Computers, Materials & Continua

Tech Science Press

https://doi.org/10.32604/cmc.2026.081203

ARTICLE

Disturbance Observers-Based Adaptive Visual Servoing for Aerial Vehicle with Trajectory Tracking Applications of Soccer

Yao-Bo Long[1], Yu-Ke Ouyang[2], Bo Zhuang[2] and Ao-Qi Liu[3,*]

[1]School of General Education, Guangdong University of Science and Technology, Dongguan, China
[2]School of Automation, Guangdong University of Technology, Guangzhou, China
[3]School of Computer Science and Engineering, Macau University of Science and Technology, Macau, China

*Corresponding Author: Ao-Qi Liu. Email: 3230006888@student.must.edu.mo

Received: 25 February 2026; Accepted: 13 April 2026; Published: 15 June 2026

ABSTRACT: This study addresses the real-time visual tracking task in edge environments by proposing a robust visual servoing control system based on a higher-order sliding mode observer, enabling a quadrotor UAV to autonomously track a moving soccer ball during outdoor sports broadcasts while relying solely on a monocular camera and an inertial measurement unit, thereby eliminating any dependency on external positioning or velocity sensors such as GPS. The system adopts a hierarchical control architecture in which the observer plays a central role: operating on resource-constrained edge devices, it leverages only visual information to estimate unknown external disturbances and target motion states in real time, significantly reducing both hardware requirements and computational overhead to meet the lightweight deployment demands of edge nodes while simultaneously providing the continuous disturbance estimation necessary to sustain a low-latency control loop. Experimental validation was conducted under simulated real soccer match scenarios featuring wind disturbances and abrupt trajectory changes, and the results demonstrate that the system restores the ball to a stable position near the center of the camera's field of view within seconds of a perturbation, thereby satisfying the stringent stability and real-time performance requirements of live broadcasting. This work confirms the feasibility of delivering low-latency, highly reliable visual intelligence services on resource-constrained edge platforms through advanced observer design, offering a technical reference for the convergence of real-time perception and control within the edge Internet of Things and providing valuable insights for future research on large-scale edge collaboration and secure, sustainable service architectures.

KEYWORDS: Image-based visual servoing; adaptive control; aerial vehicles; disturbance observers; soccer ball

1 Introduction

The increasing demand for immersive viewing experiences in live sports broadcasting—particularly in dynamic football matches—has revealed the inherent limitations of conventional static camera systems [1]. Traditional fixed-angle cameras cannot adequately capture rapid positional transitions, such as counterattacks or long-ball trajectories, leading to fragmented spectator engagement. In contrast, unmanned aerial vehicles (UAVs), with their agile mobility and autonomous tracking capabilities [2–4], provide a disruptive solution capable of delivering cinematic-level perspectives beyond the reach of terrestrial setups. In this context, the visual-based drone tracking control method can effectively address the shortcomings of the traditional live streaming approach. Football tracking represents a fundamental task in drone-assisted live broadcasting, where the primary objective is high-precision trajectory tracking of the football's motion path. This aligns with the broader control paradigm of visual servoing for unmanned aerial systems [5,6].

Within the domain of robotic vision-guided control, two principal methodologies have emerged: position-based visual servoing (PBVS), which relies on reconstructed 3D target pose, and image-based visual servoing (IBVS), which directly utilizes 2D image features for error regulation [7]. The latter approach circumvents the need for explicit pose estimation, thereby mitigating projection model inaccuracies and camera calibration errors that commonly degrade PBVS performance. In quadrotor systems, traditional depth estimation approaches often lead to increased computational complexity and measurement inaccuracies [8]. To overcome these challenges, alternative solutions have been proposed, including techniques based on spherical image moments [9] and virtual image moments [10]. The virtual image moments approach introduces an auxiliary reference frame that effectively isolates roll and pitch dynamics, consequently streamlining controller development. Experimental validation has confirmed its superior effectiveness in stationary positioning scenarios [11].

While visual servoing methods have demonstrated considerable success in UAV control, significant challenges persist in tracking unknown dynamic targets and mitigating external disturbances [12], particularly limiting their application in dynamic environments like live soccer broadcasting. Current research predominantly assumes perfect knowledge of target dynamics, which proves unrealistic in practice—for instance, the instantaneous velocity of a soccer ball during play cannot be predetermined. Several approaches have attempted to address these limitations: barrier Lyapunov function-based controllers [13] eliminate the need for explicit platform dynamics, while auxiliary dynamic systems [14] handle unknown target velocities by treating them as bounded disturbances, albeit with inherent steady-state tracking errors. More advanced solutions employ nonlinear tracking-differentiator observers [15] to estimate target motion through visual cues, with subsequent developments [16] extending virtual camera IBVS through velocity estimation for improved tracking of arbitrarily oriented planar targets. Nevertheless, these approaches still depend on velocity measurements from the UAV to indirectly determine the moving target's velocity. The noisy velocity signals propagate through the control system, causing undesirable oscillations in the attitude control loop and increased power expenditure [17].

In practical applications, UAVs conducting aerial live broadcasts over open soccer fields are susceptible to environmental disturbances such as wind, which can significantly compromise tracking accuracy and consequently degrade broadcast quality. To systematically address the challenge of disturbance rejection in quadrotor visual servoing, Xie et al. have developed a series of adaptive control solutions. Their initial work [18] established adaptive control schemes to enhance servoing performance under external disturbances. Building upon this foundation, their subsequent research [19] introduced an advanced output feedback IBVS approach with improved disturbance compensation capabilities. Nevertheless, these methods capable of compensating for external disturbances, though its effectiveness against time-varying disturbances in real-world scenarios remains limited. In work [20], a nonlinear model predictive control strategy was developed through the offline design of tightened constraint sets. Work [21] introduced a saturated control scheme for rotational dynamics to counteract external disturbances. However, model-based control approaches often rely on worst-case assumptions, trading off nominal performance for enhanced robustness [22]. Observer-based methodologies have gained prominence in addressing the critical trade-off triangle of control precision, disturbance attenuation, and closed-loop stability. Work [23] introduced an enhanced extended state observer (ESO) design exhibiting improved disturbance rejection capabilities. Finite-time estimation of external disturbances was accomplished using sliding-mode observers (SMOs) in [24]. To address the high-frequency chattering issue inherent in conventional sliding mode observers, reference [25] pioneered the introduction of a higher-order sliding mode observer (HOSMO) to achieve continuous estimation. Work [26] introduced a higher-order sliding mode observer design that leverages quadrotor position and velocity state feedback to enable both continuous and finite-time disturbance

estimation. Recently, work [27] further designed HOSMO and successfully achieved highly accurate synchronous estimation of the relative speed of moving targets and external disturbances. While this method demonstrates effectiveness, its practical implementation critically depends on high-accuracy measurements of both positional and velocity states.

Building upon these research findings, this study focuses on controlling a quadrotor UAV to track a moving soccer ball, addressing needs distinct from conventional fixed-view broadcasting in football matches. In this work, we decompose the UAV control system into two cascaded loops: an attitude loop and a position loop. For the position control loop, we develop a high-order sliding mode observer that utilizes visual information to estimate both external disturbances and the football's velocity. A backstepping-based controller is then designed to actively compensate for these estimated disturbances and unknown football velocity, enabling accurate tracking of the football's trajectory. The attitude loop derives desired angular velocities from the position loop's thrust commands, while a torque controller ensures the UAV's angular velocities precisely track these reference commands. In this work, the onboard sensing suite comprises only a downward-facing camera and a standard IMU; no external positioning systems or dedicated velocity sensors are required. Compared with existing literature, the main contributions of this work are summarized as follows:

(1) The proposed control architecture incorporates a state observer to reconstruct velocity information, providing a complete solution for vision-based football tracking that operates without physical velocity sensors and relies solely on an IMU and an onboard camera.
(2) The proposed control scheme guarantees asymptotic convergence of tracking errors for UAVs operating under time-varying external disturbances.
(3) The corresponding torque controller is designed based on the full nonlinear quadrotor dynamics rather than employing linearization techniques, thereby significantly enhancing the dynamic responsiveness of the IBVS strategy.

This paper is organized as follows. Section 2 introduces a quadrotor dynamic and a image dynamic. Section 3 presents the controller design and provides Lyapunov-based stability proofs for the closed-loop system. Section 4 presents comprehensive performance evaluation and experimental results of the proposed method, followed by concluding remarks in Section 5.

2 Problem Formulation

This section formulates the rigid-body dynamic model of a quadrotor UAV in three-dimensional space and establishes a perspective projection model for the onboard camera system. By introducing a virtual camera paradigm, we derive image moment features defined on the virtual imaging plane, followed by a formal problem statement for the visual tracking task.

