ACCEPTED VERSION

Alireza Khosravian, Tat-Jun Chin, Ian Reid, Robert Mahony

A discrete-time attitude observer on SO(3) for vision and GPS fusion

Published version at: http://dx.doi.org/10.1109/ICRA.2017.7989669

http://hdl.handle.net/2440/112400

A Discrete-Time Attitude Observer on SO(3) for Vision and GPS Fusion

Alireza Khosravian, Tat-Jun Chin, Ian Reid, Robert Mahony

Abstract— This paper proposes a discrete-time geometric attitude observer for fusing monocular vision with GPS velocity measurements. The observer takes the relative transformations obtained from processing monocular images with any visual odometry algorithm and fuses them with GPS velocity measurements. The objectives of this sensor fusion are twofold; first to mitigate the inherent drift of the attitude estimates of the visual odometry, and second, to estimate the orientation directly with respect to the North-East-Down frame. A key contribution of the paper is to present a rigorous stability analysis showing that the attitude estimates of the observer converge exponentially to the true attitude and to provide a lower bound for the convergence rate of the observer. Through experimental studies, we demonstrate that the observer effectively compensates for the inherent drift of the pure monocular vision based attitude estimation and is able to recover the North-East-Down orientation even if it is initialized with a very large attitude error.

I. Introduction

Attitude estimation is the problem of determining the orientation of a vehicle with respect to a known frame of reference. Due to its broad application in robotics, navigation, and control of mechanical systems, this problem has been extensively studied in the past decades [1]–[10]. High-end IMUs are commonly used for pose estimation in applications, such as mining, that require long term accurate localisation. However, high-end IMUs are usually expensive, heavy, and power hungry. Cameras are attractive alternatives for pose estimation due to their low cost, small size, and high accuracy. Lots of elegant visual odometry algorithms have been developed in the literature and have been successfully used for localisation of moving robots (see e.g. [11]–[15] and the references therein). Monocular visual odometry algorithms rely on matching consecutive images to estimate the relative transformation of camera between those images. This relative transformation is concatenated over time to compute the absolute pose of the camera with respect to the first camera frame. Despite their popularity, there are two potential problems with pure visual odometry in applications that require long term accurate positioning; drift over time, and lack of a global known reference frame.

The concatenated estimate of camera motion based purely on the visual measurements inevitably drifts over time, due to uncertainties in camera intrinsic parameters, aggregation or numerical errors, and measurement noise. Although attenuated, the drift still exists even if graph based bundle adjustment techniques are used to optimally estimate the relative motion over tens of images [16]. One way of mitigating the drift is to perform loop closure [17]. However, not all robot paths necessarily contain loops in practice. Moreover, effective loop closure might be difficult in some applications, such as mapping of mine sites, where there is perceptual aliasing in the environment (i.e. several distinct places look similar), or where large moving vehicles, dust, wind, etc. change the appearance of the same location [17]. Also in real time applications, even if effective loop closure could be done, the estimate of robot's path still drifts between two consecutive occurrence of the loop closure.

The lack of a global known reference frame is another problem in pure vision based navigation. Many control objectives require knowledge of the robot's attitude with respect to a fixed known frame of reference (e.g. the North-East-Down (NED) frame) to navigate the robot toward a target point in the environment. Vision based orientation estimates can not be computed with respect to the NED frame as camera alone does not have any information about the NED frame. Although the concatenation of relative transformations yields an estimate of the orientation with respect to the initial camera frame, the transformation of the initial camera frame with respect to the NED frame is still unknown.

Fusing GPS measurements with vision helps mitigating both of the problems discussed above. GPS velocity measurements contain the heading information, that help compensating for the inherent drift of visual odometry estimates. Also, GPS measures the motion directly with respect to a known reference, enabling estimation of the attitude directly with respect to the NED frame. Fusing vision with GPS measurements is particularly of interest in open cut mine sites where RTK GPS is employed to provide centimetre-level position accuracy, albeit with the aid of reference stations. GPS measurements can be quite noisy and unreliable in deep mine pits, hence require fusion with high bandwidth sensors such as IMUs in Inertial Navigation Systems (INS). Fusing the low bandwidth GPS measurements with the high bandwidth visual measurements is a promising solution for long term localisation with a comparable level of accuracy to a high-end INS at a fraction of the cost (see section VII).

Recursive stochastic filtering techniques, such as (extended) Kalman filtering, are very popular in sensor fusion for navigation [1], [18]. Nevertheless, Kalman filters usually require careful tuning, lack rigorous proof of stability when applied to non-linear systems, and are prone to divergence if not initialised carefully [3], [19]–[23]. To address these issues, geometric nonlinear observer design methods on Lie

Alireza Khosravian, Tat-Jun Chin, and Ian Reid are with School of Computer Science, University of Adelaide ({alireza.khosravian, tat-jun.chin, ian.reid} @adelaide.edu.au). Robert Mahony is with the Research School of Engineering, Australian National University (robert.mahony@anu.edu.au)

groups have been developed in the past decade [21], [24]–[26]. These methods do not provide as rich information as stochastic methods due to the lack of the state covariance. Nevertheless, continuous-time geometric nonlinear attitude observers developed based on these methods are simple to tune, computationally cheap, and have demonstrated good performance in applications where high rate IMUs are present [3]–[10]. These methods, however, can not be directly applied to the vision-GPS fusion problem as the underlying kinematics of the this problem is naturally *discrete-time* due to the low rate of camera in most applications.

Elegant methods for designing discrete-time observers on Lie groups are developed in [27], [28]. Although suitable for satellite applications, the observer of [28] relies on attitude dynamics and the knowledge of external forces applied to the rigid body, hence is not applicable to the problem considered here. The observer of [27, equ. (18)] assumes availability of at least *two* non-collinear vectorial measurements and provides a novel stability analysis. As we will explain in this paper, however, GPS velocity measurements provide *only a single* vectorial measurement which prevents application of the stability theory of [27] to the particular problem considered here. To the best of our knowledge, there is no discrete-time geometric attitude observer in the literature concerning the problem of fusing vision and GPS velocity measurements (without the aid of IMU).

In this paper, we propose a discrete-time geometric observer on SO(3) for fusing monocular visual odometry estimates with GPS velocity measurements. The structure of the proposed observer is inspired by [27] and the proof of stability provided here is an extension to that work based on inspirations from [5], [9], [22], [29]. We show that the estimates of relative translation of camera (provided by a visual odometry algorithm) can be modelled as a vectorial measurement whose associated reference direction is the GPS velocity measurement. Although this is the sole vectorial measurement available in this application, we show that the attitude can be asymptotically estimated if the direction of the velocity of vehicle varies sufficiently with time (see Assumption 1). A key contribution of this paper is to provide a rigorous stability analysis yielding a deep understanding of the underlying fundamental performance limits of the vision and GPS fusion problem in terms of the convergence rate of the observer. Our fusion method is independent of the choice of the visual odometry algorithm, giving the users an extra freedom to apply any algorithm that suites their particular application. We demonstrate the application of the developed theory by experimental studies showing excellent performance of the proposed observer in compensating for the drift in the visual odometry and recovering the orientation of vehicle with respect to the NED frame.

