TECHNICAL UNIVERSITY OF CLUJ-NAPOCA

ACTA TECHNICA NAPOCENSIS

Series: Applied Mathematics, Mechanics, and Engineering
Vol. 69, Issue Special III, July, 2026

HYBRID UAV-UGV COORDINATION FOR POINT-OF-INTEREST LOCALIZATION IN GPS-DENIED ENVIRONMENTS USING STEREO VISION, LIDAR AND GESTURE-BASED SIGNALING

Patricia PREDA, Lucian GRIGORE, Eliza MARIN

Abstract: *This paper presents a hybrid sensing and coordination framework enabling UAV-UGV teams to identify and localize points of interest in GPS-denied or communication-degraded environments. The approach is designed strictly for defensive and protective applications, where improving situational awareness and reducing human exposure to risk are primary objectives. The system integrates stereo vision, LiDAR-based clustering, and biologically inspired gesture signaling as lightweight, non-intrusive communication layer between aerial and ground platforms. ROS is used only for sensor synchronization and spatial alignment. Mechanical considerations – such as sensor mounting, vibration resilience, and field-of-view geometry – are included to ensure safe operation on uneven terrain. Experimental observations highlight the complementary strengths of stereo and LiDAR perception in supporting responsible, human-centered multi-robot collaboration.*

Keywords: *UAV-UGV collaboration, LiDAR, gesture-based signaling, GPS-denied environments, ROS, hybrid perception, point-of-interest localization.*

1. INTRODUCTION

Cooperative UAV-UGV systems are increasingly required in GPS-denied environments, where neither platform can rely exclusively on satellite signals or persistent communication channels. Conventional localization methods and RF-based control mechanisms degrade significantly in enclosed spaces, underground environments, or in operations affected by infrastructure failure, making autonomous perception a primary requirement for safe deployment.

The core problem addressed in this work is achieving spatial consistency and point-of-interest localization using onboard sensing rather than external positioning or continuous teleoperation.

Existing approaches rely predominantly on either stereo-vision-based depth estimation or LiDAR-based geometric reconstruction. Stereo methods provide dense depth maps but are sensitive to illumination and vibration; LiDAR yields stable geometric information but lacks semantic context and imposes higher computational load. Hybrid sensing strategies have been investigated in aerial robotics and mobile platforms separately, however cooperative UAV-UGV fusion, particularly under communication degraded conditions, remains insufficiently addressed [1] in the literature. Furthermore, most existing studies assume reliable communication between platforms, with coordination implemented through explicit message passing.

The present work integrated stereo vision and LiDAR clustering within a unified perception framework evaluated on real experimental data. A dataset comprising 100 LiDAR frames acquired during UAV motion was processed using geometric filtering and DBSCAN-based clustering, enabling quantitative assessment of detection stability, positional jitter, trajectory reconstruction, and dynamic behavior. The contribution of this paper is fourfold: (1) a hybrid stereo-LiDAR perception architecture tailored for UAV-UGV cooperation in GPS-denied environments, (2) a

biologically inspired gesture-based coordination concept, (3) an analysis of mechanical integration affecting sensing reliability, and (4) a quantitative experimental validation of LiDAR-based UAV tracking performance.

Unlike purely theoretical models, this work presents real-world experimental results demonstrating an 86% detection rate and a positional jitter of only 3.46 cm under dynamic conditions. These metrics support the feasibility of the proposed hybrid fusion approach and illustrate measurable performance improvements compared to the known limitations of single-sensor configurations.

2. RELATED WORK

UAVs are commonly employed for wide-area inspection and hazard assessment, while UGVs are used for ground-level verification and manipulation. Cooperative heterogeneous multi-robot systems, including combinations of aerial and ground agents, have been extensively reviewed, highlighting key challenges in perception, coordination, control, and communication for robust joint autonomy [2]. Comprehensive surveys of UAV-UGV collaborative systems have classified these architectures based on their level of autonomy and communication coupling [3]. Recent studies have further emphasized the shift toward resilient multi-modal sensor fusion in heterogeneous teams [4]. However, most approaches treat the two platforms as independent systems, relying either on aerial perception or on ground-based mapping, without tightly integrating heterogeneous sensing [1]. Recent frameworks have addressed cooperative path planning to optimize search-and-rescue missions [5], yet tightly integrating heterogeneous sensing in GPS-denied environments remain challenging, while [2] notes that specific instantiations that unify advanced perception modalities with implicit gesture signaling in GPS-denied environments remain largely unexplored. Stereo-based depth estimation has been investigated extensively in UAV inspections tasks, due to the availability of dense depth maps and low payload requirements. Nevertheless, stereo depth remains highly sensitive to illumination conditions, occlusions, and vibration, requiring explicit filtering and calibration procedures. Multi-sensor fusion methods have been widely studied in cooperative robotics, demonstrating that integrating complementary sensors can improve robustness and metric accuracy over single modality perception [6]. Recent surveys on autonomous UAV perception emphasize that robust localization and 3D understanding often require the integration of multiple sensing modalities, particularly under challenging conditions such as illumination variation or absence of GPS [7].

