Jing LU[1,2] , Yuda SHI[2] , Shihong WU[3]

School of Optoelectronic Science and Engineering, Huazhong Univ. of Science and Technology (1), Dept. of Computer Science and Technology, Hubei Univ. of Education (2), The Command of PLA Guangdong Province Zhuhai Sub-Military Region (3)

Monocular Vision-based Sensor for Autonomous Mobile Robot Localization by Circular Markers

Abstract. *To navigate reliably in indoor environments, a mobile robot must know where it is. This paper is concerned with the design of a monocular vision-based algorithm for on-line estimation of a mobile robot's location using circular markers. The algorithm is based on 3-D analytic geometry, which is capable of estimating both the orientation and the position of the camera by a single camera image. The method can be used for camera-robot calibration for eye-on-hand systems, and autonomous mobile robot guidance. Laboratory experiments using a moving cylindrical object demonstrate both the accuracy and stability of the method.*

Streszczenie: *Robot mobilny, aby wiarygodnie nawigować we wnętrzu, musi znać swoje położenie. Opracowanie, w celu oszacowania bieżącego położenia robota, koncentruje się na projekcie algorytmu opartego o obraz jedno-okularowy, z zastosowaniem markerów kołowych. Algorytm przeprowadza analizę w geometrii 3D, co umożliwia oszacowanie zarówno orientacji jak i położenia kamery na podstawie danych obrazu z jednej kamery. Metoda może być zastosowana do kalibracji robota-kamery w systemie oko-ręka i do sterowania autonomicznym robotem mobilnym. Zademonstrowano eksperymenty laboratoryjne z wykorzystaniem obiektu cylindrycznego, analizując zarówno dokładność jak i stabilność metody.* ***Jedno-okularowy czujnik wizyjny do lokalizacji autonomicznego robota mobilnego z zastosowaniem markerów kołowych.***

Keywords: vision-based sensor, monocular vision, 3-D localization, circular marker.
Słowa kluczowe: Czujnik wizyjny **,** Widok jedno-okularowy, Lokalizacja 3D, Marker kołowy

Introduction

To navigate reliably in indoor environments, an autonomous mobile robot must know where it is, but reliable localization and navigation within highly unstructured environments is a difficult task. A widely adopted approach is the vision-based technique, and there have been proposed a variety of vision-based snesing methods such as using a stereo camera[1, 2], using a single camera[3], and using a laser light projection[4, 5]. However, the sensors using a stereo camera need two power sources for two cameras. The synchronization mechanisms of two cameras are also needed. Those using a laser also need two power sources for a camera and a laser. Additionally, the whole sizes of these sensors become large, which is inconvenient for the robot to carry with. Monocular vision can be used to find 3-D position and orientation of an object if the object model is known, and estimation of 3-D information from 2-D image coordinates by monocular vision is a fundamental problem in both machine vision and computer vision. This problem exists in two forms: the direct and the inverse[6]. In the direct type, the objective is to estimate the 3-D location of objects, landmarks, and features. This type of problem occurs in many areas: for example, in automatic assembly, tracking, and industrial metrology. In the inverse type, the objective is to estimate the relevant camera parameters: for a fixed camera, all the intrinsic and extrinsic parameters; for a moving camera, only the extrinsic parameters, namely, its 3-D location. This type of problem occurs in areas such as camera calibration, camera-robot calibration for eye-on-hand systems, and autonomous mobile robot guidance.

A circular shape is the most common quadratic that has been addressed for localization due to the following reasons: 1) many manufactured objects have circular holes or circular surface contours; 2) a circle's perspective projection in any arbitrary orientation is always an exact ellipse; 3) a circle has been shown to have the property of high image-location accuracy[7]; and 4) the complete boundary or an arc of a projected circle can be used for localization without knowing the exact point correspondence. There have been some previous work dealing with accurate localization of a mobile robot using circular landmarks[8], however few work focused on robot localization by circular markers on quadratic surfaces. In this paper, the fundamental algorithmic principles of the 3-D analytical geometry solutions for the localization problem are described, and laboratory experiments are done to testify the accuracy and stability of the algorithm. Sect. 2 constructed the geometrical imaging model of a circle, and introduced the 3-D analytical geometry solutions for localization of the circles; Experiment results and analysis for applying the method to localization with a cylindrical object are shown in Sect. 3; Sect. 4 is the conclusion.

