Research on global measurement method based on multi-cameras

Huizong Feng¹a, Gaofeng Wu¹b, Mingchi Feng²c*, Ming Cen¹d, Yibo Liu²e

¹ College of Automation, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
² School of Advanced Manufacturing Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China
*Corresponding author *fenghz@cqupt.edu.cn, bgood_wugaofeng@163.com, cfengmc@cqupt.edu.cn, dmcen0104@163.com, e812358853@qq.com

Abstract. The 3D reconstruction based on multi-cameras has been widely used and studied in the field of advanced manufacturing. Many 3D reconstruction methods have been proposed, but the limitation of high-accuracy, large-scale measurement always exists. The construction and calibration of multi-cameras system are the fundamental task as well as the key for the accurate measurement. In this paper, a high-precision camera calibration method for existing multi-cameras system is put forward. The transparent glass checkerboard and the principle of least time is applied in the method. The application of the transparent glass checkerboard leads to refraction phenomenon, which will generate calibration error. The refractive projection model is proposed to eliminate the error. Then the Scale-invariant feature transform (SIFT) algorithm is used to extract the feature point and epipolar geometry constraint is used to match the feature points. Finally, the 3D points can be obtained based on the principle of the triangle. The experimental results show the ability and robustness of the proposed approach.

1. Introduction
In recent years, the 3D reconstruction has been applied in the field of advanced manufacturing. Usually, the measurement of dimension, shape, and deformation is a dynamic process, and there are multiple approaches to complete the dynamic 3D reconstruction. Some researchers apply Non-Rigid Structure from Motion approaches [1,2] and multi-Kinects system [3] to measure the target. Due to the characteristics of the monocular camera, the monocular 3D reconstruction cannot accurately estimate the geometric position. Several Kinect can increase the field of view, but these errors become more noticeable when multiple Kinects are used simultaneously to model a live and large 3D scene. Thus, the multi-cameras system is widely used and studied [4,5]. However, in [4] the process of calibration is too complication and in [5] the special calibration plate exists some problems that the feature points on both sides cannot be aligned precisely, which will affect the accuracy of the experimental result.

In order to overcome the defect of the methods mentioned above, a novel calibration method for the extrinsic parameters of multi-cameras system is proposed in this paper, in which a glass calibration plate is adopted. For all cameras in the system, the point on the pattern is the same. The glass calibration plate improves the precision of calibration parameters. Based on the experimental results of calibration and reconstruction, the proposed system of multi-cameras in this paper can achieve high-accuracy.
2. Multi-cameras model

2.1. The camera projection model

In order to obtain 3D shape of the target, the relationship among the coordinate systems must be established. The ideal pinhole imaging model is shown in figure.1.

![The pinhole camera model](image)

In figure 1, \(O_{W}X_{W}Y_{W}Z_{W}\) is the world coordinate system. \(O_{C}X_{C}Y_{C}Z_{C}\) is the camera coordinate system. \(O_{I}X_{I}Y_{I}\) indicates the image coordinate system and \(O_{P}X_{P}\) is the pixel coordinate system. \(M(X_{W},Y_{W},Z_{W})\) is 3D point on the target and \(m(u,v)\) represents the projection of \(M\) on the image plane. By applying the pinhole camera model, the transformation from the pixel coordinate system to the world coordinate system for each camera can be expressed in Eq. (1).

\[
\hat{m} = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K[R \mid T] \begin{bmatrix} X_{W} \\ Y_{W} \\ Z_{W} \end{bmatrix}
\]

(1)

where, \(R\) and \(T\) are the rotation matrix and the translation vector from the world coordinate system to the image coordinate system, and \(K\) is the camera intrinsic parameters. The 3D point \(M\) in the world coordinate system can be projected to the point \(m\) in the image coordinate system, which can use the following projection equation:

\[
m = f(K, R, T, \phi, M)
\]

(2)

where, \(\phi\) represent the vector of distortion coefficients.

2.2. The refraction projection model

Because the transparent glass target is applied in our method, there are calibration errors when the aforementioned pinhole model is applied directly. The refraction in the optical paths must be considered for when projecting 3D points into cameras through glass. To simplify the problem, one camera is taken as an example, which is shown in figure.2.

![Flat refractive geometry with two-layer of refractive interface](image)
The refractive index of air and glass is \( \mu_a \) and \( \mu_g \), respectively. Usually, the refractive index of air is equal to one \( (\mu_a = 1) \), and the refractive index of air is one of the optimized parameters. \( \vec{r}_0 \) and \( \vec{r}_1 \) are incident ray and refracted ray. \( q_0 \) is the intersection point between the incident ray \( \vec{r}_0 \) and the refractive plane. \( q_1 \) is the intersection point between the incident ray \( \vec{r}_1 \) and the plane of pattern printed. The thickness of the glass plate is \( d \). The distance between \( O \) and \( q_0 \) is denoted as \( L_0 \) and the distance between \( q_0 \) and \( q_1 \) is denoted as \( L_1 \).

