2016 IEEE International Conference on Automation Science and Engineering (CASE)
Fort Worth, TX, USA, August 21-24, 2016

# Calibration with Ball Objects for Telecentric Micro-Vision System

Fangbo Qin, De Xu, Fei Shen, Dapeng Zhang, Xilong Liu

***Abstract*—The calibration of the micro-vision system is an fundamental issue in the micro-vision based measurement and control areas. The calibration methods for conventional macro-vision system are not suitable for telecentric micro-vision system. This paper proposes a calibration method with ball objects for the telecentric micro-vision system, which is used for the 3-dimensional relative pose measurement. The ball objects are low-cost and easy to be made. The intrinsic parameters and depth of field of camera are calibrated with a one-ball object. The angle between the camera's optical axis and the linear stage's axis is considered, and the image offset matrix caused by this angle is calibrated using the one-ball object. Furthermore, the image Jacobian matrices of two cameras are obtained by calibrating the two cameras' rotation transformation using a two-ball object. The experimental results verify the calibration methods' effectiveness and accuracy.**

***Index Terms*—Calibration, microscopic vision, vision measurement.**

## I. Introduction

Micro-vision has been widely applied in the areas of microassembly, micromanipulation and 3-dimensional (3D) measurement, providing non-contact and high-resolution measurement of micro-scale objects [1-4]. Microscopes have the advantages of high magnification and neglectable lens distortion, and the limitations of small field of view (FOV) and small depth of field (DOF). Moreover, telecentric microscope, whose magnification does not change with object depth variations in a specific range, is widely used in 3D object measurement [5]. Monocular micro-vision is suitable for the measurement in a planar space, and two or more microscopic cameras can obtain the object's 3D information.

The calibration methods for micro-vision system are different from those for macro-scale vision, which are generally classified into two categories: pattern based calibration and active motion based calibration. The main differences between these methods lie in how to construct scene points in the world space. Li et al. [6] used a target plate with precise circle pattern to calibrate the telecentric camera's intrinsic parameters. Chen et al. [7] proposed a grid-plate based calibration method for a telecentric stereo micro-vision system, which calibrates both the intrinsic and extrinsic parameters. Gorpas et al. [8] calibrated the binocular telecentric vision system using a 3D object with dot-grids on its two perpendicular planes. Shen et al. [9] calibrated the telecentric cameras in a microassembly system based on manipulator's active motions and a nonlinear damping least squares method. In [10], an active motion based calibration method was designed for the microscopic camera, which included only two steps of motion of the tube's tip on the focal plane. However, the precision manufacturing of the 2D or 3D patterns with small sizes are very costly and difficult. The active motion based calibration methods are only suitable for the systems with manipulators.

Because microscope is with shallow DOF, it is necessary to mount the camera on a *linear stage* to adjust the object distance to acquire a clear image. The linear stage's axis is approximately parallel to the camera's optical axis, and is called *focus axis* in this paper. Autofocus techniques enable a camera to focus on features of interest. However, if the camera's optical axis and the focus axis are not strictly parallel, the camera's translations along the focus axis will lead to the image feature's offset. In addition, it is very helpful to obtain the actual DOF of a camera whose magnification or aperture size is adjustable, because the DOF is dependent on its magnification and aperture size.

The motivation of this work is to develop a low-cost calibration method for a 3D relative pose measurement system with two telecentric microscopic cameras. The calibration objects include a one-ball object and a two-ball object, which are with low costs and small sizes. The cameras' DOFs, intrinsic parameters, and image offset matrices are calibrated with the one-ball object. The rotation transformation between the two camera frames is calibrated using the two-ball object. The image Jacobian matrices, which are used for 3D vector measurement, are obtained from the intrinsic parameters and the rotation matrix. The proposed calibration methods are verified on a vision system with a high precision manipulator.

The rest of this paper is organized as follows. Section II describes the imaging model, the principle of DOF, and the pose measurement method based on image Jacobian matrix. In Section III, the calibration methods for DOF, image offset matrix, intrinsic parameters, and image Jacobian matrix are addressed. The experiments and results are presented in Section IV. Finally, the paper is concluded in Section V.

## II. Vision System Model

### A. Telecentric Camera Model

The object-side telecentric lens is with an aperture stop at

This work is supported in part by the National Natural Science Foundation of China under Grant 61227804, 61421004, and 61503378.

F. Qin, D. Xu, F. Shen, D. Zhang and X. Liu are with the Research Center of Precision Sensing and Control, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: de.xu@ia.ac.cn).

its image-side focal point, and only the ray cones whose principal axis is parallel to the optical axis are collected by the lens pupil. Therefore, the telecentric camera's projection center lies at infinity, and the imaging model is independent of object depth. Ignoring the distortion, the imaging model is

$$\boldsymbol{p}=\begin{bmatrix}u\\v\end{bmatrix}=\begin{bmatrix}\alpha & \gamma\\0 & \beta\end{bmatrix}\begin{bmatrix}{}^{C}x\\{}^{C}y\end{bmatrix}+\begin{bmatrix}u_0\\v_0\end{bmatrix} \tag{1}$$

