Going through a window and landing a quadrotor using optical flow

Zhiqi Tang[1] Rita Cunha[2] Tarek Hamel [3] Carlos Silvestre[4]

Abstract— This paper considers the problem of controlling a quadrotor to go through a window and land on a planar target, using an image-based controller, with only a camera and an Inertial Measurement Unit (IMU) as sensors. The maneuver is divided into two stages: crossing the window and landing on the target plane. For the first stage, a control law is proposed that guarantees that the vehicle will not collide with the wall containing the window and will go through the window with non-zero velocity along the direction orthogonal to the window, keeping at all times a safety distance with respect to the window edges. For the landing stage, the proposed control law ensures that the vehicle achieves a smooth touchdown, keeping at all time a positive height above the target plane. For control purposes, the centroids of the images of a collection of landmarks (corners) for both the window and the target are used as position measurement. The translational optical flow relative to the wall, window edges, and target plane is used as velocity cue. To achieve the proposed objective, no direct measurements of position or velocity are used and no explicit estimate of the height above the target plane or of the distance to the wall is required. Simulation results are provided to illustrate the performance of the proposed controller.

I. INTRODUCTION

Many tasks require Unmanned Aerial Vehicles (UAVs) to land on a final destination or perform intermediate landings for recharging or battery change during long missions, which typically take place in complex environments. Since GPS information is not always available, an efficient way of tackling these scenarios is to consider the integration of a vision system with an IMU[1], given that both sensors are passive, light weight, and have low power consumptions [2].

The use of vision for obstacle avoidance has been a active topic of research. The work in [3] uses methods with pushbroom stereo to automatically avoid obstacles. In [1], the authors address the problem of state estimation, control, and planning for aggressive flight through a narrow window using only a single camera and an IMU. The authors in

[4] introduce a novel setup of self-supervised learning, in which optical flow cues serve as a scaffold to learn the visual appearance of obstacles in the environment. In [5], stereo vision and optical flow are used for obstacle avoidance during the landing process, requiring estimation of velocity and depth. A method for ego-motion estimation using optical flow is proposed for UAV flying in cluttered environments[6].

In this paper, we consider the problem of landing a quadrotor on a planar target that is placed inside a room and requires the vehicle to first go through a window and only then land on the target. This scenario occurs in many practical applications, such as search and rescue in an earthquake-damaged building[7], recharging or battery change during long missions, package delivery using UAVs, etc. This work is builds on previous work reported in [8], which used optical flow to land a UAV in cluttered environments. Compared with [8], where obstacles were represented by spherical balls, the present work considers the more realistic scenario of crossing a window, while avoiding the window edges and the wall that contains the window. Compared with other vision-based methods, the proposed controller requires no estimation of 3D quantities, using directly image centroids as position measurements and optical flow as a velocity cue.

The proposed control law comprises two stages. The first accomplishes the goal of going through the window and the second the goal of landing on the target. For control purposes, collections of landmarks are used as features, which are given by the edges and corners of the window for the first stage and the corners of the landing area for the second stage. The corresponding spherical image measurements are combined into centroid vectors, which provide information about position. For velocity, it is assumed that both the target plane and the window plane have enough texture to provide reliable optical flow measurements, which are then used to compute the translational optical flow relative to the window plane, window edges, and target plane. With no direct measurement of position or velocity, the control law proposed for going through the window ensures that no collision with the wall or window edges will occur and the vehicle will align with the center line orthogonal to the window, crossing it with non-zero velocity. The control law proposed for landing on the target is an improvement with respect to the one used in [9], with the centroid vector now directly given by the image centroid. The proof of convergence is also simplified by relying on an alternative storage function that considers all directions of motion in conjunction, requiring no ultimate bound on the height above the target and keeping the same property that the height will remain positive for all time.

The body of the paper consists of six parts. Section II presents the kinematics and dynamics of a quadrotor vehicle. Section III introduces the environment and presents the image features that are used in the control law. Section IV proposes two control laws, one for landing in an obstacle-free environment and the other for flying through the window. The combination of these two control laws in a practical case is also presented in this section. Section V shows simulation results obtained with the proposed controller. Conclusions are presented in Section VI.

