Continuous Error Map Aided Adaptive Multi-Sensor Integration for Connected Autonomous Vehicles in Urban Scenarios

Feng Huang[1], Weisong Wen[1*], Guohao Zhang[1], Dongzhe Su[2], Yulong Huang[3]

***Abstract*—Precise multi-sensor integrated positioning is essential for autonomous vehicles (AVs) in urban environments. One of the key challenges in multi-sensor fusion is accurately estimating the weights of heterogeneous sensor data. With the emergence of Cellular Vehicle-to-Everything (C-V2X) technology and smart roadside infrastructure (RSIs), these systems can collaborate to provide enhanced and reliable services to connected vehicles. Motivated by this, our research explores the use of sensor error maps for heterogeneous sensor measurements under varying environmental conditions to improve the positioning accuracy of connected AVs in complex urban areas. We propose a multi-sensor integrated positioning system that utilizes error map information generated by sensor-rich connected autonomous vehicles (CAVs). This error information is shared with RSIs and then distributed to nearby CAVs within the same region. Sensors with higher estimated errors are assigned lower weights, as determined by the error maps. To validate the proposed approach, we conducted experiments both day and night in a realistic simulation environment as well as in the Hong Kong C-V2X testbed. The results demonstrate that the use of continuous error maps significantly enhances the performance of multi-sensor integrated positioning. The data and implementation of our system are available at https://github.com/DarrenWong/continuous_error_map.**

***Index Terms*—Continuous Error Map, C-V2X, Multi-Sensor Integrated Positioning, Urban Scenarios.**

NOMENCLATURE

$\mathbf{T}_k$	The transformation in 6 degrees of freedom (DoF) at timestamp k, can be divided into translational vector $\mathbf{t}_k$ and rotational matrix $\mathbf{R}_k$.
$\{\cdot\}^L$	Value expressed in the LiDAR body frame.
$\{\cdot\}^G$	Value expressed in the GNSS frame, which refers to the Earth-Centered, Earth-Fixed (ECEF) frame.
$\{\cdot\}^{EN}$	Value expressed in the East-North-Up (ENU) frame, commonly used to represent the state on the Earth's surface relative to a local reference point.
$\{\cdot\}^C$	Value expressed in the camera frame.
$\{\cdot\}^b$	Value expressed in the IMU frame.
$\{\cdot\}^W$	The world frame which aligned with the initial position of the vehicle. For example, transformation $\mathbf{T}^W_{L_{k+1}}$ representing the pose of the LiDAR sensor at time $t+1$ with respect to the world frame.
$\mathbf{r}(\cdot)$	The residual of corresponding term.
$e(\cdot)$	Error associated with the specified method, expressed in double precision.
$\omega(\cdot)$	The weighting coefficient associated with the specified method.

This work was supported in part by the Innovation and Technology Fund under the project "Safety-Certified Multi-Source Fusion Positioning for Autonomous Vehicles in Complex Scenarios (ZPE8)", in part by the Germany/Hong Kong Joint Research Scheme under the project "Maximum Consensus Integration of GNSS and LiDAR", and in part by the Research Center of Deep Space Exploration (RC-DSE) under the project "Multi-Robot Collaborative Operations".

[1]Feng Huang, Weisong Wen and Guohao Zhang are with the Department of Aeronautical and Aviation Engineering and the Research Centre for Unmanned Autonomous Systems, The Hong Kong Polytechnic University, Hong Kong SAR, China. (e-mail: darren-f.huang@connect.polyu.hk, welson.wen@polyu.edu.hk, gh.zhang@polyu.edu.hk).

[2]Dongzhe Su is with the Hong Kong Applied Science and Technology Research Institute (ASTRI), Hong Kong. (e-mail: dzsu@astri.org).

[3]Yulong Huang is with the College of Intelligent Systems Science and Engineering, Harbin Engineering University, Harbin 150001, China (e-mail: heuedu@163.com).

*Correspondence author: Weisong Wen.

I. INTRODUCTION

RECENT advancements of multi-sensor integration for autonomous vehicles (AVs) highlight the significant potential of smart mobility solutions. Simultaneous localization and mapping (SLAM) [1] is a widely used technique in most AVs. Light detection and ranging (LiDAR), an essential sensor for AVs, captures dense clouds of the surroundings, making it an indispensable module for odometry and mapping [2]. However, the performance of LiDAR-based odometry can be challenged in urban areas [3] due to dynamic objects and insufficient features. Visual-based method, another widely adopted technique, utilizes feature matching for state estimation. But its reliability is highly sensitive to lighting conditions and visual features—making it less effective at night or in environments with poor illumination [4]. The Global Navigation Satellite System (GNSS) offers absolute positioning but suffers from degraded performance due to multipath effects [5] in urban areas. Although numerous research [6]–[8] have been proposed to mitigate the multipath effects, the GNSS solutions are still unreliable in complex urban areas. Consequently, relying on a single sensor for AV localization is insufficient, prompting significant research into multi-sensor integration.

Weighting estimation for multi-sensor integration can be challenged in complex scenarios: To address the challenges

of the single sensor-based method, numerous studies have focused on developing robust sensor integration [9], [10]. However, accurately estimating the weighting coefficients for sensor fusion remains a significant challenge, particularly in complex urban scenarios. Traditional methods like the Kalman filter [11] estimate sensor weights via the Kalman gain, assuming linear dynamics and Gaussian noise, but they are less effective under sudden changes in sensor reliability. Particle filters [12] handle nonlinearities and multimodal distributions but are computationally expensive and sensitive to parameter tuning. Reference [13] integrates LiDAR, wheel odometry, and IMU data, employing a dynamic covariance matrix that adjusts according to the quantity of captured features. The MSCKF framework [14], a filter-based Visual-Inertial Odometry (VIO) method, enables rapid pose estimation but is sensitive to temporal synchronization errors. To mitigate this, graph optimization approaches such as VINS are proposed to refine historical data using sliding windows to obtain an optimized solution. Recent studies have shown the potential of deep learning-based methods [15] for efficient weighting estimation in multi-sensor integration. Reference [16] is a deep learning-based system that fuses data from cameras, GPS/IMU, and semantic maps to enhance positioning accuracy. Another study [17] introduces a tightly coupled LiDAR-visual-inertial SLAM system with progressive multi-modal semantic segmentation using KITTI [18] and nuScenes [19] dataset. However, its segmentation performance could be affected by varying illumination conditions in complex urban environments. Additionally, an end-to-end sensor fusion method [20] was introduced to learn features and relationships from visual-inertial sensor data. Another study [21] proposes an adaptive fusion framework that utilizes deep reinforcement learning. However, the effectiveness of deep learning-based methods relies on having a sufficient amount of training data and can be computationally expensive during model training. Additionally, the generalization of these methods for estimating sensor weighting coefficients in diverse and challenging scenarios remains an unresolved issue. Therefore, it is necessary to enhance the performance of multi-sensor integrated positioning in urban areas by accurately estimating the weighting of different measurements.

Error estimation for multi-sensor in urban areas: Several studies [4], [22] have explored the sensor error matrix associated with different algorithms in urban environments. For instance, the work in [23] offers a review of errors in various LiDAR SLAM methods, using the KITTI dataset. Similarly, in [24], several LiDAR-Inertial Odometry (LIO) methods are evaluated using the UrbanNav [25] dataset, which includes tall buildings and multiple dynamic objects. These LiDAR-based evaluations have demonstrated that performance degradation can occur in featureless environments, such as tunnels. A study by the University of Zurich and ETH Zurich [26] assessed pose accuracy for flying robots using publicly available VIO approaches, providing a reference for balancing accuracy and computational cost across different hardware platforms. Another work [27] evaluated the performance of various visual-based methods in urban areas. The evaluation of visual-based methods highlights challenges such as the loss of feature tracking under low illumination or overexposure conditions. The study in [28] proposed to predict UAV positioning errors in urban areas based on building models, integrating the predicted errors into an A* planning algorithm to generate safer flight paths. However, this approach relies heavily on accurate and up-to-date 3D building data, which may not always be available. In 2023, [29] proposed a random forest-based method for estimating GNSS pseudorange errors in urban areas, where non-line-of-sight (NLOS) and multipath effects are significant sources of error. Our previous work [3] highlighted the degraded performance of several open-source LiDAR odometry methods in complex urban canyons, though no further analysis was performed using additional sensors. To sum up, limited research has explored the impact of varying environmental conditions with heterogeneous sensor measurements. Additionally, the effectiveness of error estimation for multi-sensor integration to enhance smart mobility remains unclear.

