



**HAL**  
open science

# Visual-inertial sensor fusion and state estimation for the navigation of autonomous vehicles

Ninad Manerikar

► **To cite this version:**

Ninad Manerikar. Visual-inertial sensor fusion and state estimation for the navigation of autonomous vehicles. Automatic. Université Côte d'Azur, 2022. English. NNT : 2022COAZ4111 . tel-03969404

**HAL Id: tel-03969404**

**<https://theses.hal.science/tel-03969404>**

Submitted on 2 Feb 2023

**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.

# THÈSE DE DOCTORAT

## Fusion de capteurs visuels-inertiels et estimation d'état pour la navigation des véhicules autonomes

**Ninad MANERIKAR**

Laboratoire: CNRS-I3S

**Présentée en vue de l'obtention  
du grade de docteur en Automatique,  
Traitement du Signal et des Images  
de l'Université Côte d'Azur**

**Dirigée par:**

Minh Duc Hua (Co-directeur de thèse et  
Encadrant principal)

Tarek Hamel (Directeur de thèse)

**Soutenue le:** 28/01/2022

**Devant le jury, composé de:**

Ezio Malis, Directeur de Recherche INRIA, Sophia-  
Antipolis

Rita Cunha, Associate Professor, Université de Lisbonne

Julien Marzat, Ingénieur-HDR, Université Paris-Saclay

Carlos J. F. Silvestre, Professeur, Université de Macao

Vincent Creuze, MCF (HDR), LIRMM CNRS-UM

Minh Duc Hua, Chargé de recherche, I3S CNRS-UCA

Tarek Hamel, Professeur, I3S CNRS-UCA



# THÈSE DE DOCTORAT

UNIVERSITÉ CÔTE D'AZUR

ÉCOLE DOCTORALE STIC

Sciences et Technologies de l'Information et de la Communication

Laboratoire I3S-CNRS

## **Fusion de capteurs visuels-inertiels et estimation d'état pour la navigation des véhicules autonomes**

Thèse pour obtenir le grade de

Docteur en Automatique, Traitement du Signal et des Images

Présentée par

**Ninad MANERIKAR**

Soutenue publiquement le 28 Janvier 2022 devant le jury composé de:

Président:	Ezio Malis	Directeur de Recherche INRIA, Sophia-Antipolis
Rapporteur:	Rita Cunha	Associate Professor, Université de Lisbonne
Rapporteur:	Julien Marzat	Ingenieur-HDR, Université Paris-Saclay
Rapporteur:	Carlos J. F. Silvestre	Professeur, Université de Macao
Co-directeur de thèse : et Encadrant principal	Minh Duc Hua	Chargé de recherche, I3S CNRS-UCA
Directeur de thèse:	Tarek Hamel	Professeur, I3S CNRS-UCA
Invité:	Vincent Creuze	MCF (HDR), LIRMM CNRS-UM



# Abstract

Accurate state estimation is a fundamental problem for the navigation of Autonomous vehicles. This is particularly important when the vehicle is navigating through cluttered environments or it has to navigate in close proximity to its physical surroundings in order to perform localization, obstacle avoidance, environmental mapping etc. Although several algorithms were proposed in the past for this problem of state estimation, they were usually applied to a single sensor or a specific sensor suite. To this end, researchers in the computer vision and control community came up with a visual-inertial framework (Camera + Imu) that exploit the combined properties of this sensor suite to produce precise local estimates (position, orientation, velocity etc). Taking inspiration from this, my thesis focuses on developing nonlinear observers for State Estimation by exploiting the classical Riccati design framework with a particular emphasis on visual-inertial sensor fusion. In the context of this thesis, we use a suite of low-cost sensors consisting of a monocular camera and an IMU. Throughout the thesis, the assumption on the planarity of the visual target has been considered.

In the present thesis, two research topics have been considered. Firstly, an extensive study for the existing techniques for homography estimation has been carried out after which a novel nonlinear observer on the  $SL(3)$  group has been proposed with application to optical flow estimation. The novelty lies in the linearization approach undertaken to linearize a nonlinear observer on  $SL(3)$ , thus making it more simplistic and suitable for practical implementation. Then, another novel observer based on deterministic Riccati observer has been proposed for the problem of partial attitude, linear velocity and depth estimation for planar targets. The proposed approach does not rely on the strong assumption that the IMU provides the measurements of the vehicle's linear acceleration in the body-fixed frame. Again experimental validations have been carried out to show the performance of the observer. An extension to this observer has been further proposed to filter the noisy optical flow estimates obtained from the extraction of continuous homography. Secondly, two novel observers for tackling the classical problem of homography decomposition have been proposed. The key contribution here lies in the design of two deterministic Riccati observers for addressing the homography decomposition problem instead of solving it on a frame-by-frame basis like traditional algebraic approaches. The performance and robustness of the observers have been validated over simulations and practical experiments. All the observers proposed above are part of the **Homography-Lab library** that has been evaluated at the TRL 7 (Technology Readiness Level) and is protected by the French APP (Agency for the Protection of Programs) which serves as the main brick for various applications like velocity, optical flow estimation and visual homography based stabilization.

**Keywords:** Autonomous vehicles, State Estimation, Homography and Homography Decomposition, Visual-Inertial Sensor Fusion, Nonlinear observers



# Résumé

L'estimation précise de l'état du système est un problème fondamental pour la navigation des véhicules autonomes. Ceci est particulièrement important lorsque le véhicule navigue dans des environnements encombrés ou à proximité d'obstacles, afin d'effectuer la localisation, l'évitement d'obstacles, la cartographie de l'environnement, etc. Bien que plusieurs algorithmes aient été proposés dans le passé pour ce problème d'estimation d'état, ils impliquent généralement un seul capteur ou plusieurs du même type. Afin de pouvoir exploiter les propriétés de multiples capteurs dotés de caractéristiques différentes (tels que Camera, IMU, Lidar, etc.), les chercheurs de la communauté de vision et de contrôle ont mis au point des modèles mathématiques qui produisent des estimations locales précises (position, orientation, vitesse, etc.). En m'inspirant de cela, ma thèse se concentre sur le développement d'observateurs non-linéaires pour l'estimation d'état en exploitant les algorithmes classiques de type Riccati en mettant l'accent sur la fusion de capteurs visuels-inertiels. Dans le cadre de cette thèse, nous utilisons une suite de capteurs à faible coût composée d'une caméra monoculaire et d'une centrale inertielle. Dans le cadre de la vision monoculaire, nous faisons l'hypothèse que la cible est pratiquement plate. Bien que cette hypothèse soit restrictive, les solutions proposées sont pertinentes pour de nombreuses applications dans les domaines de la robotique aérienne, terrestre et sous-marine. Dans ce contexte, deux nouveaux observateurs non linéaires sont proposés, le premier pour l'estimation de l'homographie et le deuxième pour l'estimation de l'attitude partielle, de la vitesse linéaire et de la profondeur. Dans la deuxième partie de la thèse, deux nouveaux observateurs déterministes de Riccati sont proposés pour traiter le problème classique de décomposition d'homographie au lieu de le résoudre image par image comme les approches algébriques traditionnelles. Tous ces travaux sont publiés dans des conférences internationales de haute niveau. Tous les observateurs proposés ci-dessus font partie de la bibliothèque HomographyLab dont je suis l'un des principaux contributeurs. Cette bibliothèque a été évaluée au niveau TRL 7 (Technology Readiness Level) et est protégée par l'APP (Agence pour la Protection des Programmes) qui sert de brique principale pour diverses applications telles que l'estimation de vitesse et de flux optique, et la stabilisation basée sur l'homographie visuelle.

**Keywords:** Véhicules autonomes, Estimation d'état, Homographie et Décomposition d'homographie, Fusion de capteurs visuels-inertiels, Observateurs non linéaires

*Do or do not. There is no try*



## *Acknowledgements*

I owe my deepest gratitude to my supervisors, CR. Minh Duc HUA and PR. Tarek HAMEL, for their academic guidance, discussions and constant encouragement during my PhD project. I would also like to thank DR. Claude SAMSON for his invaluable advice and support throughout my Phd. I would like to particularly thank again my mentor CR. Minh Duc HUA who always believed in me, sometimes even more than I did myself.

Then, I would like to acknowledge the members of the Jury for evaluating the present work. My special thanks go to Dr. Ezio MALIS for chairing the Jury. I am grateful to PR. Carlos J. F. SILVESTRE, Prof. Rita CUNHA and Dr. Julien MARZAT for reviewing this manuscript. I am also thankful to Dr. Vincent CREUZE for being the invited member of the Jury.

I would like to thank André Anglade, Lam Hung Nguyen, Simone De Marco, Andrew Comport, Jean Martinez and Jean-Marie Kai for all their help and support to complete my work successfully. I am also extremely lucky to have made a family away from home with some of the most amazing people. I would like to thank Fernando, Miguel, Melissa, Melpo, Dimitris, Diana, Vasilina, Thomas, Lucy, Francisco, Pedro, Denis, Myriana, Majdi, Antonio for giving me constant support, friendship and a memorable life during my stay here. I would like to particularly thank Eva (my flatmate and my best friend here), with whom I have spent the most amazing moments together in these three years and who has supported me through everything in all the good times as well as the most difficult ones. Thank you all for the time we spent together.

I would also like to thank my chaddi buddies and my lifetime friends Roshni, Ankit, Dipti, Mansi, Vrushali, Lakshmi, Tanvi, Sumeel who have always been there for me, no matter what even though it's very difficult to meet them since all of us stay in different continents. I also want to thank Deeksha, Jenet and Swadha for being by my side when life got tough, always offering support where it was needed, and always listening when I needed to vent.

Finally and most importantly, I cannot thank enough my mother, father, aaji, appa and my incredible brother for all the inspiration, guidance, support, encouragement and love. I am forever indebted to them for their unconditional love, and all the sacrifices they have made to support me in doing whatever I wished to do. No words are enough to describe what you have done for me.



# Visual-Inertial Sensor Fusion and State Estimation for the Navigation of Autonomous Vehicles

**Ninad MANERIKAR**

# Contents

<b>Notation</b>	<b>xiv</b>
<b>List of Figures</b>	<b>xvi</b>
<b>1 Motivations, contributions and thesis structure</b>	<b>2</b>
1.1 Motivations and contributions . . . . .	2
1.2 Thesis structure . . . . .	6
<b>2 Theoretical recalls on State Estimation</b>	<b>8</b>
2.1 Introduction . . . . .	8
2.2 Observability of Systems . . . . .	9
2.2.1 Observability of Linear Time Invariant Systems . . . . .	9
2.2.2 Observability of Linear Time Varying Systems . . . . .	10
2.3 Observer Design . . . . .	12
2.3.1 Observer Design for Linear System . . . . .	12
2.3.1.1 Kalman Filter . . . . .	13
2.3.2 Observers for Nonlinear Systems . . . . .	13
2.3.2.1 Extended Kalman Filter . . . . .	14
2.3.2.2 Deterministic Riccati Observers . . . . .	15
<b>3 Introduction to Homography and Continuous Homography</b>	<b>17</b>
3.1 Introduction . . . . .	17
3.2 Preliminary material for Homography Estimation . . . . .	18
3.2.1 Perspective Projection . . . . .	18
3.2.2 Homographies . . . . .	19

3.2.3	Homographies as elements of the Special Linear Group $SL(3)$ . . .	21
3.2.4	Rigid-body and Homography Dynamics . . . . .	22
3.3	Existing feature-based homography estimation techniques . . . . .	23
3.3.1	A classical algebraic algorithm of homography estimation . . . . .	24
3.3.2	Nonlinear homography observer on $SL(3)$ . . . . .	25
3.4	Novel approach for Homography Observer design on $SL(3)$ . . . . .	26
3.5	Proof of Proposition 1 . . . . .	29
3.6	Application to optical flow estimation . . . . .	31
3.6.1	Experimental Results . . . . .	32
3.6.1.1	Experimental Setup . . . . .	32
3.6.1.2	Experimental Results and Conclusions . . . . .	34
<b>4</b>	<b>Riccati observers for state estimation exploiting optical flow and IMU mea-</b>	
	<b>surements</b> . . . . .	<b>35</b>
4.1	Introduction . . . . .	36
4.2	System Equations and Measurements for Observer Design . . . . .	38
4.3	Partial Attitude, Linear Velocity and Depth Estimation of a Camera ob-	
	serving a planar target using continuous homography and inertial data .	39
4.3.1	Observer Derivation . . . . .	39
4.3.2	Observability analysis . . . . .	41
4.3.3	Yaw estimation . . . . .	45
4.4	Practical implementation aspects . . . . .	46
4.4.1	Unit quaternion equivalence . . . . .	46
4.4.2	Hybrid discrete-continuous version . . . . .	46
4.4.3	Practical solutions for the boundedness of $P$ . . . . .	47
4.4.4	Simulation results . . . . .	48
4.4.5	Experimental evaluation . . . . .	51
4.4.5.1	Experimental setup . . . . .	51
4.4.5.2	Experimental results . . . . .	53
4.5	Extension to observer design for optical flow filtering . . . . .	55
4.5.1	System Equations and Measurements . . . . .	56
4.5.2	Observer derivation . . . . .	56
4.5.3	Observability analysis . . . . .	57
4.5.4	Experimental Results . . . . .	61
<b>5</b>	<b>Homography Decomposition</b> . . . . .	<b>63</b>
5.1	Introduction . . . . .	64
5.2	Problem Statement . . . . .	65

5.2.1	Homography decomposition problem . . . . .	65
5.2.2	Brief overview of the existing approaches for Homography Decomposition . . . . .	67
5.2.3	System equations and measurements for observer design . . . . .	68
5.3	Observer design for decomposing the homography matrix . . . . .	69
5.3.1	Observer derivation . . . . .	69
5.3.2	Observability and stability analysis . . . . .	72
5.4	Observer design for decomposing the inverse homography matrix . . . . .	75
5.4.1	Observer derivation . . . . .	75
5.4.2	Observability and stability analysis . . . . .	78
5.5	Simulation results . . . . .	78
5.6	Experimental validations . . . . .	84
5.6.1	Experimental setup . . . . .	85
5.6.2	Experimental results . . . . .	86
<b>6</b>	<b>Implementation Details of the HomographyLab Software Library</b>	<b>92</b>
6.1	Software Implementation Details . . . . .	93
6.2	I3S-AUV SOFTWARE ARCHITECTURE . . . . .	94
6.2.1	Software architecture in ROS . . . . .	95
6.2.2	Ros Nodes used in our Architecture . . . . .	96
6.2.3	Transfer of Data between Ros Nodes using Topics . . . . .	99
	<b>Conclusion</b>	<b>100</b>
	<b>Bibliography</b>	<b>103</b>

# Notation

## Constants, variables and symbols

$\mathbb{R}^n$	n-dimensional Euclidean space
$e_1, \dots, e_n$	the canonical basis of $\mathbb{R}^n$
$I_n$	the identity of $\mathbb{R}^{n \times n}$
$0_n$	the zero matrix of $\mathbb{R}^{n \times n}$
$\{\mathcal{A}\}$	the inertial reference system
$\{\mathcal{B}\}$	the body fixed frame attached to the vehicle
$\{\mathcal{I}\}$	the inertial fixed frame attached to the earth chosen as north-east-down frame
$\mathcal{B}_r^{n_1}$	the closed ball in $\mathbb{R}^{n_1}$ of radius $r$
$S^n$	n-dimensional sphere embedded in $\mathbb{R}^{n+1}$ with radius equal to one
$\pi_u$	projection onto the tangent space of the unit n-dimensional sphere at the point $u \in S^n$ , $\pi_u \triangleq I_{n+1} - uu^\top$
$\Omega$	angular velocity vector
$V$	linear velocity vector
$SL(3)$	the Special Linear group, the set of all real valued $3 \times 3$ matrices with unit determinant
$\mathfrak{sl}(3)$	Lie-algebra of $SL(3)$ , the set of matrices with trace equal to zero
$SO(3)$	the Special Orthogonal group of the orthogonal $3 \times 3$ matrices with unit determinant
$R$	rotation matrix, $R \in SO(3)$
$g$	the gravity constant

### Operators and functions

$(\cdot)^\top$	the transpose operator on a matrix or vector
$u_\times$	the skew-symmetric matrix associated with the cross product by vector $u \in \mathbb{R}^3$ , i.e., $u_\times v = u \times v, \forall v \in \mathbb{R}^3$
$ \cdot $	the Euclidean norm in $\mathbb{R}^n$
$\ \cdot\ $	the Frobenius norm in $\mathbb{R}^{n \times n}$

### List of Acronyms

<b>BRIEF</b>	Binary Robust Independent Elementary Features
<b>EKF</b>	Extended Kalman Filter
<b>FAST</b>	Features from Accelerated Segment Test
<b>GPS</b>	Global Positioning System
<b>IMU</b>	Inertial Measurement Unit
<b>LTI</b>	Linear Time-Invariant
<b>LTV</b>	Linear Time-Varying
<b>ORB</b>	Oriented Fast and Rotated BRIEF
<b>PnP</b>	Perspective-n-Points
<b>SIFT</b>	Scale-Invariant Feature Transform
<b>SURF</b>	Speeded Up Robust Features
<b>UAV</b>	Unmanned Aerial Vehicle
<b>AUV</b>	Autonomous Underwater Vehicle
<b>UKF</b>	Unscented Kalman Filter
<b>MEMS</b>	Micro-electromechanical Systems

# List of Figures

3.1	Euclidean homography relating the camera's pose, the distance to the plane and the plane's normal vector by $H = R + \frac{1}{d}\xi\eta^\top$ . . . . .	19
3.2	Experimental setup . . . . .	33
3.3	Textured planar horizontal ground (target) used for experiment validations	33
3.4	Optical flow components estimated from images (blue curves) and derived from ground truth pose (red curves) versus time ( $s$ ) . . . . .	35
3.5	$\phi_\perp = \frac{-\dot{d}}{d}$ estimated from images (blue curve) and derived from ground truth pose (red curve) versus time ( $s$ ) . . . . .	35
4.1	(Simulation) Estimated and ground-truth attitudes represented by roll, pitch and yaw Euler angles ( $deg$ ) versus time ( $s$ ) . . . . .	48
4.2	(Simulation) Estimated and ground-truth linear velocity components in body-fixed frame ( $m/s$ ) versus time ( $s$ ) . . . . .	49
4.3	(Simulation) Estimated and ground-truth depth inverse ( $m^{-1}$ ) versus time ( $s$ ) . . . . .	49
4.4	(Simulation) The Frobenius norm $ P $ versus time ( $s$ ) . . . . .	50
4.5	(Experiment) Optical flow components measured from images (blue curves) and derived from ground truth pose (red curves) versus time ( $s$ ) . . . . .	52
4.6	(Experiment) $\phi_\perp = -\frac{\dot{d}}{d}$ measured from images (blue curve) and derived from ground truth pose (red curve) versus time ( $s$ ) . . . . .	53
4.7	(Experiment) Estimated and ground-truth attitudes represented by roll, pitch and yaw Euler angles ( $deg$ ) versus time ( $s$ ) . . . . .	53
4.8	(Experiment) Estimated and ground-truth depth inverse ( $m^{-1}$ ) versus time ( $s$ ) . . . . .	54

4.9	(Experiment) Estimated and ground-truth linear velocity components in body-fixed frame ( $m/s$ ) versus time ( $s$ ) . . . . .	54
4.10	Optical flow components estimated from images (green curves), filtered optical flow components from the proposed observer (blue curves) and derived from ground truth pose (red curves) versus time ( $s$ ) . . . . .	62
4.11	Zoom of the optical flow components from time $t = 15s$ to $t = 25s$ . . . . .	63
5.1	Euclidean homography relating the camera's pose, the distance to the plane and the plane's normal vector by $H = R + \frac{1}{d}\xi\eta^\top$ . . . . .	66
5.2	<b>Scenario 1 (Method 1)</b> – Estimated and real attitudes represented by roll, pitch and yaw Euler angles ( $deg$ ) and the attitude estimate error versus time ( $s$ ) (LEFT) and their zooms (RIGHT). . . . .	78
5.3	<b>Scenario 1 (Method 1)</b> – – Estimated and real scaled position and norm of the scaled position error versus time ( $s$ ) (LEFT) and their zooms (RIGHT). . . . .	79
5.4	<b>Scenario 1 (Method 1)</b> – Estimation error of the normal vector estimate represented by $1 - \hat{\eta}^\top \eta$ . . . . .	79
5.5	<b>Scenario 1 (Method 2)</b> – Estimated and real attitudes represented by roll, pitch and yaw Euler angles ( $deg$ ) and the attitude estimate error versus time ( $s$ ) (LEFT) and their zooms (RIGHT). . . . .	80
5.6	<b>Scenario 1 (Method 2)</b> – – Estimated and real scaled position and norm of the scaled position error versus time ( $s$ ) (LEFT) and their zooms (RIGHT). . . . .	80
5.7	<b>Scenario 1 (Method 2)</b> – Estimation error of the normal vector estimate represented by $1 - \hat{\eta}^\top \eta$ . . . . .	81
5.8	<b>Scenario 2 (Method 1)</b> – Estimated and real attitudes represented by roll, pitch and yaw Euler angles ( $deg$ ) and the attitude estimate error versus time ( $s$ ) (LEFT) and their zooms (RIGHT). . . . .	82
5.9	<b>Scenario 2 (Method 1)</b> – – Estimated and real scaled position and norm of the scaled position error versus time ( $s$ ) (LEFT) and their zooms (RIGHT). . . . .	82
5.10	<b>Scenario 2 (Method 1)</b> – Estimation error of the normal vector estimate represented by $1 - \hat{\eta}^\top \eta$ . . . . .	83
5.11	<b>Scenario 2 (Method 2)</b> – Estimated and real attitudes represented by roll, pitch and yaw Euler angles ( $deg$ ) and the attitude estimate error versus time ( $s$ ) (LEFT) and their zooms (RIGHT). . . . .	83
5.12	<b>Scenario 2 (Method 2)</b> — Estimated and real scaled position and norm of the scaled position error versus time ( $s$ ) (LEFT) and their zooms (RIGHT). . . . .	84
5.13	<b>Scenario 2 (Method 2)</b> – Estimation error of the normal vector estimate represented by $1 - \hat{\eta}^\top \eta$ . . . . .	84

---

5.14	Experimental setup consisting of a Camera-IMU system looking at a textured planar target and OptiTrack markers for ground-truth data. . . . .	86
5.15	Attitudes estimated by Method 2 (Red) and by Method 3 (Green), and corresponding ground-truth attitude (Blue) versus time. . . . .	87
5.16	Scaled position estimated by Method 2 (Red) and by Method 3 (Green), and corresponding ground-truth attitude (Blue) versus time. . . . .	88
5.17	Normal vector estimated by Method 2 (Red) and by Method 3 (Green). . .	89
5.18	Zoom of the estimated attitude, scaled position estimates and the normal vector from time $t = 110s$ to $t = 140s$ . . . . .	90
5.19	Zoom of the estimated attitude, scaled position estimates and the normal vector from time $t = 195s$ to $t = 225s$ . . . . .	90
6.1	Flowchart of the Homography Estimation Algorithm . . . . .	93
6.2	(a) Nvidia Jetson TX1 (b) Nvidia Jetson Xavier AGX (c) Nvidia Jetson Xavier NX . . . . .	94
6.3	(a) Platform 1 (b) Platform 2 . . . . .	95
6.4	Software architecture using ROS as middleware . . . . .	96
6.5	ROS nodes and topics in I3S Architecture . . . . .	98

# 1

## Motivations, contributions and thesis structure

### 1.1 Motivations and contributions

Navigation of Autonomous vehicles in an unknown or partially known and dynamically changing environment represents a great challenge. In case of aerial vehicles or underwater autonomous vehicles, the scientific issues are strongly linked to the fact that the vehicle may have to navigate in close proximity to the physical environment or the sea bed where often the GPS signals (in case of aerial vehicles) and remote acoustic positioning systems (in case of underwater vehicles) are sometimes unavailable or insufficiently precise for safe navigation. In this case, the robot must rely on exteroceptive sensors associated with inertial sensors and specific sensor-based guidance strategies. The critical issue here is to retrieve the state of the vehicle relative to its environment via the use of sensors that do not measure this state directly. In case of underwater scenarios, two popular solutions for detecting the environment are on-board sonars and optical sensors. Although the range of a camera is significantly shorter than that of acoustic sensors and is highly dependent on the turbidity of the water, cameras offer other definite

advantages. For example, they are considerably less expensive than acoustic sensors, they provide rich information at a high rate, and they do not cause interference problems with other instruments. Similarly for aerial vehicles, combining a vision system along with an IMU has led to the development of integrated observers that exploit the optical flow measurements and IMU readings for attitude estimation. These solutions are particularly important in GPS denied environments or in applications such as hazard detection and surveillance, inspection of infrastructures etc. Safe and precise navigation remains very difficult and advanced vision-based navigation strategies are particularly relevant in this context even if they involve difficult problems to solve, both theoretical and experimental. Progress in this direction will undoubtedly have a significant impact on the development of applications related to the navigation of autonomous vehicles in a congested environment.

At the level of underwater vision, the problems are diverse. A large number of factors specific to the fact that the cameras are submerged disturb the processing conventionally carried out in terrestrial vision. We can mention the problem related to the lighting of the stage by artificial light sources. Image processing techniques are therefore necessary to compensate for light halo phenomena. The problem can be even more critical in the case of stereo vision. Indeed, the source of illumination often induces differences in illumination between the two images and therefore difficulties in matching the characteristic points between them. Underwater light attenuation models can be used to improve visibility in underwater scenes. On the other hand, the quality of image processing is very dependent on the quality of the water. Indeed, turbid water can make most algorithms completely inoperative (feature point extractions, target tracking, etc.). It is therefore necessary to provide robust treatments for this type of disturbance. Likewise when the machine is close to the surface, the effects of light on the ripples on the surface introduce moving reflections on the submerged structures. The movement of these reflections is unrelated to the robot's movement and also disrupts movement estimates or feature point pairings. It is therefore necessary to provide robust treatments for this type of disturbance.

The work of this thesis proposal is specifically dedicated to the development of state estimation algorithms which can be used as a basis to develop control strategies for autonomous vehicles. These estimation algorithms merge measurements provided by a monocular camera and more conventional inertial sensors such as inertial unit, depth gauge, magnetometer, etc. Our preferred approach relies on nonlinear observers exploiting non-minimal representations that emphasize the invariance properties of the system. Throughout the entire work carried out in the thesis, the assumption on the planarity of the visual target has been maintained.

This PhD project has been carried out by keeping in mind a more global picture of

the I3S-OSCAR team that aims to develop generic observers by fusing the data from images and inertial sensors for attitude, velocity estimation etc. and whose applications will not only be restricted to aerial vehicles but can also be applied to any robotic vehicle equipped with these sensors. Keeping this global picture in mind, my PhD project focuses on the following two challenging research topics:

- 1. Homography Estimation and its Applications:** In the field of Computer vision, Homography represents an invertible mapping between two views of the same planar surface. Essentially, this homography matrix encodes the camera pose, the distance between the camera and the scene, along with the normal direction to the scene. Although homography estimation is a relatively mature topic, it has been used in various computer vision and robotic applications where the scenarios involve man-made environments composed of (near) planar surfaces. In [71] the authors exploited the homography matrix to estimate the rigid-body pose of a robot equipped with a camera. One of the most successful visual servo control paradigms makes use of homographies [52]. Homography sequences have also been used for the navigation of robotic vehicles [64]. Navigation strategies based on homography are also well suited for applications where the camera is sufficiently far from the observed scene. This is particularly the case when ground images are taken from aerial vehicles [13, 69, 64]. Traditional algorithms for homography estimation rely on algebraic approaches by computing the homography on a frame to frame basis and thus, were not focused on improving the homography over time. This inspired the members of the I3S-OSCAR team to develop non-linear observers based on the underlying structure of the Special Linear group  $SL(3)$ , by incorporating the velocity information across a sequence of images thereby improving the homography estimates over time [11, 49]. This approach however still requires computation of individual image homographies thus making it computationally expensive. In order to overcome this problem, a new observer was developed that directly uses point correspondences without requiring the prior reconstruction of the individual homographies. Inspired from these works, we developed a novel approach for homography observer design on the Special Linear group  $SL(3)$  as well. However, the novelty of this work lies in the linearization approach undertaken to linearize a nonlinear observer on  $SL(3)$ , and what makes this observer extremely effective is its simplicity thus making it suitable for real-time implementation.
- 2. Development of Observers based on Riccati design framework:** As discussed previously homography estimation can be used for a wide range of applications such as control of aerial vehicles, pose estimation etc. where the observed scene

is planar or nearly planar. In absence of an external localization system one of the main problems faced was to estimate the attitude of the robot relative to its environment. In order to overcome this problem, researchers started developing velocity-aided attitude observers by fusing IMU measurements along with linear velocity measurements. However, due to the lack of onboard linear velocity sensors for mini UAV's and the highly expensive DVL (Doppler velocity log) for AUV's researchers started looking into the vehicle's dynamic equations instead. In previous examples we saw that most of the applications involve physical interaction of the robot with the environment. This eventually led to the development of integrated observers that use a vision based system along with an imu, which exploit optical flow measurements and the imu readings to estimate the linear velocity, attitude and distance to the planar target. Traditional approaches make use of the Extended Kalman filters and show the practical convergence of the estimation errors experimentally. However, observability and convergence analyses are either missing or incomplete in these studies. An alternative technique is to use deterministic approaches like the Riccati observer design framework in [24]. My work on this topic has been inspired from the Riccati observer design framework mentioned previously in order to address the following two problems:

- **Designing an observer to estimate the linear velocity, attitude and depth observing a planar target:** The main objective of this work was to estimate these quantities without relying on the strong assumption of an IMU providing measurements of the vehicle's linear acceleration expressed in the body-fixed frame [18, 17]. In this work, we provide a detailed observability analysis that points out the camera's motion excitation conditions whose satisfaction grants stability of the observer and convergence of the estimation errors to zero. At the end, we also provide an extension to the observer design for the filtering of optical flow along with experimental results.
- **Designing two observers for tackling the classical problem of Homography Decomposition:** The main motivation behind this work was to find an alternative solution to the traditional algebraic approaches that solve the Homography Decomposition problem on a frame-by-frame basis thus providing noisy estimates. Also, traditional algebraic approaches fail to obtain the correct solution when the camera translation is really small. In our work, we have proposed "persistence of excitation" conditions which cover such degenerate situations in which the traditional algebraic approaches fail to obtain a correct solution.

Most of the theoretical contributions and experimental validation results reported in

this thesis were published in (or submitted to) the following research papers:

- **Manerikar, N.**, Hua, M.-D. and Hamel, T. (2018). *Homography Observer Design on Special Linear Group  $SL(3)$  with Application to Optical Flow Estimation*. In proceedings of European Control Conference (ECC'18), pp. 1–5, Limassol ,Cyprus. [55]
- Hua, M.-D., **Manerikar, N.**, Hamel, T., and Samson, C. (2018). *Attitude, Linear Velocity and Depth Estimation of a Camera Observing a Planar Target Using Continuous Homography and Inertial Data..* In IEEE International Conference on Robotics and Automation (ICRA'18), pp. 1429–1435, Brisbane, Australia. [30]
- **Manerikar, N.**, Hua, M.-D., Hamel, T. and De. Marco, S. (2020). *Riccati observer design for homography decomposition*. In proceedings of European Control Conference (ECC'18), pp. 1306–1311, Saint Petersburg, Russia. [54]
- **Manerikar, N.**, Hua, M.-D., and Hamel, T. (2021). *Deterministic Riccati observers for homography decomposition*. Preparing for Submission.

In addition, I am one of the principal contributors to the development of the **HomographyLab library** that has been evaluated at the TRL 7 (Technology Readiness Level) and is protected by the French APP (Agency for the Protection of Programs).

## 1.2 Thesis structure

The present thesis is organized in two parts and partitioned in five chapters.

- **Chapter 1 - Motivations, contributions and thesis structure.** As dedicated in the title, this chapter first briefly presents the motivations and objectives of this thesis work. The main contributions are then provided. The thesis structure section briefly introduces the content of all chapters.
- **Chapter 2 - Theoretical recalls on State Estimation.** This chapter is devoted to recall some basic notions related to the topic of State Estimation. We start by talking about the classic state estimation filters for linear and nonlinear systems and discuss about their observability properties that essentially characterize the behaviour of such systems. Then, the Deterministic Riccati observer Design framework has been detailed which will be used as the main brick for the development of novel observers for state estimation during the remainder of the thesis.
- **Chapter 3 - Introduction to Homography and Continuous Homography.** This chapter recalls the preliminary details regarding homography and continuous homography, and then discusses about some relevant homography estimation algorithms. Then a novel linear approach for homography estimation on  $SL(3)$  has

been proposed with application to optical flow estimation. The problem formulation and basic ideas of observer design are first presented. Finally, the performance and robustness of the proposed observer are then verified by performing experiments using a Camera-IMU system.

- **Chapter 4 - Riccati observers for state estimation exploiting optical flow and IMU measurements.** This chapter revisits the problem of partial attitude, linear velocity and depth estimation of an IMU-Camera with respect to a planar target. The considered solution relies on the measurement of the optical flow (extracted from the continuous homography) complemented with gyrometer and accelerometer measurements. The proposed deterministic observer is accompanied with an observability analysis that points out camera's motion excitation conditions whose satisfaction grants stability of the observer and convergence of the estimation errors to zero. The performance of the observer is illustrated by performing experiments on a test-bed IMU-Camera system. Finally we provide an extension to observer design for the filtering of optical flow measurements along with experimental results.
- **Chapter 5 - Homography Decomposition.** This chapter talks about the classical problem of Homography Decomposition and the state of the art approaches for Homography Decomposition. Then two novel nonlinear Riccati observers for the decomposition of the homography and its inverse have been proposed after which a rigorous observability and convergence analysis has been carried out. The large domain of convergence and good performance of the proposed observers have then been demonstrated through both simulation results and extensive experimental validations.
- **Chapter 6 - Software Implementation Details.** The software architecture design for the HomographyLab library has been detailed in this chapter. This chapter also provides an example of the overall software architecture that is used for the I3S-AUV platforms that consists of the vision brick as well as the control brick used for experimental validations. In the context of this thesis, I have worked only on the vision brick.

# 2

## Theoretical recalls on State Estimation

### 2.1 Introduction

In control theory, the state space representation of the physical system is a mathematical model which consists of a set of inputs, outputs and state variables that are related to each other by first order differential equations. The values of these state variables evolve over time in a way that depends on the inputs to which it is subjected. The output values of the physical system thus depends on the values of these state variables. The most general state space representation of a linear time-invariant system can be expressed by the following equation

$$\begin{cases} \dot{x}(t) = A(t)x + B(t)u \\ y(t) = C(t)x + Du(t) \end{cases} \quad (2.1)$$

with  $x \in \mathbb{R}^n$  the system state vector,  $u \in \mathbb{R}^s$  the system input vector, and  $y \in \mathbb{R}^m$  the system output vector and  $A \in \mathbb{R}^{n \times n}$ ,  $B \in \mathbb{R}^{n \times s}$ ,  $C \in \mathbb{R}^{m \times n}$ ,  $D \in \mathbb{R}^{m \times s}$  denoting constant matrices of the physical system with adequate dimensions. The matrix  $D$  is the feed forward matrix and is considered to be zero when the system model doesn't have a direct

feed through. This is the case in most of the systems and that's why we will neglect this term throughout the remainder of the chapter and use the following representation

$$\begin{cases} \dot{x}(t) &= A(t)x + B(t)u \\ y(t) &= C(t)x \end{cases} \quad (2.2)$$

In case of non-linear systems the state space representation has the following generic form

$$\begin{cases} \dot{x}(t) &= f(x(t), u(t), t) \\ y(t) &= h(x(t), u(t), t) \end{cases} \quad (2.3)$$

where  $t$  denotes the time, the system state  $x \in \mathbb{X} \subseteq \mathbb{R}^n$ , the system input vector  $u \in \mathbb{U} \subseteq \mathbb{R}^s$ , and the system output vector  $y \in \mathbb{Y} \subseteq \mathbb{R}^m$ ,  $f(x(t), u(t), t)$  and  $h(x(t), u(t), t)$  are vector functions with proper dimensions.

In general, all the state variables of the system cannot be completely measured. This is because of the fact that most of the times its practically not feasible due to space limitations or cost restrictions to mount all the necessary sensors that are needed to measure the states. And thus for the vast majority of the systems, it is often considered that the dimension of the state vector is greater than that of the output vector ( $m < n$ ) meaning that at some time instance  $t$ , the state vector  $x(t)$  cannot be completely measured or deduced from the outputs. Thus in such a case, it is necessary to calculate or estimate the values of these state variables by using only the input/output relation and the outputs of the system at the starting time. This property of inferring the internal state of the system by knowledge of its outputs is referred to as Observability. The next section provides a detailed description of the observability of systems.

## 2.2 Observability of Systems

The concept of observability studies the possibility of estimating the state from the output. This was first introduced by Rudolph Kalman for the analysis of linear time-invariant (LTI) systems [39], and since then on the observability property has been extensively studied and extended for other classes such as linear time-varying (LTV) and nonlinear systems.

### 2.2.1 Observability of Linear Time Invariant Systems

The observability of a system ensures the reconstruction of the initial state from the knowledge of the inputs and outputs available on a time interval. A system is said to be observable if from the measurements of inputs and outputs we can reconstruct the initial state of the system. For a linear system described by (2.2), there are two fundamental

## 2.2. Observability of Systems

---

notions that essentially characterize the behaviour of this system. These notions are Distinguishability and Observability. The standard concept of distinguishability can be explained by the following definition

**Definition 1.** (*Distinguishability*): The two events  $(x_0, t_0)$  and  $(w_0, t_0)$  are said to be distinguishable [45] on the interval  $[t_0, t_0 + T]$  if there exists  $u \in \mathbb{R}^s$  such that the corresponding outputs  $y(t, x_0) \neq y(t, w_0)$  over the time interval  $[t_0, t_0 + T]$ .

**Definition 2.** (*Observability*): The system (2.2) is said to be observable [45] at  $x_0$  if  $x_0$  is distinguishable from any  $x \in \mathbb{R}^n$ , and is fully observable if  $\forall x_0 \in \mathbb{R}^n$ ,  $x_0$  is distinguishable.

Since for linear systems the observability is independent of the input  $u$ , the matrix  $B$  does not intervene in the observability analysis and it is dependent only on the matrices  $A$  and  $C$ . Thus we can either say that the "the pair  $(A, C)$  is observable", or that the "system is observable".

**Lemma 1.** (*Kalman's Rank Condition*): For a LTI system if the observability matrix is defined by

$$\mathcal{O} := \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \quad (2.4)$$

and the observability rank is defined by  $\text{rank}(\mathcal{O}) = k_{\mathcal{O}}$ , then the pair  $(A, C)$  is fully observable if and only if  $\mathcal{O}$  is of full rank i.e.,  $k_{\mathcal{O}} = n$

**Corollary 1.** The system (2.2) is observable in time  $t$  if and only if the matrix

$$W(t) = \int_0^t e^{-\tau A^T} C^T C e^{-\tau A} d\tau \quad (2.5)$$

is invertible.

