1-Point RANSAC Application: an Independent Positioning Method for Crawler Mobile Robots Based on Binocular Stereo Vision

Xixuan Zhao, Jiangming Kan *, Chuandong Zhan

Abstract—Independent positioning is a basic problem for realizing the autonomous navigation of a mobile robot in an unknown environment. In this paper, an independent positioning method based on binocular stereo vision is examined. This method includes two steps: extracting and matching point features based on binocular stereo vision and estimating the motion model of the crawler mobile robot. In the first step, point features are detected using a FAST algorithm and matched according to the Hamming distance. In the second step, the motion model of the crawler mobile robot is estimated by points that are matched using the visual odometry method which is based on 1-Point RANSAC (random sample consensus) framework and Kalman filter. The results of indoor and outdoor practical experiments illustrated that the proposed method can achieve the goal of real-time independent positioning and obtain accurate and robust results.

Keywords—Independent positioning, 1-Point RANSAC, Kalman Filter, crawler mobile robot, binocular stereo vision

I. INTRODUCTION

In the field of robot independent positioning, the visual odometer method is increasingly utilized. The most successful application of visual odometer technology is in the "Spirit" Mars probe and "Opportunity" Mars probe launched by NASA. Visual odometers can estimate the moving distances of a robot from a continuous image sequence, regardless of the working condition and the mode of motion. [1]

Visual odometer approaches are primarily feature-based methods. The image sequence is first acquired by camera; then, the six degrees of freedom (including the position (x, y, z) and the direction (roll, pitch, yaw) of the robot) are obtained through the process of feature extraction, feature matching, characteristic tracking and motion estimation. The final step involves navigation and positioning task. [2, 3]

Nister et al. designed a real-time visual odometer system. They detected angular points using the Harris algorithm and used some related optimization techniques such as MMX and non-maxima suppression to select feature points. By setting saturated values in local areas, the number of feature points was limited. They put forward a 5-point RANSAC and a 3-point RANSAC algorithm to estimate motion. Parallax and consistency constraints were used to reduce false matching points. Nister et al. conducted outdoor experiments for visual odometer systems in forestry, alley and meadow environments. Experimental results illustrated that it performed well compared to the differential global positioning system (DGPS). They also compared the positioning results of monocular and binocular cameras and came to the conclusion that more reliable and accurate positioning could be obtained through use of binocular cameras. [4,5]

Zhu et al. improved Nister's system by addressing sensitivity to outliers and the accumulation of errors for long distance measurements. [6] They adopted 3D-related technology to reduce 3D reconstruction errors and minimize drift error through the removal of outliers. Then, they combined visual odometers with IMU and GPS on the basis of the Kalman filter algorithm. This lead to a greater precision and robustness in both indoor and outdoor environments.

Wu et al. put forward a binocular visual odometer algorithm based on parallax space. According to the scale and rotation invariance of SIFT feature points, the accurate matching of feature points was realized in the left and right images and characteristic tracking in two consecutive frames. Under the RANSAC framework, motion estimation used the matched points to acquire the initial motion parameters. The parallax ratio of the matching points was then updated. The isotropy of the distribution of noise in the parallax space was then used to estimate the motion needed to overcome the defect of the non-uniform noise distribution in 3D space, which cannot be avoided through traditional algorithms. [7]

Xu et al. put forward a 2-point RANSAC algorithm that was combined with information regarding the EKF motion model. The RANSAC algorithm of two extraction points was used to remove outliers. A monocular vision SLAM experiment based on a 2-point RANSAC algorithm was conducted on a miniature unmanned helicopter platform. Experimental results showed that the 2-point RANSAC algorithm was reliable and that the position estimation precision of SLAM can meet the need of autonomous flight. [8]

Although the independent positioning methods mentioned above prove highly accurate, the numerous, time-consuming calculations required prove problematic and do not meet the

This work is supported by the National Natural Science Foundation of China (Grant No. 31570713) and the Fundamental Research Funds for the Central Universities (Grant No. TD2013-4)

Xixuan Zhao, Jiangming Kan are with the Beijing Forestry University, Beijing 100083, China. (corresponding author phone: 010-62336137-706; e-mail: kanjm@bjfu.edu.cn)..

demands of real-time processing. Thus, we put forward an independent positioning method based on binocular stereo vision, which combines a 1-point RANSAC framework to a Kalman filter. This allows full play of the prediction effect of the Kalman filter. Advantages include high accuracy and the fewer calculations required of 1-point RANSAC algorithms. The independent positioning can be obtained more quickly.