II. MODELING

Consider a quadrotor vehicle equipped with an IMU and a camera. To describe the motion of the vehicle, two reference frames are introduced: the inertial reference frame $\{I\}$ fixed to the earth surface and the body-fixed frame $\{B\}$ attached to the quadrotor's centre of mass. Let $R = {}_B^I R \in SO(3)$ denote the rotation matrix that transforms vectors from the frame $\{B\}$ to $\{I\}$ and $\xi \in \mathbb{R}^3$ the position vector of the vehicle expressed in $\{I\}$. Let $v \in \mathbb{R}^3$ denote the translational velocity expressed in $\{I\}$ and $\Omega \in \mathbb{R}^3$ the angular velocity expressed in $\{B\}$. The kinematics and dynamics of the quadrotor vehicle are then described as

$$\begin{cases} \dot{\xi} = v \\ \dot{v} = F \end{cases} \tag{1}$$

$$\begin{cases} \dot{R} = RS(\Omega) \\ I\dot{\Omega} = -S(\Omega)I\Omega + \Gamma \end{cases} \tag{2}$$

where $F \in \mathbb{R}^3$ expressed in $\{I\}$ denotes the force applied to the vehicle, $\Gamma \in \mathbb{R}^3$ expressed in $\{B\}$ is the torque and I is the tensor of inertia. The matrix $S(\Omega)$ denotes a skew-symmetric matrix. The force F combines thrust and gravity components and can be described as

$$F = ge_3 - \frac{1}{m}TRe_3 \tag{3}$$

where g is the gravitational acceleration, m is the mass of the vehicle, the variable T represents the thrust magnitude, and $e_3 = [1\ 0\ 0]^\top$.

III. ENVIRONMENT AND IMAGE FEATURES

Using an image-based controller to perform a landing maneuver and to avoid the wall and the window edges requires the definition of adequate images features. Some assumptions regarding the environment and the setup are established.

Assumption 1: The camera is attached to the center of mass of the vehicle so that the camera reference frame coincides with the body-fixed frame.

Assumption 2: The landing target lies on a textured plane called the target plane. The normal to the target plane, with coordinates in the inertial frame denoted by η_t is aligned with the gravitational direction ($\eta_t \equiv e_3$).

Assumption 3: The target window has the form of a square and lies on a textured wall called the window plane, with normal in the inertial frame denoted by η_w.
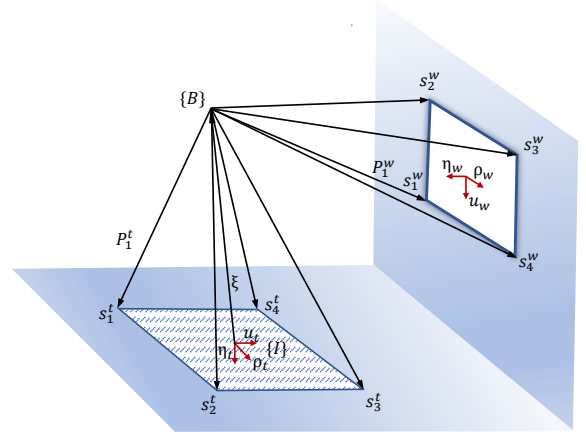

Fig. 1.   Target plane and window plane.

Assumption 4: The orientation matrix R is known in order to represent all image information as well as the dynamics of the system in the inertial frame.

Assumption 5: Assume that the control input is F or more specifically the thrust vector (TRe_3). This implicitly means that any desired orientation in the direction e_3, $R_de_3 = \frac{mge_3 - F_d}{T}$ (with $T = \|mge_3 - F_d\|$), can be obtained after a short transient. In the language of automatic control there is a high gain inner-loop controller that ensures that $Re_3 \approx R_de_3$[10]. Consequently, the control input of the translational dynamics described by (1) is directly given by $F = F_d$.

Both target plane and window plane are placed in the environment, as shown in Figure 1. It is assumed that the vehicle is able to recognize the target and the window by landmarks on the target and the edges of the window respectively. The background texture on both target plane and window plane can be used to obtain information about velocity relative to the background and also to avoid collision with both planes. With the initial position outside the room containing the target, the main objective of the controller is to drive the vehicle through the window and land it safely on the target, avoiding collisions along the way.

A. Image features of the target plane

The target on the target plane is shown in Figure 1. The axes of $\{I\}$ are given by (u_t, ρ_t, η_t), where $\rho_t = \eta_t \times u_t$, and the origin of $\{I\}$ is placed at the center of the target. As shown in Figure 1, s_i^t denotes the position of ith corner of the target relative to the inertial frame. Note that $\eta_t^\top s_i^t = 0$ and $\sum_i s_i^t = 0$. Define the position vector of ith corner of the target relative to $\{B\}$ as

$$P_i^t = s_i^t - \xi. \tag{4}$$

Using the spherical projection model for a calibrated camera, the spherical image points of target's corners can be expressed as

$$p_i^t = \frac{P_i^t}{\|P_i^t\|}.$$

A visual feature that encodes information about the position of the vehicle relative to the target plane is the centroid
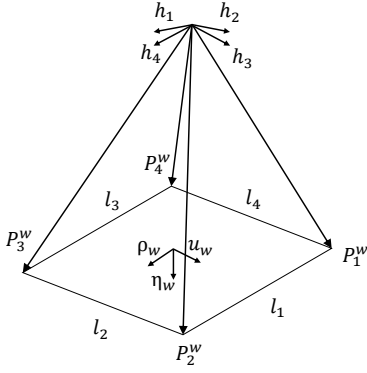
Fig. 2. Window.

