Calibrating distance mapping of non-SVP catadioptric camera of the soccer robot

Xiaoxiao Zhu and Qixin Cao

State Key Lab of Mechanical Systems and Vibration, Research Institute of Robotics, Shanghai Jiao Tong University, Shanghai, China

Abstract

Purpose – The purpose of the present paper is to propose a full model-based method for distance-mapping calibration for the non-SVP (non-single viewpoint) catadioptric camera of the soccer robot. The method should be easy to operate, efficient, accurate, and scalable to fit larger field sizes. **Design/methodology/approach** – The distance-mapping model was first constructed based on the imaging principle. The authors then calibrated the internal parameters using the mirror boundary and used the mirror center to choose the correct pose from two possible solutions. The authors then proposed a three-point method based on a unique solution case of the non-SVP P3P (perspective-three-point) problem to solve the external parameters. Lastly, they built the distance mapping by back-projection.

Findings – The simulation experimental results have shown that the authors' method is very accurate even when there is severe misalignment between the mirror and the camera and that all calibration operations, except the calibration of a standard camera, can be completed in 1 min. The result of the comparison with the traditional calibration method shows that the authors' method is superior to the traditional method in terms of accuracy and efficiency.

Originality/value – The proposed calibration method is scalable to larger fields because it only uses the boundary of the mirror and three feature points on the field, and does not need additional calibration objects. Additionally, an automatic calibration method that can be used during the game can be easily developed based on this method. Moreover, the proposed mirror-pose-selection method and a unique solution to the non-SVP P3P problem are especially useful for a non-SVP catadioptric camera.

Keywords Non-SVP catadioptric camera, Distance mapping, Calibration, Soccer robot, Robots, Football, Cameras

Paper type Research paper

1. Introduction

A catadioptric camera is a vision system consisting of a standard camera directed upwards toward a mirror (Daniilids and Gever, 2001). As a type of omni-directional camera, it has a very large field of view (FOV), typically 360° in azimuth and 90° -140° in elevation, which is a great advantage for many applications. Generally, catadioptric cameras can be divided into two types: single viewpoint (SVP) and non-SVP. The SVP catadioptric camera maintains a single viewpoint just like a regular camera, and so it has a simple mathematical model. According to Baker and Nayar (1999), to construct an SVP catadioptric camera, the mirror and camera must meet two criteria: the form of the mirror surface should be parabolic, hyperbolic, or elliptical, and the mirror and camera should be aligned in a rigid relative position. In practice, it is difficult and costly to build an accurate SVP catadioptric camera, even impossible in some special cases. In the Middle Size League of RoboCup competition (MSL), soccer robots on most of the teams are equipped with the non-SVP catadioptric cameras

The current issue and full text archive of this journal is available at www.emeraldinsight.com/0143-991X.htm



Industrial Robot: An International Journal 40/5 (2013) 462–473 © Emerald Group Publishing Limited [ISSN 0143-991X] [DOI 10.1108/IR-09-2012-404] dissatisfy those two criteria and they have complicated mathematical models.

The MSL is one of the most interesting leagues in the RoboCup (Robot Soccer World Cup Kitano *et al.* (1997)), in which several midsized (the largest robot is $50 \text{ cm} \times 50 \text{ cm} \times 80 \text{ cm}$) soccer robots play soccer autonomously. In the game, the soccer robot must perceive the surrounding environment mostly by itself, rarely by communicating with its teammates, and then make a decision as to what to do; therefore, the performance of the sensors of the soccer robot is important. During the development of the MSL, many types of sensors were used by soccer robots, such as odometers, ultrasonic sensors, laser range finders, and infrared field detectors. The non-SVP catadioptric camera is now the main sensor in soccer robots for almost all teams.

Generally, a calibration process is necessary before using any type of the vision system, and this process varies according to the model of the system and the application. The main task of a non-SVP catadioptric camera in a soccer robot is to detect surrounding objects such as the ball, other robots, and the white lines on the field, and then measure the distances to those objects. Usually, because of the plane motion properties of a soccer robot, we measure the distance between the robot base and the contact point of those objects on the field.

This paper is supported by National Natural Science Foundation of China (No. 61273331) and Research Foundation of State Key Lab of MSV, China (No. MSV201101). The authors would like to thank Shouhong Miao, Wei Wang and Aolin Tang for their help in paper writing.

Therefore, we need to calibrate the distance mapping between the real distance on the soccer field relative to the robot and the image distance in the omni-directional image captured by the catadioptric camera. In the case of the MSL, this calibration usually needs to be done frequently because of the displacement of the camera system in transportation or during the game. Therefore, the calibration needs to be accurate as well as efficient. Currently, there are two types of calibration methods for distance mapping:

- Interpolation method. The interpolation method uses a very 1 simple principle. First, it builds an initial mapping between several image points and world points on the soccer field using some calibration objects. This is done manually by most of the teams, but it also can be done automatically (Tribots code, 2005). The residual image points can then be mapped to the world points by interpolating them from the initial mapping. Interpolation method is the traditional calibration method that is widely used in soccer robot field. Implementing like a black box algorithm, it can work well without the knowledge of the surface equation of the mirror and the parameters of the camera. On the other hand, it will perform poorly when the misalignment between the mirror and the camera is severe. Another disadvantage of this method is that, it is time consuming, particularly to create the initial mapping. That is because it needs a large number of initial corresponding points evenly distributed in the omni directional image, but restricted to the size of the calibration objects not all of those points can be obtained at one time, thus usually, the robot needs to spin around, which causes a lot of time. Moreover, along with the increasing of the size of the soccer field, it requires larger calibration objects which will make it difficult to be used.
- Model-based method. The model-based method first defines 2 a mathematical model of the distance mapping, and then calibrates the unknown parameters of the model. The model can be divided into an approximate model and an accurate (or full) model. The approximate mapping model uses an empirical formula to model distance mapping, usually assuming every direction is the same. This means that the axes of the camera and the mirror must coincide. The calibration process is similar to that of the interpolation method. It first obtains corresponding points between the image and the environment and then uses those points to calibrate the model parameters. Heinemann et al. (2006) proposed to apply evolutionary algorithms to solve the mapping function automatically, but their model considered only the displacement between the axis of the camera and the symmetry axis of the mirror. The advantage of using the evolutionary algorithm is that it does not need the initial value of those model parameters. Zongjie et al. (2008) added the image center parameter into the approximate model, and proposed an error descent algorithm to optimize those model parameters. Because they do not consider the robot position in their model, they need to put the robot in a known position on the field. On the other hand, the accurate model is based on the imaging principle, so it can handle the complete non-SVP case. The CAMBADA team proposed an accurate model method (António et al., 2011; Cunha et al., 2006), and they tried to calibrate the model parameters partly by measuring the setup itself and partly by analysis of the images of objects such as the mirror boundary, the center of the

