



OPEN

DATA DESCRIPTOR

Multimodal sensor dataset from vehicle-mounted mobile mapping system for comprehensive urban scenes

Shengyu Lu^{1,2}, Sheng Bao^{1,2}, Wenzhong Shi¹✉, Yitao Wei¹, Shuyu Zhang¹ & Daping Yang¹

Mobile mapping is the research trend in the mapping field due to its superior time efficiency compared to traditional fixed mapping methods. It is an important digital base for numerous applications, such as high-definition (HD) maps, digital twins, smart cities. However, most mobile mapping datasets are based on portable platforms, such as backpacks and robotics, leading to insufficient research on large-scale mobile mapping and autonomous driving. To change the status quo, a multimodal sensor dataset from a vehicle-mounted mobile mapping system for comprehensive urban scenes (MSD-VMMS-HK) is provided. It has rich, high-precision, and large-scale multimodal sensor information, including high-precision (millimeter-level) light detection and ranging (LiDAR), the panoramic camera, and GNSS/INS. The MSD-VMMS-HK dataset features a wide variety of scenarios in Hong Kong, which is a representative urban area with diverse and comprehensive challenging urban scenes like mountain tunnels, cross-harbour tunnels, urban canyons, mountain and seaside roads. It is the first urban-level comprehensive urban scenes dataset that provides high-precision references for the validation of point clouds and image processing. Additionally, examples of various applications of the dataset, such as accurate mapping of urban canyons, urban infrastructure management and maintenance, and change detection, are provided to facilitate reference by the academic community.

Background & Summary

Mobile mapping technology¹ integrates measurement techniques with mobile platforms to acquire and analyze geographic information and environmental data. By equipping sensors and measurement devices on mobile platforms, this technology has widespread applications in various fields, including navigation systems, map creation^{2,3}, and urban planning. The advancement of mobile mapping technology has significantly improved the accuracy, real-time capabilities, and comprehensiveness of geographic information and environmental data, encompassing terrain, landforms, buildings, and roads. Notably, the dataset generated by mobile mapping systems plays a pivotal role in advancing research in the field. Over the past decade, many known datasets such as Ford Campus dataset⁴, ISPRS MIMAP⁵, UrbanNav dataset⁶, and PolyU-BPCoMa⁷ have been released to provide a standard algorithm benchmark and keep benefiting the mobile mapping community.

Mobile mapping systems (MMS) encompass a range of forms, including handheld devices, backpack devices, boats, vehicles, and robots (unmanned aerial vehicles (UAVs), unmanned ground vehicles (UGVs), and Legged robots⁸). Among various types of mapping systems, the MMS stand out with its high accuracy 3D point cloud data and high-dynamic-range images, as well as its wide range of application prospects. However, most mobile mapping datasets based on portable platforms, such as backpacks⁹ and robotics¹⁰, are limited in terms of the quantity of sensor data, the scale of measurement scenarios, and outdated in content. The vehicle-mounted mobile mapping systems provide a good solution to the inability to collect high-precision data for large scenes, by integrating cutting-edge technologies such as high-precision LiDARs, high dynamic range (HDR) panoramic cameras, Global Navigation Satellite System (GNSS), and inertial navigation systems (INS), however, the high

¹The Otto Poon C. F. Smart Cities Research Institute, Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong SAR, 999077, China. ²These authors contributed equally: Shengyu Lu, Sheng Bao. ✉e-mail: john.wz.shi@polyu.edu.hk

cost of the system limits its accessibility and results in a scarcity of high-quality vehicle-mounted mobile mapping datasets.

Regarding the limitations of existing vehicle-mounted mobile mapping datasets for research and applications, there are numerous challenges to be addressed. In available datasets, there is a notable shortage of single-line LiDAR mapping datasets for indoor environments or poor GNSS signals environments, such as parking lots, mountain tunnels, cross-harbour tunnels, urban canyons¹¹ and mountain roads. The dataset for indoor parking lots holds significant importance for applications such as intelligent parking systems and indoor navigation¹². Nevertheless, due to the enclosed GNSS-denied environments and degraded scenarios of mountain tunnels and cross-harbour tunnels, data collection and mapping become increasingly challenging¹³. One of the main challenges involves maintaining positioning accuracy and sensor reliability in such closed environments. As a result, acquiring and ensuring the accuracy of vehicle-mounted mobile mapping datasets for indoor environments still presents challenges.

Recognizing these gaps and challenges, the MSD-VMMS-HK database is motivated to be provided, to address this lack of information and provide valuable insights into the field of vehicle-mounted mobile mapping. A comprehensive vehicle-mounted mobile mapping system dataset is introduced in this paper, which encompasses a diverse range of scenes, including cross-harbour tunnels, mountain tunnels, campus roads, urban expressways, urban canyons, highways, and indoor parking lots. This dataset holds immense value for both research and practical applications. The vehicle-mounted platform is equipped with cutting-edge sensors, including a 128-channel Velodyne VLS-128, Riegl VUX-1HA, Ladybug5+, and NovAtel SPAN CPT7. Notably, all these sensors have undergone time synchronization, ensuring superior sensor synchronization performance compared to datasets without such synchronization.

Most vehicle-mounted mobile mapping datasets rely on GNSS or Inertial Navigation System (INS) fusion to provide ground truth. While GNSS ensures global consistency within the absolute world coordinate system, the resultant real-world trajectory lacks local accuracy due to the factors like satellite signal loss or reacquisition. This is particularly evident in non-exposed spaces such as indoor environments, tunnels, and densely built urban canyons¹⁴. For instance, the KITTI¹⁵ dataset's odometry benchmark excludes evaluations at length scales below 100 meters to avoid this issue. To address these challenging scenarios, a GNSS + INS + Simultaneous Localization and Mapping (SLAM) approach has been implemented, which optimizes the GNSS odometry and the LiDAR odometry by factor graph to obtain the optimized odometry, and the experiments prove that this method significantly improves the positioning accuracy in the non-exposed space.

The primary contributions of this paper are as follows:

1. A comprehensive urban multimodal mobile mapping dataset MSD-VMMS-HK is provided. This dataset encompasses a wide range of scenarios, especially non-exposed spaces and GNSS-challenged spaces such as mountain tunnels, cross-harbour tunnels, urban canyons, parking lots, mountain and seaside roads. It provides researchers with valuable data for various applications in these fields.
2. Full-element perception sensor equipments. This dataset comprises 2 TB of cutting-edge sensor data, including 128-channel LiDAR, single-line LiDAR (millimeter-level accuracy), panoramic imagery, GNSS/inertial navigation system (INS) data. All sensors have undergone time synchronization.
3. A benchmark is provided to evaluate the mapping accuracy for a multimodal vehicle-mounted mobile mapping system. Besides, a variety of sample applications are introduced to inspire how to use this dataset.

In conclusion, we provide the academic community with a multimodal, high-precision, multi-scenario, application-rich in-vehicle-mounted mobile mapping dataset to contribute to the research in the field of mobile mapping, which also near-field applications such as SLAM, localization, autonomous driving, and point cloud processing.

Methods

In recent years, the development of mobile mapping technology and the rise of autonomous driving research have led to the emergence of a multitude of vehicle-mounted mobile mapping datasets¹⁶. Previously, localization and mapping datasets were primarily focused on environments involving robots¹⁷ or small-scale vehicles¹⁸, with particular emphasis on indoor or controlled outdoor scenes. However, the increasing progress in autonomous driving and intelligent transportation systems has created a growing demand for vehicle-level localization and mapping datasets. Consequently, the latest datasets have shifted their focus towards encompassing a wider range of road and traffic scenarios, including urban roads, highways, and complex urban environments. These datasets offer more challenging scenes, enabling researchers to assess and enhance the performance of vehicle localization and mapping algorithms.

