2011 50th IEEE Conference on Decision and Control and
European Control Conference (CDC-ECC)
Orlando, FL, USA, December 12-15, 2011

Homography estimation on the Special Linear Group based on direct point correspondence

Tarek Hamel, Robert Mahony, Jochen Trumpf, Pascal Morin and Minh-Duc Hua

Abstract— **This paper considers the question of obtaining a high quality estimate of a time-varying sequence of image homographies using point correspondences from an image sequence without requiring explicit computation of the individual homographies between any two given images. The approach uses the representation of a homography as an element of the Special Linear group and defines a nonlinear observer directly on this structure. We assume, either that the group velocity of the homography sequence is known, or more realistically, that the homographies are generated by rigid-body motion of a camera viewing a planar surface, and that the angular velocity of the camera is known.**

I. INTRODUCTION

A homography is an invertible mapping relating two images of the same planar scene. Homographies play a key role in many computer vision and robotics problems, especially those that involve manmade environments typically constructed of planar surfaces, and those where the camera is sufficiently far from the scene viewed that the relief of surface features is negligible, such as the situation encountered in vision sequences of the ground taken from a flying vehicle. Computing homographies from point correspondences has been extensively studied in the last fifteen years and different techniques have been proposed in the literature that provide an estimate of the homography matrix [5]. The quality of the homography estimate depends strongly on the algorithm used and the size of the set of points considered. For a well textured scene, state of the art methods provide high quality homography estimates but require significant computational effort (see [17] and references therein). For a scene with poor texture, the existing homography estimation algorithms perform poorly. In a recent paper by the authors [13], [14] a nonlinear observer for homography estimation was proposed based on the group structure of the set of homographies, the Special Linear group $SL(3)$ [2]. This observer uses velocity information to interpolate across a sequence of images and to improve the overall homography estimate between any two given images. Although this earlier approach addresses the problem partly by using temporal information to improve instantaneous estimates, the observer still requires individual image homographies to be computed for each image in the sequence to compute the observer innovation.

In this paper, we consider the question of deriving an observer for a sequence of image homographies that directly takes point feature correspondences as input. The proposed approach is only valid in the case where a sequence of images used as data is associated with a continuous variation of the reference image. The most common case encountered is where the images are derived from a moving camera viewing a planar scene. The nonlinear proposed observer is posed on the Special Linear group $SL(3)$ that is in one-to-one correspondence with the group of homographies [2] and uses velocity measurements to propagate the homography estimate and fuse this with new data as it becomes available [13], [14]. A key advance on prior work by the authors is the formulation of a point feature innovation for the observer that incorporates point correspondences directly in the observer without requiring reconstruction of individual image homographies. The proposed approach has a number of advantages. Firstly, it avoids the computation associated with the full homography construction. This saves considerable computational resources and makes the proposed algorithm suitable for embedded systems with simple point tracking software. Secondly, the algorithm is well posed even when there is insufficient data for a full reconstruction of a homography. For example, if the number of corresponding points between two images drops below four it is impossible to algebraically reconstruct an image homography and the existing algorithms fail. In such situations, the proposed observer will continue to operate, incorporating what information is available and relying on propagation of prior estimates where necessary. Finally, even if a homography can be reconstructed from a small set of feature correspondences, the estimate is often unreliable and the associated error is difficult to characterize. The proposed algorithm integrates information from a sequence of images, and noise in the individual feature correspondences is filtered through the natural low-pass response of the observer, resulting in a highly robust estimate. As a result, the authors believe that the proposed observer is ideally suited for poorly textured scenes and real-time implementation. We initially consider the case where the group velocity is known, a situation that is rarely true in practice but provides considerable insight. The main result of

T. Hamel and M.D. Hua are with I3S UNSA-CNRS, Nice-Sophia Antipolis, France. e-mails: thamel(hua)@i3s.unice.fr.

R. Mahony and J. Trumpf are with the School of Engineering, Australian National University, ACT, 0200, Australia. e-mails: Robert.Mahony(Jochen.Trumpf)@anu.edu.au.

P. Morin is with ISIR UPMC, Paris, France. email: morin@isir.upmc.fr.

the paper considers the case of a moving camera where the angular velocity of the camera is measured. This is a practical scenario where a camera is equipped with gyrometers. The primary focus of the paper is on the presentation of the observers and analysis of their stability properties, however, we do provide simulations to indicate the performance of the proposed scheme.

The paper is organized into five sections including the introduction and the conclusion sections. Section II presents a brief recap of the Lie group structure of the set of homographies and relates it to rigid-body motion of the camera. Section III provides an initial lemma in the case where it is assumed the group velocity is known and then considers the case of a moving camera where the angular velocity of the camera is known. Simulation results are provided in Section IV to verify performance of the proposed algorithms.

