Robust Vision-Based Control of a Rotorcraft UAV for Uncooperative Target Tracking

Shijie Zhang *, Xiangtian Zhao [ORCID] **and Botian Zhou**

Research Center of Satellite Technology, Harbin Institute of Technology, Harbin 150080, China;
zhaoxiangtian@hit.edu.cn (X.Z.); btzhou.hit@gmail.com (B.Z.)
* Correspondence: sjzhang@hit.edu.cn

Abstract: This paper investigates the problem of using an unmanned aerial vehicle (UAV) to track and hover above an uncooperative target, such as an unvisited area or an object that is newly discovered. A vision-based strategy integrating the metrology and the control is employed to achieve target tracking and hovering observation. First, by introducing a virtual camera frame, the reprojected image features can change independently of the rotational motion of the vehicle. The image centroid and an optimal observation area on the virtual image plane are exploited to regulate the relative horizontal and vertical distance. Then, the optic flow and gyro measurements are utilized to estimate the relative UAV-to-target velocity. Further, a gain-switching proportional-derivative (PD) control scheme is proposed to compensate for the external interference and model uncertainties. The closed-loop system is proven to be exponentially stable, based on the Lyapunov method. Finally, simulation results are presented to demonstrate the effectiveness of the proposed vision-based strategy in both hovering and tracking scenarios.

Keywords: UAV; vision-based control; virtual image plane; uncooperative target tracking; gain-switching control

1. Introduction

Unmanned aerial vehicles (UAVs) have received growing interest, due to their advantages of vertical take off and landing, rapid maneuverability, and low cost. With improvements in sensing devices, batteries, materials, and other technologies, UAVs have sufficient payload and flight endurance, supporting many applications such as transportation, real-time monitoring, search and rescue, and security and surveillance [1–4]. There have been a variety of studies related to missions using autonomous hovering and tracking technologies [5–7]. In [5], a finite-time controller was proposed to drive a quadrotor hovering above a target with a limited duration. In [6], a novel fuzzy PID-type iterative learning control was developed for trajectory tracking of a quadrotor under the effects of external disturbances and uncertain factors of the system. The problem of energy-efficient path planning and simultaneously anticipating disturbances has been addressed in [7]. However, the available studies have mainly focused on hovering above or tracking a target with a known trajectory and definite position and velocity information. Tracking passive, uncooperative, or even unknown targets, such as vehicles in traffic accidents and fire areas in groves or forests, is still a challenging problem for UAV control. These cases usually occur suddenly and detailed geometric information (e.g., dimensions and size) of the target is not available. On the other hand, rough information (e.g., shape and structure) or a simplified model of the target can be stored onboard, which can help the vehicle to identify and lock on to the target.

A large amount of state of the art sensors has been equipped on the airborne platform, and multi-sensor fusion is a perspective trend in UAV navigation and control [8]. Inertial measurement

units (IMUs) and Global Positioning Systems (GPS) are commonly mounted on UAVs to provide position, orientation, and velocity information [9]. However, GPS cannot determine the relative position between the UAV and a target, unless the target's position information is accurately known. Moreover, the accuracy of GPS is unreliable and ineffective in the presence of obstacles or external disturbances, such as urban canyon effects or electromagnetic interference [10,11]. On the other hand, a large number of UAVs, especially small-scale quadrotors, are equipped with cameras which can provide adequate visual information and act as an alternative option to GPS. The notion of introducing visual information into the control system is referred to as visual servoing [12]. According to different feedback information, visual servoing can be classified into two approaches: position-based visual servoing (PBVS) and image-based visual servoing (IBVS).

PBVS methods rely on relative pose estimation techniques, including the PnP algorithm, epipolar geometry, ICP algorithm, or the fusion of vision and other sensors. In [13], a monocular semi-direct visual odometry algorithm was applied to obtain position information and an adaptive sliding mode controller based on the backstepping technique was presented for quadrotor tracking control. In [14], visual feedback was exploited to estimate the position and attitude of a UAV based on the triangulation method, and a classical PID control was designed for indoor UAV flight. In [15], a vision-based control scheme was proposed for a quadrotor, in order to track a moving target. To compensate for the image dynamic uncertainties, a neural network controller based on a radial basis function was introduced into the closed-loop system. However, the methods mentioned above require a priori knowledge of the target's geometric model, which does not apply to a newly discovered uncooperative target. Further, 3D reconstruction based on generic features (instead of specific artificial markers) is usually a challenging task, and the sophisticated image processing involved requires long computation times, which may result in delays in the actuation. The IBVS approach, on the other hand, allows for designing a control law directly in the 2D image space and omits the relative pose estimation process. Thus, the geometric information of the target is not required in this approach. The applications of IBVS for UAV control are largely found in recent studies, in which image features have been used as the position cue. To decouple the rotational motion of the vehicle from the change in the image features, several methods have been adopted in IBVS control of a UAV, including spherical projection [16,17], homography-based methods [18,19], virtual spring methods [20,21], and virtual camera methods [22,23]. It is worth noting that the reference image (or image trajectory) is usually designated ahead of the control process, which is typically referred to as the "teach-by-showing" framework. However, this process may not be practical in some specific missions, such as when a UAV is deployed to track a newly discovered target [24].

When tracking an uncooperative target without an accurate geometric model, one effective solution is Simultaneous Localization and Mapping (SLAM). SLAM is a real-time process by which a vehicle computes its localization and simultaneously builds a map of an unknown environment [25]. SLAM techniques are often employed to provide the pose estimation for the PBVS control of the UAV. In [26], a novel estimator based on an Extended Kalman Filter was proposed for multi-sensor fusion, including AHRS, GPS, and a monocular camera. Moreover, a visual SLAM-based scheme has been adopted to provide an accurate estimation of the translational position and velocity, as described in [27]. In [28], a strategy based on visual-inertial SLAM was employed for collision avoidance in a GPS-denied environment. In general, visual SLAM requires long computation times and large amounts of onboard resources, thus typically struggling to provide the rapid pose estimation needed by the UAV guidance system.

This paper presents a simple vision-based strategy to track an uncooperative target which is newly discovered by the UAV. The proposed strategy demonstrates a strict interaction between vision-based metrology and the resulting control. It is assumed that the target's image is acquired and identified by matching the visual data extracted from the monocular camera with similar information stored on board. Then, a set of image features are well-defined by the image processing unit for the following process. Image centroid is employed for the metrology in the horizontal direction, in spite of the fact

that the geometric model is unknown to the UAV. To eliminate the coupling effect of the UAV rotational motion and the change in the image features, we introduce a virtual camera frame. After reprojecting the image features onto the virtual image plane, we can directly estimate the relative horizontal distance between the UAV and the target, based on the perspective projection model. At the same time, the flying height of the UAV can be adjusted by specifying an optimal observation area on the virtual image plane. Motivated by the work in [29], the relative velocity is estimated by a Partial Velocity Evaluation method, which is reliable and computationally effective. Robust methods for the optic flow measurement have been reported in [30,31]. Considering the deviation of the measured mass and inertia matrix from the true values, as well as the disturbances resulting from the wind and other external factors, traditional PID algorithms may not be competent in the UAV controller design, as described in [32,33]. To compensate for the model and disturbance uncertainties and improve robust performance of the system, a gain-switching proportional-derivative (PD) control scheme is proposed in this paper. In this control strategy, the stability and robustness properties of the UAV nonlinear system are guaranteed based on the Lyapunov stability theory. Unlike the sliding mode control, which can be found in [34,35], the transient behavior of the system can be characterized analytically.

The remaining parts of this paper are organized as follows. The problem formulation is given in Section 2, the visual measurement for uncooperative target tracking is discussed thoroughly in Section 3, the gain switching PD controllers are designed in Section 4, the corresponding simulation results are depicted in Section 5 and, finally, the paper is concluded in Section 6.

2. Problem Formulation

The problem addressed in this paper corresponds to UAV tracking and hovering scenarios in which the target is newly discovered. The target, which is referred to as "uncooperative", lacks accurate dimension/size information and real-time communication with the vehicle. It can be identified by the rough information of generic features stored onboard, such as shape and structure. The UAV studied here is a quadrotor equipped with multiple sensors, including an IMU, an ultrasonic sensor, and a monocular camera. A vision-based strategy integrating the metrology and the resulting control is employed to achieve target tracking and hovering observation. The quadrotor is a typical underactuated mechanism with more degrees of freedom than actuations. To analyze the tracking maneuvers, we first define several reference coordinate frames and present the equations of the motion of the quadrotor, then give an overall control framework of tracking and hovering observation .

2.1. Quadrotor Model

The quadrotor considered in this paper consists of a rigid cross-frame equipped with four rotors, as shown in Figure 1. The two rotors on the diagonal rotate in the same direction, while the adjacent rotors rotate in opposite directions. Six degrees of freedom of the quadrotor's position and attitude can be achieved by adjusting the rotation speed of the four motors. Two coordinate frames are introduced to describe the equations of motion of the quadrotor. An inertial frame $\mathcal{I}$ is fixed to some point O_i on the earth, with a basis $\{X_i, Y_i, Z_i\}$, whose elements are oriented north, east, and down, respectively. A body-fixed frame $\mathcal{B}$ is assumed to be attached to the center of mass O_b of the quadrotor. The unit vectors of the body-fixed frame are represented by $\{X_b, Y_b, Z_b\}$, which are oriented forward, right, and down, respectively.
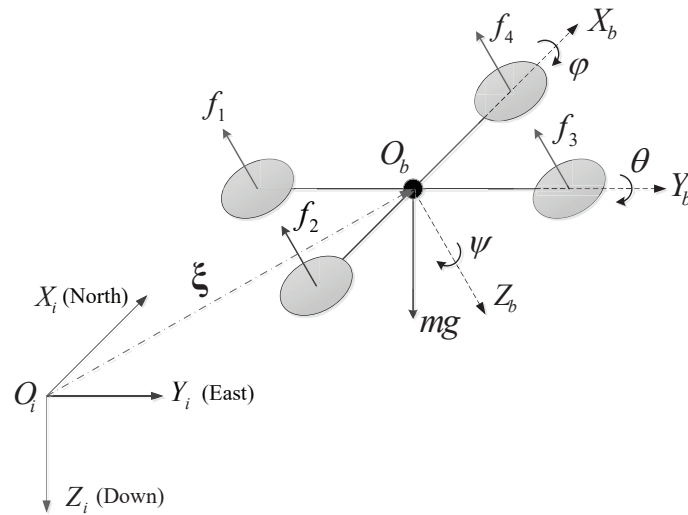
Figure 1. Quadrotor model and frame definitions.

Consider a quadrotor with mass m and inertial matrix $J \in \mathbb{R}^{3\times3}$. The translational dynamics in the inertial frame and the rotational dynamics in the body frame are given as follows.

