

Feature Based Vision Destination Recognition for Mobile Robot

Xianguang Cui, Hai Zhang* and Yue Wang

BeiHang University, No.37 Xueyuan Road, HaiDian District, Beijing, China

*Corresponding author

Abstract—This paper presents a feature based vision destination recognition approach. A simplified static calibration of camera is implemented and then based on the homography matrix calculated by coplanar matched Speeded-Up Robust Features (SURF) points the destination which is not even a direct feature point is tracked. By using the reference image and the invariance of area ratio of two triangles in affine transformation, the accumulated error is eliminated and the homography is corrected. The result of the laboratory dynamic experiment proves the stability and accuracy of the method.

Keywords—camera calibration; destination matching; visual heading

I. INTRODUCTION

Navigation system, which has been widely studied by almost all kinds of measurement means, is very important for mobile robot to accomplish complicated tasks. Vision, inertial, GPS and GPS similar systems, ranging and direction finding are the fundamental techniques and different integrated systems based on them have been developed.

Vision and inertial technology are typical independent measurements, which can still be reliable without any outside help, and are very important to cope with GPS denied and other complicated situations. Simultaneous Localization and Mapping (SLAM) is the most popular algorithm and has been widely applied besides mobile robot. Because abundant image features can be extracted as landmarks by image matching, vision SLAM [1] becomes a potential and effective mean for mobile robot and Unmanned Air Vehicle. But due to the big amounts of calculation and scale ambiguity of mono vision [2], vision SLAM still has some bottlenecks to break through.

A lot of works have been done to extract navigation information, which is classified as the product of low level and high level processing [3], other than vision SLAM. Based on planar constrains of image features, homography can be utilized to distinguish moving object and back ground [4]. For a mobile robot moving on the floor, considering the special circumstance, simplified visual odometer can be realized not only to measure speed but also 3D objects [5]. The vanishing point is very helpful for recognizing the circumstance and extracting heading when the mobile robot is moving in outdoor environment [6]. Basically, many works have been done to calibrate the camera in different situations [7], [8].

In order to realize less calculation and better applicability, we present a robust visual destination matching algorithm which can extract the destination even no direct feature of it

exists. A fast simplified camera calibration method is also developed.

II. SIMPLIFIED STATIC CALIBRATION OF LINEAR CAMERA

A low cost camera cannot be effectively calibrated by traditional methods because of its poor consistency. In order to realize quantitative measurement of heading, we implement a simplified static calibration as below.

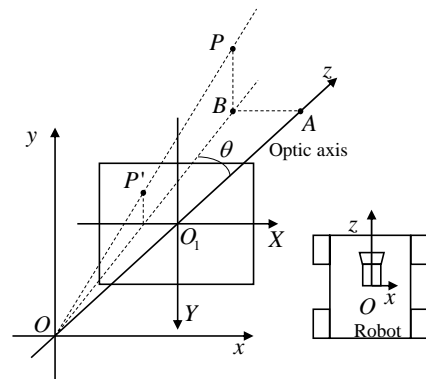


FIGURE I. IMAGE-FORMING PRINCIPLE OF LINEAR CAMERA

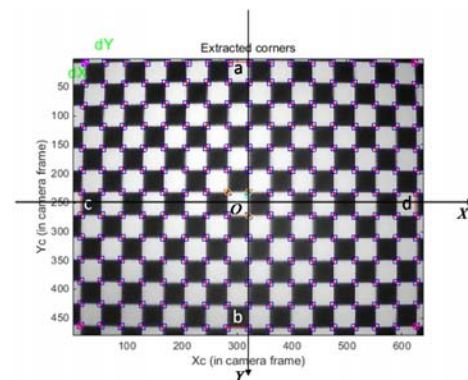


FIGURE II. IMAGE OF CHECKERBOARD

According to Fig. 1, each point in the image has a fixed azimuth and pitch angle with respect to the optic axis. A checkerboard is utilized to gird the image and extract the relative angles at corners. The distances between adjacent corners which are nearest to the middle of each image edge are utilized to determine whether the optic axis is perpendicular to the checker board by the following equation