where ${}^{C}\boldsymbol{P}=[{}^{C}x,\ {}^{C}y,\ {}^{C}z]^{T}$ is a scene point in the camera frame {C}, and $\boldsymbol{p}=[u, v]^{T}$ is its image point. $\alpha$, $\beta$, $\gamma$, $u_0$, and $v_0$ are the intrinsic parameters. $\boldsymbol{p}_{\perp}=[{}^{C}x,\ {}^{C}y]^{T}$ is defined as the scene point's orthographic projection on the XOY plane of {C}. Considering the transformation from the world frame {W} to {C}, we have

$$\boldsymbol{p}_{\perp}=\begin{bmatrix}r_{11} & r_{12} & r_{13}\\r_{21} & r_{22} & r_{23}\end{bmatrix}{}^{W}\boldsymbol{P}+\begin{bmatrix}t_1\\t_2\end{bmatrix} \tag{2}$$

where ${}^{W}\boldsymbol{P}=[{}^{W}x,\ {}^{W}y,\ {}^{W}z]^{T}$ is a scene point in the world frame {W}. $[r_{11}, r_{12}, r_{13}]$ and $[r_{21}, r_{22}, r_{23}]$ are the first two rows of the rotation matrix ${}^{C}\boldsymbol{R}_{W}$. $t_1$ and $t_2$ are the first two elements of the translation vector ${}^{C}\boldsymbol{t}_{W}$. Finally, the telecentric camera model is

$$\boldsymbol{p}=\begin{bmatrix}\alpha & \gamma\\0 & \beta\end{bmatrix}\begin{bmatrix}r_{11} & r_{12} & r_{13}\\r_{21} & r_{22} & r_{23}\end{bmatrix}{}^{W}\boldsymbol{P}+\begin{bmatrix}\alpha & \gamma\\0 & \beta\end{bmatrix}\begin{bmatrix}t_1\\t_2\end{bmatrix}+\begin{bmatrix}u_0\\v_0\end{bmatrix} \tag{3}$$

Note that the telecentric camera's projection center is immaterial. Therefore, the translation vector ${}^{C}\boldsymbol{t}_{W}$ is arbitrary, which determines image center $[u_0, v_0]^{T}$. In this paper, the ${}^{C}\boldsymbol{t}_{W}$ is established subject to $u_0=v_0=0$.

### *B. Depth of Field*

Because of lens aberrations and diffraction, the practical image of a scene point at the focus depth is not a point but a small spot, which is called the least blur circle. If the scene point is defocused, the blur circle's diameter increases gradually with the distance from the scene point to the focus plane. The blur circle is described by the point spread function (PSF). A practical image of an object's part at the same depth is the convolution of the ideally sharp image with the PSF. A 2D Gaussian model can be used to approximate the actual PSF [11], whose standard deviation $\sigma$ determines the blur circle's diameter and is dependent on the object depth. The depth range providing acceptable blur circle diameters is called the depth of field. In the autofocus process, the microscopic camera is translated along the focus axis to search for the best focus position. The DOF is an important reference for selecting the searching step.

### *C. 3D Relative Pose Measurement*

The object's 3D position changes and orientation changes can be indicated by the 3D vectors in the world frame. Two scene points $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$ form a 3D vector $\Delta\boldsymbol{P}=\boldsymbol{P}_2-\boldsymbol{P}_1$ in the world space, and their image points $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$ form a 2D vector $\Delta\boldsymbol{p}=\boldsymbol{p}_2-\boldsymbol{p}_1$ in the image space. If the telecentric camera's pose is fixed, we have the relationship

$$\Delta\boldsymbol{p}=\boldsymbol{J}\Delta\boldsymbol{P}=\begin{bmatrix}\alpha & \gamma\\0 & \beta\end{bmatrix}\begin{bmatrix}r_{11} & r_{12} & r_{13}\\r_{21} & r_{22} & r_{23}\end{bmatrix}\Delta\boldsymbol{P} \tag{4}$$

where $\boldsymbol{J}$ is the camera's image Jacobian matrix, which is determined by the camera's intrinsic parameters and attitude.

As mentioned, the telecentric microscope has a small DOF, so that the camera might have to be translated along the focus axis to two different position $l_1$ and $l_2$ to acquire the clear images of $\boldsymbol{p}_1$ and $\boldsymbol{p}_2$, respectively. The camera's translation leads to the same scene point's image offset

$$\boldsymbol{\delta}(\Delta l)=\begin{bmatrix}\alpha & \gamma\\0 & \beta\end{bmatrix}\begin{bmatrix}\tau_1\Delta l\\\tau_2\Delta l\end{bmatrix}=\begin{bmatrix}\delta_1\\\delta_2\end{bmatrix}\Delta l \tag{5}$$

where $\Delta l=l_2-l_1$ and $[\delta_1, \delta_2]^{T}$ is called the image offset matrix. $\tau_1$ and $\tau_2$ indicate the focus axis's direction in {C}. If the focus axis is strictly parallel with the optical axis, $\tau_1$ and $\tau_2$ both equal zero. Therefore, considering both the world vector $\Delta\boldsymbol{P}$ and the camera translation $\Delta l$, the image vector is

$$\Delta\boldsymbol{p}=\boldsymbol{J}\Delta\boldsymbol{P}+\boldsymbol{\delta}(\Delta l) \tag{6}$$

