Journal of Electronic Research and Application, 2026, Volume 10, Issue 9
https://bbwpublisher.com/journal/JERA
Online ISSN: 2208-3510
Print ISSN: 2208-3502

Research on Production Method of High-Definition Map for Autonomous Driving Sweeping Robots

Jian Liu

Artificial Intelligence Research Institute, Changsha Branch, Fulongma Group Co., Ltd., Changsha 421300, Hunan, China

Copyright: © 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

Abstract: Sweeping robots require centimeter-level high-definition maps to realize edge-hugging cleaning and narrow-road passage. Passenger-vehicle-oriented solutions suffer from point-cloud smearing caused by unsynchronized sensor timestamps, lack of sweeping-related semantics such as curbs and trash cans, and map deformation induced by odometer drift during long-hour operation. This paper redesigns a full-link solution: GPS-Pulse plus PTP for sub-millisecond time synchronization; improved BevFusion combined with bilateral filtering for depth completion in semantic processing; LOAM plus FPFH loop-closure detection and graph optimization for pose estimation; the editing module supports point-cloud import, curb delineation, cleaning-path generation and topology network construction. A verification test covering approximately 5 km² in Changsha achieves synchronization error < 1 ms, semantic segmentation accuracy > 95 cm, positioning error < 5 cm, coverage rate 98 %, and an efficiency improvement of around 30 %.

Keywords: Autonomous driving; Sweeping robot; High-definition map; Multi-sensor fusion; Pose optimization

Online publication: October 30, 2026

1. Introduction

Over the past five years, the annual expansion rate of road-cleaning areas has reached 6–8 %, while the average age of sanitation workers exceeds 55 years, resulting in difficulties in recruitment and staff retention [1]. Autonomous-driving sweeping robots are regarded as a solution to ease labor shortages. Different from passenger vehicles, sweeping robots need to travel close to edges and pass through narrow lanes, and their map requirements go far beyond lane-level navigation [2]. Porting passenger-vehicle solutions encounters three obstacles. Equipped only with 16-line LiDAR and monocular cameras for cost control, such systems suffer low clock-synchronization accuracy, leading to misalignment and point-cloud smearing. Passenger-vehicle maps only contain lane lines and traffic lights, whereas sweeping tasks additionally require

semantics for curbs, trash cans, speed bumps and other objects. Multi-hour continuous operation gives rise to odometer drift, producing decimeter-level deviations between the start and end of trajectories. This paper takes self-developed sweeping robots of Fulongma as the platform and conducts multi-scenario verifications in Changsha.

2. Methodology and principles

2.1. Real-time multi-sensor data synchronization technology

The sweeping robot carries a 16-line LiDAR, a forward-looking camera, a nine-axis IMU, and a GPS/RTK module, with respective sampling frequencies of 10 Hz, 30 fps, 100 Hz, and 1 Hz, each driven by an independent crystal oscillator. Without precise synchronization, time offsets exist among data from different sensors, and direct fusion produces point-cloud smearing manifested as blurred or even double-shadows on building facades.

2.1.1. Hardware clock anchoring

The GPS module outputs a PPS pulse every second, whose rising edge is aligned to UTC second with an accuracy of ± 30 ns. The PPS signal is fed into the GPIO interrupt of the main-control board. The rising-edge signal latches the local clock counter, while the GPRMC sentence is parsed to obtain UTC time, and their difference yields the clock offset. A 10-epoch sliding-window filter is adopted to eliminate gross errors, and the median value is used for correction. Error sources include nominal 30 ns GPS error, approximately 0.2 ms serial-port delay, and roughly 0.05 ms crystal-oscillator stability error. The theoretical synchronization accuracy is better than 0.3 ms.

$$\sigma_{sync} = \sqrt{\sigma_{PPS}^2 + \sigma_{GPRMC}^2 + \sigma_{local}^2}$$

2.1.2. Fine calibration of software timestamps

After hardware anchoring, fine-alignment is still required for points within each LiDAR frame, camera exposure instants, and IMU samples. The IEEE 1588 PTP protocol is introduced [3]. The main-control board acts as Grandmaster and sensors serve as Slaves. Link delay and clock offset are calculated to achieve sub-microsecond accuracy. During data acquisition, trigger pulses (100 ms, 10 Hz) generated from the synchronized clock trigger both LiDAR and cameras, yielding an alignment error below 1 μs. The approximately 8 ms rolling-shutter delay of cameras is subtracted in the calibration phase.