For the translational dynamics,

$$\begin{cases} \dot{\xi} = v \\ m\dot{v} = -u_1 R E_3 + F_g + F_d \end{cases},$$ (1)

and for the rotational dynamics,

$$\begin{cases} \dot{\Phi} = W(\Phi) \cdot \omega \\ J\dot{\omega} = -\omega \times J\omega + \tau + \tau_d \end{cases},$$ (2)

where $\xi = [x, y, z]^T$ and $v = [v_x, v_y, v_z]^T$ are position and linear velocities of the quadrotor, respectively, expressed in the inertial frame; $E_3 = [0, 0, 1]^T$ is the unit vector in the body frame; the attitude of the vehicle, Φ, is given by three Euler angles φ, θ, and ψ denoting the roll, pitch, and yaw, respectively; and $\omega = [\omega_x, \omega_y, \omega_z]^T$ is the quadrotor's angular velocity expressed in the body frame. The corresponding rotation matrix from the body frame to the inertial frame is denoted by R, and the matrix associated with the Euler angles and the angular velocity can be written as

$$W(\Phi) = \begin{bmatrix} 1 & \sin\varphi\tan\theta & \cos\varphi\tan\theta \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi/\cos\theta & \cos\varphi/\cos\theta \end{bmatrix}.$$ (3)

Assumption 1. *If the quadrotor does not perform maneuvers that are too aggressive, the roll and pitch angles will both be very small ($< 15°$). Then, the matrix $W(\Phi)$ can be replaced with a unit matrix, and $\dot{\varphi}$, $\dot{\theta}$, $\dot{\psi}$, $\ddot{\varphi}$, $\ddot{\theta}$, $\ddot{\psi}$ can be regarded as approximately equal to ω_x, ω_y, ω_z, $\dot{\omega}_x$, $\dot{\omega}_y$, $\dot{\omega}_z$, respectively.*

According to the working principle of the quadrotor (see Figure 1), by adjusting the rotational speed of each group of rotors, the quadrotor can generate a thrust force u_1 and torque vectors $\tau = [u_2, u_3, u_4]^T$, which can be described as

$$\begin{cases} u_1 = b(\omega_1^2 + \omega_2^2 + \omega_3^2 + \omega_4^2) \\ u_2 = bl(\omega_4^2 - \omega_2^2) \\ u_3 = bl(\omega_3^2 - \omega_1^2) \\ u_4 = d(\omega_2^2 + \omega_4^2 - \omega_1^2 + \omega_3^2) \end{cases},$$ (4)

where ω_1, ω_2, ω_3, and ω_4 are the speeds of four motors, respectively; b and d represent lift and drag coefficients, respectively; and l is the distance from the center of each rotor to the center of mass of the quadrotor.

The gravity vector of the quadrotor is denoted by F_g. As the quadrotor may be disturbed by the wind and other external factors, some unstructured forces and moments of the translational and rotational dynamics, which are described as F_d and τ_d, are introduced into the system.

Assumption 2. *The external disturbances F_d and τ_d are assumed to be bounded; that is, $\|F_d\| \leq (F_d)_{\max}$ and $\|\tau_d\| \leq (\tau_d)_{\max}$, where $(F_d)_{\max}$ and $(\tau_d)_{\max}$ are positive constants and $\|\cdot\|$ denotes the standard Euclidean vector norm and induced matrix norm.*

Assumption 3. *The mass is $m = \bar{m} + \Delta m$, where $\bar{m}$ and Δm are the nominal and uncertain parts of the mass, respectively. The inertia matrix is $J = \bar{J} + \Delta J$, where $\bar{J}$ and ΔJ are the nominal and uncertain parts of the inertia matrix, respectively. Δm and ΔJ satisfy the inequalities $\|\Delta m\| \leq c_m$ and $\|\Delta J\| \leq c_J$, where c_m and c_J are positive constants.*

2.2. Control Framework of Tracking and Hovering Observation

A flowchart of the proposed vision-based UAV tracking strategy is presented as follows (see Figure 2). During the cruise of the UAV, when a sensitive target is captured and identified by the onboard camera, the tracking and hovering observation process starts. First, a set of feature points on the target are extracted. By using the center point of the feature points and their pixel velocities, we can determine the relative distance and velocity, respectively, between the UAV and target. Then, the relative distance and velocity are input into the controller as error states. Based on the gain-switching PD control, the target centroid is expected to coincide with the center of the image plane, and the target image is required to be within the optimal observation area.
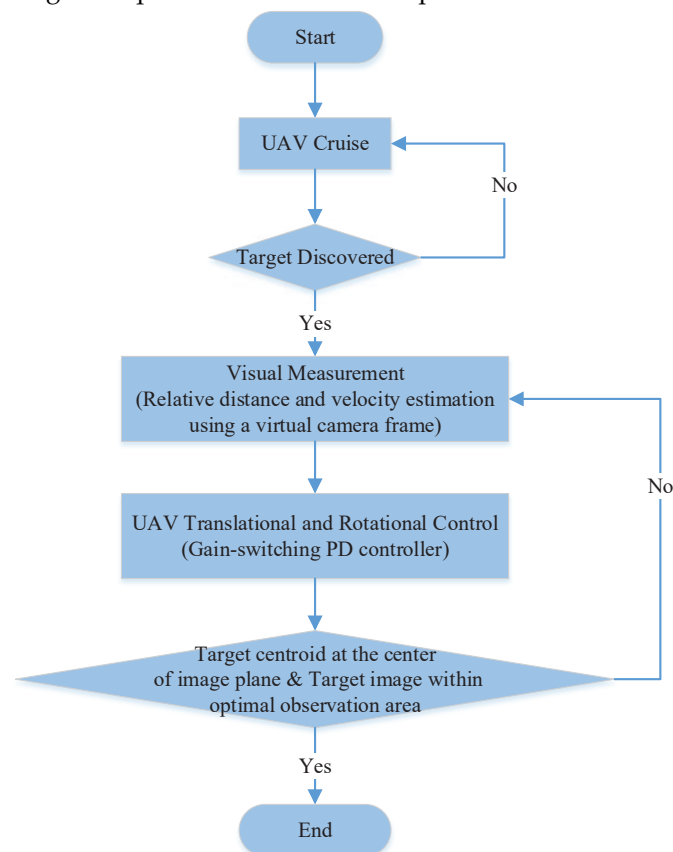

Figure 2. Flowchart of the vision-based unmanned aerial vehicle (UAV) tracking strategy.

3. Visual Measurement Using a Virtual Camera Frame

When a sensitive target is first detected and determined by the camera during the quadrotor flight, its projection on the image plane can be described by a set of parameters, including slope, curvature, area, image centroid, and some parameters related to the shape of the target [36]. The selected target for tracking is assumed to be horizontal to the X_i–Y_i plane of the inertial frame, and its projection onto the image plane is a compact area with large values of shape parameters, such as sphericity or rectangularity.

3.1. Relative Distance Estimation in Horizontal Direction

As the target is uncooperative and cannot send any position information to the quadrotor, the relative distance measurement relies entirely on the vision-based method. However, the target does not have specific markers and its detailed geometry is unknown to the quadrotor, such that it is challenging to use PnP solvers or Template Matching approaches for relative pose acquisition. In this work, we utilize an effective and simple method based on the image error to estimate the relative horizontal distance between the quadrotor and the target. On the other hand, the quadrotor is an underactuated system with only four independently controllable degrees of freedom. Rolling and translational motions along the Y_b axis are coupled, as are pitching and translational motions along the X_b axis, which means that the vehicle will inevitably tilt when maneuvering horizontally. Thus, the image features of the target will change with not only the translational, but also the rotational, motion of the quadrotor, which makes it more complicated to estimate the relative distance and velocity.

Without loss of generality, the camera frame considered in this paper, $\mathcal{C} = \{X_c, Y_c, Z_c\}$, is assumed to coincide with the body frame $\mathcal{B}$. To solve the problem mentioned above, we introduce a virtual camera coordinate frame $\mathcal{V}$, which has the same origin as the frame $\mathcal{B}(\mathcal{C})$. The corresponding virtual image plane X_v-Y_v is always parallel to the X_i–Y_i plane and with the same yaw angle as the frame $\mathcal{C}$; that is, its roll and pitch angles remain zero (see Figure 3).
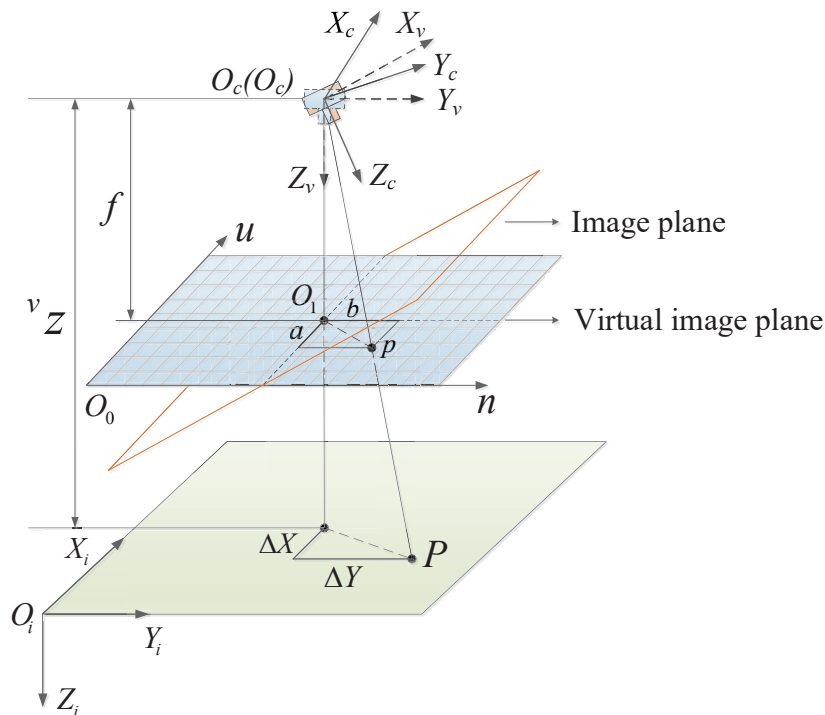

Figure 3. Camera frame $\mathcal{C}$ and virtual camera frame $\mathcal{V}$ with their corresponding image planes. The perspective projection of point P on the virtual image plane.

The coordinates of a point P in the inertial frame and camera frame are denoted by ${}^{i}\boldsymbol{P} = [{}^{i}P_x, {}^{i}P_y, {}^{i}P_z]^T$ and ${}^{c}\boldsymbol{P} = [{}^{c}P_x, {}^{c}P_y, {}^{c}P_z]^T$, respectively. They have the following geometric relationship,

$$ {}^{c}\boldsymbol{P} = \boldsymbol{R}^T({}^{i}\boldsymbol{P} - {}^{i}\boldsymbol{O}_c), \tag{5} $$

where ${}^{i}\boldsymbol{O}_c$ denotes the coordinates of the origin O_c in the inertial frame.

The pixel coordinates of the point P on the image plane are given by perspective projection equations [37]:

$$ \begin{cases} u = f\frac{{}^{c}P_x}{{}^{c}P_z} + u_0 \\ n = f\frac{{}^{c}P_y}{{}^{c}P_z} + n_0 \end{cases}, \tag{6} $$

in which f is the focal length of the camera and $[u_0, n_0]^T$ is the coordinate of the image plane center.

