



Homography Observer Design on Special Linear Group $SL(3)$ with Application to Optical Flow Estimation

Ninad Manerikar, Minh-Duc Hua, Tarek Hamel

► To cite this version:

Ninad Manerikar, Minh-Duc Hua, Tarek Hamel. Homography Observer Design on Special Linear Group $SL(3)$ with Application to Optical Flow Estimation. European Control Conference (ECC), Jun 2018, Limassol, Cyprus. hal-01878702

HAL Id: hal-01878702

<https://hal.science/hal-01878702>

Submitted on 21 Sep 2018

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 Observer Design on Special Linear Group $SL(3)$ with Application to Optical Flow Estimation

Ninad Manerikar, Minh-Duc Hua, Tarek Hamel

Abstract—In this paper we present a novel approach for homography observer design on the Special Linear group $SL(3)$. This proposed approach aims at computing online, the optical flow estimation extracted from continuous homography obtained by capturing video sequences from aerial vehicles equipped with a monocular camera. The novelty of the paper lies in the linearization approach undertaken to linearize a nonlinear observer on $SL(3)$. Experimental results have been presented to show the performance and robustness of the proposed approach.

I. INTRODUCTION

Originated from computer vision, the homography represents a mapping that relates two views of the same planar scene. It encodes the camera's pose (i.e. position and orientation), the distance between the camera and the scene and the scene's normal vector in a single matrix. The homography has played an important role in various computer vision and robotics applications, where the scenarios involve man-made environments that are composed of planar surfaces [1], [2]. It is also suitable for applications where the camera is sufficiently far from the viewed scene, such as the situations when ground images are taken from an aerial drone.

The homography estimation problem has been extensively studied in the past [3], [4], [9]. Most existing works consider the homography as an incidental variable and tend to solve a system of algebraic constraints on the frame-by-frame basis [9]. The quality of the homography estimates is strongly dependant on the algorithm used as well as the nature of the image features used for estimation. Image features for homography estimation are typically geometric (points, lines, conics, etc.) or texture. For a well textured scene, state-of-the-art methods can provide high quality homography estimates at the cost of high computational complexity (see [4]). For a scene with poor texture, while significant computational effort is still required, the poor quality of homography estimates is often an important issue. This is particularly the case in varying lighting conditions and in presence of specular reflections or moving shadows, where photogrammetric error criteria used in texture-based methods become ill-posed. Feature-based methods of image-to-image homography estimation are robust to these issues as long as good features and good matching are obtained. They have been the mainstay of robust computer vision in the past.

In a previous work [5], [6] by some authors of this paper, a nonlinear observer for homography estimation was proposed

based on the group structure of the set of homographies, the Special Linear group $SL(3)$. This nonlinear observer enhances the overall homography estimate between any two given images by using the velocity information in order to interpolate across a sequence of images. However, this observer still requires the pre-computation of the individual image homographies, which are then smoothed using filtering techniques, thus making this approach computationally expensive. Also one of the drawbacks of this approach is that in order to obtain improved homography estimates, each pair of images needs to have a sufficient amount of data features for the computation of homography. In another work [7] by some of the authors of the paper, the question of deriving an observer that directly uses point correspondences from an image sequence without requiring the prior reconstruction of the individual homographies has been addressed. This approach is not only computationally more efficient but also works for scenarios where there is lack of data features for the reconstruction of the homography matrix.

In the present paper, we address the same problem as [7] but with a different approach that directly exploits the basis of the Lie algebra of the group $SL(3)$ for observer design. The simplicity of the resulting observer makes it suitable for real-time implementation. The proposed observer is also ideal for the estimation of the so-called “continuous homography” and the optical flow by exploiting the homography estimated from every two consecutive images obtained from a combined Camera-IMU (Inertial Measurement Unit) system.

The paper is structured into six sections including the introduction and the concluding sections. Section II gives a brief description of the notation and the math related to homography. In section III a linear approach for observer design on $SL(3)$ is proposed using point correspondences and the knowledge of the group velocity. In section IV the computation of the optical flow estimate extracted from continuous homography is presented. Experimental results supporting the proposed approach are presented in section V. Finally, some concluding remarks and future works are provided in section VI.

