Calibration of Position and Orientation between Cameras without Common Field of View Using Cooperative Target

Yongzhi Min^{*} and Jie Hu[†] *Lanzhou Jiaotong University, Lanzhou, China E-mail: minyongzhi@mail.lzjtu.cn [†]Lanzhou Jiaotong University, Lanzhou, China E-mail: 2043687908@qq.com

Abstract— In online surface settlement monitoring system of image-based ballastless track, a transfer station consisting of a multi-camera without a common field of view is often used. Because measurement error will accumulate as the number of transfer increases, position and orientation relation betweeen cameras should be calibrated. Traditional calibration methods based on total stations are pretty complicated. We found that the hand-eye calibration of the robot is equivalent to this problem. Firstly, the multi-camera is moved twice in small step and the multi-camera captures the cooperative target image at three different positions and image physical coordinates of four mark points in each image are extracted. Secondly, in order to solve the extrinsic parameters of the camera, we use a special P4P algorithm for feature points distributed in a square. Lastly, the matrix rearrangement method is used to solve the hand-eye transformation matrix according to the extrinsic parameters of The experimental results show that the the cameras. measurement accuracy of this method is same as the traditional method. Moreover, the method features simple operation, less calculation task and high precision position and orientation.

I. INTRODUCTION

With the development of high-speed and heavy haul railways, settlement of subgrade has become the important factor of train operation stability. Online surface settlement monitoring system of image-based ballastless track obtains the difference of the position data before and after the settlement of the subgrade by the center position of the laser spot, in order to reflect the settlement deformation of the subgrade that is relative to the reference point. Due to the different needs of the monitoring site, the intermediate transmission monitoring stations may have multiple combinations, including multi-camera-multi-target-multi-laser combination, multi-camera-multi-target combination and so on.Different camera-target combinations in the transfer station must point to the respective indicator laser. Therefore, there is no common field of view between the camera and the camera.Because of the need for the transmission of position and orientation parameters, it is necessary to accurately calibrate the relative pose between the two cameras in the intermediate station.

Nowadays, position and orientation parameters of multi-

cameras rig in the system are generally calibrated by assistant of total station and a method based on hand-in-eye calibration in robot vision[1]. The method of using the total station to assist in calibrating a multi-camera rig requires that control points be placed around the camera to unify all control points into the total station coordinate system. However, this method requires a large number of control points, a large amount of calculation, high work intensity, so it can not meet the needs of field measurement. When we discuss the calibration method for setting relation of multi-head cameras based on hand-eye calibration, the checkerboard is often used as a template, and a large number of checkerboard corner coordinates are extracted, and there is a large deviation in the extraction. For the solution of the equation AX = XB, common method requires a multi-head camera to move a large angle to accurately calibrate. In order to solve this problem, a calibration method for relationship of position and orientation between cameras based on cooperative targets is proposed. This method uses the P4P method to greatly reduce the calculation in the case of ensuring high-precision of position and orientation parameters The matrix rearrangement method enables the multi-head camera to complete the calibration of the multi-head camera with a small angle of rotation.

II. MEASUREMENT MODEL

The calibration for relationship of position and orientation between cameras without common field of view is very similar to hand eye calibration in robot vision. The problem of hand-eye calibration is solved by using the equation AX = XB. A is the transformation relationship between camera coordinate systems at different positions. B is the transformation relationship of the robot end link before and after the movement. X is the transformation relationship between the hand and the eye of the robot. There are rigid connections between the two cameras in the intermediate station. One camera can be regarded as the hand in the handeye system, and the other can be regarded as the eye of the hand-eye system. By solving the position and orientation of the camera and the fixed target, we can further calculate the set relation between the two cameras. At the initial moment, the position and orientation relation of the main camera and the subordinative camera relative to the fixed target is expressed as V_m^0 , V_c^0 . At the other moment, they are represented as V_m^t , V_c^t . Position and orientation relation between two cameras is represented as X. Main camera coordinate system and subordinative camera coordinate system are represented as C_m , C_c . Corresponding target coordinate systems are represented as C_{t1} , C_{t2} . The position of the multihead camera at the initial time and other time is expressed as L_1 , L_2 . The relationship between variables is represented by the following "Fig. 1."



