Image-Based Visual Servoing of Rotorcrafts to Planar Visual Targets of Arbitrary Orientation

Jianan Li, Hui Xie, Kin Huat Low, Jianwen Yong, and Boyang Li

***Abstract*—This letter for the first time extends the virtual camera image-based visual servoing (IBVS) scheme to enable an underactuated rotorcraft UAV to regulate its translational motion and heading relative to a planar visual target of arbitrary orientation. The conversion from real camera images to virtual camera images of visual targets are proposed based on a set of rotation matrices. Hence, image moment features can be reused due to the simplicity and decoupled structure of the interaction matrix, and satisfactory 3D Cartesian trajectory of UAVs. In the design of the IBVS control law, the external disturbance and model uncertainties are estimated by an integral-based filter. In addition, to enable tracking of a moving visual target, a velocity estimator is developed. The global asymptotic stability of the error dynamics is proven. Both of the simulation and experimental results of tracking of a tilted moving planar target are provided to show the efficacy of the proposed IBVS scheme.**

***Index Terms*—Visual servoing, aerial systems: mechanics and control, image moment, nonlinear backstepping control.**

I. Introduction

USING visual information from onboard cameras of rotorcraft unmanned aerial vehicles (RUAVs) for closed-loop pose control is referred to as *visual servoing* [1]. Approaches for visual servoing are usually divided into two categories which are position-based visual servoing (PBVS) and image-based visual servoing (IBVS) [2]. In IBVS, kinematic errors are estimated directly from image features. It removes the requirement of pose reconstruction which needs precise camera calibration, thus it is a computationally effective algorithm [3]. In addition, a coarse estimation of the interaction matrix in IBVS will only cause perturbations in the robot trajectory but has no effect on the pose reached, and IBVS is generally more robust to image noise [2]. Considering these advantages, this paper selects IBVS approach.

Manuscript received February 22, 2021; accepted July 13, 2021. Date of publication August 4, 2021; date of current version August 17, 2021. This letter was recommended for publication by Associate Editor G. Loianno and Editor P. Pounds upon evaluation of the reviewers' comments. *(Corresponding author: Hui Xie.)*

Jianan Li is with the Department of Information Science & Electronic Engineering, Zhejiang University and the School of Engineering at Westlake University, 310024 Hangzhou, China (e-mail: lijianan@westlake.edu.cn).

Hui Xie is with the School of Electrical Engineering, Computing & Mathematical Sciences, Curtin University, Bentley, WA 6102, Australia (e-mail: hui.xie@curtin.edu.au).

Kin Huat Low and Jianwen Yong are with the School of Mechanical and Aerospace Engineering, Nanyang Technological University, Singapore 639798, Singapore (e-mail: mkhlow@ntu.edu.sg; yong0188@e.ntu.edu.sg).

Boyang Li is with the Department of Aeronautical and Aviation Engineering, Hong Kong Polytechnic University, Hong Kong (e-mail: boyang.li@polyu.edu.hk).

Digital Object Identifier 10.1109/LRA.2021.3101878

For an underactuated aerial vehicle like a traditional quadrotor whose translation is controlled by only a thrust force along a body-fixed axis [4], the dynamics of that robot must be considered in visual servoing as suggested in [5]. IBVS considering robot dynamics is referred to as *dynamic IBVS*, or *DIBVS* [6]. In addition, compared to classical kinematics-based visual servoing, DIBVS can help to achieve high speed motion control [7]. In designing DIBVS laws, the challenge is that the nonlinear perspective projection of the camera will destroy the passivity-like property of the aerial vehicle's dynamics [6]. In [8] a DIBVS controller using first order spherical image moments as image features is proposed. A rigorous proof of closed-loop stability is provided but this approach is insensitive to vertical motion due to the ill-conditioned interaction matrix of the proposed features. After the seminal work in [8], several improved DIBVS laws are proposed based on the spherical image moments, such as the works in [9]–[13]. However, the ill conditioning of the spherical moment features still has not been completely solved though some new image features based on spherical projection are proposed to alleviate the ill conditioning [6], [14].

Alternatively, the work in [15] proposes a method based on virtual camera philosophy where the roll and pitch motion of the UAV is decoupled from the image feature kinematics. Similarly, in [16] the innovative perspective image moments proposed in [17] are adopted to form image features in a virtual camera with zero roll and pitch for IBVS of RUAVs. The changing rates of image moment features are linear to the translational motion of the UAV and there is no ill conditioning in the interaction matrix, and therefore satisfactory trajectories in both image space and 3D Cartesian space can be achieved as compared to the spherical projection-based approach [6]. Because of the above-mentioned benefits, in recent years many virtual camera-based DIBVS laws are proposed [18]–[29]. These works adopt the same perspective image moments features as in [16] and propose various control laws to address specific issues, such as system uncertainties [18]–[21], removing the requirement of velocity estimation [19], [21]–[24], field of view (FoV) constraint [25], disturbance rejection [21], [26], tracking of moving targets [27], and global stability of IBVS [28], [29]. These IBVS laws in [18]–[29] are based on the assumption that the visual target lies on a horizontal plane. Both simulation and experimental results in [22] show that the proposed approach may also be applied to visual servoing of a visual target in a tilted plane, but no theoretical proof is given. To the authors' knowledge, there is no existing results of applying the virtual camera approach to visual servoing of arbitrarily oriented visual targets. Other

less popular DIBVS methods include the homography-based method as shown in [30] and the virtual spring method presented in [31][32]. Among all above-mentioned approaches, the virtual camera approach is usually considered easier to implement due to its simplicity and decoupled structure shown in the interaction matrix [6].

This paper proposes a novel virtual camera IBVS control scheme to an arbitrarily oriented planar visual target for underactuated RUAVs. This scheme relies on the conversion of real time images to virtual images of a virtual camera pointing in the direction of normal vector of the visual target. To extend the application of perspective moment features to IBVS of non-horizontal targets, practical issues encountered in [18]–[29] have to be addressed as well. In this paper, we focus on addressing the issues caused by model uncertainties and disturbance, tracking moving targets, and stability analysis. The control law is designed using the adaptive backstepping technique. An integral-based filter is introduced to estimate the non-structured model uncertainties and constant external disturbance. Another filter is also introduced to estimate the velocity of the visual target that tracking of the moving target can be achieved, while most existing virtual camera based DIBVS approaches only consider static targets [18]–[26], [28], [29]. The controller is proven to be able to uniformly asymptotically stabilize the image feature errors to zeros. In order to show the efficacy of the proposed scheme, numerical simulation and experimental results of tracking a moving target in a tilted plane are presented. To the best of authors' knowledge, this paper for the first time presents a DIBVS scheme for planar moving visual targets of arbitrary orientation with both theoretical analysis and experimental results.