Furthermore, modern datasets are not limited to localization and mapping, but provide valuable support for a variety of downstream tasks. While traditional datasets were primarily used for evaluating localization and mapping algorithms, the datasets now offer extensive annotation information. This includes object detection¹⁹ and tracking, semantic segmentation, road segmentation²⁰. Such comprehensive annotations allow researchers to explore multiple tasks using a single dataset, fostering cross-domain collaboration and driving innovation.

Here the MSD-VMMS-HK dataset is proposed, a comprehensive multimodal urban sensing resource. In the following sections, an overview of vehicle-mounted mobile mapping datasets is provided, focusing on datasets designed specifically for localization and mapping, as well as those tailored for autonomous driving applications.

The vehicle-mounted data acquisition platform. The data acquisition platform is a mobile mapping vehicle developed by our team, which is shown in Fig. 1. The platform integrates a 128-channel LiDAR Velodyne VLS-128 with accuracy of 3 cm. A single-line LiDAR, namely the Riegl VUX-1HA, whose accuracy is up to 3 mm.

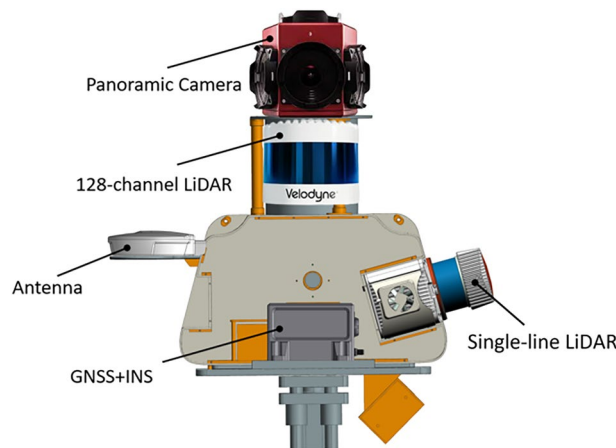


Fig. 1 Sensor positioning of the data acquisition platform. The sensor configuration combines LiDAR, panoramic cameras, and GNSS/INS for comprehensive data collection.

It is also equipped with a Ladybug5+, a 6-lens panoramic camera with outstanding image quality. The NovAtel SPAN CPT7, a high-performance GNSS positioning and Inertial INS system with accuracy of centimeter-level and output frequency up to 100 Hz, is mounted on the platform. The extrinsic parameters between the GNSS/INS and LiDAR are calibrated using the method²¹.

Sensors. The vehicle-mounted mobile mapping system is a sophisticated integrated system comprised of multiple key components and sensors. Below, the basic architecture of the vehicle-mounted mobile mapping system is enumerated and its key sensors and devices are highlighted:

Velodyne VLS-128: The Velodyne VLS-128 is a long-range 3D LiDAR sensor. The detection range is up to 300 m with 3 cm typical range accuracy. Its 128-channel laser beams provide $[-25^\circ, +15^\circ]$ vertical and 360 degree horizontal field of views (FOVs). It can collect 2,300,000 points per second in the single return mode.

Riegl VUX-1HA: The Riegl VUX-1HA High Accuracy kinematic LiDAR sensor is a high speed, non-contact profile measuring system using a narrow laser beam and a fast line scanning mechanism, enabling full 360 degree beam deflection. The Riegl VUX-1HA is a single-line LiDAR, whose accuracy is up to 3 mm. The scan speed of Riegl VUX-1HA is up to 250 and can take up to 1,800,000 measurements per second.

FLIR Ladybug5+: The Ladybug5+ is a 360 degree spherical camera which provide 30 Megapixels ($2464 \times 2048 \times 6$) image resolution with a 3.45 m pixel size and a 12-bit analog-to-digital converter. And the camera operates in the readout manner of global shutter. In this dataset, Ladybug5+ directly outputs stitched panoramic images from the six lenses at $8192 \times 4096 @ 15 \text{ Hz FPS}$.

NovAtel SPAN CPT7: A high-performance GNSS positioning and INS system. The NovAtel SPAN CPT7 integrates a GNSS receiver that receives signals from GPS, GLONASS, Galileo, and other global positioning satellites for precise vehicle positioning information. And the performance parameters of the GNSS receiver include positioning accuracy, positioning velocity, and signal reception sensitivity. The NovAtel SPAN CPT7 also integrates an INS to provide more precise positioning and attitude information. The INS utilizes sensors such as gyroscopes and accelerometers to measure the vehicle's acceleration, angular velocity, and directional changes, enabling estimation of the vehicle's position and attitude. Performance parameters of the INS include attitude accuracy, acceleration measurement accuracy, and angular velocity measurement accuracy. Fusing GNSS and INS data enables the NovAtel SPAN CPT7 to provide centimeter-level accuracy with output frequencies up to 100 Hz.

The Time Synchronization. The time synchronization functionality of sensors during data acquisition is essential to improve the accuracy of data fusion between sensors. It is worth mentioning that in our dataset, the timestamps of all sensors have been synchronized, including the Velodyne VLS-128, Riegl VUX-1HA, and Ladybug5+, with the GPS time of the NovAtel SPAN CPT7.

By utilizing the Pulse Per Second (PPS) time synchronization functionality, nanosecond level precision in the time differences between different sensors' data can be achieved²². However, using the Robot Operating System (ROS) time on the computer as the time reference for the sensors, the time synchronization accuracy between the sensors is about 100 microseconds²³. This level of synchronization would create difficulties in aligning data from sensors operating at different frequencies, such as the Riegl VUX-1HA with a 250 Hz working frequency and the NovAtel SPAN CPT7 with a 100 Hz working frequency. Non-alignment introduces systematic errors that affect the accuracy of sensor data fusion as well as the accuracy of localization and mapping. Therefore, it is crucial to synchronize the timestamps of all sensors. By ensuring accurate time synchronization of sensor data, a higher quality dataset is provided, offering a more reliable foundation for autonomous driving and localization and mapping research.

Calibration. The vehicle-mounted mobile mapping system was implemented with rigorous geometric calibration during installation to ensure measurement consistency and system reliability. As illustrated in Fig. 1, the Ladybug5+, Velodyne VLS-128, and NovAtel SPAN CPT7 sensors are all horizontally positioned, with their centers aligned along the same axis, differing only in height. Specifically, the height difference between Ladybug5+ and NovAtel SPAN CPT7 is 462.0 mm, while the height difference between Velodyne VLS-128 and NovAtel SPAN CPT7 is 277.0 mm. And the Riegl VUX-1HA sensor is inclined and positioned with a horizontal tilt angle of 15°. The vertical and horizontal distances between the Riegl VUX-1HA and NovAtel SPAN CPT7 are 107.8 mm and 270.3 mm, respectively. These parameters can be used to calibrate the sensors after the relative positions of all the sensors have been measured.

After performing initial coarse calibration of the sensors, the further calibration has been conducted using calibration algorithms for the Ladybug5+ and Velodyne VLS-128. Through experimental validation, we confirmed that the calibration accuracy of all sensors meets the required standards. By calibrating the system, the accurate relative positions and orientations between the sensors can be obtained, enabling precise data alignment and fusion. This is crucial for subsequent tasks such as localization and mapping, and environment perception. The accuracy of system calibration ensures data consistency and reliability, providing a solid foundation for the applications.

Data Records

The vehicle-mounted mobile mapping system dataset encompasses a wide range of scenarios, especially non-exposed spaces such as mountain tunnels, cross-harbour tunnels, urban canyons, and parking lots. Prior to data collection, it is vital to plan the collection path. Path planning should take into account the efficient and complete collection of data, the selection of representative scenarios, and the requirements for path length.