2.1 System Description

To characterize the kinematic relationship between the UAV and the target soccer ball, four coordinate frames are considered for the quadrotor IBVS system description. The inertial frame $F_i = \{O_i \;\; X_i \;\; Y_i \;\; Z_i\}$ is aligned with north, east, down. $F_b = \{O_b \;\; X_b \;\; Y_b \;\; Z_b\}$ represents a body-fixed coordinate system aligned with forward, right, and down axes, with its origin located at the center of mass of the quadrotor. Orign of frame B is located at $p = (x \;\; y \;\; z)^T$ with respect to the Inertia frame, its attitude is represented by three Euler angles ϕ, θ, and ψ, corresponding to roll, pitch, and yaw angles, respectively. The transformation can be represented by a rotation matrix R: $F_b \rightarrow F_i$. For the convenience of presentation, it is assumed that the camera frame $F_c = \{O_c \;\; X_c \;\; Y_c \;\; Z_c\}$ aligns with the quadrotor body frame. Consider a virtual camera frame

$F_v = \{O_v \ \ X_v \ \ Y_v \ \ Z_v\}$ that inherits the translational motion of the real quadrotor without rotation, with its origin aligned with that of the actual camera frame, and its roll and pitch angles are zero and yaw angle is the same as the actual camera frame. The geometric relationships between the four frames is shown in Fig. 1.

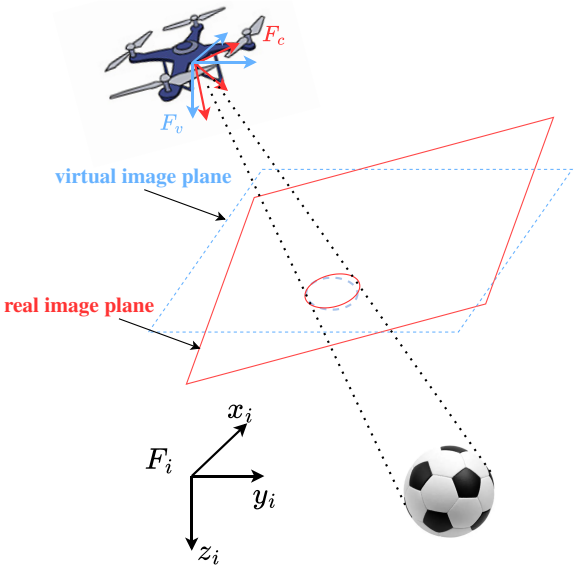


Figure 1: The geometric relationship among the camera frame, virtual image frame, and inertial frame.

Establishing the quadrotor dynamics model in the virtual camera coordinate system are as follows:

$$\dot{p} = v, \tag{1}$$

$$\dot{v} = -\frac{fRe_3}{m} + ge_3 + d_f \tag{2}$$

$$\dot{R} = Rsk(\Omega) \tag{3}$$

$$J\dot{\Omega} = \tau - \Omega \times J\Omega + d_\tau \tag{4}$$

where the mass of the quadrotor m, gravitational acceleration g, and symmetric inertia matrix J define its physical properties. v and Ω respectively represent the linear velocity of the quadrotor within the virtual camera plane and the angular velocity in the body coordinate frame. f and τ represent the thrust force and torque inputs to the quadrotor. Lumped disturbance forces d_f and torques d_τ act on the translational and rotational dynamics of the quadrotor, respectively, arising from both unmodeled dynamics and external disturbances. The unit vector is defined as $e_3 = [0 \ \ 0 \ \ 1]^T$, and $sk(\cdot)$ denotes the skew-symmetric matrix operator.

To facilitate the derivation of the quadrotor dynamic model and enable disturbance identification and compensation, this paper adopts the following fundamental assumptions:

Assumption 1: The Euler angles of the quadrotor are constrained within the following ranges:

$$\phi \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \psi \in [-\pi, \pi].$$

Assumption 2: The disturbance d_f and d_τ in dynamics are second-order differentiable, with bounded first-order and second-order derivatives.

2.2 Image Dynamics

Our target of interest is the motion of a soccer ball, which projects as a circular shape in the camera plane. We therefore derive the image dynamics equation based on the virtual imaging plane using circular image moments. For a soccer ball located at $P_c = (X, Y, Z)$ in the camera coordinate frame, the corresponding image-plane coordinates are given by

$$p_c = \begin{bmatrix} u_c \\ n_c \end{bmatrix} = \frac{f_c}{Z} \begin{bmatrix} X \\ Y \end{bmatrix} \tag{5}$$

with $f_c > 0$ is the focal length of the camera.

The soccer ball coordinates in the virtual camera frame are $P_v = (X_v, Y_v, Z_v)$. Based on the virtual imaging plane defined in the previous section, and utilizing the quadrotor angular measurements directly obtained from the IMU, we project points from the physical imaging plane to the virtual imaging plane through rotation matrix transformation:

$$p_v = \frac{1}{r_3 \bar{p}} \begin{bmatrix} r_1 \bar{p} \\ r_2 \bar{p} \end{bmatrix} \tag{6}$$

where $p_v = [u_v, n_v]^T$ denote the coordinates of the soccer ball projection in the virtual imaging plane, $\bar{p} = [u_c, n_c, 1]^T \in \mathbb{R}^3$, the $r_i \in \mathbb{R}^{1\times 3}$ is the ith row of the orientation matrix R, for $i = 1, 2, 3$.

Then, the circle-based image moments m_{ij} of order $i + j$ are defined as

$$m_{ij} = \iint u_v^i n_v^j du_v dn_v \tag{7}$$

and the image feature vector is defined by

$$q = [q_z x_g, q_z y_g, q_z]^T \tag{8}$$

The visual feature parameters are mathematically defined through image moment calculations, where $x_g = \frac{m_{10}}{a}$ and $y_g = \frac{m_{01}}{a}$ specify the centroid coordinates of the circular feature in the virtual image plane, with $a = m_{00}$ representing its observed area. The normalized area metric $q_z = z_d \sqrt{\frac{a^*}{a}}$, incorporates the expected feature area a^* at the desired quadrotor pose, while $z_d > 0$ corresponds to the target vertical coordinate in the body-fixed frame when the UAV achieves its reference position. These parameters collectively form the basis for visual servoing control, where (x_g, y_g) provides positional feedback and q_z serves as a scale-invariant feature measurement. The reference features are given by the constant vector $q_d = [x_d, y_d, z_d]$, which represents the desired target in the virtual image plane at a desired height z_d.

Define the image moment tracking error as

$$q_1 = q - q_d. \tag{9}$$

Under the virtual quadrotor rotation-independent assumption and the kinematics, the differential relation $\dot{P}_v = -(v - v_d)$ holds, where the v_d denotes the velocity of the target soccer ball. This, when integrated with the projective mapping (Eq. (5)), generates the visual dynamics as

$$\dot{q} = -(v - v_d). \tag{10}$$

Let us define the velocity error term as

$$e_v = v - v_d. \tag{11}$$

Then we express the complete dynamics, incorporating both the velocity error dynamics and the image feature dynamics, in the following manner:

$$\dot{q}_1 = -e_v \tag{12}$$

$$\dot{e}_v = -\frac{fRe_3}{m} + ge_3 + d_f - \dot{v}_d \tag{13}$$

Remark 1: Based on the perspective projection model, a visual servoing system is constructed using image moment features to establish precise mapping between image space and Cartesian space. This feature design offers two critical advantages: geometric invariance and dynamic decoupling. The virtual image plane maintains parallel alignment with the target plane during UAV attitude variations, while the scaled area feature directly correlates with depth information, thereby forming a complete image feature state representation.

3 Controller Design

This section first designs a disturbance observer and linear velocity estimator based on quadrotor dynamics and image feature dynamics. Under conditions involving external disturbances and unknown target velocity, a visual feedback controller is derived using the backstepping method. Finally, a torque controller is developed to guarantee convergence of angular tracking errors, with rigorous stability proofs provided through Lyapunov theory. The control block diagram of the system is shown in the Fig. 2.

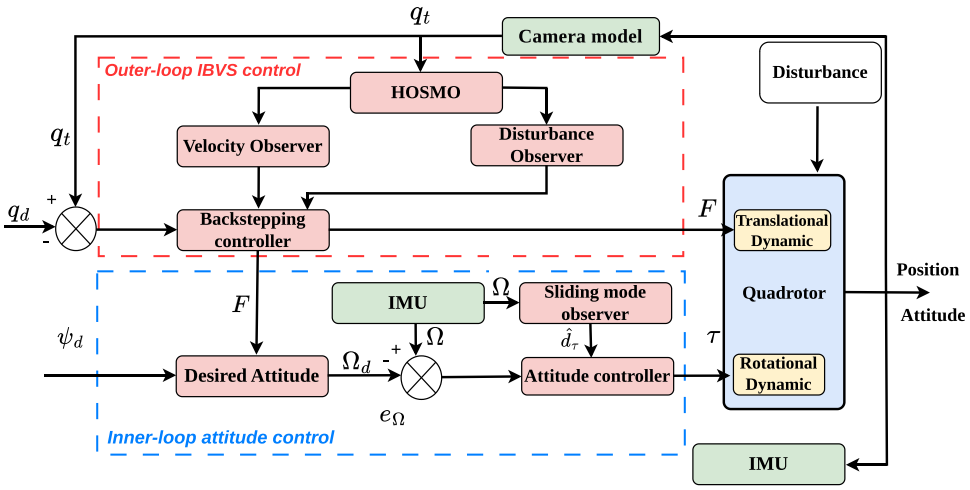


Figure 2: Control structure.

