

# Imaging Monocular Vision Localization and Application in Smart Home

Gao Junchai<sup>1\*</sup>, Yan Keding<sup>1</sup>, Han Bing<sup>2</sup>

<sup>1</sup>*School of Electronic Information Engineering, Xi'an Technological University, Xi'an, 710032, China;*

<sup>2</sup>*Information Technology Center, Xi'an Technological University, Xi'an, 710032, China*

*E-mail: lili77999@163.com*

## Abstract

*The mobile robot motion is generally on the plane, not in three-dimensional space, and its motion is in a plane constraint way, which would have a special representation. The paper models a plane motion with two-dimensional motion parameters to be solved for the plane motion, which simplifies the essential matrix model, parameters and its constraints greatly, and enhances the independence of model parameters. In order to enhance the estimation accuracy of plane motion, the paper combines Kalman filter method for localization estimation. A initial localization estimation is based on the simplified plane motion model with corresponding points projecting in two viewpoint images of a 3D point, and nonlinear optimization localization estimation is based on distance minimization criterion to iteration, more accurate can be got. For the stochastic process error in the localization estimation of sequences images, filtering localization estimation is based on Kalman filter to smooth error. Simulation results show that the method is simple and feasible for mobile robot localization.*

**Keywords:** *Mobile robot; Monocular vision; Kalman Filter Estimation Method*

## 1. Introduction

Transformation matrix between sequence images acquired from monocular vision motion is restrained by multi-view geometric [1], epipolar geometry can describe geometric relationship between two images, and the fundamental matrix and essential matrix is algebra expression of epipolar geometry, which is only relevant to internal and external parameters of the monocular imaging system, rather than relying on a three-dimensional scene structure. The fundamental matrix is the geometric constraints between two un-calibrated images, without knowing internal parameters of imaging system, which can describe the whole epipolar geometry relationship [2]. If the internal parameters are known, essential matrix constraint is modeled between two images, namely the relative pose of the monocular vision [3]. The position and orientation relationship between the views determines the motion of the mobile robot.

The traditional localization estimation system based on the classical SFM algorithm can process camera moving in six degrees [4]. Real-time large-scale monocular vision localization algorithm was proposed by Nistér [5], which removed external data with RANSAC, matched 2D- 2D feature with tracking, and estimated the camera new position and orientation with five points smallest solver for RANSAC motion assumptions. All-

---

<sup>1\*</sup> *Address correspondence to this author at NO.4 Jinhua North Road, Xi'an, China. Postcard: 710032;*

round orientation monocular vision localization was used by Corke [6], which acquired images from the catadioptric camera systems, and tracked with optical flow. Lhuillier and Mouragnon [7-9] *etc.* estimated motion and 3D maps simultaneously with bundle adjustment (Bundle Adjustment, BA) in the local window. Scaramuzza [10-12] firstly introduced kinematic model of the vehicle, Ackermann steering model and plane motion constrained into the visual localization model, estimated closed solution of vehicle motion model successfully with only one pair associated feature.

For the absence of mobile robot motion and steering models, mobile robot localization with plane motion constraints is studied, and Kalman filter localization method [13-15] is used, in which a simplified essential matrix is modeled for plane motion. That reduces the motion freedom, while reducing the minimum matching points set used to estimate; nonlinear optimization estimation with initial values and consensus associated data, more accurate localization of current frame is got; for correlative of sequences frames, Kalman filter is used to filter estimating stochastic process errors.

## 2. Multi-view Geometry Model Established with Plane Motion Constraints for Mobile Robot

During the mobile robot moving, sequence images are acquired with monocular vision. That position of corresponding point changes in sequence images implies motion of the mobile robot, so transformation model of corresponding points is established between the adjacent two images.

### 2.1. The Epipolar Geometry of Pair-view Images

Let  $I$  and  $I'$  are two images acquired at two different positions and perspectives of the same scene for the mobile robot.  $C$  and  $C'$  are respectively the optical center of the camera, as shown in Figure 1.  $e'$  is the projection point of  $C$  on the image plane  $I'$ ,  $e$  is the projection point of  $C'$  on the image plane  $I$ , which are called epipolar points. Connect line between two cameras center is called baseline. Landmark  $M$  with two cameras center determines triangular plane, which known as epipolar plane.  $m$  and  $m'$  are projection points of the landmark  $M$  on the image plane  $I$  and  $I'$ , respectively, they are the corresponding point of landmark. So, for any corresponding point  $m$  on the image plane  $I$  of landmark, its corresponding point  $m'$  on the image plane  $I'$  is bound on epipolar line  $l'$ ; For any corresponding point  $m'$  on the image plane  $I'$  of landmark, its corresponding point  $m$  on the image plane  $I$  is bound on epipolar line  $l$ . epipolar line  $l$  is radial ray on the image plane  $I$  projecting over landmark corresponding points  $m'$  and the camera's center  $C'$ , epipolar line  $l'$  is radial ray on the image plane  $I'$  projecting over landmark corresponding points  $m$  and the camera's center  $C$ .

