

Velocity-free Image-based Control of Unmanned Aerial Vehicles

Rafik Mebarki and Bruno Siciliano

Abstract—Controlling aerial vehicles motion considerably relies on the measure of their velocity. Yet, obtaining the translational part of such information from solely the on-board sensors is still an open issue to envisage fully autonomous applications. In this paper, we present a nonlinear observer based on the images from a single on-board camera to on-line estimate the translational velocity. Spherical image coordinates are adopted. Then, through Backstepping design, the result is exploited in a visual servo controller, thus endowing the vehicle with the capability of fully autonomously positioning on visual targets. The stability and convergence of the closed-loop system are established through Lyapunov synthesis. By means of computer simulations, the validity and robustness of the proposed system are shown.

I. INTRODUCTION

Nowadays, Unmanned Aerial Vehicles (UAVs) have become subject of intensive research work, due to numerous advantages they afford. They flight without a pilot onboard, and thus they might be of amenable size and can operate in hazardous environments. Especially, Vertical Take-off and Landing (VTOL) UAVs present remarkable assets. This is due to their ability of holding stationary, like aerial vessels, allowing different potential applications: moving in relatively narrow and cumbersome areas, surveillance and exploration, interaction with humans as search and rescue, aerial manipulation [1], [2], to name but a few. Different sensing modalities can be considered to relay the necessary environment information required by autonomous vehicle motion control. We can cite for example Global Positioning System (GPS), Inertial Measurement Units (IMUs), and vision. GPS however suffers from signal loss at low altitudes, does not operate in urban areas [3], [4], and does not provide measures with enough accuracy to envisage fine motion control. Vision nevertheless presents noticeable advantages. Besides providing wealthy information with a resolution of pixel order, it is passive, light-weight, and cheap [5]. In this paper we adopt this modality as the main source of information. Vision has been extensively investigated in the field of robot manipulators and ground vehicles control, referred to as *visual servoing* [6].

The research leading to these results has been supported by the ARCAS collaborative project, which has received funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement ICT-287617. The authors are solely responsible for its content. It does not represent the opinion of the European Community and the Community is not responsible for any use that might be made of the information contained therein.

The authors are with PRISMA Lab, Dipartimento di Ingegneria Elettrica e Tecnologie dell'Informazione, Università degli Studi di Napoli Federico II, Via Claudio 21, 80125 Naples, Italy, rafik.mebarki@unina.it and bruno.siciliano@unina.it

To control UAV motion, a reliable measure of its translational velocity is needed to close the servo loop (e.g. see [7]). Yet, existing sensors do not provide a measure of the velocity. The issue sparked different research works. In [8], virtual states are introduced in the controller for UAV stabilisation. Vision was not considered. In [9], scanning laser range sensor measurements and IMU are fused through simultaneous localization and mapping (SLAM) and Extended Kalman Filter (EKF) algorithms for pose and velocity estimation and control. Adopting laser sensors restricts the height at which the UAV might operate, in addition to the limited bandwidth. As for EKF, its performance heavily relies on the tuned parameters, such as the covariance matrix and the initial conditions, with no guarantee of convergence though. In [10], vision-based guidance and control of UAV for indoor environment is presented. However, an external Vicon localisation system is employed for vehicle state determination, thus yielding non-autonomous flight. In [11] air pressure sensors, vision, and IMU are used for UAV hovering control. Vision is used to localize the vehicle using the SLAM technique, while air pressure sensor is used to obtain its height. Data are fused with an EKF algorithm. A linear model is adopted to predict the UAV behavior. Linear models in contrast to nonlinear ones present a restricted reactivity and domain of convergence.

A technique for velocity determination of a moving camera equipped with an IMU is presented in [12]. The camera image coordinates are combined with the IMU accelerometer readings through three consecutive positions to perform the estimation. Robustness to image noise is however not reported. Similarly, but using spherical image coordinates, the velocity of a camera-attached UAV is extracted in [13]. In [14]–[17] IMU accelerometer readings are used in nonlinear observers for velocity, pose, and pose and velocity estimation, respectively. In [14] and [15], vision is not adopted. In this work we focus on vision information rather than IMU accelerometers, whose measures are known to be very noisy and biased. Only rotational parts of the IMU data are considered. Work [18] presents a velocity-free controller for a fleet of UAVs without vision. It needs the position of each vehicle, a requirement though not readily met by on-board sensing. Other works considered optical flow. In [19], the *continuous homography constraint* presented in [20] is applied to estimate UAV velocity, which is then used for motion control. It is assumed that the ground terrain is planar and textured. The stability of the closed-loop system is not provided. In [21], velocity measures are replaced by optical flow in an image-based UAV controller. Optical flow suffers from image noise, is not easy to compute from real images,

and is prone to drifts [11], in addition to the constraint of a well-textured environment for image matching.

In this paper, we propose a nonlinear observer that uses vision as the main source of information to perform translational velocity estimation. IMU is used only to obtain the attitude and angular velocity of the vehicle. These informations are known as being reliable. The scenario consists of a camera facing downward, attached to the UAV, observing motionless visual targets, consisting of 3D points in this study. Spherical image coordinates are adopted. As shown hereinafter, two points are enough to ensure the stability of the observer. Then, we consider automatic control, where the objective is to stabilize the vehicle with respect to these visual targets. We make use of the observer results, thus building upon previous work [22] by endowing the vehicle with the capability of fully autonomously tracking visual targets, without the need of any exteroceptive sensor to obtain the velocity. The system design and stability proof are derived through Lyapunov synthesis using Backstepping design [23].

