ICAS 2021 SHANGHAI
International Council of the Aeronautical Sciences
32nd Congress of the International Council of the Aeronautical Sciences
September 6-10, 2021 Pudong Shangri-La, Shanghai, China

A Robust and Real-time Visual-Inertial Pose Estimation for Fixed-wing Aircraft Landing

Guanfeng Yu[1,2], Lei Zhang*[2], Changhao Zou[2], Yuecheng Liu[3], Yue Cheng[2]

[1]School of Software, Northwestern Polytechnical University

[2]AVIC Xi'an Aeronautics Computing Technique Research Institute

[3]Beijing Aerospace Automatic Control Institute

Abstract

To assist fixed-wing aircraft landing, a precise, robust and real time vision-inertial based navigation method is proposed. The work presented in this paper focuses on two points: (1) Runway detection. An inertial aided runway detection algorithm is improved, which can efficiently extract the corners of the runway contour from airborne forward-looking near-infrared images. (2) Pose estimation. Based on geographic coordinates and corresponding image features of runway corners, the aircraft pose relative to runway can be calculated by EPnP precisely. In order to improve the robustness and the output precision of EPnP in case of rare image features, some synthetic image features derived by inertial pose and airport geographic information also be considered as EPnP inputs. Experiments on real flight data show that accurate pose estimation for aircraft landing can be achieved, and the proposed method demonstrates the superiority in terms of speed, precision and robustness.

Keywords: Aircraft landing; Forward-looking near-infrared image; Pose estimation; Runway detection

1. Introduction

Approach and landing of fixed-wing aircraft is one of the most sophisticated flight phases, which is highly dependent on navigation precision, weather and surroundings. Precise position and attitude are significant to control aircraft land accurately and safely, then navigation method is essential to guide aircraft landing.

Recent years many available landing navigation methods have been used for civil or military aircrafts landing. Instrument landing system (ILS) and distance measuring equipment (DME) are widely used, but they are vulnerable to several factors, such as radio, bad weather and high-rise buildings. Furthermore, Microwave landing system (MLS) is precise, but difficult to keep balance between performance, procurement and maintenance costs. Global navigation satellite system (GNSS) is lacking in many scenarios and vulnerable to jamming and spoofing[1], which makes it unavailable or unreliable. Moreover, Differential GNSS is accurate enough but expensive, which is not available for most general airports. Inertial navigation is limited due to its cumulate error and unable to provide precise pose information for landing solely. Ground control station is susceptible to man-made factors. These above issues could cause catastrophic flight accidents easily.

Vision-based navigation provides a new perspective in aircrafts landing. Vision sensors acquire image by passive sensor then obtain position and attitude by image processing, which has advantages of high-precision, non-contact detection, anti-jamming, low cost, small size and low power consumption. Owing to independent of land and air devices, it has great superiority in anti-electronic countermeasures and improvement of autonomy.

Extensive work has been done in the area of vision-based navigation. Simultaneous localization and mapping (SLAM) is the research hotspot at present[2][3][4][5]. State-of-the-art SLAM methods have the virtues of self-localization and navigation under unknown circumstances, but it is not employed in our system due to the following reasons. (1) Accurate vision-SLAM requires high-quality feature tracking, matching and loop closure, high speed and maneuverability of fixed-wing aircraft arise critical challenges for finding and recognition of environment features, and it is difficult to ensure

pose data of updated real-timely. (2) Current typical SLAM algorithms are only applicable to indoor location or relatively small-scale outdoor scenes for robot or small rotorcraft, they are inaccurate in large-scale scene of several kilometers. (3) Airport surroundings is flat and lack of texture which cannot provide sufficient features for SLAM. Therefore, SLAM has poor adaptability in this case which is unable to provide precise, robust and real-time pose information to meet the requirement of fixed-wing aircraft landing.

Beyond that, multifarious methods have been used for fixed-wing aircraft landing. Gui et al.[6] install four infrared lamps near the runway, extracting the exact position of lamps by airborne monocular camera image, which can be used to calculate the pose of aircraft for landing assistance. Nevertheless, the method is not universal due to the additional infrastructure. Besides, Ruchanurucks et al.[7] present a robust positioning method for fixed-wing unmanned aerial vehicles (UAV). By setting ground markers can be observed during landing, the system can achieve pose information with the help of inertial moment unit (IMU) and monocular camera. However, the method depends on artificial markers, lacking of autonomy. Besides, it is discussed in this article that the rotary matrix can be efficiently solved by high-precision IMU, but it also introduces inertial error. And a robust vision-based runway detection and tracking method with an effect energy function is proposed in[8]. Nevertheless, its performance is restricted by the high complexity. In addition, it can only detect the parallel lines of runway, which cannot provide accurate pose estimation. In addition, most of the existing algorithms are carried out under visible image, which is not applicable under the complex scenes, including fog, haze, rain.

Traditional navigation systems (GNSS, INS) work well when cruising. However, aircrafts require accurate pose information with sub-meter precision during approach and landing, as shown in figure.1. Reliable information of position and attitude can be given by vision-based navigation for fixed-wing aircraft landing. Only a minority of scenes and markers can be observable by airborne camera in landing, where runway is the most conspicuous and reliable signs of the scene. Therefore, it has great practical value and promising prospect of application to research on pose estimation based on runway features.

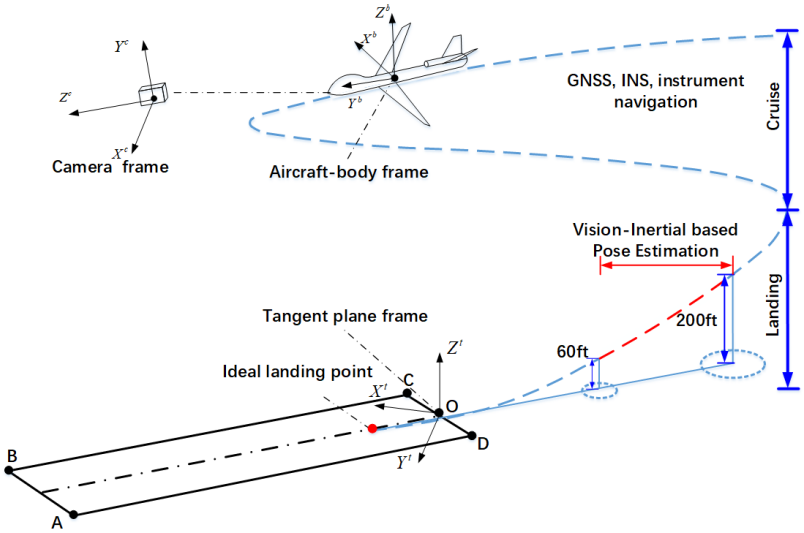


Figure 1 – The available landing scenarios and frame definition of the proposed method

In this paper, airborne inertial data is used fully to assist monocular forward-looking infrared (FLIR) camera to extract runway features from a relatively small region of interest (ROI) rather than an entire image. Comparing with tracking algorithms such as KCF[9], it has higher processing efficiency and higher reliability without introducing extra inertial errors. Then accurate pose of aircraft can be calculated by efficient perspective-n-point (EPnP), which lay a basis for robust, precise and real-time navigation for landing assistance system.