Method for Localization

When put the image plane and a circular marker by the same side of the camera, camera imaging of the circle can be described as a cone model (Fig. 1), where vertex P is the center of the camera lens, and ellipse is the perspective projection of the circle onto the image plane. Note the circle as ξ, the 3-D localization problem can be defined as: given a perspective projection of ξ (a 2-D ellipse), and the focal length of the camera, estimate ξ's 3-D location, or in other words, estimate the coordinates of the circle's center C_1 and the unit surface-normal vector d_1.

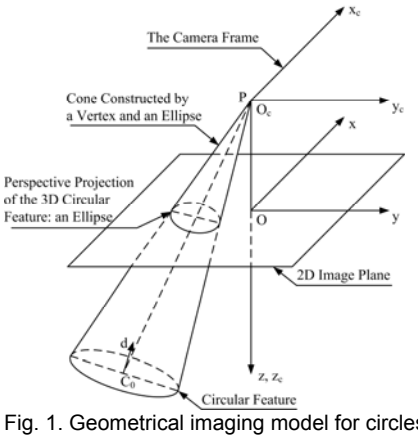

Fig. 1. Geometrical imaging model for circles: a cone

The cone equation can be derived from the image ellipse[9]. Note it as:

(1) $$ax^2 + by^2 + cz^2 + 2fyz + 2gzx + 2hxy + 2ux + 2vy + 2wz + d = 0.$$

3-D analytical geometry solutions are applied to solve the problem of location estimation of circles. The method is based on the idea of firstly estimating the location of the

circle in a frame where the cone equation is reduced to its central form, then determine the location with respect to the camera $x_cy_cz_c$ frame by transforming the frame to $x_cy_cz_c$. Let *XYZ* be the canonical frame[10] with Z be the principal axis of the cone, and *X′Y′Z′* be the frame with *Z′*-axis vertical to the circle surface by rotating *XYZ*, the total transformation from *X′Y′Z′* to $x_cy_cz_c$ can be written as

(2) $X'Y'Z' \xrightarrow{T_3} XYZ \xrightarrow{T_2} x'y'z' \xrightarrow{T_1} xyz \xrightarrow{T_0} x_cy_cz_c$

where $\mathbf{T}_0$, $\mathbf{T}_2$ are translational, and $\mathbf{T}_1$, $\mathbf{T}_3$ are rotational transformations. Thus the total transformation can be written as

$$(3)\qquad \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \mathbf{T}_0\mathbf{T}_1\mathbf{T}_2\mathbf{T}_3 \begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & e & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{T}_1\mathbf{T}_2\mathbf{T}_3 \begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix}$$

In the canonical (*XYZ*) frame, (1) is reduced to a more compact form

$$(4)\qquad K_1X^2 + K_2Y^2 + K_3Z^2 = \mu$$

It has been proven in [10] that K_i are the roots (with K_1, K_2 be positive, K_3 be negative) of the discriminating cubic equation

$$(5)\qquad K^3 - K^2(a+b+c) + K(bc+ca+ab-f^2-g^2-h^2) - (abc+2fgh-af^2-bg^2-ch^2) = 0$$

$\mathbf{T}_1$, $\mathbf{T}_2$ and the circle's unit surface-normal vector (*l*, *m*, *n*) with respect to *XYZ* frame are derived from K_i. Thus the unit surface-normal vector **d**: (l_0, m_0, n_0) with respect to the camera $x_cy_cz_c$ frame is obtained by rotating (*l*, *m*, *n*) as:

$$\begin{bmatrix} l_0 & m_0 & n_0 & 1 \end{bmatrix}^{\mathrm{T}} = \mathbf{T}_1 \begin{bmatrix} l & m & n & 1 \end{bmatrix}^{\mathrm{T}}$$

