A seamless LiDAR/IMU/RTK fused localization method for UAV-Based bridge inspection

Anbang Liang [1], Yirong Pan [1], Yuelong Huo [1] ,Qingquan Li [1], Baoding Zhou [1], Zhipeng Chen [1,*]

[1] School of Architecture & Urban Planning, Shenzhen University, Shenzhen 518060, China – liangab@szu.edu.cn; 2453044006@mails.szu.edu.cn; 2200432087@email.szu.edu.cn; liqq@szu.edu.cn; bdzhou@szu.edu.cn; chenzp1990@szu.edu.cn
[*] Corresponding author – chenzp1990@szu.edu.cn

Keywords: Simultaneous Localization and Mapping (SLAM), LiDAR/IMU/RTK fusion, Unmanned aerial vehicle (UAV), bridge inspection, seamless localization.

Abstract

The Simultaneous Localization and Mapping (SLAM) technology is fundamental to the autonomous navigation of Unmanned Aerial Vehicles (UAVs) and holds significant value for the realization of UAV-based bridge inspections. However, conventional SLAM methods for UAV face challenges related to low continuity and weak reliability across different scenes, making it difficult to meet the requirements for comprehensive bridge localization and mapping. To address the limitations of existing UAV-based SLAM approaches, we propose a seamless SLAM system that integrates IMU, LiDAR, and RTK. In open scenes (such as the top and sides of a bridge), high-precision absolute localization is achieved by fusing IMU and RTK through an iterative error-state Kalman filter (IESKF). In occluded environments (such as the underside of a bridge), an IMU/LiDAR odometry is used to recursively estimate the UAV's pose. In cross-scene situations (when the UAV passes through a bridge arch), the quality of sensor data is evaluated based on an interactive multi-model (IMM), and an adaptive switching mechanism is employed between two localization modes—IMU/RTK mode and IMU/LiDAR mode—ensuring smooth and seamless multi-source fusion localization even in the presence of sensor signal fluctuations. To validate the effectiveness of our method, extensive tests were conducted on several real-world bridge scenarios. The results show that our method can achieve centimetre-level cross-scene localization accuracy in bridge inspection applications, which indicates its feasibility and effectiveness.

1. Introduction

Bridge inspection is crucial for the maintenance of public transportation, yet traditional manual inspection methods are inadequate in meeting the demands for efficiency and safety. UAV-based inspection technologies (Ri et al., 2024) provide an efficient solution by enabling rapid and comprehensive scanning of bridges. Nevertheless, due to GNSS signal blockage or interference beneath large bridges (Jiang et al., 2021), UAV-based simultaneous localization and mapping (SLAM) methods face challenges in achieving continuous and reliable localization, which hampers inspection efficiency. Therefore, further research on seamless localization techniques is crucial for enhancing the intelligence level of bridge maintenance.

Current UAV-based SLAM methods for infrastructure inspections typically incorporate sensors such as IMU, LiDAR, and RTK. IMU/RTK fused methods can achieve centimetre-level absolute accuracy in open spaces with good RTK signal availability (Han et al., 2017; Farrell et al., 2000). However, GNSS signals may lost under bridge, and IMU errors cannot be effectively corrected. LiDAR-inertial odometry (LIO) integrate point clouds with high-frequency IMU motion (Wu et al., 2023), enables efficient localization in the absence of GNSS. However, in large bridges, LIO errors tend to accumulate due to the sparse point clouds in open areas (Xu et al., 2022). LiDAR/IMU/RTK fused methods can achieve high precision (Li et al., 2023), but most existing LiDAR/IMU/RTK fused frameworks adopt fixed optimization model (Shan et al., 2020; Liu et al., 2024.), which cannot cope with the sensor signal fluctuations in complex environments. As a result, it is difficult to maintain continuity and reliability of localization when UAV traverse large bridges.

To overcome the limitations of existing SLAM methods in cross-scene situations, an intuitive idea is to explore the dynamic fusion strategies for multi-modality data (Meng et al., 2022.). On the one hand, it is necessary to explore how to intelligently select reliable information from different sensors for complementary fusion, thereby improving localization accuracy and robustness. On the other hand, it is essential to investigate how to maintain the continuity of pose estimation through adaptive switching between multi-source data when a single sensor degrades.

(a) A bride located in Shenzhen (b) The UAV under the bridge

Figure 1. UAV cross-scene localization under bridge. Due to sensor signal fluctuations when UAV traverse large bridges, it is difficult to maintain continuity and reliability of localization.

Based on the above ideas of dynamic fusion and adaptive switching, we propose a seamless SLAM framework based on the integration of IMU, LiDAR and RTK, which achieves reliable cross-scene localization performance in several bridges. The main contributions of our work are as follows:

(1) We propose a seamless LiDAR/IMU/RTK fused SLAM method. The method can dynamically switch the localization mode and adjust the multi-sensor fusion weights according to the specific scenes, avoiding the jump in localization accuracy caused by the failure of a single sensor, and ensuring the

continuity and stability of UAV localization trajectory in complex environments.

(2) We design a novel multi-sensor switching strategy, which employs Interactive Multiple Model (IMM) to monitor the residuals of each sensor's state and evaluate its reliability, and then adjusts the confidence of each sensor to achieve adaptive switching between LiDAR/IMU/RTK, thus providing reliable pose estimation in both open and obstructed areas.

