Research on solving heading attitude of airdrop cargo platform based on line features

Xia Li1,2, Bin Zhang1,2, Hongying Zhang3, Ronghua Xu4 and Yalei Bai3

Abstract
The present study envisages the development of an improved line features method to accurately estimate the attitude of the airdrop cargo platform during airdrop landing. Therefore, this article uses the geometric characteristics of the line features to improve the traditional line features extraction and removes the locally dense line features in the image, which greatly reduces the number of line features in the image. Then, the improved random sample consensus is used to remove the mismatching of line features, which improves the real-time performance of the algorithm and the accuracy of the attitude angle, and makes up for the problem of difficult extraction of point features or low matching accuracy in the airdrop environment. Finally, a constraint equation is established for the line features that are successfully matched, and using homography to obtain attitude of the airdrop cargo platform. This article also meets the requirements of accurate calculation attitude of airdrop cargo platform. The experiment shows the significance and feasibility of the airdrop cargo platform heading and attitude calculation technology based on the line feature, and it has a good application prospect.

Keywords
Airdrop cargo platform, line features, attitude calculation, aircraft landing

Date received: 18 November 2021; accepted: 3 February 2022

Topic: Vision Systems
Topic Editor: Antonio Fernandez-Caballero
Associate Editor: Shengyong Chen

Introduction
With the development in the large-scale military transport aircrafts, the airdrop of heavy equipment and materials is the basic capability of large-scale transport aircraft. Airdrop equipment uses airdrop cargo platform as carrier. To achieve the smooth landing of the airdrop cargo platform, it is important to calculate the three-axis attitude angle of the airdrop cargo platform in real time to avoid the overturning accidents.1–4 Attitude estimation is to measure the position and attitude of the aircraft through one or more sensors such as inertial measurement unit (IMU), global positioning system (GPS), vision, and laser. Research scholars have

1 Tianjin Key Laboratory for Advanced Mechatronic System Design and Intelligent Control, School of Mechanical Engineering, Tianjin University of Technology, Tianjin, China
2 National Demonstration Center for Experimental Mechanical and Electrical Engineering Education, Tianjin University of Technology, Tianjin, China
3 College of Aerospace Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China
4 Aviation Industry Aerospace Lifesaving Equipment Limited Liability Company, Hubei, China

Corresponding author:
Hongying Zhang, College of Aerospace Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China.
Email: zhanghy19760406@163.com
done a lot of research on the problem of multi-sensor data fusion to estimate the aircraft’s attitude. With the development of computer vision, it has attracted the attention of research scholars to estimate the attitude of aircraft by vision. During the landing of the airdrop cargo platform, haze or clouds may appear at high altitudes, the clarity of the collected images is reduced, and the ground point features are masked and difficult to identify. However, the line features can still be identified according to the local tracking direction in the images, and the line features can better represent the image information. Therefore, the line features are more stable in complex environment, so as to the algorithm based on line features has attracted more and more attention.\(^{12-15}\) Xu et al.\(^{16-18}\) independently estimated the landing attitude of unmanned aerial vehicle (UAV) by using the actual shape of the ground cooperative target and shadow line features and verified that the attitude estimation using line features has higher accuracy and reliability than point features. Li et al.\(^{19}\) proposed a point-line stereo visual simultaneous localization and mapping (SLAM) system with semantic invariants. The system improved the accuracy of line features matching by fusing line features with image semantic invariant information. Yu et al.\(^{18}\) proposed a non-iterative method to solve the perspective-n-line problem. Firstly, the objective function of nonlinear least squares was first expressed by Cayley as the parameterized rotation matrix. A third-order equation was derived from the optimal conditions, and then the camera attitude was directly solved based on Grobner technique. Zuo et al.\(^{20}\) tried to use Plücker coordinates to represent the line features in the front-end algorithm of the SLAM system and to express parameterized line features in the back-end optimization algorithm by minimum orthogonality. Pumarola et al.\(^{21}\) added line features on the basis of ORB-SLAM, and put forward the method of system initialization by using line features. Wang et al.\(^{22}\) added the line angle to the construction of line features re-projection error. He et al.\(^{23}\) put forward point and line features visual-inertial odometry (PL-VIO) system by adding line features on the basis of visual inertial navigation system (VINS)-Mono. Gomez-Ojeda et al.\(^{24}\) proposed to use geometric constraints to construct a method of minimizing the L1 norm to complete line feature matching. Compared with the traditional method based on descriptors, this method improved the tracking efficiency and matching quality of line feature. Wang et al.\(^{24,25}\) used point and line features to fuse in SLAM and established two projection models to estimate the motion state of the camera.

For the attitude acquisition of airdrop cargo platform, only the obvious features in the scene need to be extracted as the ground reference, and there is no need to extract all the features in the scene carefully. Using line segment detector (LSD)\(^{26}\) and line band descriptor (LBD)\(^{27}\) to detect line features in the image is a more mainstream method. LSD directly groups adjacent pixels into line segment regions according to the consistency of pixel gradient directions and can obtain sub-pixel-level precision detection results in linear time. Although LSD has achieved good results in terms of speed and accuracy, it also has its own shortcomings. Because it does not have a screening and merging mechanism when extracting features, it collects a large number of similar line features in local dense areas of the image, and it is prone to over-segmentation of line features. Too many short line features greatly increase the time for line features detection and matching and also increase the probability of mismatching line features, which inevitably affect the operation efficiency of algorithm.

