2020 International Conference on Unmanned Aircraft Systems (ICUAS)
Athens, Greece. September 1-4, 2020

Image-Based Visual Servo Control for Ground Target Tracking Using a Fixed-Wing UAV with Pan-Tilt Camera

Lingjie Yang[1], Zhihong Liu*[1], Guanzheng Wang[1], Xiangke Wang[1]

Abstract— **This paper proposes a control framework to achieve the tracking of the moving target by a fixed-wing unmanned aerial vehicle (UAV) with a monocular pan-tilt camera. This control framework is based on the image-based visual servoing (IBVS) method, which takes the target feature point on the captured image as the input and outputs the control signals directly with the aid of image Jacobian matrix. However, the image is affected by the attitude of both the UAV and the pan-tilt, and the attitude of the pan-tilt is coupled with that of the UAV simultaneously. To solve this problem, we present an *"Ideal State"* as the reference state, and make sure the coordinates of the feature point in the state are only affected by the change of the yaw angle of the UAV. In this way, we can integrate the attitude control of the UAV and the pan-tilt. By using this control framework, the fixed-wing UAV can track the ground target continuously on the one hand, and the target will tend to locate at image center on the other hand. This prevents the target from moving toward to the edge of the image or even disappearing. Besides, we prove the controller is exponentially convergent by the Lyapunov method. In order to evaluate the performance of our controller, we build a hardware-in-the-loop (HIL) simulation platform and a prototype platform. Based on these platforms, extensive experiments including simulations and real flight tests are conducted. The results show that our controller can achieve continuous and robust tracking of the target with a speed of 20km/h when the speed of the UAV is 16m/s.**

I. INTRODUCTION

The past decade has witnessed a rapid development of UAVs. They can execute dull, dirty or dangerous tasks instead of humans and have been widely used in both the civilian and military fields. Due to the excellent maneuverability of the UAVs, applying them in ground target tracking has tremendous benefits [1], [2], [3], [4]. Considering that the fixed-wing UAVs have longer navigation time, larger payload and faster flight speed than rotor UAVs, the fixed-wing UAVs are more suitable for the long endurance tracking tasks.

Visual servo control is a commonly used tracking method [5], [6], [7]. There are mainly two categories [8], [9], [10]: position-based visual servoing (PBVS) and image-based visual servoing (IBVS). PBVS requires the controller to be designed in 3D Cartesian space, thus the camera needs to be calibrated for the transformation of coordinate frame. Basically, most of the existing work of target tracking for the UAVs adopt PBVS [11], [12], [13]. In this approaches, it only allows an optimal trajectory in 3D Cartesian space to be followed theoretically, instead of in the image space. As a result, even small errors in the image measurements can lead to errors in the pose [9], which may affect the tracking performance. Compared to PBVS, IBVS directly defines the difference between the current and desired image characteristics as the control error, thereby eliminating the process of 3D modeling, and insensitive to the calibration errors of the sensors, cameras and robots. Therefore, this kind of methods can better ensure that the target is in the camera field of view [14].

Recently, applying the IBVS technique to tackle the vision based control problem in UAVs is more and more common. Serra et al. [15], [16] use a quadrotor to track a moving platform and then land on it with the aid of IBVS. Srivastava et al. [17] control the quadrotor to keep following a target drone while maintaining a fixed distance from it in the absence of GPS. Lyu et al. [18] propose a framework to realize vision-based multi-UAV cooperative mapping and control based on IBVS. Although a number of approaches based on IBVS for UAVs emerge, most of the existing work is proposed for rotor UAVs [19], [20], [21], [22]. Compared to rotor UAVs, the fixed-wing UAVs have much more dynamical constraints, such as it cannot move omni-direction on the one hand, and its minimum speed is limited by the stalling speed on the other hand. Therefore, applying IBVS in target tracking for fixed-wing UAVs is more challenging.

To this end, Florent et al. [23] design a controller for the fixed-wing UAV with a fixed camera to track a ground target. Their work enforces the trajectory of the UAV to converge to a cone and a plane at the same time, and only the tracking of stationary target is studied. Pietro et al. [24] achieve the target tracking for the fixed-wing UAV equipped with a pan-tilt camera. They require the pan-tilt to tend to be perpendicular to the body, which causes the yaw capability of the pan-tilt cannot be fully utilized. As a result, the target may be out of sight when it moves fast. Wang et al. [25] present a framework for tracking a mobile ground target using a fixed-wing UAV. Yet the design of guidance law in this approach is based on target localization, which is easily affected by the calibration of the camera. In our previous work [26], we have studied the tracking of ground target by fixed-wing UAV with fixed camera. Nevertheless, due to the limited sight of view and the fixed attitude of the camera, it is easy to lost the target when the target moves fast.

In this work, we propose a control framework based on IBVS for tracking a moving target by a fixed-wing UAV with a pan-tilt camera. This approach integrates the attitude

This work was funded by the National Natural Science Foundation of China (61906209) and (61973309)

[1]authors are with the College of Intelligence Science and Technology, National University of Defense Technology, Changsha, China. ljyang13@163.com, zhliu@nudt.edu.cn, guanzhengw@163.com, xkwang@nudt.edu.cn

* corresponding author

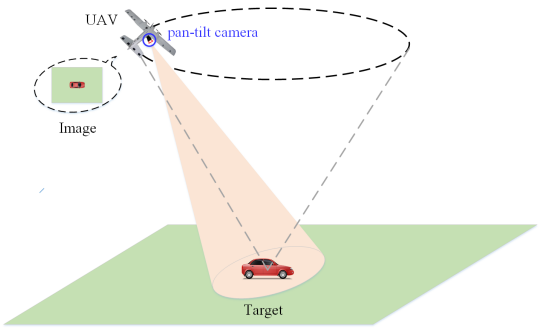


Fig. 1: The ground target tracking problem for a fixed-wing UAV with a pan-tilt camera.