II. SYSTEM DYNAMICS

We adopt the model used in [24], [25] to characterize the dynamic behaviors of the vehicle, namely VTOL micro-aerial craft. Consider a quadrotor of mass m and inertia $\mathbf{I}$. The equations of motions of the vehicle are thus given by

$$\begin{cases} \dot{\mathbf{v}} &= -\boldsymbol{\omega} \times \mathbf{v} + \frac{1}{m} \mathbf{f} \\ \mathbf{I} \dot{\boldsymbol{\omega}} &= -\boldsymbol{\omega} \times (\mathbf{I} \boldsymbol{\omega}) + \boldsymbol{\Gamma}, \end{cases} \quad (1)$$

with

$$\mathbf{f} = -u \mathbf{e}_3 + m g \mathbf{R}^\top \mathbf{e}_3, \quad (2)$$

where $\times$ denotes the cross product. $\mathbf{v} = (v_x, v_y, v_z)^\top$ and $\boldsymbol{\omega} = (\omega_x, \omega_y, \omega_z)^\top$ represent respectively the translational and rotational velocity of vehicle-fixed-frame $\{R_s\}$ relative to fixed Cartesian base frame $\{R_b\}$ (see Fig. 1). Note that $\mathbf{v}$ and $\boldsymbol{\omega}$ are expressed in Cartesian frame $\{R_s\}$. The latter is assumed to be located at the center of mass of the vehicle, with Z axis fixing downward. $\mathbf{f} \in \mathbb{R}^3$ and $\boldsymbol{\Gamma} \in \mathbb{R}^3$ correspond to the total linear forces and total torques acting on the craft, respectively. Scalar u represents the thrust control input to the vehicle, g the gravity term, and $\mathbf{e}_3 = (0, 0, 1)^\top$, i.e. the Z axis of the vehicle. $\mathbf{R}$ denotes the rotation matrix defining the orientation of $\{R_s\}$ with respect to inertial frame $\{R_i\}$. The latter has its z -axis fixing downward. We assume that $\mathbf{R}$ can be estimated using literature techniques, e.g. [26].

III. VISUAL INFORMATION

The choice of the type of visual features plays a predominant role in such a way as to condition the stability and behavior of the designed image-based robotic system. In this paper, we build upon previous work [22], where we employ spherical visual features. More precisely, we adopt spherical centroid visual information. Such features lead to a passivity property, in such a way as to make the system amenable from the control design point of view [25]. In addition, they do

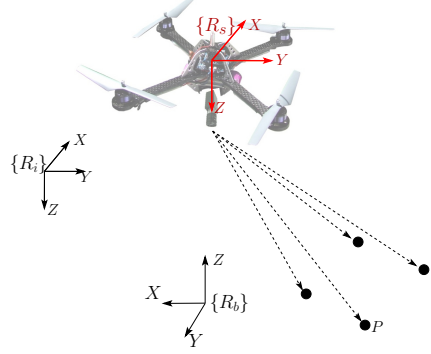


Fig. 1. Vehicle associated frames.

not necessitate a matching of the observed points through the successive acquired images, because of the summation (integration), and are robust to image noise.

We assume without loss of generality that $\{R_s\}$ coincides with the camera frame. Consider an object (target) characterized by a set of n stationary points P_i , whose coordinates in $\{R_s\}$ are ${}^s P_i = (X_i \ Y_i \ Z_i)^\top$ (see Fig. 2). The unnormalized spherical centroid is

$$\mathbf{q} = \sum_{i=1}^n \mathbf{s}_i \quad \text{with} \quad \mathbf{s}_i = \frac{{}^s P_i}{\|{}^s P_i\|}, \quad (3)$$

where $\mathbf{s}_i$ represents the spherical image coordinates of point P_i , when considering without loss of generality a unitary sphere radius. Term $\|\cdot\|$ denotes the Euclidean norm. Considering $\mathbf{p}_i \in \mathbb{R}^2$ as the perspective image coordinates of P_i (Fig. 2), one can notice that $\mathbf{s}_i$ can be solely expressed as a function of $\mathbf{p}_i$ as $\mathbf{s}_i = \bar{\mathbf{p}}_i / \|\bar{\mathbf{p}}_i\|$, where $\bar{\mathbf{p}}_i^\top = [\mathbf{p}_i^\top \ 1]$. This result is in fact obtained by dividing by Z_i both the numerator and denominator of $\mathbf{s}_i$ in (3). The feedback visual error used in [25] is $\boldsymbol{\delta} = \mathbf{q} - \mathbf{q}^*$, where $\mathbf{q}^*$ is the desired spherical feature. It is expressed as $\mathbf{q}^* = \mathbf{R}^\top \mathbf{b}$, where $\mathbf{b}$ represents the known desired direction of the visual feature in the inertial frame. Feedback error $\boldsymbol{\delta}$ as formulated above suffers from the fact that it leads to a system which is not well-conditioned. To cope with this, another version $\boldsymbol{\delta}_1$ is used [22]:

$$\boldsymbol{\delta}_1 = \mathbf{q}_0^* \times \mathbf{q} + \lambda \mathbf{q}_0^* \mathbf{q}_0^{*\top} \boldsymbol{\delta}, \quad (4)$$

where $\mathbf{q}_0^* = \mathbf{q}^* / \|\mathbf{q}^*\|$, and $\lambda > 0$ is a constant scalar. One can verify that time variation of $\boldsymbol{\delta}_1$ can be obtained as follows:

$$\dot{\boldsymbol{\delta}}_1 = -\boldsymbol{\omega} \times \boldsymbol{\delta}_1 - \mathbf{A} \mathbf{Q} \mathbf{v}, \quad (5)$$

where

$$\mathbf{Q} = \sum_{i=1}^n \frac{1}{\|{}^s P_i\|} (\mathbf{I}_3 - \mathbf{s}_i \mathbf{s}_i^\top), \quad \text{and} \quad \mathbf{A} = [\mathbf{q}_0^*]_\times + \lambda \mathbf{q}_0^* \mathbf{q}_0^{*\top}. \quad (6)$$

$[a]_\times$ denotes the skew-symmetric matrix of vector a , and $\mathbf{I}_3$ the 3×3 identity matrix. Note that matrix $\mathbf{Q} \in \mathbb{R}^{3 \times 3}$ is positive definite as long as more than one point is observed ($n \geq 2$) [25]. Note also that it encloses the image point depth