**System overview**

To reduce the number of line features, gradient density filter is used to eliminate local dense gradient regions to avoid a large number of invalid features, thus improving the efficiency of features extraction and matching, and reducing the computational complexity and mismatching rate of the algorithm. The improved LSD restricts the length of line features, which improves the quality of extracted line features, increases the probability of the same features appearing between adjacent frames, and saves the time of detecting too many short line features. The improved LSD groups the line features that have not been eliminated, and then roughly screens them according to the length, angle, horizontal and vertical distance, and divides the adjacent line features into a group. In the line segment merging stage, line features angle, endpoints of line, and angle are introduced as merging standards, and line features merging is completed. The LBD matches the line features. The same line has corresponding geometric relationship (parallelism, length approximation, and overlap) in the matching of two adjacent frames. This relationship can effectively reduce the mismatching rate of line features and save the calculation time of random sample consensus (RANSAC) to eliminate mismatching. Finally, a constraint equation is established for the successfully matched line features, and the homography matrix is solved and then to obtain the triaxial attitude angle of the airdrop cargo platform. The system overview is shown in Figure 1.

**Eliminate local dense line features**

Line features usually appear in dense gradient areas. If the number of pixels with high gradient in the target area is too high, it means that the target area is too dense, so it is a dense line features area, which is eliminated to reduce feature mismatch.

In the airdrop environment, there may be dense grass or soil texture on the ground. A large number of line features are detected in the dense texture areas. Excessive density of line features in the same area usually lead to
mismatching. To eliminate the line features of the local dense gradient region, the local gradient filter\textsuperscript{28} is used to filter out the region. According to different airdrop environments, the adaptively pixel gradient density threshold is set as the filtering standard.

The pixel gradient density is defined as the percentage of pixels with higher gradients in the unit pixel gradient area, which is used to measure whether the area is feature-intensive. Set the gradient value $IG_{ij}$ of pixel point $(i, j)$, when the gradient value $IG_{ij}$ is greater than or equal to the preset intensity threshold $IG$, the value $o_{ij}$ is marked as 1, otherwise it is 0

$$IG = \frac{\sum IG_{ij}}{(i-2) \times (j-2)}$$

$$o_{ij} = \begin{cases} 0 & |IG_{ij}| \leq IG \\ 1 & |IG_{ij}| > IG \end{cases}$$

The dense gradient of the local target area $r_{ij}$ is the sum of the values $\sum_{i=1}^{m} \sum_{j=1}^{m} o_{ij}$ in the $M \times M$ area centered on the pixel $(i, j)$, which represents the percentage of pixel points whose gradient is greater than the set threshold. Initially set $M = 5$, look for a small range of high-gradient dense areas in the image, and then gradually expand the value of $M$ until $M_l = 30$

$$r_{ij} = \frac{\sum_{i=1}^{M_l} \sum_{j=1}^{M_l} o_{ij}}{M_l \times M_l}$$

To reduce pixel areas with too high-gradient density, filter is performed to reduce the dense feature areas in the image. $\rho_{gd}$ is the average gradient density threshold of the whole image, and adaptive threshold $\rho_{gd}$ can be obtained according to different scenes to measure the line density of the whole image. The areas with gradient density $\rho_{ij}$ greater than $\rho_{gd}$ are regarded as invalid areas and the invalid areas are not processed. Only the line features areas where the gradient density $\rho_{ij}$ is less than $\rho_{gd}$ in the image are retained

$$\rho_{gd} = \sum_{i=1}^{n} \sum_{j=1}^{m} o_{ij} \quad \frac{i \times j}{i \times j}$$

$$GD_{ij} = \begin{cases} 1 & \rho_{ij} \leq \rho_{gd} \\ 0 & \rho_{ij} > \rho_{gd} \end{cases}$$

After the filter is added, all line features in the area are eliminated and replaced by the contour edges of the area, so as to avoid a large number of invalid features, which reduces the calculation cost, and improve system accuracy.

**Improved line features extraction algorithm**

The line features detected by LSD contain a large number of short line features, so the length of short line features needs to be constrained. Short line features are filtered out, the line features that occupy more pixels are retained, the remaining line features are grouped, and merge and connect the adjacent effective short line features in the group, as shown in Figure 2. These improved the quality of line features and made the line features evenly distributed in the image, so as to provide stable line features.

**Line features grouping**

Line features grouping needs to filter the length of the line feature, and then filter the angle and distance. The angle calculation of adjacent line features is relatively simple, so it is used as the first screening condition. The horizontal distance and vertical distance only involve the addition and subtraction of absolute values of formulas, which make the
screening process more efficient. The process is shown in Figure 3. Geometric relationship between Group GL1 inner line features and longest line feature L1 is shown in Figure 4.

Line features length screening. Long line features could be detected by multi-frame images and have a longer life span. Moreover, the length of the line features between two adjacent frames is almost unchanged, which is considered as a more stable feature. The basis for screening the length of line features is constrained according to the size of different images collected by airdrop cargo platform, the depth of scenes, and the number of extracted line features and sets an appropriate threshold to constrain the length of the line feature. The difference in image size determines the length of effective line features. The higher the resolution of images, the more pixels should be included. Remove short line features based on the number of pixels contained in the line feature. Set \( L_{\text{min}} \) as the lower limit of the length of the line features in the detection image, the length and width of the image are \( h_i \) and \( w_i \), respectively, and take the shortest side of the image size as a reference to determine the length of different effective line features.

\( \eta \) is the ratio coefficient of the long line features in the entire image and is set to 0.245. Ensure that the number of extracted line features in each frame of image is between 20 and 300, so as to meet the line feature merging in the subsequent steps

\[
L_i \geq \left( L_{\text{min}} = \eta \times \min(h_i, w_i) \right) \quad i \in (1, k)
\]

Figure 2. Line features merging.

Line features angle screening. The line features retained after length filtering are sorted in descending order by the pixel length to obtain \( L_n = \{L_1, L_2, \ldots, L_n\} \), where \( L_n \) represents the \( n \)th line feature. Longer line features come from areas with large continuous gradient differences, which made the extracted line features more accurate and stable, so the line features grouping starts from the longest line feature \( L_1 \).

