This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Generation of Horizontally Curved Driving Lines in HD Maps Using Mobile Laser Scanning Point Clouds

Lingfei Ma, *Student Member, IEEE*, Ying Li, Jonathan Li, *Senior Member, IEEE*, Zilong Zhong, and Michael A. Chapman

***Abstract*—This paper presents the development of a semiautomated driving line generation method using point clouds acquired by a mobile laser scanning system. Horizontally curved driving lines are a critical component for high-definition maps that are required by autonomous vehicles. The proposed method consists of three steps: Road surface extraction, road marking extraction, and driving line generation. First, the points covering road surfaces are extracted using the curb-based road surface extraction algorithms depending on both the elevation and slope differences. Then, road markings are identified and extracted according to a variety of algorithms consisting of georeferenced intensity imagery generation, multithreshold road marking extraction, and statistical outlier removal. Finally, the conditional Euclidean clustering algorithm is employed, followed by the cubic spline curve-fitting algorithm and equidistant line-based driving line generation algorithms for horizontally curved driving line generation. Our method is evaluated by six MLS point cloud datasets collected from various types of horizontally curved road corridors. Quantitative evaluations demonstrate that the proposed road marking extraction algorithm achieves an average recall, precision, and F1-score of 90.79%, 92.94%, and 91.85%, respectively. The generated driving lines are assessed by overlaying them on the manually interpreted reference buffers from 4-cm resolution unmanned aerial vehicle orthoimagery, and a 15 cm level navigation and localization accuracy is achieved.**

***Index Terms*—Cubic spline, driving line, equidistant line, high-definition (HD) map, mobile laser scanning (MLS), point clouds, road marking, road surface.**

I. Introduction

In RECENT years, many prominent multinational automotive manufacturers (e.g., General Motors, BMW, Mercedes-Benz, Audi, Toyota, and Ford) and information and communication technology companies (e.g., Google, Apple, Uber, Tesla, Baidu, and Nvidia), are investing heavily, adjusting their development strategies, and indicating their ambitions to participate in the emerging market of autonomous vehicles (AVs) [1]. An AV is capable of determining the best navigation routes, driving itself on the most challenging road environments, and avoiding collisions with fixed or moving road users (e.g., pedestrians, cyclists, and cars) without direct driver intervention [2].

According to the updated policy on automated vehicle development, released in May 2016 by the U.S. Department of Transportation and the National Highway Traffic Safety Administration (NHTSA), six levels are defined for vehicle automation: No automation (Level 0), driver assistance (Level 1), partial automation (Level 2), conditional automation (Level 3), high automation (Level 4), and full automation (Level 5) [3]. In order to achieve fully autonomous driving (Level 5), several subsystems, including sensing and perception systems, decision support systems, operating systems, hardware platforms, and cloud platforms, are highly integrated to work in parallel [4].

These subsystems are essential components to control and navigate an AV, depending on two sources of data: Real-time sensing and perception data from multiple onboard sensors, and preloaded high-definition (HD) maps [5]. The onboard sensors, including Global Navigation Satellite System (GNSS) units, RAdio Detection And Ranging sensors, video cameras, and Light Detection And Ranging (LiDAR) sensors, are capable of supporting the navigation, localization, object detection, and object tracking missions. Moreover, a centralized computing system analyzes all of the data obtained from the multiple sensors to manipulate the steering, acceleration, and braking. However, such onboard sensors cannot provide efficient and reliable navigation services for AVs under complex urban street canyons or uncertain road conditions, including the limited sight distances and ambiguous visual clues [6]. Therefore, detailed and updated HD maps are necessary to provide precise vehicle localization and real-time route navigation services for autonomous driving in all road environments [7]. These preloaded maps can provide AVs with an extended monitoring range, allowing cars to anticipate in turns, roundabouts, and intersections far beyond the reach of onboard sensors.

Compared with conventional navigation maps, the lane-level accuracy performance of HD maps can reach 30 cm [8]. These HD maps, composed of detailed sublane level road information, are commonly developed by remote sensing and surveying

Manuscript received April 7, 2018; revised January 22, 2019 and February 18, 2019; accepted March 6, 2019. This work was supported in part by the National Natural Science Foundation of China under Grant 41471379 and in part by the Natural Sciences and Engineering Research Council of Canada under Grant 50503-10284. *(Corresponding author: Jonathan Li.)*

L. Ma, Y. Li, and Z. Zhong are with the Department of Geography and Environmental Management, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: l53ma@uwaterloo.ca; y2424li@uwaterloo.ca; z26zhong@uwaterloo.ca).

J. Li is with the Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University, Xiamen 361005, China, and also with the Departments of Geography and Environmental Management and System Design Engineering, University of Waterloo, Waterloo, ON N2L 3G1, Canada (e-mail: junli@xmu.edu.cn; junli@uwaterloo.ca).

M. A. Chapman is with the Department of Civil Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada (e-mail: mchapman@ryerson.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JSTARS.2019.2904514

1939-1404 © 2019 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.
See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

techniques. Laser scanning techniques are capable of providing highly accurate and georeferenced 3-D data with higher time efficiency and better cost savings in comparison to aerial imagery [9]. Mobile laser scanning (MLS) systems are the vehicle-based mobile mapping systems that produce three-dimensional (3-D) point clouds from the surrounding objects in urban road environments [10]. Furthermore, the point density captured by MLS platforms can reach up to 8000 points/m^2. It is challenging for both airborne laser scanning (ALS) and terrestrial laser scanning (TLS) platforms to achieve such precision and flexibility [11]. Accordingly, due to its high flexibility and quick data collection rate in large urban areas, MLS point clouds are more capable than other surveying and remote sensing techniques for driving line generation in this paper.

However, processing a massive amount of 3-D point clouds is a great challenge. These 3-D point clouds contain highly dense points with 3-D georeferenced intensity and detailed geographic information, which also causes challenges. Additionally, the complex patterns of curved road markings at horizontally curved road sections make detection and extraction challenging. Object variations, density variations, outliers, and the incompleteness of objects caused by occlusions and different sensor-scanning patterns must be checked and corrected. Moreover, other limitations arise from prior knowledge (e.g., road design regulations) in lane centerline extraction and driving line generation. Such challenges add to the difficulty in generating driving lines using MLS point clouds.

In this paper, we propose a semiautomated method for the generation of horizontally curved driving lines in urban curved road corridors using MLS point clouds. The main contributions of this paper are as follows: 1) several optimized algorithms to extract road surfaces and road markings using MLS point clouds at horizontally curved road sections are developed; and 2) a novel semiautomated algorithm to generate horizontally curved driving lines is proposed according to equidistant lines and road design regulations.