Industrial Robot: An International Journal

Volume 40 · Number 5 · 2013 · 462-473

mirror image, the center of the robot image, as well as the radius, distance, and eccentricity of the game field lines, mainly the midfield circle, lateral, and area lines. Their method is difficult to operate.

Outside the domain of the RoboCup, the calibration methods of non-SVP catadioptric cameras for general use, most of which are 3D reconstruction applications (Mičušík and Pajdla, 2004), also provide much inspiration. Strelow et al. (2001) presented a full model of the imaging process, which includes the rotation and translation between the camera and mirror, and an algorithm that determines this relative position from observations of a batch of known points in a single image. This algorithm can tackle various amounts of misalignment between the mirror and camera, but it is difficult to prepare those 3D marker points. Mashita et al. (2005) proposed a method using the mirror boundary to calibrate the camera-to-mirror relation, which is a conic-based analytical method that can avoid the initial value and local minimum problems arising from nonlinear optimization. Zhiyu et al. (2012) also use the mirror boundary to calibrate the camera-to-mirror relation, but they propose to select the correct pose by the image of the lens. Maxime (2008) calibrated the catadioptric camera using multi view geometry (Hartley, 2003). Their method needs a rich-featured environment and needs many images to be taken, and it only calibrates the internal parameters.

The present paper proposes an improved accurate model method. We also used the boundary of the mirror to calibrate the mirror-to-camera pose similar to Mashita *et al.* (2005) and Zhiyu *et al.* (2012), but a simple and practical pose selection method is proposed to select the correct solution from the two possible solutions. Then, a unique solution non-SVP P3P problem is proposed and used for calibrating the mirror-to-world transformation. Last, according to the model and the parameters, we built a distance-mapping matrix using a back-projection method. In the experiment, we built a simulation environment, and based on synthetic data, we tested the calibration accuracy.

The paper is organized as follows. In Section 2, we construct the imaging principle model of a non-SVP catadioptric camera. Section 3 describes the calibration procedure. Section 4 provides the experimental results, and in Section 5, we present our conclusion.

2. Distance mapping model

In this section, we model the distance mapping between the image points and the real world points based on the model developed by Mashita *et al.* (2005), extend it to an arbitrary type of revolving mirror, and add a robot-centered coordinate frame.

2.1 Notation

A 2D vector $p = \begin{bmatrix} u & v \end{bmatrix}^T$ represents the image coordinate of an image point p, and a 3D vector $P^* = \begin{bmatrix} x & y & z \end{bmatrix}^T$ represents the coordinate of a point *P* in the coordinate frame **, which can be the *r* (robot), *w* (world), *c* (camera), or *m* (mirror) coordinate frame. We use \bar{x} to denote the homograph vector by adding 1 as the last element: $\bar{p} = \begin{bmatrix} u & v & 1 \end{bmatrix}^T$ and $\bar{P}^* = \begin{bmatrix} x & y & z & 1 \end{bmatrix}^T$. A 3 × 3 matrixes \mathbf{R}_{ij} and a 3 × 1 vector T_{ij} represent the coordinate transformation from coordinate frame i to j (i, j are the same as *), with the transformation equation (1):

$$P_j = \mathbf{R}_{ij} P_i + T_{ij} \tag{1}$$

A camera is modeled using the usual pinhole model, and image distortion is not considered, since the distortion can be corrected before calibration. The camera calibration matrix **K** and the projection relationship between a 3D point P and its image point p are given by equation (2):

$$\mathbf{s}\bar{p} = \mathbf{K}[\mathbf{I}|0]\bar{P}_c = \mathbf{K}P_c \quad \mathbf{K} = \begin{bmatrix} \alpha & \mathbf{c} & \mathbf{u}_0 \\ 0 & \beta & \mathbf{v}_0 \\ 0 & 0 & 1 \end{bmatrix}$$
(2)

where s is a scale factor, α and β is the focus length, c is the skew value, u_0 and v_0 is the coordinates of the image center.

We use the abbreviations ICS, CCS, WCS, and RCS for the coordinate systems of the image plane, camera, world, and robot, respectively.

The mirror surface in the MCS is represented by $f_m(X) = 0$, which is assumed to be known in our method.

2.2 Modeling

Figure 1 shows the structure of the non-SVP catadioptric camera. To map from an image point to a world point, we assume that one ray emanates from the camera center O_c , passes through the image point p, meets the mirror at M, and is then reflected to point P. We can obtain the mapping through the following steps.

