IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 14 October 2022, accepted 31 October 2022, date of publication 4 November 2022, date of current version 10 November 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3219415

RESEARCH ARTICLE

Mobile Mapping System for Automatic Extraction of Geodetic Coordinates for Traffic Signs Based on Enhanced Point Cloud Reconstruction

CHENG-WEI PENG, CHEN-CHIEN HSU, (Senior Member, IEEE), AND WEI-YEN WANG, (Fellow, IEEE)

Department of Electrical Engineering, National Taiwan Normal University, Taipei 10610, Taiwan

Corresponding author: Chen-Chien Hsu (jhsu@ntnu.edu.tw)

This work was financially supported by the "Chinese Language and Technology Center" of National Taiwan Normal University (NTNU) from The Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education (MOE) in Taiwan, and National Science and Technology Council (NSTC), Taiwan, under Grants no. 111-2221-E-003-025, 110-2221-E-003-019, 110-2221-E-003-020-MY2, and 110-2634-F-A49-004 under the program of AI Thematic Research Program to Cope with National Grand Challenges through Pervasive Artificial Intelligence Research (PAIR) Labs of the National Yang Ming Chiao Tung University. We are also grateful to the National Center for High-Performance Computing for computer time and facilities to conduct this research.

ABSTRACT Lidar sensors are commonly equipped on a mobile mapping system (MMS) to establish point clouds for HD map creation. However, the point clouds themselves do not contain object attributes. Therefore, human operators have to manually obtain objects' position to assign attributes for further HD map conversion, inevitably resulting in time-consuming processes and significant labor costs. To solve the above problems, in this paper, we present an MMS equipped with non-survey grade Lidar, commercial grade camera, and entry level GNSS/INS, which incorporates ground control points (GCPs) with a Normal Distribution Transform Simultaneously Localization and Mapping (NDT SLAM) refinement and fluctuation adjustment to secure both absolute position accuracy and relative position accuracy of the reconstructed point cloud. Meanwhile, a deep neural network for image detection is employed to obtain the bounding box of traffic signs from each image frame. By applying the translation and rotation transformation between Lidar points and camera pixels, intersection of the detected object in the image and Lidar scan points can be found. By accumulating extracted Lidar points of the traffic sign in several detection frames, we can then obtain an accurate 3D geodetic coordinate of the traffic signs. Experimental results show that point clouds can be reconstructed with an average 3D RMSE of only 8.6cm, and center geodetic coordinates of traffic signs can be further extracted in sub-meter accuracy to significantly reduce labor work in HD map creation.

INDEX TERMS HD map, point cloud, GIS, autonomous driving.

I. INTRODUCTION

High Definition Map (HD Map) is essential for localization, route planning, and environment perception in autonomous driving. The higher a vehicle autonomy level (i.e., L1-L5) is, the tighter the sensor fusion results are to be coupled with the HD Map. Considering the interoperability for various autonomy levels, HD map data should therefore be exchangeable and convertible. Over the past years, specifications of HD map are springing up due to the high growth of autonomous driving demand. Although proprietary HD map specifications are available for specific industries and organizations, the importance of open HD map formats has become more evident. For example, Lanelet2 [1] and OPENDRIVE [2] are the two open and commonly used HD map formats in autonomous driving development. The well-known autonomous driving platform, Autoware [3], has supported Lanelet2 in its latest release as shown in Figure 1. Simulation Software CarSim [4] and VTD [5] always rapidly support the latest version of OPENDRIVE in their products.

The essential components of a HD map for autonomous driving usage include topological information, geometric

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

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

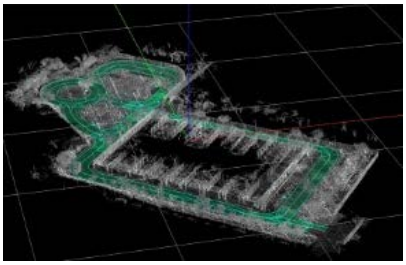


FIGURE 1. HD map Lanelet2 format used in Autoware [3].

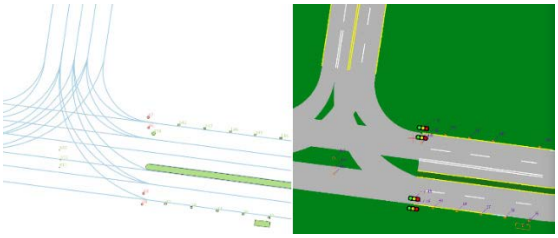


FIGURE 2. HD map conversion from raw map data (left) to OPENDRIVE format (right).

information, and semantic information [6]. The mobile mapping systems (MMS) equipped with Lidar are the mainstream solution for data collection to contribute to high accuracy geodetic information from a point cloud. Professional Geographic Information System (GIS) software like ArcGIS [7] and MicroStation [8] are used in dealing with point clouds for semantic content extraction and validation. Through the GIS software, extracted objects with geodetic information are typically stored into a cloud database or a standalone file (i.e., *.shp). Map providers then follow the HD map specifications to convert the raw map data to meet different demands.

As shown in Figure 2, points are the fundamental elements in representing 3D objects like poles, traffic lights, traffic signs, etc. Road centerlines and lane centerlines containing road/lane width and traffic regulation information are stored in line type. Stop lines, road markings, roadside parking, and pedestrian refuges are stored in polygon type. It is mandatory to assign an ID to each road/lane segment to correlate all other points, polygons, and lines. Furthermore, junctions are defined as a virtual area in polygon type to fulfill the demand of a specific HD map format.

With the use of conventional methods, depicting the road/lane information from a point cloud is not time exhausting, but extracting an object from the point cloud is. The massive time spent for object extraction starts from finding and picking an exact center point of an object. Next, the width and height are measured from the point cloud. For example, a traffic sign appearing in the point cloud obtained by a GIS software can be easily seen. When a human operator inspects the traffic sign, its background points also exist in the view. As soon as the operator clicks the center position of the traffic sign, a background point can often be wrongly picked in practice. Therefore, designing a suitable method to pick the correct center point of objects from an accurate point cloud for HD map creation is highly expected.

Based on the above-mentioned discussions, problems that we encountered for HD creation can be summarized as follows:

1. High-accuracy point cloud reconstruction based on a cost-effective hardware configuration is desired for practical real-world applications.

2. The position accuracy of extracted objects by a cost-effective MMS should be secured for HD map creation.

3. Human labor intervention should be adequately reduced in the HD map creation process.

In this paper, we present an MMS equipped with non-survey grade sensors, which incorporates ground control points (GCPs) with a Normal Distribution Transform Simultaneously Localization and Mapping (NDT SLAM) [9] refinement and fluctuation adjustment, to secure both absolute position accuracy and relative position accuracy of the reconstructed point cloud in the first place. Secondly, thanks to the advance of Artificial Intelligence (AI), Deep Learning (DL) deep neural networks (DNN) have amply demonstrated their prowess to sense, reason, act, and adapt to imitate human ability or behavior for various applications over the past years. An architecture CDLSTM [10] was proposed for accurate climate change forecasting, and a DNN model SMOTEDNN [11] were presented for efficient classification and forecasting of air pollution, to name a few of them. In this paper, a DNN image detection model DriveNet is conducted to detect and classify the objects. By applying the rotation and translation parameters between camera and Lidar, the intersection of the detected object in the image and Lidar scan points can be found. Finally, a proposed algorithm is then implemented to remove the outlier Lidar points according to an intensity threshold. By accumulating extracted Lidar points of the traffic sign in several detection frames, we can then deduce an accurate 3D position of the traffic sign center. This method is simple, much straightforward, and time efficient in HD map creation. According to the experimental results, the traffic sign extraction has an absolute position error in sub-meter accuracy, providing road regulation information with high position accuracy for autonomous driving, road asset management [12], and related intelligent transportation system usages. To this end, the contributions of the proposed work are summarized as follows:

1. The quality of point cloud is enhanced by the proposed MMS equipped with non-survey grade sensors to deal with wide multi-lane roads, urban areas, and other complex environments.

