 https://doi.org/10.3311/PPtr.41522

Periodica Polytechnica Transportation Engineering, 54(3), pp. 364–375, 2026

Synchronous Localization and Mapping Model of Autonomous Driving Based on Multi-Sensor Fusion

Meng Wang[1*]

[1] Intelligent Technology Department, Tianjin Polytechnic College, No. 38, Xuehai Road, Beichen District, 300400 Tianjin, China
[*] Corresponding author, e-mail: MWwmWang@outlook.com

Received: 04 July 2025, Accepted: 07 November 2025, Published online: 20 March 2026

Abstract

With the development of intelligent transportation and autonomous driving technology, how to achieve precise and robust synchronous positioning and mapping has become a research focus. To improve the navigation accuracy and environmental perception capability of multi-sensor fusion systems in complex road environments, this study constructs a high-precision synchronous positioning and mapping model that integrates inertial measurement units, LiDAR, cameras, and wheel speed sensors. In addition, a closed-loop detection mechanism and graph optimization method have been introduced to enhance trajectory consistency and drift correction capability. The experiment showed that the positioning accuracy of the proposed method has increased from 85% to 97%, and the false alarm rate has decreased from 22% to 7%. The initial heading deviation was controlled within 0.8 degrees, with a root mean square error of 0.53 m and an average processing time of 60 milliseconds per frame. Further simulation showed that the model had an average positioning accuracy of 93.8%, detection coverage of over 95%, trajectory smoothness better than 0.07 m, and cumulative error drift controlled within 0.37 m/km in three typical road scenarios. The research shows that the model has significant performance in improving positioning accuracy, mapping consistency and environmental adaptability, and has the potential to be popularized and applied in actual autonomous driving systems.

Keywords

autonomous driving, synchronize positioning and mapping, revisit testing, graph optimization, multi-sensor fusion

1 Introduction

With the rapid development of intelligent transportation and autonomous driving technology, vehicles have put forward higher requirements for high-precision and highly robust positioning and environmental perception capabilities. In complex urban environments such as roads, tunnels, and highways, the positioning method of a single sensor often faces problems of sudden drops in accuracy and perception failure due to factors such as field of view obstruction, signal interruption, and environmental interference. Especially in scenarios where Global Navigation Satellite System (GNSS) signals are weak or lost, traditional systems that rely on inertial or visual navigation often experience trajectory drift or positioning failure due to error accumulation, seriously affecting vehicle decision-making and safety performance (Pan et al., 2025). In recent years, autonomous driving positioning and mapping technology has undergone an evolutionary path from single sensor navigation to multimodal sensor collaborative fusion. Among them, methods such as Extended Kalman Filter (EKF), graph optimization, visual inertial odometry, and LiDAR mapping have been widely studied and applied in synchronous positioning and mapping tasks. However, these technologies still face various challenges in complex dynamic scenarios. For example, the sampling frequency and delay of different sensors are inconsistent, the bias drift of Inertial Measurement Unit (IMU) sensors is serious, the visual feature matching is vulnerable to occlusion and light changes, and the laser radar mapping is difficult to achieve global optimization. These issues can make it difficult for the system to balance robustness, real-time performance, and accuracy (Trzeciak et al., 2023). Therefore, this study proposes a multi-sensor fusion-based automatic driving synchronous positioning and mapping model. This model constructs a unified spatiotemporal reference through joint state estimation and multi-source information fusion. It also introduces a closed-loop detection mechanism for loop optimization,

Cite this article as: Wang, M. (2026) "Synchronous Localization and Mapping Model of Autonomous Driving Based on Multi-Sensor Fusion", Periodica Polytechnica Transportation Engineering, 54(3), pp. 364–375. https://doi.org/10.3311/PPtr.41522

thereby improving the overall environmental adaptability and spatial consistency of the system. During the initialization phase, dead reckoning is used for short-term attitude estimation, and error correction and trajectory correction modules are designed to improve system continuity and accuracy. The purpose of this study is to provide new ideas on improving trajectory accuracy, navigation robustness and environmental adaptability, and lay a solid foundation for the practical deployment of the auto-drive system. The innovation of the research lies in proposing a complete multi-sensor fusion mapping framework and constructing a high-precision closed-loop detection and global graph optimization mechanism.

2 Related works

In recent years, with the rapid iteration of autonomous driving technology, the requirements for autonomous vehicles' perception and high-precision positioning have been continuously increasing. Wenzel et al. (2025) proposed a new visual autonomous driving and long-term positioning benchmark based on the large-scale 4Seasons dataset to promote the development of autonomous driving vision and long-term positioning under challenging conditions. This benchmark could jointly evaluate multiple performances, which was crucial for autonomous driving. The experiments provided new insights for current methods, demonstrating enormous potential for future research (Wenzel et al., 2025). Su et al. (2024) proposed a multi-sensor adaptive fusion Simultaneous Localization and Mapping (SLAM) framework based on degradation detection and deep reinforcement learning to address the difficulty in ensuring accuracy and performance of SLAM technology in complex dynamic environments. This framework could adaptively and collaboratively adjust fusion weights, adapt to different tightly coupled SLAM algorithms, and effectively improve the accuracy and performance of SLAM fusion in complex dynamic environments (Su et al., 2024). Liu et al. (2024) introduced a SLAM benchmark and a new method for improving map accuracy in order to objectively evaluate positioning and surveying errors and improve map accuracy. The proposed benchmark could be comprehensively and objectively evaluated, with an average improvement of 15% in map accuracy. The proposed algorithm outperformed traditional algorithms, especially when dealing with large-scale environments, providing new insights for obtaining ground truth maps and robot navigation.

Bai et al. (2024) proposed a new global registration technique to address the SLAM problem in deformable environments. By defining deformable transformations and utilizing kernel methods, this technique could globally solve deformable transformations and environmental mappings, and effectively solve scale ambiguity problems through optimized formulas. Experimental demonstrations and validations have shown that this method has superiority on multiple datasets. Guo et al. (2024) proposed an event-based rotation only bundling adjustment method to systematically study the problem of using event cameras for rotational motion estimation, and constructed the first event-based rotation only SLAM system. This method performed well on both synthetic and real datasets, and the published resources would help drive further research. Wang et al. (2024) proposed a new method that integrates an enhanced SLAM framework with a lightweight You Only Look Once Version 5 (YOLOv5) model to improve the robustness and accuracy of visual SLAM systems in indoor dynamic environments. This algorithm optimized the performance of feature point extraction, significantly improved the performance of dynamic indoor scenes, outperformed the original algorithm in accuracy, and had better real-time capabilities.

In summary, the current challenges in the field of autonomous driving include insufficient positioning and mapping accuracy, accumulation of sensor errors, and difficulties in multi-source information fusion. Although some research has achieved certain results in the fusion of single or dual sensors, there are still problems such as high False Alarm Rate (FAR), poor trajectory continuity, and insufficient closed-loop detection capability in complex road scenes. Therefore, this study constructs a synchronous localization and mapping model based on multi-sensor fusion. This model integrates inertial measurement, LiDAR, and visual and wheel speed information, and introduces closed-loop detection and graph optimization mechanisms for global constraint adjustment. Its purpose is to improve the navigation stability and mapping accuracy of the auto-drive system in the dynamic environment, and provide more reliable space support and path awareness for the high-level autopilot perception system.