Expanding deployment of C-V2X and Vehicle-Infrastructure Cooperation: In recent years, the significant development in the C-V2X has drawn considerable interest because of its capability to enable low-latency information sharing for ITS. According to the CAV Development Study [30] conducted by Deloitte China and Hong Kong ASTRI in April 2024, roadside infrastructure and sensors are fundamental components in the CAV ecosystems. With the advancement of 5G technology and RSIs, vehicle-infrastructure collaboration has attracted considerable attention for its potential in intelligent transportation systems. The study in [31] focuses on vehicle reidentification using both vehicle-mounted and infrastructure-based cameras to capture shared vehicle features within the CARLA simulation. Another work proposed a V2X planning strategy that jointly optimizes the system using a block coordinate descent framework. However, both studies were only evaluated on simulated data may limit generalizability to real-world scenarios. In 2024, research [32] demonstrated improved V2X cooperative perception performance through Transformer-based models, evaluated in both simulated V2X environments and the real-world DAIR-V2X [33]. However, the approach relies on accurate pose exchange between CAVs, which is difficult to achieve in complex urban environments due to signal degradation and localization uncertainty. Meanwhile, study [34] investigates vehicle-to-vehicle (V2V) localization using multi-modal data collected in intercity and rural environments based on the DeepSense 6G dataset [35]. But in those scenarios, 90% of the GNSS RTK positioning results achieve errors below 0.5 meters, an accuracy level that is challenging to maintain in complex urban areas. Our previous work [36] leveraged roadside LiDAR constraints to enhance the performance of onboard LIO in real-world

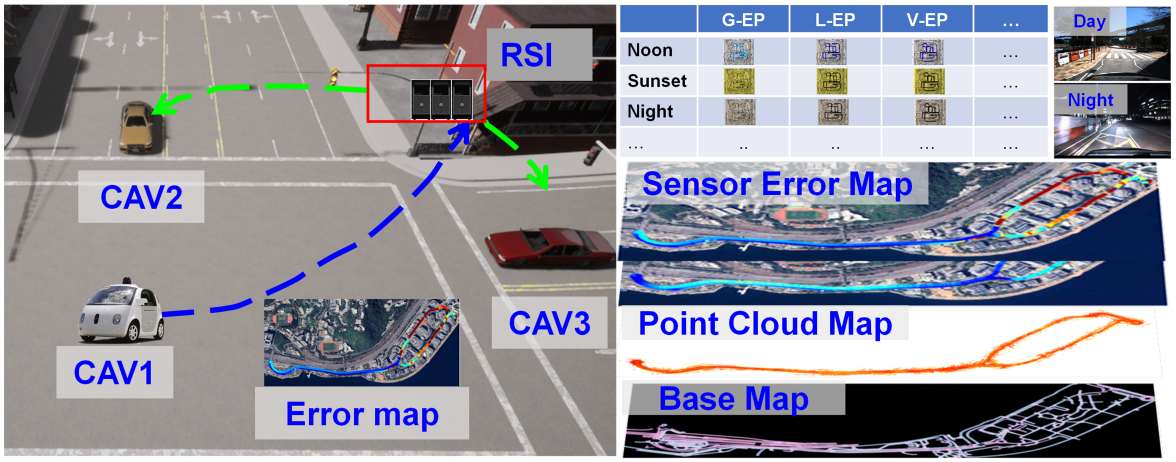

Fig. 1. *Left*: Illustration of the error map being transmitted through the RSIs in urban areas. CAV1, mounted with a full sensor-rich suite (e.g., LiDAR, camera, GNSS, and high-end sensors capable of providing ground truth solutions), periodically evaluates sensor errors. CAV2 and CAV3, the nearby CAVs, receive this real-time sensor error information to assist with their navigation using their available sensors. *Right*: Illustration of the error map functioning as an additional layer over the base vector map and point cloud map for intelligent transportation applications. These error maps are generated in real-time by CAV1. G-EP, L-EP, and V-EP represent the GNSS error map, LiDAR error map, and Visual error map, respectively.

scenarios. As smart mobility evolves, there is significant potential to explore more efficient V2X services within the context of vehicle-infrastructure collaboration in urban areas.

Inspired by the efficient data transmission capabilities of C-V2X technology among CAVs [37], we investigate the generation of continuous sensor error maps and their application in V2X-enabled multi-sensor integration. The concept behind this work is illustrated in Fig. 1. To estimate errors in heterogeneous sensor measurements, a fully sensor-equipped, high-accuracy vehicle, such as a map update vehicle or Robobus (CAV1, shown on the left of the figure), is periodically deployed to operate and collect multi-sensor data in urban environments. The vehicle uses high-precision ground truth [25] (obtained via centimeter-level integrated GNSS/INS systems) to evaluate the performance of individual sensors (e.g., GNSS, LIO, VIO) by comparing their outputs against this reference. The resulting sensor-specific errors are aggregated into an error matrix, which is uploaded to intelligent edge platforms. These platforms then process and integrate the data into heterogeneous sensor error maps that reflect spatially varying performance across modalities. Connected vehicles (CAV2 and CAV3, shown on the left of Fig. 1) operating within the same area can access these error maps broadcast by RSIs. Connected vehicles mounted with limited sensors can enhance their navigation capabilities by referencing the error map, while autonomous vehicles with comprehensive sensors can improve sensor integration by incorporating the measurement weightings extracted by the error information. These sensor error maps generated periodically by sensor-rich CAVs, serve as an additional layer upon the base map and point cloud map (right of Fig. 1) for intelligent transportation applications.

To validate our proposed method, heterogeneous onboard sensor data in urban day and night conditions is collected both in high-fidelity simulated platform [38], [39] and the Hong Kong C-V2X testbed. The major contributions of this study are threefold,

1) This paper presents the continuous error map estimation using various onboard sensors with varying confidence levels across different urban environments. These sensor error maps can support a range of downstream navigation tasks in vehicle-infrastructure cooperation.
2) This paper introduces an adaptive multi-sensor integration system that leverages error information from the error map estimation. The system computes weighting coefficients based on the errors identified in the corresponding error maps.
3) Extensive simulated data and Hong Kong C-V2X testbed data are conducted across both day and night conditions, demonstrating the enhanced positioning performance aided by the error map estimation. The implementation is open-sourced at Github https://github.com/DarrenWong/continuous_error_map.

The structure of the remaining sections is outlined as follows. Section II presents the methodology of the error map estimation and its utilization to aid sensor integration. A detailed description of the experimental results, including both simulation and C-V2X testbed validations is shown in Section III, followed by the conclusion and future work discussed in Section IV.

II. Methodology

A. System Overview

Fig. 2 provides an overview of the proposed pipeline. The pipeline consists of two modules: (1) error map generation using a sensor-rich CAV, and (2) multi-sensor fusion enhanced by adaptive weighting based on the sensor error maps. The system processes the input of LiDAR scans, IMU measurements, visual images, and GNSS data from CAVs. First, sensor data from the CAV1 (e.g., map update vehicle) are processed. To produce the sensor error maps. These error maps are then sent to the RSI via the V2X network. The error maps are subsequently aggregated and distributed to other CAVs (e.g., CAV2 and CAV3 in Fig. 1) in the same area through the roadside unit (RSU). By utilizing error map data, CAVs adaptively adjust the weighting of heterogeneous sensor modalities to enhance sensor fusion accuracy and ensure more robust positioning. Although sensor configurations may vary between vehicles, the error maps predominantly capture environment-dependent factors, such as GNSS signal occlusion and multipath, which consistently impact CAVs operating within the same geographic area [40]. The coordinate systems used in this work are depicted in Fig. 3. The extrinsic parameters among sensors are pre-calibrated, while the extrinsic matrix between the ENU and IMU body frames is estimated using the Attitude and Heading Reference System (AHRS). For details on the conversion between the ECEF and ENU frames, please refer to [41].

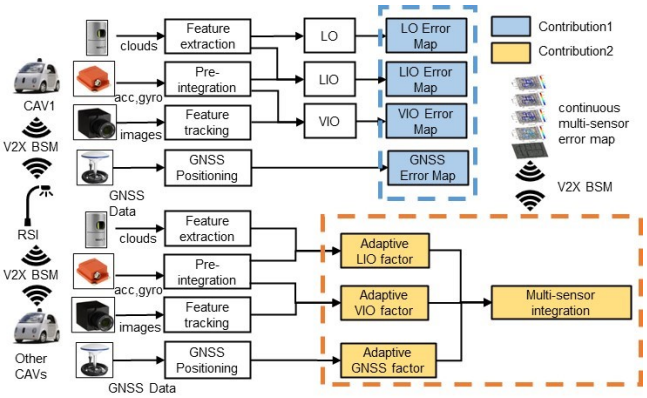

Fig. 2. The system overview. BSM is short for basic safety message which is defined in V2X message.

B. Error Map Estimation

This section presents an overview of representative methods used in error map estimation in heterogeneous sensor measurements, including LiDAR-based LOAM [2] and LIO-SAM [42], vision-based VINS [43], and GNSS positioning [44].

LOAM [2] is a commonly used method for LiDAR odometry. The example of feature extraction is illustrated in Fig. 4 (a). The scan-to-modal scheme is then adapted to

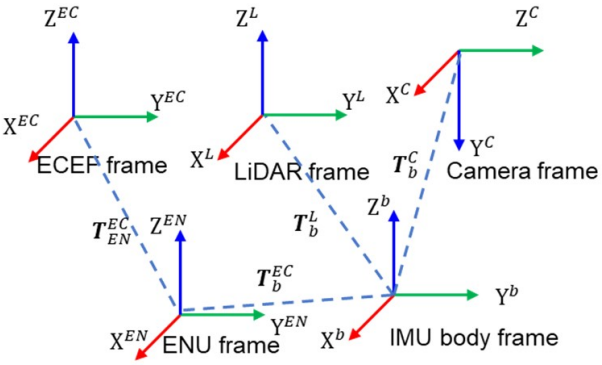

Fig. 3. Overview of the coordinate system.