Now, reproject the image coordinates $[u, n]^T$ onto the virtual image plane using a matrix, $\boldsymbol{R}_{\theta\varphi}$, associated with a rotation in the roll angle around X_i and in the pitch angle around Y_i:

$$ \begin{bmatrix} {}^{v}u \\ {}^{v}n \end{bmatrix} = \frac{f}{\bar{R}^3_{\theta\varphi}\boldsymbol{s}} \begin{bmatrix} \bar{R}^1_{\theta\varphi}\boldsymbol{p} \\ \bar{R}^2_{\theta\varphi}\boldsymbol{p} \end{bmatrix}, \tag{7} $$

in which $\boldsymbol{p} = [u, n, f]^T$ and $\bar{R}^1_{\theta\varphi}$, $\bar{R}^2_{\theta\varphi}$, and $\bar{R}^3_{\theta\varphi}$ are the row vectors of the matrix $\boldsymbol{R}_{\theta\varphi}$.

It is assumed that $N(N \geq 3)$ non-collinear points are fixed on the selected target. The image centroid of the target is computed as

$$ {}^{v}u_g = \frac{1}{N} \cdot \sum_{k=1}^{N} {}^{v}u_k, \qquad {}^{v}n_g = \frac{1}{N} \cdot \sum_{k=1}^{N} {}^{v}n_k. \tag{8} $$

In this paper, we need to drive the quadrotor directly above the target, such that the desired image feature is determined as the center of the virtual image plane, $[{}^{v}u_0, {}^{v}n_0]^T$, as shown in Figure 3. The reprojection of the feature points onto the virtual camera frame decouples the pitch and roll motion of the vehicle through the change of coordinates $[{}^{v}u, {}^{v}n]^T$. Therefore, the relative horizontal distance can be estimated directly, using the deviation of the image centroid of the target from the center of the image plane:

$$ \Delta x = {}^{v}z\frac{a}{f}, \qquad \Delta y = {}^{v}z\frac{b}{f}, \tag{9} $$

where $a = {}^{v}u_g - {}^{v}u_0$ and $b = {}^{v}n_g - {}^{v}n_0$ are image errors defined in image space, and ${}^{v}z$ is the vertical distance of the virtual camera frame obtained by the ultrasonic sensor measurement.

3.2. Relative Distance Estimation in Vertical Direction

When tracking an uncooperative target, the relative horizontal distance estimated in the above section should converge to and remain at zero. The height control is relatively flexible, depending on different observation requirements. The easiest way which comes to our mind is to keep the quadrotor flying at a constant height; however, this cannot guarantee the target being observed in an optimal area on the image plane. Intuitively, the image size of the target varies with the height of the camera. The quadrotor is required to fly neither too high nor too low, avoiding making the target image invisible or leaving the field of view.

Now, introduce a circle as the optimal observation area on the virtual image plane centered at O_1 with radius r_{opt}. Then, construct a circle that passes through the farthest feature point from the image centroid of the target. The radius of the constructed circle, denoted by r, is defined as

$$ r = \max\{\sqrt{({}^{v}u_k - {}^{v}u_g)^2 + ({}^{v}n_k - {}^{v}n_g)^2}\}, k = 1, ..., N. \tag{10} $$

According to Equation (10), the constructed circle will cover all of the feature points of the selected compact target. Given that the relative horizontal error converges to zero, the vertical distance of the quadrotor can be controlled by adjusting r to be equal to or less than r_{opt}, which guarantees the target image being kept in the optimal observation area. Therefore, the desired value of radius r satisfies the following conditions,

$$r_d = \sigma r_{opt}, \quad \sigma \leq 1, \tag{11}$$

where σ is the radius scaling factor and $\sigma = 1$ indicates that the optimal observation area circumscribes the target image.

From Equations (10) and (11), the desired flying height of the quadrotor can be written as

$$z_d = {}^v z \cdot \frac{r}{r_d} = {}^v z \cdot \frac{r}{\sigma r_{opt}}. \tag{12}$$

Then, we can define the position error in the vertical direction as

$$\Delta z = {}^v z - z_d. \tag{13}$$

To control the translational motion of the quadrotor, we define the full position error as $\Delta \boldsymbol{\xi} = [\Delta x, \Delta y, \Delta z]^T$, which is desired to converge to $[0, 0, 0]^T$.

3.3. Relative Velocity Estimation

Generally, the control actions require knowledge of not only position error but also velocity error. The quadrotor is equipped with an IMU providing the angular velocity and acceleration of the vehicle. If the target is stationary, the relative velocity can be obtained directly by integrating the measured acceleration information. However, while the precision of the gyro is satisfactory for the needs of the vehicle maneuvers, UAV-equipped accelerometers are usually not accurate enough for evaluating the platform velocity [29]. Velocity error acquirement is more complicated when the uncooperative target is moving. In this paper, we use the optic flow and gyro measurements to estimate the relative quadrotor-to-target velocity, which is recorded by Partial Velocity Evaluation.

Taking the first derivative of Equation (7), we obtain the dynamics of an image feature ${}^v p_k = [{}^v u_k, {}^v n_k]^T$ [22]:

$$ {}^v \dot{\boldsymbol{p}}_k = \boldsymbol{L}_{vk}({}^v \boldsymbol{v} - {}^v \boldsymbol{v}_{pk}) + \boldsymbol{L}_{\psi k} \dot{\psi}, \tag{14}$$

in which ${}^v \boldsymbol{v}$ and ${}^v \boldsymbol{v}_{pk}$ are the velocities of the point P and the quadrotor expressed in the virtual camera frame, and the matrices $\boldsymbol{L}_{vk}$ and $\boldsymbol{L}_{\psi k}$ are defined as

$$\boldsymbol{L}_{vk} = \begin{bmatrix} -\frac{f}{{}^v P_z} & 0 & \frac{{}^v u_k}{{}^v P_z} \\ 0 & -\frac{f}{{}^v P_z} & \frac{{}^v n_k}{{}^v P_z} \end{bmatrix}, \boldsymbol{L}_{\psi k} = \begin{bmatrix} {}^v n_k \\ -{}^v u_k \end{bmatrix}.$$

Define a vector of image features, ${}^v \boldsymbol{p} = [{}^v \boldsymbol{p}_1^T, \cdots, {}^v \boldsymbol{p}_N^T]^T \in \mathbb{R}^{2N}$, and assume that the velocities of N points are approximately equal to the target velocity, ${}^v \boldsymbol{v}_{pk} = {}^v \boldsymbol{v}_t$. Then, Equation (14) can be extended as

$$ {}^v \dot{\boldsymbol{p}} = \boldsymbol{L}_v({}^v \boldsymbol{v} - {}^v \boldsymbol{v}_p) + \boldsymbol{L}_\psi \dot{\psi}, \tag{15}$$

in which $\boldsymbol{L}_v = [\boldsymbol{L}_{v1}^T, \cdots, \boldsymbol{L}_{vN}^T]^T \in \mathbb{R}^{2N \times 3}$ and $\boldsymbol{L}_\psi = [\boldsymbol{L}_{\psi 1}^T, \cdots, \boldsymbol{L}_{\psi N}^T]^T \in \mathbb{R}^{2N}$.

The finite time difference of image features, ${}^v \dot{\boldsymbol{p}}$, can be computed by directly measuring the optic flow of the visual features in images. Denoting by $\Delta \boldsymbol{v}$ the relative quadrotor-to-target velocity expressed in the inertial frame, we have

$$\Delta \boldsymbol{v} = \boldsymbol{R}_\psi \boldsymbol{L}_v^+({}^v \dot{\boldsymbol{p}} - \boldsymbol{L}_\psi \dot{\psi}), \tag{16}$$

in which $\boldsymbol{L}_v^+ \in \mathbb{R}^{3 \times 2N}$ is the Moore–Penrose pseudo-inverse of the matrix $\boldsymbol{L}_v$ and $\boldsymbol{R}_\psi$ is the rotation matrix corresponding to the yaw angle by the relation $\boldsymbol{R} = \boldsymbol{R}_\psi \boldsymbol{R}_{\theta\varphi}$.

The relative translational velocity (expressed in the inertial frame) Δv and the pixel velocities $^v\dot{\boldsymbol{p}}$ are related by the interaction matrix $\boldsymbol{L}_{vk}$ and $\boldsymbol{L}_{\psi k}$. At least two points on the target are required to determine the three components of vector Δv. Due to the fact that the virtual image plane is introduced, the depth information of each feature point is not necessary, which is an advantage over traditional visual odometry.

4. Gain-Switching PD Controller Design

The control objective in this work is to drive a quadrotor tracking an uncooperative target at an appropriate height for optimal observation. The quadrotor, as an underactuated system, has six degrees of freedom with only four control inputs. Therefore, the control of a quadrotor is typically divided into an outer position loop and an inner attitude loop. The outer loop provides the reference attitude signal to the inner loop, while the inner loop tracks the orientation reference of the vehicle. A block diagram of the UAV control loop structure is shown in Figure 4. The objective is equivalent to designing a control input for translational motion based on visual feedback, such that the controlled underactuated system with model uncertainties can guarantee $\Delta \boldsymbol{\xi} \to \boldsymbol{0}$ and $\Delta v \to \boldsymbol{0}$, then designing a control input for rotational motion with the knowledge of reference Euler angles derived from the outer loop.

PD controllers are commonly used in UAV control. Given the uncertainties in the dynamics of the vehicle, a traditional PD controller can produce large steady-state errors or even affect the overall stability of the system. To improve the robustness of the PD controller, we introduce a gain-switching term to deal with the uncertainties by avoiding using large gains.
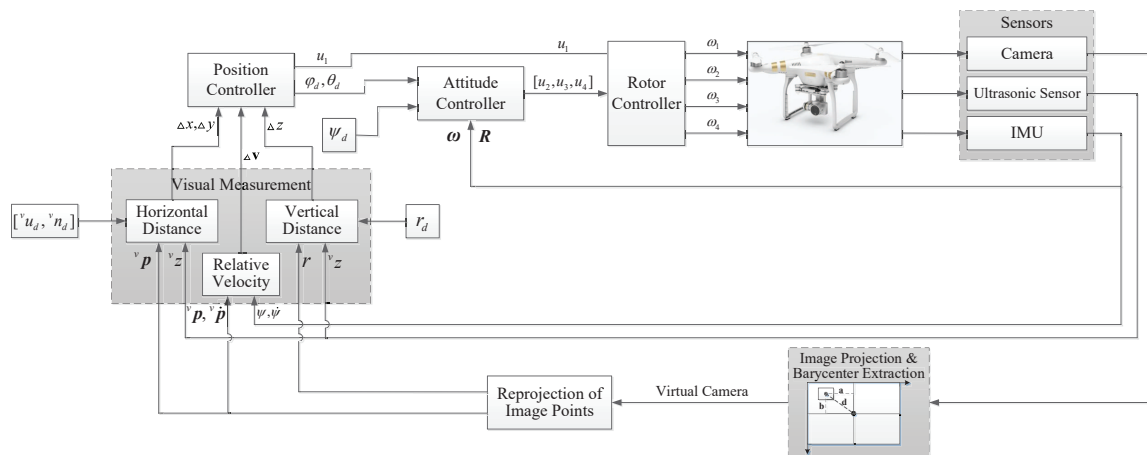

Figure 4. Block diagram of the UAV control system.

