IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 7 August 2023, accepted 29 August 2023, date of publication 1 September 2023, date of current version 8 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3311359

RESEARCH ARTICLE

Improved Multi-Sensor Fusion Positioning System Based on GNSS/LiDAR/Vision/IMU With Semi-Tight Coupling and Graph Optimization in GNSS Challenging Environments

JIAMING ZHU, HAN ZHOU, ZIYI WANG, AND SULI YANG
School of Earth Sciences and Engineering, Hohai University, Nanjing 211100, China

Corresponding author: Jiaming Zhu (2009030215@hhu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 42004014, in part by the Natural Science Foundation of Jiangsu Province under Grant BK20200530, and in part by the Fundamental Research Funds for the Central Universities under Grant B230201015.

ABSTRACT With the development of autonomous driving, precise positioning capabilities are becoming increasingly important. GNSS (Global Navigation Satellite System) is normally utilized for vehicle positioning, but is susceptible to factors such as urban canyons, especially in increasingly urbanized scenario nowadays. The interpretation of relative positioning information by means of multi-source sensors such as LiDAR (Light Detection And Ranging) or camera, has also been widely investigated, but there are deficiencies in the precision and reliability of the sensors due to their operating principles. For GNSS challenging environments, we give a semi-tightly coupled sensor fusion system based on factor graph optimization. The system is tightly coupled with the raw observations from the LiDAR, camera, and inertial measurement unit. Based on this, the back end is optimized by means of the factor graphs, which ultimately give optimized positional information of the vehicle, and the GNSS sensor is decoded to reduce the cumulative error, which is regarded as the loosely coupled component of graph optimization. We selected three representative datasets from publicly available datasets in different categories and conducted experiments respectively. The results indicate that this semi-tightly coupled system outperforms both the GNSS strategy and the typical SLAM (Simultaneous Localization And Mapping) strategies in terms of precision and possesses very high reliability, exhibiting the best performance.

INDEX TERMS GNSS, SLAM, LiDAR/Vision/IMU, semi-tightly coupling, factor graph optimization.

I. INTRODUCTION

Navigation and positioning systems have become the key to autonomous driving [1]. While autonomous driving depends on continuous and precise localization anywhere anytime during its whole course for simultaneous localization and environmental perception [2]. In outdoor environments, the GNSS (Global Navigation Satellite System) is currently the main tool for vehicle simultaneous position and navigation. However, due to the error sources such as multipath effects and other interference, it is easy to produce insufficient accuracy or positioning errors [3]. Single GNSS sensors make it difficult to guarantee the precision and reliability of the input information in GNSS challenging environments. At the same time, other sensors such as LiDAR (Light Detection And Ranging) and camera are also being used for SLAM (Simultaneous Localization And Mapping). The precision and dependability of location and navigation are further enhanced by the redundant and complementary information about the same environment that is provided by the fusion of multi-source sensors. Multi-sensor fusion is becoming one of the major trends in the field of navigation, positioning, and mapping today [4].

The associate editor coordinating the review of this manuscript and approving it for publication was Riccardo Carotenuto.

Since GNSS is a commonly utilized technology that effectively provides precise positioning regardless of time and location, there are many studies based on RTK (real-time kinematic) positioning. RTK positioning is a relative positioning procedure, where the unknown position of the receiver is determined with respect to a stationary base station of known coordinates [5]. RTK positioning obtains position information by performing differential calculations on carrier phase observations. But along with the increased urbanization, the signals are easily reflected, refracted, diffracted, and even blocked, which will usually cause site-specific observation errors [6], [7]. There are some main reasons for this, one is the presence of substantial multipath error, and the second one is a decrease in the number of satellites that are visible [8]. Furthermore, the limitations of receiver hardware are also a significant factor as well, taking the most common smartphone as an example, the carrier-to-noise ratio (C/N0) of smartphone GNSS signals is lower than that of geodesic receivers by 7.5–10 dB Hz [9]. It is seriously inconsistent with the expectation of high-precision positioning.

Meanwhile, relative positioning techniques based on the processing of information from multi-source sensors are also emerging. Contrary to the GNSS approach, which only occasionally provides positioning coordinates that satisfy precision requirements in the global coordinate system, multi-source sensors solve the issue of continuous positional acquisition of subjects in the local coordinate system by collecting several types of environmental data. One of the most important features and advantages of multi-source sensor positioning is the possibility of SLAM [10], [11]. LiDAR, camera (also called Vision) and IMU (Inertial Measurement Unit) are the most common sensors in the positioning systems. Although the LiDAR could immediately capture depth information, it is vulnerable to degradation and even failure in regions without evident structural features. The camera can quickly capture colours and textures in the environment but is susceptible to drift. The IMU provides high accuracy over a short period, but inevitably there are cumulative errors over a long period [12]. Depending on the use of major sensors, the positioning algorithms of SLAM can be divided into two categories: Vision-based positioning and LiDAR-based positioning [13]. A common vision-based positioning algorithm is the monocular VINS (visual-inertial system). VINS is a tightly coupled method, recovering the scale via a monocular camera and IMU [14]. Once the picture is blurred or the image is lost, the camera positioning will fail; On the other hand, a common LiDAR-based positioning algorithm is LOAM (LiDAR Odometry and Mapping) [15], which utilizes LiDAR information to calculate odometry information for position acquisition by aligning the laser point cloud between frames, but at the same time it is difficult to perform loop closure detection to correct for drift.

As the research progressed, scholars found that multi-source sensor fusion can improve the precision and reliability of SLAM with its advantages of multi-sensor heterogeneity and complementarity [16], [17], [18], [19]. The general combination of sensors can be divided into loosely and tightly coupled forms. For global positioning, as a loosely coupled factor, GNSS sensors are usually used to provide absolute position information, i.e. global position information [20], [21], [22], [23]. VINS-Fusion [24] uses the position information provided by GNSS sensors as a global constraint. LIO-SAM [25] incorporates GNSS sensor information as one of the constraints for optimization. And at the local positioning level, algorithms such as LVI-SAM [26] tightly couple LiDAR with Vision while adding GNSS factors to complement the advantages of each sensor, achieving great precision and reliability, but it depends on the initialization parameters, and failure of the initialization can easily lead to system failure. Another algorithm of the LVIO system, R3LIVE, implements the colouring of the point cloud on the basis of positioning and map building.

The purpose of adding GNSS sensors as described above is mainly to eliminate cumulative errors. Further, the cumulative errors can be reduced by using graph optimization [27]. In recent years, the graph optimization-based fusion approach has received more and more attention, as it can integrate diverse measurements from different types of sensors, such as GNSS, wheel odometer, LiDAR, camera, and IMU [2], through nonlinear optimization.

In this paper, we give an improved multi-sensor fusion positioning system based on GNSS/LiDAR/Vision/IMU with semi-tightly coupling and graph optimization and also compare it experimentally with a conventional sensor (fusion) strategy. First, to compare with the semi-tightly sensor fusion strategy, we chose four different types of sensor strategies: GNSS, GNSS/LiDAR/IMU, GNSS/Vision/IMU, and LiDAR/Vision/IMU (The strategy proposed in this paper for GNSS-free modules) according to the variations in sensor fusion approaches. Then, we tried each of these five strategies using three datasets of various sorts and roles from various sources. Finally, we determine the precision and reliability properties of the system proposed in this paper and assess the remaining conventional strategies by analyzing and contrasting the experimental results.

II. METHODOLOGY

In this section, we provide a detailed explanation of the semi-tightly sensor fusion system and integrate the other four strategies.

A. POSITIONING AND NAVIGATION WITH GNSS SENSOR

The GNSS sensor, also called GNSS receiver, can simultaneously receive code pseudorange, carrier phase, doppler shift, signal-to-noise ratio (SNR) observations, as well as various error corrections. Therefore, the raw code pseudo-range and carrier phase observation equation between satellite s and receiver r can be expressed as [28], [29]:

$$\begin{aligned} P^{s}_{r,i} = \rho^{s}_{r} + c\delta t_{r} - c\delta t^{s} + \xi_{r,i} - \xi^{s,i} + T^{s}_{r} \\ + I^{s}_{r,i} + M^{s}_{r,i} + \varepsilon^{s}_{r,i} \end{aligned} \quad (1)$$

$$\begin{aligned}\phi^s_{r,i} &= \rho^s_r + c\delta t_r - c\delta t^s + \zeta_{r,i} - \zeta^{s,i} + T^s_r - I^s_{r,i} \\ &\quad + \lambda N^s_{r,i} + m^s_{r,i} + e^s_{r,i}\end{aligned} \tag{2}$$

where $P^s_{r,i}$ and $\phi^s_{r,i}$ denote the code pseudorange and carrier phase observation value between satellite s and receiver r in unit of length; i denotes the frequency; ρ^s_r denotes the geometric distance between satellite s and receiver r at the moment of signal transmission; c denotes the light speed in a vacuum; λ denotes the wavelength; δt_r and δt^s denote the receiver and satellite clock offsets; N denotes the ambiguity of carrier wavelength; ξ and ζ denote the code and phase hardware delays; I and T denote the ionospheric and tropospheric delays; M and m denote the code and phase multipath; ε and e denote the noise of code and phase measurements. The other systematic errors are assumed to be corrected previously, including the phase center offset and variation, phase windup, and earth rotation, etc.

