

Addressing GNSS Vulnerabilities in AAM: A Multi-Modal UAV Testbed for Redundant and Reliable Navigation

Felix Rocco Matzke¹, Tushar Jayesh Barot², Shreeja Sridharan³, Nikolaus Ammann⁴,
Christoph Kessler⁴, Hans-Gerd Maas¹, Frank H. P. Fitzek³, Anette Eltner¹

¹Institute of Photogrammetry and Remote Sensing, ²Institute of Logistics and Aviation,

³Deutsche Telekom Chair of Communication Networks, ^{1,2,3}Dresden University of Technology, Germany

⁴Institute of Flight Systems, German Aerospace Center (DLR)

(felix_rocco.matzke, tushar_jayesh.barot, shreeja.sridharan, hans-gerd.maas, frank.fitzek, anette.eltner)@tu-dresden.de,
(nikolaus.ammann, christoph.kessler)@dlr.de

Keywords: Uncrewed Aerial Vehicle (UAV), Multi-modal Navigation, Advanced Air Mobility (AAM), SLAM, unified map representation, sensor calibration

Abstract

Advanced Air Mobility (AAM) aims to revolutionize transportation by enabling on-demand, low-altitude movement of people and goods using Vertical Take-Off and Landing (VTOL)-capable vehicles (VCAs), but achieving precise and reliable navigation remains a key challenge. The navigation of UAVs typically relies on the integration of GNSS and IMU measurements. When GNSS positioning is degraded or unavailable due to factors such as jamming, spoofing, and multipath effects, these approaches become unreliable. Perception-based methods, such as SLAM based on a combination of RGB camera, event camera, LiDAR, and IMU can lead to resilient navigation with loop closure for drift correction. However, loop closure is often impractical for a point-to-point AAM flight. To tackle these challenges, this ongoing research work proposes a multi-modal navigation framework that integrates map-based navigation and semantic SLAM. For the development of this framework, we designed a conceptual setup, which will be presented in this paper. Additionally, we focus on the selection of sensors, their spatio-temporal calibration, and their integration into a holistic onboard system.

1. Introduction

Advanced Air Mobility (AAM)¹ is a comprehensive concept introduced by the National Aeronautics and Space Administration (NASA) that signifies a transformative evolution in aviation. It centers on the integration of next-generation aircrafts, including remotely piloted and autonomous Vertical Take-Off and Landing (VTOL)-capable vehicles. These aircrafts, collectively known as VTOL-Capable Aircrafts (VCAs)², are broadly categorized into two groups: Standard-Performing Vehicles (SPVs) - typically small Uncrewed Aerial Systems (UASs), and High-Performing Vehicles (HPVs) - such as air taxis and larger cargo aircraft (EASA, 2024b, Altun et al., 2023). AAM's advanced aircraft technologies redefine aviation and create efficient air services that connect diverse communities.

AAM aims to provide consumers with access to on-demand air travel, package delivery, and emergency services by leveraging a connected, multi-modal transportation network (Goyal et al., 2021). It is designed to operate in variety of environments, including urban, suburban, and rural - with local missions typically ranging within 50 miles (80 km), and regional missions extending to several hundred miles (Goyal et al., 2021). Urban Air Mobility (UAM)³ and Regional Air Mobility (RAM)⁴ are both key components under the broader AAM umbrella. The overarching vision of AAM is to develop a safe, accessible, and sustainable air transportation ecosystem capable of reaching

areas that are currently under-served or congested by traditional transport modes. Realizing this vision requires collaborative efforts across industry, academia, and regulatory authorities to advance enabling technologies and infrastructure (Xiang et al., 2024).

AAM is reshaping how low-altitude airspace is utilized and managed (Konert and Kasprzyk, 2024). Traditionally, airspace below 3,000 m served as a transition zone, mainly near airports during takeoff and landing. AAM use cases now operate directly within this layer. A key focus is the Very Low Level (VLL) airspace—up to 150 m above ground level (Pongsakornsathien et al., 2025). Its close proximity to people and ground infrastructure adds further safety and regulatory challenges, especially for autonomous systems (EASA, 2021a). To maximize the capacity of the sub-120 m airspace, tighter separations between UAS are desirable. Classifying UAS under traditional Visual Flight Rules (VFR) or Instrument Flight Rules (IFR) becomes less relevant in this context (Konert and Kasprzyk, 2024). Additionally, the existing technological infrastructure of traditional aviation faces challenges in this new and highly complex operating zone. Furthermore, given the increased risk due to high proximity to terrestrial infrastructure, dependence on radio signals prone to shadowing can make the system vulnerable. Thus, to realize safe autonomous UAS operations, there is a need for a real-time and onboard navigation system with sufficient redundancy (Siciliano and Khatib, 2016, p. 660).

A redundant onboard navigation system must therefore not only ensure operational reliability in the face of potential failures, but also align with the evolving regulatory emphasis on performance-based criteria. Civil aviation has evolved over the decades with radio, inertial, and satellite navigation, leading to a regulation

¹ AAM: <https://www.nasa.gov/mission/aam/>

² VCA: <https://www.easa.europa.eu/en/document-library/easy-access-rules/easy-access-rules-small-category-vca>

³ UAM: <https://www.easa.europa.eu/en/what-is-uam>

⁴ RAM: <https://sacd.larc.nasa.gov/ram/>

that focuses on the required performance of the used system. Performance Based Navigation (PBN) enables using multiple navigation systems, until the metrics like availability, continuity, integrity, and accuracy are met for 95% of the flight time (ICAO, 2012). For UAS operations in urban scenarios, radio navigation with ground beacons is not an option due to the upcoming obsolescence, required obstacle clearance or altitude, and insufficient ground beacons in urban areas (EUROCONTROL, 2021). Inertial navigation remains a viable but only short-term positioning source. Global Navigation Satellite System (GNSS)-based satellite navigation is susceptible to spoofing, jamming, and space weather effects (EASA, 2021b, EASA, 2024a). Additionally, in UAM operations, urban canyons have a detrimental effect on GNSS accuracy due to multipath propagation and satellite occlusion. PBN provides a framework to achieve the required navigation performance, independent of the navigation system used. However, the accuracy and integrity requirements should be adapted for the dense AAM airspace.

