Vol.31, No.3 ACTA AUTOMATICA SINICA May, 2005

Study on Visual Positioning Based on Homography for Indoor Mobile Robot[1)]

XU De TU Zhi-Guo TAN Min

(*Laboratory of Complex Systems and Intelligence Science, Institute of Automation, Chinese Academy of Sciences, Beijing* 100080)

(E-mail: xude@compsys.ia.ac.cn)

Abstract Monocular visual positioning for indoor mobile robot is concerned in this paper. A new visual positioning method based on homography matrix in Euclidean space is proposed. It can calculate the position and pose of the mobile robot according to the intrinsic parameters of camera and two position-known points in a plane. It is very simple and low cost in computation. The experimental results show its effectiveness.

Key words Mobile robot, visual positioning, homography matrix

1 Introduction

The positioning methods for mobile robots could be classified into two categories, relative and absolute. Relative positioning is also known as dead reckoning (DR), which mainly includes odometry and inertial navigation (IN) positioning. Absolute positioning mainly includes magnetic compasses (MC), active beacons (AB), global positioning systems (GPS), landmark navigation (LN), model matching (MM) positioning and so on[1]. Vision plays an important role in some methods above. For example, both LN and MM employ it to realize positioning. LN realizes positioning through landmark recognition with vision. MM extracts the environment features with vision and locates the robot by comparing them with the pre-known environment map[1]. MM has high positioning precision in global range, but it has disadvantages such as large calculation cost, worse capability in real time and high probability of incorrect feature matching. All of these restrict its application. LN technique is easy to be realized, but landmark searching is inevitable when it is used alone. It slows down the positioning process. So LN combined with DR is usually a good selection to reduce the landmark searching time and to correct accumulated errors of relative positioning. For example, El-Hakim *et al.*[2,3] gave a positioning method using LN combined with DR. It uses DR to determine the approximate initial position of the mobile robot. And it employs 8 CCD cameras on the robot to obtain the accurate position and pose of the robot, and the accurate position of the measurement point in the landmark. The errors of DR positioning are corrected at the same time. El-Hakin′s method has high positioning precision for slow moving robot. The visual positioning method must be improved in the respects of time cost and computational complexity especially for satisfying the demand of quick moving robot.

Recently Malis *et al.* presented 2.5D visual servo method based on homography matrix. The method was discussed, and its validity and real time capability were verified in their series papers[4∼10]. Fang *et al* applied 2.5D visual servo method on industrial robots that have redundant kinematics[11]. They also realized visual servo for mobile robot with a single camera. The homography matrix between the camera and object was calculated as the control variable according to the image coordinates of three points at least on a plane[12]. In fact, the transform from the object to the camera described as homography matrix can be taken as the positioning of the camera in the object′s coordinates. A new visual positioning method for mobile robot with a single camera is discussed in this paper, which is developed based on [4∼12].

2 Visual positioning based on homography matrix

2.1 Homography matrix in Euclidean space

The image homogeneous coordinates of a point P on a plane Π can be denoted as $\boldsymbol{I}_1 = [u_1, v_1, 1]^{\mathrm{T}}$ and $\boldsymbol{I}_2 = [u_2, v_2, 1]^{\mathrm{T}}$ in two different coordinates O_1 and O_2. It is well-known that there exists a

1) Supported by the National High Technology Research and Development Program of P. R. China (2002AA422160)
Received August 26, 2003; in revised form May 28, 2004

homography matrix H_i satisfying the following equation.

$$\alpha \boldsymbol{I}_2 = H_i \boldsymbol{I}_1 \tag{1}$$

where α is a nonzero constant.

The homography matrix H_i in the image coordinates is unique if the constant factor is not considered.

