Integrated Visual Servoing Solution to Quadrotor Stabilization and Attitude Estimation Using a Pan and Tilt Camera

David Cabecinhas, Sérgio Brás, Rita Cunha, Carlos Silvestre and Paulo Oliveira.

***Abstract*—This paper presents an integrated visual servoing solution to the problem of quadrotor attitude estimation and stabilization based solely on image feedback from a pan and tilt camera and biased rate gyros. The solution comprises a control law for the quadrotor position and attitude, a control law for the pan and tilt camera platform, and an estimator for the vehicle attitude, whose interconnection stabilizes the quadrotor in hover above certain terrain landmarks. Lateral-longitudinal stabilization is achieved with a nested saturation control law by feedback of the image measurements, estimated attitude, and corrected angular velocity measurements and the vehicle is stabilized vertically by means of image feedback. Resorting to the input-to-state stability properties of the controllers, the quadrotor's position and attitude are shown to converge to the desired equilibrium point and, in the end, the interconnected system's closed loop is proven robust to observer estimation errors. Additionally, the pan and tilt camera is actively actuated to keep the relevant features visible in the image field of view for most operating conditions. The robustness and performance of the proposed control and estimation architecture is illustrated through both simulation and experimental results.**

I. Introduction

Over the last decades, the development of Unmanned Aerial Vehicles (UAVs) has witnessed a remarkable growth. Currently, UAVs have applications that range from inspection and maintenance of critical structures to surveillance and coastal patrol. The increasing operation of autonomous vehicles in indoor environments and occluded locations (e.g. in the vicinity of tall structures and buildings), where GPS signals are degraded or simply unavailable, calls for alternative solutions to the positioning problem. Under such constraints, local sensor measurements such as captured images [1], [2] can be exploited to provide reliable position information. Typical attitude and heading reference systems (AHRS) provide attitude estimates based on a triad of gyros, accelerometers and magnetometers. The estimates are, for the most part, accurate but degrade when the vehicle is submitted to large accelerations or when the sensors are located near strong magnetic fields, such as the ones experienced near the fast-rotating motors that equip aerial vehicles. In these particularly adverse settings, visual servoing solutions can be a suitable alternative to full state measurement via GPS and AHRS. Measurements from the cameras are not subject to electromagnetic interferences and do not rely on very feeble radio signals as is the case of GPS.

Computer vision has long been recognized as an extremely flexible technique for sensing the environment and acquiring valuable information for pose estimation and control. Awareness of this potential has brought about a widespread interest in the field of vision-based control and localization. The literature on vision-based rigid-body stabilization and estimation highlights important questions and indicates possible solutions to i) keep feature visibility along the system's trajectories for a large region of attraction [3], [4], ii) minimize the required knowledge about the 3-D model of the observed object [5], iii) guarantee convergence in the presence of camera parametric uncertainty and image measurement noise [5], iv) establish observability conditions for attitude estimation [6]. A variety of algebraic and iterative estimation methods based on point and line correspondences have been proposed (see for example [7]). In [8], an inertial navigation system aided by computer vision is used to estimate the relative position, attitude, and velocity. Algorithms for attitude estimation greatly benefit from the integration with inertial sensors, namely rate gyros and accelerometers as well as from the use of dynamic filtering and observer techniques [8], [9], [10].

Apart from rigid-body stabilization, vision-based control has been used to accomplish other tasks relying on different image features, such as straight line and curve representations [11], [12], and image centroids or higher order image moments [13]. For example, in [11], the authors propose an image-based controller to track parallel linear features for an underactuated vehicle. A controller for point stabilization based on backstepping and optical flow is presented in [14]. A follow-the-leader problem for mobile robots equipped with panoramic cameras is addressed in [15]. In [12], the authors consider the problem of steering a mobile robot to track a ground curve by controlling the shape of the curve in the image plane. In both [12] and [15], the two-dimensional nature of the problem removes depth ambiguity from the image measurements, which indicates that an extension to 3-D space

This work was supported by the Macao Science and Technology, Development Fund under Grants FDCT/048/2014/A1 and FDCT/026/2017/A1, by the University of Macau, Macao, China, under Project MYRG2015-00126-FST by the Fundação para a Ciência e a Tecnologia (FCT) through ISR under Grant LARSyS UID/EEA/50009/2013 and through IDMEC under Grant LAETA Pest-OE/EME/LA0022, and by the European Union's Horizon 2020 research and innovation programme under grant agreement No 731667 (MULTIDRONE). The work of Rita Cunha was supported by the FCT Investigator Programme IF/00921/2013.

David Cabecinhas is with the Department of Electrical and Computer Engineering, Faculty of Science and Technology of the University of Macau, Macau, China, and also with the Institute for Systems and Robotics, Instituto Superior Técnico, Universidade de Lisboa, Portugal (e-mail: dcabecinhas@umac.mo).

Sérgio Brás and Rita Cunha are with the Department of Electrical Engineering and Computer Science, and Institute for Systems and Robotics, Instituto Superior Técnico, Universidade de Lisboa, Portugal (e-mail: rita@isr.ist.utl.pt, sbras@isr.ist.utl.pt).

C. Silvestre is with the Department of Electrical and Computer Engineering of the Faculty of Science and Technology of the University of Macau, Macao, China, on leave from Instituto Superior Técnico, Universidade de Lisboa, Portugal (e-mail: csilvestre@umac.mo)

Paulo Oliveira is with LAETA - Associated Laboratory for Energy, Transports and Aeronautics, IDMEC - Institute of Mechanical Engineering and collaborator of ISR/LARSyS, both at Instituto Superior Técnico, Universidade de Lisboa, Portugal (e-mail: pjcro@isr.ist.utl.pt).

may not be straightforward. In [16] the authors proposed a visual servoing approach where the depth is observed and made available for servoing. The depth observer converges exponentially but the stability region of the closed-loop with the IBVS controller is not analytically characterized.

The quadrotor is a typical example of an underactuated vehicle ideally suited for the development and application of new control strategies due to its mechanical simplicity and maneuverability. The simplified model commonly employed for quadrotors is a 6-DoF rigid body actuated in force and torque with four inputs given by a scalar thrust and three torque inputs. These rotorcraft have drawn the attention of the control community in recent years and several approaches use nonlinear techniques, such as backstepping [17], [18], and feedforward control [19], [20], to solve the trajectory tracking problem of single vehicles.

The main contribution of this paper is an integrated visual servoing solution to position stabilization and attitude estimation for an aerial vehicle with a closed-loop stability guarantees. The attitude observer is based on rate gyro measurements and visual information about a set of landmarks placed on the terrain, which is retrieved by a camera mounted on a pan and tilt platform. The vehicle controller then uses the attitude estimate, the camera pan and tilt angles and the image coordinates of the landmarks for stabilization over the desired location.

The proposed nonlinear observer estimates the quadrotor attitude and the rate gyros bias, driving the estimation error exponentially fast to the origin. The pan and tilt camera controller compensates body rotations and has a proportional feedback term to center the features in the image. It differs from other solutions present in the literature (see e.g. [2]) as it does not require explicit estimation of the camera's position and velocity. The controller for quadrotor position follows the approach proposed in [19] and [20] and imposes a two-time scale dynamics, decoupling the vertical from the lateral-longitudinal subsystem. The vertical controller can be viewed as a time-varying proportional-derivative (PD) controller. A nested saturations control scheme is used to stabilize the lateral-longitudinal subsystem, which has a feedforward structure. In the proposed controller, only measurements from the available sensors (3-axis rate gyros and image features) and estimates from the attitude observer are used for feedback, instead of classical full-state feedback. Notwithstanding, the overall stability of the interconnected system is established based on the robustness and input-to-state stability (ISS) properties of each individual controller and the convergence rates guaranteed by the proposed observer.

A preliminary version of these results has been presented at the IEEE Conference on Decision and Control [21]. In comparison with the conference version, the present paper incorporates a more thorough literature review and revised design for the controller and observer, wherein some results have been elaborated upon and are presented more clearly. The vertical controller dependency on an external altitude sensor has been eliminated. The results section has been expanded to include experimental results and an in-depth analysis of more realistic simulation results, compared to those of the conference version.

The remainder of this paper is organized as follows. The problem formulation is presented in Section II, together with the quadrotor and the pan and tilt camera models. The attitude observer is discussed in Section III and the nonlinear controller for the pan and tilt camera platform is described in Section IV. The quadrotor vehicle controller based on image measurements is introduced in Section V, where the stability properties of the proposed feedback control architecture that includes the quadrotor and the camera controllers are shown. Simulation and experimental results are presented in Sections VI and VII, respectively, attesting the robustness and feasibility of the proposed estimation and control architecture. Finally, concluding remarks are given in Section VIII.

NOMENCLATURE

The following summarizes the notation used throughout the remainder of the paper. Bold letters (e.g. $\mathbf{a}$, $\boldsymbol{\omega}$) denote column vectors in $\mathbb{R}^n$ and bold digits are used to represent column vectors of the appropriate dimensions such as $\mathbf{0} = [0 \dots 0]^T$ and $\mathbf{1} = [1 \dots 1]^T$. The Special Orthogonal group is denoted by $\mathrm{SO}(3) = \left\{R \in \mathbb{R}^{3\times 3} : R^T R = I_3, \det(R) = 1\right\}$, where I_3 denotes the 3×3 identity matrix. The Lie algebra associated with $\mathrm{SO}(3)$ is denoted by $\mathfrak{so}(3)$ and is composed by the 3×3 skew-symmetric matrices $\mathfrak{so}(3) = \left\{K \in \mathbb{R}^{3\times 3} : K^T = -K\right\}$. The Special Euclidean group $\mathrm{SE}(3) = \mathrm{SO}(3) \times \mathbb{R}^3$ is used to express rigid body motions. The skew-symmetric operator is denoted as $S(\mathbf{x}) : \mathbb{R}^3 \to \mathfrak{so}(3)$ such that $S(\mathbf{x})\mathbf{y} = \mathbf{x} \times \mathbf{y}$, where $\mathbf{x}, \mathbf{y} \in \mathbb{R}^3$, whereas the inverse map $S^{-1}(.) : \mathfrak{so}(3) \to \mathbb{R}^3$ is defined such that $S^{-1}(S(\mathbf{x})) = \mathbf{x}$, where $\mathbf{x} \in \mathbb{R}^3$. The notation $\mathrm{diag}(\mathbf{a})$ describes a diagonal matrix formed by placing the elements of $\mathbf{a} \in \mathbb{R}^n$ in the main diagonal. The Frobenius norm of matrices is denoted as $\|M\|_F$ such that $\|M\|_F = \sqrt{\mathrm{tr}\,(M^T M)}$, where $M \in \mathbb{R}^{m\times n}$, $m, n \neq 1$, and the Euclidean norm of vectors denoted as $\|\mathbf{x}\|$, $\mathbf{x} \in \mathbb{R}^n$. We introduce $\sigma : \mathbb{R} \to \mathbb{R}$ as a *saturation function* that satisfies $|\sigma'(x)| = |\frac{d\sigma(x)}{dx}| \leq 2$ for all x, $x\sigma(x) > 0$ for all $x \neq 0$, $\sigma(0) = 0$, $\sigma(x) = \mathrm{sign}(x)$ for $|x| > 1$, and $|x| < |\sigma(x)| < 1$ for $|x| < 1$. The definition corresponds to an odd sigmoid function ranging in the interval $[-1, 1]$. Throughout this work, the time dependence of variables is omitted unless when required for the sake of clarity.

II. PROBLEM FORMULATION

In this paper, we design a control law for the quadrotor position and attitude, a control law for the pan and tilt camera platform and an estimator for the vehicle's attitude, whose interconnection stabilizes the quadrotor in hover above certain terrain landmarks.

The problem setup is illustrated in Fig. 1, where the reference frames used to derive the quadrotor and camera models are depicted. We consider a fixed inertial frame $\{\mathcal{I}\}$ and a body frame $\{\mathcal{B}\}$ attached to the vehicle's center of mass. The pose of $\{\mathcal{B}\}$ with respect to $\{\mathcal{I}\}$ is given by the pair $(R, \mathbf{p}) = ({}^{I}_{B}R\,, {}^{I}\mathbf{p}_{B})$. Attached to the aerial vehicle is a pan and tilt camera with reference frame denoted by $\{\mathcal{C}\}$. Its origin coincides with the camera's center of projection, and the z-axis

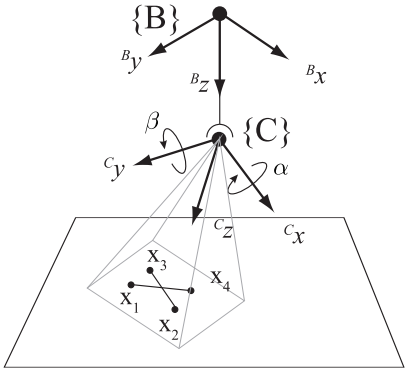

Fig. 1. Diagram of the quadrotor (frame $\{\mathcal{B}\}$), camera (frame $\{\mathcal{C}\}$), and landmarks setup ($\mathbf{x_1}$ to $\mathbf{x_4}$).

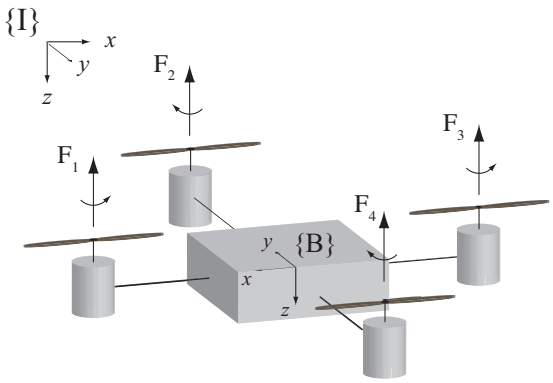

Fig. 2. Quadrotor vehicle setup showing the body frame $\{\mathcal{B}\}$ and the individual motor rotations and forces.

is aligned with the camera optical axis. The observed scene consists of four points, whose coordinates in $\{\mathcal{I}\}$ are denoted by ${}^I\mathbf{x}_i \in \mathbb{R}^3$, $i \in \{1, \ldots, 4\}$.

