

Experimentation of Monocular Visual-Aided Inertial Navigation on Fixed-Wing Unmanned Aerial Vehicle [†]

Baheerathan Sivalingam ^{*} and Ove Kent Hagen

Norwegian Defence Research Establishment (FFI), 2027 Kjeller, Norway; ove-kent.hagen@ffi.no

^{*} Correspondence: baheerathan.sivalingam@ffi.no[†] Presented at the European Navigation Conference 2023, Noordwijk, The Netherlands, 31 May–2 June 2023.

Abstract: This paper presents an experimental study utilizing the vision-aided inertial navigation system (VaINS) on a fixed-wing unmanned aerial vehicle (UAV). VaINS operates as a data-logging box, collecting sensor data, processing them, and transmitting the navigation solution to an operator terminal in real time. The flight data contain buildings, vegetation, a river, a lake, and mountains. The camera's pitch angle was tested at both a grazing angle of 22° and a nadir angle of 90°. The aim of this experiment was twofold: to explore the application of VaINS on a fixed-wing UAV and to assess the performance of the image-aided navigation system. The image-aided inertial navigation system (INS) solution was compared with a reference solution during GNSS-denied time intervals. The camera's pitch angle at 90° exhibited a slight advantage over the grazing angle at 22°. Moreover, in the linear flight segments, the image-aided navigation system demonstrated significantly better performance compared to the free inertial solution during the GNSS-denied periods.

Keywords: navigation; tracking; image feature points; INS; fixed-wing UAV

1. Introduction

Unmanned aerial systems (UAS) predominantly utilize the global navigation satellite system (GNSS) in conjunction with an INS for navigation. If the GNSS signal is denied, it results in an unacceptable position error, even over a brief time span. One potential solution to the problem is to use additional aiding sensors, e.g., LiDAR, radar, radio, or visual-based navigation [1]. In the VaINS, we use image-aided navigation, which can be implemented in various forms. A comprehensive literature review of the image-assisted methods for UAV navigation is presented in [2]. The techniques are categorized based on using a map as prior information, building the map while navigating as in visual simultaneous localization and mapping (SLAM) techniques, or operating without maps, such as with feature tracking and optical flow techniques.

The Norwegian Defense Research Establishment (FFI) has been investigating an image-aided inertial navigation system to cope with GNSS failure that is operating without maps. The system [3] is based on tightly integrating inertial sensor data with pixel position data of image feature points over an image sequence in an error-state extended Kalman filter (EKF).

The system's earlier versions underwent testing using image data obtained from a manned helicopter [4], an octocopter [3,5], and a hex-copter. The collected data were processed and analyzed offline. Later, a real-time navigation demonstration box called a visual-aided inertial navigation system (VaINS) was developed [6]. This VaINS box, as shown in Figure 1a, collects sensor data, processes them, and transmits the navigation solution to an operator terminal in real time.

An extensive experiment was conducted at Elvenes airstrip in the Salangen municipality area, Norway, using VaINS. It was used as a data-logging box on a fixed-wing UAV, Scout (Figure 1b) from NORCE [7]. NORCE was responsible for mounting the VaINS box and operating the aircraft.



Citation: Sivalingam, B.; Hagen, O.K. Experimentation of Monocular Visual-Aided Inertial Navigation on Fixed-Wing Unmanned Aerial Vehicle. *Eng. Proc.* **2023**, *54*, 42. <https://doi.org/10.3390/ENC2023-15476>

Academic Editors: Tom Willems and Okko Bleeker

Published: 29 October 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/>).