Rotate *XYZ* frame until *Z*-axis is vertical to the plane of the circle, and define the new frame as *X′Y′Z′*. The rotational transformation $\mathbf{T}_3$ is:

$$(6)\qquad \mathbf{T}_3 = \begin{bmatrix} -\frac{m}{\sqrt{l^2+m^2}} & -\frac{ln}{\sqrt{l^2+m^2}} & l & 0 \\ \frac{l}{\sqrt{l^2+m^2}} & -\frac{mn}{\sqrt{l^2+m^2}} & m & 0 \\ 0 & \sqrt{l^2+m^2} & n & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \equiv \begin{bmatrix} u_1 & v_1 & w_1 & 0 \\ u_2 & v_2 & w_2 & 0 \\ u_3 & v_3 & w_3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The center coordinates (X_0', Y_0', Z_0') with respect to *X′Y′Z′* are obtained as follows:

$$(7)\qquad \begin{bmatrix} X_0' & Y_0' & Z_0' \end{bmatrix} = \begin{bmatrix} -\frac{B}{A}Z_0' & -\frac{C}{A}Z_0' & \pm\frac{Ar}{\sqrt{B^2+C^2-AD}} \end{bmatrix}$$

where

$$A = K_1u_1^2 + K_2u_2^2 + K_3u_3^2,\ B = K_1u_1w_1 + K_2u_2w_2 + K_3u_3w_3,$$
$$C = K_1v_1w_1 + K_2v_2w_2 + K_3v_3w_3,\ D = K_1w_1^2 + K_2w_2^2 + K_3w_3^2$$

and *r* is the radius of the circle. Therefore, the center coordinates C_0: (x_{c0}, y_{c0}, z_{c0}) with respect to $x_cy_cz_c$ is obtained by substituting (7) into (3).

Experiment Results and Analysis

The experiment was accomplished in a laboratory environment. As shown in Fig. 2, a cylinder object marked with paralleled circles was put on the table. A camera took pictures of the cylinder object with an elevation angle of α. We changed the camera's 3-D orientation and position four times, and took four groups of images. The distance d between the camera lens and each of the circles are measured as reference values.

In our experiment, we took *d* as an indicator of the circle's location because the center is inside the cylinder, which makes it impossible to measure. This is reasonable since *d* is strictly determined by the circle's location.

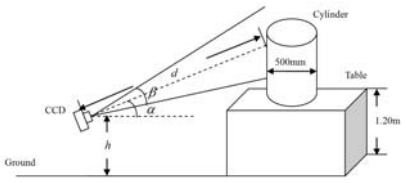

Fig. 2. Experiment setup

The experiment setup consisted of the following major components: a color camera: Nikon (model D80) with 23.6 × 15.8 mm image sensor, 3872 × 2592 resolution and 10.2 million effective pixels; a Nikon 135mm f/2D AF-DC Nikkor lens; an 500mm(R) × 1m(H) cylinder object marked with black and orange circles that paralleled to the cylinder underside; an 1.2m(W) × 0.9m(H) calibration board contained with 8×6 squares.

The experiment procedure is as follows:

1. Camera calibration. Construct the experiment setup as Fig. 2. The camera is calibrated by applying the flexible plane-based calibration technique[11]. As a result, the five intrinsic parameters of the camera (the coordinates of the principal point, the scale factors in image *u* and *v* axes, and the skewness of the two image axes) and the radial distortion coefficients are obtained.

2. Take a few images of the object. Then change the camera's 3-D orientation and position and return to step 1. Repeat step 1 and 2 four times and get four groups of images. One image is picked up from each group, shown as Fig. 3.

