Computers, Materials & Continua

Tech Science Press

https://doi.org/10.32604/cmc.2026.085054

REVIEW

Recent Advances in UAV-Based SLAM: A Survey

Yaolei Wang[1], Wangyan Li[1,*] and Guoliang Wei[2]

[1]School of Mathematics, University of Shanghai for Science and Technology, Shanghai, China
[2]Business School, University of Shanghai for Science and Technology, Shanghai, China

*Corresponding Author: Wangyan Li. Email: wangyan_Li@usst.edu.cn

Received: 04 May 2026; Accepted: 01 July 2026; Published: 13 August 2026

ABSTRACT: With the rapid development of *unmanned aerial vehicle* (UAV) technologies, *simultaneous localization and mapping* (SLAM) has emerged as a key enabling paradigm for autonomous navigation and environmental perception. This paper presents a comprehensive survey of recent trends in UAV-based SLAM. First, we review the fundamental components of UAV-based SLAM systems, including commonly used onboard sensors and front-end odometry methods such as visual odometry, visual-inertial odometry, and LiDAR-inertial odometry, which provide reliable ego-motion estimation. Next, we summarize back-end methodologies that enhance estimation accuracy and global consistency, covering pose graph optimization, 3D reconstruction techniques, filter-based SLAM, fusion-based multi-UAV SLAM, and emerging blockchain-enabled frameworks for secure and distributed mapping. Furthermore, representative application scenarios are discussed, including a newly proposed category of collaborative UAV-SLAM platforms, termed the UAV-plus system. We also identify several remaining research gaps and highlight future research directions, including observability challenges, operation in imperfect and dynamic environments, and the miniaturization of UAV-based SLAM systems.

KEYWORDS: UAV-based SLAM; multi-UAV systems; distributed UAV-based SLAM; front-end odometry; pose graph optimization; 3D reconstruction; sensor fusion; blockchain; back-end optimization

1 Introduction

With the rapid development of aerial robotic platforms and sensing technologies [1], *unmanned aerial vehicles* (UAVs) have been increasingly deployed across a wide range of application domains, including environmental monitoring, infrastructure inspection, precision agriculture, disaster response, and urban management [2]. This widespread adoption has driven growing interest in more autonomous, robust, and intelligent UAV systems. In many practical scenarios, UAVs are required to operate in complex and *global navigation satellite systems* (GNSS)-denied environments, where GNSS is a general term referring to satellite-based positioning systems, while GPS is the most widely used system in UAV-based SLAM applications [3]. Such as indoor spaces, urban, forests, and tunnels [3], where satellite-based positioning is often unreliable due to occlusion and interference [4]. As a result, accurate self-localization and environment perception become fundamental requirements for autonomous UAV operation. In this context, *simultaneous localization and mapping* (SLAM) has emerged as a key enabling paradigm to address these challenges [5–8], and has become a central research trend in UAV-based autonomous systems [9–12].

To enable reliable SLAM systems in real-world UAV applications, perception and motion estimation must operate under stringent platform and environmental constraints [2,13–15]. UAV systems are inherently limited in payload, energy, and sensing range, while often deployed in dynamic, unstructured, and GNSS-denied environments [3], which imposes fundamental challenges on sensing reliability and state estimation accuracy [4]. In this context, the integration of heterogeneous sensors has become a necessary design paradigm, as individual sensing modalities are insufficient to ensure robust perception across diverse conditions [11,16]. Within the UAV-based SLAM pipeline, the front-end plays a central role in transforming raw sensor measurements into motion constraints, directly affecting the stability and accuracy of the overall system [17,18]. However, front-end estimation remains particularly sensitive to factors such as feature degradation, motion blur, and aggressive platform dynamics, leading to error accumulation and potential system failure [19,20]. These limitations have driven the evolution of UAV-based SLAM from single-sensor perception toward tightly coupled multi-modal odometry frameworks [21], with increasing focus on improving the resilience and adaptability of front-end estimation under adverse conditions, laying the foundation for robust UAV navigation in complex real-world scenarios [11,22].

The front-end of UAV-based SLAM provides reliable local motion estimation from sensor measurements, and achieves long-term accuracy and consistency requires integrating information across time and space, while the back-end serves as a critical component that refines state estimation by incorporating global constraints and performing consistent inference over accumulated observations [23–26]. In particular, optimization-based approaches, such as pose graph optimization [20,27], have become fundamental tools for mitigating drift and enforcing global geometric consistency [28], while also supporting high-quality 3D reconstruction for environment representation and downstream perception tasks [29,30]. In parallel, the need to explicitly model uncertainty and nonlinear system dynamics has led to the development of filtering-based SLAM frameworks [3,31,32], which provide recursive state estimation under noisy and partially observable conditions [31]. As UAV-based SLAM systems evolve toward large-scale and cooperative deployments, the role of the back-end has further expanded to multi-agent fusion and distributed mapping [22,33,34], where maintaining estimation consistency under communication constraints becomes a central challenge. More recently, the increasing deployment of networked and cooperative UAV systems has introduced new requirements for data reliability and secure information sharing, motivating the exploration of blockchain-enabled UAV-based SLAM frameworks [35,36], where local maps and sensory data generated at the UAV end are securely recorded and verified through distributed blockchain ledgers, enabling tamper-resistant data sharing and reliable map fusion at the server end.

Driven by the growing demand for autonomous perception in complex environments, UAV-based SLAM systems are increasingly deployed in real-world applications. They are often operated in collaborative settings, as illustrated in Fig. 1d. In such UAV-plus paradigms [37–40], UAVs cooperate with *unmanned ground vehicles* (UGVs) [40], *unmanned underwater vehicles* (UUVs) [37], and human operators to accomplish shared tasks, relying on continuous spatial information exchange for coordination and decision-making. In addition, UAV-based SLAM systems are also deployed as standalone agents in diverse real-world scenarios, including agriculture, infrastructure inspection, environmental monitoring, and disaster response [2,41–44].

Although several surveys on SLAM and UAV navigation [2,4,45–47] have been reported in recent years, most of them focus on certain aspects of UAV-based visual SLAM, for example semantic SLAM, UAV mechanical architectures and control techniques. Recently, reference [47] have reviewed learning-based feature extraction, loop detection, semantic segmentation, and neural scene representation in visual SLAM. However, these surveys do not comprehensively discuss UAV-based SLAM from a systematic level. This motivates us to provide a comprehensive review on topics from data sources (Topic A, Table 1) to front-end odometry (Topic B, Table 1), and back-end optimization (Topic C, Table 1), multi-UAV collaboration (Topic D, Table 1). In particular, emerging topics such as security-enabled SLAM (Topic E, Table 1) and the UAV-plus paradigm (Topic F, Table 1) remain largely unexplored in existing surveys. Their growing importance further motivates this review.

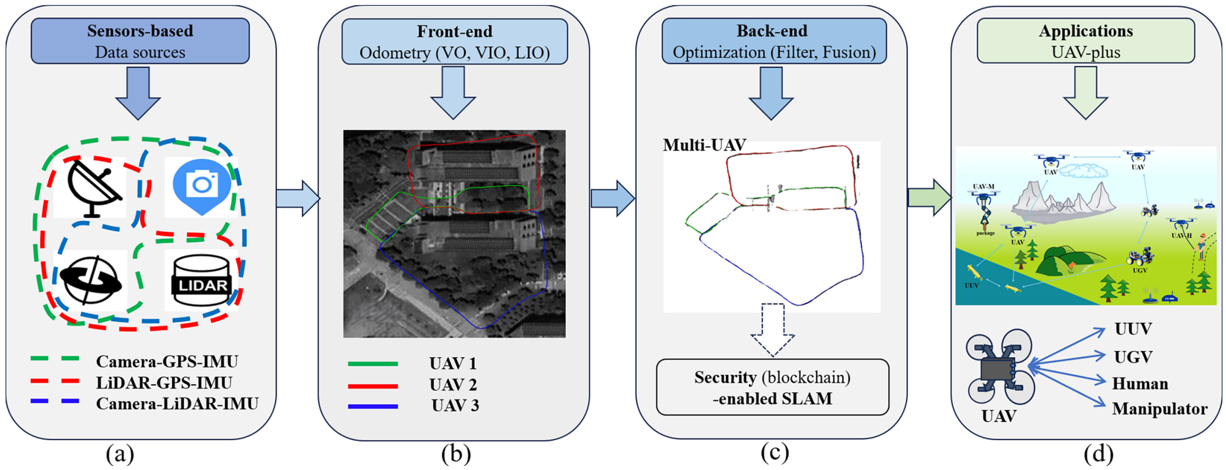


Figure 1: Schematic diagram of UAV-based SLAM platform.

Table 1: Comparison between this work and referenced UAV-SLAM survey papers.

Surveys Ref.	Topics*						Focus
	A	B	C	D	E	F	
[4]-2015	✓	✓					Visual SLAM
[2]-2016	✓	✓	✓				Semantic SLAM
[46]-2022						✓	UAV mechanical structure
[45]-2023			✓	✓			RL-based SLAM**
[47]-2023		✓		✓		✓	DL-based SLAM***
This work	✓	✓	✓	✓	✓	✓	Topics **A~F**

Note: *A: Sensors-based data sources; B: Front-end odometry; C: Back-end optimization; D: Multi-UAV collaboration; E: Security-enabled SLAM; F: Applications. **RL: Reinforcement learning. ***DL: Deep learning.

Motivated by the aforementioned developments, this review provides a systematic survey of recent advances in UAV-based SLAM systems, structured along the pipeline illustrated in Fig. 1. In addition, Fig. 2 presents a chronological overview of representative UAV-based SLAM technologies, highlighting the progression from early monocular and visual SLAM frameworks to tightly coupled visual–inertial and LiDAR–inertial fusion systems, as well as recent intelligent SLAM approaches. This survey primarily focuses on representative UAV-based SLAM studies published in recent years, including both classical and emerging frameworks related to visual, visual–inertial, LiDAR–inertial, collaborative, and learning-assisted SLAM systems. The reviewed literature was collected from major scientific databases, including IEEE Xplore, ScienceDirect, and Google Scholar, using keywords such as “UAV-SLAM”, “visual SLAM”, “VIO”, “LIO”, “multi-UAV SLAM”, and “UAV mapping”. Priority was given to highly relevant and influential studies closely related to UAV autonomous navigation, perception, and mapping.