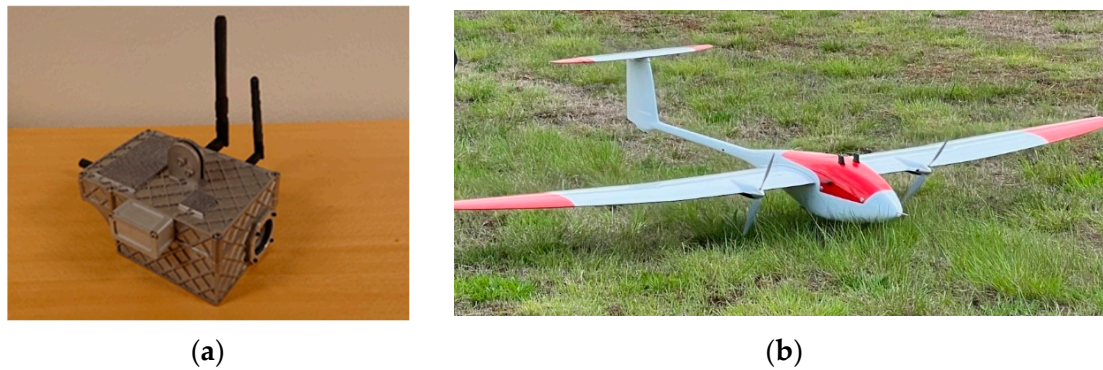


Figure 1. (a) A real-time navigation box named VaINS. (b) The Scout UAV.

There is related work on large-scale tests with image-aided navigation on UAVs. In [8], an autonomous route-following system of visual teach and repeat is coupled with a downward-facing camera on a fixed-wing UAV as a backup navigation when GNSS is unavailable. The system was evaluated using the data from the fixed-wing UAV flying a 1200 m trajectory at an altitude of 80 m. An optical-aided navigation method employing a camera and a LIDAR or radar is used in [9] when the GNSS is unavailable. It was tested with data sets from crewed helicopters and fixed-wing UAVs and achieved an accumulation error of 1–6% of the distance traveled, depending on the flight scenario.

This work herein builds on the technical report of the extensive experiment [10]. The purpose of this experiment was twofold: first, to explore the application of the VaINS box on a fixed-wing UAV, and second, to assess the performance of the image-aided INS. This document describes the developed image-aided inertial navigation system in brief, experimental setup, data collection, and results and summarizes the experiences and the performance.

2. Image-Aided Inertial Navigation System

The concept of the image-aided inertial navigation system is illustrated in Figure 2. An image feature point is the projection of a landmark, which is a static point in the scene, onto the image plane via a camera model. At each time step, the INS computes an estimate of the UAV's state, and this estimate is used together with the estimated landmark positions by the Kalman filter to predict the positions of the image feature points.

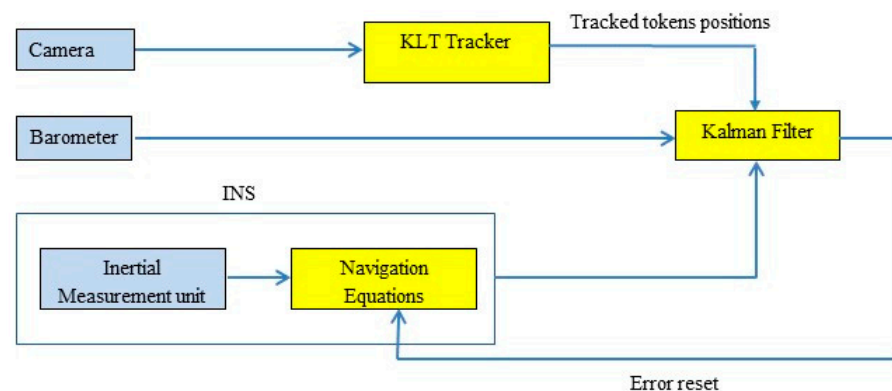


Figure 2. Image-aided inertial navigation system.

Image feature points are tracked by using the KLT algorithm [11]. The difference between the predicted and tracked positions are used as measurements by the Kalman filter, which estimates the INS errors. This process is repeated for each image time stamp. Details of the system are described in [3,5].

When using a monocular camera system, as in our case, the range to a tracked landmark is initially unknown and has to be guessed or externally estimated somehow. The

tight integration with the EKF, therefore, uses a variant of the inverse depth parameterization [3,12], which is robust to large initial range errors. With the combination of accelerated motion and feature tracking, these ranges become observable through the EKF [13], which results in a lower position error drift compared to a free inertial INS. There still is, however, a sensitivity to the level of error of our initial estimate for the range.

