



HAL
open science

Homography estimation of a moving planar scene from direct point correspondence

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

► **To cite this version:**

Simone de Marco, Minh-Duc Hua, Robert Mahony, Tarek Hamel. Homography estimation of a moving planar scene from direct point correspondence. 57th IEEE Conference on Decision and Control (CDC 2018), Dec 2018, Miami Beach, FL, United States. hal-01878699

HAL Id: hal-01878699

<https://hal.science/hal-01878699>

Submitted on 21 Feb 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Homography estimation of a moving planar scene from direct point correspondence

S. de Marco, M-D. Hua, R. Mahony and T. Hamel

Abstract—Homographies provide a robust and reliable cue for visual servo control of robots. Some nonlinear observers have been recently developed for the estimation of temporal sequences of homographies associated with rigid-body motion of a camera observing a stationary planar scene. However, these algorithms do not model well time-varying changes in the homography velocity and tend to perform poorly when the camera or the scene moves fast. In this paper, an internal model-based observer posed on $SL(3)$ for homography estimation is proposed allowing for dealing with complex camera-scene trajectories such as circular and sinusoidal motions of the camera and/or the scene. Rigorous proof of local asymptotic stability is established and excellent performance of the proposed observer is justified by experiments using an IMU-Camera prototype observing an oscillating planar target.

I. INTRODUCTION

The civil and commercial usage of robotic vehicles has become very popular thanks to the availability of low-cost MEMS sensor systems, the high computational power of modern embedded micro-controllers/autopilots and also to the utility of robots to replace human beings in undertaking dull, dangerous and dirty tasks. Modern unmanned vehicles are used in a wide range of scenarios including infrastructure/architectural inspections, surveillance, crop supervision, seafloor mapping, environmental sensing, and search and rescue missions. When unmanned vehicles approach and interact actively with the environment, the precision and reliability of Global Position System (GPS) measurements are no longer sufficient for position control. Therefore, there has been a growing interest on developing alternative localization systems and due to their low cost, light weight, small size and high resolution, vision systems remain the most promising technology. Control algorithms based on visual data are known as visual servo control [3], [4], [11].

One of the most popular visual cues used in visual servo control is the relative homography between an observed scene and a reference image of the scene. This approach is effective when the scene is planar or roughly planar with respect to the camera motions. The resulting homography matrix can be decomposed to recover position and orientation of the camera [6], [13] and classical state-space servo control can be used for position regulation of the vehicle. Estimating homography from pairs of images has been extensively studied in the computer vision literature [1], [9]. For robotic

vehicles, the rigid-body motion of the camera generates a temporal correlated sequence of images and it is possible to derive nonlinear observers for homography estimation [12], [8]. These observers are expressed on the Special Linear group $SL(3)$, a Lie-group isomorphic to the group of homographies [2], and they have very good robustness and convergence properties. As opposed to the observer proposed in [12] that requires to compute individual image homographies (using algebraic approaches for instance) for the computation of the observer innovation term, the observer proposed in [8] takes image point-feature correspondences directly as input, making it simpler and more efficient for real-time applications. However, a weakness of these observers lies in the fact that they require an estimate of the homography velocity input, which is an element of the Lie algebra $\mathfrak{sl}(3)$ and encodes the rigid-body velocity of the camera with respect to the scene. This homography velocity, however, cannot be fully measured from an Inertial Measurement Unit (IMU) or other non-vision based sensors. In the absence of direct measures, these works [12], [8] use an integral correction term to estimate the unknown part of the homography velocity that is assumed to be constant. However, this restrictive assumption is only valid for some very specific trajectories of the camera. Internal Model principles [7], [10] are ideally suited to controlling a wide class of systems with unmeasured inputs. For internal model-based control, the unknown input is modelled as belonging to a family of trajectories generated as solutions of an autonomous system (referred to as exo-system). In a recent work, the authors applied these techniques to control of systems defined on the Special Euclidean group $SE(3)$ [5]. To the authors knowledge there is presently no other work using internal model principles for nonlinear observer design for homography estimation.