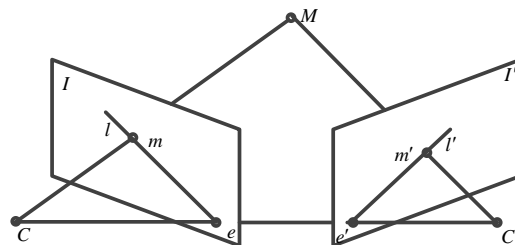


Figure 1. The Epipolar Geometric Relationships

From the above, it can be seen from the epipolar geometry, corresponding points meet certain constraints between two images [16-17], and mathematical model can be modeled for the corresponding points of landmarks.

**2.1.1. The Fundamental Matrix Constraint Model of Pair-view Images:** The epipolar geometry [18] as shown in Figure 1, according to the non-uniqueness of projection, landmark  $M$  located in the line of camera center  $C$  and projection point  $m$  connecting, and in front of the camera, could be projected onto the same point  $m = PM$ . The condition that the space landmark set meets can be expressed as:

$$M(\lambda) = P^{-1}m + \lambda C \quad (1)$$

Here  $P^{-1}$  is the pseudo-inverse of projection matrix  $P$ , In order to obtain the relationship between projection points  $m \leftrightarrow m'$  of landmark, firstly the epipolar line  $l'$  corresponding to  $m$  is determined. Two points are selected from the line  $M(\lambda)$ , whose projection on the image plane  $I'$  can determined epipolar line  $l'$ . The first point is the camera center  $C$ , whose projection on the image plane  $I'$  is epipolar point  $e' = P'C$ ; another point is  $P^{-1}m$ , its projection on the image plane  $I'$  is  $P'P^{-1}m$ . It can be seen:

$$l' = e' \times P'P^{-1}m \quad (2)$$

For the vector  $X = [x_1 \quad x_2 \quad x_3]$ , its anti-symmetric matrix is defined as:

$$[X]_{\times} = \begin{bmatrix} 0 & -x_3 & x_2 \\ x_3 & 0 & -x_1 \\ -x_2 & x_1 & 0 \end{bmatrix} \quad (3)$$

Then, the cross product of formula (2) can be written in the form of matrix multiplication,  $l' = Fm$ . Here:

$$F = [e']_{\times} P'P^{-1} \quad (4)$$

Formula (4) is the fundamental matrix between two images, so the corresponding point  $m \leftrightarrow m'$  of the landmark projection must satisfy the epipolar line constraint:

$$m'^T Fm = 0 \quad (5)$$

It is a condition that all corresponding points are satisfied, from the formula (5) fundamental matrix  $F$  can be solved with all corresponding points of landmarks, which is a  $3 \times 3$  matrix, rank is 2, has seven independent parameters. So the motion of the mobile robot resolved with at least seven pairs corresponding points.

The fundamental matrix  $F$  includes both camera internal parameter matrix  $K$ , also includes camera external motion parameters  $(R, t)$ . If the intrinsic parameters matrix  $K$  is known, and unchanged in the pair viewpoints, the second viewpoint pose to the first viewpoint is  $(R, t)$ , here  $R$  is rotation matrix,  $t$  is translation vector. The fundamental matrix between the two images can be expressed as:

$$F = K^{-T} \begin{bmatrix} t \\ \times \end{bmatrix} RK^{-1} \quad (6)$$

If matrix  $F$  transformed with matrix  $K$ , only motion parameters matrix can be obtained. Rotation parameters and translation parameters can be obtained by decomposing between two adjacent images.

**2.1.2. Essential Matrix Constraint Model of Pair-view Images:** In a fundamental matrix  $F$ , a part of the parameters are the camera intrinsic parameters, others related to the camera external parameters, which are defined as the essential matrix:

$$E = \begin{bmatrix} t \\ \times \end{bmatrix} R \quad (7)$$

It only has a relationship with motion parameters of monocular vision system. The epipolar constraint with the essential matrix can be expressed as:

$$m'^T K^{-T} EK^{-1}m = 0 \quad (8)$$

For a known intrinsic parameter matrix of camera, two images can be transformed with normalized:

$$\tilde{m} = K^{-1}m; \tilde{m}' = K^{-1}m' \quad (9)$$

So epipolar constraint equation of normalized corresponding points is:

$$\tilde{m}'^T E \tilde{m} = 0 \quad (10)$$

Essential matrix  $E$  represents the epipolar geometry of normalized corresponding points, which is fundamental matrix of normalized corresponding points. It is a matrix of  $3 \times 3$ , and rank is 2, has five independent parameters. So the motion of the mobile robot resolved with at least five pairs of corresponding points. Once essential matrix  $E$  is solved, it can be decomposed into rotational motion parameters  $R$  and translational motion parameter  $t$  with a constant factor between two viewpoints.

## 2.2. The Plane Motion Model Designed for Mobile Robot

In the natural environment, the mobile robot motion coincide the essential matrix  $E$  constraints; there are six degrees of freedom. If the mobile robot moves in plane of indoor floor, workbench, or outdoor road, its motion also abides by the plane motion constraints, then there are only heading angle rotational and position translational changes on the plane motion between two adjacent pose. The motion has three degrees of freedom, and the motion parameters can be expressed as:

$$R = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad T = \begin{bmatrix} T_x & T_y & 0 \end{bmatrix}^T \quad (11)$$

Here  $\theta \in R$  is the rotation angle of the second viewpoint to the first viewpoint,  $T_x$ ,  $T_y$  is the translation vector of the second viewpoint to the first viewpoint, the anti-symmetric matrix of the translation vector  $T$  is:

$$[T]_x = \begin{bmatrix} 0 & 0 & T_y \\ 0 & 0 & -T_x \\ -T_y & T_x & 0 \end{bmatrix} \quad (12)$$