II. Preliminary material

A. Projection

Visual data is obtained via a projection of observed images onto the camera image surface. The projection is parameterised by two sets of parameters: intrinsic ("internal" parameters of the camera such as the focal length, the pixel aspect ratio, etc.) and extrinsic (the pose, i.e. the position and orientation of the camera). Let $\mathring{\mathcal{A}}$ (resp. $\mathcal{A}$) denote projective coordinates for the image plane of a camera $\mathring{A}$ (resp. A), and let $\{\mathring{A}\}$ (resp. $\{A\}$) denote its (right-hand) frame of reference. Let $\xi \in \mathbb{R}^3$ denote the position of the frame $\{A\}$ with respect to $\{\mathring{A}\}$ expressed in $\{\mathring{A}\}$. The orientation of the frame $\{A\}$ with respect to $\{\mathring{A}\}$, is given by a rotation matrix, element of the Special Orthogonal group, $R \in SO(3) : \{A\} \to \{\mathring{A}\}$. The pose of the camera determines a rigid body transformation from $\{A\}$ to $\{\mathring{A}\}$ (and visa versa). One has

$$\mathring{P} = RP + \xi \tag{1}$$

as a relation between the coordinates of the same point in the reference frame ($\mathring{P} \in \{\mathring{A}\}$) and in the current frame ($P \in \{A\}$). The camera internal parameters, in the commonly used approximation, define a 3×3 matrix K so that we can write[1]:

$$\mathring{p} \cong K\mathring{P}, \quad p \cong KP, \tag{2}$$

where $p \in \mathcal{A}$ is the image of a point when the camera is aligned with frame $\{A\}$, and can be written as $(x, y, w)^T$ using the homogeneous coordinate representation for that 2D image point. Likewise, $\mathring{p} \in \mathring{\mathcal{A}}$ is the image of the same point viewed when the camera is aligned with frame $\{\mathring{A}\}$.

If the camera is calibrated (the intrinsic parameters are known), then all quantities can be appropriately scaled and the equation written in a simple form:

$$\mathring{p} \cong \mathring{P}, \quad p \cong P. \tag{3}$$

B. Homography

Assumption 2.1: Assume a calibrated camera and that there is a planar surface π containing a set of n target points ($n \geq 4$) so that

$$\pi = \left\{ \mathring{P} \in R^3 : \mathring{\eta}^T \mathring{P} - \mathring{d} = 0 \right\},$$

where $\mathring{\eta}$ is the unit normal to the plane expressed in $\{\mathring{A}\}$ and $\mathring{d}$ is the distance of the plane to the origin of $\{\mathring{A}\}$.
From the rigid-body relationships (1), one has $P = R^T\mathring{P} - R^T\xi$. Define $\zeta = -R^T\xi$. Since all target points lie in a single planar surface one has

$$P_i = R^T\mathring{P}_i + \frac{\zeta\mathring{\eta}^T}{\mathring{d}}\mathring{P}_i, \quad i = \{1, \ldots, n\}, \tag{4}$$

and thus, using (3), the projected points obey

$$p_i \cong \left(R^T + \frac{\zeta\mathring{\eta}^T}{\mathring{d}} \right) \mathring{p}_i, \quad i = \{1, \ldots, n\}. \tag{5}$$

The projective mapping $H : \mathcal{A} \to \mathring{\mathcal{A}}$, $H :\cong \left(R^T + \frac{\zeta\mathring{\eta}^T}{\mathring{d}} \right)^{-1}$ is termed a homography and it relates the images of points on the plane π when viewed from two poses defined by the coordinate systems $\mathcal{A}$ and $\mathring{\mathcal{A}}$. It is straightforward to verify that the homography H can be written as follows:

$$H \cong \left(R + \frac{\xi\eta^\top}{d} \right) \tag{6}$$

where η is the normal to the observed planar surface expressed in the frame $\{A\}$ and d is the orthogonal distance of the plane to the origin of $\{A\}$. One can verify that [2]:

$$\eta = R^T\mathring{\eta} \tag{7}$$
$$d = \mathring{d} - \mathring{\eta}^T\xi = \mathring{d} + \eta^T\zeta. \tag{8}$$

The homography matrix contains the pose information (R, ξ) of the camera from the frame $\{A\}$ (termed current frame) to the frame $\{\mathring{A}\}$ (termed reference frame). However, since the relationship between the image points and the homography is a projective relationship, it is only possible to determine H up to a scale factor (using the image points relationships alone).

C. Homography versus element of the Special Linear Goup $SL(3)$

Recall that the Special Linear Lie-group $SL(3)$ is defined as the set of all real valued 3×3 matrices with unit determinant

$$SL(3) = \{S \in \mathbb{R}^3 \mid \det S = 1\}.$$

[1] Most statements in projective geometry involve equality up to a multiplicative constant denoted by $\cong$.

Since a homography matrix H is only defined up to scale then any homography matrix is associated with a unique matrix $\bar{H} \in SL(3)$ by re-scaling

$$\bar{H} = \frac{1}{\det(H)^{\frac{1}{3}}} H \tag{9}$$

such that $\det(\bar{H}) = 1$. Moreover, the map

$$w\colon SL(3) \times \mathbb{P}^2 \longrightarrow \mathbb{P}^2,$$
$$(H, p) \mapsto w(H, p) \cong \frac{Hp}{|Hp|}$$

is a group action of $SL(3)$ on the projective space $\mathbb{P}^2$ since

$$w(H_1, w(H_2, p)) = w(H_1 H_2, p), \; w(I, p) = p$$

where H_1, H_2 and $H_1 H_2 \in SL(3)$ and I is the identity matrix, the unit element of $SL(3)$. The geometrical meaning of the above property is that the 3D motion of the camera between views $\mathcal{A}_0$ and $\mathcal{A}_1$, followed by the 3D motion between views $\mathcal{A}_1$ and $\mathcal{A}_2$ is the same as the 3D motion between views $\mathcal{A}_0$ and $\mathcal{A}_2$. As a consequence, we can think of homographies as described by elements of $SL(3)$.

