITATIONS

SEE PROFILE

Yang Linghui
Tianjin University
87 PUBLICATIONS **364** CITATIONS

SEE PROFILE

Jiarui Lin
Tianjin University
86 PUBLICATIONS **349** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project   Dynamic Error Analysis and Compensation of Multi-station Measurement Network for Space GNC Simulation View project

Project   Novel Method for Dynamic Measurement of Large-scale 3-D Shape Based on 1-D Image sequence View project

applied optics

Stereo line-scan sensor calibration for 3D shape measurement

Bo Sun, 🔟 Jigui Zhu, Linghui Yang,* Yin Guo, and Jiarui Lin

State Key Laboratory of Precision Measuring Technology and Instruments, Tianjin University, Tianjin 300072, China
*Corresponding author: icelinker@tju.edu.cn

The stereo line-scan sensor opens up new potentialities for 3D measurement owing to the ultra-high resolution and acquisition rate. Calibration is a crucial key technology for a stereo line-scan sensor. This paper presents a precise calibration method for the stereo line-scan sensor. Several reference points are installed onto the sensor's body as an intermediary. The calibration turns into a two-step process: calibrating the cameras in the laboratory prior to measurement and locating the sensor in an actual measurement field. A mobile apparatus that comprises a planar pattern and extra reference points is designed. By incorporating the apparatus in combination with an auxiliary instrument, an optimal calibration configuration is created by placing the apparatus into multiple positions. A robust algorithm is proposed to enhance the stability of the parameter estimation. The quality of the calibration method is experimentally tested, and the performance is further investigated. Experimental results demonstrate that the proposed method offers a practical solution to calibrate a stereo line-scan sensor for 3D shape measurement.   © 2017 Optical Society of America

OCIS codes: (120.0120) Instrumentation, measurement, and metrology; (150.3045) Industrial optical metrology; (150.1488) Calibration; (330.1400) Vision - binocular and stereopsis; (150.6910) Three-dimensional sensing.

https://doi.org/10.1364/AO.56.007905

1. INTRODUCTION

The field of 3D shape measurement has been studied extensively for many decades. Numerous methods based on matrix cameras (2D imaging sensor) have been developed, which can be categorized into point-wise structured light (usually in combination with the flying dot technology using a rapidly rotating mirror [1]), structured laser stripe [2–4], and fringe projection [5–9]. Although these techniques have been widely used in manufacturing industries, they still have limitations and are therefore experiencing challenges, especially in high-speed applications. The major limitation of the point-wise structured light technique is its low-efficiency scanning mode where the laser beam must scan over the surface point by point. Similarly, the scanning speed of the structured laser stripe technique is constrained by the camera's frame rate and the image processing speed. Comparatively, the fringe projection technique can provide full-field 3D coordinates with a relatively high scanning throughput for each individual measurement, but still has drawbacks. On one hand, when the size of the measured object cannot be covered by a single capture stage, the surface must be measured in parts from multiple standpoints, leading to overlap effects, dead zones, and low efficiency. On the other hand, when the measured object is moving, the result can hardly be truly continuous because seamless image acquisition using a matrix camera in this case requires ultimate synchronization.

In contrast, the innovative line-scan technology has characteristics of superior high resolution and line rate (up to 16384 pixels at 27 kHz), and freedom of movement during image acquisition. Measurement can be efficiently accomplished in a motion process. Because of that, line-scan technology is one of the most promising techniques for fast measurement of large-scale objects and for continuous measurement of fast-moving objects. There are therefore increasing stereo line-scan sensors for 3D shape measurement more recently [10–14]. Compared with the point-wise structured light technique and structured laser stripe technique, a stereo line-scan sensor has higher resolution and sample rate. High-density point clouds can be produced without compromising on resolution and data rate. Compared with the fringe projection technique, a stereo line-scan sensor has advantages of seamless acquisition, greater and easily extendable measurement volume, freedom of movement during measurement, low cost, and simple structure.

Calibration is crucial for a stereo line-scan sensor because errors in calibration will directly affect the quality of shape measurement. The objective of calibration is to determine all geometric parameters of the complete sensor system, i.e., the interior and exterior orientation parameters of both line-scan cameras. Over the past few decades, camera calibration has represented the combination and competition of topics on computer vision, machine vision, and photogrammetry.

Unfortunately, the well-known calibration approaches for matrix cameras (e.g., the plane-based calibration proposed by Zhang [15]) cannot be used for calibrating a stereo line-scan sensor. The reason lies in that the imaging of line-scan cameras is not as intuitive as that of matrix cameras. More specifically, it is impossible to directly establish a correspondence between a specific space point and a captured image point in a captured one-dimensional image.

Up to now, several calibration methods for line-scan cameras have become available. The most commonly used method is based on the cross-ratio invariance principle, initiatively proposed by Horaud in 1993 [16]. In this proposal, a planar calibration apparatus with feature lines is used. The apparatus has to translate with exact distances in the normal direction. The calibration quality depends on the accuracy of the displacements. To overcome this limitation, modified methods using 3D calibration apparatuses are developed, such as using two parallel planes placed with some displacement [17,18], or using two orthogonal planes hinged together [19]. Such configurations remove the necessity to exactly move the apparatuses during calibration because a single capture is enough to perform the calibration. Although these proposals can be used for calibrating a stereo line-scan sensor, there exist several limitations. First, the apparatus provides limited calibration points and scarce depth range, leading to lack of geometric information, causing the mathematical solution to be easily perturbed. Although increasing the dimension is partly helpful, a larger apparatus is more expensive to fabricate, more sensitive to the environment (e.g., ambient temperature), harder to transport, and more inconvenient to use. In these approaches, the two cameras have to be calibrated simultaneously. Otherwise, the spatial relationship between the cameras cannot be established. Thus the used apparatus's dimension cannot exceed the overlapped field of view (FOV) of the stereo line-scan sensor. Second, another problem with the above mentioned methods is that the obtained parameters are with respect to the apparatus coordinate frame. In general, a relative movement between the stereo line-scan sensor and the measured object is essential to obtain the entire surface by successive acquisition of multiple scans. The individual scans are merged into a unified coordinate frame based on the movement. Therefore, the exterior orientation parameters should be able to be easily transformed to other coordinate frames. Unfortunately, it is difficult for these calibration methods to achieve. Apart from these methods, several scanning calibration methods are also reported [20–22]. A linear motion for the apparatus is required. The captured 2D images have recognizable characteristic points. Thus, these techniques have the advantage of easy determination of the correspondences between the points on the apparatus and their projected image points. However, in addition to the similar disadvantage of a limited apparatus's dimension, the calibration quality relies on the precision of the uniform linear motion.

