Homography Estimation of a Moving Planar Scene from Direct Point Correspondence

Simone de Marco, Minh-Duc Hua, Robert Mahony, Tarek Hamel

Homography Estimation of a Moving Planar Scene from Direct Point Correspondence

Simone de Marco, Minh-Duc Hua, Robert Mahony, Tarek Hamel

***Abstract*—Homography estimation is an important task in robotics applications such as landing on moving platforms, docking and refueling, building inspection, etc. Non-linear observers depend on an estimate of the group algebra velocity as a feed-forward term to minimize lag in the filter response. When both the camera and the scene are moving, and for perspective image constructs such as the homography, it is often impossible to directly measure the required group velocity. This paper proposes a solution for the case where the motion is periodic, or approximately periodic, with known period. The approach is based on the internal model principle, where the internal model can be expanded to include a sufficient harmonics of the desired period in order to model complex periodic motions. The novelty of the work lies in formalizing the internal model for observer design to the non-compact Lie group $\mathrm{SL}(3)$ and providing a demonstration of the effectiveness of the approach with real-world examples.**

***Index Terms*—Homography, Internal Model Observer, Non-linear Observer.**

I. Introduction

Recent years have witnessed an increase in the use of autonomous robotic vehicles for a wide range of civilian and commercial applications. There is a growing interest in applications where the autonomous vehicle physically interacts with the surrounding environment. These operations include landing on moving platforms, docking and refueling, building inspections and search and rescue missions. Obtaining a good measure of the position of the vehicle during such missions is a crucial to the system performance. Global Positioning Systems (GPS) are widely used for navigation and control in open areas, however, reliable and accurate GPS localization in the applications of interest is not possible due to tall buildings, urban canyons, dense material such as concrete and steel and many other factors that can block or heavily degrade the GPS signals. Moreover, many missions require localization with respect to some particular features in the environment instead of an absolute position. For these reasons, the study of alternative localization systems has become popular in the robotics and control community. Examples of such localization systems are 2D laser scanner (LIDAR), radar, cameras (monocular, stereo, RGB-D, event), etc. Active sensor systems such as LIDAR are well suited for large-scale vehicles, however since they are usually heavy, expensive and power hungry they cannot be used on small-scale vehicles. In contrast, cameras are cheap, lightweight, small and convenient to mount and provide an information rich measurement. Therefore, vision-based systems remain one of the most promising technologies especially for small-scale unmanned aerial vehicles (UAV). Vision-based control algorithms are known in literature as visual servoing [8], [9], [30].

Any two images of the same planar scene taken from two different camera poses are related by a projective mapping known as a homography. Homographies play a major role in rigid-body pose estimation [10], [37], unmanned vehicle navigation [32], [33], UAV autonomous approach and landing [18], [39], to name only a few applications. Using homographies is effective when the surrounding environment is composed mainly of planar surfaces for example in man-made environments, urban scenes and also when the relief of the scene viewed is negligible compared to the distance of the camera from the scene. Estimating homography from pairs of images has been extensively studied in the computer vision literature [20]. Classical algorithms estimate homographies by minimizing a suitable cost function given a set of image feature (points, lines, conics) correspondences between image pairs. For such algorithms, there is no difference whether the images considered are generated by a temporal correlated sequence (such as in a video) or are taken independently. When a robotic vehicle moves, however, the camera captures a temporal sequence of images for which the homography varies continuously. This has lead to the development of a number of non-linear continuous-time observers [19], [21], [22], [31] that exploit the structure of the Special Linear group $\mathrm{SL}(3)$, a Lie-group isomorphic to the group of homographies [4]. A disadvantage of non-linear observers is that they rely on group velocity measurements to propagate the state estimate with time. Although robotic systems are usually provided with an inertial measurement unit, and other velocity measurements such as doppler GPS and optical flow, the homography velocity depends also on scene parameters and is difficult to reconstruct. Furthermore, if the scene itself is moving, such as for landing on a moving platform, ego motion estimation of the robot by itself cannot estimate the velocity of the relative homography. The only algorithms the authors are aware of that address this issue use an integral adaptive term [19], [22], [31] to estimate 'constant' unknown velocity. This assumption is restrictive since constant group velocity motions for the homography group are not intuitive motions for a rigid-body robotic system with a mounted camera moving over a moving scene, and it may introduce significant lag in the filter response when used in real-world scenarios [19], [22], [31].

S. de Marco and M.-D. Hua are with I3S, Université de Nice, Sophia-Antipolis, CNRS, France sdemarco(hua)@i3s.unice.fr

R. Mahony is with the Australian Centre for Robotic Vision, Australian National University, Robert.Mahony@anu.edu.au.

T. Hamel is with I3S, Université Côte d'Azur, Institut Universitaire de France, CNRS, Sophia Antipolis, France, thamel@i3s.unice.fr.

The problem of regulating the output of a system in order to track an unknown reference trajectory while rejecting disturbances is known as an output regulation problem (or servomechanism problem) [11], [15]. The case where the desired reference can be modeled as the solution to an ordinary differential equation with unknown initial condition can be solved by internal model-based regulators. This problem has been extensively studied over the last thirty years in both linear [11], [15], [16] and non-linear context [24], [6], [38], leading to the so-called *Internal Model Principle*. The central part of the regulator design incorporates an internal model unit, which is a dynamical autonomous system that generates in steady state the feed-forward input that keeps the tracking error identically to zero. Recently, the theory of non-linear observers has been shown to play a central role in the output regulation problem. In particular, adaptive observers [2] and high-gain observers [17], [1] have been successfully used for the design of adaptive internal models [34] and robust internal models [25], [5].

In this paper, we take advantages of recent results [13] on output regulation for systems on matrix Lie-groups to develop an internal model-based observer design for the homography estimation problem. The unknown group velocity, that acts like an unknown reference trajectory, is modeled as belonging to a family of trajectories generated as solutions of an autonomous linear system on the group Lie algebra. The proposed observer, in the same spirit of linear internal model principle, embeds a copy of the exogenous dynamics (posed on the Lie algebra) and incorporates an innovation to synchronize these dynamics with the exogenous plant providing a driving term for the velocity reference of a standard kinematic observer on the Lie group. The novelty of the paper, which is based on the direct point formulation proposed in [19], is its ability to learn complex feed-forward velocity profiles to reduce lag in the homography tracking where motion is periodic. This has application in a wide range of applications such as landing an aerial robot on an oscillating platform or stabilizing periodic camera shake. A preliminary version of the theoretical results of this paper was presented in [12]. Here we provide a more extensive stability analysis extending the local asymptotic results in [12] to local uniform exponential results as well as providing a more detailed experimental study and full details of the derivations. The experimental results provided demonstrate excellent performance and robustness in presence of challenging conditions and outperform the state-of-the-art results documented in [19].

The paper is organized into five sections including the present introduction and an appendix. Section II introduces the notation and formulation. The non-linear observer is detailed in Section III. Simulation results are provided in Section IV. Section V provides a detailed description of the software implementation of the observer. In the same section an application of the proposed approach for the tracking of an oscillating platform, with comparisons of the observer proposed in [19], is presented. A video link of the experiment (https://goo.gl/9Y3KsB) is provided as supplementary material. The Appendix contains proofs of some useful properties for the observer design, a rigorous analysis of the local exponential stability of the proposed observer, and an important lemma that is not in the direct line of argument for the main result.

II. Problem formulation

A. Notation and mathematical identities

The set of all $m \times n$ matrices whose entries are real numbers is denoted $\mathbb{R}^{m\times n}$. For any $A \in \mathbb{R}^{n\times n}$, $\det(A)$ and $\mathrm{tr}(A)$ denote the determinant and the trace of A respectively.

The Special Orthogonal group is denoted $\mathrm{SO}(3)$. The Lie algebra associated to the Special Orthogonal group, denoted by $\mathfrak{so}(3)$, is the set of 3×3 skew symmetric matrices

$$\mathfrak{so}(3) := \left\{\Omega_\times \in \mathbb{R}^{3\times 3} | \Omega_\times + \Omega_\times^\top = 0\right\}.$$

The Lie algebra $\mathfrak{so}(3)$ with the matrix commutator (Lie bracket) $[\cdot,\cdot]$ is isomorphic to $\mathbb{R}^3$ with the cross product. Let $\Omega = [\Omega_1, \Omega_2, \Omega_3]^\top \in \mathbb{R}^3$, then the matrix

$$\Omega_\times = \begin{bmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{bmatrix}$$

is the skew symmetric matrix associated to the cross product $\Omega_\times v = \Omega \times v$, for any $v \in \mathbb{R}^3$.

The Special Linear group $\mathrm{SL}(3)$ and its associated Lie algebra $\mathfrak{sl}(3)$ are defined by

$$\begin{array}{rcl} SL(3) & := & \left\{H \in \mathbb{R}^{3\times 3} | \det(H) = 1\right\}, \\ \mathfrak{sl}(3) & := & \left\{U \in \mathbb{R}^{3\times 3} | \mathrm{tr}(U) = 0\right\}. \end{array}$$

For $H \in \mathrm{SL}(3)$ and $U \in \mathfrak{sl}(3)$, the adjoint operator is a mapping $Ad : \mathrm{SL}(3) \times \mathfrak{sl}(3) \to \mathfrak{sl}(3)$ defined by

$$Ad_H U := HUH^{-1}.$$

For any $A, B \in \mathbb{R}^{n\times n}$, $\langle\langle A, B\rangle\rangle = \mathrm{tr}(A^\top B)$ defines an inner product on $\mathbb{R}^{n\times n}$, and the Frobenius norm $\|\cdot\|_F$ is defined by

$$\|A\|_F := \sqrt{\langle\langle A, A\rangle\rangle}.$$

Let $\mathbb{P}_{\mathfrak{sl}3}(\cdot)$ denote the unique orthogonal projection of $\mathbb{R}^{3\times 3}$ onto $\mathfrak{sl}(3)$ with respect to the trace inner product, one has

$$\mathbb{P}_{\mathfrak{sl}3}(A) = \left(A - \frac{1}{3}\mathrm{tr}(A) I\right) \in \mathfrak{sl}(3), \quad \forall A \in \mathbb{R}^{3\times 3},$$

with I the identity matrix.