II. Related Work

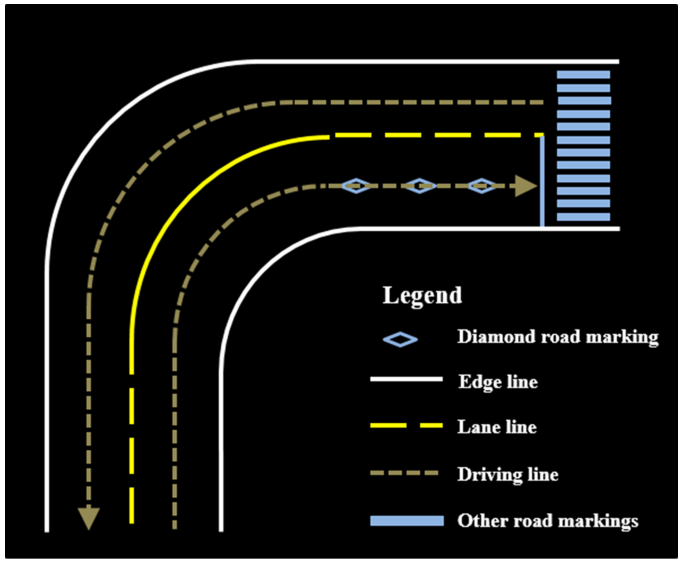
MLS systems are capable of achieving a high acquisition rate of three geospatial point clouds, which can be used for an extended range of intelligent transportation system applications [12]. Such accurate and high-density 3-D MLS data contributes to HD maps [13], road asset inventory [14], and road scene structure [15]. MLS systems have indicated great potential for rapid commercialization [16]. The following subsections present a systematic literature review of related techniques and existing methods for HD map development, road surface extraction, road marking extraction, and driving line generation using MLS point clouds.

A. Introduction to HD Maps and Horizontal Curves

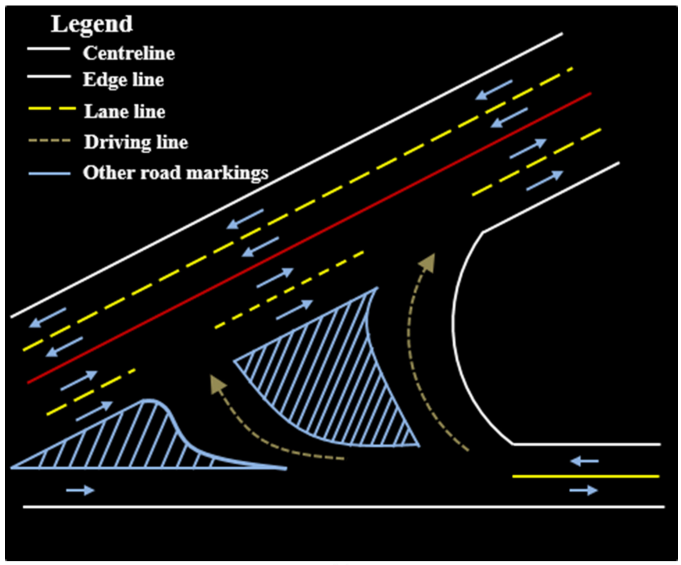
HD maps provide a highly accurate and realistic representation of the current road networks, which are capable of updating in real time and recording traffic changes, such as accidents, lane closures, traffic congestion, and updated speed limitations at cm level of accuracy [17]. Additionally, such maps are a significant element of highly autonomous driving technology, assisting AVs to precisely localize themselves on the road, and providing them with an extended monitoring range to anticipate in turns and intersections far beyond the reach of onboard sensors [18]. Furthermore, the HD Live Map established by HERE consists of dynamic content layers to provide detailed and real-time road information, including detailed sublane level information, dynamic road conditions, roadside infrastructure (e.g., traffic signs, roadside trees, and light poles), and speed profile data [13].

Horizontal curves have a considerable impact on traffic safety and efficiency due to gradual transitions to or from the curves and limited sight distances for both drivers and onboard sensors [19]. According to the Preliminary 2016 Ontario Road Safety Annual Reports Selected Statistics submitted to the Ministry of Transportation Ontario, Canada, 11.9% of fatal collisions occurred in Ontario were related to horizontal curves [20].

Fig. 1 shows typical line-shaped curved road markings defined in HD maps at a two-way and two-lane horizontal curve



(a)



(b)

Fig. 1. Lines in HD maps defined in this paper. (a) Circular horizontal curve and straight lines. (b) Complex road section.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[see Fig. 1(a)], and a two-way and four-lane road corridor with horizontal curves [see Fig. 1(b)], respectively. Several assumptions are made in the process of generating driving lines for HD maps. First, each horizontally curved road section is composed of many line-shaped curved road markings, such as lane lines, edge lines, and centerlines, which describe the geometric details of horizontal curves. Second, driving lines are generated to represent the real driving routes for AVs within the allowable positioning and navigating errors. Third, it is assumed that a majority of curved road sections are designed as simple circular curves in order to ensure a steady driving transition and reduce road hazards. This paper mainly focuses on developing reliable driving lines at horizontally circular curved road sections. The fully autonomous driving function requires positional accuracy of better than 30 cm for a 3 m wide lane, and provides AVs with a lateral 20-cm offset from the lane centerlines [21].

B. Road Surface Extraction

A variety of methods and algorithms have been developed to detect and extract road surfaces from MLS data. Such methods are mainly classified into three categories based on the following data formats: 3-D raw point clouds, 2-D feature images, and other data sources [14].

Road surface detection and extraction can be implemented on either 3-D point clouds or 2-D feature images derived from 3-D points. In [22], the authors performed a road segmentation method using a curvature analysis directly interpreted from MLS data. According to the parametric active contour or snake model, an automated algorithm for road edge extraction was proposed from MLS point clouds [23]. In [24], road edges and road surfaces were extracted by determining the angular distance to the ground normal from MLS point clouds. In [12], [25], and [26], a voxel-based upward growing algorithm was proposed to directly filter out ground points from raw MLS data. In order to enhance computational efficiency, trajectory data was widely applied [27]–[29]. In [15] and [30], a supervoxel generation method was applied to automatically extract road boundaries and road surfaces from MLS data. A computationally efficient method integrating supervoxel with Hough forest framework was used in [31] to filter out pavement from 3-D MLS points for further object detection. In [32], the raw MLS point clouds along the trajectory of a vehicle was vertically partitioned by first. The RANdom SAmple Consensus algorithm was then implemented to extract ground points based on the average height of ground. Additionally, the spatial configuration of the scan line relies on related parameters (e.g., driving speed, sensor trajectory, and scanner orientation) of a specific MLS system. Thus, the high-density pavement points boost the computational efficiency in road segmentation by processing scan lines [28], [29], [33], [34]. Furthermore, some studies concentrated on extracting road surfaces directly from MLS data, based on the smoothness of road surfaces [35]. These methods are computationally expensive and time consuming. Moreover, the majority of previous methods require much prior knowledge about data characteristics and road properties. It is thus challenging to achieve fully automated road surface extraction.

