Semi-automated approach towards efficient HD Maps generation and verification with Lanelet2 formats

Yi-Feng Chang [1], Yen-En Huang [2], Meng-Lun Tsai [3], Hatem Darweesh [4], Kai-Wei Chiang [5], Mengchi Ai [6], Naser El-Sheimy [7]

[1] Department of Geomatics, National Cheng Kung University, Tainan, Taiwan (R.O.C.) - f64066185@gs.ncku.edu.tw
[2] Department of Geomatics, National Cheng Kung University, Tainan, Taiwan (R.O.C.) - jimmy4345@gmail.com
[3] Department of Geomatics, National Cheng Kung University, Tainan, Taiwan (R.O.C.) - taurusbryant@geomatics.ncku.edu.tw
[4] Nagoya University, Furo-cho, Chikusa-ku, Nagoya, 464-8601, Japan - hatem.darweesh@zatitech.com
[5] Department of Geomatics, National Cheng Kung University, Tainan, Taiwan (R.O.C.) - kwchiang@geomatics.ncku.edu.tw
[6] Department of Geomatics Engineering, University of Calgary, Calgary Alberta T2N 1N4, Canada - mengchi.ai@ucalgary.ca
[7] Department of Geomatics Engineering, University of Calgary, Calgary Alberta T2N 1N4, Canada - elsheimy@ucalgary.ca

Keywords: High-Definition Maps, Mobile Laser Scanning, Autonomous Driving, Road Surface Markings, OpenDRIVE, Lanelet2.

Abstract

HD Maps (High-Definition Maps) serve as crucial resources for the domain of autonomous vehicle. Because HD Maps can provide detailed and accurate road information, the generation of HD Maps has been a labour-intensive and high cost. This research presents an innovative and semi-automated approach for efficient HD Maps generation by using assure mapping tool with deep learning techniques and mobile laser scanned point cloud geometry. The proposed method starts with data collection from various sources such as images, LiDAR point clouds, and integrated INS/GNSS trajectory data. These data are labelled by using a pre-trained model. After finishing post-labelling, these data are subjected to deep learning training by using VoxelNet and Yolact++ framework and leading to the generation of an AI model. The tool effectively recognizes and categorizes features such as road surface markings, traffic signs, and traffic lights, which can be further expanded as per requirements. Finally, the output format can be converted to OpenDRIVE, Lanelet2, and other else. Hence, the extracted lane lines can compare to the manual mapping data for verifying the accuracy. This study demonstrates that the proposed approach can be instrumental in streamlining the HD Maps generation procedure, reducing manual labour, and enhancing efficiency. The assure mapping tool proves to be an effective instrument, particularly when powered by deep learning algorithms and point cloud geometries, in the creation of reliable, comprehensive, and application-ready HD Maps.

1. Introduction

With the development of autonomous vehicle technology, HD Maps (High-Definition Maps) are gaining considerable relevance for various applications such as intelligent transportation systems, autonomous driving and route optimization (Liu et al., 2019). In response to the demand for high-resolution mapping solutions, many companies have invested in HD Maps creation. For example, TomTom has pioneered the RoadDNA technology, a constituent of the TomTom HD Maps, which provides the precise positioning and robustness required for autonomous driving. RoadDNA achieves accurate lateral and longitudinal location data by processing sensor input in real time (TomTom, 2018). Similarly, HERE offers the HD Live Maps, a continuously updated cloud-based service encompassing numerous map layers. This map includes a wealth of information such as road topology, geometry of the road centerline, road-level attributes, lane topology data, lane-level attributes, and localization strategies (HERE, 2017).

However, there are significant challenges by using traditional methods of HD Maps creation because of their labour-intensive and time-consuming nature, struggling to fulfil large-scale data production needs. In the creation process, the MLS (Mobile Laser Scanning) system has the ability to accumulate high-precision environmental data quickly. Despite this, subsequent stages such as map construction, digitization, and additional mapping tasks largely remain manual processes that consume considerable time and manpower (Mi et al., 2021). Consequently, cost reduction in HD Maps production emerges as an essential topic in the furtherance of autonomous vehicle technology (Chiang et al., 2022).

