2023 IEEE 26th International Conference on
Intelligent Transportation Systems (ITSC)
24-28 September 2023. Bilbao, Bizkaia, Spain

Roadside Infrastructure assisted LiDAR/Inertial-based Mapping for Intelligent Vehicles in Urban Areas

Feng Huang*[1], Hang Chen*[2], Alpamys Urtay[2], Dongzhe Su[2], Weisong Wen†[1] and Li-Ta Hsu[1]

***Abstract*— Vehicle-infrastructure cooperation is an emerging technology towards fully autonomous driving. LiDAR/Inertial odometry (LIO) based on onboard sensors can provide precise state estimation and mapping locally but is subjected to drift accumulation over time. Limited by the level of intelligence from a single intelligent vehicle, the cellular vehicle-to-everything (C-V2X) opened a new window for the realization of fully autonomous driving. *How can the roadside intelligent infrastructure assist intelligent vehicles in mapping and localization?* In this paper, we propose a roadside infrastructure (RSI) assisted LIO for reliable odometry and mapping, which benefits from the global constraint provided by RSIs. Specifically, RSI estimates the coarse vehicle state in the RSI-based LiDAR point cloud using a deep learning-based method. Based on the initial guess of the positioning state and the point cloud provided by the RSI, our system extracts a local LiDAR map that can refine registration with the RSI LiDAR point cloud. The accurately registered state then served as the global constraint in the graph-based optimization. To evaluate our approach, we collect the multi-view RSI and vehicle sensor data in the Hong Kong representative C-V2X testbed. Experimental results show that the absolute positioning accuracy of the proposed RSI-assisted LIO was significantly improved by 84.3%. To benefit the research community, the data of this work is available on our project page[3].**

I. Introduction

Recent advancements in autonomous driving show the great potential for future intelligent transportation systems (ITS) [1]. Reliable mapping and localization are of great importance for the realization of fully autonomous driving. Three-dimensional (3D) light detection and ranging (LiDAR) provides accurate and dense 3D point clouds, which is extensively studied [2, 3] for navigation tasks in autonomous systems. An inertial measurement unit (IMU) provides high-frequency acceleration and orientation of the ego-vehicle, which is essential for sensor fusion. Integration of LiDAR and IMU has been widely utilized to provide mapping and localization solutions [4] for intelligent vehicles. The work in [5] proposed tightly coupled LiDAR and IMU to provide state estimation and IMU bias correction simultaneously. However, the performance of LiDAR-based odometry can be affected by complex urban scenarios such as dynamic objects [6] and structure-less tunnels [7], which result in accumulation drift over time.

To mitigate the error accumulation by the LiDAR-based odometry, numerous methods [8, 9] were proposed to correct the accumulated drift. One common way is to loosely integrate Global Navigation Satellite System (GNSS) with LIO [10] using factor graph optimization to provide global constraint poses. However, the performance of such methods is hard to guarantee in urban areas because GNSS is significantly degraded due to the Non-Line-of-Sight (NLOS) and multipath [11] in complex urban areas. The work in [12] explored loop closure factor constraints to reduce drift accumulation. Unfortunately, the revisiting route is not always available for mapping tasks in urban areas. Another team [13] proposed the use of globally consistent LiDAR matching to minimize the residuals between frames over the entire maps. But it is still impacted by unexpected outliers such as dynamic objects. In short, the existing LiDAR-based mapping and localization solution based on the onboard sensors from a single vehicle is still challenged in complex urban scenarios.

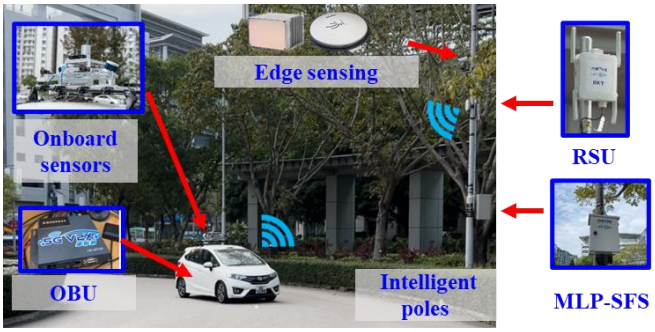


Fig. 1 Illustration of the vehicle-infrastructure cooperation system in Hong Kong C-V2X testbed. The multi-sensory vehicle platform is equipped with the onboard unit (OBU) to communicate with the roadside unit (RSU) through a 5G vehicle-to-everything (V2X) network, which operates at 5905-5925 MHz. MLP-SFS represents multi lamp poles sensor fusion server, which processes the edge sensing data in real time. The sensing information is shared between the vehicle and the infrastructure in a low-latency network. The communication range for RSU and OBU is 500 meters and 200 meters, respectively. Direct C-V2X (PC5) communication delay is less than 20ms.

With the rapid development of the roadside unit (RSU) and 5G vehicle-to-everything (V2X), vehicle-infrastructure (V2I) cooperation gained significant attention as it can enhance the sensing information to support various downstream applications for autonomous driving. Fig. 1 demonstrates the idea of the proposed V2I system. The intelligent street poles are equipped with the RSU devices, edge sensors and multi-lamp poles sensor fusion server (MLP-SFS), while the multi-sensory vehicle platform is equipped with the onboard unit (OBU) for V2X communication. As roadside infrastructure is a key component of smart cities which will be deployed at scale to investigate V2X cooperation according to the smart

*Equal contribution

[1]Feng Huang, Weisong Wen and Li-Ta Hsu are with the Hong Kong Polytechnic University, Hong Kong (†correspondence e-mail: welson.wen@polyu.edu.hk).

[2]Hang Chen, Alpamys Urtay and Dongzhe Su are with the Hong Kong Applied Science and Technology Research Institute (ASTRI), Hong Kong.