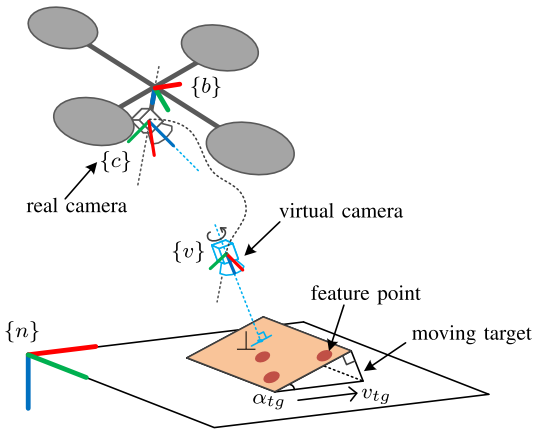

Fig. 1. A typical scenario for dynamic IBVS for quadrotors. A quadrotor tracks a tilted planar target moving with velocity v_{tg}. Frame $\{c\}$ and $\{v\}$ share the same origin, though they are depicted separately for clarity. The x, y, z axes of the coordinates are in red, green and blue (RGB), respectively.

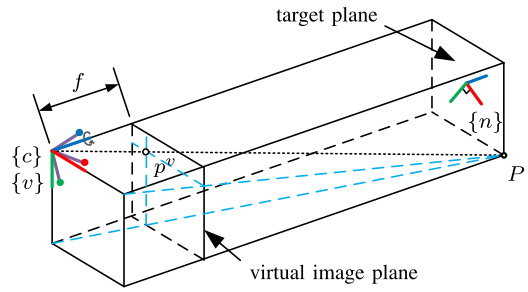

Fig. 2. Pinhole camera model. A point P is projected onto the virtual image plane with its projection p^v. The real camera frame $\{c\}$ is in light purple with colored dots. The virtual camera frame $\{v\}$ has the x, y and z axes in red, green and blue (RGB) respectively. The frame $\{v\}$ only rotates around its z-axis, which is defined to be always perpendicular to the target plane.

II. Modeling With Virtual Camera

A. Notations

Throughout the paper, for a function $f(x)$, its time derivative is denoted as $\dot{f}(x(t)) = f'(x(t))\dot{x}(t)$. For a quantity x, its estimate is denoted as $\hat{x}$, and the estimate error is $\tilde{x} = x - \hat{x}$. A rotation matrix $R_a(\alpha) \in SO(3)$ means rotating about a-axis by α. For instance, $R_z(\psi)$ means rotating about z-axis by ψ. The notation of a coordinate system is usually enclosed in curly braces, for example, the navigation frame $\{n\}$. The matrix $R^\beta_\alpha \in SO(3)$ means a rotation matrix from frame $\{\alpha\}$ to frame $\{\beta\}$. Vector $e_3 = [0, 0, 1]^T$. The map $S(\cdot) : \mathbb{R}^3 \mapsto \mathbb{R}^{3\times3}$ yields a skew symmetric matrix that verifies $S(x)y = x \times y$, for x, $y \in \mathbb{R}^3$.

B. Frame Definition

A typical scenario of dynamic IBVS is shown in Fig. 1. A navigation frame $\{n\}$ is fixed rigidly on the ground with its basis oriented north, east, and down. As in [18]–[29] for IBVS of RUAVs $\{n\}$ is assumed to be inertial considering the Earth's slow rotational motion. A body frame $\{b\}$ is oriented forward, right and down relative to the UAV with its origin at the center of mass. The attitude of the UAV is described by a set of conventional Euler angles in Z-Y-X sequence, denoted as ψ, θ and ϕ for yaw, pitch and roll, respectively. A camera frame $\{c\}$ is with its origin at the optical center of the onboard camera and basis oriented forward, right and down, respectively. A virtual camera frame $\{v\}$ is introduced with its origin fixed at the optical center of the real camera, and the virtual image plane parallel to the target plane. This means $\{v\}$ can only rotate about its own z-axis. The orientation of the planar target is described by two angles, ψ_{tg} and α_{tg}, thus its attitude relative to $\{n\}$ is $R^n_{tg} = R_z(\psi_{tg})R_y(\alpha_{tg})$.

C. Image Feature and Kinematics

As shown in Fig. 2, a commonly used pinhole camera model is used that the mapping between a point $P \in \mathbb{R}^3$ in $\{n\}$ and its pixel coordinate of p as $[u_x, u_y]^T$ is

$$Z\left[u_x, u_y, 1\right]^T = K\left[R_c^{nT} - R_c^{nT}t\right]\left[P^T, 1\right]^T \tag{1}$$

where Z is the depth, f is the focal length, $K \in \mathbb{R}^{3\times3}$ is the intrinsic matrix, R^n_c is the rotation matrix from $\{c\}$ to $\{n\}$ and t is the position of the camera in $\{n\}$. In Fig. 2, p^v is the projection point by P on the virtual image plane with coordinates as $[u_{xv}, u_{yv}]^T$ and the conversion from real image pixel coordinates p to p^v can be found in [20].

We make the following assumptions to facilitate the analysis. First, the planar visual target is assumed to be composed of $N > 1$ coplanar feature points with an inclination of $0 \leq \alpha_{tg} < \pi/2$ with respect to level ground plane. This assumption is different from the ones made in [18]–[29] where the visual target is limited on a horizontal plane. The angle $\alpha_{tg} = \pi/2$ is a critical point that the yaw of UAV has no effect on the yaw motion of $\{v\}$ resulting that the virtual image plane has only translational motion. In the case of $\alpha > \pi/2$, the virtual camera concept is applicable in the same way but the controller in section III needs slight modification in frame transformations. Second, assume we have obtained the orientation information of the target. The orientation can be estimated through using various methods such as homography decomposition [3] combining with the onboard IMU measurement and extrinsic camera calibration.

Denote the image moment features for translational motion as $s = [s_1, s_2, s_3]^T \in \mathbb{R}^3$, and for yaw motion as $s_4 \in \mathbb{R}$. The expressions for s and s_4 are [17]

$$s_1 = s_3 x_g \tag{2a}$$

$$s_2 = s_3 y_g \tag{2b}$$

$$s_3 = \sqrt{a^*/a} \tag{2c}$$

$$s_4 = \arctan(2\mu_{11}/(\mu_{20} - \mu_{02}))/2 \tag{2d}$$

where $\mu_{ij} = \sum_{k=1}^{N}(x_k - x_g)^i(y_k - y_g)^j$, $x_g = m_{10}/m_{00}$ and $y_g = m_{01}/m_{00}$, $m_{ij} = \sum_{k=1}^{N} x_k^i y_k^j$, $a = \mu_{02} + \mu_{20}$ and a^* is the desired value of a, x_k and y_k are the x- and y-coordinates of k-th point, respectively. Assume the desired centroid of the feature points as the image origin without rotation, then the desired values for s and s_4 are $s_d = [0, 0, 1]^T$ and $s_{4\,d} = 0$, respectively. Actually, the desired image moments can be set at any constant value or even a function as in [24] where the desired image moment in the height subsystem is a function of the image moments in lateral subsystem. The relation between the pixel coordinates of p^v and $[x_k, y_k]^T$ is a linear mapping defined by the camera intrinsic matrix K, that is, $[u_{xv}, u_{yv}, 1]^T = K[x_k, y_k, 1]^T$.