Let $\wedge$ denotes the mapping $\wedge : \mathbb{R}^8 \to \mathfrak{sl}(3)$ that maps the vector $v \in \mathbb{R}^8$ to an element of $\mathfrak{sl}(3)$

$$v_\wedge := \sum_{j=1}^{8} v_j B_j$$

where $\{B_1, \ldots, B_8\}$ is a basis of $\mathfrak{sl}(3)$. For instance, denoting by $b_1 = [1,0,0]^\top$, $b_2 = [0,1,0]^\top$, $b_3 = [0,0,1]^\top$, the following eight generators form a basis for the lie algebra $\mathfrak{sl}(3)$

$$\begin{array}{rclrclrcl} B_1 & := & b_1 b_1^\top - I_3, & B_4 & = & b_1 b_2^\top, & B_7 & = & b_1 b_3^\top, \\ B_2 & := & b_2 b_1^\top, & B_5 & = & b_2 b_2^\top - I_3, & B_8 & = & b_2 b_3^\top, \\ B_3 & := & b_3 b_1^\top, & B_6 & = & b_3 b_2^\top. & & & \end{array} \tag{1}$$

The operator $\text{vec}^\vee : \mathfrak{sl}(3) \to \mathbb{R}^8$ denotes the inverse of the $(\cdot)_\wedge$ operator, namely

$$\text{vec}^\vee(v_\wedge) = v, \qquad \forall v \in \mathbb{R}^8.$$

For any $A \in \mathbb{R}^{n\times n}$, $\text{vect}(A) \in \mathbb{R}^{n^2}$ denotes the column vector obtained by the concatenation of columns of the matrix A as follows

$$\text{vect}(A) = [a_{1,1}, \cdots, a_{n,1}, a_{1,2}, \cdots, a_{n,2}, \cdots, a_{1,n}, \cdots, a_{n,n}]^\top.$$

The matrix representation of the composition of the linear maps $(\text{vect} \circ \wedge) : \mathbb{R}^8 \to \mathfrak{sl}(3) \to \mathbb{R}^9$ is denoted by the full columns rank matrix $[\![\wedge]\!] \in \mathbb{R}^{9\times 8}$. Let $\mathcal{Q}_{\mathfrak{sl}3} \in \mathbb{R}^{8\times 8}$ denote the symmetric positive matrix $Q_{\mathfrak{sl}3} := [\![\wedge]\!]^\top [\![\wedge]\!]$.

For any $p \in \mathbb{S}^2$, the projection $\pi_p := (I - pp^\top)$ onto the tangent space of the unit sphere $\mathbb{S}^2$ at point p satisfies $\pi_p p = 0$.

B. Perspective Projection

Let $\mathcal{B}$ denote projective coordinates for the image plane of a camera, and $\{B\}$ its frame of reference. Let $\xi \in \mathbb{R}^3$ denote the position of the body-fixed frame $\{B\}$ with respect to the inertial frame $\{I\}$ expressed in $\{I\}$. The orientation of the frame $\{B\}$ with respect to the inertial frame $\{I\}$ is given by the rotation matrix $R \in \text{SO}(3)$. Let $\mathcal{T}$ denote the projective coordinates of reference image of the target plane, and let $\{T\}$ denote its right-hand frame of reference. Let $(\xi_T, R_T) \in \mathbb{R}^3 \times \text{SO}(3)$ denote the configuration of the frame $\{T\}$ with respect to the inertial frame $\{I\}$.

The coordinates of a single point in the inertial frame $P_I \in \{I\}$, in the camera frame $P_B \in \{B\}$ and in the target frame $\mathring{P} \in \{T\}$ (see Figure 1) are related by

$$P_I = R_T \mathring{P} + \xi_T, \tag{2a}$$

$$P_I = R P_B + \xi. \tag{2b}$$

The orientation of the body-fixed frame $\{B\}$ with respect to the target frame $\{T\}$, is denoted by the rotation matrix $\bar{R} : \{B\} \to \{T\}$, while the relative position of the frame $\{B\}$ with respect to the target frame $\{T\}$ expressed in the frame $\{T\}$ is denoted by $\bar{\xi}$, one verifies that

$$\bar{R} = R_T^\top R, \tag{3a}$$

$$\bar{\xi} = -R_T^\top(\xi_T - \xi). \tag{3b}$$

Substituting (2a) into (2b) one has

$$P_B = \bar{R}^\top \mathring{P} - \bar{R}^\top \bar{\xi} \tag{4}$$

as relation between the coordinates of the same point in the target frame and in the body-fixed frame.

Considering the classical pinhole camera model [14], [29], $3D$ points can be projected on the image planes $\mathcal{B}$ and $\mathcal{T}$. Let $p \in \mathcal{B}$ denote the image of a point when the camera is aligned with the body-fixed frame $\{B\}$ and let $\mathring{p} \in \mathcal{T}$ denote the image of the same point when the camera is aligned with the target frame $\{T\}$, then[1]

$$\mathring{p} \cong K\mathring{P}, \quad p \cong K P_B$$

[1]The notation "$\cong$" is defined in the homogeneous sense, i.e. up to a non-zero scalar factor [29].

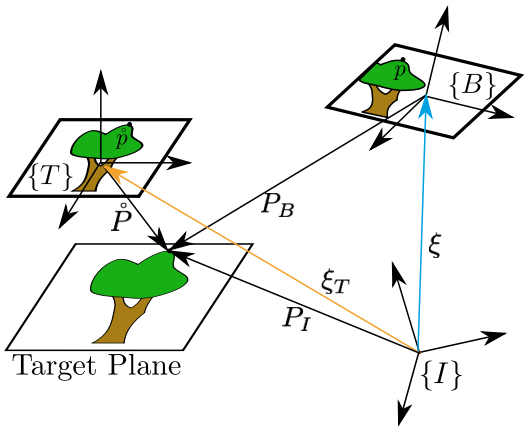


Fig. 1. Representation of a point of the planar surface in the inertial frame $\{I\}$, body-fixed frame $\{B\}$ and target frame $\{T\}$.

where K is the upper triangular camera matrix which depends on intrinsic parameters of the camera, namely focal length, pixels aspect ratio and principal point offset.

C. Homography matrix and the Special Linear Group SL(3)

Let $\mathring{\eta}$ denote the unit normal vector pointing towards the target plane expressed in $\{T\}$, and let $\mathring{d}$ denote the orthogonal distance of the plane to the origin of $\{T\}$.

Due to the fact that all target points $\mathring{P}_i$ for $i = \{1, \ldots, n\}$ lie on a single planar surface one has

$$\frac{\mathring{\eta}^\top \mathring{P}_i}{\mathring{d}} = 1,$$

it follows from (4) that

$$P_{Bi} = \left(\bar{R}^\top - \bar{R}^\top \bar{\xi} \frac{\mathring{\eta}^\top}{\mathring{d}}\right) \mathring{P}_i, \tag{5}$$

and thus, using the perspective projection, one obtains

$$\mathring{p}_i \cong K \left(\bar{R}^\top - \bar{R}^\top \bar{\xi} \frac{\mathring{\eta}^\top}{\mathring{d}}\right)^{-1} K^{-1} p_i. \tag{6}$$

The projective mapping $G : \mathcal{B} \to \mathcal{T}$ that maps pixel coordinates from $\mathcal{B}$ to $\mathcal{T}$ given by

$$G = \zeta K \left(\bar{R}^\top - \bar{R}^\top \bar{\xi} \frac{\mathring{\eta}^\top}{\mathring{d}}\right)^{-1} K^{-1},$$

where ζ represents a scaling factor, is known in literature as homography matrix.

Let η (resp. η_I) denote the normal to the target plane expressed in the frame $\{B\}$ (resp. $\{I\}$), and let d (resp. d_I) denote the orthogonal distance of the planar surface to the origin of the body-fixed frame $\{B\}$ (resp. inertial frame $\{I\}$). One verifies that:

$$\eta = R^\top \eta_I = \bar{R}^\top \mathring{\eta}, \tag{7}$$

$$d = d_I - \eta_I^\top \xi = \mathring{d} - \mathring{\eta}^\top \bar{\xi}. \tag{8}$$

It follows that the homography matrix can be written as

$$G = \zeta K \left(\bar{R} + \frac{\bar{\xi}\eta^\top}{d}\right) K^{-1}.$$

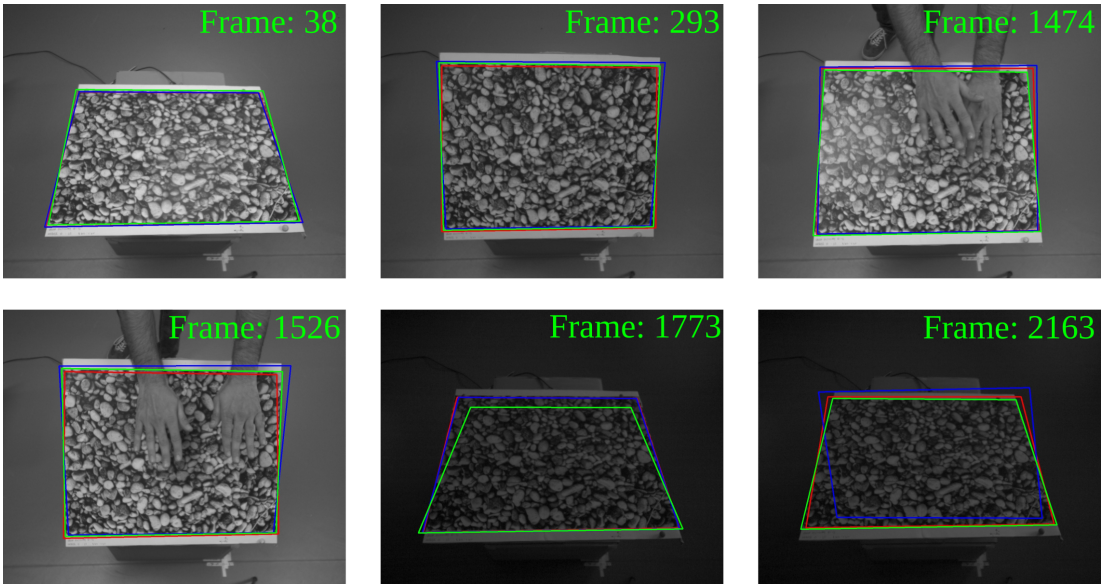


Fig. 2. The red, blue and green polygon represents the region of interest tracked by the proportional, proportional-integral and internal model-based filter respectively.