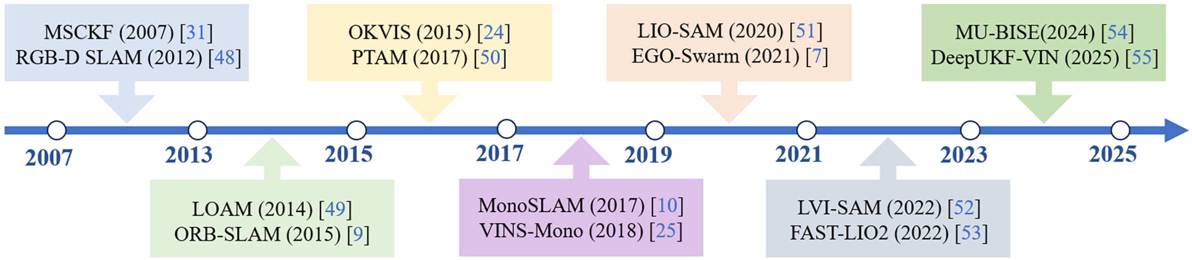


Figure 2: Chronological development of UAV-based SLAM technologies [7,9,10,24,25,31,48–55].

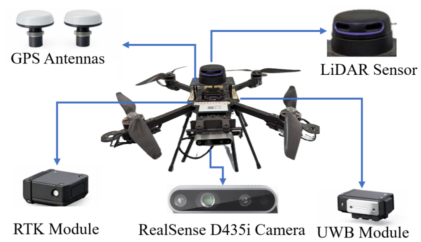
The main contributions of this survey are summarized as follows: 1) A system-level taxonomy and pipeline review of UAV-based SLAM systems is provided, covering sensors-based data sources, front-end odometry, back-end optimization, and UAV-plus applications; 2) Emerging UAV-SLAM trends are discussed, including 3D reconstruction, fusion-based multi-UAV SLAM, and a novel interesting security (blockchain)-enabled SLAM systems; 3) A new application-oriented categorization, termed *UAV-plus*, is introduced to summarize UAV-centered heterogeneous collaborative SLAM systems, together with emerging societal UAV-SLAM applications. The remainder of this paper is organized as follows. Section 2 (corresponding to Fig. 1a,b) presents the overall framework of UAV-based SLAM systems, including heterogeneous sensing configurations and front-end odometry methods. This is followed by Section 3 (corresponding to Fig. 1c), where back-end estimation and optimization techniques are examined, with particular emphasis on pose graph optimization, 3D reconstruction, filter-based SLAM, fusion-based multi-UAV SLAM, and blockchain-enabled UAV-based SLAM Methods. Section 4 (corresponding to Fig. 1d) highlights several representative real-world applications such as UAV-plus collaborative systems as well as societal scenarios. Finally, Section 5 summarizes the key insights and outlining several future research directions.

2 UAV-Based SLAM Systems

Building upon the background discussed in the previous section, this section introduces the key components of UAV-based SLAM systems. In particular, we first review commonly used onboard sensors for UAV platforms, followed by an overview of typical front-end odometry frameworks.

2.1 Sensors for UAV-Based SLAM Platforms

UAV platforms typically integrate multiple onboard sensors to acquire both motion and environmental information for SLAM. These sensors either provide direct measurements of the UAV state or support its reconstruction through indirect observations. To achieve accurate and robust state estimation, multi-sensor fusion is widely adopted to compensate for the limitations of individual sensing modalities, as illustrated in Fig. 3a. Among these sensors, the **IMU** serves as a fundamental component by providing high-frequency measurements of angular velocity and linear acceleration [18], which are essential for short-term motion propagation and dynamic state prediction. Based on this inertial estimate, the **GPS** provides absolute position information in a global coordinate frame, while *real-time kinematic* (**RTK**) further improves positioning accuracy through differential carrier-phase (CP) positioning with GNSS corrections. The combination of IMU and GPS/RTK enables both high-rate motion estimation and globally consistent localization in outdoor environments. However, in GPS-denied or degraded environments, exteroceptive sensing becomes crucial for indirectly reconstructing the UAV state through interactions with the surrounding environment. In such cases, **cameras** (including RGB-D and thermal cameras) provide rich visual cues, including geometric features and semantic information, enabling environment perception in a manner analogous to human vision [25,48]. RGB-D cameras provide direct depth measurements for indoor and short-range scenarios, while thermal cameras are less sensitive to illumination variations in night-time and smoke-filled environments. To further enhance geometric consistency and robustness, **LiDAR** actively scans the environment to generate accurate range measurements and dense point clouds, offering reliable geometric structure for mapping and localization. When fused with high-frequency **IMU** measurements, these sensors further refine motion estimation, constrain drift, and provide a more complete description of the environment in complex scenarios [56]. As a complementary sensing source, *ultra-wideband* (**UWB**) provides robust short-range relative positioning by radio frequency time-of-flight (ToF) and is particularly useful in indoor or infrastructure-supported environments where global navigation signals are unreliable. A qualitative comparison of the commonly used onboard sensors for UAV-based SLAM is summarized in Fig. 3b, including their sensing characteristics, deployment cost, weight, power consumption, accuracy ranges, operating environments, and practical limitations.



(a)

Sensors	Costs (USD)	Weight (g)	Power (W)	Accuracy ranges (m)	Environments	Characteristics
GPS	100-1000	100-500	1-5	0.1-2	Outdoor	Satellite positioning
IMU	< 100	< 100	< 1	0.1-2	Indoor/Outdoor	Inertial measurements
Camera	< 100	< 100	< 1	0.1-2	Indoor/Outdoor	Visual data
LiDAR	> 1000	> 500	> 5	< 0.1	Indoor/Outdoor	Point cloud
UWB	100-1000	< 100	< 1	0.1-2	Indoor	Radio frequency ToF
RTK	100-1000	100-500	1-5	< 0.1	Outdoor	CP positioning

(b)

Figure 3: (**a**): Typical UAV platform; (**b**): advantages and limitations of common onboard sensors [57].

In complex UAV application scenarios, the use of multiple sensors has proven effective in improving estimation robustness and accuracy. Based on these onboard sensors in Fig. 3, several typical sensor configurations (e.g., Camera-GPS-IMU, LiDAR-GPS-IMU, and Camera-LiDAR-IMU) have been widely adopted in UAV-based SLAM systems, as shown in the Fig. 1a. For instance, Camera-GPS-IMU systems combine camera and IMU measurements for motion estimation and navigation [31,55,58]. LiDAR-GPS-IMU systems integrate LiDAR with IMU data to achieve accurate geometric perception and robust state estimation [16,53,56]. Furthermore, Camera-LiDAR-IMU fusion frameworks exploit the complementary characteristics of multiple sensing sources to enhance robustness and adaptability, particularly under highly dynamic conditions and in perceptually degraded or GPS-denied environments [4,52]. These fusion strategies improve front-end odometry performance in terms of accuracy, robustness, and long-term consistency. With the onboard sensors providing raw measurements, the focus shifts to the processing of these data within the front-end odometry for motion estimation, which will be discussed in the following subsection.

Neuromorphic vision sensors (including event camera) have recently emerged as a new class of visual sensing devices, which continuously monitor scene illumination changes and output high-temporal-resolution asynchronous neuromorphic signals, (such as spikes or events), making them suitable for high-speed UAV motion and challenging illumination conditions [59]. In addition, radar sensors, especially millimeter-wave radar, provide robust ranging and localization capabilities under adverse weather and low-visibility conditions, thereby complementing vision-based sensing in UAV-based SLAM [60].

2.2 Front-End Odometry in UAV-Based SLAM

Front-end odometry plays a critical role in UAV-based SLAM systems, as it estimates the relative motion of the platform based on sequential sensor measurements. From the perspective of sensor configuration and motion constraint modeling, UAV-based SLAM odometry methods can be broadly categorized into three types: *visual odometry* (VO) [17,18], *visual-inertial odometry* (VIO) [9–11], and *LiDAR-inertial odometry* (LIO) [16,53,56]. In the following, we will introduce each of these odometry methods in turn.

1) VO: It estimates the ego-motion of a camera by tracking visual features across consecutive image frames and recovering the relative pose through geometric constraints, such as geometry or reprojection error minimization [61]. A typical VO pipeline consists of feature extraction, feature matching or tracking, and motion estimation [17]. VO provides accurate short-term motion estimation with relatively low computational cost [62]. However, in the absence of absolute scale and global reference constraints, it is inherently susceptible to drift accumulation over time and suffers from scale ambiguity in monocular settings [50]. These limitations become more evident in UAV applications, where rapid motion and frequent viewpoint changes can degrade feature tracking reliability [18]. To address these issues, inertial measurements are often incorporated with visual observations, giving rise to VIO systems.
2) VIO: To mitigate the drift and scale ambiguity of VO, inertial measurements are fused with visual observations to form VIO systems, which provides accurate and high-frequency ego-motion estimation by tightly coupling visual and inertial data. Early approaches such as MSCKF [31] achieve real-time performance using a sliding-window formulation without explicitly estimating landmarks. More recent methods adopt tightly-coupled nonlinear optimization, as demonstrated by *open keyframe-based visual-inertial SLAM* (OKVIS), which formulates VIO as a keyframe-based nonlinear least-squares problem for accurate real-time odometry, and VINS-Mono, which further enhances odometry robustness by integrating loop closure. To improve long-term consistency and robustness beyond local odometry, the ORB-SLAM series [9–11] further integrates visual-inertial odometry into a unified SLAM framework, with ORB-SLAM3 supporting multiple sensor configurations while maintaining consistent long-term performance.
3) LIO: In a typical LIO framework, the IMU provides high-frequency state propagation, while LiDAR measurements are used to perform scan matching or point cloud registration against local maps to estimate relative motion. The fusion is often implemented in a tightly coupled framework, where LiDAR geometric constraints are directly incorporated into filtering and optimization-based estimators to correct drift accumulated during inertial propagation [53]. Compared with vision-based methods, LIO utilizes explicit geometric structure from point clouds, making it less sensitive to illumination changes and texture deficiency [51]. As a result, LIO demonstrates strong robustness in challenging environments such as texture-poor scenes, low-light conditions, and large-scale outdoor scenarios [16,49].