The rest of the paper is structed as follows. In section II, the algorithm frame of visual-inertial pose estimation is proposed. Furthermore, inertial aided runway features extraction and pose estimation are presented in detail. Section III provides a flight data experiment results and a discussion. In section IV, the conclusion is drawn and future work is described.

2. System Architecture and Implementation

In this section, we proposed an algorithm frame for fixed-wing aircraft pose estimation during landing, as shown in figure 2. The complete frame includes three blocks. (1) Inertial aided ROI recognition; (2) Runway features extraction using inertial parameters; (3) Aircraft position and attitude estimation by EPnP Algorithm.

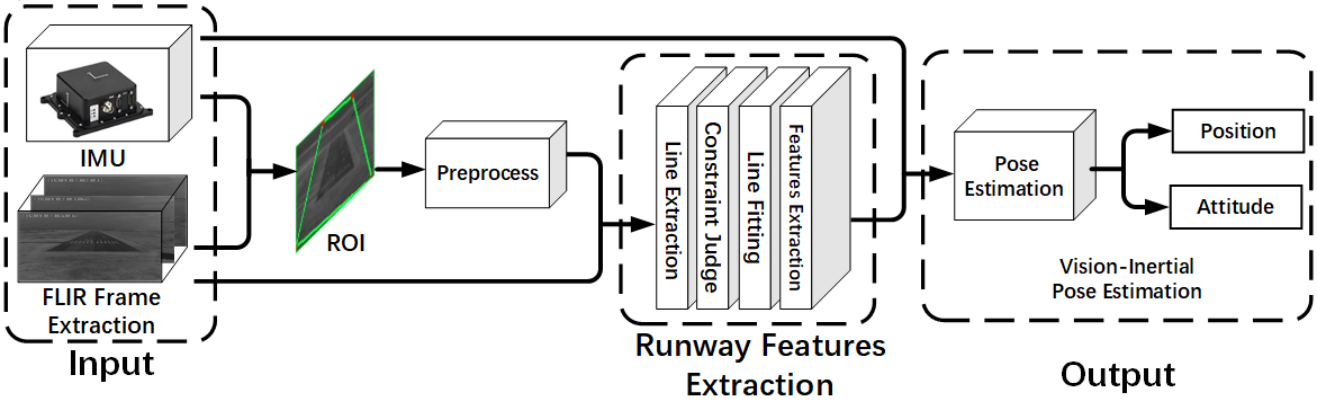


Figure 2 – Vision-Inertial based runway detection and pose estimation algorithm framework

2.1 ROI recognition

ROI extraction is essential for the runway detection process. A huge amount of image data is acquired by the forward-looking camera mounted on fixed-wing aircraft, in which the runway is only a small part. Therefore, appropriate ROI can not only determine runway location in images, but also avoid processing entire image with high computational complexity. It is great significant to reduce computation consuming and ensure real time.

In order to improve the detection precision and the robustness, a coarse-to-fine strategy is used in the system. At the coarse layer, a vision projection model from the world frame to the image frame is built by combining airborne navigation system with image sensor[10], and then a relatively accurate runway region of interest (ROI) can be extracted.

Geodetic coordinates $(L_p,\ \lambda_p,\ h_p)$ (longitude, latitude and altitude) of runway corners are measurable in prior. Using the inertial information of the flight, we can recognize runway and obtain its position in the image roughly.

2.1.1 Projection model

According to the inertial navigation parameter, the transformation process from the geodetic frame to the pixel frame[11] is shown in figure 3. $\mathbf{P}_p^G=(L_p,\ \lambda_p,\ h_p)$ represents the position in geodetic frame, $\mathbf{P}_a^G=(L_a,\ \lambda_a,\ h_a)$ is the aircraft position coordinate in geodetic frame. The symbol $\mathbf{P}_p^e=\begin{bmatrix}x_p^e & y_p^e & z_p^e\end{bmatrix}^T$ and $\mathbf{P}_a^e=\begin{bmatrix}x_a^e & y_a^e & z_a^e\end{bmatrix}^T$ are the feature point coordinate and the aircraft position coordinate in Earth-Centered Earth-Fixed (ECEF) frame respectively. The feature point ECEF coordinate is mapped to the point $\mathbf{P}_p^g=\begin{bmatrix}x^g & y^g & z^g\end{bmatrix}^T$ in geography frame through matrix transformation $[\mathbf{R}_e^g \quad \mathbf{T}_e^g]$, and equivalent to coordinate $\mathbf{P}_p^n=\begin{bmatrix}x^n & y^n & z^n\end{bmatrix}^T$ in navigation frame, $\mathbf{P}_p^b=\begin{bmatrix}x^b & y^b & z^b\end{bmatrix}^T$ and $\mathbf{P}_p^c=\begin{bmatrix}x^c & y^c & z^c\end{bmatrix}^T$ denote the feature point coordinates in the body frame and the camera frame through the transformation matrix $[\mathbf{R}_n^b \quad \mathbf{T}_n^b]$, $[\mathbf{R}_b^c \quad \mathbf{T}_b^c]$ individually. Based on pinhole camera model, point $\mathbf{P}_p^{img}=\begin{bmatrix}r & c & 1\end{bmatrix}^T$ in image coordinate system can be translated from $\mathbf{P}_p^c$.This can be obtained by off-the shelf toolboxes as for instance.