vector

$$q_t := -\frac{1}{n_t} \sum_{i=1}^{n_t} p_i^t \tag{5}$$

where n_t is the number of the corners.

B. Image features of the window plane

A rectangular window with r of width is placed on a textured wall as shown in the Figure 1. The axes representing the window are given by (η_w, ρ_w, u_w), where $\rho_w = u_w \times \eta_w$. As shown in Figure 1, s_i^w denotes the position of ith corner of the target relative to the inertial frame. Similar to Section A, the spherical image points of the corners of the window can be expressed as

$$p_i^w = \frac{P_i^w}{\|P_i^w\|} \tag{6}$$

where $P_i^w = s_i^w - \xi$. Let l_i denotes the ith edge of window as shown in Figure 2. Using the image of l_i and by exploiting the fact that the window has a rectangular shape, it is straightforward to get the directions u_w and ρ_w and consequently η_w. As described in [11], in binormalized Euclidean Plucker coordinates, l_i can be represented by its unit direction u_w (resp. ρ_w) and the unit direction h_i, normal to the plane defined by the origin of the camera/body-fixed frame and the line l_i. The unit vector h_i can be obtained directly from the images[11]. Using the fact that l_1 and l_3 (resp. l_2 and l_4) are parallel in the inertial frame, we get the measure of the direction u_w (resp. ρ_w) from the following relationships:

$$\rho_w = \pm\frac{h_1 \times h_3}{\|h_1 \times h_3\|} \qquad u_w = \pm\frac{h_2 \times h_4}{\|h_2 \times h_4\|}. \tag{7}$$

Define a visual centroid vector $q := \frac{1}{n_w}\sum_{i=1}^{n_w} p_i^w$, then the normal vector to the window plane can be obtained by

$$\eta_w = \pm\frac{u_w \times \rho_w}{\|u_w \times \rho_w\|}. \tag{8}$$

The direction of η_w is decided such that $\eta_w^\top q(0) > 0$.
Let

$$d_o := \eta_w^\top P_w^i \tag{9}$$

denote the distance along the direction η_w between the vehicle and the window plane.

We can also extract the direction to the nearest point to each edge

$$p_e^i = \pm(h_i \times \rho_w), i = \{1,3\}, \quad p_e^i = \pm(h_i \times u_w), i = \{2,4\}.$$

Define the closest point from the vehicle to the edges of the window as P_e. Then, its direction $p_e = \frac{P_e}{\|P_e\|}$ is given by

$$p_e = \arg\max_{\substack{p_e^i \\ i=1,..,4}} \{|\eta_w^\top p_e^i|\}. \tag{10}$$

Let

$$d_e := \|P_e\| = \sqrt{d_o^2 + \|\pi_{\eta_w} P_e\|^2} \tag{11}$$

denote the closest distance from the vehicle to the edges of the window, where π_x denotes the projection operator defined for any $x \in \mathbb{S}^2 := \{x \in \mathbb{R}^3 : x^\top x = 1\}$ as $\pi_x = I_3 - xx^\top$. Then define the region $\mathscr{W}$ such that

$$\mathscr{W} := \{\xi : \|q(\xi)\| \leq \varepsilon\}, \tag{12}$$

where $\varepsilon > 0$. The bound ε must be such that for $\xi \in \mathscr{W}$, the condition $\|\frac{1}{n_w}\sum_{i=1}^{n_w} P_w^i\| < \frac{r}{2} - \varepsilon$ also holds, meaning that the region defined by $\mathscr{W}$ does not contain the window edges.

