EuroGNC Navigation Guidance Control CEAS Bristol, UK June 11th-13th 2024
University of BRISTOL ROYAL AERONAUTICAL SOCIETY AIAA

Landing Pad Design for Optical-aided Automatic Vertical Take-off and Landing

Jun Shi Research Associate, Technical University of Munich, Institute of Flight System Dynamics, 85748, Garching, Germany. jun.shi@tum.de

Xiang Fang Research Associate, Technical University of Munich, Institute of Flight System Dynamics, 85748, Garching, Germany. xiang.fang@tum.de

Nils Schlautmann Research Associate, Technical University of Munich, Institute of Flight System Dynamics, 85748, Garching, Germany. nils.schlautmann@tum.de

Michael Zintl Research Associate, Technical University of Munich, Institute of Flight System Dynamics, 85748, Garching, Germany. michael.zintl@tum.de

Johann Dambeck Professor, Technical University of Munich, Institute of Flight System Dynamics, 85748, Garching, Germany. dambeck@tum.de

Florian Holzapfel Professor, Technical University of Munich, Institute of Flight System Dynamics, 85748, Garching, Germany. florian.holzapfel@tum.de

ABSTRACT

Automatic landing of VTOL aircraft under low illumination conditions, e.g., night operations, is challenging. To address this challenge, this paper proposes a novel design of ground-based infrastructure comprising a landing pad equipped with optical markers including finder markers and active markers. The design enables a VTOL aircraft equipped with an onboard camera to obtain real-time 6-degree-of-freedom relative pose during the landing phase under low illumination conditions since the optical markers are composed of active light sources. The well-structured layout geometry of the finder markers facilitates robust landing pad detection from aerial images without relying on complicated feature-matching technology. The active markers, employing specific blinking frequencies, contribute to the unique identification of the landing pad, ensuring a secure landing. To validate the feasibility of relative pose estimation with the proposed layout of optical markers, a modified four-point algorithm is applied for pose estimation in this paper. Compared with the classic four-point algorithm, our solution avoids the explicit ambiguity resolution. It uniquely determines the relative pose by fully leveraging prior knowledge of the landing pad surface. Simulations with a landing trajectory demonstrate that the proposed landing pad design is theoretically feasible for automatic take-off and landing in terms of recognition and pose estimation.

Keywords: VTOL Aircraft; Low Illumination; Landing Pad; Optical Marker; Acitve Light Source; Camera; Relative 6DoF Pose Estimation; Four-point Algorithm

Nomenclature

b	=	Body-fixed frame
l	=	Landing pad frame
l'	=	Auxiliary frame
c	=	Camera frame
o_b	=	Origin of the body frame
o_l	=	Origin of the landing pad frame
$o_{l'}$	=	Origin of the auxiliary frame
o_c	=	Origin of the camera frame
$\mathbf{R}$	=	Atitude in the form of direction cosine matrix
$\mathbf{r}$	=	Position
$\mathbf{R}_{lb}$	=	Atitude of the aircraft with respect to the landing pad frame
$\mathbf{r}^l_{o_b}$	=	Position of the aircraft with respect to the landing pad frame
$\mathbf{R}_{l'b}$	=	Atitude of the aircraft with respect to the auxiliary frame
$\mathbf{r}^{l'}_{o_b}$	=	Position of the aircraft with respect to the auxiliary frame
$\mathbf{R}_{bc}$	=	Relative attitude between the aircraft and the camera
$\mathbf{r}^b_{o_c}$	=	Lever arm between the aircraft and the camera
$\mathbf{R}_{cl'}$	=	Relative attitude between the camera frame and the auxiliary frame
$\mathbf{r}^c_{o_{l'}}$	=	Position of the auxiliary frame origin in the camera frame
$\mathbf{r}^l_s$	=	Position of Point s in the landing pad frame
$\mathbf{r}^{l'}_s$	=	Position of Point s in the auxiliary frame
$\mathbf{r}^c_s$	=	Position of Point s in the camera frame
$\mathbf{p}_s$	=	Pixel coordinates of the projection of Point s onto the image
$\bar{\mathbf{p}}_s$	=	Homogeneous representation of the pixel $\mathbf{p}_s$
$\mathbf{r}^m_s$	=	Normalized image coordinates of the projection of Point s onto the image
$\bar{\mathbf{r}}^m_s$	=	Homogeneous representation of $\mathbf{r}^m_s$
π_c	=	Projection function for the camera
$\mathbf{K}$	=	Intrinsic matrix of the camera
$\mathbf{n}$	=	Normal vector of the landing pad surface
$\mathbf{H}$	=	Homography matrix
est	=	Estimation
ref	=	Reference
r^l_x, r^l_y, r^l_z	=	Coordinates for the position $\mathbf{r}^l_{o_b}$
ϕ_x, ϕ_y, ϕ_z	=	Errors in the estimate of the attitude of the aircraft with respect to the landing pad frame
$\delta r^l_x, \delta r^l_y, \delta r^l_z$	=	Errors in the estimate of the position of the aircraft with respect to the landing pad frame
h_{avl}	=	Height above the landing pad

1 Introduction

Over last decade, the academic community and industry have shown great interest in vertical take-off and landing (VTOL) aircraft because of its various potential applications, e.g., commercial air taxi and parcel delivery services [1]. During low-speed operations, VTOL aircraft typically exhibit complex flight dynamics, limited maneuverability [2], and heightened power demands [3]. Given the criticality nature of VTOL operations, regulatory authorities have introduced specific requirements for VTOL aircraft [4]. Additionally, future VTOL operations may forego onboard pilots to enhance range or payload capacity [5]. Therefore, the development of a highly reliable automatic landing system is imperative to ensure the safety of VTOL aircraft. Furthermore, in the vision of urban air mobility (UAM), a reliable, high-precision, and economically efficient navigation solution for vertical landing is vital.

Recent advancements in computer vision have positioned cameras to play a vital role in aircraft landing. Angermann et al. [6] adopted cameras to identify the runway based on line and contour features

and then estimated the position for a fixed-wing aircraft. Liu et al. [7] proposed a new landing platform with multiple optical markers of different sizes, enabling robust recognition at different distances. To land an aircraft under extremely low illumination conditions, Lin et al. [8] utilized a model-based enhancement strategy to improve the brightness of images captured by the onboard camera. Besides, they also adopted a hierarchical-based method comprising a decision tree and an associated lightweight convolutional neural network to localize the marker with a coarse-to-fine strategy.

For the sake of robust recognition of the landing pad and precise estimation of the aircraft's position and attitude with respect to the landing pad, some artificial markers are proposed for automatic landing tasks, such as ARTag [9], ARToolKit [10], AprilTag [11], and ArUco [12]. These markers can be uniquely recognized with their encoded binary code. The four corners of these markers and their projections are utilized to estimate the relative pose. However, passive markers used in these solutions cannot provide reliable pose under low illumination conditions. Some researchers adopted optical markers containing active light sources to realize local navigation in low illumination. Xu et al. [13] utilized a platform with a T-shaped marker that emits infrared lights for tracking. Springer et al. [14] adopted a landing platform equipped with an active infrared AprilTag to allow tracking in the nighttime. However, due to the small size of these optical markers, the functional ranges of these systems are insufficient for positioning in a large-scale space (e.g., 30mx30mx30m).

To facilitate the safe automatic landing of a VTOL aircraft in an ample space under low illumination conditions, this paper presents an optical landing system capable of providing an accurate and reliable relative pose of a VTOL aircraft during the landing phase. The system consists of an onboard camera system, an onboard computer, and a landing pad equipped with optical markers. Unlike the passive markers, these optical markers containing active light sources enable landing at night or under low illumination conditions. To achieve a reliable and precise pose estimation, a new design of the layout of active markers on the landing pad is proposed. Seven finder markers are strategically placed on the landing for easy and reliable detection of the landing pad by the airborne camera, eliminating the need for a complicated feature-matching algorithm. Besides, considering the safety of landing, three blinking optical markers with different frequencies are also installed to identify the landing pad.

In addition to the aspects of landing pad recognition, this paper also investigates the feasibility of the proposed landing pad design in terms of relative pose estimation. Since all the optical markers are mounted on the same planar surface, we consider estimating the relative pose from the homography matrix between the image plane and the landing pad plane. The classic pose estimation method for planar homography, as outlined in [15] and [16], typically involves recovering the homography matrix through the Direct Linear Transform (DLT) using at least four coplanar correspondences. Subsequently, the homography matrix is normalized, and the pose is obtained through numerical decomposition techniques. Further, Malis et al. [17] proposed an analytical method to extract the relative pose from the homography matrix. However, these methods yield two physically possible solutions, necessitating additional information to resolve ambiguity. In our study, the prior knowledge of the planar surface can be leveraged to address the ambiguity. Rather than resolving ambiguity explicitly, we adjust the homography decomposition in the classic four-point method [15] with the knowledge of the planar surface. This adjustment allows us to uniquely determine the relative pose without the need of explicit ambiguity resolution.

The remainder of this paper is organized as follows. Section 2 introduces the mathematical model in the VTOL landing scenarios. In Section 3, the mapping between the landing pad plane and the image plane is proven to be a projective transformation. Then, a layout for optical markers of the landing pad is proposed based on the invariant measures of the projective transformation. Section 4 presents the corresponding landing pad detection and identification solutions to recognize the landing pad along with optical markers deployed in the proposed layout. Section 5 proposes the modified four-point algorithm, enabling relative pose estimation with the proposed landing pad. Section 6 demonstrates the simulation

results of the landing pad recognition algorithm and the modified four-point algorithm. Finally, Section 7 concludes the work of this paper.

2 Model Formulation

In this section, we describe the mathematical model in the VTOL aircraft landing scenario for further discussion.

2.1 Coodinate System

Fig. 1 presents the concept of the proposed optical landing system. An aircraft along with an airborne camera flies over a landing pad equipped with optical markers of active light source. The attitude and the position of the aircraft with respect to the landing pad can be denoted as $\mathbf{R}_{lb} \in \mathbf{SO}(3)$ and $\mathbf{r}^l_{o_b} \in \mathbb{R}^3$. In practice, the body frame of the aircraft b does not coincide with the camera frame c. Taking into account the relative pose $\left[\mathbf{R}_{bc}|\mathbf{r}^b_{o_c}\right]$ between the camera and the aircraft, the pose of the camera relative to the landing pad is

$$\mathbf{R}_{lc} = \mathbf{R}_{lb}\mathbf{R}_{bc}, \tag{1a}$$

$$\mathbf{r}^l_{o_c} = \mathbf{r}^l_{o_b} + \mathbf{R}_{lb}\mathbf{r}^b_{o_c}, \tag{1b}$$

where l is the landing pad frame.