3.1 Disturbance and Velocity Observers Design

The designed hierarchical control framework employs a dual-loop structure: the vision-based outer-loop controller calculates the necessary thrust force F according to visual servoing requirements and generates the target angular velocity command, which functions as the reference input for the inner-loop

attitude control system. The inner-loop controller maintains stable quadrotor attitude control by producing accurate torque τ outputs.

Rewrite the velocity error dynamics in a concise form:

$$\dot{e}_v = F + D_f \tag{14}$$

$$F = -\frac{fRe_3}{m} + ge_3, D_f = d_f - \dot{v}_d \tag{15}$$

The control objective necessitates trajectory tracking of a moving soccer ball by the quadrotor. Thus, D_f represents a lumped disturbance term governing the velocity error dynamics, aggregating parametric uncertainties, exogenous disturbances, and target-induced motion perturbations. Since each component of the lumped disturbance is twice differentiable with bounded second derivatives, a constant A satisfying $||D_f|| \leq A$ can be deduced from Assumptions 2. In practice, an exact value of A is not required; a sufficiently large conservative estimate can be used to initialize the observer gains, which are then tuned experimentally to achieve a balance between estimation accuracy and robustness to measurement noise.

To facilitate subsequent derivations, for any real-valued vector $\mathbf{x} \in \mathbb{R}^n$, we define the element-wise function:

$$\begin{aligned}\text{sign}(\mathbf{x}) &\triangleq [\text{sign}(x_1), \text{sign}(x_2), \ldots, \text{sign}(x_n)]^T, \\ |\mathbf{x}| &\triangleq [|x_1|, |x_2|, \ldots, |x_n|]^T, \\ \mathbf{x}^a &\triangleq [x_1^a, x_2^a, \ldots, x_n^a]^T.\end{aligned}$$

For each component:

$$\text{sign}(x_i) = \begin{cases} +1 & \text{if } x_i > 0 \\ 0 & \text{if } x_i = 0 \\ -1 & \text{if } x_i < 0 \end{cases}, \quad \forall i = 1, 2, \ldots, n.$$

To estimate the external disturbance force D_f and velocity error e_v, the high-order sliding mode observer in the position loop are designed based on Eqs. (12) and (14) as follows:

$$\begin{aligned}\dot{\sigma}_1 &= -\sigma_2 + z_1, \\ \dot{\sigma}_2 &= F + \sigma_3 + z_2, \\ \dot{\sigma}_3 &= \sigma_4 + z_3, \\ \dot{\sigma}_4 &= z_4, \\ \hat{e}_v &= \sigma_2, \\ \hat{D}_f &= \sigma_3.\end{aligned} \tag{16}$$

The formulation incorporates four correction terms $z_1 - z_4$ and an auxiliary parameter α. The auxiliary variable α is initialized to zero for computational convenience. The correction terms are design as

$$\begin{aligned}z_1 &= k_{z_1}|q_1 - \sigma_1|^{\frac{3}{4}} sign(q_1 - \sigma_1), \\ z_2 &= k_{z_2}|z_1|^{\frac{2}{3}} sign(z_1), \\ z_3 &= k_{z_3}|z_2|^{\frac{1}{2}} sign(z_2),\end{aligned} \tag{17}$$

$$z_4 = k_{z_4}|z_3|sign(z_3),$$

where the k_{z_i}, $i = 1, 2, 3, 4$, are some positive constant.

The error terms are given by: $\tilde{\sigma}_1 = q_1 - \sigma_1$, $\tilde{\sigma}_2 = e_v - \sigma_2$, $\tilde{\sigma}_3 = D_f - \sigma_3$, $\tilde{\sigma}_4 = \dot{D}_f - \sigma_4$. Then, the error dynamics can be written as

$$\Xi : \begin{cases} \dot{\widetilde{\sigma}}_1 = -k_{z,1}|\widetilde{\sigma}_1|^{\frac{3}{4}}\text{sign}(\widetilde{\sigma}_1) - \widetilde{\sigma}_2, \\ \dot{\widetilde{\sigma}}_2 = -k_{z,2}|z_1|^{\frac{2}{3}}\text{sign}(z_1) + \widetilde{\sigma}_3, \\ \dot{\widetilde{\sigma}}_3 = -k_{z,3}|z_2|^{\frac{1}{2}}\text{sign}(z_2) + \widetilde{\sigma}_4, \\ \dot{\widetilde{\sigma}}_4 = -k_{z,4}\text{sign}(z_3) + \ddot{D}_f. \end{cases}$$

With appropriately chosen observer gains

$$k_{z,1} = 5A^{1/4},\ k_{z,2} = 3A^{1/2},$$
$$k_{z,3} = 1.5A^{3/4},\ k_{z,4} = 1.1A.$$

Based on Lemma 8 in [25], we obtain that the system Ξ will converge to zero in finite time. After the convergence of the errors, $\hat{D}_f$ and $\hat{e}_v$ approaches to D_f and e_v in finite time, respectively.

Remark 2: Owing to the pronounced dynamic coupling inherent in quadrotor systems, the observer-based intermediate input necessitates availability of its first and second time derivatives. Verified herein, $\hat{D}_f$ produces smooth estimations of D_f, ensuring second-order differentiability of control input. As established in [25], the HOSMO inherently exhibits finite-time convergence. This property inherently guarantees satisfaction of the separation principle, enabling decoupled design of controller and observer.

3.2 Outer-Loop Controller Design

Based on the image feature errors, the following Lyapunov function is defined:

$$V_1 = \frac{1}{2}q_1^T q_1. \tag{18}$$

By substituting Eq. (12), the time derivative of V_1 is obtained as

$$\dot{V}_1 = -q_1^T e_v = -q_1^T(\hat{e}_v + \tilde{e}_v) \tag{19}$$

Let us consider $\hat{e}_v$ as control input, the desired value $\hat{e}_v^d$ of $\hat{e}_v$ is set as $\hat{e}_v^d = k_1 q_1$, with the constant parameter $k_1 > 0$. While direct assignment of $\hat{e}_v$ as a control input is infeasible, the backstepping methodology introduces a secondary error variable:

$$q_2 = q_1 - \frac{1}{k_1}\hat{e}_v. \tag{20}$$

The incorporation of the new error term modifies the time derivatives of both (12) and (19), resulting in

$$\dot{q}_1 = -\tilde{e}_v - k_1(q_1 - q_2) \tag{21}$$
$$\dot{V}_1 = -q_1^T\tilde{e}_v - q_1^T k_1(q_1 - q_2). \tag{22}$$

Using Eqs. (21) and (16), time derivative of q_2 is

$$\dot{q}_2 = -\tilde{e}_v - k_1(q_1 - q_2) - \frac{1}{k_1}(z_2 + F + \hat{D}_f) \tag{23}$$

Define a new Lyapunov function

$$V_2 = V_1 + \frac{1}{2}q_2^T q_2 \tag{24}$$

By time-differentiating V_2 yield

$$\dot{V}_2 = -q_1^T \tilde{e}_v - k_1 q_1^T(q_1 - q_2) - q_2^T \tilde{e}_v - k_1 q_2^T(q_1 - q_2) - \frac{1}{k_1}q_2^T(z_2 + F + \hat{D}_f) \tag{25}$$

Since the vector F serves as the actual control input, it is designed to satisfy

$$F = -z_2 - \hat{D}_f + k_2 q_2 \tag{26}$$

where k_2 is a positive constant.

By substituting (26) into (25), $\dot{V}_2$ will be

$$\dot{V}_2 = -\left(q_1^T + q_2^T\right)\tilde{e}_v - q_1^T k_1 q_1 - q_2^T\left(\frac{k_2}{k_1} - k_1\right)q_2. \tag{27}$$

Given the finite-time convergence property of the HOSMO, there exists a finite time constant $T > 0$ such that $\tilde{e}_v(t) = 0$ and $\tilde{D}_f(t) = 0$ for all $t \geq T$. During the transient phase $t \in [0, T)$, the observer estimation errors remain bounded and can be treated as bounded disturbances in the Lyapunov analysis of the controller. By the input-to-state stability (ISS) property of the closed-loop system, these bounded observer errors lead to bounded tracking errors. After the observer converges, i.e., for $t \geq T$, the nominal Lyapunov analysis ensures asymptotic convergence of all tracking errors. This cascaded systems approach rigorously justifies the separation between the observer and controller designs [12].