For the urban road environments, the prominent Central District in Hong Kong was selected as it is characterized by towering buildings, which can pose challenges due to limited sky view obstructing GNSS signals. This, in turn, affects the mapping accuracy and presents an interesting scenario for analysis. Tunnels, due to their enclosed nature, experience extended periods without GNSS signal reception and often exhibit reduced scene complexity, making mapping using SLAM methods challenging. To address this, the data from both mountain tunnels and cross-harbour tunnels has been collected to create a diverse dataset for researchers to use. Indoor parking lots were also selected as representative areas for mapping. The lack of GNSS signal reception in these indoor environments makes it challenging to rely primarily on LiDAR and IMU data for localization and mapping.

In summary, the data collection paths in this dataset encompass various significant locations, which is shown in Fig. 2, including Central District, Tsim Sha Tsui, Cross-Harbour Tunnel, Eastern Harbour Crossing, Western Harbour Crossing, Ho Man Tin, Fat Kwong Street, How Ming Street, Hong Kong High Speed Way, Stubbs Road, Wan Chai Bypass, Western Kowloon, The Hong Kong Polytechnic University on-campus parking lot, Hong Kong Science Park and Tate's Cairn Tunnel. It is worth mentioning that the data collection paths in Central District encompass multiple temporal acquisitions, including Central District_2, and Central District_3, captured respectively on December 5, 2023; July 25, 2024; and September 19, 2024. The distinct variations in building facades, road infrastructure, vehicular traffic, and pedestrian flow among these three datasets render them highly suitable for change detection analysis. The Table 1 provides specific details on capture length, acquisition time, and data obtained from Panoramas, Velodyne VLS-128, and Riegl VUX-1HA for partial data collection paths.

Data structure. The data types within a dataset are crucial for application domains. The more diverse the data types in a dataset, the higher the potential for cross-domain applications. It's worth noting that the vehicle-mounted mobile mapping system dataset in this paper primarily comprises four types of data: point cloud, IMU data, GNSS data, and panoramic images, shown in Fig. 3.

Point cloud data is further divided into point cloud data from the 128-channel Velodyne VLS-128 LiDAR and point cloud data from the single-line Riegl VUX-1HA LiDAR. The point cloud data from Velodyne VLS-128, namely `Velo-dyne_Points/Points.bag`. Acquired through the horizontally mounted laser scanner, represents a collection of three-dimensional coordinate points that depict the geometric shape and structure of the surrounding environment. This point cloud collection can be utilized for mapping, localization, object recognition and classification, semantic segmentation. This paper focuses on the application of the 128-channel LiDAR point cloud for aiding in localization and mapping in areas with poor GNSS signal reception.

The point cloud data from Riegl VUX-1HA, obtained through a tilted laser scanner, also represents a collection of three-dimensional coordinate points, namely `Riegl_Points/Points.csv`. And the scan data and mapping data are supplied as `Riegl_Points/Scans`, `Riegl_Points/Mapping`, respectively. This point cloud is able to be employed for mapping, and the generated point clouds can be utilized for object recognition and classification, semantic segmentation, creation of colored point clouds. In this paper, the utilization of single-line LiDAR point cloud for mapping and semantic segmentation is mainly discussed.

Panoramic image data is captured using the Ladybug5+ panoramic camera, which captures 360-degree panoramic images. The raw data captured by the Ladybug5+ (i.e. `Ladybug.pgr`) and GPS information for each image (i.e. `Ladybug_frame_gps_info.txt`) is provided. These data provide detailed visual information and can be used to colorize the point clouds. Additionally, the images captured by the Ladybug5+ panoramic camera can also be applied in applications such as object detection and semantic segmentation.

IMU data and GNSS data are acquired through the NovAtel SPAN CPT7, providing three-axis attitude data and GNSS positioning data. The Riegl VUX-1HA point cloud data is precisely mapped using the data from NovAtel SPAN CPT7. GNSS and IMU data are additionally stored in `.txt` files. For GNSS data, both WGS84 and local ENU coordinates are provided with a corresponding timestamp and a covariance matrix. The covariance

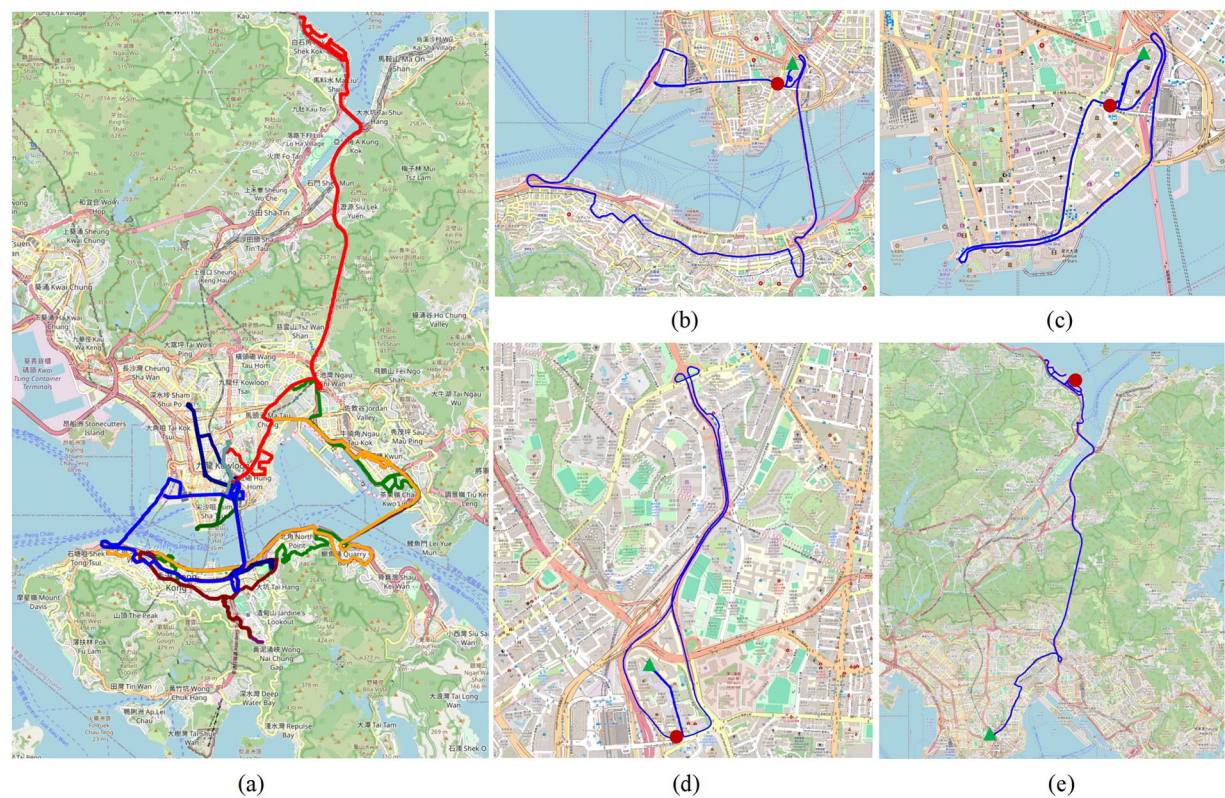


Fig. 2 Recording Zone. This figure shows the GPS traces of our recordings in the MSD-VMMS-HK dataset. (a), the GPS traces of the Hong Kong Hung Hom cross-harbour Tunnel, Western District cross-harbour Tunnel, Central District, (b), the GPS traces of Tsim Sha Tsui, (c), the GPS traces of Ho Man Tin, (d), the GPS traces of Tate's Cairn Tunnel, and Science Park, where the green triangle in each image indicates the start of the track and the red circle indicates the end.