Fig. 1 Geometric relationship of the camera frame and target frame before and after the platformmovement

The point $P_{\rm m}$ in the target surface corresponding to the main camera is represented as P_{m0} , $P_{\rm m1}$ at the initial time and other time in camera coordinate systems. The point $P_{\rm c}$ in the target surface corresponding to the subordinative camera is represented as $P_{\rm c0}$, $P_{\rm c1}$ at the initial time and other time in camera coordinate systems. Hence, wen can get

$$\begin{cases}
P_{m1} = HP_{m0} \\
P_{c1} = GP_{c0} \\
P_{m0} = XP_{C0} \\
P_{m1} = XP_{c1}
\end{cases}$$
(1)

where $H = V'_{m} V^{0^{-1}}_{m}$, $G = V'_{c} V^{0^{-1}}_{c}$, combine these equations, we can get

$$HX = XG \tag{2}$$

Where : H and G is a known 4×4 matrix. X is the position and orientation matrix between the cameras to be figured out. H, G and X are composed of a rotation matrix and a translation vector.For example, we can write

$$X = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix}$$
(3)
Then,Equation(2) can be rewritten as

$$R_{\mu}R - RR_{\alpha} = 0 \tag{4}$$

$$(R_H - I)t = Rt_G - t_H \tag{5}$$

III. CAMERA YRAJECTORY ESTIMATION

Before solving setting relation of a multi camera, it is necessary to obtain the camera pose transformation matrix G, H at different moments through camera trajectory estimation [2]. Camera trajectory estimation is defined as the position of the camera at a certain moment relative to the initial moment in a series of images. We use P4P algorithm to obtain extrinsic parameter of camera which refers to the pose relationship between the camera and the fixed target at different positions, and then calculate the camera trajectory G, H through a series of rigid body transformations.

A. Solving extrinsic parameter of camera.

we assume that the camera coordinate system is obtained by first rotating and then translate the target coordinate system. At any point in space, the relationship between camera coordinates $P_1(x_1, y_1, z_1)$ and target coordinates $P_2(x_2, y_2, z_2)$ is

$$P_1 = RP_2 + t \tag{6}$$

where $R(ij)_{3\times 3}$ is rotation matrix, $t(t_x, t_y, t_z)$ is threedimensional translation vector.

As shown in "Fig. 2," the target coordinate system is established with Q_0 as the origin, and four marker points Q_0 , Q_1 , Q_2 and Q_3 are set at four vertices of the square. Taking the coordinates of the four marker points in the target coordinate system and the image coordinates of the projection points as input, we can solve extrinsic parameter of camera.



Fig. 2 Issue of 3D position and orientation measurement

The coordinates of the marker points on the image plane in the camera coordinate system are $q_0(x_0, y_0, f)$, $q_1(x_1, y_1, f)$, $q_2(x_2, y_2, f)$, $q_3(x_3, y_3, f)$ respectively. The plane formed by the center of light O_c and segment q_2q_3 is represented as π_1 , whose normal vector is represented

$$n_1 = \overrightarrow{Oq_2} \times \overrightarrow{Oq_3}$$
 represented as $n_1 = \{n_{X1}, n_{Y1}, n_{Z1}\}$.