To express the line features conveniently, the line features in the image are represented by vectors. The starting point \( A_i(x_{Ai}, y_{Ai}) \) of a line features \( L_i \) points to the endpoints \( B_i(x_{Bi}, y_{Bi}) \), and the angle with the horizontal is \( \theta_i \). The starting point \( A_j(x_{Aj}, y_{Aj}) \) of a line features \( L_j \) points to the endpoints \( B_j(x_{Bj}, y_{Bj}) \), and the angle with the horizontal is \( \theta_j \). After angle screening, the candidate line features group \( L_n \) with close angles in the group is obtained as

\[
L_n = \{\forall L_i \in L : (|\theta_i - \theta_j| < \theta_s)\}
\]

\( \theta_s \) is the set angle threshold.

Line features plane distance screening. Screen the plane distance by the endpoints of different line features in the group. After horizontal distance screening, the candidate line features group \( L_{\text{hd}} \) are obtained as

\[
L_{\text{hd}} = \{\forall L_i, L_j \in L_n : \min \left( \frac{|x_{Ai} - x_{Aj}|}{|x_{Bi} - x_{Bj}|}, \frac{|x_{Ai} - x_{Aj}|}{|y_{Bi} - y_{Bj}|}, \frac{|y_{Ai} - y_{Aj}|}{|x_{Bi} - x_{Bj}|}, \frac{|y_{Ai} - y_{Aj}|}{|y_{Bi} - y_{Bj}|} \right) < d_h\}
\]

Through vertical distance screening, the candidate line features \( L_{\text{vd}} \) are obtained as

\[
L_{\text{vd}} = \{\forall L_i, L_j \in L_{\text{hd}} : \min \left( \frac{|y_{Ai} - y_{Aj}|}{|x_{Bi} - x_{Bj}|}, \frac{|y_{Ai} - y_{Aj}|}{|y_{Bi} - y_{Bj}|}, \frac{|x_{Ai} - x_{Aj}|}{|x_{Bi} - x_{Bj}|}, \frac{|x_{Ai} - x_{Aj}|}{|y_{Bi} - y_{Bj}|} \right) < d_v\}
\]

\( d_h \) is a screening threshold that measures the closeness of the horizontal distance and the vertical distance of the line feature. We set the distance threshold \( d_h \) to 2 to obtain a line segment group \( G_{L_1} = \{L_2, L_3, \ldots, L_n\} \) that is close to the long line feature \( L_1 \), so that the filtered
Line segment is guaranteed to be near $L_1$, which provides high-quality line features for the subsequent merging of line features.

**Line features merging**

After line features grouping, we get a line features grouping $G_{L_1}$, that matches the longest line features $L_1$ in the image. It is necessary to merge the line features in this grouping to reduce the number of line features in the image and improve the quality of line features. In the process of merging, further screening of line features in the group are required. The screening steps included angle correction of line features, endpoint screening, merging, and finally rechecking the angle of the newly merged line features, as shown in Figure 5. The geometric relationship of merging L1 and Li into Lz is shown in Figure 6.

**Angle correction.** The angle deviation, distance and length between Group $G_{L_1}$ inner line features and $L_1$ are important factors for the new line features to be merged. These factors need to modify the angle threshold in $G_{L_1}$, add different weighting coefficients, balance the influence of different factors on the angle correction, and finally get the angle correction threshold

$$
\lambda_i = \frac{1}{u} |\theta_i - \theta_1| + \frac{1}{v} \frac{d_{ii}}{L_1} + \frac{L_i}{L_1}
$$

(10)

The distance between two line features is expressed as follows

$$
d_{ii} = \sqrt{\left(\frac{(x_{A_i} - x_{B_i})x_{B_i} + (x_{B_i} - x_{A_i})y_{B_i}}{\sqrt{(x_{B_i} - x_{A_i})^2 + (y_{A_i} - y_{B_i})^2}}\right)^2 + \left(\frac{y_{A_i} - y_{B_i}}{L_1}\right)^2}
$$

(11)

Figure 4. Adjacent line feature geometric relationship.

where $\frac{1}{u}$ is the weighting coefficient of the angle of two line features, $0 < u < 1$, $u = \frac{\theta_1}{\theta}$ is the proportional relationship between the included angle of the new composite line feature $\theta_i$ and $\theta_1$, $\frac{1}{v}$ is the weighting coefficient of the distance between two line features, $0 < v < 1$, $v = \frac{d_{ii}}{L_1}$ is the proportional relationship between the longest line feature length $L_1$ and the distance between two line features. Through the analysis of the above factors, the angle correction value of the line features grouping $G_{L_1}$ relative to the long line features is

$$
\delta = 1 - \frac{1}{1 + e^{-\frac{5\lambda_i}{7}}}
$$

(12)

$\delta$ is the adaptive proportional coefficient function, where equation (12) is fitted by experimental curve data. It can be seen from the formula that the smaller the angle between the line features to be merged and $L_1$, the shorter the length, the smaller the distance to $L_1$, the larger the adaptive coefficient $\delta$, the greater the merge-ability. $\theta_2$ is the angle screening threshold that measures the similarity of the line feature angles, and its value is set to $\pi/90$ through experimental analysis. After angle screening, candidate line feature groupings $G_{L_{1,i}}$ are obtained as

$$
G_{L_{1,i}} = \{\forall L_i \in G_{L_{1}} : (|\theta_i - \theta| \leq \delta \* \theta_2)\}
$$

(13)

**Endpoint screening.** Restore $L_1$ into the screening line feature group $G_{L_{1,i}}$, to become a new line feature group $G_i = \{L_1, G_{L_{1,i-1},L_{1-i}}\}$ and filter the endpoints of all line features in the group. First, select the line features $L_1$ closest to $L_1$ and calculate the average position of the endpoints of the two line features, so as to obtain the head endpoints $(x_{A_1}, y_{A_1})$ and the tail endpoints $(x_{B_1}, y_{B_1})$. Calculate and compare the distance between all the endpoints and the origin, and select the minimum for the first endpoint, and the maximum for the endpoints, so as to obtain the endpoints of the synthesized new line feature. According to experience, at least one of the endpoints of the newly synthesized line features comes from the longest line features.