In order to achieve the adapted requirements, GNSS positioning performance enhancements play a crucial role for safe operations in complex AAM airspace. GNSS positioning is improved when receivers can access multiple satellite constellations (e.g., GPS, GLONASS, Galileo, BeiDou) and multiple frequency bands (e.g., L1, L2, L5) (Zidan et al., 2020). The application of differential GNSS (DGNSS), which uses pseudo-range corrections, can further improve the accuracy to a submeter-level. Real-Time Kinematics (RTK) positioning, a more advanced differential technique, is able to provide centimeter-level accuracy by resolving the carrier phase ambiguities. However, RTK positioning depends on continuous correction streams from a local base station or a Continuously Operating Reference Station (CORS) network. These corrections must be transmitted to the aircraft via a reliable communication link (e.g., radio or cellular), and any interruption or latency in this link can limit overall usability and degrade real-time performance (Nex et al., 2022).

An independent (isolated) navigation system faces challenges in AAM scenarios. This research addresses the limitations of traditional GNSS and inertial navigation systems by incorporating perception-based solutions, such as Simultaneous Localization and Mapping (SLAM) and map-based navigation, which are effective in complex flight environments. We propose a multi-modal navigation system to improve accuracy, redundancy and operational resilience to promote safe and reliable autonomous flight in the AAM ecosystem. By leveraging a robust wireless communication system in dense airspace, we fulfill multi-modal navigation system requirements for Uncrewed Aerial Vehicle (UAV) and enable effective ground-based monitoring.

The remainder of this paper is organized as follows. Firstly, Section 2 outlines the background of perception-based navigation along with our contributions to the field. Section 3 discusses in detail the sensor choices for the multi-modal navigation systems, highlighting the advantages and limitations of each modality. The integration strategy, calibration procedures, and the design of our conceptual setup are detailed in Section 4. Section 5 concludes with a concise discussion and potential future steps of the multi-modal navigation system for AAM operations.

2. Background: Perception-based Navigation

Safe UAV operations demand a perception-based onboard navigation system for localization and environment sensing (Siciliano and Khatib, 2016). Navigation systems, absolute or relative,

depend on their reference coordinate system, e.g. GNSS uses the Earth's coordinate system (namely WGS84) as an absolute reference. Conversely, dead reckoning approaches like inertial, visual, or LiDAR odometry deliver position estimates relative to the known start position. However, relative localization approaches suffer from the estimates drifting over time due to the accumulated errors, which are commonly corrected using an absolute position estimate. In this paper, we discuss real-time onboard navigation approaches for the multi-modal UAV testbed with map-based navigation and semantic SLAM.

2.1 Simultaneous Localization and Mapping (SLAM)

SLAM provides a navigational foundation for higher-level task planning in unknown environments where absolute localization methods are degraded or unavailable. Unlike odometry, SLAM incorporates a globally optimized map and loop closure to correct drift. With the created map, loop closure can alleviate odometry drift errors by detecting and validating location revisits, ensuring consistency in observations. The SLAM front-end estimates relative poses using the sensor observations, while the back-end encompasses loop closure detection as well as mapping (Cadena et al., 2016). Depending on the combination of sensors, some pivotal SLAM algorithms are discussed below.

Visual (inertial) SLAM (VSLAM) algorithms combine monocular (or stereo, depth) cameras with an IMU for increased robustness, notably VINS-Mono (Qin et al., 2018) and ORB-SLAM3 (Campos et al., 2021). Using feature correspondences or pixel intensities for pose estimation and tracking, VSLAM provides good performance with enough illumination, even on-line on a Central processing Unit (CPU). However, they do not perform well in low illumination, long-term scenarios, or dynamic environments due to tracking failure and scale drift (Macario Barros et al., 2022).

The forte of LiDAR SLAM algorithms lies in the accurate distance measurements to map the geometry of the environment and due to active sensing, LiDAR-based approaches outperform visual methods in low illumination environments. LiDAR-based PIN-SLAM (Pan et al., 2024) is robust to slow-moving dynamic objects and KISS-SLAM (Guadagnino et al., 2025) can sustain in large outdoor scenarios, while accurately running in real-time on a CPU. However, rapid motion and environments with sparse or repetitive structure still pose a challenge (Yin et al., 2021).

In contrast to visual-only approaches, LiDAR-Visual (Inertial) SLAM approaches have performed well in large outdoor scenarios, profiting from LiDAR's accurate distance measurements combined with texture information from the camera. This multi-modal information helps to withstand unstructured or feature-poor scenarios, as shown with LVI-SAM (Shan et al., 2021) and FAST-LIVO2 (Zheng et al., 2025). In this study, depth cameras, stereo cameras and also monocular depth estimation are not considered as they are limited by their range for AAM operations. Furthermore, LiDAR-Visual approaches are promising in large-scale outdoor scenarios, but they are ill-suited for high-speed flights and fast dynamic objects (Macario Barros et al., 2022, Cheng et al., 2025).

The SLAM approaches developed in the last decade depended on features, texture, or structure between data samples, also known as short-term data association. This works on the assumption of having a sufficiently static environment for reliable tracking, while the dynamic objects are discarded or tracked (Cadena et al., 2016). Using deep learning-based segmentation models, the

semantic information about object can add meaningful context for odometry, loop closure (Pan et al., 2024, Conclusion) and also help to deal with dynamic objects (Chen et al., 2019, Our Approach), giving rise to semantic SLAM approaches. Thus, the shift in focus, from feature or geometry-only methods towards semantics aligned with human perception, is crucial for future SLAM implementations.

