# Accurate self-localization of mobile robots based on vision sensors

**Xiaoming Dong***

*Department of Information and Technology Anqing Teachers college, 128 Linghu South Road, Anqing, Anhui province, China*

*Received 6 July 2014, www.cmnt.lv*

**Abstract**

Robot localization is a challenging problem in indoor environment since no GPS information is available. In this paper, an algorithm was proposed for accurate localization, which designed a delicate way to extract the feature points at first, then the position of the robot was determined using the relation of the features in different images, finally, the Kalman filer was designed to decrease the error caused by robot's moving. Experiments showed the accuracy and robustness of this algorithm.

*Keywords:* mobile robot, multi-cues fusion, vision sensor, Kalman filer, localization

## 1 Introduction

Self-localization of Mobile robots is one of the most important tasks in the various applications. GPS is a commonly used technique in outer environments. However, GPS cannot be used in some special conditions such as indoor environments [1-4]. What's more, GPS cannot achieve the high accuracy required in some applications of mobile robots. Vision sensors based on cameras have long been used to implement the localization of mobile robots, which is a promising technique that improves the capability of a mobile robot to estimate its position and the speed of motion. Extensive researches have been done to make use of vision sensors in practice, monocular image sequences or stereo image sequences are used in several algorithms [5-9]; however, because of the inaccurate feature matching, there are still many difficulties to estimate the robots position effectively. The aim of this paper is to design a good method to achieve the task of localization in indoor environments.

The paper is organized as follows. In section 2, we designed a delicate way to detect the position of the distinguished feature points. Then the outlier rejection is discussed, section 3 discussed the way we used to estimate the position of robots, section 4, to decrease the error further in the time that robots move, Kalman filter is designed. In Section 4, the algorithm is verified in various experiments. Finally, the conclusions are given in section 5.

## 2 Feature extractions and outlier rejection

The primary idea in vision sensors is to select the distinguished features to describe the environments. How to gain the accurate point features are the most important problem in robot localization [10-12]. In our proposed algorithm, considering the real environment characters, we designed a hierarchical framework to obtain the accurate feature points. The framework mainly contains two parts, which are the global search and local search. In global search, we extract feature points as many as possible, while in local search, we only choose the feature points that have a good effect in pose estimation.

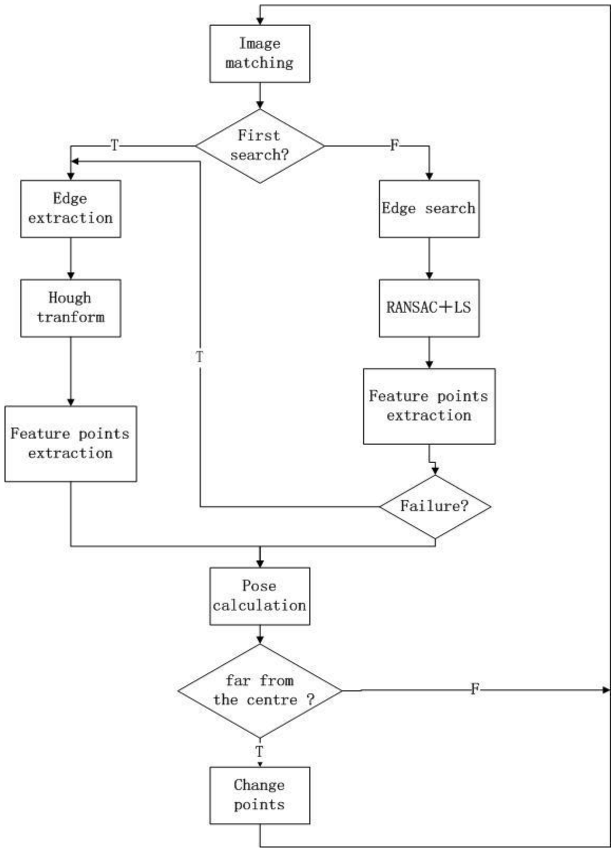


FIGURE 1 The flow graph of the feature extraction

**Corresponding author* e-mail: dongxiaoming008@126.com

Figure 1 shows the detailed process, which can be described as follow:

1) When the robot is the first time to run, the edge points is to be extracted in the global scope.

2) The parameter of the lines is obtained by using Hough Transform.

3) The optimization is implemented by using least square.

4) The cross of the lines is chosen as the feature points.

Considering the speed of the calculation and the changes in the continuous images are little, so we firstly extract the lines in the neighboring area of the last image.

When the robot is not the first time to run, and the Hough transform plus the least square are implemented similarly. If it had a great change in environments, which contributes to the results that the local search loses effectiveness, the robot turns to the global search.