(3) Comprehensive and concrete evaluations are conducted under a self-built dataset of complex bridge environments. The experimental results show that our method outperforms state-of-the-art LiDAR/IMU/RTK fused methods.

The rest of this article is organized as follows. Section II presents an overview of UAV-based SLAM methods using multi-source data. Section III elaborates the architecture of our proposed method. Experiments and evaluations are shown in section IV. Finally, we conclude this work in Section V.

2. Related Work

This paper primarily discusses UAV-based SLAM methods involving IMU/GNSS/LiDAR. The IMU/GNSS fused methods suffer from GNSS signal blockage, the IMU/LiDAR fused methods are prone to long-term trajectory drifts, while the current IMU/LiDAR/GNSS fused methods show decreased reliability in cross-scene conditions. Based on the different fusion frameworks, we present a brief review and discussion.

2.1 IMU/GNSS Fused Methods

IMU/GNSS fusion technology can provide centimetre-level positioning services in open areas (Han et al., 2017.) and has been widely used in in the field of drone navigation. However, GNSS or RTK is not usable in environments with physical obstructions (Farrell et al., 2000). While GNSS position can provide short-term corrections for IMU, when GNSS signals are lost for long periods or frequently, the cumulative error of IMU cannot be corrected. Therefore, in cross-scene conditions, due to GNSS signal fluctuations, relying solely on IMU/GNSS combination is insufficient to meet long-term stable positioning requirements.

2.2 IMU/LiDAR Fused Methods

In recent years, IMU/LiDAR fused methods have been widely applied in the fields of robot navigation. LOCUS (Palieri et al., 2020) is a robust odometry centred on LiDAR, and employs a loosely-coupled switching mechanism, allowing it to provide reliable, high-precision localization services even in cases where one or more sensor fail. LOCUS 2.0 (Reinke et al., 2022) significantly improves the system's real-time performance and memory efficiency by introducing an adaptive voxel grid filter and a sliding window map storage data structure. However, due to the relative independence of LiDAR and IMU data processing in the loosely-coupled frameworks, the complementary relationship between the two sensors cannot be fully utilized, which limits further improvements in localization accuracy.

In tightly-coupled methods, FAST-LIO2 (Xu et al., 2022) integrates LiDAR and IMU data tightly using iterative error-state Kalman filtering (IESKF), ensuring the robustness of pose estimation. Moreover, FAST-LIO2 improves its accuracy by directly registering raw LiDAR scans to the map, and enhancing computational speed with the ikd-Tree data structure. Faster-LIO (Bai et al., 2022) uses a sparse voxel-based neighbour structure, iVox (incremental voxels), as the point cloud spatial data structure, replacing the ikd-Tree structure in FAST-LIO2 and effectively reducing the time cost of point cloud registration. The tightly-coupled LIO systems are relatively complex and requires more refined parameter tuning. Furthermore, due to the system's heavy reliance on precise calibration and data synchronization between LiDAR and IMU, it may face additional challenges in practical applications.

When drones are used for large-scale bridge inspections, issues like point cloud sparsity or degradation may lead to increased registration errors, and IMU errors will accumulate over time, especially in the absence of external absolute position information (such as GNSS) for correction. Therefore, in practical applications, the LIO system needs to be adjusted according to the specific scenarios and requirements to better meet the needs of seamless localization.

2.3 IMU/LiDAR/GNSS Fused Methods

To address the limitations of IMU/GNSS or IMU/LiDAR methods, researchers investigate how to integrate all the sensors. (Gao et al., 2024) propose an IMU/LiDAR/GNSS fused methods which integrate two subsystems, GNSS-IMU and LiDAR-IMU. The former subsystem provides initial state estimation and global localization information, while the latter subsystem optimizes state and constructs 3D map. This framework is capable of continuously performing state estimation through iterative Kalman filtering even when some sensors fail. (Chen et al., 2024) propose weighted GNSS/IMU/LiDAR fusion method, which uses a laser error model based on Gaussian Process Regression and combines laser-assisted lateral constraints and non-integrity constraints to improve its accuracy in urban environments, achieving notable accuracy improvements over traditional EKF-based SLAM approaches. (Shen et al. 2024) introduced a real-time SLAM method integrating LiDAR/IMU/GNSS, through joint optimization of LiDAR/IMU and LiDAR/IMU/GNSS, combined with degradation handling and ground point constraints, to achieve drift-free, globally consistent real-time mapping. LIO-SAM (Shan et al., 2020) is a representative work that employs sliding window and factor graph optimization (FGO), and achieves precise pose estimation and mapping. FGO-GIL (Li et al., 2023) is a SLAM system developed for high-precision continuous navigation in urban environments, which utilizes a factor graph optimization framework to tightly fuse GNSS measurements, IMU, and LiDAR data. FGO-GIL employs a keyframe-based nonlinear optimization strategy that effectively uses hybrid heterogeneous sensor data to enhance the accuracy and frequency of state predictions. GLIO (Liu et al., 2024) leverages factor graph optimization to tightly integrate multiple measurement data and achieves global consistency and continuous pose estimation in two stages. The first stage employs sliding window FGO for efficient odometry estimation, while the second stage adopts a scan-to-multi-scan scheme to maintain global consistency and improve robustness to GNSS outliers. Experimental results indicate that GLIO achieves more than 70% improvement in positioning accuracy in urban areas compared to traditional IMU/GNSS methods and LIO methods.