Thus the state equation stated in (2.2) is said to be observable if for any unknown initial state  $x(0)$ , there exists a finite  $t > 0$  such that by having the knowledge of the input and the output  $y$  over the time interval  $[0, t]$  it is sufficient to determine uniquely the initial state  $x(0)$ . Otherwise, the equation is said to be unobservable.

### 2.2.2 Observability of Linear Time Varying Systems

In this part we talk about the observability properties for another class of systems specifically the linear time varying systems. Consider a general Linear Time Varying system

(LTV) given by

$$\begin{cases} \dot{x}(t) &= A(t)x + B(t)u \\ y(t) &= C(t)x \end{cases} \quad (2.6)$$

There exist different types of observability properties for LTV systems, for e.g., instantaneous, or uniform observability. All these properties are detailed in [8], but have been briefly mentioned below to recall them to facilitate the reader. Also for LTV systems, the following assumption holds:

**Assumption 1.** *The matrix-valued functions  $A$ ,  $B$ , and  $C$  of the LTV system (2.6) are continuous and bounded on  $[0, +\infty]$ .*

For instance, [8] classifies the observability of a linear time-varying system according to the following definitions.

**Definition 3.** *(Instantaneous observability) A system is instantaneously observable if,  $\forall t$ , the state  $x(t)$  can be computed from the inputs  $u(t)$ , outputs  $y(t)$  and time derivatives  $u^k(t)$ ,  $y^k(t)$  with  $k \leq n + 1$ .*

**Lemma 2.** *Define the observation space of the LTV system at the time-instant  $t$  as the space generated by*

$$\mathcal{O}(t) := \begin{bmatrix} N_0(t) \\ N_1(t) \\ \vdots \end{bmatrix} \quad (2.7)$$

with  $N_0(t) = C$ ,  $N_{k+1} = N_k A + \dot{N}_k$ ,  $k = 1, \dots$ . Then, system (2.6) is instantaneously observable if  $\text{rank}(\mathcal{O}) = n$

**Definition 4.** *(Uniform observability) System (2.6) is uniformly observable if there exist  $\delta, \mu > 0$  such that (s.t.)  $\forall t \geq 0$*

$$W(t, t + \delta) \triangleq \frac{1}{\delta} \int_t^{t+\delta} \Phi^\top(t, \tau) C^\top(\tau) C(\tau) \Phi(t, \tau) d\tau \geq \bar{\mu} I_d \quad (2.8)$$

with  $W(t, t + \delta)$  is called the Observability Gramian of the system and  $\Phi(t, \tau)$  the transition matrix associated with  $A(t)$ , i.e. such that  $\frac{d}{dt} \Phi(t, \tau) = A(t) \Phi(t, \tau)$  with  $\Phi(t, t) = I_n$ .

Define  $N_0(t) = C$ ,  $N_{k+1} = N_k A + \dot{N}_k$ ,  $k = 1, \dots$  and the set  $\mathcal{M}_K$  of matrix-valued functions  $M(\cdot)$  of dimension  $(q \times n)$  ( $q \geq 1$ ) composed of row vectors of  $N_0(\cdot)$ ,  $N_1(\cdot)$ ,  $\dots$ . By replacing the matrix  $C$  by  $M$  the following lemma has been proposed in [22],

**Lemma 3.** *(See [22]) The existence of a matrix  $M \in \mathcal{M}_K$  satisfying the following property*

$$W(t, t + \delta) \triangleq \frac{1}{\delta} \int_t^{t+\delta} \Phi^\top(t, \tau) M^\top(\tau) M(\tau) \Phi(t, \tau) d\tau \geq \bar{\mu} I_d \quad (2.9)$$

## 2.3. Observer Design

---

implies the satisfaction of (2.6), and thus uniform observability of the corresponding LTV system.

We can also say that the pair  $A(t), C(t)$  is uniformly observable when the Lemma 2.9 is satisfied. Since the calculation of the Gramian requires the integration of the solutions of  $\dot{x} = A(t)x$ , checking the uniform observability of the LTV systems can be a difficult task. Hence in general, the uniform observability for LTV systems cannot be characterized by only rank conditions. The following lemma establishes [58] establishes a sufficient condition for uniform observability

**Lemma 4.** *If there exists a matrix  $M \in \mathcal{M}_K$  such that for some positive numbers  $\bar{\delta}, \bar{\mu}$  and  $\forall t \geq 0$*

$$\frac{1}{\bar{\delta}} \int_t^{t+\bar{\delta}} \det(M^\top(\tau)M(\tau))d\tau \geq \bar{\mu} \quad (2.10)$$

The following lemma, taken from [22], gives a sufficient condition for uniform observability in terms of the properties of the matrices  $A(t)$  and  $C(t)$  and their time-derivatives:

**Lemma 5.** *(See [22]) If  $A$  is a constant matrix with real eigenvalues, and there exists  $M \in \mathcal{M}_K$  such that*

$$\frac{1}{\bar{\delta}} \int_t^{t+\bar{\delta}} M^\top(\tau)M(\tau)d\tau \geq \bar{\mu}I_d > 0 \quad (2.11)$$

*then the observability Gramian of System (2.6) satisfies condition (2.11)*

Definition 4 and the Lemma 4 are exploited for proving the uniform observability of the systems in Chapter 4 and Chapter 5, respectively.

## 2.3 Observer Design

### 2.3.1 Observer Design for Linear System

Many methods have been proposed to estimate the state of a linear dynamical system. Luenberger [47] was one of the first ones to develop deterministic state observers for continuous linear systems. Since his works, a notable amount of research has been devoted to the problem of observer design for linear systems. The main developments are detailed in [63], [2], [74] and, in the recent books [43] and [75] and the references therein. Kalman [39] also formulated an observer by considering a deterministic or stochastic linear system. For the system stated in (2.6) the observer equations can be written as

$$\begin{cases} \dot{\hat{x}}(t) &= A(t)\hat{x} + B(t)u(t) + K(t)(y(t) - \hat{y}(t)) \\ \hat{y}(t) &= C(t)\hat{x} \end{cases} \quad (2.12)$$

where  $\hat{x}(t)$  is the estimate of the state  $x(t)$  and  $K$  is the observer gain related to the error of the output reconstruction. Then the observer error  $e = x - \hat{x}$  satisfies the equation

$$\dot{e} = (A - KC)e \quad (2.13)$$

The observer gain  $K$  can then be chosen appropriately so that the observer error  $e$  converges to zero when  $t \rightarrow \infty$ . When the observer gain  $K$  is high, the linear Luenberger observer converges to the system states very quickly and vice-versa.

### 2.3.1.1 Kalman Filter

The Kalman filter (KF) is one of the most widely used methods for estimation due to its simplicity, optimality and robustness. Since the Kalman filter is a recursive estimator it means that it takes into account the estimated state from the previous time step and the current measurement in order to compute the estimate of the current state. Typically a discrete version of the Kalman filter is divided into two phases: "Predict" and "Update". In the Prediction phase, it uses the state estimate from the previous timestep to produce an estimate of the state at the current timestep while in the Update phase it uses the information from the current timestep to refine the state estimate at that instant. For the Kalman filter in case of stochastic systems, one of the important things is to compute the optimal gain  $K$  that minimizes the residual error and is given by the following expression

$$K = PC^T D(t) \quad (2.14)$$

where  $P(t)$  is the solution to the Continuous Riccati Equation (CRE)

$$\dot{P} = AP + PA^T - PC^T D(t)CP + S(t) \quad (2.15)$$

where  $D$  and  $S$  are positive definite matrices that represent the covariance matrices of the state and measurement noises and are assumed to be Gaussian.

### 2.3.2 Observers for Nonlinear Systems

For the nonlinear system expressed in (2.3) the observer can be expressed in the generic form shown below:

$$\begin{cases} \dot{\hat{x}} = f(\hat{x}, u) - K(\hat{x}, u(t), t) \cdot (h(x, u(t)) - h(\hat{x}, u(t))) \\ \hat{y} = h(\hat{x}, u) \end{cases} \quad (2.16)$$

Thus the estimation error for the nonlinear system is given by

$$\dot{e} = f(x, u(t)) - f(\hat{x}, u(t)) + K(\hat{x}, u(t), t) \cdot (h(x, u(t)) - h(\hat{x}, u(t))) \quad (2.17)$$

### 2.3. Observer Design

---

Since in this case  $f$  and  $h$  are nonlinear, the problem of finding a gain matrix  $K$  such that error converges asymptotically to zero cannot be solved like in the linear case. Thus one of the solutions in order to solve this problem is to linearize the system around an equilibrium point  $(\bar{x}, \bar{u}, \bar{y})$  characterized by  $f(\bar{x}, \bar{u}) = 0$  and  $\bar{y} = h(\bar{x}, \bar{u})$ . Considering  $\delta_x = x - \bar{x}$ ,  $\delta_u = u - \bar{u}$  and  $\delta_y = y - \bar{y}$ , the linearized model is given by

$$\begin{cases} \delta\dot{x}(t) &= A\delta x(t) + B\delta u(t) \\ \delta y(t) &= C\delta x(t) \end{cases} \quad (2.18)$$

where  $A = \frac{\partial f}{\partial x}(\hat{x}, u)$ ,  $C = \frac{\partial h}{\partial x}(\hat{x}, u)$

#### 2.3.2.1 Extended Kalman Filter

The Extended Kalman filter is probably the most widely used estimator for nonlinear systems. The Extended Kalman filter applies the Kalman filter to nonlinear systems by simply linearising all the nonlinear models so that the traditional linear Kalman filter equations can be applied. Basically a matrix of partial derivatives i.e the Jacobian is computed at each time step and evaluated with the current predicted states. This procedure essentially linearizes the nonlinear function around the current estimate. Thus the idea of the EKF is to replace the state and output matrices  $A$  and  $C$  of the linear system by the Jacobian of the nonlinearities of the system. In this case as well, the gain matrix  $K$  is computed as

$$\begin{aligned} K &= PC^T D \\ \dot{P} &= AP + PA^T - PC^T DCP + S \end{aligned} \quad (2.19)$$

with  $A = \frac{\partial f}{\partial x}(\hat{x}, u)$ ,  $C = \frac{\partial h}{\partial x}(\hat{x}, u)$ .

In case of the EKF, when the models are highly nonlinear, the first order linearization can introduce large errors that can lead to sub-optimal performance and sometimes divergence of the filter. To this end, Julien and Uhlman [38] proposed the Unscented Kalman Filter (ULF) which uses a deterministic sampling approach. Thus, instead of linearizing which is the case in EKF, it specifies the Gaussian state distribution using a set of points, referred to as sigma points, and propagates them through the true nonlinear system. It is able to capture the posterior mean and covariance accurately to the third order (Taylor series expansion) for any nonlinearity.

From all the works mentioned above it is clear that nonlinear observers have increasingly become alternative solutions to the classical filtering techniques such as Extended Kalman filters, Unscented Kalman filters, particle filters etc mainly because of their simplicity and large domain of stability and convergence. In the next section we talk about the Deterministic Riccati observer design framework that relies on the solutions to the Continuous Riccati Equation (CRE) and encompasses the EKF solutions.

### 2.3.2.2 Deterministic Riccati Observers

In this subsection we provide details about the deterministic Riccati observer framework developed in [24], which exploit first-order approximations of a class of nonlinear systems, and may formally be viewed as a generalisation of the so-called multiplicative extended Kalman filter (MEKF) [42], applies to a number of applications involving proprioceptive sensors and monocular vision. The observers proposed in Chapter 4 and Chapter 5 are based on this framework. We modify slightly the notations provided in [24], in order to be coherent with the rest of the work carried out in the thesis. The following nonlinear system (a particular case studied in [24]) is investigated:

$$\begin{cases} \dot{x} = A(x, t)x + U + O(|x|^2) + O(|x||U|) \\ y = C(x, t)x + O(|x|^2) \end{cases} \quad (2.20)$$

with state  $x = [x_1^\top, x_2^\top]^\top$ ,  $x_1 \in \mathcal{B}_r^{n_1}$  (the closed ball in  $\mathbb{R}^{n_1}$  of radius  $r$ ),  $x_2 \in \mathbb{R}^{n_2}$ , output  $y \in \mathbb{R}^m$ ,  $C(x, t) \in \mathbb{R}^{m \times (n_1+n_2)}$  a continuous matrix-valued function uniformly continuous with respect to (w.r.t.)  $x$  and uniformly bounded w.r.t.  $t$ , and  $A(x, t)$  in the form

$$A(x, t) = \begin{bmatrix} A_{1,1}(t) & 0_{n_1 \times n_2} \\ A_{2,1}(x, t) & A_{2,2}(x_1, t) \end{bmatrix} \quad (2.21)$$

with  $A_{1,1}(t)$  and  $A_{2,2}(x_1, t)$  continuous matrix-valued functions uniformly bounded w.r.t.  $t$  and uniformly continuous w.r.t.  $x$ , and  $A_{2,1}(x, t)$  satisfying

$$A_{2,1}(x, t) = \bar{A}_{2,1}(x_1, t)x_1 + O(|x_1||x_2|)$$

with  $\bar{A}_{2,1}(x_1, t)$  a continuous matrix-valued function uniformly bounded w.r.t.  $t$  and uniformly continuous w.r.t.  $x_1$ .

Then, apply the input

$$U = -PC^\top Dy \quad (2.22)$$

with  $P \in \mathbb{R}^{(n_1+n_2) \times (n_1+n_2)}$  a symmetric positive definite matrix solution to the following continuous Riccati equation (CRE):

$$\dot{P} = AP + PA^\top - PC^\top D(t)CP + S(t) \quad (2.23)$$

with  $P(0) \in \mathbb{R}^{(n_1+n_2) \times (n_1+n_2)}$  a symmetric positive definite matrix,  $D(t) \in \mathbb{R}^{m \times m}$  bounded continuous symmetric positive semi-definite, and  $S(t) \in \mathbb{R}^{(n_1+n_2) \times (n_1+n_2)}$  bounded continuous symmetric positive definite.

Then, from Theorem 3.1 and Corollary 3.2 in [24],  $x = 0$  is locally exponentially stable (LES) when both matrices  $D(t)$  and  $S(t)$  are larger than some constant positive

### 2.3. Observer Design

---

matrix and the pair  $(A^*(t), C^*(t))$ , with  $A^*(t) \triangleq A(0, t)$ ,  $C^*(t) \triangleq C(0, t)$ , is uniformly observable.

# 3

## Introduction to Homography and Continuous Homography

### 3.1 Introduction

In this chapter, we begin by unveiling the basic geometrical concepts that relates two views of the camera while observing a planar scene. This chapter serves to introduce the basic building blocks for the estimation of Homography. Then we talk about some existing feature-based approaches for homography estimation with a particular focus on nonlinear homography observers on  $SL(3)$  where a notable contribution has been made by the members of the I3S-OSCAR team [34]. Finally, we propose a new observer that addresses the same problem as [21] but follows a different approach by directly exploiting the basis of the Lie algebra of the group  $SL(3)$ . 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. Also, the proposed observer is extremely effective due to its simplicity thus making it ideal for real-time implementation.

## 3.2. Preliminary material for Homography Estimation

---

This chapter is divided into 3 sections. Section 3.2 gives a brief description of the notation and the math related to homography. In Section 3.3, various feature-based approaches (algebraic as well as geometric) for homography estimation have been discussed. In Section 3.4 a novel linear approach for observer design on  $SL(3)$  is proposed using point correspondences and the knowledge of the group velocity. In Section 3.6 the computation of the optical flow estimate extracted from continuous homography as well as experimental results supporting the proposed approach are presented.

## 3.2 Preliminary material for Homography Estimation

### 3.2.1 Perspective Projection

Visual information about the environment is obtained by projecting an observed scene onto the camera image surface. Typically two types of parameters are needed in order to reconstruct the 3D structure of a scene: intrinsic ("internal" parameters of the camera such as the principal point, the pixel aspect ratio, focal length etc.) and extrinsic (parameters that defined the location and orientation of the reference camera frame w.r.t a known world reference frame).

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  $R \in SO(3)$  (see Fig. 3.1). The coordinates of a point in the reference frame ( $\mathring{P} \in \{\mathring{A}\}$ ) are related to its coordinates in the current frame ( $P \in \{A\}$ ) using the relation

$$\mathring{P} = RP + \xi \quad (3.1)$$

In the commonly used approximation [25], the intrinsic camera parameters define a  $3 \times 3$  matrix  $K$  so that one can write

$$\mathring{p}^{im} \cong K\mathring{P}, \quad p^{im} \cong KP \quad (3.2)$$

where  $p^{im} \in \mathcal{A}$  denotes the image of the considered point when the camera is aligned with the frame  $\{A\}$  and  $\mathring{p}^{im} \in \mathring{\mathcal{A}}$  denotes the image of the considered point when the camera is aligned with the frame  $\{\mathring{A}\}$ . Both these points have the form  $(u, v, 1)^T$  using the homogeneous coordinate representation. If the intrinsic parameters of the camera

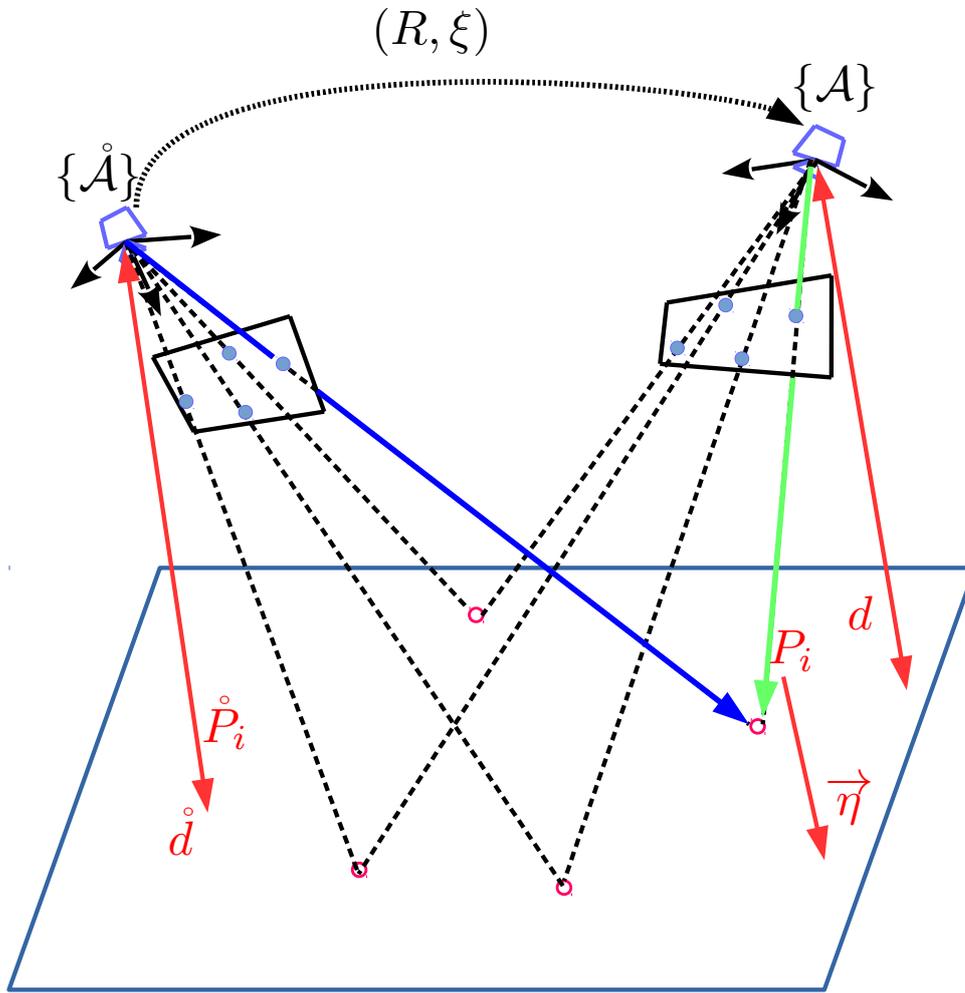


Figure 3.1: Euclidean homography relating the camera's pose, the distance to the plane and the plane's normal vector by  $H = R + \frac{1}{d}\xi\eta^\top$ .

are known meaning that the camera is calibrated one can write <sup>1</sup>:

$$\hat{p} = \frac{K^{-1}\hat{p}^{im}}{|K^{-1}\hat{p}^{im}|}, \quad p = \frac{K^{-1}p^{im}}{|K^{-1}p^{im}|} \quad (3.3)$$

### 3.2.2 Homographies

In the field of Computer Vision, the so called homography is an invertible mapping that relates two camera views of the same planar scene by encoding in a single matrix the

<sup>1</sup>Most statements in projective geometry involve equality up to a multiplicative constant denoted by  $\cong$

### 3.2. Preliminary material for Homography Estimation

---

camera pose, the distance between the camera and the scene, along with the normal direction to the scene (e.g., [25]). For thoroughly understanding this part, more details about the homography estimation are presented below by borrowing some material from [34]. Let  $\mathring{d}$  (resp.  $d$ ) and  $\mathring{n}$  (resp.  $n$ ) denote the distance from the origin of  $\{\mathring{A}\}$  (resp.  $\{\mathcal{A}\}$ ) to the observed planar scene and the coordinate normal vector pointing towards the scene expressed in  $\{\mathring{A}\}$  (resp.  $\{\mathcal{A}\}$ ), respectively. One can easily verify that

$$\begin{cases} \eta = R^\top \mathring{n} \\ d = \mathring{d} - \mathring{\eta}^\top \xi \end{cases} \quad (3.4)$$

Now by rearranging the terms in (3.1) one obtains

$$P = R^\top (\mathring{P} - \xi) \quad (3.5)$$

Since the target points belong to the same observed planar scene

$$\Pi := \{\forall P \in \mathbb{R}^3 : \eta^\top P - d = 0\} = \{\forall \mathring{P} \in \mathbb{R}^3 : \mathring{\eta}^\top \mathring{P} - \mathring{d} = 0\}$$

one derives from the planar constraint  $\frac{\mathring{\eta}^\top \mathring{P}}{\mathring{d}} = 1$  and (3.5) that the projected point obeys the relation

$$P = R^\top \left( I_3 - \frac{\xi \mathring{\eta}^\top}{\mathring{d}} \right) \mathring{P} \quad (3.6)$$

Using (3.3) and (3.6), the projected point satisfies

$$p \cong R^\top \left( I_3 - \frac{\xi \mathring{\eta}^\top}{\mathring{d}} \right) \mathring{p} \cong H \mathring{p} \quad (3.7)$$

where the projective mapping

$$H := R^\top - R^\top \frac{\xi \mathring{\eta}^\top}{\mathring{d}} \quad (3.8)$$

is defined as the Euclidean homography that maps Euclidean coordinates of the scene's points from  $\{\mathring{A}\}$  to  $\{\mathcal{A}\}$ . Using (3.4) one verifies that

$$H^{-1} \cong R + \frac{\xi n^\top}{d} \quad (3.9)$$

Depending on literature, either  $H$  given by (3.8) or  $H^{-1}$  given by (3.9) is referred to as **Homography**. Since a non-degenerate homography matrix  $H$  (i.e.  $\det(H) \neq 0$ ) is only defined up to a scale factor, it has 8 degrees of freedom while it has 9 entries. An additional constraint is thus required. Several possibilities have been proposed in

literature. For instance, a simple constraint of fixing the third diagonal element of  $H$  equal to 1 (i.e.  $h_{3,3} = 1$ ) is proposed in [25]. Another possibility consists in fixing the Frobenius norm of  $H$  equal to 1 [25]. Finally, as any non-degenerate homography matrix is associated with a unique matrix  $\bar{H} \in \text{SL}(3)$  by re-scaling  $\bar{H} = \det(H)^{-\frac{1}{3}}H$  such that  $\det(\bar{H}) = 1$ , without loss of generality it can be assumed that  $H$  is an element of  $\text{SL}(3)$  as originally proposed in [5].

Recall that the scale factor  $\gamma$  such that  $H = \gamma R^\top \left( I - \frac{\xi \hat{\eta}^\top}{d} \right)$  is equal  $(\hat{d}/d)^{\frac{1}{3}}$  and corresponds to the second singular value of  $H$  [32].

The so-called “image” homography matrix  $H_{img} \in \text{SL}(3)$  that maps pixel coordinates from  $\mathcal{A}$  to  $\mathcal{C}$  (i.e.  $p_{img} \cong H_{img} \hat{p}_{img}$ ) then satisfies  $H_{img} = KHK^{-1}$ .

Expression (3.7) provides the transformation by the homography  $H$  of point-feature correspondences between two image frames. Analogously, one can find the transformation by  $H$  of the correspondences of line features in [35] and conic features (i.e ellipses, hyperbolas) [28].

Homography plays a crucial role in various computer vision and robotics applications where the working environment consists of man made structures that are composed of nearly planar surfaces. In one of the works by [71], the rigid body pose of a robot equipped with a monocular camera was estimated by exploiting the homography matrix. Homography sequences have been also used for the navigation of robotic vehicles [64] and one of the most successful visual servo control paradigms [52] exploits the homography estimation. Apart from the applications mentioned above, homography-based methods can also be used for image stabilization, image mosaicing as well as control of autonomous underwater vehicles.

### 3.2.3 Homographies as elements of the Special Linear Group $\text{SL}(3)$

The special linear group  $\text{SL}(3)$  is defined as the set of all real valued  $3 \times 3$  matrices with unit determinant

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

Since any homography matrix is associated with a unique matrix  $\bar{H} \in \text{SL}(3)$  by re-scaling

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

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

$$w : \text{SL}(3) \times \mathbb{P} \rightarrow \mathbb{P}^2$$

$$(H, p) \mapsto w(H, p) \cong \frac{Hp}{|Hp|}$$

### 3.2. Preliminary material for Homography Estimation

---

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

$$w(H_1, w(H_2, p)) = w(H_1, H_2, p), w(I_3, p) = p, \quad p \in \mathbb{P}^2$$

where  $H_1, H_2$  and  $H_1H_2 \in SL(3)$  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  $SL(3)$ .

For completeness, let us recall on the Lie algebra  $\mathfrak{sl}(3)$  of  $SL(3)$ . It is defined by

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

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

$$\Delta = \sum_{i=1}^8 \delta_i B_i \tag{3.10}$$

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

$$\begin{aligned} B_1 &= e_1 e_2^\top, & B_2 &= e_2 e_1^\top, & B_3 &= e_2 e_3^\top \\ B_4 &= e_3 e_2^\top, & B_5 &= e_3 e_1^\top, & B_6 &= e_1 e_3^\top \\ B_7 &= e_1 e_1^\top - \frac{1}{3} I_3, & B_8 &= e_2 e_2^\top - \frac{1}{3} I_3 \end{aligned}$$

with  $I_3$  the identity element of  $\mathbb{R}^{3 \times 3}$  and  $\{e_1, e_2, e_3\}$  the canonical basis of  $\mathbb{R}^3$ .

#### 3.2.4 Rigid-body and Homography Dynamics

Consider a camera attached to the moving frame  $\{\mathcal{A}\}$  moving with kinematics

$$\begin{cases} \dot{R} &= R\Omega_\times \\ \dot{\xi} &= RV \end{cases} \tag{3.11}$$

viewing a stationary planar scene, where  $\Omega$  and  $V$  are the angular and linear velocities of  $\{\mathcal{A}\}$  with respect to  $\{\mathring{A}\}$  expressed in  $\{\mathcal{A}\}$ , respectively. Then the kinematics of the associated homography matrix  $H \in SL(3)$  are given by

$$\dot{H} = HU \tag{3.12}$$

The group velocity  $U \in \mathfrak{sl}(3)$  induced by the camera motion, and such that the dynamics of  $H$  are in the form (3.12), then satisfies [50, Lem. 5.3]

$$U = \Omega_{\times} + \frac{V\eta^{\top}}{d} - \frac{\eta^{\top}V}{3d}I_3 \quad (3.13)$$

The group velocity  $U$  given by 3.13 is often referred to as ‘‘Continuous Homography’’ in the literature [48].

### 3.3 Existing feature-based homography estimation techniques

Classical algorithms for homography estimation taken from the computer vision community consist of computing the homography on a frame-by-frame basis by solving algebraic constraints related to correspondences of image features (points, lines, conics, contours, etc.) [25, 1, 37, 40, 9]. These algorithms only considered the homography as an incidental variable and were not focused on improving (or filtering) the homography over time. In recent years, advances have been made in homography estimation algorithms by exploiting the temporal correlation of data across a video sequence rather than computing algebraically individual raw homography for each image. Powerful methodologies for nonlinear observer design on Lie groups (e.g. [51]) have been instrumental for the derivation of these algorithms.

A nonlinear observer was proposed in [50] based on the underlying structure of the Special Linear group  $SL(3)$ , which is isomorphic to the group of homographies [5]. Velocity information was exploited to interpolate across a sequence of images and improve the individual homography estimates. The observer, however, still requires individual image homographies (previously computed using an algebraic technique) as the feedback information. Thus, it needed both a classical homography algorithm and a temporal filter algorithm, and only functions if each pair of images provides sufficient features to algebraically compute a raw homography.

In order to overcome these drawbacks, the question of deriving an observer for a sequence of image homographies, which takes image point-feature correspondences directly as input has been considered [21, 34]. The previous observer is extended by also incorporating image line-feature correspondences (in addition to point-feature correspondences) directly as input in the design of observer innovation [33]. In line with this effort, conic-feature correspondences (i.e. non-degenerate second-order features such as ellipses and hyperbolas) are considered for the construction of observer innovation [28]. Without requiring any prior step for reconstruction of individual homographies for feeding the observer innovation, these algorithms are suitable for real-time applications using an embedded computer. In contrast with algebraic techniques, these observers are

### 3.3. Existing feature-based homography estimation techniques

---

also well posed even when there is insufficient data for full reconstruction of a homography. In such situations, these algorithms continue to operate by incorporating available information and relying on propagation of prior estimates. Thereafter, a classical algebraic algorithm and a state-of-the-art nonlinear observer on  $SL(3)$  for homography estimation that exploit the simplest feature correspondences – the point correspondences – are recalled for the purpose of understanding.

#### 3.3.1 A classical algebraic algorithm of homography estimation

Homography estimation is a topic well developed and discussed in classical computer vision books [25, 65]. In this section, the so-called Direct Linear Transformation (DLT) method which employs point correspondences for homography estimation is briefly revisited.

Given a set of four 2D to 2D point correspondences,  $p_i \leftrightarrow \hat{p}_i$ , where  $\hat{p}_i$  (resp.  $\hat{p}_i$ ) is the re-normalized point of  $P_i$  (resp.  $\hat{P}_i$ ), as shown in (3.3). Denote  $[u_i, v_i, w_i]^\top$  coordinates of  $p_i$ . Equation (3.7) implies that  $p_i \times (H\hat{p}_i) = 0$  which in turn yields

$$\begin{bmatrix} 0 & -w_i\hat{p}_i^\top & v_i\hat{p}_i^\top \\ w_i\hat{p}_i^\top & 0 & -u_i\hat{p}_i^\top \\ -v_i\hat{p}_i^\top & u_i\hat{p}_i^\top & 0 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} = 0 \quad (3.14)$$

with  $h_j$  ( $j = 1, 2, 3$ ) the  $j^{th}$  column of  $H$ . Equation (3.14) contains three equations, however only two of them are linearly independent. By omitting, for instance, the third equation, each point correspondence  $p_i \leftrightarrow \hat{p}_i$  gives two equations in the entries of  $H$  as

$$\begin{bmatrix} 0 & -w_i\hat{p}_i^\top & v_i\hat{p}_i^\top \\ w_i\hat{p}_i^\top & 0 & -u_i\hat{p}_i^\top \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} = 0$$

These equations have the form  $L_i h = 0$  where  $L_i$  is a  $2 \times 9$  matrix and  $h = [h_1^\top \ h_2^\top \ h_3^\top]^\top$  the vector of 9 unknown entries of  $H$ . From a set of four point correspondences on the observed plane, a set of 8 equations in form of  $Lh = 0$  is obtained, where  $L$  is the matrix of dimension  $8 \times 9$  obtained by stacking the rows of  $L_i$  contributed from each correspondence. One observes that  $h = 0$  is an obvious solution.

For a set of four consistent points (in the sense that all triplets of these four points are linearly independent),  $L$  has rank 8, and thus with an additional constraint of the norm  $|h| > 0$ , the obvious solution is avoided and  $h$  is defined up to scale. For simplification, one can choose  $|h| = 1$  which is equivalent to having the Frobenius norm of  $H$  equal to 1.

Solving these algebraic equations on a frame-by-frame basis requires computation power. It can only be carried out if the number of point correspondence is not less than 4 and these point correspondences are consistent. Insufficient number of feature correspondence leads to calculation corruption. The above-presented algorithm is the basis of the cv :: findHomography function of OpenCV<sup>2</sup>.

### 3.3.2 Nonlinear homography observer on SL(3)

The homography defined by (3.7) in this work maps Euclidean coordinates of the scene's points from  $\{\mathring{A}\}$  to  $\{\mathcal{A}\}$ . In the reverse direction,  $H := H^{-1}$  mapping Euclidean coordinates of the scene's points from  $\{\mathcal{A}\}$  to  $\{\mathring{A}\}$  satisfies

$$\mathring{p}_i \cong H p_i \quad (3.15)$$

The re-normalized point  $p_i$  is thus given by

$$p_i = \frac{H^{-1} \mathring{p}_i}{|H^{-1} \mathring{p}_i|} \quad (3.16)$$

Now, the basic ideas of observer design proposed in [34], [36] for  $H$  on SL(3) based on direct point correspondence are recalled. To expose the underlying ideas of observer design, in this part we consider the simplified case where the group velocity  $U$  is known. Assume that a set of  $n$  measurements  $p_i = h(H, \mathring{p}_i) \in \mathbb{P}^2, i = \{1 \dots n\}$  in form of (3.16) in the camera frame  $\{\mathring{A}\}$  is available, where  $\mathring{p}_i \in \mathbb{P}^2$  are constant and known.

**Definition 5.** (Consistency) Assume that a set  $\mathcal{M}_n$  of  $n \geq 4$  vector directions  $p_i^* \in \mathbb{P}^2$ , with  $i = \{1 \dots n\}$  contains a subset  $\mathcal{M}_4 \subset \mathcal{M}_n$  of 4 constant vector directions such that all vector triplets in  $\mathcal{M}_4$  are linearly independent. In this case  $\mathcal{M}_n$  is called consistent.

Let  $\hat{H} \in \text{SL}(3)$  denote the estimate of  $H$ . Then the goal is to drive the error term  $\tilde{H} = \hat{H}H^{-1}$  to the identity matrix  $I_3$ . The output errors i.e the estimates  $e_i$  of  $\mathring{p}_i$  are defined as follows,

$$e_i = \frac{\hat{H} p_i}{|\hat{H} p_i|} = \frac{\tilde{H} \mathring{p}_i}{|\tilde{H} \mathring{p}_i|} \quad (3.17)$$

The proposed observer takes the form

$$\dot{\hat{H}} = \hat{H}U + \Delta \hat{H}, \quad \hat{H}(0) \in \text{SL}(3) \quad (3.18)$$

---

<sup>2</sup><https://docs.opencv.org/3.4.1/d9/dab/tutorial-homography.html>

### 3.4. Novel approach for Homography Observer design on SL(3)

---

where  $\Delta \in \mathfrak{sl}(3)$  is the innovation term designed as

$$\Delta = \sum_{i=1}^n \pi_{e_i} \dot{p}_i e_i^\top \quad (3.19)$$

with  $\pi_x := (I_3 - xx^\top)$ ,  $\forall x \in \mathbb{S}^2$ .

Differentiating  $e_i$  from equation (3.17) we obtain

$$\dot{e}_i = k_P \pi_{e_i} \Delta e_i \quad (3.20)$$

For the stability analysis the following Lyapunov function is considered:

$$\mathcal{L} = \sum_{i=1}^n \frac{1}{2} |e_i - \dot{p}_i|^2 \quad (3.21)$$

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

$$\dot{\mathcal{L}} = -k_P \left\| \sum_{i=1}^n e_i \dot{p}_i^\top \pi_{e_i} \right\|^2 \quad (3.22)$$

From the above equation we can see that the derivative is negative semi-definite ensuring that  $\tilde{H}$  is locally bounded and the equilibrium  $\tilde{H} = I_3$  is asymptotically stable [21, Th. 3.2].

For the interested reader, the observer design with partial knowledge of the group velocity has also been discussed in [21].

### 3.4 Novel approach for Homography Observer design on SL(3)

The equation of the proposed homography estimator taking into account system kinematics in (3.11) can be expressed as a kinematic filter system on SL(3) as

$$\dot{\hat{H}} = \hat{H}U + \Delta \hat{H} \quad (3.23)$$

where the innovation term  $\Delta \in \mathfrak{sl}(3)$  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  $p_j = \frac{H^{-1}\dot{p}_j}{|H^{-1}\dot{p}_j|} \in S^2$  ( $j = 1, \dots, n$ ), with  $\dot{p}_j \in S^2$  known and constant. Here  $p_j$  and  $\dot{p}_j$  represent calibrated image points normalized onto the unit sphere and can be

computed as

$$p_j = \frac{K^{-1}p_j^{im}}{|K^{-1}p_j^{im}|}, \quad \hat{p}_j = \frac{K^{-1}\hat{p}_j^{im}}{|K^{-1}\hat{p}_j^{im}|}$$

The output errors  $e_j$  are defined as

$$e_j := \frac{\hat{H}p_j}{|\hat{H}p_j|} = \frac{\tilde{H}\hat{p}_j}{|\tilde{H}\hat{p}_j|} \quad (3.24)$$

which thus can be viewed as the estimates of  $\hat{p}_j$ .

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

$$\dot{\tilde{H}} = \Delta\tilde{H} = \sum_{i=1}^8 \delta_i B_i \tilde{H} \approx \sum_{i=1}^8 \delta_i B_i \quad (3.25)$$

where the linear representation  $\delta \in \mathbb{R}^8$  of  $\Delta$  via the relation (3.10) is used ( $\Delta = \sum_{i=1}^8 \delta_i B_i$ ).

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

$$\tilde{H} = \exp\left(\sum_{i=1}^8 x_i B_i\right) \approx \left(I_3 + \sum_{i=1}^8 x_i B_i\right) \quad (3.26)$$

Using (3.26) one obtains

$$\dot{\tilde{H}} \approx \sum_{i=1}^8 \dot{x}_i B_i \quad (3.27)$$

From (3.25) and (3.27) it is obvious that

$$\dot{x} \approx \delta \quad (3.28)$$

Now we focus on the linearization of the output vectors. From (3.24) and (3.26) the output errors  $e_j$  can be expressed as

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

### 3.4. Novel approach for Homography Observer design on $SL(3)$

---

Neglecting high order terms one gets

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

with  $\pi_y := I_3 - yy^\top, \forall y \in S^2$ , the projection operator on the plane orthogonal to  $y$  and

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

Stacking all  $n$  measurements in a vector as follows

$$y := \begin{bmatrix} e_1 - \hat{p}_1 \\ \vdots \\ e_n - \hat{p}_n \end{bmatrix} \in \mathbb{R}^{3n} \quad (3.29)$$

one obtains

$$y \approx Cx \quad (3.30)$$

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

From here, the innovation term  $\delta$  (i.e.  $\Delta$ ) can be directly designed on the linear approximation system (3.28) using the linear approximation (3.30) 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 [21]), then the design of the innovation term  $\delta$  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 (3.10).

