Research
Unmanned Intelligent Cluster—Article

Dynamic Target Tracking of Unmanned Aerial Vehicles Under Unpredictable Disturbances

Yanjie Chen [a,b,e], Yangning Wu [a], Limin Lan [a], Hang Zhong [c,e], Zhiqiang Miao [d,e], Hui Zhang [c,e], Yaonan Wang [d,e,∗]

[a] School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou 350108, China
[b] Department of Computer Science, Aberystwyth University, Aberystwyth SY23 3DB, UK
[c] School of Robotics, Hunan University, Changsha 410082, China
[d] College of Electrical and Information Engineering, Hunan University, Changsha 410082, China
[e] National Engineering Research Center of Robot Visual Perception and Control Technology, Hunan University, Changsha 410082, China

ARTICLE INFO

ABSTRACT

This study proposes an image-based visual servoing (IBVS) method based on a velocity observer for an unmanned aerial vehicle (UAV) for tracking a dynamic target in Global Positioning System (GPS)-denied environments. The proposed method derives the simplified and decoupled image dynamics of underactuated UAVs using a constructed virtual camera and then considers the uncertainties caused by the unpredictable rotations and velocities of the dynamic target. A novel image depth model that extends the IBVS method to track a rotating target with arbitrary orientations is proposed. The depth model ensures image feature accuracy and image trajectory smoothness in rotating target tracking. The relative velocities of the UAV and the dynamic target are estimated using the proposed velocity observer. Thanks to the velocity observer, translational velocity measurements are not required, and the control chatter caused by noise-containing measurements is mitigated. An integral-based filter is proposed to compensate for unpredictable environmental disturbances in order to improve the anti-disturbance ability. The stability of the velocity observer and IBVS controller is analyzed using the Lyapunov method. Comparative simulations and multistage experiments are conducted to illustrate the tracking stability, anti-disturbance ability, and tracking robustness of the proposed method with a dynamic rotating target.

1. Introduction

Unmanned aerial vehicles (UAVs) are widely used in aerial photography and patrol inspection due to their flexible flying performance and high maneuverability [1]. Autonomous control of a UAV generally requires feedback on position, velocity, and attitude information. Global Positioning System (GPS) and vision technologies are widely adopted as solutions for UAV position and velocity acquisition [2]. In particular, GPS relies on a sufficient number of available satellites to provide global positional information, which may become challenging in harsh environments due to discontinuous signal feedback from satellites. Vision technology, which provides abundant pixel information and image features, is more suitable for position information acquisition in harsh environ-

ments than GPS. Therefore, vision-based methods have been developed to allow UAVs to implement autonomous control in GPS-denied environments [3].

Vision-based control has been widely adopted for UAVs in various fields, such as positioning [4], target tracking [5], and flight landing [6]. In particular, visual servoing has received significant attention [7–9]. Visual servoing methods focus on the UAV control system, which controls the UAV to maintain a precise relative position with respect to the target object, based on image information. In contrast to simultaneous localization and mapping navigation technologies, visual servoing relies on low-cost sensors to achieve UAV positioning. For UAV target tracking in GPS-denied environments, visual servoing methods provide an efficient and low-cost solution for UAV control tasks. Visual servoing methods are generally divided into position-based visual servoing (PBVS) and image-based visual servoing (IBVS) [10,11]. In PBVS, the UAV position estimated by image information is taken as the control

---

feedback. The estimated position relies on precise camera parameters. In IBVS, feature pixels are taken as the control feedback, and the kinematics is established directly from image pixels. IBVS does not require pose reconstruction and thus yields a controller with low computational consumption. In addition, IBVS is robust against image noise and camera parameters. Therefore, IBVS methods have become popular and are widely used for the visual servoing control of UAVs. An efficient visual servoing method holds engineering significance for controlling low-cost UAVs in GPS-denied environments.

In IBVS methods for UAV, the translational velocity is essential information for assisting the controller design [12]. The accuracy of translational velocity measurements can affect the performance of the IBVS controller. The chatter phenomenon may occur in the attitude motion as noise-containing velocity measurements are introduced. Hence, velocity observers have been proposed for UAVs to avoid the need for translational velocity measurements. Model- and image-based velocity observers have been adopted in recent studies. For model-based velocity observers, such as those used in Refs. [13,14], the position information of the UAV must be introduced as observer feedback. The use of positional information limits the accuracy of translational velocity estimation in GPS-denied environments. However, image-based velocity observers utilize image information as observer feedback, making it possible to achieve robust and accurate velocity estimation in GPS-denied environments. Therefore, image-based velocity observers are widely applied in IBVS control designs. In particular, a virtual camera-based velocity observer is an effective method for estimating the translational velocities of UAVs. For example, in Ref. [15], an adaptive observer is proposed to estimate translational velocity. This method shows a reliable performance, since the observer does not introduce noisy measurements. In Ref. [16], the image depth and translational velocities are simultaneously estimated by means of a nonlinear observer using image feedback from an onboard camera.

In recent research, IBVS methods have been extended to complex target-tracking tasks. For an unpredictably moving target, it is challenging to obtain position trajectories using onboard sensors [17,18]. Hence, image-based velocity observers have been generalized to estimate tracked target translational velocities. For example, the IBVS method combined with an adaptive sliding-mode control method was designed for the moving platform tracking of quadrotors [19]. In Ref. [20], an unknown target motion is considered in the visual servoing, and an observer was utilized for the target velocity estimation. A tracking differentiator is implemented for the trajectory parameters estimation of an unpredictable dynamic target in Ref. [21]. The uncertainties of the unpredictable dynamic target are considered in the abovementioned methods, yielding stable target tracking of the UAV. However, these methods depend on the assumption that the UAV or tracked target velocities are measurable. For a UAV that has onboard sensors with limited accuracy, inaccurate measurements induce noise in the control loop, resulting in attitude chatter.

For target-tracking tasks, it is necessary to address the inaccuracy of image depth estimation caused by the attitude rotation of the tracked target. Generally, in monocular virtual camera-based IBVS methods, image depth is estimated by constructing a virtual camera plane parallel to the target plane [22]. The estimation accuracy of the image depth may be affected by the rotation of the dynamic target. For this case, an improved virtual camera-based IBVS method has been proposed to enable a UAV to track a target with an arbitrary orientation. For example, an inclined target plane is included in a virtual camera-based IBVS controller in Ref. [23]. In Ref. [24], a transformation between real and virtual image planes is proposed based on the kinematic rotation relationship, which

extends the virtual camera-based IBVS scheme to enable an under-actuated UAV to track a target with an arbitrary orientation. The aforementioned methods extend the virtual camera method and are applicable to situations where the tracked target is arbitrarily inclined. However, in the aforementioned methods, the dynamic target is assumed to be static or quasi-static. A parallel constraint between the virtual camera plane and the target plane may not be guaranteed for rotating targets with arbitrary orientations. Rotating targets may lead to inaccurate image depth estimation and affect the control accuracy. Consequently, the image depth estimation of rotating targets is a practical and challenging field for monocular IBVS control.

Since UAVs usually operate in environments with complex disturbances, their anti-disturbance ability is critical for the stability of target tracking [25–27]. Anti-disturbance control methods permit the IBVS controller for a UAV to operate in environments with uncertain disturbances. For example, in Ref. [24], an integral filter is proposed for an IBVS controller to compensate for unpredictable disturbances. In Ref. [28], an adaptive IBVS control scheme is proposed to compensate for the constant force disturbance in translational dynamics. These methods realize the stable IBVS control of UAVs under unpredictable disturbances. Nevertheless, they assume that the translational velocities of the UAV or tracked target are measurable. Translational velocity measurements are introduced into the proposed integral filter in Ref. [24], resulting in a control that is limited in GPS-denied environments. In Ref. [28], the tracked target is assumed to be static, resulting in a limitation in the target-tracking control. Adopting translational velocity measurements may lead to control limitations of the IBVS method in GPS-denied environments. Noise-containing velocity measurements may result in chatter in attitude control. Therefore, the visual servoing target-tracking control of a UAV under external disturbances is a critical problem that must be addressed.

Based on the above analysis, the challenges of IBVS control for UAVs primarily include: ① the estimation of unpredictable tracked target translational velocities, ② the image depth estimation of a dynamic rotating target, and ③ tracking stability under uncertain external disturbances. To address these challenges, we further improve the IBVS method for UAVs. To deal with an unpredictable target velocity, the uncertainty caused by a rotating target, and external disturbances in dynamic target-tracking control, we propose corresponding improvement methods. The primary contributions of this study are summarized below.

- A novel image depth model is proposed for the construction of image dynamics. The image depth can be accurately estimated using the proposed image depth model without the rotation information of the tracked target. Based on the image depth model, the proposed IBVS controller can be applied to a UAV to track a dynamic rotating target with an arbitrary orientation.

- A velocity observer is designed for the relative velocity estimation of the UAV and the dynamic target, which enables the proposed method to be utilized in GPS-denied environments. The robustness of the proposed IBVS method is enhanced, as the control chatter caused by noise-containing velocity measurements is mitigated.

- An integral-based filter is designed to estimate unpredictable disturbances, including the acceleration of the dynamic target and environmental disturbances. This enhances the disturbance rejection ability of the UAV in handling unknown movements of the dynamic target and external disturbances.

The remainder of this paper is organized as follows. The image dynamics of the UAV and a novel image depth model are introduced in Section 2. In Section 3, we design a velocity observer, integral-based filter, and dynamic IBVS controller to realize target tracking

under unpredictable disturbances. Sections 4 and 5 present the comparative simulation and experimental results, respectively. Finally, concluding remarks are presented in Section 6.

2. Image dynamics of the UAV system

This section introduces a virtual camera image to derive image dynamics. More specifically, the pixels of the dynamic target in the virtual camera plane are selected as the image features that decouple the roll/pitch motions and the image motions. Subsequently, an image depth model suitable for a dynamic target with an arbitrary orientation is derived based on the geometric properties of the dynamic target. Finally, image dynamics with a simple and decoupled form are derived based on image moments constructed using multiple feature points.

2.1. UAV dynamics

A UAV system with an inertial measurement unit (IMU) and monocular camera is shown in Fig. 1. The position and attitude dynamic models of the UAV are provided in Ref. [29].

$$\dot{\boldsymbol{v}} = \frac{f}{M}\boldsymbol{R}_b\boldsymbol{n}_3 - g\boldsymbol{n}_3 + \boldsymbol{d}_1 \tag{1}$$

$$\boldsymbol{J}\dot{\boldsymbol{\omega}}_b = -\boldsymbol{\omega}_b \times \boldsymbol{J}\boldsymbol{\omega}_b + \boldsymbol{\tau} \tag{2}$$

where $\boldsymbol{v} \in \mathbb{R}^3$ is the UAV translational velocity with respect to the inertial frame O_a and $\dot{\boldsymbol{v}}$ denotes the time derivative of $\boldsymbol{v}$. $\mathbb{R}^3$ represents the set of three-dimensional real vector. $M \in \mathbb{R}$ and $\boldsymbol{J} \in \mathbb{R}^{3 \times 3}$ denote the UAV mass and inertia matrix, respectively. $\mathbb{R}$ represents the set of real number. $\mathbb{R}^{3 \times 3}$ represents the set of three-dimensional real matrix. Both M and $\boldsymbol{J}$ are constants. g denotes the gravity acceleration, f denotes the UAV thrust, and $\boldsymbol{n}_3 = [0, 0, 1]^T \in \mathbb{R}^3$ is the unit vector along the z-direction. $\boldsymbol{R}_b \in \mathbb{R}^{3 \times 3}$ is the rotation matrix of the body-fixed frame O_b with respect to O_a, and $\boldsymbol{\omega}_b \in \mathbb{R}^3$ and $\boldsymbol{\tau}$ are the UAV angular velocity and torque, respectively. $\dot{\boldsymbol{\omega}}_b$ denotes the time derivative of $\boldsymbol{\omega}_b$. $\boldsymbol{d}_1 \in \mathbb{R}^3$ represents the environmental disturbance, and f and $\boldsymbol{\tau}$ are provided by the rotors of the UAV.