Environment	Length	Acquisition time	Panoramas (frame)	Velodyne VLS-128 points (million)	Riegl VUX-1HA points (million)
Central District	1.5 km	Daytime	6, 043	882.2	513.8
Central District_2	2.6 km	Daytime	3, 268	—	752.7
Central District_3	1.4 km	Daytime	563	991.1	482.7
Tsim Sha Tsui	1.5 km	Daytime	4, 915	730.7	442.1
Cross-Harbour Tunnel	1.8 km	Daytime	5, 146	975.8	570.7
Eastern Harbour Crossing	2.3 km	Daytime	7, 965	1, 241.8	631.2
Western Harbour Crossing	2.0 km	Daytime	2, 312	619.3	363.3
Ho Man Tin	2.5 km	Daytime	6, 088	637.8	426.8
Fat Kwong Street	3.0 km	Daytime	—	1, 880.0	873.4
How Ming Street	1.3 km	Daytime	2, 929	1, 584.0	766.3
Hong Kong High Speed Way	1.6 km	Daytime	4, 063	730.5	317.9
Stubbs Road	1.5 km	Daytime	3, 135	789.1	359.6
Wan Chai Bypass	6.4 km	Daytime	3, 520	—	690.5
Western Kowloon	0.9 km	Daytime	1, 712	361.6	145.8
The Hong Kong Polytechnic University on-campus parking lot	0.3 km	Daytime	2, 303	305.1	166.1
Hong Kong Science Park and Tate's Cairn Tunnel	8.3 km	At night fall	22, 520	9, 028.7	1, 884.3

Table 1. Details of the MSD-VMMS-HK dataset. The Table provides specific details on capture length, acquisition time, and data obtained from Panoramas, Velodyne VLS-128, and Riegl VUX-1HA for partial data collection paths.

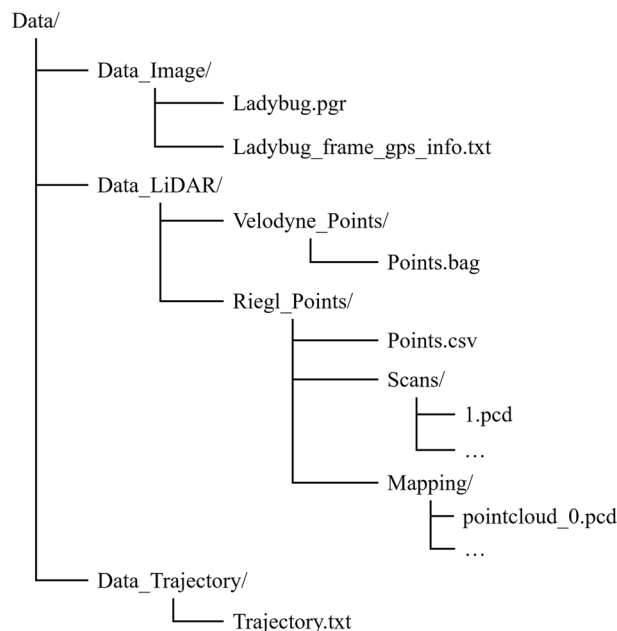


Fig. 3 Data structure of the MSD-VMMS-HK dataset.

Location	Error Type	RMSE/m	Mean/m	STD/m	Max/m
Fat Kwong Street	Horizontal position error	0.370	0.329	0.169	0.728
	Vertical position error	0.382	0.344	0.111	0.612
How Ming Street	Horizontal position error	0.656	0.629	0.342	1.302
	Vertical position error	2.842	2.793	0.528	3.342

Table 2. Horizontal and vertical position errors of point cloud data.

matrix is used to represent the accuracy of the RTK positioning. For the IMU measurement and GNSS data, timestamp, orientation in quaternion, angular velocity, linear acceleration, their corresponding covariance matrices, and raw data (i.e. Trajectory.txt) are provided.

Technical Validation

In the MSD-VMMS-HK dataset, the ground truth of positioning is provided by a NovAtel SPAN-CPT²⁴, a GNSS (GPS, GLONASS, and BeiDou) RTK/INS (fiber-optic gyroscope [FOG]) integrated navigation system. The baseline between the rover (SPAN-CPT) and the GNSS base station is within 7 km. According to the specifications of the NovAtel SPAN-CPT, centimeter-level accuracy can be obtained when the RTK correction is available with the correct fixed solution. However, accuracy is not guaranteed in urban canyons with inferences from building reflections. Therefore, the raw measurements is collected from the SPAN-CPT and postprocess the data using state-of-the-art Inertial Explorer software from NovAtel, which maximizes the accuracy of the trajectory by processing forward and backward in time, performing a backward smoothing step, and combining the results. Inertial Explorer can significantly improve the overall accuracy of the ground truth of positioning.

Ground truth of positioning in GNSS-denied scenarios: One of our datasets involves scenarios in a tunnel, where GNSS positioning is not available. As a result, the NovAtel SPAN-CPT cannot collect reliable GNSS measurements for the tunnel dataset. Fortunately, the NovAtel SPAN-CPT involves high-accuracy INS, which can provide low-drift dead-reckoning. To this end, the NovAtel SPAN-CPT is initialized from an open area in which the GNSS-RTK can easily obtain a fixed solution. As a result, the INS bias can be effectively calibrated. Then, the data collection vehicle enters the tunnel and exits the tunnel within 7 min. Next, the fixed GNSS-RTK solution can be obtained again by the NovAtel SPAN-CPT. In short, high-accuracy positioning can be achieved immediately before and after the data collection vehicle enters and exits the tunnel. Finally, the Inertial Explorer software is used to postprocess the collected raw data from the NovAtel SPAN-CPT to obtain reliable ground truth positioning. We carefully check the ground truth positioning of the tunnel data from the NovAtel SPAN-CPT based on the geodetic map from the Hong Kong government. At least meter-level accuracy can be guaranteed for the challenging urban roads and tunnel dataset. As demonstrated in the Table 2, the positioning accuracy of the point cloud data was rigorously evaluated using approximately 40 government-provided control points along both Fat Kwong Street and How Ming Street. The assessment revealed that the point cloud map of Fat Kwong Street achieved a mean horizontal positioning error of 0.329 m and a mean vertical positioning error of 0.344 m. Similarly, the point cloud data for How Ming Street exhibited comparable accuracy levels, with mean

Dataset	Year	Envrnmnt	LiDAR	Camera	IMU	Ground Truth
Darpa Urban ³⁸	2010	outdoor	12 SICK LMS291@75 Hz Velodyne HDL-64E@15 Hz	Point Grey: $4 \times 376 \times 240$ @10 Hz $1 \times 752 \times 480$ @22.8 Hz	Applanix POS-LV 220	GPS + INS
Ford Campus ⁴	2011	outdoor	Velodyne HDL-64E@10 Hz 2 Riegl LMS-Q120	Ladybug 3: $6 \times 1600 \times 600$ @8 Hz	Applanix POS-LV 420	GPS + INS
UrbanNav ⁶	2021	outdoor	Velodyne HDL 32E@10 Hz	Fisheye camera: 1920×1200 @10 Hz Monocular camera	GNSS/INS: NovAtel SPAN-CPT IMU: Xsens MTI10@100 Hz	GPS + INS
Our Dataset	2024	outdoor + indoor	Velodyne VLS-128@10 Hz Riegl VUX-1HA@250 Hz	Ladybug 5 P+: 8192×4096 @15 Hz	GNSS/INS: NovAtel SPAN-CPT7	GPS + INS + SLAM

Table 3. Datasets for localization and mapping.