With n≥2 telecentric cameras observing the same 3D vector $\Delta\boldsymbol{P}$ from different directions, we have

$$\begin{bmatrix}\Delta\boldsymbol{p}_1-\boldsymbol{\delta}_1(\Delta l_1)\\\Delta\boldsymbol{p}_2-\boldsymbol{\delta}_2(\Delta l_2)\\\vdots\\\Delta\boldsymbol{p}_n-\boldsymbol{\delta}_n(\Delta l_n)\end{bmatrix}=\begin{bmatrix}\boldsymbol{J}_1\\\boldsymbol{J}_2\\\vdots\\\boldsymbol{J}_n\end{bmatrix}\Delta\boldsymbol{P} \tag{7}$$

where $\boldsymbol{J}_i$, $\boldsymbol{\delta}_i(\Delta l)$, $\Delta\boldsymbol{p}_i$, and $\Delta l_i$ are the image Jacobian matrix, image offset, image vector and translation length of the camera $i$ ($i$=1,2,…,$n$). Eq. (7) can be solved by the linear least squares method (LSM).

$$\Delta\boldsymbol{P}=\begin{bmatrix}\boldsymbol{J}_1\\\boldsymbol{J}_2\\\vdots\\\boldsymbol{J}_n\end{bmatrix}^{\dagger}\begin{bmatrix}\Delta\boldsymbol{p}_1-\boldsymbol{\delta}_1(\Delta l_1)\\\Delta\boldsymbol{p}_2-\boldsymbol{\delta}_2(\Delta l_2)\\\vdots\\\Delta\boldsymbol{p}_n-\boldsymbol{\delta}_n(\Delta l_n)\end{bmatrix} \tag{8}$$

where † is the pseudo-inverse symbol.

## III. CALIBRATION METHODS

### *A. DOF Calibration*

The DOF can be calibrated by actively translating the camera along the focus axis, and identifying the camera position range $[l_n, l_f]$ providing the acceptable blur circle diameter $\Phi\le\Phi_t$, where $\Phi_t$ is the maximum acceptable blur circle diameter. Then the DOF is approximately the range length of camera position. The main problem in DOF calibration is to calculate the blur circle diameter $\Phi$ of an image.

A one-ball object is used as the calibration object, whose silhouette is a circle with sharp step edge. The ball image's edge is the convolution of the step edge with a 2D Gaussian PSF, as shown in Fig 1(a). The blurred edge's width is called edge spread width (ESW), which is equal to the blur circle diameter $\Phi$.

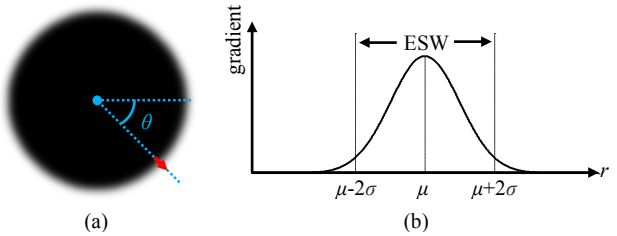



Fig. 1. Ball image and gray gradient distribution near its egde. (a) The ball image with blurred edges; (b) The gray gradient distribution along the line segment which is labeled by the red arrow in (a).

Since 2D Gaussian PSF is isotropic, the 1D gradient distribution near the edge along the circle's radial direction, as shown in Fig. 1(b), is a 1D Gaussian model

$$G(r)=\frac{a}{\sigma}e^{-(r-\mu)^2/2\sigma^2} \tag{9}$$

where $a$ determines the intensity contrast, $r$ is the radial distance to the circle center, $\mu$ indicates the ideal sharp edge's radial distance to the circle center, and the standard deviation $\sigma$ determines the ESW by

$$w_{es}=4\sigma \tag{10}$$

so that 95.45% of the light intensity lies within the zone $[\mu\text{-}w_{es}/2, \mu+ w_{es}/2]$.

To obtain the ESW from an image, the image is first smoothed by a 2D Gaussian kernel with the standard deviation $\sigma_0$. Then $m$ local line segments $L_{si}$ ($i$=1,2,…,$m$) near the ball edges are sampled, which are along the radial directions. The direction angle of $L_{si}$ is $\theta_i$, as shown in Fig. 2(a). $n$ image points are sampled with equivalent interval on the $L_{si}$, whose gray values are $I_i$ ($r_j$) ($i$=1,2,…,$n$). The gray gradients $g_i(r_j)$ can be calculated by centered difference of these gray values. Then the 1D Gaussian model of $L_{si}$ can be obtained from these gray gradient samples based on the maximum-likelihood estimation. The average value of the standard deviation is

$$\bar{\sigma}=\frac{1}{m}\sum_{i=1}^{m}\hat{\sigma}_i \tag{11}$$

where $\hat{\sigma}_i$ is the estimated standard deviation of $L_{si}$. Finally, considering the smoothing convolution at the beginning, the estimated ESW of the raw image is $4\sqrt{\bar{\sigma}^2-\sigma_0^2}$.

### B. Intrinsic Parameters Calibration

A one-ball object is placed in the camera's DOF and FOV. Its projection on the orthographic projection on the XOY plane of {C} is a circle $C_\perp$,

$$(x-x_c)^2+(y-y_c)^2=r_b^2 \tag{12}$$

where $(x_c, y_c)$ is the circle's center, and $r_b$ is the ball's radius. According to Eq. (1), the points $(x, y)$ on $C_\perp$ is transformed to the image points $(u, v)$ by

$$\begin{cases} u=\alpha x+\gamma y, \\ v=\beta y. \end{cases} \tag{13}$$

The ball's image is the affine transformation of $C_\perp$, generally an ellipse

$$Au^2+Buv+Cv^2+Du+Ev+F=0 \tag{14}$$

where $B^2$-4$AC$<0. The ellipse equation can be obtained from the ball's image. Canny edge detection algorithm is used to get the edge points on the ellipse. Then the ellipse equation is fitted by these edge points based on LSM.