By denoting the visual servoing control input $\boldsymbol{u} = \frac{f}{M}\boldsymbol{R}_b\boldsymbol{n}_3 - g\boldsymbol{n}_3$, Eq. (1) can be expressed as follows:

$$\dot{\boldsymbol{v}} = \boldsymbol{u} + \boldsymbol{d}_1 \tag{3}$$

Since the environmental disturbance $\boldsymbol{d}_1$ is typically produced by wind, the following assumption can be made.
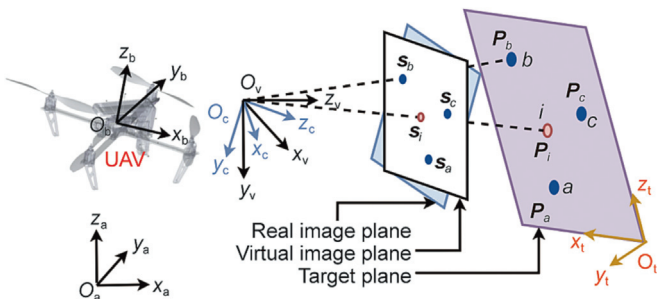

Fig. 1. Reference coordinate frames of the UAV system. O_a, O_b: inertial and body-fixed frames, respectively; O_c, O_v: real camera frame and virtual camera frame, respectively; O_t: tracked target frame; a, b, and c: noncollinear feature points in the target plane; i: the feature points inside the triangle enclosed by a, b, and c; P_a, P_b, P_c, and P_i: positions of a, b, c, and i with respect to O_t, respectively; s_a, s_b, s_c, and s_i: virtual image pixels of a, b, c, and i, respectively; x_a, y_a, z_a: x-axis, y-axis, and z-axis of O_a, respectively; x_b, y_b, z_b: x-axis, y-axis, and z-axis of O_b, respectively; x_v, y_v, z_v: x-axis, y-axis, and z-axis of O_v, respectively; x_c, y_c, z_c: x-axis, y-axis, and z-axis of O_c, respectively; x_t, y_t, z_t: x-axis, y-axis, and z-axis of O_t, respectively.

Assumption 1. The environmental disturbances $\boldsymbol{d}_1$ are bounded: $\|\boldsymbol{d}_1\| \leq D_1$, where $\|\boldsymbol{y}\|$ represents the L^2-norm of vector $\boldsymbol{y}$ and D_1 is the upper bound of the disturbances $\boldsymbol{d}_1$.

The attitude controller of the UAV is assumed to be available, based on existing commercial UAVs [30]. The desired $\boldsymbol{R}_b$ and f are adopted as inputs for the IBVS control. Denoting $\boldsymbol{R}_b = [\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3]$, where $\boldsymbol{r}_1$, $\boldsymbol{r}_2$, and $\boldsymbol{r}_3$ are the first, second, and third column elements of the rotation matrix $\boldsymbol{R}_b$, respectively. The designed visual servoing control input $\boldsymbol{u}$ can be transformed into the control input of a commercial UAV using the following equations [30]:

$$f = M\|\boldsymbol{u} + g\boldsymbol{n}_3\| \tag{4}$$

$$\begin{cases} \boldsymbol{r}_1 = \boldsymbol{r}_2 \times \boldsymbol{r}_3 \\ \boldsymbol{r}_2 = \frac{\boldsymbol{r}_3 \times \boldsymbol{l}}{\|\boldsymbol{r}_3 \times \boldsymbol{l}\|} \\ \boldsymbol{r}_3 = \frac{\boldsymbol{u} + g\boldsymbol{n}_3}{f} \end{cases} \tag{5}$$

where $\boldsymbol{l} = [\cos\phi, \sin\phi, 0]^T$ is the heading vector of the UAV and ϕ denotes the UAV yaw angle.

2.2. Virtual camera

The pixel motions and translational velocities are coupled with the roll and pitch motions for the underactuated UAV. The image dynamics are related to the roll and pitch motions, presenting a challenge in the IBVS controller design. Virtual camera methods have been proposed to decouple the image and roll/pitch motions [8,15]. Inspired by Ref. [15], the attitudes of the virtual camera are constructed in such a way that they remain fixed with the inertial frame, thereby permitting independence of the image motion and UAV rotation.

As shown in Fig. 1, O_c and O_v denote the camera frame and virtual camera frame, respectively. The tracked target frame is defined as O_t. The tracked target is a rotating plane with arbitrary orientation. The rotation of O_v remains fixed with O_a, and the rotation matrix of O_v with respect to O_a is defined as $\boldsymbol{R}_v$. The origin of O_v is set to coincide with that of O_c. The rotation matrix from O_v to O_c is defined as $\boldsymbol{R}_c^v$ and can be derived using the following equation:

$$\boldsymbol{R}_c^v = \boldsymbol{R}_v^T \boldsymbol{R}_b \boldsymbol{R}_c^b \tag{6}$$

where $\boldsymbol{R}_c^b \in SO(3)$ denotes the rotation matrix of O_c with respect to O_b. $\boldsymbol{R}_c^b$ is a constant matrix, as the camera is fixed to the UAV.

We denote $\boldsymbol{s}_k = [w_k, h_k, \lambda]^T \in \mathbb{R}^3$ and $\bar{\boldsymbol{s}}_k = \left[\bar{w}_k, \bar{h}_k, \lambda\right]^T$ as the homogeneous pixel coordinates of the kth point in the virtual camera plane and the real image plane, respectively. w_k and $\bar{w}_k$ are the pixel width coordinates of virtual camera and real camera, respectively. h_k and $\bar{h}_k$ are the pixel height coordinates of virtual camera and real camera, respectively. λ is the focal length of the camera. k is the serial number of feature points in the target plane. Adopting the perspective projection model of the camera, the following equations can be derived:

$$\boldsymbol{s}_k = \frac{\lambda}{z_k}\boldsymbol{p}_k \tag{7}$$

$$\bar{\boldsymbol{s}}_k = \frac{\lambda}{\bar{z}_k}\bar{\boldsymbol{p}}_k \tag{8}$$

where $\boldsymbol{p}_k = [x_k, y_k, z_k]^T$ and $\bar{\boldsymbol{p}}_k = \left[\bar{x}_k, \bar{y}_k, \bar{z}_k\right]^T$ denote the kth point positions expressed in O_v and O_c, respectively. z_k is image depth of the kth point. x_k, y_k, and z_k are the three components of $\boldsymbol{p}_k$, respectively. $\bar{x}_k$, $\bar{y}_k$, and $\bar{z}_k$ are the three components of $\bar{\boldsymbol{p}}_k$.

The relationship between $\boldsymbol{p}_k$ and $\bar{\boldsymbol{p}}_k$ can be expressed as follows:

$$\boldsymbol{p}_k = \boldsymbol{R}_c^v \bar{\boldsymbol{p}}_k \tag{9}$$

Subsequently, by combining Eqs. (7)–(9), s_k can be expressed as follows:

$$s_k = \frac{\overline{z}_k}{z_k} R_c^v \overline{s}_k \tag{10}$$

where $\frac{\overline{z}_k}{z_k} = \frac{\lambda}{n_3^T R_c^v \overline{s}_k}$ is calculated using the third row of the matrix form of Eq. (10). By applying Eq. (10), the feature pixel coordinates of the virtual camera can be derived from real images.

Based on the defined virtual camera, the image kinematics describing the relationship between the translational velocities and virtual camera image pixel motions can be derived as follows [10]:

$$\begin{bmatrix} \dot{w}_k \\ \dot{h}_k \end{bmatrix} = \frac{1}{z_k} \begin{bmatrix} -\lambda & 0 & w_k \\ 0 & -\lambda & h_k \end{bmatrix} (v_v - v_e) \tag{11}$$

where $v_v = [v_{v1}, v_{v2}, v_{v3}]^T$ denotes the UAV translational velocity with respect to O_v, and $v_e = [v_{e1}, v_{e2}, v_{e3}]^T$ is the translational velocity of the dynamic target. $\dot{w}_k$ and $\dot{h}_k$ represent the time derivative of w_k and h_k, respectively.

Since the rotation matrix R_v is defined as a constant, the following equation can be obtained:

$$\begin{cases} \dot{z}_k = -(v_{v3} - v_{e3}) \\ v_v = R_v^T v \\ \dot{v}_v = R_v^T \dot{v} \end{cases} \tag{12}$$

where $\dot{z}_k$ and $\dot{v}_v$ represent the time derivative of z_k and v_v, respectively.

Based on the defined virtual camera, the pixel motion is related only to the translational velocities of the UAV. The image kinematics with a simple form is established as Eq. (11), which can simplify the derivation of the image dynamics and IBVS controller design.

2.3. Depth information acquisition

Since the image depth z_k in Eq. (11) cannot be directly measured using a monocular camera, the estimation of z_k needs to be addressed. The virtual camera image moment is widely utilized as an effective method for image depth estimation. In the virtual camera method, a virtual plane is constructed parallel to the target plane. Based on this parallel constraint, a reliable depth estimation can be derived from the proportional relationship between the image moment and image depth. For a rotating target with unpredictable attitudes, it is challenging to maintain the parallel constraint of the virtual plane, resulting in chatter in the image depth estimation. To overcome this limitation, an image depth model based on the geometric properties of the rotating target is proposed in order to estimate the image depth. The depth model is independent of the parallel constraints of the virtual image plane. The image depth of a rotating target with an arbitrary orientation can be accurately obtained.

As shown in Fig. 1, a, b, and c are three noncollinear feature points in the target plane. $i = \{1, 2, 3, \ldots\}$ denote the feature points inside the triangle enclosed by a, b, and c. a, b, c, and i are assumed to be identifiable. P_k is defined as the positions of multiple points expressed in O_t, where $k=\{a, b, c, i\}$ denotes the serial number of feature points in the target plane. P_a, P_b, P_c, and P_i are the positions of a, b, c, and i with respect to O_t, respectively. The virtual image pixels of a, b, c, and i are s_a, s_b, s_c, and s_i, respectively. Then, P_i can be expressed as the following equation by using barycentric interpolation.

$$P_i = \alpha_i P_a + \beta_i P_b + \gamma_i P_c \tag{13}$$

where $\alpha_i, \beta_i, \gamma_i \geq 0$ are the interpolation coefficients and satisfy the equation $\alpha_i + \beta_i + \gamma_i = 1$. The values of α_i, β_i, and γ_i are constant and can be uniquely determined by the geometric information of a dynamic target.