Because of $P_0P_1//\pi_1$, we can get $P_0P_1 \cdot n_1 = 0$. The ratio of the light center to the point of the marker and the distance from the light center to the point in image plane is represented

as
$$k_0 = \frac{|OQ_0|}{|Oq_0|}, k_1 = \frac{|OQ_1|}{|Oq_1|}$$
. Hence, we can write
 $(k_1x_1 - k_0x_0) \cdot n_{x_1} + (k_1y_1 - k_0y_0) \cdot n_{y_1} + (k_1y_1 - k_0y_0) \cdot n_{z_1} = 0$ (7)

becauce of $|P_0P_1| = d$, we can write

$$\sqrt{(k_1x_1 - k_0x_0)^2 + (k_1y_1 - k_0y_0)^2 + (k_1f - k_0f)^2} = d$$
(8)

By Eq(7) and Eq(8), we can get k_0 , k_1 . Further, we get the coordinates of the marker point in the camera coordinate system $Q_0(k_0x_0, k_0y_0, k_0z_0)$, $Q_1(k_1x_1, k_1y_1, k_1z_1)$. In the same way, they can find coordinates of Q_2 , Q_3 in the camera coordinate system.

In fact, the coordinates of Q_0 in the camera coordinate system are the translation vectors. The direction of he X axis of the target coordinate system is in the camera coordinate system

$$Q_0 Q_1 = O Q_1 - O Q_0 = (k_1 x_1 - k_0 x_0, k_1 y_1 - k_0 y_0, k_1 f - k_0 f)$$
(9)

Normalize the vector and record it as (r_{11}, r_{21}, r_{31}) . It represents the first column vector of the rotation matrix. Similarly, the Y axis of the target coordinate system in the camera coordinate system is expressed as (r_{12}, r_{22}, r_{32}) . By the orthogonality of the rotation matrix, we can get (r_{13}, r_{23}, r_{33}) . Finally, we get the rotation matrix R.

B. Solving calibration of position and orientation between cameras.

First, let's discuss the solution of the rotation matrix in X. Let $T = R_H R - R R_G$. Any entryof T can be represented as the product of R_H , R_G and R

Thus, we can get

$$Fv = 0$$

$$\mathbf{v} = \begin{bmatrix} R_{11} & R_{12} & R_{13} & R_{21} & R_{22} & R_{23} & R_{31} & R_{32} & R_{33} \end{bmatrix}^{I} \\ = \begin{bmatrix} R_{in1} - R_{ci1} & -R_{ci2} & -R_{ci} & R_{in2} & 0 & 0 & R_{in3} & 0 & 0 \\ -R_{ci2} & R_{in1} - R_{ci2} & -R_{ci} & 0 & R_{in2} & 0 & 0 & R_{in3} & 0 \\ -R_{ci3} & -R_{ci3} & R_{in1} - R_{ci3} & 0 & 0 & R_{in2} & 0 & 0 & R_{in3} \\ R_{ir21} & 0 & 0 & R_{ir22} - R_{ci} & -R_{ci} & R_{ci31} & R_{ir23} & 0 & 0 \\ 0 & R_{ir21} & 0 & -R_{ci32} & R_{ir22} - R_{ci} & R_{ci32} & 0 & 0 & R_{ir33} \\ R_{ir31} & 0 & 0 & R_{ir22} & -R_{ci} & R_{ci32} & 0 & 0 & R_{ir23} \\ 0 & 0 & R_{ir21} & -R_{ci13} & -R_{ci} & R_{ir22} - R_{ci33} & 0 & 0 & R_{ir23} \\ R_{ir31} & 0 & 0 & R_{ir32} & 0 & 0 & R_{ir33} - R_{ci1} - R_{ci31} & -R_{ci31} \\ 0 & R_{ir31} & 0 & 0 & R_{ir32} & 0 & 0 & R_{ir33} - R_{ci31} - R_{ci32} - R_{ci32} \\ 0 & 0 & R_{ir11} & 0 & 0 & R_{ir32} & 0 & 0 & R_{ir3} - R_{ci3} - R_{ci3} \\ 0 & 0 & R_{ir11} & 0 & 0 & R_{ir32} & 0 & 0 & R_{ir3} - R_{ci3} - R_{ci3} \\ 0 & 0 & R_{ir11} & 0 & 0 & R_{ir32} & 0 & 0 & R_{ir3} - R_{ci3} - R_{ci3} \\ 0 & 0 & R_{ir11} & 0 & 0 & R_{ir32} & 0 & R_{ci3} - R_{ci3} - R_{ci3} - R_{ci3} \\ 0 & 0 & R_{ir11} & 0 & 0 & R_{ir32} & 0 & R_{ci3} - R_{ci3} - R_{ci3} - R_{ci3} \\ 0 & 0 & R_{ir11} & 0 & 0 & R_{ir32} & 0 & R_{ci3} - R_{ci3} - R_{ci3} - R_{ci3} \\ 0 & 0 & R_{ir11} & 0 & 0 & R_{ir32} - R_{ci3} - R_{ci3} - R_{ci3} - R_{ci3} \\ 0 & 0 & R_{ir11} & 0 & 0 & R_{ir32} - R_{ci3} - R$$