2. The accurate geodetic information for raw map data can be automatically obtained through the proposed geodetic coordinates extraction approach.

3. The presented comprehensive workflow achieving high-accuracy geodetic coordinates extraction could relieve human labor intervention when creating the HD map.

This paper is organized as follows. Section II discusses related works for HD map creation. Section III describes the hardware/software configuration of the proposed MMS and methods to reconstruct the point cloud in a wide road scene.

Section IV presents the proposed approach for automatic traffic sign extraction from a point cloud, where key extraction methods are described in detail. Section V shows the experimental results that we conducted using the proposed hardware, software, and methods. Finally, we conclude this paper in Section VI.

II. RELATED WORKS

For HD map creation, Jiao [13] summarized a complete workflow from the autonomous driving aspect. This paper concluded that the HD map creation requires an accurate 3D object extraction process for execution as efficient as possible. Poggenhans et al. [14] gave a detailed review of the open HD map format Lanelet2, which consists of Points, Linestrings, Lanelets, Areas, and Regulatory elements. The paper enumerated several examples in the real world to illustrate how mandatory and optional items in Lanelet2 are to be filled in. It can be seen that correctly connecting the relationship of objects is indispensable. The other popular open HD map format OpenDRIVE is also widely discussed in the autonomous driving field. Becker et al. [15] introduced the geometric expression in OpenDRIVE format in detail, pointing out the elements in OpenDrive, including Reference Lines and Segments, are represented by curvature (composed of spirals and arcs). The authors contributed a detailed guideline for geometry conversion in representing an OpenDRIVE road network topology. To consider the interoperability for those open HD map formats, Althoff et al. [16] presented a converter from OpenDRIVE to Lanelets (predecessor of Lanelet2). It shows that open HD map formats can be exchangeable and convertible. From the above survey, the development of HD map formats is quite evident nowadays. Hence it is worthwhile to investigate the methods to reduce data collection cost and increase 3D object extraction efficiency in HD map creation.

The erroneous scenarios of localization based on 3D pint cloud were investigated by Javanmardi et al. [17]. This paper concluded that feature sufficiency, layout, local similarity, and representation quality are factors that highly affect the accuracy in self-localization. To obtain an accurate 3D point cloud, MMSs are commonly used in data collection. Traditional filter-based positioning systems are the mainstream methods in outdoor terrain survey. Yao et al. [18] presented an MMS by integrating 2D Lidars, panoramic images and GNSS/INS to obtain an accurate point cloud. Zeng and Zhong [19] presented an algorithm to assign the color attributes from panoramic images to a 3D laser point cloud based on a proposed MMS. Ilci et al. [20] utilized multiple well calibrated 3D Lidars to reconstruct the point cloud for autonomous driving localization. In the traditional filter-based MMS, the quality of the point cloud highly depends on the grade of the MMS positioning systems employed. Hence there come out the graph-based optimization methods to mitigate the costs of the MMS hardware deployment. For graph-based optimization, the key concept is to align each different scan on the same static feature. By performing several iterations to find the best features alignment, a refined position and orientation of the Lidar can be found. Hence, a refined trajectory can be subsequently reconstructed from the whole Lidar scans as a result. Hamieh et al. [21] presented a system that integrated entry level GNSS/INS with Lidar odometry localization and mapping method LEGO-LOAM to predict the position and orientation to produce the point cloud. Peng et al. [22] utilized GCPs with NDT SLAM to refine the position trajectory to effectively secure the absolute/relative position of the reconstructed point cloud. Koide et al. [23] presented an interactive method that allows users to correct point cloud mapping failures to improve the quality of a 3D point cloud. Based on an accurate and quality 3D point cloud, Kim [24] proposed a framework for updating the changes in point cloud through crowd sourcing. The experimental result shows that incremental update of point cloud via a non-survey grad Lidar can be expected. In map attributes extraction, Joshi and James [25] proposed a graph-based refinement approach to extract the lane geometry information. The research result showed that the lane information can be obtained on a large scale. Kumar et al. [26] presented an automated algorithm for extracting road markings based on which the intensity can be distinguished from road surfaces. Zhou et al. [27] proposed a method to align and extract pavements, boundaries, lane markings through point cloud data. Niijima et al. [28] further applied road centerline estimation with pose graph optimization to reconstruct the point cloud on a large scale. By adding a more trustable road centerline reference, the error of the point cloud can be minimized. From the above-mentioned MMS research, it can be seen that graph-based optimization methods are beneficial for incorporating with the filter-based localization. For autonomous driving, it is believed the trend would be more concerned about the point cloud reconstruction in complex urban environments.

From the perspective of deep learning, a state-of-the-art architecture PointNet++ [29] was designed to deal with 3D point cloud segmentation and classification tasks. A hierarchical point set feature learning was then applied in the proposed architecture, so that objects in the point cloud can be meticulously classified. On the other hand, image-based object detection and recognition are getting more and more popular on the basis of concrete deep learning architectures. For internet of things (IoT) applications, Bi et al. [30] proposed a distributed PCANN architecture achieving high accuracy with minimized trained model size for image recognition, which is suitable for use on most IoT devices with constrained resource. For traffic sign recognition, Tabernik and Skocaj [31] addressed the problem of detecting and recognizing many traffic sign categories for traffic sign inventory management. Gerhardt and Broll [32] developed a neural network-based traffic sign recognition system for semiautomatic road infrastructure management. Yan et al. [33] demonstrated a network for detecting lane and lane-changing information. The network can segment the lane line instances and get each lane line's color and type attributes. Heo et al. [34] integrated the image detection and HD map

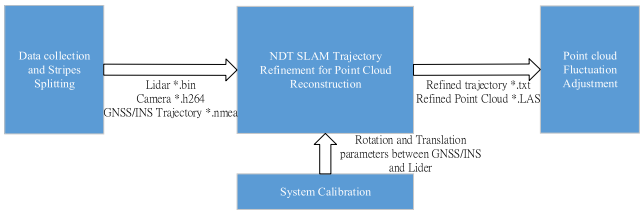


FIGURE 3. Main flow of the proposed enhanced point cloud reconstruction method.

to detect the environment change. Jo et al. [35] presented an algorithm to process, detect, and update the HD map when the environment changes. From the above-mentioned studies, we can see that researchers have been working hard to obtain geodetic information by deep learning-based image detection approaches. However, there is room for further investigation on the accuracy of geodetic coordinates for detected objects in HD map creation.

From the above discussions, the performance of available MMSs has been significantly improved in various aspects over the past years. The image detection using deep learning neural network is also widely utilized in many fields through either open mature pre-trained models or private self-trained models. Taking advantages of precise 3D geodetic point cloud reconstruction and accurate image-based object detection, we can create an accurate HD map in an automatic way. We will provide feasible solutions by using traffic signs as an example to demonstrate the extraction of their precise 3D position for HD map creation.

III. ENHANCED POINT CLOUD RECONSTRUCTION OF THE PROPOSED MMS

In our previous work [22], the MMS mainly integrated a non-survey grade Lidar sensor and non-tactical grade GNSS/INS for accurate point cloud reconstruction. The presented hardware and method can secure a quality point cloud by overcoming the dispersion issue with the aid of ground control points (GCPs) under good NDT SLAM refined trajectory condition. However, the NDT localization might have high uncertainty if the features in the point cloud are not sufficient. Hence, the point cloud reconstruction quality would be unacceptable because of the poorly refined trajectory. It turns out that a graph-based fine-tuning process could be further used for optimization. Figure 3 depicts the main flow of the proposed cost-effective MMS for enhanced point cloud reconstruction. Detailed descriptions will be given in the following subsections.