The structure of paper is as follows. In Sections II and III formulate, for the first time, the attitude estimation problem using vision and GPS velocity fusion as a discrete-time geometric observer design problem on SO(3) with an associated vectorial measurement belonging to the unit sphere. The attitude observer is proposed in Section IV and the stability analysis is performed in Section V (see Theorem 1). We provide simulation studies as well as experimental results in Sections VI and VII, respectively, and we finish the paper with conclusions in Section VIII.

II. PROBLEM FORMULATION

Consider a vehicle equipped with a vision system and a GPS receiver moving in an environment with an unknown map. Assume that the GPS receiver and the camera are rigidly connected to the vehicle. Denote the North-East-Down (NED) reference frame (i.e. inertial frame) by $\{A\}$ and suppose that the reference frame $\{C\}$ is attached to the camera and moves with it. For simplicity, we refer to the camera reference frame when the k-th image was captured as the k-th camera frame. Denote the rotation of the i-th camera frame with respect to (wrt.) the j-th camera frame by $R_{C_i}^{C_j} \in$ SO(3). By matching corresponding feature points in two consecutive images k and $k+1$, visual odometry algorithms are capable of computing the relative rotation between the camera frames of those two images [11], [12] We assume that $R_{C_{k+1}}^{C_k}$ is provided by a visual odometry algorithm, and we treat this variable as a known measurement. One can obtain the rotation of the $(k+1)$-th camera frame wrt. the initial camera frame C_0 by concatenating previous relative rotations, as follows.

$$R_{C_{k+1}}^{C_0} = R_{C_k}^{C_0} R_{C_{k+1}}^{C_k}, \qquad R_{C_0}^{C_0} = I_3. \qquad (1)$$

Visual odometry techniques are also able to compute the relative translation of the consecutive camera frames. For monocular cameras, however, such a relative translation is valid only up to a constant scale since the depth can not be measured in the monocular setup [30]. Denote the translation of the i-th camera frame wrt. the j-th camera frame (and expressed in the j-th camera frame) by $p_{C_i}^{C_j} \in \mathbb{R}^3$. Assuming that the camera and GPS receiver are placed sufficiently close together[1], it is shown in Appendix I that

$$p_{C_{k+1}}^{C_k} = \frac{1}{d} R_k^\top (p_{k+1}^A - p_k^A) \qquad (2)$$

where $p_{k+1}^A \in \mathbb{R}^3$ is the k-th camera position (wrt. and expressed in the NED frame), $p_{C_{k+1}}^{C_k}$ is the relative translation of the k-th camera frame wrt. the $(k+1)$-th camera frames (expressed in the k-th camera frame) computed by the visual odometry algorithm and treated here as a known measurement, and $R_k \in$ SO(3) is the rotation of the k-th camera frame wrt. the NED frame. This rotation is related to $R_{C_k}^{C_0}$ via $R_k = R_{C_0}^A R_{C_k}^{C_0}$ where $R_{C_0}^A \in$ SO(3) is the unknown initial rotation of the camera frame wrt. the NED frame. One can use the GPS position measurements to compute $p_{k+1}^A - p_k^A$ in (2), nevertheless, numerical differencing the noisy GPS position measurements further increases the level of noise. For this reason, it is recommended to use the GPS

---

[1] It is possible to extend the derivations of this paper to the case where the camera and the GPS receiver are not placed close together. We consider that as a possible future work.

velocity measurements to approximate $p_{k+1}^A - p_k^A$ as follows[2].

$$p_{k+1}^A - p_k^A = \int_{t_k}^{t_{k+1}} v^A(s)\mathrm{d}s \approx (t_{k+1} - t_k)\bar{v}_k^A. \tag{3}$$

If the vehicle's velocity is approximately constant from t_k to t_{k+1}, we simply use $\bar{v}_k^A = v_k^A \in \mathbb{R}^3$ where v_k^A is the GPS velocity measurement at time t_k. Otherwise, if the velocity changes linearly with time from t_k to t_{k+1}, we use the approximation $\bar{v}_k^A = 0.5(v_{k+1}^A + v_k^A)$.

The aim of this paper is to estimate R_k using the discrete-time dynamics model (1) and the output measurement model (2). That is to fuse the outputs of visual odometry algorithm $R_{C_{k+1}}^{C_k}$ and $p_{C_{k+1}}^{C_k}$ with the GPS velocity measurement v_k (or equivalently the difference of consecutive position measurements if the velocity is unavailable). There are several challenges in such an estimation problem, some of which have been listed below.

- As opposed to classical linear estimation problem where the state space is a vector space, here the underlying system (1) evolves on the manifold SO(3).
- The underlying dynamics is discrete-time, which prevents direct application of previous continuous-time observer design techniques on Lie groups [21], [24]–[26].
- The scale d of the translation measurements provided by the visual odometry algorithm is unknown.
- The GPS measurements do not immediately provide information about the camera rotation. More specifically, the GPS measurements are related to the camera orientation via (2) where the initial camera orientation $R_{C_0}^A$ is unknown. Also, since (2) only provides a single 3D vectorial measurement, the rotation is not instantaneously observable.

Note that we do not consider stochastic measurement noise in our modeling as our design framework is based on the deterministic geometric setup [21], [24]–[26].

III. VISION-GPS FUSION AS A GEOMETRIC OBSERVER DESIGN PROBLEM

In this section, we transform the underlying system kinematics (1) and the output model (2) to a form more suitable for observer design purpose. We normalize the measurements involved in (2) to eliminate the unknown scale d. Defining

$$\bar{p}_k^A := \frac{p_{k+1}^A - p_k^A}{\|p_{k+1}^A - p_k^A\|} \approx \frac{\bar{v}_k^A}{\|\bar{v}_k^A\|}, \tag{4}$$

$$\bar{p}_k^C := \frac{p_{C_{k+1}}^{C_k}}{\|p_{C_{k+1}}^{C_k}\|}, \tag{5}$$

we have

$$\bar{p}_k^C = R_k^\top \bar{p}_k^A. \tag{6}$$

[2]Velocity measurements are computed in many GPS chips from the Doppler shift of the pseudo-range signals, yielding far more accurate velocity measurements compared to numerically differencing the range based position measurements. This accuracy is further enhanced if time-differenced carrier phase (TDCP) based GPS is used.

The variables $\bar{p}_k^C$ and $\bar{p}_k^A$ are unit vectors expressed in the camera frame and the NED frame, respectively. The normalization is particularly important since the scale of the visual odometry estimates drifts over time. The normalization not only isolates the effect of the scale drift, but also eliminates the need for knowledge of the velocity time stamp t_k. Equation (6) formulates the output measurement model of the system as a right action of SO(3)(the state space) on S^2 (the output manifold) where $\bar{p}_k^A$ represent a *time-varying reference output* in this context [22], [24], [29], [31]. This model directly relates the measurements $\bar{p}_k^C$ and $\bar{p}_k^A$ to the unknown state R_k and is used in the observer design.

