Adaptive Multi-Sensor Integrated Navigation System Aided by Continuous Error Map from RSU for Autonomous Vehicles in Urban Areas

Feng Huang[1], Weisong Wen[1*], Guohao Zhang[1], Dongzhe Su[2] and Li-Ta Hsu[1]

***Abstract*— There has been a widespread study on multi-sensor integration to achieve precise and robust odometry for autonomous vehicles (AVs) in urban areas. LiDAR odometry and visual odometry can be affected by structureless scenarios and numerous dynamic objects. GNSS positioning can be degenerated due to the multipath and non-line-of-sight signals by buildings. Therefore, selecting appropriate weighing for heterogeneous sensors is a challenge for multi-sensor fusion. With the advancements in cellular vehicle-to-everything (C-V2X) and intelligent roadside units (RSUs), vehicles and the RSUs can collaborate to deliver reliable service. Inspired by this, this paper investigates continuous error maps for available sensors under different time conditions (noon, sunset, and night) to improve the positioning performance of surrounding AVs in complex urban environments. In particular, this paper presents an error-map-aided multi-sensor integrated system, which benefits from the error information collected by a sensor-rich AV. Then the error information is uploaded to the RSUs which is then distributed to the AVs. A smaller weight is assigned if a larger error is queried from the error map. To validate our approach, experiments were performed using the realistic CARLA simulator and our self-developed GNSS RUMS simulator. To benefit the research community, we open-sourced the implementation on our project page[3].**

I. Introduction

Recent advances in autonomous systems show the great potential for smart mobility. Simultaneous localization and mapping (SLAM) are [1] fundamental for most autonomous systems. Three-dimensional (3D) light detection and ranging (LiDAR) provides dense 3D point clouds of the surrounding information, which is widely utilized to provide position and map solutions [2] for autonomous systems. However, the performance of LiDAR-based odometry can be affected by the numerous dynamic objects [3] and the structureless environments [4]. Visual odometry [5] is a popular technique to provide state estimation by feature matching. But the performance of the visual-based method is sensitive to illumination conditions and the available features [6]. The global navigation satellite system (GNSS) provides absolute positioning services. Unfortunately, its performance can be degenerated due to the non-line-of-sight (NLOS) and multipath [7]. A single sensor is hard to meet the reliable navigation requirements for AV, thus multi-sensor integration has received significant attention because of their complementary and redundancy.

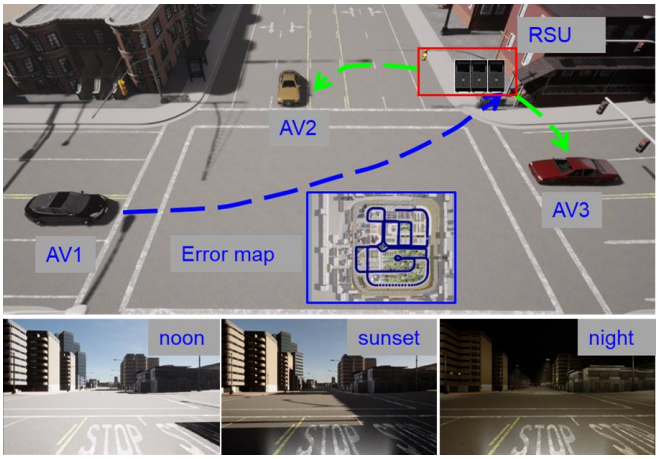

Fig. 1 ***Top:*** Illustration of the error map broadcast through the roadside unit (RSU). AV1 is equipped with a sensor-rich (e.g., LiDAR, camera, GNSS, and high-end devices which can provide ground truth positioning) to evaluate the sensor error periodically, while the AV2 and AV3 are the autonomous vehicles that receive the error map information to aid their navigation with their available sensors. ***Bottom:*** The collected RGB images day and night.

Many efforts have been made to achieve reliable sensor integration [8-11]. The most common way is to combine inertial sensors with LiDAR [9, 12] to provide an initial guess for scan matching. Visual-inertial odometry (VIO) [10] can provide robust pose estimation by integrating high-frequency inertial sensors. However, both LiDAR-inertial (LIO) and VIO suffer from accumulated errors [13, 14] under the degenerated scenes. With the help of GNSS positioning, the drift accumulation problem can be eliminated [13]. But the GNSS positioning performance is significantly degraded in dense urban areas [7, 11]. In short, each sensor has its advantage and limitation in different scenes. How to adaptively fuse the different types of sensors appropriately for autonomous vehicles in urban areas is still an open question.

Roadside Unit (RSU) is a critical infrastructure for the next-generation vehicular network [15], which is mounted along the road or sidewalk to perform real-time communication with vehicles and pedestrians. It is a key component of smart mobility that will be deployed at scale to explore the vehicle-to-everything (V2X) capabilities according to the Hong Kong Transport Department [16]. With the rapid development of the 5G and RSUs, vehicle-infrastructure cooperation [17] gained significant attention for

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

[2]Dongzhe Su is with the Hong Kong Applied Science and Technology Research Institute (ASTRI), Hong Kong.

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

intelligent transportation systems. Inspired by the data transmission efficiency [18] between vehicle infrastructure and real-time traffic density estimation [19], we introduce continuous error map estimation and its downstream application for multi-sensor integration. Fig. 1 illustrates the basic idea of the proposed system. To generate the error maps for different sensors in a specific urban area, we leverage the full-sensor-enabled autonomous high definition (HD) map update [20] vehicle or autonomous bus (AV1 in Fig. 1) for data collection and error evaluation periodically. The error data generated by the vehicle is then transmitted to the roadside edge system which merges the error information into a real-time error map. Intelligent vehicles (AV2 and AV3 in Fig. 1) in the same region can share the error information broadcasted by RSUs. Vehicles equipped with a single sensor can optimize their navigation with the error map, while other AVs equipped with a full-sensor suit can perform sensor fusion with the prior weighting provided by the error maps.

