TELKOMNIKA Indonesian Journal of Electrical Engineering
Vol.12, No.3, March 2014, pp. 2138 ~ 2144
DOI: http://dx.doi.org/10.11591/telkomnika.v12i3.4425

Cayley Transform of Position and Orientation Measurement with Monocular Vision

Li Mingsuo[1], Yan Zhanjun*[2], Li Wenqiang[3]
[1]College of Automation, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
[2,3]Science and Technology on Electro-Optical Laboratory, Luoyang 471000, China
Corresponding author, e-mail: yavzc@sina.com[1], yanzj672@gmail.com[2], wql@126.com[3]

Abstract
In monocular vision-based position and attitude measurement, the non-linear iteration method based on feature-point geometric constraints is usually used to obtain results with higher accuracy. However, due to its strict geometric constraints to the feature points of target object and the requirement for highly-accurate initial value by iteration process, this algorithm is relatively poor in reliability. In this paper, Cayley transform is applied to the monocular vision-based position and attitude measurement; through an equivalent transformation from the unit orthogonal rotation matrix to 3D vectors by using Cayley transform, the computational process can be linearized, such that the complicated iterations caused by non-linear feature-point geometric constraints are eliminated, and the number of required feature points is reduced. Test data show that Cayley transform is practical and effective in the monocular vision-based position and attitude measurement.

Keywords: *monocular vision, position and attitude measurement, linear algorithm, Cayley transform*

Copyright © 2014 Institute of Advanced Engineering and Science. All rights reserved

1. Introduction

3D position and attitude measurement is to figure out the relative position and attitude relationship between two coordinate systems. Geometric features adopted by visual image-based 3D position and attitude measuring technique can be classified into point, straight-line and high-level geometric features, etc. For line feature or high-level geometric feature, during the measuring process, what're participated in operation are the edges of these features, which proposes higher requirements for the fabrication of camera's depth of field, FOV and labeling point, thus increasing the cost of measurement. Therefore, lots of researches focus on the method of 3D attitude measurement based on point feature. For example, M. Ferri proposed a method to realize object 3D attitude measurment based on straight-line feature, quadratic curve and quadratic surface [1], and Chen Zen proposed a method to realize object 3D attitude measurement based on circular label [2], etc.

The method of monocular position and attitude measurement based on point feature is to image target through single camera [3]. Given camera intrinsic parameters, it's to estimate the camera's rotation matrix **R** and translation vector **t** from the corresponding relationship between space point and its image point, which is usually called "PnP problem". The methods to solve PnP problem generally fall into two categories: one is linear algorithm, such as the position and attitude measurement model based on Lie algebra proposed by Drummond T [4] and Ortegon-aguilar J [5]; this algorithm is relatively high in speed, but easy to be affected by error and of low accuracy. The other is non-linear iteration algorithm, which expresses the PnP problem as a constrained non-linear optimization problem, such as the iteration solution algorithm for target object position and attitude based on POSIT method proposed by Dementhond F [6] and Gramegna T [7]; this algorithm is of higher accuracy, but there exist two problems: (1) at the beginning of iteration solution, it needs to provide a better estimated value of the true value, i.e., the initial value of object position and attitude; (2) the iteration solution algorithm is more complicated, and thus not applicable to some cases which have a higher requirement for speed.

Research on PnP problem is generally classified into two types: $2<n<6$ and $n\geq 6$. For P2P problem, there are always infinite groups of solutions. For P3P problem, there are at most 4 solutions, and the upper limit of solution can be reached. For P4P problem, when the 4

Received August 15, 2013; Revised October 5, 2013; Accepted October 14, 2013

feature points are coplanar, there is a only solution; when the 4 feature points are non-coplanar, there are possibly 4 solutions at most, and the upper limit of solution can be reached. For P5P problem, there are possibly 2 solutions at most, and the upper limit of solution can be reached [8]. When n≥6, the position and attitude of target object can be derived from direct linear computations, but it's very sensitive to image noise and doesn't consider the orthogonal constraints of rotation matrix.

Based on the unit orthogonal constraint of rotation matrix, Chen Shan has directly solved the rotation and translation matrix of object by using least square method [9]; the number of feature points solved by this method is somewhat a lot, causing a complicated computational process. Quan Long has conducted polynomial decomposition of the position and attitude measurement model, taking the even-order terms in equation as variables, to realize the linearization of position and attitude measuring process; by increasing the number of feature points from 4 to 5, the number of polynomial expressions is effectively increased; and rotation matrix is derived by using singular value decomposition, thus improving the accuracy in position and attitude computation process.