The  $R$  and  $[T]_x$  of equation (11) (12) are substituted into the equation (7), Then:

$$E = \begin{bmatrix} 0 & 0 & T_y \\ 0 & 0 & -T_x \\ -T_y \cos \theta + T_x \sin \theta & T_y \sin \theta + T_x \cos \theta & 0 \end{bmatrix} \quad (13)$$

The above essential matrix containing three unknown parameters of uncoupled. So, if the essential matrix is estimated,  $T_y$  and  $T_x$  can be got directly, and the angle  $\theta$  can be calculated. If each item of the essential matrix is represented by a parameter, then the essential matrix has four unknown parameters, a pair of corresponding points determines an equation, then matrix parameters can be linear resolved with four pairs of corresponding points by the least squares method. Because of equation homogeneity,  $T_y$  and  $T_x$  are determined by a scale factor, so the essential matrix can be normalized, which divided by  $T_y$ , as following formula:

$$E = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & e_1 \\ e_2 & e_3 & 0 \end{bmatrix} \quad (14)$$

$$\text{Here, } e_1 = -\frac{T_x}{T_y}, \quad e_2 = \frac{-T_y \cos \theta + T_x \sin \theta}{T_y}, \quad e_3 = \frac{T_y \sin \theta + T_x \cos \theta}{T_y}$$

Through the above transformation there are three unknown parameters in the essential matrix, because it is also the constraint matrix of natural scenes, it can be solved linearly in principle with only three pairs of corresponding points, whereby the motion parameters of the mobile robot can be got. The three unknown parameters has two degrees of freedom, namely: a direction angle and the translation ratio of x and y. Due to the direction angle parameter, the constraint relationship of three unknown parameters is non-linear relationship, and essential matrix can be resolved in nonlinear with two pairs of corresponding points.

### 3. Mobile Robot Plane Motion Model Estimation Method

From multi-view geometric model of mobile robot, we can know that data association between the adjacent two images is needed firstly. For the plane motion two viewpoints fast localization parameter estimation method is proposed. In the simplified essential matrix estimation, parameters constraints are not taken into account, so motion parameters are estimated by nonlinear optimization with the tracking features, to get more accurate motion parameters of two viewpoints. There is always random error in estimation, while the actual motion of the mobile robot is relatively smooth, and finally the whole motion estimates with the Kalman filter.

### 3.1. Data Associated Algorithm Designed based on Essential Matrix Model and Plane Constraint

For pyramid KLT tracking features, due to the smooth areas and similar areas, there will be error tracking corresponding points, they will lead to an error or large deviation motion estimated, so the motion estimation methods need process error data mechanism to filter out false corresponding points. How to distinguish association data, and remove the wrong tracking corresponding points are important. Based on the pair-view geometric, the correct tracking corresponding points imply the same essential matrix transform model, which is determined linearly with minimum five pairs corresponding points while no plane motion constraints, otherwise at least three pairs corresponding points, and two pairs corresponding points can determine in nonlinear. Data association is the premise and foundation for essential matrix estimation. According to the clustering properties of proper tracking and randomness of error tracking between corresponding points, using the implicit essential matrix for data association and multiple sets of associated data is found guided by different planes, to remove the error tracking corresponding point. Due to the essential matrix variable is three, for the constraints between variables, the real motion parameters are only two variables, so in order to avoid the uncertainty caused by probability MAPSAC method, the two-dimensional histogram voting method is used for data associated.

By the formula (14), the relationship between the corresponding points tracking can be expressed with motion parameters variables  $\theta$  and  $e_1$  as:

$$v \sin \theta - u \cos \theta - e_1 (u \sin \theta + v \cos \theta) + u' + e_1 v' = 0 \quad (15)$$

For  $N$  pairs normalized coordinate tracking features  $\{\tilde{m}_{k-1}^i \leftrightarrow \tilde{m}_k^i\}$ ,  $i = 1:N$ , the specific implementation steps of histogram voting method are as follows:

1) For normalized coordinate tracking features,  $\{\tilde{m}_{k-1}^i \leftrightarrow \tilde{m}_k^i\}$ ,  $i = 1:N$ , nonlinear equation is construct as form (15);

2) Motion parameters variables  $\theta$  and  $e_1$  are discrete in certain resolution and range (the step of  $\theta$  is 0.05 degrees with range  $\pm 10$  degrees; the step of  $e_1$  is 0.1, with range  $\pm 10$ ) motion parameter variables  $\theta$  and  $e_1$ , and two-dimensional histogram is constructed;

3)  $\theta^j$  is calculated and voted for histogram in accordance with the linear equation (15) for each pair of tracking corresponding feature and discrete  $e_1^i$ ;

4) If the votes number of model is highest, whose tracking corresponding features are selected as associated data;

5) According to preliminary associated data, the motion parameters rotation matrix and translation vector are estimated;

6) Map is build with pair viewpoint,  $R$  and  $t$  of motion model for associated data  $\{\tilde{m}_{k-1}^i \leftrightarrow \tilde{m}_k^i\}$ , which is three-dimensional coordinate estimate  $\hat{M}^i$  without scale. Outside data is determined with re-projection error threshold value.