[3] https://sites.google.com/view/v2x-cooperative-navigation

2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC) | 979-8-3503-9946-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ITSC57777.2023.10422552

mobility roadmap for Hong Kong [14] announced by the Hong Kong Transport Department. Inspired by this, this paper proposes an RSI-assisted LIO method for urban mapping and localization, which leverages the global constraint provided by the RSIs based on the point registration between the ego-vehicle and RSU point clouds. The drift of the LIO is corrected by the registered absolute positioning with the help of the RSIs. To evaluate the performance of our proposed method, we collect multi-view RSIs and vehicle sensor data in the Hong Kong C-V2X testbed [15]. The contributions of this work are highlighted as follows:

(1) We present the deep learning-based vehicle detection framework using multi-RSI LiDARs, which provides an accurate initial guess for aligning the ego-vehicle and RSI LiDAR clouds.

(2) We propose an RSI-assisted LIO to reduce global drift using the global constraint provided by the point cloud registration between the vehicle and RSI.

(3) We extensively validate the effectiveness of the proposed method using the real-world data collected in Hong Kong C-V2X testbed. The results demonstrate that the positioning performance is significantly improved aided by the RSI in challenging urban areas. We also open-sourced our data[4] to benefit the research society.

The rest of the paper is structured as follows. Firstly, the related works are presented in Section II. The detail of the proposed method is elucidated in Section III. The performance evaluation is performed in Section IV. Finally, the conclusions and future perspectives are presented in Section V.

II. Related Work

LiDAR odometry (LO) estimates the relative pose between successive frames using scan-matching methods such as NDT [16] and LOAM [2]. Satisfactory results can be obtained in the widely evaluated KITTI dataset [17]. Unfortunately, according to our previous research [6], the performance of the LO is degenerated in challenging urban areas with a large number of dynamic objects. The deep learning-based method [18, 19] can identify the moving outlier in the LiDAR point cloud thus improving the positioning performance of the LO. However, it still suffers incorrect registration due to inefficient initial guess [20] of scan matching caused by the rapid movements. Therefore, IMU is usually used to cooperate with LiDAR for state estimation. Based on how to model the LiDAR and IMU measurements, the typically LiDAR-inertial fusion can be divided into two groups: the loosely-coupled and the tightly-coupled methods. Loosely-coupled LIO methods estimate the LiDAR point cloud and IMU data separately and fuse them in the positioning domain. In LOAM [2] and Lego-LOAM [21], the pose integration from the IMU measurements is employed as the initial guess for LiDAR scan matching. The work in [22] used an extended Kalman filter (EKF) to fuse the measurements from rotational velocity IMU and multi-LiDARs. However, EKF is subject to the first-order Markov assumption, thus historical information cannot be fully utilized in current state estimation [23]. The tightly-coupled method directly fuses the raw measurement of LiDAR and IMU sensors. For example, LINS adopted an iterated error-state Kalman filter (ESKF) with robocentric formulation. LIO-mapping [5] proposed to joint-optimize the state of IMU and LiDAR within a sliding window. However, the performance of such methods depends on the quality of the IMU as it is directly fused with the scan-matching residuals for state estimation.

Recent advancements in RSU and 5G-V2X have significantly enhanced the capabilities of autonomous vehicles. In [24], roadside-assisted received signal strength (RSS) is fused with GNSS signal to improve the V2I-based localization via a simulated platform. The author [25] proposed a background filtering method that integrated with 3D object detectors [26] to detect vehicles from RSI LiDAR data in mine environments. A work in [27] introduced a vehicle-to-everything (V2X) perception framework using a visual transformer via CARLA [28] simulator. Another team [29] proposed to use of RSI LiDAR for the detection and tracking of pedestrians and vehicles. Recent work [30] proposed to align the point cloud between the ego-vehicle and the roadside infrastructures using semantic traffic elements in campus scenarios. Our previous work [31] showed the positioning performance of multi-agent can be improved by sensor measurement sharing in dense urban areas. However, there is no existing research on using vehicle infrastructure to improve the positioning of LIO in challenging urban areas such as Hong Kong, which is the aim of this paper.

III. Methodology

A. Overview of the Proposed Architecture

An overview of the proposed system is depicted in Fig. 2. The proposed system comprises two major components: (1) vehicle detection and tracking by multi-RSI point clouds; (2) LIO aided by RSI global constraint. The system input includes the vehicle LiDAR point cloud, IMU measurements, and roadside LiDAR point cloud by multi-RSUs. Firstly, the point clouds captured by the multi-RSIs are fused for vehicle detection and tracking. Then the detected vehicle poses served as the initial guess for point cloud alignment between the onboard local map and the multi-RSIs. Finally, the registered vehicle pose is used as the absolute constraint in factor graph optimization for accurate positioning.

The following symbols are defined and followed by the rest of the paper.

a) The vehicle LiDAR body coordinate $\{\cdot\}^L$: fixed at the center of the onboard LiDAR sensor.
b) The vehicle IMU coordinate $\{\cdot\}^b$: fixed at the center of the onboard IMU sensor.
c) The RSI coordinate $\{\cdot\}^R$: fixed at the bottom center of the base roadside infrastructure. $\{\cdot\}^{R1}$and $\{\cdot\}^{R2}$ denotes the RSI1 frame and RSI2 frame, respectively. The origin of the RSI1 frame is the same as the RSI base frame in this paper.
d) The world coordinate $\{\cdot\}^W$: For example, $\mathbf{T}_{L_k}^W \in SE(3)$ is denoted as the LiDAR pose at timestamp k in the world coordinate,

[4] https://github.com/DarrenWong/RSI-aided_LIO

$$\mathbf{T}_{L_k}^{W} = \begin{bmatrix} \mathbf{R}_{L_k}^{W} & \mathbf{t}_{L_k}^{W} \\ 0 & 1 \end{bmatrix} \quad (1)$$

e) The k-th frame of the vehicle LiDAR cloud and RSI LiDAR cloud $\mathcal{P}^{L_k}$ and $\mathcal{P}^{R,L_k}$: The i-th points in $\mathcal{P}^{L_k}$ and $\mathcal{P}^{R,L_k}$ is denoted as $\mathbf{p}_i^{L_k}$ and $\mathbf{p}_i^{R,L_k}$. The k-th frame of the accumulated vehicle local cloud and multi-RSU cloud is represented as $\mathcal{M}^{W,L_k}$ and $\mathcal{M}^{W,RL_k}$.