$$\frac{|a-b|}{a} \leq 0.05, \quad \frac{|c-d|}{c} \leq 0.05. \quad (1)$$

As shown in Fig. 2, the middle pixel of the image is selected as the origin O of the coordinate system XOY , which is not necessary on a corner. Then the relative position in the world coordinate system (WCS), corresponding to any given pixel, on the checkerboard can be calculated by bilinear interpolation according to the detected corners. Based on the focal length f and the known distance D from the camera to the checkerboard, the relative azimuth angle referring to optic axis is

$$\theta(P') = \arctan\left(\frac{AB}{OA}\right) = \arctan\left(\frac{x_{P'}^w - x_O^w}{D + f}\right). \quad (2)$$

where $(x_{P'}^w, y_{P'}^w, z_{P'}^w)$ and (x_O^w, y_O^w, z_O^w) denote respectively the coordinates of pixels P' and O in the WCS.

Therefore, an azimuth angle reference table can be constructed and utilized to determine the relative horizontal angle between any pixel and the optic axis by interpolation. Experiments have demonstrated the error of the calibration is less than 0.1 degree.

III. DESTINATION TRACKING

For a low cost vision/INS system, visual destination recognition can effectively reduce the heading error caused by the low accuracy of gyroscope, especially in long straight moving where the mobile robot can be considered as moving in a two dimensional plane. How to locate and track the destination where there is no direct image feature in the area of destination is the key problem.

A. Location Scheme of Destination

According to the invariance of homography transformation, a non-feature destination can be located by those surrounding features in the same plane which is far enough from the camera. The proper plane selection is based on matched feature points between the image $I(k_1)$ and $I(k_2)$. In this paper P denotes the coordinates in pixel (x_p, y_p) of point in image in the coordinate system XOY .

$$\mathbb{P}^d(k_1) = \left\{ P_i^m(k_1) \mid \left| P_i^m(k_1) - P'(k_1) \right| < T_{dis}, i \in [1, N_m] \right\} \quad (3)$$

where $P_i^m(k_1)$ and $P'(k_1)$ denote the i th matched feature and the known location of destination of image $I(k_1)$ respectively, $\mathbb{P}^d(k_1)$ denotes the features surrounding destination, T_{dis} and N_m are the distance threshold and the total number of features.

Because only T_{dis} cannot guarantee all the points in

$\mathbb{P}^d(k_1)$ are on the same plane, MSAC with a 1.6 pixels distance threshold is further utilized to get a more accurate plane and the corresponding homography matrix H_1 based on which the destination $P'(k_2)$ in $I(k_2)$ is acquired.

$$\mathbb{P}^M(k_1), H_1 = \left\{ P_j^d(k_1), H_1 \mid MSAC[P_j^d(k_1), P_j^d(k_2)], j \in [1, N_d] \right\} \quad (4)$$

$$[P'(k_2) \ 1] = H_1 [P'(k_1) \ 1] \quad (5)$$

B. Selection of Regions of Interest

In order to accelerate the calculation, two different size regions of interest, ROI_1 and ROI_2 which are all centered at the predicted destination, are selected according to the measurements of gyroscope as bellow.

$$P_c(k+1) = P'(k) + \left(\frac{W_l}{\alpha} \omega \Delta t, 0 \right) \quad (6)$$

$$\begin{cases} W_{R_i}(k+1) = \eta_w(k+1)W_l \\ \eta_w(k+1) = \begin{cases} 0.6, |\omega \Delta t / \alpha| \in [0, 0.1) \\ 0.2, |\omega \Delta t / \alpha| \in [0.1, 0.3) \\ 1.0, |\omega \Delta t / \alpha| \in [0.3, 0.5) \end{cases} \end{cases} \quad (7)$$

where $P_c(k+1)$ denotes the predicted destination in $I(k+1)$, W_l is the width of image, α is the horizontal field angle of view, ω denotes the angular rate measured by the gyroscope, Δt denotes the time difference between two matching images, $W_{R_i}(k+1)$ is the width of $ROI_1(k+1)$ and $\eta_w(k+1)$ is the ratio of width between ROI_1 and image. When $|\omega \Delta t / \alpha| \geq 0.5$, the features in those two images cannot be matched.