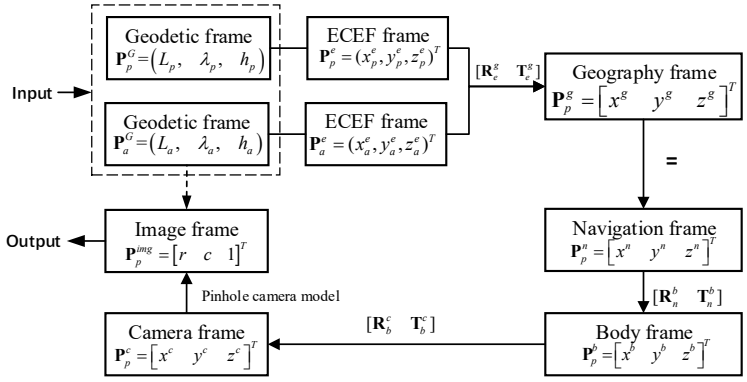


Figure 3 – Vision projection from geodetic frame to pixel frame

2.1.2 ROI detection

In this section, we illustrate the projection procedure from runway plane to airborne camera pixel plane. Then we can get the coordinates of runway corners in pixel plane. Because of cumulative errors, the calculated runway in the image does not coincide with the real runway image. Accordingly, we zoom it and obtain a rough contour of runway, and take the contour as the ROI for precise runway detection. The ROI detection process is shown in figure 4.

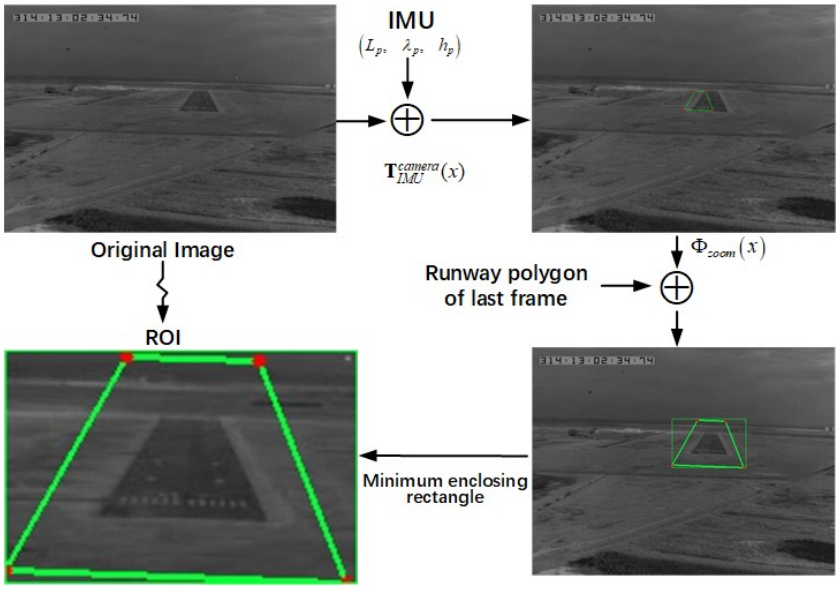


Figure 4 – Inertial aided ROI recognition

2.2 Runway Features Extraction

Differing from rotorcrafts, fixed-wing aircraft cannot hover, then it demands high efficiency in vision navigation. Furthermore, in order to ensure the accuracy and stability of pose estimation, it is required that feature points coordinates are accurate enough. Thus, runway features extraction requires high accuracy and real-time performance.

Due to the thermography principle of infrared camera, FLIR images with more details can detect distinct runway features in terrible weather. Figure 5 shows the visible and near-infrared images taken under the condition of haze.

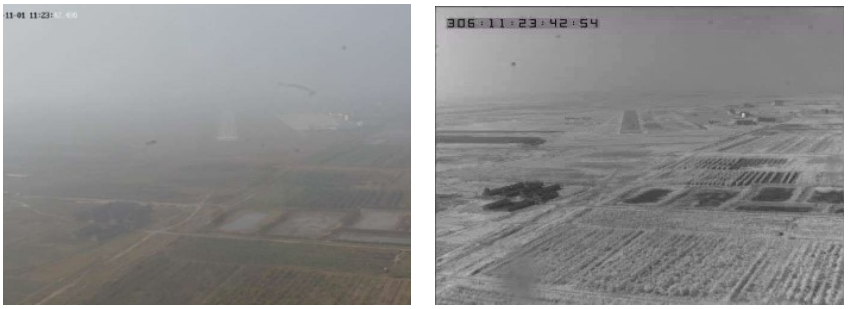


Figure 5 – Contrast of near-infrared (left) and visible (right) images in the low visibility weather

2.2.1 Synthetic Runway Contour

Although inertial sensors have high refreshing frequency and low measurement delay, IMU is insusceptible to exotic environment. Current Visual–Inertial Navigation not only use vision information to correct the cumulative error of IMU unidirectionally, but also optimize inertial information. We utilize inertial parameters of aircraft to calculate ROI of camera images, and take IMU information as a constraint to detect edges of runway, which overcome most existing algorithms' shortcomings of blind traversal.

The outlier inertial aided runway features extraction scheme is as follow:

- Compute ROI of images and inertial runway contour;
- Preprocess the ROI;
- Detect lines of ROI by LSD;
- Make full use of inertial runway's line angle and distance constraint to eliminate interference lines;
- Eliminate non-edged lines by color variance.

2.2.2 Line Segments Detection

Generally, shooting distance of airborne camera is relative far away runway. Moreover, the weather is usually undesirable. Thus, the line segments are hard to be extracted. We preprocess the ROI images with guided filtering algorithm[12], then detect line segments by Line Segments Detection (LSD) algorithm[13].

Comparing with the line segment detection algorithm based on the Hough Transform[14] and Freeman chain code line extraction[15], there are three advantages for LSD: (1) Less time-consuming to obtain a sub-pixel level result. (2) Strong robustness. It can detect line segments in images with texture and much noise. (3) Independent detection without toning parameters. The LSD algorithm is to merge pixels and control error. Bases on gradient features, a structure called Line Support Region (LSR) are used to model line edges. Traversing rectangle regions of image, the edges of region are conformed by angle tolerance and contrario[16][17].The false detection ratio can be automatically regulated by this approach, without toning parameters manually.

2.2.3 Line Fitting

The line fitting is based on the M-estimator[18] technique that iteratively fits the line using the weighted least-squares algorithm. Compared with least square estimation, the method is more robust, besides, abnormal points will have less impact for line fitting. Distance function of fitting is as following:

$$\rho(r) = \frac{C^2}{2} \cdot \left(1 - \exp(-(\frac{r}{C})^2) \right) \tag{1}$$

where $C = 2.9846$.

We dynamically detect ROI by attitude, filter line segments by color and inertial information, and fit line with endpoints, improving detection efficiency greatly. The work flow of runway detection is shown in figure 6.