2.2. Semantic-information extraction and depth-information fusion

Sweeping robots have distinctly different semantic requirements compared with passenger vehicles [4]. Passenger vehicles focus on lane lines and traffic signs; sweeping robots additionally need to identify curbs (determining edge-hugging distance), trash cans (determining detour strategies), speed bumps (determining brush lifting/lowering), and manhole covers (determining avoidance needs). General-purpose models require targeted adjustments in both network architecture and training datasets.

2.2.1. Data preprocessing

Raw point clouds contain noise from dust, raindrops, and multipath reflection. Statistical filtering is applied

first: for each point, compute its average distance to its 30 nearest-neighbor points; points exceeding the global mean plus twice the standard deviation are removed. Voxel down-sampling is then performed, replacing original points with centroids. Gaussian-kernel denoising is applied to camera images, and histogram equalization is used to handle backlight interference. Camera intrinsic parameters are obtained via Zhang Zhengyou's chessboard-grid method, and extrinsic parameters are acquired through hand-eye calibration, with reprojection error below 1.5 pixels.

2.2.2. Semantic-segmentation network

BevFusion is selected as the base architecture to support BEV-based fusion of point clouds and images [5]. Improvements include upgrading the backbone network from ResNet-50 to ResNet-101 and adding ASPP modules (dilation rates 6, 12, 18) in the decoder. The training dataset consists of Cityscapes (approximately 5 000 frames) and self-collected sweeping-scene data (approximately 2 300 frames with four newly-added categories: curbs, trash cans, speed bumps and manhole covers). Class-weighted cross-entropy loss, Adam optimizer, and cosine annealing for 120 epochs are adopted. Morphological closing operation fills holes in segmentation outputs, and connected-component analysis removes fragments smaller than 50 pixels.

2.2.3. Depth-information extraction and enhancement

Although LiDAR point clouds deliver accurate depth values, 16-line scanners produce sparse point-cloud distributions at long ranges, resulting in numerous voids in depth maps. First, calculate the distance from each point to the robot's center according to three-dimensional point-cloud coordinates, and project points onto a two-dimensional polar-coordinate plane to generate a depth map. Bilateral-filter interpolation fills void regions. This filter suppresses noise while preserving depth-discontinuity edges such as curbs and walls.

2.2.4. Point-cloud semantic labeling and multimodal fusion

Project each point in the point cloud onto the camera-image plane using the extrinsic-parameter matrix to read and assign semantic labels [6]. For occluded or out-of-boundary points, fall back to rule-based classification using geometric features. For finer-grained scenes, a lightweight multimodal-fusion network is designed. It extracts point-cloud features via PointNet++ and image features via ResNet, and performs adaptive fusion with channel attention in the BEV space. Compared with pure projection-based methods, edge-semantic accuracy increases by approximately four percentage points.

2.3. Global pose optimization

Sweeping robots travel trajectories of more than ten kilometers in a single operation. Pure frame-by-frame pose accumulation from odometry produces drift increasing at several centimeters per kilometer, leading to misalignment between the start and end of maps [7]. This paper constructs a pose-optimization framework consisting of front-end odometry, back-end graph optimization, and loop-closure detection [8].

2.3.1. Front-end pose estimation

LOAM decomposes point clouds into edge and planar features and computes inter-frame transformations via point-to-line and point-to-plane matching against the previous frame. To improve robustness under sharp turns and uphill-downhill conditions, an EKF fuses IMU measurements as predictions and LOAM outputs as observations. The 15-dimensional state vector contains position, attitude, velocity, and biases. Fusion

significantly reduces pose jitter.

2.3.2. Point-Cloud Registration and Local Mapping

The NDT algorithm performs fine registration of adjacent-frame point clouds. It voxelizes reference point clouds and describes local normal distributions with mean μ and covariance Σ. It seeks a transformation T that maximizes the probability of source-point-cloud samples falling into reference distributions.

Registered point clouds are fused into the global map through voxel grids.

$$J(T)=\sum_{i=1}^{n}N(p_i \mid \mu_{T(p_i)},\Sigma_{T(p_i)})$$

2.3.3. Loop-closure detection

FPFH descriptors are extracted and matched against historical frames within a 500-frame sliding window, selecting the top-5 candidates. RANSAC eliminates false matches. Loop closure is confirmed when reprojection error < 0.3 m, followed by a 60-second cooldown period. All frame poses serve as nodes, while odometry and loop-closure constraints serve as edges for graph optimization. The sum-of-squared residuals is minimized via the LM solver in the g2o library, with convergence achieved within 5-8 iterations [8]. In field tests, terminal drift over a 4-km trajectory is reduced from 18 cm to 4 cm.