In the camera coordinates, axis Z is defined as the direction of optical axis of the camera. The optical center of the camera is selected as the origin. The coordinates of point P on plane Π in the camera coordinates are denoted as $\boldsymbol{P}_c = [x_c, y_c, z_c]^{\mathrm{T}}$. The camera′s intrinsic parameter matrix is indicated by T_c.

$$I_i = T_c P_{ci}/z_{ci} \tag{2}$$

From (1) and (2),

$$\boldsymbol{P}_{c2} = T_c^{-1} H_i T_c \boldsymbol{P}_{c1} z_{c2}/z_{c1} = H_e \boldsymbol{P}_{c1} \tag{3}$$

where H_e is called homography matrix in Euclidean space, $H_e = R + \dfrac{p}{\|\boldsymbol{P}_{c1}\|^2 \boldsymbol{P}_{c1}^2} \boldsymbol{P}_{c1}^{\mathrm{T}}$. R is the rotation transformation from coordinates O_1 to O_2. $\boldsymbol{p}$ is the position vector of the origin of coordinates O_1 in coordinates O_2.

In fact, H_e is the transformation from the coordinates O_1 to O_2, with different positions of the camera. The parameters in H_e can be determined according to three points on a plane Π[12].

2.2 Visual positioning

A plane vertical to the floor is selected as the plane Π, and a point on Π as the origin of coordinates O_1. The floor is taken as reference. Axes X_1 and Y_1 selected on Π are orthogonal. Axis X_1 is vertical to the floor, and Y_1 parallel to it. For coordinates O_2, the optical center of the camera is selected as its origin, and its axis X_2 is parallel to X_1. Axis Z_2 is defined as the direction of optical axis. The coordinates O_1 and O_2 are shown in Fig. 1. The camera is mounted on a mobile robot.

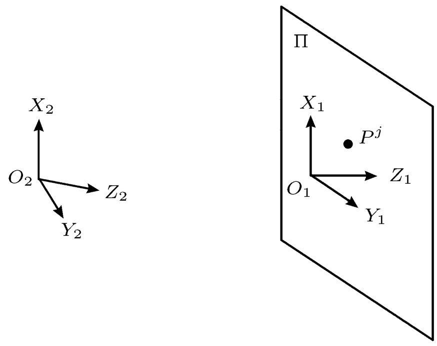

Fig. 1 Camera coordinates

When the mobile robot moves, the origin of coordinates O_2 translates relative to that of O_1 while the coordinates rotate around the axis X_2. So, from (3) we have

$$\boldsymbol{P}_{c2} = R\boldsymbol{P}_{c1} + \boldsymbol{p} \tag{4}$$

$$\begin{bmatrix} x_{c2} \\ y_{c2} \\ z_{c2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x_{c1} \\ y_{c1} \\ z_{c1} \end{bmatrix} + \begin{bmatrix} p_x \\ p_y \\ p_z \end{bmatrix} \tag{5}$$

where θ is the angle between axes Z_1 and Z_2. $[p_x, p_y, p_z]^{\mathrm{T}}$ is the coordinates of position vector $\boldsymbol{p}$.

Considering an arbitrary point P^j on plane Π, denoted as $\boldsymbol{P}_{c1}^j = [x_{c1}^j \quad y_{c1}^j \quad 0]^{\mathrm{T}}$ in the coordinates O_1, we have formula (6) form (5).

$$\begin{cases} x_{c2}^j = x_{c1}^j + p_x \\ y_{c2}^j = y_{c1}^j \cos\theta + p_y \\ z_{c2}^j = y_{c1}^j \sin\theta + p_z \end{cases} \tag{6}$$

where $[x_{c1}^j, y_{c1}^j, 0]$ and $[x_{c2}^j, y_{c2}^j, z_{c2}^j]$ are the coordinates of P^j in the frame O_1 and O_2, respectively.