Event cameras are bio-inspired, neuromorphic sensors that function like the human retina. Rather than frames of a common RGB camera, it delivers asynchronous pixel-level event streams for brightness changes at a high temporal rate of microseconds. It brings immense potential to perception-based systems due to its temporal resolution, high dynamic range (alleviates under/over-exposure) and resilience to motion blur (Gehrig and Scaramuzza, 2024). UltimateSLAM (Vidal et al., 2018) demonstrates the potential of frame-event SLAM pipelines in flight scenarios to combat large illumination differences and enable high-speed flights. Thus, the UAV testbed in this paper is designed with the combination of LiDAR, frame-based RGB camera, event camera, and Inertial Measurement Unit (IMU) for semantic SLAM.

2.2 Map-based Navigation

In AAM scenarios, the drone will point-to-point fly as fast as possible to its destination. A loop closure for drift correction is therefore difficult to realize. Drift correction with GNSS measurements is one solution, but is not always available. Map-based navigation aims to obtain a global position or pose by localizing local sensor observations to a map of the environment, similar to how we would try to match what we perceive with an analog, digital or internal map. This requires a prior map of the scene, which can be 2 D or 3 D. While many approaches require a rough initial pose estimate, (Kinnari et al., 2023) propose a solution that works from an uncertain prior position. They matched orthoprojected UAV images to satellite imagery by transforming the UAV observations and satellite image patches into descriptor vectors. Through these transformations, their method also works even when significant seasonal changes are present. Other approaches transform 2 D geodata such as OpenStreetMap maps (Sarlin et al., 2023a), Street View images or aerial orthophotos (Sarlin et al., 2023b) into neural maps to enable cross-modal alignment. Other methods use digital elevation models as a 2 D map representation so that corresponding topographic features must be found (Wan et al., 2022).

These approaches determine the pose only in SE(2). Other approaches try to determine the pose in SE(3) through the use of 3D map representations. The authors (Merat et al., 2024) use the point-to-plane ICP algorithm to find the transformation that aligns a local point cloud created with their VSLAM system and a global point cloud representing the environment. Instead of aligning sensor observations and the map in 3D space, (Zhang et al., 2024) render RGB and depth images from a 3D colored map represented in the form of a mesh, point cloud or NeRF to establish 2D-3D correspondences. The rendered images are converted into global descriptors, which are stored in a visual database to enable fast image retrieval. Once an image is found covering the same part of the scene, matching is done in 2D while the corresponding depth images provide the necessary 3D information to determine the pose via a PnP algorithm. A similar concept is used by (Vultaggio et al., 2024), who render images of aerial meshes to find corresponding points and additionally use GNSS positioning as a pose prior. Integration with SLAM (Merat et al., 2024) would help to narrow down the search space and accelerate these approaches.

2.3 Need for a UAV Testbed

The motivation for the UAV testbed arises from the lack of a dataset for AAM navigation which ideally covers:

1. Very Low Level (VLL) perspective in dynamic urban scenarios (<120 m)
2. The complete sensor suite (RGB and event camera, LiDAR, IMU, RTK-GNSS) and
3. Semantic annotations.

Another motivation is that the real-time system performance can be evaluated using the onboard computing resources. Ample of datasets are available for the autonomous driving domain or for ground robots with a wide variety of sensors, particularly LiDAR, multi-view cameras, IMU, and GNSS; like KITTI (Geiger et al., 2013), nuScenes (Caesar et al., 2020). Also, M2DGR (Yin et al., 2021) and VECtor (Gao et al., 2022) include the event camera, but these datasets only have the sensor suite and lack (1) and (3). Most datasets with UAVs are indoors or close to the ground, which lacks the required aerial perspective for AAM, like EuRoC MAV (Burri et al., 2016) or UZH FPV (Delmerico et al., 2019). To enable more research with semantic labels and panoptic tracking, Cityscapes (Cordts et al., 2016) and Panoptic Nuscenes (Fong et al., 2022) exist for autonomous driving ((2) unsatisfied), whereas UAVid (Lyu et al., 2020) and Ruralscapes (Marcu et al., 2019) don't include the event camera (see (2)). To train segmentation models without needing a large number of real images, (Hinniger and Rüter, 2023) demonstrate how models trained on a synthetic dataset, created using a game engine, require a small set of real images to eliminate the domain gap with fine-tuning. Our UAV testbed aims to facilitate algorithmic development, real-time onboard evaluation, and to generate a small, real-world dataset for fine-tuning.

2.4 Research Contribution

Safe operation of the UAS in the VLL airspace relies on the precise positioning performance of the underlying navigation system. The system must be sophisticated and resilient to enable autonomous flight operation. The contributions of this paper are as follows:

1. We highlight the existing navigation challenges of AAM pertaining to complex VLL airspace. Then, we analyse the difficulties involved in selecting suitable sensors for perception-based navigation. A comprehensive overview is presented, along with key parameters that meet the performance requirements.
2. We propose a reliable, spoof-resistant multi-modal navigation architecture along with an associated methodology for integration and calibration.
3. We address the communication and synchronization requirements for UAS operations encompassing reliable data exchange and coordination in dense AAM environments.