In [14], they suggest using a near range to the camera, setting the initial range to twice the near range with a 95% confidence level from the near range to infinity. Another strategy uses some initial feature measurements to pre-estimate the range before actually using them in the estimator [15]. In the tests herein, we estimated an effective near-range from coarse estimates of the UAV's height above ground and the camera view angle, assuming nothing was tracked in between.

3. Experimental Setup

VaINS consists of three sensors and two processing units. The sensors are the MTi-100 MEMS Inertial Measurement Unit (IMU), a u-blox GPS, and an embedded Basler da1600 monochrome camera. The processing units are a Jetson TX2 running the KLT-tracker and a TE0720 with a Xilinx Zynq 7020 running the INS on its CPU and interfacing and time stamping the sensors, including the camera through an FPGA. The technical documentation of VaINS is described in [6].

This experiment was a collaborative experiment involving both FFI and NORCE. NORCE was in charge of mounting the VaINS box on the aircraft (see Table 1 for aircraft specification) and planning the flight operations, while FFI took responsibility for devising the test plan, conducting data logging, and performing data analysis.

Table 1. Aircraft specification.

Parameter	Value
Cruise speed	15–20 m/s
Wingspan	2.62 m
Maximum endurance	3 h
Standard empty weight	4.3 kg
Batteries—Li-Ion	2.6 kg
MTOW	10 kg
Maximum useful load	3.1 kg

For practical implementation, the camera was securely mounted outside the VaINS box at the front of the payload bay, using a fixture that regulated various pitch angles. The VaINS box found its position at the back of the bay. After the aircraft was launched from a launching device, it was monitored from a mobile ground control station. More details are given in [10].

4. Data Collection

Data collection took place in the vicinity of an airstrip located in Salangen municipality, Norway. The flight trajectory remained consistent within a predefined area, viewing buildings, vegetation, a river, a lake, and mountains. The camera's pitch angle was set at a grazing angle of 22° and 90°.

The flight details are tabulated in Table 2. The initial segment of the flight altitude is approximately 80 m, while the altitude of the primary portion of the flight is around 180 m. The camera captures monochrome images at an 8-bit depth, with a frame rate of 10 Hz and a resolution of 1600 × 1200 pixels. Figure 3 shows an example of an image for each pitch angle of the camera.

Table 2. Flight details.

Flight #	Camera's Pitch	Time after Launch (s)	Total Time (s)
1	22°	903	1653
2	90°	742	992

**Figure 3.** Example image from each pitch angle. (a) 22° (b) 90°.

Flight time, after launch, for the flights was around 16–28 min. However, data collection stopped around 12–15 min after launch due to a system setup error.

5. Results

An INS post-processing tool developed by FFI, NavLab4 [16], is used to process and analyze the data from VaINS offline and gives an indication of the expected real-time performance since the INS algorithms in VaINS are based on the implementation in NavLab4. The aircraft's reference position and orientation are determined using the optimal smoother in NavLab4, which utilizes IMU and GPS data for the whole flight. To test the image-aided INS, image feature point data (12 points at 10 Hz) is combined with the IMU data and GPS height during the simulated GPS-denied intervals. Alternatively, barometer height measurements from NORCE can replace GPS height.

The performance of the image-aided INS solution is assessed by comparing it separately with the reference solution during linear flight (straight-line) and banking maneuvers. The root-mean-square (RMS) error between the image-aided INS and the reference solution at the end of each GPS-denied interval serves as the evaluation metric. For the straight-line portions, the comparison considers both the percent of distance traveled (%DT) and the average speed error (ASE), while for the banking portions, only the ASE is taken into account.

From the initial analysis in NavLab4, when including all the sensors, we identified and coarsely estimated a significant misalignment of the camera with respect to the IMU [10]. The camera and the box were remounted prior to each flight and without guaranteeing the same mounting alignment. Thus, we estimated the misalignment for each flight.

5.1. Flight 1

The camera's pitch angle of this flight is 22 degrees, and the flight time after launching is 903 s. This flight has four linear flights and four banking maneuvers.

Figure 4 presents the trajectory of the first two linear flights and the four banking maneuvers. The image-aided navigation when GPS data are disabled is presented with a green line. The red line presents the reference trajectory. The rest of the maneuvers are similar to the presented ones.