All computation models used in position and attitude measurement are generally non-linear; therefore, in order to realize the linear solution for position and attitude measurement, it needs to conduct mathematical transform in accordance with the characteristics of position and attitude measurement model. Tang Jianliang proposed a geometric theory method based on partial differential equation, to realize the stable linear solution for position and attitude measurement model with 5 feature points [10].

2. Position and Attitude Measurement Model Based on 3 Feature Points

Position and attitude measurement model is constructed on the basis of camera pinhole imaging model, as shown in Figure 1. O_w-$X_wY_wZ_w$ is the world coordinate system, and O_c-$X_cY_cZ_c$ is the camera coordinate system.

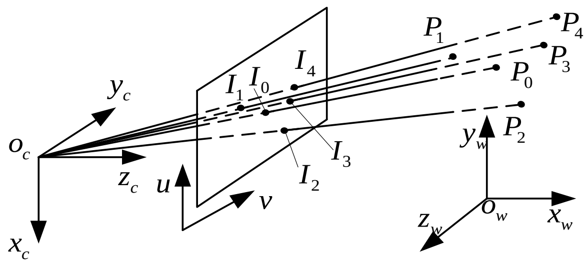


Figure 1. Point Pinhole Imaging Model

Given the coordinates of 3 reference points of P_i (i=0, 1, 2) in world coordinate system Ow-$X_wY_wZ_w$, distance f from the camera coordinate system center O_c to image plane is derived from camera calibration (camera intrinsic parameters are derived from camera calibration, including the focal length of camera, radial distortion coefficient of lens, center coordinate of image plane, and uncertain horizontal proportional factor). In consideration of radial distortion, the ideal image coordinate of feature point $I_i(u_i,v_i)^T$ (i=0, 1, 2) can be derived by the transformation formula (1)between ideal coordinate and actual coordinate:

$$\begin{cases} u_i = u_d\,(1+k_1r^2) \\ v_i = v_d\,(1+k_1r^2) \end{cases} \tag{1}$$

Where, k_1 is the radial distortion coefficient of camera's lens.wher

$$r = \sqrt{u_d^2 + v_d^2}$$

According to camera pinhole imaging model, coordinate of $I_{ci}(i=0,1,2)$ for the ideal image point of feature point in camera coordinate system can be expressed as:

$$s_i \begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = K \begin{bmatrix} x_i^c \\ y_i^c \\ z_i^c \end{bmatrix} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i^c \\ y_i^c \\ z_i^c \end{bmatrix}$$

Where, **K** is the intrinsic parameter matrix, s_i is the proportional factor; f_x, f_y are the effective focal length,$(u_0\ v_0)$ is the center coordinate of image plane, $(x_i^c\ y_i^c\ z_i^c)$ is the coordinate of I_{ci} in camera coordinate system.

In camera coordinate system, after camera imaging, the relationship between the feature point's camera coordinate and its world coordinate can be expressed as:

$$\begin{bmatrix} x_i^c \\ y_i^c \\ z_i^c \\ 1 \end{bmatrix} = \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} x_i^w \\ y_i^w \\ z_i^w \\ 1 \end{bmatrix} = M \begin{bmatrix} x_i^w \\ y_i^w \\ z_i^w \\ 1 \end{bmatrix} \tag{2}$$

Where, **R** is 3×3 unit orthogonal matrix; **t** is 3D translation vector; $\left(x_i^w\ y_i^w\ z_i^w\right)$ is the coordinate of feature point P_i in world coordinate system.

3. Linear Solution Algorithm for Position and Attitude Measurement Based on Cayley Transform

For Cayley expression of rotation matrix, **M** (n) is the set of all components of n-order matrix; **LN** (n) is the set of n-order invertible matrix. Suppose matrix $A \in M(n)$ and $I + A \in LN(n)$, of which, I is the unit matrix, the following transformation is called as Cayley transform of matrix A:

$$B = \varphi(A) = (I - A)/(I + A)$$

Given $ME(n) = \{A \in M(n) : I + A \in LN(n)\}$, Cayley transform is the one-to-one mapping on **ME** (n). For any 3-order rotation matrix **R**, by using the method of matrix multiplication decomposition, the rotation axis, rotation angle of rotation matrix can be quantified by matrix feature parameters. Unit vector of $\vec{e} = (x \quad y \quad z)$ is taken as the direction vector of rotation axis, and transformation matrix of the rotation transform with rotation angle of θ is:

$$R_\theta = \begin{bmatrix} x^2 + (1-x^2)\cos\theta & xy(1-\cos\theta) - z\sin\theta & xz(1-\cos\theta) + y\sin\theta \\ xy(1-\cos\theta) + z\sin\theta & y^2 + (1-y^2)\cos\theta & yz(1-\cos\theta) - x\sin\theta \\ xz(1-\cos\theta) - y\sin\theta & yz(1-\cos\theta) + x\sin\theta & z^2 - (1-z^2)\cos\theta \end{bmatrix} \tag{3}$$

From Equation (3), it is known that any 3-order rotation matrix **R** can be uniquely expressed by the unit direction vector and rotation angle of rotation axis, as shown in Figure 2, of which, point **w** is the point that rotates around the direction vector **e**.

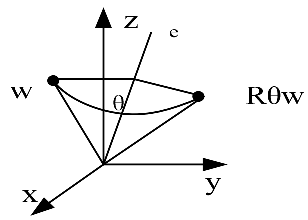


Figure 2. Coordinate System of Rotation

For linear solution for position and attitude measurement model, Given a pair of points corresponding relationship from 3D space to 2D image:

$$\overrightarrow{p_i^{\,w}} = \begin{pmatrix} x_i^w & y_i^w & z_i^w \end{pmatrix}^T \longleftrightarrow \overrightarrow{u_i} = (u_i \quad v_i \quad 1)^T \qquad i = 1, 2, \cdots N$$

The mapping from space point to image point can be expressed as:

$$s_i \overrightarrow{u_i} = R(\overrightarrow{p_i^{\,w}} - \vec{t}\,), \qquad i = 1, 2 \cdots N \tag{4}$$

In this paper, 3 feature points are adopted for target, and feature points *Pi*(*i*=0, 1, 2) are non-collinear. Let

$$\begin{pmatrix} c_0^j & c_1^j & c_2^j \end{pmatrix}^T = \begin{pmatrix} \overrightarrow{p_0^w} & \overrightarrow{p_1^w} & \overrightarrow{p_2^w} \end{pmatrix}^{-1} \overrightarrow{p_j^w}, \qquad j = 3, 4, \cdots N \tag{5}$$

The other points on target can be expressed by feature points, and Equation (5) can be transformed into:

$$\overrightarrow{p_j^w} = \sum_{i=0}^{2} c_i^j \overrightarrow{p_i^w}, \qquad j = 3, 4, \cdots N \tag{6}$$

With Equation (6) inserted into Equation (4), it can be derived that

$$s_j \overrightarrow{u_j} = \sum_{i=0}^{2} \left(c_i^j \overrightarrow{u_i} \right) s_i - \left(1 - \sum_{i=0}^{2} c_i^j \right) \vec{t}, \quad j = 3, 4, \cdots N \tag{7}$$

Then, homogeneous linear equations about $\begin{pmatrix} s_0 & s_1 & s_2 & \vec{t} \end{pmatrix}$ are derived:

$$\sum_{i=0}^{2} \left(c_i^j \left[\overrightarrow{u_j} \right]_\times \overrightarrow{u_i} \right) s_i - \left(1 - \sum_{i=0}^{2} c_i^j \right) \left[\overrightarrow{u_j} \right]_\times \vec{t} = 0, \qquad j = 3, \cdots N \tag{8}$$

Least-square solution for Equation (8) is

$$\begin{cases} s_i = \alpha s_i^*, \qquad i = 0, 1, 2 \\ \qquad \vec{t} = \alpha \vec{t}^{\,*} \end{cases} \tag{9}$$

Due to that vector norm is kept unchanged in rotation transformation, $\alpha > 0$ is unknown scalar. With Equation (9) inserted into Equation (4), it can be derived that:

$$\alpha = \frac{1}{3}\sum_{i=0}^{2}\left(\left\|\overrightarrow{p_i^w}\right\| \Big/ \left\|s_i^*\overrightarrow{u_i} + \overrightarrow{t^*}\right\|\right) = \alpha^* \tag{10}$$