II. NOTATION AND THEORETICAL BACKGROUND

A. Mathematic background

The special linear group $SL(3)$ and its Lie algebra $\mathfrak{sl}(3)$ are defined as

$$SL(3) := \{H \in \mathbb{R}^{3 \times 3} \mid \det(H) = 1\}$$
$$\mathfrak{sl}(3) := \{U \in \mathbb{R}^{3 \times 3} \mid \text{tr}(U) = 0\}$$

N. Manerikar, M.-D. Hua, T. Hamel are with I3S, Université Côte d'Azur, CNRS, Sophia Antipolis, France, *manerika(hua, hamel)@i3s.unice.fr*.

Since the Lie algebra $\mathfrak{sl}(3)$ is of dimension 8, it can be spanned by 8 generators so that for any $\Delta \in \mathfrak{sl}(3)$ there exists a unique vector $\delta \in \mathbb{R}^8$ such that

$$\Delta = \sum_{i=1}^8 \delta_i B_i \quad (1)$$

where the basis of $\mathfrak{sl}(3)$ are chosen as follows:

$$\begin{aligned} B_1 &= e_1 e_2^\top, & B_2 &= e_2 e_1^\top, & B_3 &= e_2 e_3^\top \\ B_4 &= e_3 e_2^\top, & B_5 &= e_3 e_1^\top, & B_6 &= e_1 e_3^\top \\ B_7 &= e_1 e_1^\top - \frac{1}{3}I, & B_8 &= e_2 e_2^\top - \frac{1}{3}I \end{aligned}$$

with I the identity element of $\mathbb{R}^{3 \times 3}$ and $\{e_1, e_2, e_3\}$ the canonical basis of \mathbb{R}^3 .

B. Homographies

Let \mathring{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 frame of reference. The position of the frame $\{A\}$ with respect to $\{\mathring{A}\}$ expressed in $\{\mathring{A}\}$ is denoted by $\xi \in \mathbb{R}^3$. The orientation of the frame $\{A\}$ with respect to $\{\mathring{A}\}$ is represented by a rotation matrix $R \in SO(3)$. Let \mathring{d} (resp. d) and $\mathring{\eta} \in S^2$ (resp. $\eta \in S^2$) denote the distance from the origin of $\{\mathring{A}\}$ (resp. $\{A\}$) to the planar scene and the normal vector pointing to the scene expressed in $\{\mathring{A}\}$ (resp. $\{A\}$). Hence the coordinates of the same point in the reference frame ($\mathring{P} \in \{\mathring{A}\}$) and in the current frame ($P \in \{A\}$) can be expressed by the relation

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

If the intrinsic parameters of the camera are known meaning that the camera is calibrated one can write¹

$$\mathring{P} \cong K^{-1} \mathring{p}, \quad P \cong K^{-1} p$$

where K is the camera calibration matrix and $\mathring{p} \in \mathring{\mathcal{A}}$ (resp. $p \in \mathcal{A}$) is the image of the point \mathring{P} (resp. P) when the camera is aligned with frame $\{\mathring{A}\}$ (resp. $\{A\}$) and can be written in the form $(u, v, 1)^\top$ using the homogeneous coordinate representation for that 2D image point. The image homography matrix that maps pixel coordinates from the current frame to the reference frame is given by

$$H_{im} = \gamma K \left(R + \frac{\xi \eta^\top}{d} \right) K^{-1}$$

where γ is a scale factor that, without loss of generality, can be chosen so that $H_{im} \in SL(3)$. In this case, γ is equal to $(d/\mathring{d})^{\frac{1}{3}}$ and corresponds to the second singular value of H_{im} [8]. The Euclidean Homography $H \in SL(3)$ is related to H_{im} by the following relation

$$H = K^{-1} H_{im} K = \gamma \left(R + \frac{\xi \eta^\top}{d} \right)$$

¹Equalities in projective geometry are defined up to a multiplicative constant using the notation \cong .

C. Homography kinematics

Let Ω and V denote the rigid body angular velocity and linear velocity of the current frame $\{A\}$ w.r.t the reference frame $\{\mathring{A}\}$, both expressed in $\{A\}$. Then the rigid body kinematics are

$$\begin{cases} \dot{R} &= R\Omega_\times \\ \dot{\xi} &= RV \end{cases} \quad (2)$$

with $[\cdot]_\times$ denoting the skew-symmetric matrix associated with the cross product, i.e. $u_\times v = u \times v$, $\forall u, v \in \mathbb{R}^3$.

