

DOI 10.24412/1829-0450-fm-2026-1-28-35  
УДК 519.688

Поступила: 09.03.2026г.  
Сдана на рецензию: 10.03.2026г.  
Подписана к печати: 19.03.2026г.

## VISUAL RELATIVE LOCALIZATION FOR GNSS-FREE CONTROL OF UAV SWARM

*V. Sahakyan*

*Russian-Armenian University  
vardan.sahakyan@student.rau.am*

### ABSTRACT

Operation of UAV formations in GNSS-denied environments requires alternative sources of position information that can be integrated directly into the flight-control loop. This paper presents a leader–follower approach in which GNSS is replaced for follower UAVs using visual relative localization computed onboard a single leader platform. The leader estimates the relative three-dimensional positions of follower UAVs using monocular vision and injects these estimates into the followers' PX4 flight controllers through the external-vision interface. The followers rely exclusively on the injected relative position measurements for state estimation and closed-loop position control, without using GNSS or performing onboard localization. The proposed system is fully onboard, requires no external infrastructure, and is implemented using an open-source ROS 2–PX4 software stack. The approach is evaluated in a multi-UAV software-in-the-loop simulation under GNSS-denied conditions. Two experiments are conducted: nominal formation flight to assess steady-state behavior, and a leader velocity sweep to analyze the effect of increased dynamic excitation on tracking accuracy. Results show stable follower position control using relative localization alone, with bounded error across the tested operating range, demonstrating the feasibility of visual relative localization as a practical GNSS replacement for leader–follower UAV formations.

**Keywords:** GNSS-denied navigation; UAV formations; leader–follower architecture; visual relative localization; PX4 autopilot

### Introduction

Most multi-UAV formation and swarm systems implicitly assume the availability of absolute position estimates on each vehicle, typically provided by GNSS. In practice, this assumption strongly constrains deployability: follower UAVs must carry full navigation sensors, run onboard localization pipelines, and maintain global consistency even when their role in the formation is purely reactive. This paper adopts the opposite perspective and asks a simpler question: can GNSS be removed entirely from follower UAVs while preserving closed-loop position control using a standard autopilot?

We address this question using a leader–follower architecture in which all sensing and relative localization are performed onboard a single leader UAV. The leader estimates the three-dimensional relative positions of follower vehicles using monocular vision and onboard state information, and streams these estimates directly into the followers’ flight controllers through the PX4 [1] external-vision interface. The followers do not perform localization, do not fuse GNSS, and rely exclusively on the injected relative position measurements for state estimation and control.

Unlike approaches that require custom controllers or SLAM-based backends [2],[3], the proposed method integrates relative localization directly into the standard PX4 estimation and position-control loop. This allows follower UAVs to operate using unmodified autopilot control laws while replacing GNSS with externally provided relative pose information. The system is evaluated in multi-UAV software-in-the-loop simulation under GNSS-denied conditions, using nominal formation flight and a leader velocity sweep to characterize stability and dynamic limits.

## **Related Work**

GNSS-denied multi-UAV operation has been addressed through visual-inertial SLAM, cooperative localization, and relative sensing–based formation control. Collaborative SLAM frameworks enable multiple UAVs to estimate globally consistent trajectories by sharing visual and inertial information, often using centralized or distributed backends [3],[4]. These systems typically require onboard perception and estimation on each vehicle and rely on persistent environmental features. Relative localization approaches avoid global mapping by estimating inter-UAV geometry directly. Vision-based flocking and formation-control systems use bearing-only or relative pose measurements to maintain coordinated motion under GNSS-denied conditions, either in decentralized configurations [5][6] or hierarchical leader–follower architectures [7]. Such methods reduce sensing requirements but commonly employ custom control laws or operate outside standard autopilot estimation pipelines.

Leader–follower visual collaboration has also been demonstrated in indoor and outdoor environments, where a sensing-capable leader provides relative position information to follower UAVs [8]. These systems illustrate the feasibility of relative localization–based coordination while differing in sensing assumptions, estimator design, and control integration.