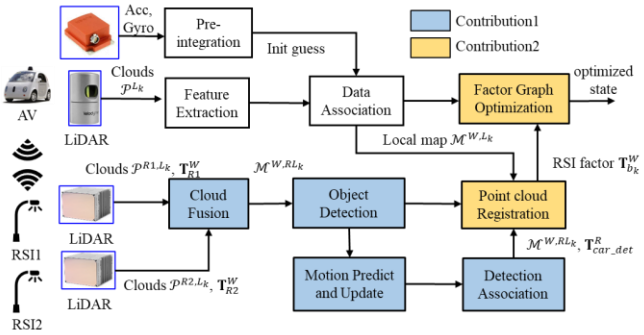


Fig. 2 Overview of the proposed pipeline. $\mathrm{T}_{car\_det}^{R}$ indicates object detection in the RSI base frame. T_{R1}^{W}, T_{R2}^{W} and $T_{b_k}^{W}$ denote the RSI1 pose, RSI2 pose and vehicle IMU body in the world frame.

B. *RSI-based Object Detection and Tracking*

To minimize the "blind zones" caused by the static traffic elements and moving dynamic objects, it is necessary to deploy a set of RSIs and LiDAR sensors optimally throughout to generate a precise and comprehensive point cloud map of the specific traffic area. By merging point clouds from different RSI LiDAR, a dense map of the area is created for various downstream tasks related to autonomous vehicles. The initial translation matrix of the multi-RSI LiDARs can be derived by adjusting the alignment of numerous landmark points located in overlapping field of view (FOV). The extrinsic matrix between two RSI LiDARs can be further optimized using the iterative closest point (ICP)-based [32] method,

$$\min_{\mathbf{T}_{R2}^{R1}} \sum_{i=1,\mathbf{p}_i^{R1,L_k}\epsilon\mathcal{P}^{R1,L_k},\mathbf{p}_i^{R2,L_k}\epsilon\mathcal{P}^{R2,L_k}}^{N^{R1,2}} \left\| \mathbf{p}_i^{R1,L_k} - \mathbf{T}_{R2}^{R1}\mathbf{p}_i^{R2,L_k} \right\|^2 \quad (2)$$

where $\mathbf{T}_{R2}^{R1}$ is the optimized registration between RSI1 and RSI2 point clouds. $N^{R1,2}$ indicate the total number of corresponding points of RSIs. Then the merged RSI points cloud $\mathcal{M}^{R,RL_k}$ in the RSI base frame can be computed,

$$\mathcal{M}^{R,RL_k} = \mathbf{T}_{R2}^{R1}\mathcal{P}^{R2,L_k} + \mathcal{P}^{R1,L_k} \quad (3)$$

Fig. 3 illustrates the cloud fusion results in the base frame with the optimal extrinsic parameters. It can be observed that the points that belong to the road curb (Fig. 3(A)) and roundabout (Fig. 3(B)) are partially covered by different RSI LiDAR that forms a single coherent model. Given the pre-calibrated RSI pose in the world coordinate, the merged cloud or single RSI cloud can be transformed into the global frame,

$$\mathcal{M}^{W,RL_k} = \mathbf{T}_R^W \mathcal{M}^{R,RL_k} \quad (4)$$

Object detection and tracking algorithm is then applied on the fused point cloud $\mathcal{M}^{R,RL_k}$. The tracking pipeline requires the object to be detected first and then utilizes the detection results as inputs for the tracking algorithm.

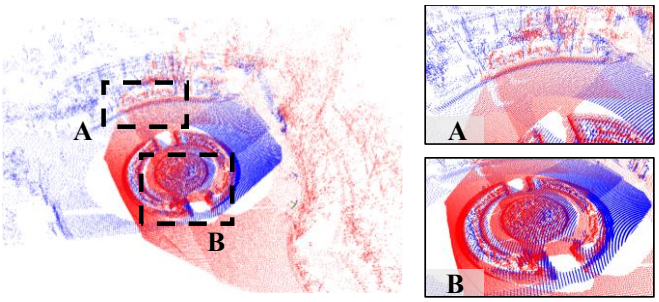


Fig. 3 Illustration of the point cloud from multi-RSI LiDAR. It is merged by transforming the cloud from RSI1(blue) and RSI2 (red) into the base frame. The marked areas indicate the point cloud of road curb (A) and roundabout (B) are aligned well.

***3D object detection*:** The CenterPoint [33] was adopted for RSI LiDAR object detection because of its efficiency and effectiveness. The two-stage detector first detects the center of the object and then estimates attributes such as the size and rotation of the 3D box. Additional point features are then used to refine the initial estimations. The detector demonstrates robust performance in identifying and localizing objects in 3D space in the Waymo Open Dataset [34] and our customized RSI data.

***3D object tracking*:** 3D Object tracking is to track objects through multiple point cloud frames over time that include the position, velocity, and other attributes of the objects. Each object is tracked using a bounding box with a unique ID. To achieve this, we leverage the Simpletrack algorithm [35], which implements the concept of "tracking-by-detection" to track the 3D box states with lower computational costs while retaining tracking performance. Simpletrack [35] applies non-maximum suppression (NMS) to remove extra boxes and preserve authentic boxes with low scores after receiving the object detection bounding boxes. Generalized IoU (GIOU) [36] is used as association metrics and the Hungarian algorithm [37] matches bounding boxes. Then Kalman Filter is applied to estimate the bounding boxes motion in the next frames. To manage the life cycle of tracklets [35], initiate a tracklet after 3 consecutive matches, and terminate a tracklet after 2 misses is implemented. Finally, the specific ego-vehicle can be identified by the RSI with object tracking and V2X communication.

C. *RSI Assisted LiDAR-Inertial Odometry*

The factor graph of the proposed RSI-assisted LIO is depicted in Fig. 4. Noted that the RSI factor is a unary constraint with the state. The process of RSI-assisted odometry includes two parts, the point cloud registration aided by object tracking and a factor graph incorporating the RSI factor.