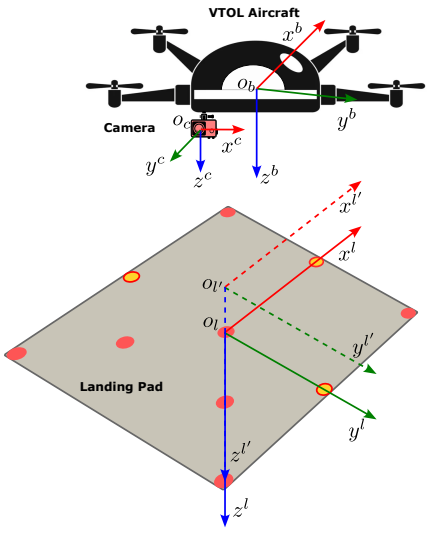


Fig. 1 Concept of the optical landing system. The yellow spots on the landing pad indicate the optical markers blinking at specific frequencies and duty cycles; the red spots indicate the optical markers are constantly on during the landing phase.

The optical markers are designed to be installed on the landing pad surface. Their coordinates in the landing pad frame are defined as $\mathbf{r}^l_{s_i} = \left[x^l_{s_i}, y^l_{s_i}, 0\right]$, which fulfills

$$\mathbf{n}^\top \mathbf{r}^l_{s_i} = d, \tag{2}$$

where $\mathbf{n} = [0, 0, 1]^\top$ is the normal vector of the landing pad surface, $d = 0$. For later mathematical derivation, it is expected that Eq. (2) can be converted in the form of $\frac{\mathbf{n}^\top \mathbf{r}^l_{s_i}}{d} = 1$. In this case, d should not equal 0. Therefore, an auxiliary frame denoted by l' is introduced to avoid $d = 0$. It can be obtained by shifting the frame l by 1m along its negative z-axis as shown in Fig. 1. Thus, we have $\mathbf{R}_{l'l} = \mathbf{I}_{3\times3}$ and $\mathbf{r}^{l'}_{o_l} = [0, 0, 1]^\top$. The coordinates of optical markers in the frame l' can be expressed as:

$$\mathbf{r}^{l'}_{s_i} = \mathbf{r}^{l'}_{o_l} + \mathbf{R}_{l'l}\mathbf{r}^l_{s_i} = \left[x^{l'}_{s_i}, y^{l'}_{s_i}, 1\right]^\top. \tag{3}$$

Obviously, it satisfies

$$\mathbf{n}^\top \mathbf{r}^{l'}_{s_i} = d', \text{ where } d' = 1. \tag{4}$$

Further, we have

$$\frac{\mathbf{n}^\top \mathbf{r}_{s_i}^{l'}}{d'} = 1 \tag{5}$$

With the attitude $\mathbf{R}_{lc}$ and the position $\mathbf{r}_{o_c}^{l}$ of the camera relative to the landing pad, $\mathbf{R}_{l'c}$ and $\mathbf{r}_{o_c}^{l'}$ are achieved as follows

$$\mathbf{R}_{l'c} = \mathbf{R}_{l'l}\mathbf{R}_{lc} = \mathbf{R}_{lc}, \tag{6a}$$

$$\mathbf{r}_{o_c}^{l'} = \mathbf{r}_{o_l}^{l'} + \mathbf{R}_{l'l}\mathbf{r}_{o_c}^{l} = \mathbf{r}_{o_l}^{l'} + \mathbf{r}_{o_c}^{l}. \tag{6b}$$

Given the coordinates $\mathbf{r}_s^{l'}$ of a 3D point s in the frame l', its coodinates in the frame c are

$$\mathbf{R}_{cl'} = \mathbf{R}_{l'c}^{-1}, \tag{7a}$$

$$\mathbf{r}_{o_{l'}}^{c} = -\mathbf{R}_{l'c}^{-1}\mathbf{r}_{o_c}^{l'}, \tag{7b}$$

$$\mathbf{r}_s^c = \mathbf{R}_{cl'}\mathbf{r}_s^{l'} + \mathbf{r}_{o_{l'}}^{c}. \tag{7c}$$

2.2 Camera Model

Considering a calibrated pespective camera, any 3D point s on a visiable surface maps to the pixel coordinates $\mathbf{p}_s = [u_s, v_s]^\top$ with the projection function π_c:

$$\mathbf{p}_s = \pi_c\left(\mathbf{r}_s^c\right) = \left[\frac{f_x x_s^c}{z_s^c} + c_x, \frac{f_y y_s^c}{z_s^c} + c_y\right]^\top, \tag{8}$$

where $\mathbf{r}_s^c = \left[x_s^c, y_s^c, z_s^c\right]^\top$ is the coordinates of point s expressed in the camera frame c, f_x and f_y the focal length of the camera in terms of pixels, c_x and c_y the principle point in the standard pinhole camera model.

2.3 Homogeneous Representation

For a pixel on an image, it can be represented in the homogeneous form with $\bar{\mathbf{p}}_{s_i} = \left[\bar{x}_{s_i}^p, \bar{y}_{s_i}^p, \bar{z}_{s_i}^p\right]^\top$. Its pixel coordinates are

$$\mathbf{p}_{s_i} = \left[\bar{x}_{s_i}^p / \bar{z}_{s_i}^p, \bar{y}_{s_i}^p / \bar{z}_{s_i}^p\right]^\top. \tag{9}$$

With the homogeneous representation of the pixel, the projection model of the camera can be re-written as

$$\bar{\mathbf{p}}_{s_i} = \mathbf{K}\left(\mathbf{R}_{cl'}\mathbf{r}_{s_i}^{l'} + \mathbf{r}_{o_{l'}}^{c}\right), \tag{10}$$

where $\mathbf{K} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$.

3 Layout of the Optical Markers on the Landing Pad

The basic idea behind the landing pad design is based on the invariant properties of the projective transformation. In this section, we initially prove that the mapping between the landing pad plane and the image plane is a projective transformation, namely homography. The proof is crucial for designing the optical marker layout. Then, we introduce the invariant measures under projective transformation, which guide our layout design. Finally, one layout of the optical markers based on these invariant measures is presented.

3.1 Projective Transformation

For the 3D points lying on the landing pad plane, the projection model in Eq. (10) can be updated by substituting Eq. (5) into Eq. (10) as follows:

$$\bar{\mathbf{p}}_{s_i} = \mathbf{K}\left(\mathbf{R}_{cl'}\mathbf{r}_{s_i}^{l'} + \mathbf{r}_{o_{l'}}^{c} \cdot 1\right) = \mathbf{K}\left(\mathbf{R}_{cl'}\mathbf{r}_{s_i}^{l'} + \mathbf{r}_{o_{l'}}^{c}\frac{\mathbf{n}^\top \mathbf{r}_{s_i}^{l'}}{d'}\right) = \mathbf{K}\left(\mathbf{R}_{cl'} + \frac{1}{d'}\mathbf{r}_{o_{l'}}^{c}\mathbf{n}^\top\right)\mathbf{r}_{s_i}^{l'}. \tag{11}$$

Since $d' = 1$, it can be simplified as

$$\bar{\mathbf{p}}_{s_i} = \mathbf{K}\left(\mathbf{R}_{cl'} + \mathbf{r}_{o_{l'}}^{c}\mathbf{n}^\top\right)\mathbf{r}_{s_i}^{l'}. \tag{12}$$

Here, we define a 3×3 matrix $\mathbf{H} = \mathbf{K}\left(\mathbf{R}_{cl'} + \mathbf{r}_{o_{l'}}^{c}\mathbf{n}^\top\right)$, and then we have $\bar{\mathbf{p}}_{s_i} = \mathbf{H}\mathbf{r}_{s_i}^{l'}$. With the matrix inversion lemma [18], the inverse matrix of $\mathbf{H}$ is:

$$\mathbf{H}^{-1} = \left(\mathbf{R}_{cl'}^{-1} - \frac{\mathbf{R}_{cl'}^{-1}\mathbf{r}_{o_{l'}}^{c}\mathbf{n}^\top\mathbf{R}_{cl'}^{-1}}{1 + \mathbf{n}^\top\mathbf{R}_{cl'}^{-1}\mathbf{r}_{o_{l'}}^{c}}\right)\mathbf{K}^{-1}. \tag{13}$$

Substitution of $\mathbf{r}_{o_c}^{l'} = -\mathbf{R}_{cl'}^{-1}\mathbf{r}_{o_{l'}}^{c}$ into the denominator in Eq. (13) gives us

$$1 + \mathbf{n}^\top\mathbf{R}_{cl'}^{-1}\mathbf{r}_{o_{l'}}^{c} = 1 - \mathbf{n}^\top\mathbf{r}_{o_c}^{l'}. \tag{14}$$

Since the camera can be considered not on the landing plane in this paper, the projection of $\mathbf{r}_{o_c}^{l'}$ on the normal vector of the planar surface $\mathbf{n}$ does not equal to 1. Therefore, we have $1 - \mathbf{n}^\top\mathbf{r}_{o_c}^{l'} \neq 0$, i.e., $\mathbf{H}$ is a non-singular matrix. Although $\mathbf{r}_{s_i}^{l'}$ is of 3D inhomogeneous coordinates, it can also be regarded as homogeneous coordinates, which represent a 2D point $\left[x_{s_i}^{l'}, y_{s_i}^{l'}\right]^\top$ on the landing pad plane. Therefore, $\mathbf{H}$ is a planar projective transformation, which realizes the mapping of any 2D point from the landing pad plane to the image plane of the airborne camera.

3.2 Invariant Measures of Projective Transformation

Several measures are invariant under the projective transformation, including collinearity and cross ratio [19]. Here, the invariant measures employed in the design of the layout of the optical markers are outlined below.

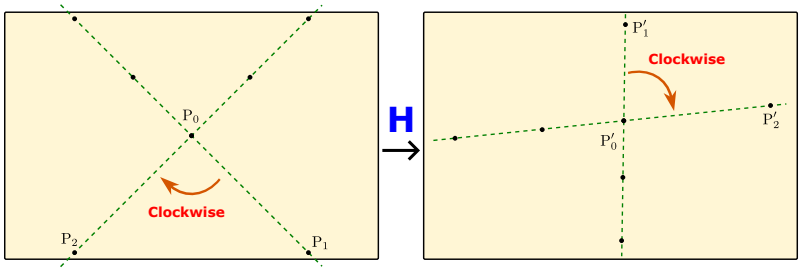
Collinearity Through a camera with an arbitrary pose, the projections on the image plane of a set of collinear points on the landing pad plane remain collinear.

Direction of Two-dimensional Rotation As long as the camera stays above the landing pad, the direction of the projections onto the image plane of two 2D vectors on the landing pad plane remains unchanged (see Fig. 2a).