To overcome the limitations of current approaches in calibrating a stereo line-scan sensor, a practical calibration method is proposed in this paper. By attaching several reference points to the stereo line-scan sensor, the calibration turns into a two-step process: the first step is to calibrate the line-scan cameras in the laboratory prior to measurement, where time

and configuration condition permits more precise calibration; and the second step is to locate the sensor in the world coordinate frame of an actual measurement field, with the reference points used as common points. In addition, a hybrid calibration apparatus is designed. By incorporating an ancillary measurement system, either the sensor or the apparatus could be freely moved during calibration. The cameras are permitted to be calibrated individually. The amount and distribution of the calibration points are unrestricted, which contributes to the reliability and robustness of the parameter estimation. Since all the imaging configurations can be created by placing the apparatus in multiple positions with respect to the stereo line-scan sensor, it is also applicable for extremely large FOV applications. The dimension of the apparatus can be small, which brings benefits of low cost, easy to fabricate, and convenient to use. Furthermore, a robust parameter estimation algorithm is proposed, which is less prone to noises.

The remainder of the paper is organized as follows: Section 2 introduces the structure and principle of the stereo line-scan sensor briefly. Section 3 first describes the calibration problem and presents the proposed calibration approach subsequently. The experiments and results are reported in Section 4. Finally, conclusions are summarized in Section 5.

2. STEREO LINE-SCAN SENSOR

As depicted in Fig. 1(a), the stereo line-scan sensor is composed of dual line-scan cameras along with a projector. The viewing planes of the cameras are aligned coplanar. In addition, four distinct sphere mounts are rigidly installed on the sensor, as illustrated in Fig. 1(b). The sphere mounts are compatible with
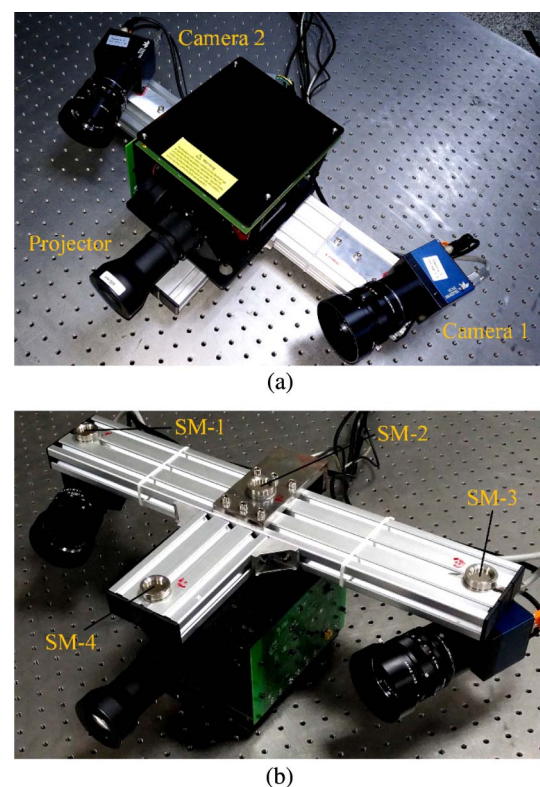

(a)

(b)

Fig. 1. Stereo line-scan sensor. (a) Top view. (b) Bottom view.

the cooperative targets of several standard portable measuring systems, such as a laser tracker [23] or an optical coordinate measuring machine (CMM) [24]. The sensor coordinate frame is defined as follows: the origin is located at SM-1, the x axis is given by the vector from SM-1 to SM-3, the z axis is perpendicular to the plane passing through SM-1, SM-3, and SM-4, and the y axis completes the right-handed orthogonal set. Compared with defining the stereo sensor's coordinate frame coinciding with either camera's coordinate frame, it is more explicit and more convenient to get in touch with other systems or coordinate frames.

The projector projects coded patterns onto the object's surface to enhance the texture. Single-shot patterns are preferred [13]. In this study, a static pseudo random binary pattern is adopted. The overlapped FOV intersects the pattern at a stripe that represents the measured surface. By capturing the image pair of the stripe and establishing the correspondences, the 3D coordinates of all points (all pixels to the limit) in the stripe can be calculated by a triangulation [13]:

$$\begin{cases} r_{11}(^{S}X - X_0) + r_{12}(^{S}Y - Y_0) + r_{13}(^{S}Z - Z_0) = 0 \\ v = v_c + \Delta v + F_y \frac{r_{21}(^{S}X-X_0)+r_{22}(^{S}Y-Y_0)+r_{23}(^{S}Z-Z_0)}{r_{31}(^{S}X-X_0)+r_{32}(^{S}Y-Y_0)+r_{33}(^{S}Z-Z_0)} \\ r'_{11}(^{S}X - X'_0) + r'_{12}(^{S}Y - Y'_0) + r'_{13}(^{S}Z - Z'_0) = 0 \\ v' = v'_c + \Delta v' + F'_y \frac{r'_{21}(^{S}X-X'_0)+r'_{22}(^{S}Y-Y'_0)+r'_{23}(^{S}Z-Z'_0)}{r'_{31}(^{S}X-X'_0)+r'_{32}(^{S}Y-Y'_0)+r'_{33}(^{S}Z-Z'_0)} \end{cases} \quad \textbf{(1)}$$

Here, the first and third equations parameterize the viewing planes of both cameras. The other two equations describe the projected rays. The coefficients v_c and F_y are, respectively, the principal point and the focal length, r_{ij} are the entries of the rotation matrix from the camera to the sensor coordinate frame, X_0, Y_0, and Z_0 are the 3D coordinates of the perspective center, and Δv is the lens distortion. The corresponding parameters with superscript apostrophes belong to the second camera. The superscript S means the 3D coordinates are expressed in the sensor coordinate frame. Only two-order radial distortions are taken into account, which are the dominant factors that influence the accuracy performance of line-scan imaging [25]. The distortion is given by

$$\Delta v = k_1(v - v_c)^3 + k_2(v - v_c)^5. \quad \textbf{(2)}$$

Typically, k_2 is needed only for wide-angle lenses.

The stereo line-scan sensor behaves as a local line scanner that can provide a stripe of surface profile on each capture. The coordinates of the points on each scan profile are obtained by solving Eq. (1). To reach the full range of the measured object, either the sensor or the object must move along a scanning trajectory. Subsequently, the local 3D point clouds, scanned from individual captures, are merged to build up a complete model according to the movement. In the case of a moving object and stationary sensor [11–13], the sensor should first be located in the world coordinate frame. Then the individual scans are merged according to the object's movement (in general, the movement is known or measurable). Alternatively, in the case of a moving sensor and stationary object [14], the sensor's position and orientation in the world coordinate frame should be measured in real time. Subsequently, the individual captured stripes are merged according to the sensor's scanning trajectory.

3. CALIBRATION APPROACH

Based on the surface reconstruction principle of the stereo line-scan sensor, the calibration problem involves two steps. The first step is obtaining the unknown parameters (v_c, F_y, r_{11}, r_{12}, ..., r_{33}, X_0, Y_0, Z_0, k_1, k_2 and v'_c, F'_y, r'_{11}, r'_{12}, ..., r'_{33}, X'_0, Y'_0, Z'_0, k'_1, k'_2) in Eq. (1). Owing to the structural configuration of the sensor, i.e., the mechanical installation as well as the focus and aperture settings, these parameters always stay constant. Therefore, this step is allowed to be carried out prior in the laboratory where various auxiliary instruments can be used to facilitate calibration. The second step is locating the sensor in the world coordinate frame of an actual measurement field.