To summarize the characteristics of different front-end odometry approaches, Table 2 compares representative VO, VIO, and LIO systems that have significantly influenced the development of UAV-based SLAM in recent years, including both classical frameworks and recent tightly coupled multi-sensor approaches. The comparison includes sensor configuration, scale observability, motion robustness, illumination sensitivity, performance in texture-poor environments, drift characteristics, computational complexity, typical estimation frameworks. Apart from the system-level comparison of VO, VIO, and LIO, the practical performance of front-end odometry also depends heavily on several key implementation modules, including feature extraction, feature tracking, data association, and robustness enhancement strategies.

2.3 Benchmark Datasets Evaluation Metrics

In this section, we compare five widely used benchmark datasets, including the indoor EuRoC dataset [63] (stereo + IMU), indoor/outdoor environment TUM-VI [64] (camera + IMU), outdoor environment KITTI [65] (camera + LiDAR), simulation environment AirSim [66] (synthetic multi-sensor), and large-scale industrial environment Hilti SLAM Challenge [67] (multi-sensor + LiDAR), from the perspectives of sensor type, environment, and typical usage. Table 3 summarizes representative evaluation results of VIO, VO, and LIO methods using commonly adopted metrics, including *absolute trajectory error* (ATE)-RMSE, *relative pose error* (RPE)-RMSE, drift, runtime, and memory usage. These metrics jointly evaluate global accuracy, local consistency, accumulated drift, computational efficiency, and onboard deployment feasibility. As shown in Table 3, VO achieves lower trajectory errors on several visual benchmarks, such as EuRoC, TUM-VI, and AirSim, but usually requires higher memory usage than VIO. In contrast, LIO shows clear advantages on the Hilti industrial dataset, where LiDAR-based geometric constraints provide more reliable localization in complex structured environments. These results further confirm that no single method is universally optimal; instead, the selection of VO, VIO, or LIO should depend on sensing payload, environmental texture, illumination conditions, motion intensity, and onboard computational resources.

Table 2: Comparison of VO, VIO, and LIO for UAV-based SLAM.

Categories	**VO**	**VIO**	**LIO**
Sensor configuration	Camera	Camera + IMU	LiDAR + IMU
Motion robustness	Low	Medium	High
Sensitivity to illumination	High	High	Low
Texture-poor scenes	Poor	Medium	High
Drift characteristics	High drift	Moderate drift	Low drift
Computational complexity	Low	Medium	Medium–High
Representative systems	DSO [62], PTAM [50], MSCKF [31]	OKVIS [24], VINS-Mono [25], ORB-SLAM3 [11]	LIO-SAM [51], FAST-LIO2 [53], Swarm-LIO [56]

Table 3: Comparison of VIO, VO, and LIO methods on different datasets.

Dataset**	**ATE-RMSE (m)**			**RPE-RMSE (m)**			**Drift (m)**			**Runtime (s)**			**Memory Usage (MB)**		
	VIO	**VO**	**LIO***	**VIO**	**VO**	**LIO**	**VIO**	**VO**	**LIO**	**VIO**	**VO**	**LIO**	**VIO**	**VO**	**LIO**
EuRoC	0.273	0.110	N/A	0.016	0.008	N/A	0.069	0.066	N/A	199.0	227.4	N/A	100.1	694.8	N/A
TUM-VI	1.692	0.073	N/A	0.037	0.008	N/A	0.073	0.001	N/A	152.9	376.5	N/A	82.9	810.1	N/A
KITTI	52.333	52.442	48.524	1.137	1.138	1.171	17.348	17.401	17.435	69.2	78.7	72.3	80.4	639.5	350.4
AirSim	0.527	0.015	N/A	0.254	0.006	N/A	0.119	0.002	N/A	46.6	43.4	N/A	77.5	561.2	N/A
Hilti	2759.2	N/A	0.11	852.5	N/A	0.053	20.81	N/A	0.001	79.17	2267.9	1762.6	N/A	594.4	178.8

Note: *: VIO: Vins-fusion [25], VO: ORB-SLAM2 [10], LIO: FAST-LIO2 [53]; **: EuRoC: MH_01 (easy); TUM-VI: room3; KITTI_Raw: 2011_09_26_drive_0093_sync; AirSim: tartanair-shibuya01; Hilti SLAM Challenge (Hilti): IC_Office_1.

Overall, the front-end odometry provides local motion estimation from sequential sensor measurements. However, long-term autonomous UAV operation requires globally consistent estimation and drift correction, which motivates the development of back-end optimization and fusion frameworks.

3 Back-End for UAV-Based SLAM

Back-end plays a critical role in improving the global consistency and long-term accuracy of SLAM systems. While front-end odometry provides local motion estimates. In the following, the SLAM back-end is reviewed from several perspectives, including *pose graph optimization* (PGO), 3D reconstruction, filter-based SLAM, fusion-based multi-UAV SLAM, and blockchain-enabled UAV-based SLAM methods.

3.1 Pose Graph Optimization (PGO)

PGO estimates a set of UAV poses represented by $(R_i, t_i)_{i=1}^N$, where $R_i \in SO(3)$ denotes the attitude of the i-th pose and $t_i \in \mathbb{R}^3$ denotes its position. The pose graph is defined as $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where each node $i \in \mathcal{V}$ corresponds to a UAV pose and each edge $(i, j) \in \mathcal{E}$ represents a relative pose constraint obtained from odometry or loop closure measurements. Let $\tilde{R}_{ij} \in SO(3)$ and $\tilde{t}_{ij} \in \mathbb{R}^3$ denote the measured relative rotation and translation from pose i to pose j, respectively [20]. Mathematically, PGO can be formulated as a *maximum likelihood estimation* (MLE) problem over the set of poses [68] as follows:

Problem 1: *Given a pose graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and relative pose measurements $(\tilde{R}_{ij}, \tilde{t}_{ij})_{(i,j)\in\mathcal{E}}$, the objective of PGO is to estimate the set of poses $(R_i, t_i)_{i=1}^N$ that best satisfies these relative constraints under measurement noise. A typical formulation is given by:*

$$(\boldsymbol{R}_i^*, \boldsymbol{t}_i^*)_{i=1}^N = \arg \min_{\substack{\boldsymbol{t}_i \in \mathbb{R}^d \\ \boldsymbol{R}_i \in \mathbb{SO}(d)}} \sum_{(i,j)\in\mathcal{E}} \left(\omega_{ij}^t \left\| \boldsymbol{R}_i^\top (\boldsymbol{t}_j - \boldsymbol{t}_i) - \tilde{\boldsymbol{t}}_{ij} \right\|_2^2 + \frac{\omega_{ij}^R}{2} \left\| \boldsymbol{R}_i^\top \boldsymbol{R}_j - \tilde{\boldsymbol{R}}_{ij} \right\|_F^2 \right), \tag{1}$$

where the weights ω_{ij}^t and ω_{ij}^R reflect the confidence of the translation and rotation constraints, respectively.

As shown in Problem 1, due to the non-convexity of the rotation group $\mathbb{SO}(d)$ [69], the PGO solution of Problem 1 is a challenging non-convex problem. As a consequence, most existing approaches rely on various relaxation strategies, including iterative SLAM optimization methods [70,71]. Although this type of methods are efficient and widely adopted, they do not guarantee global convergence to the MLE solution. To address this limitation, several works have explored global optimality verification and convex relaxation techniques. For instance, reference [70] proposed a verification technique to assess whether a given 2D SLAM solution is globally optimal. However, the back-end optimization in such systems is typically solved via iterative nonlinear methods, which exhibit strong empirical performance but provide no guarantees on solution quality. To bridge this gap, reference [72] showed that Lagrangian duality can be used to certify candidate PGO solutions by computing tight lower bounds through convex programming, even when obtaining the global optimum is intractable.

Another major challenge in PGO problems is robustness to outliers and mismodeled measurement noise [20]. In particular, incorrect loop-closure constraints can cause catastrophic failures in least-square-based SLAM. To mitigate this problem, reference [73] proposed a robust SLAM framework that jointly estimates poses and measurement information matrices during optimization. Similarly, reference [74] introduced an *expectation maximization* (EM) based robust pose-graph optimization method that detects and reduces the influence of erroneous loop closures. Robust PGO methods have also been extended to multi-robot scenarios. For example, reference [33] applied the EM framework to the localization of the multi-UAV pose graph and the association of data without requiring prior knowledge of the initial relative poses. Additional sensing information can further improve robustness. In [75], planar markers were used to assist PGO and enhance the reliability of the systems. More recently, several works have focused on improving the robustness and efficiency of PGO in large-scale SLAM systems. For instance, reference [28] improved the robustness of PGO for SLAM by detecting and suppressing outlier loop-closure constraints in the front-end. Building on this, reference [27] proposed a weighted multi-node joint refinement strategy that further enhances robustness and accelerates back-end convergence. Furthermore, reference [71] developed a quasi-Newton-based iteratively reweighted least squares method operating in the

Lie algebra of the 3D rotation group, enabling robust and efficient relative rotation averaging for large-scale structure-from-motion applications.