3. Design of Multi-Modal Sensor Setup

The impact of sensor selection on the fidelity of available measurements and consequently the algorithm efficacy, cannot be

Device	Type	Specifications	Freq. (Hz)	Weight (g)
RGB camera	Sony IMX304 (16 mm C-Mount lens)	12.3 MP, 3.45 μm , 48.2° × 36.2°, 71.4 dB	22.6	90 (+89.8)
Event camera	Sony, Prophesee IMX636 (6 mm C-Mount lens)	0.9 MP, 4.86 μm , 54.6°, >120 dB	1.06 giga events	90 (+67)
LiDAR	Hesai XT32M2X	Mechanical, 32-Channel, 360° × 40.3°, 0.5 to 300 m (80 m @10% reflectivity)	20	490
IMU (Pixhawk 6X Pro)	ADIS16470	$\pm 2000^\circ/\text{sec}$, 8°/hr, 13 μg , $\pm 40\text{g}$	250	50
RTK GNSS Receiver	DroneCAN H-RTK F9P Helical	Galileo, BeiDou, GLONASS, GPS (L1/L2), 1.5 m (3D FIX) / 0.01 m (RTK)	20 (raw), 8 (RTK)	58
RTK GNSS-aided INS	Advanced Navigation Certus Mini D	Galileo, BeiDou, GLONASS, GPS (L1/L5), 0.1° Roll, Pitch and Heading	1000 (processed)	55

Table 1. Specifications of onboard sensors.

overstated. RGB cameras offer high spatial resolution and mature semantic models, but lack direct depth sensing, are sensitive to illumination, have a low temporal resolution, and introduce scale ambiguity. Conversely, event cameras excel in high dynamic range and temporal resolution with sparse data throughput, but offer low spatial resolution compared to frame-based RGB cameras. These complementary features make fusion with these two sensors appealing. LiDAR sensors provide accurate distance measurements and are, due to their active sensing, independent of illumination. Drawbacks entail the sparsity, the low sample rate, and the algorithmic challenges in scenes with repetitiveness and few structures. While all the sensors mentioned above are exteroceptive, IMUs are proprioceptive sensors that provide high-frequency updates for ego-motion estimation, but accumulate drift errors over time.

As pointed out each sensor and the associated approaches have their weaknesses. Therefore, we seek a multi-modal (multi-sensor) approach which fuses data from cameras, LiDAR, IMU, and GNSS. While well-engineered commercial-off-the-shelf drones are increasingly available, they are not designed to integrate this multitude of sensors, nor are they designed to process large volumes of data in real-time directly onboard. However, designing UAVs entirely from scratch is a complex and time-intensive endeavour. As a practical compromise, we augment commercially available development kits with additional peripherals, enabling both flexibility in sensor integration (selection and placement) and sufficient onboard computational capabilities. The selected sensors and their key specifications are summarized in Table 1. Hardware choices were made in accordance with the Size, Weight, and Power-Cost (SWaP-C) principle to account for payload constraints and to increase the achievable flight time.

For the RGB camera, a balanced compromise was sought between resolution and frame rate to facilitate vision-based navigation. While a higher resolution provides more image detail and increases the number of detectable features, it also imposes greater demands on data bandwidth and processing, and therefore limits the frame rate, which is detrimental in high-speed UAV operations. Additionally, there is a tradeoff between resolution and pixel size. Larger pixels improve the signal-to-noise ratio and enable shorter exposure times, which are advantageous in dynamic flight scenarios but in turn reduce the resolution. The eventually chosen camera features a fixed 16 mm focal length lens. The resulting Field of View (FOV) is relatively narrow, which reduces scene coverage and feature overlap but the larger focal length leads to a sufficiently high resolution for distant

objects. If the camera with the given specifications is mounted in nadir direction, then for a flight height of 50 m, the ground sampling distance on a flat surface would be approximately 1.1 cm and for the maximum allowed flight height of 120 m, according to European Union Aviation Safety Agency (EASA)’s open category, 2.6 cm. Furthermore, the camera has a global shutter and a moderate dynamic range.

The event camera was primarily chosen for its high resolution compared to the popular DAVIS cameras in the literature. Furthermore, the event camera as well as the RGB camera have Power over Ethernet (PoE) and Robot Operating System (ROS2) support, while only the latter applies to the chosen mechanical LiDAR. Additionally, the LiDAR is lightweight, has a moderate number of channels, offers a range that suffices for flight heights up to 120 m, and therefore offers a good price-performance ratio. The main drawback is the missing internal IMU, which increases the requirements for an accurate LiDAR-IMU calibration. For this, we use the embedded IMUs of the Pixhawk flight controller. More specifically, we used the Pixhawk 6X Pro, which features a high-quality miniature vibration-isolated and temperature-controlled MEMS IMU. It has a wide dynamic range and medium bias drift for the triaxial gyroscope and accelerometer. With the built-in u-blox ZED-F9P module, the RTK GNSS receiver offers concurrent reception of multiple constellations and dual-frequency RTK. Due to the helical antenna, the receiver is lightweight and promises better signal reception. In addition to the IMU built into the flight controller, a tactical-grade GNSS-aided Inertial Navigation System (INS) will be tested, which is able to receive GNSS Signals from L1 and L5 Bands. The latter are less susceptible to multipath effects and signal attenuation and therefore better suited in urban environments (u-blox AG, 2023). Furthermore, the INS offers high-quality roll, pitch and heading.

The mounting of the sensors and other relevant components such as the flight controller, companion computer, and the battery on the UAV will be realized as shown in Figure 1.

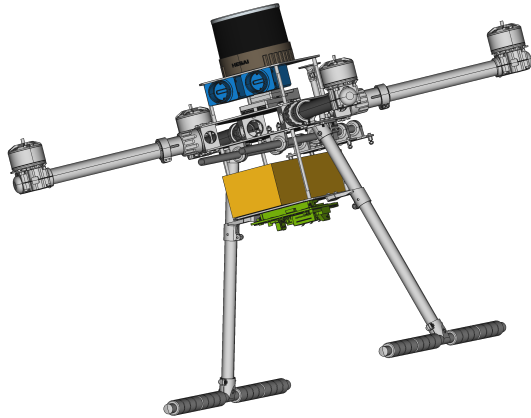


Figure 1. CAD Design of the multi-modal sensor setup on the UAV (blue - cameras, yellow - battery, green - companion computer).

4. Integration

Our conceptual UAV setup, shown in Figure 2, is based on open source projects supported by the Dronecode Foundation⁵ and consists of four main blocks. The Multi-Modal Navigation system encompasses a frame-based RGB camera, a Event camera, a LiDAR, a GNSS receiver, a IMU (embedded in the flight controller), and optionally a GNSS-aided INS. The Pixhawk flight controller⁶ of our platform runs the PX4 Autopilot (Meier et al., 2015) and together with the companion computer offers sufficient computational resources for onboard processing. The communication system handles the information exchange between the UAV and the Ground Control Station (GCS), while the latter serves as the central interface for the human operator for mission planning, monitoring, and control.