Substituting Eq. (13) into Eq. (14), we have

$$\begin{aligned} &A\alpha^2x^2+(2A\alpha\gamma+B\alpha\beta)xy+(A\gamma^2+B\beta\gamma+C\beta^2)y^2 \\ &+D\alpha x+(D\gamma+E\beta)y+F=0 \end{aligned} \tag{15}$$

Based on the equivalence of Eq. (12) and Eq. (15), the intrinsic parameters are obtained by

$$\begin{cases} \alpha=(1/r_b)\sqrt{(CD^2+AE^2-BDE)/(4A^2C-AB^2)-F/A}, \\ \beta=\alpha\sqrt{4A^2/(4AC-B^2)}, \\ \gamma=-B\beta/2A. \end{cases} \tag{16}$$

### C. Image Offset Matrix Calibration

A one-ball object is placed in the camera's FOV and DOF. The ball is fixed, and the camera is translated to n≥2 positions along the focus axis. When the camera's position is $l_i$, the ball center's image position is $(u_i, v_i)^T$ ($i$=1,2,…,n). Then we have,

$$\underbrace{\begin{bmatrix}\delta_1 & c_1\\ \delta_2 & c_2\end{bmatrix}}_{J_o}\underbrace{\begin{bmatrix}l_1 & l_2 & \cdots & l_n\\ 1 & 1 & \cdots & 1\end{bmatrix}}_{L}=\underbrace{\begin{bmatrix}u_1 & u_2 & \cdots & u_n\\ v_1 & v_2 & \cdots & v_n\end{bmatrix}}_{U} \tag{17}$$

where $c_1$ and $c_2$ are the constants. Using the linear LSM, the solution of Eq. (17) is $\boldsymbol{J}_o=\boldsymbol{U}\boldsymbol{L}^T(\boldsymbol{L}\boldsymbol{L}^T)^{-1}$. The image offset matrix $[\delta_1, \delta_2]^T$ is the first column of $\boldsymbol{J}_o$.

### D. Image Jacobian Matrices Calibration

A two-ball object is used in the image Jacobian matrices calibration. The two balls $\Omega_1$ and $\Omega_2$ are tangent to each other and their radius $R_1$ and $R_2$ are known, as depicted in Fig. 2. The 3D vector $\Delta\boldsymbol{P}_b=[\Delta x_b, \Delta y_b, \Delta z_b]^T$ pointing from $\Omega_1$'s center $\boldsymbol{P}_{b1}$ to $\Omega_2$'s center $\boldsymbol{P}_{b2}$ is formed in the world frame {W}, whose length is $L_b$= $R_1$+$R_2$. The camera focuses on $\Omega_i$'s outline at the position $l_{bi}$, and the ball center's image position is acquired as $\boldsymbol{p}_{bi}$ ($i$=1,2). Considering the camera's translation, the image projection of $\Delta\boldsymbol{P}_b$ is $\Delta\boldsymbol{p}_b=[\Delta u_b, \Delta v_b]^T$= $\boldsymbol{p}_{b2}$-$\boldsymbol{p}_{b1}$-$\boldsymbol{\delta}$($l_{b2}$-$l_{b1}$). The orthographic projection of $\Delta\boldsymbol{P}_b$ is given by

$$\Delta \boldsymbol{p}_{\perp b} = \begin{bmatrix} \Delta x_b \\ \Delta y_b \end{bmatrix} = \begin{bmatrix} \alpha & \gamma \\ 0 & \beta \end{bmatrix}^{-1} \begin{bmatrix} \Delta u_b \\ \Delta v_b \end{bmatrix} \tag{18}$$

The absolute value $|\Delta z_b|$ equals $\sqrt{L_b^{\,2} - \Delta x_b^{\,2} - \Delta y_b^{\,2}}$. Its sign can be determined by the camera's best-focus positions $l_{b1}$ and $l_{b2}$. As shown in Fig. 2, if the camera focuses on the ball outline, the ball center lies on the focus plane. The focus axis's positive direction has a small or zero angle with the $z$-axis of {C}. Therefore, if $l_{b1}<l_{b2}$, the depth of $\boldsymbol{P}_{b2}$ is relatively larger than $\boldsymbol{P}_{b1}$, and vice versa. Note that if $|\Delta z_b|$ is smaller than the DOF, it is not possible to distinguish the two ball centers' depth. Thus, the relative depth is given by

$$\Delta z_b = \begin{cases} \sqrt{L_b^{\,2} - \Delta x_b^{\,2} - \Delta y_b^{\,2}}, & \text{if } l_{b1} < l_{b2}, \\ -\sqrt{L_b^{\,2} - \Delta x_b^{\,2} - \Delta y_b^{\,2}}, & \text{if } l_{b1} > l_{b2}. \end{cases} \tag{19}$$

Therefore, the 3D vector in {C} can be obtained from the two images which focus on the two ball's outline respectively.