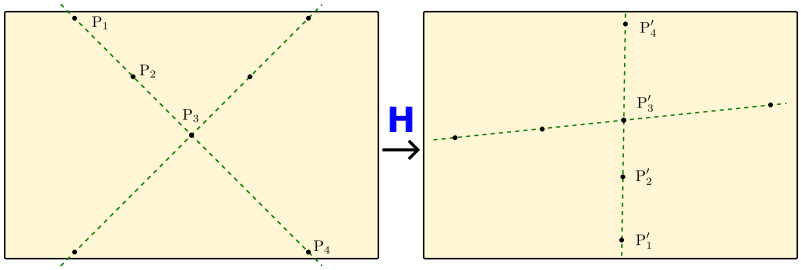
Cross Ratio of Four Collinear Points The cross ratio of lengths on a line is invariant under projective transformation (see Fig. 2b).

3.3 Layout of Optical Markers

Based on these invariant measures, we design a layout of optical markers depicted in Fig. 3. This layout contains seven finder markers and three active markers. All optical markers emit light within the same wavelength band. The distinct colors assigned to the optical markers in Fig. 3 are solely for distinguishing between finder markers and active markers, not indicative of variations in emitted light colors. The finder markers are kept on throughout the approach and landing phases. The active markers



(a) Rotation of two 2D vectors.



(b) Cross ratio.

Fig. 2 Invariant measures of projective transformation.

are configured to blink with specific frequencies and duty circles. In order to determine the blinking frequencies of the active markers, the frequencies of the active markers shall be less than half of the frame rate of the airborne camera according to the Nyquist–Shannon sampling theorem. In our layout design, all the optical markers are placed on a square with a side length of $4a$. Two sets of four collinear finder markers share the finder marker D. Set 1 comprises finder markers A, B, C, and D, while Set 2 finder markers D, E, F, and G. The active marker X is collinear with the finder markers A and G. Analogically, the active marker Y is collinear with the finder markers C and E. The active marker X is collinear with the finder markers C and G. It is essential to note that while a square is employed in our design, it is not a prerequisite; the layout can adopt the shape of any quadrangle. Different layouts result in different performance of pose estimation.

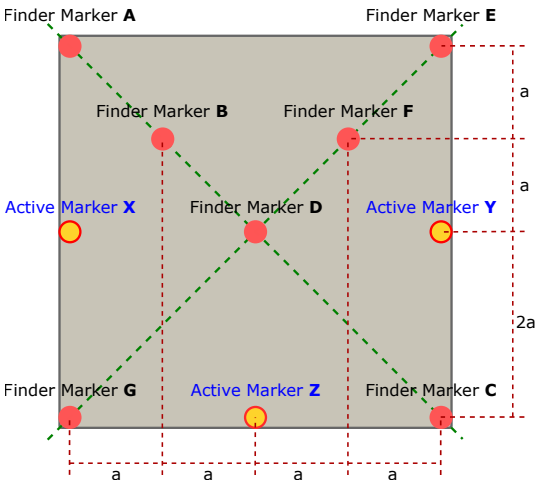


Fig. 3 Layout of optical markers.

4 Recognition of the Landing Pad

This section explains how the onboard system recognizes the designed landing pad from images captured by the airborne camera. The landing pad is detected with every single image. Once the onboard system finds the seven markers of the specific geometry, it is regarded as a candidate for the landing pad

specified by the flight mission. After a successful continuous detection in a sequence of images, the landing pad candidate can be verified through behavior analysis of the active markers, e.g., frequency and duty cycle. In this paper, the process of figuring out the candidate landing pad with seven finder markers in single image is the detection of the landing pad. The process of validating the candidate landing pad by analyzing active markers within a short period is identification. Thus, the whole process of landing pad recognition is split into two parts: detection and identification.

4.1 Detection of Landing Pad with Finder Markers

By analyzing the geometry of finder markers' projections on the image, some measures (i.e., collinearity, rotation of two vectors, cross ratio) can be extracted and compared with those calculated from the geometry of the finder markers before the mission (serving as *a priori* knowledge). If they match each other, the system is considered to have successfully detected the landing pad. Considering the integrity monitoring issue, the measures we can obtain from images are redundant so that some are used for detection and others for self-check.

The landing pad detection consists of five steps.

1) Detect markers: In this paper, the optical markers are assumed as LED spotlights. Their projections on the image are assumed as blobs as shown in Fig. 4a. Therefore, a blob detector based on the Laplacian of the Gaussian is performed to find the projection of the optical markers. Then, it returns the pixel coordinates of the projection center of the blobs as shown in Fig. 4b. For brevity, we call the center of these unidentified projections "point" in the rest of this section. Note that although LED spotlights are assumed in our design, it is not a must. The choice of optical marker type is open as long as the corresponding feature detection algorithm is available.

2) Find two sets of four collinear optical markers: A brutal force traversal of all the possible four-point sets is applied to find the two sets of four collinear points. If the search result is not exactly two sets of exact four collinear points, a message "invalid detection" is returned. If yes, the shared finder marker D can be determined (see Fig. 4c). However, the identity (ABCD or DEFG) of those two sets of four collinear points remains unknown.

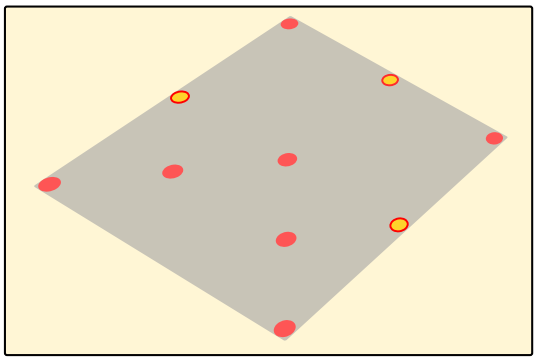
3) Identify marker sets by the rotation direction of two vectors: As shown in Fig. 4d, the direction from the vector $\overrightarrow{DA}$ to the vector $\overrightarrow{DE}$ shall be clockwise as long as the camera is above the landing pad. This conclusion enables the distinction between markers A and E, facilitating the identification of the marker sets.

4) Identify each finder marker: Since the markers B and C should lie on the line joining the identified markers A and D, they can be easily distinguished. Similarly, the markers B and C can also be identified. At this stage, the identities of all the finder markers are determined as shown in Fig. 4e.

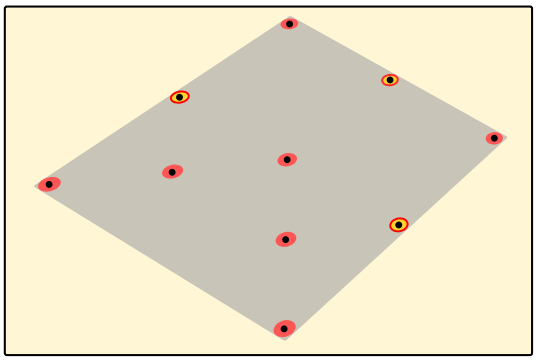
5) Self-check with cross ratio: After identifying all finder markers, the cross ratios of the two sets of four collinear markers as shown in Fig. 4f can be calculated. By comparing these cross ratios against the prior knowledge of the cross ratios extracted from the landing pad, the validity of the landing pad detection can be checked.

4.2 Identification of the Landing Pad with Active Markers

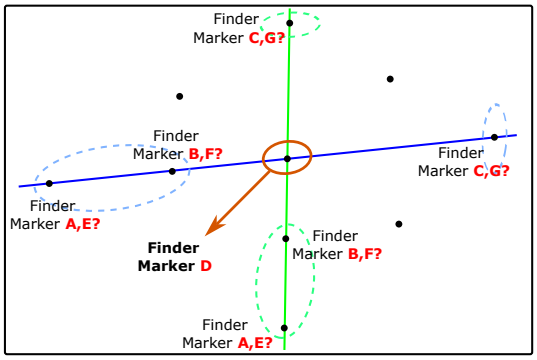
In case we have several ground-based landing pads whose layouts of finder markers are identical, it is hard to distinguish between them without additional information or sensors. However, we expect a stand-alone optical-based position and orientation system, resilient to potential failures of the external information or sensors. Thereby, the active markers with encoded information, i.e., several specific frequencies and duty cycles, are introduced to the landing pad design to uniquely identify the landing pad specified by the mission. Unlike the detection solution of the landing pad, the encoded information



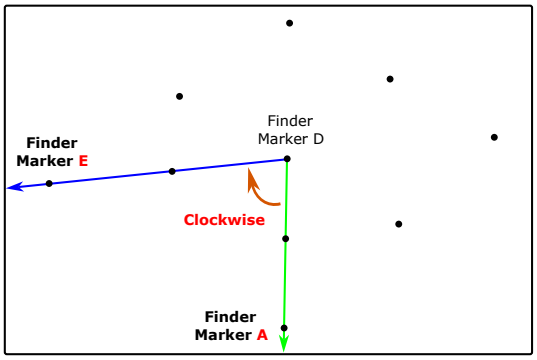
(a) Captured aerial image.



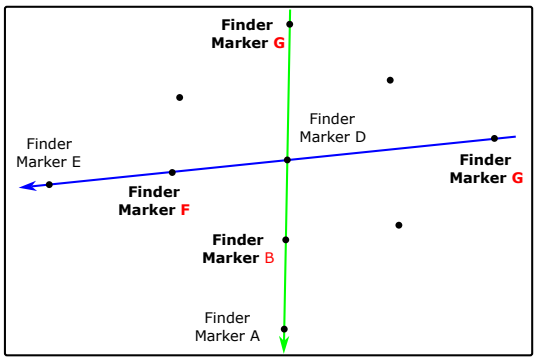
(b) Detect markers with blob detectors.



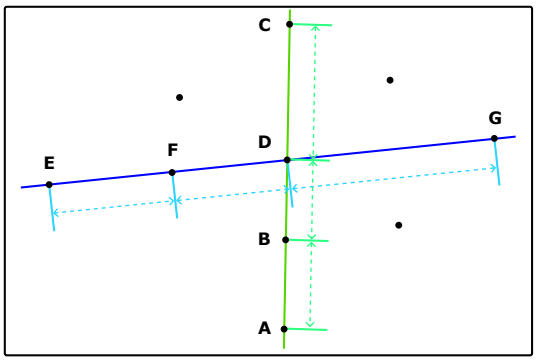
(c) Detect two sets of 4 collinear optical markers.



(d) Identify the optical marker sets.



(e) Identify each optical marker.



(f) Self-check with cross ratio.

Fig. 4 Landing pad detection.

of the active markers cannot be extracted from a single image: rather, it becomes available after a period of successful detection of the landing pad.