$|P_i|$. One also has

$$\mathbf{A}^{-1} = -[\mathbf{q}_0^*]_{\times} + \frac{1}{\lambda} \mathbf{q}_0^* \mathbf{q}_0^{*\top} \text{ and } \frac{d}{dt} \mathbf{A}^{-1} = \mathbf{A}^{-1} [\boldsymbol{\omega}]_{\times} - [\boldsymbol{\omega}]_{\times} \mathbf{A}^{-1}. \quad (7)$$

IV. NONLINEAR OBSERVER AND CONTROL SYSTEM

In this section, we propose a nonlinear observer to estimate online the translational velocity of a quadrotor, then derive a low-level controller to automatically stabilize it with respect to visual targets. It is assumed that the vehicle is equipped with an onboard camera, which has in view a stationary object. We propose to exploit δ_1 , introduced by (4), as the main information to update the observer, as well as feedback to the controller. The stability and convergence of the system is demonstrated by means of Lyapunov analysis.

A. Observer and controller

Let $\hat{\mathbf{v}}$ be the estimate of vehicle velocity $\mathbf{v}$. We propose the following nonlinear observer:

$$\begin{cases} \tilde{\delta}_{1o} &= \delta_1 - \hat{\delta}_1 \\ \tilde{\mathbf{v}}_o &= \mathbf{v} - \hat{\mathbf{v}} \\ \dot{\tilde{\delta}}_1 &= -\boldsymbol{\omega} \times \delta_1 - \mathbf{A} \mathbf{Q} \hat{\mathbf{v}} + k_{1o} \tilde{\delta}_{1o} \\ \dot{\hat{\mathbf{v}}} &= \hat{\mathbf{z}} - k_{2o} \mathbf{A}^{-1} \tilde{\delta}_{1o} \\ \dot{\hat{\mathbf{z}}} &= -\boldsymbol{\omega} \times \hat{\mathbf{v}} + \frac{1}{m} \mathbf{f} - k_{2o} \mathbf{G} \tilde{\delta}_{1o} - \boldsymbol{\xi} \\ \mathbf{G} &= k_{1o} \mathbf{A}^{-1} - \dot{\mathbf{A}}^{-1} \text{ and } \boldsymbol{\xi} = \mathbf{Q}^{\top} \mathbf{A}^{\top} \tilde{\delta}_{1o}. \end{cases} \quad (8)$$

Gains k_{1o} and k_{2o} are positive scalars to tune the observer convergence rate. $\mathbf{A}^{-1}$ and $\dot{\mathbf{A}}^{-1}$ are expressed as (7). We show afterwards that the above observer yields the estimation error $\tilde{\mathbf{v}}_o$, as well as $\tilde{\delta}_{1o}$, to converge to zero. Note that in the proof it is assumed that the image depth, namely Z_i , is known since $|^s P_i|$ is enclosed in matrix $\mathbf{Q}$. The latter is involved in (8). Such a depth measure can in fact be retrieved by employing stereo vision and using algorithms such as triangulation [27]. Nevertheless, we found that the control algorithm we propose is robust with respect to the image depth. This is reported and more detailed later in Sect. V.

Then let us propose the following controller:

$$\begin{aligned} \bar{\mathbf{u}} &= (u \omega_y - u \omega_x \dot{u})^{\top} \\ &:= m c_1^2 c_2 \mathbf{A}^{\top} \left(- (c_1 c_2 + c_3) \delta_3 - \frac{1}{c_1} \mathbf{A}_3 \tilde{\delta}_{1o} \right. \\ &\quad \left. + c_1 \mathbf{H} (\delta_1 - \delta_2) \right) \in \mathbb{R}^3, \end{aligned} \quad (9)$$

where c_1 , c_2 , and c_3 are positive scalar gains to tune the controller. $(\delta_2, \delta_3) \in \mathbb{R}^3$ are Backstepping errors as follows:

$$\delta_2 = \delta_1 - \frac{1}{c_1} \mathbf{A}^{-\top} \hat{\mathbf{v}}, \text{ and } \delta_3 = \delta_2 - \frac{1}{m c_1^2 c_2} \mathbf{A}^{-\top} \mathbf{f}, \quad (10)$$

where $\mathbf{A}_3 = \mathbf{A}^{-\top} \mathbf{Q}^{\top} \mathbf{A}^{\top}$. The above controller yields feedback visual error δ_1 to converge to zero using only the on-board camera and IMU angular data, as long as

$$\begin{cases} k_{1o} > \mu_{1o}, & k_{2o} > \mu_{2o} q_{min}^{-1}, & c_1 > \mu_{1c} h_{min}^{-1}, \\ c_2 > h_{max} + c_1^{-1} \mu_{2c}, & c_3 > \mu_{3c}, \end{cases} \quad (11)$$

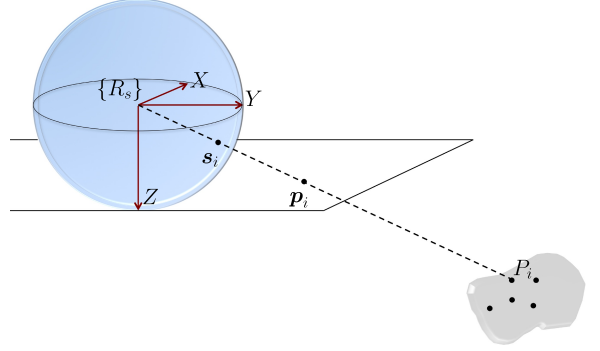


Fig. 2. Spherical projection: spherical $\mathbf{s}_i$ and perspective $\mathbf{p}_i$ image coordinates of 3D point P_i .