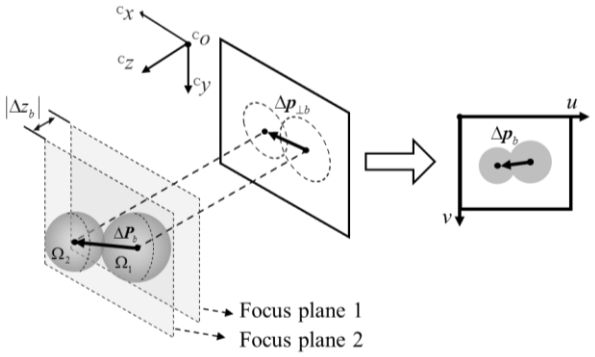



Fig. 2. The 3D vector pointing from one ball's center to the other's, its orthographic projection and image.

To obtain two camera's rotation relationship, the two-ball object is placed in the two cameras' FOVs under $n\geq 3$ different attitudes. The camera frames are labeled as $\{C_1\}$ and $\{C_2\}$ for the camera 1 and 2, respectively. The 3D vectors $\Delta \boldsymbol{P}_{1bi}$ and $\Delta \boldsymbol{P}_{2bi}$ ($i$=1,2,…,$n$) are obtained by the camera 1 and 2 using the aforementioned method, whose relationship is expressed as

$$\left[\Delta \boldsymbol{P}_{2b1}, \Delta \boldsymbol{P}_{2b2}, ..., \Delta \boldsymbol{P}_{2bn}\right] = {}^{C2}\boldsymbol{R}_{C1}\left[\Delta \boldsymbol{P}_{1b1}, \Delta \boldsymbol{P}_{1b2}, ..., \Delta \boldsymbol{P}_{1bn}\right] \tag{20}$$

where ${}^{C2}\boldsymbol{R}_{C1}$ is the rotation matrix from $\{C_1\}$ to $\{C_2\}$. Eular angles $\omega$, $\varphi$, and $\kappa$ can determine a rotation matrix $\boldsymbol{R}(\omega,\varphi,\kappa)$, which is guaranteed to be orthogonal [12]. Then we build the non-linear least squares formed error function

$$E(\omega,\varphi,\kappa) = \frac{1}{n}\sum_{i=1}^{n} e_i^2 \tag{21}$$

$$e_i = [\boldsymbol{R}(\omega,\varphi,\kappa)\Delta \boldsymbol{P}_{1bi}]\cdot \Delta \boldsymbol{P}_{2bi} / L_b^2 - 1 \tag{22}$$

which equals zero only when Eq. (20) is satisfied. The optimal solutions of $\omega$, $\varphi$, and $\kappa$ which minimize $E(\omega,\varphi,\kappa)$ can be solved by Levenberg-Marquardt algorithm [13], which forms the optimal estimation of ${}^{C2}\boldsymbol{R}_{C1}$.

In the vision system with two cameras, the world frame {W} is established to be identical with the camera 1 frame $\{C_1\}$ whose position along the focus axis is $l_0$=0. Thus, the rotation matrix from {W} to $\{C_1\}$ is an identity matrix. According to Eq. (4), the image Jacobian matrix of the camera 1 can be obtained by

$$\boldsymbol{J}_1 = \begin{bmatrix} \alpha_1 & \gamma_1 \\ 0 & \beta_1 \end{bmatrix}\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \alpha_1 & \gamma_1 & 0 \\ 0 & \beta_1 & 0 \end{bmatrix} \tag{23}$$

where $\alpha_1$, $\beta_1$, and $\gamma_1$ are the intrinsic parameters of the camera 1. The rotation matrix ${}^{C2}\boldsymbol{R}_{C1}$ equals ${}^{C2}\boldsymbol{R}_{W}$, so the image Jacobian matrix of the camera 2 is

$$\boldsymbol{J}_2 = \begin{bmatrix} \alpha_2 & \gamma_2 \\ 0 & \beta_2 \end{bmatrix}\boldsymbol{\Gamma} \tag{24}$$

where $\alpha_2$, $\beta_2$, and $\gamma_2$ are the intrinsic parameters of the camera 2. $\boldsymbol{\Gamma}$ is formed by the first two rows of ${}^{C2}\boldsymbol{R}_{C1}$.

## IV. Experiments and Results

### A. Experimental System

The telecentric micro-vision system for 3D relative pose measurement is shown in Fig. 3, which mainly consisted of two cameras with telecentric microscopes, two linear stages. The system is with the high precision 3 degrees-of-freedom manipulator, so that the proposed method can be verified and compared to the traditional active motion based method.

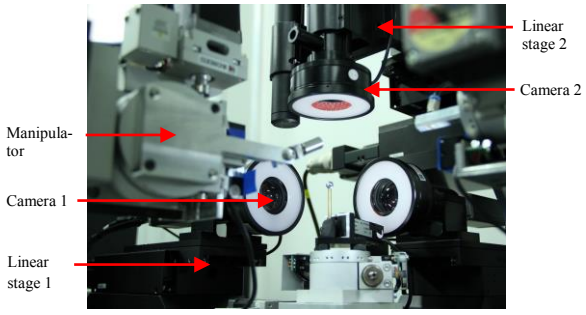



Fig. 3. Telecentric micro-vision system.