The Lie-algebra $\mathfrak{sl}(3)$ for $SL(3)$ is the set of matrices with trace equal to zero: $\mathfrak{sl}(3) = \{X \in \mathbb{R}^{3\times 3} \mid \mathrm{tr}(X) = 0\}$. The adjoint operator is a mapping $\mathrm{Ad} : SL(3) \times \mathfrak{sl}(3) \to \mathfrak{sl}(3)$ defined by

$$\mathrm{Ad}_H X = HXH^{-1}, \quad H \in SL(3), X \in \mathfrak{sl}(3).$$

For any two matrices $A, B \in \mathbb{R}^{3\times 3}$ the Euclidean matrix inner product and Frobenius norm are defined as

$$\langle\langle A, B \rangle\rangle = \mathrm{tr}(A^\top B)\,, \qquad ||A|| = \sqrt{\langle\langle A, A \rangle\rangle}$$

Let $\mathbb{P}$ denote the unique orthogonal projection of $\mathbb{R}^{3\times 3}$ onto $\mathfrak{sl}(3)$ with respect to the inner product $\langle\langle \cdot, \cdot \rangle\rangle$. It is easily verified that

$$\mathbb{P}(H) := \left(H - \frac{\mathrm{tr}(H)}{3} I \right) \in \mathfrak{sl}(3). \tag{10}$$

For any matrices $G \in SL(3)$ and $B \in \mathfrak{sl}(3)$ then $\langle\langle B, G \rangle\rangle = \langle\langle B, \mathbb{P}(G) \rangle\rangle$ and hence

$$\mathrm{tr}(B^\top G) = \mathrm{tr}(B^\top \mathbb{P}(G)). \tag{11}$$

Since any homography is defined up to a scale factor, we assume from now that $H \in SL(3)$:

$$H = \gamma \left(R + \frac{\xi \eta^\top}{d} \right) \tag{12}$$

There are numerous approaches for determining H, up to this scale factor, cf. for example [6]. Note that direct computation of the determinant of H in combination with the expression of d (8) and using the fact that $\det(H) = 1$, shows that $\gamma = (\frac{d}{\mathring{d}})^{\frac{1}{3}}$.

Extracting R and $\frac{\xi}{d}$ from H is in general quite complex [2], [20], [19], [4] and is beyond the scope of this paper.

D. Homography kinematics from a camera moving with rigid-body motion

Assume that a sequence of homographies is generated by a moving camera viewing a stationary planar surface. Thus, any group velocity (infinitesimal variation of the homography) must be associated with an instantaneous variation in measurement of the *current* image $\mathcal{A}$ and not with a variation in the *reference* image $\mathring{\mathcal{A}}$. This imposes constraints on two degrees of freedom in the homography velocity, namely those associated with variation of the normal to the reference image, and leaves the remaining six degrees of freedom in the homography group velocity depending on the rigid-body velocities of the camera.

Denote the rigid-body angular velocity and linear velocity of $\{A\}$ with respect to $\{\mathring{A}\}$ expressed in $\{A\}$ by Ω and V, respectively. The rigid body kinematics of (R, ξ) are given by

$$\dot{R} = R\Omega_\times \tag{13}$$
$$\dot{\xi} = RV \tag{14}$$

where $\Omega_\times$ is the skew symmetric matrix associated with the vector cross-product, i.e. $\Omega_\times y = \Omega \times y$, for all y.

Recalling (8), it is easily verified that

$$\dot{d} = -\eta^\top V, \qquad \frac{d}{dt}\mathring{d} = 0$$

This constraint on the variation of η and $\mathring{d}$ is precisely the velocity constraint associated with the fact that the reference image is stationary.

Lemma 2.2: Consider a camera attached to the moving frame $\mathcal{A}$ moving with kinematics (13) and (14) viewing a stationary planar scene. Let $H : \mathcal{A} \to \mathring{\mathcal{A}}$ denote the calibrated homography (12). The group velocity $U \in \mathfrak{sl}(3)$ induced by the rigid-body motion and such that

$$\dot{H} = HU, \tag{15}$$

is given by

$$U = \left(\Omega_\times + \frac{V\eta^\top}{d} - \frac{\eta^\top V}{3d} I \right) \tag{16}$$

Proof: Consider the time derivative of (12). One has

$$\dot{H} = \gamma \left(\dot{R} + \frac{\dot{\xi}\eta^\top + \xi\dot{\eta}^\top}{d} - \frac{\dot{d}\xi\eta^\top}{d^2} \right) + \frac{\dot{\gamma}}{\gamma} H \tag{17}$$

Recalling Equations (13) and (14) one has

$$\begin{aligned}
\dot{H} &= \gamma \left(R\Omega_\times + \frac{RV\eta^\top + \xi\eta^\top\Omega_\times}{d} + \frac{\eta^\top V \xi \eta^\top}{d^2} \right) + \frac{\dot{\gamma}}{\gamma} H \\
&= \gamma \left(\left[R + \frac{\xi\eta^\top}{d} \right] \Omega_\times + \left[R + \frac{\xi\eta^\top}{d} \right] \frac{V\eta^\top}{d} \right) + \frac{\dot{\gamma}}{\gamma} H \\
&= H \left(\Omega_\times + \frac{V\eta^\top}{d} + \frac{\dot{\gamma}}{\gamma} I \right)
\end{aligned}$$

