A new microscopic telecentric stereo vision system - Calibration, rectification, and three-dimensional reconstruction

Yan Hu [a,b], Qian Chen [a], Shijie Feng [a,b], Tianyang Tao [a,b], Anand Asundi [c], Chao Zuo [a,b,*]

[a] *Jiangsu Key Laboratory of Spectral Imaging & Intelligent Sense, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China*
[b] *Smart Computational Imaging (SCI) Laboratory, Nanjing University of Science and Technology, Nanjing, Jiangsu Province 210094, China*
[c] *School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798, Singapore*

ARTICLE INFO

ABSTRACT

In stereo vision based three-dimensional (3D) measurements, calibration and stereo matching are the most challenging tasks for accurate 3D reconstruction. In this paper, a new microscopic telecentric stereo vision system is proposed to retrieve 3D data of micro-level objects by direct triangulation from two accurately calibrated telecentric cameras. The complex projector calibration procedure commonly seen in traditional structured-light based systems can be avoided. Besides, an improved and practical calibration framework of telecentric cameras is presented. Compared with existing calibration methods which retrieve the distortion parameters without reasonable initial guesses, our proposed approach derives reliable initial guesses for all the camera parameters to avoid the local minima problem based on the estimation algorithms before iterations. To realize precise sub-pixel stereo matching, we propose an effective searching algorithm based on the epipolar rectification of the absolute phase maps obtained from fringe projection profilometry. Experimental results indicate that a measurable volume of 10 mm (L) × 7 mm (W) × 7 mm (H) is achieved with the standard deviation of 1.485 μm.

1. Introduction

In recent years, advances in precision manufacturing demand micro-level three-dimensional (3D) metrology to guarantee accurate fabrication and optimal designs. As a classic passive 3D measurement technique, stereo vision measurement has been extensively used for topometry of a wide range of surfaces [1]. However, correspondence (stereo matching), which tries to find the corresponding parts in multi-views remains as one main challenge in passive stereo vision measurement because the physical features are usually not readily detectable and distinguishable [2]. A practical solution is to use active markers for accurate stereo matching.

As a widely used non-contact method for surface profiling, structured-light projection technique can provide full-field active markers [3,4]. In traditional structured-light based systems, a digital projector is employed to create required artificial fringe patterns and considered as one view for stereo-measurement [5]. Appropriate calibration of a projector based multi-view system is vital for accurate 3D reconstruction [6,7]. For micro-3D measurement, the smaller field of view (FOV) in the microscopic multi-view system necessitates auxiliary lenses. One approach is to adopt a stereo-microscope for this task [8]. However, because of the complicated optics inside a stereo-microscope, calibration of such an optical system with a flexible method [6] is difficult. The sec-

ond approach is to use telecentric lenses owing to their increased depth of field and constant magnification along the optical axis [9]. Initial studies using telecentric lenses for micro-3D topometry were based on a single-camera configuration [10–14]. However, the projector in these systems needs calibration with the help of a calibrated camera, which is a complex process. Besides, the calibration accuracy cannot be guaranteed due to the unpredictability caused by the gamma effects, lens distortion, and the misalignment of the additional auxiliary lens in the projector. To utilize the advantage of full-field active markers provided by the projector and avoid being bothered by the issues related to the projector, a potential solution is to design a stereo vision system that comprises two well-calibrated telecentric cameras for stereo matching and one projector for active marker projection.

In this paper, a microscopic telecentric stereo vision system is introduced to overcome the challenges of system calibration and stereo matching. The calibration procedure only involves two telecentric cameras without the necessity to calibrate the projector; therefore, the complexity of the projector calibration and other issues in single-camera based systems can be effectively avoided. For the telecentric camera calibration, many approaches have been proposed. Li and Tian [15] proposed a planar calibration technique which however does not consider the ambiguity of the extrinsic parameters. Chen et al. [16] addressed

this shortcoming by capturing image pairs along the depth axis for each posture of the calibration board to unambiguously recover the rotation matrix. However, the distortion center is not calibrated for the intrinsic parameters of the camera. Rao et al. [17] also proposed a method that directly uses the pinhole model to describe the imaging process of the telecentric camera, which may fail for lenses with good telecentricity. Different from the pinhole cameras, the distortion center (projection center) of a telecentric camera locates at infinite so that it is hidden in the distortion-free camera model and can be only derived through iterations.

To better calibrate the distortion of a telecentric camera, we have to know the position of the distortion center correctly. In Yao's work [18], a novel two-step calibration procedure to get the full-scale parameters including the distortion center by two non-linear optimization processes is proposed. However, in their method, the detector's center is assumed as the initial value of the distortion center. Because the full-scale camera parameters are all estimated together in each optimization step, the search process from the detector's center is prone to find a local minimum when the actual distortion center is not close enough to the detector's center. To solve this problem, we propose an improved and efficient calibration framework for telecentric cameras in this paper. Compared with the state-of-the-art methods that ignore or assume initial guesses for the distortion parameters before the iterations, the proposed method uses a two-step estimation algorithm to decouple the distortion center calculation from the full-scale parameters optimization. The initial guesses of distortion coefficients are also effectively estimated before the iteration. With these trustworthy initial guesses, a more robust calibration framework is achieved, and thus the probability of being trapped in local minima is significantly decreased.

In the stereo matching step, telecentric epipolar rectification of the calibrated telecentric stereo vision system is performed to facilitate the matching process. Firstly full-field phase maps are obtained as in the traditional manner. After the distortion compensation and epipolar rectification of the phase maps, all the points of the measured scene lie in the same vertical position in the two images. Thus by searching the pixel with the closest phase value and inverse linear interpolation, sub-pixel stereo matching of the two telecentric cameras can be efficiently achieved.

In this paper, the overall processes of the system calibration, epipolar rectification, and 3D reconstruction using our microscopic profilometry system are detailed in Sections 2–4, respectively. The simulation result shows that with the reliable camera parameters acquired based on trustworthy initial guesses, the re-projection error can be reduced by one order of magnitude or so. The experiments also demonstrate that the proposed stereo matching method based on epipolar rectification of the full-field and high-resolution phase maps accommodates the measurements for various kinds of objects with a sizeable measurable depth. Finally, the conclusion and suggestion for future work are summarized to close this paper.

2. Telecentric camera model and telecentric epipolar rectification

2.1. Telecentric camera model

Unlike the perspective projection in the imaging process of a pinhole camera, telecentric cameras provide purely orthographic projections of the scene, which makes it easier to measure physical size independently from depths. The imaging model of a distortion-free telecentric camera is illustrated in Fig. 1. (x_w, y_w, z_w) and (x_c, y_c, z_c) are the coordinates of an arbitrary object point $\mathbf{P}$ in the world coordinate system $O - X_W Y_W Z_W$ and camera coordinate system $O_C - X_C Y_C Z_C$, respectively. Because of the affine projection from the camera coordinate to the image coordinate, the camera center O_C of a telecentric camera locates at infinite. $\mathbf{p}(u, v)$ is the image coordinate of point $\mathbf{P}$ and $\mathbf{e}(u_0, v_0)$ is the image coordinate of the optical center, which is also known as the distortion center and usually does not coincide with the detector's cen-
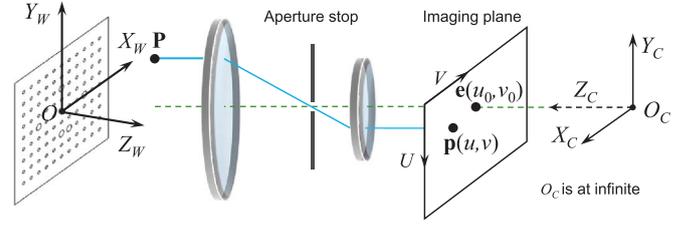

Fig. 1. Simplified schematic of the imaging process and the coordinate systems of a telecentric camera.

ter (image center). The aperture stop is used to assure the telecentricity of the lens [19].