3 Methods

3.1 Combined localization model based on IMU

IMU, as a high-frequency and low latency inertial sensor, can maintain continuous estimation of vehicle attitude and displacement in the event of GNSS signal obstruction or failure. Traditional pure IMU positioning suffers from error accumulation. This study adopts a combined navigation strategy of GNSS and IMU, and achieves complementary fusion of the two through EKF (Chen et al., 2024). The acceleration and

angular velocity data provided by IMU are used to predict the motion state of the vehicle, which is calculated based on the geocentric inertial coordinate system, as shown in Fig. 1. Fig. 1 not only illustrates the definition of different coordinate systems but also clarifies the transformation relationship between the IMU body frame, the geocentric system, and the navigation reference. It provides the mathematical foundation for subsequent attitude estimation and error correction, ensuring that the reader can clearly follow how sensor data are aligned in a unified reference space.

In Fig. 1, the coordinate system takes the center of the Earth as the origin, with the *X*-axis pointing towards the equator and the *Z*-axis pointing towards the north pole of the Earth's axis of rotation. The *Y*-axis forms a 3D coordinate system with the *X* and *Z* axes according to the right-hand rule. The geocentric coordinate system also takes the center of the Earth as its origin and shares the *Z*-axis with the geocentric inertial system, but its *X* and *Y* axes change direction with the rotation of the Earth (Dor et al., 2024). The world coordinate system is generally used to describe the motion state of objects that are independent of the Earth's motion, usually set at a stationary reference point. The IMU coordinate system takes the IMU body as the coordinate center, with the *X*-axis pointing forward and the *Y*-axis and *Z*-axis cooperating to form a 3D right-handed system, which is used to obtain the attitude information and dynamic parameters of the carrier. IMU typically consists of a three-axis accelerometer and a three-axis gyroscope, used to measure the linear acceleration and angular velocity of the carrier separately. The observation model of gyroscope is shown in Eq. (1).

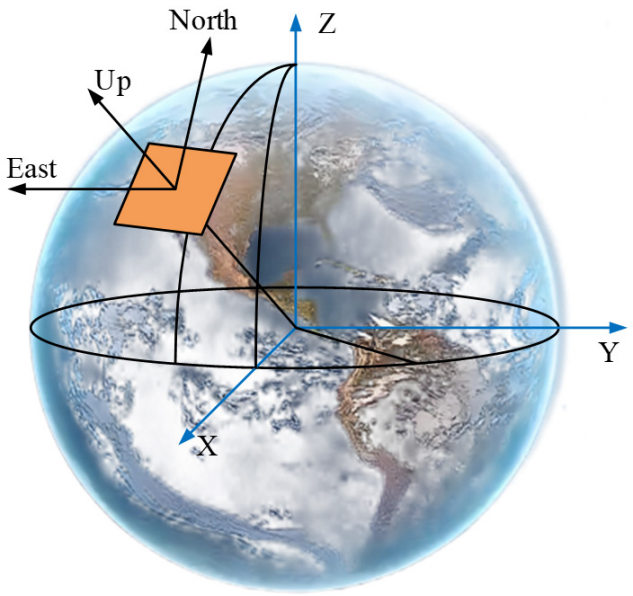

Fig. 1 Analysis of geocentric inertial coordinate system

$$\omega_m = \omega_b + b_g + n_g \tag{1}$$

In Eq. (1), ω_m is the measured angular velocity of the gyroscope, ω_b is the true angular velocity output of the gyroscope, b_g is the bias of the gyroscope, reflecting the inherent zero bias error of the sensor, n_g is the Gaussian white noise of the gyroscope. The expression of the accelerometer observation model is shown in Eq. (2).

$$a_m = \boldsymbol{R}\ \left(a_b - \boldsymbol{g}\right) + b_a + n_a \tag{2}$$

In Eq. (2), a_m is the acceleration measured by the accelerometer, a_b is the true linear acceleration of the carrier in the world coordinate system, $\boldsymbol{g}$ is the vector of gravitational acceleration, $\boldsymbol{R}$ is the rotation matrix estimated from the current pose, used to project the world coordinate acceleration onto the IMU coordinate system, b_a is the bias of the accelerometer, n_a is the measurement noise of the accelerometer (Zhao et al., 2025). The formula for attitude propagation is shown in Eq. (3).

$$q' = \frac{1}{2} q \otimes \omega_b \tag{3}$$

In Eq. (3), q is the unit quaternion of the attitude, $\otimes$ is the quaternion multiplication, and ω_b is the angular velocity after removing the bias. The formula for calculating the attitude change of the carrier based on the output of the gyroscope is the core formula for IMU attitude update (Du et al., 2024). Then, the expression for predicting the state of the target is shown in Eq. (4).

$$\begin{cases} P_{k+1} = p_k + v_k \Delta t + \dfrac{1}{2} a_b \Delta t^2 \\ v_{k+1} = v_k + a_b \Delta t \end{cases} \tag{4}$$

In Eq. (4), p_k and v_k are the current position and velocity, respectively, Δt is the sampling period, and a_b is the acceleration after removing the gravity term and rotating to the world coordinate system (Islam et al., 2024). However, collecting too much IMU data can result in a large computational load. Therefore, to achieve real-time estimation of the carrier's position, velocity, and attitude, this study uses EKF for state propagation and error correction of IMU measurement data. The system state vector includes quaternions of position, velocity, and attitude, and its propagation model is shown in Eq. (5).

$$\begin{aligned} p(t+\Delta t) &= \boldsymbol{p}(t) + \boldsymbol{v}(t) \times \Delta t \\ v(t+\Delta t) &= \boldsymbol{v}(t) + \left[R(q) \times \left(a_m - b_a\right) \right] \times \Delta t + g \times \Delta t \end{aligned} \tag{5}$$

In Eq. (5), $\boldsymbol{p}(t)$ is the position vector at the current time, $\boldsymbol{v}(t)$ is the current velocity vector, a_m is the acceleration value measured by IMU, and b_a is the accelerometer zero bias error. $R(q)$ is a rotation matrix composed of attitude

quaternions q. EK can improve the positioning continuity and robustness of the system in high-speed and weak GNSS environments. The process of data time synchronization and state estimation is shown in Fig. 2. This figure highlights the different sampling frequencies of IMU, LiDAR, camera, and wheel sensors, and demonstrates how multi-source data are aligned on a unified temporal axis. By presenting the interpolation of high-frequency IMU data with lower-frequency LiDAR and camera data, it visually supports the claim that the proposed method achieves high-precision, low-latency state estimation.

In Fig. 2, the timeline extends horizontally, representing the continuous acquisition process of data over time. Fig. 2 includes IMU frame, laser radar frame, camera frame and wheel speed meter frame. The data sampling frequency of different sensors is different. Among them, IMU has the highest sampling frequency, so the inter frame interval is the shortest, while other sensors such as LiDAR and camera frames appear intermittently at lower frequencies. The green semi-transparent rectangular area represents the sampling time interval of the LiDAR. Within this interval, IMU will provide continuous data for interpolation and attitude extrapolation, ensuring that each frame of radar point cloud can be registered at a unified reference time. The system updates the state at each keyframe moment, integrating multi-source information such as IMU, LiDAR, camera, and wheel speed sensor within the corresponding time window to achieve high-precision and low latency state estimation and sensor data alignment (Wang et al., 2023; Xu et al., 2024). Due to issues such as zero bias drift, high measurement noise, and error accumulation in IMU, a Strapdown Inertial Navigation System (SINS) is introduced to model and compensate for them. Through SINS, the problem of decreased accuracy of IMU during long-term operation can be effectively improved, and high-frequency continuous estimation of attitude, velocity, and position can be achieved, providing support for stable navigation in complex environments. Its structure is shown in Fig. 3.

