Study on the position and orientation measurement method with monocular vision system

Peng Wang (王 鹏)*, Xu Xiao (肖 旭), Zimiao Zhang (张子淼), and Changku Sun (孙长库)

State Key Laboratory of Precision Measuring Technology and Instruments,

Tianjin University, Tianjin 300072, China

**E-mail: wangpeng810@gmail.com*

Received February 27, 2009

In order to estimate the position and orientation of an object with a single camera, a novel measurement method based on pinhole camera model with five reference points is presented. Taking the specially designed planar target with the monocular vision system, the projection line of the reference points is built. According to the projection model, the coordinates of the reference points in the camera coordinate system are estimated with the least-squares algorithm. Thus the position and orientation of the target are worked out. Experimental result shows that the measurement precision of angle is less than 0.2°, and that of displacement is less than 0.1 mm.

OCIS codes: 150.1135, 330.4060, 150.0155.

doi: 10.3788/COL20100801.0055.

To estimate the position and orientation between two objects is an important technique, which can be widely applied in the fields as robot navigation[1], surgery[2], human motion estimate[3], and electro-optic aiming system[4,5], etc. Compared with the magnetic position and orientation measurement method[6,7], the machine vision method which is free from the influence of the electromagnetic surroundings, is widely studied[8−10]. But the method discussed in Refs. [8] and [10] uses two or more cameras to realize the position and orientation measurement. And in Ref. [9], the measurment is realized with the assitance of a laser range-finder. In this letter, a novel effective algorithm is proposed to recover the three-dimentional (3D) position and orientation of a planar target with a single camera view. A target pattern with five reference points is specially designed to ensure that the target special coordinate system can be fixed in any orientation. With the camera perspective projection model, the coordinate of the reference points in the camera coordinate system can be estimated by considering the relationship of them on the target pattern. The transformation matrix between the camera coordinate system and target coordinate system, which is built by the reference point coordinates, is worked out to realize the position and orientation measurement.

The planar target pattern with five reference points is designed for position and orientation estimation, as shown in Fig. 1. The colinearity of the vector, composed of every two reference points, is used to judge whether the three points are in a straight line. The two points out of the straight line are marked as No. 0, which is closer to the line, and No. 1 (see Fig. 2). The angles between the vector composed of No. 0 and No. 1 points and the vectors composed of No. 0 and other points are calculated. According to the angle size, the other reference points are marked as No. 2, No. 3, and No. 4. The origin of the target coordinate system is the No. 0 point. The $o_\mathrm{t}x_\mathrm{t}$ axis is parallel to the vector composed of No. 2 and No. 4 points. The $o_\mathrm{t}y_\mathrm{t}$ axis coincides with the vector composed of No. 0 and No. 1 points.

The measurement method is based on the pinhole camera model (Fig. 3). Here o_c denotes the optical center of the camera lens. $o_\mathrm{c}-x_\mathrm{c}y_\mathrm{c}z_\mathrm{c}$ is the camera coordinate system, which denotes the camera frame. uv is the CCD image plane with the original point at the image center of the charge coupled device (CCD) plane. $o_\mathrm{t}-x_\mathrm{t}y_\mathrm{t}z_\mathrm{t}$ is the target coordinate system, built by the five reference points P_i (i = 0, 1, $\cdots$, 4) as mentioned above. The distance from o_c to image plane is f, which is obtained from camera calibration algorithms[11,12]. On account of the lens radial distortion[13], the undistorted coordinate point on the image plane is $I_i = (u_i, v_i)^\mathrm{T}$ (i = 0, 1, $\cdots$, 4), which are different from the actual image reference points obtained by image processing.

Considering the pinhole camera model, the coordinates of the undistorted image reference point are $I_{\mathrm{c}i}=(u_i, v_i, f)^\mathrm{T}$ (i = 0, 1, $\cdots$, 4) in the camera coordinate system, in which the projecting line of the reference point P_i (i = 0, 1, $\cdots$, 4) is

$$\frac{x}{u_i} = \frac{y}{v_i} = \frac{z}{f} = t_i. \tag{1}$$

The coordinates of the reference points in the camera coordinate system are

$$P_{\mathrm{c}i} = \begin{pmatrix} x_{\mathrm{c}i} \\ y_{\mathrm{c}i} \\ z_{\mathrm{c}i} \end{pmatrix} = \begin{pmatrix} t_i u_i \\ t_i v_i \\ t_i f \end{pmatrix}, \tag{2}$$

where t_i is the scale factor which defines the position of P_i.