The visual feature that encodes information about the position of the vehicle relative to the window and is used for feedback is given by

$$q_w =: -\frac{1}{n_w}\sum_{i=1}^{n_w} p_i^w \left[(1 - \alpha_w(t))\frac{1}{\eta_w^\top p_i^w} + \alpha_w(t)\frac{\eta_w^\top p_e}{\eta_w^\top p_i^w}\right], \tag{13}$$

where n_w ($n_w = 4$) is the number of corners of the window and the weight $\alpha_w(t)$ is given by

$$\alpha_w(t) = \begin{cases} 0 & \text{, if } \|q\| > \varepsilon \\ \frac{1}{\delta}(\varepsilon - \|q\|), & \text{, if } \varepsilon - \delta \leq \|q\| \leq \varepsilon \\ 1 & \text{, if } \|q\| < \varepsilon - \delta, \end{cases} \tag{14}$$

which ensures that there is continuity in the visual feature $q_w(t)$. Noting that $\eta_w^\top p_e = \frac{d_o}{d_e}$, and substituting (6) into (13), q_w can be rewritten as

$$q_w(t) = (1 - \alpha_w(t))\frac{\chi(t)}{d_o(t)} + \alpha_w(t)\frac{\chi(t)}{d_e(t)} \tag{15}$$

where χ denotes the position of the vehicle relative to the window's center, such that

$$\chi = -\frac{1}{n_w}\sum_{i=1}^{n_w} P_i^w = \xi - \sum_{i=1}^{n_w} s_i^w, \tag{16}$$

and $\sum_{i=1}^{n_w} s_i^w$ is constant.

C. Image Kinematics and Translational Optical Flow

The kinematics of any observed points on the target plane can be written in the inertial frame as:

$$\dot{P}_t = -v \tag{17}$$

where P_t denotes a point on the textured ground of the target plane, not to be confused with P_t^i, which refers to the ith corner of the target. So the kinematics of the corresponding image point $p_t = \frac{P_t}{\|P_t\|}$ can be written as

$$\dot{p}_t = -\pi_{p_t}\frac{v}{\|P_t\|} \tag{18}$$

For the point on the textured ground of target plane, define the height of the vehicle above the target plane $d_t := \eta_t^\top P_t$ and, equation (18) can be rewritten as

$$\dot{p}_t = -\cos\theta_t \pi_p \frac{v}{d_t} \qquad (19)$$

where $\cos\theta_t = \frac{d_t}{\|P\|} = \eta_t^\top p_t$. It can be shown that the translational optical flow used for landing is

$$\phi_t(t) = \frac{v(t)}{d_t(t)}, \qquad (20)$$

which is obtained from the integral of $\dot{p}_t$ in the normal direction to the observed surface over a solid angle [9].

Similarly, we can exploit the kinematics of observed image points from the textured surface of the window plane, to obtain the following optical flow measurement

$$\phi_w = (1 - \alpha_w(t))\frac{v(t)}{d_o(t)} + \alpha_w(t)\frac{v(t)}{d_e(t)} \qquad (21)$$

which will be used as velocity cue for going through the window.

IV. CONTROLLER DESIGN

A. Landing in obstacle free environment

Theorem 1: Consider the system (1) subject to the control input

$$F_t = -k_1^t q_t - k_2^t \phi_t. \qquad (22)$$

Then for any initial condition such that $d_t(0) > 0$ and for some positive gains k_1^t and k_2^t the following assertions hold
1) the height $d_t(t)$ is positive and bounded for all the time,
2) $d_t(t)$ and $\dot{d}_t(t)$ converge to zero asymptotically,
3) the states $(\xi(t), v(t))$ are bounded and converge to zero asymptotically.
Proof: See Appendix A. ∎

B. Going through the center of the window

To accomplish the goal of going through the window, while avoiding the wall and window edges, we propose to use the following control law

$$F_w = \sigma(t)(-k_1^w \pi_{\eta_w} q_w - k_2^w \pi_{\eta_w} \phi_w - k_3^w \eta_w(\eta_w^\top \phi_w + \phi^*)), \qquad (23)$$

with

$$\sigma(t) = \begin{cases} 0 & , \text{ if } \eta_w^\top q \le 0 \\ 1 & , \text{ if } \eta_w^\top q > 0, \end{cases} \qquad (24)$$

yielding $F_w = 0$ for $d_o \le 0$, which indicates the vehicle already crossed the window. Note that for $\eta_w^\top q > 0$, the resulting closed-loop system can be written as

$$\begin{cases} \dot{\chi} = v \\ \dot{v} = -k_1^w \pi_{\eta_w} \frac{\chi}{d_w} - k_2^w \pi_{\eta_w} \frac{v}{d_w} - k_3^w \eta_w(\eta_w^\top \frac{v}{d_w} + \phi^*), \end{cases} \qquad (25)$$

where d_w is such that

$$\frac{1}{d_w} = ((1 - \alpha_w)\frac{1}{d_o} + \alpha_w \frac{1}{d_e}) \qquad (26)$$

which according to the definition of α_w given in (14) yields

$$d_w = \begin{cases} d_o, & \text{if } \|q\| > \varepsilon \\ \frac{d_o d_e}{(1-\alpha_w)d_e + \alpha_w d_o}, & \text{if } \varepsilon - \delta \le \|q\| \le \varepsilon \\ d_e, & \text{if } \|q\| < \varepsilon - \delta \end{cases} \qquad (27)$$

