An Algorithm of Determining the Plane Based on Monocular Vision and Laser Loop

Xinglong Zhu, Ying Zhang, Luyang Li, Longqin Gao, and Jiping Zhou*

School of Mechanical Engineering, Yangzhou University,
Huayang West Road 196, 225127, Yangzhou, Jiangsu, China
{xlzhu,yzhang,lyli,lqgao,jpzhou}@yzu.edu.cn

Abstract. Since the spot is elliptical when the cylindrical laser irradiates on the spatial plane, the pose of the spatial plane can be described by the elliptical spot. At the same time, the laser spot in the CCD plane is also elliptical. The information of the spatial plane will include in the image ellipse. The monocular vision has been established, and the boundary equation is obtained by image processing, and the relationship is derived between the boundary equation and the pose parameters of the spatial plane by minimum mean-square method. In order to obtain the depth information of the spatial plane, the boundary equation of the cylinder laser is introduced as the constrained condition. Because the constrained condition is transcendental equation set which includes trigonometric function, SWIFT (sequential weight increasing factor technique) is adopted for solving the parameters of the spatial plane. The simulation results show that the algorithm proposed is effective and feasible.

Keywords: monocular vision, cylinder laser spot, searching algorithm, position and pose.

1 Introduction

The stereo vision measurement technology based on the triangulation measurement principle has some special advantages, such as non-contact, fast, good flexibility, medium precision and so on. So it is widely used in online testing of modern manufacturing, contour measurement of three-dimensional objects, etc [1].

Pose determination of parameterized object models plays an important role in verification of model based recognition systems and real-time tracking of objects in images. Described herein is a data structure that models 3D parametric objects. Algorithms are presented for obtaining analytical partial derivatives of distance functions from projected model edges and endpoints to corresponding image segments, with respect to changes in model and camera parameters. These parameters include rotational, translational, scale and dilation in camera and object model. Solving for camera and model parameters in a 2D image (pose determination) is a nonlinear least squares paradigm. Weights are considered for line-to-line and point-to-point matching

* Xinglong Zhu: Prof., Ying Zhang: Master, Luyang Li: Associate Prof., Longqin Gao: Associate Prof., Jiping Zhou: Prof.

H. Deng et al. (Eds.): AICI 2011, Part III, LNAI 7004, pp. 1–11, 2011.

to allow for the applicability of the methods to images with noisy data [2]. The orientation disparity field from two orthographic views of an inclined planar surface patch (covered by straight lines) is analyzed, and a new tool to extract the patch orientation is provided. That is, the function coupling the average orientation of each pair of corresponding surface contours with their orientation disparity. This function allows identifying the tilt of the surface, and two indeterminacy functions describing the set of surface inclinations (around the vertical and horizontal axes) over convergence angle values compatible with the orientation disparity field [3]. K Achour presents a new approach for 3D scene reconstruction based on projective geometry without camera calibration. Previous works use at least six points to build two projective reference planes. Their contribution is to reduce the number of reference points to four by exploiting some geometrical shapes contained in the scene. The first implemented algorithm allows the reconstruction of a fourth point on each reference plane. The second algorithm is devoted to the 3D reconstruction [4]. Perceived depth was measured for three-types of stereogram with the color/texture of half-occluded (monocular) regions either similar to or dissimilar to that of binocular regions or background. In a two-panel random dot stereogram the monocular region was filled with texture either similar or different to the far panel or left blank. In unpaired background stereogram the monocular region either matched the background or was different in color or texture and in phantom stereogram the monocular region matched the partially occluded object or was a different color or texture. In all three cases depth was considerably impaired when the monocular texture did not match either the background or the more distant surface. The content and context of monocular regions as well as their position are important in determining their role as occlusion cues and thus in three-dimensional layout [5]. A model-based method for indoor mobile robot localization is presented herein; this method relies on monocular vision and uses straight-line correspondences. A classical four-step approach has been adopted (i.e. image acquisition, image feature extraction, image and model feature matching, and camera pose computing). These four steps will be discussed with special focus placed on the critical matching problem. An efficient and simple method for searching image and model feature correspondences, which has been designed for indoor mobile robot self-location, will be highlighted: this is a three-stage method based on the interpretation tree search approach. During the first stage, the correspondences pace is reduced by virtue of splitting the navigable space into view-invariant regions. In making use of the specificity of the mobile robotics frame of reference, the global interpretation tree is divided into two sub-trees; two low-order geometric constraints are then defined and applied directly on 2D–3D correspondences in order to improve pruning and search efficiency. During the last stage, the pose is calculated for each matching hypothesis and the best one is selected according to a defined error function [6]. Li-Juan QIN etc presented a method that can 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, they 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, they propose a method to deal with multi-solution phenomenon and a