For the object point $\mathbf{P}(x_c, y_c, z_c)$ in the camera coordinate system, its homogeneous image coordinate $\tilde{\mathbf{p}}$ is projected in an affine form as

$$\begin{bmatrix} \mathbf{p} \\ 1 \end{bmatrix} = \begin{bmatrix} m & 0 & u_0 \\ 0 & m & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ 1 \end{bmatrix}, \tag{1}$$

where m is the effective magnification of the telecentric lens. For an ideal telecentric camera, u_0 and v_0 can be set as zeros. Each calibration pattern has it own world coordinate system. The pattern on the calibration board determines X_W and Y_W as well as the original point O. The world and camera coordinate systems are related by a rotation matrix $\mathbf{R}$ and a translation vector $\mathbf{t}$ as

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \mathbf{R} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \mathbf{t}. \tag{2}$$

Here, $\mathbf{R} = [\mathbf{r}_x \ \mathbf{r}_y \ \mathbf{r}_z]^T$ and $\mathbf{t} = [t_x \ t_y \ t_z]^T$. The whole projection of a point $\mathbf{P}(x_w, y_w, z_w)$ in the world coordinate to an image point $\mathbf{p}(u, v)$ can be expressed as [15]

$$\begin{bmatrix} \mathbf{p} \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{R}_{2 \times 3} & \mathbf{t}_{2 \times 1} \\ \mathbf{0}_{1 \times 3} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{P} \\ 1 \end{bmatrix} = \mathbf{H} \begin{bmatrix} \mathbf{P} \\ 1 \end{bmatrix}, \tag{3}$$

Here,

$$\mathbf{A} = \begin{bmatrix} m & 0 & u_0 \\ 0 & m & v_0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{4}$$

and $\mathbf{H}$ is the homography matrix, which transforms the world coordinates of objects into their corresponding image coordinate.

2.2. Telecentric epipolar rectification

The epipolar geometry of two telecentric cameras is similar to that of two pinhole cameras. A pixel $\mathbf{p}_L(u_L, v_L)$ in the left view corresponds to an epipolar line in the other view, on which the matched pixel $\mathbf{p}_R(u_R, v_R)$ meets an the affine epipolar constraint equation as [20]

$$au_R + bv_R + cu_L + bv_L + e = 0, \tag{5}$$

where, $a \sim e$ are five constants. As shown in Fig. 2(a), pixel $\mathbf{p}_{L_1}$ on the left camera corresponds to an epipolar line l_R, on which a pixel $\mathbf{p}_{R_1}$ also corresponds to an epipolar line l_L. The telecentric stereo images can be rectified to make the matched pixel pairs in the same vertical position to facilitate the stereo matching of the stereo vision system as shown in Fig. 2(b). The original images need to be transformed into new ones, and thus a new set of camera parameters is acquired. Different from calculating the fundamental matrix between two views [21], we first calibrate the cameras with two set of parameters for each camera and then rectify them. Here we use a prime to represent the new parameters and add subscript "L" or "R" to distinguish the left and right cameras.
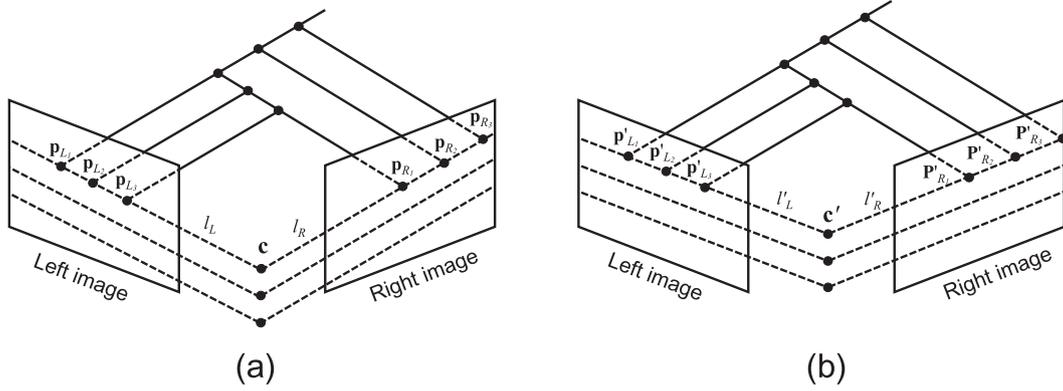
Fig. 2. Epipolar constraints for (a) unrectified telecentric stereo vision system and (b) rectified stereo vision system.

Then we describe the original projection process for both cameras as

$$
\begin{cases}
\begin{bmatrix} \mathbf{p}_L \\ 1 \end{bmatrix} = \mathbf{A}_L \begin{bmatrix} \mathbf{R}_{L2\times3} & \mathbf{t}_{L2\times1} \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{P} \\ 1 \end{bmatrix} = \mathbf{H}_L \begin{bmatrix} \mathbf{P} \\ 1 \end{bmatrix} \\[2em]
\begin{bmatrix} \mathbf{p}_R \\ 1 \end{bmatrix} = \mathbf{A}_R \begin{bmatrix} \mathbf{R}_{R2\times3} & \mathbf{t}_{R2\times1} \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{P} \\ 1 \end{bmatrix} = \mathbf{H}_R \begin{bmatrix} \mathbf{P} \\ 1 \end{bmatrix}
\end{cases}
\tag{6}
$$

and the rectified as

$$
\begin{cases}
\begin{bmatrix} \mathbf{p}'_L \\ 1 \end{bmatrix} = \mathbf{A}'_L \begin{bmatrix} \mathbf{R}'_{L2\times3} & \mathbf{t}'_{L2\times1} \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{P} \\ 1 \end{bmatrix} = \mathbf{H}'_L \begin{bmatrix} \mathbf{P} \\ 1 \end{bmatrix} \\[2em]
\begin{bmatrix} \mathbf{p}'_R \\ 1 \end{bmatrix} = \mathbf{A}'_R \begin{bmatrix} \mathbf{R}'_{R2\times3} & \mathbf{t}'_{R2\times1} \\ \mathbf{0}_{1\times3} & 1 \end{bmatrix} \begin{bmatrix} \mathbf{P} \\ 1 \end{bmatrix} = \mathbf{H}'_R \begin{bmatrix} \mathbf{P} \\ 1 \end{bmatrix}
\end{cases}
\tag{7}
$$

According to Liu et al. [22], we can use the same intrinsic matrix for both cameras as

$$
\mathbf{A}'_L = \mathbf{A}'_R = (\mathbf{A}_L + \mathbf{A}_R)/2,
\tag{8}
$$

and the new rotation matrix $\mathbf{R}'_L$ and $\mathbf{R}'_R$ is derived through

$$
\begin{cases}
\mathbf{r}'_{zL} = \mathbf{r}_{zL} \\
\mathbf{r}'_{zR} = \mathbf{r}_{zR}
\end{cases},
\tag{9}
$$

$$
\begin{cases}
\mathbf{r}'_{yL} = \mathrm{norm}(\mathbf{r}'_{zL} \times \mathbf{r}'_{zR}) \\
\mathbf{r}'_{yR} = \mathbf{r}'_{yL}
\end{cases},
\tag{10}
$$

and

$$
\begin{cases}
\mathbf{r}'_{xL} = \mathbf{r}'_{zL} \times \mathbf{r}'_{yL} \\
\mathbf{r}'_{xR} = \mathbf{r}'_{zR} \times \mathbf{r}'_{yR}
\end{cases},
\tag{11}
$$

where, function "norm" means normalization of a matrix and operator "$\times$" means cross product. The new translation vectors $\mathbf{t}'_L$ and $\mathbf{t}'_R$ are then obtained through

$$
\begin{cases}
\begin{bmatrix} \tau_{x_L} & \tau_{y_L} & \tau_{z_L} \end{bmatrix}^{\mathrm{T}} = \mathbf{R}'_L \mathbf{R}_L^{-1} \begin{bmatrix} \mathbf{t}_{L2\times1}^{\mathrm{T}} & 1 \end{bmatrix}^{\mathrm{T}} \\
\begin{bmatrix} \tau_{x_R} & \tau_{y_R} & \tau_{z_R} \end{bmatrix}^{\mathrm{T}} = \mathbf{R}'_R \mathbf{R}_R^{-1} \begin{bmatrix} \mathbf{t}_{R2\times1}^{\mathrm{T}} & 1 \end{bmatrix}^{\mathrm{T}}
\end{cases},
\tag{12}
$$

and

$$
\begin{cases}
\mathbf{t}'_{L2\times1} = \begin{bmatrix} \tau_{x_L} & (\tau_{y_L} + \tau_{y_R})/2 \end{bmatrix}^{\mathrm{T}} \\
\mathbf{t}'_{R2\times1} = \begin{bmatrix} \tau_{x_R} & (\tau_{y_L} + \tau_{y_R})/2 \end{bmatrix}^{\mathrm{T}}
\end{cases}
\tag{13}
$$