Integrating the two policies to extract the point features, our algorithm can get the distinguished points to pose estimation and acceptable speed. Figure 2 shows the points that robot has obtained from ceiling.

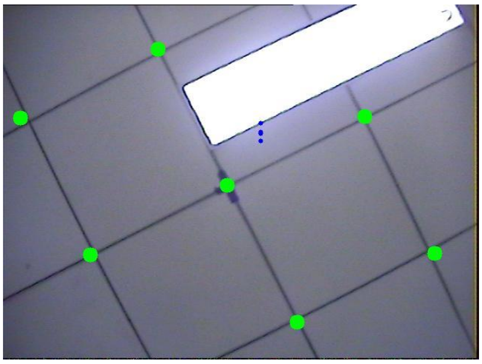

FIGURE 2 the feature points detected

For the accuracy of the point features, the method designs two steps for outlier's rejection. Firstly, an initial estimation of the homography matrix $H$ is achieved by the RANSAC algorithm, secondly, the unmatched point features are picked out according to the homography matrix $H$. the process is described as follows in calculating the homography matrix based on the correspondent features: it is assumed that the correspondent features are $P_i(x_i, y_i)(i=1,...,5)$, and then we can compute $\mathbf{H}$ as follows. $P_i$ and $V_i$ satisfy:

$$k\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} = \mathbf{H}\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}, \quad \mathbf{H} = \begin{bmatrix} m_1 & m_2 & m_3 \\ m_4 & m_5 & m_6 \\ m_7 & m_8 & m_9 \end{bmatrix}. \tag{1}$$

Eliminating $k$ from equation (1), we can get two equations. Combining these two equations, we get:

$$\mathbf{KM} = m_9\mathbf{U}, \tag{2}$$

where $\mathbf{K}$, $\mathbf{M}$ and $\mathbf{U}$ are defined as follows $(i=1,...,5)$:

$$\mathbf{K} = \begin{bmatrix} x_i & y_i & 1 & 0 & 0 & 0 & -x_iu_i & -y_iu_i \\ 0 & 0 & 0 & x_i & y_i & 1 & -x_iv_i & -y_iv_i \\ & & & & \cdots & & & \\ & & & & \cdots & & & \end{bmatrix}, \tag{3}$$

$$\mathbf{M} = (\mathbf{K}^T\mathbf{K})^{-1}\mathbf{K}^T\mathbf{U}. \tag{4}$$

The process can be concluded as follows:

1) Capturing the current frame and rectifying it to remove the distortion.

2) Finding out the correspondent features between the current frame and the last frame, then computing the homography matrix H to reject outliers.

3) Capturing the next frame and repeat this process.

## 3 Pose estimation

### 3.1 POSE ESTIMATION

Assuming that the coordinate of point $P$ is $[x_w, y_w, z_w, 1]^T$ in the world coordinate system, then the Equation (5) is satisfied, the Equation (5) is called external parameter model of camera.

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} {}^cn_{wx} & {}^co_{wx} & {}^ca_{wx} & {}^cp_{wx} \\ {}^cn_{wy} & {}^co_{wy} & {}^ca_{wy} & {}^cp_{wy} \\ {}^cn_{wz} & {}^co_{wz} & {}^ca_{wz} & {}^cp_{wz} \\ 0 & 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \begin{bmatrix} R & p \\ 0^T & 1 \end{bmatrix}\begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = {}^cT_w\begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}. \tag{5}$$

As the Figure 3 shows, on the time $i$ when the robot is moving, the reference coordinate system $O_{ri}$ is set up in the cross point between the ceiling and the optic axes, And the axis direction of the reference coordinate system is the same as the world coordinate system, Similarly, the reference coordinate system $Or_{i+1}$ is set up too.

When the camera moves from point $Oc_i$ to $Oc_{i+1}$, the process can be decomposed into two parts: the one is the rotation of the robot around axis $Zc_i$ and the other one is the translation from point $Oc_i$ to point $Oc_{i+1}$. Then the relative position will be calculated if the position of the feature point $P_j$ in the reference coordinate system is obtained between time $i$ and time $i$+1.