### 3.2. Rapid Localization Parameter Estimation Algorithm Based on Simplified Essential Matrix Model

Parameters of essential matrix itself has not physical sense, but it is composite result of the mobile robot motion parameters, so there is a constraint relationship between the essential matrix parameters described as the following theorem, according to the constraint relationships, localization parameters can be estimated with low computational complexity and fast speed.

Theorem 1: Determinant of essential matrix must be zero, that is  $\det(E) = 0$

For Theorem 1, Homography model of formula (14) is automatically satisfied, which can not be a constraint.

Theorem 2:  $EE^T E - \frac{1}{2} \text{trace}(EE^T)E = 0$

Theorem 2 can form constraint of formula (14), so the essential matrix can be solved with only two pairs features. Once the essential matrix obtained, then  $T_x/T_y = -e_1$ , translation proportion was available directly, which was substituted in the following formula:

$$\begin{cases} \frac{-T_y \cos \theta + T_x \sin \theta}{T_y} = e_2 \\ \frac{T_y \sin \theta + T_x \cos \theta}{T_y} = e_3 \end{cases} \quad (16)$$

The rotation angle  $\theta$  can be calculated without SVD decomposition, moreover a determining unique solution, which eludes the solving ambiguity of five pairs features.

The indirection motion estimation based on the essential matrix decomposition, firstly essential matrix is estimated, and then the motion parameters are calculated. The indirection motion estimation method is simple, but the results are not satisfactory, even if there are many associated data, the analytical solution is still much affected by noise. There are mainly two reasons: Firstly, the constraint of essential matrix is ignoring, while the every elements of the essential matrix  $E$  is not independent between; secondly is that the minimize equation has not much physical meaning, so the unknown motion parameters can be estimated directly with Longuet-Higgins criteria:

$$LH(R, t) = \sum_{i=1}^n (\tilde{m}'_i (t \times R \tilde{m}_i))^2 \quad (17)$$

Rotation matrix is represented with three parameters, the translation vector can be represented with its spherical coordinates, but the size of the translation can not be determined with these guidelines. However, which has not resolving solution; three motion parameters must be estimated with this criterion by nonlinear optimization. Its initial solution can be got by the front fast motion estimation.

Longuet-Higgins Criterion has not much physical sense too, that the distance of point  $m'_i$  to the straight line  $l'_i$  as small as possible is a guideline of more physical sense.

Let  $l'_i \equiv (\alpha'_i, \beta'_i, \gamma'_i)$ , the distance of point  $m'_i$  to the straight line  $l'_i$  is:

$$d'_i(m'_i, l'_i) = \frac{|\tilde{m}'_i{}^T [t]_x R \tilde{m}'_i|}{\sqrt{\alpha_i'^2 + \beta_i'^2}} \quad (18)$$

In order to make the two images act with symmetry, it can also calculate the distance  $d_i$  of point  $m'_i$  to the straight line  $l'_i$ . Let  $l_i \equiv (\alpha_i, \beta_i, \gamma_i)$ , the

$$d_i(m_i, l_i) = \frac{|\tilde{m}_i{}^T ([t]_x R)^T \tilde{m}'_i|}{\sqrt{\alpha_i^2 + \beta_i^2}} \quad (19)$$

Therefore, the motion parameters can be solve by minimized the following equation:

$$\sum_{i=1}^n (d_i^2 + d_i'^2) = \sum_{i=1}^n (\tilde{m}'_i{}^T [t]_x R \tilde{m}'_i)^2 \left( \frac{1}{\alpha_i'^2 + \beta_i'^2} + \frac{1}{\alpha_i^2 + \beta_i^2} \right) \quad (20)$$

### 3.3. Mobile Robot Kalman Filter Localization Estimation Algorithm

Assuming the state variables of the discrete time process  $r \in R^n$  is estimated with Kalman filter. For the motion model of the mobile robot system, because the camera is fixed on the mobile robot platform, it is no longer moving and rotating, so their motion model is consistent. The key problem of monocular vision localization is that its motion state is expressed in a reasonable way. The most common motion model is 13 variables model: location coordinates  $s^W$ , pose  $q^W$ , line speed  $v^W$  and angular velocities  $\omega^C$ . In order to simplify the noise model, the linear velocity and angular velocity are both constant, that is, the mean of linear acceleration  $\alpha^W$  and angular acceleration  $\beta^C$  is zero, and independence mutual with Gauss normal distribution. The state model of the camera is:

$$r_k = f(r_{k-1}, u) = \begin{pmatrix} s_k^W \\ q_k^W \\ v_k^W \\ \omega_k^W \end{pmatrix} = \begin{pmatrix} s_{k-1}^W + (v_{k-1}^W + V_{k-1}^W) \Delta t \\ q_{k-1}^{WC} \times q((\omega_{k-1}^C + \Omega_{k-1}^C) \Delta t) \\ v_{k-1}^W + V_{k-1}^W \\ \omega_{k-1}^W + \Omega_{k-1}^C \end{pmatrix} \quad (21)$$

Here,  $r_k^W = (x_k^W, y_k^W)^T$ ,  $q_k^{WC} = (q_{0,k}^{WC}, q_{1,k}^{WC}, q_{2,k}^{WC}, q_{3,k}^{WC})^T$ ,  $v_k^W = (v_{x,k}^W, v_{y,k}^W)^T$ ,  $\omega_k^W = (\omega_{z,k}^W)$ .