For now, we have derived all the parameters to generate $\mathbf{H}'_L$ and $\mathbf{H}'_R$. Therefore, the transformation between the rectified and the original image coordinates can be executed by

$$
\begin{cases}
\begin{bmatrix} \mathbf{p}'_L \\ 1 \end{bmatrix} = \mathbf{H}'_L \mathbf{H}_L^{+} \begin{bmatrix} \mathbf{p}_L \\ 1 \end{bmatrix} \\[2em]
\begin{bmatrix} \mathbf{p}'_R \\ 1 \end{bmatrix} = \mathbf{H}'_R \mathbf{H}_R^{+} \begin{bmatrix} \mathbf{p}_R \\ 1 \end{bmatrix}
\end{cases}
\tag{14}
$$

The symbol }}+'' means pseudo-inverse of a matrix.

However, the rectification discussed above is valid only under the distortion-free situation. Thus, before the rectification, the lens distortion should be correctly removed by the accurate calibration of the cameras. The following section details our calibration process for the telecentric cameras.

3. Calibration of telecentric cameras

3.1. Homography matrix

We use a plane with equally spaced circles as the calibration board as shown in Fig. 1. By placing the board in the common field of view of the two cameras, one needs to capture several different calibration patterns in any posture to calibrate the cameras. Each calibration posture corresponds to a unique 3D coordinate system which sets the pattern plane as its $z_w = 0$ plane. For each calibration pattern, the circular markers are on the pattern plane, thus their z_w are all set as zero in the calculation. The pixel coordinates (u, v) of the circles' centers can be extracted from the captured images, which make up an over-determined set of equations based on Eq. (3) as

$$
\begin{bmatrix}
x_w & y_w & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & x_w & y_w & 1 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{bmatrix}
\begin{bmatrix}
h_{11} \\ h_{12} \\ h_{14} \\ h_{21} \\ h_{22} \\ h_{24}
\end{bmatrix}
=
\begin{bmatrix}
u \\ v \\ \vdots
\end{bmatrix}.
\tag{15}
$$

Here, $h_{ij}(i = 1, 2; j = 1, 2, 4)$ are the elements of $\mathbf{H}$, which can be calculated for each calibration posture using least square method. For h_{13} and h_{23}, we have to calculate them after the rotation matrix is obtained.

3.2. Effective magnification

Based on Eqs. (3) and 4, it is easy to get

$$
r_{ij} = h_{ij}/m
\tag{16}
$$

with $i = 1, 2; j = 1, 2, 3$. Because the rotation matrix $\mathbf{R}$ is a unitary and orthogonal matrix, we can get

$$
\begin{cases}
\langle \mathbf{r}_i, \mathbf{r}_j \rangle = 1, i = j \\
\langle \mathbf{r}_i, \mathbf{r}_j \rangle = 0, i \neq j
\end{cases}
\tag{17}
$$

Combining Eqs. (16) and 17, the following equation can be derived [18]

$$m^4 + Bm^2 + C = 0, \qquad (18)$$

where, $B = -(h_{11}^2 + h_{12}^2 + h_{21}^2 + h_{22}^2)$ and $C = (h_{11}h_{22} - h_{12}h_{21})^2$. Basically, we get two different m^2 as the roots for Eq. (18). Theoretically, the bigger root is a constant for any $\mathbf{H}$ while the smaller root changes for different $\mathbf{H}$. Thus the unique solution for m^2 is always the bigger root. Without loss of generality, we assume the effective magnification as a positive value so that we get

$$m = \sqrt{0.5\left(-B + \sqrt{B^2 - 4C}\right)}. \qquad (19)$$

Then $r_{ij}(i = 1, 2; j = 1, 2)$ follows based on Eq. (16).

3.3. Rotation matrix and translation vetor

Although the effective magnification can be derived from each calibration image, the extrinsic matrix **[R t]** of different calibration posture still varies. From Section 3.2, $\mathbf{R}_{2\times2}$ is obtained. Because $\mathbf{R}$ is unitary and orthogonal, the remaining elements of $\mathbf{R}$ can be recovered by

$$\begin{cases} r_{x3} = \pm\sqrt{1 - r_{x1}^2 - r_{x2}^2} \\ r_{y3} = \pm\sqrt{1 - r_{y1}^2 - r_{y2}^2} \\ \mathbf{r}_z = \mathbf{r}_x \times \mathbf{r}_y \end{cases} \qquad (20)$$

Obviously there still exists uncertainty of the signs for r_{x3} and r_{y3}. To solve the sign ambiguity problem, Chen et al. [16] proposed a method that uses a micro-positioning stage to provide a translational displacement z_d along the Z_W axis of the calibration pattern. Together with the captured image before the displacement, the signs of r_{x3} and r_{y3} can be unambiguously determined. However, this method is effective only provided that (u_0, v_0) and (t_x, t_y) are known. Then in Yao's method [18], the detector's center is temporally preset as (u_0, v_0), and then (t_x, t_y) can be presented as

$$\begin{cases} t_x = (h_{14} - u_0)/m \\ t_y = (h_{24} - v_0)/m \end{cases} \qquad (21)$$

Accordingly, extrinsic parameters in $[\mathbf{R}_{2\times3}\ \mathbf{t}_{2\times1}]$ can be completely recovered. However, when the lens distortion is considered, the distortion center is required because the imaging properties are radially symmetric around the distortion center. Assuming the detector's center as the optical center is anyhow a compromise that could result in inconsistent results with Eq. (17). Therefore, it is necessary to estimate a reasonable initial guess of the distortion center not only for the accuracy and correctness of the extrinsic parameters but also for the better convergence of the iteration to avoid local minima situation.

3.4. Distortion center estimation

In the last subsection, we have discussed and concluded that an initial guess of the distortion center is important to unambiguously and accurately recover the rotation matrix $\mathbf{R}$ as well as the translation vector $\mathbf{t}_{2\times1}$. Besides, reasonable initial guesses can effectively decrease the probability of being trapped into local minima. Here, we present a two-step method to accurately retrieve the distortion center of a telecentric camera before the full-scale parameters optimization.

Step 1. Radial distortion center estimation In this step, we ignore the tangential distortion and only consider the radial distortion to estimate an initial value of the distortion center $\mathbf{e}(u_0, v_0)$. The relation between a distorted point $\mathbf{p}_d(u_d, v_d)$ and its undistorted point $\mathbf{p}_u(u_u, v_u)$ on the image plane can be expressed as [23]

$$\begin{cases} u_d = \lambda(u_u - u_0) + u_0 \\ v_d = \lambda(v_u - v_0) + v_0 \end{cases} \qquad (22)$$
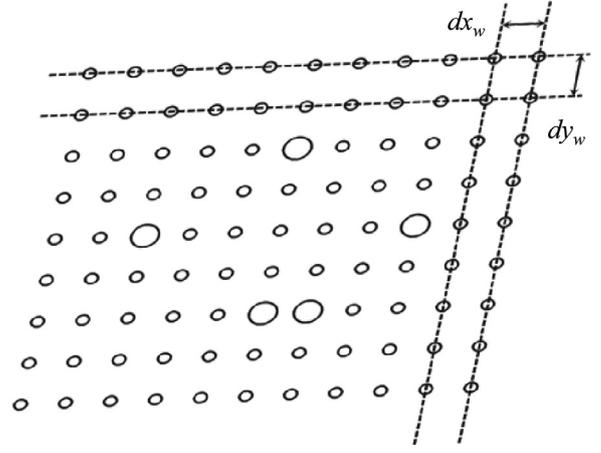

