

Abstract

Alessandro Minervini

July 2026

Autonomous navigation in GNSS-denied environments remains a fundamental challenge in robotics. Heterogeneous multi-robot formations composed of UAVs and UGVs have gained considerable interest within the scientific community, as they are particularly advantageous for accomplishing missions in real-world scenarios. One of the most critical aspects to enable multi-robotic cooperation in GNSS-denied environments is robot localization. The presence of both ground and aerial platforms can improve the design of the platforms in the formation by equipping each robot with a sensor setup tailored to its physical and operational characteristics. Additionally, shared information among the robots can introduce additional constraints to improve localization accuracy.

A typical configuration in service robotics consists of UGVs cooperating with rotary-wing UAVs. These aerial platforms are significantly constrained by strict limitations in weight, size, and power consumption; thus, some sensor setups are inconvenient in applications requiring small dimensions and high flight endurance. For instance, 3D LiDAR-based localization algorithms can provide high accuracy and robustness, representing some of the most comprehensive solutions for navigation purposes. However, their applicability to UAVs is limited due to their weight, size, and power consumption.

On the other hand, vision-based approaches such as Visual-Inertial Odometry (VIO) offer optimal solutions for their installation simplicity, low weight, and power consumption. Additionally, in many scenarios, image data represent the final output of the mission, while LiDAR sensors are often limited to support navigation and obstacle avoidance. Therefore, exploiting cameras for both perception and navigation purposes reduces sensor redundancy and enables a more compact, lightweight, and energy-efficient sensor setup. However, vision-based localization algorithms suffer from drift accumulation and poor robustness under challenging conditions, such as low-texture environments, high-speed motion, and brightness changes. These limitations effectively reduce their applicability in real-world scenarios, with substantial consequences on system design and overall efficiency.

This dissertation focuses on improving the localization performance of UAVs relying on VIO algorithms. The limitations of visual-based navigation are identified, and different methodologies are introduced to address the typical failures experienced by these algorithms. To improve robustness in challenging environments, state-of-the-art pipelines are analyzed, and the solutions are implemented

within popular existing frameworks.

First, monocular setups are analyzed, and methodologies to mitigate common issues related to these setups are introduced. The inherent limitations of these configurations are addressed through a geometric initialization strategy, and adjustments to the optimization process are introduced to mitigate solution divergence during pure rotational motions.

Then, the focus is shifted to the problem of insufficient feature tracking, which substantially reduces the reliability of visual-based systems. To address this issue, an IMU-based prediction of the feature locations across consecutive frames is included within an open-source VIO pipeline to improve performance during high-speed motions. Subsequently, a generalized multi-camera setup is implemented within the same framework to enhance localization in challenging environments.

The drift accumulated by VIO algorithms is addressed with two sensor-fusion techniques. To avoid significant increases in the payload, lightweight sensors are utilized for this scope. In particular, VIO estimates are combined with the measurements from a downward-facing range finder to regularly correct the altitude drift accumulated during the navigation. In the context of multi-robotic cooperation, a sensor fusion method leveraging the UGV sensor setup is introduced. The approach effectively reduces the UAV drift by combining the UAV detections obtained by the UGV cameras with the 3D LiDAR point cloud acquired by the ground platform.

Experimental results demonstrate that the proposed methodologies significantly improve localization accuracy, robustness, and global consistency of VIO algorithms across a range of challenging scenarios. Moreover, the results confirm that these improvements can be achieved while preserving the compactness and limited computational requirements of small aerial platforms.