The calibration approach is presented in the following subsections. Subsection 3.A describes the structure and characteristics of the calibration apparatus. Subsections 3.B and 3.C present the approach of calibrating a separate camera with respect to the sensor coordinate frame. Subsection 3.D outlines establishing the spatial relationship between the sensor and the world. At the end, the whole calibration process is summarized.

A. Calibration Apparatus

Obtaining the 3D points that correspond to the captured image points is in general a key problem for calibration of a line-scan camera. In theory, a line-scan camera can be calibrated as long as the calibration points reside in two different depths. However, in order to achieve precise calibration, the distribution of the calibration points must satisfy certain preferable criteria. The first criterion is that a comparatively large number of calibration points with a suitable distribution should be offered. The accuracy and robustness of parameter estimation will be increased because each point contributes to constraining the parameters. The second criterion is that the calibration points should cover the entire FOV so that the distortion parameters can be estimated more accurately [26]. For these purposes, a hybrid calibration apparatus is designed. By incorporating an ancillary instrument, either the camera or the apparatus could be freely moved during calibration to obtain sufficient calibration points. Furthermore, benefitting from the reference points on both the sensor and the apparatus, it is possible to calibrate each camera individually, without having to place the apparatus being sighted by both cameras simultaneously.

As shown in Fig. 2, the calibration apparatus comprises a photolithographic pattern and four distinct sphere mounts attached on the corners. The planar pattern is printed on a transparent glass board. By applying backlight, the pattern can be easily separated from the background. Inside the pattern, there exist nine straight lines and a grid of ring markers. Five lines are vertical and the other four are oblique, connecting each pair of vertical lines. The apparatus coordinate frame is defined where the origin is located at the bottom-left marker; the x axis is given by the vector from the origin to the top-left marker, the z axis is perpendicular to the pattern plane, and the y axis completes the right-handed orthogonal set. The coordinates of the markers and the equations of the lines are manually designed to be known in the apparatus coordinate frame. In addition, the coordinates of the sphere mounts are calibrated in advance, which are used as common reference points for establishing the spatial relationship between the apparatus and the sensor.
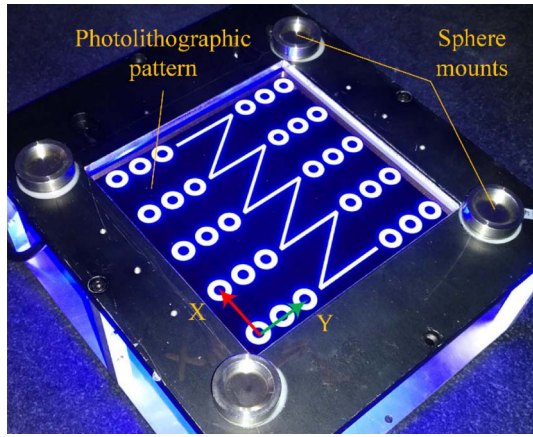
Fig. 2. Structure and frame definition of the calibration apparatus.

B. Obtaining 3D Coordinates of the Calibration Points

The FOV of a line-scan camera is a plane that is also called the viewing plane. When the camera is stationary, only the objects within the viewing plane project onto the image sensor array. Therefore, the intersection points of the viewing plane with the straight lines on the pattern are taken for calibration. The centers of the projected image points can be detected. However, it is difficult to distinguish which points match the captured image points exactly, because the captured image is a one-dimensional line without recognizable information (e.g., circles or linear grids). In this subsection, the corresponding 3D points are located indirectly according to the pattern geometry and the cross ratio invariant principle.

As depicted in Fig. 3, the cross-ratio of the 4-tuple of image points (p_1, p_2, p_3, p_5) is given by

$$(v_1, v_2; v_3, v_5) = \frac{(v_1 - v_3)(v_2 - v_5)}{(v_2 - v_3)(v_1 - v_5)}, \tag{3}$$

where v_1, v_2, v_3, and v_5 are the pixel coordinates of the image points. $(v_1, v_2; v_3, v_5)$ describes their cross ratio. On the basis of the cross-ratio invariability, the cross ratio of the corresponding
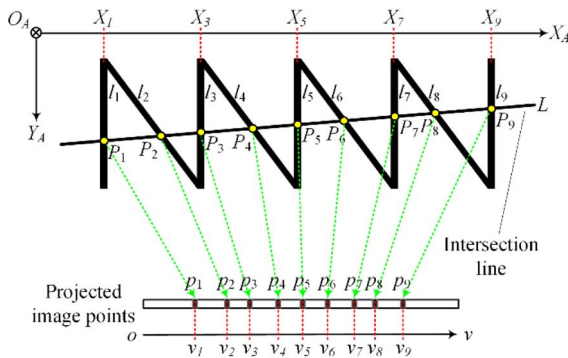

Fig. 3. Imaging of the pattern. l_1 to l_9 are the feature lines with known equations with respect to the apparatus coordinate frame. L is the intersection line between the viewing plane and the pattern plane. P_1 to P_9 are the intersection points that are used for calibration. p_1 to p_9 are the homologous projected image points. v_1 to v_9 are the corresponding pixel coordinates.

intersection points P_1, P_2, P_3, and P_5 has the same value. Similar to Eq. (3), the cross-ratio invariability becomes

$$(X_1, X_2; X_3, X_5) = \frac{(X_1 - X_3)(X_2 - X_5)}{(X_2 - X_3)(X_1 - X_5)} = (v_1, v_2; v_3, v_5), \tag{4}$$

where X_1, X_2, X_3, and X_5 are X coordinates of the intersection points. $(X_1, X_2; X_3, X_5)$ describes the cross ratio in x axis. In Eq. (4), v_1, v_2, v_3, and v_5 are measured from the image with sub-pixel accuracy. In addition, the x coordinates (X_1, X_3, X_5) can be obtained according to the pattern geometry. The only unknown coordinate X_2 can be calculated by combining Eqs. (3) and (4):

$$X_2 = \frac{X_5 - \rho X_3}{1 - \rho}, \tag{5}$$

where $\rho = \frac{(v_1 - v_3)(v_2 - v_5)(X_1 - X_5)}{(v_2 - v_3)(v_1 - v_5)(X_1 - X_3)}$.

According to the pattern geometry, the equation of l_2 in the x-y plane is known, given by

$$y = a_2 x + b_2. \tag{6}$$

Substituting Eq. (5) into Eq. (6), the complete coordinates of P_2 are achieved, expressed as (X_2, Y_2). Theoretically, the calculation of P_2 can be implemented by using arbitrary three vertical lines. To better neutralize the distortion effect, the nearest ones are employed. The calculation procedures for P_4, P_6, and P_8 are similar.