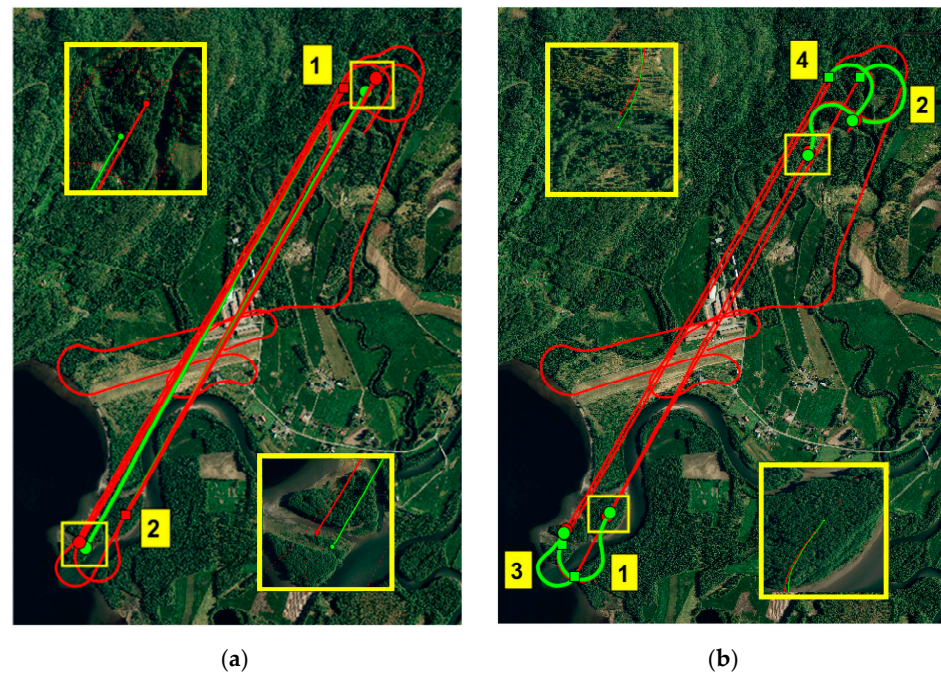


Figure 4. The trajectories of the first flight with image-aided navigation when GPS data are disabled (green line) and the reference trajectory (red line). (a) Linear flights (about 2400 m) (b) and banking maneuvers.

Tables 3 and 4 tabulate the duration, distance traveled, and drift compared to the reference data in the straight-line parts and banking parts, respectively. When comparing the drift (ASE), the drift in linear flights is slightly lower than the drift in banking maneuvers.

Table 3. Duration of image-aided navigation (flight 1) when GPS is disabled, distance traveled within the duration, and the drift in RMS error, in ASE, and in %DT.

Duration (s)	Distance (m)	RMS (m)	ASE (m/s)	%DT
136	2439	37.6	0.28	1.54
130	2338	82.4	0.63	3.52
122	2231	50.3	0.41	2.25
133	2411	56.7	0.43	2.35

Table 4. Duration of image-aided navigation (flight 1) and the drift in RMS error and ASE in banking maneuvers.

Duration	RMS (m)	ASE (m/s)
33	18.4	0.56
53	14.1	0.27
26	18.9	0.73
24	15.0	0.63

5.2. Flight 2

Flight 2 has the camera’s pitch angle at 90° nadir, and flight time after launching is 742 s. This flight has three linear flights and two banking maneuvers. The first two trajectories of the linear flights and the banking maneuvers are shown in Figure 5.