Dataset	Year	Envrnmnt	LiDAR	Camera	IMU	Ground Truth
Darpa Urban ³⁸	2010	outdoor	12 SICK LMS291@75 Hz Velodyne HDL-64E@15 Hz	Point Grey: $4 \times 376 \times 240$ @10 Hz $1 \times 752 \times 480$ @22.8 Hz	Applanix POS-LV 220	GPS + INS
Ford Campus ⁴	2011	outdoor	Velodyne HDL-64E@10 Hz 2 Riegl LMS-Q120	Ladybug 3: $6 \times 1600 \times 600$ @8 Hz	Applanix POS-LV 420	GPS + INS
KITTI ¹⁵	2013	outdoor	Velodyne HDL 64E@10 Hz	Point Grey (2 gray + 2 RGB): $4 \times 1392 \times 512$ @10 Hz	OXTS RT3003	GPS + INS
Oxford RoboCar ³⁹	2017	outdoor	2 SICK LMS151@50 Hz SICK LD-MRS@12.5 Hz	BumbleBee XB3: $2 \times 1280 \times 960$ @16 Hz 3 Grasshopper2: $3 \times 1024 \times 1024$ @11.1 Hz	GNSS/INS: NovAtel SPAN-CPT	GPS + INS + VO
SemanticKITTI ⁴⁰	2019	outdoor	Velodyne HDL 64E	N/A	N/A	N/A
Oxford Radar RobotCar ⁴¹	2020	outdoor	2 HDL-32E@20 Hz 2 SICK LMS151@50 Hz	BumbleBee XB3: $2 \times 1280 \times 960$ @16 Hz 3 Grasshopper2: $3 \times 1024 \times 1024$ @11.1 Hz	GNSS/INS: NovAtel SPAN-CPT	GPS + INS + VO
NuScenes Dataset ⁴²	2019	outdoor	Velodyne HDL 32E@20 Hz	6 cameras Basler acA1600-60gc @12 Hz @12 Hz	IMU + GPS	GPS + INS + Monte Carlo Localization scheme
ApolloScape ⁴³	2018	outdoor	Riegl VMX-1HA	VMX-CS6 camera system (3384×2710)	IMU/GNSS	GPS + INS
WHU-Urban3D ⁴⁴	2024	outdoor	N/A	N/A	N/A	N/A
Our Dataset	2024	outdoor + indoor	Velodyne VLS-128@10 Hz Riegl VUX-1HA@250 Hz	Ladybug 5 P+: 8192×4096 @15 Hz	GNSS/INS: NovAtel SPAN-CPT7	GPS + INS + SLAM

Table 4. Datasets for autonomous driving.

errors remaining within acceptable thresholds. This meter-level precision provides a reliable spatial reference framework suitable for various applications including: high-definition map generation, urban infrastructure inspection and autonomous vehicle localization.

GNSS reference station data: The Hong Kong government provides GNSS reference station service, named the Hong Kong Satellite Positioning Reference Station Network (SatRef). SatRef consists of 16 reference stations and 2 integrity monitoring stations evenly distributed throughout Hong Kong. SatRef provides raw data for postprocessing in receiver independent exchange format (RINEX) format for web or file transfer protocol (FTP) download; details can be found at <https://www.geodetic.gov.hk/en/satref/satref.htm>. This resource provides both RINEX 2 and 3 versions, with a file length of 1 h or 24 h. For RINEX 3.02, the file length of 1 h provides a data interval of 1 s or 5 s; for 24 h, only a data interval of 30 s is provided.

Usage Notes

The introduction and usage of the MSD-VMMS-HK dataset. The MSD-VMMS-HK dataset offers the comprehensive data source among the listed categories of vehicle-mounted mobile mapping datasets in Tables 3, 4. By utilizing two different type of laser scanners, 360-degree coverage of the full scene is realized and provides point cloud data of varying accuracy for different applications in positioning and mapping. The INS measurement frequency for acceleration, angular velocity, and orientation is set at 100 Hz. The RTK-GNSS data is also provided for georeferencing and benchmarking mileage measurements. One of the most impressive features of our dataset is the collection of panoramic images at a frequency of 8192×4096 @15 Hz, capturing nearly half billion RGB pixels per second. This aspect proves highly valuable for various downstream applications such as

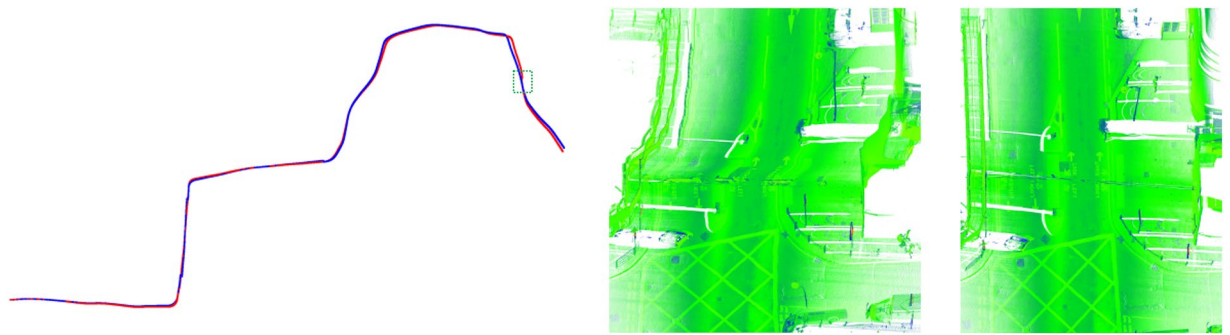


Fig. 4 (a) The trajectory comparison between the IMU-only and IMU + SLAM approaches is depicted in the trajectory diagram. (b) The point cloud map constructed based on IMU trajectory and single-line LiDAR point cloud data. (c) The point cloud map constructed based on IMU + SLAM trajectory and single-line LiDAR point cloud data.

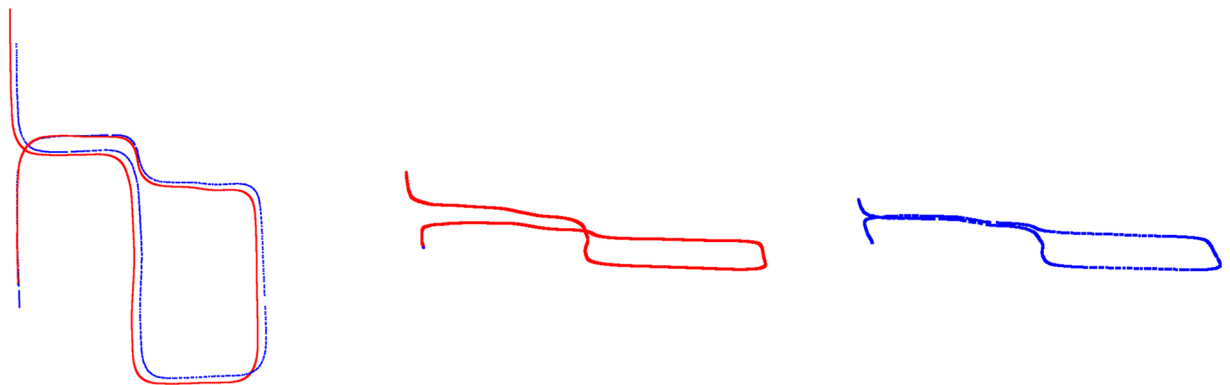


Fig. 5 (a) The trajectory comparison between the IMU-only and IMU + SLAM approaches is depicted in the trajectory diagram. (b) The IMU trajectory. (c) The IMU + SLAM trajectory.

color mapping, modeling, semantic segmentation, and object detection. To enhance localization accuracy, we employ a GNSS + INS + SLAM approach. This method combines GNSS, INS, and SLAM technologies, effectively improving the accuracy and stability of vehicle positioning.

In addition, the dataset of vehicle-mounted mobile mapping system presented in this paper can be applied to wide range of domains to solve potentially challenging problems in various fields due to its accuracy of localization and reliability of point cloud and image data. For example, it can be applied in areas such as mapping of urban roads, autonomous driving, high-definition maps, semantic segmentation, change detection, and digital twins. The following section describes in detail the application areas of the dataset and the potential problems it can contribute to solving.