A. DATA COLLECTION AND STRIPES SPLITTING

Table 1 shows the proposed MMS equipped with non-survey grade sensors. In the data collection phase, the GNSS/INS trajectory consisting of a series positions (x, y, z) and orientations (roll, pitch, heading) is configured in 200Hz via a SPAN-IGM-A1 (NovAtel, Calgary, AB, Canada) GNSS/INS receiver. The virtual reference station - real time kinematics (VRS-RTK) correction data is offered through a 4G mobile phone. A distance measurement instrument (DMI) [36] is installed and connected with the GNSS/INS directly to secure the trajectory is stable under a weak GNSS signal environment.

TABLE 1. The proposed MMS equipment list.

Equipment	Model	Configuration Notes
GNSS/INS	NovAtel SPAN-IGM-A1	VRS-RTK correction data is offered by a 4G mobile phone. The receiver outputs 200 HZ position and orientation trajectory.
Distance Measurement Instrument (DMI)	KISTLER Wheel Pulse Transducers	Connected with GNSS/INS directly to provide up to 5000 pulses per wheel rotation while vehicle is moving.
Computing Platform	NVIDIA Drive PX2	System time is corrected by GNSS/INS pulse per second signal for reference by the sensors.
Lidar Sensor	Velodyne-32E	Runs on 15Hz and only captures points within 20 meters.
Camera Sensor	Sekonix sf3324-10x	Provides 120° FOV and 30 fps with frame timestamp from the computing platform.

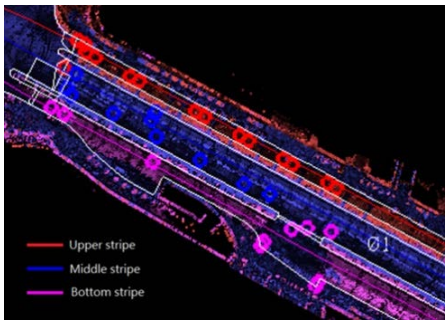


FIGURE 4. Multi-stripe scan performed in a multi-lane road with two refuge islands.

To have a desired trajectory adjustment, redundant observations are required. Thus, a multi-stripe scan is conducted in this paper as exemplified in Figure 4, where the scan is performed in a multi-lane road with two refuge islands. The multi-stripe scan deals with extensive area surveys for further trajectory adjustment, which improves the previous work [22] one step further to better fit in practical situations. Lidar localization refers to the use of static object features for use to deduce the origin position of the Lidar. By applying the transformation between Lidar and GNSS/INS, the GNSS/INS trajectory can be refined and further constrained by precise DMI measurement. It implies that a multi-stripe scan is mandatory in the last step of the proposed data collection phase for both higher accuracy and larger scan coverage. The scanned data of each stripe are subsequently split into several segments for further data processing.

(a) (b)

FIGURE 5. Calibration data collection: (a) A new land consolidation area is selected, (b) MMS calibration data collection by 2 scan missions (blue, purple).

B. SYSTEM CALIBRATION

Given a Lidar scan point $p_{L^{xyz}}$ in the Cartesian Lidar coordinate, the transformation from a Lidar coordinate to a local coordinate system, 2-degree transverse Mercator (TM2), can be expressed as:

$$p_{LTM2} = p_{TM2} + (R_{LTM2} \cdot p_{L^{xyz}}) + T_{LTM2}, \tag{1}$$

where

p_{LTM2} : Lidar point in TM2 coordinate format based on TWD97 System,

p_{TM2} : GNSS/INS trajectory in TM2 coordinate format based on Taiwan Datum 1997 (TWD97) System,

R_{LTM2} : Rotation matrix from Lidar coordinate to TM2 coordinate,

$p_{L^{xyz}}$: Lidar scan point in Lidar coordinate,

T_{LTM2} : Offset between Lidar and GNSS/INS positions.

For the sake of data validation and visual presentation, in this paper, the coordinates of Lidar points p_{LTM2} are converted into TM2 based on TWD97, which is the formal coordinate system used in Taiwan.

From (1), to accurately obtain the elements in the rotation matrix (roll, pitch, and heading) and translation matrix (x, y, and z), we utilize a commercial software TerraMatch [37] for system calibration. In the calibration process, a good GNSS/INS trajectory with a large number of scanned static features is required. According to the calibration guideline, the MMS scan mission is recommended to keep straight to obtain scans of opposite directions to overlap the scanned features. Therefore, the low traffic flow with plenty static object features of new land consolidation area would be beneficial toward a better calibration result as shown in Figure 5.

After MMS calibration data collection, the data collected during a sharp turn, including a U-turn, should be excluded for further processing. Hereafter, each direction scan is divided into individual trajectories. By giving coarse initial rotation and translation parameters to (1), the point clouds in yellow, green, blue, and red can then be reconstructed for each trajectory as illustrated in Figure 6.

The initial coarse rotation and translation parameters inevitably result in certain absolute/relative position errors for points in the point cloud. As a result, the tie line search function and GCP matching in the TerraMatch calibration tool are applied to adjust the initial coarse parameters to

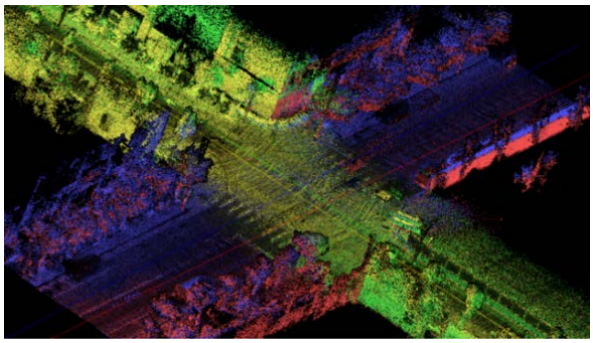


FIGURE 6. Point clouds reconstructed by individual trajectories are indicated in yellow, green, blue, and red.

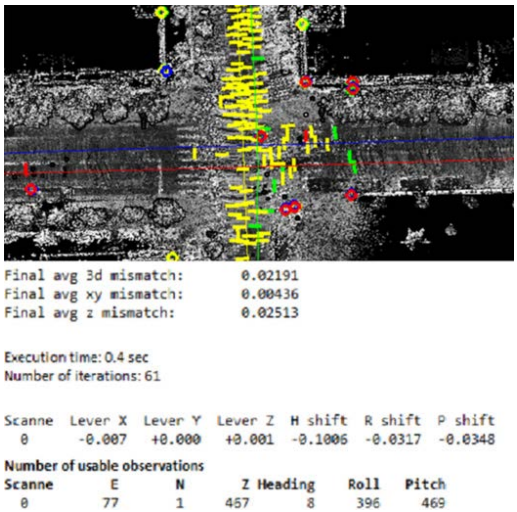


FIGURE 7. Partial view of the multi-stripe point cloud reconstruction by tie line search function (top) and final calibration results (bottom).

obtain the calibrated parameters. Figure 7 shows the partial view of a multi-stripe point cloud reconstruction by tie line search function and final calibration results, where color bars indicate the same features between trajectories are observed and matched, and color circles indicate the GCPs assigned to specific features for absolute position reference. When an incorrect feature is observed from the calibration tool, it is necessary to manually remove the mismatched feature. According to the calibration result, the average 3D mismatch can be reduced to 2.1 cm. We can then apply those adjusted rotation and translation values to the initial coarse parameters for further point cloud reconstruction.