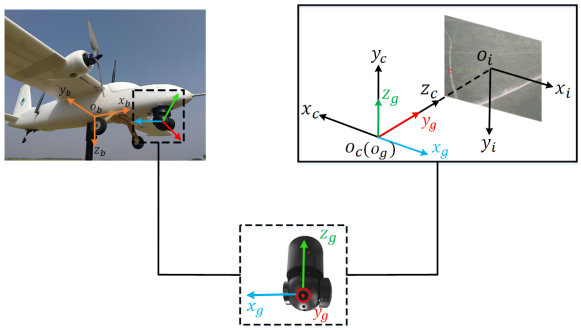


Fig. 2: The relationships among four coordinate frames.

control of the pan-tilt camera and the UAV. Hence, the UAV can loiter over the moving target while maintaining the target at the image center. In this case, the target will not be out of sight even if it moves fast. Besides, we build a HIL simulation platform and a prototype platform. Based on these platforms, extensive experiments including simulations and real flight tests are conducted to evaluate our controller. The results show that our controller can achieve continuous robust tracking of the target, and the speed of the target can be as fast as 20km/h when that of the UAV is 16m/s. The contributions of our work are:

- Proposing a framework to integrate the attitude control of the pan-tilt camera and the fixed-wing UAV, which ensures that the UAV can loiter over the moving target while maintaining the target at the image center.
- Conducting extensive experiments including HIL simulations and field experiments. To the best of our knowledge, the field experiments for the moving target tracking by the fixed-wing UAV based on IBVS are barely performed in the existing work. This highlights the feasibility and superiority of our controller.

The paper is organized as follows. Section II describes the problem of target tracking for the fixed-wing UAV with a pan-tilt camera. Section III introduces the main work of the paper, including the design of the controller and proof of the algorithm. Section IV provides the results and analysis for the HIL simulations and field experiments. Section V gives the conclusion of the paper.

II. PROBLEM STATEMENT

In this section, we provide the description of the problem to be addressed and the overall framework of the proposed controller.

A. Problem Description

We consider the ground target tracking problem for a miniature fixed-wing UAV equipped with a pan-tilt camera in this work. As Fig. 1 shows, the target lies in the image center with the aid of the pan-tilt when the UAV is tracking it. It is known that the fixed-wing UAVs have a minimum stall speed limit, and can only track the target by loitering, rather than hovering of rotor UAVs. Therefore, the design of the controller for the fixed-wing UAVs is more complicated than the rotor UAVs.

To illustrate the relationships among the UAV, the pan-tilt and the camera, four right-hand Cartesian frames are constructed, shown in Fig. 2.

(a) The body coordinate frame ($\mathbb{F}_B$)

The origin of $\mathbb{F}_B$ is located at the center of mass of the fixed-wing UAV. Besides, x_b-axis is along the heading direction of the UAV, and y_b-axis points to the right side of the UAV.

(b) The pan-tilt coordinate frame ($\mathbb{F}_G$)

The top of the pan-tilt is fixed to the body, and we denote the center of the connecting plane as the origin of $\mathbb{F}_G$. Besides, y_g-axis is along the optic axis and x_g-axis points to the right side of y_g-axis.

(c) The camera coordinate frame ($\mathbb{F}_C$)

The optical center of the camera is defined as the origin of $\mathbb{F}_c$, and the optic axis is denoted by z_c-axis. Besides, x_c-axis is opposite to x_g-axis.

(d) The image coordinate frame ($\mathbb{F}_I$)

The origin of $\mathbb{F}_I$ lies in the center of the image, x_i-axis and y_i-axis point to the right and bottom of the image, respectively.

Assume that the UAV autopilot has a speed controller that holds the speed constant (or nearly so) and has a height holding controller for constant UAV altitude. Actually, the assumption is commonly used in related tasks performed by the fixed-wing UAVs [27]. Therefore, a simplified kinematic model in two dimensional plane is considered. Besides, the paper adopts a pan-tilt with two limited degrees of freedom for the reason of low cost and weight, and its yaw and pitch rate are denoted by $\dot{\theta}_p$ and $\dot{\theta}_t$, respectively. Furthermore, we denote current speed of the UAV (denoted by V_t) along the inertial frame by $\{V_x, V_y\}$, and yaw angle of the UAV by ψ. Therefore, the "UAV & pan-tilt" model can be represented by Eq. (1).

$$\begin{cases} V_x = V_t \cdot cos\psi \\ V_y = V_t \cdot sin\psi \\ \dot{\psi} = u_\psi \\ \dot{\theta}_p = u_p, \quad \theta_p \in (\gamma_{p_1}, \gamma_{p_2}) \\ \dot{\theta}_t = u_t, \quad \theta_t \in (\gamma_{t_1}, 0) \end{cases}, \tag{1}$$

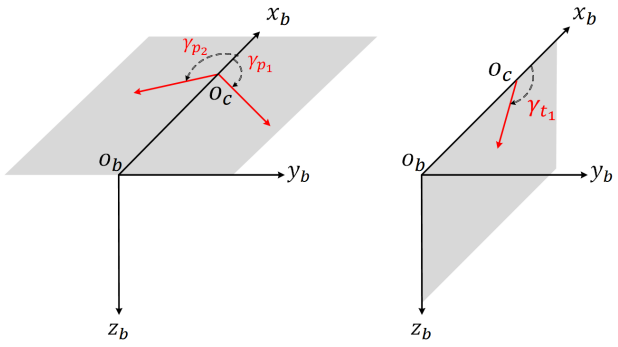


Fig. 3: The range of the pan-tilt attitude.