3.2 3D Reconstruction

UAV-based SLAM technologies are often employed for 3D reconstruction of the surrounding environment. In classical computer vision literature, 3D reconstruction is generally understood as a recovery of the three-dimensional structure of an object from a collection of multiple views [29]. To this end, multiple UAVs first capture multi-view images from different viewpoints during flight to provide raw visual observations of the scene, which are then processed in the front-end matching stage through feature extraction, feature matching, and outlier rejection to establish reliable correspondences across views, and subsequently fed into the back-end optimization stage, where camera poses are estimated and scene geometry is reconstructed via triangulation to recover a sparse 3D point cloud, which is further refined through bundle adjustment and PGO to improve global consistency, and finally used to generate a more complete 3D representation of the environment, enabling dense or semi-dense reconstruction for downstream applications.

From a mathematical perspective [76,77], 3D reconstruction in SLAM relies on multi-view geometry and pose consistency across observations. In this settings, the UAV pose defines the transformation between the body frame $\{\mathcal{B}\}$ and the reference frame $\{\mathcal{I}\}$, while landmarks are represented in a global coordinate system. The platform attitude can be recovered from vector measurements in both frames, where the rotation matrix describes the nonlinear transformation between coordinate systems. However, ensuring globally consistent pose estimation from multiple observations remains a challenging problem due to noise accumulation and inconsistent pairwise measurements. To address this, reference [78] studied rigid-motion synchronization (motion averaging) in the special Euclidean group $\mathbb{SE}(3)$, which is essential for recovering globally consistent camera poses from multiple views. Meanwhile, to address the challenge of UAV visual reconstruction under low-light conditions, reference [79] developed a novel feature detector that operates directly on image bursts, thereby enhancing reconstruction performance in extremely low-light environments.

Based on accurate pose estimation and multi-view observations, UAV-based SLAM technologies can reconstruct detailed 3D structures of complex environments. Such techniques have been widely applied in infrastructure inspection [43], terrain mapping [80] and cultural heritage preservation [30]. For example, reference [43] demonstrated that dense and high-fidelity 3D reconstructions generated from image data can support precise condition assessment of large-scale infrastructure. Meanwhile, reference [80] showed that integrating 3D reconstruction with terrain matching enables reliable pose estimation and environment mapping for aerial platforms, even in GPS-denied scenarios. In particular, to address limited accessibility in the digital documentation of large indoor historical structures, a cooperative UAV system for autonomous exploration and documentation of the interior of buildings is proposed [30].

3.3 Filter-Based SLAM

As discussed in Section 2, the UAV-based SLAM systems typically rely on the fusion of heterogeneous sensors. While the front-end focuses on sensor configuration and measurement generation, the back-end is responsible for fusing these measurements within a unified estimation framework to improve accuracy and robustness. In this context, filter-based SLAM has served as a representative paradigm for single-UAV multi-sensor SLAM, aiming to combine complementary observations for reliable state estimation. Among these, *extended Kalman filter* (EKF) type methods including *multi-state constraint Kalman filter* (MSCKF) [31] have been widely adopted, which can achieve accurate and consistent visual-inertial state estimation by maintaining a sliding window of camera poses [21]. To better capture nonlinear system dynamics, *unscented*

Kalman filter (UKF) variants have also been explored. For instance, reference [3] proposed a quaternion-based UKF for GPS-denied navigation, while reference [55] further incorporated deep learning to enhance estimation performance in complex environments.

Instead of purely probabilistic filtering formulations, recent research has explored geometry-aware and symmetry-preserving estimation frameworks for UAV-based visual SLAM. These approaches typically rely on intrinsic system representations on Lie groups and exploit equivariance properties to design observers with improved stability and consistency guarantees. For instance, reference [81] proposed a lifted observer design by embedding the SLAM system into a higher-dimensional Lie group, where the system output is equivariant. This formulation enables the construction of a nonlinear observer with almost semi-global asymptotic stability by defining estimation errors in intrinsic coordinates, thereby avoiding reliance on local linearization. Building upon this idea, references [76,77] further developed symmetry-preserving observer designs that relax strong observability conditions and achieve almost global convergence. In addition, references [82,83] introduced geometric stochastic filtering approaches for visual SLAM, further enriching the theoretical foundation of state estimation. More recently, such geometry-based estimation frameworks have been extended beyond pure state estimation toward integrated estimation and control. For example, reference [32] proposed a hybrid attitude observer on $\mathbb{SO}(3)$ for multirate sensor fusion, and reference [84] further incorporated this framework into closed-loop task-space control for robotic manipulation based on visual SLAM.

3.4 Fusion-Based Multi-UAV SLAM

A key capability in multi-UAV autonomous systems is collaborative localization and mapping in challenging, partially unknown environments. By fusing information between among UAVs, the performance of individuals in the group can be significantly improved, allowing for cooperatively performing complicated tasks in different domains including surveillance, search and rescue, and object manipulation [33,85]. However, in the context of data sharing among UAVs for multi-UAV association, maintaining estimation consistency becomes a critical challenge, primarily due to the presence of unknown cross-correlations among local estimates. To address this issue, *covariance intersection* (CI)-based sensor fusion methods [23,86] have been widely used to address this problem. For example, reference [87] first introduced CI into SLAM within a *split covariance intersection* (SCI)-SLAM framework to avoid maintaining full cross-correlations between vehicle and landmark estimates. The method reduces computational complexity while ensuring consistency under unknown correlations. Building on this idea, reference [19] proposed an observability Gramian-based CI-SLAM framework, which performs offline optimizing fusion weights using observability metrics, thereby less computing burden. Furthermore, when visual measurements become unreliable, SCI can be employed to handle unknown correlations between model-based estimation and learning-based predictions [88], where the learning component adopts a ResNet-18 architecture to predict the position and velocity dynamics.

Recently, multi-UAV SLAM systems have increasingly evolved toward distributed fusion frameworks, in which distributed SLAM eliminates the need for a central coordinator, enabling each UAV to perform perception, estimation, and planning locally, while exchanging compact information (e.g., poses or submaps) with neighboring agents. For instance, Swarm-LIO [56] achieved fully distributed LiDAR-inertial odometry through local ego-state estimation and inter-UAV exchange of mutual observations and state information, whereas Swarm-LIO2 [16] further improves plug-and-play capability, scalability, and bandwidth efficiency for large-scale aerial swarms. From a sensor fusion perspective, reference [89] proposed a decentralized vehicle localization architecture based on SCI, where each agent propagates its state using local motion measurements, and fuses relative observations and shared state information from neighboring vehicles. Building upon the Rao-Blackwellized *labeled multi-Bernoulli* (LMB) [90] framework, reference [91] extended this paradigm to distributed multi-robot SLAM. Each agent performs local estimation while exchanging probabilistic map representations with neighbors, enabling reliable data association and consistent global mapping under communication constraints.

3.5 Blockchain-Enabled UAV-Based SLAM Methods

In distributed UAV-based SLAM scenarios, multi-UAVs simultaneously explore the environment and exchange mapping information through wireless communication. As a result, issues such as data sharing, distributed mapping, and system security have become increasingly critical. From a security perspective [45], frequent data transmission over wireless channels exposes the system to threats such as eavesdropping, malware injection, and replay attacks. Consequently, ensuring the security and trustworthiness of shared data has become an urgent challenge in multi-UAV SLAM systems. Fortunately, *blockchain technology* [35] provides a decentralized and tamper-resistant framework for secure map sharing and data integrity of distributed multi-UAV SLAM by periodic legality verification and blocks hash registry. A typical blockchain-enabled architecture for collaborative UAV-based SLAM is illustrated in Fig. 4 [12], which consists of the UAV end, the blockchain network, and the server end. At the UAV end, multiple UAVs collect sensor data and generate local maps, which are transmitted to the blockchain network for secure storage and verification through hashing and timestamping. The validated data are then forwarded to the server end, where local maps are optimized and fused into a globally consistent representation.

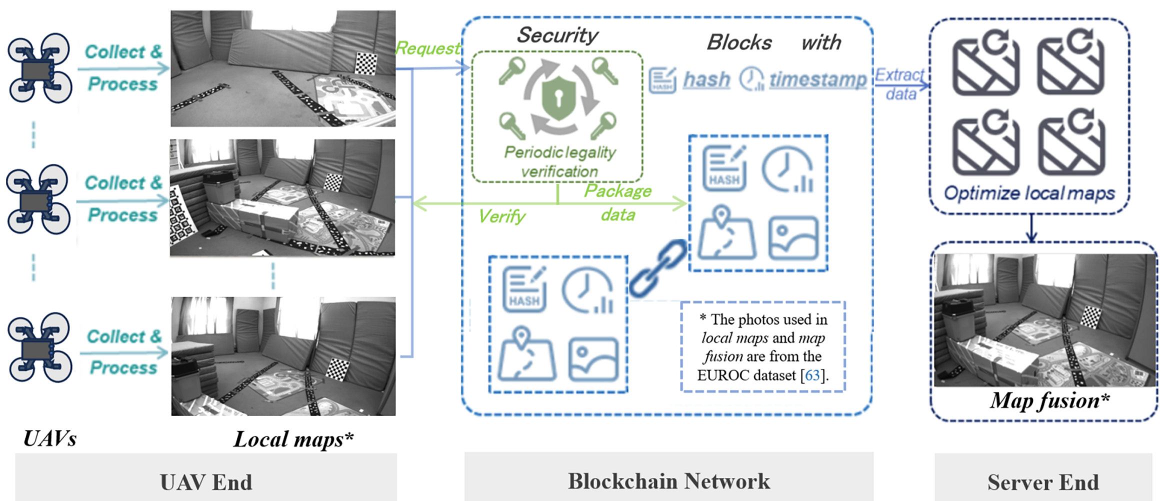


Figure 4: Blockchain-enabled architecture between the UAV end and server end [12].