Special attention must be payed when defining the landmark feature's positions so as to ensure that all the rotational degrees of freedom are observable. This property is lost, for instance, when all the landmarks are collinear. The following assumption is a necessary condition to ensure that we can obtain a correct attitude estimation based on image measurements, as discussed in [22] and references therein.

Assumption 1: There are at least four landmarks of which no three are collinear.

An additional assumption is made regarding the planarity of the landmarks, necessary for the position controller.

Assumption 2: The landmarks are all coplanar and they define a plane parallel to $x - y$ plane.

Without loss of generality, the origin of $\{\mathcal{I}\}$ is assumed to coincide with the centroid of the feature points so that $\sum_{i=1}^{4} {}^I\mathbf{x}_i = 0$.

We develop our control architecture taking into account that an external state measurement solution, such as a motion capture system, is not available to the vehicle, modeling realistic conditions in GPS-denied environments. As such, the control laws are required to make use solely of the camera images and onboard sensor measurements. We consider that a triad of rate gyros is installed onboard the vehicle and that it is aligned with $\{\mathcal{B}\}$, providing measurements of the body angular velocity $\boldsymbol{\omega}_B$ corrupted by a constant unknown bias term $\mathbf{b}$, such that

$$\boldsymbol{\omega}_r = \boldsymbol{\omega}_B + \mathbf{b}, \qquad \dot{\mathbf{b}} = \mathbf{0}.$$

In summary, the available measurements are the landmarks' positions in the camera images, the pan and tilt camera angles, and the angular rate measurements corrupted by bias.

A. Camera model

As shown in Fig. 1, the camera can perform pan and tilt motions corresponding to the angles α and β, respectively. As such, the rotation matrix from $\{\mathcal{C}\}$ to $\{\mathcal{B}\}$ is given by

$$\begin{aligned} {}^B_C R &= R_{\text{pan}} R_{\text{tilt}}, \\ R_{\text{pan}} = R_x(\alpha), &\quad R_{\text{tilt}} = R_y(\beta), \end{aligned} \tag{1}$$

where $R_x(\cdot)$ and $R_y(\cdot)$ denote rotation matrices about the camera x-axis and y-axis, respectively. We denote the pose of $\{\mathcal{C}\}$ with respect to $\{\mathcal{I}\}$ by $({}^I_C R, {}^I\mathbf{p}_C) \in \text{SE}(3)$, where ${}^I_C R$ is the rotation matrix from $\{\mathcal{C}\}$ to $\{\mathcal{I}\}$ and ${}^I\mathbf{p}_C$ the position of the origin of $\{\mathcal{C}\}$ with respect to $\{\mathcal{I}\}$. Then, the 3-D coordinates of the feature points expressed in $\{\mathcal{C}\}$ can be written as

$$\mathbf{r}_i = {}^I_C R^{\,T} {}^I\mathbf{x}_i + {}^C\mathbf{p}_I,$$

for $i \in 1, \ldots, 4$ and where ${}^C\mathbf{p}_I = -{}^C_I R\, {}^I\mathbf{p}_C$. Using the perspective camera model [9], the 2-D image coordinates $\mathbf{y}_i \in \mathbb{R}^2$ of the landmark points are expressed as

$$\begin{bmatrix} \mathbf{y}_i \\ 1 \end{bmatrix} = d_i A \mathbf{r}_i, \tag{2}$$

where $A \in \mathbb{R}^{3\times3}$ is the camera calibration matrix, assumed to be known, and d_i is an unknown scalar encoding depth information and given by $d_i = (\mathbf{u}_3^T A \mathbf{r}_i)^{-1}$, where $\mathbf{u}_3 = [0\ 0\ 1]^T$.

B. Quadrotor Model

The quadrotor vehicle is modeled as a rigid body actuated in force and torque. The kinematic and dynamic equations of motion for the rigid body can be written as

$$\dot{R} = RS(\boldsymbol{\omega}_B) \tag{3}$$

$$\dot{\mathbf{p}} = R\mathbf{v} \tag{4}$$

$$\dot{\boldsymbol{\omega}}_B = -\mathbb{J}^{-1} S(\boldsymbol{\omega}_B)\mathbb{J}\boldsymbol{\omega}_B + \mathbb{J}^{-1}\mathbf{n} \tag{5}$$

$$\dot{\mathbf{v}} = -S(\boldsymbol{\omega}_B)\mathbf{v} + \frac{1}{m}\mathbf{f}, \tag{6}$$

where the position $\mathbf{p}$ is expressed in the inertial frame $\{\mathcal{I}\}$, R is the rotation matrix from $\{\mathcal{B}\}$ to $\{\mathcal{I}\}$, and the angular velocity $\boldsymbol{\omega}_B \in \mathbb{R}^3$ and the linear velocity $\mathbf{v} \in \mathbb{R}^3$ are expressed in the body frame $\{\mathcal{B}\}$. The scalar m and the matrix $\mathbb{J} \in \mathbb{R}^{3\times3}$ represent the quadrotor's mass and moment of inertia, respectively. Vectors $\mathbf{f} \in \mathbb{R}^3$ and $\mathbf{n} \in \mathbb{R}^3$ denote respectively the external force and torque expressed in the body frame. Aerodynamic drag forces due to the fuselage are neglected in light of the low speeds at which the quadrotor typically operates.

Given the geometry of the quadrotor and assuming that the forces and moments generated by each of the four rotors are

approximately given by the thrust and torque components perpendicular to the rotor disk plane, we can consider a quadrotor model such that torques can be generated in any direction and the actuation force is always aligned with the body z-axis. Figure 2 shows a sketch of the quadrotor setup, together with the force generated by each motor F_i and the direction of rotation for each propeller. The bijective correspondence between the motor forces and the total thrust T and torque $\mathbf{n} = [n_1\, n_2\, n_3]^T$ is given by (7)-(10), where k_{n_i} are constants intrinsic to the vehicle.

$$T = F_1 + F_2 + F_3 + F_4, \tag{7}$$
$$n_1 = k_{n_1}(F_4 - F_2), \tag{8}$$
$$n_2 = k_{n_2}(F_1 - F_3), \tag{9}$$
$$n_3 = k_{n_3}(F_1 + F_3 - F_2 - F_4). \tag{10}$$

The external force in body coordinates is given by

$$\mathbf{f} = -T\mathbf{u}_3 + mgR^T\mathbf{u}_3 \tag{11}$$

where g is the gravitational acceleration. The quadrotor is thus an underactuated vehicle, as evidenced by (6) and (11), making the control problem much more difficult to address when compared with the control problem for a fully-actuated vehicle. In this particular case, the force actuation only has one degree of freedom in the body frame and we are required to control the position of the vehicle $\mathbf{p} \in \mathbb{R}^3$.

III. Attitude observer

In this section, we present a nonlinear observer for the vehicle attitude and angular velocity based on the image coordinates of the landmarks and angular velocity measurements corrupted with constant bias. The nonlinear attitude observer follows [22] and is designed to match the rigid body attitude kinematics (3) by taking the form

$$\dot{\hat{R}} = \hat{R}S(\hat{\boldsymbol{\omega}}_s), \tag{12}$$

where $\hat{\boldsymbol{\omega}}_s$ is the feedback term designed to compensate for the estimation errors. The attitude and bias estimation errors are defined as $\tilde{R} = \hat{R}R^T$ and $\tilde{\mathbf{b}} = \hat{\mathbf{b}} - \mathbf{b}$, respectively. Using (3) and (12), the rotation error dynamics are given by

$$\dot{\tilde{R}} = \tilde{R}S(R(\hat{\boldsymbol{\omega}}_s - \boldsymbol{\omega}_B)). \tag{13}$$

The feedback law is a function of the angular rate measurements and the image coordinates of the landmarks. To derive it, we start by defining the following matrices

$$X = \begin{bmatrix} {}^I\mathbf{x}_1 & \cdots & {}^I\mathbf{x}_4 \end{bmatrix}, \quad Y = \begin{bmatrix} \mathbf{y}_1 & \cdots & \mathbf{y}_4 \\ 1 & \cdots & 1 \end{bmatrix},$$

where ${}^I\mathbf{x}_i$ are the 3-D coordinates of the feature points expressed in $\{\mathcal{I}\}$ and $\mathbf{y}_i$ are the corresponding 2-D image coordinates. Recall that, as discussed in Section II, without loss of generality, the origin of $\{\mathcal{I}\}$ coincides with the centroid of the feature points so that $X\mathbf{1} = \mathbf{0}$ and the landmarks belong to the $x-y$ plane. The following result allows us to establish a relationship between the image coordinates and the camera attitude.

Lemma 3: Let $\boldsymbol{\sigma} = [\sigma_1\ \sigma_2\ \sigma_3\ \sigma_4]^T \in \mathbb{R}^4 \setminus \{\mathbf{0}\}$ and $\boldsymbol{\rho} = [\rho_1\ \rho_2\ \rho_3\ \rho_4]^T \in \mathbb{R}^4 \setminus \{\mathbf{0}\}$ be such that $Y\boldsymbol{\sigma} = \mathbf{0}$, $X\boldsymbol{\rho} = \mathbf{0}$, and $\mathbf{1}^T\boldsymbol{\rho} = 0$, where $\mathbf{1} = [1\ 1\ 1\ 1]^T$. Consider that the landmarks satisfy Assumption 1 and the camera pose is such that the image is not degenerate (neither a point nor a line). Then, the depth variables d_i can be written as

$$d_i = \mu\frac{\rho_i}{\sigma_i},$$

where $\mu \in \mathbb{R}$, $\rho_i \neq 0$, and $\sigma_i \neq 0$ for $i \in \{1, \ldots, 4\}$.

Proof: See proof of [22, Lemma 1] ∎

Writing (2) in matrix form and using Lemma 3, we have

$$Y = A({}^I_C R^T X - {}^C\mathbf{p}_I\mathbf{1}^T)\mu D_\sigma^{-1} D_\rho,$$

where $D_\rho = \operatorname{diag}(\boldsymbol{\rho})$. From the feature centroid constraint $X\mathbf{1} = 0$, it follows that

$$\mu {}^I_C R^T X = A^{-1} Y D_\rho^{-1} D_\sigma (I_4 - \frac{1}{4}\mathbf{1}\mathbf{1}^T),$$

which encodes information about the attitude of the camera up to the scale factor μ. We can use the properties of the rotation matrix and the positive depth constraint $d_i > 0$ to obtain the normalized vector readings

$${}^C\bar{\mathbf{x}}_i = {}^I_C R^T\, {}^I\bar{\mathbf{x}}_i = \operatorname{sign}(\mu)\frac{\mu {}^I_C R^T\, {}^I\mathbf{x}_i}{\|\mu {}^I_C R^T\, {}^I\mathbf{x}_i\|}, \tag{14}$$

where $\operatorname{sign}(\mu) = \operatorname{sign}(\rho_i/\sigma_i)$ and ${}^I\bar{\mathbf{x}}_i = {}^I\mathbf{x}_i/||{}^I\mathbf{x}_i||$, $i = \{1, \ldots, 4\}$. Note that no discontinuity is introduced by the use of the $\operatorname{sign}(.)$ function.

We now proceed with the observer design where (14) and previous definitions are instrumental in allowing to write the observer dynamics in terms of the image location of the landmarks.

A. Observer Design

Recall that the bias in angular velocity measurements is assumed to be constant, i.e. $\dot{\mathbf{b}} = \mathbf{0}$, and consider the proposed Lyapunov function

$$V_R = \frac{\|\tilde{R} - I_3\|_F^2}{2} + \frac{1}{2k_b}\|\tilde{\mathbf{b}}\|^2 = \operatorname{tr}(I_3 - \tilde{R}) + \frac{1}{2k_b}\|\tilde{\mathbf{b}}\|^2,$$

where $k_b > 0$. From the attitude error dynamics (13) and noting that, by the properties of skew matrices,

$$\operatorname{tr}(NS(\mathbf{a})) = -S^{-1}(N - N^T)^T\mathbf{a}, \text{ for } N \in \mathbb{R}^{3\times 3},\ \mathbf{a} \in \mathbb{R}^3,$$

we obtain the time derivative

$$\dot{V}_R = \mathbf{s}_\omega^T(\hat{\boldsymbol{\omega}}_s - \boldsymbol{\omega}_B) + \frac{1}{k_b}\dot{\tilde{\mathbf{b}}}^T\tilde{\mathbf{b}}, \tag{15}$$

where

$$\mathbf{s}_\omega = R^T S^{-1}(\tilde{R} - \tilde{R}^T).$$

Consider now the attitude feedback law

$$\begin{aligned} \hat{\boldsymbol{\omega}}_s &= \boldsymbol{\omega}_r - \hat{\mathbf{b}} - k_\omega\mathbf{s}_\omega \\ &= \boldsymbol{\omega}_B - \tilde{\mathbf{b}} - k_\omega\mathbf{s}_\omega, \end{aligned} \tag{16}$$

where $k_\omega > 0$. Substituting the estimator (16) in (15) and defining the bias estimator update law

$$\dot{\hat{\mathbf{b}}} := k_b\mathbf{s}_\omega, \tag{17}$$

the Lyapunov function derivative becomes $\dot{V}_R = -k_\omega ||\mathbf{s}_\omega||^2$. Taking into account the feedback law (16) and the update law (17), the closed loop attitude error dynamics can be written as

$$\begin{aligned} \dot{\tilde{R}} &= -k_\omega \tilde{R}(\tilde{R} - \tilde{R}^T) - \tilde{R}S(R\tilde{\mathbf{b}}), \\ \dot{\tilde{\mathbf{b}}} &= k_b R^T S^{-1}(\tilde{R} - \tilde{R}^T). \end{aligned} \tag{18}$$

Exploiting results derived for linear time-varying (LTV) systems in [23], it can be shown that the trajectories of the system (18) converge exponentially fast to the desired equilibrium point. Global asymptotic stability is however precluded by topological limitations associated with the points that satisfy $||\tilde{R} - I_3||_F^2 = 8$ (for further information see e.g. [24]). The stabilization result for the proposed controller is formally stated in the sequel.

