A New Method for Pose Estimation from Line Correspondences

QIN Li-Juan[1, 2] ZHU Feng[2]

Abstract We can usually determine the pose of objects from three lines in a general position. The configuration of three non-coplanar lines that intersect at two points has some particular characteristics, which three lines in a general position do not have. Here, we present a new method of determining object pose using this particular line configuration. In theory, this method enriches the pose estimation methods from three line correspondences. In addition, it provides guidance for practical applications. Furthermore, we propose a method to deal with multi-solution phenomenon and a new iterative method. Simulation results demonstrate that our algorithm works speedily and robustly.

Key words Line correspondences, pose estimation, multi-solution phenomenon, iterative method

The way to determine object pose from 2D to 3D feature correspondences[1−3] is an important problem in computer vision. It has wide applications in such realms as robot self-positioning[4−5], robot navigation, robot obstacle avoidance, object recognition, object tracking, and camera calibration.

Points, lines, and high-level features (circles, ellipses, conic sections, etc.) are the most popular configurations for pose estimation. Lines have the following advantages over points[6]: 1) Images of man-made environments contain many lines; 2) Lines are easy to extract and the accuracy of line detection is high; 3) Lines are superior in dealing with occluding and ambiguous situations, which are very likely to be encountered in many practical applications. In contrast to high-level features, lines have been largely preferred for the following reasons: 1) Images of man-made environment contain more lines than high-level features; 2) The mathematical expressions for lines are easy; thus they work efficiently. Therefore, lines outperform points and high-level features for pose estimation in some respects.

Much work has been done on the pose estimation from three line correspondences. Most researchers used the mathematical model in which the normal vectors of the explanation planes are perpendicular to object lines in the camera frame. Usually, a minimum of three lines is required to solve the problem of pose estimation. Then, three lines can produce three nonlinear equations. Thus, the pose estimation problem is converted into a problem of solving a non-linear equation sets. The pose estimation method determines the object pose from three line correspondences in a general position; however, one of the weaknesses of this technique is that its mathematical expression is too complex to solve.

At present, most researchers focused on how to solve the pose estimation algorithm from line correspondences. The approaches to this question fall into two categories: analytical methods and numerical methods. To simplify the analytical methods, Dhome[7] and Chen[8] set up a particular model coordinate system and a particular viewer coordinate system. At last, they developed an 8th-degree polynomial to determine the closed-form solutions of object pose from three lines in a general position. Horaud[9] derived a biquadratic polynomial in one unknown for the case of three non-coplanar lines. The roots of such an equation can be found using analytical or numerical method. For real-time applications, we are interested in analytical solutions free of initialization. However, the drawback of analytical techniques is the presence of multiple solutions and large errors in determining the pose. For numerical methods, Yuan[10] has used Newton′s method to estimate the orientation and location of an object with respect to a camera. Liu[11] used alternative iterative approaches to solve the viewing parameters. Christy[12] computed the pose with iterative method, which started with a solution close to true solution. Numerical approaches usually have better pose estimation accuracy than analytical methods. However, they require a good initial estimation of true solution and are time-consuming. Therefore, such approaches cannot be used in such tasks as they require high speed performance (visual servoing, object tracking).

For three non-coplanar lines that intersect at two points, it has some particular characteristics that three lines in general position do not have. For this particular line configuration, we present a new method to determine object poses. Its mathematical expression is easy to solve. To get the solution of this method, we propose a new iterative method, which is based on geometric relationship and step acceleration method. This method is free of the optimization process of complex nonlinear equation and the process of solution is fast; thus, it can ensure the real-time performance of system.

Multi-solution phenomenon, the most puzzling question for pose computation, has been widely studied from the angle of point correspondences[13−15]; however, little is done from line correspondences except for [7]. The method proposed in the paper is to compute all the solutions and use judgment rules to get rid of wrong solutions. This method has too much computation. If we can determine the unique solution before computing all the solutions, then we can avoid unnecessary iterative computation and thus save time. As for the algorithm in the paper, we have designed a new method to determine unique solution by geometric information of 3D lines. This method does not need to compute all the solutions to find the unique solution, and therefore it is time-saving.