## **Proposed Method**

### **A. System Overview**

The proposed method replaces GNSS-based positioning for follower UAVs using visual relative localization computed onboard a single leader UAV. The leader performs all sensing and estimation, while followers operate using standard PX4

estimation and position control driven exclusively by externally provided relative position measurements. No global localization, mapping, or GNSS fusion is performed on follower vehicles.

The system follows a centralized leader–follower structure in which relative pose estimates are generated on the leader and streamed to the followers at a fixed rate. These estimates are injected directly into the follower flight controllers via PX4’s external-vision interface and treated as the primary position source by the onboard estimator.

### **B. Visual Relative Localization**

The leader UAV is equipped with a downward-facing monocular camera and onboard state estimation. Follower UAVs are detected in the image plane using a lightweight YOLO-based object detector with an explicit UAV target class [9] and tracked over time using the Norfair multi-object tracking framework [10], producing pixel-space measurements for each visible follower. Each detection is back-projected into a three-dimensional direction vector using known camera intrinsics.

Depth is recovered using the vertical separation between the leader and follower, obtained from onboard altitude measurements. The resulting three-dimensional relative position is first expressed in the camera frame and then transformed into the leader body frame using calibrated camera-to-body extrinsics. Finally, the relative position is rotated into the local NED frame using the leader’s attitude estimate. The output of this process is a relative position vector expressed in the same coordinate frame used by the flight controller.

### **C. GNSS Replacement via External Vision Injection**

Relative position estimates computed on the leader are transmitted to each follower UAV over a wireless link and injected into the PX4 estimator using the `vehicle_visual_odometry[11]` interface. Each message contains the relative position expressed in the local NED frame, a synchronized timestamp, and appropriate validity flags.

On the follower UAVs, GNSS fusion is disabled and the PX4 estimator is configured to rely solely on externally provided position updates. The injected relative positions are treated as ground-truth measurements for state estimation, enabling closed-loop position control without GNSS or onboard localization. Followers use their standard PX4 position controller to track commanded setpoints based on the estimated state.

### **D. Control loop integration**

The proposed method integrates relative localization directly into the standard PX4 estimation–control loop. No modifications to the autopilot controller structure are required. As the leader moves, updated relative position estimates are continuously streamed to the followers, allowing them to maintain their assigned relative positions using unmodified PX4 position control laws. This tight coupling

between perception, estimation, and control enables stable closed-loop operation while replacing GNSS entirely for follower UAVs.

## Results

### A. Simulation environment

The proposed method was evaluated in a multi-UAV software-in-the-loop simulation using PX4 SITL[12], Gazebo[13], and ROS 2 [14]. Each UAV operated in an isolated namespace with an independent PX4 instance and uXRCE-DDS[15] session. One UAV acted as the leader and five as followers. GNSS fusion was disabled on all followers, which relied exclusively on externally injected relative position estimates for state estimation and control. Gazebo served as the global time reference to ensure synchronized operation across vehicles. All experiments were conducted under GNSS-denied conditions.

### B. Metrics

Follower performance was evaluated using two complementary metrics.

The primary metric is the root-mean-square (RMS) relative position error, defined as the Euclidean distance between the estimated relative position of each follower and its desired relative offset with respect to the leader, expressed in the leader-aligned NED frame. RMS error was computed over the full trajectory for each follower and averaged across all followers.

To characterize deformation of the formation under dynamic excitation, formation distortion (FD) was also evaluated. Formation distortion measures deviations in pairwise distances between follower UAVs relative to their initial spacing, averaged over all follower pairs and over time. This metric captures changes in formation shape that may not be reflected by leader–follower error alone, particularly under high-speed motion.

### C. Results Nominal operation

In the nominal experiment (Figure 1), the leader UAV executed two representative motion patterns at a constant speed of 3 m/s:

- (i) a straight-line trajectory of 50 m length, and
- (ii) a planar figure-eight trajectory composed of smooth circular arcs.