Converting 3-D laser point clouds into 2-D georeferenced feature (GRF) images can effectively extract road surfaces using the existing image processing approaches, and achieve computational efficiency enhancement [36]. Road segmentation of range scan lines was performed in 2-D. An effective method was proposed in [33] to differentiate geometric features (e.g., buildings, trees, and road surfaces) by analyzing the height deviation. In [37], with the assistance of the scan line segment between vehicles trajectory points, road edges and road surfaces were detected and extracted by taking height differences, altitude mean values, and altitude variances into consideration. In the study conducted by [38], high elevation points were filtered out from the profiles, and rapid slope changes in the spline were then determined. Subsequently, road surfaces were completely extracted in [39] by generating GRF images to separate road surface points from the entire MLS point clouds. 3D–2D dimensional reduction methods reduces the spatial distances between points belonging to the same classes and thus significantly decreases computational cost. However, it is still challenging to handle undulating terrain environments by using 2-D GRF images.

With the assistance of other data sources, such as UAV imagery, high-resolution satellite imagery, and ALS or TLS point clouds, the accuracy of extraction results can be improved by using additional road texture information. In [40], 3-D surface geometric models were generated with the fusion of MLS data and camera data. Moreover, multiple aerial scans and TLS high-density point clouds of an urban road scenario were combined together to effectively extract road surfaces [41]. However, these data obtained from multiple sensors at several times, in various weather situations, under different lighting conditions, and with diverse sampling densities, make the data calibration and data fusion challenging [14]. Since point clouds are not in a regular format, the deficiencies of massive MLS points (e.g., large volume, distortions, and occlusions) bring in dilemmas and uncertainties for accurate road segmentation. Therefore, it is necessary to propose an improved and robust road surface extraction method from MLS point clouds to deal with the aforementioned problems.

C. Road Marking Extraction

Road markings, as significant elements in traffic management systems, play a critical role in providing guidance, warning, and bans for all road users. Typically, road markings are highly retro-reflective paintings on concrete pavement. Therefore, the relatively high intensity is regarded as a unique characteristic to identify and extract road markings from MLS point clouds [42], [43]. According to semantic knowledge (e.g., shape and size) and laser intensity characteristics, road marking extraction methods are mainly categorized into two types: 2-D GRF image-based extraction, and 3-D point-based extraction methods [44].

Most studies extracted road markings from 2-D GRF images interpreted from 3-D point clouds. Therefore, the existing image processing algorithms, including multiscale threshold segmentation, Hough Transform, and multiscale tensor voting (MSTV) were applied with regard to the semantic information

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

of road markings [39], [45], [46]. For instance, in [43], a range-dependence thresholding algorithm was employed to identify and extract road markings from intensity and range images. Based on the generated 2-D feature images from raw MLS data, a global intensity filtering method was performed to roughly detect road markings [45]. In addition, in [39], a georeferenced reflectance intensity image was generated by first, and a Hough Transform approach was then employed in four connected regions of the image to extract broken lane line markings and continuous road edge lines. However, the Hough Transform method has a difficulty in processing complex road markings (e.g., words) while specifying the number of road markings to be extracted. In contrast, the MSTV algorithm is capable of noise suppression and road marking preservation. In [46], a dynamic multiple thresholding method was implemented first by determining the relationships between scanning range and intensity values, followed by a morphological nearest operation with linear road structures. Then, a remarkable improvement was achieved by using the MSTV algorithm. Extracting road markings by generating 2-D GRF images from MLS data is persuasive to deal with the intensity inconsistency issue caused by scanning patterns. Nevertheless, extracting complicated types of road markings (e.g., digits and hatchings) is a very difficult task using 2-D feature image processing algorithms.

Meanwhile, many studies focused on directly extracting road markings from 3-D point clouds. In [12], road markings were directly extracted and classified from raw 3-D point clouds into edge lines, stop lines, zebra crossing lines, arrow markings, rectangular markings, and centerlines. Based on road curbs and trajectory data, the large-size road markings through spatial density filtering and multisegment thresholding methods were extracted by first, while Otsus thresholding algorithm was adopted to determine optimal thresholds [47]. Subsequently, the small-size road markings were extracted according to the principal component analysis (PCA) and machine learning methods. In [48], profile-based intensity analysis algorithms were proposed to directly extract painted markings from raw point clouds. First, raw 3-D points were segmented into point cloud slices with the assistance of trajectory data. Next, road surfaces were identified, based on the geometric properties of road edges, barriers, and boundary lines. Finally, linear road markings were successfully extracted by analyzing the peak value of intensity within each scan line. Compared with 2-D image-based extraction methods, MLS point-based extraction methods aiming to directly detect and extract road markings using raw MLS data, which is capable to improve completeness and correctness in extraction results within a short computational time. Additionally, their geospatial information of road markings is preserved after extraction, which can be utilized in further applications. However, automated extraction of road markings from a mass of 3-D laser points especially with huge concavo-convex features and unevenly distributed point clouds is still a very challenging task. Thus, a refined multithreshold road marking extraction algorithm is proposed in this paper to improve the robustness and effectiveness for road marking extraction using MLS point clouds.

D. Road Horizontal Parameter Estimation

Driver behavior, including misperceptions of driving speed and poor visibility at horizontal curves, increases the potential risks of traffic accidents. Thus, detecting traffic conditions, especially at horizontally curved road sections, and estimating road horizontal parameters (e.g., curvature) are significant for AVs to determine reliable driving lines and prevent collisions [49].

Many studies have been performed to extract geometric parameters at horizontally curved road sections by using MLS systems [50]–[54]. In [50], the software for precise estimations of road horizontal geometric features was released using a suitably equipped vehicle moving along the road on a two-way trip. Road centerlines were extracted at first with the assistance of trajectory data. Subsequently, the parameters of horizontal features were determined according to a least-squares optimization of characteristic curves. Additionally, in [51], an automated solution was proposed to estimate curvature diagrams and analyze horizontal geometric features by using trajectory data. In [52], road geometry of horizontal elements was defined, based on the dynamic measurements of GNSS systems. To begin with, the centerlines were regarded as the middle points between two trajectories in opposite directions. Then, a least-squares adjustment was applied to estimate the horizontal elements (e.g., straight lines, circle arcs, and clothoid curves). Additionally, Holgado-Barco *et al.* [54] proposed a semiautomatic method to extract road horizontal alignment from a mobile LiDAR system, such as road centerlines, straight lines, circular arcs, and clothoids according to road axis modeling and transportation design standards. Meanwhile, in [55], a cost-effective system was developed to calculate vertical alignments of a road, based on data acquisition from GNSS sensors. In [56], the road geometry of horizontal alignments was semiautomatically extracted from MLS data. The related features were first extracted to model the road axis from MLS data. The geometric design features of the horizontal alignments were afterward calculated using azimuth and curvature information. However, these methods mainly depend on the localization precision of GNSS signals, which can be impacted in GNSS weak/denied environments. Therefore, in this paper, estimating such road horizontal parameters based on the generated mathematical equations can remarkably improve the estimation precision and reduce the labor cost.