where

$$\begin{cases} \mu_{1o} = \frac{1}{2} c_1^{-1} \|\mathbf{A}_3\|_F, & \mu_{2o} = \frac{1}{2} \|\mathbf{Q}\|_F (\|\mathbf{A}\|_F + 2 \|\mathbf{A}_2\|_F) \\ \mu_{1c} = \frac{1}{2} \|\mathbf{Q}\|_F \|\mathbf{A}\|_F, & \mu_{2c} = \frac{1}{2 c_1} \|\mathbf{A}_3\| + \frac{1}{2} \|\mathbf{A}_2\|_F \|\mathbf{Q}\|_F \\ \mu_{3c} = \frac{1}{2} \|\mathbf{A}_2\|_F \|\mathbf{Q}\|_F, \end{cases} \quad (12)$$

and such that q_{min} represents a bound on the minimal eigenvalue of $\mathbf{Q}$, while h_{min} and h_{max} are bounds on respectively the minimal and maximal eigenvalues of matrix $\mathbf{H} = \mathbf{A} \mathbf{Q} \mathbf{A}^{\top}$. Note that $\mathbf{H}$ is positive definite, since it is the case for $\mathbf{Q}$. Matrix $\mathbf{A}_2$ is given by $\mathbf{A}_2 = \mathbf{A} + (k_{2o}/c_1) \mathbf{A}^{-\top}$. Finally, $\|\cdot\|_F$ denotes the Frobenius norm.

B. Proof

The convergence of both observer and controller is now established. First, consider the following Lyapunov candidate function related to the observer:

$$V_o = \frac{1}{2} |\tilde{\delta}_{1o}|^2 + \frac{1}{2} |\tilde{\mathbf{v}}_o|^2. \quad (13)$$

The time derivative of (13) writes

$$\dot{V}_o = \tilde{\delta}_{1o}^{\top} \dot{\tilde{\delta}}_{1o} + \tilde{\mathbf{v}}_o^{\top} \dot{\tilde{\mathbf{v}}}_o. \quad (14)$$

Time differentiating the first equation of (8) and taking into account kinematics (5) of visual error δ_1 yields

$$\dot{\tilde{\delta}}_{1o} = -k_{1o} \tilde{\delta}_{1o} - \mathbf{A} \mathbf{Q} \tilde{\mathbf{v}}_o. \quad (15)$$

Similarly for the second equation of (8), and taking into account result (15), yields

$$\dot{\tilde{\mathbf{v}}}_o = -\boldsymbol{\omega} \times \tilde{\mathbf{v}}_o + k_{2o} \mathbf{Q} \tilde{\mathbf{v}}_o + \frac{1}{m} \mathbf{f} - \boldsymbol{\xi}. \quad (16)$$

Then, subtracting the above relationship from the first equation of (1), it follows that:

$$\dot{\tilde{\mathbf{v}}}_o = -\boldsymbol{\omega} \times \tilde{\mathbf{v}}_o - k_{2o} \mathbf{Q} \tilde{\mathbf{v}}_o + \boldsymbol{\xi}. \quad (17)$$

Plugging (15) and (17) into (14), we have

$$\dot{V}_o = -k_{1o} |\tilde{\delta}_{1o}|^2 - k_{2o} \tilde{\mathbf{v}}_o^{\top} \mathbf{Q} \tilde{\mathbf{v}}_o, \quad (18)$$

which is negative definite ($\dot{V}_o < 0$), since $\mathbf{Q} > 0$, provided that more than one point is observed in the image. This shows that the estimation error $\tilde{\mathbf{v}}_o$, along with $\tilde{\delta}_{1o}$, converges to

zero. Therefore, the observer we propose by (8) ensures the estimation of the vehicle translational velocity.

Now we also show the stability and convergence of the velocity-free visual servo controller presented above. Consider first the following Lyapunov candidate function related to feedback visual error δ_1

$$V_1 = \frac{1}{2}|\delta_1|^2. \quad (19)$$

Time differentiating (19), taking into account (5), and noticing that $\mathbf{x}^\top(\mathbf{x} \times \mathbf{y}) = 0$, it follows that:

$$\dot{V}_1 = \delta_1^\top \dot{\delta}_1 = -\delta_1^\top \mathbf{A} \mathbf{Q} \mathbf{v}. \quad (20)$$

Now, substituting $\mathbf{v}$ according to the second expression of (8) yields

$$\dot{V}_1 = -\delta_1^\top \mathbf{A} \mathbf{Q} \hat{\mathbf{v}} - \delta_1^\top \mathbf{A} \mathbf{Q} \tilde{\mathbf{v}}_o. \quad (21)$$

We have shown earlier that the observation errors converge to zero. Thus assume, without loss of generality, and only for analysis purpose, that $\tilde{\mathbf{v}}_o \approx 0$. We thus would have $\dot{V}_1 = -\delta_1^\top \mathbf{A} \mathbf{Q} \hat{\mathbf{v}}$. Therefore, reaching

$$\hat{\mathbf{v}} := c_1 \mathbf{A}^\top \delta_1 \quad (22)$$

would yield $\dot{V}_1 := -c_1 \delta_1^\top \mathbf{A} \mathbf{Q} \mathbf{A}^\top \delta_1 = -c_1 \delta_1^\top \mathbf{H} \delta_1$, a negative definite function, because of $\mathbf{H}$. Accordingly a new Lyapunov error, namely δ_2 , is considered. It has already been introduced by (10). Taking into account δ_2 , $\dot{V}_1$ of (21) is re-written as

$$\dot{V}_1 = -c_1 \delta_1^\top \mathbf{H} \delta_1 + c_1 \delta_1^\top \mathbf{H} \delta_2 - \delta_1^\top \mathbf{A} \mathbf{Q} \tilde{\mathbf{v}}_o. \quad (23)$$

Similarly, $\dot{\delta}_1$ becomes

$$\dot{\delta}_1 = -\omega \times \delta_2 - \frac{1}{c_1}[\omega]_\times \mathbf{A}^{-\top} \hat{\mathbf{v}} - \mathbf{A} \mathbf{Q} \hat{\mathbf{v}} - \mathbf{A} \mathbf{Q} \tilde{\mathbf{v}}_o. \quad (24)$$

