



HAL
open science

Relative pose estimation from bearing measurements of three unknown source points

Minh-Duc Hua, Simone de Marco, Tarek Hamel, Randal W Beard

► **To cite this version:**

Minh-Duc Hua, Simone de Marco, Tarek Hamel, Randal W Beard. Relative pose estimation from bearing measurements of three unknown source points. IEEE Conference on Decision and Control (CDC), 2020, Jeju Island (virtual), South Korea. hal-03052573

HAL Id: hal-03052573

<https://hal.science/hal-03052573>

Submitted on 10 Dec 2020

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.

Relative pose estimation from bearing measurements of three unknown source points

Minh-Duc Hua, Simone De Marco, Tarek Hamel and Randal W. Beard

Abstract—This paper unveils a novel discovery that the full relative pose of a monocular camera moving in a three dimensional space can be estimated exploiting bearing measurements of only 3 unknown source points (together with velocity measurements) without any additional knowledge if the camera translational motion is sufficiently exciting. The epipolar constraint commonly used in Computer Vision algebraic algorithms for the determination of the so-called essential matrix (all of them require at least 5 source points) is here exploited in the design of the proposed Riccati observer for pose estimation. One remarkable feature of this work is the determination of an explicit persistence of excitation condition that guarantees uniform observability and, subsequently, (local) exponential stability of the proposed observer. Convincing simulation results are provided to support the proposed approach.

I. INTRODUCTION

Estimating camera motion from a video sequence has many applications in robotics including target tracking, visual odometry, and 3D scene reconstruction. One method to estimate motion from a video sequence is to calculate the essential matrix between consecutive frames. The essential matrix relates the homogeneous image coordinates between frames using the epipolar constraint. After the essential matrix has been determined, it can be decomposed into a rotation and a normalized translation to determine the relative motion of the camera between frames. In order to be robust to noise and feature mismatches, the essential matrix is typically estimated by generating a large number of hypotheses from five-point minimum subsets of matching features, and selecting the best hypothesis using either Random Sample Consensus (RANSAC) [4] or Least Median of Squares (LMedS) [10].

State of the art methods calculate essential matrix hypotheses directly from each five-point minimum subset. One of the best known methods is Nister’s algorithm [9]. Nister showed that for five matching points, there are potentially ten essential matrices that satisfy the constraints, each corresponding to a real root of a tenth-order polynomial generated from the data. There are many open-source implementations of Nister’s five-point algorithm including OpenCV’s `findEssentialMat` function [1]. However, constructing, solving, and extracting the essential matrix from this tenth-order polynomial is complex and can be computationally expensive. Furthermore, since each minimum subset produces up to ten hypotheses, it can be time consuming to score them.

As an alternative to directly calculating essential matrix solutions, some authors [2], [3], [7], [8], [11] propose solving for the essential matrix using nonlinear optimization algorithms such as Gauss-Newton (GN) and Levenberg-Marquardt (LM). Since the essential matrix has nine entries but only five degrees of freedom, the optimization is performed on the five dimensional essential matrix manifold. There are a number of ways to define the essential matrix manifold. Some authors define the manifold using a rotation and a translation unit vector, which are elements of $SO(3)$ and S^2 respectively [3], [8]. Others define the manifold using two elements of $SO(3)$ [7], [11]. The computational requirements of the resulting scheme are significantly less than Nister’s five point algorithm. However, one weakness of optimization-based solvers is that they only find one of the ten possible essential matrices at a time. Finding all solutions requires additional optimization runs with different initialization points. The optimization method is also sensitive to initial conditions, which can cause the optimizer to fail to produce a valid solution.

After the essential matrix between images is found, it must then be decomposed into a rotation and a normalized translation. Given an essential matrix, there are four possible rotation-translation pairs [6]. The correct rotation-translation pair is typically determined using the Cheirality check that ensures that matching features are in front of both cameras. However, the Cheirality check is sensitive to noise in the image and frequently returns the wrong decomposition.

In this paper we propose a novel observer to estimate the 3D relative pose using the recently introduced Riccati observer design framework [5]. We show that the proposed observer is locally exponentially stable if the motion of the camera is sufficiently exciting. The key contribution consists in using the relative pose of the camera as system state and the epipolar constraints (involving the essential matrix) of three unknown source points. This approach is different from existing works that estimate first the essential matrix and then decompose it to obtain the relative orientation and only the relative normalized translation.

The remainder of the paper is organized as follows. The problem is formally stated in Section II. The Riccati observer and its theoretical properties are reviewed in Section III. The application of the observer to relative pose estimation and the associated stability analysis are provided in Section IV. Simulation results are presented in Section V and the conclusions are in Section VI.

M.-D. Hua, S. De Marco, T. Hamel are with the laboratory I3S, CNRS, Université Côte d’Azur, Sophia Antipolis, France. T. Hamel is also with the Institute Universitaire de France (IUF). Emails: [hua\(sdemarco,thamel\)@i3s.unice.fr](mailto:hua(sdemarco,thamel)@i3s.unice.fr).

R. W. Beard is with Brigham Young University. Email: beard@byu.edu.

II. PROBLEM STATEMENT

A. Mathematical notation

- $\{e_1, e_2, e_3\}$ denotes the canonical basis of \mathbb{R}^3 and the identity matrix and the null matrix of $\mathbb{R}^{n \times n}$ are denoted as I_n and 0_n , respectively. The closed ball in \mathbb{R}^n of radius r is denoted as \mathcal{B}_r^n . Let $(\cdot)_\times$ denote the skew-symmetric matrix associated with the cross product, i.e. $u_\times v = u \times v, \forall u, v \in \mathbb{R}^3$. Let $\pi_u \triangleq I_3 - uu^\top, \forall u \in S^2$, denote the projection operator onto the plane orthogonal to u .
- We write $f(x, y, t) = O(|x|^{k_1}|y|^{k_2})$ with $k_1 \geq 0$ and $k_2 \geq 0$ if $\forall t : |f(x, y, t)|/(|x|^{k_1}|y|^{k_2}) \leq \gamma < \infty$ in the neighbourhood of $(x = 0, y = 0)$. If f depends only on x and t then we write $f(x, t) = O(|x|^k)$ if $\forall t : |f(x, t)|/|x|^k \leq \gamma < \infty$ in the neighbourhood of $x = 0$.