Multiplying the sides of (1) by the constant rotation $R_{C_0}^A$ from the left and recalling $R_k = R_{C_0}^A R_{C_k}^{C_0}$, we have

$$R_{k+1} = R_k R_{C_{k+1}}^{C_k}, \qquad R_0 = R_{C_0}^A. \tag{7}$$

Now, the observer design problem is simplified to estimate R_k governing the dynamics (7) using the measurement of $R_{C_{k+1}}^{C_k}$ (as the input of the dynamics (7)) and the vectorial measurements $\bar{p}_k^C$ and $\bar{p}_k^A$ modeled via (6). The main advantage of formulating the problem as above compared to Section II is that the initial camera rotation $R_{C_0}^A$ now appear as an unknown initial condition of the estimation problem rather than as unknown variables in the system equation.

IV. OBSERVER DESIGN ON SO(3)

The underlying kinematics of the system (7) is left-invariant and the considered output model (6) is right-invariant (in the sense of [24], [31]). Inspired by the novel work of [27], we consider the following structure for our observer.

$$\hat{R}_{k+1} = \hat{R}_k \exp(-\hat{R}_k^\top \mathrm{grad}_1\phi(\hat{R}_k, \bar{p}_k^C))R_{C_{k+1}}^{C_k}, \tag{8}$$

with $\hat{R}_0 = I_3$, where $\phi : SO(3) \times S^2 \rightarrow \mathbb{R}^+$ is a right-invariant cost function (in the sense of [24]) and $\mathrm{grad}_1\phi(\hat{R}_k, \bar{p}_k^C) \in T_{\hat{R}_k}SO(3)$ is the gradient of $\phi(., \bar{p}_k^C) :$ $SO(3) \rightarrow \mathbb{R}^+$ wrt. a right-invariant Riemannian metric (see [24, Proposition 6.1.]). The motivation behind the observer structure (8) is that it generates an autonomous gradient-like estimation error dynamics if the reference output is constant [27]. We naturally choose the simple cost function

$$\phi(\hat{R}_k, \bar{p}_k^C) = (\hat{R}_k\bar{p}_k^C - \bar{p}_k^A)^\top L(\hat{R}_k\bar{p}_k^C - \bar{p}_k^A) \tag{9}$$

where $L \in \mathbb{R}^{3\times3}$ is a constant positive definite matrix. Observe that $\hat{R}_k = R_k$ implies $\phi = 0$. Employing standard differential geometry derivations, the gradient of ϕ wrt. the induced Riemannian metric $\langle \Omega_1 R, \Omega_2 R \rangle_R = \mathrm{tr}(\Omega_2^\top \Omega_1)$ is given by [32, Section VI]

$$\mathrm{grad}_1\phi(\hat{R}_k, \bar{p}_k^C) = -\left(\left(L(\hat{R}_k\bar{p}_k^C - \bar{p}_k^A)\right)_\times \hat{R}_k\bar{p}_k^C\right)_\times \hat{R}_k. \tag{10}$$

where $(.)_\times$ denotes the isomorphism from $\mathbb{R}^3$ to $\mathfrak{so}(3)$ such that $a_\times b$ yields the vector cross product $a \times b$ for any

$a, b \in \mathbb{R}^3$. Using (8) and (10), and employing the property $\exp(\hat{R}^\top a_\times \hat{R}) = \hat{R}^\top \exp(a_\times)\hat{R}$, we propose the observer

$$\hat{R}_{k+1} = \exp\left(\left(\left(L(\hat{R}_k \bar{p}_k^C - \bar{p}_k^A)\right)_\times \hat{R}_k \bar{p}_k^C\right)_\times\right)\hat{R}_k R_{C_{k+1}}^{C_k}, \quad \hat{R}_0 = I_3. \tag{11}$$

Remark 1: The observer (11) assumes GPS measurements are synchronized with images. In practice, the sampling rate of GPS is usually lower than that of the camera. In this case, one can introduce the auxiliary state $\hat{R}_k^+$ and break the observer (11) into two stages; prediction and update. In the prediction stage, the auxiliary state is updated by $\hat{R}_{k+1}^+ = \hat{R}_k R_{C_{k+1}}^{C_k}$ each time a new relative orientation $R_{C_{k+1}}^{C_k}$ is provided by visual odometry. In the update stage, the state $\hat{R}$ is updated via $\hat{R}_{k+1} = \exp\left(\left(\left(L(\hat{R}_k^+ \bar{p}_k^C - \bar{p}_k^A)\right)_\times \hat{R}_k^+ \bar{p}_k^C\right)_\times\right)\hat{R}_k^+$ when GPS provides a new measurement. It is possible to show that the stability results of Section V still hold when the two-stage implementation of the observer is employed.

V. STABILITY ANALYSIS

Although the structure of the proposed observer (11) is inspired by the novel work of [27], the stability analysis presented in [27, Proposition 4] does not apply here as the cost function (9) does not have a unique global minimum. Instead, we provide a new stability analysis by extending the results of [22], [29]. Consider the attitude estimation error.

$$E_k = \hat{R}_k R_k^\top. \tag{12}$$

To prove $\hat{R}_k \to R_k$, we should show $E_k \to I_3$. Using (7), (11), (6), and (12), the dynamics of E_k are given by

$$E_{k+1} = \hat{R}_{k+1} R_{k+1}^\top = \exp\left(\left(\left(L(\hat{R}_k \bar{p}_k^C - \bar{p}_k^A)\right)_\times \hat{R}_k \bar{p}_k^C\right)_\times\right)\hat{R}_k R_k^\top$$

$$= \exp\left(\left(\left(L(E_k \bar{p}_k^A - \bar{p}_k^A)\right)_\times E_k \bar{p}_k^A\right)_\times\right)E_k. \tag{13}$$

One can verify that $E_k = I_3$ is an equilibrium point of the error dynamics (13). If $\bar{p}_k^A$ were constant and if we had two or more vector measurements, the elegant stability analysis of [27, Proposition 4] would be applied. However, only one vector measurement is available here. Since $\bar{p}_k^A$ is time-varying, the stability analysis is tedious even in the continuous-time case [22], [29]. Inspired by the continuous-time case, here we propose a novel stability analysis for the discrete-time case. The following assumption formally formulates the time-varying property of $\bar{p}_k^A$.

Assumption 1 (Persistency of excitation): There exist a positive integer T and a constant $\beta > 0$ such that

$$\frac{1}{T+1}\sum_{i=k}^{k+T}(I_3 - \bar{p}_i^A \bar{p}_i^{A\top}) \geq \beta I_3, \tag{14}$$

for all non-negative integers k. □

Assumption 1 could be written in other equivalent forms, see [22, Lemma 1]. This assumption is satisfied if $\bar{p}_k^A$ varies with time, i.e. if the vehicle's direction of motion is not constant all the time (see [22, Lemma 1]). The only case where this assumption does not hold is when the vehicle's trajectory is a perfectly straight line. It is worth mentioning

that in the closely related problem of velocity-aided attitude estimation where an IMU is present instead of a camera, a similar condition as (14) is required unless an extra sensor (such as magnetometer) is available [33]. The more the direction of the motion of the vehicle changes, the larger β gets. Nevertheless, since $\bar{p}_k^A$ is a unit vector, the largest singular value of the matrix $I_3 - \bar{p}_i^A \bar{p}_i^{A\top}$ is one and we have $\beta < 1$ [22]. The following theorem summarizes the stability of the error dynamics (13). To simplify presentation of the results, we consider the observer gain $L = lI_3$ where l is a scalar. It is straight-forward to extend the results to a more general positive definite gain matrix.