To the aim of nullifying δ_2 , define the following Lyapunov candidate storage function:

$$V_2 = \frac{1}{2}|\delta_2|^2. \quad (25)$$

From (10), time differentiating δ_2 , replacing $\frac{d}{dt} \mathbf{A}^{-\top}$ according to (7), while $\dot{\delta}_1$ and $\hat{\mathbf{v}}$ are substituted by (24) and (16), respectively, we have

$$\begin{aligned} \dot{\delta}_2 = & -\omega \times \delta_2 - c_1 \mathbf{H}(\delta_1 - \delta_2) - \frac{1}{m} c_1^{-1} \mathbf{A}^{-\top} \mathbf{f} \\ & + c_1^{-1} \mathbf{A}^{-\top} \xi - \mathbf{A}_2 \mathbf{Q} \tilde{\mathbf{v}}_o. \end{aligned} \quad (26)$$

Plugging (26) into time variation $\dot{V}_2$ of V_2 given by (25) yields

$$\begin{aligned} \dot{V}_2 = & -c_1 \delta_2^\top \mathbf{H}(\delta_1 - \delta_2) - \frac{1}{m} c_1^{-1} \delta_2^\top \mathbf{A}^{-\top} \mathbf{f} \\ & + c_1^{-1} \delta_2^\top \mathbf{A}^{-\top} \xi - \delta_2^\top \mathbf{A}_2 \mathbf{Q} \tilde{\mathbf{v}}_o. \end{aligned} \quad (27)$$

A command candidate to make $\dot{V}_2$ negative is $\frac{1}{m} \mathbf{A}^{-\top} \mathbf{f} := c_1^2 c_2 \delta_2$. Thus another control error is adopted, i.e. δ_3 previously introduced in (10). Therefore, consider the Lyapunov candidate function

$$V_3 = \frac{1}{2}|\delta_3|^2. \quad (28)$$

From (10), time variation of δ_3 is

$$\dot{\delta}_3 = \dot{\delta}_2 - \frac{1}{m}(c_1^2 c_2)^{-1} \left(\frac{d}{dt}(\mathbf{A}^{-\top}) \mathbf{f} + \mathbf{A}^{-\top} \dot{\mathbf{f}} \right). \quad (29)$$

Note that the classical relationship $\dot{\mathbf{R}} = \mathbf{R}[\omega]_\times$ is used. With the new error δ_3 introduced, from (26) $\dot{\delta}_2$ becomes:

$$\begin{aligned} \dot{\delta}_2 = & -\omega \times \left(\delta_3 + \frac{1}{m}(c_1^2 c_2)^{-1} \mathbf{A}^{-\top} \mathbf{f} \right) - c_1 \mathbf{H}(\delta_1 - \delta_2) \\ & + c_1 c_2 (\delta_3 - \delta_2) + c_1^{-1} \mathbf{A}^{-\top} \xi - \mathbf{A}_2 \mathbf{Q} \tilde{\mathbf{v}}_o. \end{aligned} \quad (30)$$

Similarly, $\dot{V}_2$ becomes

$$\begin{aligned} \dot{V}_2 = & -c_1 \delta_2^\top \mathbf{H}(\delta_1 - \delta_2) + c_1^{-1} \delta_2^\top \mathbf{A}^{-\top} \xi \\ & - c_1 c_2 |\delta_2|^2 + c_1 c_2 \delta_2^\top \delta_3 - \delta_2^\top \mathbf{A}_2 \mathbf{Q} \tilde{\mathbf{v}}_o. \end{aligned} \quad (31)$$

Updating (29) according to (30) yields

$$\dot{V}_3 = -c_1 c_2 \delta_3^\top \delta_2 - \delta_3^\top \mathbf{A}_2 \mathbf{Q} \tilde{\mathbf{v}}_o - c_3 |\delta_3|^2. \quad (32)$$

Finally, consider the following Lyapunov function that encloses the errors of both the observer and controller:

$$\begin{aligned} V = & V_o + V_1 + V_2 + V_3 \\ = & \frac{1}{2}|\tilde{\delta}_{1o}|^2 + \frac{1}{2}|\tilde{\mathbf{v}}_o|^2 + \frac{1}{2}|\delta_1|^2 + \frac{1}{2}|\delta_2|^2 + \frac{1}{2}|\delta_3|^2. \end{aligned} \quad (33)$$

Time differentiating (33), and replacing results (18), (23), (31), and (32), we have

$$\begin{aligned} \dot{V} = & -k_{1o}|\tilde{\delta}_{1o}|^2 - k_{2o} \tilde{\mathbf{v}}_o^\top \mathbf{Q} \tilde{\mathbf{v}}_o - \delta_1^\top \mathbf{A} \mathbf{Q} \tilde{\mathbf{v}}_o \\ & - c_1 \delta_1^\top \mathbf{H} \delta_1 - c_1 \delta_2^\top (c_2 \mathbf{I}_3 - \mathbf{H}) \delta_2 - \delta_2^\top \mathbf{A}_2 \mathbf{Q} \tilde{\mathbf{v}}_o \\ & + c_1^{-1} \delta_2^\top \mathbf{A}^{-\top} \mathbf{Q}^\top \mathbf{A}^\top \tilde{\delta}_{1o} - \delta_3^\top \mathbf{A}_2 \mathbf{Q} \tilde{\mathbf{v}}_o - c_3 |\delta_3|^2. \end{aligned} \quad (34)$$