The relationship between P_k and p_k can be expressed as the following equation:

$$p_k = R_t^v P_k + p_t^v, k = \{a, b, c, i\} \tag{14}$$

where R_t^v and p_t^v are the rotation matrix and the translation from O_v to O_t, respectively.

Combining Eqs. (7), (13), and (14) yields the following:

$$z_i s_i = \alpha_i z_a s_a + \beta_i z_b s_b + \gamma_i z_c s_c \tag{15}$$

where z_a, z_b, and z_c are the image depth of a, b, and c, respectively. z_i is the image depth of feature point i.

By taking the vector operation on both sides of Eq. (15), the equation can be rewritten as follows:

$$z_i s_a \times s_i = \beta_i z_b s_a \times s_b + \gamma_i z_c s_a \times s_c \tag{16}$$

$$z_i s_b^T (s_a \times s_i) = \gamma_i z_c s_b^T (s_a \times s_c) \tag{17}$$

Then, z_c can be expressed as follows:

$$z_c = \frac{1}{\gamma_i} \frac{s_b^T (s_a \times s_i)}{s_b^T (s_a \times s_c)} z_i = \frac{1}{\gamma_i} \frac{s_i^T (s_a \times s_b)}{s_c^T (s_a \times s_b)} z_i = r_c z_i \tag{18}$$

where $r_c = \frac{1}{\gamma_i} \frac{s_i^T (s_a \times s_b)}{s_c^T (s_a \times s_b)}$ is the image depth parameter of c.

Similarly, z_a and z_b can be expressed as follows:

$$z_a = \frac{1}{\alpha_i} \frac{s_i^T (s_b \times s_c)}{s_a^T (s_b \times s_c)} z_i = r_a z_i \tag{19}$$

$$z_b = \frac{1}{\beta_i} \frac{s_i^T (s_a \times s_c)}{s_b^T (s_a \times s_c)} z_i = r_b z_i \tag{20}$$

where $r_a = \frac{1}{\alpha_i} \frac{s_i^T (s_b \times s_c)}{s_a^T (s_b \times s_c)}$ and $r_b = \frac{1}{\beta_i} \frac{s_i^T (s_a \times s_c)}{s_b^T (s_a \times s_c)}$ are the image depth parameters of a and b, respectively.

Assuming that the geometric features of the dynamic target are invariant, the area enclosed by P_a, P_b, and P_c is constant and can be expressed as follows:

$$S_{a,b,c} = \frac{1}{2} \|(p_a - p_b) \times (p_a - p_c)\| \tag{21}$$

where $S_{a,b,c}$ is the area enclosed by feature points a, b, and c. p_a, p_b, and p_c are the positions of feature points a, b, and c with respect to O_c, respectively.

Applying Eqs. (7), (18), (19), and (21), Eq. (21) can be rewritten as follows:

$$S_{a,b,c} = \frac{z_i}{2\lambda} \|(r_a s_a - r_b s_b) \times (r_a s_a - r_c s_c)\| \tag{22}$$

Then, the image depth of P_i can be expressed as follows:

$$z_i = 2\lambda S_{a,b,c} \rho_i \tag{23}$$

where $\rho_i = \frac{1}{\|(r_a s_a - r_b s_b) \times (r_a s_a - r_c s_c)\|}$ is the image depth parameter of i.

For the points within the area enclosed by P_a, P_b, and P_c, the image depth can be estimated by means of Eq. (23) without the influence of target rotation.

2.4. Virtual camera image dynamics

Consider N visible feature pixels in the target plane. The image features are defined as follows:

$$q = \frac{1}{N} \sum_{i=1}^{N} \rho_i s_i \tag{24}$$

where q is the image moment.

The time derivative of $\boldsymbol{q}$ is derived by

$$\dot{\boldsymbol{q}} = \frac{1}{N} \sum_{i=1}^{N} \left(\dot{\rho}_i \boldsymbol{s}_i + \rho_i \begin{bmatrix} \dot{w}_i \\ \dot{h}_i \\ 0 \end{bmatrix} \right) \tag{25}$$

where $\dot{\rho}_i = \frac{\dot{z}_i}{2\lambda S_{a,b,c}} = -\frac{(v_{v3} - v_{e3})}{2\lambda S_{a,b,c}}$ is time derivative of ρ_i, which can be derived using Eq. (23). w_i and h_i are the virtual camera pixel width and height coordinates of feature points i, respectively. $\dot{w}_i$ and $\dot{h}_i$ are the time derivative of w_i and h_i, respectively.

Combining Eqs. (11), (12), (23), and (25), $\dot{\boldsymbol{q}}$ is obtained as follows:

$$\begin{aligned} \dot{\boldsymbol{q}} &= \frac{1}{N} \sum_{i=1}^{N} \left(-\frac{(v_v - v_{e3})}{2\lambda S_{a,b,c}} \begin{bmatrix} w_i \\ h_i \\ \lambda \end{bmatrix} + \frac{z_i}{2\lambda S_{a,b,c}} \begin{bmatrix} \dot{w}_i \\ \dot{h}_i \\ 0 \end{bmatrix} \right) \\ &= \frac{1}{2\lambda S_{a,b,c}} \frac{1}{N} \sum_{i=1}^{N} \left(\begin{bmatrix} -\lambda & 0 & 0 \\ 0 & -\lambda & 0 \\ 0 & 0 & -\lambda \end{bmatrix} (\boldsymbol{v}_v - \boldsymbol{v}_e) \right) \\ &= -\frac{1}{2 S_{a,b,c}} (\boldsymbol{v}_v - \boldsymbol{v}_e) \end{aligned} \tag{26}$$

Denote $m = \frac{1}{2 S_{a,b,c}} > 0$ (where m is the tracked target geometric parameter) and $\boldsymbol{\delta} = -m(\boldsymbol{v}_v - \boldsymbol{v}_e)$ (where $\boldsymbol{\delta}$ is the UAV velocity relative to the tracked target). The image dynamics can be expressed by applying Eqs. (3), (12), and (26), as follows:

$$\dot{\boldsymbol{q}} = \boldsymbol{\delta} \tag{27}$$

$$\dot{\boldsymbol{\delta}} = -m \boldsymbol{R}_v^T \boldsymbol{u} + \boldsymbol{d} \tag{28}$$

where $\boldsymbol{d} = -m\left(\boldsymbol{R}_v^T \boldsymbol{d}_1 - \dot{\boldsymbol{v}}_e \right)$ is the total uncertainty of the image dynamics system.

Since the maneuverability of a dynamic target is usually limited, the following assumption can be adopted.

Assumption 2. The accelerations of the tracked target are bounded: $\|\dot{\boldsymbol{v}}_e\| \leq D_2$, where D_2 is the upper bound of the tracked target accelerations.

Denote $D = m(D_1 + D_2)$. $\|\boldsymbol{d}\| \leq m\left(\|\boldsymbol{R}_v^T \boldsymbol{d}_1\| + \|\dot{\boldsymbol{v}}_e\| \right) \leq D$ is the total uncertainty upper bound and can be obtained by Assumptions 1 and 2, as $\boldsymbol{R}_v$ and m are constants.

Thanks to the constructed virtual camera and image depth model, the image dynamics can be simplified to a decoupled form, as shown in Eqs. (27) and (28).

3. Dynamic visual servoing control method

This section derives and analyzes the dynamic visual servoing controller design based on the Lyapunov theory. More specifically, a velocity observer is proposed to estimate the relative velocity between the UAV and the dynamic target. The bounds of the unpredictable disturbances, including the external disturbances, and the dynamic target are considered in the design of the observer and are estimated using an integral-based filter. Subsequently, the velocity observer outputs are adopted in the visual servoing controller. The control flow of the proposed IBVS method is illustrated in Fig. 2.

3.1. Adaptive velocity observer

The estimated values of $\boldsymbol{q}$ and $\boldsymbol{\delta}$ are defined as $\hat{\boldsymbol{q}}$ and $\hat{\boldsymbol{\delta}}$, respectively. $\tilde{\boldsymbol{q}} = \boldsymbol{q} - \hat{\boldsymbol{q}}$ and $\tilde{\boldsymbol{\delta}} = \boldsymbol{\delta} - \hat{\boldsymbol{\delta}}$ are the estimated errors of $\boldsymbol{q}$ and $\boldsymbol{\delta}$,
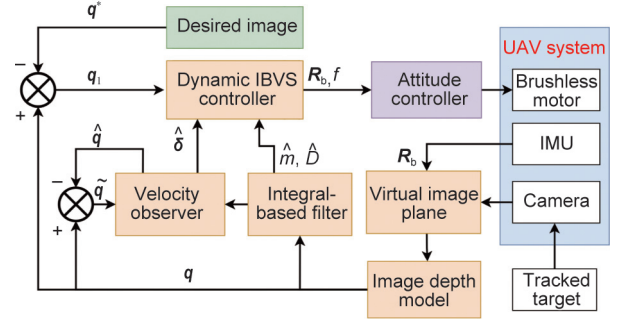

Fig. 2. Architecture of the proposed visual servoing scheme. $\boldsymbol{q}^*$: desired image moment; $\hat{\boldsymbol{q}}$: auxiliary variable for the velocity observer; $\hat{\boldsymbol{\delta}}$: estimated values of $\boldsymbol{\delta}$; $\boldsymbol{q}_1$, $\tilde{\boldsymbol{q}}$: image and estimated errors of $\boldsymbol{q}$, respectively; $\hat{m}$, $\hat{D}$: estimated values of m and D, respectively.

respectively. As $\boldsymbol{q}$ is an obtainable state, $\hat{\boldsymbol{q}}$ acts as an auxiliary variable for the velocity observer.