Several recent works presented an initial investigation into blockchain-based collaborative mapping for multi-UAV systems. For example, references [35,92] explored blockchain-enabled frameworks for multi-UAV systems, enabling secure and decentralized data sharing through distributed ledgers and smart contracts, while improving coordination, scalability, and fault tolerance in large-scale applications. Reference [93] further provided a systematic guideline for fully autonomous multi-UAV teamwork using layer-2 blockchain, emphasizing secure data logging, trust management, and privacy-preserving collaboration for reliable multi-UAV operations. More recently, a blockchain-based multi-UAV collaborative mapping frame is initially explored in [54]. However, its encryption-heavy design posed computational bottlenecks and increased node failure risks in resource constrained UAVs. To overcome these limitations, reference [12] proposed a blockchain-based data security mechanism for multi-UAV visual SLAM systems, which shifts symmetric encryption toward lightweight identity authentication and integrates distributed map data storage based on a one-way hash function with multi-source map fusion, providing a secure, efficient, and SLAM-compatible solution for dynamic multi-UAV environments.

When blockchain mechanisms are introduced into UAV-SLAM, their limitations become more critical because SLAM-related information, such as keyframes, local maps, loop-closure constraints, and state estimates, must be exchanged and updated in a timely manner [94]. Frequent block generation, consensus verification, cryptographic authentication, and ledger synchronization may increase communication latency, computational load, and energy consumption, thereby affecting the real-time feasibility of collaborative UAV-SLAM [95]. These observations suggest that blockchain should be integrated as a lightweight security layer for selected map records, identity authentication, and integrity verification, rather than being directly applied to the high-frequency SLAM estimation pipeline.

4 Applications

The back-end optimization techniques provide effective tools for deploying UAV-based SLAM systems in complex real-world scenarios, which drives the systems increasingly integrated into various domains. Several popular UAV-based SLAM frameworks (e.g., MSCKF, ORB-SLAM, Swarm-LIO and EGO-Swarm) have been introduced in the preceding sections. In general, these methods are first validated using public benchmark datasets or simulation environments, such as: 1) the EuRoC micro aerial vehicle datasets [63]; 2) the TUM visual-inertial benchmark [64]; 3) the AirSim simulation platform, which is widely used for developing and testing aerial AI algorithms [66]. Beyond those benchmark evaluations, validating algorithm performance in real-world scenarios is equally important. Accordingly, the following sections focus on several representative and innovative application scenarios to further demonstrate the popularity of the UAV-based SLAM systems.

4.1 UAV-Plus: UAV-Based Collaborative SLAM

In order to facilitate more extensive multi-machine cooperation, recent research has extended UAV-based SLAM systems to heterogeneous collaborative frameworks, where UAVs can operate jointly with other intelligent platforms, such as *UAV & unmanned underwater vehicle* (UAV-UUV), *UAV & unmanned ground vehicle* (UAV-UGV), *UAV-Manipulator* (UAV-M), and *UAV-Human* (UAV-H). In this survey, this kind of plarform is termed as UAV-plus system, which enables complementary sensing and motion capabilities, allowing complex tasks to be performed more efficiently in challenging environments.

4.1.1 UAV-UUV

The operation of a single type of unmanned system (e.g., UAVs, USVs, or UUVs) has often been insufficient to accomplish complex tasks. Therefore, cross-domain cooperation has become a natural solution for handling such challenges. For instance, with the assistance of *unmanned surface vehicles* (USVs), UAVs have been able to land on USVs equipped with onboard recharging platforms, thereby effectively extending their battery life and operational range [96]. Furthermore, to accomplish complex underwater tasks, reference [37] has proposed a cooperative UAV-USV-UUV framework for underwater target hunting by jointly optimizing the UAV's position, the UUV's trajectory, and their communication connectivity.

4.1.2 UAV-UGV

Efficient warehouse management relies on real-time inventory visibility, enabling accurate tracking of stock levels, locations, and movements [40,97]. UAV-UGV collaboration leverages the complementary capabilities of heterogeneous robotic systems, where UAVs are responsible for aerial perception and rapid inspection, while UGVs provide ground mobility, extended endurance, and system-level support. For example, in traditional UAV-UGV systems [97], wheeled UGVs act as mobile platforms that carry UAVs and determine deployment locations, allowing UAVs to perform localized inspections while compensating for their limited flight endurance. This cooperative strategy has effectively improved coverage efficiency under resource constraints. However, wheeled UGVs struggle to operate effectively in multi-level environments and complex terrains, such as stairs and confined spaces. To address these limitations, recent works [40] have extended this paradigm by introducing more versatile ground platforms, such as legged robots and robotic dogs. Therefore, this transition from wheeled to legged UGVs have significantly improved the applicability of UAV-UGV systems.

4.1.3 UAV-M

UAV-based aerial manipulation has imposed higher requirements on coordination among multiple agents. Inspired by human dexterity, recent research has explored collaborative and perception-aware aerial manipulation frameworks for grasping, maintenance, and aerial assembly tasks [98]. These platforms have been capable of performing versatile grasping actions, such as palm gripping and fingertip pinching, as well as executing complex operations including tree perching, door opening, object transportation, and human interaction. Such UAV-based manipulators (UAV-M) have demonstrated agile and precise operation in challenging scenarios such as disaster response and field applications [98]. Furthermore, successful outdoor manipulation of a floating target by a UAV-M has also been achieved in [39]. In this context, UAV-M systems have become a core component of a torch-relay platform during the Beijing 2022 Winter Olympics, where they have cooperated autonomously with UGVs in ice and snow environments.

4.1.4 UAV-H

UAV-based SLAM system can also be applied to human-centered assistance tasks. For instance, systems in which a human is physically (or wirelessly) connected to a UAV via a cable have been investigated in [38,99]. Such applications are well suited for visually or hearing-impaired users or people in low-visibility environments, overcoming the limitations of perception-dependent interaction from human. For example, in [38], an admittance-based control approach inspired by flexible manipulator theory was proposed to compute desired cable forces for guiding the human. However, this method relied solely on the robot state; as a result, the guiding force tended to decrease with increasing human walking speed, which could lead to unstable interaction and loose cable. To address this issue, reference [99] extended the framework by incorporating human-state awareness, introducing human velocity feedback into the control loop, resulting in more stable, continuous, and comfortable interaction in real-world assistive navigation scenarios.

4.2 Society Applications

UAV have been widely deployed across various societal applications, with particularly growing importance in agricultural monitoring and decision-making. Agricultural environments-particularly orchards-pose significant challenges for UAV-based SLAM system due to complex spatial structures, irregular obstacle distributions, and large operational areas. These factors make reliable navigation and obstacle avoidance essential. To address this, reference [8] extends conventional 2D path planning to realistic 3D scenarios using an improved ant colony optimization algorithm, incorporating terrain variation, obstacle distribution, and flight safety to enable safe and feasible UAV navigation. In addition, UAV-based robotic systems have been further explored for sustainable agricultural applications, especially water resource optimization. As noted in [41], traditional methods for assessing crop water status (e.g., stem water potential) are labor-intensive and difficult to scale. To overcome this, integrated UAV-robotic systems combining sensing, planning, and automation enable efficient data collection and analysis, thereby improving irrigation management and supporting sustainable agriculture.

Beyond agricultural applications, UAVs have also demonstrated substantial potential in civil engineering and public safety. For example, in infrastructure inspection, reference [42] proposed a lightweight attention-based convolutional network for UAV-based road damage detection, achieving a balance between accuracy and real-time performance with strong generalization capability. In disaster response, reference [44] developed a fixed-time cooperative control strategy for multi-UAV systems to monitor the dynamic spread of wildfires, enabling timely situational awareness and enhancing emergency response effectiveness. In addition, when communication infrastructures are destroyed by disasters, UAVs can be employed to perform immediate rescue missions in destroyed areas and assist data sharing for ground vehicles by a lightweight UAV-assisted vehicular blockchain framework [36]. Moreover, in response to the COVID-19 pandemic, reference [92] presented a blockchain-enabled multi-UAV collaboration framework for large-scale applications, such as door-to-door delivery, detecting and identifying infected cases, broadcasting news, and monitoring people who break the quarantine rules in quarantined areas.

To provide a clearer comparison of representative application scenarios, Table 4 summarizes UAV-based SLAM applications in terms of application domains, SLAM methods, sensors configuration, experimental platforms, benefits, and limitations. Overall, UAV-based SLAM system has been extended from conventional navigation and mapping toward heterogeneous collaboration in diverse domains, as summarized in Table 4.

Table 4: Summary of representative UAV-based SLAM application scenarios.

Application Domains	SLAM Methods	Sensors*	Platforms	Benefits	Limitations
UAV-plus	Deep Q-learning [37], tag-based [40], marker-based [98], path-following [99]	Sonar, LiDAR, depth camera, manipulator, force-cable sensor	UAV–USV–UUV, UAV-(Legged)UGV, UAV-manipulator, UAV-human	Cross-platform heterogeneous collaboration	Communication delay, lose-range perception occlusion, motion uncertainty and interaction latency
Agriculture	Learning-based [41]	Multispectral sensor	Orchards-UAV,	Low-altitude agility	Occlusion, and illumination changes
Civil	Learning-based [42]	RGB-D camera, LiDAR	UAV-based, inspection platform	High-resolution, structural assessment	Weak texture and lighting variation
Healthcare	Blockchain-enabled [92]	Thermal camera	Multi-UAV	Secured operation	Blockchain lifecycle management

Note: *: Common sensors such as traditional camera, IMU, and GNSS are used across four application domains.

5 Conclusion and Future Perspectives

With the rapid advancement of UAV technologies, UAV-based SLAM has become a fundamental enabling paradigm for autonomous navigation and environmental perception. This review has presented a systematic survey of recent developments in UAV-based SLAM from a system-level perspective, covering heterogeneous sensing and front-end odometry, back-end estimation and optimization, as well as emerging multi-agent and application-driven frameworks. More importantly, this survey reveals several new trends in UAV-based SLAM. At the system level, UAV-based SLAM is evolving toward tightly integrated multi-modal sensing and front-end odometry frameworks to improve robustness under diverse operating conditions. At the back-end level, estimation methods are increasingly focused on achieving global consistency, scalability, and distributed multi-agent cooperation. At the application level, SLAM is progressively transitioning from an isolated estimation module to a task-driven component that supports UAV-plus collaboration and real-world deployment in complex scenarios.

