

Article

# Long-Distance GNSS-Denied Visual Inertial Navigation for Autonomous Fixed-Wing Unmanned Air Vehicles: $SO(3)$ Manifold Filter Based on Virtual Vision Sensor

Eduardo Gallo \*  and Antonio Barrientos 

Centro de Automática y Robótica, Universidad Politécnica de Madrid—Consejo Superior de Investigaciones Científicas, 28006 Madrid, Spain; antonio.barrientos@upm.es

\* Correspondence: e.gallo@alumnos.upm.es or edugallo@yahoo.com; Tel.: +34-600643104

**Abstract:** This article proposes a visual inertial navigation algorithm intended to diminish the horizontal position drift experienced by autonomous fixed-wing UAVs (unmanned air vehicles) in the absence of GNSS (Global Navigation Satellite System) signals. In addition to accelerometers, gyroscopes, and magnetometers, the proposed navigation filter relies on the accurate incremental displacement outputs generated by a VO (visual odometry) system, denoted here as a virtual vision sensor, or VVS, which relies on images of the Earth surface taken by an onboard camera and is itself assisted by filter inertial estimations. Although not a full replacement for a GNSS receiver since its position observations are relative instead of absolute, the proposed system enables major reductions in the GNSS-denied attitude and position estimation errors. The filter is implemented in the manifold of rigid body rotations or  $SO(3)$  in order to minimize the accumulation of errors in the absence of absolute observations. Stochastic high-fidelity simulations of two representative scenarios involving the loss of GNSS signals are employed to evaluate the results. The authors release the C++ implementation of both the visual inertial navigation filter and the high-fidelity simulation as open-source software.

**Keywords:** GNSS-denied; visual inertial navigation; autonomous navigation; EKF; autonomy; VIO



**Citation:** Gallo, E.; Barrientos, A. Long-Distance GNSS-Denied Visual Inertial Navigation for Autonomous Fixed-Wing Unmanned Air Vehicles:  $SO(3)$  Manifold Filter Based on Virtual Vision Sensor. *Aerospace* **2023**, *10*, 708. <https://doi.org/10.3390/aerospace10080708>

Academic Editors: Hailong Huang, Gokhan Inalhan and Yan Xu

Received: 26 May 2023

Revised: 7 July 2023

Accepted: 4 August 2023

Published: 14 August 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction and Outline

The main objective of this article is to develop a navigation system capable of diminishing the position drift inherent to the flight in GNSS (Global Navigation Satellite System)-denied conditions of an autonomous fixed-wing aircraft, so in case GNSS signals become unavailable during flight, the vehicle has a higher probability of reaching the vicinity of a distant recovery point, from where it can be landed by remote control. The proposed approach combines two different navigation systems (inertial, which is based on the outputs of all onboard sensors except the camera, and visual, which relies exclusively on the images of the Earth surface taken from the aircraft) in such a way that both simultaneously assist and control each other, resulting in a positive feedback loop with major improvements in horizontal position estimation accuracy compared with either system by itself.

The fusion between the inertial and visual systems is a two-step process. The first one, described in [1], feeds the visual system with the inertial attitude and altitude estimations to assist with its nonlinear pose optimizations, resulting in major horizontal position accuracy gains with respect to both a purely visual system (without inertial assistance) and a standalone inertial one such as that described in [2]. The second step, which is the focus of this article, feeds back the visual horizontal position estimations into the inertial navigation filter as if they were the outputs of a virtual vision sensor, or VVS, replacing those of the GNSS receiver, and results in additional horizontal position accuracy improvements.

Section 2 describes the mathematical notation employed throughout the article. An introduction to GNSS-denied navigation and its challenges, together with a review of the state of the art in visual inertial navigation, is included in Section 3. Sections 4 and 5 discuss the article novelty and main applications, respectively. The proposed architecture to combine the inertial and visual navigation algorithms is discussed in Section 6, which relies on the VVS described in Section 7. The VVS outputs are employed by the navigation filter, which is described in detail in Section 8. Section 9 introduces the stochastic high-fidelity simulation employed to evaluate the navigation results by means of Monte Carlo executions of two scenarios representative of the challenges of GNSS-denied navigation. The results obtained when applying the proposed algorithms to these scenarios are described in Section 10, comparing them with those achieved by standalone inertial and visual systems. Last, Section 11 summarizes the results, while Appendix A describes the obtainment of various Jacobians employed in the proposed navigation filter.

### 2. Mathematical Notation

The meaning of all variables is specified on their first appearance. Any variable with a hat accent  $\langle \hat{\cdot} \rangle$  refers to its estimated value, with a tilde  $\langle \tilde{\cdot} \rangle$  to its measured value, and with a dot  $\langle \dot{\cdot} \rangle$  to its time derivative. In the case of vectors, which are displayed in bold (e.g.,  $\mathbf{x}$ ), other employed symbols include the skew-symmetric form  $\langle [\cdot]_{\times} \rangle$  and the double vertical bars  $\langle \|\cdot\| \rangle$ , which refer to the norm. The superindex T denotes the transpose of a vector or matrix. In the case of scalars, the vertical bars  $\langle |\cdot| \rangle$  refer to the absolute value. The left arrow  $\leftarrow$  represents an update operation, in which the value on the right of the arrow is assigned to the variable on the left.

Four different reference frames are employed in this article: the ECEF (Earth centered Earth fixed) frame  $F_E$  (centered at the Earth center of mass  $O_E$ , with  $\mathbf{i}_3^E$  pointing towards the geodetic north along the Earth rotation axis,  $\mathbf{i}_1^E$  contained in both the equator and zero longitude planes, and  $\mathbf{i}_2^E$  orthogonal to  $\mathbf{i}_1^E$  and  $\mathbf{i}_3^E$  forming a right-handed system), the NED (north east down) frame  $F_N$  (centered at the aircraft center of mass  $O_N$ , with axes aligned with the geodetic north, east, and down directions), the body frame  $F_B$  (centered at the aircraft center of mass  $O_B = O_N$ , with  $\mathbf{i}_1^B$  contained in the plane of symmetry of the aircraft pointing forward along a fixed direction,  $\mathbf{i}_3^B$  contained in the plane of symmetry of the aircraft, normal to  $\mathbf{i}_1^B$  and pointing downward, and  $\mathbf{i}_2^B$  orthogonal to both in such a way that they form a right-hand system), and the inertial frame  $F_I$  (centered at the Earth center of mass with axes that do not rotate with respect to any stars other than the Sun). The first three frames are graphically depicted in Figure 1.

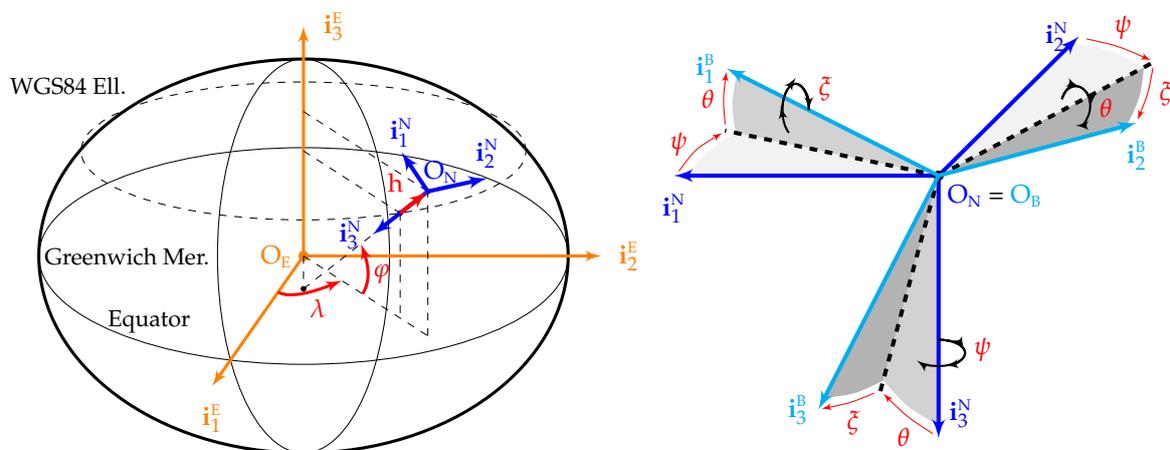


Figure 1. ECEF ( $F_E$ ), NED ( $F_N$ ), and body ( $F_B$ ) reference frames.

This article makes use of the space of rigid body rotations and, hence, relies on the Lie algebra of the special orthogonal group of  $\mathbb{R}^3$ , known as  $\mathbb{SO}(3)$ ; refs. [3–5] are recommended as references. Generic rotations are represented by  $\mathcal{R}$  and usually parameterized by the

unit quaternion  $\mathbf{q}$ , although the direct cosine matrix  $\mathbf{R}$  is also employed as required; tangent space representations include the rotation vector  $\mathbf{r}$  and the angular velocity  $\boldsymbol{\omega}$ . Related symbols include the quaternion conjugate  $\langle \cdot^* \rangle$ , the quaternion product  $\otimes$ , and the  $\mathbb{SO}(3)$  plus  $\oplus$  and minus  $\ominus$  operators.

Superindexes are employed over vectors to specify the frame in which they are viewed (e.g.,  $\mathbf{v}^N$  refers to ground velocity viewed in  $F_N$ , while  $\mathbf{v}^B$  is the same vector but viewed in  $F_B$ ). Subindexes may be employed to clarify the meaning of a variable or vector, but may also indicate a given vector component; e.g.,  $v_2^N$  refers to the second component of  $\mathbf{v}^N$ . In addition, where two reference frames appear as subindexes to a vector, it means that the vector goes from the first frame to the second. For example,  $\boldsymbol{\omega}_{NB}^B$  refers to the angular velocity from  $F_N$  to  $F_B$  viewed in  $F_B$ . Jacobians are represented by a  $\mathbf{J}$  combined with a subindex and a superindex; the former identifies the function domain, while the latter applies to the function image or codomain.

### 3. GNSS-Denied Navigation and Visual Inertial Odometry

The number, variety, and applications of unmanned air vehicles (UAVs) have grown exponentially in the last few years, and the trend is expected to continue in the future [6,7]. A comprehensive review of small UAV navigation systems and the problems they face is included in [8]. The use of GNSS constitutes one of the main enablers for autonomous inertial navigation [9–11], but it is also one of its main weaknesses, as in the absence of GNSS signals, inertial systems are forced to rely on dead reckoning, which results in position drift, with the aircraft slowly but steadily deviating from its intended route [2,12]. The availability of GNSS signals cannot be guaranteed; a thorough analysis of GNSS threats and reasons for signal degradation is presented in [13].

At this time, there are no comprehensive solutions to the operation of autonomous UAVs in GNSS-denied scenarios, for which the permanent loss of the GNSS signals is equivalent to losing the airframe in an uncontrolled way. A summary of the challenges of GNSS-denied navigation and the research efforts intended to improve its performance is provided by [2,14,15]. Two promising techniques for completely eliminating the position drift are the use of *signals of opportunity* [16–18] (existing signals originally intended for other purposes, such as those of television and cellular networks, can be employed to triangulate the aircraft position) and *georegistration* [19–27] (the position drift can be eliminated by matching landmarks or terrain features as viewed from the aircraft to preloaded data), also known as *image registration*.

Visual odometry (VO) consists in employing the ground images generated by one or more onboard cameras without the use of prerecorded image databases or any other sensors, incrementally estimating the vehicle pose (position plus attitude) based on the changes that its motion induces on the images [28–32]. It requires sufficient illumination, dominance of static scene, enough texture, and scene overlap between consecutive images or frames. Modern standalone algorithms, such as semidirect visual odometry (SVO) [33,34], direct sparse odometry (DSO) [35], large-scale direct simultaneous localization and mapping (LSD-SLAM) [36], and large-scale feature-based SLAM (ORB-SLAM) [37–39] (ORB stands for oriented FAST and rotated BRIEF, a type of blob feature), are robust and exhibit a limited drift. The incremental concatenation of relative poses results in a slow but unbounded pose drift that disqualifies VO for long-distance autonomous UAV GNSS-denied navigation.

A promising research path to GNSS-denied navigation appears to be the fusion of inertial and visual odometry algorithms into an integrated navigation system. Two different trends coexist in the literature, although an implementation of either group can sometimes employ a technique from the other:

- The first possibility, known as *filtering*, consists in employing the visual estimations as additional observations with which to feed the inertial filter. In the case of attitude estimation exclusively (momentarily neglecting the platform position), the GNSS signals are helpful and enable the inertial navigation system to obtain more accurate and less noisy estimations, but they are not indispensable, as driftless attitude estimations

can be obtained based on the inertial measurement unit (IMU) without their use [2]. Proven techniques for attitude estimation in all kinds of platforms that do not rely on GNSS signals include Gaussian filters [40], deterministic filters [41], complimentary filters [42], and stochastic filters [43,44]. In the case of the complete pose estimation (both attitude and position), it is indispensable to employ the velocity or incremental position observations obtained with VO methods in the absence of the absolute references provided by GNSS receivers. The literature includes cases based on Gaussian filters [45] and, more recently, nonlinear deterministic filters [46], stochastic filters [47], Riccati observers [48], and invariant extended Kalman filters, or EKF's [49].

- The alternative is to employ VO optimizations with the assistance of the inertial navigation estimations, reducing the pose estimation drift inherent to VO. This is known as visual inertial odometry (VIO) [50,51], which can also be combined with image registration to fully eliminate the remaining pose drift. A previous article by the authors, [1], describes how the nonlinear VO optimizations can be enhanced by adding priors based on the inertial attitude and altitude estimations. VIO has matured significantly in the last few years, with detailed reviews available in [50–54].

There exist several open-source VIO packages, such as the multistate constraint Kalman filter (MSCKF) [55], the open keyframe visual inertial SLAM (OKVIS) [56,57], the robust visual inertial odometry (ROVIO) [58], the monocular visual inertial navigation system (VINS-Mono) [59], SVO combined with multisensor fusion (MSF) [33,34,60,61], and SVO combined with incremental smoothing and mapping (iSAM) [33,34,62,63]. All these open-source pipelines are compared in [50], and their results, when applied to the EuRoC micro air vehicles (MAV) data sets, ref. [64] are discussed in [65]. There also exist various other published VIO pipelines with implementations that are not publicly available [66–72], and there are also others that remain fully proprietary.

The existing VIO schemes to fuse the visual and inertial measurements can be broadly grouped into two paradigms: *loosely coupled* pipelines process the measurements separately, resulting in independent visual and inertial pose estimations, which are then fused to obtain the final estimate [73,74]; on the other hand, *tightly coupled* methods compute the final pose estimation directly from the tracked image features and the IMU outputs [50,51]. Tightly coupled approaches usually result in higher accuracy, as they use all the information available and take advantage of the IMU integration to predict the feature locations in the next frame. Loosely coupled methods, although less complex and more computationally efficient, lose information by decoupling the visual and inertial constraints, and are incapable of correcting the drift present in the visual estimator.