In figure 2, the coordinate of camera center \( O \) in the world coordinate system is \((x_c, y_c, z_c)\). In the same way, the coordinates of \( q_0 \) and \( q_1 \) are \((x_0, y_0, d)\) and \((x_1, y_1, 0)\). Thus, \( L_0 \) and \( L_1 \) can be expressed as:

\[
\begin{align*}
L_0 &= \sqrt{(x_c - x_0)^2 + (y_c - y_0)^2 + (z_c - d)^2} \\
L_1 &= \sqrt{(x_0 - x_c)^2 + (y_0 - y_c)^2 + d^2}
\end{align*}
\] (3)

where, \( O_c \) can be obtained by the extrinsic parameters of the camera, which is expressed as \( O_c = R_{w1}^T \). \( q_1 \) is the 3D point on the surface of the calibration pattern. Based on Eq. (2), the optical path between \( O \) and \( q_1 \) can be expressed as:

\[
L = \mu_0 L_0 + \mu_1 L_1
\] (4)

Based on the principle of least time, when light travels from one point to another in any medium, optical path must be the shortest, the point \( q_0 \) can be acquired. Thus, a point \( M' \) on the target may be transformed into a 2D point \( m' \) on the image with refraction using the following equation.

\[
m' = f'(K, R, T, \phi, M', \mu_t)
\] (5)

2.3. Geometrical relationship of multi-cameras

Without losing the generality, 4 cameras are used in the system. Figure 3 shows the geometrical relationship of four-cameras.

Fig.3. Geometrical transformations involved in the modeling of the four-cameras system

\( O_{w1}X_{w1}Y_{w1}Z_{w1} \) represents the world coordinate of the front subsystem, and \( O_{w2}X_{w2}Y_{w2}Z_{w2} \) indicates the world coordinate of the back subsystem. The relationship between two world coordinates can be expressed:

\[
M_{w2} = R_{w1 \rightarrow w2} M_{w1} + T_{w1 \rightarrow w2}
\] (6)

where, \( R_{w1 \rightarrow w2} \) and \( T_{w1 \rightarrow w2} \) are the rotation matrix and translation vector between the two coordinate system. Since the information of the calibration plate is known, \( R_{w1 \rightarrow w2} \) and \( T_{w1 \rightarrow w2} \) can be obtained. The camera C1 and C2 belongs to the front subsystem, the other two belongs to the back subsystem. In this paper, the coordinate system of the camera C1 is taken as the reference coordinate system, the
coordinates of all the other three cameras can be unified to the reference coordinate based on Eq.(6). When there are more cameras in the system, the relationship between them can be obtained by the above principles.

3. Multi-cameras calibration and 3D reconstruction

3.1. Multi-cameras calibration based on refraction model

The calibration of multi-camera system is a fundamental task as well as the key for the accurate measurement. The intrinsic parameters of four cameras can be obtained by the Zhang’s method [6]. In this paper, a novel calibration method for the extrinsic parameters of multi-camera system is proposed, in which a glass calibration plate is adopted, and the calibration pattern is only printed on single side. As is shown in figure 4, two cameras (C1 and C2) are placed in front of the object, and the other two (C3 and C4) are placed in the back side. When the 3D point \( \mathbf{M} \) is projected to all cameras at the same time, the two cameras (C1 and C2) without refraction are based on the Eq. (2), and the other two (C3 and C4) apply the Eq.(5).

![Fig 4. Four-cameras calibration](image)

3.2. Initial Value and Nonlinear Optimization.

There is refraction phenomenon because of the glass calibration target. Nevertheless, the thickness of the target is small, the refraction can be ignored for initial estimation. Therefore, the initial extrinsic parameters can be obtained by the linear method. The camera parameters can be optimized by using the bundle adjustment method to minimize the reprojection error which is a geometric error corresponding to the image distance between a projected and measured point. The objective function is express as:

\[
\text{ArgMin}_{\mathbf{R}, \mathbf{t}} \sum \sum \sum ||x_{ij} - m_{ij}||^2
\]

where, \( x_{ij} \) and \( m_{ij} \) are the measured image point and the predicted image point of the \( i_{th} \) 3D point on the \( k_{th} \) camera.