If the camera is calibrated the homography matrix can be scaled

$$H = K^{-1}GK = \zeta\left(\bar{R} + \frac{\bar{\xi}\eta^\top}{d}\right). \tag{9}$$

This mapping is referred to as Euclidean homography since it maps Euclidean coordinates of the scene's points from $\{B\}$ to $\{T\}$. Indeed, the Euclidean homography contains the relative pose information of the camera body-fixed frame $\{B\}$ with respect to the target frame $\{T\}$. From now on we restrict our attention to the Euclidean homography H.

Since the homography H is only defined up to a scale factor, it is possible to represent it uniquely as an element of the Special Linear group $\mathrm{SL}(3)$ by choosing ζ such that $\det(H) = 1$, namely

$$H' = H\det(H)^{-\frac{1}{3}} \in \mathrm{SL}(3).$$

For the remainder of this paper, we assume that the calibrated homograpy is appropriately scaled in such a way $H \in \mathrm{SL}(3)$.

D. Homography Kinematics and Measurements

Denote the angular velocity and linear velocity of the target frame $\{T\}$ with respect to the inertial frame $\{I\}$ expressed in the target frame $\{T\}$ by Ω_T and V_T. The target kinematics of (R_T, ξ_T) are given by:

$$\dot{R}_T = R_T\Omega_{T\times}, \quad \dot{\xi}_T = R_T V_T. \tag{10}$$

Let Ω denote the angular velocity of the body-fixed frame expressed in $\{B\}$ and let V denote the linear velocity of the body-fixed frame $\{B\}$ with respect to the inertial frame $\{I\}$ expressed in the frame $\{B\}$. The kinematics of (R, ξ) are given by:

$$\dot{R} = R\Omega_\times, \quad \dot{\xi} = RV. \tag{11}$$

It follows that the kinematics of the relative attitude and position of the body-fixed frame $\{B\}$ with respect to the target frame $\{T\}$ expressed in $\{T\}$ can be written as

$$\dot{\bar{R}} = \bar{R}\bar{\Omega}_\times, \quad \dot{\bar{\xi}} = \bar{R}\bar{V} \tag{12}$$

where

$$\bar{\Omega}_\times = \Omega_\times - (\bar{R}^\top\Omega_T)_\times, \tag{13a}$$
$$\bar{V} = \bar{R}^\top\Omega_{T\times}\bar{\xi} + V - \bar{R}^\top V_T. \tag{13b}$$

Lemma 1. *Consider a camera moving with kinematics* (11) *viewing a moving planar scene whose kinematics are defined by* (10). *Let* $H : \{B\} \to \{T\}$ *denote the Euclidean homography matrix. The group velocity* $\bar{U} \in \mathfrak{sl}(3)$ *induced by the relative motion between the camera and the target plane is such that*

$$\dot{H} = H\bar{U}, \text{ with } \bar{U} = \bar{\Omega}_\times + \frac{\bar{V}\eta^\top}{d} - \frac{\eta^\top\bar{V}}{3d}I. \tag{14}$$

Proof: See Appendix A. ■

Note that the group velocity $\bar{U}$ induced by the relative motion between the camera and the target plane depends on the normal to the surface η and the orthogonal distance d, that define the scene geometry at time t, as well as the angular velocity of the target plane Ω_T. The scene parameters and the velocity of the target are usually unmeasurable and they have to be taken into account during the observer design. To this purpose we rewrite the group velocity as

$$\bar{U} = \Omega_\times + \Gamma_\wedge, \tag{15}$$

where Ω is typically the measurable part, which can be obtained from the set of embedded gyros and can be directly used in the observer as feed-forward term. While the remainder of the group velocity $\Gamma_\wedge$ represents the unmeasurable part that has to be estimated.

In order to cope with the unmeasurable group velocity, we assume that $\Gamma_\wedge$ can be immersed into a finite-dimensional and observable linear system ([7], [26]) in such a way it can be written in the form

$$\Gamma_\wedge = (C\gamma)_\wedge, \quad \dot{\gamma} = S\gamma \tag{16}$$

where the known matrix $C \in \mathbb{R}^{8\times m}$ is a full row rank matrix with $m \geq 8$, $\gamma \in \mathbb{R}^m$ and the known matrix $S \in \mathbb{R}^{m\times m}$ is skew symmetric, namely $S = -S^\top$.

Remark. *The eigenvalues of a real skew-symmetric matrix are all purely imaginary and they come in conjugate pairs. In case of odd dimension there is an additional unpaired null eigenvalue. The system in* (16) *is stable in the Lyapunov sense. Indeed considering the Lyapunov candidate* $\mathcal{L} = \|\gamma\|^2$*, and differentiating it along the solutions of* (16) *one obtains* $\dot{\mathcal{L}} = 2\gamma^\top S\gamma = 0$*. It follows that any autonomous observable system of the form*

$$\begin{array}{rcll} \dot{x} & = & Ax, & A \in \mathbb{R}^{m\times m},\ x \in \mathbb{R}^m \\ y & = & C_y x, & C_y \in \mathbb{R}^{8\times m},\ y \in \mathbb{R}^8 \end{array}$$

is equivalent to (16) *iff all eigenvalues of* A *have zero real part and multiplicity one in the minimal polynomial. The considered assumption on* $\Gamma_\wedge$ *is of course not valid for all types of relative motions between the camera and the target plane. It, however, allows us to cover many possible camera/target trajectories encountered in practice such as (near) circular or sinusoidal motions.*

We consider a set of n point measurements $p_i \in \mathbb{P}^2$ associated to a group action of $\mathrm{SL}(3)$ onto the projective space $\mathbb{P}^2$

$$p_i = \frac{H^{-1}\mathring{p}_i}{|H^{-1}\mathring{p}_i|}, \quad i = \{1,\dots,n\}, \tag{17}$$

representing calibrated image points projected onto the unit sphere.

III. Non-Linear Observer Design On SL(3)

The goal of the non-linear filter is to provide an estimate $\hat{H} \in \mathrm{SL}(3)$ given a collection of n measurements p_i to drive the error $\tilde{H} := \hat{H}H^{-1}$ to the identity element of the group. To this purpose, we define the estimates $\hat{p}_i \in \mathbb{P}^2$ of p_i as

$$\hat{p}_i := \frac{\hat{H}^{-1}\mathring{p}_i}{\left|\hat{H}^{-1}\mathring{p}_i\right|}, \quad i = \{1,\dots,n\}. \tag{18}$$

The estimates e_i of $\mathring{p}_i$ then is defined as

$$e_i := \frac{\hat{H}p_i}{\left|\hat{H}p_i\right|} = \frac{\tilde{H}\mathring{p}_i}{\left|\tilde{H}\mathring{p}_i\right|}, \quad i = \{1,\dots,n\}. \tag{19}$$

Definition 1. *A set* $\mathcal{M}_n$ *of* $n(\geq 4)$ *vector directions* $\mathring{p}_i \in \mathbb{P}^2$ *is called consistent, if it contains a subset* $\mathcal{M}_4 \subset \mathcal{M}_n$ *of 4 constant vector directions such that all its vector triplets are linearly independent.*

Theorem 1. *Let* H *denote the Euclidean homography* (9) *and consider the kinematic system in* (14) *along with* (15) *and* (16)*. Consider the following nonlinear filter*

$$\dot{\hat{H}} = \hat{H}(\Omega_\times + \hat{\Gamma}_\wedge) + k_p\Delta\hat{H} \tag{20a}$$

$$\hat{\Gamma}_\wedge = (C\hat{\gamma})_\wedge \tag{20b}$$

$$\dot{\hat{\gamma}} = S\hat{\gamma} + k_I C^\top Q_{\mathfrak{sl}3}\sum_{i=1}^{n}\varphi_i \tag{20c}$$

with k_p *and* k_I *some positive gains and*

$$\begin{array}{rcl} \Delta & := & \sum_{i=1}^{n}\pi_{e_i}\mathring{p}_i e_i^\top \quad \in \mathfrak{sl}(3) \\ \varphi_i & := & \mathrm{vec}^\vee\left(\mathbb{P}_{sl3}\left((\hat{H}^\top\mathring{p}_i)(\hat{H}^{-1}e_i)^\top\right)\right) + \\ & & -e_i^\top\mathring{p}_i\,\mathrm{vec}^\vee\left(\mathbb{P}_{\mathfrak{sl}3}\left((\hat{H}^\top e_i)(\hat{H}^{-1}e_i)^\top\right)\right). \end{array} \tag{21}$$

Assume that the measured angular velocity Ω *of the planar target, the group velocity* Γ *(equivalently* γ*) in* (16) *and homography matrix* H *are bounded.*

Then, if the set $\mathcal{M}_n$ *of measured directions* $\mathring{p}_i$ *is consistent, the equilibrium* $(\tilde{H},\tilde{\gamma}) = (I,0)$*, with* $\tilde{H} := \hat{H}H^{-1}$*,* $\tilde{\gamma} := \gamma - \hat{\gamma}$*, of the error system is locally exponentially stable.*

The proof of Theorem 1 is given in Appendix B.

The positive gains k_p and k_I in (20) share the same effects of a proportional gain and an integral gain for a linear system, respectively. In particular k_p directly relates the dynamics of $\hat{H}$ to the innovation, while the gain k_I represents the integral term of the innovation projected into $\mathbb{R}^8$ in the inertial frame.