B. Problem formulation

Consider a robotic vehicle equipped with a monocular camera observing n ($n \geq 3$) source points with unknown 3D coordinates. Without loss of generality and to make the presentation clear we focus only on the most difficult case of $n = 3$.

Let us now introduce some notation. Let $\{\mathcal{A}\}$ (resp. $\{\mathcal{B}\}$) denote the reference (resp. current or body) frame attached to the camera at the reference (resp. current) view. The orientation of frame $\{\mathcal{B}\}$ w.r.t. frame $\{\mathcal{A}\}$ is represented by a rotation matrix $R \in \text{SO}(3)$. Let $\xi \in \mathbb{R}^3$ (resp. $\bar{\xi} \in \mathbb{R}^3$) denote the position of frame $\{\mathcal{B}\}$ w.r.t. frame $\{\mathcal{A}\}$ expressed in the body frame $\{\mathcal{B}\}$ (resp. frame $\{\mathcal{A}\}$). One verifies that $\bar{\xi} = R\xi$. The dynamics of the camera pose (R, ξ) are given by

$$\begin{cases} \dot{R} = R\Omega_\times \\ \dot{\xi} = -\Omega_\times \xi + V \end{cases} \quad (1)$$

where $V \in \mathbb{R}^3$ and $\Omega \in \mathbb{R}^3$ denote the camera's linear and angular velocities expressed in the body frame $\{\mathcal{B}\}$.

Assume that the vehicle is equipped with a linear velocity sensor (e.g., a Doppler sensor) that measures the linear velocity V together with an Inertial Measurement Unit (IMU) that measures of the angular velocity Ω .

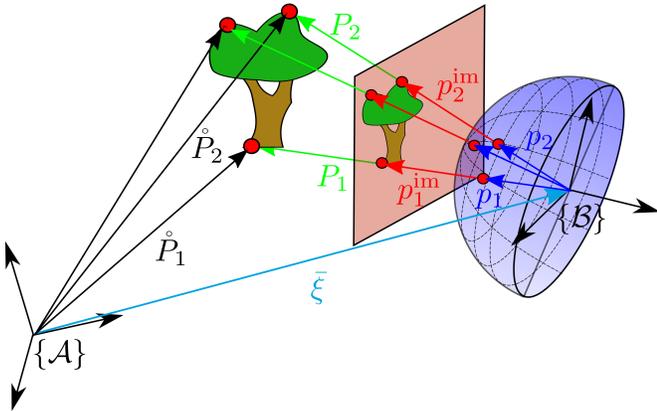


Fig. 1. Intuitive representation of 3D coordinates \bar{P}_i and P_i ($i = 1, 2, 3$), of planar projective coordinates \bar{p}_i^{im} and p_i^{im} , and of spherical projective coordinates \bar{p}_i and p_i of the 3 source points.

We also consider the situation where a collection of 3 unknown source points is always observed by the camera so that their bearings can be directly obtained from the camera images. Let \bar{p}_i^{im} ($i = 1, 2, 3$) (resp. p_i^{im}) denote the calibrated projective coordinates of the 3 source points onto the camera plane expressed w.r.t. the frame $\{\mathcal{A}\}$ (resp. frame $\{\mathcal{B}\}$) (see Fig. 1). Denoting $\bar{P}_i \in \mathbb{R}^3$ (resp. $P_i \in \mathbb{R}^3$) the 3D coordinates of the i -th source point w.r.t. the frame $\{\mathcal{A}\}$ (resp. frame $\{\mathcal{B}\}$) expressed in $\{\mathcal{A}\}$ (resp. frame $\{\mathcal{B}\}$), and by \bar{z}_i (resp. z_i) its third component, one verifies that $\bar{p}_i^{\text{im}} = \bar{P}_i/\bar{z}_i$ (resp. $p_i^{\text{im}} = P_i/z_i$).

Instead of using the perspective outputs typically used in computer vision algorithms, we use bearing outputs

$$\bar{p}_i := \frac{\bar{p}_i^{\text{im}}}{|\bar{p}_i^{\text{im}}|} = \frac{\bar{P}_i}{|\bar{P}_i|} \in S^2, \quad p_i := \frac{p_i^{\text{im}}}{|p_i^{\text{im}}|} = \frac{P_i}{|P_i|} \in S^2 \quad (2)$$

that correspond to the projection onto a virtual unit spherical image plane and differ from the perspective outputs only by the scaling. Using the relations $P_i = R^\top \bar{P}_i - \xi = R^\top (\bar{P}_i - \bar{\xi})$, the following epipolar constraint can be deduced

$$\bar{p}_i^\top R \xi_\times p_i = 0, \quad (i = 1, 2, 3) \quad (3)$$

which, using the essential matrix definition $E := R t_\times$ with $t := \xi/|\xi|$, can be expressed as $\bar{p}_i^\top E p_i = 0$. Instead of estimating the essential matrix E from the epipolar constraints and then decomposing it into a rotation R and a normalized translation t like in traditional algebraic approaches, we will estimate directly the pose (i.e. R and ξ) by also exploiting the dynamic equations (1). Note that in contrast with the ill-definition of the essential matrix when the translation vector ξ vanished (i.e. $\xi = 0$), the pose is always well defined.