Theorem 1: Consider the observer (11) for the system (7) with the output (4)-(6). Suppose that Assumption 1 holds. Assume moreover that $L = lI_3$ where $0 < l < 2$. Then, the error dynamics (13) is locally exponentially stable to $E_k = I_3$. Moreover, there exists a compact region around $E_k = I$ in which

$$\|I_3 - E_k\|_F \leq \|I_3 - E_0\|_F \alpha^k, \tag{15}$$

where $0 < \alpha := \left(1 - \frac{\beta(T+1)(2l-l^2)}{2+l^2 T(T+1)}\right)^{\frac{1}{2(1+T)}} < 1$ indicates a lower bound for the convergence rate and $\|I_3 - E_k\|_F^2 = \mathrm{tr}((I_3 - E_k)^\top(I_3 - E_k)) = 2\mathrm{tr}(I_3 - E_k)$ is the induced Frobenius norm on SO(3). □

Proof of Theorem 1: the proof is mainly inspired by [22, Section V] which investigates the continuous-time scenario. There are subtle technical differences in the discrete-time case that prevent direct application of the continuous-time case. For this reason, here we provide a stability proof for the discrete-time case in detail. The first order approximation of the error E_k around the identity is given by $E_k \approx I + \epsilon_{k\times}$ where $\epsilon_k \in \mathbb{R}^3$. Substituting for this approximation into the error dynamics (13) and resorting to Appendix II, we obtain the linearized error dynamics

$$\epsilon_{k+1} = \epsilon_k - lP_k \epsilon_k, \tag{16}$$

where

$$P_k := -\bar{p}_k^A {}_\times \bar{p}_k^A {}_\times = I_3 - \bar{p}_k^A \bar{p}_k^{A\top} \tag{17}$$

is a time-varying positive semi-definite matrix whose rank is two. The Assumption 1 imposes a persistency of excitation condition on P_k which can increase the rank of its summation over time to three. Consider the following Lyapunov candidate.

$$\mathcal{L}_k = \epsilon_k^\top \epsilon_k. \tag{18}$$

Using (16) and noting that $P_k^\top = P_k$ and $P_k^2 = P_k$ (due to (17) and the fact that $\bar{p}_k^A$ has a unit norm), we have

$$\mathcal{L}_{k+1} - \mathcal{L}_k = (\epsilon_k - lP_k \epsilon_k)^\top (\epsilon_k - lP_k \epsilon_k) - \epsilon_k^\top \epsilon_k \tag{19}$$

$$= -\gamma \epsilon_k^\top P_k \epsilon_k. \tag{20}$$

where $\gamma := 2l - l^2$. Since $0 < l < 2$, we have $0 < \gamma < 1$. Since P_k is positive semi-definite, we have $\mathcal{L}_{k+1} - \mathcal{L}_k \leq$

0 which indicates that the Lyapunov candidate is non-increasing along the system trajectories. Summing the sides of (20) from k to $k+T$ yields

$$\mathcal{L}_{k+1+T} - \mathcal{L}_k = -\gamma \sum_{i=k}^{k+T} \epsilon_i^\top P_i \epsilon_i \qquad (21)$$

Since $P_i \geq 0$, there exist $N_i \in \mathbb{R}^{2\times 3}$ (with rank 2) such that $P_i = N_i^\top N_i$. Choose $a = N_i\epsilon_k$ and $b = N_i(\epsilon_i - \epsilon_k)$ and employ the property $(a+b)^\top(a+b) \geq \frac{1}{2}a^\top a - b^\top b$ to obtain

$$\sum_{i=k}^{k+T} \epsilon_i^\top P_i \epsilon_i \geq \frac{1}{2}\sum_{i=k}^{k+T} \epsilon_k^\top P_i \epsilon_k - \sum_{i=k}^{k+T} (\epsilon_i^\top - \epsilon_k^\top)P_i(\epsilon_i - \epsilon_k). \qquad (22)$$

Using the derivations of Appendix III, we have

$$\sum_{i=k}^{k+T} (\epsilon_i^\top - \epsilon_k^\top)P_i(\epsilon_i - \epsilon_k) \leq \frac{l^2 T(T+1)}{2}\sum_{i=k}^{k+T} \epsilon_i^\top P_i \epsilon_i \quad (23)$$

Also, (14) and (18) yield

$$\sum_{i=k}^{k+T} \epsilon_k^\top P_i \epsilon_k \geq \beta(T+1)\epsilon_k^\top \epsilon_k = \beta(T+1)\mathcal{L}_k \qquad (24)$$

Substituting for (23) and (24) into (22), we obtain

$$\sum_{i=k}^{k+T} \epsilon_i^\top P_i \epsilon_i \geq \frac{\beta(T+1)}{2}\mathcal{L}_k - \frac{l^2 T(T+1)}{2}\sum_{i=k}^{k+T} \epsilon_i^\top P_i \epsilon_i, \qquad (25)$$

which implies $\sum_{i=k}^{k+T} \epsilon_i^\top P_i \epsilon_i \geq \frac{\beta(T+1)}{2+l^2 T(T+1)}\mathcal{L}_k$. Substituting this into (21) yields

$$\mathcal{L}_{k+1+T} \leq \bar{\alpha}\mathcal{L}_k. \qquad (26)$$

where $\bar{\alpha} := 1 - \frac{\beta(T+1)(2l-l^2)}{2+l^2 T(T+1)}$. Recalling that $0 < l < 2$ and $T \geq 1$, it is straight-forward to show that $1 - \frac{\sqrt{2}}{2}\beta < \bar{\alpha} < 1$. Employing (17) and recalling that $\bar{p}_k^A$ is a unit vector, we conclude that the largest singular value of P_k is one. This, together with (14) implies that $0 < \beta \leq 1$. Hence, $0 < \frac{\sqrt{2}}{2} < \bar{\alpha} < 1$. Taking $k = n(1+T)$, where n is a non-negative integer, and successively employing (26) yields

$$\mathcal{L}_k \leq \mathcal{L}_{n(1+T)} \leq \bar{\alpha}^n \mathcal{L}_0, \quad \forall k \geq n(1+T) \qquad (27)$$

Hence, $\mathcal{L}_k$ decays exponentially to zero as n goes to infinity. Substituting $\mathcal{L}_k$ for (18) and considering $n = \frac{k}{1+T}$ yields

$$\|\epsilon_k\| \leq \|\epsilon_0\|\alpha^k \qquad (28)$$

where α is defined in the statement of Theorem 1. Consequently, the equilibrium point $\epsilon_k = 0$ of the linearized error dynamics (16) is uniformly globally exponentially stable [34]. Hence, by extending the results of [35, Theorem 4.15] to the discrete-time case according to the guidelines of [34], one concludes that the equilibrium point $E_k = I$ of the error dynamics (13) is uniformly locally exponentially stable. The exponential bound (15) follows from (28) directly. ∎