By taking time derivatives of Eq. (2) and rearrange, the image kinematics can be obtained as

$$\dot{s} = -S(\dot{\psi}^v e_3)s - v^v/Z^* \tag{3a}$$

$$\dot{s}_4 = -\dot{\psi}^v \tag{3b}$$

where Z^* is the desired depth, ψ^v is the rotation of $\{v\}$ around its z-axis and v^v is the velocity in $\{v\}$.

D. Dynamic IBVS Modeling

Combining the six degrees of freedom rigid body dynamics of quadrotors [33] and Eq. (3), we have the DIBVS model as

$$\dot{s} = -S(e_3)\dot{\psi}^v s - (v^v - v^v_{tg})/Z^* \tag{4a}$$

$$\dot{s}_4 = -(\dot{\psi}^v - \dot{\psi}^v_{tg}) \tag{4b}$$

$$\dot{v}^v = -S(e_3)\dot{\psi}^v v^v + R^v_n g e_3 + T^v/m - R^v_b D v^b + R^v_b b^b_1 + R^v_n b^n_2 \tag{4c}$$

where $v^v \in \mathbb{R}^3$ and $v^v_{tg} \in \mathbb{R}^3$ are the velocities of the UAV and the visual target expressed in $\{v\}$; ψ^v_{tg} is the yaw angle of the target in $\{v\}$; R^v_n and R^v_b are rotation matrices from $\{n\}$ and $\{b\}$ to $\{v\}$, respectively; $T^v = -R^n_b T e_3 \in \mathbb{R}^3$ is the thrust vector in $\{v\}$ where T is a non-negative scalar assumed to be always in the negative z-axis direction of $\{b\}$; m is the mass and $J \in \mathbb{R}^{3\times3}$ is the moment of inertia; g is the gravitation constant; D is the drag coefficient; $b^b_1 \in \mathbb{R}^3$ is the model uncertainty of the UAV in $\{b\}$ and $b^n_2 \in \mathbb{R}^3$ is the constant external disturbance in $\{n\}$. The model in Eq. (4) is similar to the ones in [18]–[29] whereas the difference is that the input signals for the dynamics or kinematics in Eq. (4) are expressed in $\{v\}$.

To extend DIBVS of rotorcraft to arbitrarily oriented planar target, we need to convert the input signals in Eq. (4) into real physical signals which is achieved by finding the relationship between ψ^v and ψ. Recall the definition of $\{v\}$ described in section II-B that it can only rotate about its z-axis, therefore, its orientation R^n_v can be expressed as

$$R^n_v = R^n_{tg} R_z(\psi^v) \tag{5}$$

Now consider the following equality

$$R_z(\Delta\psi)R^n_{tg} = R^n_{tg} R_z(\psi^v) R_y(\theta^v) R_x(\phi^v) \tag{6}$$

where $\Delta\psi = \psi - \psi_{tg}$ and θ^v and ϕ^v are the Euler angles. The pre-multiplication of $R_z(\Delta\psi)$ in the left-hand-side of Eq. (6) represents a rotation about z-axis of $\{n\}$ which can be decomposed to three sequential rotations as shown in the right-hand-side. The relation between ψ and ψ^v are thus built. By rearranging Eq. (6), we can define a new rotation matrix R_e as

$$R_e = {R^n_{tg}}^T R_z(\Delta\psi) R^n_{tg}$$

and ψ^v can be obtained as the 'Z' element of the Z-Y-X Euler angles of R_e.

In practice, we first obtain the pixel coordinates of the target in the real image plane and then convert it into $\{v\}$ as

$$\left[u_{xv}, u_{yv}, 1\right]^T = K R^{nT}_v R^n_b R^b_c K^{-1} k_z \left[u_x, u_y, 1\right]^T$$

where k_z is the depth ratio, and the fact $R^{nT}_v = R^{nT}_v R^n_c R^{nT}_c = R^{nT}_v R^n_b R^b_c R^{nT}_c$ is used. The matrix R^b_c denotes the rotation from $\{c\}$ to $\{b\}$. For a rigidly fixed camera, R^b_c is a constant matrix during IBVS. For a pan-tilt camera, R^b_c can be easily obtained through the roll and pitch angles of the camera or from the gimbal system.

III. Dynamic IBVS Control

Two main control loops are adopted for the dynamic IBVS system including the outer visual servoing loop and the inner attitude control loop as shown in Fig. 3. The outer loop receives the image moment features s and s_4, and outputs the thrust command T to the quadrotor and attitude reference R_d to the inner loop. The inner loop takes R_d and outputs the desired torque τ to the quadrotor. The real camera captures the image of the target and exports the coordinates of the feature points $[u_x, u_y]^T$ to the virtual camera. Combined with the attitude of the quadrotor, we compute the image moment features on the

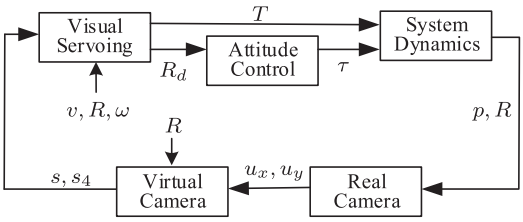

Fig. 3. Control Structure.

virtual image plane, and finally transmit it to the visual servoing controller.

The outer-loop controller is designed using adaptive backstepping technique based on Lyapunov stability criterion. Set the first error as the image moment feature error in translational direction denoted as z_1, thus we have $z_1 = s - s_d$ with its time derivative given by $\dot{z}_1 = \dot{s}$, where $s_d = [0, 0, 1]^T$. Assume the target velocity is v_{tg} in $\{n\}$ with its estimate denoted as $\hat{v}_{tg}$, then the estimate error is $\tilde{v}_{tg} = v_{tg} - \hat{v}_{tg}$. Define the first Lyapunov function as $V_1 = z_1^T z_1/2$. By substituting the kinematics of image features in Eq. (4a) and the estimate error, the time derivative of V_1 can be obtained as $\dot{V}_1 = z_1^T z_2 - z_1^T k_1 z_1 + z_1^T R_n^v \tilde{v}_{tg}/Z^*$ where $k_1 \in \mathbb{R}^{3\times3}$ is a positive definite gain matrix, z_2 is the new backstepping error defined as $z_2 = k_1 z_1 - S(e_3)\dot{\psi}^v s - (v^v - R_n^v \hat{v}_{tg})/Z^*$. Employing Eq. (4a), the time derivative of z_2 is

$$\begin{aligned}\dot{z}_2 = &(k_1 - S(e_3)\dot{\psi}^v)\left[z_2 - k_1 z_1 + R_n^v \tilde{v}_{tg}/Z^*\right] \\ &- (\dot{v}^v + S(e_3)\dot{\psi}^v R_n^v \hat{v}_{tg} - R_n^v \dot{\hat{v}}_{tg})/Z^* - S(e_3)\ddot{\psi}^v s\end{aligned} \tag{7}$$