new iterative method [7]. The present state-of-the-art in computing the error statistics in three-dimensional reconstruction from video concentrates on estimating the error covariance. A different source of error which has not received much attention is the fact that the reconstruction estimates are often significantly statistically biased. In this paper, they derive a precise expression for the bias in the depth estimate, based on the continuous (differentiable) version of structure from motion (SfM). Many SfM algorithms, or certain portions of them, can be posed in a linear least-squares (LS) frame-work. Examples include initialization procedures for bundle adjustment or algorithms that alternately estimate depth and camera motion. It is a well-known fact that the LS estimate is biased if the system matrix is noisy. In SfM, the matrix contains point correspondences, which are always difficult to obtain precisely; thus, it is expected that the structure and motion estimates in such a formulation of the problem would be biased. Existing results on the minimum achievable variance of the SfM estimator are extended by deriving a generalized Cramer–Rao lower bound [8]. S Frintrop and P Jensfelt are centered around landmark detection, tracking, and matching for visual simultaneous localization and mapping using a monocular vision system with active gaze control. They present a system that specializes in creating and maintaining a sparse set of landmarks based on a biologically motivated feature-selection strategy. A visual attention system detects salient features that are highly discriminative and ideal candidates for visual landmarks that are easy to redetect. Features are tracked over several frames to determine stable landmarks and to estimate their 3-D position in the environment. Matching of current land-marks to database entries enables loop closing. Active gaze control allows us to overcome some of the limitations of using a monocular vision system with a relatively small field of view. It supports 1) the tracking of landmarks that enable a better pose estimation, 2) the exploration of regions without landmarks to obtain a better distribution of landmarks in the environment, and 3) the active redetection of landmarks to enable loop closing in situations in which a fixed camera fails to close the loop [9]. Xinglong Zhu presents a quickly matching method, which enables to solve the corresponding points about the same point on the objective matter in the different CCD images by adding the laser on the cameras. The deepness message can be conveniently obtained by laser token [10].

In this paper, as the application of industrial robot vision system in the background, a new spatial plane determining algorithm is proposed. This paper is divided into the following sections: (1) Introduction of monocular vision model. (2) The spacial plane is described by ellipse; (3) The algorithm design is introduced. Finally, the numerical example verifies the feasibility of the method and it gives a summary of the study.

2 Monocular Vision Model

When the actor is performing in the stage, the spotlight is turned on. When the stage is irradiated by the spotlight, the spot shape is elliptical. With movement, the elliptical shape is changing. It shows the relative between the stage and the spotlight. Therefore, if the spotlight is fixed, and the pose of the stage plane is changed, the spot shape will take place change. When laser loop irradiates on the spacial plane, a few of shape will appear on the plane irradiated, sees Fig.1.

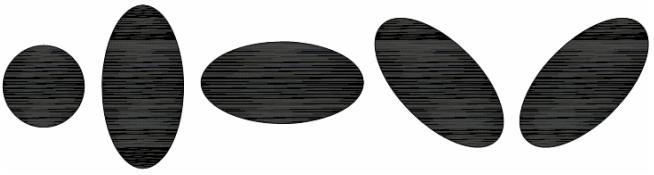

Fig. 1. The spot shape

The model, which uses monocular vision system to determine the pose of spatial plane, is shown in Fig.2. In this model, the relationship between the camera coordinate system and the world coordinate system and the camera's intrinsic parameters are known, the pinhole camera model is a useful model that enables simple mathematical formulation. The axis of cylindrical light source is parallel to cameras' optical axis.

