



# Lightweight and accurate infrared rangefinder-fused LiDAR-inertial localization for UAV-based bridge inspection in GPS-denied environments

Anna Kong<sup>✉</sup>, Yulu Wang<sup>✉</sup>, Jiankang Zhao<sup>✉\*</sup>

Department of Instrument Science and Engineering, Shanghai Jiao Tong University, Shanghai, 200240, China

## ARTICLE INFO

### Keywords:

Localization  
Unmanned Aerial Vehicles (UAVs)  
LiDAR-inertial odometry  
Bridge inspection  
GPS-denied environments  
Error-State Kalman Filter  
Sensor fusion

## ABSTRACT

Effective Structural Health Monitoring (SHM) relies on high-precision and continuous data acquisition to support bridge damage detection, condition assessment, and maintenance decision-making. Unmanned aerial vehicles (UAVs) provide an efficient and flexible means for such data collection, but their reliability is constrained by localization accuracy. In practical bridge scenarios, GPS signals are often unavailable or severely degraded due to structural occlusion or electromagnetic interference, leading to significant localization errors. To address this challenge, this paper presents a lightweight and accurate infrared rangefinder-fused LiDAR-inertial localization (IRF-LIL) system designed for GPS-denied environments. The system incorporates residual modeling based on statistical features and grid-based map management into LiDAR-inertial odometry (LIO) to achieve efficient pose estimation, thereby reducing computational complexity and enabling real-time performance on resource-constrained platforms. An infrared-based altitude estimation module is further integrated with LIO through an Error-State Kalman Filter (ESKF) to suppress vertical drift and enhance altitude stability. Experiments conducted in a real-world bridge environment demonstrate that the proposed IRF-LIL system achieves high-precision localization, with an absolute position error (APE) of 0.17 m, outperforming two state-of-the-art LiDAR-inertial SLAM algorithms.

## 1. Introduction

Bridges are important components of civil infrastructure system, whose structural functionality is affected by deterioration as time progress without correct maintenance [1]. Thus, it is necessary to monitor and estimate the condition of bridge termly to extend their lifecycle. Structural Health Monitoring (SHM) is not only an effective approach for identifying potential structural damage and ensuring the safety and integrity of bridge networks [2], but has also become a critical strategy for guaranteeing safety, durability, and service performance throughout the entire life cycle of bridges [3]. SHM typically involves long-term monitoring of structural responses, such as strain, displacement, and vibration, as well as the detection of cracks, corrosion, and material degradation, thereby enabling performance evaluation, damage identification, and maintenance decision support [4]. However, conventional SHM methods often rely on distributed sensor networks and manual inspection, which are resource-intensive and challenging to achieve comprehensive coverage for large-span bridge structures [5].

In recent years, UAVs have gradually become an important tool for SHM due to their flexibility, efficiency, and ability to access hard-to-reach areas. By carrying high-resolution cameras, LiDAR, and infrared sensors, UAVs can rapidly acquire multi-scale visual and geometric data

of bridges, thereby supporting defect detection and three-dimensional structural modeling [6]. Compared with traditional approaches that rely on manual inspections and fixed sensor deployments, UAVs not only significantly reduce labor and time costs but also provide more comprehensive and high-resolution multimodal monitoring information [7]. Through multi-sensor data fusion, UAVs enable accurate identification and quantitative assessment of bridge defects, offering strong support for structural performance analysis and long-term service life prediction [8]. The effective application of UAVs in SHM tasks fundamentally depends on accurate and reliable localization. Only when monitoring results are strictly aligned with the structural coordinates of the bridge can detected damages achieve spatial traceability and temporal comparability, thereby supporting long-term monitoring and damage evolution analysis [9].

Since SHM data analysis typically relies on high-precision sensor information (e.g., image data), UAVs are often required to capture critical bridge components at close range [10]. This operational requirement inevitably compels UAVs to frequently enter complex environments such as under-bridge areas where GPS signals are obstructed or unavailable. In such scenarios, localization accuracy may degrade significantly

\* Corresponding author.

E-mail addresses: [anna.kong@sjtu.edu.cn](mailto:anna.kong@sjtu.edu.cn) (A. Kong), [yulu.wang@sjtu.edu.cn](mailto:yulu.wang@sjtu.edu.cn) (Y. Wang), [zhaojiankang@sjtu.edu.cn](mailto:zhaojiankang@sjtu.edu.cn) (J. Zhao).

or even be completely lost, making it difficult for UAVs to perform autonomous navigation and thereby severely limiting their applicability in complex environments [11]. Traditional approaches that rely on manual operation are not only time-consuming but also constrained by limited operational coverage, which often results in missed detections [12]. Furthermore, manual operations are susceptible to communication noise or electromagnetic interference [13], which may affect both UAV hardware and software, thereby degrading measurement accuracy and compromising flight safety. Under such interference, sensor errors and image distortions can also occur, as demonstrated in [14]. To overcome these challenges, UAV autonomous navigation and localization methods for SHM data acquisition in GPS-denied conditions have attracted increasing attention in recent years.

Multi-sensor fusion has emerged as a core approach to enhancing UAV localization and navigation performance [15]. Single sensors are often constrained by environmental factors such as illumination, texture, occlusion, or electromagnetic interference, making it difficult to maintain stable accuracy in complex bridge scenarios [16]. By integrating LiDAR, vision, inertial measurement units (IMU), and other auxiliary sensors, the limitations of individual sensors can be overcome, while significantly enhancing localization accuracy, system robustness, and adaptability to complex environments [17]. However, since SHM data acquisition is characterized by high frequency and long-term operation, sensor lifespans are often rapidly depleted. In practical engineering applications, the use of low-cost sensors can not only reduce maintenance expenses but also provide greater economic benefits. Due to payload and power constraints, UAV platforms can only carry compact, low-power sensors and computationally limited onboard computers, which typically exhibit constrained performance [17]. Therefore, it is imperative to develop UAV localization algorithms tailored for the complex environments beneath bridges. Such algorithms should strike a balance between low cost, robustness, high accuracy, and low computational complexity, ensuring safe and autonomous UAV operations in SHM data acquisition tasks.

This paper proposes IRF-LIL, a high-precision localization algorithm for complex under-bridge environments, which integrates LiDAR, IMU, and infrared rangefinder. By utilizing LiDAR with a built-in IMU to construct the LIO module and introducing a residual modeling strategy based on statistical features together with a grid-based map management approach, the proposed algorithm effectively reduces computational complexity and significantly improves efficiency. For altitude estimation, an infrared rangefinder is employed to establish height constraints, thereby addressing the vertical drift issue commonly observed in conventional LIO systems during prolonged horizontal motion. By fusing these constraints with LIO through an ESKF framework, the stability of altitude estimation is substantially enhanced. Experimental results demonstrate that the proposed method can effectively suppress positional drift, verifying its effectiveness and robustness in real bridge environments. The main contributions of this work are as follows:

- (1) **An optimized LIO algorithm tailored for bridge environments.** Conventional LIO algorithms are designed for general-purpose scenarios and often rely on computationally intensive registration techniques based on point cloud surface and edge feature extraction, as well as nearest-neighbor search methods using tree structures or brute-force strategies. These approaches impose high computational burdens and are unsuitable for real-time localization on UAV platforms with limited onboard processing capabilities. To address this issue, this paper proposes an LIO algorithm that employs Normal Distributions Transform (NDT)-based point cloud registration and voxel map-based nearest-neighbor search. In bridge environments, the proposed method achieves faster pose estimation while maintaining comparable accuracy to leading LIO algorithms.
- (2) **An adaptive absolute altitude estimation module based on infrared rangefinder and RTK.** During UAV visual inspection tasks beneath bridges, conventional LIO algorithms often suffer from rapid altitude drift due to the LiDAR observation degradation in the vertical direction. To mitigate this effect, an adaptive altitude estimation module is constructed by integrating onboard infrared rangefinder and RTK data. The algorithm dynamically adjusts the weighting of sensor inputs based on their measurement conditions, enabling accurate and reliable absolute altitude estimation.
- (3) **A lightweight multi-sensor fusion algorithm.** The proposed algorithm fuses outputs from the LIO and altitude estimation modules, providing accurate, stable, and low-latency absolute pose estimation for UAV navigation beneath bridge decks. Experimental results in real bridge environments demonstrate that the method achieves an absolute position error (APE) of less than 0.17 m and reduces altitude drift to within 0.1 m, significantly enhancing localization accuracy and robustness in complex scenarios.
- (4) **A high-precision, highly integrated, low-cost UAV localization platform.** A compact, low-power LiDAR sensor with a built-in IMU is horizontally mounted on the UAV to ensure robust and accurate horizontal localization in complex environments. A vertically upward-facing infrared rangefinder, together with an RTK positioning system, provides reliable absolute altitude estimation. External parameter calibration of all onboard sensors is performed to enhance overall localization accuracy. Owing to the algorithm's low computational complexity, the system can operate on resource-constrained onboard processors, thereby reducing hardware cost.

The remainder of this article is organized as follows. Section 2 gives a brief introduction to the related work. Section 3 provides a detailed explanation of the algorithms employed in each module of the IRF-LIL method, as well as the interactions among these modules. Section 4 describes the hardware architecture of the system and outlines the design of field experiments conducted to test the algorithm. Section 5 evaluates the performance of the IRF-LIL algorithm across the experiments presented in Section 4, compares it with state-of-the-art methods, and analyzes the results. The final section concludes the paper and discusses future research directions.

## 2. Related work

This section provides an overview of recent advancements in UAV localization for bridge inspection applications in GPS-denied environments.

Recent studies demonstrate that placing offboard sensors such as ultrasonic beacon system (UBS), ultra-wideband beacon system (UWB) support localization in GPS-denied areas [18]. Beacon-based localization systems operate through wireless communication between mobile beacons mounted on the UAV and fixed beacons installed within the bridge environment [19]. Ali et al. [20] employed a UBS-based approach to achieve UAV localization and navigation beneath bridges; however, its applicability was constrained by a limited effective range. To address this limitation, Jiang et al. [21] proposed a UWB-based method, which significantly extended the localization range and improved accuracy. Nevertheless, the inability of UWB signals to penetrate obstacles restricts its effectiveness in multi-span bridge inspections [22], and extending coverage by deploying additional base stations inevitably increases deployment costs and operational complexity. Moreover, wireless signals are susceptible to interference from other communication systems [23], potentially compromising UAV flight safety and stability. Although these methods have made notable progress in UAV localization within GPS-denied under-bridge environments, there remains an urgent need for a localization approach that

is accurate, convenient, and cost-effective for large-scale, multi-span bridge inspection.

Instead of relying on supplementary stations, a promising technique called odometry has been developed for the estimation of pose for real-time UAV localization. Visual-Inertial odometry (VIO) is based on the fusion of data from visual sensors (e.g., monocular cameras, stereo cameras, or RGB-D (Red, Green, Blue, and Depth) cameras) and an IMU. One of the most prominent approaches in the field of VIO is the VINS-fusion [24] method which employs a pose graph optimization in order to achieve accurate state estimation by using the combination of a stereo camera and an IMU. However, the pose estimated by odometry is relative to the initial frame, whereas the necessity for bridge detection is absolute position. To narrow this gap, Wang et al. [11] proposed an enhanced AprilTag2-based measurement method founded on the stereo VIO for accurate global positioning of UAV. Wang et al. [25] using VIO by the aide of intermittent GPS data. Nonetheless, visual-based methods are susceptible to illumination, and the perception range is restricted by the baseline of stereo camera, which results in low robustness in complex environments such as under the bridge.

The latest advances in LiDAR technologies have met the requirements of mass-deployment on cost-effective, lightweight and reliability [26]. LiDAR sensors are inherently insensitive to variations in lighting conditions, which enables LiDAR-based odometry methods to exhibit higher robustness in environments with low texture or significant illumination changes [27]. Zhang et al. [28] proposed the LOAM algorithm, a seminal work in the field of LiDAR odometry. It decouples high-frequency odometry estimation from low-frequency map optimization and performs point cloud registration by extracting sparse edge and planar features, thereby achieving high localization accuracy. However, LOAM does not incorporate inertial information, resulting in limited robustness in dynamic environments and high-speed motion scenarios. Building upon LOAM, the LIO-SAM algorithm [29] introduces a tightly-coupled LiDAR-inertial fusion framework and employs factor graph optimization for backend constraint management. One of its notable advantages lies in enhanced system robustness through IMU integration, enabling reliable localization even under aggressive motion or degraded LiDAR data. However, due to the high computational overhead associated with its backend graph optimization, real-time performance on resource-constrained platforms remains a challenge [30]. FAST-LIO2 [31] further optimizes the LiDAR-inertial fusion pipeline by employing an ESKF combined with a recursive map update mechanism, resulting in a highly real-time LIO system. Moreover, it introduces a novel Kalman gain computation strategy that significantly reduces matrix inversion operations, thereby accelerating the filter update process. These enhancements substantially improve overall computational efficiency, making FAST-LIO2 well-suited for deployment on resource-constrained systems with stringent real-time requirements.

In under-bridge scenarios, LiDAR ranging noise remains relatively low when scanning vertical surfaces such as bridge piers, whereas it increases significantly when the beams strike horizontal surfaces like the bridge deck. Since LIO algorithms estimate the pose in a frame-by-frame manner, each estimation introduces incremental errors. As point cloud noise increases, these errors accumulate over time, leading to a degradation in overall pose estimation accuracy [32]. Furthermore, as under-bridge operations are predominantly conducted in horizontal flight, the absence of sufficient vertical motion excitation leads to poor observability in the vertical direction during point cloud registration. This results in weak constraints on altitude estimation, ultimately causing drift in the vertical component of the estimated pose and degrading the overall localization accuracy [33]. To ensure localization accuracy and robustness, Zheng et al. [32] proposed the Fast-LIVO2 algorithm, which introduces visual constraints into the inertial LiDAR odometry by incorporating image data, thereby reducing vertical drift and improving overall pose estimation accuracy. Although this method

demonstrates excellent performance in terms of computational efficiency and accuracy, its reliance on vision makes it susceptible to failure under unstable illumination conditions commonly encountered in under-bridge environments. Václav Pritzl et al. [33] proposed an inertial LiDAR-SLAM framework that integrates a laser rangefinder and a barometer to measure the UAV's relative altitude from the ground. These measurements are fused with LiDAR odometry using a loosely coupled adaptive filter, effectively mitigating vertical drift. However, this approach is designed for indoor UAV localization and only provides relative positioning with respect to the initial location, which is insufficient for bridge inspection scenarios that require real-time absolute pose estimation.