Given these 4-tuple points (P_2, P_4, P_6, and P_8), the intersection line L is fitted, whose parameters (a_L, b_L) are estimated via a least squares method. Since the x coordinates of the intersection points (P_1, P_3, P_5, P_7, and P_9) are known from the pattern geometry, the y coordinates are calculated by

$$\begin{cases} Y_1 = a_L X_1 + b_L \\ Y_3 = a_L X_3 + b_L \\ Y_5 = a_L X_5 + b_L. \\ Y_7 = a_L X_7 + b_L \\ Y_9 = a_L X_9 + b_L \end{cases} \tag{7}$$

So far, the 3D coordinates of all intersection points with respect to the apparatus coordinate frame are known because z coordinate is always zero.

Nine collinear calibration points are obtained for each capture. Two captures are the minimum condition mandatory for calibration. As mentioned previously, in order to ensure good ray intersections and filling the image format, sufficient images should be taken, either by placing the apparatus at multiple positions or by moving the sensor to multiple standpoints. Moving which one depends on the experimental conditions, e.g., for a sensor mounted on a robot-end, moving the sensor is preferable. In this paper, the sensor is fixed on an optical table, thus moving the apparatus is advisable. As shown in Fig. 4, the apparatus is placed at N positions for each camera. All images are flexibly captured with the optimal configurations (e.g., imaging distance and pose) for both cameras. A total of $9 \times N$ calibration points are obtained. Their coordinates are expressed as
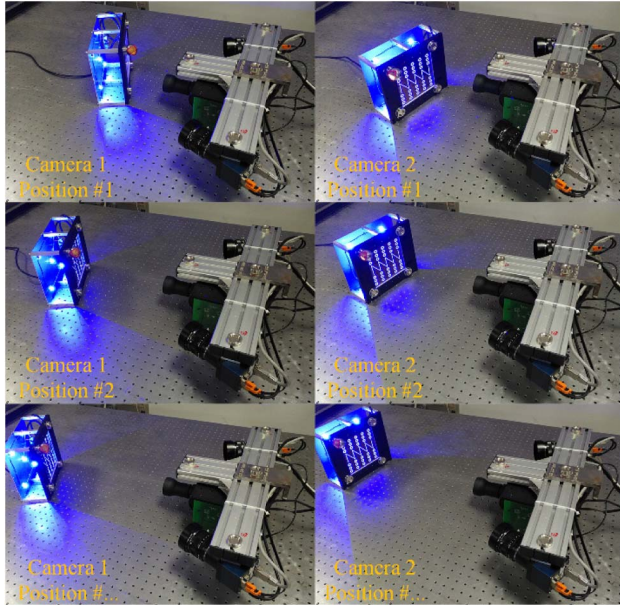
Fig. 4. Calibration by placing the apparatus at N positions for each camera. A positioning system in combination with a cooperative target is used to align the calibration points to the sensor coordinate frame.

$$\{{}^A\boldsymbol{P}_{m,n} = ({}^A X_{m,n}, {}^A Y_{m,n}, 0)^T \,|\, m = 1, 2, 3 \ldots 9; n = 1, 2, 3, \ldots N\},$$
(8)

where ${}^A\boldsymbol{P}_{m,n}$ denotes the intersection point on the mth line at the nth position, expressed in the apparatus coordinate frame.

In order to align all the points into the sensor coordinate frame, an auxiliary positioning system, such as a laser tracker, an optical CMM, or an ordinary dual-camera system, is employed to monitor the position and orientation of the apparatus. At each position, a cooperative target is mounted into the sphere mounts in sequence (see Fig. 4). The positioning system measures the coordinates of the cooperative target so that the coordinate transformation from the apparatus to the positioning system is determined. In a similar way, the transformation from the positioning system to the sensor is determined by measuring the reference points on the sensor body. Taking the positioning system coordinate frame as an intermediary, the rotation matrix and translation vector from the apparatus coordinate frame to the sensor coordinate frame are calculated as follows:

$$\begin{cases} {}^S_A\boldsymbol{R}_n = {}^S_P\boldsymbol{R}_n {}^P_A\boldsymbol{R}_n \\ {}^S_A\boldsymbol{T}_n = {}^S_P\boldsymbol{R}_n {}^P_A\boldsymbol{T}_n + {}^S_P\boldsymbol{T}_n \end{cases} \quad (n = 1, 2, 3, \ldots N),$$
(9)

where a subscript for a transformation represents the source frame, and the corresponding superscript represents the

destination frame. The letters "A", "P", and "S" represent coordinate frames of the apparatus, positioning system, and the sensor frames, respectively. The subscript "n" means the nth position.

Finally, all the calibration points are transformed to the sensor coordinate frame according to Eqs. (8) and (9):

$${}^S\boldsymbol{P}_{m,n} = ({}^S X_{m,n}, {}^S Y_{m,n}, {}^S Z_{m,n})^T = {}^S_A\boldsymbol{R}_n {}^A\boldsymbol{P}_{m,n} + {}^S_A\boldsymbol{T}_n, \quad \textbf{(10)}$$

where ${}^S\boldsymbol{P}_{m,n}$ denotes the intersection point on the mth line at the nth position, expressed in the sensor coordinate frame.

C. Parameter Estimation

In essence, the imaging model of a line-scan camera consists of two parts: the first part is the constraint of the viewing plane, and the second part is the perspective projection. The pending parameters are described by a vector $\boldsymbol{\varphi} = (v_c, F_y, \alpha, \beta, \gamma, X_0, Y_0, Z_0, k_1, k_2)^T$, where the coefficients α, β, and γ parameterize the rotation matrix according to Rodrigues rotation formula [15]. In this subsection, the 10 unknown parameters of a line-scan camera are estimated using a two-step algorithm. The first step is obtaining an initial guess, and the second step is refining the parameters using a nonlinear optimization.

1. Initial Guess

Given $9 \times N$ calibration points and their corresponding image points, an initial guess can be obtained by a Lagrange multiplier algorithm in Ref. [27]. Experimentally, the Lagrange multiplier algorithm is occasionally not robust enough. The result may be perturbed by noises owing to the numerical instability, particularly in v_c and F_y. In the worst-case scenario, the extracted parameters considerably diverge from real experimental configuration, resulting in a slow convergence or even failure of the optimization algorithm. In effect, the optimization algorithm converges to the global minimum only if the initial guess is already close to the final solution. Therefore, it is crucial to provide a reasonable initial guess for the nonlinear optimization algorithm. To enhance the robustness of the initial guess estimation, a simplified optimization procedure is first formulated to improve the initial guess, where the distortion parameters are purposely omitted.