To address this common issue, some studies have tried to use deep learning to extract road objects (Elhousni et al., 2020). Some methods in the literature use road surface extraction method based on curb structures using the height differences between sidewalks and roads to find the lane and further obtain the road surface (Yang et al., 2013). There are also research methods that use the intensity differences between road and sidewalk point clouds to extract road edge information (Zeng, 2020). In summary, this study presents a semi-automated tool designed for the production of HD Maps. By leveraging MLS point clouds and machine learning algorithms, this tool aims to expedite the process of HD Maps generation and editing. First, this study outlines the central technologies and principles underpinning the automated HD Maps production tool; following this, the extraction accuracy of map objects and the outcomes of map production will be evaluated. Finally, this study delves into the strengths and limitations of tool and discuss its potential future directions.

2. Methodology

2.1 Extraction architecture of assure mapping tools

Figure 1 illustrates the structure employed when creating HD Maps, mainly divided into five iterative steps. Initially, the process starts with data collection, where the gathered information comprises images, LiDAR (Light Detection And Ranging) point clouds, trajectory data, and more. To streamline the process, a pre-set model is applied to automatically label the collected data, thereby accelerating the labelling speed. Human intervention is then introduced to refine the results of the automated labelling, ensuring that the labels are as precise and accurate as possible. Once the labelling phase is complete, the data is introduced to AI (Artificial Intelligence) for deep learning, resulting in the formation of an AI model. This model serves as a crucial tool in the process of map creation. The concluding step utilizes the newly trained AI model to conduct an automated extraction on the test data, effectively evaluating the performance of the AI model's extraction capabilities. The insights from this evaluation phase are harnessed to refine the model, driving continuous improvements in the accuracy and precision of the maps produced. This cyclical process is an integral part of creating high-precision maps. Each step is intricately connected, and improvements in one phase directly contribute to enhancing the entire process, thus refining the final output, i.e., the HD Maps.

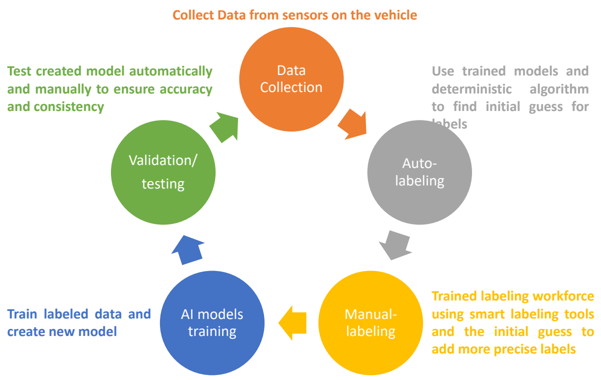

Figure 1. HD Maps generation process.

As demonstrated in Figure 2, the annotation stage begins by uploading both the LiDAR point cloud and image data to the annotation tool server developed by Dr.Hatem Darweesh. Each dataset is separately annotated, and appropriate categories are assigned to the labelled data. Post annotation, the point cloud dataset is trained with VoxelNet, while the image dataset is trained using Yolact++.

The utilization of these two advanced deep learning models allows for the effective extraction of features from both the point cloud and image data. The use of VoxelNet for the point cloud data takes advantage of its ability to manage 3D data, and its capability of detecting objects in three dimensions. Similarly, the employment of Yolact++ for image data allows for high-speed instance segmentation, making it an excellent tool for handling 2D image data.

Following the deep learning training phase, the resultant pretrained models can be utilized to extract data from other unlabelled datasets. These pretrained models hold considerable value as they possess learned features that can be leveraged to identify and categorize similar instances in new, unseen data. This capability significantly expedites the extraction process, contributing to the efficiency and accuracy of the HD Maps creation process. The interface of the 2D annotation tool for images is illustrated in Figure 3, while the interface of the 3D annotation tool for point clouds is illustrated in Figure 4.

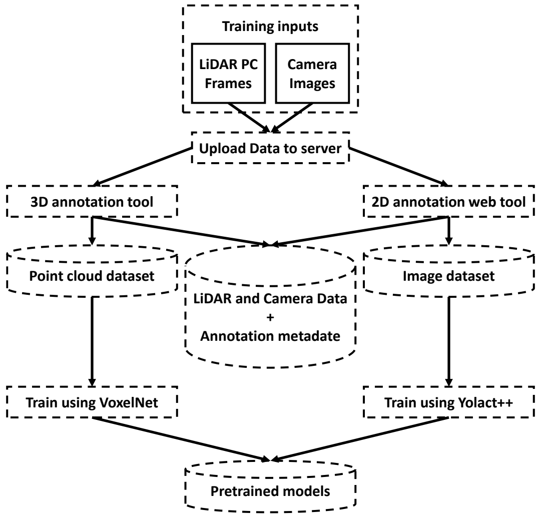

Figure 2. Annotation process.