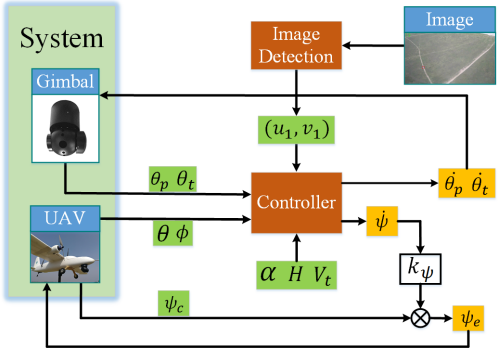


Fig. 4: The framework of the controller.

where u_ψ, u_p and u_t are control inputs. The range of the pan-tilt attitude is shown as Fig. 3, $\gamma_{p_1}(<0)$, $\gamma_{p_2}(>0)$ and $\gamma_{t_1}(<0)$ are the thresholds of the pan-tilt attitude.

Different from the conventional position-based tracking method, we intend to propose an IBVS control law for the UAV and the pan-tilt together by using the image information directly, without target positioning processes in $\mathbb{F}_C$. However, there exist two challenges in the work as follows.

- A small jitter of the UAV attitude will cause a dramatical jump for the position of the target in the image, and the attitude of the fixed-wing UAV is dynamically changed all the time during the mission. Therefore, how to implement the tracking under this circumstance is challenging.
- Both the attitude of the pan-tilt and that of the UAV are coupled, how to integrate the attitude control of them is challenging as well.

B. Overall Framework

In this subsection, we design the overall framework of the visual control scheme (Fig .4). The pitch, roll and yaw angle of the UAV in current state are denoted by $\{\theta, \phi, \psi_c\}$ in order, and the desired yaw angle of the UAV is denoted by ψ_e. Besides, H represents the flight height, α represents a desired pitch angle of the pan-tilt, and k_ψ is a coefficient.

Firstly, an *image detection* module is used to obtain the target information, which is the target position (u_1, v_1) in $\mathbb{F}_I$. And then, the *servo control* module is performed with the aid of the (u_1, v_1) information and the UAV's state, to obtain the desired u_ψ, u_p and u_t.

1) Image Detection: UAV-based image detection is in charge of processing image and detecting target. It takes the image captured by the airborne camera as the input, by locating the position of the target on the image, and outputs the coordinates of the target feature point (denoted by (u_1, v_1) in $\mathbb{F}_I$). It should be noted that high requirement of accuracy and real-time performance is necessary in field experiments. To the best of our knowledge, there exist two kinds of image detection methods [28], which can be divided into traditional method and neural network method. Compared to the former, the latter can obtain (u_1, v_1) directly from the image, which is more efficient. Besides, considering that this part of image detection is not the focus of the paper, we adopt existing method named YOLOv3 [29] to solve the problem directly.

2) Servo Control: Image-based visual servoing is able to directly map the desired velocity of feature point into the motion velocity of the UAV. With the aid of it, there is no need to calculate the relative position of the target to the UAV. As a result, the generated error in this process can be eliminated.

We denote the desired velocity of the centroid coordinates by $\dot{S}(\dot{u}, \dot{v})$ in $\mathbb{F}_C$, and it is related with the altitude of the UAV and the pan-tilt. Simultaneously, the linear velocity of the UAV is denoted by $T = (V_x, V_y, V_z)^T$, and angular velocity of it is denoted by $\Omega = (\omega_x, \omega_y, \omega_z)^T$. There exists image Jacobian matrix J_v that associates $\dot{S}$ with $\{T, \Omega\}$, which is shown as Eq. (2).

$$\begin{pmatrix} \dot{u} \\ \dot{v} \end{pmatrix} = J_v \cdot \begin{pmatrix} T \\ \Omega \end{pmatrix}, \tag{2}$$

where $J_v \in \mathbb{R}^{2\times 6}$ [9], and we can obtain the control output directly from $\dot{S}$ through the inverse of the equation. Simultaneously, we need to adjust θ_p and θ_t to make (u_1, v_1) tend to $(0, 0)$.

III. MAIN WORK

A. Control Law Design

During the target tracking process, the attitude of the UAV $\{\psi, \theta, \phi\}$ is coupled with that of the pan-tilt $\{\theta_p, \theta_t\}$. As a result, the centroid coordinates of the target are affected by the attitude of both the UAV and the pan-tilt. Furthermore, the coordinates of the desired feature point changes at runtime before the UAV reaches a stable loitering state, which causes the unknown of $\dot{S}$.

In order to make $\dot{S}$ unaffected by the change of $\{\theta, \phi, \theta_p, \theta_t\}$, we propose a reference state named *"Ideal State"* (Definition 1). Through this method, we map the current system state into an reference state, and $\dot{S}$ in the latter can be determined uniquely. Moreover, $\dot{S}$ is only affected by the change of ψ, then u_ψ can be obtained on the basis.

Definition 1 (Ideal State): The "UAV & pan-tilt" system that the model corresponding to must simultaneously satisfy the following four conditions: (a) The pitch angle of the UAV is 0°; (b) The roll angle of the UAV is 0°; (c) The pitch angle of the pan-tilt is α; (d) The yaw angle of the pan-tilt is 90°.

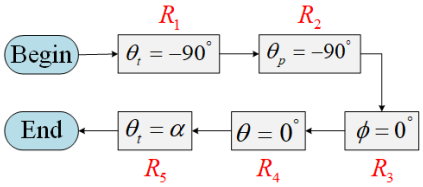


Fig. 5: Flow chart of the transformation for the camera coordinate frame.

When the fixed-wing UAV loiters clockwise over the target, according to the signals of the angles in Fig. 3, the ideal state for the pan-tilt can be represented as: $\theta_p = -90°$, $\theta_t = \alpha(\alpha < 0)$. In order to obtain $\dot{S}$, we need to transform current state of the system into the reference state. The transformation process is shown as Fig. 5, where $R_i(i \in \{1,2,3,4,5\})$ represents the rotation matrix that changes the current state into the state inside the corresponding rectangle, with the other state unchanged. Besides, all matrices are based on the rotation of $\mathbb{F}_C$.