Because point M on the mirror surface is projected onto the image plane at point p, from equation (2), we obtain M_c with a scale factor as:

$$s\bar{p} = \mathbf{K}M_c \Rightarrow M_c = s\mathbf{K}^{-1}\bar{p}$$
 (3)

Figure 1 Imaging principle of non-SVP catadioptric camera

Industrial Robot: An International Journal

Then coordinate of M in the MCS is expressed as the following equation, in which vi_m is the direction vector of the incident light:

$$M_m = \mathbf{R}_{\rm cm} M_c + T_{cm} = {\rm s} \mathbf{R}_{\rm cm} \mathbf{K}^{-1} \bar{p} + T_{cm} = {\rm s} v i_m + T_{cm} \quad (4)$$

Point M expressed in the WCS is:

$$M_w = \mathbf{R}_{\mathrm{mw}} M_m + T_{mu}$$

The normal vector of the mirror at point M can be represented as $n_m = [(\partial f_m / \partial x)_{M_m} \quad (\partial f_m / \partial y)_{M_m} \quad (\partial f_m / \partial z)_{M_m}]^{\mathrm{T}}$. If we normalize the normal vector as $\hat{n}_m = n_m / ||n_m||$, we can compute the reflected ray vo_m as:

$$vo_m = vi_m - 2(vi_m \cdot \hat{n}_m) \cdot \hat{n}_m$$

Vector vo_m can be expressed in the WCS, i.e.:

$$vo_w = \mathbf{R}_{mw} vo_m$$

The point *P*lies on the reflected ray *vo*, so we can present *P* in the WCS as:

$$P_{w} = M_{w} + tvo_{w}$$

$$= (\mathbf{R}_{mw}(\mathbf{s}\mathbf{R}_{cm}\mathbf{K}^{-1}\bar{p} + T_{cm}) + T_{mw})$$

$$+ t\mathbf{R}_{mw}(\mathbf{R}_{cm}\mathbf{K}^{-1}\bar{p} - 2(\mathbf{R}_{cm}\mathbf{K}^{-1}\bar{p}\cdot\hat{n}_{m})\cdot\hat{n}_{m})$$

$$(t > 0)$$

$$(5)$$

where the t is a scale factor.



Finally, expressing the point *P* in the RCS, we have:

$$P_{r} = \mathbf{R}_{wr} P_{w} + T_{wr}$$

$$= \mathbf{R}_{wr} ((\mathbf{R}_{mw}(\mathbf{s}\mathbf{R}_{cm}\mathbf{K}^{-1}\bar{p} + T_{cm}) + T_{mw})$$

$$+ \mathbf{t}\mathbf{R}_{mw}(\mathbf{R}_{cm}\mathbf{K}^{-1}\bar{p} - 2(\mathbf{R}_{cm}\mathbf{K}^{-1}\bar{p}\cdot\hat{n}_{m})\cdot\hat{n}_{m})) + T_{wr}$$
With $\hat{n}_{m} = \frac{n_{m}}{\|n_{m}\|},$

$$n_{m} = \left[\left(\frac{\partial f_{m}}{\partial \mathbf{x}}\right)_{M_{m}} \quad \left(\frac{\partial f_{m}}{\partial \mathbf{y}}\right)_{M_{m}} \quad \left(\frac{\partial f_{m}}{\partial \mathbf{z}}\right)_{M_{m}} \right]^{\mathrm{T}}, \quad \mathbf{t} > 0$$
(6)

The final equation (6) shows the relation between the image point \bar{p} and the world point P_r . Note that in equation (6), the unknown variables are **K**, \mathbf{R}_{cm} , T_{cm} , s, \mathbf{R}_{mw} , T_{mw} , \mathbf{R}_{wr} , T_{wr} and t. However, because we assume that the surface function $f_m(X) = 0$ is known so that s can be solved by letting $f_m(M_m) = 0$ and because P lies in the x-y plane so that the z-coordinate of P_w is zero, t can also be solved. Finally, the remaining unknown variables are **K**, \mathbf{R}_{cm} , T_{cm} , \mathbf{R}_{mw} , T_{mw} , \mathbf{R}_{wr} and T_{wr} . As usual, **K**, \mathbf{R}_{cm} and T_{cm} are defined as the internal parameters, and \mathbf{R}_{mw} , T_{mw} , \mathbf{R}_{wr} and T_{wr} are defined as the external parameters.

3. Calibration

This section provides details on how to solve the unknown variables in equation (6). We will calibrate the internal and external parameters separately.

3.1 Internal parameter calibration

As mentioned above, the internal parameters contain the camera calibration matrix **K** and the camera-to-mirror transformations \mathbf{R}_{cm} and T_{cm} . The calibration of **K** has been widely researched as a basic problem in the computer vision domain, such as Zhang's (1999) method, and in the implementation the Matlab camera calibration toolbox provided by Bouguet (2008). The problem calibrating \mathbf{R}_{cm} and T_{cm} is the same as a mirror pose estimation problem. In our case, we used the circle contour of the mirror base to determine the pose of the mirror referring to Chen *et al.* (2004).

Assuming the quadratic curve equation of the image of the circular boundary of the mirror base in the ICS is:

$$Ax^{2} + 2Bxy + Cy^{2} + 2Dx + 2Ey + F = 0$$
 (7)

Denote:

$$\mathbf{Q} = \begin{pmatrix} A & B & D \\ B & C & E \\ D & E & F \end{pmatrix},$$

and let $\lambda = (\lambda_1 \ \lambda_2 \ \lambda_3)$ and $\mathbf{V} = (v_1 \ v_2 \ v_3)$ be the eigenvalues and normalized eigenvectors of \mathbf{Q} . Let z_0 be the distance from the projected center to the base of the mirror, and r be the radius of the base boundary. Then we have:

Industrial Robot: An International Journal

Volume 40 · Number 5 · 2013 · 462-473

$$\begin{cases} z_0 = S_3 \frac{\lambda_2 r}{\sqrt{-\lambda_1 \lambda_3}} \\ C = z_0 \mathbf{V} \left(S_2 \frac{\lambda_3}{\lambda_2} \sqrt{\frac{\lambda_1 - \lambda_2}{\lambda_1 - \lambda_3}} & 0 & -S_1 \frac{\lambda_1}{\lambda_2} \sqrt{\frac{\lambda_2 - \lambda_3}{\lambda_1 - \lambda_3}} \right)^{\mathrm{T}} \\ N = \mathbf{V} \left(S_2 \sqrt{\frac{\lambda_1 - \lambda_2}{\lambda_1 - \lambda_3}} & 0 & -S_1 \sqrt{\frac{\lambda_2 - \lambda_3}{\lambda_1 - \lambda_3}} \right)^{\mathrm{T}} \\ \begin{cases} C \cdot (0 & 0 & 1)^{\mathrm{T}} < 0 \\ N \cdot (0 & 0 & 1)^{\mathrm{T}} > 0 \end{cases}$$
(9)

where C is the center of the boundary circle, N is the normal vector of the mirror base, and S_1 , S_2 and S_3 are undetermined signs.