Specifically, since the velocity estimation error $\tilde{e}_v$ achieves finite-time convergence to zero, there exists a finite time constant t_1 such that $\tilde{e}_v(t) = 0$ for all $t \geq t_1$. For $t \geq t_1$, the closed-loop dynamics of $\dot{V}_2$ is governed by the following:

$$\dot{V}_2 = -q_1^T k_1 q_1 - q_2^T\left(\frac{k_2}{k_1} - k_1\right)q_2. \tag{28}$$

By choosing the constants satisfying

$$k_1 > 0, \frac{k_2}{k1} - k_1 > 0$$

We obtain that $\dot{V}_2 \leq 0$. Analysis of Eq. (24) reveals that $\dot{V}_2 = 0$ is satisfied if and only if $q_1 = 0$, $q_2 = 0$. The designed controller is uniformly asymptotically stable, and the system errors q_1, q_2 will converge to zero. According to Eq. (20), when the terms $q_1 = q_2 = 0$, it follows that $e_v = 0$.

3.3 Inner-Loop Attitude Controller

In the previous subsection, we derived the control input u_f required for the quadrotor to complete the visual servoing task. The objective of the inner-loop attitude controller is to design the control input τ that

enables the quadrotor to adjust its attitude angles and track the desired force vector. This section will focus on torque controller design based on unknown parameter and disturbance.

It follows from Assumption 2 that there exist positive constant B such that: $||\dot{d}_\tau|| \leq B$. A sliding-mode disturbance observer is designed for the external disturbances acting on the quadrotor during rotational motion. Let $\hat{d}_\tau$ represent the estimated value of d_τ, $\tilde{d}_\tau = \hat{d}_\tau - d_\tau$ denote the estimated error, and to observe the system disturbance signals, the following auxiliary variables are designed as

$$e_{rev} = z_{rev} - J\Omega, \tag{29}$$

$$\dot{z}_{rev} = -\Omega \times J\Omega + \tau + \hat{d}_\tau - \lambda_a e^{\delta_1} - \lambda_b e^{\delta_2}, \tag{30}$$

where λ_a and λ_b are positive constant, $0 < \delta_1 < 1$, $\delta_2 > 1$.

The sliding mode surface and the reaching law are designed to make the error converge to zero in a finite time, which is shown as follows:

$$\tilde{d}_\tau = \hat{d}_\tau - d_\tau = \dot{e}_{rev} + \lambda_a e^{\delta_1} + \lambda_b e^{\delta_2}, \tag{31}$$

$$\dot{\tilde{d}}_\tau = -c_1 \tilde{d}_\tau - c_2 sign(\tilde{d}_\tau). \tag{32}$$

A new Lyapunov candidate function is defined as

$$L_1 = \frac{1}{2} \tilde{d}_\tau^T \tilde{d}_\tau. \tag{33}$$

The time derivative of L_1 yields

$$\begin{aligned} \dot{L}_1 &= \tilde{d}_\tau^T (-c_1 \tilde{d}_\tau - c_2 sign(\tilde{d}_\tau) - \dot{d}_\tau) \\ &= -\tilde{d}_\tau^T c_1 \tilde{d}_\tau - c_2 \tilde{d}_{sum}^\tau - \tilde{d}_\tau^T \dot{d}_\tau \\ &\leq -\tilde{d}_\tau^T c_1 \tilde{d}_\tau - (c_2 - B) \tilde{d}_{sum}^\tau. \end{aligned} \tag{34}$$

Vector $d_{sum}^\tau = (|d_{\tau 1}| + |d_{\tau 2}| + |d_{\tau 3}|) \geq 0$, $d_{\tau i}$ $(i = 1, 2, 3)$ represents the i-th component value of vector d_τ. With $c_1 > 0$, $c_2 > B$, hence $\dot{L}_1 \leq 0$. The disturbance estimation error is guaranteed to converge to zero. We now design a torque controller for the quadrotor.

Let $\Omega_d = [\Omega_{1d} \ \ \Omega_{2d} \ \ \Omega_{3d}]$ denote the desired angular velocity expressed in the body frame. Recall the expression of F in Eq. (15), the control input admits the following reformulation:

$$F = \frac{-RfE_3}{m} + ge_3 \tag{35}$$

By time-differentiating (26), we can obtain the desired attitude matrix $R_d = (r_{1d}, r_{2d}, r_{3d})$ as

$$\begin{cases} r_{1,d} = r_{2,d} \times r_{3,d} \\ r_{2,d} = \frac{r_{3,d} \times \Psi}{\|r_{3,d} \times \Psi\|} \\ r_{3,d} = -\frac{F}{\|F\|} \end{cases} \tag{36}$$

where the $\Psi = [\cos\psi_d, \sin\psi_d, 0]^T$, ψ_d denotes the desired yaw angle.

For a time-varying R_d, the desired angular velocity and its derivative can be derived as

$$sk(\Omega_d) = R_d^T \dot{R}_d, \quad sk(\dot{\Omega}_d) = sk(\Omega_d)^T R_d^T \dot{R}_d + R_d^T \ddot{R}_d.$$

By solving for $\dot{F}$ and $\ddot{F}$, the desired angular velocity Ω_d and angular acceleration $\dot{\Omega}_d$ can be obtained.

Note that the intermediary input F designed in Eq. (26) is twice differentiable, ensuring the well-defined nature of Ω_d and $\dot{\Omega}_d$. However, higher-order derivatives of F cannot be reliably computed through difference methods due to image noise and external disturbances. To address this, we propose employing an HOSM differentiator to estimate $\dot{F}$ and $\ddot{F}$. Let $\hat{F}$ represent the estimated value of F, define the estimated error $\tilde{F} = \hat{F} - F$.

$$\begin{aligned}
\dot{x}_1 &= k_a \|\widetilde{F}\|^{\frac{2}{3}} \operatorname{sign}(\widetilde{F}) + x_2 \\
\dot{x}_2 &= k_b \|\widetilde{F}\|^{\frac{1}{3}} \operatorname{sign}(\widetilde{F}) + x_3 \\
\dot{x}_3 &= k_c \operatorname{sign}(\widetilde{F})
\end{aligned} \tag{37}$$

The positive constant gains $k_a > 0$, $k_b > 0$, $k_c > 0$ are selected, with the estimator outputs $\hat{F} = x_1$, $\dot{\hat{F}} = x_2$, and $\ddot{\hat{F}} = x_3$. As established in [28], finite-time stability guarantees $F = \hat{F}$, $\dot{F} = \dot{\hat{F}}$, and $\ddot{F} = \ddot{\hat{F}}$ within finite time. Consequently, both the desired angular velocity Ω_d and its derivative $\dot{\Omega}_d$ are obtained.

The angular velocity tracking error is defined as follows:

$$e_\Omega = \Omega - \Omega_d. \tag{38}$$

A new Lyapunov function is selected as follows:

$$V_3 = L_1 + \frac{1}{2} e_\Omega^T J e_\Omega. \tag{39}$$

Differentiating this function yields:

$$\dot{V}_3 = \dot{L}_1 + e_\Omega^{\mathrm{T}} (-\Omega \times J\Omega + \tau + d_\tau - J\dot{\Omega}_d) \tag{40}$$

The torque controller is designed as:

$$\tau = \Omega \times J\Omega + J\dot{\Omega}_d - \hat{d}_\tau - k_3 e_\Omega. \tag{41}$$

Substituting Eq. (41) into (40) yields

$$\begin{aligned}
\dot{V}_3 &= -\tilde{d}_\tau^T c_1 \tilde{d}_\tau - c_2 \tilde{d}_{sum}^\tau - \tilde{d}_\tau^T \dot{d}_\tau - e_\Omega^T k_3 e_\Omega + e_\Omega^T \tilde{d}_\tau \\
&\le -\tilde{d}_\tau^T \left(c_1 - \frac{1}{2} \right) \tilde{d}_\tau - (c_2 - \xi) \tilde{d}_{sum}^\tau - e_\Omega^T \left(k_3 - \frac{1}{2} \right) e_\Omega
\end{aligned} \tag{42}$$

Then, the control parameters c_1, c_2 and k_3 are selected to satisfy the following relationship:

$$\begin{cases}
c_1 > \dfrac{1}{2}, \\
c_2 > B, \\
k_3 > \dfrac{1}{2}.
\end{cases} \tag{43}$$

Substituting Eq. (43) into (42), we obtain that $\dot{V}_3 \le 0$, the closed-loop system exhibits uniform asymptotic stability, satisfying that $\lim_{t \to \infty} (e_\Omega) = 0$.

Remark 3: The desired angular velocity is derived from Eq. (15). Unlike conventional PD controllers that track Ω_d, the proposed torque controller directly incorporates the quadrotor's dynamic model characteristics. This approach provides more accurate dynamic compensation while addressing model uncertainties and environmental variations, resulting in improved robustness and real-time performance compared to traditional methods. Moreover, since conventional differentiation methods introduce unacceptable errors in computing higher-order derivatives of F, a high-order sliding mode observer is employed to accurately estimate $\dot{F}$ and $\ddot{F}$, thereby enhancing the controller's precision.