(6) can be rewritten as (7), in which x_{c2}^j/z_{c2}^j and y_{c2}^j/z_{c2}^j can be obtained from (2) according to its image coordinates.

$$\begin{cases} \frac{x_{c2}^j}{z_{c2}^j}(y_{c1}^j \sin\theta + p_z) = x_{c1}^j + p_x \\ \frac{y_{c2}^j}{z_{c2}^j}(y_{c1}^j \sin\theta + p_z) = y_{c1}^j \cos\theta + p_y \end{cases} \tag{7}$$

Proposition 1. A camera is mounted on the mobile robot, whose optical axis is parallel to the floor. Suppose the intrinsic parameters of the camera are known, and plane Π is parallel to axis X or Y in the camera coordinates. If the coordinates of two arbitrary points on plane Π are pre-known, the pose and position of the mobile robot can be calculated according to their image coordinates.

Proof. Formula (7) is formed by one point′s coordinates on plane Π. It has two individual equations with four unknown parameters. Obviously, if there are two pre-known points on plane Π we can have four equations with four unknowns. Then the parameters p_x, p_y, p_z and θ can be obtained. The pose and position of the mobile robot represented as the coordinates O_2 in the coordinates O_1 could be known. Thus the mobile robot positioning is realized.

The method to calculate the pose and position of the robot from the image coordinates of two pre-known points on plane Π is discussed below in detail. If the origin of coordinates O_1 is selected as point P^1, then $x_{c1}^1 = 0, y_{c1}^1 = 0$. So,

$$\begin{cases} p_x = \frac{x_{c2}^1}{x_{c2}^1} p_z \\ p_y = \frac{y_{c2}^1}{z_{c2}^1} p_z \end{cases} \tag{8}$$

Submitting formula (8) to (7), we have formula (9) for point P_2,

$$\begin{cases} y_{c1}^2 \sin\theta = \frac{1}{x_{c2}^2/z_{c2}^2}[x_{c1}^2 + (\frac{x_{c2}^1}{z_{c2}^1} - \frac{x_{c2}^2}{z_{c2}^2})p_z] = a_1 + b_1 p_z \\ y_{c1}^2 \cos\theta = \frac{y_{c2}^2/z_{c2}^2}{x_{c2}^2/z_{c2}^2}[x_{c1}^2 + (\frac{x_{c2}^1}{z_{c2}^1} - \frac{x_{c2}^2}{z_{c2}^2})p_z] + \frac{y_{c2}^2}{z_{c2}^2}p_z - \frac{y_{c2}^1}{z_{c2}^1}p_z = a_2 + b_2 p_z \end{cases} \tag{9}$$

where $\begin{cases} a_1 = \frac{1}{x_{c2}^2/z_{c2}^2}x_{c1}^2, \quad b_1 = \frac{1}{x_{c2}^2/z_{c2}^2}(\frac{x_{c2}^1}{z_{c2}^1} - \frac{x_{c2}^2}{z_{c2}^2}) \\ a_2 = \frac{y_{c2}^2/z_{c2}^2}{x_{c2}^2/z_{c2}^2}x_{c1}^2, \quad b_2 = \frac{y_{c2}^2/z_{c2}^2}{x_{c2}^2/z_{c2}^2}(\frac{x_{c2}^1}{z_{c2}^1} - \frac{x_{c2}^2}{z_{c2}^2}) + \frac{y_{c2}^2}{z_{c2}^2} - \frac{y_{c2}^1}{z_{c2}^1} \end{cases}$.

(10) is obtained through making square sum of two equations in (9). Then the solution of p_z can be given as expression (11):

$$(y_{c1}^2)^2 = (a_1 + b_1 p_z)^2 + (a_2 + b_2 p_z)^2 \tag{10}$$

$$p_z = (-c_2 \pm \sqrt{c_2^2 - 4c_1 c_3})/2c_1 \tag{11}$$

where $c_1 = (b_1^2 + b_2^2),\ c_2 = 2(a_1 b_1 + a_2 b_2),\ c_3 = a_1^2 + a_2^2 - (y_{c1}^2)^2$.

p_x and p_y can be resolved according to (8) once p_z is known. And then the angle θ can be obtained from (12):

$$\theta = \arctan(a_1 + b_1 p_z, a_2 + b_2 p_z) \tag{12}$$

We can obtain a group of parameters p_x, p_y, p_z and θ from a pair of points on the plane Π. Many groups of parameters obtained from pairs of points can be fused to a group of accurate parameters[13].