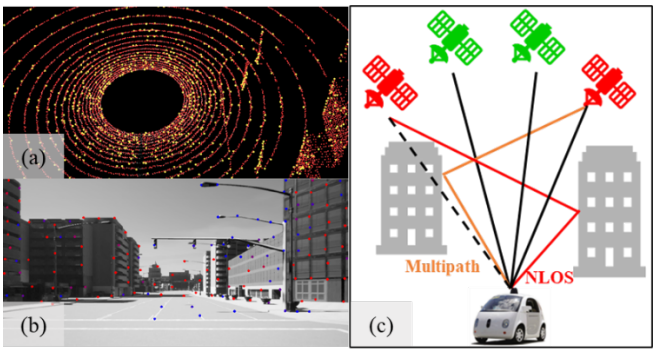

Fig. 4. Illustration of (a) Features extracted by LOAM; (b) VINS-Mono's feature tracking process, where red features indicate longer tracking duration compared to blue features; (c) GNSS NLOS and multipath interference in urban environments due to reflections of surrounding buildings.

optimize the residuals between the scan and the map. The objective function for the scan-to-modal are as follows,

$$\min_{\mathbf{T}^W_{L_{k+1}}} \left\{ \sum \left\| \mathbf{r_e} \left(\mathcal{F}^e_{k+1}, \mathbf{T}^W_{L_{k+1}} \right) \right\|^2 + \sum \left\| \mathbf{r_p} \left(\mathcal{F}^p_{k+1}, \mathbf{T}^W_{L_{k+1}} \right) \right\|^2 \right\} \tag{1}$$

where $\mathbf{r_e}\left(\mathcal{F}^e_{k+1},\ \mathbf{T}^W_{L_{k+1}}\right)$ and $\mathbf{r_p}\left(\mathcal{F}^p_{k+1},\ \mathbf{T}^W_{L_{k+1}}\right)$ represent the residuals for the edge features in $\mathcal{F}^e_{k+1}$ and planar points $\mathcal{F}^p_{k+1,i}$, respectively.

LIO-SAM [42] utilizes high-frequency IMU measurements to estimate an initial guess for LiDAR-based scan mapping, which is then refined based on graph optimization.

VINS-Mono [43] is an optimization-based visual odometry that integrates data from visual and inertial sensors using a sliding window-based graph. Its feature tracking process is shown in Fig. 4-(b). The objective function is designed to minimize the residuals from IMU pre-integration, visual features, and marginalization.

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

where χ refer the state within the sliding windows. The residual $\mathbf{r}_p$, $\mathbf{r}_{\mathcal{B}}$ and $\mathbf{r}_C$ correspond to marginalization, IMU pre-integration, and visual reprojection, respectively. $\mathbf{H}_p$ represents the matrix for marginalization estimation, and $\mathbf{P}(.)$ is the covariance matrix of each residuals. $\widehat{\mathbf{Z}}_{b_{k+1}}^{b_k}$ and $\widehat{\mathbf{Z}}_l^{\mathbf{c_j}}$ indicate the IMU and camera measurements, respectively.

The weighted least squares (WLS) method [44] is a widely used technique used to resolve the position of a GNSS receiver. This method incorporates both pseudorange and carrier phase measurements to compute the GNSS WLS solution, aiming to optimize the weighted sum of residual values from the GNSS observations.

$$\mathbf{P}_G=\left(\mathbf{H}^T\mathbf{W}\mathbf{H}\right)^{-1}\mathbf{H}^T\mathbf{W}\mathbf{z} \tag{3}$$

where $\mathbf{P}_G$ denotes the GNSS state to be resolved, which normally represents the vehicle positioning. $\mathbf{H}$ refers to the Jacobian matrix derived from the observation model. $\mathbf{W}$ denotes the weighted matrix for the observations, which is based on the satellite measurement's standard deviation. $\mathbf{z}$ denotes the observed measurements. GNSS-RTK positioning enhances WLS solutions by eliminating systematic errors using the double-differenced (DD) technique [44], which compares observations from the reference and user stations. The GNSS-RTK process involves two steps: the first step estimates the float solution using WLS or Extended Kalman Filter [44] to handle dynamic system states, and the second step resolves integer ambiguity using the LAMBDA method [45] to obtain a fixed solution. For implementation details on GNSS-RTK positioning, refer to RTKLIB [46].

In summary, the methods mentioned above perform effectively in ideal conditions. However, both LIO and VIO experience degraded positioning accuracy in featureless environments or when confronted with unmodeled outliers such as moving vehicles in urban areas. Similarly, GNSS positioning is degraded by NLOS and multipath interference, as illustrated in Fig. 4 (c). Since the influence of dynamic objects has been intensively evaluated in our previous work [3], [47], this study focuses on sensor performance degradation due to varying environments in urban areas across heterogeneous sensors.

Given LiDAR odometry and LIO which provides high-rate relative measurements, LO/LIO error e_L denotes the translation part of the Relative Pose Error (RPE) of the ground truth state $\mathbf{T}_{j,k,gt}$ and estimated state $\mathbf{T}_{j,k,est}$ between timestamp j and k.

$$e_L=\left\|trans\left(\left(\mathbf{T}_{L_j,L_k,gt}\right)^{-1}\left(\mathbf{T}_{L_j,L_k,est}\right)\right)\right\| \tag{4}$$

Similarly, the VIO error e_C can be expressed as,

$$e_C=\left\|trans\left(\left(\mathbf{T}_{C_j,C_k,gt}\right)^{-1}\left(\mathbf{T}_{C_j,C_k,est}\right)\right)\right\| \tag{5}$$

Regarding the GNSS positioning which provides the absolute positioning, GNSS error e_G denotes the absolute position error (APE) of the state at timestamp j,

$$e_G=\left\|trans(\mathbf{T}_{j,\ gt})-trans(\mathbf{T}_{j,est})\right\| \tag{6}$$

$trans(\cdot)$ denotes the translation component of the state to be estimated.

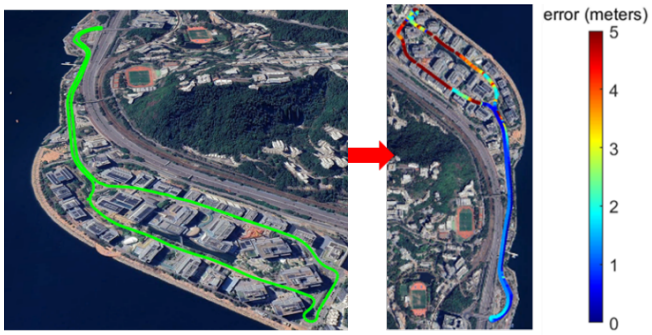

Fig. 5. Illustration of the GNSS error maps. The left green curve indicates the ground truth estimation by the sensor-rich map update vehicle while the right color curve indicates the estimated error of GNSS positioning. Note that the orientation of the trajectory is rotated to align with the ENU coordinates.

By utilizing the map update vehicle, which collects data at regular intervals, we can periodically generate error maps for the listed algorithms based on the data gathered from the available sensors. Fig. 5 demonstrates the GNSS positioning error map generated in the Hong Kong C-V2X testbed at a single route, which includes the timestamp, geographic information, and corresponding error. The matrix form $\mathbf{M}_t$ of the error maps at timestamp *t* can be expressed as,

$$\mathbf{M}_t=\begin{bmatrix}s_L & t_1 & x_1 & y_1 & z_1 & e_{L,1} & \dots\\ s_C & t_1 & x_1 & y_1 & z_1 & e_{C,1} & \dots\\ s_G & t_1 & x_1 & y_1 & z_1 & e_{G,1} & \dots\\ s_L & s_L & x_2 & y_2 & z_2 & e_{L,1} & \dots\\ \dots & \dots & \dots & \dots & \dots & \dots & \dots\end{bmatrix} \tag{7}$$

where s_L, s_C, s_G denotes the sensor type LiDAR, camera and GNSS, respectively. $e_{L,1}$, $e_{C,1}$, $e_{G,1}$ are the corresponding errors derived from different sensors using Equations 4, 5, and 6 at the timestamp t_1 and geographic position $\mathbf{T}_{1,gt}=[x_1,\ y_1,\ z_1\ ]$ in the world frame, respectively.

C. C-V2X Communication for Error Maps

C-V2X is developed to improve road safety, optimize traffic efficiency, and facilitate the emergence of new applications and services within the context of CAVs and ITS. However, there is currently no existing C-V2X message

specifically designed for transmitting error map information. To address this, the SAE J2735-defined basic safety message (BSM) [48] is introduced, as it contains essential data such as the timestamp, vehicle position, accuracy, component status, and other relevant travel information necessary for error maps. Additional information can also be included in the BSM message Part 1 & 2 if required. Fig. 1 illustrates the pipeline of C-V2X data communication incorporating error maps. In this system, sensor-rich vehicles transmit packed BSM data through the V2X OBU device to the RSI, which includes the RSU and edge computing capabilities, integrates the error maps, and subsequently broadcasts them to nearby connected vehicles that have an OBU installed, utilizing the V2X communication channel.

D. Error Map-aided Adaptive Sensor Integration

Since the error information is received from the RSI, nearby connected vehicle can utilize it to aid adaptive multi-sensor fusion in the same region. Fig. 6 depicts the process of obtaining the closest error matrix from the error information, considering the estimated trajectory under the world frame.

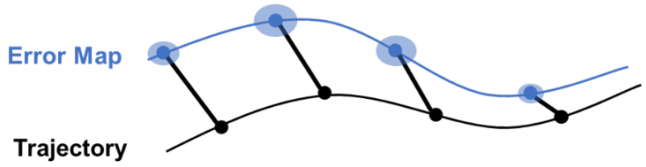

Fig. 6. Illustration of the error information extracted from the error map under world frame. For each estimated pose, it selects the nearest point in Euclidean space from the updated error map. The corresponding error value is then used to adaptively adjust the sensor fusion process for the current pose.

The weighting coefficient for the LIO factor is computed as,

$$\omega_{L_{k,k+1}} = \frac{e_{L_{k,k+1}} + e_{C_{k,k+1}}}{e_{L_{k,k+1}}} \tag{8}$$

$e_{L_{k,k+1}}$ and $e_{C_{k,k+1}}$ indicate the estimated RPE for LIO and VIO, respectively, extracted from error maps.