LiDAR-based point-cloud processing has been widely adopted in UGV navigation and environmental mapping. Such methods provide reliable metric information independently of lighting but incur higher computational load and do not intrinsically provide semantic context. Hybrid perception solutions have been proposed in mobile robotics, typically combining vision and LiDAR for improved robustness; however cooperative UAV-UGV hybrid sensing under GPS-denied and communication-degraded conditions remains comparatively under-explored [1].

Prior studies predominantly assume stable communication links and centralized coordination, which hinders applicability in GPS-denied or communication-constrained environments. Recent reviews emphasize that decentralized and sensor-based cooperation strategies remain under-explored in practical UAV-UGV frameworks [5]. In contrast, biologically inspired signaling approaches, based on motion cues rather than explicit radio communication, have been discussed mainly in theoretical contexts [8] and rarely combined with hybrid perception for multi-platform cooperation. To the best of the authors' knowledge, no previous work addresses the integration of stereo vision, LiDAR clustering, indirect gestures cues, and mechanical sensor considerations in a unified UAV-UGV cooperation framework for defensive, human-centered applications.

3. SYSTEM OVERVIEW

The proposed framework comprises three main components: an aerial platform (UAV),

used for scouting and gesture-based signaling, a ground platform (UGV) responsible for local action and validation, and a perception subsystem based on a ZED Mini stereo camera and rotating LiDAR sensor. The UAV is equipped with its integrated RGB camera and executes predefined motion patterns to indicate points of interest or areas requiring intervention. The UGV operates in the vicinity of the perception subsystem and uses the fused stereo-LiDAR information to detect, localize, and interpret the UAV's position and motion cues.

In the considered configuration, the ZED Mini stereo camera and the LiDAR unit form a ground-level hybrid perception layer associated with the UGV side. Stereo vision provides RGB images and dense depth maps used for UAV detection and approximate 3D localization, while the LiDAR delivers a 3D point cloud from which clusters corresponding to the UAV can be extracted and metrically localized. The two sensing modalities are processed under a ROS-based pipeline [9] that handles time synchronization, coordinate frame management, and data visualization. No teleoperation or high-bandwidth communication between UAV and UGV is assumed; instead, guidance emerges from the perception of UAV motion patterns observed by the ground sensors.

From a mechanical standpoint, the stereo camera and LiDAR are mounted on a rigid support linked to the UGV platform, with their relative poses calibrated in a common reference frame. Sensor mounting height, tilt angle, and baseline directly influence the robustness of depth estimation under vibration and motion. These factors affect the accuracy with which the UAV and its gesture-like trajectories can be reconstructed in 3D, and are therefore explicitly considered when analyzing hybrid perception performance and cooperative behavior in GPS-denied environments.

4. STEREO VISION SUBSYSTEM

The stereo vision subsystem employs a ZED Mini camera [10] rigidly mounted on the UGV platform, acquiring synchronized left-right image pairs for disparity estimation and metric depth reconstruction.

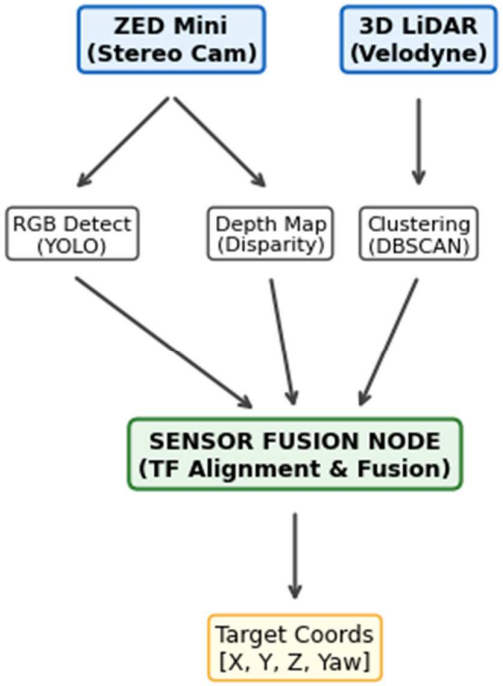


Fig. 1. Hybrid perception architecture integrating stereo vision (ZED Mini) and 3D LiDAR for UAV detection and coordinate estimation in a common reference frame via ROS TF alignment.