A different classification involves the number of images involved in each estimation [50,51,75], which is directly related with the resulting accuracy and computing demands. *Batch algorithms*, also known as *smoothers*, estimate multiple states simultaneously by solving a large nonlinear optimization problem or bundle adjustment, resulting in the highest possible accuracy. Valid techniques to limit the required computing resources include the reliance on a subset of the available frames (known as *keyframes*), the separation of tracking and mapping into different threads, and the development of incremental smoothing techniques based on factor graphs [63]. Although employing all available states (*full smoothing*) is sometimes feasible for very short trajectories, most pipelines rely on *sliding window* or *fixed lag smoothing*, in which the optimization relies exclusively on the measurements associated with the last few keyframes, discarding both the old keyframes and other frames that have not been cataloged as keyframes. On the other hand, *filtering algorithms* restrict the estimation process to the latest state; they require less resources but suffer from permanently dropping all previous information and a much harder identification and removal of outliers, both of which lead to error accumulation or drift.

Additional VIO pipelines are also being developed to take advantage of *event cameras* [76,77], which measure per-pixel luminosity changes asynchronously instead of capturing images at a fixed rate. Event cameras hold very advantageous properties for visual odometry, including low latency, very high frequency, and excellent dynamic range, which

make them well suited to deal with high-speed motion and high dynamic range, which are problematic for traditional VO and VIO pipelines. On the negative side, new algorithms are required to cope with the sequence of asynchronous events they generate [50].

#### 4. Novelty

The main novelty of this article (second phase of the fusion between the inertial and visual navigation systems) lies in the following two topics:

- The use of a modified EKF scheme within the navigation filter (Section 8) based on Lie algebra to ensure that the estimated aircraft body attitude is propagated along its tangent space and never deviates from its  $\mathbb{SO}(3)$  manifold, reducing the error growth inherent to concatenating estimations over a long period of time.
- The transformation within the VVS (Section 7) of the position estimations obtained by the visual system into incremental displacement measurements, which are unbiased and hence more adequate to be used as observations within an EKF.

When comparing the integrated system (which includes both [1] and this article) with the state of the art discussed in Section 3, the proposed approach combines the high accuracy representative of tightly coupled smoothers with the lack of complexity and reduced computing requirements of filters and loosely coupled systems. The existence of two independent solutions (inertial and visual) hints to a loosely coupled approach, but in fact, the proposed solution shares most traits with tightly coupled pipelines:

- No information is discarded because the two solutions are not independent, as each simultaneously feeds and is fed by the other. The visual estimations depend on the inertial outputs, while the inertial filter uses the visual incremental displacement estimations as observations. Unlike loosely coupled solutions, the visual estimations for attitude and altitude are not allowed to deviate from the inertial ones above a certain threshold, so they do not drift.
- The two estimations are never fused together, as the inertial solution constitutes the final output, while the sole role of the visual outputs is to act as the previously mentioned incremental sensor.

With respect to the last criterion found in Section 3, the proposed approach contains both a filter within the inertial system and a keyframe-based sliding window smoother within the visual one, obtaining the best properties of both categories.

#### 5. Application

The visual inertial solutions listed in Section 3 are quite generic with respect to the platforms on which they are mounted, with most applications focused on the short distance trajectories of ground vehicles, indoor robots, and multirotors, as well as with respect to the employed sensors, which are usually restricted to the IMU and one or more cameras. The proposed approach differs in its more restricted application, as it focuses on the specific case of long-distance GNSS-denied turbulent flight of fixed-wing autonomous aircraft:

- Being restricted to aerial vehicles, it takes advantage of the extra sensors already present on board these platforms, such as magnetometers and barometer.
- The fixed-wing limitation is caused by the visual system relying on the pitot tube to continue navigating when overflying texture-poor terrain [1].
- As exemplified by the two scenarios employed to evaluate the algorithms, the proposed approach can cope with heavy turbulence, which increases the platform accelerations as well as the optical flow between consecutive images. Although not discussed in the results, the turbulence level has a negligible effect on the navigation accuracy as long as it remains below a sufficiently elevated threshold.
- With the exception of atmospheric pressure changes, which accumulate as vertical position estimation errors (Section 10.2), weather changes such as wind and atmospheric temperature have no effect on the navigation accuracy.

- It focuses more on GNSS-denied environments of a different nature than those experienced by other platforms, as it can always be assumed that GNSS signals are present at the beginning of the flight, and if they disappear, the reason is likely to be technical error or intentional action, so the vehicle needs to be capable of flying for long periods of time in GNSS-denied conditions.
- The reliance on visual navigation imposes certain restrictions, such as the need for sufficient illumination, lack of cloud cover below the aircraft, and impossibility to navigate over large bodies of water. The use of infrared cameras, although out of the scope of this article, is a promising research area for the first two restrictions, but the lack of static features makes visual systems inadequate for navigation over water.

### 6. Proposed Visual Inertial Navigation Architecture

The navigation system processes the measurements of the various onboard sensors ( $\tilde{\mathbf{x}} = \mathbf{x}_{\text{SENSED}}$ ) and generates the estimated state ( $\hat{\mathbf{x}} = \mathbf{x}_{\text{EST}}$ ) required by the guidance and control systems, which contains an estimation of the true aircraft state  $\mathbf{x} = \mathbf{x}_{\text{TRUTH}}$ . Note that in this article, the images  $\mathbf{I}$  are considered part of the sensed states  $\tilde{\mathbf{x}} = \mathbf{x}_{\text{SENSED}}$  as the camera is treated as an additional sensor. Hardware and computing power limitations often require the camera and visual algorithms to operate at a lower frequency than those of inertial algorithms and the remaining sensors. Table 1 includes the operating frequencies employed to generate the Section 10 results. Note that the index  $t$  identifies a discrete time instant ( $t_t$ ) for the aircraft trajectory,  $s$  ( $t_s$ ) refers to the sensor outputs,  $n$  ( $t_n$ ) identifies an inertial estimation,  $c$  ( $t_c$ ) is a control system execution,  $i$  ( $t_i$ ) applies to a camera image and the corresponding visual estimation, and  $g$  ( $t_g$ ) refers to a GNSS receiver observation.

Table 1. Working frequencies of the different systems.

Discrete Time	Frequency	Variable	System
$t_t = t \cdot \Delta t_{\text{TRUTH}}$	500 Hz	$\mathbf{x} = \mathbf{x}_{\text{TRUTH}}$	Flight physics
$t_s = s \cdot \Delta t_{\text{SENSED}}$	100 Hz	$\tilde{\mathbf{x}} = \mathbf{x}_{\text{SENSED}}$	Sensors
$t_n = n \cdot \Delta t_{\text{EST}}$	100 Hz	$\hat{\mathbf{x}} = \mathbf{x}_{\text{EST}}$	Inertial navigation
$t_c = c \cdot \Delta t_{\text{CNTR}}$	50 Hz	$\delta_{\text{TARGET}}, \delta_{\text{CNTR}}$	Guidance and control
$t_i = i \cdot \Delta t_{\text{IMG}}$	10 Hz	$\mathbf{x}_{\text{IMG}}$	Visual navigation and camera
$t_g = g \cdot \Delta t_{\text{GNSS}}$	1 Hz		GNSS receiver

The proposed strategy to fuse the inertial and visual navigation systems is based on their performances as standalone systems [1,2], as summarized in Table 2. Inertial systems are clearly superior to visual ones in GNSS-denied attitude and altitude estimation since their estimations are bounded and do not drift. In the case of horizontal position, both systems drift and are hence inappropriate for long-term GNSS-denied navigation, but for different reasons; while the inertial drift is the result of integrating the bounded ground velocity estimations without absolute position observations, the visual one originates on the slow but continuous accumulation of estimation errors between consecutive images.

Table 2. Summary of inertial and visual standalone navigation systems.

GNSS-Denied	Inertial	Visual
Attitude	Bounded by sensor quality Yaw worse than pitch and bank	Drifts Yaw better than pitch and bank
Vertical	Bounded by atmospheric physics	Drifts
Horizontal	Drifts	Drifts

It is also worth noting that inertial estimations are significantly noisier than visual ones, so the latter, although worse on an absolute basis, are often more accurate when evaluating the estimations incrementally, this is, from one image to the next or even for the time interval corresponding to a few images. When applied to the horizontal position, the

high accuracy of the incremental image-to-image displacement observations (equivalent to ground velocity observations) constitutes the basis for the second phase of the fusion described in the following sections.

This article proposes an inertial navigation filter (Section 8) that, in the absence of GNSS signals, takes advantage of the incremental displacement outputs generated by a visual system as if they were the observations generated by an additional sensor (Section 7), denoted as the virtual vision sensor, or VVS. The proposed EKF relies on the observations provided by the onboard accelerometers, gyroscopes, and magnetometers, combined with those of the GNSS receiver when signals are available, or those of the VVS when they are not. The inertial navigation filter and the visual system continuously assist and control each other, as depicted in Figure 2. As the integrated system needs to estimate the state  $\hat{\mathbf{x}} = \mathbf{x}_{EST}$  at every  $t_n = t_s = n \cdot \Delta t_{EST} = s \cdot \Delta t_{SENSED}$ , it has three different operation modes, depending not only on whether GNSS signals are available or not, but also on whether there exist a position and ground velocity observation at the execution time  $t_n = t_s$ :

- Given the operating frequencies of the different sensors listed in Table 1, most navigation filter cycles (99 out of 100 for GNSS-based navigation, 9 out of 10 in case of GNSS-denied conditions) can not rely on the position and ground velocity observations provided by either the GNSS receiver or the VVS. These cycles successively apply the EKF time update, measurement update, and reset equations (Section 8), noting that the filter observation system needs to be simplified by removing the ground velocity and position observations, which are not available.
- GNSS-denied filter cycles for which the VVS position and ground velocity observations are available at the required time ( $t_n = n \cdot \Delta t_{EST} = t_i = i \cdot \Delta t_{IMG}$ ) rely on the Figure 3 scheme. Note that this case only occurs in 1 out of 10 executions for GNSS-denied conditions. The first part of the process is exactly the same as in the previous case, estimating first the a priori state and covariance ( $\hat{\mathbf{x}}_n^-, \mathbf{P}_n^-$ ), followed by the a posteriori ones ( $\hat{\mathbf{x}}_{PRE,n}^+, \mathbf{P}_{PRE,n}^+$ ), and finally the estimated state ( $\hat{\mathbf{x}}_{PRE}(t_n) = \mathbf{x}_{EST,PRE}(t_n)$ ). Note that as these estimations do not employ the VVS measurements (that is, do not employ the  $I_i$  image corresponding to  $t_i$ ), they are preliminary and hence denoted with the subindex PRE to avoid confusion.

The preliminary estimation  $\hat{\mathbf{x}}_{PRE,n} = \mathbf{x}_{EST,PRE,n}$  together with the  $I_i$  image is then passed to the visual system [1] to obtain a visual state  $\mathbf{x}_{IMG}(t_i)$ , which, as explained in Section 7, is equivalent to the VVS position and ground velocity outputs. Only then it is possible to apply the complete navigation filter measurement update and reset equations for a second time to obtain the final state estimation  $\hat{\mathbf{x}}(t_n) = \mathbf{x}_{EST}(t_n)$ .

- GNSS-based filter cycles for which the GNSS receiver position and ground velocity observations are available at the required time ( $t_n = n \cdot \Delta t_{EST} = t_g = g \cdot \Delta t_{GNSS}$ ). This case, which only occurs in one out of a hundred executions when GNSS signals are available, is in fact very similar to the first one above, with the only difference being that the navigation filter observation system does not need any simplifications.

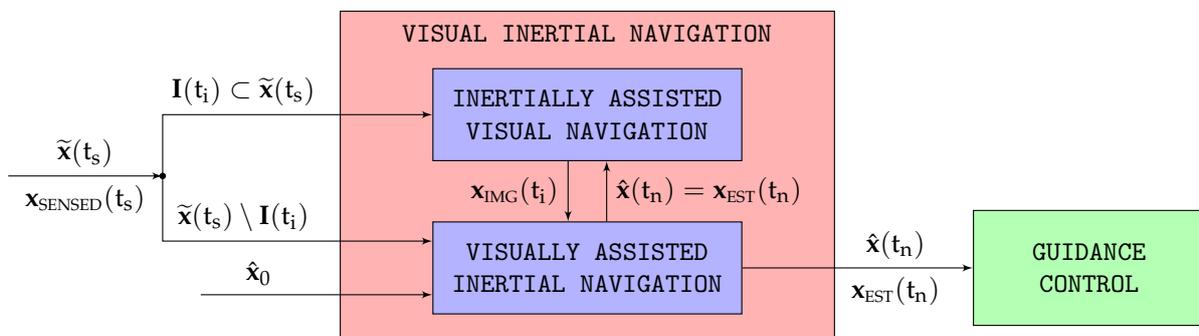


Figure 2. Proposed visual inertial navigation system flow diagram.

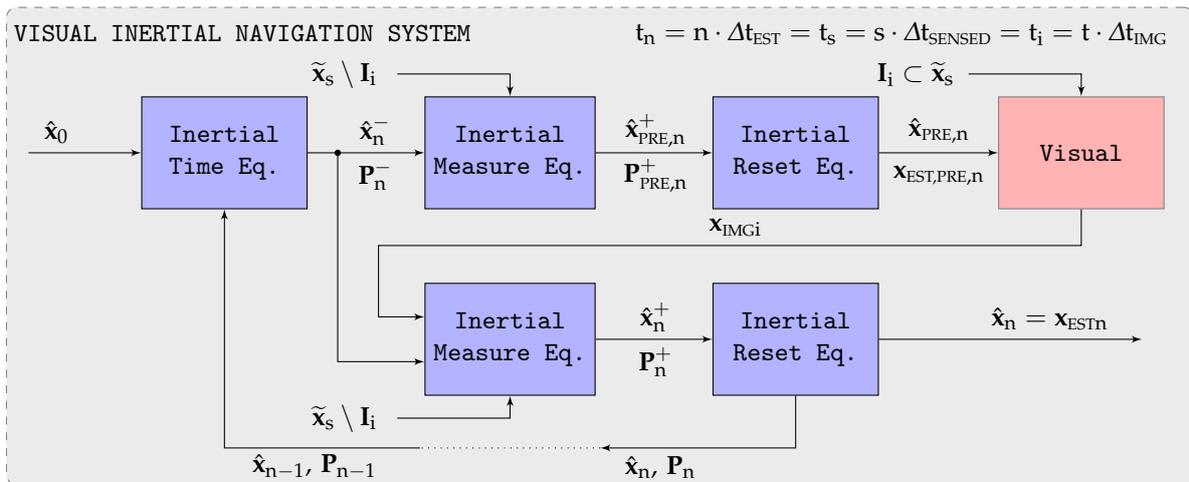


Figure 3. Navigation filter flow diagram when  $t_n = t_s = t_i$ .