The follower UAVs maintained predefined circular, V-shaped, and line formation geometries, with fixed relative offsets with respect to the leader, using only externally provided localization. For both trajectories, the formation remained stable throughout the flight, with RMS relative position errors on the order of 1–2 m and low formation distortion. Error distributions remained bounded, and no estimator instability or oscillatory behavior was observed, indicating that the injected relative position estimates were sufficient to support sustained closed-loop position control without GNSS under nominal motion.

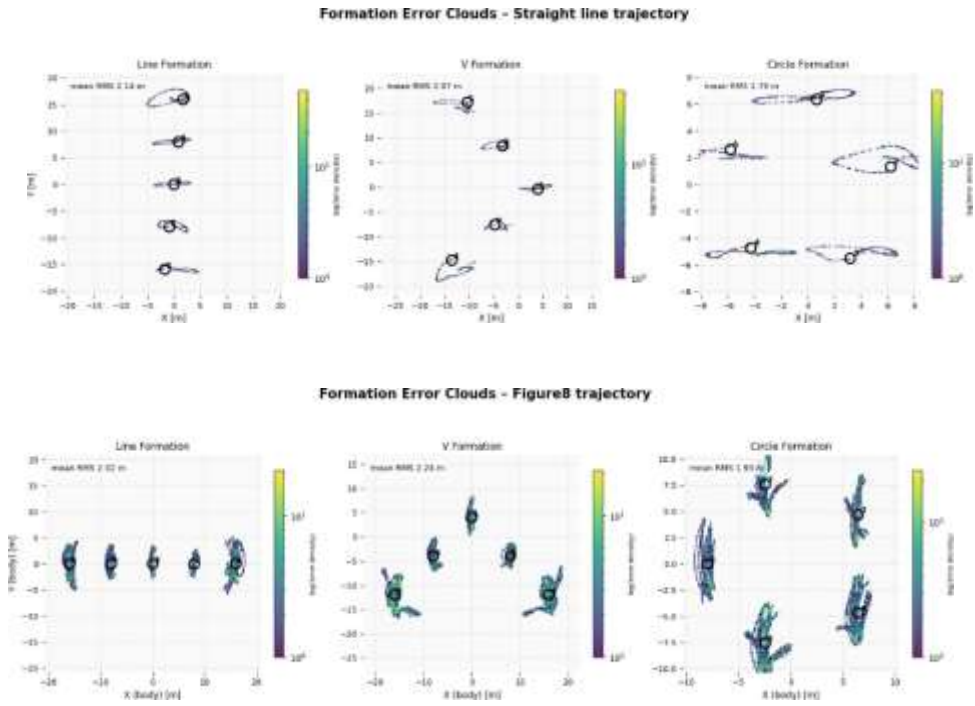


Figure 1. Error-cloud plots for the straight-line, figure-8, and circular missions, shown for line, V, and circle formations. Each subplot depicts the instantaneous follower relative-position error in the leader's body frame, sampled at 30 Hz. The horizontal and vertical axes correspond to the leader body-frame  $x$  and  $y$  directions.

## Velocity Sweep

To evaluate dynamic limits, the leader executed the straight-line trajectory while its speed was increased from 2 m/s to 7 m/s in 1 m/s increments, with the follower UAVs maintaining a circular formation. Followers maintained the same relative configuration throughout all runs (Figure 2). At low to moderate speeds (2–4 m/s), both RMS relative error and formation distortion remained close to nominal values, indicating near-rigid formation behavior. As the leader speed exceeded approximately 5 m/s, the RMS error increased gradually, and the formation distortion became more pronounced, elongating along the direction of motion. Despite this degradation, the system remained stable across the entire tested range, with no divergence or loss of control observed. The observed performance degradation reflects bandwidth and latency limitations of the perception–estimation–control pipeline rather than instability of the control loop.