In recent years, various filtering and optimization frameworks have been proposed to address the problem of multi-sensor fusion-based localization. These approaches range from classical recursive methods, such as the Extended Kalman Filter (EKF) and the Unscented Kalman Filter (UKF), to nonlinear optimization frameworks based on graph optimization and factor graphs [15]. Such methods have played a critical role in improving localization accuracy, enhancing system robustness, and enabling adaptability in complex environments. The EKF and its variants, such as the UKF, are capable of providing higher-order accuracy in state propagation for nonlinear systems [34]. Nevertheless, their computational complexity increases substantially, especially in high-frequency LiDAR-IMU fusion scenarios, thereby imposing stringent requirements on real-time performance [35]. The Multi-State Constraint Kalman Filter (MSCKF) improves camera-IMU fusion accuracy through sliding-window constraints [36], but is computationally demanding and less effective in LiDAR-dominated scenarios. The backend of LIO-SAM is based on graph optimization, achieving global consistency and high-precision state estimation through factor graph construction [29]. However, its optimization process is computationally demanding, making it difficult to guarantee real-time performance on resource-constrained embedded UAV platforms [37]. In contrast, the Error-State Kalman Filter (ESKF) partitions the system state into nominal and error components and performs filtering operations only in the low-dimensional error space, thereby enhancing numerical stability while effectively reducing computational complexity [38]. Therefore, different multi-sensor fusion methods exhibit significant differences in accuracy and computational cost, and balancing these trade-offs to select a framework suitable for UAV platforms has become one of the central research challenges.

Although existing LIO algorithms have demonstrated promising performance across diverse scenarios and platforms, several limitations remain.

1. Prolonged horizontal motion often results in severe drift in altitude estimation.
2. The LiDAR registration and multi-sensor fusion processes involve high computational complexity, making them unsuitable for UAV platforms with constrained onboard resources.

To address these challenges, this study introduces an infrared rangefinder-based altitude estimation module, which is fused into the LIO framework through an ESKF, thereby achieving localization without altitude drift. Furthermore, within the LiDAR registration component of the LIO module, we construct residuals using statistical characteristics of the point cloud and adopt a grid-based mapping strategy. Compared with traditional LIO methods that rely on geometric feature extraction for residual construction and tree-structured map representations, the proposed approach offers faster computational performance. At the data fusion level, ESKF is employed to integrate measurements from multiple modules, which provides significantly higher computational efficiency than graph optimization-based fusion methods.

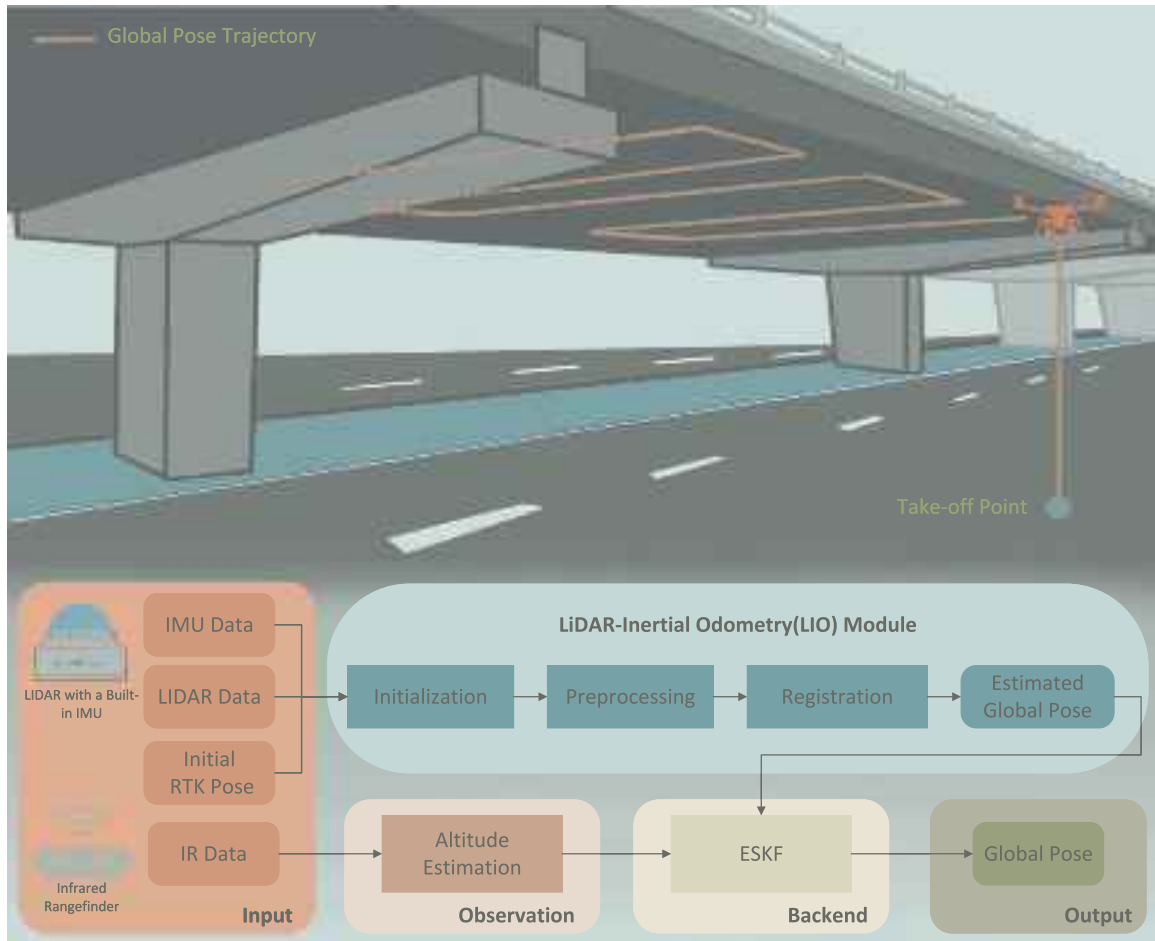


Fig. 1. Overall framework of the IRF-LIL method.

### 3. Methodology

In this section, the details of the IRF-LIL method are described. Fig. 1 shows the overall framework of the IRF-LIL which consists of four modules: LIO, Initialization, AE (Altitude Estimate), and ESKF. The LIO module employs IMU and LiDAR measurements to register the current point cloud frame with the global point cloud map, delivering a preliminary global pose estimation. The initialization module determinate the initial state of ESKF. Stable altitude estimation in the AE module is achieved through the fusion of infrared rangefinder data and absolute altitude measurements provided by an RTK GNSS receiver. The ESKF module is applied to fuse the pose estimation from the LIO module with the absolute altitude estimation from the AE module for the generation of the continuous optimal global pose estimation of the UAV. Section 3.1 introduce the mathematical fundamentals. Sections 3.2–3.5 explain the details of LIO, Initialization, AE, ESKF modules respectively.

#### 3.1. Mathematical fundamentals

In this subsection, a concise overview of the notations and fundamental operators of the UAV pose is introduced firstly. Subsequently the reference frames and notions involved in the problem are described.

##### 3.1.1. UAV pose expression

As stated in [39], the UAV localization problem can be described as estimating its six-degrees-of-freedom, including position and orientation. Mathematically, this can be represented by a  $4 \times 4$  matrix within

the three-dimensional Special Euclidean group  $SE(3)$  in Eq. (1).

$$SE(3) = \left\{ T = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \in \mathbb{R}^{4 \times 4} \mid \mathbf{R} \in SO(3), \mathbf{t} \in \mathbb{R}^3 \right\} \quad (1)$$

where  $\mathbf{R}$  is a standard  $3 \times 3$  rotation matrix in the special orthogonal group of  $SO(3)$  that represents rotation in 3D and is defined as Eq. (2).  $\mathbf{t}$  is a 3D translation ( $3 \times 1$ ) vector.

$$SO(3) = \left\{ \mathbf{R} \in \mathbb{R}^{3 \times 3} \mid \mathbf{R}\mathbf{R}^T = \mathbf{I}, \det(\mathbf{R}) = 1 \right\} \quad (2)$$

The  $SO(3)$  and the  $SE(3)$  are Lie groups that exhibit continuous and smooth properties in real spaces. Each group has a corresponding Lie algebra, denoted as  $\mathfrak{so}(3)$  and  $\mathfrak{se}(3)$ , respectively. The Lie algebra represents the tangent space at the identity element of the group and is a linear vector space. In this study, we represent the pose of the UAV using the rotation matrix  $\mathbf{R}$  from the  $SO(3)$  and the translation vector  $\mathbf{t}$  in  $\mathbb{R}^3$ . As defined in [40], the transformation between the Lie group  $SO(3)$  and its corresponding Lie algebra  $\mathfrak{so}(3)$  is represented by the following formulas:

$$\mathfrak{so}(3) = \{ \phi \in \mathbb{R}^3, \Phi = \phi^\wedge \in \mathbb{R}^{3 \times 3} \} \quad (3)$$

$$\mathbf{R} = \exp(\phi^\wedge) \quad (4)$$

in which,  $\phi^\wedge$  represents the skew-symmetric matrix form of the Lie algebra element  $\phi$  corresponding to  $\mathbf{R}$ , with its computation defined in Eq(5).

$$\phi^\wedge = \begin{bmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \end{bmatrix} = \begin{bmatrix} 0 & -\phi_3 & \phi_2 \\ \phi_3 & 0 & -\phi_1 \\ -\phi_2 & \phi_1 & 0 \end{bmatrix} \in \mathbb{R}^{3 \times 3} \quad (5)$$



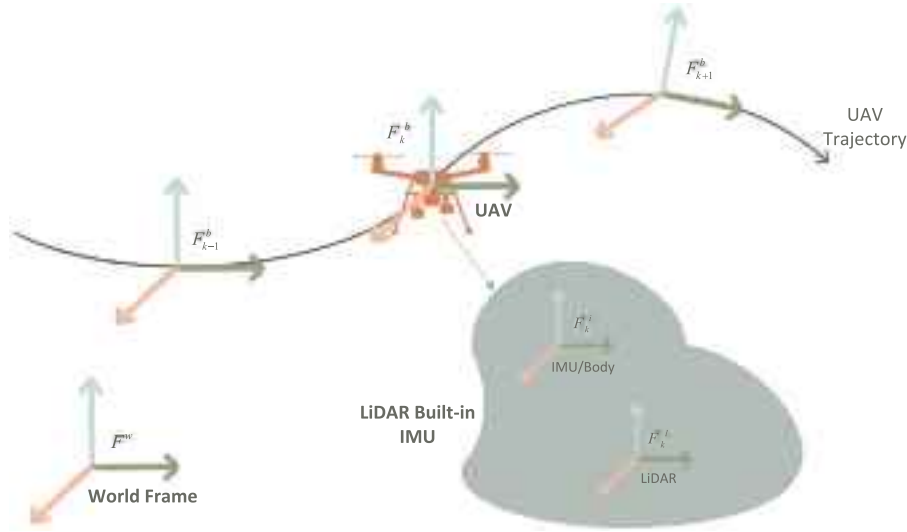


Fig. 2. References frames involved in the IRF-LIL method.

In this paper, we employ both  $\mathbf{R}$  and  $\phi$  to represent rotation.  $\mathbf{R}$  is used to transform three-dimensional points.  $\phi$  as a three-dimensional vector, allows for parameter reduction in state estimation (with the rotation matrix being nine-dimensional and the quaternion four-dimensional) for computational efficiency.

### 3.1.2. References frames and notions

In this subsection, all coordinates frames used in the localization problem and their notions are introduced. As shown in Fig. 2,  $\vec{F}^w$  is a fixed frame, known as the world frame, with its origin at the UAV's initial position in IRF-LIL, the  $x$ -axis points east, the  $y$ -axis points north, and the  $z$ -axis is opposite to the direction of gravity, oriented upward, forming a right-handed coordinate system. The reference frame of the UAV at time  $k$  is denoted as  $\vec{F}_k^b$ . Similarly,  $\vec{F}_{k-1}^b$  and  $\vec{F}_{k+1}^b$  are UAV reference frames used to represent the trajectory of the UAV. In our system, the LiDAR and IMU are rigidly attached to the UAV.  $\vec{F}^l$  represents the LiDAR frame, and  $\vec{F}^i$  represents the IMU frame, which is typically considered equivalent to the UAV's body frame. The transformation matrix  ${}^b_l\mathbf{T}$ , which converts the LiDAR frame to the IMU frame, is fixed and already determined by calibration. In this study, similar as [41], the UAV state denote as  $\chi_k$  in Eq. (6), where  $\mathbf{p}_k$  is the position of UAV,  $\mathbf{v}_k$  is the velocity,  ${}^w_b\phi_k$  is the rotation matrix from the UAV body frame to the world frame at time  $k$ ,  $\mathbf{b}_{ak}$  and  $\mathbf{b}_{gk}$  are accelerometer and gyroscope biases respectively,  $\mathbf{g}_k$  is the gravity vector.

$$\chi_k = [\mathbf{p}_k, \mathbf{v}_k, {}^w_b\phi_k, \mathbf{b}_{ak}, \mathbf{b}_{gk}, \mathbf{g}_k] \quad (6)$$

The UAV localization problem is defined as determining the transformation matrix  ${}^w_b\mathbf{T}_k$ , which is embedded within the state  $\chi_k$ . Consequently, the problem can be formally described as the estimation of the state  $\chi_k$ .

### 3.2. LIO module

Conventional LIO algorithms estimate the relative pose between the current LiDAR frame to the initial LiDAR frame by fusing points cloud data from the LiDAR with acceleration and angular velocity measurements from the IMU [42,43]. Proposed LIO module estimate a preliminary global pose by incorporating the initialization module that integrates an initial RTK pose. Fig. 3 illustrates the workflow details of the LIO module. Firstly, synchronization between a LiDAR points cloud frame and IMU measurements is required. Typically, LiDAR operates at a frequency of 10 Hz, whereas the IMU produces data at a frequency of 100 Hz or higher. Consequently, multiple IMU measurements are available between two consecutive LiDAR frames, necessitating the

temporal alignment of IMU data between the current and previous LiDAR frames. During a LiDAR scanning, UAV motion induces distortions in the point cloud data [44]. Specifically, discrepancies arise in the measured distances of the same object at the beginning and end of a single LiDAR scan. The magnitude of these distortions increases with the UAV's velocity. By leveraging IMU measurements acquired during the LiDAR scan, relative pose estimation can be performed to correct these distortions. Following distortion correction, the refined point cloud data is registered against the global map to estimate the preliminary global pose.

#### 3.2.1. LiDAR point cloud distortion rectification

For a point cloud scan at time step  $k$ , denoted as  $S_k = \{\mathbf{p}_0, \mathbf{p}_1, \dots, \mathbf{p}_n\}$ , where  $\mathbf{p}_n = (x, y, z)^T$  and  $n \in (k-1, k)$ , as shown in Fig. 4, the objective is to transform the coordinates of the points  $\mathbf{p}_n$  recorded at different time instances to the unified coordinate frame at LiDAR scan end time  $k$ , denoted as  $\mathbf{p}'_k$ , in order to correct point cloud distortion. Assuming the pose transformation  ${}^w_l\mathbf{T}_n$  from the LiDAR frame to the world frame at time  $n$  is known, according to [30], the motion compensation for point  $\mathbf{p}_n$  at time  $n$  can be expressed using the following transformation formula:

$$\mathbf{p}'_k = {}^l_k\mathbf{T}_k {}^w_l\mathbf{T}_n \mathbf{p}_n \quad (7)$$

The calculation of  ${}^l_k\mathbf{T}_k$  and  ${}^w_l\mathbf{T}_n$  necessitates the utilization of IMU data synchronized with the LiDAR point cloud. The IMU measurement model is represented as follows [45]:

$$\tilde{\mathbf{a}}_m^b = {}^b_w\mathbf{R}_m(\mathbf{a}_m^w - \mathbf{g}) + \mathbf{b}_{am} + \eta_a \quad (8)$$

$$\tilde{\omega}_m^b = \omega_m^b + \mathbf{b}_{gm} + \eta_b \quad (9)$$

$\tilde{\mathbf{a}}_m^b$  and  $\tilde{\omega}_m^b$  represent the acceleration and angular velocity measurement in the body frame at time  $m$ , which is considered equivalent to the IMU frame. The accelerometer and gyroscope biases are denoted as  $\mathbf{b}_{am}$  and  $\mathbf{b}_{gm}$ , respectively, with their first-order derivatives modeled as Gaussian white noise.  $\eta_a$  and  $\eta_b$  represent the measurement noise of the accelerometer and gyroscope, respectively, both of which are modeled as Gaussian white noise.