Despite substantial progress in recent years, UAV-based SLAM systems still face significant challenges when deployed in real-world, unstructured environments. Among these, estimation and control for increasingly complex and tightly coupled UAV platforms remain a well-recognized difficulty; the reader is referred to [46] for comprehensive surveys and detailed discussions. In the following, we instead highlight several critical yet underexplored challenges in the existing literature:

1) *Observability issues*. The UAV-based SLAM problem is fundamentally one of state estimation, and as such, is inherently subject to observability limitations. Prior work [100] has shown that standard 2D planar, world-centric SLAM formulations with odometry and range–bearing measurements are intrinsically unobservable in the absence of additional prior information, with observability attainable only when sufficient absolute feature references are available. Furthermore, reference [76] establishes a connection between uniform observability and persistency of excitation, highlighting the critical role of sufficiently rich motion and informative measurements in ensuring consistent state estimation. However, in practical UAV-based SLAM, these conditions are routinely violated by limited sensing, constrained motion, and environmental uncertainty, leading to fundamental observability loss. Future work should therefore augment observability using learned priors, structural or semantic constraints (such as object categories, geometric regularities, and scene-level semantic cues), multi-sensor fusion, and occasional external references to suppress unobservable modes and maintain consistent, drift-bounded estimation.
2) *Imperfect and dynamic environments*. Imperfect and dynamic environments remain a fundamental challenge for UAV-based SLAM systems. Unlike controlled settings, UAVs are typically deployed in highly dynamic and unstructured scenarios, where assumptions of static scenes and ideal sensing conditions are frequently violated [101]. In addition, UAV applications often involve challenging adverse weather conditions, such as fog, rain, strong wind, low-light or nighttime environments encountered in delivery, mining, and ecological monitoring-where noise limited imagery further degrades SLAM performance [44]. As a result, both environmental dynamics and sensing imperfections can significantly impair localization and reconstruction accuracy, particularly over lifelong mapping [102]. Therefore, it calls for more effective robust UAV-SLAM methods that take into consideration of adverse weather condition.
3) *Miniaturization*. UAV-based SLAM platforms are increasingly trending toward miniaturization (commonly referred to as *micro aerial vehicles* (MAVs)) to enable agile operation in confined or complex environments [5,103]. Recent efforts have explored lightweight perception, control, and navigation strategies tailored to resource-constrained MAV platforms [104]. However, for MAVs with severely limited payload capacity (typically ≤100 g), a fundamental limitation is the inability to accommodate sensors capable of providing high-resolution metric depth, such as LiDAR or stereo cameras [61], or batteries that support power demanding tasks, such as edge intelligence, federated learning, privacy-preserving collaboration. Consequently, future research should focus on resource-aware UAV-based SLAM frameworks that compensate for sensing and energy deficiencies using monocular cues,

learning-based depth priors, or tightly coupled sensor fusion, while ensuring robust, accurate, and real-time operation under strict onboard constraints.

Acknowledgement: The authors gratefully acknowledge the financial support provided by the Natural Science Foundation of Shanghai and National Natural Science Foundation of China.

Funding Statement: This work was sponsored by the Natural Science Foundation of Shanghai through Grant (No. 25ZR1401264) and National Natural Science Foundation of China through Grant (No. 62273239).

Author Contributions: Conceptualization: Wangyan Li; validation: Wangyan Li and Guoliang Wei; methodology and formal analysis: Wangyan Li and Yaolei Wang; investigation and resources: Wangyan Li and Guoliang Wei; writing—original draft: Yaolei Wang and Wangyan Li; writing—review and editing: Yaolei Wang, Wangyan Li and Guoliang Wei; supervision and project administration: Wangyan Li; funding acquisition: Wangyan Li and Guoliang Wei. All authors reviewed and approved the final version of the manuscript.

Availability of Data and Materials: Not applicable.

Ethics Approval: Not applicable.

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

References

1. Thrun S, Burgard W, Fox D. Probabilistic robotics. Cambridge, MA, USA: MIT Press; 2005.
2. Cadena C, Carlone L, Carrillo H, Latif Y, Scaramuzza D, Neira J, et al. Past, present, and future of simultaneous localization and mapping: toward the robust-perception age. IEEE Trans Robot. 2016;32(6):1309–32. doi:10.1109/tro.2016.2624754.
3. Ghanizadegan K, Hashim HA. Quaternion-based unscented Kalman filter for 6-DoF vision-based inertial navigation in GPS-denied regions. IEEE Trans Instrum Meas. 2025;74(9):8500713. doi:10.1109/tim.2024.3509582.
4. Fuentes-Pacheco J, Ruiz-Ascencio J, Rendón-Mancha JM. Visual simultaneous localization and mapping: a survey. Artif Intell Rev. 2012;43(1):55–81. doi:10.1007/s10462-012-9365-8.
5. Thumiger N, Deghat M. A multi-agent deep reinforcement learning approach for practical decentralized UAV collision avoidance. IEEE Control Syst Lett. 2022;6:2174–9. doi:10.1109/lcsys.2021.3138941.
6. Zhou X, Wang Z, Ye H, Xu C, Gao F. EGO-planner: an ESDF-free gradient-based local planner for quadrotors. IEEE Robot Autom Lett. 2021;6(2):478–85.
7. Zhou X, Zhu J, Zhou H, Xu C, Gao F. EGO-Swarm: a fully autonomous and decentralized quadrotor swarm system in cluttered environments. In: Proceedings of 2021 IEEE International Conference on Robotics and Automation (ICRA); 2021 May 30–Jun 5; Xi'an, China. Piscataway, NJ, USA: IEEE; 2021. p. 4101–7.
8. Shao Y, Yang H, Gao R, Li F. Three-dimensional obstacle avoidance path planning for agricultural UAV based on improved ant colony algorithm. Int J Netw Dyn Intell. 2025;4(4):100028. doi:10.53941/ijndi.2025.100028.
9. Mur-Artal R, Montiel JMM, Tardos JD. ORB-SLAM: a versatile and accurate monocular SLAM system. IEEE Trans Robot. 2015;31(5):1147–63.
10. Mur-Artal R, Tardos JD. ORB-SLAM2: an open-source SLAM system for monocular, stereo, and RGB-D cameras. IEEE Trans Robot. 2017;33(5):1255–62.
11. Campos C, Elvira R, Rodriguez JJG, Montiel JMM, Tardos JD. ORB-SLAM3: an accurate open-source library for visual, visual-inertial, and multimap SLAM. IEEE Trans Robot. 2021;37(6):1874–90. doi:10.1109/tro.2021.3075644.
12. Xiao W, Jiang S, Hong D, Jiang H, Chen Z, Li W. BBDS-vSLAM: multi-UAV visual SLAM system with blockchain-based data security. IEEE Embed Syst Lett. 2026. doi:10.1109/LES.2025.3640718.
13. Mellinger D, Kumar V. Minimum snap trajectory generation and control for quadrotors. In: 2011 IEEE International Conference on Robotics and Automation (ICRA); 2011 May 9–13; Shanghai, China. Piscataway, NJ, USA: IEEE; 2011. p. 2520–5.

14. Gu Y, Yu X, Guo K, Qiao J, Guo L. Detection, estimation, and compensation of false data injection attack for UAVs. Inf Sci. 2021;546(2):723–41. doi:10.1016/j.ins.2020.08.055.
15. Fu L, An L, Zhang L. Attitude-position obstacle avoidance of trajectory tracking control for a quadrotor UAV using barrier functions. Int J Syst Sci. 2024;55(16):3337–54. doi:10.1080/00207721.2024.2367719.
16. Zhu F, Ren Y, Yin L, Kong F, Liu Q, Xue R, et al. Swarm-LIO2: decentralized efficient LiDAR-inertial odometry for aerial swarm systems. IEEE Trans Robot. 2025;41:960–81.
17. Guan Q, Wei G, Wang Y, Liu Y. A dual-mode automatic switching feature points matching algorithm fusing IMU data. Measurement. 2021;185(8):110043. doi:10.1016/j.measurement.2021.110043.
18. Guan Q, Wei G, Wang L, Song Y. A novel feature points tracking algorithm in terms of IMU-aided information fusion. IEEE Trans Industr Inform. 2021;17(8):5304–13. doi:10.1109/tii.2020.3024079.
19. Hu Y, Li W, Wu L, Wei G. A novel observability gramian-based SLAM. In: Proceedings of 2023 42nd Chinese Control Conference (CCC);2023 Jul 24–26; Tianjin, China. Piscataway, NJ, USA: IEEE; 2023. p. 3421–6.
20. Carlone L, Calafiore GC. Convex relaxations for pose graph optimization with outliers. IEEE Robot Autom Lett. 2018;3(2):1160–7. doi:10.1109/lra.2018.2793352.
21. Sun K, Mohta K, Pfrommer B, Watterson M, Liu S, Mulgaonkar Y, et al. Robust stereo visual inertial odometry for fast autonomous flight. IEEE Robot Autom Lett. 2018;3(2):965–72. doi:10.1109/lra.2018.2793349.
22. Hu Y, Li W, Wu L, Yu Z. An attack-resilient distributed extended Kalman consensus filtering algorithm with applications to multi-UAV tracking problems. Front Inf Technol Electron Eng. 2024;25(8):1110–22. doi:10.1631/fitee.2300621.
23. Li W, Yang F. Information fusion over network dynamics with unknown correlations: an overview. Int J Netw Dyn Intell. 2023;2(2):100003. doi:10.53941/ijndi0201003.
24. Leutenegger S, Lynen S, Bosse M, Siegwart R, Furgale P. Keyframe-based visual-inertial odometry using nonlinear optimization. Int J Robot Res. 2014;34(3):314–34. doi:10.1177/0278364914554813.
25. Qin T, Li P, Shen S. VINS-Mono: a robust and versatile monocular visual-inertial state estimator. IEEE Trans Robot. 2018;34(4):1004–20.
26. Bondy JA, Murty USR. Graph theory. London, UK: Springer; 2008.
27. Cai J, Wei G, Hu H. Weighted multi-node joint refinement for VSLAM pose-graph robustification. Meas Sci Technol. 2025;36(5):056303. doi:10.1088/1361-6501/adcce2.
28. Cai J, Wei G, Li W, Wang Y. Robust pose graph optimization against outliers using consistency credibility factor. IEEE-CAA J Autom Sin. 2025;12(5):1044–6. doi:10.2139/ssrn.4272436.
29. Schönberger JL, Frahm JM. Structure-from-motion revisited. In: Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); 2016 Jun 27–30; Las Vegas, NV, USA. Piscataway, NJ, USA: IEEE; 2016. p. 4104–13.
30. Petracek P, Kratky V, Baca T, Petrlik M, Saska M. New era in cultural heritage preservation: cooperative aerial autonomy for fast digitalization of difficult-to-access interiors of historical monuments. IEEE Robot Autom Magaz. 2024;31(2):8–25.
31. Mourikis AI, Roumeliotis SI. A multi-state constraint Kalman filter for vision-aided inertial navigation. In: Proceedings 2007 IEEE International Conference on Robotics and Automation; 2007 Apr 10–14; Rome, Italy. Piscataway, NJ, USA: IEEE; 2007. p. 3565–72.
32. Wang M, Tayebi A. Nonlinear attitude estimation using intermittent and multirate vector measurements. IEEE Trans Autom Control. 2024;69(8):5231–45. doi:10.1109/tac.2023.3346799.
33. Indelman V, Nelson E, Michael N, Dellaert F. Multi-robot pose graph localization and data association from unknown initial relative poses via expectation maximization. In: Proceedings of 2014 IEEE International Conference on Robotics and Automation (ICRA); 2014 May 31–Jun 7; Hong Kong, China. Piscataway, NJ, USA: Hong Kong; 2014. p. 593–600.
34. Liu J, Chen K, Liu R, Yang Y, Wang Z, Zhang J. Robust and accurate multi-agent SLAM with efficient communication for amart mobiles. In: Proceedings of 2022 International Conference on Robotics and Automation (ICRA); 2022 May 23–27; Philadelphia, PA, USA. Piscataway, NJ, USA: IEEE; 2022. p. 2782–8.