The remainder of this paper is organized as follows. In Section 1, a new pose estimation algorithm is presented. Section 2 presents a new method to find the unique solution by geometric information. In Section 3, a new iterative method based on geometric relationship and step acceleration method is presented. Section 4 presents experimental results. Section 5 is the conclusion.

Received November 30, 2006; in revised form May 30, 2007
Supported by National Natural Science Foundation of China (60575024, 60705011)
1. Shenyang Institute of Automation, Chinese Academy of Sciences, Shenyang 110116, P. R. China 2. Graduate University of Chinese Academy of Sciences, Beijing 100080, P. R. China
DOI: 10.3724/SP.J.1004.2008.00130

1 A new pose estimation algorithm

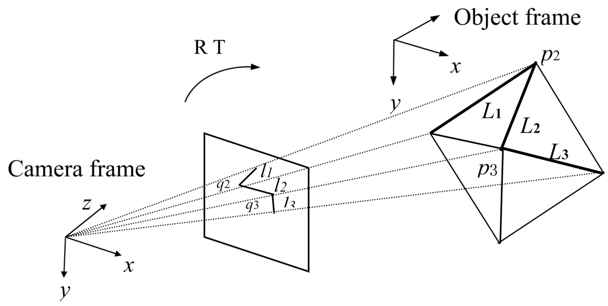


Fig. 1 Perspective projection of three lines

We consider a pin-hole camera model, and assume that the intrinsic camera parameters are known. As shown in Fig. 1, we assume that $\boldsymbol{v}_i = (A_i, B_i, C_i)$ is the director vector of space line L_i in the camera frame. Line L_1 and line L_2 intersect at p_2. Line L_2 and line L_3 intersect at p_3. Image points q_2 and q_3 corresponding to p_2 and p_3 have coordinates of (x_2, y_2, z_2) and (x_3, y_3, z_3). The coordinates of point, p_2 and p_3, are (k_2x_2, k_2y_2, k_2z_2) and (k_3x_3, k_3y_3, k_3z_3), where k_2 and k_3 are unknown. We consider an image line l_i characterized by a vector $(-b_i, a_i, f)$ and a point of coordinates of (x_i, y_i, f). At the same time, we also know that 1) $|p_2p_3| = l$; 2) The angle between line L_1 and line L_2 is α, the angle between line L_2 and line L_3 is β, and the angle between line L_1 and line L_3 is γ. The perspective projection model confines line L_i, image line l_i and the origin of the camera frame in a plane. This plane is called the explanation plane. The vector, $\boldsymbol{n}_i = (N_{i1}, N_{i2}, N_{i3})$, normal to this plane is equal to the cross product of vector $\boldsymbol{v}_i$ and vector $\boldsymbol{ot}_i$

$$\boldsymbol{n}_i = \boldsymbol{v}_i \times \boldsymbol{ot}_i = (a_if, b_if, c_i) \tag{1}$$

For line L_1, the norm of the director vector of line L_1 is 1, line L_1 is perpendicular to the normal vector N_1, and the angle between line L_1 and line L_2 is α. We obtain

$$\left\{\begin{array}{l} A_1^2 + B_1^2 + C_1^2 = 1 \\ A_1N_{11} + B_1N_{12} + C_1N_{13} = 0 \\ A_1A_2 + B_1B_2 + C_1C_2 = \cos\alpha \end{array}\right. \tag{2}$$

Thus, A_1, B_1, and C_1 can be expressed by A_2, B_2, and C_2:

$$\begin{array}{l} A_1 = \frac{-(mn+pq)\pm\sqrt{(mn+pq)^2-(n^2+q^2+1)(p^2+m^2-1))}}{(n^2+q^2+1)} \\ B_1 = m + nA_1 \\ C_1 = p + qA_1 \end{array} \tag{3}$$

where,

$$m = \frac{\cos\alpha N_{13}}{B_2N_{13} - N_{12}C_2}, \quad n = \frac{N_{11}C_2 - A_2N_{13}}{B_2N_{13} - N_{12}C_2}$$

$$p = \frac{-\cos\alpha N_{12}}{B_2N_{13} - C_2N_{12}}, \quad q = \frac{A_2N_{12} - B_2N_{11}}{B_2N_{13} - C_2N_{12}}$$

