Accurate Calibration of Stereo Cameras for Machine Vision

Liangfu Li, Zuren Feng, Yuanjing Feng
Institute of System Engineering, Xi'an Jiaotong University
Xi'an, ShanXi Province, China
longford@xjtu.edu.cn

ABSTRACT

Camera calibration is an important task for machine vision, whose goal is to obtain the internal and external parameters of each camera. With these parameters, the 3D positions of a scene point, which is identified and matched in two stereo images, can be determined by the triangulation theory. This paper presents a new accurate estimation of CCD camera parameters for machine vision. We present a fast technique to estimate the camera center with special arrangement of calibration target and the camera model is aimed at efficient computation of camera parameters considering lens distortion. Built on strict geometry constraint, our calibration method has compensated the error for distortion cased by circular features on calibration target, which gets over the relativity influence of every unknown parameters of traditional calibration way and make the error distributed among the constraint relation of parameters, in order to guarantee the accuracy and consistency of calibration results. Experimental results are provided to show that the calibration accuracy is high.

Keywords: Calibration, Camera, Parameter, Vision, Measurement.

1. INTRODUCTION

Reconstructing 3D scene information from 2D images has been an important task for many machine vision applications such as vision guided robots for automated assembly [1], visual servo control, robot navigation, vision based industrial detection and recognition, etc. The goal of machine vision is making machine simulate the vision function of mankind, to reconstruct 3D information with stereo images. Vision based sensing, also called optical gauging, is a technique for making displacement measurements based on the relative position of some type of pattern or feature in the field of view of a vision sensor such as a CCD camera [2]. Camera is the main tool for getting original 3D information of machine vision. Vision based systems with camera must determine camera parameter, so camera calibration is a crucial step to obtaining 3D information with camera-based vision systems. Camera calibration is to determine the internal camera geometric and optical characteristics, along with the 3D position and orientation of the camera frame relative to a certain world coordinate system, which are called extrinsic parameters, in order to establish the connection between the position of image pixel and scene point. For any vision applications where more accurate 3D world coordinates should be determined from their 2D image coordinate, the camera calibration with higher accuracy must be obtained first [3].

Due to various applications, there are different requirements for camera calibration. In such applications as robot guidance, the calibration procedure should be fast and automatic, but in industrial metrology applications, the precision is typically a more important factor. Much work has been done in the camera calibration field [4], [5], [6], [7], and we can categorize those calibration methods generally into three classes:

The traditional methods using the 3D calibration target:
Camera calibration is performed by observing a calibration target consisting of two or three planes orthogonal to each other, whose 3D geometry dimension is known with good precision [4], but this approach requires expensive equipment and elaborate installation.

2D plan based calibration:
Camera calibration requires observing a planar pattern shown at a few different orientations, while the knowledge of the plane motion is not necessary. Since almost anyone can make such a calibration device by himself, the installation is very easier than the traditional method, but the camera parameters can't be estimated reliably [5], [6].

Self-calibration method:
Camera calibration does not use any calibration object, just by moving a camera in a static scene, so only image point correspondences are required [7]. Although no calibration object is necessary, a large number of parameters require to be estimated, resulting in a much harder mathematical problem and unreliable results.

Our current research is focused on obtaining the six DOF pose of our movable robot. In the most general case, an object's position and orientation can be described by six-DOF like the robot named Stewart platform [8], so we use stereo vision based system as shown in Fig.1 to measure its accurate pose, accordingly achieve visual servo control. We deal with the calibration of cameras mounted on the wall of our experimented room, to perform real time localizations of our movable robot. After we obtained the movement information and error information of robot with vision based system, we can achieve precisely close-loop feedback servo control.

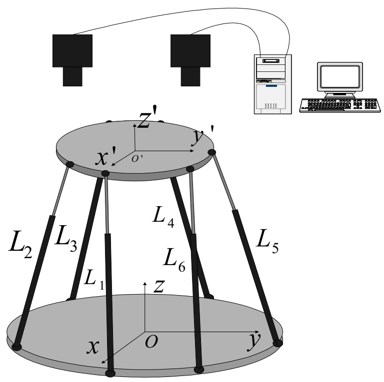

Fig. 1 Vision based Stewart platform

In order to carry out high precision and high-speed visual servo control of robot, constructing an exact and proper model is very important to the accuracy and efficiency of CCD camera parameters. In this paper we present a new camera calibration method based on accurate model and rapid algorithm. Compared with calibration ways in the literature, our method can get definite, reliable and high accuracy parameter, while the algorithm is very simple.