The point cloud data  $S_k$  at time  $k$  is associated with multiple IMU measurements, with timestamps distributed across the interval  $[k-1, k]$ . For any time  $m$  within the interval  $[k-1, k]$ , the transformation matrix  ${}^w_b\mathbf{T}_m$  from the IMU frame to the world frame can be recursively derived from the known state  $\chi_{k-1}$ .  ${}^w_b\mathbf{T}_m$  can be computed using

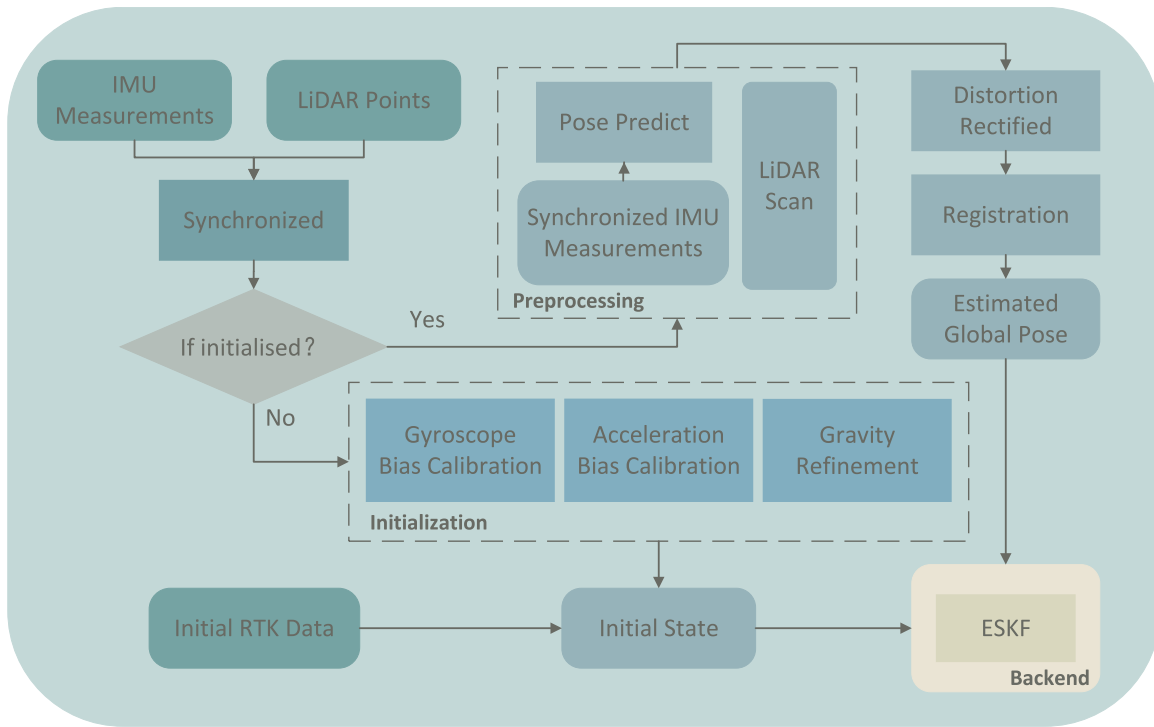


Fig. 3. The workflow of LIO modules.

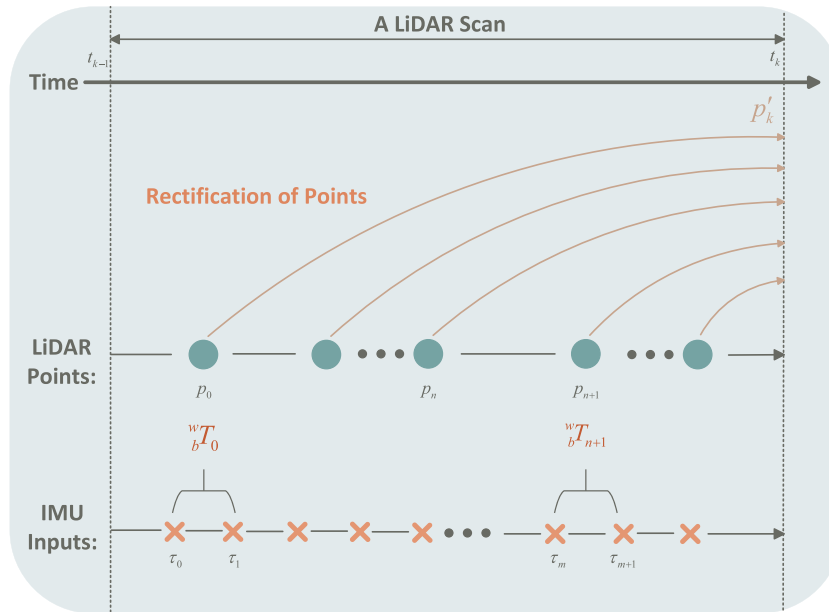


Fig. 4. The schematic of LiDAR point cloud distortion rectification.

Eqs. (10)–(13) [39]. Where  $\Delta t$  represents the time interval between two consecutive IMU measurements.

$${}^w_b \mathbf{R}_m = {}^w_b \mathbf{R}_{k-1} \prod_{j=k-1}^{m-1} \exp((\hat{\omega}_j^b - \mathbf{b}_{g_j})^\wedge \Delta t) \quad (10)$$

$$\mathbf{p}_m = \mathbf{p}_{k-1} + \sum_{j=k-1}^{m-1} \left[ \mathbf{v}_j \Delta t + \frac{1}{2} \mathbf{g} \Delta t^2 \right] + \frac{1}{2} \sum_{j=k-1}^{m-1} {}^w_b \mathbf{R}_j (\hat{\mathbf{a}}_j^b - \mathbf{b}_{a_j}) \Delta t^2 \quad (11)$$

$$\mathbf{v}_m = \mathbf{v}_{k-1} + \sum_{j=k-1}^{m-1} [{}^w_b \mathbf{R}_j (\hat{\mathbf{a}}_j^b - \mathbf{b}_{a_j}) \Delta t + \mathbf{g} \Delta t] \quad (12)$$

$${}^w_b \mathbf{T}_m = \begin{bmatrix} {}^w_b \mathbf{R}_m & \mathbf{p}_m \\ \mathbf{0} & 1 \end{bmatrix} \quad (13)$$

For each point in the point cloud  $S_k$ , the transformation matrix  ${}^w_b \mathbf{T}_n$  at the corresponding timestamp  $n$  can be obtained through linear interpolation of  ${}^w_b \mathbf{T}_m$ . The calculation for point cloud distortion rectification, as expressed in Eq. (7), can then be reformulated as:

$$\mathbf{p}'_k = {}^l_b \mathbf{T}_w {}^b_b \mathbf{T}_k {}^w_b \mathbf{T}_n {}^l_b \mathbf{T}_p \mathbf{p}_n \quad (14)$$

### 3.2.2. Point cloud registration

LiDAR point cloud registration is commonly performed using two primary approaches: scan-to-scan and scan-to-map. The scan-to-scan method directly aligns the current point cloud scan with the preceding scan by employing point cloud registration algorithms to estimate the relative pose. Chen et al. [46] and Zhang et al. [28] utilize this method for a fast pose estimation in their front-end odometry module. This method is computationally efficient, offers high real-time performance, and exhibits robustness to dynamic objects in the environment. However, it estimates only the relative pose between consecutive frames, leading to rapid accumulation of drift over time. In contrast, the scan-to-map approach aligns the current point cloud scan with a global map constructed from historical keyframes to determine the UAV's pose [47–49]. By leveraging the extensive information within the global map, this approach provides greater stability in the registration process and mitigates drift. However, as the size of the global map increases, the computational complexity grows, and its performance degrades in environments with high levels of dynamic interference. Bridge inspection scenarios typically involve UAVs operating at slow speeds within a single-span bridge, characterized by prolonged inspection times, limited movement ranges, and minimal dynamic disturbances. These scenarios are distinguished by frequent registration requirements, small-scale point cloud maps, and low environmental dynamics. Given these characteristics, this study employs the scan-to-map approach for point cloud registration, as it aligns more effectively with the demands of bridge inspection tasks.

Scan matching algorithms are generally classified into feature-based and statistics-based methods. Feature-based approaches extract feature points from the current point cloud and align them with the target point cloud using algorithms such as Iterative Closest Point (ICP) [50] to estimate the pose. Shan et al. [51] and Lin et al. [52] computed LiDAR odometry by solving point-to-plane and point-to-edge residual equations constructed from features extracted from the point cloud. These methods demonstrate high accuracy in pose estimation and rapid convergence in environments with abundant features. However, their performance significantly degrades in feature-sparse or low-density environments. Moreover, the computational cost associated with feature extraction results in longer pose estimation times, limiting their applicability in real-time scenarios.

In contrast, the statistics-based Normal Distributions Transform (NDT) [53] algorithm splits the target point cloud into multiple voxels based on resolution and calculates the Gaussian distribution of points within each voxel. During registration, the pose is estimated by constructing residual equations based on the relationship between points in the current point cloud and their corresponding voxels. The NDT algorithm [54] eliminates the need for feature extraction and nearest-neighbor searches, leading to lower computational cost. However, in feature-rich environments, its pose estimation accuracy is marginally lower compared to feature-based ICP. Considering the characteristics of bridge inspection scenarios, which involve sparse point clouds, limited scene features, and computational constraints on UAV platforms, this study employs the NDT algorithm for scan matching.

Eqs. (15) and (16) are the residual equations,  $\mathbf{p}_j$  represents a point in the point cloud,  $\mu_j$  and  $\Sigma_j$  denote the mean and variance of the Gaussian distribution of points in the corresponding voxel, respectively. The Gauss-Newton method is used to iteratively solve the residual equations for the optimal estimation of rotation  ${}^w\mathbf{R}_b^*$  and translation  ${}^w\mathbf{t}_b^*$ .

$$\mathbf{e}_j = {}^w\mathbf{R}_l^b \mathbf{R}_{p_j} + {}^w\mathbf{t}_l - \mu_j \quad (15)$$

$$({}^w\mathbf{R}_b, {}^w\mathbf{t}_b)^* = \arg \min_{{}^w\mathbf{R}_b, {}^w\mathbf{t}_b} \sum_j \mathbf{e}_j^T \Sigma_j^{-1} \mathbf{e}_j \quad (16)$$

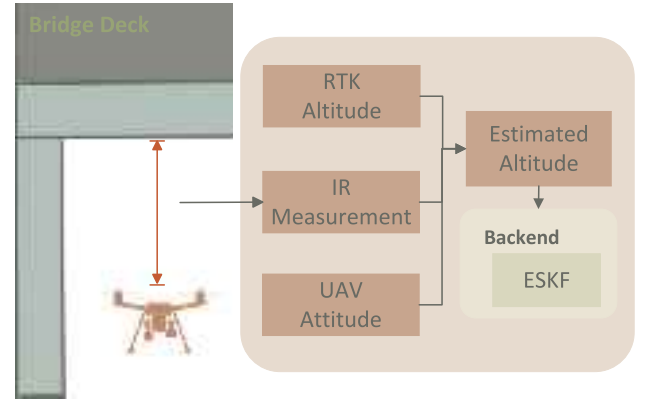


Fig. 5. The workflow of altitude estimate module.

### 3.3. Initialization

The initialization process aims to determine the system's initial state, including the initial position, orientation, and IMU biases. In this study, initialization is conducted under conditions where the RTK signal is reliable, and the UAV remains stationary. Under stationary conditions, the gyroscope of the IMU is assumed to measure only its bias, while the accelerometer captures the combined effect of its bias and gravity. To estimate the biases, the mean values of the accelerometer and gyroscope measurements over a stationary period are computed and denoted as  $\bar{\mathbf{a}}$  and  $\bar{\boldsymbol{\omega}}$ , respectively. The gyroscope bias is given by  $\mathbf{b}_g = \bar{\boldsymbol{\omega}}$ . Then the gravity vector  $\mathbf{g}$  is calculated as:

$$\mathbf{g} = \frac{\bar{\mathbf{a}}}{|\bar{\mathbf{a}}|} G \quad (17)$$

where  $G$  represents the gravitational acceleration constant. The accelerometer bias  $\mathbf{b}_a$  is then determined by subtracting the estimated gravity vector  $\mathbf{g}$  from the accelerometer measurements and recomputing the mean value over the stationary period.

The initial orientation is represented by three attitude angles: roll, pitch, and yaw. Roll and pitch are estimated based on the computed gravity vector  $\mathbf{g}$ , while yaw is determined from the average of dual-antenna RTK readings over a stationary period. The initial position is obtained by converting latitude and longitude values provided by the RTK system into Universal Transverse Mercator (UTM) coordinates. After the initialization process, the system's position is estimated within a world coordinate frame, where the takeoff point serves as the origin. The coordinate axes are defined as follows: the positive  $x$ -axis points east, the positive  $y$ -axis points north, and the positive  $z$ -axis points upward and perpendicular to the ground plane.