Theorem 2: Consider the system (1) with the control input given by (23). Then for any initial condition such that $d_w(0) > 0$, desired optical flow $\phi^* > 0$, and positive gains k_1^w, k_2^w and k_3^w such that $\frac{k_2^{w2}}{k_1^w} > \frac{r}{2}$, the following assertions hold:
1) the height $d_w(t)$ never crosses zero as long as $\|q\| \ge \varepsilon - \delta$;
2) the states $(\pi_{\eta_w}\chi(t), \pi_{\eta_w}v(t))$ are bounded and converge to zero exponentially as long as $\|q\| \ge \varepsilon - \delta$.
3) there is a finite time $T > 0$ when the system crosses the window $d_o(T) = 0$, with strictly negative velocity $\dot{d}_o(T)$ and inside the region where $\|q(T)\| < \varepsilon - \delta$.
Proof: See Appendix B. ∎

C. Landing on the target plane after crossing the window

The double goal of crossing the window and landing on the target can be achieved by simply applying the control laws F_w and F_t in sequence, with an adequate trigger to switch from F_w to F_t. Under the assumption that the initial position of the vehicle is such that it must cross the window to reach the landing target, the trigger to start using F_t is given by the condition $\eta_w^T q(t) \le 0$, which is equivalent of having $d_o(t) \le 0$.

By Theorem 2, we know that $d_o(t) = 0$ can only be reached when the vehicle position is inside $\mathcal{W}$ and consequently inside the window area. Moreover, when the crossing occurs the velocity $\dot{d}_o(t)$ is strictly negative, which implies that these are adequate initial conditions to start the landing maneuver using F_t. To avoid any chattering behavior due to noise or external disturbances, the switching from F_w to F_t is triggered only once by the first occurrence of $\eta_w^T q(t) \le 0$.

V. SIMULATION

In this section, simulation results are shown to illustrate the behavior of the closed-loop system using the designed controller. A high-gain inner-loop controller based on backstepping is used to control the attitude dynamics [12]. It generates the torque inputs in order for the vehicle to follow a desired thrust direction $R_d e_3$, which is provided by the outer-loop controller. In the simulation, the position of the center of the window is $[-5\ -4\ 20]$ m and the position of the center of the target is $[0\ 0\ 0]$. Different initial conditions are set as $\xi(0) = [-5.5\ -2.5\ 21]$ m with $v(0) = [0.5\ 0\ 0]$ m/s, $\xi(0) = [-4\ -5.5\ 18]$ m/s with $v(0) = [2\ 0\ 0]$ m/s and $\xi(0) = [-4.5\ -3.5\ 17]$ m with $v(0) = [2\ 1\ 2]$m/s, which means the quadrotor starts outside of the room containing the target. And gains are chosen following values $k_1^t = 3$, $k_2^t = 3$, $k_1^w = 3$, $k_2^w = 3$, $k_3^w = 2$, and $\phi^* = 0.1$ s^{-1}.

As shown in Fig. 3, with different initial positions and initial velocities, the vehicle avoids the wall successfully and aligns with the center line of the window for crossing. Finally the quadrotor lands on the center of the target.
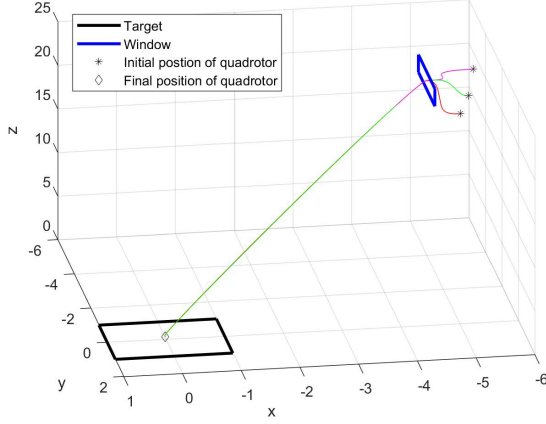
Fig. 3.   3-D plot of the quadrotor trajectories.