Fig. 3. Geometric properties of the feature points on the calibration pattern. dx_w and dy_w are the intervals of the feature points in X_W and Y_W directions.

where, $\lambda = (1 + k_1 r^2 + k_2 r^4 + k_3 r^6)$ with $r^2 = (u_u - u_0)^2 + (v_u - v_0)^2$. $k_{1,2,3}$ are the radial distortion coefficients. If we use homogeneous coordinates, we can represent Eq. (22) as $\tilde{\mathbf{p}}_d = \lambda(\tilde{\mathbf{p}}_u - \tilde{\mathbf{e}}) + \tilde{\mathbf{e}}$ and multiply $[\tilde{\mathbf{e}}]_\times$ (the 3×3 matrix representing the cross product) on the left of this equation, resulting in $[\tilde{\mathbf{e}}]_\times\tilde{\mathbf{p}}_d = \lambda([\tilde{\mathbf{e}}]_\times\tilde{\mathbf{p}}_u)$. Since $\tilde{\mathbf{p}}_u = \mathbf{H}\tilde{\mathbf{P}}$, we have

$$[\tilde{\mathbf{e}}]_\times\tilde{\mathbf{p}}_d = \lambda([\tilde{\mathbf{e}}]_\times\mathbf{H}\tilde{\mathbf{P}}). \qquad (23)$$

Finally, multiplying on the left by $\tilde{\mathbf{p}}_d^T$, and observing that $\tilde{\mathbf{p}}_d^T[\tilde{\mathbf{e}}]_\times\tilde{\mathbf{p}}_d = 0$ we get

$$\tilde{\mathbf{p}}_d^T([\tilde{\mathbf{e}}]_\times\mathbf{H})\tilde{\mathbf{P}} = 0. \qquad (24)$$

Writing $\mathbf{F}_a = [\tilde{\mathbf{e}}]_\times\mathbf{H}$ as a 3×3 matrix, we have the affine fundamental matrix relation, which can be solved using singular value decomposition (SVD). Note that $\tilde{\mathbf{e}}^T[\tilde{\mathbf{e}}]_\times\mathbf{H} = \mathbf{0}^T$. Then we have

$$\mathbf{F}_a^T\tilde{\mathbf{e}} = \mathbf{0}. \qquad (25)$$

By solving Eq. (25), the radial distortion center can be obtained. In order to acquire a more trustworthy result, $\mathbf{F}_a$ that derived from multi-views of the calibration pattern can be integrated together into a single $\mathbf{F}_A$.

Step 2. Optimization in the image-plane After estimating the radial distortion center, we then apply another optimization step considering both tangential and radial distortion. In this optimization, we rewrite the relation between the ideal and distorted pixel coordinates in another commonly used form as [24]

$$\begin{cases} u_u = \lambda' u_n + 2p_1' u_n v_n + p_2'(r^2 + 2u_n^2) + u_0 \\ v_u = \lambda' v_n + 2p_2' u_n v_n + p_1'(r^2 + 2v_n^2) + v_0 \end{cases} \qquad (26)$$

Here, $\lambda' = (1 + k_1' r'^2 + k_2' r'^4 + k_3' r'^6)$ with $r'^2 = u_n^2 + v_n^2$ and $u_n = u_d - u_0$ and $v_n = v_d - v_0$. $k_{1,2,3}'$ and $p_{1,2}'$ are the radial and tangential distortion coefficients, respectively. Note that because the relation in Eq. (26) does not coincide with Eq. (22), it is only used in the optimization to facilitate the calculation.

In this step, the optimization is based on the property of the affine projection from the calibration pattern to the captured image. As shown in Fig. 3, the feature points on the calibration pattern make up a 2D grid with the following geometric properties: all the points lie on a line in either X_W or Y_W direction, and these lines are equally spaced by dx_w and dy_w in X_W and Y_W directions, respectively. However, when captured by a camera through a lens with distortion, the feature points shift from the lines and all the adjacent points are no more equally spaced. The purpose of the optimization is to find the best solution of (u_0, v_0) and the appropriate distortion coefficients to make the compensated points after using Eq. (26) meet the geometric properties of the 2D grid again with the help of the cost function detailed below.

The initial guess of the distortion coefficients is set to zeros. After each re-projection, the lines in both directions are fitted using the extracted feature points. There are totally $N = 20$ lines on the calibration board: 11 lines along Y_W direction and 9 lines along X_W direction (as shown in Fig. 3). Writing the ith fitted line as l^i_{fit}. The distances from the line to its corresponding feature points $\mathbf{p}^i_d$ are denoted as $\mathbf{D}_i = \text{dist}(\mathbf{p}^i_d, l^i_{fit})$. Because of the affine transformation, the intervals of the feature points in X_W or Y_W directions should still be equal. Write all the intervals of the extracted points in X_W and Y_W directions as $\mathbf{I}_X$ and $\mathbf{I}_Y$. Then average values of $\mathbf{I}_X$ and $\mathbf{I}_Y$ are subtracted from them to get $\bar{\mathbf{I}}_X$ and $\bar{\mathbf{I}}_Y$. Finally, we have the cost function being minimized in the optimization:

$$F_{opt} = \underset{(u_0, v_0)}{\text{argmin}} \sum \left\| [\mathbf{D}^T_1 \cdots \mathbf{D}^T_N \ \bar{\mathbf{I}}^T_X \ \bar{\mathbf{I}}^T_Y] \right\|^2, \tag{27}$$

which is solved using *Levenberg − Marquardt* algorithm. After the distortion center is accurately retrieved, the closed-form solutions for all the parameters in Eq. (3) can be derived and used as the reliable initial guesses for the full-scale optimization.

3.5. Non-linear full-scale optimization

Without considering the lens distortion, we have got the closed-form solutions of the parameters in Eq. (3). However, two problems still need taking into account. The first one is that the derived rotation matrix from Eq. (20) is not strictly orthogonal. The second one is that lens distortion has not been calibrated. In this subsection, two non-linear optimization processes are conducted in sequence to solve these problems.

The first non-linear optimization is to get the unique intrinsic matrix $\mathbf{A}$ from different calibration postures as well as the orthogonal rotation matrix $\mathbf{R}$ and translation vector $\mathbf{t}$ by minimizing the following function:

$$F_1 = \underset{\mathbf{A}, \mathbf{R}, \mathbf{t}}{\text{argmin}} \sum_j \sum_k \left\| \mathbf{p}_{jk} - \hat{\mathbf{p}}(\mathbf{A}, \mathbf{R}_j, \mathbf{t}_j, \mathbf{P}_{jk}) \right\|^2. \tag{28}$$

Here, j is the number of calibration postures, k is the number of feature points on the calibration pattern, $\mathbf{p}_{jk}$ and $\mathbf{P}_{jk}$ are the control points on the captured images and the calibration pattern, respectively. $\hat{\mathbf{p}}$ is projection of feature points $\mathbf{P}_{jk}$ according to Eq. (3). In the optimization process, the rotation matrix $\mathbf{R}$ is firstly transformed into an orthogonal matrix using SVD and then parameterized by three scalers using *Rodrigues' rotation formula*.