Depth is obtained by correlating pixel correspondents across the stereo baseline and converting disparity values into distance measurements based on calibrated intrinsic and extrinsic parameters. The resulting depth map provides dense per-pixel ranging information, enabling the detection and approximate localization of the UAV within the camera's field of view.

In the proposed configuration, depth acquisition is performed using the manufacturer's SDK, while RGB images are processed for UAV detection and bounding box extraction. The detected bounding box is associated with the corresponding depth region to estimate 3D position relative to the camera coordinate frame. Depth reconstruction accuracy depends primarily on baseline stability, feature richness, illumination, and the disparity search range. Under limited texture or high-exposure conditions, depth values may become unreliable, requiring filtering and confidence thresholding.

Stereo-based localization is sensitive to vibration and platform motion. When the UGV

operates on uneven terrain, local oscillations may affect disparity estimation and cause temporal depth fluctuations. Mechanical mounting thus plays an important role, as camera rigidness and stabilizing structures directly influence measurement robustness. Furthermore, the stereo field of view determines the spatial volume in which the UAV can be accurately detected using the current sensor geometry. These factors are considered during experimental evaluation and when comparing stereo performance with LiDAR-based ranging.

5. LIDAR SUBSYSTEM

A rotating 3D LiDAR sensor is employed on the UGV side to acquire metric point-cloud data of the surrounding environment. Unlike image-based sensing, LiDAR measures direct time-of-flight information, providing a dense set of 3D points that are independent of illumination, texture, or color. Each scan produces a spatial distribution of returns that can be processed to identify clusters corresponding to the UAV and to estimate its relative position with respect to the ground platform.

In the proposed configuration, raw point clouds are published over ROS as sensor_msg/PointCloud2 messages and processed using a DBSCAN-based clustering algorithm [11]. DBSCAN groups spatially consistent points while rejecting sparse outliers, enabling robust UAV segmentation even under partial occlusions. Once a UAV cluster is identified, its centroid is extracted and expressed in the LiDAR reference frame, which is subsequently transformed into the common coordinate system using ROS TF.

LiDAR returns are less affected by illumination conditions and vibration-induced motion blur; however, the sensor is sensitive to reflective materials, range saturation, and moving background elements. In addition, rotational LiDAR units produce non-instantaneous scans, meaning that rapidly moving points-of-interest may appear slightly distorted due to the sequential sampling process. From a mechanical standpoint, LiDAR mounting height and tilt angle directly influence near-field occlusion from the UGV chassis, whereas a higher tilt enhances aerial detection at the expense of local geometry resolution.

LiDAR measurements therefore complement stereo depth by providing reliable ranging under illumination variations and over longer distances, particularly when visual texture is limited. These properties motivate the combination of the two sensing modalities in the subsequent hybrid perception analysis, where the relative strengths and limitations of each sensor are compared in a unified reference frame.

6. HYBRID PERCEPTION AND COMPLEMENTARITY

Stereo vision and LiDAR provide fundamentally different sensing characteristics, and their complementary behavior can be exploited to obtain more reliable point-of-interest localization. Stereo depth estimation produces dense ranging information associated with image features, enabling detailed UAV segmentation and pixel-level depth reconstruction. However, stereo performance degrades under limited texture, strong illumination, or vibration-induced disparity noise. In contrast, LiDAR delivers accurate geometric estimates that are largely invariant to illumination, but lacks semantic awareness and produces lower spatial resolution in regions with sparse returns.

Hybrid perception is achieved by fusing stereo and LiDAR measurements in a common coordinate frame. UAV detection and bounding-box regions obtained from stereo images are associated with LiDAR clusters extracted through DBSCAN, allowing both sensing modalities to confirm the UAV's approximate location. When stereo depth is uncertain, LiDAR measurements act as fallback ranging, while stereo data provides spatial detail and appearance-based segmentation when visual conditions are favorable. The fusion requires consistent calibration and accurate sensor mounting, as geometric alignment errors propagate into the combined estimate.

Experimental observations indicate that the stereo offers reliable depth values at short and medium distances when sufficient visual texture is present, whereas LiDAR provides stable

returns over longer ranges and under high exposure. Low-light environments and vibration disturbances particularly affect stereo disparity, while LiDAR remains largely unaffected. Conversely, LiDAR suffers from reduced near-field resolution and occlusions induced by the UGV structure, which can be mitigated using stereo depth. These complementary characteristics motivate the hybrid perception strategy adopted in this work and are further analyzed with respect to sensor geometry and mechanical integration.