Remark. *The constant image points* $\mathring{p}_i$ *represent known reference points which can be extracted from a stored reference image of the planar scene or extracted directly during the initialization process of the observer, namely the first image taken by the camera. In the latter case one has* $\mathring{p}_i = p_i(0)$*,* $\forall i = \{1,\dots,n\}$*, whereas* e_i *represent the estimates of the reference points* $\mathring{p}_i$ *based on the current measurements* p_i*. Roughly speaking, in a classical linear Luenberger observer the error* $\sum_i(\mathring{p}_i - e_i)$ *would represent the output prediction error (innovation).*

IV. Simulation Result

In this section we illustrate the performance of the proposed observer through simulations. The camera is assumed to be a stationary camera. The target plane trajectory, depicted in Figure 3, is chosen as a Lemniscate of Gerono with zero angular velocity and linear velocity given by:

$$V_T = \begin{bmatrix} -12\pi\sin(6\pi t) \\ 12\pi\cos(12\pi t) \\ 24\pi\cos(12\pi t) \end{bmatrix}.$$

Note that the group velocity $\Gamma_\wedge$ in (16) induced by the Gerono trajectory cannot be immersed into a finite linear system. However, due to the fact that $\Gamma_\wedge$ is a periodic function it is possible to decompose it into Fourier series (see Figure 4).

The initial homography H is initialized as

$$H(0) = \begin{bmatrix} 2.28 & 0 & 0 \\ 0 & 2.28 & 0 \\ 0 & 0 & 0.19 \end{bmatrix}$$

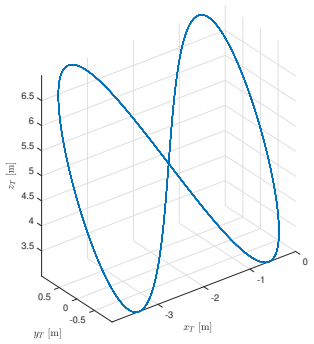


Fig. 3. Target Plane Trajectory.

while the initial estimated homography $\hat{H}(0)$ is set to the identity element of the group. The filter parameters are $k_p = 50$ and $k_I = 10$.

Considering the spectrum of the group velocity induced by the motion of the planar target, depicted in Figure 4, we implemented in the internal model unit the first four harmonics for the basis B_1, B_5, B_7, B_8 defined in (1).

For comparison purpose a second simulation was run considering the observer proposed in [19], using the same value $k_p = 50$ for the proportional term.

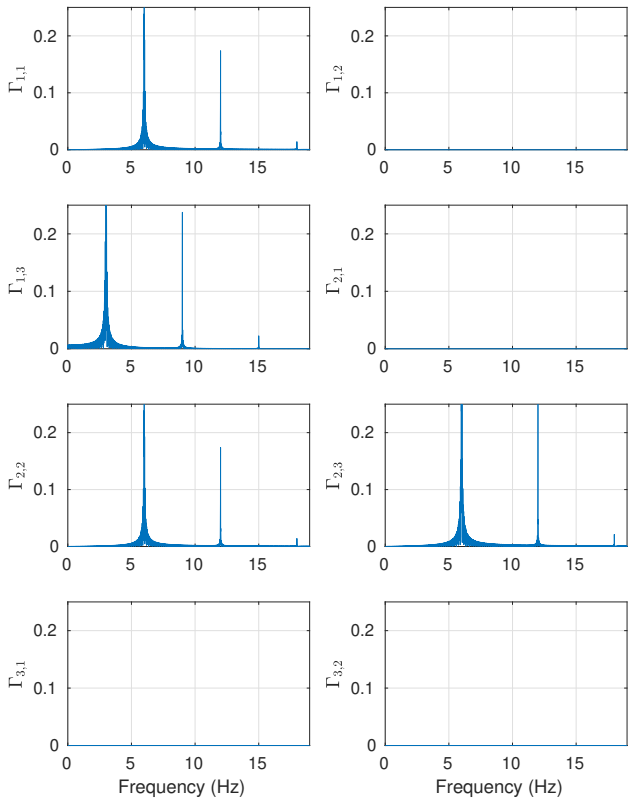


Fig. 4. Frequency spectrum of the group velocity $\Gamma_\wedge$.

The results reported in Figures 5-6 show respectively the time behavior of the Frobenious norm $\|I - \tilde{H}\|_F$ for the observer presented in [19] and the proposed observer (termed OSC in order to highlight the oscillatory nature of the terms in the Lie algebra), the Frobenious norm $\|\Gamma_\wedge - \hat{\Gamma}_\wedge\|_F$, and the group velocity vs. the estimated group velocity. Figure 5 clearly shows that the proposed observer outperform the observer in [19]. Plots show that, in steady state, the velocity $\hat{\Gamma}_\wedge$ of the observer practically converges to $\Gamma_\wedge$ and the estimated homography $\hat{H}$ practically converges to the real homography H.

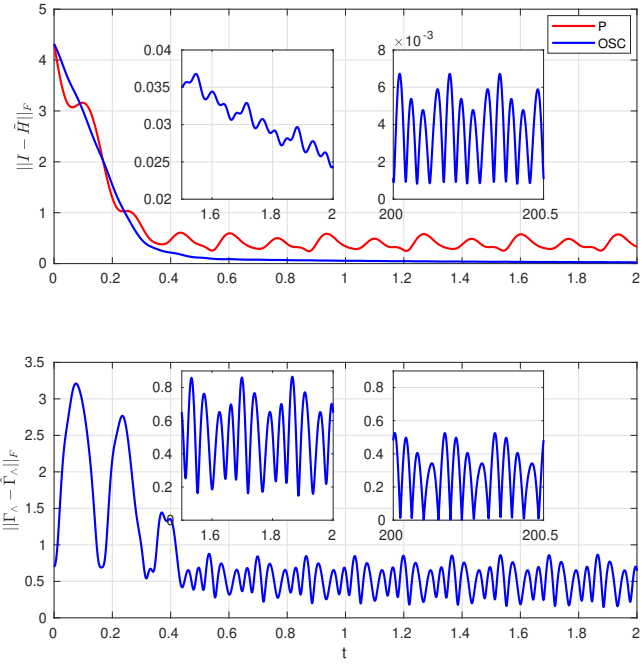


Fig. 5. Time behaviour of the Frobenious norm $\|I-\tilde{H}\|_F$ and $\|\Gamma_\wedge - \hat{\Gamma}_\wedge\|_F$. In red the proportional observer (P) proposed in [19] and in blue the observer (OSC) proposed in this paper.

In Figure 7 it is shown the state of the internal model unit relative to the basis B_8 in order to track the desired group velocity $\Gamma_\wedge$.

V. Experimental Result

A. Experimental implementation aspects

The proposed non-linear observer has been implemented in C++ with OpenCV library. Due to real-time constraint, the time consuming image processing part of the observer has been implemented taking advantages of modern GPU parallel computing. In particular, in view of embedding the filter on small-scale unmanned vehicles on the Jetson NVIDIA family, CUDA API have been used. The algorithm, whose flowchart is depicted in Figure 8, can be summarized as follows. The process is initialized detecting feature and extracting the descriptor of a reference image (red boxes Figure 8). As soon as a new image arrives, it is transformed with a perspective transformation (OpenCV's *cuda::warpPerspective*) based on the homography estimate. The algorithm detects key-points on the current warped image and extracts the features' descriptor. The

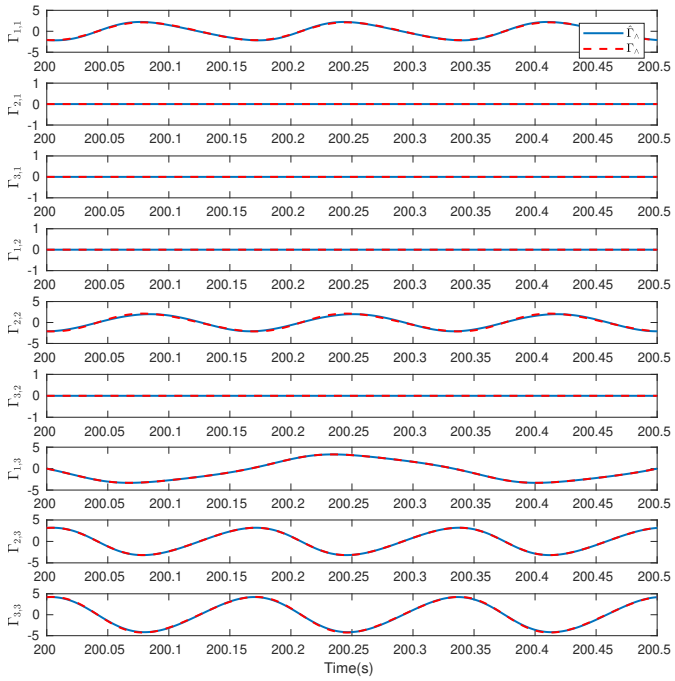


Fig. 6. Estimated group velocity $\hat{\Gamma}_\wedge$ (blue lines) and true group velocity $\Gamma_\wedge$ (red lines). Note that each plot represent an element of the group velocity matrix.