By Theorem 1, the convergence rate of the observer depends on the value of β which in turn depends on the
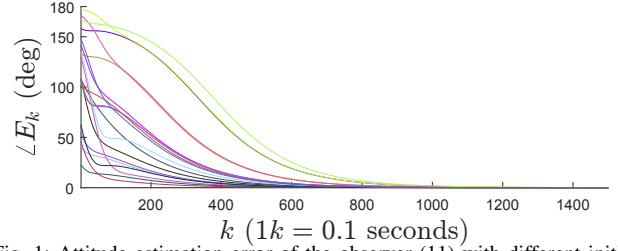

Fig. 1: Attitude estimation error of the observer (11) with different initialization error. Note that 180 (deg) indicates the maximum possible error.

motion of the vehicle (equ. (14)). Having an analytical solution for the convergence rate (provide by Theorem 1) makes it possible to obtain the value of the observer gain l that yields the fastest convergence rate. This can be done by considering a moving horizon of time and numerically computing T and β satisfying (14) for that horizon. Given β and T, one can analytical compute the value of the observer gain l that maximizes the convergence rate α. This yields a time-varying gain as T and β should be updates as the time horizon changes. It is straight-forward to extend the stability results of Theorem 1 to a time-varying gain.

VI. SIMULATION RESULTS

Theorem 1 proves local stability of the error dynamics around the identity, but it does not provide an estimate of the domain of attraction. As the initial orientation of the camera wrt. to the NED frame is unknown, one cannot always choose the initial condition of the observer (11) close to the actual value of $R_{C_0}^A$. The aim of this section is to examine the domain of attraction of the error dynamics (13). Consider a camera moving with a constant velocity of $2\pi \approx 6.3$ (m/s) in a circular trajectory with the radius of 50 (m) in the North-East plane. The NED frame is set to be at the centre of the circle. The camera starts its motion from the far North part of the path and moves clockwise to East. Assuming the sampling interval of 0.1 seconds, the corresponding relative motions $R_{C_{k+1}}^{C_k}$ and $p_{C_{k+1}}^{C_k}$ are computed to model the output of visual odometry and passed to the observer (11) together with the linear velocity of the camera wrt. the NED frame. We neither model camera images nor we add noise to the measurements as the aim of this section is to only investigate the effect of initial condition on the attitude estimation error $\angle E_k = \text{acos}(1 - \frac{1}{4}\|I - E_k\|_F^2)^3$. We perform 20 simulations where in each simulation the observer is initialised with a random rotation corresponding to a random axis of rotation in a unit spare and a random angle of rotation between 0 and 179 degrees. Fig. 1 shows the results of the simulations, demonstrating that the attitude estimation error converges to zero even when a very large initial estimation error is chosen.

VII. EXPERIMENTAL RESULTS

In this section, we demonstrate the performance of the proposed observer (11) with real world experiments. The experiments have been done in collaboration with Maptek

---

[3] The error $\angle E_k$ corresponds to the angle of rotation in the angle-axis decomposition of E_k which models how far $\hat{R}_k$ is from R_k [22], [25].
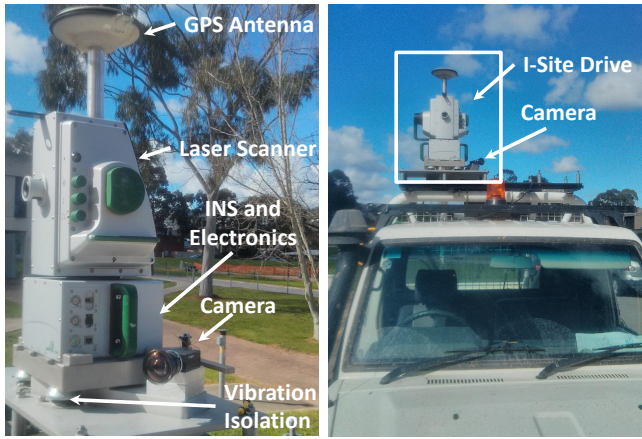
Fig. 2: Right: the experimental setup mounted on a truck. Left: enlarged photo of the experimental setup showing its components.
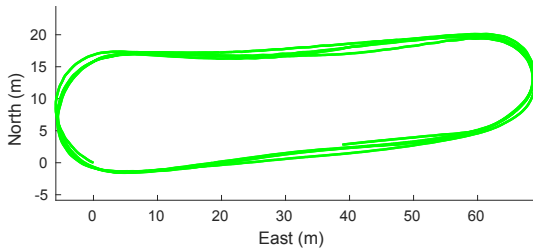

Fig. 3: Top view of the ground truth trajectory of the vehicle. The vehicle starts at the origin and moves clockwise.

Pty Ltd[4]. The experimental setup consists of Maptek's I-Site Drive system [5] together with a Pointgrey Grasshopper3 GS3-U3-23S6C-C camera[6] mounted on a truck (see Fig 2). The I-Site Drive includes the NovAtel SPAN-CPT[7] high precision Inertial Navigation System (INS) consist of a RTK GPS and an inertial measurement unit, providing accurate estimates of the 6DoF transformation of the I-Site Drive system wrt. the inertial frame which is used as the ground truth (R_k, p_k) in our experiments. The extrinsic camera to I-Site Drive calibration transformation have been carefully computed using a hand-eye calibration method that enables computing ground truth transformation of the camera wrt. the NED frame [36]. The camera is configured to capture 1920×1200 pixels images at approximately 22 fps. The camera is hardware synced to the I-Site Drive system to provide precise time stamping of measurements for accurate evaluation of the results. In our experiments, we use ORB-SLAM [13] as the visual odometry algorithm. ORB-SLAM is a state-of-the-art large scale simultaneous localisation and mapping software. Direct outputs of ORB-SLAM are estimates of the translation of the k-th camera frame wrt. the initial camera frame. Given two consecutive transformation estimates $(\check{R}_{C_k}^{C_0}, \check{p}_{C_k}^{C_0})$ and $(\check{R}_{C_{k+1}}^{C_0}, \check{p}_{C_{k+1}}^{C_0})$ as direct outputs of ORB-SLAM, we compute $R_{C_{k+1}}^{C_k} = \check{R}_{C_k}^{C_0\top} \check{R}_{C_{k+1}}^{C_0}$ and

[4] http://www.maptek.com/
[5] http://www.maptek.com/products/i-site/i-site_drive.html
[6] https://www.ptgrey.com/grasshopper3-23-mp-color-usb3-vision-sony-pregius-imx174
[7] http://www.novatel.com/products/span-gnss-inertial-systems/span-combined-systems/span-cpt/
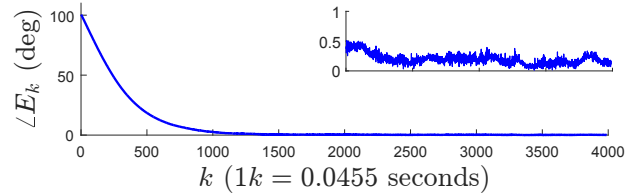

Fig. 4: Attitude estimation error. The enlarged portion of the figure shows the steady-state estimation error.