In addition to sensing complementarities, cooperative behavior relies on interpreting simple trajectory patterns performed by the UAV. The primary inspiration originates from the waggle dance observed in honey bees [8], where motion patterns encode directional information without explicit communication symbols. In the proposed framework, UAV motion can convey a coarse indication of the point of interest, which the UGV interprets using the hybrid perception output. The minimalist strategy reduces the dependence on persistent radio links and remains aligned with defensive and human-centered objectives.

6.1 Data Association and Fusion Logic

While modern 3D LiDAR sensors can provide accurate X, Y, Z, and Yaw coordinates, they inherently lack the semantic ability to distinguish the cooperative UAV from the dynamic aerial objects (e.g., birds, debris) or static overhangs. The hybrid approach leverages stereo vision to confirm the target identity and LiDAR to lock its precise metric position, ensuring the UGV follows the correct agent.

To mitigate the limitations of single-sensor configurations, a complementary fusion strategy based on gated association and confidence-driven switching is employed.

LiDAR provides high metric stability (3.46 cm positional jitter observed experimentally) and illumination – independent ranging, yet lacks semantic descriptors. Conversely, stereo vision supplies dense visual context but exhibits quadratic depth uncertainty with increasing distance. The fusion pipeline operates at two levels:

(1) Spatial Gating

LiDAR-extracted clusters are projected into the camera image plane using the calibrated extrinsic transformation. A valid association is confirmed if the projected centroid lies within the stereo-detected bounding box and satisfies an overlap threshold IoU > 0.3. This geometric constraint reduces false positives and enforces cross-modal consistency.

(2) Confidence-Based Switching

Depth estimation prioritization is determined by measurement reliability. For targets beyond 3 m, LiDAR depth is preferred due to lower range-dependent variance. For near-field interaction (<3 m) or when LiDAR cluster dispersion exceeds a threshold ($\sigma > 0.15$ m), the system re-weights toward stereo depth, which provides denser sampling in close-range scenarios.

This strategy enables graceful degradation under adverse sensing conditions while maintaining continuous tracking capability.

7. MECHANICAL AND MATHEMATICAL MODELING

7.1 UAV Kinematics and Gesture Modeling

The aerial platform communicates directional cues through specific motion patterns, biologically inspired by the waggle dance. The UAV state at time *t* is defined by its position and orientation vector in the local North-East-Down frame:

$$X_{UAV}(t) = [x(t), y(t), z(t), \psi(t)]^T \tag{1}$$

To convey a directional vector V_{target} towards the point of interest, the UAV executes a periodic trajectory. We model the "figure-eight" signaling pattern using parametric equations of a lemniscate, centered at the hovering point (x_0, y_0, z_0):

$$x(t) = x_0 + \frac{(\alpha \cdot cos(\omega t))}{1 + sin^2(\omega t)} cos(\theta) \tag{2}$$

$$y(t) = y_0 + \frac{(\alpha \cdot sin(\omega t) \cdot cos(\omega t))}{1 + sin^2(\omega t)} sin(\theta) \tag{3}$$

Where α represents the scaling amplitude of the gesture, ω is the angular frequency determining the signaling speed, and θ encodes the azimuth angle towards the target. The UGV perceives this trajectory over a temporal window,

extracting the dominant axis of motion to estimate θ.

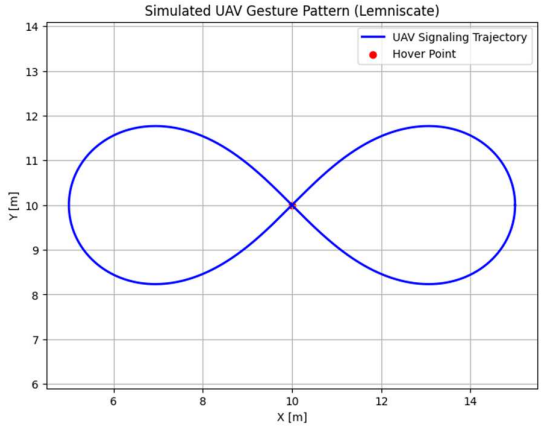


Fig. 2. Simulated UAV Gesture trajectory modeled as a lemniscate for directional encoding of the point-of-interest azimuth.