### 3.4. Altitude estimate module

The altitude estimation module, which fuses measurements from RTK positioning and an infrared rangefinder, enables the UAV to obtain accurate absolute altitude estimates during bridge inspection tasks. The infrared rangefinder operates by emitting infrared rays toward the target and measuring the reflected signals to determine the distance between the sensor and the object. As shown in Fig. 5, in the UAV inspection platform proposed in this study, an upward-facing infrared rangefinder mounted on top of the UAV measures the distance from objects above to the UAV, with a maximum effective detection range denoted as  $D_{max}$ . The infrared rangefinder employed exhibits the typical noise characteristics of low-cost time-of-flight sensors. Its measurement noise can be modeled as zero-mean Gaussian white noise, with variance increasing as the target distance grows. Within the specified measurement range, the sensor provides reliable output. However, measurements beyond this range are highly uncertain. Multipath reflections

on highly reflective surfaces and signal attenuation on low-reflective or transmissive materials further increase measurement uncertainty. To mitigate these effects, the proposed algorithm incorporates adaptive weighting strategies, outlier rejection mechanisms, and fusion with LIO through an ESKF, thereby enhancing overall robustness and accuracy. Given the RTK height measurement as  $H_{RTK}$ , and the distance from the object above to the UAV as  $D$  the estimated absolute altitude of the UAV at time  $k$ , denoted as  $H_k$ , can be expressed as follows:

$$H_k = c_1 H_{RTK} + (1 - c_1)(c_2(D_{k-1} \cos \theta_{k-1} \cos \psi_{k-1} - D_k \cos \theta_k \cos \psi_k + \chi_{k-1,3}) + (1 - c_2)\chi_{k,3})$$

$$c_1 = \begin{cases} 1, & \text{if RTK fix} \\ 0, & \text{if RTK not fix} \end{cases} \quad (18)$$

$$c_2 = \begin{cases} 1 - \frac{D_k}{D_{\max}} c_3, & D_k \leq D_{\max} \text{ and } D_{k-1} \leq D_{\max} \\ 0, & D_k > D_{\max} \text{ or } D_{k-1} > D_{\max} \end{cases}$$

$\theta$  and  $\psi$  are the pitch and roll angles of the UAV, respectively, which are computed from Lie algebra elements of the state vector via Rodrigues' rotation formula and subsequent calculations.  $\chi_{k,3}$  is the third component of state vector which represents UAV altitude. Attitude angles and state vectors at timestamp  $k-1$  are posterior estimates, whereas those at timestamp  $k$  are prior estimates. The coefficients  $c_1$ ,  $c_2$ , and  $c_3$  are designed to regulate the contributions of RTK, the infrared rangefinder, and LIO in altitude estimation. Specifically,  $c_1$  controls the weighting between RTK-derived altitude and infrared-LIO estimates. When RTK measurements are valid, the system relies exclusively on RTK values, otherwise, altitude is derived solely from the infrared rangefinder and LIO outputs. The parameter  $c_2$  adjusts weighting between altitudes obtained by the infrared rangefinder and the LIO system. Height measurements from the infrared rangefinder are discarded when beyond the effective detection range. Within detection range, the weight of rangefinder-derived altitude decreases linearly with increasing distance to reflect the error characteristics of the rangefinder. The rate of this decrease is determined by the empirical coefficient  $c_3$ , ranging between 0 and 1, which is derived from the noise characteristics of the specific infrared sensor used. For sensors whose measurement noise increases slowly with distance, a smaller value of  $c_3$  is adopted, while a larger value is used when the noise grows more rapidly. In the experiments presented in this study,  $c_3$  was set to 0.1.

In practical applications, the infrared rangefinder may be affected by interference, leading to abnormal measurements such as sudden jumps or short-term omissions. To prevent such anomalies from introducing errors into the system, the altitude estimation module monitors the difference between the current and previous measurements. If the absolute change is significantly larger than the altitude increment estimated by the LIO system, the measurement is classified as a jump, and the coefficient  $c_2$  is set to zero, thereby excluding the current reading from altitude estimation. In cases of short-term omissions, the system applies linear regression based on the most recent  $N$  valid measurements to extrapolate the current range value and estimate altitude. To balance computational efficiency and fitting accuracy,  $N$  is set to 5 in this study.

At this stage, a preliminary altitude estimate has been obtained. This estimate will be integrated into the back-end ESKF for further optimization, resulting in a more accurate and stable altitude estimation.

### 3.5. ESKF

In integrated navigation systems, multi-sensor data fusion is typically performed using either filtering-based or optimization-based methods to achieve optimal pose estimation. Among filtering-based methods, the EKF and the ESKF are widely adopted due to their capability to handle nonlinear systems efficiently while maintaining computational simplicity. The EKF directly models and estimates the full state of the system, with both prediction and update operations performed in the complete state space. In contrast, the ESKF separates the

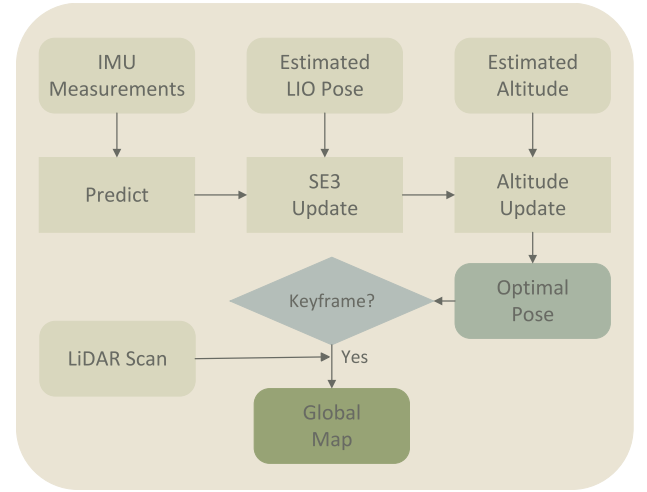


Fig. 6. The pipeline of the ESKF applied in IRF-LIL method.

system state into nominal and error states, focusing the filtering process on the error state, which simplifies computation. Addressing rotational nonlinearity in the EKF often requires high-dimensional representations such as quaternions or rotation matrices, or low-dimensional but singular-prone representations like Euler angles, increasing the complexity of modeling and computation. In comparison, the ESKF employs a three-dimensional rotational error vector, facilitating more efficient modeling and computation. Additionally, the linearization error in EKF tends to amplify with increased system complexity, reducing estimation accuracy. The ESKF, however, offers greater stability and improved performance in such scenarios.

As shown in Fig. 6. The ESKF consists of two main steps: prediction and update. During the prediction step, when IMU data is received, the nominal state is propagated based on the motion model, while the mean and covariance of the error state are computed. In the update step, the pose estimation obtained from LiDAR odometry and the altitude value from the altitude estimation module are utilized to update the mean and covariance of the error state. Subsequently, the updated error state values are incorporated into the nominal state variables to produce the optimal pose estimation. After obtaining the optimal pose estimation, the system performs keyframe identification for the current frame. If identified as a keyframe, the system transforms the LiDAR point cloud of the current body frame into the world coordinate system using the optimal pose estimation and integrates it into the global map. Keyframe identification criteria involve comparisons between the current and previous optimal pose estimations, specifically assessing whether translational displacement or rotational change exceeds predefined thresholds. This approach reduces map size and enhances LiDAR registration efficiency.

#### 3.5.1. ESKF prediction step

The complete state variable is defined in Eq. (6), while the nominal state variable at time step  $k$  is denoted as  $\hat{\chi}_k$ , and the error state variable is denoted as  $\delta\chi_k$ . Their relationship is described in Eq. (19). The symbol  $\oplus$  denotes a generalized addition operation [40]. When performing addition on rotational variables within the state, it is computed using a right multiplication form, as shown in Eq. (20), for other variables, standard addition is applied.

$$\chi_k = \hat{\chi}_k \oplus \delta\chi_k \quad (19)$$

$${}^u_b \mathbf{R}_k = {}^u_b \hat{\mathbf{R}}_k \exp(\delta {}^u_b \phi_k^\wedge) \quad (20)$$

$$\delta\chi_{t+\Delta t} = \mathbf{H}\delta\chi_t + \mathbf{w}, \mathbf{w} \sim \mathcal{N}(0, \mathbf{Q}) \quad (21)$$



$$\mathbf{H} = \begin{bmatrix} \mathbf{I} & \mathbf{I}\Delta t & 0 & 0 & 0 & 0 \\ 0 & \mathbf{I} & -{}^w\mathbf{R}_l(\hat{\mathbf{a}}_l^b - \mathbf{b}_{al})^\wedge \Delta t & 0 & -{}^w\mathbf{R}_l \Delta t & \mathbf{I}\Delta t \\ 0 & 0 & \exp(-(\hat{\omega}_l^b - \mathbf{b}_{gl})^\wedge \Delta t) & -\mathbf{I}\Delta t & 0 & 0 \\ 0 & 0 & 0 & \mathbf{I} & 0 & 0 \\ 0 & 0 & 0 & 0 & \mathbf{I} & 0 \\ 0 & 0 & 0 & 0 & 0 & \mathbf{I} \end{bmatrix} \quad (22)$$

When IMU data is received, the nominal state is computed using Eqs. (10)–(13). The error state is propagated according to the ESKF motion model Eq. (21) [55], as defined in Eqs. (23) and (24).  $\mathbf{w}$  represents the process noise in the motion model, modeled as Gaussian noise with zero mean and a covariance matrix  $\mathbf{Q}$ , whose values are determined by the characteristics of the IMU sensor. According to the definition in [30], the  $\mathbf{H}$  matrix in this study is specified in Eq. (22), represents the Jacobian of the motion model with respect to the error state.

$$\delta\chi_{t,pred} = \mathbf{H}\delta\chi_{t-1} \quad (23)$$

$$\mathbf{P}_{t,pred} = \mathbf{H}\mathbf{P}_{t-1}\mathbf{H}^T + \mathbf{Q} \quad (24)$$

### 3.5.2. ESKF update step

In the ESKF update process, sensor measurements are used to update the error state. Assuming the sensor observation model is given by:

$$\mathbf{z} = f(\mathbf{x}) + \mathbf{v}, \mathbf{v} \sim \mathcal{N}(\mathbf{0}, \mathbf{V}) \quad (25)$$

$$\mathbf{J} = \left. \frac{\partial f}{\partial \delta\chi} \right|_{\chi_{k,pred}} \quad (26)$$

where  $\mathbf{z}$  represents the observation data, and  $\mathbf{v}$  is the observation noise modeled as Gaussian noise with zero mean and a covariance matrix  $\mathbf{V}$ . The observation process begins by calculating the Jacobian matrix  $\mathbf{J}$  of the observation model with respect to the error state, which is defined in Eq. (26). Subsequently, the Kalman gain  $\mathbf{K}$  is calculated, which is then used to update the error state (Eqs. (27)–(30)) [55].  $\chi_k$  represents the optimal estimation of the system state, which is used for UAV localization and control.  $\mathbf{P}_k$  is the covariance matrix of the system, reflecting the confidence level of the estimated state.

$$\mathbf{K} = \mathbf{P}_{k,pred}\mathbf{J}^T(\mathbf{J}\mathbf{P}_{k,pred}\mathbf{J}^T + \mathbf{V})^{-1} \quad (27)$$

$$\delta\chi_k = \mathbf{K}(\mathbf{z} - f(\chi_{k,pred})) \quad (28)$$

$$\chi_k = \chi_{k,pred} \oplus \delta\chi_k \quad (29)$$

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{K}\mathbf{J})\mathbf{P}_{k,pred} \quad (30)$$

The data obtained from LiDAR point cloud registration provides observations of both position and orientation. The corresponding observation model and its Jacobian matrix are defined in Eqs. (31)–(32).  $\text{Log}()^\vee$  denotes the operation that converts a rotation matrix into its corresponding rotation vector [40],  ${}^w\mathbf{t}_{lidar}$  and  ${}^w\mathbf{R}_{lidar}$  represent translation and rotation obtained from LIO measurements as defined by Eq. (16).

$$\mathbf{z}_{lidar} - f(\chi_{k,pred})_{lidar} = \begin{bmatrix} {}^w\mathbf{t}_{lidar} - {}^w\mathbf{t}_{k,pred}, \text{Log}({}^w\mathbf{R}_{k,pred}^T {}^w\mathbf{R}_{lidar})^\vee \end{bmatrix}^T \quad (31)$$

$$\mathbf{J}_{lidar} = \begin{bmatrix} \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 3} & \mathbf{0}_{3 \times 12} \\ \mathbf{0}_{3 \times 3} & \mathbf{I}_{3 \times 3} & \mathbf{0}_{3 \times 12} \end{bmatrix} \quad (32)$$

The altitude estimation module provides observations for the third dimension of the position state variable. The corresponding observation model and Jacobian matrix are defined in Eqs. (33)–(34).  $h_{alt}$  represents

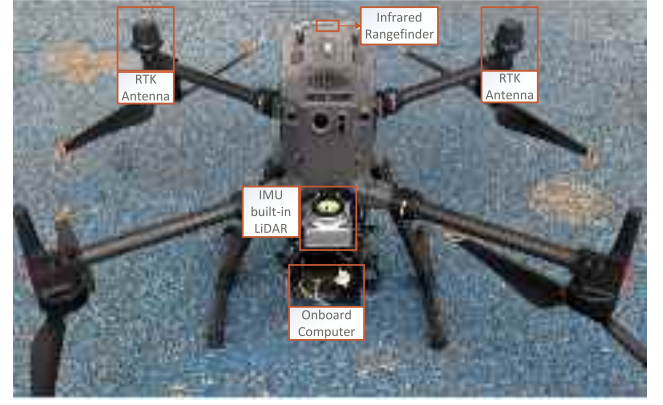


Fig. 7. UAV system for evaluating the proposed algorithm.

the altitude measurement data obtained by the altitude estimation module, as calculated using Eq. (18).  ${}^w\mathbf{h}_{k,pred}$  denotes the third-dimensional component of the translation state  ${}^w\mathbf{t}_{k,pred}$ .

$$z_{alt} - f(\chi_{k,pred})_{alt} = h_{alt} - {}^w\mathbf{h}_{k,pred} \quad (33)$$

$$\mathbf{J}_{alt} = \begin{bmatrix} 0 & 0 & 1 & \mathbf{0}_{1 \times 15} \end{bmatrix} \quad (34)$$

## 4. Experimental validation

The proposed algorithm was first evaluated on publicly available datasets and compared against leading SLAM methods, thereby enhancing the credibility of the approach.

Subsequently, to ensure flight safety in real bridge environments, the proposed localization algorithm was first evaluated in an outdoor setting with stable RTK signals. The experimental setting simulated an under-bridge scenario, where RTK-based positioning data served as the ground truth. In this setup, the UAV operated in an environment without overhead obstructions, surrounded by buildings and flat ground below. The LiDAR and infrared rangefinder were mounted horizontally in a downward-facing configuration beneath the UAV, enabling scanning of the ground and surrounding structures to simulate inspection conditions under a bridge. A subsequent experiment was conducted in an actual bridge environment, in which the UAV's ground truth positions were measured using a total station. In this experiment, the LiDAR was horizontally mounted in a forward-facing direction to accommodate the operational requirements of bridge inspection tasks. To assess its applicability in real-world UAV bridge inspection scenarios, the proposed IRF-LIL algorithm is evaluated in comparison with two state-of-the-art LIO methods, Fast-LIO2 [31], LIO-SAM [29] and LIO-SAM with loop closure, both of which are extensively adopted in practical engineering applications.

The under-bridge environment simulation experiment was conducted within a campus setting. The real-world bridge experiment was performed beneath a highway bridge located on an expressway. These experiments aim to evaluate the robustness and positioning accuracy of the UAV localization algorithm. The specifics of the systems are outlined in the following section.