And the height of ROI_1 is double of the diameter of ROI_2 where the radius of $ROI_2(k+1)$ $T_{dis}(k+1)$ is determined by

$$T_{dis}(k+1) = 1.5 \max\left(\left| P_i^M(k) - P'(k) \right|, P_i^M(k) \in \mathbb{P}^M(k) \right). \quad (8)$$

C. Error Correction

As the destination tracking between adjacent images in sequential images brings accumulated error and not all of the coplanar points selected are in the ideal plan at infinity, the homography is not accurate. The solutions are as follows.

First, a previous and can be effectively matched image, containing the determined destination, is selected as the reference image to suppress the accumulated error. The reference image $I_r(k)$ can be updated adaptively as bellow.

$$I_r(k) = I(k+v), \text{ if } \frac{N(k+1)}{N(k+v)} < 0.5 \quad (9)$$

where $N(k+1)$ and $N(k+v)$ denote the amount of matched feature point pairs between $I_r(k)$ and the image captured at the corresponding moment.

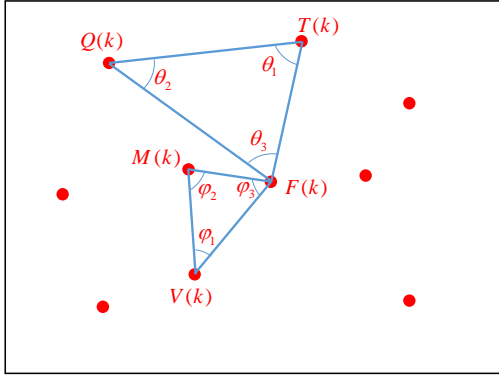


FIGURE III. TRIANGLES SELECTED IN $I_r(k)$

Second, based on the invariance of area ratio of two triangles in affine transformation, the stable ideal coplanar matched feature points are selected to correct the destination in $I(k+v)$ when $I_r(k)$ is updated. As shown in Fig. 3, the ratio of area of coplanar triangles, $S_{\square_{TQF}}(k)$ and $S_{\square_{VMF}}(k)$, is constant in affine transformation. According to the stable matched feature points $\mathbb{P}^S(k)$ and $\mathbb{P}^S(k+v)$, the ideal coplanar matched feature points in $\mathbb{P}^{SP}(k)$ are selected by

$$\left| \frac{S_{\square_{TQF}}(k)}{S_{\square_{VMF}}(k)} - \frac{S_{\square_{TQF}}(k+v)}{S_{\square_{VMF}}(k+v)} \right| < T_m, \begin{cases} \frac{S_{\square_{TQF}}(k)}{S_{\square_{VMF}}(k)} > T_r > 1 \\ C_{VMF}(k) > T_c \\ \theta_1, \theta_2, \theta_3, \varphi_1, \varphi_2, \varphi_3 > T_a \end{cases} \quad (10)$$

$$T(k), Q(k), V(k), M(k), F(k) \in \mathbb{P}^S(k) = \mathbb{P}^d(k) \cap \mathbb{P}^d(k+1) \cap \dots \cap \mathbb{P}^d(k+v)$$

where T_m is the threshold of difference of area ratio, T_r is the area ratio threshold, T_c and T_a denote respectively the perimeter threshold and the threshold of interior angle. In order to guarantee the sensitivity of the area ratio to the location of common vertex, the two triangles are selected first by T_r , furthermore, the robustness of calculation is guaranteed by T_c and T_a .