7.2 UGV Hybrid Perception Model

The ground platform integrates data from the stereo camera and the LiDAR unit. The stereo depth Z_s for a matched pixel pair with disparity d is given the standard triangulation model:

$$Z_s = \frac{f \cdot B}{d} \tag{4}$$

Where f is the focal length and B is the calibrated baseline between the left and the right optical centers. Since stereo depth uncertainty increases quadratically with distance, the system fuses this data with LiDAR measurements.

The LiDAR point cloud P_L is processed to extract the UAV cluster C_{UAV}. A point $p_i \subset P_L$ belongs to the cluster if it satisfies the density reachability condition defined by the DBSCAN parameters (radius and minimum points).

Finally, to fuse the two modalities, the LiDAR coordinates are transformed into the camera reference frame using the extrinsic calibration matrix composed of a rotation R and translation T, determined during the mechanical integration phase:

$$P_{cam} = R \cdot P_{lidar} + T \tag{5}$$

This transformation ensures that the geometric centroid of the UAV detected by the LiDAR is spatially aligned with the semantic bounding box provided by the stereo vision system, allowing for a unified estimation of the UAV's position relative to the UGV.

8. EXPERIMENTAL ANALYSIS

To validate the proposed perception framework and mechanical integration, a real-world test sequence consisting of 100 LiDAR frames was recorded during UAV flight within the sensing volume of the UGV-mounted sensor array.

8.1 Data Processing Methodology

The experimental metrics were computed from 100 consecutive LiDAR frames exported as CSV files containing multiple parameters including XYZ point coordinates. Timestamp information was partially inconsistent; therefore, a fixed sampling interval $\Delta t = 0.1$ s (≈ 10 Hz) was assumed for temporal derivative calculations.

For each frame: (1) the raw point clouds were filtered using geometric constraints, (2) DBSCAN clustering was applied using empirically tuned parameters (ε and minimum points), (3) the UAV cluster was identified based on spatial coherence and centroid continuity and (4) the cluster centroid was extracted for temporal analysis.

Subsequent kinematic quantities were derived using finite difference approximations over the centroid trajectory.

8.2 Quantitative Assessment

The following quantitative metrics were computed to evaluate detection robustness, positional stability, dynamic behaviour, and geometric consistency of the LiDAR-based UAV tracking pipeline. All indicators were derived from centroid trajectories extracted from 100 consecutive LiDAR frames.

Detection rate was defined as the ratio between successfully tracked frames and the total number of recorded frames. In the present experiment, 86 out of 100 frames resulted in valid cluster association, yielding a detection reliability of 86%. Missed detections occurred primarily during rapid trajectory transitions that temporarily reduced cluster coherence.

$$DetectionRate = \frac{N_{tracked}}{N_{total}} \tag{6}$$

where $N_{tracked} = 86$ and $N_{total} = 100$.

Positional Jitter was computed as the mean standard deviation of centroid coordinates along X, Y and Z axes over time. This metric quantifies spatial stability of the tracked UAV position and reflects the intrinsic measurement noise and clustering variability. The resulting value of 3.46 cm indicates centimeter-level localization consistency under dynamic UAV motion.

$$J = \frac{\sigma_x + \sigma_y + \sigma_z}{3} \tag{7}$$

where

$$\sigma_x = std(x(t)) \tag{8}$$

Velocity was estimated using first-order finite differences of consecutive centroid position, assuming a constant sampling interval of approximately 0.1 seconds (10Hz). The analysis yielded an average speed of 0.236 m/s, with peak values reaching 2.14 m/s during rapid directional changes.

$$v(t) = \frac{\|C(t) - C(t-1)\|}{\Delta t} \tag{9}$$

Acceleration was derived from successive velocity variations. To mitigate noise amplification inherent to discrete differentiation, acceleration values were smoothed prior to statistical evaluation. The mean absolute acceleration was measured at 0,287 m/s^2, confirming stable tracking under moderate dynamic conditions.

$$a(t) = \frac{v(t) - v(t-1)}{\Delta t} \tag{10}$$

The reconstructed trajectory length was computed as the cumulative displacement between consecutive centroid positions. The total measured path length was 2.008 meters, confirming geometric continuity of the tracking process.

$$L = \sum_{i=1}^{N} \|c_i - c_{i-1}\| \tag{11}$$

Cluster dispersion was defined as the mean Euclidean distance between LiDAR points belonging to the UAV cluster and the corresponding centroid. This metric evaluates geometric compactness of the detected target and serves as an indicator of clustering robustness. The average dispersion remained within 0.108 ± 0.017 meters, demonstrating consistent spatial grouping.

$$D = \frac{1}{n}\sum_{i=1}^{n} \|p_i - C\| \tag{12}$$

Together, these indicators provide a multi-dimensional evaluation of tracking reliability, spatial precision, dynamic responsiveness, and geometric stability. The observed detection rate and centimeter-level positional jitter confirm the suitability of the proposed pipeline for short-range cooperative UAV guidance in GPS-denied environments.