### 4.1. Platforms

In this study, to guarantee the safety and stability of flying, the DJI M300 UAV system was selected as the flight platform, which integrates IMU, Dual-antenna RTK module and infrared rangefinders, as shown in Fig. 7. The data generated by the sensors is available to users via the

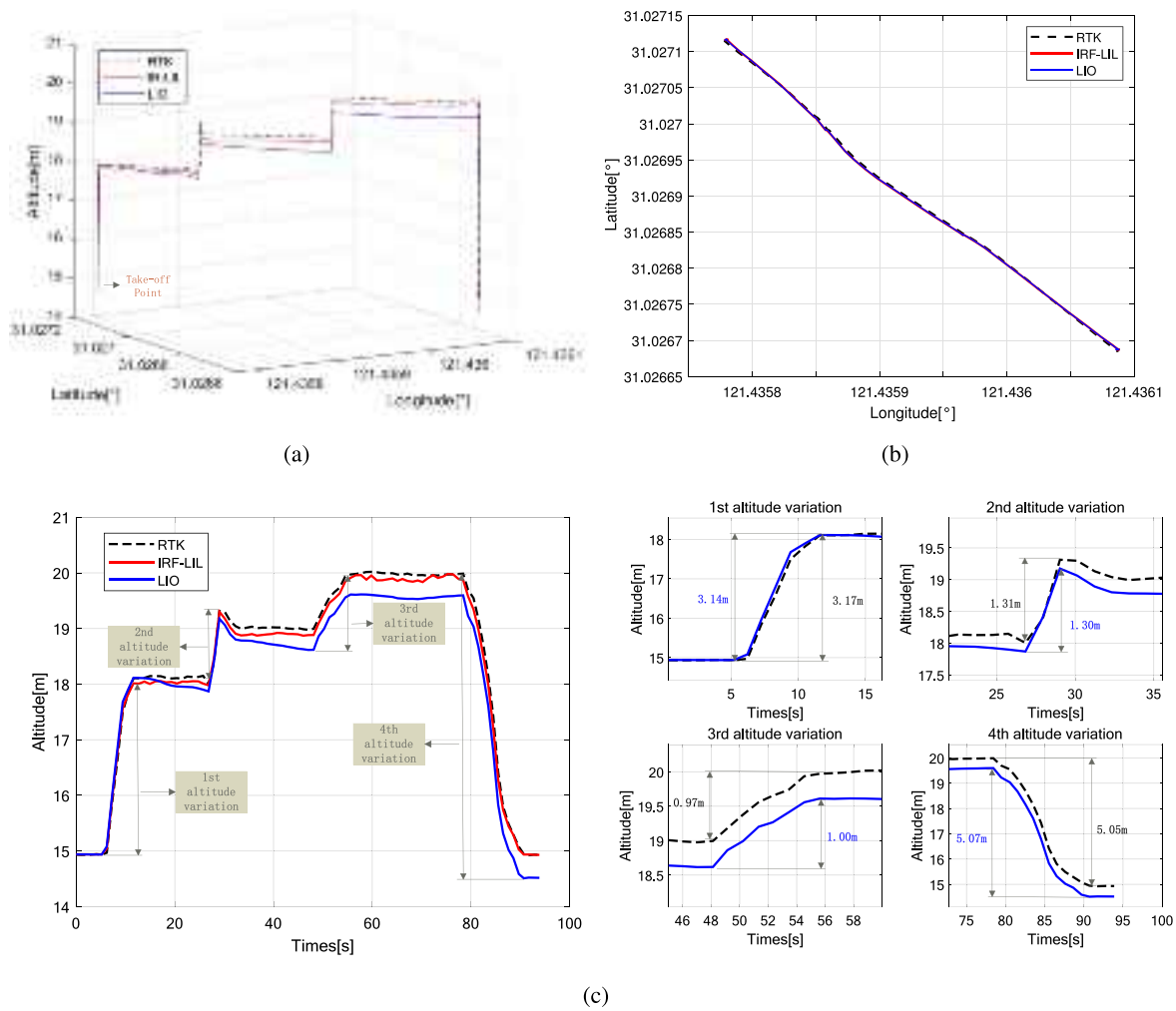


Fig. 8. Trajectories of straight stepwise ascent flight test: (a) 3D trajectory, (b) X-Y plane trajectory, (c) Vertical trajectory.

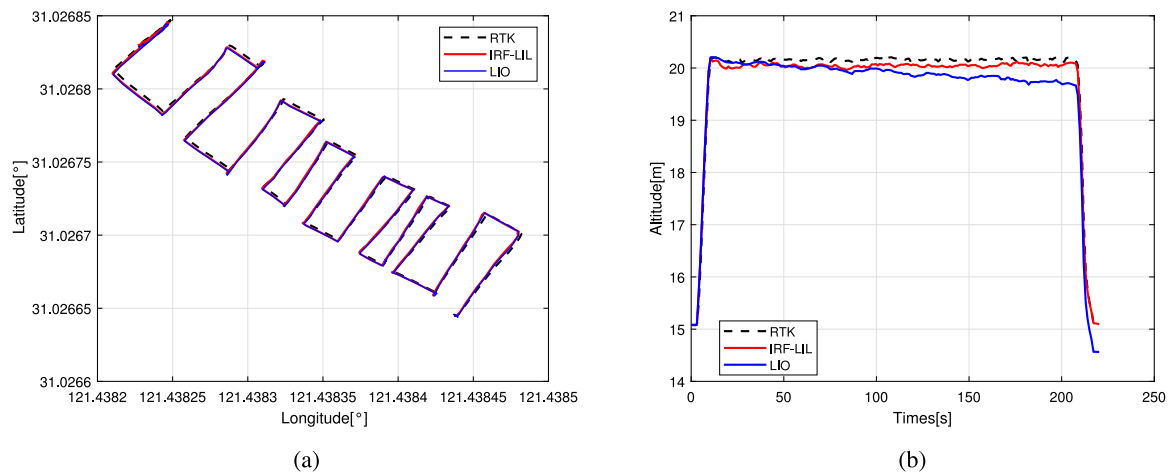


Fig. 9. Trajectories of zigzag flight test: (a) X-Y plane trajectory, (b) Vertical trajectory.

official interface. Moreover, the UAV was equipped with two removable 4920mAh batteries, which afford a flight duration of approximately 40 min, and a payload capacity of 9 kg.

The external LiDAR sensor is Livox Mid-360, which integrates an IMU module. The LiDAR is capable of scanning 200000 points per second, with a detection range of approximately 20 m. The size of LiDAR is 60 mm \* 65 mm \* 60 mm, with a mass of only 265 g, which

is well-suited for the limited payload of drone. The data processing unit utilizes an onboard computer equipped with the Rockchip RK3588 chipset, which integrates an octa-core CPU comprising four Cortex-A76 cores and four Cortex-A55 cores. The system operates on the Ubuntu 20.04 platform. The aforementioned external appliances were integrated and mounted on the UAV by means of a customized frame, as shown in Fig. 7.

**Table 1**

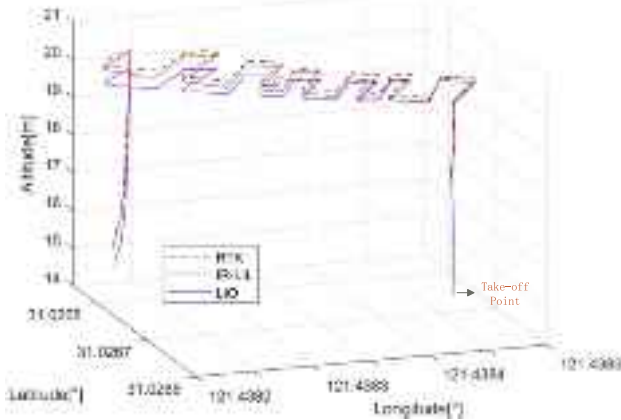
Average processing time per scan (unit: ms) and absolute translation error (unit: m) of different algorithms on public datasets. Mean represents the average value of APE, and std denotes the standard deviation of APE.

Trajectory	Method	APTS	Mean	Std	X-Y mean	X-Y Std	Altitude mean	Altitude Std
eee01	LIO	<b>3.12</b>	<b>0.066</b>	<b>0.033</b>	0.052	0.034	0.029	0.026
	Fast-LIO2	9.33	0.070	0.043	0.053	0.037	0.032	0.037
	LIO-SAM	86.35	0.081	0.042	0.060	0.028	0.042	0.044
eee02	LIO	<b>2.53</b>	0.067	0.033	0.059	0.037	0.020	0.019
	Fast-LIO2	8.40	<b>0.065</b>	<b>0.029</b>	0.058	0.030	0.021	0.020
	LIO-SAM	77.83	0.073	0.034	0.058	0.029	0.035	0.033
eee03	LIO	<b>2.10</b>	0.108	0.052	0.096	0.050	0.027	0.040
	Fast-LIO2	9.25	<b>0.099</b>	0.045	0.091	0.040	0.022	0.030
	LIO-SAM	85.30	0.105	<b>0.040</b>	0.090	0.036	0.041	0.041
street04	LIO	<b>12.71</b>	<b>0.324</b>	0.173	0.303	0.179	0.095	0.044
	Fast-LIO2	19.17	0.333	<b>0.152</b>	0.309	0.158	0.104	0.050
	LIO-SAM	87.69	0.630	0.336	0.538	0.325	0.287	0.178
street05	LIO	<b>8.05</b>	<b>0.311</b>	0.141	0.191	0.079	0.238	0.133
	Fast-LIO2	16.40	0.346	<b>0.132</b>	0.275	0.113	0.188	0.112
	LIO-SAM	84.01	0.362	0.146	0.294	0.121	0.192	0.119
street06	LIO	<b>8.66</b>	<b>0.324</b>	<b>0.107</b>	0.252	0.099	0.173	0.113
	Fast-LIO2	17.01	0.418	0.130	0.325	0.107	0.233	0.143
	LIO-SAM	88.06	0.387	0.158	0.370	0.164	0.082	0.065

**Table 2**

Statistics of the absolute position error between the estimated trajectories and RTK trajectory. Mean represents the average value of APE, and std denotes the standard deviation of APE.

Trajectory	Method	Mean (m)	Std (m)	X-Y mean (m)	X-Y Std (m)	Altitude mean (m)	Altitude Std (m)
Stepwise	IRF-LIL	0.293	0.159	0.268	0.169	0.090	0.053
	LIO	0.436	0.184	0.275	0.177	0.302	0.163
Zigzag	IRF-LIL	0.262	0.105	0.228	0.118	0.109	0.048
	LIO	0.366	0.167	0.236	0.128	0.255	0.160

**Fig. 10.** 3D trajectory of zigzag flight test.

#### 4.2. Experimental setup and grand truth acquisition

The NTU-VIRAL [56] and M2DGR [57] public datasets were employed to validate the proposed algorithm. The NTU-VIRAL dataset was collected on an UAV platform equipped with a 3D LiDAR, stereo cameras, and an IMU, with ground truth provided by an optical motion capture system or RTK. The experimental scenarios cover indoor corridors and outdoor open areas, where the eee01–03 sequences selected in this study correspond to an outdoor campus building environment. The M2DGR dataset was recorded using a ground robot platform carrying a 3D LiDAR, cameras, an IMU, and GNSS receivers, with ground truth trajectories obtained from a high-precision GNSS-RTK system. The street04–06 sequences selected in this study represent campus environments with complex structures. Since the public datasets do not provide infrared rangefinder measurements, the experiments focus solely on the proposed LIO algorithm, with comparisons against FAST-LIO2 and LIO-SAM to demonstrate feasibility and strengthen credibility.

The under-bridge simulation experiment was conducted on a street within the campus, surrounded by buildings and trees. The area of the site is approximately 300 square meters, with dimensions of 30 m in length and 10 m in width. In an unobstructed open environment, the RTK signal remained continuous and stable throughout the experiment. The dual-antenna RTK system provided centimeter-level positioning accuracy, making it a reliable reference for ground truth. In the experiment, the predefined flight path was completed by manually operating the UAV through a remote controller.

The site of real-world bridge experiment is a part of Xiangjiang Bridge which is located in Changsha, Hunan Province, China. The bridge has a width of 38 m, with each one span measuring approximately 66 m, and the height of the bridge pier is about 22 m. In experimental trial, the drone was operated manually to fly beneath a single bridge span. The ground truth for this experiment was obtained using a total station set up on the exterior of the bridge, as shown in Fig. 11. In order to facilitate a comparative analysis with the other leading LiDAR SLAM algorithm, the LiDAR points cloud, IMU data and infrared rangefinder data were recorded as ROS bag files for post-processing.

To ensure the accuracy of multi-sensor data fusion, all sensors were carefully calibrated prior to the experiments. Specifically, a total station was employed as a high-precision measurement instrument to calibrate the installation position and orientation relationships among the LiDAR, IMU, infrared rangefinder, and RTK antenna. The total station provides millimeter-level spatial accuracy and, by observing predefined markers on each sensor, allows computation of the extrinsic transformation matrices (rotation and translation) between sensors. The extrinsic parameters between the LiDAR and IMU were obtained directly from the product specification manual. During the calibration procedure, the LiDAR center was defined as the global reference point, after which the installation positions of the other sensors were measured sequentially. Four corner points on the LiDAR XY-plane were additionally measured as a reference plane to align the coordinate systems. The final calibration yielded the complete set of extrinsic transformation matrices used for subsequent data fusion.



Fig. 11. Real-world bridge experiment setup.

## 5. Performance evaluation

Absolute Position Error (APE) is a commonly used evaluation metric to evaluate the accuracy of position estimation. The APE is defined as the Euclidean distance between the true and estimated positions. In this study, the APE calculated by computing the square root of the sum of the squared differences between the corresponding coordinates of the true and estimated positions.

Average Processing Time Per Scan (APTS) refers to the average amount of time required to process a single scan of sensor data, such as a LiDAR or radar scan. This metric denotes the duration taken by the algorithm during the analysis, processing, and extraction of relevant information from each individual scan. It is typically used to evaluate the computational efficiency of algorithms in real-time applications and is usually measured in milliseconds (ms).

### 5.1. The public datasets experiment

In the public datasets, the proposed LIO algorithm was evaluated in terms of APE and APTS, with comparisons against FAST-LIO2 and LIO-SAM. As shown in Table 1, the proposed algorithm consistently achieved the fastest pose estimation speed across all sequences, significantly lower than that of FAST-LIO2 and LIO-SAM. This efficiency stems from the LiDAR registration strategy in the LIO module, where residuals are constructed from statistical characteristics of point cloud and a voxel-based map representation is adopted, resulting in reduced computation time. These results highlight the advantage of the proposed LIO algorithm on computation-constrained platforms.

For trajectory error evaluation, since the ground-truth trajectories and the outputs of all LIO algorithms lie in different coordinate frames, Umeyama [58] alignment was employed to unify the trajectories before error computation. The proposed algorithm exhibited the best overall localization accuracy in most sequences, with deviations from the best accuracy in other sequences remaining below 1 cm. In the street04 and street05 sequences, the proposed algorithm achieved the lowest mean localization error but showed larger standard deviations than competing methods. This indicates that the algorithm can deliver higher accuracy in certain scenarios, though its localization stability requires further improvement.

With respect to altitude errors, all street sequences demonstrated noticeable drift, particularly in street05 where the error increased significantly. This can be attributed to the long, straight-line ground vehicle motion in that sequence, which reduced vertical observability and aggravated drift. In contrast, the eee sequences involved frequent vertical UAV motions, which enhanced observability in the vertical direction and thereby mitigated altitude drift.

### 5.2. The simulated bridge-underneath experiment

In this scenario, two experimental trajectories were designed: one is a unidirectional horizontal path with stepwise increases in altitude, while the other is a constant altitude trajectory that simulates a zigzag path for bridge inspection. To demonstrate the contribution of the altitude estimation module to the positioning accuracy of the proposed algorithm, the output of the position estimation without this module was used as a comparison and is denoted as LIO in the results. In the simulation experiment, the coefficient  $c_1$  in Eq. (18) was set to zero to evaluate the performance of the altitude estimation module based on infrared rangefinder measurements.