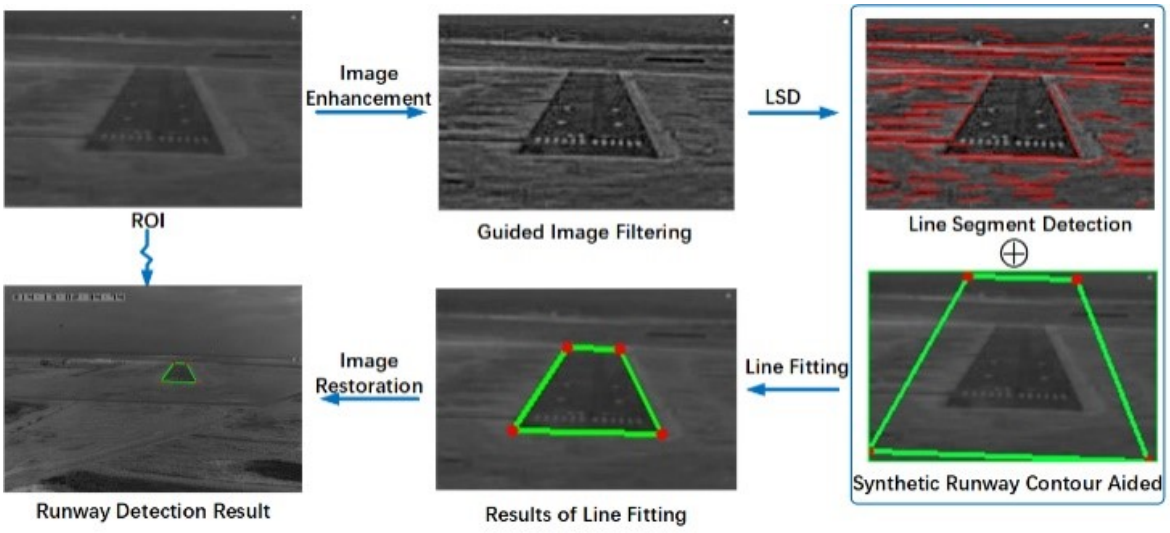


Figure 6 – Inertia aided runway features extraction scheme

2.3 Relative Pose Estimation

Position and attitude solution algorithm is the core of visual-inertial pose estimation. The accurate and efficient measurement of the position and attitude is premise to land successfully. Precise pose estimation has great effects upon (1) Autonomous landing for fixed-wing UAV. (2) Rotorcraft landing. (3) Landing assistance system of manned fixed-wing aircraft.

2.3.1 EPnP Principle

There is now a general consensus that EPnP algorithm is one of the most efficient way for camera pose estimation, as shown in figure 7.

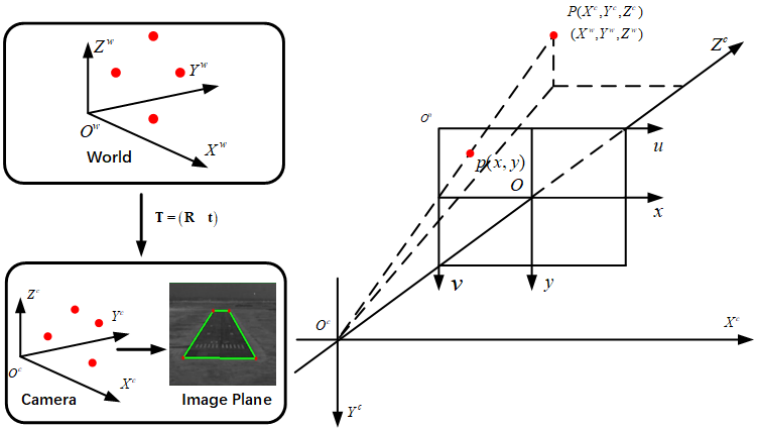


Figure 7 – EPnP principle

F^c stands for camera frame, and F^w stands for world frame. So, for every point p, we define control point p_i^w, p_i^c as the world coordinate and camera coordinate, respectively.

$$\mathbf{p}_i^w=\sum_{j=1}^{4}a_{ij}\mathbf{c}_j^w,\ \ with\sum_{j=1}^{4}a_{ij}=1,\mathbf{p}_i^c=\sum_{j=1}^{4}a_{ij}\mathbf{c}_j^c,\ \ with\sum_{j=1}^{4}a_{ij}=1 \tag{2}$$

where the a_{ij} is homogeneous barycentric coordinate. Because the equations are underdetermined, a solution exists.

Theoretically, any control point is acceptable. Considering the stability of algorithm, we choose central point of control points, other points are main axis' unit length points, derived from principal component analysis (PCA). So, we can translate camera coordinates into image coordinates by pinhole camera model.

$$\forall i,w_i\begin{bmatrix}u_i\\v_i\\1\end{bmatrix}=\begin{bmatrix}f_x&0&u_0\\0&f_y&v_0\\0&0&1\end{bmatrix}\sum_{j=1}^{4}a_{ij}\begin{bmatrix}x_j^c\\y_j^c\\z_j^c\end{bmatrix} \tag{3}$$

The last row of (3) implies $w_i = \sum_{j=1}^{4} a_{ij} c_j^w z_j^c$, substituting this expression in the first two rows yields two linear equations for each control point:

$$\sum_{j=1}^{4} a_{ij} f_x x_j^c + a_{ij}(u_c - u_i) z_j^c = 0 \tag{4}$$

$$\sum_{j=1}^{4} a_{ij} f_y y_j^c + a_{ij}(v_c - v_i) z_j^c = 0 \tag{5}$$

We rewrite Eq. (4) and (5) in matrix form.

$$\mathbf{M}\mathbf{x} = 0, \quad \text{with} \quad \mathbf{x} = \left[\mathbf{c}_1^{cT}, \mathbf{c}_2^{cT}, \mathbf{c}_3^{cT}, \mathbf{c}_4^{cT} \right]^T \tag{6}$$

where $\mathbf{M}$ is a *2n×12* matrix, generated by arranging the coefficients of (4) and (5) for each control point, and $\mathbf{x}$ is a 12-vector made of the unknowns, consisting of coordinates of control points, which can be calculated by solving the zero-space eigen value of $\mathbf{M}^{\mathrm{T}}\mathbf{M}$ with computational complexity of $O(n)$.

$$\mathbf{x} = \sum_{i=1}^{N} \beta_i \mathbf{v}_i \tag{7}$$

where v_i are the columns of the right-singular vectors of **M** corresponding to the null singular values of **M**.