When F_1 is minimized, the generated intrinsic matrix $\mathbf{A}$ and extrinsic matrix $[\mathbf{R}_{2 \times 3} \ \mathbf{t}_{2 \times 1}]$ are utilized to calculate the initial guesses of distortion coefficients in the second non-linear optimization. In this work, the lens distortion is represented by five coefficients denoted as $\mathbf{kp} = [k_1 \ k_2 \ p_1 \ p_2 \ k_3]^T$ and modeled in the camera coordinates system as

$$\begin{bmatrix} x^d_c \\ y^d_c \end{bmatrix} = \begin{bmatrix} x^u_c \\ y^u_c \end{bmatrix} + \begin{bmatrix} \delta_x \\ \delta_y \end{bmatrix},$$

$$\begin{bmatrix} \delta_x \\ \delta_y \end{bmatrix} = \begin{bmatrix} x^u_c r^2_c & x^u_c r^4_c & 2x^u_c y^u_c & r^2_c + 2x^u_c & x^u_c r^6_c \\ y^u_c r^2_c & y^u_c r^4_c & r^2_c + 2y^u_c & 2x^u_c y^u_c & y^u_c r^6_c \end{bmatrix} \mathbf{kp}. \tag{29}$$

Here, (x^d_c, y^d_c) and (x^u_c, y^u_c) are the distorted and undistorted position in the camera coordinate system. $r^2_c = x^{u2}_c + y^{u2}_c$. Based on the optimized results above, we can derive the initial value of $\mathbf{kp}$ by solving Eq. (29) with

$$\begin{cases} \begin{bmatrix} x^d_c & y^d_c & 1 \end{bmatrix}^T = \mathbf{A}^{-1} \tilde{\mathbf{p}} \\ \begin{bmatrix} x^u_c & y^u_c \end{bmatrix}^T = \begin{bmatrix} \mathbf{R}_{2 \times 3} & \mathbf{t}_{2 \times 1} \end{bmatrix} \tilde{\mathbf{P}} \end{cases} \tag{30}$$

Then the second non-linear optimization can be performed to refine all the parameters with the following cost function:

$$F_2 = \underset{\mathbf{A}, \mathbf{R}, \mathbf{t}, \mathbf{kp}}{\text{argmin}} \sum_j \sum_k \left\| \mathbf{p}_{jk} - \hat{\mathbf{p}}(\mathbf{A}, \mathbf{R}_j, \mathbf{t}_j, \mathbf{kp}, \mathbf{P}_{jk}) \right\|^2. \tag{31}$$

Here, $\hat{\mathbf{p}}$ is a projection of the feature point $\mathbf{P}_{jk}$ according to Eqs. (3) and (29). After F_2 is minimized, the calibration of a telecentric camera is thoroughly completed.

4. Stereo matching and 3D reconstruction

Another complex task within a stereo vision system is the stereo matching, which in our technique is accomplished with the help of fringe projection technique. The structure model of the system is illustrated in Fig. 4(a). Sinusoidal patterns encoded by horizontally increased phase map are projected in sequence. The fringes are deformed by the object and then captured by the cameras. The captured fringe image can be expressed as

$$I(u, v) = A(u, v) + B(u, v) \cos[\varphi(u, v)]. \tag{32}$$

Here, A is the background item, B is the modulation and $\varphi(u, v) = 2\pi u f_u + \phi_\delta(u, v)$. f_u is the carrier frequency and ϕ_δ is the phase value modulated by the object. By phase-shifting technique [25] and multi-frequency phase unwrapping [26], φ from both cameras can be obtained and used to help the stereo matching after epipolar rectification.

Telecentric epipolar rectification is performed on the unwrapped phase maps using Eq. (14) after the cameras' lens distortion is compensated with the calibrated parameters [27]. Without loss of generality, the left camera is considered as the main camera. It can be seen from Fig. 4(b), for a pixel (u'_L, v'_L) on the left camera with phase value $\varphi(u'_L, v'_L)$, the task is to find the corresponding pixel u'_R in the v'_L th row on the right image. Because the fringe direction is vertical so that the unwrapped phase value increases along the horizontal direction. The integral pixel u'^I_R that has the closest phase value to $\varphi(u'_L, v'_L)$ in the v'_L th row is firstly obtained with its phase being $\varphi(u'^I_R)$. Then sub-pixel coordinate u'_R is thereby calculated based on linear interpolation:

$$u'_R = u'^I_R + \begin{cases} \dfrac{\varphi(u'_L, v'_L) - \varphi(u'^I_R)}{\varphi(u'^I_R + 1) - \varphi(u'^I_R)}, \varphi(u'_L, v'_L) > \varphi(u'^I_R) \\[12pt] \dfrac{\varphi(u'_L, v'_L) - \varphi(u'^I_R)}{\varphi(u'^I_R) - \varphi(u'^I_R - 1)}, \varphi(u'_L, v'_L) \le \varphi(u'^I_R) \end{cases} \tag{33}$$

After completing the stereo matching, we have the pixel pair maps $(\mathbf{p}'_L, \mathbf{p}'_R)$. For more accurate sub-pixel searching, u'_R can be interpolated with more complex fitting method that involves more neighboring pixels but more time will be consumed. Together with the new camera parameters after the epipolar rectification, the world coordinates $\mathbf{P}$ can be directly calculated from Eq. (7) with least square method.

5. Experiments and discussion

5.1. Simulation of distortion center estimation

In the simulation, the controlled variables contain the distortion center (u_0, v_0) and different combinations of distortion coefficients (k_1, k_2, k_3). The simulated image size is 500×600 with pixel size of $4 \ \mu m$. For (u_0, v_0), we simulated five positions along the diagonal line of the image with the distance from the image center increased: (300,250), (240,200), (180,150), (120,100), and (60,50). At each position, three distortion situations are simulated (only k_1, only k_1, k_2, and k_1, k_2, k_3). The simulated coefficients are $k_1 = 3.5 \times 10^{-4}$, $k_2 = 2 \times 10^{-5}$, and $k_3 = 2.5 \times 10^{-6}$, which are set according to common experience. Gaussian noise with the standard deviation $\sigma_n = 0.01$ is added to each image.

We first verify the effectiveness of the proposed distortion center estimation method for each situation and then the full-scale calibration parameters are derived using our method and Yao's method [18], respectively. The retrieved distortion center position and re-projection error are compared, and the simulation result is summarized in Fig. 5. As shown in Fig. 5(a) and (b), the overall trend is that the error of the estimation increases as the position deviation between the actual distortion center and the image center increases. When the deviation is
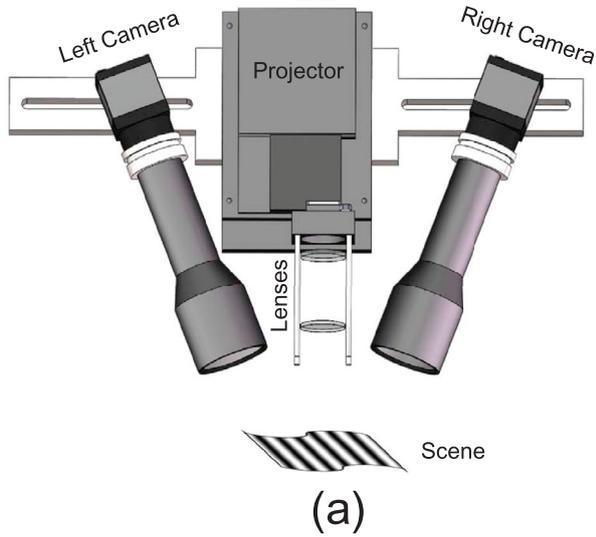
Fig. 4. (a) Simplified structure model of the system. (b) Illustration of the bilocular matching based on the unwrapped phase map.
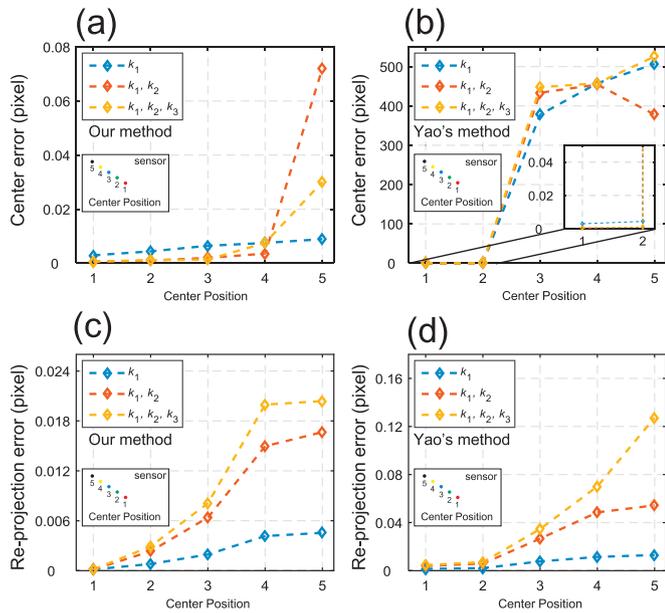