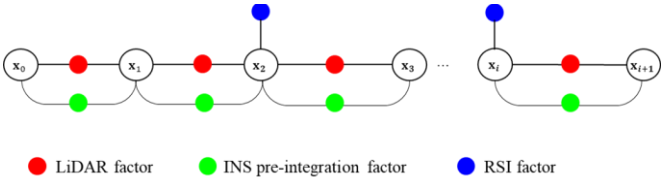


Fig. 4 The graph of the RSI-assisted LIO. $\mathbf{x}_i$ denotes the estimated state.

The accumulated vehicle-side local map $\mathcal{M}_{L_k}^{W}$ at timestamp k can be expressed as,

$$\mathcal{M}^{W,L_k} = \sum_{s=1}^{N^s} \mathbf{T}_{L_s}^{W} \mathcal{P}^{L_k,s} \quad (5)$$

where N^s denotes the extracted N surrounding keyframes. $\mathbf{T}_{L_s}^{W}$ and $\mathcal{P}^{L_k,s}$ represent the pose and LiDAR point cloud of the corresponding keyframe s, respectively.

The vehicle local point cloud map $\mathbf{T}_{L_s}^{W}$ can be transformed into the RSI frame $\mathbf{T}_{L_s}^{R}$ based on the vehicle tracking pose $\mathbf{T}_{V\_det}^{R}$ and the $\mathbf{T}_{L_k}^{W}$,

$$\mathcal{M}^{R,L_k} = \mathbf{T}_{V\_det}^{R} \mathbf{T}_{L}^{V} (\mathbf{T}_{L_k}^{W})^{-1} \mathcal{M}^{W,L_k} \quad (6)$$

where $\mathbf{T}_{L}^{V}$ indicates the extrinsic parameter of vehicle center and its LiDAR. The point clouds of vehicle and RSI can be further aligned using point-based ICP [32] after removing the ground points,

$$\min_{\mathbf{T}_{L_k}^{R}} \sum_{i=1,\mathbf{p}_i^{R,RL_k} \epsilon \mathcal{M}^{R,RL_k}, \mathbf{p}_i^{R,L_k} \epsilon \mathcal{M}^{R,L_k}}^{N^{RL}} \left\| \mathbf{p}_i^{R,RL_k} - \mathbf{T}_{L_k}^{R} \mathbf{p}_i^{R,L_k} \right\|^2 \quad (7)$$

where $\mathbf{T}_{L_k}^{R}$ is the optimized alignment between RSI and vehicle LiDAR point clouds at timestamp k. N^{RL}, $\mathbf{p}_i^{R,RL_k}$ and $\mathbf{p}_i^{R,L_k}$ indicate the total number of corresponding points, points in $\mathcal{M}^{R,RL_k}$ and $\mathcal{M}^{R,L_k}$, respectively. The details of the point cloud registration process are described in Algorithm 1.

The output of Algorithm 1 is the registered vehicle pose in the RSI frame. With the pre-calibrated RSI position $\mathbf{T}_{R}^{W}$ in the world frame, the registered accurate vehicle pose $\widehat{\mathbf{T}}_{b_k}^{W}$ can be further transformed into world frame to form the RSI factor,

$$\widehat{\mathbf{T}}_{b_k}^{W} = \mathbf{T}_{b}^{L} \mathbf{T}_{R}^{W} \mathbf{T}_{L_k}^{R} \quad (8)$$

where $\mathbf{T}_{b}^{L}$ indicates the extrinsic parameter of LiDAR and IMU. Finally, the cost function of the proposed method aided by the RSI factor can be expressed as,

$$\min_{\mathbf{T}_{b_k}^{W}} \left(\left\| \mathbf{r}_L(\mathbf{T}_{b_k}^{W}) \right\|_{\mathbf{P}_L}^2 + \left\| \mathbf{r}_R(\mathbf{T}_{b_k}^{W}) \right\|_{\mathbf{P}_R}^2 \right) \quad (9)$$

where $\mathbf{r}_L$, $\mathbf{r}_R$ indicate the residuals of the LIO factor and RSI factor, respectively. $\mathbf{P}_{(\cdot)}$ present the covariance matrix of each term. IMU pre-integration is used to provide the initial guess and deskew point cloud. The accumulated drift of the LIO method can be corrected effectively with the RSI factor.

Algorithm 1: Vehicle-RSI Point cloud registration

Inputs: multi-RSI clouds $\mathcal{M}^{R,RL_k}$, vehicle local map $\mathcal{M}^{W,L_k}$

Outputs: $\mathbf{T}_{L_k}^{R}$

Step 1: Transform the $\mathcal{M}^{W,L_k}$ from the world frame to the RSI frame $\mathcal{M}^{R,L_k}$ based on the detection information $\mathbf{T}_{V\_det}^{R}$ using (6).

Step 2: Remove the ground points of the $\mathcal{M}^{R,RL_k}$ and $\mathcal{M}^{R,L_k}$ based on the ground height.

Step 3: Search the corresponding points between $\mathcal{M}_{L_k}^{W}$ and $\mathcal{M}_{L_k}^{R}$ based on the kd tree.

Step 4: Iteratively compute the optimized $\mathbf{T}_{L_k}^{R}$ by minimizing the residuals between the clouds from the vehicle and RSI using (7).