The whole process of the landing pad identification can be split into four steps:

1) Identify active markers by their geometry: In order to explain how the active markers are identified in every single image, a function of two 2D vectors $\vec{v}_1$ and $\vec{v}_2$ is defined as

$$\text{isClockwise}\left(\vec{v}_1, \vec{v}_2\right) = \begin{cases} 1, & \text{the rotation from } \vec{v}_1 \text{ to } \vec{v}_2 \text{ is clockwise;} \\ 0, & \text{the rotation from } \vec{v}_1 \text{ to } \vec{v}_2 \text{ is counterclockwise.} \end{cases} \tag{15}$$

In addition, the identity code of an active marker Q=X, Y, or Z is defined as

$$\text{ID}\left(\text{Q}\right) = \begin{bmatrix} \text{isClockwise}\left(\overrightarrow{\text{DQ}}, \overrightarrow{\text{DA}}\right) \\ \text{isClockwise}\left(\overrightarrow{\text{DQ}}, \overrightarrow{\text{DE}}\right) \\ \text{isClockwise}\left(\overrightarrow{\text{DQ}}, \overrightarrow{\text{DC}}\right) \\ \text{isClockwise}\left(\overrightarrow{\text{DQ}}, \overrightarrow{\text{DG}}\right) \end{bmatrix}. \tag{16}$$

Since the rotation between two 2D vectors is invariant under the projective transformation when the camera stays over the landing pad, the identity code for each active marker is also invariant. With analysis of the layout of optical markers, the identity codes of the active markers X, Y, and Z are $[1, 1, 0, 0]^\top$, $[0, 0, 1, 1]^\top$, and $[1, 0, 0, 1]^\top$ respectively.

After identifying all the finder markers on an aerial image, all other detected markers are regarded as potential active markers (see Fig. 5a). By comparing their identity codes with prior knowledge, these potential active markers become candidates for each active marker. For each active marker, there shall be 0 or 1 candidate. If the result violates this principle, return "invalid identification".

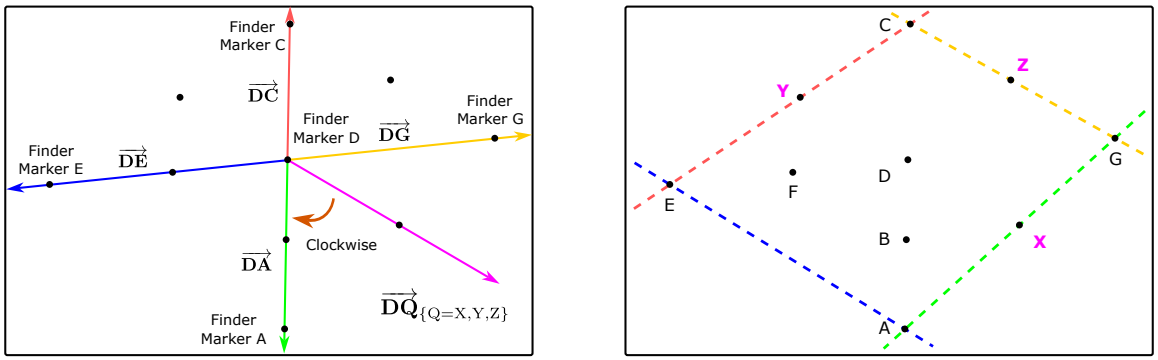
2) Self-check with collinearity: The candidate of the active marker X shall be on the line joining the finder marker A and G as shown in Fig. 5b. Similarly, the candidates of the active marker Y and Z have collinear constraints. If the candidates do not fulfill this assumption, return "invalid identification".

3) Determine the frequency and duty cycle of active marker signal: The status of the identified active marker over time, i.e., "visible", "invisible", and "invalid", can be recorded and then analyzed as a digital signal. The orange dots in Fig. 6a and Fig. 6b indicate the signals of one active marker. The sampling frequency of the signal is the frame rate of the camera. Through detecting two consecutive rise edges or trailing edges as shown in Fig. 6a, a complete period of the signal can be found. At least one complete period of signal beginning with a rising edge or trailing edge should be detected to calculate the frequency and duty cycle of the signal. Of course, analyzing more complete periods of signal can make the calculation of frequencies and duty cycle safer, but at the cost of longer delay of the landing pad identification. As shown in Fig. 6b, the analysis needs to be restarted once encountering "invalid" status.

4) Identify the landing pad: By comparing the frequencies and duty cycle with prior knowledge, the system can determine whether the landing pad it detects and tracks is the designated one for the mission.

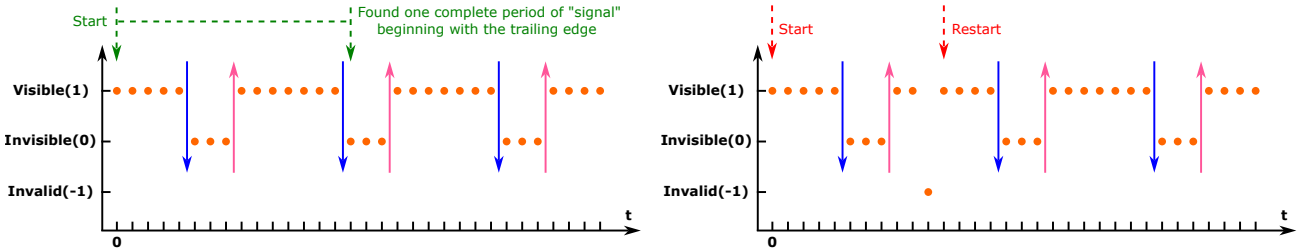
5 Relative Pose Estimation

In order to show the feasibility of the relative pose estimation with the proposed layout of the landing pad, this section explains how the relative pose of the camera relative to the landing pad is estimated.



(a) Identify active markers. **(b) Self-check with collinearity.**

Fig. 5 Landing pad identificaion.



(a) Successful case. **(b) Failed case.**

Fig. 6 Determination of the frequency and duty cycle from active marker signal.

5.1 Problem Formulation

With successful landing pad recognition, seven to ten 2D-3D correspondences can be established, contingent on the status of active markers. Since all 3D points lie on the same planar surface, the four-point algorithm is a classic method to compute the relative pose. However, this classic method yields two physically possible solutions. Ambiguity resolution typically requires additional information. For example, in our application, the normal vector of the planar surface is known and can serve as *a priori* knowledge to resolve ambiguity. Instead, in this paper, we adjust the homography decomposition in the classic four-point algorithm with the knowledge of the planar surface to uniquely determine the relative pose without ambiguity resolution.

The optical markers A, C, E, and G shown in Fig. 7 and their projections on the image are selected to estimate the relative pose since any three of them are not collinear. For concise notation, the 2D-3D correspondences for the optical markers A, C, E, and G are denoted as $\langle \bar{\mathbf{p}}_{s_i}, \mathbf{r}^l_{s_i} \rangle$, $i = 1, 2, 3, 4$. Here, we define normalized image coordinates in the homogeneous form

$$\bar{\mathbf{r}}^m_{s_i} = \mathbf{K}^{-1} \bar{\mathbf{p}}_{s_i}. \tag{17}$$

With Eq. (3) and (17), the 2D-3D correspondences can be transformed to $\langle \bar{\mathbf{r}}^m_{s_i}, \mathbf{r}^{l'}_{s_i} \rangle$, $i = 1, 2, 3, 4$. Substitution of Eq. (17) into Eq. (12) yields

$$\bar{\mathbf{r}}^m_{s_i} = \left(\mathbf{R}_{cl'} + \mathbf{r}^c_{o_{l'}} \mathbf{n}^\top \right) \mathbf{r}^{l'}_{s_i} = \bar{\mathbf{H}} \mathbf{r}^{l'}_{s_i}, \tag{18}$$

where $\bar{\mathbf{H}} = \mathbf{R}_{cl'} + \mathbf{r}^c_{o_{l'}} \mathbf{n}^\top$, is also called projective tranformation or homography. Obviously, $\bar{\mathbf{r}}^m_{s_i}$ is linear to $\left(\mathbf{R}_{cl'} + \mathbf{r}^c_{o_{l'}} \mathbf{n}^\top \right) \mathbf{r}^{l'}_{s_i}$. Further, we have

$$\bar{\mathbf{r}}^m_{s_i\,[\times]} \left(\mathbf{R}_{cl'} + \mathbf{r}^c_{o_{l'}} \mathbf{n}^\top \right) \mathbf{r}^{l'}_{s_i} = \mathbf{0}_{3\times 1}, \tag{19}$$

where $(\cdot)_{[\times]}$ indicates the skew symmetric matrix of a given vector.

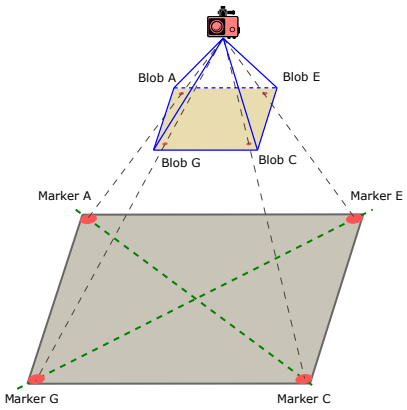


Fig. 7 Relative pose estimation.

To sum up, the relative pose estimation problem can be abstracted as: given a set of 2D-3D correspondences $\langle \bar{\mathbf{r}}^m_{s_i}, \mathbf{r}^{l'}_{s_i} \rangle$, $i = 1, 2, 3, 4$ and the 3D points are on the plane $\mathbf{n}^\top \mathbf{r}^{l'}_{s_i} = 1$, find $\left\{ \mathbf{R}_{cl'}, \mathbf{r}^{l'}_{s_i} \right\}$ that fulfills

$$\bar{\mathbf{r}}^m_{s_i\,[\times]} \bar{\mathbf{H}} \mathbf{r}^{l'}_{s_i} = \mathbf{0}_{3\times 1}, \; i = 1, 2, 3, 4. \tag{20}$$

Similar to the conventional four-point algorithm for the above problem, our solution can be factored into three steps: 1) recover the homography matrix up to a scale; 2) normalize the homography matrix; and 3) decompose the homography matrix. The difference between our solution and the conventional is in the step of homography decomposition, where the *a priori* knowledge of the planar surface is directly involved in calculating the direction cosine matrix instead of ambiguity resolution. The details of our modified four-point algorithm are explained based on the derivation of the classic four-point algorithm [15] as follows.