Specifically, finding optimal $\mathbf{R}/\mathbf{t}$ is implemented through Gauss-Newton algorithm to minimize $\boldsymbol{\beta} = [\beta_1, \beta_2, \beta_3, \beta_4]^T$, objective function is

$$Error(\boldsymbol{\beta}) = \sum_{(i,j)s.t.i<j} \left(\left\| \mathbf{c}_i^c - \mathbf{c}_j^c \right\|^2 - \left\| \mathbf{c}_i^w - \mathbf{c}_j^w \right\|^2 \right) \tag{8}$$

In the preceding step, we get zero-space errors of different dimension and choose minimum error of $\boldsymbol{\beta}$, then calculate $\mathbf{x}$ accordingly. Therefore, we can recovery 3D points in camera coordinate, and figure out control points' camera coordinates according to the homogeneous barycentric coordinates. Furthermore, coordinate transformation from world coordinate frame to camera coordinate frame can be calculated by known coordinates in two coordinate systems.

The method has the attractive property of efficiency, but a small number of reference points in the algorithm has a direct impact on the accuracy for pose estimation. Therefore, we calculate deviation of runway corners between real image pixel and project position increase the reference points to improve accuracy and stability by using offset information, and it has a little impact on efficiency.

2.3.2 Aircraft Pose Estimation by EPnP

According to the projection model, camera coordinate of point D is:

$$\begin{bmatrix} x_O^c \\ y_O^c \\ z_O^c \end{bmatrix} = \mathbf{R}_m^f \cdot \mathbf{R}_b^m \cdot \mathbf{R}_n^b \cdot \mathbf{R}_{eI}^t \cdot \begin{bmatrix} x_O^e - x_I^e \\ y_O^e - y_I^e \\ z_O^e - z_I^e \end{bmatrix} + \mathbf{R}_m^f \cdot (\mathbf{t}_{IMU}^m - \mathbf{t}_{camera}^m) \tag{9}$$

$$\begin{bmatrix} x_O^e \\ y_O^e \\ z_O^e \end{bmatrix} = \mathbf{R} \cdot \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + \mathbf{t} = \mathbf{t} = \mathbf{R}_m^f \cdot \mathbf{R}_b^m \cdot \mathbf{R}_n^b \cdot \mathbf{R}_{eI}^t \cdot \begin{bmatrix} x_O^e - x_I^e \\ y_O^e - y_I^e \\ z_O^e - z_I^e \end{bmatrix} + \mathbf{R}_m^f \cdot (\mathbf{t}_{IMU}^m - \mathbf{t}_{camera}^m) \tag{10}$$

where $\mathbf{R}_{eI}^t$ is coordinate transformation with IMU's center as origin from the ECEF frame to the tangent plane frame.

$$\begin{bmatrix} x_D^c \\ y_D^c \\ z_D^c \end{bmatrix} = \mathbf{R}_m^f \cdot \mathbf{R}_b^m \cdot \mathbf{R}_n^b \cdot \mathbf{R}_{eI}^t \cdot \begin{bmatrix} x_D^e - x_I^e \\ y_D^e - y_I^e \\ z_D^e - z_I^e \end{bmatrix} + \mathbf{R}_m^f \cdot (\mathbf{t}_{IMU}^m - \mathbf{t}_{camera}^m) \tag{11}$$

Equally, camera coordinate of point D can be expressed by $\mathbf{R}$ and $\mathbf{t}$:

$$\begin{bmatrix} x_D^c \\ y_D^c \\ z_D^c \end{bmatrix} = \mathbf{R} \cdot \mathbf{R}_{eO}^t \begin{bmatrix} x_D^e - x_O^e \\ y_D^e - y_O^e \\ z_D^e - z_O^e \end{bmatrix} + \mathbf{t} = \mathbf{R} \cdot \begin{bmatrix} x_D^t \\ y_D^t \\ z_D^t \end{bmatrix} + \mathbf{t} \tag{12}$$

where $\mathbf{R}_{eO}^t$ is coordinate transformation with point $\mathbf{O}$ as origin from ECEF coordinate system to tangent plane coordinate system.

$$\begin{bmatrix} x_D^c \\ y_D^c \\ z_D^c \end{bmatrix} = \mathbf{R} \cdot \mathbf{R}_{eO}^t \begin{bmatrix} x_D^e - x_O^e \\ y_D^e - y_O^e \\ z_D^e - z_O^e \end{bmatrix} + \mathbf{R}_m^f \cdot \mathbf{R}_b^m \cdot \mathbf{R}_n^b \cdot \mathbf{R}_{eI}^t \begin{bmatrix} x_O^e - x_I^e \\ y_O^c - y_I^e \\ z_O^e - z_I^e \end{bmatrix} + \mathbf{R}_m^f \cdot (\mathbf{t}_{IMU}^m - \mathbf{t}_{camera}^m) \tag{13}$$

(12) and (13) are physically equivalent, so

$$\mathbf{R} \cdot \mathbf{R}_{eO}^t \begin{bmatrix} x_D^e - x_O^e \\ y_D^e - y_O^e \\ z_D^e - z_O^e \end{bmatrix} + \mathbf{R}_m^f \cdot \mathbf{R}_b^m \cdot \mathbf{R}_n^b \cdot \mathbf{R}_{eI}^t \begin{bmatrix} x_O^e - x_I^e \\ y_O^c - y_I^e \\ z_O^e - z_I^e \end{bmatrix} = \mathbf{R}_m^f \cdot \mathbf{R}_b^m \cdot \mathbf{R}_n^b \cdot \mathbf{R}_{eI}^t \begin{bmatrix} x_D^e - x_I^e \\ y_D^c - y_I^e \\ z_D^e - z_I^e \end{bmatrix} \tag{14}$$

where $\mathbf{R}_m^f \cdot \mathbf{R}_b^m \cdot \mathbf{R}_n^b = \mathbf{R}_n^f$, and

$$\mathbf{R} \cdot \mathbf{R}_{eO}^t \begin{bmatrix} x_D^e - x_O^e \\ y_D^e - y_O^e \\ z_D^e - z_O^e \end{bmatrix} + \mathbf{R}_n^f \cdot \mathbf{R}_{eI}^t \begin{bmatrix} x_O^e - x_I^e \\ y_O^c - y_I^e \\ z_O^e - z_I^e \end{bmatrix} = \mathbf{R}_n^f \cdot \mathbf{R}_{eI}^t \begin{bmatrix} x_D^e - x_I^e \\ y_D^c - y_I^e \\ z_D^e - z_I^e \end{bmatrix} \tag{15}$$

$$(\mathbf{R} \cdot \mathbf{R}_{eO}^t - \mathbf{R}_n^f \cdot \mathbf{R}_{eI}^t) \begin{bmatrix} x_D^e - x_O^e \\ y_D^e - y_O^c \\ z_D^e - z_O^e \end{bmatrix} = \mathbf{0} \tag{16}$$

The equality (16) follows in condition of $\mathbf{R} \cdot \mathbf{R}_{eO}^t - \mathbf{R}_n^f \cdot \mathbf{R}_{eI}^t = \mathbf{0}$. There is some difference in longitude and latitude between point O and center of IMU. Approximately, we assume that $\mathbf{R}_{eO}^t = \mathbf{R}_{eI}^t$.