IMU/LiDAR/GNSS fused methods integrates the strengths of multiple sensors to provide high-precision localization solution. However, these methods also face challenges such as system complexity, difficulty in parameter tuning, and low scalability. In cross-scene conditions with fluctuating or failure of multiple sensor signals, IMU/LiDAR/GNSS methods with fixed fusion model struggle to maintain long-term stable trajectory and map. Therefore, to fulfil the need for seamless localization across complex bridge scenarios, there is an urgent need to develop more reliable multi-sensor fused SLAM frameworks with environmental adaptability.

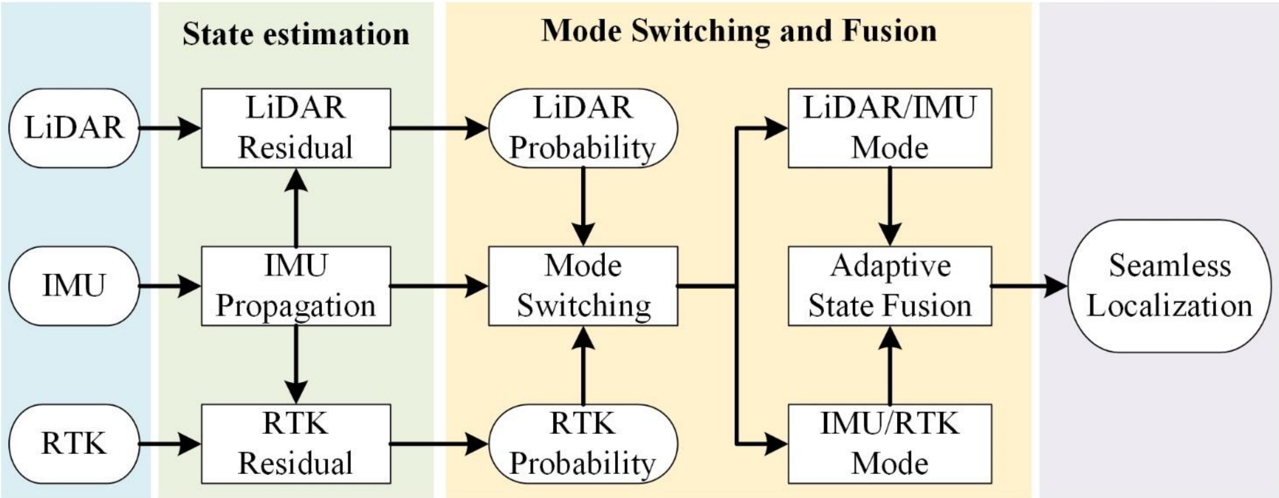


Figure 2. The seamless LiDAR/IMU/RTK localization method.

3. Method

This paper proposes a seamless UAV localization method based on IMU/LiDAR/RTK fusion: In open areas, high-precision localization is achieved by integrating all sensor data through an iterative error-state Kalman filter (IESKF); in obstructed environments, IMU/LiDAR odometry recursively estimates pose; when transitioning across different scenarios, the quality of sensor data is evaluated using an interactive multi-model (IMM), and adaptive switching between IMU/RTK and IMU/LiDAR localization modes ensures smooth, seamless multi-source fusion localization during sensor measurement fluctuations under cross-scene conditions.

This method is suitable for a variety of complex environments, offering stable localization services in both open and obstructed areas, preventing abrupt changes and degradation in localization accuracy. The method also supports flexible configuration, allowing for adjustment of parameters and fusion weights based on specific scenarios and requirements, meeting the needs of diverse application. The multi-sensor fusion strategy enhances the system's fault tolerance, ensuring the continuity and stability of localization results even in the event of sensor failure, thus improving the UAV's autonomous inspection capabilities in complex environments.

The process of our IMU/LiDAR/RTK seamless localization method is shown in Figure 2, which takes LiDAR point clouds, IMU measurements, and RTK position information as inputs, and then evaluates the quality of different sensors using IMM model, enabling seamless switching between two localization modes: IMU/RTK mode in open areas with GNSS signals and IMU/LiDAR mode in GNSS-denied regions, thus achieving continuous and reliable localization and mapping across all scenarios of the bridge.

3.1 State Estimation

The state estimation process is based on the Iterative Error-State Kalman Filter (IESKF), which utilizes IMU for propagation and LiDAR/RTK measurements for state update. The system state $\mathbf{X}$ is defined as follows:

$$\mathbf{X} \triangleq \left[{}^{G}\mathbf{R}_I^T,\ {}^{G}\mathbf{P}_I^T,\ {}^{G}\mathbf{v}_I^T, \mathbf{b}_a^T, \mathbf{b}_\omega^T,\ {}^{G}\mathbf{g}^T \right]^{\mathrm{T}} \tag{1}$$

where ${}^{G}\mathbf{R}_I^T$, ${}^{G}\mathbf{P}_I^T$, ${}^{G}\mathbf{v}_I^T$ denote the IMU attitude, position and velocity in the global frame, ${}^{G}\mathbf{g}^T$ is the gravity vector in the global frame, $\mathbf{b}_a^T, \mathbf{b}_\omega^T$ are IMU biases.