For line L_3, the norm of the director vector of line L_3 is 1, line L_3 is perpendicular to the normal vector $\boldsymbol{N}_3$, and the angle between line L_3 and line L_2 is β. We have

$$\left\{\begin{array}{l} A_3^2 + B_3^2 + C_3^2 = 1 \\ A_3N_{31} + B_3N_{32} + C_3N_{33} = 0 \\ A_3A_2 + B_3B_2 + C_3C_2 = \cos\beta \end{array}\right. \tag{4}$$

In the above equations, A_3, B_3, and C_3 can similarly be expressed by A_2, B_2, and C_2:

$$\begin{array}{l} A_3 = \frac{-(gh+wl)\pm\sqrt{(gh+wl)^2-(h^2+l^2+1)(w^2+g^2-1))}}{(h^2+l^2+1)} \\ B_3 = g + hA_1 \\ C_3 = w + lA_1 \end{array} \tag{5}$$

where,

$$g = \frac{\cos\beta N_{33}}{B_2N_{33} - N_{32}C_2}, \quad h = \frac{N_{31}C_2 - A_2N_{33}}{B_2N_{33} - N_{32}C_2}$$

$$w = \frac{-\cos\beta N_{32}}{B_2N_{33} - C_2N_{32}}, \quad l = \frac{A_2N_{32} - B_2N_{31}}{B_2N_{33} - C_2N_{32}}$$

The director vectors of line L_2 can be expressed by points p_2 and p_3:

$$\begin{pmatrix} A_2 \\ B_2 \\ C_2 \end{pmatrix} = \begin{pmatrix} k_3x_3 - k_2x_2 \\ k_3y_3 - k_2y_2 \\ k_3z_3 - k_2z_2 \end{pmatrix} \tag{6}$$

From known condition 1) $|p_2p_3| = l$, we get

$$(k_3x_3 - k_2x_2)^2 + (k_3y_3 - k_2y_2)^2 + (k_3z_3 - k_2z_2)^2 = l^2 \tag{7}$$

From (7), we obtain

$$k_3 = \frac{-f_1k_2 \pm \sqrt{(f_2k_2)^2 - 4f_1(f_3k_2^2 - l^2)}}{2f_1} \tag{8}$$

where, $f_1 = x_2^2 + y_2^2 + z_2^2$, $f_2 = -2(x_2x_3 + y_2y_3 + z_2z_3)$, and $f_3 = x_3^2 + y_3^2 + z_3^2$.

Substituting (8) into (6), we can get the director vectors of line L_2. They have two expressions. Substituting the director vectors of line L_2 into (3) and (5), we can obtain the vectors of line L_1 and line L_3. Line L_1 and line L_3 both have two expressions.

From known condition 2) that the angle between line L_1 and line L_3 is γ , we have

$$A_1A_3 + B_1B_3 + C_1C_3 = \cos\gamma \tag{9}$$

Substituting the director vectors of line L_1 and line L_3 into (9), we can obtain eight equations for one variant k_2. Therefore, there exist eight solutions for pose estimation from three lines. In practical applications, one needs to find the correct one from eight solutions.

2 Solving multi-solution problem

As shown in Fig. 2 (a), we assume that p_i $(i = 1, 2, 3, 4)$ are four points on a quadrangular pyramid. In addition, we assume that the director vector of line L_1 is from point p_1 to point p_2, the director vector of line L_2 is from point p_3 to point p_2, and the director vector of line L_3 is from point p_3 to point p_4.

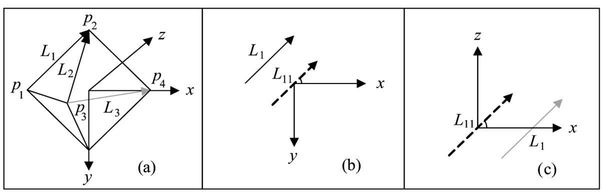


Fig. 2 Judgment method to find correct solution

First, the position relationship of the intersection points p_2, p_3 on the model is invariant because the view angle of the camera is very small. As for the quadrangular pyramid, the third point p_3 is always nearer to the camera than the second point p_2 when the model moves within the view angle of the camera. As demonstrated in Section 1, k_2 and k_3 express the distances between point p_2, p_3 and the optical center of the camera. Thus, the value of k_3 is smaller than that of k_2. Hence, from (8), we obtain $k_3 = \frac{-f_1 k_2 - \sqrt{(f_2 k_2)^2 - 4f_1(f_3 k_2^2 - l^2)}}{2f_1}$. Therefore, there is one expression for line L_2 from (6) and there are two expressions for each line of L_2, L_3 from expressions (3) and (5). Thus, there exist four solutions for pose estimation from lines L_i ($i = 1, 2, 3$) at this time.