In order to make the number of independent vectors in the basic solution system of the homogeneous equation keep one, the rank of the matrix should be increased by increasing the dimension of the column vectors in the matrix.

We can make the multi-camera rig rotate twice, which meet the need. Therefore, we can get the following two equations.

$$\begin{cases} R_{H1}R = RR_{G1} \\ R_{H2}R = RR_{G2} \end{cases}$$
(11)

By combining these two equations, we can get 18×9 matrix with rank of 8.By methods of singular value decomposition (*SVD*), we can solve equations Fv = 0. We decompose the matrix $F^T F$ with *SVD*. The eigenvector corresponding to the minimum eigenvalue differs only by a multiple from the solution of the equation. We can rearrange the obtained eigenvecto into a 3×3 matrix and normalize the schmidt units. So, we get rotation matrix R.

Next, we discuss the solution translation vector in X.

Let's rearrange formula 1, we can write

$$[R_H - I]t = [Rt_G - t_H]$$
(12)

We need to make the multi-camera rig rotate twice to make the rank of the equation large enough. Hence, we can get the following two equations.

$$\begin{bmatrix} R_{H1} - I \\ R_{H2} - I \end{bmatrix} t = \begin{bmatrix} Rt_{G1} - t_{H1} \\ Rt_{G2} - t_{H2} \end{bmatrix}$$
(13)

t of the above equations can be solve by method of least squares.

By recombining the rotation matrix and the translation matrix, we can get matrix X.

C. Nonlinear optimization

When solving the pose relationship between cameras, we assume that the camera trajectory is accurate and do not optimize external parameters of the camera. So it needs to be optimized. In nonlinear optimization, we establish a minimizing objective function according to the deviation between the reprojection result and coordinates of the actual image point. The sum of the squares of the errors of the reprojected image point coordinates and the actual image point coordinates is represented as F_{\min} . The actual image point coordinates is represented as f_{xi} and f_{yi} . The re-projected image point coordinates is represented as f_{xi} and f_{yi} . Hence, wen can get

$$F_{\min} = \sum_{i=1}^{4} \left[\left(f_{xi} - f_{xi}'(R, t, K, \delta) \right)^2 + \left(f_{yi} - f_{yi}'(R, t, K, \delta) \right)^2 \right]$$
(14)

We choose the Levenberg-Marquardt (LM) method which has characteristics of fast convergence rate and high stability to solve it. Next, we discuss the selection of iterative initial value. The iterative initial values of the rotation parameters R and translation parameters t are the results obtained by P4P algorithm. The initial value of the camera

(10)

distortion parameter δ is 0. And the initial value of the camera intrinsic parameters K is the camera factory parameter.