Similarly, the weighting coefficient for the VIO factor is determined as,

$$\omega_{C_{k,k+1}} = \frac{e_{L_{k,k+1}} + e_{C_{k,k+1}}}{e_{C_{k,k+1}}} \tag{9}$$

which $\omega_{C_{k,k+1}}$ denotes the weighting coefficient associated with the VIO factor, where a higher weight is applied as $e_{C_{k,k+1}}$ decreases. Meanwhile, the absolute GNSS measurement is incorporated into the graph optimization process when the prior positional error falls below an experimental threshold value m.

$$\omega_{G,k+1} = \begin{cases} 1, & e_{G_{k+1}} < m \\ 0, & e_{G_{k+1}} > m \end{cases} \tag{10}$$

Subsequently, the relative LIO and VIO factors, along with the absolute GNSS factor, are loosely integrated within a factor graph framework, with weights assigned based on the sensor error map. The factor graph of the proposed approach is depicted in Fig. 7. Note that the error map is continuously updated to reflect varying environmental conditions, using high-accuracy vehicles, such as map update vehicles or Robobuses, that are periodically deployed to operate and collect multi-sensor data in urban environments.

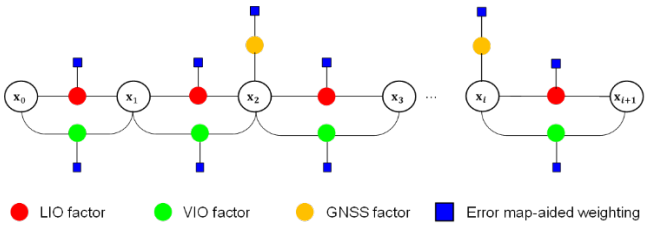

Fig. 7. Factor graph of the proposed method.

The cost function of the proposed method can be formulated as,

$$\begin{aligned} \min_{\mathbf{T}^W_{b_{k+1}}} \Bigg(& \left\| \mathbf{r}_L \left(\omega_{L_{k,k+1}}, \mathbf{T}^W_{b_{k+1}} \right) \right\|^2_{\mathbf{\Sigma}_L} \\ & + \left\| \mathbf{r}_C \left(\omega_{C,k,k+1}, \mathbf{T}^W_{b_{k+1}} \right) \right\|^2_{\mathbf{\Sigma}_C} \\ & + \left\| \mathbf{r}_G \left(\omega_{G,k+1}, \mathbf{T}^W_{b_{k+1}} \right) \right\|^2_{\mathbf{\Sigma}_G} \Bigg) \end{aligned} \tag{11}$$

where $\omega_{L_{k,k+1}}$, $\omega_{C,k,k+1}$, and $\omega_{G,k+1}$ denote the weighting coefficients assigned to the LIO, VIO, and GNSS factors, respectively. $\mathbf{r}_L(\bullet)$, $\mathbf{r}_C(\bullet)$, and $\mathbf{r}_G(\bullet)$ represent the residuals of the LIO, VIO, and GNSS factors after weighting. $\mathbf{\Sigma}_L$, $\mathbf{\Sigma}_C$, and $\mathbf{\Sigma}_G$ are the corresponding covariance matrices for each factor. This proposed method can adaptively mitigate the effect of degenerated sensor inputs and rely on more reliable ones in real time. The weighting strategy automatically downscales the contribution of unreliable readings in situations when one sensor shows a noticeably big error, such as in tunnel conditions where LO/LIO suffers from featureless and GNSS becomes unavailable. Furthermore, the system reverts to employing the original covariance matrices when all sensor modalities produce moderate residuals. Algorithm 1 summarizes the details of the error map-assisted multi-sensor fusion.

III. Performance Evaluation

A. Experiment Setup

Sensor Setups: The error map estimation and the proposed method are evaluated in simulated platforms and the data collected from the Hong Kong C-V2X testbed.

Simulated sensor data is generated using the CARLA [38] autonomous driving platform in combination with the GNSS Realistic Urban Measurement Simulator (RUMS) [39]. The

Algorithm 1 Error map-assisted sensor integrated positioning

Input: Point cloud $\hat{\mathcal{P}}^L$, image $\hat{\mathbf{I}}^C$, GNSS measurements $\hat{\phi}^G$, IMU data $\hat{\mathbf{Z}}^b$

Output: The optimal state $\mathbf{T}_{b_k}^W$

1: **Step 1:** Obtain the factor from different sensors.
2: **Step 1-1:** Obtain the LIO factor using $\hat{\mathcal{P}}^L$ and $\hat{\mathbf{Z}}^b$ based on (1).
3: **Step 1-2:** Obtain the VIO factor using $\hat{\mathbf{I}}^C$ and $\hat{\mathbf{Z}}^b$ based on (2).
4: **Step 1-3:** Obtain the GNSS unary factor using $\hat{\phi}^G$ based on (3).
5: **Step 2:** Extract error information from the error map and compute sensor weighting coefficients based on (8), (9) and (10).
6: **Step 3:** Minimize the cost function to obtain the optimal $\mathbf{T}_{b_k}^W$ based on (11).

GNSS simulator utilizes a ray-tracing technique to generate code measurements considering building models and receiver noise. To examine the effects of temporal changes in urban environments, data were generated for different times of day (noon, sunset, and night) within urban areas. Dynamic objects are excluded from this study, as their impact has been extensively analyzed in previous work [47]. The collected data include *10 Hz* LiDAR point clouds (Velodyne HDL-32E with a 100 m range, 5cm STD noise), *20 Hz* images (90° FoV, 960×600 resolution), *100 Hz* IMU measurements (9-axis, using the Xsens MTi noise model from [25]), with an average vehicle speed of 30 km/h. Additionally, GNSS code measurements were generated using the RUMS simulator, which accounts for signal reflection and diffraction effects based on 3D building models, as CARLA only simulates GNSS positioning with random noise. Detailed information on the simulated sensor measurements and trajectories is provided in Fig. 8.

In the Hong Kong C-V2X testbed evaluation, UrbanNav vehicle platform [25] is employed for onboard data collection, which includes GNSS, INS, cameras, and LiDARs. NovAtel SPAN-CPT and Inertial Explorer software [49] are used for GT positioning. The accuracy of our ground truth system is centimeter level continuously, plus 1 ppm, with RTK enabled. Each RSI platform [50] includes GNSS measurements, high-resolution LiDAR, and low-latency V2X communication. We collected *1Hz* U-blox F9P GNSS code and carrier phase measurements, *10 Hz* Velodyne HDL 32-line LiDAR data, *20 Hz* Point Grey camera data, and *400 Hz* Xsens Mti-10 sensor data from our vehicle platform under an average speed of 50 km/h. RSIs are evenly distributed throughout the urban environment, each with an effective range of 500 meters. These RSIs are capable of providing error map services through direct C-V2X (PC5) communication, ensuring latency below 20 milliseconds. Details of the sensor setup can be found in Fig. 9. The collected sequences from the experimental vehicles were divided into two subsets: one for generating the error map and the other for evaluating the effectiveness of the proposed error map-aided method within a closely aligned time period. To minimize the impact of incidental factors such as parking or unexpected stops, we excluded data segments where the vehicle remained stationary for extended durations. The details of the simulated and real-world sequences are listed in Table I.

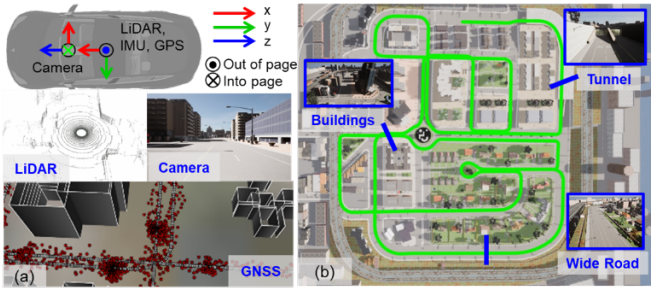

Fig. 8. (a) Sensors used in the simulated platform; (b) Evaluated trajectory and its corresponding ground truth trajectory in the simulated urban environment (modified Town03 map), which includes diverse urban scenarios such as medium-rise buildings, wide roads, and tunnels.

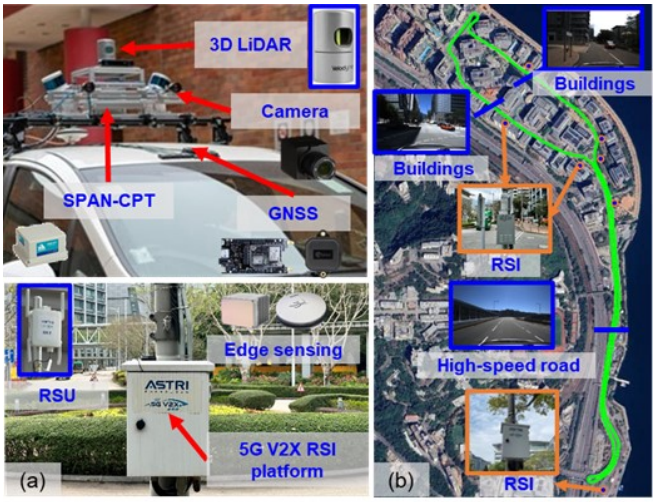

Fig. 9. (a) Sensor used in the vehicle-infrastructure platform; (b) Evaluated trajectory and its corresponding ground truth in the Hong Kong C-V2X testbed. The distributed RSIs are marked in orange.