In this paper we derive an internal model-based observer for the homography estimation problem. We base our approach on the direct point correspondence formulation proposed in [8] and replace the simple integral velocity estimation term with an internal model construction. The exo-system considered is a collection of linear oscillators defined on the Lie algebra $\mathfrak{sl}(3)$ of the homography group $SL(3)$. This design is of significant practical interest since it covers many possible camera/target trajectories such as circular and sinusoidal motion of the camera. A key aspect of the approach is the estimation of the homography velocity. Thus, the proposed approach applies to scenarios where both camera and target are moving, for example estimation of homography between two moving rigid-bodies such as the

S. de Marco, M-D. Hua are with I3S, Université de Nice, Sophia-Antipolis, France [sdemarco\(hua\)@i3s.unice.fr](mailto:sdemarco(hua)@i3s.unice.fr).

² 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

situation encountered in landing of an aerial vehicle on a marine platform in high-seas.

The paper is organized into four sections including the present introduction. Section II introduces the notation and formulation. Section III derives the observer and provides a rigorous analysis of the local asymptotic stability of the proposed observer. Section IV provides experimental results that demonstrate the performance of the observer.

II. PROBLEM FORMULATION

A. Notation and mathematical identities

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

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

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, denoted by $SL(3)$, is the set of 3×3 real valued matrices with unit determinant

$$SL(3) := \{H \in \mathbb{R}^{3 \times 3} | \det(H) = 1\}.$$

The Lie algebra associated to the Lie-group $SL(3)$, denoted by $\mathfrak{sl}(3)$, is the set of 3×3 traceless matrices

$$\mathfrak{sl}(3) := \{U \in \mathbb{R}^{3 \times 3} | \text{tr}(U) = 0\}.$$

For $H \in SL(3)$ and $U \in \mathfrak{sl}(3)$ the mapping $Ad : SL(3) \times \mathfrak{sl}(3) \rightarrow \mathfrak{sl}(3)$ is defined by

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

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

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

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

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

where $B = \{B_1, \dots, B_8\}$ is a basis for the Lie algebra $\mathfrak{sl}(3)$. The operator $\text{vec}^\vee : \mathfrak{sl}(3) \rightarrow \mathbb{R}^8$ denotes the inverse of the $(\cdot)_\wedge$ operator, namely

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

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

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

The matrix representation of the composition of the linear maps $(\text{vect} \circ \wedge) : \mathbb{R}^8 \rightarrow \mathfrak{sl}(3) \rightarrow \mathbb{R}^9$ is denoted by the full columns rank matrix $[[\wedge]] \in \mathbb{R}^{9 \times 8}$, namely

$$\text{vect}(v_\wedge) = (\text{vect} \circ \wedge)(v) = [[\wedge]]v, \quad v \in \mathbb{R}^8$$

Let $Q_{\mathfrak{sl}(3)} \in \mathbb{R}^{8 \times 8}$ denotes the symmetric positive matrix $Q_{\mathfrak{sl}(3)} := [[\wedge]]^\top [[\wedge]]$.

For any $A, B \in \mathbb{R}^{n \times n}$, $\langle\langle A, B \rangle\rangle = \text{tr}(A^\top B)$ defines an inner product on $\mathbb{R}^{n \times n}$. We denote by $\|A\|$, with $A \in \mathbb{R}^{n \times n}$, the Frobenius norm defined by

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

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

B. Modeling the target plane and the camera frame

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 vehicle $\{B\}$ with respect to the inertial frame $\{I\}$ is given by the rotation matrix $R \in SO(3)$. Let $(\xi_T, R_T) \in \mathbb{R}^3 \times SO(3)$ denote the configuration of a frame $\{T\}$, integral with the target plane, with respect to the inertial frame $\{I\}$. The coordinate of a point in the target plane can