The second Lyapunov function is defined as $V = V_1 + (z_2^T z_2 + \tilde{b}_1^T k_{b1}^{-1}\tilde{b}_1 + \tilde{b}_2^T k_{b2}^{-1}\tilde{b}_2 + \tilde{v}_{tg}^T k_{tg}^{-1}\tilde{v}_{tg})/2$, where $k_{b1}, k_{b2}, k_{tg} \in \mathbb{R}^{3\times3}$ are gain matrices defined to be positive definite. Applying Eq. (7) and the dynamic equation in Eq. (4c), the time derivative of V_2 is obtained as

$$\begin{aligned}\dot{V} = &- z_1^T k_1 z_1 - z_2^T k_2 z_2 - \tilde{b}_1^T (k_{b1}^{-1}\dot{\hat{b}}_1 + R_b^{v\,T} z_2/Z^*) \\ &+ z_2^T \Big\{ z_1 + k_2 z_2 - S(e_3)\ddot{\psi}^v s \\ &+ (k_1 - S(e_3)\dot{\psi}^v)(z_2 - k_1 z_1) \\ &- \Big[-S(e_3)\dot{\psi}^v v^v + R_n^v g e_3 + T^v/m - R_b^v D v^b \\ &+ R_b^v \hat{b}_1 + R_n^v \hat{b}_2 + S(e_3)\dot{\psi}^v R_n^v \hat{b}_{tg} \\ &- R_n^v \dot{\hat{v}}_{tg}\Big]/Z^* \Big\} - \tilde{b}_2^T (k_{b2}^{-1}\dot{\hat{b}}_2 + R_n^{v\,T} z_2/Z^*) \\ &+ \tilde{v}_{tg}^T \left[\frac{R_n^{v\,T} z_1 R_n^{v\,T} (k_1 - S(e_3)\dot{\psi}^v)^T z_2}{Z^*} - k_{tg}^{-1}\dot{\hat{v}}_{tg}\right]\end{aligned} \tag{8}$$

where again $k_2 \in \mathbb{R}^{3\times3}$ is a positive definite gain matrix. At this stage, we can set the control law of the virtual thrust vector T_d^v and the integral-based estimators for $\hat{v}_{tg}$, $\hat{b}_1$ and $\hat{b}_2$ to render $\dot{V}_2$ negative semi-definite, so as to stabilize the error system. By applying the Lyapunov-like analysis using Barbalat's lemma [34], we can prove the uniform asymptotic stability of the error system. This is summarized in the following theorem.

Theorem 1: The origin of the system described in Eq. (4) is uniformly asymptotically stable with the following desired thrust in the virtual camera frame $\{v\}$:

$$\begin{aligned}T_d^v = m\Big\{ &Z^*\left[z_1 + k_2 z_2 + (k_1 - S(e_3)\dot{\psi}^v)(z_2 - k_1 z_1)\right] \\ &+ S(e_3)\dot{\psi}^v v^v - R_n^v g e_3 + R_b^v D v^b - R_b^v \hat{b}_1 - R_n^v \hat{b}_2\Big\} \\ &- S(e_3)\dot{\psi}^v R_n^v \hat{v}_{tg} + R_n^v \dot{\hat{v}}_{tg}\end{aligned} \tag{9}$$

and the integral-based estimators :

$$\dot{\hat{v}}_{tg} = k_{tg} R_n^{v\,T}\left[z_1 + (k_1 - S(e_3)\dot{\psi}^v)\right]/Z^* \tag{10a}$$

$$\dot{\hat{b}}_1 = -k_{b1} R_b^{v\,T} z_2/Z^* \tag{10b}$$

$$\dot{\hat{b}}_2 = -k_{b2} R_n^{v\,T} z_2/Z^* \tag{10c}$$

Proof: Consider the lower bounded Lyapunov function, $V = (z_1^T z_1 + z_2^T z_2 + \tilde{b}_1^T k_{b1}^{-1}\tilde{b}_1 + \tilde{b}_2^T k_{b2}^{-1}\tilde{b}_2 + \tilde{v}_{tg}^T k_{tg}^{-1}\tilde{v}_{tg})/2$. The time derivative of V is obtained as in Eq. (8). Apply the actuation law in Eq. (9) and the estimators in Eq. (10), we have $\dot{V}$ in the the closed-loop as $\dot{V} = -z_1^T k_1 z_1 - z_2^T k_2 z_2 \leq 0$, which is a negative semi-definite function of error states and estimation errors. As long as there exist errors z_1 and z_2, $\dot{V}$ is strictly negative definite. Since the non-autonomous property of the whole system, we apply Lyapunov analysis using Barbalat's lemma to analyze the stability. The positive semi-definiteness of V and negative semi-definiteness of $\dot{V}$ imply that $V(t) \leq V(0)$, and therefore, the error states z_1 and z_2, and the estimate errors $\tilde{b}_1$, $\tilde{b}_1$ and $\tilde{v}_{tg}$ are bounded. To check the uniform continuity of $\dot{V}$, we take the time derivative of $\dot{V}$ which is $\ddot{V} = -z_1^T(k_1^T + k_1)\dot{z}_1 - z_2^T(k_2^T + k_2)\dot{z}_2$. This implies $\ddot{V}$ is bounded since z_1 and z_2 are shown above to be bounded. Therefore, $\dot{V}$ is uniformly continuous. Together with the properties of V and $\dot{V}$, we apply the Barbalat's lemma to prove the convergence of $\dot{V}$ to zero as $t \to \infty$. Observing that $\dot{V}$ is a function of the error states, consequently, we conclude the error states z_1 and z_2 tend to the origin. ■

The control law in Theorem 1 gives the desired thrust in $\{v\}$, which needs to be further transformed into $\{b\}$ as

$$T_d^b = -T e_3 = R_v^b T_d^v = R_b^{n\,T} R_v^n T_d^v \tag{11}$$

where R_v^n is obtained from Eq. (5). Rearranging Eq. (11), the desired attitude of the quadrotor $(R_b^n)_d$ can be obtained as

$$(R_b^n)_d e_3 = -R_v^n T_d^v/T \tag{12}$$

Note the Eq. (12) only constrains the desired attitude $(R_b^n)_d$ to two degrees of freedom. To fully constrain the attitude, we add the desired yaw angle to the constraints. The controller for yaw angle is in Eq. (13). Together with the orthogonal property of the rotation matrix we are able to find the desired attitude through the value of $(R_b^n)_d e_3$.