III. Method for Driving Line Generation

In this section, we present the theoretical and technical details for the proposed method of semiautomated generation of horizontally curved driving lines. Such generated driving lines can be further used to support the development of HD maps and autonomous driving.

A. Workflow of the Proposed Method

To guarantee solid and precise navigation solutions for AVs in complex urban road environments, this semiautomated method endeavors to generate driving lines from MLS point clouds. This method consists of three modules: Road surface extraction, road marking extraction, and driving line generation (see Fig. 2).

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

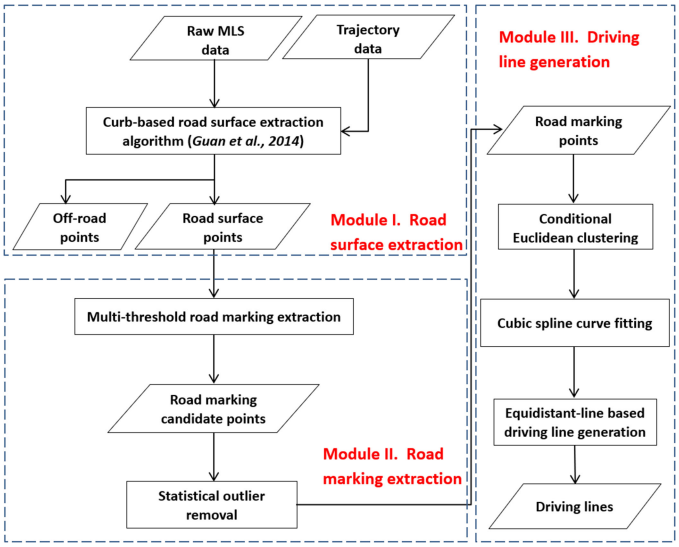


Fig. 2. Workflow of the proposed method.

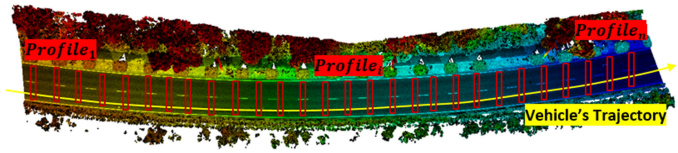


Fig. 3. Profiling process on an MLS point cloud dataset.

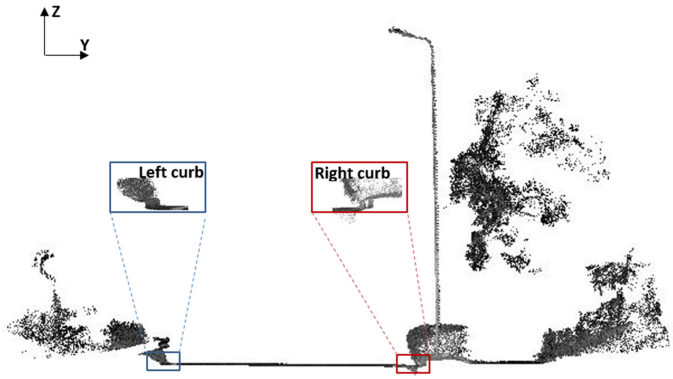


Fig. 4. Curb identification in a profile image.

In Module I, with the assistance of trajectory data, the curb-based road surface extraction algorithm [28] is performed to detect and extract road surfaces from raw MLS point clouds. In Module II, the multithreshold road marking segmentation method [12] is conducted for road marking extraction. A statistical outlier removal (SOR) algorithm is afterward to filter out noise points (e.g., isolated points and outliers). Module III, in order to determine the best-fitting horizontal curves, the conditional Euclidean clustering (CEC) algorithm is used first, followed by a cubic spline curve-fitting algorithm. The equidistant-line based driving line generation algorithms are subsequently proposed to generate driving lines.

B. Module I: Road Surface Extraction

In this paper, we employ a curb-based road surface extraction algorithm [28] to extract road surfaces from the raw MLS point clouds. Depending on the vehicle's trajectory data, the raw MLS point clouds are partitioned into a sequence of point cloud data blocks, in each of which a corresponding profile is sectioned with a certain width. Next, the point clouds in each profile are projected onto the plane perpendicular to the direction in which the vehicle moves forward. Each profile is afterward gridded to generate a pseudo scan line, and a principal point is determined within each grid cell. Then, road curb points are extracted from each pseudo scan line, based on both elevation and slope differences. Finally, a cubic B-spline interpolation algorithm is employed to fit the curb points into two smooth road edges. All point clouds located between two smooth edge lines are considered road surface points.

The curb-based road surface extraction algorithm is adopted and revised to improve the performance in road surface extraction. However, the majority of road surface extraction methods based on trajectory data generally used a fixed-size data block to segment road surfaces. Such methods cannot accurately extract road surfaces for curved road sections. Moreover, point densities are various due to different scanning patterns and varying scanning distances between scanning objects and onboard laser scanners. In order to overcome the variations of point density and enhance the computational efficiency, trajectory points cannot be directly used, since the curvatures of horizontal curves have an influence on the size of a data block to be partitioned. Therefore, compared to the previous study conducted in [28], which utilized a fixed size of data blocks, our revised algorithm can dynamically determine this size based on the differences of horizontal curvatures. Fig. 3 presents a test sample of the raw MLS point clouds. The red rectangles represent the sliced profiles, and the yellow line indicates the vehicle's trajectory data. Such data, regarded as a driving route of a moving vehicle, play a critical role in the process of MLS point clouds profiling. To this end, three critical parameters, block width (B_g), grid width (S_p), and profile width (P_g), are involved. The B_g determines the size of a data block to be sectioned, S_p is the grid size of a pseudo scan line, and the P_g controls the number of points to be counted at the stage of pseudo scan line generation. According to a collection of experiments, the values of B_g range from 0.5 to 2.0 m, and S_p range from 0.05 to 0.10 m in this paper, based on the trajectory data and curvatures of horizontal curves. The larger curvature of the horizontally curved road section is, the smaller the values of B_g and S_g are determined, and the more data blocks are to be generated. Based on [28], the P_g is predefined as 25 cm.