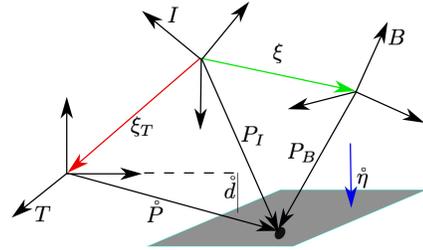


Fig. 1. Representation of a point of the planar surface in the inertial frame $\{I\}$, body-fixed frame $\{B\}$ and target frame $\{T\}$.

be expressed in the inertial frame as follow

$$P_I = R_T \dot{P} + \xi_T \quad (1a)$$

$$P_I = R P_B + \xi. \quad (1b)$$

Denoting by

$$\bar{R} := R_T^\top R \quad (2)$$

$$\bar{\xi} := -R_T^\top (\xi_T - \xi), \quad (3)$$

and substituting (1a) into (1b) one gets

$$P_B = \bar{R}^\top \dot{P} - \bar{R}^\top \bar{\xi}. \quad (4)$$

We denote by \dot{p}_i (resp. p_i) the image of the point i -th, in homogeneous coordinate, when the camera is aligned with

the frame $\{T\}$ (resp. $\{B\}$) as depicted in figure 2. Using the classical pinhole camera model one has

$$\mathring{p}_i \cong K \mathring{P}_i, \quad p_i \cong K P_{Bi}$$

where K is the upper triangular camera matrix which depends on intrinsic parameters of the camera such as the focal length and skew. If the camera is calibrated then all quantities can be appropriately scaled and the equation above can be written in the following form:

$$\mathring{p}_i \cong \mathring{P}_i, \quad p_i \cong P_{Bi}. \quad (5)$$

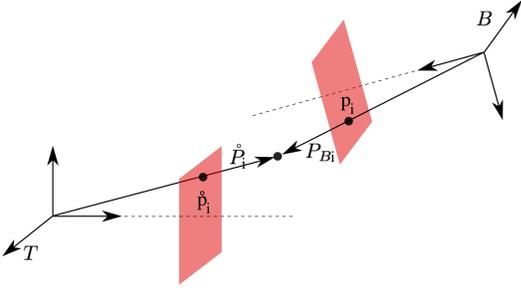


Fig. 2. Image of the point i -th when the camera is aligned with respect to the frame $\{T\}$ and with respect to the frame $\{B\}$.

Let $\mathring{\eta}$ denote the unit normal to 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 lie on a planar surface one has

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

it follows from (4) that

$$P_{Bi} = \left(\bar{R}^\top - \bar{R}^\top \bar{\xi} \frac{\mathring{\eta}^\top}{\mathring{d}} \right) \mathring{P}_i, \quad (6)$$

thus using image points (5) it yields

$$p_i \cong \left(\bar{R}^\top - \bar{R}^\top \bar{\xi} \frac{\mathring{\eta}^\top}{\mathring{d}} \right) \mathring{p}_i, \quad i = \{1, \dots, n\}. \quad (7)$$

The projective mapping $H \cong \left(\bar{R}^\top - \bar{R}^\top \bar{\xi} \frac{\mathring{\eta}^\top}{\mathring{d}} \right)^{-1}$ is called homography matrix and it relates the images points on the target plane when viewed from two different poses. The homography matrix contains the relative pose information of the camera fixed-body frame $\{B\}$ with respect to the target frame $\{T\}$. Let η_I denote the unit normal to the plane expressed in the inertial frame $\{I\}$ and let d_I denote the orthogonal distance of the planar surface to the origin of the inertial frame $\{I\}$, one has

$$\eta_I = R_T \mathring{\eta} \quad (8)$$

$$d_I = \eta_I^\top \xi_T + \mathring{d}. \quad (9)$$

Let η denote the normal to the target plane expressed in the body-fixed frame $\{B\}$, and d the orthogonal distance of the target plane to the origin of $\{B\}$, one verifies

$$\eta = R^\top \eta_I = \bar{R}^\top \mathring{\eta} \quad (10)$$

$$d = d_I - \eta_I^\top \xi = \mathring{d} - \mathring{\eta}^\top \bar{\xi}, \quad (11)$$

it follows that the homography matrix can be written as

$$H \cong \left(\bar{R} + \frac{\bar{\xi} \mathring{\eta}^\top}{\mathring{d}} \right). \quad (12)$$