**Proposition 1.** *Assume that the group velocity  $U \in \mathfrak{sl}(3)$  is known. Consider the following nonlinear filter*

$$\dot{\hat{H}} = \hat{H}U + \Delta \hat{H} \quad (3.31)$$

with

$$\Delta = \sum_{i=1}^8 \delta_i B_i, \delta = -kC^\top y$$

with  $y$  given by (3.29). Then, if the set of measured directions  $\hat{p}_i$  is consistent, the equilibrium  $\tilde{H} = I_3$  of the error system is locally exponentially stable.

### 3.5 Proof of Proposition 1

We will introduce some notations and mathematical properties which are instrumental for the proof of the main Proposition 1 in this chapter.

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

$$\delta_\wedge := \sum_{i=1}^8 \delta_i B_i$$

The operator  $vec^\vee : \mathfrak{sl}(3) \rightarrow \mathbb{R}_8$  denotes the inverse of the  $(\cdot)_\wedge$  operator, namely

$$vec^\vee(\delta_\wedge) = \delta, \forall \delta \in \mathbb{R}^8$$

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

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

The matrix representation of the composition of the linear maps  $(vect \circ \wedge) : \mathbb{R}^8 \rightarrow \mathfrak{sl}(3) \rightarrow \mathbb{R}^9$  is denoted by the full columns rank matrix  $[[\wedge]] \in \mathbb{R}^{9 \times 8}$ . We recall that the matrix representation of the composition of the linear maps  $(vect \circ \wedge)$  is denoted by the full columns rank matrix  $[[\wedge]] \in \mathbb{R}^{9 \times 8}$ . Let  $[[vec^\vee]] \in \mathbb{R}^{8 \times 9}$  denotes the matrix representation of the inverse of the map  $(vect \circ \wedge)$ , namely

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

for any  $v \in \mathbb{R}^8$ . The operator  $\otimes$  denotes the usual Kronecker product.

**Remark 1.** *The proof is inspired from Theoram 1 proved in [10].*

**Proof.** By considering the Euclidean homography (3.12), with kinematics (3.11), (3.13) along with the observer (3.23), the dynamics of the error system are given by

$$\dot{\tilde{H}} = \Delta \tilde{H}$$

### 3.5. Proof of Proposition 1

To prove that the origin of the error system  $\tilde{H} = I_3$  is locally exponentially stable, it is sufficient to show that the origin of the linearized error system is exponentially stable. Let us define  $x_\wedge$ , with  $x \in R^8$ , the first order approximation of  $\tilde{H}$  around the equilibrium  $I_3$

$$\tilde{H} \approx (I_3 + x_\wedge) = I_3 + \sum_{i=1}^8 x_i B_i$$

A first-order approximation of the output errors  $e_i$  given by (3.17), considering the equation above, can be written as

$$\begin{aligned} e_i - \dot{p}_i &= \pi_{\dot{p}_i} x_\wedge \dot{p}_i \\ &= \text{vect}(\pi_{\dot{p}_i} x_\wedge \dot{p}_i) \end{aligned}$$

Using the property  $\text{vect}(AXB) = (B^\top \otimes A)\text{vect}(X)$ , one obtains

$$\begin{aligned} e_i - \dot{p}_i &= (\dot{p}_i^\top \otimes \pi_{\dot{p}_i}) \text{vect}(x_\wedge) \\ &= (\dot{p}_i^\top \otimes \pi_{\dot{p}_i}) \llbracket \wedge \rrbracket \end{aligned} \tag{3.32}$$

Now in first order approximation one has

$$\dot{x} = -kC^\top Cx \tag{3.33}$$

with  $C_i = (\dot{p}_i^\top \otimes \pi_{\dot{p}_i})$  and  $C := \begin{bmatrix} C_1 \\ \vdots \\ C_n \end{bmatrix}$ .

From (3.32), (3.33) and using the property  $(A \otimes B)^\top = A^\top \otimes B^\top$  we obtain

$$\begin{aligned} \dot{x} &= -k \llbracket \wedge \rrbracket^\top \sum_{i=1}^n (\dot{p}_i^\top \otimes \pi_{\dot{p}_i})^\top (\dot{p}_i^\top \otimes \pi_{\dot{p}_i}) \llbracket \wedge \rrbracket x \\ &= -k \llbracket \wedge \rrbracket^\top \sum_{i=1}^n (\dot{p}_i \otimes \pi_{\dot{p}_i}) (\dot{p}_i^\top \otimes \pi_{\dot{p}_i}) \llbracket \wedge \rrbracket x \end{aligned} \tag{3.34}$$

Using the property  $(A \otimes B)(C \otimes D) = AC \otimes BD$  it yields,

$$\begin{aligned} \dot{x} &= -k \llbracket \wedge \rrbracket^\top (\dot{p}_i \dot{p}_i^\top \otimes \pi_{\dot{p}_i} \pi_{\dot{p}_i}) \llbracket \wedge \rrbracket x \\ &= -k \llbracket \wedge \rrbracket^\top (\dot{p}_i \dot{p}_i^\top \otimes \pi_{\dot{p}_i}) \llbracket \wedge \rrbracket x \end{aligned} \tag{3.35}$$

Consider the Lyapunov function

$$L_0 = \frac{|x|^2}{2k} \quad (3.36)$$

with  $\dot{L}_0 = -x^\top P x$  and the matrix  $P$  defined as follows

$$P := \llbracket \wedge \rrbracket^\top \sum_{i=1}^n (\dot{p}_i \dot{p}_i^\top \otimes \pi_{\dot{p}_i}) \llbracket \wedge \rrbracket \quad (3.37)$$

Then as shown in Theorem 1 in [10] and due to the consistency of the measurement set (see definition 5) it is proved that  $P$  is positive definite by contradiction. Therefore we can conclude that the equilibrium  $\tilde{H} = I_3$  of the error system is locally exponentially stable which in turn concludes the proof.  $\square$

### 3.6 Application to optical flow estimation

In this section we talk about the estimation of optical flow that can be obtained from the decomposition of the continuous homography and provide some experimental results. 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 in section 3.4 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 (3.13) 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:

$$\log(H) = \log(I_3 - W) \approx \widetilde{\log}(H) := -W - \frac{W^2}{2} - \frac{W^3}{3} - \dots$$

with  $W := I_3 - H$ . However, such an approximation no longer ensures that both  $\widetilde{\log}(H)$  and the resulting  $U$  will remain in  $\mathfrak{sl}(3)$ . Hence a reprojection on  $\mathfrak{sl}(3)$  is needed

$$U \approx \frac{1}{T} \left( \widetilde{\log}(H) - \frac{1}{3} \text{tr}(\widetilde{\log}(H)) I_3 \right)$$

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  $\eta$ . Our objec-

### 3.6. Application to optical flow estimation

---

tive consists in obtaining the estimation of both  $\phi$  and  $\phi_{\perp}$  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

$$\bar{U} := \frac{V\eta^{\top}}{d} = U - \frac{1}{2}\gamma_2(U + U^{\top})I_3 - \Omega_{\times}$$

with  $\gamma_2(U + U^{\top})$  the second largest eigenvalue of  $U + U^{\top}$  [48]. Taking into account the fact that  $\eta \in S^2$  one deduces

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

Defining  $\beta$  as the vector of the diagonal elements of  $\Phi$  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  $\phi$  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  $\phi_{\perp}$  is straightforwardly obtained by

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

#### 3.6.1 Experimental Results

##### 3.6.1.1 Experimental Setup

A Visual-Inertial (VI) sensor [62] developed by the company Skybotix and the Autonomous Systems Lab (ETH Zurich) has been used to perform experimental validation (see Fig. 3.2). 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

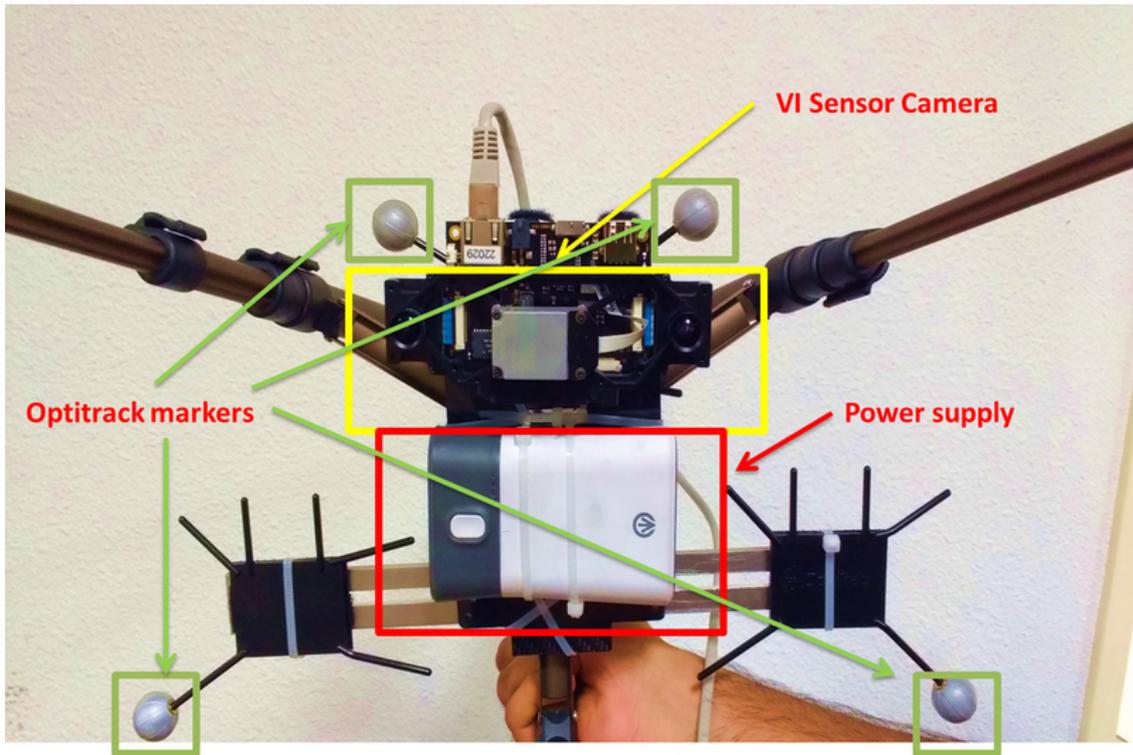


Figure 3.2: Experimental setup



Figure 3.3: Textured planar horizontal ground (target) used for experiment validations

### 3.6. Application to optical flow estimation

---

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  $\phi_{\perp}(= -\frac{\dot{d}}{d})$  in real-time.

#### 3.6.1.2 Experimental Results and Conclusions

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. 3.3). A video showing this experiment is provided as a supplementary material and is also available at:

<https://goo.gl/i8zGj2>

From Figs. 3.4 and 3.5, one can observe that the estimation of the translational optical flow  $\phi$  and its perpendicular component  $\phi_{\perp}$  (optical flow divergence) obtained from the decomposition of the estimated continuous homography are pretty accurate when compared to the corresponding ground-truth data. The experimental results thus show the performance and the robustness of the proposed approach in real time. However we also notice the fact that the estimates of the optical flow and the optical flow divergence are a bit noisy. Thus, in order to exploit these measurements e.g for control of autonomous vehicles it is extremely important to either filter these values or another possible solution could be to estimate the velocity directly.

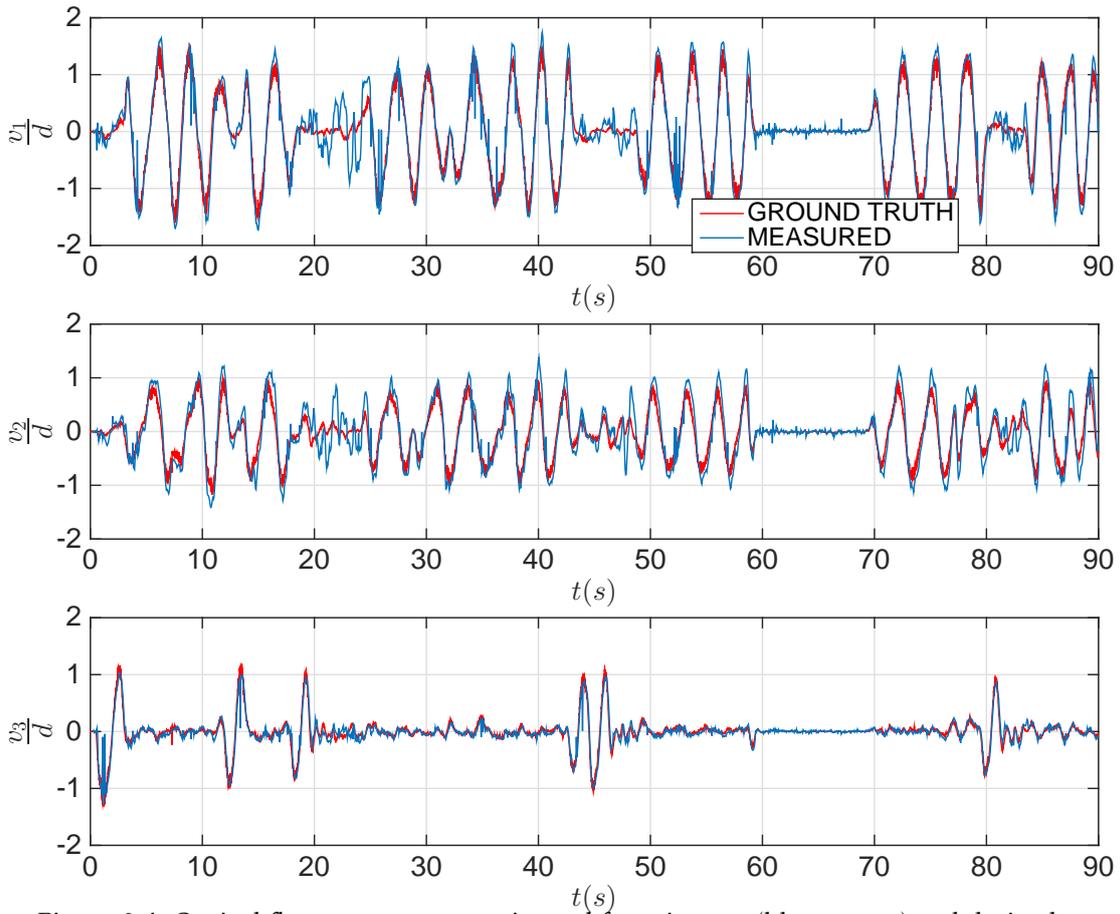


Figure 3.4: Optical flow components estimated from images (blue curves) and derived from ground truth pose (red curves) versus time (s)

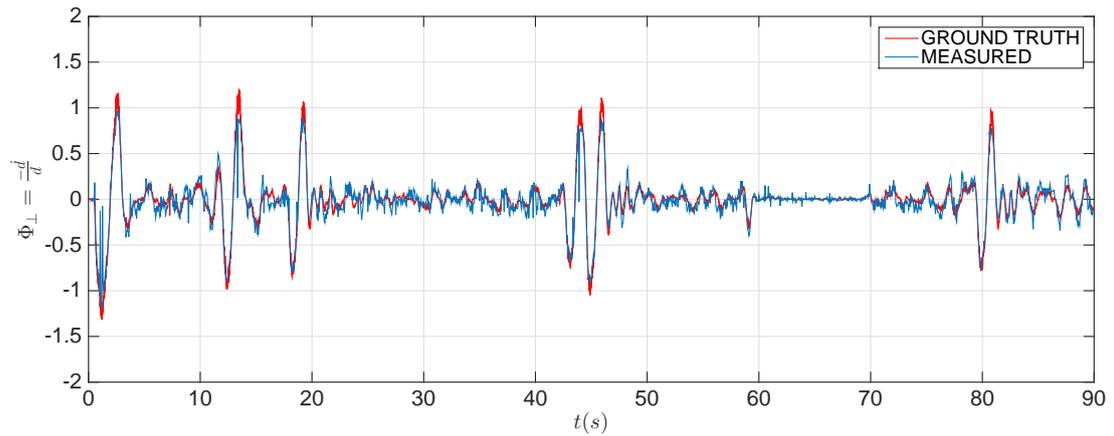


Figure 3.5:  $\phi_{\perp} = \frac{-\dot{d}}{d}$  estimated from images (blue curve) and derived from ground truth pose (red curve) versus time (s)

# 4

## **Riccati observers for state estimation exploiting optical flow and IMU measurements**

### **4.1 Introduction**

The recent proliferation of mini Unmanned Aerial Vehicles (UAVs), the emergence of modern embedded computing, and the availability of low-cost MEMS sensor systems have opened a vast range of new civil applications such as traffic congestion monitoring, environmental sensing, infrastructure inspection, real estate photography, and hazard detection and surveillance. In practice, most of these applications require the UAVs to fly in close proximity to the physical environment with GPS signals that are sometimes unavailable or unreliable. A number of research groups work actively on associated technical issues. Significant advances have also been obtained in the last few years in the domain of controlling aerial robotic vehicles. We may cite, for instance, quadrotor landing on inclined surfaces [73], quadrotor manoeuvring with a cable-suspended payload [72], inverted pendulum balancing and catching [7]. However, in these examples

the local environment is not taken into account. The control algorithms rely on full state measurements obtained from an external 3D localization system, which limits their applicability to a suitably equipped experimental flight area. Documented results for aerial robotic vehicles in a GPS-denied environment using onboard exteroceptive sensors, such as laser range finders or vision, involve far less aggressive manoeuvres [20, 70]. A central issue is thus the estimation of the vehicle's state relatively to its environment via the use of sensors that do not measure this state directly. It is only in the last five years or so that researchers have begun to tackle this issue by focusing primarily on the problem of attitude estimation when the vehicle undergoes sustained accelerations. This led to the development of velocity-aided attitude observers that fuse Inertial Measurement Unit (IMU) readings with the linear velocity measurements [19, 26, 56, 66, 41, 6, 31]. The lack of an onboard linear velocity sensor for mini UAVs led other researchers to exploit the vehicle's dynamics equations instead [44, 57, 3, 29]. Previously mentioned examples exploit proprioceptive sensor modalities whereas typical tasks involve interaction of the robotic vehicle with the environment. Combining a vision system with an IMU recently led to the development of integrated observers that exploit optical flow measurements and IMU readings to estimate the camera's attitude, linear velocity, and its distance to a planar target [76, 18, 17]. The standard approach consists in applying extended Kalman filters and showing experimentally the practical convergence of the estimation errors [76]. However, observability and convergence analyses are either missing or incomplete in these studies. An alternative solution is to use deterministic observer design techniques, alike those reported in [18, 17] except that the algorithms proposed in these references rely on the strong assumption of an IMU providing measurements of the vehicle's linear acceleration expressed in the body-fixed frame<sup>1</sup>.

In this chapter, the problem of attitude, linear velocity and depth estimation is revisited by also adopting a deterministic observer point of view, but without relying on the assumption used in [18, 17]. The proposed observer is adapted from the deterministic Riccati observer design framework derived in [23]. In contrast with most existing works on the same topic [76], the structural question of observability, on which the exponential stability and convergence of the observer rely, is here addressed with the derivation of an explicit (and simple) observability condition based on the persistence of excitation granted by the camera-IMU linear velocity. Several practical algorithmic and implementation issues are also discussed. In the end we also propose an extension to observer design for optical flow filtering.

---

<sup>1</sup>In fact, accelerometers only provide measurements of the so-called specific acceleration.

## 4.2 System Equations and Measurements for Observer Design

The vehicle's attitude satisfies the differential equation

$$\dot{R} = R\Omega_{\times} \quad (4.1)$$

It is assumed that the vehicle is equipped with an IMU comprising a 3-axis gyrometer that measures the angular velocity  $\Omega \in \mathbb{R}^3$  and a 3-axis accelerometer that measures the so-called *specific acceleration*  $a_{\mathcal{B}} \in \mathbb{R}^3$ , expressed in  $\{\mathcal{B}\}$ . Using the flat non-rotating Earth assumption, one has [6]

$$\dot{V} = -\Omega_{\times}V + a_{\mathcal{B}} + gR^{\top}e_3 \quad (4.2)$$

A 3-axis magnetometer is also assumed to be available to measure the normalized Earth's magnetic field vector expressed in  $\{\mathcal{B}\}$ . Let  $m_{\mathcal{I}} \in S^2$  denote the known normalized Earth's magnetic field vector expressed in  $\{\mathcal{I}\}$ . The vectors  $m_{\mathcal{I}}$  and  $e_3$  are usually assumed to be non-collinear so that  $R$  can be estimated from the observation (measurements) in the body-fixed frame of the gravity vector and of the Earth's magnetic field vector. The magnetometer thus measures  $m_{\mathcal{B}} = R^{\top}m_{\mathcal{I}}$ . The need for using the magnetometer is optional and is only required for yaw estimation.

We further assume that the vehicle is equipped with a monocular camera that observes a planar scene so that we can obtain an estimation of the continuous homography from which we extract the optical flow ( $\phi := \frac{V}{d}$ ) and the optical flow divergence ( $\phi_{\perp} := \frac{V^{\top}n}{d} = \frac{-\dot{d}}{d}$ ) which serve as measurements for the observer design. We further define

$$s \triangleq \frac{1}{d} \quad (4.3)$$

From (4.3) one deduces

$$\dot{s} = -\frac{\dot{d}}{d^2} = \phi_{\perp} s \quad (4.4)$$

In summary, the observer design for  $R$ ,  $V$  and  $s$  (i.e.  $d$ ) will be based on Eqs. (4.1)–(4.4) and the measured quantities  $\frac{V}{d}$ ,  $\phi$ ,  $\omega$ ,  $a_{\mathcal{B}}$  and  $m_{\mathcal{B}}$ . We will also show thereafter that the magnetometer measurements  $m_{\mathcal{B}}$  are not required if we only need to estimate the gravity direction (i.e.  $R^{\top}e_3$ ) instead of the whole attitude  $R$ .

## 4.3 Partial Attitude, Linear Velocity and Depth Estimation of a Camera observing a planar target using continuous homography and inertial data

### 4.3.1 Observer Derivation

Let  $\hat{R} \in SO(3)$ ,  $\hat{V} \in \mathbb{R}^3$ ,  $\hat{s} \in \mathbb{R}$  denote the estimates of  $R$ ,  $V$  and  $s$ , respectively. In view of the equations (4.1), (4.2) and (4.4), the following observer form is considered

$$\begin{cases} \dot{\hat{R}} = \hat{R}\Omega_{\times} - [\sigma_R]_{\times}\hat{R} \\ \dot{\hat{V}} = -\Omega_{\times}\hat{V} + a_B + g\hat{R}^{\top}e_3 - \sigma_v \\ \dot{\hat{s}} = \phi_{\perp}\hat{s} - \sigma_s \end{cases} \quad (4.5)$$

where  $\sigma_R, \sigma_v \in \mathbb{R}^3$ ,  $\sigma_s \in \mathbb{R}$  are innovation terms to be designed thereafter. Defining the observer error variables

$$\tilde{R} \triangleq R\hat{R}^{\top}, \quad \tilde{V} \triangleq V - \hat{V}, \quad \tilde{s} \triangleq s - \hat{s}$$

then the objective of the observer consists in stabilizing  $(\tilde{R}, \tilde{V}, \tilde{s})$  around  $(e_3, 0, 0)$  when the estimation of the gravity direction is concerned instead of the whole attitude estimation). One verifies from (4.1), (4.2), (4.5) that the error dynamics are given by

$$\begin{cases} \dot{\tilde{R}} = \tilde{R}[\sigma_R]_{\times} \\ \dot{\tilde{V}} = -\Omega_{\times}\tilde{V} + g\hat{R}^{\top}(\tilde{R}^{\top} - I_3)e_3 + \sigma_v \\ \dot{\tilde{s}} = \phi_{\perp}\tilde{s} + \sigma_s \end{cases} \quad (4.6)$$

We will work out next first order approximations of the error system (4.6) complemented with first order approximations of the measurement equations. The application to these approximations of the Riccati observer design framework reported in [23] (see Section II-D) will then provide us with the equations of the innovation terms of the proposed observer.

For this application the following technical (but non-restrictive) assumption is made.

**Assumption 2.**  $V(t)$ ,  $\dot{V}(t)$ ,  $\Omega(t)$  and  $\phi$  are bounded in norm by some positive numbers  $V_{max}$ ,  $\dot{V}_{max}$ ,  $\Omega_{max}$  and  $\phi_{max}$ , respectively. The distance  $d$  is lower- and upper-bounded by some positive numbers  $d_{min}$  and  $d_{max}$ , respectively.

First order approximations of the attitude error equations are derived using a (local) minimal parametrization of the rotation group  $SO(3)$ . The parametrization here chosen is the vector part  $\tilde{q}_v$  of the Rodrigues unit quaternion  $\tilde{q} = (\tilde{q}_0, \tilde{q}_v)$  associated with  $\tilde{R}$ .

### 4.3. Partial Attitude, Linear Velocity and Depth Estimation of a Camera observing a planar target using continuous homography and inertial data

---

Rodrigues formula relating  $\tilde{q}$  to  $\tilde{R}$  is

$$\tilde{R} = I_3 + 2[\tilde{q}_v]_\times(\tilde{q}_0 I_3 + [\tilde{q}_v]_\times)$$

From this relation, one deduces

$$\tilde{R} = I_3 + [\tilde{\lambda}]_\times + O(|\tilde{\lambda}|^2),$$

with  $\tilde{\lambda} \in \mathcal{B}_2^3$  equal to twice the vector part of the quaternion associated with the attitude error matrix  $\tilde{R}$ . Then, in view of the dynamics of  $\tilde{R}$  in (4.6) one verifies (see also [23]) that the derivative of  $\tilde{\lambda}$  is given by

$$\dot{\tilde{\lambda}} = \sigma_R + O(|\tilde{\lambda}||\sigma_R|)$$

As for the dynamics of  $\tilde{V}$  one obtains

$$\begin{aligned} \dot{\tilde{V}} &= -\Omega_\times \tilde{V} + g\hat{R}^\top [e_3]_\times \tilde{\lambda} + \sigma_v + O(|\tilde{\lambda}|^2) \\ &= -\Omega_\times \tilde{V} + g\hat{R}^\top e_2 \tilde{\lambda}_1 - g\hat{R}^\top e_1 \tilde{\lambda}_2 + \sigma_v \end{aligned}$$

with  $\tilde{\lambda}_1, \tilde{\lambda}_2$  the first and second components of  $\tilde{\lambda}$ .

Concerning the measurement of  $\frac{V}{d}$  one has

$$\begin{aligned} \frac{V}{d} - \hat{V}\hat{s} &= (\tilde{V} + \hat{V})(\tilde{s} + \hat{s}) - \hat{V}\hat{s} \\ &= (\hat{s}I_3)\tilde{V} + \hat{V}\tilde{s} + O(|\tilde{V}||\tilde{s}|) \end{aligned}$$

By setting the system output vector equal to

$$y = \frac{V}{d} - \hat{V}\hat{s} \tag{4.7}$$

one obtains LTV first order approximations in the form (2.20) with

$$\left\{ \begin{array}{l} x = \begin{bmatrix} \tilde{\lambda}_1 \\ \tilde{\lambda}_2 \\ \tilde{s} \\ \tilde{V} \end{bmatrix}, \quad x_1 = \begin{bmatrix} \tilde{\lambda}_1 \\ \tilde{\lambda}_2 \end{bmatrix}, \quad x_2 = \begin{bmatrix} \tilde{s} \\ \tilde{V} \end{bmatrix}, \quad u = \begin{bmatrix} \sigma_{R,1} \\ \sigma_{R,2} \\ \sigma_s \\ \sigma_v \end{bmatrix} \\ A = \begin{bmatrix} 0 & 0 & 0 & 0_{1 \times 3} \\ 0 & 0 & 0 & 0_{1 \times 3} \\ 0 & 0 & \phi_\perp & 0_{1 \times 3} \\ g\hat{R}^\top e_2 & -g\hat{R}^\top e_1 & 0_{3 \times 1} & -\Omega_\times \end{bmatrix} \in \mathbb{R}^{6 \times 6} \\ C = \begin{bmatrix} 0_{3 \times 1} & 0_{3 \times 1} & \hat{V} & \hat{s}I_3 \end{bmatrix} \in \mathbb{R}^{3 \times 6} \end{array} \right. \tag{4.8}$$

with  $\sigma_{R,1}, \sigma_{R,2}$  the first and second components of  $\sigma_R$ . The third component  $\sigma_{R,3}$  is considered to be zero. From there the proposed observer is given by (4.5) with  $\sigma_{R,1}, \sigma_{R,2}$  and  $\sigma_v$  determined from the input  $u$  calculated according to (2.22) and (2.23).

### 4.3.2 Observability analysis

According to [23, Corollary 3.2], good conditioning of the solutions  $P(t)$  to the CREs and exponential stability of the proposed observer rely on the uniform observability of the pair  $(A^*(t), C^*(t))$  obtained by setting  $x = 0$  in the expressions of the matrices  $A$  and  $C$  given by (4.8). One verifies that

$$\begin{cases} A^* = \begin{bmatrix} 0 & 0 & 0 & 0_{1 \times 3} \\ 0 & 0 & 0 & 0_{1 \times 3} \\ 0 & 0 & \phi_{\perp} & 0_{1 \times 3} \\ gR^{*\top} e_2 & -gR^{*\top} e_1 & 0_{3 \times 1} & -\Omega_{\times} \end{bmatrix} \\ C^* = \begin{bmatrix} 0_{3 \times 1} & 0_{3 \times 1} & V & sI_3 \end{bmatrix} \end{cases} \quad (4.9)$$

with  $s = \frac{1}{d}$ ,  $\phi_{\perp} = -\dot{d}/d$ , and  $R^* \in \text{SO}(3)$  satisfying

$$R^{*\top} e_3 = R^{\top} e_3, \quad \dot{R}^* = R^* \Omega_{\times}$$

**Proposition 2.** *The transient matrix  $\Phi(\tau, t)$  associated with  $A^*(t)$ , for all  $\tau \geq t$ , is given by*

$$\Phi(\tau, t) = \begin{bmatrix} 1 & 0 & 0 & 0_{1 \times 3} \\ 0 & 1 & 0 & 0_{1 \times 3} \\ 0 & 0 & \frac{s(\tau)}{s(t)} & 0_{1 \times 3} \\ g(\tau - t)R^*(\tau)^{\top} e_2 & -g(\tau - t)R^*(\tau)^{\top} e_1 & 0_{3 \times 1} & R^*(\tau)^{\top} R^*(t) \end{bmatrix} \quad (4.10)$$

**Proof:** The transient matrix  $\Phi(\tau, t)$  is the solution to the following equation

$$\frac{d\Phi(\tau, t)}{d\tau} = A^*(\tau)\Phi(\tau, t), \quad \Phi(t, t) = I_6 \quad (4.11)$$

From the above relation and the expression of  $A^*(\tau)$ , one easily verifies that  $\Phi(\tau, t)$  has the form

$$\Phi(\tau, t) = \begin{bmatrix} 1 & 0 & 0 & 0_{1 \times 3} \\ 0 & 1 & 0 & 0_{1 \times 3} \\ 0 & 0 & \varphi_{33}(\tau, t) & 0_{1 \times 3} \\ \varphi_{41}(\tau, t) & \varphi_{42}(\tau, t) & 0_{3 \times 1} & \bar{R}(\tau, t)^{\top} \end{bmatrix} \quad (4.12)$$

### 4.3. Partial Attitude, Linear Velocity and Depth Estimation of a Camera observing a planar target using continuous homography and inertial data

---

where  $\bar{R}(\tau, t)$ , with  $\tau \geq t$ , is the solution to the equation

$$\frac{d\bar{R}(\tau, t)}{d\tau} = \bar{R}(\tau, t)\Omega(\tau)_{\times}, \quad \bar{R}(t, t) = I_3$$

It remains to compute  $\varphi_{33}(\tau, t)$ ,  $\varphi_{41}(\tau, t)$ ,  $\varphi_{42}(\tau, t)$ . From (4.11), (4.12), and the definition of  $\phi_{\perp}$  (i.e.  $\phi_{\perp} = -\dot{d}/d$ ) one gets

$$\begin{aligned} \frac{d\varphi_{33}(\tau, t)}{d\tau} &= -\frac{dd(\tau)}{d\tau} \frac{1}{d(\tau)} \varphi_{33}(\tau, t) \\ \Rightarrow \frac{d\varphi_{33}(\tau, t)}{\varphi_{33}(\tau, t)} &= -\frac{dd(\tau)}{d(\tau)} \\ \Rightarrow \varphi_{33}(\tau, t) &= \frac{d(t)}{d(\tau)} = \frac{s(\tau)}{s(t)} \end{aligned} \tag{4.13}$$

From (4.11) and (4.12), one has

$$\frac{d\varphi_{41}(\tau, t)}{d\tau} = -\Omega(\tau)_{\times}\varphi_{41}(\tau, t) + gR^*(\tau)^{\top}e_2$$

Then, by change of variable  $\bar{\varphi}_{41}(\tau, t) := R^*(\tau)\varphi_{41}(\tau, t)$  the above equation yields

$$\frac{d\bar{\varphi}_{41}(\tau, t)}{d\tau} = ge_2$$

from which one deduces

$$\varphi_{41}(\tau, t) = R^*(\tau)^{\top}\bar{\varphi}_{41}(\tau, t) = g(\tau - t)R^*(\tau)^{\top}e_2 \tag{4.14}$$

Similarly, one gets

$$\varphi_{42}(\tau, t) = R^*(\tau)^{\top}\bar{\varphi}_{41}(\tau, t) = -g(\tau - t)R^*(\tau)^{\top}e_1 \tag{4.15}$$

Finally, using the fact that  $\bar{R}(\tau, t)^{\top} = R^*(\tau)^{\top}R^*(t)$  along with relations (4.13), (4.14) and (4.15), one deduces the explicit form (4.10) of the transient matrix  $\Phi(\tau, t)$  (**End of Proof**).

The following technical (but non-restrictive) assumption is made.

**Assumption 3.**  $V(t)$ ,  $\dot{V}(t)$ ,  $\Omega(t)$  and  $\phi_{\perp}$  are bounded in norm by some positive numbers  $V_{max}$ ,  $\dot{V}_{max}$ ,  $\Omega_{max}$  and  $\phi_{\perp max}$ , respectively. The distance  $d$  is lower- and upper-bounded by some positive numbers  $d_{min}$  and  $d_{max}$ , respectively.

**Proposition 3.** Assume that Assumption 3 holds. Assume that **one** of the following ‘‘persistent excitation’’ conditions is satisfied:

(C.1) For all  $R_\gamma \in \text{SO}(3)$  of the form

$$R_\gamma = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.16)$$

there exist  $\delta, \rho > 0$  and  $t_0 \geq 0$  such that

$$\frac{1}{\delta} \int_t^{t+\delta} U(\tau)^\top U(\tau) d\tau \geq \rho I_6, \quad \forall t \geq t_0 \quad (4.17)$$

with

$$U(\tau) := s(\tau) \begin{bmatrix} g(\tau-t)e_1 & g(\tau-t)e_2 & R_\gamma^\top v(\tau) & I_3 \end{bmatrix} \quad (4.18)$$

where  $v(\tau) := R(\tau)V(\tau)$  the velocity vector expressed in the inertial frame.

(C.2) For all  $R_\gamma \in \text{SO}(3)$  of the form (4.16), there exists  $\delta, \rho > 0$  and  $t_0 \geq 0$  such that

$$\int_t^{t+\delta} W(\tau)^\top W(\tau) d\tau - \frac{1}{\delta} \left( \int_t^{t+\delta} W(\tau) d\tau \right)^\top \left( \int_t^{t+\delta} W(\tau) d\tau \right) \geq \rho I_3, \quad \forall t \geq t_0 \quad (4.19)$$

with

$$W(\tau) := \begin{bmatrix} g(\tau-t)e_1 & g(\tau-t)e_2 & R_\gamma^\top v(\tau) \end{bmatrix} \quad (4.20)$$

Then, the pair  $(A^*, C^*)$  given by (4.9) is uniformly observable and the equilibrium  $(\hat{R}^\top e_3, \hat{V}, \hat{s}) = (R^\top e_3, V, s)$  of the proposed Riccati observer is locally exponentially stable.

**Proof:** From definition 4 in Chapter 2 the pair  $(A^*, C^*)$  is uniformly observable if  $\exists \delta, \mu > 0$  and  $t_0 \geq 0$  such that

$$\frac{1}{\delta} \int_t^{t+\delta} \Phi(\tau)^\top C^{*\top}(\tau) C^*(\tau) \Phi(\tau) d\tau \geq \mu I_6, \quad \forall t > t_0 \quad (4.21)$$

We now prove (4.21). In fact, it is verified that

$$U(\tau) = s(\tau) \begin{bmatrix} W(\tau) & I_3 \end{bmatrix}$$

Thus, one has

$$U(\tau)^\top U(\tau) \leq \frac{1}{d_{min}^2} \begin{bmatrix} W(\tau)^\top W(\tau) & W(\tau)^\top \\ W(\tau) & I_3 \end{bmatrix}$$

Then, using the fact that  $W$  is bounded, it is straightforward to verify that condition (C.2) implies condition (C.1). Therefore, we only need to prove (4.21) for the case where the persistent excitation condition (C.1) is satisfied.