To evaluate our approach, we collect the multi-sensor data from multi-vehicles, including LiDAR, camera, inertial measurement unit (IMU), and GNSS, that considers the real-world noise in the urban day and night conditions using the high-fidelity simulated platform CARLA [21] and GNSS RUMS [22]. The contributions of this work are summarized as follows,

(1) We study the continuous error map estimation for vehicle-Infrastructure cooperation. The up-to-date error map can support various downstream navigation tasks.

(2) We present the error map-aided adaptive multi-sensor integrated system, which benefited from the prior information from the error map estimation.

(3) We evaluate the performance of the proposed system using extensively simulated data that considers real-world noise. The implementation is open-sourced at Github [1] to benefit the research community.

The rest of the paper is structured as follows. Firstly, the related works are presented in Section II. The design of the proposed system is described in Section III. Numerous experiments were conducted to verify the effectiveness of the proposed method in Section IV. Finally, the conclusions and future perspectives are presented in Section V.

II. Related Work

A. Error Map Estimation

Several studies [23-25] have investigated the positioning error from different sensors in complex scenes. In [26], the error of various LiDAR SLAM methods is reviewed and summarized using the KITTI dataset [27]. Another team [28] investigated the pose accuracy using publicly available VIO approaches. The work in [29] presented a 3D map and ray-tracing technique to predict the positioning error of unmanned aerial vehicles (UAVs) in urban areas. Our previous research [3] demonstrated the degenerated positioning performance of several open-sourced LiDAR odometry in dense urban areas. However, no further investigation is performed on the degeneration using available sensors. Furthermore, few studies examined the impact of different illumination conditions such as days and nights, which is challenging for the VIO methods.

B. Multi-sensor Integration

Multiple sensors used in the integrated navigation system make it robust for handling challenging scenarios [30] where sensor failures or degradation partially. MSCKF [31] is a filter-based VIO framework for fast pose estimation. However, it is sensitive to time synchronization. Optimization-based methods such as VINS [10] can mitigate the time synchronization and optimize the history frame by using local maps or sliding windows to achieve the globally optimized solution. Another team [32] proposed an integrated sensor fusion strategy considering the GNSS geometric dilution of precision (GDOP) and K-mean clustering algorithm. Recent research has shown the effectiveness of the deep learning-based method [33] for multi-sensor localization and mapping. DeLS-3D [34] is a deep-learning framework that integrates a camera, GPS/IMU, and a semantic map to improve robustness and accuracy. An unsupervised method is proposed [35] to estimate the camera motion from stereo images and IMU data. Lvio-Fusion [36] adopted the reinforcement learning method to tune the weighting of different sensors by learning with the ground truth positioning. Unfortunately, it is still an open issue for the generalization of the deep learning-based method to various challenging scenarios.

C. Vehicle-Infrastructure Cooperation

Vehicle-infrastructure cooperation [37] is an emerging technology for future mobility. The work in [38] proposed a connected multi-agent positioning pipeline that utilizes information sharing and loop closure to enhance positioning accuracy. Another team introduced a V2X perception framework using a visual transformer via CARLA [21] simulator. A V2X collaborative motion planning [39] for autonomous vehicles is proposed to improve pedestrian safety in intersection areas. As the growth develops in smart cities, more efficient V2X services can be investigated for vehicle-infrastructure cooperation.

III. Methodology

A. Overview of the Proposed System

An overview of the proposed system in this paper is depicted in Fig. 2. The pipeline comprises two main components: (1) the error map estimation by the existing sensor-rich map update [20] vehicle; (2) the multi-sensor integration aided by the adaptive weighting from the latest error map. The system input includes the LiDAR point cloud, IMU measurements, camera images and GNSS measurements. Firstly, various sensor data is processed in the map update vehicle to generate the error map. Then the error maps are merged and transmitted to the other AVs in the same region through the RSU. With the help of the error information, the weighting of different odometry factors is further applied to the multi-sensor integration system to achieve robust estimation.

The following coordinates are clarified and followed by the remainder of the paper,

[1] https://github.com/DarrenWong/continuous_error_map

a) The LiDAR body coordinate $\{\cdot\}^L$: fixed at the center of the LiDAR sensor.
b) The IMU coordinate $\{\cdot\}^b$: fixed at the center of the IMU sensor.
c) The GNSS measurement coordinate $\{\cdot\}^G$: fixed at the center of the GNSS receiver.
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,

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

where $\mathbf{R}_{L_k}^W$ and $\mathbf{t}_{L_k}^W$ stand for the rotation matrix and translation matrix in the world coordinate, respectively.

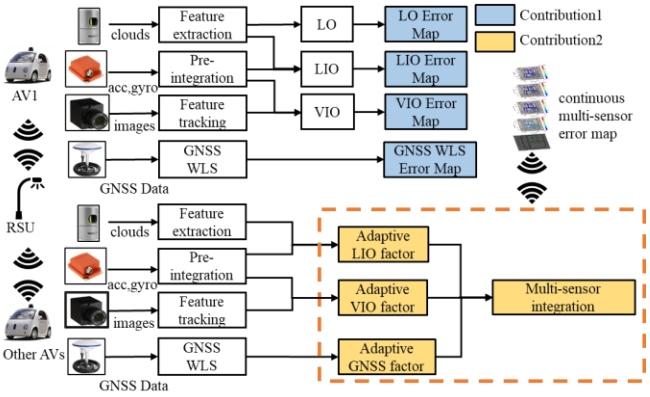

Fig. 2 Overview of the proposed system. LO, LIO, and VIO represent the LiDAR odometry, LiDAR-inertial odometry, and visual-inertial odometry, respectively.

B. Error Map Estimation

The section presents a brief introduction to the algorithms covered in the error map estimation for different sensors.