In the vision system, the two Basler piA2400-17gm cameras (Resolution: 2456×2058 pixel; Sensor size: 2/3"; Pixel size: 3.45μm×3.45μm) were mounted with the Myutron MGTL10VC telecentric microscopes (Magnification: ×1.0; Optical resolution: 2.5μm; TV distortion: <0.01%; Aperture: Adjustable). The two Sugura KS102-100 linear stages provided the translations with the 1μm positioning resolution. The manipulator was constructed by the Kohzu YA10A-L1 and ZA10A-X1T, whose resolution was 1μm.

### B. DOF Calibration

The one-ball object was placed in the camera 1's FOV. The camera's position was manually adjusted to focus on the ball outline. The camera was translated backward by 1000μm, and then stepped forward through 2000μm with the step of 50μm.

At each position, the ESW was estimated using the method in Section III.A. The Gaussian kernel for smoothing was with the standard deviation $\sigma_0$=1.0 pixel. On each image, 282 line segments $L_{si}$ ($i$=1,2,…,282) were sampled with the equivalent angle interval of 1° and the length of 100 pixel, as shown in Fig. 4.

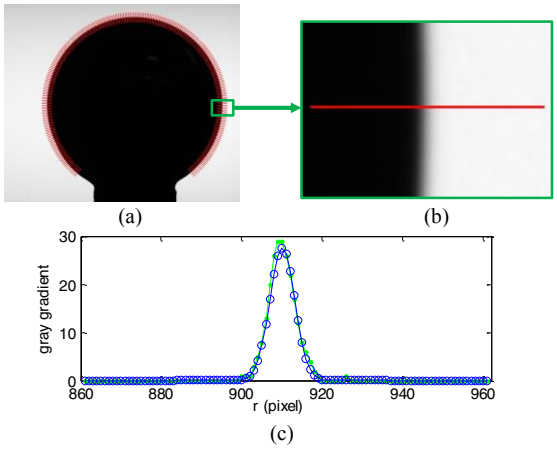



Fig. 4. Sampled line segments and their gray gradient distributions. (a) The ball's image and the sampled line segments (red); (b) The local image at the position labeled by the green window in (a), and the sample line segment (red) on it; (c) The raw gradient distribution (green dots) on the line segment in (b) and the fitted Gaussian model (blue circles).

The maximum acceptable blur circle diameter $\Phi_t$ was 40 μm on the CCD plane, and 11.6 pixel in the image frame. The DOF calibration was finished under three different aperture size settings: aperture size 1> aperture size 2> aperture size 3. The ESW curve changing with the camera's relative position is shown in Fig. 5. The calibrated DOFs, i.e. the length of the camera position range satisfying ESW≤11.6 pixel, under the aperture size 1, 2, and 3, were 450μm, 550μm, and 950μm, respectively. Therefore, the aperture size's increasing will decrease the DOF.

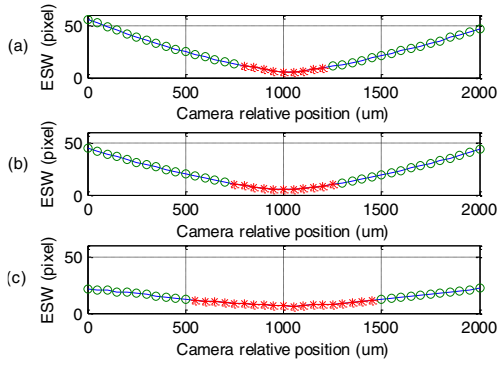



Fig. 5. The DOF calibration results with (a) aperture size 1, (b) aperture size 2, and (c) aperture size 3, respectively. The points with ESW≤11.6 pixel are labeled by "*", and the other points are labeled by "o".

## C. Image Offset Matrices Calibration

The one-ball object was placed in the FOVs of two cameras. The two cameras' aperture sizes were adjusted to the minimum value to maximize their DOFs. The two cameras automatically focused on the ball's outline. The two cameras were first translated backward by 400μm, and then stepped forward through 800μm with the step of 50μm. At each position $l_i$ ($i$=1,2,…9), the position $\boldsymbol{p}_i=(u_i, v_i)$ of the ball image's center was obtained. The image offsets $\Delta\boldsymbol{p}_i= p_i- p_1$ along with the camera position changes $\Delta l_i= l_i- l_1$ are shown in Fig 6. The calibrated image offset matrix of the camera 1 was $[0.000860294, -0.00359510]^T$ pixel/μm, and the root mean square (RMS) of the residuals was $[0.0218014, 0.0624679]^T$ pixel. The calibrated image offset matrix of the camera 2 was $[-0.000154902, 0.000296078]^T$ pixel/μm, and the RMS of the residuals was $[0.0728826, 0.105943]^T$ pixel.

## D. Intrinsic Parameters Calibration

The one-ball object was used to calibrate the two cameras' intrinsic parameters. The fitted ellipse equations in the image of camera 1 and 2 were $u^2$-0.004061$uv$+0.994663$v^2$ -2599.151075$u$-1955.627501$v$+1827875.874307=0 and $u^2$-0.001306$uv$+1.000378$v^2$-2548.430456$u$-2097.143104$v$+ 1890820.292536 = 0, respectively. According to Eq. (16), the intrinsic parameters of the camera 1 were $\alpha_1$=0.286505, $\beta_1$=0.287274, $\gamma_1$=0.000583377, and the intrinsic parameters of the camera 2 were $\alpha_2$=0.287571, $\beta_2$=0.287516, $\gamma_2$=0.000187692.