Mapping of urban scenes. To begin with, the vehicle-mounted mobile mapping dataset can be utilized for the creation and updating of three-dimensional digital maps, encompassing information such as road networks, buildings, terrain, and other geographical features. These maps play a vital role in various applications, including traffic planning²⁵, navigation systems, urban planning, and emergency management, high definition (HD) mapping²⁶, virtual reality (VR)²⁷, digital twin²⁸, and so on.

Inaccurate IMU odometry calculations were frequently encountered in the datasets of Hong Kong Central District and the Hong Kong Polytechnic University campus parking lots. This issue arises from the poor or interrupted GNSS reception, there are many tall buildings in Central, Hong Kong, which block the GNSS signals, resulting in poor GNSS reception. Consequently, prolonged periods of GNSS data loss contribute to the gradual accumulation of errors in the IMU odometry. Without GNSS data for trajectory correction, the IMU odometry drifts over time, such as distorted road curvatures (as depicted in Fig. 4) and height discrepancies in the trajectory (as shown in Fig. 5). To overcome this challenge, the SLAM techniques were employed to rectify the IMU odometry. By utilizing SLAM-derived LiDAR odometry for mapping single-line LiDAR point clouds, significant improvements were observed in these situations. For example, in the Hong Kong Central District dataset, the optimization of the odometry using SLAM techniques improved the accuracy of the positioning from the meter level to the decimeter level.

Autonomous driving and intelligent traffic systems. The vehicle-mounted mobile mapping system dataset provides valuable input data for autonomous driving and intelligent traffic systems, such as obstacle

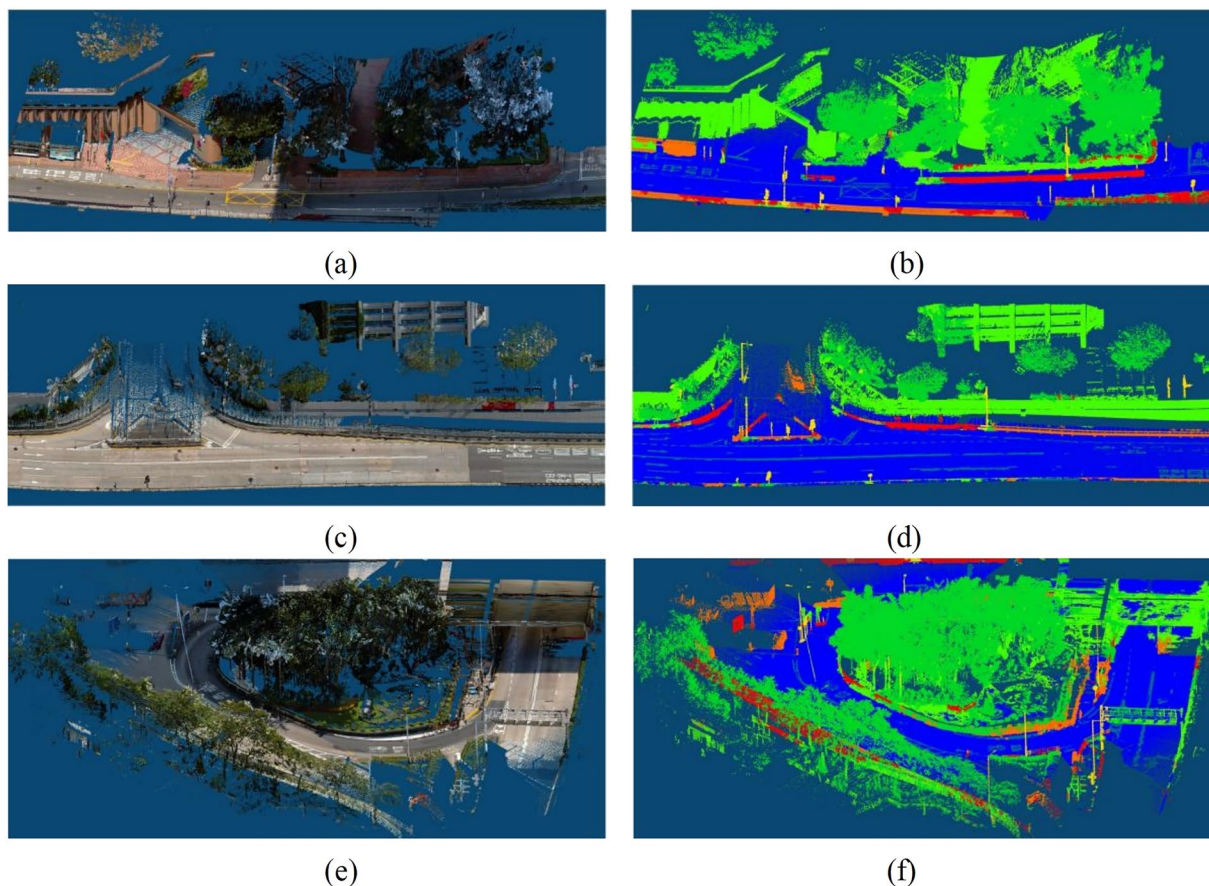


Fig. 6 (a) The colorized point cloud map of campus area. (b) The corresponding semantic segmentation image of the colorized point cloud. (c) The colorized point cloud map of high speed road. (d) The corresponding semantic segmentation image of the colorized point cloud. (e) The colorized point cloud map of circular interchange highway. (f) The corresponding semantic segmentation image of the colorized point cloud.

detection²⁹, road sign recognition³⁰, road surface extraction³¹, vegetation extraction³², building extraction³³, and traffic flow analysis.

As shown in the Fig. 6, the colorized point cloud of our vehicle-mounted mobile mapping dataset can be classified using a well-trained deep learning network, such as Randla-net³⁴, the semantic segmentation network Randla-net was trained using the point cloud data in our dataset. Subsequently, the performance of the trained network was evaluated using the point cloud data from a distinct subset of our dataset. The results affirm that the point cloud data from our dataset is highly effective in achieving precise semantic segmentation. Different colors are used to represent the categories of the classified objects. The available labels include vehicles, road surfaces, lane markings, sidewalks, guardrails, traffic signs, streetlights, vegetation, buildings, and more. This classification process enhances the understanding and analysis of the environment, and these above-mentioned research results are effectively applied in the fields of autonomous driving and intelligent traffic systems.

Infrastructure management and maintenance. The street-level dataset generated by the vehicle-mounted mobile mapping system proves valuable for monitoring and managing infrastructure, including the identification of changes in buildings and billboards³⁵, as well as road condition assessment³⁶. This aids in timely detection of potential issues, facilitating maintenance and repair work to ensure the safety and reliability of infrastructure. As shown in Fig. 7, panoramic images within the street-level dataset can be leveraged to identify road defects such as cracks and potholes, thus contributing to road maintenance efforts.

Change detection. The point cloud data generated by the vehicle-mounted mobile mapping system can also be applied to detect changes in building facilities³⁷. In the fields of urban planning and building monitoring, point cloud change detection plays a crucial role in monitoring changes and damages in buildings. By comparing point cloud data from different time, it becomes possible to detect various changes in buildings, such as expansions, demolitions, and structural modifications. In the manufacturing and industrial automation sectors, point cloud change detection can be used to identify changes in object positions and statuses along the production line. Real-time analysis of point cloud data allows for the timely detection of anomalies, such as missing components or assembly errors. In the provided Fig. 8, two point cloud maps collected on the campus of the Hong Kong



Fig. 7 Schematic illustration of urban road defects captured by the panoramic camera of a vehicle-mounted mobile mapping system.