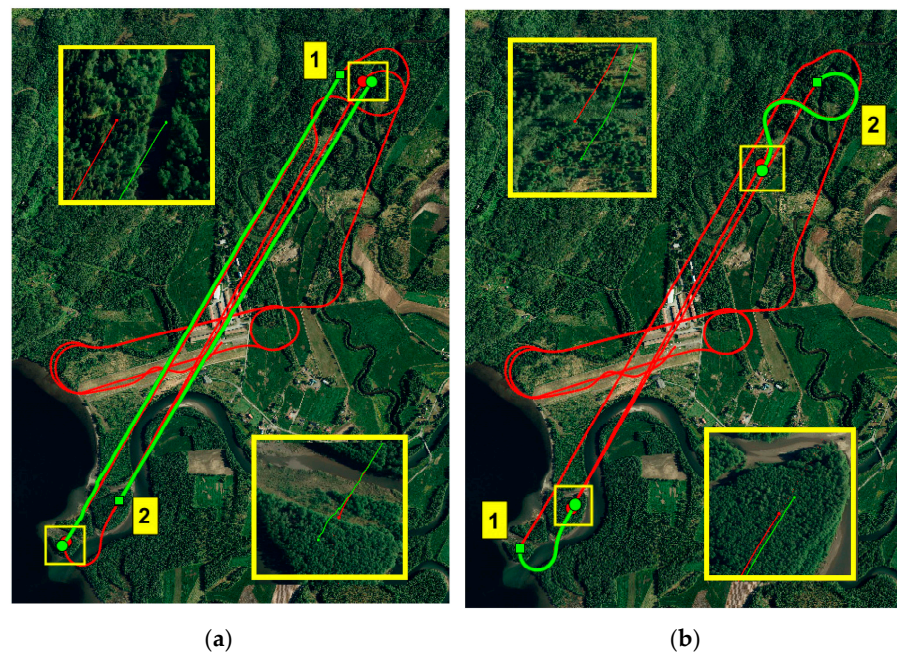


Figure 5. The trajectories of the second flight with image-aided navigation (green line) and the reference trajectory (red line). (a) Linear flights (about 2400 m) (b) and banking maneuvers.

The green line represents the image-aided navigation when the GPS is disabled, and the red lines represent the reference trajectories. Table 5 tabulates the duration, distance traveled, and the drift of the linear flights, and Table 6 tabulates the drift data for the banking maneuvers. In the linear parts, the ASE is slightly lower than the banking parts.

Table 5. Duration of image-aided navigation (flight 2), distance traveled within the duration, and the drift in RMS, in ASE, and in %DT.

Duration (s)	Distance (m)	RMS (m)	ASE (m/s)	%DT
158	2576	20.0	0.13	0.78
123	2310	36.9	0.30	1.60
67	1090	20.5	0.31	1.88

Table 6. Duration of image-aided navigation (flight 2) and the drift in RMS error and in ASE.

Duration	RMS (m)	ASE (m/s)
30	16.3	0.54
57	24.4	0.43

5.3. Free Inertial

Within the simulated GPS-denied time intervals, the performance of free inertial is analyzed using the RMS error and ASE for straight-line laps and for the banking laps separately (see Tables 7 and 8).

In the linear parts, the free inertial ASE ranges from 0.20 to 8.31 m/s, while image-aided navigation ASE ranges from 0.13 to 0.73 m/s. In the banking parts, however, the results are more similar: from 0.37 to 5.32 m/s for free inertial and from 0.27 to 0.73 m/s for image-aided navigation.

Table 7. The RMS error and ASE at the end of the GPS-denied intervals for straight-line laps under free inertial.

Flight	Lap	Duration (s)	Distance (m)	RMS (m)	ASE (m/s)
1	1	136	2439	340	2.50
	2	130	2338	855	6.58
	3	122	2231	627	5.14
	4	133	2411	1105	8.31
2	1	158	2576	960	6.08
	2	123	2310	26	0.20
	3	67	1090	59	0.88

Table 8. The RMS error and ASE at the end of the GPS-denied intervals for banking laps under free inertial.

Flight	Lap	Duration (s)	RMS (m)	ASE (m/s)
1	1	33	68.9	2.12
	2	53	39.2	0.74
	3	26	9.5	0.37
	4	24	10.3	0.43
2	1	30	159.7	5.32
	2	57	39.5	0.69

6. Discussion

The results obtained from flights 1 and 2, as discussed here, are indicative of the other flights in the experiment. For further details, refer to [10]. It was observed that the image-aided INS exhibited slightly larger drift during banking maneuvers compared to linear flights. This phenomenon can be attributed to the image feature points being selected from distant regions in the scene during banking maneuvers due to the combination of low pitch angle and high roll angle. Consequently, these feature points rapidly disappeared from the camera's field of view, leading to increased drift in the image-aided navigation.