### 4.3. Partial Attitude, Linear Velocity and Depth Estimation of a Camera observing a planar target using continuous homography and inertial data

---

One verifies that

$$\begin{aligned} C^*(\tau)\Phi(\tau) &= s(\tau) \begin{bmatrix} g(\tau-t)R^*(\tau)^\top e_2 & -g(\tau-t)R^*(\tau)^\top e_1 & \frac{V(\tau)}{s(t)} & R^*(\tau)^\top R^*(t) \end{bmatrix} \\ &= s(\tau)R(\tau)^\top \begin{bmatrix} g(\tau-t)\gamma_2 & -g(\tau-t)\gamma_1 & \frac{v(\tau)}{s(t)} & R(t) \end{bmatrix} \end{aligned} \quad (4.22)$$

where  $v(\tau) := R(\tau)V(\tau)$  the velocity vector expressed in the inertial frame. Also

$$\gamma_{1,2} := R(\tau)R^*(\tau)^\top e_{1,2} = R(0)R^*(0)^\top e_{1,2} \quad (4.23)$$

are constant and, thus, do not depend on  $\tau$ . Using the relation  $R^*(\tau)^\top e_3 = R^\top e_3$  one verifies that

$$e_3^\top \gamma_{1,2}(\tau) = e_3^\top R(\tau)R^*(\tau)^\top e_{1,2} = 0$$

meaning that  $\gamma_{1,2}$  is spanned by  $\{e_1, e_2\}$ . Therefore, there exists a constant rotation matrix  $R_\gamma$  of the form (4.16) such that

$$R_\gamma \gamma_2 = e_1, R_\gamma \gamma_1 = -e_2, R_\gamma \in \text{SO}(3) \quad (4.24)$$

For all  $x = [x_1, x_2, x_3, x_4]^\top$ , with  $x_{1,2,3} \in \mathbb{R}, x_4 \in \mathbb{R}^3$ , using (4.22) one deduces

$$\begin{aligned} C^*(\tau)\Phi(\tau)x &= s(\tau)R(\tau)^\top R_\gamma \begin{bmatrix} g(\tau-t)e_1 & g(\tau-t)e_2 & \frac{R_\gamma^\top v(\tau)}{s(t)} & R_\gamma^\top R(t) \end{bmatrix} x \\ &= s(\tau)R(\tau)^\top R_\gamma \begin{bmatrix} g(\tau-t)e_1 & g(\tau-t)e_2 & R_\gamma^\top v(\tau) & I_3 \end{bmatrix} \bar{x} \\ &= R(\tau)^\top R_\gamma U(\tau) \bar{x} \end{aligned}$$

with  $\bar{x} := [x_1, x_2, \frac{x_3}{s(t)}, (R_\gamma^\top R(t)x_4)^\top]^\top$  (which ensures that  $|\bar{x}| \geq \min(1, d_{min})|x|$ ) and  $U$  defined by (4.18). One then obtains

$$\begin{aligned} x^\top \Phi(\tau)^\top C^{*\top}(\tau)C^*(\tau)\Phi(\tau)x &= \bar{x}^\top U(\tau)^\top R_\gamma^\top R(\tau)R(\tau)^\top R_\gamma U(\tau)\bar{x} \\ &= \bar{x}^\top U(\tau)^\top U(\tau)\bar{x} \end{aligned} \quad (4.25)$$

From (4.25) and the persistent excitation condition C.1, one straightforwardly deduces inequality (4.21), with  $\mu = (\min(1, d_{min}))^2 \rho$ , which concludes the proof (**End of proof**).

Now the persistent excitation condition give by Proposition (3) deserves some comments.

- First, this condition is violated if the vehicle's linear velocity is null for all time (i.e.  $|V(t)| \equiv 0$ ). In fact, in this situation it is impossible to recover the depth from

monocular vision without any prior knowledge of the observed planar scene.

- If the vehicle does move “persistently” so that  $\exists \delta_V, \rho_V > 0$  such that

$$\frac{1}{\delta_V} \int_t^{t+\delta_V} |V(\tau)| d\tau \geq \rho_V, \forall t \geq 0 \quad (4.26)$$

and that Assumption 2 holds, then condition (3) is satisfied for almost all types of motion, excepts some very particular cases. For instance, such is the case where the vehicle moves, with constant linear velocity and constant attitude, in a straight-line path parallel to the observed plane (i.e.  $V(t)$  and  $d(t)$  remain constant).

### 4.3.3 Yaw estimation

For completeness, the third component  $\sigma_{R,3}$  of the innovation term  $\sigma_R$  is now independently designed for yaw estimation.

**Corollary 2.** *In addition to the innovations  $\sigma_{R,1}$ ,  $\sigma_{R,2}$  and  $\sigma_V$  specified previously, define*

$$\sigma_{R,3} = -k_m e_3^\top (\hat{R} m_B \times m_I) \quad (4.27)$$

where  $k_m \in \mathbb{R}$  is either a positive number or  $k_m = (m_1^2 + m_2^2) D_m P_m$ , with  $P_m \in \mathbb{R}$  solution to the following CRE:

$$\dot{P}_m = -(m_1^2 + m_2^2)^2 D_m P_m^2 + S_m, P_m(0) > 0$$

and  $D_m, S_m$  positive numbers. Then, the equilibrium  $(\hat{R}, \hat{V}, \hat{s}) = (R, V, s)$  of the proposed Riccati observer is locally exponentially stable.

**Proof.** As a result of Proposition 3, it suffices to prove the local exponential stability of  $\tilde{\lambda}_3 = 0$  at the local zero-dynamics of  $\tilde{\lambda}_3$  by setting  $\tilde{\lambda}_1 \equiv \tilde{\lambda}_2 \equiv 0$  and  $\tilde{V} \equiv 0$ . One verifies that the zero-dynamics of  $\tilde{\lambda}_3$  are locally given by

$$\dot{\tilde{\lambda}}_3 = \sigma_{R,3}$$

while the “conditioned” magnetometer measurement  $e_3^\top (\hat{R} m_B \times m_I)$  in first order approximations and with  $\tilde{\lambda}_1 \equiv \tilde{\lambda}_2 \equiv 0$  is approximately given by

$$e_3^\top (\hat{R} m_B \times m_I) = (m_1^2 + m_2^2) \tilde{\lambda}_3$$

so that

$$\dot{\tilde{\lambda}}_3 = -k_m (m_1^2 + m_2^2) \tilde{\lambda}_3$$

From here the proof straightforwardly follows.  $\square$

## 4.4 Practical implementation aspects

### 4.4.1 Unit quaternion equivalence

Although the attitude estimate is designed on  $SO(3)$ , it can be directly lifted to an equivalent algorithm on the unit quaternion group (see, e.g., [36]). Let  $\hat{q}$  denote the unit quaternion associated with  $\hat{R}$ . Then, the proposed observer (5.11) can be rewritten as

$$\begin{cases} \dot{\hat{q}} = \frac{1}{2}(\Gamma_1(\Omega) - \Gamma_2(\sigma_R))\hat{q} \\ \dot{\hat{V}} = -\Omega_{\times}\hat{V} + a_B + g\hat{R}^{\top}e_3 - \sigma_v \\ \dot{\hat{s}} = \phi_{\perp}\hat{s} - \sigma_s \end{cases} \quad (4.28)$$

where the mappings  $\Gamma_1, \Gamma_2 : \mathbb{R}^3 \rightarrow \mathbb{R}^{4 \times 4}$  are defined as

$$\Gamma_1(x) = \begin{bmatrix} 0 & -x^{\top} \\ x & -[x]_{\times} \end{bmatrix}, \quad \Gamma_2(x) = \begin{bmatrix} 0 & -x^{\top} \\ x & [x]_{\times} \end{bmatrix}, \quad \forall x \in \mathbb{R}^3$$

and the term  $\hat{R}$  is calculated from  $\hat{q}$  using the Rodrigues formula.

### 4.4.2 Hybrid discrete-continuous version

In practice, the IMU measurements can be obtained at a very high frequency while the continuous homography is often estimated at a much lower frequency. This fact should be carefully taken into account in the implementation process. Inspired by existing hybrid continuous-discrete Kalman or extended Kalman filters, we propose thereafter a hybrid continuous-discrete version of the proposed observer, where for the sake of simplicity the gain  $k_m$  involved in the expression (4.27) of the innovation component  $\sigma_{R,3}$  is a positive number.

Let  $\{t_k\}$  denote the suite of time-instants that the continuous homography estimates are obtained. Then, the prediction and correction steps of the proposed observer are described below.

• **Prediction step:** At each step  $k$ , integrate during  $t \in [t_{k-1}, t_k]$  the following equations

$$\begin{cases} \dot{\hat{q}} = \frac{1}{2}(\Gamma_1(\Omega) - \Gamma_2(\sigma_{R,3}e_3))\hat{q} \\ \dot{\hat{V}} = -\Omega_{\times}\hat{V} + a_B + g\hat{R}^{\top}e_3 \\ \dot{\hat{s}} = \phi_{\perp}\hat{s} \\ \dot{P} = AP + PA^{\top} + S \end{cases}$$

with  $\hat{q}(t_{k-1}) = \hat{q}_{k-1|k-1}$ ,  $\hat{V}(t_{k-1}) = \hat{V}_{k-1|k-1}$ ,  $\hat{s}(t_{k-1}) = \hat{s}_{k-1|k-1}$ ,  $P(t_{k-1}) = P_{k-1|k-1}$  to obtain

$$\begin{cases} \hat{q}_{k|k-1} = \hat{q}(t_k) \\ \hat{V}_{k|k-1} = \hat{V}(t_k) \\ \hat{s}_{k|k-1} = \hat{s}(t_k) \\ P_{k|k-1} = P(t_k) \end{cases}$$

- **Correction step:** First, compute the innovation terms as

$$\begin{cases} K_k = P_{k|k-1} C_k^\top (C_k P_{k|k-1} C_k^\top + D^{-1})^{-1} \\ u_k = \begin{bmatrix} \sigma_{Rk,1} \\ \sigma_{Rk,2} \\ \sigma_{sk} \\ \sigma_{vk} \end{bmatrix} = -K_k y_k \end{cases}$$

with  $C_k = \begin{bmatrix} 0_{3 \times 1} & 0_{3 \times 1} & \hat{V}_{k|k-1} & \hat{s}_{k|k-1} I_3 \end{bmatrix}$ ,  $y_k = \left(\frac{V}{d}\right)_k - \hat{V}_{k|k-1} \hat{s}_{k|k-1}$ . Then, update the state estimates and the Riccati matrix as

$$\begin{cases} \hat{q}_{k|k} = \exp\left(-\frac{1}{2} \Gamma_2(\sigma_{Rk,1} e_1 + \sigma_{Rk,2} e_2)\right) \hat{q}_{k|k-1} \\ \hat{V}_{k|k} = \hat{V}_{k|k-1} - \sigma_{V_k} \\ \hat{s}_{k|k} = \hat{s}_{k|k-1} - \sigma_{s_k} \\ P_{k|k} = (I_6 - K_k C_k) P_{k|k-1} \end{cases}$$

with  $\text{sinc}(x) = \sin(x)/x, \forall x \in \mathbb{R}$ .

#### 4.4.3 Practical solutions for the boundedness of $P$

As mentioned in Section 4.3.2 the Riccati matrix  $P$  is well conditioned provided that the pair  $(A^*, C^*)$  given in (4.9) is uniformly observable (i.e. the persistent excitation condition (3) is satisfied). However, when this uniform observability condition is violated (as discussed in the end of Section 4.3.2)  $P$  may grow arbitrarily large or even explode. Some “practical” solutions to that issue are proposed next. For instance, when the measured quantity  $\frac{V}{d}$  is not null, it is likely that condition (3) is satisfied. Therefore, when the norm of  $\frac{V}{d}$  is smaller than some small threshold, one can simply inactivate the correction step and also the integration of  $P$  within the prediction step. Another solution consists in saturating  $P$  after every correction step so that its Frobenius norm remains always smaller than a given threshold. The latter should be chosen large enough so that the saturation of  $P$  will not occur when the system is uniformly observable (i.e. sufficiently excited).

#### 4.4. Practical implementation aspects

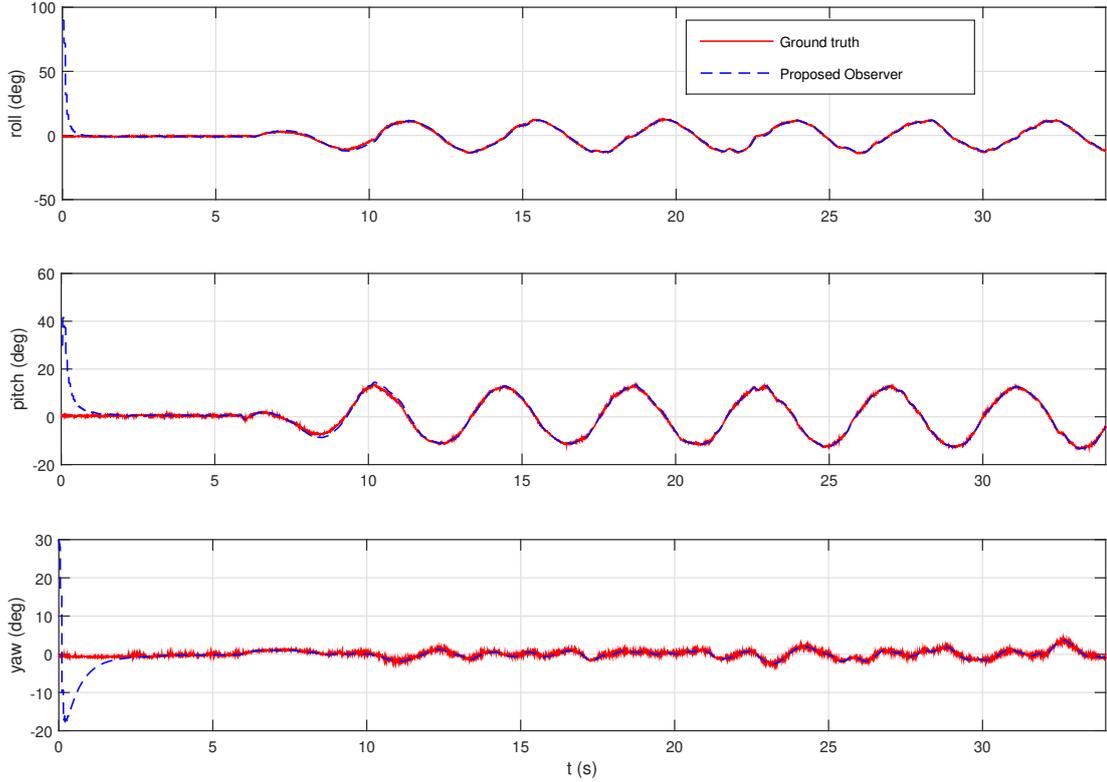


Figure 4.1: (Simulation) Estimated and ground-truth attitudes represented by roll, pitch and yaw Euler angles (*deg*) versus time (*s*)

#### 4.4.4 Simulation results

Simulations have been carried out using a dataset<sup>2</sup> of a quadrotor UAV performing a take-off followed by an aggressive circular flight so that its linear velocity and attitude vary rapidly and in large proportions. The dataset contains time-stamped measurements from the quadrotor’s IMU ( $180Hz$ ) and “ground truth” poses (i.e. position and attitude) provided by an external motion capture system ( $200Hz$ ). This dataset is ideal to validate the proposed observer for an extreme condition (i.e. aggressive flight) but without the need for image processing and continuous homography estimation. Note that the latter is not the focus of the present paper. As a matter of fact, the quantities such as  $\frac{V}{a}$ ,  $\phi_{\perp}$  and  $m_B$  can be emulated from the ground-truth pose measurements. Whereas  $m_B$  can be easily computed from the ground-truth attitude measurements (i.e.  $m_B = R^T m_I$ ), obtaining  $\frac{V}{a}$  and  $\phi_{\perp}$  is more involved. For instance, using the position measurements a high-gain observer is applied for the estimation of the linear velocity expressed in the inertial frame that is then converted to the linear velocity expressed in the body-fixed

<sup>2</sup>Available at [http://rpg.ifi.uzh.ch/software\\_datasets.html](http://rpg.ifi.uzh.ch/software_datasets.html)

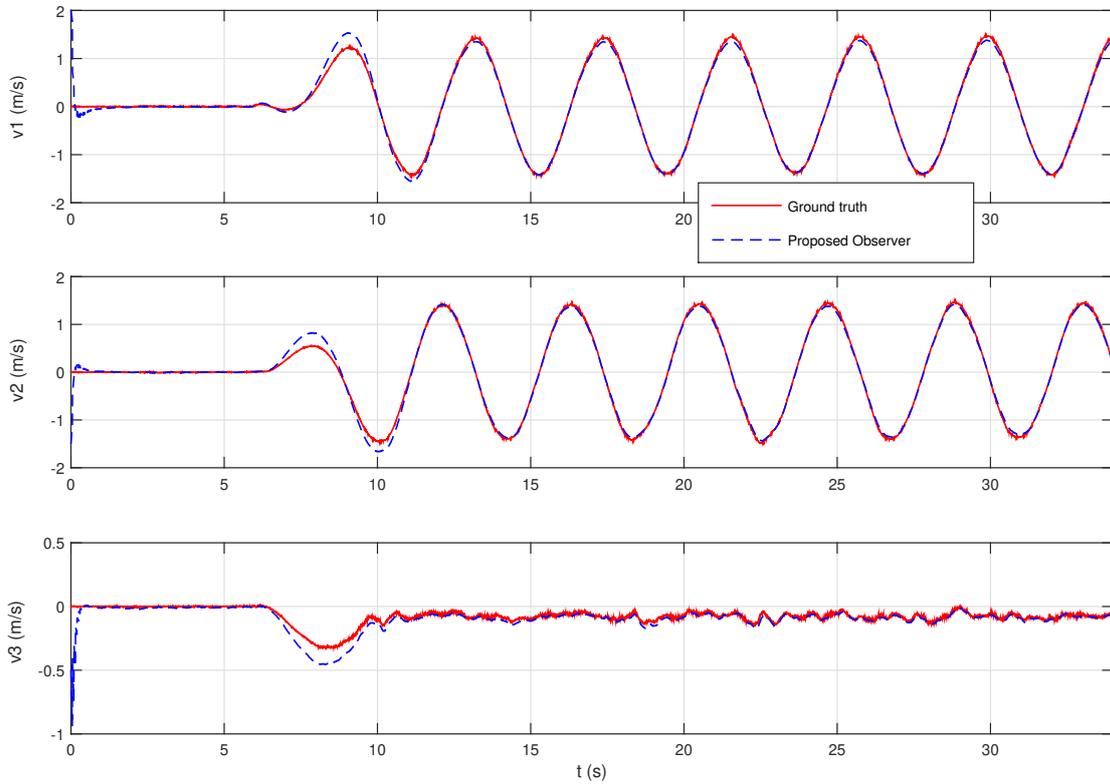


Figure 4.2: (Simulation) Estimated and ground-truth linear velocity components in body-fixed frame ( $m/s$ ) versus time ( $s$ )

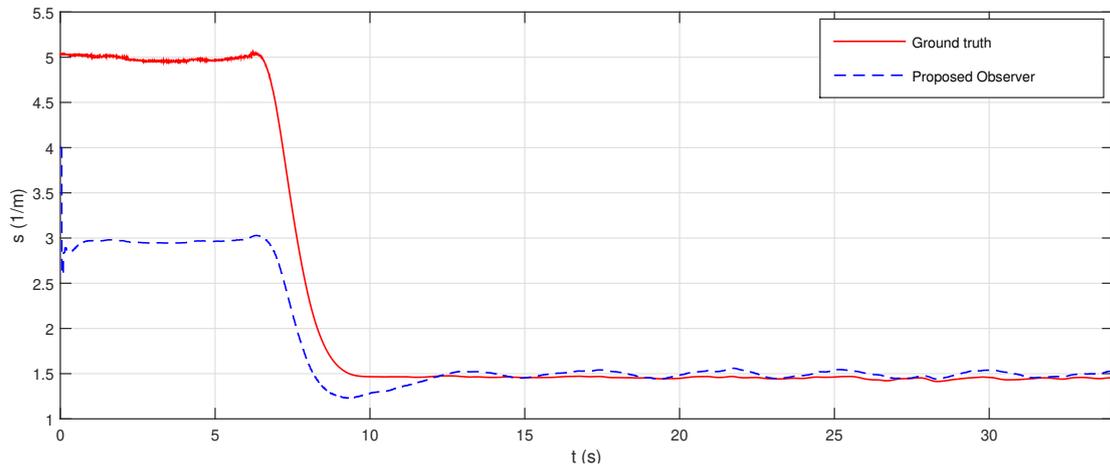


Figure 4.3: (Simulation) Estimated and ground-truth depth inverse ( $m^{-1}$ ) versus time ( $s$ )

frame using the ground-truth attitude. Then, by considering the situation where the UAV carries a downward-looking camera to observe a planar horizontal ground, it is

#### 4.4. Practical implementation aspects

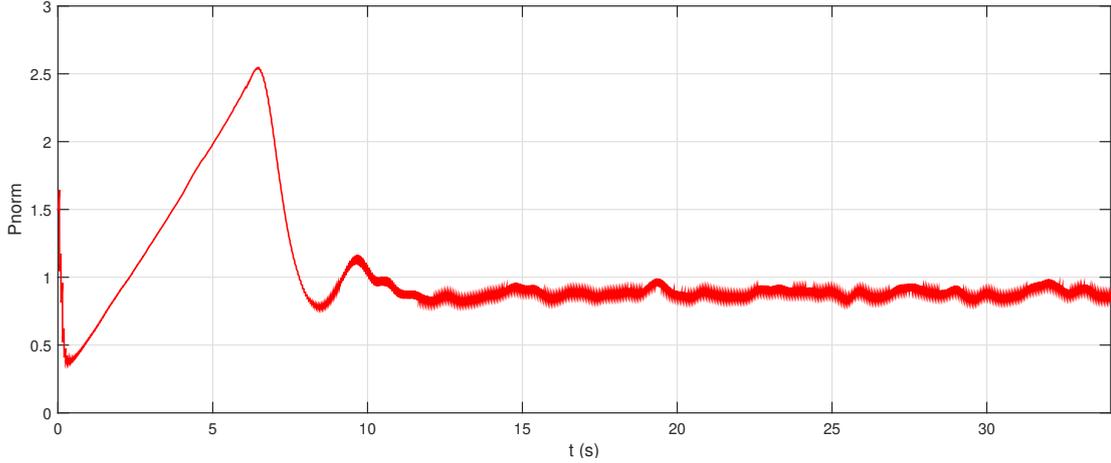


Figure 4.4: (Simulation) The Frobenius norm  $|P|$  versus time (s)

not difficult to construct the “measurements” of  $\frac{V}{d}$  and  $\phi_{\perp}$  (at  $20Hz$ ). For the reported data-based simulations, the hybrid continuous-discrete version of the proposed Riccati observer has been performed with  $D = I_3$ ,  $S = \text{diag}\{0.01I_2, 0.25, 0, 36I_3\}$  and  $P(0) = \text{diag}\{0.5I_2, 1, 0.5I_3\}$ . The initial estimates are given by  $\hat{q}(0) = [\frac{\sqrt{2}}{2}, \frac{\sqrt{6}}{4}, \frac{\sqrt{2}}{4}, 0]$ ,  $\hat{V}(0) = [2, -1.5, -0.5]^{\top}$  (m/s),  $\hat{s}(0) = 4$  (m $^{-1}$ ), whereas the initial state values are  $R(0) \approx I_3$ ,  $V(0) \approx 0$  (m/s),  $s(0) \approx 5$  (m $^{-1}$ ). Note that the attitude error corresponds to relatively large Euler angle errors of  $90(deg)$ ,  $30(deg)$  and  $30(deg)$  approximately in roll, pitch and yaw.

The time evolutions of the estimated and ground-truth attitudes, represented by Euler angles, along with the estimated and ground-truth body-fixed linear velocities and depths are shown in Figs. 4.1–4.3, respectively. During the first 6 seconds, the quadrotor is motionless on the ground, leading to the violation of the persistent excitation condition (3). In contrast to the depth estimate  $\hat{d}$  (or equivalently  $\hat{s}$ ) that does not converge to the ground-truth value due to the lack of excitations, both the estimated attitude and linear velocity still converge exponentially near to the corresponding ground-truth values during that time period despite the large initial estimated errors. This is an interesting and desirable feature of the proposed observer although the convergence and observability analysis for this particular “unobservable” case (i.e.  $|V| \equiv 0$ ) remains open. It can be observed from Fig. 4.4 that during that period the norm of the Riccati matrix  $P$  grows almost linearly that in turn highlights the discussed preoccupation and its associated solutions in Section 4.4.3.

After the second 6 when the UAV takes off and carries out the circular flight (i.e. the persistent excitation condition (3) is guaranteed), all the estimated variables  $\hat{R}$ ,  $\hat{V}$ ,  $\hat{s}$  converge almost perfectly to the corresponding ground-truth data (see Figs. 4.1–4.3) and the norm of  $P$  remains bounded and varying around a constant number ( $\approx 0.9$ ) (see Fig. 4.4). In summary, we find that the performance of the proposed observer is quite

satisfactory.

## 4.4.5 Experimental evaluation

### 4.4.5.1 Experimental setup

For experimental validations, we make use of a Visual-Inertial (VI) sensor developed by the Autonomous Systems Lab (ETH Zurich) and the company Skybotix. Among the two cameras and two IMUs of the VI-sensor, only one camera and one IMU (composed of a 3-axis gyrometer and a 3-axis accelerometer) are used to validate the proposed algorithm. The main reason for using the VI-sensor in this experimental setup is the possibility of obtaining perfectly time-synchronized images and IMU readings ( $20Hz$  for the camera and  $200Hz$  for the IMU). On the other hand, the OptiTrack motion capture system available at I3S is used to obtain the ground truth data for comparison purposes. This highly accurate OptiTrack system provides the full pose of the Camera-IMU system at  $120Hz$ . As a matter of fact, the quantities such as  $\frac{V}{d}$  and  $\phi_{\perp}$  can be emulated from the ground-truth pose measurements. For instance, using the position measurements a high-gain observer is applied for the estimation of the linear velocity expressed in the inertial frame that is then converted to the linear velocity expressed in the body-fixed frame using the ground-truth attitude. Then, by considering the situation where the UAV carries a downward-looking camera to observe a planar horizontal ground, it is not difficult to construct the “measurements” of  $\frac{V}{d}$  and  $\phi_{\perp}$  (at  $20Hz$ ).

The hybrid discrete-continuous version of the proposed observer has been implemented in C++, combined with OpenCV for image processing, on an Intel Core i7-6400 CPU running at 3.40Ghz. The transmission of data from the camera to the PC is carried out through a high speed ethernet cable. The PC has a Linux based operating system and is responsible for two major software tasks:

- Interface with the camera hardware and acquisition of images and IMU data from the VI-sensor.
- Estimation of the continuous homography based on two consecutive images which is then decomposed to obtain the measurements of  $\frac{V}{d}$  and  $\phi_{\perp}$  in real-time.

Due to real-time constraint for the continuous homography estimation, feature detection and descriptor extraction in images are carried out using the FAST Feature Detector<sup>3</sup> and ORB Descriptor Extractor algorithms already implemented in the OpenCV library. Since the quality of the continuous homography estimation depends heavily

---

<sup>3</sup>Although FAST is less robust than other algorithms such as SIFT or SURE, it is much faster and more suitable for real-time implementation.

#### 4.4. Practical implementation aspects

on the capability of rejecting outliers of point matchings, we have implemented an M-estimator-like observer for the estimation of the homography between every two consecutive images, which is then used to compute the continuous homography via a logarithm conversion. This M-estimator-like homography observer is a modified version of the homography observer proposed in [34] but is not presented here due to space limitation. However, the reader can appreciate its performance and robustness via the following video link:

<https://youtu.be/x75RpjoJ9HM>

Although the combined implementation of image processing and estimation algorithm runs at about  $50Hz$ , the continuous homography estimate is only obtained at every  $50ms$  ( $20Hz$ ) due to the lower acquisition frequency of the VI-sensor camera.

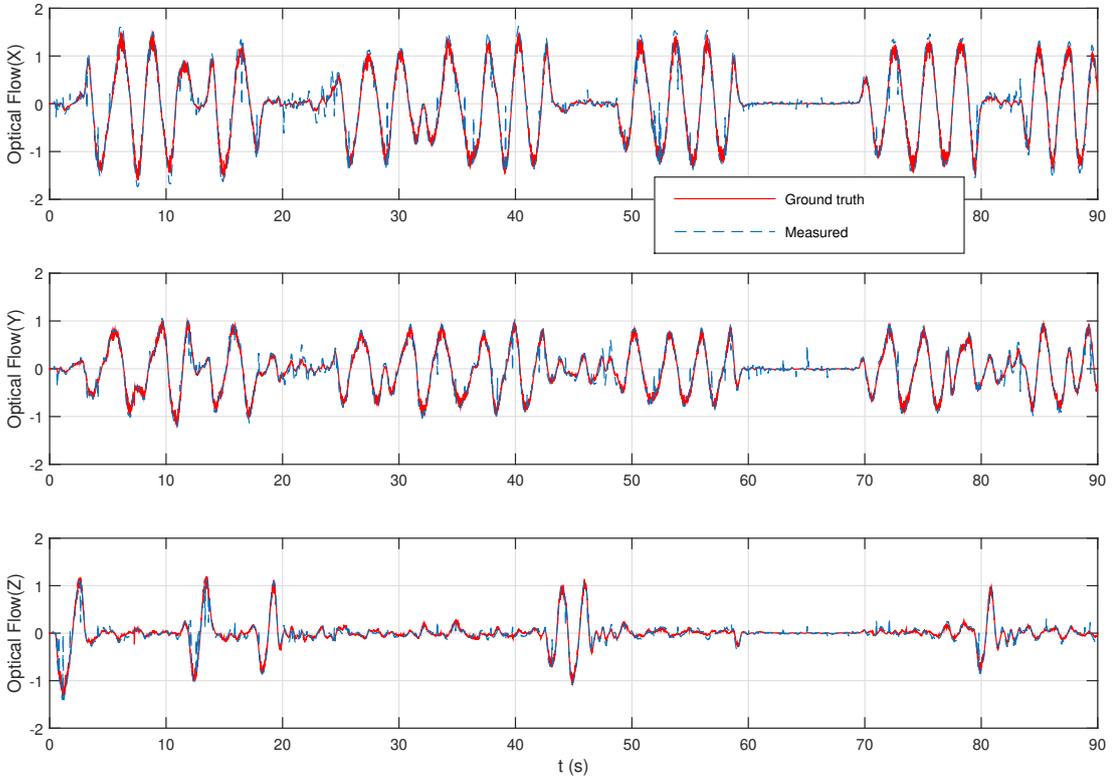


Figure 4.5: (Experiment) Optical flow components measured from images (blue curves) and derived from ground truth pose (red curves) versus time ( $s$ )

The parameters involved in the proposed observer are chosen as follows:  $D = \text{diag}\{8, 8, 24\}$ ,  $S = \text{diag}\{0.02^2 I_2, 0.1^2, 0.2^2 I_3\}$  and  $P(0) = 1.7I_6$ . The initial estimates are given by  $\hat{q}(0) = [1, 0, 0, 0]$ ,  $\hat{V}(0) = [0, 0, 0](m/s)$ ,  $\hat{s}(0) = 4(m^{-1})$ .

Figs. 4.5 and 4.6 show a good quality of the optical flow  $\frac{V}{d}$  as well as  $\phi_{\perp}(= -\frac{\dot{d}}{d})$  obtained from the decomposition of the continuous homography estimate when compared

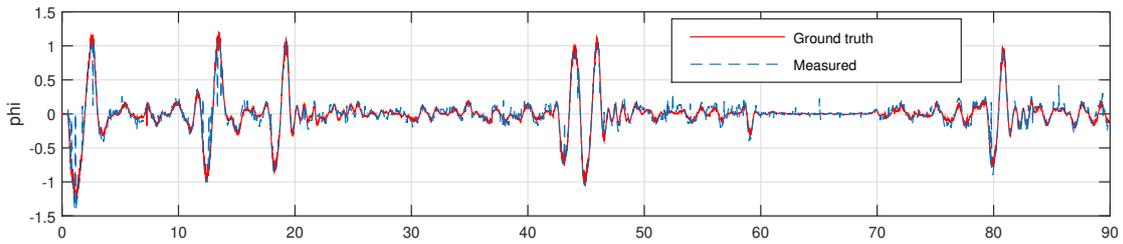


Figure 4.6: (Experiment)  $\phi_{\perp} = -\frac{d}{a}$  measured from images (blue curve) and derived from ground truth pose (red curve) versus time ( $s$ )

to the corresponding ground-truth data.

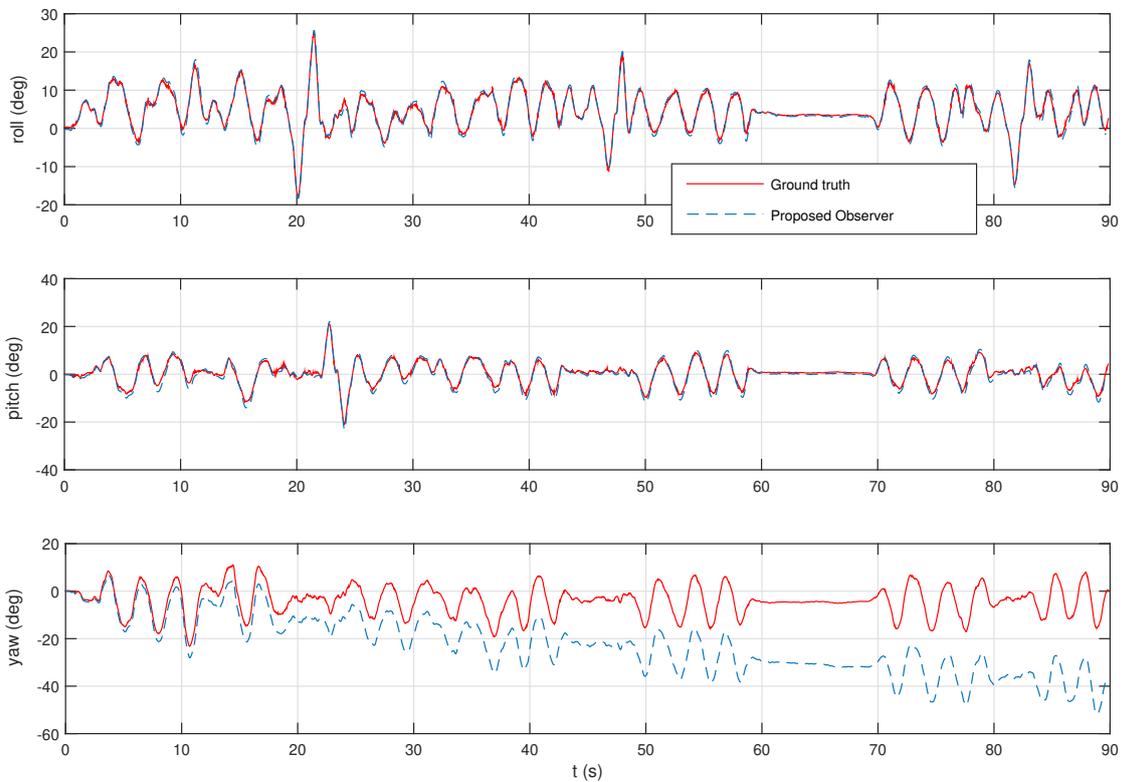


Figure 4.7: (Experiment) Estimated and ground-truth attitudes represented by roll, pitch and yaw Euler angles ( $deg$ ) versus time ( $s$ )

#### 4.4.5.2 Experimental results

The reported experiment has been performed online with the VI-sensor camera looking downward to observe a well textured planar horizontal ground. A demo video is provided as a supplemental material and is also available at

<https://youtu.be/R09oTjr4s40>

#### 4.4. Practical implementation aspects

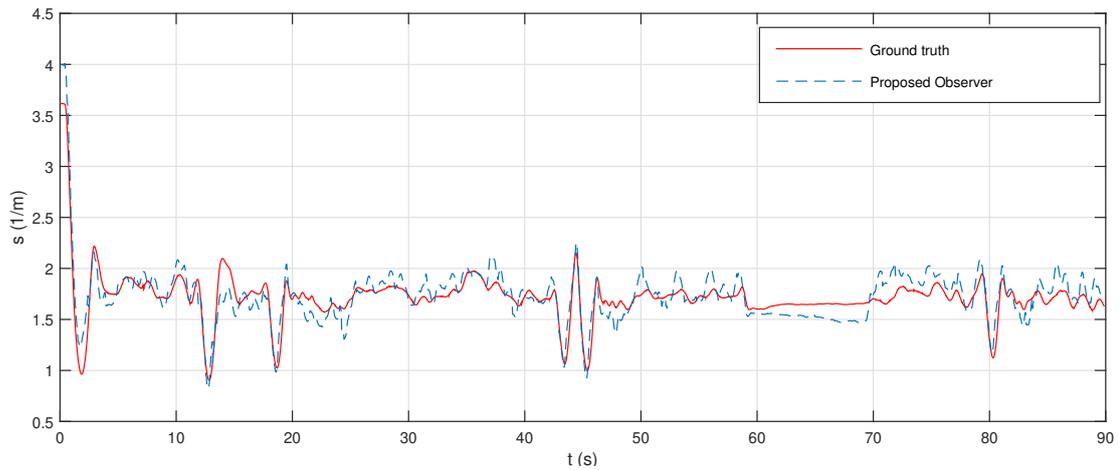


Figure 4.8: (Experiment) Estimated and ground-truth depth inverse ( $m^{-1}$ ) versus time

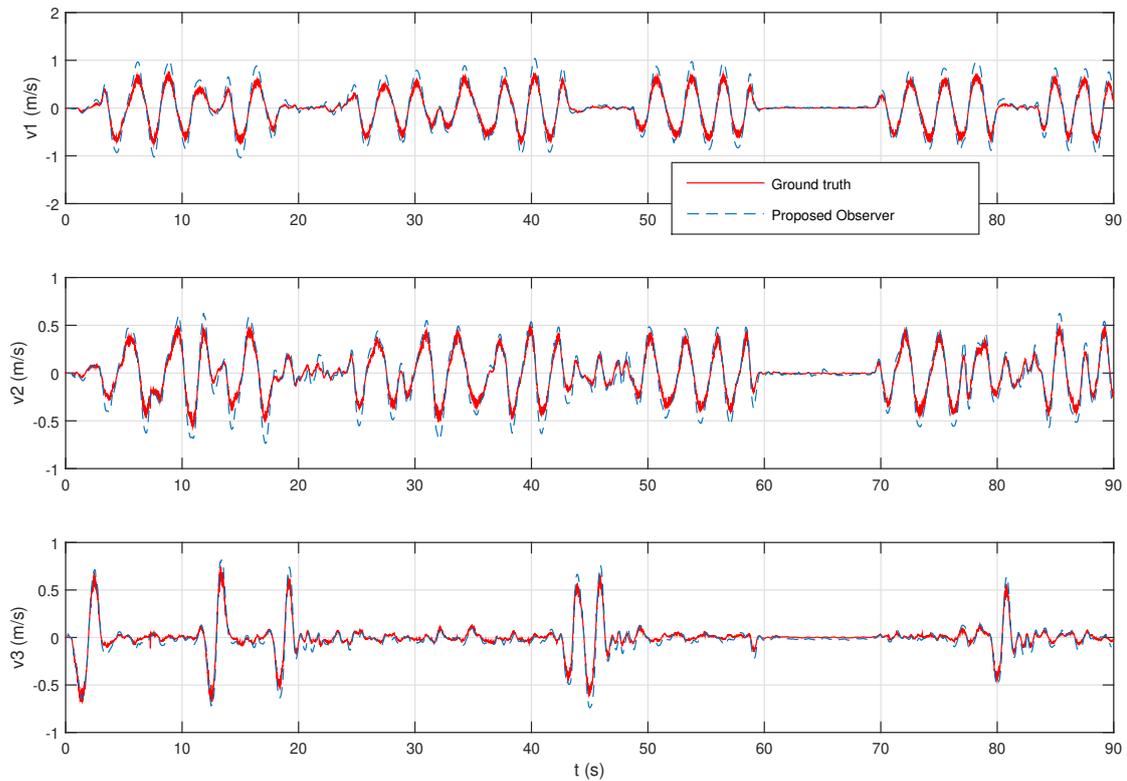


Figure 4.9: (Experiment) Estimated and ground-truth linear velocity components in body-fixed frame ( $m/s$ ) versus time ( $s$ )

The time evolutions of the estimated and ground-truth attitudes, body-fixed linear velocities and depths are depicted in Figs. 4.7–4.8. During the first 60 seconds, it can be observed that the depth estimate as well as both the estimated gravity direction (i.e.

roll and pitch Euler angles) and the estimated linear velocity converge near to the corresponding ground-truth values since the condition of persistent excitation is preserved. In contrast, during the period of 60 to 69 seconds when the camera is kept still thus violating the condition of persistent excitation, it can be seen that the depth estimate slightly drifts away from the ground-truth depth, whereas both the estimated gravity direction and linear velocity always remain close to the corresponding ground-truth values. Once the condition of persistent excitation is revoked again by moving the camera from 69 to 89 seconds, the depth estimate follows closely again the ground-truth value along with the gravity direction and linear velocity estimates. From Fig. 4.7 it can also be observed that the yaw angle estimate drifts away from the ground-truth value. This is normal since it is simply an integration of the gyros (i.e. magnetometer measurements are not used for yaw estimation correction in this experiment). In conclusion, the reported experiment shows that whereas the (practical) convergence of the gravity direction and linear velocity estimates is always achieved, the convergence of the depth estimate is additionally obtained only when the condition of persistent excitation is guaranteed.