Due to the fact that the homography matrix H is defined up to a scale factor we can represent uniquely the homography matrix as an element of the special linear group $SL(3)$ by re-scaling

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

From now on we assume that the homography matrix H is an element of $SL(3)$

$$H = \zeta \left(\bar{R} + \frac{\bar{\xi} \mathring{\eta}^\top}{\mathring{d}} \right) \in SL(3). \quad (13)$$

C. Homography Kinematics

Let Ω_T denote the angular velocity of the target frame expressed in $\{T\}$ and V_T the linear velocity of the frame $\{T\}$ with respect to the inertial frame $\{I\}$ expressed in $\{T\}$. Then, the kinematics of the target plane can be written as

$$\dot{R}_T = R_T \Omega_{T \times}, \quad \dot{\xi}_T = R_T V_T. \quad (14)$$

Denote the angular velocity and the linear velocity of the body-fixed frame $\{B\}$ with respect to the inertial frame $\{I\}$ expressed in the body-fixed frame $\{B\}$, respectively, by Ω and V . The rigid body kinematics are given by

$$\dot{R} = R \Omega_{\times}, \quad \dot{\xi} = R V. \quad (15)$$

The times derivatives of (2) and (3) are given by

$$\dot{\bar{R}} = \bar{R} \bar{\Omega}_{\times}, \quad \dot{\bar{\xi}} = \bar{R} \bar{V} \quad (16)$$

where

$$\bar{\Omega}_{\times} = \Omega_{\times} - (\bar{R}^\top \Omega_T)_{\times} \quad (17a)$$

$$\bar{V} = \bar{R}^\top \Omega_{T \times} \bar{\xi} + V - \bar{R}^\top V_T. \quad (17b)$$

Lemma 1: Consider a camera attached to the rigid-body with kinematics (15) viewing a moving planar scene whose kinematics is defined by (14). Let $H : \{B\} \rightarrow \{T\}$ denote the calibrated 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

$$\dot{H} = H \bar{U}, \quad \text{with } \bar{U} = \bar{\Omega}_{\times} + \frac{\bar{V} \mathring{\eta}^\top}{\mathring{d}} - \frac{\mathring{\eta}^\top \bar{V}}{3\mathring{d}} I. \quad (18)$$

Proof: The proof is based on the time derivative of the homography (13), and here omitted due to lack of space. ■

Note that the group velocity \bar{U} in (18) is defined with respect to the camera body-fixed frame. Induced from the relative motion between the camera and the target plane,

it depends usually on unmeasurable variables such as the angular velocity of the target plane Ω_T , the orthogonal distance d , the normal to the surface η and the linear velocity \bar{V} . The group velocity \bar{U} can be rewritten as

$$\bar{U} := \Omega_{\times} + \Gamma_{\wedge} \quad (19)$$

where Ω is the known part that can be obtained from the set of embedded gyros and $\Gamma_{\wedge} = -(\bar{R}^{\top} \Omega_T)_{\times} + \frac{\bar{V} \eta^{\top}}{d} - \frac{\eta^{\top} \bar{V}}{3d} I$ is typically unknown. In this work, we assume that the velocity part Γ_{\wedge} is a periodic matrix generated as solution of an autonomous linear system so that we can write Γ_{\wedge} in the form

$$\Gamma_{\wedge} = (C\gamma)_{\wedge}, \quad \dot{\gamma} = S\gamma \quad (20)$$

where $C \in \mathbb{R}^{8 \times m}$ is a full row rank matrix with $m \geq 8$, $\gamma \in \mathbb{R}^m$ and $S = -S^{\top} \in \mathbb{R}^{m \times m}$. The considered assumption on Γ_{\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 circular or sinusoidal motions.