The next phase entails using the observations along with the broadcast ephemeris and the base station file after the GNSS sensor receiver gets the data. Through the combination of these data sources, we can derive precision GNSS absolute positioning information.

For the short and medium distance (e.g., 0∼30km) baseline, inaccuracies in the satellite and receiver clocks are eliminated, and the ionospheric and tropospheric delays can be ignored. Then the between-receiver and between-satellite double differenced observation equation can be expressed as [30], [31]:

$$\Delta\nabla P = \Delta\nabla\rho + \Delta\nabla e \tag{3}$$

$$\Delta\nabla\varphi = \Delta\nabla\rho + \lambda\Delta\nabla N + \Delta\nabla\varepsilon \tag{4}$$

where $\Delta\nabla$ denotes the double-difference operator. $\Delta\nabla P$ denotes the code pseudo-range; $\Delta\nabla\varphi$ denotes the carrier phase double-difference observation values; $\Delta\nabla\rho$ denotes the double-difference geometric distance between the satellite and the GNSS receiver; $\Delta\nabla N$ denotes the double-difference integer ambiguity; $\Delta\nabla e$ denotes the double-difference code pseudo-range; $\Delta\nabla\varepsilon$ denotes the carrier phase measurement noise. It's necessary to note that the resulting wavelength ambiguity is an integer.

Finally, we can obtain the raw GNSS coordinates in the ECEF (Earth-Centered, Earth-Fixed) geocentric coordinate system. After the conversion, the GNSS measurement $\boldsymbol{p}^{\text{GNSS}}_t$ in the ENU world coordinate system (or other coordinate system) expressed as follows:

$$\boldsymbol{p}^{\text{GNSS}}_t = [x^{\text{w}}_t, y^{\text{w}}_t, z^{\text{w}}_t]^{\text{T}} \tag{5}$$

where x^{w}_t, y^{w}_t, and z^{w}_t denote the coordinate components in three directions in the ENU world, respectively.

B. MULTI-SENSOR FUSION POSITIONING SYSTEM BASED ON LiDAR/VISION/IMU TIGHTLY COUPLING

We gather information using LiDAR, Vision, and IMU, and we enhance the integration to finish the positioning and mapping function with tightly-coupled method. This process can also be considered as a tightly coupled form of LiDAR/IMU system and Vision/IMU system.

First, we write the robot state as:

$$\boldsymbol{X} = [\boldsymbol{R}^{\text{T}}, \boldsymbol{P}^{\text{T}}, \boldsymbol{V}^{\text{T}}, \boldsymbol{B}^{\text{T}}]^{\text{T}} \tag{6}$$

where $\boldsymbol{R} \in \text{SO(3)}$ denotes the rotation matrix, SO(3) denotes the three-dimensional rotating groups. $\boldsymbol{P}$ denotes the robot position, $\boldsymbol{V}$ denotes the robot speed, and $\boldsymbol{B}$ denotes the IMU bias.

The system uses IMU to extrapolate the optimization in real time and estimate the correction value of the optimization result. We define that $\boldsymbol{Q}^{\text{IMU}}_{\text{t}}$ as the IMU observation at time t. The velocity, position, rotation and IMU observation of the robot at time t + Δt can be calculated as follows:

$$\boldsymbol{v}_{t+\Delta t} = \boldsymbol{v}_t + \boldsymbol{g}\Delta t + \boldsymbol{R}_t(a_t - b^a_t - n^a_t)\Delta t \tag{7}$$

$$\boldsymbol{p}_{t+\Delta t} = \boldsymbol{p}_t + \boldsymbol{v}_t\Delta t + \frac{1}{2}\boldsymbol{g}\Delta t^2 + \frac{1}{2}\boldsymbol{R}_t(a_t - b^a_t - n^a_t)\Delta t^2 \tag{8}$$

$$\boldsymbol{R}_{t+\Delta t} = \boldsymbol{R}_t\exp((\omega_t - b^\omega_t - n^\omega_t)\Delta t) \tag{9}$$

$$\boldsymbol{Q}^{\text{IMU}}_{t+\Delta t} = [\boldsymbol{v}_{t+\Delta t}, \boldsymbol{p}_{t+\Delta t}, \boldsymbol{R}_{t+\Delta t}] \tag{10}$$

where $\boldsymbol{\omega}_t$ and $\boldsymbol{a}_t$ denote the raw IMU accelerometer and gyroscope measurements in B at time t. $\boldsymbol{v}_t$ and $\boldsymbol{v}_{t+\Delta t}$ denote the velocity at time t and $t + \Delta t$, respectively; $\boldsymbol{p}_t$ and $\boldsymbol{p}_{t+\Delta t}$ denote the position at time t and $t + \Delta t$, respectively; $\boldsymbol{R}_t$ and $\boldsymbol{R}_{t+\Delta t}$ denote the rotation at time t and $t + \Delta t$, respectively.

Similarly, we obtain information from the LiDAR and camera sensors for a certain rate, and the observation equation is shown as follows:

$$\boldsymbol{Q}^{\text{LiDAR}}_t = [\theta^{\text{LiDAR}}_t, \boldsymbol{E}^{\text{LiDAR}}_t] \tag{11}$$

$$\boldsymbol{Q}^{\text{Vision}}_t = [\theta^{\text{Vision}}_t, \boldsymbol{E}^{\text{Vision}}_t] \tag{12}$$

where $\boldsymbol{Q}^{\text{LiDAR}}_t$ and $\boldsymbol{Q}^{\text{Vision}}_t$ denote the LiDAR and camera observations, respectively; θ^{LiDAR}_t and θ^{Vision}_t denote the laser point cloud data and image data acquired at time t, respectively; $\boldsymbol{E}^{\text{LiDAR}}_t$ and $\boldsymbol{E}^{\text{Vision}}_t$ denote the respective parameters of the sensors at time t, respectively.

The tightly coupled system combines the physical models and measurement data of each sensor, and jointly optimizes the measurement data from all sensors. Therefore, we obtain the pose estimation $\boldsymbol{X}$. Detailed fusion principles and processes will be shown in subsection C and D.

C. MULTI-SENSOR FUSION POSITIONING SYSTEM BASED ON GNSS/LiDAR/VISION/IMU SEMI-TIGHTLY COUPLING

The architecture of the proposed semi-tightly coupled positioning strategy is described, which combines GNSS, LiDAR, Vision, and IMU, and is depicted in Fig. 1. The system prototypes utilized GNSS system, LIS system (LiDAR/IMU) and VINS system (Vision/IMU). Specifically, the proposed approach integrates three components: GNSS position measurement, LiDAR odometry measurement, and vision odometry measurement.

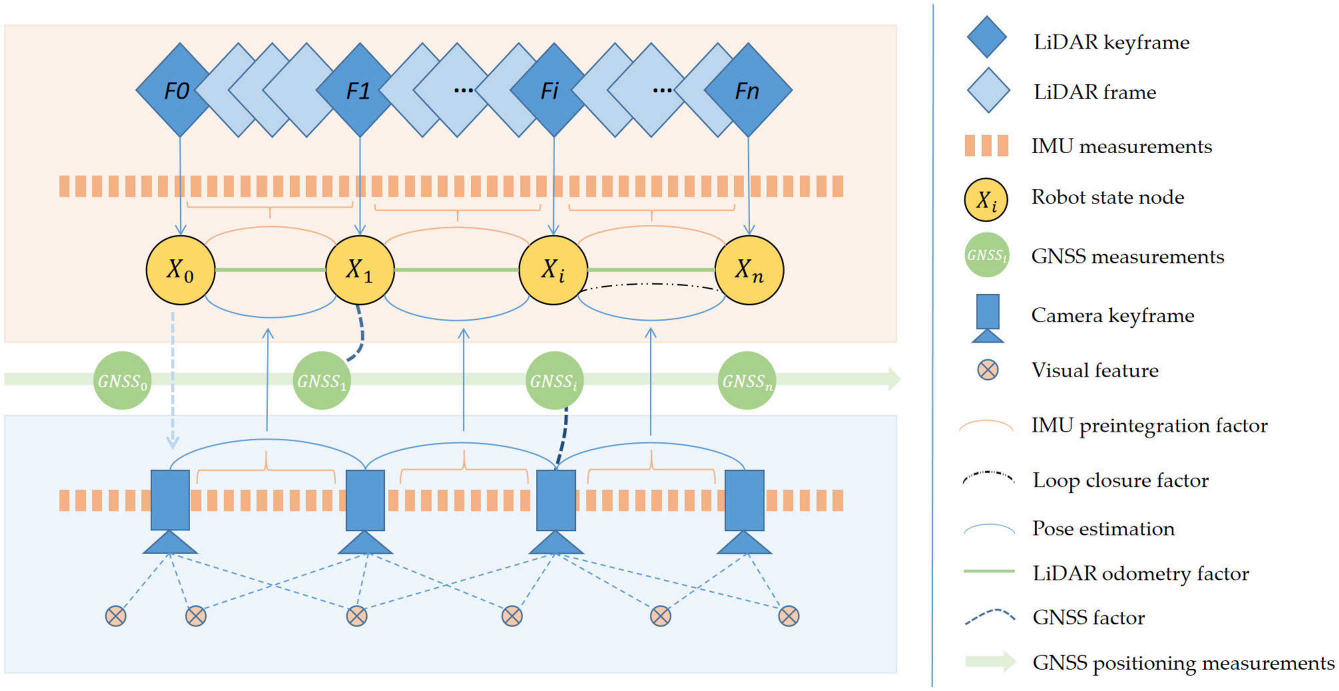


FIGURE 1. **Schematic diagram of the proposed method. The orange, blue, and green areas represent the schematic of the LiDAR, Vision and GNSS components, respectively.**