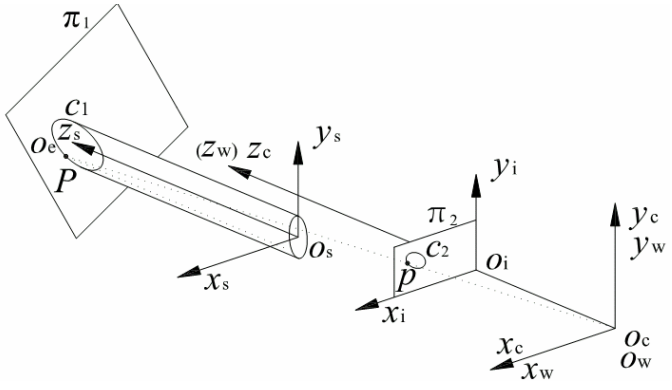

Fig. 2. Monocular vision system model

Define the camera coordinate system $o_c x_c y_c z_c$, the light source system $o_s x_s y_s z_s$, the imaging coordinate system $o_i x_i y_i z_i$ and its imaging plane π_2. Among them, z_c axis coincides with the camera optical axis, o_i is located in the camera optical axis. In general, it takes the camera coordinate system origin o_c as the absolute origin. So, the imaging coordinate system origin o_i is $[0 \quad 0 \quad f]^T$, f is the focal length. The center o_s coordinates of the cylindrical light source relative to camera coordinate system are $\begin{bmatrix} e_x & e_y & 0 \end{bmatrix}^T$. e_x and e_y are the relative offsets of the center of the cylindrical light source in *x* and *y* direction.

When the spatial plane π_1 is irradiated by cylindrical laser, the spot shape is elliptical. Its contours is expressed by c_1. Long half-axis length is *R*, short half-axis

length is *r*, the center o_e coordinates is $\begin{bmatrix} e_x & e_y & z_L \end{bmatrix}^T$, z_L is the depth from ellipse center to the right camera coordinate system origin.

3 Spacial Plane Described with Ellipse

The spatial plane π_1 is described as follows,

$$\frac{x}{a}+\frac{y}{b}+\frac{z}{c}=1. \tag{1}$$

In (1), *a*, *b*, and *c* are intercept at the x_c axis, y_c axis, and z_c axis, respectively.

From (1), determining the spatial plane need obtain three unknown parameters *a*, *b*, and *c*. When the spatial plane π_1 is irradiated by cylindrical laser, the spot shape is elliptical, its contour is denoted by c_1 curve, and its center is o_e. The spatial plane described is convenient by elliptical parameters, so we select laser spot to denote the pose of the spatial plane.

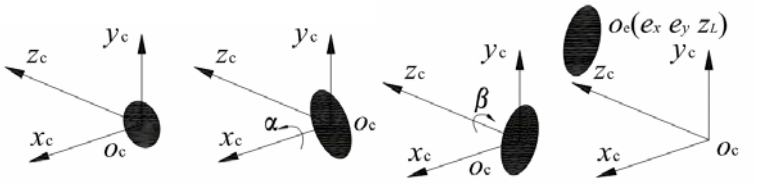

Fig. 3. The formation process of elliptical spot

The formation process of elliptical spot in spatial plane π_1: elliptical spot rotates α degrees around x_c axis firstly, then β degrees around z_c axis and finally ellipse center will pan $\begin{bmatrix} e_x & e_y & z_L \end{bmatrix}^T$, see Fig.3.

Therefore, the problem which determines the parameters of *a*, *b* and *c* may be transform into solving the elliptical pose and depth information parameters of α, β and z_L. That is,

$$(a,b,c)\rightarrow(\alpha,\beta,z_L). \tag{2}$$

In (2), elliptical pose α, β can be characterized on the CCD plane.

Select an arbitrary point *P* on the edge of ellipse, after rotation and translation, the coordinates are as follows:

$$\begin{bmatrix} x_e \\ y_e \\ z_e \end{bmatrix}=\begin{bmatrix} e_x+R\mathrm{c}\theta\mathrm{c}\beta-r\mathrm{s}\theta\mathrm{s}\beta\mathrm{c}\alpha \\ e_y+R\mathrm{c}\theta\mathrm{s}\beta+r\mathrm{s}\theta\mathrm{c}\beta\mathrm{c}\alpha \\ z_L+r\mathrm{s}\theta\mathrm{s}\alpha \end{bmatrix}. \tag{3}$$

Where $[x_e \quad y_e \quad z_e]^{\mathrm{T}}$ express an arbitrary point P's coordinates of the ellipse edge relative to the camera coordinate system. c , s denotes cos and sin respectively. The value of θ is $0-2\pi$. This shows that the elliptic laser spot edge can be expressed by (3).