Figure 3. 2D annotation tool in semantic segmentation.

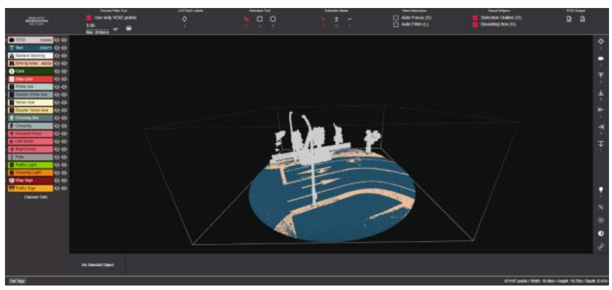

Figure 4. 3D annotation tool in semantic segmentation.

Once our model is trained, we will produce an initial map by processing Rosbag and PCD point cloud data through APIs (Application Programming Interfaces). As shown in Figure 5, the Rosbag contains data from LiDAR, camera, GNSS (Global Navigation Satellite System), and CAN (Controller Area Network). The PCD point cloud map undergoes colorization through camera imagery, infusing it with more context and visual details. However, achieving perfect accuracy in map generation through automation is challenging. Therefore, even after acquiring the initial map, manual inspection and editing are necessary to ensure its correctness and fidelity. We utilize assure mapping tools for this purpose. Upon completion of this manual editing process, we obtain the final map and an updated dataset. This iterative refinement process, combining automation and human supervision, aims to optimize the quality and reliability of our mapping results.

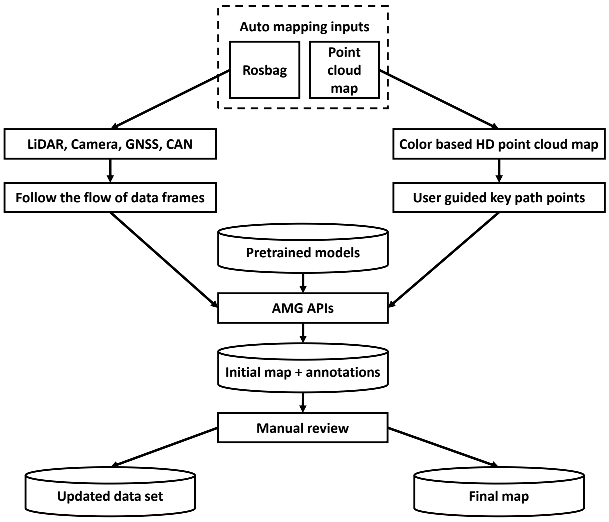

Figure 5. Semi-automated map generation process.

2.2 The extraction type of road elements

This study concentrates primarily on the extraction of certain critical elements of the road infrastructure, as listed in Table 1 and visually illustrated in Figure 6. Specifically, we target road surface markings, traffic signs, and traffic lights due to their fundamental role in providing guidance and ensuring safe navigation for autonomous vehicles.

Road surface markings offer valuable information about the nature of the road, including lanes, turns, and other critical driving guidelines. Traffic signs, on the other hand, communicate important regulatory, warning, and informative instructions that influence driving decisions. Similarly, traffic lights regulate the flow of vehicles and pedestrians, providing clear signals to control movement at intersections.

However, it's essential to note that the categories of extraction objects are not limited to these classes. The strength of our approach lies in its flexibility and scalability, as it can be readily expanded to include a wider array of objects as required. For instance, it could be further extended to extract more complex elements such as road barriers, pedestrian crossings, or different types of vehicles, among others. By doing so, we can continue to enrich the detail and comprehensiveness of our HD Maps, further enhancing their utility for autonomous driving systems.