For the relative position of the reference points P_i in the camera coordinate system and the target coordinate system, t_i can be estimated by solving

$$\begin{aligned} &\|P_{\mathrm{c}i}P_{\mathrm{c}j}\| = \|P_{\mathrm{t}i}P_{\mathrm{t}j}\|, \;\; i = 0,\ 1,\ 2,\ 3,\ j = i + 1, \\ &\angle P_{\mathrm{c}i}P_{\mathrm{c}j}P_{\mathrm{c}k} = \angle P_{\mathrm{t}i}P_{\mathrm{t}j}P_{\mathrm{t}k}, \\ &i = 0,\ 1,\ 2,\ j = i + 1,\ k = j + 1, \end{aligned} \tag{3}$$

1671-7694/2010/010055-04

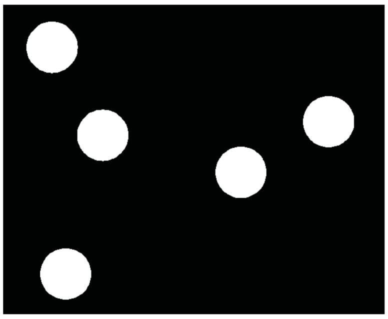

Fig. 1. Planar pattern for the position and orientation measurement.

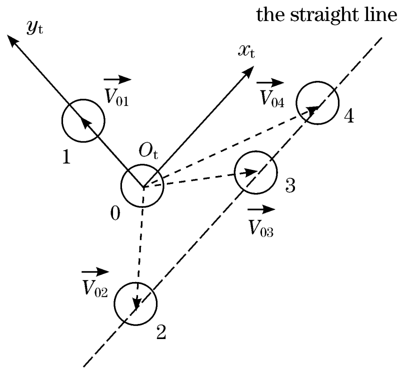

Fig. 2. Topological relationship of the five reference points.

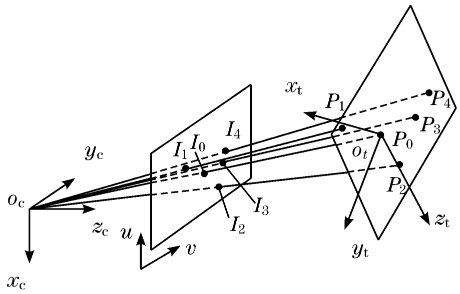

Fig. 3. Pinhole camera model.

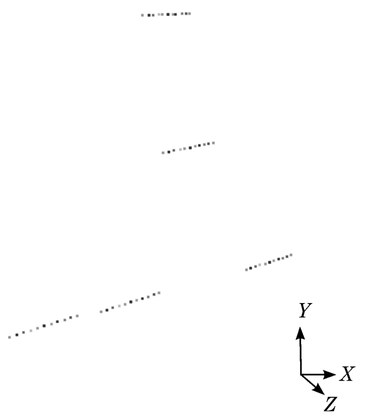

Fig. 4. Coordinates of reference points in the camera coordinate system with translation.

where P_{ti} denotes the coordinate of the reference points in the target coordinate system which is fixed by the target design. With the least-squares algorithm, the

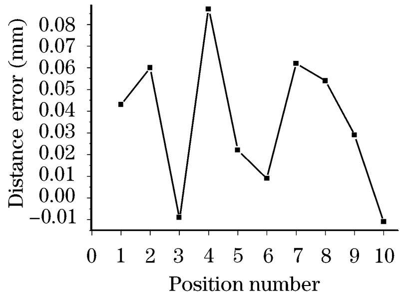

Fig. 5. Distance measurement error.