Second, the angles between three lines L_i ($i = 1, 2, 3$) and x, y, z axes of camera frame change little because the view angle of the camera is very small. For the quadrangular pyramid, the angle between the first line L_1 and x axis always ranges between 0 and $\pi/2$ (see Fig. 2 (b)). At the same time, the angle between line L_3 and x axis also ranges between 0 and $\pi/2$ (see Fig. 2 (c)). From (3), A_1 expresses the cosine of the angle between line L_1 and x axis; thus the value of A_1 is positive. Therefore, we can get the director vectors of line L_1 as follows

$$\begin{aligned} A_1 &= \frac{-(mn+pq) + \sqrt{(mn+pq)^2 - (n^2+q^2+1)(p^2+m^2-1))}}{(n^2+q^2+1)} \\ B_1 &= m + nA_1 \\ C_1 &= p + qA_1 \end{aligned} \tag{10}$$

From (5), A_3 expresses the cosine of the angle between line L_3 and x axis. This angle always ranges between 0 and $\pi/2$; thus the value of A_3 is positive. Therefore, we can get the director vectors of line L_3 as follows

$$\begin{aligned} A_3 &= \frac{-(gh+wl) + \sqrt{(gh+wl)^2 - (h^2+l^2+1)(w^2+g^2-1))}}{(h^2+l^2+1)} \\ B_3 &= g + hA_1 \\ C_3 &= w + lA_1 \end{aligned} \tag{11}$$

From (9), we obtain one equation $f(k_2) = 0$. Thus, there exists one solution for pose estimation from lines L_i ($i = 1, 2, 3$) at this time.

We determine the unique solution from geometric meaning of the model. This method judges which is the correct answer without computing all the solutions. Thus, it can avoid unnecessary computation and is time-saving.

3 A new iterative method

We obtain $f(k_2) = 0$ from (9). In this section, we present a new method which is based on geometric relationship and step acceleration method to compute k_2.

We suppose that the optical center is O, and the explanation planes formed by image line l_i and O is S_i ($i = 1, 2, 3$). The intersection line formed by explanation plane S_1 and S_2 is J_1, and the intersection line formed by explanation planes S_2 and S_3 is J_2 (Fig. 3).

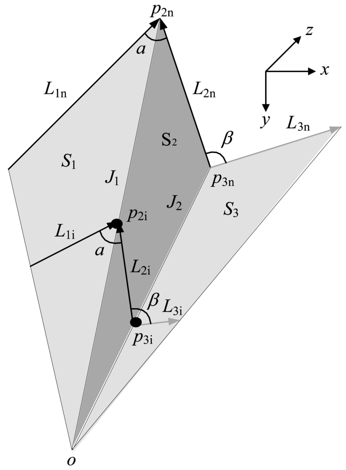


Fig. 3 Searched method based on geometric relationship

We start to search the correct value of k_2 based on geometric relationship (Fig. 3). First, we select a point p_{2i} on line J_1 near the optical center. Second, we continue searching the point p_{3i}, which is at a distance of l from point p_{2i} on line J_2. Point p_{2i} and point p_{3i} form the line L_{2i}. Third, we find a line L_{1i} on the explanation plane S_1, which forms an angle of α with line L_{2i}. Fourth, we find the line L_{3i} on the explanation plane S_3, which makes an angle of β with line L_{2i}. Line L_{1i}, and line L_{3i} form an angle of γ_i. With the increase of the distance between point p_{2i} and the optical center of camera, the angle between line L_{1i} and line L_{3i} becomes larger and larger. After n times of iterative process, this angle satisfies the known value γ. At this time, the whole iterative process finishes and point p_{2i} and lines L_{1i}, L_{2i}, and L_{3i} reach the correct position.