III. NON LINEAR OBSERVER ON SL(3)

We assume to have a collection of n measurements $p_i \in \mathbb{S}^2$ associated to a group action of SL(3) onto the projective space \mathbb{P}^2 of the form

$$p_i = \frac{H^{-1} \hat{p}_i}{|H^{-1} \hat{p}_i|}, \quad i = \{1, \dots, n\}. \quad (21)$$

The goal of the homography estimator is to provide an estimate $\hat{H} \in \text{SL}(3)$ 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 $\hat{p}_i \in \mathbb{S}^2$ of p_i

$$\hat{p}_i := \frac{\hat{H}^{-1} \hat{p}_i}{|\hat{H}^{-1} \hat{p}_i|}, \quad i = \{1, \dots, n\}. \quad (22)$$

The estimates e_i of \hat{p}_i , then can be written as

$$e_i := \frac{\hat{H} p_i}{|\hat{H} p_i|} = \frac{\hat{H} \hat{p}_i}{|\hat{H} \hat{p}_i|}, \quad i = \{1, \dots, n\}. \quad (23)$$

Definition 1: A set \mathcal{M}_n of $n \geq 4$ vector directions $\hat{p}_i \in \mathbb{S}^2$ 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 calibrated homography (13) and consider the kinematic system in (18) along with (19) and (20). Consider the following nonlinear filter

$$\dot{\hat{H}} = \hat{H}(\Omega_{\times} + \hat{\Gamma}_{\wedge}) + k_p \beta \hat{H} \quad (24a)$$

$$\dot{\hat{\Gamma}}_{\wedge} = (C\hat{\gamma})_{\wedge} \quad (24b)$$

$$\dot{\hat{\gamma}} = S\hat{\gamma} + k_I C^{\top} Q_{\mathfrak{sl}(3)} \sum_{i=1}^n \varphi_i \quad (24c)$$

with k_p and k_I some positive gains, where

$$\begin{aligned} \beta &:= \sum_{i=1}^n \pi_{e_i} \hat{p}_i e_i^{\top} \in \mathfrak{sl}(3) \\ \varphi_i &:= \text{vec}^{\vee} \left(\mathbb{P}_{\mathfrak{sl}(3)} \left((\hat{H}^{\top} \hat{p}_i) (\hat{H}^{-1} e_i)^{\top} \right) \right) + \\ &\quad - e_i^{\top} \hat{p}_i \text{vec}^{\vee} \left(\mathbb{P}_{\mathfrak{sl}(3)} \left((\hat{H}^{\top} e_i) (\hat{H}^{-1} e_i)^{\top} \right) \right). \end{aligned} \quad (25)$$

Assume that the measured angular velocity Ω of the planar target, the group velocity Γ (equivalently γ) in (20) and homography matrix H are bounded.

Then, if the set \mathcal{M}_n of measured directions \hat{p}_i is consistent, the equilibrium $(\tilde{H}, \tilde{\Gamma}_{\wedge}) = (I, 0)$ is locally asymptotically stable.

Proof: By considering the calibrated homography (18), with kinematics (19), (20) along with the filter (24), the error system can be written as follows

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

Consider the following Lyapunov candidate

$$\mathcal{L}(\tilde{e}_i, \tilde{\gamma}) = \underbrace{\sum_{i=1}^n (1 - e_i^{\top} \hat{p}_i)}_{\mathcal{L}_1} + \underbrace{\frac{1}{2k_{I1}} \|\tilde{\gamma}\|^2}_{\mathcal{L}_2}$$

where $\tilde{\gamma} := \gamma - \hat{\gamma}$, which using the consistency of the set \mathcal{M}_n is a locally positive definite function of \tilde{H} and $\mathcal{L}(0, 0) = 0$, the proof here is omitted due to lack of space.