According to two magnified views, Fig. 4 shows that the road curbs are sharp height jumps, and vertical to road surfaces. Thus, curb points can be identified by analyzing both elevation-jump and slope-difference thresholds. The proposed algorithms mathematically define the slope between two adjacent principal points (highlighted in red dots) within a generated pseudo scan line, and the elevation difference of a certain point to its nearest point in the pseudo scan line (see Fig. 5). According to

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

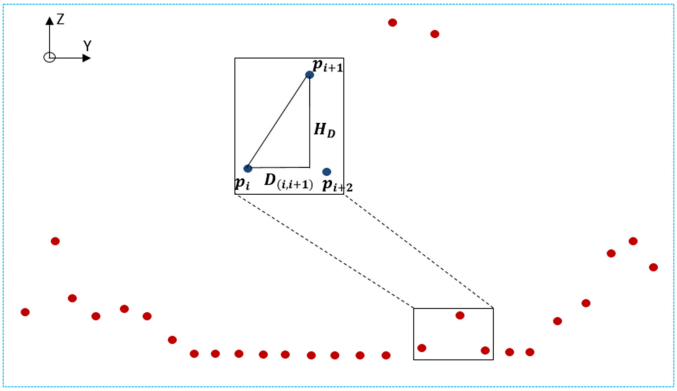


Fig. 5. Pseudo scan line of a profile image.

both slope and elevation differences, a point can be identified whether it is a curb point or not. These two criteria are therefore defined by

$$\forall p_i: \begin{cases} \text{if}((S_{\text{slope}} \geq S_T)\&(H_{\text{min}} \leq H_i \leq H_{\text{max}})), \text{candidates} \\ \text{otherwise, noncurb points} \end{cases} \tag{1}$$

where S_{slope} represents the slope of two consecutive neighbour points. S_T denotes a predefined slope threshold (i.e., S_T is $\Pi/3$ in this paper). H_i is the elevation difference of a specified point to its adjacent point. H_{min} and H_{max} indicate the minimal and maximal curb height thresholds (i.e., $H_{\text{min}} = 5$ cm and $H_{\text{max}} = 30$ cm in this paper). Furthermore, S_{slope} can be calculated by

$$\begin{cases} S_{\text{slope}} = \arctan\left(\frac{Z_{i+1}-Z_i}{\sqrt{(X_{i+1}-X_i)^2+(Y_{i+1}-Y_i)^2}}\right) \\ S_{\text{slope}} \in \left[\frac{-\Pi}{2}, \frac{\Pi}{2}\right] \end{cases} \tag{2}$$

where (X_i, Y_i, Z_i) and $(X_{i+1}, Y_{i+1}, Z_{i+1})$ denote the coordinates of two adjacent MLS points within a pseudo scan line, while *X* and *Y* coordinates located on the *YZ*-plane and *Z* coordinate represented by the elevation direction. Considering that there are both positive and negative values in (2), a positive slope value represents the point queue adding an off-road point from the road at the curb edge; whereas a negative slope value indicates the point queue switching an off-road point to the road at the curb edge.

Accordingly, an MLS point will be regarded as a curb point candidate if the slope of this point is equal or larger than the predefined slope threshold S_T (i.e., $S_{\text{slope}} \geq S_T$). If the elevation difference of a road curb point candidate is within the range of $[H_{\text{min}}, H_{\text{max}}]$, this candidate is labeled as a real curb point. Otherwise, it will be grouped into the noncurb points. Based on a prior knowledge of the road design and construction standards, these curb point candidates nearest to the vehicle's moving trajectory, are determined as the road curbs. As demonstrated in this study, such curb-based road surface extraction strategies by embedding the curvatures of horizontal curves and dynamic sizes of point cloud data blocks, can achieve greater computational efficiency than directly using fixed-sized data blocks.

C. Module II: Road Marking Extraction

Road markings, painted at horizontally curved road sections, such as solid edge lines, dashed centerlines, and broken lane lines, should be completely and effectively extracted from road surfaces for driving line generation. Based on the inverse distance weighting method [39], in combination with intensity information and local-global elevation data, the MLS road surface points are interpolated to generate 2-D GRF intensity images. There are two crucial rules for such images generation: 1) the greater intensity value of a certain point, the greater weight is assigned; and 2) the closer from a certain point to the center of the grid, the greater weight is assigned. A grid resolution is defined as 4 cm in this paper to ensure the accuracy and computational efficiency.

Road markings are brightly painted with higher reflectance materials than surrounding road surfaces. Such road markings provide higher intensities than other pavement points. Therefore, the multithreshold extraction method [12], and the Otsus thresholding [47] method are conducted to extract road markings from the generated GRF intensity imagery. The Otsus thresholding approach supposes that an image is bimodal, and its illumination is consistent. The bimodal brightness is thus calculated by analyzing the different properties of surface materials. Furthermore, the generated intensity imagery is divided into two classes: Road markings as foreground, and others (e.g., road surfaces and cracks) as background. Then, their cumulative probabilities and mean levels are determined, accordingly. Consequently, the intensity imagery is automatically segmented by selecting optimum thresholds to minimize the within-class variance.

However, noises (e.g., isolated points and outliers) inevitably exist that reduce the accuracy of extraction results and affect the completeness of the road markings. Therefore, in order to minimize the effect of noises, an SOR filter in the PCL package [57] is employed to remove noises from the extracted road marking point clouds. First, the SOR algorithm defines the number of the nearest searching points k of a certain point and determines their corresponding distances from this point to its neighbors. Next, the average distance $d_i (i = 1, 2, \ldots, n)$ of each point $p_i (i = 1, 2, \ldots, n)$ to its neighbors is calculated, whereas n is the total number of road marking points. Furthermore, with an assumption that the distribution of average distances should be fitted to a Gaussian distribution with the mean μ and standard deviation σ, points located outside a thresholding interval is described as noise points and is removed afterward from the road marking point clouds. In this way, road markings are completely extracted.

D. Module III: Driving Line Generation

3-D MLS points, pertaining to the same semantic objects, are still unorganized and isolated after noise removal. In order to successfully generate driving lines at horizontal curves, the sparse and unorganized road marking points are clustered into topological and semantic objects using the CEC method. Compared to the existing clustering methods, such as Fuzzy C-Means clustering, the advantage of using CEC method is that the constraints for clustering are customizable by the users. First, a predefined

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Euclidean distance threshold d_e is determined by considering the point density and average spacing of the generated road marking point clouds. Subsequently, two adjacent points will be assigned into the same cluster if their Euclidean distance d_i is less than or equal to d_e (i.e., $d_i \leq d_e$). Otherwise, these two points will be grouped into different clusters. The clustering targets in this paper are road markings, particularly for lane lines, centerlines, and edge lines. *A prior* knowledge (e.g., shape, scanning distance, and urban street design regulations) are utilized for completely filtering based on [56]. Accordingly, the final road marking clusters are optimized using the following criteria:

$$\begin{cases} \text{reversed clusters, if } d_r \leq d_c \ \& \ w_r \geq w_c \\ \text{removed clusters, if } d_r \geq d_c \ \& \ w_r \leq w_c \end{cases} \tag{3}$$

where d_c denotes a scanning distance threshold, w_c indicates a clustered road marking width threshold, and d_r is the scanning distance. The clusters will be removed if either the scanning distance d_r is less or equal than the given d_c (i.e., $d_r \leq d_c$), or their width w_r is larger or equal than the given w_c (i.e., $w_r \geq w_c$).