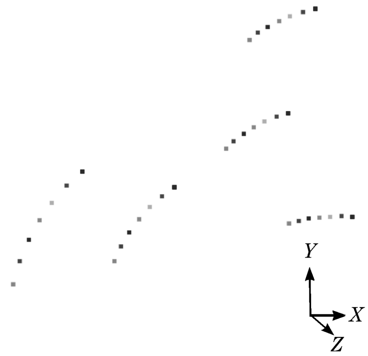

Fig. 6. Coordinates of reference points in camera coordinate system with rotation.

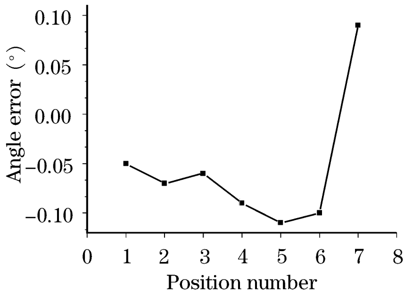

Fig. 7. Angle measurement error.

Table 1. Camera Calibration Result

Focal Length f (mm)	12.233
Distortion Factor	0.000374
Scale Factor	0.999272
Frame Center (pixel)	(640.499,465.306)

estimated value is obtained to let the object function:

$$F=\sum_{i=0}^{3}\sum_{j=i+1}^{4}\left(\|P_{ci}P_{cj}\|-\|P_{ti}P_{tj}\|\right)^2+\sum_{i=0}^{3}\sum_{j=i+1}^{4}\left(\angle P_{ci}P_{cj}P_{ck}-\angle P_{ti}P_{tj}P_{tk}\right)^2 \quad (4)$$

get the minimal value.

In the 3D environment, the target position in the camera coordinate system is represented with the origin

Table 2. Translation Movement Measurement Result

No.	t_x (mm)	t_y (mm)	t_z (mm)	Distance between No. i and No. i-1 (mm)
0	−88.558,	36.333	−13.354	
1	−85.542	36.312	−12.947	3.043
2	−82.569	36.302	−12.224	3.060
3	−79.579	36.283	−12.179	2.991
4	−76.619	36.280	−11.303	3.087
5	−73.621	36.257	−10.929	3.022
6	−70.615	36.233	−10.803	3.009
7	−67.615	36.225	−10.192	3.062
8	−64.569	36.193	−10.419	3.054
9	−61.571	36.184	−9.989	3.029
10	−58.595	36.176	−9.702	2.989

Table 3. Rotation Measurement Result

No.	r_1	r_2	r_3	Angle between No. i and No. i-1 (°)
0	0.675	−0.009	−0.738	
1	0.609	−0.008	−0.793	4.95
2	0.539	−0.007	−0.843	4.93
3	0.464	−0.006	−0.886	4.94
4	0.387	−0.005	−0.922	4.91
5	0.307	−0.004	−0.952	4.89
6	0.224	−0.002	−0.974	4.90
7	0.137	−0.004	−0.991	5.09

coordinates of the target in the camera coordinate system:

$$\mathbf{T} = [\ t_x \quad t_y \quad t_z\ ]^{\mathrm{T}} = [\ x_{P_{c0}} \quad y_{P_{c0}} \quad z_{P_{c0}}\ ]^{\mathrm{T}}. \tag{5}$$

For that the original point P_0 is critical to position estimation, two penalty factors M_1 and M_2 are applied to Eq. (4):

$$\begin{aligned} F = {} & M_1 \sum_{j=1}^{4} (\|P_{c0}P_{cj}\| - \|P_{t0}P_{tj}\|)^2 \\ & + M_2 \sum_{j=1}^{3} \sum_{k=j+1}^{4} (\angle P_{c0}P_{cj}P_{ck} - \angle P_{t0}P_{tj}P_{tk})^2 \\ & + \sum_{i=1}^{3} \sum_{j=i+1}^{4} (\|P_{ci}P_{cj}\| - \|P_{ti}P_{tj}\|)^2 \\ & + \sum_{i=1}^{2} \sum_{j=i+1}^{3} \sum_{k=j+1}^{4} (\angle P_{ci}P_{cj}P_{ck} - \angle P_{ti}P_{tj}P_{tk})^2. \end{aligned} \tag{6}$$