Based on the position values $\boldsymbol{p}_t^{\mathrm{GNSS}} = [x_t^{\mathrm{w}}, y_t^{\mathrm{w}}, z_t^{\mathrm{w}}]^{\mathrm{T}}$ from GNSS sensor solution and the observations of LiDAR ($\boldsymbol{Q}_t^{\mathrm{LiDAR}}$) and camera ($\boldsymbol{Q}_t^{\mathrm{Vision}}$), we can now obtain the robot pose through the GNSS/LiDAR/Vision/IMU semi-tightly coupled system. We can describe it as:

$$\boldsymbol{X}_t = \boldsymbol{X}_{t,1} + (\boldsymbol{X}_{t,2} + \boldsymbol{X}_{t,3}) \tag{13}$$

$$\boldsymbol{X}_{t,1} = \boldsymbol{A}(\boldsymbol{A}(\boldsymbol{Q}_t^{\mathrm{LiDAR}}, \boldsymbol{Q}_t^{\mathrm{IMU}}), \boldsymbol{A}(\boldsymbol{Q}_t^{\mathrm{Vision}}, \boldsymbol{Q}_t^{\mathrm{IMU}})) \tag{14}$$

$$\boldsymbol{X}_{t,2} = \boldsymbol{B}(\boldsymbol{p}_t^{\mathrm{GNSS}}, \boldsymbol{A}(\boldsymbol{Q}_t^{\mathrm{LiDAR}}, \boldsymbol{Q}_t^{\mathrm{IMU}})) \tag{15}$$

$$\boldsymbol{X}_{t,3} = \boldsymbol{B}(\boldsymbol{p}_t^{\mathrm{GNSS}}, \boldsymbol{A}(\boldsymbol{Q}_t^{\mathrm{Vision}}, \boldsymbol{Q}_t^{\mathrm{IMU}})) \tag{16}$$

where $\boldsymbol{A}(\boldsymbol{Q}_t^a, \boldsymbol{Q}_t^b)$ denotes the tight coupling of the sensors a and b with their original observations; $\boldsymbol{B}(\boldsymbol{Q}_t^a, \boldsymbol{Q}_t^b)$ denotes the loose coupling of the sensors a and b with their position values.

In Fig. 1, the GNSS strategy is represented by the green arrow. Through the processing in subsection A, the GNSS sensor outputs position information $\boldsymbol{p}_t^{\mathrm{GNSS}}$ at a certain frequency, which can be described as the GNSS measurements. Based on (13), the GNSS/LiDAR/IMU sensor fusion strategy can be abstractly expressed as:

$$\boldsymbol{X}_t = \boldsymbol{A}(\boldsymbol{Q}_t^{\mathrm{LiDAR}}, \boldsymbol{Q}_t^{\mathrm{IMU}}) + \boldsymbol{B}(\boldsymbol{p}_t^{\mathrm{GNSS}}, \boldsymbol{A}(\boldsymbol{Q}_t^{\mathrm{LiDAR}}, \boldsymbol{Q}_t^{\mathrm{IMU}}) \tag{17}$$

For LIS system, the LiDAR odometry corrects laser point cloud motion distortion by using IMU raw data and pre-integrated data. The current frame pose is updated by scan-to-map matching. When the pose change exceeds a certain threshold, it is seen as a new keyframe. For the second part of (17), the loose coupling of the GNSS is mainly reflected as a factor of the graph optimization in part D. Likewise, the GNSS/Vision/IMU sensor fusion strategy can be abstractly expressed as:

$$\boldsymbol{X}_t = \boldsymbol{A}(\boldsymbol{Q}_t^{\mathrm{Vision}}, \boldsymbol{Q}_t^{\mathrm{IMU}}) + \boldsymbol{B}(\boldsymbol{p}_t^{\mathrm{GNSS}}, \boldsymbol{A}(\boldsymbol{Q}_t^{\mathrm{Vision}}, \boldsymbol{Q}_t^{\mathrm{IMU}})) \tag{18}$$

For VINS system, the observation pre-processing section contains image data tracking and IMU data pre-integration. After that, the initialization part based on image feature points and IMU data, provides all necessary values, including scale and bias. For the scale, the system uses IMU pre-integrated values so that the scale information in pure-vision can be matched to the IMU measurements. The initial value of the LIS scan-to-map matching from LiDAR/IMU also can be used as an initial value for the VINS system to optimize the visual reprojection error and IMU measurement error, which can speed up the LIS system and improve its accuracy. After the initialization, the parameters enter the tightly coupled monocular VINS system based on a sliding window.

We define the state within the sliding window as:

$$\chi = [\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n, \lambda_0, \lambda_1, \ldots, \lambda_m] \tag{19}$$

$$\boldsymbol{x}_k = [\boldsymbol{p}_k^{\omega}, \boldsymbol{v}_k^{\omega}, \boldsymbol{q}_k^{\omega}, \boldsymbol{b}_a, \boldsymbol{b}_g], k \in \hat{\mathrm{E}}[0, n] \tag{20}$$

where $\boldsymbol{x}_k$ denotes the corresponding state of the IMU at frame k, λ_l denotes the value of the inverse depth of the feature point relative to its first observation; n denotes the number of keyframes. m denotes the number of all feature points in the sliding window; $\boldsymbol{p}_k^{\omega}$ denotes the pose of IMU in the world

TABLE 1. **Basic information for the three datasets.**

Dataset	Distance（m）	Times（s）	Environment	Source
Data I	694.697 m	115s	Short-range rural area	KITTI
Data II	1705.051 m	165s	Medium-range rural area	KITTI
Data III	4.51km	1536s	Long-range urban area	UrbanNav

(a) (b)

FIGURE 2. **Overview of the trajectories on Google Maps of the Data I (a) and Data II (b).**

coordinate system, $\boldsymbol{v}_k^{\omega}$ denotes the velocity of IMU, and $\boldsymbol{q}_k^{\omega}$ denotes the orientation of the body coordinate system at this time; $\boldsymbol{b}_a$ and $\boldsymbol{b}_g$ denote accelerometer bias and gyroscope bias, respectively.

Then the optimization function is written in the form of an equation as follows:

$$\min_{x}\left\{\begin{array}{c}\left\|\boldsymbol{r}_p-\boldsymbol{H}_p\chi\right\|^2+\sum_{k\in B}\left\|\boldsymbol{r}_B(\hat{z}_{b_{k+1}}^{b_k},\chi)\right\|^2\\+\sum_{(i,j)\in C}\rho(\left\|\boldsymbol{r}_C(\hat{z}_l^{c_j},\chi)\right\|^2)\end{array}\right\} \tag{21}$$

where the optimization cost function consists of three components. In the first part, i.e., the priori term, $\boldsymbol{r}_p$ and $\boldsymbol{H}_p$ denote the information obtained after marginalization; In the second part, b_k denotes the local frame, $\boldsymbol{r}_B(\hat{z}_{b_{k+1}}^{b_k},\chi)$ denotes the measured residuals of the IMU estimator $\hat{z}_{\mathrm{b}_{k+1}}^{\mathrm{b}_k}$; In the third part, $\boldsymbol{r}_C(\hat{z}_l^{c_j},\chi)$ denotes the visual reprojection residuals of the visual estimator $\hat{z}_l^{c_j}$, $\rho()$ denotes robust huber norm; B denotes the set of IMU data and C denotes the set of feature points in the sliding window that has at least two observations.

FIGURE 3. **Overview of the trajectory on Google Maps of the Data III.**

By relocalization, the frame of the sliding window is associated with the past pose. Then the pose graph optimization gets started to ensure that past poses are loaded correctly into the global map. The camera keyframes processed by VINS are added to the global pose graph. On the other hand, the sum of errors is reduced, because the difference between GNSS poses and VINS poses at the same timestamp is evenly distributed between every two adjacent VINS keyframes. In the ENU world coordinate system, the first GNSS measurement is used as the origin, and the measurement is made with the GNSS factor expressed as follows:

$$z_t^{\mathrm{GNSS}}-\boldsymbol{h}_t^{\mathrm{GNSS}}(\chi)=z_t^{\mathrm{GNSS}}-\boldsymbol{h}_t^{\mathrm{GNSS}}(\boldsymbol{x}_t)=\boldsymbol{p}_t^{\mathrm{GNSS}}-\boldsymbol{p}_t^{\mathrm{w}} \tag{22}$$

where z denotes the GNSS measurement, χ denotes the state prediction and $\boldsymbol{h}$ denotes the observation equation. $\boldsymbol{p}_t^{\mathrm{w}}$ denotes the estimated position value.