VI. CONCLUSION

This paper considers the problem of controlling a quadrotor to go through a window and land on planar target, using an image-based controller that relies on optical flow measurements. For control purposes, the centroids of the images of a collection of landmarks (corners and edges) for the window and target are used as position measurement. The translational optical flow relative to the wall, window edges, and target plane is used as velocity measurement, meaning that no direct measurements or explicit estimates of 3-D position or velocity are used. With the initial position outside the room containing the target, the proposed control law guarantees that the quadrotor aligns with the center line orthogonal to the window, crosses it with non-zero velocity and finally lands on the planar target successfully avoiding collisions with the wall and edges of the window. A proof of convergence of the overall control scheme is provided and simulation results show the effectiveness of the proposed controller. Future work will focus on implementing and experimentally testing the proposed solution, taking into consideration the presence of disturbances and measurement noise.

APPENDIX

A. Proof of Theorem 1

Proof: The proof follows a reasoning very similar to that of Theorem 1 in [8]. Recalling (1) and applying the control input (22), we can write the closed-loop system as

$$\begin{cases} \dot{\xi} = v \\ \dot{v} = -k_1^t q_t(\xi) - k_2^t \dfrac{v}{d_t}. \end{cases} \tag{28}$$

Before proceeding to the proof of item 1), we define a positive definite storage function $\mathscr{L}_2(\xi, v)$ and show that if $d_t(t)$ remains positive, $\mathscr{L}_2$ is negative semi-definite, which implies that the solutions remain bounded for all $t \geq 0$. Let $\mathscr{L}_2$ be given by

$$\mathscr{L}_2(\xi, v) = k_1^t \mathscr{L}_1(\xi) + \frac{1}{2}\|v\|^2 \tag{29}$$

where $\mathscr{L}_1(\xi)$ is the radially unbounded function

$$\mathscr{L}_1(\xi) = \frac{1}{n_t} \sum_{i=1}^{n_t} (\|P_i^t(\xi)\| - \|P_i^t(0)\|). \tag{30}$$

To show that $\mathscr{L}_1(\xi)$ is positive definite, we note that

$$\frac{\partial \mathscr{L}_1}{\partial \xi} = q_t^\top \qquad\qquad \frac{\partial^2 \mathscr{L}_1}{\partial \xi^2} = Q \tag{31}$$

where $Q = \frac{1}{n_t} \sum_{i=1}^{n_t} \frac{1}{\|P_i^t\|} \pi_{p_i^t}$ is positive definite, as long as at least two of the vectors p_i^t are non-collinear. It follows that $\mathscr{L}_1$ is a convex function of ξ, with a global minimum attained when $\frac{\partial \mathscr{L}_1}{\partial \xi} = q_t^T = 0$, or equivalently, when $\xi = 0$. Since $\mathscr{L}_1(0) = 0$ is the global minimum of the function, we can conclude that $\mathscr{L}_1(\xi)$ is positive definite. Noting that, $\dot{\mathscr{L}}_1 = k_1^t q_t^\top v$, it follows that

$$\dot{\mathscr{L}}_2 = -\frac{k_2^t}{d_t}\|v\|^2 \tag{32}$$

which is negative semi-definite as long as d_t remains positive and implies that the states $\xi(t)$ and $v(t)$ remain bounded for all $t \geq 0$.

Proof of Item 1: Using (28) and the fact that $d_t(t) = -\eta_t^\top \xi$ and $\dot{d}_t = -\eta_t^\top v$ yields

$$\ddot{d}_t = -k_2^t\left(\frac{\dot{d}_t}{d_t} + \alpha\right) \tag{33}$$

with

$$\alpha(t) = -\frac{k_1^t}{k_2^t}\eta_t^\top q_t = \frac{k_1^t}{k_2^t}\frac{1}{n_t}\sum_{i=1}^{n_t}\frac{d_t}{\|P_i^t\|}, \quad \alpha(t) \in (0,1), \forall t \tag{34}$$

Taking the integral of (33) and applying a similar analysis as in [8] (Th. 1, Item 1), we can conclude that d_t remains positive for all time.

Proof of Item 2:

Since the time evolution of $d_t(t)$ is described by (33) and $\alpha(t)$ is bounded, we can conclude that $d_t(t)$ and $\dot{d}_t(t)$ converge to zero asymptotically by direct application of the result in [8] (Th. 1, Item 2).

Proof of Item 3:

To show that $v(t)$ converges to zero, note that

$$\dot{v} = -k_2^t\frac{v}{d_t} - k_1^t q_t. \tag{35}$$

Since $d_t(t)$ is converging to zero and $q_t(t)$ is bounded, we can immediately conclude that $v(t)$ converges to zero by direct application of [8] (Lemma 3).