$\mathbf{R} = \mathbf{R}_n^f = \mathbf{R}_m^f \cdot \mathbf{R}_b^m \cdot \mathbf{R}_n^b$, attitude matrix $\mathbf{R}_n^b = (\mathbf{R}_b^m)^T \cdot (\mathbf{R}_m^f)^T \cdot \mathbf{R}$, therefore, angles of roll, pitch and yaw can be calculated by decomposing matrix $\mathbf{R}_n^b$. The position of IMU in tangent plane coordinate system is:

$$\begin{bmatrix} x_I^t \\ y_I^t \\ z_I^t \end{bmatrix} \approx \mathbf{R}_{eI}^g \cdot \begin{bmatrix} x_I^e - x_O^e \\ y_I^e - y_O^e \\ z_I^e - z_O^e \end{bmatrix} = (\mathbf{R}_n^f)^T \cdot (-\mathbf{t} + \mathbf{R}_n^f \cdot (\mathbf{t}_{IMU}^m - \mathbf{t}_{camera}^m)) \tag{17}$$

3. Experiments and Result

Flight test data acquired by a modified Y12 (a general aviation fixed-wing aircraft) from Pucheng airport is used to verify our method for precise pose estimation. All experiments are performed on a computer with Intel Core i7-8550U processor clocked at 2.00GHz and 16.0 GB memory. All codes are run in Visual Studio 2010 with OpenCV 3.40.

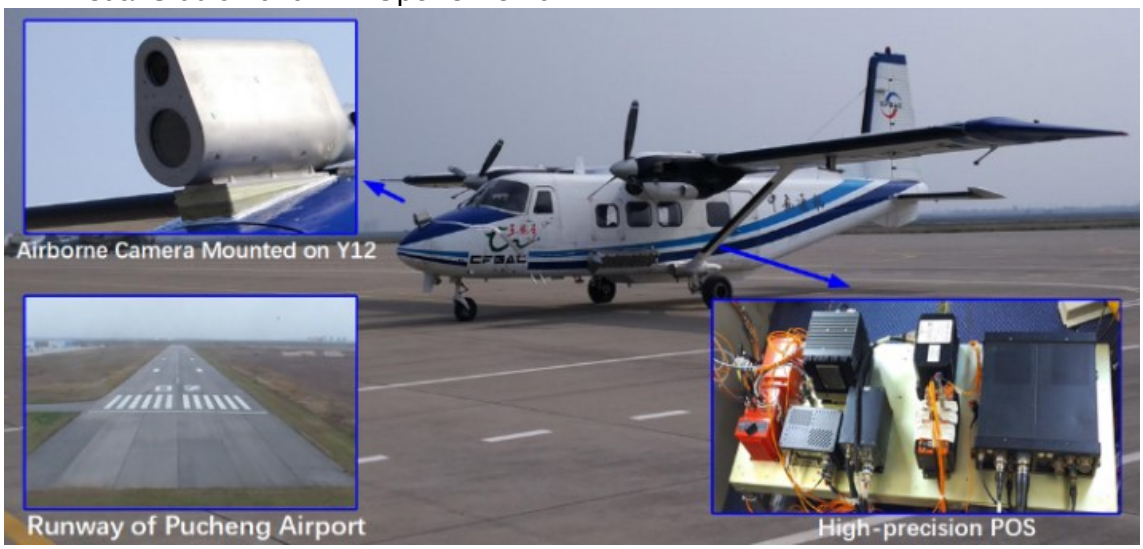


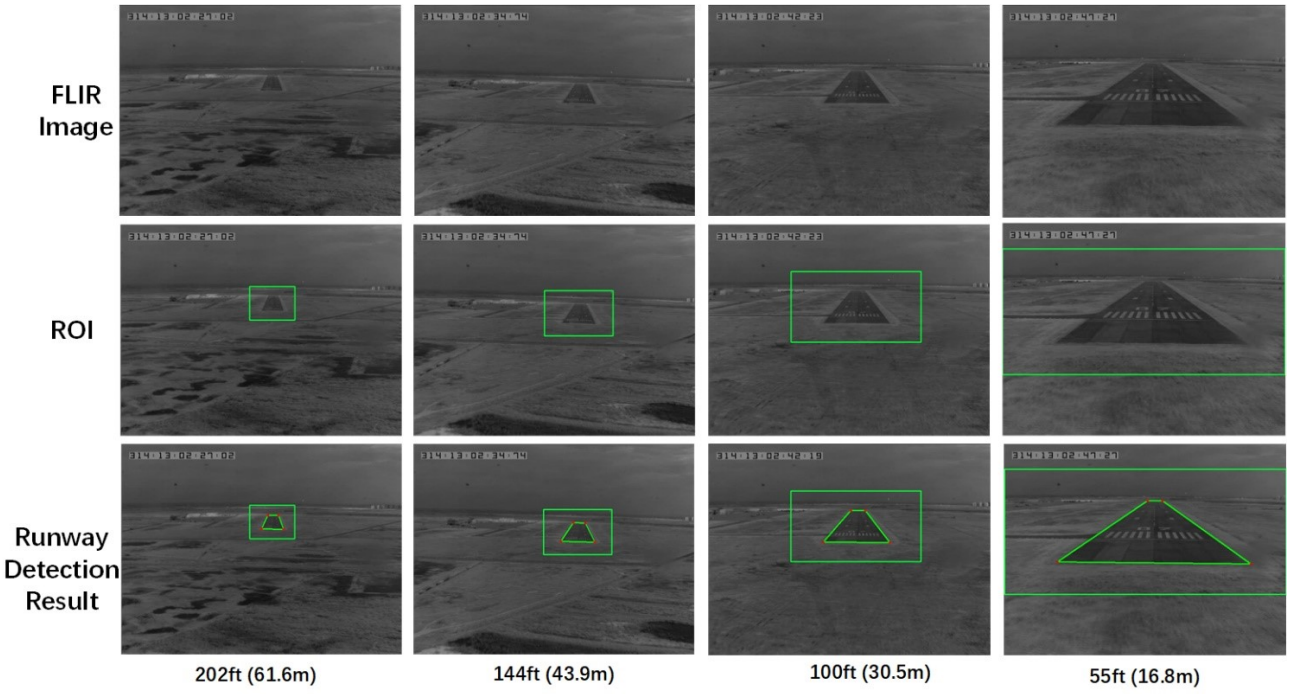
Figure 8 – Instruments mounted byY12 and runway for data acquisition