Theorem 1. The velocity observer is asymptotically stable and yields the observer errors $\tilde{\boldsymbol{q}}$ and $\tilde{\boldsymbol{\delta}}$ to converge to zero:

$$\dot{\hat{\boldsymbol{q}}} = \hat{\boldsymbol{\delta}} + k_1 \tilde{\boldsymbol{q}} \tag{29}$$

$$\hat{\boldsymbol{\delta}} = \boldsymbol{\delta}_o + k_2 \tilde{\boldsymbol{q}} \tag{30}$$

$$\dot{\boldsymbol{\delta}}_o = -\hat{m} \boldsymbol{R}_v^T \boldsymbol{u} + (k_1 k_2 + 1) \tilde{\boldsymbol{q}} + \hat{D} \mathrm{sgn}\left(\dot{\tilde{\boldsymbol{q}}} + k_1 \tilde{\boldsymbol{q}} \right) \tag{31}$$

where $\dot{\hat{\boldsymbol{q}}}$ is the time derivative of $\hat{\boldsymbol{q}}$. $\boldsymbol{\delta}_o$ is the auxiliary variable for the velocity observer. $\dot{\boldsymbol{\delta}}_o$ is the time derivative of $\boldsymbol{\delta}_o$. $k_1 > 0 \in \mathbb{R}$ and $k_2 > 0 \in \mathbb{R}$ are the observer gains. $\hat{D}$ and $\hat{m}$ are the estimated values of D and m, respectively. The integral-based filters of $\hat{D}$ and $\hat{m}$ are designed as follows:

$$\dot{\hat{D}} = \left(\dot{\tilde{\boldsymbol{q}}} + k_1 \tilde{\boldsymbol{q}} \right)^T \mathrm{sgn}\left(\dot{\tilde{\boldsymbol{q}}} + k_1 \tilde{\boldsymbol{q}} \right) \tag{32}$$

$$\dot{\hat{m}} = -\left(\dot{\tilde{\boldsymbol{q}}} + k_1 \tilde{\boldsymbol{q}} \right)^T \boldsymbol{R}_v^T \boldsymbol{u} \tag{33}$$

Proof. Applying Eqs. (27)–(30), $\dot{\tilde{\boldsymbol{q}}}$ and $\dot{\tilde{\boldsymbol{\delta}}}$ can be expressed as follows:

$$\dot{\tilde{\boldsymbol{q}}} = -k_1 \tilde{\boldsymbol{q}} + \tilde{\boldsymbol{\delta}} \tag{34}$$

$$\dot{\tilde{\boldsymbol{\delta}}} = -\tilde{m} \boldsymbol{R}_v^T \boldsymbol{u} - k_2 \tilde{\boldsymbol{\delta}} - \hat{D} \mathrm{sgn}\left(\tilde{\boldsymbol{\delta}} \right) - \tilde{\boldsymbol{q}} + \boldsymbol{d} \tag{35}$$

As D and m are constants, $\dot{D} = 0$ and $\dot{m} = 0$ can be derived. Denote $\tilde{D} = D - \hat{D}$ and $\tilde{m} = m - \hat{m}$. $\dot{\tilde{D}}$ and $\dot{\tilde{m}}$ can then be expressed as follows:

$$\dot{\tilde{D}} = -\tilde{\boldsymbol{\delta}}^T \mathrm{sgn}\left(\tilde{\boldsymbol{\delta}} \right) \tag{36}$$

$$\dot{\tilde{m}} = \tilde{\boldsymbol{\delta}}^T \boldsymbol{R}_v^T \boldsymbol{u} \tag{37}$$

The velocity observer Lyapunov function V_o is defined as follows:

$$V_o = \frac{1}{2} \tilde{\boldsymbol{q}}^T \tilde{\boldsymbol{q}} + \frac{1}{2} \tilde{\boldsymbol{\delta}}^T \tilde{\boldsymbol{\delta}} + \frac{1}{2} \tilde{D}^2 + \frac{1}{2} \tilde{m}^2 \tag{38}$$

The time derivative of V_o can be derived as follows:

$$\dot{V}_o = \tilde{\boldsymbol{q}}^T\dot{\tilde{\boldsymbol{q}}} + \tilde{\boldsymbol{\delta}}^T\dot{\tilde{\boldsymbol{\delta}}} - \tilde{D}\dot{D} - \tilde{m}\dot{m} - k_1\tilde{\boldsymbol{q}}^T\tilde{\boldsymbol{q}} + \tilde{\boldsymbol{q}}^T\tilde{\boldsymbol{\delta}} - \tilde{m}\tilde{\boldsymbol{\delta}}^T\boldsymbol{R}_v^T\boldsymbol{u}$$

$$= -k_2\tilde{\boldsymbol{\delta}}^T\tilde{\boldsymbol{\delta}} - \hat{D}\tilde{\boldsymbol{\delta}}^T\text{sgn}\left(\tilde{\boldsymbol{\delta}}\right) - \tilde{\boldsymbol{\delta}}^T\tilde{\boldsymbol{q}} + \tilde{\boldsymbol{\delta}}^T\boldsymbol{d} - \tilde{D}\tilde{\boldsymbol{\delta}}^T\text{sgn}\left(\tilde{\boldsymbol{\delta}}\right)$$

$$\quad + \tilde{m}\tilde{\boldsymbol{\delta}}^T\boldsymbol{R}_v^T\boldsymbol{u}$$

$$\leq -k_1\tilde{\boldsymbol{q}}^T\tilde{\boldsymbol{q}} - k_2\tilde{\boldsymbol{\delta}}^T\tilde{\boldsymbol{\delta}} \tag{39}$$

By applying Barbalat's lemma, $\tilde{\boldsymbol{q}}$ and $\tilde{\boldsymbol{\delta}}$ converge to zero asymptotically as time t approaches ∞.

By applying Eqs. (29)–(33), the estimated $\hat{\boldsymbol{\delta}}$ is adopted instead of $\boldsymbol{\delta}$ for the IBVS controller design. $\hat{D}$ is utilized to compensate for external disturbances in the IBVS controller.

3.2. Dynamic IBVS controller design

The desired image moment is defined as $\boldsymbol{q}^*$, and the image error of $\boldsymbol{q}$ are defined as $\boldsymbol{q}_1 = \boldsymbol{q} - \boldsymbol{q}^*$. For target-tracking tasks, the UAV is expected to maintain a stable position with respect to the dynamic target; $\dot{\boldsymbol{q}}$ is configured as $\dot{\boldsymbol{q}}^* = 0$, and $\dot{\boldsymbol{q}} = \dot{\boldsymbol{q}}_1$ can be derived. Subsequently, the backstepping control method is utilized for the IBVS controller design.

Consider the Lyapunov function:

$$V_1 = \frac{1}{2}\boldsymbol{q}_1^T\boldsymbol{q}_1 \tag{40}$$

where V_1 is preliminary select Lyapunov function in backstepping techniques.

The time derivative of V_1 can be derived as follows:

$$\dot{V}_1 = \boldsymbol{q}_1^T\dot{\boldsymbol{q}}_1 = \boldsymbol{q}_1^T\boldsymbol{\delta} = \boldsymbol{q}_1^T\left(\tilde{\boldsymbol{\delta}} + \hat{\boldsymbol{\delta}}\right) \tag{41}$$

The error term $\boldsymbol{q}_2$ is defined as $\boldsymbol{q}_2 = \boldsymbol{q}_1 + \frac{1}{k_3}\hat{\boldsymbol{\delta}}$, where $k_3 > 0 \in \mathbb{R}$ is the IBVS controller gain. $\dot{V}_1$ can be written as follows:

$$\dot{V}_1 = -k_3\boldsymbol{q}_1^T\boldsymbol{q}_1 + k_3\boldsymbol{q}_1^T\boldsymbol{q}_2 + \boldsymbol{q}_1^T\tilde{\boldsymbol{\delta}} \tag{42}$$

Then, the time derivative of $\boldsymbol{q}_2$ can be expressed as follows:

$$\dot{\boldsymbol{q}}_2 = \dot{\boldsymbol{q}}_1 + \frac{1}{k_3}$$

$$\quad \times \left[-\hat{m}\boldsymbol{R}_v^T\boldsymbol{u} + k_2\left(\dot{\tilde{\boldsymbol{q}}} + k_1\tilde{\boldsymbol{q}}\right) + \hat{D}\text{sgn}\left(\dot{\tilde{\boldsymbol{q}}} + k_1\tilde{\boldsymbol{q}}\right) + \tilde{\boldsymbol{q}}\right]$$

$$= \hat{\boldsymbol{\delta}} + \frac{1}{k_3}\left[-\hat{m}\boldsymbol{R}_v^T\boldsymbol{u} + \hat{D}\text{sgn}\left(\dot{\tilde{\boldsymbol{q}}} + k_1\tilde{\boldsymbol{q}}\right) + \tilde{\boldsymbol{q}}\right] + \left(\frac{k_2}{k_3} + 1\right)\tilde{\boldsymbol{\delta}} \tag{43}$$

Theorem 2. The dynamic system provided by Eqs. (27) and (28) with the visual servoing controller input $\boldsymbol{u}$ is asymptotically stable and yields the error terms $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, $\tilde{\boldsymbol{q}}$, and $\tilde{\boldsymbol{\delta}}$, which converge to zero:

$$\boldsymbol{u} = \frac{1}{\hat{m}}\boldsymbol{R}_v\left[\hat{D}\text{sgn}\left(\dot{\tilde{\boldsymbol{q}}} + k_1\tilde{\boldsymbol{q}}\right) + \tilde{\boldsymbol{q}} + k_3\hat{\boldsymbol{\delta}} + k_3k_4\boldsymbol{q}_2 + k_3^2\boldsymbol{q}_1\right] \tag{44}$$

where $k_4 > 0 \in \mathbb{R}$ is the IBVS controller gain.

Proof. Applying Eqs. (43) and (44), $\dot{\boldsymbol{q}}_2$ can be expressed as follows:

$$\dot{\boldsymbol{q}}_2 = -k_4\boldsymbol{q}_2 - k_3\boldsymbol{q}_1 + \left(\frac{k_2}{k_3} + 1\right)\tilde{\boldsymbol{\delta}} \tag{45}$$

Consider the Lyapunov function:

$$V_2 = V_1 + \frac{1}{2}\boldsymbol{q}_2^T\boldsymbol{q}_2 + V_o \tag{46}$$

where V_2 is final select Lyapunov function in backstepping techniques.

The time derivative of V_2 can be rewritten as follows:

$$\dot{V}_2 = -k_3\boldsymbol{q}_1^T\boldsymbol{q}_1 + k_3\boldsymbol{q}_1^T\boldsymbol{q}_2 + \boldsymbol{q}_1^T\tilde{\boldsymbol{\delta}} - k_4\boldsymbol{q}_2^T\boldsymbol{q}_2 + \left(\frac{k_2}{k_3} + 1\right)\boldsymbol{q}_2^T\tilde{\boldsymbol{\delta}} - k_3\boldsymbol{q}_2^T\boldsymbol{q}_1$$

$$\quad - k_1\tilde{\boldsymbol{q}}^T\tilde{\boldsymbol{q}} - k_2\tilde{\boldsymbol{\delta}}^T\boldsymbol{\delta} \leq -\left(k_3 - \frac{1}{2}\right)\boldsymbol{q}_1^T\boldsymbol{q}_1 \tag{47}$$

$$\quad - \left(k_4 - \frac{k_2 + k_3}{2k_3}\right)\boldsymbol{q}_2^T\boldsymbol{q}_2 - k_1\tilde{\boldsymbol{q}}^T\tilde{\boldsymbol{q}} - \left(k_2 - \frac{k_2}{2k_3} - 1\right)\tilde{\boldsymbol{\delta}}^T\tilde{\boldsymbol{\delta}}$$

From Eq. (47), the controller and observer gains are selected as follows:

$$\begin{cases} k_1 > 0, \quad k_2 > \frac{2k_3}{2k_3 - 1} \\ k_3 > \frac{1}{2}, \quad k_4 > \frac{k_2 + k_3}{2k_3} \end{cases} \tag{48}$$

According to Barbalat's lemma, $\boldsymbol{q}_1$, $\boldsymbol{q}_2$, $\tilde{\boldsymbol{q}}$, and $\tilde{\boldsymbol{\delta}}$ converge asymptotically to zero as time t approaches ∞.