Multi-modal navigation requires a companion computer with enough computing power to run deep learning models in real-time, while simultaneously managing other essential processes such as sensor fusion, map optimization, and communication. We selected the NVIDIA Jetson Orin family for sufficient AI inference performance. The Jetson's Linux operating system enables the execution of ROS2, which is used to interact and manage the perceptual sensors that have ROS2 driver support. In addition, ROS2 enables a deep integration with PX4 via a bridge realized by the uXRCE-DDS middleware⁷, facilitating efficient information exchange between the companion computer and the flight controller.

4.1 Temporal Synchronization

Precise temporal synchronization is critical for real-time, tightly-coupled systems. Time synchronization between our system components is realized using the Precision Time Protocol (PTP). The companion computer acts as the grandmaster clock, distributing time via Ethernet to the sensors, which all support the PTP protocol. The PX4-ROS 2 bridge (uXRCE-DDS) can be used directly for convenient time synchronization between the companion computer and the Pixhawk flight controller. Furthermore, this bridge is also used to read the data of Pixhawk's internal IMU. Next to time synchronization with PTP, the use of ethernet

allows fast data transfer and bypasses possible GNSS interference problems that can occur with USB 3.0 (Intel Corporation, 2012).

4.2 Communication Systems: Multi-link, RTK, and GCS

Reliable wireless communication is essential for safe and efficient data exchange among components in UAS operating within AAM environments (Tomaszewski and Kołakowski, 2023). To ensure continuous connectivity, a heterogeneous multi-link approach is recommended, leveraging Wi-Fi for local, low-latency connections and 5G/LTE for high-bandwidth, real-time mobile operations. Furthermore, we consider Low Earth Orbit (LEO) satellite systems, which provide reliable global communication in areas lacking terrestrial infrastructure (Becker and Schalk, 2024, Stouffer et al., 2020). The fundamental Internet Protocol (IP) underpins all three technologies: Wi-Fi, 5G/LTE, and Satcom, enabling seamless interoperability. Implementing relevant computed redundancy mechanism across these communication links is both practical and ideal for effective bandwidth utilization. Several Ethernet- and IP-based network protocols fundamentally support the exchange of control signals, data, and feedback between UAV onboard systems and the GCS. Therefore, a multi-link approach aids the navigation system in our setup.

RTK-based localization requires real-time correction data, which can be determined on the basis of CORS networks. Operators of these networks provide correction data via the Internet using NTRIP (Networked Transport of RTCM via Internet Protocol) caster, which then can be received via NTRIP clients. In our experimental setup, the correction data can be received by the GCS, equipped with a communication modem and can then be further transmitted to the flight controller using MAVLink⁸. Another option is to use the 5G/LTE modem on board the UAV, which allows the companion computer to retrieve correction data directly and bypass the GCS. Instead of using an NTRIP caster, correction data can also be provided by a local base station. Thus, communication plays a vital role in using RTK-GNSS as ground truth for the evaluation of our navigation experiments.

The GCS serves as a central node in the multi-modal navigation ecosystem, seamlessly integrating telemetry, control, and data management functions for UAV operations. Our GCS setup encompasses three main elements: communication system, GCS application platform and ROS2 map visualization. Modern GCS application platforms such as QGroundControl (QGC)⁹, Mission Planner (MP)¹⁰. ROS-based clients provide intuitive interfaces for monitoring and managing UAV missions. Equipped with a unified map representation, it supports advanced multi-modal navigation by aggregating and visualizing data from diverse sources. The GCS also functions as a robust communication hub, managing bidirectional data exchange between the UAV and ground infrastructure. It continuously receives comprehensive telemetry streams consisting of position, altitude, velocity, battery health, and low-bandwidth sensor outputs. Thereby enabling real-time situational awareness and informed decision-making for operators.

4.3 Calibration

For real-time sensor fusion applications, spatiotemporal calibration plays a crucial role. Inaccurate sensor calibration in-