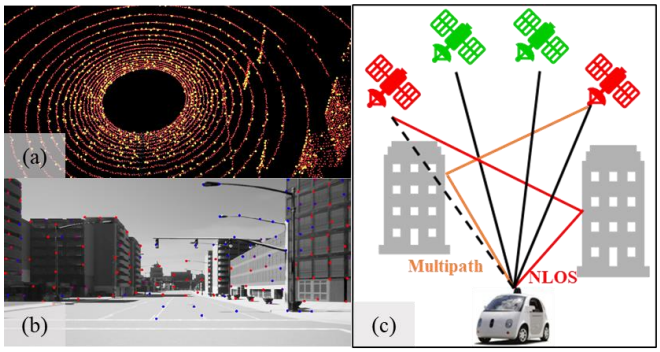

Fig. 3 Illustration of (a) Extracted edges and planar via LOAM. The red circles indicate the planar features, while the yellow points represent the edge features; (b) Feature tracking process of VINS-mono. The red feature indicates more tracked time compared to the blue feature; (c) GNSS NLOS and multipath in urban areas caused by the buildings.

LiDAR Odometry and Mapping (LOAM) [2] is a popular LiDAR odometry method. It first extracts edges and planar features from the raw point cloud according to the point curvature value. An illustration of the feature extraction can be shown in Fig. 3 (a). Scan-to-scan and scan-to-map are then conducted to minimize the residual of point-to-line and point-to-plane of the consecutive LiDAR frame and the map point cloud, respectively. The cost function of the scan-to-map can be simplified as,

$$\min_{\mathbf{T}_{L_{k+1}}^W} \left\{ \sum_{i=1}^{N_{k+1}^e} \left\| \mathbf{r}\left(\mathbf{X}_{k+1,i}^e, \mathbf{T}_{L_{k+1}}^W\right) \right\|^2 + \sum_{i=1}^{N_{k+1}^p} \left\| \mathbf{r}\left(\mathbf{X}_{k+1,i}^p, \mathbf{T}_{L_{k+1}}^W\right) \right\|^2 \right\} \quad (2)$$

where N_{k+1}^e and N_{k+1}^p indicate the numbers of edge points and plane points. $\mathbf{r}\left(\mathbf{X}_{k+1,i}^e, \mathbf{T}_{L_{k+1}}^W\right)$ and $\mathbf{r}\left(\mathbf{X}_{k+1,i}^p, \mathbf{T}_{L_{k+1}}^W\right)$ represent the residuals for the edge $\mathbf{X}_{k+1,i}^e$ and planar feature $\mathbf{X}_{k+1,i}^p$, respectively.

The idea of LiDAR inertial odometry via smoothing and mapping (LIO-SAM) [9] algorithm is to leverage the high-frequency update from IMU measurement to provide an initial guess for LiDAR odometry using factor graph optimization.

VINS-Mono [10] is a non-linear-optimization-based estimator for monocular visual-based odometry. It effectively integrates the measurement from the monocular camera and IMU using a sliding window factor graph. Fig. 3 (b) illustrates the feature tracking process using VINS-mono. The cost function of the VINS-mono is to minimize the residual from the IMU pre-integration, visual measurements and marginalization, respectively.

$$\min_{\chi} \left\{ \left\| \mathbf{r}_p - \mathbf{H}_p \chi \right\|^2 + \sum_{k \in \mathcal{B}} \left\| \mathbf{r}_{\mathcal{B}} \left(\hat{\mathbf{Z}}_{b_{k+1}}^{b_k}, \chi \right) \right\|_{\mathbf{P}_{b_{k+1}}^{b_k}}^2 + \sum_{(l,j) \in \mathcal{C}} \left\| \mathbf{r}_C \left(\hat{\mathbf{Z}}_l^{c_j}, \chi \right) \right\|_{\mathbf{P}_l^{c_j}}^2 \right\} \quad (3)$$

where χ is the state vector in the sliding windows. $\mathbf{r}_p$, $\mathbf{r}_{\mathcal{B}}$ *and* $\mathbf{r}_C$ denotes the residuals for marginalization, IMU pre-integration and visual reprojection, respectively. $\mathbf{H}_p$ represents the marginalization estimation matrix, while $\mathbf{P}(.)$ presents the covariance matrix of each term. $\hat{\mathbf{Z}}_{b_{k+1}}^{b_k}$ and $\hat{\mathbf{Z}}_l^{c_j}$ denotes the observations of IMU and features, respectively.

GNSS Weighted Least Squares (WLS) [40] is a well-known method that uses the WLS to estimate the positioning of the GNSS receiver. The pseudorange and carrier phase measurement can be used to compute the GNSS WLS solution. The object of the WLS is to minimize the weighted sum of the residuals from GNSS observation,

$$\mathbf{x}^G = (\mathbf{H}^T \mathbf{W} \mathbf{H})^{-1} \mathbf{H}^T \mathbf{W} \mathbf{z} \quad (4)$$

where $\mathbf{x}^G$ is the GNSS state to be estimated. $\mathbf{H}$ denotes the design matrix. $\mathbf{W}$ is the weighted matrix for observation, which is given based on the standard deviation of the satellite measurement. $\mathbf{z}$ is the measurement vector.

In short, the above-mentioned methods work well in the idea scenarios. However, the positioning performance of the LIO or VIO would suffer from the degeneration from the featureless environment or the unmodelled outliers such as dynamic objects in urban areas. Meanwhile, the performance of GNSS positioning is degraded due to NLOS and multipath, as shown in Fig. 3 (c). As the impact of dynamic objects is investigated in our previous research [14, 41], we focus on the sensor degeneration caused by various static scenes for heterogeneous sensors in urban areas. With the help of the map update vehicle which will collect the data regularly, we can generate the error map for the listed algorithms using the available sensors periodically.