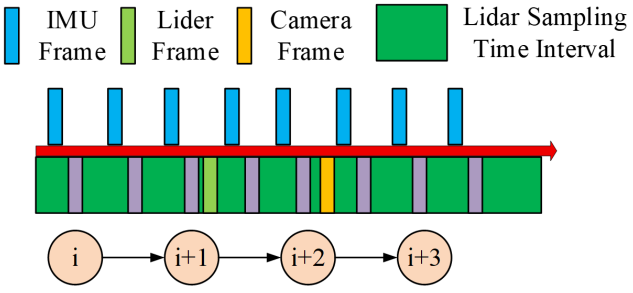

Fig. 2 Data time synchronization and state estimation process

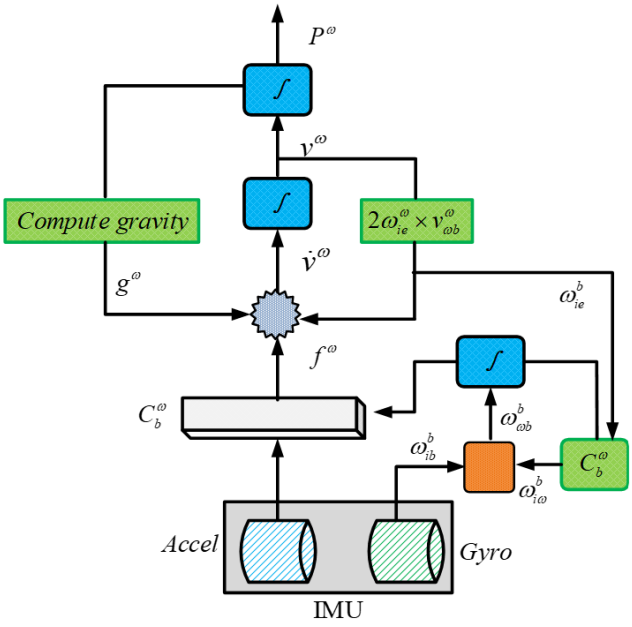

Fig. 3 Strapdown inertial machinery organization algorithm model

In Fig. 3, the IMU section includes an accelerometer and a gyroscope, which are used to measure linear acceleration and angular velocity, respectively. The angular velocity output by the gyroscope is integrated and used to calculate the attitude change of the carrier, and obtain the attitude transformation relationship from the body coordinate system to the navigation coordinate system. Subsequently, the acceleration measured by the accelerometer is converted to the navigation coordinate system based on attitude transformation, and combined with the gravitational acceleration of the Earth and the Coriolis acceleration caused by the Earth's rotation to obtain the total acceleration in the navigation system. The acceleration is first integrated once to obtain velocity, and then integrated twice to obtain position, thus achieving the complete navigation calculation process.

To quantitatively analyze the influence of IMU zero-offset drift and random noise, this study employed Allan variance analysis to characterize the stochastic error components of the gyroscope and accelerometer. The bias instability of the gyroscope was measured at approximately 0.012 °/s, while the accelerometer bias instability was about 0.0009 m/s². The random walk coefficients were 0.18 °/√h for angular velocity and 0.07 m/s/√h for linear acceleration. These parameters were incorporated into the Strapdown Inertial Navigation System (SINS) error model to simulate cumulative drift in long-term operation. The error compensation strategy introduced in this study dynamically corrected bias terms through online estimation, and attenuated high-frequency noise by integrating a low-pass filtering step within the Extended Kalman Filter (EKF).

This ensured that systematic drift and random disturbances were effectively mitigated before state propagation.

3.2 Mapping model based on multi sensor fusion

After completing the construction of the IMU-based combined positioning model, although a single sensor can achieve short-term and high-frequency state estimation, it is still difficult to ensure the global consistency and environmental cognitive accuracy of the system in long-term operation or complex environments. To this end, this study further introduces a multi-sensor fusion strategy, combining multi-source observation information such as LiDAR, cameras, and wheel speed sensors to jointly map and constrain environmental features, and construct a more robust and accurate synchronous positioning and mapping system. The specific fusion mapping algorithm framework is shown in Fig. 4.

In Fig. 4, the system mainly consists of a Laser Radar Inertial Odometer Subsystem (LRIOS), a Vision Inertial Odometer Subsystem (VIOS), and an IMU module, which achieve high-precision positioning and mapping through multi-sensor data fusion. LRIOS, with LiDAR as its core, sequentially completes B-spline motion compensation, feature extraction, adaptive feature constraints, and plane cylinder residual calculation, and establishes a connection with the map through absolute constraints. VIOS is mainly based on camera input, sequentially performing feature extraction, image block construction, sliding window optimization, and backlight residual calculation, and using the results for state extension and map point projection to form relative constraint information. In addition, data from wheel speed sensors or other sensors can also participate in the pre-integration process to assist VIOS in pre-propagation (Yin et al., 2025). The IMU module runs through the entire system process, first aligning with LiDAR and camera data in the initialization stage, then providing high-frequency pose prediction information through the state propagation module, and participating in joint optimization of multi-source data in the subsequent Kalman filter. To further enhance the adaptability of the proposed model under adverse weather conditions, this study introduces a millimeter-wave radar (mmWave) fusion layer. Millimeter-wave radar has strong penetration capability in fog, rain, and snow, which effectively compensates for the degradation of LiDAR performance under low-visibility scenarios. In the fusion architecture, mmWave radar data are incorporated into the mapping subsystem through a probabilistic occupancy grid representation, and the extracted range-Doppler features are aligned with LiDAR point cloud features via a Kalman-based sensor fusion strategy. By introducing this layer, the system improves target detection reliability and robustness of environmental perception in challenging weather conditions. To improve the consistency and robustness of mapping, the system introduces a closed-loop detection (LCD) mechanism to identify whether the current frame forms a closed loop with historical frames. After detecting similar frames, the system constructs a closed-loop error term and introduces a Graph Optimization (GO) module for overall graph structure adjustment. To evaluate the computational requirements of the proposed multi-sensor fusion framework, the time complexity of the main modules was analyzed. The LiDAR-Inertial subsystem involves feature extraction and point cloud registration, with a computational complexity of approximately $O(n \log n)$, where