⁵ Dronecode Foundation : <https://dronecode.org/>

⁶ Pixhawk: <https://pixhawk.org/>

⁷ uXRCE-DDS: https://docs.px4.io/main/en/middleware/uxrce_dds.html

⁸ PX4-RTK: https://docs.px4.io/main/en/advanced/rtk_gps.html

⁹ QGC: <https://docs.qgroundcontrol.com/master/en/>

¹⁰ MP: <https://ardupilot.org/planner/>

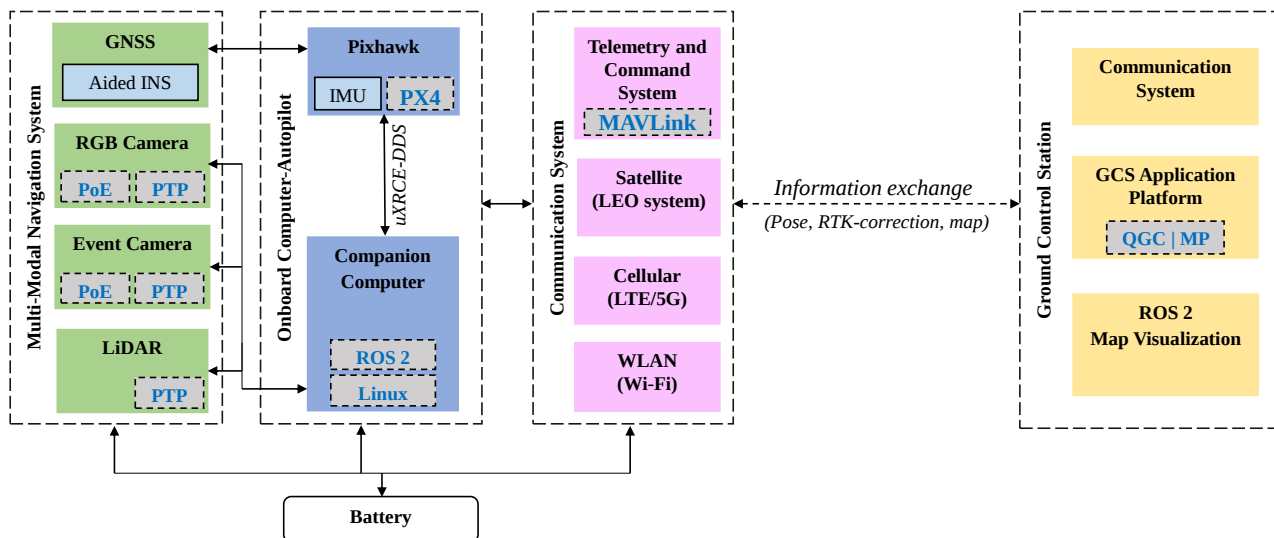


Figure 2. System design representation of the conceptual setup.

duces a large error due to sample mismatch (temporal) or misalignment errors (spatial) and results in a large absolute trajectory error (ATE) in sensor fusion-based navigation approaches. E2CALIB (Muglikar et al., 2021) and the DSEC dataset (Gehrig et al., 2021, Synchronization and Calibration) provide an intrinsic and extrinsic calibration approach for the frame-based and event camera. They propose using E2VID (Rebecq et al., 2021) for event-to-frame reconstruction to get frames from the event camera data. This allows the use of the Kalibr toolbox for multi-camera calibration (Furgale et al., 2013). With the cameras and IMU calibrated with Kalibr, LiDAR-camera calibration can be performed with FAST-Calib¹¹, as recommended by FAST-LIVO2 (Zheng et al., 2025). To enable onboard operations with the UAV testbed, temporal calibration is taken care of by PTP and the uXRCE-DDS middleware.

5. Conclusion and Outlook

This paper highlights the unique navigation challenges faced by UAS in AAM. Based on existing research, we proposed a navigation pipeline that includes map-based navigation and semantic SLAM on the basis of various sensors to address these challenges. While datasets for ground vehicles that combine multiple sensors already exist, there is a need for an aerial dataset to develop and evaluate the proposed approaches. To fill this gap, we presented a conceptual setup focused on both, individual sensor performance and the comprehensive integration of system components, laying the groundwork for advancing UAS navigation in AAM contexts.

Future work will address key open challenges, including holistic system calibration and a unified map representation capable of seamlessly integrating data from diverse sensors. The creation of such a common map framework is crucial for enabling reliable, interoperable navigation for UAS Traffic Management (UTM) surveillance. In the near future, we plan to perform flight tests and to publish a comprehensive dataset to support open science initiatives. We plan to continue refining our proposed methods to advance the state of multi-sensor integration and mapping for reliable navigation in AAM.

¹¹ FAST-Calib: <https://github.com/hku-mars/FAST-Calib>

Acknowledgements

This research was funded by the German Research Foundation (DFG) as part of RTG-AirMetro, grant number GRK 2947/1, 508591287 of Technische Universität Dresden.

We would like to thank Holybro, Connect Tech, LUCID Vision Labs Inc., and Hesai for kindly providing free access to 3D CAD models of their products, which facilitated the accurate design and integration of our system components.

References

- Altun, A. T., Hasanzade, M., Saldiran, E., Guner, G., Uzun, M., Fremond, R., Tang, Y., Bhundoo, P., Su, Y., Xu, Y., Inalhan, G., Hardt, M. W., Fransoy, A., Modha, A., Tena, J. A., Nieto, C., Vilaplana, M., Tojal, M., Gordo, V., Menendez, P., Gonzalez, A., 2023. The Development of an Advanced Air Mobility Flight Testing and Simulation Infrastructure. *Aerospace*, 10(8). <https://doi.org/10.3390/aerospace10080712>.
- Becker, D., Schalk, L. M., 2024. Toward robust and efficient communications for urban air mobility. *CEAS Aeronautical Journal*, 1–28. <https://doi.org/10.1007/s13272-024-00738-6>.
- Burri, M., Nikolic, J., Gohl, P., Schneider, T., Rehder, J., Omari, S., Achtelik, M. W., Siegwart, R., 2016. The EuRoC micro aerial vehicle datasets. *The International Journal of Robotics Research*. <https://doi.org/10.1177/0278364915620033>.
- Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., Leonard, J. J., 2016. Past, Present, and Future of Simultaneous Localization and Mapping: Toward the Robust-Perception Age. 32(6), 1309–1332. <https://doi.org/10.1109/TRO.2016.2624754>.
- Caesar, H., Bankiti, V., Lang, A. H., Vora, S., Liong, V. E., Xu, Q., Krishnan, A., Pan, Y., Baldan, G., Beijbom, O., 2020. Nuscen: A multimodal dataset for autonomous driving. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 11621–11631.