Fig. 5. Summarized result of the simulation of distortion center estimation: Retrieved center error using (a) our method and (b) Yao's method in five positions; Re-projection error using (c) our method and (d) Yao's method in five positions.
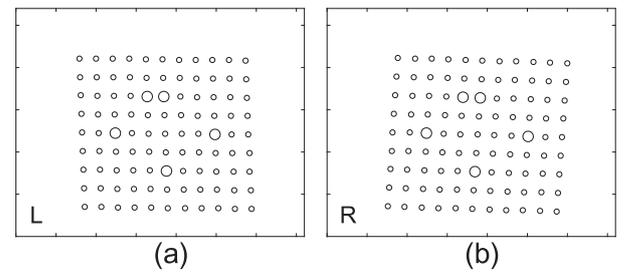

Fig. 6. One image pair of the captured patterns from (a) left camera and (b) right camera in the calibration.

field: 7 mm; Spatial resolution: 22μm). The calibration board (CC-008-G-0.65 from Calib Optics) used here is made of diffused reflection ceramic substrates. The pattern printed on it is white circles, which can be seen in Fig. 8. The center-to-center spacing between every two adjacent circles is 0.65 mm, and the diameters of big and small circles are 0.4 mm and 0.175 mm, respectively. The calibration is based on this pattern with an additional stage. Ten pairs of the calibration patterns are captured with different locations and angles. Edge detection and ellipse fitting can accurately extract these white circles' centers after binarization processing of the image. Fig. 6 shows one pair of the extracted calibration patterns.

Firstly, the distortion center is estimated for both cameras. Because the distortion of a telecentric lens is not as obvious as a pinhole lens, more calibration patterns are required to achieve a reliable estimation of the distortion center. By increasing the calibration pattern number from 1 to 10, we conclude that at least six patterns are required. However, this number is empirical, which means it only can be obtained from experiments. Then the full-scale parameters follow with distortion considered. The calibration results of the intrinsic parameters are presented in Table 1. Based on the calibration results, the re-projection errors (δ_u, δ_v) of both cameras are calculated and presented in Fig. 7. The standard deviations of the re-projection errors are (0.07859, 0.07811) for the left camera and (0.07969, 0.07554) for the right camera.

within 60 pixels like position 1 and 2, the results from both methods are very close. However, the error by using our method remains very small even though the deviation exceeds 120 pixels like position 3, 4, and 5, while the error using Yao's method becomes unstable because the initial distortion center is not close enough to the real one. Correspondingly, Fig. 5(c) and (d) reveal that if the distortion center is wrongly derived, the re-projection error can be ten times that when a reasonable distortion center is initially estimated.

5.2. Real-world camera calibration

As shown in Fig. 4(a). The system is composed of a digital projector (DLP® LightCrafter™ Evaluation Module with its original lens being replaced by a lens with much smaller magnification) and two cameras (BFS-U3-04S2M-CS, Pixel size: 6.9 μm; Resolution: 720 × 540) mounted with telecentric lenses (XF-T0.5X110, Magnification: 0.5×; Depth of

*5.3. Rectification considering len distortion*

Note that the rectification is valid only for the distortion-free images so that the captured images need to be undistorted based on the distortion coefficients [27] in Table 1 before being rectified using Eq. (14).

Table 1
Intrinsic parameters of left and right cameras.

	Estimated center (u_0, v_0)	Optimized center (u_0, v_0)	Effective magnification $m(m \times \text{Pixel size})$	Distortion coefficients $(k_1, k_2, k_3, p_1, p_2)$
L	(389.1, 259.2)	(383.0, 269.4)	72.20 (0.4982)	(−1.215 e−4, 1.852 e−5, −5.900 e−7, 4.2303 e−6, 1.821 e−5)
R	(358.3, 267.6)	(356.2, 268.6)	72.49 (0.5002)	(−1.870 e−4, 2.472 e−5, −8.122 e−7, −9.301 e−7, 6.670 e−6)
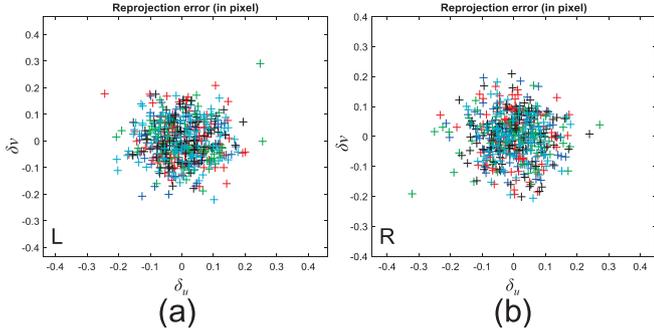
Fig. 7. Re-projection errors (δ_u, δ_v) of (a) left camera and (b) right camera.
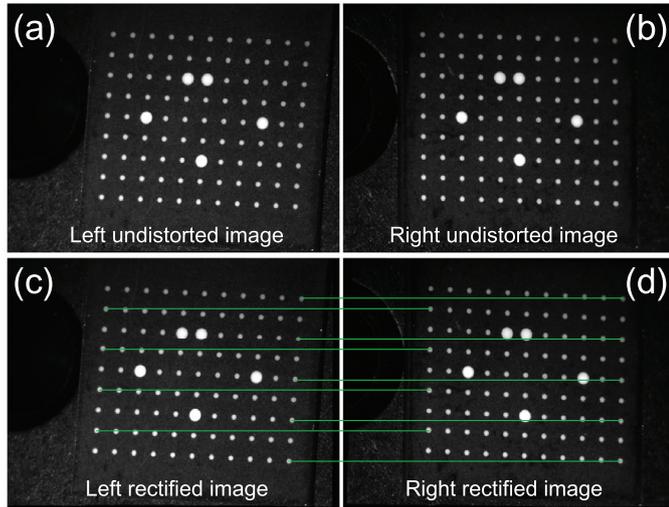
Fig. 8. Rectification example: (a) - (b) Undistorted image pair of the calibration pattern; (c) - (d) Rectified image pair of the calibration pattern.

The new derived calibrated homography matrices are as follows:

$$
\begin{cases}
\mathbf{H}'_L = \begin{bmatrix} -68.081 & -2.5486 & -23.844 & 362.54 \\ -1.3196 & 72.431 & -3.0225 & 301.20 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
\mathbf{H}'_R = \begin{bmatrix} -68.488 & -0.2762 & 22.457 & 357.90 \\ -1.3196 & 72.431 & -3.0225 & 301.20 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{cases}
\tag{34}
$$

with the same second row parameters in $\mathbf{H}'_L$ and $\mathbf{H}'_R$. Here we present one example of the rectification of one pair of images as shown in Fig. 8. The calibration board is randomly positioned and clearly that the corresponding control points are not in the same vertical position in Fig. 8(a) and (b). After the rectification using Eq. (14), all the pixels in the left and right image are strictly transformed on the same vertical position, which can be clearly verified though the horizontal lines in Fig. 8(c) and (d).
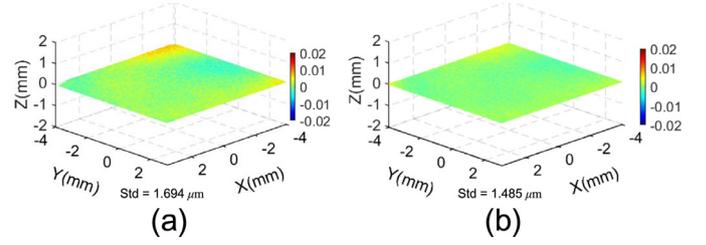
Fig. 9. Error distribution of the reconstructed result of the flat plane : (a) With distortion unremoved; (b) With distortion removed.
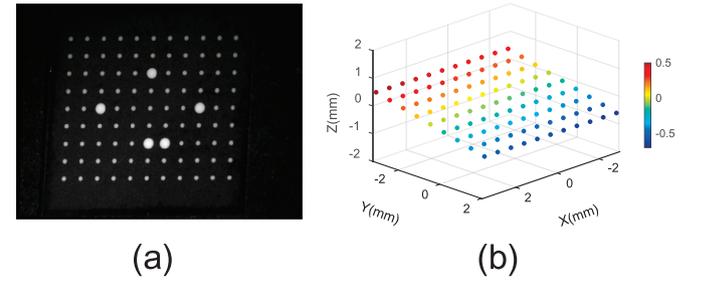

Fig. 10. 3D reconstruction of the markers' centers on the calibration board: (a) Image of the calibration board; (b) 3D distribution of the reconstructed markers' centers with the depth color-coded. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5.4. Real-world 3D reconstruction

After image rectification, the matched pixels in the image pair are in the same vertical position. With the help of the unwrapped phase maps acquired using fringe projection, the horizontal position of the matched pixel in the right camera can be uniquely determined. Then 3D reconstruction follows by solving Eq. (7). To visually show the calibration results, we conducted two 3D measurements of a flat plane and the calibration board, respectively. The flat plane is first measured. Fig. 9(a) is the result without considering the lens distortion and Fig. 9(b) is the result of the distortion removed. The standard deviation of both results is calculated to reflect the accuracy quantitatively. The error distribution in the outside area in Fig. 9(a) is larger than that in Fig. 9(b) because of the lens distortion. Although the standard deviation only improves 0.209 μm, the difference between Fig. 9(a) and (b) reveals that the distortion is effectively removed. Note that when conducting the calibration step, one should render the pattern fill as more field of view of the camera as possible. If only the center part of the field of view observes the patter, the distortion can be hardly retrieved precisely.