4.1. Control of the Translational Motion

The translational dynamics (1) can be rewritten, in matrix form, as

$$
\begin{bmatrix} \ddot{x} \\ \ddot{y} \\ \ddot{z} \end{bmatrix} = -\frac{u_1}{m} \begin{bmatrix} \cos\varphi \sin\theta \cos\psi + \sin\varphi \sin\psi \\ \cos\varphi \sin\theta \sin\psi - \sin\varphi \cos\psi \\ \cos\varphi \cos\theta \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ g \end{bmatrix} + \begin{bmatrix} f_{dx} \\ f_{dy} \\ f_{dz} \end{bmatrix}.
\tag{17}
$$

The translational motion of the quadrotor is actually accomplished by both the thrust and orientation of the body. Let us define a set of virtual control inputs for the translational motion,

$$\begin{cases} u_x = u_1(\cos\varphi\sin\theta\cos\psi + \sin\varphi\sin\psi) \\ u_y = u_1(\cos\varphi\sin\theta\sin\psi - \sin\varphi\cos\psi) \\ u_z = u_1(\cos\varphi\cos\theta) \end{cases} . \tag{18}$$

From Equations (17) and (18), we have

$$m\dot{v} = u + F_g + F_d. \tag{19}$$

The control law for the translational motion is proposed as

$$u = u_R + u_P + u_D + u_\varepsilon = -\bar{F}_g + (-K_P\bar{m}\Delta\xi) + (-K_D\bar{m}\Delta v) + (-K_\varepsilon\bar{m}n(\varepsilon,s)), \tag{20}$$

where quantities with overbar symbols indicate that *a priori* estimates are used, which may deviate from the true value; $\Delta\xi$ and Δv are obtained by visual measurement, as detailed in Section 3; $-\bar{F}_g$ is the model compensation term designed to eliminate nonlinear elements of the dynamics; and K_P, K_D, and K_ε represent the proportional, differential, and gain-switching coefficient matrices, respectively. The switching function $n(\varepsilon,s)$ can be any piecewise-continuous function with the following properties,

$$\begin{cases} \|s\|\,n(\varepsilon,s) = \|n(\varepsilon,s)\|\,s \\ \|n(\varepsilon,s)\| \ge 1 - \varepsilon/s, \ s \ne 0 \end{cases} . \tag{21}$$

In this paper, the switching function is chosen as

$$n = \begin{cases} -s/\varepsilon, & \|s\| < \varepsilon \\ -s/\|s\|, & \|s\| \ge \varepsilon, \end{cases} \tag{22}$$

where ε is the gain switching threshold and s is the error feedback term given by

$$s = k_P\Delta\xi + k_D\Delta v. \tag{23}$$

Substituting Equation (20) into (19) yields the error dynamics for the translational motion:

$$m\Delta\dot{v} = -K_P m\Delta\xi - K_D m\Delta v - K_\varepsilon n + h, \tag{24}$$

where the disturbance function, denoted by h, can be written as:

$$h = -K_P\delta m\Delta\xi - K_D\delta m\Delta v - \delta F_g - m\dot{v}_t + F_d, \tag{25}$$

where $\delta m = \bar{m} - m$ and $\delta F_g = \bar{F}_g - F_g$ denote the deviation between the measured values and the true values, $\dot{v}_t$ is the target acceleration, which is assumed to be unknown but has an upper bound.

Remark 1. *The disturbance function h represents all sources of uncertainties in translational error dynamics, including model uncertainty resulting from the inaccurately measured mass of the quadrotor $-K_P\delta m\Delta\xi$ and $-K_D\delta m\Delta v$, gravitational error $-\delta F_g$, unknown target motion $-m\dot{v}_t$, and unstructured forces F_d. Under Assumption 2–3, the disturbance function h is bounded.*

Now, the properties of the gain-switching coefficient will be addressed. The matrix K_ε is symmetric and positive definite, and must satisfy the following conditions,

$$\begin{cases} \|K_\varepsilon n\|/(\bar{m}+c_m) \ge k_\varepsilon\|n\| \\ k_\varepsilon \ge \|h\|/(\bar{m}+c_m) \end{cases}, \tag{26}$$

where k_ε is a positive bounding scalar designated to ensure that the term u_ε provides greater acceleration than that resulting from the disturbances in h.

The proportional and differential coefficient matrices $\boldsymbol{K}_P$ and $\boldsymbol{K}_D$ are selected to satisfy the following constraints with k_P and k_D,

$$\begin{cases} \|\boldsymbol{K}_P \bar{m} \Delta \boldsymbol{\xi}\| / (\bar{m} + c_m) \geq k_P \|\Delta \boldsymbol{\xi}\| \\ \|\boldsymbol{K}_D \bar{m} \Delta \boldsymbol{v}\| / (\bar{m} + c_m) \geq k_D \|\Delta \boldsymbol{v}\| \end{cases}, \tag{27}$$

in which k_P and k_D are feedback gains designated as the following functions of the desired rate of convergence, α:

$$\begin{cases} k_P = 2\alpha^2 \\ k_D = 2\alpha \end{cases}, \quad \alpha > 0. \tag{28}$$

These equations guarantee that the norm of $\boldsymbol{K}_P$ and $\boldsymbol{K}_D$ is large enough, such that the commanded force specified by $\boldsymbol{u}_P$ and $\boldsymbol{u}_D$ delivers—at a minimum—the acceleration specified by the vector $\boldsymbol{u}_\varepsilon$.

Remark 2. *Compared with traditional PD control, the proposed PD control consists of not only the proportional and differential terms $\boldsymbol{u}_P$ and $\boldsymbol{u}_D$, which can eliminate the position and velocity errors, but also a gain-switching term $\boldsymbol{u}_\varepsilon$, which acts as a robustifying term based on the switching function $\boldsymbol{n}(\varepsilon, \boldsymbol{s})$. By using Equation (26), appropriate robust control parameters can be selected to restrain the uncertainties $\boldsymbol{h}$. In addition, the transient behavior of the system can be characterized analytically, based on the desired rate of convergence in feedback gains (28).*

Referring to Equation (24) and defining the system state $\boldsymbol{e}_t = \left[\Delta \boldsymbol{\xi}^T, \Delta \boldsymbol{v}^T \right]^T$, the tracking dynamics model for translational motion can be written as

$$\begin{bmatrix} \Delta \dot{\boldsymbol{\xi}} \\ \Delta \dot{\boldsymbol{v}} \end{bmatrix} = \begin{bmatrix} \Delta \boldsymbol{v} \\ (\frac{1}{m}) \cdot (-\boldsymbol{K}_P m \Delta \boldsymbol{\xi} - \boldsymbol{K}_D m \Delta \boldsymbol{v} - \boldsymbol{K}_\varepsilon \boldsymbol{n}(\varepsilon, \boldsymbol{s}) + \boldsymbol{h}) \end{bmatrix}. \tag{29}$$

Lemma 1. *Consider a system $\dot{\boldsymbol{x}}(t) = f(t, \boldsymbol{x}(t))$. If there exist a continuously differentiable function $V(\boldsymbol{x})$ and scalars $c_1, c_2 > 0$, which satisfy [38]*

$$\begin{array}{ll} \text{(i)} & c_1 \|\boldsymbol{x}\|^2 \leq V(\boldsymbol{x}) \leq c_2 \|\boldsymbol{x}\|^2 \\ \text{(ii)} & \forall \boldsymbol{x} \in \{\boldsymbol{x} | \ V(\boldsymbol{x}) > V^*, V^* > 0\}, \\ & \exists \dot{V}(\boldsymbol{x}) \leq -2\alpha [V(\boldsymbol{x}) - V^*], \end{array}$$

then the system is (globally and uniformly) exponentially convergent to $\boldsymbol{B}(q)$ with rate α, where

$$\begin{array}{l} \boldsymbol{B}(q) \overset{\Delta}{=} \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}\| \leq q\} \\ q \overset{\Delta}{=} (V^*/c_1)^{1/2}. \end{array}$$

Theorem 1. *Consider the closed-loop system for the translational motion in Equation (29) with the controller designed by Equation (20). If the corresponding parameters are assigned as in Equations (26)–(28), then the state $\boldsymbol{e}_t$ will exponentially converge to zero.*

Proof. The Lyapunov function candidate is designated as

$$V(\boldsymbol{e}_t) = \frac{1}{2} \boldsymbol{e}_t^T \boldsymbol{A} \boldsymbol{e}_t, \tag{30}$$

where

$$\boldsymbol{A} = \begin{bmatrix} k_L \boldsymbol{I} & k_P \boldsymbol{I} \\ k_P \boldsymbol{I} & k_D \boldsymbol{I} \end{bmatrix} \tag{31}$$

and $\boldsymbol{I}$ is a unit matrix. The constant parameter k_L is also a function of α:

$$k_L = 4\alpha^3. \tag{32}$$

By using the values k_P, k_D, and k_L in Equations (28) and (32), the matrix $\boldsymbol{A}$ is positive definite. Thus, $V(\boldsymbol{e_t})$ is a positive definite function satisfying

$$\lambda_{\min}\|\boldsymbol{e_t}\|^2 \leq V(\boldsymbol{e_t}) \leq \lambda_{\max}\|\boldsymbol{e_t}\|^2. \tag{33}$$

Condition (i) in Lemma 1 is satisfied with $c_1 = \lambda_{\min}$ and $c_2 = \lambda_{\max}$, where $\lambda_{\min}$ and $\lambda_{\max}$ are the smallest and largest eigenvalues of $\boldsymbol{A}$.

Taking the time derivative of $V(\boldsymbol{e_t})$ and substituting Equation (29) into it, we have

$$\dot{V}(\boldsymbol{e_t}) = k_L \Delta\boldsymbol{\xi}^T \Delta\boldsymbol{v} + k_P \Delta\boldsymbol{v}^T \Delta\boldsymbol{v} - k_P[\Delta\boldsymbol{\xi}^T \boldsymbol{K_P} \Delta\boldsymbol{\xi} + \Delta\boldsymbol{\xi}^T \boldsymbol{K_D} \Delta\boldsymbol{v}] - \\ k_D[\Delta\boldsymbol{v}^T \boldsymbol{K_D} \Delta\boldsymbol{v} + \Delta\boldsymbol{v}^T \boldsymbol{K_P} \Delta\boldsymbol{\xi}] + (1/m)\boldsymbol{s}^T(\boldsymbol{h} - \boldsymbol{K}_\varepsilon \boldsymbol{n}(\varepsilon, \boldsymbol{s})). \tag{34}$$

Using the parameter constraints in Equations (26) and (27), it follows that

$$\dot{V}(\boldsymbol{x}) \leq -k_P^2 \Delta\boldsymbol{\xi}^T \Delta\boldsymbol{\xi} + (k_P - k_D^2)\Delta\boldsymbol{v}^T \Delta\boldsymbol{v} + (k_L - 2k_P k_D)\Delta\boldsymbol{\xi}^T \Delta\boldsymbol{v} + \\ \boldsymbol{s}^T[(1/m)\boldsymbol{h} - k_\varepsilon \boldsymbol{n}(\varepsilon, \boldsymbol{s})]. \tag{35}$$

Substituting Equations (28) and (32) into (35) leads to the following,

$$\dot{V}(\boldsymbol{e_t}) \leq -2\alpha V(\boldsymbol{e_t}) + E, \tag{36}$$

where

$$E = \boldsymbol{s}^T[(1/m)\boldsymbol{h} - k_\varepsilon \boldsymbol{n}(\varepsilon, \boldsymbol{s})]. \tag{37}$$

By using the properties of $\boldsymbol{n}$ in (21), Equation (37) yields the following expression,

$$E \leq \|\boldsymbol{s}\| h^* - k_\varepsilon \|\boldsymbol{n}\| \|\boldsymbol{s}\|, \tag{38}$$

where $h^* = \|\boldsymbol{h}\|/m$.