In the above formula (21), the superscript  $C$  represents that the reference system is the camera coordinate system, and the superscript  $W$  represents that the reference system is the world coordinate system, and the superscript  $WC$  represents that the camera coordinate system is converted to the world coordinate system.

And,  $V_{k-1}^W$  and  $\Omega_{k-1}^C$  represent the size of the additional linear velocity and angular velocity;  $q((\omega_k^C + \Omega_k^C) \Delta t)$  represents the quaternion change obtained by the camera rotation in time  $\Delta t$ . In addition, the  $u$  of the state transition function indicates the noise caused by the control input of the system. Here is defined as:



$$u = \begin{pmatrix} V_{k-1}^W \\ \Omega_{k-1}^W \end{pmatrix} = \begin{pmatrix} a_{k-1}\Delta t \\ b_{k-1}\Delta t \end{pmatrix} \quad (22)$$

Here,  $a_{k-1}$  and  $b_{k-1}$  are the Gauss distribution white noise of three dimensional known, which is used to simulate the uncertainty of camera motion.

The motion estimated of the adjacent key frames is the motion measurement value of the essential matrix constraint measurement equation. That is the relative pose changes, so:

$$Z_k = H_k r_k + W_k \quad (23)$$

Here,  $H_k = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

From  $H_k$  matrix parameters of equation (23), Observations express the rotation angle and translation of adjacent frames, which can be estimated by the foregoing essential matrix constraint. Random signal  $W_k$  respectively indicate process noise and observation noise, assuming them as independent, normalized distributed white noise. In the actual system, the observation noise covariance matrix  $Q$  may change when each iteration.

Kalman filter estimates process state with feedback control method, for every moment, then obtained feedback with (noisy) variable measurements. Therefore, the Kalman filter can be divided into two parts: prediction and update.

Prediction is to reckon the next time state variables and error covariance estimate, to construct a priori estimates of the next time state.

$$r_{k,k-1} = \Phi_{k,k-1} r_{k-1} \quad (24)$$

Here,  $\Phi_{k,k-1}$  is one step transfer matrix from time T1 to T2 of equation (24),  $\Gamma_{k-1}$  is the system noise drive matrix.

Update construction improve posterior estimation combined with priori estimate and the new variables measurement.

$$r_k = r_{k,k-1} + K_k (Z_k - H_k r_{k,k-1}) \quad (25)$$

Here,  $K_k = P_{k,k-1} H_k^T (H_k P_{k,k-1} H_k^T + R_k)^{-1}$ , and  $Cov[W_k, W_j] = R_k \delta_{kj}$ ,

$$P_{k,k-1} = \Phi_{k,k-1} P_{k-1} \Phi_{k,k-1}^T + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^T, \quad P_k = P_{k,k-1}^{-1} + H_k^T R_k^{-1} H_k$$

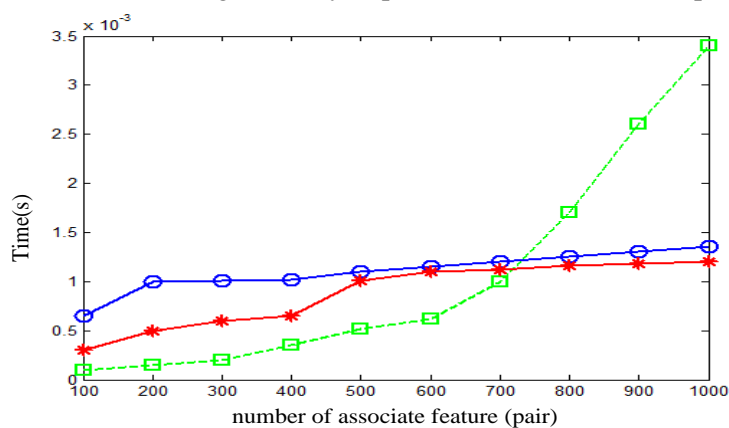
#### 4. Localization Simulation Analysis of the Mobile Robot Essential Matrix

By localization error analysis before with the essential matrix constraint, since the measurement error is of random distribution, a lot of associated feature is needed to more accurately locate, but that will potentially increase the complexity of the calculations, so the number of associated features and sample size are tested with three different localization of the classic eight pairs feature method, two pairs feature method and the Kalman filter method, to provides a analysis basis for improving the efficiency of localization system, and localization performance of three different algorithms is compared with plane motion constraints.

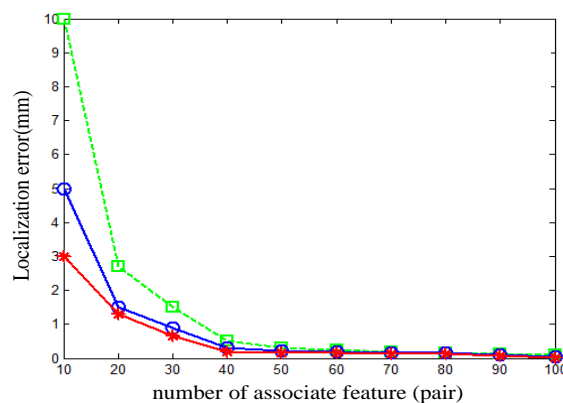
#### 4.1. Influence Test of the Correlation Feature on the Localization Estimation:

In front of the camera, 100 sets of 3D points are generated randomly, which are used as the basis for estimating the motion of the camera in the first viewpoint to the second viewpoint, and 1 pixel random Gauss error of the feature is added with sampling 1000 times.