Figs. 8(a) and 8(b) respectively illustrate the 3D trajectory and the trajectory in the X-Y plane of the proposed algorithm and LIO during the stepwise ascent flight route. It can be observed that both methods exhibit a good alignment with the ground truth in the X-Y plane. In terms of altitude, as shown in Fig. 8(c), the inclusion of the altitude estimation module significantly contributes to improving accuracy. As shown in Fig. 8(c), during the four stages of altitude variation, the altitude estimation of the LIO module aligns with the ground truth of altitude changes. However, during horizontal flight, there is a significant deviation from the ground truth.

In long-duration horizontal flight experiments, altitude errors gradually increased. This issue primarily arises from insufficient observational constraints in the vertical direction during purely horizontal motion, which reduces observability and weakens measurement updates along the altitude axis. Consequently, altitude estimation relies more heavily on the IMU-predicted priors, leading to larger accumulated errors. In addition, when the LiDAR beam strikes the ground at large incidence angles, the resulting ground point clouds exhibit high positional uncertainty, further increasing the uncertainty of altitude estimation. The proposed infrared rangefinder-based altitude estimation module alleviates this drift by introducing additional vertical observational constraints.

In the context of bridge inspection, where the majority of the UAV's trajectory is horizontal, the zigzag path designed for this experiment resembles the flight trajectory of a UAV during bridge inspection, with a constant flight altitude. As illustrated in Fig. 9(a), the inclusion of the altitude estimation module does not result in a significant improvement in positioning accuracy within the X-Y plane. However, the vertical trajectory depicted in Fig. 9(b) shows that the altitude estimation error of LIO increases progressively as the horizontal flight distance increases. This leads to a degradation in overall localization accuracy, as observed in the 3D trajectory in Fig. 10. In contrast, the trajectory estimated by the proposed IRF-LIL method exhibits a high degree of alignment with the ground truth.

The experimental results demonstrate that the proposed localization algorithm is capable of providing accurate absolute positioning. During the operation of the LIO algorithm, state propagation is initially performed using IMU measurements to obtain a predicted pose, which is subsequently refined through point-cloud registration based on the prior estimate. The resulting observation is then integrated into the filter for state correction. Due to the relatively high noise levels in IMU measurements, accurate pose estimation can only be maintained over short durations. Consequently, the LIO system relies heavily on LiDAR-based point-cloud registration to achieve long-term and accurate localization.

The accuracy of pose estimates derived from point cloud registration depends on both the positional uncertainty of individual points and the availability of sufficient structural constraints along the direction of motion. During horizontal UAV flight, surrounding objects provide effective constraints in the horizontal direction, enabling accurate estimation of horizontal pose components. In contrast, limited motion in the vertical direction results in reduced constraint residuals during point-cloud registration, thereby weakening the observability and limiting the ability to constrain vertical motion. Furthermore, due to the



**Table 3**

Statistics of the absolute position error between the proposed method, leading SLAM algorithms, and ground truth. Mean represents the average value of APE, and std denotes the standard deviation of APE, LC denotes Loop Closure.

Method	Mean (m)	Std (m)	X-Y mean (m)	X-Y Std (m)	Altitude mean (m)	Altitude Std (m)
IRF-LIL	0.169	0.084	0.105	0.075	0.112	0.081
Fast-LIO2	0.436	0.297	0.117	0.073	0.404	0.309
LIO-SAM	0.477	0.293	0.176	0.102	0.411	0.322
LIO-SAM-LC	0.466	0.402	0.181	0.335	0.388	0.288

large incident angles of laser beams on the ground surface, the positional uncertainty of ground points increases, leading to accumulated errors in altitude estimation [59].

During takeoff and landing, however, ground point clouds provide effective vertical constraints, facilitating accurate estimation of altitude variations. Since LIO estimates poses incrementally, it lacks the ability to correct accumulated drift in absolute altitude. To address this limitation, the proposed altitude estimation module, based on infrared rangefinder measurements, provides a stable estimation of absolute altitude. By incorporating this module into the filter as a direct altitude observation, altitude drift is effectively corrected, resulting in improved positioning accuracy.

To further quantify the precision of the estimated trajectory, the APE was calculated based on the ground truth provided by RTK. Table 2. presents the statistical results of APE, including the mean and variance. The proposed localization algorithm exhibits higher accuracy and stability on the zigzag trajectory compared to the stepwise trajectory. This improvement is attributed to two main factors. First, the zigzag trajectory involves motion in multiple directions within the X-Y plane, providing more effective structural constraints for point cloud registration than the stepwise trajectory, which involves motion primarily along a single direction. Second, the zigzag trajectory results in more LiDAR points being acquired in close proximity to surrounding buildings. Based on the LiDAR noise model, points located nearer to the laser emission center exhibit lower positional uncertainty. These two factors contribute to higher pose estimation accuracy from point cloud registration, which, when integrated into the filter, leads to improved overall localization performance.

In contrast, the IRF-LIL method achieves better altitude estimation accuracy on the stepwise trajectory than on the zigzag trajectory. The zigzag trajectory maintains an approximate and constant altitude of 5 meters above the ground, whereas the stepwise trajectory involves a gradual ascent to the same altitude. According to the infrared rangefinder noise model, greater measurement distances are associated with increased noise, resulting in higher overall range uncertainty for the zigzag trajectory. Moreover, the zigzag trajectory covers a wider area, introducing larger variations in ground flatness within the infrared sensor's field of view. This variation further degrades the performance of the altitude estimation module.

According to Table 2, the IRF-LIL algorithm exhibits smaller standard deviations than the LIO algorithm without the altitude estimation module in both the Stepwise and Zigzag sequences. This indicates that the incorporation of the altitude estimation module enhances not only localization accuracy but also the stability of pose estimation.

The proposed IRF-LIL algorithm demonstrates an average positioning accuracy of less than 0.3 m, with a particularly notable improvement in altitude accuracy, which is approximately 0.1 m. The error is negligible, making it sufficiently accurate to support reliable positioning for the UAV in bridge inspection tasks. The integration of the altitude estimation module effectively mitigates the height drift issue in the LIO.

### 5.3. The real bridge experiment

In this section, the proposed algorithm is evaluated in a real-world bridge inspection scenario and compared with two state-of-the-art LIO methods, Fast-LIO2 and LIO-SAM, both of which are extensively adopted in practical engineering applications, to assess their performance differences and suitability for real-world deployment.

**Table 4**

Comparison of average processing time per scan (unit: ms) between the proposed method and leading SLAM algorithms. Here, FE denotes Feature Extraction and LC denotes Loop Closure.

Method	Total	FE	Registration	LC
IRF-LIL	1.479	\	0.842	\
Fast-LIO2	4.507	\	3.338	\
LIO-SAM	48.056	0.912	36.390	\
LIO-SAM-LC	63.613	1.190	49.144	383.014

The flight trajectory is a zigzag pattern covering the underside of the selected span. A total of 44 hover points were selected along the route. This flight mode is designed to simulate the typical trajectory of a UAV during bridge inspection, specifically when hovering for photo capture. The coordinates of the UAV at the hover point, measured using a total station, were used as the ground truth for the experiment.

Fig. 12(a) presents the trajectories of IRF-LIL, LIO-SAM and Fast-LIO2 algorithm in the 3D space. It can be observed that, due to the lack of global coordinate alignment, the trajectories estimated by the Fast-LIO2 and LIO-SAM algorithms exhibit significant discrepancies compared to the ground truth. By comparison, the trajectory estimated by the proposed IRF-LIL method corresponds closely to the ground truth.

To evaluate the precision of IRF-LIL in comparison with two leading algorithms, the trajectories were transformed into the coordinate system of the total station, which serves as the ground truth, as shown in Fig. 12(b). The Fig. 12(c) and Fig. 12(d) display the trajectories in the XY plane and the vertical direction, respectively. As depicted in Fig. 12(d), Fast-LIO2 and LIO-SAM demonstrate noticeable altitude drift. In contrast, the trajectory estimated by the proposed IRF-LIL method aligns closely with the ground truth in both horizontal and vertical directions.

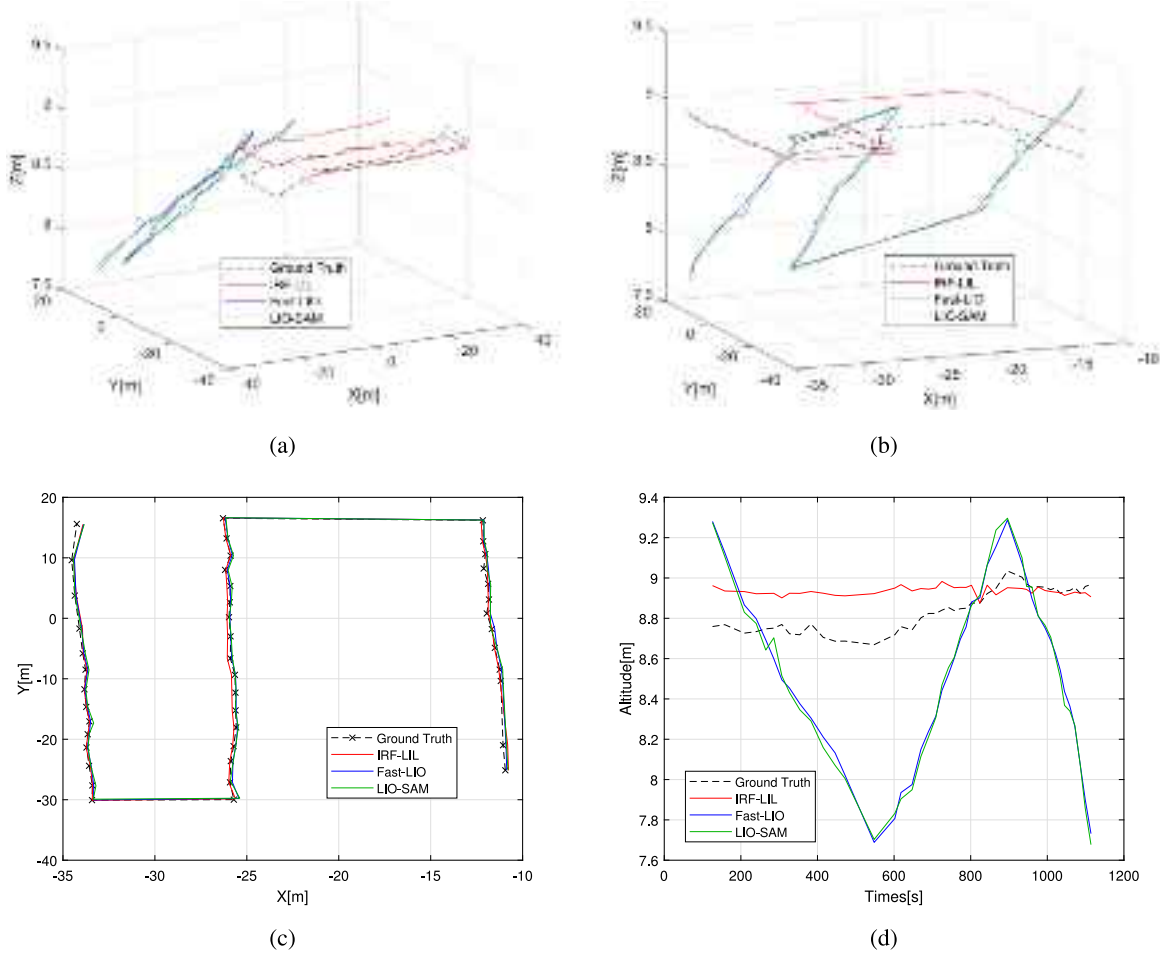
Table 3 presents the APE results of the experiment, which serve as a quantifiable representation of the accuracy. The proposed IRF-LIL algorithm achieves the lowest overall mean position error and the smallest standard deviation, demonstrating high-accuracy and high-stability localization in bridge environments. In the X-Y plane, although IRF-LIL achieves the lowest mean position error, Fast-LIO2 exhibits a smaller standard deviation, indicating that the localization stability of IRF-LIL in the horizontal direction requires further improvement. To improve the balance and comprehensiveness of the evaluation, LIO-SAM with loop closure was also analyzed. The results show that enabling loop closure yields smaller mean errors and standard deviations in altitude estimation, while both the mean error and standard deviation in the horizontal plane increase, with the latter showing a substantial rise. This indicates that point-cloud-based loop closure is less effective in bridge environments. Structural similarity between bridge components, such as piers, increases the probability of false loop closure matches, amplifies loop-closure uncertainty, and ultimately leads to greater horizontal localization error and instability.

LIO-SAM is built upon a factor graph optimization framework that integrates IMU pre-integration factors and LiDAR frame factors to construct residuals, and estimates poses through nonlinear least-squares optimization. For each LiDAR frame, edge and planar features are extracted and registered against a local submap of historical features to form residuals. The optimized pose is subsequently utilized to update the submap using the features extracted from the current LiDAR frame.

**Table 5**

Comparison of performance metrics across different algorithms. APE = Absolute Position Error, Std = Standard Deviation, APTS = Average Processing Time per Scan.

Method	APE (m)	APE Std (m)	Altitude Drift (m)	Robustness	APTS (ms)	Computation hardware cost
LIO-SAM	0.477	0.293	0.411	Medium	48.06	High
Fast-LIO2	0.436	0.297	0.404	Medium	4.51	Medium
IRF-LIL	<b>0.169</b>	<b>0.084</b>	<b>0.112</b>	<b>1.48</b>	<b>High</b>	<b>Low</b>



**Fig. 12.** Trajectories of real bridge experiment test: (a) 3D trajectory, with raw trajectories of leading SLAM algorithms, (b) 3D trajectory, with aligned trajectories of leading SLAM algorithms, (c) X-Y plane trajectory, with aligned trajectories of leading SLAM algorithms, (d) Vertical trajectory.

To improve computational efficiency, LIO-SAM employs a keyframe selection strategy in which only frames with motion exceeding a certain threshold are included in the factor graph. However, in bridge inspection scenarios, the UAV often performs hovering maneuvers for imaging, leading to large temporal gaps between keyframes. As a result, pose propagation between keyframes relies heavily on IMU integration, which introduces significant uncertainty and degrades the convergence of subsequent optimizations. Moreover, the map construction strategy in LIO-SAM, which uses sparse feature points, can lead to structural discontinuities such as layered artifacts, where planar or edge features are incorrectly reconstructed as separate due to pose estimation errors. This degrades the quality of nearest-neighbor search during residual construction and increases the likelihood of incorrect point-to-feature associations.

In contrast, both Fast-LIO2 and the proposed IRF-LIL algorithm bypass feature extraction and use all points from each LiDAR frame to construct residuals. Fast-LIO2 performs point-to-plane registration by fitting a plane to the nearest neighbors of each point. Given the abundance of planar structures in bridge environments, this assumption holds for most points. Fast-LIO2 adopts a tightly coupled filter-based

framework that fuses IMU and LiDAR measurements at each frame, thereby mitigating drift from IMU pose propagation.