Theorem 4: Assume that $\boldsymbol{\omega}_B$ is bounded and $\dot{\mathbf{b}} = \mathbf{0}$. Then, for any initial condition satisfying

$$\frac{\|\tilde{\mathbf{b}}(t_0)\|^2}{8 - \|\tilde{R}(t_0) - I_3\|_F^2} < k_b, \tag{19}$$

the estimation error $\tilde{\mathbf{x}} = (\tilde{R}, \tilde{\mathbf{b}})$ is bounded and $\|\tilde{R}(t) - I_3\|_F^2 < 8$ for all $t \geq t_0$. Moreover, the attitude and bias estimation errors converge exponentially fast to the equilibrium point $(\tilde{R}, \tilde{\mathbf{b}}) = (I_3, 0)$ for any initial condition satisfying (19). The proof of this theorem follows a similar reasoning to the one used in the proof of [22, Theorem 1] and is therefore not presented.

Remark 5: Note that the conditions of Theorem 4 are not restrictive, since $\boldsymbol{\omega}_B$ is intrinsically bounded due to the practical limitation on the energy of the system and the condition (19) can always be satisfied inside the almost global domain of attraction by tuning the gains.

We now detail how to express the estimation laws (18) solely as a function of the image measurements and the biased gyro measurements. Consider the identity

$$QS^{-1}(N - N^T) = S^{-1}(QNQ^T - QN^TQ^T),$$

where $N \in \mathbb{R}^{3\times 3},\ Q \in \mathrm{SO}(3)$, and the relation

$${}^I_C R = R^B_C R = {}^I\bar{X}\,{}^C\bar{X}^\dagger,$$

where ${}^B_C R$ is given by (1), ${}^C\bar{X} = [{}^C\bar{\mathbf{x}}_1,\ \ldots,\ {}^C\bar{\mathbf{x}}_4,\ {}^C\bar{\mathbf{x}}_i \times {}^C\bar{\mathbf{x}}_j]$, ${}^I\bar{X} = [{}^I\bar{\mathbf{x}}_1,\ \ldots,\ {}^I\bar{\mathbf{x}}_4,\ {}^I\bar{\mathbf{x}}_i \times {}^I\bar{\mathbf{x}}_j]$, for any linearly independent ${}^I\bar{\mathbf{x}}_i$ and ${}^I\bar{\mathbf{x}}_j$, and where ${}^C\bar{X}^\dagger = {}^C\bar{X}^T({}^C\bar{X}\,{}^C\bar{X}^T)^{-1}$ is the Moore-Penrose inverse of ${}^C\bar{X}$. Using the following derivation, the feedback term $\mathbf{s}_\omega$ can be expressed as an explicit function of the sensor readings and known quantities

$$\begin{aligned} \mathbf{s}_\omega &= R^T S^{-1}(\tilde{R} - \tilde{R}^T) \\ &= S^{-1}(R^T\tilde{R}R - R\tilde{R}^TR^T) \\ &= S^{-1}(R^T\hat{R} - \hat{R}^TR) \\ &= S^{-1}({}^B_C R({}^C\bar{X}^\dagger)^T\,{}^I\bar{X}^T\hat{R} - \hat{R}^T\,{}^I\bar{X}\,{}^C\bar{X}^\dagger\,{}^B_C R^T). \end{aligned}$$

IV. Pan and tilt controller

The camera frame attitude kinematics can be described by

$${}^I_C\dot{R} = {}^I_C R S(\boldsymbol{\omega}_C),$$

where $\boldsymbol{\omega}_C \in \mathbb{R}^3$ denotes the camera angular velocity. Taking the time derivative of (1), and noting that ${}^I_C R = {}^I_B R^B_C R$, straightforward computations show that $\boldsymbol{\omega}_C$ can be written as

$$\boldsymbol{\omega}_C = {}^C_B R\boldsymbol{\omega}_B + R^T_{\text{tilt}}[\dot{\alpha}\ \dot{\beta}\ 0]^T, \tag{20}$$

where $\dot{\alpha}$ and $\dot{\beta}$ are the time derivatives of the camera pan and tilt angles, respectively.

In summary, to develop an active vision system using the camera pan and tilt degrees of freedom, we let $\bar{\mathbf{y}}$ be the image of the landmarks' centroid given by $[\bar{\mathbf{y}}^T\ 1]^T = \bar{\delta}A\bar{\mathbf{r}}$, where $\bar{\mathbf{r}} = -{}^I_C R^T\,{}^I\mathbf{p}_C$ denotes the position of $\{\mathcal{I}\}$ expressed in $\{\mathcal{C}\}$ and $\bar{\delta} = (\mathbf{u}_3^T\bar{\mathbf{r}})^{-1}$. The control objective is then to design a control law for inputs $\dot{\alpha}$ and $\dot{\beta}$ based on the measurements of $\boldsymbol{\omega}_B$ and $\mathbf{y}_i$, $i \in \{1, \ldots, 4\}$, such that $\bar{\mathbf{y}}$ approaches the center of the image plane.

A. Camera Pan and Tilt Controller

We resort to Lyapunov theory and consider the following candidate Lyapunov function

$$W = \frac{1}{2}\bar{\mathbf{r}}^T\Pi_{\mathbf{u}_3}\bar{\mathbf{r}} = \frac{1}{2}(r_x^2 + r_y^2), \tag{21}$$

where $\bar{\mathbf{r}} = [r_x\ r_y\ r_z]^T$ and $\Pi_{\mathbf{u}_3} = I - \mathbf{u}_3\mathbf{u}_3^T$ denotes the matrix representing the projection operator into the plane orthogonal to $\mathbf{u}_3$. Using the expression for $\boldsymbol{\omega}_C$ given in (20), the camera pose kinematics can be written as

$$\begin{aligned} \dot{\bar{\mathbf{r}}} &= S(\bar{\mathbf{r}})\boldsymbol{\omega}_C - \mathbf{v}_\mathbf{C} \\ &= S(\bar{\mathbf{r}})(R^T_{\text{tilt}}R^T_{\text{pan}}\boldsymbol{\omega}_B + R^T_{\text{tilt}}[\dot{\alpha}\ \dot{\beta}\ 0]^T) - \mathbf{v}_\mathbf{C}, \end{aligned} \tag{22}$$

where $\mathbf{v}_\mathbf{C}$ is the camera linear velocity. Recall that by definition $\bar{\mathbf{r}}$ coincides with the position of the landmarks' centroid and its image is given by $\bar{\mathbf{r}}$. Therefore, by guaranteeing that the Lyapunov function W is decreasing, or equivalently $[r_x\ r_y]$ is approaching the origin, we can ensure that $\bar{\mathbf{y}}$ is approaching the center of the image plane. To simplify the computations and without loss of generality, assume from now on that the camera calibration matrix $A = I$ so that $\bar{\mathbf{y}} = \begin{bmatrix} r_x & r_y \end{bmatrix}^T / r_z$.

Lemma 6: Let the camera pose kinematics be described by (22) and assume that the rigid body and camera motions are such that $r_z > 0$ and $\cos\beta \neq 0$. Consider the control law for the camera pan and tilt angular velocities given by

$$\begin{bmatrix} \dot{\alpha} \\ \dot{\beta} \end{bmatrix} = k_c \begin{bmatrix} 0 & -\frac{1}{\cos\beta} \\ 1 & 0 \end{bmatrix} \bar{\mathbf{y}} - \begin{bmatrix} 1 & 0 & -\tan\beta \\ 0 & 1 & 0 \end{bmatrix} R^T_{\text{pan}}\hat{\boldsymbol{\omega}}_B, \tag{23}$$

where $k_c > 0$ and $\hat{\boldsymbol{\omega}}_B = \boldsymbol{\omega}_r - \hat{\mathbf{b}}$. Then, the time derivative of the Lyapunov function W along the system trajectories satisfies

$$\dot{W} \leq -(k_c - \epsilon)W, \quad \forall\ \|\Pi_{\mathbf{u}_3}\bar{\mathbf{r}}\| \geq \frac{1}{\epsilon}\|\Pi_{\mathbf{u}_3}(r_z S(\mathbf{u}_3){}^C_B R(-\tilde{\mathbf{b}}) + \mathbf{v}_\mathbf{C})\|, \tag{24}$$

and $0 < \epsilon < k_c$.

Proof: By taking the time derivative of (21) and using the expressions for $\dot{\bar{\mathbf{r}}}$ given in (22), we obtain

$$\dot{W} = \bar{\mathbf{r}}^T\Pi_{\mathbf{u}_3}(r_z S(\mathbf{u}_3)\boldsymbol{\omega}_C - \mathbf{v}_\mathbf{C}) \tag{25}$$

Let $\hat{\boldsymbol{\omega}}_C$ denote the estimate of $\boldsymbol{\omega}_C$ obtained using $\hat{\boldsymbol{\omega}}_B$ in (20). Choosing $\dot{\alpha}$ and $\dot{\beta}\boldsymbol{\omega}_C$ such that

$$S(\mathbf{u}_3)\hat{\boldsymbol{\omega}}_C = -k_c\Pi_{\mathbf{u}_3}\bar{\mathbf{y}} \tag{26}$$

yields $\dot{W} = -k_c W - \bar{\mathbf{r}}^T \Pi_{\mathbf{u}_3}(r_z S(\mathbf{u}_3) {}^C_B R(-\tilde{\mathbf{b}}) + \mathbf{v_C})$ and consequently (24) holds. Substituting (20) in (26) and solving for $\dot{\alpha}$ and $\dot{\beta}$ we obtain the control law (23). ∎

Remark 7: If we apply the control law (23) to the system with state $\Pi_{\mathbf{u}_3}\bar{\mathbf{r}} = [r_x \; r_y]^T$ and interpret $\mathbf{v_C}, \tilde{\boldsymbol{\omega}}_B$ and r_z as inputs, it follows from (24) that the system is exponentially input-to-state stable (ISS). As such, the distance between the image of the centroid $\bar{\mathbf{y}}$ and the origin is ultimately bounded by $\|\Pi_{\mathbf{u}_3}(r_z S(\mathbf{u}_3) {}^C_B R(-\tilde{\mathbf{b}}) + \mathbf{v_C})/r_z\|$ and converges exponentially fast to that bound. Moreover, if $\Pi_{\mathbf{u}_3}(r_z S(\mathbf{u}_3) {}^C_B R(-\tilde{\mathbf{b}}) + \mathbf{v_C})/r_z$ converges to zero so does $\bar{\mathbf{y}}$.

The proposed control law (23) has a significant advantage over classical eye-in-hand system controllers, which are based on the inversion of the error Jacobian matrix to achieve an exponential decrease of the error. The inverse of the error Jacobian matrix for the present pan and tilt camera system is

$$J_{\mathbf{e}}^{-1} = \begin{bmatrix} 0 & \frac{1}{r_z \cos\beta - r_x \sin\beta} \\ -\frac{1}{r_z} & \frac{r_y \sin\beta}{r_z^2 \cos\beta - r_x r_z \sin\beta} \end{bmatrix},$$

where the Jacobian is computed from the equality

$$\dot{\mathbf{e}} = \dot{\bar{\mathbf{r}}} = J_{\mathbf{e}} \begin{bmatrix} \dot{\alpha} \\ \dot{\beta} \end{bmatrix} + \frac{\partial \bar{\mathbf{r}}}{\partial t},$$

with the error given by $\mathbf{e} = \bar{\mathbf{r}} - \mathbf{0}$ and $\bar{\mathbf{r}} = \begin{bmatrix} r_x & r_y \end{bmatrix}^T$. Clearly, the error Jacobian matrix has singularities not present in the proposed control law. By exploiting the structure of $\dot{W}$ in (25), we obtain a controller that still achieves exponential decay of the error for the zero-disturbance case $\mathbf{v_C} = \tilde{\mathbf{b}} = \mathbf{0}$, and whose only singularity does not depend on the vehicle's position but is intrinsic to the camera's geometry. Moreover, the control law given in (23) is effectively image-based in the sense that it uses solely the image coordinates $\bar{\mathbf{y}}$ and rate gyros measurements, whereas using the inverse of the Jacobian matrix would require reconstruction of the depth coordinate r_z.

V. Quadrotor controller

The quadrotor control objective consists of designing a control law for the quadrotor actuations T and $\mathbf{n}$, which ensures the convergence of the horizontal position in frame $\{\mathcal{I}\}$ to zero with the largest possible basin of attraction, while maintaining the landmarks visible in the image field of view and the vehicle's vertical position stable.

As sensor measurements, the image coordinates of the landmarks are available for feedback in addition to the vehicle's attitude and angular velocity estimates from the proposed observer and the pan and tilt angles of the camera. Recall also that for the attitude observer it is assumed that the relative locations of the landmarks are known.

In order to achieve the stabilization goal, the proposed controller makes use of partial information on the position and body linear velocity, which are not directly measured. The 2-D image coordinates of the landmarks' $\mathbf{y}_i$ together with the rotation matrices ${}^B_I R$ and ${}^C_B R$ provide us with means of obtaining, up to a scale factor (the altitude z), the position $\mathbf{p}$ and the body linear velocity ${}^I\mathbf{v} = [{}^I v_x \; {}^I v_y \; {}^I v_z]^T$, both expressed in the inertial frame. For that purpose, we first determine the direction of the landmarks, or more precisely their position up to a scale factor with respect to the body, expressed in the inertial frame. To simplify the necessary notation, let us introduce a new reference frame $\{\mathcal{L}\}$, with the same origin as $\{\mathcal{B}\}$ but with the orientation of $\{\mathcal{I}\}$. This new frame can be thought as of a *virtual camera* that moves along with the vehicle but keeps a steady attitude, always pointing vertically towards the ground. Recall also that, without loss of generality, we assume the camera calibration matrix $A = I$ in order to simplify the computations. This avoids cluttering the math derivations with offsets (and offset compensations) of $[c_x \; c_y]^T$ and gains of f and $1/f$, which are known intrinsic camera calibration parameters.