With the least squares, the five reference points $P_{ci}(x_{P_{ci}}, y_{P_{ci}}, z_{P_{ci}})$ are fitted into a plane $Ax_n+By_n+Cz_n+D=0$. The vector of the $o_t z_t$ axis is

$$\mathbf{u} = [r_1\ r_2\ r_3]^{\mathrm{T}} = \left[\frac{A}{\sqrt{A^2+B^2+C^2}} \quad \frac{B}{\sqrt{A^2+B^2+C^2}} \quad \frac{C}{\sqrt{A^2+B^2+C^2}}\right]^{\mathrm{T}}. \tag{7}$$

Since the camera is fixed during measurement, this directional vector can represent the orientation of the target plane.

In the experiment, a Basler A631f digital CCD camera is used to test the measurement accuracy of the position and orientation estimate method. With the camera calibration method[14], the intrinsic parameters of the camera are shown in Table 1.

Using a high precision translation stage, the position of the target is changed 10 times with 3 mm each step in the field of the camera's view. The coordinates of the five reference points in the camera coordinate system are obtained with the position and orientation measurement method, as shown in Fig. 4. With Eq. (5), the position of the target plane, which denotes the movement of the translation stage, can be figured out. The result is shown in Table 2. The distance measurement error is less than 0.1 mm, which is shown in Fig. 5.

With a turntable, the position of the target is changed 7 times with 5° each step in the field of the camera's view. The coordinate of the five reference points in the camera coordinate system is shown in Fig. 6. With Eq. (7), the vector of $o_t z_t$ axis which denotes the target orientation is calculated out. Angles between the adjacent positions are shown in Table 3. The angle measurement error is better than 0.2°, which is shown in Fig. 7.

In conclusion, a novel and effective position and orientation measurement algorithm based on the view of a single camera is introduced. With a designed target, the relationship between the reference points is fixed. According to the pinhole camera model, the coordinates of the reference points are estimated by considering the same relative position in the camera coordinate system and target coordinate system, respectively. The position and orientation of the target in the camera coordinate system are derived. In the experiment, the position and orientation measurement method is tested, with the distance measurement precision better than 0.1 mm, and the angle measurement precision better than 0.2°.

This work was supported by the Program for New Century Excellent Talents in University (No. NCET-06).

References

1. Y. Fang, W. Dixon, D. Dawson, and P. Chawda, *IEEE Trans. System, Man, and Cybernetic–Part B: Cybernetics* **35,** 1041 (2005).
2. F. Tatar, J. Mollinger, and A. Bossche, in *Proceedings of IEEE on Sensors 2003* **2,** 987 (2003).
3. I. Karaulova, P. Hall, and A. Marshall, in *Proceedings of British Machine Vision Conference* 262 (2000).
4. S. Hao, Z. Chai, Z. Zhang, and Y. Hu, Electron. Opt. Control (in Chinese) **9,** 19 (2002).
5. Z. Li, Z. Ding, and F. Yuan, Chin. Opt. Lett. **6,** 564 (2008).

6. E. Paperno, I. Sasada, and E. Leonovich, *IEEE Trans. Magnetic* **37,** 1938 (2001).
7. D. Roetenberg, J. Slycke, and H. Veltink, *IEEE Trans. Bio. Engin.* **54,** 883 (2007).
8. M. Maeda, T. Oqawa, K. Kiyokawa, and H. Takemura, *ISWC* **1,** 77 (2004).
9. X. Gao, K. Xu, H. Zhang, and X. Liu, Chinese J. Sci. Instru. (in Chinese) **28,** 1479 (2007).
10. Y. Xu, Y. Jiang, and F. Chen, Acta. Opt. Sin. (in Chinese) **29,** 72 (2009).
11. C. Sun, X. Zhang, and Y. Qu, Chin. Opt. Lett. **3,** 585 (2005).
12. Q. Wu, X. Su, L. Xiang, and Y. Li, Chinese. J. Lasers (in Chinese) **34,** 259 (2007).
13. C. Sun, L. Tao, P. Wang, and L. He, Chin. Opt. Lett. **4,** 282 (2006).
14. R. Y. Tsai, *IEEE J. Robotics and Automation* **3,** 323 (1987).