III. THEORETICAL BACKGROUND: BRIEF RECALL OF A RICCATI OBSERVER DESIGN FRAMEWORK

The observer proposed in this paper is based on the Riccati design framework recently developed in [5]. The following nonlinear system (a particular case of systems studied in [5]) is investigated:

$$\begin{cases} \dot{X} = A(t)X + U + O(|X|^2) + O(|X||U|) \\ Y = C(X, t)X + O(|X|^2) \end{cases} \quad (4)$$

with state $X = [X_1^\top, X_2^\top]^\top$, $X_1 \in \mathcal{B}_r^n$, $X_2 \in \mathbb{R}^n$, output $Y \in \mathbb{R}^m$, $A(t) \in \mathbb{R}^{2n \times 2n}$ a continuous matrix-valued function uniformly bounded w.r.t. t in the form

$$A(t) = \begin{bmatrix} A_{1,1}(t) & 0_n \\ A_{2,1}(t) & A_{2,2}(t) \end{bmatrix} \quad (5)$$

and $C(X, t) \in \mathbb{R}^{m \times 2n}$ a continuous matrix-valued function uniformly continuous w.r.t. X and uniformly bounded w.r.t. t . Then, apply the input

$$U = -PC^\top D(t)Y \quad (6)$$

with $P \in \mathbb{R}^{2n \times 2n}$ a symmetric positive definite matrix solution to the following continuous Riccati equation (CRE):

$$\dot{P} = AP + PA^\top - PC^\top D(t)CP + S(t) \quad (7)$$

with $P(0) \in \mathbb{R}^{2n \times 2n}$ a symmetric positive definite matrix, $D(t) \in \mathbb{R}^{m \times m}$ bounded continuous symmetric positive

semi-definite, and $S(t) \in \mathbb{R}^{2n \times 2n}$ bounded continuous symmetric positive definite.

Then, from Theorem 3.1 and Corollary 3.2 in [5], $X = 0$ is locally exponentially stable when both matrices $D(t)$ and $S(t)$ are larger than some constant positive matrix and the pair $(A^*(t), C^*(t))$, with $A^*(t) \triangleq A(t)$, $C^*(t) \triangleq C(0, t)$, is uniformly observable.

Definition 1 (uniform observability) *The pair $(A(t), C(t))$ is called uniformly observable if there exist $\delta, \mu > 0$ such that $\forall t \geq 0$*

$$W(t, t + \delta) \triangleq \frac{1}{\delta} \int_t^{t+\delta} \Phi^\top(s, t) C^\top(s) C(s) \Phi(s, t) ds \geq \mu I_{2n}$$

with $\Phi(s, t)$ the transition matrix associated with A , i.e. $\frac{d}{ds} \Phi(s, t) = A(s) \Phi(s, t)$ with $\Phi(t, t) = I_{2n}$.

The matrix $W(t, t + \delta)$ is the so-called observability Gramian.

IV. POSE OBSERVER DESIGN FROM BEARING MEASUREMENTS OF THREE UNKNOWN SOURCE POINTS

A. Observer derivation

The proposed observer has the following form

$$\begin{cases} \dot{\hat{R}} = \hat{R} \Omega_\times - \hat{R} \sigma_{R \times} \\ \dot{\hat{\xi}} = -\Omega_\times \hat{\xi} + V - \sigma_\xi \end{cases} \quad (9)$$

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

The following error variables are defined

$$\tilde{R} := \hat{R}^\top R, \quad \tilde{\xi} := \xi - \hat{\xi} \quad (10)$$

Then, the objective of observer design consists in stabilizing $(\tilde{R}, \tilde{\xi})$ about $(I_3, 0)$.

From (1), (9) and (10) one verifies that the error system is given by

$$\begin{cases} \dot{\tilde{R}} = -\Omega_\times \tilde{R} + \tilde{R} \Omega_\times + \sigma_{R \times} \tilde{R} \\ \dot{\tilde{\xi}} = -\Omega_\times \tilde{\xi} + \sigma_\xi \end{cases} \quad (11)$$

For analysis purposes let us assume that ξ, Ω and V remain bounded for all time, which is a completely reasonable assumption for the considered applications.

The following step involves developing first order approximations of the error system (11) and of the measurement equations (2) and (3) in order to obtain the system in the form (4). From the Rodrigues' formula, the first order approximation of \tilde{R} is given by

$$\tilde{R} = I + \lambda_\times + O(|\lambda|^2) \quad (12)$$

with $\lambda \in \mathcal{B}_2^3$ equal to twice the vector part of the quaternion associated with the attitude error matrix \tilde{R} . One then deduces from the first equation of (11), (12) and the identity

$$a_\times b_\times - b_\times a_\times = (a \times b)_\times, \quad \forall a, b \in \mathbb{R}^3$$

that in first order approximations

$$\dot{\lambda} = -\Omega_\times \lambda + \sigma_R + O(|\lambda|^2) + O(|\lambda| |\sigma_R|) \quad (13)$$

Now let us develop first order approximations of the measurement equations (2)–(3). From the epipolar constraint (3) one deduces

$$\begin{aligned} 0 &= \hat{p}_i^\top \hat{R} \tilde{R} (\tilde{\xi} + \hat{\xi})_\times p_i \\ &= \hat{p}_i^\top \hat{R} (I + \lambda_\times) (\tilde{\xi} + \hat{\xi})_\times p_i + O(|\lambda|^2) \\ &= -\hat{p}_i^\top \hat{R} (\hat{\xi} \times p_i)_\times \lambda - \hat{p}_i^\top \hat{R} p_i \times \tilde{\xi} + \hat{p}_i^\top \hat{R} \hat{\xi} \times p_i \\ &\quad + O(|\lambda|^2) + O(|\lambda| |\tilde{\xi}|) \\ &= \hat{p}_i^\top \hat{R} \hat{\xi} \times p_i - [\hat{p}_i^\top \hat{R} (\hat{\xi} \times p_i)_\times \quad \hat{p}_i^\top \hat{R} p_i \times] \begin{bmatrix} \lambda \\ \tilde{\xi} \end{bmatrix} \\ &\quad + O(|\lambda|^2) + O(|\lambda| |\tilde{\xi}|) \end{aligned}$$

or equivalently

$$\begin{aligned} \hat{p}_i^\top \hat{R} \hat{\xi} \times p_i &= [\hat{p}_i^\top \hat{R} (\hat{\xi} \times p_i)_\times \quad \hat{p}_i^\top \hat{R} p_i \times] \begin{bmatrix} \lambda \\ \tilde{\xi} \end{bmatrix} \\ &\quad + O(|\lambda|^2) + O(|\lambda| |\tilde{\xi}|) \end{aligned} \quad (14)$$