Let $\begin{bmatrix} x_i & y_i & z \end{bmatrix}^T = {}^I_C R \, \mathbf{r}_i$ be the coordinates of the landmarks' centroid expressed in $\{\mathcal{L}\}$. Choosing $1/z$ as the scale factor, the direction of the landmarks in frame $\{\mathcal{L}\}$ can be obtained from

$$\begin{bmatrix} x_i/z \\ y_i/z \\ 1 \end{bmatrix} = \frac{{}^L_C R \begin{bmatrix} \mathbf{y}_i^T & 1 \end{bmatrix}^T}{\mathbf{u}_3^T \, {}^L_C R \begin{bmatrix} \mathbf{y}_i^T & 1 \end{bmatrix}^T} \triangleq \begin{bmatrix} \mathbf{s}_i \\ 1 \end{bmatrix}.$$

The 2D points $\mathbf{s}_i$ can be thought of as the image coordinates of the landmarks in the *virtual camera*. Due to the landmarks positioning and orientation of the virtual camera the centroid of the landmark's virtual image points are the virtual image of the landmarks' centroid, that is,

$$\bar{\mathbf{s}} = \begin{bmatrix} \bar{x}/z \\ \bar{y}/z \end{bmatrix}. \tag{27}$$

Taking the time derivative of (27), the following relation is obtained for the vehicle velocities, expressed either in $\{\mathcal{L}\}$ or $\{\mathcal{I}\}$,

$$\begin{bmatrix} \frac{{}^I v_x}{z} - \frac{\bar{x} \, {}^I v_z}{z^2} \\ \frac{{}^I v_y}{z} - \frac{\bar{y} \, {}^I v_z}{z^2} \end{bmatrix} = \dot{\bar{\mathbf{s}}}, \tag{28}$$

where the right-hand-side time derivative is a function of variables for which measurements or estimators exist ${}^L_C R = {}^I_C R$, $\boldsymbol{\omega}_C$, $\mathbf{y}_i$, and $\dot{\mathbf{y}}_i$.

For the vertical distance we rely on the distance between any two of the features in the *virtual camera* image. Let $\mathbf{s}_{ij} = \|\mathbf{s}_i - \mathbf{s}_j\|$, for $i, j \in \{1, 2, 3, 4\}$ with $i \neq j$, be the virtual image distance between two feature points. The distance between image features is related to the height z by

$$\mathbf{s}_{ij} = \frac{\mathbf{X}_{ij}}{z}. \tag{29}$$

Differentiating both sides and rearranging one gets

$$\dot{z} = -\mathbf{X}_{ij} \frac{\dot{\mathbf{s}}_{ij}}{\mathbf{s}_{ij}^2},$$

leading to

$$\frac{\dot{z}}{z} = -\frac{\dot{\mathbf{s}}_{ij}}{\mathbf{s}_{ij}}. \tag{30}$$

Using these relations of image features with altitude and their derivatives one can recover the scaled lateral velocities from (28) through only image feature feedback as

$$\begin{bmatrix} {}^I v_x / z \\ {}^I v_y / z \end{bmatrix} = \dot{\mathbf{s}}_i - \mathbf{s}_i \frac{\dot{\mathbf{s}}_{ij}}{\mathbf{s}_{ij}}, \tag{31}$$

for any of the image features $\mathbf{s}_i$.

The proposed controller makes use of the *unit quaternions* to represent the attitude, in contrast with the rotation matrix parametrization used previously. Unit quaternions $\mathbf{q} \in \mathbb{S}^4$, are written in the form $\mathbf{q} = \begin{bmatrix} q_s & \mathbf{q}_\nu^T \end{bmatrix}^T$, where the *scalar* part $q_s \in \mathbb{R}$ is related to the rotation angle $\theta \in [0, \pi)$ and the *vector* part $\mathbf{q}_\nu = [q_{\nu 1}\ q_{\nu 2}\ q_{\nu 3}]^T \in \mathbb{R}^3$ to the axis of rotation $\boldsymbol{\nu} \in \mathbb{S}^3$ through

$$\mathbf{q}(\theta, \boldsymbol{\nu}) = \begin{bmatrix} q_s \\ \mathbf{q}_\nu \end{bmatrix} = \begin{bmatrix} \cos(\theta/2) \\ \boldsymbol{\nu} \sin(\theta/2) \end{bmatrix}.$$

In general, there is an ambiguity in the unit quaternion parametrization as $\mathbf{q}$ and $-\mathbf{q}$ represent the same attitude. However, in the present case, as the attitude controller guarantees $q_s(t) > \epsilon > 0$ for all time. Then, for all system trajectories, there is a bijective correspondence of the quaternion representation and the rotation matrix representation

The methodology adopted to address the quadrotor vehicle control problem is in line with the state feedback controller proposed in [19]. However, as the full system state is not directly available for feedback, the controller is modified to exploit the image measurements and attitude estimates to stabilize the quadrotor position at the desired location. The controller comprises a vertical stabilization law together with a lateral-longitudinal-attitude stabilization law. The latter law enforces two-time scale dynamics and decouples the lateral-longitudinal dynamics from the attitude dynamics.

A. Stabilization of the vertical error dynamics

The control objective of the vertical stabilization law is to exponentially stabilize the vehicle at some arbitrary altitude. Since the specific altitude is not relevant for the application, and given the relation (29), the reference altitude can be specified through a desired distance between image features $\mathbf{s}_{ij}^\star$ in the *virtual camera*.

The dynamic equation for the altitude,

$$m\ddot{z} = (1 - 2q_{\nu 1}^2 - 2q_{\nu 2}^2)T - mg, \tag{32}$$

is derived from the altitude definition and the linear dynamics of the vehicle system represented in (4) and (6). We propose a control law for the thrust T that drives the vehicle to a fixed altitude $z^\star$ through

$$T = \frac{mg - k_1(z - z^\star) - k_2 \frac{\dot{z}}{z}}{1 - (2q_{\nu 1}^2 + 2q_{\nu 2}^2)}, \tag{33}$$

where k_1 and k_2 are positive parameters. Most importantly, the law can be expressed in the form of an image feedback law as

$$T = \frac{mg - k_1'(1/\mathbf{s}_{ij} - 1/\mathbf{s}_{ij}^\star) + k_2' \frac{\dot{\mathbf{s}}_{ij}}{\mathbf{s}_{ij}}}{1 - (2q_{\nu 1}^2 + 2q_{\nu 2}^2)},$$

for some $k_1' > 0, k_2' > 0$. The resulting closed-loop altitude dynamics are

$$m\,\ddot{z} = -k_1(z - z^\star) - \frac{k_2}{z}\dot{z}, \tag{34}$$

which amount to a double integrator system driven by a PD controller with a time-varying derivative gain due to the nature of $z(t)$. As proved in the sequel, the closed-loop is asymptotically stable for initial conditions $z(t_0) > 0$ and it is ensured that the quadrotor altitude is always positive, preventing crashes against the ground. A subsequent choice of the attitude control law guarantees that the quadrotor never overturns, and thus $2q_{\nu 1}^2 + 2q_{\nu 2}^2 < 1$ for all time, precluding the loss of altitude control through thrust actuation. For now, we take that fact as assumption and state the following lemma, regarding the altitude control.

Lemma 8: Consider the quadrotor altitude dynamic system described by the closed-loop system comprising (32) and (33) with $k_1, k_2 > 0$. If the initial conditions fulfill $z(0) > 0$, then the control law is well defined and $z(t) > 0$ for all time, even in the presence of errors in the attitude estimates provided by the observer. Additionally, the cascade of the attitude observer and the altitude controller is exponentially stable.

Proof: Let us define the auxiliary state

$$\xi = z \exp\left(\frac{1}{k_2}\left(m\dot{z} + \int_0^t k_1(z(\tau) - z^\star)\mathrm{d}\tau\right)\right)$$

and notice that, with the imposed closed-loop dynamics (34), it has a constant value as $\dot{\xi} = 0$. Since $\xi(t) = \xi(0)$ is positive and the exponential of a number is always positive, it results that $z(t) > 0$ for all time and thus collisions with the ground are always avoided.

Asymptotic stability of $(z, \dot{z}) = (z^\star, 0)$ is established from LaSalle's invariance principle and the Lyapunov function

$$V_z = \frac{1}{2}k_1(z - z^\star)^2 + \frac{1}{2}m\dot{z}^2,$$

which has negative semi-definite time derivative

$$\dot{V}_z = -k_2\frac{\dot{z}^2}{z} \leq 0.$$

An additional consequence of the convergence of $(z, \dot{z})$ to $(z^\star, 0)$ and the constancy of the auxiliary state ξ is that the altitude is lower and upper bounded for all time and $z(t) > \epsilon$ for some $\epsilon > 0$.

Furthermore, LTV system theory asserts that the convergence is indeed exponential. Let the state $\mathbf{x} = \begin{bmatrix} \dot{z} & z - z^\star \end{bmatrix}^T$ and compute

$$\dot{\mathbf{x}} = \begin{bmatrix} A(t) & B \\ -C & 0 \end{bmatrix}\mathbf{x} = \begin{bmatrix} -\frac{k_2}{z(t)} & -k_1 \\ 1 & 0 \end{bmatrix}\mathbf{x}. \tag{35}$$

to obtain an LTV system equivalent to the altitude dynamics (34). Let $P = \frac{1}{k_1}$ and notice that

$$A^T(t)P + PA(t) = -2\frac{k_2}{k_1 z(t)} = -Q(t),$$

with $Q(t)$ bounded as $0 < q_m < Q(t) < q_M$. Under these conditions, the LTV system (35) is uniformly exponentially stable [23].

The interconnection of the attitude observer and the altitude subsystem can be regarded as a cascade of two exponentially stable systems with $A(t) = -\frac{k_2}{z(t)}$ bounded for all trajectories. In these circumstances, the cascade is also exponentially stable [25, Proposition 2.1]. Finally, impact with the ground is also avoided when the altitude subsystem is perturbed by the orientation errors. This can be established by letting $\Delta_1(t)$ be

the perturbations due to the estimation errors and considering that the state

$$\xi = z \exp\left(\tfrac{1}{k_2}\left(m\dot{z} + \int_0^t k_1(z(\tau) - z^\star) - \Delta_1(\tau)\mathrm{d}\tau\right)\right) \quad (36)$$

is constant for the perturbed vertical dynamics

$$m\,\ddot{z} = -k_1(z - z^\star) - \frac{k_2}{z}\dot{z} + \Delta_1(t).$$

To complete the proof it must be determined that the pathological cases of the finite time escape and convergence to zero or divergence to infinity are impossible for the altitude. This is established by taking into account that the state ξ in (36) is constant and $z^\star$ is strictly positive. The case of finite time escape is impossible since it implies that $z(t)$, $\dot{z}(t)$, and consequently ξ all diverge to infinity, which is in contradiction to the previously established result that the state ξ is constant. Likewise, were $z(t)$ to converge to zero, it would result on ξ converging to zero. Recalling that the attitude error converges exponentially we can conclude that $\int_0^t \Delta_1(\tau)\mathrm{d}\tau$ converges and is bounded. In that situation, the argument of the exponential would be dominated by the integral term $\int_0^t k_1(z(\tau) - z^\star)\mathrm{d}\tau$, whose argument is a strictly negative number in the event $z(t)$ tends to zero. Divergence of the altitude to infinity with time is also impossible since, if we take it as a premise, it results in a ξ state that diverges to infinity. Since ξ is proven to be constant all the aforementioned situations are impossible for the dynamic system at hand and thus, even in the presence of attitude disturbances, $z(t)$ has a lower and upper bound and converges to the desired $z^\star$. ■

B. Stabilization of the lateral and longitudinal dynamics

To stabilize the quadrotor in hover, the proposed vertical stabilizer (33) needs to be combined with a controller for the torque actuation $\mathbf{n}$, which stabilizes both the attitude and the lateral-longitudinal dynamics. In order to achieve these goals, we interpret the attitude as a *virtual control input* for the lateral-longitudinal dynamics. In this setting, the attitude follows the virtual control law with fast dynamics and a slower outer control loop generates the virtual control for the attitude so as to stabilize the lateral-longitudinal dynamics. The proposed attitude-lateral-longitudinal closed-loop simultaneously stabilizes the lateral and longitudinal dynamics and ensures the quadrotor does not overturn, that is, $q_s(t) > \epsilon$ for all time.