Evaluated Methods: The error matrix of the following methods is analyzed using the RPE to generate the error maps while the performance of the following methods is evaluated,

(1) **LOAM** [2]: The widely-used LiDAR odometry method.
(2) **LIO-SAM** [42]: The popular LIO method that can be loosely coupled with GPS if available.
(3) **VINS-Mono** [43]: The popular VIO system that uses camera and IMU to estimate the state.

(4) **GNSS Positioning** [46]: GNSS WLS positioning is used for simulated data, while GNSS Standard Point Positioning (SPP) and GNSS-RTK positioning are evaluated for real-world data, as the simulated data contains only pseudorange measurements.
(5) **LVI-SAM** [51]: The tightly coupled method integrates the LIO and VIO into a factor graph. The depth information provided by LiDAR aids in recovering the scale of visual odometry.
(6) **EMA-MS**: The error map-assisted sensor integrated positioning algorithm. The sensor weighting is determined using the continuous error estimates provided based on continuous error estimates provided by periodically updated maps, which are generated by map update vehicles or robobuses.

TABLE I. The evaluation is conducted using simulated urban data and real-world data from the Hong Kong C-V2X testbed.

Data		Description
Simulated Urban	*sim-noon 1*	data collected at noon is used to estimate error maps
	sim-noon 2	data for evaluation of proposed method at noon
	sim-sunset 1	data collected at sunset is used to estimate error maps
	sim-sunset 2	data for evaluation of proposed method at sunset
	sim-night 1	data collected at night is used to estimate error maps
	sim-night 2	data for evaluation of proposed method at night
Hong Kong C-V2X Testbed	*testbed-day 1*	data collected at daytime is used to estimate error maps
	testbed-day 2	data for evaluation of proposed method at daytime
	testbed-night 1	data collected at night is used to estimate error maps
	testbed-night 2	data for evaluation of proposed method at night

B. Experiment in Simulated Urban Areas

The data used for evaluation was collected in a simulated urban scenario covering 3.5 km, as shown in Fig. 8-(a). As detailed in Table I, three data were collected using a single sensor-rich CAV across temporal varying temporal conditions (*sim-noon 1*, *sim-sunset1*, and *sim-night 1*) for error map estimation. Subsequently, three additional data were collected with the same CAV (*sim-noon2*, *sim-sunset2*, and *sim-night2*) during the corresponding periods to assess the effectiveness of the error map-aided method, which leverages the error maps generated from the initial data.

1) Results of Sensor Error Maps

APE is employed as the metric for estimating the sensor error maps. Table II reports the error metric of the listed algorithms. The LOAM performance is inferior to that of the other methods, with an error of 0.520 meters in *sim-noon 1*. The RPE is reduced to 0.318 meters in the same scenario aided by inertial sensors using LIO-SAM. However, the RPE and APE are still inadequate for autonomous driving due to degradation in the tunnel environments (as shown in Fig. 10) and the accumulation of drift. VINS-Mono demonstrates average accuracy regarding RPE and APE in this scene. Nevertheless, the RPE increases from 0.066 in *sim-noon 1* to 0.077 meters in *sim-night 1* due to illumination affecting the VINS feature extraction process. GNSS positioning yields improved APE results as it provides absolute positioning without the accumulation of drift. However, in urban areas, as shown in Fig. 10 and shortage observations in tunnels fail to meet the robustness requirements for CAVs. In summary, each sensor has its limitations, and combining sensors effectively can enhance positioning accuracy.

TABLE II. Error metric of listed methods using heterogeneous sensor measurements.

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

Fig. 10 and Fig. 11 illustrate the positioning accuracy and trajectories for the four methods, respectively. The LiDAR-based methods, LOAM and LIO-SAM, experience degradation in tunnel environments (as illustrated in Fig. 10-(A)), making it challenging to maintain accurate positioning. In contrast, the visual-based method, VINS, delivers consistent performance across various time periods but experiences reduced accuracy in low-light conditions, as depicted in Fig. 10-(C). GNSS positioning delivers average absolute performance among these methods. However, it experiences significant degenerated performance near tall buildings, as illustrated in Fig. 10-(B).

Fig. 12 illustrates the error map results for the four methods based on *sim-noon 1*. The error maps for *sim-sunset 1* and *sim-night 1* are available in the supplementary materials[4]. Similar to Figs. 10 and 11, , LOAM and LIO-SAM exhibit significant errors in tunnel environments. Interestingly, error maps reveal that pose estimations from visual and LiDAR-based methods complement each other in typical tunnel environments. Additionally, the GNSS error map can be used to identify areas with more accurate GNSS

[4]https://github.com/DarrenWong/continuous_error_map/blob/main/doc/error_map_supplementary.pdf

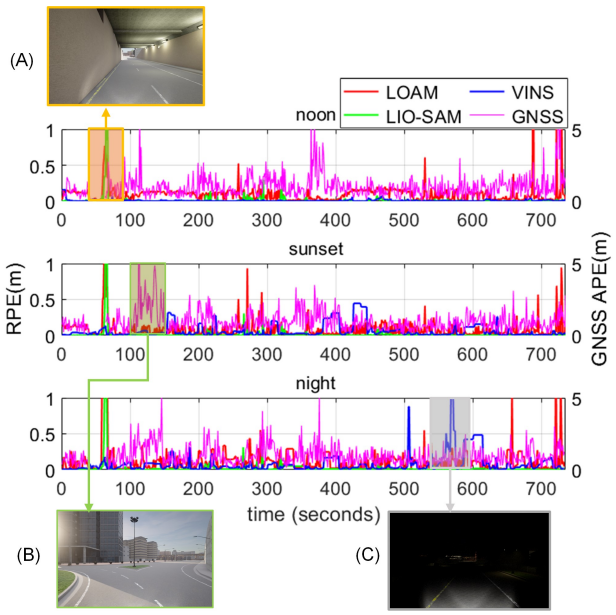

Fig. 10. Results of the listed methods (GNSS APE is in 2D). The horizontal axis represents epochs, while the vertical axis represents the corresponding RPE (left axis, LOAM marked in red, LIO-SAM marked in green, VINS marked in blue and GNSS APE (right axis, marked in pink). Scenes (A), (B), and (C) highlight the challenging scenes by different sensors in state estimation.

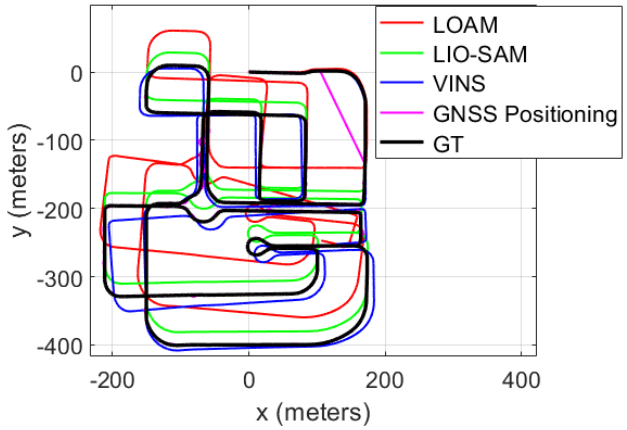

Fig. 11. Trajectories of the listed methods using *sim-noon 1*. The trajectories for *sim-sunset 1* and *sim-night 1* are provided in the supplementary[4].

solutions, which can help improve sensor integration by selecting good GNSS solutions with less NLOS and multipath effects. Consequently, CAVs mounted with LiDAR, cameras, or GNSS sensors can take advantage of real-time error map estimations transmitted by the RSI to improve the integrated positioning performance. Additionally, CAVs can utilize state estimation errors to improve autonomous planning and control in regions with higher error rates.

2) Comparison Results

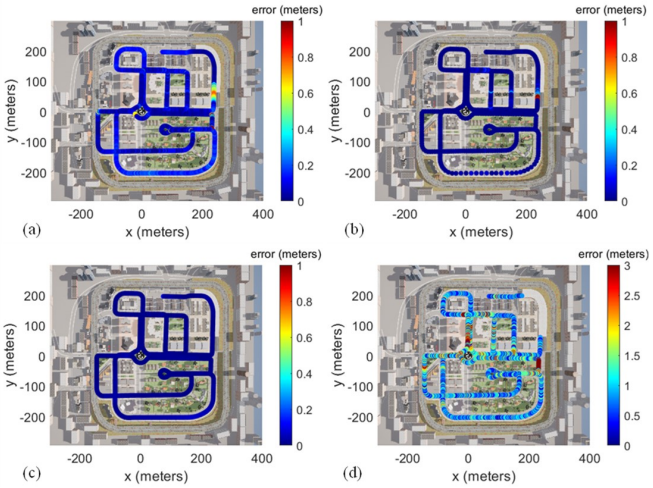

Fig. 12. Illustration of error map for the listed methods using data from *sim-noon 1*: (a) LOAM; (b) LIO-SAM; (c) VINS;(d) GNSS Positioning. The error maps of *sim-sunset 1* and *sim-night 1* can be found in supplementary[4]. Note that the error color that is larger than the color legend upper limit will be the same red color as the upper limit.

TABLE III. Performance evaluation of listed methods.