In this paper, considering that the most common type of mobile robot is the crawler mobile robot, we used our algorithm on the Voyager-II crawler mobile robot to perform the abovementioned verification tests.

II. MOTION MODEL

When considering the independent positioning of robots, we should first analyze the motion feature of the crawler mobile robots. Fig.1(a) shows the practical Voyager-II crawler mobile robot used in our research. The motion diagram of a crawler mobile robot in a global reference system $X_I O Y_I$ is shown in Fig.1(b). For the robot and its movement, we could reasonably give the following hypotheses: the tracks on both sides have the same grounding length and width; the attachment of the center of the two tracks is vertical to the direction of motion; the robot has a rigid shell; and the tracks are well-contacted with the ground. [9]

Based on the assumptions above, the movement of the crawler mobile robot in the horizontal plane can be treated as a rotating motion around a certain point—the instantaneous center. Assuming the projection coordinate of the center of mass in the global reference system is (x_c, y_c) , θ is the attitude angle (the included angle of the X_R axis in local reference frame $X_R CY_R$ of the robot and the X_I axis in global reference frame $X_I OY_I$). Then, the position of the robot can be expressed by the following vectors:

$$\mathbf{x}_{C} = \begin{bmatrix} x_{C} & y_{C} & \theta \end{bmatrix}^{T}$$
(1)



Fig.1 (a): Voyager-II tracked mobile robot (b)Robot locomotion under the global reference frame

Our Voyager-II crawler mobile robot is differential driven. The position can be estimated by calculating the integral of movement (the sum of the increment of the walking distance). For discrete systems with fixed sampling interval Δt , the increment of the walking distance is:

$$\Delta x = \Delta u \cdot \cos(\theta + \Delta \theta / 2) \tag{2}$$

$$\Delta y = \Delta u \cdot \sin(\theta + \Delta \theta / 2) \tag{3}$$

$$\Delta \theta = \frac{\Delta u_R - \Delta u_L}{D + 2b} \tag{4}$$

$$\Delta u = \frac{\Delta u_R + \Delta u_L}{2} \tag{5}$$

In the above equations, $(\Delta x, \Delta y, \Delta \theta)$ is the walking distance in the previous sampling interval; $\Delta u_L, \Delta u_R$ are the walking distances of the left and the right track, respectively; D is the distance between the two tracks of the Voyager-II crawler mobile robot; and b is the width of the tracks. Therefore, we can obtain the state equation of position estimation based on the odometer update:

$$\mathbf{x}_k = \mathbf{F}_k \mathbf{x}_{k-1} + \mathbf{B}_k \mathbf{u}_k \tag{6}$$

In the equation, the state vector \mathbf{X}_k is

$$\mathbf{x}_{k} = \begin{bmatrix} x_{k} & y_{k} & \theta_{k} \end{bmatrix}^{\mathrm{T}}$$
⁽⁷⁾

The control vector \mathbf{u}_k is

$$\mathbf{u}_{k} = \begin{bmatrix} \Delta u_{L,k} & \Delta u_{R,k} \end{bmatrix}^{\mathrm{I}}$$
⁽⁸⁾

The state transition matrix \mathbf{F}_k is

$$\mathbf{F}_{k} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
(9)

The control matrix \mathbf{B}_k is

$$B_{k} = \frac{1}{2} \begin{bmatrix} \cos(\theta_{k-1} + \varphi_{k}) & \cos(\theta_{k-1} + \varphi_{k}) \\ \sin(\theta_{k-1} + \varphi_{k}) & \sin(\theta_{k-1} + \varphi_{k}) \\ -\frac{2}{D+2b} & \frac{2}{D+2b} \end{bmatrix}$$
(10)