Based on sets of stable ideal coplanar matched feature points $\mathbb{P}^{SP}(k)$ and $\mathbb{P}^{SP}(k+v)$, the accurate homography matrix H_2 is calculated and according to H_2 the latest destination is corrected.

$$H_2 = \left\{ H_2 \left[MSAC \left[P_t^{SP}(k), P_t^{SP}(k+v) \right], t \in [1, Z] \right] \right\} \quad (11)$$

$$[P'(k+v) \ 1]^t = H_2 [P'(k) \ 1]^t \quad (12)$$

D. Steps of Algorithm

The steps of visual heading calculation algorithm are as follows:

1. Select the initial destination in the first captured image $I(1)$, extract and describe SURF points in $ROI_1(1)$ in and store $I(1)$ as the first reference image;
2. Capture new image, adjust the ROI_1 and ROI_2 by Eq. 6, Eq.7 and Eq.8, calculate the homography by Eq. 4 and locate the latest destination;
3. Update the reference image and correct the latest destination if the condition of renewal is satisfied.
4. Based on Eq. 2, calculate the visual heading;
5. Repeat step 2) to 4) until the navigation service is stopped.

IV. LABORATORY DYNAMIC EXPERIMENT

As shown in Fig. 5, a wheeled mobile robot (WMR) with a single-axis gyroscope ADIS16265 and a camera was used in the laboratory dynamic experiment. The inertial heading is calculated by STM32F103VET6 and transferred to a laptop using an Intel Core i5 CPU with a constant frequency of 3.30GHz. The 640×480 pixels image captured by the camera was transferred to the laptop through USB. The program running in MATLAB took charge of image processing. The speed of image processing was 0.2s per frame on average.

In this experiment, the WMR moved from the beginning to the end of a corridor. 20 images were used for calculating visual headings. In this experiment, $f = 2cm$, $D = 26cm$, $T_{dis} = 0.15 \times 480 \text{ pixels}$, $T_m = 10^{-10}$, $T_r = 1.2$, $T_c = 25 \text{ pixels}$ and $T_a = 30 \text{ degrees}$.

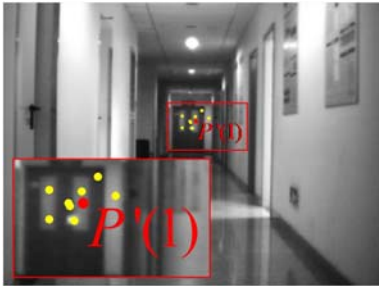
Fig. 4 shows parts of the process of destination tracking. As shown in Fig. 4.(a), The initial destination is selected in the first image $I(1)$. Fig. 4.(b) shows the matched SURF points between $ROI_1(1)$ and $ROI_1(2)$ in $I(1)$. And the coplanar matched SURF points $\mathbb{P}^M(1)$ and $\mathbb{P}^M(6)$ selected by the homography H_1 are shown in Fig. 4.(c) and Fig 4.(d). When $I(9)$ is updated into $I_r(9)$, the triangles are selected by Eq. 10 for acquiring $\mathbb{P}^{SP}(9)$ as shown in Fig. 4.(e), In Fig. 4.(f), the point denotes the final destination located. Compared with the initial destination, the result of tracking is approximately at the same location of the initial destination, which proves that the accuracy of the destination tracking can satisfy the demands of practical applications.