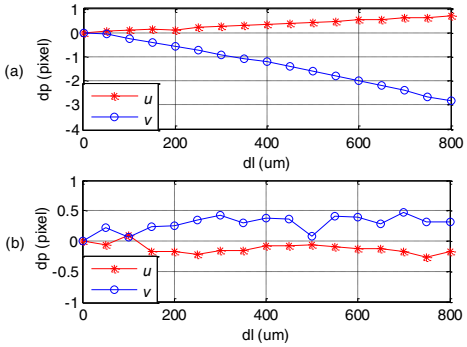



Fig. 6. Image offset vs. camera translation length. (a) camera 1, (b) camera 2.

## E. Image Jacobian Matrices Calibration

The two-ball object was assembled with two balls A and B, respectively, so that $L_b$ equaled 3950μm. The two-ball object was fixed in the two cameras' FOVs, whose images acquired by the two cameras are shown in Fig. 7. The two-ball object was adjusted to 7 different attitudes. Using the method in Section III.D and the calibrated results of image offset matrices and intrinsic parameters, the Eular angles were obtained as $\omega$=91.0107°, $\varphi$=-0.167998°, $\kappa$=-90.6705°. The minimized $E(\omega,\varphi,\kappa)$ was $9.0\times10^{-9}$. The rotation matrix determined by the Eular angles was given by

$$ {}^{C2}R_{C1} = \begin{bmatrix} -0.0117020 & -0.0176036 & -0.999777 \\ -0.999927 & 0.00313862 & 0.0116485 \\ 0.0116485 & 0.999840 & -0.0176391 \end{bmatrix}. $$

According to Eq. (23) and Eq. (24), the image Jacobian matrices calibration result were

$$ \boldsymbol{J}_1 = \begin{bmatrix} 0.286505 & 0.000583377 & 0 \\ 0 & 0.287274 & 0 \end{bmatrix}, $$

$$ \boldsymbol{J}_2 = \begin{bmatrix} -0.00355414 & -0.00506563 & -0.287504 \\ -0.287495 & 0.000897149 & 0.00335052 \end{bmatrix}. $$

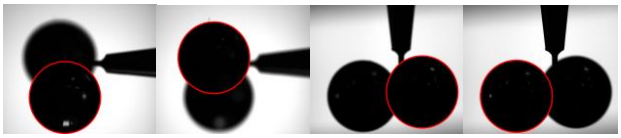
For the comparison, the image Jacobian calibration based on the active motion was conducted. The ball object, which was fixed on the manipulator's end, was translated for 8 times,

and the corresponding image position changes were recorded. The calibrated image Jacobian matrices were

$$\boldsymbol{J}_1' = \begin{bmatrix} -0.00155526 & -0.286613 & -0.00122268 \\ 0.00350263 & 0.00168493 & -0.286206 \end{bmatrix},$$

$$\boldsymbol{J}_2' = \begin{bmatrix} -0.288742 & 0.00114311 & 0.000651609 \\ 0.00101322 & 0.288763 & -0.000478239 \end{bmatrix}.$$

Note that in the proposed two-ball object based method, the world frame is identical with the camera frame $\{C_1\}$. Differently, in the active motion based method, the world frame is identical with the manipulator frame $\{M\}$.



(a) (b) (c) (d)

Fig. 7. Two-ball objects' images acquired by: (a) the camera 1 focusing on ball A, (b) the camera 1 focusing on ball B, (c) the camera 2 focusing on ball A, (d) the camera 2 focusing on ball B.

### *F. 3D Relative Pose Measurement*

To verify the calibration accuracy, the 3D vector measurement experiments were conducted with the calibration results. A ball object was fixed on the manipulator's end, and was translated by the 3D vectors $\Delta\boldsymbol{P}_i$ ($i$=1,2,…,8) for 8 times in the manipulator frame $\{M\}$. Using the method in Section II.C, the 3D vectors were measured by the vision system. The measured 3D vectors based on $\boldsymbol{J}_1$ and $\boldsymbol{J}_2$, which were calibrated with the proposed two-ball object based method, are $\Delta\boldsymbol{P}_{C1i}$ in the camera frame $\{C_1\}$. The measured 3D vectors based on $\boldsymbol{J}'_1$ and $\boldsymbol{J}'_2$, which were calibrated with the traditional active motion based method, are $\Delta\boldsymbol{P}_{mi}$ in the manipulator frame $\{M\}$. The length errors were $e_{li}=\|\Delta\boldsymbol{P}_{C1i}\|_2-\|\Delta\boldsymbol{P}_i\|_2$ and $e'_{li}=\|\Delta\boldsymbol{P}_{Mi}\|_2-\|\Delta\boldsymbol{P}_i\|_2$, which indicated the relative position measurement accuracies. The angle between two vectors $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ is given by $\psi(\boldsymbol{v}_1,\boldsymbol{v}_2)$, which indicates the relative orientation. Then the relative orientation errors were $e_{\psi i}=\psi(\Delta\boldsymbol{P}_{C1i}, \Delta\boldsymbol{P}_{C1i\text{-}1})- \psi(\Delta\boldsymbol{P}_i, \Delta\boldsymbol{P}_{i\text{-}1})$ and $e'_{\psi i}=\psi(\Delta\boldsymbol{P}_{mi}, \Delta\boldsymbol{P}_{mi\text{-}1})- \psi(\Delta\boldsymbol{P}_i, \Delta\boldsymbol{P}_{i\text{-}1})$ ($i$=2,…,8). The 3D vectors' actual and measured values and their errors are shown in the Table I.