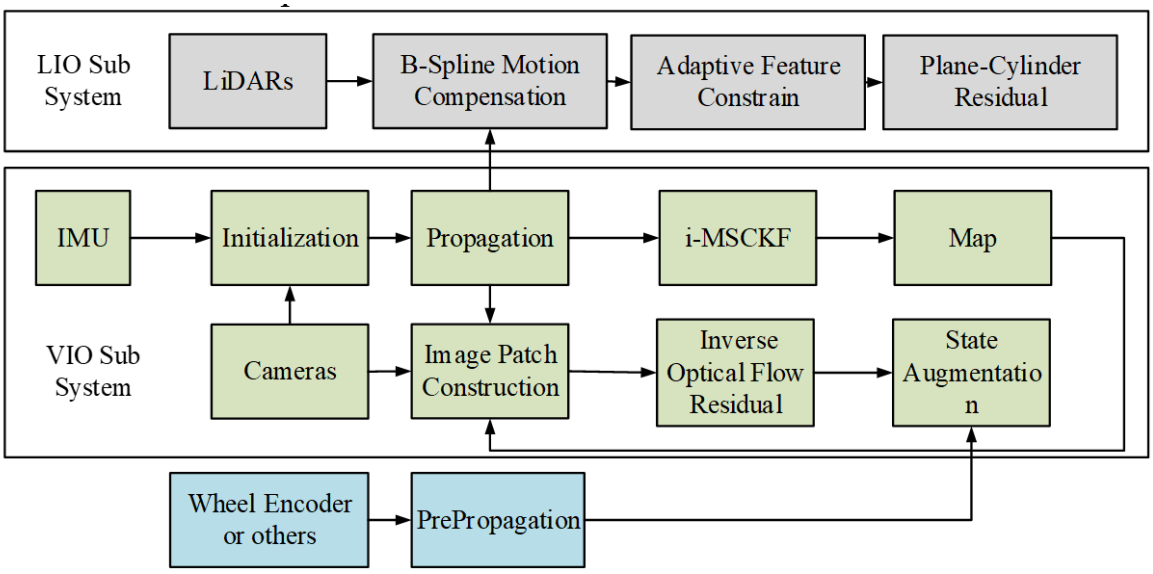

Fig. 4 Structure of mapping model based on multi-sensor fusion

n is the number of point cloud features. The Vision-Inertial subsystem performs sliding window optimization and bundle adjustment, which typically scale as $O(m^2)$, where m denotes the number of keyframes in the optimization window. The graph optimization module for loop closure has a worst-case complexity of $O(k^3)$ with respect to the number of pose nodes k, but in practice the sparsity of the optimization graph reduces the effective complexity to nearly linear. Under the experimental hardware configuration, the overall pipeline achieves real-time performance at 60 ms per frame, which is consistent with the theoretical complexity estimates when sensor data volumes are controlled by adaptive sampling strategies. This process essentially involves adding new constraint edges in the SLAM framework to achieve global correction of drift trajectories. The error residual term is shown in Eq. (6).

$$\boldsymbol{e}_{\text{loop}} = \boldsymbol{T}_{ij} \ominus \left(\boldsymbol{T}_j^{-1} \cdot \boldsymbol{T}_i \right) \circ \tag{6}$$

In Eq. (6), $\boldsymbol{T}_j$ and $\boldsymbol{T}_i$ are pose estimates for the current frame and historical frames, respectively $\boldsymbol{T}_{ij}$ is the relative transformation calculated through revisit matching between two frames, $\ominus$ is an algebraic residual operation. The goal of closed-loop optimization is to minimize the weighted sum of all residual terms in the entire graph, and the optimization problem can be formalized as shown in Eq. (7).

$$\min \sum \boldsymbol{w}_k \left\| e_k \right\|^2 \tag{7}$$

In Eq. (7), e_k denotes various types of observation residuals, and $\boldsymbol{w}_k$ is the corresponding information matrix or weight. The optimization method adopts nonlinear least squares implementation, and its structure is shown in Fig. 5.

In Fig. 5, firstly, the initial revisit area is located based on historical trajectory data, and the pose information is preprocessed and the trajectory endpoint status is monitored in real-time. If the trajectory has not reached the end, the system will extract the local pose waveform features of the current area and lock the equivalent mileage points in the measurement trajectory through mileage similarity comparison. Then it enters the fine matching stage, where the peak correlation coefficient between the two waveforms is calculated. If the correlation coefficient meets the threshold condition, it is judged as an accurate match, and the recognition result is output. If the optimal value is not reached, the waveform window length will be dynamically extended, and the matching operation will be re-executed within the extended interval until the global optimal matching position is obtained, ultimately achieving accurate calibration of revisit points. In the initialization stage of the IMU system, to obtain reliable initial attitude and position information, this study introduces Dead Reckoning (DR) method for short-term initial alignment. DR utilizes the continuous angular velocity and acceleration data output by IMU after startup to estimate the initial motion state of the carrier through integration. The process is shown in Fig. 6.

In Fig. 6, firstly, the three-axis angular velocity and linear acceleration signals are obtained by IMU, and the error correction unit eliminates device bias and random noise interference, improving the reliability of the original data. For closed-loop path scenarios, a closed-loop calibration module has been added to suppress cumulative drift errors and dynamically adjust the zero bias parameters of the closed-loop interval. The next step is to perform initial heading calculation, determine the reference horizontal attitude angle and reset the initial azimuth angle to zero, and establish an inertial calculation coordinate system. The dynamic DR mechanism is based on corrected angular velocity data

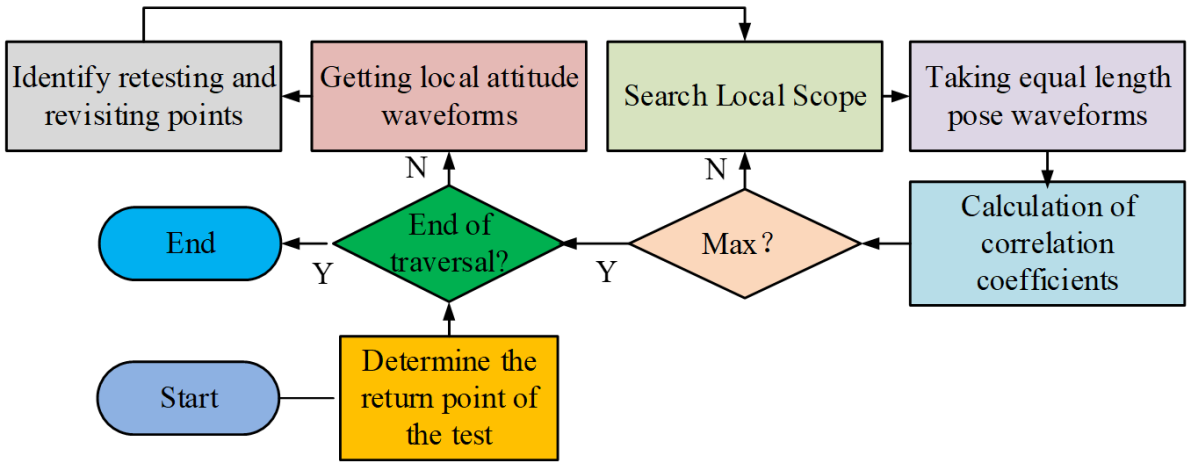

Fig. 5 Mechanism for closed-loop detection

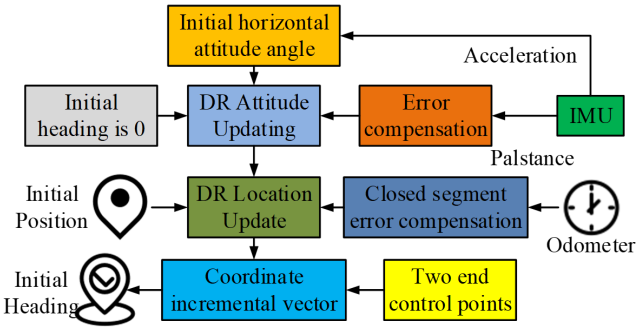

Fig. 6 Heading initialization model

to achieve real-time attitude calculation, and synchronously integrates odometer displacement data to complete position recursive estimation (Gheisari et al., 2022). To ensure trajectory continuity, the system incorporates the displacement increment output by the odometer into the coordinate update system and achieves real-time position iteration through the fusion operation of coordinates and directional increments. Finally, after global heading optimization compensation, high-precision initial azimuth parameters and spatial coordinate information are output.