Applying the constraint that $\mathrm{tr}(U) = 0$ for any element of $\mathfrak{sl}(3)$, one obtains

$$0 = \mathrm{tr}\left(\Omega_\times + \frac{V\eta^\top}{d} + \frac{\dot\gamma}{\gamma}I\right) = \frac{\eta^\top V}{d} + \frac{3\dot\gamma}{\gamma}.$$

The result follows by substitution. ■

Note that the group velocity U induced by camera motion depends on the additional variables η and d that define the scene geometry at time t as well as the scale factor γ. Since these variables are unmeasurable and cannot be extracted directly from the measurements, in the sequel, we rewrite:

$$U := (\Omega_\times + \Gamma), \quad \text{with } \Gamma = \left(\frac{V\eta^\top}{d} - \frac{\eta^\top V}{3d}I\right). \quad (18)$$

Since $\{\mathring{A}\}$ is stationary by assumption, the vector Ω can be directly obtained from the set of embedded gyros. The term Γ is related to the translational motion expressed in the current frame $\{A\}$. If we assume that $\frac{\dot\xi}{d}$ is constant (*e.g.* the situation in which the camera moves with a constant velocity parallel to the scene or converges exponentially toward it), and using the fact that $V = R^T\dot\xi$, it is straightforward to verify that

$$\dot\Gamma = [\Gamma, \Omega_\times] \quad (19)$$

where $[\Gamma, \Omega_\times] = \Gamma\Omega_\times - \Omega_\times\Gamma$ is the Lie bracket.

However, if we assume that $\frac{V}{d}$ is constant (the situation in which the camera follows a circular trajectory over the scene or performs an exponential convergence towards it), it follows that

$$\dot\Gamma_1 = \Gamma_1\Omega_\times, \text{ with } \Gamma_1 = \frac{V}{d}\eta^T. \quad (20)$$

III. Nonlinear observer on $SL(3)$ based on direct measurements

The system considered is the kinematics of an element of $SL(3)$ given by (15) along with (18).

A general framework for non-linear filtering on the Special Linear group is introduced for the case where the group velocity is known. The theory is then extended to the situation for which a part of the group velocity is not available. In particular, we extend the result to the case where the camera is attached to a current frame $\{A\}$ while the reference frame $\{\mathring{A}\}$ is assumed to be a Galilean frame such that Ω represents the measurements of embarked gyrometers. In this case, it is also assumed that the matrix Γ is slowly time varying in the inertial frame and its time derivative obeys either (19) or (20).

A. Observer with known group velocity

The goal of the estimation of $H(t) \in SL(3)$, is to provide an estimate $\hat H \in SL(3)$ and to drive the error term $\tilde H = \hat H H^{-1}$ to the identity matrix I while assuming that $U \in \mathfrak{sl}(3)$ is known.

The estimator equation is posed directly as a kinematic filter system on $SL(3)$ with state $\hat H$, based on a collection of n measurements $p_i \in S^2$.

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

The observer includes a correction term derived from the measurements and depends implicitly on the error $\tilde H$. The general form of the estimator filter is

$$\dot{\hat H} = \hat H U + k_P \omega \hat H \quad (22)$$

The innovation or correction term $\omega \in \mathfrak{sl}(3)$ is thought of as an error function of the measurements p_i and their estimates $\hat p_i$ (or as an error function of the measured points $\mathring{p}_i$ and their estimates e_i). It depends implicitly on $\tilde H$. The estimates $\hat p_i$ of p_i are defined as follows,

$$\hat p_i = \frac{\hat H^{-1}\mathring{p}_i}{|\hat H^{-1}\mathring{p}_i|} \quad (23)$$

Equivalently, the estimates e_i of $\mathring{p}_i$ are defined as follows,

$$e_i = \frac{\hat H p_i}{|\hat H p_i|} = \frac{\tilde H \mathring{p}_i}{|\tilde H \mathring{p}_i|}, \quad \tilde H = \hat H H^{-1} \quad (24)$$

Definition 3.1: A set $\mathcal{M}_n$ of $n \geq 4$ vector directions $\mathring{p}_i \in S^2$ $(i = \{1\ldots n\})$ 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.

This definition implies that if the set $\mathcal{M}_n$ is consistent then, for all $\mathring{p}_i \in \mathcal{M}_4$ there exist a unique set of three non vanishing scalars $b_j \neq 0$ $(j \neq i)$ such that

$$\mathring{p}_i = \frac{y_i}{|y_i|} \text{ where } y_i = \sum_{j=1(j\neq i)}^{4} b_j \mathring{p}_j$$

Theorem 3.2: Let $H : \mathcal{A} \to \mathring{\mathcal{A}}$ denote the calibrated homography (12) and consider the kinematics (15) along with (18). Assume that $U \in \{\mathring{A}\}$ is known. Consider the nonlinear estimator filter defined by (22) along with the innovation $\omega \in \mathfrak{sl}(3)$ given by

$$\omega = \sum_{i=1}^{n} \pi_{e_i}\mathring{p}_i e_i^T \quad (25)$$

where $\pi_x = (I - xx^T)$, for all $x \in S^2$. Then, if the set $\mathcal{M}_n$ of the measured directions $\mathring{p}_i$ is consistent, the equilibrium $\tilde H = I$ is asymptotically stable.