The parameters are determined by minimizing the reprojection errors, which are the differences between the reprojected image points and the observed values. In accordance with Eqs. (1) and (10), each reprojected image point $(\hat{u}_{m,n}, \hat{v}_{m,n})$ can be described as a function of the camera parameters and its corresponding 3D point:

$$\begin{cases} \hat{u}_{m,n}(\boldsymbol{\varphi}_0, {}^S\boldsymbol{P}_{m,n}) = r_{11}({}^S X_{m,n} - X_0) + r_{12}({}^S Y_{m,n} - Y_0) + r_{13}({}^S Z_{m,n} - Z_0) \\ \hat{v}_{m,n}(\boldsymbol{\varphi}_0, {}^S\boldsymbol{P}_{m,n}) = v_c + F_y \frac{r_{21}({}^S X_{m,n} - X_0) + r_{22}({}^S Y_{m,n} - Y_0) + r_{23}({}^S Z_{m,n} - Z_0)}{r_{31}({}^S X_{m,n} - X_0) + r_{32}({}^S Y_{m,n} - Y_0) + r_{33}({}^S Z_{m,n} - Z_0)} \end{cases}, \quad \textbf{(11)}$$

where $\boldsymbol{\varphi}_0 = (v_c, F_y, \alpha, \beta, \gamma, X_0, Y_0, Z_0)^T$. The objective of the maximum likelihood estimation is to find the parameters $\boldsymbol{\varphi}_0$ so that the sum of the squares of the reprojection errors is minimized. The cost function is modeled as follows:

$$\boldsymbol{\varphi}_0^* = \min_{\boldsymbol{\varphi}_0} \sum_{n=1}^{N} \sum_{m=1}^{9} ([\hat{u}_{m,n}(\boldsymbol{\varphi}_0, {}^S\boldsymbol{P}_{m,n})]^2 + [v_{m,n} - \hat{v}_{m,n}(\boldsymbol{\varphi}_0, {}^S\boldsymbol{P}_{m,n})]^2).$$

$$(12)$$

Here, $v_{m,n}$ is the homologous detected pixel coordinate. This is a nonlinear optimization problem. The Levenberg–Marquardt algorithm is employed to solve the problem. To start a minimization, the linear solution of the Lagrange multiplier approach serves as the starting parameters.

2. Final Optimization

For general usage, the initial guess is already sufficient. However, for 3D shape measurement, the sensor should be calibrated as precisely as possible. In solving the initial guess, the distortion is not taken into account. As a consequence, the 3D calibration points are contaminated because they are derived from the distorted image points. Therefore, the parameters should be further refined by avoiding using the derived 3D coordinates. In this subsection, the parameters are solved by a nonlinear optimization using a complete model, where the objective function is formed by incorporating the only known image information and intersection conditions for the unknown 3D calibration points. Such optimization does not require known coordinates of the calibration points. Instead, the intersection points of the viewing plane and the feature lines are calculated from current parameters.

The intersection point on the mth line at the nth position is given by

$$\begin{cases} {}^S X_{m,n} = p_{m,n} + \lambda_{m,n} a_{m,n} \\ {}^S Y_{m,n} = q_{m,n} + \lambda_{m,n} b_{m,n} \\ {}^S Z_{m,n} = r_{m,n} + \lambda_{m,n} c_{m,n} \\ r_{11}({}^S X_{m,n} - X_0) + r_{12}({}^S Y_{m,n} - Y_0) + r_{13}({}^S Z_{m,n} - Z_0) = 0 \end{cases}.$$

$$(13)$$

The first three equations represent a feature line equation in a parametric form, where $\lambda_{m,n}$ is a scale factor. The coefficients $p_{m,n}$, $q_{m,n}$, $r_{m,n}$, $a_{m,n}$, $b_{m,n}$, and $c_{m,n}$ are calculated according to the pattern geometry and the 6D positions of the apparatus in Eq. (9). The last equation denotes the current viewing plane in the current iteration.

By substituting Eq. (13) into Eq. (11) and including the distortion parameters, a complete calibration model is formulated as a function of only the parameters. The cost function is given by

$$\boldsymbol{\varphi}^* = \min_{\boldsymbol{\varphi}} \sum_{n=1}^{N} \sum_{m=1}^{9} ([\hat{u}_{m,n}(\boldsymbol{\varphi})]^2 + [v_{m,n} - \hat{v}_{m,n}(\boldsymbol{\varphi})]^2), \quad (14)$$

where $(\hat{u}_{m,n}, \hat{v}_{m,n})$ is the updated reprojection calculated by combining Eqs. (11) and (13). In Eq. (14), the only known quantities are the observed pixel coordinates $v_{m,n}$. The initial guess from Eq. (12) is used for initializing the iteration. In addition, the initial guesses of the distortion coefficients are set to zero because they are very close to zero.

D. Locating the Sensor

So far, the parameters of the dual cameras with respect to the sensor coordinate frame have been achieved. When the sensor is taken to an actual measurement field, all that remains is to locate the sensor in the world coordinate frame, which means obtaining the location and orientation of the sensor. It should be noted that, in moving-sensor applications, the sensor should be located throughout the measurement process, just like the typical integrated scanning systems, e.g., ScanTRAK [28] and MetraScan [29].

Since the sensor's frame is defined by the sphere mounts installed on its body, the coordinates of the sphere centers in the sensor coordinate frame are known. As long as their corresponding coordinates in the world coordinate frame are measured, the optimal transformation from the sensor coordinate frame to the world coordinate frame is determined by taking the sphere centers as common reference points [30]. In an actual measurement field, a portable measuring system (also called optical tracker or optical CMM) is commonly used. The world coordinate frame can be either created or measured by it. Anyway, the world frame can be represented by the portable measuring system. As mentioned above, the sphere mounts are compatible with several portable measuring systems. By mounting adaptable cooperative targets into the sphere mounts, the position and orientation of the sensor can be resolved in the world coordinate frame effortlessly.

In summary, the whole calibration process can be outlined by the following steps, in which the first five steps are done in the laboratory and the last step is performed in an actual measurement field.

(1) Place the calibration apparatus at multiple positions within the FOV of Camera 1 to obtain a suitable distribution of calibration points.
(2) Locate the image coordinates of the calibration points in the captured images.
(3) Calculate the corresponding 3D coordinates of the calibration points (Section 3.B).
(4) Estimate the parameters of Camera 1 with respect to the sensor coordinate frame (Section 3.C).
(5) Repeat steps (1)–(4) for Camera 2.
(6) Locate the sensor in the world coordinate frame of an actual measurement field (Section 3.D).