Noticing that $\mathbf{x}^\top \mathbf{y} \leq |\mathbf{x}| |\mathbf{y}|$, $|\mathbf{M} \mathbf{x}| \leq \|\mathbf{M}\|_F |\mathbf{x}|$, and $a b \leq \frac{1}{2}(a^2 + b^2)$, for any matrix $\mathbf{M}$, vectors $(\mathbf{x}, \mathbf{y})$, and scalars (a, b) , we can upper bound $\dot{V}$ as follows:

$$\begin{aligned} \dot{V} \leq & -k_{1o}|\tilde{\delta}_{1o}|^2 + \frac{1}{2c_1} \|\mathbf{A}_3\|_F (|\delta_2|^2 + |\tilde{\delta}_{1o}|^2) - c_1 \delta_1^\top \mathbf{H} \delta_1 \\ & + \|\mathbf{A}\|_F \|\mathbf{Q}\|_F \left(\frac{1}{2}|\delta_1|^2 + \frac{1}{2}|\tilde{\mathbf{v}}_o|^2 \right) - k_{2o} \tilde{\mathbf{v}}_o^\top \mathbf{Q} \tilde{\mathbf{v}}_o \\ & + \frac{1}{2} \|\mathbf{A}_2\|_F \|\mathbf{Q}\|_F (|\delta_2|^2 + |\tilde{\mathbf{v}}_o|^2) - c_1 \delta_2^\top (c_2 \mathbf{I}_3 - \mathbf{H}) \delta_2 \\ & + \frac{1}{2} \|\mathbf{A}_2\|_F \|\mathbf{Q}\|_F (|\delta_3|^2 + |\tilde{\mathbf{v}}_o|^2) - c_3 |\delta_3|^2. \end{aligned} \quad (35)$$

The above inequality is finally recast as

$$\begin{aligned} \dot{V} \leq & -(k_{1o} - \mu_{1o})|\tilde{\delta}_{1o}|^2 - \tilde{\mathbf{v}}_o^\top (k_{2o} \mathbf{Q} - \mu_{2o} \mathbf{I}_3) \tilde{\mathbf{v}}_o \\ & - \delta_1^\top (c_1 \mathbf{H} - \mu_{1c} \mathbf{I}_3) \delta_1 - (c_3 - \mu_{3c})|\delta_3|^2 \\ & - \delta_2^\top \left(c_1 (c_2 \mathbf{I}_3 - \mathbf{H}) - \mu_{2c} \mathbf{I}_3 \right) \delta_2, \end{aligned} \quad (36)$$

where μ_{1o} , μ_{2o} , μ_{1c} , μ_{2c} , and μ_{3c} have been introduced by (12). Taking into account the lower bounds of observer and control gains set by (11), one has $\dot{V}$ negative definite ($\dot{V} < 0$). Accordingly, the convergence statement of the controller presented earlier in this section follows.

V. RESULTS

The velocity-free controller proposed in the presented paper is tested via computer simulations using C++ programming language. The robotic task consists in automatically positioning the UAV with respect to an observed object. The object is represented by a set of four points. In this section, we report results of both ideal and error-noisy environments.

The quadrotor considered for the study is of mass $m = 0.55$ kg and inertia $\mathbf{I} = 9e^{-3} \text{ diag}(1, 1, 2)$ kg · m², operating in an environment of gravity $g = 9.8$ m/s². As for the communication, the vehicle accepts control commands at a rate of 1 KHz, while the images are provided by the on-board camera at a rate of 25 Hz. The target points are placed in the 3D space at the following positions with respect to base frame $\{R_b\}$: ${}^bP_1 = (0.5, 0.5, 2)^\top$ m; ${}^bP_2 = (0.5, -0.5, 2)^\top$ m; ${}^bP_3 = (-0.5, -0.5, 2)^\top$ m; and ${}^bP_4 = (-0.5, 0.5, 2)^\top$ m. As for the UAV locations, its desired position with respect to $\{R_b\}$ is set to ${}^b\mathbf{t}_s = (0, 0, 4)^\top$ m, while the initial one to $(-1.3, 1, 7)^\top$ m. We empirically tune the controller gains to $k_{1o} = 7$, $k_{2o} = 0.5$, $c_1 = 0.4$, $c_2 = 9$, and $c_3 = 10$. Constant λ involved in definition (4) of visual feedback δ_1 is set to 7. The observer initial values on the velocity are randomly set to $\hat{\mathbf{v}}|_{t=0} = (10, 10, 10)^\top$ cm/s. As for the initial values of $\hat{\delta}_1$, they are set to $\delta_1|_{t=0}$, since the latter is obtained from the camera image.

We first consider an ideal case, where neither error nor noise is introduced. This is interesting to deduce the main outcome of the controller. The corresponding results are shown in Fig. 3. In Fig. 3(f), estimated translational velocity $\hat{\mathbf{v}}$ follows actual one $\mathbf{v}$, besides the observer runs 40 times slower than the vehicle control update rate. Note that by increasing gains k_{1o} and k_{2o} we obtained even better tracking of velocity $\mathbf{v}$. Nevertheless, we preferred to keep the reported gain values to make a trade-off with robustness to image noise and modeling errors. Especially, we can see that the feedback visual errors converge to zero and that the UAV automatically reaches the desired position [Figs. 3(a) and 3(b)]. The vehicle trajectory in the 3D space is depicted in Fig. 3(d), while the variation of the image points in 3(c). The control inputs, thrust u and torque Γ to the UAV, are acceptable [see Fig. 3(e)]. These results thus validate the proposed controller in the ideal case.

In the simulation above, the feature depth is assumed known. This holds when adopting stereo vision, as highlighted earlier. It is not the case when using only one on-board camera. To deal with this it is common to replace the actual depth with the desired one Z^* ; that measured at the desired configuration [6]. We do so in the following simulation in order to test the robustness of the system to the corresponding error. In addition, the on-board images are corrupted with a White Gaussian noise (WGN) of 1 cm standard deviation (SD), the measures of angular velocity ω are corrupted with a WGN of 2 deg/s SD, and the roll, pitch, and yaw angles measures are corrupted with WGNs of 2 deg SD respectively. Corresponding results are shown in Fig. 4. From Fig. 4(f), we can notice that velocity estimate $\hat{v}_z$