State Propagation: We take the first IMU frame (I) as the global frame (G). The propagation is performed upon the arrival of an IMU measurement.

$$\widehat{\mathbf{X}}_{k+1} = \widehat{\mathbf{X}}_k \boxplus \left(\mathbf{f}\left(\widehat{\mathbf{X}}_k, \mathbf{u}_k, \mathbf{w}_k\right) \cdot \Delta t\right), \mathbf{X} \in \mathcal{M}_s \tag{2}$$

$$\mathbf{u} \triangleq [\mathbf{a}^{\mathrm{T}} \quad \boldsymbol{\omega}^{\mathrm{T}}]^{\mathrm{T}},\ \boldsymbol{w} \triangleq \left[\mathbf{n}_a^{\mathrm{T}} \quad \mathbf{n}_\omega^{\mathrm{T}} \quad \mathbf{n}_{\mathbf{b}_a}^{\mathrm{T}} \quad \mathbf{n}_{\mathbf{b}_\omega}^{\mathrm{T}}\right]^{\mathrm{T}} \tag{3}$$

where $\mathbf{f}$ denotes the state transition model, $\mathbf{u}$ denote the IMU measurements, $\mathbf{w}$ is the noises. $\boxplus$ and $\boxminus$ denotes the plus and minus operation on a manifold $\mathcal{M}_s$.

Residual Computation: The LiDAR residual is defined as the distance between the measured point to the nearest plane in the map. The residuals $\mathbf{z}_{L_{kj}}^\tau$ of point ${}^{L}\boldsymbol{p}_j$ and its noise $\mathbf{r}_{L_{kj}}$ are:

$$\mathbf{z}_{L_{kj}}^\tau = \mathbf{h}_{L_k}\left(\widehat{\mathbf{X}}_k^\tau, \mathbf{0}\right) \triangleq {}^{G}\boldsymbol{\mu}_j^T\left({}^{G}\widehat{\mathbf{T}}_{L_k}^\tau\ {}^{I}\widehat{\mathbf{T}}_L^\tau\ {}^{L}\boldsymbol{p}_j - {}^{G}\boldsymbol{q}_j \right) \tag{4}$$

where $\boldsymbol{\mu}$ denotes the normal vector of corresponding plane. τ is the step of iterative state update process of IESKF. ${}^{G}\widehat{\mathbf{T}}_{L_k}^\tau$ denotes the transformation matrix for projecting the current point cloud frame to the global map, and ${}^{I}\widehat{\mathbf{T}}_L^\tau$ denotes the extrinsic matrix between LiDAR and IMU.

The RTK residual $\mathbf{z}_{R_k}^\tau$ is defined as the difference between the IMU predicted position ${}^{G}\mathbf{P}_I$ and the RTK position ${}^{N}\mathbf{P}_R$:

$$\mathbf{z}_{R_k}^\tau = \mathbf{h}_{R_k}\left(\widehat{\mathbf{X}}_k^\tau, \mathbf{0}\right) \triangleq {}^{G}\mathbf{P}_I - {}^{G}\mathbf{T}_N\ {}^{N}\mathbf{P}_R \tag{5}$$

where ${}^{G}\mathbf{T}_N$ denotes the transformation matrix for projecting the current RTK position to the global coordinates.

State Update: Combining the prior distribution of $\widehat{\mathbf{X}}_k$ with the LiDAR/RTK measurement model from (4-5) yields a posteriori distribution of the state $\mathbf{X}_k$, and the maximum a posteriori estimate (MAP) are:

$$\min\left(\left\|\mathbf{X}_k \boxminus \widehat{\mathbf{X}}_k\right\|^2 + \textstyle\sum_{j=1}^{m}\left\|\mathbf{r}_{L_{kj}}\right\|^2 + \left\|\mathbf{r}_{R_k}\right\|^2\right) \tag{6}$$

The posteriori probability distribution $\widehat{\mathbf{X}}_k^{\tau+1}$ is described as:

$$\bar{X}_{k+1} = \widehat{X}_k^{\tau+1} = \widehat{\mathbf{X}}_k^\tau \boxplus \left(\mathbf{K}\mathbf{z}_m^\tau - (\mathbf{I} - \mathbf{K}\mathbf{H})(\mathbf{J}^\tau)^{-1}\right)\left(\widehat{\mathbf{X}}_k^\tau - \widehat{\mathbf{X}}_k\right) \tag{7}$$

where $\mathbf{K}$ denotes the Kalman gain and $\mathbf{H}$ denotes the observation matrix of all sensor data. For the detailed derivation process, please refer to the literature (Xu et al., 2022). When the iterations converge, the optimal state of IESKF at stage k is updated.

3.2 IMU/RTK localization mode for open scenes

IMU provides high-frequency motion information unaffected by external environments, which can be used for short-term pose propagation, but IMU errors will accumulate over time. RTK provides low-frequency, high-precision long-term localization benchmarks when the signal is good. These two sensors are highly complementary, so in open areas with good RTK signal, IMU/RTK fused localization mode is used to provide precise global localization.

3.3 IMU/LiDAR localization mode for occluded scenes

In occluded areas with poor RTK signal, the IMU/LiDAR odometry is used to propagate the motion state.