Proof: Based on Eqn's (15) and (22), it is straightforward to show that the derivatives of (21) and (23) fulfill

$$\dot p_i = -\pi_{p_i} U p_i \text{ and } \dot{\hat p}_i = -\pi_{\hat p_i}(U + k_P \mathrm{Ad}_{\hat H^{-1}}\omega)\hat p_i$$

Differentiating $\tilde H$, it yields

$$\dot{\tilde H} = \hat H(U + k_P \mathrm{Ad}_{\hat H^{-1}}\omega)H^{-1} - \hat H U H^{-1} = k_P \omega \tilde H$$

Differentiating e_i defined by (24), we get

$$\dot{e}_i = k_P \pi_{e_i} \omega e_i$$

Define the following candidate Lyapunov function.

$$\mathcal{L}_0 = \sum_{i=1}^{n} \frac{1}{2} |e_i - \mathring{p}_i|^2 \tag{26}$$

Using the consistency of the set $\mathcal{M}_n$, one can ensure that $\mathcal{L}_0$ is locally a definite positive function of $\tilde{H}$. Differentiating $\mathcal{L}_0$, it yields:

$$\dot{\mathcal{L}}_0 = \sum_{i=1}^{n} (e_i - \mathring{p}_i)^T \dot{e}_i$$

Introducing the above expression of $\dot{e}_i$, it follows:

$$\dot{\mathcal{L}}_0 = \sum_{i=1}^{n} k_P (e_i - \mathring{p}_i)^T \pi_{e_i} \omega e_i = -k_P \mathrm{tr}\left(\sum_{i=1}^{n} e_i \mathring{p}_i^T \pi_{e_i} \omega \right)$$

Introducing the expression of ω (25), we get:

$$\dot{\mathcal{L}}_0 = -k_P \left\| \sum_{i=1}^{n} e_i \mathring{p}_i^T \pi_{e_i} \right\|^2$$

The derivative of the Lyapunov function is negative and equal to zero when $\omega = 0$, and therefore one can ensure that $\tilde{H}$ is locally bounded. From the definitions of ω (25) and e_i (24), one deduces that

$$\omega \tilde{H}^{-\top} = \sum_{i=1}^{n} \left(I_3 - \frac{\tilde{H}\mathring{p}_i \mathring{p}_i^\top \tilde{H}^\top}{|\tilde{H}\mathring{p}_i|^2} \right) \frac{\mathring{p}_i \mathring{p}_i^\top}{|\tilde{H}\mathring{p}_i|}$$

Computing the trace of $\omega \tilde{H}^{-\top}$, it follows:

$$\mathrm{tr}(\omega \tilde{H}^{-\top}) = \sum_{i=1}^{n} \frac{1}{|\tilde{H}\mathring{p}_i|^3} \left(|\tilde{H}\mathring{p}_i|^2 |\mathring{p}_i|^2 - ((\tilde{H}\mathring{p}_i)^\top \mathring{p}_i)^2 \right)$$

Define $X_i = \tilde{H}\mathring{p}_i$ and $Y_i = \mathring{p}_i$, it is straightforward to verify that

$$\mathrm{tr}(\omega \tilde{H}^{-\top}) = \sum_{i=1}^{n} \frac{1}{|X_i|^3} \left(|X_i|^2 |Y_i|^2 - (X_i^T Y_i)^2 \right) \geq 0$$

Using the fact that $\omega = 0$ at the equilibrium and therefore $\mathrm{tr}(\omega \tilde{H}^{-\top}) = 0$, as well as the Cauchy-Schwarz inequality, it follows that $X_i^T Y_i = \pm |X_i||Y_i|$ and consequently one has:

$$(\tilde{H}\mathring{p}_i)^\top \mathring{p}_i = \pm |\tilde{H}\mathring{p}_i||\mathring{p}_i|, \ \forall i = \{1 \dots n\},$$

which in turn implies the existence of some non-null constants $\lambda_i = \pm |\tilde{H}\mathring{p}_i|$ such that

$$\tilde{H}\mathring{p}_i = \lambda_i \mathring{p}_i \tag{27}$$

Note that the fact that $\lambda_i \neq 0$ can be easily verified. For instance, if $\lambda_i = 0$, then $\mathring{p}_i = \lambda_i \tilde{H}^{-1} \mathring{p}_i = 0$ which contradicts the fact that $\mathring{p}_i \in S^2$. Relation (27) indicates that all λ_i are eigenvalues of $\tilde{H}$ and all $\mathring{p}_i \in S^2$ are the associated eigenvectors of $\tilde{H}$.

Since $\mathcal{M}_n$ is a consistent set, it follows at the limit (and without loss of generality) that $(\mathring{p}_1, \mathring{p}_2, \mathring{p}_3)$ are three independent vectors and therefore they represent three non collinear eigenvectors of $\tilde{H}$ associated with the eigenvalues λ_i for $i = \{1, 2, 3\}$ such that $\tilde{H}\mathring{p}_i = \lambda_i \mathring{p}_i$.