2. CAMERA MODEL

Camera model is the mathematic description of the physics process from scene imaging to image plane. The perspective projection of eye and camera can be approximately looked as a pinhole model. Here we will use pinhole model considering lens distortion to describe the image forming in the camera.

Let us first construct an ideal pinhole camera model as illustrated in Fig.2 to describe perspective transform, where (x_w, y_w, z_w) is the coordinate of the target point P in the 3D world coordinate system, and (x_c, y_c, z_c) is the coordinate of the target point P in the 3D camera coordinate system, whose corresponding image coordinate is (u,v). The image of a 3D point $P=[x_w, y_w, z_w, 1]^T$ on the target of camera $p=[u,v,1]^T$ can be described as a perspective projection of P on p through the optics center o_c, whose distance to the image plane is the effective focal length f, expressed in homogenous notation in the following matrix equation:

$$\lambda p = C\begin{bmatrix} R & T \end{bmatrix} P \tag{1}$$

where λ is a scale factor, R is a 3×3 orthonormal rotational matrix, and $T=[t_x\ t_y\ t_z]^T$ is a 3×1 translation vector, which represent the relative rotation and translation between the world reference frame and the camera coordinate frame respectively. C is the camera intrinsic matrix and can be denoted as:

$$C=\begin{bmatrix} N_x f & 0 & u_0 \\ 0 & N_y f & \mathrm{v}_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \alpha & 0 & u_0 \\ 0 & \beta & \mathrm{v}_0 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

where N_x and N_y are the number of pixels per unit distance in the row and column respectively, u_0 and v_0 are the coordinates of the principal point in pixels.

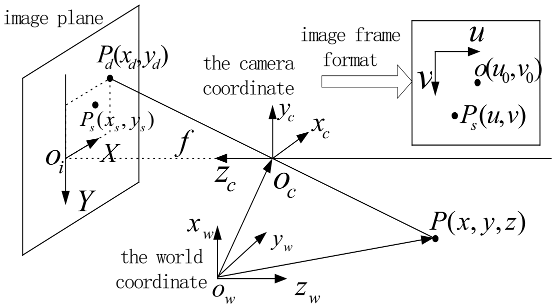

Fig.2. Camera model

The pinhole model is only an approximation of the real camera projection, however, it is not valid for the requirement of high accuracy. In the industrial machine vision application, radial distortion is the main factor that has effect on its accuracy [4], so only radial distortion needs to be considered, and it can be expressed approximately as:

$$\begin{cases} x_s = x_d(1+kr_s^2) \\ y_s = y_d(1+kr_s^2) \end{cases} \tag{3}$$

where (x_s, y_s) is the actual image coordinate which differs from (x_d, y_d) due to lens distortion, $r_s^2 = x_s^2 + y_s^2$, and k is the lens distortion coefficient.

3. CALIBRATION STRATEGY

This section provides the strategy on how to effectively solve the camera calibration problem for machine vision. To improve the calibration accuracy, the circular features have been used as calibration pattern for its property of the noise immunity [9], so we use calibration board with circular feature points. To model the way that cameras project the 3D world into 2D image we need to first know where the camera's image center is, whose accuracy has effect on the accuracy of camera model. R.G.Wilson and S.A.Shafer [10] have summarized 15 techniques for measuring image center, but almost all of them required expensive equipment and precise setup, while the results is unreliable. We use a fast and accurate technique to estimate the camera center and other calibration parameters with error compensation for circular feature detection.

Predetermine the camera center:

To predetermine the camera center, we can use a special calibration way. We arrange the calibration target and take two images at two different depths d_1 and d_2 from the optics center, both of which are perpendicular to the optical axis. As shown in Fig.3, this arrangement of calibration target can eliminate the unknown intrinsic parameters except the camera center by the geometry constraint. The origin of the camera coordinate (x_c, y_c, z_c) is assumed to be located at the optics center o_c. If two calibration points whose distance to the optics axis are h_1 and h_2 have the image point height b and a respectively, we can get the following equation through triangle geometry theorem:

$$\begin{cases} \dfrac{a}{f} = \dfrac{h_1}{d_1} \\ \dfrac{b}{f} = \dfrac{h_1}{d_2} \end{cases} \tag{4}$$