Another experiment measuring the absolute 3D position of the markers' centers on the calibration board is conducted to further verify the reconstruction accuracy. The pattern on the board has white circle spots with their centers equally distributed like a 2D comb function as shown in Fig. 10(a). The centers' pixel coordinates can be extracted and their absolute 3D distribution are reconstructed as presented in Fig. 10(b) with their depths being color-coded. The distance between two neighboring centers is calculated by $d = \left\| (x_1, y_1, z_1) - (x_2, y_2, z_2) \right\|^2$, where, d is the distance between two points: (x_1, y_1, z_1) and (x_2, y_2, z_2) in space.

(a)   (b)   (c)   (d)
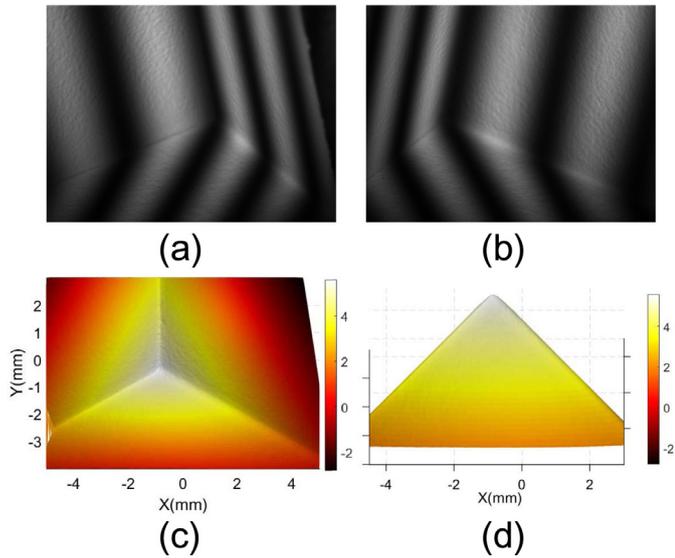
Fig. 11. 3D reconstruction a cube's corner with mutually perpendicular sides: (a) - (b) Fringe images from two cameras; (c) Top view of the reconstructed profile of the corner; (d) Rotated view to examine the perpendicularity.

There are 90 pairs of neighboring centers in *X* direction and 88 pairs in *Y* direction. The calculated distances between the reconstructed centers have the mathematical expectation of 0.6506 mm and the RMSE of 0.0014 mm, which is pretty close to the standard data (0.65 mm, 0.0015 mm) of the calibration board.

Benefiting from the increased depth of field of telecentric lenses, the measurable volume of our system extends significantly. As shown in Fig. 11, a corner of a cube with side length of 10 mm is measured. Fig. 11(a) and (b) are the fringe images from the left camera and the right camera, respectively. A top view of the reconstructed profile of the corner is given in Fig. 11(c), from which we can see that the measurable depth easily achieves the depth of field of the lens (7 mm). As we know that any two sides of a cube's corner are mutually perpendicular, we rotate the reconstructed 3D profile as shown in Fig. 11(d) to examine this property. The results from the two experiments manifest that both the calibration of the system and 3D reconstruction process are successful.

So far we have proved that the calibration of our system and the 3D reconstruction is feasible based on the method described in Section 4. Furthermore, our system can be applied in quality inspection of industrial products. The first sample we measured is one side of an integrated circuit chip (package: QFP, pitch: 0.65mm). As shown in Fig. 12(a), the 4th and 13th pin from left to right are out of shape and a slightly higher than the others. Usually, we cannot easily detect these defects because of the large pin number. The 3D data of the target is reconstructed as shown in Fig. 12(b), from which we can easily find out the deformed pins. The right side view of the data in Fig. 12(c) provides a more intuitive observation. Quantitatively, we plot three 3D lines corresponding to the 4th (red), the 9th (green), and the 13th (blue) pin as shown in Fig. 12(d) for better observation.

Another application of our technique shown here is detecting the electronic circuit and metal pads on a PCB (Printed Circuit Board). The proper performance of PCB is assured by unerringly printing the electronic circuit and the metal pads on the substrate. Any small fault may cause irreparable harm. Traditional inspection methods mainly rely on 2D images, which cannot reveal the height information of the printed metal. Here we give the measured 3D results of some wires and pads on a PCB of a graphics card as shown in Fig. 13(a) and (b). The normal thickness of the printed copper is 1 ounce (1 ounce here represents the thickness of 1 ounce of copper rolled out to an area of 1 square foot,
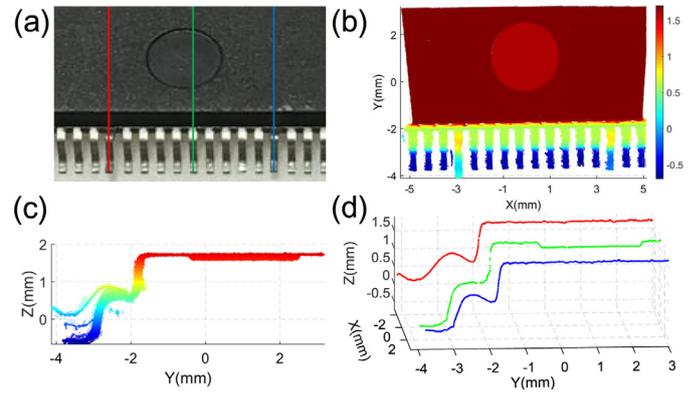
Fig. 12. Application of the quality inspection of one side of an integrated circuit chip: (a) Image of the measured area; (b) Reconstructed 3D data of the target; (c) Right side view of (b); (d) Three 3D lines corresponding to the 4th (red), the 9th (green), and the 13th (blue) pin in (a) from left to right. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
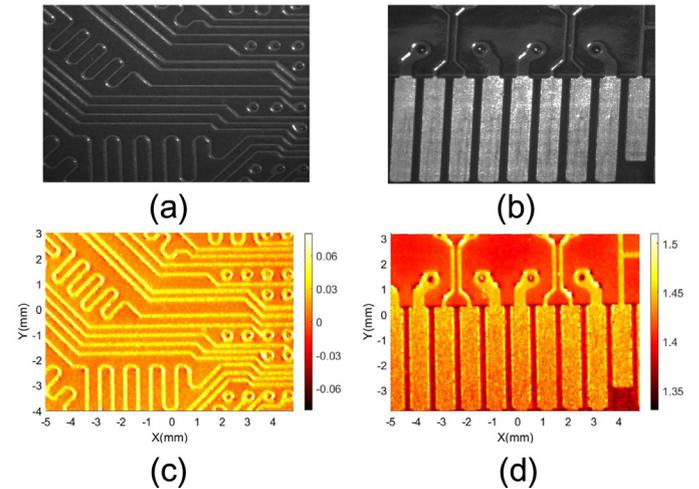
Fig. 13. Application of detecting the electronic circuit and metal pads on a PCB: (a) - (b) Images of the measured areas. (c) - (d) Reconstructed 3D profile of the targets.