4 Results

4.1 Performance analysis of synchronous positioning and mapping model for autonomous driving based on multi-sensor fusion

The hardware configuration uses Intel Core i5-8750H CPU, NVIDIA Geforce GTX2080Ti GPU, 8GB VRAM, and 16GB RAM. This study choosed the KITTI Dataset public dataset. The classic public dataset for autonomous driving research, released by the Karlsruhe Institute of Technology in Germany and the Toyota Joint Research Center, covers various sensor data in real urban road environments. The dataset uses Velodyne HDL-64E LiDAR to collect 3D point cloud data and is equipped with forward facing binocular cameras for image acquisition. IMU is used to measure vehicle acceleration and angular velocity. GPS/INS provides true values of geographic location and attitude. The data undergo high-precision time synchronization and spatial calibration. The data collection scenarios include urban streets, highways, rural roads, etc. The driving environment is real and complex, including multiple types of targets such as pedestrians, vehicles, and buildings. In order to further evaluate the contribution of IMU errors to system accuracy, an ablation experiment was conducted. Without error compensation, the average Root Mean Square Error (RMSE) of the trajectory increased from 0.53 m to 0.87 m, and the initial heading deviation expanded from 0.8° to 1.9°. Meanwhile, the false alarm rate rose to 0.15, showing the significant adverse impact of uncompensated bias drift and noise. After applying the proposed compensation strategy, the Allan variance-derived bias terms were suppressed by over 60%, and the positioning accuracy improved correspondingly. This quantitative evidence highlights that IMU zero-drift and noise are primary contributors to accumulated error and demonstrates that the designed compensation mechanism is effective in enhancing the stability and reliability of the system. This study selects traditional positioning models and GO-based positioning models as comparative models, as shown in Fig. 7.

Figs. 7 (a) and (b) show the accuracy changes and FAR performance of three methods under different dataset sizes. These results directly demonstrate that the proposed model consistently outperforms both traditional and GO-based approaches. In particular, the steadily higher accuracy curve and the significantly reduced false alarm rate visually validate the effectiveness of the multi-sensor fusion strategy, thereby reinforcing the core conclusion that the method provides superior robustness in large-scale datasets. In Fig. 7 (a), as the dataset size increases from 200 to 1,600, the accuracy of all three methods shows an upward trend. Among them, the research method has always maintained a leading position, with its accuracy steadily increasing from 0.85 to 0.97, significantly better than GO and traditional methods. The accuracy of GO is 0.76 when the data volume is 200, gradually increasing to 0.91. The traditional method starts at 0.72 and eventually increases to 0.90. In Fig. 7 (b), the overall FAR decreases with the increase of dataset size, indicating that more data helps reduce the model's judgment errors. The research method still performs the best throughout the entire process, with its FAR decreasing from 0.22 to 0.07, while GO's FAR decreases from 0.36 to 0.11, and the traditional method's FAR decreases from 0.34 to 0.13. This indicates that the proposed method not only enhances recognition accuracy but also effectively controls the occurrence of false positives, possessing higher robustness and credibility. The results of comparing the offsets of each model are shown in Fig. 8.

Figs. 8 (a) and (b) show the initial heading deviation angle comparison and accuracy performance of three methods in four tests. The initial heading deviation curves confirm the stability of the method in reducing attitude errors, while the accuracy comparison charts show that the performance advantage is maintained across multiple tests. Together, these visual results provide strong evidence supporting the claim that the proposed system achieves higher trajectory consistency and error correction capability than baseline methods. In Fig. 8 (a), in the first test, the traditional method has an

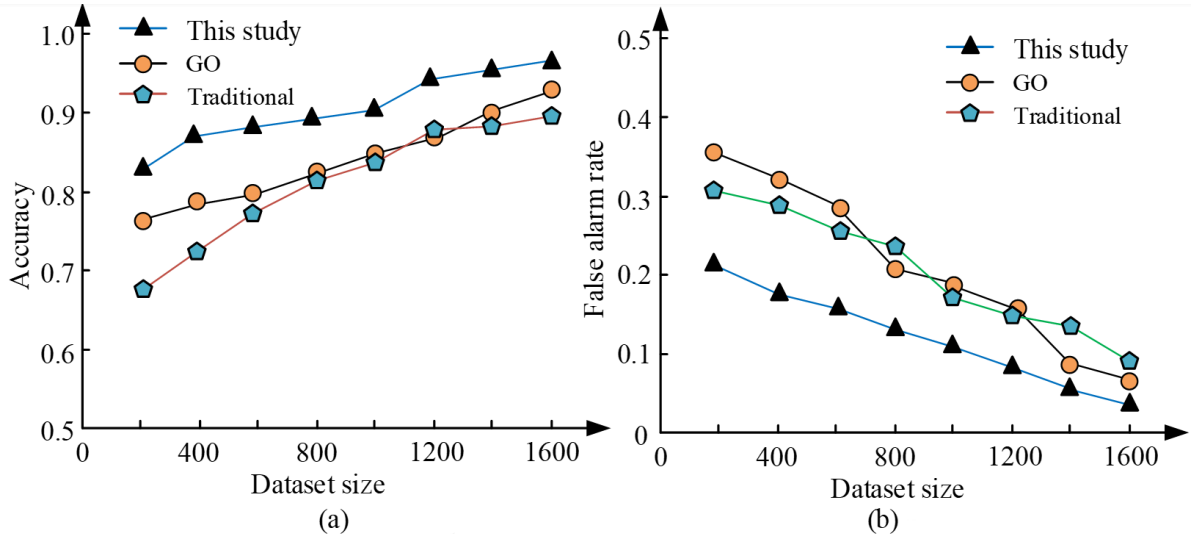

Fig. 7 Accuracy and FAR analysis of each model: (a) Accuracy; (b) False alarm rate

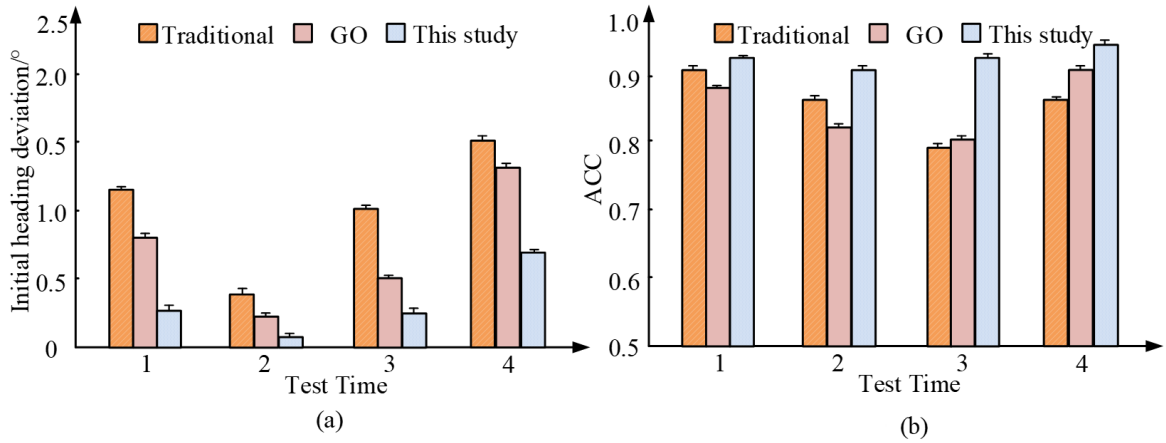

Fig. 8 Comparison of deviations among various models: (a) Initial heading deviation; (b) ACC

offset of nearly 2.1 degrees, the GO method has an offset of 1.2 degrees, while the research method only has an offset of 0.8 degrees. In the fourth test, the traditional method has a deviation of nearly 2.0 degrees, the GO method has a deviation of 1.7 degrees, and the research method has a deviation of less than 1.0 degree. In Fig. 8 (b), in the first test, there is not much difference in accuracy among the three methods, with that of the traditional method being 0.91 and the GO method being 0.94, while the proposed method achieves 0.96. In the second and third tests, the difference gradually widens. The accuracy of the traditional method is 0.86 and 0.81, while the GO is 0.90 and 0.85. The proposed method achieves 0.95 and 0.91. In the fourth test, the accuracy of the research method reaches 0.96, significantly better than GO's 0.92 and the traditional method's 0.89. The experiment shows that the research method has stronger dynamic robustness and optimization mechanism. Table 1 analyzes the comprehensive performance of each model.