When comparing the two pitch angles (90° nadir and 22° grazing), it was observed that the nadir configuration performed slightly better than the grazing configuration, as depicted in Figure 6. The advantage of the nadir configuration lies in the utilization of the entire image scene, with similar ranges to all landmarks, and the ease of obtaining a good initial estimate of altitude. However, a drawback is that the image feature points are not tracked for an extended period.

On the other hand, the grazing configuration has certain downsides, such as a portion of the image above the horizon that is never used, resulting in significant variation in the range to landmarks and making it more challenging to establish a good initial estimate. To address this, an approach is proposed to dynamically estimate the initial range from the first few frames of each image-feature-point track before initializing it in the EKF, as described in [13]. The advantage of the grazing configuration is the potential for tracking landmarks over a longer duration. Nonetheless, during banking maneuvers, almost all landmarks are lost.

Comparing image-aided navigation with free inertial, image-aided navigation gives a great overall improvement in linear flights. Over greater time intervals, the free inertial error drift becomes non-linear and increases faster, as seen in the long straight-line parts. In the banking turns, there are two effects that make the free inertial error lower. The time period is shorter, so the non-linear parts of the error drift are less, and errors may cancel due to the turn. In these parts, the image-aided navigation seems to primarily introduce increased robustness (see Figure 6).

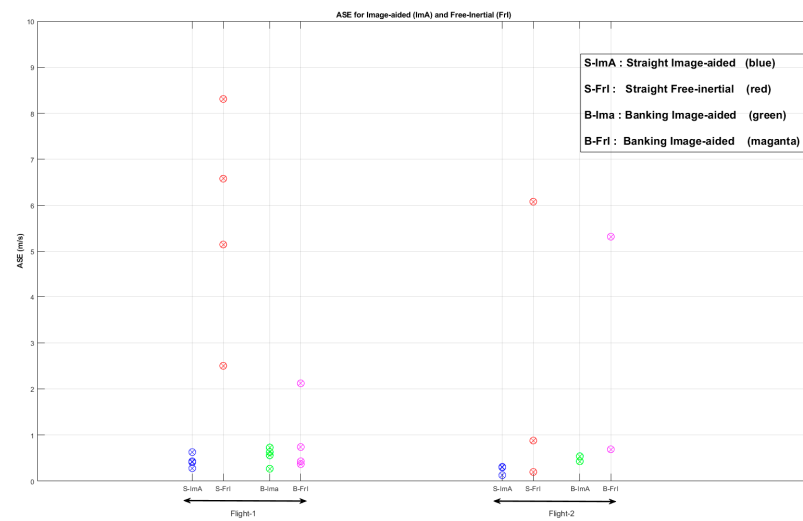


Figure 6. Comparison of ASE between image-aided INS and free-inertial for flight 1 and flight 2 under straight-line and banking laps.

The navigation solution is sensitive to the alignment between the IMU and the camera. Thus, the current installation, with an unknown misalignment for each flight, is not ideal. An improvement is to fix the IMU and the camera on a bracket placed outside the VaINS box. By doing so, the alignment can be estimated once, streamlining the setup process.

7. Conclusions

This document outlines an experiment involving the VaINS box. The VaINS box provides a navigation solution in real time. It integrates inertial sensor data with position data of image feature points over an image sequence employing an error-state extended Kalman filter. This experiment's aim was to explore the application of the VaINS box on a fixed-wing UAV and to assess the image-aided navigation solution.

The data were collected from flights with camera pitch angles of 22° and 90° and viewing buildings, vegetation, a river, a lake, and mountains. The aircraft's reference navigation solution is calculated using the optimal smoother, utilizing IMU and GPS data. For the image-aided navigation solution, image-feature-point data are combined with IMU data and GPS height during GPS-denied intervals. The flights with the camera pitched at the nadir demonstrated an overall better performance compared to the flights with a grazing angle. The drift during banking maneuvers was slightly larger than during straight-line maneuvers. During banking maneuvers, the image-aided navigation performed slightly better than the free inertial navigation within the same GPS-denied time intervals. Remarkably, it performed significantly better during straight-line maneuvers.