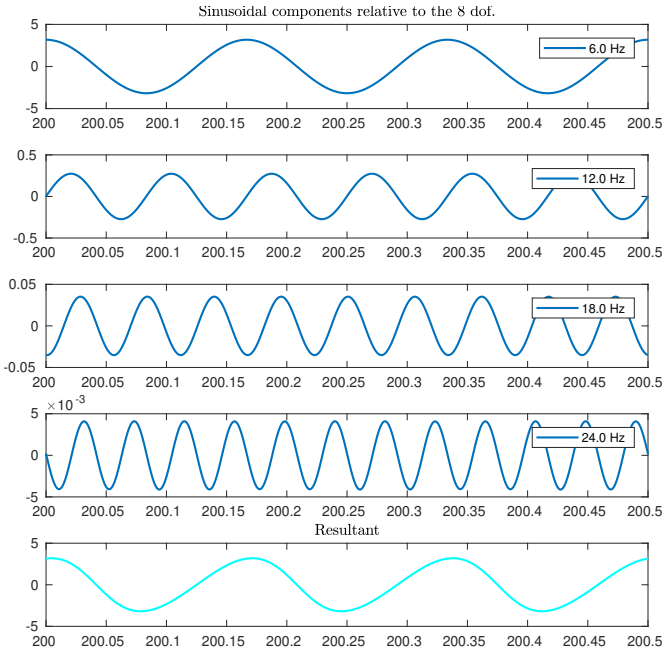


Fig. 7. Bank of oscillators relative to the eighth generator of the Lie algebra $\mathfrak{sl}(3)$. In blue the first four harmonics $(6,\ 12,\ 18,\ 24\ Hz)$ implemented in the observer and in cyan the resultant.

descriptor of the current image is matched with the descriptor of the reference image (Figure 9). Those matches, however, may contain many outliers which are removed with the outliers algorithm presented in [23]. The remaining inliers, that lie on the image plane, are projected onto the unit sphere (17). From these measurements the homography estimate is updated by iterating the observer equations 250 times per video frame. In this correction step, special attention has been paid on the integration of the observer equations in (20). In particular a leapfrog integrator has been implemented for the numerical integration of the oscillators bank in the lie algebra $\mathfrak{sl}(3)$ of the observer. Note that feature detection and description are essential components for the algorithm. Several feature detectors, such as scale invariant feature transform (SIFT)[28], features from accelerated segment test (FAST)[35], oriented FAST and rotated BRIEF (ORB)[36], speeded-up robust features (SURF)[3] and descriptors such as ORB, SIFT and SURF have been proposed in literature. Choosing the detector-descriptor combination depends on the particular application and is usually a trade-off between accuracy, repeatability, robustness and speed. Since the major constraint for this application is the time it has been decided to use the combination FAST-ORB.

B. Experimental setup

In order to experimentally validate the proposed observer we present an application for the homography estimate of an oscillating cardboard target in challenging conditions such as partial target occlusions, changing in light condition and specular reflections. The platform used for the experiment, depicted in Figure 10, is made of a cardboard box whose

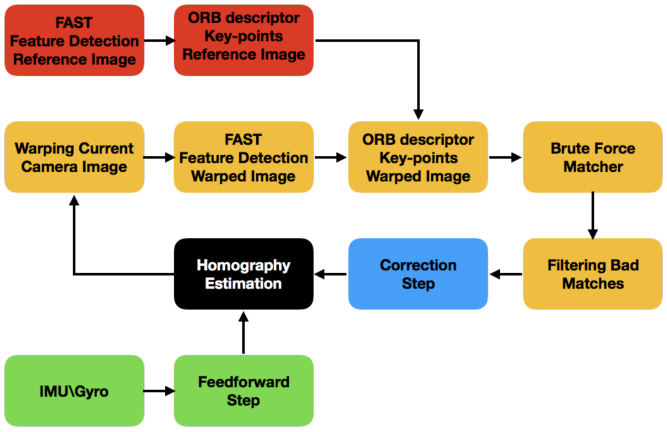


Fig. 8. Flowchart of the implemented algorithm.

upper flap is actuated by a small servo motor via an *Arduino Mega 2560*. The camera used is a *Basler acA 1200-200uc* featuring a $6.1\,\text{mm} \times 4.9\,\text{mm}$ sensor, maximum resolution $1280\,\text{px} \times 1024\,\text{px}$ and maximum frame rate of $203\,\text{fps}$. An Optitrack motion capture system is used together with markers mounted to the camera and to the cardboard in order to provide homography ground truth measurements starting from the knowledge of the full pose of the camera and the planar target.

C. Experimental results

The experiment has been performed using the *Basler* camera which is held stationary recording at $50\,\text{fps}$ while looking at the oscillating target. In an ideal scenario, the target platform should oscillate with a frequency of $0.83\,\text{Hz}$ only on the pitch

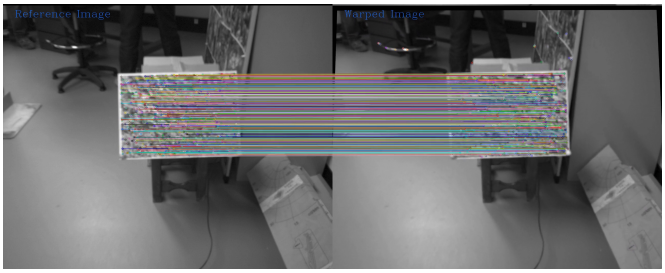


Fig. 9. Key-points detection, description and matching.

Fig. 10. Oscillating Target Platform.

angle. However, due to the simple design of the handmade platform, the target oscillates along all the three axes, as depicted in Figure 11. Since the platform is tracking the desired sinusoidal trajectory by means of position feedback only, and due to the fact that the center of mass of the target is located slightly below the rotation axis, the rising and the falling edge of the sinusoidal trajectory are not symmetrical. As a consequence, the motion of the target is still periodical but with a much richer frequency spectrum (see Figure 11, cyan box).

We proceed by presenting a comparison of the homography observer proposed in [19] with respect to the filter formulated in this paper. In order to have a fair comparison between the algorithms, the proportional term of the three observers is chosen as $k_p = 80$, the integral term of the proportional-integral filter is chosen equal to the oscillator gain of the internal model-based observer and set to $k_I = 80$. In the internal model-based observer only the nominal oscillation frequency $f_n = 0.83\,\text{Hz}$ of the target has been implemented in the bank of oscillators defined on the Lie algebra $\mathfrak{sl}(3)$. In particular denoting by

$$S_1 := \begin{bmatrix} 0 & -2\pi f_n \\ 2\pi f_n & 0 \end{bmatrix}, \quad C_1 := \begin{bmatrix} 0 & 1 \end{bmatrix},$$

the matrices of the implemented observer have the following form

$$\begin{aligned} S &= \text{diag}(S_1, S_1, S_1, S_1, S_1, S_1, S_1, S_1) \in \mathbb{R}^{16\times 16}, \\ C &= \text{diag}(C_1, C_1, C_1, C_1, C_1, C_1, C_1, C_1) \in \mathbb{R}^{8\times 16}. \end{aligned}$$

Figures 2-12 show the comparison between the three algorithms: namely proportional (P), proportional-integral (PI), and internal model-based (OSC). To this regard, we performed an experiment where:

- From 0-25 seconds (Figure 2 frames 38-293) the conditions are ideal, except for the presence of specular reflections.

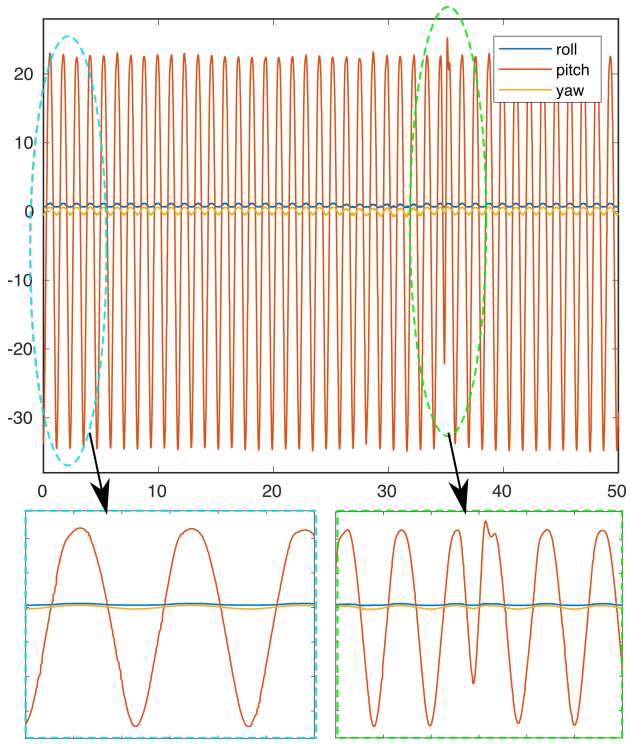


Fig. 11. Roll, pitch and yaw Euler angles of the oscillating platform.

- From 25-32 seconds (Figure 2 frame 1474-1526) the target is partially occluded by hands.
- From 34-36 seconds (Figure 2 frame 1773, Figure 11 green box) the periodical assumption is violated.
- From 32-50 seconds (Figure 2 frame 2163) there is a sudden change in light intensity conditions.

The video showing the above experiment is available in the supplementary material and from the link https://goo.gl/9Y3KsB.

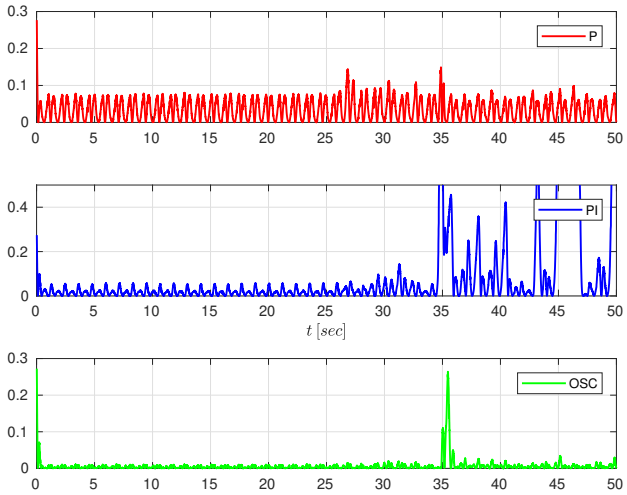


Fig. 12. Time behaviour of the Frobenious norm $\|I-\tilde{H}\|_F$ of the proportional observer (red line), proportional-integral observer (blue line) and internal model based observer (green line).

Figure 12 shows the time behavior of the Frobenious norm $\|I - \tilde{H}\|_F$ for the three implemented algorithms. The figure clearly shows better performance of the internal model based observer with respect to [19] in all the conditions except when the assumption on the periodicity of the target trajectory is strongly violated.