4 Experimental Results and Discussion

This section experimentally validates the proposed observer-based IBVS controller under conditions of unknown ball velocity and external disturbances. Two representative disturbance scenarios are investigated: (i) constant disturbances to evaluate steady-state disturbance rejection capability, and (ii) time-varying disturbances to validate robustness under realistic environmental conditions such as wind fluctuations. Correspondingly, two distinct target motion trajectories are considered: (i) a curvilinear-to-linear ground trajectory (xy-plane motion with $v_z = 0$), modeling receiving-and-kicking maneuvers; and (ii) a parabolic aerial trajectory (xyz-axis motion), representing long-kicking situations.

4.1 Hardware Platform

A comprehensive hardware setup was implemented to experimentally validate the proposed disturbance-observer-based visual control strategy. The platform was designed to meet the specific demands of dynamic soccer ball tracking, emphasizing real-time image processing, robust communication, and precise flight control. The core components are detailed in the following subsections.

For visual perception, a downward-facing global shutter camera was mounted on a vibration-damped plate to ensure clear image capture during flight. The critical task of real-time image processing, feature extraction, and executing the high-order sliding mode observer and backstepping controller was handled by a Raspberry Pi 4B single-board computer, which served as the onboard computational unit. This choice provides a compact, low-power solution capable of meeting the processing demands of the vision-based control system. A Pixhawk flight controller managed the low-level attitude stabilization and motor control. The entire system was integrated through the Robot Operating System (ROS) framework, with the signal architecture illustrated in Fig. 3, ensuring robust and synchronized communication between all components.

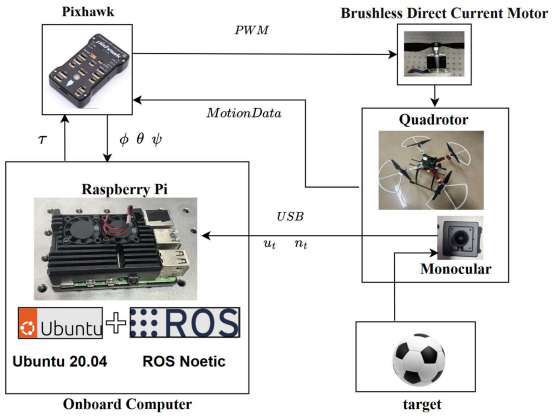


Figure 3: Signal architecture of quadrotor control system.

Physical parameters of the quadrotor were set as: mass $m = 2\,\text{kg}$, gravitational acceleration $g = 9.81\,\text{m/s}^2$, and moment of inertia $J = \text{diag}(0.0745, 0.0745, 0.113)\,\text{kg}\cdot\text{m}^2/\text{rad}^2$. A downward-facing camera with a focal length of 1.8 mm and resolution of 640 × 480 pixels was mounted beneath the quadrotor. The reference image moment and height were specified as $q_d = [0\ \ 0\ \ 2]^T$ and $z_d = 2$. Controller and observer parameters were set as: $k_1 = 2$, $k_2 = 6$, $c_1 = 4$, $c_2 = 12$, $\lambda_a = \lambda_b = 6$, $\delta_1 = 0.5$, $\delta_2 = 2$, $k_a = 10$, $k_b = 20$, $k_c = 30$.

4.2 Image Processing and Real-Time Implementation Strategy

To ensure high-frequency visual servoing on the resource-constrained embedded platform (Raspberry Pi 4B), it is possible to implement a Region of Interest (ROI) extraction strategy based on spatio-temporal continuity.

Vision Algorithm Pipeline: The algorithm initiates with a global search in the first frame to lock onto the target soccer ball. Once the target is identified, subsequent processing is confined to a dynamic ROI window centered on the feature coordinates from the previous frame. Within this localized window, feature extraction is executed through HSV color space segmentation, morphological filtering, and the Hough Circle Transform. To strike an optimal balance between the field of view (FOV) and computational overhead, the image acquisition resolution is configured at 640 × 480 pixels.

Latency and Update Frequency: The vision processing pipeline is optimized using the OpenCV library on the Raspberry Pi, achieving an average per-frame processing time of approximately 18 ms. Coupled with a camera acquisition rate of 30 FPS, the overall update frequency of the visual control loop is maintained at a consistent 30 Hz.

Communication and Architectural Advantages: Communication between the Raspberry Pi and the Pixhawk flight controller is facilitated via a TTL serial port using the MAVLink protocol at a high baud rate of 921,600 bps. Since all visual computations are executed locally on the onboard edge device, the system entirely circumvents the bandwidth fluctuations and multi-second latencies inherent in traditional wireless image transmission. This edge-computing architecture ensures that the total end-to-end latency is strictly controlled within 50 ms, providing the necessary real-time responsiveness for tracking high-dynamic targets in outdoor environments.

4.3 Results and Discussion

4.3.1 Curvilinear-to-Linear Ground Trajectory

The football's velocity is initially set to $\mathbf{v} = [-3sin(0.25\,t), 3cos(0.25\,t) - 3, 0]^\top$ m/s for the first 20 s, then switched to $\mathbf{v} = [2.8, -1.9, 0]^\top$ m/s from 20 to 40 s. The system is subject to constant external disturbances including a linear acceleration disturbance $d_f = [0.6, 0.5, 0.3]^\top\,\text{m/s}^2$ and an angular acceleration disturbance $d_\tau = [0.4, 0.5, 0.6]^\top\,\text{rad/s}^2$ throughout the entire duration. This configuration enables evaluation of the system's dynamic response under changing motion conditions while maintaining persistent disturbance inputs. The initial positions expressed in the initial frame are set to $[3, 2, 3]^\top$ m for the unmanned aerial vehicle (UAV) and $[0, 0, 0]^\top$ m for the soccer ball. To evaluate the performance advantages of the proposed control scheme, the IBVS method without disturbance compensation was employed as a baseline for comparison. The test results are shown in Figs. 4–7.

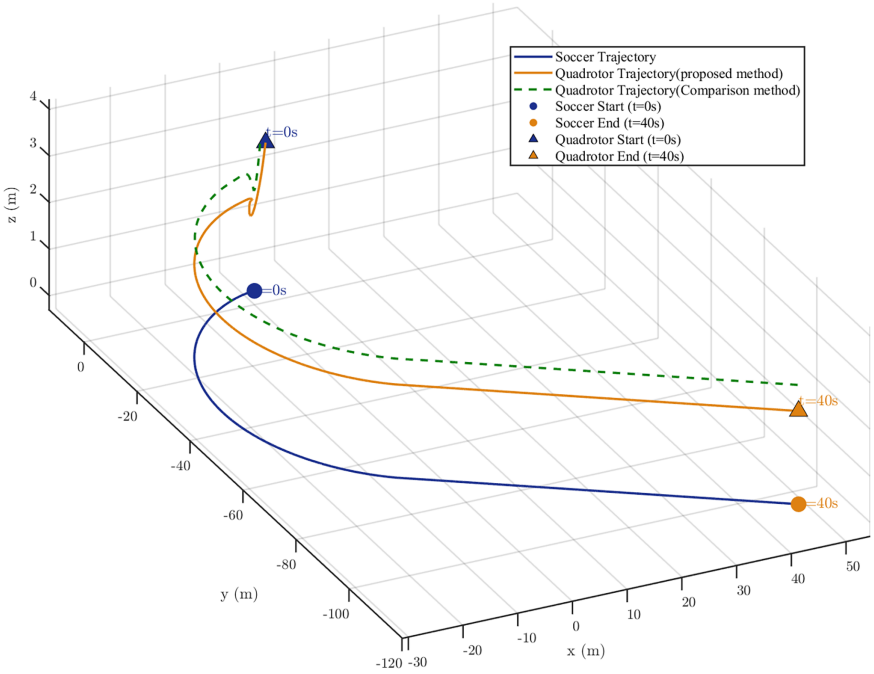


Figure 4: The spatial trajectories of the drone and the soccer ball.

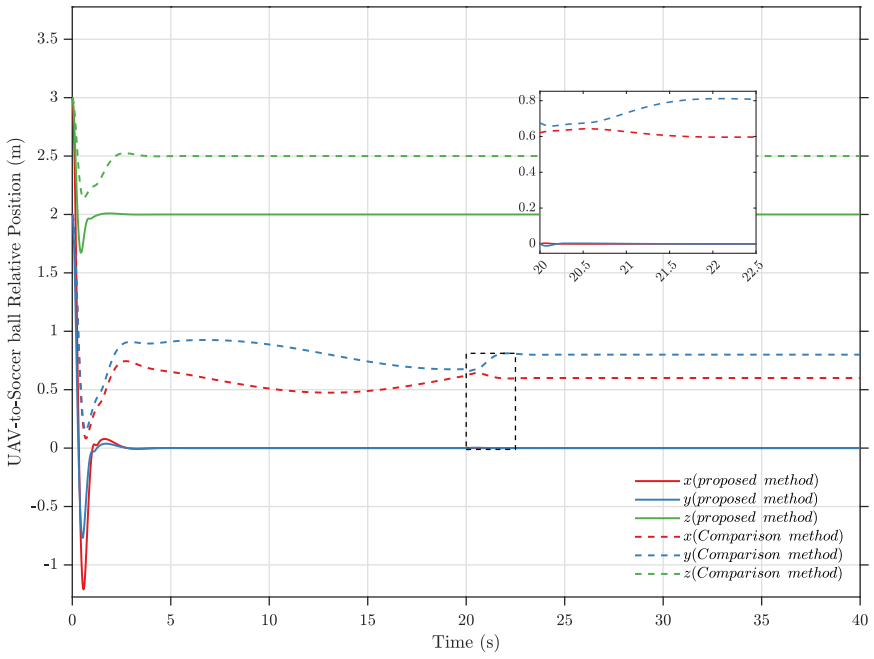


Figure 5: The relative position of the quadrotor with respect to the soccer ball.