The lateral-longitudinal-attitude subsystem dynamics for a quadrotor vehicle, with equations of motion (4)-(5) and the thrust defined by (33), are described by the following system of equations, where quaternions are used to represent the vehicle attitude,

$$\begin{aligned}
\dot{y} &= v_y, \\
m\dot{v}_y &= d(\mathbf{q})q_{\nu 1} + m(\mathbf{q})q_{\nu 2}q_{\nu 3} + \delta_y, \\
\dot{x} &= v_x, \\
m\dot{v}_x &= -d(\mathbf{q})q_{\nu 2} + m(\mathbf{q})q_{\nu 1}q_{\nu 3} + \delta_x, \\
\dot{q}_s &= -\frac{1}{2}q^T\boldsymbol{\omega}_B, \\
\dot{\mathbf{q}}_\nu &= \frac{1}{2}(q_s I_4 + S(\mathbf{q}_\nu))\boldsymbol{\omega}_B, \\
\mathbb{J}\dot{\boldsymbol{\omega}}_B &= -S(\boldsymbol{\omega}_B)\mathbb{J}\boldsymbol{\omega}_B + \mathbf{n}.
\end{aligned} \quad (37)$$

The components x, y, v_x and v_y are written in frame $\{\mathcal{I}\}$, the attitude functions $d(\mathbf{q})$ and $m(\mathbf{q})$ are given by

$$\begin{aligned}
d(\mathbf{q}) &= \frac{2\,m\,g\,q_s}{1 - (2q_{\nu 1}^2 + 2q_{\nu 2}^2)}, \\
m(\mathbf{q}) &= -\frac{2\,m\,g}{1 - (2q_{\nu 1}^2 + 2q_{\nu 2}^2)},
\end{aligned} \quad (38)$$

and δ_x, δ_y are asymptotically vanishing signals (see Lemma 8) defined as

$$\begin{aligned}
\delta_x &= \frac{2q_{\nu 1}q_{\nu 3} + 2q_s q_{\nu 2}}{1 - (2q_{\nu 1}^2 + 2q_{\nu 2}^2)}(-k_1(z - z^\star) - k_2\frac{\dot{z}}{z}), \\
\delta_y &= \frac{2q_{\nu 2}q_{\nu 3} - 2q_s q_{\nu 1}}{1 - (2q_{\nu 1}^2 + 2q_{\nu 2}^2)}(-k_1(z - z^\star) - k_2\frac{\dot{z}}{z})).
\end{aligned}$$

The control law for the attitude subsystem is chosen as the proportional-derivative law

$$\mathbf{n} = K_P(\boldsymbol{\eta} - K_D\hat{\boldsymbol{\omega}}_B) \quad (39)$$

where $K_P > 0$ and $K_D > 0$ are design parameters and

$$\boldsymbol{\eta} = \mathbf{q}_\nu^\star - \hat{\mathbf{q}}_\nu$$

is the attitude error with $\mathbf{q}_\nu^\star$ defined as the *virtual control* input for the $x-y$ system and $\hat{\mathbf{q}}_\nu$ the vectorial part of the quaternion estimate obtained with proposed observer system (18). The quadrotor attitude estimation subsystem in closed-loop with the control feedback (39) results in the following dynamics, written in quaternion representation

$$\dot{q}_0 = -\frac{1}{2}q^T\boldsymbol{\omega}_B \quad (40)$$

$$\dot{q} = \frac{1}{2}(q_0 I_4 + S(q))\hat{\boldsymbol{\omega}}_B \quad (41)$$

$$\mathbb{J}\dot{\boldsymbol{\omega}}_B = -S(\boldsymbol{\omega}_B)\mathbb{J}\boldsymbol{\omega}_B + K_P((\mathbf{q}_\nu^\star - q) - K_D\boldsymbol{\omega}_B) + \Delta_2(t), \quad (42)$$

where the external input

$$\Delta_2(t) = K_P\tilde{\mathbf{q}}_\nu - K_D\tilde{\mathbf{b}}$$

includes the errors resulting from the observer measurements

$$\tilde{\mathbf{q}}_\nu = \mathbf{q}_\nu - \hat{\mathbf{q}}_\nu$$

and vanishes exponentially fast. According to Proposition 5.7.1 in [19], which we restate for the sake of completeness, proper tuning of the torque control law (39) ensures boundedness of the attitude subsystem trajectories and consequent stabilization of the vertical error dynamics, even in the presence of attitude and bias estimation errors.

Proposition 9: ([19, Proposition 5.7.1]) For some $0 < \epsilon < 1$, fix compact sets of initial conditions $\mathcal{Q}$, Ω for the observer estimates $\hat{\mathbf{q}}_\nu(t)$ and $\hat{\boldsymbol{\omega}}_B(t)$, respectively, such that

$$\mathcal{Q} \subset \{\hat{\mathbf{q}}_\nu \in \mathbb{R}^3 : \|\hat{\mathbf{q}}_\nu\| < \sqrt{1 - \epsilon^2}\}.$$

Then there exist $K_D^\star(\|\Delta(t)\|_\infty) > 0$ and positive numbers $K_P^\star(K_D^\star)$, $\lambda^\star(K_D^\star)$ such that, for any initial conditions $(\hat{\mathbf{q}}_\nu(0), \hat{\boldsymbol{\omega}}_B(0)) \in \mathcal{Q} \times \Omega$ and $\|\hat{\mathbf{q}}_\nu^\star(t)\| < \lambda^\star$, the trajectories of the attitude subsystem (40)-(42) are bounded and satisfy $\hat{q}_s(t) > \epsilon$ for all time.

Remark 10: A corollary of this Proposition is that the quadrotor does not overturn for initial attitude and bias estimation errors satisfying $\sqrt{\frac{1}{2}\tilde{q}_s(0)^2 + \frac{1}{2}\tilde{\mathbf{b}}(0)^T\tilde{\mathbf{b}}(0)} < \epsilon$.

To achieve convergence of the overall system, the virtual control input $\mathbf{q}_\nu^\star$ is generated from the quadrotor position and velocities by a *nested saturation* control law. Consider the new state variables

$$\boldsymbol{\zeta}_1 = \frac{1}{z}\begin{bmatrix} y \\ x \end{bmatrix},$$
$$\boldsymbol{\zeta}_2 = \frac{1}{z}\begin{bmatrix} v_y \\ v_x \end{bmatrix} + \lambda_1 \sigma\left(\frac{K_1}{\lambda_1}\boldsymbol{\zeta}_1\right) - \frac{v_z}{z}\boldsymbol{\zeta}_1,$$

where $\sigma(\mathbf{x}) = (\sigma(x_1), \ldots, \sigma(x_n))$ is a saturation function and $v_z = \dot{z}$. Notice that the states $\boldsymbol{\zeta}_1$ and $\boldsymbol{\zeta}_2$ are readily obtained from $x/z,\ y/z$ and ${}^I\mathbf{v}/z$, whose estimates can be derived from the camera sensor and attitude estimate through (27), (30) and (31).

Fix for $\mathbf{q}_\nu^\star$ the nested saturation structure

$$\mathbf{q}_\nu^\star = -P_2 \lambda_2 \sigma\left(\frac{K_2}{\lambda_2}\hat{\boldsymbol{\zeta}}_2\right), \tag{43}$$

where

$$P_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \\ 0 & 0 \end{bmatrix}$$

and $\hat{\boldsymbol{\zeta}}_2$ is the estimate for $\boldsymbol{\zeta}_2$ obtained by applying (27) with the virtual image centroid $\bar{\mathbf{s}}$ obtained using the attitude estimate $\hat{R}$. The time derivatives of the states are then

$$\dot{\boldsymbol{\zeta}}_1 = \boldsymbol{\zeta}_2 - \lambda_1 \sigma(\frac{K_1}{\lambda_1}\boldsymbol{\zeta}_1),$$
$$m\dot{\boldsymbol{\zeta}}_2 = \frac{D}{z}\left(-P_2\lambda_2\sigma(\frac{K_2}{\lambda_2}\boldsymbol{\zeta}_2) + \boldsymbol{\eta}\right) + mK_1\sigma'(\frac{K_1}{\lambda_1}\boldsymbol{\zeta}_1)\dot{\boldsymbol{\zeta}}_1 + \delta_1 + \delta_2 + \Delta_3,$$

where

$$D = \begin{bmatrix} d(\mathbf{q}) & m(\mathbf{q})q_{\nu 3} & 0 \\ m(\mathbf{q})q_{\nu 3} & -d(\mathbf{q}) & 0 \end{bmatrix},$$

the exogenous inputs δ_1 and δ_2 are given by

$$\delta_1 = \begin{bmatrix} \delta_x \\ \delta_y \end{bmatrix}/z,$$
$$\delta_2 = \frac{k_1(z - z^\star) + k_2\frac{\dot{z}}{z})}{z}\boldsymbol{\zeta}_1 + m\frac{v_z^2}{z^2}\boldsymbol{\zeta}_1 - m\frac{v_z}{z}\dot{\boldsymbol{\zeta}}_1,$$

and the errors due to the attitude estimator are encapsulated in

$$\Delta_3 = \frac{D}{z}\left(-P_2\lambda_2\left(\sigma(\tfrac{K_2}{\lambda_2}\boldsymbol{\zeta}_2) - \sigma(\tfrac{K_2}{\lambda_2}\hat{\boldsymbol{\zeta}}_2)\right) + \tilde{\mathbf{q}}_\nu\right).$$

From the exponential convergence of the altitude $(z - z^\star)$ and the attitude estimation error $\tilde{q}$ to the origin and by noting that the growth of $\|\boldsymbol{\zeta}_1\|$, $\|\boldsymbol{\zeta}_2\|$ is, at most, quadratic, we can establish that the exogenous inputs and estimation induced errors are asymptotically vanishing and converge exponentially fast to zero. From definition (38) and the attitude and vertical controllers, we have the bounds $0 < d^L \leq d(\mathbf{q}, t) \leq d^U$ and $0 < z^L < z(t) < z^U$. The following result is an adaptation of Proposition 5.7.2 and Theorem 5.7.5 in [19] and gives guarantees for the proposed quadrotor stabilization law.

Theorem 11: Let K_D be fixed according to Proposition 5.7.1 in [19] and let $K_i^\star$ and $\lambda_i^\star,\ i = 1, 2$, be such that the following inequalities are satisfied

$$\frac{\lambda_2^\star}{K_2^\star} < \frac{\lambda_1^\star}{4}, \quad 4\lambda_1^\star K_1^\star < \frac{1}{m}\frac{d^L}{z^U}\frac{\lambda_2^\star}{8}, \quad 24\frac{K_1^\star}{K_2^\star} < \frac{1}{6}\frac{d^L}{d^U}\frac{z^L}{z^U}. \tag{44}$$

Then, there exist positive numbers $K_P^\star$ and $\epsilon^\star$ such that, taking

$$\lambda_i = \epsilon^i \lambda_i^\star \text{ and } K_i = \epsilon K_i^\star,\ i = 1, 2, \tag{45}$$

for all $K_P > K_P^\star$ and $0 < \epsilon \leq \epsilon^\star$, the state trajectories of the system (37) in closed-loop with the controller defined by (33), (39) and (43) converge asymptotically to the origin for any initial condition such that $z(0) > 0$, $(x(t), v_x(t), y(t), v_y(t)) \in \mathbb{R}^4$, $(\hat{q}(0), \hat{\boldsymbol{\omega}}_B(0)) \in \mathcal{Q} \times \Omega$ and $|\tilde{q}_s(0)| < q_s(0)$.

Proof: The proof follows from the arguments in [19] where the statement is proven for constant $z(t) = Z$ and exogenous disturbances $\delta_2(t) = 0$, $\Delta(t) = [\Delta_1(t)\, \Delta_2(t)\, \Delta_3(t)]^T = 0$. The statement of Theorem 11 is shown by noting that the additional disturbances $\delta_2(t)$ and $\Delta(t)$ are asymptotically vanishing. The lateral-longitudinal subsystem does not have finite escape time and the trajectory $(\boldsymbol{\zeta}_1(t), \boldsymbol{\zeta}_2(t))$ exists and is bounded for any $t > 0$. Since the disturbance $\delta_2(t)$ is asymptotically vanishing, there exists a finite time $T^\star$ such that for $t > T^\star$ the disturbances are within the bounds for which the convergence of $(\boldsymbol{\zeta}_1, \boldsymbol{\zeta}_2)$ to the origin is ensured by using the gains in (45), satisfying (44). The remainder of the claims in the theorem statement follows identically from [19]. ■

Remark 12: At this point, it is important to notice that the vehicle controller can be obtained by feedback of the image coordinates and their derivatives, vertical coordinate, and vehicle attitude estimate for the thrust, and by feedback of the image coordinates and their derivatives, attitude estimates and angular velocities of the camera and vehicle for the torque.

Gathering the previous results regarding the pan and tilt camera, stabilization of the vertical position, attitude and lateral-longitudinal subsystems, we can now state the following theorem which summarizes the main results of the paper and corresponds to the control architecture represented in Fig. 3. The camera controller K_{camera} is given by (23) and the

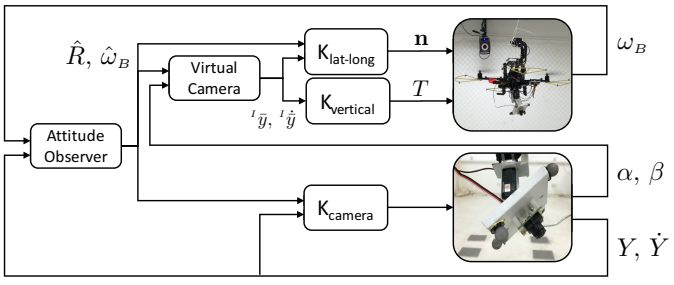

Fig. 3. Block diagram showing the interconnections of the proposed control architecture.

vertical position controller K_{vertical} by (33), respectively. The lateral-longitudinal controller $K_{\text{lat-long}}$ is described by (43) and (39), respectively. Lastly, the attitude observer estimates $\hat{R}$ and $\hat{\boldsymbol{\omega}}_B$ through (12), (16), and (17).

Theorem 13: Consider a quadrotor described by the dynamic system (3)-(6) equipped with a pan and tilt camera modeled by (1) with dynamics (23), and apply the set of camera and quadrotor controllers (23), (33), (39) and (43), using the quadrotor attitude and rate gyro bias estimator (16)-(17). Then, for any initial condition $z(0) > 0$, $(x(t), v_x(t), y(t), v_y(t)) \in \mathbb{R}^4$, $(\hat{q}(0), \hat{\boldsymbol{\omega}}_B(0)) \in \mathcal{Q} \times \Omega$ and $|\tilde{q}_s(0)| < q_s(0)$ such that the landmarks are visible in the image plane of the camera, the vehicle's position, attitude, velocities converge asymptotically to ${}^I\mathbf{p}_B = [0\ 0\ z^\star]^T$, ${}^I_BR = I_3$, $\mathbf{v}_B = \mathbf{0}$, and $\boldsymbol{\omega}_B = \mathbf{0}$, respectively, whereas the camera's velocity and image coordinates converge to $\boldsymbol{\omega}_C = \mathbf{0}$ and $\bar{\mathbf{y}} = \mathbf{0}$, respectively.