$p_{C_{k+1}}^{C_k} = \check{R}_{C_k}^{C_0\top}(\check{p}_{C_{k+1}}^{C_0} - \check{p}_{C_k}^{C_0})$ (recalling (1) and (29)) and we use them in the observer (11). For our experiments, we disable the loop closure thread of ORB-SLAM to observe the inherent drift of the pure visual odometry[8].

After careful calibration, the camera intrinsic parameters are estimated as $(1061.01, 1062.22)$ for the focal length and $(951.28, 604.86)$ for the principal point. The truck follows the path shown in Fig. 3 during which 4000 images are captured. Since raw velocity measurement log was not available in our experiment, we use the numerical difference of the position estimates of the INS system as estimates of the linear velocity in our observer. This yields noisy estimates of the linear velocity, but is good enough for the purpose of demonstration in this section. We choose $L = 0.004 I_3$ and we initialise the observer (11) with $\hat{R}_0 = I_3$ since the vision-GPS system does not initially have any information about the rotation of the camera wrt. the NED frame. Fig. 4 shows that despite the very large initial attitude estimation error, $\angle E_k$ converges to a very close neighborhood of zero. The observer is capable of accumulating the information of velocity measurements over time to asymptotically estimate the camera's orientation wrt. the NED frame. Fig. 5 shows the Euler angle decomposition of the ground truth orientation R_k versus its estimate $\hat{R}_k$. The Euler angles estimates follow the ground truth angles very closely, demonstrating excellent performance of the observer. The covariance of the INS attitude estimate is about 0.2 degrees in our experiment. Fig. 4 and Fig. 5 show that the vision-GPS system is capable of providing the level of accuracy close to a high-end INS system, albeit at a fraction of its cost.

As explained in Section II, one purpose of fusing GPS measurements with vision is to mitigates the inherent drift due to accumulation of the visual odometry error. In order to show this effect, we use the ground truth initial orientation $R_{C_0}^A$ to transform the direct orientation estimates of ORB-SLAM $\check{R}_{C_k}^{C_0}$ to the NED frame via $\check{R}_k := R_{C_0}^A \check{R}_{C_k}^{C_0}$. This gives the estimates of R_k with pure visual odometry if the initial attitude was known. We also initialise the observer (11) with $\hat{R}_0 = R_{C_0}^A$ and obtain the attitude estimates $\hat{R}_k$. Fig. 6a shows that the attitude estimation error of the pure visual odometry $\angle(\check{R}_k R_k^\top)$ (red curve) drifts over time while the proposed observer (11) successfully compensates for this drift by fusing the velocity such that the magnitude of the attitude error $\angle E_k$ (blue curve) remains bounded for all

[8] We recall that effective loop closure is not possible in some mining applications due to the restrictions of vehicle's path (not necessarily containing loops), high perceptual aliasing, large moving vehicles, and change in the appearance of the environment due to dust, wind, etc.
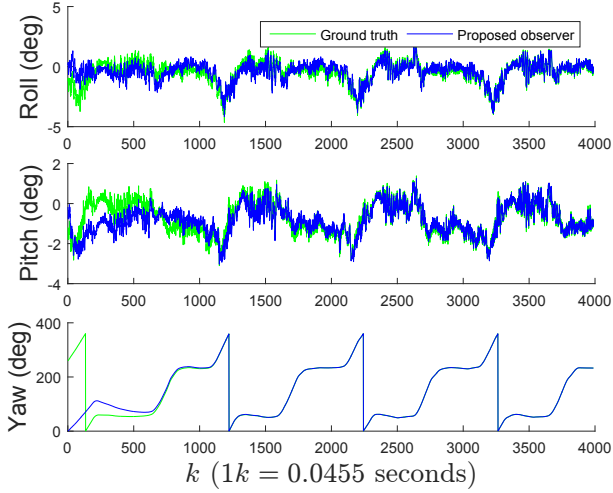
Fig. 5: Ground truth Euler angle vs their estimates via the observer (11).

Fig. 6: Attitude estimation error of the pure visual odometry versus the proposed observer (11); (a) 1920×1200 images are used, (b) 960×600 images are used, (c) 1% error is added to the intrinsic camera calibration.

times. All of the results presented so far are obtained by processing the full resolution images. The visual odometry algorithm might not yield real time performance with full resolution of the camera. Here, we resized the images to 960×600 and perform the same experiment as Fig. 6a. Fig. 6b compares the attitude estimation error of the pure visual odometry versus that of the observer (11). Compared to Fig. 6a, the accumulation of the visual odometry error is significantly higher while the observer error remains almost unchanged. Lastly, we point out that the presented results so far are obtained after very careful calibration of the camera intrinsic parameters. Even slight calibration error may significantly increase the drift of the pure visual odometry results while such drifts are compensated when the observer (11) is employed. To demonstrate this, we add 1 percent error to the camera intrinsic parameters used in ORB-SLAM and reprocess the full resolution images. Fig. 6c compares the attitude estimation of the pure visual odometry versus the observer (11) both initialised with the ground truth orientation. Compared to Fig. 6a, the attitude estimation error of the pure visual odometry significantly increases while the observer (11) successfully compensates for such a drift.

VIII. CONCLUSION

This paper formulates the vision and GPS velocity fusion as an observer design problem on the Lie group SO(3)with a single vectorial output measurement whose associated reference output is time-varying. Inspired by [5], [9], [22], [27], [29], we propose a discrete-time attitude observer and provide a rigorous stability analysis showing the exponential convergence of the attitude estimation error to zero. We present experimental studies demonstrating that the observer is effectively capable of mitigating the inherent drift of pure visual odometry estimates. The observer also enables direct attitude estimation wrt. the NED frame, making it ideal for applications involving navigation and control of robots in the inertial frame. Where depth (i.e. scale) is available through stereo vision, extending the proposed observer to a estimate the pose on SE(3) is a potential future work.
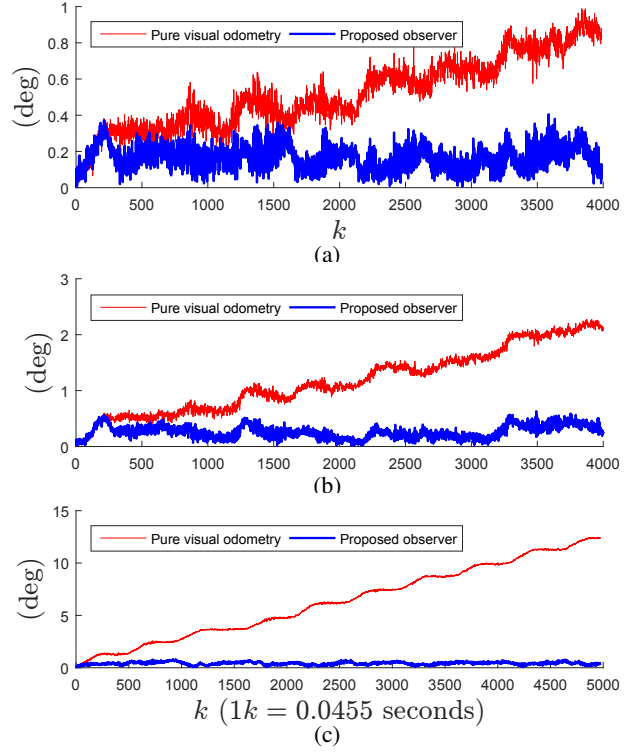
APPENDIX I