Differentiating \mathcal{L}_1 , and bearing in mind (26), it yields

$$\dot{\mathcal{L}}_1 = -\sum_{i=1}^n \hat{p}_i^{\top} \pi_{e_i} \left[k_p \beta - \text{Ad}_{\tilde{H}} \tilde{\Gamma}_{\wedge} \right] e_i$$

where $\tilde{\Gamma}_{\wedge} = \Gamma_{\wedge} - \hat{\Gamma}_{\wedge}$. Introducing the expression of β (25) in $\dot{\mathcal{L}}_1$, one obtains

$$\dot{\mathcal{L}}_1 = -k_p \left\| \sum_{i=1}^n e_i \hat{p}_i^{\top} \pi_{e_i} \right\|^2 + \text{tr} \left(\tilde{\Gamma}_{\wedge}^{\top} \sum_{i=1}^n \text{Ad}_{\tilde{H}^{\top}} \pi_{e_i} \hat{p}_i e_i^{\top} \right)$$

Let us focus on the last term of the equation above, recalling the orthogonal projection onto $\mathfrak{sl}(3)$ with respect to the inner product one has

$$\begin{aligned} \text{Ad}_{\tilde{H}^{\top}} \pi_{e_i} \hat{p}_i e_i^{\top} &= \mathbb{P}_{\mathfrak{sl}(3)} \left((\hat{H}^{\top} \hat{p}_i) (\hat{H}^{-1} e_i)^{\top} \right) + \\ &\quad - e_i^{\top} \hat{p}_i \mathbb{P}_{\mathfrak{sl}(3)} \left((\hat{H}^{\top} e_i) (\hat{H}^{-1} e_i)^{\top} \right). \end{aligned}$$

Introducing the above expression in the $\dot{\mathcal{L}}_1$ derivative, it follows:

$$\begin{aligned} \dot{\mathcal{L}}_1 &= -k_p \left\| \sum_{i=1}^n e_i \hat{p}_i^{\top} \pi_{e_i} \right\|^2 + \\ &\quad + \text{tr} \left(\tilde{\Gamma}_{\wedge}^{\top} \sum_{i=1}^n \mathbb{P}_{\mathfrak{sl}(3)} \left((\hat{H}^{\top} \hat{p}_i) (\hat{H}^{-1} e_i)^{\top} \right) \right) + \\ &\quad - \text{tr} \left(\tilde{\Gamma}_{\wedge}^{\top} \sum_{i=1}^n e_i^{\top} \hat{p}_i \mathbb{P}_{\mathfrak{sl}(3)} \left((\hat{H}^{\top} e_i) (\hat{H}^{-1} e_i)^{\top} \right) \right). \end{aligned}$$

Consider the time derivatives of the second term of the Lyapunov candidate

$$\dot{\mathcal{L}}_2 = \frac{1}{k_{I1}} \tilde{\gamma}^{\top} \dot{\tilde{\gamma}} = \frac{1}{k_{I1}} \tilde{\gamma}^{\top} (S\tilde{\gamma} - \dot{\hat{\gamma}})$$

and recalling the expression of $\dot{\hat{\gamma}}$ (24) one obtains

$$\dot{\hat{\mathcal{L}}}_2 = -\text{tr} \left(\tilde{\Gamma}^\top \left(\sum_{i=1}^n \varphi_i \right) \wedge \right).$$

Introducing the expression of the innovation term φ_i (24) one has

$$\begin{aligned} \dot{\hat{\mathcal{L}}}_2 &= -\text{tr} \left(\tilde{\Gamma}^\top \sum_{i=1}^n \mathbb{P}_{sl3} \left((\hat{H}^\top \hat{p}_i) (\hat{H}^{-1} e_i)^\top \right) \right) \\ &+ \text{tr} \left(\tilde{\Gamma}^\top \sum_{i=1}^n e_i^\top \hat{p}_i \mathbb{P}_{sl3} \left((\hat{H}^\top e_i) (\hat{H}^{-1} e_i)^\top \right) \right). \end{aligned}$$