Using equation (9), the condition that the mirror is in front of the camera and the direction of the z-axis of the MCS is consistent with the CCS, we can only determine two of the three undetermined signs in equation (8), and so we have two solutions remaining. Mashita et al. (2005) selected the true solution from the two possible solutions by the line at infinity, which is not suited for an indoor environment because it cannot obtain an image of the line at infinity. And different with Zhiyu et al. (2012), we propose to use the center of the mirror as an additional point to choose the correct pose. When we designed the mirror, we added a platform to the mirror, as in Figure 2(a), and glued a small color marker to the center of the mirror to easily to recognize the center, as in Figure 2(b). We predict the idea image coordinate of the center marker point for each possible pose, and then compared them to the real image coordinate of the center to choose the right pose. The distinguish ability of the mirror center is explained as follows.

Without losing generality, we assume that S_1 is the last undetermined sign and $\lambda_1 \neq \lambda_2 \neq \lambda_3$. Therefore, according to the matrix theory, **V** is a orthogonal matrix, and the column vectors can be treated as a basis vector group of the R^3 space; hence, from equation (8), so the coordinates of *C* and *N* under this group of base are:

$$\begin{array}{ccc} \left(\begin{array}{ccc} z_0 S_2 \frac{\lambda_3}{\lambda_2} \sqrt{\frac{\lambda_1 - \lambda_2}{\lambda_1 - \lambda_3}} & 0 & -z_0 S_1 \frac{\lambda_1}{\lambda_2} \sqrt{\frac{\lambda_2 - \lambda_3}{\lambda_1 - \lambda_3}} \end{array} \right)^T & \text{and} \\ \\ \left(\begin{array}{ccc} S_2 \sqrt{\frac{\lambda_1 - \lambda_2}{\lambda_1 - \lambda_3}} & 0 & -S_1 \sqrt{\frac{\lambda_2 - \lambda_3}{\lambda_1 - \lambda_3}} \end{array} \right)^T. \end{array}$$

The two possible results are shown in Figure 3. In Figure 3, the two possible mirror centers are P_1 and P_2 . Because P_1 and P_2 are symmetrical with the axis v_3 , the projection of them in the image plane are distinguishable unless they are coincident. Therefore, the mirror center point can be used to select the true solution.

3.2 External parameter calibration

The external parameters contain the mirror-to-world transformation \mathbf{R}_{mw} and T_{mw} and the world-to-robot transformation \mathbf{R}_{wr} and T_{wr} . We will focus on the mirror-to-world calibration.

3.2.1 Mirror-to-world calibration

If we know the pose of the WCS's *x-y* plane in the MCS, the mirror-to-world transformation can be solved. This idea leads us to the classical three-point space resection problem or perspective-three-point problem (review by Haralick *et al.* (1994)): if given

Volume 40 · Number 5 · 2013 · 462-473

Figure 2 Illustration of mirror center marker





Figure 3 Illustration of two possible solutions



the perspective projection (the image coordinates) of three points constituting the vertices of a known triangle in 3D space, it is possible to determine the position of each of the vertices in the CCS. Nistér and Stewénius (2007) extended the perspective camera to a general camera, which can be a non-SVP camera, and pointed out that this problem generally has eight solutions. Gao *et al.* (2003) proved that, in reality, if the angles between any two of the back-projected rays corresponding to the three control points are obtuse, the P3P problem can have only one solution. In this paper, we extend Gao's special case to the non-SVP P3P problem and use this unique solution three-point method to determinate the pose of the *x*-*y* plane of the WCS in the MCS.

First, we will construct the non-SVP P3P model and then prove the uniqueness of the solution under the obtuse condition that is the same as that of Gao's. We built the non-SVP P3P model as follows; the model is shown in Figure 4. After calibrating the internal parameters, if we know three image points p^i (i = 1, 2, 3), we can obtain three rays $ray_m^i = M_m^i + tvo_m^i$ (i = 1, 2, 3) (where M_m^i is the intersection point of the rays in the surface of the mirror, and vo_m^i is the direction vector of the reflection rays), which pass through three world points P^i (i = 1, 2, 3). Let d_{ij} be the distance from P^i to P^i , θ_{ij} be the angle between ray_i and ray_j , and h_{ij} be the distance between ray_i and ray_j . The line segment $D_{ij} D_{ji}$ is perpendicular to both ray_i and ray_j and meets ray_i on D_{ij} and ray_j on D_{ji} . We define the lengths from D_{12} to P^1 as t_1 , from D_{21} to P^2 as t_2 and from D_{32} to P^3 as t_3 .

Figure 4 Three-point model for non-SVP catadioptric camera



Variables a,b and c represent the distances of D_{21} to D_{23} , D_{32} to D_{31} , and D_{12} to D_{13} , respectively. By a simple geometric relationship and the cosine theorem, we can obtain equation (10):

$$\begin{cases} d_{12}^2 - h_{12}^2 = t_1^2 + t_2^2 - 2t_1t_2\cos\theta_{12} \\ d_{23}^2 - h_{23}^2 = (t_2 + a)^2 + t_3^2 - 2(t_2 + a)t_3\cos\theta_{23} \\ d_{13}^2 - h_{13}^2 = (t_1 + c)^2 + (t_3 + b)^2 - 2(t_1 + c)(t_3 + b)\cos\theta_{13} \end{cases}$$
(10)

The obtuse condition is:

$$\begin{cases} t_1 > 0, & (i = 1, 2, 3) \\ \theta_{12} > \frac{\pi}{2}, & \theta_{23} > \frac{\pi}{2}, & \theta_{13} > \frac{\pi}{2} \end{cases}$$
(11)