To control the yaw motion, we define the image feature error in yaw subsystem as $z_\psi = s_4 - s_{4\,d}$, where $s_{4\,d} = 0$. Denote $\dot{\psi}_{tg}^v$ as h, and the controller for yaw motion is

$$\psi^v = \int \left(-k_\psi z_\psi - \hat{h}\right) dt \tag{13}$$

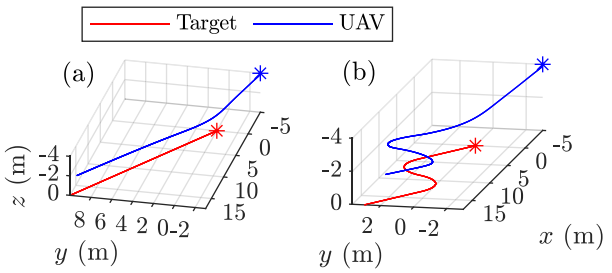

Fig. 4. Trajectories of dynamic IBVS for (a) constant moving target and (b) maneuvering target. The symbol * represents starting positions.

with the estimator as $\dot{\hat{h}} = k_h z_\psi$ where $k_\psi \in \mathbb{R}$ and $k_h \in \mathbb{R}$ are positive gains. This can be proved by simply using the Lyapunov theory. The Lyapunov function is constructed as $V = (1/2)z_\psi^2 + (1/2k_h)\tilde{h}^2$ and its time derivative is computed as $\dot{V} = z_\psi \dot{z}_\psi + (1/k_h)\tilde{h}\dot{\tilde{h}}$. Substituting the controller for yaw motion in Eq. (13) and its estimator, we can have $\dot{V} \leq 0$. Applying the Barbalat's lemma, it can be proved similarly to Theorem 1 that the error dynamics of z_ψ and $\tilde{h}$ are uniformly asymptotically stable.

IV. Simulation Results

A simulation system is built to validate the proposed control algorithm. To make the simulation environment closer to reality, the following major points are considered. (i) Electric motor hysteresis is modeled by a first order transfer function as $1/(T_m s + 1)$ where T_m is a time constant. (ii) The thrust for each rotor is saturated and the total thrust cannot exceed 1.5 times the gravity. (iii) Drag, mainly coming from the blade flapping and induction, is represented by an effective lumped parameter model as $-TK_r v^b$ where $K_r = \text{diag}(\bar{c}, \bar{c}, 0)$ and $\bar{c}$ is a linear drag coefficient [35]. (iv) Target orientation estimate noise is added, thus, R_{tg}^n becomes $R_{tg}^n = R_z(\psi_{tg} + \sigma_\psi)R_y(\alpha_{tg} + \sigma_\alpha)$ where the estimate noise is assumed to be normally distributed as $\sigma_\psi, \sigma_\alpha \sim \mathcal{N}(0, 0.1)$. (v) Noise is also added in the feedback of quadrotor's velocity as $\sigma_v \sim \mathcal{N}(0, 0.05)$. The controller gains are kept the same for all the simulations to show the robustness of the algorithm. The performance of the controller can be improved if the gains are finely tuned.

A. For Constant Moving Target

In this scenario, the quadrotor is commanded to follow a constant moving visual target with velocity $v_{tg} = [\sqrt{3}/2, 1/2, 0]^T$ m/s starting from the origin. The heading of the target is $\psi_{tg} = \pi/6$ and the inclination is $\alpha_{tg} = \pi/6$. The quadrotor's initial position is $[-5, -3, -4]$ m with velocity and attitude both being zeros. The desired depth is set to be $Z^* = 2$ m. The gain matrices for the IBVS controller are $k_1 = 0.5I$, $k_2 = 3I$, $k_{b1} = I$, $k_{b2} = 0.5I$, $k_{tg} = 0.5I$, where $I \in \mathbb{R}^{3\times 3}$ is an identity matrix.

The trajectories for the target and the UAV are shown in Fig. 4(a). As expected, the UAV has successfully tracked behind and above the target and the path is sufficiently smooth. Quantitatively, from the desired depth Z^* we can compute the desired relative position using simple trigonometric function

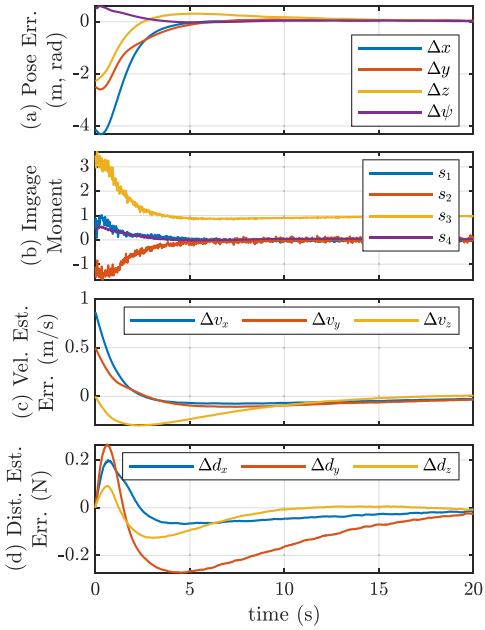

Fig. 5. Time evolution of the variables in dynamic IBVS for a constant moving target: (a) relative pose errors; (b) image moments; (c) estimation errors of target velocity and (d) constant disturbance estimation error.

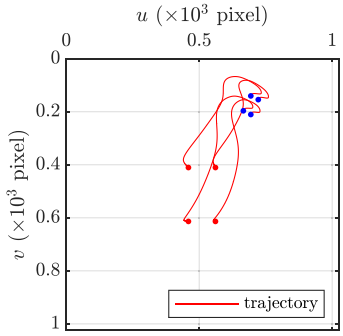

Fig. 6. Paths for the four image feature points in dynamic IBVS for a constant moving target. The blue and red dots represent starting and ending points, respectively.

as $Z^*[\sin\alpha_{tg}\cos\psi_{tg}, \sin\alpha_{tg}\sin\psi_{tg}, \cos\alpha_{tg}]^T$. Thus, we can compute the difference between the desired relative position and the actual ones as shown in Fig. 5(a). The UAV has almost reached the desired relative position after 5 s. The time evolution of the image moments is shown in Fig. 5(b). The variables s_1, s_2 and s_4 gradually converge to zero and s_3 converges to 1 which validates the IBVS law in Eq. (9) and shows the stability of the IBVS system. The jitter of the curves are mainly caused by the target orientation estimation errors. Figure. 5(c) and (d) show the estimation errors of target velocity and disturbance, respectively, where the disturbance is set as 0.2 N in x-axis of $\{n\}$. The convergence to zeros implies the efficacy and accuracy of the estimators. Note for high-speed tasks the target velocity estimator may absorb part of the aerodynamic drag modeled in Eq. (4c) whereas for low speed tasks, such absorption can be neglected. Finally, the paths of the four image points in $\{v\}$-frame are depicted in Fig. 6. The image resolution is set as 1024×1024 pixel. At the beginning, the feature points are

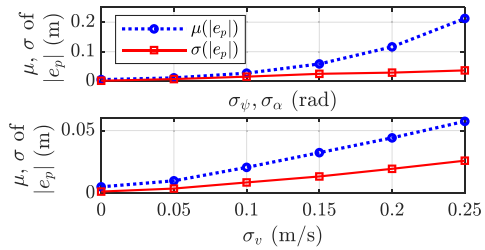

Fig. 7. Robustness analysis with varying standard deviation of noise.