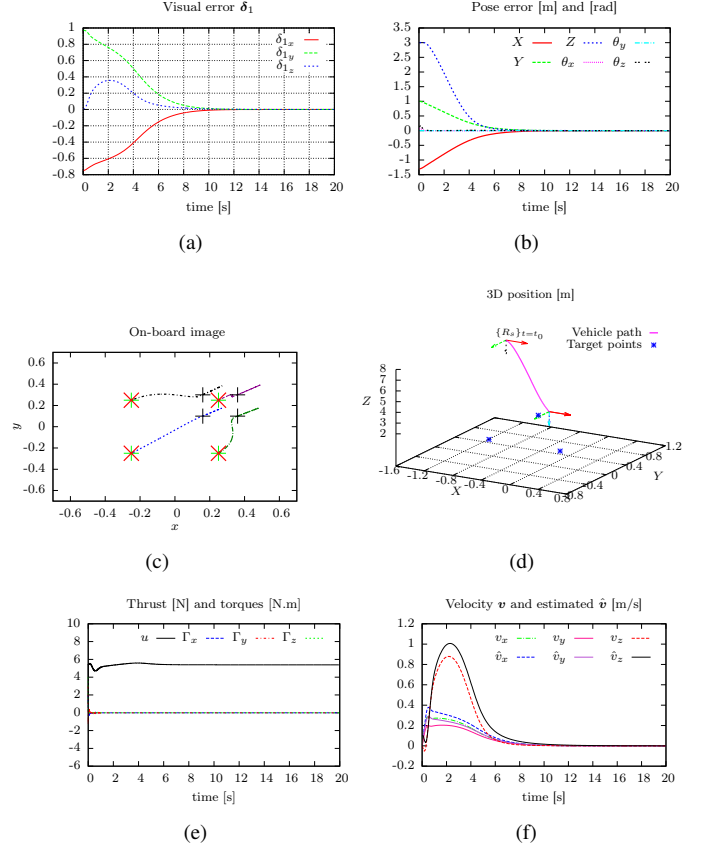


Fig. 3. Velocity-free visual control of the UAV in an ideal case: (c) evolution of the image points. The desired, initial, and final locations are respectively denoted by red "x", black "+", and green "+" symbols - (d) the UAV moves from up to down. The UAV body-fixed frame $\{R_s\}$ is depicted with the red, green, and blue axes referring respectively to its X, Y, and Z axes.

follows well the actual one. Especially, the feedback visual errors decrease to zero and the UAV reaches the desired location [Fig. 4(a) and 4(b)], which is the main objective of this work. The vehicle trajectory is depicted in Fig. 4(d), while the evolution of the points in the image in Fig. 4(c). Note that we tested the system for errors on Z up to for example 300% (overestimation). The system still converges with acceptable behavior. These results clearly show the robustness of the proposed velocity-free controller to both errors on image depth and sensor noise.

VI. CONCLUSION

In this paper we presented a velocity-free visual servo controller to automatically stabilize rotary wing aerial vehicles with respect to visual targets. Our algorithm does not employ exteroceptive sensors nor IMU accelerometers. The former yield to non-autonomy of the system, where the latter suffer from noisy and biased measures and are considered not reliable. Instead, in the present work only vision, along with IMU angular readings, is employed as sensing modality. To do so, a nonlinear observer for velocity estimation was first proposed. Then, a low-level nonlinear controller was derived that exploits the observer outcomes. The system

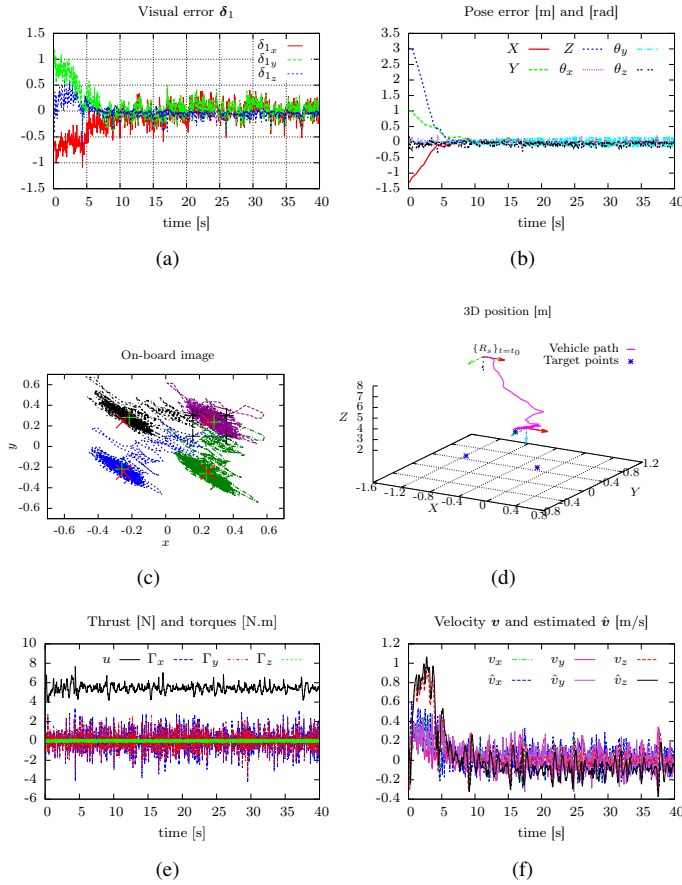


Fig. 4. Velocity-free visual servoing in noisy case with desired Z^* instead of actual image depth Z in the controller.

stability proof has been established with Lyapunov theory, where the validation has been performed with computer simulations. Note that the system performed well albeit the observer is updated with a slow rate compared to that at which the vehicle operates and thus the velocity changes; 40 times slower in this case study. Although the image depth is assumed known in the stability proof, we noticed that the system is very robust to this measure, besides the robustness to sensor noises. Notice nonetheless that such a depth can in fact be retrieved using existing vision techniques of the literature. Future works consists in testing the algorithms on a real hardware system.