TABLE I
3D VECTOR MEASUREMENT RESULTS

| Actual value | Measured value with the proposed method | | Measured value with the active motion based method | |
|---|---|---|---|---|
| $\Delta\boldsymbol{P}$ (μm) | $e_l$ (μm) | $e_\psi$ (°) | $e'_l$ (μm) | $e'_\psi$ (°) |
| $[-387,-1013,-1125]^T$ | -0.7 | —— | 0.5 | —— |
| $[-93,-1209,93]^T$ | 1.7 | 0.14 | -2.2 | 0.01 |
| $[-744,1302,-744]^T$ | 0.5 | 0.09 | -1.4 | 0.01 |
| $[-97,-307,1084]^T$ | -4.7 | -0.18 | -2.2 | 0.01 |
| $[-269,-995,659]^T$ | 4.2 | 0.05 | 0.1 | 0.02 |
| $[-1110,-222,0]^T$ | 4.0 | 0.27 | -2.0 | 0.04 |
| $[-12,2361,-1836]^T$ | 4.1 | -0.16 | -2.1 | 0.04 |
| $[-321,-1786,0]^T$ | 7.3 | -0.15 | 0.6 | 0.06 |

The proposed two-ball object based method provided acceptable calibration accuracy, based on which the maximum errors of relative pose measurement were 7.3μm in position and 0.27° in orientation. In comparison, the active motion based method provided the higher accuracy, based on which the maximum errors of relative pose measurement were 2.2μm in position and 0.06° in orientation. However, the active motion method is not suitable for some applications, because it is difficult to install a manipulator. The two-ball object can be easily made and placed under the desired poses.

## V. CONCLUSION

The low-cost calibration methods with ball objects for the telecentric micro-vision system are presented in this paper. In the DOF calibration, the edge spread width of the ball image is estimated to judge whether the ball center is within the DOF. The intrinsic parameters are calculated from the relationship between the ball's orthographic projection and image projection. The angle between the camera's optical axis and the linear stage's axis is considered, and the image offset matrix caused by this angle is calibrated. The rotation matrix between two cameras is calibrated with a two-ball object placed under several different attitudes. The image Jacobian matrices, used for relative pose measurement, are obtained from the intrinsic parameters and the rotation matrix.

## REFERENCES


[1] M. Greminger, B.J. Nelson, "Vision-based force measurement," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 3, pp. 290-298, 2004.

[2] S. Liu, D. Xu, D. Zhang, et al., "High precision automatic assembly based on microscopic vision and force information," *IEEE Trans. Automat. Sci. and Eng.*, vol. 13, no. 1, pp. 382-393, 2016.

[3] H. Yamamoto and T. Sano, "Study of micromanipulation using stereoscopic microscope," *IEEE Trans. Instrum. Meas.*, vol. 51, no. 2, pp.182-187, 2002.

[4] Z. Lu, C. Moraes, Y. Zhao, et al., "A micromanipulation system for single cell deposition," *IEEE Int. Conf. Robot. Autom.*, pp. 494-499, 2010.

[5] M. Watanabe, S. K. Nayar, "Telecentric optics for constant magnification imaging," *IEEE Trans. Pattern Anal. Mach. Intell,* vol. 19, pp. 1360-1365, 1995.

[6] D. Li, J. Tian, "An accurate calibration method for a camera with telecentric lenses," *Optics and Lasers in Engineering*, vol. 51, no. 5, pp. 538-541, 2013.

[7] Z. Chen, H. Liao, X. Zhang, "Telecentric stereo micro-vision system: Calibration method and experiments," *Optics and Lasers in Engineering*, vol. 57, pp. 82-92, 2014.

[8] D. S. Gorpas, K. Politopoulos, D. Yova, "Development of a computer vision binocular system for non-contact small animal model skin cancer tumour imaging," in *Proc. SPIE Diffuse Optical Imaging of Tissue*, Munich, 2007, vol. 66291J.

[9] F. Shen, W. Wu, D. Yu, et al., "High-precision automated 3-D assembly with attitude adjustment performed by LMTI and vision-based control," IEEE Trans," *IEEE/ASME Trans. Mechatron.*, vol. 20, no. 4, pp. 1777-1789, 2015.

[10] Fu. Li, D. Xu, Y. Shi, et al., "Motion-based microscopic camera calibration and application on micro tube-hole insertion," *Opt. Eng*., vol. 53, no. 5, pp. 053103, 2014.

[11] J. Bela, "A new sense for depth of field," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-9, no. 4, pp. 523-530, 1987.

[12] J. Heikkila and O. Silven, "A four-step camera calibration procedure with implicit image correction," in *Proc. IEEE CVPR*, San Juan, Puerto Rica, 1997, pp. 1106-1112.

[13] J. More, "The Levenberg-Marquardt algorithm, implementation and theory," in *Numerical Analysis, Lecture Notes in Mathematics*, vol. 630, G. A. Watson, Ed. New York, NY, USA: Springer-Verlag, 1977, pp. 105-116.