The coordinates of the target in current $\mathbb{F}_C$ are denoted by (X_1, Y_1, Z_1), then the corresponding coordinates under the reference state can be expressed as:

$$\begin{pmatrix} X_2 \\ Y_2 \\ Z_2 \end{pmatrix} = R_5 \cdot R_4 \cdot R_3 \cdot R_2 \cdot R_1 \begin{pmatrix} X_1 \\ Y_1 \\ Z_1 \end{pmatrix} \tag{3}$$

We define $R = R_5 \cdot R_4 \cdot R_3 \cdot R_2 \cdot R_1$. According to the principle of pinhole imaging, the relationship between (X_1, Y_1, Z_1) and (u_1, v_1) satisfies the triangle similarity. The corresponding similarity coefficient is defined as k. We denote the focal length of the camera by f, which can be obtained directly from the adopted camera parameters. After letting $Z_2 = -f$, we can get the corresponding coordinates (u_2, v_2) under the ideal state.

$$\begin{pmatrix} u_2 \\ v_2 \end{pmatrix} = k \cdot R^{(1-2,\cdot)} \begin{pmatrix} u_1 \\ v_1 \end{pmatrix}, \tag{4}$$

where $R^{(1-2,\cdot)}$ represents first two rows of R, and the coefficient k can be calculated by the third line of Eq. (3).

In the *"Ideal State"*, the desired feature point is the center of the image, then the desired rate of the feature point can be obtained:

$$\begin{pmatrix} \dot{u}_2 \\ \dot{v}_2 \end{pmatrix} = -\begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \begin{pmatrix} u_2 \\ v_2 \end{pmatrix} = -\lambda \begin{pmatrix} u_2 \\ v_2 \end{pmatrix}, \tag{5}$$

where $\{\lambda_1, \lambda_2\} \in \mathbb{R}^+$.

During the control, we design u_ψ based on Eq. (5) to make the feature point tend to locate at the image center, thereby achieving the tracking. However, the target is moving and the speed is unknown, then the actual $(\dot{u}_2, \dot{v}_2)$ are different, and the feature point may even deviate from the image center. Therefore, we adopt the pan-tilt camera to adjust the angle of view, so that the target can be kept near the image center during the tracking.

When analyzing the control of the pan-tilt, we hope the target can be as close as possible to the center of the image

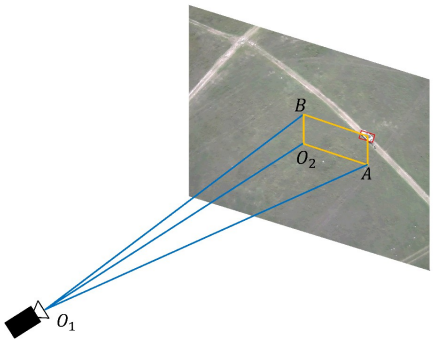


Fig. 6: Relationships between the position of the pan-tilt camera and coordinates of the feature point.

by adjusting θ_p and θ_t. The reason is that it makes poor detection at the edge of the image, which causes the target to be lost easily. As Fig. 6 shows, O_1O_2 is the optic axis, O_2 is the center of the image, and $(|O_2A|, -|O_2B|)$ represents the coordinates of the target in $\mathbb{F}_I$. With the pan-tilt rotating, the changes of θ_p and θ_t will affect the abscissa and ordinate of the target, respectively. Since $\Delta\theta_p < 0$ when the pan-tilt deflects to the right and $\Delta\theta_t < 0$ when it rotates downward, the relationship between them can be expressed as:

$$\begin{cases} tan(\Delta\theta_p) = -|O_2A|/|O_1O_2| \\ tan(\Delta\theta_t) = |O_2B|/|O_1O_2| \end{cases} \tag{6}$$

Since the feature point needs to tend to locate at the image center, $\dot{u}_1$ and $\dot{v}_1$ can be expressed as:

$$\begin{pmatrix} \dot{u}_1 \\ \dot{v}_1 \end{pmatrix} = -\begin{pmatrix} \lambda_u & \\ & \lambda_v \end{pmatrix} \begin{pmatrix} u_1 \\ v_1 \end{pmatrix} = -\lambda_{uv} \begin{pmatrix} u_1 \\ v_1 \end{pmatrix}, \tag{7}$$

where $\{\lambda_u, \lambda_v\} \in \mathbb{R}^+$.

We denote the depth from the target to the camera by z, the element in i-th row and j-th column of J_v by $J_v^{(i,j)}$. Based on the previous work (see detail in Theorem 3 in [26]), we have

$$\begin{pmatrix} M_1 \\ M_2 \end{pmatrix} \dot{\psi} = \begin{pmatrix} \dot{u}_2 - \frac{f}{z} V_t \\ \dot{v}_2 \end{pmatrix}, \tag{8}$$

where

$$M_1 = -J_v^{(1,5)} \cdot sin\alpha + J_v^{(1,6)} \cdot cos\alpha,$$
$$M_2 = -J_v^{(2,5)} \cdot sin\alpha + J_v^{(2,6)} \cdot cos\alpha.$$

By using the least square method, we are able to obtain u_ψ. Furthermore, combining Eqs. (6) (7), the control law of the "UAV & pan-tilt" can be represented as

$$\begin{cases} u_\psi = \frac{M_1(\dot{u}_2 - \frac{f}{z}V_t) + M_2\dot{v}_2}{M_1^2 + M_2^2} \\ u_p = arctan(\frac{-\lambda_u u_1}{f}) \\ u_t = arctan(\frac{-\lambda_v v_1}{f}) \end{cases}, \tag{9}$$

Furthermore, we analyze the effect of the parameter z on u_ψ. Compared with u_ψ when z is measured in real time, it will only cause a small deviation of each yaw amplitude when z is a constant, without affecting the trend of the UAV circling and tracking the target.