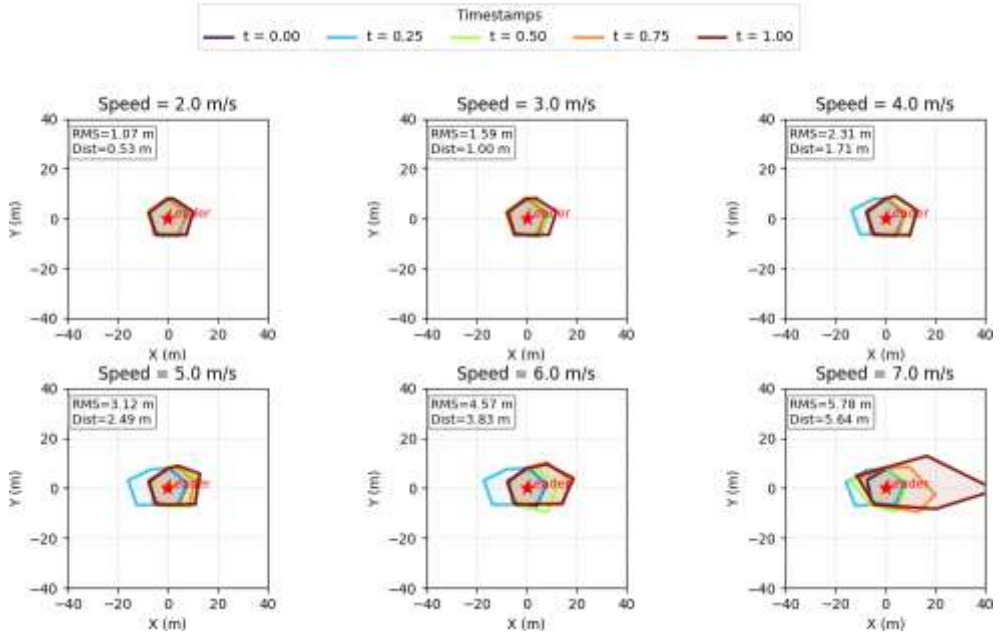


Figure 2. Leader and follower trajectories for straight-line missions at speeds of 2, 3, 4, 5, 6, and 7 m/s, shown in the leader's camera frame. Each subplot depicts follower positions at normalized timestamps  $t=\{0.00, 0.25, 0.50, 0.75, 1.00\}$  for the circular formation. The reported values in each subplot are the time-averaged RMS drift and formation-distortion metrics computed over the full trajectory.

## Conclusion

This paper presented a practical GNSS replacement strategy for follower UAVs in a leader–follower formation, based on visual relative localization and tight integration with the PX4 flight-control stack. Relative position estimates computed onboard a single leader UAV were injected directly into follower estimators via the PX4 external-vision interface, enabling closed-loop position control without GNSS or onboard localization on follower vehicles.

Simulation results under GNSS-denied conditions demonstrated stable follower control during nominal motion and predictable, gradual performance degradation as leader speed increased. The velocity-sweep experiments highlighted the dynamic limits of the perception–estimation–control pipeline while confirming the absence of instability or divergence across the tested operating range. Together, these results indicate that visual relative localization can serve as a viable positioning source for follower UAVs when integrated directly into standard autopilot estimation and control loops.

Future work will focus on experimental validation in real-world multi-UAV flights, improved robustness to visual occlusions, and extension toward more distributed or redundant architectures that relax the single-leader assumption.