The bag-of-words model DBoW2 for loop closure detection is used by the visual odometry to enable closed-loop detection as well. For each new picture keyframe, the BRIEF descriptor is retrieved and compared with the previous descriptor for the bag-of-words. The timestamp of the loop closure frames returned by DBoW2 is sent to the LIS, which performs further loop closure verification.

After processing by motion bundle adjustment, we can get the robot pose at different rates through the VINS system.

It should be noted that we employ the principle of partial belief in selecting sensor parameters in this study. Specifically, upon initialization, the sensor parameters are independently computed and then integrated with the provided parameters. This approach effectively mitigates the influence of incorrect parameter assumptions to a certain extent, thereby enhancing the system's robustness. Notably, this constitutes a significant advantage of the proposed system discussed in this paper.

D. MULTI-SENSOR FUSION POSITIONING SYSTEM BASED ON GRAPH OPTIMIZATION

In the back-end optimization section, we use the factor graph optimization method to optimize the obtained pose information which is depicted in Fig. 1. The factor graph is an undirected graph consisting of two types of nodes: variable nodes representing optimization variables and factor nodes representing factors. Since the state estimation and sensor integration problem can be formulated as a typical Maximum A Posteriori (MAP) problem [32], subject to a Gaussian noise

FIGURE 4. Typical urban canyon environment (a) and repetitive structural roads (b) in Data III.

distribution model, we are able to transform the problem into a nonlinear least-squares problem. In this paper, the final cost function to be optimized is as follows:

$$\delta \boldsymbol{X}_i = \underset{\chi}{\arg\min} \sum_i \begin{pmatrix} \|\mathrm{F_{GNSS}}\|^2_{\boldsymbol{\Omega}_{\mathrm{GNSS},i}} + \|\mathrm{F_{IMU}}\|^2_{\boldsymbol{\Omega}_{\mathrm{IMU},i}} \\ + \|\mathrm{F_{LiDAR}}\|^2_{\boldsymbol{\Omega}_{\mathrm{LiDAR},i}} + \|\mathrm{F_{loop}}\|^2_{\boldsymbol{\Omega}_{\mathrm{loop},i}} \end{pmatrix} \tag{23}$$

where the $\mathrm{F_{GNSS}}$, $\mathrm{F_{IMU}}$, $\mathrm{F_{LiDAR}}$, and $\mathrm{F_{loop}}$ denote the error function of the four factors for the constraints (the IMU pre-integration factor, the LiDAR odometer factor, the GNSS factor, and the loop closure factor, which can also be found in Fig. 1); $\boldsymbol{\Omega}_{\{sensors\},i}$ denotes the information matrix which indicates the weighting of the sensors, which is adjustable, and will be adapted to the specific type of sensor and expectations; The robot states $\boldsymbol{X}_i$, i.e., the positions of driving vehicles, are also drawn in Fig. 1 as the yellow nodes.

The poses of all keyframes are updated by the optimization of the factor graph, which includes the IMU pre-integration factor and LiDAR odometry factor. Additionally, when the calculated position covariance is greater than the received GNSS position covariance, the GNSS factor is added to the factor graph. The closed-loop constraint factor is a new constraint factor that is added to the factor graph as VINS provides loop closure information to the LIS system.

III. EXPERIMENTAL SETUP

In order to test the performance of this semi-tightly coupled system, also compare the characteristics and effects of different sensor fusion strategies, we selected several different sensor systems for comparison (hereafter collectively referred to as strategies). Following the coupling form of typical sensors as well as the theory of the methodology above, we obtain five different systems: the GNSS strategy (single GNSS receiver solver), the GNSS/LiDAR/IMU strategy (Refer to (17)), the GNSS/Vision/IMU strategy (Refer to (18)), LiDAR/Vision/IMU strategy (the tight strategy of this paper without GNSS module), and GNSS/LiDAR/Vision/IMU strategy (the semi-tightly strategy proposed in this paper).

We chose three datasets for our experiments in order to examine the precision and reliability of the semi-tightly coupled system strategy in comparison to other four strategies.

As shown in Table 1, Data I and Data II are selected from the KITTI typical dataset which is collected in and around Karlsruhe, Germany [33]. Data I and Data II are primarily intended to test the accuracy characteristics of the strategy. We can therefore categorize the two datasets as being from rural areas by looking at the top view of the two datasets in Fig. 2 (a) and Fig. 2 (b) on Google Maps with the approximate trajectories aligned with yellow lines. Furthermore, Data I is utilized as the short-range rural region and Data II is used as the medium-range rural area. Together, they have a combined length of 1705.051 m and 694.697 m. The difference in distance can be used to test the performance of the relative positioning sensors.

Data III is selected from UrbanNav's Hong Kong Datasets [34], which is a localization dataset collected in Asian urban canyons. The overview of the Data III trajectory embedded in Google Maps is shown in Fig. 3. Since its long-range distance, is also harsh and complex urban environments (Fig. 4) for sensors, especially GNSS, as a GNSS challenging environment, it is primarily used to test the characteristics and strength of each strategy and is one of the main datasets we need to analyze. Like the treatment of Data I and Data II, Data III is classified as a long-range urban area.

The ground truth values (reference data) of Data I and Data II are directly given by the output of the GPS/IMU localization, which has a resolution of 0.02 m / 0.1°. For Data III, the ground truth of the vehicle is provided by high-precision RTK/INS processing using the SPAN combination system and IE software of NovAtel. To ensure the consistency of the experiments, the three datasets are treated in the same

TABLE 2. Summary of information on the configuration and processing of each data sensor.

Dataset	Info.	GNSS	LiDAR	Vision	IMU
Data I	Type	OXTS RT3003	HDL-64E Velodyne	FL2-14S3 M-C	OXTS RT3003
	Spec.	L1/L2 RTK correction	64 beams, field of view: 360° horizontal, 26.8° vertical, range: 120m	rectified images, 1226*370	6 axis
	Freq.	10 Hz	10 Hz	10 Hz	100 Hz
Data II	Type	OXTS RT3003	HDL-64E Velodyne	FL2-14S3M-C	OXTS RT3003
	Spec.	L1/L2 RTK correction	64 beams, field of view: 360° horizontal, 26.8° vertical, range: 120m	rectified images, 1226*370	6 axis
	Freq.	10 Hz	10 Hz	10 Hz	100 Hz
Data III	Type	EVK-M8T (x1)	HDL 32E Velodyne	ZED2 Stereo	Xsens Mti 10
	Spec.	Kinematic, L1 SPP	360 HFOV, +10~ -30 VFOV, range: 80 m	Single left camera	AHRS
	Freq.	1 Hz	10 Hz	15 Hz	400 Hz

way, and the specific setups are listed in Table 2. Each set of experiments was run independently with the five strategies. Note that as the focus of the monocular vision module in this paper is on positional determination, LiDAR has an advantage in map building. The ability to build maps is partly reflected in the accuracy of positioning. As a result, this paper focuses on the simultaneous positioning capabilities of each strategy.

The metric used in this paper is the ATE (Absolute Trajectory Error), which is used to measure the absolute trajectory pose error. The absolute trajectory error is the direct difference between the estimated and true poses and is a very intuitive reflection of the accuracy of the algorithm and the global consistency of the trajectory. For the monocular camera used in the strategy of this paper, the similarity transformation matrix $\boldsymbol{S} \in \mathrm{Sim}(3)$ from the estimated poses to the true poses needs to be calculated, defining the ATE for frame i as follows:

$$\boldsymbol{F}_i = \boldsymbol{Q}_i^{-1} \boldsymbol{S} \boldsymbol{P}_i \tag{24}$$

where $\boldsymbol{Q}_i \in Sim(3)$ denotes the true pose and $\boldsymbol{P}_i \in \mathrm{Sim}(3)$ denotes the estimated pose; Sim (3) denotes the similar transformation group.