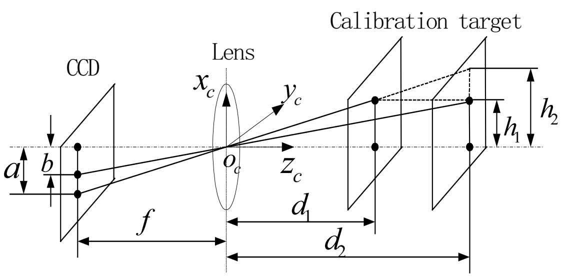

Fig.3 Predetermine camera center

Then we can conclude that:

$$\frac{d_2}{d_1} = \frac{a}{b} \tag{5}$$

Let (u_0, v_0) be the image center and $a = u_1 - u_0$, $b = u_2 - u_0$, then we have:

$$\frac{a}{b} = \frac{u_1 - u_0}{u_2 - u_0} \tag{6}$$

And the same we have:

$$\frac{a}{b} = \frac{v_1 - v_0}{v_2 - v_0} \tag{7}$$

So we can obtain (u_0, v_0) by solving the following equation with least-squares algorithm:

$$\begin{bmatrix} v_2 - v_1 & u_1 - u_2 \end{bmatrix} \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} = \begin{bmatrix} u_1 & -u_2 \end{bmatrix} \begin{bmatrix} v_2 \\ v_1 \end{bmatrix} \tag{8}$$

Calibration other camera parameters:
Extract some feature points from a 3D target, detect each calibration point (u_i, v_i), $i = 1, \cdots, N$, whose responding world coordinate is (x_{wi}, y_{wi}, z_{wi}), and take the calibration points the number is bigger than five and its corresponding image points, we can get a out-stable equation, then solve the variables through the least-square solution:

$$\rho_x = t_x / t_y, \quad \rho_i = r_i / t_y, \quad i = 1,2,4,5 \tag{9}$$

So we can obtain t_y with:

$$t_y^2 = \frac{\eta - [\eta^2 - 4(\rho_1\rho_5 - \rho_4\rho_2)^2]^{1/2}}{2(\rho_1\rho_5 - \rho_4\rho_2)^2} \tag{10}$$

where $\eta = \rho_1^2 + \rho_2^2 + \rho_4^2 + \rho_5^2$.
For each calibration point, we can obtain:

$$\begin{cases} f \cdot x_c \cdot (1 + kr_s^2) = x_s \cdot z_c \\ f \cdot y_c \cdot (1 + kr_s^2) = y_s \cdot z_c \end{cases} \tag{11}$$

where $r_s^2 = x_s^2 + y_s^2$, and k is the lens distortion coefficient. Without loss of the generality, let $z_w = 0$, suppose that:

$$\begin{cases} E_x = r_1 x_w + r_2 y_w + t_x \\ E_y = r_4 x_w + r_5 y_w + t_y \\ E_z = r_7 x_w + r_8 y_w + t_z \end{cases} \tag{12}$$

and let $m = f \cdot k$, $G = r_7 x_w + r_8 y_w$, that is: $x_c = E_x$, $y_c = E_y$, $z_c = E_z$, $G = E_z - t_z$; then we have:

$$\begin{cases} E_x \cdot f + E_x \cdot r_s^2 \cdot m - x_s \cdot t_z = x_s \cdot G \\ E_y \cdot f + E_y \cdot r_s^2 \cdot m - y_s \cdot t_z = y_s \cdot G \end{cases} \tag{13}$$

For N feature points whose (u, v) and its corresponding (x_w, y_w, z_w) are known, we can get the optimized f, k, t_z with the least-squares method.

Error compensation of circular feature:
Circular features as circle and ellipse are not only the basic element in nature but also very common shapes in many man-made objects, which have been commonly used in robot vision fields. The accurate center of projected circles provides exact correspondences between ellipses in 2D image plane and circles in 3D plane. The perspective projection of a spatial circle is commonly an ellipse, however, the center of projected circle does not correspond to the ellipse center, as illustrated in Fig.4. Let both the world coordinate system $\Omega_1(x_w, y_w, z_w)$ and the camera coordinate system $\Omega_2(x_c, y_c, z_c)$ be centered in the optics center o, and let z_w axis be orthogonal to the object plane Π_1, also let z_c axis be perpendicular to the CCD plane Π_2, whose corresponding image axis u and v are parallel to x_c and y_c respectively. If the direction angle of vector $\overline{oo_1}$ is α, β, γ in Ω_1, and the distance between o_1 and o_2 equal to d, then the projection of the circle Γ_1 on the $x_w o y_w$ plane is a circle Γ_3 whose radius r is known, which can be given by:

$$(x_w - z_w \cdot \frac{\cos\alpha}{\cos\gamma})^2 + (y_w - z_w \cdot \frac{\cos\beta}{\cos\gamma})^2 = (r)^2 \tag{14}$$