**Acknowledgment:** This work was supported by the Science Committee of RA (Research project No 23AA-1B007).

## REFERENCES

1. PX4 Autopilot Development Team, “PX4 Flight Stack Architecture,” PX4 User Guide. [Online]. Available: <https://docs.px4.io/main/en/concept/architecture#flight-stack>
2. Xie J., He X., Mao J., Zhang L., & Hu X. (2022). C2VIR-SLAM: centralized collaborative visual-inertial-range simultaneous localization and mapping. *Drones*, 6 (11), 312.
3. Xu H., Liu P., Chen X., & Shen S. (2024).  $\$ D^{\wedge}\{2\} \$$  SLAM: Decentralized and Distributed Collaborative Visual-Inertial SLAM System for Aerial Swarm. *IEEE Transactions on Robotics*, 40, 3445–3464.
4. Patel M., Karrer M., Bänninger P. & Chli M. (2023). Covins-g: A generic back-end for collaborative visual-inertial slam. *arXiv preprint arXiv:2301.07147*.
5. Schilling F., Schiano F., Floreano D. Vision-Based Drone Flocking in Outdoor Environments,” *IEEE Robotics and Automation Letters*, vol. 6, no. 2. PP. 2953–2960, 2021.
6. Horyna J., Krátký V., Pritzl V., Băča T., Ferrante E. & Saska M. (2024). Fast swarming of uavs in gnss-denied feature-poor environments without explicit communication. *IEEE Robotics and Automation Letters*, 9 (6), 5284–5291.
7. Ma L., Meng D., Huang X., & Zhao S. (2023). Vision-based formation control for an outdoor UAV swarm with hierarchical architecture. *IEEE Access*, 11, 75134–75151.
8. Evangeliou N. et al., Visual Collaboration Leader-Follower UAV-Formation for Indoor Exploration, *Frontiers in Robotics and AI*, 2022, DOI: 10.3389/frobt.2021.777535.
9. Kjelberg M. Drone Detector, Hugging Face, 2023. [Online]. Available: <https://huggingface.co/marie-kjelberg/drone-detector>. Accessed: Jan. 23, 2026.
10. Tryolabs, “Norfair: Lightweight Python library for real-time multi-object tracking,” GitHub repository, 2023. [Online]. Available: <https://github.com/tryolabs/norfair>. Accessed: Jan. 23, 2026.
11. PX4 Development Team. VehicleOdometry (uORB message). [https://docs.px4.io/main/en/msg\\_docs/VehicleOdometry](https://docs.px4.io/main/en/msg_docs/VehicleOdometry)
12. PX4 Development Team. PX4 Autopilot User Guide – Simulation (SITL). <https://docs.px4.io/main/en/simulation/>
13. Open Source Robotics Foundation (OSRF). . <https://gazebosim.org/>
14. ROS 2: <https://docs.ros.org/en/rolling/index.html>
15. PX4 uXRCE-DDS Middleware: <https://docs.px4.io/main/en/middleware/micrortps.html>

## ВИЗУАЛЬНАЯ ОТНОСИТЕЛЬНАЯ ЛОКАЛИЗАЦИЯ ДЛЯ GNSS-НЕЗАВИСИМОГО УПРАВЛЕНИЯ РОЕМ БПЛА

*В. Саакян*

*Российско-Армянский университет*

## АННОТАЦИЯ

Эксплуатация формаций беспилотных летательных аппаратов в условиях отсутствия GNSS требует альтернативных источников по-

зиционной информации, которые могут быть напрямую интегрированы в контур управления полетом. В данной работе представлен подход типа «лидер-ведомый», в котором GNSS для ведомых БПЛА заменяется визуальной относительной локализацией, вычисляемой на борту одного ведущего аппарата. Ведущий БПЛА оценивает относительные трехмерные положения ведомых аппаратов с использованием монокулярного зрения и передает эти оценки в полетные контроллеры PX4 ведомых через интерфейс внешнего зрения. Ведомые БПЛА используют исключительно внедряемые относительные измерения положения для оценки состояния и замкнутого управления положением, без применения GNSS и без выполнения бортовой локализации.

Предложенная система полностью реализована на борту летательных аппаратов, не требует внешней инфраструктуры и основана на открытом программном стеке ROS 2-PX4. Подход был оценен в симуляторе с несколькими БПЛА (software-in-the-loop) в условиях отсутствия GNSS. Проведены два эксперимента: номинальный полет в формации для оценки установившегося режима работы и варьирование скорости ведущего БПЛА для анализа влияния возросшей динамической нагрузки на точность слежения. Полученные результаты демонстрируют устойчивое управление положением ведомых БПЛА на основе одной лишь относительной локализации, с ограниченной ошибкой во всем исследованном диапазоне режимов, что подтверждает практическую применимость визуальной относительной локализации в качестве замены GNSS для формаций БПЛА с архитектурой «лидер-ведомый».

**Ключевые слова:** навигация без GNSS; формации БПЛА; архитектура «лидер-ведомый»; визуальная относительная локализация; автопилот PX4.