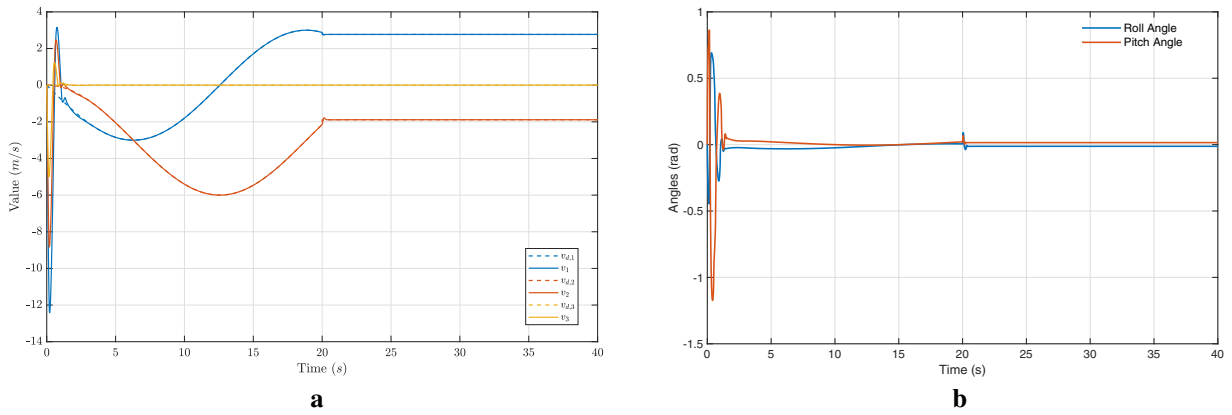


Figure 6: The state variables of the quadrotor. (**a**) The velocity of quadrotor (The dashed line represents the soccer ball's velocity, while the solid line indicates the quadrotor velocity). (**b**) The Euler angles of the quadrotor.

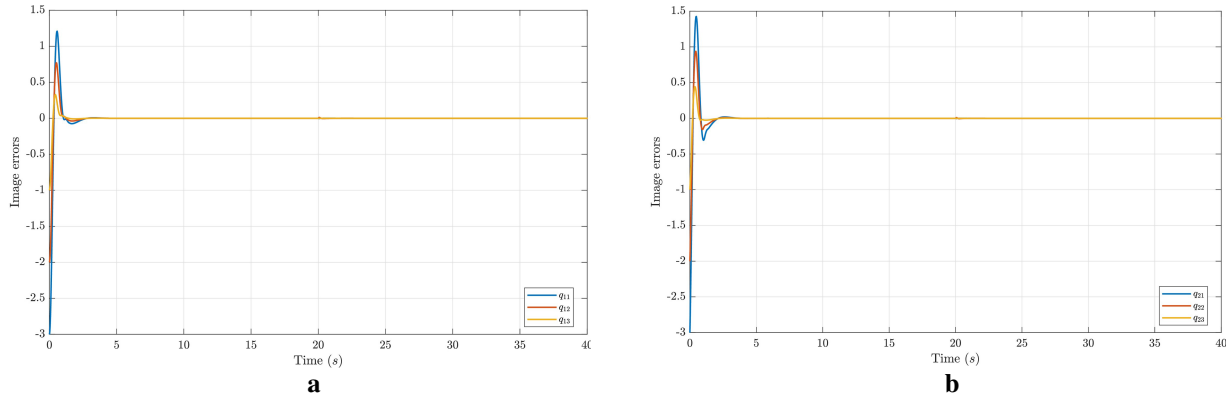


Figure 7: Time evolution of the Image errors. (**a**) Time evolution of q_1. (**b**) Time evolution of q_2.

Figs. 4 and 5 demonstrate that the proposed control algorithm achieves stable tracking of a moving soccer ball in the presence of constant external disturbances and unknown ball velocity, delivering satisfactory tracking performance and accuracy, thereby validating our theoretical results. To illustrate the effectiveness of the proposed approach, it is observed that the proposed method (solid orange line) achieves higher fidelity in tracking the soccer ball trajectory (blue line) compared to the baseline method (dashed green line). Starting from the same initial position, the proposed controller rapidly adjusts the UAV's path to maintain the desired relative pose directly above the moving target. In contrast, the method without disturbance compensation exhibits a significant and persistent offset throughout the flight. As can be seen from Fig. 5, due to the sudden change in the velocity of the soccer ball, a slight chattering occurs in the drone's relative position at $t = 20$ s. However, the amplitude of chattering is very small, and the relative position re-converges to stability within 1.5 s, demonstrating the robustness of the proposed control algorithm. The proposed method (solid lines) successfully drives the relative x and y positions to zero and the z position to the desired altitude of 2.0 m within approximately 4 s. Conversely, the comparison method (dashed lines) fails to eliminate steady-state errors, resulting in constant offsets of approximately 0.5 m in x and 0.8 m in y and a bias in altitude.

As can be observed from Fig. 6a, since the drone is initially far from the soccer ball at $t = 0$ s, it needs to accelerate to converge to the desired position directly above the ball. Upon reaching the target position, the drone successfully tracks the ball's velocity with high precision, maintaining a constant relative position

to the moving ball. Fig. 6b demonstrates that when the soccer ball's velocity changes abruptly, the quadrotor Euler angles exhibit minor oscillations but rapidly regain convergence. This validates both the effectiveness and robustness of our proposed attitude-loop controller. Fig. 7 demonstrates the convergence of image feature errors, which ensures simultaneous convergence of both the UAV's spatial positioning and velocity tracking errors.

4.3.2 Parabolic Aerial Trajectory

The football's velocity is initially set to $\mathbf{v} = [3sin(0.25\,t), 3 - cos(0.25\,t), 2sin(0.314\,t)]^\top$ m/s for the first 20 s, then switched to $\mathbf{v} = [2t, t, 0]^\top$ m/s from 20 to 25 s. The system is subject to external disturbances including a variational disturbance $d_f = [0.2sin\,t, 0.5cos\,t, 0.4sin(1.5\,t)]^\top$ m/s^2 and an angular acceleration disturbance $d_\tau = [0.4sint, 0.5cost, 0.6sin2t]^\top$ rad/s^2 throughout the entire duration. This configuration enables to validate the effectiveness of the proposed algorithm under time-varying disturbances and varying soccer ball depth information. The initial positions expressed in the initial frame are set to $[2, 4, 3]^\top$ m for the unmanned aerial vehicle and $[0, 0, 0]^\top$ m for the soccer ball. To evaluate the superiority of the proposed HOSMO, we introduced a nonlinear Extended State Observer (ESO) as a baseline, which is a widely recognized benchmark for disturbance estimation. The test results are shown in Figs. 8–12.

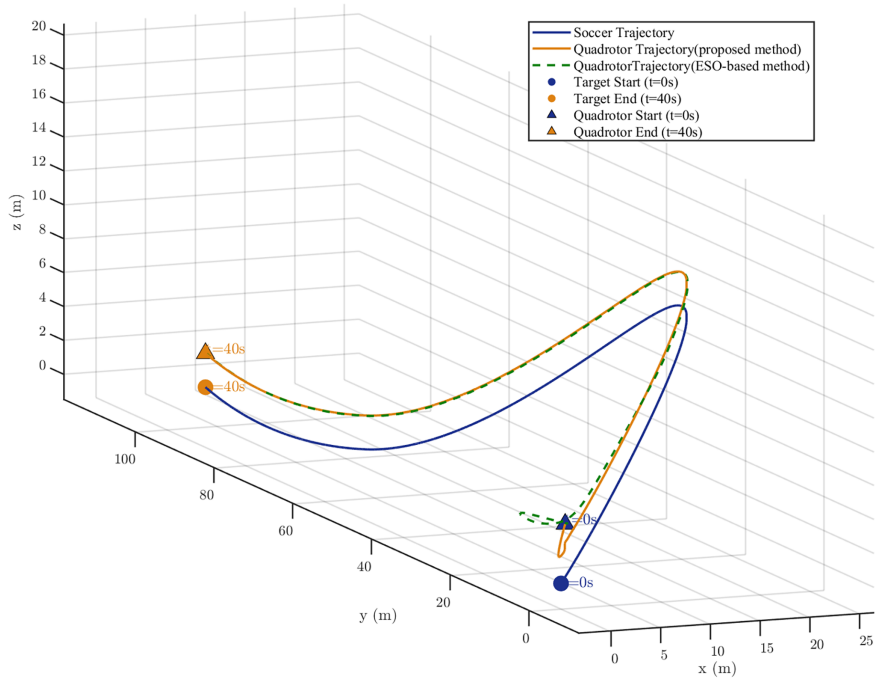


Figure 8: The spatial trajectories of the drone and the soccer ball.