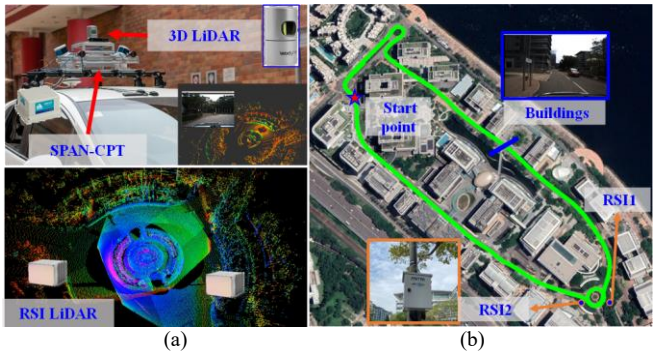
IV. Experimental Results and Discussions

A. Experiment Setup

Sensor Setups: We use the UrbanNav [38] vehicle platform to conduct the experiments. GNSS, INS, cameras, and LiDARs are equipped on the vehicle platform. In addition, NovAtel SPAN-CPT integrates a fiber optics gyroscope (FOG) and GNSS-RTK to provide the ground truth (GT) positioning. Furthermore, the measurements from SPAN-CPT are further tightly coupled using the NovAtel Inertial Explorer to maximize the accuracy of the GT. Each RSI in the Hong Kong testbed [15] contains measurements from GNSS, 300-line LiDAR, high-performance V2X communication, and edge computing. In this experiment, we collect 10 Hz LiDAR data and 400 Hz IMU from one vehicle platform and 10 Hz LiDAR data from two RSIs simultaneously in a challenging urban scenario. Fig. 5 presents the sensor setup and evaluated trajectory with a total length of 1.85 km. The RSIs are located in the roundabout area (around half) of the trajectory. The evaluated route revisits the start point to investigate the performance between the LiDAR loop closure constraint [39] and the global constraint from the RSI. PPS time synchronization with the GPS source is performed on our vehicle platform while PTP time synchronization with the GPS source is conducted on the RSI side.

The specification for the data process is provided below:

- NVIDIA Jetson AGX Orin as RSI edge device
- An AMD Ryzen 5950X CPU, 64 GB DDR4 RAM and Nvidia 2080TI graphic card for onboard computing



(a) (b)

Fig. 5 (a) Setup for Vehicle-infrastructure platform. A Velodyne 32-line LiDAR, and SPAN-CPT for the vehicle side while two Innovusion Jaguar LiDARs are integrated with two RSIs; (b) The evaluated scene of Hong Kong C-V2X testbed and its ground truth trajectory.

Evaluation Metrics: The effectiveness of the listed method is evaluated via the absolute translation error (ATE) to investigate the global accuracy of the pose estimation.

$$ATE_k = \left\| trans(\mathbf{T}_{k,gt}) - trans(\mathbf{T}_{k,est}) \right\| \quad (10)$$

where $\mathbf{T}_{k,gt}$ and $\mathbf{T}_{k,est}$ represents the absolute ground truth pose and estimate pose at k-th frame. $trans(.)$ denotes the translation part of the pose.

Evaluated Methods: To verify the performance of the proposed algorithms, we use the following methods,

(1) LIO-SAM [4]: The state-of-the-art LiDAR-inertial odometry.

(2) RSIA-LIO: The proposed RSI-assisted LiDAR-inertial method. The RSI factor will constrain the state in the

factor graph once available.

B. Experimental Validation in Urban Areas

1) Results of Object Detection by RSIs

To conduct the evaluation, the models for our self-collected RSI data are trained with 80% of the available data, and the remaining 20% of the data is used for validation. All detection results are measured by the average precision (AP) [40] at the 3D-GIoU [36] threshold of 0.7 for the *cars* class. The AP on the validation set is computed with 40 recall positions.

Table I demonstrates the performance evaluation of CenterPoint [35] on the validation set of our self-collected dataset. The trained model achieves 57.17% 3D detection AP and 90.09%average performance of orientation similarity (AOS) AP. Most of the vehicles are successfully predicted with satisfactory accuracy from Table I. Examples of the testing results are illustrated in Fig. 6. Moreover, vehicles can be tracked continuously [35] throughout different frames as shown by their ID number (shown in Fig. 6).

TABLE I. Performance evaluations of the CenterPoint on our validation set.

Results	Self-collected RSI data
3D detection AP (%)	57.17
AOS AP (heading angle) (%)	90.09

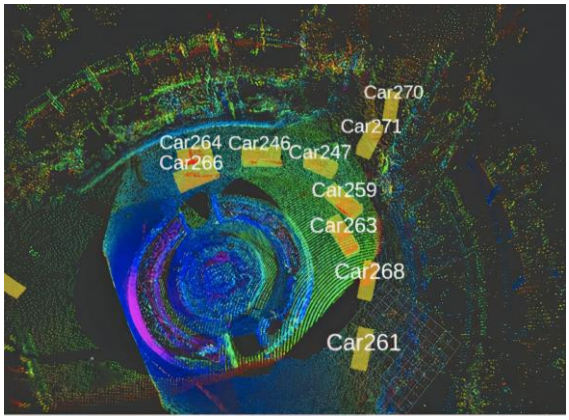
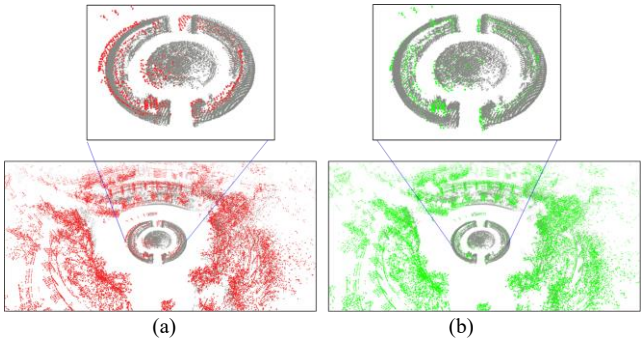


Fig. 6 Illustration of the RSI detection results.



(a) (b)

Fig. 7 (a) The point cloud transformation based on the detection information provided by RSI; (b) The results of fine registration. The RSI cloud is marked in grey. The transformed vehicle point cloud map using the detection information is marked in red. The registered vehicle points using ICP are marked in green.

2) Results of Point Registration

Fig. 7 demonstrates the point cloud alignment results between the onboard local map and the multi-RSIs. It can be observed that server meters error (shown in Fig. 7 (a)) if only align the point cloud based on the detection information provided by the RUSs. With the help of the ICP-based [32] point cloud registration, we can successfully align the vehicle-RSI point cloud at high accuracy, as demonstrated in Fig. 7-(b).