REFERENCES

- [1] V. Lippiello and B. Siciliano, "Wall inspection control of a VTOL unmanned aerial vehicle based on a stereo optical flow," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 4296–4302, 2012.
- [2] V. Lippiello and F. Ruggiero, "Exploiting redundancy in cartesian impedance control of UAVs equipped with a robotic arm," in *IEEE/RSJ Int. Conf. on Intelligent Robots and Systems*, pp. 3768 – 3773, 2012.
- [3] A. D. Wu, E. N. Johnson, and A. A. Proctor, "Vision-aided inertial navigation for flight control," in *Journal of Aerospace Computing, Information, and Communication*, vol. 2, pp. 348–360, 2005.
- [4] R. J. Sattigeri, E. Johnson, and A. J. Calise, "A vision-based target tracking with adaptive target state," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, August 2008.
- [5] J. Prasad, A. J. Calise, E. N. Johnson, R. Sattigeri, and J. Moon, "Flight demonstration of an adaptive guidance controller for autonomous formation flight," in *American Helicopter Society 64th, Annual Forum*, April–May 2008.
- [6] F. Chaumette and S. Hutchinson, "Visual servo control, part I: Basic approaches," *IEEE Robotics and Automation Magazine*, vol. 13, pp. 82–90, 2006.
- [7] P. Castillo, A. Dzul, and R. Lozano, "Real-time stabilization and tracking of a four-rotor mini rotorcraft," *IEEE Trans. on Control Systems and Technology*, vol. 12, pp. 510 – 516, 2004.
- [8] S. Bertrand, T. Hamel, and H. Piet-Lahanier, "Stabilization of a small unmanned aerial vehicle model without velocity measurement," in *IEEE Int. Conf. on Robotics and Automation*, pp. 724 – 729, April 2007.
- [9] S. Shen, N. Michael, and V. Kumar, "Autonomous multi-floor indoor navigation with a computationally constrained MAV," in *IEEE Int. Conf. on Robotics and Automation*, pp. 20–25, May 2011.
- [10] S. Ahrens, D. S. Levine, G. Andrews, and J. P. How, "Vision-based guidance and control of a hovering vehicle in unknown, GPS-denied environments," in *IEEE Int. Conf. on Robotics and Automation*, pp. 2643–2648, May 2009.
- [11] M. Achtelik, M. Achtelik, S. Weiss, and R. Siegwart, "Onboard IMU and monocular vision based control for MAVs in unknown in- and outdoor environments," in *IEEE Int. Conf. on Robotics and Automation*, May 2011.
- [12] L. Kneip, A. Martinelli, S. Weiss, D. Scaramuzza, and R. Siegwart, "Closed-form solution for absolute scale velocity determination combining inertial measurements and a single feature correspondence," in *IEEE Int. Conf. on Robotics and Automation*, pp. 4546–4553, May 2011.
- [13] V. Lippiello, G. Loianno, and B. Siciliano, "MAV indoor navigation based on a closed-form solution for absolute scale velocity estimation using optical flow and inertial data," in *IEEE Int. Conf. on Decision and Control/European Control Conf.*, pp. 3566–3571, December 2011.
- [14] K. Benzemrane, G. L. Santosuosso, and G. Damm, "Unmanned aerial vehicle speed estimation via nonlinear adaptive observers," in *American Control Conf.*, pp. 985–990, July 2007.
- [15] M. Boutayeb, R. Richard, H. Rafaralahy, H. S. Ali, and G. Zaloylo, "A simple time-varying observer for speed estimation of UAV," in *17th World Congress. IFAC*, pp. 1760–1765, July 2008.
- [16] G. Baldwin, R. E. Mahony, and J. Trumpf, "A nonlinear observer for 6 DOF pose estimation from inertial and bearing measurements," in *IEEE Int. Conf. on Robotics and Automation*, pp. 2237–2242, 2009.
- [17] T. Chevion, T. Hamel, R. E. Mahony, and G. Baldwin, "Robust nonlinear fusion of inertial and visual data for position, velocity and attitude estimation of UAV," in *IEEE Int. Conf. on Robotics and Automation*, pp. 2010–2016, 2007.
- [18] A. Abdessameud and A. Tayebi, "Velocity-free consensus algorithms for double-integrator dynamics with input saturations constraints," in *49th IEEE Conf. on Decision and Control*, pp. 4486–4491, 2010.
- [19] V. Grabe, H. H. Bühlhoff, and P. R. Giordano, "On-board velocity estimation and closed-loop control of a quadrotor UAV based on optical flow," in *IEEE Int. Conf. on Robotics and Automation*, pp. 491–497, 2012.
- [20] Y. Ma, S. Soatto, J. Kosecka, and S. S. Sastry, *An Invitation to 3-D Vision: From Images to Geometric Models*. Springer–Verlag, 2003.
- [21] R. Mahony, P. Corke, and T. Hamel, "Dynamic image-based visual servo control using centroid and optic flow features," *Journal of Dynamic Systems, Measurements, and Control*, vol. 130, 2008.
- [22] N. Guenard, T. Hamel, and R. E. Mahony, "A practical visual servo control for an unmanned aerial vehicle," *IEEE Trans. on Robotics*, vol. 24, pp. 331–340, 2008.
- [23] M. Krstic, P. V. Kokotovic, and I. Kanellakopoulos, *Nonlinear and Adaptive Control Design*. New York, NY, USA: John Wiley & Sons, Inc., 1st ed., 1995.
- [24] O. Shakhnia, T. Koo, and S. Sastry, "Landing an unmanned air vehicle: vision based motion estimation and nonlinear control," *Asian Journal of Control*, vol. 1, pp. 128–145, 1999.
- [25] T. Hamel and R. Mahony, "Visual servoing of an under-actuated rigid body system: An image based approach," *IEEE Trans. on Robotics and Automation*, vol. 18, pp. 187–198, 2002.
- [26] R. J. Baerveldt and R. Klang, "A low-cost and low-weight attitude estimation system for an autonomous helicopter," in *Proc. Intelligent Engineering Systems*, pp. 391–395, 1997.
- [27] R. I. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2nd ed., 2004.