Exploiting again the consistency of the set $\mathcal{M}_n$, it follows that there exists a constant direction $\mathring{p}_k$ from the set $\{\mathring{p}_4, \dots \mathring{p}_n\}$ such that:

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

Since $\mathring{p}_k$ can be seen as a forth eigenvector for $\tilde{H}$ associated to the eigenvalue $\lambda_k = \pm |\tilde{H}\mathring{p}_k|$, this yields

$$\lambda_k \mathring{p}_k = \tilde{H}\mathring{p}_k = \frac{1}{|y_k|} \tilde{H} \sum_{i=1}^{3} b_i \mathring{p}_i = \frac{1}{|y_k|} \sum_{i=1}^{3} b_i \tilde{H}\mathring{p}_i$$

$$\lambda_k \sum_{i=1}^{3} b_i \mathring{p}_i = \sum_{i=1}^{3} b_i \lambda_i \mathring{p}_i$$

Using the fact that the measured directions form a consistent set, it follows that $b_i \neq 0, \ i = \{1, 2, 3\}$ and invoking the fact that $\det(\tilde{H}) = 1$, a straightforward identification shows that $\lambda_k = \lambda_1 = \lambda_2 = \lambda_3 = 1$. Consequently, $\tilde{H}$ converges asymptotically to the identity I. ∎

Remark 3.3: Note that the characterization of the stability domain remains an open problem and not addressed in the paper. Although, simulation results that we have performed tend to indicate that the stability domain is sufficiently large, this issue along with convergence property towards the equilibrium should be thoroughly analysed. △

B. Observer with partially known velocity of the rigid body

In this section we assume that U (18) is not available and the goal consists in providing an estimate $\hat{H}(t) \in SL(3)$ to drive the error term $\tilde{H} = \hat{H}H^{-1}$ to the identity matrix I and the error term $\tilde{\Gamma} = \Gamma - \hat{\Gamma}$ (resp. $\tilde{\Gamma}_1 = \Gamma_1 - \hat{\Gamma}_1$) to 0 if Γ (resp. Γ_1) is constant or slowly time varying. The observer when Γ is constant in $\{\mathring{A}\}$, is chosen as follows:

$$\dot{\hat{H}} = \hat{H}(\Omega_\times + \hat{\Gamma}) + k_P \omega \hat{H}, \tag{28}$$

$$\dot{\hat{\Gamma}} = [\hat{\Gamma}, \Omega_\times] + k_I \mathrm{Ad}_{\hat{H}^T} \omega \tag{29}$$

The observer when $\frac{V}{d}$ is constant in $\{A\}$, is defined as follows,

$$\dot{\hat{H}} = \hat{H}(\Omega_\times + \hat{\Gamma}_1 - \frac{1}{3}\mathrm{tr}(\Gamma_1)I) + k_P \omega \hat{H}, \tag{30}$$

$$\dot{\hat{\Gamma}}_1 = \hat{\Gamma}_1 \Omega_\times + k_I \mathrm{Ad}_{\hat{H}^T} \omega \tag{31}$$

Proposition 3.4: Consider a camera moving with kinematics (13) and (14) viewing a planar scene. Assume that $\mathring{A}$ is stationary and that the orientation velocity $\Omega \in \{A\}$ is measured and bounded. Let $H : \mathcal{A} \to \mathring{\mathcal{A}}$ denote the calibrated homography (12) and consider the kinematics (15) along with (18). Assume that H is bounded and Γ (resp. Γ_1) is constant in $\{\mathring{A}\}$ (resp. in {A}) such that it obeys (19) (resp. (20)).

Consider the nonlinear estimator filter defined by (28-29), (resp. (30-31)) along with the innovation $\omega \in \mathfrak{sl}(3)$ given by (25). Then, if the set $\mathcal{M}_n$ of the measured directions $\mathring{p}_i$ is consistent, the equilibrium $(\tilde{H}, \tilde{\Gamma}) = (I, 0)$ (resp. $(\tilde{H}, \tilde{\Gamma}_1) = (I, 0)$) is asymptotically stable.

Sketch of the Proof 3.5: We will consider only the situation where the estimate of Γ is used. The same arguments are used for the case where the estimate of Γ_1 is considered. Differentiating e_i (24) and using (28) yields

$$\dot{e}_i = \pi_{e_i}(k_P \omega - \mathrm{Ad}_{\hat{H}} \tilde{\Gamma}) e_i$$

Define the following candidate Lyapunov function

$$\mathcal{L} = \sum_{i=1}^{n} \frac{1}{2} |e_i - \mathring{p}_i|^2 + \frac{1}{2k_I} ||\tilde{\Gamma}||^2 \tag{32}$$

Differentiating $\mathcal{L}$ and using the fact that $\mathrm{tr}\left(\tilde{\Gamma}^T \left(\left[\tilde{\Gamma}, \Omega\right]\right)\right) = 0$, it follows that

$$\dot{\mathcal{L}} = \sum_{i=1}^{n} (e_i - \mathring{p}_i)^T \dot{e}_i - \mathrm{tr}\left(\tilde{\Gamma}^T \mathrm{Ad}_{\hat{H}^T} \omega\right)$$