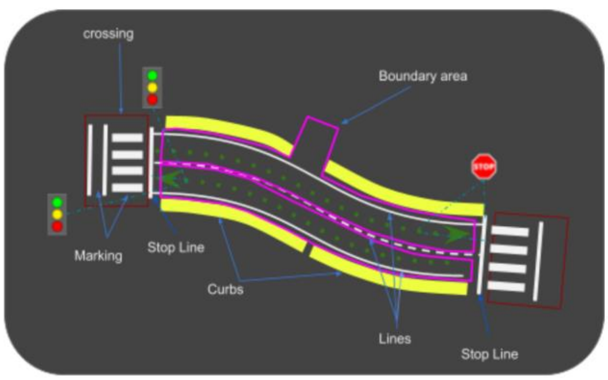

Figure 6. Example of map semantics drawn.

Table 1. Map semantic components.

Type	Description
Lane	■ The basic road network structure ■ Consists of waypoints ■ Points to branching lanes to form the navigational road network
Line	■ White line as lane boundaries
Marking	■ Other signs on the road surface ■ Such as arrows, speed limits, and driving directions
Stop Line	■ Stopping indicator for signs and lights
Curb	■ Roadside and central curbs
Boundary	■ Boundary area border of defined space
Crossing	■ Boundary to a crosswalk area
Traffic light	■ Traffic lights can be (red, green yellow, left arrow, right arrow, forward arrow, crossing red, crossing green)
Traffic sign	■ Most common signs will be represented ■ Such as stop signs, speed limit, and no parking signs

2.2.1 Yolact++: The main focus of this section is to introduce and evaluate Yolact++, an advanced real-time instance segmentation algorithm, and its application in tools for the automated creation of HD Maps.

Yolact++ is an enhanced instance segmentation algorithm based on Yolact. It employs complicated image processing and deep learning techniques to perform object detection and segmentation on input images at a faster and more accurate rate. The advent of Yolact++ not only resolves the challenges that Yolact faced when segmenting complex images, but it also furthers the ability to identify and locate object shapes. This significant advancement offers substantive changes when applied to the tools for automated production of HD Maps (Bolya et al., 2019).

Underlying Yolact++ is a carefully designed deep learning network architecture. This architecture integrates efficient object detection techniques, such as FPN (Feature Pyramid Network) and Fast(er) R-CNN (Region-based Convolutional Neural Networks), and employs a neural network known as Protonet to generate 'prototypes' of objects. This further enhances the speed and precision of segmentation. When compared to other methods, such as Mask R-CNN and Yolact, Yolact++ showcases noticeable advantages. First, Yolact++ substantially accelerates segmentation speed while maintaining high precision, made possible by the use of advanced integration techniques such as DCNv2 (Deformable Convolutional Networks v2) and Fast NMS (Non-Maximum Suppression). Furthermore, Yolact++ provides not only the bounding box for each object but also the complete segmentation map, thus offering more detailed information and aiding the creation of more accurate maps (Bolya, D. et al., 2019).

2.2.2 VoxelNet: VoxelNet was introduced in 2018. It employs a VFE (Voxel Feature Encoding) layer to learn local spatial features and subsequent volumetric representation. It employs a volumetric feature encoding layer to convert the unstructured point cloud data into a structured 3D feature map, which can be more easily processed by standard 3D convolutional layers.

Compared to other 3D object detection networks like PointNet and AVOD (Aggregate View Object Detection), VoxelNet shows a distinct advantage. PointNet operates directly on raw point clouds but struggles to capture local spatial features, and AVOD requires expensive pre-processing and conversion of LiDAR data to pseudo-images.

In contrast, VoxelNet, with its voxel feature encoding layer, learns complex spatial features and eliminates the need for manual feature engineering, leading to improved accuracy and efficiency (Zhou et al., 2018). VoxelNet's ability to operate directly on raw point clouds makes it an effective tool for map production. As point cloud data often form the basis for such maps, VoxelNet's capabilities for high-accuracy object detection and feature extraction are a critical asset in the creation of HD Maps. Furthermore, the model's end-to-end architecture facilitates a streamlined workflow, from raw data input to object detection.

2.2.3 Properties of Lane: The point clouds extracted through our automated HD Maps mapping tool subsequently undergo a modelling process. This allows us to acquire the necessary properties required for the map, as displayed in Table 2, which shows the properties of lanes recorded by assure mapping tools. These properties include elements like ID of lanes, speed limit, width, and so forth. These pieces of information can also be visually represented when utilizing the tool's graphical interface. As depicted in Figure 7, when a lane is selected, the relevant details of that lane are displayed in an information panel. This feature not only enriches the user experience but also provides immediate and comprehensive insight into the selected object's properties. The inclusion of such properties is essential as it allows users to understand more intricate aspects of the mapped environment, fostering greater precision and utility in autonomous driving applications. The visual display of these properties also enhances user interaction, making the HD Maps mapping tool more intuitive and user-friendly.