### 7. Virtual Vision Sensor

The *virtual vision sensor* (VVS) constitutes an alternative denomination for the incremental displacement outputs obtained by the inertial assisted visual system described in [1]. Its horizontal position estimations drift with time due to the absence of absolute observations, but when evaluated on an incremental basis, that is, from one image to the next, the incremental displacement estimations are quite accurate. This section shows how to convert these incremental estimations into measurements for the geodetic coordinates ( $\hat{\mathbf{T}}^{E,GDT}$ ) and the  $F_N$  ground velocity ( $\hat{\mathbf{v}}^N$ ), so the VVS can smoothly replace the GNSS receiver within the navigation filter in the absence of GNSS signals (Section 8).

Their obtainment relies on the current visual state  $\mathbf{x}_{IMG}(t_i) = \mathbf{x}_{IMG,i}$  generated by the visual system, the one obtained with the previous frame  $\mathbf{x}_{IMG}(t_{i-1}) = \mathbf{x}_{IMG,i-1}$ , as well as the estimated state  $\hat{\mathbf{x}}(t_{n-\delta t}) = \mathbf{x}_{EST}(t_{n-\delta t}) = \hat{\mathbf{x}}_{n-\delta t}$  corresponding to the previous image generated by the inertial system. Note that as  $t = t_i = i \cdot \Delta t_{IMG} = t_n = n \cdot \Delta t_{EST}$ , the relationship between  $i$  and  $n$  is as follows, where  $\delta t = 10$  is the number of inertial executions for every image being processed (Table 1):

$$n = i \cdot \frac{\Delta t_{IMG}}{\Delta t_{EST}} = i \cdot \delta t \tag{1}$$

- To obtain the VVS velocity observations ( $\hat{\mathbf{v}}^N$ ), it is necessary to first compute the geodetic coordinate (longitude  $\lambda$ , latitude  $\varphi$ , altitude  $h$ ) time derivative based on the difference between their values corresponding to the last two images (2), followed by their transformation into  $\hat{\mathbf{v}}^N$  per (3), in which  $M$  and  $N$  represent the WGS84 radii of curvature of meridian and prime vertical, respectively. Note that  $\hat{\mathbf{v}}^N$  is very noisy given how it is obtained.

$$\hat{\mathbf{T}}_{IMG,i}^{E,GDT} = [\lambda_{IMG,i} \ \varphi_{IMG,i} \ h_{IMG,i}]^T = \frac{\mathbf{T}_{IMG,i}^{E,GDT} - \mathbf{T}_{IMG,i-1}^{E,GDT}}{\Delta t_{IMG}} \tag{2}$$

$$\hat{\mathbf{v}}^N = \begin{bmatrix} [M(\tilde{\varphi}_{n-\delta t}) + \tilde{h}_{n-\delta t}] \dot{\varphi}_{IMG,i} \\ [N(\tilde{\varphi}_{n-\delta t}) + \tilde{h}_{n-\delta t}] \cos \tilde{\varphi}_{n-\delta t} \dot{\lambda}_{IMG,i} \\ -\dot{h}_{IMG,i} \end{bmatrix} \tag{3}$$

- With respect to the VVS geodetic coordinates, the sensed longitude and latitude can be obtained per (5) and (6) based on propagating the previous inertial estimations (those corresponding to the time of the previous VVS reading) with their visually obtained time derivatives. To avoid drift, the geometric altitude  $\tilde{h}$  is estimated based on the

barometer observations assuming that the atmospheric pressure offset remains frozen from the time the GNSS signals are lost [2].

$$\tilde{\mathbf{T}}_n^{\text{E,GDT}} = \begin{bmatrix} \tilde{\lambda}_n & \tilde{\varphi}_n & \tilde{h}_n \end{bmatrix}^T \quad (4)$$

$$\tilde{\lambda}_n = \hat{\lambda}_{n-\delta t} + \dot{\lambda}_{\text{IMG},i} \Delta t_{\text{IMG}} \quad (5)$$

$$\tilde{\varphi}_n = \hat{\varphi}_{n-\delta t} + \dot{\varphi}_{\text{IMG},i} \Delta t_{\text{IMG}} \quad (6)$$

With respect to the covariances, the authors have assigned the VVS an ad hoc position standard deviation one order of magnitude lower than that employed for the GNSS receiver so the navigation filter (Section 8) closely tracks the position observations provided by the VVS. Given the noisy nature of the virtual velocity observations, the authors have preferred to employ a dynamic evaluation for the velocity standard deviation, which coincides with the absolute value (for each of the three dimensions) of the difference between each new velocity observation and the running average of the last 20 readings (equivalent to the last 2 s). In this way, when GNSS signals are available, the covariance of the position observations is higher, so the navigation filter slowly corrects the position estimations obtained from integrating the state equations instead of closely adhering to the noisy GNSS position measurements; when GNSS signals are not present, the VVS low position observation covariance combined with the high-velocity one implies that the navigation EKF (Section 8) closely tracks the VVS position readings (4), continuously adjusting the aircraft velocity estimations with little influence from the VVS velocity observations in order to complement its position estimations.

## 8. Proposed Navigation Filter

As graphically depicted in Figure 2, the navigation filter estimates the aircraft pose  $\hat{\mathbf{x}} = \mathbf{x}_{\text{EST}}$  based on the observations provided by the onboard sensors with the exception of the camera,  $\tilde{\mathbf{x}} \setminus \mathbf{I} = \mathbf{x}_{\text{SENSED}} \setminus \mathbf{I}$ , complemented if necessary (in the absence of GNSS signals) by the visual observations  $\mathbf{x}_{\text{IMG}}$  provided by the VVS. Implemented as an EKF, the objective of the navigation filter is the estimation of the aircraft attitude between the  $F_N$  and  $F_B$  frames represented by its unit quaternion  $\hat{\mathbf{q}}_{\text{NB}}$  [3] together with the aircraft position represented by its geodetic coordinates  $\hat{\mathbf{T}}^{\text{E,GDT}}$ . Additional state vector members include the aircraft angular velocity  $\hat{\omega}_{\text{NB}}^{\text{B}}$ ; the ground velocity  $\hat{\mathbf{v}}^{\text{N}}$ ; the specific force  $\hat{\mathbf{f}}_{\text{IB}}^{\text{B}}$  or nongravitational acceleration from  $F_I$  to  $F_B$ ; the full gyroscope, accelerometer, and magnetometer errors ( $\hat{\mathbf{E}}_{\text{GYR}}$ ,  $\hat{\mathbf{E}}_{\text{ACC}}$ ,  $\hat{\mathbf{E}}_{\text{MAG}}$ ) that include all sensor error sources except system noise; and the difference  $\hat{\mathbf{B}}_{\text{DEV}}^{\text{N}}$  between the Earth magnetic field provided by the onboard model  $\mathbf{B}_{\text{MOD}}^{\text{N}}$  and the real one  $\mathbf{B}_{\text{REAL}}^{\text{N}}$ .

$$\mathbf{x}(t) = \left[ \mathbf{q}_{\text{NB}} \quad \omega_{\text{NB}}^{\text{B}} \quad \mathbf{T}^{\text{E,GDT}} \quad \mathbf{v}^{\text{N}} \quad \mathbf{f}_{\text{IB}}^{\text{B}} \quad \mathbf{E}_{\text{GYR}} \quad \mathbf{E}_{\text{ACC}} \quad \mathbf{E}_{\text{MAG}} \quad \mathbf{B}_{\text{DEV}}^{\text{N}} \right]^T \quad (7)$$

Rigid body rotations are characterized by their orthogonality (they preserve the norm of vectors and the angle between vectors) and their handedness (they preserve the orientation or cross product between vectors); when encoded as quaternions, these two constraints converge into a single one, the unitary norm, as only unit quaternions are valid parameterizations of rigid body rotations [3–5]. As the body attitude  $\mathbf{q}_{\text{NB}}$  belongs to the non-Euclidean *special orthogonal group* or  $\mathbb{S}\mathbb{O}(3)$ , the classical EKF scheme [78], which relies on linear algebra, shall be replaced by the alternative formulation proposed here to ensure that the estimated attitude never deviates from its  $\mathbb{S}\mathbb{O}(3)$  manifold, that is, to ensure that  $\mathbf{q}_{\text{NB}}$  remains unitary at all times [79]. Relying on linear algebra implies treating the unit quaternion  $\mathbf{q}_{\text{NB}}$  as a 4-vector in both the time update and measurement update equations of each EKF cycle, and then normalizing the resulting nonunit quaternion at the end of each cycle; although this is an established practice that provides valid results when operating at the high frequencies typical of inertial sensors, the slower frequencies characteristic of visual pipelines (refer to  $\Delta t_{\text{SENSED}}$  and  $\Delta t_{\text{IMG}}$  in Table 1) when combined with the lack of absolute position observations in the absence of GNSS signals result in

slight but non-negligible inaccuracies in each cycle, which can accumulate into significant errors when the aircraft needs to navigate without absolute position observations for long periods of time.

The use of Lie theory, with its manifolds and tangent spaces, enables the construction of rigorous calculus techniques to handle uncertainties, derivatives, and integrals of non-Euclidean elements with precision and ease [4]. Lie theory can hence be employed to generate an alternative EKF scheme suitable for state vectors containing non-Euclidean components [5]. This alternative EKF scheme differs in four key points from the classical Euclidean one; these four differences are discussed in the following paragraphs. Note that the GNSS-denied navigation results discussed in Section 10 cannot be obtained without the use of Lie algebra and the alternative EKF formulation.

The most rigorous and precise way to adapt the EKF scheme to the presence of non-Euclidean elements, such as rigid body rotations, is to exclude the rotation element  $\mathbf{q}_{NB} \in \mathbb{SO}(3)$  from the state vector, replacing the unit quaternion  $\mathbf{q}_{NB}$  by a local ( $F_B$ ) tangent space perturbation represented by the rotation vector  $\Delta\mathbf{r}_{NB}^B \in \mathbb{R}^3$  [4,5]. Note that the state vector (9) includes the velocity of  $\mathbf{q}_{NB}$  as it moves along its manifold (its angular velocity  $\boldsymbol{\omega}_{NB}^B$ ), also contained in the local ( $F_B$ ) tangent space. Each filter step now consists in estimating the rotation element  $\hat{\mathbf{q}}_{NBn} = \hat{\mathbf{q}}_{NB}(t_n) \in \mathbb{SO}(3)$ , the state vector  $\hat{\mathbf{x}}_n = \hat{\mathbf{x}}(t_n) = [\Delta\mathbf{r}_{NBn}^B \ \boldsymbol{\omega}_{NBn}^B \ \hat{\mathbf{z}}_n]^T \in \mathbb{R}^{27}$ , and its covariance  $\mathbf{P}_n = \mathbf{P}(t_n) \in \mathbb{R}^{27 \times 27}$ , based on their values at  $t_{n-1} = (n - 1) \Delta t$  and the observations at  $t_n = n \Delta t$ .

$$\mathcal{R}(t) = \mathbf{q}_{NB} \in \mathbb{SO}(3)(t) \tag{8}$$

$$\begin{aligned} \mathbf{x}(t) &= [\Delta\mathbf{r}_{NB}^B \ \boldsymbol{\omega}_{NB}^B \ \mathbf{z}]^T = [\Delta\mathbf{r}_{NB}^B \ \mathbf{p}]^T \\ &= [\Delta\mathbf{r}_{NB}^B \ \boldsymbol{\omega}_{NB}^B \ \mathbf{T}^{E,GDT} \ \mathbf{v}^N \ \mathbf{f}_{IB}^B \ \mathbf{E}_{GYR} \ \mathbf{E}_{ACC} \ \mathbf{E}_{MAG} \ \mathbf{B}_{DEV}^N]^T \in \mathbb{R}^{27}(t) \end{aligned} \tag{9}$$

The observations  $\mathbf{y}_n$  are provided by the gyroscopes, which measure the inertial angular velocity  $\tilde{\boldsymbol{\omega}}_{IB}^B$ , the accelerometers that provide the specific force  $\tilde{\mathbf{f}}_{IB}^B$  or nongravitational acceleration, the magnetometers that measure the magnetic field  $\tilde{\mathbf{B}}^B$ , and either the GNSS receiver or the VVS (Section 7) that provides the position  $\tilde{\mathbf{T}}^{E,GDT}$  (geodetic coordinates) and ground velocity  $\tilde{\mathbf{v}}^N$  observations. Note that the measurements are provided at different frequencies, as listed in Table 1; most are available every  $\Delta t_{SENSED} = 0.01$  s, but the ground velocity and position measurements are generated every  $\Delta t_{GNSS} = 1$  s if provided by the GNSS receiver or every  $\Delta t_{IMG} = 0.1$  s when the GNSS signals are not available and the readings are instead supplied by the VVS.

$$\mathbf{y}_n = [\tilde{\boldsymbol{\omega}}_{IB}^B \ \tilde{\mathbf{f}}_{IB}^B \ \tilde{\mathbf{B}}^B \ \tilde{\mathbf{T}}^{E,GDT} \ \tilde{\mathbf{v}}^N]^T \in \mathbb{R}^{15} \tag{10}$$

### 8.1. Time Update Equations

The variation with time of the state variables (11) comprises a continuous time non-linear state system composed of three exact equations (with no simplifications), (12), (14), and (15), together with six other differential Equations (13) and (16) in which the filter does not possess any knowledge about the time derivatives of the state variables, which are hence set to zero. The three exact equations are the following:

- The rotational motion time derivative (12) [4,5];
- The geodetic coordinates kinematics (14);
- The aircraft velocity dynamics (15) [2].

$$\dot{\mathbf{x}}(t) = [\Delta\dot{\mathbf{r}}_{NB}^B \ \dot{\boldsymbol{\omega}}_{NB}^B \ \dot{\mathbf{z}}]^T = \mathbf{f}(\mathbf{q}_{NB}(t) \oplus \Delta\mathbf{r}_{NB}^B(t), \boldsymbol{\omega}_{NB}^B(t), \mathbf{z}(t), \mathbf{w}(t), t) \tag{11}$$

$$\dot{\mathbf{x}}_{1-3} = \Delta\dot{\mathbf{r}}_{NB}^B = \boldsymbol{\omega}_{NB}^B = \mathbf{x}_{4-6} \tag{12}$$

$$\dot{\mathbf{x}}_{4-6} = \dot{\boldsymbol{\omega}}_{NB}^B = \mathbf{0}_3 \tag{13}$$

$$\dot{\mathbf{x}}_{7-9} = \dot{\mathbf{T}}^{E,GDT} = \begin{bmatrix} v_2^N & v_1^N & -v_3^N \\ [N(\varphi) + h] \cos \varphi & M(\varphi) + h & -v_3^N \end{bmatrix}^T \tag{14}$$

$$\dot{\mathbf{x}}_{10-12} = \dot{\mathbf{v}}^N = (\mathbf{q}_{NB} \oplus \Delta \mathbf{r}_{NB}^B) \otimes \mathbf{f}_{IB}^B \otimes (\mathbf{q}_{NB} \oplus \Delta \mathbf{r}_{NB}^B)^* - [\boldsymbol{\omega}_{EN}^N]_{\times} \mathbf{v}^N + \mathbf{g}_{C,MOD}^N - \mathbf{a}_{COR}^N \quad (15)$$

$$\dot{\mathbf{x}}_{13-27} = \begin{bmatrix} \mathbf{f}_{IB}^B & \dot{\mathbf{E}}_{GYR} & \dot{\mathbf{E}}_{ACC} & \dot{\mathbf{E}}_{MAG} & \dot{\mathbf{B}}_{DEV}^N \end{bmatrix}^T = \mathbf{0}_{15} \quad (16)$$

Note that in expression (15),  $\boldsymbol{\omega}_{EN}^N$  represents the  $F_N$  motion angular velocity experienced by any object that moves without modifying its attitude with respect to the Earth surface,  $\mathbf{g}_{C,MOD}^N$  constitutes the gravity acceleration modeled by the navigation system, and  $\mathbf{a}_{COR}^N$  represents the Coriolis acceleration also viewed in  $F_N$ . In expression (11),  $\mathbf{w}$  is the process noise that, although present, is not shown in Equations (12) through (16).