The spatial straight line equation is constructed by point p and o_{c}:

$$\frac{x-x_e}{x_e-0}=\frac{y-y_e}{y_e-0}=\frac{z-z_e}{z_e-0}. \tag{4}$$

Let us denote $z=f$ and an arbitrary point's coordinates of imaging edge curve c_2 are obtained from (4):

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix}=\begin{bmatrix} x_e f / z_e \\ y_e f / z_e \\ f \end{bmatrix}. \tag{5}$$

Then, the trigonometric functions about θ can be derived by (3) and (5).

$$\cos\theta=\frac{a_1x_i+a_2y_i+a_3}{R(a_7x_i+a_8y_i+a_9)}$$

$$\sin\theta=\frac{a_4x_i+a_5y_i+a_6}{r(a_7x_i+a_8y_i+a_9)}$$

$$a_1=z_L\cos\beta\cos\alpha-e_x\sin\alpha \quad a_2=z_L\sin\beta\cos\alpha+e_y\sin\alpha$$

$$a_3=-f(e_x\cos\beta+e_y\sin\beta)\cos\alpha$$

$$a_4=-z_L\sin\beta$$

$$a_5=z_L\cos\beta$$

$$a_6=f(e_x\sin\beta-e_y\cos\beta)$$

$$a_7=\sin\beta\sin\alpha$$

$$a_8=-\cos\beta\sin\alpha$$

$$a_9=f\cos\alpha$$

From $\sin^2\theta+\cos^2\theta=1$, two variables and 2nd order equation can be obtained:

$$b_0x_i^2+b_1x_iy_i+b_2y_i^2+b_3x_i+b_4y_i+b_5=0. \tag{6}$$

$$b_0=a_1^2r^2+a_4^2R^2-a_7^2R^2r^2$$

$$b_1=2(a_1a_2r^2+a_4a_5R^2-a_7a_8R^2r^2)$$

$$b_2 = a_2^2 r^2 + a_5^2 R^2 - R^2 r^2 a_8^2$$
$$b_3 = 2(a_1 a_3 r^2 + a_4 a_6 R^2 - a_7 a_9 R^2 r^2)$$
$$b_4 = 2(a_2 a_3 r^2 + a_5 a_6 R^2 - a_8 a_9 R^2 r^2)$$
$$b_5 = a_3^2 r^2 + a_6^2 R^2 - a_9^2 R^2 r^2$$

Equation (6) is the ellipse equation.

From (3) again, we can obtain

$$x_e = e_x + R\mathrm{c}\theta\mathrm{c}\beta - r\mathrm{s}\theta\mathrm{s}\beta\mathrm{c}\alpha . \tag{7}$$

$$y_e = e_y + R\mathrm{c}\theta\mathrm{s}\beta + r\mathrm{s}\theta\mathrm{c}\beta\mathrm{c}\alpha . \tag{8}$$

Form (7) and (8),

$$\cos\theta = \frac{(x_e - e_x)\cos\beta + (y_e - e_y)\sin\beta}{R}$$
$$\sin\theta = \frac{(y_e - e_y)\cos\beta - (x_e - e_x)\sin\beta}{r\cos\alpha}$$

Then,

$$d_0 x_e^2 + d_1 x_e y_e + d_2 y_e^2 + d_3 x_e + d_4 y_e + d_5 = 0 . \tag{9}$$

In (9),

$$d_0 = R^2 \sin^2\beta + r^2 \cos^2\alpha \cos^2\beta$$
$$d_1 = (r^2 \cos^2\alpha - R^2)\sin 2\beta$$
$$d_2 = R^2 \cos^2\beta + r^2 \cos^2\alpha \sin^2\beta$$
$$d_3 = [e_y(R^2 - r^2\cos^2\alpha)\sin 2\beta - 2e_x(R^2 \sin^2\beta + r^2\cos^2\alpha\cos^2\beta)]$$
$$d_4 = [e_x(R^2 - r^2\cos^2\alpha)\sin 2\beta - 2e_y(R^2 \cos^2\beta + r^2\cos^2\alpha\sin^2\beta)]$$
$$d_5 = (r^2\cos^2\alpha\cos^2\beta + R^2\sin^2\beta)e_x^2 + (r^2\cos^2\alpha\sin^2\beta + R^2\cos^2\beta)e_y^2 + (r^2\cos^2\alpha - R^2)e_x e_y \sin 2\beta - R^2 r^2 \cos^2\alpha$$

Substituting (7) and (8) to (10),

$$(x - e_x)^2 + (y - e_y)^2 = r_1^2 . \tag{10}$$

In (10), r_1 is radius of the laser cylinder, it equals *r*.