Figure 5. Merging of line features.
\[
(x_t, y_t) = \min \left\{ \sqrt{x_{A1}^2 + y_{A1}^2}, \sqrt{x_{A2}^2 + y_{A2}^2}, \right\}
\]

\[
(x_b, y_b) = \max \left\{ \sqrt{x_{B1}^2 + y_{B1}^2}, \sqrt{x_{B2}^2 + y_{B2}^2}, \right\}
\]

Check the characteristic included angle of synthetic line. The angle between the newly synthesized line feature \(L_Z\) and the horizontal direction is \(\theta_Z\). The absolute angle difference between the angle \(\theta_Z\) and the included angle \(\theta_i\) is less than or equal to \(\frac{\theta_i}{2}\), it is considered that the composite line feature is successfully merged and replaces the positions of the original line features \(L_i\) and \(L_1\). If the absolute difference of the angles is greater than \(\frac{\theta_i}{2}\), it means that the synthesized line feature \(L_Z\) deviates from the longer line feature \(L_1\), so \(L_Z\) is discarded.

\[
G_{L_Z} = \left\{ \forall L_i \in G_{L_{1-i}} : |\theta_Z - \theta_i| \leq \frac{\theta_i}{2} \right\} \quad (16)
\]

Repeat the above steps until all line features in the group are filtered.

To verify the performance of this algorithm, the LSD and the improved algorithm are used to extract the number of line features in two different scenes for comparison, as shown in Figure 7. And for scene 1, the number of line features extracted from images with different heights is counted, which proves the effectiveness of improved algorithm, as shown in Table 1.

It can be seen from Figure 7 that there are more long line features (roads and riversides) in scene 1 and dense weeds
in scene 2. In the two scenarios, this article uses the line geometric characteristics to filter the line features, which could merge the shorter lines into long line features, and at the same time uses the local dense filter to filter out the short line features around them (such as the box area of Figure 7(e) and (f)), which effectively reduces the number of line features extracted from the image, and makes the line features more stable. Although the line features are still discontinuous in the image, it is due to the removal of too many dense short line features in this area, which causes the distance between similar long line features and could not be merged. However, the merged long line features are stable enough in the image to match successfully and effectively improve the accuracy of the heading attitude of the airdrop cargo platform.

According to Table 1, the number of line features extract by improved algorithm of different heights has decreased, and the maximum decrease relative to the number of line features extract by LSD can reach 27.24%. When the

---

**Table 1.** Features extraction quantity of different height lines.

| Height from ground (m) | LSD line features number | The improved algorithm line features number | Compared with LSD, improved algorithm reduces the number of line features (%) |
|------------------------|--------------------------|------------------------------------------|--------------------------------------------------------------------------------|
| 30                     | 1101                     | 889                                      | 19.26                                                                         |
| 60                     | 1248                     | 923                                      | 26.04                                                                         |
| 90                     | 1315                     | 1025                                     | 22.05                                                                         |
| 125                    | 1196                     | 891                                      | 24.96                                                                         |
| 160                    | 1270                     | 924                                      | 27.24                                                                         |

LSD: line segment detector.
The accuracy of airdrop cargo platform attitude estimation based on line features depends on the accuracy of features matching. If there are many wrong matches, there are errors in the attitude calculation of the airdrop cargo platform. Therefore, before the attitude calculation of the airdrop cargo platform, the RANSAC\textsuperscript{29,30} is used to remove the wrong matching in the image.

RANSAC eliminates outliers through step-by-step iteration. If there are too many wrong matching line features in the observed data, the iteration accuracy decreases. The geometric constraints between the line features are used to prescreen out the line features that could not be matched successfully. The movement of the airdrop camera between two adjacent frames is small, and the constraint is completed by the geometric characteristics between the line features. For example, the position, length, and angle of the line features of two adjacent frames should be approximately the same.

We use the parallel relationship between the two line features, the length approximation and the overlap to eliminate the mismatch of line features, so as to remove the unmatched line features in advance and reduce the number of iterations of the RANSAC.

Matching line features have a parallel relationship

\[ |\theta_a - \theta_i| \leq \alpha \]  

(17)

Matching line features have length similarity

\[ \rho_{uv} = \frac{\min(|L_u|, |L_v|)}{\max(|L_u|, |L_v|)} \]  

(18)

Matching line features have length overlap

\[ n, m = 1, 2, \ldots, +\infty \]  

(19)

To test whether the improved algorithm can effectively delete the mismatching, the line features of the image of shooting scene 1 are extracted and matched, as shown in Figure 8.

It can be seen from Figure 8 that the improved algorithm in this article can reduce the number of matching line features and effectively reduce mismatches during the matching process. Using LBD to directly match the screened effective line features, it can be seen that the line features are distributed in all directions. Figure 8(b) shows the result of using the RANSAC to remove mismatches. Most mismatches can be deleted, but there are a few mismatches with inconsistent line feature directions. In this article, the improved algorithm constrains the geometric characteristics of the line feature and further removes the mismatch. It can be seen from Figure 8(c) that the matched line feature pairs appear in the same direction, which greatly reduces the number of mismatches.