It is discussed in articles[19][20] that small number of reference points will cause large deviation for pose estimation. The effect of number of reference points on accuracy in our system is show in table 1. To improve the stability and precision of EPnP, 16 points in Pucheng airport have been measured precisely.

Table 1 The estimation error for **different number of reference points**

Estimation error(%)	8 points	10 points	13 points	16 points
Distance error (200ft)	0.05	0.03	0.03	0.03
Distance error (100ft)	0.48	0.42	0.42	0.39
Distance error (60ft)	1.75	2.16	1.98	2.65
Rotation error (200ft)	0.14	0.14	0.14	0.14
Rotation error (100ft)	0.47	0.42	0.42	0.40
Rotation error (60ft)	1.49	1.49	1.48	1.48

We obtain pixel coordinates of measured points by calculating offset of runway corners between detected corners and projective corner points by inertial parameters, then accurate position and attitude of aircraft is achieved.



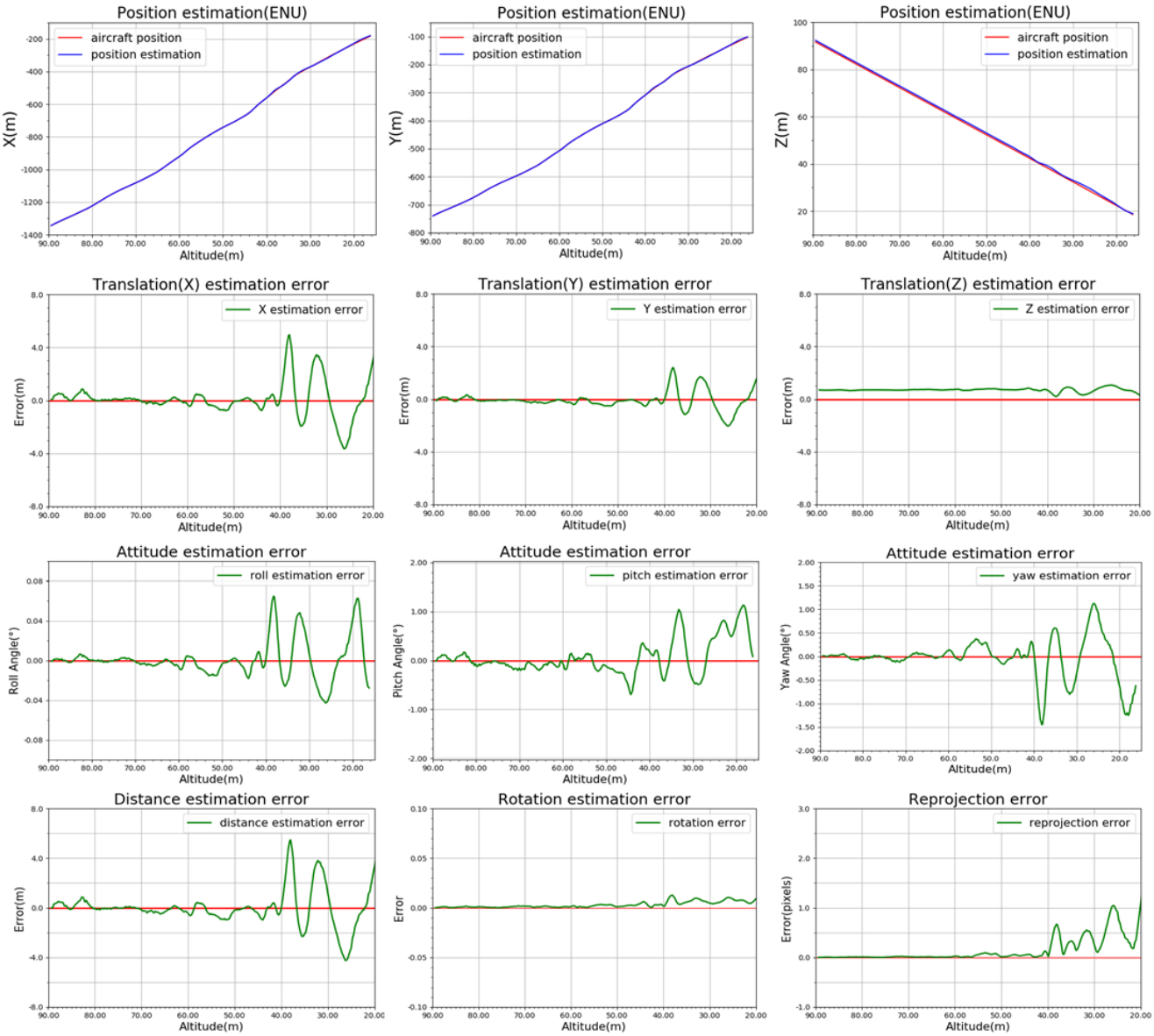


Figure 9 – Runway detection result and pose estimation error in Local East-North-Up (ENU) frame

As shown in figure 9, the experimental result indicates that the proposed method has advantages in both runway detection and pose estimation, errors of translation and distance are within 1.0 m at the altitude of 40m-90m.But the accuracy of the pose estimation has fluctuated below 40m due to runway detection error. Although line segments detection is more precise at a lower elevation, interferences also occur increasingly, which have brought on difficulty in runway corner detection. And the errors are amplified in small-scale, any minimal disturbance leads to inaccuracy of pose estimation. By contraries, features change is marginal on larger scales.