APPENDIX A
PROOF OF LEMMA 1

Considering (9), the time derivative of the homography is given by

$$\dot{H} = \zeta \left[\dot{\bar{R}} + \frac{\dot{\bar{\xi}}\eta^\top + \bar{\xi}\dot{\eta}^\top}{d} - \dot{d}\frac{\bar{\xi}\eta^\top}{d^2} \right] + \frac{\dot{\zeta}}{\zeta} H.$$

Recalling (7), (8) it is straightforward to verify that

$$\dot{d} = -\eta^\top \bar{V}, \quad \dot{\eta}^\top = \eta^\top \bar{\Omega}_\times$$

and bearing in mind (13a), (13b) one has

$$\begin{aligned} \dot{H} &= \zeta \left[\bar{R}\bar{\Omega}_\times + \frac{\bar{R}\bar{V}\eta^\top + \bar{\xi}\eta^\top\bar{\Omega}_\times}{d} + \frac{\eta^\top\bar{V}\bar{\xi}\eta^\top}{d^2} \right] + \frac{\dot{\zeta}}{\zeta} H \\ &= \left[\zeta \left(\bar{R} + \frac{\bar{\xi}\eta^\top}{d} \right) \bar{\Omega}_\times + \zeta \left(\bar{R} + \frac{\bar{\xi}\eta^\top}{d} \right) \frac{\bar{V}\eta^\top}{d} \right] + \frac{\dot{\zeta}}{\zeta} H \\ &= H \left(\bar{\Omega}_\times + \frac{\bar{V}\eta^\top}{d} + \frac{\dot{\zeta}}{\zeta} I \right). \end{aligned} \tag{22}$$

Since $\bar{U} \in \mathfrak{sl}(3)$ it follows

$$\operatorname{tr}\left(\bar{\Omega}_\times + \frac{\bar{V}\eta^\top}{d} + \frac{\dot{\zeta}}{\zeta} I \right) = 0$$

which in turn implies

$$\frac{\dot{\zeta}}{\zeta} = -\frac{\eta^\top \bar{V}}{3d} I.$$

APPENDIX B
PROOF OF THEOREM 1

We recall that the matrix representation of the composition of the linear maps $(\operatorname{vect} \circ \wedge)$ is denoted by the full columns rank matrix $[\![\wedge]\!] \in \mathbb{R}^{9\times 8}$. Let $[\![\operatorname{vec}^\vee]\!] \in \mathbb{R}^{8\times 9}$ denotes the matrix representation of the inverse of the map $(\operatorname{vect} \circ \wedge)$, namely

$$\begin{aligned} \operatorname{vect}(v_\wedge) &= (\operatorname{vect} \circ \wedge)(v) = [\![\wedge]\!] v, \\ \operatorname{vec}^\vee(v_\wedge) &= (\operatorname{vect} \circ \wedge)^{-1}(\operatorname{vect}(v_\wedge)) = [\![\operatorname{vec}^\vee]\!] \operatorname{vect}(v_\wedge), \end{aligned} \tag{23}$$

for any $v \in \mathbb{R}^8$. The operator $\otimes$ denotes the usual Kroneker product. Let us introduce the following properties which are instrumental for the proof of the main theorem of this paper.

A. Property 1

For any $H \in \mathrm{SL}(3)$ and $v \in \mathbb{R}^8$, the matrix representation of the composition of the linear maps $(\operatorname{vec}^\vee \circ \operatorname{Ad}_H)$

$$\operatorname{vec}^\vee(\operatorname{Ad}_H(v_\wedge)) := [\![\operatorname{Ad}_H]\!] v$$

is given by

$$[\![\operatorname{Ad}_H]\!] = [\![\operatorname{vec}^\vee]\!](H^{-\top} \otimes H)[\![\wedge]\!]. \tag{24}$$

Proof. From the definition of $[\![\wedge]\!]$ and $[\![\operatorname{vec}^\vee]\!]$ in (23) one has

$$\begin{aligned} \operatorname{vec}^\vee(Hv_\wedge H^{-1}) &= [\![\operatorname{vec}^\vee]\!] \operatorname{vect}(Hv_\wedge H^{-1}) \\ &= [\![\operatorname{vec}^\vee]\!](H^{-\top} \otimes H) \operatorname{vect}(v_\wedge) \\ &= [\![\operatorname{vec}^\vee]\!](H^{-\top} \otimes H)[\![\wedge]\!] v \\ &= [\![\operatorname{Ad}_H]\!] v \end{aligned}$$

□

B. Property 2

For any $A \in \mathbb{R}^{3\times 3}$, $B \in \mathbb{R}^{3\times 3}$, one verifies

$$\begin{aligned} &[\![\wedge]\!][\![\operatorname{vec}^\vee]\!](A \otimes B)[\![\wedge]\!] \operatorname{vec}^\vee(U) = \\ &\quad = [\![\wedge]\!][\![\operatorname{vec}^\vee]\!](A \otimes B) \operatorname{vect}(U) \\ &\quad = [\![\wedge]\!][\![\operatorname{vec}^\vee]\!] \operatorname{vect}(BUA^\top) \\ &\quad = \operatorname{vect}(BUA^\top) \\ &\quad = (A \otimes B) \operatorname{vect}(U) \\ &\quad = (A \otimes B)[\![\wedge]\!] \operatorname{vec}^\vee(U) \end{aligned}$$

which implies

$$[\![\wedge]\!][\![\operatorname{vec}^\vee]\!](A \otimes B)[\![\wedge]\!] = (A \otimes B)[\![\wedge]\!]. \tag{25}$$

C. Theorem 1 Proof

By considering the Euclidean homography (14), with kinematics (15), (16) along with the observer (20), the dynamics of the error system are given by

$$\begin{aligned} \dot{\tilde{H}} &= -[\operatorname{Ad}_{\hat{H}}(C\tilde{\gamma})_\wedge]\tilde{H} + k_p \Delta \tilde{H} \\ \dot{\tilde{\gamma}} &= S\tilde{\gamma} - k_I C^\top Q_{\mathfrak{sl}3} \sum_{i=1}^{n} \operatorname{vec}^\vee(Ad_{\hat{H}^\top}\Delta). \end{aligned} \tag{26}$$

To prove that the origin of the error system $(\tilde{H}, \tilde{\gamma}) = (I, 0)$ is locally exponentially stable, it suffices to show that the origin of the linearized error system is exponentially stable. The proof is based on Theorem 1 in [27] which establishes sufficient conditions for the uniform exponential stability of the origin of a linear time-varying system having the following standard form

$$\begin{bmatrix} \dot{x} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \mathcal{A}(t) & \mathcal{B}(t)^\top \\ -\mathcal{C}(t) & 0 \end{bmatrix} \begin{bmatrix} x \\ \theta \end{bmatrix}. \tag{27}$$

In order to write the linearized system in the standard form (27) we consider the following time-varying change of coordinates

$$\tilde{\gamma}' = \bar{Q}\tilde{\gamma}, \quad \dot{\bar{Q}} = -\bar{Q}S,$$

with $\bar{Q} \in \mathrm{SO}(m)$, yielding

$$\begin{aligned} \dot{\tilde{H}} &= -[\operatorname{Ad}_{\hat{H}}(C\bar{Q}^\top\tilde{\gamma}')_\wedge]\tilde{H} + k_p \Delta \tilde{H}, \\ \dot{\tilde{\gamma}}' &= -k_I \bar{Q} C^\top Q_{\mathfrak{sl}3} \sum_{i=1}^{n} \operatorname{vec}^\vee(Ad_{\hat{H}^\top}\Delta). \end{aligned} \tag{28}$$

Let us define $x_\wedge$, with $x \in \mathbb{R}^8$, and $\theta \in \mathbb{R}^m$ the first order approximation of $\tilde{H}$, $\tilde{\gamma}'$ around the equilibrium point $(I, 0)$

$$\tilde{H} \simeq (I + x_\wedge) = (I + \sum_{k=1}^{8} x_k B_k), \quad \tilde{\gamma}' = \theta. \tag{29}$$

A first-order approximation of the output errors e_i given by (19), considering the equation above, can be written as

$$e_i \simeq \frac{\mathring{p}_i + x_\wedge}{|\mathring{p}_i + x_\wedge \mathring{p}_i|} \simeq (\mathring{p}_i + x_\wedge)\left(1 - \mathring{p}_i^\top x_\wedge \mathring{p}_i\right)$$

and by neglecting high order terms one has

$$e_i \simeq \mathring{p}_i + x_\wedge \mathring{p}_i - \mathring{p}_i \mathring{p}_i^\top x_\wedge \mathring{p}_i. \tag{30}$$

Substituting the expression of the output errors e_i (30) in the innovation term Δ one obtains

$$\begin{aligned}\Delta &= \sum_{i=1}^n \pi_{e_i} \mathring{p}_i e_i^\top = \sum_{i=1}^n (I - e_i e_i^\top)\mathring{p}_i e_i^\top \\ &\simeq \sum_{i=1}^n \big[(\mathring{p}_i\mathring{p}_i^\top + \mathring{p}_i\mathring{p}_i^\top x_\wedge^\top - \mathring{p}_i\mathring{p}_i^\top x_\wedge^\top \mathring{p}_i\mathring{p}_i^\top) + \\ &\quad -(\mathring{p}_i + x_\wedge\mathring{p}_i - \mathring{p}_i\mathring{p}_i^\top x_\wedge \mathring{p}_i)(\mathring{p}_i + x_\wedge\mathring{p}_i - \mathring{p}_i\mathring{p}_i^\top x_\wedge \mathring{p}_i)^\top \\ &\quad (\mathring{p}_i\mathring{p}_i^\top + \mathring{p}_i\mathring{p}_i^\top x_\wedge^\top - \mathring{p}_i\mathring{p}_i^\top x_\wedge^\top \mathring{p}_i\mathring{p}_i^\top)\big] \\ &\simeq -\sum_{i=1}^n \pi_{\mathring{p}_i} x_\wedge \mathring{p}_i \mathring{p}_i^\top.\end{aligned} \tag{31}$$