Let us consider that the stationary planar scene is viewed by a camera attached to the rigid body with kinematics (2). Let $H \in SL(3)$ denote the associated homography matrix. Then, its kinematics are given by

$$\dot{H} = HU \quad (3)$$

where the group velocity $U \in \mathfrak{sl}(3)$ induced by the relative motion between the camera and the stationary planar scene satisfies [5]

$$U = \Omega_\times + \frac{V\eta^\top}{d} - \frac{\eta^\top V}{3d} I \quad (4)$$

The group velocity U given by (4) is often referred to as ‘‘Continuous Homography’’ in the literature [8].

III. OBSERVER DESIGN ON $SL(3)$ FOR HOMOGRAPHY ESTIMATION

The equation of the proposed homography estimator can be expressed as a kinematic filter system on $SL(3)$ as

$$\dot{\hat{H}} = \hat{H}U + \Delta \hat{H} \quad (5)$$

where the innovation term $\Delta \in (3)$ has to be designed in order to drive the group error $\tilde{H} := \hat{H}H^{-1}$ to identity, based on the assumption that we have a collection of n measurements $y_j = \frac{H^{-1}\mathring{y}_j}{|H^{-1}\mathring{y}_j|} \in S^2$ ($j = 1, \dots, n$), with $\mathring{y}_j \in S^2$ known and constant. Here y_j and \mathring{y}_j represent calibrated image points normalized onto the unit sphere and can be computed as

$$y_j = \frac{K^{-1}p_j}{|K^{-1}p_j|}, \quad \mathring{y}_j = \frac{K^{-1}\mathring{p}_j}{|K^{-1}\mathring{p}_j|}$$

The output errors e_j are defined as

$$e_j := \frac{\hat{H}y_j}{|\hat{H}y_j|} = \frac{\tilde{H}\mathring{y}_j}{|\tilde{H}\mathring{y}_j|} \quad (6)$$

which thus can be viewed as the estimates of \mathring{y}_j .

In order to design the innovation term Δ , we first develop linear approximations of both the dynamics of \tilde{H} and the system output errors e_i . Taking the time derivative of \tilde{H} ($= \dot{\hat{H}}H^{-1}$) and using first order approximation $\tilde{H} \approx I$ one obtains

$$\dot{\tilde{H}} = \Delta \tilde{H} = \sum_{i=1}^8 \delta_i B_i \tilde{H} \approx \sum_{i=1}^8 \delta_i B_i \quad (7)$$

where the linear representation $\delta \in \mathbb{R}^8$ of Δ via the relation (1) has been used.

Let $x \in \mathbb{R}^8$ denote the linear representation of \tilde{H} . One then deduces the following approximation

$$\tilde{H} = \exp\left(\sum_{i=1}^8 x_i B_i\right) \approx \left(I + \sum_{i=1}^8 x_i B_i\right) \quad (8)$$

Using (8) one obtains

$$\dot{\tilde{H}} \approx \sum_{i=1}^8 \dot{x}_i B_i \quad (9)$$

From (7) and (9) it is obvious that

$$\dot{x} \approx \delta \quad (10)$$

Now we focus on the linearization of the output vectors. From (6) and (8) the output errors e_j can be expressed as

$$\begin{aligned} e_j &\approx \frac{\dot{y}_j + \sum_{i=1}^8 x_i B_i \dot{y}_j}{|\dot{y}_j + \sum_{i=1}^8 x_i B_i \dot{y}_j|} \\ &\approx \left(\dot{y}_j + \sum_{i=1}^8 x_i B_i \dot{y}_j\right) \left(1 - \dot{y}_j^\top \sum_{i=1}^8 x_i B_i \dot{y}_j\right) \end{aligned}$$