5.2 Recovery of the Homography Matrix up to a Scale

For ith correspondence, we have the constraint expressed in Eq. (20). It can be converted to a convenient expression with the Kronecker product:

$$\mathbf{a}_i \bar{\mathbf{H}}^{\mathrm{s}} = \mathbf{0}_{3\times 1}, \tag{21}$$

where

$$\mathbf{a}_i = \mathbf{r}^{l'\,\top}_{s_i} \otimes \bar{\mathbf{r}}^m_{s_i\,[\times]} \in \mathbb{R}^{3\times 9},$$

and $\bar{\mathbf{H}}^{\mathrm{s}} \in \mathbb{R}^9$ denotes a single column vector formed by stacking the columns of $\bar{\mathbf{H}}$. Note that the rank of the matrix $\bar{\mathbf{r}}^m_{s_i\,[\times]}$ is two, so is the matrix $\mathbf{a}_i$. Thus, although Eq. (21) has three rows, it provides only two independent constraints.

In this paper, four correspondences, of which no three are collinear, are established. Then, the constraints from the four pairs of correspondences can be constructed as

$$\mathbf{A} \bar{\mathbf{H}}^{\mathrm{s}} = \mathbf{0}_{12\times 1}, \tag{22}$$

where $\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^\top & \mathbf{a}_2^\top & \mathbf{a}_3^\top & \mathbf{a}_4^\top \end{bmatrix}^\top$. Then, by 1) performing the SVD of $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^\top$, 2) taking the ninth column of $\mathbf{V}$ as $\bar{\mathbf{H}}^{\mathrm{s}}_{\mathrm{L}}$ which satisfies $\mathbf{A}\bar{\mathbf{H}}^{\mathrm{s}}_{\mathrm{L}} = \mathbf{0}$, and 3) unstacking the the nine elements of $\bar{\mathbf{H}}^{\mathrm{s}}_{\mathrm{L}}$ into a 3×3 matrix $\bar{\mathbf{H}}_{\mathrm{L}}$, the homography $\mathbf{H}$ can be recovered in the following form with an unknown scale λ:

$$\bar{\mathbf{H}}_{\mathrm{L}} = \lambda \bar{\mathbf{H}} = \lambda \left(\mathbf{R}_{cl'} + \mathbf{r}^c_{o_{l'}} \mathbf{n}^\top \right). \tag{23}$$

5.3 Normalization of the Homography Matrix

Define $\mathbf{u} = \mathbf{R}_{cl'}^{\top}\mathbf{r}_{ol'}^{c}$, we have

$$\bar{\mathbf{H}}^{\top}\bar{\mathbf{H}} = \mathbf{I}_{3\times3} + \mathbf{u}\mathbf{n}^{\top} + \mathbf{n}\mathbf{u}^{\top} + \|\mathbf{u}\|^2\mathbf{n}\mathbf{n}^{\top}. \tag{24}$$

Here, we introduce the vector $\mathbf{u}\times\mathbf{n}$, which is orthogonal to $\mathbf{u}$ and $\mathbf{n}$ repectively. Multiply $\bar{\mathbf{H}}^{\top}\bar{\mathbf{H}}$ and the vector $\mathbf{u}\times\mathbf{n}$. Then, we have

$$\bar{\mathbf{H}}^{\top}\bar{\mathbf{H}}\,(\mathbf{u}\times\mathbf{n}) = \mathbf{u}\times\mathbf{n}. \tag{25}$$

Obviously, one eigenvalue of $\bar{\mathbf{H}}^{\top}\bar{\mathbf{H}}$ is 1 and the corrreponding eigenvector is $\mathbf{u}\times\mathbf{n}$. Besides, one singular value of $\bar{\mathbf{H}}$ is also 1.
To be concise, we define $\mathbf{v} = \|\mathbf{u}\|\mathbf{n}$ and $\mathbf{w} = \dfrac{1}{\|\mathbf{u}\|}\mathbf{u}$. Then, we have

$$\mathbf{Q} = \bar{\mathbf{H}}^{\top}\bar{\mathbf{H}} - \mathbf{I}_{3\times3} = \mathbf{u}\mathbf{n}^{\top} + \mathbf{n}\mathbf{u}^{\top} + \|\mathbf{u}\|^2\mathbf{n}\mathbf{n}^{\top} = (\mathbf{w}+\mathbf{v})(\mathbf{w}+\mathbf{v})^{\top} - \mathbf{w}\mathbf{w}^{\top}. \tag{26}$$

For the general case, the 3×3 matrix $\mathbf{Q}$ has three eigenvalues: one positive, one negative, and one equals zero. In the special case that $\mathbf{u}$ and $\mathbf{n}$ are collinear, $\mathbf{Q}$ has two repeated zero eigenvalues. Accordingly, the eigenvalues of $\bar{\mathbf{H}}^{\top}\bar{\mathbf{H}}$ fulfills

$$\sigma_1^2 \geq \sigma_2^2 = 1 \geq \sigma_3^2. \tag{27}$$

As the $\bar{\mathbf{H}}^{\top}\bar{\mathbf{H}}$ is a symmetric matrix, it can be decomposed with a diagonal matrix $\mathbf{\Sigma}$ and an orthogonal matrix $\mathbf{V}$:

$$\bar{\mathbf{H}}^{\top}\bar{\mathbf{H}} = \mathbf{V}\mathbf{\Sigma}\mathbf{V}^{\top}, \tag{28}$$

where

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1^2 & & \\ & \sigma_2^2 & \\ & & \sigma_3^2 \end{bmatrix}, \text{ and } \mathbf{V} = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix}.$$

Note that the corresponding eigenvalue of $\mathbf{v}_2$ is 1. Thus, we know $\mathbf{v}_2 = \mathbf{u}\times\mathbf{n}$.

For the homography matrix up to a scale, we have

$$\bar{\mathbf{H}}_{\mathrm{L}}^{\top}\bar{\mathbf{H}}_{\mathrm{L}} = \lambda^2\bar{\mathbf{H}}^{\top}\bar{\mathbf{H}} = \mathbf{V}\left(\lambda^2\mathbf{\Sigma}\right)\mathbf{V}^{\top}. \tag{29}$$

Obviously, the second largest eigenvalue of $\bar{\mathbf{H}}_{\mathrm{L}}^{\top}\bar{\mathbf{H}}_{\mathrm{L}}$ is λ^2, namely $\lambda^2 = \sigma_2^2\left(\bar{\mathbf{H}}_{\mathrm{L}}^{\top}\bar{\mathbf{H}}_{\mathrm{L}}\right)$. Thus, the absolute value of the unknown scale can be determined by computing the second largest singular value of $\bar{\mathbf{H}}_{\mathrm{L}}^{\top}\bar{\mathbf{H}}_{\mathrm{L}}$:

$$\|\lambda\| = \sqrt{\sigma_2^2\left(\bar{\mathbf{H}}_{\mathrm{L}}^{\top}\bar{\mathbf{H}}_{\mathrm{L}}\right)}. \tag{30}$$

Further, the homography can be recovered up to an unknown sign:

$$\bar{\mathbf{H}} = \pm\frac{\bar{\mathbf{H}}_{\mathrm{L}}}{\|\lambda\|} = \pm\frac{\bar{\mathbf{H}}_{\mathrm{L}}}{\sqrt{\sigma_2^2\left(\bar{\mathbf{H}}_{\mathrm{L}}^{\top}\bar{\mathbf{H}}_{\mathrm{L}}\right)}}. \tag{31}$$

The final step of the homography matrix determination is to find the correct sign. It is realized by imposing the positive depth constraint.

$$\bar{\mathbf{r}}_{s_i}^{m\,\top}\bar{\mathbf{H}}\mathbf{r}_{s_i}^{l'} > 0, \quad \forall i = 1,2,3,4. \tag{32}$$

5.4 Decomposition of the Homography Matrix

From the SVD of $\bar{\mathbf{H}}_{\mathrm{L}}^{\top}\bar{\mathbf{H}}_{\mathrm{L}}$ in the normalization step, $\mathbf{v}_2$ is also obtained. Notice that $\mathbf{v}_2 = \mathbf{u} \times \mathbf{n}$, which is orthogonal to $\mathbf{n}$. Here, we construct two vectors of unit length with $\mathbf{v}_2$ and $\mathbf{n}$, which satisfy $\mathbf{b}_i \perp \mathbf{n},\ i = 1, 2$:

$$\mathbf{b}_1 = \mathbf{n} \times \mathbf{v}_2 \times \mathbf{n}, \text{ and } \mathbf{b}_2 = \mathbf{n} \times \mathbf{v}_2. \tag{33}$$

For any vector $\mathbf{b}$, which satisfies $\mathbf{b} \perp \mathbf{n}$, we have

$$\bar{\mathbf{H}}\mathbf{b} = \left(\mathbf{R}_{cl'} + \mathbf{r}^{c}_{o_{l'}}\mathbf{n}^{\top}\right)\mathbf{b} = \mathbf{R}_{cl'}\mathbf{b}. \tag{34}$$

Thus,

$$\mathbf{R}_{cl'}\mathbf{b}_1 = \bar{\mathbf{H}}\mathbf{b}_1, \quad \mathbf{R}_{cl'}\mathbf{b}_2 = \bar{\mathbf{H}}\mathbf{b}_2. \tag{35}$$

Since $\mathbf{b}_1$ and $\mathbf{b}_2$ are orthogonal to $\mathbf{n}$, $\mathbf{n}$ can also be expressed as $\mathbf{n} = \mathbf{b}_1 \times \mathbf{b}_2$. Then, we have

$$\mathbf{R}_{cl'}\mathbf{n} = \mathbf{R}_{cl'}\left(\mathbf{b}_1 \times \mathbf{b}_2\right) = \left(\bar{\mathbf{H}}\mathbf{b}_1\right) \times \left(\bar{\mathbf{H}}\mathbf{b}_2\right). \tag{36}$$

Define the matrices $\mathbf{U} = \begin{bmatrix}\mathbf{b}_1, & \mathbf{b}_2, & \mathbf{n}\end{bmatrix}$ and $\mathbf{W} = \begin{bmatrix}\bar{\mathbf{H}}\mathbf{b}_1, & \bar{\mathbf{H}}\mathbf{b}_2, & \left(\bar{\mathbf{H}}\mathbf{b}_1\right) \times \left(\bar{\mathbf{H}}\mathbf{b}_2\right)\end{bmatrix}$. Then we obtain

$$\mathbf{R}_{cl'}\mathbf{U} = \mathbf{W}. \tag{37}$$

Therefore, the relative attitude and position can be determined with

$$\mathbf{R}_{cl'} = \mathbf{W}\mathbf{U}^{-1}, \tag{38a}$$

$$\mathbf{r}^{c}_{o_{l'}} = \left(\bar{\mathbf{H}} - \mathbf{R}_{cl'}\right)\mathbf{n}. \tag{38b}$$

The relative pose of the camera w.r.t. the landing pad can be obtained as follows:

$$\mathbf{R}_{lc} = \mathbf{R}^{-1}_{cl'}, \tag{39a}$$

$$\mathbf{r}^{l}_{o_c} = -\mathbf{R}^{-1}_{cl'}\mathbf{r}^{c}_{o_{l'}} - \mathbf{r}^{l'}_{o_l}. \tag{39b}$$

Further, the relative pose of the aircraft w.r.t. the landing pad can be achieved with

$$\mathbf{R}_{lb} = \mathbf{R}_{lc}\mathbf{R}_{cb}, \tag{40a}$$

$$\mathbf{r}^{l}_{o_b} = \mathbf{r}^{l}_{o_c} + \mathbf{R}_{lc}\mathbf{r}^{c}_{o_b}. \tag{40b}$$

6 Simulation and Analysis

In this section, the results of the landing pad detection and identification algorithms based on the simulated data are presented.