Thus, the first-order error dynamics for the error system are given by

$$\begin{aligned}\dot{x}_\wedge &\simeq -\left(\mathrm{Ad}_{\hat{H}}(C\bar{Q}^\top\theta)_\wedge\right)(I + x_\wedge) + \Delta(I + x_\wedge) \\ &\simeq -\mathrm{Ad}_{\hat{H}}(C\bar{Q}^\top\theta)_\wedge - k_p\sum_{i=1}^n \pi_{\mathring{p}_i} x_\wedge \mathring{p}_i\mathring{p}_i^\top\end{aligned}$$

which in turn implies

$$\dot{x} \simeq -\operatorname{vec}^\vee\left(\mathrm{Ad}_{\hat{H}}(C\bar{Q}^\top\theta)_\wedge\right) - k_p\sum_{i=1}^n \operatorname{vec}^\vee\left(\pi_{\mathring{p}_i} x_\wedge \mathring{p}_i\mathring{p}_i^\top\right).$$

Bearing in mind Property 1 in (24), it yields

$$\begin{aligned}\dot{x} &= -[\![\mathrm{Ad}_{\hat{H}}]\!]C\bar{Q}^\top\theta - k_p\sum_{i=1}^n [\![\operatorname{vec}^\vee]\!]\operatorname{vect}\left(\pi_{\mathring{p}_i} x_\wedge \mathring{p}_i\mathring{p}_i^\top\right) \\ &= -[\![\mathrm{Ad}_{\hat{H}}]\!]C\bar{Q}^\top\theta - k_p\sum_{i=1}^n [\![\operatorname{vec}^\vee]\!](\mathring{p}_i\mathring{p}_i^\top \otimes \pi_{\mathring{p}_i})\operatorname{vect}(x_\wedge) \\ &= -[\![\mathrm{Ad}_{\hat{H}}]\!]C\bar{Q}^\top\theta - k_p\sum_{i=1}^n [\![\operatorname{vec}^\vee]\!](\mathring{p}_i\mathring{p}_i^\top \otimes \pi_{\mathring{p}_i})[\![\wedge]\!]x.\end{aligned}$$

Proceeding in a similar way, the first order approximation of $\tilde{\gamma}'$ is given by

$$\begin{aligned}\dot{\theta} &\simeq -k_I\bar{Q}C^\top Q_{\mathfrak{sl}3}\operatorname{vec}^\vee(Ad_{\hat{H}^\top}\Delta) \\ &= -k_I\bar{Q}C^\top Q_{\mathfrak{sl}3}[\![\operatorname{vec}^\vee]\!]\operatorname{vect}(\hat{H}^\top\Delta\hat{H}^{-\top}) \\ &= -k_I\bar{Q}C^\top Q_{\mathfrak{sl}3}[\![\operatorname{vec}^\vee]\!](\hat{H}^{-1}\otimes\hat{H}^\top)\operatorname{vect}(\Delta) \\ &= k_I\bar{Q}C^\top Q_{\mathfrak{sl}3}\sum_{i=1}^n [\![\operatorname{vec}^\vee]\!](\hat{H}^{-1}\otimes\hat{H}^\top)(\mathring{p}_i\mathring{p}_i^\top\otimes\pi_{\mathring{p}_i})[\![\wedge]\!]x,\end{aligned}$$

and recalling the definition of $Q_{\mathfrak{sl}3}$ and bearing in mind Property 2 in (25) one obtains

$$\dot{\theta} \simeq k_I\bar{Q}C^\top[\![\wedge]\!]^\top\sum_{i=1}^n(\hat{H}^{-1}\otimes\hat{H}^\top)(\mathring{p}_i\mathring{p}_i^\top\otimes\pi_{\mathring{p}_i})[\![\wedge]\!]x.$$

It follows that the linearized system is in standard form (27) with

$$\begin{aligned}\mathcal{A}(t) &:= -k_p\sum_{i=1}^n [\![\operatorname{vec}^\vee]\!](\mathring{p}_i\mathring{p}_i^\top\otimes\pi_{\mathring{p}_i})[\![\wedge]\!], \\ \mathcal{B}(t) &:= -\bar{Q}C^\top[\![\mathrm{Ad}_{\hat{H}}]\!]^\top, \\ \mathcal{C}(t) &:= -k_I\bar{Q}C^\top[\![\wedge]\!]^\top\sum_{i=1}^n(\hat{H}^{-1}\otimes\hat{H}^\top)(\mathring{p}_i\mathring{p}_i^\top\otimes\pi_{\mathring{p}_i})[\![\wedge]\!].\end{aligned}$$

It is straightforward to verify that the matrices

$$\begin{aligned}\mathcal{P} &:= k_I[\![\wedge]\!]^\top\sum_{i=1}^n(\mathring{p}_i\mathring{p}_i^\top\otimes\pi_{\mathring{p}_i})^\top[\![\wedge]\!], \\ \mathcal{Q} &:= 2k_Ik_p[\![\wedge]\!]^\top\sum_{i=1}^n(\mathring{p}_i\mathring{p}_i^\top\otimes\pi_{\mathring{p}_i})\sum_{i=1}^n(\mathring{p}_i\mathring{p}_i^\top\otimes\pi_{\mathring{p}_i})[\![\wedge]\!]\end{aligned}$$

are symmetric and they satisfy the required relations $\mathcal{P}\mathcal{B}^\top = \mathcal{C}^\top$ and $\mathcal{Q} = -\mathcal{A}^\top\mathcal{P} - \mathcal{P}\mathcal{A}$ of Theorem 1 in [27]. It remains to prove that the matrices $\mathcal{P}$ and $\mathcal{Q}$ are positive definite, to this purpose note that the matrix $\mathcal{P}$ is associated to the quadratic form $\|e_i - \mathring{p}_i\|^2$ at the origin, indeed one has

$$\begin{aligned}\frac{1}{k_I}x^\top\mathcal{P}x &= x^\top[\![\wedge]\!]^\top\sum_{i=1}^n(\mathring{p}_i^\top\otimes\pi_{\mathring{p}_i})^\top(\mathring{p}_i^\top\otimes\pi_{\mathring{p}_i})[\![\wedge]\!]x \\ &= \operatorname{vect}(x_\wedge{}^\top)\sum_{i=1}^n(\mathring{p}_i^\top\otimes\pi_{\mathring{p}_i})^\top(\mathring{p}_i^\top\otimes\pi_{\mathring{p}_i})\operatorname{vect}(x_\wedge) \\ &= \sum_{i=1}^n\mathring{p}_i^\top x_\wedge^\top\pi_{\mathring{p}_i}x_\wedge\mathring{p}_i = \sum_{i=1}^n\|e_i - \mathring{p}_i\|^2\Big|_{\tilde{H}=I},\end{aligned}$$

which implies $x^\top\mathcal{P}x \geq 0$. We show that $\mathcal{P}$ is positive definite by contradiction. Assume that there exists a non null vector $\bar{x}$ such that

$$\bar{x} \neq 0 \in \ker P, \tag{32}$$

which in turn implies

$$\pi_{\mathring{p}_i}\bar{x}_\wedge\mathring{p}_i = 0.$$

It follows that for each $i = \{1,\dots,n\}$ one has

$$\bar{x}_\wedge\mathring{p}_i \in \ker(\pi_{\mathring{p}_i})$$

and recalling that $\pi_{\mathring{p}_i}$ is a projector it yields

$$\bar{x}_\wedge\mathring{p}_i = \lambda_i\mathring{p}_i.$$

Since the measurement set is consistent, one can consider (without loss of generality) that $(\mathring{p}_1, \mathring{p}_2, \mathring{p}_3)$ are three non collinear eigenvectors of $\bar{x}_\wedge$ associated with the eigenvalues λ_i for $i = \{1,2,3\}$. Moreover due to the consistency of the set it can be shown that there exists a constant direction $\mathring{p}_k$ from the set $\{\mathring{p}_4,\dots,\mathring{p}_n\}$ such that:

$$\mathring{p}_k = \frac{\mathring{y}_k}{|\mathring{y}_k|}, \text{ where } \mathring{y}_k = \sum_{i=1}^3\alpha_i\mathring{p}_i,\ \alpha_i \in \mathbb{R}/\{0\},\ i = \{1,2,3\}$$

which implies that

$$\bar{x}_\wedge\sum_{i=1}^3\alpha_i\mathring{p}_i = \sum_{i=1}^3\lambda_i\alpha_i\mathring{p}_i = \sum_{i=1}^3\lambda_k\alpha_i\mathring{p}_i.$$

From the equation above it is straightforward to verify that $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_k$.

Since $\bar{x}_\wedge \in \mathfrak{sl}(3)$ one has $\operatorname{tr}(\bar{x}_\wedge) = \sum_{i=1}^3\lambda_i = 0$, which along with the equation above implies $\lambda_1 = \lambda_2 = \lambda_3 = \lambda_k = 0$, which in turn implies $\bar{x} = 0$ and this contradicts $\bar{x} \neq 0$ in (32). It follows that the kernel of the matrix $\mathcal{P}$ is trivial and $\mathcal{P}$ is positive definite.

For the matrix $\mathcal{Q}$, one verifies

$$\begin{aligned}\frac{x^\top \mathcal{Q} x}{2k_I k_p} &= x^\top [\![\wedge]\!]^\top \sum_{i=1}^{n} (\mathring{p}_i \mathring{p}_i^\top \otimes \pi_{\mathring{p}_i})^\top \sum_{i=1}^{n} (\mathring{p}_i \mathring{p}_i^\top \otimes \pi_{\mathring{p}_i}) [\![\wedge]\!] x \\ &= \operatorname{vect}\left(\sum_{i=1}^{n} (\pi_{\mathring{p}_i} x_\wedge \mathring{p}_i \mathring{p}_i^\top)\right)^\top \operatorname{vect}\left(\sum_{i=1}^{n} (\pi_{\mathring{p}_i} x_\wedge \mathring{p}_i \mathring{p}_i^\top)\right) \\ &= \operatorname{tr}(\Delta^\top \Delta)\Big|_{\tilde{H}=I},\end{aligned}$$

which implies $x^\top \mathcal{Q} x \geq 0$, and using the fact that the set $\mathcal{M}_n$ is consistent one can ensure that $\mathcal{Q}$ is positive definite.