Proof: The stated result follows from Theorems 4 and 11. Theorem 11 states that convergence of the vehicle position and velocity to zero is achieved, even in the presence of attitude estimation errors. Convergence of the landmarks' centroid image coordinates to zero is achieved if the vehicle velocity and bias error converge to zero, which is guaranteed by Theorem 4. ■

VI. Simulation Results

In this section, we present the results from a simulation run of the proposed control architecture. The quadrotor is hovering to the side of four landmarks and the camera is aimed so that all the landmarks are visible but not centered in the camera image. The objective of the simulation is for the quadrotor to perform a stable hover flight above the centroid of the landmarks and for the centroid of the landmarks to be centered on the camera image.

The vehicle parameters, the control and the estimation gains are $m = 1.47\,\text{kg}$, $\mathbb{J} = \text{diag}(3.2, 3.2, 1.4)\cdot 10^{-3}\,\text{kg m}^2$, $\lambda_1 = 2$, $\lambda_2 = 0.2$, $K_1 = 4.5$, $K_2 = 1.5$, $K_P = 30$, $K_D = 6$, $k_1 = 0.024$, $k_2 = 0.024$, $k_b = 0.01$ and $k_\omega = 0.1$. The camera is simulated with an image size of 752x480 pixels, its output is quantized to integer values and its intrinsic calibration matrix is

$$A = \begin{bmatrix} 420 & 0 & 376 \\ 0 & 420 & 240 \\ 0 & 0 & 1 \end{bmatrix},$$

corresponding to $f = 420$, $c_x = 376$, $c_y = 240$, based on the calibration for a real camera with a 2.4 mm lens. The camera pan and tilt gain is $k_c = 5/f$. The landmarks were defined as the vertices of a 30x30 cm square centered at the origin. The constant additive bias for the angular velocity is $\mathbf{b} = \begin{bmatrix}1.7 & 1.2 & 2.9\end{bmatrix}^T \cdot 10^{-2}\,\text{deg}/s$. The initial quadrotor position is $\mathbf{p}(0) = \begin{bmatrix}-1 & 0.5 & -1\end{bmatrix}^T$ m and the initial pan and tilt angles are $\alpha(0) = 0\,\text{deg}$ and $\beta(0) = 22.9\,\text{deg}$, respectively.

The sample rate for angular rotation measurements and for image acquisitions was set to 100Hz. During the first 5 seconds the quadrotor is at rest and the observer is warming up. Once the quadrotor and the camera controller are turned on the vehicle moves from its initial position towards hovering at the origin at a height of approximately 1.4 m. From the time evolution of the quadrotor horizontal position we can see that the convergence is fast, occurring in around 3 seconds, from 5 s to 8 s.

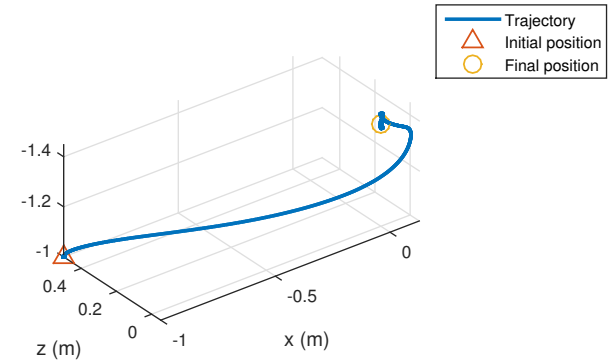

Fig. 4. Quadrotor trajectory with the visual servoing controller.

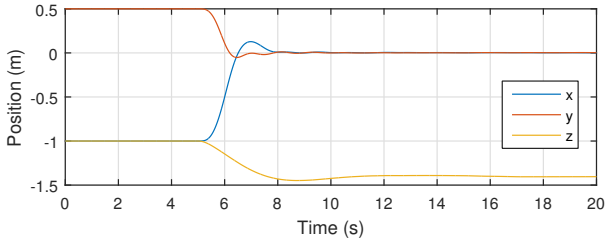

Fig. 5. Quadrotor trajectory with the visual servoing controller. The quadrotor stabilizes faster in x and y than in z.

The camera image features and their centroid are shown in Figs. 6 and 7 for the *real* moving camera and the fixed orientation *virtual* camera. We can see that both converge to the camera optical center, as expected. The disturbance effect of the quadrotor linear velocity on the convergence of the features in the *real* camera can be observed in Fig. 6 by noting that the convergence to the origin is not monotonic.

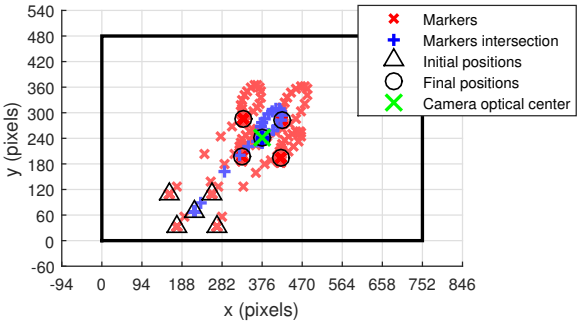

Fig. 6. Image of the landmark features as viewed by the pan and tilt camera during the simulation. The image sensor corresponds to the 752x480 pixels black rectangle.

A more detailed representation of the image errors during the stabilization maneuver is shown in Fig. 8. The pan and tilt camera controller action can be seen in the convergence of the camera feature's centroid ${}^C\bar{\mathbf{y}}$ and the simultaneous action of lateral-longitudinal controller is visible in the convergence of the *virtual* camera image feature centroid ${}^I\bar{\mathbf{y}}$ error to the origin, due to the quadrotor hovering above the landmarks.

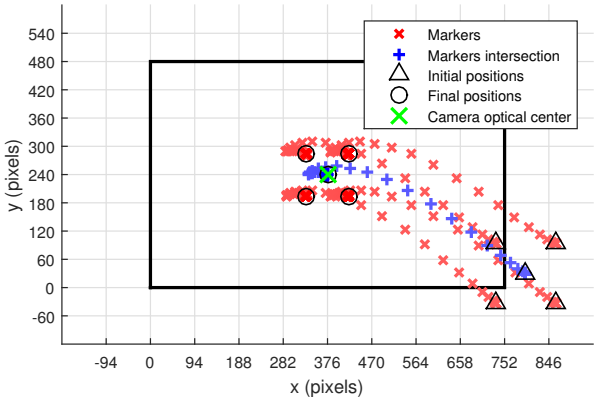

Fig. 7. Image of the landmark features as viewed by the fixed orientation *virtual* camera during the simulation. The image sensor corresponds to the 752x480 pixel black rectangle and, were the camera fixed, some of the features would not be visible.

The vertical coordinate is controlled by feedback of the distance between image features in the virtual camera. The convergence of the pixel distance $\mathbf{s}_{ij}$ to the desired one is shown in Fig. 9, together with its time derivative. The relatively large noise in the time derivative is due to the pixel quantization for the image feature measurements and is well tolerated by the proposed controller.

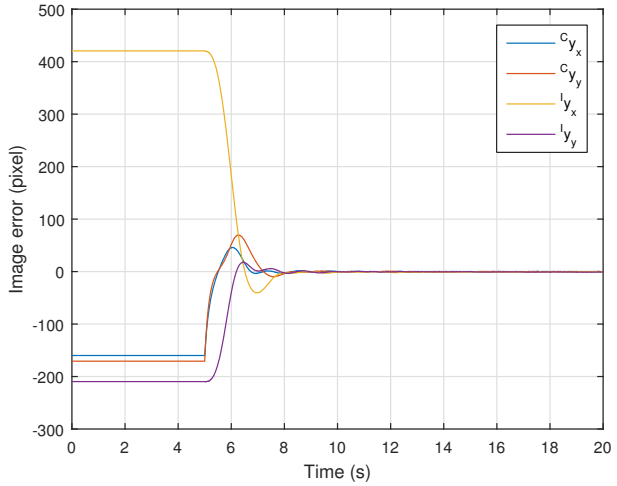

Fig. 8. Pixel error of the feature centroid in both *real* and *virtual* cameras displaying a fast and clear convergence to zero.

The pan and tilt angles are directly related to the location of the quadrotor and converge to zero when the quadrotor is hovering above the landmarks and the landmarks' centroid is centered in the image frame. In Fig. 10, we see that the pan and tilt angles have a fast initial transient corresponding to the cancellation of the quadrotor's initial tilt towards the landmarks. Once the quadrotor stabilizes the pan and tilt angles also quickly converge to zero.

The attitude observer estimates of the angular velocity bias

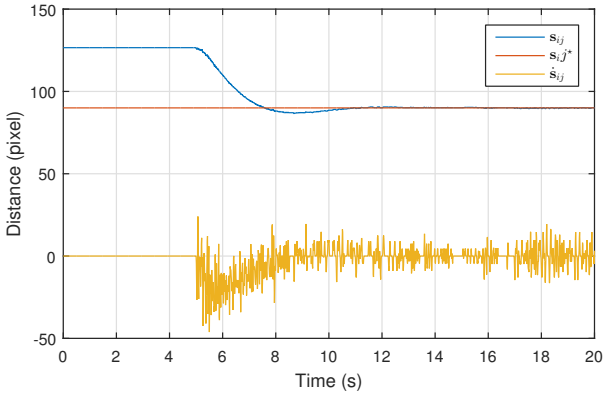

Fig. 9. Distance between image features in the *virtual* camera, the desired distance, and the image features time derivative (showing noise amplification).

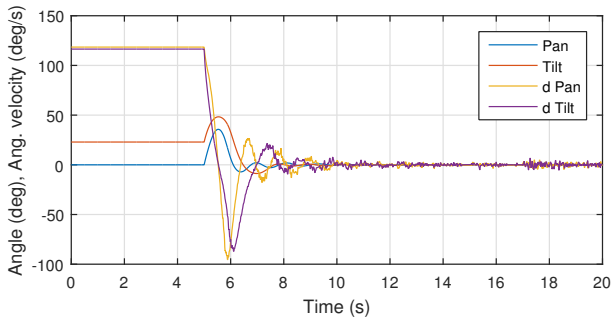

Fig. 10. Evolution of the pan and tilt angles, and their time derivatives, during the stabilization maneuver. The noise in the derivatives is minimal compared to the noise in the image features derivative.

are shown in Fig.11 and clearly converge to the vicinity of the simulated sensor bias.

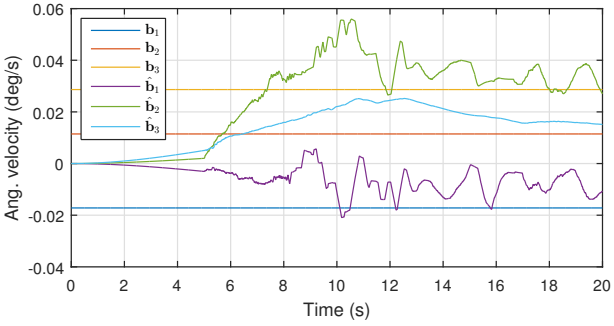

Fig. 11. Evolution of the estimated angular velocity bias $\hat{\mathbf{b}}$ during the stabilization maneuver together with the simulated bias $\mathbf{b}$.

The estimated attitude error is shown in Fig. 12 as the angle between the actual and the estimated attitudes. The convergence of the error to zero is not perfect primarily due to the quantization error introduced by simulating the discrete camera pixels. Accurate measurement of the feature direction is not possible, there is almost always a residual quantization error, and thus precludes the convergence of the attitude error exactly to zero. Nonetheless, the performance and robustness of the proposed observer can be perceived by

the clear reduction of the observation error after it is turned on, and the quadrotor and camera are moving, in the timespan between $t = 5\,\mathrm{s}$ and $t = 10\,\mathrm{s}$.

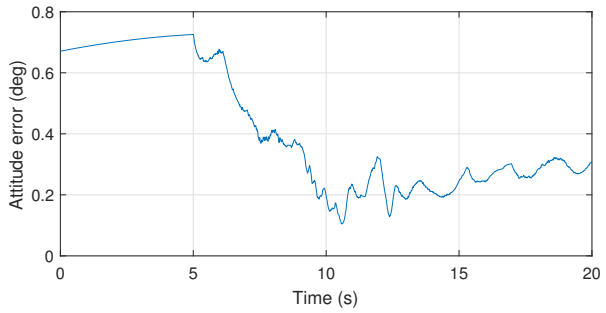

Fig. 12. Angular error between the actual and estimated attitude rotation matrices showing that the error diminishes as the quadrotor moves towards its target location.

The desired attitude reference from the lateral-longitudinal controller is presented in Fig. 13. The effective quadrotor actuations, thrust force, and torque, are shown in Fig. 14. The initial thrust is higher than the gravitational force, thereby making the quadrotor move up, closer to the desired altitude (prescribed by the desired distance between features). When the altitude finally stabilizes, the thrust also stabilizes to a steady-state value where it compensates gravity. The high initial torque actuation drives the thrust vector to point in the direction of the landmarks. Once this is accomplished, the torque actuation necessary to counteract the vehicle's movement is gradually smaller until the quadrotor comes to a full stop over the landmarks. The remaining steady state noise in the actuation is due to the simulated quantization imposed by the camera pixels. The proposed controller is robust to such noise and the features centroid in both the *real* and *virtual* cameras does not deviate more than a couple of pixels from the optical center.

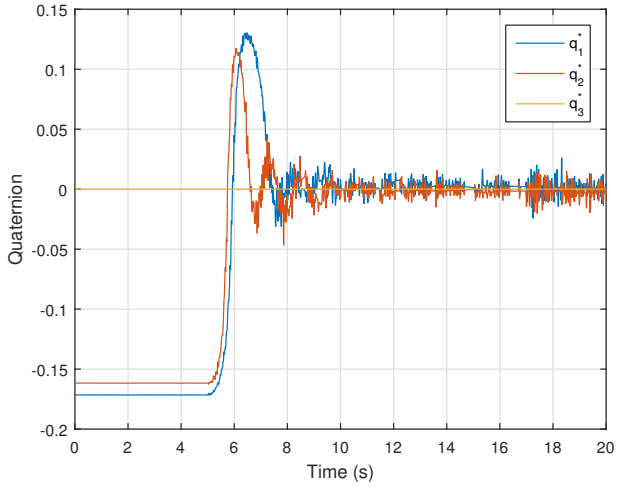

Fig. 13. Desired attitude from the quadrotor lateral-longitudinal controller showing a large initial transient, required for a fast maneuver.