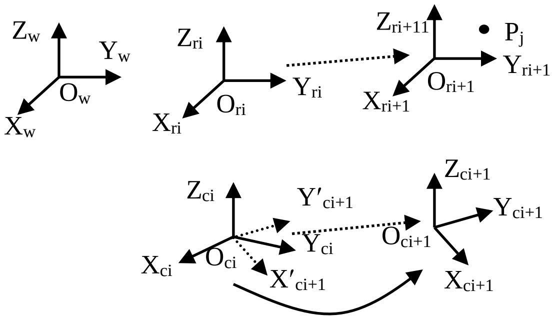


FIGURE 3 The coordinate system in different time

The transform of the reference coordinate can be described as equation 6 in the time *i*.

$$
{}^iT_{ri} = \begin{bmatrix} {}^in_{wx} & {}^io_{wx} & {}^ia_{wx} & 0 \\ {}^in_{wy} & {}^io_{wy} & {}^ia_{wy} & 0 \\ {}^in_{wz} & {}^io_{wz} & {}^ia_{wz} & {}^cp_{rz} \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{6}
$$

The distance between the camera axis and the feature points is required in the localization. If the camera plane is parallel to the ceiling, it would be a constant. Considering the camera axis $O_c$ is the same both in the reference coordinate system and in the world coordinate system, this can be calculated based on the Equations (5) and (6).

$$
\begin{cases} ({}^in_{wx} - {}^ix_{1cj}\,{}^in_{wz})\,{}^ix_{rj} + ({}^io_{wx} - {}^ix_{1cj}\,{}^io_{wz})\,{}^iy_{rj} = {}^ix_{1cj}\,{}^cp_{rz} \\ ({}^in_{wy} - {}^iy_{1cj}\,{}^in_{wz})\,{}^ix_{rj} + ({}^io_{wy} - {}^iy_{1cj}\,{}^io_{wz})\,{}^iy_{rj} = {}^iy_{1cj}\,{}^cp_{rz} \end{cases}. \tag{7}
$$

In the above equations, (${}^ix_{rj}$, ${}^iy_{rj}$) represents the $P_j$ coordinate and (${}^ix_{1cj}$, ${}^iy_{1cj}$) represents the normalized coordinate of the image focus.

Then, the coordinate of feature $P_j$ can be calculated in the reference coordinate system of different time.

Because in time *i* and time *i*+1, the transform matrix between the references coordinate system and the camera ordinate system can be represented as Equation (8):

$$
{}^{ri}T_i = \begin{bmatrix} {}^in_{wx} & {}^in_{wy} & {}^in_{wz} & -{}^in_{wz}\,{}^cp_{rz} \\ {}^io_{wx} & {}^io_{wy} & {}^io_{wz} & -{}^io_{wz}\,{}^cp_{rz} \\ {}^ia_{wx} & {}^ia_{wy} & {}^ia_{wz} & -{}^ia_{wz}\,{}^cp_{rz} \\ 0 & 0 & 0 & 1 \end{bmatrix}. \tag{8}
$$

The localization of the robots can be calculated by iteration based on the equation (9):

$$
\begin{cases} p_{xi+1} = p_{xi} + {}^ix_{rj} - {}^{i+1}x_{rj} - {}^{i+1}n_{wz}\,{}^cp_{rz} + {}^in_{wz}\,{}^cp_{rz} \\ p_{yi+1} = p_{yi} + {}^iy_{rj} - {}^{i+1}y_{rj} - {}^{i+1}o_{wz}\,{}^cp_{rz} + {}^io_{wz}\,{}^cp_{rz} \end{cases}. \tag{9}
$$

## 3.2 ERROR ANALYSIS AND FILTERING

To analyze the error, we set up the moving model of the mobile robot at first, and then the Kalman filter is designed to decrease the error. The robot used in this paper is a wheel robot, its model is shown in Figure 4, assuming that SL represents the distance of the left wheel moves and SR is that of right wheel. The rotation angle of the wheel axis is $\Delta\theta$, $\theta$ is the intersection angle, then the relation between the current pose and the last pose can be referred, assuming that $X_i(x_i, y_i, \theta_i)$ is the current pose and $X_{i-1}(x_{i-1}, y_{i-1}, \theta_{i-1})$ is the last pose.

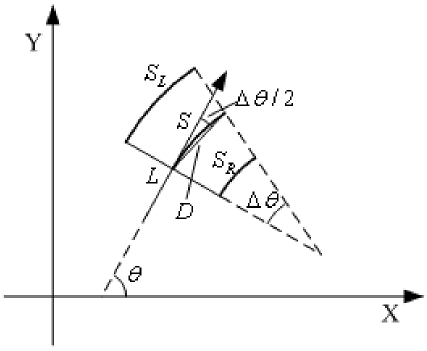


FIGURE 4 The moving model of robot wheels

Then the relation between the rotation angle and the distance the robot moves is:

$$
\Delta\theta = (S_L - S_R)/L. \tag{10}
$$

The arc length, which is the distance the robot centre has moved, can be represented as follows:

$$
S = (S_L + S_R)/2. \tag{11}
$$

And *D* can be referred as:

$$
D/2 = S/\Delta\theta \times \sin(\Delta\theta/2). \tag{12}
$$

Considering that $S_L$ and $S_R$ are very close, it can be assumed that:

$$
\Delta\theta = 0, \tag{13}
$$

$$
\begin{aligned} \Delta x &= S \times \cos\theta, \\ \Delta y &= S \times \sin\theta. \end{aligned} \tag{14}
$$