## 4.5 Extension to observer design for optical flow filtering

The design of the observer follows exactly in the same manner like in the previous section, the only difference here being that instead of estimating directly the linear velocity we filter the optical flow extracted from the continuous homography. One obtains the dynamics for  $\phi = \frac{V}{d}$  as follows:

$$\begin{aligned}\dot{\phi} &= \frac{\dot{V}}{d} - \frac{V\dot{d}}{d^2} \\ &= s\dot{V} + \phi_{\perp}\phi\end{aligned}$$

where  $s = \frac{1}{d}$  and  $\phi_{\perp} = \frac{-\dot{d}}{d}$ . Now using the dynamics of  $V$  from (4.2) one obtains

$$\begin{aligned}\dot{\phi} &= -\Omega_{\times}\phi + sa_{\mathcal{B}} + gsR^{\top}e_3 + \phi_{\perp}\phi \\ &= (-\Omega_{\times} + \phi_{\perp}I_3)\phi + sa_{\mathcal{B}} + gsR^{\top}e_3\end{aligned}\tag{4.29}$$

### 4.5.1 System Equations and Measurements

Thus from (4.29), (4.4) and (4.1) the rigid-body dynamics for the system and the measurements for the observer are given below:

$$\begin{cases} \dot{R} = R\Omega_\times \\ \dot{\phi} = -(\Omega_\times + \phi_\perp I_3)\phi + sa_{\mathcal{B}} + gsR^\top e_3 \\ \dot{s} = \phi_\perp s \end{cases} \quad (4.30)$$

### 4.5.2 Observer derivation

Let  $\hat{R} \in \text{SO}(3)$ ,  $\hat{\phi} \in \mathbb{R}^3$ ,  $\hat{s} \in \mathbb{R}$  denote the estimates of  $R$ ,  $\phi$ ,  $s$ , respectively. The proposed observer is given by

$$\begin{cases} \dot{\hat{R}} = \hat{R}\Omega_\times - [\sigma_R]_\times \hat{R} \\ \dot{\hat{\phi}} = -(\Omega_\times + \phi_\perp I_3)\hat{\phi} + \hat{s}a_{\mathcal{B}} + g\hat{s}\hat{R}^\top e_3 - \sigma_\phi \\ \dot{\hat{s}} = \phi_\perp \hat{s} - \sigma_s \end{cases} \quad (4.31)$$

where  $\sigma_R, \sigma_v \in \mathbb{R}^3$ ,  $\sigma_s \in \mathbb{R}$  are innovation terms to be designed thereafter. Defining the observer errors

$$\tilde{R} \triangleq R\hat{R}^\top, \quad \tilde{\phi} \triangleq \phi - \hat{\phi}, \quad \tilde{s} \triangleq s - \hat{s}$$

then the observer's objective can be stated as the exponential stability of  $(\tilde{R}, \tilde{\phi}, \tilde{s}) = (I_3, 0, 0)$  (or of  $(\tilde{R}e_3, \tilde{\phi}, \tilde{s}) = (e_3, 0, 0)$  when the estimation of the gravity direction is concerned instead of the whole attitude estimation). From (4.30) and (4.31), one verifies that the dynamics of  $(\tilde{R}, \tilde{\phi}, \tilde{s})$  satisfy

$$\begin{cases} \dot{\tilde{R}} = \tilde{R}[\sigma_R]_\times \\ \dot{\tilde{\phi}} = -(\Omega_\times + \phi_\perp I_3)\tilde{\phi} + \tilde{s}(a_{\mathcal{B}} + g\hat{R}^\top e_3) + g\hat{s}\hat{R}^\top(\tilde{R}^\top - I_3)e_3 + \sigma_\phi \\ \dot{\tilde{s}} = \phi_\perp \tilde{s} + \sigma_s \end{cases} \quad (4.32)$$

We will work out next first order approximations of the error system (4.32) complemented with first order approximations of the measurement equations. The first order approximations of the attitude error equation is carried out exactly in the same manner as for the previous observer and is given as

For the dynamics of  $\tilde{\phi}$  one obtains

$$\begin{aligned} \dot{\tilde{\phi}} &= -(\Omega_\times + \phi_\perp I_3)\tilde{\phi} + \tilde{s}(a_{\mathcal{B}} + g\hat{R}^\top e_3) + g\hat{s}\hat{R}^\top[e_3]_\times \tilde{\lambda} + \sigma_\phi + O(|\tilde{\lambda}|^2) \\ &= -(\Omega_\times + \phi_\perp I_3)\tilde{\phi} + \tilde{s}(a_{\mathcal{B}} + g\hat{R}^\top e_3) + g\hat{s}\hat{R}^\top e_2 \tilde{\lambda}_1 - g\hat{s}\hat{R}^\top e_1 \tilde{\lambda}_2 + \sigma_\phi \end{aligned}$$

with  $\tilde{\lambda}_1, \tilde{\lambda}_2$  the first and second components of  $\tilde{\lambda}$ .

By setting the system output vector equal to

$$y = \phi - \hat{\phi} \quad (4.33)$$

one obtains LTV first order approximations in the form (2.20) with

$$\left\{ \begin{array}{l} x = \begin{bmatrix} \tilde{\lambda}_1 \\ \tilde{\lambda}_2 \\ \tilde{s} \\ \tilde{\phi} \end{bmatrix}, \quad x_1 = \begin{bmatrix} \tilde{\lambda}_1 \\ \tilde{\lambda}_2 \end{bmatrix}, \quad x_2 = \begin{bmatrix} \tilde{s} \\ \tilde{\phi} \end{bmatrix}, \quad u = \begin{bmatrix} \sigma_{R,1} \\ \sigma_{R,2} \\ \sigma_s \\ \sigma_\phi \end{bmatrix} \\ A = \begin{bmatrix} 0 & 0 & 0 & 0_{1 \times 3} \\ 0 & 0 & 0 & 0_{1 \times 3} \\ 0 & 0 & \phi_\perp & 0_{1 \times 3} \\ g\hat{s}\hat{R}^\top e_2 & -g\hat{s}\hat{R}^\top e_1 & a_B + g\hat{R}^\top e_3 & -\Omega_\times + \phi_\perp I_3 \end{bmatrix} \in \mathbb{R}^{6 \times 6} \\ C = \begin{bmatrix} 0_{3 \times 1} & 0_{3 \times 1} & 0_{3 \times 1} & I_3 \end{bmatrix} \in \mathbb{R}^{3 \times 6} \end{array} \right. \quad (4.34)$$

with  $\sigma_{R,1}, \sigma_{R,2}$  the first and second components of  $\sigma_R$ . From there the proposed observer is given by (5.11) with  $\sigma_{R,1}, \sigma_{R,2}$  and  $\sigma_\phi$  determined from the input  $u$  calculated according to (2.22) and (2.23).

As for the innovation component  $\sigma_{R,3}$ , it can be independently designed for estimating the remaining degree of freedom of the attitude (i.e. yaw). For instance, without loss of generality  $\sigma_{R,3}$  is assumed to be bounded for all time.

**Remark 2.** *The hybrid discrete-continuous version of this observer as well as the part of yaw estimation proceeds in exactly the same manner as described in section 4.4.2 and subsection 4.3.3.*

### 4.5.3 Observability analysis

The observability analysis proceeds in a similar fashion to the one shown in section 4.3.2. The exponential stability of the proposed observer depends on the uniform observability of the pair  $(A^*(t), C^*(t))$  obtained by setting  $x = 0$  in the expression of the matrices  $A$  and  $C$  given by (4.34). Thus, one verifies that,

$$\left\{ \begin{array}{l} A^* = \begin{bmatrix} 0 & 0 & 0 & 0_{1 \times 3} \\ 0 & 0 & 0 & 0_{1 \times 3} \\ 0 & 0 & \phi_\perp & 0_{1 \times 3} \\ g\hat{s}\hat{R}^{*\top} e_2 & -g\hat{s}\hat{R}^{*\top} e_1 & \hat{R}^\top \dot{v} & -\Omega_\times + \phi_\perp I_3 \end{bmatrix} \\ C^* = \begin{bmatrix} 0_{3 \times 1} & 0_{3 \times 1} & 0_{3 \times 1} & I_3 \end{bmatrix} \end{array} \right. \quad (4.35)$$

#### 4.5. Extension to observer design for optical flow filtering

---

with  $s = \frac{1}{d}$ ,  $\phi_\perp = -\dot{d}/d$ , and  $R^* \in \text{SO}(3)$  satisfying

$$R^{*\top} e_3 = R^\top e_3, \quad \dot{R}^* = R^* \Omega_\times$$

Also the expression  $a_B + gR^\top e_3$  which is the component  $A(4, 3)$  can be replaced by

$$\begin{aligned} a_B + g\hat{R}^\top e_3 &= \dot{V} + \Omega_\times V \\ &= R^\top \dot{v} \end{aligned} \quad (4.36)$$

which is thereby the component  $A^*(4, 3)$  in (4.35).

**Proposition 4.** *The transient matrix  $\Phi(\tau, t)$  associated with  $A^*(t)$ , for all  $\tau \geq t$ , is given by*

$$\Phi(\tau, t) = \begin{bmatrix} 1 & 0 & 0 & 0_{1 \times 3} \\ 0 & 1 & 0 & 0_{1 \times 3} \\ 0 & 0 & \frac{s(\tau)}{s(t)} & 0_{1 \times 3} \\ gs(\tau)(\tau - t)R(\tau)^\top R_\gamma e_1 & gs(\tau)(\tau - t)R(\tau)^\top R_\gamma e_2 & \frac{s(\tau)R(\tau)^\top [v(\tau) - v(t)]}{s(t)} & \frac{s(\tau)R(\tau)^\top R(t)}{s(t)} \end{bmatrix} \quad (4.37)$$

where  $R_\gamma, \gamma_{1,2}$  are defined in (4.24) and (4.23) respectively.

**Proof:** The transient matrix  $\Phi(\tau, t)$  is the solution to the following equation

$$\frac{d\Phi(\tau, t)}{d\tau} = A^*(\tau)\Phi(\tau, t), \quad \Phi(t, t) = I_6 \quad (4.38)$$

From the above relation and the expression of  $A^*(\tau)$ , one easily verifies that  $\Phi(\tau, t)$  has the form

$$\Phi(\tau, t) = \begin{bmatrix} 1 & 0 & 0 & 0_{1 \times 3} \\ 0 & 1 & 0 & 0_{1 \times 3} \\ 0 & 0 & \varphi_{33}(\tau, t) & 0_{1 \times 3} \\ \varphi_{41}(\tau, t) & \varphi_{42}(\tau, t) & \varphi_{43}(\tau, t) & \varphi_{44}(\tau, t) \end{bmatrix} \quad (4.39)$$

Now we need to compute  $\varphi_{33}(\tau, t) \in \mathbb{R}$ ,  $\varphi_{41}(\tau, t) \in \mathbb{R}^3$ ,  $\varphi_{42}(\tau, t) \in \mathbb{R}^3$ ,  $\varphi_{43}(\tau, t) \in \mathbb{R}^3$ ,  $\varphi_{44}(\tau, t) \in \mathbb{R}^{3 \times 3}$ . The expression of  $\varphi_{33}(\tau, t)$  is exactly the same as shown in (4.13). From (4.38) and (4.39), one has

$$\frac{d\varphi_{41}(\tau, t)}{d\tau} = (-\Omega(\tau)_\times + \phi_\perp(\tau)I_3)\varphi_{41}(\tau, t) + gs(\tau)R^*(\tau)^\top e_2$$

Then, by change of variable  $\bar{\varphi}_{41}(\tau, t) := \frac{R^*(\tau)\varphi_{41}(\tau, t)}{s(\tau)}$  the above equation yields

$$\begin{aligned}
 \frac{d\bar{\varphi}_{41}(\tau, t)}{d\tau} &= \frac{-\dot{s}(\tau)R^*(\tau)\varphi_{41}(\tau)}{s^2(\tau)} + \frac{R^*(\tau)\Omega(\tau)_{\times}\varphi_{41}(\tau)}{s(\tau)} + \frac{R^*(\tau)(gs(\tau)R^*e_2 + (-\Omega(\tau)_{\times} + \phi(\tau)_{\perp}I_3))\varphi_{41}(\tau)}{s(\tau)} \\
 &= \frac{-\phi(\tau)_{\perp}s(\tau)R^*(\tau)\varphi_{41}(\tau)}{s(\tau)^2} + \frac{R^*(\tau)\Omega(\tau)_{\times}\varphi_{41}(\tau)}{s(\tau)} + ge_2 - \frac{R^*(\tau)\Omega(\tau)_{\times}\varphi_{41}(\tau)}{s(\tau)} + \frac{R^*(\tau)\phi(\tau)_{\perp}I_3\varphi_{41}(\tau)}{s(\tau)} \\
 &= -\phi(\tau)_{\perp}\bar{\varphi}_{41}(\tau) + ge_2 + \phi(\tau)_{\perp}\bar{\varphi}_{41}(\tau) \\
 &= ge_2
 \end{aligned} \tag{4.40}$$

Thus we obtain

$$\bar{\varphi}_{41}(\tau, t) = g(\tau - t)e_2 \tag{4.41}$$

from which one deduces

$$\begin{aligned}
 \varphi_{41}(\tau, t) &= s(\tau)R^*(\tau)^{\top}\bar{\varphi}_{41}(\tau, t) \\
 &= gs(\tau)(\tau - t)R(\tau)^{\top}\gamma_2 \\
 &= gs(\tau)(\tau - t)R(\tau)^{\top}R_{\gamma}e_1
 \end{aligned} \tag{4.42}$$

with  $\gamma_2$  defined in (4.23). Proceeding in a similar fashion, one has

$$\frac{d\varphi_{42}(\tau, t)}{d\tau} = (-\Omega(\tau)_{\times} + \phi_{\perp}(\tau)I_3)\varphi_{42}(\tau, t) - gs(\tau)R^*(\tau)^{\top}e_1$$

Then by change of variable  $\bar{\varphi}_{42}(\tau, t) := \frac{R^*(\tau)\varphi_{42}(\tau, t)}{s(\tau)}$  the above equation yields

$$\frac{d\bar{\varphi}_{42}(\tau, t)}{d\tau} = -ge_1$$

after which one obtains the expression for  $\bar{\varphi}_{42}(\tau, t)$  as follows

$$\bar{\varphi}_{42}(\tau, t) = -g(\tau - t)e_1$$

Then from the above equation one deduces  $\varphi_{42}(\tau, t)$

$$\begin{aligned}
 \varphi_{42}(\tau, t) &= s(\tau)R^*(\tau)^{\top}\bar{\varphi}_{42}(\tau, t) \\
 &= -gs(\tau)(\tau - t)R(\tau)^{\top}\gamma_1 \\
 &= gs(\tau)(\tau - t)R(\tau)^{\top}R_{\gamma}e_2
 \end{aligned} \tag{4.43}$$

with  $\gamma_1$  defined in (4.23). From (4.38) and (4.39), one has

$$\frac{d\varphi_{43}(\tau, t)}{d\tau} = R(\tau)^{\top}\dot{v}(\tau)\varphi_{33}(\tau) + (-\Omega(\tau)_{\times} + \phi_{\perp}(\tau)I_3)\varphi_{43}(\tau, t)$$

#### 4.5. Extension to observer design for optical flow filtering

---

Then, by change of variable  $\bar{\varphi}_{43}(\tau, t) := \frac{R^*(\tau)\varphi_{43}(\tau, t)}{s(\tau)}$  and from (4.13) the above equation yields

$$\begin{aligned}\frac{d\bar{\varphi}_{43}(\tau, t)}{d\tau} &= \frac{R^*(\tau)R(\tau)^\top \dot{v}(\tau)s(\tau)}{s(\tau)s(t)} \\ &= \frac{R^*(\tau)R(\tau)^\top \dot{v}(\tau)}{s(t)}\end{aligned}$$

Since  $R^*(\tau)R(\tau)^\top = R^*(0)R(0)^\top$  is constant, one obtains

$$\bar{\varphi}_{43}(\tau, t) = \frac{R^*(\tau)R(\tau)^\top (v(\tau) - v(t))}{s(t)} \quad (4.44)$$

from which one deduces

$$\begin{aligned}\varphi_{43}(\tau, t) &= s(\tau)R^*(\tau)^\top \bar{\varphi}_{43}(\tau, t) \\ &= \frac{s(\tau)R(\tau)^\top [v(\tau) - v(t)]}{s(t)}\end{aligned} \quad (4.45)$$

Now again from (4.38) and (4.39), one has

$$\frac{d\varphi_{44}(\tau, t)}{d\tau} = (-\Omega(\tau)_\times + \phi_\perp(\tau)I_3)\varphi_{44}(\tau, t)$$

Then, by change of variable  $\bar{\varphi}_{44}(\tau, t) = \frac{R^*(\tau)\varphi_{44}(\tau, t)}{s(\tau)}$  the above equation yields

$$\begin{aligned}\frac{d\bar{\varphi}_{44}(\tau, t)}{d\tau} &= \frac{-\dot{s}(\tau)R^*(\tau)\varphi_{44}(\tau)}{s^2(\tau)} + \frac{R^*(\tau)\Omega(\tau)_\times\varphi_{44}(\tau)}{s(\tau)} + \frac{R^*(\tau)(-\Omega(\tau)_\times + \phi(\tau)_\perp I_3)\varphi_{44}(\tau)}{s(\tau)} \\ &= \frac{-\phi(\tau)_\perp R^*(\tau)\varphi_{44}(\tau)}{s(\tau)} + \frac{R^*(\tau)\Omega(\tau)_\times\bar{\varphi}_{44}(\tau)}{s(\tau)} - \frac{R^*(\tau)\Omega(\tau)_\times\varphi_{44}(\tau)}{s(\tau)} + \frac{R^*(\tau)\phi(\tau)_\perp I_3\varphi_{44}(\tau)}{s(\tau)} \\ &= -\phi(\tau)_\perp \bar{\varphi}_{44}(\tau) + \phi(\tau)_\perp \bar{\varphi}_{44}(\tau) \\ &= 0\end{aligned} \quad (4.46)$$

Thus we obtain

$$\bar{\varphi}_{44}(\tau, t) = s(\tau)R^*(\tau)^\top \varphi_{44}(\tau, t) \quad (4.47)$$

From the above equation one deduces

$$\begin{aligned}\varphi_{44}(\tau, t) &= s(\tau)R^*(\tau)^\top \bar{\varphi}_{44}(\tau, t) \\ &= \frac{s(\tau)R(\tau)^\top R(t)}{s(t)}\end{aligned} \quad (4.48)$$

Finally, with the relations (4.13), (4.42), (4.43), (4.45) and (4.48) one deduces the explicit form (4.37) of the transient matrix  $\Phi(\tau, t)$  (**End of Proof**).

**Proposition 5.** *The persistent excitation conditions C.1 and C.2 defined in Proposition (3) in section 4.3.2 are the same, however with the matrices  $U(\tau)$  and  $W(\tau)$  defined as follows:*

$$U(\tau) := s(\tau) \begin{bmatrix} g(\tau - t)e_1 & g(\tau - t)e_2 & R_\gamma^\top[v(\tau) - v(t)] & I_3 \end{bmatrix} \quad (4.49)$$

$$W(\tau) := \begin{bmatrix} g(\tau - t)e_1 & g(\tau - t)e_2 & R_\gamma^\top v(\tau) \end{bmatrix} \quad (4.50)$$

The proof for proving that the pair  $(A^*(t), C^*(t))$  is uniformly observable proceeds in a similar manner again as shown in section 4.3.2. From the proof in section 4.3.2 we know that it is needed only to prove (4.21) for the case where the persistent excitation condition (C.1) is satisfied. Accordingly, one verifies that

$$C^*(\tau)\Phi(\tau) = s(\tau)R(\tau)^\top R_\gamma \begin{bmatrix} g(\tau - t)e_1 & g(\tau - t)e_2 & \frac{R_\gamma^\top[v(\tau) - v(t)]}{s(t)} & \frac{R_\gamma^\top R(t)}{s(t)} \end{bmatrix} \quad (4.51)$$

Thus for all  $x = [x_1, x_2, x_3, x_4^\top]^\top$ , with  $x_{1,2,3} \in \mathbb{R}$ ,  $x_4 \in \mathbb{R}^3$ , using (4.51) one deduces

$$\begin{aligned} C^*(\tau)\Phi(\tau)x &= s(\tau)R(\tau)^\top R_\gamma \begin{bmatrix} g(\tau - t)e_1 & g(\tau - t)e_2 & \frac{R_\gamma^\top[v(\tau) - v(t)]}{s(t)} & I_3 \end{bmatrix} \bar{x} \\ &= R(\tau)^\top R_\gamma U(\tau)\bar{x} \end{aligned}$$

with  $\bar{x} := [x_1, x_2, \frac{x_3}{s(t)}, \frac{R_\gamma^\top R(t)x_4}{s(t)}]^\top$  (which ensures that  $|\bar{x}| \geq \min(1, d_{min})|x|$ ) and  $U(\tau)$  defined by (4.49). One then obtains

$$\begin{aligned} x^\top \Phi(\tau)^\top C^{*\top}(\tau)C^*(\tau)\Phi(\tau)x &= \bar{x}^\top U(\tau)^\top R_\gamma^\top R(\tau)R(\tau)^\top R_\gamma U(\tau)\bar{x} \\ &= \bar{x}^\top U(\tau)^\top U(\tau)\bar{x} \end{aligned} \quad (4.52)$$

From (4.52) and the persistent excitation condition C.1, one straightforwardly deduces inequality (4.21), with  $\mu = (\min(1, d_{min}))^2 \rho$ , which concludes the proof (**End of proof**).

#### 4.5.4 Experimental Results

As seen in the section 3.6.1 the optical flow estimates were quite noisy. This motivated us to the development of the nonlinear observer 4.31 that filters out these noisy estimates. In this section we use the same data set as used in section 3.6.1, the only difference being the observer for homography estimation is the one taken from [34] and not the one proposed in [55]. Using the position measurements obtained from the Optitrack system a high-gain observer is applied for the estimation of the optical flow. The figure 4.10

#### 4.5. Extension to observer design for optical flow filtering

---

shows the comparison between the optical flow obtained from the mocap optitrack system (in red), the raw optical flow estimates extracted from continuous homography (in green) and the filtered values of these estimates (in blue) from the proposed observer. The ground truth shown by the red curve is slightly noisy as well due to high gain observers used for its estimation. It can be seen that as soon as there is a relatively large displacement of the camera the estimates are quite good but are noisy thereby demonstrating the need of a filter. The effect of the filtering can be clearly seen in the Figure 4.11 when there is minimal amount of displacement, thus making the optical flow estimates suitable for exploitation in control strategies. It can also be noticed that the filter does not introduce any significant delay in the system. The experimental results thus show the effectiveness of the proposed observer.

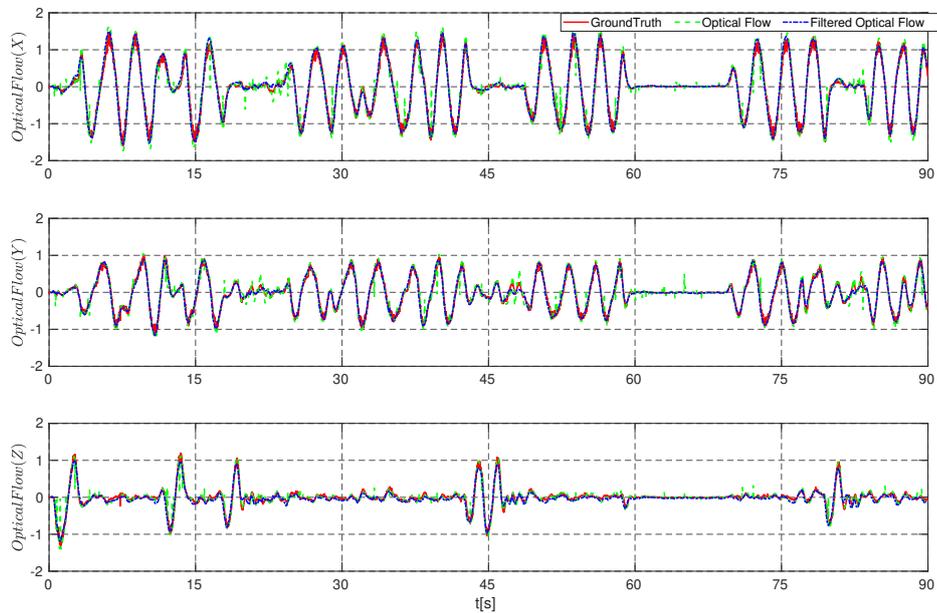


Figure 4.10: Optical flow components estimated from images (green curves), filtered optical flow components from the proposed observer (blue curves) and derived from ground truth pose (red curves) versus time ( $s$ )

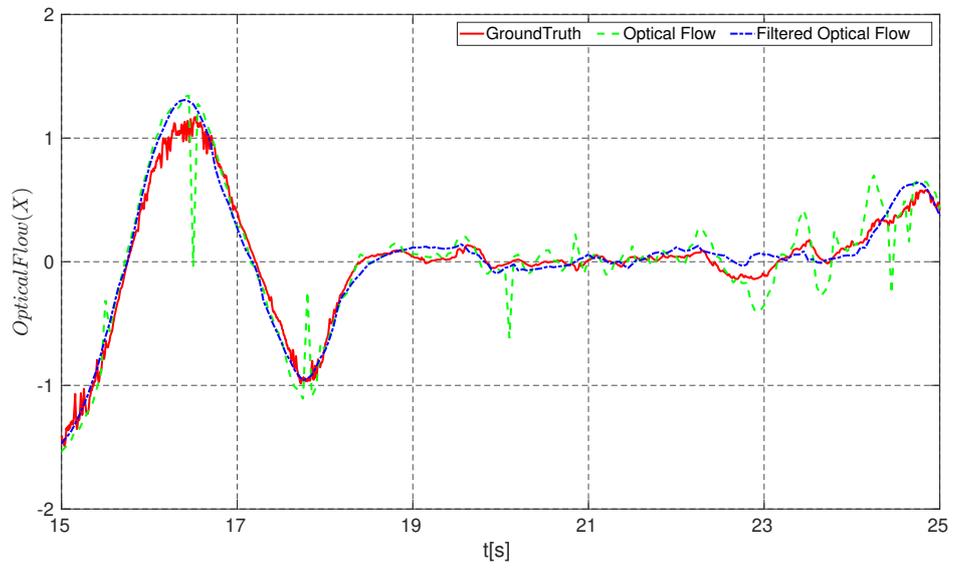


Figure 4.11: Zoom of the optical flow components from time  $t = 15s$  to  $t = 25s$

# 5

## Homography Decomposition

### 5.1 Introduction

The notion of homography and various techniques used for homography estimation have been already detailed in Chapter 3. We briefly recall some of the applications mentioned in Chapter 3 which will serve as motivation for the work presented in this chapter. Homography has been instrumental in various computer vision and robotic applications where the scene involves (near) planar surfaces as in the case of man-made environments or in the case of an unmanned aerial vehicle flying sufficiently far from the observed ground. Homography has been used to perform image stabilization [34] and image mosaicing [13, 69]. Homography has also been exploited for the estimation of the relative pose (up to a scale factor) of a robotic vehicle equipped with a camera [14, 15, 59, 71]. Homography has been widely used as visual feedback in robot control [12, 60, 52, 5] and one of the most successful visual servo control paradigms is the  $2\frac{1}{2}$ D approach [52] that relies on the extraction of the camera displacement (i.e. orientation and translation up to a scale factor) and the scene's normal vector from the Euclidean homography. Such a process of extraction is referred to as *Euclidean homography decomposition* or *Euclidean reconstruction from homography*.

Classical approaches for homography decomposition such as Faugeras SVD-based [16], Zhang SVD-based [77] algorithms use singular value decomposition to obtain numerical solutions. Malis and Vargas [53] lately solved the homography decomposition problem with an analytical approach, making it more suitable for real-time robot control applications. However, to our knowledge all existing homography decomposition methods belong to the “algebraic category” that only focuses on solving the homography decomposition problem on a frame-by-frame basis, but not on filtering measurement noise. The precision of the decomposition elements is thus highly prone to noise, especially when the camera's translation is small [53]. The problem even becomes de-

generate when the latter vanishes. Robotic vehicle applications, however, provide temporal sequences of images together with inertial measurements from, e.g., an embedded Inertial Measurement Unit (IMU). It, thus, seems natural to exploit the temporal correlation rather than to solve the homography decomposition problem for each pair of image frames.

In the current chapter, a novel direction for solving the homography decomposition by exploiting the system dynamics is explored. The proposed solutions are developed in the form of nonlinear observers derived from the recent deterministic Riccati observer design framework proposed in [24]. We believe that the resulting estimated quantities would be less noisy since the noise can be filtered through a natural low-pass response of the observer. Moreover, the challenging theoretical issue related to the uniform observability, under which local exponential stability is granted, has been carefully addressed. The proposed “persistence of excitation” conditions even cover the degenerate situation for which all existing algebraic algorithms fail to provide the correct solution. Finally, simulation and experimental results demonstrate a good performance and a large attraction domain of the proposed observers.

The chapter is organized as follows. Section 5.2.1 talks about the classical problem of homography decomposition, existing approaches for the decomposition of homography and details the system dynamics and measurements used for the observer design. In Sections 5.3 and 5.4, we propose two nonlinear Riccati observers for the decomposition of the homography and its inverse, respectively. Simulation results are provided in Section 5.5. Experimental results have been presented to show the performance and robustness of the proposed observers in Section 5.6.

## 5.2 Problem Statement

### 5.2.1 Homography decomposition problem

The so-called *Euclidean homography* that maps Euclidean coordinates of the scene’s points from  $\{\mathcal{A}\}$  to  $\{\mathring{\mathcal{A}}\}$  is given by (see Fig. 5.1)

$$H = R + \frac{1}{d}\xi\eta^\top \quad (5.1)$$

Assume that the camera is well calibrated, then the Euclidean homography can be directly computed from the so-called *projective homography* or *image homography* estimated from the image point correspondences (see Chapter 3 for details).

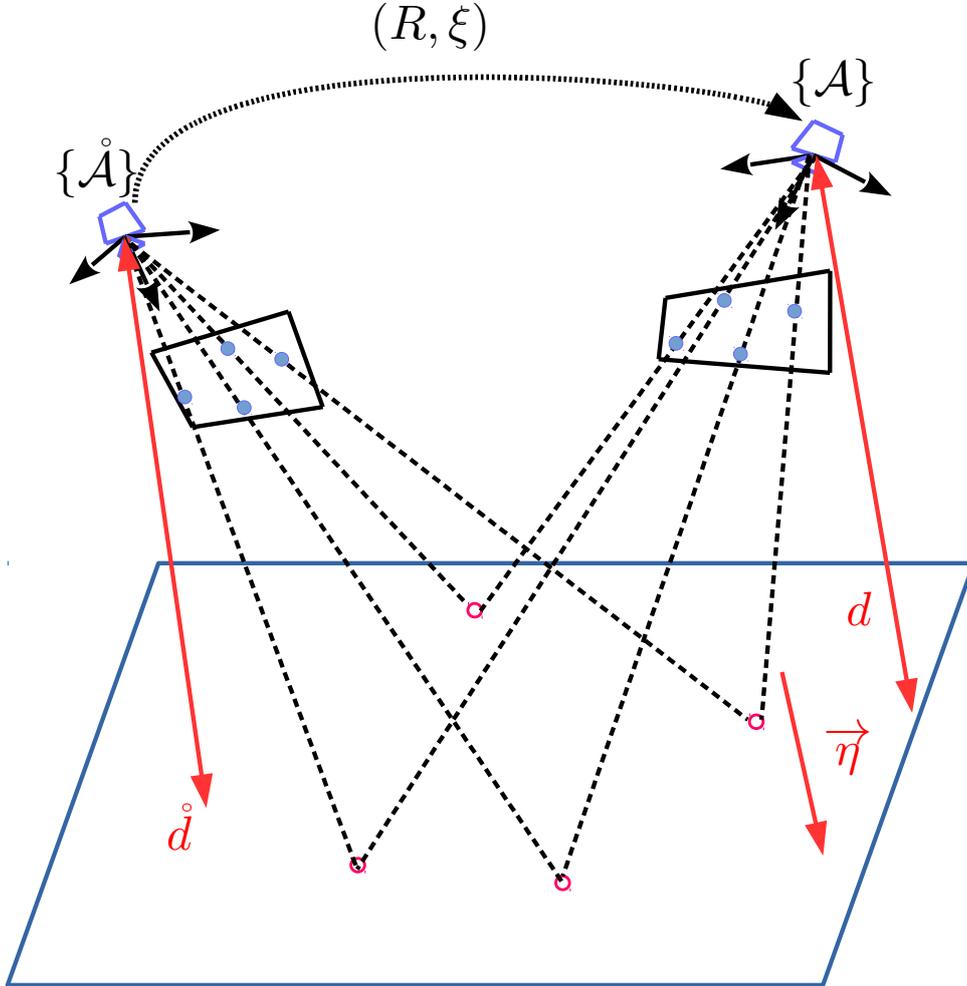


Figure 5.1: Euclidean homography relating the camera's pose, the distance to the plane and the plane's normal vector by  $H = R + \frac{1}{a}\xi\eta^\top$ .

Using the relations in (3.4), one verifies that the inverse of  $H$  is given by

$$G := H^{-1} = R^\top - \frac{1}{d}R^\top\xi\hat{\eta}^\top \quad (5.2)$$

The well-known problem of *Euclidean homography decomposition* consists in decomposing the matrix  $H$  into the separate elements  $(R, \xi/d, \eta)$ . There exist few algorithms for this homography decomposition problem such as Faugeras SVD-based [16], Zhang SVD-based [77], Malis-Vargas analytical [53] decomposition algorithms and have been

detailed in the next subsection. These algorithms, when applied to the homography inverse  $G$ , would yield the separate elements  $(R^\top, -R^\top \xi / \dot{d}, \dot{\eta})$  from which one can obtain  $(R, \xi / \dot{d}, \dot{\eta})$ .

### 5.2.2 Brief overview of the existing approaches for Homography Decomposition

Faugeras and Lustman [16] were the first to solve the problem of homography decomposition. They proved that the problem of homography decomposition gives up to eight different solutions except in some special cases (for e.g when there is a pure rotation). The proof provided by Faugeras and Lustman to compute the set of solutions is constructive and exhaustively considers all the cases, thus providing a practical method to realize the solutions whenever possible. This method uses the singular value decomposition  $H = U\Sigma V^\top$  by considering the diagonal matrix as an homography matrix. This problem can be solved analytically, and then after obtaining the values for  $U$  and  $V$  we can compute the final decomposition elements  $(R, \xi/d, \eta)$ . As already stated, out of the eight mathematical solutions obtained only two of them are physically possible. These two solutions are obtained by taking into consideration several constraints (for e.g all the viewed scene points should be in front of the camera).

A similar approach to obtain the homography decomposition was proposed by Zhang and Hanson [77]. In this method as well the solutions are obtained numerically, again from SVD decomposition of the homography matrix  $H^\top H = V\Sigma V^\top$ . This method however gives a slightly easier way to handle the special cases as compared to Faugeras and Lustman [16] and is also claimed to be computationally cheaper. From the two methods mentioned above it is clear that they don't provide an analytical expression of the decomposition elements in terms of the Homography matrix  $H$ .

A more recent method which was proposed by Malis and Vargas [53] was able to provide analytical expressions for the decomposition elements  $(R, \xi/d, \eta)$ , in terms of the components of the Homography matrix  $H$ . Again with this method proposed by Malis we obtain four solutions. The procedure for obtaining these four solutions has been detailed in [53]. After obtaining these 4 solutions, they are reduced to only 2 by verifying the positive depth constraint (i.e. third component of the normal vector should be positive). From the remaining two solutions, it is difficult to predict the correct solution. However since for our experiments, the planar target is almost perpendicular to the camera then if we have a good estimate of the normal vector, it should be approximately equal to  $e_3$ . Hence we choose the estimated normal vector closest to  $\dot{\eta} \cong e_3$  as the valid solution.

The main drawback of all classical (algebraic) homography decomposition approaches

## 5.2. Problem Statement

---

is that they only try to solve the homography decomposition problem on a frame-by-frame basis and, thus, do not provide any filtering effect on the resulting variables. The precision of the decomposition elements (particularly of the normal vector estimate) is highly prone to noise, especially when the camera's translation is small [53]. The problem even becomes degenerate when the latter vanishes (i.e.  $\xi = 0$ ), since in such a case any unit vector  $\eta \in S^2$  can be a solution. In this paper, a novel approach for solving the homography decomposition problem (i.e. decomposing either  $H$  or  $G$ ) is explored by exploiting the differential equations guiding the decomposed elements via the design of deterministic nonlinear observers.

Define

$$\zeta := \frac{R^\top \xi}{d}, \quad \dot{\zeta} := \frac{\dot{\xi}}{\dot{d}} \quad (5.3)$$

One verifies that the expressions (5.1) and (5.2) of  $H$  and  $G$ , respectively, can be rewritten as

$$H = R(I_3 + \zeta \eta^\top) \quad (5.4)$$

$$G = R^\top(I_3 - \dot{\zeta} \dot{\eta}^\top) \quad (5.5)$$

The objective of observer design then consists in providing the estimation of  $(R, \zeta, \eta)$  (resp.  $(R, \dot{\zeta}, \dot{\eta})$ ) using  $H$  (resp.  $G$ ) as measurement quantity. Depending on the considered application, it is preferable to decompose  $H$  or  $G$ .