Table 2. Attributes of lane.

Attribute	Description
LaneID	Unique number identifier
LaneNumber	Order from left to right in case of multiple parallel lanes
Action	Semantic lane direction
Action cost	Cost of choosing certain action (actions and costs are pairs)
Width	General width of the lane
Speed	Maximum driving speed of the lane in km/hr
Type	Lane type
Lane change	If lane change is possible of not
From Lanes	List of lanes ids that lead to this lane
To Lanes	List of lanes ids that branch from this lane
LeftID	ID of the parallel lane on the left
RightID	ID of the parallel lane on the right
OppositeID	ID of the opposite lane on the same road when no road is exists
RoadID	Related road section ID

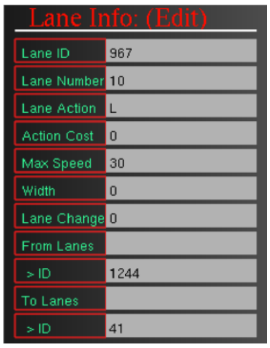

Figure 7. Lane supported properties.

3. Results and discussion

In this research, data acquisition utilizes the RIEGL VMX-250 system, a cutting-edge technology graciously supplied by a professional surveying firm. The testing grounds for this exploratory study are strategically positioned within the boundaries of the Taiwan CAR Lab, nestled in Shalun, Tainan, Taiwan. As Taiwan's premier and pioneering closed-field testing site, this facility delivers a well-controlled environment for examining and calibrating algorithms pertinent to the future of autonomous transportation.

The distinctive characteristics of the Taiwan CAR Lab make it an optimum hub for the progress and evolution of autonomous navigation technologies. The establishment is equipped with thirteen simulated traffic situations, which simulate challenging scenarios including railway level crossings, curved roads, and tunnels. This diversified simulation facilitates the production of invaluable data, contributing immensely to the respective research.

A key attribute of this testing environment is the presence of clearly markings and the exclusion of interference from external vehicles. These factors significantly enhance the accuracy and reliability of the testing process. Additionally, the simulated traffic scenarios faithfully reproduce real-world conditions, ensuring the experiments' relevance and applicability.

In short, the Taiwan CAR Lab offers an unparalleled platform for researchers to investigate and refine technologies associated with autonomous driving and map creation under controlled yet realistic circumstances. Refer to Figure 8 for a visual representation of this innovative site.

Figure 8. The test field: Taiwan CAR Lab (CARTURE, 2019).

3.1 Efficient extraction of road elements through frame-based processing

As depicted in Figure 9, our HD Maps creating tool is designed to effectively identify and extract road elements utilizing a preset model. To achieve efficiency and computational manageability, this process implements a unique frame-based approach to dissect the large-scale data set.

Initially, the data set, consisting of rich environmental details captured by Lidar sensors or cameras, is partitioned into numerous frames. This partitioning is executed based on the trajectory data, which represents the motion path of the sensor-equipped vehicle. By breaking down the extensive data set into smaller, manageable frames, the tool significantly reduces the computational burden. This method is highly beneficial when dealing with massive amounts of data, as is typically the case with HD Maps mapping tasks.

The radius or the size of these frames is not fixed but can be altered by adjusting specific parameters. This customizable feature allows for greater flexibility in data handling and analysis. It empowers users to tune the frame size according to the specific needs of their mapping project, offering a balance between computational efficiency and detail extraction precision.

Once the frame partitioning is accomplished, the tool sets about extracting the targeted road elements sequentially from each frame. The targets include road surface markings, traffic signs, traffic lights, and any other features defined in the preset model. The extraction process employs sophisticated algorithms to recognize and segregate these features from the point cloud data.