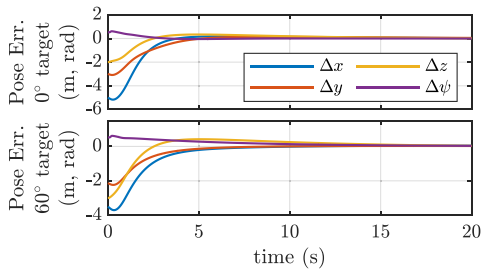

Fig. 8. Time evolution of the relative pose errors in dynamic IBVS for constant moving tilted planar targets with tilt angles of 0 and $\pi/3$, respectively.

very close to each other because the UAV is far away from the target. Meanwhile, the feature points seem rotated due to the misalignment of the headings. At the end, the aircraft reaches the desired relative pose, thus the four feature points are almost at the center of the image.

To show the robustness of the algorithm, an array of simulations were conducted using varying standard deviation of noise of target orientation estimation and quadrotor's velocity feedback, as shown in Fig. 7. The norm of the position error $|e_p|$ from 40 to 50 s is selected as performance index. As expected, larger noise results in larger tracking error but the position error is within 0.2 m even with noise with 0.25 rd standard deviation. Also, $|e_p|$ is less sensitive to the accuracy of velocity estimate.

B. For Maneuvering Target

The scenario for tracking a maneuvering target is shown in Fig. 4(b). In this scenario, the target is assumed to be a nonholonomic ground vehicle with the same initial velocity but a lateral acceleration with a magnitude of 0.2 m/s^2 is applied to the normal to the velocity from 5 to 10 s. Then the vehicle turns right with the same acceleration from 10 to 15 s behaving like an "s" maneuver. The yaw of the vehicle is in the same direction as the velocity. The simulation result reveals that though the IBVS controller is designed for constant moving target, it endows UAVs the ability to track a maneuvering target. Even when the acceleration of the target is larger than 0.2 m/s^2, the quadrotor is still able to track as long as the target is in the FoV of the camera.

C. For Arbitrarily Tilted Target

Finally, we demonstrate the efficacy of the controller for arbitrarily oriented targets. Two more cases are shown in Fig. 8 where the scenario is the same as in section IV-A except for the tilt angles being 0 and $\pi/3$, respectively. The convergence

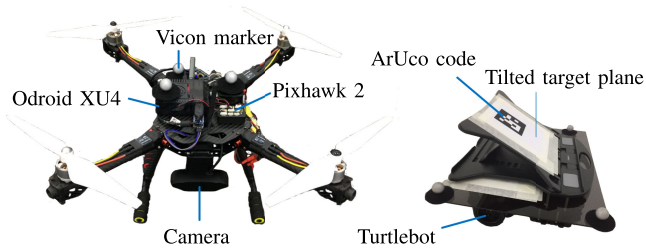

Fig. 9. Built-up quadrotor system and the tilted planar target (ArUco marker).

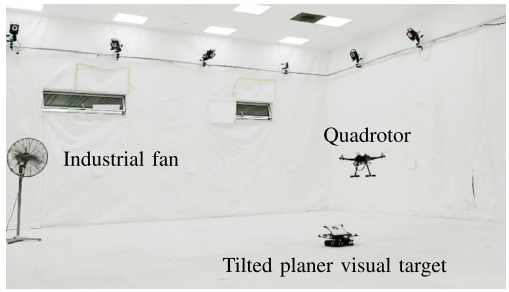

Fig. 10. Experiment scenario.

pattern for the relative pose errors are similar to the ones shown in Fig. 5(a) where the tilt angle is set as $\pi/6$, and similar behaviors can be observed for other cases. These results validate the proposed algorithm in the perspective of adaptation to the change of tilt angles.

V. Experimental Results

A. Experimental Setup

To evaluate the efficacy of the proposed IBVS scheme, we have built a quadrotor UAV and a UGV as shown in Fig. 9. The quadrotor has a motor-to-motor distance of 450 mm with weight of 1.75 kg. A Pixhawk with PX4 firmware [36] is selected as the autopilot. A monocular camera is mounted at the lower front of the quadrotor with a focal length of 3.67 mm and FoV of 70.42° horizontally and 43.3° vertically. The Odroid XU4 is selected as the companion computer of the autopilot as it is lightweight and has sufficient computational power. The visual target is an ArUco marker placed on a ground robot known as the Turtlebot with a tilt angle of 30°. An industrial standing fan is used to generate wind disturbance during the motion of the quadrotor in order to evaluate the robustness of the system. The whole experiment scenario is shown in Fig. 10.

The graphical representation of the overall architecture is illustrated in Fig. 11. The whole system is implemented in ROS framework. The companion computer plays a role as an agent to disseminate the UAV state information and visual information from the camera to the ground control computer. At the same time, it sends the control commands including the reference attitude and thrust computed in the ground control computer to the autopilot. The position information of the UAV and UGV from the motion capture system are used as ground truth for evaluating the performance of the proposed algorithm. It is also used to replace the GPS to aid linear velocity estimation

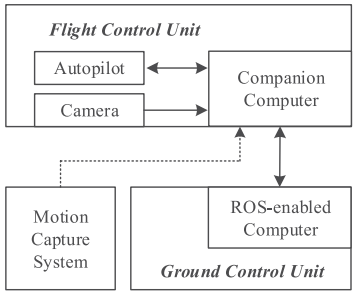

Fig. 11. Experiment architecture.

of the quadrotor through an extended Kalman filter, which is required for implementing the controller in Eq. (9). Actually, the velocity estimation using a motion capture system can be replaced by other velocity estimation methods, such as monocular visual-inertial state estimation [37] and dynamics-inertial state estimation [38], that the proposed visual servoing scheme can be implemented purely based on an onboard inertial measurement unit and a monocular vision sensor.

B. Experimental Results

The task of the UAV is to track a 30° tilted planar target with a relative distance of 1 m. The corresponding desired pixel coordinates of the four feature points are selected as

$$(u_x u_y)_d = \begin{bmatrix} 293 & 292 & 341 & 341 \\ 215 & 264 & 264 & 216 \end{bmatrix}$$

where each column represents the pixel coordinate of one feature point. The desired relative position can be calculated by trigonometry which is $[-0.5, 0, 0.5\sqrt{3}]^T$ m. The positive definite gain matrices are tuned to be $k_1 = 0.1I$, $k_2 = 3I$, $k_{b1} = k_{b2} = 0.1I$ and $k_{tg} = 0.5I$. The gains in the yaw subsystem are $k_\psi = 0.5$ and $k_h = 0.1$.