3 Experiment and results

The camera intrinsic parameter, T_c, is calibrated firstly.

$$T_c = \begin{bmatrix} 2030.0 & 0 & 809.6 \\ 0 & 2083.4 & 279.3 \\ 0 & 0 & 1 \end{bmatrix}$$

The grid landmark is pasted on the wall. The position and orientation of the camera are adjusted so that the grid sides of the landmark image are parallel to axis X_1 or Y_1 separately when the camera is facing to the landmark directly. Fig. 2 shows the images captured by the camera when the mobile robot is at different positions. The nine intersection points are selected as the feature points. In the coordinates O_1, they are denoted as

$$\boldsymbol{P}_{c1} = \begin{bmatrix} 0 & 0 & 0 & 51 & 51 & 51 & 101.5 & 101.5 & 101.5 \\ 0 & 57 & 113.5 & 0 & 57 & 113.5 & 0 & 57 & 113.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

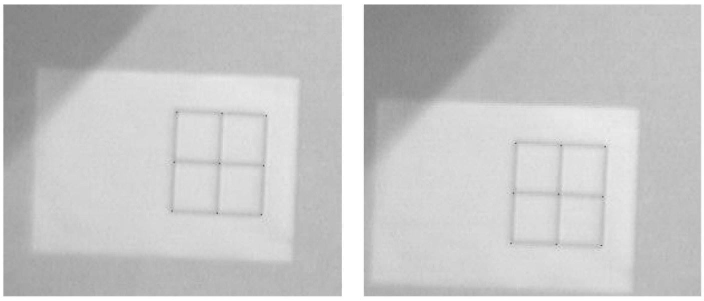

Fig. 2 Images captured for positioning

Eight pairs of points are formed from eight nonzero points with the origin of coordinates O_1. Then eight groups of positioning results can be obtained from them. When the mobile robot is at the initial position, the positioning results are

$$\theta = [-0.1686 \quad -0.1706 \quad -0.0260 \quad -0.1485 \quad -0.1468 \quad -0.0298 \quad -0.1893 \quad -0.2304]$$

$$P = \begin{bmatrix} -808.0 & -816.7 & -784.5+5.4i & -791.2 & -799.5 & -790.9+7.7i & -790.1+36.3i & -781.7 \\ 52.1 & 52.7 & 50.6-0.3i & 51.0 & 51.6 & 51.0-0.5i & 51.0-2.3i & 50.4 \\ 2805.7 & 2835.9 & 2724.0-18.8i & 2747.4 & 2776.2 & 2746.5-26.9i & 2743.6-126.2i & 2714.5 \end{bmatrix}$$

The positioning results after fusion are $\theta = -0.1587$, $\boldsymbol{p} = [-804\ 52\ 2791]^{\mathrm{T}}$.

After the robot moves 100mm in the direction of axis Y_2, the positioning results are $\theta = -0.1652$, $\boldsymbol{p} = [-796\ \ -48\ 2781]^{\mathrm{T}}$.

After the robot moves 50mm from the initial position along axis Y_2, and moves 50mm along axis Z_2, the positioning results are $\theta = -0.1669$, $\boldsymbol{p} = [-788\ 1\ 2731]^{\mathrm{T}}$.

The homogeneous transformation matrix T from coordinates O_1 to O_2 can be obtained using parameters p and θ above. $T = \begin{bmatrix} R & p \\ 0 & 1 \end{bmatrix}$. And T^{-1} represents the position and orientation of coordinates O_2 established on the robot in the landmark coordinates O_1. The results shown in Table 1 include the yawing angle and landmark position in coordinates O_2, and the position of the camera in coordinates O_1. The mean of three measurement results of yawing angle is considered as its desired value. The desired values of position are computed from the desired yawing angle and movement values.