3) Results of RSI-assisted LIO

To validate the contribution of the proposed RSI-assisted LIO methods, experiments were conducted under the challenging urban scene. Table II demonstrates the positioning performance using the evaluated methods. An RMSE of 17.004 meters was obtained using the existing LIO-SAM, with a standard deviation of 10.784 meters. The RMSE error of LIO-SAM with loop closure enabled was 18.992 meters, which was unable to correct the large drift accumulation by LiDAR odometry in urban areas. With the help of the proposed RSI-aided method, both RMSE and standard deviation were reduced to 5.614 meters and 4.318 meters. Interestingly, the proposed RSIA-LIO method with loop closure outperforms other methods with an RMSE of 2.676 meters. The combination of the LiDAR loop closure constraint [39] and the global constraint from the RSI is complementary for reliable odometry and mapping in urban areas.

Figs. 8-9 show the positioning error and trajectories using the listed methods while mapping results of the proposed method are presented in Fig. 10. It can be seen that the positioning accuracy of LIO-SAM cannot be benefited by the loop closure factor in the grey area of Fig. 8. The possible reason for this is that it can only eliminate the positioning error in the revisit points, but it introduces more errors in the estimated trajectory due to the large accumulation errors. After applying the RSI global constraint, the positioning performance is improved by 67% in the urban area, especially in the yellow area in Fig. 8 and Fig. 9. Interestingly, the error of RSIA-LIO can be further eliminated by including the loop closure factor in the revisiting point (around epoch 250). The RMSE is further reduced from 5.614 meters (RSIA-LIO) to 2.676 meters (RSIA-LIO w/ loop) with an improvement of 84.3% compared to the existing LIO method. Moreover, Fig. 10 presents the mapping result using the proposed method is well-aligned with Google Maps. Unfortunately, unexpected positioning accuracy is still observed around epoch 200 (shown in Fig. 8) due to the drift accumulation. Therefore, it is important to include more RSIs to help improve the performance of state estimation for autonomous driving.

TABLE II. Performance evaluation of listed methods.

Results	**LIO-SAM**	**LIO-SAM w/ loop**	**RSIA-LIO**	**RSIA-LIO w/ loop**
RMSE(m)	17.004	18.992	5.614	**2.676**
MEAN (m)	13.147	14.077	3.588	**2.392**
STD (m)	10.784	12.748	4.318	**1.200**
Max (m)	33.152	36.177	19.293	**5.781**
Improvement	/	/	67.0%	**84.3%**

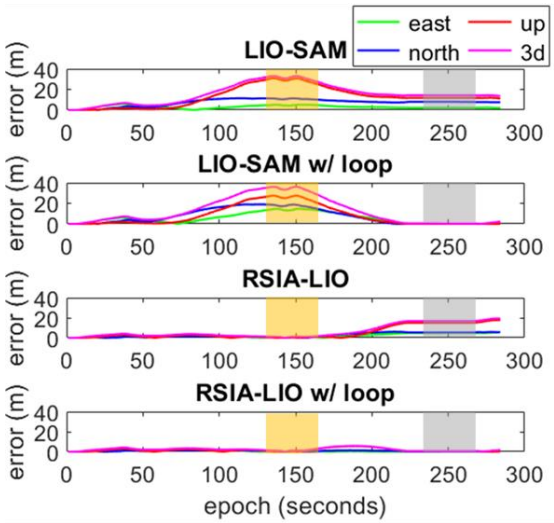


Fig. 8. Positioning errors in the east-north-up (ENU) directions. The x-axis indicates the timestamp while the y-axis represents errors in meters, respectively. The area marked in yellow indicates the error eliminated with the RSI-assisted global constraints while the area marked in grey denotes the loop closure factor when revisiting the start points.

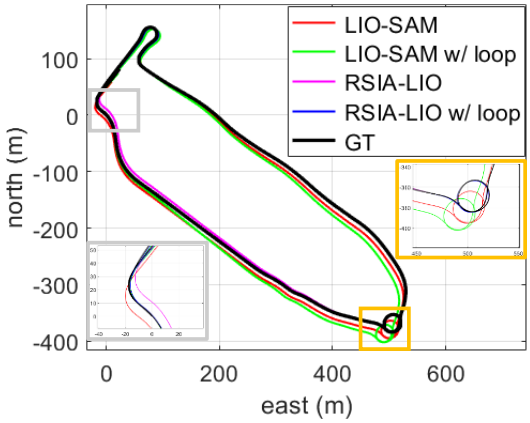


Fig. 9. 2D positioning of the listed methods in the urban area. The x-axis indicates the east direction while the y-axis represents the north direction. The areas marked in yellow and grey indicate the corresponding yellow and grey areas in Fig. 8, respectively.

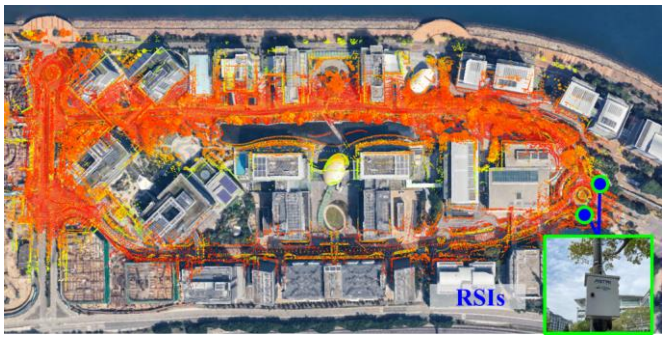


Fig. 10. Mapping results of the proposed RSI-assisted LIO with loop closure enabled, which showed high consistency with Google Maps. Noted that the ENU orientation of the map is rotated for better demonstration.

V. Conclusion

With the growth development in smart cities, it is a huge demand for vehicle-infrastructure cooperation. This paper presented a complete pipeline to enhance the performance of the LIO aided by the global constraint provided by the RSIs. Evaluation results on Hong Kong C-V2X testbed show that our proposed method outperforms the state-of-the-art method regarding absolute positioning accuracy.