According to the road design and construction standards [58], the majority of horizontally curved road segments are designed as the combination of circular curves with straight lines [see Fig. 1(a)]. A cubic spline curve fitting algorithm is thus employed to determine the best-matching mathematical functions of these horizontal curves. Since the extracted road marking points have been projected on the *XY*-plane in Module II, the cubic spline function is defined as

$$\begin{aligned} S(x) = \Phi_0\left[\frac{x-x_i}{h_i}\right]y_i + \Phi_1\left[\frac{x-x_i}{h_i}\right]y_{i+1} \\ + h_i\Psi_0\left[\frac{x-x_i}{h_i}\right]m_i + h_i\Psi_1\left[\frac{x-x_i}{h_i}\right]m_{i+1} \end{aligned} \tag{4}$$

where $S(x)$ represents the spline function, $(x_i,\ y_i)$ denotes the coordinate information in *XY*-plane, $i = 0,\ 1,...,\ n$, n is the total number of clustered road marking points to be fitted, and m_i is the parameters defined in the cubic spline function. In addition, $x_i \leq x \leq x_{i+1}$, $h_i = x_{i+1} - x_i$, $\Phi_0(x) = (x+1)^2(2x+1)$, $\Phi_1(x) = x^2(-2x+3)$, $\Psi_0(x) = x(x-1)^2$, and $\Psi_1(x) = x^2(x-1)$ are predefined in (4), respectively. Subsequently, in order to determine the parameters m_i, the second derivative of (4) is further calculated as follows:

$$\begin{aligned} S''(x) = \frac{6}{h_i^2}\left[2\frac{(x-x_i)}{h_i} - 1\right]y_i - \frac{6}{h_i^2}\left[2\frac{(x-x_i)}{h_i} - 1\right]y_{i+1} \\ + \frac{1}{h_i}\left[6\frac{(x-x_i)}{h_i} - 4\right]m_i + \frac{1}{h_i}\left[6\frac{(x-x_i)}{h_i} - 2\right]m_{i+1}. \end{aligned} \tag{5}$$

To ensure the continuity of the second derivative, $S''(x_i - 0)$ should be equal to $S''(x_i + 0)$, where $i = 1, 2, \ldots, n-1$, which get the following:

$$\frac{(m_{i-1} + 2m_i)}{h_{i-1}} + \frac{(m_{i+1} + 2m_i)}{h_i} = 3\left[\frac{y_i - y_{i-1}}{h_{i-1}^2} + \frac{y_{i+1} - y_i}{h_i^2}\right]. \tag{6}$$

TABLE I
VARIABLES DEFINED IN A SIMPLE CIRCULAR CURVE

Symbols	Descriptions	Units
D	Degree of curvature	degree
R	Radius of curve (measured to centerline)	ft
L	Length of curve (measured along centerline)	ft
PC	Point of curvature, start of the horizontal curve	-
PT	Point of tangency, end of the horizontal curve	-
PI	Point of tangent intersection	-
Δ	Subtended angle of curve (PC to PT)	degree
T	Tangent length	ft
M	Middle ordinate	ft
LC	Length of long chord (from PC to PT)	ft
E	External distance	ft

Assume that

$$\begin{cases} \alpha_i = \dfrac{h_{i-1}}{h_{i-1} + h_i} \\ \beta_i = 3\left[(1-\alpha_i)\dfrac{(y_i - y_{i-1})}{h_{i-1}} + \alpha_i\dfrac{(y_{i+1} - y_i)}{h_i^2}\right]. \end{cases} \tag{7}$$

Thus, based on (4)–(7), the cubic spline functions with parameters $m_i (i = 1, 2, \ldots, n-1)$ for extracted road marking points can be obtained

$$\begin{cases} 2m_i + \alpha_1 m_2 = \beta_1 - (1-\alpha_1)y_0^{'} \\ (1-\alpha_2)m_1 + 2m_2 + \alpha_2 m_3 = \beta_2 \\ \vdots \\ (1-\alpha_{n-2})m_{n-3} + 2m_{n-2} + \alpha_{n-2}m_{n-1} = \beta_{n-2} \\ (1-\alpha_{n-1})m_{n-2} + 2m_{n-1} = \beta_{n-1} - \alpha_{n-1}y_n^i \end{cases} \tag{8}$$

where $y^{'} = S^{'}(x)$, it denotes the first derivative of the spline function. Therefore, the piecewise cubic spline functions about the clustered road marking point clouds at horizontal curve road section are determined. Accordingly, the curvatures of generated cubic spline functions [i.e., (8)] are calculated as follows:

$$\rho = \frac{\left|S''(x)\right|}{(1 + S'(x)^2)^{\frac{3}{2}}}. \tag{9}$$

Accordingly, if $\rho = 0$, the output of (8) is a straight line; if ρ is a constant, the output is a circular curve; and if ρ ranges from zero to a constant, the output is a transition curve. Based on (9), the normals slope of any point on the curve and corresponding normal equations are determined. Subsequently, the control points on the equidistant lines are ascertained according to the given width, slope, and direction. Table I presents all variables of a circular horizontal curve, such variables are calculated based on the obtained best-matching functions, to support the urban street design and maintenance.

According to driving behaviors and prior knowledge, the generated driving lines are parallel to the road centerlines and edge lines to ensure gradual transitions and traffic safety. Additionally, assumed that the curvatures of the generated driving lines are equal to curvatures of road centerlines or edge lines, and then the centers of these horizontal curves are the same (see Fig. 6). Meanwhile, equidistant lines play a significant role in road

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

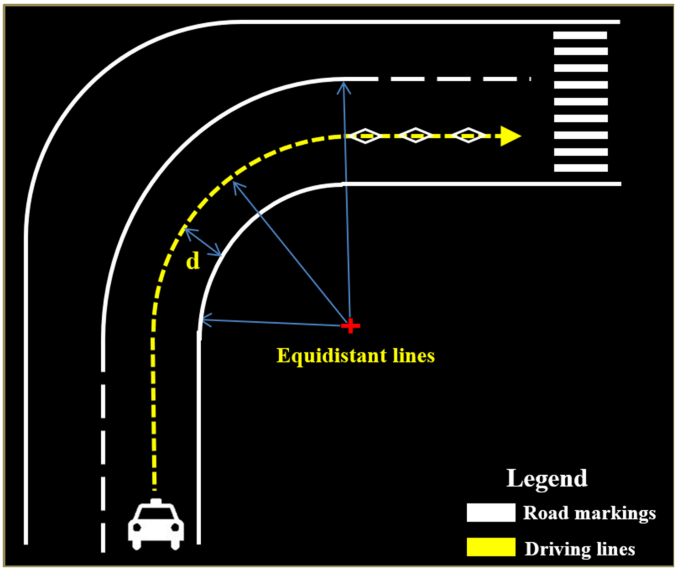


Fig. 6. Determining a driving line using the proposed method.