To enhance the reliability of point cloud registration in complex environments, we use a frame-to-frame registration method introduced in our prior work (Chen et al., 2022b). The proposed method uses IMU propagation to provide initial guess for frame-to-frame ICP pre-registration, thus avoiding the risk of degradation that occurs in frame-to-map association methods in homogeneous structured scenes, and achieving high-precision frame-to-frame point cloud registration.

3.4 Mode switching and fusion

We design a novel localization mode switching strategy for cross-scene seamless localization. When the RTK signal is good, RTK/IMU is used as the primary localization mode; whereas in the absence of RTK signals, LiDAR/IMU mode is prioritized; During transitions across different scenarios, as the sensors signal quality fluctuates, the fusion weights of different sensors is adaptively adjusted to ensure smooth switching between the two localization modes, thus ensuring continuous localization of the UAV in the bridge environment.

The switching and fusion model is implemented based on the Interactive Multi-Model (IMM) (Kirubarajan et al., 2003) and Bayesian probability theory. IMM model first adopts parallel sub-modes to process different sensor data separately, then evaluates the mode probability according to its residual and adjust the weights of mode; finally, the parallel modes are fused using the adaptive weight, which leads to smooth switching and adaptive fusion between LiDAR/IMU/RTK.

At the module switching stage, the likelihood of different mode is calculated using the residual $\mathbf{r}_k$ and its covariance $\mathbf{S}_k$. The likelihood function is related to the observation residuals, and the fluctuations in residuals in complex environments reflect changes in sensor reliability. The form of likelihood function is:

$$l_k^{(i)} = p\left[\mathbf{z}_k^{(i)} | \mathbf{M}_k^{(i)}, \mathbf{Z}_k^{(i)}\right] \sim \mathcal{N}\left(\mathbf{r}_k^{(i)}; \mathbf{0}, \mathbf{S}_k^{(i)}\right) \tag{8}$$

where $\mathbf{M}$ denotes the localization mode, $\mathbf{Z}$ denote the measurements. Based on Bayesian probability theory, we can derive the likelihood function:

$$l_k \triangleq= \frac{1}{\sqrt{\det(2\mathbf{S}_k)}} \exp\left[-\frac{1}{2}\mathbf{r}_k^T \mathbf{S}_k^{-1} \mathbf{r}_k\right] \tag{9}$$

Besides the likelihood function, we also calculate the mode transition probability π based on the results of each mode in the previous stage. Since the reliability of each mode differs, directly propagating the states with low reliability will lead to a decrease in localization accuracy. To address this issue, we define the transition probability π, which denotes the probability of mode j switching to mode i in the Markov chain, and represents the strength of the constraints between different modes. If the credibility of mode j is high, the corresponding transition probability from j to i becomes larger, thus enabling interaction between modes. The transition probability is calculated as:

$$\pi_{k+1}^{(j)|(i)} = \frac{\pi_k^{(j)|(i)} \exp\left[\sigma_{k+1}^{(i)} - \sigma_k^{(i)}\right]}{\sum_{j=1}^{N} \pi_k^{(j)|(i)} \exp\left[\sigma_{k+1}^{(i)} - \sigma_k^{(i)}\right]} \tag{10}$$

Algorithm 1: The Mode switching and fusion model

Input: The state $\bar{\mathbf{X}}_k^{(i)}$ of each mode; current measurement $\mathbf{Z}_k^{(i)}$.
Begin
1: Calculate the likelihood function of each mode $l_k^{(i)}$
2: Update the transition probability $\pi_k^{(j)|(i)}$ between modes
3: Calculate the priori interaction weights $\mu_k^{(i)}$
4: Combine $\pi_k^{(j)|(i)}$ and $\mu_k^{(i)}$, and get the fusion weight $\sigma_k^{(i)}$.
5: Adaptive fusion of modes to obtain an optimal state $\bar{\mathbf{X}}_k^*$
End
Output: The optimal state $\bar{\mathbf{X}}_k^*$.

Figure 3. The UAV localization system.

Based on the transition probability, we can compute the interaction weights μ, which can be serve as priori knowledge to measure the credibility of localization modes:

$$\mu_{k+1}^{(i)} = \sum_{j=1}^{N} \pi_k^{(j)|(i)} \sigma_k^{(j)} \tag{11}$$

By comprehensively considering the likelihood function l and prior weight μ, we can calculate the posterior fusion weight σ of each mode. The likelihood reflects the reliability assessed in the current stage. The prior weight reflects the prior reliability of each sensor in the previous stage. By integrating them, the fusion weight of mode (i) is updated:

$$\sigma_k^{(i)} = \mu_k^{(i)} l_k^{(i)} \Big/ \sum_{j=1}^{M} \mu_k^{(j)} l_k^{(j)} \tag{12}$$

At the fusion stage, the IMU/LiDAR mode and IMU/RTK modes are fused according to the weights:

$$\bar{\mathbf{X}}_k^* = \sum_{i=1}^{N} \sigma_k^{(i)} \bar{\mathbf{X}}_k^{(i)} \tag{13}$$

where $\bar{\mathbf{X}}_k^*$ denotes the optimal localization state.

When the mode weight is less than the threshold (e.g., RTK in occluded environments), this mode is not involved in state estimation. When the data quality fluctuates, IMM model can dynamically adjust the fusion weight of modes, realizing the seamless fusion of LiDAR/IMU/RTK across scenes, which can effectively improve the reliability of localization.