In the future, we will evaluate our proposed system in various challenging urban scenarios in Hong Kong with the increasing deployment of intelligent RSI. We will also investigate the calibration problems of multi-RSIs to guarantee the global position of RSIs are always optimal over time. With the aid of the accurate point cloud map using our proposed method, the potential GNSS NLOS can be detected [41] which is one of our future works. The integration of GNSS [42, 43] is another interesting topic to reduce positioning errors if RSI is unavailable in urban areas.

Acknowledgment

This work is supported by the Smart Traffic Fund (STF) on the project "Advanced C-V2X Applications to Enhance Hong Kong's Mobility Competence and Road Safety (PSRI/19/2109/RA_ACAE)", granted by Hong Kong Transport Department. This research was also supported by the Faculty of Engineering, The Hong Kong Polytechnic University under the project "Perception-based GNSS PPP-RTK/LVINS integrated navigation system for unmanned autonomous systems operating in urban canyons". The authors would like to express their thanks to Hoi-Fung Ng, Xikun Liu, and Yihan Zhong for their kind help in the data collection.

References

[1] L. Wei, Z. Li, J. Gong, C. Gong, and J. Li, "Autonomous Driving Strategies at Intersections: Scenarios, State-of-the-Art, and Future Outlooks," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 19-22 Sept. 2021 2021, pp. 44-51, doi: 10.1109/ITSC48978.2021.9564518.

[2] J. Zhang and S. Singh, "Low-drift and real-time lidar odometry and mapping," *Autonomous Robots,* vol. 41, no. 2, pp. 401-416, 2017.

[3] Y. Chen, S. Sun, H. Yin, and M. H. Ang, "Exploring the Effect of 3D Object Removal using Deep Learning for LiDAR-based Mapping and Long-term Vehicular Localization," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, 2022: IEEE, pp. 1730-1735.

[4] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and 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)*, 24 Oct.-24 Jan. 2021 2020, pp. 5135-5142, doi: 10.1109/IROS45743.2020.9341176.

[5] H. Ye, Y. Chen, and M. Liu, "Tightly Coupled 3D Lidar Inertial Odometry and Mapping," in *2019 International Conference on Robotics and Automation (ICRA)*, 20-24 May 2019 2019, pp. 3144-3150, doi: 10.1109/ICRA.2019.8793511.

[6] F. Huang, W. Wen, J. Zhang, and L. T. Hsu, "Point Wise or Feature Wise? A Benchmark Comparison of Publicly Available Lidar Odometry Algorithms in Urban Canyons," *IEEE Intelligent Transportation Systems Magazine,* vol. 14, no. 6, pp. 155-173, 2022, doi: 10.1109/MITS.2021.3092731.

[7] X. Xu *et al.*, "A review of multi-sensor fusion slam systems based on 3D LIDAR," *Remote Sensing,* vol. 14, no. 12, p. 2835, 2022.

[8] J. Zhang, W. Wen, F. Huang, Y. Wang, X. Chen, and L.-T. Hsu, "GNSS-RTK Adaptively Integrated with LiDAR/IMU Odometry for Continuously Global Positioning in Urban Canyons," *Applied Sciences,* vol. 12, no. 10, p. 5193, 2022. [Online]. Available: https://www.mdpi.com/2076-3417/12/10/5193.

[9] J. Zhang, W. Wen, F. Huang, X. Chen, and L.-T. Hsu, "Coarse-to-Fine Loosely-Coupled LiDAR-Inertial Odometry for Urban Positioning and Mapping," *Remote Sensing,* vol. 13, no. 12, p. 2371, 2021. [Online]. Available: https://www.mdpi.com/2072-4292/13/12/2371.

[10] L. Chang, X. Niu, T. Liu, J. Tang, and C. Qian, "GNSS/INS/LiDAR-SLAM Integrated Navigation System Based on Graph Optimization," *Remote Sensing,* vol. 11, no. 9, p. 1009, 2019. [Online]. Available: https://www.mdpi.com/2072-4292/11/9/1009.

[11] L.-T. Hsu, "Analysis and modeling GPS NLOS effect in highly urbanized area," *GPS solutions,* vol. 22, no. 1, p. 7, 2018.

[12] C. Xu, H. Zhang, and J. Gu, "Scan Context 3D Lidar Inertial Odometry via Iterated ESKF and Incremental K-Dimensional Tree," in *2022 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, 18-20 Sept. 2022 2022, pp. 21-27, doi: 10.1109/CCECE49351.2022.9918525.

[13] K. Koide, M. Yokozuka, S. Oishi, and A. Banno, "Globally consistent and tightly coupled 3D LiDAR inertial mapping," in *2022 International Conference on Robotics and Automation (ICRA)*, 2022: IEEE, pp. 5622-5628.

[14] Transport Department of the Government of Hong Kong. "Smart Mobility Roadmap for Hong Kong " https://www.td.gov.hk/filemanager/en/publication/smr_roadmap_hk.pdf (accessed April 27, 2023).

[15] Hong Kong Applied Science and Technology Research Institute Company Limited. "C-V2X Technology." https://www.astri.org/tdprojects/connected-vehicle-v2x-technology/ (accessed May 02, 2023).

[16] P. Biber and 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)*, 2003, vol. 3: IEEE, pp. 2743-2748.

[17] A. Geiger, P. Lenz, C. Stiller, and R. Urtasun, "Vision meets robotics: The kitti dataset," *The International Journal of Robotics Research,* vol. 32, no. 11, pp. 1231-1237, 2013.

[18] P. Pfreundschuh, H. F. C. Hendrikx, V. Reijgwart, R. Dubé, R. Siegwart, and A. Cramariuc, "Dynamic Object Aware LiDAR SLAM based on Automatic Generation of Training Data," presented at the IEEE International Conference on Robotics and Automation (ICRA), 2021.

[19] F. Huang, D. Shen, W. Wen, J. Zhang, and L.-T. Hsu, "A Coarse-to-Fine LiDAR-Based SLAM with Dynamic Object Removal in Dense Urban Areas," in *Proceedings of the 34th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2021)*, 2021, pp. 3162-3172.