Figure 8 illustrates a point cloud reconstruction by applying the initial coarse parameters and calibrated parameters, respectively. Through comparison of the point clouds of a wall, plate, and pole, it is apparent that the calibrated parameters obtained from the mentioned calibration tool result in a desired point cloud reconstruction.

C. NDT SLAM TRAJECTORY REFINEMENT FOR POINT CLOUD RECONSTRUCTION

It has been found that the point clouds reconstructed by scans of opposite directions based on GNSS/INS trajectory do not

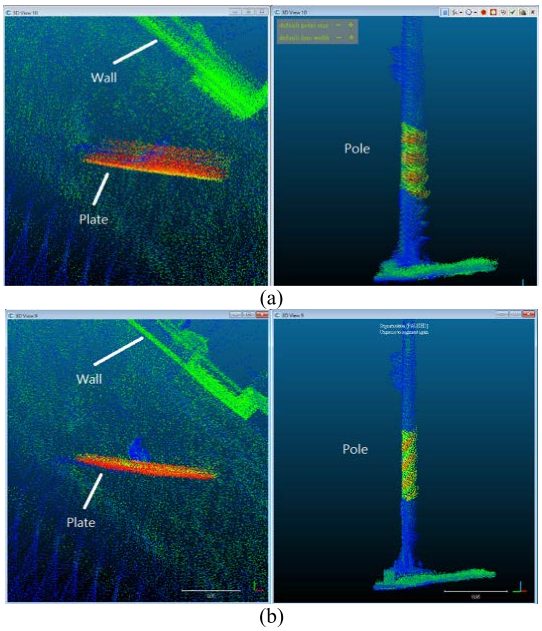


FIGURE 8. Point cloud of a wall, plate, and pole is reconstructed by: (a) applying the initial coarse parameters, (b) applying the calibrated parameters.

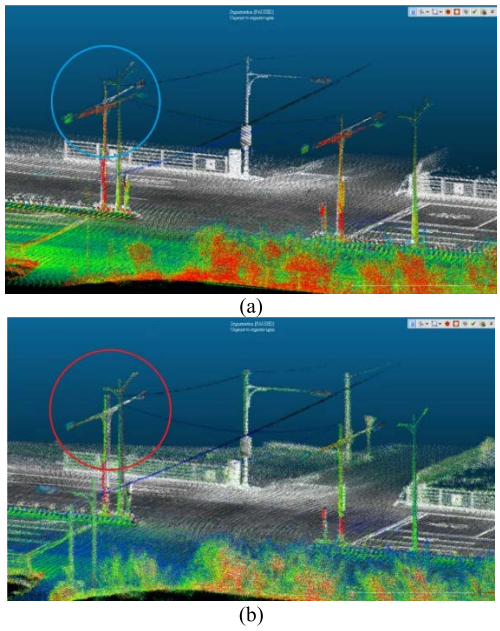


FIGURE 9. Main stripe (white) and sub stripes (color) point cloud reconstruction by (a) applying GNSS/INS trajectories, (b) applying NDT-refined trajectories.

merge well as illustrated in Figure 9(a), where the top portions of a pole indicated by a blue circle are mismatched. This is because the point clouds reconstructed by each direction have their individual uncertain initial attitude of the MMS. This will cause a dispersion problem and incur uncertain object positions to be manually selected for HD map creation.

To address the dispersion problem, an NDT SLAM trajectory refinement is designed in the process. At the beginning of the work, a main stripe is selected for point cloud reconstruction by using GNSS/INS trajectory. The main stripe is selected to cover the entire scan area as comprehensive as possible. Next, GCPs are manually assigned to their corresponding feature points (for example, anchors of dash line or zebra lines on the ground) in the point cloud. Each corresponding feature point is forcibly moved to the GCPs position. The remaining points in the point cloud between the GCPs are then linearly adjusted according to the distance from the GCPs. The shorter the distance between a point and GCP, the more the effect of position adjustment is applied. Afterward, a NDT-refined trajectory of the main stripe can be generated by applying the NDT localization based on the GCPs-adjusted point cloud. The refined trajectory can be further obtained as:

$$\begin{aligned}&\overrightarrow{P_{RFTM2}}(t_{NDT}+t_{offset})\\&\quad=\overrightarrow{P_{TM2}}(t_{NDT}+t_{offset})+D_{(x,y,z)}(t_{NDT}),\\&\qquad if\ \left\|D_{(x,y,z)}(t_{NDT})\right\|>threshold,\end{aligned}\tag{2}$$

where

$\overrightarrow{P_{RFTM2}}$: NDT-refined trajectory,

$\overrightarrow{P_{TM2}}$: GNSS/INS trajectory at 200 Hz,

$D_{(x,y,z)}$: Distance difference between NDT trajectory and GNSS/INS trajectory at matched timestamp,

t_{NDT} : Timestamp of NDT localization points,

t_{offset} : Time offset since last position has been obtained.

In this paper, the threshold in (2) is set as 10cm. When the absolute value of $D_{(x,y,z)}$ is greater than the threshold, the current and successive points of the GNSS/INS trajectory in $\overrightarrow{P_{TM2}}$ are adjusted until the subsequent NDT localization is performed. After successfully obtaining the refined trajectory of the main stripe, the main-stripe point cloud can then be reconstructed by replacing $\overrightarrow{P_{TM2}}$ with $\overrightarrow{P_{RFTM2}}$ through (1). Afterward, the refined point cloud of the main stripe serves as the base reference for all sub-stripe trajectories refinement for doing sub-stripe point cloud reconstruction. Figure 9(b) shows that the point cloud reconstruction result is well aligned between the main stripe and sub-stripe. For example, the top portions of the pole are well merged as indicated by a red circle.

D. POINT CLOUD FLUCTUATION ADJUSTMENT

It is clear that NDT trajectory refinement has a significant advantage in dealing with multi-stripe point cloud reconstruction. However, the accuracy of NDT localization results cannot be secured if there are unclear scanned features or there is a lack of static features. In those cases, a graph-based fluctuation adjustment function in the TerraMatch can be employed for final adjustment between stripes. In Figure 10, 3 scan stripes (blue, red, and pink) by MMS of NDT-refined trajectories and their corresponding point clouds are loaded into the tool, where the main stripe is indicated in blue (Line 1) and sub stripes (Line 3, Line 5) are indicated in red and pink, respectively.

Figure 11(a) shows that Line 1 is adjusted with Line 3 and Line 5. In the fluctuation adjustment result (Figure 11(b)), the

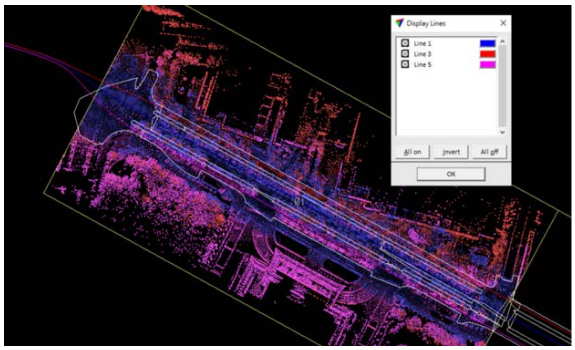
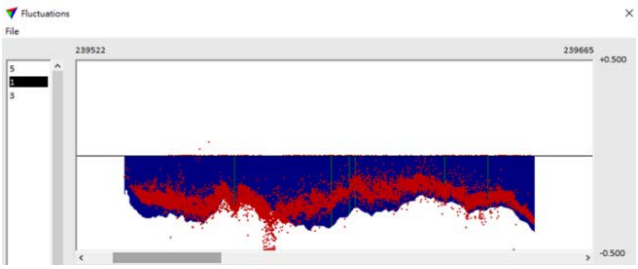
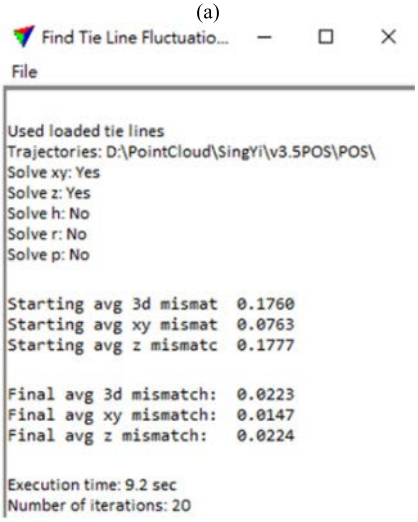


FIGURE 10. 3 scan stripes (blue, red, and pink) by MMS of NDT-refined trajectories and their corresponding point clouds are loaded into the tool.