IV. ANALOG EXPERIMENT

First, we set the rotation paramete (φ , θ , ϕ) and translation parameters (t_x , t_y , t_z) from the main camera to the subordinative camera to be (0°,180°,0°) and (0m,0m,-0.05m). Next, We add gaussian distribution random measurement error to the image point coordinates. The mean of the measurement error varies from 0 pixel to 1 pixel with an interval of 0.1 pixel. Through the algorithm in section 3, we get the linear initial value and nonlinear optimization results. For each noise condition, we simulated 100 times, and the errors of rotation parameters and displacement parameters were obtained as shown in "Table 1" and "Table 2."

Tab.1 Errors in rotation versus image noise ()

	linear initialization			noolin	ear optin	nization
σ	α	β	γ	α	β	γ
0.0	0.041	0.029	0.036	0.031	0.019	0.022
0.1	0.129	0.136	0.137	0.071	0.063	0.076
0.2	0.193	0.201	0.197	0.106	0.115	0.105
0.3	0.271	0276	0.281	0.160	0.139	0.147
0.4	0.360	0.357	0.362	0.190	0.214	0.216
0.5	0.491	0.476	0.479	0.261	0.281	0.260
0.6	0.561	0.563	0.572	0.326	0.337	0.325
0.7	0.631	0.647	0.638	0.379	0.391	0.383
0.8	0.710	0.692	0.703	0.435	0.437	0.418
0.9	0.781	0.761	0.769	0.490	0.501	0.473
1.0	0.827	0.833	0.830	0.537	0.526	0.507

T 1 A	-	•			•	•	< >	
Inh?	Lrrorg	11	rototion	Vorcuo	1100000	noice	(mm)	
1 a. U. Z	1211015		TOLALION	VEISUS	IIIIage	HUISE		1
					8-		()	

	linear initialization			noolinear		
σ	t_x	t_y	t,	t_x	t_y	t,
0.0	0.061	0.067	0.073	0.026	0.037	0.031
0.1	1.571	1.531	1.597	0.371	0.417	0.393
0.2	2.364	2.537	2.539	0.972	0.901	0.819
0.3	3.903	3.571	3.612	1.560	1.734	1.801
0.4	4.697	4.913	5.173	2.447	2.509	2.772
0.5	6.072	5.961	6.303	3.627	3.561	3.603
0.6	7.671	7.806	7.891	4.301	4.671	4.470
0.7	9.307	9.176	9.239	5.162	5.507	5.132
0.8	10.691	10.732	10.621	5.761	5.939	5.910
0.9	11.337	11.109	11.532	6.210	6.360	6.307
1.0	11.903	11.761	11.956	6.671	6.813	6.771

V. EXPERIMENTAL RESULTS AND ANALYSISL

The experiment used Microvision MV3000UC cameras with a focal length of 8mm and Allied Vision Technologies GE680 camera with a focal length of 12mm, and the intrinsic parameters of the two cameras before the experiment have been calibrated. The two cameras are fixed on the NT305WM three-dimensional high-precision displacement platform with the precision of 0.01°, and can be freely rotated within the range by fine-tuning the bolts. The diameter of the feature point on the cooperative target is 15mm. Marker points are distributed in a square with 150mm side length. The experiment device is shown in "Fig. 3."



Fig. 3 Experiment facility

We control the platform to rotate 5° around the Z axis, the turntable moves from first position to second position, then we control the platform to rotate 5° around the Y axis and translate 30 cm along the axis Z of the main camera coordinate system. And the turntable moves to third position. We took 50 pictures in each of three positions, and one of them is shown in "Fig. 4."