Equation (22) requires that $\boldsymbol{n} = \boldsymbol{s}/\varepsilon$ for $\|\boldsymbol{s}\| < \varepsilon$. Moreover, recalling (26), which states that $k_\varepsilon \geq h^*$, Equation (38) shows, in this case, that

$$E \leq k_\varepsilon \cdot \Delta, \tag{39}$$

where $\Delta = (\varepsilon^2 - \|\boldsymbol{s}\|^2)\big/\varepsilon > 0$.

When $\|\boldsymbol{s}\| \geq \varepsilon$, $\boldsymbol{n} = -\boldsymbol{s}/\|\boldsymbol{s}\|$, it follows that

$$E \leq 0. \tag{40}$$

Combining Equations (39) and (40), we can obtain a global upper bound

$$\bar{E} \leq k_\varepsilon \cdot \Delta. \tag{41}$$

Substituting Equation (41) into the time derivative of $\dot{V}(\boldsymbol{e_t})$ in Equation (38) leads to

$$\dot{V}(\boldsymbol{e_t}) \leq -2\alpha[V(\boldsymbol{e_t}) - V^*], \tag{42}$$

in which $V^* = \bar{E}/2\alpha$.

Inequality (42) ensures that the state of the system e_t can exponentially converge to a small ball around the origin defined by $V(e_t) < V^*$. Thus, the condition (*ii*) of Lemma 1 is satisfied.　□

It is worth noting that the control law given by Equation (20) yields virtual control inputs for translational motion. Using Equation (18), we can compute the command thrust u_1, as well as the desired roll and pitch angles for the attitude controller:

$$\begin{cases} u_1 = \sqrt{u_x^2 + u_y^2 + u_z^2} \\ \varphi_d = \arcsin\left(\frac{u_x \sin\psi_d - u_y \cos\psi_d}{u_1}\right) \\ \theta_d = \arctan\left(\frac{u_x \cos\psi_d + u_y \sin\psi_d}{u_z}\right) \end{cases}, \tag{43}$$

where ψ_d is the reference value of yaw. As the yaw angle is independent of the outer loop, we can prescribe it as its initial value or another constant.

4.2. Control of the Rotational Motion

In the inner attitude loop, a similar gain-switching PD control law is presented:

$$\begin{aligned} \boldsymbol{\tau} &= \boldsymbol{\tau}_R + \boldsymbol{\tau}_P + \boldsymbol{\tau}_D + \boldsymbol{\tau}_\varepsilon \\ &= \boldsymbol{\omega} \times \bar{\boldsymbol{J}}\boldsymbol{\omega} + (-\boldsymbol{K}_P'\bar{\boldsymbol{J}}\Delta\boldsymbol{\Phi}) + (-\boldsymbol{K}_D'\bar{\boldsymbol{J}}\Delta\boldsymbol{\omega}) + (-\boldsymbol{K}_\varepsilon'\bar{\boldsymbol{J}}\boldsymbol{n}'(\varepsilon',\boldsymbol{s}')), \end{aligned} \tag{44}$$

where $\Delta\boldsymbol{\Phi} = \boldsymbol{\Phi} - \boldsymbol{\Phi}_d$ and $\Delta\boldsymbol{\omega} = \boldsymbol{\omega} - \boldsymbol{\omega}_d$ denote the quadrotor attitude and the equivalent rate error, respectively, and $\boldsymbol{\Phi}_d = [\varphi_d, \theta_d, \psi_d]^T$ are the desired Euler angles output from the outer loop. The desired angular velocity $\boldsymbol{\omega}_d$ is taken to be zero.

The switching function $\boldsymbol{n}'(\varepsilon',\boldsymbol{s}')$ is chosen the same as before, and the error feedback term $\boldsymbol{s}'$ is given by

$$\boldsymbol{s}' = -k_P'\Delta\boldsymbol{\Phi} - k_D'\Delta\boldsymbol{\omega}. \tag{45}$$

Referring to Equation (2) and defining the system states $\boldsymbol{e}_r = \left[\Delta\boldsymbol{\Phi}^T, \Delta\boldsymbol{\omega}^T\right]^T$, the tracking dynamics model for rotational motion, under Assumption 1, can be written as

$$\begin{bmatrix} \Delta\dot{\boldsymbol{\Phi}} \\ \Delta\dot{\boldsymbol{\omega}} \end{bmatrix} = \begin{bmatrix} \Delta\boldsymbol{\omega} \\ \boldsymbol{J}^{-1}(-\boldsymbol{K}_P'\boldsymbol{J}\Delta\boldsymbol{\Phi} - \boldsymbol{K}_D'\boldsymbol{J}\Delta\boldsymbol{\omega} - \boldsymbol{K}_\varepsilon'\boldsymbol{J}\boldsymbol{n}'(\varepsilon',\boldsymbol{s}') + \boldsymbol{h}') \end{bmatrix}, \tag{46}$$

where the disturbance function $\boldsymbol{h}'$ is given by

$$\boldsymbol{h}' = -\boldsymbol{K}_P'\delta\boldsymbol{J}\Delta\boldsymbol{\Phi} - \boldsymbol{K}_D'\delta\boldsymbol{J}\Delta\boldsymbol{\omega} + (\boldsymbol{\omega} \times \bar{\boldsymbol{J}}\boldsymbol{\omega} - \boldsymbol{\omega} \times \boldsymbol{J}\boldsymbol{\omega}) - \boldsymbol{J}\dot{\boldsymbol{\omega}}_t + \boldsymbol{\tau}_d. \tag{47}$$

Theorem 2. *Consider the closed-loop system for the rotational motion in Equation (46) with the controller designed by Equation (44). If the corresponding parameters are assigned by*

$$\begin{cases} \|\boldsymbol{K}_P'\bar{\boldsymbol{J}}\Delta\boldsymbol{\Phi}\|/\lambda_{\max}(\boldsymbol{J}) \geq k_P' \|\Delta\boldsymbol{\Phi}\| \\ \|\boldsymbol{K}_D'\bar{\boldsymbol{J}}\Delta\boldsymbol{\omega}\|/\lambda_{\max}(\boldsymbol{J}) \geq k_D' \|\Delta\boldsymbol{\omega}\| \end{cases}, \tag{48}$$

$$\begin{cases} \|\boldsymbol{K}'_\varepsilon\|/\lambda_{\max}(\boldsymbol{J}) \geq k_\varepsilon' \|\boldsymbol{n}'(\varepsilon',\boldsymbol{s}')\| \\ k_\varepsilon' \geq \|\boldsymbol{h}'\|/\lambda_{\min}(\boldsymbol{J}) \end{cases}, \tag{49}$$

$$\begin{cases} k_P' = 2\alpha'^2 \\ k_D' = 2\alpha' \end{cases}, \tag{50}$$

then, the state $\boldsymbol{e}_r$ will exponentially converge to zero.

Proof. The stability analysis of the system for rotational motion is similar to that for translational motion, so it is not described in detail here. □

5. Simulation Results

In this section, MATLAB simulations are presented to validate the performance of the proposed vision-based control scheme. We considered two scenarios in this work: hovering above a stationary target and tracking a moving target. In both scenarios, the target to be observed was assumed to be uncooperative and without detailed geometry information. The physical parameters of the simulated quadrotor were $m = 2.1$ kg and $J = \text{diag}\{0.0096, 0.0098, 0.016\}$ kg $\cdot$ m^2/rad. The nominal part of the quadrotor's mass and moment of inertia were $\bar{m} = 2.1$ kg and $\bar{J} = \text{diag}\{0.0081, 0.0081, 0.0142\}$ kg $\cdot$ m^2/rad. The focal length divided by pixel size of the camera was set as 213, and the image resolution was 160×120 pixels with principal point located at $[80, 60]$. Considering that the quadrotor may be affected by wind disturbances in the environment, we applied some sinusoidal (cosinusoidal) forces and torques to the vehicle with the following values,

$$
\begin{aligned}
F_d &= 0.5 \cdot \begin{bmatrix} \sin(2\pi t) & \cos(\pi t) & -\sin(2\pi t) \end{bmatrix}^T \\
\tau_d &= 0.01 \cdot \begin{bmatrix} -\sin(2\pi t) & -\cos(\pi t) & \sin(2\pi t) \end{bmatrix}^T
\end{aligned}
.
$$

5.1. Scenario 1: Hovering Observation

In some cases, such as traffic and fire accidents, we usually drive a UAV equipped with optical sensors to approach the scene of the accident and hover over the damaged vehicle or burning object. To provide distinct images and effective air support for the subsequent rescue, it is required to adjust the height of the UAV, keeping the designated target within the optimal area of the camera's field of view. The proposed vision-based control scheme in this work is applicable to the above missions.

In the simulation, the target was set as a rectangular object. Its visual features included its four vertexes with the following coordinates, $[0.2, 0.25, 0]$ m, $[0.2, 0.25, 0]$ m, $[-0.2, -0.25, 0]$ m, and $[0.2, -0.25, 0]$ m, all expressed in the inertial frame. The quadrotor initial position and attitude were $[1, 0.8, -4]$ m and $[0, 0, 0.2]$ rad, respectively. For the optimal observation of the target, we defined a reference circle of radius 40 pixels on the virtual image plane, and the radius scaling factor σ was designated to be 1. The corresponding desired height was 1.7048 m, which can be computed using Equation (12). The control gains used in the simulation are listed in Table 1.

Table 1. Control gains in the simulation.

Gains	Values	Gains	Values
k_P	1	k'_P	8
k_D	1.44	k'_D	4
k_ε	0.3	k'_ε	1
K_P	diag{1.2,1.2,1.3}	K'_P	diag{10,10,10}
K_D	diag{1.5,1.5,1.6}	K'_D	diag{5,5,5}
K_ε	$m \cdot$diag{0.3,0.3,0.5}	K'_ε	$k'_\varepsilon \cdot J$
ε	0.1	ε'	0.3

The transient response of the system mainly depends on the feedback gains k_P, k_D and k'_P, k'_D, which are determined by the constraint Equations (28) and (50). Based on the trial and error method, the desired rates of convergence for translational and rotational motion were selected as $\sqrt{2}/2$ and 2, respectively. Then, the proportional and differential matrices K_P, K_D and K'_P, K'_D could be determined by using Equations (27) and (48). To guarantee the robustness properties of the system in the presence of wind and modelling errors, the control gains in the gain-switching term k_ε, K_ε and k'_ε, K'_ε must be

selected such that they are larger than the effect resulting from the unstructured disturbances h, h', based on Equations (26) and (49). The gain switching thresholds ε and ε' were tuned repeatedly to restrain the chattering phenomenon of the control efforts.