From (13), the second equation of (11), and (14), one obtains in first order approximations the system in compact form (4) with

$$\begin{cases} X := \begin{bmatrix} \lambda \\ \tilde{\xi} \end{bmatrix}, \quad U := \begin{bmatrix} \sigma_R \\ \sigma_\xi \end{bmatrix}, \quad Y := \begin{bmatrix} \hat{p}_1^\top \hat{R} \hat{\xi} \times p_1 \\ \hat{p}_2^\top \hat{R} \hat{\xi} \times p_2 \\ \hat{p}_3^\top \hat{R} \hat{\xi} \times p_3 \end{bmatrix} \\ A(t) := \begin{bmatrix} -\Omega_\times & 0_3 \\ 0_3 & -\Omega_\times \end{bmatrix} \\ C(X, t) = \begin{bmatrix} \hat{p}_1^\top \hat{R} (\hat{\xi} \times p_1)_\times & \hat{p}_1^\top \hat{R} p_1 \times \\ \hat{p}_2^\top \hat{R} (\hat{\xi} \times p_2)_\times & \hat{p}_2^\top \hat{R} p_2 \times \\ \hat{p}_3^\top \hat{R} (\hat{\xi} \times p_3)_\times & \hat{p}_3^\top \hat{R} p_3 \times \end{bmatrix} \end{cases} \quad (15)$$

The fact that $X = [X_1^\top, X_2^\top]^\top$ with $X_1 := \lambda \in \mathcal{B}_2^3$, $X_2 := \tilde{\xi} \in \mathbb{R}^3$, together with the particular form of the matrix A as (5), allows one to obtain the expression of the innovation terms from the input U calculated according to (6) and (7) where the matrices D and S , involved in the CRE (7), are chosen larger than some constant positive matrix.

B. Observability and stability analysis

According to [5] 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 (15), and using $\tilde{\xi} = R\xi$ and $Rp_i = \hat{p}_i - \tilde{\xi}$ with $i = 1, 2, 3$, one obtains

$$\begin{aligned} C^* &= \begin{bmatrix} \hat{p}_1^\top (\tilde{\xi} \times Rp_1)_\times R & \hat{p}_1^\top (Rp_1)_\times R \\ \hat{p}_2^\top (\tilde{\xi} \times Rp_2)_\times R & \hat{p}_2^\top (Rp_2)_\times R \\ \hat{p}_3^\top (\tilde{\xi} \times Rp_3)_\times R & \hat{p}_3^\top (Rp_3)_\times R \end{bmatrix} \\ &= \begin{bmatrix} \frac{1}{|\hat{P}_1| |\hat{P}_1|} \hat{P}_1^\top (\tilde{\xi} \times \hat{P}_1)_\times R & -\frac{1}{|\hat{P}_1| |\hat{P}_1|} \hat{P}_1^\top \tilde{\xi} \times R \\ \frac{1}{|\hat{P}_2| |\hat{P}_2|} \hat{P}_2^\top (\tilde{\xi} \times \hat{P}_2)_\times R & -\frac{1}{|\hat{P}_2| |\hat{P}_2|} \hat{P}_2^\top \tilde{\xi} \times R \\ \frac{1}{|\hat{P}_3| |\hat{P}_3|} \hat{P}_3^\top (\tilde{\xi} \times \hat{P}_3)_\times R & -\frac{1}{|\hat{P}_3| |\hat{P}_3|} \hat{P}_3^\top \tilde{\xi} \times R \end{bmatrix} \\ &= \begin{bmatrix} -|\hat{P}_1| \frac{\tilde{\xi}^\top}{|\hat{P}_1|} \pi_{\hat{p}_1} R & \frac{\tilde{\xi}^\top}{|\hat{P}_1|} \hat{p}_1 \times R \\ -|\hat{P}_2| \frac{\tilde{\xi}^\top}{|\hat{P}_2|} \pi_{\hat{p}_2} R & \frac{\tilde{\xi}^\top}{|\hat{P}_2|} \hat{p}_2 \times R \\ -|\hat{P}_3| \frac{\tilde{\xi}^\top}{|\hat{P}_3|} \pi_{\hat{p}_3} R & \frac{\tilde{\xi}^\top}{|\hat{P}_3|} \hat{p}_3 \times R \end{bmatrix} \end{aligned} \quad (16)$$

The following theorem settles a persistence of excitation condition for the uniform observability of the matrix pair (A^*, C^*) and, subsequently, the local exponential stability of the error dynamics.

Theorem 1 *Assume that the bearings $\hat{p}_1, \hat{p}_2, \hat{p}_3$ of the 3 observed source points are linearly independent. Assume that the camera translational motion is sufficiently exciting in the sense that for all time there exist $\delta, \beta > 0$ such that $\forall i = 1, 2, 3$*

$$\Pi_i(t, t + \delta) := \frac{1}{\delta} \int_t^{t+\delta} \frac{\bar{\xi}(s)\bar{\xi}^\top(s)}{|\bar{P}_i - \bar{\xi}(s)|^2} ds \geq \beta I_3 \quad (17)$$

Assume also that ξ, Ω and V remain uniformly bounded for all time. Then, the pair (A^*, C^*) is uniformly observable. By choosing the matrices D and S involved in the CRE (7) larger than some constant positive matrix, one ensures that the equilibrium $(\bar{R}, \bar{\xi}) = (I_3, 0)$ of the error system is locally exponentially stable.