Using equation (11) and representing t_1 and t_3 by t_2 in equation (10):

$$\begin{cases} t_1 = t_2 \cos\theta_{12} + \sqrt{d_{12}^2 - t_2^2 \sin\theta_{12}^2} \\ t_3 = (t_2 + a) \cos\theta_{23} + \sqrt{d_{32}^2 - (t_2 + a)^2 \sin\theta_{23}^2} \\ g(t_2) = (t_1 + c)^2 + (t_3 + b)^2 - 2(t_1 + c)(t_3 + b) \cos\theta_{13} \\ -(d_{13}^2 - h_{13}^2) = 0 \end{cases}$$
(12)

The derivations are:

$$\begin{cases} \frac{d_{t_1}}{d_{t_2}} = \cos\theta_{12} - \frac{t_2}{\sqrt{d_{12}^2 - t_2^2 \sin\theta_{12}^2}} < 0\\ \frac{d_{t_3}}{d_{t_2}} = \cos\theta_{23} - \frac{t_2 + a}{\sqrt{d_{32}^2 - (t_2 + a)^2 \sin\theta_{23}^2}} < 0\\ \frac{d_{g(t_2)}}{d_{t_2}} = 2(t_1 + c)\frac{d_{t_1}}{d_{t_2}} + 2(t_3 + b)\frac{d_{t_3}}{d_{t_2}}\\ -2\cos\theta_{13}\left((t_1 + c)\frac{d_{t_3}}{d_{t_2}} + (t_3 + b)\frac{d_{t_1}}{d_{t_2}}\right) < 0 \end{cases}$$
(13)

In equation (12), we obtain an equation $g(t_2) = 0$, $(t_2 > 0)$, which is a nonlinear function. From the derivation equation (13), $g(t_2)$ is a monotonic decreasing function of t_2 , when $t_2 > 0$. Therefore, if $g(t_2) = 0$, $(t_2 > 0)$ has a solution, that is physically guarded, it must be the unique solution. We can easily obtain a numerical solution for t_2 using a nonlinear function solver with an arbitrary positive initial value or use a simple root search method (Haralick *et al.*, 1994). After we solve for t_1 , t_2 and t_3 , we can obtain the coordinates of the three world points in the MCS, P_m^i , (i = 1, 2, 3).

Then, because P_w^i , (i = 1, 2, 3) is given, we can obtain the three point pairs (P_m^i, P_w^i) , (i = 1, 2, 3) that correspond between the MCS and the WCS in equation (14), which can be solved using the method of Arun *et al.* (1987):

$$P_{w}^{i} = \mathbf{R}_{mw}P_{m}^{i} + T_{mw}$$
 (i = 1, 2, 3) (14)

3.2.2 World-to-robot calibration

Like the solution of \mathbf{R}_{mw} and T_{mw} we calibrate \mathbf{R}_{wr} and T_{wr} with the point pair (P_w^i, P_r^i) . In practice, if we assume that the *x-y* planes of the WCS and the RCS are coincident, then, the unknown variables decrease to three, two from translation and one from rotation; hence, two corresponding point pairs are sufficient.

3.3 Nonlinear optimization

Although the three-point method can obtain a good calibration result of the mirror-to-world parameters, using the result as the initial value, higher accuracy can be obtained by some nonlinear optimization methods. We adapt the optimization method of Zongjie et al. (2008), which uses the back-projection error of the white line points on the soccer field as the optimization objects. The back-projection error is defined as follows. Assuming that a white point p in the ICS is transformed to point P on the soccer field in WCS according to equation (5), we can say that because point P is a white point, it should be located on one of the white lines on the soccer field. Then, the back-projection error of p, notated as e_{k} , is the distance from P to the nearest white line. Thus, the back-projection error function of all the white points in the image is the sum of e_k and can be represented by equation (15). However, according to Lauer et al. (2005), using equation (16) can be more robust when considering the noise, where c is a constant value. The value of c is empirical, and in the original paper by Lauer et al., it is set as 250 mm. At present, however, the soccer field is larger than when the paper was written. Thus, we accordingly increased the value of c to 500 mm. Because, at present, we take the image for calibration statically (the robot is not moving), there is negligible noise and there is little effect from the value of c. However, if we develop an automatic calibration method in the future, which allows **Industrial Robot: An International Journal**

Volume 40 · Number 5 · 2013 · 462-473

calibration when the robot is moving, this value should be checked carefully.

Because we have a very good initial value from the threepoint method, we can use the general Levenberg-Marquardt optimization algorithm to optimize equation (16). Notice that, theoretically, we can optimize both the external and internal parameters simultaneously, but we found that the optimization result is not very good. Moreover, because the calibration of the internal parameters has already been optimized and is very accurate, in this study, we only optimize the external parameters:

$$totalError = \sum_{k=1}^{n} e_k^2$$
(15)

$$totalError = \sum_{k=1}^{n} \left(1 - \frac{c^2}{c^2 + e_k^2} \right) \tag{16}$$

3.4 Building distance mapping

After calibrating all those parameters, it is easy to build the distance mapping by back-projecting the image points onto the RCS based on equation (6). We construct the distance mapping as a matrix with the same size as the image. The index of the matrix element is consistent with the image coordinate of the image points, and the context of the element is the coordinate of the world point in the RCS. This distance-mapping matrix makes the transformation from the image point to the RCS very easy and efficient.

3.5 Summary

The working flow of the proposed calibration method is as follows:

- calibrate the camera by the conventional method;
- use the contour of the mirror to calibrate the mirror-tocamera position and select the correct mirror center;
- use the three-point method to calibrate the mirror-toworld position and optimize the result;
- calibrate the world-to-robot position; and
- build the distance mapping matrix using equation (6).

4. Experimental results

In this section, we report on the experiments conducted to test the accuracy and efficiency of the proposed calibration method, and we compare our method with the traditional calibration method. Computer simulated data were used because the ground truth is easy to obtain in a simulation environment. Notice that the world-to-robot transformation was not calibrated.