The implementation of the algorithm is shown as Algorithm 1.

Algorithm 1 Using coordinates of the feature point to obtain the control of the UAV and the pan-tilt

Require: the image captured by camera
Ensure: the control of UAV and pan-tilt

1: **while** discover the target **do**
2: detect the centroid coordinates (u_1, v_1);
3: calculate the rotation matrices R_1, R_2, R_3, R_4, R_5;
4: obtain the transformed centroid coordinates:

$$\begin{pmatrix} u_2 \\ v_2 \end{pmatrix} = k \cdot (R_5 \cdot R_4 \cdot R_3 \cdot R_2 \cdot R_1)^{(1-2,\cdot)} \begin{pmatrix} u_1 \\ v_1 \\ -f \end{pmatrix}$$

5: obtain the velocity of the transformed target:

$$\begin{pmatrix} \dot{u_2} \\ \dot{v_2} \end{pmatrix} = - \begin{pmatrix} \lambda_1 & \\ & \lambda_2 \end{pmatrix} \begin{pmatrix} u_2 \\ v_2 \end{pmatrix}$$

6: calculate M_1 and M_2:

$$M_1 = -J_v^{(1,5)} \cdot sin\alpha + J_v^{(1,6)} \cdot cos\alpha$$
$$M_2 = -J_v^{(2,5)} \cdot sin\alpha + J_v^{(2,6)} \cdot cos\alpha$$

7: obtain the yaw rate of the UAV:

$$u_\psi = \frac{M_1(\dot{u_2} - \frac{f}{z}V_t) + M_2\dot{v_2}}{M_1^2 + M_2^2}$$

8: obtain the velocity of the target

$$\begin{pmatrix} \dot{u_1} \\ \dot{v_1} \end{pmatrix} = - \begin{pmatrix} \lambda_u & \\ & \lambda_v \end{pmatrix} \begin{pmatrix} u_1 \\ v_1 \end{pmatrix}$$

9: obtain the deflection rate of the pan-tilt:

$$\begin{cases} u_p = arctan(\frac{\dot{u_1}}{f}) \\ u_t = arctan(\frac{\dot{v_1}}{f}) \end{cases}$$

10: **end while**

B. Stability Analysis

We denote the error between (u_1, v_1) and image center by e_1, and the error between (u_2, v_2) and image center by e_2. The convergence of e_1 to 0 indicates that the feature point tends to locate at the image center with the aid of the pan-tilt. Simultaneously, the UAV will circle around the stationary target when e_2 converges to 0. Furthermore, we denote the error state by $e = (e_1, e_2)^T$.

Theorem 1: With the controller designed as Eq. (9), when we define

$$\dot{e} = P_e \cdot e, \qquad P_e = -diag\{\lambda_u, \lambda_v, \lambda_1, \lambda_2\}$$

the UAV is able to asymptotically track the stationary target while maintaining it at the image center.

Proof: According to Eqs. (6) (8), we can construct the following equations to obtain the control output:

$$\dot{e} = \begin{pmatrix} f & & \\ & f & \\ & & M_1 \\ & & M_2 \end{pmatrix} \begin{pmatrix} tan(u_p) \\ tan(u_t) \\ u_\psi \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ \frac{fV_t}{z} \\ 0 \end{pmatrix}$$

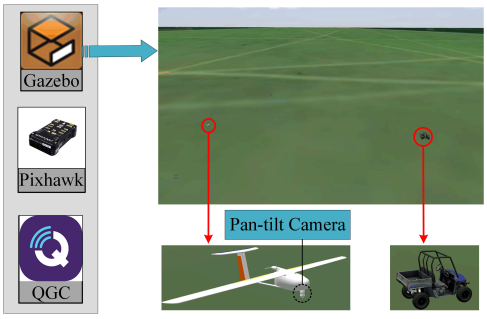


Fig. 7: The HIL simulation environment.

Set a Lyapunov function

$$V = \frac{1}{2} e^T \cdot e,$$

it is obvious that $V > 0$, and the time derivative of V can be represented as

$$\dot{V} = e^T \cdot \dot{e} = e^T \cdot P_e \cdot e, \tag{10}$$

Since P_e is negative, and we choose

$$\eta = ||P_e||_\infty,$$

then Eq. (10) can be changed into

$$\dot{V} < \eta \cdot e^T \cdot e < 0, \tag{11}$$

The result shows that e is exponentially convergent. Therefore, the algorithm is exponentially asymptotically stable. Q.E.D. ■

IV. Experiments And Results

In this section, we build a HIL simulation platform based on Gazebo and complete the preliminary validations of the controller. Furthermore, we conduct field experiments to evaluate the performance of our controller.

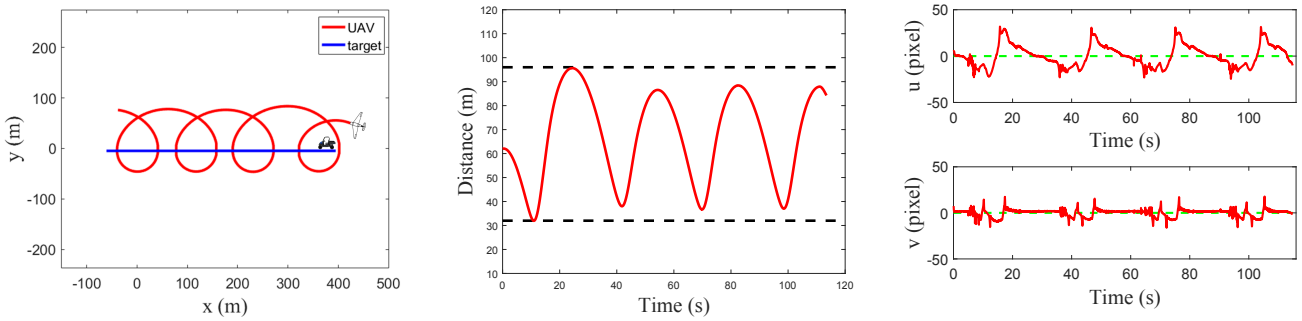
A. HIL Simulations

1) Simulation Setup: The HIL simulation platform is composed of three parts: Gazebo, Pixhawk and QGroundControl (QGC), as Fig. 7 shows.