2.4. High-definition-map editing workflow

Global point-cloud maps cannot be directly consumed by planning modules, which require structured semantic maps containing road boundaries, cleaning paths, and intersection connections [9]. Accordingly, an editing workflow converting raw point clouds into semantic maps is implemented in self-developed software.

2.4.1. Data import and visualization

Optimized global point clouds and semantic-label files [10] are imported into the editing tool. Rendered by the OpenGL engine, the tool supports rotation, zooming, and panning. Data are colored by semantic category and managed in independently toggle-able layers.

2.4.2. Road-boundary annotation

Road contours are automatically extracted from semantic labels and further refined with the Canny operator. Due to point-cloud noise and semantic misclassification, automatic boundaries often contain saw-teeth and gaps, which are fitted segment-wise with cubic Bézier curves and subject to manual fine-tuning by operators. Polygons are automatically inferred for intersections; roundabouts and irregular junctions support manual drawing.

2.4.3. Generation of cleaning guidance lines

Cleaning-guidance lines represent the actual sweeping trajectories: they maintain a 10–15 cm offset from curbs; turning radii are ≥ 1.2 m; spacing between adjacent travel paths equals the effective width of sweeping brushes [11]. Equally-offset lines are first generated from road boundaries and chained into continuous paths according to intersection connectivity, then reviewed and corrected by operators.

2.4.4. Road-object annotation

Road facilities are identified from point-cloud semantic labels and geometric shapes. Trash cans appear as isolated clusters approximately 0.8–1.2 m high with near-circular cross-sections; speed bumps are long raised strips crossing road surfaces with heights of 2–5 cm. Automatically-detected objects are manually verified one-by-one before database storage. Each record includes object category, center coordinates, orientation, and dimensions.

2.4.5. Topology construction and map export

Connectivity among road segments and associations between road segments and road objects are established to form a topological network [12]. After editing, maps are exported in OpenDRIVE format together with metadata files.

3. Experiments and result analysis

3.1. Experimental platform and scenarios

Experiments are carried out in Yuelu District, Changsha, using the self-developed Fulongma sweeping robot equipped with a 16-line LiDAR, binocular camera, and RTK receiver. Test scenarios cover arterial roads, secondary roads, sidewalks (2-3 m wide), and public squares, with a total area of approximately 5 km^2.

3.2. Synchronization-accuracy verification

A high-precision timestamp LED screen is placed along the data-collection route. LiDAR and cameras capture images simultaneously, and time offsets are obtained by comparing timestamps displayed on the LED screen. Among 200 consecutive frames, the maximum time deviation reaches 0.7 ms and RMS error equals 0.3 ms, satisfying the < 1 ms specification.

3.3. Semantic-segmentation accuracy

Three-hundred frames are randomly sampled from self-collected data as the test set. Two annotators independently label frames, and their intersection serves as ground truth. Pixel-wise accuracy reaches 95.3 %. Large-area categories such as road curbs achieve accuracy above 97 %, while small objects including trash cans and fire hydrants achieve approximately 91 %. Compared with the baseline model without ASPP modules, small-object IoU increases by 3.2 percentage points.

3.4. Map accuracy

The RTK check-point method is adopted: forty-two evenly-distributed check-points are laid out across the test area with RTK horizontal accuracy of ± 2 cm. Corresponding coordinates are measured on the finished map. Results show the median horizontal deviation is 3.1 cm with a maximum value of 4.8 cm (both < 5 cm); median elevation deviation is 2.4 cm. Residual error drops from 14.6 cm to 3.2 cm after loop-closure optimization.

3.5. Sweeping-operation performance

Five repeated test runs are performed under two modes: “with high-definition map” and “manual teaching only”. (**Table 1**)

Table 1. Repeated test run

Indicator	Manual teaching	High-definition map	Improvement
Cleaning coverage rate	89.2 %	98.1 %	+ 9 %
Single-run operation time	47 min	36 min	-23.4 %
Edge-hugging-cleaning compliance rate (distance to curb < 15 cm)	72.5 %	96.3 %	+ 23.8 %
Number of manual interventions (per run)	3.6	0.4	-88.9 %

Results: coverage rises from 89.2 % to 98.1 %; single-run time decreases from 47 min to 36 min; edge-hugging compliance rate increases from 72.5 % to 96.3 %; manual-intervention frequency drops from 3.6 to 0.4 times per run. High-definition maps deliver the most prominent improvement for edge-hugging cleaning.