35. Abichandani P, Lobo D, Kabrawala S, McIntyre W. Secure communication for multiquadrotor networks using ethereum blockchain. IEEE Internet Things J. 2021;8(3):1783–96. doi:10.1109/jiot.2020.3015716.
36. Su Z, Wang Y, Xu Q, Zhang N. LVBS: lightweight vehicular blockchain for secure data sharing in disaster rescue. IEEE Trans Depend Secur Comput. 2022;19(1):19–32.
37. Wei W, Wang J, Fang Z, Chen J, Ren Y, Dong Y. 3U: joint design of UAV-USV-UUV networks for cooperative target hunting. IEEE Trans Veh Technol. 2023;72(3):4085–90.
38. Tognon M, Alami R, Siciliano B. Physical human-robot interaction with a tethered aerial vehicle: application to a force-based human guiding problem. IEEE Trans Robot. 2021;37(3):723–34.
39. Zhang G, He Y, Yang L, Huang C, Chang Y, Li S. An aerial manipulator for robot-to-robot torch-relay task: system design and control scheme. IEEE Robot Autom Magaz. 2024;31(2):54–65.
40. Dibattista MA, Frericks J, Garcia CI, Camelio JA. SCOUT: an autonomous UHF RFID-equipped robot dog for flexible inventory monitoring. Manuf Lett. 2025;44(S):1525–32.
41. Dechemi A, Chatziparaschis D, Chen J, Campbell M, Shamshirgaran A, Mucchiani C, et al. Robotic assessment of a crop's need for watering: automating a time-consuming task to support sustainable agriculture. IEEE Robot Autom Magaz. 2023;30(4):52–67.
42. Chen L, Wu P, Tan W, Li H, Chen H, Zeng N. A novel UAV-based road damage detection algorithm with lightweight convolution and attention mechanism. Int J Netw Dyn Intell. 2025;4(4):100025. doi:10.53941/ijndi.2025.100025.
43. Khaloo A, Lattanzi D. Hierarchical dense structure-from-motion reconstructions for infrastructure condition assessment. J Comput Civ Eng. 2017;31(1):04016047. doi:10.1061/(asce)cp.1943-5487.0000616.
44. Mawanza JT, Manganda M, Ijiga OE. A novel fixed-time control scheme for multi-UAV wildfire tracking with dynamic reconfiguration and fault tolerance. Syst Sci Control Eng. 2025;13(1):04016047. doi:10.1080/21642583.2025.2587840.
45. Bai Y, Zhao H, Zhang X, Chang Z, Jantti R, Yang K. Toward autonomous multi-UAV wireless network: a survey of reinforcement learning-based approaches. IEEE Commun Surv Tutor. 2023;25(4):3038–67.
46. Idrissi M, Salami M, Annaz F. A review of quadrotor unmanned aerial vehicles: applications, architectural design and control algorithms. J Intell Robot Syst. 2022;104(2):22.
47. Pu H, Luo J, Wang G, Huang T, Liu H, Luo J. Visual SLAM integration with semantic segmentation and deep learning: a review. IEEE Sens J. 2023;23(19):22119–38. doi:10.1109/jsen.2023.3306371.
48. Endres F, Hess J, Engelhard N, Sturm J, Cremers D, Burgard W. An evaluation of the RGB-D SLAM system. In: Proceedings of 2012 IEEE International Conference on Robotics and Automation; 2012 May 14–18; Saint Paul, MN, USA. Piscataway, NJ, USA: IEEE; 2012. p. 1691–6.
49. Zhang J, Singh S. LOAM: lidar odometry and mapping in real-time. Robot: Sci Syst. 2014;2(9):1–9. doi:10.15607/rss.2014.x.007.
50. Pire T, Fischer T, Castro G, De Cristóforis P, Civera J, Berlles JJ. S-PTAM: stereo parallel tracking and mapping. Robot Auton Syst. 2017;93:27–42.
51. Shan T, Englot B, Meyers D, Wang W, Ratti C, Rus D. LIO-SAM: lightly-coupled lidar inertial odometry via smoothing and mapping. In: Proceedings of 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2020 Oct 25–29; Las Vegas, NV, USA. Piscataway, NJ, USA: IEEE; 2020. p. 5135–42.
52. Shan T, Englot B, Ratti C, Rus D. LVI-SAM: tightly-coupled LiDAR-visual-inertial odometry via smoothing and mapping. In: Proceedings of 2021 IEEE International Conference on Robotics and Automation (ICRA); 2021 May 30–Jun 5; Xi'an, China. Piscataway, NJ, USA: IEEE; 2021. p. 5692–8.
53. Xu W, Cai Y, He D, Lin J, Zhang F. FAST-LIO2: fast direct LiDAR-inertial odometry. IEEE Trans Robot. 2022;38(4):2053–73.
54. Xiao W, Hong D, Jiang S, Li W. MU-BISE: multi-UAV collaborative mapping based on blockchain information security enhancement. In: Proceedings of 2024 IEEE 3rd International Conference on Robotics, Artificial Intelligence and Intelligent Control (RAIIC); 2024 Jul 5–7; Mianyang, China. Piscataway, NJ, USA: IEEE; 2024. p. 241–4.
55. Ghanizadegan K, Hashim HA. DeepUKF-VIN: adaptively-tuned deep unscented Kalman filter for 3D visual-inertial navigation based on IMU-vision-net. Expert Syst Appl. 2025;271:126656.

56. Zhu F, Ren Y, Kong F, Wu H, Liang S, Chen N, et al. Swarm-LIO: decentralized swarm LiDAR-inertial odometry. In: Proceedings of 2023 IEEE International Conference on Robotics and Automation (ICRA); 2023 May 29–Jun 2; London, UK. Piscataway, NJ, USA: IEEE; 2023. p. 3254–60.
57. Kuutti S, Fallah S, Katsaros K, Dianati M, Mccullough F, Mouzakitis A. A survey of the state-of-the-art localization techniques and their potentials for autonomous vehicle applications. IEEE Internet Things J. 2018;5(2):829–46. doi:10.1109/jiot.2018.2812300.
58. Schmuck P, Ziegler T, Karrer M, Perraudin J, Chli M. COVINS: visual-inertial SLAM for centralized collaboration. In: Proceedings of 2021 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct); 2021 Oct 4–8; Bari, Italy. Piscataway, NJ, USA: IEEE; 2021. p. 171–6.
59. Ghosh S, Gallego G. Event-based stereo depth estimation: a survey. IEEE Trans Pattern Anal Mach Intell. 2025;47(10):9130–49. doi:10.1109/tpami.2025.3586559.
60. Hong Z, Petillot Y, Wallace A, Wang S. RadarSLAM: a robust simultaneous localization and mapping system for all weather conditions. Int J Robot Res. 2022;41(5):519–42.
61. Simon N, Majumdar A. MonoNav: MAV navigation via monocular depth estimation and reconstruction. In: Proceedings of 2024 Conference on Experimental Robotics; 2024 Jun 18–21; Venice, Italy. Cham, Switzerland: Springer Nature; 2024. p. 415–26.
62. Wang R, Schworer M, Cremers D. Stereo DSO: large-scale direct sparse visual odometry with stereo cameras. In: Proceedings of the IEEE International Conference on Computer Vision; 2017 Oct 22–29; Venice, Italy. Piscataway, NJ, USA: IEEE; 2017. p. 3903–11.
63. Burri M, Nikolic J, Gohl P, Schneider T, Rehder J, Omari S, et al. The EuRoC micro aerial vehicle datasets. Int J Robot Res. 2016;35(10):1157–63. doi:10.1177/0278364915620033.
64. Schubert D, Goll T, Demmel N, Usenko V, Stueckler J, Cremers D. The TUM VI benchmark for evaluating visual-inertial odometry. In: Proceedings of 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2018 Oct 1–5; Madrid, Spain. Piscataway, NJ, USA: IEEE; 2018. p. 1680–7.
65. Geiger A, Lenz P, Urtasun R. Are we ready for autonomous driving? The KITTI vision benchmark suite. In: Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition; 2012 Jun 16–21; Providence, RI, USA. Piscataway, NJ, USA: IEEE; 2012. p. 3354–61.
66. Shah S, Dey D, Lovett C, Kapoor A. Airsim: high-fidelity visual and physical simulation for autonomous vehicles. In: Field and Service Robotics: Results of the 11th International Conference. Cham, Switzerland: Springer; 2017. p. 621–35.
67. Helmberger M, Morin K, Berner B, Kumar N, Cioffi G, Scaramuzza D. The Hilti SLAM challenge dataset. IEEE Robot Autom Lett. 2022;7(3):7518–25. doi:10.1109/lra.2022.3183759.
68. Rosen DM, Carlone L, Bandeira AS, Leonard JJ. A certifiably correct algorithm for synchronization over the special euclidean group. In: Algorithmic foundations of robotics XII. Cham, Switzerland: Springer; 2020. p. 64–79.
69. Lajoie PY, Hu S, Beltrame G, Carlone L. Modeling perceptual aliasing in SLAM via discrete–continuous graphical models. IEEE Robot Autom Lett. 2019;4(2):1232–9. doi:10.1109/lra.2019.2894852.
70. Carlone L, Dellaert F. Duality-based verification techniques for 2D SLAM. In: Proceedings of 2015 IEEE International Conference on Robotics and Automation (ICRA); 2015 May 26–30; Seattle, WA, USA. Piscataway, NJ, USA: IEEE; 2015. p. 4589–96.
71. Chatterjee A, Govindu VM. Robust relative rotation averaging. IEEE Trans Pattern Anal Mach Intell. 2018;40(4):958–72. doi:10.1109/tpami.2017.2693984.
72. Carlone L, Rosen DM, Calafiore G, Leonard JJ, Dellaert F. Lagrangian duality in 3D SLAM: verification techniques and optimal solutions. In: Proceedings of 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2015 Sep 28–Oct 2; Hamburg, Germany. Piscataway, NJ, USA: IEEE; 2015. p. 125–32.
73. Graham MC, How JP. Robust simultaneous localization and mapping via information matrix estimation. In: Proceedings of 2014 IEEE/ION Position, Location and Navigation Symposium (PLANS); 2014 May 5–8; Monterey, CA, USA. Piscataway, NJ, USA: IEEE; 2014. p. 937–44.