C. *Error Map-aided Adaptive Sensor Integration*

This section presents error map-aided adaptive multi-sensor integration. LIO and VIO can provide locally accurate pose estimation while GNSS WLS can provide absolute positioning. Therefore, we loosely couple the relative LIO factor, relative VIO factor and absolute GNSS factor together with the corresponding weighting computed from the error map. The factor graph of the proposed method is shown in Fig. 4. The weighting of the LIO factor is obtained as,

$$\omega_{L_{k,k+1}} = \frac{e_{L_{k,k+1}} + e_{C_{k,k+1}}}{e_{L_{k,k+1}}} \tag{5}$$

which $\omega_{L_{k,k+1}}$ denotes the weighting parameters of the LIO factor. $e_{L_{k,k+1}}$ and $e_{C_{k,k+1}}$ represent the estimated relative translation error corresponding to the LIO and VIO at the same location.

Similarly, the weighting of the VIO factor is calculated as,

$$\omega_{C_{k,k+1}} = \frac{e_{L_{k,k+1}} + e_{C_{k,k+1}}}{e_{C_{k,k+1}}} \tag{6}$$

which $\omega_{V_{k,k+1}}$ indicates the weighting coefficient of the VIO factor. A larger weighting is assigned if the $e_{C_{k,k+1}}$ decreases.

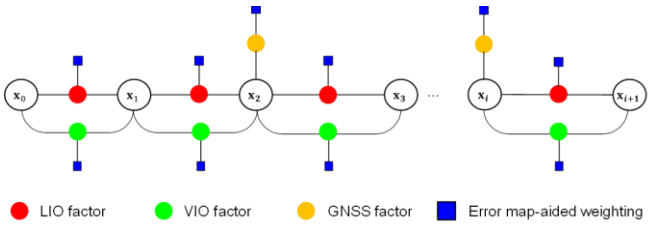

Fig. 4 Factor graph of the proposed method.

The work in [42] demonstrates that a similar positioning error can be observed in the same location. Therefore, we integrate the absolute GNSS measurement into the graph optimization if the prior error of the estimated position is less than 10 centimeters.

The cost function of the proposed error-map-aided multi-sensor integration can be expressed as:

$$\min_{\mathbf{T}^W_{b_{k+1}}} \left(\left\| \mathbf{r}_L\left(\omega_{L_{k,k+1}}, \mathbf{T}^W_{b_{k+1}}\right) \right\|^2_{\mathbf{P}_L} + \left\| \mathbf{r}_C\left(\omega_{C,k,k+1}, \mathbf{T}^W_{b_{k+1}}\right) \right\|^2_{\mathbf{P}_C} + \left\| \mathbf{r}_G\left(\omega_{G,k+1}, \mathbf{T}^W_{b_{k+1}}\right) \right\|^2_{\mathbf{P}_G} \right) \tag{7}$$

where $\omega_{L_{k,k+1}}$, $\omega_{C,k,k+1}$, $\omega_{G,k+1}$ indicate the weighting of the LIO factor, VIO factor and GNSS factor, respectively. $\mathbf{r}_C(\cdot)$ and $\mathbf{P}_{(\cdot)}$ present the residual and covariance matrix of each term.

IV. Experiment Results and Discussions

A. *Experiment Setup*

Sensor Setups: The effectiveness of the proposed method is validated using the CARLA [21] autonomous driving simulator and RUMS [22] urban GNSS simulator. This paper does not use existing datasets, such as KITTI [27] or UrbanNav [43], since they contain only one single agent under specific conditions. To investigate the impact of time conditions, we simulated the data in different time slots (noon, sunset and night) based on the urban environments. We do not include dynamic objects in this research as the impact of dynamic objects is extensively evaluated in our previous work [14, 41]. We collect 10 Hz LiDAR point cloud, 20 Hz images, 100 Hz IMU and 100 Hz ground truth positioning that explicitly considers the real-world noise in urban scenarios (example data are shown in Fig. 5) under a maximum speed of 30 km/h. Furthermore, we simulated the GNSS measurements using the realist RUMS simulator, which considers the impact of signal reflection and diffraction based on the 3D building model, as the CARLA can only simulate GNSS solutions with Gaussian noise. Assume the RSUs are distributed equally in this urban environment with an effective range of 500 meters. The RSUs are able to provide the error map service using the direct C-V2X (PC5) with a communication delay of less than 20 ms. The details of the sensor setup are provided in Table I and Fig. 5.

TABLE I. Sensor description.

Sensor	*Description*
1x LiDAR	360° Horizontal FOV, +10°~-30° vertical FOV, 100 meters, Velodyne[2] HDL 32 noise model, 5cm standard deviation on range measurement
1 x RGB Camera	FoV of 90°, 960x600, forward, default noise setting[3] in CARLA [21]
1 x IMU	100 Hz, 9-axis, Xsens MTi 10[4] noise model according to UrbanNav [43]
1 x GNSS	10 Hz, Ublox M8T[5] noise model with sign reflection from buildings

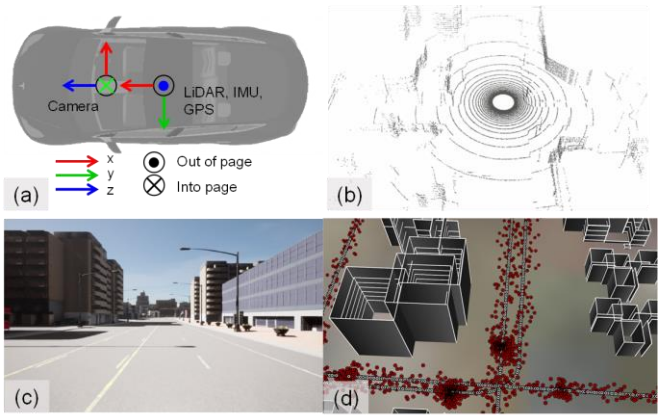

Fig. 5 (a) Sensor layout and coordinate; (b) The LiDAR point cloud; (c) The image captured by the camera; (d) The GNSS positioning near buildings simulated by RUMS. The red dot is the WLS solution while the white dot is the ground truth positioning.

The data is processed with a desktop with the specification of AMD Ryzen 5950X CPU, 64GB RAM and Nvidia 2080TI.