Table 1 Experiment results of visual positioning

Test positions of the robot	Yawing angle θ (rad)		The position of O_1 in O_2 (mm)		The position of O_2 in O_1 (mm)	
	Measured	Desired	Measured	Desired	Measured	Desired
Initial	-0.1587		$[-804\ 52\ 2791]^{\mathrm{T}}$		$[804\ 390\ -2764]^{\mathrm{T}}$	
100mm moved along axis Y_2	-0.1652	-0.1636	$[-796\ -48\ 2781]^{\mathrm{T}}$	$[-804\ -48\ 2791]^{\mathrm{T}}$	$[796\ 505\ -2735]^{\mathrm{T}}$	$[804\ 502\ -2746]^{\mathrm{T}}$
50mm moved along axis Y_2, Z_2 separately	-0.1669	-0.1636	$[-788\ 1\ 2731]^{\mathrm{T}}$	$[-804\ 2\ 2741]^{\mathrm{T}}$	$[788\ 453\ -2693]^{\mathrm{T}}$	$[-804\ 445\ -2705]^{\mathrm{T}}$

Error analysis. The errors in measurement results are caused by many factors, including axes X_2 and X_1 are not parallel, the calibrated intrinsic parameters of the camera have errors, the results of image processing have errors, and so on. Especially, formulas (4∼12) are based on that axes X_2 and X_1 are parallel. If there is apparent error in the directions of axes X_2 and X_1, the positioning results will have obvious errors. It is easy to find that the grid sides of the landmark image are not parallel to the sides of picture. It is the main reason to cause the errors in measurement results.

Based on the initial position and neglecting the positioning error in the direction of axis X, the absolute and relative errors in visual measurement results are listed in Table 2. The absolute errors of yawing angle are the absolute difference between measured and desired values. Its relative errors are from the absolute errors divided by the absolute value of desired angle. The absolute errors in position are formed with the absolute difference between measured and desired position along axes Y and Z respectively. They are divided by the absolute coordinates in axes Y and Z individually, in which the larger result is taken as relative error.

Table 2 The errors of visual positioning based on initial position

Test positions of the robot	Yawing angle errors		Position errors of O_1 in O_2		Position errors of O_2 in O_1	
	Absolute	Relative	Absolute	Relative	Absolute	Relative
100mm moved along axis Y_2	0.0016rad	0.98%	[0,10]mm	0.29%	[3,11]mm	0.60%
50mm moved along axis Y_2, Z_2 separately	0.0033rad	2.02%	[1,10]mm	0.32%	[8,12]mm	1.80%

It can be found from Table 2 that the position errors of O_2 in O_1 are much larger than those of O_1 in O_2. The yawing angle errors have influence on transformation so that the position errors are enlarged. Therefore, the improvement of measuring precision of the yawing angle is the key of reducing the positioning errors. How to improve the precision of angle measurement is our work in future.

In addition, the position of the origin of coordinates O_2 in coordinates O_1 represents the position of optical center of the camera relative to the grid landmark. The robot coordinates can be established at a fixed point on the robot. The extrinsic parameters matrix T_m, the camera relative to the robot coordinates, could be calibrated. Then the position and pose of the landmark in the robot coordinates could be expressed as T_mT. And the position and pose of the robot in the landmark coordinates could be represented with $T^{-1}T_m^{-1}$.

4 Conclusion

A new visual positioning method for indoor mobile robot is proposed. It can calculate the position and pose of the landmark by use of the intrinsic parameters of a camera and two position-known points in a plane landmark, which is parallel to axis X or Y in the camera coordinates. The position and pose of the landmark in the robot coordinates can be computed with the extrinsic parameters of the camera. Thus the position and pose of the robot in the landmark coordinates could be obtained. The method is very simple and low cost in computation. The experimental results show its effectiveness.