The proposed IRF-LIL method employs a NDT-based registration strategy. The map is divided into 3D voxels, and the point distribution within each voxel is modeled using Gaussian distributions. During residual construction, the similarity between each incoming point and the Gaussian distribution assigned to the corresponding voxel is computed, and the pose that maximizes this similarity is estimated. Compared to nearest-neighbor-based registration, the probabilistic approach of NDT method is more robust to irregular point distributions and reduces the likelihood of incorrect correspondences, leading to improved pose estimation accuracy.

A comparison between the real bridge experiment and the simulated environment reveals that the proposed localization algorithm achieves higher accuracy in the X-Y plane under the bridge scenario. This improvement is primarily attributed to the relatively static nature of the bridge environment, which contains no moving objects and features structural elements located in close proximity to the LiDAR sensor. In contrast, the campus environment introduces dynamic objects and exhibits a wider variation in the distances of structural elements, which

contributes to degraded localization accuracy. Regarding altitude estimation, the bridge experiment involves longer flight durations and greater travel distances, while also lacking sufficient vertical structural constraints. These factors result in increased cumulative errors in height estimation. As illustrated in Fig. 12(d), after an extended flight, the altitude estimation of the LIO algorithms converges back to the ground truth. This behavior arises from the scan-to-map registration mechanism used by both LIO methods. When the UAV returns to a location near the starting point, the current LiDAR frame is aligned with the initial map, yielding a height estimate close to the original value. In contrast, the IRF-LIL algorithm incorporates a dedicated height estimation module based on infrared range measurements, which continuously constrains the altitude to remain close to the ground truth throughout the entire flight.

The experimental results demonstrate that the integrated altitude estimation module is effective in real-world bridge scenarios, significantly mitigating altitude drift and validating the accuracy and robustness of the IRF-LIL localization system for UAV positioning in GPS-denied bridge environments.

### 5.3.1. Assessment of computational cost

In order to evaluate the computational efficiency of the proposed IRF-LIL method, a comparison of the APTS was conducted between the proposed method and two leading methods. During the comparison of APTS, the LIO-SAM algorithm demonstrated delays when executed on the onboard computer equipped with the RK3588 chip. To ensure experimental fairness, all algorithm programs were executed on the same laptop. The laptop is powered by a 13th Gen Intel(R) Core(TM) i9-13980HX @ 2.20 GHz processor and is equipped with 32 GB of RAM.

Table 4 shows the APTS for the three algorithms. It can be observed that the APTS of the LIO-SAM algorithm is the highest, which is due to the significant computational resource demands of the feature extraction module (it is labeled as FE in the table) in the front-end processing. Compared to Fast-LIO2, the APTS of the IRF-LIL algorithm is reduced by 67%, indicating that the IRF-LIL algorithm is more efficient in real-time data processing. This performance improvement can be attributed to the point cloud registration strategy implemented in the algorithm, where the nearest-neighbor search within the point clouds is confined to a limited grid area, thereby avoiding extensive range searches. The average point cloud registration time per frame was calculated for both algorithms, with the proposed algorithm demonstrating a 75% improvement over Fast-LIO2. The results demonstrate that the IRF-LIL algorithm exhibits lower computational complexity and higher efficiency, making it suitable for UAV onboard platforms with limited computational resources.

When loop closure is enabled in LIO-SAM, each loop detection requires a relatively long execution time of approximately 383 ms. However, since loop detection is not performed for every LiDAR frame, the increase in APTS is about 15 ms on average. Moreover, the additional constraints introduced between LiDAR frames by loop closure further increase the registration time by approximately 13 ms. This demonstrates that although loop closure enhances long-term consistency, it imposes additional computational overhead, which may reduce the suitability of LIO-SAM for resource-constrained UAV platforms.

### 5.4. Discussion

This section highlights several critical aspects regarding the applicability, robustness, and limitations of the proposed IRF-LIL system.

Robustness against sensor noise and environmental variability also represents a key consideration. In the altitude estimation module, the measurement noise of the infrared rangefinder increases with target distance. To mitigate this effect, the module adaptively adjusts the weighting between infrared-based and LIO-based altitude estimates according to the measured range, thereby reducing noise-induced deviations. The positions of the points in LiDAR point cloud are influenced

by range noise, timestamp misalignment, and scanning distortions, leading to inconsistent measurements of the same static object. To suppress such errors, the proposed method extracts statistical features from multiple points within a voxel to construct residuals for pose estimation, thereby attenuating the influence of individual noisy points and improving localization accuracy. In addition to sensor noise, environmental variability also introduces disturbances. The presence of dynamic objects increases the variance of LiDAR measurements. This study employs scan-to-map registration and statistical feature-based mapping to alleviate the impact of moving objects on localization. When LiDAR beams encounter low-reflectivity surfaces or multipath reflections, range errors may abruptly increase. To address these issues, outlier rejection is incorporated during LiDAR registration by discarding points with excessively large residuals, and pose estimation is refined through Gauss–Newton optimization, collectively enhancing both accuracy and robustness of the system.

Although the altitude estimation module demonstrated satisfactory performance in the experiments presented in this study, it is worth noting that the selection of the empirical coefficient  $c_3$  in Eq. (18) may vary for infrared rangefinder sensors with different noise characteristics and performance specifications. Future work will involve systematic evaluation of  $c_3$  across various infrared rangefinder sensors to investigate the relationship between coefficient selection and sensor characteristics, thereby improving the generalizability of the altitude estimation module.

Importantly, the proposed framework transcends a simple functional aggregation of LiDAR, IMU, and infrared sensors. It introduces methodological contributions at both the fusion and registration levels. In altitude estimation, the infrared rangefinder provides a vertical constraint that directly addresses the long-standing observability deficiency of conventional LIO systems under prolonged horizontal motion. In LiDAR registration, the proposed statistical residual modeling and voxel-based mapping strategy reduces computational complexity relative to geometric feature-based approaches, enabling efficient execution on low-power hardware. Furthermore, the fusion framework, built upon an ESKF, ensures dynamic consistency among LiDAR, IMU, and infrared measurements, while also offering extensibility for additional modalities such as barometers and vision sensors. These design choices emphasize not only robustness and accuracy but also modularity and platform adaptability, distinguishing the system from conventional LIO pipelines.

Error propagation remains an inherent challenge in multi-sensor architectures. Initialization errors originating from RTK-based position and attitude estimation propagate through subsequent modules. Calibration inaccuracies between IMU and RTK introduce fixed position and attitude biases, while attitude misalignment amplifies positional drift with increasing trajectory length. In the LIO module, registration errors accumulate through map construction, progressively degrading pose estimation. Similarly, altitude estimation, derived from recursive integration of range measurements, is also subject to cumulative error. Addressing this issue requires additional global constraints. Future research will investigate visual landmarks, fiducial markers, and lightweight loop-closure mechanisms, as well as adaptive weighting schemes in the ESKF to mitigate single-module error dominance and suppress long-term drift.

To achieve absolute localization, the proposed method employs RTK data for initialization and compensates for measurement failures of the infrared rangefinder during the takeoff phase when no effective targets are present above the sensor. However, in certain scenarios, RTK data may be unavailable throughout the entire operation, in which case the algorithm outputs only relative position estimates. To enhance the generalizability and robustness of the system, future research will focus on reducing reliance on RTK and exploring alternative solutions. Potential directions include the integration of visual data (e.g., monocular or stereo cameras), which can not only provide supplementary

constraints in GPS-denied environments but also support loop closure detection and environmental perception through image features, thereby improving global consistency. In addition, other lightweight sensors such as barometers, fiducial markers, and upward-facing solid-state LiDAR may be incorporated. Specifically, barometers can provide stable atmospheric pressure measurements to impose continuous constraints for absolute altitude estimation; fiducial markers can serve as external reference points, introducing global coordinate information for pose estimation and thereby improving absolute localization accuracy; and upward-facing solid-state LiDAR can enhance altitude estimation in environments with significant terrain undulations or elevation variations. By fusing multi-source information, the accuracy and stability of altitude estimation and absolute localization can be further improved. These enhancements are expected to significantly extend the applicability of the proposed method to diverse bridge inspection tasks, particularly in large-span or electromagnetically challenging environments.

Current experimental results demonstrate that the IRF-LIL algorithm can achieve high-precision localization in real bridge scenarios. However, further validation under more diverse environmental conditions is still required to enhance its generalizability and adaptability. Conducting experiments in real bridge environments often involves complex approval procedures and strict safety requirements, which to some extent limit the scale and diversity of testing. Future research will involve testing in more diverse scenarios, such as adverse weather conditions with reduced visibility (e.g., rain or fog), as well as bridges constructed with different structural materials (e.g., steel, concrete, or composites) and varying levels of structural complexity (e.g., multi-span, large slope, and high-altitude arch bridges). Addressing these challenges through multi-scenario testing will enable a more comprehensive assessment of the proposed method's applicability and provide stronger support for its deployment in practical UAV-based bridge inspection tasks.

Although the IRF-LIL algorithm effectively suppresses altitude drift in UAV localization under complex GPS-denied environments, challenges remain in scenarios involving prolonged horizontal motion or uneven and sloped surfaces for the infrared rangefinder measurements. Future work will consider integrating visual constraints to further enhance vertical observability. For example, incorporating a forward-facing camera to extract visual features from fixed bridge structures and construct a global map, followed by frame-to-map matching, would provide additional vertical constraints for altitude estimation. This integration is expected to further improve localization accuracy and robustness in complex environments.

The feasibility of real-time operation has been demonstrated through deployment on an RK3588 CPU, validating the algorithm's suitability for embedded UAV platforms with constrained computational resources. Nevertheless, dense point clouds from higher-resolution LiDARs or structurally complex environments significantly increase registration complexity, threatening real-time performance. Given the independence of per-point operations, future work will explore parallelization strategies, including GPU-based acceleration, to reduce pose estimation latency and maintain scalability.

## 6. Conclusion

An integrated UAV localization method that fuses infrared rangefinder measurement with LiDAR and IMU is proposed, specifically designed to support UAV-assisted SHM data acquisition in GPS-denied environments. As shown in Table 5, compared with leading localization algorithms such as LIO-SAM and Fast-LIO2, the proposed approach offers advantages for SHM applications in terms of high accuracy, strong robustness, low computational complexity, and cost-effectiveness. The main contribution of this study is the use of a computationally lightweight NDT method for registering sequential LiDAR point cloud scans in the front-end of an LIO framework. A loosely

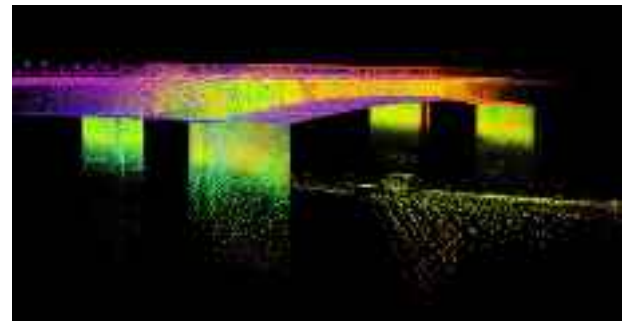


Fig. 13. The LiDAR point cloud map.

coupled ESKF is then employed to fuse the registered LiDAR data with IMU data and RTK data obtained during initialization, resulting in global pose estimation. Additionally, an absolute altitude estimation module based on infrared rangefinder data and RTK positioning is used to correct altitude drift in the pose estimation, thereby achieving a more accurate and robust global position. The method was evaluated and validated through multiple real-world experiments and compared with several state-of-the-art localization methods. From the experimental results, the following conclusions can be derived:

1. In a real bridge environment, when the UAV operates within a horizontal range of 38 m by 66 m, the average positioning accuracy is 0.17 m, with an average processing time per scan of 1.47 s. (Table 3)
2. Compared to the original LiDAR and IMU-based odometry (LIO), the IRF-LIL algorithm with the inclusion of the altitude estimation module achieves a 28% improvement in global localization accuracy and a 57% improvement in altitude localization accuracy. (Table 2)
3. The application of the enhanced statistics-based point cloud registration algorithm significantly reduces computational complexity and improves processing efficiency. (Table 4)
4. The proposed lightweight localization algorithm is designed to operate on computationally constrained small UAV platforms, demonstrating superior performance compared to two leading algorithms. (Table 4)

The primary contributions of this paper can be outlined as follows:

1. The proposed IRF-LIL algorithm provides accurate global position and attitude information for UAVs in GPS-denied environments, ensuring safe and stable flight in complex under-bridge scenarios.
2. An improved lightweight localization algorithm is proposed, which can efficiently operate on UAV onboard computing platforms with constrained computational resources, thereby offering robust support for real-time localization in UAV-based bridge inspection.
3. The proposed localization algorithm, based on ESKF fusion of LiDAR, IMU, and infrared rangefinder measurement, incorporates an altitude fusion module that enhances the observability of pose estimation and reduces drift in the vertical direction. Compared to other LiDAR-inertial SLAM algorithms, it achieves higher localization accuracy and is well-suited for data acquisition tasks.
4. An integrated UAV localization system is proposed, which operates independently of external devices, significantly optimizing the operational complexity of localization in GPS-denied areas and improving operational efficiency.



Future work will aim to further improve the robustness and accuracy of the proposed approach. To address the inspection requirements of large-span bridges, such as sea-crossing bridges, the implementation of loop closure techniques within the SLAM backend can effectively mitigate the accumulation of errors that may occur during operation in extensive environments. Furthermore, the integration of an upward-facing solid-state LiDAR will enhance the system's ability to estimate height more accurately in environments with significant variations in ground elevation. For bridge defects captured by UAVs, maintenance personnel require precise information on the location of defects within the bridge structure. The proposed IRF-LIL algorithm, by generating a point cloud map during the LiDAR point cloud registration process, can effectively assist in the localization of defect positions, as illustrated in Fig. 13.

### CRedit authorship contribution statement

**Anna Kong:** Writing – original draft, Validation, Methodology, Data curation, Conceptualization. **Yulu Wang:** Writing – review & editing, Visualization, Software, Formal analysis, Data curation. **Jiankang Zhao:** Writing – review & editing, Supervision.

### Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

### Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

### Data availability

Data will be made available on request.