It is important to remark that this alternative EKF implementation, which removes the rotation element  $\mathbf{q}_{NB}$  from the state vector and instead relies on the local tangent space perturbation  $\Delta \mathbf{r}_{NB}^B$ , avoids the use of (17), whose integration in the classical EKF scheme considers the unit quaternion as a Euclidean 4-vector without respecting its  $\mathbb{S}\mathbb{O}(3)$  unitary constraint, requiring a normalization after each filter step. Avoiding these normalizations and their associated errors constitutes the first significant advantage of employing the alternative EKF scheme presented in this article.

$$\dot{\mathbf{q}}_{NB} = \frac{1}{2} \mathbf{q}_{NB} \otimes \boldsymbol{\omega}_{NB}^B \quad (17)$$

The linearization of the (11) system results in (18), which makes use of the system matrix  $\mathbf{A}(t) = \partial \mathbf{f} / \partial \mathbf{x}(\mathbf{w} = \mathbf{0}) \in \mathbb{R}^{27 \times 27}$ , where the obtainment of the various Jacobians  $\mathbf{J}$  is described in Appendix A. Those components of  $\mathbf{A}$  not individually listed in (19) through (23) are zero. Note that the linearization neglects the influence of the geodetic coordinates  $\mathbf{T}^{E,GDT}$  on four different terms ( $\dot{\mathbf{T}}^{E,GDT}$ ,  $\boldsymbol{\omega}_{EN}^N$ ,  $\mathbf{g}_{C,MOD}^N$ , and  $\mathbf{a}_{COR}^N$ ) present in the (11) differential equations, but includes the influence of the aircraft velocity  $\mathbf{v}^N$  on three of them (gravity does not depend on velocity). The reason is that the  $\mathbf{v}^N$  variation amount that can be experienced in a single state estimation step can have a significant influence on the values of the variables considered, while that of the geodetic coordinates  $\mathbf{T}^{E,GDT}$  does not and hence can be neglected.

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t) \mathbf{x}(t) + \tilde{\mathbf{w}}(t) \quad (18)$$

$$\mathbf{A}_{1-3,4-6} = \frac{d\Delta \mathbf{r}_{NB}^B}{d\boldsymbol{\omega}_{NB}^B} = \mathbf{I}_{3 \times 3} \quad (19)$$

$$\mathbf{A}_{7-9,10-12} = \frac{d\dot{\mathbf{T}}^{E,GDT}}{d\mathbf{v}^N} = \mathbf{J}_{\mathbf{v}^N} \dot{\mathbf{T}}^{E,GDT} \quad (20)$$

$$\mathbf{A}_{10-12,1-3} = \frac{d\dot{\mathbf{v}}^N}{d\Delta \mathbf{r}_{NB}^B} = \mathbf{J}_{\mathcal{R}_{NB}} \mathbf{q}_{NB} \otimes \mathbf{f}_{IB}^B \otimes \mathbf{q}_{NB}^* \quad (21)$$

$$\mathbf{A}_{10-12,10-12} = \frac{d\dot{\mathbf{v}}^N}{d\mathbf{v}^N} = -[\boldsymbol{\omega}_{EN}^N]_{\times} + [\mathbf{v}^N]_{\times} \mathbf{J}_{\mathbf{v}^N} \boldsymbol{\omega}_{EN}^N - \mathbf{J}_{\mathbf{v}^N} \mathbf{a}_{COR}^N \quad (22)$$

$$\mathbf{A}_{10-12,13-15} = \frac{d\dot{\mathbf{v}}^N}{d\mathbf{f}_{IB}^B} = \mathbf{J}_{\mathbf{f}_{IB}^B} (\mathbf{q}_{NB} \oplus \Delta \mathbf{r}_{NB}^B) \otimes \mathbf{f}_{IB}^B \otimes (\mathbf{q}_{NB} \oplus \Delta \mathbf{r}_{NB}^B)^* \quad (23)$$

The use of Lie Jacobians within (21) and (23), as well as in (31), (33), and (37), (39) below, constitutes the second difference between the proposed EKF scheme and the Euclidean one (refer to Appendix A for details on how these Jacobians are obtained). Although not intuitive, Lie Jacobians adhere to the concept of matrix derivative (Jacobian) as the infinitesimal variations in their formulation are contained within the corresponding tangent spaces and do not deviate from them [4,5]. In contrast, employing the classical EKF scheme with  $\mathbf{q}_{NB}$  as part of the state vector implies establishing independent partial derivatives with respect to each of the four quaternion coefficients, which lack any physical meaning in the absence of the unitary constraint that links their values together.

### 8.2. Measurement Update Equations

The discrete time nonlinear observation system (24) contains the observations provided by the different sensors, without any simplifications.  $\omega_{IE}^N$  represents the Earth angular velocity (viewed in  $F_N$ ) caused by its rotation around the  $i_3^E$  axis at a constant rate  $\omega_E$ , and  $\mathbf{v}_n$  represents the measurement or observation noise. Note that although present, measurement noise  $\mathbf{v}_n$  is not shown in Equations (25) through (29).

$$\mathbf{y}_n = \mathbf{h}(\mathbf{q}_{NBn} \oplus \Delta \mathbf{r}_{NBn}^B, \omega_{NBn}^B, \mathbf{z}_n, \mathbf{v}_n, t_n) \quad (24)$$

$$\tilde{\omega}_{IB}^B = \omega_{NB}^B + (\mathbf{q}_{NB} \oplus \Delta \mathbf{r}_{NB}^B)^* \otimes (\omega_{IE}^N + \omega_{EN}^N) \otimes (\mathbf{q}_{NB} \oplus \Delta \mathbf{r}_{NB}^B) + \mathbf{E}_{GYR} \quad (25)$$

$$\tilde{\mathbf{f}}_{IB}^B = \mathbf{f}_{IB}^B + \mathbf{E}_{ACC} \quad (26)$$

$$\tilde{\mathbf{B}}^B = (\mathbf{q}_{NB} \oplus \Delta \mathbf{r}_{NB}^B)^* \otimes (\mathbf{B}_{MOD}^N - \mathbf{B}_{DEV}^N) \otimes (\mathbf{q}_{NB} \oplus \Delta \mathbf{r}_{NB}^B) + \mathbf{E}_{MAG} \quad (27)$$

$$\tilde{\mathbf{T}}^{E,GDT} = \mathbf{T}^{E,GDT} \quad (28)$$

$$\tilde{\mathbf{v}}^N = \mathbf{v}^N \quad (29)$$

Its linearization results in the following output matrix  $\mathbf{H}_n = \partial \mathbf{h} / \partial \mathbf{x}_n (\mathbf{v}_n = \mathbf{0}) \in \mathbb{R}^{15 \times 27}$ , in which Appendix A describes the obtainment of the various Jacobians  $\mathbf{J}$  present in (31), (33), (37), and (39). Those components of  $\mathbf{H}$  not individually listed in (31) through (41) are zero. As in the state system case above, the linearization also neglects the influence of the geodetic coordinates  $\mathbf{T}^{E,GDT}$  on three different terms ( $\omega_{IE}^N$ ,  $\omega_{EN}^N$ , and  $\mathbf{B}_{MOD}^N$ ) present in the (24) observation equations.

$$\mathbf{y}_n \approx \mathbf{H}_n \mathbf{x}_n + \tilde{\mathbf{v}}_n \quad (30)$$

$$\mathbf{H}_{1-3,1-3} = \frac{d\tilde{\omega}_{IB}^B}{d\Delta \mathbf{r}_{NB}^B} = \mathbf{J} \mathbf{q}_{NB}^* \otimes (\omega_{IE}^N + \omega_{EN}^N) \otimes \mathbf{q}_{NB} \quad (31)$$

$$\mathbf{H}_{1-3,4-6} = \frac{d\tilde{\omega}_{IB}^B}{d\omega_{NB}^B} = \mathbf{I}_{3 \times 3} \quad (32)$$

$$\mathbf{H}_{1-3,10-12} = \frac{d\tilde{\omega}_{IB}^B}{d\mathbf{v}^N} = \mathbf{J} (\mathbf{q}_{NB} \oplus \Delta \mathbf{r}_{NB}^B)^* \otimes \omega_{EN}^N \otimes (\mathbf{q}_{NB} \oplus \Delta \mathbf{r}_{NB}^B) \mathbf{J} \omega_{EN}^N \quad (33)$$

$$\mathbf{H}_{1-3,16-18} = \frac{d\tilde{\omega}_{IB}^B}{d\mathbf{E}_{GYR}} = \mathbf{I}_{3 \times 3} \quad (34)$$

$$\mathbf{H}_{4-6,13-15} = \frac{d\tilde{\mathbf{f}}_{IB}^B}{d\mathbf{f}_{IB}^B} = \mathbf{I}_{3 \times 3} \quad (35)$$

$$\mathbf{H}_{4-6,19-21} = \frac{d\tilde{\mathbf{f}}_{IB}^B}{d\mathbf{E}_{ACC}} = \mathbf{I}_{3 \times 3} \quad (36)$$

$$\mathbf{H}_{7-9,1-3} = \frac{d\tilde{\mathbf{B}}^B}{d\Delta \mathbf{r}_{NB}^B} = \mathbf{J} \mathbf{q}_{NB}^* \otimes (\mathbf{B}_{MOD}^N - \mathbf{B}_{DEV}^N) \otimes \mathbf{q}_{NB} \quad (37)$$

$$\mathbf{H}_{7-9,22-24} = \frac{d\tilde{\mathbf{B}}^B}{d\mathbf{E}_{MAG}} = \mathbf{I}_{3 \times 3} \quad (38)$$

$$\mathbf{H}_{7-9,25-27} = \frac{d\tilde{\mathbf{B}}^B}{d\mathbf{B}_{DEV}^N} = -\mathbf{J} (\mathbf{q}_{NB} \oplus \Delta \mathbf{r}_{NB}^B)^* \otimes \mathbf{B}_{DEV}^N \otimes (\mathbf{q}_{NB} \oplus \Delta \mathbf{r}_{NB}^B) \quad (39)$$

$$\mathbf{H}_{10-12,7-9} = \frac{d\tilde{\mathbf{T}}^{E,GDT}}{d\mathbf{T}^{E,GDT}} = \mathbf{I}_{3 \times 3} \quad (40)$$

$$\mathbf{H}_{13-15,10-12} = \frac{d\tilde{\mathbf{v}}^N}{d\mathbf{v}^N} = \mathbf{I}_{3 \times 3} \quad (41)$$

### 8.3. Covariances

Based on the above linearized continuous state system (18) and discrete observations (30), the time and measurement update equations within each EKF cycle (there are no differences here between the proposed and classical EKF schemes) [5,78] result in estimations for both

the state vector  $\hat{\mathbf{x}}_n = [\Delta \hat{\mathbf{r}}_{\text{NBn}}^{\text{B}} \quad \hat{\boldsymbol{\omega}}_{\text{NBn}}^{\text{B}} \quad \hat{\mathbf{z}}_n]^T$  as well as its covariance  $\mathbf{P}_n$ . Note that the definition of the covariance matrix in (42) combines that of its Euclidean components (46) with its local Lie counterparts (43) [5], with additional combined members. The mean or expect value is represented by  $\mu$ .

$$\mathbf{P}_n = \begin{bmatrix} \mathbf{C}_{\mathcal{R}\mathcal{R},n}^{\mathcal{R}} & \mathbf{C}_{\mathcal{R}p,n}^{\mathcal{R}} \\ \mathbf{C}_{p\mathcal{R},n}^{\mathcal{R}} & \mathbf{C}_{pp,n}^{\mathcal{R}} \end{bmatrix} \in \mathbb{R}^{27 \times 27} \tag{42}$$

$$\mathbf{C}_{\mathcal{R}\mathcal{R},n}^{\mathcal{R}} = \mathbb{E} [\Delta \mathbf{r}_{\text{NBn}}^{\text{B}} \Delta \mathbf{r}_{\text{NBn}}^{\text{B},T}] = \mathbb{E} \left[ \left( \mathcal{R}_{\text{NBn}} \ominus \mu_{\mathcal{R},\text{NBn}} \right) \left( \mathcal{R}_{\text{NBn}} \ominus \mu_{\mathcal{R},\text{NBn}} \right)^T \right] \in \mathbb{R}^{3 \times 3} \tag{43}$$

$$\mathbf{C}_{\mathcal{R}p,n}^{\mathcal{R}} = \mathbb{E} \left[ \Delta \mathbf{r}_{\text{NBn}}^{\text{B}} \left( \mathbf{p}_n - \mu_{p,n} \right)^T \right] = \mathbb{E} \left[ \left( \mathcal{R}_{\text{NBn}} \ominus \mu_{\mathcal{R},\text{NBn}} \right) \left( \mathbf{p}_n - \mu_{p,n} \right)^T \right] \in \mathbb{R}^{3 \times 24} \tag{44}$$

$$\mathbf{C}_{p\mathcal{R},n}^{\mathcal{R}} = \mathbb{E} \left[ \left( \mathbf{p}_n - \mu_{p,n} \right) \Delta \mathbf{r}_{\text{NBn}}^{\text{B},T} \right] = \mathbb{E} \left[ \left( \mathbf{p}_n - \mu_{p,n} \right) \left( \mathcal{R}_{\text{NBn}} \ominus \mu_{\mathcal{R},\text{NBn}} \right)^T \right] \in \mathbb{R}^{24 \times 3} \tag{45}$$

$$\mathbf{C}_{pp,n}^{\mathcal{R}} = \mathbb{E} \left[ \left( \mathbf{p}_n - \mu_{p,n} \right) \left( \mathbf{p}_n - \mu_{p,n} \right)^T \right] \in \mathbb{R}^{24 \times 24} \tag{46}$$

Uncertainty around a given  $\mathbb{S}\mathbb{O}(3)$  attitude is properly expressed as a covariance on the tangent space [4]. The definition of  $\mathbf{C}_{\mathcal{R}\mathcal{R},n}^{\mathcal{R}}$  in (43) represents the third advantage of the proposed EKF scheme, as it implements the concept of covariance but applied to  $\mathbb{S}\mathbb{O}(3)$  attitudes instead of Euclidean vectors, as in (46). Its size is  $\mathbb{R}^{3 \times 3}$ , as rotations contain three degrees of freedom, and each coefficient contains the covariance when rotating about a specific axis viewed in the local ( $F_B$ ) tangent space; in contrast, introducing the  $\mathbf{q}_{\text{NB}}$  unit quaternion in the state vector demands an  $\mathbb{R}^{4 \times 4}$  covariance matrix in which the coefficients do not have any physical meaning. Note that if required, it is possible to obtain the covariance for each of the three Euler angles (yaw, pitch, roll) by rotating  $\mathbf{C}_{\mathcal{R}\mathcal{R},n}^{\mathcal{R}}$  to view it in frames other than  $F_B$  [5].