74. Lee GH, Fraundorfer F, Pollefeys M. Robust pose-graph loop-closures with expectation-maximization. In: Proceedings of 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems; 2013 Nov 3–7; Tokyo, Japan. Piscataway, NJ, USA: IEEE; 2013. p. 556–63.
75. Muñoz-Salinas R, Marín-Jimenez MJ, Yeguas-Bolivar E, Medina-Carnicer R. Mapping and localization from planar markers. Pattern Recognit. 2018;73(3):158–71. doi:10.1016/j.patcog.2017.08.010.
76. Yi B, Jin C, Wang L, Shi G, Manchester IR. An almost globally convergent observer for visual SLAM without persistent excitation. In: Proceedings of 2021 60th IEEE Conference on Decision and Control (CDC); 2021 Dec 13–15; Austin, TX, USA. Piscataway, NJ, USA: IEEE; 2021. p. 5441–6.
77. Yi B, Jin C, Manchester IR. Globally convergent visual-feature range estimation with biased inertial measurements. Automatica. 2022;146(7):110639. doi:10.1016/j.automatica.2022.110639.
78. Arrigoni F, Rossi B, Fusiello A. Spectral synchronization of multiple views in SE(3). SIAM J Imag Sci. 2016;9(4):1963–90. doi:10.1137/16m1060248.
79. Ravendran A, Bryson M, Dansereau DG. BuFF: burst feature finder for light-constrained 3D reconstruction. IEEE Robot Autom Lett. 2023;8(12):8438–45.
80. Zhu Z, Su A, Liu H, Shang Y, Yu Q. Vision navigation for aircrafts based on 3D reconstruction from real-time image sequences. Sci China Technol Sci. 2015;58(7):1196–208. doi:10.1007/s11431-015-5828-x.
81. van Goor P, Mahony R, Hamel T, Trumpf J. Constructive observer design for visual simultaneous localisation and mapping. Automatica. 2021;132(11):109803. doi:10.1016/j.automatica.2021.109803.
82. Hashim HA. A geometric nonlinear stochastic filter for simultaneous localization and mapping. Aerosp Sci Technol. 2021;111(5):106569. doi:10.1016/j.ast.2021.106569.
83. Hashim HA, Abouheaf M, Abido MA. Geometric stochastic filter with guaranteed performance for autonomous navigation based on IMU and feature sensor fusion. Control Eng Pract. 2021;116(4):104926. doi:10.1016/j.conengprac.2021.104926.
84. Hashemi SH, Mattila J. Hybrid controller for robot manipulators in task-space with visual-inertial feedback. IEEE Trans Control Syst Technol. 2023;34(2):737–46. doi:10.1109/tcst.2025.3631985.
85. Bernreiter L, Khattak S, Ott L, Siegwart R, Hutter M, Cadena C. Collaborative robot mapping using spectral graph analysis. In: Proceedings of 2022 International Conference on Robotics and Automation (ICRA); 2022 May 23–27; Philadelphia, PA, USA. Piscataway, NJ, USA: IEEE; 2022. p. 3662–8.
86. Julier SJ, Uhlmann JK. A non-divergent estimation algorithm in the presence of unknown correlations. In: Proceedings of the 1997 American Control Conference (Cat. No. 97CH36041); 1997 Jun 6; Albuquerque, NM, USA. Piscataway, NJ, USA: IEEE; 1997. p. 2369–73.
87. Julier SJ, Uhlmann JK. Using covariance intersection for SLAM. Robot Auton Syst. 2007;55(1):3–20. doi:10.1016/j.robot.2006.06.011.
88. Jin Y, Zhang WA, Sun H, Yu L. Learning-aided inertial odometry with nonlinear state estimator on manifold. IEEE Trans Intell Transp Syst. 2023;24(9):9792–803. doi:10.1109/tits.2023.3273391.
89. Li H, Nashashibi F, Yang M. Split covariance intersection filter: theory and its application to vehicle localization. IEEE Trans Intell Transp Syst. 2013;14(4):1860–71.
90. Deusch H, Reuter S, Dietmayer K. The labeled multi-Bernoulli SLAM filter. IEEE Signal Process Lett. 2015;22(10):1561–5. doi:10.1109/lsp.2015.2414274.
91. Gao L, Battistelli G, Chisci L. Efficient distributed multi-robot SLAM under missed detections and clutter with labeled multi-Bernoulli map. Robot Auton Syst. 2026;198(2):105358. doi:10.1016/j.robot.2026.105358.
92. Alsamhi SH, Lee B, Guizani M, Kumar N, Qiao Y, Liu X. Blockchain for decentralized multi-drone to combat COVID-19 and future pandemics: framework and proposed solutions. Trans Emerg Telecommun Technol. 2021;32(9):1–19. doi:10.1002/ett.4255.
93. Buzcu B, Ozgun M, Gürcan O. Fully autonomous trustworthy unmanned aerial vehicleteamwork: a research guideline using level 2 blockchain. IEEE Robot Autom Magaz. 2024;31(2):78–88.
94. Mehta P, Gupta R, Tanwar S. Blockchain envisioned UAV networks: challenges, solutions, and comparisons. Comput Commun. 2020;151:518–38.

95. Hafeez S, Khan AR, Al-Quraan MM, Mohjazi L, Zoha A, Imran MA, et al. Blockchain-assisted UAV communication systems: a comprehensive survey. IEEE Open J Veh Technol. 2023;4:558–80.
96. Gupta PM, Pairet E, Nascimento T, Saska M. Landing a UAV in harsh winds and turbulent open waters. IEEE Robot Autom Lett. 2023;8(2):744–51. doi:10.1109/lra.2022.3231831.
97. Khazetdinov A, Aleksandrov A, Zakiev A, Magid E. RFID-based warehouse management system prototyping using a heterogeneous team of robots. In: Proceedings of 23rd International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines; 2020 Sep 8–10; Moscow, Russia. Singapore: World Scientific; 2020. p. 263–70.
98. Wu Y, Yang F, Jin R, Zhong Y, Wang J, Wu X, et al. Hand-like autonomous flying robot for airborne grasping and interaction. Nat Commun. 2026;17(1):2200. doi:10.1038/s41467-026-68967-3.
99. Allenspach M, Vyas Y, Rubio M, Siegwart R, Tognon M. Human-state-aware controller for a tethered aerial robot guiding a human by physical interaction. IEEE Robot Autom Lett. 2022;7(2):2827–34. doi:10.1109/lra.2022.3143574.
100. Lee KW, Wijesoma WS, Guzman JI. On the observability and observability analysis of SLAM. In: Proceedings of 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems; 2006 Oct 9–15; Beijing, China. Piscataway, NJ, USA: IEEE; 2006. p. 3569–74.
101. Bescos B, Fácil JM, Civera J, Neira J. DynaSLAM: tracking, mapping, and inpainting in dynamic scenes. IEEE Robot Autom Lett. 2018;3(4):4076–83.
102. Deng T, Xie H, Wang J, Chen W. Long-term visual simultaneous localization and mapping: using a Bayesian persistence filter-based global map prediction. IEEE Robot Autom Magaz. 2023;30(1):36–49.
103. Dai B, Meng G. Tracking control and disturbance rejection of micro quadrotor via adaptive internal model. Syst Sci Control Eng. 2024;12(1):2313322. doi:10.1080/21642583.2024.2313322.
104. Sui D, Deghat M, Sun Z, Eskandari M. Adaptive bearing-only target localization and circumnavigation under unknown wind disturbance: theory and experiments. IEEE Robot Autom Lett. 2024;9(12):11321–8. doi:10.1109/lra.2024.3487483.