Table 1 Comprehensive performance analysis

Model	Traditional	GO	Proposed
Accuracy (%)	90	91	96
False Alarm Rate	0.13	0.11	0.07
Initial Heading Deviation (°)	2.0	1.7	0.8
Mean Offset Error (m)	0.75	0.62	0.41
RMSE (m)	0.82	0.71	0.53
Processing Time (ms/frame)	48	55	60
Robustness Score	6.8	7.5	9.2
Memory Usage (MB)	610	680	725

In Table 1, from the perspective of accuracy, the proposed method achieves 96%, significantly better than GO's 91% and traditional method's 90%, indicating its stronger recognition ability in complex traffic scenarios. The FAR of the research method is 0.07, which is lower than the GO's 0.11 and the traditional method's 0.13, indicating that it effectively reduces misidentification while improving accuracy.

The initial heading deviation is an important indicator for evaluating the initial positioning accuracy. The accuracy of the research method is 0.8 degrees, which effectively reduces the risk of initial error accumulation compared to GO's 1.7 degrees and traditional methods' 2.0 degrees. In terms of Root Mean Square Error (RMSE), the research method has a value of 0.53 m, which is better than GO's 0.71 m and the traditional method's 0.82 m, indicating that it has higher stability and continuity in the entire positioning process. Although the processing time of the research method is slightly higher at 60 ms/frame, it is within an acceptable range and the performance improvement is sufficient to compensate for its slight disadvantage in computational efficiency. To sum up, the proposed method shows stronger comprehensive advantages in a number of key performance indicators, and is suitable for autonomous driving systems with high accuracy and reliability. In order to further contextualize the reported processing time, a comparative time analysis was conducted. The proposed model maintained an average of 60 ms/frame, while the GO-based baseline required 55 ms/frame and the traditional model 48 ms/frame. Although the complexity of the proposed approach is slightly higher due to multi-source fusion and global optimization, the additional overhead corresponds to a relative increase of only 9.1% compared to the GO method. Considering that the accuracy improved from 91% to 96% and the false alarm rate dropped from 0.11 to 0.07, the trade-off between complexity and performance is justified. Moreover, the modular design of the system allows for optimization of graph sparsification and adaptive keyframe selection, which can further reduce time complexity toward $O(n)$ in practical deployments.

4.2 Simulation analysis

To further validate the performance of the proposed model, this study selects actual situations to simulate and analyze the performance of the model, as shown in Fig. 9.

Fig. 9 (a) shows the variation of the north-south error of the model under two testing conditions, Test 1 and Test 2. Fig. 9 (b) shows the error variation in the east-west direction. In Fig. 9 (a), within a total length of 200 m, the error curves of the two sets of tests show an overall consistent trend, with values mainly distributed between −3 and 3 m, and most errors concentrated within ± 1.5 m, with small fluctuations. For example, at distances of about 80 m and 160 m, there are significant peaks in the errors of both tests, but they do not exceed ± 3 m, indicating that the model's positioning accuracy in the north-south direction was stable and consistent. In Fig. 9 (b), the error fluctuation range is slightly larger. Although it is still within ± 4 m overall, there are local abnormal peaks close to 5 m, which are evident at around 120 m and 180 m. This indicates that the two tests still show good consistency in overall trends and can maintain relatively consistent localization performance in real-world scenarios. Fig. 10 analyzes the positioning effects of different routes.

Figs. 10 (a) and (b) show the synchronized positioning and mapping trajectories constructed by autonomous vehicles in two typical urban road network environments. By showing smoother path trends and accurate obstacle avoidance, the figures intuitively confirm that the model not only achieves stable pose estimation but also ensures reliable navigation in both narrow and wide road networks. These visual demonstrations directly support the conclusion that the method is adaptable to real urban driving conditions. In Fig. 10 (a), the path trend shows that the model can maintain continuous and stable pose estimation even in the presence of multiple path bends and narrow passages, and successfully avoid most impassable areas, ultimately reaching the target position accurately. This indicates that the model performs well in terms of environmental perception accuracy and spatial mapping ability, and can effectively handle actual road conditions with high occlusion and complex structures. In Fig. 10 (b), the obstacle distribution is more dispersed, the road space is more spacious, the trajectory generated by the model is overall smoother, the number of path turns is reduced, and the trajectory is closer to the actual traffic route. To sum up, the model can operate stably in two different environments and can provide reliable navigation and mapping support for the auto-drive system. Table 2 presents the simulation performance analysis of the model.

In Table 2, in terms of positioning accuracy, the model achieves accuracies of 92.4%, 93.8%, and 95.1% in Test 1 to Test 3, with an average of 93.8%, demonstrating high environmental adaptability and stability. The average error gradually decreases from 0.53 m to 0.44 m. In terms of time efficiency, the model takes an average of 0.99 seconds to process each frame of data, which is controlled within the real-time application requirements. The robustness score increased from 8.7 to 9.1, indicating that the model still has stable output ability in the face of environmental disturbances. The mapping related indicators also reflect the excellent performance of the model, with detection coverage rates of 94.3%, 95.1%, and 96.2% in the three tests, indicating that the model can accurately perceive most environmental areas. The smoothness of the trajectory has