#### 8.4. Reset Equations

The fourth and final difference between the suggested EKF scheme and the Euclidean one is that after each filter estimation cycle, it is necessary to reset the estimated tangent space perturbation  $\Delta \hat{\mathbf{r}}_{\text{NBn}}^{\text{B}}$  to zero while updating the estimations for the rotation element  $\hat{\mathbf{q}}_{\text{NBn}}$  and the error covariance  $\mathbf{P}_n$  accordingly; the estimations for angular velocity  $\hat{\boldsymbol{\omega}}_{\text{NBn}}^{\text{B}}$  and the Euclidean components  $\hat{\mathbf{z}}_n$  are not modified. Note that the accuracy of the linearizations required to obtain the  $\mathbf{A}$  and  $\mathbf{H}$  system and output matrices is based on first-order Taylor expansions, which are directly related to the size of the tangent space perturbations. Although it is not strictly necessary to reset the perturbations in every EKF cycle, the accuracy of the whole state estimation process is improved by maintaining the perturbations as small as possible, so it is recommended to never bypass the reset step.

Taking into account that the rotation element is going to be updated per (49), the error covariance is propagated to the new rotation element as follows [5], where the Lie Jacobian is provided by (A7) within Appendix A:

$$\mathbf{P}_n \longleftarrow \mathbf{D} \mathbf{P}_n \mathbf{D}^T \tag{47}$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{J}_{\mathcal{R}_{\text{NB}}} \hat{\mathbf{q}}_{\text{NBn}} \oplus \Delta \hat{\mathbf{r}}_{\text{NBn}}^{\text{B}} & \mathbf{0}_{3 \times 24} \\ \mathbf{0}_{24 \times 3} & \mathbf{I}_{24 \times 24} \end{bmatrix} \in \mathbb{R}^{27 \times 27} \tag{48}$$

Last, the rotation element is propagated by means of the local rotation vector perturbation, which is itself reset to zero:

$$\hat{\mathbf{q}}_{\text{NBn}} \longleftarrow \hat{\mathbf{q}}_{\text{NBn}} \oplus \Delta \hat{\mathbf{r}}_{\text{NBn}}^{\text{B}} \tag{49}$$

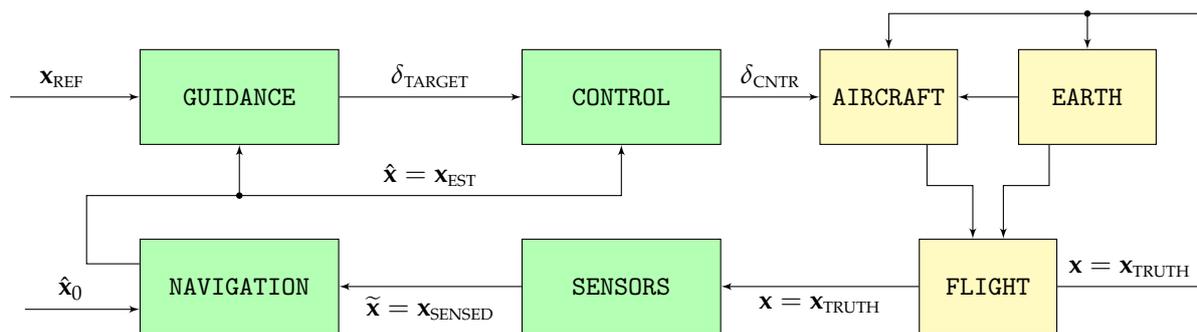
$$\Delta \hat{\mathbf{r}}_{\text{NBn}}^{\text{B}} \longleftarrow \mathbf{0}_3 \tag{50}$$

Different position (geodetic coordinates) and ground velocity system noise values are employed depending on whether the measurements are provided by the GNSS receiver or the VVS:

- Coupled with the observation covariances discussed in Section 7, lower geodetic coordinate system noise values are employed when GNSS signals are available, as the objective is for the solution to avoid position jumps by smoothly following the state equations and only slightly updating the position based on the GNSS observations to avoid any position drift in the long term.
- When the position observations are instead supplied by the VVS, higher system noise values are employed so the EKF relies more on the observations and less on the integration. Note that the VVS velocity observations are very noisy because of (2), so it is better if the EKF closely adheres to the position observations that originate at the visual system, correcting in each step as necessary.

### 9. Testing: High-Fidelity Simulation and Scenarios

To evaluate the performance of the proposed navigation algorithms, this article relies on Monte Carlo simulations consisting of one hundred runs each of two different scenarios based on the high-fidelity stochastic flight simulator graphically depicted in Figure 4. The simulator, whose open source C++ implementation is available in [80], models the flight of a fixed-wing piston engine autonomous UAV.



**Figure 4.** Components of the high-fidelity simulation.

The simulator consists of two distinct processes:

- The first, represented by the yellow blocks on the right of Figure 4, models the physics of flight and the interaction between the aircraft and its surroundings, which results in the real aircraft trajectory  $\mathbf{x} = \mathbf{x}_{\text{TRUTH}}$ .
- The second, represented by the green blocks on the left, contains the aircraft systems in charge of ensuring that the resulting trajectory adheres as much as possible to the mission objectives. It includes the different sensors whose output comprises the sensed trajectory  $\tilde{\mathbf{x}} = \mathbf{x}_{\text{SENSED}}$ , the navigation system in charge of filtering it to obtain the estimated trajectory  $\hat{\mathbf{x}} = \mathbf{x}_{\text{EST}}$ , the guidance system that converts the reference objectives  $\mathbf{x}_{\text{REF}}$  into the control targets  $\delta_{\text{TARGET}}$ , and the control system that adjusts the position of the throttle and aerodynamic control surfaces  $\delta_{\text{CNTR}}$  so the estimated trajectory  $\hat{\mathbf{x}}$  is as close as possible to the reference objectives  $\mathbf{x}_{\text{REF}}$ . Table 1 lists the working frequencies of the various blocks represented in Figure 4.

All components of the flight simulator have been modeled with as few simplifications as much as possible to increase the realism of the results. With the exception of the aircraft performances and its control system, which are deterministic, all other simulator components are treated as stochastic and hence vary from one execution to the next, enhancing the significance of the Monte Carlo simulation results.

Most VIO packages discussed in Section 3 include in their release articles an evaluation when applied to the EuRoC MAV data sets [64], and so do independent evaluations, such

as [65]. These data sets contain perfectly synchronized stereo images, IMU measurements, and ground truth readings obtained with laser, for eleven different indoor trajectories flown with a MAV, each with a duration in the order of 2 min and a total distance in the order of 100 m. This fact by itself indicates that the target application of exiting VIO implementations differs significantly from the main focus of this article, as there may exist accumulating errors that are completely nondiscernible after such short periods of time, but that grow nonlinearly and have the capability of inducing significant pose errors when the aircraft remains aloft for long periods of time. The algorithms presented in this article are hence tested through simulation under two different scenarios designed to analyze the consequences of losing the GNSS signals for long periods of time; these two scenarios coincide with those employed in [1,2] to evaluate standalone visual and inertial algorithms. Most parameters comprising the scenarios are defined stochastically, resulting in different values for every execution. Note that all results shown in Section 10 are based on Monte Carlo simulations comprising one hundred runs of each scenario, testing the sensitivity of the proposed navigation algorithms to a wide variety of values in the parameters.

- Scenario #1 has been defined with the objective of adequately representing the challenges faced by an autonomous fixed-wing UAV that suddenly cannot rely on GNSS and hence changes course to reach a predefined recovery location situated at approximately 1 h of flight time. In the process, in addition to executing an altitude and airspeed adjustment, the autonomous aircraft faces significant weather and wind field changes that make its GNSS-denied navigation even more challenging.

With respect to the mission, the stochastic parameters include the initial airspeed, pressure altitude, and bearing ( $v_{TAS,INI}$ ,  $H_{P,INI}$ ,  $\varnothing_{INI}$ ); their final values ( $v_{TAS,END}$ ,  $H_{P,END}$ ,  $\varnothing_{END}$ ); and the time at which each of the three maneuvers is initiated (turns are executed with a bank angle of  $\zeta_{TURN} = \pm 10^\circ$ . Altitude changes employ an aerodynamic path angle of  $\gamma_{TAS,CLIMB} = \pm 2^\circ$ . Additionally, airspeed modifications are automatically executed by the control system as set-point changes). The scenario lasts for  $t_{END} = 3800$  s, while the GNSS signals are lost at  $t_{GNSS} = 100$  s.

The wind field is also defined stochastically, as its two parameters (speed and bearing) are constant both at the beginning ( $v_{WIND,INI}$ ,  $\varnothing_{WIND,INI}$ ) and conclusion ( $v_{WIND,END}$ ,  $\varnothing_{WIND,END}$ ) of the scenario, with a linear transition in between. The specific times at which the wind change starts and concludes also vary stochastically among the different simulation runs.

A similar linear transition occurs with the temperature and pressure offsets that define the atmospheric properties, as they are constant both at the start ( $\Delta T_{INI}$ ,  $\Delta p_{INI}$ ) and end ( $\Delta T_{END}$ ,  $\Delta p_{END}$ ) of the flight.

The turbulence remains strong throughout the whole scenario, but its specific values also vary stochastically from one execution to the next.

- Scenario #2 represents the challenges involved in continuing with the original mission upon the loss of the GNSS signals, executing a series of continuous turn maneuvers over a relatively short period of time with no atmospheric or wind variations. As in scenario #1, the GNSS signals are lost at  $t_{GNSS} = 100$  s, but the scenario duration is shorter ( $t_{END} = 500$  s). The initial airspeed and pressure altitude ( $v_{TAS,INI}$ ,  $H_{P,INI}$ ) are defined stochastically and do not change throughout the whole scenario; the bearing, however, changes a total of eight times between its initial and final values, with all intermediate bearing values as well as the time for each turn varying stochastically from one execution to the next. Although the same turbulence is employed as in scenario #1, the wind and atmospheric parameters ( $v_{WIND,INI}$ ,  $\varnothing_{WIND,INI}$ ,  $\Delta T_{INI}$ ,  $\Delta p_{INI}$ ) remain constant throughout scenario #2.

## 10. Results: Navigation System Error in GNSS-Denied Conditions

This section presents the results obtained with the proposed fully integrated visual inertial navigation system when executing Monte Carlo simulations of the two GNSS-

denied scenarios introduced in Section 9, each consisting of one hundred executions. They are compared with those obtained with two other navigation systems:

- A standalone inertial system specifically designed to lower the GNSS-denied horizontal position drift, for which the results obtained with the same two scenarios are described in [2]. The attitude estimation error does not drift and is bounded by the quality of the onboard sensors, ensuring that the aircraft can remain aloft for as long as there is fuel available. The vertical position and ground velocity estimation errors are also bounded by atmospheric physics and do not drift; they depend on the atmospheric pressure offset and wind field changes that occur since the GNSS signals are lost. On the other hand, the horizontal position drifts as a consequence of integrating the ground velocity without absolute observations. Note that a standalone inertial system is hence capable of successfully estimating four of the six  $\mathbb{S}\mathbb{E}(3)$  degrees of freedom (attitude plus altitude).
- A visual navigation system (which relies exclusively on the images generated by an onboard camera) aided by the attitude and altitude estimated by the above inertial system, for which the results are described in [1]. This system slowly drifts in all six degrees of freedom. Although its attitude and altitude estimation capabilities are qualitatively inferior to those of the inertial system, its horizontal position error is just a fraction of what can be achieved without the use of images.

The tables below contain the *navigation system error*, or NSE (difference between the real states  $\mathbf{x}$  and their estimations  $\hat{\mathbf{x}}$ ), incurred by the three navigation systems at the conclusion of both scenarios, represented by the mean  $\mu$ , standard deviation  $\sigma$ , and maximum value of the estimation errors. In addition, the figures depict the variation with time of the NSE mean ( $\mu$ , solid lines) and standard deviation ( $\sigma$ , dashed lines) for the one hundred executions.

### 10.1. Body Attitude Estimation

Table 3 shows the NSE at the conclusion of the scenarios for the norm of the rotation vector between the real body attitude  $\mathbf{q}_{\text{NB}}$  and its estimation  $\hat{\mathbf{q}}_{\text{NB}}$ , which can be formally written as  $\|\Delta\mathbf{r}_{\text{NB}}^{\text{B}}\| = \|\hat{\mathbf{q}}_{\text{NB}} \ominus \mathbf{q}_{\text{NB}}\|$ . In addition, Figure 5 depicts the variation with time of the body attitude NSE.

**Table 3.** Aggregated final body attitude NSE (100 runs).

Scenario	$\ \Delta\hat{\mathbf{r}}_{\text{NB}}^{\text{B}}\ (\mathbf{t}_{\text{END}})[^{\circ}]$	Inertial	Visual	Proposed
#1	mean	0.158	0.218	0.100
	std	0.114	0.103	0.059
	max	0.611	0.606	0.328
#2	mean	0.128	0.221	0.107
	std	0.078	0.137	0.068
	max	0.369	0.788	0.377

The proposed visual inertial navigation system shows bounded body attitude estimations, with significant improvements with respect to both the inertial and the inertially assisted visual systems, which can be attributed to various factors. On the one hand, the navigation filter successfully merges the incremental horizontal displacement observations supplied by the VVS with the bearing provided by the magnetometers, resulting in better yaw estimations. In addition, the absence of simplifications in the EKF equations ([2] describes how, without both GNSS receiver and VVS observations, the filter equations need to be simplified to protect the filter behavior from the negative effects of inaccurate velocity and position estimations) ensures that the benefits also reach the already-accurate inertial pitch and bank angle estimations. Additional body attitude estimation improvements are achieved by closing the loop (Figure 2) and directly employing the increasingly accurate inertial estimations to feed the visual system (Section 6).