(a)



(b)

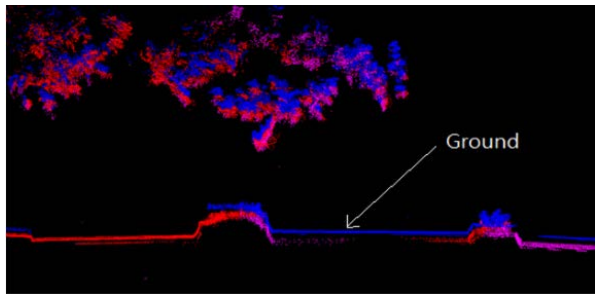
FIGURE 11. Fluctuation adjustment: (a) Line 1 is adjusted with Line 3 and Line 5, where x-axis indicates the timestamp and y-axis indicates the 3D distance adjustment in meter; (b) fluctuation adjustment report.

3D mismatch is improved from 0.176m to 0.0223m. Through the vertical view of the point cloud in Figure 12(a), it can be seen the gaps between Line 1 (Blue) and Line 3 (Red) point clouds have been improved by the fluctuation adjustment in addition to the NDT trajectory refinement (Figure 12(b)), as indicated by a white solid line pointing to the ground.

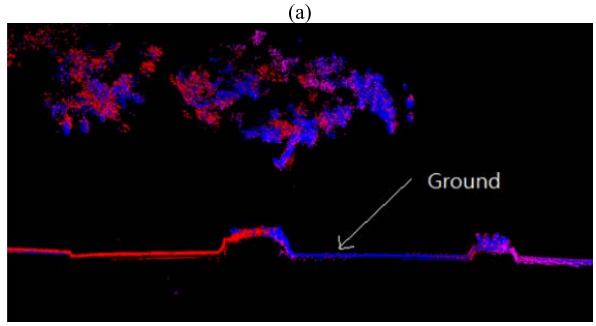
IV. EXTRACTION OF CENTER GEODETIC COORDINATES FOR TRAFFIC SIGNS FROM POINT CLOUD

A. TRAFFIC SIGN DETECTION

For image detection, many mature frameworks and pre-trained models have been developed and further improved



(a)



(b)

FIGURE 12. Vertical view of Line 1 (blue) point cloud, Line 3 (red) point cloud and Line 5 (pink) point cloud: (a) before fluctuation adjustment, (b) after fluctuation adjustment.

till now. The NVIDIA DriveWorks SDK [38] was designed for developers to implement autonomous vehicle solutions by providing a library of models, developer tools, and reference applications. Since the DriveWorks SDK is mainly dealing with road related information, therefore, the perception module of DriveNet in DriveWorks is suitable to serve as the detection core for traffic sign detection in the proposed work. The DriveNet is trained by front and side camera images, where the front camera supports a 30/60/120 degree field of view. The detected objects are represented by a bounding box defined by the top-left and bottom-right corners denoted as (u,v) pixels for 5 different classes, including car, bicycle, pedestrian, road sign, and traffic light.

Through the SDK reference documents and sample codes, a utility is developed to output bounding box position, object class, matched timestamp into a json file. According to the test in the local fields as shown in Figure 13, the detection result is acceptable and true negative detection of the road signs rarely happens, which fulfills the goals of leveraging feasible and straightforward concepts for local traffic sign detection.

B. GEODETIC TRANSFORMATION BETWEEN CAMERA PIXEL AND LIDAR POINT

Each Lidar point $p_{L^{xyz}}$ can be transformed to the local geodetic coordinate p_{LTM2} by (1). The next step is to figure out the transformation between Lidar points and camera pixels. According to the pinhole camera model, the approximated projection model can be expressed as:

$$s\begin{bmatrix} u \\ v \end{bmatrix} = A_C\,[R_{LC}|T_{LC}]\begin{bmatrix} X_L \\ Y_L \\ Z_L \end{bmatrix} \tag{3}$$

FIGURE 13. Object detection by DriveNet neural network using the proposed MMS camera in Table 1.

FIGURE 14. Lidar scan points (colored dots) mapped to their corresponding pixels of the detected traffic sign by applying rotation and translation parameters.

where

s: Scale factor between pixels represented in the real world,

u, v : 2 D position in the image,

A_C : Camera intrinsic parameter matrix,

R_{LC} : Rotation matrix from Lidar to camera,

T_{LC} : Translation matrix from Lidar to camera,

X_L, Y_L, Z_L: Point in Lidar coordinate.

Note that A_C contains focal length and sensor center parameters of a camera, which can be obtained by intrinsic parameter calibration through OPENCV. The rotation matrix R_{LC} and translation matrix T_{LC} between Lidar and camera can be derived through Calibration Toolkit [39] by manually clicking the center of the calibration board in various angles and distances. The normal vectors of anchors on the calibration board in the Lidar coordinate and camera coordinate can be used to deduce the relationship of the transformations. Thus, the scale factor s can be found, given X_L, Y_L, Z_L, and u, v during the calibration process.

After the translation and rotation parameters are applied in (3), the Lidar scan points can be well mapped to their corresponding pixels in the image. Figure 14 shows the overlapped result of Lidar scan points and detected traffic sign, where the colored dots indicate the Lidar points in the nearest (Light blue), near (Green), middle (Yellow), far (Orange), and the farthest (Red). The red bounding boxes indicate the objects detected by the utility. Since the Lidar points in the bounding boxes for each traffic sign are well distinguished, therefore the feasibility of the calibrated parameters for geodetic transformation between camera pixel and Liar point is confirmed.

C. CENTER GEODETIC COORDINATES OF TRAFFIC SIGNS

The Lidar point set of a traffic sign can be extracted and expressed by:

$$P_{TS} = P_{L|u,v;i} \cap P_{u,v}|_{i>TH;u,v\in Bounding\ box}, \tag{4}$$

where

$P_{L|u,v;i}$: Lidar points mapped to the image, i is the intensity of each Lidar point,

$\cap$: Intersection operator,

$P_{u,v}$: Pixels of the detected traffic sign within the bounding box,

An intensity threshold is set to eliminate the background points scanned by Lidar, where intensity $i < TH$. Next, the current center $C_{current-center}$ in the local geodetic location can be calculated by:

$$C_{current-center} = (avg(P_{TS-X}), avg(P_{TS-Y}), avg(P_{TS-Z})), \tag{5}$$

where

$C_{current-center}$: Average center of Lidar point projected on the current image,

$P_{TS-X}, P_{TS-Y}, P_{TS-Z}$: Lidar point set of a detected traffic sign in geodetic local coordinate.