Evaluation Metrics: We first analyzed the performance of LIO, and VIO via the relative pose error (RPE) to local consistency of the trajectory with standard practice [44].

$$RPE_{j,k} = \left(\mathbf{T}_{j,k,gt}\right)^{-1}\left(\mathbf{T}_{j,k,est}\right) \tag{8}$$

where $\mathbf{T}_{j,k,gt}$ and $\mathbf{T}_{j,k,est}$ represent the relative ground truth pose and estimate pose between consecutive frames. The

[2] https://velodynelidar.com/
[3] https://carla.readthedocs.io/en/latest/ref_sensors/#rgb-camera
[4] https://www.movella.com/products/sensor-modules
[5] https://www.u-blox.com/en/product/neolea-m8t-series

performance of GNSS positioning is evaluated via the absolute trajectory error (APE).

$$ATE_k = \left\| trans(\mathbf{T}_{k,gt}) - trans(\mathbf{T}_{k,est}) \right\| \quad (9)$$

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

Second, the performance of the proposed method was evaluated via the ATE to investigate the global accuracy of the trajectory.

Evaluated Methods: To evaluate the performance of algorithms with different types of sensors, we use the following methods,

(1) LOAM [2]: The popular LiDAR odometry.

(2) LIO-SAM [9]: The state-of-the-art LiDAR-inertial odometry.

(3) VINS-Mono [10]: The popular visual-inertial odometry.

(4) GNSS WLS [40]: The weighted-lease square (WLS) GNSS positioning.

(5) LVI-SAM [45]: The state-of-the-art tightly coupled LiDAR-visual-inertial odometry. The depth information from LiDAR will help recover the scale of visual odometry.

(6) EMA-MS: The proposed error map-aided multi-sensor method with adaptive weighting. The weighting of the different sensors is computed using the continuous error estimated by the map update vehicle periodically.

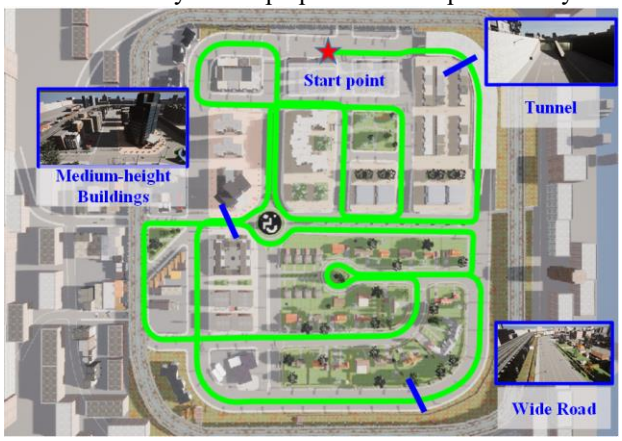

Fig. 6 The evaluated trajectory in the typical urban areas (modified Town03 in CARLA). It contains diverse urban scenarios, including medium-height buildings, wide roads, and tunnels.

B. Experimental Validation in Urban Areas

The evaluated data is collected in the typical urban areas (modified Town03 in CARLA) with a total length of 3.5km, as shown in Fig. 6. It involves medium-height buildings, wide roads, and tunnels. As shown in Table II, we collected three sets of data using one sensor-rich AV under different time conditions (*urban-noon1, urban-sunset1,* and *urban-night1*) for error map estimation. Afterward, we collected another three sets of data using one tested AV(*urban-noon2, urban1-sunset2,* and *urban1-night2*) with the same time slots to verify the effectiveness of the proposed method aided by the error map generated by the previous data.

TABLE II. The collected dataset for evaluation. Example scenarios can be found at the bottom the Fig. 1.

Data	*Description*
urban-noon 1	simulation data to generate error maps at noon
urban-noon 2	test data for error map-aided sensor fusion at noon
urban-sunset 1	simulation data to generate error maps at sunset
urban-sunset 2	test data for error map-aided sensor fusion at sunset
urban-night 1	simulation data to generate error maps at night
urban-night 2	test data for error map-aided sensor fusion at night

1) Error Map Estimation

Translation error is used as the criteria of the error map estimation. Table III demonstrates the positioning performance of the evaluated methods. The performance of LiDAR-only LOAM is worse than other methods, up to 0.520 meters in *urban-noon1*. With the help of the inertial sensors, the RPE error is reduced to 0.318 in *urban-noon1*. However, the RPE and ATE error is not enough for autonomous driving due to the degeneration in the tunnel (shown in Fig. 6) and accumulated drift. VINS achieves a better performance in terms of RPE and ATE without dynamic objects. However, the RPE error is increased from 0.066 to 0.077 meters as the illumination is worse from *urban-noon1* to *urban-night1* for the visual-based method. GNSS WLS obtained the best results in terms of ATE as the GNSS can provide absolute positioning without drift accumulation. But the noisy GNSS positioning (shown in Fig. 7) in urban areas and unavailable measurements in the tunnel cannot satisfy the requirements of the AV. In short, each sensor has its limitation, and the positioning results could be improved by integrating the sensors in a proper way.

Fig. 7 and Fig. 8 show the positioning error and the trajectories using four methods, respectively. The black curve represents the ground truth. The LiDAR-based method LOAM and LIO-SAM are degraded in the tunnel (shown in Fig. 7 (A)) thus it is hard to obtain a satisfactory positioning. The visual-based method VINS are smoother among different time slots but suffer a worse performance if under ill illumination, as shown in Fig. 7 (C). The GNSS WLS method provides average absolute performance compared to the LiDAR-based method and visual-based method. However, it suffers from large errors near high buildings, as shown in Fig. 7 (B).

TABLE III. Error evaluation of listed methods using different sensors.