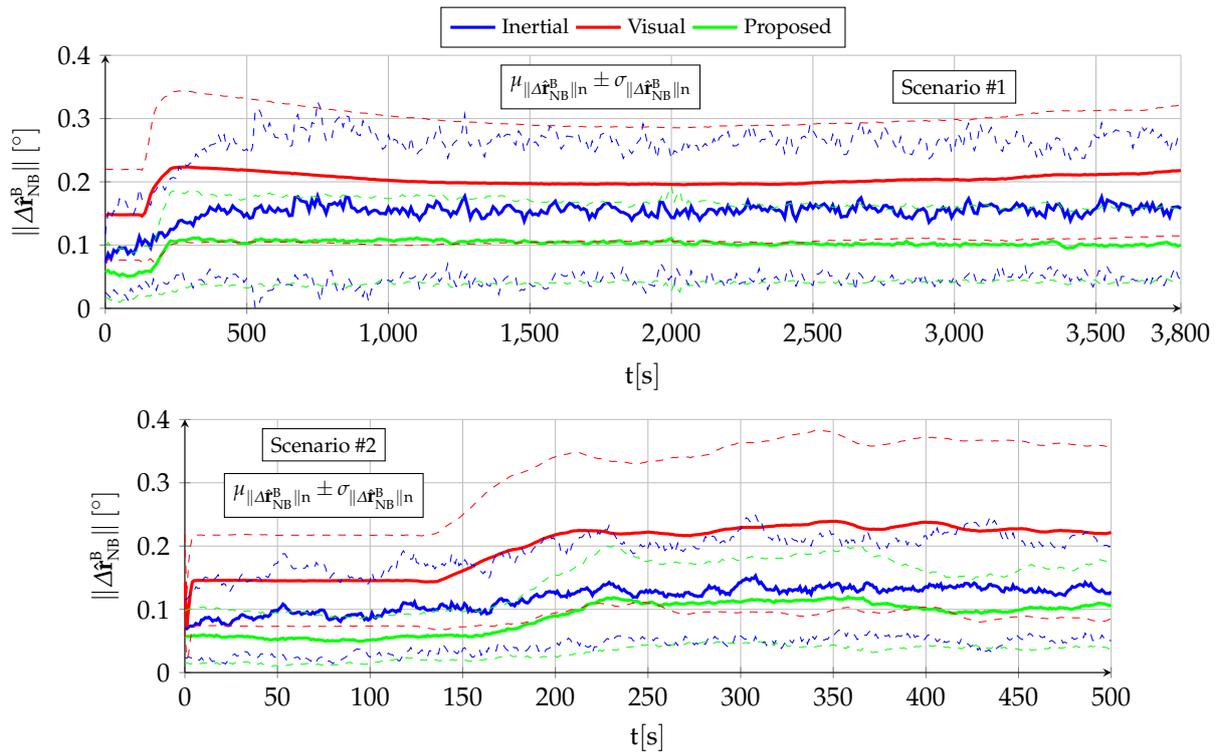


Figure 5. Aggregated body attitude NSE (100 runs).

10.2. Vertical Position Estimation

Table 4 shows the vertical position NSE ( $\Delta\hat{h} = \hat{h} - h$ ) at the conclusion of both scenarios. The estimation errors are unbiased in all cases as the mean is always significantly lower than both the standard deviation and the maximum error value.

Table 4. Aggregated final vertical position NSE (100 runs).

Scenario	$\Delta\hat{h}(t_{END})$ [m]	Inertial	Visual	Proposed
#1	mean	−4.18	+22.86	−3.97
	std	25.78	49.17	26.12
	max	−70.49	+175.76	−70.47
#2	mean	+0.76	+3.59	+0.74
	std	7.55	13.01	7.60
	max	−19.86	+71.64	−18.86

The evolution with time of the geometric altitude NSE is depicted in Figure 6. Note that in contrast with Sections 10.1 and 10.3, the altitude error is a scalar that can be either positive or negative, so the dashed lines (standard deviation) are more indicative of the relative performance of the various navigation systems, while the solid lines (mean) stay relatively close to zero when averaging the one hundred executions.

The vertical position estimation accuracy of the proposed system is virtually the same as that of the standalone inertial one because the navigation filter relies on freezing the pressure offset estimation when the GNSS signals are lost, exactly the same as the inertial filter [2]. The resulting altitude estimations are unbiased and bounded, with an error that depends on the change in pressure offset since the time the GNSS are lost.

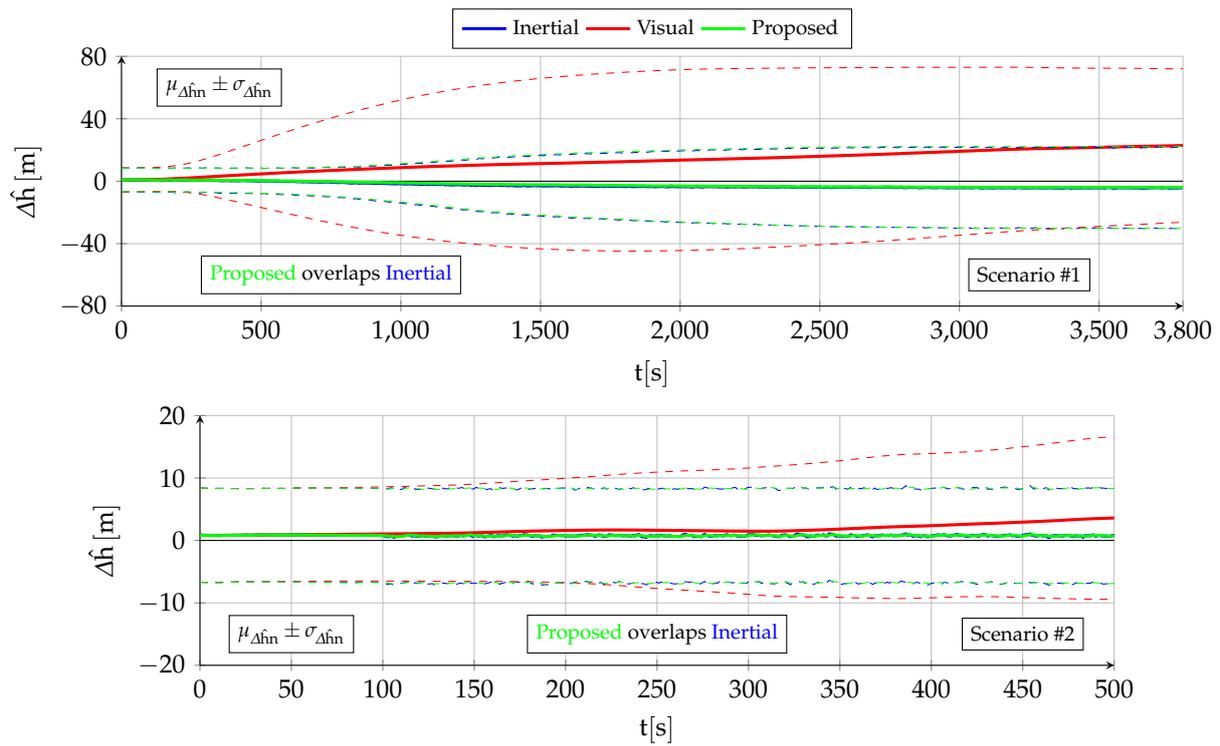


Figure 6. Aggregated vertical position NSE (100 runs).

10.3. Horizontal Position Estimation

Table 5 lists the mean, standard deviation, and maximum value of the horizontal position NSE at the conclusion of the two scenarios, both in length units and as a percentage of the total distance flown in GNSS-denied conditions. The three navigation systems considered exhibit an unrestrained horizontal position drift or error growth with time, as shown in Figure 7.

Table 5. Aggregated final horizontal position NSE (100 runs).

Scenario		Distance [m]	Inertial		Visual		Proposed	
			$\Delta\hat{x}_{HOR}(t_{END})$ [m]	[%]	$\Delta\hat{x}_{HOR}(t_{END})$ [m]	[%]	$\Delta\hat{x}_{HOR}(t_{END})$ [m]	[%]
#1	mean	107,873	7276	7.10	488	0.46	207	0.19
	std	19,756	4880	5.69	350	0.31	185	0.15
	max	172,842	25,288	32.38	1957	1.48	1257	1.09
#2	mean	14,198	216	1.52	33	0.23	18	0.13
	std	1176	119	0.86	26	0.18	9	0.07
	max	18,253	586	4.38	130	0.98	43	0.33

The major reduction in horizontal position drift that occurs when comparing the standalone inertial system with the proposed one is obtained in two distinct phases. The first one, described in [1], is caused by employing a visual navigation system that can rely on the inertial attitude and altitude estimations to improve its nonlinear pose optimizations.

The objective when supplying the proposed navigation filter with the VVS observations is to employ more accurate equations within the EKF as there is no longer a need to protect the filter from the negative effects of inaccurate velocity and position values, as in the case of the standalone inertial filter [2]. This improved design enables significant improvements in body attitude accuracy (Section 10.1) because the filter is capable of incorporating the horizontal position accuracy improvements into the EKF filter. In a

positive feedback loop, the proposed system then capitalizes on its more accurate attitude estimations to further improve its internal horizontal position estimations (Figure 2).

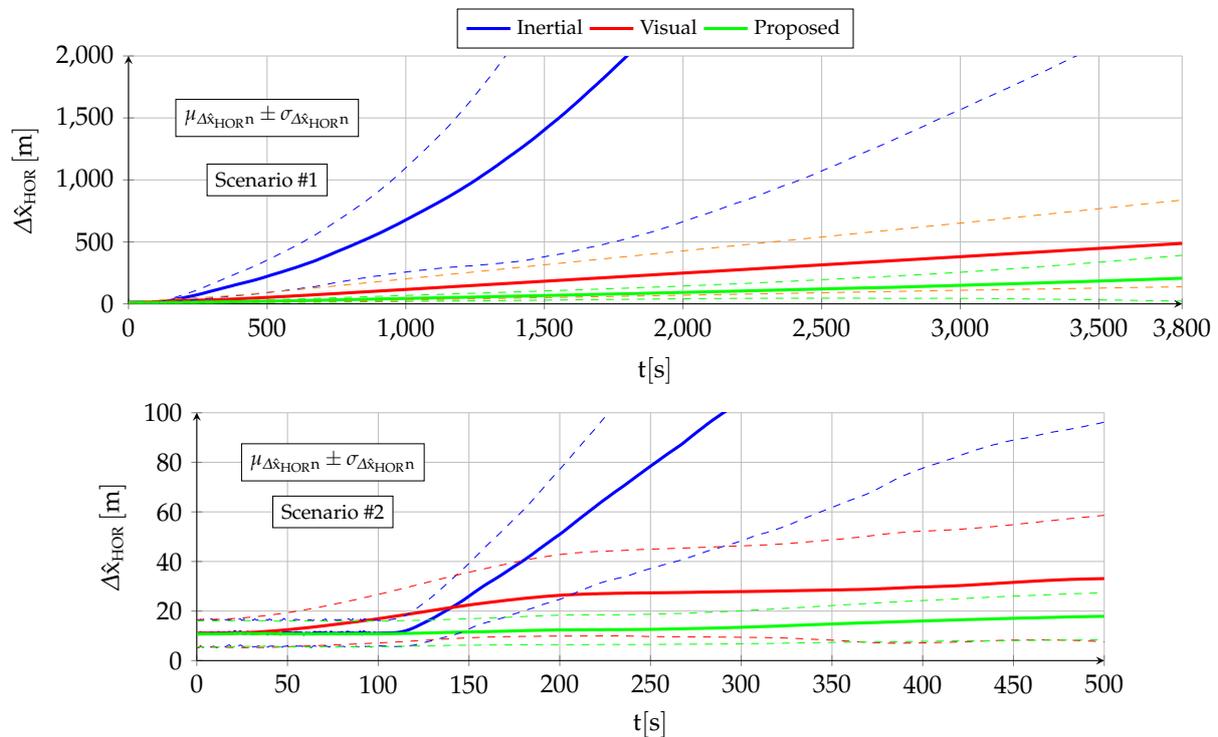


Figure 7. Aggregated horizontal position NSE (100 runs).

## 11. Summary and Conclusions

This article proves that the inertial and visual navigation systems of an autonomous UAV can be combined in such a way that the resulting long-term GNSS-denied horizontal position drift is only a small fraction of what can be achieved by either system individually. The proposed system, however, does not constitute a full GNSS replacement as it relies on incremental instead of absolute position observations and, hence, can only reduce the position drift but not eliminate it.

The proposed visual inertial navigation filter, specifically designed for the challenges faced by autonomous fixed-wing aircraft that encounter GNSS-denied conditions, merges the observations provided by onboard accelerometers, gyroscopes, and magnetometers with those of the virtual vision sensor, or VVS. The VVS is the denomination of the outputs generated by a visual inertial odometry pipeline that relies on the images of the Earth surface generated by an onboard camera as well as on the navigation filter outputs. The filter is implemented in the manifold of rigid body rotations or  $\mathbb{SO}(3)$  in order to minimize the accumulation of errors in the absence of the absolute position observations provided by the GNSS receiver. The results obtained when applying the proposed algorithms to high-fidelity Monte Carlo simulations of two scenarios representative of the challenges of GNSS-denied navigation indicate the following:

- The **body attitude** estimation is qualitatively similar to that obtained by a standalone inertial filter without any visual aid [2]. The bounded estimations enable the aircraft to remain aloft in GNSS-denied conditions for as long as it has fuel. Quantitatively, the VVS observations and the associated more accurate filter equations result in significant accuracy improvements when compared with the [2] baseline.
- The **vertical position** estimation is qualitatively and quantitatively similar to that of the standalone inertial filter [2]. In addition to ionospheric effects (which also apply when GNSS signals are available), the altitude error depends on the amount of

pressure offset variation since entering GNSS-denied conditions, being unbiased (zero mean) and bounded by atmospheric physics.

- The **horizontal position** estimation exhibits drastic quantitative improvements over the baseline standalone inertial filter [2], although from a qualitative point of view, the estimation error is not bounded as the drift cannot be fully eliminated.

Future work will focus on addressing some of the limitations of the proposed navigation system, such as the restriction to fixed-wing vehicles (which originates when using the pitot tube airspeed readings to replace visual navigation when not enough features can be extracted from a texture-poor terrain) and the inability to fly at night or above clouds (infrared images could potentially be used for this purpose).

**Author Contributions:** Conceptualization, E.G.; methodology, E.G.; software, E.G.; validation, E.G.; formal analysis, E.G.; investigation, E.G.; resources, E.G.; data curation, E.G.; writing—original draft preparation, E.G.; writing—review and editing, A.B.; visualization, E.G.; supervision, A.B.; project administration, A.B.; funding acquisition, A.B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work has received funding from RoboCity2030-DIH-CM, Madrid Robotics Digital Innovation Hub, S2018/NMT-4331, funded by R&D Activity Programs in the Madrid Community and cofinanced by the EU Structural Funds.