**4.1.1. Computational Complexity and Localization Error Test of Three Kinds of Localization Algorithms:** Simulated testing with three different localization of the classic eight pairs feature method, two pairs feature method and the Kalman filter method. As shown in Figure 2, eight pairs feature method as shown in green line, two pairs feature method as shown in blue line, the Kalman filter method as shown in red line, with the increasing of the correlation feature, the computation complexity of three algorithms is increasing, and the computational complexity of the eight pairs of feature localization estimation method is increasing with nonlinear. Although the plane motion model has been simplified, when the associated feature is less than 700, the two pairs feature localization estimation method is slightly larger than eight pairs feature. Only when the associated features are more, the computational complexity of two pairs feature localization estimation method has appearance advantages. With the increase of the associated feature, the error of all the estimation is decreased, and the localization accuracy of estimation is not significantly improved after more than 40 pairs feature.



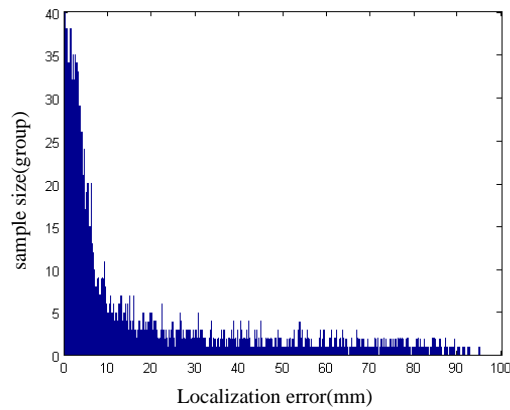
(a) Computational Complexity Varies with the Number of Associated Feature



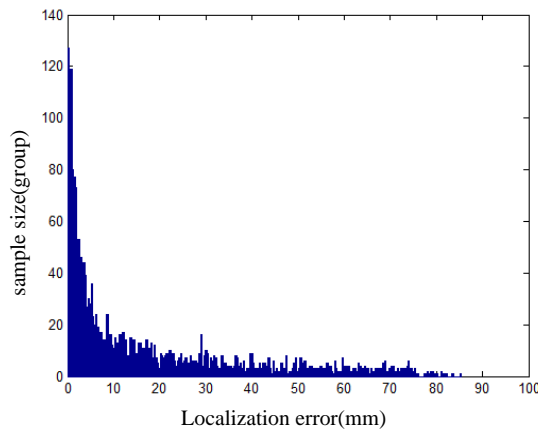
(b) Localization Error Varies with the Number of Associated Features

**Figure 2. Computational Complexity and Localization Error of Three Algorithms with the Number of Associated Features**

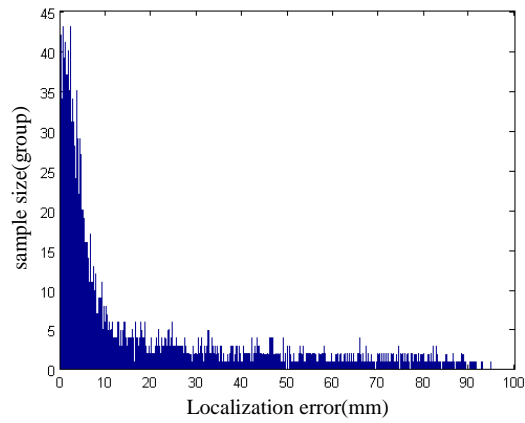
**4.1.2. Localization Error Distribution of Three Localization Algorithms with the Number of Associated Feature Samples:** Localization error distribution of every localization method with the minimum associated feature samples and 50 pairs associated feature sample is analyzed, shown in Figure 3. When located with the minimum sample, error distribution of the Kalman filtering localization method is similar to the eight pairs feature method, and the two pairs features has the larger localization error. When located with 50 pairs associated feature, error distributions of three methods are similar, but Kalman filter localization error is minimized.



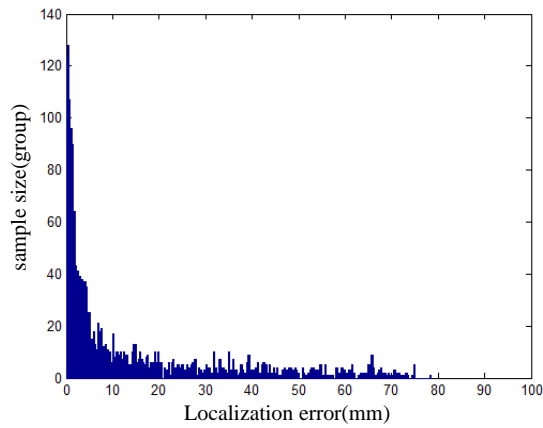
(a) Localization Error Distribution of Eight Pairs Feature Method with Minimum Sample