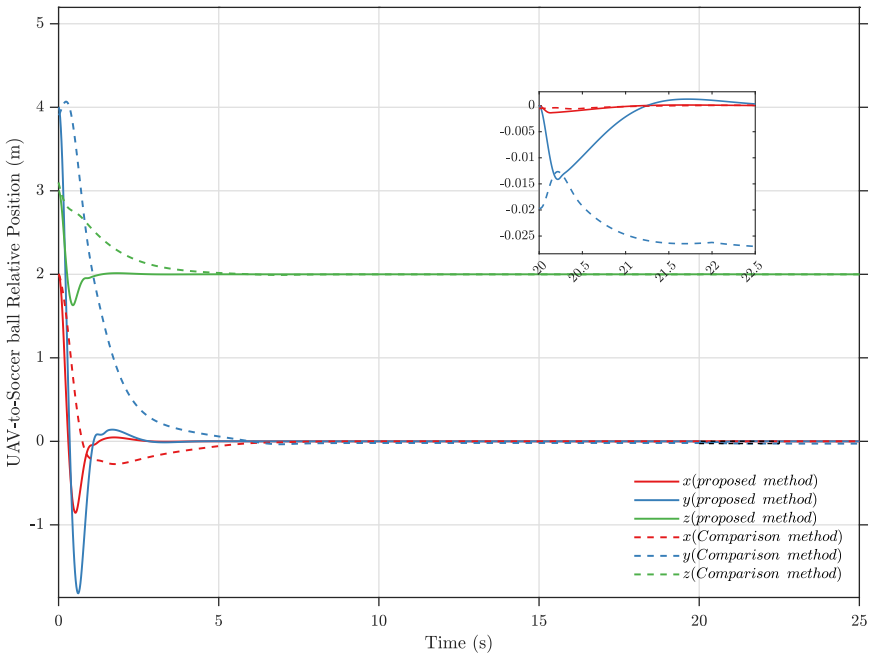
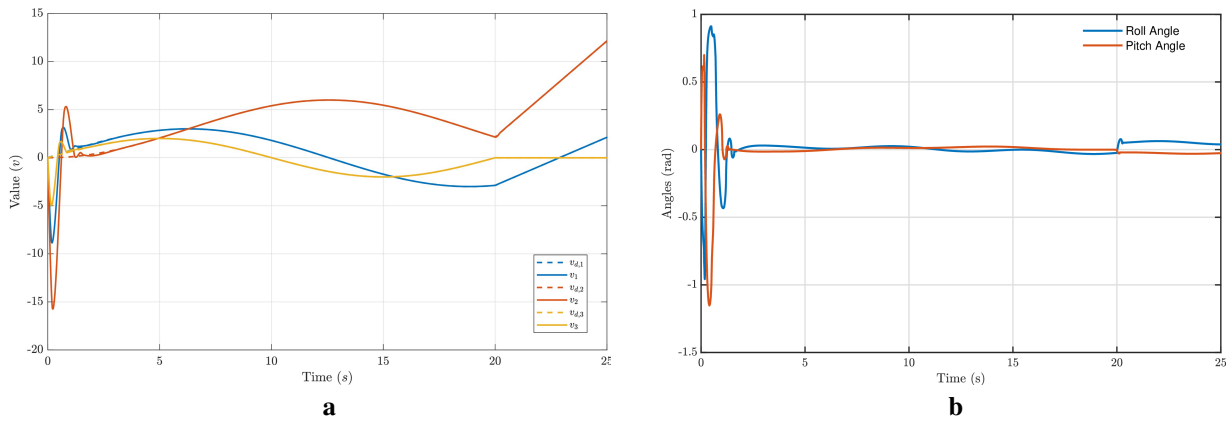


Figure 9: The relative position of the quadrotor with respect to the soccer ball.



a **b**

Figure 10: The state variables of the quadrotor. (**a**) The velocity of quadrotor (The dashed line represents the soccer ball's velocity, while the solid line indicates the quadrotor velocity). (**b**) The Euler angles of the quadrotor.

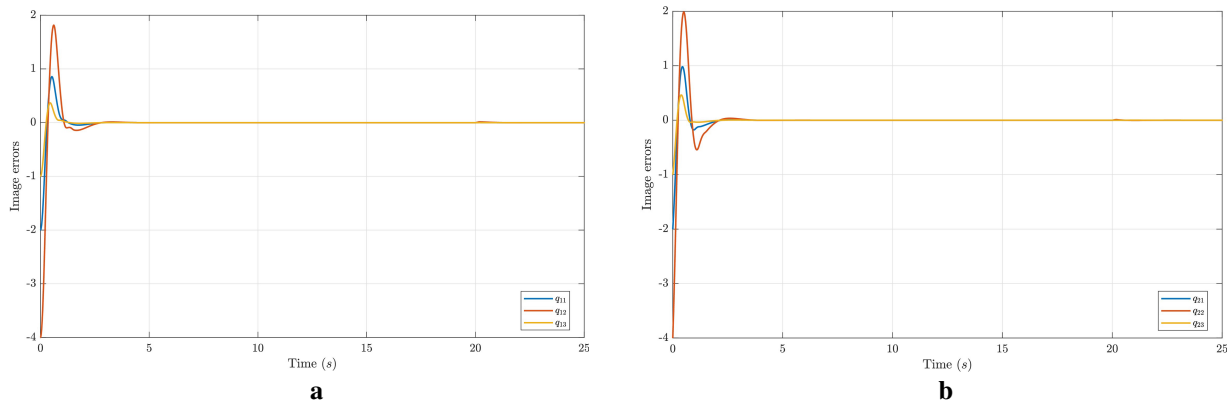


Figure 11: Time evolution of the Image errors. (**a**) Time evolution of q_1. (**b**) Time evolution of q_2.

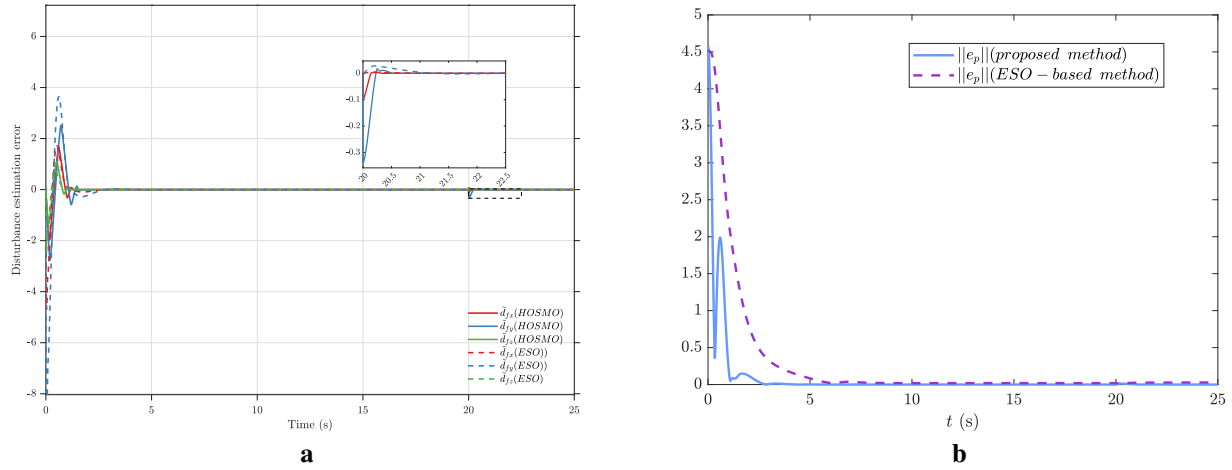


Figure 12: Error convergence situation. (**a**) Position-loop disturbance estimation error (The solid line represents HOSMO, and the dotted line represents ESO). (**b**) Evolution of the Euclidean norm of tracking errors over time.

As shown in Figs. 8 and 9, the proposed control strategy successfully maintains accurate tracking of the motion of the soccer ball despite time-varying external disturbances and variations in ball altitude, demonstrating satisfactory tracking precision. Unlike the conventional ESO, which only provides asymptotic convergence, the proposed HOSMO guarantees finite-time convergence of the estimation errors. This allows the observer to reconstruct the lumped disturbances and target velocities with a significantly superior convergence rate.