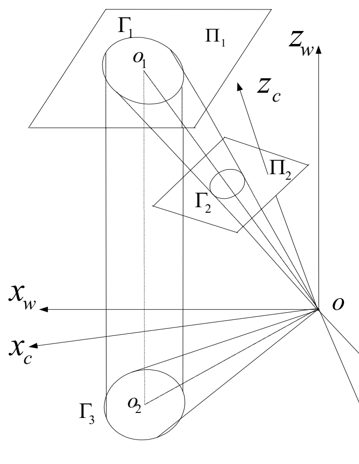

Fig.4. Perspective projection of circular feature

So the rays coming from the circle Γ_1 located on the plane Π_1 form a skewed cone, whose boundary curve can be expressed as:

$$(x_w - z_w \cdot a_1)^2 + (y_w - z_w \cdot a_2)^2 = (z_w a_3)^2 \tag{15}$$

where $a_1 = \frac{\cos\alpha}{\cos\gamma}$, $a_2 = \frac{\cos\beta}{\cos\gamma}$, $a_3 = \frac{r}{d}$

The relationship between the world coordinate system Ω_1 and the camera coordinate system Ω_2 can be given by:

$$\begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = \begin{bmatrix} b_1 & b_2 & b_3 \\ b_4 & b_5 & b_6 \\ b_7 & b_8 & b_9 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \tag{16}$$

where the vectors $[b_1, b_4, b_7]^T$, $[b_2, b_5, b_8]^T$, $[b_3, b_6, b_9]^T$ form an orthonormal basis. The orthogonal distance between the optics center o and the image plane Π_2 is the focal length f, so we can express Eq. (15) in the camera coordinate system:

$$
\begin{aligned}
&(g^2+l^2-p^2)x_c^2+2(gh+lm-pq)x_cy_c \\
&+(h^2+m^2-q^2)y_c^2+2(gk+ln-ps)x_c \\
&+2(hk+mn-qs)y_c+k^2+n^2-s^2=0
\end{aligned}
\tag{17}
$$

where $g=b_1-a_1b_7$

$h=b_2-a_1b_8$

$k=(b_3-a_1b_9)f$

$l=b_4-a_2b_7$

$m=b_5-a_2b_8$

$n=(b_6-a_2b_9)f$

$p=a_3b_7$

$q=a_3b_8$

$s=a_3b_9f$.

As we know that a common quadratic curve can be expressed as the form of:

$$Ax_c^2+Bx_cy_c+Cy_c^2+Dx_c+Ey_c+F=0 \tag{18}$$

and the projection of a common circle on the image plane is an ellipse, then the image of the circle Γ_1 is an ellipse Γ_2 located on the image plane Π_2. Thus the center point coordinates of the ellipse can be calculated using the following formula:

$$
\begin{cases}
u_c=\dfrac{c_1c_2-c_3c_4}{(gq-hp)^2+(pm-ql)^2-(gm-lh)^2} \\
v_c=\dfrac{c_1c_5-c_2c_3}{(gq-hp)^2+(pm-ql)^2-(gm-lh)^2}
\end{cases}
\tag{19}
$$

where: $c_1=(h^2+m^2-q^2)$

$c_2=(gk+ln-ps)$

$c_3=(gh+lm-pq)$

$c_4=(hk+mn-qs)$

$c_5=(g^2+l^2-p^2)$.

In the camera coordinate system the equation of the line oo_1 can be give by:

$$
\begin{aligned}
&\frac{b_1x_c+b_2y_c+b_3z_c}{\cos\alpha}=\frac{b_4x_c+b_5y_c+b_6z_c}{\cos\beta} \\
&=\frac{b_7x_c+b_8y_c+b_9z_c}{\cos\gamma}
\end{aligned}
\tag{20}
$$

The real coordinate of the ellipse center is the intersection of the line oo_1 and the plane $z_c=f$ is:

$$
\begin{cases}
\tilde{u}_c=\dfrac{d_3d_5-d_2d_6}{d_1d_5-d_2d_4} \\
\tilde{v}_c=\dfrac{d_1d_6-d_3d_4}{d_1d_5-d_2d_4}
\end{cases}
\tag{21}
$$

where: $d_1=b_1\cos\gamma-b_7\cos\alpha$

$d_2=b_2\cos\gamma-b_8\cos\alpha$

$d_3=(b_9\cos\alpha-b_3\cos\gamma)f$

$d_4=b_4\cos\gamma-b_7\cos\beta$

$d_5=b_5\cos\gamma-b_8\cos\beta$

$d_6=b_9\cos\beta-b_6\cos\gamma$.