6.1 Simulation Environment

6.1.1 Image Generation and 3D Model of Landing Pad

In order to validate the landing pad detection and identification algorithms, X-Plane is utilized to generate images since it has a perfect 3D model of the real world, which gives us diverse real-world scenarios. We create a 3D model for the landing pad based on the layout design in Fig. 3. The side length of the landing pad is 6m, i.e., a=1.5m. The 3D model is introduced to X-Plane as a plugin. Besides, one

additional plugin is developed to control the position and orientation of the camera and active markers. The blinking frequencies of the active markers X, Y, and Z are set to 2Hz, 1Hz, and 0.5Hz respectively. The duty cycles of them are 60%, 70%, and 75%, respectively. The RGB images rendered by X-Plane are recorded via the shared memory. Since we would like to use a monochrome camera and a bandpass optical filter in the application, the RGB images from X-Plane are converted to greyscale images to simulate the monochrome camera.

6.1.2 Camera

In the real world, the coarse intrinsic parameters of the camera depend on the image sensor parameters (including the image sensor size and the pixel size) and the lens's focal length. However, X-Plane does not include the concept of the actual camera. Thus, we use the field of view and the image size to calculate the intrinsic parameters. In our simulation with X-Plane, the horizontal field of view is set to 90°, and the image size is set to 1024×768. The frame rate of the camera is set to 10 frames per second. The camera is assumed to be mounted on the bottom of the aircraft without losing field of view. The relative pose between the camera and the aircraft is set to

$$\mathbf{R}_{bc} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \text{ and } \mathbf{r}_{o_c}^{b} = \mathbf{0}_{3\times 1},$$

which enables the camera to look downwards to the ground when the aircraft is leveled.

6.1.3 Trajectory of Aircraft

Since the whole system, including the onboard camera and the optical markers, is designed to navigate the VTOL aircraft to land on the specific site, a trajectory mimicking the behavior of the VTOL aircraft landing is generated to verify the algorithms of the landing pad recognition and relative pose estimation. Fig. 8 shows the trajectory we used for validation. Before landing, the VTOL aircraft cruises to the center of the landing pad at a height of 30m with a velocity of 5m/s. Then, it vertically lands on the landing pad with a velocity of 2m/s. In the simulation, the attitudes of the aircraft are set to a constant.

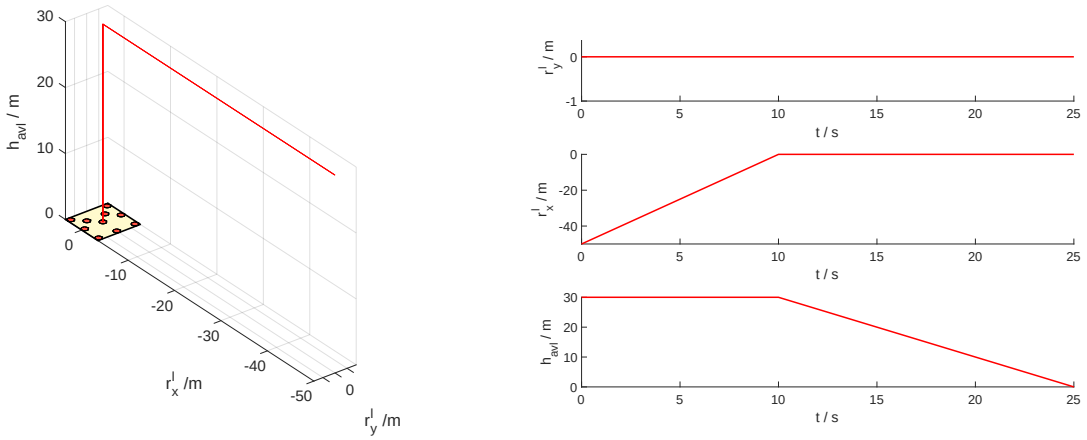
1) Attitude settings for the validation of the modified four-point algorithm: the [roll, pitch, heading] of the aircraft are constantly set to $[0°, 2°, 0°]$, which are closer to the practical attitude during landing.

2) Attitude settings for the validation of the recognition algorithm: if the attitude is set to $[0°, 2°, 0°]$, the appearance of the landing pad in the image sequence are pretty close to a square instead of an arbitrary quadrilateral, which cannot manifest generality. However, if the aircraft is tilted too much, the camera can only capture the landing pad for a short period. After considering the trade-off of the availability and the generality, the attitude is set to $[0°, 15°, 0°]$ in the validation of the landing pad recognition algorithm.

6.2 Validation of Landing Pad Recognition

The results of the landing pad recognition algorithm are divided into two parts for discussion. From 0s to 10s, the VTOL aircraft horizontally cruises to the position above the designated landing site. Throughout the cruising process, the appearance of the landing pad in the airborne images remains the same. In the subsequent phase of the simulation, spanning from 10s to 25s, the aircraft executes a vertical descent onto the landing pad. The appearance of the landing pad captured by the camera undergoes proportional scaling.

During the approach phase, the landing pad appears in the image sequence for the first time at 2.4s, as shown in Fig. 9a. However, the onboard system cannot detect the landing pad since it does not entirely appear in this image. The detection algorithm cannot find two sets of four collinear points. At 2.8s,



(a) Trajectory and the landing pad. **(b) Relative position over time.**

Fig. 8 Reference trajectory of the VTOL aircraft with respect to the landing pad.

the landing pad has the first complete appearance in the simulation, as shown in Fig. 9b. The two sets of four collinear points are found, and thus, the onboard system distinguishes all the optical markers. The cross ratios of the point set ABCD and DEFG (i.e., 1.49700 and 1.48996) are computed to verify the detection of the landing pad. So, it successfully detects the landing pad for the first time. Besides, the identification algorithm starts to track the status of each subsequent image from this image. Fig. 9c shows that the frequency and the duty cycle of the active marker X are calculated for the first time in the simulation (2.0Hz and 60%) at 3.8s with the status track of the marker X for 11 images. At 4.8s, the system has already tracked the active marker Y over 21 images. With analysis of the status over time, the active marker Y's frequency and duty cycle are also obtained for the first time(1.0Hz and 70%). Since the blinking frequency of marker Z is the lowest, the active marker Z's frequency and duty cycle (0.5Hz and 75%) are determined after the system tracks it for 40 images. As of the present frame, recorded at 6.7s, successful recognition of the landing pad has been achieved because the estimated frequencies and duty cycle of the markers X, Y, and Z match the settings of the specific landing pad. We can find that the landing pad is recognized until the active marker with the lowest frequency (i.e., the marker Z) is thoroughly analyzed.

(a) In the image at 2.4s, the landing pad appears in the image sequence. However, the onboard system cannot detect the landing pad since the landing pad is not entirely shown in this image.

(b) In the image at 2.8s, the landing pad shows entirely for the first time, and the onboard system detects the landing pad successfully. In addition, the identification algorithm starts to track the status of each subsequent image from this image.

(c) In the image at 3.8s, the system has already tracked the active marker X over 11 images. The active marker X's frequency and duty cycle are calculated for the first time in the simulation (2Hz and 60%).

(d) In the image at 4.8s, the system has already tracked the active marker Y over 21 images. The active marker Z's frequency and duty cycle are calculated for the first time in the simulation (1.0Hz and 70%).

(e) In the image at 6.7s, the system has already tracked the active marker Z over 40 images. The active marker Z's frequency and duty cycle are calculated for the first time in the simulation (0.5Hz and 75%). Up until the current image, the landing pad is successfully recognized.

Fig. 9 Results of the landing pad recognition algorithm during the approach phase.

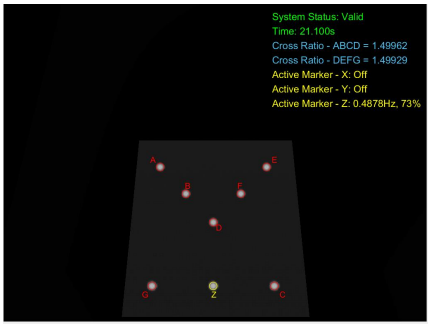
At 10.0s, the VTOL aircraft reaches a point exactly above the landing pad, as shown in Fig. 10a, and starts to vertically land onto the landing pad. Fig. 10b and Fig. 10c shows the results of the landing pad recognition algorithm at 17.7s and 21.1s. Fig. 10d demonstrates the last image (22.3s) in the simulation where the landing pad is recognized. We can find that the markers C and G are pretty close to the image border in this image. At 22.4s, as shown in Fig. 10e, only half of the marker C and G projections (the two at the bottom) are within the image. The blob detector does not successfully extract these two markers from the image. Thus, the system fails to recognize the landing pad since it cannot find two sets of four collinear points from this image.

If the onboard camera is replaced with a camera with larger field of view, the duration of successful detection of the landing pad can be extended under the same trajectory. Generally, we have two ways to increase the field of view, i.e., a larger image sensor and a lens with smaller focal length. Besides a higer economic cost, a larger image senosr with the same pixel size means larger resolution of image, which brings 1) larger communication cost between the camera and the onboard computer and 2) larger computational cost for the whole system. A lens with smaller focal length also involves larger distortion towards the edges of the image, which cannot be easily rectified by camera calibration. In the design stage of the optical landing system, there should be a tradeoff between camera's field of view and the cost in economics, communication, computation, and image distortion.

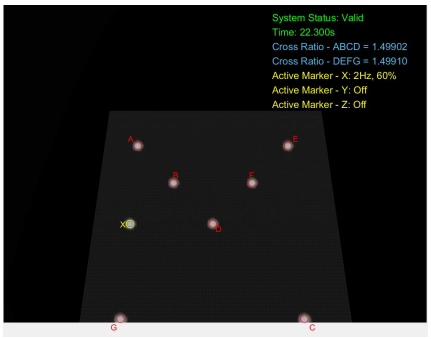
The results show that the landing pad detection algorithm can find the landing pad from a single image as long as the image has a complete appearance of the landing pad. In addition to the finder markers, the active markers are also identified. After a continuous track of the status of the active markers, the landing pad can be identified. The duration of a successful recognition depends on the lowest frequency of active markers. If we need a faster identification of the landing pad, the lowest frequency of the active markers should be increased. To have sufficient distinction, the frequencies of the other active markers should also be increased. Consequently, the frame rate of the camera is required to be increased in order to fulfill the requirements of the Nyquist–Shannon sampling theorem.