Proof: To prove that the pair $(A^*(t), C^*(t))$ is uniformly observable, let us compute the observability Gramian $W(t, t + \delta)$ as defined by (8). First, in view of the expression of A in (15) one verifies that the transition matrix associated with the state matrix A^* (i.e. A) is of the form

$$\Phi(s, t) = \begin{bmatrix} \bar{R}^\top(s) & 0_3 \\ 0_3 & \bar{R}^\top(s) \end{bmatrix} \quad (18)$$

with $\bar{R}(s) \in SO(3)$ the solution to

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

One deduces from (16) that

$$C^{*\top} C^* = \begin{bmatrix} R^\top & 0_3 \\ 0_3 & R^\top \end{bmatrix} M \begin{bmatrix} R & 0_3 \\ 0_3 & R \end{bmatrix} \quad (19)$$

with

$$M := \begin{bmatrix} \sum_{i=1,2,3} |\dot{P}_i|^2 \pi_{\hat{p}_i} \frac{\bar{\xi}\bar{\xi}^\top}{|\bar{P}_i|^2} \pi_{\hat{p}_i} & \sum_{i=1,2,3} |\dot{P}_i| \pi_{\hat{p}_i} \frac{\bar{\xi}\bar{\xi}^\top}{|\bar{P}_i|^2} \hat{p}_{i \times} \\ - \sum_{i=1,2,3} |\dot{P}_i| \hat{p}_{i \times} \frac{\bar{\xi}\bar{\xi}^\top}{|\bar{P}_i|^2} \pi_{\hat{p}_i} & - \sum_{i=1,2,3} \hat{p}_{i \times} \frac{\bar{\xi}\bar{\xi}^\top}{|\bar{P}_i|^2} \hat{p}_{i \times} \end{bmatrix}$$

Using (18), (19) and the fact that $\bar{R}^\top(s) = R^\top(s)R(t)$ one then deduces the observability Gramian

$$\begin{aligned} W(t, t + \delta) &= \frac{1}{\delta} \int_t^{t+\delta} \Phi^\top(s, t) C^{*\top}(s) C^*(s) \Phi(s, t) ds \\ &= \begin{bmatrix} R^\top(t) & 0_3 \\ 0_3 & R^\top(t) \end{bmatrix} \left(\frac{1}{\delta} \int_t^{t+\delta} M(s) ds \right) \begin{bmatrix} R(t) & 0_3 \\ 0_3 & R(t) \end{bmatrix} \end{aligned} \quad (20)$$

with

$$\begin{aligned} &\frac{1}{\delta} \int_t^{t+\delta} M(s) ds \\ &= \begin{bmatrix} \sum_{i=1,2,3} |\dot{P}_i|^2 \pi_{\hat{p}_i} \Pi_i(t, t + \delta) \pi_{\hat{p}_i} & \sum_{i=1,2,3} |\dot{P}_i| \pi_{\hat{p}_i} \Pi_i(t, t + \delta) \hat{p}_{i \times} \\ - \sum_{i=1,2,3} |\dot{P}_i| \hat{p}_{i \times} \Pi_i(t, t + \delta) \pi_{\hat{p}_i} & - \sum_{i=1,2,3} \hat{p}_{i \times} \Pi_i(t, t + \delta) \hat{p}_{i \times} \end{bmatrix} \end{aligned} \quad (21)$$

and $\Pi_i(t, t + \delta)$ defined in (17).

From now on, the shortened notation Π_i is used in the place of $\Pi_i(t, t + \delta)$ for convenience. Since Π_i are symmetric and positive definite by assumption (c.f. Eq. (17)), they can be decomposed as $\Pi_i = Q_i^\top \Delta_i Q_i$, with $\Delta_i = \text{diag}([\beta_{i,1}, \beta_{i,2}, \beta_{i,3}])$ and $\beta_{i,j} \geq \beta$ ($\forall j = 1, 2, 3$) using (17), and some $Q_i \in SO(3)$. Thus,

$$\Pi_i - \beta I_3 = Q_i^\top (\Delta_i - \beta I_3) Q_i = \Gamma_i^\top \Gamma_i$$

with $\Gamma_i = (\Delta_i - \beta I_3)^{\frac{1}{2}} Q_i, \forall i = 1, 2, 3$. Thus, one obtains

$$\begin{aligned} &\begin{bmatrix} |\dot{P}_i|^2 \pi_{\hat{p}_i} (\Pi_i - \beta I_3) \pi_{\hat{p}_i} & |\dot{P}_i| \pi_{\hat{p}_i} (\Pi_i - \beta I_3) \hat{p}_{i \times} \\ - |\dot{P}_i| \hat{p}_{i \times} (\Pi_i - \beta I_3) \pi_{\hat{p}_i} & - \hat{p}_{i \times} (\Pi_i - \beta I_3) \hat{p}_{i \times} \end{bmatrix} \\ &= \begin{bmatrix} |\dot{P}_i|^2 \pi_{\hat{p}_i} \Gamma_i^\top \Gamma_i \pi_{\hat{p}_i} & |\dot{P}_i| \pi_{\hat{p}_i} \Gamma_i^\top \Gamma_i \hat{p}_{i \times} \\ - |\dot{P}_i| \hat{p}_{i \times} \Gamma_i^\top \Gamma_i \pi_{\hat{p}_i} & - \hat{p}_{i \times} \Gamma_i^\top \Gamma_i \hat{p}_{i \times} \end{bmatrix} \\ &= \begin{bmatrix} |\dot{P}_i| \pi_{\hat{p}_i} \Gamma_i^\top \\ - \hat{p}_{i \times} \Gamma_i^\top \end{bmatrix} \begin{bmatrix} |\dot{P}_i| \Gamma_i \pi_{\hat{p}_i} & \Gamma_i \hat{p}_{i \times} \end{bmatrix} \\ &\geq 0 \end{aligned} \quad (22)$$