As shown in Table 2, because there are many line features in the scene, the matching accuracy without mismatching screening is 41.18\%, and the number of mismatching is large. The matching accuracy of RANSAC to delete mismatching is 60.43\%, and the matching accuracy rate after the geometric constraint screening can reach 75\%. The improved RANSAC algorithm can obtain accurate matching results, which greatly improves the accuracy of attitude calculation of airdrop cargo platform.
It can be further expressed as

\[ q^T x_{i+1} = 0 \quad q^T = [e \ f \ 1] \quad (23) \]

According to homography, the two line features should meet the constraints of \( \lambda x_{i+1} = H_L x_i \). The two ends of the equation are left multiplied by the normal vector \( q^T \)

\[ q^T \lambda x_{i+1} = q^T H_L x_i \quad (24) \]

\[ q^T x_{i+1} = \lambda^{-1} q^T H_L x_i \quad (25) \]

According to \( q^T x_{i+1} = 0 \), it can be known that formula (25) can be equal to

\[ \lambda^{-1} q^T H_L x_i = 0 \quad (26) \]

According to formulas (21) and (26), it can be deduced that

\[ \lambda^{-1} q^T H_L = p^T \quad (27) \]

The two ends of formula (27) are transformed to obtain

\[ H_L^T q = \lambda p \quad (28) \]
To eliminate the scale factor $\lambda$, cross multiply the two ends of formula (28) and get

$$p \times H^T q = 0$$ (29)

The homography $H$ directly describes the transformation between image coordinates $p_1$ and $p_2$, and the middle part is denoted as $H$. The definition of homography is related to rotation matrix, translation matrix, camera internal parameters, and plane parameters

$$p_2; K(RP + t)$$

$$: K\left(\begin{array}{c} RP + t \cdot \left(-\frac{n^T p}{d}\right) \end{array}\right)$$

$$: K\left(R - \frac{tn^T}{d}\right)p$$ (30)

$$: K\left(R - \frac{tn^T}{d}\right)K^{-1}p_1$$

$$: KHK^{-1}p_1$$

$$H = R - \frac{tn^T}{d}$$ (31)

Here $H_L$ is to describe the conversion relationship of the same line features in different frames $H_L = KHK^{-1}$. The camera internal parameters need to be included between the homography matrices:

$$(KHK^{-1})^T q = \lambda p$$ (32)

$$H^T K^T q = \lambda K^T p$$ (33)

Let $K^T q = q'$ and $K^T p = p'$, simplify

$$H^T q' = \lambda p'$$ (34)

Cross multiplication is available

$$p' \times H^T q' = 0$$ (35)

Expand the matrix $p'$, write it in vector form $p' = [p_1, p_2, 1]^T$, and expand the matrix $q'$ into vector form $q' = [q_1, q_2, 1]^T$ to get the constraint equation of line features

$$\begin{bmatrix} -p_2 & p_1 & q_1p_1 & q_2p_1 & q_2 \\ p_1 & p_2 & -q_1p_1 & -q_1p_2 & -q_1 \\ p_2 & p_1 & -q_2^2 & -p_2^2 & q_2 \\ -p_1 & p_2 & q_1^2 & p_2^2 & q_1 \\ p_1 & p_2 & q_2^2 & p_2^2 & -q_2 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ h_4 \\ h_5 \end{bmatrix} = 0$$ (36)

### Decompose homography

The homography $H$ is regarded as a vector, which is recovered by solving the linear equation of the vector. The homography $H$ can be decomposed to get the corresponding rotation matrix $R$ and translation vector $t$. According to the form of homography $H$, it can be expressed as

$$H = \begin{bmatrix} h_1 & -h_2 & h_3 \\ h_2 & h_1 & h_4 \\ 0 & 0 & h_5 \end{bmatrix}$$ (37)

There are five unknowns, and if there are more than five line features that are successfully matched, then each element of the homography $H$ can be solved by the above equation and decomposed to get the corresponding rotation matrix $R$ and translation vector $t$

$$R_L = \begin{bmatrix} h_1 & -h_2 & 0 \\ h_2 & h_1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$ (38)

$$t_L = [h_3, h_4, h_5, -1]^T$$ (39)

### Experiments

In this part, the experiment is to verify the effectiveness and real-timeness of the method for solving the heading attitude of airdrop cargo platform based on line features. The improved algorithm in this article, traditional line features LSD extraction, and LBD matching are compared with point features. And calculate the deviation between the estimated value output by the vision algorithm and the real value output by the IMU. Finally, the performance of the vision algorithm in terms of processing time is discussed in the Time performance experience section.

### Materials and equipment

Since the drop of the airdrop cargo platform requires a military assistance and is confidential, it is not convenient to disclose the test data. Therefore, the DJI Mavic Air 2 UAV is used to simulate the landing process of the airdrop cargo platform after the parachute is opened. The DJI MAVIC 2 UAV is equipped with a camera (model FC6310; DJ-Innovations), GPS/GLONASS dual-mode satellite positioning module, IMU, and barometric altimeter sensors. The angular rate sensor integrated in the IMU module uses the (model ADXRS620; Analog Devices), and the most important feature of this series of sensors is that it consists of vibration suppression and high-impact resistant properties. Furthermore, it has a full-range gyroscope of 300°/s, which could provide accurate UAV three-axis attitude values for the experiments. The altitude barometer and IMU output data are provided by DJI + Assistant software. The position of the IMU in the UAV relative to the camera.
Table 3. The parameters used in the experiments.

| Parameter                                      | Value                                                                 |
|------------------------------------------------|-----------------------------------------------------------------------|
| Camera intrinsic parameters                    | Focal length \( f_x = 3655.62 \) pixel, \( f_y = 3648.68 \) pixel |
| Principal point                                 | \( C_x = 354.571 \) pixel, \( C_y = 408.870 \) pixel               |
| Radial distortion                               | \( K_1 = -0.2407, K_2 = 0.1225 \)                                  |
| Camera resolution                               | 20 m: \( 1280 \times 720 \)                                         |
| Pixel size                                      | \( 2.41 \) \( \mu \) m                                            |
| Camera and IMU transformation matrix \( T_{\text{CAM-IMU}} \) | \[
\begin{bmatrix}
0.0124 & -0.8864 & 0.0365 & -0.012 \\
-0.8873 & -0.0186 & 0.0135 & -0.018 \\
-0.0159 & -0.0292 & -0.8878 & -0.058 \\
0 & 0 & 0 & 1
\end{bmatrix}
\] |