Finally, it is straightforward to verify that $\mathcal{B}$ is a full rank matrix, which implies that $\mathcal{B}^\top \mathcal{B}$ is positive definite which in turn implies that the term $\mathcal{B}$ is persistent exciting. Therefore, we conclude that all the conditions of Theorem 1 in [27] are satisfied and thus the set $\mathcal{S}$ is locally exponentially stable. This in turn completes the proof.

ACKNOWLEDGMENT

This research was supported by the French DGA-Rapid project "Alcyon", the French EQUIPEX Robotex project, the French Astrid CONGRE project ANR-18-ASTR-0006 and the Australian Research Council through the "Australian Centre of Excellence for Robotic Vision" CE140100016.

The authors also thank C. Samson (INRIA and I3S) for his constructive feedback and advice on this work.

REFERENCES

[1] D. Astolfi and L. Marconi. A high-gain nonlinear observer with limited gain power. *IEEE Transactions on Automatic Control*, 60(11):3059–3064, 2015.

[2] G. Bastin and M. R. Gevers. Stable adaptive observers for nonlinear time-varying systems. *IEEE Transactions on Automatic Control*, 33(7):650–658, 1988.

[3] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-up robust features (SURF). *Computer Vision and Image Understanding*, 110(3):346 – 359, 2008. Similarity Matching in Computer Vision and Multimedia.

[4] S. Benhimane and E. Malis. Homography-based 2D visual tracking and servoing. *Int. J. of Robotics Research*, 26(7):661–676, 2007.

[5] M. Bin, D. Astolfi, and L. Marconi. Robust internal model design by nonlinear regression via low-power high-gain observers. In *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pages 4740–4745. IEEE, 2016.

[6] C.I. Byrnes, F. Delli Priscoli, and A. Isidori. *Output regulation of Uncertain Nonlinear Systems*. Birkhäuser Boston, 1997.

[7] C.I. Byrnes, F. Delli Priscoli, A. Isidori, and W. Kang. Structurally stable output regulation of nonlinear systems. *Automatica*, 33(3):369 – 385, 1997.

[8] F. Chaumette and S. Hutchinson. Visual servo control part I: Basic approaches. *IEEE Robotics and Automation Magazine, Institute of Electrical and Electronics Engineers*, 13(4):82–90, 2006.

[9] F. Chaumette and S. Hutchinson. Visual servo control, part II: Advanced approaches. *IEEE Robotics and Automation Magazine, Institute of Electrical and Electronics Engineers*, 14(1):109–118, 2007.

[10] J. Chen, W.E. Dixon, D.M. Dawson, and M.L. McIntyre. Homography-based visual servo tracking control of a wheeled mobile robot. *IEEE Transactions on Robotics*, 22:406–415, 2003.

[11] E.J. Davison. The robust control of a servomechanism problem for linear time-invariant multivariable systems. *IEEE Trans. on Automatic Control*, 1976.

[12] S. de Marco, M.D. Hua, R. Mahony, and T. Hamel. Homography estimation of a moving planar scene from direct point correspondence. *IEEE Conference on Decision and Control (CDC)*, pages 565–570, 2018.

[13] S. de Marco, L. Marconi, R. Mahony, and T. Hamel. Output regulation for systems on matrix lie-group. *Automatica*, 87:8–16, 2018.

[14] O. Faugeras. *Three-dimensional Computer Vision: A Geometric Viewpoint*. MIT Press, Cambridge, MA, USA, 1993.

[15] B.A. Francis. The linear multivariable regulator problem. *SIAM J. Contr. Optimiz.*, 1977.

[16] B.A. Francis and W.M. Wonham. The internal model principle of control theory. *Automatica*, 12(5):457–465, 1976.

[17] J.P. Gauthier and I. Kupka. *Deterministic Observation Theory and Applications*. Cambridge University Press, 2001.

[18] T.F. Gonçalves, J.R. Azinheira, and P. Rives. Homography-based visual servoing of an aircraft for automatic approach and landing. *2010 IEEE International Conference on Robotics and Automation*, pages 9–14, 2010.

[19] T. Hamel, R. Mahony, J. Trumpf, P. Morin, and M.D. Hua. Homography estimation on the special linear group based on direct point correspondence. *Decision and Control and European Control Conference (CDC-ECC)*, pages 7902–7908, 2011.

[20] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.

[21] M.D. Hua, T. Hamel, R. Mahony, and G. Allibert. Explicit complementary observer design on special linear group SL(3) for homography estimation using conic correspondences. In *2017 IEEE 56th Annual Conference on Decision and Control (CDC)*, pages 2434–2441, 2017.

[22] M.D. Hua, J. Trumpf, T. Hamel, R. Mahony, and P. Morin. Point and line feature-based observer design on SL(3) for Homography estimation and its application to image stabilization. Research report, I3S, Université Côte d'Azur ; CNRS, University of Nice Sophia Antipolis, I3S, UMR 7271, COATI, Inria, 06900 Sophia Antipolis, France, 2017.

[23] M.D. Hua, J. Trumpf, T. Hamel, R. Mahony, and P. Morin. Feature-based Recursive Observer Design for Homography Estimation and its Application to Image Stabilization. hal-01764913, 2018.

[24] A. Isidori and C.I. Byrnes. Output regulation for nonlinear systems. *IEEE Trans. on Automatic Control*, 35(2):131–140, 1990.

[25] A. Isidori, L. Marconi, and L. Praly. Robust design of nonlinear internal models without adaptation. *Automatica*, 48(10):2409–2419, 2012.

[26] H.K. Khalil. Robust servomechanism output feedback controllers for feedback linearizable systems. *Automatica*, 30(10):1587 – 1599, 1994.

[27] A. Loria and E. Panteley. Uniform exponential stability of linear time-varying system: revisited. *Systems & Control Letters*, 47(1):13–24, 2002.

[28] D.G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60(2):91–110, 2004.

[29] Y. Ma, S. Soatto, J. Kosecka, and S.S. Sastry. *An Invitation to 3-D Vision: From Images to Geometric Models*. SpringerVerlag, 2003.

[30] E. Malis, F. Chaumette, and S. Boudet. 2-1/2D visual servoing. *IEEE Trans. Robot. Automat.*, 15(2):238–250, 1999.

[31] E. Malis, T. Hamel, R. Mahony, and P. Morin. Dynamic estimation of homography transformations on the special linear group for visual servo control. *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 1498–1503, 2009.

[32] L.H. Nguyen, M.D. Hua, G. Allibert, and T. Hamel. Inertial-aided homography-based visual servo control of autonomous underwater vehicles without linear velocity measurements. In *2017 21st International Conference on System Theory, Control and Computing (ICSTCC)*, pages 9–16, 2017.

[33] H. Plinval, P. Morin, P. Mouyon, and T. Hamel. Visual servoing for underactuated VTOL UAVs: a linear, homography-based framework. *International Journal of Robust and Nonlinear Control*, 24(16):2285–2308.

[34] F. Delli Priscoli, L. Marconi, and A. Isidori. A new approach to adaptive nonlinear regulation. *SIAM J. Contr. Optimiz.*, 2006.

[35] E. Rosten and T. Drummond. Fusing points and lines for high performance tracking. In *Proceedings of the Tenth IEEE International Conference on Computer Vision - Volume 2*, ICCV '05, pages 1508–1515, Washington, DC, USA, 2005.

[36] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski. Orb: An efficient alternative to SIFT or SURF. In *Proceedings of the 2011 International Conference on Computer Vision*, ICCV '11, pages 2564–2571, Washington, DC, USA, 2011.

[37] D. Scaramuzza and R. Siegwart. Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles. *Trans. Rob.*, 24(5):1015–1026, 2008.

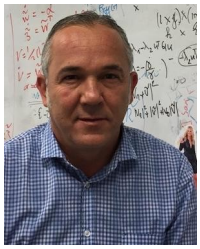
[38] A. Serrani, A. Isidori, and L. Marconi. Semi-global nonlinear output regulation with adaptive internal model. *IEEE Transactions on Automatic Control*, 46(8):1178–1194, 2001.

[39] O. Shakernia, R. Vidal, C.S. Sharp, Y. Ma, and S.S. Sastry. Multiple view motion estimation and control for landing an unmanned aerial vehicle. In *Proceedings IEEE International Conference on Robotics and Automation (ICRA)*, 2002.

Simone de Marco received the B.S. and M.S. degrees in Automation Engineering from the University of Bologna, Italy, in 2010 and 2013, respectively, and the Ph.D. degree in Automatic control and Operational research from the University of Bologna, Bologna, Italy, in 2017. He is currently a post-doctoral researcher at the laboratory I3S UNS-CNRS, France. His research interests are focused on nonlinear control, output regulation, estimation and vision-based control with applications in robotics.

Minh-Duc Hua received the engineer's degree from Ecole Polytechnique, France in 2006 and the Ph.D. degree from the University of Nice Sophia-Antipolis, France in 2009. After two postdoctoral years with I3S laboratory in France and a short research stay with Australian National University, he joined the French National Center for Scientific Research (CNRS) in 2011, where he currently holds a tenured research position. He was assigned to the Institute for Intelligent Systems and Robotics (ISIR UPMC-CNRS) in Paris from 2011 to 2016. Currently, he is with I3S UCA-CNRS Laboratory. His research interests include nonlinear control, sensor-based control, nonlinear observer, and computer vision, with applications to Unmanned Aerial Vehicles and Autonomous Underwater Vehicles.

Robert Mahony is a Professor in the Research School of Engineering at the Australian National University. He received his PhD in 1995 in systems engineering from the Australian National University after originally training in geophysics and working in marine seismology. His research interests are in non-linear systems theory and optimization with applications in robotics and computer vision.



Tarek Hamel is full Professor at the University of Nice Sophia Antipolis since 2003. He received his Ph.D. degree in Robotics from the University of Technology of Compiègne (UTC), France, in 1996. After two years as a research assistant at the UTC, he joined the Centre d'Etudes de Mécanique d'Ile de France in 1997 as an associate professor. His research interests include nonlinear control theory, estimation and vision-based control with applications to Unmanned Aerial Vehicles. Prof. T. Hamel is senior member of the Institut Universitaire de France. He served as Associate Editor for IEEE Transactions on Robotics and for Control Engineering Practice.