Using (21) and (22) one deduces

$$\frac{1}{\delta} \int_t^{t+\delta} M(s) ds \geq \beta \dot{N} \quad (23)$$

with

$$\begin{aligned} \dot{N} &:= \begin{bmatrix} \sum_{i=1,2,3} |\dot{P}_i|^2 \pi_{\hat{p}_i} & \sum_{i=1,2,3} |\dot{P}_i| \hat{p}_{i \times} \\ - \sum_{i=1,2,3} |\dot{P}_i| \hat{p}_{i \times} & \sum_{i=1,2,3} \pi_{\hat{p}_i} \end{bmatrix} = \sum_{i=1,2,3} \dot{N}_i \\ \dot{N}_i &:= \begin{bmatrix} |\dot{P}_i|^2 \pi_{\hat{p}_i} & |\dot{P}_i| \hat{p}_{i \times} \\ - |\dot{P}_i| \hat{p}_{i \times} & \pi_{\hat{p}_i} \end{bmatrix} = \dot{Z}_i^\top \dot{Z}_i \\ \dot{Z}_i &:= [|\dot{P}_i| \pi_{\hat{p}_i} \quad \hat{p}_{i \times}] \end{aligned} \quad (24)$$

In view of (20) and (23), the uniform observability condition (8) involving the observability Gramian matrix $W(t, t + \delta)$ is satisfied if the constant symmetric matrix \dot{N} given by (24) is positive definite. Thus, one only needs to prove that the equation $\nu^\top \dot{N} \nu = 0$, with $\nu = [\nu_1, \nu_2]^\top \in \mathbb{R}^6$, implies that $\nu = 0$ (i.e. $\nu_1 = \nu_2 = 0$) is the unique solution. One has

$$\begin{aligned} 0 &= \nu^\top \dot{N} \nu = \sum_{i=1}^3 \nu^\top \dot{N}_i \nu = \sum_{i=1}^3 |\dot{Z}_i \nu|^2 \\ &\Rightarrow \dot{Z}_i \nu = 0, \quad \forall i = 1, 2, 3 \\ &\Rightarrow |\dot{P}_i| \pi_{\hat{p}_i} \nu_1 + \hat{p}_i \times \nu_2 = 0, \quad \forall i = 1, 2, 3 \\ &\Rightarrow \hat{p}_i \times \left(-|\dot{P}_i| \hat{p}_{i \times} \nu_1 + \nu_2 \right) = 0, \quad \forall i = 1, 2, 3 \\ &\Rightarrow \exists \beta_i \in \mathbb{R} \text{ s.t. } \nu_2 = \beta_i \dot{P}_i + \dot{P}_i \times \nu_1, \quad \forall i = 1, 2, 3 \\ &\Rightarrow \beta_i \dot{P}_i - \beta_j \dot{P}_j + (\dot{P}_i - \dot{P}_j) \times \nu_1 = 0, \quad \forall i \neq j, \quad i, j \in \{1, 2, 3\} \\ &\Rightarrow \nu_1^\top \left((\dot{P}_i - \dot{P}_j) \times (\dot{P}_i \times \dot{P}_j) \right) = 0 \\ &\Rightarrow \nu_1^\top \left(|\dot{P}_i| \pi_{\hat{p}_i} \hat{p}_j + |\dot{P}_j| \pi_{\hat{p}_j} \hat{p}_i \right) = 0 \end{aligned}$$

Since the triplet $(|\dot{P}_1| \pi_{\hat{p}_1} \hat{p}_2 + |\dot{P}_2| \pi_{\hat{p}_2} \hat{p}_1, |\dot{P}_2| \pi_{\hat{p}_2} \hat{p}_3 + |\dot{P}_3| \pi_{\hat{p}_3} \hat{p}_2, |\dot{P}_3| \pi_{\hat{p}_3} \hat{p}_1 + |\dot{P}_1| \pi_{\hat{p}_1} \hat{p}_3)$ is linearly independent (as a consequence of linear independence of the triplet

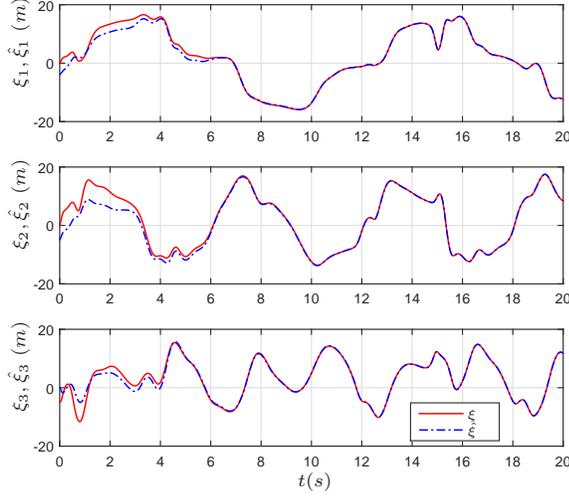


Fig. 2. (Simulation 1) Real and estimated positions (expressed in frame $\{B\}$) ξ , $\hat{\xi}$ (m) versus time (s).