In this experiment, the IMU and the camera were loosely mounted together, which led to additional uncertainty in misalignment. This could be solved by fixing the camera and IMU on a rigid fixture out of VaINS.

Author Contributions: B.S. and O.K.H. together contributed in all aspects of the article. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available when requested.

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

References

1. Gyagenda, N.; Hatilima, J.V.; Roth, H.; Zhmud, V. A review of GNSS-independent UAV navigation techniques. *Robot. Auton. Syst.* **2022**, *152*, 104069. [[CrossRef](#)]
2. Lu, Y.; Xue, Z.; Xia, G.; Zhang, L. A survey on vision-based UAV navigation. *Geo-Spat. Inf. Sci.* **2018**, *21*, 21–32. [[CrossRef](#)]
3. Baheerathan, S.; Hagen, O.K. Image-aided inertial navigation for an Octocopter. In Proceedings of the SPIE Defense + Security 2018, Orlando, FL, USA, 3 May 2018.
4. Baheerathan, S.; Hagen, O.K. Image-aided inertial navigation system based on image tokens. In Proceedings of the SET-168 NATO Symposium on Navigation Sensors and Systems in GNSS Denied Environments, Imzir, Trucky, 8–9 October 2012.
5. Baheerathan, S.; Hagen, O.K. Experiment of Image-aided inertial navigation with multiple cameras in Octocopter. In Proceedings of the SPIE Defense + Security 2020, Anaheim, CA, USA, 26–30 April 2020.
6. Baheerathan, S.; Hagen, O.K. *ValNS Technical Documentation*; FFI Report 22/01877; Norwegian Defense Research Establishment: Kjeller, Norway, 2022.
7. NORCE Norwegian Research Center. Available online: <https://www.norcereasearch.no> (accessed on 15 August 2023).
8. Warren, M.; Paton, M.; MacTavish, K.; Schoellig, A.P.; Barfoot, T.D. *Towards Visual Teach and Repeat for GPS-Denied Flight of a Fixed-Wing UAV, Field and Service Robotics*; Springer: Berlin/Heidelberg, Germany, 2018; pp. 481–498.
9. Andert, F.; Ammann, N.; Krause, S.; Lorenz, S.; Bratanov, D.; Mejias, L. Optical aided aircraft navigation using decoupled visual SLAM with range sensor augmentation. *J. Intell. Robot. Syst.* **2017**, *88*, 547–565. [[CrossRef](#)]
10. Baheerathan, S.; Hagen, O.K. *Experimentation of Vision-Aided Inertial Navigation System on a Small Fixed-Wing UAV*; FFI Report 23/00148; Norwegian Defense Research Establishment: Kjeller, Norway, 2023.
11. Thomasi, C.; Kanade, T. *Detection and Tracking of Point Features*; Tech. Rep. CMU-CS-91-132; Carnegie Mellon University: Pittsburg, PA, USA, 1991.
12. George, M.; Sukkariéh, S. Inertial navigation aided by monocular camera observations of unknown features. In Proceedings of the 2007 IEEE International Conference on Robotics and Automation, Rome, Italy, 10–14 April 2007.
13. Martinelli, A. Vision and IMU data fusion: Closed form solutions for attitude, speed, absolute scale and bias determination. *IEEE Trans. Robot.* **2012**, *28*, 44–60. [[CrossRef](#)]
14. Civera, J.; Davison, A.J.; Montiel, J.M.M. Inverse depth parametrization for monocular SLAM. *IEEE Trans. Robot.* **2008**, *24*, 932–945. [[CrossRef](#)]
15. Munguia, R.; Antoni, G. Delayed inverse depth monocular SLAM. *IFAC Proc.* **2008**, *41*, 2365–2370. [[CrossRef](#)]
16. Gade, K. NAVLAB: A generic simulation and post-processing tool for navigation. *Eur. J. Navig.* **2004**, *2*, 51–59. [[CrossRef](#)]

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.