Gazebo is a simulator that offers the ability to accurately and efficiently simulate populations of robots in complex environments. The weight of the fixed-wing UAV is 2.65kg, the wing surface and wingspan are 0.47m^2 and 2.59m, respectively. With a 60° field of view, the camera captures images of 1280 × 720 in pixels. Besides, the weight of the car is 771kg, and its front wheels can be deflected to change the motion direction.

Pixhawk is a high performance autopilot that is widely used in small UAVs. With the aid of MAVLink protocol, we implement the communication between Pixhawk and the UAV in Gazebo.

QGC is a ground control station that is adapted to the Pixhawk autopilot. Similarily, it communicates with the autopilot via the MAVLink protocol. We use it to observe the flight state (flight speed, altitude, trajectory, etc.) of the UAV in real time.



(a) (b) (c)

Fig. 8: The fixed-wing UAV tracks the target of uniform linear motion in the HIL simulation. The speed of the target is 4m/s, the flight height and speed of the UAV are 100m and 16m/s, respectively. (a) Tracking trajectories; (b) Horizontal distance between the target and the fixed-wing UAV; (c) Image stabilizing precision.

2) Simulation Results and Analysis: To evaluate the performance of our controller, we record the position information of the UAV and the target in Gazebo. After that, the results are presented and analyzed through plotting the trajectories in MATLAB.

In this experiment, the fixed-wing UAV tracks the target with uniform linear motion at a constant speed. The velocity and flight height of the fixed-wing UAV is 16m/s and 100m, respectively. Simultaneously, the target keeps moving forward at a constant speed of 4m/s. The values of $\{\lambda_1, \lambda_2, \lambda_u, \lambda_v\}$ in our work are all chosen as 0.5.

Fig. 8a shows the tracking trajectory under this circumstance. Note that we conduct the experiments under this circumstance several times and similar results are found. It is obvious that the fixed-wing UAV can circle and track the moving target.

Besides, we depict the horizontal distance between the UAV and the target during the experiment in Fig. 8b. When the target moves, the distance will fluctuate within a small range. The two black dashed lines represent the maximum range of the distance is approximately [32, 96].

Fig. 8c shows the image stabilizing precision for the experiment. When the target makes a uniform linear motion, both u and v will fluctuate regularly around 0. It can be easily found that the maximum of the variation will not exceed 50 pixels.

Furthermore, in order to embody the superiority of the pan-tilt camera in target tracking, we compare the tracking performance of the fixed-wing UAV with a pan-tilt camera and that with a fixed camera. We set the flight speed and height of the UAV as 16m/s and 100m, respectively. Besides, the motion speed of the target is 4m/s. The result is shown in Fig. 9 indicates that with the aid of the adjustment of θ_p and θ_t, the distance between the feature point and image center in the image coordinate will not exceed 50 pixels.

However, since the camera is fixed as Fig. 9b shows ($\theta_p = -90°$, $\theta_t = -45°$), the distance between the feature point and the image center in the image coordinate will change greatly. As the green line shows, the distance remains at about 230 pixels when the target is stationary, what accounts for it not near 0 is that the attitude of the camera is fixed. When the target begins to move after 25s, the distance changes and starts to increase sharply after 50s. As a result, the target is out of sight at about 60s, and we denote the distance after that by (0, 0) to make the UAV circle around where the target disappears.

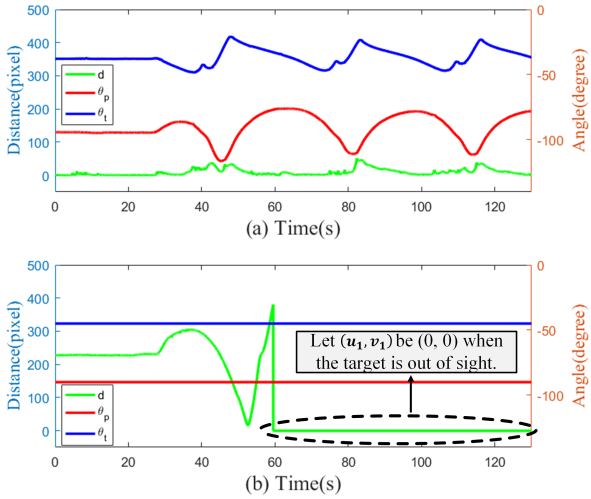


Fig. 9: The comparison of tracking performance between the fixed-wing UAV with a pan-tilt camera (figure a) and that with a fixed camera (figure b). Green line represents the distance between the feature point and image center in the image coordinate. Red line and blue line represent the change of θ_p and θ_t, respectively.

B. Field Experiments

1) Field Experimental Setup: The results of the HIL simulation verify the effectiveness of our controller in virtual environment. However, there still exist disturbances (e.g. wind disturbance and inaccurate airspeed measurement) during the real flight tests. Therefore, we set up the field experiment to

Fig. 10: The fixed-wing UAV and the car (as the ground target) we use in the field experiments.