Updating the current pose:

$$
X_i = X_{i-1} + (\Delta x, \Delta y, \Delta\theta). \tag{15}
$$

Considering the noise, the formula of the pose can be changed as:

$$
X_i = F(X_{i-1}, S_{Li-1}, S_{Ri-1}) + W_i, \tag{16}
$$

where $W_{i-1}$ is assumed as gauss noise and its average value and the variance are:

$$
\mathrm{E}[W_i] = 0,\ \mathrm{E}[W_i, W_i^{\mathrm{T}}] = Q_i. \tag{17}
$$

Covariance can be modeled as diagonal matrix, the diagonal entries are:

$$Q_{11} = K_x \left| S\cos\theta \right|,$$
$$Q_{22} = K_y \left| S\sin\theta \right|, \quad (18)$$
$$Q_{33} = K_{S\theta} \left| S \right| + K_{\theta\theta} \left| \Delta\theta \right|,$$

where $K_x$ and $K_y$ are drift coefficients the robot moves along the axis X and axis Y, similarly, $K_{S\theta}$ and $K_{\theta\theta}$ is the drift coefficients of angles. The values of the coefficients can represent the error. The more the coefficients, the more the errors.

For the more accurate localization, we choose the extended Kalman filter, in EKF framework; the moving model can be represented as:

$$x_t = f(x_{t-1}, u_{t-1}, w_{t-1}) . \quad (19)$$

And the observation model is:

$$z_t = h(x_t, v_t) . \quad (20)$$

Then, state prediction and observation prediction can be determined as:

$$\hat{x}_t^- = f(\hat{x}_{t-1}, u_{t-1}, 0) , \quad (21)$$

$$\hat{z}_t^- = h(\hat{x}_t^-, 0). \quad (22)$$

Using tailor Equation to linear the models, the updating equation in EKF is:

$$\hat{x}_t^- = f(\hat{x}_{t-1}, u_{t-1}, 0),$$
$$P_t^- = A_t P_{t-1} A_t^T + W_t Q_{t-1} W_t^T . \quad (23)$$

And the observation model updating equation is:

$$K_t = P_t^- H_t^T (H_t P_t^- H_t^T + V_t R_t V_t^T)^{-1},$$
$$\hat{x}_t = \hat{x}_t^- + K_t (z_t - h(\hat{x}_t^-, 0)), \quad . \quad (24)$$
$$P_t = (I - K_t H_t) P_t^- .$$

In the design, we set up the observation model according to the character of the vision sensors and decrease the error further.

## 4 Experiments and analysis

Various experiments have been carried out using the mobile robot. The camera is precisely calibrated beforehand and is mounted on the center of the robot fixedly.

The inner calibration parameter is shown in Table 1 and Table 2. They represent the focus and the center respectively of left camera and right camera.

TABLE 1 The inner parameter of calibration (mm)

| Inner parameter | kx | | ky | |
|---|---|---|---|---|
| | *result* | *variants* | *result* | *variants* |
| Left camera | 568.27 | 2.44 | 566.21 | 2.76 |
| Right camera | 569.35 | 3.18 | 569.58 | 2.98 |

TABLE 2 The center of the camera (mm)

| u0 | | v0 | |
|---|---|---|---|
| result | variants | result | variants |
| 228.77 | 2.19 | 183.25 | 2.35 |
| 232.12 | 2.91 | 177.98 | 1.98 |

The external parameter is shown in Tables 3 and 4, which are the parameter of the rotation and the translation respectively. And the unit is mm. the values represent the relative angle and translation between the two cameras.