Finally, collecting all the terms of the Lyapunov candidate we get:

$$\dot{\hat{\mathcal{L}}} = -k_p \left\| \sum_{i=1}^n e_i \hat{p}_i^\top \pi_{ei} \right\|^2.$$

The derivative of the Lyapunov function is negative semi-definite, and equal to zero when $\beta = 0$. Since the Lyapunov function is locally positive definite and (H, γ) is bounded is straightforward to verify that $(\tilde{H}, \tilde{\gamma})$ is bounded and a solution exists for all time. One then verifies that \tilde{H} is bounded by the boundedness of $(\tilde{H}, \tilde{\gamma})$ and of H , which in turn ensures the boundedness of $\dot{\hat{\mathcal{L}}}$ and, subsequently, the uniform continuity of $\dot{\hat{\mathcal{L}}}$. Then by application of Barbalat's lemma one deduces that $\lim_{t \rightarrow \infty} \dot{\hat{\mathcal{L}}} = 0$.

From the definition of the ‘‘injection’’ term β (25) and e_i (23), and exploiting the consistency of the set is possible to show, in a similar way of Theorem 3.2 in [8], that $\dot{\hat{\mathcal{L}}} = 0$ implies $\tilde{H} = I$. To prove that \tilde{H} tends to zero using Barbalat's lemma, one needs to prove that \ddot{H} is bounded, which can be verified using the boundedness of H , Ω and Γ . From this combining (24a) with (18), (19), (20) it yields

$$0 = \lim_{t \rightarrow \infty} \dot{\hat{H}} = -[\text{Ad}_{\tilde{H}}(C\tilde{\gamma})] \wedge \tilde{H}$$

which implies $\lim_{t \rightarrow \infty} \tilde{\gamma} = 0$, and this completes the proof. ■

IV. EXPERIMENTAL RESULTS

A. Experimental setup

We present an application of the proposed approach for the tracking of an oscillating platform in presence of specular reflections, partial target occlusions and severe changing in light conditions. The oscillating platform used for experimental validation, depicted in figure 3, is composed of a simple cardboard box whose upper flap is actuated by a servo motor. The servo motor is controlled by an *Arduino Mega 2560* only in position without relying on a velocity feedback loop. The experiments have been performed with a *Basler acA 1300-200uc* camera, that provides 50 frames per second with a resolution of 1280×1024 pixels. For comparison purposes, the *OptiTrack* motion capture system has been used to obtain the homography ground truth starting from the full pose of the target and the camera.

The observer proposed in this paper has been implemented in C++ with OpenCV library. In particular, the detection of images features and descriptor extraction is carried out using the *FAST Feature Detector* and *ORB Descriptor Extractor*

respectively. Then, the matching of descriptors is carried out using OpenCV's brute-force matcher routine.



Fig. 3. Oscillating Target Platform.

B. Experimental results

The experiment has been carried out keeping the camera still while looking at the oscillating platform. The oscillating platform is tracking a sinusoidal signal with a frequency of $0.83Hz$, however due to the handmade design and due to the servo motor limitations the platform trajectory exhibit a much richer frequency spectrum. The observer parameters are chosen to be $k_p = 80$ and $k_I = 80$. Regarding the bank of oscillators in the lie algebra $\mathfrak{sl}(3)$ of the filter, it has been taken into account only the nominal frequency $f_n = 0.83Hz$ of the oscillating platform, namely

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

where

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

In Figure 4 some screenshot of the experiment are depicted.

The experiment, whose full video is available at <https://goo.gl/AbRyHs>, has been conducted as follow:

- From 0-25 seconds (Figure 4 frames 35) the conditions are ideal, some reflection from the room neons appears.
- From 25-32 seconds (Figure 4 frame 1470) the target is partially occluded by hands.
- From 34-36 seconds (Figure 4 frame 1775) the periodical assumption is violated.
- From 32-50 seconds (Figure 4 frame 2517) we turned off the room lights.