In order to obtain a more accurate $C_{current-center}$, a series of P_{TS} in each image are extracted by a moving MMS then accumulated to calculate the average center geodetic coordinates for the detected traffic sign. When there exists an object center $C_{exist-center}$ within 30cm from $C_{current-center}$, the current P_{TS} will merge into existed P_{TS} to calculate a new $C_{exist-center}$. Otherwise, a new ID will be registered to each point in current P_{TS} (Figure 15). Finally, according to each ID, the width and height dimensions set by local traffic authority regulations can be applied to every $C_{exist-center}$ of the detected traffic signs.

This section describes the transformation between Lidar point and camera pixel, where the rotation matrix and translation matrix are obtained by utilizing both OPENCV and Calibration Tool. According to the Lidar to image overlapped verification, each Lidar point can be well mapped to the detected bounding box. In accumulating the scan points distinguished by object ID, the traffic sign center coordinates can be finally determined.

V. EXPERIMENTAL RESULTS

Experiments are conducted in the Dan-Shui area, New Taipei City, Taiwan, which provides an open street scene for autonomous shuttle service operation test. The shuttle service operates between Kan-Ding light rail station to a mall nearby (600m) to carry out the last-mile transportation concept of mobility as a service vision in Taiwan. Figure 16 shows the HD map incorporating point cloud and vector data produced via the proposed method in this paper

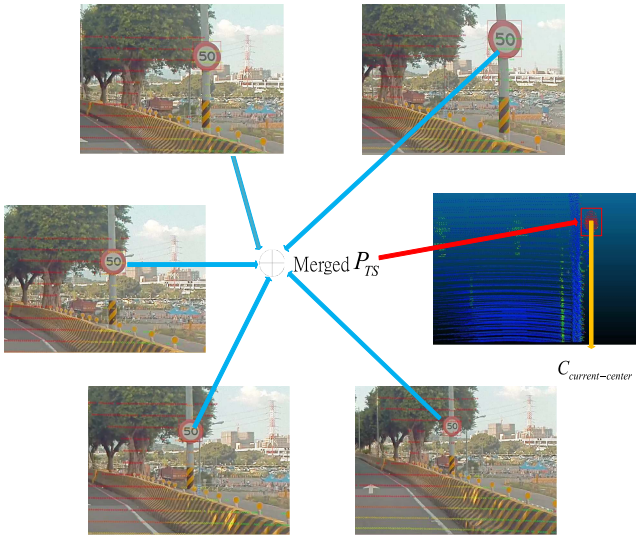


FIGURE 15. Series of P_{TS} in each image are extracted by a moving MMS then accumulated to calculate the average center geodetic coordinates for the detected traffic sign.

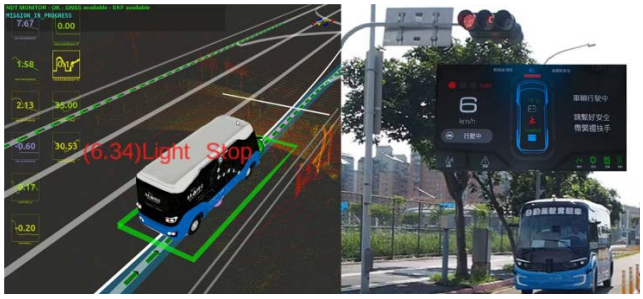


FIGURE 16. HD map produced by the proposed method for use in an autonomous shuttle service.

for use by the autonomous shuttle service. During the test period, real-time traffic light information is provided and appropriately linked with vector information in the HD map under local government permission.

To make a fair comparison, the total distance of the road survey in the experiment has been enlarged to 2.5km long. Because of the large area considered, 33 control points and 20 check points were measured for HD map correction and verification, respectively. Figure 17 shows the top view of the point cloud reconstructed by the proposed method in the test area. We conducted extensive measurements to evaluate the thickness, absolute position, and extracted traffic sign center position, respectively.

A. QUALITY ASSESSMENT OF THE RECONSTRUCTED POINT CLOUD

For point cloud quality assessment, the thickness of each check point location evenly spread in the test area is measured. Thinner measurement result is desired because it implies higher accuracy of relative position for further vector map depiction. Figure 18 shows the thickness measurement result. It can be seen that the quality of the reconstructed

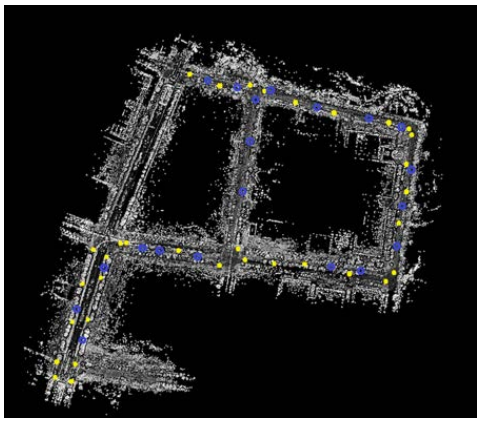


FIGURE 17. Top view of point cloud reconstruction in the test area, where check points and control points are indicated in blue and yellow, respectively.

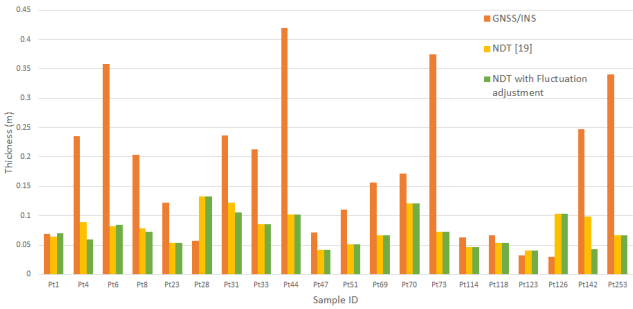


FIGURE 18. Point cloud thickness assessment in the test area.

point cloud has been gradually improved from the GNSS/INS method (Avg. 17.875 cm), NDT method [22] (Avg. 7.85 cm) to the newly proposed NTD with fluctuation adjustment method (Avg. 7.35 cm). From the newly proposed method, in addition to the improved average thickness, the fluctuation adjustment is added and aimed to find the optimal result of an area. By doing so, the existing inaccurate GCPs measurement can be leveraged to guarantee that the previous work [22] can be performed in realistic open traffic flow scenes.

For absolute position assessment, a total of 20 check points have been measured and compared to their position in the point cloud as shown in Figure 19, where the XY plane and 3D root mean square errors (RMSE) are considered in the evaluation. According to the evaluation results, we can see that with the aid of sufficiently precise control points, the 2D average absolute position error can be limited to 4.7 cm, and 3D average absolute position error can be controlled within 8.6 cm by the proposed method.

B. TRAFFIC SIGN CENTER POSITION ASSESSMENT

Figure 20 shows the measurement of the traffic sign center positions by a human operator and the proposed automatic extraction method. A total 23 desired traffic signs have been taken into account in the experiment. Due to the limited vertical view angle of Lidar, we omitted the false positive detection results and low-density Lidar points of each traffic sign in our assessment. By calculating the difference between

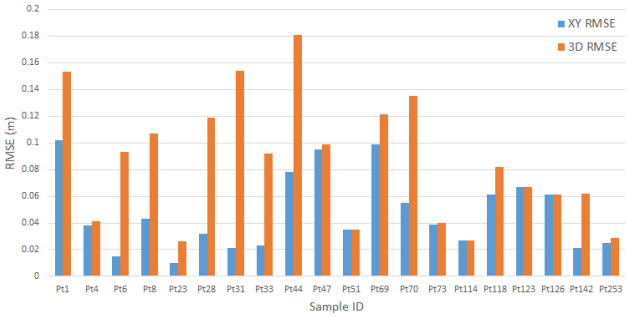


FIGURE 19. Point cloud absolute position assessment in the test area.