**Data Availability Statement:** An open-source C++ implementation of the described algorithms can be found at [80].

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

BRIEF	binary robust independent elementary features
DSO	direct sparse odometry
ECEF	Earth centered Earth fixed
EKF	extended Kalman filter
FAST	features from accelerated segment test
GNSS	Global Navigation Satellite System
IMU	inertial measurement unit
iSAM	incremental smoothing and mapping
LSD	large-scale direct
MAV	micro air vehicle
MSCKF	multistate constraint Kalman filter
MSF	multisensor fusion
NED	north east down
NSE	navigation system error
OKVIS	open keyframe visual inertial SLAM
ORB	oriented FAST and rotated BRIEF
ROVIO	robust visual inertial odometry
SLAM	simultaneous localization and mapping
$\mathbb{SE}(3)$	special Euclidean group of $\mathbb{R}^3$
$\mathbb{SO}(3)$	special orthogonal group of $\mathbb{R}^3$
SVO	semidirect visual odometry
UAV	unmanned aerial vehicle
VINS	visual inertial navigation system
VIO	visual inertial odometry
VO	visual odometry
VVS	virtual vision sensor
WGS84	World Geodetic System 1984

## Appendix A. Required Jacobians

This appendix groups together the various Jacobians employed in this article. They are the following:

- The time derivative of the geodetic coordinates  $\mathbf{T}^{\text{E,GDT}} = [\lambda \ \varphi \ h]^T$  (longitude, latitude, and altitude) depends on the ground velocity  $\mathbf{v}^N$  per (A1), where M and N represent the WGS84 ellipsoid radii of curvature of meridian and prime vertical, respectively. The Jacobian with respect to  $\mathbf{v}^N$ , given by (A2) and employed in (20), is hence straightforward:

$$\dot{\mathbf{T}}^{\text{E,GDT}} = \begin{bmatrix} \frac{v_2^N}{[N(\varphi) + h] \cos \varphi} & \frac{v_1^N}{M(\varphi) + h} & -v_3^N \end{bmatrix}^T \quad (\text{A1})$$

$$\mathbf{J}_{\mathbf{v}^N} \dot{\mathbf{T}}^{\text{E,GDT}} = \begin{bmatrix} 0 & \frac{1}{(N + h) \cos \varphi} & 0 \\ \frac{1}{M + h} & 0 & 0 \\ 0 & 0 & -1 \end{bmatrix} \in \mathbb{R}^{3 \times 3} \quad (\text{A2})$$

- The motion angular velocity  $\boldsymbol{\omega}_{\text{EN}} = \dot{\lambda} \mathbf{i}_3^{\text{E}} - \dot{\varphi} \mathbf{i}_2^{\text{N}}$  represents the rotation experienced by any object that moves without modifying its attitude with respect to the Earth surface. It is caused by the curvature of the Earth, and its expression when viewed in  $F_N$  is given by (A3). Its Jacobian with respect to  $\mathbf{v}^N$ , provided by (A4) and employed in (22) and (33), is also straightforward:

$$\boldsymbol{\omega}_{\text{EN}}^N = \begin{bmatrix} \frac{v_2^N}{N(\varphi) + h} & \frac{-v_1^N}{M(\varphi) + h} & \frac{-v_2^N \tan \varphi}{N(\varphi) + h} \end{bmatrix}^T \quad (\text{A3})$$

$$\mathbf{J}_{\mathbf{v}^N} \boldsymbol{\omega}_{\text{EN}}^N = \begin{bmatrix} 0 & \frac{1}{N + h} & 0 \\ -1 & 0 & 0 \\ \frac{1}{M + h} & 0 & 0 \\ 0 & \frac{-\tan \varphi}{N + h} & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 3} \quad (\text{A4})$$

- The Coriolis acceleration  $\mathbf{a}_{\text{COR}} = 2 [\boldsymbol{\omega}_{\text{IE}}]_{\times} \mathbf{v}$  is the double cross product between the Earth angular velocity caused by its rotation around the  $\mathbf{i}_3^{\text{E}}$  axis at a constant rate  $\omega_{\text{E}}$  and the aircraft velocity. Its expression when viewed in  $F_N$  is provided by (A5), resulting in the (A6) Jacobian with respect to  $\mathbf{v}^N$ , which appears in (22):

$$\mathbf{a}_{\text{COR}}^N = 2 [\boldsymbol{\omega}_{\text{IE}}^N]_{\times} \mathbf{v}^N = 2 \omega_{\text{E}} [v_2^N \sin \varphi \ (-v_1^N \sin \varphi - v_3^N \cos \varphi) \ v_2^N \cos \varphi]^T \quad (\text{A5})$$

$$\mathbf{J}_{\mathbf{v}^N} \mathbf{a}_{\text{COR}}^N = 2 \omega_{\text{E}} \begin{bmatrix} 0 & \sin \varphi & 0 \\ -\sin \varphi & 0 & -\cos \varphi \\ 0 & \cos \varphi & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 3} \quad (\text{A6})$$

- The Lie Jacobian  $\mathbf{J}_{\mathcal{R}}^{\mathbf{q} \oplus \Delta \mathbf{r}}$  represents the derivative of the function  $f(\mathbf{q}, \Delta \mathbf{r}) = \mathbf{q} \oplus \Delta \mathbf{r}$ , that is, the concatenation between the  $\mathbb{S}\mathbb{O}(3)$  attitude  $\mathbf{q}$  and its local perturbation  $\Delta \mathbf{r}$ , with respect to the  $\mathbb{S}\mathbb{O}(3)$  attitude  $\mathcal{R} \equiv \mathbf{q}$ , when the increments are viewed in their respective local tangent spaces, that is, tangent respectively at  $\mathbf{q}$  and  $\mathbf{q} \oplus \Delta \mathbf{r}$ . Interested readers should refer to [4,5] for the obtainment of (A7), where  $\mathbf{R}(\Delta \mathbf{r})$  represents the direct cosine matrix corresponding to a given rotation vector  $\Delta \mathbf{r}$ . This Jacobian is employed in (48).

$$\begin{aligned} \mathbf{J}_{\mathcal{R}}^{\mathbf{q} \oplus \Delta \mathbf{r}} &= \mathbf{R}^T(\Delta \mathbf{r}) = \mathbf{R}(-\Delta \mathbf{r}) \\ &= \mathbf{I}_3 - \frac{[\Delta \mathbf{r}]_{\times}}{\|\Delta \mathbf{r}\|} \sin \|\Delta \mathbf{r}\| + \frac{[\Delta \mathbf{r}]_{\times}^2}{\|\Delta \mathbf{r}\|^2} (1 - \cos \|\Delta \mathbf{r}\|) \in \mathbb{R}^{3 \times 3} \quad (\text{A7}) \end{aligned}$$

- The Lie Jacobian  $J_{\mathcal{R}}^{\mathbf{q} \otimes \mathbf{v} \otimes \mathbf{q}^*}$  represents the derivative of the function  $f(\mathbf{q}, \mathbf{v}) = \mathbf{q} \otimes \mathbf{v} \otimes \mathbf{q}^*$ , that is, the rotation of  $\mathbf{v}$  according to the  $\mathbb{S}\mathbb{O}(3)$  attitude  $\mathbf{q}$ , with respect to the  $\mathbb{S}\mathbb{O}(3)$  attitude  $\mathcal{R} \equiv \mathbf{q}$ , when the  $\mathcal{R}$  increment is viewed in its local tangent space and that of the resulting  $\mathbb{R}^3$  vector is viewed in its Euclidean space.  $J_{\mathbf{v}}^{\mathbf{q} \otimes \mathbf{v} \otimes \mathbf{q}^*}$  is the derivative of the same function with respect to the  $\mathbb{R}^3$  unrotated vector  $\mathbf{v}$ , in which the increments of the unrotated vector are also viewed in the  $\mathbb{R}^3$  Euclidean space. Refer to [4,5] for the obtainment of (A8) and (A9), which appear on (21) and (23), respectively, and where  $\mathbf{R}(\mathbf{q})$  represents the direct cosine matrix equivalent to the unit quaternion  $\mathbf{q}$ :

$$J_{\mathcal{R}}^{\mathbf{q} \otimes \mathbf{v} \otimes \mathbf{q}^*} = -\mathbf{R}(\mathbf{q}) [\mathbf{v}]_{\times} \in \mathbb{R}^{3 \times 3} \quad (\text{A8})$$

$$J_{\mathbf{v}}^{\mathbf{q} \otimes \mathbf{v} \otimes \mathbf{q}^*} = \mathbf{R}(\mathbf{q}) \in \mathbb{R}^{3 \times 3} \quad (\text{A9})$$

- The Lie Jacobians  $J_{\mathcal{R}}^{\mathbf{q}^* \otimes \mathbf{v} \otimes \mathbf{q}}$  and  $J_{\mathbf{v}}^{\mathbf{q}^* \otimes \mathbf{v} \otimes \mathbf{q}}$  are similar to the previous ones but refer to the inverse rotation action  $f(\mathbf{q}, \mathbf{v}) = \mathbf{q}^* \otimes \mathbf{v} \otimes \mathbf{q}$ . Expressions (A10) and (A11), which appear on (31), (33), (37), and (39), respectively, are also obtained in [4,5]

$$J_{\mathcal{R}}^{\mathbf{q}^* \otimes \mathbf{v} \otimes \mathbf{q}} = [\mathbf{R}^T(\mathbf{q}) \mathbf{v}]_{\times} \in \mathbb{R}^{3 \times 3} \quad (\text{A10})$$

$$J_{\mathbf{v}}^{\mathbf{q}^* \otimes \mathbf{v} \otimes \mathbf{q}} = \mathbf{R}^T(\mathbf{q}) \in \mathbb{R}^{3 \times 3} \quad (\text{A11})$$

## References

1. Gallo, E.; Barrientos, A. GNSS-Denied Semi Direct Visual Navigation for Autonomous UAVs Aided by PI-Inspired Priors. *Aerospace* **2023**, *10*, 220. [CrossRef]
2. Gallo, E.; Barrientos, A. Reduction of GNSS-Denied Inertial Navigation Errors for Fixed Wing Autonomous Unmanned Air Vehicles. *Aerosp. Sci. Technol.* **2021**, *120*, 107237. [CrossRef]
3. Sola, J. Quaternion Kinematics for the Error-State Kalman Filter. *arXiv* **2017**, arXiv:1711.02508. [CrossRef]
4. Sola, J.; Deray, J.; Atchuthan, D. A Micro Lie Theory for State Estimation in Robotics. *arXiv* **2018**, arXiv:1812.01537. [CrossRef]
5. Gallo, E. The  $\mathbb{S}\mathbb{O}(3)$  and  $\mathbb{S}\mathbb{E}(3)$  Lie Algebras of Rigid Body Rotations and Motions and their Application to Discrete Integration, Gradient Descent Optimization, and State Estimation. *arXiv* **2023**, arXiv:2205.12572. [CrossRef]
6. Hassanalian, M.; Abdelkefi, A. Classifications, Applications, and Design Challenges of Drones: A Review. *Prog. Aerosp. Sci.* **2017**, *91*, 99–131. [CrossRef]
7. Shakhathreh, H.; Sawalmeh, A.H.; Al-Fuqaha, A.; Dou, Z.; Almaita, E.; Khalil, I.; Othman, N.S.; Khreishah, A.; Guizani, M. Unmanned Aerial Vehicles (UAVs): A Survey on Civil Applications and Key Research Challenges. *IEEE Access* **2019**, *7*, 48572–48634. [CrossRef]
8. Bijjahalli, S.; Sabatini, R.; Gardi, A. Advances in Intelligent and Autonomous Navigation Systems for Small UAS. *Prog. Aerosp. Sci.* **2020**, *115*, 100617. [CrossRef]
9. Farrell, J.A. *Aided Navigation, GPS with High Rate Sensors*; Electronic Engineering Series; McGraw-Hill: New York, NY, USA, 2008.
10. Groves, P.D. *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*; GNSS Technology and Application Series; Artech House: Norwood, MA, USA, 2008.
11. Chatfield, A.B. *Fundamentals of High Accuracy Inertial Navigation*; American Institute of Aeronautics and Astronautics, Progress in Astronautics and Aeronautics: Reston, VA, USA, 1997; Volume 174.
12. Elbanhawi, M.; Mohamed, A.; Clothier, R.; Palmer, J.; Simic, M.; Watkins, S. Enabling Technologies for Autonomous MAV Operations. *Prog. Aerosp. Sci.* **2017**, *91*, 27–52. [CrossRef]
13. Sabatini, R.; Moore, T.; Ramasamy, S. Global Navigation Satellite Systems Performance Analysis and Augmentation Strategies in Aviation. *Prog. Aerosp. Sci.* **2017**, *95*, 45–98. [CrossRef]
14. Tippitt, C.; Schultz, A.; Procino, W. *Vehicle Navigation: Autonomy Through GPS-Enabled and GPS-Denied Environments*; State of the Art Report DSIAC-2020-1328; Defense Systems Information Analysis Center: Belcamp, MD, USA, 2020.
15. Gyagenda, N.; Hatilima, J.V.; Roth, H.; Zhmud, V. A Review of GNSS Independent UAV Navigation Techniques. *Robot. Auton. Syst.* **2022**, *152*, 104069. [CrossRef]
16. Kapoor, R.; Ramasamy, S.; Gardi, A.; Sabatini, R. UAV Navigation using Signals of Opportunity in Urban Environments: A Review. *Energy Procedia* **2017**, *110*, 377–383. [CrossRef]
17. Coluccia, A.; Ricciato, F.; Ricci, G. Positioning Based on Signals of Opportunity. *IEEE Commun. Lett.* **2014**, *18*, 356–359. [CrossRef]
18. Goh, S.T.; Abdelkhalik, O.; Zekavat, S.A. A Weighted Measurement Fusion Kalman Filter Implementation for UAV Navigation. *Aerosp. Sci. Technol.* **2013**, *28*, 315–323. [CrossRef]