The soccer field used for the test had a standard size of $18 \text{ m} \times 12 \text{ m}$, as shown in Figure 5(a). We used a hyperboloid mirror with the surface equation of:

$$f_m(X) = \frac{x^2 + y^2}{548.1140} - \frac{z^2}{789.3274} + 1 = 0$$

in MCS, which was used by the Tribots team, as shown in Figure 2. The catadioptric camera was about 1m above the field, as per the rules of the MSL. The simulation results are shown in Figure 5(b), and the resolution of the omni directional image is 640×480 , which is a common resolution in current competition. In Figure 5(a), it can be seen that the three red marker points are used by the three-point method,

Volume 40 · Number 5 · 2013 · 462-473

Figure 5 (a) Illustration of the soccer field, three marker points, and camera position, (b) simulated omni directional image, (c) the simulated omni directional image, which is closer to the real image, (d) the result of the color segment, (e) the result of Canny edge detection and (f) the white line points for nonlinear optimization



and the catadioptric camera is placed on the black point. This configuration of the three points provided good results, which has been proven in our previous paper (Xiaoxiao and Qixin, 2012) – the further the three points are from the robot, the better the results will be.

4.1 The calibration result of our method

For the implementation, we first need to determine the outlines of the mirror and the central mark. In practice, the color of the background of the omni directional image and the central mark can be chosen appropriately to simplify finding of the outline, as shown in Figure 5(c). However, in the simulation system, we failed to change the background color, so we generated Figure 5(c) using Photoshop. A simple color segmentation method was used to obtain the results shown in Figure 5(d), and then, the Canny edge detection method was used to detect the outlines shown in Figure 5(e). We then used an ellipse-fitting algorithm (Andrew *et al.*, 1999) to compute the **Q** matrix. The image coordinates of the three marker points were obtained manually to make the experiment more like the practical calibration. The white line points used for the nonlinear optimization were also found by the simple color-segment method and are shown in Figure 5(f).

The calibration results of the internal and external parameters are summarized in Tables I and II, respectively. In Table I, the comparison of the two predicted image coordinates Mapping of non-SVP catadioptric camera of the soccer robot

Xiaoxiao Zhu and Qixin Cao

Industrial Robot: An International Journal

Volume 40 · Number 5 · 2013 · 462-473

Tab	le	L	Calibration	result	for	internal	parameters
-----	----	---	-------------	--------	-----	----------	------------

	<i>C</i> (mm)	п	Perspective center 1 (pixel)	Perspective center 2 (pixel)
Ground truth	(-0.7027, 0.8343, 99.9326)	(-0.0789, 0.0944, 0.9924)	(325.8, 231.9)	(325.8, 231.9)
Calibration result	(-0.787, 0.7785, 100.1089)	(-0.0797, 0.0928, 0.9925)	(325.8, 232)	(293.1, 270.6)
Absolute error	(0.0843, 0.0558, 0.1763)	(0.0008, 0.0016, 0.0001)	(0, 0.1)	(32.7, 38.7)

Table II Calibration result for external parameters

θ _x (°)	θ _y (°)	θ _z (°)	T _x (mm)	T _y (mm)	T _z (mm)
5.4163	4.5455	175	- 3,383.2334	3.2515	754.1425
5.6419	4.3868	175.0366	- 3,413.3504	6.3483	717.7123
0.2256	0.1587	0.0366	30.117	3.0968	36.4302
5.5920	4.4576	174.9355	- 3,342.1218	7.5052	736.117
0.1717	0.0911	0.0685	41.1115	4.2536	18.0255
	θ _x (°) 5.4163 5.6419 0.2256 5.5920 0.1717	θ _x (°) θ _y (°) 5.4163 4.5455 5.6419 4.3868 0.2256 0.1587 5.5920 4.4576 0.1717 0.0911	θ_x (°) θ_y (°) θ_z (°)5.41634.54551755.64194.3868175.03660.22560.15870.03665.59204.4576174.93550.17170.09110.0685	$\begin{array}{c c c c c c c c c c c c c c c c c c c $	θ_x (°) θ_y (°) θ_z (°) T_x (mm) T_y (mm)5.41634.5455175 $-3,383.2334$ 3.25155.64194.3868175.0366 $-3,413.3504$ 6.34830.22560.15870.036630.1173.09685.59204.4576174.9355 $-3,342.1218$ 7.50520.17170.09110.068541.11154.2536

of the mirror center with the true value indicates that the image coordinate of the mirror center is distinguishable for choosing the right pose. From the comparison of C and N with the ground truth, we can see that the accuracy of the mirror pose is very high because, theoretically, a one-pixel error in the ICS corresponds to 0.13 mm in the mirror base plane. Table II summarizes the calibration results of the mirror-to-world parameters, using three Euler angles θ_x , θ_y and θ_z to represent the rotation and T_x , T_v and T_z to represent the translation components. We can see that the three-point method has a small error in the rotation components but a slightly larger error in the translation components. Moreover, the nonlinear optimization reduces the error of most of those parameters, except those of θ_z and T_x . There are two reasons for these results; one is the error of the internal parameter and the other is the discretization error of the image, which make the ground truth unreachable by calibration.

To illustrate the result more intuitively, we built the distance mapping from the ICS to the WCS using equation (5) and then unwrapped the omni directional image. Additionally, some crossed lines were added to the soccer field to make the result clearer, as shown in Figure 6(a) and (b). Figure 7(a) shows the unwrapped image using the results of the threepoint method, and Figure 7(b) is drawn with the correct lines. In Figure 7(b), we can see that the unwrapped result has some error even in the place of the three calibration points, such as the center point of the middle circle (the black point in the zoomed part of Figure 7(b)), and this is because the image coordinates of the calibration points are manually selected with a random error. In other places on the field, there is also some error because those back-projected lines do not correspond to the correct lines very well. However, the back-projection result is still acceptable. Figure 7(c) and (d) are the results of the nonlinear optimization. From the comparison of Figure 7(b) and (d), it is clear that the back-projection error is reduced, and the maximum error is less than 300 mm in a $12 \text{ m} \times 12 \text{ m}$ area around the camera. The average back-projection errors of those white line points totalError/*n* are 112.3 and 63.8 mm before and after the nonlinear optimization, respectively.