3.3. 3D reconstruction

The algorithm of SIFT is used to extract feature points and feature points are matched based on the descriptor similarity and the epipolar geometry across two images. When the matching has been established between each pair of images, the corresponding 3D points of each two image can be obtained based on the triangulation principle by the matching points. According to the triangulation principle, the deduced formula is as follows:
\[
\begin{align*}
X &= \frac{x_l(f_l t_x - x_i t_x)}{(r_l y_i + r_l y_i + r_l f_l)x_r - f_r(r_l x_i + r_l y_i + r_l f_l)} \\
Y &= \frac{y_l(f_l t_y - x_i t_y)}{(r_l y_i + r_l y_i + r_l f_l)x_r - f_r(r_l x_i + r_l y_i + r_l f_l)} \\
Z &= \frac{f_l(f_l t_z - x_i t_z)}{(r_l y_i + r_l y_i + r_l f_l)x_r - f_r(r_l x_i + r_l y_i + r_l f_l)}
\end{align*}
\]  

where, \((x_l, y_l)\), \((x_r, y_r)\) are the coordinates of the corresponding points \(m_i\) and \(m_r\) in the image coordinate system, respectively. \(f_l\) and \(f_r\) are the focal length of the left camera and right camera. \((X, Y, Z)\) is the coordinates of the 3D point \(M\). \((r_l, r_r, r_l, r_r, r_l, r_r, t_l, t_r, t_l)\) are the elements of \(R_l\) and \(T_l\). The 3D points of stereo vision system can be obtained by Eq.(8).

4. Experiments

The four-cameras system used in the experiment is shown in figure 5. The calibration results of four cameras are relatively complex, so the intrinsic and extrinsic parameters of the cameras are not shown.

In order to verify the accuracy and feasibility of the proposed system, the reconstruction experiment of four-cameras system is implemented using the intrinsic and extrinsic parameters. In figure 6 (a), the 3D point cloud is generated by the two subsystems, which have been transformed into the reference coordinate system. The least squares algorithm is used to synthesize 3D points into a plane, which is shown in figure 6 (b). Most of the 3D points are on the plane of the fitting, but there are some deviations in some 3D points, leading to flat irregularities. Because the target object used in the experiment is a thin iron product, the surface of which is prone to deformation and the surface has been uneven. Object surface is smooth and easy to reflect, which will result in mismatching of feature points. But overall, the two planes are approximately parallel, and the distance between the two planes is approximately equal to the width of the object. To better demonstrate the experimental results, we obtain Delaunay triangulation and texture mapping of the 3D point clouds, which is shown in figure 7. It shows that the 3D reconstruction of multi-camera system proposed in this paper has strong accuracy and practicability.
5. Conclusion
In this paper, we presented a method to calibrate multiple cameras with the help of a glass plate. This method is applied to the multi-cameras measurement system, by which the dynamic shape of whole part can be accurately measured and monitored. The proposed method involves calibrating the intrinsic and extrinsic parameters, solving a set of transformation relationships between the other cameras and the refractive model. The experiment results show that the calibrated parameters can well describe the imaging phenomenon of all cameras and the target object can well reconstruction. The method has a more extensive application in the field of advanced manufacturing.

Acknowledgements
National Natural Science Foundation of China (NSFC) (51505054); Chongqing Municipal Education Commission (KJ1500405, CYS17233); Chongqing Science and Technology Commission (cstc2015zdcy-ztzx60002, cstc2016jcyjA1613, cstc2015zdcy-ztzx30001).

References
[1] Bronte, S., Paladini, M., Bergasa, L. M., & Agapito, L. (2014). Real-time sequential model-based non-rigid SFM. Ieee/rsj International Conference on Intelligent Robots and Systems (pp.1026-1031). IEEE.
[2] Paladini, M., Bartoli, A., & Agapito, L. (2010). Sequential Non-Rigid Structure-from-Motion with the 3D-Implicit Low-Rank Shape Model. European Conference on Computer Vision (Vol.6312, pp.15-28). Springer-Verlag.
[3] Yang, R. S., Chan, Y. H., Gong, R., & Nguyen, M. (2013). Multi-Kinect scene reconstruction: Calibration and depth inconsistencies. International Conference on Image and Vision Computing New Zealand (pp.47-52). IEEE.
[4] Orteu, J. J., Bugarin, F., Harvent, J., Robert, L., & Velay, V. (2011). Multiple-camera instrumentation of a single point incremental forming process pilot for shape and 3d displacement measurements: methodology and results. Experimental Mechanics, 51(4), p. 625-639.

[5] Li, J., Xie, X., Yang, G., Zhang, B., Siebert, T., & Yang, L. (2017). Whole-field thickness strain measurement using multiple camera digital image correlation system. Optics & Lasers in Engineering, 90, 19-25.

[6] Zhang Z. A flexible new technique for camera calibration[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2000, 22(11):1330-1334.