4. Simulation of target tracking

In the simulation, the proposed IBVS controller is compared with the controller in Ref. [24]. A comparative simulation is designed using CoppeliaSim (Coppelia Robotics, ltd., Switzerland) to simulate actual flight situations. The simulation results are provided to illustrate the effectiveness of the proposed method.

4.1. Simulation setup

The UAV model is built using CoppeliaSim software to implement the comparative simulations. The proposed IBVS controller and the controller in Ref. [24] are respectively implemented in the simulation platform. The structure of the simulation platform is depicted in Fig. 3. The simulation parameters are as follow: M = 0.8 kg; control time step Δt = 0.02 s; noise of velocities $N\sim$ (0, 0.02) m·s^{-1}; $\boldsymbol{d} = N\sim$(0, 0.1); attitude of the tracked target $\text{Rot}_{zyx}\left(\left[\frac{\sin(t)}{6}, 0, \frac{\cos(t)}{12}\right]\right)$ rad; position trajectories of the tracked target $\boldsymbol{p}_t^a = [1, 0.3\cos(0.25t), 0]$ m; desired position of the UAV relative to the target $\boldsymbol{p}_t^c = [1, 0, 0]$ m, where $N\sim(\cdot)$ denotes Gaussian distribution and $\text{Rot}_{zyx}(\cdot)$ represents the rotation matrix corresponding to Euler angle. The desired pixels of $\boldsymbol{s}_a$, $\boldsymbol{s}_b$, $\boldsymbol{s}_c$, and $\boldsymbol{s}_i$ are $\boldsymbol{s}_a^* = [0, -27]$, $\boldsymbol{s}_b^* = [-24, 15]$, $\boldsymbol{s}_c^* = [24, 15]$, and $\boldsymbol{s}_i^* = [0, 0]$, respectively. The controller parameters of the proposed method are set as $k_1 = 2, k_2 = 3, k_3 = 1$, and $k_4 = 3$. The controller parameters of the comparison method in Ref. [24] are set as $k_{c1} = 1, k_{c2} = 2, k_{tg} = 1$, and $k_{b1} = k_{b2} = 0.05$. In the simulation, the geometric parameters of the dynamic target are $\alpha_i = \frac{1}{3}$, $\beta_i = \frac{1}{3}$, and $\gamma_i = \frac{1}{3}$, respectively. Translational velocity measurements are utilized to estimate the dynamic target velocity in the comparison method, while these measurements are avoided in the proposed method. Moreover, in the proposed method, the virtual camera-based IBVS method is extended to track rotating targets with arbitrary orientations. The effect of the noise-containing measurements and rotating target on the control are compared in the simulation.

The comparative simulation is divided into three stages: ① target tracking from t = 0 to 15 s, ② tracking under a disturbance from t = 15 to 30 s, and ③ rotating target tracking from t = 30 to 45 s. The execution process of the dynamic target-tracking task is illustrated in Fig. 4. In the target-tracking stage, the UAV is set to track a nonuniform moving target to verify the effectiveness of the proposed velocity observer. Subsequently, in the tracking under the disturbance stage, external disturbances
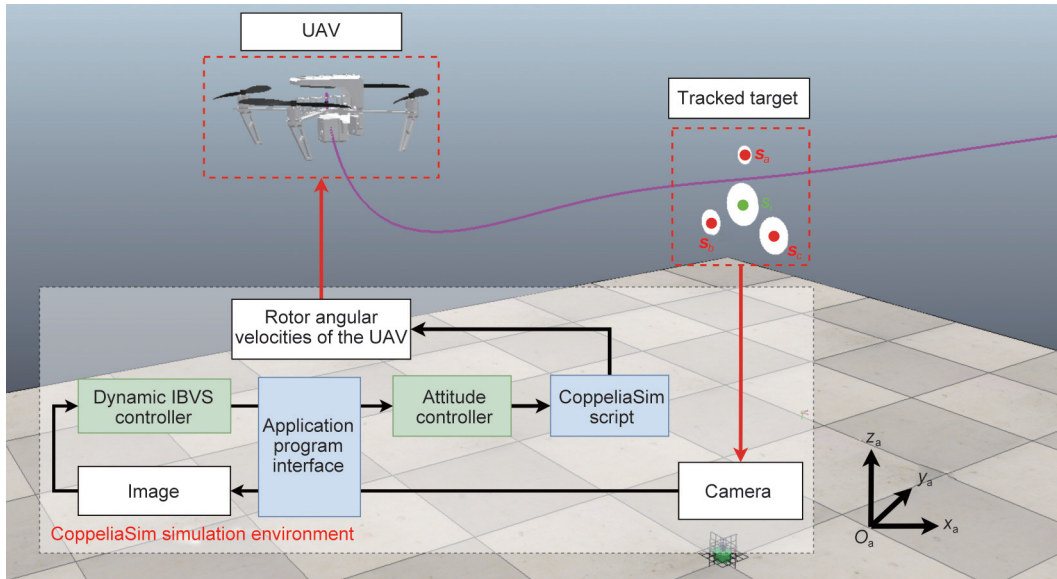
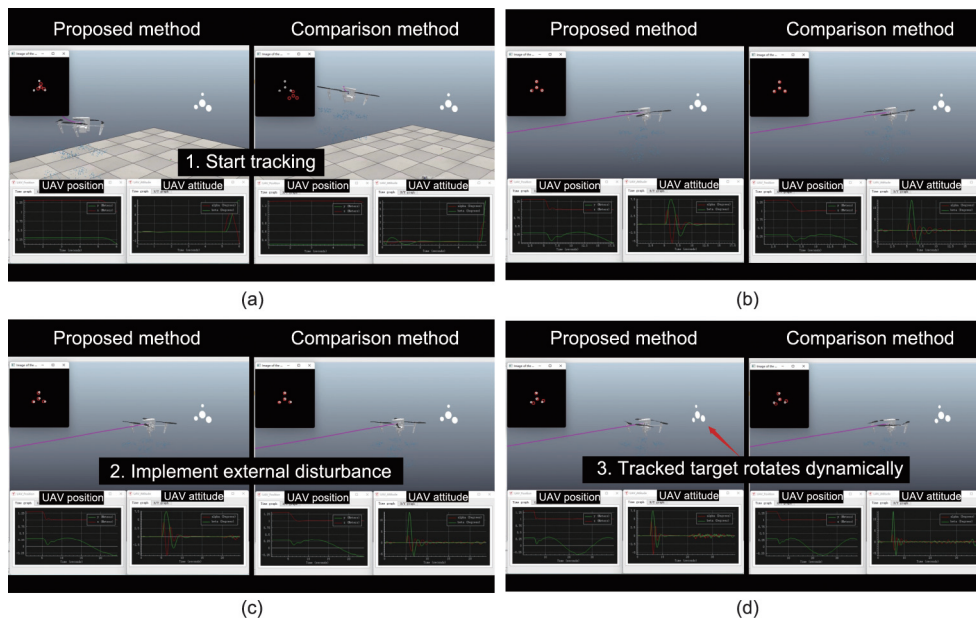**Fig. 3.** The CoppeliaSim simulation environment used for the tracking performance comparison.

Fig. 4. Snapshots of the simulation task execution. (a) Start tracking; (b) stable target tracking; (c) tracking during the disturbance; (d) rotating target tracking.

are applied to the UAV to demonstrate the anti-disturbance capability of the proposed method. Finally, in the rotating target-tracking stage, the attitude of the dynamic target is configured to rotate dynamically to demonstrate the tracking robustness of the rotating target.

4.2. Simulation results and analysis

The comparative simulation results are shown in Figs. 5–10. The position trajectories of the UAV and the tracked target are shown in Figs. 5 and 6, respectively. The image errors, translational velocities, and Euler angles are respectively shown in Figs. 7, 9, and 10. The feature pixel convergence trajectories are shown in Fig. 8. In the target-tracking stage from $t = 0$ to 15 s, the image errors converge to zero in 5 s and stabilize around zero afterward. Compared with the method in Ref. [24], the image, attitude, and translational velocity trajectories of the proposed method are smooth. In the

tracking during the disturbance stage from $t = 15$ to 30 s, although external disturbances lead to a slight chatter phenomenon in the translational velocity and attitude motions, the UAV can maintain a stable tracking trajectory. The image errors of the proposed IBVS controller and the controller in Ref. [24] are within ±10 pixels. The attitude chatter of the proposed IBVS controller is within ±0.02 rad, whereas that of the comparison method is within ±0.025 rad. The proposed method realizes an anti-disturbance performance similar to that of the comparison method without translational velocity measurements. In the rotating target-tracking stage from $t = 30$ to 45 s, the image error in z-direction $q_{1,z}$ of the method in Ref. [24] is increased to ±20 pixels and chatter. For comparison, the image errors in the proposed method are maintained within ±10 pixels. The chatter phenomenon in the translational velocities and attitudes of the method in Ref. [24] is also evident. In comparison, the attitude and translational velocity trajectories in the proposed method are smooth. The image errors of the proposed IBVS
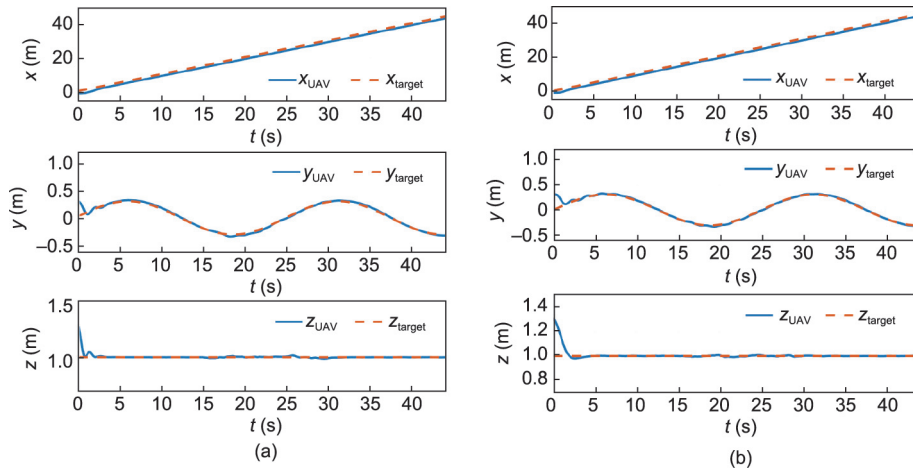
Fig. 5. Positions of the UAV and tracked target in the comparative simulation. (a) Proposed IBVS controller; (b) IBVS controller from Ref. [24]. x, y, z: components of position vector in the x, y, and z directions, respectively; x_{UAV}, y_{UAV}, z_{UAV}: components of UAV position in the x, y, and z directions, respectively; x_{target}, y_{target}, z_{target}: components of tracked target position in the x, y, and z directions, respectively.
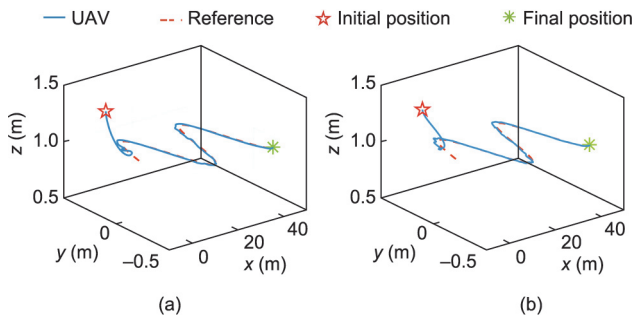

Fig. 6. 3D trajectories of the UAV and tracked target in the comparative simulation. (a) Proposed IBVS controller; (b) IBVS controller from Ref. [24].

controller are stable around zero and are bounded within ±10 pixels.