Data	Methods	APE RMSE (m)	APE MEAN (m)
sim-noon 2	LIO-SAM	19.749	17.723
	VINS	16.103	15.202
	LVI-SAM	1.414	2.095
	LIO-SAM w/ GNSS	11.077	7.246
	EMA-MS	**1.019**	**0.828**
sim-sunset 2	LIO-SAM	24.172	21.678
	VINS	29.565	28.06
	LVI-SAM	2.549	2.456
	LIO-SAM w/ GNSS	12.040	8.238
	EMA-MS	**1.732**	**1.568**
sim-night 2	LIO-SAM	21.376	19.168
	VINS	12.425	11.382
	LVI-SAM	3.658	3.523
	LIO-SAM w/ GNSS	6.715	4.787
	EMA-MS	**2.909**	**1.852**

Table III reports the APE RMSE and MEAN of the test data. The proposed method outperforms existing techniques, achieving an RMSE of 1.019 meters in *sim-noon 2*. In the same scenario, LIO-SAM recorded an RMSE of 19.749 meters, with a mean error of 17.723 meters. VINS-Mono demonstrated comparable results, yielding an RMSE of 16.103 meters. LVI-SAM performed accurately on the test data because of its tightly integrated framework. However, its error increased under poor illumination conditions, particularly in *sim-night 2*, where the APE RMSE increased from 1.414 meters to 3.658 meters. The proposed method, on the other hand, maintained better performance by adaptively

weighting LIO, VIO, and GNSS factors aided by error maps in *sim-noon 2*.

Fig. 13 illustrates the trajectories produced by the listed methods. It is clear that the trajectory estimated by the proposed EMA-MS method closely matches the ground truth, particularly in the zoomed-in U-turn area. The trajectories for *sim-sunset 2* and *sim-night 2* are available in the supplementary materials[4].

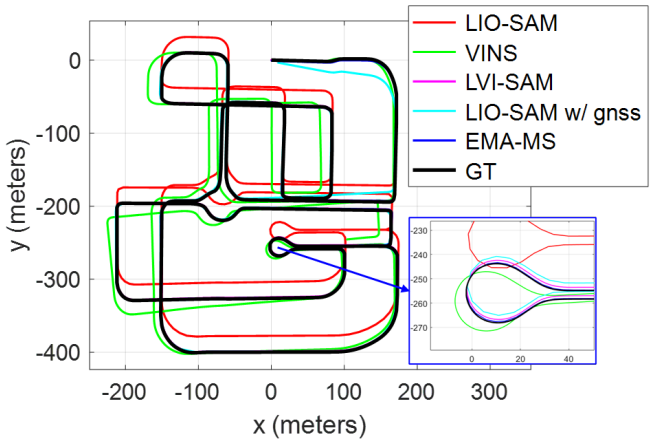

Fig. 13. Trajectories of the listed methods in *sim-noon 2*, with the blue-marked area showing a zoom-in of the U-turn region. The trajectories for *sim-sunset 2* and *sim-night 2* are provided in the supplementary materials[4].

C. Experiment in Hong Kong C-V2X Testbed

To further evaluate the proposed method in complex real-world scenarios, another experimental validation is conducted in the Hong Kong C-V2X testbed with a total length of 4.19 km, as shown in Fig. 9-(b). It involved various challenging scenarios, including dense urban buildings, high-speed roads, and areas prone to sensor degradation. As detailed in Table I, two datasets were collected using a sensor-rich CAV in days and night conditions (*testbed-day 1* and *testbed-night 1*) for error map estimation. Subsequently, two additional datasets (*testbed-day 2* and *testbed-night 2*) were collected during the same periods to assess the performance of the error map-aided method, with the error map estimated from the previous data assisting in the sensor integration process.

1) Results of Sensor Error Maps

Table IV summarizes the performance of the listed methods. The LiDAR-only approach shows lower accuracy, with an RPE RMSE of up to 0.452 meters in *HK C-V2X testbed day 1*. With the aid of the inertial sensors, LIO-SAM reduces the RPE to 0.233 meters. VINS achieves an average performance of 0.438 meters in terms of RPE. The APE of GNSS-SPP and GNSS-RTK can reach up to 20.986 meters and 9.5 meters, respectively, due to severe multi-path effects in urban environments. Since GNSS-RTK can provide higher accuracy by the DD technique and is widely available in urban areas [19], [34], [52], we adopt GNSS-RTK positioning results as the GNSS solutions in our real-world experiments.

Fig. 14 and Fig. 15 present the positioning results and the trajectories using the listed methods, respectively. The ground truth positioning is denoted in the black curve. The LiDAR-based methods are degenerated in the featureless scene (shown in Fig. 14-(A) and (E)). VINS suffers a degraded performance if under ill illumination and featureless scenes, as shown in Fig. 14 (D) and (E). The GNSS positioning method provided better absolute performance compared to the LIO-SAM. Similar to the simulated scene, it suffers from large errors entering the dense urban areas, as shown in Fig. 14 (C) and 14 (F).

Fig. 16 illustrates the error map results from the listed methods using *testbed-day 1*. Similar to Figs. 14 and 15, significant errors are observed in VINS, particularly in high-speed scenarios. GNSS positioning performance degraded in areas with tall buildings and beneath bridges.

TABLE IV. Error metric of listed methods using heterogeneous sensor measurements.

Data	Methods	RPE RMSE (m)	APE RMSE (m)
testbed-day 1	LOAM	0.452	223.142
	LIO-SAM	0.233	9.546
	VINS	0.438	556.614
	GNSS-SPP	*	20.986
	GNSS-RTK	*	9.597
testbed-night 1	LOAM	0.159	47.848
	LIO-SAM	0.357	16.501
	VINS	0.134	258.772
	GNSS-SPP	*	11.307
	GNSS-RTK	*	9.306

2) Comparison Results

The positioning performance for the Hong Kong C-V2X data is summarized in Table V. The existing LIO-SAM method achieved an APE RMSE of 12.476 meters in the same scenario, with a mean error of 10.875 meters. The performance of VINS deteriorated as illumination conditions worsened, particularly in *testbed-night 2*. Similarly, LVI-SAM exhibited comparable performance to VINS, primarily due to the challenges that VINS faced in these dynamic illumination scenes. With the aid of the weighted constraint, the error of the proposed method is reduced to 4.743 meters in *HK C-V2X day 2* and 4.120 meters in *HK C-V2X night 2*, respectively. Fig. 17 demonstrated the trajectories generated by the five methods. It is evident that the trajectory estimated by the proposed method closely matches the ground truth, particularly in the zoomed-in U-turn zone. The results for *HK C-V2X night 2* are available in the supplementary materials[4]. In short, incorporating the sensor error maps greatly enhances the multi-sensor integrated positioning in urban environments.

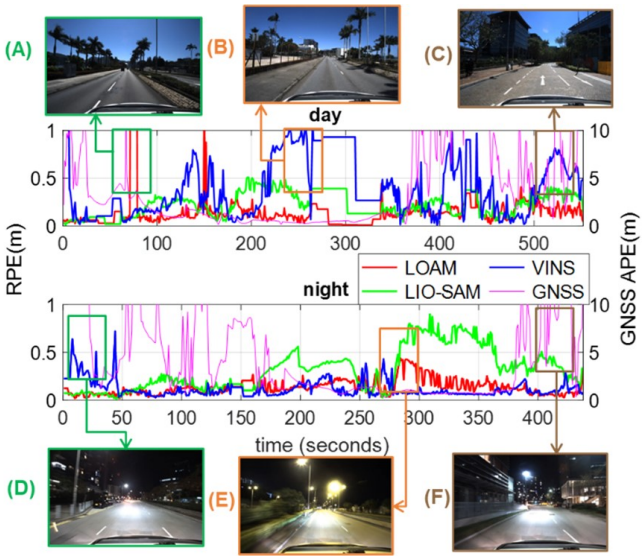

Fig. 14. Results of the listed methods (GNSS APE is in 2D). The horizontal axis represents epochs, while the vertical axis represents the corresponding RPE (left axis, LOAM marked in red, LIO-SAM marked in green, VINS marked in blue and GNSS APE (right axis, marked in pink), respectively. Scene (A)-(F) highlights the challenging scenes by different sensors in state estimation

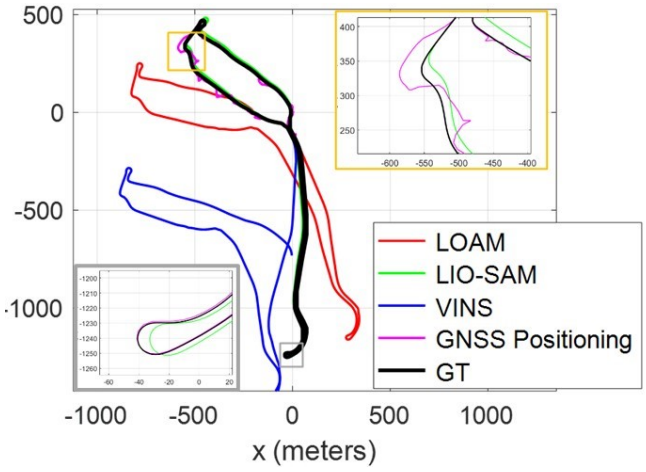

Fig. 15. Trajectories of the listed methods using *testbed-day 1*. The trajectories of *testbed-night 1* can be found in the supplementary[4]

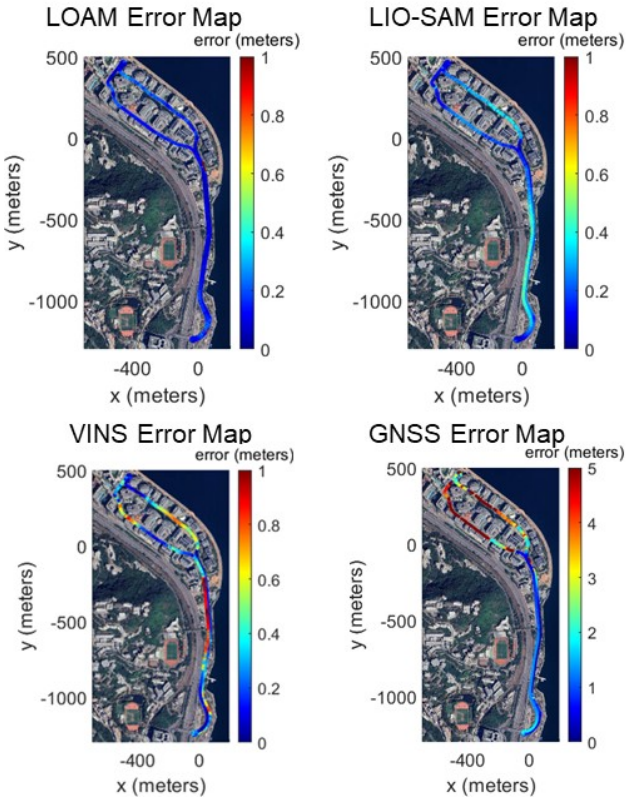

Fig. 16. Illustration of the error maps for the listed methods using data from *HK C-V2X day 1*. The error maps of *HK C-V2X night 1* can be found in the supplementary[4]. Note that the error color that is larger than the color legend upper limit will be the same red color as the upper limit.