Neglecting high order terms one gets

$$\begin{aligned} e_j &\approx \dot{y}_j + \sum_{i=1}^8 x_i B_i \dot{y}_j - \sum_{i=1}^8 x_i \dot{y}_j \dot{y}_j^\top B_i \\ \Rightarrow e_j - \dot{y}_j &\approx \sum_{i=1}^8 x_i \pi_{\dot{y}_j} B_i \dot{y}_j = C_j x \end{aligned}$$

with $\pi_y := I - yy^\top$, $\forall y \in S^2$, the projection operator on the plane orthogonal to y and

$$C_j := [\pi_{\dot{y}_j} B_1 \dot{y}_j \mid \cdots \mid \pi_{\dot{y}_j} B_8 \dot{y}_j] \in \mathbb{R}^{3 \times 8}$$

Stacking all n measurements in a vector as follows

$$Y := \begin{bmatrix} e_1 - \dot{y}_1 \\ \vdots \\ e_n - \dot{y}_n \end{bmatrix} \in \mathbb{R}^{3n}$$

one obtains

$$Y \approx Cx \quad (11)$$

with $C := \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix} \in \mathbb{R}^{3n \times 8}$.

From here, the innovation term δ (i.e. Δ) can be directly designed on the linear approximation system (10) using the linear approximation (11) of the output vector Y . In fact, if the matrix C is of rank 8 (it is well known that the homography is observable from the measurements of at least 4 linearly independent points [7]), then the design of the innovation term δ is straightforward. An obvious solution is

$$\delta = -kC^\top Y$$

with $k > 0$, resulting in the following stable closed-loop system, in first order approximations,

$$\dot{x} = \delta = -kC^\top Cx$$

Finally, $\Delta \in \mathfrak{sl}(3)$ is computed from $\delta \in \mathbb{R}^8$ using relation (1).

IV. APPLICATION TO OPTICAL FLOW ESTIMATION

Using a moving Camera-IMU (i.e. a combined system composed of a Camera and an Inertial Measurement Unit) that observes a stationary planar scene, the previously proposed algorithm can be applied to estimate the homography matrix H related to every two consecutive images. If the camera frequency is fast enough, then the continuous homography U defined by (4) can be approximately computed via logarithm operator as

$$U \approx \frac{1}{T} \log(H)$$

with T the camera sample time. Since H is normally close to the identity matrix, $\log(H)$ can then be approximated using Taylor expansions as follows:

$$\log(H) = \log(I - W) \approx \widetilde{\log}(H) := -W - \frac{W^2}{2} - \frac{W^3}{3} - \dots$$

with $W := I - H$. However, such an approximation no longer ensures that both $\log(H)$ and the resulting U will remain in $\mathfrak{sl}(3)$. Hence a reprojection on $\mathfrak{sl}(3)$ is needed

$$U \approx \frac{1}{T} \left(\widetilde{\log}(H) - \frac{1}{3} \text{tr}(\widetilde{\log}(H)) I \right)$$

Denoting $\phi := \frac{V}{d}$ and $\phi_\perp := \frac{V^\top \eta}{d} = \frac{-\dot{d}}{d}$, which respectively correspond to the so-called translational optical flow and its projection along the normal vector η . Our objective consists in obtaining the estimation of both ϕ and ϕ_\perp from the decomposition of the already computed continuous homography U .

Using the fact that the angular velocity of the Camera-IMU is measured by the gyrometers, one deduces

$$\bar{U} := \frac{V\eta^\top}{d} = U - \frac{1}{2} \gamma_2 (U + U^\top) I - \Omega_\times$$

with $\gamma_2(U + U^\top)$ the second largest eigenvalue of $U + U^\top$ [8]. Taking into account the fact that $\eta \in S^2$ one deduces

$$\Phi := \phi\phi^\top = \bar{U}\bar{U}^\top$$

Defining β as the vector of the diagonal elements of Φ as follows

$$\beta = \begin{bmatrix} \Phi_{11} \\ \Phi_{22} \\ \Phi_{33} \end{bmatrix}$$

one verifies that

$$\bar{U}^\top \beta = \eta(\phi_1^3 + \phi_2^3 + \phi_3^3)$$

From here, the estimate of ϕ is calculated as follows

$$\hat{\phi} = \begin{cases} \bar{U} \frac{\bar{U}^\top \beta}{|\bar{U}^\top \beta|} & \text{if } |\bar{U}^\top \beta| > \epsilon_\eta \\ 0 & \text{if } |\bar{U}^\top \beta| < \epsilon_\eta \end{cases}$$