Introducing the above expression of $\dot{e}_i$ and using the fact that $\mathrm{tr}(AB) = \mathrm{tr}(B^T A^T)$, it follows:

$$\begin{aligned}
\dot{\mathcal{L}} &= \sum_{i=1}^{n} (e_i - \mathring{p}_i)^T \pi_{e_i}(k_P \omega - \mathrm{Ad}_{\hat{H}} \tilde{\Gamma}) e_i - \mathrm{tr}\left(\mathrm{Ad}_{\hat{H}^{-1}} \omega^T \tilde{\Gamma}\right) \\
&= -\sum_{i=1}^{n} \mathring{p}_i^T \pi_{e_i}(k_P \omega - \mathrm{Ad}_{\hat{H}} \tilde{\Gamma}) e_i - \mathrm{tr}\left(\mathrm{Ad}_{\hat{H}^{-1}} \omega^T \tilde{\Gamma}\right) \\
&= -\,\mathrm{tr}\left(\sum_{i=1}^{n} e_i \mathring{p}_i^T \pi_{e_i}(k_P \omega - \mathrm{Ad}_{\hat{H}} \tilde{\Gamma}) + \mathrm{Ad}_{\hat{H}^{-1}} \omega^T \tilde{\Gamma}\right) \\
&= -\,\mathrm{tr}\Bigg(k_P \sum_{i=1}^{n} e_i \mathring{p}_i^T \pi_{e_i} \omega \\
&\qquad + \left(\mathrm{Ad}_{\hat{H}^{-1}} (\omega^T - \sum_{i=1}^{n} e_i \mathring{p}_i^T \pi_{e_i}) \tilde{\Gamma}\right)\Bigg)
\end{aligned}$$

Finally, introducing the expression of ω (25), we get:

$$\dot{\mathcal{L}} = -k_P \left\| \sum_{i=1}^{n} e_i \mathring{p}_i^T \pi_{e_i} \right\|^2$$

The derivative of the Lyapunov function is negative semi-definite, and equal to zero when $\omega = 0$. Given that Ω is bounded, it is easily verified that $\dot{\mathcal{L}}$ is uniformly continuous and Barbalat's Lemma can be used to prove asymptotic convergence of $\omega \to 0$. Using the same arguments used in the proof of theorem 3.2, it is straightforward to verify that $\tilde{H} \to I$. Consequently the left hand side of the Lyapunov expression (32) converges to zero and $\|\tilde{\Gamma}\|^2$ converges to a constant.

Computing the time derivative of $\tilde{H}$ and using the fact that ω converges to zero and $\tilde{H}$ converges to I, it is straightforward to show that:

$$\lim_{t \to \infty} \dot{\tilde{H}} = -\mathrm{Ad}_{\hat{H}} \tilde{\Gamma} = 0$$

Using boundedness of H, one can insure that $\lim_{t \to \infty} \tilde{\Gamma} = 0$.

IV. Simulation results

In this section, we illustrate the performance and robustness of the proposed observers through simulation results. The camera is assumed to be attached to an aerial vehicle moving in a circular trajectory which stays in a plane parallel to the ground. The reference camera frame $\{\mathring{A}\}$ is chosen as the NED (North-East-Down) frame situated above four observed points on the ground. The four observed points form a square whose center lies on the Z-axis of the NED frame $\{\mathring{A}\}$. The vehicle's trajectory is chosen such that the term Γ_1 defined by (20) remains constant, and the observer (30-31) is applied with the following gains: $k_P = 4$, $k_I = 1$. Distributed noise of variance 0.01 is added on the measurement of the angular velocity Ω. The chosen initial estimated homography $\hat{H}(0)$ corresponds to *i)* an error of $\pi/2$ in both pitch and yaw angles of the attitude, and *ii)* an estimated translation equal to zero. The initial value of $\hat{\Gamma}_1$ is set to zero. From $40s$ to $45s$, we assumed that the measurements of two observed points are lost. Then, from $45s$ we regain the measurements of all four points as previously.

The results reported in Fig. 1 show a good convergence rate of the estimated homography to the real homography (see from 0 to $40s$ and from $45s$). The loss of point measurements marginally affects the global performance of the proposed observed. Note that in this situation, no existing method for extracting the homography from measurements of only two points is available.

V. Concluding remarks

In this paper we developed a nonlinear observer for a sequence of homographies represented as elements of the Special Linear group $SL(3)$. More precisely, the observer directly uses point correspondences from an image sequence without requiring explicit computation of the individual homographies between any two given images. The stability of the observer has been proved for both cases of known full group velocity and known rigid-body velocities only. Even if the characterization of the stability domain still remains an open issue, simulation results have been provided as a complement to the theoretical approach to demonstrate a large domain of stability.