Based on step acceleration method, concrete iterative process is as follows.

We first initialize k_2 as k_0 and assume the iterative step to be Δk. We compute the value of $f(k_2)$ when $k_2 = k_0$.

The variant k_2 increases at the step of Δk and we compute the value of $f(k_2)$. This process finishes if $f(k_2) < 0$. At this time, we go back to the last iterative point and continue searching at half of the step Δk. Then, we continue searching iteratively. If the value of $f(k_2)$ is still smaller than zero, then we go back to the last iterative point and continue searching at half of the step Δk. If the value of $f(k_2)$ is bigger than zero, then we continue searching at the step of Δk. The process finishes when the value of $f(k_2)$ is small enough.

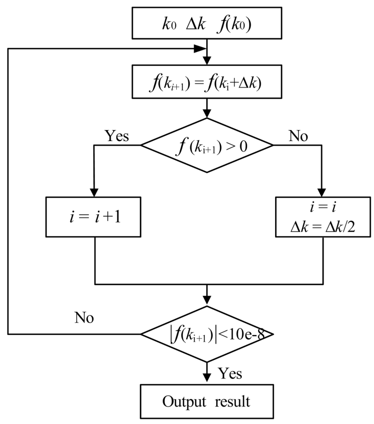


Fig. 4 Structure chart of step acceleration method

The iterative method can solve complex nonlinear equations. It is simple and time-saving. Unlike previous solution methods, the iterative method we introduce gives insight into the geometric characteristics of pose estimation problem. This method can avoid from being trapped in local minima. At the same time, it is free of the choice of the initial estimation and it can find the true value even though the initial estimation is far from the true value.

4 Experimental results

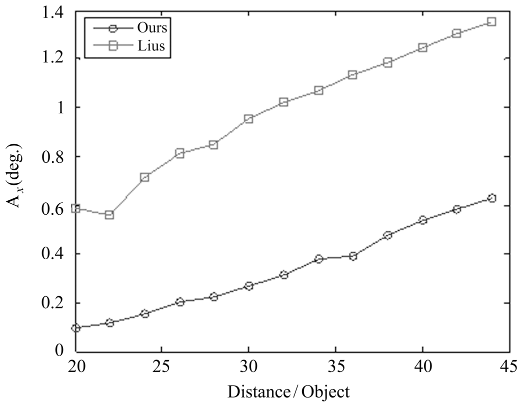
In our simulation experiments, the intrinsic parameters of camera are $\begin{pmatrix} 787.887 & 0 & 0 \\ 0 & 787.887 & 0 \\ 0 & 0 & 1 \end{pmatrix}$. The size of image is 512×512 and the view angle of camera is about $36^\circ \times 36^\circ$.

We add 0.25 pixel random noises to image lines. We choose 13 groups of test points at a distance of $20 \sim 45$ times of model size. We compute the error for 16 000 different poses at each test point. In Fig. 5, the horizontal coordinates are the ratio of the distance between the optical center and the model to the model size. The vertical coordinates of Fig. 5 (a)$\sim$(c) are degrees. The vertical coordinates of Fig. 5 (d)$\sim$(f) is three times of the RMS position error to the distance between the optical center and the model. The model size is 12.0208 mm.

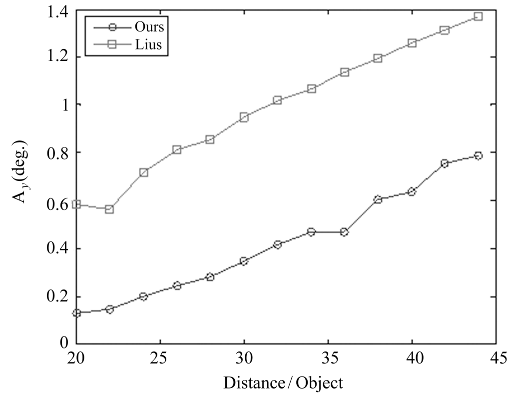
To examine the characteristics of the algorithm in this paper, we do contrastive experiments with the numerical algorithm in [11] at corresponding test points. Experimental results are as follows:

Experimental results (see Fig. 5) show the precision of yaw and pitch angles that rotate around the axes of x, y is two times more accurate than that of the algorithm in [11], and the translation precision along axes of x, y, and z is five times more accurate than that of the algorithm in [11]. Thus, the algorithm in this paper has good precision.