IMU: inertial measurement unit.

reference system is represented by the transformation matrix \( T_{\text{CAM-IMU}} \), as shown in Table 3. Since the current research is not fully implemented the integration of vision algorithms with flight control and is under further study, the following experimental procedure only captures the aerial images through the UAV onboard monocular camera, followed by simulation experiments by the computer. All the experiments are conducted on a computer equipped with an AMD Ryzen 7 4800H 4.20 GHz CPU (AMD, California, USA), and NVIDIA GeForce RTX2060 GPU (Nvidia, California, USA), and memory 16G devices. The CPU programs are executed in a single thread. The image processing is based on OpenCV3.3.0, which is programmed in Microsoft Visual Studio2015.

In the experiment, the image sequence of scene 1 is evaluated. Scene 1 contains a large number of linear features, such as riversides, roads, and house edges, which can provide a large number of stable linear features. During the landing of the UAV, data statistics are made on different landing heights and the experimental results are recorded. The roll \( \phi \), pitch \( \theta \), and yaw \( \psi \) of the UAV calculated by the visual are used as the estimated values, and attitude measured by airdrop IMU is used as the true value. Finally, the root-mean-square error (RMSE) between the solution attitude and the three-axis attitude of the IMU is calculated to measure the deviation between the observed value and the true value. The RMSE is taken as the standard to evaluate the accuracy, and then the average value of RMSE after five experiments is used as the final angle error result of the image sequence

\[
\text{RMSE}(\hat{\theta}_i, \theta_i) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (\text{trans}(\hat{\theta}_i) - \text{trans}(\theta_i))^2}
\]  

(40)

Equation (39) gives the calculation formula of RMSE, where \( i \) represents \( n \) frames of images to be processed, \( \theta_i \) represents the attitude of the real airdrop cargo platform, and \( \hat{\theta}_i \) represents the three-axis attitude calculation based on visual algorithm. And in the formula, trans is the root-mean-square of Euclidean distance.

Validity experiment of image geometric transformation

During the landing process of the airdrop cargo platform, the acquired images are subject to affine transformation, perspective transformation, and scale transformation. Setting up this experiment is to prove the effectiveness of this attitude calculation in geometric transformation. The experimental image is taken at a height of 120 m from the ground. The affine transformation and perspective transformation function in OpenCV are used to transform the image. The scale transformation image is taken at a height of 160 m from the ground. The image processing result is shown in Figure 9. It can be seen from Figure 9 that the affine transformation and perspective transformation of the image affect the distribution of line features (angle, length, and position in the images), which makes some line features unstable, resulting in a small reduction in the number of line features detected in the image. However, the longer line features are not affected by geometric transformation. In the image geometric transformation, the improved algorithm can still detect abundant line features, which is enough to support the attitude calculation of the airdrop cargo platform. It can verify the robustness of the improved line features in the scene and make the detection line features more reliable.

Attitude estimation experiment based on line features

To verify the effectiveness and accuracy of the visual algorithm for UAV attitude estimation based on line features, UAV landed at a height of 160 m above the ground, the flight distance is long, and a vertical landing speed is about 1.6 m/s. The camera’s sampling frequency is low, with a resolution of \( 960 \times 540 \). Compare the visually calculated attitude angle and the real attitude of the airborne IMU output over time, and calculate the RMSE value, as shown in Figure 10.

Figure 10 illustrates the relationship between UAV attitude error and landing time. The red line represents the fitting curve of the real-time attitude angle of UAV landing.
estimated based on line features, the blue line represents the fitting curve of real value output by IMU, and the yellow line represents the fitting curve of UAV attitude angle estimated by LSD and LBD. It can be seen from the observation that the RMSE values of roll, pitch, and yaw of UAV are 1.715°, 1.698°, and 0.886°, respectively, based on the

Figure 9. Line features extraction results in geometric transformation: (a) scene image, (b) perspective transformation, (c) affine transformation, and (d) scale transformation.

Figure 10. (a)–(c) are graphs UAV’s roll $\phi$, pitch $\theta$, and yaw $\psi$ calculated based on line features.
improved line features. The RMSE values of roll, pitch, and yaw of UAV are 3.524°, 3.165°, and 1.848°, respectively, by LSD and LBD. Compared with the traditional algorithm, the accuracy of the improved algorithm is more accurate and closer to the real value of IMU measurement.

**Line features and point features comparison experiment**

In scene 1, the UAV is used to collect the image sequence of landing. For the improved algorithm in this article, the traditional line feature extraction (LSD + LBD) and the oriented fast and rotated brief (ORB) feature point extraction, and performing the brute force matching is calculated to obtain the angle accuracy for comparison. The results of point feature extraction and matching are shown in Figure 11. The RMSE values of roll $\phi$, pitch $\theta$, and yaw $\psi$ at different landing heights of the airdrop cargo platform are counted, as shown in Figure 12.

As shown in Figure 12, the improved line feature extraction algorithm in this article is compared with the algorithm that only extracts point features, RMSE is reduced by 48.71%. Compared with LSD and LBD, RMSE is reduced by 22.85%. The accuracy of calculating the attitude based on the geometric characteristics of the line features is significantly improved compared with other methods, which proves that the algorithm in this article has a higher accuracy and can be used for the attitude calculation during the landing of the airdrop cargo platform.