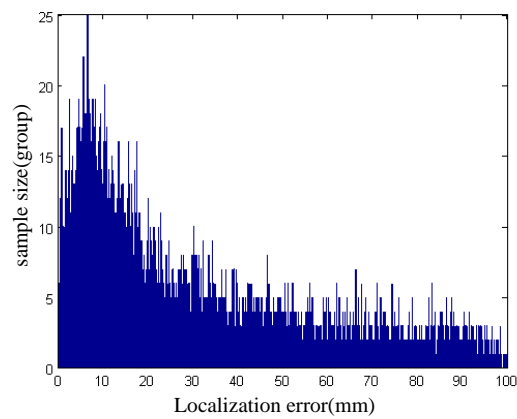
(b) Localization Error Distribution of Eight Pairs Feature Method with 50 Pairs Feature Sample



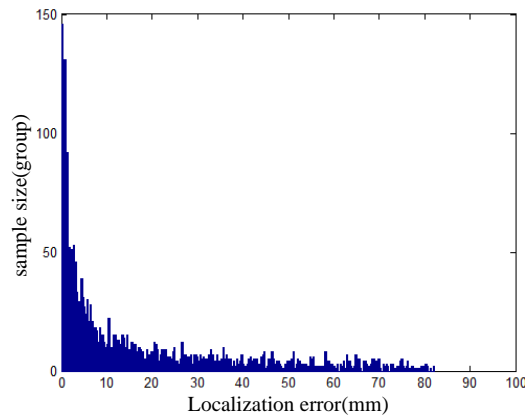
(c) Localization Error Distribution of Kalman Filter Method with Minimum Sample



(d) Localization Error Distribution of Kalman Filtering Method with 50 Pairs Feature Samples



(e) Localization Error Distribution of Two Pairs Feature Method with Minimum Sample

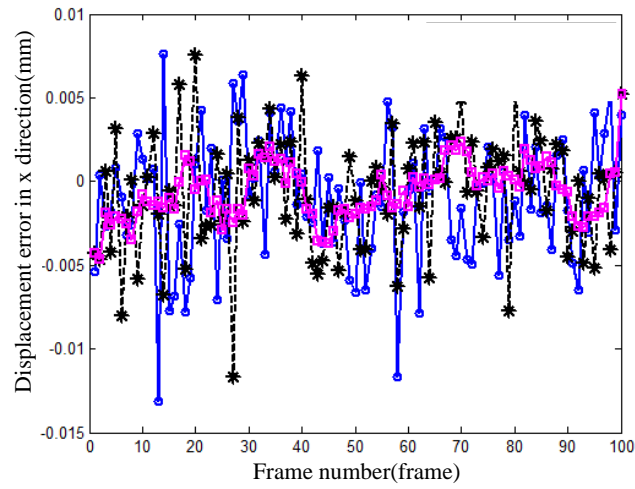


(f) Localization Error Distribution of Two Pairs Feature Method with 50 Pairs Feature Samples

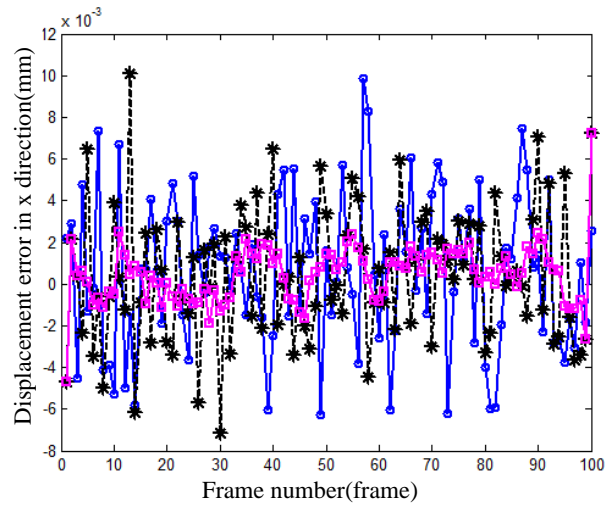
**Figure 3. Localization Error Distribution of Different Localization Method with Different Number Samples**

#### 4.2. Localization Estimation Test of Plane Motion Constrained

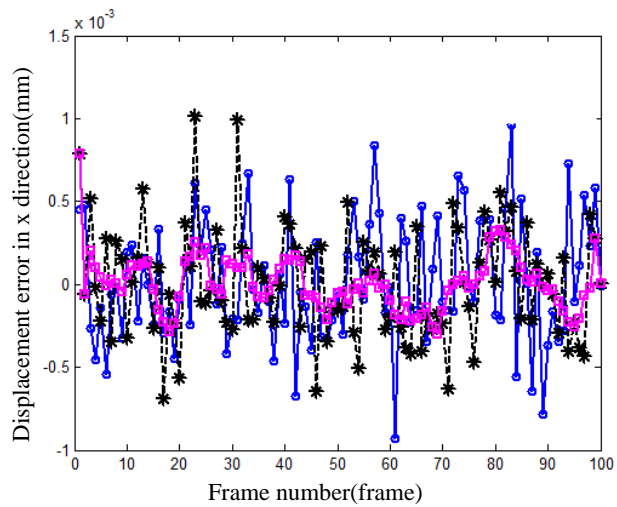
In order to compare the localization performance of different algorithms with plane motion constraints, the motion model of the mobile robot is in accordance with the constant velocity model. The relative displacement and steering angle of X direction and Y direction is predicted in accordance with the constant velocity model from two continuous frames, the number of observation frames is 100 frames. Under the given condition, with the plane motion constraints, the location and pose estimation are compared and simulated with the eight pairs feature method, two pairs feature method and Kalman filtering method, the simulation results are shown in Figure 4. Compared from the figure, the X and Y direction displacement estimate error of eight pairs feature method is maximum, up to 15mm; relative steering angle estimate error is the biggest too, which can reach 0.001 arc, whose accuracy is less than two pairs feature method. Therefore, if we know a priori motion constraints, it is best to use more accurate model for modeling; more accurate localization can be got. The estimated trajectory of Kalman filter method is smoother than that of two pairs feature method, localization accuracy is similar to two pairs feature method.