- Campos, C., Elvira, R., Rodríguez, J. J. G., Montiel, J. M., Tardós, J. D., 2021. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial, and Multimap SLAM. *IEEE Transactions on Robotics*, 37(6), 1874–1890. <https://doi.org/10.1109/TRO.2021.3075644>.
- Chen, X., Milioto, A., Palazzolo, E., Giguère, P., Behley, J., Stachniss, C., 2019. SuMa++: Efficient LiDAR-based Semantic SLAM. *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 4530–4537.
- Cheng, J., Cai, Z., Zhang, Z., Yin, W., Muller, M., Paulitsch, M., Yang, X., 2025. Romeo: Robust metric visual odometry. <https://arxiv.org/abs/2412.11530>.
- Cordts, M., Omran, M., Ramos, S., Rehfeld, T., Enzweiler, M., Benenson, R., Franke, U., Roth, S., Schiele, B., 2016. The Cityscapes Dataset for Semantic Urban Scene Understanding. <https://doi.org/10.48550/arXiv.1604.01685>.
- Delmerico, J., Cieslewski, T., Rebecq, H., Faessler, M., Scaramuzza, D., 2019. Are we ready for autonomous drone racing? the UZH-FPV drone racing dataset. *IEEE Int. Conf. Robot. Autom. (ICRA)*.
- EASA, 2021a. AMC and GM to Regulation (EU) 2019/947 on the rules and procedures for the operation of unmanned aircraft. Technical report, EASA. Accessed: 2025-02-20.
- EASA, 2021b. EASA SIB 2012-09R1 : Effects of Space Weather on Aviation. <https://ad.easa.europa.eu/ad/2012-09R1>.
- EASA, 2024a. EASA SIB 2022-02R3 : Global Navigation Satellite System Outage and Alterations Leading to Communication / Navigation / Surveillance Degradation. <https://ad.easa.europa.eu/ad/2022-02R3>.
- EASA, 2024b. Easy Access Rules for small category VCA. Technical report, EASA. Accessed: 2025-02-20.
- EUROCONTROL, 2021. European PBN Implementation and Transition Planning Handbook (Ed.2). Technical report, 2021. Accessed: 2025-02-20.
- Fong, W. K., Mohan, R., Hurtado, J. V., Zhou, L., Caesar, H., Beijbom, O., Valada, A., 2022. Panoptic Nuscenes: A Large-Scale Benchmark for LiDAR Panoptic Segmentation and Tracking. *IEEE Robotics and Automation Letters*, 7(2), 3795–3802. <https://doi.org/10.1109/LRA.2022.3148457>.
- Furgale, P., Rehder, J., Siegwart, R., 2013. Unified temporal and spatial calibration for multi-sensor systems. *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1280–1286.
- Gao, L., Liang, Y., Yang, J., Wu, S., Wang, C., Chen, J., Kneip, L., 2022. VECtor: A Versatile Event-Centric Benchmark for Multi-Sensor SLAM. *IEEE Robotics and Automation Letters*, 7(3), 8217–8224. <https://doi.org/10.1109/LRA.2022.3186770>.
- Gehrig, D., Scaramuzza, D., 2024. Low Latency Automotive Vision with Event Cameras. *Nature*, 629(8014), 1034–1040. <https://doi.org/10.1038/s41586-024-07409-w>.
- Gehrig, M., Aarents, W., Gehrig, D., Scaramuzza, D., 2021. DSEC: A Stereo Event Camera Dataset for Driving Scenarios. *IEEE Robotics and Automation Letters*, 6(3), 4947–4954. <https://doi.org/10.1109/LRA.2021.3068942>.
- Geiger, A., Lenz, P., Stiller, C., Urtasun, R., 2013. Vision meets Robotics: The KITTI Dataset. *International Journal of Robotics Research (IJRR)*.
- Goyal, R., Reiche, C., Fernando, C., Cohen, A., 2021. Advanced Air Mobility: Demand Analysis and Market Potential of the Airport Shuttle and Air Taxi Markets. *Sustainability*, 13(13). <https://doi.org/10.3390/su13137421>.
- Guadagnino, T., Mersch, B., Gupta, S., Vizzo, I., Grisetti, G., Stachniss, C., 2025. Kiss-slam: A simple, robust, and accurate 3d lidar slam system with enhanced generalization capabilities. <https://arxiv.org/abs/2503.12660>.
- Hinniger, C., Rüter, J., 2023. Synthetic Training Data for Semantic Segmentation of the Environment from UAV Perspective. *Aerospace*, 10(7). <https://doi.org/10.3390/aerospace10070604>.
- ICAO, 2012. Performance-based Navigation (PBN) Manual (Advanced fourth edition). Technical report, ICAO. Accessed: 2025-02-20.
- Intel Corporation, 2012. USB 3.0 Radio Frequency Interference Impact on 2.4 GHz Wireless Devices White Paper. <https://www.intel.com/content/www/us/en/content-details/841692/usb-3-0-radio-frequency-interference-impact-on-2-4-ghz-wireless-devices-white-paper.html>. Accessed: 2025-07-06.
- Kinnari, J., Renzulli, R., Verdoja, F., Kyrki, V., 2023. LSVL: Large-scale Season-Invariant Visual Localization for UAVs. *Robotics and Autonomous Systems*, 168, 104497. <https://doi.org/10.1016/j.robot.2023.104497>.
- Konert, A., Kasprzyk, P., 2024. Very Low Level Flight Rules for Manned and Unmanned Aircraft Operations. *Journal of Intelligent & Robotic Systems*, 110.
- Lyu, Y., Vosselman, G., Xia, G.-S., Yilmaz, A., Yang, M. Y., 2020. UAVid: A semantic segmentation dataset for UAV imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 165, 108 - 119. <https://doi.org/10.1016/j.isprsjprs.2020.05.009>.
- Macario Barros, A., Michel, M., Moline, Y., Corre, G., Carrel, F., 2022. A Comprehensive Survey of Visual SLAM Algorithms. 11(1), 24. <https://doi.org/10.3390/robotics11010024>.
- Marcu, A., Costea, D., Licaret, V., Leordeanu, M., 2019. Towards Automatic Annotation for Semantic Segmentation in Drone Videos. *arXiv preprint arXiv:1910.10026*. <https://doi.org/10.48550/arXiv.1910.10026>.
- Meier, L., Honegger, D., Pollefeys, M., 2015. PX4: A node-based multithreaded open source robotics framework for deeply embedded platforms. *2015 IEEE International Conference on Robotics and Automation (ICRA)*. <https://doi.org/10.1109/ICRA.2015.7140074>.
- Merat, R., Cioffi, G., Bauersfeld, L., Scaramuzza, D., 2024. Drift-Free Visual SLAM Using Digital Twins. *IEEE Robotics and Automation Letters*. <https://doi.org/10.1109/LRA.2024.3518840>.
- Muglikar, M., Gehrig, M., Gehrig, D., Scaramuzza, D., 2021. How to calibrate your event camera. *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 1403–1409.