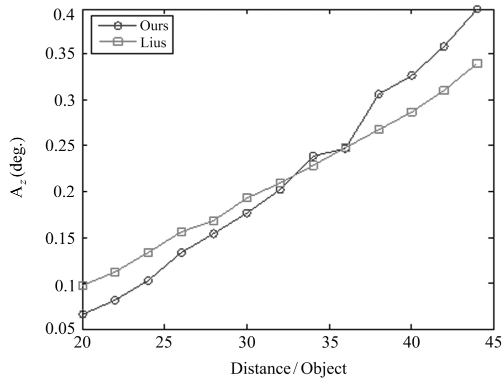
In Fig. 6, the horizontal coordinate is the ratio of the distance between the optical center and the model to the model size. The vertical coordinate is second. Experimental results demonstrate that the runtime of our algorithm is a little shorter than that of the algorithm in [11]. Therefore, our algorithm has good real-time performance.



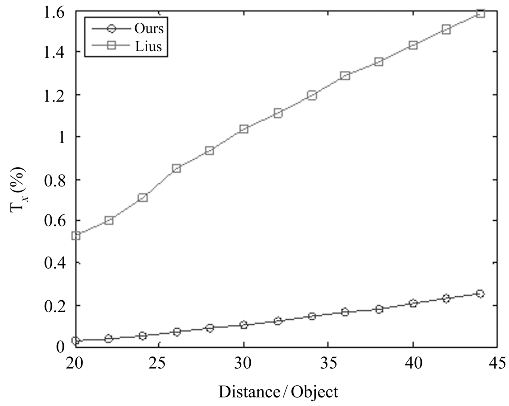
(a) Orientation error in x axis



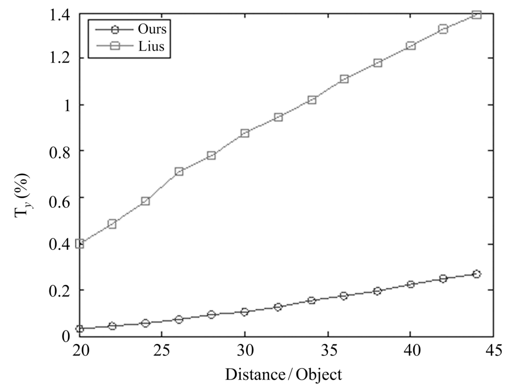
(b) Orientation error in y axis



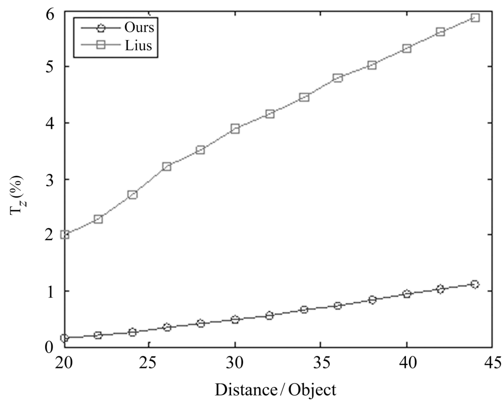
(c) Orientation error in z axis



(d) Position error in x axis



(e) Position error in y axis



(f) Position error in z axis

Fig. 5 Comparison results of errors

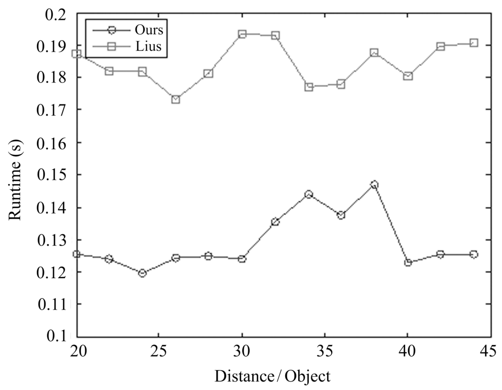


Fig. 6 Comparison results of runtimes

5 Conclusion

We discuss a new method to locate objects for the particular configuration of three non-coplanar lines, which intersect at two points. This configuration has some particular characteristics that three random lines do not have. Thus, its mathematical expression is easy to solve if we use this particular line configuration to locate objects. We give a detailed description of this new algorithm. Multi-solution phenomenon is the important problem that limits the application of pose estimation algorithm. This paper uses the geometric information of model to obtain the unique solution when the model moves within the view angle of camera. This method does not need to compute all the solutions to find the unique answer. Thus, it is time-saving and can improve the real-time performance of pose estimation system. We introduced a new iterative solution method based on the geometry and step acceleration method. Experimental results show that the iterative method works fast. What is more, it can avoid from being trapped into local minima and ensure the algorithm to find the right solution. At last, the simulation results show that our method has high pose estimation accuracy and good real-time characteristic. It can be applied to practical engineering applications.

References

1 Hu Zhan-Yi, Lei Cheng, Wu Fu-Chao. A short note on P4P problem. *Acta Automatica Sinica*, 2001, **27**(6): 770−776 (in Chinese)

2 Fang Shuai, Cao Yang, Xu Xin-He. A new vision location algorithm for uncalibrated camera. *Chinese Journal of Scientific Instrument*, 2005, **26**(8): 845−848 (in Chinese)

3 Huang Feng-San, Liu Shu-Gui, Peng Kai. Distance tracking 3D coordinates vision measuring system. *Opto-Electronic Engineering* , 2006, **33**(2): 107−110 (in Chinese)