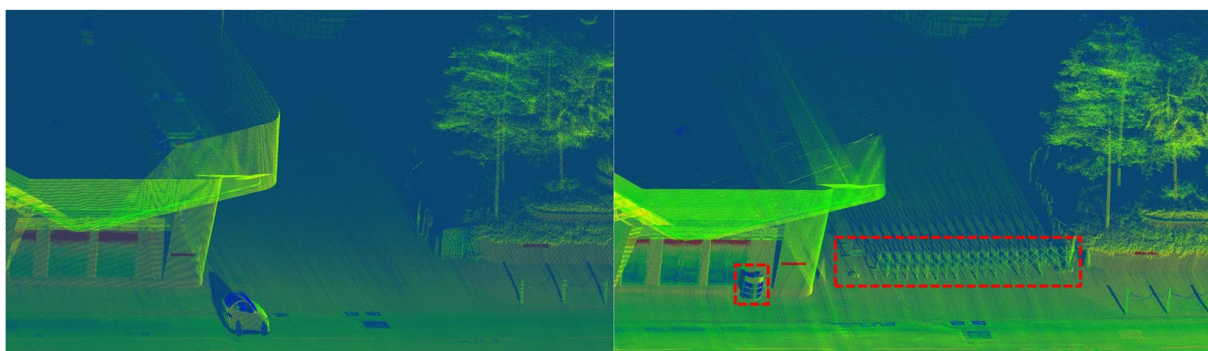


Fig. 8 (a) Initial acquisition of the point cloud map of road infrastructure. (b) Updated point cloud of road facilities captured in the later stage, along with annotated schematic illustrations highlighting the changed road facilities.

Polytechnic University at different times reveal significant changes in the roadside facilities. For example, there are noticeable differences in the presence of a reception desk and the availability of retractable road barriers.

In summary, our dataset, employing the GNSS + INS + SLAM approach and featuring diverse scene data, offers robust support for improving localization accuracy. This contribution accelerates the development of autonomous driving technology, enhancing vehicle positioning performance across various complex driving scenarios.

The link to the vehicle-mounted mobile mapping dataset presented in this paper is: <https://scri-platform.org/vmms>. Any updates or expansion plans will also be documented at this website, providing researchers and users with the latest information and resources related to this dataset.

Code availability

A specialized data processing software and code tailored for the MSD-VMMS-HK dataset is accessible on the website: <https://github.com/yutoulu/The-code-for-MSD-VMMS-HK-dataset>. This software can efficiently handle diverse data types mentioned in the paper, including point cloud data, IMU data, and GNSS data. Note that the Ladybug5+ and NovAtel SPAN CPT7 data need to be pre-processed using their respective software: Teledyne LadybugCapPro and Waypoint-Inertial Explorer 8.90.

Received: 8 January 2025; Accepted: 26 June 2025;

Published online: 13 August 2025

References

- Puente, I., González-Jorge, H., Martínez-Sánchez, J. & Arias, P. Review of mobile mapping and surveying technologies. *Measurement* **46**, 2127–2145, <https://doi.org/10.1016/j.measurement.2013.03.006> (2013).
- Chiang, K. W., Tsai, G.-J. & Zeng, J. C. Mobile mapping technologies. *Urban Informatics* 439–465, https://link.springer.com/chapter/10.1007/978-981-15-8983-6_25 (2021).
- Bao, S., Shi, W., Yang, D., Xiang, H. & Yu, Y. Global principal planes aided lidar-based mobile mapping method in artificial environments. *Adv. Eng. Informatics* **61**, 102472, <https://doi.org/10.1016/j.aei.2024.102472> (2024).
- Pandey, G., McBride, J. R. & Eustice, R. M. Ford campus vision and lidar data set. *The Int. J. Robotics Res.* **30**, 1543–1552, <https://doi.org/10.1177/0278364911400640> (2011).
- Wang, C. *et al.* Isprs benchmark on multisensory indoor mapping and positioning. *ISPRS Annals Photogramm. Remote. Sens. Spatial Inf. Sci.* **5**, 117–123, <https://doi.org/10.5194/isprs-annals-V-5-2020-117-2020> (2020).
- Hsu, L.-T. *et al.* Urbannav: An open-sourced multisensory dataset for benchmarking positioning algorithms designed for urban areas. In *Proceedings of the 34th international technical meeting of the satellite division of the institute of navigation (ION GNSS + 2021)*, 226–256, <https://doi.org/10.33012/2021.17895> (2021).
- Shi, W. *et al.* Polyu-bpcoma: A dataset and benchmark towards mobile colorized mapping using a backpack multisensorial system. *Int. J. Appl. Earth Obs. Geoinformation* **112**, 102962, <https://doi.org/10.1016/j.jag.2022.102962> (2022).
- Zhang, X., Huang, Z., Li, Q., Wang, R. & Zhou, B. Legged robot-aided 3 d tunnel mapping via residual compensation and anomaly detection. *ISPRS J. Photogramm. Remote. Sens.* **214**, 33–47, <https://doi.org/10.1016/j.isprsjprs.2024.05.025> (2024).
- Wen, C. *et al.* Toward efficient 3-d colored mapping in gps-/gnss-denied environments. *IEEE Geosci. Remote. Sens. Lett.* **17**, 147–151, <https://doi.org/10.1109/LGRS.2019.2916844> (2019).
- Carlevaris-Bianco, N., Ushani, A. K. & Eustice, R. M. University of michigan north campus long-term vision and lidar dataset. *The Int. J. Robotics Res.* **35**, 1023–1035, <https://doi.org/10.1177/0278364915614638> (2016).
- Gu, Y., Hsu, L.-T. & Kamijo, S. Gns/onboard inertial sensor integration with the aid of 3-d building map for lane-level vehicle self-localization in urban canyon. *IEEE Transactions on Veh. Technol.* **65**, 4274–4287, <https://doi.org/10.1109/TVT.2015.2497001> (2015).
- Gong, Z. *et al.* Mapping and semantic modeling of underground parking lots using a backpack lidar system. *IEEE Transactions on Intell. Transp. Syst.* **22**, 734–746, <https://doi.org/10.1109/TITS.2019.2955734> (2019).
- Kim, K., Im, J. & Jee, G. Tunnel facility based vehicle localization in highway tunnel using 3 d lidar. *IEEE Transactions on Intell. Transp. Syst.* **23**, 17575–17583, <https://doi.org/10.1109/TITS.2022.3160235> (2022).
- Yang, C. & Soloviev, A. Mobile positioning with signals of opportunity in urban and urban canyon environments. In *2020 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, 1043–1059, <https://doi.org/10.1109/PLANS46316.2020.9109876> (IEEE, 2020).
- Geiger, A., Lenz, P. & Urtasun, R. Are we ready for autonomous driving? the kitti vision benchmark suite. In *2012 IEEE conference on computer vision and pattern recognition*, 3354–3361, <https://doi.org/10.1109/CVPR.2012.6248074> (IEEE, 2012).
- Bresson, G., Alsayed, Z., Yu, L. & Glaser, S. Simultaneous localization and mapping: A survey of current trends in autonomous driving. *IEEE Transactions on Intell. Veh.* **2**, 194–220, <https://doi.org/10.1109/TIV.2017.2749181> (2017).
- Chebroli, N. *et al.* Agricultural robot dataset for plant classification, localization and mapping on sugar beet fields. *The Int. J. Robotics Res.* **36**, 1045–1052, <https://doi.org/10.1177/0278364917720510> (2017).
- Ruiz-Sarmiento, J. R., Galindo, C. & González-Jiménez, J. Robot@ home, a robotic dataset for semantic mapping of home environments. *The Int. J. Robotics Res.* **36**, 131–141, <https://doi.org/10.1177/0278364917695640> (2017).
- Xiao, J. *et al.* Tiny object detection with context enhancement and feature purification. *Expert. Syst. with Appl.* **211**, 118665, <https://doi.org/10.1016/j.eswa.2022.118665> (2023).
- Tan, W. *et al.* Toronto-3d: A large-scale mobile lidar dataset for semantic segmentation of urban road- ways. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops*, 202–203, https://openaccess.thecvf.com/content_CVPRW_2020/html/w11/Tan_Toronto-3D_A_Large-Scale_Mobile_LiDAR_Dataset_for_Semantic_Segmentation_of_CVPRW_2020_paper.html?ref=https://githubhelp.com (2020).
- Bao, S., Shi, W., Chen, P. & Xiang, H. Sole gnss sensor extrinsic calibration methods for mobile mapping systems. *IEEE Robotics Autom. Lett.* **7**, 11338–11345, <https://doi.org/10.1109/LRA.2022.3193632> (2022).
- Niu, X. *et al.* Quality evaluation of the pulse per second (pps) signals from commercial gnss receivers. *GPS solutions* **19**, 141–150, <https://link.springer.com/article/10.1007/s10291-014-0375-7> (2015).
- Gutiérrez, C. S. V. *et al.* Time synchronization in modular collaborative robots. *arXiv preprint arXiv:1809.07295* <https://doi.org/10.48550/arXiv.1809.07295> (2018).
- Kennedy, S., Hamilton, J. & Martell, H. Architecture and system performance of span—novatel's gps. *INS Solution, Tech. Pap. NovAtel* (2006).
- Li, X. *et al.* Traffic management and forecasting system based on 3 d gis. In *2015 15th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing*, 991–998, <https://doi.org/10.1109/CCGrid.2015.62> (IEEE, 2015).
- Liu, R., Wang, J. & Zhang, B. High definition map for automated driving: Overview and analysis. *The J. Navig.* **73**, 324–341, <https://doi.org/10.1017/S0373463319000638> (2020).
- Anthes, C., García-Hernández, R. J., Wiedemann, M. & Kranzlmüller, D. State of the art of virtual reality technology. In *2016 IEEE aerospace conference*, 1–19, <https://doi.org/10.1109/AERO.2016.7500674> (IEEE, 2016).
- Tao, F., Zhang, H., Liu, A. & Nee, A. Y. Digital twin in industry: State-of-the-art. *IEEE Transactions on industrial informatics* **15**, 2405–2415, <https://doi.org/10.1109/TII.2018.2873186> (2018).
- Yu, X. & Marinov, M. A study on recent developments and issues with obstacle detection systems for automated vehicles. *Sustainability* **12**, 3281, <https://doi.org/10.3390/su12083281> (2020).
- Greenhalgh, J. & Mirmehdi, M. Real-time detection and recognition of road traffic signs. *IEEE transactions on intelligent transportation systems* **13**, 1498–1506, <https://doi.org/10.1109/TITS.2012.2208909> (2012).
- Wei, Y., Zhang, K. & Ji, S. Simultaneous road surface and centerline extraction from large-scale remote sensing images using cnn-based segmentation and tracing. *IEEE Transactions on Geosci. Remote. Sens.* **58**, 8919–8931, <https://doi.org/10.1109/TGRS.2020.2991733> (2020).