The root mean square error (RMSE) and standard deviation (STD) are presented in Table 1. The RMSE statistics in Table 1 indicate that the steady-state errors of the proposed IBVS controller are smaller than those of the controller in Ref. [24]. In addition, the STD statistics indicate that the attitude motion of the proposed IBVS controller is smoother than that of the comparison controller.

In general, the image, translational velocity, and attitude motions of the proposed IBVS controller are smooth, resulting in the stable tracking performance of the UAV. The proposed IBVS controller achieves proximity anti-disturbance capability as the comparison controller without the condition of translational velocity measurements. The proposed method can achieve satisfactory tracking accuracy for tracking rotating targets.

5. Experimental results

In this section, a custom-built UAV is used to demonstrate the feasibility of the proposed method. The experimental setup and process are described, and experimental data are provided to illustrate the control performance of the proposed method. The experimental video is available from Appendix A.

5.1. Experimental setup

In the experiments, a quadrotor platform is adopted to verify the feasibility of the proposed dynamic IBVS controller. Based on
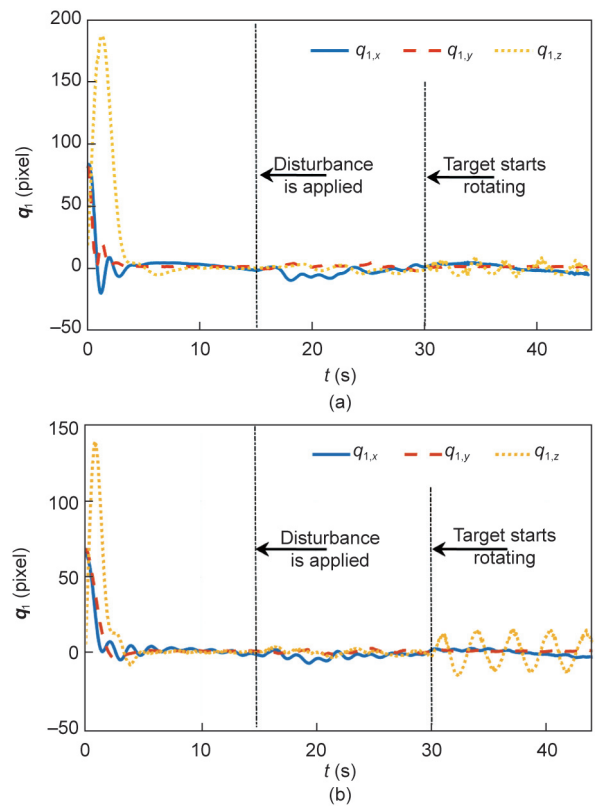
Fig. 7. Image errors in the comparative simulation. (a) Proposed IBVS controller; (b) IBVS controller from Ref. [24] $q_{1,x}$, $q_{1,y}$, $q_{1,z}$: component of q_1 in the x, y, and z directions, respectively.

our previous work [31], a custom-built UAV with an onboard computer, an Ardupilot flight controller, and a monocular camera is used as the primary platform. The hardware structure of the UAV system is shown in Fig. 11. In particular, the embedded Ardupilot provides accurate angular velocities and roll–pitch angles for implementing low-level attitude control by means of an extended Kalman filter (EKF). Optical flow and distance sensors are utilized for the loiter mode to ensure flight safety. An onboard computer is utilized for image processing and controller calculating; it connects with the flight controller through micro air vehicle
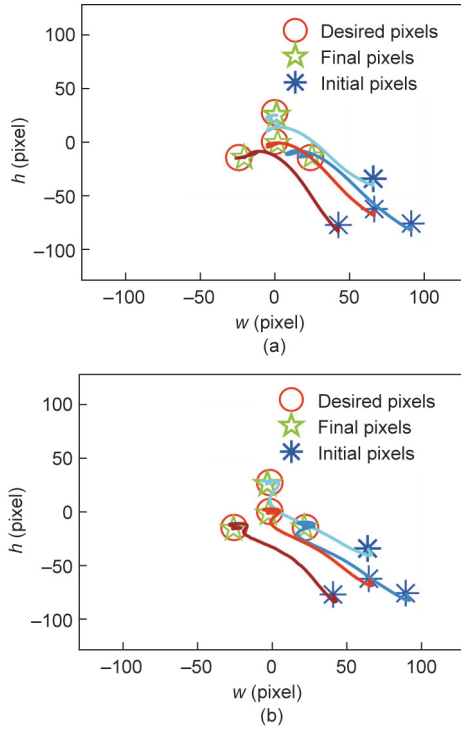
Fig. 8. Convergence trajectories of the feature pixels in the comparative simulation. (a) Proposed IBVS controller; (b) IBVS controller from Ref. [24]. w, h: components of image pixels in the width and height directions, respectively.

link (Mavlink) for attitude measurement acquisition and control signal transmission. The ground station communicates with the onboard computer through wireless fidelity (WiFi) and is utilized to switch the IBVS controller. The control time step is set to 0.025 s, and the resolution and sampling frequency of the camera are 320 × 240 and 60 Hz, respectively. The open-source tag detection AprilTag2 [32] is adopted to acquire image features and UAV relative positions. The selected feature points $\boldsymbol{s}_a$, $\boldsymbol{s}_b$, $\boldsymbol{s}_c$, and $\boldsymbol{s}_i$ are shown in Fig. 11. The geometric parameters of the dynamic target

are $\alpha_i = 0$, $\beta_i = 0.5$, and $\gamma_i = 0.5$, respectively. A mobile robot is utilized as the dynamic target, and wheel odometry is adopted to obtain the position of the mobile robot. A propeller providing wind with a velocity of 10 m·s^{-1} is used to simulate external disturbances. The experiments are implemented without external positioning conditions, such as GPS or a Vicon system.

The experiments are executed in four stages: ① visual positioning from $t = 0$ to 12 s, ② target tracking from $t = 12$ to 32 s, ③ tracking under a disturbance from $t = 32$ to 42 s, and ④ positioning under a disturbance from $t = 42$ to 68 s. First, the IBVS controller is activated to approach and maintain a fixed relative distance from the target. Second, the tracked target moves with a trajectory of $\boldsymbol{p}_t^a = [0.1, 0.05\cos(0.42t)\tanh(0.05t), 0]$, and the tracked target rotates correspondingly with a yaw angle of $\arcsin[0.5\cos(0.42t)\tanh(0.05t)]$. Then, the UAV enters the area disturbed by the wind and continues tracking the target. Finally, the mobile robot stops moving and the UAV maintains its positioning under the disturbance. Detailed experimental configurations are as follow: $M = 0.8$ kg; control time step $\Delta t = 0.02$ s; desired position of the UAV relative to the target $\boldsymbol{p}_t^c = [1, 0, 0]$ m. The desired pixels of $\boldsymbol{s}_a$, $\boldsymbol{s}_b$, $\boldsymbol{s}_c$, and $\boldsymbol{s}_i$ are $\boldsymbol{s}_a^* = [-20, -20]$, $\boldsymbol{s}_b^* = [20, -20]$, $\boldsymbol{s}_c^* = [-20, 20]$, and $\boldsymbol{s}_i^* = [0, 0]$, respectively. The controller parameters are set as $k_1 = 2, k_2 = 3, k_3 = 1$, and $k_4 = 3$.

5.2. Experimental results and analysis

Snapshots of the experiment are shown in Fig. 12, and the experimental results are shown in Figs. 13–19. The Euler angle of the tracked target is shown in Fig. 20. The UAV position and three-dimensional (3D) trajectories obtained by the wheel odometer of the mobile robot and AprilTag are depicted in Figs. 13 and 14. The image moments and pixel convergence trajectories are presented in Figs. 15 and 16, respectively. The outputs of the velocity observer and integral-based filters are shown in Figs. 17 and 18.

In the visual positioning stage, the quadrotor converges to the desired position within 5 s when the IBVS controller is activated and hovers at a relative position of approximately 1 m behind the tracked target. The estimation error of the relative velocity $\tilde{\boldsymbol{\delta}}$ converges to zero and remains within ±0.05 m·s^{-1}. The positioning
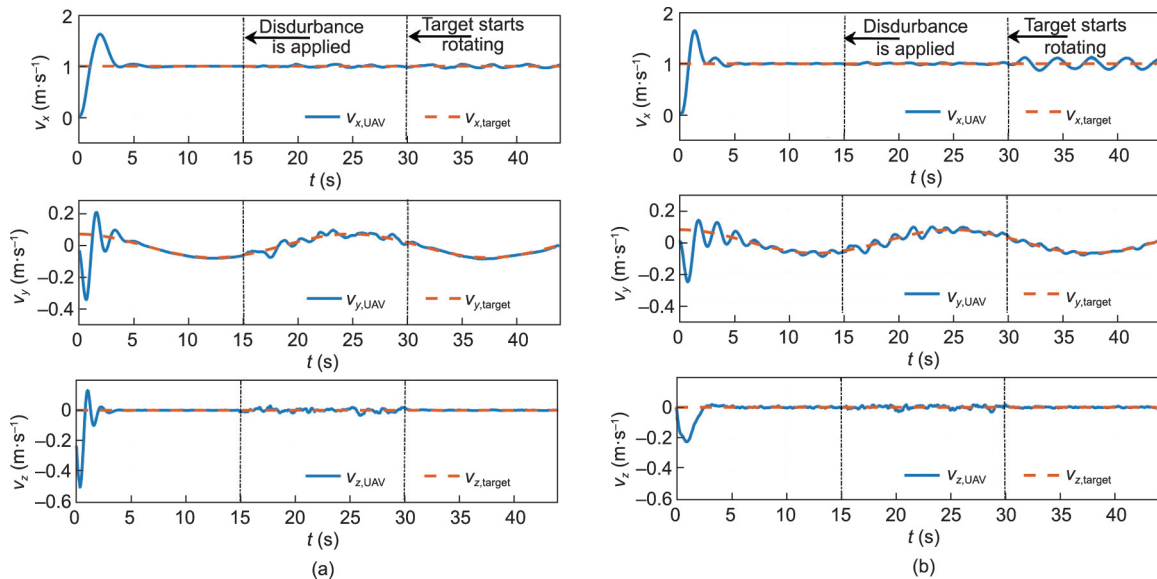