4. EXPERIMENTAL RESULTS

In order to check the performance of the proposed calibration method, a series of experiments is carried out. The used cameras have a resolution of 4096 pixels, with a pixel size of 10 μm. The focal length of the lenses is 50 mm. The pattern is illuminated using blue light because its narrowband enables precise measurements, independent of environmental lighting conditions. The pattern accuracy can reach 1 μm, and the flatness of glass is better than 5 μm. The coordinates of the sphere mounts in the apparatus coordinate frame are measured by exploiting a multisensor coordinate measuring machine (VISIO 300-DCC, TESA Technology). As shown in Fig. 5, by mounting an adaptable spherical retro-reflector (SMR), the center of each sphere mount is measured by the stylus. The ring markers are measured by the video sensor. Taking these markers as common reference points, the coordinates of the sphere mounts are expressed in the apparatus coordinate frame. An accuracy of 10 μm can be guaranteed for the coordinate frame transformation.
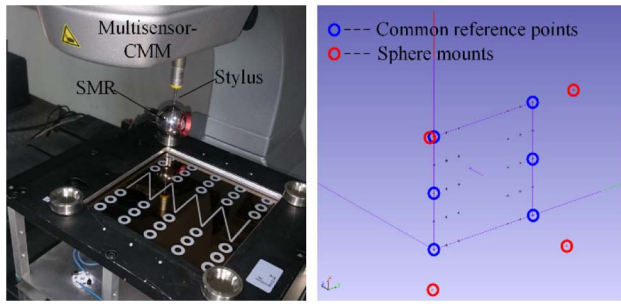
Fig. 5.   Geometry determination of the calibration apparatus.

Fig. 7.   Example of the captured images.

The *x* range of the feature lines is 120 mm, and the FOVs of both cameras at a working distance of 700 mm are approximately 400 mm. In one certain depth, the apparatus needs to be placed at three positions to fulfill the entire FOV. For each camera, the apparatus is placed at 21 positions with seven different depths. A total of 189 calibration points is obtained for each camera. The points for each camera cover a depth of about 180 mm, which is nearly the sensor's depth of field. At each position, a laser tracker (Leica AT-901 LR [23]) monitors the position and orientation of the apparatus by measuring a 1.5-inch (38.1 mm) SMR. According to the obtained transformations, the points are aligned to the sensor coordinate frame, as illustrated in Fig. 6.

To reduce the noises introduced during the image acquisition, 500 lines of repetitive images are captured at each position (e.g., Fig. 7 shows the first captured image for Camera 1). The exposure is set to an appropriate time so that the maximum intensity is between 200 and 255 for accurate center detection. The average images are taken for center detection.

All algorithms are implemented in MATLAB running on a laptop (with a Core i7-6700HQ CPU). By using full calibration points, the stereo line-scan sensor is calibrated. The results are listed in Table 1.

The upper rows are the parameters, and the last two rows are, respectively, the number of iterations and root-mean-square (RMS) values of the reprojection errors. For each camera, the initial guess and the optimal solution are recorded. As can be seen from the table, the results of the initial estimation are fairly close to the optimal estimation. Consequently, both
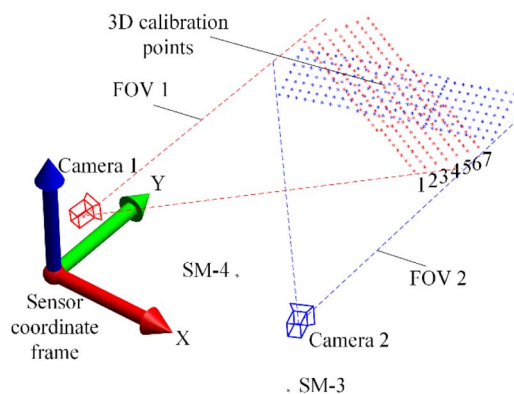
Table 1.   Calibration Results for the Stereo Line-Scan Sensor

Parameter	Camera 1		Camera 2	
(unit)	Initial	Optimal	Initial	Optimal
v_c (pixel)	2030.98	2030.98	2043.20	2043.88
F_y (pixel)	5561.98	5562.33	5570.48	5570.77
α (rad)	−0.913853	−0.913869	−1.488678	−1.488688
β (rad)	1.380095	1.380089	1.005891	1.005886
γ (rad)	−0.910521	−0.910539	−1.491256	−1.491284
X_0 (mm)	20.101	20.589	479.449	479.078
Y_0 (mm)	58.374	59.494	57.953	58.904
Z_0 (mm)	95.226	95.235	94.854	94.857
k_1	NaN	−3.76e−10	NaN	−3.02e−10
k_2	NaN	2.23e−17	NaN	4.32e−18
Iterations	51	5	38	4
RMS (pixel)	0.33	0.12	0.23	0.08

cameras converge to the optimal value quickly in the final optimization process. However, the initial estimation needs more iterations to be terminated because the linear solution of the Lagrange multiplier is not accurate enough in this case. The RMS errors reveal that the parameter estimation algorithm performs quite well on both cameras. The reprojection errors are further reduced by the final optimization. The final RMS errors are 0.12 and 0.08 pixels for Camera 1 and Camera 2, respectively.

The absolute calibration accuracy of the stereo line-scan sensor is verified by scanning a reference surface. The calibrated sensor is taken to the actual measurement field and is then mounted in a fixed position. The reference surface [see Fig. 8(a)], made of zirconia (with an accuracy of 1 μm), is attached onto a motorized stage by which a linear motion with constant velocity and orientation along a straight line is

Fig. 6.   Distribution of the calibration points. All points are aligned to the sensor coordinate frame, where the red points correspond to Camera 1 and the blue points are for Camera 2.
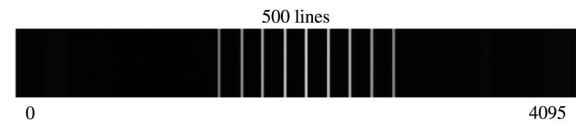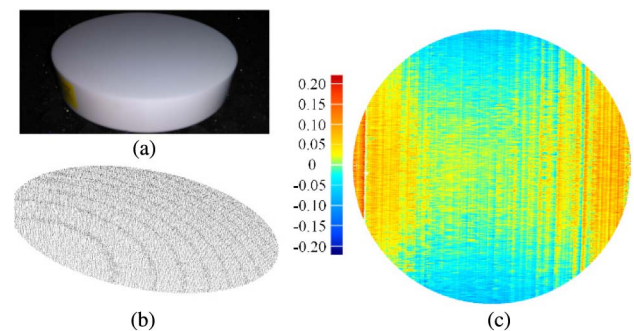

Fig. 8.   Verification of the absolute calibration accuracy by scanning a reference surface. (a) Measured reference surface. (b) Obtained point clouds. (c) Colormap of the deviations from the obtained point clouds to the reference surface.

provided. The surface moves over the stereo line-scan sensor with a speed of 60 mm/s. The frame rate of the sensor is set to 500 Hz. For simplicity, the sensor coordinate frame is used as the world coordinate frame so that the rotation matrix between them is an identity matrix and the translation is a zero vector. The moving direction is calibrated prior to measurement. More details about the experimental setup are described in Ref. [13]. The optimal parameters in Table 1 are used for triangulation of the stereo line-scan sensor. The obtained point clouds are shown in Fig. 8(b). A best-fit reference surface is fitted from the obtained point clouds. The deviations from all points to the reference surface are calculated. The color maps of the deviations are shown in Fig. 8(c). The statistical result of the deviations is counted. The RMS deviation is 0.067 mm.