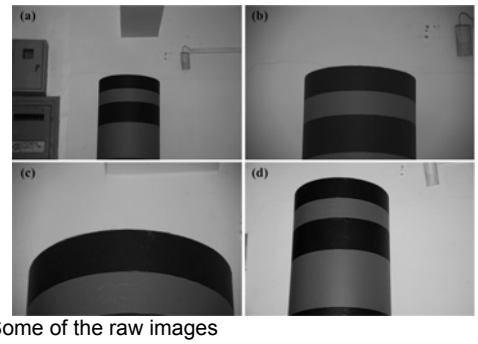

Fig. 3. Some of the raw images

3. Image preprocess and elliptical-shape edge detection. A subpixel edge detector based on the principal axis analysis and the moment-preserving principle[12] was applied to detect the elliptical-shape edge in the image, which resulted in a set of subpixel edge-point data. Edge detection results of Fig. 3(b) are shown as Fig. 4(a). Since only part of the elliptical shape is visible, direct least square fitting of ellipse[13] is done to estimate all the parameters of the actual image ellipse. As a result, three image ellipses detected from Fig. 4(a) are shown as Fig. 4, from which we can see the edge points and the ellipse equations fit well.

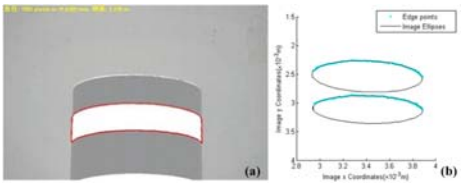

Fig. 4. Elliptical-shape edge detection results

4. Coordinate transformation and lens radial distortion compensation. Using the camera's intrinsic parameters and the radial distortion coefficients, a coordinate transformation between the image (*u*, *v*) coordinates that expressed in terms of pixel units, and the physical (*x*, *y*) coordinates that expressed in terms of absolute length is formed.

5. Apply the coordinate transformation to the ellipse equation, so that all the parameters in (1) are obtained.

6. Estimate the circles' orientation and position by the closed-form solution as described in Sect. 2.2. Use them to calculate the distance between the camera and the circles.

7. Compare the estimated values with reference values.

For each of the four images in Fig. 3, three circles were estimated. The application of the above procedure to the twelve circles resulted in two sets of data, as shown in Fig. 5 and Table Ⅰ. Note that in each group, the circles must have the same orientation angles since they were paralleled, we define the average orientation angles as the mean value of the orientation angles of the circles. Through camera calibration, the orientation angles of the surface normal of the circles were estimated as reference angles. The deviations are defined as the absolute value of the difference between the average angles and the reference angles. Note α, β, and γ as the angles that the surface normal of a circle makes with the x_c, y_c and z_c axis of the camera frame, respectively. As can be seen in Fig. 5, no matter what 3-D orientation the cylindrical object had, the deviations were relatively stable within the range between 0.44° and 1.38°, and the average deviations for the three orientation angles were determined as 1.22°, 0.75°, and 0.72°, respectively. The results show only a small error indicative of the good performance of the total process.

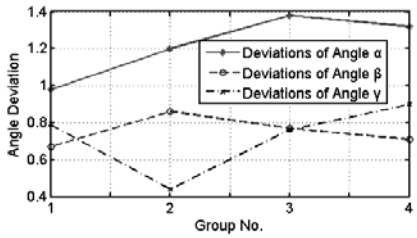

Fig. 5. Deviations of the estimated orientation

Table 1. Distance between the camera lens and each circle (measurement unit: m)

Circle No.	1	2	3	4	5	6
Estimated Distance	4.71	4.90	5.00	3.26	3.24	3.24
Reference Distance	4.78	4.85	4.91	3.19	3.31	3.38
Deviations	0.07	0.05	0.09	0.08	0.07	0.14
Circle No.	7	8	9	10	11	12
Estimated Distance	2.58	2.45	2.70	3.68	3.56	3.51
Reference Distance	2.47	2.54	2.60	3.53	3.62	3.60
Deviations	0.11	0.09	0.10	0.15	0.06	0.09
Average Deviation	0.09					

In Table 1, the location-estimation results of our method are presented. The measured distance of the twelve circles to the camera lens are given in the row "Reference Distance", while the distance calculated by our method are given in the row "Estimated Depth". The absolute value of the difference between the reference and estimated distance of all the circles are calculated and given as "Deviations", and the means of these values are given as "Average Deviation". As can be seen, the average deviation of the distance is only 0.09m, which is only 2.54% of the average reference distance. The results indicate the high accuracy of the total process.