In the equation,

$$\varphi_k = \frac{\Delta u_{R,k} - \Delta u_{L,k}}{2(D+2b)} \tag{11}$$

Parameter Estimation of Planar Motion Model with Nonholonomic Constraints

According to Section 2.1, any instant planar motion can be locally described as circular motion around a point, and linear motion can be regarded as infinite radius circular motion. This section describes the parameterized process of the essential matrix of circular motion, the effect of the rotation matrix and translation vector, and the parameter estimation of the nonholonomic constraint motion model.

In the planar circular motion, the camera's relative position can be described by deflection angle θ and polar coordinates (ρ, φ) , as shown in Fig.2.



Fig.2 Relative camera positions in circular motion

Axis *z* of the camera coordinate system is vertical with R_{ICR} , the radius of gyration. $\varphi = \theta / 2$, and with linear motion, it satisfies $\theta = 0, \varphi = 0, R_{ICR} = \infty$. In the camera coordinate system, through analysis, we can deduce that

$$\varphi = \tan^{-1} \left(\frac{x'y - xy'}{yz' + y'z} \right)$$
(12)

Given the expressions in the disparity space of $\mathbf{P} = [x, y, z]^{\mathsf{T}}$

and $\mathbf{P}' = [x', y', z']^{\mathsf{T}}$ as $\mathbf{w} = [u, v, d]^{\mathsf{T}}$ and $\mathbf{w}' = [u', v', d']^{\mathsf{T}}$, (u, v) and (u', v') are the pixel coordinates of projection points on the image, and d and d' are the corresponding disparity:

$$\begin{cases} x = \frac{B \cdot (u - u_0)}{d} \\ y = \frac{B \cdot (v - v_0)}{d}, \\ z = \frac{B \cdot f}{d} \end{cases} \begin{cases} x' = \frac{B \cdot (u' - u_0)}{d'} \\ y' = \frac{B \cdot (v' - v_0)}{d'} \\ z' = \frac{B \cdot f}{d'} \end{cases}$$
(13)

According to the Equations (12) and (13), we can obtain the expression of deflection angle θ :

$$\varphi = \tan^{-1} \left(\frac{(u' - u_0) \cdot (v - v_0) - (u - u_0) \cdot (v' - v_0)}{f \cdot (v + v' - 2v_0)} \right), \theta = 2\varphi$$
(14)

In the equation, (u_0, v_0) is the pixel coordinates of the main

points. According to equation (14), deflection angle θ is only related to the equivalent focal length of the camera, the pixel coordinates of the main points and the pixel coordinates of the matching points. Therefore, there is no need to calculate the three-dimensional coordinate of the matching points in the camera coordinate system. According to equation (14), we can obtain the value of the deflection angle θ . In addition, it is worth noting that when the disparity of the matching points is 0 and the depth infinite, equation (14) can be calculated as well.

Rotation matrix **R**, translation vector **T**, $\mathbf{P} = [x, y, z]^{\mathsf{T}}$ and

$$= [x', y', z']^{\mathrm{i}} \text{ satisfy } \mathbf{RP} + \mathbf{T} = \mathbf{P}', \text{ namely,}$$

$$\begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} - \rho \cdot \begin{bmatrix} \sin\phi \\ 0 \\ \cos\phi \end{bmatrix} = \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$
(15)

Rewrite equation (15) as

P'

$$\rho = \begin{bmatrix} -\sin\varphi & 0 & \cos\varphi \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} \sin\varphi & 0 & \cos\varphi \end{bmatrix} \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix}$$
(16)

According to equations (16), (14) and (13), we can obtain

$$\rho = \frac{B}{d} \cdot \left[-\sin\varphi \quad 0 \quad \cos\varphi \right] \begin{bmatrix} u - u_0 \\ v - v_0 \\ f \end{bmatrix} + \frac{B}{d'} \cdot \left[\sin\varphi \quad 0 \quad \cos\varphi \right] \begin{bmatrix} u' - u_0 \\ v' - v_0 \\ f \end{bmatrix}$$
(17)

By back substituting θ and (ρ, φ) , we can obtain the rotation matrix **R** and translation vector **T** namely, the planar motion model parameters of the crawler-type mobile robot.