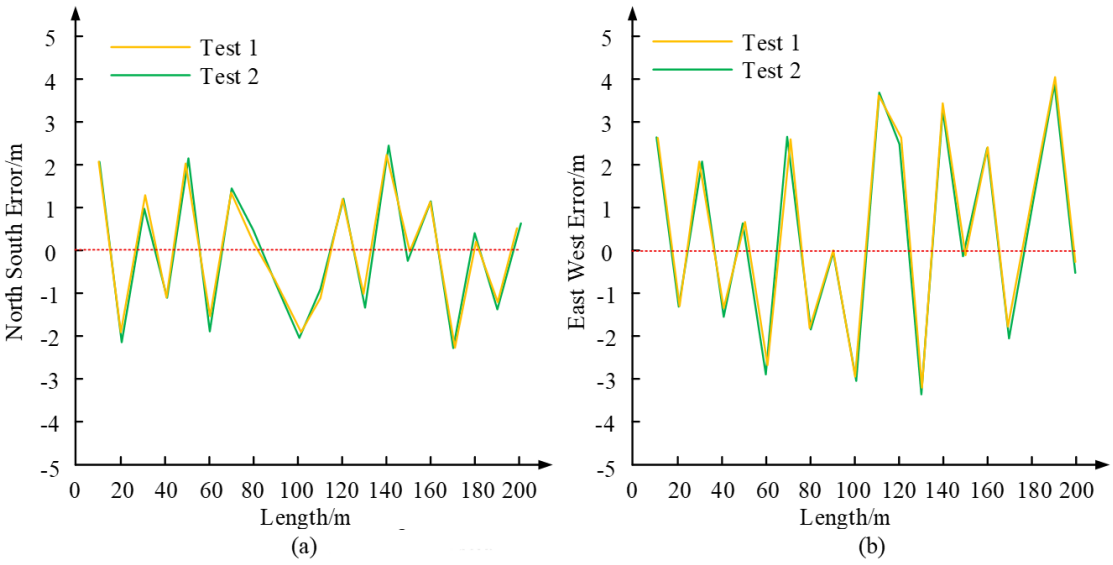

Fig. 9 Model error analysis: (a) North South error; (b) East West error

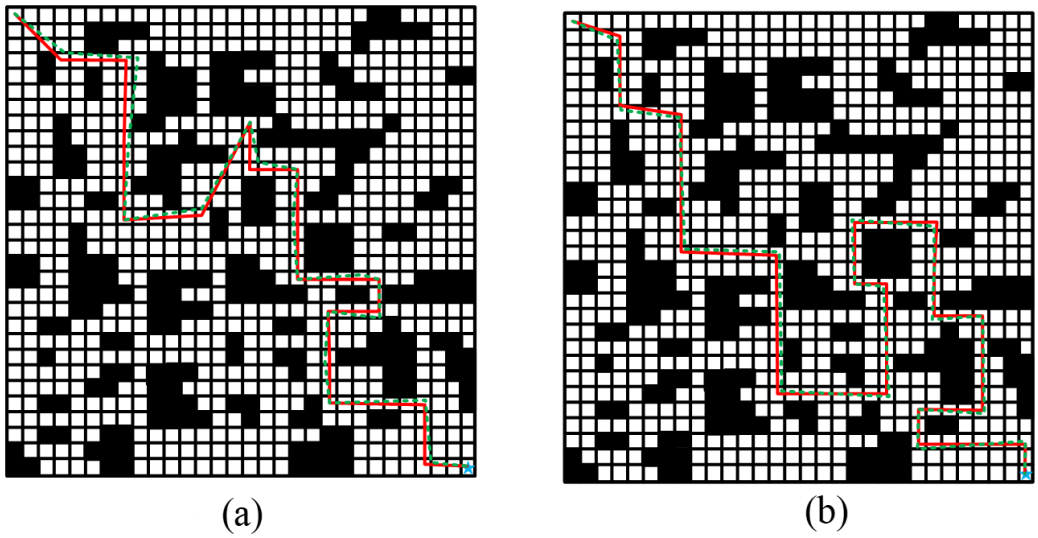

Fig. 10 Synchronous positioning and mapping trajectory analysis: (a) Route 1; (b) Route 2

Table 2 Simulation performance analysis table

Test dimension	Test 1	Test 2	Test 3	Average value
Positioning accuracy (%)	92.4	93.8	95.1	93.8
Average error (m)	0.53	0.47	0.44	0.48
Time efficiency (s)	1.05	0.98	0.94	0.99
Robustness	8.7	8.9	9.1	8.9
Detection coverage (%)	94.3	95.1	96.2	95.2
Track smoothness (m)	0.09	0.07	0.06	0.07
Re-visit detection accuracy (%)	87.5	88.9	90.2	88.9
Cumulative error drift (m/km)	0.41	0.37	0.33	0.37

been optimized from 0.09 m to 0.06 m, and the coherence of the path has been significantly enhanced. The accuracy of follow-up testing remains stable at around 88.9%, indicating that the system has strong loop recognition capabilities, which is beneficial for improving the completeness of mapping. The cumulative error drift decreases from 0.41 m/km to 0.33 m/km, further proving that the model can still effectively suppress error diffusion during long-distance operation. The overall performance is stable and reliable, suitable for automatic driving positioning and mapping tasks in complex environments. Further analysis of the performance of the model is shown in Table 3.

According to Table 3, on the KITTI dataset, the accuracy of traditional methods is only 90%, the false alarm rate is 0.13, the root mean square error is 0.82 m, the robustness score is 6.8, and the number of mutations per km is as high as 6.1, indicating that their trajectory fluctuates significantly and their stability is poor in complex road environments. In contrast, the accuracy of the GO method on KITTI has slightly improved to 91%, the false positive rate has decreased to

Table 3 Robustness analysis including mutation frequency

Model	Dataset	Accuracy (%)	FAR	RMSE (m)	Robustness score	Mutations/km
Traditional	KITTI	90	0.13	0.82	6.8	6.1
GO-based	KITTI	91	0.11	0.71	7.5	4.8
Proposed	KITTI	96	0.07	0.53	9.2	2.3
Proposed	NuScenes	94.7	0.09	0.59	8.9	2.6

0.11, the RMSE has decreased to 0.71 m, the robustness score has increased to 7.5, and the number of mutations has decreased to 4.8/km. This indicates that the introduction of graph optimization mechanism has improved trajectory continuity, but there is still a certain degree of mutation phenomenon. Further observation of the proposed model shows an accuracy rate of 96% on KITTI, far higher than the previous two. The false alarm rate is controlled at 0.07, the RMSE is reduced to 0.53 m, and the robustness score is as high as 9.2. At the same time, the number of mutations is significantly reduced to 2.3/km, indicating that the model can effectively suppress trajectory mutations and maintain higher stability and continuity during long-distance driving. In the complex urban traffic scenario of NuScenes, the proposed model has an accuracy rate of 94.7%, a false alarm rate of 0.09, an RMSE of 0.59 m, a robustness score of 8.9, and a mutation frequency of 2.6/km. Although slightly lower than its performance on KITTI, it still maintains a good level, indicating that the model can maintain strong robustness and anti-interference ability in urban environments with multiple traffic elements and frequent dynamic disturbances.

According to Table 4, on the KITTI dataset, MSCKF and VINS Fusion, as commonly used advanced algorithms, have accuracies of 92.1% and 93.5%, root mean square errors of 0.68 m and 0.61 m, and robustness scores of 7.8 and 8.3, respectively. However, the proposed model achieves an accuracy of 96% in the same environment, with an RMSE of only 0.53 m and a robustness score of 9.2. At the same time, the false positive rate is also controlled at 0.07, which is significantly better than the compared algorithms, reflecting the progress of this method in accuracy and stability. In the complex urban environment of NuScenes, the proposed model achieved an accuracy of 94.7% under clear conditions, an RMSE of 0.59 m, and a robustness score of 8.9, indicating that it still maintains high performance in multi-target and multi-interference scenarios. In rainy weather conditions, the accuracy decreased to 91.8%, the RMSE increased to 0.72 ms, and the robustness score was 8.5, indicating that severe weather has a certain impact on laser and visual sensors, but the overall performance is still acceptable; In nighttime scenarios, the accuracy is 92.6%, the RMSE is 0.69 m, and the robustness score is 8.6, which is better than rainy conditions, indicating that the model still has good stability in low light environments. Overall, the proposed model not only outperforms MSCKF and VINS Fusion on standard datasets but also maintains high robustness and accuracy in extreme weather and nighttime scenarios, demonstrating its practical value in complex environments.