In order to guarantee a sound analysis of the proposed calibration method, several further analyses are performed by experiments.

The first experiment is to compare the performances with different point distributions; Camera 1 and its data are taken for this analysis. As labeled in Fig. 6, the calibration points with the same depth are bundled as a row. Several different point distributions are generated via various combinations of the rows, as shown in the first column of Table 2. The corresponding calibration results are shown in the second column. Only the results of v_c are listed, and actually the other parameters have similar performances. The combinations are divided into five groups according to their depth coverages, as shown in the third column. The average value (AV) and standard deviation (SD) of each group are given in the fourth and fifth columns. The RMS reprojection errors are also counted in the last column.

To highlight the differences, the reprojection errors are also visualized using a box-and-whiskers plot in Fig. 9. The statistics in the figure are computed from the errors of all five groups.
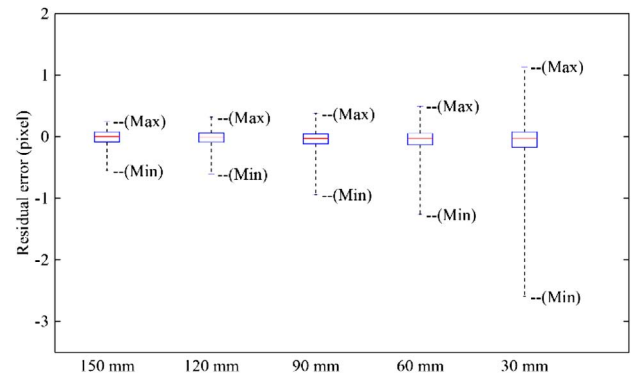

Fig. 9. Box-and-whiskers plot showing error statistics of various groups that have different depth coverage. The red line shows the median, the blue box is drawn between the first and the last quadrants, and the whiskers are drawn to the maximum and minimum values.

As revealed in Table 2 and Fig. 9, the average value of v_c shows no significant difference. However, when fewer rows of calibration points are used, the standard deviation of v_c increases, the RMS reprojection error becomes larger, and the distribution of the reprojection errors gets divergent. Results of this experiment indicate that the parameter estimation is more stable and reliable if more calibration points that are spatially distributed in a larger depth are used. In conclusion, this experiment reveals the importance of a good point distribution.

The second experiment is to test the robustness of the proposed parameter estimation algorithm. The camera (Camera 1) is calibrated using the proposed algorithm as well as the algorithm in Ref. [27] as a comparison. Similarly, various point distributions are used. In most cases, the results of the two algorithms are the same. However, in some poor conditions, as a negative example demonstrated in Fig. 10, only four positions are taken for calibration, which cover just a small portion of the entire FOV and occupy a restricted depth range. The projected rays are almost overlapping. As can be seen, the captured images have no significant difference. Such a poor ray intersection condition will lead to less constraint of parameters. The parameters will be highly vulnerable to noises. Making use of the illustrated configuration, the parameters are estimated via both algorithms. Figure 11 shows the comparative results.
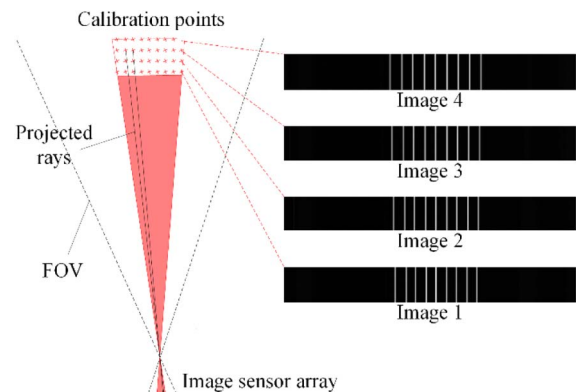
Table 2. Calibration Results with Various Point Distributions

Combinations	v_c (Pixel)	Depth (mm)	AV (Pixel)	SD (Pixel)	RMS (Pixel)
1, 2, 3, 4, 5, 6	2031.64	150	2031.74	0.14	0.13
2, 3, 4, 5, 6, 7	2031.84				
1, 2, 3, 4, 5	2031.55	120	2032.11	0.57	0.14
2, 3, 4, 5, 6	2032.69				
3, 4, 5, 6, 7	2032.10				
1, 2, 3, 4	2031.24	90	2032.36	0.77	0.17
2, 3, 4, 5	2032.65				
3, 4, 5, 6	2033.02				
4, 5, 6, 7	2032.53				
1, 2, 3	2031.86	60	2032.60	0.88	0.22
2, 3, 4	2032.34				
3, 4, 5	2032.63				
4, 5, 6	2034.09				
5, 6, 7	2032.08				
1, 2	2029.20	30	2032.42	2.40	0.41
2, 3	2034.92				
3, 4	2030.48				
4, 5	2034.70				
5, 6	2033.86				
6, 7	2031.34				

Fig. 10. Example of poor conditions. The points cover a small portion of the entire FOV and occupy a restricted depth range.
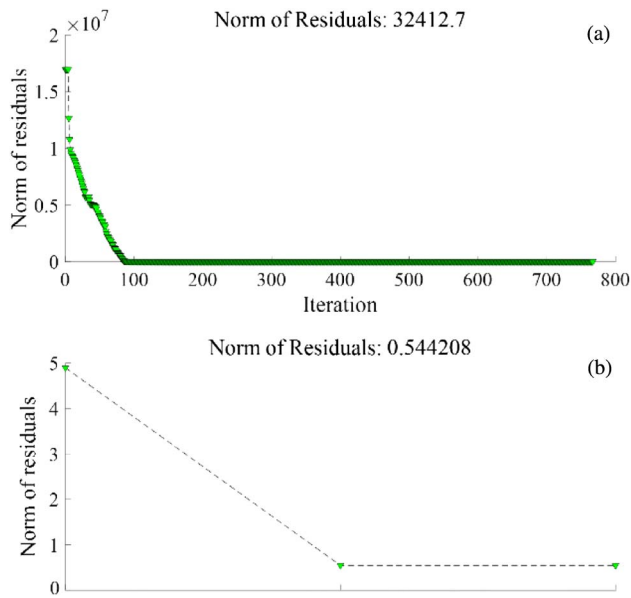
Fig. 11. Comparative results using a poor point distribution. (a) Final optimization process of the comparison algorithm. (b) Final optimization process of the proposed algorithm.

The comparison algorithm performs 768 iterations and converges to a local minimum. Judging from the norm of residuals, the solution is obviously not correct. Comparatively, the proposed algorithm performs just two iterations and is able to converge to a true local minimum. The reason for this is the fact that the proposed algorithm is able to provide a good initial estimation even in a poor point distribution condition. The result of this experiment indicates that the proposed parameter estimation algorithm is more robust than the comparison algorithm and is able to cope with various point distributions.