Fig. 9 reveals a transient response at t = 20 s following the velocity of the ball discontinuity, where the drone exhibits minor positional chattering. The system demonstrates rapid recovery, achieving restabilization within 2.5 s-evidence of the disturbance rejection capability of the controller. Fig. 10a demonstrates the drone's convergence from an initial offset position at t = 0 s, achieving precise velocity tracking and stable relative positioning above the moving ball. Fig. 10b confirms the attitude controller's robustness through Euler angle dynamics during velocity transitions, with transient oscillations dampening within 1 s. Notably, Fig. 11 establishes the image feature error convergence, which guarantees simultaneous stabilization of both spatial positioning and velocity tracking errors through the visual servoing framework.

As shown in Fig. 12a, the disturbance estimation errors of the HOSMO (solid lines) exhibit a significantly faster convergence rate compared to the ESO (dotted lines). At the initial phase and during the disturbance occurrence at $t = 20$ s, the HOSMO-based observer demonstrates its finite-time stability by reconstructing the unknown perturbations almost instantaneously. Fig. 12b illustrates the evolution of the Euclidean norm of tracking errors $||e_p||$. The proposed method (solid blue line) achieves a rapid settling time of approximately 4 s, while the ESO-based method (dashed purple line) requires nearly 7 s to reach a comparable steady state.

5 Conclusion and Discussion

This research delivers a robust vision-based solution for quadrotor tracking of soccer balls during high-dynamic matches, directly addressing sports filming challenges. The virtual camera projection enables consistent ball imaging regardless of drone attitude changes, while moment features decouple depth and position sensing—ensuring the target remains centered during free kicks or aerial passes. The integrated control architecture compensates for unpredictable ball motions and environmental disturbances through finite-time disturbance observers and adaptive backstepping, achieving sub-meter tracking precision (as quantified in Table 1) even in parabolic flight tests with time-varying wind interference. Crucially, the system operates using only desired flight height and yaw orientation, simplifying deployment in stadium environments.

Table 1: Quantitative comparison of tracking performance (Unit: m).

RMSE	Proposed Method		Comparison Method	
	Experiment 1	Experiment 2	Experiment 1	Experiment 2
x (m)	0.0787	0.1780	0.5864	0.2274
y (m)	0.0501	0.3876	0.8068	0.7586
z (m)	0.0008	0.0035	0.4977	0.0631

Future work will focus on computational optimization for onboard vision systems and multi-drone cooperative filming strategies to expand coverage for large-scale sporting events. Additionally, a comparative evaluation of resource utilization (e.g., memory footprint, power consumption) across different edge platforms will be conducted to further validate the lightweight deployment capability.

Acknowledgement: Not applicable.

Funding Statement: This research was supported by the School-Level Fund Project of Guangdong University of Science and Technology (No. GKY-2025BSQDW-15).

Author Contributions: The authors confirm contribution to the paper as follows: Conceptualization, Yao-Bo Long; methodology, Yao-Bo Long, Yu-Ke Ouyang; software, Yu-Ke Ouyang; validation, Yu-Ke Ouyang, Bo Zhuang, Ao-Qi Liu; formal analysis, Yao-Bo Long, Yu-Ke Ouyang; investigation, Yao-Bo Long; resources, Yao-Bo Long; data curation, Yao-Bo Long; writing—original draft preparation, Yu-Ke Ouyang; writing—review and editing, Ao-Qi Liu; visualization, Bo Zhuang; supervision, Ao-Qi Liu; project administration, Yao-Bo Long. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: Data available within the article.

Ethics Approval: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Ding R, Zhu Y, Feng X, Chai B, Ji R. A multi-objective assignment algorithm for UAVs in multiple task with different workloads. In: Proceedings of the 2023 Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC); 2023 Apr 14–16; Dalian, China. p. 354–9.
2. Hulens D, Goedemé T. Autonomous flying cameraman with embedded person detection and tracking while applying cinematographic rules. In: Proceedings of the 2017 14th Conference on Computer and Robot Vision (CRV); 2017 May 16–19; Edmonton, AB, Canada. p. 56–63.
3. Pueyo P, Dendarieta J, Montijano E, Murillo A, Schwager M. CineMPC: a fully autonomous drone cinematography system incorporating zoom, focus, pose, and scene composition. IEEE Trans Robot. 2024;40:1740–57.
4. Mademlis I, Mygdalis V, Nikolaidis N, Montagnuolo M, Negro F, Messina A, et al. High-level multiple-UAV cinematography tools for covering outdoor events. IEEE Trans Broadcast. 2019;65(3):627–35. doi:10.1109/tbc.2019.2892585.
5. Lin J, Miao Z, Wang Y, Wang H, Wang X, Fierro R. Vision-based safety-critical landing control of quadrotors with external uncertainties and collision avoidance. IEEE Trans Control Syst Technol. 2024;32(4):1310–22. doi:10.1109/tcst.2024.3363372.
6. Chen Y, Wu Y, Lan L, Zhong H, Miao Z, Zhang H, et al. Dynamic target tracking of unmanned aerial vehicles under unpredictable disturbances. Engineering. 2024;35(5):74–85. doi:10.1016/j.eng.2023.05.017.
7. Janabi-Sharifi F, Deng L, Wilson W. Comparison of basic visual servoing methods. IEEE/ASME Trans Mechatron. 2011;16(5):967–83. doi:10.1109/tmech.2010.2063710.
8. Lu M, Chen H, Lu P. Perception and avoidance of multiple small fast moving objects for quadrotors with only low-cost RGBD camera. IEEE Robot Autom Lett. 2022;7(4):11657–64. doi:10.1109/lra.2022.3205114.
9. Guo D, Leang K. Image-based estimation, planning, and control for high-speed flying through multiple openings. Int J Robot Res. 2020;39(9):1122–37. doi:10.1177/0278364920921943.
10. 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. doi:10.1109/tie.2019.2917420.
11. 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. doi:10.1109/tmech.2016.2639531.
12. Lin J, Wang Y, Miao Z, Fan S, Wang H. Robust observer-based visual servo control for quadrotors tracking unknown moving targets. IEEE/ASME Trans Mechatron. 2023;28(3):1268–79. doi:10.1109/tmech.2022.3217034.
13. Asl H, Yoon J. Bounded-input control of the quadrotor unmanned aerial vehicle: a vision-based approach. Asian J Control. 2017;19(3):840–55. doi:10.1002/asjc.1420.
14. Asl H. Robust vision-based tracking control of VTOL unmanned aerial vehicles. Automatica. 2019;107:425–32.
15. 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. doi:10.1109/tsmc.2017.2720173.
16. Li J, Xie H, Low K, 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. doi:10.1109/lra.2021.3101878.
17. Chen Y, Wu Y, Zhang Z, Miao Z, Zhong H, Zhang H, et al. Image-based visual servoing of unmanned aerial manipulators for tracking and grasping a moving target. IEEE Trans Ind Inform. 2023;19:8889–99.
18. Xie H, Low K, He Z. Adaptive visual servoing of unmanned aerial vehicles in GPS-denied environments. IEEE/ASME Trans Mechatron. 2017;22(6):2554–63. doi:10.1109/tmech.2017.2755669.
19. Xie H, Lynch A, Low K, Mao S. Adaptive output-feedback image-based visual servoing for quadrotor unmanned aerial vehicles. IEEE Trans Control Syst Technol. 2020;28(3):1034–41. doi:10.1109/tcst.2019.2892034.
20. 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. doi:10.1109/tmech.2021.3053267.
21. Jabbari Asl H, Yoon J. Robust image-based control of the quadrotor unmanned aerial vehicle. Nonlinear Dyn. 2016;85(3):2035–48. doi:10.1007/s11071-016-2813-2.
22. Chen W, Yang J, Guo L, Li S. Disturbance-observer-based control and related methods—an overview. IEEE Trans Ind Electron. 2015;63(2):1083–95. doi:10.1109/tie.2015.2478397.

23. Hua H, Fang Y, Zhang X, Lu B. A novel robust observer-based nonlinear trajectory tracking control strategy for quadrotors. IEEE Trans Control Syst Technol. 2021;29(5):1952–63. doi:10.1109/tcst.2020.3024805.
24. Li B, Gong W, Yang Y, Xiao B, Ran D. Appointed fixed time observer-based sliding mode control for a quadrotor UAV under external disturbances. IEEE Trans Aerosp Electron Syst. 2022;58(1):290–303. doi:10.1109/taes.2021.3101562.
25. Levant A. Higher-order sliding modes, differentiation and output-feedback control. Int J Control. 2003;76:924–41. doi:10.1080/0020717031000099029.
26. Zhao Z, Cao D, Yang J, Wang H. High-order sliding mode observer-based trajectory tracking control for a quadrotor UAV with uncertain dynamics. Nonlinear Dyn. 2020;102:2583–96. doi:10.1007/s11071-020-06050-2.
27. Hu S, Wang Q, Wang F, Li Y. Finite-time dynamic visual servo control for quadrotor tracking unknown motion target. Nonlinear Dyn. 2025;113(7):6959–77. doi:10.1007/s11071-024-10713-9.
28. Moreno J. A Lyapunov approach to output feedback control using second-order sliding modes. IMA J Math Control Inf. 2012;29(3):291–308. doi:10.1093/imamci/dnr036.