4 Hao Ying-Ming, Wu Qing-Xiao, Zhou Chuan, Li Shuo, Zhu Feng. Technique and implementation of underwater vehicle station keeping based on monocular vision. *Robot*, 2006, **28**(6): 656−661 (in Chinese)

5 Zhou Chuan, Hao Ying-Ming, Wu Qing-Xiao, Li Suo, Zhu Feng. Visual station keeping based on constrained motion for underwater robot. *Chinese Journal of Scientific Instrument*, 2006, **27**(z3): 1840−1843 (in Chinese)

6 Lin H T. Autonomous Recovery of Exterior Orientation of Imagery Using Free-form Linear Features [Ph. D. dissertation], Ohio State University, 2002

7 Dhome M, Richetic M, Lapreste J T, Rives G. Determination of the attitude of 3-D objects from a single perspective view. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1989, **11**(12): 1265−1278

8 Chen H. Pose determination from line-to-plane correspondences: existence condition and closed-form solutions. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1991, **13**(6): 530−541

9 Horaud R P, Conio B, Leboullcux O, Lacolle B. An analytic solution for the perspective 4-point problem. *Computer Vision, Graphics, and Image Processing*, 1989, **47**(1): 33−44

10 Yuan J S C. A general photogrammetric method for determining object position and orientation. *IEEE Transactions on Robotics and Automation*, 1989, **5**(2): 129−142

11 Liu Y C, Huang T S, Faugeras O D. Determination of camera location from 2-D to 3-D line and point correspondences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1990, **12**(1): 28−37

12 Christy S, Horaud R. Iterative pose computation from line correspondences. *Computer Vision and Image Understanding*, 1999, **73**(1): 137−144

13 Gao X S, Hou X R, Tang J L, Cheng H F. Complete solutions classification for the perspective-three-point problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2003, **25**(8): 930−943

14 Zhou Xin, Zhu Feng. A note on unique solution conditions of the p3p proble. *Chinese Journal of Computers*, 2003, **26**(12): 1696−1701 (in Chinese)

15 Tang Jian-Liang. On the number of solutions for the P4P problem. *Journal of Mathematics*, 2006, **26**(2): 137−141 (in Chinese)

QIN Li-Juan Ph. D. candidate at Shenyang Institute of Automation, Chinese Academy of Sciences. Her research interest covers computer vision and virtual reality. Corresponding author of this paper. E-mail: qinlijuan@sia.ac.cn

ZHU Feng Professor at Shenyang Institute of Automation, Chinese Academy of Sciences. His research interest covers computer vision, image processing, and virtual reality. E-mail: fzhu@sia.ac.cn