In comparison with (9) and (10), we may obtain the following equations,

$$d_1/d_0 = 0 . \tag{11}$$

$$d_2/d_0 = 1. \tag{12}$$

$$d_3/d_0 = -2e_x. \tag{13}$$

$$d_4/d_0 = -2e_y. \tag{14}$$

$$d_5/d_0 = -r^2. \tag{15}$$

4 Algorithm Design

Figure 4 shows that the spot shape is elliptical when the spatial plane is irradiated by cylindrical laser, which is equivalent to the formation: the plane $o_c x_c y_c$ rotates 60° around the x-axis firstly, then rotates 30° around the z-axis and finally pans $[30 \quad 40 \quad 75]^T$ mm. Take the camera coordinate system origin o_c as the absolute origin. The origin o_i coordinates of the imaging coordinate system are $[0 \quad 0 \quad 25]^T$ mm. The center o_s coordinates of the cylindrical laser relative to camera coordinate system are $[30 \quad 40 \quad 0]^T$ mm. The ellipse center o_e coordinates are $[30 \quad 40 \quad 75]^T$ mm. The radius of cylindrical laser is 10 mm. Figure 5 shows the images in the imaging plane.

Fig. 4. Simulative spot

Fig. 5. Images in imaging plane

According to the above known parameters and (3)-(5), we can obtain the data of image edge on CCD plane. The data shows in table 1.

Table 1. CCD Imaging Ellipse fitting

θ/°	X_i/mm	Y_i/mm	θ/°	X_i/mm	Y_i/mm
0	12.8868	15.0000	180	7.1132	11.6667
18	11.4158	14.7573	198	8.3666	11.6905
36	9.9986	14.4209	216	10.0019	11.9041
54	8.7194	14.0276	234	11.8691	12.3201
72	7.6309	13.6055	252	13.7026	12.9079
90	6.7699	13.1770	270	15.1700	13.5836
108	6.1681	12.7610	288	15.9889	14.2279
126	5.8600	12.3767	306	16.0424	14.7295
144	5.8859	12.0465	324	15.4064	15.0244
162	6.2906	11.7983	342	14.2795	15.1043

Assuming that elliptical equation on CCD plane is as follows,

$$x^2 + B_1 xy + B_2 y^2 + B_3 x + B_4 y + B_5 = 0 . \tag{16}$$

According to the data in table 1, we can determine the elliptical parameters on CCD plane by minimum mean-square elliptical algorithm. After fitting process, the elliptical parameters on CCD plane shows in table 2.

Table 2. Elliptical Parameters on Ccd Plane

B_1	B_2	B_3	B_4	B_5
-3.9899	8.8868	31.4459	-194.0465	1111.1111

Therefore, $b_i / b_0 = B_i$, (i=1,2,3,4,5), that is, the following of five equal equation can be obtained,

$$b_1 / b_0 = -3.9899 . \tag{17}$$

$$b_2 / b_0 = 8.8868 . \tag{18}$$

$$b_3 / b_0 = 31.4459 . \tag{19}$$

$$b_4 / b_0 = -194.0465 . \tag{20}$$

$$b_5 / b_0 = 1111.1111 . \tag{21}$$

Through (11)-(15) and (17)-(21), we can solving the unknown parameters R, z_L, α and β. Because (11)-(15) and (17)-(21) are equation set which include trigonometric function, it is difficult to solve the R, z_L, α and β by analytic

solution. Therefore, we adopt a searching method to solve the above problem. The original problem transforms to solve the minimum question,

$$\min J = 0 . \tag{22}$$

s.t.

$$h_i = b_i / b_0 - B_i . \quad (i=1,2,3,4,5) \tag{23}$$

That is, it must satisfy (17)-(21).And at the same time, it must satisfy (11)-(15), that is,

$$h[6] = d_1 / d_0 . \tag{24}$$

$$h[7] = d_2 / d_0 - 1 . \tag{25}$$

$$h[8] = d_3 / d_0 + 2e_x . \tag{26}$$

$$h[9] = d_4 / d_0 + 2e_y . \tag{27}$$

$$h[10] = d_5 / d_0 + r^2 . \tag{28}$$

The above solving the minimum problem will translate into the following penalty function question,

$$P(\Theta, r_k) = \min J + r_w \sum_{i=1}^{10} h_i^2 . \tag{29}$$

In (29), r_w is the penalty factor. $\Theta = [R \quad z_L \quad \alpha \quad \beta]$ is optimize variable.