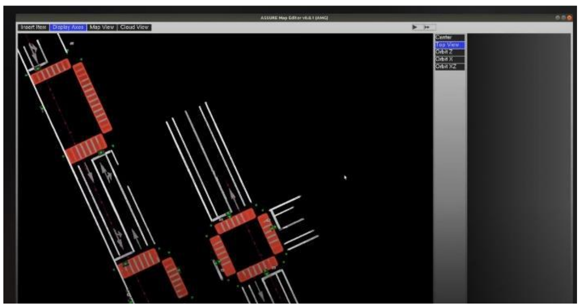

Figure 9. The procedure for flexible HD Maps production.

3.2 Classification and Recognition of Road Elements

Once the extraction process is completed, the tool assigns categories to the identified point clouds, reflecting their corresponding real-world features, as shown in Figure 10. The model's classification algorithm is good at identifying various road elements, enhancing the richness and accuracy of the extracted data.

Crucial traffic infrastructure such as lane lines, crossing lines, and traffic signals, among others, are successfully detected and categorized. Lane lines, which serve as crucial guidance for vehicular movement, are detected with high precision, their positions and orientations accurately reflected in the categorized point cloud data.

The detection and classification of crossing lines are equally accurate. Traffic lights, crucial for controlling vehicular flow and ensuring road safety, are also successfully identified, with each individual light signal accurately represented in the classified point cloud data.

In conclusion, the tool effectively transforms raw sensor data into structured, classified point cloud data, where each point cloud represents a specific type of road element. This classification enhances the utility and interpretability of the extracted data, paving the way for automatic creating detailed, accurate HD Maps.

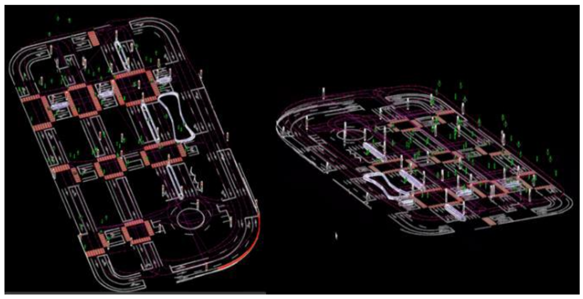

Figure 10. Extraction of road elements by using assure mapping tools.

3.3 Convert and output map into different kinds of format

The assure mapping tool supports the map format conversion function. As Figure 11 shown, the tool can convert and out the map into vector map, OpenDRIVE, Lanelet2, and kml format. These maps can be used for autonomous vehicle simulators such as CARLA, VTD, Autoware Universe, and others. Figure 12 to Figure 14 shows the simulation results. The map can be successfully input by simulator and provide the usability by using this method.

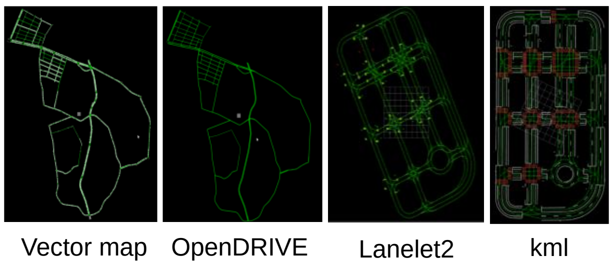

Figure 11. Output map with different kinds format.

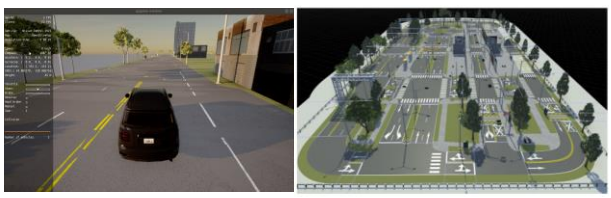

Figure 12. CARLA simulator with OpenDRIVE format.

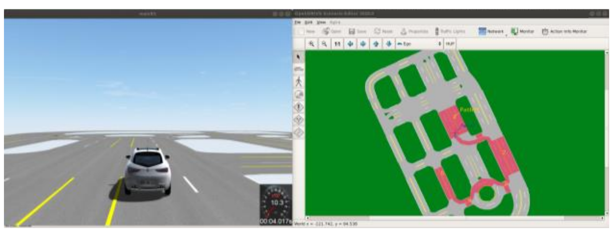

Figure 13. VTD simulator with OpenDRIVE format.

https://doi.org/10.5194/isprs-archives-XLVIII-1-2024-79-2024 | © Author(s) 2024. CC BY 4.0 License.