The trajectories of the positions of the UGV (Turtlebot), the UAV, and their relative position are shown in Fig. 12. From the time period of 0-14 s, the quadrotor is manually controlled to a random position such that the UAV is able to capture the visual target. At the time of $t = 14$ s, the quadrotor enters a fully autonomous mode where the dynamic IBVS controller is activated . The target is kept stationary from 14-29 s, the quadrotor gradually moves to the desired relative position. At $t = 29$ s, the Turtlebot starts to move mainly in x-direction with a speed around 7.5 cm/s which can be roughly estimated from the first sub-figure of Fig. 12. The low speed of the Turtlebot is to ensure it always stays in the limited FoV. Nevertheless, this issue can be mitigated by using omnidirectional cameras. The last sub-figure of Fig. 12 shows the relative position. Comparing with the desired one $[-0.5, 0, 0.5\sqrt{3}]^T$ m, the error is within around 13 cm. The velocity estimate of the target in x- and y-direction is shown in Fig. 13. The estimated velocity in x-direction $\hat{v}_{tg}$ converges to around 7.5 cm/s and that in y-direction stays around zero which is close to the ground truth shown in Fig. 12.

The time evolution of the image moments s and s_4 are shown in Fig. 14. In the time period of 0-8 s, the camera does not catch the view of the target because UAV stays on the ground and

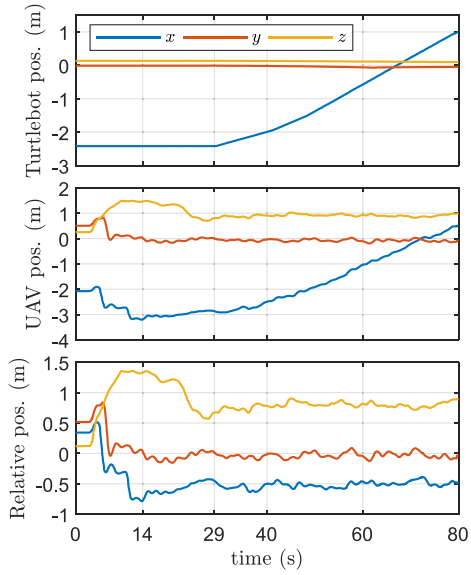

Fig. 12. Time evolution of the Turtlebot's position, UAV's position and their relative position. IBVS controller is activated at time $t = 14$ s. The visual target (Turtlebot) starts to move at time $t = 29$ s.

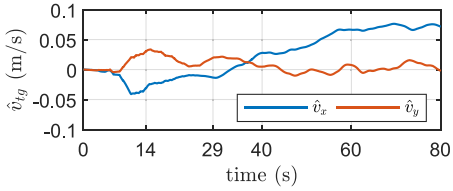

Fig. 13. Estimate of the target velocity in x and y direction.

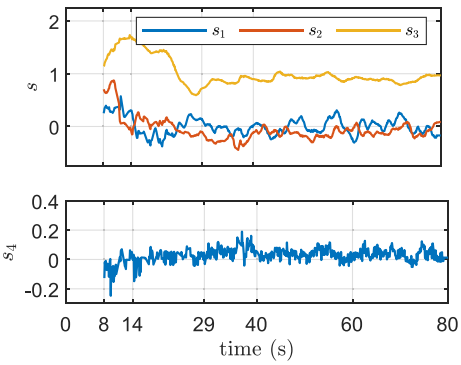

Fig. 14. Time trajectory of the image moment $s = [s_1, s_2, s_3]^T$ and s_4 where the desired values are $s_d = [0, 0, 1]^T$ and $s_{4\,d} = 0$.

with a large horizontal distance from the target. Thus the image moment features does not exist and are not shown in Fig. 14. After the IBVS controller is activated at time $t = 14$ s and later at $t = 29$ s when the target starts to move, the image moments are evolving as expected where for s_3 it converges to one, and for s_1, s_2 and s_4 they stay around zero. Hence, we can conclude that the motion control objective has been achieved for a moving target despite a noisy environment. The video of the experiment is available at https://youtu.be/22V6hpXLhJ0.

VI. Conclusion

By introducing a conversion of real images to virtual camera images, for the first time the perspective image moment features have been extended to visual servoing of underactuated rotorcraft UAVs to arbitrarily oriented planar target with theoretical analysis. The proposed visual servoing method is robust to system uncertainties and also able to track moving targets with integral-based filters. The proposed IBVS scheme has been applied in tracking a visual planar target. Expected performance has been observed both from the simulation and experimental results. Future works include robust visual feature detection and tracking, aggressive motion control by using the proposed virtual camera concept, and image feature path planning in the 2-D virtual image space to enable the aerial vehicle to fly in a designed 3D motion profile, for example, obstacle avoidance motion profile.

References

[1] S. Hutchinson, G. D. Hager, and P. I. Corke, "A tutorial on visual servo control," *IEEE Trans. Robot. Automat.*, vol. 12, no. 5, pp. 651–670, Oct. 1996.

[2] F. Chaumette and S. Hutchinson, "Visual servo control. i. basic approaches," *IEEE Robot. Automat. Mag.*, vol. 13, no. 4, pp. 82–90, Dec. 2006.

[3] P. Corke, *Robotics, Vision and Control: Fundamental Algorithms in MATLAB*. Berlin, Germany: Springer, 2017.

[4] M. Hua, T. Hamel, P. Morin, and C. Samson, "A control approach for thrust-propelled underactuated vehicles and its application to VTOL drones," *IEEE Trans. Autom. Control*, vol. 54, no. 8, pp. 1837–1853, Aug. 2009.

[5] F. Chaumette and S. Hutchinson, "Visual servo control. *II. Advanced approaches [Tutorial]*," *IEEE Robot. Automat. Mag.*, vol. 14, no. 1, pp. 109–118, Mar. 2007.

[6] H. Xie, "Dynamic visual servoing of rotary wing unmanned aerial vehicles," Ph.D. dissertation, Dept. Electrical and Computer Engineering, University of Alberta, Edmonton, AB, 2016.

[7] P. Corke and M. Good, "Dynamic effects in high-performance visual servoing," in *Proc. IEEE Int. Conf. Robot. Automat.*, 1992, pp. 1838–1843.

[8] T. Hamel and R. Mahony, "Visual servoing of an under-actuated dynamic rigid-body system: An image-based approach," *IEEE Trans. Robot. Automat.*, vol. 18, no. 2, pp. 187–198, Apr. 2002.