Then the RMSE (Root Mean Square Error) is calculated as follows:

$$\mathrm{RMSE} = \left(\frac{1}{m} \sum\nolimits_{i=1}^{m} \|\mathrm{trans}(\boldsymbol{F}_i)\|^2\right)^{\frac{1}{2}} \tag{25}$$

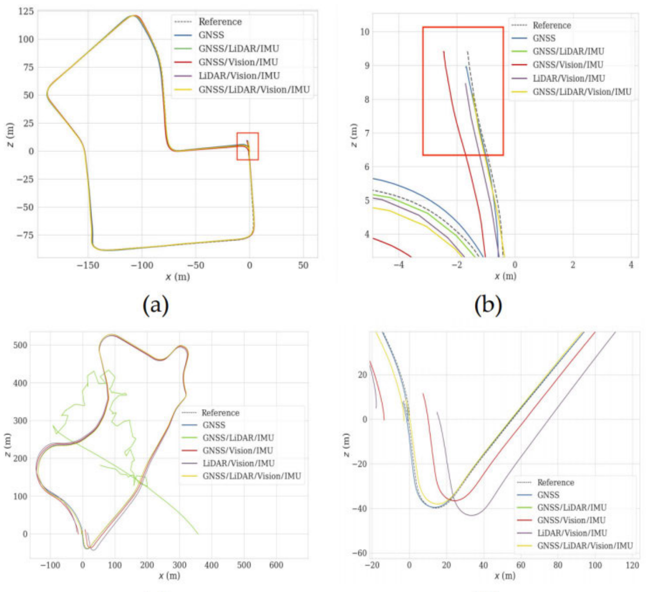
where $\mathrm{trans}(\boldsymbol{F}_i)$ represents the translational part of the $\boldsymbol{F}_i$.

IV. RESULTS AND DISCUSSION

We use the three separate datasets: Data I, Data II, and Data III to produce the experimental results for each of the five strategies (GNSS, GNSS/LiDAR/IMU, GNSS/Vision/IMU, LiDAR/Vision/IMU, and GNSS/LiDAR/Vision/IMU) in this part, which are then comprehensively described and analyzed.

A. PERFORMANCE ANALYSIS IN RURAL AREAS

First, we experiment with the semi-tightly coupled strategy proposed in this paper (represented as GNSS/LiDAR/Vision/IMU in the figures and tables below) and the rest of the typical strategies through Data I and Data II. Fig. 5 (a) and Fig. 5 (c) show the trajectories of these five strategies in the two datasets. In Fig. 5 (a), all five sensor strategies show good accuracy in this scenario when the sensors are functioning normally, meanwhile, the GNSS/LiDAR/IMU sensor fusion strategy fails at the coordinates (−110.67, −80.72) in Fig. 5 (a), which only shows the trajectory of the strategies before it fails. As shown in Fig. 5 (c), the GNSS/LiDAR/IMU sensor fusion strategy fails again and becomes unavailable. One of the causes of the error is that the point cloud's distortion removal fails due to the IMU's pose estimation error. To demonstrate the accuracy, Fig. 5 (b) shows the relatively discrete part of the trajectory during the running process. We can see the sensor's influence on the positioning effect by looking at the loop closure endpoint in Fig. 5 (b) and Fig. 5 (d). The worst loop closure effect is seen with the LiDAR/Vision/IMU sensor fusion strategy. Usually, position and rotational drift increase significantly with the distance travelled. With the addition of the GNSS factor to the factor graph, the cumulative drift error is significantly reduced, as evidenced by the fact that compared with the LiDAR/Vision/IMU sensor fusion strategy, the positioning trajectories of the semi-tightly coupled strategy in this paper are significantly more closely matched to the ground truth after the addition of GNSS loose coupling, as well as the



(a) (b)

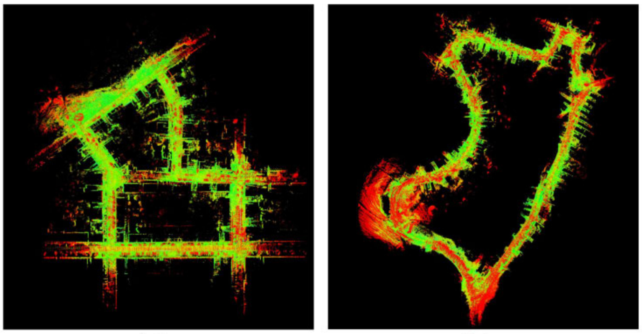
(c) (d)

FIGURE 5. Positioning trajectory of each strategy and enlarged view. The panel (a) is the graph of the global trajectories of the five strategies shown in the labels of the figure under Data I, and the panel (b) is the graph of the global trajectories of the five strategies shown in the labels of the figure under Data II; The panels (c) and (d) correspond to zoomed-in views of the final stages of the trajectories for data I and Data II, respectively.

semi-tightly coupled strategy has a positioning effect comparable to absolute GNSS positioning and a close alignment to the ground truth. As for the deviation at the starting point, it is related to the alignment used for the trajectory plotting. This does not affect the difference in distance between the head and tail of these trajectories.

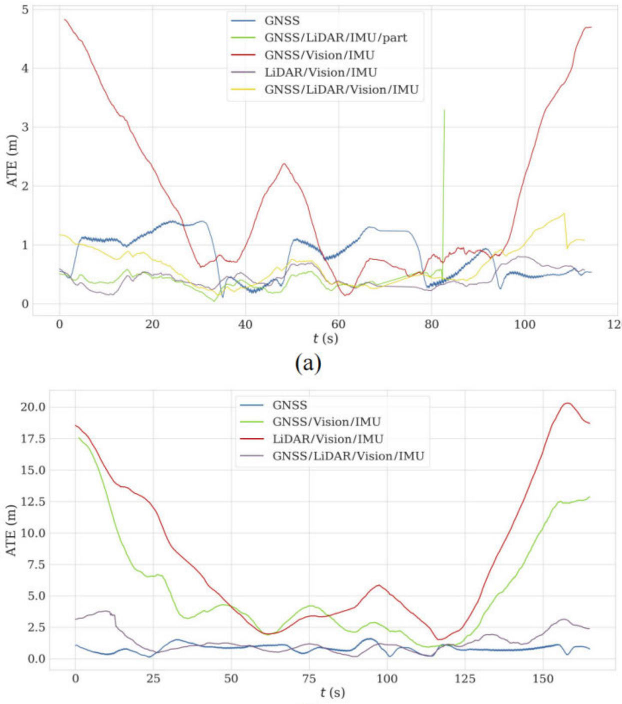
Besides, we are also able to obtain accurate laser point cloud maps with this semi-tight coupling strategy as shown in Fig. 6, thanks to the high precision and reliability positioning results obtained from the coupled multi-sensor data information. This will provide more operational space and foundation for subsequent work such as navigation and positioning based on a priori maps.

For the purpose of better demonstrating the effect of positioning in the overall scope, and also for more specific analysis and comparison, we plotted the ATE line graph of the semi-tightly coupled strategy proposed in this paper with other typical strategies over time as shown in Fig. 7. On Data I, we only show the data of the GNSS/LiDAR/IMU sensor fusion strategy for the effective part and which is named with "GNSS/LiDAR/IMU/part". In Data II, the trajectory of the GNSS/LiDAR/IMU sensor fusion strategy is eliminated because it is not available. From the graph, it is clear that even though the GNSS/Vision/IMU sensor fusion strategy is successfully implemented, its accuracy is the worst, especially at the beginning and ending parts, while the GNSS/LiDAR/IMU sensor fusion strategy has the highest precision among the five strategies when it is in effect.



(a) (b)

FIGURE 6. Plots of the point cloud mapping results for the semi-tightly fusion strategy of Data I (a) and Data II (b).



(a)

(b)

FIGURE 7. The absolute trajectory error line graph for each moment of the four strategies based on Data I (a) and Data II (b).

The beginning or end of the vehicle's movement has no bearing on GNSS, which exhibits a general accuracy globally. Meanwhile, the LiDAR/Vision/IMU sensor fusion strategy, the solution that removes the GNSS factor with general accuracy, shows the best results, achieving decimeter-level positioning over the full-time period on Data I; but on Data II, it is in the worst state and shows poor stability. In 84.34% of the travelled time in Data I and 39.09% in Data II, the semi-tightly coupled strategy (GNSS/LiDAR/Vision/IMU) achieves decimeter-level positioning, which guarantees high precision and excellent reliability. Furthermore, we also counted the RMSE and other statistical data in Table 3. With this table, it is easy to see the overall accuracy of each strategy when it is in effect.

TABLE 3. Statistics of each sensor fusion strategy for Data I and Data II. (m).

Data	Sensor Fusion Strategies	Max	Mean	Median	Min	RMSE
Data I	GNSS	1.397	0.786	0.764	0.137	0.889
	GNSS/LiDAR/IMU/part	70.833	24.395	18.848	1.647	0.431
	GNSS/Vision/IMU	3.295	0.379	0.363	0.062	2.231
	LiDAR/Vision/IMU	0.611	0.403	0.415	0.091	0.459
	GNSS/LiDAR/Vision/IMU	1.539	0.629	0.600	0.146	0.695
Data II	GNSS	1.626	1.626	0.806	0.124	0.860
	GNSS/LiDAR/IMU	-	-	-	-	-
	GNSS/Vision/IMU	17.542	5.417	3.8150	0.915	6.827
	LiDAR/Vision/IMU	20.304	7.973	5.480	1.513	9.804
	GNSS/LiDAR/Vision/IMU	3.793	1.327	1.095	0.175	1.589

B. PERFORMANCE ANALYSIS IN LONG-RANGE URBAN AREA

In this experiment, the environment is much more challenging, especially for GNSS, and we will test the precision and reliability of the semi-tightly coupled strategy proposed in this paper (GNSS/LiDAR/Vision/IMU) against other typical SLAM strategies in this experiment.