Additionally, the calibration process is very efficient, and it takes less than 1 min.

4.2 Comparison with traditional method

As mentioned in the introduction, the traditional method utilizes the interpolating method to calibrate the distance mapping between the image coordinate and the world coordinate. Obtaining the initial mapping of several points is the core of the interpretation method. The common scene of the traditional calibration is shown in Figure 8, in which the blueand-white board is the calibration board and the edges between two colors are used as the calibration marks (with known distances to the robot used to build the initial mapping).





Figure 7 (a) Unwrapped result of omni directional image, (b) unwrapped result with correct cross lines drawn in, (c) unwrapped result of omni directional image using nonlinear optimization and (d) unwrapped result of nonlinear optimization with correct cross lines drawn in



(a)



(b)





Industrial Robot: An International Journal

Volume 40 · Number 5 · 2013 · 462-473

Figure 8 The calibration scene of the traditional calibration method



Notice that because the omni directional camera used by the soccer robot is non-SVP, which means that the distortion of the omni directional image usually is not symmetrical, all of the directions of the omni directional camera need to be calibrated, and in the practical calibration processing, the robot needs to rotate to obtain the initial correspondence in every direction. In our experiment, we made a perfect calibration board, as shown in Figure 9(a), which is not possible in practice. Using this calibration board, we do not need to rotate the robot and can obtain the initial correspondence of all the directions in one image. During implementation, we apply the code of the Tribots team (Tribots code, 2005) for interpolation, which uses linear interpolation. After the distance mapping is computed, we unwrap the omni directional image to show the accuracy of the distance mapping. Because the distances of the initial mark points to the robot are 500, 1,000, 2,000, 3,000, 4,000, 6,000 and 10,000 mm, we will unwrap the soccer field in the range of 10 m to the camera.

Two configurations of the catadioptric camera were used to compare our method with the traditional method. The first one is a severe misalignment configuration, which is the same configuration as used in the last section, having both a transformation and a rotation between the mirror and the camera. The second one is a slight misalignment configuration, which only has an offset along the axis causing a non-confocal condition between the mirror and camera. Because we cannot obtain the internal and external parameters explicitly by using the traditional method, we will use the unwrapped result and back-projection error as the comparison object.





The severe misalignment case

The unwrapped results of our method are shown in the last section as Figure 7. Figure 10(a) and (b) are the original and unwrapped omni directional images of the calibration board, respectively. Figure 10(b) reflects the precision of the initial distance mapping of those mark points. It can be seen that the back-projection error is large in the upper right part of the largest blue ring, which is not available in the severe misalignment configuration and is caused by the error that occurs when we obtain the image coordinates of the edge of a color change. Figure 10(c) shows the unwrapped result of the soccer field with cross lines (the original image is Figure 6(b)), in which we can see that the interpolation result is not very good. This is because the large misalignment between the mirror and the camera caused serious nonlinear distance mapping. Moreover, the further the distance to the camera, the larger the error is.

From the comparison of the results shown in Figures 7(d) and 10(b), it is easy to conclude that our method is more accurate than the traditional method for the non-SVP omni directional camera with large misalignment between the mirror and the camera.

The slight misalignment case

Figure 11(a) is the omni directional image of the soccer field with cross lines, and the red point on it is used to calculate the

Industrial Robot: An International Journal

Volume 40 · Number 5 · 2013 · 462-473

back-projection error. Figure 11(b) shows the omni directional image of the calibration board. It is easy to see that the distortion of the image is smaller than the distortion in Figure 10(a). After calibration, we unwrapped the omni directional image using the calibration results of our method, as shown in Figure 11(c), and of the traditional method, as shown in Figure 11(d). Although it is not as clear as in the severe misalignment case, the result of our method is still better, from the comparison of Figure 11(c) and (d). For a quantitative comparison, we use the average back-projection error of the red points in Figure 11(a). Those points are divided into several groups according to their distances to the catadioptric camera. Table III summarizes the comparison results. At every distance interval, the average error of our method is smaller than that of the traditional method. Moreover, at a distance interval of 6-10 m, the error of the traditional method is three times the error in a distance interval of 0.5-6 m, while the error using our method with the nonlinear optimization is only double. Additionally, at a distance farther than 10 m, our method still has good precision, while the error value is meaningless for the traditional method, because it is just calibrated in the 10m range. This means that our method is scalable to a larger soccer field.

Additionally, our method is more efficient, because in the practical calibration, the most time-consuming part is preparing the calibration board, which is not needed

Figure 10 (a) The omni directional image of the calibration board, (b) unwrapped result of calibration board and (c) unwrapped result of the soccer field with cross lines drawn in





(b)



(c)

471

Figure 11 (a) Simulated omni directional image of the soccer field, (b) omni directional image of the calibration board, (c) unwrapped result of (a) using the calibration result of our method and (d) unwrapped result of (a) using the calibration result of the traditional method



c)

Table III Comparison of the average back-projection errors

(d)

Distance nterval (m)	Average error of three-point method (mm)	Average error after nonlinear optimization (mm)	Average error of traditional method (mm)
0.5-6	48.1	28.9	74.8
6-10	114.5	61.2	256.6
10-14.7	296.3	94.1	Meaningless
0.5-10	74.5	41.8	147.2

in our method. Furthermore, the robot needs to rotate at least 360° in the traditional calibration, which also increases calibration time.