32. Xiaoqin, W., Miaomiao, W., Shaoqiang, W. & Yundong, W. Extraction of vegetation information from visible unmanned aerial vehicle images. *Transactions Chin. Soc. Agric. Eng.* **31**, <https://doi.org/10.3969/j.issn.1002-6819.2015.05.022> (2015).
33. Du, S. *et al.* Automatic building extraction from lidar data fusion of point and grid-based features. *ISPRS journal photogrammetry remote sensing* **130**, 294–307, <https://doi.org/10.1016/j.isprsjprs.2017.06.005> (2017).
34. Hu, Q. *et al.* Randla-net: Efficient semantic segmentation of large-scale point clouds. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 11108–11117, https://openaccess.thecvf.com/content_CVPR_2020/html/Hu_RandLA-Net_Efficient_Semantic_Segmentation_of_Large-Scale_Point_Clouds_CVPR_2020_paper.html (2020).
35. Qin, R. & Gruen, A. 3d change detection at street level using mobile laser scanning point clouds and terrestrial images. *ISPRS J. Photogramm. Remote. Sens.* **90**, 23–35, <https://doi.org/10.1016/j.isprsjprs.2014.01.006> (2014).
36. Shahi, K., Shafri, H. Z. M. & Hamedianfar, A. Road condition assessment by obia and feature selection techniques using very high-resolution worldview-2 imagery. *Geocarto Int.* **32**, 1389–1406, <https://doi.org/10.1080/10106049.2016.1213888> (2017).
37. Lu, Q., Xie, X., Parlikad, A. K. & Schooling, J. M. Digital twin-enabled anomaly detection for built asset monitoring in operation and maintenance. *Autom. Constr.* **118**, 103277, <https://doi.org/10.1016/j.autcon.2020.103277> (2020).
38. Huang, A. S. *et al.* A high-rate, heterogeneous data set from the darpa urban challenge. *The Int. J. Robotics Res.* **29**, 1595–1601, <https://doi.org/10.1177/0278364910384295> (2010).
39. Maddern, W., Pascoe, G., Linegar, C. & Newman, P. 1 year, 1000 km: The oxford robotcar dataset. *The Int. J. Robotics Res.* **36**, 3–15, <https://doi.org/10.1177/0278364916679498> (2017).
40. Behley, J. *et al.* Semantickitti: A dataset for semantic scene understanding of lidar sequences. In *Proceedings of the IEEE/CVF international conference on computer vision*, 9297–9307, https://openaccess.thecvf.com/content_ICCV_2019/html/Behley_Semantickitti_A_Dataset_for_Semantic_Scene_Understanding_of_LiDAR_Sequences_ICCV_2019_paper.html (2019).
41. Barnes, D., Gadd, M., Murcutt, P., Newman, P. & Posner, I. The oxford radar robotcar dataset: A radar extension to the oxford robotcar dataset. In *2020 IEEE international conference on robotics and automation (ICRA)*, 6433–6438, <https://doi.org/10.1109/ICRA40945.2020.9196884> (IEEE, 2020).
42. Caesar, H. *et al.* nuscenes: A multimodal dataset for autonomous driving. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 11621–11631, https://openaccess.thecvf.com/content_CVPR_2020/html/Caesar_nuScenes_A_Multimodal_Dataset_for_Autonomous_Driving_CVPR_2020_paper.html (2020).
43. Huang, X. *et al.* The apolloscape dataset for autonomous driving. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 954–960, https://openaccess.thecvf.com/content_cvpr_2018_workshops/w14/html/Huang_The_ApolloScape_Dataset_CVPR_2018_paper.html (2018).
44. Han, X. *et al.* Whu-urban3d: An urban scene lidar point cloud dataset for semantic instance segmentation. *ISPRS J. Photogramm. Remote. Sens.* **209**, 500–513, <https://doi.org/10.1016/j.isprsjprs.2024.02.007> (2024).

Acknowledgements

We thank the Shenzhen Park of Hetao Shenzhen-Hong Kong Science and Technology Innovation Cooperation Zone and this research has been supported by the “Theories for Spatiotemporal Intelligence and Reliable Data Analysis” (Project ID: HZQSW-S-KCCYB-2024058). This research has also been supported by Otto Poon Charitable Foundation Smart Cities Research Institute, the Hong Kong Polytechnic University (Work Program: CD06); The Hong Kong Polytechnic University (E-RB0Y). The authors are grateful to the anonymous reviewers for their insightful comments on improving this article.

Author contributions

S.L. and S.B. wrote the manuscript. S.L. and S.B. conceived the experiments, S.L., S.B. and Y.W. conducted the experiments, S.L., S.B., S.Z. and D.Y. analysed the results. W.S. provided the supervision of this project. All authors reviewed the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to W.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025