19. Couturier, A.; Akhloufi, M.A. A Review on Absolute Visual Localization for UAV. *Robot. Auton. Syst.* **2020**, *135*, 103666. [[CrossRef](#)]
20. Goforth, H.; Lucey, S. GPS-Denied UAV Localization using Pre Existing Satellite Imagery. In Proceedings of the IEEE International Conference on Robotics and Automation, Montreal, QC, Canada, 20–24 May 2019; IEEE: Piscataway Township, NJ, USA, 2019. [[CrossRef](#)]
21. Ziaei, N. Geolocation of an Aircraft using Image Registration Coupling Modes for Autonomous Navigation. *arXiv* **2019**, arXiv:1909.02875. [[CrossRef](#)]
22. Wang, T. Augmented UAS Navigation in GPS Denied Terrain Environments Using Synthetic Vision. Master's Thesis, Iowa State University, Ames, IA, USA, 2018. [[CrossRef](#)]
23. Kinnari, J.; Verdoja, F.; Kyrki, V. Season-Invariant GNSS-Denied Visual Localization for UAVs. *IEEE Robot. Autom. Lett.* **2022**, *7*, 10232–10239. [[CrossRef](#)]
24. Ren, Y.; Wang, Z. A Novel Scene Matching Algorithm via Deep Learning for Vision-Based UAV Absolute Localization. In Proceedings of the International Conference on Machine Learning, Cloud Computing and Intelligent Mining, Xiamen, China, 5–7 August 2022. [[CrossRef](#)]
25. Liu, K.; He, X.; Mao, J.; Zhang, L.; Zhou, W.; Qu, H.; Luo, K. Map Aided Visual Inertial Integrated Navigation for Long Range UAVs. In Proceedings of the International Conference on Guidance, Navigation, and Control, Sopot, Poland, 12–16 June 2023. [[CrossRef](#)]
26. Zhang, Q.; Zhang, H.; Lan, Z.; Chen, W.; Zhang, Z. A DNN-Based Optical Aided Autonomous Navigation System for UAV Under GNSS-Denied Environment. In Proceedings of the International Conference on Autonomous Unmanned Systems, Warsaw, Poland, 6–9 June 2023. [[CrossRef](#)]
27. Jurevicius, R.; Marcinkevicius, V.; Seibokas, J. Robust GNSS-Denied Localization for UAV using Particle Filter and Visual Odometry. *Mach. Vis. Appl.* **2019**, *30*, 1181–1190. [[CrossRef](#)]
28. Scaramuzza, D.; Fraundorfer, F. Visual Odometry Part 1: The First 30 Years and Fundamentals. *IEEE Robot. Autom. Mag.* **2011**, *18*, 80–92. [[CrossRef](#)]
29. Fraundorfer, F.; Scaramuzza, D. Visual Odometry Part 2: Matching, Robustness, Optimization, and Applications. *IEEE Robot. Autom. Mag.* **2012**, *19*, 78–90. [[CrossRef](#)]
30. Scaramuzza, D. *Tutorial on Visual Odometry*; Robotics & Perception Group, University of Zurich: Zurich, Switzerland, 2012.
31. Scaramuzza, D. *Visual Odometry and SLAM: Past, Present, and the Robust Perception Age*; Robotics & Perception Group, University of Zurich: Zurich, Switzerland, 2017.
32. Cadena, C.; Carlone, L.; Carrillo, H.; Latif, Y.; Scaramuzza, D.; Neira, J.; Reid, I.; Leonard, J.J. Past, Present, and Future of Simultaneous Localization and Mapping: Towards the Robust Perception Age. *IEEE Trans. Robot.* **2016**, *32*, 1309–1332. [[CrossRef](#)]
33. Forster, C.; Pizzoli, M.; Scaramuzza, D. SVO: Fast Semi-Direct Monocular Visual Odometry. In Proceedings of the IEEE International Conference on Robotics and Automation, Hong Kong, China, 31 May–7 June 2014. [[CrossRef](#)]
34. Forster, C.; Zhang, Z.; Gassner, M.; Werlberger, M.; Scaramuzza, D. SVO: Semidirect Visual Odometry for Monocular and Multicamera Systems. *IEEE Trans. Robot.* **2016**, *33*, 249–265. [[CrossRef](#)]
35. Engel, J.; Koltun, V.; Cremers, D. Direct Sparse Odometry. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 611–625. [[CrossRef](#)]
36. Engel, J.; Schops, T.; Cremers, D. LSD-SLAM: Large Scale Direct Monocular SLAM. *Eur. Conf. Comput. Vis.* **2014**, 834–849. [[CrossRef](#)]
37. Mur-Artal, R.; Montiel, J.M.M.; Tardos, J.D. ORB-SLAM: A Versatile and Accurate Monocular SLAM System. *IEEE Trans. Robot.* **2015**, *31*, 1147–1163. [[CrossRef](#)]
38. Mur-Artal, R.; Tardos, J.D. ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras. *IEEE Trans. Robot.* **2017**, *33*, 1255–1262. [[CrossRef](#)]
39. Mur-Artal, R. Real-Time Accurate Visual SLAM with Place Recognition. Ph.D. Thesis, University of Zaragoza, Zaragoza, Spain, 2017. [[CrossRef](#)]
40. Crassidis, J.L.; Markley, F.L. Unscented Filtering for Spacecraft Attitude Estimation. *J. Guid. Control Dyn.* **2003**, *26*, 536–542. [[CrossRef](#)]
41. Grip, H.F.; Fossen, T.I.; Johansen, T.A.; Saberi, A. Attitude Estimation Using Biased Gyro and Vector Measurements with Time Varying Reference Vectors. *IEEE Trans. Autom. Control* **2012**, *57*, 1332–1338. [[CrossRef](#)]
42. Kottah, R.; Narkhede, P.; Kumar, V.; Karar, V.; Poddar, S. Multiple Model Adaptive Complementary Filter for Attitude Estimation. *Aerosp. Sci. Technol.* **2017**, *69*, 574–581. [[CrossRef](#)]
43. Hashim, H.A. Systematic Convergence of Nonlinear Stochastic Estimators on the Special Orthogonal Group SO(3). *Int. J. Robust Nonlinear Control* **2020**, *30*, 3848–3870. [[CrossRef](#)]
44. Hashim, H.A.; Brown, L.J.; McIsaac, K. Nonlinear Stochastic Attitude Filters on the Special Orthogonal Group SO(3): Ito and Stratonovich. *IEEE Trans. Syst. Man Cybern.* **2019**, *49*, 1853–1865. [[CrossRef](#)]
45. Batista, P.; Silvestre, C.; Oliveira, P. On the Observability of Linear Motion Quantities in Navigation Systems. *Syst. Control Lett.* **2011**, *60*, 101–110. [[CrossRef](#)]
46. Hashim, H.A.; Brown, L.J.; McIsaac, K. Nonlinear Pose Filters on the Special Euclidean Group SE(3) with Guaranteed Transient and Steady State Performance. *IEEE Trans. Syst. Man Cybern.* **2019**, *51*, 2949–2962. [[CrossRef](#)]

47. Hashim, H.A. GPS Denied Navigation: Attitude, Position, Linear Velocity, and Gravity Estimation with Nonlinear Stochastic Observer. In Proceedings of the 2021 American Control Conference, Online, 25–28 May 2021. [CrossRef]
48. Hua, M.D.; Allibert, G. Riccati Observer Design for Pose, Linear Velocity, and Gravity Direction Estimation Using Landmark Position and IMU Measurements. In Proceedings of the IEEE Conference on Control Technology and Applications, Copenhagen, Denmark, 21–24 August 2018. [CrossRef]
49. Barrau, A.; Bonnabel, S. The Invariant Extended Kalman Filter as a Stable Observer. *IEEE Trans. Autom. Control* **2017**, *62*, 1797–1812. [CrossRef]
50. Scaramuzza, D.; Zhang, Z. Visual-Inertial Odometry of Aerial Robots. *arXiv* **2019**, arXiv:1906.03289. [CrossRef]
51. Huang, G. Visual-Inertial Navigation: A Concise Review. *arXiv* **2019**, arXiv:1906.02650. [CrossRef]
52. von Stumberg, L.; Usenko, V.; Cremers, D. Chapter 7—A Review and Quantitative Evaluation of Direct Visual Inertial Odometry. In *Multimodal Scene Understanding*; Yang, M.Y., Rosenhahn, B., Murino, V., Eds.; Academic Press: Cambridge, MA, USA, 2019. [CrossRef]
53. Feng, X.; Jiang, Y.; Yang, X.; Du, M.; Li, X. Computer Vision Algorithms and Hardware Implementations: A Survey. *Integr. Vlsi J.* **2019**, *69*, 309–320. [CrossRef]
54. Al-Kaff, A.; Martin, D.; Garcia, F.; de la Escalera, A.; Maria, J. Survey of Computer Vision Algorithms and Applications for Unmanned Aerial Vehicles. *Expert Syst. Appl.* **2017**, *92*, 447–463. [CrossRef]
55. Mourikis, A.I.; Roumeliotis, S.I. A Multi-State Constraint Kalman Filter for Vision-aided Inertial Navigation. In Proceedings of the 2007 IEEE International Conference on Robotics and Automation, Roma, Italy, 10–14 April 2007. [CrossRef]
56. Leutenegger, S.; Furgale, P.; Rabaud, V.; Chli, M.; Konolige, K.; Siegwart, R. Keyframe Based Visual Inertial SLAM Using Nonlinear Optimization. *Robot. Sci. Syst.* **2013**. [CrossRef]
57. Leutenegger, S.; Lynen, S.; Bosse, M.; Siegwart, R.; Furgale, P. Keyframe Based Visual Inertial SLAM Using Nonlinear Optimization. *Int. J. Robot. Res.* **2015**, *34*, 314–334. [CrossRef]
58. Bloesch, M.; Omari, S.; Hutter, M.; Siegwart, R. Robust Visual Inertial Odometry Using a Direct EKF Based Approach. In Proceedings of the International Conference of Intelligent Robot Systems, Hamburg, Germany, 28 September–3 October 2015. [CrossRef]
59. Qin, T.; Li, P.; Shen, S. VINS-Mono: A Robust and Versatile Monocular Visual Inertial State Estimator. *IEEE Trans. Robot.* **2018**, *34*, 1004–1020. [CrossRef]
60. Lynen, S.; Achtelik, M.W.; Weiss, S.; Chli, M.; Siegwart, R. A Robust and Modular Multi Sensor Fusion Approach Applied to MAV Navigation. In Proceedings of the International Conference of Intelligent Robot Systems, Tokyo, Japan, 3–7 November 2013. [CrossRef]
61. Faessler, M.; Fontana, F.; Forster, C.; Mueggler, E.; Pizzoli, M.; Scaramuzza, D. Autonomous, Vision Based Flight and Live Dense 3D Mapping with a Quadrotor Micro Aerial Vehicle. *J. Field Robot.* **2015**, *33*, 431–450. [CrossRef]
62. Forster, C.; Carlone, L.; Dellaert, F.; Scaramuzza, D. On Manifold Pre Integration for Real Time Visual Inertial Odometry. *IEEE Trans. Robot.* **2017**, *33*, 1–21. [CrossRef]
63. Kaess, M.; Johannsson, H.; Roberts, R.; Ila, V.; Leonard, J.; Dellaert, F. iSAM2: Incremental Smoothing and Mapping Using the Bayes Tree. *Int. J. Robot. Res.* **2012**, *31*, 216–235. [CrossRef]
64. Burri, M.; Nikolic, J.; Gohl, P.; Schneider, T.; Rehder, J.; Omari, S.; Achtelik, M.W.; Siegwart, R. The EuRoC MAV Datasets. *IEEE Int. J. Robot. Res.* **2016**, *35*, 1157–1163. [CrossRef]
65. Delmerico, J.; Scaramuzza, D. A Benchmark Comparison of Monocular Visual-Inertial Odometry Algorithms for Flying Robots. In Proceedings of the IEEE International Conference on Robotics and Automation, Brisbane, Australia, 21–25 May 2018. [CrossRef]
66. Mur-Artal, R.; Montiel, J.M.M. Visual Inertial Monocular SLAM with Map Reuse. *IEEE Robot. Autom. Lett.* **2017**, *2*, 796–803. [CrossRef]
67. Clark, R.; Wang, S.; Wen, H.; Markham, A.; Trigoni, N. VINet: Visual-Inertial Odometry as a Sequence-to-Sequence Learning Problem. In Proceedings of the AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017. Available online: <https://ojs.aaai.org/index.php/AAAI/article/view/11215> (accessed on 25 May 2023).
68. Paul, M.K.; Wu, K.; Hesch, J.A.; Nerurkar, E.D.; Roumeliotis, S.I. A Comparative Analysis of Tightly Coupled Monocular, Binocular, and Stereo VINS. In Proceedings of the IEEE International Conference on Robotics and Automation, Singapore, 19 May–3 June 2017. [CrossRef]
69. Song, Y.; Nuske, S.; Scherer, S. A Multi Sensor Fusion MAV State Estimation from Long Range Stereo, IMU, GPS, and Barometric Sensors. *Sensors* **2017**, *17*, 11. [CrossRef]
70. Solin, A.; Cortes, S.; Rahtu, E.; Kannala, J. PIVO: Probabilistic Inertial Visual Odometry for Occlusion Robust Navigation. In Proceedings of the IEEE Winter Conference on Applications of Computer Vision, Lake Tahoe, NV, USA, 12–15 March 2018. [CrossRef]
71. Houben, S.; Quenzel, J.; Krombach, N.; Behnke, S. Efficient Multi Camera Visual Inertial SLAM for Micro Aerial Vehicles. In Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Daejeon, Republic of Korea, 9–14 October 2016. [CrossRef]
72. Eckenhoff, K.; Geneva, P.; Huang, G. Direct Visual Inertial Navigation with Analytical Preintegration. In Proceedings of the IEEE International Conference on Robotics and Automation, Singapore, 29 May–3 June 2017. [CrossRef]

73. Negru, S.A.; Geragersian, P.; Petrunin, I.; Zolotas, A.; Grech, R. GNSS/INS/VO Fusion using Gated Recurrent Unit in GNSS-Denied Environments. *AIAA SciTech Forum* **2023**. [[CrossRef](#)]
74. Geragersian, P.; Petrunin, I.; Guo, W.; Grech, R. An INS/GNSS Fusion Architecture in GNSS-Denied Environment using Gated Recurrent Unit. *AIAA SciTech Forum* **2022**. [[CrossRef](#)]
75. Strasdat, H.; Montiel, J.M.M.; Davison, A.J. Real Time Monocular SLAM: Why Filter? In Proceedings of the IEEE International Conference on Robotics and Automation, Anchorage, AK, USA, 3–7 May 2010. [[CrossRef](#)]
76. Gallego, G.; Delbruck, T.; Orchard, G.; Bartolozzi, C.; Tabá, B.; Censi, A.; Leutenegger, S.; Davison, A.; Conradt, J.; Daniilidis, K.; et al. Event Based Cameras: A Survey. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *44*, 154–180. [[CrossRef](#)]
77. Mueggler, E.; Gallego, G.; Rebecq, H.; Scaramuzza, D. Continuous Time Visual Inertial Odometry for Event Cameras. *IEEE Trans. Robot.* **2018**, *34*, 1425–1440. [[CrossRef](#)]
78. Simon, D. *Optimal State Estimation*; John Wiley & Sons: Hoboken, NJ, USA, 2006; ISBN 0-471-70858-5.
79. Blanco, J.L. A Tutorial on SE(3) Transformation Parameterizations and On-Manifold Optimization. *arXiv* **2020**, arXiv:2103.15980. [[CrossRef](#)]
80. Gallo, E. High Fidelity Flight Simulation for an Autonomous Low SWaP Fixed Wing UAV in GNSS-Denied Conditions. C++ Open Source Code. 2020. Available online: [https://github.com/edugallogithub/gnssdenied\\_flight\\_simulation](https://github.com/edugallogithub/gnssdenied_flight_simulation) (accessed on 14 August 2023).

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.