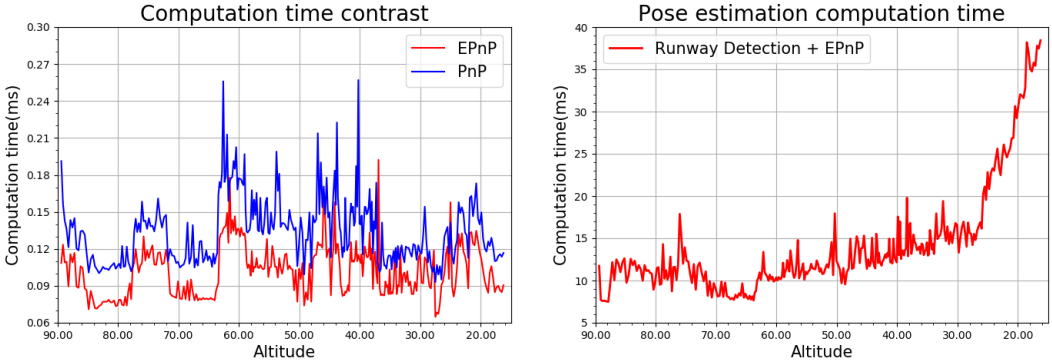


Figure 10 – Time consume

In addition, comparison of time consume with 16 feature points between EPnP and DLS PnP[21] leads to the conclusion that accuracies of two method are equivalent, but EPnP algorithm can save more than 25.08%computation time on average. Besides, total time of runway detection and pose estimation shown in figure 13, which is related to the size of ROI. The ROI size arises with the altitude decreased, and features of ROI increase resulting in more time consume for line segments detection and line fitting.

4. Conclusions and Future Work

In this paper, a real-time and accurate visual-inertial pose estimation for fixed-wing aircraft landing is proposed. The improved algorithm implements runway detection in real time by adjusting ROI with altitude, optimizing line segments filter and line fitting with endpoints. Furthermore, Increasing the number of reference points artificially compensates weakness that EPnP is inaccurate with less feature points. The experimental results indicate that the proposed method has advantages in both the real-time and the robustness, which takes full advantages of the airborne navigation information to estimate aircraft pose.

In future works, the authors will use deep learning methods to extract runway feature. Besides, projection error can be calculated, further reducing the ROI size and reducing computation time. Furthermore, the proposed method will be applied to multithreading and run on the embedded computer to test its efficiency.

5. Copyright Statement

References

[1] Kerns, A. J., Shepard, D. P. Unmanned aircraft capture and control via GPS spoofing. *Journal of Field Robotics*, Vol. 31, No. 4, pp 617-636, 2014.

[2] Mur-Artal R, Tardós J D. ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras. *IEEE Transactions on Robotics*, Vol. 33, No. 5, pp 1255-1262, 2017.

[3] Engel J, Schöps T, Cremers D. LSD-SLAM: Large-Scale Direct Monocular SLAM. *European conference on computer vision (ECCV)*, Zurich, Switzerland, Vol. 1, 8690, pp 834-849, 2014.

[4] Tong Qin, Peiliang Li, Shaojie Shen. VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator. *IEEE Transactions on Robotics*, Vol 34, No.4, pp 1004-1020, 2018.

[5] Engel J, Koltun V, Cremers D. Direct Sparse Odometry. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol. 40, No. 3, pp611-625, 2016.

[6] Gui Y, Guo P, Zhang H, et al. Airborne Vision-Based Navigation Method for UAV Accuracy Landing Using Infrared Lamps. *Journal of Intelligent & Robotic Systems*, Vol. 72, No. 2, pp 197-218, 2013.

[7] Ruchanurucks M, Rakprayoon P, Kongkaew S. Automatic Landing Assist System Using IMU + PnP for Robust Positioning of Fixed-Wing UAVs. *Journal of Intelligent & Robotic Systems*, Vol. 90, No. 1-2, pp 1-11, 2017.

[8] Abu-Jbara K, Alheadary W, Sundaramorthi G, et al. A robust vision-based runway detection and tracking algorithm for automatic UAV landing. *International Conference on Unmanned Aircraft Systems (ICUAS)*, Denver, CO, USA, Vol.1, pp 1148-1157, 2015.

[9] Henriques J F, Rui C, Martins P, et al. High-Speed Tracking with Kernelized Correlation Filters. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol. 37, No. 3, pp 583-596, 2015.

[10]Zhang L, Cheng Y, Zhai Z, et al. Real-time accurate runway detection based on airborne multi-sensors fusion. *Defence Science Journal*, Vol. 67, No. 5, pp 542-550, 2017.

[11]Britting K R. *Inertial navigation systems analysis*. 1st edition, Wiley-Interscience, 1971.

[12]HE K, SUN J, TANG X. Guided image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 35, No. 6, pp 1397-1409, 2013.

[13]Gioi R G V, Jakubowicz J, Morel J M, et al. LSD: A Fast Line Segment Detector with a False Detection Control. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol. 32, No. 4 pp 722-732, 2010.

[14]Illingworth J, Kittler J. A survey of the Hough transform. *Computer Vision Graphics & Image Processing*, Vol. 43, No. 2, pp 280-280, 1988.

[15]Li C, Wang Z, Li L. An Improved HT Algorithm on Straight Line Detection Based on Freeman Chain Code. *International Congress on Image and Signal Processing*.Tianjin, China,Vol. 1 pp 1-4, 2019.

[16]Desolneux A, Moisan L, Morel J M. Meaningful Alignments. *International Journal of Computer Vision*, Vol. 40, No. 1, pp 7-23, 2000.

[17]Desolneux A, Moisan L, Morel J M. *From Gestalt Theory to Image Analysis*. 1st edition, Springer New York, 2008.

[18]Xue G, Song L, Sun J. Foreground Estimation Based on Linear Regression Model With Fused Sparsity on Outliers. *IEEE Transactions on Circuits & Systems for Video Technology*, Vol. 23, No. 8, pp 1346-1357, 2013.

[19]Lepetit V, Moreno-Noguer F, Fua P. EP n P: An Accurate O(n) Solution to the PnP Problem. *International Journal of Computer Vision*, Vol. 81, No. 2, pp 155-166, 2009.

[20]Li S, Xu C, Xie M. A Robust O(n) Solution to the Perspective-n-Point Problem. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, Vol. 34, No. 7, pp 1444-1450, 2012.

[21]Hesch, Joel A., and Stergios I. Roumeliotis. A direct least-squares (DLS) method for PnP. *International Conference on Computer Vision(ICCV)*. Barcelona, Spain, Vol.1, pp 383-390, 2011.