3.6. Discussion

The proposed method still has limitations. First, semantic-segmentation accuracy degrades significantly under heavy occlusion (> 70 %), and multi-frame temporal fusion may be introduced. Second, approximately 30 % of map-editing workload still requires manual intervention. Third, data collection is performed under low-dynamic conditions in early-morning hours; semantic segmentation and point-cloud registration become considerably more challenging in highly-dynamic scenarios.

4. Conclusion and future work

4.1. Conclusion

Targeting practical operation requirements of sweeping robots, this paper designs and implements a high-definition-map production pipeline covering data acquisition, semantic fusion, pose optimization, and map editing. Time synchronization is realized with low-cost hardware; semantic categories are extended to include curbs and trash cans; pose optimization mitigates long-term drift; cleaning-guidance lines are added to the map-editing workflow. Field tests yield synchronization error < 1 ms, semantic-segmentation accuracy > 95 %, positioning error < 5 cm. All key indicators meet design expectations.

4.2. Future work

Future-stage research will focus on three directions: first, incremental map updating for dynamic scenarios [13]; second, multi-robot collaborative mapping [14]; third, introducing large-language-models into map editing to explore interactive-editing paradigms [15].

Disclosure statement

The author declares no conflict of interest.

References

[1] Ministry of Housing and Urban-Rural Development, 2021, China Urban-Rural Construction Statistical Yearbook (2020). Beijing: China Statistics Press.

[2] Duan X, Zhou Y, Tian D, et al., 2021, Review of Deep-Learning Applications in Autonomous Driving. Unmanned Systems Technology, 4(6): 27.

[3] Xiu X, Wang K, Wu J, 2025, Design and Implementation of Optimized PTP Synchronization Based on MAC. Journal of Time and Frequency, 48(2): 101–108.

[4] Du M, 2022, Research on Image Semantic Segmentation of Road Scenes Based on Deep Learning, thesis, Henan Polytechnic University.

[5] Ruder S, 2016, An Overview of Gradient Descent Optimization Algorithms. arXiv.

[6] Zhang J, Huang Y, Zhang R, 2021, CNN-Based Object Detection via Point-Cloud-Image Fusion. Opto-Electronic Engineering.

[7] Yang W, 2026, Research on Multi-Sensor-Fusion-Based SLAM Algorithm for Mobile Robots, thesis, University of Electronic Science and Technology of China.

[8] Wang M, Wei G, Cai J, et al., 2024, Robust Pose-Graph-Optimization Algorithm for 3D SLAM Under Heavy Noise. Control Engineering of China, 31(7): 1305–1313.

[9] Kong T, Gao H, Chen X, 2024, Full-Coverage Path Planning for Sweeping Robots Based on Improved RRT. Computer Engineering and Applications, 60(13): 311–318.

[10] Xia X, Wang D, Cao J, et al., 2022, LiDAR Point-Cloud Semantic-Segmentation Method for Vehicles Based on Sparse Convolutional Neural Networks. Automotive Engineering, 44(1): 10.

[11] Li B, Guo Y, Zhou J, et al., 2024, Development and Prospect of High-Definition Maps for Intelligent Driving. Geomatics and Information Science of Wuhan University, 49(4): 491–505.

[12] Xu J, Hou F, Cao G, 2022, Production Methods and Key Technologies for High-Definition Road Maps. Bulletin of Surveying and Mapping, 2022(1): 129–133.

[13] Yang M, Jiang K, Wen T, et al., 2023, Status and Challenges of Crowdsourced Updating Technology for Autonomous-Driving High-Definition Maps. China Journal of Highway and Transport, 36(5): 244–259.

[14] Chen L, Liu K, Zhou B, et al., 2021, Research on Key Technologies for Multi-Agent Collaborative High-Definition-Map Construction. Acta Geodaetica et Cartographica Sinica, 50(11): 10.

[15] Shomer H, Xu J, 2025, Automated Label Placement on Maps via Large Language Models, arXiv preprint arXiv:2507.22952.

Publisher's note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.