[20] Y. Zhong, F. Huang, J. Zhang, W. Wen, and L.-T. Hsu, "Low-cost solid-state LiDAR/inertial-based localization with prior map for autonomous systems in urban scenarios," *IET Intelligent Transport Systems,* vol. 17, no. 3, pp. 474-486, 2023, doi: https://doi.org/10.1049/itr2.12273.

[21] T. Shan and B. Englot, *LeGO-LOAM: Lightweight and Ground-Optimized Lidar Odometry and Mapping on Variable Terrain*. 2018, pp. 4758-4765.

[22] M. Demir and K. Fujimura, "Robust Localization with Low-Mounted Multiple LiDARs in Urban Environments," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 27-30 Oct. 2019 2019, pp. 3288-3293, doi: 10.1109/ITSC.2019.8916995.

[23] W. Wen, T. Pfeifer, X. Bai, and L.-T. Hsu, "Factor graph optimization for GNSS/INS integration: A comparison with the extended Kalman filter," *NAVIGATION: Journal of the Institute of Navigation,* vol. 68, no. 2, pp. 315-331, 2021, doi: 10.1002/navi.421.

[24] J. Li, J. Gao, H. Zhang, and T. Z. Qiu, "RSE-assisted lane-level positioning method for a connected vehicle environment," *IEEE Transactions on Intelligent Transportation Systems,* vol. 20, no. 7, pp. 2644-2656, 2018.

[25] G. Wang, J. Wu, T. Xu, and B. Tian, "3D Vehicle Detection With RSU LiDAR for Autonomous Mine," *IEEE Transactions on Vehicular Technology,* vol. 70, no. 1, pp. 344-355, 2021, doi: 10.1109/TVT.2020.3048985.

[26] A. H. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang, and O. Beijbom, "Pointpillars: Fast encoders for object detection from point clouds," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 12697-12705.

[27] R. Xu, H. Xiang, Z. Tu, X. Xia, M.-H. Yang, and J. Ma, "V2X-ViT: Vehicle-to-Everything Cooperative Perception with Vision Transformer," in *Computer Vision – ECCV 2022*, Cham, S. Avidan, G. Brostow, M. Cissé, G. M. Farinella, and T. Hassner, Eds., 2022// 2022: Springer Nature Switzerland, pp. 107-124.

[28] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," in *Conference on robot learning*, 2017: PMLR, pp. 1-16.

[29] J. Zhao, H. Xu, H. Liu, J. Wu, Y. Zheng, and D. Wu, "Detection and tracking of pedestrians and vehicles using roadside LiDAR sensors," *Transportation Research Part C: Emerging Technologies,* vol. 100, pp. 68-87, 2019/03/01/ 2019, doi: https://doi.org/10.1016/j.trc.2019.01.007.

[30] Y. He, L. Ma, Z. Jiang, Y. Tang, and G. Xing, "VI-eye: semantic-based 3D point cloud registration for infrastructure-assisted autonomous driving," in *Proceedings of the 27th Annual International Conference on Mobile Computing and Networking*, 2021, pp. 573-586.

[31] W. Wen, X. Bai, G. Zhang, S. Chen, F. Yuan, and L.-T. Hsu, "Multi-agent collaborative GNSS/Camera/INS integration aided by inter-ranging for vehicular navigation in urban areas," *IEEE Access,* vol. 8, pp. 124323-124338, 2020.

[32] P. J. Besl and N. D. McKay, "A method for registration of 3-D shapes," *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 14, no. 2, pp. 239-256, 1992, doi: 10.1109/34.121791.

[33] T. Yin, X. Zhou, and P. Krahenbuhl, "Center-based 3d object detection and tracking," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021, pp. 11784-11793.

[34] P. Sun *et al.*, "Scalability in perception for autonomous driving: Waymo open dataset," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 2446-2454.

[35] Z. Pang, Z. Li, and N. Wang, "Simpletrack: Understanding and rethinking 3d multi-object tracking," in *Computer Vision–ECCV 2022 Workshops: Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part I*, 2023: Springer, pp. 680-696.

[36] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, and S. Savarese, "Generalized intersection over union: A metric and a loss for bounding box regression," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019, pp. 658-666.

[37] M. Wright, "Speeding up the Hungarian algorithm," *Computers & Operations Research,* vol. 17, no. 1, pp. 95-96, 1990.

[38] L.-T. Hsu *et al.*, "Hong Kong UrbanNav: An Open-Source Multisensory Dataset for Benchmarking Urban Navigation Algorithms," *NAVIGATION: Journal of the Institute of Navigation,* vol. 70, no. 4, p. navi.602, 2023, doi: 10.33012/navi.602.

[39] Z. Jin, Y. Shao, M. So, C. Sable, N. Shlayan, and D. M. Luchtenburg, "A Multisensor Data Fusion Approach for Simultaneous Localization and Mapping," in *2019 IEEE Intelligent Transportation Systems Conference (ITSC)*, 27-30 Oct. 2019 2019, pp. 1317-1322, doi: 10.1109/ITSC.2019.8916930.

[40] M. Everingham, L. Van Gool, C. K. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (voc) challenge," *International journal of computer vision,* vol. 88, no. 2, pp. 303-338, 2010.

[41] W. W. Wen and L. T. Hsu, "3D LiDAR Aided GNSS NLOS Mitigation in Urban Canyons," *IEEE Transactions on Intelligent Transportation Systems,* vol. 23, no. 10, pp. 18224-18236, 2022, doi: 10.1109/TITS.2022.3167710.

[42] X. Liu, W. Wen, F. Huang, H. Gao, Y. Wang, and L.-T. Hsu, "3D LiDAR Aided GNSS NLOS Mitigation for Reliable GNSS-RTK Positioning in Urban Canyons," *arXiv preprint arXiv:2212.05477,* 2022.

[43] F. Huang, W. Wen, H.-F. Ng, and L.-T. Hsu, "LiDAR Aided Cycle Slip Detection for GNSS Real-Time Kinematic Positioning in Urban Environments," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, 2022: IEEE, pp. 1572-1578.