Data	*Methods*	RPE RMSE (m)	ATE RMSE (m)
urban-noon 1	**LOAM**	0.520	54.672
	LIO-SAM	0.318	16.935
	VINS	0.045	12.076
	GNSS WLS	*	1.357
urban-sunset 1	**LOAM**	0.643	43.619
	LIO-SAM	0.381	26.324
	VINS	0.036	11.312
	GNSS WLS	*	1.171
urban-night 1	**LOAM**	0.536	43.982
	LIO-SAM	0.297	17.809
	VINS	0.077	12.914
	GNSS WLS	*	1.2205

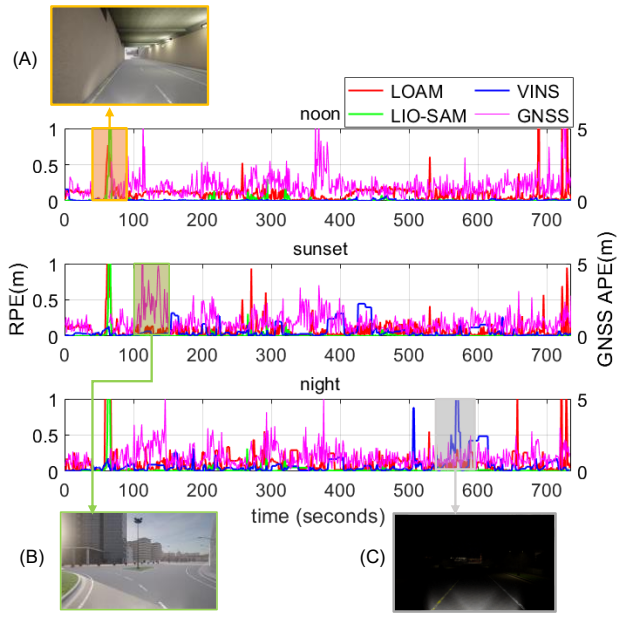

Fig. 7. Positioning errors of the listed methods (GNSS APE is a 2D error). The x-axis denotes the timestamp while the y-axis indicates the error, respectively. (A)-(C) are the marked area of the challenging scenarios for the LiDAR-based method, GNSS WLS and VINS, respectively.

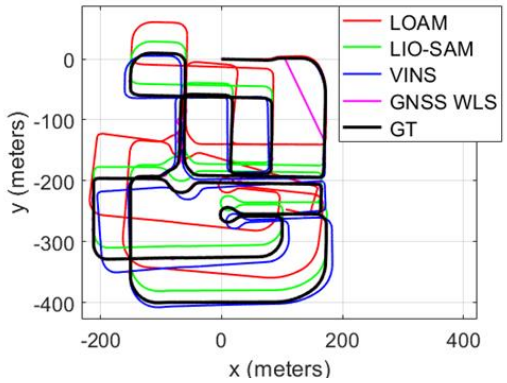

Fig. 8 Illustration of trajectories of the four methods under *urban-noon 1*. The trajectories of *urban-sunset 1* and *urban-night 1* can be found in the supplementary[6].

Fig. 9 demonstrates the results of the error maps from the listed methods using *urban-noon 1*. The error maps of *urban-sunset 1* and *urban-night 1* can be found in supplementary[6]. Similar to Figs. 7-8, it can observe large errors of LOAM and LIO-SAM in tunnel scenarios. Interestingly, it is complementary if we fuse the state estimation of VINS and the LiDAR-based method in the typical tunnel scene according to the error map of LIO-SAM and VINS. And the GNSS error map can be adopted to determine the areas with small errors, which can be used for sensor integration. As a result, each AV which is equipped with LiDAR, visual or GNSS sensors can benefit from the real-time error map estimation broadcasted by the RSU. One potential application for this is that other AVs could consider the error from the state estimation to optimize planning and control in the larger error areas. Furthermore, we can obtain better performance by utilizing the error map for the adaptive weighting of different sensors if the AV is equipped with all sensor-suit, as shown in the following section.

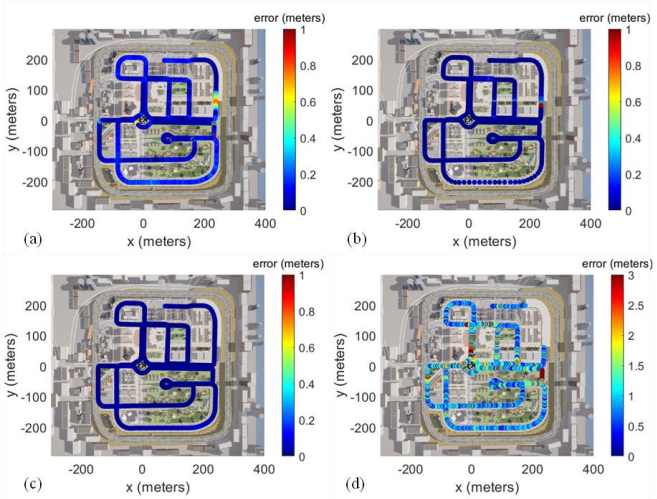

Fig. 9 Illustration of error map of the listed methods in data *urban-noon 1*. (a) LOAM; (b) LIO-SAM; (c) VINS;(d) GNSS WLS. The error maps of sunset and night can be found in supplymentary[6].

TABLE IV. Performance evaluation of listed methods. RMSE Improvement is calculated based on the evaluated method compared to the LIO-SAM.