Table 1

Quantitative performance metrics of LiDAR-based UAV tracking over 100 frames.

Metric	Value
Detection Rate	86%
Frames Tracked	86/100
Sampling Rate	10 Hz
Positional Jitter	3.46 cm
Total Path Length	2.008 m
Average Speed	0.236 m/s
Peak Speed	2.14 m/s
Average Acceleration	0.287 m/s^2
Average Points per cluster	167 ± 42.6
Active Beams	10.9/16
Cluster Dispersion	0.108 ± 0.017

The static UGV configuration isolates sensor performance from vehicle-induced vibration, providing a controlled baseline for perception evaluation.

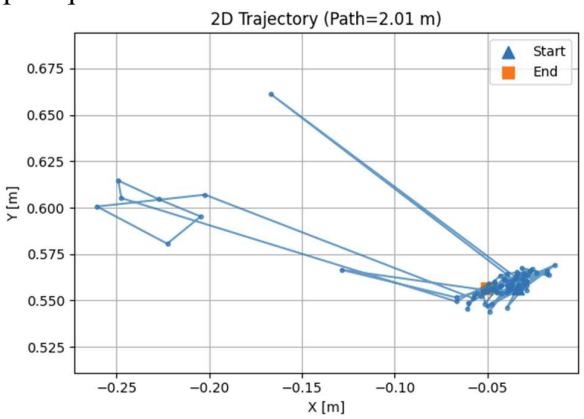


Fig. 3. Reconstructed 2D centroid trajectory of the UAV cluster extracted from 100 consecutive LiDAR frames.

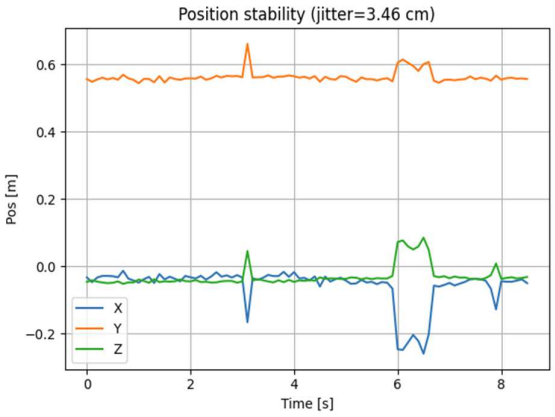


Fig. 4. Temporal stability of LiDAR centroid estimation along X, Y, Z axes (mean positional jitter = 3.46 cm)

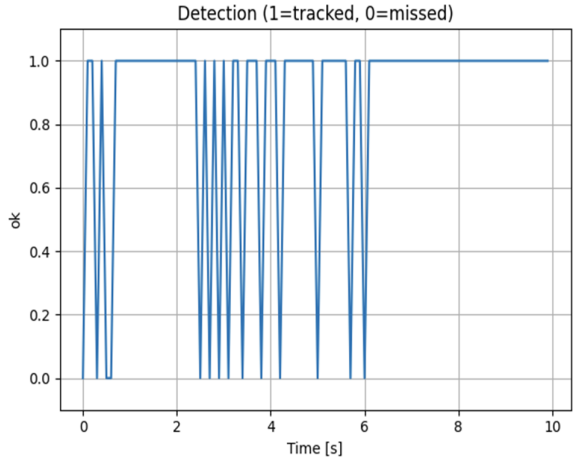


Fig. 5. Frame-wise detection outcome over time (1 = valid cluster association, 0 = missed detection).

The fixed sampling assumption ($\Delta t = 0.1s$) introduces a bounded temporal uncertainty in velocity and acceleration estimation.

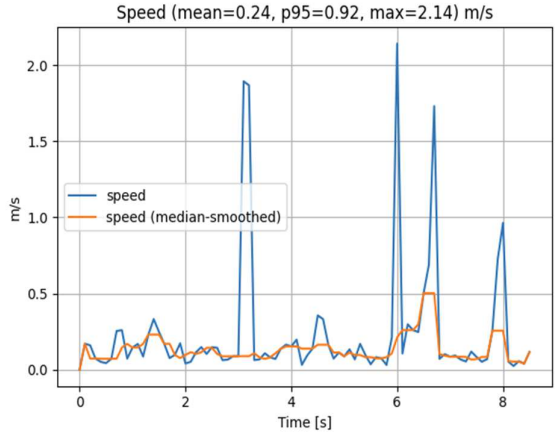


Fig. 6. Estimated UAV centroid speed derived from finite difference approximation ($\Delta t = 0.1s$).