As to GNSS, it faces challenges in difficult urban environments since the reception of complex multipaths will seriously deteriorate the performance of precise positioning. For the GNSS strategy, the results of its Standard Point Positioning (SPP) are presented in Fig. 8. The red dots represent the position coordinates of the solution, and the green lines represent the trajectories that connect the solution coordinates by time series. It is easy to see that the vast majority of GNSS positioning results in urbanized areas do not match the actual needs. Compared to the GNSS positioning results of Data I and Data II in the rural areas by RTK calibration, more tall buildings in Data III cause high deterioration of GNSS signals with error phenomena.

FIGURE 8. GNSS position track after RTK processing. The red line indicates the reference trajectory.

To further test the effect of GNSS accuracy on the semi-tightly coupled strategy in this paper as well as other typical sensor strategies, we use GNSS signals based on GPS/INS high-precision RTK/INS processing in the four strategies other than GNSS as processing objects. The estimated trajectories of each strategy are shown in Fig. 9. As shown, the GNSS strategy positioning results are almost unusable. For the GNSS/LiDAR/IMU sensor fusion strategy, when passing through the tunnel as shown in Fig. 4 (a), LiDAR degrades since laser matching in this scene has an extra degree of freedom in one direction, i.e. the laser point cloud acquired on this path is the same, making it difficult for the matching algorithm to accurately estimate the motion in this direction. For the GNSS/Vision/IMU sensor fusion strategy, when the vehicle is stationary, the bias of the IMU continues to diverge, which can easily lead to the occurrence of drift phenomena, such as the drift of the vehicle's positioning trajectory at coordinates A (−366.805, 605.068) and B (−277.248, 577,591) in Fig. 9. And in the city, it is common for vehicles to stop moving. So, the GNSS/Vision/IMU sensor fusion strategy of positioning is not recommended. It is worth noting that because of the inclusion of the GNSS sensor, the strategy recovers the relative direction of motion at coordinates C (-399.145, 566.199), suggesting that the inclusion of GNSS can overcome errors in the relative positioning of the sensor through absolute positioning information. Lacking an absolute positioning sensor, the LiDAR/Vision/IMU sensor fusion strategy, despite running successfully, was unable to form a closed loop due to cumulative errors when the vehicle returned to the starting point at the end. The semi-tightly coupled strategy as proposed in this paper brings together the precision of LiDAR and uses Vision to improve the reliability of the solution, with the highest accuracy in the entire range, achieve the posi-

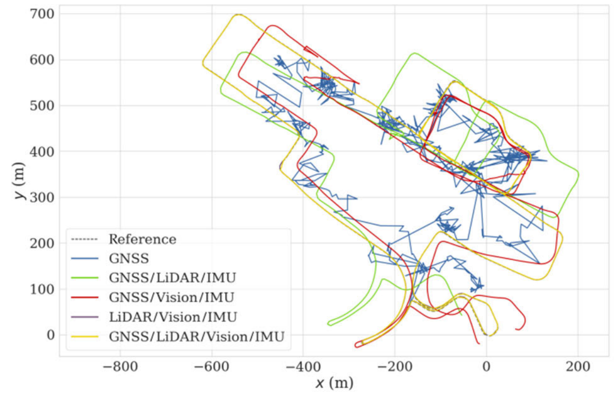


FIGURE 9. Trajectories of the five kinds of sensor fusion strategies using Data I. The blue, green, red, purple, and yellow lines denote the corresponding five strategies marked in the figure, respectively. The dashed line indicates reference trajectories.

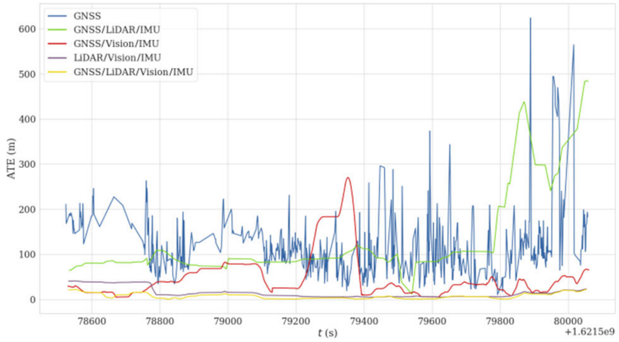


FIGURE 10. ATE line graph for each moment of the four strategies based on Data III over time.

tioning effect successfully based on the GNSS challenging environments. This confirms that the semi-tightly coupled strategy as proposed in this paper is precise and reliable in GNSS challenging environments, which is difficult to surpass with other typical sensor strategies. Note that due to the low frequency of the ground truth, we set the upper limit of the aligned timestamp parameter to 0.1, which is to better present comparisons between experimental results while ensuring that it is numerically significant.

At the data level, Fig. 10 shows the line graph of ATE over time, and we can see that GNSS performs the worst. Due to varying degrees of degradation of the LiDAR during the forward, middle and backward stages, a greater degree of error occurs in the alignment of the trajectory, resulting in the loss of high-precision characteristics. The GNSS/Vision/IMU sensor fusion strategy drift between the aforementioned points (A, B, and C), is represented in Fig. 10 as a sharp rise and fall in the time horizon (79200, 79400). The LiDAR/Vision/IMU sensor fusion strategy is slightly weaker than the semi-tightly coupled strategy of this paper at all global scales and the semi-tightly sensor fusion strategy has the highest precision and reliability.

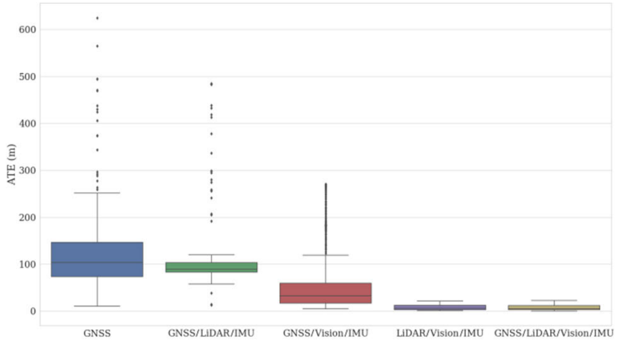


FIGURE 11. Box plot of overall ATE of the five strategies of Data III.

TABLE 4. Statistics of each sensor fusion strategy for Data II. (m).

Sensor Fusion Strategies	Max	Mean	Min	RMSE
GNSS	624.354	116.865	10.361	137.345
GNSS/LiDAR/ IMU	484.644	102.762	12.471	119.108
GNSS/Vision/ IMU	270.122	48.098	4.453	69.772
LiDAR/Vision/ IMU	41.150	13.777	4.981	17.537
GNSS/LiDAR/ Vision/IMU	22.295	7.393	0.235	9.593

To show the overall error, we calculated the cumulative ATE and plotted the corresponding box plots in Fig. 11, the specific indicators of which can be found in Table 4. From the specific data, we can see that the error decreases as the strategy changes from left to right.

Once again, we can get the point cloud maps through algorithms as Fig. 12. The top three panels of Fig. 12 show the global laser point cloud maps constructed under Data III for the three types of strategies containing LiDAR (GNSS/LiDAR/IMU sensor fusion strategy, LiDAR/Vision/IMU sensor fusion strategy, and the semi-tightly sensor fusion strategy of this paper). Panel (d) displays the point cloud created by the GNSS/LiDAR/IMU strategy for the tunnel section; As can be seen, the point cloud only includes the tunnel's entrance and exit points; the middle portion is compressed due to repetitive structures on both sides of the tunnel, which degrades the LiDAR. In panels (b) and (c), a comparison shows that the semi-tight combination strategy with the addition of the GNSS loose combination in the factor graph optimization builds better graphs, such as the tunnel region pointed to by the white arrows in the two figures, and the results in panel (c) are significantly better. In panel (e), the point cloud created by the final stage of vehicle operation is shown in the white box, and comparing it to panel (f), we can easily see the role of GNSS factor of graph optimization in reducing the global cumulative error, even if it is loosely coupled.