The simulation results are illustrated in Figures 5–7. The translational motion of the quadrotor is shown in Figure 5. The horizontal position errors converged to less than 0.05 m with transient response time enduring at about 5 s. The error in the vertical direction was relatively large, about 0.12 m, mainly resulting from the uncertainty of the quadrotor mass. Figure 6 describes the rotational motion of the quadrotor. The Euler angles were kept within a small range, satisfying Assumption 1. The properties of the target centroid in image space are illustrated in Figure 7. It is shown that the vehicle could rapidly and smoothly approach the target, and the control accuracy of the target centroid on the image plane was kept within 5 pixels. The proposed system, therefore, is satisfactory for continuous hovering observation of the target.
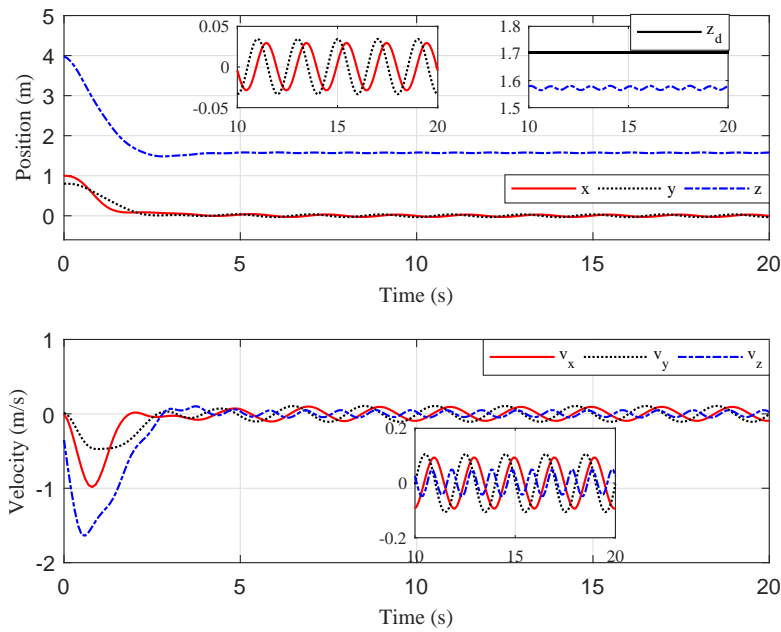
Figure 5. Scenario 1: Time evolution of the quadrotor Cartesian coordinates and velocity.
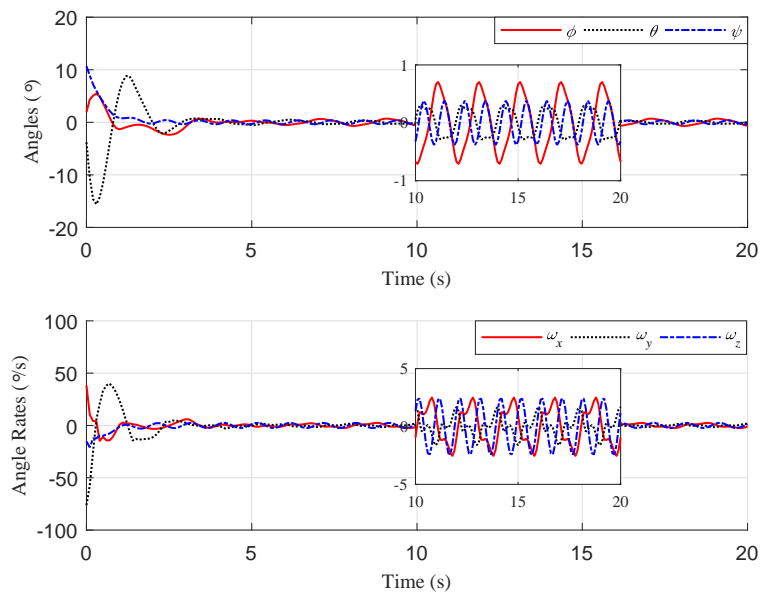
**Figure 6.** Scenario 1: Time evolution of the quadrotor attitude and angular velocity.
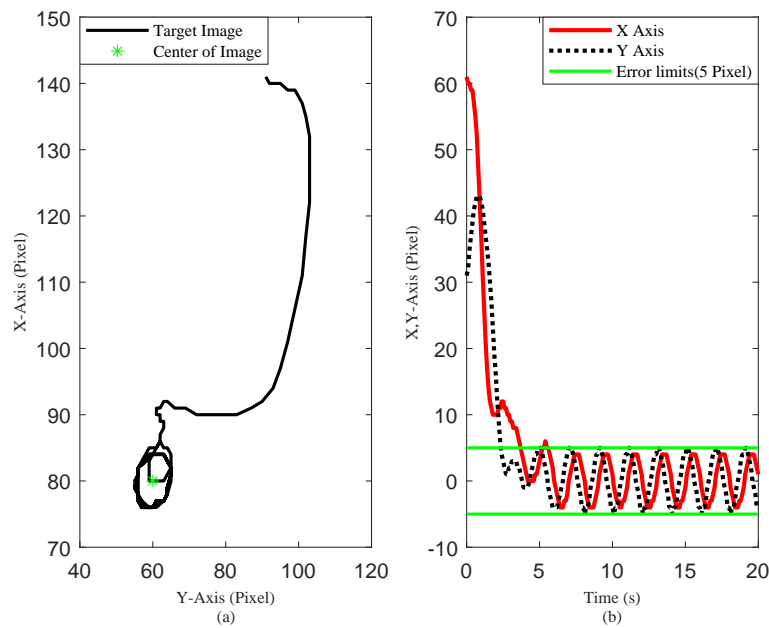
Figure 7. Scenario 1: (**a**) Target centroid trajectory on the image plane. (**b**) Target centroid deviation from the image center.

For comparison, we ran a simulation of traditional PD control using the same initial conditions and gains. The results are reported in Figures 8 and 9. The results indicate that the proposed gain-switching PD controller had better performance than the traditional PD controller in steady-state performance, with lower oscillation, and the proposed method had a smaller error limit of the target centroid in image space. Further, we relocated the quadrotor at two different initial positions to compare the system steady-state behavior between the gain-switching PD controller with the traditional PD controller. The mean, amplitude, and accuracy (3σ) of the position error were illustrated in Table 2. The results showed that the proposed controller was effective in different initial conditions. Compared with the traditional PD controller, the performance of the proposed controller was greatly improved, with the position-control accuracy increasing by a factor of 4–5.
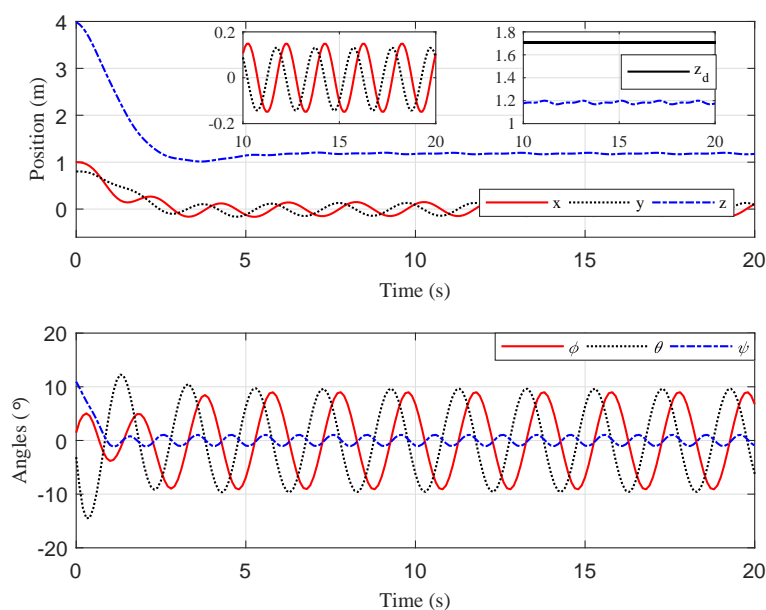
Figure 8. Scenario 1: Time evolution of the quadrotor Cartesian coordinates and attitude (proportional-derivative (PD) controller).
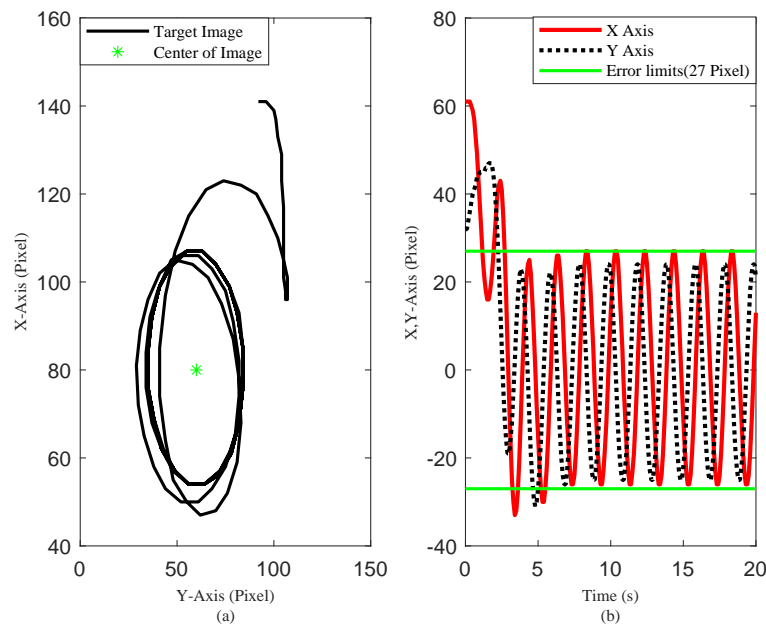
Figure 9. Scenario 1: (**a**) Target centroid trajectory on the image plane; and (**b**) Target centroid deviation from the image center (PD controller).

Table 2. The steady-state performance (position error) of the two methods.

	Traditional PD Control			Gain switching PD Control		
	Mean	**Amplitude**	**Accuracy**	**Mean**	**Amplitude**	**Accuracy**
Initial Position (*m*) [1, 0.8, −4]	8.9×10^{-4}	0.1475	0.3209	3.7×10^{-4}	0.0290	0.0628
	−0.0051	0.1360	0.2985	1.5×10^{-4}	0.0345	0.0731
	−0.5211	0.0164	0.5495	−0.1323	0.0082	0.1484
Initial Position (*m*) [1, 2, −5]	9.0×10^{-4}	0.1475	0.3209	3.7×10^{-4}	0.0291	0.0628
	−0.0051	0.1360	0.2984	1.6×10^{-4}	0.0345	0.0731
	−0.5212	0.0165	0.5495	−0.1324	0.0083	0.1485
Initial Position (*m*) [0.6, 0.8, −1.5]	8.8×10^{-4}	0.1475	0.3208	3.7×10^{-4}	0.0290	0.0627
	−0.0052	0.1360	0.2985	1.6×10^{-4}	0.0345	0.0731
	−0.5211	0.0164	0.5494	−0.1322	0.0082	0.1483