### 5.2.3 System equations and measurements for observer design

The rigid body kinematics of  $(R, \xi)$  are given by

$$\begin{cases} \dot{R} = R\Omega_\times \\ \dot{\xi} = RV \end{cases} \quad (5.6)$$

with  $V \in \mathbb{R}^3$  and  $\Omega \in \mathbb{R}^3$  denoting the vectors of coordinates of the camera's linear and angular velocities expressed in  $\{\mathcal{A}\}$ .

Since the scene is stationary the normal vector  $\dot{\eta} \in S^2$  expressed in the reference frame  $\{\dot{\mathcal{A}}\}$  is constant. Since  $\eta = R^\top \dot{\eta}$  one thus deduces

$$\dot{\eta} = -\Omega_\times \eta \quad (5.7)$$

Using (3.4), (5.3) and (5.6), one verifies that the dynamics of  $\zeta$  and  $\dot{\zeta}$  are given by

$$\dot{\zeta} = -\Omega_\times \zeta + \frac{V}{d} - \frac{\dot{d}}{d} \zeta = (-\Omega_\times + \phi_\perp I_3) \zeta + \phi \quad (5.8)$$

$$\dot{\zeta} = \frac{RV}{\dot{d}} = \frac{d}{\dot{d}}R\phi = (1 - \hat{\eta}^\top \hat{\zeta})R\phi \quad (5.9)$$

where  $\phi := \frac{V}{d}$  and  $\phi_\perp := -\frac{\dot{d}}{d} = \frac{V^\top \eta}{d}$  are the so-called translational optical flow and optical flow divergence, respectively. Both the translational optical flow and the optical flow divergence are assumed to be measured (see, e.g., section 3.6 in Chapter 3). Assume also that the angular velocity  $\Omega$  is measured using an embedded 3-axis gyrometer.

**Remark 3.** Note that observer design for the case where the linear velocity measurement is available (using a Doppler velocity sensor or a GPS) can be addressed differently, but it is out of scope of the present paper. Using optical flow and optical flow divergence as measurement in turn makes our solution more appealing in practice because of its simplicity of sensor requirement, bearing in mind that a Doppler velocity sensor is very expensive while a GPS does not work in indoor or GPS-denied environments. On the contrary, the quality of optical flow and optical flow divergence measurements depends much on the texture of the scene along with its planarity approximation –the assumption used to extract these quantities from images.

## 5.3 Observer design for decomposing the homography matrix

### 5.3.1 Observer derivation

Inspired by [27], we avoid using minimal parametrization techniques such as *spherical coordinates* to parametrize the normal unit vector  $\eta$  – an element of  $S^2$ . Instead, an auxiliary rotation matrix  $Q \in \text{SO}(3)$  is introduced such that

$$\eta = Q^\top e_3$$

The underlying idea is to over-parameterize an element of  $S^2$  (dimension 2) by an element of  $\text{SO}(3)$  (dimension 3). The advantage of such type of parameterization is that it reduces the complexities of the error system in first-order approximations that arise due to minimal parameterization techniques for elements on  $S^2$  (see [27] for more thorough discussions).

In view of the dynamics (5.7) of  $\eta$ , one deduces a possibility of the dynamics of  $Q$  as

$$\dot{Q} = Q\Omega_\times \quad (5.10)$$

Let  $\hat{Q} \in \text{SO}(3)$ ,  $\hat{R} \in \text{SO}(3)$ ,  $\hat{\zeta} \in \mathbb{R}^3$  denote the estimates of  $Q$ ,  $R$ ,  $\zeta$ , respectively. The estimated normal vector is then given by

$$\hat{\eta} := \hat{Q}^\top e_3$$

### 5.3. Observer design for decomposing the homography matrix

---

In view of the first equation of (5.6), (5.8) and (5.10), the following general form of observer is proposed

$$\begin{cases} \dot{\hat{Q}} = \hat{Q}\Omega_\times - \sigma_Q \times \hat{Q} \\ \dot{\hat{R}} = \hat{R}\Omega_\times - \hat{R}\sigma_{R\times} \\ \dot{\hat{\zeta}} = (-\Omega_\times + \phi_\perp I_3)\hat{\zeta} + \phi - \sigma_\zeta \end{cases} \quad (5.11)$$

with initial conditions  $\hat{Q}(0), \hat{R}(0) \in \text{SO}(3)$ ,  $\hat{\zeta}(0) \in \mathbb{R}^3$  and with innovation terms  $\sigma_Q, \sigma_R, \sigma_\zeta \in \mathbb{R}^3$  to be designed thereafter.

The following error variables are defined:

$$\tilde{Q} := Q\hat{Q}^\top, \tilde{R} := \hat{R}^\top R, \tilde{\zeta} := \zeta - \hat{\zeta} \quad (5.12)$$

Then the objective of observer design consists in stabilizing  $(\tilde{Q}e_3, \tilde{R}, \tilde{\zeta})$  about  $(e_3, I_3, 0)$ . From (5.6), (5.11) and (5.12), one verifies that the error system is

$$\begin{cases} \dot{\tilde{Q}} = \tilde{Q}\sigma_{Q\times} \\ \dot{\tilde{R}} = \tilde{R}\Omega_\times - \Omega_\times \tilde{R} + \sigma_{R\times} \tilde{R} \\ \dot{\tilde{\zeta}} = (-\Omega_\times + \phi_\perp I_3)\tilde{\zeta} + \sigma_\zeta \end{cases} \quad (5.13)$$

For analysis purposes let us assume that  $\zeta, \Omega$  and  $\phi$  remain uniformly bounded for all time, which is a completely reasonable assumption in practice.

The following step involves developing first-order approximations of the error system (5.13) and of the measurement equation (5.4). From the Rodrigues' formula, one deduces the following first-order approximations of  $\tilde{Q}$  and  $\tilde{R}$

$$\begin{aligned} \tilde{Q} &= I_3 + \lambda_{\tilde{Q}\times} + O(|\lambda_{\tilde{Q}}|^2) \\ \tilde{R} &= I_3 + \lambda_{\tilde{R}\times} + O(|\lambda_{\tilde{R}}|^2) \end{aligned} \quad (5.14)$$

with  $\lambda_{\tilde{Q}}, \lambda_{\tilde{R}} \in \mathcal{B}_2^3$  equal to twice the vector part of the quaternion associated with the attitude error matrix  $\tilde{Q}$  and  $\tilde{R}$ , respectively.

One then deduces from the first two equations of (5.13) and (5.14) that in first-order approximations

$$\dot{\lambda}_{\tilde{Q}} = \sigma_Q + O(|\lambda_{\tilde{Q}}||\sigma_Q|) \quad (5.15)$$

and

$$\begin{aligned} \dot{\lambda}_{\tilde{R}\times} &= \lambda_{\tilde{R}\times}\Omega_\times - \Omega_\times \lambda_{\tilde{R}\times} + \sigma_{R\times} + O(|\lambda_{\tilde{R}}|^2) + O(|\lambda_{\tilde{R}}||\sigma_R|) \\ &= (\lambda_{\tilde{R}} \times \Omega)_\times + \sigma_{R\times} + O(|\lambda_{\tilde{R}}|^2) + O(|\lambda_{\tilde{R}}||\sigma_R|) \end{aligned}$$

which yields

$$\dot{\lambda}_{\tilde{R}} = -\Omega_{\times} \lambda_{\tilde{R}} + \sigma_R + O(|\lambda_{\tilde{R}}|^2) + O(|\lambda_{\tilde{R}}| |\sigma_R|) \quad (5.16)$$

As for the measurement equation (5.4), this homography expression can be developed in first-order approximations as follows

$$\begin{aligned} R^{\top} H - I_3 &= \zeta \eta^{\top} = \zeta (Q^{\top} e_3)^{\top} \\ \Rightarrow \tilde{R}^{\top} \hat{R}^{\top} H - I_3 &= (\tilde{\zeta} + \hat{\zeta}) (\hat{Q}^{\top} \tilde{Q}^{\top} e_3)^{\top} \\ \Rightarrow (I_3 - \lambda_{\tilde{R} \times}) \hat{R}^{\top} H - I_3 & \\ &= (\tilde{\zeta} + \hat{\zeta}) (\hat{Q}^{\top} (I_3 - \lambda_{\tilde{Q} \times}) e_3)^{\top} + O(|\lambda_{\tilde{Q}}|^2) + O(|\lambda_{\tilde{R}}|^2) \\ \Rightarrow (\hat{R}^{\top} H - I_3) - \lambda_{\tilde{R} \times} (\hat{R}^{\top} H) & \\ &= \hat{\zeta} (\hat{Q}^{\top} e_3)^{\top} + \tilde{\zeta} (\hat{Q}^{\top} e_3)^{\top} + \hat{\zeta} (\hat{Q}^{\top} e_3 \times \lambda_{\tilde{Q}})^{\top} \\ &\quad + O(|\lambda_{\tilde{Q}}|^2) + O(|\lambda_{\tilde{R}}|^2) + O(|\tilde{\zeta}| |\lambda_{\tilde{Q}}|) \\ \Rightarrow (\hat{R}^{\top} H - I_3) - \hat{\zeta} (\hat{Q}^{\top} e_3)^{\top} & \\ &= \lambda_{\tilde{R} \times} (\hat{R}^{\top} H) + \tilde{\zeta} (\hat{Q}^{\top} e_3)^{\top} - \lambda_{\tilde{Q}, 2} \hat{\zeta} (\hat{Q}^{\top} e_1)^{\top} + \lambda_{\tilde{Q}, 1} \hat{\zeta} (\hat{Q}^{\top} e_2)^{\top} \\ &\quad + O(|\lambda_{\tilde{Q}}|^2) + O(|\lambda_{\tilde{R}}|^2) + O(|\tilde{\zeta}| |\lambda_{\tilde{Q}}|) \end{aligned}$$

Note that the last equality only involves the first two components of  $\lambda_{\tilde{Q}}$  (i.e.  $\lambda_{\tilde{Q}, 1}$  and  $\lambda_{\tilde{Q}, 2}$ ) and can be equivalently written as

$$\begin{aligned} y &:= \begin{bmatrix} (\hat{R}^{\top} H - I_3) (\hat{Q}^{\top} e_3) - \hat{\zeta} \\ (\hat{R}^{\top} H - I_3) (\hat{Q}^{\top} e_2) \\ (\hat{R}^{\top} H - I_3) (\hat{Q}^{\top} e_1) \end{bmatrix} \\ &= \begin{bmatrix} -((\hat{R}^{\top} H) (\hat{Q}^{\top} e_3))_{\times} \lambda_{\tilde{R}} + \tilde{\zeta} \\ -((\hat{R}^{\top} H) (\hat{Q}^{\top} e_2))_{\times} \lambda_{\tilde{R}} + \lambda_{\tilde{Q}, 1} \hat{\zeta} \\ -((\hat{R}^{\top} H) (\hat{Q}^{\top} e_1))_{\times} \lambda_{\tilde{R}} - \lambda_{\tilde{Q}, 2} \hat{\zeta} \end{bmatrix} \\ &\quad + O(|\lambda_{\tilde{Q}}|^2) + O(|\lambda_{\tilde{R}}|^2) + O(|\tilde{\zeta}| |\lambda_{\tilde{Q}}|) \end{aligned} \quad (5.17)$$

From (5.15), (5.16), the third equation of (5.13), and (5.17), one obtains in first-order

### 5.3. Observer design for decomposing the homography matrix

---

approximations the system in compact form (2.20) with output  $y$  defined in (5.17) and

$$\left\{ \begin{array}{l} x := \begin{bmatrix} \lambda_{\tilde{Q},1} \\ \lambda_{\tilde{Q},2} \\ \lambda_{\tilde{R}} \\ \tilde{\zeta} \end{bmatrix}, \quad U := \begin{bmatrix} \sigma_{Q,1} \\ \sigma_{Q,2} \\ \sigma_R \\ \sigma_{\zeta} \end{bmatrix} \\ A(t) := \begin{bmatrix} 0 & 0 & 0_{1 \times 3} & 0_{1 \times 3} \\ 0 & 0 & 0_{1 \times 3} & 0_{1 \times 3} \\ 0_{3 \times 1} & 0_{3 \times 1} & -\Omega_{\times} & 0_3 \\ 0_{3 \times 1} & 0_{3 \times 1} & 0_3 & -\Omega_{\times} + \phi_{\perp} I_3 \end{bmatrix} \\ C(x, t) := \begin{bmatrix} 0_{3 \times 1} & 0_{3 \times 1} & -((\hat{R}^{\top} H)(\hat{Q}^{\top} e_3))_{\times} & I_3 \\ \hat{\zeta} & 0_{3 \times 1} & -((\hat{R}^{\top} H)(\hat{Q}^{\top} e_2))_{\times} & 0_3 \\ 0_{3 \times 1} & -\hat{\zeta} & -((\hat{R}^{\top} H)(\hat{Q}^{\top} e_1))_{\times} & 0_3 \end{bmatrix} \end{array} \right. \quad (5.18)$$

The fact that  $x = [x_1^{\top}, x_2^{\top}]^{\top}$  with  $x_1 := [\lambda_{\tilde{Q},1}, \lambda_{\tilde{Q},2}, \lambda_{\tilde{R}}^{\top}]^{\top} \in \mathcal{B}_{2\sqrt{2}}^5$  (the closed ball in  $\mathbb{R}^5$  of radius equal to  $2\sqrt{2}$ ),  $x_2 := \tilde{\zeta} \in \mathbb{R}^3$ , together with the particular form of the matrix  $A$  as (2.21), allows one to obtain the expression of the innovation terms from the input  $U$  calculated according to (2.22) and (2.23) where the matrices  $D$  and  $S$  (involved in (2.23)) are chosen larger than some constant positive matrix.

**Remark 4.** *Interestingly and as also discussed in [27], although we have used a 3-dimensional variable  $Q \in \text{SO}(3)$  to over-parametrize a 2-dimensional variable  $\eta \in S^2$ , the resulting first-order approximated system only involves a minimum number of components of  $\lambda_{\tilde{Q}}$  so that the dimension of  $X$  is equal to the dimension of the state  $(\eta, R, \bar{\xi})$ . Moreover, only the first two components of the innovation term  $\sigma_Q$  (i.e.  $\sigma_{Q,1}$  and  $\sigma_{Q,2}$ ) are involved in the Riccati observer design process. Its last component (i.e.  $\sigma_{Q,3}$ ) can thus be set equal to zero for the sake of simplicity. Of course, the above reasoning together with the associated local stability and convergence properties is only valid in the first order approximation.*

#### 5.3.2 Observability and stability analysis

According to [24] the equilibrium  $x = 0$  is locally exponentially stable, provided that the pair  $(A^*(t), C^*(t))$  with  $A^*(t) = A(t)$  and  $C^*(t) := C(0, t)$  is uniformly observable.

By setting  $x = 0$  in the expression of  $C(x, t)$  in (5.18) one obtains

$$\begin{aligned} C^* &= \begin{bmatrix} 0_{3 \times 1} & 0_{3 \times 1} & -((I_3 + \zeta \eta^\top)(Q^{*\top} e_3))_\times & I_3 \\ \zeta & 0_{3 \times 1} & -((I_3 + \zeta \eta^\top)(Q^{*\top} e_2))_\times & 0_3 \\ 0_{3 \times 1} & -\zeta & -((I_3 + \zeta \eta^\top)(Q^{*\top} e_1))_\times & 0_3 \end{bmatrix} \\ &= \begin{bmatrix} 0_{3 \times 1} & 0_{3 \times 1} & -(\eta + \zeta)_\times & I_3 \\ \zeta & 0_{3 \times 1} & -(Q^{*\top} e_2)_\times & 0_3 \\ 0_{3 \times 1} & -\zeta & -(Q^{*\top} e_1)_\times & 0_3 \end{bmatrix} \end{aligned} \quad (5.19)$$

with  $Q^* \in \text{SO}(3)$  satisfying  $Q^{*\top} e_3 = \eta$  and  $\dot{Q}^* = Q^* \Omega_\times$ .

For later use, let  $q_i^* \in S^2$ ,  $i = 1, 2, 3$ , denote the  $i$ -th row of  $Q^*$  (i.e.  $q_i^* = Q^{*\top} e_i$ ).

**Proposition 6.** *Assume that there exists a positive number  $\nu$  such that  $\forall t > 0$*

$$\frac{1}{\delta} \int_t^{t+\delta} |\zeta(\tau) \times \eta(\tau)| d\tau \geq \nu \quad (5.20)$$

*Then, the pair  $(A^*, C^*)$  is uniformly observable. Consequently, the equilibrium  $(\tilde{Q}e_3, \tilde{R}, \tilde{\zeta}) = (e_3, I_3, 0)$  of the error system is locally exponentially stable.*

**Proof.** According to Lemma 3 in chapter 2 by choosing  $M = C^*$ , then the pair  $(A^*, C^*)$  is uniformly observable if  $\exists \delta, \mu > 0$  such that

$$\frac{1}{\delta} \int_t^{t+\delta} \det(C^{*\top}(\tau)C^*(\tau)) d\tau \geq \mu, \quad \forall t > 0 \quad (5.21)$$

We show thereafter that condition (5.20) is sufficient to guarantee (5.21). One verifies that

$$C^{*\top}C^* = |\zeta|^2 \begin{bmatrix} I_2 & \frac{1}{|\zeta|^2} B \\ \frac{1}{|\zeta|^2} B^\top & \frac{1}{|\zeta|^2} G \end{bmatrix}$$

with  $B \in \mathbb{R}^{2 \times 6}$  and  $G \in \mathbb{R}^{6 \times 6}$  defined by

$$\begin{aligned} B &:= \begin{bmatrix} -\zeta^\top q_{2 \times}^* & 0_{1 \times 3} \\ \zeta^\top q_{1 \times}^* & 0_{1 \times 3} \end{bmatrix} \\ G &:= \begin{bmatrix} -(\eta + \zeta)_\times^2 - q_{1 \times}^{*2} - q_{2 \times}^{*2} & (\eta + \zeta)_\times \\ -(\eta + \zeta)_\times & I_3 \end{bmatrix} \end{aligned}$$

Thus, one deduces

### 5.3. Observer design for decomposing the homography matrix

$$\begin{aligned}
\det(C^{\star\top}C^{\star}) &= |\zeta|^4 \det\left(G - \frac{1}{|\zeta|^2} B^{\top}B\right) \\
&= |\zeta|^4 \det\begin{bmatrix} -(\eta+\zeta)_{\times}^2 - q_{1\times}^{\star} \pi_{\frac{\zeta}{|\zeta|}} q_{1\times}^{\star} - q_{2\times}^{\star} \pi_{\frac{\zeta}{|\zeta|}} q_{2\times}^{\star} & (\eta+\zeta)_{\times} \\ & -(\eta+\zeta)_{\times} & I_3 \end{bmatrix} \\
&= |\zeta|^4 \det\left(-q_{1\times}^{\star} \pi_{\frac{\zeta}{|\zeta|}} q_{1\times}^{\star} - q_{2\times}^{\star} \pi_{\frac{\zeta}{|\zeta|}} q_{2\times}^{\star}\right) \\
&= |\zeta|^4 \det\left(Q^{\star\top} \begin{pmatrix} -e_{1\times} \pi_{\frac{Q^{\star}\zeta}{|\zeta|}} e_{1\times} & -e_{2\times} \pi_{\frac{Q^{\star}\zeta}{|\zeta|}} e_{2\times} \end{pmatrix} Q^{\star}\right) \\
&= |\zeta|^4 \det\left(-e_{1\times} \pi_{\eta_{\zeta}^{\star}} e_{1\times} - e_{2\times} \pi_{\eta_{\zeta}^{\star}} e_{2\times}\right) \\
&= |\zeta|^4 \det\left(e_3 e_3^{\top} + \eta_{\zeta}^{\star} \eta_{\zeta}^{\star\top} + (\eta_{\zeta}^{\star} \times e_3)(\eta_{\zeta}^{\star} \times e_3)^{\top}\right) \\
&= |\zeta|^4 |\eta_{\zeta}^{\star} \times e_3|^4 \\
&= |\zeta \times \eta|^4
\end{aligned}$$

with  $\eta_{\zeta}^{\star} := \frac{Q^{\star}\zeta}{|\zeta|}$ . From the Cauchy-Schwarz integral inequality, one deduces (using (5.20))

$$\begin{aligned}
\int_t^{t+\delta} |\zeta(\tau) \times \eta(\tau)|^4 d\tau &\geq \frac{1}{\delta} \left( \int_t^{t+\delta} |\zeta(\tau) \times \eta(\tau)|^2 d\tau \right)^2 \\
&\geq \frac{1}{\delta^3} \left( \int_t^{t+\delta} |\zeta(\tau) \times \eta(\tau)| d\tau \right)^4 \geq \delta \nu^4
\end{aligned}$$

One finally deduces (5.21) with  $\mu = \nu^4$ . The remainder of the proof then directly follows by application of Theorem 3.1 and Corollary 3.2 in [24].  $\square$

**Remark 5.** The sufficient uniform observability condition (5.20) excludes the degenerate case where the camera's translation w.r.t. the reference frame is identically equal to zero (i.e.  $\xi(t) \equiv 0, \forall t \geq 0$ ). It also excludes the restrictive case where the camera translation motion is always parallel to the normal vector to the plane, i.e.  $\xi(t) \times \dot{\eta}(t) \equiv 0, \forall t \geq 0$ , which can however still be uniformly observable. The non-trivial observability analysis for such situations (requiring another formulation of  $M$  that involves  $C^{\star}A^{\star} + \dot{C}^{\star}$ ) is left to the interested reader.

**Remark 6.** Since  $\zeta = R^{\top} \mathring{\zeta}$  and  $\eta = R^{\top} \mathring{\eta}$ , condition (5.20) can be equivalently rewritten as

$$\frac{1}{\delta} \int_t^{t+\delta} |\mathring{\zeta}(\tau) \times \mathring{\eta}(\tau)| d\tau \geq \nu$$

## 5.4 Observer design for decomposing the inverse homography matrix

### 5.4.1 Observer derivation

Similarly to the previous section, we introduce an auxiliary matrix  $\mathring{Q} \in \text{SO}(3)$  to over-parametrize  $\mathring{\eta} \in S^2$  such that  $\mathring{\eta} = \mathring{Q}^\top e_3$ . Since  $\frac{d}{dt}\mathring{\eta} = 0$ , one deduces

$$\mathring{Q} = 0_{3 \times 3} \quad (5.22)$$

Let  $\hat{Q} \in \text{SO}(3)$ ,  $\hat{R} \in \text{SO}(3)$ ,  $\hat{\zeta} \in \mathbb{R}^3$  denote the estimates of  $\mathring{Q}$ ,  $R$ ,  $\mathring{\zeta}$ , respectively. The estimated normal vector is obtained as  $\hat{\eta} = \hat{Q}^\top e_3$ . In view of the first equation of (5.6), (5.9) and (5.22), the following observer form is considered

$$\begin{cases} \dot{\hat{Q}} = -\sigma_{\hat{Q} \times} \hat{Q} \\ \dot{\hat{R}} = \hat{R} \Omega_{\times} - \sigma_{R \times} \hat{R} \\ \dot{\hat{\zeta}} = (1 - e_3^\top \hat{Q} \hat{\zeta}) \hat{R} \phi - \sigma_{\zeta} \end{cases} \quad (5.23)$$

with  $\hat{Q}(0), \hat{R}(0) \in \text{SO}(3), \hat{\zeta}(0) \in \mathbb{R}^3$  and with innovation terms  $\sigma_{\hat{Q}}, \sigma_R, \sigma_{\zeta} \in \mathbb{R}^3$  to be designed thereafter.

Define the error variables as

$$\tilde{Q} := \mathring{Q} \hat{Q}^\top, \quad \bar{R} := R \hat{R}^\top, \quad \tilde{\zeta} := \mathring{\zeta} - \hat{\zeta} \quad (5.24)$$

Then the objective of observer design consists in stabilizing  $(\tilde{Q} e_3, \bar{R}, \tilde{\zeta})$  about  $(e_3, I_3, 0)$ . One verifies from (5.6), (5.23) and (5.24) that the error dynamics are given by

$$\begin{cases} \dot{\tilde{Q}} = \tilde{Q} \sigma_{\tilde{Q} \times} \\ \dot{\bar{R}} = \bar{R} \sigma_{R \times} \\ \dot{\tilde{\zeta}} = (1 - e_3^\top \tilde{Q} \tilde{\zeta}) \bar{R} \phi - (1 - e_3^\top \tilde{Q} \tilde{\zeta}) \hat{R} \phi + \sigma_{\zeta} \end{cases} \quad (5.25)$$

We now develop first-order approximations of the error system (5.25) and of the measurement equation (5.5). The first-order approximations of  $\tilde{Q}$  and  $\bar{R}$  are given by

$$\begin{aligned} \tilde{Q} &= I_3 + \lambda_{\tilde{Q} \times} + O(|\lambda_{\tilde{Q}}|^2) \\ \bar{R} &= I_3 + \lambda_{\bar{R} \times} + O(|\lambda_{\bar{R}}|^2) \end{aligned} \quad (5.26)$$

#### 5.4. Observer design for decomposing the inverse homography matrix

---

with  $\lambda_{\tilde{Q}}, \lambda_{\tilde{R}} \in \mathcal{B}_2^3$ . One then deduces from (5.25) and (5.26) that in first-order approximations

$$\dot{\lambda}_{\tilde{Q}} = \sigma_{\hat{Q}} + O(|\lambda_{\tilde{Q}}| |\sigma_{\hat{Q}}|) \quad (5.27)$$

$$\dot{\lambda}_{\tilde{R}} = \sigma_R + O(|\lambda_{\tilde{R}}| |\sigma_R|) \quad (5.28)$$

First-order approximations of the third equation of (5.25) yield

$$\begin{aligned} \dot{\zeta} &= \left( e_3^\top \hat{Q} \hat{\zeta} \right) \hat{R} \phi - \left( e_3^\top \tilde{Q} \hat{Q} (\hat{\zeta} + \tilde{\zeta}) \right) \bar{R} \hat{R} \phi + (I_3 - \bar{R}) \hat{R} \phi + \sigma_{\hat{\zeta}} \\ &= \left( e_3^\top \hat{Q} \hat{\zeta} \right) \hat{R} \phi - \left( e_3^\top (I_3 + \lambda_{\tilde{Q} \times}) \hat{Q} (\hat{\zeta} + \tilde{\zeta}) \right) (I_3 + \lambda_{\tilde{R} \times}) \hat{R} \phi \\ &\quad - \lambda_{\tilde{R} \times} \hat{R} \phi + \sigma_{\hat{\zeta}} + O(|\lambda_{\tilde{Q}}|^2) + O(|\lambda_{\tilde{R}}|^2) \\ &= - \left( \hat{R} \phi (e_2^\top \hat{Q} \hat{\zeta}) \right) \lambda_{\tilde{Q},1} + \left( \hat{R} \phi (e_1^\top \hat{Q} \hat{\zeta}) \right) \lambda_{\tilde{Q},2} \\ &\quad + \left( 1 + e_3^\top \hat{Q} \hat{\zeta} \right) (\hat{R} \phi) \times \lambda_{\tilde{R}} - \left( \hat{R} \phi e_3^\top \hat{Q} \right) \tilde{\zeta} + \sigma_{\hat{\zeta}} \\ &\quad + O(|\lambda_{\tilde{Q}}|^2) + O(|\lambda_{\tilde{R}}|^2) + O(|\lambda_{\tilde{Q}}| |\lambda_{\tilde{R}}|) \\ &\quad + O(|\lambda_{\tilde{Q}}| |\tilde{\zeta}|) + O(|\lambda_{\tilde{R}}| |\tilde{\zeta}|) \end{aligned} \quad (5.29)$$

As for the measurement equation (5.5), one deduces

$$\begin{aligned} I_3 - RG &= \hat{\zeta} \hat{\eta}^\top = \hat{\zeta} (\hat{Q}^\top e_3)^\top \\ \Rightarrow I_3 - \bar{R} \hat{R} G &= (\tilde{\zeta} + \hat{\zeta}) (\hat{Q}^\top \tilde{Q}^\top e_3)^\top \\ \Rightarrow I_3 - (I_3 + \lambda_{\tilde{R} \times}) \hat{R} G &= (\tilde{\zeta} + \hat{\zeta}) (\hat{Q}^\top (I_3 - \lambda_{\tilde{Q} \times}) e_3)^\top + O(|\lambda_{\tilde{Q}}|^2) + O(|\lambda_{\tilde{R}}|^2) \\ \Rightarrow (I_3 - \hat{R} G) - \lambda_{\tilde{R} \times} \hat{R} G &= \hat{\zeta} (\hat{Q}^\top e_3)^\top + \tilde{\zeta} (\hat{Q}^\top e_3)^\top + \hat{\zeta} (\hat{Q}^\top e_3 \times \lambda_{\tilde{Q}})^\top \\ &\quad + O(|\lambda_{\tilde{Q}}|^2) + O(|\lambda_{\tilde{R}}|^2) + O(|\tilde{\zeta}| |\lambda_{\tilde{Q}}|) \\ \Rightarrow (I_3 - \hat{R} G) - \hat{\zeta} (\hat{Q}^\top e_3)^\top &= \lambda_{\tilde{R} \times} \hat{R} G + \tilde{\zeta} (\hat{Q}^\top e_3)^\top + \lambda_{\tilde{Q},1} \hat{\zeta} (\hat{Q}^\top e_2)^\top - \lambda_{\tilde{Q},2} \hat{\zeta} (\hat{Q}^\top e_1)^\top \\ &\quad + O(|\lambda_{\tilde{Q}}|^2) + O(|\lambda_{\tilde{R}}|^2) + O(|\tilde{\zeta}| |\lambda_{\tilde{Q}}|) \end{aligned}$$

which can be rewritten as

$$\begin{aligned}
 y &:= \begin{bmatrix} (I_3 - \hat{R}G)(\hat{Q}^\top e_3) - \hat{\zeta} \\ (I_3 - \hat{R}G)(\hat{Q}^\top e_2) \\ (I_3 - \hat{R}G)(\hat{Q}^\top e_1) \end{bmatrix} \\
 &= \begin{bmatrix} -((\hat{R}G)(\hat{Q}^\top e_3))_\times \lambda_{\hat{R}} + \hat{\zeta} \\ -((\hat{R}G)(\hat{Q}^\top e_2))_\times \lambda_{\hat{R}} + \lambda_{\hat{Q},1} \hat{\zeta} \\ -((\hat{R}G)(\hat{Q}^\top e_1))_\times \lambda_{\hat{R}} - \lambda_{\hat{Q},2} \hat{\zeta} \end{bmatrix} \\
 &\quad + O(|\lambda_{\hat{Q}}|^2) + O(|\lambda_{\hat{R}}|^2) + O(|\hat{\zeta}| |\lambda_{\hat{Q}}|)
 \end{aligned} \tag{5.30}$$

From (5.27), (5.28), (5.29), and (5.30), one obtains in first-order approximations the system in compact form (2.20) with output  $y$  defined in (5.30) and

$$\left\{ \begin{array}{l}
 x := \begin{bmatrix} \lambda_{\hat{Q},1} \\ \lambda_{\hat{Q},2} \\ \lambda_{\hat{R}} \\ \hat{\zeta} \end{bmatrix}, \quad U := \begin{bmatrix} \sigma_{\hat{Q},1} \\ \sigma_{\hat{Q},2} \\ \sigma_{\hat{R}} \\ \sigma_{\hat{\zeta}} \end{bmatrix} \\
 A(x, t) := \begin{bmatrix} 0 & 0 & 0_{1 \times 3} & 0_{1 \times 3} \\ 0 & 0 & 0_{1 \times 3} & 0_{1 \times 3} \\ 0_{3 \times 1} & 0_{3 \times 1} & 0_3 & 0_3 \\ -\hat{R}\phi(e_2^\top \hat{Q} \hat{\zeta}) & \hat{R}\phi(e_1^\top \hat{Q} \hat{\zeta}) & (1 + e_3^\top \hat{Q} \hat{\zeta})(\hat{R}\phi)_\times & -\hat{R}\phi e_3^\top \hat{Q} \end{bmatrix} \\
 C(x, t) := \begin{bmatrix} 0_{3 \times 1} & 0_{3 \times 1} & -((\hat{R}G)(\hat{Q}^\top e_3))_\times & I_3 \\ \hat{\zeta} & 0_{3 \times 1} & -((\hat{R}G)(\hat{Q}^\top e_2))_\times & 0_3 \\ 0_{3 \times 1} & -\hat{\zeta} & -((\hat{R}G)(\hat{Q}^\top e_1))_\times & 0_3 \end{bmatrix}
 \end{array} \right. \tag{5.31}$$

From there, one deduces the expression of the innovation terms from the input  $U$  calculated according to (2.22) and (2.23).

### 5.4.2 Observability and stability analysis

Similarly to (5.19), by setting  $x = 0$  in the expression of  $C(x, t)$  in (5.31) one obtains

$$\begin{aligned}
 C^* &= \begin{bmatrix} 0_{3 \times 1} & 0_{3 \times 1} & -((I_3 - \dot{\zeta} \dot{\eta}^\top)(\dot{Q}^{*\top} e_3))_\times & I_3 \\ \dot{\zeta} & 0_{3 \times 1} & -((I_3 - \dot{\zeta} \dot{\eta}^\top)(\dot{Q}^{*\top} e_2))_\times & 0_3 \\ 0_{3 \times 1} & -\dot{\zeta} & -((I_3 - \dot{\zeta} \dot{\eta}^\top)(\dot{Q}^{*\top} e_1))_\times & 0_3 \end{bmatrix} \\
 &= \begin{bmatrix} 0_{3 \times 1} & 0_{3 \times 1} & -(\dot{\eta} - \dot{\zeta})_\times & I_3 \\ \dot{\zeta} & 0_{3 \times 1} & -(\dot{Q}^{*\top} e_2)_\times & 0_3 \\ 0_{3 \times 1} & -\dot{\zeta} & -(\dot{Q}^{*\top} e_1)_\times & 0_3 \end{bmatrix}
 \end{aligned} \tag{5.32}$$

with  $\dot{Q}^* \in \text{SO}(3)$  constant and satisfying  $\dot{Q}^{*\top} e_3 = \dot{\eta}$ .

One observes that the expression of  $C^*$  in (5.32) is very similar to the one in (5.19). Therefore, the following proposition can be directly stated where its proof proceeds identically to the one of Proposition 6 and is thus left to the interested reader.

**Proposition 7.** *Assume that there exists a positive number  $\nu$  such that  $\forall t > 0$  inequality (5.20) holds. Then, the pair  $(A^*, C^*)$  is uniformly observable. Consequently, the equilibrium  $(\tilde{Q} e_3, \tilde{R}, \tilde{\zeta}) = (e_3, I_3, 0)$  of the error system is locally exponentially stable.*

## 5.5 Simulation results

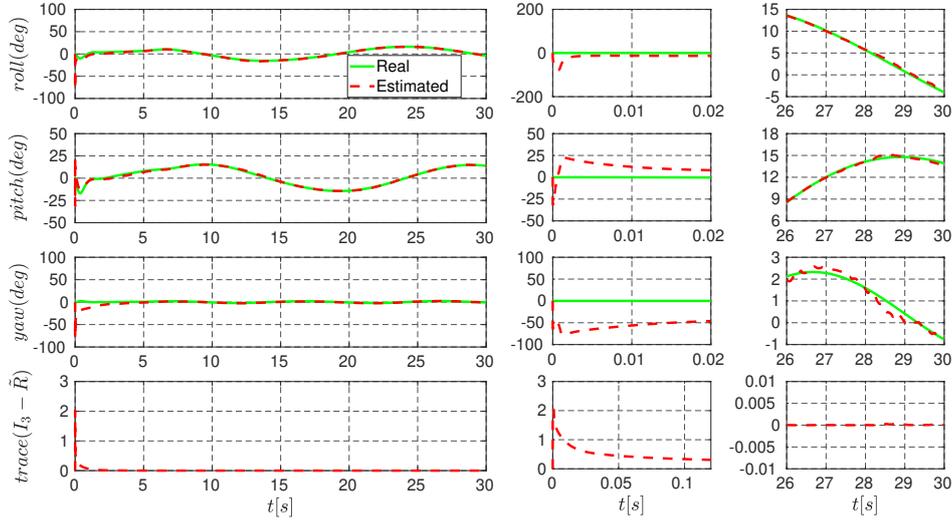


Figure 5.2: **Scenario 1 (Method 1)** – Estimated and real attitudes represented by roll, pitch and yaw Euler angles ( $deg$ ) and the attitude estimate error versus time ( $s$ ) (LEFT) and their zooms (RIGHT).

In this section the robustness and performance of the proposed observers are demonstrated through simulation by considering two scenarios. For the sake of simplicity, the

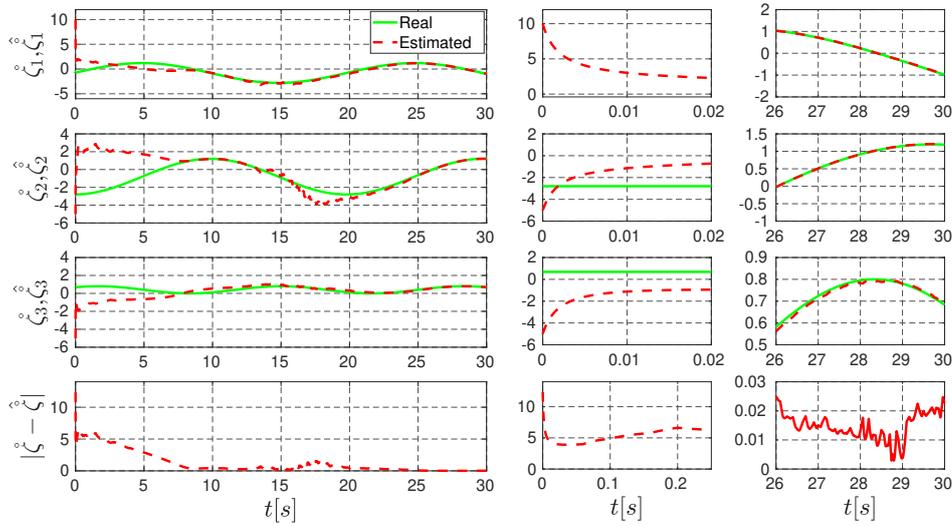


Figure 5.3: **Scenario 1 (Method 1)** – Estimated and real scaled position and norm of the scaled position error versus time (s) (LEFT) and their zooms (RIGHT).