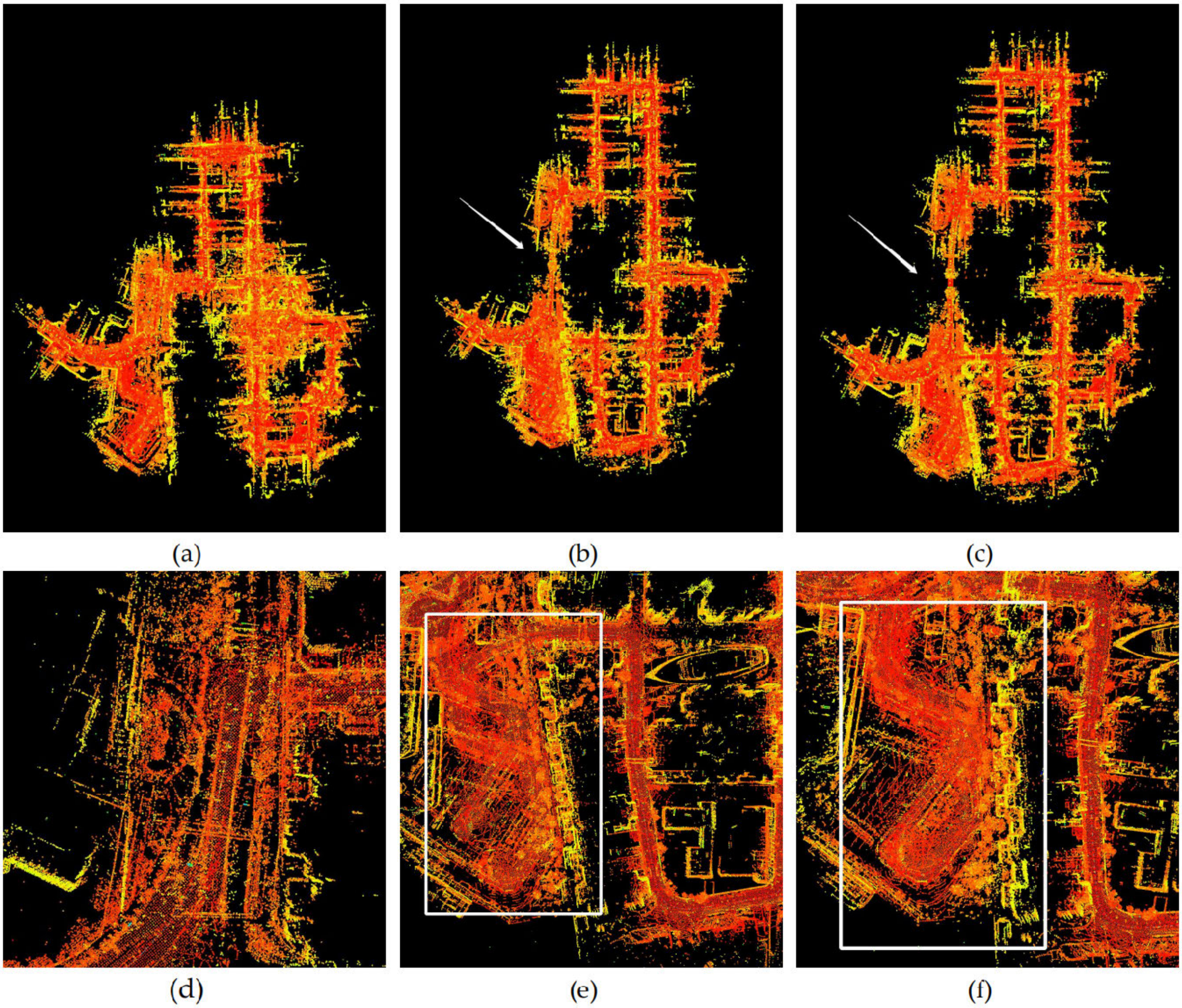


FIGURE 12. Plots of the point cloud mapping results for the three different strategies. The panels (a) , (b), and (c) denote the mapping effects of the GNSS/LiDAR/IMU sensor fusion strategy, LiDAR/Vision/IMU sensor fusion strategy, and the semi-tightly sensor fusion strategy, respectively. The panel (d) is the LiDAR degradation point cloud of (a); The panels (e) and (f) denote the map building effect in the final stage of (b) and (c), respectively.

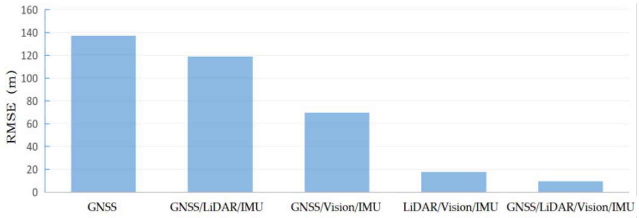


FIGURE 13. RMSE histogram for the five strategy of Data III. From left to right, GNSS, GNSS/LiDAR/IMU, GNSS/Vision/IMU, LiDAR/Vision/IMU and the GNSS/LiDAR/Vision/IMU sensor fusion strategies.

We estimated the RMSE of each approach on Data III, the same as how Data I was processed, and the results are displayed in Fig. 13. As can be seen, the worst performance was obtained by the GNSS strategy which only resulted in a RMSE value of 137.345. Meanwhile, the RMSE of the GNSS/LiDAR/IMU sensor fusion strategy, GNSS/Vision/IMU sensor fusion strategy, LiDAR/Vision/IMU sensor fusion strategy, and the semi-tightly sensor fusion strategy all demonstrate improvements when compared to the GNSS strategy. Specifically, the RMSE of the GNSS/LiDAR/IMU sensor fusion strategy was reduced by 13.28%, the GNSS/Vision/IMU sensor fusion strategy was reduced by 49.20%, the LiDAR/Vision/IMU sensor fusion strategy was reduced by 87.23%, and the semi-tightly sensor fusion strategy reduced by 93.02%. These improvements signify a significant enhancement in the quality of location data in GNSS challenging environments and provide robust support for our conclusions.

In conclusion, through three different data experiments, the different strategies in this paper show different characteristics. The GNSS strategy is not ideally accurate in most cases and is highly influenced by the environments, making it a great challenge in urban areas. The GNSS/LiDAR/IMU sensor fusion strategy is highly precision but extremely prone to failure and requires the fusion of other sensors to improve stability. The GNSS/Vision/IMU sensor fusion strategy is prone to drift in frequent stop scenarios and the incorporation of GNSS helps with absolute orientation recovery. The LiDAR/Vision/IMU sensor fusion strategy inherits the high precision of LiDAR, while Vision plays a crucial role in improving stability. The semi-tightly sensor fusion strategy adds a GNSS factor to the LiDAR/Vision/IMU sensor fusion strategy, which can improve the accuracy of the strategy very well during the loosely coupled process, can therefore be used as an additional complementary element.

V. CONCLUSION AND OUTLOOK

In this paper, we give an improved multi-sensor fusion positioning system based on GNSS/LiDAR/Vision/IMU with semi-tightly coupling and graph optimization. With three different levels of GNSS challenging environments, we have experimented and thoroughly analyzed the five sensor strategies mentioned in this paper in order to compare the characteristics of the semi-tightly coupled strategy with various conventional sensors strategies in SLAM and investigate the performance of all those strategies.

This semi-tightly coupled system is based on tight coupling of LiDAR, Vision, and IMU with a factor graph optimization approach at the back end, uses the absolute coordinate information generated by the GNSS sensor as the GNSS factor to join the factor graph for loose coupling. To compare the performance of this system with other strategies, in addition to the traditional GNSS strategy, according to the different forms of sensor combinations, we also give the GNSS/LiDAR/IMU sensor fusion strategy, GNSS/Vision/IMU sensor fusion strategy, and LiDAR/Vision/IMU sensor fusion strategy, in which LiDAR/Vision/IMU sensor fusion strategy is the system proposed in this paper that removes the loosely-coupled module of GNSS.

Each of these five strategies processes a short-range rural data, a medium-range rural data, and a long-range urban data. As the environments changes from rural to urban, the challenges for GNSS increase. Although GNSS has been certified to be able to provide absolute position information, urban canyon environments tend to limit its usefulness; Although having high precision, the GNSS/LiDAR/IMU sensor fusion system used in this paper is reliant on the IMU's proper operation and is prone to malfunction in repetitive structural environments easily; The GNSS/Vision/IMU sensor fusion system is more stable than the GNSS/LiDAR/IMU sensor fusion system, but its precision overall is average and it is subject to drift in urban environments; the addition of GNSS can correct for this drift and cumulative error; While the LiDAR/Vision/IMU sensor fusion system, which tightly couples LiDAR, Vision and IMU, achieves the medium precision and reliability by complementing the advantages of each sensor; With the addition of the factor graph optimization with GNSS, the semi-tightly coupled strategy has the highest precision and the best stability. In other words, this semi-tightly coupled strategy provides as much stability and feasibility as possible compared to the rest of the SLAM strategies, while significantly improving the positioning accuracy of the vehicle body compared to the GNSS strategy, leading to precise and reliable positioning results. Furthermore, compared to other multi-source sensor fusion algorithms, such as R3LIVE and LVI-SAM, R3LIVE does not incorporate GNSS for optimization, whereas the method in this paper has been tested to improve the reliability and overall robustness of GNSS compared to LVI-SAM.

By the way, we can see that currently, the integration of GNSS with other sensors is still mostly limited to loosely coupled method and the processing of GNSS information is relatively simple. It is still a challenging problem to realize high-precision GNSS positioning in most environments such as urban canyons, as well as the better fusion strategies and global constraint effects of GNSS in SLAM. We will attempt to achieve a tightly coupled effect of GNSS based on the semi-tightly sensor fusion strategy in this paper to further explore the methods of SLAM.