Table 2
Information on used patterns to calculate phase maps.

Items	Values		
i (Fringe set)	1	2	3
N_i (Phase-shifting step)	2	2	8
P_i (Pixel number)	608	96	12
K_i (Frequency ratio)	1	6.3	8
Period number	1	6.3	50.6

equals $\sim 36\mu$mm), which can be examined by the results in Fig. 13(c) and (d). Note that the rightmost pad in Fig. 13(b) is deliberately designed shorter than the others to ensure safe power during hot-swap operation, and this property is also observable in Fig. 13(d).

The number of images used to calculate phase maps is 12. Specific information on the projected fringe patterns is as shown in Table 2. More fringe patterns used, the more accurate the 3D result is [28]. To capture the images within less time, we use trigger wires to electrically connect the projector and the cameras to provide a time-saving image acquisition process. All the settings of the cameras are the same to make sure the phase maps based stereo matching is credible. The total time used to get one final 3D model is typically around 1 second in Matlab,

depending on the computing power. By using a laptop with an Intel Core i7-4720 CPU and Matlab R2015b, the time used for Fig. 12 is 1361.3 ms.

6. Conclusions

In this paper, we present a new microscopic telecentric stereo vision system with its calibration and applications. The complexity of the dense stereo matching process is greatly simplified by applying accurate telecentric rectification for the stereo vision system. However, the rectification is valid only if the lens distortion is correctly compensated. From the simulation result of the distortion center estimation, we find that when the optical center is not close enough to the detector's center, local minima problem can happen in the iteration-based optimization and lead to inaccurate solutions. In our proposed calibration framework, the initial guesses of distortion center and coefficients are effectively estimated before the iteration, which significantly reduces the possibility of being trapped in local minima, and thus the robustness of the calibration of telecentric cameras is improved considerably.

Since the projector does not need calibrating, the calibration of the system is quite convenient and efficient. Small FOV of the projector is realized by using two low-cost lenses, and the FOV can also be easily adjusted by changing the relative position of the lenses. Based on the rectification of the telecentric stereo vision system, sub-pixel matching is realizable with the help of the phase maps generated from the fringe projection technique. The measurement uncertainty is determined by the noise level of the phase maps and also depends on the interpolation method used for stereo matching. Our approach is resistant to slight gamma effect since it affects the phase values of both cameras simultaneously. The noise in the phase maps can be decreased by using denser fringe patterns and more substantial phase-shifting steps, but more patterns are required. Also, the interpolation method involving more data would give more precise matching result but more computationally costly.

The overall process of the calibration for our system is detailed, and the 3D reconstruction is demonstrated in the applications for microsurface profilometry. The experimental results show that our technique can be successfully applied in industrial applications, such as quality control and on-line inspection for micro-scale products with a measurable volume of $10\,\mathrm{mm}\,(L) \times 7\,\mathrm{mm}\,(W) \times 7\,\mathrm{mm}\,(H)$. It is worth mentioning that the system also has the potential for high-speed 3D measurement [29] provided that the unwrapped phase maps can be obtained using fewer images, which is the possible direction for future work.

Funding

Acknowledgments

References

[1] Lazaros N. Review of stereo vision algorithms: from software to hardware. Int J Optomechatr 2008;2(4):435–62. doi:10.1080/15599610802438680.

[2] Tippetts B, Lee DJ, Lillywhite K, Archibald J. Review of stereo vision algorithms and their suitability for resource-limited systems. J Real-Time Image Process 2016;11(1):5–25.

[3] Tao T, Chen Q, Da J, Feng S, Hu Y, Zuo C. Real-time 3-d shape measurement with composite phase-shifting fringes and multi-view system. Opt Express 2016;24(18):20253–69.

[4] Li Z, Zhong K, Li YF, Zhou X, Shi Y. Multiview phase shifting: a full-resolution and high-speed 3d measurement framework for arbitrary shape dynamic objects. Opt Lett 2013;38(9):1389–91.

[5] Zhang S. Recent progresses on real-time 3d shape measurement using digital fringe projection techniques. Opt Lasers Eng 2010;48(2):149–58.

[6] Zhang Z. A flexible new technique for camera calibration. Pattern Anal Mach Intell IEEE Trans Pattern Anal Mach Intell 2000;22(11):1330–4.

[7] Zhang S, Huang PS. Novel method for structured light system calibration. Opt Eng 2006;45(8). 083601–083601

[8] Hu Y, Chen Q, Tao T, Li H, Zuo C. Absolute three-dimensional micro surface profile measurement based on a greenough-type stereomicroscope. Meas Sci Technol 2017;28(4):045004.

[9] Quan L. Self-calibration of an affine camera from multiple views. Int J Comput Vis 1996;19(1):93–105.

[10] Li D, Liu C, Tian J. Telecentric 3d profilometry based on phase-shifting fringe projection. Opt Express 2014;22(26):31826–35.

[11] Li B, Zhang S. Flexible calibration method for microscopic structured light system using telecentric lens. Opt Express 2015;23(20):25795–803.

[12] Yin Y, Wang M, Gao BZ, Liu X, Peng X. Fringe projection 3d microscopy with the general imaging model. Opt Express 2015;23(5):6846–57.

[13] Liu H, Lin H, Yao L. Calibration method for projector-camera-based telecentric fringe projection profilometry system. Opt Express 2017;25(25):31492–508. doi:10.1364/OE.25.031492.

[14] Wang M, Yin Y, Deng D, Meng X, Liu X, Peng X. Improved performance of multi-view fringe projection 3d microscopy. Opt Express 2017;25(16):19408–21.

[15] Li D, Tian J. An accurate calibration method for a camera with telecentric lenses. Opt Lasers Eng 2013;51(5):538–41.

[16] Chen Z, Liao H, Zhang X. Telecentric stereo micro-vision system: calibration method and experiments. Opt Lasers Eng 2014;57:82–92.

[17] Rao L, Da F, Kong W, Huang H. Flexible calibration method for telecentric fringe projection profilometry systems. Opt Express 2016;24(2):1222–37.

[18] Yao L, Liu H. A flexible calibration approach for cameras with double-sided telecentric lenses. Int J Adv Rob Syst 2016;13(3):82. doi:10.5772/63825.

[19] Watanabe M, Nayar SK. Telecentric optics for focus analysis. IEEE Trans Pattern Anal Mach Intell 1997;19(12):1360–5.

[20] Hartley R, Zisserman A. Multiple view geometry in computer vision. Cambridge University Press; 2003.

[21] Kudryavtsev AV, Dembl S, Piat N. Stereo-image rectification for dense 3d reconstruction in scanning electron microscope. In: International Conference on Manipulation, Automation and Robotics at Small Scales; 2017. p. 1–6.

[22] Liu H, Zhu Z, Yao L, Dong J, Chen S, Zhang X, et al. Epipolar rectification method for a stereovision system with telecentric cameras. Opt Lasers Eng 2016;83:99–105.

[23] Hartley R, Kang SB. Parameter-free radial distortion correction with center of distortion estimation. IEEE Trans Pattern Anal Mach Intell 2007;29(8):1309–21.

[24] Willson RG, Shafer SA. What is the center of the image? JOSA A 1994;11(11):2946–55.

[25] Zhang S, Yau S-T. High-resolution, real-time 3d absolute coordinate measurement based on a phase-shifting method. Opt Express 2006;14(7):2644–9.

[26] Zuo C, Huang L, Zhang M, Chen Q, Asundi A. Temporal phase unwrapping algorithms for fringe projection profilometry: a comparative review. Opt Lasers Eng 2016;85:84–103.

[27] Huang L, Chua PS, Asundi A. Least-squares calibration method for fringe projection profilometry considering camera lens distortion. Appl Opt 2010;49(9):1539–48.

[28] Zuo C, Feng S, Huang L, Tao T, Yin W, Chen Q. Phase shifting algorithms for fringe projection profilometry: a review. Opt Lasers Eng 2018;109:23–59.

[29] Zuo C, Tao T, Feng S, Huang L, Asundi A, Chen Q. Micro fourier transform profilometry (μftp): 3d shape measurement at 10,000 frames per second. Opt Lasers Eng 2018;102:70–91.