When the airdrop cargo platform reaches the initial landing height, imaging at a high altitude, and the slight changes of attitude angle and the terrain fluctuation affect the imaging bottom point to generate a large deviation, which led to poor point features extraction and track failure, and the accuracy error of solving attitude angle is too large to ensure accurate landing. There are a lot of interference information such as short line features in the airdrop scene, and there are too many line features with inconspicuous gradient in the high altitude, which increases the mismatching of line features and the amount of calculation. The improved algorithm in this article can eliminate short line features in local dense areas and use geometric features of line features to group and merge the remaining line features, which makes the long line features more stable in the image, avoids the interference of short line features in the matching process, and effectively improves the matching accuracy of line features.

**Time performance experience**

To verify the real-time performance of the improved algorithm, the time performance of the vision algorithm is tested to avoid the imbalance of UAV. The traditional line feature extraction (LSD + LBD) and the improved algorithm are used to calculate the landing time of the UAV, respectively. The real-time performance of the algorithm is measured according to the single-frame pose solving time, as shown in Table 4.

It can be seen from Table 4 that in processing the image resolution of 960 × 540, the average single-frame line features extraction takes 301.56 ms, the line features matching takes 38.49 ms, and the total single-frame attitude angle settlement takes 348.69 ms, including line feature extraction, matching, and attitude angle resolution time of 13.7 ms. Compared with the traditional line feature extraction (LSD + LBD), the single-frame solution time is saved by 314.54 ms, and 287 frames of images can be processed...
during the landing of the airdrop cargo platform. At the image resolution of 3840 × 2160, the average single-frame line feature extraction takes 4009.89 ms, the line feature matching takes 46.10 ms, and the total single-frame attitude angle settlement takes 4076.45 ms, including line feature extraction, matching, and attitude angle resolution time of 20.46 ms. Compared with the traditional line feature extraction (LSD + LBD), the single-frame solution time is saved by 4673.7 ms, and 24 frames of images can be processed during the landing of the airdrop cargo platform. Obviously, the improved algorithm in this article is faster in solving attitude based on the geometric features of the line feature and saves a lot of calculation time. It has certain reference significance for the study of the attitude solution of the landing airdrop cargo platform.

In the aspect of time-consuming for line features matching, this article uses the geometric characteristics of line features between adjacent frames to further eliminate mismatching. Although it does not save too much time, it reduces mismatches and improves the accuracy of attitude solution. In the aspect of time-consuming for line features extraction, more line features will be processed in high-resolution images, and the extraction of line features accounts for most of the time spent in single-frame attitude solution. A large number of line features increase the time for extracting line time per frame, which makes the airdrop system unable to meet the real-time requirements. At a resolution of 960 × 540, some short line features cannot be displayed in the image, which naturally reduces the number of line features. Through the screening of line features, a large number of short line features are eliminated, and some short line features are merged, which reduces the number of line features to be extracted in the image again, saves the time of attitude solution, and improves the real-time performance of the system.

**Conclusions and future work**

This article proposes a vision algorithm for airdrop cargo platform attitude estimation and autonomous landing based on line features. The line features in the airdrop landing scene are extracted as visual goals, but too many invalid line features in the scene increase the computational complexity of the algorithm. Therefore, this article uses the geometric characteristics of line features to improve the traditional line feature extraction and eliminates the locally dense line features in the image, which greatly reduces the number of line features in the image. Then, the improved RANSAC algorithm is used to remove the mismatching of line features, which improves the accuracy of attitude angle, and makes up for the problems of difficult point features extraction or low matching accuracy in airdrop environment. Finally, a constraint equation is established for the successfully matched line features, and the homography matrix is solved and decomposed to obtain the realtime triaxial attitude angle of the airdrop cargo platform.

To illustrate the performance of the airdrop cargo platform attitude estimation system based on line features, outdoor experiments are conducted and analyzed. Comprehensive experimental results show that the maximum reduction of the number of line features extracted by the improved line features algorithm compared with LSD can reach 27.24%, the matching accuracy after geometric constraint screening can reach 75%, and the line features extraction and matching effects have been significantly improved. The RMSE values of roll, pitch, and yaw of UAV are 1.715°, 1.698°, and 0.886°, respectively, based on line features, and the average single-frame attitude angle resolution time is 348.69 ms, which saves 314.54 ms compared with the traditional line feature extraction (LSD + LBD). Although this article is a preliminary work, it demonstrates the feasibility of the attitude angle based on visual algorithm of the landing process of airdrop cargo platforms.

Future work will consider the poor real-time processing of images for the airdrop cargo platform, optimize the algorithm, and divide the two processes of feature detection and pose estimation into two independent threads. The segmentation/line/transformation prediction model based on convolutional neural network will be studied later. The combination of the vision algorithm and the flight control system further proves its reliability and accuracy during airdrop landing.

**Declaration of conflicting interests**

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.
Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