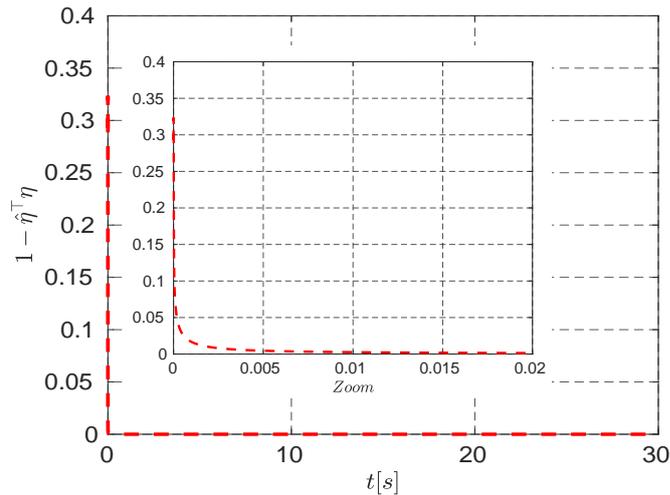


Figure 5.4: **Scenario 1 (Method 1)** – Estimation error of the normal vector estimate represented by  $1 - \hat{\eta}^\top \eta$ .

observers proposed in Sections 5.3 and 5.4 are referred to as **Method 1** and **Method 2**, respectively. For the simulations, we consider that a monocular camera is attached to an aerial drone performing some specific trajectories and observing a planar target. In the first scenario, the drone realizes an aggressive periodical trajectory and noise is introduced to the measurements used in the proposed observer, namely  $H$ ,  $\phi$ ,  $\phi_\perp$  and  $\Omega$ . While in the second scenario, we consider the case where the camera trajectory passes through the reference position (i.e. the camera translation vanishes) again with the introduction of noise in the measurements. With Matlab Simulink, we try

## 5.5. Simulation results

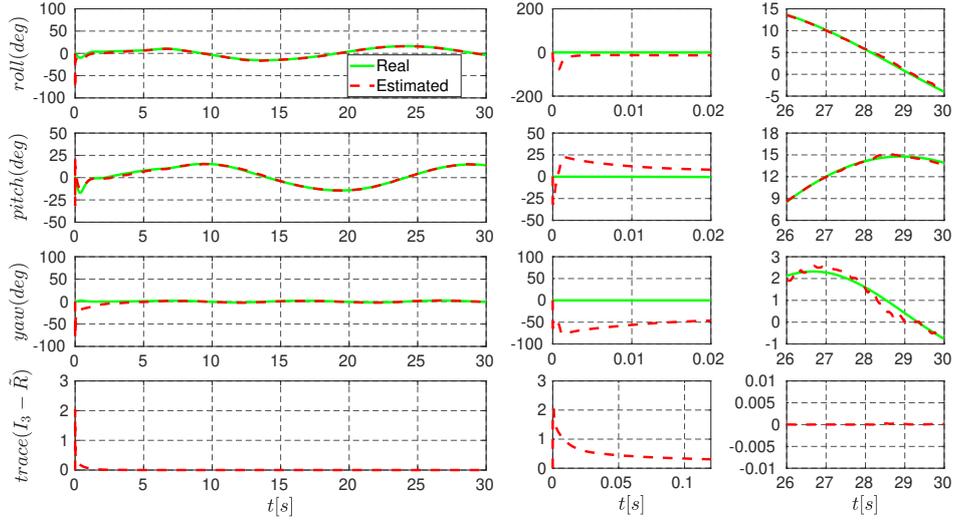


Figure 5.5: **Scenario 1 (Method 2)** – Estimated and real attitudes represented by roll, pitch and yaw Euler angles ( $deg$ ) and the attitude estimate error versus time ( $s$ ) (LEFT) and their zooms (RIGHT).

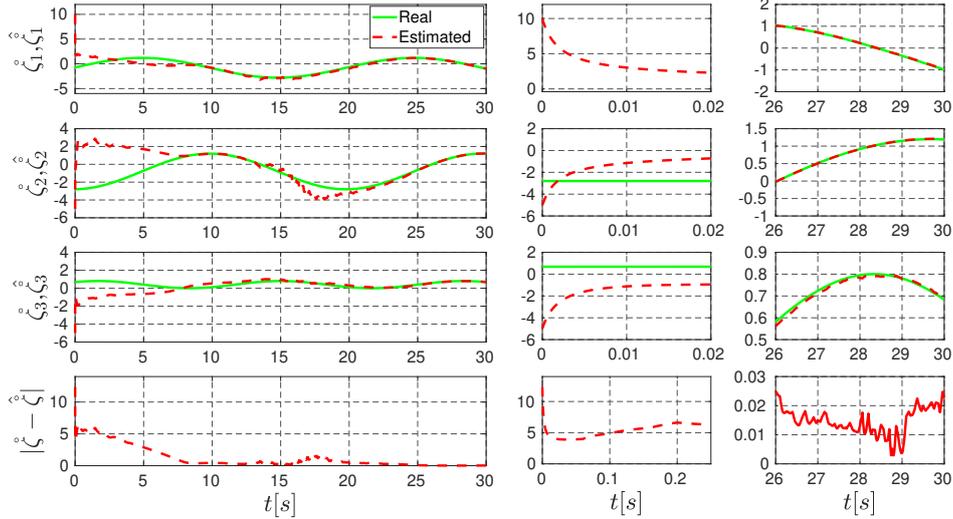


Figure 5.6: **Scenario 1 (Method 2)** – Estimated and real scaled position and norm of the scaled position error versus time ( $s$ ) (LEFT) and their zooms (RIGHT).

to simulate the image noise by introducing white Gaussian noise of the level of about 10 percent of the real values on each individual component of the homography matrix  $H$ . We also introduce white Gaussian noise of variance of  $1(deg/s)$ ,  $0.1$  and  $0.1$  on the measurements of  $\Omega$ ,  $\phi$  and  $\phi_{\perp}$ , respectively. In both simulated scenarios, the matrices  $S$  and  $D^{-1}$  involved in the CRE (2.23) are interpreted as covariance matrices of the additive noise on the system state and output respectively, and the observer is tuned in a similar way like Kalman-Bucy filters. Thus, the following parameters are

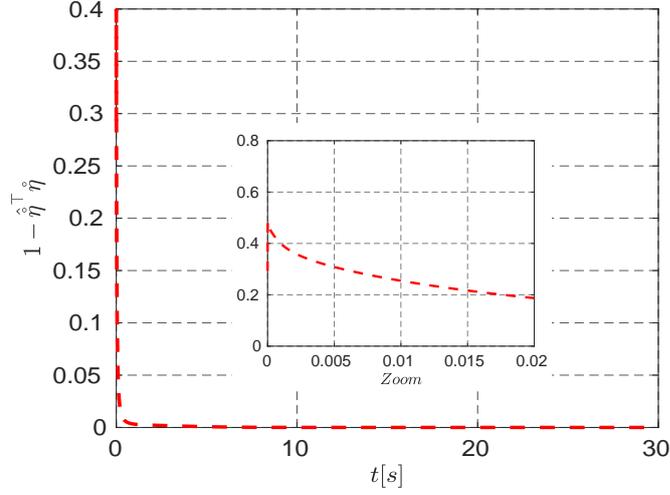


Figure 5.7: **Scenario 1 (Method 2)** – Estimation error of the normal vector estimate represented by  $1 - \hat{\eta}^\top \eta$ .

chosen:  $P(0) = 50I_9$ ,  $D = 100I_9$ ,  $S = \text{diag}(0.1I_2; 0.1I_3; 0.1I_3)$ . The following initial estimation errors for Method 1 (resp. Method 2) are considered for both scenarios<sup>1</sup>:  $\tilde{\zeta}(0) = \hat{\zeta}(0) = [10, -5, 5]^\top$ ,  $\text{quat}_{\tilde{R}}(0) = [0.0436, 0.2586, 0.965, 0]^\top$  (corresponding to errors in roll, pitch and yaw Euler angles of  $178.7(deg)$ ,  $-4.8(deg)$ ,  $-150(deg)$ , respectively),  $\text{quat}_{\hat{Q}}(0) = \text{quat}_{\hat{Q}}(0) = [0.9239, 0.3827, 0, 0]^\top$  (corresponding to an angle error of  $45(deg)$  between  $\eta(0)$  and  $\hat{\eta}(0)$  (resp. between  $\dot{\eta}$  and  $\hat{\dot{\eta}}(0)$ )), where the scene is chosen such that  $\dot{\eta} = e_3$  and  $\dot{d} = 5(m)$ .

**Scenario 1:** For this particular case the drone is commanded to perform a complex periodical trajectory in the inertial frame given by

$$\xi = [10\cos(t/\sqrt{10}) - 4, 10\sin(t/\sqrt{10}) - 4, 2\sin(0.3\pi t/2) - 4]^\top$$

The time evolutions of the estimated and real attitudes (represented by Euler angles), the attitude error estimate (represented by  $\text{trace}(I_3 - \tilde{R})$  for Method 1 and by  $\text{trace}(I_3 - \bar{R})$  for Method 2) as well as the scaled position error estimate and the estimation error of the normal vector to the planar scene (represented by  $1 - \hat{\eta}^\top \eta$  for Method 1 and by  $1 - \hat{\dot{\eta}}^\top \dot{\eta}$  for Method 2) are shown in Figs. 5.2–5.7 for both the methods, respectively. From these figures, it can be clearly seen that all the estimated variables converge to the real ones after a short transition period of a few milliseconds despite the large initial estimation errors. The above results also show that the proposed observers are robust to noisy measurements as the latter marginally affects the overall performance of the proposed observers.

<sup>1</sup>The  $\text{quat}_{(\cdot)}$  notation is used for the unit quaternions of the associated rotation matrix.

## 5.5. Simulation results

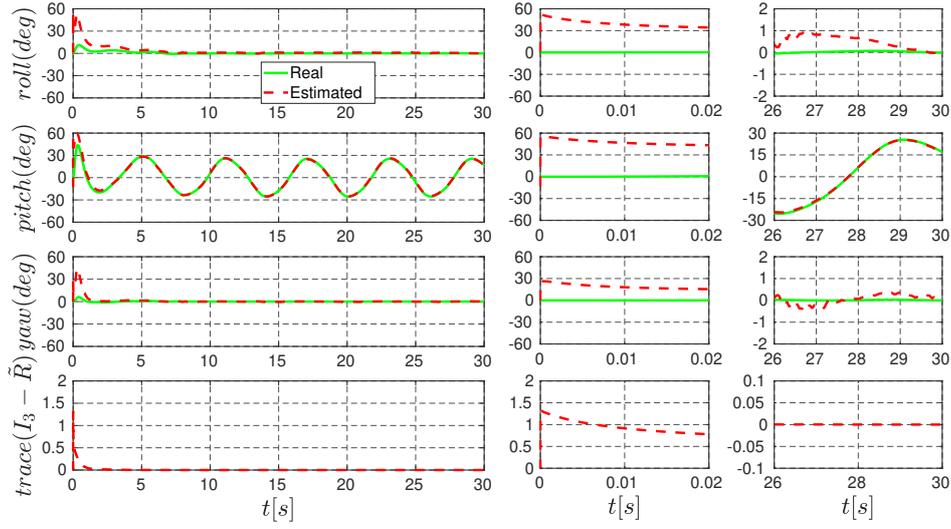


Figure 5.8: **Scenario 2 (Method 1)** – Estimated and real attitudes represented by roll, pitch and yaw Euler angles ( $deg$ ) and the attitude estimate error versus time ( $s$ ) (LEFT) and their zooms (RIGHT).

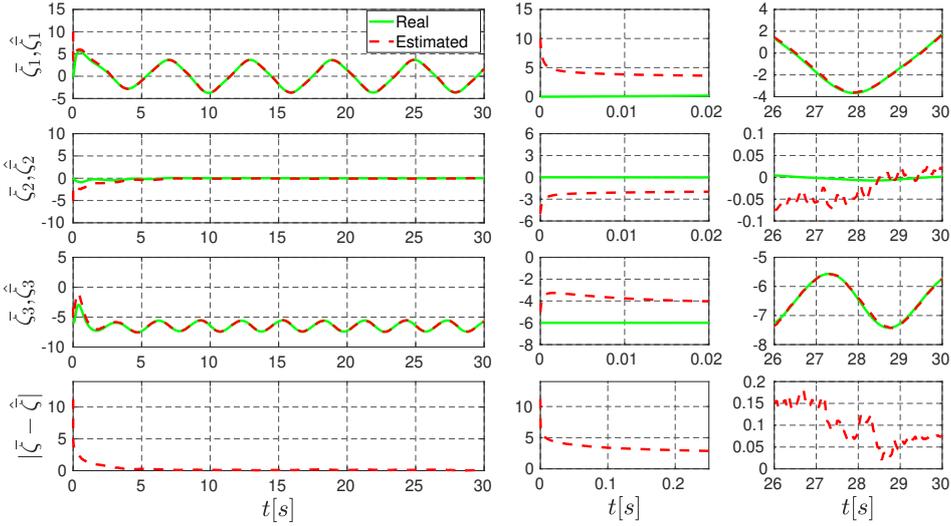


Figure 5.9: **Scenario 2 (Method 1)** – Estimated and real scaled position and norm of the scaled position error versus time ( $s$ ) (LEFT) and their zooms (RIGHT).

**Scenario 2:** In this case the drone performs a rectilinear sinusoidal trajectory that passes through the reference position (i.e. the camera translation vanishes) with added noise on the measurements of  $H$ ,  $\phi$ ,  $\phi_{\perp}$ , and  $\Omega$  as in the previous simulation. The reference trajectory in the inertial frame is given by  $\xi = [5\sin(\pi t/3), 0, 0]^T (m)$ . It is worth noting that in such a degenerate case (i.e. when the camera translation vanishes) all traditional algebraic approaches fail to obtain a correct estimate of the normal vector (there exists in fact an infinity of solutions) whereas our approach still works since the uniform observ-

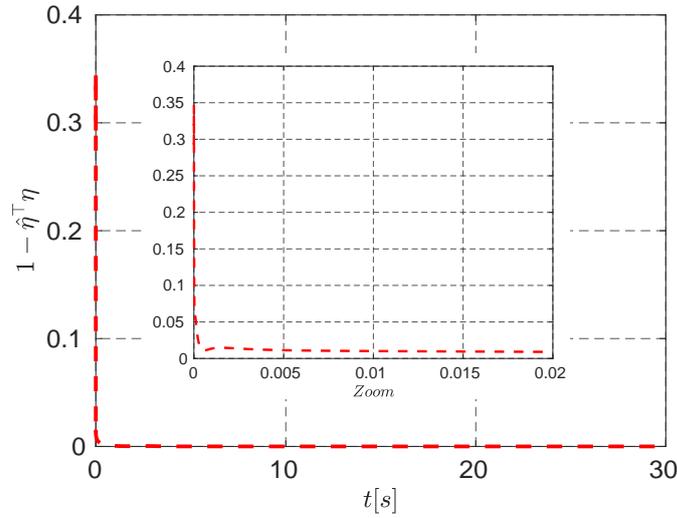


Figure 5.10: **Scenario 2 (Method 1)** – Estimation error of the normal vector estimate represented by  $1 - \hat{\eta}^T \eta$ .

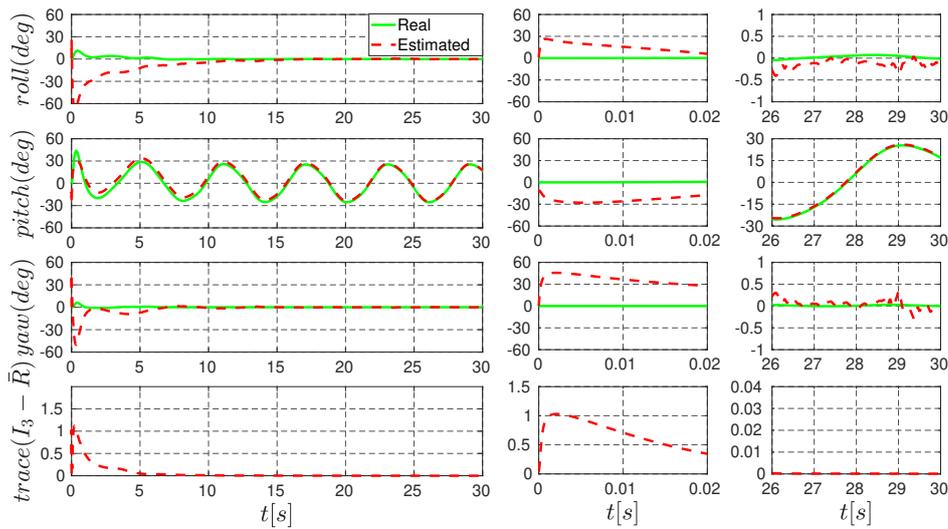


Figure 5.11: **Scenario 2 (Method 2)** – Estimated and real attitudes represented by roll, pitch and yaw Euler angles (*deg*) and the attitude estimate error versus time (*s*) (LEFT) and their zooms (RIGHT).

ability conditions (5.20) for Method 1 and (6) for Method 2 are always satisfied. From Figs. 5.8–5.13 it can be clearly seen that for this degenerate case and even with the addition of noise to the system measurements, the proposed observers are robust enough to provide a convincing performance in terms of convergence rate, smooth transient phase, and filtering of measurement noise.

## 5.6. Experimental validations

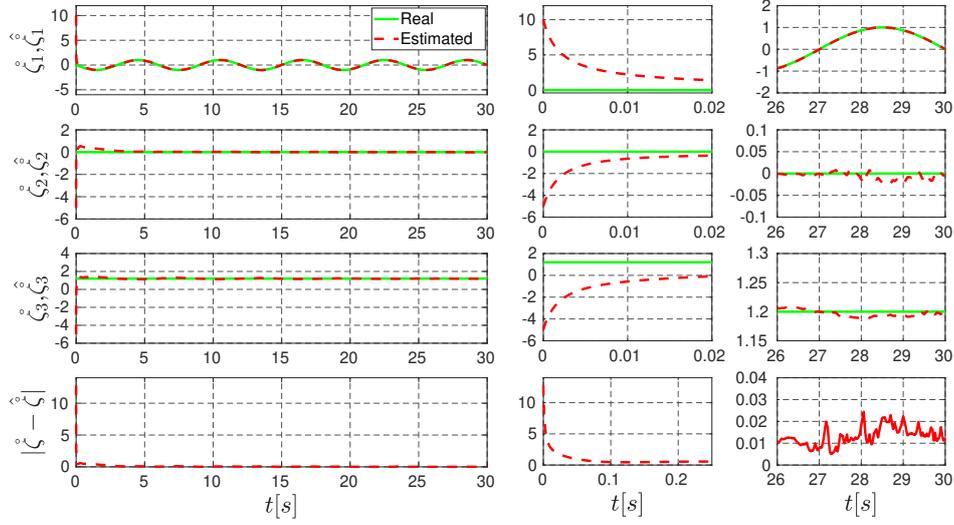


Figure 5.12: **Scenario 2 (Method 2)** — Estimated and real scaled position and norm of the scaled position error versus time (s) (LEFT) and their zooms (RIGHT).

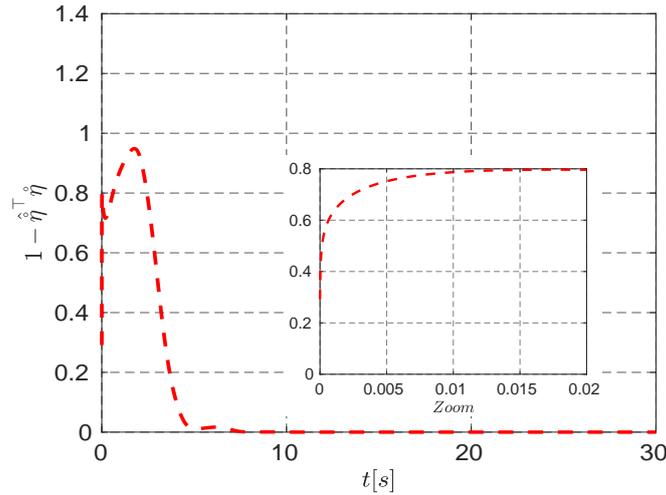


Figure 5.13: **Scenario 2 (Method 2)** – Estimation error of the normal vector estimate represented by  $1 - \hat{\eta}^\top \eta$ .

## 5.6 Experimental validations

In this section, experimental results are reported showing the comparative performance of the proposed observer (i.e. Method 2) w.r.t. the algebraic algorithm proposed by Malis and Vargas [53] that we call **Method 3**. The two methods used for comparison purposes have two principal parts:

- the vision part providing the homography estimate;
- and the homography decomposition part applied to the inverse of the homography

estimate.

Since our main interest consists in performing a comparative analysis between the two homography decomposition algorithms, the vision part are common for both of them.

**Remark 7.** *For experiments we use a hybrid continuous-discrete version of the proposed observer which is similar to the one in Chapter 4 in section 4.4.*

### 5.6.1 Experimental setup

The experimental setup makes use of a Camera-IMU system consisting of a Basler Ace Pylon camera and a MPU-9250 IMU (see Fig. 5.14). The camera provides images at a frequency of 25 frames per second with a resolution of  $1280 \times 1040$  pixels, and the IMU output data rate is 100 Hz. Data transmission from the camera as well as the IMU to the PC is carried out using a USB 3.0 cable. The highly accurate OptiTrack motion capture system is also used to obtain ground-truth data (i.e position and attitude at a frequency of 200 Hz) for comparison purposes. All data are time synchronized using ROS.

The homography matrix is acquired by using the **HomographyLab**<sup>2</sup> library developed by our team based on the homography observer proposed in [34]. The HomographyLab library has been implemented in C++ combined with OpenCV for image processing (i.e., *FAST Feature Detector* and *ORB Descriptor Extractor* functions for feature detection and descriptor extraction in images). Real-time and robustness (w.r.t. fast camera motions, occlusions, image blurs, sudden changes in light intensity, etc.) are the main advantages of HomographyLab and the implemented algorithm with respect to the state-of-the-art codes and algorithms (see [34] for more details).

The proposed observer for the decomposition of the homography inverse (i.e. Method 2) has been implemented in C++ on an Intel Core i7-6400 CPU running at 3.4 GHz, and the entire implementation that includes the image processing as well as the estimation part runs online at a frequency of 25 Hz. For the sake of simplicity, the measurement of optical flow divergence and translational optical flow were not used for the proposed observer in the reported experiment, but this fact only had a negligible effect on the observer's performance as attested by the experiment reported hereafter.

The decomposition solution of Method 3 are obtained using the existing OpenCv's `cv::decomposeHomographyMat` function that, however, gives 4 possible solutions. We then reduce these set of solutions to only 2 by verifying the positive depth constraint (i.e. third component of the normal vector should be positive). From the remaining two solutions, the one with the estimated normal vector closest to  $\hat{\eta} \cong e_3$  is chosen. This choice is made based on the way the experiments are performed by keeping the planar

<sup>2</sup><http://sdb3.i3s.unice.fr/homographyLab/>

## 5.6. Experimental validations

---

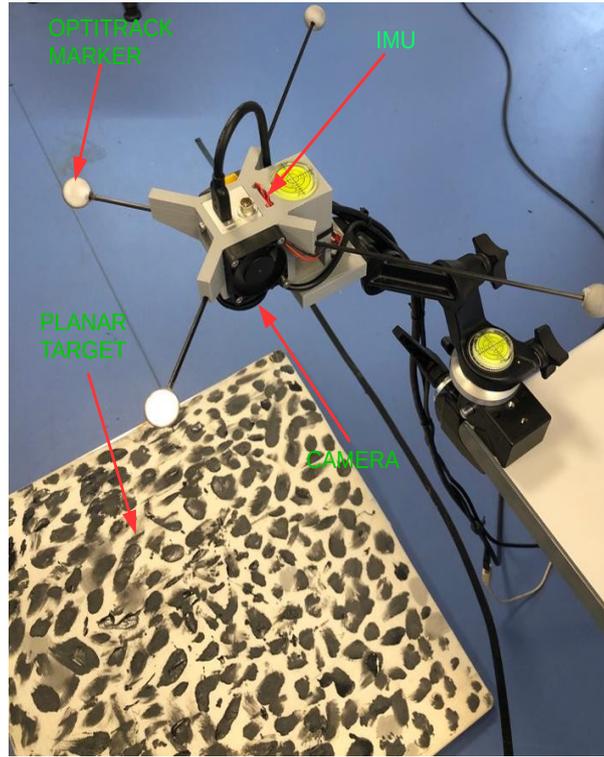


Figure 5.14: Experimental setup consisting of a Camera-IMU system looking at a textured planar target and OptiTrack markers for ground-truth data.

target almost perpendicular to the camera. This implies that if we have a good estimate of the normal vector it should be approximately equal to  $e_3$

### 5.6.2 Experimental results

The reported experiment has been performed in real time with the Camera-IMU system pointing downwards and observing a well-textured horizontal planar target (with the normal vector  $\hat{\eta}$  approximately equal to  $e_3$ ), and by moving this system along various directions and orientations. The gains involved in the proposed observer were chosen as follows  $P(0) = 2I_8$ ,  $D = 100I_8$ ,  $S = \text{diag}(0.005^2 I_2, 0.025^2 I_3, I_3)$ . The following video link <https://youtu.be/SID2JDe4cZI>

showing this experiment is provided as a supplementary material.

The plots in Figs. 5.15 and 5.16 (and their zooms in Figs. 5.18 and 5.19) and in Fig. 5.17 show the comparison of the attitude estimates (represented by the Euler angles), the scaled position estimates, and the normal vector estimates using Method 2 and Method 3 and the corresponding ground-truth variables.

It can be observed from these figures that although the algebraic algorithm (i.e. Method 3) provides rather good estimation of the attitude and the scaled position in

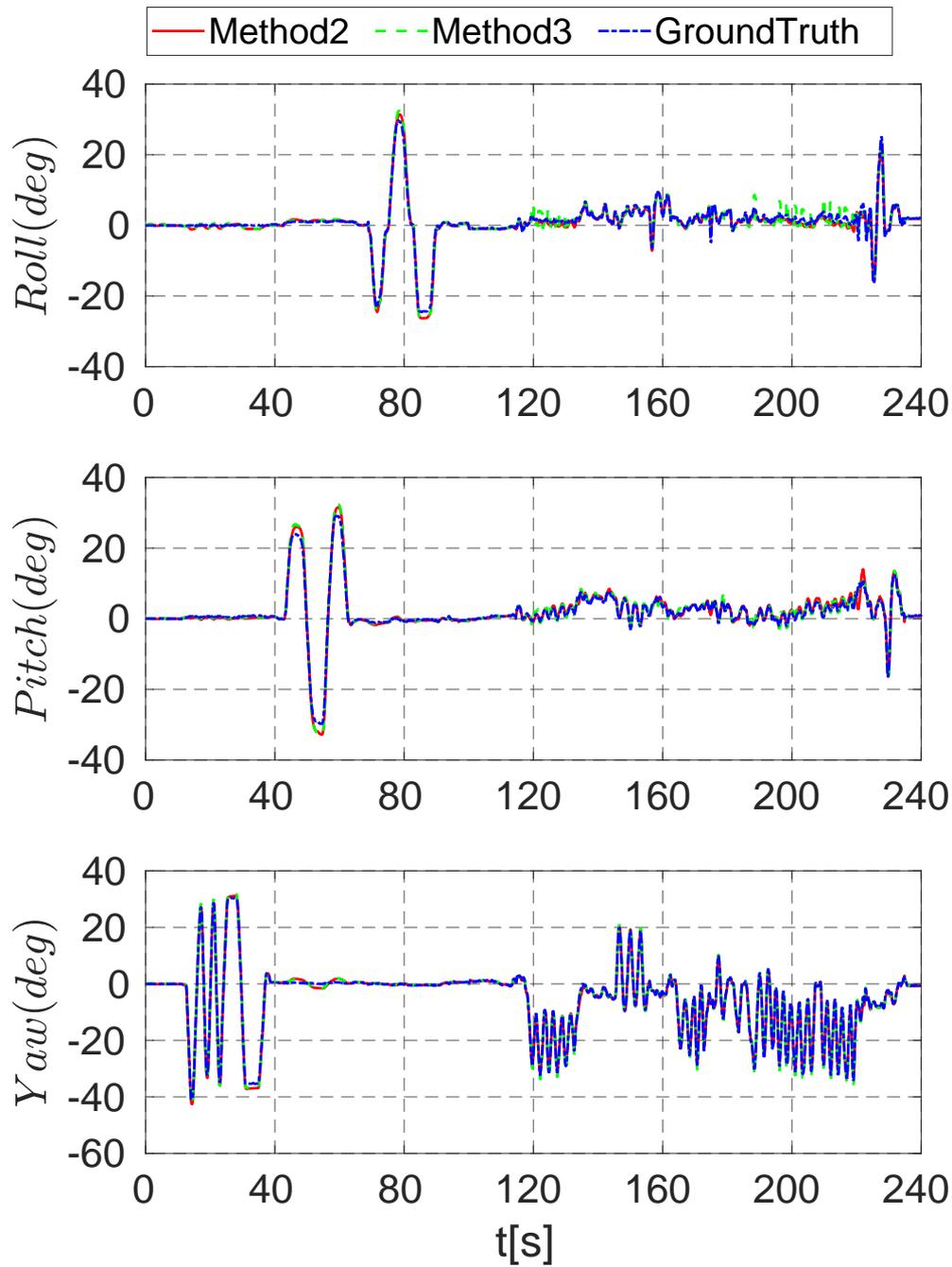


Figure 5.15: Attitudes estimated by Method 2 (Red) and by Method 3 (Green), and corresponding ground-truth attitude (Blue) versus time.

comparison with ground-truth data when the camera's translation is large enough, it tends to give erroneous estimation when the latter evolves near to zero (i.e. the degenerate case). Indeed, the zoomed plots from the time 115s to 130s shown in Fig. 5.18 and

## 5.6. Experimental validations

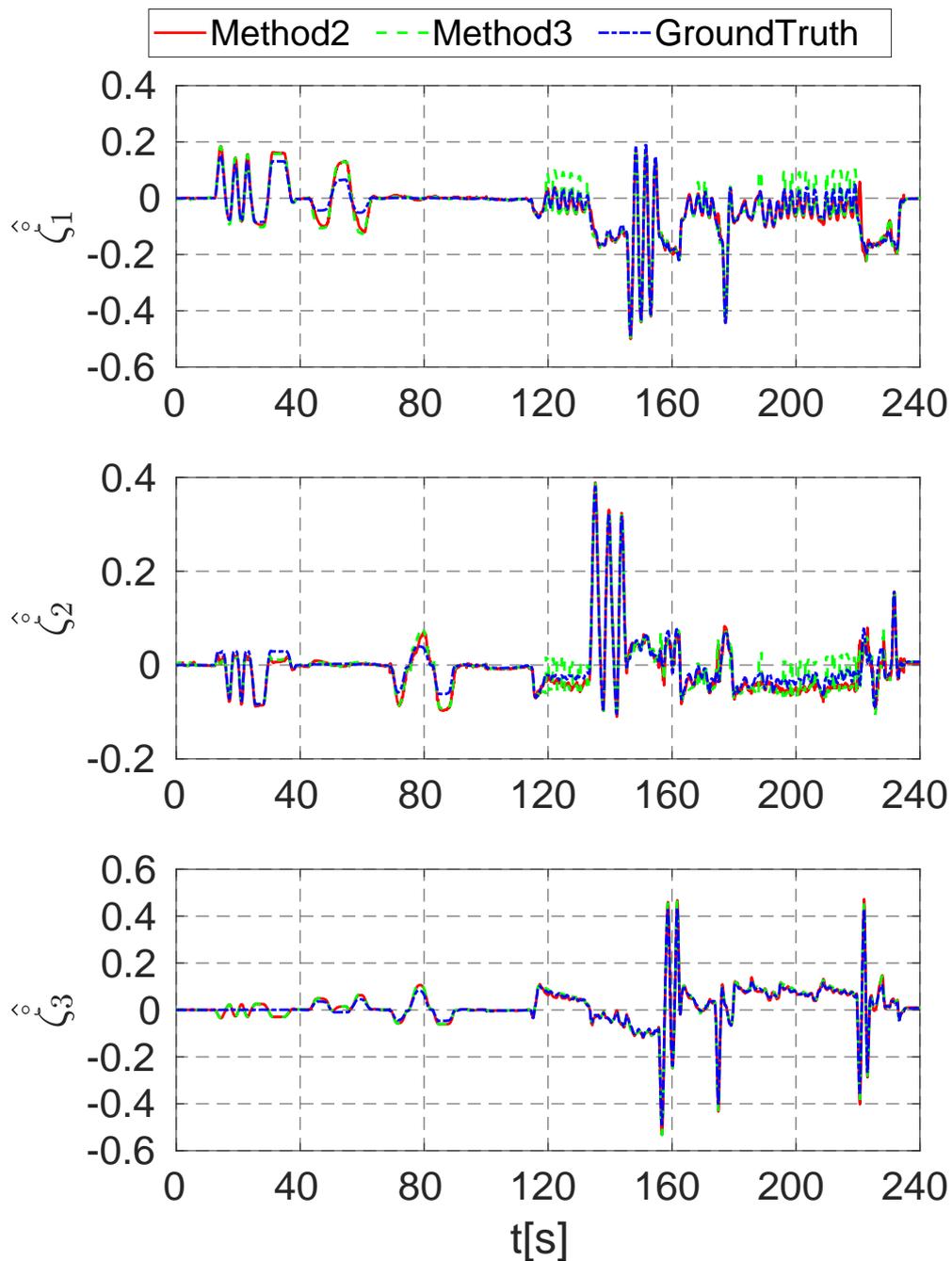


Figure 5.16: Scaled position estimated by Method 2 (Red) and by Method 3 (Green), and corresponding ground-truth attitude (Blue) versus time.

from the time 195s to 220s shown in Fig. 5.19 clearly highlight such poor behaviour. This can be explained by the fact that the choice of the correct solution amongst the 4

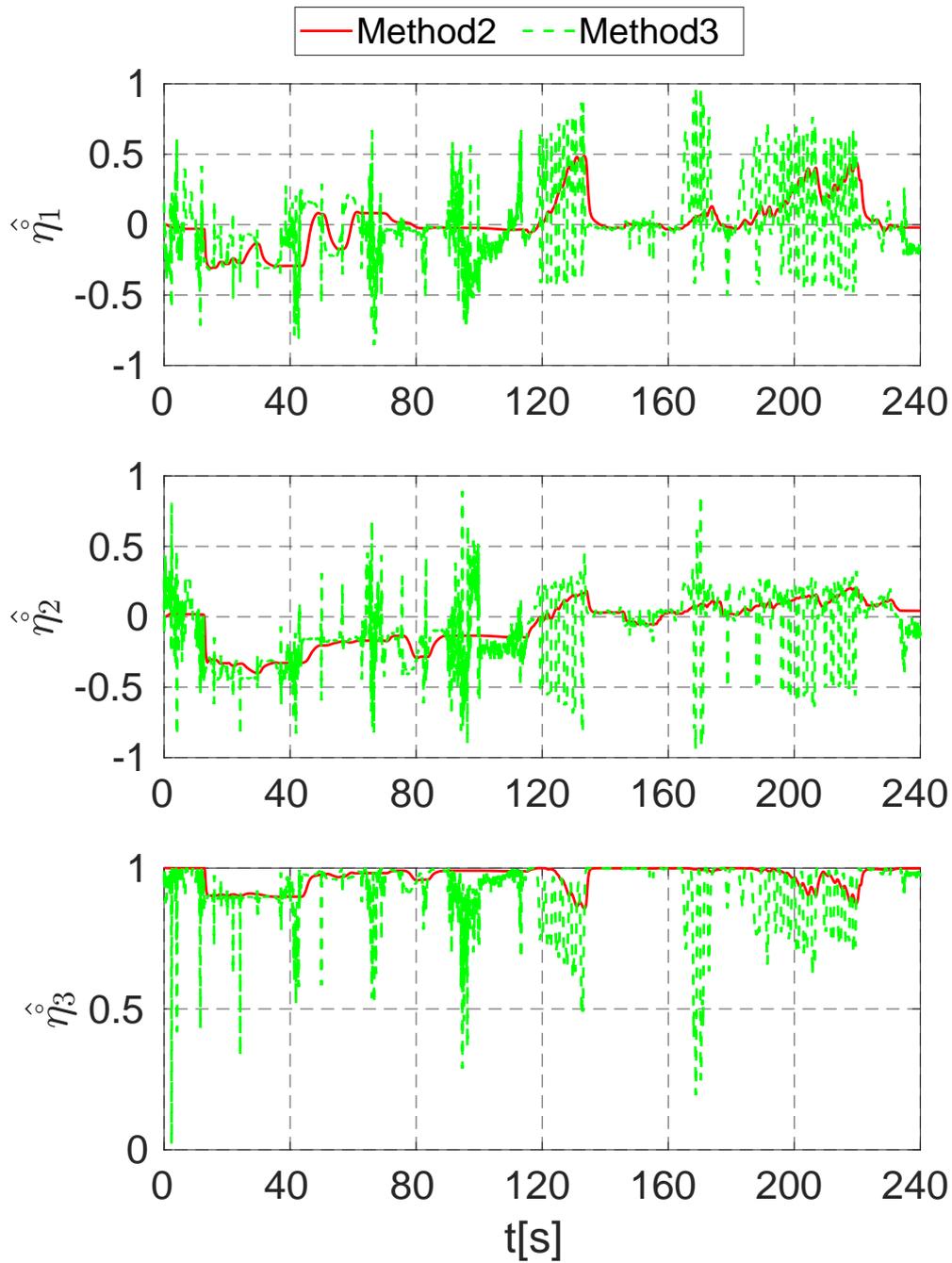


Figure 5.17: Normal vector estimated by Method 2 (Red) and by Method 3 (Green).

possible algebraic ones relies heavily on the precision of the normal vector estimate. The latter obtained from algebraic approaches is, however, very sensitive to noise and can be wrongly estimated when camera motion evolves near to the degenerate situation (see

## 5.6. Experimental validations

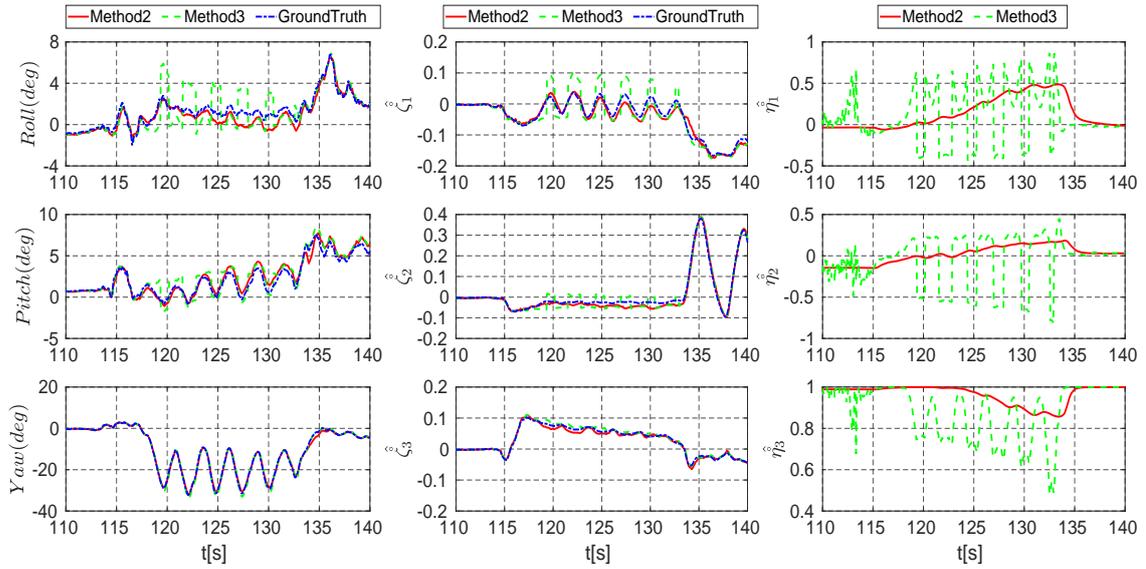


Figure 5.18: Zoom of the estimated attitude, scaled position estimates and the normal vector from time  $t = 110s$  to  $t = 140s$ .