The detailed methodology of this switching and fusion model is summarized in Algorithm 1.

4. Results

4.1 UAV localization system

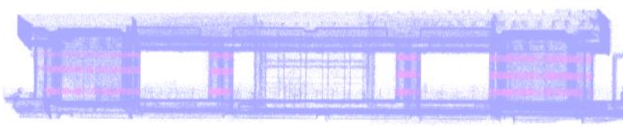
We designed and constructed a UAV localization system integrated with LiDAR/IMU/RTK, as shown in Figure 3. We used a DJI M350 drone as platform. The system is equipped with a Livox Mid360 solid LiDAR, capable of providing point clouds and 6-axis IMU measurement data. In addition, real-time RTK data is read through the DJI M350 SDK.

(a) Building scenario (b) The first bridge (c) The second bridge

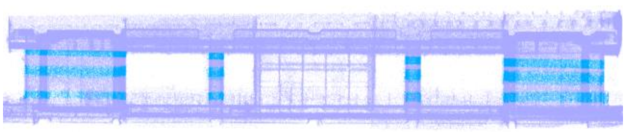
Figure 4. Experimental scenarios.

(a) Z+F 5010 fixed station scanner (b) LiGrip H120 scanner

Figure 5. Equipment used for collecting ground-truth



(a) Ground-truth of laser scanner



(b) Map of our method

Figure 6. Map comparisons at occluded building scene

4.2 Experimental scenarios

4.2.1 Building scenarios: We chose a Building of Shenzhen University as the experimental scenario, as shown in Figure 4 (a). The building consists of three areas, with an open space on the periphery and a GNSS-denied space inside. The centre area of the building can receive fluctuating RTK signals. The experimental site allows UAVs to shuttle between open and occluded areas, which can provide ideal conditions for testing the seamless LiDAR/IMU/RTK switching. In open space, a high-precision INS system is used to obtain ground-truth trajectory. In GNSS-denied area, the point cloud acquired by Z+F 3D laser scanner is used as ground-truth.

4.2.2 Bridge scenarios: We selects two bridges in Shenzhen as experimental scenarios, as shown in Figure 4 (b-c). We use our UAV system to collect cross-scene data by flying along the sides and underneath the bridge for localization and mapping, and test the localization performance of our method with state-of-the-art multi-sensor fusion SLAM methods. We utilize a LiGrip H120 handheld laser scanner to collect point cloud map as ground-truth.

Methods	ATE (m)
DLIO	0.157
Faster-LIO	0.179
Fast-LIO2	0.152
LIO-SAM	0.103
Ours	**0.087**

Table 1 ATE (m) of SLAM Methods

Horizontal cross-section	RMS (m)
cross-section 1	0.066
cross-section 2	0.064
cross-section 3	0.072
Average	0.067

Table 2 RMS (m) of horizontal building cross-sections

Vertical cross-section	RMS (m)
cross-section 1	0.043
cross-section 2	0.048
cross-section 3	0.043
Average	0.045

Table 3 RMS (m) of vertical cross-sections

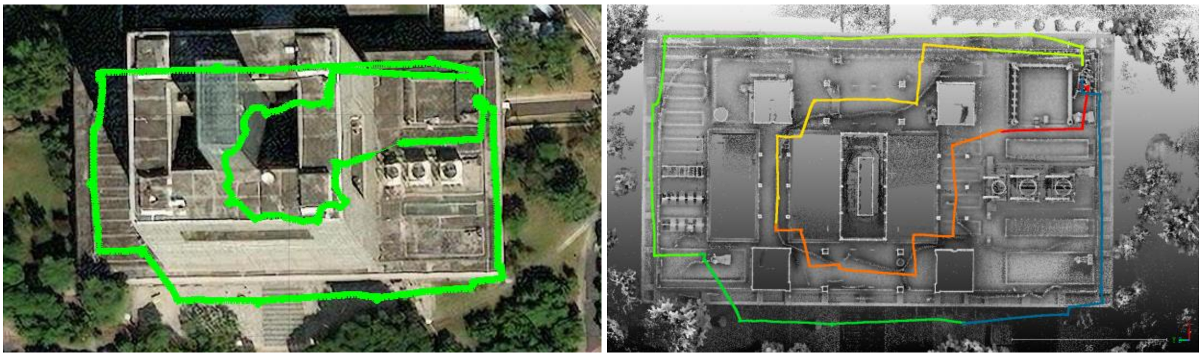
4.3 Accuracy evaluation at building scenarios

4.3.1 Accuracy evaluation at open areas: When the UAV is at open space where GNSS is available, the trajectory collected by INS system is used as ground-truth. We choose the Absolute Trajectory Error (ATE) as metric, and we compare our method with LIO odometry DLIO (Chen et al., 2022a), Faster-LIO (Bai et al., 2022) and Fast-LIO2 (Xu et al., 2022), IMU/LiDAR/GNSS fused method LIO-SAM (Shan et al., 2020), as shown in Table 1. Our method shows optimal performance in terms of ATE, indicating that our method's superiority in maintaining global trajectory consistency.