To evaluate the imaging effects of the target and the corresponding flying height of the quadrotor under different observation requirements, we ran another simulation ignoring the external disturbance and model uncertainties. In the simulation, three criteria were set: flying at a constant height (i.e., $\Delta z = 0$), radius scaling factor $\sigma = 1$, and $\sigma = 0.8$. As plotted in Figure 10, the results showed that introducing a reference circle on the virtual image plane and adjusting the radius scaling factor can regulate the imaging effect, according to different observation requirements. To illustrate that the proposed strategy is applicable to different kinds of targets, we ran another simulation to hover above an irregularly shaped target. The Cartesian coordinates of the feature points that enclose the target were: [0.25, −0.05, 0] m, [0.10, 0.20, 0] m, [−0.20, 0.15, 0] m, [−0.25, −0.10, 0] m, [−0.10, −0.20, 0] m, [0.15, −0.15, 0] m. The imaging effect under different observation requirements was shown in Figure 11. Similar to the rectangular target, the proposed control strategy can meet the hovering and observation needs of the target.
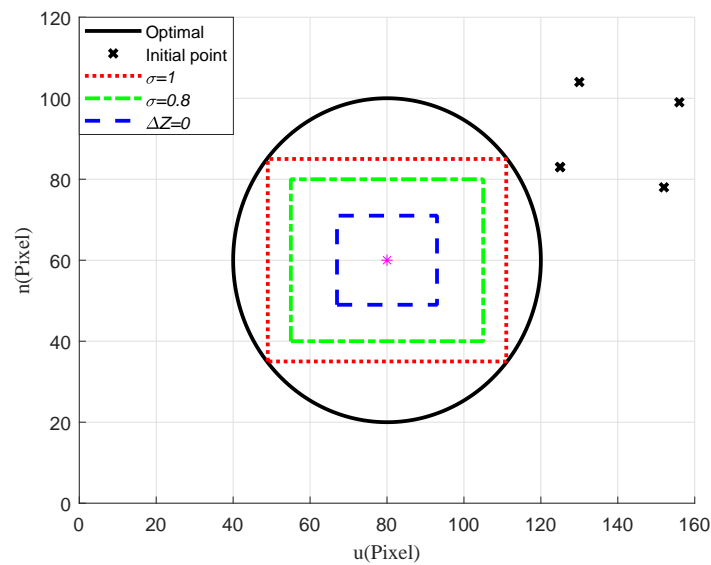
Figure 10. Scenario 1: Imaging effect of the rectangular target under different observation requirements.
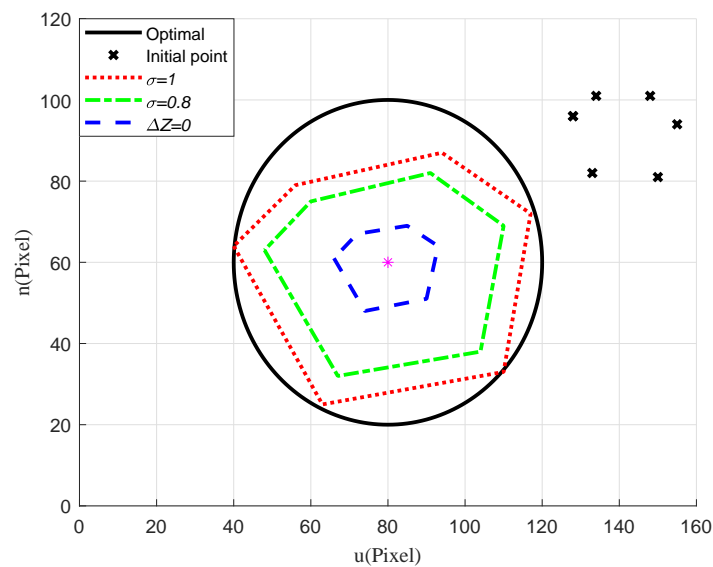

Figure 11. Scenario 1: Imaging effect of the irregularly shaped target under different observation requirements.

5.2. Scenario 2: Tracking a Moving Target

We also considered some more challenging cases, such as when the target vehicle is moving or when a fire area changes in real-time. Therefore, not only positioning the UAV over the target is required, but also following its (unknown) trajectory. Compared with the traditional PD control, the robust controller proposed in this paper was more competent in this challenging mission. In this simulation, the target had the same geometric characteristics as in SCENARIO 1 but followed three different trajectories on the flat ground: circular movement, S-type movement, and linear movement. The quadrotor initial position was set to be $[1, 1.5, -3]$ m. To evaluate the performance of the proposed controllers in a more realistic environment, we added some noise to the measurement information. White noise with covariances of 0.5 and 10^{-4} was augmented to the visual data (image features and their pixel velocities) and angular rates, respectively.

The tracking performance of the circular, S-type, and linear movements were illustrated in Figures 12–17, respectively. Trajectories of motion of the quadrotor and the target in a 3D environment

were plotted in Figures 12, 14 and 16, which showed satisfactory tracking performance in spite of the presence of disturbance and noise. Meanwhile, Figures 13, 15 and 17 depicted the position tracking error of the two controllers in different target movements. The results indicated that the proposed control strategy applied to different target maneuvers, and showed better robustness and higher tracking accuracy than the traditional PD controller. The only exception was that there existed some fluctuations during the linear tracking (Figures 16 and 17). It was because of the vehicle's reaction time when the target suddenly changed its direction.
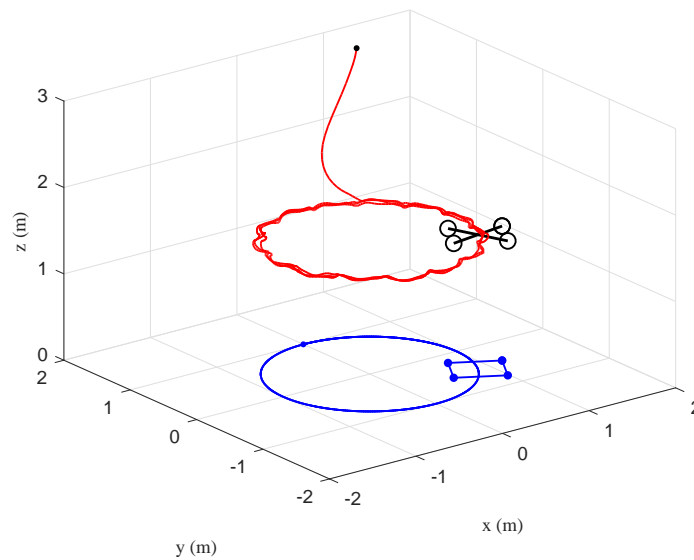
Figure 12. Scenario 2: Circular trajectory of the quadrotor and the target in a 3D environment under the proposed controller.
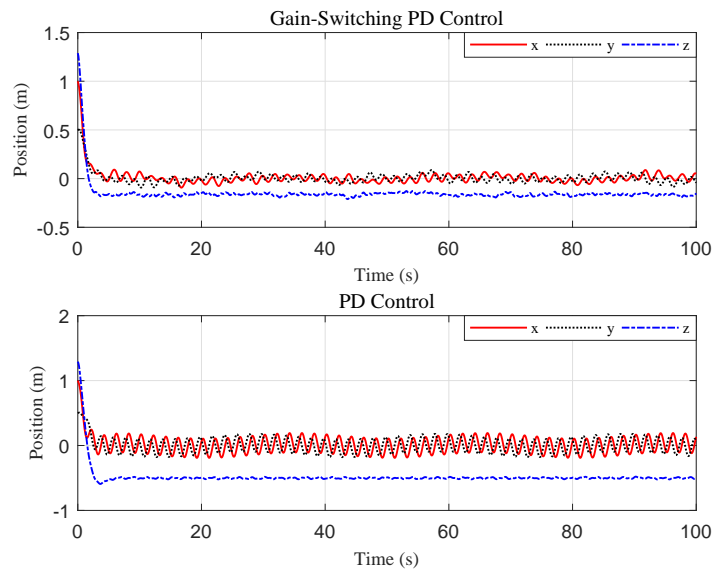
Figure 13. Scenario 2: Time evolution of the position tracking error under the two controllers (circular movement).
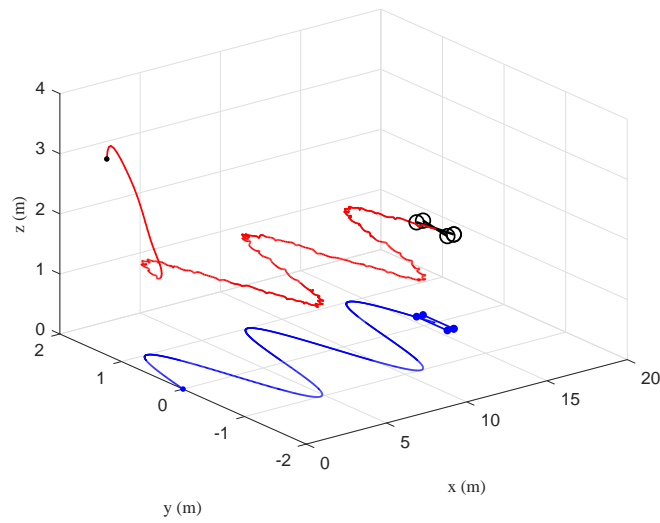
Figure 14. Scenario 2: S-type trajectory of the quadrotor and the target in a 3D environment under the proposed controller.
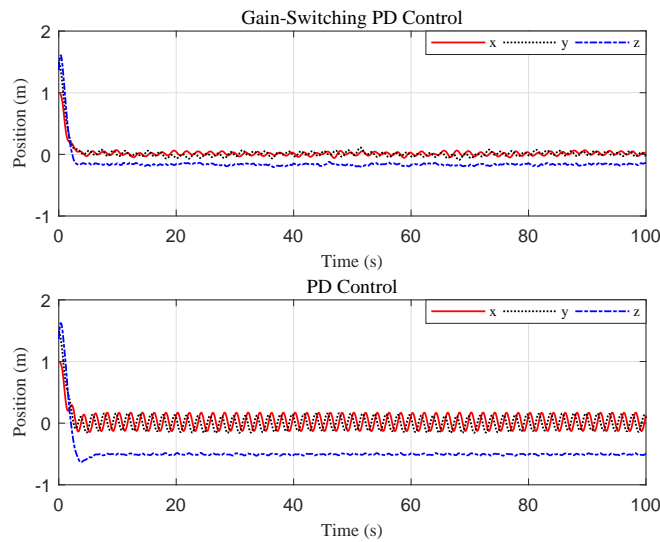
Figure 15. Scenario 2: Time evolution of the position tracking error under the two controllers (S-type movement)
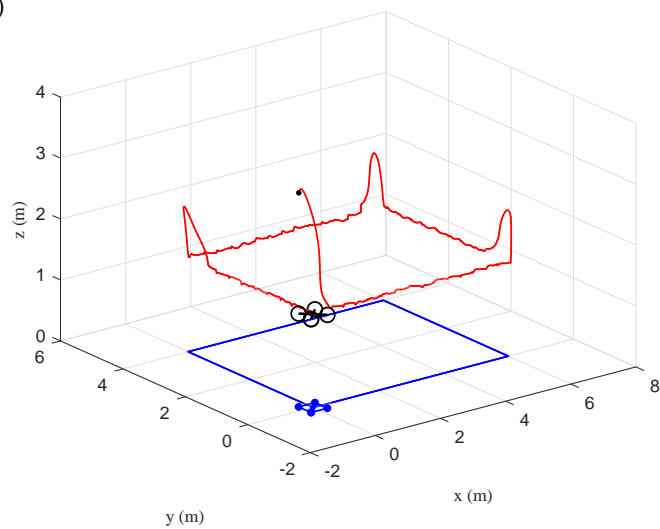

Figure 16. Scenario 2: Linear trajectory of the quadrotor and the target in a 3D environment under the proposed controller.
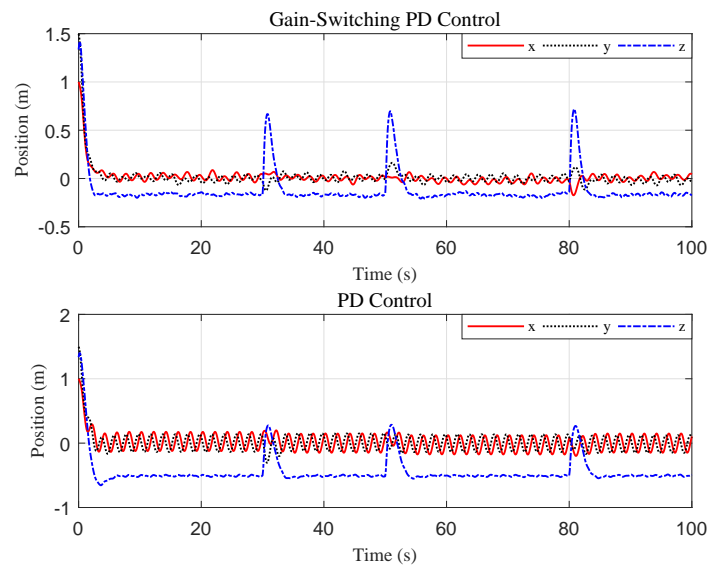
Figure 17. Scenario 2: Time evolution of the position tracking error under the two controllers (linear movement).