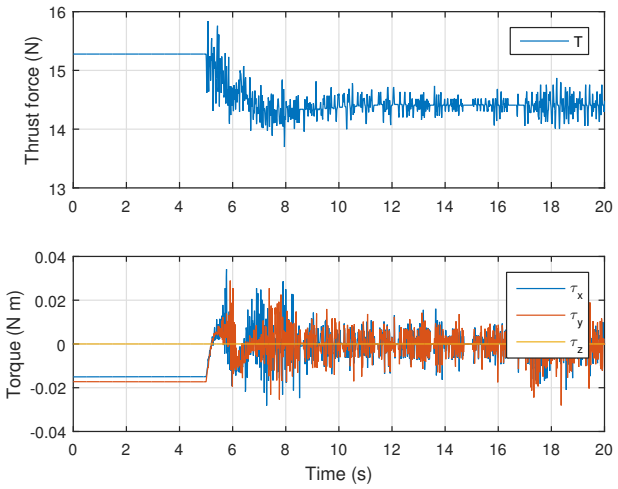

Fig. 14. The quadrotor actuation during the stabilization maneuver is noisy but remains within realistic bounds.

VII. Experimental Results

We evaluated the performance and robustness of the proposed controllers and estimators at the SCORE lab flying arena at the University of Macau. The facilities are equipped with a motion capture system that provides ground-truth data for the vehicle and camera poses, allowing to assess the real-world performance of the proposed estimator, as well as of the controllers.

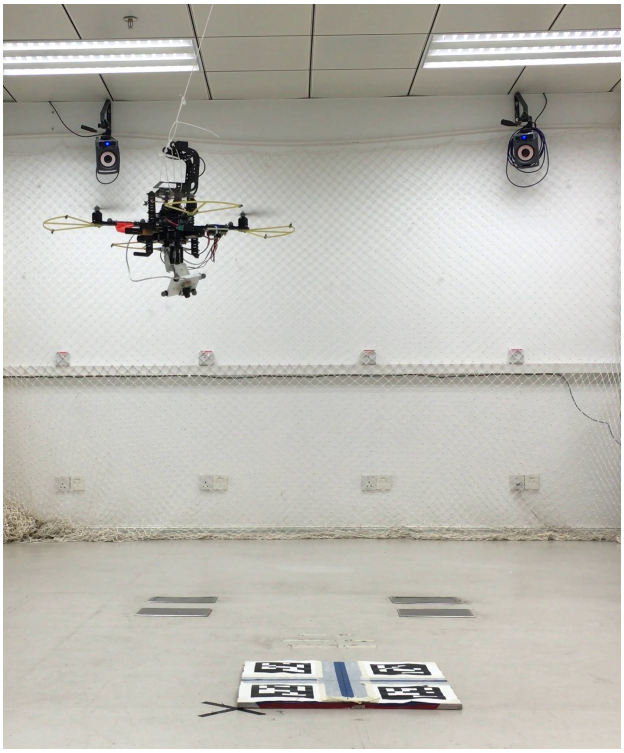

Fig. 15. Experimental setup at the University of Macau. The motion capture cameras, the quadrotor equipped with the pan and tilt camera, and the landmark markers are visible.

The setup is depicted in Fig. 15, in which the motion

capture cameras, the quadrotor and the landmark markers are visible. For the sake of simplicity we used artificial landmarks and fiduciary markers since accurate recognition, processing, and identification, of natural features and landmarks is a computer vision field of study in itself. Our aerial vehicle is an Asctec Pelican quadrotor with an onboard Atom computer. The onboard computer communicates with the low-level Asctec autopilot to obtain gyro measurements at 100Hz and acquires and processes camera images at 60 Hz to identify the four markers. Only the raw gyro data and retinal coordinates (in units of meters with respect to the principal point) of the markers are used for feedback and estimation. The nonlinear estimator and controllers are designed in Matlab/Simulink, from which C code is generated and compiled to run on the Intel Atom computer onboard the quadrotor. The final computed commands for the quadrotor are issued to the low-level Asctec autopilot. Due to restrictions in the experimental setup, it was not possible to use the attitude controller (39) to drive the vehicle. Instead of commanding directly the quadrotor motors, we relied on an inner control-loop for the attitude, which was implemented onboard the vehicle and receives attitude commands, such as the ones given by the proposed lateral-longitudinal controller (43). As long as the exponential convergence property of the attitude controller is maintained the overall quadrotor controller is still stable.

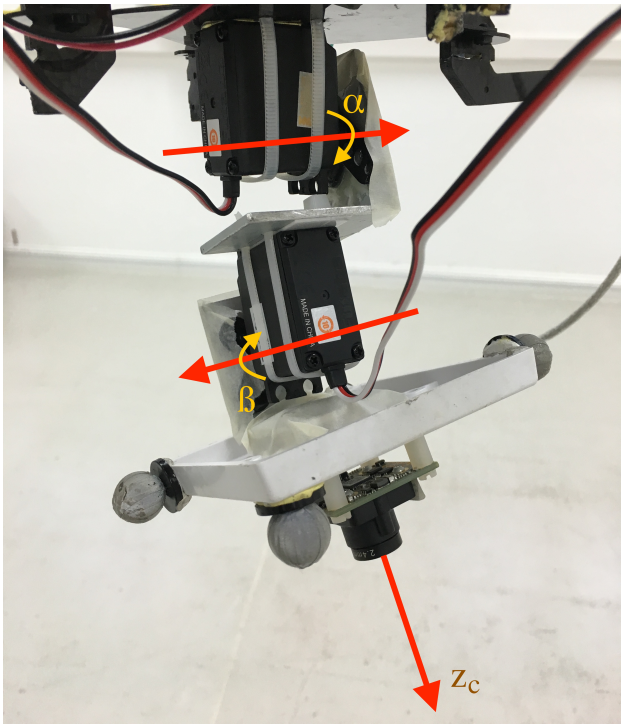

Fig. 16. The camera setup with pan and tilt rotation axes signaled.

The steerable platform connects the camera to the quadrotor through two servo motors (see Fig. 16), controlling each the pan and tilt angles of the platform. The platform is adorned with reflective markers, so that ground-truth data can be obtained using the motion capture system for calibration purposes.

The camera lens is an M12 mount lens with focal length of 2.4 mm and its calibration matrix was determined *a priori*. The relevant parameters are the focal length $f = 420$ pixels and the calibrated camera optical center $(c_x, c_y) = (371.9, 236.3)$ pixels. The imaging sensor area is 752 by 480 pixels. The camera feature measurements are lens-corrected for radial and tangential distortions.

The controller and observer gains are $\lambda_1 = 2$, $\lambda_2 = 0.2$, $K_1 = 1$, $K_2 = 0.8$, $k_1 = 0.021$, $k_2 = 0.019$, $k_b = 0.01$ and $k_\omega = 0.1$.

We now proceed with the analysis of a representative run of the proposed controller and estimator, focusing first on the pan and tilt camera control. The camera coordinates of the markers and centroid are depicted in Fig. 17 and are clearly far from the image plane limits throughout the whole maneuver. The camera controller performs well and, despite abrupt motions of the quadrotor, the image features are always in the vicinity of the image center. Fig. 18, on the other hand, shows that the stabilization maneuver would be impossible with the camera fixed to the quadrotor body, as the landmarks would be outside the camera's field of view. The pan and tilt platform is particularly useful to extend the range of positions where visual-servoing can be used to control the quadrotor.

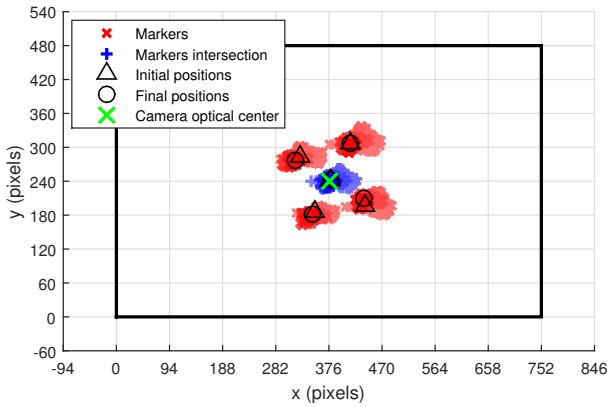

Fig. 17. Image features in the camera sensor during the stabilization maneuver. The features remain within a small region around the center, attesting to the tracking properties of the proposed pan and tilt camera controller.

The detailed representation of the image errors during the stabilization maneuver is shown in Fig. 19. The pan and tilt camera controller maintains the camera feature's centroid ${}^{c}\bar{\mathbf{y}}$ near the optical center and the simultaneous action of lateral-longitudinal controller is visible in the convergence of the *virtual* camera image feature centroid ${}^{I}\bar{\mathbf{y}}$ error to the origin, particularly along the y-axis, due to the quadrotor hovering above the landmarks.

The time evolution of the vertical stabilization based on the distance between image features is shown in Fig. 9 where we can observe the convergence of $\mathbf{s}_{ij}$ to the desired pixel distance $\mathbf{s}^{\star}_{ij}$.

The corresponding camera pan and tilt angles are presented in Fig. 21. The pan and tilt angles are obtained directly from the actuation servos, assuming that the servo references are correctly followed with a small delay — 0.08 s for the servos used in the experiment. The pan angle is initially high due

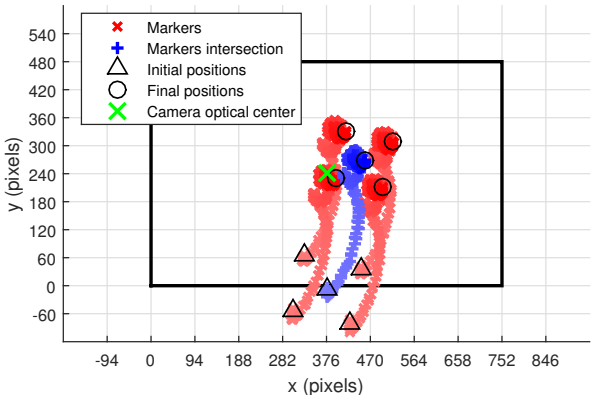

Fig. 18. Image features in the *virtual* camera sensor during the stabilization maneuver. The features would have left the image sensor if a fixed camera had been used.

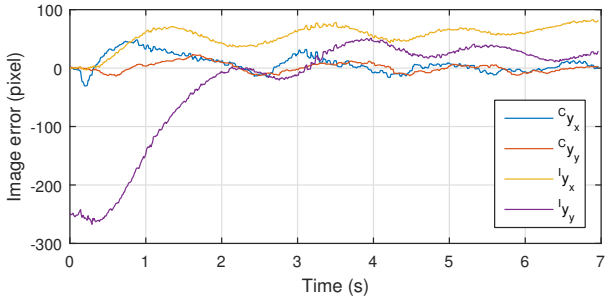

Fig. 19. Pixel error of the feature centroid in both *real* and *virtual* cameras. The markers are not centered in the virtual camera due to modeling and calibration errors but the features remain fixed in the center of the real camera.

to the quadrotor displacement along the y-axis but quickly converges to zero, together with the tilt angle, as the quadrotor approximates the landmarks' location.

The overall 3D view of the maneuver is plotted in Fig. 22, and the quadrotor location during the maneuver is depicted in Fig. 23.

The attitude error is shown in Fig. 24 as the angle between the VICON attitude measurements and the estimated attitude. The convergence is not perfect due to unmodeled dynamics and imperfections in the camera setup but the reduction of the

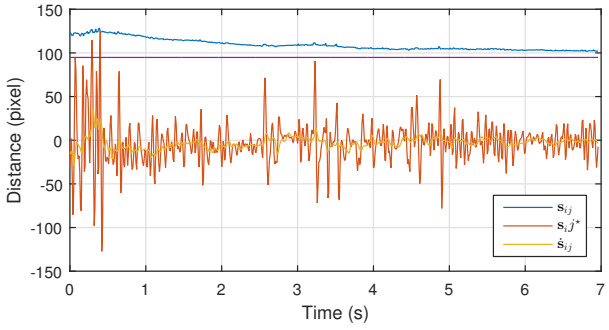

Fig. 20. Distance between image features in the *virtual* camera, the desired distance, and the time derivative of the image features. The convergence of the features distance is slower than in simulation.

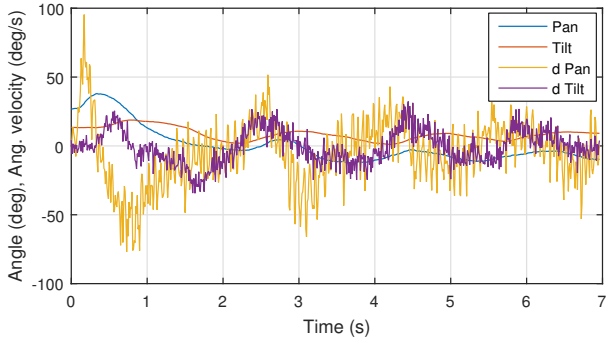

Fig. 21. Evolution of the pan and tilt angles, and their time derivatives, during the stabilization maneuver. A small oscillation subsides in steady state due time delays and other unmodeled dynamics of the system.

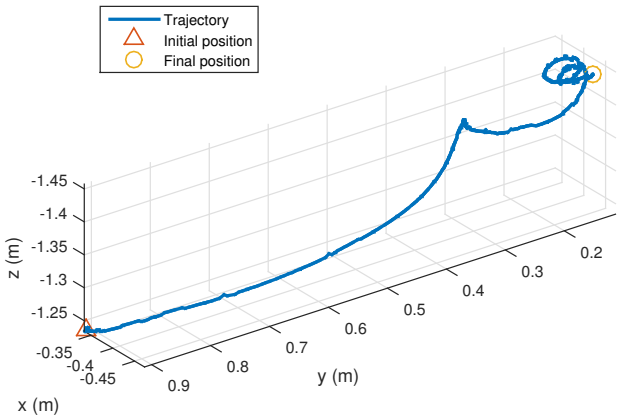

Fig. 22. Three dimensional visualization of the quadrotor trajectory with the visual servoing controller. The quadrotor converges first in $x-y$ and later in z emulating the behavior observed in simulation.

observation error is noticeable.