4.3.2 Accuracy evaluation at occluded areas: When the UAV enters GNSS-denied areas, the INS system cannot collect the ground-truth data. To evaluate the performance of our method in occluded areas, we utilize point cloud map from a fixed-station 3D laser scanner as ground-truth, and then evaluate the accuracy by calculating the root mean square error (RMSE) of ICP registration between the ground-truth and the mapping results of our method.

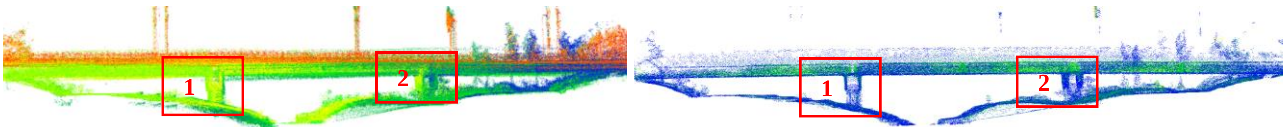
We first collect the ground-truth map in the building scene using Z+F 5010 scanner, as shown in Figure 5(a), which has a measurement range of 0.3 to 187.3 meters and a distance resolution of 0.1 millimetres, with a registration accuracy within 0.01 meters. We then select fixed objects (non-glass facades) in the test area for accuracy evaluation. These fixed objects are stable over time, and their 3D shapes can be clearly captured in both data, so they can serve as suitable benchmarks for accuracy evaluating.

We extract point cloud cross-sections in the horizontal and vertical directions and calculate their corresponding registration errors. The results are shown in Table 2,3 and Figure 6, the average RMS of the horizontal cross-sections is 0.067 m, and the average RMS for the vertical cross-sections is 0.045 m, which achieve centimetre-level mapping accuracy in GNSS-denied areas, indicating the effectiveness of our method.



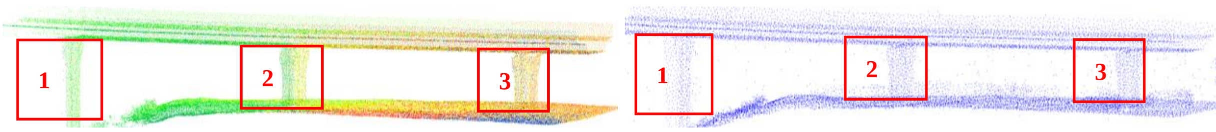
(a) IMU/RTK Trajectory (b) Our Method

Figure 7. Trajectory comparisons in building cross-scene experiment



(a) Ground-truth collected by 3D laser scanner (b) Point cloud map of our method

Figure 8. Map comparisons at the first bridge scene



(a) Ground-truth collected by 3D laser scanner (b) Point cloud map of our method

Figure 9. Map comparisons at the second bridge scene

Cross-sections	RMS (m)
bridge pier 1	0.079
bridge pier 2	0.075
Average	0.077

Table 4 RMS (m) of cross-sections at the first bridge

Cross-sections	RMS (m)
bridge pier 1	0.076
bridge pier 2	0.083
bridge pier 3	0.077
Average	0.079

Table 5 RMS (m) of cross-sections at the second bridge

4.4 Performance of cross-scene localization

To validate the switching capability of our method in cross-scene situations, a qualitative analysis was carried out in the building scenario. The comparison between the IMU/RTK trajectory and the trajectory of our method is shown in Figure 7. The trajectory using only the IMU/RTK mode drifts when the RTK signal quality decreases in the occluded areas, while our method maintains a continuous global trajectory by switching to the IMU/LiDAR mode, which demonstrates its better reliability. If the IMU/RTK and IMU/LiDAR mode is not adjusted using the switching model, direct fusion of erroneous RTK positions can lead to localization failures. Additionally, the accuracy of our method in this scene has been discussed in detailed in Section 4.3. Overall, our method can effectively perform adaptive switching and fusion of localization modes when RTK signals fluctuate, thereby achieving precise and reliable seamless localization across different scenes.

4.5 Applications at bridge scenarios

We select two bridges in the Longgang District of Shenzhen University as test scenarios. The first bridge is 125 meters in length and 27 meters in width. The second bridge is 76 meters in length and 10 meters in width. Figures 8 and 9 qualitatively show the mapping results of our method during cross-scene flight.

We first collect the ground-truth map in the bridges scenes using LiGrip H120 handheld scanner, as shown in Figure 5(b), which has a measurement range of 120 meters and a relative mapping accuracy within 1 centimetre.

Furthermore, we extracted the bridge piers from the point cloud and evaluated the mapping accuracy. The selected bridge piers are identified with red rectangles in Figures 8 and 9. The UAV needs to fly back and forth under the bridge when measurement the bridge piers, so the mapping accuracy of these objects can reflect the cross-scene performance of our method.

We evaluate the mapping accuracy by calculating the root mean square error (RMSE) of ICP registration between ground-truth and the map of our method. The results are shown in Tables 4 and 5. Our method achieves mapping accuracies of 0.077m and 0.079m for the two bridges, demonstrating good consistency and precision, indicating that our method can achieve centimetre-level seamless localization and mapping in bridge scenarios.