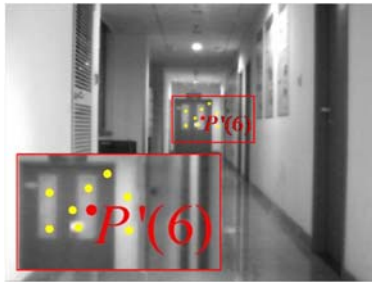
(a) Initial destination in $I(1)$



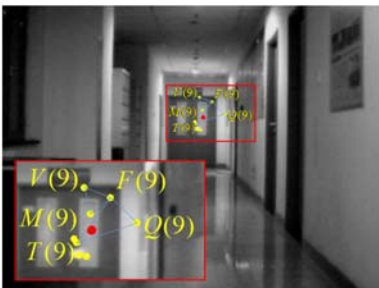
(b) Matched SURF points in $ROI_1(1)$



(c) Coplanar points $\mathbb{P}^M(1)$ in $I(1)$



(d) Coplanar points $\mathbb{P}^M(6)$ in $I(6)$



(e) Ideal coplanar points $\mathbb{P}^{SP}(9)$ in $I(9)$



(f) Final destination in $I(20)$

FIGURE IV. EXPERIMENTAL PROCESS

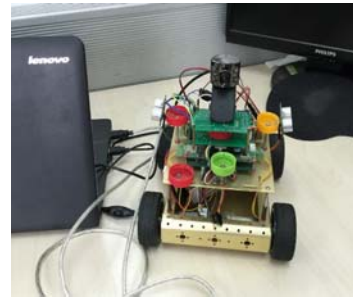


FIGURE V. THE EXPERIMENTAL FACILITIES

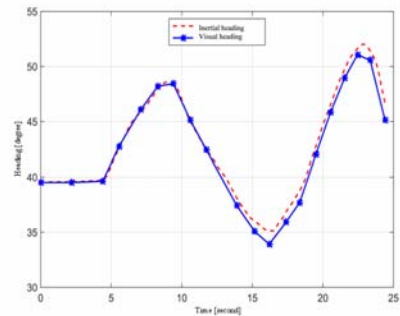


FIGURE VI. THE INERTIAL HEADINGS AND THE VISUAL HEADINGS

As shown by the result in Fig. 6, the trend of two curves are basically the same. Visual headings are accurate and do not diverge over time. However, the accuracy of inertial headings are as same as the visual headings at the beginning, but they diverge fast due to the accumulated error. Therefore, it can be concluded that the visual heading calculation algorithm is effective and robust.

V. CONCLUSIONS

In this paper, a robust visual heading calculation method is presented for the mobile robot. The method is based on the destination recognition and the result of simplified static camera calibration. The destination which has no direct feature is located due to the use of the homography calculated by the coplanar matched SURF points. In order to eliminate the accumulative error, the reference image was used in the destination matching. And based on the invariance of area ratio of two triangles in affine transformation, the latest destination was corrected.

Experiment shows that, compared with the low cost INS, the vision navigation system which utilized our method is robust and accurate.

ACKNOWLEDGEMENT

This research was financially supported by the International (regional) cooperation and exchange program of natural foundation. The number of the program is 61320106010.

REFERENCE

- [1] Won D H, Lee Y J, Sung S, et al. Integration of vision based SLAM and nonlinear filter for simple mobile robot navigation[C]//2008 IEEE National Aerospace and Electronics Conference. IEEE, 2008: 373-378.
- [2] Huang R, Tan P, Chen B M. Monocular vision-based autonomous navigation system on a toy quadcopter in unknown environments[C]//Unmanned Aircraft Systems (ICUAS), 2015 International Conference on. IEEE, 2015: 1260-1269.
- [3] Hrach D, Brandner M, Fossati P, et al. Intelligent vision-sensor for robot-sensing applications[C]//International Workshop on Robotic Sensors: Robotic and Sensor Environments, 2005. IEEE, 2005: 37-42.
- [4] Rodrigo R, Chen Z, Samarabandu J. Feature motion for monocular robot navigation[C]//2006 International Conference on Information and Automation. IEEE, 2006: 201-205.
- [5] Cobos J, Pacheco L, Cufi X, et al. Integrating visual odometry and dead-reckoning for robot localization and obstacle detection[C]//Automation Quality and Testing Robotics (AQTR), 2010 IEEE International Conference on. IEEE, 2010, 1: 1-6.
- [6] Siagian C, Chang C K, Itti L. Mobile robot navigation system in outdoor pedestrian environment using vision-based road recognition[C]//Robotics and Automation (ICRA), 2013 IEEE International Conference on. IEEE, 2013: 564-571.
- [7] Ji Y, Yamashita A, Asama H. Automatic calibration and trajectory reconstruction of mobile robot in camera sensor network[C]//2015 IEEE International Conference on Automation Science and Engineering (CASE). IEEE, 2015: 206-211.
- [8] Petersen A, Koch R. Video-based realtime IMU-camera calibration for robot navigation[C]//SPIE Photonics Europe. International Society for Optics and Photonics, 2012: 843706-843706-10.