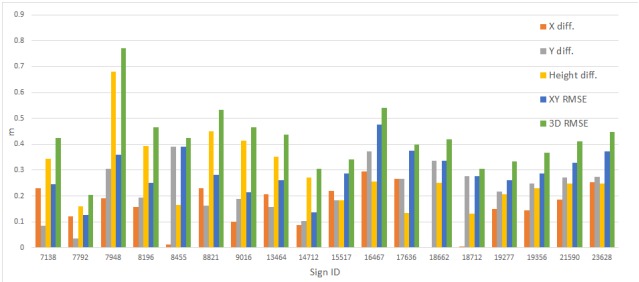


FIGURE 20. Center position comparison of traffic sign extraction.

each center position obtained by a human operator and the proposed automatic extraction method, the average differences of traffic sign positions in terms of XY RMSE and 3D RMSE are 29 cm and 42cm, respectively. The result shows the proposed method is able to address the time exhausting issue in HD map creation and would further benefit HD map updating.

In this section, we have presented the performance of the proposed method in point cloud reconstruction and position accuracy of automatic traffic sign extraction. Through the thickness assessment, we've shown the proposed method benefits the reconstruction of point cloud in comparison to GNSS/INS approach and further reduces the error by NDT and fluctuation adjustment solution. The produced point cloud absolute position is secured with high accuracy in few centimeters; hence the center position of each traffic sign can be extracted in sub-meter accuracy level to reduce labor work for HD map creation.

VI. CONCLUSION

In this paper, we have presented an MMS with an enhanced workflow to further improve the accuracy of point clouds reconstructed by the previously revealed MMS and deal with a large scan area for practical applications. To solve the dispersion problem of the point cloud, a multi-stripe scan with NDT refinement and fluctuation adjustment have been proposed in the point cloud reconstruction process with the aid of GCPs. According to the experimental results, it can be seen that the point cloud accuracy is secured in open wide road scenes by the cost-effective MMS configuration proposed in this paper. Furthermore, the proposed automatic object extraction based on the reconstructed point cloud with centimeter accuracy has successfully compensated the lack of depth information via traditional image object detection. This work proves the concept of combining image object detection with accurate point cloud is feasible for efficient 3D object extraction to fulfill the goal of significantly reducing labor expenses in HD map creation. It is worth noting that the designed post-processing workflow allows an easy integration of various detection models. Any state-of-the-art models (i.e., YOLO or SSD) for object detection can be easily and straightforwardly incorporated into the proposed framework to further enhance the detection accuracy. We are now heading into an integration stage to determine object detection models which are suitable for use in geodetic coordinates extraction for complex road scenarios. In the future, we will investigate a challenging topic to automatically link the relationship among roads, lanes, and 3D objects to achieve an even more efficient process in HD map creation.

ACKNOWLEDGMENT

The authors are grateful for all assistance of field works to Kingwaytek Technology Co., Ltd.

REFERENCES

[1] *The Lanelet2 Map Format Open Library*. Accessed: Jun. 18, 2022. [Online]. Available: https://github.com/fzi-forschungszentrum-informatik/Lanelet2

[2] *The OPENDRIVE Map Format*. Accessed: Jun. 2022. [Online]. Available: https://www.asam.net/standards/detail/opendrive/

[3] *The Autoware Foundation*. Accessed: Jun. 18, 2022. [Online]. Available: https://www.autoware.org/

[4] *Simulator CarSim*. Accessed: Jun. 18, 2022. [Online]. Available: https://www.carsim.com/products/carsim/

[5] *VIRES Simulation Technologie GmbH VTD*. Accessed: Jun. 18, 2022. [Online]. Available: https://vires.mscsoftware.com/

[6] K. Wong, Y. Gu, and S. Kamijo, "Mapping for autonomous driving: Opportunities and challenges," *IEEE Intell. Transp. Syst. Mag.*, vol. 13, no. 1, pp. 91–106, Spring 2021.

[7] *ESRI GIS Tool ArcGIS*. Accessed: Jun. 18, 2022. [Online]. Available: https://www.arcgis.com/index.html

[8] *Bentley Systems GIS Tool MicroStation*. Accessed: Jun. 18, 2022. [Online]. Available: https://www.bentley.com/en/products/brands/microstation

[9] M. Magnusson, A. Lilienthal, and T. Duckett, "Scan registration for autonomous mining vehicles using 3D-NDT," *J. Field Robot.*, vol. 24, no. 10, pp. 803–827, Oct. 2007.

[10] M. A. Haq, "CDLSTM: A novel model for climate change forecasting," *Comput., Mater. Continua*, vol. 71, no. 2, pp. 2363–2381, 2022.

[11] A. Attaallah and R. A. Khan, "SMOTEDNN: A novel model for air pollution forecasting and AQI classification," *Comput., Mater. Continua*, vol. 71, no. 1, pp. 1403–1425, 2022.

[12] Y. Yue, M. Gouda, and K. El-Basyouny, "Automatic detection and mapping of highway guardrails from mobile lidar point clouds," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2021, pp. 2520–2523, doi: 10.1109/IGARSS47720.2021.9553055.

[13] J. Jiao, "Machine learning assisted high-definition map creation," in *Proc. IEEE 42nd Annu. Comput. Softw. Appl. Conf. (COMPSAC)*, Tokyo, Japan, Jul. 2018, pp. 23–27.

[14] F. Poggenhans, J.-H. Pauls, J. Janosovits, S. Orf, M. Naumann, F. Kuhnt, and M. Mayr, "Lanelet2: A high-definition map framework for the future of automated driving," in *Proc. 21st Int. Conf. Intell. Transp. Syst. (ITSC)*, Maui, HI, USA, Nov. 2018, pp. 4–7.

[15] D. Becker, F. Rus, C. Geller, and L. Eckstein, "Generation of complex road networks using a simplified logical description for the validation of automated vehicles," in *Proc. IEEE 23rd Int. Conf. Intell. Transp. Syst. (ITSC)*, Rhodes, Greece, Sep. 2020, pp. 20–23.

[16] M. Althoff, S. Urban, and M. Koschi, "Automatic conversion of road networks from OpenDRIVE to lanelets," in *Proc. IEEE Int. Conf. Service Oper. Logistics, Informat. (SOLI)*, Singapore, Jul. 2018, pp. 157–162.

[17] E. Javanmardi, M. Javanmardi, Y. Gu, and S. Kamijo, "Factors to evaluate capability of map for vehicle localization," *IEEE Access*, vol. 6, pp. 49850–49867, 2018.

[18] L.-B. Yao, H.-B. Wu, Y.-Y. Li, B. Meng, J.-F. Qian, C. Liu, and H.-C. Fan, "Registration of vehicle-borne point clouds and panoramic images based on sensor constellations," *Sensors*, vol. 17, no. 4, p. 837, Apr. 2017.

[19] F.-Y. Zeng and R.-F. Zhong, "The algorithm to generate color point-cloud with the registration between panoramic image and laser point-cloud," in *Proc. 35th Int. Symp. Remote Sens. Environ. (ISRSE)*, Beijing, China, vol. 17, Apr. 2013, pp. 22–26.

[20] V. Ilci and C. Toth, "High definition 3D map creation using GNSS/IMU/LiDAR sensor integration to support autonomous vehicle navigation," *Sensors*, vol. 20, no. 3, p. 899, Feb. 2020.

[21] I. Hamieh, R. Myers, and T. Rahman, "Construction of autonomous driving maps employing LiDAR odometry," in *Proc. IEEE Can. Conf. Electr. Comput. Eng. (CCECE)*, Edmonton, AB, Canada, May 2019, pp. 5–8.

[22] C.-W. Peng, C.-C. Hsu, and W.-Y. Wang, "Cost effective mobile mapping system for color point cloud reconstruction," *Sensors*, vol. 20, no. 22, p. 6536, Nov. 2020.