Fig. 7. Surveyed areas in the City of Xiamen, China. (a) Ring Road. (b) Haicang Industrial Park.

design and construction, particularly for horizontally curved road sections. In this study, the driving lines to be generated can be regarded as equidistant lines of road centerlines and edge lines. Furthermore, the equation of equidistant lines is accordingly determined as follows:

$$P_d(t) = (x(t), y(t)) \pm d\frac{(-y'(t), x'(t))}{\sqrt{x'(t)^2 + y'(t)^2}} \tag{10}$$

where $P_d(t)$ is the curve equation of a driving line to be created, $P_d(t) = (x(t), y(t))$ denotes the curve equation of the extracted road markings obtained using (8), and d is the equidistant distance. Consequently, the equidistant line located between two extracted road markings is determined, as shown in the yellow dashed line in Fig. 6. Therefore, the relative coordinates of all 3-D MLS points belonging to the driving lines, are calculated.

IV. Results and Discussion

A. MLS Point Cloud Datasets

In this study, 3-D MLS point clouds at multiple road corridors were collected on Ring Road and in Haicang Industrial Park in the City of Xiamen, Fujian, China (see Fig. 7). These MLS point clouds were collected using a RIEGL VMX-450 system.

TABLE II
Specifications of the RIEGL VMX-450 System

Components	Specifications	
Laser scanner component	Laser scanner	RIEGL VQ-450 (2)
	Laser wavelength	near infrared
	Measurement range	1.5-800 m
	Measurement precision	5 mm(1σ)
	Absolute accuracy	8 mm
	Scan frequency	400 lines/sec
	Angle measurement resolution	0.001°
	Scanner FOV	360°
POS component	GNSS types	POS LV-520
	Roll & Pitch	0.005°
	Heading	0.015°
Imagery component	Camera types	500 MP (6)
	Lens size	2/3" CCD
	Field of view	80 × 65°
	Exposure	8

This system comprises two fully calibrated RIEGL VQ-450 laser scanners, four RIGEL VMX-450-CS6 digital cameras with the pixel array of 2452H × 2056V, and one integrated Applanix POS LV 520 processing system with one GNSS antenna, one inertial measurement unit (IMU), one DMI, and one POS computing system. Based on a point-of-sale synthetic computer system, main components are assembled within a case and mounted on the roof of a motorized vehicle. Table II indicates the detailed specifications of all sensors in a VMX-450 system. The absolute accuracy of the collected MLS point clouds is 8 mm, and the measurement precision is 5 mm.

Six test datasets were selected from the Ring Road and Haicang Industrial Park datasets [see Fig. 8(a)]. Dataset I is an MLS point cloud dataset of a typical two-way and two-lane horizontally curved road section with 65 68 656 MLS points, containing road-side trees, light poles, and fences. Dataset II is a two-way horizontal curve with 20 504 262 MLS points, which includes both straight and curved circular road segments. Dataset III presents a two-way road section with 41 688 329 MLS points, which consists of two reversed curved road segments. Dataset IV indicates a two-way and two-lane horizontal curve with 2 063 626 MLS points, where it contains typical types of road markings, including lane lines, centerlines, and zebra crossings. Dataset V is a two-way and two-lane horizontal curve with 3 661 745 MLS points with a high curvature. Dataset VI is a one-way horizontal curve with 18 561 253 MLS points, and its length is approximately 200 m. Since these six datasets cover the most types of horizontal curves and road markings, they are used to validate the flexibility and computational efficiency of the proposed algorithms. Additionally, the workstation used in this paper is a Dell Alienware x51 desktop with an Intel Quad Core i5-6400 CPU, and an 8-GB RAM. The performance evaluations about road marking extraction and driving line generation were conducted on these test datasets.

B. Driving Line Generation

In order to completely extract road markings and further boost computational efficiency, road surface points were first extracted using the revised curb based road surface extraction algorithms.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