VIII. Concluding Remarks

This paper proposed a cascaded architecture comprising a nonlinear attitude observer and a nonlinear controller for the stabilization of a quadrotor vehicle based on image measurements of a set of landmarks obtained from a pan and tilt camera and biased rate gyros. The vehicle was stabilized vertically to a given altitude by resorting to a proportional-derivative control law based on image measurements. The lateral-longitudinal stabilization was achieved with a nested saturation control law using feedback of the image measurements, estimated body attitude and angular rate.

Both controllers were shown to be input-to-state-stable with respect to the attitude and rate gyros bias estimation errors, which guarantees the closed-loop stability of the overall cascaded architecture. During the whole stabilization procedure the pan and tilt camera was actuated so as to keep the image of the landmarks' centroid at the center of the image plane. Experimental and simulation results exhibited good performance and attested the applicability of the proposed technique, even in non-ideal conditions.

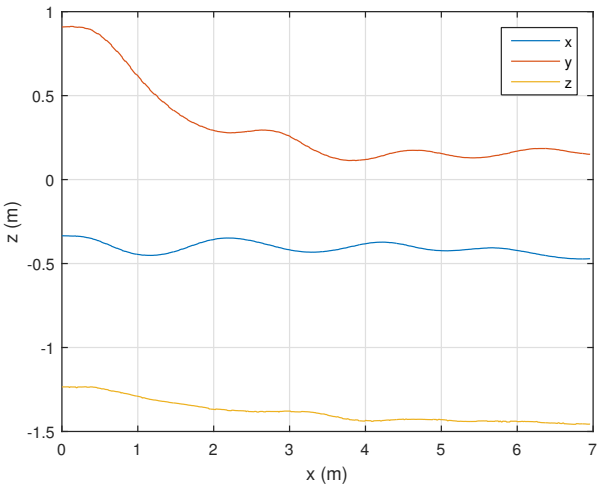

Fig. 23. Quadrotor trajectory with the visual servoing controller. The quadrotor stabilized over the landmarks with a small offset and steady state oscillation.

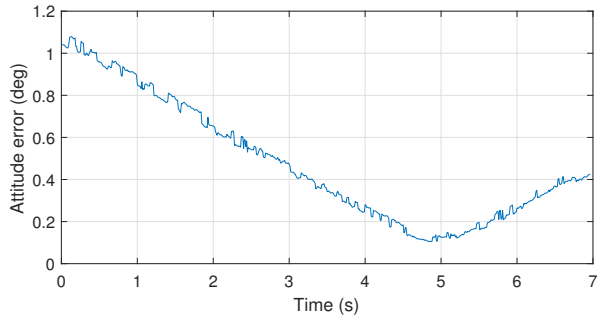

Fig. 24. Angular error between the attitude from VICON measurements and the estimated attitude rotation matrices. The error is small and typically less than 1 degree during the stabilization maneuver.

Future work regarding the controller and observer systems can be developed focusing on the extension of the proposed control and estimation architecture to allow for richer quadrotor control objectives, such as trajectory tracking or path following, always with a focus on using explicitly the image sensor measurements without recovering the quadrotor state.

The major obstacles in transposing the proposed integrated visual servoing solution from a laboratory environment to a real-world setting are Assumption 2, the required knowledge of the landmark locations and their visual identification. Artificial landmarks and fiduciary markers were used to facilitate the computer vision tasks and focus on the controller design and analysis since a full scale implementation and experiment in an outdoor setting is outside of the scope of the present manuscript. The proposed controller, however, can be brought to a real-world setting by leveraging the recent promising results in accurate recognition, processing, identification, and tracking, of natural features in computer vision fueled Augmented Reality applications [26], [27], [28]. Using natural features poses the challenge of knowing, *a priori*, the feature's real world locations and the possibility of breaking Assumption 2. For artificial structure inspection (buildings, bridges, chimneys, etc.) it is possible to know the feature locations from blueprints and use these data to match with the detected visual features and construct the attitude observer. Assumption 2 is only necessary for position control and its relaxation, in particular, the extension of the proposed visual servoing to at least inclined slopes, is promising. Ideally the errors introduced due to the different perspective scaling factors for the features can be mitigated or, as a worse case, kept within a neighborhood of zero.

The identification and tracking of natural features is more computationally intensive than of artificial ones such as fiduciary markers. However, recent hardware advances in graphics processing units (GPUs) and application specific integrated chips (ASICs), motivated by deep learning research, could allow for onboard natural feature processing and tracking in real time at the high rates necessary for accurate vehicle attitude estimation and position tracking. Furthermore, adding new features while navigating around the mission scenario can be also dynamically achieved by complementing the presented technique with a Sensor-based Simultaneous Localization and Mapping (SLAM) algorithm like the one presented in [29] where authors present a globally asymptotically stable solution for the SLAM problem.

REFERENCES

[1] F. Chaumette and S. Hutchinson, "Visual servo control, part I: Basic approaches," *IEEE Robotics and Automation Magazine*, vol. 13, no. 4, pp. 82–90, Dec. 2006.

[2] ——, "Visual servo control, part II: Advanced approaches," *IEEE Robotics and Automation Magazine*, vol. 14, no. 1, pp. 109–118, Mar. 2007.

[3] N. Cowan, J. Weingarten, and D. Koditschek, "Visual servoing via navigation functions," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 4, pp. 521–533, Aug. 2002.

[4] R. Cunha, C. Silvestre, J. Hespanha, and A. Aguiar, "Vision-based control for rigid body stabilization," in *Proceedings of the IEEE Conference on Decision and Control*, Dec. 2007, pp. 2345–2350.

[5] E. Malis and F. Chaumette, "Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 2, pp. 176–186, Apr. 2002.

[6] A. Aguiar and J. Hespanha, "Minimum-energy state estimation for systems with perspective outputs," *IEEE Transactions on Automatic Control*, vol. 51, no. 2, pp. 226–241, Feb. 2006.

[7] Yi Ma, S. Soatto, J. Kosecka, and S. Sastry, *An Invitation to 3-D Vision From Images to Geometric Models*. Springer, 2004, volume 26 of Interdisciplinary Applied Mathematics.

[8] A. I. Mourikis, N. Trawny, S. I. Roumeliotis, A. E. Johnson, A. Ansar, and L. Matthies, "Vision-aided inertial navigation for spacecraft entry, descent, and landing," *IEEE Transactions on Robotics*, vol. 25, no. 2, pp. 264–280, 2009.

[9] H. Rehbinder and B. K. Ghosh, "Pose estimation using line-based dynamic vision and inertial sensors," *IEEE Transactions on Automatic Control*, vol. 48, no. 2, pp. 186–199, 2003.

[10] J. Lobo and J. Dias, "Vision and inertial sensor cooperation using gravity as a vertical reference," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, no. 12, pp. 1597–1608, Dec. 2003.

[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] Y. Ma, J. Kosecka, and S. Sastry, "Vision guided navigation for a nonholonomic mobile robot," *IEEE Transactions on Robotics and Automation*, vol. 15, no. 3, pp. 521–536, Jun 1999.

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

[14] R. Mahony, P. Corke, and T. Hamel, "Dynamic image-based visual servo control using centroid and optic flow features," *Journal of Dynamic Systems, Measurement, and Control*, vol. 130, no. 1, p. 011005, Jan. 2008. [Online]. Available:
[15] R. Vidal, O. Shakernia, and S. Sastry, "Following the flock [formation control]," *IEEE Robotics and Automation Magazine*, vol. 11, no. 4, pp. 14–20, Dec. 2004.
[16] A. D. Luca, G. Oriolo, and P. R. Giordano, "Feature depth observation for image-based visual servoing: Theory and experiments," *The International Journal of Robotics Research*, vol. 27, no. 10, pp. 1093–1116, 2008.
[17] E. Frazzoli, M. Dahleh, and E. Feron, "Trajectory tracking control design for autonomous helicopters using a backstepping algorithm," in *Proceedings of the American Control Conference*, vol. 6, 2000, pp. 4102–4107.
[18] D. Cabecinhas, R. Cunha, and C. Silvestre, "A nonlinear quadrotor trajectory tracking controller with disturbance rejection," *Control Engineering Practice*, vol. 26, no. 0, pp. 1–10, 2014.
[19] A. Isidori, L. Marconi, and A. Serrani, *Robust autonomous guidance: an internal model approach*, ser. Advances in industrial control. London: Springer, 2003.
[20] L. Marconi and R. Naldi, "Robust full degree-of-freedom tracking control of a helicopter," *Automatica*, vol. 43, no. 11, pp. 1909–1920, 2007.
[21] D. Cabecinhas, S. Brás, Silvestre, C., P. Oliveira, and R. Cunha, "Integrated solution to quadrotor stabilization and attitude estimation using a pan and tilt camera," in *Proceedings of the IEEE Conference on Decision and Control*, Dec 2012, pp. 3151–3156.
[22] S. Brás, R. Cunha, J. Vasconcelos, C. Silvestre, and P. Oliveira, "A nonlinear attitude observer based on active vision and inertial measurements," *IEEE Transactions on Robotics*, vol. 27, no. 4, pp. 664–677, Aug. 2011.
[23] A. Loría and E. Panteley, "Uniform exponential stability of linear time-varying systems: revisited," *Systems & Control Letters*, vol. 47, no. 1, pp. 13–24, Sep. 2002.
[24] S. P. Bhat and D. S. Bernstein, "A topological obstruction to continuous global stabilization of rotational motion and the unwinding phenomenon," *Systems & Control Letters*, vol. 39, no. 1, pp. 63–70, Jan. 2000.
[25] S. Sastry and A. Isidori, "Adaptive control of linearizable systems," *IEEE Transactions on Automatic Control*, vol. 34, no. 11, pp. 1123 –1131, nov 1989.
[26] X. Li, W. Hu, C. Shen, Z. Zhang, A. Dick, and A. V. D. Hengel, "A survey of appearance models in visual object tracking," *ACM Trans. Intell. Syst. Technol.*, vol. 4, no. 4, pp. 58:1–58:48, Oct. 2013. [Online]. Available:
[27] A. W. M. Smeulders, D. M. Chu, R. Cucchiara, S. Calderara, A. Dehghan, and M. Shah, "Visual tracking: An experimental survey," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 7, pp. 1442–1468, July 2014.
[28] A. Thangarajah, J. Wu, B. Madon, and A. K. Chowdhury, "Vision-based registration for augmented reality-a short survey," in *2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)*, Oct 2015, pp. 463–468.
[29] B. J. N. Guerreiro, P. Batista, C. Silvestre, and P. Oliveira, "Globally asymptotically stable sensor-based simultaneous localization and mapping," *IEEE Transactions on Robotics*, vol. 29, no. 6, pp. 1380–1395, Dec 2013.

David Cabecinhas received the Licenciatura and Ph.D. degrees in electrical and computer engineering from the Instituto Superior Técnico (IST), Universidade de Lisboa, Portugal, in 2006 and 2014, respectively. He has been a Researcher with the Institute for Systems and Robotics, LarSyS, Lisbon, since 2007. He is currently a Post-Doctoral Fellow with the Faculty of Science and Technology, University of Macau, Macau, China. His current research interests include nonlinear control, sensor-based and vision-based control with applications to autonomous aerial and surface vehicles, and modeling and identification of aerial and surface vehicles.

Sérgio Brás received the B.Sc., M.Sc., and Ph.D. degrees in electrical and computer engineering in 2006, 2008, and 2015 respectively, from Instituto Superior Técnico (IST), University of Lisbon (UL), Portugal. Since 2016, he is an AOCS Performance Engineer at HESpace for the European Space Agency in Noordwijk, The Netherlands. His research interests include nonlinear systems, fault detection and isolation, attitude estimation and control, and inertial navigation systems.

Rita Cunha received the Licenciatura degree in Information Systems and Computer Engineering and the Ph.D. degree in Electrical and Computer Engineering from the Instituto Superior Técnico (IST), Universidade de Lisboa, Portugal, in 1998 and 2007, respectively. She is currently an Assistant Researcher with the Institute for Systems and Robotics, LARSyS, Lisbon, and an Invited Assistant Professor with the Department of Electrical and Computer Engineering of IST. Her research interests include nonlinear systems and control, vision-based and laser-based control with application to autonomous air vehicles.

Carlos Silvestre (M'95) received the Licenciatura degree in Electrical Engineering from the Instituto Superior Tecnico (IST) of Lisbon, Portugal, in 1987 and the Master degree in Electrical Engineering and the Ph.D. degree in Control Science from the same school in 1991 and 2000, respectively. In 2011 he received the Habilitation in Electrical Engineering and Computers also from IST. Since 2000, he is with the Department of Electrical Engineering of the Instituto Superior Tecnico, where he is currently on leave. Since 2015 he is a professor of the Department of Electrical and Computer Engineering of the Faculty of Science and Technology of the University of Macau. His current research interests include linear and nonlinear control theory; hybrid systems; multi-agent control systems; networked control systems; inertial navigation systems and real time architectures for complex autonomous systems with application to unmanned air and underwater vehicles.

Paulo Oliveira (S'91–M'95–SM'11) received the Licenciatura, M.Sc., and Ph.D. degrees in electrical and computer engineering from Instituto Superior Técnico (IST), Lisbon, Portugal, in 1987, 1991, and 2002, respectively. Currently, he is an Associate Professor (2010) with Habilitation (2016) in the Department of Mechanical Engineering, IST. His research interests are in the area of autonomous robotic vehicles with a focus on the fields of estimation, sensor fusion, navigation, positioning, and mechatronics. He is the author or coauthor of more than 65 journal papers and 180 conference communications. He participated in more than 30 European and Portuguese research projects, over the last 25 years.