Let w is Cayley expression of R, then $R = \left(I - [\vec{w}]_\times\right) / \left(I - [\vec{w}]_\times\right)$, then, by equivalent transformation, it can be derived that

$$\left[\overrightarrow{p_i^w} + \alpha^*\left(s_i^*\overrightarrow{u_i} + \overrightarrow{t^*}\right)\right]_\times \vec{w} = \alpha^*\left(s_i^*\overrightarrow{u_i} + \overrightarrow{t^*}\right) - \overrightarrow{p_i^w}, \quad i = 0,1,2 \tag{11}$$

By least square method, it is derived that:

$$\overrightarrow{w^*} = -\left(\sum_{i=0}^{2}\left[\overrightarrow{p_i^w} + \alpha^*\left(s_i^*\overrightarrow{u_i} + \overrightarrow{t^*}\right)\right]_\times^2\right)^{-1}\left(\sum_{i=0}^{2}\left[\overrightarrow{p_i^w} + \alpha^*\left(s_i^*\overrightarrow{u_i} + \overrightarrow{t^*}\right)\right]_\times^T\left(\alpha^*\left(s_i^*\overrightarrow{u_i} + \overrightarrow{t^*}\right) - \overrightarrow{p_i^w}\right)\right) \tag{12}$$

From Equation (12), it can be known that the rotation matrix of this position and attitude measurement model is $R = \left(I - \left[\overrightarrow{w^*}\right]_\times\right) \Big/ \left(I + \left[\overrightarrow{w^*}\right]_\times\right)$, translation vector is $\overrightarrow{t^*} = \alpha^* R^{*T}\overrightarrow{t^*}$.

4. Solution Procedure for Position and Attitude Measurement Algorithm

Based on the above theoretical derivation, solution procedure for position and attitude measurement algorithm can be summarized as follows:

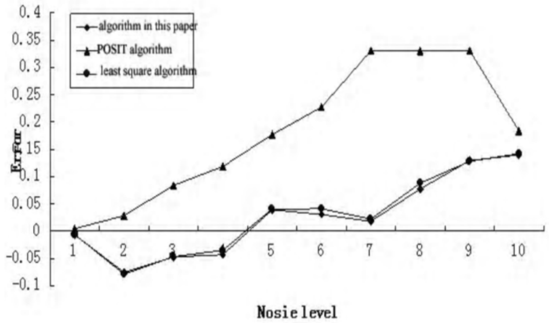
(1) According to the pinhole imaging model, build a position and attitude measurement system with 3 feature points.
(2) Based on the picture taken by camera and the radial distortion coefficient of lens k_1, the ideal image coordinate of feature point *Ici*(*ui*, *vi*) is derived, i.e., i = 1, 2, 3.
(3) Cayley transform is used to linearly transform the rotation matrix between world coordinate system and camera coordinate system to 3D vectors.
(4) Select 3 points other than feature points from the picture taken by camera, and express them with feature points set.
(5) Insert the 3 points expressed by the feature point set into Equation (9), and obtain the least-square solution for **Si** and translation vector **t**.
(6) Due to least-square solution and vector rotation transformation norm are not changed, unknown scalar can be expressed by given data.
(7) Insert Equation (12) and Cayley transform of rotation matrix **R** into Equation (11), and the least-square solution for **w** is derived.
(8) Based on the derived least-square solution for rotation matrix **R** and translation vector **t**, the position and attitude of target relative to camera is figured out.

5. Accuracy Verification of Position and Attitude Measurement Algorithm

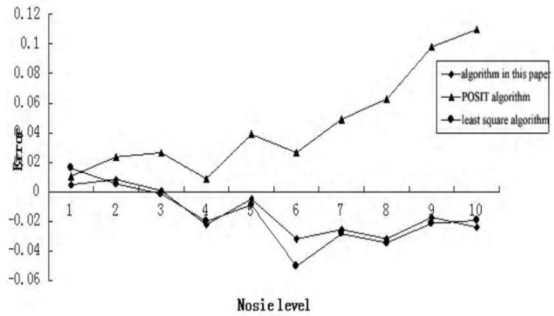
In accordance with the above theoretical analysis, the linear algorithm for Cayley transform-based position and attitude measurement model is verified through experiments.

In measuring process, 3-axis rotation platform is used to control the position and attitude of target; camera is used to take pictures of target at different measuring positions, and to extract feature points of target. The algorithm of this paper, POSIT algorithm and least square algorithm are used respectively to calculate the position and attitude of target; the accuracy of position and attitude measurement can be derived by comparing the measured results with the position and attitude control results of 3-axis rotation platform.