(a) Estimation of Relative Displacement between Two Frames in x Direction



(b) Estimation of Relative Displacement between Two Frames in y Direction



(c) Estimation of Relative Steering Angle between Two Frames

**Figure 4. Relative Pose Estimation Curve between Two Frames**

## 5. Conclusion

For a natural scene, this paper studied motion estimation localization principle with image sequences based on essential matrix, and analyzed the error. For estimated computational complexity of classic eight pairs feature, it proposed monocular vision Kalman filter positioning method based on two pairs feature, established monocular vision relative motion equations and observation equations to estimate filtering motion parameters. For plane motion constraints, it designed the plane motion model and the corresponding fast motion estimation algorithm, which optimized the estimation results. Practice simulation show that two pairs feature Kalman filter localization methods is more effective.

## Acknowledgement

Foundation item: Principal Scientific Research Funded by The Xi'an Technological University (Program No. XAGDXJJ1212), and Special Scientific Research Program of Shaanxi Provincial Education Department (Program No.15JK1365). Authors are grateful to the Xi'an Technological University and Shaanxi Provincial Education Department for financial support to carry out this work.

## References

- [1] Kundu A., Krishna K. M. and Sivaswamy J., "Moving object detection by multi-view geometric techniques from a single camera mounted robot[C]", 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, (2009), pp. 4306-4312.
- [2] Luong Q T. and Faugeras O. D., "The fundamental matrix: Theory, algorithms, and stability analysis[J]", International journal of computer vision, vol. 17, no. 1, (1996), pp. 43-75.
- [3] Yang J., Rao D. and Chung S., "Monocular vision based navigation in GPS-denied riverine environments[C]", Proceedings of the AIAA Info tech Aerospace Conference, St. Louis, MO, (2011), pp. 1-12.
- [4] Oliensis J., "A critique of structure-from-motion algorithms[J]", Computer Vision and Image Understanding, vol. 80, no. 2, (2000), pp. 172-214.
- [5] Pollefeys M., Nistér D. and Frahm J. M., "Detailed real-time urban 3d reconstruction from video[J]", International Journal of Computer Vision, vol. 78, no. 2-3, (2008), pp. 143-167.
- [6] Corke P. and Hutchinson S. A., "Real-time vision, tracking and control[J]", (2000), pp. 622-629.
- [7] Mouragnon E., Lhuillier M. and Dhome M., "Generic and Real-Time Structure from Motion[C]", BMVC, (2007), vol. 7, no. 4, pp. 1-10.
- [8] Lhuillier M., "Incremental fusion of structure-from-motion and GPS using constrained bundle adjustments[J]", IEEE transactions on pattern analysis and machine intelligence, vol. 34, no. 12, (2012), pp. 2489-2495.
- [9] Lhuillier M., "Fusion of gps and structure-from-motion using constrained bundle adjustments[C]", 2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, (2011), pp. 3025-3032.
- [10] Scaramuzza D. and Fraundorfer F., "Visual odometry [tutorial][J]", IEEE Robotics & Automation Magazine, vol. 18, no. 4, (2011), pp. 80-92.
- [11] Scaramuzza D., Martinelli A. and Siegwart R., "A flexible technique for accurate omnidirectional camera calibration and structure from motion[C].Fourth IEEE International Conference on Computer Vision Systems (ICVS'06)", IEEE, (2006), pp. 45-45.
- [12] Scaramuzza D. and Siegwart R., "Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles [J]", IEEE transactions on robotics, vol. 24, no. 5, (2008), pp. 1015-1026.
- [13] Jetto L., Longhi S. and Venturini G., "Development and experimental validation of an adaptive extended Kalman filter for the localization of mobile robots[J]", IEEE Transactions on Robotics and Automation, vol. 15, no. 2, (1999), pp. 219-229.
- [14] Rezaei S. and Sengupta R., "Kalman filter-based integration of DGPS and vehicle sensors for localization [J]", IEEE Transactions on Control Systems Technology, vol. 15, no. 6, (2007), pp. 1080-1088.
- [15] Kwon S., Yang K. W. and Park S., "An effective kalman filter localization method for mobile robots[C]", 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, (2006), pp. 1524-1529.
- [16] Zhang Z., Deriche R. and Faugeras O., "A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry[J]", Artificial intelligence, vol. 78, no. 1, (1995), pp. 87-119.

- [17] Zhang Z., “Determining the epipolar geometry and its uncertainty: A review[J]”, International journal of computer vision, vol. 27, no. 2, (1998), pp. 161-195.
- [18] Zhang Z., “On the epipolar geometry between two images with lens distortion[C]”, Proceedings of the 13th International Conference on Pattern Recognition. IEEE, vol. 1, (1996), pp. 407-411.

### **Authors**



**Gao Junchai**, works in Xi’an Technological University from 2006, and receive associate professor in 2012. Her research interests include image processing, machine vision and visual navigation.