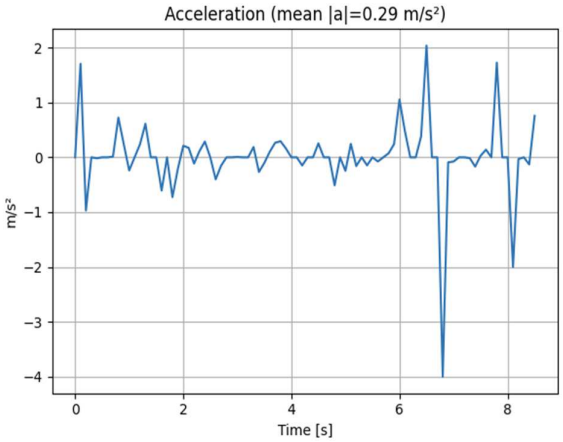


Fig. 7. Smoothed centroid acceleration profile computed from consecutive velocity estimates.

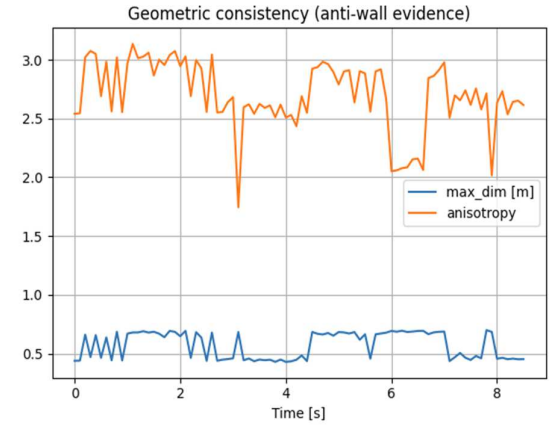


Fig. 8. Geometric consistency indicators of the detected UAV cluster, including maximum dimension and anisotropy over time.

Detection failures correlated with temporary increases in cluster dispersion and reduced beam participation.

8.3 Sensor Mounting Stability Under Static Conditions

During the experimental campaign, the UGV platform remained stationary while the UAV performed dynamic maneuvers within the sensing volume. This configuration isolates sensing ability from platform-induced vibrations and allows assessment of intrinsic sensor and clustering consistency.

The observed centroid positional jitter of 3.46 cm reflects measurement noise and clustering variability rather than mechanical oscillations. Given that the LiDAR and stereo sensors were rigidly mounted, this low jitter value confirms geometric consistency of the

sensing pipeline under static mechanical conditions.

It is important to note that dynamic UGV motion may introduce additional vibration-related disturbances, which were not evaluated in the present study and remain subject to future investigation.

8.4 Sensor Complementarity Validation

While LiDAR provides accurate geometric coordinates, it cannot inherently discriminate between UAV targets and other dynamic objects with similar spatial dimensions. The stereo subsystem provides semantic validation through visual detection, reducing the risk of false association and enhancing overall system reliability.

9. TECHNOLOGY READINESS LEVEL

The proposed system currently corresponds to Technology Readiness Level (TRL) 4, defined as component and subsystem validation in a laboratory environment. The presented results demonstrate experimental verification of the sensing architecture using real hardware and dynamic UAV tracking, without extending to operational field deployment.

10. CONCLUSION

This paper presented a hybrid perception framework for UAV-UGV cooperation in GPS-denied and communication-degraded environments, focused on safe and defensive applications. The approach combines stereo vision and LiDAR sensing on the UGV side to achieve robust point-of-interest localization through complementary depth and geometric information. Experimental observations highlight the strengths and limitations of each sensing modality and motivate their combined use under adverse environmental conditions.

Beyond sensing, the proposed framework explores a minimalist coordination strategy inspired by the waggle dance observed in honey bees. By encoding coarse spatial cues through UAV motion patterns rather than explicit communication messages, cooperation can be achieved with reduced dependence on continuous radio links. This biologically inspired signaling mechanism supports basic coordination while remaining aligned with human-centered and non-intrusive operational objectives.

The main contribution of this work lies in integrating hybrid stereo-LiDAR perception, biologically inspired gestural cues, and system-level considerations into a unified UAV-UGV cooperation framework. The results suggest that such an approach can enhance situational awareness and reduce human exposure to risk in environments where conventional localization and communication methods are unreliable. Future work will focus on refining mechanical integration aspects, improving sensing robustness under dynamic conditions, and extending the framework toward more complex cooperative behaviors.