GPS measurement model (2): similar to (1), one can concatenate the camera positions $p_{C_{k+1}}^{C_0}$ to obtain

$$p_{C_{k+1}}^{C_0} - p_{C_k}^{C_0} = R_{C_k}^{C_0} p_{C_{k+1}}^{C_k}, \qquad p_{C_0}^{C_0} = 0. \qquad (29)$$

On the other hand, we have

$$p_{C_{k+1}}^{C_0} - p_{C_k}^{C_0} = R_{C_0}^{A\top}(p_{C_{k+1}}^A - p_{C_k}^A). \qquad (30)$$

where $R_{C_0}^{A\top}$ transforms the vector $p_{C_{k+1}}^A - p_{C_k}^A$ from the reference frame $\{A\}$ to the frame $\{C_0\}$. Substituting for (30) into (29) and noting that GPS positions are related to $p_{C_{k+1}}^A - p_{C_k}^A$ via $p_{k+1}^A - p_k^A = d(p_{C_{k+1}}^A - p_{C_k}^A)$, yields (2).

APPENDIX II

Linearizion of error dynamics (13): using $E_k \approx I_3 + \epsilon_{k\times}$, we have $\left(L(E_k \bar{p}_k^A - \bar{p}_k^A)\right)_\times E_k \bar{p}_k^A \approx \left(L((I_3 + \epsilon_{k\times})\bar{p}_k^A - \bar{p}_k^A)\right)_\times (I_3 + \epsilon_{k\times})\bar{p}_k^A \approx \left(L\epsilon_{k\times}\bar{p}_k^A\right)_\times (I_3 + \epsilon_{k\times})\bar{p}_k^A \approx \left(L\epsilon_{k\times}\bar{p}_k^A\right)_\times \bar{p}_k^A$ where we ignore the terms that are of order two or higher on ϵ_k. Using the Taylor series expansion of the exponential map, we have $\exp\left(\left(\left(L(E_k\bar{p}_k^A - \bar{p}_k^A)\right)_\times E_k\bar{p}_k^A\right)_\times\right)E_k \approx (I_3 + \left(\left(L\epsilon_{k\times}\bar{p}_k^A\right)_\times\bar{p}_k^A\right)_\times + \text{HOT})(I_3 + \epsilon_{k\times}) \approx I_3 + \epsilon_{k\times} + \left(\left(L\epsilon_{k\times}\bar{p}_k^A\right)_\times\bar{p}_k^A\right)_\times + \text{HOT}$. Substituting this into (13) and considering $E_{k+1} \approx I_3 + \epsilon_{k+1\times}$, we have $I_3 + \epsilon_{k+1\times} = I_3 + \epsilon_{k\times} + \left(\left(L\epsilon_{k\times}\bar{p}_k^A\right)_\times\bar{p}_k^A\right)_\times$. Canceling I_3 from the sides and taking the inverse of the operator $(.)_\times$ from the sides, we obtain $\epsilon_{k+1} = \epsilon_k + (L\epsilon_{k\times}\bar{p}_k^A)_\times\bar{p}_k^A$. Observing $(L\epsilon_{k\times}\bar{p}_k^A)_\times\bar{p}_k^A = \bar{p}_k^A{}_\times L\bar{p}_k^A{}_\times\epsilon_k$, noting $L = lI_3$, and using the property $a_\times b_\times = ba^\top - a^\top bI_3$, yields (16).

APPENDIX III

Proof of (23): using (16), we have

$$\epsilon_i - \epsilon_k = \sum_{j=k}^{i-1} \epsilon_{j+1} - \epsilon_j = -l \sum_{j=k}^{i-1} P_j \epsilon_j. \tag{31}$$

Since the maximum singular value of P_i is one, and using (31) we have $\sum_{i=k}^{k+T} (\epsilon_i^\top - \epsilon_k^\top) P_i (\epsilon_i - \epsilon_k) \leq \sum_{i=k}^{k+T} (\epsilon_i^\top - \epsilon_k^\top)(\epsilon_i - \epsilon_k) = l^2 \sum_{i=k}^{k+T} \left(\sum_{j=k}^{i-1} \epsilon_j^\top P_j \right) \left(\sum_{j=k}^{i-1} P_j \epsilon_j \right)$. Using the Cauchy-Schwarz inequality [37] and noting $P_j P_j = P_j$, we obtain $\sum_{i=k}^{k+T} (\epsilon_i^\top - \epsilon_k^\top) P_i (\epsilon_i - \epsilon_k) \leq l^2 \sum_{i=k}^{k+T} (i - k) \sum_{j=k}^{i-1} \epsilon_j^\top P_j P_j \epsilon_j = l^2 \sum_{i=k+1}^{k+T} (i - k) \sum_{j=k}^{i-1} \epsilon_j^\top P_j \epsilon_j$, where we used the fact that the value of the first summand evaluated at $i = k$ is zero and hence we can start the first summation from $i = k + 1$. Changing the order of summations we obtain

$$\sum_{i=k}^{k+T} (\epsilon_i^\top - \epsilon_k^\top) P_i (\epsilon_i - \epsilon_k) \leq l^2 \sum_{j=k}^{k+T-1} \epsilon_j^\top P_j \epsilon_j \sum_{i=j+1}^{k+T} (i - k). \tag{32}$$

We have $\sum_{i=j+1}^{k+T} (i - k) = \frac{T(T+1)}{2} - \frac{(j-k+1)(j-k)}{2}$. Using the limits of the first summation we have $k \leq j \leq k+T-1$ which implies $0 \leq (j - k + 1)(j - k) \leq T(T - 1)$. Hence $\sum_{i=j+1}^{k+T} (i-k) \leq \frac{T(T+1)}{2}$. Using this and adding the positive term $\frac{l^2 T(T+1)}{2} \epsilon_{k+T}^\top P_{k+T} \epsilon_{k+T}$ to the right hand side of (32) yields (23) and completes the proof.

REFERENCES

[1] F. L. Markley and J. L. Crassidis, *Fundamentals of Spacecraft Attitude Determination and Control.* Springer, 2014.

[2] J. L. Crassidis, F. L. Markley, and Y. Cheng, "Survey of nonlinear attitude estimation methods," *Journal of guidance, control, and dynamics*, vol. 30, no. 1, pp. 12–28, 2007.

[3] R. Mahony, T. Hamel, and J.M. Pflimlin, "Nonlinear complementary filters on the special orthogonal group," *IEEE Trans. Autom. Control*, vol. 53, no. 5, pp. 1203–1218, 2008.