- Nex, F., Armenakis, C., Cramer, M., Cucci, D. A., Gerke, M., Honkavaara, E., Kukko, A., Persello, C., Skaloud, J., 2022. UAV in the Advent of the Twenties: Where We Stand and What Is Next. 184, 215–242. <https://doi.org/10.1016/j.isprsjprs.2021.12.006>.
- Pan, Y., Zhong, X., Wiesmann, L., Posewsky, T., Behley, J., Stachniss, C., 2024. PIN-SLAM: LiDAR SLAM Using a Point-Based Implicit Neural Representation for Achieving Global Map Consistency. *IEEE Transactions on Robotics*.
- Pongsakornsathien, N., Safwat, N. E.-D., Xie, Y., Gardi, A., Sabatini, R., 2025. Advances in low-altitude airspace management for uncrewed aircraft and advanced air mobility. *Progress in Aerospace Sciences*, 154, 101085. <https://doi.org/10.1016/j.paerosci.2025.101085>.
- Qin, T., Li, P., Shen, S., 2018. VINS-mono: A Robust and Versatile Monocular Visual-Inertial State Estimator. *IEEE Transactions on Robotics*, 34(4), 1004–1020. <https://doi.org/10.1109/TRO.2018.2853729>.
- Rebecq, H., Ranftl, R., Koltun, V., Scaramuzza, D., 2021. High Speed and High Dynamic Range Video with an Event Camera. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(6), 1964–1980. <https://doi.org/10.1109/TPAMI.2019.2963386>.
- Sarlin, P.-E., DeTone, D., Yang, T.-Y., Avetisyan, A., 2023a. Cut-paste: Self-supervised learning for anomaly detection with imperfect data. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 21495–21504.
- Sarlin, P.-E., Trulls, E., Pollefeys, M., Hosang, J., Lynen, S., 2023b. SNAP: Self-Supervised Neural Maps for Visual Positioning and Semantic Understanding. <http://arxiv.org/abs/2306.05407>.
- Shan, T., Englot, B., Ratti, C., Rus, D., 2021. Lvi-sam: Tightly-coupled lidar-visual-inertial odometry via smoothing and mapping. *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 5692–5698.
- Siciliano, B., Khatib, O., 2016. Robotics and the handbook. *Springer Handbook of Robotics*, Springer.
- Stouffer, V. L., Cotton, W. B., DeAngelis, R. A., Devasirvatham, D. M., Irvine, T. B., Jennings, R. E., Lehmer, R. D., Nguyen, T. C., Shaver, M., 2020. Reliable, secure, and scalable communications, navigation, and surveillance (cns) options for urban air mobility (uam). Technical report, NASA Glenn Research Center.
- Tomaszewski, L., Kołakowski, R., 2023. Advanced Air Mobility and Evolution of Mobile Networks. *Drones*, 7(9). <https://doi.org/10.3390/drones7090556>.
- u-blox AG, 2023. Modern gnss/gps signals: Moving from single-band to dual-band. <https://www.u-blox.com/en/publication/white-paper/gps-signals-evolution>. Accessed: 2025-07-03.
- Vidal, A. R., Rebecq, H., Horstschaefer, T., Scaramuzza, D., 2018. Ultimate SLAM? Combining events, images, and IMU for robust visual SLAM in HDR and high-speed scenarios. *IEEE Robotics and Automation Letters*, 3(2), 994–1001. <https://doi.org/10.1109/LRA.2018.2793357>.
- Vultaggio, F., Fanta-Jende, P., Schörghuber, M., Kern, A., Gerke, M., 2024. Investigating Visual Localization Using Geospatial Meshes. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLVIII-2-W8-2024, 447–454. <https://doi.org/10.5194/isprs-archives-XLVIII-2-W8-2024-447-2024>.
- Wan, X., Shao, Y., Zhang, S., Li, S., 2022. Terrain Aided Planetary UAV Localization Based on Geo-referencing. *IEEE Transactions on Geoscience and Remote Sensing*, 60, 1–18. <https://doi.org/10.1109/TGRS.2022.3198745>.
- Xiang, S., Xie, A., Ye, M., Yan, X., Han, X., Niu, H., Li, Q., Huang, H., 2024. Autonomous eVTOL: A summary of researches and challenges. *Green Energy and Intelligent Transportation*, 3(1), 100140. <https://doi.org/10.1016/j.geits.2023.100140>.
- Yin, J., Li, A., Li, T., Yu, W., Zou, D., 2021. M2dgr: A multi-sensor and multi-scenario slam dataset for ground robots. *IEEE Robotics and Automation Letters*, 7(2), 2266–2273. <https://doi.org/10.1109/LRA.2021.3138527>.
- Zhang, L., Tao, Y., Lin, J., Zhang, F., Fallon, M., 2024. Visual Localization in 3D Maps: Comparing Point Cloud, Mesh, and NeRF Representations. *arXiv preprint arXiv:2408.11966*. <https://doi.org/10.48550/arXiv.2408.11966>.
- Zheng, C., Xu, W., Zou, Z., Hua, T., Yuan, C., He, D., Zhou, B., Liu, Z., Lin, J., Zhu, F., Ren, Y., Wang, R., Meng, F., Zhang, F., 2025. FAST-LIVO2: Fast, Direct LiDAR–Inertial–Visual Odometry. *IEEE Transactions on Robotics*, 41, 326–346. <https://doi.org/10.1109/TRO.2024.3502198>.
- Zidan, J., Adegoke, E. I., Kampert, E., Birrell, S. A., Ford, C. R., Higgins, M. D., 2020. GNSS Vulnerabilities and Existing Solutions: A Review of the Literature. *IEEE Access*, 9, 153960–153976. <https://doi.org/10.1109/ACCESS.2020.2973759>.