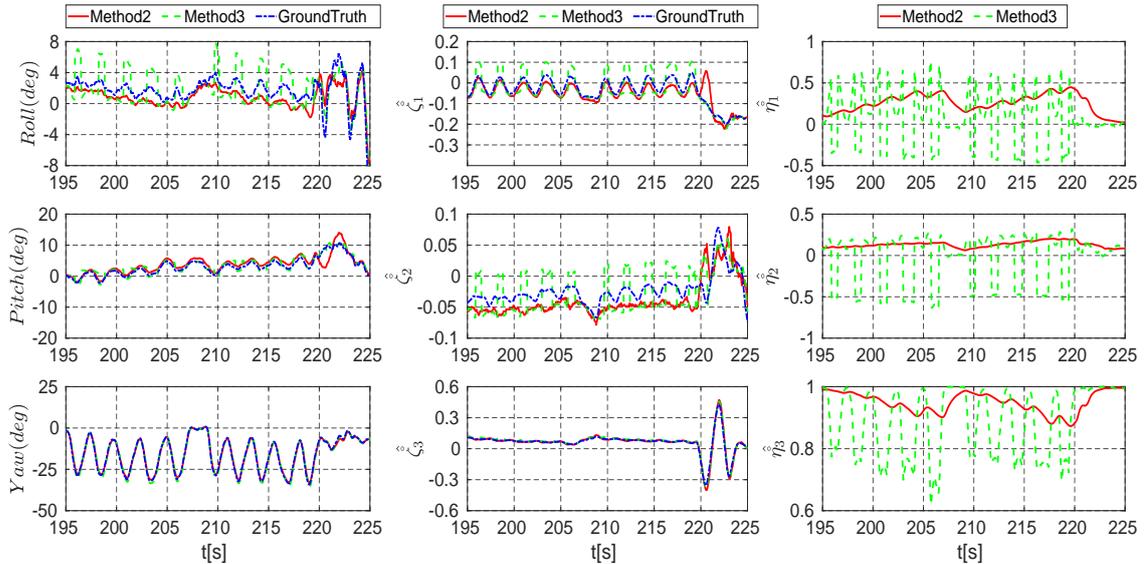


Figure 5.19: Zoom of the estimated attitude, scaled position estimates and the normal vector from time  $t = 195s$  to  $t = 225s$ .

the green dashed curves in Fig. 5.17 and in the corresponding zoomed plots in Figs. 5.18 and 5.19).

In contrast, the normal vector estimate of the proposed observer (i.e. Method 2) is much less noisy than the one obtained from the algebraic algorithm (i.e. Method 3) and

may be seen as the filtered signal of the latter, thus showing the filtering interest of our approach w.r.t. traditional algebraic algorithms. One observes that non-negligible but bounded error of the normal vector estimate has occurred during the time periods  $[10s, 30s]$ ,  $[120s, 130s]$  and  $[195s, 115s]$  when the camera's translation is very small but its yaw motion is significant. We believe that this bounded error is rather more related to the imperfection of the camera calibration and to the sensitivity of the estimated homography (and its decomposed normal vector solution) w.r.t. image noise and low resolution when the camera's translation is not large enough, than a fast drift due to measurement noise (the proof is that from the time  $90s$  to  $100s$  when the camera's pose nearly superpose to the reference one, the normal vector estimate always remains near to  $e_3$ ). Moreover, whenever the camera's translation is large enough, the decomposed normal vector estimate becomes quite precise. Most importantly, the attitude and scaled position estimated by our algorithm always evolve closely to the corresponding ground-truths, even in the degenerate situation.

The reported experimental results have thus illustrated the convincing performance and robustness of the proposed approach, showing its advantage with respect to the state-of-the-art algebraic approach.

# 6

## Implementation Details of the HomographyLab Software Library

HomographyLab (Lab is an abbreviation of LABORatory) is a library for Homography Estimation written in C++ combined with OpenCV (both CPU and GPU implementations are available). This library has been evaluated at the TRL 7 (Technology Readiness Level) and is protected by the French APP (Agency for the Protection of Programs). This library implements the homography observers proposed in the papers [34, 35]. Real-time and robustness (with respect to fast camera motions, occlusions, image blurs, sudden changes in light intensity, poor image quality, etc.) are the two principle distinguished features of HomographyLab and the implemented algorithm with respect to the state-of-the-art codes and algorithms. The library has been designed with various sub modules written in separate C++ classes like feature extraction and detection, feature matching, nonlinear homography observer, observers for homography decomposition. Therefore, the user can also use these individual libraries to write their own custom code and thus save significant time and resources. It has also been successfully implemented in various robotic and computer vision applications such as image stabilization, station-keeping of an AUV, and linear velocity estimation.

## 6.1 Software Implementation Details

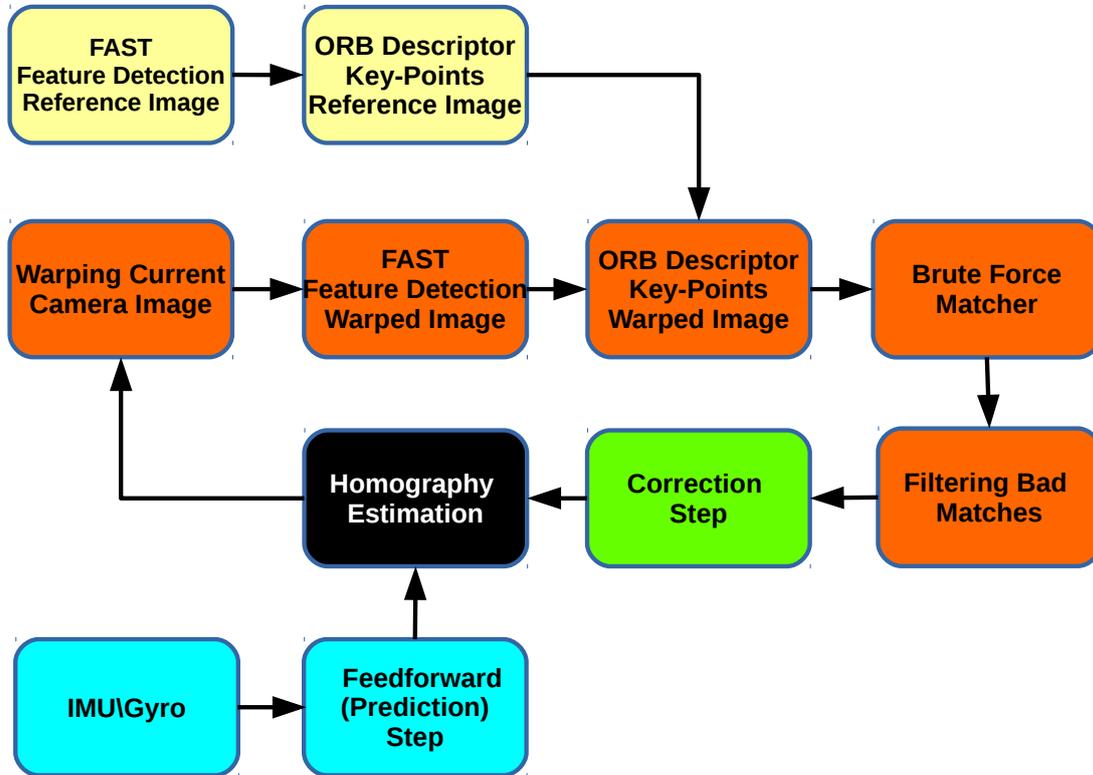


Figure 6.1: Flowchart of the Homography Estimation Algorithm

The proposed algorithm has been implemented using C++ with OpenCV library. Figure 6.1 shows the schema of the algorithm which can be summarized as follows. The process is initialized by detecting the features and the descriptors from the reference image. As soon as a new image arrives, it is transformed with a perspective transformation (warped using the OpenCV's `warpPerspective` function) based on the predicted homography estimate obtained from the gyro measurements. The algorithm detects the key-points from this current warped image and then extracts the features descriptor. The descriptors of the current image are matched with the descriptors of the reference image. We use the brute-force matching algorithm since it is more adept to translational motion than rotational motion, and most of the rotational motion has already been compensated for by forward integrating the angular velocity during the prediction step. The outliers are removed using the M-estimator like observer proposed in [34] which greatly enhances the precision as well as the robustness of the homography estimate. After performing the feature detection and matching, the homography estimate is then updated (correction step of the observer) by iterating the observer equations 300

## 6.2. I3S-AUV SOFTWARE ARCHITECTURE

---

times per video frame. The steps of feature detection and matching which are essential parts for the algorithm are computationally quite expensive. In the literature, there exists several feature detectors such as speeded-up robust features(SURF)[4], features from accelerated segment test (FAST)[67], scale invariant feature transform (SIFT)[46], oriented FAST and rotated BRIEF (ORB)[68] and descriptors such as ORB, SIFT and SURF. We decided to use the combination FAST-ORB for the simple reason that they are very time-efficient and thus more suited for real-time implementation. HomographyLab software library has a CPU as well as a GPU version. The library was tested across several hardware platforms some of which are shown in the Figure 6.2 to compare the performances. The main goal behind creating the GPU version of the library was not only to improve the computational performance but also to fully exploit the nvidia cards shown in the figure 6.2 that can be easily mounted on aerial vehicles or AUV's (autonomous underwater vehicles) for onboard computations. In the next subsection, I will briefly introduce the software architecture in one of the UAV platforms developed by members of the I3S-OSCAR team to show how exactly the vision module is used in the overall architecture.



Figure 6.2: (a) Nvidia Jetson TX1 (b) Nvidia Jetson Xavier AGX (c) Nvidia Jetson Xavier NX

## 6.2 I3S-AUV SOFTWARE ARCHITECTURE

The platforms showed in the figure 6.3 have been completely developed in the laboratory for experimental validations of the vision as well as the control algorithms. With the HomographyLab software library we perform homography based stabilization and positioning of our vehicles for both forward looking and downward looking configurations as can be seen in the videos

<https://www.i3s.unice.fr/oscar/node/11>

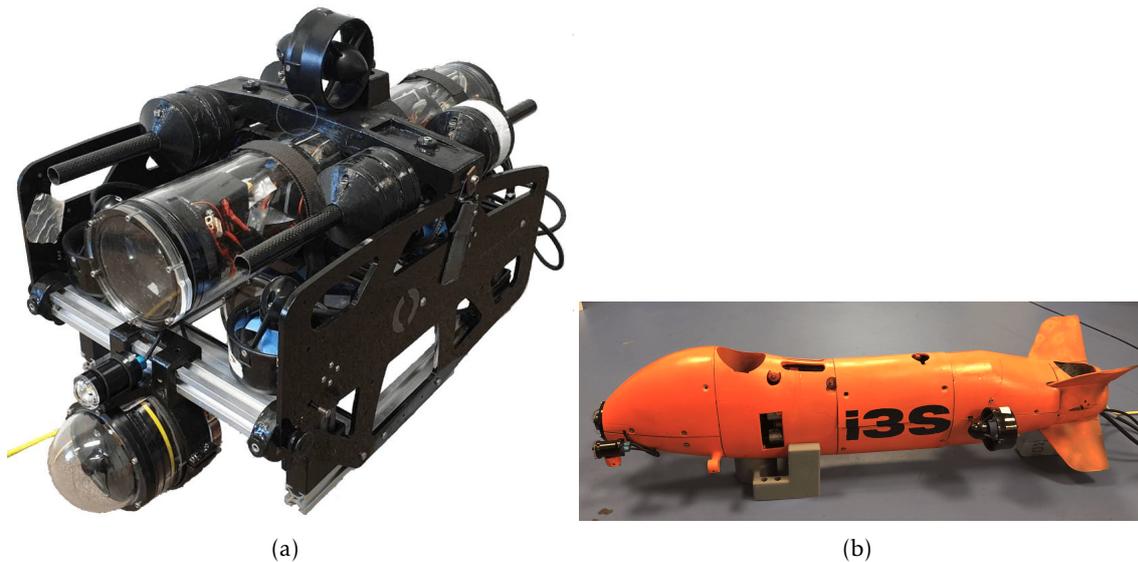


Figure 6.3: (a) Platform 1 (b) Platform 2

### 6.2.1 Software architecture in ROS

Currently we are using ROS drivers for the camera and the IMU. The pressure sensor is connected to the Pixhawk over I2C connection and it is operated using a driver running in the Pixhawk. The drivers for the camera and the Pixhawk are currently running on the companion computer (Jetson Xavier NX in our case) in order to obtain the images, IMU data as well as the pressure data. Apart from these drivers, there are two main programs running on the Xavier NX:

- **ROS node for estimating the homography and for performing homography decomposition:** This is the node highlighted in red in the figure 6.4. The input to this node are the images and the IMU data (either coming from the IMU attached to the camera or the IMU of the pixhawk). The synchronization of the images and the IMU reading is done via ROS. Internally this node calls the functions from the HomographyLab library. The output of this node are two matrices: Inverse of the Euclidean Homography  $\mathbf{G}$  and the transpose of the rotation matrix estimate  $\hat{\mathbf{R}}^\top$ . One important thing to note here is that the HomographyLab library is written purely in C++.
- **ROS node for control:** This node runs the control algorithms whose inputs are the two matrices (i.e. Inverse of the Euclidean Homography and the tranpose of the rotation matrix), the IMU data and the pressure data in order to generate the control forces and torques that are further sent to the Pixhawk. Apart from this we also send the control status (manual mode, emergency stop, etc.) to the Pixhawk.

## 6.2. I3S-AUV SOFTWARE ARCHITECTURE

This control status input is read through the joystick. The joystick is connected to the ground station on which we have a ROS driver running. The inputs of the joystick are read in this node and then sent to the Pixhawk.

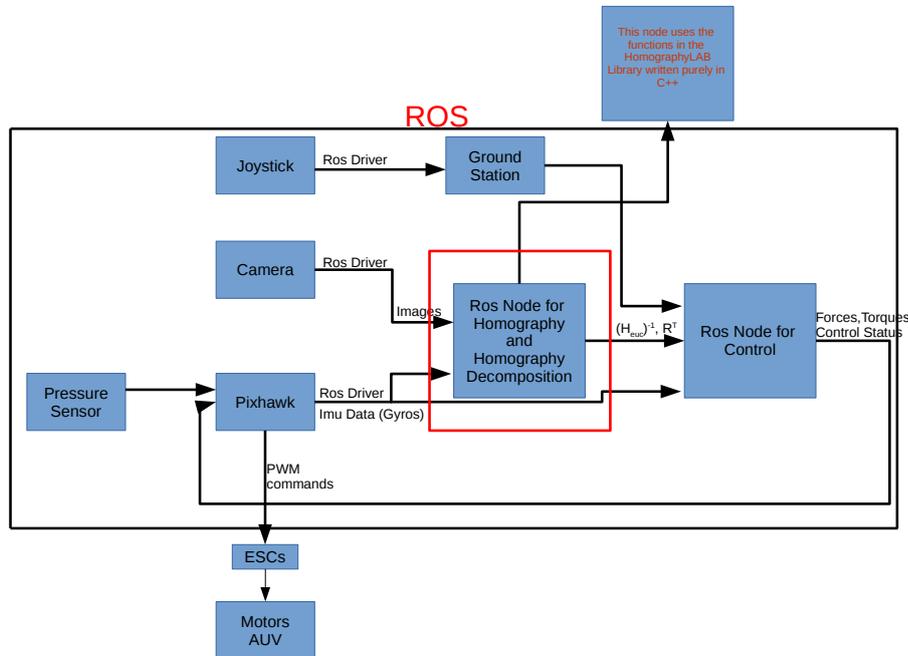


Figure 6.4: Software architecture using ROS as middleware

The communication between the two nodes is carried out using Publisher/Subscriber service which is an inbuilt functionality in ROS.

**Pixhawk:** The Pixhawk is also used for the control allocation. The control forces and torques that are obtained from the ROS node for control are converted into pwm commands and then sent to the ESCs.

### 6.2.2 Ros Nodes used in our Architecture

Note that all nodes are in oval while topics are in rectangular. From the figure 6.5 you can see that we have five principal nodes:

- */pylon\_camera\_node*: Camera node for receiving the images.
- */homdecompobserver*: Node for computing the Euclidean homography inverse and the transpose of the Rotation matrix which is sent as an input to the Control node
- */joy\_node*: Node for getting the inputs from the joystick used for manual control
- */control\_in\_ros\_node*: Control node used for generating the control forces and the control torques

- */mavros*: Pixhawk node for getting the imu data, receiving the data from the pressure sensor and then finally converting the control forces and torques into pwm commands which are sent to the motors.

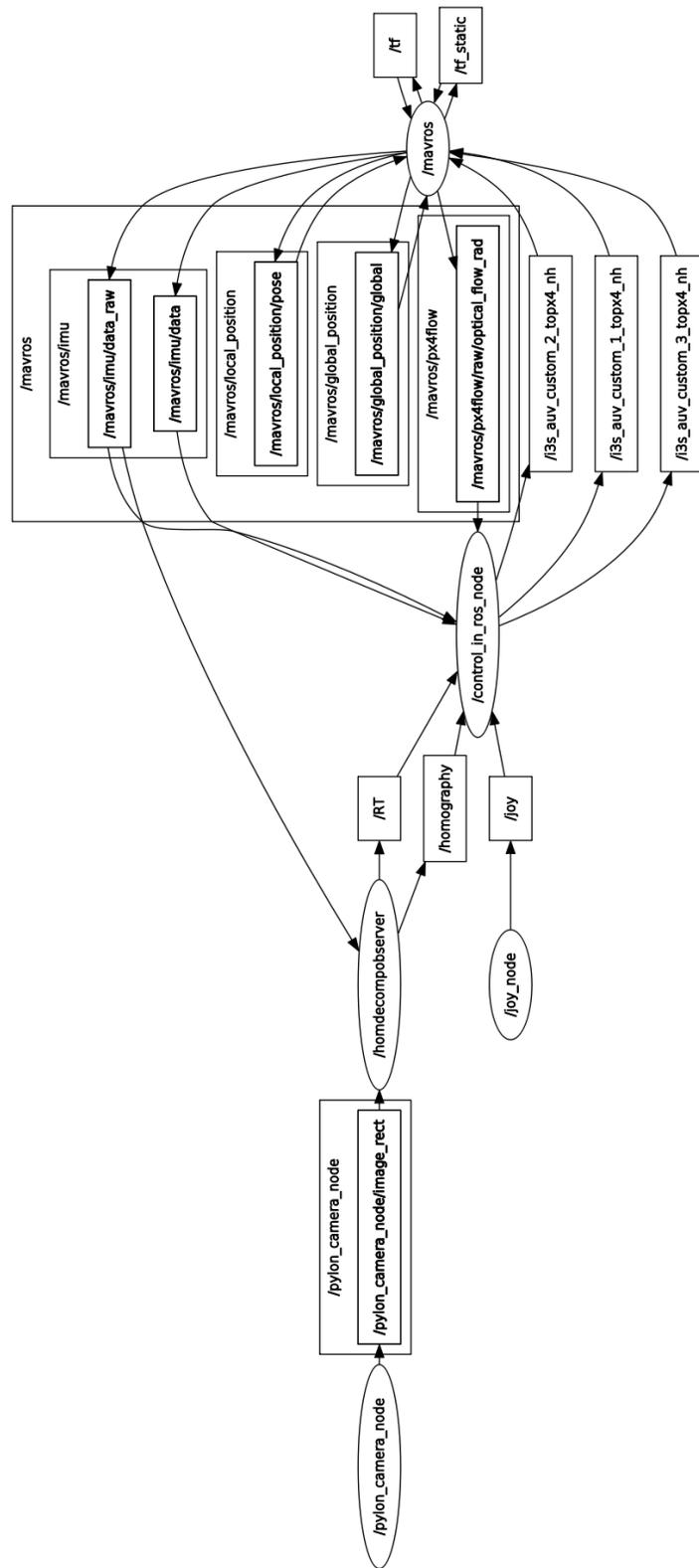


Figure 6.5: ROS nodes and topics in I3S Architecture

### 6.2.3 Transfer of Data between Ros Nodes using Topics

In order to transfer data between different nodes we use the Publisher and Subscriber service in ROS. Normally the data is published on "Topics" and then you can read this data from other Ros nodes. From the figure 6.5 you can see that the input to the node */homdecompobserver* are the images published on the topic */pylon\_camera\_node/image\_rect* and the data coming from the Pixhawk IMU published on the topic */mavros\_imu/data\_raw*. The output of this node is the inverse of the Euclidean Homography matrix that is published on the topic */homography* and the transpose of the Rotation matrix published on the topic */RT*. On the other hand the */control\_in\_ros\_node* subscribes to the topics */RT*, */homography*, */joy\_node* and also */mavros\_imu/data\_raw* in order to receive the data and then publishes the control forces and the control torques on the topics */i3s\_auv\_custom\_1\_topx4\_nh* and */i3s\_auv\_custom\_2\_topx4\_nh*. Of course in the figure you can see other ros topics as well but the most important ones are mentioned in the explanation above.

# Conclusion

**T**HIS thesis addresses many aspects which are particularly related to the state estimation and navigation of autonomous vehicles. They range from visual-inertial sensor fusion, homography estimation and decomposition, design of nonlinear observers based on Lie groups for the estimation of attitude, linear velocity etc. that can be further used for safe and precise navigation for a wide range of autonomous vehicles.

## **Context and contributions of the thesis:**

The contributions reported in the first part of the thesis constitute a continuation of prior work of the I3S-OSCAR team [31] on the topic of nonlinear observer design for homography estimation with application to image stabilization, image mosaicing etc. Furthermore, we use Visual-Inertial sensor fusion for the development of integrated nonlinear observers based on the Riccati observer design framework [24] that exploit optical flow estimates and IMU readings to estimate the camera's attitude, linear velocity, and its distance to a planar target. A key thing to note here is the use of a suite of low-cost sensors consisting of a monocular camera and a MEMS imu. In the context of monocular vision, the assumption on the planarity of the visual target is considered here, resulting in the meaningful involvement of the homography in the observer design. Despite such a restrictive assumption, the proposed solutions are still relevant for a number of applications in the fields of aerial, ground as well as underwater robotics. Two nonlinear observers have been proposed in this part, the first one for the estimation of homography and the second one for the estimation of partial attitude, linear velocity and depth estimation and the results constitute the subject of the following publications or submissions [55], [30].

The work presented in Part 3 presents a novel approach for tackling the classical problem of Homography Decomposition. The novelty of this contribution lies in the

design of two deterministic Riccati observers for addressing the homography decomposition problem instead of solving it on a frame-by-frame basis like traditional algebraic approaches. The large domain of convergence and good performance of the proposed observers have been demonstrated through both simulation results and extensive experimental validations.

The work presented in Part 4 talks about the software implementation details of the HomographyLab library which is partially an accumulation of the observers presented in the previous parts. This library has served as a key basis in particular for the development of homography based visual servoing control for fully/under actuated AUV's [61]. The development of the HomographyLab library and of the I3S-AUV platform has been a defining factor for successfully demonstrating the developed theories and concepts as an appealing success factor to the public, the specialized robotics community and industrial companies.

### **Perspectives:**

The work, already done in the thesis, on the design of nonlinear Riccati observers for state estimation has allowed us to get a really good and thorough understanding of this powerful framework that can be used for a wide range of applications. This constitutes the first step for us to address in the near future other challenging applications such as autonomous landing of aerial vehicles and obstacle avoidance amongst many others. One of the solutions to deal with this problem is to exploit the filtered optical flow estimates proposed in Chapter 3. Also the observer proposed to estimate the linear velocity, partial attitude and depth could be thought of as a low cost sensor to provide these measurements. It can be utilized in a wide range of applications in which linear velocity measurements are not available and a reliable estimate of the linear velocity and the pose is need for control purposes. This is the case of low cost small-scale vehicles that operates in GPS denied environment for which the cost of body velocity sensors such as sensors based on Dopplereffect or ad-hoc inertial velocity sensor systems are usually prohibitive. This would also be particularly useful in the field of underwater robotics where the sensors used to measure the linear velocity (Doppler Velocity Log) cost thousands of euros and are relatively bulky. Thus, one of the future works would be to perform experimental validations to have an idea of the effectiveness of the proposed approach using the platforms available in the I3S laboratory.

One of the other problems we would like to address is to take into account the accelerometer bias in the observer design. Thus, we would like to design a novel observer based on the Ricatti design framework for the estimation of linear velocity and pose of a rigid-body along with accelerometer bias correction. We believe that this would certainly improve the quality of the estimates. Lastly with regards to homography and ho-

mography decomposition, we would like to combine directly the point correspondences obtained from images for the estimation of the variables of relative pose and normal vector instead of passing through the entire estimation process. This is a really interesting research direction and will be explored in the new thesis starting shortly in the I3S team.

# Bibliography

- [1] A. Agarwal, C.V. Jawahar, and P.J. Narayanan. A survey of planar homography estimation techniques. Technical report, IIIT/TR/2005/12, IIIT, India, 2015.
- [2] M. Aldeen and H. Trinh. Reduced-order linear functional observers for linear systems. 146(5):399–405, 1999.
- [3] G. Allibert, R. Mahony, and M. Bangura. Velocity aided attitude estimation for aerial robotic vehicles using latent rotation scaling. In *IEEE Int. Conf. on Robotics and Automation*, pages 1538–1543, 2016.
- [4] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool. Speeded-up robust features (surf). *Computer Vision and Image Understanding*, 110(3):346–359, 2008. Similarity Matching in Computer Vision and Multimedia.
- [5] S. Benhimane and E. Malis. Homography-based 2d visual tracking and servoing. *International Journal of Robotics Research*, 26:661–676, 2007.
- [6] S. Bonnabel, P. Martin, and P. Rouchon. Symmetry-preserving observers. *IEEE Trans. on Autom. Control*, 53(11):2514–2526, 2008.
- [7] D. Brescianini, M. Hehn, and R. D’Andrea. Quadcopter pole acrobatics. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3472–3479, 2013.
- [8] C.-T. Chen. *Linear system theory and design*. Oxford University Press, Inc., 1995.

- [9] C. Conomis. Conics-based homography estimation from invariant points and pole-polar relationships. pages 908–915, 2006.
- [10] S. De Marco, M.-D. Hua, R. Mahony, and T. Hamel. Homography estimation of a moving planar scene from direct point correspondence. pages 565–570, 2018.
- [11] E. Malis et al. Dynamic estimation of homography transformations on the special linear group for visual servo control. pages 1498–1503, 2009.
- [12] E. Malis et al. Dynamic estimation of homography transformations on the special linear group for visual servo control. *IEEE International Conference on Robotics and Automation (ICRA), Kobe, Japan*, pages 1498–1503, 2009.
- [13] F. Caballero et al. Homography based kalman filter for mosaic building. applications to uav position estimation. *IEEE international conference on robotics and automation (ICRA), Roma, Italy*, pages 2004–2009, 2007.
- [14] H. Wang et al. Visual odometry based on locally planar ground assumption. *IEEE International Conference on Information Acquisition Hong Kong, China*, pages 59–64, 2005.
- [15] O. Saurer et al. Homography based egomotion estimation with a common direction. *IEEE Trans. Pattern Anal. Mach. Intell.*, 39(2):327–341, 2017.
- [16] O. D. Faugeras and F. Lustman. Motion and structure from motion in a piecewise planar environment. *International Journal of Pattern Recognition and Artificial Intelligence*, 2(03):485–508, 1988.
- [17] V. Grabe, H.H. Bühlhoff, D. Scaramuzza, and P.R. Giordano. Nonlinear ego-motion estimation from optical flow for online control of a quadrotor uav. *The International Journal of Robotics Research*, 34(8):1114–1135, 2015.
- [18] V. Grabe, H.H. Bühlhoff, and P.R. Giordano. On-board velocity estimation and closed-loop control of a quadrotor uav based on optical flow. pages 491–497, 2012.
- [19] H.F. Grip, T.I. Fossen, T.A. Johansen, and A. Saberi. A nonlinear observer for integration of GNSS and IMU measurements with gyro bias estimation. In *American Cont. Conf.*, pages 4607–4612, 2012.

- [20] B.J. Guerreiro, C. Silvestre, R. Cunha, and D. Cabecinhas. Lidar-based control of autonomous rotorcraft for the inspection of pierlike structures. *IEEE Transactions on Control Systems Technology*, 2017.
- [21] T. Hamel, R. Mahony, J. Trumpf, P. Morin, and M.-D. Hua. Homography estimation on the special linear group based on direct point correspondence. pages 7902–7908, 2011.
- [22] T Hamel and C. Samson. Position estimation from direction or range measurements. *Automatica*, 82:137–144, 2017.
- [23] T. Hamel and C. Samson. Riccati observers for the non-stationary PnP problem. *IEEE Trans. on Automatic Control (in press)*, 2017.
- [24] T. Hamel and C. Samson. Riccati observers for the nonstationary PnP problem. volume 63, pages 726–741, 2017.
- [25] R. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge Univ. Press, 2003.
- [26] M.-D. Hua. Attitude estimation for accelerated vehicles using GPS/INS measurements. *Control Eng. Pract.*, 18(7):723–732, 2010.
- [27] M.-D. Hua and G. Allibert. Riccati observer design for pose, linear velocity and gravity direction estimation using landmark position and IMU measurements. In *IEEE Conference on Control Technology and Applications (CCTA)*, pages 1313–1318, 2018.
- [28] M.-D. Hua, T. Hamel, R. Mahony, and G. Allibert. Explicit complementary observer design on special linear group  $sl(3)$  for homography estimation using conic correspondences. pages 2434–2441, 2017.
- [29] M.-D. Hua, T. Hamel, and C. Samson. Riccati nonlinear observer for velocity-aided attitude estimation of accelerated vehicles using coupled velocity measurements. pages 2428–2433, 12 2017.
- [30] M.-D. Hua, N. Manerikar, T. Hamel, and C. Samson. Attitude, linear velocity and depth estimation of a camera observing a planar target using continuous homography and inertial data. pages 1429–1435, 05 2018.

- [31] M.-D. Hua, P. Martin, and T. Hamel. Stability analysis of velocity-aided attitude observers for accelerated vehicles. *Automatica*, 63:11–15, 2016.
- [32] M.-D. Hua, J. Trumpf, T. Hamel, R. Mahony, and P. Morin. An invitation to 3-d vision: From images to geometric models. *Springer Verlag*, 2003.
- [33] M.-D. Hua, J. Trumpf, T. Hamel, R. Mahony, and P. Morin. Point and line feature-based observer design on  $sl(3)$  for homography estimation and its application to image stabilization. Technical report, Technical Report hal-01628175, I3S, 2017.
- [34] M.-D. Hua, J. Trumpf, T. Hamel, R. Mahony, and P. Morin. Feature-based recursive observer design for homography estimation and its application to image stabilization. *Asian Journal of Control*, pages 1–16, 2019.
- [35] M.-D. Hua, J. Trumpf, T. Hamel, R. Mahony, and P. Morin. Nonlinear observer design on  $sl(3)$  for homography estimation by exploiting point and line correspondences with application to image stabilization. *Automatica*, 2019.
- [36] M.-D. Hua, M. Zamani, J. Trumpf, R. Mahony, and T. Hamel. Observer design on the Special Euclidean group  $SE(3)$ . In *IEEE Conference on Decision and Control*, pages 8169–8175, 2011.
- [37] C.V. Jawahar and P.K. Jain. Homography estimation from planar contours. pages 877–884, 2006.
- [38] S.J. Julier and J.K. Uhlmann. New extension of the kalman filter to nonlinear systems. 3068:182–193, 1997.
- [39] R. Kalman. On the general theory of control systems. *IRE Transactions on Automatic Control*, 4(3):110–110, 1959.
- [40] J.Y. Kaminski and A. Shashua. Multiple view geometry of general algebraic curves. *International Journal of Computer Vision*, 56(3):195–219, 2004.
- [41] A. Khosravian, J. Trumpf, R. Mahony, and T. Hamel. State estimation for invariant systems on lie groups with delayed output measurements. *Automatica*, 68:254–265, 2016.
- [42] D.P. Koch, R.W. Wheeler, D.O. amd Beard, T.W. McLain, and K.M. Brink. Rela-

- tive multiplicative extended kalman filter for observable gps-denied navigation. *39(9):1085–1121*, 2020.
- [43] S. K. Korovin and V. V. Fomichev. *State observers for linear systems with uncertainty*. 2009.
- [44] R.C. Leishman, J.C. Macdonald, R.W. Beard, and T.W. McLain. Quadrotors and accelerometers: State estimation with an improved dynamic model. *IEEE Control Systems*, 34(1):28–41, 2014.
- [45] H. Lou and P. Si. The distinguishability of linear control systems. *Nonlinear Analysis: Hybrid Systems*, 3(1):21–38, 2009.
- [46] G. LoweDavid. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 2004.
- [47] D.-G. Luenberger. Observing the state of a linear system. *IEEE Transactions on Military Electronics*, 8(2):74–80, 1964.
- [48] Y. Ma, S. Soatto, J. Kosecka, and S.S. Shankar. *An invitation to 3-D vision: from images to geometric models*, volume 26. Springer Science & Business Media, 2012.
- [49] R. Mahony, T. Hamel, P. Morin, and E. Malis. Nonlinear complementary filters on the special linear group. *International Journal of Control*, 85:1–17, 10 2012.
- [50] R. Mahony, T. Hamel, P. Morin, and E. Malis. Nonlinear complementary filters on the special linear group. *International Journal of Control*, 85(10):1557–1573, 2012.
- [51] R. Mahony, J. Trumpf, and T. Hamel. Observers for kinematic systems with symmetry. pages 617–633, 2013.
- [52] E. Malis, F. Chaumette, and S. Boudet. 2 1/2 D visual servoing. *IEEE Transactions on Robotics and Automation*, 15(2):238–250, 1999.
- [53] E. Malis and M. Vargas. Deeper understanding of the homography decomposition for vision-based control. *INRIA Research report*, RR-6303:90, 2007.
- [54] Hua M.-D. Hamel T Manerikar, N. and S. De. Marco. Riccati observer design for homography decomposition. In *European Control Conference (ECC)*, pages 1306–

1311, 2020.

- [55] N. Manerikar, M.-D. Hua, and T. Hamel. Homography observer design on special linear group  $SL(3)$  with application to optical flow estimation. In *European Control Conference (ECC)*, pages 1–5, 2018.
- [56] P. Martin and E. Salaün. An invariant observer for Earth-Velocity-Aided attitude heading reference systems. In *IFAC World Congr.*, pages 9857–9864, 2008.
- [57] P. Martin and I. Sarras. A semi-global model-based state observer for the quadrotor using only inertial measurements. In *IEEE Conference on Decision and Control (CDC)*, pages 7123–7128, 2016.
- [58] P. Morin, A. Eudes, and G. Scandaroli. Uniform observability of linear time-varying systems and application to robotics problems. In *International Conference on Geometric Science of Information*, pages 336–344, 2017.
- [59] F. Mufti, R. Mahony, and J. Kim. Super resolution of speed signs in video sequences. *Digital image computing: Techniques and applications (DICTA)*, Glenelg, South Australia, Australia, pages 278–285, 2007.
- [60] L. Nguyen, M.-D. Hua, G. Allibert, and T. Hamel. A homography-based dynamic control approach applied to station keeping of autonomous underwater vehicles without linear velocity measurement. *IEEE Transactions on Control Systems Technology, Institute of Electrical and Electronics Engineers*, 29(5):2065–2078, 2020.
- [61] L-H Nguyen, M-D Hua, G Allibert, and T. Hamel. A homography-based dynamic control approach applied to station keeping of autonomous underwater vehicles without linear velocity measurements. *IEEE Transactions on Control Systems Technology*, PP:1–14, 10 2020.
- [62] J. Nikolic, J. Rehder, M. Burri, P. Gohl, S. Leutenegger, P.T. Furgale, and R. Siegwart. A synchronized visual-inertial sensor system with fpga pre-processing for accurate real-time slam. pages 431–437, 05 2014.
- [63] J. O’Reilly. *Observers for linear systems*. 1983.
- [64] H. Plinval, P. Morin, and P. Mouyon. Stabilization of a class of underactuated vehicles with uncertain position measurements and application to visual servoing.

*Automatica*, 77:155–169, 2017.

- [65] S.J.D. Prince. *Computer Vision: Models Learning and Inference*. Cambridge Univ. Press, 2012.
- [66] A. Roberts and A. Tayebi. On the attitude estimation of accelerating rigid-bodies using GPS and IMU measurements. In *IEEE Conf. on Dec. and Cont.*, pages 8088–8093, 2011.
- [67] Edward Rosten and T. Drummond. Fusing points and lines for high performance tracking. volume 2, pages 1508–1515, 11 2005.
- [68] Ethan Rublee, Vincent Rabaud, Kurt Konolige, and Gary Bradski. Orb: an efficient alternative to sift or surf. pages 2564–2571, 11 2011.
- [69] J. Ruiz, F. Caballero, and L. Merino. Mgraph: A multigraph homography method to generate incremental mosaics in real-time from uav swarms. *IEEE Robot. Autom. Lett.* 3, (4):2838–2845, 2018.
- [70] D. Scaramuzza, M.C. Achtelik, L. Doitsidis, F. Friedrich, and et al. Vision-controlled micro flying robots: from system design to autonomous navigation and mapping in gps-denied environments. *IEEE Robotics & Automation Magazine*, 21(3):26–40, 2014.
- [71] D. Scaramuzza and R. Siegwart. Appearance-guided monocular omnidirectional visual odometry for outdoor ground vehicles. *IEEE Trans. Robot.*, 24(5):1015–1026, 2008.
- [72] T. Tang, K. Sreenath, and V. Kumar. Aggressive maneuvering of a quadrotor with a cable-suspended payload. In *Robotics: Science and Systems, Workshop on Women in Robotics*, 2014.
- [73] J. Thomas, M. Pope, G. Loianno, E.L. Elliot, M.A. Estrada, H. Jiang, M.R. Cutkosky, and V. Kumar. Aggressive flight with quadrotors for perching on inclined surfaces. *Journal of Mechanisms and Robotics*, 8(5):051007, 2016.
- [74] H. Trinh and T. Fernando. On the existence and design of functional observers for linear systems. pages 1974–1979, 2007.

- [75] H. Trinh and T. Fernando. *Functional observers for dynamical systems*. 2012.
- [76] S. Weiss, R. Brockers, and L. Matthies. 4DoF drift free navigation using inertial cues and optical flow. In *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 4180–4186, 2013.
- [77] Z. Zhang and A. R. Hanson. 3D reconstruction based on homography mapping. *Proc. ARPA96*, pages 1007–1012, 1996.