III. INDEPENDENT POSITIONING

A. Visual odometry (VO) based on binocular stereo vision

Visual odometry is a method of motion estimation based on monocular or stereo vision input which can be used in robot independent positioning. In this paper, we estimated motion in two steps: extracting and matching feature points and estimating motion based on matching feature points. A key process is to maximize the inliers set and minimize the re-projection error of the inliers set. The RANSAC algorithm frame is commonly used in this process.

In this paper, we put forward a visual odometry method based on a 1-Point RANSAC frame to achieve independent positioning. According to the planar motion of the nonholonomic constraint, we established a restrictive camera motion model. After adopting priori-state information acquired by the Kalman filter in the prediction stage to estimate the parameters of constructing a 1-Point RANSAC framework, we integrated the 1-Point RANSAC algorithm into the Kalman filter framework. The basic procedure of the proposed VO algorithm is shown below in Fig.3.

1) Binocular Stereo Vision Processing

A basic process of binocular stereo vision processing is show below in Fig.4. Given the matching points and disparity, the three-dimensional information of the corresponding points in the camera reference coordinate system can be restored.



Fig.3 The basic procedure diagram of the proposed VO algorithm



(a)



Fig.4 (a) Stereo color images after distortion rectification (400*300 resolutions); (b) Results of edge detection; (c) disparity image

2) Prediction and Observation Matching of the Kalman Filter

The prediction of the Kalman Filter includes state prediction, estimated covariance prediction and observation prediction. Matching of points in two consecutive images is achieved in the observation prediction period.

3) Parameter Estimation Based on 1-Point RANSAC

The planar motion model of nonholonomic constraint, namely, the estimation of deflection angle θ and parameter ρ , is needed in the process of parameter estimation based on 1-Point RANSAC [10,11,12].

In the model validation phase of estimating the deflection angle θ in the RANSAC process, the process of calculating the re-projection error of the data is shown in Fig.5. According to Fig.5, all data satisfy the following equation:

$$d \cdot \tan \alpha = d' \cdot \tan \alpha', d \cdot \sin \gamma = d' \cdot \sin \gamma' \tag{18}$$

Rewrite Equation (18) as

$$\frac{d}{d'} = \frac{\tan \alpha'}{\tan \alpha}, \frac{d}{d'} = \frac{\sin \gamma'}{\sin \gamma}$$
(19)

$$error_{\theta} = \left| \frac{\tan \alpha'}{\tan \alpha} - \frac{\sin \gamma'}{\sin \gamma} \right|$$
(20)

If the value of $error_{\theta}$ is less than the given threshold th_yaw_angle , then the point is considered an inlier.

In the model validation phase of estimating the parameter ρ in the RANSAC process, the re-projection error of the data is

$$error_{\rho} = \left\| \mathbf{P}' - (\mathbf{R}\mathbf{P} + \mathbf{T}) \right\| = \left\| \begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} - \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \hat{\rho} \cdot \begin{bmatrix} \sin\varphi \\ 0 \\ \cos\varphi \end{bmatrix} \right\|$$
(21)

If the value of $error_{\theta}$ is less than threshold value th_polar_radius , then the data are considered inliers.

4) Update of Kalman Filter

The update of the Kalman filter includes the update of the state estimation and the update of the covariance matrix estimation.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The experiments include indoor and outdoor tests on the Voyager-II crawler mobile robot with a Bumblebee® 2 binocular stereo camera. Outdoor tests aim to test the robustness of the algorithm when encountering such circumstances as numerous extracted feature points, open local area and low parallax phenomenon in stereo matching.

A. Indoor Environment

1) Processing Time and Stability Test

In this experiment, image sequences of the same scene with different resolutions are utilized. The details of the average elapsed time of sub-procedures under different resolutions are shown in Table 1.