TABLE V. 2D Positioning evaluation of listed methods.

Data	Methods	APE RMSE (m)	APE MEAN (m)
testbed-day 2	LIO-SAM	12.476	10.875
	VINS	424.171	378.136
	LVI-SAM	633.347	575.513
	LIO-SAM w/ GNSS	5.536	4.289
	EMA-MS	**4.743**	**3.911**
testbed-night 2	LIO-SAM	8.871	8.275
	VINS	600.720	422.203
	LVI-SAM	410.649	310.221
	LIO-SAM w/ GNSS	6.091	4.471
	EMA-MS	**4.120**	**3.498**

D. Results of Computational Efficiency

Regarding computational efficiency, the error map-aided adaptive sensor integration introduces two additional components compared to the original LIO or VIO methods. First, the extraction of error information for the error maps per epoch takes less than 1 millisecond using the nearest neighbor search. Second, the backend optimization process for sensor integration takes approximately 0.17 seconds, satisfying real-time requirements. Overall, including error maps does not significantly impact computational efficiency, making them suitable for most CAV applications, as demonstrated by our proposed method in this section.

E. Results of Transmission Latency

The error map message is packed in the standard V2X basic safety message (BSM) of 254 bytes per message [48] and broadcast to the OBU which is located in the vehicle platform. The average latency of one message or in one

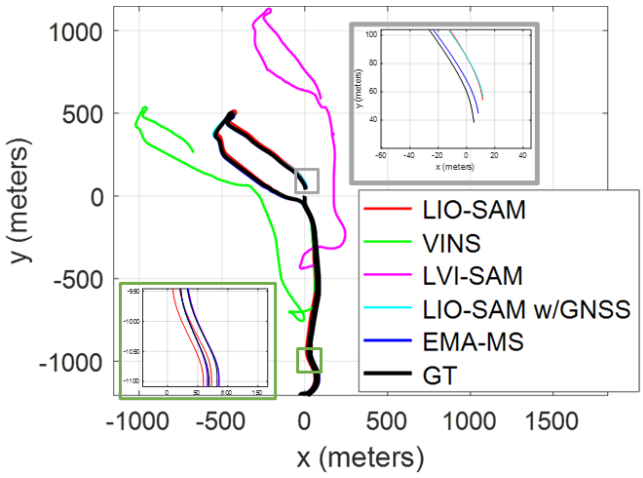

Fig. 17. Trajectories of the listed methods in Data *HK C-V2X day 2*.

message group sent from RSU and then received by OBU is 21.3 milliseconds.

IV. Conclusion

This study presents the integration of heterogeneous sensor error maps as an additional semantic layer atop high-definition maps, tailored for V2X-enabled urban environments. By leveraging real-time error information shared via roadside infrastructure, the proposed system dynamically adjusts sensor weighting in multi-sensor fusion, resulting in improved localization performance. Extensive validation in both high-fidelity simulation and a real-world C-V2X testbed in Hong Kong demonstrates the effectiveness of the approach in enhancing positioning accuracy under complex urban conditions.

Our future work will focus on further improving the robustness and scalability of the error maps through a deeper investigation of sensor behavior across diverse vehicles and deployment scenarios. This includes:

1) Assessing cross-vehicle sensor correlations by investigating how spatial and temporal alignment of sensors across different vehicles affects error propagation.
2) Evaluating the impact of sensor accuracy and configuration by analyzing how variations in LiDAR, GNSS, and IMU specifications influence the accuracy of shared error maps.
3) Examining operational duration effects by exploring the impact of long-term operation and accumulated drift on error consistency across multiple vehicles.

Acknowledgments

The authors thank Yihan Zhong from PolyU IPNL lab, Chen Hang, and Alpamys Urtay from ASTRI for their kind support in this research data collection.

References

[1] B. Song, X. Yuan, Z. Ying, B. Yang, Y. Song, and F. Zhou, "Dgm-vins: Visual–inertial slam for complex dynamic environments with joint geometry feature extraction and multiple object tracking," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–11, 2023.

[2] J. Zhang and S. Singh, "Low-drift and real-time lidar odometry and mapping," *Autonomous robots*, vol. 41, 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.

[4] 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.

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

[6] L. Zhang, H.-F. Ng, G. Zhang, and L.-T. Hsu, "Ray-tracing correction for gnss velocity estimation using doppler frequency: A feasibility analysis," *IEEE Transactions on Instrumentation and Measurement*, 2024.

[7] H.-F. Ng and L.-T. Hsu, "3d mapping database-aided gnss rtk and its assessments in urban canyons," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 57, no. 5, pp. 3150–3166, 2021.

[8] Y. Zhong, R. Hu, X. Bai, X. Li, L.-T. Hsu, and W. Wen, "Enhancing gnss positioning accuracy for road monitoring systems: A factor graph optimization approach aided by geospatial information," *IEEE Transactions on Instrumentation and Measurement*, vol. 73, pp. 1–12, 2024.

[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.

[10] C. Xue, Y. Huang, C. Zhao, X. Li, L. Mihaylova, Y. Li, and J. A. Chambers, "A gaussian-generalized-inverse-gaussian joint-distribution-based adaptive msckf for visual-inertial odometry navigation," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 59, no. 3, pp. 2307–2328, 2022.

[11] G. Welch, "An introduction to the kalman filter," 1995.

[12] Z. Chen *et al.*, "Bayesian filtering: From kalman filters to particle filters, and beyond," *Statistics*, vol. 182, no. 1, pp. 1–69, 2003.

[13] G. P. C. Júnior, A. M. C. Rezende, V. R. F. Miranda, R. Fernandes, H. Azpúrua, A. A. Neto, G. Pessin, and G. M. Freitas, "Ekf-loam: An adaptive fusion of lidar slam with wheel odometry and inertial data for confined spaces with few geometric features," *IEEE Transactions on Automation Science and Engineering*, vol. 19, no. 3, pp. 1458–1471, 2022.

[14] 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*. IEEE, 2027, Conference Proceedings, pp. 3565–3572.

[15] 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.

[16] 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.

[17] H. Xiao, Z. Hu, C. Lv, J. Meng, J. Zhang, and J. You, "Progressive multi-modal semantic segmentation guided slam using tightly-coupled lidar-visual-inertial odometry," *IEEE Transactions on Intelligent Transportation Systems*, vol. 26, no. 2, pp. 1645–1656, 2025.

[18] 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.

[19] H. Caesar, V. Bankiti, A. H. Lang, S. Vora, V. E. Liong, Q. Xu, A. Krishnan, Y. Pan, G. Baldan, and O. Beijbom, "nuscenes: A multimodal dataset for autonomous driving," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 11 621–11 631.

[20] C. Chen, S. Rosa, Y. Miao, C. X. Lu, W. Wu, A. Markham, and N. Trigoni, "Selective sensor fusion for neural visual-inertial odometry," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 10 542–10 551.