5. Conclusion

Continuous and seamless UAV localization in complex environments is essential for improving the intelligence of bridge maintenance and management. However, existing UAV-based SLAM methods face challenges in maintaining a continuous trajectory across entire bridge due to limited sensor signals in occluded area. To address the issues of low continuity and weak reliability of existing SLAM methods for bridge cross-scene situations, we propose a seamless UAV SLAM framework based on the integration of IMU, LiDAR and RTK information. We design a novel multi-sensor switching and fusion strategy based on Interactive Multi-Model, which enables seamless switching between IMU/LiDAR mode and IMU/RTK mode, thus providing continuous localization services in both open and obstructed areas. Experimental results in real-word demonstrate that our method can achieve centimetre-level accuracy, outperforming state-of-the-art LiDAR-inertial odometry and LiDAR-IMU-RTK fused methods. The proposed method can effectively improve the fault tolerance of SLAM system, ensuring the continuity of the trajectory even if a single sensor fails, which is valuable for enhancing the autonomous navigation capabilities of UAV.

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.

Acknowledgements

This study was supported in part by the National Key Research and Development Program of China with project No.2022YFB3904601, and the Shenzhen Higher Education Stability Support Program with project No.20220810184822001.

References

Bai, C., Xiao, T., Chen, Y., et al., 2022. Faster-LIO: Lightweight Tightly Coupled Lidar-Inertial Odometry Using Parallel Sparse Incremental Voxels. *IEEE Robotics and Automation Letters*, vol. 7, no. 2, pp. 4861-4868.

Chen, H., Sun, R., Cheng, Q., et al., 2024a. GNSS/IMU/LO integration with a new LO error model and lateral constraint for navigation in urban areas. *Satellite Navigation*, 5, 30.

Chen, Z., Li, Q., Li, J., et al., 2022b. IMU-Aided Registration of MLS Point Clouds Using Inertial Trajectory Error Model and Least Squares Optimization. *Remote Sensing*, 14(6):1365.

Chen K, Nemiroff R, Lopez B T. 2022. Direct LiDAR-Inertial Odometry: Lightweight LIO with Continuous-Time Motion Correction[C]// IEEE International Conference on Robotics and Automation (ICRA): 3983-3989.

Farrell, J, A., Givargis, T, D., Barth, M, J., 2000. Real-time differential carrier phase GPS-aided INS. *IEEE Transactions on Control Systems Technology*, vol. 8, no. 4, pp. 709-721.

Gao, J., Sha, J., Li, H., Wang, Y., 2024. A Robust and Fast GNSS-Inertial-LiDAR Odometry With INS-Centric Multiple Modalities by IESKF. *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1-12

Han, H., Wang, J., Wang, J, et al., 2017. Reliable partial ambiguity resolution for single-frequency GPS/BDS and INS integration[J]. *GPS Solutions*., 21, 251–264.

Jiang, H., Shi, C., Li, T., et al., 2021. Low-cost GPS/INS integration with accurate measurement modelling using an extended state observer. *GPS Solutions*., 25, 17.

Kirubarajan, T., Bar-Shalom, Y., 2003. Kalman filter versus IMM estimator: when do we need the latter. *IEEE Transactions on Aerospace and Electronic Systems*, vol. 39, no. 4, pp. 1452-1457.

Li, X., Yu, H., Wang, X., et al., 2023. FGO-GIL: Factor Graph Optimization-Based GNSS RTK/INS/LiDAR Tightly Coupled Integration for Precise and Continuous Navigation. *IEEE Sensors Journal*, vol. 23, no. 13, pp. 14534-14548.

Liu, X., Wen, W., Hsu, L, T., 2024. GLIO: Tightly-Coupled GNSS/LiDAR/IMU Integration for Continuous and Drift-Free State Estimation of Intelligent Vehicles in Urban Areas. *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 1, pp. 1412-1422.

Palieri, M., Morrell, B., Thakur, A., et al., 2020. Locus: A multi-sensor lidar-centric solution for high-precision odometry and 3d mapping in real-time. *IEEE Robotics and Automation Letters*, 6(2): 421-428.

Reinke, A., Palieri, M., Morrell, B., et al., 2022. Locus 2.0: Robust and computationally efficient lidar odometry for real-time 3d mapping. *IEEE Robotics and Automation Letters*, 7(4): 9043-9050.

Ri, S., Ye, J., Toyama, N., Ogura, N., 2024. Drone-based displacement measurement of infrastructures utilizing phase information. *Nature Communications.*, 15(1), 395.

Shan, T., Englot, B., Meyers, D., et al., 2020. LIO-SAM: Tightly-coupled Lidar Inertial Odometry via Smoothing and Mapping. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 5135-5142.

Shen, Z., Wang, J., Pang, C., et al., 2024. A LiDAR-IMU-GNSS fused mapping method for large-scale and high-speed scenarios. *Measurement*, vol. 225pp. 1-11,

Wu, W., Li, J., Chen, C., Yang, B., et al., 2023. AFLI-Calib: Robust LiDAR-IMU extrinsic self-calibration based on adaptive frame length LiDAR odometry. *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 199, pp. 157-181.

Xu, W., Zhang, F., 2021. FAST-LIO: A Fast, Robust LiDAR-Inertial Odometry Package by Tightly-Coupled Iterated Kalman Filter. *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 3317-3324.

Xu W, Cai Y, He D, et al., 2022. FAST-LIO2: Fast Direct LiDAR-Inertial Odometry[J]. *IEEE Transactions on Robotics*, vol. 38, no. 4, pp. 2053-2073.