Table 1 Average elapsed time of sub-procedures under different resolutions

	Resolution	Binocular stereo vision processing (ms)	Parameter estimation (ms)	Kalman filtering (ms)
1	640*480	48.969	0.874	0.161
2	512*384	39.851	0.664	0.157
3	400*300	31.970	0.754	0.158
4	320*240	22.165	0.662	0.168
5	256*192	22.460	0.610	0.153

From Table 1, we find that binocular stereo vision processing is the most time-consuming (approximately 85.13%), while motion model parameter estimation and Kalman filtering processing account for 1.95% and 0.45%, respectively. Trivial tasks such as grabbing the image, image gray processing, smooth filtering, image distortion correction, feature point extraction and matching of preprocessing during the binocular stereo vision processing period contribute to this result. The average number of stereo matching features under different resolutions is shown in Fig.6. By adopting the 1-Point RANSAC framework, motion model parameter estimation consumes little time.



Fig.6 Average number of stereo matching features under different resolutions

Next, we conducted long-time motion experiments on resolution images with 10000 frames. We found that the average elapsed time is 26.5511 ms and that the standard deviation of elapsed time is 4.551 ms.

The results of this experiment show that the proposed method operates very quickly and can meet the speed demands of practical applications. When selecting an upper threshold limit of 64 for the number of matching points, the real-time processing speed can reach 30 frames per second.

2) Linear Motion experiments

Five sets of experiments were implemented. Each set consisted of 200 frames of forward motion and 200 frames of backward motion. The graphs of the trends in the x-direction displacement and y-direction displacement of one set are shown in Fig.7. The relative error is less than 5%.

For results calculated in the x-direction with an average error of -31.5 mm, the standard deviation is 22.6 mm. In the y-direction with an average error of -65.2 mm, the standard deviation is 16.0 mm. The accuracy of the linear motion test is relatively high, and the average absolute error is less than 10 cm.



Fig.7 Trend Graphs of x-direction and y-direction displacementFree Motion experiments

The estimation result of the key positions of the free motion experiment is shown in Fig.8. By calculating the obtained results, the x-direction average error is -1.5 mm, and the standard deviation is 47.4 mm. In the y-direction, the average error is 7.4 mm, and the standard deviation is 8.5 mm. The average error of the deflection angle is 0.05 deg, and the standard deviation is 0.56 deg.

INTERNATIONAL JOURNAL OF CIRCUITS, SYSTEMS AND SIGNAL PROCESSING



Fig.8 Experimental results of unconstrained planar motion test

B. Outdoor environment

1) Linear Motion experiments

In the practical linear motion experiment, we conducted five sets of tests. Each set consisted of 600 frames of forward motion and 600 frames of backward motion. The test result of the linear motion experiment is shown in Fig.9. By calculating the obtained results, the x-direction average error is 15.8 mm, and the standard deviation is 53.0 mm. In the y-direction, the average error is 124.5 mm, and the standard deviation is 40.3 mm.





Fig.9 Trend Graphs of the error in the x-direction and y-direction displacement

2) Spot-turn Motion experiments

Two sets of spot-turn motion experiments were performed: the length of the video sequence was 1800 frames in the first set and 2400 for the second. The first set of experiments was conducted on asphalt pavement. We chose a relatively open site—a parking lot. The second set of experiments was conducted on a cement floor, an artificial garden.

The trend graph of the deflection angle is shown in Fig.10. For the first set, the average error is 6.38 deg, and the standard deviation is 6.38 deg. In the second set, the average error is 4.09 deg, and the standard deviation is 8.17 deg.

In this experiment, the large deflection angle error is observed when the practical motion of the Voyager-II robot does not meet the hypothesis. The track on either side has a different grounding length, that is, the tracks of Voyager-II were not well contacted with the ground. Thus, the parameter estimation of the planar motion model is invalid and leads to a poor positioning result for the deflection angle.