[21] W. Su, T. Huang, F. Liu, and H. Wang, "Aslam-fd: a multi-sensor adaptive collaborative fusion slam framework based on degradation detection and deep reinforcement learning," *Measurement Science and Technology*, vol. 35, no. 12, p. 126312, 2024.
[22] 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*, vol. 21, no. 2, pp. 277–292, 2023.
[23] M. Elhousni and X. Huang, "A survey on 3d lidar localization for autonomous vehicles," in *2020 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, 2020, pp. 1879–1884.
[24] X. Xu, L. Zhang, J. Yang, C. Cao, W. Wang, Y. Ran, Z. Tan, and M. Luo, "A review of multi-sensor fusion slam systems based on 3d lidar," *Remote Sensing*, vol. 14, no. 12, p. 2835, 2022.
[25] L.-T. Hsu, F. Huang, H.-F. Ng, G. Zhang, Y. Zhong, X. Bai, and W. Wen, "Hong Kong UrbanNav: An open-source multisensory dataset for benchmarking urban navigation algorithms," *NAVIGATION: Journal of the Institute of Navigation*, vol. 70, no. 4, 2023.
[26] 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)*. IEEE, 2018, pp. 2502–2509.
[27] P. Chen, W. Guan, F. Huang, Y. Zhong, W. Wen, L.-T. Hsu, and P. Lu, "Ecmd: An event-centric multisensory driving dataset for slam," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 1, pp. 407–416, 2024.
[28] 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.
[29] R. Sun, L. Fu, Q. Cheng, K.-W. Chiang, and W. Chen, "Resilient pseudorange error prediction and correction for gnss positioning in urban areas," *IEEE Internet of Things Journal*, vol. 10, no. 11, pp. 9979–9988, 2023.
[30] Deloitte and ASTRI, "Hong kong connected & autonomous vehicle (cav) development study," https://www2.deloitte.com/cn/zh/pages/technology/articles/hk-cav-development-study.html, Deloitte, 2024, accessed: April 18, 2024.
[31] H. Wang, X. Yuan, Y. Cai, L. Chen, and Y. Li, "V2i-carla: A novel dataset and a method for vehicle reidentification-based v2i environment," *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1–9, 2022.
[32] R. Xu, C.-J. Chen, Z. Tu, and M.-H. Yang, "V2x-vitv2: Improved vision transformers for vehicle-to-everything cooperative perception," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2024.
[33] H. Yu, Y. Luo, M. Shu, Y. Huo, Z. Yang, Y. Shi, Z. Guo, H. Li, X. Hu, J. Yuan, *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. 21 361–21 370.
[34] J. Morais, G. Charan, N. Srinivas, and A. Alkhateeb, "Deepsense-v2v: A vehicle-to-vehicle multi-modal sensing, localization, and communications dataset," *arXiv preprint arXiv:2406.17908*, 2024.
[35] A. Alkhateeb, G. Charan, T. Osman, A. Hredzak, J. Morais, U. Demirhan, and N. Srinivas, "Deepsense 6g: A large-scale real-world multi-modal sensing and communication dataset," *IEEE Communications Magazine*, vol. 61, no. 9, pp. 122–128, 2023.
[36] F. Huang, H. Chen, A. Urtay, D. Su, W. Wen, and L.-T. Hsu, "Roadside infrastructure assisted lidar/inertial-based mapping for intelligent vehicles in urban areas," in *2023 IEEE 26th International Conference on Intelligent Transportation Systems (ITSC)*, 2023, pp. 5831–5837.
[37] 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.
[38] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "Carla: An open urban driving simulator," in *Conference on robot learning*. PMLR, 2017, pp. 1–16.
[39] 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.
[40] 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.
[41] 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, pp. 1572–1578.
[42] T. Shan, B. Englot, D. Meyers, W. Wang, C. Ratti, and R. Daniela, "Lio-sam: Tightly-coupled lidar inertial odometry via smoothing and mapping," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2020, pp. 5135–5142.
[43] 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.
[44] P. D. Groves, *Principles of GNSS, Inertial, and Multisensor Integrated Navigation Systems*. Artech House, jan 2008.
[45] P. Teunnissen, "The least-square ambiguity decorrelation adjustment: a method for fast gps integer ambiguity estimation," *J. Geodesy*, vol. 70, no. 1, pp. 65–82, 1995.
[46] T. Takasu and A. Yasuda, "Development of the low-cost rtk-gps receiver with an open source program package rtklib," in *International symposium on GPS/GNSS*, vol. 1. International Convention Center Jeju Korea Seogwipo-si, Republic of Korea, 2009, pp. 1–6.
[47] F. Huang, W. Wen, J. Zhang, C. Wang, and L.-T. Hsu, "Dynamic object-aware LiDAR odometry aided by joint weightings estimation in urban areas," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 2, pp. 3345–3359, 2024.
[48] M.-S. Kang, J.-H. Ahn, J.-U. Im, and J.-H. Won, "Lidar-and v2x-based cooperative localization technique for autonomous driving in a gnss-denied environment," *Remote Sensing*, vol. 14, no. 22, p. 5881, 2022.
[49] N. Inc, "Waypoint inertial explorer 8.80 post processing software," 2020.
[50] H. K. A. Science and T. R. I. C. Limited, "C-v2x technology," https://www.astri.org/tdprojects/connected-vehicle-v2x-technology/, 2023, accessed: September 15, 2024.
[51] 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)*. IEEE, 2021, pp. 5692–5698.
[52] X. Wang, X. Li, H. Chang, S. Li, Z. Shen, and Y. Zhou, "Give: A tightly coupled rtk-inertial–visual state estimator for robust and precise positioning," *IEEE Transactions on Instrumentation and Measurement*, vol. 72, pp. 1–15, 2023.

Huang Feng received his bachelor's degree from Shenzhen University in Automation in 2014 and MSc in Electronic Engineering at Hong Kong University of Science and Technology in 2016. He is a Ph.D. student in the Department of Aeronautical and Aviation Engineering, at Hong Kong Polytechnic University. His research interests include localization and sensor fusion for autonomous driving.

Wen Weisong (Member, IEEE) received a BEng degree in Mechanical Engineering from Beijing Information Science and Technology University (BISTU), Beijing, China, in 2015, and an MEng degree in Mechanical Engineering from the China Agricultural University, in 2017. After that, he received a PhD degree in Mechanical Engineering from The Hong Kong Polytechnic University (PolyU), in 2020. He was also a visiting PhD student with the Faculty of Engineering, University of California, Berkeley (UC Berkeley) in 2018. Before joining PolyU as an Assistant Professor in 2023, he was a Research Assistant Professor at AAE of PolyU since 2021. He has published 30 SCI papers and 40 conference papers in the field of GNSS (ION GNSS+) and navigation for Robotic systems (IEEE ICRA, IEEE ITSC), such as autonomous driving vehicles. He won the innovation award from TechConnect 2021, the Best Presentation Award from the Institute of Navigation (ION) in 2020, and the First Prize in Hong Kong Section in Qianhai-Guangdong-Macao Youth Innovation and Entrepreneurship Competition in 2019 based on his research achievements in 3D LiDAR aided GNSS positioning for robotics navigation in urban canyons. The developed 3D LiDAR-aided GNSS positioning method has been reported by top magazines such as Inside GNSS and has attracted industry recognition with remarkable knowledge transfer.

Zhang Guohao (Member, IEEE) received his bachelor's degree in mechanical engineering and automation from University of Science and Technology Beijing, China. He received his master's degree in Mechanical Engineering and his Ph.D. degree in Aeronautical and Aviation Engineering, the Hong Kong Polytechnic University. He is currently a research assistant professor at Department of Aeronautical and Aviation Engineering of Hong Kong Polytechnic University. He won the Best Presentation Award in 2018 and the ION GNSS+ Student Paper Award in 2019 from the Institute of Navigation (ION). His research interests include GNSS urban positioning, collaborative positioning, machine learning aided GNSS, and multi-sensor integrated navigation.

Su Dongzhe received a bachelor's degree from the Huazhong University of Science and Technology, and the M.Phil. degree in computer science from The Hong Kong University of Science and Technology. He is currently a Chief Engineer with the Communication Technologies Group, Hong Kong Applied Science and Technology Research Institute. He has been leading the system architecture in research and development of vehicle-to-everything (V2X) communication and application systems, connected autonomous vehicles (CAV) systems, and the Internet of Things (IoT). His role has been to define the technical scope and overall system design for ASTRI's V2X networking system. In 2021, ASTRI launched one of the world's largest C-V2X public road tests in Hong Kong, covering a 14 km route with various road environments of Hong Kong.

Yulong Huang (M'19-SM'24) received the B.S. degree in Automation from the College of Automation, Harbin Engineering University, Harbin, China, in 2012. He received the Ph.D degree in control science and engineering from the College of Automation, Harbin Engineering University, Harbin, China, in 2018. From Nov. 2016 to Nov. 2017, he was a visiting graduate researcher at the Electrical Engineering Department of Columbia University, New York, USA. From Dec. 2019 to Dec. 2021, he was associated with the Department of Mechanical Engineering, City University of Hong Kong, as a Hong Kong Scholar.

He is a Full Professor of navigation, guidance, and control, and Vice Dean of College of Intelligent Systems Science and Engineering at Harbin Engineering University (HEU) in China. His current research interests include state estimation, intelligent information fusion and their applications in navigation technology, such as inertial navigation, integrated navigation, intelligent navigation, and cooperative navigation. Currently, he serves as an Associate Editor for the IEEE TRANSACTIONS ON AUTOMATIC CONTROL, for the IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS, for the IEEE TRANSACTIONS ON AUTOMATION SCIENCE AND ENGINEERING, for the IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT, and for the IEEE SENSORS JOURNAL, and a Youth Editor for the IEEE/CAA JOURNAL OF AUTOMATICA SINICA (JAS), for the JOURNAL OF MARINE SCIENCE AND APPLICATION, for the UNMANNED SYSTEMS TECHNOLOGY, for the JOURNAL OF UNMANNED UNDERSEA SYSTEMS, for the NAVIGATION POSITIONING AND TIMING.

He was the recipient of the 2018 IEEE Barry Carlton Award from IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS in 2022, the Honorable Mention of 2017 IEEE Barry Carlton Award from IEEE TRANSACTIONS ON AEROSPACE AND ELECTRONIC SYSTEMS in 2021, the Best Student Paper Award of 2023 IEEE International Conference on Mechatronics and Automation (IEEE ICMA 2023), and the Best Paper Award of 2021 International Conference on Autonomous Unmanned Systems (ICAUS 2021). He was also the recipient of the 11th China Youth Science and Technology Innovation Awards in 2018, the excellent doctoral thesis from Chinese Association of Automation (CAA) in 2019, the Wu Wen-Jun AI Excellent Youth Scholar Award in 2021, the First Prize of Natural Science Award of Chinese Association of Automation (Ranked the 2nd) in 2021, and he was selected into the sixth Young Elite Scientists Sponsorship Program by China Association for Science and Technology in 2020. He was an outstanding Associate Editor and Reviewer for the IEEE TRANSACTIONS ON INSTRUMENTATION AND MEASUREMENT.