Data	*Methods*	*ATE RMSE(m)*	ATE MEAN(m)	*RMSE Improvement*
urban-noon 2	**LIO-SAM**	19.749	17.723	/
	VINS	16.103	15.202	/
	LVI-SAM	1.414	2.095	92.84%
	EMA-MS w/o gnss	2.633	2.41	86.67%
	EMA-MS w/ gnss	**1.019**	**0.828**	93.67%
urban-sunset 2	**LIO-SAM**	24.172	21.678	/
	VINS	29.565	28.06	/
	LVI-SAM	2.549	2.456	89.45%
	EMA-MS w/o gnss	**1.507**	**1.398**	76.66%
	EMA-MS w/ gnss	1.732	1.568	92.83%
urban-night 2	**LIO-SAM**	21.376	19.168	/
	VINS	12.425	11.382	/
	LVI-SAM	3.658	3.523	82.89%
	EMA-MS w/o gnss	4.388	3.986	79.47%
	EMA-MS w/ gnss	**2.909**	**1.852**	86.39%

2) Positioning Results

The positioning performance for the test data is shown in Table IV. The proposed EMA-MS outperforms state-of-the-art methods with an ATE RMSE of 1.019 meters in *urban-noon 2*. An ATE RMSE of 19.749 meters was obtained using the existing LIO-SAM in the same scene, with a mean error of 17.723 meters. VINS achieves similar performance with an ATE RMSE of 16.103 meters. LVI-SAM demonstrates accurate performance in the test data because of the tightly coupled LIO and VIO systems. However, the error of LVI-SAM increases as the illumination condition worsens, especially in *urban-night 2*, which ATE RMSE from 1.414

[6]https://drive.google.com/file/d/1oU5vTzdirTi-G47VZLJpOLkvixxIrxeK/view?usp=share_link

meters to 3.658 meters. EMA-MS retained a better performance by loosely coupling the LIO and VIO factors with adaptive weighting. With the aid of the weighted GNSS constraint, the error is further reduced from 2.633 meters to 1.019 meters in *urban-noon 2*. Fig. 10 shows the trajectories using five methods. It can be seen that the estimated trajectory of the proposed EMA-MS method aligns well with the ground truth, especially in the zoom-in U-turn area. The trajectories of *urban-sunset 2* and *urban-night 2* can be found in supplementary[6].

In summary, the use of the continuous error map can significantly improve the positioning performance of the multi-sensor integration in urban areas. We believe that we can investigate more autonomous applications aided by sensor error maps.

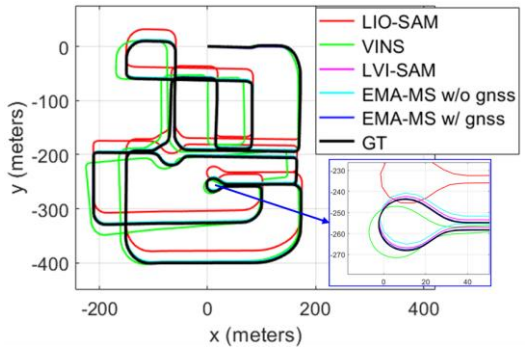

Fig. 10. Trajectories of the listed methods in Data *urban-noon 2*. The area marked in blue is the zoom-in of the U-turn area. The trajectories of *urban-sunset 2* and *urban-night 2* can be found in supplymentary[6].

V. Conclusion

With the growth development in smart cities, it is a huge demand for vehicle-infrastructure cooperation. This paper proposes sensor error maps that can be deployed as an additional layer on top of the high-definition map for the V2X application in urban areas. We demonstrate the multi-sensor integrated system can benefit from the prior error information obtained from error map estimation using the realistic simulator.

In the future, we will study how to improve the performance of autonomous driving by sensor sharing among the multi-vehicles and the RSUs. We will also evaluate our proposed method in Hong Kong C-V2X testbed [46] by collaborating with our industrial partner ASTRI, the largest C-V2X team in Hong Kong.

Acknowledgment

This work was supported by funded by the Guangdong Basic and Applied Basic Research Foundation (2021A1515110771) and the University Grants Committee of Hong Kong under the scheme Research Impact Fund (R5009-21). This work is also supported by the Innovation and Technology Fund (ITF) on the project "Adaptation of Cellular Vehicle-to-Everything (C-V2X) Technology in Hong Kong (ART/297CP_CV2XHK)", granted by the Hong Kong Innovation and Technology Commission.

References

[1] Z. Peng, S. Cheng, X. Li, K. Li, M. Cai, and L. You, "Dynamic Visual SLAM Integrated with IMU for Unmanned Scenarios," in *2022 IEEE 25th International Conference on Intelligent Transportation Systems (ITSC)*, 8-12 Oct. 2022 2022, pp. 4247-4253, doi: 10.1109/ITSC55140.2022.9922151.

[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] 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.

[4] 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.

[5] P. Geneva, K. Eckenhoff, W. Lee, Y. Yang, and G. Huang, "OpenVINS: A Research Platform for Visual-Inertial Estimation," in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, 31 May-31 Aug. 2020 2020, pp. 4666-4672, doi: 10.1109/ICRA40945.2020.9196524.

[6] X. Bai, W. Wen, and L.-T. Hsu, *Performance Analysis of Visual/Inertial Integrated Positioning in Typical Urban Scenarios of Hong Kong*. 2020.

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

[8] 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.

[9] 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.

[10] T. Qin, P. Li, and S. Shen, "VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator," *IEEE Transactions on Robotics,* vol. 34, no. 4, pp. 1004-1020, 2018, doi: 10.1109/TRO.2018.2853729.

[11] W. Wen, X. Bai, Y. C. Kan, and L. T. Hsu, "Tightly Coupled GNSS/INS Integration via Factor Graph and Aided by Fish-Eye Camera," *IEEE Transactions on Vehicular Technology,* vol. 68, no. 11, pp. 10651-10662, 2019, doi: 10.1109/TVT.2019.2944680.

[12] 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.

[13] 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.

[14] X. Bai, W. Wen, and L.-T. Hsu, "Degeneration-Aware Outlier Mitigation for Visual Inertial Integrated Navigation System in Urban Canyons," *IEEE Transactions on Instrumentation and Measurement,* vol. 70, pp. 1-15, 2021.

[15] A. Y. Alhilal, B. Finley, T. Braud, D. Su, and P. Hui, "Street Smart in 5G: Vehicular Applications, Communication, and Computing," *IEEE Access,* vol. 10, pp. 105631-105656, 2022.

[16] 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).