Figure 5 shows the time behavior of the Frobenious norm $\|I - \tilde{H}\|_F$ for the implemented observer. Figures 5-4 show a very fast convergence of the estimated homography to the real one, robustness with respect to lighting conditions and target occlusion and quick recover after the violation of the periodical assumption.

V. CONCLUSIONS

A nonlinear observer designed on the Special Linear group $SL(3)$ for the estimation of a sequence of homographies has been proposed. It extends our prior works on this topic by replacing the simple integral velocity estimation term with

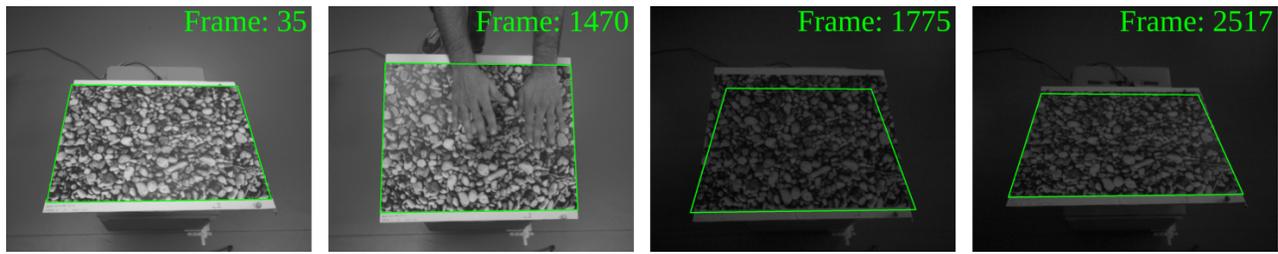


Fig. 4. The green polygon represents the region of interest tracked by the filter.

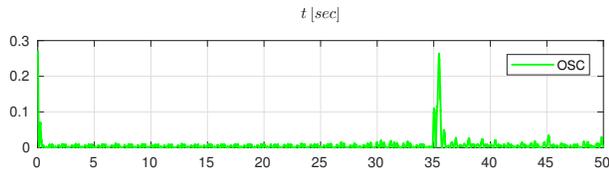


Fig. 5. Time behaviour of the Frobenius norm $\|I - \tilde{H}\|_F$ of the proposed observer.

an internal model construction. Experimental results have been provided as a complement to the theoretical approach showing the robustness of the observer proposed in this paper. The proposed approach is well suited for scenarios in which both the camera and target are moving simultaneously.

ACKNOWLEDGMENTS

This research was supported by the French DGA-Rapid project “Alcyon”, the French EQUIPEX Robotex project, 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] A. Agarwal, C. V. Jawahar, and P. J. Narayanan. A survey of planar homography estimation techniques. *IEEE Third Int. Symp. on 3D Data Processing, Visualization and Transmission*, pages 908–915, 2006.
- [2] S. Benhimane and E. Malis. Homography-based 2d visual tracking and servoing. *Int. J. of Robotics Research*, 26(7):661–676, 2007.
- [3] 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.
- [4] 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.
- [5] S. de Marco, L. Marconi, R. Mahony, and T. Hamel. Output regulation for systems on matrix lie-group. *Automatica*, 87:8–16, 2018.
- [6] O. Faugeras and F. Lustman. Motion and structure from motion in a piecewise planar environment. *International Journal of Pattern Recognition and Artificial Intelligence*, 2(3):485–508, 1988.
- [7] B.A. Francis and W.M. Wonham. The internal model principle of control theory. *Automatica*, 12(5):457–465, 1976.
- [8] 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.
- [9] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, 2003.
- [10] A. Isidori and C.I. Byrnes. Output regulation for nonlinear systems. *IEEE Trans. on Automatic Control*, 35(2):131–140, 1990.

- [11] E. Malis, F. Chaumette, and S. Boudet. 2-1/2d visual servoing. *IEEE Trans. Robot. Automat.*, 15(2):238–250, 1999.
- [12] 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.
- [13] E. Malis and M. Vergas. Deeper understanding of the homography decomposition for vision-based control. Technical report, INRIA, 2007.