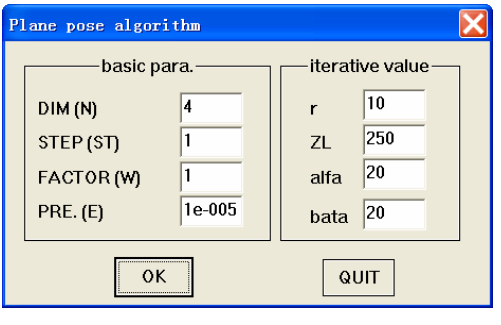

Fig. 6. Dialogue box

Fig. 7. The searching result

Figure 6 is the dialogue box. The initial iterative values $\Theta = [R \quad z_L \quad \alpha \quad \beta] = [10\text{mm} \quad 250\text{mm} \quad 20^\circ \quad 20^\circ]$. Figure 7 is the searching results.

$\Theta^* = [R^* \ z_L^* \ \alpha^* \ \beta^*] = [20\text{mm} \ 75\text{mm} \ 60.000001^\circ \ 30.00001^\circ]$. The searching result is next to the known values. It test that the algorithm proposed is feasible.

5 Conclusions

An algorithm that can determine the depth and pose information of the spatial plane is proposed. The algorithm adopts the monocular vision image in CCD plane and the boundary equation of the cylinder laser spot. Because of only the monocular vision, the vision model is simpler, the arithmetic speed is faster, and the system cost is cheaper. Because the boundary equation of the cylinder laser is introduced, the algorithm can obtain the depth information, as well as the pose parameters of the spatial plane. The simulation results show that the method is effective and feasible.

Acknowledgment. This work is supported by National Science Foundation of China under Project 60977071, by Science and Technology Department of Jiangsu Province under Project BK2010323, by "333 Talents Engineering" of Jiangsu Province under Project BRA2010131 and by Innovation Fund for Small and Medium-Sized Technology-Based Firms of Jiangsu Province under Project BC2009242.

References

1. Yamada, H., Togasaki, T., Kimura, M., et al.: High density 3-D packaging technology based on the sidewall interconnection method and its application for CCD micro camera visual inspection system. IEEE Transactions on Advanced Packaging 26(2), 113–121 (2003)
2. Goldberg, R.R.: Pose determination of parameterized object models from a monocular image. Image and Vision Computing 11(1), 49–62 (1993)
3. Fantoni, C.: 3D surface orientation based on a novel representation of the orientation disparity field. Vision Research 48(25), 2509–2522 (2008)
4. Achour, K., Benkhelif, M.: A new approach to 3D reconstruction without camera calibration. Pattern Recognition 34(12), 2467–2476 (2001)
5. Grovea, P.M., Gillamb, B., Onoa, H.: Content and context of monocular regions determine perceived depth in random dot, unpaired background and phantom stereograms. Vision Research 42(15), 1859–1870 (2002)
6. Aider, O.A., Hoppenot, P., Colle, E.: A model-based method for indoor mobile robot localization using monocular vision and straight-line correspondences. Robotics and Autonomous Systems 52(2,3), 229–246 (2005)
7. Qin, L.-J., Zhu, F.: A New Method for Pose Estimation from Line Correspondences. Acta Automatic Sinica 34(2), 130–134 (2008)
8. Roy-Chowdhury, A.K., Chellappa, R.: Statistical Bias in 3-D Reconstruction From a Monocular Video. IEEE Transactions on Image Processing 14(8), 1057–1062 (2005)
9. Frintrop, S., Jensfelt, P.: Attentional Landmarks and Active Gaze Control for Visual SLAM. IEEE Transactions on Robotics 24(5), 1054–1064 (2008)
10. Zhu, X.: Quick measurement algorithms of position and pose of spatial object by CCD images based on stereo vision and laser token. Chinese Journal of Mechanical Engineering 40(7), 161–165 (2004)