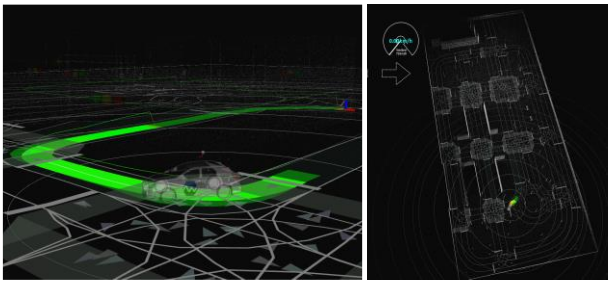

Figure 14. Autoware Universe with Lanelet2 format.

4. Conclusion

In response to the development of autonomous vehicle and the requirement of HD Maps, the car company in advanced countries have invested in the research for self-driving maps. Currently, many foreign countries prioritize the development of maps for highways and expressways because the complexity of road infrastructure with highway and expressway is lower than urban roads. When map companies try to make HD Maps in Taiwan, they have to referenced the HD Maps guidelines and standards to make sure that the quality and accuracy of map data are the same. The procedure costs lots of times and resources. In order to increase the capability and mileage of HD Maps, this study proposes two flexible surveying and mapping methods as diverse sources of GCP (Ground Control Point) and autonomous vehicle grade data with multi-sensor fusion engine system. The preliminary results show that the virtual ground control points and integrated LC-INS (Loosely Couple-Inertial Navigation System)/GNSS/LiDAR for SLAM (Simultaneous Localization And Mapping) can be used for HD Maps construction. Then, the proposed methods use for making National expressway in Taiwan. The total mileages are around 458 km. The point cloud map also can be used as a base map to accomplish the concept of ground control point cloud. The ground control point cloud can be considered as ground control point so that the new collecting point cloud can implement the establishment, updating, and increasing density from original point cloud. On the other hand, integrated multi-sensor and point cloud map can provide SAE level 3 navigation of autonomous vehicle service. The positioning accuracy can achieve "where in lane (0.5 m)" level. These experiences will improve the efficiency and integrity for making map.

Acknowledgements

The authors would like to thank the financial support by the Ministry of the Interior (MOI), R.O.C. (Taiwan).

References

Bolya, D., Zhou, C., Xiao, F., Lee, Y. J., 2019. YOLACT++: Better real-time instance segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(2), 1108-1121.

CARTURE, 2019. 政府一小步，台灣智慧駕駛科技一大步. Retrieved from https://www.carture.com.tw/topic/article/5663.

Chiang, K.W., Zeng, J.C., Tsai, M.L., Darweesh, H., Chen, P.X., Wang, C.K., 2022. Bending the curve of HD Maps production for autonomous vehicle applications in Taiwan. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 15, 8346–8359.

Elhousni, M., Lyu, Y., Zhang, Z., Huang, X., 2020. Automatic building and labeling of HD Maps with deep learning. *Proceedings of the AAAI Conference on Artificial Intelligence*, 13255-13260.

HERE, 2017. HERE HD Live Map. Retrieved from https://www.here.com/sites/g/files/odxslz166/files/2018-11/HERE_HD_Live_Map_one_pager_0.pdf, 2017.

Liu, R., Wang, J., Zhang, B., 2019. High Definition Map for automated driving: Overview and analysis. *Journal of Navigation*, 1-18.

Mi, L., Zhao, H., Nash, C., Jin, X., Gao, J., Sun, C., Schmid, C., Shavit, N., Chai, Y., Anguelov, D., 2021. HDMapGen: A hierarchical graph generative model of High Definition Maps. *Computer vision foundation*.

TomTom, 2018. HD Map with roadDNA. Retrieved from http://download.tomtom.com/open/banners/HD_Map_with_RoadDNA_Product_Info_Sheet.pdf.

Yang, B., Fang, L., Li, J., 2013. Semi-automated extraction and delineation of 3D roads of street scene from mobile laser scanning point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing*, 80-93.

Zeng, J.C., 2020. Automated road-elements modelling and centerline generation for high-definition maps utilizing 3D point cloud. Department of Geomatics, National Cheng Kung University.

Zhou, Y., Tuzel,O., 2018. VoxelNet: End-to-End Learning for Point Cloud Based 3D Object Detection. *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 4490-4499.

https://doi.org/10.5194/isprs-archives-XLVIII-1-2024-79-2024