Conclusions

Knowing the position and distance a mobile robot has moved is critical to effective operation, but reliable localization and navigation within highly unstructured environments is a difficult task. In this paper we applied the monocular model-based vision to estimate the 3-D location of circles. Advantages of the monocular model-based method are: (1) small system size and simple system structure; (2) closed-form solution using simple mathematics; (3) high accuracy; (4) handles all kinds of objects with circles, such as objects with circular surface contours, objects that have holes, objects with circular markers, etc. Laboratory experiments demonstrate that the method is capable of estimating both orientation and position of a cylindrical object with radius of 500mm within 5m away from the camera lens. The high accuracy and stability of the method are also demonstrated by moving and relocating the cylindrical object.

Acknowledgments

This work was supported by Program for Innovative Research Team in Hubei University of Education, as well as the No.2011A001 Grant project of Hubei University of Education.

REFERENCES

[1] P. Corke, et al., "Experiments with underwater robot localization and tracking," in Proceedings of the 2007 Ieee International Conference on Robotics and Automation, Vols 1-10, ed New York: Ieee, 2007, 4556-4561.

[2] R. Cipolla and N. Hollinghurst, "Visually guided grasping in unstructured environments," Robotics and Autonomous Systems, vol. 19(1997), no. 3-4, 337-346.

[3] A. Yamashita, Y. Shirane, and T. Kaneko, "Monocular Underwater Stereo-3D Measurement Using Difference of Appearance Depending on Optical Paths," in Ieee/Rsj 2010 International Conference on Intelligent Robots and Systems, ed New York: Ieee, 2010, 3652-3657.

[4] K. D. Moore, "Intercalibration method for underwater three-dimensional mapping laser line scan systems," Appl. Opt., vol. 44(2001), no. 33, 5991-6004.

[5] C.-C. Wang and M. S. Cheng, "Nonmetric camera calibration for underwater laser scanning system," IEEE J. Oceanic Eng., vol. 32(2007), no. 2, 383-399.

[6] R. Safaee-Rad, K. C. Smith, and B. Benhabib, "Three-dimensional location estimation of circular features for machine vision," IEEE Trans. Robotic Autom., vol. 8(1992), no. 5, 624-639.

[7] R. A. Young, "Locating industrial parts with subpixel accuracies," presented at the SPIE Proc. Optics, Illumination, Image Sensing Machine Vision, vol. 728(1986), no.

[8] C. J. Wu and W. H. Tsai, "Location estimation for indoor autonomous vehicle navigation by omni-directional vision using circular landmarks on ceilings," Robotics and Autonomous Systems, vol. 57(2009), no. 5, 546-55.

[9] Y. C. Shiu and S. Ahmad, "3D Location of Circular and Spherical Features by Monocular Model-Based Vision," presented at the Proceedings of the 1989 IEEE. Systems, Man and Cybernetics, vol. 1989), no. Cambridge.

[10] R. J. T. Bell, An Elementary Treatise on Coordrnated Geometry of Three Drmensrons, 3rd ed.: London: Macmillan, 1944.

[11] Z. Zhang, "A Flexible New Technique for Camera Calibration," IEEE Trans. Pattern Anal., vol. 22(2000), no. 11, 1330-1334.

[12] S. C. Cheng and T. L. Wu, "Subpixel edge detection of color images by principal axis analysis and moment-preserving principle," Pattern Recogn., vol. 38(2005), no. 4, 527-537.

[13] A. Fitzgibbon, M. Pilu, and R. B. Fisher, "Direct least square fitting of ellipses," IEEE Trans. Pattern Anal., vol. 21(1999), no. 5, 476-490.

***Authors**: Dr. Jing Lu is now in Wuhan National Laboratory for Optoelectronics, School of Optoelectronic Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China, E-mail: lujing8102@126.com.*