11. REFERENCES

[1] Michael, N., Shen, S., Mohta, K., et al., *Collaborative mapping of an earthquake-damaged building via ground and aerial robots*, Journal of Field Robotics, Vol. 29, Issue 5, pp. 832-841, 2012.

[2] Rizk, Yara & Awad, Mariette & Tunstel, E.. (2019). *Cooperative Heterogeneous Multi-Robot Systems: A Survey*.

[3] Çaşka, Serkan & Gayretli, Ahmet. (2014). *A survey of UAV/UGV collaborative systems*. 453-463.

[4] Chen, P., Zhao, X., Zeng, L., Liu, L., Liu, S., Sun, L., Li, Z., Chen, H., Liu, G., Qiao, Z., Qu, Y., Xu, D., Li, L., & Li, L. (2025). *A Review of Research on SLAM Technology Based on the Fusion of LiDAR and Vision*. Sensors, 25(5), 1447. https://doi.org/10.3390/s25051447

[5] Munasinghe I, Perera A, Deo RC. *A Comprehensive Review of UAV-UGV Collaboration: Advancements and Challenges*. Journal of Sensor and Actuator Networks. 2024; 13(6):81. https://doi.org/10.3390/jsan13060081

[6] A. Bewley, Z. Ge, L. Ott, F. Ramos and B. Upcroft, *Simple online and realtime tracking*, 2016 IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 2016, pp. 3464-3468, doi: 10.1109/ICIP.2016.7533003.

[7] Liu, H., Long, Q., Yi, B. et al. *A survey of sensors based autonomous unmanned aerial vehicle (UAV) localization techniques*. Complex Intell. Syst. 11, 371 (2025). https://doi.org/10.1007/s40747-025-01961-2

[8] Von Frisch, K., *The Dance Language and Orientation of Bees*, Harvard University Press, 1967.

[9] Quigley, M., *ROS: an open-source Robot Operating System*, ICRA workshop on open source software, Vol. 3, No. 3.2, pp. 5, Kobe, Japan, 2009.

[10] Stereolabs, *ZED Mini - Mixed Reality Camera* Technical Documentation, https://www.stereolabs.com/zed-mini/

[11] Ester, M., Kriegel, H. P., Sander, J., Xu, X., *A density-based algorithm for discovering clusters in large spatial databases with noise*, Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining, pp. 226-231, Portland, 1996.

Coordonare hibridă UAV-UGV pentru localizarea punctelor de interes în medii fără semnal GPS utilizând vedere stereoscopică, LiDAR și semnalizare bazată pe gesturi

Rezumat: Această lucrare prezintă un cadru hibrid de percepție și coordonare care permite echipelor mixte UAV-UGV să identifice și să localizeze puncte de interes în medii lipsite de semnal GPS sau cu comunicații degradate. Abordarea este proiectată strict pentru aplicații defensive și de protecție, unde îmbunătățirea conștientizării situaționale și reducerea expunerii umane la risc constituie obiective primare. Sistemul integrează vederea stereoscopică, clusterizarea bazată pe LiDAR și semnalizarea prin gesturi bio-inspirată ca un nivel de comunicare minimalist și neintruziv între platformele aeriene și cele terestre. Mediul ROS este utilizat exclusiv pentru sincronizarea senzorilor și alinierea spațială. Sunt incluse, de asemenea, considerente mecanice – precum montarea senzorilor, reziliența la vibrații și geometria câmpului vizual – pentru a asigura operarea în siguranță pe teren accidentat. Observațiile experimentale evidențiază atuurile complementare ale percepției stereo și LiDAR în susținerea unei colaborări multi-robot responsabile, centrată pe factorul uman.

Patricia PREDA, Eng., Master's student, National University of Science and Technology POLITEHNICA Bucharest, patriciapreda03@gmail.com, +40 730 123 803, 313 Splaiul Independentei, Bucharest, Romania.

Lucian-Ștefăniță GRIGORE, CS-I PhD.eng.habil, Doctoral supervisor, "Ferdinand I" Military Technical Academy, Center of Excellence in Robotics and Autonomous Systems, Faculty of Communications and Electronic Systems for Defense and Security, lucian.grigore@mta.ro, +40 733 910 730, 39-49 George Cosbuc Blvd., Bucharest, Romania.

Eliza MARIN, Eng., R&D department at UM 01616 Bucharest, meliza@logmil.ro, +40 757 311 664, 33-35 Bucuresti-Magurele Road, Bucharest, Romania.