Fig.10 Trend Graph of deflection angle of spot-turn motion experiment

3) Free Motion experiments

Three sets of free motion experiments were conducted. The length of the video sequence was 1403 frames in the first set of experiments, 2419 frames in the second set, and 2670 frames in the third set. The estimation results of key positions are shown in Fig.11 with corresponding statistical error results in Table 2.

Set	1			2		3			
	x	у	θ	x	у	θ	x	у	θ
	(mm)	(mm)	(deg)	(mm)	(mm)	(deg)	(mm)	(mm)	(deg)
AVG	47.1	268.4	2.32	-326.2	-1057.3	16.68	-25.4	-131.9	2.43
S.D.	36.4	287.7	2.57	818.1	950.6	5.32	135.8	115.1	3.68

Table 2 Error of unconstrained planar motion test



Fig.11 Experimental results of the unconstrained planar motion test

V. CONCLUSIONS

In this paper, we proposed an independent positioning method based on a visual odometer. According to the results of the indoor and outdoor experiments, our proposed method achieved real-time independent positioning. The positioning results were considerately accurate, reached the preconceived positioning task and were processed quickly.

REFERENCES

- Kostavelis I, Boukas E, Nalpantidis L, et al. Stereo-based Visual Odometry for Autonomous Robot Navigation[J]. International Journal of Advanced Robotic Systems, 2016.
- [2] Yujiao Ma. 'Research on the localization of mobile robot based on stereo vision odometer'. PhD thesis, Wuhan University, 2009
- Yubo Li, Xiaozhou Zhu, Huimin Lu, Hui Zhang. 'The overview of visual odometer technology'. Application Research Of Computers, 2012, 29, (8), pp.2801-2804
- [4] Nistér D, Naroditsky O, Bergen J. 'Visual odometry'. Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Washington, DC, USA, June 27 to July 2, 2004, pp. 652 ~ 659.
- [5] Nistér D, Naroditsky O, et al. 'Visual Odometry for Ground Vehicle Applications'. Journal of Field Robotics, 2006, 23, (1), pp. 3~20.
- [6] Zhu Z, Oskiper T, Naroditsky O, Samarasekera S, et al. 'An Improved Stereo-based Visual Odometry System'. Performance Metrics for Intelligent Systems Workshop with the IEEE Safety, Security, and Rescue Robotics Conference, Maryland, USA, August 21-23, 2006
- [7] Gongwei Wu, Wenhui Zhou, Weikang Gu. 'Binocular Vision Odometer Based on Disparity Space'. Journal of Transduction Technology, 2007, pp.1432~1436.
- [8] Weijie Xu, Ping Li, Bo Han. 'Monocular Visual SLAM of Unmanned Helicopter Based on 2-point RANSAC'. Robot, 2012, 34, (1), pp.65-71
- [9] De Xu, Wei Zou. 'The perception of the indoor service mobile robot, Positioning and control'. (Science Press, 2nd edn, 2008)
- [10] caramuzza D., Fraundorfer F., Siegwart R. 'Real-time monocular visual odometry for on-road vehicles with 1-point RANSAC'. International Conference on Robotics and Automation, Kobe Japan, 12-17 May 2009, pp. 4293-4299.
- [11] Civera J., Grasa O. G., Davison A. J., Montiel J. M. M.. 1-Point RANSAC for extended Kalman filtering: Application to real-time structure from motion and visual odometry[J]. Journal of Field Robotics, 2010, 27(5): 609–631.
- [12] Kostavelis I, Boukas E, Nalpantidis L, et al. Visual Odometry for autonomous robot navigation through efficient outlier rejection[C]// IEEE International Conference on Imaging Systems and Technique. 2013:45 - 50.

Xixuan Zhao born in 1992, received her Bachelor's degree from China Jiliang University, Hangzhou, China, at 2011~2015. Now she is pursuing her doctoral degree at Beijing Forestry University, Beijing, China. Her research interests include image processing and computer vision

Jiangming Kan born in 1976, Ph.D. degree, professor at Beijing Forestry University. His research interests include computer vision and intelligent control.

Chuandong Zhan born in 1990, received his master's degree from Beijing Forestry University. His research interests include independent positioning and map reconstruction.