### References

- [1] S. Dorafshan, M. Maguire, Bridge inspection: human performance, unmanned aerial systems and automation, *J. Civ. Struct. Heal. Monit.* (2018).
- [2] B.T. Svendsen, G.T. Froseth, O. Øiseth, A. Rønnequist, A data-based structural health monitoring approach for damage detection in steel bridges using experimental data, *J. Civ. Struct. Heal. Monit.* 12 (1) (2022) 101–115, <http://dx.doi.org/10.1007/s13349-021-00530-8>.
- [3] G. Wang, J. Ke, Literature review on the structural health monitoring (SHM) of sustainable civil infrastructure: An analysis of influencing factors in the implementation, *Buildings* 14 (2) (2024) <http://dx.doi.org/10.3390/buildings14020402>.
- [4] Z. Deng, M. Huang, N. Wan, J. Zhang, The current development of structural health monitoring for bridges: A review, *Buildings* 13 (6) (2023) <http://dx.doi.org/10.3390/buildings13061360>.
- [5] L. Sun, Z. Shang, Y. Xia, S. Bhowmick, S. Nagarajiah, Review of bridge structural health monitoring aided by big data and artificial intelligence: From condition assessment to damage detection, *J. Struct. Eng.* 146 (2020) 04020073, [http://dx.doi.org/10.1061/\(ASCE\)ST.1943-541X.0002535](http://dx.doi.org/10.1061/(ASCE)ST.1943-541X.0002535).
- [6] S. Feroz, S. Abu Dabous, UAV-based remote sensing applications for bridge condition assessment, *Remote. Sens.* 13 (9) (2021) <http://dx.doi.org/10.3390/rs13091809>.
- [7] H. Li, Y. Chen, J. Liu, Z. Zhang, H. Zhu, Unmanned aircraft system applications in damage detection and service life prediction for bridges: A review, *Remote. Sens.* 14 (17) (2022) <http://dx.doi.org/10.3390/rs14174210>.
- [8] I.-H. Kim, S. Yoon, J.H. Lee, S. Jung, S. Cho, H.-J. Jung, A comparative study of bridge inspection and condition assessment between manpower and a UAS, *Drones* 6 (11) (2022) <http://dx.doi.org/10.3390/drones6110355>.
- [9] J. Zhou, L. He, H. Luo, Real-time positioning method for UAVs in complex structural health monitoring scenarios, *Drones* 7 (3) (2023) <http://dx.doi.org/10.3390/drones7030212>.
- [10] M. Maboudi, J. Backhaus, I. Mai, Y. Ghassoun, Y. Khedar, D. Lowke, B. Riedel, U. Bestmann, M. Gerke, Very high resolution bridge deformation monitoring using UAV-based photogrammetry, *J. Civ. Struct. Heal. Monit.* (2025) <http://dx.doi.org/10.1007/s13349-025-01001-0>.
- [11] F. Wang, Y. Zou, C. Zhang, J. Buzzatto, M. Liarokapis, E.D. Castillo, J.B.P. Lim, UAV navigation in large-scale GPS-denied bridge environments using fiducial marker-corrected stereo visual-inertial localisation, *Autom. Constr.* 156 (2023) <http://dx.doi.org/10.1016/j.autcon.2023.105139>.
- [12] H. Liang, S.-C. Lee, S. Seo, UAV-based low altitude remote sensing for concrete bridge multi-category damage automatic detection system, *Drones* 7 (6) (2023) <http://dx.doi.org/10.3390/drones7060386>.
- [13] H. Jie, Z. Zhao, Y. Zeng, Y. Chang, F. Fan, C. Wang, K.Y. See, A review of intentional electromagnetic interference in power electronics: Conducted and radiated susceptibility, *IET Power Electron.* 17 (2024) 1487–1506, <http://dx.doi.org/10.1049/pel2.12685>.
- [14] H. Jie, Z. Zhao, H. Li, T.H. Gan, K.Y. See, A systematic three-stage safety enhancement approach for motor drive and gimbal systems in unmanned aerial vehicles, *IEEE Trans. Power Electron.* 40 (7) (2025) 9329–9342, <http://dx.doi.org/10.1109/TPEL.2025.3549964>.
- [15] H. Liu, Q. Long, B. Yi, W. Jiang, A survey of sensors based autonomous unmanned aerial vehicle (UAV) localization techniques, *Complex Intell. Syst.* 11 (8) (2025) 371, <http://dx.doi.org/10.1007/s40747-025-01961-2>.
- [16] I. Jarraya, A. Al-Batati, M.B. Kadri, M. Abdelkader, A. Ammar, W. Boulila, A. Koubaa, Gns-denied unmanned aerial vehicle navigation: analyzing computational complexity, sensor fusion, and localization methodologies, *Satell. Navig.* 6 (1) (2025) 9, <http://dx.doi.org/10.1186/s43020-025-00162-z>.
- [17] Z. Fan, L. Zhang, X. Wang, Y. Shen, F. Deng, LiDAR, IMU, and camera fusion for simultaneous localization and mapping: a systematic review, *Artif. Intell. Rev.* 58 (6) (2025) 174, <http://dx.doi.org/10.1007/s10462-025-11187-w>.
- [18] S. Jiang, Y. Wu, J. Zhang, Bridge coating inspection based on two-stage automatic method and collision-tolerant unmanned aerial system, *Autom. Constr.* 146 (2023) 104685, <http://dx.doi.org/10.1016/j.autcon.2022.104685>.
- [19] A. Phadke, A. Hadimlioglu, An overview of using external beacons with UAV for navigation, localization, and information exchange, in: 2024 6th International Conference on Control and Robotics, ICCR, 2024, pp. 93–100, <http://dx.doi.org/10.1109/ICCR64365.2024.10927537>.
- [20] R. Ali, D. Kang, G. Suh, Y.J. Cha, Real-time multiple damage mapping using autonomous UAV and deep faster region-based neural networks for GPS-denied structures, *Autom. Constr.* (130-Oct) (2021).
- [21] S. Jiang, Y. Wu, J. Zhang, Bridge coating inspection based on two-stage automatic method and collision-tolerant unmanned aerial system, *Autom. Constr.* 146 (2023) 104685, <http://dx.doi.org/10.1016/j.autcon.2022.104685>.
- [22] D. Kang, Y.J. Cha, Autonomous UAVs for structural health monitoring using deep learning and an ultrasonic beacon system with geo-tagging, *Computer-Aided Civ. Infrastruct. Eng.* (2018).
- [23] F.J. Perez-Grau, F. Caballero, L. Merino, A. Viguria, Multi-modal mapping and localization of unmanned aerial robots based on ultra-wideband and RGB-D sensing, in: 2017 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, 2017, pp. 3495–3502, <http://dx.doi.org/10.1109/IROS.2017.8206191>.
- [24] T. Qin, S. Cao, J. Pan, S. Shen, A general optimization-based framework for global pose estimation with multiple sensors, 2019, arXiv preprint [arXiv:1901.03642](https://arxiv.org/abs/1901.03642).
- [25] Z.X. Wang, X.D. Li, Y.Z. Zhang, P.F. Huang, Localization, planning, and control of a UAV for rapid complete coverage bridge inspection in large-scale intermittent GPS environments, *IEEE Trans. Control. Syst. Technol.* 32 (4) (2024) 1357–1369, <http://dx.doi.org/10.1109/tcst.2024.3366695>.
- [26] Z. Liu, F. Zhang, X. Hong, Low-cost retina-like robotic lidars based on incommensurable scanning, *IEEE/ASME Trans. Mechatronics* (2021).
- [27] M. Petrlík, T. Krajník, M. Saska, LiDAR-based stabilization, navigation and localization for UAVs operating in dark indoor environments, in: 2021 International Conference on Unmanned Aircraft Systems, ICUAS.
- [28] J. Zhang, S. Singh, LOAM: Lidar odometry and mapping in real-time, in: *Robotics: Science and Systems*, vol. 2, Berkeley, CA, 2014, pp. 1–9.
- [29] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, D. Rus, Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping, in: 2020 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, IEEE, pp. 5135–5142.
- [30] W. Xu, F. Zhang, FAST-LIO: A fast, robust LiDAR-inertial odometry package by tightly-coupled iterated Kalman filter, *IEEE Robot. Autom. Lett.* 6 (2) (2021) 3317–3324, <http://dx.doi.org/10.1109/lra.2021.3064227>.
- [31] W. Xu, Y. Cai, D. He, J. Lin, F. Zhang, FAST-LIO2: Fast direct LiDAR-inertial odometry, *IEEE Trans. Robot.: A Publ. IEEE Robot. Autom. Soc.* (4) (2022) 38.
- [32] C. Zheng, W. Xu, Z. Zou, T. Hua, C. Yuan, D. He, B. Zhou, Z. Liu, J. Lin, F. Zhu, Y. Ren, R. Wang, F. Meng, F. Zhang, FAST-LIVO2: Fast, direct LiDAR-Inertial-Visual odometry, *IEEE Trans. Robot.* 41 (2025) 326–346, <http://dx.doi.org/10.1109/TRO.2024.3502198>.
- [33] W. Wen, L.-T. Hsu, AGPC-SLAM: Absolute ground plane constrained 3D LiDAR SLAM, *Navig., J. Inst. Navig.* 69 (3) (2022) <http://dx.doi.org/10.33012/navi.527>, Publisher Copyright: © 2022 Institute of Navigation.
- [34] G. Hu, Z. Zhang, A. Armaou, Z. Yan, Robust extended Kalman filter based state estimation for nonlinear dynamic processes with measurements corrupted by gross errors, *J. Taiwan Inst. Chem. Eng.* 106 (2020) 20–33, <http://dx.doi.org/10.1016/j.jtice.2019.10.015>.

- [35] L. Alfeqy, H.E.H. Abdelmunim, S.A. Maged, D. Emad, Kalman filter-based fusion of LiDAR and camera data in bird's eye view for multi-object tracking in autonomous vehicles, *Sensors* 24 (23) (2024) <http://dx.doi.org/10.3390/s24237718>.
- [36] A.I. Mourikis, S.I. Roumeliotis, A Multi-State Constraint Kalman Filter for Vision-Aided Inertial Navigation, *IEEE*, 2007.
- [37] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, J.J. Leonard, Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age, *IEEE Trans. Robot.* 32 (6) (2016) 1309–1332, <http://dx.doi.org/10.1109/TRO.2016.2624754>.
- [38] W. Zhen, S. Zeng, S. Soberer, Robust localization and localizability estimation with a rotating laser scanner, in: 2017 IEEE International Conference on Robotics and Automation, ICRA, *IEEE*, pp. 6240–6245.
- [39] C. Forster, L. Carlone, F. Dellaert, D. Scaramuzza, On-manifold preintegration for real-time visual-inertial odometry, *IEEE Trans. Robot.* 33 (1) (2017) 1–21, <http://dx.doi.org/10.1109/tro.2016.2597321>.
- [40] C. Hertzberg, R. Wagner, U. Frese, L. Schröder, Integrating generic sensor fusion algorithms with sound state representations through encapsulation of manifolds, *Inf. Fusion* 14 (1) (2013) 57–77.
- [41] M. George, S. Sukkarieh, Tightly coupled INS/GPS with bias estimation for UAV applications, in: Proceedings of Australasian Conference on Robotics and Automation, ACRA, 2005, pp. 1–7.
- [42] T. Shan, B. Englot, C. Ratti, D. Rus, Lvi-sam: Tightly-coupled lidar-visual-inertial odometry via smoothing and mapping, in: 2021 IEEE International Conference on Robotics and Automation, ICRA, *IEEE*, 2021, pp. 5692–5698.
- [43] J. Lin, F. Zhang, R 3 LIVE: A robust, real-time, RGB-colored, lidar-inertial-visual tightly-coupled state estimation and mapping package, in: 2022 International Conference on Robotics and Automation, ICRA, *IEEE*, 2022, pp. 10672–10678.
- [44] K. Chen, R. Nemiroff, B.T. Lopez, Direct lidar-inertial odometry: Lightweight lio with continuous-time motion correction, in: 2023 IEEE International Conference on Robotics and Automation, ICRA, *IEEE*, 2023, pp. 3983–3989.
- [45] T. Qin, P.L. Li, S.J. Shen, VINS-mono: A robust and versatile monocular visual-inertial state estimator, *IEEE Trans. Robot.* 34 (4) (2018) 1004–1020, <http://dx.doi.org/10.1109/tro.2018.2853729>.
- [46] S.B. Chen, H. Ma, C.H. Jiang, B.D. Zhou, W.X. Xue, Z.Z. Xiao, Q.Q. Li, NDT-LOAM: A real-time lidar odometry and mapping with weighted NDT and LFA, *IEEE Sensors J.* 22 (4) (2022) 3660–3671, <http://dx.doi.org/10.1109/jsen.2021.3135055>.
- [47] X. Xia, N.P. Bhatt, A. Khajepour, E. Hashemi, Integrated inertial-LiDAR-based map matching localization for varying environments, *IEEE Trans. Intell. Veh.* 8 (10) (2023) 4307–4318, <http://dx.doi.org/10.1109/tiv.2023.3298892>.
- [48] G.H. Xie, Q. Zong, X.W. Zhang, B.L. Tian, Loosely-coupled lidar-inertial odometry and mapping in real time, *Int. J. Intell. Robot. Appl.* 5 (2) (2021) 119–129, <http://dx.doi.org/10.1007/s41315-021-00164-5>.
- [49] X.Y. Ji, S.H. Yuan, P.Y. Yin, L.H. Xie, LIO-GVM: An accurate, tightly-coupled lidar-inertial odometry with Gaussian voxel map, *IEEE Robot. Autom. Lett.* 9 (3) (2024) 2200–2207, <http://dx.doi.org/10.1109/lra.2024.3354616>.
- [50] P.J. Besl, N.D. McKay, A method for registration of 3-D shapes, *IEEE Trans. Pattern Anal. Mach. Intell.* 14 (2) (1992) 239–256, <http://dx.doi.org/10.1109/34.121791>.
- [51] T.X. Shan, B. Englot, LeGO-LOAM: Lightweight and ground-optimized lidar odometry and mapping on variable terrain, in: 25th IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS, in: IEEE International Conference on Intelligent Robots and Systems, 2018, pp. 4758–4765.
- [52] J. Lin, F. Zhang, Loam livox: A fast, robust, high-precision LiDAR odometry and mapping package for LiDARs of small FoV, in: 2020 IEEE International Conference on Robotics and Automation, ICRA, 2020, pp. 3126–3131, <http://dx.doi.org/10.1109/ICRA40945.2020.9197440>.
- [53] M. Magnusson, The Three-Dimensional Normal-Distributions Transform: An Efficient Representation for Registration, Surface Analysis, and Loop Detection (Ph.D. thesis), 2009.
- [54] P. Biber, W. Straßer, The normal distributions transform: A new approach to laser scan matching, in: Proceedings 2003 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2003)(Cat. No. 03CH37453), vol. 3, *IEEE*, pp. 2743–2748.
- [55] D. He, W. Xu, F. Zhang, Symbolic representation and toolkit development of iterated error-state extended Kalman filters on manifolds, *IEEE Trans. Ind. Electron.* 70 (12) (2023) 12533–12544, <http://dx.doi.org/10.1109/TIE.2023.3237872>.
- [56] T.-M. Nguyen, S. Yuan, M. Cao, Y. Lyu, T.H. Nguyen, L. Xie, NTU VIRAL: A visual-inertial-ranging-lidar dataset, from an aerial vehicle viewpoint, *Int. J. Robot. Res.* 41 (3) (2022) 270–280.
- [57] J. Yin, A. Li, T. Li, W. Yu, D. Zou, M2dgr: A multi-sensor and multi-scenario slam dataset for ground robots, *IEEE Robot. Autom. Lett.* 7 (2) (2021) 2266–2273.
- [58] K.S. Arun, T.S. Huang, S.D. Blostein, Least-squares fitting of two 3-D point sets, *IEEE Trans. Pattern Anal. Mach. Intell.* PAMI-9 (5) (1987) 698–700, <http://dx.doi.org/10.1109/TPAMI.1987.4767965>.
- [59] J.H. Chen, H.W. Wang, S. Yang, Tightly coupled LiDAR-inertial odometry and mapping for underground environments, *Sensors* 23 (15) (2023) <http://dx.doi.org/10.3390/s23156834>.