TABLE 3 Translation vector of the external parameter (relative to the left camera)

| | | | | |
|---|---|---|---|---|
| **Translation vector P(mm)** | result | 445.83 | 6.54 | 8.25 |
| | variants | 4.23 | 3.15 | 2.86 |

TABLE 4 Rotation vector of the external parameter (relative to the left camera)

| | | |
|---|---|---|
| **Rotation matrix R** | **results** | 0.0127; 0.9875; -0.0643<br>0.9876; -0.0354; -0.1768<br>-0.1920; -0.0675; -0.9326 |
| | **variance** | 0.05342<br>0.03425<br>0.00978 |

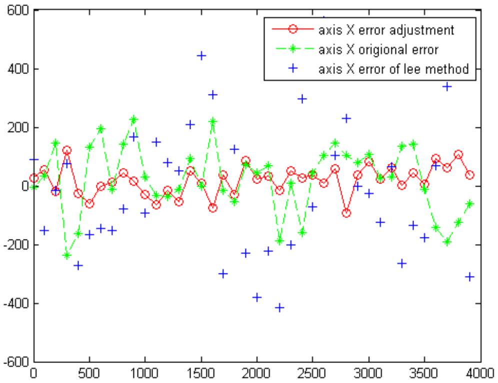


FIGURE 5 The error comparison of axis X

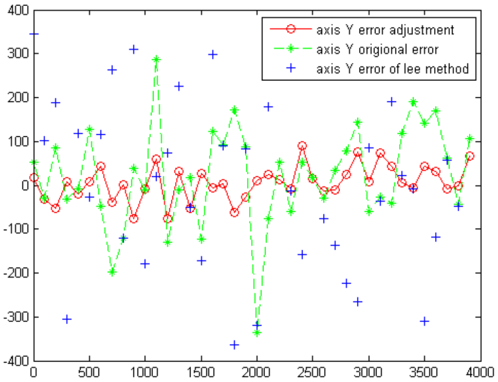


FIGURE 6 The error comparison of axis Y

The comparison is implemented between the method we proposed and the famous Lee method [13], because the accurate feature extraction, the error of our method is about 100mm or so, while the Lee algorithm is about 200mm, the error is decreased further through the Kalman filter, which is about 50mm after filter. Figure 5 gives the error of axis *X* and Figure 6 gives that of axis Y. The figure shows that the effect of robot localization has been greatly improved.

## 5 Conclusions and future work

In the various applications, Vision sensor is one of the most important sensors because it can provide larger amount of information comparing with other traditional sensors. In this paper, an effective method is proposed for accurate robot localization. The experiments show its accurateness. Vision sensor has its shortcoming that the features in image would be affected when the environment has great changes. Future work will consider more about the robustness of the Vision sensors, the combination of the traditional sensor and the Vision sensors maybe a possible way to resolve this problem.

## References


[1] Kiryati N, Kälviäinen H, Alaoutinen S 2000 Randomized or probabilistic Hough transform: unified performance evaluation *Pattern Recognition Letters* **21**(13-14) 1157-64.

[2] Nistér D, Naroditsky O, Bergen J 2006 Visual odometry for ground vehicle applications *Journal of Field Robotics* **23**(1) 3-20

[3] Yang Z-F, Tsai W-H 1999 *IEEE Transactions on Industrial Electronics* **46**(3) 653-61

[4] Olson C G, Matthies L H Schoppers M, Maimone M W 2003 Rover navigation using stereo ego-motion *Robotics and Autonomous Systems* **43**(4) 215-29

[5] Cumani A, Guiducci A 2010 Comparison of feature detectors for rover navigation *Mathematical Modelling and Applied Computing* **1**(1) 245-50

[6] Schweighofer G, Pinz A 2006 *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28**(12): 2024-30.

[7] Benhimane S, Malis E 2004 Real-time image-based tracking of planes using efficient second-order minimization *Intelligent Robots and Systems* **41**(4) 1943-8

[8] Jepson A D, Fleet D J, El-Maraghi T F 2003 *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25**(10) 1296-311

[9] Kalal Z, Mikolajczyk K, Matas J 2012 *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34**(7) 1409-22

[10] Mei X, Ling H 2011 *IEEE Transactions on Pattern Analysis and Machine Intelligence* **33**(11) 2259-72

[11] Avidan S 2007 *IEEE Transactions on Pattern Analysis and Machine Intelligence* **29**(2) 261-71

[12] Ross D A, Lim J, Lin R-S, Yang M-H 2008 Incremental learning for robust visual tracking *International Journal in Computer Vision* **21**(3) 125-41

[13] Jang G, Kim S, Lee W, Kweon I 2003 Robust self-localization of mobile robots using artificial and natural landmarks *Proceedings of the IEEE International Symposium on Computational Intelligence in Robotics and Automation 2003* **1** 412-7



**Author**

**Xiaoming Dong, born in November, 1977, Anqing County, Anhui Province, P.R. China**

**Current position, grades:** Associate Professor at the School of Anqing teachers college, China.
**University studies:** MSc at Beijing science and technology University in China.
**Scientific interests:** machine vision, pattern recognition.
**Publications:** more than 10 papers.
**Experience:** teaching experience of 9 years, 2 scientific research projects.