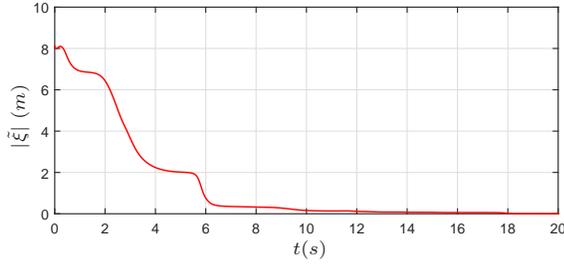


Fig. 3. (Simulation 1) Norm of the position error $|\tilde{\xi}|$ (m) versus time (s).

($\dot{p}_1, \dot{p}_2, \dot{p}_3$)), one deduces that $\nu_1 = 0$, which in turn implies that $\dot{p}_i \times \nu_2 = 0$, $\forall i = 1, 2, 3$. Using again the fact that the triplet ($\dot{p}_1, \dot{p}_2, \dot{p}_3$) is linearly independent, one deduces that $\nu_2 = 0$. The remainder of the proof then directly follows by application of Theorem 3.1 and Corollary 3.2 in [5]. ■

V. SIMULATION RESULTS

The performance of the proposed observer is demonstrated through simulation results performed with Matlab Simulink. In the simulated scenario, the camera's position $\bar{\xi}$ expressed in frame $\{A\}$ follows a periodical trajectory varying in all 3 components:

$$\bar{\xi}(t) = [15\sin(\pi t/6), 15\sin(\pi t/3), -5 + 2\sin(\pi t/2)]^\top$$

The camera's attitude is generated so that it is also strongly varying as shown in Figure 4. The (unknown) 3D coordinates expressed in frame $\{A\}$ of the 3 observed source points are given by $\dot{P}_1 = [2, 4, 2.5]^\top$, $\dot{P}_2 = [-4.5, 1, 1.5]^\top$, $\dot{P}_3 = [-1, -1.5, 0.6]^\top$.

The matrices S and D^{-1} involved in the CRE (7) 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. The following parameters are chosen: $P(0) = 0.1I_6$, $D = 100I_3$, $S = \text{diag}(0.1I_3, I_3)$.

Two simulations will be reported next, where the first one corresponds to the case of ideal measurements (i.e. no noise)

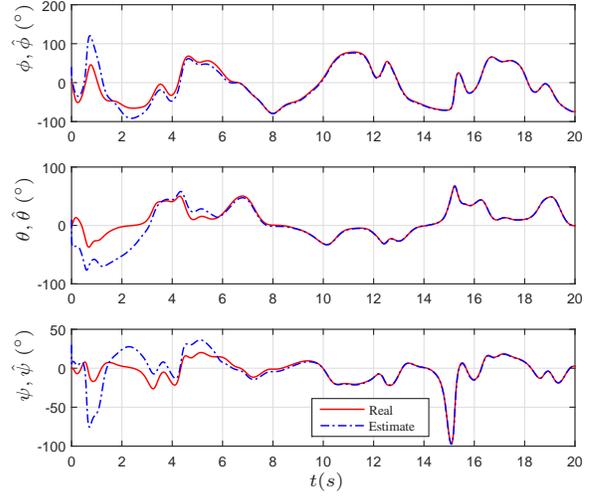


Fig. 4. (Simulation 1) Real and estimated attitudes represented by roll, pitch and yaw Euler angles (deg) versus time (s).

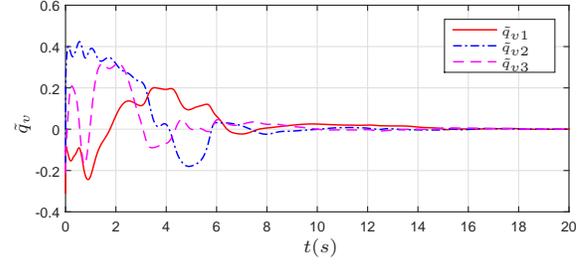


Fig. 5. (Simulation 1) Imaginary part \tilde{q}_v of the unit quaternion associated with the attitude error \tilde{R} versus time (s).

and the second one evaluates the effects of measurement noise. For both simulations, the following initial estimation errors are considered: $\tilde{\xi}(0) = [4, 5, -5]^\top$, $q_{\tilde{R}}(0) = [\tilde{q}_0, \tilde{q}_v^\top]^\top(0) = [0.9119, -0.3079, -0.1673, -0.2135]^\top$ (the unit quaternion associated with $\tilde{R}(0)$, corresponding to errors in roll, pitch and yaw Euler angles of $40(\text{deg})$, $10(\text{deg})$, $30(\text{deg})$, respectively).

A. Simulation 1 – ideal measurements

Simulation results for this noise-free case are illustrated by Figures 2–5. Figures 2 and 4 show the fast convergence of the estimated position and attitude (represented by Euler angles) to the real values. In a different perspective, Figures 3 and 5 illustrate the observer performance by showing the convergence to zero of the estimation errors in terms of the norm of the position error $|\tilde{\xi}|$ and the imaginary part \tilde{q}_v of the unit quaternion associated with the attitude error \tilde{R} .

B. Simulation 2 – noisy measurements

The same scenario as in Simulation 1 is simulated but now with noise added in the measurements of Ω, V and of the source point bearings. More precisely, white Gaussian noise with variance of $2(\text{deg/s}) \approx 0.035(\text{rad/s})$ and $0.2(\text{m/s})$ are introduced on the measurements of Ω and V , respectively. We also try to simulate the image noise by first adding uncorrelated white Gaussian noise with variance of 0.01 to all components of the source point bearings and

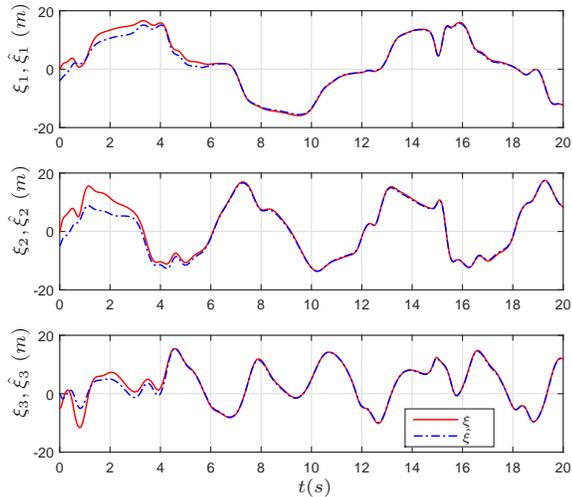


Fig. 6. (Simulation 2) Real and estimated positions (expressed in frame $\{B\}$) ξ , $\hat{\xi}$ (m) versus time (s).