[23] K. Koide, J. Miura, M. Yokozuka, S. Oishi, and A. Banno, "Interactive 3D graph SLAM for map correction," *IEEE Robot. Autom. Lett.*, vol. 6, no. 1, pp. 40–47, Jan. 2021.

[24] C. Kim, S. Cho, M. Sunwoo, P. Resende, B. Bradai, and K. Jo, "Updating point cloud layer of high definition (HD) map based on crowd-sourcing of multiple vehicles installed LiDAR," *IEEE Access*, vol. 9, pp. 8028–8046, 2021.

[25] A. Joshi and M. R. James, "Generation of accurate lane-level maps from coarse prior maps and lidar," *IEEE Intell. Transp. Syst. Mag.*, vol. 7, no. 1, pp. 19–29, Spring 2015.

[26] P. Kumar, C. P. McElhinney, P. Lewis, and T. McCarthy, "Automated road markings extraction from mobile laser scanning data," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 32, pp. 125–137, Oct. 2014.

[27] Y. Zhou, R. Huang, T. Jiang, Z. Dong, and B. Yang, "Highway alignments extraction and 3D modeling from airborne laser scanning point clouds," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 102, Oct. 2021, Art. no. 102429.

[28] S. Niijima, J. Nitta, Y. Sasaki, and H. Mizoguchi, "Generating 3D fundamental map by large-scale SLAM and graph-based optimization focused on road center line," *Proc. Int. Symp. Robot Hum. Interact. Commun. (RO-MAN)*, Lisbon, Portugal, Sep. 2017, pp. 1188–1193.

[29] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "PointNet++: Deep hierarchical feature learning on point sets in a metric space," in *Proc. Conf. Neural Inf. Process. Syst. (NIPS)*, San Francisco, CA, USA, Dec. 2017, pp. 5100–5109.

[30] T. Bi, Q. Liu, T. Ozcelebi, D. Jarnikov, and D. Sekulovski, "PCANN: Distributed ANN architecture for image recognition in resource-constrained IoT devices," in *Proc. 15th Int. Conf. Intell. Environments (IE)*, Jun. 2019, pp. 1–8, doi: 10.1109/IE.2019.000-3.

[31] D. Tabernik and D. Skocaj, "Deep learning for large-scale traffic-sign detection and recognition," *IEEE Trans. Intell. Transp. Syst.*, vol. 21, no. 4, pp. 1427–1440, Apr. 2020.

[32] C. Gerhardt and W. Broll, "Neural network-based traffic sign recognition in 360° images for semi-automatic road maintenance inventory," in *Proc. IEEE 23rd Int. Conf. Intell. Transp. Syst. (ITSC)*, Rhodes, Greece, Sep. 2020, pp. 20–23.

[33] C. Yan, C. Zheng, C. Gao, W. Yu, Y. Cai, and C. Ma, "Lane information perception network for HD maps," in *Proc. IEEE 23rd Int. Conf. Intell. Transp. Syst. (ITSC)*, Rhodes, Greece, Sep. 2020, pp. 20–23.

[34] M. Heo, J. Kim, and S. Kim, "HD map change detection with cross-domain deep metric learning," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Las Vegas, NV, USA, Oct. 2020, pp. 10218–10224.

[35] K. Jo, C. Kim, and M. Sunwoo, "Simultaneous localization and map change update for the high definition map-based autonomous driving car," *Sensors*, vol. 18, no. 9, p. 3145, Sep. 2018.

[36] *The Kistler Distance Measurement Instrument*. Accessed: Jun. 18, 2022. [Online]. Available: https://www.kistler.com/en/product/type-cwpta/

[37] *Calibration Tool: TerraMatch*. Accessed: Jun. 18, 2022. [Online]. Available: https://terrasolid.com/

[38] *NVIDIA DriveWorks SDK*. Accessed: Jun. 18, 2022. [Online]. Available: https://developer.nvidia.com/drive/driveworks

[39] *CalibrationToolkit Instruction Video*. Accessed: Jun. 18, 2022. [Online]. Available: https://www.youtube.com/watch?v=pfBmfgHf6zg

CHENG-WEI PENG received the B.S. degree in electronic engineering from St. Johns University, Taipei, Taiwan, in 2002, and the master's degree in communication and navigation control engineering from the National Taiwan Ocean University, Keelung, Taiwan, in 2004. He is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, National Taiwan Normal University, Taipei. He has been studying GNSS/INS, geographic information systems, and telematics platform for years. His expertise includes sensor integration of mobile mapping systems and utilizing map data to develop location-based services. His research interests include mobile mapping systems, HD map creation, and autonomous shuttle service development.

CHEN-CHIEN HSU (Senior Member, IEEE) was born in Hsinchu, Taiwan. He received the B.S. degree in electronic engineering from the National Taiwan University of Science and Technology, Taipei, Taiwan, in 1987, the M.S. degree in control engineering from the National Chiao-Tung University, Hsinchu, in 1989, and the Ph.D. degree from the School of Microelectronic Engineering, Griffith University, Brisbane, Australia, in 1997.

He was a Systems Engineer at IBM Corporation, Taipei, for three years, where he was responsible for information systems planning and application development, before commencing the Ph.D. studies. He joined the Department of Electronic Engineering, St. Johns University, Taipei, as an Assistant Professor, in 1997, and appointed as an Associate Professor, in 2004. From 2006 to 2009, he was at the Department of Electrical Engineering, Tamkang University, Taipei. He is currently a Professor with the Department of Electrical Engineering, National Taiwan Normal University, Taipei. He is the author or coauthor of more than 180 refereed journal and conference papers. His current research interests include digital control systems, evolutionary computation, vision-based measuring systems, sensor applications, and mobile robot navigation. He is an IET Fellow.

WEI-YEN WANG (Fellow, IEEE) received the Diploma degree in electrical engineering from the National Taipei Institute of Technology, in 1984, and the M.S. and Ph.D. degrees in electrical engineering from the National Taiwan University of Science and Technology, Taipei, Taiwan, in 1990 and 1994, respectively.

From 1990 to 2006, he worked concurrently as a Patent Screening Member of the National Intellectual Property Office, Ministry of Economic Affairs, Taiwan. Since 2003, he has been certified as a Patent Attorney in Taiwan. In 1994, he was appointed as an Associate Professor at the Department of Electronic Engineering, St. John's and St. Mary's Institute of Technology, Taiwan. From 1998 to 2000, he worked at the Department of Business Mathematics, Soochow University, Taiwan. From 2000 to 2004, he was at the Department of Electronic Engineering, Fu-Jen Catholic University, Taiwan. In 2004, he became a Full Professor at the Department of Electronic Engineering, Fu-Jen Catholic University. In 2006, he was a Professor and the Director of the Computer Center, National Taipei University of Technology, Taiwan. From 2007 to 2014, he was a Professor at the Department of Applied Electronics Technology, National Taiwan Normal University, Taiwan. From 2011 to 2013, he was the Director of the Information Technology Center, National Taiwan Normal University. He is currently a Professor with the Department of Electrical Engineering, National Taiwan Normal University. His current research interests and publications include the areas of fuzzy logic control, robust adaptive control, neural networks, computer-aided design, digital control, and CCD camera based sensors. He has authored or coauthored over 200 refereed conference and journal papers in the above areas.

Dr. Wang is a fellow of IET, AAIA, CACS, RST, and TFSA. He was a recipient of the Best Associate Editor Award of IEEE Transactions on Cybernetics. He is currently serving as an Associate Editor of the IEEE Transactions on Cybernetics and the Editor-in-Chief of the *International Journal of Fuzzy Systems*.