(a) In the image at 10.0s, the aircraft positions itself above the specific landing pad and starts to descend vertically onto the landing pad.

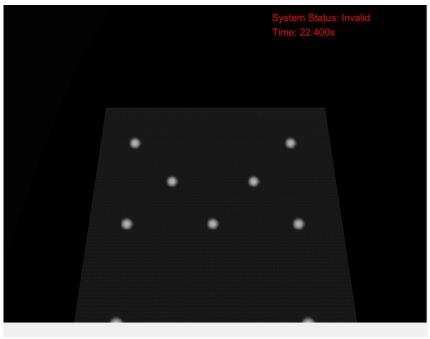
(b) In the image at 17.7s, the landing pad is under track during the landing phase.



(c) In the image at 21.1s, the landing pad is under track during the landing phase.



(d) In the image at 22.3s, the system recognizes the landing pad for the last time.



(e) In the image at 22.4s, the blob detector fails to extract the marker C and G projections (the two at the bottom) from this image since only half of the marker C and G appear in the image. Thus, the system fails to recognize the landing pad since it cannot find two sets of four collinear points from the image.

Fig. 10 Results of the landing pad recognition algorithm during the landing phase.

6.3 Validation of Relative Pose Estimation

To validate the modified four-point algorithm, we do a simulation with the trajectory depicted in Fig. 8b. Zero-mean white noise is added to the ground truth of the measurement, i.e., pixel coordinates achieved by the blob detector, to simulate the blob detector's detection error. In the simulation, we assume that the blob detector's detection error is consistent during the approach and landing phases and the standard deviation of the white noise is 0.5 pixels.

Fig. 11 demonstrates the estimated position and reference trajectory. During the landing phase, the error of the relative position estimate decreases as the VTOL aircraft gradually reaches the landing pad.

In Fig. 12, we can find that the system provides estimates of relative pose from 6s to 23s since the whole landing pad is only observed during that period. During the approach phase (0-10s, height: 30m), the maximum attitude and position errors reach 7.5° and 4m. In the landing phase, the attitude and position errors gradually reduce as the VTOL aircraft reaches the landing pad. The position error reaches the centimeter level just before the system loses track of the landing pad (23s, height: 5m). Therefore, the modified four-point algorithm can find the relative position and attitude. Although the error at the height of 30m is relatively large, it gradually decreases as the height decreases, which can still fulfill the positioning accuracy requirements of the VTOL aircraft landing before touchdown.

We can find that the error at the height of 30m is relatively large. It gradually decreases as the height decreases. During the landing phase, the errors come from the same source at the same quantity, i.e., blob detection error. The reason for the decreasing error over the decreasing height is the geometry between the optical markers and the camera. In the extreme ideal case, the aircraft is at an infinite height from the landing pad. Since all the markers on the image lie on the same point, we have no clues about the position and attitude of the aircraft. In other words, the uncertainty of the estimate is infinite. In another extreme case, the VTOL aircraft touches down the landing pad. The optical center of the airborne camera is located right at the landing pad center. The camera is also assumed to be an omnidirectional one, which is able to observe all the markers. In this case, the uncertainty of the estimate is at the minimum.

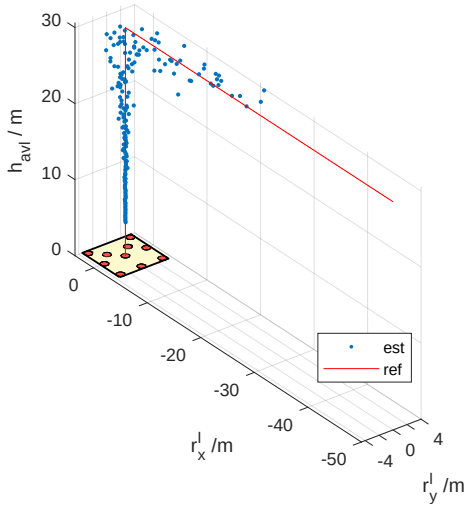
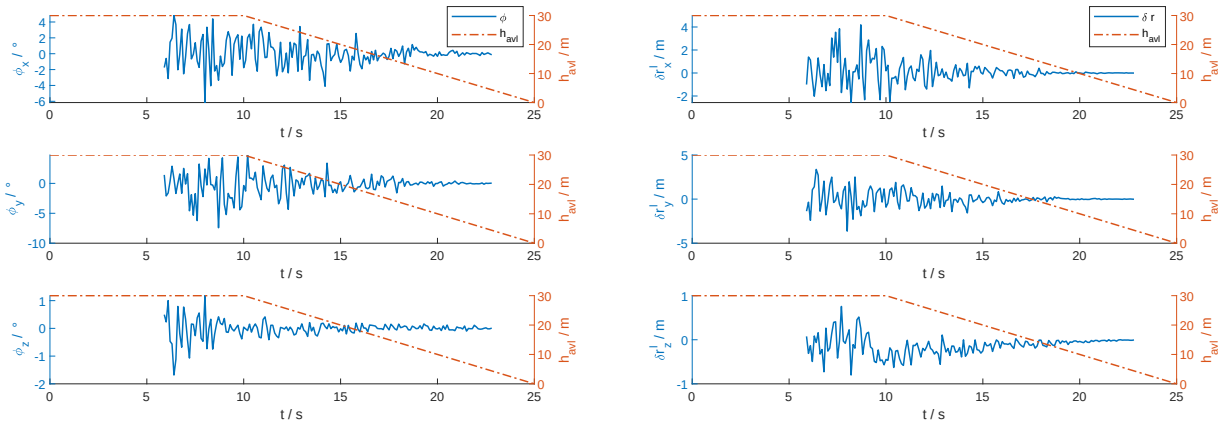


Fig. 11 Reference trajectory and estimated position with the modified four-point algorithm.



(a) Estimation error of relative attitude (vertical speed during approach: 0 m/s).

(b) Estimation error of relative position (vertical speed during approach: 0 m/s).

Fig. 12 Simulation results of the modified four-point algorithm with a landing trajectory.

6.4 Applicability over Different Approach Trajectories

To investigate the applicability over different approach trajectories, we simulate our solution with three additional trajectories. These trajectories are modified based on the trajectory designed for validation of the four-point algorithm as shown in Fig. 8a. During the approach phase, the vertical velocities in these three trajectories are set to 1 m/s, 2m/s, and 3m/s instead of 0 m/s. The VTOL aircraft cruises to the center of the landing pad at a height of 30m and then enters the landing phase. The other settings, i.e., attitude, camera, and landing pad settings, are the same. Fig. 13 illustrates these three trajectories and the original trajectory.

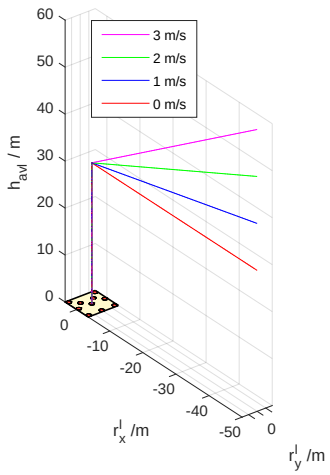
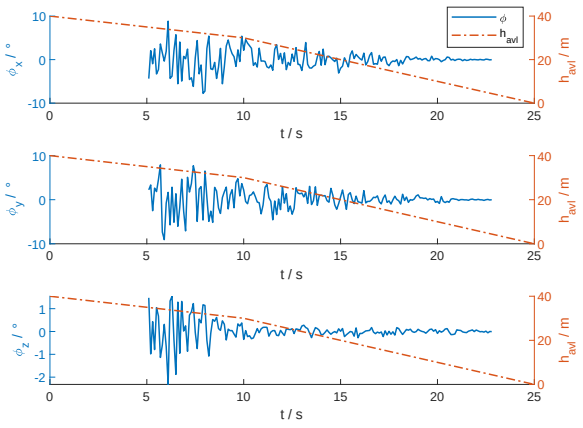


Fig. 13 Trajectories with vertical velocities of 0m/s, 1m/s, 2m/s, and 3m/s during approach.

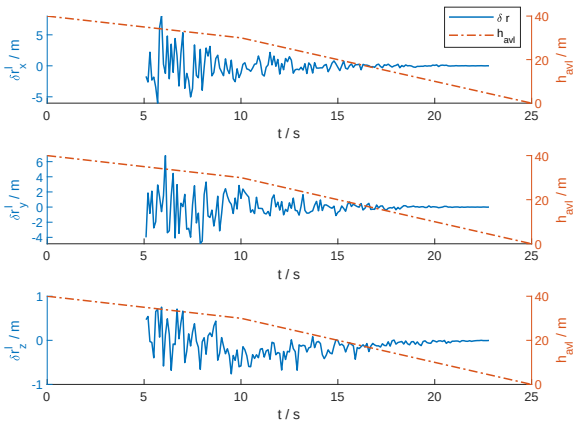
The attitude and position errors from our solution over the additional three trajectories are shown in Fig. 14. For brevity, we call the trajectories with vertical velocities of 0m/s, 1m/s, 2m/s, and 3m/s during approach Trajectory (0m/s), Trajectory (1m/s), Trajectory (2m/s), and Trajectory (3m/s) respectively. We can find that the position and attitude obtained from the optical landing system are still available. However, during the approach phase, the maximal position errors in Trajectory (0m/s), Trajectory (1m/s), Trajectory (2m/s), and Trajectory (3m/s) reach 4m, 7m, 10m, and 30m respectively. The maximal attitude errors also reach 7.5°, 9°, 14°, and 20° respectively. When the VTOL aircraft is far from the landing pad, the pose information with large errors is still available to be delivered. Obviously, the accuracy of position and attitude estimated by the optical landing system during the approach phase is insufficient for the flight controller. In contrast, in all the simulations, the errors are limited to a small range when the aircraft flies closer to the landing pad. Therefore, a flag of validity based on the height above the landing pad, as shown in Eq. (41), can be delivered to the flight controller to indicate whether the error of the position and attitude is good enough.

$$\text{isValid}\left(h_{\text{avl}}\right) = \begin{cases} 1, & h_{\text{avl}} \leq h_{\text{th}}; \\ 0, & h_{\text{avl}} > h_{\text{th}}, \end{cases} \tag{41}$$

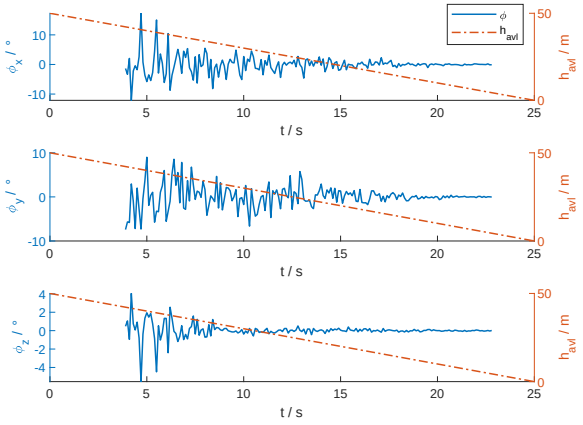
where h_{avl} is the height above the landing pad, h_{th} the height threshold for reasonable position and attitude accuracy. The height above the landing pad can come from a radar altimeter. If no radar altimeter is equipped on the VTOL aircraft, the average height from the optical landing system over a short period can be utilized. The validity flag for the pose information can enhance the safety and robustness of the optical landing system.