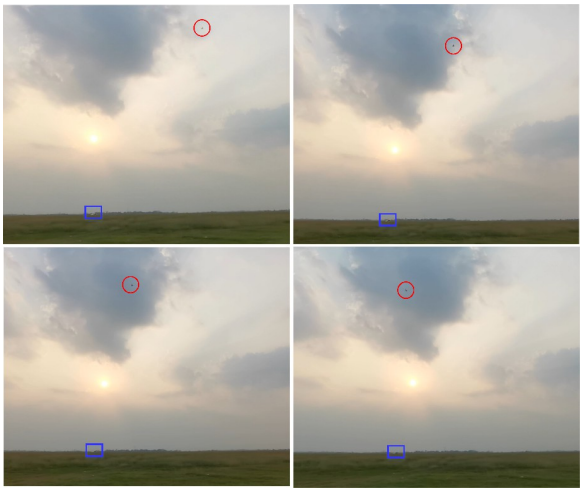


Fig. 11: The scenes of the real flight tests. The red circle represents the flying fixed-wing UAV, and the blue box shows the target, which is a white car.

further evaluate our controller, and it is developed based on the existing architecture of our team [30], [31].

The fixed-wing UAV and the car we use in the field experiments are shown in Fig. 10. The UAV has a wingspan of 1.8m and a weight of 5kg. Besides, the length and height of the UAV are 1.23m and 0.35m, respectively. For the pan-tilt camera, it is the same as we use in the HIL simulation platform.

Fig. 11 displays the scenes of our field experiments. As the figure shows, our experimental site is a meadow, and the target we use is a white car (the blue box). The four pictures represent a sequence of scenes from the same experiment.

2) Experimental Results and Analysis: We make the fixed-wing UAV track target with uniform linear motion in the experiment. The flight airspeed and height are 16m/s and 100m, respectively. Besides, the target is moving at a constant speed of 20km/h.

The tracking trajectories are shown as Fig. 12a, and the horizontal distance between the UAV and the target is shown as Fig. 12b. We can find that the horizontal distance keeps at about 95m when the target is stationary, and it will vary within the range of [23, 210] (denoted by two black dashed lines).

Fig. 12c shows the image stabilizing precision for the experiment. The green part indicates that even the target is stationary, the angles of the pan-tilt cannot keep constant for the unstable state of the fixed-wing UAV in field experiments. Besides, the points inside blue circles deviate significantly, which is caused by misdetection. And there are also some points missing inside blue rectangles. The reasons accounting for this can be divided into two kinds, one is that the target is too small in the image edge to be detected, the other is the angle of the pan-tilt reaches threshold, which makes the target out of sight briefly. Under the circumstance, we use the u_ψ calculated from (u_2, v_2) at the last moment before the target is lost for control. After that, the UAV will yaw to the direction where the target disappears from the image. As a result, the UAV can still track the target despite these situations, which further verifies the feasibility of our controller in real flight tests.

V. CONCLUSION

This paper has presented a control framework for a fixed-wing UAV with a monocular pan-tilt camera to track a moving target. More specifically, this control framework uses the target detection algorithm based on YOLOv3 to obtain the centroid coordinates of the target firstly. Then, this framework uses a reference state called *"Ideal State"* to make the coordinates of the feature point only affected by the change of the yaw angle of the UAV. After that, the controller is designed based on the image Jacobian matrix. In order to verify the feasibility of the controller, we have conducted the HIL simulation experiments based on Gazebo, and then have carried out field experiments based on a prototype system. The results show that our controller can achieve continuous and robust tracking of the target by the UAV.

Future work will focus on target tracking by multiple fixed-wing UAVs. We believe that by using the images of diverse perspectives and surrounded trajectories from different UAVs cooperatively, the challenge of tracking the high speed maneuvering ground target can be met.

REFERENCES

[1] Wu J, Wang H, Li N, et al. Distributed trajectory optimization for multiple solar-powered UAVs target tracking in urban environment by Adaptive Grasshopper Optimization Algorithm[J]. Aerospace Science and Technology, 2017, 70: 497-510.

[2] Zhang M, Liu H H T. Cooperative tracking a moving target using multiple fixed-wing UAVs[J]. Journal of Intelligent & Robotic Systems, 2016, 81(3-4): 505-529.

[3] Vanegas F, Campbell D, Roy N, et al. UAV tracking and following a ground target under motion and localisation uncertainty[C]//2017 IEEE Aerospace Conference. IEEE, 2017: 1-10.

[4] Vanegas F, Gonzalez F. Uncertainty based online planning for UAV target finding in cluttered and GPS-denied environments[C]//2016 IEEE Aerospace Conference. IEEE, 2016: 1-9.

[5] Larouche B P, Zhu Z H. Autonomous robotic capture of non-cooperative target using visual servoing and motion predictive control[J]. Autonomous Robots, 2014, 37(2): 157-167.

[6] Pestana J, Sanchez-Lopez J L, Campoy P, et al. Vision based gps-denied object tracking and following for unmanned aerial vehicles[C]//2013 IEEE International Symposium on Safety, Security, and rescue Robotics (SSRR). IEEE, 2013: 1-6.

[7] Olivares Mendez M A, Mondragon I, Campoy P, et al. Aerial object following using visual fuzzy servoing[C]//First Workshop on Research, Development and Education on Unmanned Aerial Systems (RED-UAS 2011). 2011.