References

1 Borenstein J, Everett H R, Feng L, Wehe D. Mobile robot positioning: Sensors and techniques. *Journal of Robotic Systems*, 1997, **14**(4): 231∼249

2 El-Hakim S F, Boulanger P, Blais F, Beraldin J-A, Roth G. A mobile system for indoors 3-D mapping and positioning. In: Proceedings of Optical 3-D Measurement Techniques IV. Zurich: Wichmann Press, 1997. IV: 275∼282

3 El-Hakim S F, Boulanger P, Blais F, Beraldin J-A. Sensor based creation of indoor virtual environment models. In: Proceedings of the 1997 International Conference on Virtual Systems and Multimedia. Geneva: IEEE Computer Society Press, 1997. 50∼58

4 Malis E, Chaumette F. 2 1/2 D visual servoing with respect to unknown objects through a new estimation scheme of camera displacement. *International Journal of Computer Vision*, 2000, **37**(1): 79∼97

5 Chaumette F, Malis E. 2 1/2 D visual servoing: a possible solution to improve image-based and position-based visual servoings. In: Proceedings of IEEE International Conference on Robotics and Automation. San Francisco: IEEE Press. 2000. **1**: 630∼635

6 Malis E, Chaumette F, Boudet S. 2D 1/2 visual servoing. *IEEE Transactions on Robotics and Automation*, 1999, **15**(2): 234~246

7 Malis E, Chaumette F, Boudet S. 2D 1/2 visual servoing stability analysis with respect to camera calibration errors. In: Proceedings of IEEE/RSJ International Conference on Intelligent Robots Systems. Victoria: IEEE Press, 1998. **2**: 691~697

8 Malis E, Chaumette F, Boudet S. Positioning a coarse-calibrated camera with respect to an unknown object by 2D 1/2 visual servoing. In: Proceedings of IEEE International Conference on Robotics and Automation. Leuven: IEEE Press, 1998. **2**: 1352~1359

9 Malis E, Chaumette F, Boudet S. Positioning a coarse-calibrated camera with respect to an unknown planar object by 2D 1/2 visual servoing. In: Proceedings of IFAC Symposium on Robot Control. Nantes: IFAC Press, 1997. **2**: 517~523

10 Chaumette F, Malis E, Boudet S. 2D 1/2 visual servoing with respect to a planar object. In: Proceedings of Workshop on New Trends in Image-based Robot Servoing. Grenoble: IEEE Press, 1997. 45~52

11 Fang Y, Behal A, Dixon W E, Dawson D M. Adaptive 2.5D Visual Servoing of Kinematically Redundant Robot Manipulators. In: Proceedings of the 41st IEEE Conference on Decision and Control. Las Vegas: IEEE Press, 2002. 2860~2865

12 Fang Y, Dawson D M, Dixon W E, De Queiroz M S. Homography-Based Visual Servoing of Wheeled Mobile Robots. In: Proceedings of the 41st IEEE Conference on Decision and Control, Las Vegas: IEEE Press, 2002. 2866~2871

13 Xu De, Tan Min. Accurate positioning in real time for mobile robot. *Acta Automatica Sinica*, 2003, **29**(5): 716~725

XU De Received his bachelor and master degrees from Shandong University of Technology in 1985 and 1990, respectively. He received his Ph. D. degree from Zhejiang University in 2001. He is an associate professor with Institute of Automation, Chinese Academy of Sciences (CASIA). His research interests include robot control, visual measurement and control.

TU Zhi-Guo Received his bachelor degree from Central South University in 2000. He is a postgraduate with CASIA. His research interests include robot control and remote control.

TAN Min Received his bachelor degree from Tsinghua University in 1986 and his Ph. D. degree from CASIA in 1990. He is a professor with CASIA. His research interests include robotics and complex system theory.