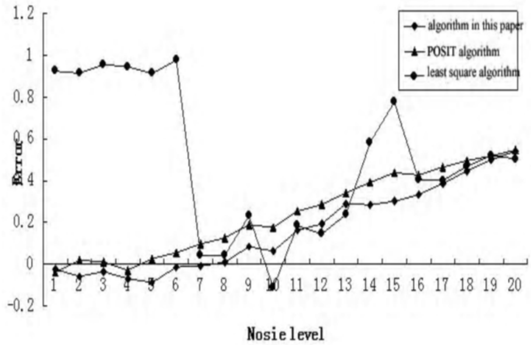
With an interval of 3 degrees, the 3-axis rotation platform of position and attitude measurement system drives the target to rotate respectively in 3 DOF, i.e., azimuth, pitch and roll; camera is used to take a picture at each position, which is compared with the picture of initial position (position of 0°), and then the relationship between these two positions are figured out. The measuring range of azimuth angle is set to 0°~60°, that of pitch angle is set to 0°~30° , and that of roll angle is set to 0°~30°. The measured results obtained are shown as in Figure 3.



(a) Error of pitch angle



(b) Error of roll angle



(c) Error of azimuth angle

Figure 3. Position and Attitude Measurement Test Errors

Of 3 DOF (pitch, roll and azimuth), azimuth angle is rotated in a large degree and makes target off camera much faster. In the measuring process of azimuth DOF, a better iteration initial value isn't derived from least square method, and a big oscillation is existing during measurement. Overall error of the algorithm proposed in this paper is lower than that of the least square algorithm and POSIT algorithm. Test results demonstrate that the algorithm in this paper can reduce errors in position and attitude measurement significantly.

6. Conclusion

This paper studies on 3 non-collinear feature points tracking the position and attitude of moving target. The method is based on that the rotation matrix of moving target position and attitude matrix can be expressed by 3 variables: 1 rotation axis and 2 rotation angles. This expression is unique, with definite geometric definition and calculation formula. In this paper, based on such expression, Cayley transform of matrix is introduced, through equivalent linear transformation from 3-order rotation matrix to 3D vectors in the position and attitude measuring process, linear solution for position and attitude measurement algorithm is realized. Actual measurement test shows that the algorithm in this paper is of higher accuracy. By comparison to POSIT algorithm and least square algorithm, it can be found that with the algorithm based on Cayley transform, the number of feature points is reduced, the non-linear geometric constraints among feature points are weakened, the solving process for position and attitude measurement algorithm is optimized, and the accuracy of algorithm is improved.

References

[1] Ferri M, Mangili F, Viano G. Projective pose estimation of linear and quadratic primitives in monocular computer vision. *CVGIP: Image Understanding*. 1993; 58(1): 66-84.

[2] Chen Zen, Huang Jenbin. A Vision-Based Method for the Circle Pose Determination With a Direct Geometric Interpretation. *IEEE Transactions on Robotics and Automation*. 1999; 15(6): 1135-1140.

[3] Huang Guiping, Wang Baofeng, Li Guanyun. *Principle of coordinate's acquisition using one single camera*. Proceedings of 6th Conference on Optical 3D Measurement Techniques(C), Zurich: ETH. 2003.

[4] Drummond T, Cipollar R. Application of lie algebras to visual servoing. *International Journal of Computer Vision*. 2000; 37(1): 21-41.

[5] Ortegon-aguilar J, Bayro-corrochano E. Lie. Algebra and system identification technique for 3D rigid motion estimation and monocular tracking. *Journal of mathematical Imaging and Vision.* 2006; 25: 173-185.

[6] Dementhond F, Davisl S. Model-based object pose in 25 lines of code. *International Journal on Computer Vision*. 1995; 15: 123-141.

[7] Gramegna T, Venturino L, Cicirelli G, et al. Optimization of the POSIT algorithm for indoor autonomous navigation. *Robotics and Autonomous Systems*. 2004; 48: 145-162.

[8] Chen Shan, Zhou Tao, Zhang Xiaodong, et al. Monocular vision measurement system of the position and attitude of the object. *Chinese Journal of Sensors and Actuators*. 2009; 9: 2011-2015 (in Chinese).

[9] Quan Long, Lan Zhongdan. Linear n-point camera poses determination. *IEEE Transactions on pattern analysis and machine intelligence*. 1999; 21(8): 774-780.

[10] Tang Jianliang, Chen Wensheng, Wang Jie. A novel linear algorithm for P5P problem. *Applied Mathematics and Computation*. 2008; 205: 628-634.