5. Conclusion

In this paper, we proposed a full model-based method for the distance-mapping calibration of a non-SVP catadioptric camera used by soccer robots. This method uses only one image of three feature points on the soccer field and does not need additional calibration objects. The procedure for this method consists of the internal parameter calibration, a three-point method and nonlinear optimization to obtain external parameters, and a back-projection process. The comparison with the traditional method shows that the weaknesses of our method are that the surface equation of the mirror must be known and the outline of the mirror base needs to be in the view of the camera (or at least part of the outline needs to be viewable). However, those two requirements usually can be satisfied for the soccer robot. On the other hand, the strengths of our method are that because we do not need additional calibration objects, it is easy to operate and is more efficient. At the same time, it is more accurate than the traditional method, especially when the catadioptric camera has a large misalignment configuration. Our method also maintains good precision when the target is far from the camera, so it can be used for larger soccer fields. Moreover, a possible future improvement is to develop an automatic calibration method based on our method to allow in-game calibration.

References

- Andrew, F., Maurizio, P. and Robert, B.F. (1999), "Direct least squares fitting of ellipses", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 21 No. 5, pp. 476-480.
- António, J.R.N., Armando, J.P., Daniel, A.M. and Bernardo, C. (2011), "An efficient omni directional vision system for soccer robots: from calibration to object detection", *Mechatronics*, Vol. 21 No. 2, pp. 399-410.
- Arun, K.S., Huang, T.S. and Blostein, S.D. (1987), "Leastsquares fitting of two 3-D point sets", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, No. 5, pp. 698-700.
- Baker, S. and Nayar, S.K. (1999), "A theory of singleviewpoint catadioptric image formation", *Proceeding of IJCV*, Vol. 32 No. 1, pp. 175-196.
- Bouguet (2008), "Matlab toolbox of camera calibration", available at: www.vision.caltech.edu/bouguetj/index.html (accessed April 26, 2012).
- Chen, Q., Wu, H.Y. and Wada, T. (2004), "Camera calibration with two arbitrary coplanar circles", *Proceeding* of ECCV'2004, pp. 521-532.
- Cunha, B., Azevedo, J., Lau, N. and Almeida, L. (2006), "Obtaining the inverse distance map from a non-SVP hyperbolic catadioptric robotic vision system", *RoboCup-2007: Robot Soccer World Cup XI*, LNAI, Springer, Berlin, pp. 417-427.
- Daniilids, K. and Geyer, C. (2001), "Catadioptric projective geometry", *International Journal of Computer Vision*, Vol. 45 No. 3, pp. 223-243.
- Gao, X., Hou, X., Tang, J. and Cheng, H. (2003), "Complete solution classification for the perspective-three-point problem", *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 25 No. 8, pp. 930-943.

- Haralick, R.M., Lee, C.N., Ottenberg, K. and Nolle, M. (1994), "Review and analysis of solutions of the three point perspective pose estimation problem", *International Journal of Computer Vision*, Vol. 13 No. 3, pp. 331-426.
- Hartley, R. (2003), Multiple View Geometry in Computer Vision, United Kingdom University Press, Edinburgh.
- Heinemann, P., Streichert, F., Sehnke, F. and Zell, A. (2006), "Automatic calibration of camera to world mapping in robocup using evolutionary algorithms", *Proceedings of the IEEE Congress on Evolutionary Computation (CEC 2006).*
- Kitano, H., Minoru, A., Yasuo, K.S., Itsuki, N. and Eiichi, O. (1997), "RoboCup: the robot world cup initiative", *Proceedings of the First International Conference on Autonomous Agents*, ACM Press, New York, NY, pp. 340-347.
- Lauer, M., Lange, S. and Riedmiller, M. (2005), Calculating the Perfect Match: An Efficient and Accurate Approach for Robot Self-localization, Lecture Notes in Computer Science, Vol. 4020, pp. 142-153.
- Mashita, T., Iaw, Y. and Yachida, M. (2005), "Calibration method for misaligned catadioptric camera", *Proceeding of OMNIVIS*'2005.
- Maxime, L. (2008), "Automatic sense structure and camera motion using a catadioptric system", *Computer Vision and Image Understanding*, Vol. 109 No. 2, pp. 186-203.
- Mičušík, B. and Pajdla, T. (2004), "Autocalibration & 3D reconstruction with non-central catadioptric cameras", *Proceedings of International Conference on Computer Vision and Pattern Recognition CVPR'04, Washington, DC, USA.*
- Nistér, D. and Stewénius, H. (2007), "A minimal solution to the generalised 3-point pose problem", *Journal of Mathematical Imaging and Vision*, Vol. 27, pp. 67-79 (Manufactured in The Netherlands).

- Volume 40 · Number 5 · 2013 · 462–473
- Strelow, D., Mishler, J., Koes, D. and Singh, S. (2001), "Precise omni directional camera calibration", *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 689-695.
- Tribots code (2005), available at: http://ml.informatik.unifreiburg.de/research/tribots/code
- Xiaoxiao, Z. and Qixin, C. (2012), "A special unique solution case of the perspective-three-point problem for external parameter calibration of an omni directional camera", *International Journal of Advanced Robotic Systems*, Vol. 9.
- Zhang, Z. (1999), "Flexible camera calibration by viewing a plane from unknown orientations", *ICCV*'99, pp. 666-667.
- Zhiyu, X., Bo, S. and Xing, D. (2012), "The camera itself as a calibration pattern: a novel self-calibration method for noncentral catadioptric cameras", *Sensors*, Vol. 12 No. 6, pp. 7299-7317.
- Zongjie, X., Wanmi, C. and Haiting, F. (2008), "Automatic calibration of camera to world mapping in omni directional vision system using error descent algorithm", paper presented at ICAI08, Multi-conference on Advanced Intelligence, Beijing, October.

Further reading

More, J. (1978), "The levenberg-marquardt algorithm, implementation and theory", *Numerical Analysis*, Lecture Notes in Mathematics, Vol. 630, Springer, Berlin, pp. 105-111.

Corresponding author

Xiaoxiao Zhu can be contacted at: zhuxiaoxiao_nuaa@ yahoo.com

Or visit our web site for further details: www.emeraldinsight.com/reprints