ORCID iD
Hongying Zhang https://orcid.org/0000-0002-5655-2934

References
1. Liu R, Sun XX, and Wang D. Heavyweight airdrop flight control design using feedback linearization and adaptive sliding mode. Trans Inst Meas Control 2016; 38(10): 1155–1164.
2. Desabrais KJ, Riley J, Sadeck J, et al. Low-cost high-altitude low-opening cargo airdrop systems. J Air Cr 2012; 49(1): 349–354.
3. Al-Kaff A, Martin D, Garcia F, et al. Survey of computer vision algorithms and applications for airdrop cargo platform. Expert Syst Appl 2018; 92: 447–463.
4. Kendoul F. Survey of advances in guidance, navigation, and control of unmanned rotorcraft systems. J Field Robot 2012; 29(2): 315–378.
5. Meng Y, Wang W, Han H, et al. A visual/inertial integrated landing guidance method for UAV landing on the ship. Aerosc Sci Technol 2019; 85: 474–480.
6. Tzoumanikas D, Li WB, Grimm M, et al. Fully autonomous micro air vehicle flight and landing on a moving target using visual-inertial estimation and model-predictive control. J Field Robot 2019; 36(1): 49–77.
7. Araar O, Aouf N, and Vitanov I. Vision based autonomous landing of multicopter UAV on moving platform. J Intell Robot Syst 2017; 85(2): 369–384.
8. Yang QQ and Sun LL. A fuzzy complementary Kalman filter based on visual and IMU data for UAV landing. Optik 2018; 173: 279–291.
9. Zhang L, Zhai ZJ, He L, et al. Infrared-inertial navigation for commercial aircraft precision landing in low visibility and GPS-denied environments. Sensors 2019; 19(2): 408.
10. Xu XB, Wang Z, and Deng YM. A software platform for vision-based UAV autonomous landing guidance based on markers estimation. Sci China-Technol Sci 2019; 62(10): 1825–1836.
11. Lopez E, Garcia S, Barea R, et al. A multi-sensorial simultaneous localization and mapping (SLAM) system for low-cost micro aerial vehicles in GPS-denied environments. Sensors 2017; 17(4): 802.
12. Meng Z, Kong X, Meng L, et al. Camera motion estimation and optimization approach. In: 2019 International conference on advanced mechatronic systems, ICAMechS 2019, Kusatsu, Japan, 26–28 August 2019. pp. 212–217. IEEE Computer Society.
13. Lee SJ and Hwang SS. Elaborate monocular point and line SLAM with robust initialization. In: 17th IEEE/CVF international conference on computer vision, ICCV 2019, Seoul, Korea, 27 October–2 November 2019. Republic of Korea: Institute of Electrical and Electronics Engineers Inc.
14. Li K, Yao J, Lu XH, et al. Hierarchical line matching based on line-junction-line structure descriptor and local homography estimation. Neurocomputing 2016; 184: 207–220.
15. Zhang FK, Rui T, Yang CS, et al. LAP-SLAM: a line-assisted point-based monocular VSLAM. Electronics 2019; 8(2): 243.
16. Teng XC, Yu QF, Luo J, et al. Aircraft pose estimation based on geometry structure features and line correspondences. Sensors 2019; 19(9): 2165.
17. Yu QD, Xu GL, and Cheng YH. An efficient and globally optimal method for camera pose estimation using line features. Mach Vis Appl 2020; 31(6): 48.
18. Li X, Xu G, and Wu Q. An infrared small target detection method based on local contrast measure and gradient property. In: 2020 Fifth international seminar on computer technology, mechanical and electrical engineering, ISCMIE 2020, Shenyang, Virtual, China, 30 October–1 November 2020, 2021. IOP Publishing Ltd.
19. Li G, Zeng YW, Huang HL, et al. A multi-feature fusion slam system attaching semantic invariant to points and lines. Sensors 2021; 21(4): 1196.
20. Zuo X, Xie X, Liu Y, et al. Robust visual SLAM with point and line features. In: 2017 IEEE/RSJ international conference on intelligent robots and systems, IROS 2017, Vancouver, BC, Canada, 24–28 September 2017. Institute of Electrical and Electronics Engineers Inc.
21. Pumarola A, Vakhitov A, Agudo A, et al. PL-SLAM: real-time monocular visual SLAM with points and lines. In: 2017 IEEE international conference on robotics and automation, ICRA 2017, Singapore, Singapore, 29 May–3 June 2017. Institute of Electrical and Electronics Engineers Inc.
22. Wang RZ, Di KC, Wan WH, et al. Improved point-line feature based visual SLAM method for indoor scenes. Sensors 2018; 18(10): 3559.
23. He YJ, Zhao J, Guo Y, et al. PL-VIO: tightly-coupled monocular visual-inertial odometry using point and line features. Sensors 2018; 18(4): 1159.
24. Gomez-Ojeda R and Gonzalez-Jimenez J. Geometric-based line segment tracking for HDR stereo sequences. In: 2018 IEEE/RSJ international conference on intelligent robots and systems, IROS 2018, Madrid, Spain, 1–5 October 2018. Institute of Electrical and Electronics Engineers Inc.
25. Zhou F, Zhang LM, Deng CL, et al. Improved point-line feature based visual SLAM method for complex environments. Sensors 2021; 21(13): 4604.
26. Grompone von Gioi R, Jakubowicz J, Morel J-M, et al. LSD: a fast line segment detector with a false detection control. IEEE Trans Pattern Anal Mach Intell 2010; 32(4): 722–732.
27. Zhang LL and Koch R. An efficient and robust line segment matching approach based on LBD descriptor and pairwise geometric consistency. J Vis Commun Image Represent 2013; 24(7): 794–805.
28. Salaun Y, Marlet R, and Monasse P. The multiscale line segment detector. In: First workshop on reproducible research in pattern recognition, RRPR 2016, Cancun, Mexico, 4 December 2016, 2017. Springer Verlag.
29. Lebeda K, Matas J, and Chum O. Fixing the locally optimized RANSAC. In: 2012 23rd British machine vision conference, *BMVC 2012*, Guildford, Surrey, United Kingdom, 3–7 September 2012. British Machine Vision Association, BMVA.

30. Sattler T, Leibe B, and Kobbelt L. SCRAMSAC: improving RANSAC’s efficiency with a spatial consistency filter. In: *12th international conference on computer vision, ICCV 2009*, Kyoto, Japan, 29 September–2 October 2009. Institute of Electrical and Electronics Engineers Inc.

31. Rublee E, Rabaud V, Konolige K, et al. ORB: an efficient alternative to SIFT or SURF. In: *2011 IEEE international conference on computer vision, ICCV 2011*, Barcelona, Spain, 6–13 November 2011. Institute of Electrical and Electronics Engineers Inc.