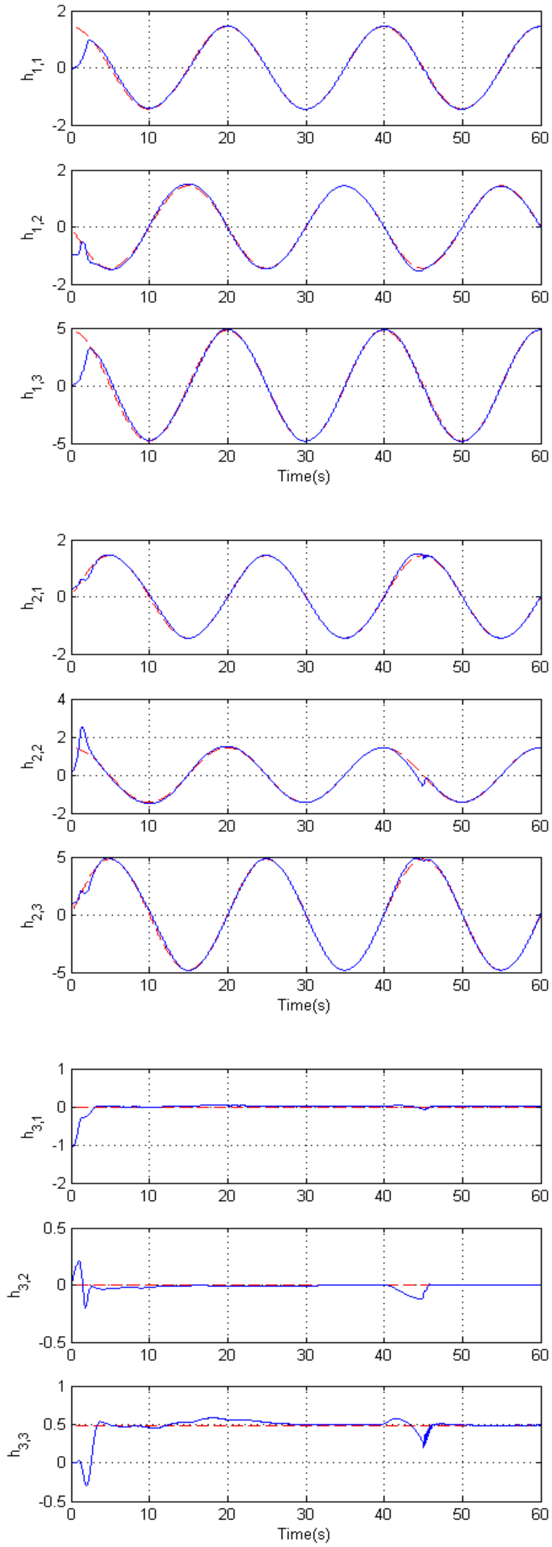

Fig. 1. Estimated homography (solid line) and true homography (dashed line) vs. time.

REFERENCES

[1] Faugeras, O. and Lustman, F., "Motion and structure from motion in a piecewise planar environment". International Journal of Pattern Recognition and Artificial Intellihence, 2(3) : 485-508, 1988.

[2] S. Benhimane and E. Malis. Homography-based 2d visual tracking and servoing. *International Journal of Robotic Research*, 26(7): 661–676, 2007.

[3] S. Bonnabel and P. Rouchon. *Control and Observer Design for Nonlinear Finite and Infinite Dimensional Systems*, volume 322 of *Lecture Notes in Control and Information Sciences*, chapter On Invariant Observers, pages 53–67. Springer-Verlag, 2005.

[4] Faugeras, O. and Lustman, F., "Motion and structure from motion in a piecewise planar environment". International Journal of Pattern Recognition and Artificial Intellihence, 2(3) : 485-508, 1988.

[5] R. Hartley and A. Zisserman. *Multiple View Geomerty in Computer Vision*. Cambridge University Press, second edition, 2003.

[6] E. Malis and M. Vargas, Deeper understanding of the homography decomposition for vision-based control, INRIA, Tech. Rep. 6303, 2007, available at http://hal.inria.fr/inria-00174036/fr/.

[7] C. Lageman, R. Mahony, and J. Trumpf. State observers for invariant dynamics on a lie group. *Conf. on the Mathematical Theory of Networks and Systems*, 2008.

[8] C. Lageman, J. Trumpf, and R. Mahony. Gradient-like observers for invariant dynamics on a lie group. *IEEE Trans. on Automatic Control*, to appear.

[9] R. Mahony, T. Hamel, and J.-M. Pflimlin. Non-linear complementary filters on the special orthogonal group. *IEEE Trans. on Automatic Control*, 53(5): 1203–1218, 2008.

[10] E. Malis and F. Chaumette. Theoretical improvements in the stability analysis of a new class of model-free visual servoing methods. *IEEE Trans. on Robotics and Automation*, 18(2): 176–186, 2002.

[11] E. Malis, F. Chaumette, and S. Boudet. 2 1/2 d visual servoing. *IEEE Trans. on Robotics and Automation*, 15(2):234–246, 1999.

[12] E. Malis and M. Vargas. Deeper understanding of the homography decomposition for vision-based control. Research Report 6303, INRIA, 2007.

[13] 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)*, 2009.

[14] E. Malis, T. Hamel, R. Mahony, and P. Morin. Estimation of Homography Dynamics on the Special Linear Group, in *Visual Servoing via Advanced Numerical Methods*, Eds G. Chesi and K. Hashimoto, Springer. Chapter 8, pages 139–158, 2010.

[15] P. Martin and E. Salaün. Invariant observers for attitude and heading estimation from low-cost inertial and magnetic sensors. *IEEE Conf. on Decision and Control*, pages 1039–1045, 2007.

[16] P. Martin and E. Salaün. An invariant observer for earth-velocity-aided attitude heading reference systems. *IFAC World Congress*, pages 9857–9864, 2008.

[17] C. Mei, S. Benhimane, E.Malis and P. Rives. Efficient Homography-Based Tracking and 3-D Reconstruction for Single-Viewpoint Sensors. *IEEE Trans. On Robotics*, VOL. 24, NO. 6, December 2008.

[18] J. Thienel and R. M. Sanner. A coupled nonlinear spacecraft attitude controller and observer with an unknow constant gyro bias and gyro noise. *IEEE Trans. on Automatic Control*, 48(11): 2011–2015, 2003.

[19] Weng, J., Ahuja, N. and Huang, T. S., "Motion and structure from point correspondences with error estimation: planar surfaces" IEEE Trans. Pattern Aanalysis and Machine Intelligence, 39(12) : 2691-2717, 1991.

[20] Zhang, Z. and Hanson, A. R., "3D Reconstruction Based On Homography Mapping" Proc. Image nderstanding Workshop, Morgan Kaufman, pages 1007-1012, 1996.