5. CONCLUSION

In this paper, a calibration method for a stereo line-scan sensor is proposed. Several reference points are also rigidly mounted onto the sensor. By doing this, the connection between the sensor and the world can be easily established. The only remaining task is to calibrate the two cameras with respect to the sensor frame. The sensor is allowed to be calibrated in the laboratory prior to measurement. The significant benefit is that sufficient time and optimum experimental condition permit more precise calibration. Meantime, a mobile calibration apparatus with extra reference points is designed. The sensor is calibrated by calibrating the cameras individually. By employing an ancillary instrument, a suitable point distribution is created by placing the apparatus in multiple positions with respect to the sensor. It is not restricted by the apparatus's dimension and therefore is able to deal with large FOV conditions. All calibration points are aligned into the sensor frame by utilizing the reference points on both the apparatus and the sensor. A robust parameter estimation algorithm is proposed. The initial guess is obtained using a simplified optimization model. The quality of the calibration method is experimentally tested, and the performance is further investigated. As indicated by the experimental

results, the proposed calibration method is suitable for 3D shape measurement, in which an accuracy of tens of micrometers is demanded. Furthermore, the proposed parameter estimator shows more robust behavior against various point distributions as compared to that of state-of-the art algorithms.

REFERENCES

1. H. L. Zhang, Y. J. Ren, C. J. Liu, and J. G. Zhu, "Flying spot laser triangulation scanner using lateral synchronization for surface profile precision measurement," Appl. Opt. **53**, 4405–4412 (2014).
2. Z. Z. Wei, F. Q. Zhou, and G. J. Zhang, "3D coordinates measurement based on structured light sensor," Sens. Actuators A Phys. **120**, 527–535 (2005).
3. B. Wu, F. Zhang, and T. Xue, "Monocular-vision-based method for online measurement of pose parameters of weld stud," Measurement **61**, 263–269 (2015).
4. L. He, S. Wu, and C. Wu, "Robust laser stripe extraction for three-dimensional reconstruction based on a cross-structured light sensor," Appl. Opt. **56**, 823–832 (2017).
5. J. Salvi, S. Fernandez, T. Pribanic, and X. Llado, "A state of the art in structured light patterns for surface profilometry," Pattern Recogn. **43**, 2666–2680 (2010).
6. S. Zhang, "Recent progresses on real-time 3D shape measurement using digital fringe projection techniques," Opt. Laser Eng. **48**, 149–158 (2010).
7. S. Heist, P. Kühmstedt, A. Tünnermann, and G. Notni, "Theoretical considerations on aperiodic sinusoidal fringes in comparison to phase-shifted sinusoidal fringes for high-speed three-dimensional shape measurement," Appl. Opt. **54**, 10541–10551 (2015).
8. Y. An, J. S. Hyun, and S. Zhang, "Pixel-wise absolute phase unwrapping using geometric constraints of structured light system," Opt. Express **24**, 18445–18459 (2016).
9. B. Li, Y. Wang, J. Dai, W. Lohry, and S. Zhang, "Some recent advances on superfast 3D shape measurement with digital binary defocusing techniques," Opt. Laser Eng. **54**, 236–246 (2014).
10. T. Ilchev, E. Lilienblum, B. Michaelis, B. Joedicke, and M. Schnitzlein, "A stereo line sensor system to high speed capturing of surfaces in color and 3d shape," in *International Conference on Computer Graphics Theory and Applications and International Conference on Information Visualization Theory and Applications* (2012).
11. E. Lilienblum and A. Al-Hamadi, "A structured light approach for 3-D surface reconstruction with a stereo line-scan system," IEEE Trans. Instrum. Meas. **64**, 1258–1266 (2015).
12. P. Zhang, T. Jay Arre, and A. Ide-Ektessabi, "A line scan camera-based structure from motion for high-resolution 3D reconstruction," J. Cult. Herit. **16**, 656–663 (2015).
13. B. Sun, J. Zhu, L. Yang, S. Yang, and Y. Guo, "Sensor for in-motion continuous 3D shape measurement based on dual line-scan cameras," Sensors **16**, 1949 (2016).
14. M. V. Ortiz Segovia, P. Urban, J. P. Allebach, P. Zhang, T. Takeda, J. A. Toque, Y. Murayama, and A. Ide-Ektessabi, "A line scan camera based stereo method for high resolution 3D image reconstruction," Proc. SPIE **9018**, 901807 (2014).
15. Z. Zhang, "A flexible new technique for camera calibration," IEEE Trans. Pattern Anal. Mach. Intell. **22**, 1330–1334 (2000).
16. R. Horaud, R. Mohr, and B. Lorecki, "On single-scanline camera calibration," IEEE Trans. Robot. Autom. **9**, 71–75 (1993).

17. C. A. Luna, M. Mazo, J. L. Lazaro, and J. F. Vazquez, "Calibration of line-scan cameras," IEEE Trans. Instrum. Meas. **59**, 2185–2190 (2010).
18. E. Lilienblum, A. Al-Hamadi, and B. Michaelis, "A coded 3d calibration method for line-scan cameras," in *German Conference on Pattern Recognition (GCPR)*, J. Weickert, M. Hein, and B. Schiele, eds. (Springer-Verlag, 2013), pp. 81–90.
19. D. D. Li, G. J. Wen, B. W. Hui, S. H. Qiu, and W. F. Wang, "Cross-ratio invariant based line scan camera geometric calibration with static linear data," Opt. Laser Eng. **62**, 119–125 (2014).
20. J. Drareni, S. Roy, and P. Sturm, "Plane-based calibration for linear cameras," Int. J. Comput. Vis. **91**, 146–156 (2011).
21. B. Hui, G. Wen, Z. Zhao, and D. Li, "Line-scan camera calibration in close-range photogrammetry," Opt. Eng. **51**, 053602 (2012).
22. B. Hui, G. Wen, P. Zhang, and D. Li, "A novel line scan camera calibration technique with an auxiliary frame camera," IEEE Trans. Instrum. Meas. **62**, 2567–2575 (2013).
23. Hexagon Manufacturing Intelligence, http://www.hexagonmi.com/products/laser-tracker-systems.
24. Nikon Metrology, https://www.nikonmetrology.com/en-gb/product/k-cmm.
25. S. P. Fang, X. H. Xia, and Y. Xiao, "A calibration method of lens distortion for line scan cameras," Optik **124**, 6749–6751 (2013).
26. T. Luhmann, S. Robson, S. Kyle, and J. Boehm, *Close-Range Photogrammetry and 3D Imaging* (De Gruyter, 2013).
27. B. Sun, J. Zhu, L. Yang, S. Yang, and Z. Niu, "Calibration of line-scan cameras for precision measurement," Appl. Opt. **55**, 6836–6843 (2016).
28. Northern Digital, https://www.ndigital.com/msci/products/.
29. Creaform, http://www.creaform3d.com/en/metrology-solutions/optical-3d-scanner-metrascan.
30. D. W. Eggert, A. Lorusso, and R. B. Fisher, "Estimating 3-D rigid body transformations: a comparison of four major algorithms," Mach. Vision Appl. **9**, 272–290 (1997).