with $\epsilon_\eta > 0$ being a small given threshold. Finally, the estimate of ϕ_\perp is straightforwardly obtained by

$$\phi_\perp = \text{tr}(\bar{U})$$

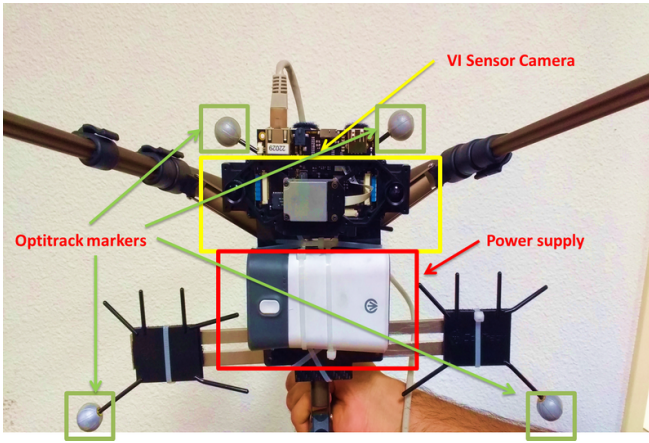


Fig. 1. Experimental setup



Fig. 2. Textured planar horizontal ground (target) used for experiment validations

V. EXPERIMENTAL RESULTS

A. Experimental Setup

A Visual-Inertial (VI) sensor developed by the company Skybotix and the Autonomous Systems Lab (ETH Zurich) has been used to perform experimental validation (see Fig. 1). This VI-sensor consists of two cameras and two IMU's (composed of a 3-axis gyrometer and a 3-axis accelerometer). However, in order to validate the proposed approach only one camera and one IMU are sufficient. One of the main reasons for using the VI-sensor is the possibility to obtain perfectly time-synchronized images and IMU measurements ($20Hz$ for camera and $200Hz$ for IMU). For validation purposes, the ground truth is obtained by using the highly accurate Optitrack Motion Capture system available at the I3S lab that provides the full pose of the Camera-IMU system at $120Hz$.

The proposed algorithm has been implemented using C++ on an Intel Core i7 CPU running at $3.40GHz$. A high speed ethernet cable is used to carry out the transmission of data from the camera to the PC. The Linux based PC is in charge of carrying out two principal software tasks:

- Acquisition of data (images as well as IMU data) by interfacing with the camera hardware.

- Continuous homography estimation based on two consecutive images that is further decomposed to obtain the estimation of $\phi(= \frac{V}{d})$ and $\phi_{\perp}(= -\frac{d}{d})$ in real-time.

The detection of features and extraction of descriptors from the images has been carried out by using FAST Feature Detector and ORB Descriptor Extractor algorithms that have been already implemented in the OpenCV library. These algorithms are used for the simple reason that they are very time-efficient and thus more suited for real-time implementation.

Remark 1 The continuous homography estimate is obtained at $20Hz$ due to the fact that the VI-sensor only provides images at that frequency. However, our offline tests have shown that the continuous homography and the optical flow ϕ and ϕ_{\perp} estimates can be obtained at a higher frequency of $50Hz$. This has suggested us to develop our proper hardware Camera-IMU using a fast camera in the near future.

B. Experimental Results

The experiment reported below has been performed online with the VI-sensor camera looking downward to observe a well textured planar horizontal ground (see Fig. 2). A video showing this experiment is provided as a supplementary material and is also available at <https://goo.gl/i8zGj2>.

From Figs. 3 and 4, one can observe that the estimation of the optical flow ϕ and its perpendicular component ϕ_{\perp} obtained from the decomposition of the estimated continuous homography are pretty accurate when compared to the corresponding ground-truth data.