To show that $\xi(t)$ is also converging to zero, we can use an analysis similar to that in [8] (Th. 1, Item 4) applying it to $\|q_t(t)\|$ instead of $\|\xi(t)\|$. ∎

B. Proof of Theorem 2

Proof: We consider only the case where $\sigma = 1$, that is, the situation in which the vehicle is going through the window without collision.

Proof of Item 1: We first recall that while $\|q(t)\| > \varepsilon$ (and $\sigma = 1$), the dynamics of d_o can be written as

$$\ddot{d}_o = -k_3^w\left(\frac{\dot{d}_o}{d_o} + \phi^*\right). \tag{36}$$

Proceeding analogously to the Proof of Item 1 in Theorem 1, it is straightforward to verify that, while $\|q(t)\| > \varepsilon$, $d_o(t)$ remains positive and bounded and never crosses zero, and $d_o(t)$ and $\dot{d}_o(t)$ are converging to zero.

When $\varepsilon - \delta \leq \|q(t)\| \leq \varepsilon$, the dynamics of $d_o(t)$ becomes

$$\ddot{d}_o = -k_3^w\left((1-\alpha_w)\frac{\dot{d}_o}{d_o} + \alpha_w\frac{\dot{d}_o}{d_e} + \phi^*\right) \tag{37}$$

with α_w a continuous valued function in $(0,1)$ defined by (14). The above equation can be rewritten as

$$\ddot{d}_o(t) = -k_3^w b(t)\left(\frac{\dot{d}_o(t)}{d_o(t)} + \phi(t)\right) \tag{38}$$

with $b(t) = \frac{(1-\alpha_w(t))d_e(t)+\alpha_w(t)d_o(t)}{d_e(t)}$ and $\phi(t) = \frac{\phi^*}{b(t)}$. Define the following storage function:

$$\mathscr{S} = \frac{1}{2}\dot{d}_o^2, \text{ with } \dot{\mathscr{S}} = -k_3^w b(t)\dot{d}_o\left(\frac{\dot{d}_o(t)}{d_o(t)} + \phi(t)\right)$$

that implies that $\dot{d}_o$ is bounded as long as $d_o(t)$ is positive (*i.e.* $\sigma = 1$). Using the fact that for $\sigma = 1$ one has $d_o(t)$ is positive, and therefore $b(t)$ and $\phi(t)$ are uniformly continuous and upper and lower bounded by positive constants. Now by a direct application of [10]-(Th. 5.1), we can prove that $d_o(t)$ and $\dot{d}_o(t)$ are still converging to zero and $d_o(t)$ never crosses zero.

Proof of Item 2:
Let $F_1 := \pi_{\eta_w}F_w$, $\chi_1 := \pi_{\eta_w}\chi$ and $v_1 := \pi_{\eta_w}v$ denote the components orthogonal to η_w. From (1) and (23)

$$\begin{cases} \dot{\chi}_1 = v_1 \\ \dot{v}_1 = F_1 = -\frac{k_2^w}{d_w}\left(v_1 + \frac{k_1^w}{k_2^w}\chi_1\right). \end{cases} \tag{39}$$

Define a new state $z_2 = v_1 + \frac{k_1^w}{k_2^w}\chi_1$, and a positive definite storage function

$$\mathscr{L}_4 = \frac{1}{2}\|z_2\|^2 + \frac{1}{2}\frac{k_1^w}{k_2^w}\|\chi_1\|^2,$$

with time derivative is given by

$$\dot{\mathscr{L}}_4 = -\left(\frac{k_1^w}{k_2^w}\right)^2\|\chi_1\|^2 - \left(\frac{k_2^w}{d_w} - \frac{k_1^w}{k_2^w}\right)\|z_2\|^2. \tag{40}$$

It follows that $\mathscr{L}_4$ is decreasing provided that $0 < d_w < \frac{k_2^{w2}}{k_1^w}$.

By the choice of gains k_1^w and k_2^w such that $\frac{(k_2^w)^2}{k_1^w} > \frac{r}{2}$ and recalling that $\|\chi(t)\| < \frac{r}{2} - \varepsilon$ when $\|q(t)\| \leq \varepsilon$, we have that the condition on d_w holds when $\|q(t)\| \leq \varepsilon$. From the proof of item 1), it is already proved that as long as $\|q(t)\| > \varepsilon$, d_o (and therefore d_w) is decreasing and converging to zero which in turn implies that there is a finite time after which $d_w(t) = d_o(t) < \frac{k_2^{w2}}{k_1^w}$ and hence $\mathscr{L}_4$ becomes negative definite. This implies that the state (χ_1, v_1) is bounded and converges to zero exponentially.