[4] H. Rehbinder and B. K. Ghosh, "Pose estimation using line-based dynamic vision and inertial sensors," *IEEE Trans. Automatic Control*, vol. 48, no. 2, pp. 186–199, 2003.

[5] H. F. Grip, T. I. Fossen, T. A. Johansen, and A. Saberi, "Attitude estimation using biased gyro and vector measurements with time-varying reference vectors," *IEEE Trans. Autom. Control*, vol. 57, no. 5, pp. 1332–1338, 2012.

[6] S. Bonnabel, P. Martin, and P. Rouchon, "A non-linear symmetry-preserving observer for velocity-aided inertial navigation," in *Proc. American Control Conf.*, 2006, pp. 2910–2914.

[7] J. Vasconcelos, C. Silvestre, and P. Oliveira, "A nonlinear observer for rigid body attitude estimation using vector observations," in *Proc. IFAC World Congr.*, Korea, July 2008.

[8] D. E. Zlotnik and J. R. Forbes, "A nonlinear attitude estimator with desirable convergence properties," in *European Control Conference*, 2015, pp. 2103–2107.

[9] D. Seo and M.R. Akella, "Separation property for the rigid-body attitude tracking control problem," *J. Guid., Control, Dynam.*, vol. 30, no. 6, pp. 1569–1576, 2007.

[10] A. Tayebi, S. McGilvray, A. Roberts, and M. Moallem, "Attitude estimation and stabilization of a rigid body using low-cost sensors," in *Proc. IEEE Conf. on Decision and Control*, USA, December 2007.

[11] F. Fraundorfer and D. Scaramuzza, "Visual odometry: Part I: The first 30 years and fundamentals," *IEEE Robotics and Automation Magazine*, vol. 18, no. 4, pp. 80–92, 2011.

[12] ——, "Visual odometry: Part II: Matching, robustness, optimization, and applications," *IEEE Robotics & Automation Magazine*, vol. 19, no. 2, pp. 78–90, 2012.

[13] R. Mur-Artal, J. Montiel, and J. D. Tardós, "ORB-SLAM: a versatile and accurate monocular SLAM system," *IEEE Trans. Robotics*, vol. 31, no. 5, pp. 1147–1163, 2015.

[14] J. Engel, J. Sturm, and D. Cremers, "Semi-dense visual odometry for a monocular camera," in *Proc. IEEE International Conf. Computer Vision*, 2013, pp. 1449–1456.

[15] T. Pire, T. Fischer, J. Civera, P. De Cristóforis, and J. Jacobo berlles, "Stereo Parallel Tracking and Mapping for robot localization," in *Proc. IEEE/RSJ International Conf. Intelligent Robots and Systems*, 2015, pp. 1373–1378.

[16] B. Triggs, P. F. McLauchlan, R. I. Hartley, and A. W. Fitzgibbon, "Bundle adjustment —a modern synthesis," in *International workshop on vision algorithms*, 1999, pp. 298–372.

[17] A. Angeli, D. Filliat, S. Doncieux, and J.-A. Meyer, "Fast and incremental method for loop-closure detection using bags of visual words," *IEEE Trans. Robotics*, vol. 24, no. 5, pp. 1027–1037, 2008.

[18] E. J. Lefferts, F. L. Markley, and M. D. Shuster, "Kalman filtering for spacecraft attitude estimation," *Journal of Guidance, Control, and Dynamics*, vol. 5, no. 5, pp. 417–429, 1982.

[19] M. Izadi, E. Samiei, A. K. Sanyal, and V. Kumar, "Comparison of an attitude estimator based on the lagrange-d'alembert principle with some state-of-the-art filters," in *IEEE International Conf. Robotics and Automation*, 2015, pp. 2848–2853.

[20] J. M. Roberts, P. I. Corke, and G. Buskey, "Low-cost flight control system for a small autonomous helicopter," in *Proc. IEEE International Conf. Robotics and Automation*, vol. 1, 2003, pp. 546–551.

[21] M. Zamani, "Deterministic attitude and pose filtering, an embedded Lie groups approach," Ph.D. dissertation, Australian National University, 2013.

[22] A. Khosravian and M. Namvar, "Globally exponential estimation of satellite attitude using a single vector measurement and gyro," in *Proc. IEEE Conf. Decision and Control*, 2010.

[23] A. Barrau *et al.*, "Invariant filtering for pose EKF-SLAM aided by an imu," in *IEEE Conference on Decision and Control*, 2015, pp. 2133–2138.

[24] A. Khosravian, J. Trumpf, R. Mahony, and C. Lageman, "Observers for invariant systems on lie groups with biased input measurements and homogeneous outputs," *Automatica*, vol. 55, pp. 19–26, 2015.

[25] C. Lageman, J. Trumpf, and R. Mahony, "Gradient-like observers for invariant dynamics on a Lie group," *IEEE Trans. Automa. Control*, vol. 55, no. 2, pp. 367–377, 2010.

[26] S. Bonnabel, P. Martin, and P. Rouchon, "Non-linear symmetry-preserving observers on Lie groups," *IEEE Trans. Autom. Control*, vol. 54, no. 7, pp. 1709–1713, 2009.

[27] A. Barrau and S. Bonnabel, "Intrinsic filtering on lie groups with applications to attitude estimation," *IEEE Transactions on Automatic Control*, vol. 60, no. 2, pp. 436–449, 2015.

[28] T. Lee, M. Leok, N. H. McClamroch, and A. Sanyal, "Global attitude estimation using single direction measurements," in *Proc. American Control Conf.*, USA, July 2007.

[29] J. Trumpf, R. Mahony, T. Hamel, and C. Lageman, "Analysis of nonlinear attitude observers for time-varying reference measurements," *IEEE Trans. Autom. Control*, vol. 57, no. 11, pp. 2789–2800, 2012.

[30] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision.* Cambridge university press, 2003.

[31] R. Mahony, J. Trumpf, and T. Hamel, "Observers for kinematic systems with symmetry," in *Proc. IFAC Symposium on Nonlinear Control Systems*, September 2013, pp. 617–633.

[32] A. Khosravian, J. Trumpf, R. Mahony, and C. Lageman, "Bias estimation for invariant systems on Lie groups with homogeneous outputs," in *Proc. IEEE Conf. on Decision and Control*, December 2013.

[33] M.-D. Hua, "Attitude estimation for accelerated vehicles using GPS/INS measurements," *Control Engineering Practice*, vol. 18, no. 7, pp. 723–732, 2010.

[34] R. Kalman and J. Bertram, "Control system analysis and design via the second method of lyapunov: Ii discrete-time systems," *Journal of Basic Engineering*, vol. 82, no. 2, pp. 394–400, 1960.

[35] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice Hall, 2002.

[36] R. Horaud and F. Dornaika, "Hand-eye calibration," *The international journal of robotics research*, vol. 14, no. 3, pp. 195–210, 1995.

[37] M. Abramowitz and I. A. Stegun, *Handbook of mathematical functions: with formulas, graphs, and mathematical tables.* Courier Corporation, 1964, vol. 55.