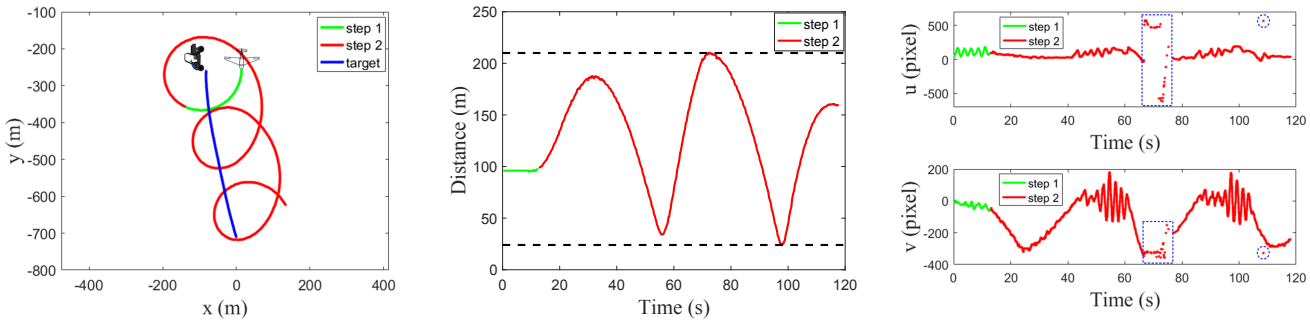


Fig. 12: The fixed-wing UAV tracks the target of uniform linear motion in the field experiment. The speed of the target is 20km/h, the flight height and speed of the UAV are 100m and 16m/s, respectively. Besides, the target is stationary in step 1 and moving in step 2. (a) Tracking trajectories; (b) Horizontal distance between the target and the fixed-wing UAV; (c) Image stabilizing precision.

[8] Hutchinson S, Hager G D, Corke P I. A tutorial on visual servo control[J]. IEEE Transactions on Robotics and Automation, 1996, 12(5): 651-670.

[9] Chaumette F, Hutchinson S. Visual servo control. I. Basic approaches[J]. IEEE Robotics & Automation Magazine, 2006, 13(4): 82-90.

[10] Chaumette F, Hutchinson S. Visual servo control. II. Advanced approaches [Tutorial][J]. IEEE Robotics & Automation Magazine, 2007, 14(1): 109-118.

[11] Teulire C, Eck L, Marchand E, et al. 3D model-based tracking for UAV position control[C]//2010 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE, 2010: 1084-1089.

[12] Wise R, Rysdyk R. UAV coordination for autonomous target tracking[C]//AIAA Guidance, Navigation, and Control Conference and Exhibit. 2006: 6453.

[13] Chen J, Dawson D M. UAV tracking with a monocular camera[C]//Proceedings of the 45th IEEE Conference on Decision and Control. IEEE, 2006: 3873-3878.

[14] Yongchun F. A survey of robot visual servoing [J]. CAAI Transactions on Intelligent Systems, 2008, 3(2): 109-114.

[15] Serra P , Cunha R , Hamel T , et al. Landing of a Quadrotor on a Moving Target Using Dynamic Image-Based Visual Servo Control[J]. IEEE Transactions on Robotics, 2017, 32(6):1524-1535.

[16] Araar O , Aouf N , Vitanov I . Vision Based Autonomous Landing of Multirotor UAV on Moving Platform[J]. Journal of Intelligent & Robotic Systems, 2017, 85(2):369-384.

[17] Srivastava R, Lima R, Das K, et al. Least Square Policy Iteration for IBVS based Dynamic Target Tracking[C]//2019 International Conference on Unmanned Aircraft Systems (ICUAS). IEEE, 2019: 1089-1098.

[18] Lyu Y, Pan Q, Zhang Y, et al. Simultaneously multi-UAV mapping and control with visual servoing[C]//2015 International Conference on Unmanned Aircraft Systems (ICUAS). IEEE, 2015: 125-131.

[19] Popova M G. Visual Servoing for a Quadrotor UAV in Target Tracking Applications[M]. University of Toronto, 2015.

[20] Zarudzki M, Shin H S, Lee C H. An image based visual servoing approach for multi-target tracking using an quad-tilt rotor UAV[C]//2017 International Conference on Unmanned Aircraft Systems (ICUAS). IEEE, 2017: 781-790.

[21] Hinas A, Roberts J M, Gonzalez F. Vision-based target finding and inspection of a ground target using a multirotor UAV system[J]. Sensors, 2017, 17(12): 2929.

[22] Wei Y, Lin Z. Vision-based Tracking by a Quadrotor on ROS[J]. IFAC-PapersOnLine, 2017, 50(1): 11447-11452.

[23] Le Bras F, Hamel T, Mahony R. Image-based visual servo control for circular trajectories for a fixed-wing aircraft[C]//Proceedings of the 48h IEEE Conference on Decision and Control (CDC) held jointly with 2009 28th Chinese Control Conference. IEEE, 2009: 3430-3435.

[24] Peliti P, Rosa L, Oriolo G, et al. Vision-based loitering over a target for a fixed-wing UAV[J]. IFAC Proceedings Volumes, 2012, 45(22): 51-57.

[25] Wang X, Zhu H, Zhang D, et al. Vision-based detection and tracking of a mobile ground target using a fixed-wing UAV[J]. International Journal of Advanced Robotic Systems, 2014, 11(9): 156.

[26] Yang L, Liu Z, Wang X, et al. An optimized image-based visual servo control for fixed-wing unmanned aerial vehicle target tracking with fixed camera[J]. IEEE Access, 2019, 7: 68455-68468.

[27] Zhao S, Wang X, Lin Z, et al. Integrating Vector Field Approach and Input-to-State Stability Curved Path Following for Unmanned Aerial Vehicles[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2018:1-8.

[28] Zhiqiang W, Jun L. A review of object detection based on convolutional neural network[C]//2017 36th Chinese Control Conference (CCC). IEEE, 2017: 11104-11109.

[29] Redmon J, Farhadi A. Yolov3: An incremental improvement[J]. arXiv preprint arXiv:1804.02767, 2018.

[30] Wang X, Shen L, Liu Z, et al. Coordinated flight control of miniature fixed-wing UAV swarms: methods and experiments[J]. Science China Information Sciences, 2019, 62(11): 212204.

[31] Liu Z, Wang X, Shen L, et al. Mission Oriented Miniature Fixed-wing UAV Swarms: A Multi-layered and Distributed Architecture[J]. arXiv preprint arXiv:1912.06285, 2019.