6. Conclusions

In this paper, we have developed a vision-based control scheme of a quadrotor for target tracking in the absence of location information and geometric features of the target. After transforming the image features to a virtual camera frame, optical-based metrology is exploited to estimate the relative distance and velocity. At the same time, the height of the quadrotor and image size can be adjusted by regulating the optimal observation area and radius scaling factor. Considering the presence of external interference and model uncertainties, we presented a gain-switching proportional-derivative (PD) control strategy to improve the robustness of the system. Two case studies, corresponding to hovering and tracking scenarios, are presented in this work. The simulation results indicated that the proposed vision-based scheme performed better in both hovering and tracking missions, compared with the traditional PD control.

In future work, we are going to add a field of view constraint to the system, as the proposed algorithm can not guarantee that all visual features are always kept inside the field of view of the camera. We also plan to implement the proposed control scheme in a real quadrotor.

Author Contributions: S.Z. provided the main idea and supervised the whole process. X.Z. performed the simulation and finished the draft manuscript. B.Z. reviewed and edited the manuscript. All authors have read and agreed to the published version of the manuscript.

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

References

1. Liang, X.; Fang, Y.; Sun, N.; Lin, H. Nonlinear hierarchical control for unmanned quadrotor transportation systems. *Aerosp. Sci. Technol.* **2018**, *65*, 3395–3405. [CrossRef]
2. Kim, D.; Liu, M.Y.; Lee, S.; Kamat, V.R. Remote proximity proximity monitoring between mobile construction resources using camera-mounted UAVs. *Autom. Constr.* **2019**, *99*, 168–182. [CrossRef]
3. Yao, P.; Xie, Z.; Ren, P. Optimal UAV route planning for coverage search of stationary target in river. *IEEE Trans. Control Syst. Technol.* **2019**, *27*, 822–829. [CrossRef]
4. Kim, I.H.; Jeon, H.; Baek, S.C.; Hong, W.H.; Jung, H.J. Application of crack identification techniques for an aging concrete bridge inspection using an unmanned aerial vehicle. *Sensors* **2018**, *18*, 1881. [CrossRef]

5. Zhu, W.W.; Du, H.B.; Cheng, Y.Y.; Chu, Z.B. Hovering control for quadrotor aircraft based on finite-time control algorithm. *Nonlinear Dyn.* **2017**, *88*, 2359–2369. [CrossRef]

6. Dong, J.; He, B. Novel fuzzy PID-type iterative learning control for quadrotor UAV. *Sensors* **2019**, *19*, 24. [CrossRef]

7. Wai, R.J.; Prasetia, A.S. Adaptive neural network control and optimal path planning of UAV surveillance system with energy consumption prediction. *IEEE Access* **2019**, *7*, 126137–126153. [CrossRef]

8. Jin, X.B.; Sun, S.; Wei, H.; Yang, F.B. Advances in Multi-Sensor Information Fusion: Theory and Applications 2017. *Sensors* **2018**, *18*, 1162. [CrossRef]

9. Mahony, R.; Kumar, V.; Corke, P. Multirotor aerial vehicles: Modeling, estimation, and control of quadrotor. *IEEE Robot. Autom. Mag.* **2012**, *19*, 20–32. [CrossRef]

10. Agrawal, P.; Ratnoo, A.; Ghose, D. Inverse optical flow based guidance for UAV navigation through urban canyons. *Aerosp. Sci. Technol.* **2017**, *68*, 163–178. [CrossRef]

11. Hrabar, S.; Sukhatme, G.S.; Corke, P.; Usher, K.; Roberts, J. Combined optic-flow and stereo-based navigation of urban canyons for a UAV. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Edmonton, AB, Canada, 2–6 August 2005; pp. 3309–3316.

12. Chaumette, F.; Hutchinson, S. Visual servo control, part I: Basic approaches. *IEEE Robot. Autom. Mag.* **2006**, *13*, 82–90. [CrossRef]

13. Zhao, B.; Tang, Y.; Wu, C.; Du, W. Vision-based tracking control of quadrotor with backstepping sliding mode control. *IEEE Access* **2018**, *6*, 72439–72448. [CrossRef]

14. Oh, H.; Won, D.Y.; Huh, S.S.; Shim, D.H.; Tahk, M.J.; Tsourdos, A. Indoor UAV control using multi-camera visual feedback. *J. Intell. Robot. Syst.* **2011**, *61*, 57–84. [CrossRef]

15. Shirzadeh, M.; Asl, H.J.; Amirkhani, A.; Jalali, A.A. Vision-based control of a quadrotor utilizing artificial neural networks for tracking of moving targets. *Eng. Appl. Artif. Intell.* **2017**, *58*, 34–48. [CrossRef]

16. Herisse, B.; Hamel, T.; Mahony, R.; Russotto, F.X. Landing a VTOL unmanned aerial vehicle on a moving platform using optical flow. *IEEE Trans. Robot.* **2012**, *28*, 77–89. [CrossRef]

17. 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*, 1524–1535. [CrossRef]

18. Benhimane, S.; Malis, E. Homography-based 2D visual tracking and servoing. *IEEE Trans. Robot.* **2007**, *26*, 661–676. [CrossRef]

19. De Plinval, H.; Morin, P.; Mouyon, P.; Hamel, T. Visual servoing for underactuated VTOL UAVs: A linear, homography-based framework. *Int. J. Robot. Res.* **2014**, *24*, 2285–2308. [CrossRef]

20. Ozawa, R.; Chaumette, F. Dynamic visual servoing with image moments for a quadrotor using a virtual spring approach. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA), Shanghai, China, 9–13 May 2011; pp. 5670–5676.

21. Ozawa, R.; Chaumette, F. Dynamic visual servoing with image moments for an unmanned aerial vehicle using a virtual spring approach. *Adv. Robot.* **2013**, *27*, 683–696. [CrossRef]

22. Asl, H.J.; Oriolo, G.; Bolandi, H. An adaptive scheme for image-based visual servoing of an underactuated UAV. *Int. J. Robot. Autom.* **2014**, *29*, 92–104. [CrossRef]

23. Li, J.A.; Xie, H.; Ma, R.; Low, K. H. Output feedback image-based visual servoing of rotorcrafts. *J. Intell. Robot. Syst.* **2019**, *93*, 277–287. [CrossRef]

24. Guo, D.; Wang, H.; Leang, K.K. Nonlinear vision-based observer for visual servo control of an aerial robot in global positioning system denied environments. *J. Mech. Robot.* **2018**, *10*, 061018. [CrossRef]

25. Opromolla, R.; Fasano, G.; Rufino, G.; Grassi, M. A review of cooperative and uncooperative spacecraft pose determination techniques for close-proximity operations. *Prog. Aerosp. Sci.* **2017**, *93*, 53–72. [CrossRef]

26. Munguía, R.; Urzua, S.; Bolea, Y.; Grau, A. Vision-based SLAM system for unmanned aerial vehicles. *Sensors* **2016**, *16*, 372. [CrossRef] [PubMed]

27. Zhang, X.; Xian, B.; Zhao, B.; Zhang, Y. Autonomous flight control of a nano quadrotor helicopter in a gps-denied environment using on-board vision. *IEEE Trans. Ind. Electron.* **2015**, *62*, 6392–6403. [CrossRef]

28. Fu, C.H.; Olivares-Mendez, M.A.; Suarez-Fernandez, R.; Campoy, P. Monocular visual-inertial SLAM-based collision avoidance strategy for fail-safe UAV using fuzzy logic controllers. *J. Intell. Robot. Syst.* **2014**, *73*, 513–533. [CrossRef]

29. Sabatini, M.; Palmerini, G.B.; Monti, R.; Gasbarri, P. Image based control of the "PINOCCHIO" experimental free flying platform. *Acta Astronaut.* **2014**, *94*, 480–492. [CrossRef]

30. Chen, Z.Y.; Chen, W.C.; Liu, X.M.; Song, C. Fault-tolerant optical flow sensor/SINS integrated navigation scheme for MAV in a GPS-denied environment. *J. Sens.* **2018**, *2018*, 9678505. [CrossRef]

31. Cheng, H.W.; Chen, T.L.; Tien, C. H. Motion estimation by hybrid optical flow technology for UAV landing in an unvisited area. *Sensors* **2019**, *19*, 1380. [CrossRef]

32. Moreno-Valenzuela, J.; Perez-Alcocer, R.; Guerrero-Medina, M.; Dzul, A. Nonlinear PID-type controller for quadrotor trajectory tracking. *IEEE-ASME Trans. Mechatron.* **2018**, *23*, 2436–2447. [CrossRef]

33. Goodarzi, F.; Lee, D.; Lee T. Geometric nonlinear PID control of a quadrotor UAV on SE (3). In Proceedings of the European Control Conference (ECC), Zurich, Switzerland, 17–19 July 2013; pp. 3845–3850.

34. Huang, T.; Huang, D.; Wang Z.; Shah, A. Robust tracking control of a quadrotor UAV based on adaptive sliding mode controller. *Complexity* **2019**, *2019*, 7931632. [CrossRef]

35. Tan, J.; Fan, Y.; Yan, P.; Wang, C.; Feng, H. Sliding mode fault tolerant control for unmanned aerial vehicle with sensor and actuator faults. *Sensors* **2019**, *19*, 643. [CrossRef] [PubMed]

36. Gonzalez, R.C.; Woods, R.E. *Digital Image Processing*, 3rd ed.; Prentice Hall: Upper Saddle River, NJ, USA, 2008.

37. Corke, P. *Robotics, Vision and Control: Fundamental Algorithms in MATLAB*, 2nd ed.; Springer: Berlin, Germany, 2017.

38. Thurman, S.W.; Flashner, H. Robust digital autopilot design for spacecraft equipped with pulse-operated thrusters. *J. Guid. Control Dyn.* **1996**, *19*, 1047–1055. [CrossRef]