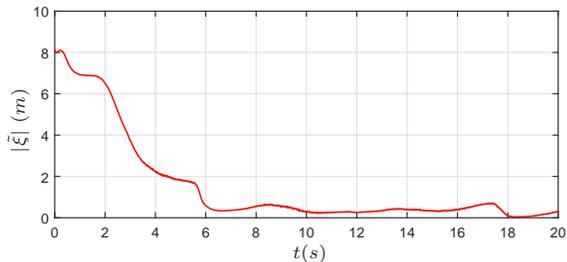


Fig. 7. (Simulation 2) Norm of position error $|\tilde{\xi}|$ (m) versus time (s).

then re-normalizing them onto S^2 to obtain noisy bearing measurements.

Simulation results reported in Figures 6–9 show that the ultimate time evolution of the estimated variables follows very near to the corresponding real variables and the ultimate estimation errors remain relatively small.

VI. CONCLUSIONS

The classical problem of pose estimation from bearing measurements is re-visited. However, unlike the classical problem of *essential matrix* estimation that requires at least 5 source points, we have shown that 3 source points would be sufficient to estimate the camera full pose if the camera translational motion is sufficiently exciting. To our knowledge, this discovery is a new result. By adopting the recent Riccati observer design framework developed in [5], we have proposed a nonlinear Riccati observer with the support of rigorous observability and stability analysis together with convincing simulation results.

Acknowledgment: This work was supported by the French ANR Astrid CONGRE (ANR-18-ASTR-0006) and FUI GreenExplorer projects.

REFERENCES

[1] OpenCV 3.1: Open source computer vision library. <https://github.com/opencv/opencv/releases/tag/3.1.0>, 2015.

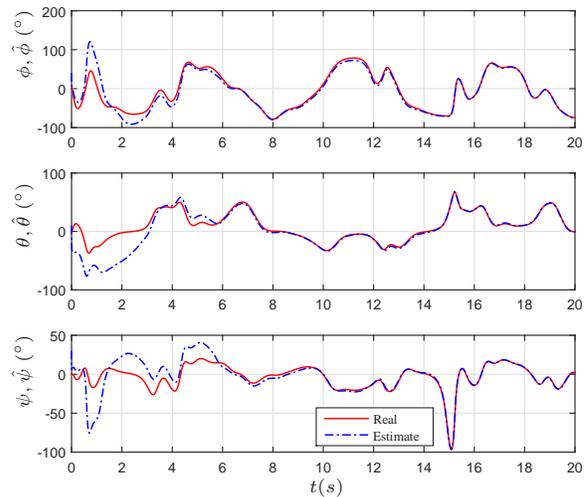


Fig. 8. (Simulation 2) Real and estimated attitudes represented by roll, pitch and yaw Euler angles ($^\circ$) versus time (s).

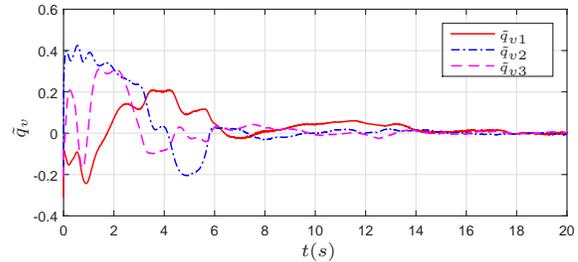


Fig. 9. (Simulation 2) Imaginary part \tilde{q}_v of the unit quaternion associated with the attitude error \tilde{R} versus time (s).

- [2] T. Botterill, S. Mills, and R. Green. Fast RANSAC hypothesis generation for essential matrix estimation. In *International Conference on Digital Image Computing Techniques and Applications (DICTA)*, pages 561–566, 2011.
- [3] T. Botterill, S. Mills, and R. Green. Refining essential matrix estimates from RANSAC. In *Proceedings Image and Vision Computing New Zealand*, pages 1–6, 2011.
- [4] M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6):381–395, 1981.
- [5] T. Hamel and C. Samson. Riccati observers for the nonstationary PnP problem. *IEEE Transactions on Automatic Control*, 63(3):726–741, 2017.
- [6] R. Hartley and A. Zisserman. *Multiple View Geometry in computer vision*. Cambridge University Press, 2003.
- [7] U. Helmke, K. Hüper, P. Y. Lee, and J. Moore. Essential matrix estimation using gauss-newton iterations on a manifold. *International Journal of Computer Vision*, 74(2):117–136, 2007.
- [8] Y. Ma, J. Košecká, and S. Sastry. Optimization criteria and geometric algorithms for motion and structure estimation. *International Journal of Computer Vision*, 44(3):219–249, 2001.
- [9] D. Nistér. An efficient solution to the five-point relative pose problem. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(6):756–770, 2004.
- [10] P. J. Rousseeuw. Least median of squares regression. *Journal of the American statistical association*, 79(388):871–880, 1984.
- [11] M. Sarkis, K. Diepold, and K. Hüper. A fast and robust solution to the five-point relative pose problem using gauss-newton optimization on a manifold. In *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, volume 1, pages 1–681, 2007.