[9] T. Hamel and R. Mahony, "Image based visual servo control for a class of aerial robotic systems," *Automatica*, vol. 43, no. 11, pp. 1975–1983, 2007.

[10] N. Guenard, T. Hamel, and R. Mahony, "A practical visual servo control for an unmanned aerial vehicle," *IEEE Trans. Robot.*, vol. 24, no. 2, pp. 331–340, Apr. 2008.

[11] R. Mahony, P. Corke, and T. Hamel, "Dynamic image-based visual servo control using centroid and optic flow features," *J. Dyn. Syst., Meas., Control*, vol. 130, no. 1, 2008, Art. no. 011005.

[12] F. L. Bras, R. Mahony, T. Hamel, and P. Binetti, "Dynamic image-based visual servo control for an aerial robot: Theory and experiments," *Int. J. Optomechatronics*, vol. 2, no. 3, pp. 296–325, 2008.

[13] O. Bourquardez, R. Mahony, N. Guenard, F. Chaumette, T. Hamel, and L. Eck, "Image-based visual servo control of the translation kinematics of a quadrotor aerial vehicle," *IEEE Trans. Robot.*, vol. 25, no. 3, pp. 743–749, Jun. 2009.

[14] D. Guo and K. K. Leang, "Image-based estimation, planning, and control for high-speed flying through multiple openings," *Int. J. Robot. Res.*, vol. 39, no. 9, pp. 1122–1137, 2020.

[15] D. Lee, T. Ryan, and H. J. Kim, "Autonomous landing of a VTOL UAV on a moving platform using image-based visual servoing," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2012, pp. 971–976.

[16] H. Jabbari, G. Oriolo, and H. Bolandi, "Dynamic IBVS control of an underactuated UAV," in *Proc. IEEE Int. Conf. Robot. Biomimetics*, 2012, pp. 1158–1163.

[17] O. Tahri and F. Chaumette, "Point-based and region-based image moments for visual servoing of planar objects," *IEEE Trans. Robot.*, vol. 21, no. 6, pp. 1116–1127, Dec. 2005.

[18] H. Jabbari, G. Oriolo, and H. Bolandi, "An adaptive scheme for image-based visual servoing of an underactuated UAV," *Int. J. Robot. Automat.*, vol. 29, no. 1, pp. 92–104, 2014.

[19] H. Xie, K. H. Low, and Z. He, "Adaptive visual servoing of unmanned aerial vehicles in GPS-denied environments," *IEEE/ASME Trans. Mechatronics*, vol. 22, no. 6, pp. 2554–2563, Dec. 2017.

[20] H. Xie and A. F. Lynch, "State transformation-based dynamic visual servoing for an unmanned aerial vehicle," *Int. J. Control*, vol. 89, no. 5, pp. 892–908, 2016.

[21] H. Xie, A. F. Lynch, K. H. Low, and S. Mao, "Adaptive output-feedback image-based visual servoing for quadrotor unmanned aerial vehicles," *IEEE Trans. Control Syst. Techonol.*, vol. 20, no. 3, pp. 1034–1041, May 2019.

[22] D. Zheng, H. Wang, J. Wang, S. Chen, W. Chen, and X. Liang, "Image-based visual servoing of a quadrotor using virtual camera approach," *IEEE/ASME Trans. Mechatronics*, vol. 22, no. 2, pp. 972–982, Apr. 2017.

[23] X. Zhang, Y. Fang, X. Zhang, J. Jiang, and X. Chen, "Dynamic image-based output feedback control for visual servoing of multirotors," *IEEE Trans. Ind. Informat.*, vol. 16, no. 12, pp. 7624–7636, Dec. 2020.

[24] J. Li, H. Xie, R. Ma, and K. Low, "Output feedback image-based visual servoing of rotorcrafts," *J. Intell. Robotic Syst.*, vol. 93, no. 1-2, pp. 277–287, 2019.

[25] H. Xie and A. F. Lynch, "Input saturated visual servoing for unmanned aerial vehicles," *IEEE/ASME Trans. Mechatronics*, vol. 22, no. 2, pp. 952–960, Apr. 2017.

[26] H. Xie, Z. He, and D. Veitch, "Disturbance observer-based visual servoing for multirotor unmanned aerial vehicles," *at-Automatisierungstechnik*, vol. 66, no. 3, pp. 258–267, 2018.

[27] H. J. Asl, M. Yazdani, and J. Yoon, "Vision-based tracking control of quadrotor using velocity of image features," *Int. J. Robot. Automat.*, vol. 31, no. 4, pp. 301–309, 2016.

[28] A. Abdessameud and F. Janabi-Sharifi, "Image-based tracking control of VTOL unmanned aerial vehicles," *Automatica*, vol. 53, pp. 111–119, 2015.

[29] X. Zhang, Y. Fang, X. Zhang, J. Jiang, and X. Chen, "A novel geometric hierarchical approach for dynamic visual servoing of quadrotors," *IEEE Trans. Ind. Electron.*, vol. 67, no. 5, pp. 3840–3849, May 2020.

[30] H. De Plinval, P. Morin, P. Mouyon, and T. Hamel, "Visual servoing for underactuated VTOL UAVs: A linear, homography-based framework," *Int. J. Robust Nonlinear Control*, vol. 24, no. 16, pp. 2285–2308, 2014.

[31] R. Ozawa and F. Chaumette, "Dynamic visual servoing with image moments for a quadrotor using a virtual spring approach," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2011, pp. 5670–5676.

[32] R. Ozawa and F. Chaumette, "Dynamic visual servoing with image moments for an unmanned aerial vehicle using a virtual spring approach," *Adv. Robot.*, vol. 27, no. 9, pp. 683–696, 2013.

[33] J. J. Craig, *Introduction to Robotics: Mechanics and Control*. Pearson Education India, 2009.

[34] J.-J. E. Slotine and W. Li, Applied nonlinear control. Prentice hall Englewood Cliffs, *NJ*, 1991.

[35] M. Bangura, F. Kuipers, G. Allibert, and R. Mahony, "Non-linear velocity aided attitude estimation and velocity control for quadrotors," in *Proc. 16th Australas. Conf. Robot. Automat.*, 2015.

[36] L. Meier, D. Honegger, and M. Pollefeys, "PX4: A node-based multithreaded open source robotics framework for deeply embedded platforms," in *Proc. IEEE Int. Conf. Robot. Automat.*, 2015, pp. 6235–6240.

[37] D. Abeywardena, S. Kodagoda, G. Dissanayake, and R. Munasinghe, "Improved state estimation in quadrotor MAVs: A novel drift-free velocity estimator," *IEEE Robot. Automat. Mag.*, vol. 20, no. 4, pp. 32–39, Dec. 2013.

[38] T. Qin, P. Li, and S. Shen, "VINS-Mono: A robust and versatile monocular visual-inertial state estimator," *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 1004–1020, Aug. 2018.