REFERENCES

[1] Y. Zhuang, X. Sun, Y. Li, J. Huai, L. Hua, X. Yang, X. Cao, P. Zhang, Y. Cao, L. Qi, J. Yang, N. El-Bendary, N. El-Sheimy, J. Thompson, and R. Chen, "Multi-sensor integrated navigation/positioning systems using data fusion: From analytics-based to learning-based approaches," *Inf. Fusion*, vol. 95, pp. 62–90, Jul. 2023.

[2] J. Liu, Y. Liang, D. Xu, X. Gong, and J. Hyyppä, "A ubiquitous positioning solution of integrating GNSS with LiDAR odometry and 3D map for autonomous driving in urban environments," *J. Geodesy*, vol. 97, no. 4, p. 39, Apr. 2023.

[3] L. Bofeng, *Theory and Method of Parameter Estimation in Mixed Integer GNSS Model*. Beijing, China: Surveying and Mapping Press, 2014.

[4] Z. Zhang, B. Li, and X. He, "Geometry-free single-epoch resolution of BDS-3 multi-frequency carrier ambiguities," *Acta Geodaetica Cartographica Sinica*, vol. 49, no. 9, pp. 1139–1148, Sep. 2020.

[5] D. Medina, H. Li, J. Vilà-Valls, and P. Closas, "Robust filtering techniques for RTK positioning in harsh propagation environments," *Sensors*, vol. 21, no. 4, p. 1250, Feb. 2021.

[6] Z. Zhang, Y. Li, X. He, and L. Hsu, "Resilient GNSS real-time kinematic precise positioning with inequality and equality constraints," *GPS Solutions*, vol. 27, no. 3, p. 116, Jul. 2023.

[7] P. Gakne and K. O'Keefe, "Tightly-coupled GNSS/vision using a sky-pointing camera for vehicle navigation in urban areas," *Sensors*, vol. 18, no. 4, p. 1244, Apr. 2018.

[8] N. Kubo, "Advantage of velocity measurements on instantaneous RTK positioning," *GPS Solutions*, vol. 13, no. 4, pp. 271–280, Sep. 2009.

[9] R. Gao, L. Xu, B. Zhang, and T. Liu, "Raw GNSS observations from Android smartphones: Characteristics and short-baseline RTK positioning performance," *Meas. Sci. Technol.*, vol. 32, no. 8, Aug. 2021, Art. no. 084012.

[10] D. L. Hall and J. Llinas, "An introduction to multisensor data fusion," *Proc. IEEE*, vol. 85, no. 1, pp. 6–23, Jan. 1997.

[11] C. Cadena, L. Carlone, H. Carrillo, Y. Latif, D. Scaramuzza, J. Neira, I. Reid, and J. J. Leonard, "Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age," *IEEE Trans. Robot.*, vol. 32, no. 6, pp. 1309–1332, Dec. 2016.

[12] 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 Sens.*, vol. 14, no. 12, p. 2835, Jun. 2022.

[13] T. Taketomi, H. Uchiyama, and S. Ikeda, ‘‘Visual SLAM algorithms: A survey from 2010 to 2016,’’ *IPSJ Trans. Comput. Vis. Appl.*, vol. 9, no. 1, Dec. 2017, doi: 10.1186/S41074-017-0027-2.

[14] T. Qin, P. Li, and S. Shen, ‘‘VINS-mono: A robust and versatile monocular visual-inertial state estimator,’’ *IEEE Trans. Robot.*, vol. 34, no. 4, pp. 1004–1020, Aug. 2018.

[15] J. Zhang and S. Singh, ‘‘LOAM: LiDAR odometry and mapping in real-time,’’ in *Proc. Robot., Sci. Syst.*, vol. 2, no. 9, 2014, pp. 1–9.

[16] Y.-T. Liu, R.-Z. Sun, X.-N. Zhang, L. Li, and G.-Q. Shi, ‘‘An autonomous positioning method for fire robots with multi-source sensors,’’ *Wireless Netw.*, vol. 2021, pp. 1–13, Mar. 2021.

[17] Z. Chen, A. Xu, X. Sui, C. Wang, S. Wang, J. Gao, and Z. Shi, ‘‘Improved-UWB/LiDAR-SLAM tightly coupled positioning system with NLOS identification using a LiDAR point cloud in GNSS-denied environments,’’ *Remote Sens.*, vol. 14, no. 6, p. 1380, Mar. 2022.

[18] C. Qian, H. Liu, J. Tang, Y. Chen, H. Kaartinen, A. Kukko, L. Zhu, X. Liang, L. Chen, and J. Hyyppä, ‘‘An integrated GNSS/INS/LiDAR-SLAM positioning method for highly accurate forest stem mapping,’’ *Remote Sens.*, vol. 9, no. 1, p. 3, Dec. 2016.

[19] X. He, W. Gao, C. Sheng, Z. Zhang, S. Pan, L. Duan, H. Zhang, and X. Lu, ‘‘LiDAR-visual-inertial odometry based on optimized visual point-line features,’’ *Remote Sens.*, vol. 14, no. 3, p. 622, Jan. 2022.

[20] T. Albrecht, K. Lüddecke, and J. Zimmermann, ‘‘A precise and reliable train positioning system and its use for automation of train operation,’’ in *Proc. IEEE Int. Conf. Intell. Rail Transp.*, Beijing, China, Aug. 2013, pp. 134–139, doi: 10.1109/ICIRT.2013.6696282.

[21] X. Liu, D. Li, J. Shi, A. Li, and L. Jiang, ‘‘A framework for low-cost fusion positioning with single frequency RTK/MEMS-IMU/VIO,’’ *J. Physics, Conf. Ser.*, vol. 1738, no. 1, Jan. 2021, Art. no. 012007.

[22] R. Mascaro, L. Teixeira, T. Hinzmann, R. Siegwart, and M. Chli, ‘‘GOMSF: Graph-optimization based multi-sensor fusion for robust UAV pose estimation,’’ in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2018, pp. 1421–1428.

[23] W. Lee, K. Eckenhoff, P. Geneva, and G. Huang, ‘‘Intermittent GPS-aided VIO: Online initialization and calibration,’’ in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2020, pp. 5724–5731.

[24] T. Qin, J. Pan, S. Cao, and S. Shen, ‘‘A general optimization-based framework for local odometry estimation with multiple sensors,’’ 2019, *arXiv:1901.03638*.

[25] 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 *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Oct. 2020, pp. 5135–5142.

[26] T. Shan, B. Englot, C. Ratti, and D. Rus, ‘‘LVI-SAM: Tightly-coupled LiDAR-visual-inertial odometry via smoothing and mapping,’’ in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, May 2021, pp. 5692–5698.

[27] W. Xuanbin, L. I. Xingxing, and L. Jianchi, ‘‘Tightly-coupled stereo visual-inertial-LiDAR SLAM based on graph optimization,’’ *Acta Geodaetica et Cartographica Sinica*, vol. 51, no. 8, pp. 1744–1756, 2022.

[28] Z. Zhang, H. Yuan, X. He, B. Chen, and Z. Zhang, ‘‘Unmodeled-error-corrected stochastic assessment for a standalone GNSS receiver regardless of the number of tracked frequencies,’’ *Measurement*, vol. 206, Jan. 2023, Art. no. 112265.

[29] A. Leick, L. Rapoport, and D. Tatarnikov, *GPS Satellite Surveying*. Hoboken, NJ, USA: Wiley, 2015.

[30] S. Wang, X. Dong, G. Liu, M. Gao, G. Xiao, W. Zhao, and D. Lv, ‘‘GNSS RTK/UWB/DBA fusion positioning method and its performance evaluation,’’ *Remote Sens.*, vol. 14, no. 23, p. 5928, Nov. 2022.

[31] G. Xu, Y. Xu, and G. Xu, ‘‘Applications of GPS theory and algorithms,’’ in *GPS: Theory, Algorithms and Applications*, 2016, pp. 313–340.

[32] T. D. Barfoot, *State Estimation for Robotics*. Cambridge, U.K.: Cambridge Univ. Press, 2017.

[33] A. Geiger, P. Lenz, and R. Urtasun, ‘‘Are we ready for autonomous driving? The KITTI vision benchmark suite,’’ in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.* IEEE, 2012, pp. 3354–3361.

[34] 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, J. Inst. Navigat.*, vol. 70, no. 4, 2023, Art. no. navi.602.

JIAMING ZHU received the B.Sc. degree in surveying and mapping engineering from Hohai University, China, where he is currently pursuing the M.Sc. degree in surveying and mapping engineering. His research interests include multi-source sensors and GNSS for positioning and map building in complex environments.

HAN ZHOU received the B.Sc. degree in surveying and mapping engineering from Hohai University, China. He has been engaged in some research under the guidance of his mentor for a long time. His research interests include GNSS and SLAM.

ZIYI WANG received the B.Sc. degree in surveying and mapping engineering and in computer science and technology from Hohai University, China. He has been engaged in relevant research under the guidance of his mentor for a long time. His research interests include computer vision, SLAM, and multi-source sensor fusion.

SULI YANG received the B.Sc. degree in surveying and mapping engineering from Hohai University, China. His research interests include SLAM and GNSS.

• • •