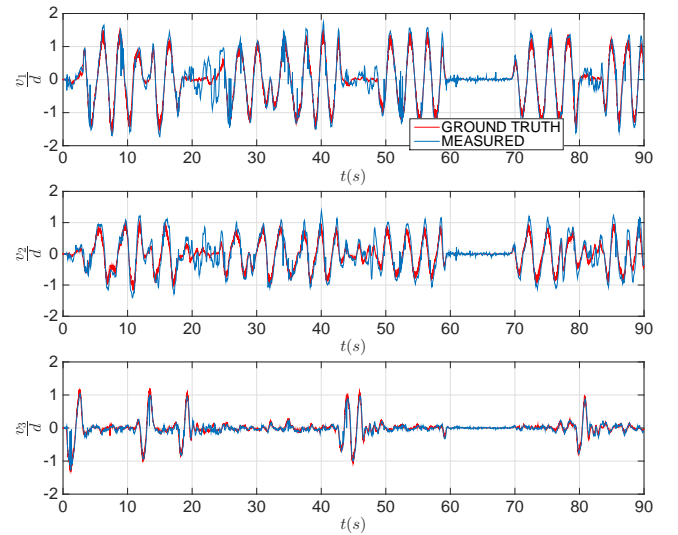


Fig. 3. Optical flow components estimated from images (blue curves) and derived from ground truth pose (red curves) versus time (s)

VI. CONCLUSION

This paper presents a general analysis on a novel linear approach for observer design posed directly on the special linear group $SL(3)$ and its application for translational optical flow estimation. In order to support the theoretical approach, the performance of this observer was demonstrated

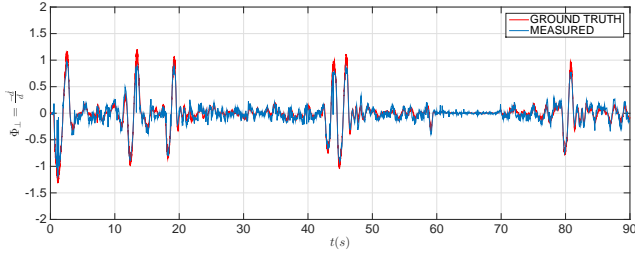


Fig. 4. $\phi_{\perp} = -\frac{\dot{d}}{d}$ estimated from images (blue curve) and derived from ground truth pose (red curve) versus time (s)

with a set of real world experiments. The experiments showed the good quality of estimates that could be obtained in real time. The proposed approach for the estimation of optical flow can also be used as a cue for haptic force feedback [10]. In the near future, the idea is to perform this experiment with a camera which gives images at a much higher frequency (for e.g. $100Hz$) in order to obtain the optical flow estimates at a much faster rate ($50Hz$).

ACKNOWLEDGMENT

This work was supported by the FUI GreenExplorer project and the EQUIPEX Robotex project.

REFERENCES

- [1] Y. Fang, W. Dixon, D. Dawson, and P. Chawda. Homography-based visual servoing of wheeled mobile robots. *IEEE Trans. on Systems, Man, and Cybernetics - Part B*, 35: 10411050, 2005
- [2] F. Fraundorfer, C. Engels, and D. Nister. Topological mapping, localization and navigation using image collections. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 38723877, 2007.
- [3] 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
- [4] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*, Cambridge University Press, second edition, 2003
- [5] R. Mahony, T. Hamel, P. Morin, and E. Malis. Nonlinear complementary filters on the special linear group. *International Journal of Control*, 85(10): 15571573, 2012.
- [6] E. Malis, T. Hamel, R. Mahony, and P. Morin. Dynamic estimation of homography transformations on the special linear group for visual servo control. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 14981503, 2009.
- [7] T. Hamel, R. Mahony, J. Trumpf, P. Morin, and M.-D. Hua. Homography estimation on the Special Linear group based on direct point correspondence. In *IEEE Conference on Decision and Control (CDC)*, pages 79027908, 2011
- [8] Y. Ma, S. Soatto, J. Kosecka, and S. S. Sastry. *An Invitation to 3-D Vision: From Images to Geometric Models*, Springer Verlag, 2003.
- [9] A. Agarwal, C.V. Jawahar, and P.J. Narayanan. A survey of planar homography estimation techniques. Centre for Visual Information Technology, Tech. Rep. IIIT/TR/2005/12, 2005
- [10] S. Omari, M.-D. Hua, G. Ducard, and T. Hamel. Bilateral Haptic Teleoperation of VTOL UAVs. In *IEEE International Conference on Robotics and Automation (ICRA)*, pages 23932399, 2013.