Proof of Item 3: Consider now the case for which $\|q(t)\| < \varepsilon - \delta$ while $\sigma = 1$. In this case the dynamics of d_o takes the form

$$\ddot{d}_o = -k_3^w\left(\frac{\dot{d}_o}{d_e} + \phi^*\right), \tag{41}$$

Recalling that in this case the vehicle is inside the window region and hence one can ensure that d_e is lower bounded, $d_e \geq c > 0$, implying that $\ddot{d}_o \leq -\frac{k_3^w}{d_e}(\dot{d}_o + c\phi^*)$. Defining a new variable $z := \dot{d}_o + c\phi^*$, we obtain

$$\dot{z} \leq -\frac{k_3^w}{d_e}z, \tag{42}$$

and by the Comparison Lemma one has $z(t) \leq \exp(-k_3^w\int_0^t \frac{1}{d_e(\tau)}d\tau)z(0)$, and $\dot{d}_o(t) \leq -c\phi^* + o(t)$ where

$$o(t) = \exp(-k_3^w\int_0^t \frac{1}{d_e(\tau)}d\tau)(\dot{d}_o(0) + c\phi^*) \tag{43}$$

is converging to zero. This implies that there exists a finite time T after which $d_0(t) = 0$ with $\dot{d}_o(t) < 0$. ∎

REFERENCES

[1] G. Loianno, C. Brunner, G. McGrath, and V. Kumar, "Estimation, control, and planning for aggressive flight with a small quadrotor with a single camera and imu," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 404–411, 2017.

[2] S. Zingg, D. Scaramuzza, S. Weiss, and R. Siegwart, "Mav navigation through indoor corridors using optical flow," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2010, pp. 3361–3368.

[3] A. J. Barry, P. R. Florence, and R. Tedrake, "High-speed autonomous obstacle avoidance with pushbroom stereo," *Journal of Field Robotics*, vol. 35, no. 1, pp. 52–68, 2018.

[4] H. Ho, C. De Wagter, B. Remes, and G. De Croon, "Optical-flow based self-supervised learning of obstacle appearance applied to mav landing," *Robotics and Autonomous Systems*, vol. 100, pp. 78–94, 2018.

[5] K. McGuire, G. de Croon, C. De Wagter, K. Tuyls, and H. Kappen, "Efficient optical flow and stereo vision for velocity estimation and obstacle avoidance on an autonomous pocket drone," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 1070–1076, 2017.

[6] A. Briod, J.-C. Zufferey, and D. Floreano, "A method for ego-motion estimation in micro-hovering platforms flying in very cluttered environments," *Autonomous Robots*, vol. 40, no. 5, pp. 789–803, Jun 2016. [Online]. Available: https://doi.org/10.1007/s10514-015-9494-4

[7] N. Michael, S. Shen, K. Mohta, Y. Mulgaonkar, V. Kumar, K. Nagatani, Y. Okada, S. Kiribayashi, K. Otake, K. Yoshida *et al.*, "Collaborative mapping of an earthquake-damaged building via ground and aerial robots," *Journal of Field Robotics*, vol. 29, no. 5, pp. 832–841, 2012.

[8] L. Rosa, T. Hamel, R. Mahony, and C. Samson, "Optical-flow based strategies for landing vtol uavs in cluttered environments," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 3176–3183, 2014.

[9] P. Serra, R. Cunha, T. Hamel, D. Cabecinhas, and C. Silvestre, "Landing of a quadrotor on a moving target using dynamic image-based visual servo control," *IEEE Transactions on Robotics*, vol. 32, no. 6, pp. 1524–1535, 2016.

[10] B. Hérissé, T. Hamel, R. Mahony, and F.-X. Russotto, "Landing a vtol unmanned aerial vehicle on a moving platform using optical flow," *IEEE Transactions on Robotics*, vol. 28, no. 1, pp. 77–89, feb. 2012.

[11] R. Mahony and T. Hamel, "Image-based visual servo control of aerial robotic systems using linear image features," *IEEE Transactions on Robotics*, vol. 21, no. 2, pp. 227–239, April 2005.

[12] Z. Tang, L. Li, P. Serra, D. Cabecinhas, T. Hamel, R. Cunha, and C. Silvestre, "Homing on a moving dock for a quadrotor vehicle," in *TENCON 2015-2015 IEEE Region 10 Conference.* IEEE, 2015, pp. 1–6.