Table 4 Comparative analysis with state-of-the-art algorithms and adverse scenarios

Model	Dataset/ Scenario	Accuracy (%)	FAR	RMSE (m)	Robustness score
MSCKF	KITTI	92.1	0.12	0.68	7.8
VINS-Fusion	KITTI	93.5	0.10	0.61	8.3
Proposed	KITTI	96.0	0.07	0.53	9.2
Proposed	NuScenes (clear)	94.7	0.09	0.59	8.9
Proposed	NuScenes (rain)	91.8	0.11	0.72	8.5
Proposed	NuScenes (night)	92.6	0.10	0.69	8.6

5 Conclusions

In response to the limited positioning accuracy, sensor error accumulation, and weak mapping consistency of autonomous vehicles in complex road environments, this study constructed a synchronous positioning and mapping model based on multi-sensor fusion. This model combined LiDAR, vision, inertial measurement, and wheel speed information to achieve high robustness, high-precision state estimation, and dynamic mapping capabilities. In the experiment, the overall accuracy of the model on the KITTI dataset reached 96%, significantly better than the traditional method's 90% and GO's 91%. FAR was controlled at 0.07, which was 6% and 4% lower than other methods, respectively. The initial heading deviation was compressed from the traditional method's 2.0 degrees and GO's 1.7 degrees to 0.8 degrees, significantly reducing the impact of initial attitude error on the path. In terms of RMSE, the research model had a value of 0.53 m, which was significantly better than GO's 0.71 m and traditional methods' 0.82 m, and the positioning process was more continuous and accurate. At the same time, the average

offset error was also controlled within 0.41 m, and the system robustness score was as high as 9.2, indicating that stable performance can still be maintained in high dynamic environments. In further simulation analysis, the average positioning accuracy of the model in three typical scenarios was 93.8%, the detection coverage was 95.2%, the trajectory smoothness remained within 0.07 m, the cumulative error drift was as low as 0.37 m/km, and the accuracy of follow-up detection remained stable at 88.9%. Research has shown that the model performs well in multiple key performance indicators and is suitable for autonomous driving tasks that require high accuracy, stability, and adaptability. Despite the issue of slightly higher computational complexity, it is expected to achieve wider engineering applications in the future through further model lightweighting and efficient optimization strategies. Future research can further expand the generalization ability of the model in dynamic obstacle-dense, severe weather, and nighttime scenarios, and improve its processing efficiency in large-scale map construction at the city level. Future work will further integrate mm-wave radar to enhance perception robustness in adverse weather conditions and extend the model's applicability to all-weather autonomous driving. In addition, optimizing the timestamp alignment algorithm will continue to be a focus, as it is essential to minimize cross-sensor temporal drift and ensure reliable spatiotemporal fusion in dynamic environments.

References

Bai, F., Wu, K., Bartoli, A. (2024) "Kernel-GPA: A globally optimal solution to deformable SLAM in closed-form", International Journal of Robotics Research, 43(4), pp. 456–484. https://doi.org/10.1177/02783649231195380

Chen, K., Liu, J., Tong, Q., Cheng, X., Zhang, J. (2024) "Accurate Perception and Association of Objects for Humanoid Robots Under Dynamic Visual SLAM", International Journal of Humanoid Robotics, 21(3), 2350019. https://doi.org/10.1142/S0219843623500196

Dor, M., Driver, T., Getzandanner, K., Tsiotras P. (2024) "AstroSLAM: Autonomous monocular navigation in the vicinity of a celestial small body - Theory and experiments", International Journal of Robotics Research, 43(11), pp. 1770–1808. https://doi.org/10.1177/02783649241234367

Du, X., Zhang, C., Gao, K., Liu, J., Yu, X., Wang, S. (2024) "YPL-SLAM: A Simultaneous Localization and Mapping Algorithm for Point-line Fusion in Dynamic Environments", Sensors, 24(14), pp. 4517. https://doi.org/10.3390/s24144517

Gheisari, M., Hamidpour, H. ., Liu, Y. ., Saedi, P., Raza, A., Jalili, A., Rokhsati, H., Amin, R. (2022) "Data Mining Techniques for Web Mining: A Survey", Artificial Intelligence and Applications, 1(1), pp. 3–10. https://doi.org/10.47852/bonviewAIA2202290

Guo, S., Gallego, G. (2024) "CMax-SLAM: Event-Based Rotational-Motion Bundle Adjustment and SLAM System using Contrast Maximization", IEEE Transactions on Robotics, 40, pp. 2442–2461. https://doi.org/10.1109/TRO.2024.3378443

Islam, Q. U., Ibrahim, H., Chin, P. K., Lim, K., Abdullah, M. Z. (2024) "MVS-SLAM: Enhanced multiview geometry for improved semantic RGBD SLAM in dynamic environment", Journal of Field Robotics, 41(1), pp. 109–130. https://doi.org/10.1002/rob.22248

Liu, S., Sun, E., Dong, X. (2024) "SLAMB&MAI: a comprehensive methodology for SLAM benchmark and map accuracy improvement", Robotica, 42(4), pp. 1039–1054. https://doi.org/10.1017/S0263574724000079

Pan, G., Cao, S., Lv, S., Yi, Y. (2025) "DEG-SLAM: a dynamic visual RGB-D SLAM based on object detection and geometric constraints for degenerate motion", Measurement Science and Technology, 36(2), 026302. https://doi.org/10.1088/1361-6501/ada39c

Su, W., Huang, T., Liu, F., Wang, H. (2024) "ASLAM-FD: a multi-sensor adaptive collaborative fusion SLAM framework based on degradation detection and deep reinforcement learning", Measurement Science and Technology, 35(12), 126312. https://doi.org/10.1088/1361-6501/ad7bdd

Trzeciak, M., Pluta, K., Fathy, Y., Alcalde, L., Chee, S., Bromley, A., Brilakis, I., Alliez, P. (2023) "ConSLAM: Construction Data Set for SLAM", Journal of Computing in Civil Engineering, 37(3), pp. 13–16. https://doi.org/10.1061/JCCEE5.CPENG-5212

Wang, W., Wang, Y., Wu, Z. (2024) "Optimized feature extraction and object detection for indoor dynamic environment visual SLAM study", International Journal of Advanced Robotic Systems, 21(5), pp. 1–5. https://doi.org/10.1177/17298806241279610

Wang, Y., Zhao, L., Gong, L., Chen, X., Zou, S. (2023) "A monocular SLAM system based on SIFT features for gastroscope tracking", Medical & Biological Engineering & Computing, 61(2), pp. 511–523. https://doi.org/10.1007/s11517-022-02739-1

Wenzel, P., Yang, N., Wang, R., Zeller, N., Cremers, D. (2025) "4Seasons: Benchmarking Visual SLAM and Long-Term Localization for Autonomous Driving in Challenging Conditions", International Journal of Computer Vision, 133(4), pp. 1564–1586. https://doi.org/10.1007/s11263-024-02230-4

Xu, M., Wang, Y., Xu, B., Zhang, J., Ren, J., Huang, Z., Poslad, S., Xu, P. (2024) "A critical analysis of image-based camera pose estimation techniques", Neurocomputing, 570, 127125 https://doi.org/10.1016/j.neucom.2023.127125

Yin, H., Sun, M., Zhang, L., Rigoll, G. (2025) "Online dynamic object removal for LiDAR-inertial SLAM via region-wise pseudo-occupancy and two-stage scan-to-map optimization", Displays, 88, 103030. https://doi.org/10.1016/j.displa.2025.103030

Zhao, Y., Wang, C., Ouyang, Y., Zhong, J., Lis, Y., Zhao, N. (2025) "DHDP-SLAM: Dynamic Hierarchical Dirichlet Process based data association for semantic SLAM", Displays, 86, 102892. https://doi.org/10.1016/j.displa.2024.102892