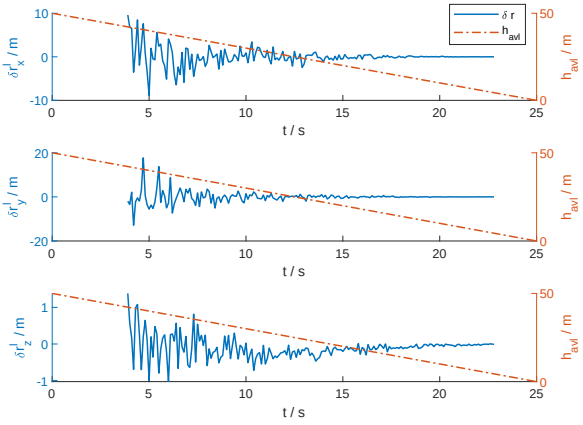
(a) Estimation error of relative attitude (vertical speed during approach: 1 m/s).



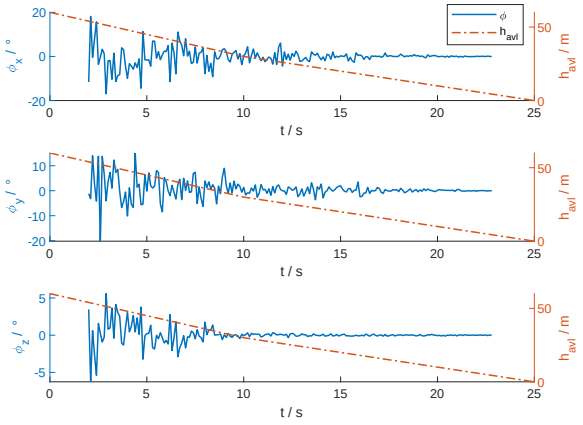
(b) Estimation error of relative position (vertical speed during approach: 1 m/s).



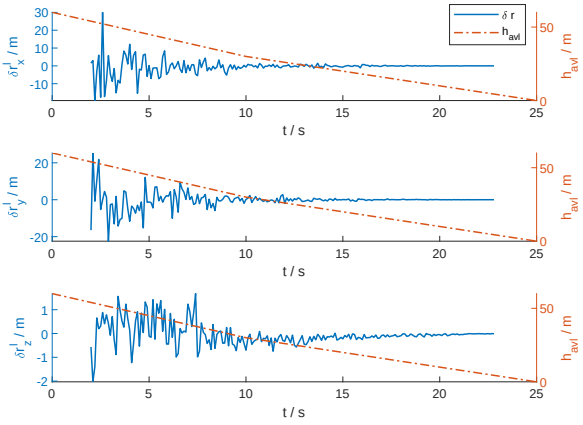
(c) Estimation error of relative attitude (vertical speed during approach: 2 m/s).



(d) Estimation error of relative position (vertical speed during approach: 2 m/s).



(e) Estimation error of relative attitude (vertical speed during approach: 3 m/s).



(f) Estimation error of relative position (vertical speed during approach: 3 m/s).

Fig. 14 Simulation results with additional three approach trajectories.

7 Conclusion

This paper presents a new concept of an optical landing system consisting of an airborne camera system and a landing pad equipped with optical markers. We proves that the mapping between the landing pad plane and the image plane is a projective transformation. Leveraging the invariant measures of the

projective transformation, we propose a layout of optical markers along with the landing pad recognition algorithm. The characteristics of this proposed system are that it can robustly and uniquely recognize the landing pad. A modified four-point algorithm is applied to estimate the position and attitude of the VTOL aircraft. Considering that the normal vector of the landing pad surface is known in our case, we adapt the classic four-point algorithm in this paper to avoid ambiguity resolution. This system is able to provide accurate relative position and attitude of the VTOL aircraft with respect to the landing pad in real time. The use of active light sources in the optical markers enables the system to operate effectively even under low illumination conditions, such as at night. Simulation results demonstrate that the proposed recognition algorithm is able to recognize the landing pad, and the modified four-point algorithms can determine the relative position and attitude of the camera with respect to the landing pad. In summary, the simulations verify the feasibility of the optical marker layout.

In the current study, the optical landing system provides position and attitude outputs without accompanying information regarding their uncertainties. Future research should prioritize the incorporation of uncertainty quantification to enhance safety and fault tolerance for flight navigation. In addition, the VTOL trajectory composed of line segments was employed in this paper to validate algorithms for landing pad recognition and relative pose estimation. For future research, incorporating more dynamically feasible and energy-efficient trajectories could enhance the realism and utility of these simulations [20]. Furthermore, conducting real flight tests is imperative to provide crucial insights into the algorithm's performance under real-world conditions, thereby enhancing its applicability and robustness in practical aviation settings.

References

[1] Christoph Krammer, Markus Rosenbauer, and Florian Holzapfel. Flight guidance for vision-augmented automatic landing of electric vertical take-off and landing vehicles. *IFAC-PapersOnLine*, 55(22):248–254, 2022.

[2] Zhidong Lu, Haichao Hong, Johannes Diepolder, and Florian Holzapfel. Maneuverability set estimation and trajectory feasibility evaluation for evtol aircraft. *Journal of Guidance, Control, and Dynamics*, 46(6):1184–1196, 2023. ISSN: 0731-5090. DOI: 10.2514/1.G007109.

[3] Valentin A. Marvakov and Florian Holzapfel. Defining robust transition and re-transition procedures for unmanned fixed-wing vtol aircraft. In *AIAA Scitech 2021 Forum*, Reston, Virginia, 01112021. American Institute of Aeronautics and Astronautics. ISBN: 978-1-62410-609-5. DOI: 10.2514/6.2021-1634.

[4] European Union Aviation Safety Agency. Special condition for small-category vtol aircraft, 2019.

[5] Christoph Krammer, Simon Scherer, Chinmaya Mishra, and Florian Holzapfel. Concept for a vision-augmented automatic landing system for vtol aircraft. In *AIAA AVIATION 2021 FORUM*, page 3217, 2021.

[6] M. Angermann, S. Wolkow, A. Dekiert, U. Bestmann, and P. Hecker. Fusion of dual optical position solutions for augmentation of gnss-based aircraft landing systems. In *Proceedings of the 2019 International Technical Meeting of The Institute of Navigation*, The International Technical Meeting of the The Institute of Navigation, pages 283–295. Institute of Navigation, 2019. DOI: 10.33012/2019.16694.

[7] Liu, Zhang, and Tian. An onboard vision-based system for autonomous landing of a low-cost quadrotor on a novel landing pad. *Sensors*, 19(21):4703, 2019. DOI: 10.3390/s19214703.

[8] Shanggang Lin, Lianwen Jin, and Ziwei Chen. Real-time monocular vision system for uav autonomous landing in outdoor low-illumination environments. *Sensors (Basel, Switzerland)*, 21(18), 2021. DOI: 10.3390/s21186226.

[9] M. Fiala. Artag, a fiducial marker system using digital techniques. In Cordelia Schmid, Stefano Soatto, and Carlo Tomasi, editors, *CVPR 2005*, pages 590–596, Los Alamitos Calif., 2005. IEEE Computer Society. ISBN: 0-7695-2372-2. DOI: 10.1109/CVPR.2005.74.

[10] Zongyi He, Zongguo Xia, Yong Chang, Weirong Chen, Jinshan Hu, and Xiuqin Wei. Research on underground pipeline augmented reality system based on artoolkit. In Huayi Wu and Qing Zhu, editors, *Geoinformatics 2006: Geospatial Information Technology*, SPIE Proceedings, page 642112. SPIE, 2006. DOI: 10.1117/12.713123.

[11] Edwin Olson. Apriltag: A robust and flexible visual fiducial system. In *2011 IEEE International Conference on Robotics and Automation*, pages 3400–3407. IEEE, 5/9/2011 - 5/13/2011. ISBN: 978-1-61284-386-5. DOI: 10.1109/ICRA.2011.5979561.

[12] S. Garrido-Jurado, R. Muñoz-Salinas, F. J. Madrid-Cuevas, and M. J. Marín-Jiménez. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47(6):2280–2292, 2014. ISSN: 00313203. DOI: 10.1016/j.patcog.2014.01.005.

[13] Guili Xu, Xiaopeng Qi, Qinghua Zeng, Yupeng Tian, Ruipeng Guo, and Biao Wang. Use of land's cooperative object to estimate uav's pose for autonomous landing. *Chinese Journal of Aeronautics*, 26(6):1498–1505, 2013. ISSN: 10009361. DOI: 10.1016/j.cja.2013.07.049.

[14] Joshua Springer, Gylfi Þór Guðmundsson, and Marcel Kyas. A precision drone landing system using visual and ir fiducial markers and a multi-payload camera, 2024.

[15] Yi Ma, Stefano Soatto, Jana Košecká, and Shankar Sastry. *An invitation to 3-d vision: from images to geometric models*, volume 26. Springer, 2004.

[16] Z. ZHANG. 3d reconstruction based on homography mapping. *Proc. ARPA96*, pages 1007–1012, 1996.

[17] Ezio Malis and Manuel Vargas Villanueva. Deeper understanding of the homography decomposition for vision-based control. `https://idus.us.es/handle/11441/128196`.

[18] Nicholas J. Higham. *Accuracy and stability of numerical algorithms*. Society for Industrial and Applied Mathematics, Philadelphia, 2nd ed. edition, 2002. ISBN: 0898715210.

[19] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision*. Cambridge University Press, Cambridge, second edition edition, 2004. ISBN: 9780511811685. DOI: 10.1017/CBO9780511811685.

[20] Zhidong Lu, Haichao Hong, and Florian Holzapfel. Multi-phase vertical take-off and landing trajectory optimization with feasible initial guesses. *Aerospace*, 11(1):39, 2023. DOI: 10.3390/aerospace11010039.