(c) Position3 Fig.4 Target images for three different position

By solving the position and orientation relation between the camera and the fixed target through the P4P algorithm, we calculate the trajectory of the two cameras H_1, G_1, H_2, G_2 .

ine anaci	e me majeero	i y oi uie two		$[, 0], 11_2, 0_2$
	0.996133	-0.087850	0.001150	0.087625
$H_{1} =$	0.087851	0.996132	-0.010012	2 0.083908
	-0.001047	0.001221	0.996133	0.961237
	0	0	0	1
	0.996083	0.088406	0.000978	0.215178
<i>C</i> –	-0.088407	0.996084	0.000561	0.088021
$G_1 =$	- 0.000925	- 0.000645	0.996084	0.172316
	0	0	0	1
	0.996110	-0.008439	0.087667	0.041961
11 _	0.009214	0.999948	-0.004251	0.230351
$\Pi_{2} =$	-0.087626	0.005041	0.999944	30.013243
	0	0	0	1
	0.996113	0.004625 -	0.087986	0.231417
<i>G</i> ₂ =	-0.004868	0.999985	0.002537	-0.048139
	0.087973	0.002955	0.999983	-30.601321
	0	0	0	1

Hence, we get the position and orientation relation of the two cameras matrix \boldsymbol{X} .

	-0.999130	-0.030504	-0.028434	0.616143
<i>v</i> _	0.029828	-0.999269	0.023895	0.314279
л =	-0.029143	0.023026	- 0.999289	8.013021
	0	0	0	1

The Euler angle (Φ_X, Φ_Y, Φ_Z) is used to describe the angle transformation instead of the rotation matrix, so we can know the main camera coordinate system and the subordinate camera coordinate system rotating about the X, Y, Z axes by 1.32° , 178.29° , 1.67° .

Let $err_i = ||H_i X - XG_i|| (i = 1, 2)$. By the method of

this paper and the literature [2-3], we calculate the matrix X. The relative error results of the two methods are shown in "Table 3."

Tab.3 Experimental error analysis						
	Literature [2]	Literature [3]	This paper			
<i>err</i> ₁	0.064238	0.082319	0.052143			
err ₂	0.066921	0.084321	0.055237			

It can be seen that the accuracy of the algorithm in this paper is better than that in literature [2] and literature [3]. In the experiment, the angle of rotation of the platform is very small, but the precision of calibration is very ideal.

VI. CONCLUSIONS

Aiming at the calibration of relative pose between cameras, this paper proposes a method for the calibration of position and orientation relation between cameras based on fixed targets according to the hand-eye calibration in robot vision and P4P algorithm. Multi-cameras only need to perform 2 small motion to complete the calibration, which avoid the inconvenience of large motion. It only needs to process 24 points in 6 pictures, which is fast and requires less calculation. The error meets the requirements of ballastless track settlement monitoring.

REFERENCES

- Zhuoman Wen, Yanjie Wang, Nan Di, et al, "On-orbit hand-eye calibration using cooperative target," *Chinese Journal of Scientific Instrument. China*, vol. 35(05), pp. 1005-1012, 2014.
- [2] Gang Wang, Yang Shang, Banglei Guan, et al. "Flexible calibration of setting relation of a multi-Camera Rig for nonoverlapping views," *Chinese Journal of Lasers. China*, vol. 44(06), pp. 207-213,2017.
- [3] LIU J B, GUO P Y, LI X, et al. "Evaluation strategy fpr camera pose estimation algorithm based on point correspondences," *Acta optica sinica, China*, vol.36(05), pp. 129-138,2016.
- [4] MIAO X K, ZHU F, HAO Y M, XIA R B, "A new pose estimation method based on uncertainty-weighted errors of the feature points," *Journal of optoelectronics laser. China*, vol. 23(07), pp. 1348-1355, 2012.
- [5] YANG W H, LIN J R, GAO Y, et al. "Pose measurement system of double shield universal compact TBM," Acta optica sinica, China, vol. 35(11), pp. 47-155, 2015.
- [6] WANG T Y, DONG W B, WANG Z Y, "Position and orientation measurement system based on monocular vision and fixed target," *Infrared & laser engineering. China*, vol. 46(04), pp. 153-160, 2017.
- [7] WANG X J, CAO Y, ZHOU K, "Methods of monocular pose measurement based on planar objects," *Opt. Precision Eng.China*, vol,25(01),pp. 274-280,2017.