Fig. 9. Translational velocities of the UAV in the comparative simulation. (a) Proposed IBVS controller; (b) IBVS controller from Ref. [24]. v_x, v_y, v_z: components of velocity vector in the x, y, and z directions, respectively; $v_{x,\text{UAV}}$, $v_{y,\text{UAV}}$, $v_{z,\text{UAV}}$: components of UAV velocity in the x, y, and z directions, respectively; $v_{x,\text{target}}$, $v_{y,\text{target}}$, $v_{z,\text{target}}$: components of tracked target velocity in the x, y, and z directions, respectively.
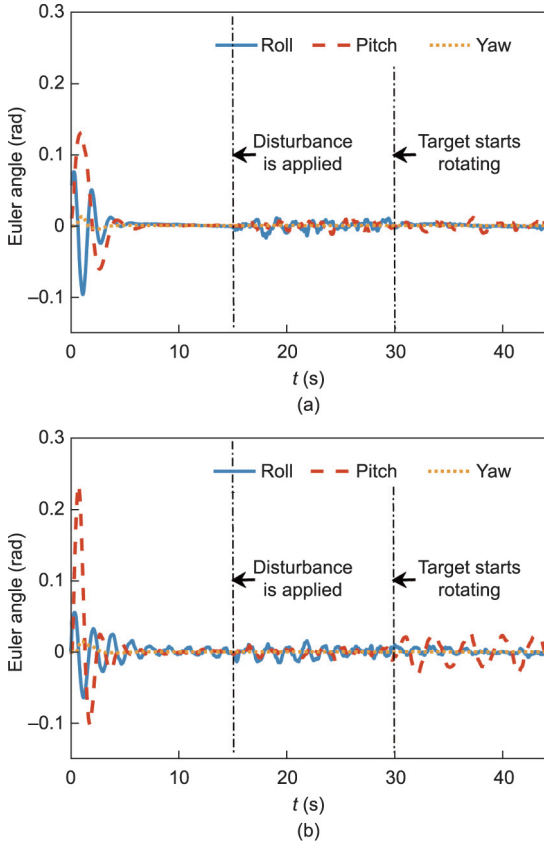
(a)

(b)

Fig. 10. Euler angles of the UAV in the comparative simulation. (a) Proposed IBVS controller; (b) IBVS controller from Ref. [24].

Table 1
Steady-state error statistics of the positions, image moments, and attitudes in the comparative simulation.

State	Proposed IBVS controller		IBVS controller in Ref. [24]	
	RMSE	STD	RMSE	STD
x (cm)	1.5728	0.4870	2.0902	1.3462
y (cm)	0.3789	0.2992	0.3261	0.2306
z (cm)	0.3031	0.1216	0.3634	0.1337
q_x (pixel)	1.6103	1.0695	1.5109	0.9011
q_y (pixel)	1.6136	1.4057	1.8316	1.8294
q_z (pixel)	3.3810	3.3355	7.8623	7.8364
ϕ (rad)	0.0039	0.0037	0.0052	0.0042
θ (rad)	0.0044	0.0041	0.0113	0.0107
ψ (rad)	0.0008	0.0008	0.0009	0.0009

q_x, q_y, q_z: component of $\mathbf{q}$ in x, y, and z directions, respectively; ϕ, θ, ψ: roll, pitch, and yaw angle of the UAV, respectively.

and image errors remain within ±0.1 m and ±10 pixels, respectively. The estimated disturbance of the upper bound converges at approximately 1.1 N. The image depth estimation error remains within ±0.02 m.

In the target-tracking stage, the mobile robot is set to move along the given trajectory, and the UAV starts tracking the target. Due to the unknown acceleration of the tracked target, the tracking errors of the UAV are bounded within ±0.15 m. The estimation errors of the relative velocities and image depth are maintained within ±0.05 m·s⁻¹ and ±0.02 m, respectively. Thanks to the accurate velocity estimation, the UAV can maintain stable dynamic target tracking.

During tracking in the disturbance stage, the UAV enters the disturbance area. Due to the disturbance, the estimation of the
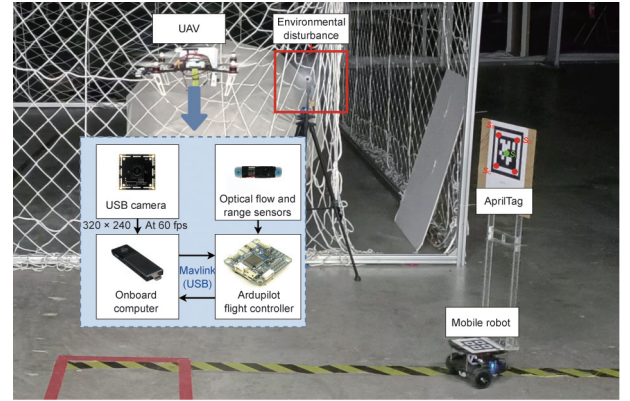
Fig. 11. Experimental hardware configuration for the dynamic target tracking of the UAV. USB: universal serial bus; fps: frames per second.
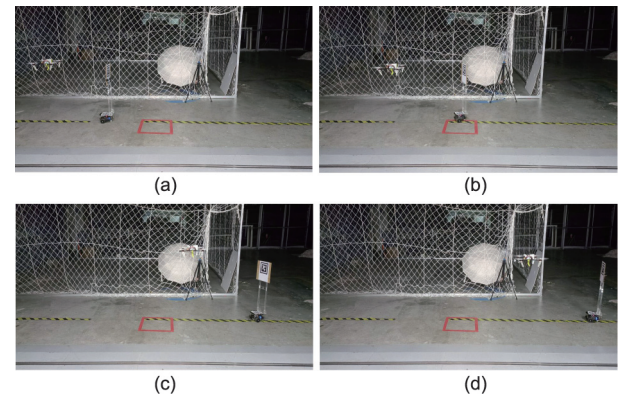

Fig. 12. Snapshots of the dynamic target-tracking experiments. (a) t = 0 s: takeoff; (b) t = 25 s: target tracking; (c) t = 35 s: tracking under a disturbance; (d) t = 45 s: positioning under a disturbance.
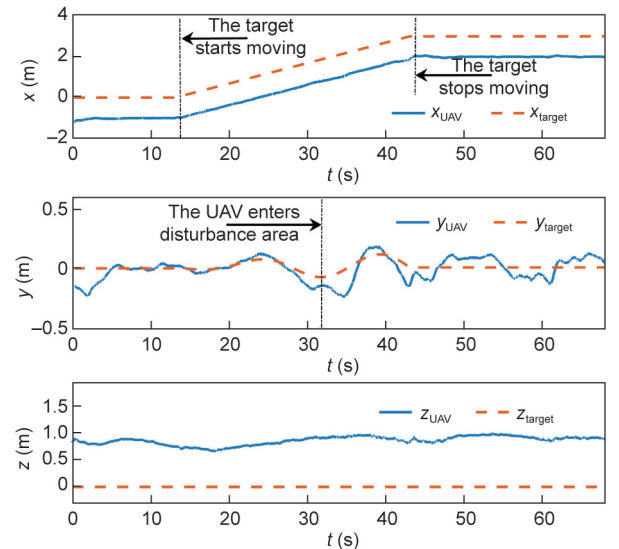

Fig. 13. Positions of the UAV and tracked target.

disturbance upper bound increases and converges to 1.35 N. The estimation error of the image depth converges to a bounded value within ±0.06 m. The UAV can remain stable and maintain a tracking error of ±0.2 m by compensating for the integral-based filter.
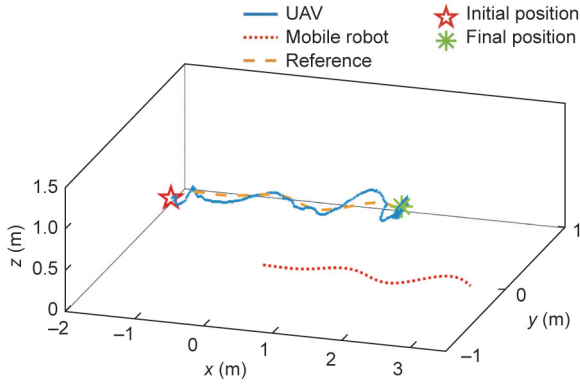
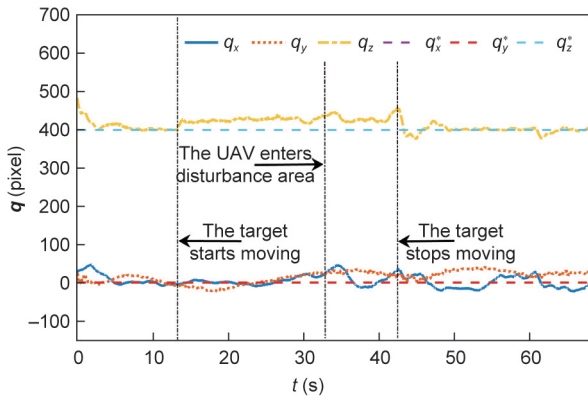**Fig. 14.** 3D trajectories of the UAV, mobile robot, and tracked target.

Fig. 15. Time trajectories of the image moment. q_x^*, q_y^*, q_z^*: component of $\boldsymbol{q}^*$ in x, y, and z directions, respectively.
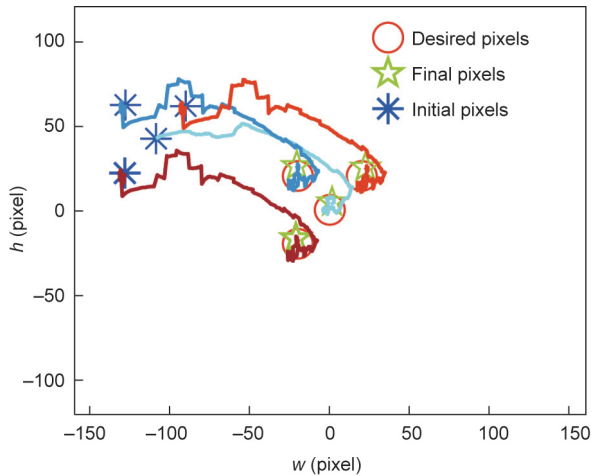

Fig. 16. Pixel trajectories of the feature points in the convergence stage of the IBVS controller.

When positioning during the disturbance stage, the UAV stops tracking as the tracked target stops moving. The UAV hovers under the disturbance and remains at a distance of 1 m from the target. Although the tracking error of the UAV increases compared with that in the visual positioning stage, the UAV still hovers with a stable attitude and maintains a positioning accuracy within ±0.2 m.