The observed image coordinate (u,v) should be corrected with the error compensation:

$$
\begin{cases}
u=u-N_x(\tilde{u}_c-u_c) \\
v=v-N_y(\tilde{v}_c-v_c)
\end{cases}
\tag{22}
$$

After the error compensation, the camera parameters are computed again, so we can get more accurate camera parameter, to guarantee high accuracy for machine vision.

4. EXPERIMENT RESULTS

In this system, we use the camera named TM1400, which is made in Pulnix Company of United States with high resolution 1392*1040, and digital image board named Pc2-camlink, which is made in Coreco Company of Canada, to form a high accuracy measurement system. We use 64 calibration points of circular feature, to achieve these camera parameters as shown in table.1.

Table 1. Camera parameters using real images

Camera parameters		First camera	Second camera
Rotation matrix *R*	*r1*	0.999849	0.999988
	r2	0.000203	-0.002123
	r3	0.017351	-0.004424
	r4	-0.004127	-0.000880
	r5	-0.968312	-0.966059
	r6	0.249710	0.258321
	r7	0.016852	-0.004822
	r8	-0.249744	-0.258314
	r9	-0.968165	-0.966049
Translation *T*	*tx*	-281.46484	-848.46375
	ty	83.014908	52.595356
	tz	5130.63525	5137.75879
Intrinsic parameter	*u0*	687.386	692.437
	v0	540.894	551.932
	f	13.988143	13.896669
	k	-0.000362	-0.001011

To check the accuracy of our system, we use 32 test points to measure the difference between their real coordinate and reconstructed coordinate with:

$$\varepsilon=\frac{\sum_{i=1}^{N}\left|(X_s-X_d)\right|}{N} \tag{23}$$

where X_s is real 3D coordinate while X_d is the measurement value, and *N* is the number of test points. The depth from the baseline to the calibration board is 4.618m, and the average error of our measurement system is ε =0.16mm, so that the calibration accuracy is very high. We also have conducted other experiments. General speaking, this method is quite accurate.

8. CONCLUSIONS

In this paper, an accurate estimation of CCD camera parameters was presented for machine vision applications where high accuracy is needed. We use a fast technique to estimate the camera center with proper arrangement of calibration target, and obtain other parameters through logical organization of solving order. Built on strict geometry constraint, our calibration method has compensated the error for distortion cased by circular features on calibration target, which gets over the relativity influence of every unknown parameters of traditional calibration way and make the error distributed among the constraint relation of parameters. This representation makes the decomposing of all camera parameters possible, and leads to the parameter estimation one by one. Compared with classical calibration techniques that use expensive equipment and complicated mathematical.

9. REFERENCES

[1] F.G. King, G.V. Puskorius and F. Yuan, "Vision guided robots for automated assembly", Proc. 1988 IEEE Int. Conf. Robotics and Automation, Vol. 3, April 1988, pp. 1611-1616.
[2] Harry Douglas Garner, "Development of a real-time vision based absolute orientation sensor", Ph.D dissertation, Georgia Institute of Technology, 2001.
[3] L. Jong-Soo and J. Yu-Ho, "CCD camera calibrations and projection error analysis", Proc. 2000 IEEE Int. Conf. Science and Technology, Korea-Russia, Vol. 2, 2000, pp. 50-55.
[4] R.Y. Tsai, "A versatile camera calibration technique for high-accuracy 3D machine vision metrology using off-the-shelf TV cameras and lenses", IEEE J. Robotics and Automation, Vol. 3, No. 4, Aug. 1987, pp. 323-344.
[5] Z.Y. Zhang, "A flexible new technique for camera calibration", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 22, No. 11, Nov. 2000, pp.1330-1334.
[6] J. Heikkilä, "Geometric camera calibration using circular control points", IEEE Trans. Pattern Analysis and Machine Intelligence, Vol. 22, No. 10, Oct. 2000, pp. 1066-1077.
[7] R.I. Hartley, "Self-calibration from multiple views with a rotating camera", Proc. 1994 IEEE Int. Conf. Computer Vision and Pattern Recognition, May 1994, pp. 471-478.
[8] D. Stewart, "A platform with six degree of freedom", Proc. Inst. Mech. Eng., Vol. 180, No. 5, 1965, pp. 371-378.
[9] J. Heikkilä, "A four-step camera calibration procedure with implicit image correction", Proc. 1997 IEEE Int. Conf. Computer Vision and Pattern Recognition, June 1997, pp. 1106-1112.
[10] R.G. Wilson and S.A. Shafer, "What is the center of the image", IEEE J. Optical Society of America, Vol. 11, No. 11, Nov. 1994, pp. 2946-2955.