[17] D. Xue, N. Yang, X. Zhao, and Z. Wang, "Point-cloud map update for connected and autonomous vehicles based on vehicle infrastructure cooperation: framework and field experiments," in *2021 IEEE International Intelligent Transportation Systems Conference (ITSC)*, 2021: IEEE, pp. 2062-2067.

[18] A. Alhilal, B. Finley, T. Braud, D. Su, and P. Hui, "Distributed vehicular computing at the dawn of 5G: A survey," *arXiv preprint arXiv:2001.07077,* 2020.

[19] J. Barrachina *et al.*, "A V2I-based real-time traffic density estimation system in urban scenarios," *Wireless Personal Communications,* vol. 83, pp. 259-280, 2015.

[20] A. Boubakri, S. M. Gammar, M. B. Brahim, and F. Filali, "High definition map update for autonomous and connected vehicles: A survey," in *2022 International Wireless Communications and Mobile Computing (IWCMC)*, 30 May-3 June 2022 2022, pp. 1148-1153, doi: 10.1109/IWCMC55113.2022.9825276.

[21] 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.

[22] G. Zhang, B. Xu, H.-F. Ng, and L.-T. Hsu, "GNSS RUMS: GNSS realistic urban multiagent simulator for collaborative positioning research," *Remote Sensing,* vol. 13, no. 4, p. 544, 2021.

[23] Q. Zou, Q. Sun, L. Chen, B. Nie, and Q. Li, "A Comparative Analysis of LiDAR SLAM-Based Indoor Navigation for Autonomous Vehicles," *IEEE Transactions on Intelligent Transportation Systems,* vol. 23, no. 7, pp. 6907-6921, 2022, doi: 10.1109/TITS.2021.3063477.

[24] K. Kobayashi and N. Kubo, "Prediction of Real-Time Kinematic Positioning Availability on Road Using 3D Map and Machine Learning," *International Journal of Intelligent Transportation Systems Research,* pp. 1-16, 2023.

[25] J. Chang, L. Zhang, L.-T. Hsu, B. Xu, F. Huang, and D. Xu, "Analytic Models of a Loosely Coupled GNSS/INS/LiDAR Kalman Filter Considering Update Frequency Under a Spoofing Attack," *IEEE Sensors Journal,* vol. 22, no. 23, pp. 23341-23355, 2022.

[26] M. Elhousni and X. Huang, "A survey on 3d lidar localization for autonomous vehicles," in *2020 IEEE Intelligent Vehicles Symposium (IV)*, 2020: IEEE, pp. 1879-1884.

[27] 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.

[28] J. Delmerico and D. Scaramuzza, "A Benchmark Comparison of Monocular Visual-Inertial Odometry Algorithms for Flying Robots," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, 21-25 May 2018 2018, pp. 2502-2509, doi: 10.1109/ICRA.2018.8460664.

[29] G. Zhang and L.-T. Hsu, "A new path planning algorithm using a GNSS localization error map for UAVs in an urban area," *Journal of Intelligent & Robotic Systems,* vol. 94, pp. 219-235, 2019.

[30] R. Jin, Y. Wang, Z. Gao, X. Niu, L.-T. Hsu, and J. Liu, *DynaVIG: Monocular Vision/INS/GNSS Integrated Navigation and Object Tracking for AGV in Dynamic Scenes*. 2022.

[31] A. I. Mourikis and S. I. Roumeliotis, "A multi-state constraint Kalman filter for vision-aided inertial navigation," in *Proceedings 2007 IEEE international conference on robotics and automation*, 2007: IEEE, pp. 3565-3572.

[32] Y. Wang, R. Sun, Q. Cheng, and W. Y. Ochieng, "Measurement quality control aided multi-sensor system for improved vehicle navigation in urban areas," *IEEE Transactions on Industrial Electronics,* pp. 1-10, 2023, doi: 10.1109/TIE.2023.3288188.

[33] C. Chen, B. Wang, C. X. Lu, N. Trigoni, and A. Markham, "A survey on deep learning for localization and mapping: Towards the age of spatial machine intelligence," *arXiv preprint arXiv:2006.12567,* 2020.

[34] P. Wang, R. Yang, B. Cao, W. Xu, and Y. Lin, "Dels-3d: Deep localization and segmentation with a 3d semantic map," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 5860-5869.

[35] L. Han, Y. Lin, G. Du, and S. Lian, "Deepvio: Self-supervised deep learning of monocular visual inertial odometry using 3d geometric constraints," in *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2019: IEEE, pp. 6906-6913.

[36] Y. Jia *et al.*, "Lvio-fusion: A self-adaptive multi-sensor fusion slam framework using actor-critic method," in *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2021: IEEE, pp. 286-293.

[37] H. Yu *et al.*, "Dair-v2x: A large-scale dataset for vehicle-infrastructure cooperative 3d object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 21361-21370.

[38] 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, doi: 10.1109/ACCESS.2020.3006210.

[39] Y. Shi, Y. Pan, Z. Zhang, Y. Li, and Y. Xiao, "A 5G-V2X based collaborative motion planning for autonomous industrial vehicles at road intersections," in *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 2018: IEEE, pp. 3744-3748.

[40] P. Groves, *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems, Second Edition*. 2013.

[41] 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.

[42] G. Zhang, H.-F. Ng, W. Wen, and L.-T. Hsu, "3D mapping database aided GNSS based collaborative positioning using factor graph optimization," *IEEE Transactions on Intelligent Transportation Systems,* vol. 22, no. 10, pp. 6175-6187, 2020.

[43] 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.

[44] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of RGB-D SLAM systems," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2012: IEEE, pp. 573-580.

[45] T. Shan, B. Englot, C. Ratti, and D. Rus, "LVI-SAM: Tightly-coupled Lidar-Visual-Inertial Odometry via Smoothing and Mapping," in *2021 IEEE International Conference on Robotics and Automation (ICRA)*, 30 May-5 June 2021 2021, pp. 5692-5698, doi: 10.1109/ICRA48506.2021.9561996.

[46] 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).