In general, the proposed IBVS method is effective for UAV tracking dynamic targets in GPS-denied environments. The control uncertainty generated by external disturbances and the dynamic
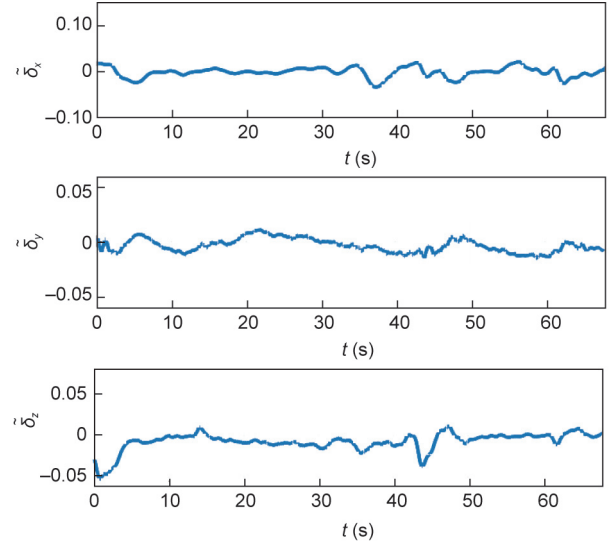
Fig. 17. Estimation error of the relative velocities between the UAV and the tracked target. $\tilde{\delta}_x$, $\tilde{\delta}_y$, $\tilde{\delta}_z$: component of $\tilde{\boldsymbol{\delta}}$ in x, y, and z directions, respectively.
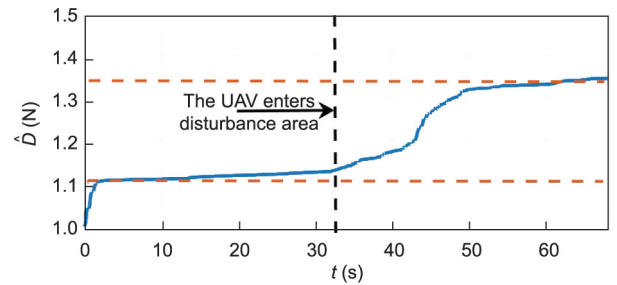

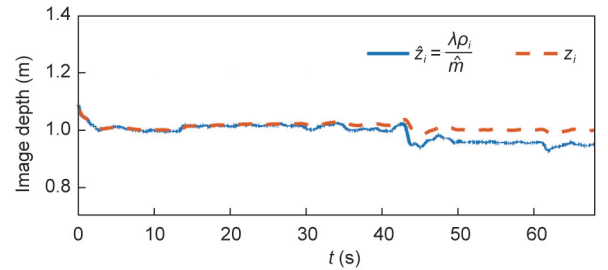**Fig. 18.** Disturbance upper bound estimation of the integral-based filter.

Fig. 19. Image depth estimation $\hat{z}_i = \frac{\lambda \rho_i}{\hat{m}}$ and actual image depth z_i of feature point i.
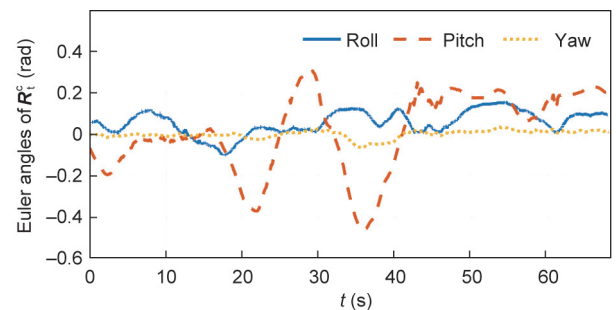

Fig. 20. Euler angles of the tracked target relative to the camera. $\boldsymbol{R}_t^c$: rotation matrix of the tracked target relative to the camera.

rotation of the tracking target can be mitigated using the proposed IBVS method.

6. Conclusions

This study proposed a dynamic IBVS method based on a velocity observer that improves the tracking performance of a UAV under unpredictable disturbances. The proposed method derives the image dynamics using multiple feature points of a virtual camera. The image dynamics of an underactuated UAV can be decoupled and simplified based on a virtual camera. A novel image depth model was proposed to enable a UAV to track a rotating target with an arbitrary orientation, and a velocity observer was proposed to estimate the relative velocity between the UAV and a dynamic target. Furthermore, unpredictable disturbances were estimated by means of an integral-based filter and were compensated for in the observer and IBVS controller design. The stabilities of the proposed velocity observer and IBVS method were analyzed based on the Lyapunov theory. A comparative simulation and an experiment with multiple stages were conducted to verify the effectiveness and robustness of the proposed dynamic IBVS method. In our future research, the proposed dynamic IBVS method will be applied in practical application scenarios of UAV, for purposes such as capturing dynamic targets and autonomous landing.

Compliance with ethics guidelines

Yanjie Chen, Yangning Wu, Limin Lan, Hang Zhong, Zhiqiang Miao, Hui Zhang, and Yaonan Wang declare that they have no conflict of interest or financial conflicts to disclose.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.eng.2023.05.017.

References

[1] Chung HM, Maharjan S, Zhang Y, Eliassen F, Strunz K. Placement and routing optimization for automated inspection with unmanned aerial vehicles: a study in offshore wind farm. IEEE Trans Industr Inform 2021;17(5):3032–43.

[2] Kendoul F. Survey of advances in guidance, navigation, and control of unmanned rotorcraft systems. J Field Robot 2012;29(2):315–78.

[3] Scaramuzza D, Achtelik MC, Doitsidis L, Friedrich F, Kosmatopoulos E, Martinelli A, et al. Vision-controlled micro flying robots: from system design to autonomous navigation and mapping in GPS-denied environments. IEEE Robot Autom Mag 2014;21(3):26–40.

[4] McFadyen A, Jabeur M, Corke P. Image-based visual servoing with unknown point feature correspondence. IEEE Robot Autom Lett 2017;2(2):601–7.

[5] Rafique MA, Lynch AF. Output-feedback image-based visual servoing for multirotor unmanned aerial vehicle line following. IEEE Trans Aerosp Electron Syst 2020;56(4):3182–96.

[6] Serra P, Cunha R, Hamel T, Cabecinhas D, Silvestre C. Landing of a quadrotor on a moving target using dynamic image-based visual servo control. IEEE Trans Robot 2016;32(6):1524–35.

[7] Xie H, Low KH, He Z. Adaptive visual servoing of unmanned aerial vehicles in GPS-denied environments. IEEE/ASME Trans Mechatron 2017;22(6):2554–63.

[8] Zhang X, Fang Y, Zhang X, Jiang J, Chen X. A novel geometric hierarchical approach for dynamic visual servoing of quadrotors. IEEE Trans Ind Electron 2020;67(5):3840–9.

[9] Zhang K, Shi Y, Sheng H. Robust nonlinear model predictive control based visual servoing of quadrotor UAVs. IEEE/ASME Trans Mechatron 2021;26(2):700–8.

[10] Chaumette F, Hutchinson S. Visual servo control. I. Basic approaches. IEEE Robot Autom Mag 2006;13(4):82–90.

[11] Chaumette F, Hutchinson S. Visual servo control. II. Advanced approaches. IEEE Robot Autom Mag 2007;14(1):109–18.

[12] Zhong H, Miao Z, Wang Y, Mao J, Li L, Zhang H, et al. A practical visual servo control for aerial manipulation using a spherical projection model. IEEE Trans Ind Electron 2020;67(12):10564–74.

[13] Sun N, Yang T, Chen H, Fang Y. Dynamic feedback antiswing control of shipboard cranes without velocity measurement: theory and hardware experiments. IEEE Trans Industr Inform 2019;15(5):2879–91.

[14] Abdessameud A, Tayebi A. Global trajectory tracking control of VTOL-UAVs without linear velocity measurements. Automatica 2010;46(6):1053–9.

[15] Zhang X, Fang Y, Zhang X, Jiang J, Chen X. Dynamic image-based output feedback control for visual servoing of multirotors. IEEE Trans Industr Inform 2020;16(12):7624–36.

[16] Wang H, Zheng D, Wang J, Chen W, Yuan J. Ego-motion estimation of a quadrotor based on nonlinear observer. IEEE/ASME Trans Mechatron 2018;23(3):1138–47.

[17] Shao X, Zhang J, Zhang W. Distributed cooperative surrounding control for mobile robots with uncertainties and aperiodic sampling. IEEE Trans Intell Transp Syst 2022;23(10):18951–61.

[18] Zhang J, Shao X, Zhang W, Na J. Path-following control capable of reinforcing transient performances for networked mobile robots over a single curve. IEEE Trans Inst Meas 2023;72:1–12.

[19] Lee D, Ryan T, Kim HJ. Autonomous landing of a VTOL UAV on a moving platform using image-based visual servoing. In: Proceedings of 2012 IEEE International Conference on Robotics and Automation (ICRA); 2012 May 14–18; Saint Paul, MN, USA. IEEE; 2012. p. 971–6.

[20] Jabbari Asl H, Oriolo G, Bolandi H. Output feedback image-based visual servoing control of an underactuated unmanned aerial vehicle. Proc Inst Mech Eng Part I J Syst Control Eng 2014;228(7):435–48.

[21] Cao Z, Chen X, Yu Y, Yu J, Liu X, Zhou C, et al. Image dynamics-based visual servoing for quadrotors tracking a target with a nonlinear trajectory observer. IEEE Trans Syst Man Cybern Syst 2020;50(1):376–84.

[22] Lin J, Wang Y, Miao Z, Zhong H, Nie J, Fierro R. Robust image-based landing control of a quadrotor on an unknown moving platform using circle features. In: Proceedings of 2021 IEEE International Conference on Real-Time Computing and Robotics (RCAR); 2021 Jul 15–19; Xining, China. IEEE; 2021. p. 177–82.

[23] Zheng D, Wang H, Wang J, Chen S, Chen W, Liang X. Image-based visual servoing of a quadrotor using virtual camera approach. IEEE/ASME Trans Mechatron 2017;22(2):972–82.

[24] Li J, Xie H, Low KH, Yong J, Li B. Image-based visual servoing of rotorcrafts to planar visual targets of arbitrary orientation. IEEE Robot Autom Lett 2021;6(4):7861–8.

[25] Liang J, Chen Y, Lai N, He B, Miao Z, Wang Y. Low-complexity prescribed performance control for unmanned aerial manipulator robot system under model uncertainty and unknown disturbances. IEEE Trans Industr Inform 2022;18(7):4632–41.

[26] Zhang W, Shao X, Zhang W, Qi J, Li H. Unknown input observer-based appointed-time funnel control for quadrotors. Aerosp Sci Tech 2022;126:107351.

[27] Zhang Z, Chen Y, Wu Y, Lin L, He B, Miao Z, et al. Gliding grasping analysis and hybrid force/position control for unmanned aerial manipulator system. ISA Trans 2022;126:377–87.

[28] Xie H, Lynch AF, Low KH, Mao S. Adaptive output-feedback image-based visual servoing for quadrotor unmanned aerial vehicles. IEEE Trans Control Syst Technol 2020;28(3):1034–41.

[29] Kumar V, Michael N. Opportunities and challenges with autonomous micro aerial vehicles. Int J Robot Res 2012;31(11):1279–91.

[30] Fresk E, Nikolakopoulos G. Full quaternion based attitude control for a quadrotor. In: Proceedings of 2013 European Control Conference (ECC); 2013 Jul 17–19; Zurich, Switzerland. IEEE; 2013. p. 3864–9.

[31] Lai N, Chen Y, Liang J, He B, Zhong H, Wang Y. An onboard-eye-to-hand visual servo and task coordination control for aerial manipulator based on a spherical model. Mechatronics 2022;82:102724.

[32] Wang J, Olson E. AprilTag 2: efficient and robust fiducial detection. In: Proceedings of 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2016 Oct 9–14; Daejeon, Republic of Korea. IEEE; 2016. p. 3400–7.