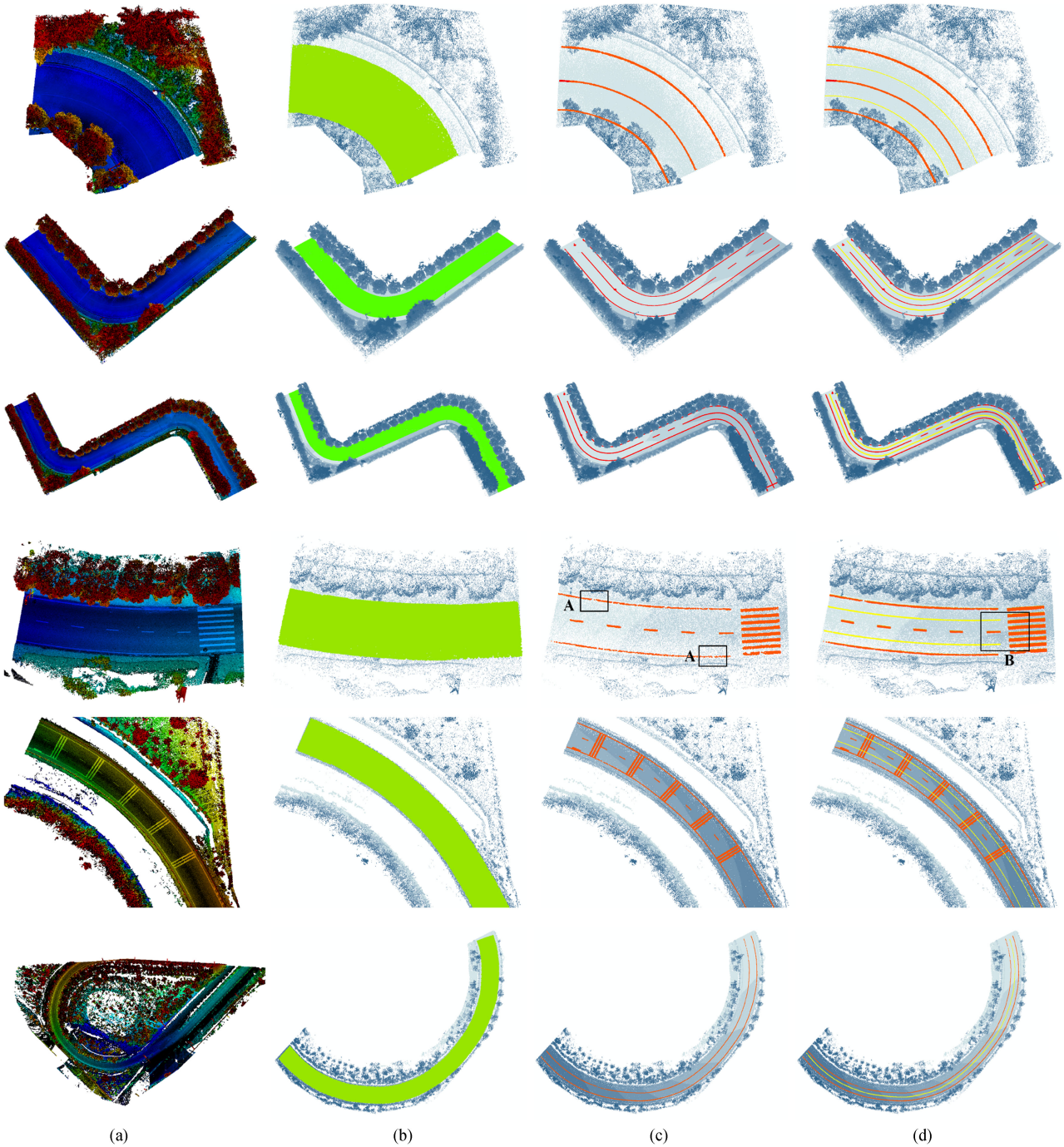


Fig. 8. Driving line generation results from six test datasets. Column (a): Raw MLS point clouds of six test datasets. Column (b): Road surface points (green) obtained after road surface extraction. Column (c): Road marking extraction results (red) obtained by using the revised multithreshold road marking extraction algorithm. Column (d): Generated driving lines (yellow) in this paper.

According to [28], $H_{\min} = 5$ cm, $H_{\max} = 30$ cm, $E_T = 5$ cm, and $S_T = \Pi/3$ were predefined. To determine the optimal parameters of B_g, P_g, and S_p, a series of experiments were conducted based on prior knowledge about road surfaces. Accordingly, the width of a block was set ranging from $B_g = 0.5$ m to $B_g = 2.0$ m, depending on the curvatures of the horizontal curves. The value of P_g was set to be 0.25 m. Each profile was then gridded into a number of grids with an adaptive grid width S_p ranging from 0.05 to 0.10 m to acquire more road segments and road profiles. Subsequently, more road curb points can be extracted for precisely detecting road curbs based on different input datasets for pseudo scan line generation. Fig. 8(b) shows the road surface extraction results, which demonstrates that our algorithm is able to deal with roads with various curvatures. Though the curvatures of horizontal curves vary in complex urban environments, the extraction results indicate a great number of nonground points were effectively excluded, and the road surfaces were thoroughly extracted. On the whole, the

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

revised method can achieve promising road surface extraction results on the six selected test datasets, and reduce the time complexity of the followup algorithms. However, some road surface points failed to be extracted occurred in Dataset IV. The health condition of road curbs has a great influence on the performance of the revised road surface extraction algorithms. According to the visual inspection, the settlement, decay, and fragile structure of road curbs in Dataset IV could conduce to the false extraction of road surfaces. In conclusion, first, the extracted road surface points are used to improve the computational efficiency and completeness in the process of road marking extraction. Additionally, the smooth boundaries of road surfaces are extracted for creating HD maps. Finally, small cracks still exist, due to the damages to road surfaces.

Road markings indicate higher illumination than pavement with respect to intensity information. Accordingly, the GRF intensity imagery with 4-cm resolution was generated from extracted road surface points using the IDW interpolation algorithm. The multithreshold method was afterward performed to extract road markings. A series of experiments were carried out to refine related parameters, including the numbers of bins B_n, the intensity values of road markings I_i, and the distances between vehicle trajectories to the road markings D_i. In this paper, $k = 20$ was set in the SOR filter based on a collection of experiments. Fig. 8(c) shows the road marking extraction results after noise removal. Consequently, most of the discrete noises were successfully eliminated, and the road markings were completely and effectively extracted. However, some painted road markings covered by obstacles (e.g., cars, cyclists, and pedestrians), have low reflectance and intensities. Such road markings were difficult to be detected and extracted. Besides, due to the inevitable damage of painted markings caused by moving heavy-duty trucks and weather conditions (e.g., salt-fog corrosion), there exist small gaps in the extracted road markings [see the boxes labeled A in Fig. 8(c)]. Such defects of the road markings result in a failure at the stage of curve fitting, thus causing errors in driving line generation.

Based on the Gaussian point density distributions and the resolution of the generated noise-removed point clouds, a collection of experiments by adjusting two critical parameters: The range of scanning distances d_e, and the width of a road marking w_c, were implemented using the CEC algorithm. According to [56], $w_c \geq 15$ cm and $d_e \geq 100$ cm were determined. In this paper, all of horizontal curves in MLS test datasets were regarded as circular horizontal curves, and the mean curvatures of horizontal curves were smaller than 0.15 using a kernel size of 0.2 m. Due to the existence of noncurved road markings (e.g., hatch markings) and the discontinuity of broken centerlines, the mean curvature of each test dataset was overestimated. In order to increase computational efficiency at the stage of driving line generation, the noncurved road markings (e.g., zebra crossings) were successfully removed by using the cubic spline curve fitting algorithm.

Driving lines at horizontally curved road sections were generated, based on the best-fitting functions of curved road markings. The equidistant distance from the edge line, lane line, or centerline to the driving line d was used, d was set to be ranging from 2.0 to 2.5 m in this study according to the road design regulations. Fig. 8(d) indicates the driving line generation results. The final results demonstrate that all driving lines at horizontal curves can be successfully generated on the six test datasets by using the proposed algorithms. As can be perceived, most of the generated driving lines are located at central locations between lane lines and edge lines (or centerlines), which provides the best-matching driving lines that meet the requirements of HD maps and ensure traffic safety for AVs. Table III shows the computed variables of a generated driving line in the test Dataset I. However, these algorithms are still affected by disconnectivity of edge lines or centerlines. As shown by the box labeled B in Fig. 8(d), two generated driving lines interruption occur, due to the presence of a zebra crossing and the discontinuities of edge lines and centerlines.

TABLE III
VARIABLES OF A GENERATED DRIVING LINE OF TEST DATASET I

Symbols	Descriptions	Units	Values
D	Degree of curvature	degree	12.4
Δ	Subtended angle of curve (PC to PT)	degree	10.49
R	Radius of curve (measured to centerline)	ft	462.06
L	Length of curve (measured along centerline)	ft	84.65
T	Tangent length	ft	42.42
M	Middle ordinate	ft	5.94
LC	Length of long chord (from PC to PT)	ft	84.48
E	External distance	ft	5.83

C. Accuracy Assessment of Road Marking Extraction

In this paper, the accuracy assessment mechanism of road marking extraction is based on the differences between the extracted road markings and the manually created reference data. According to [59], the following three criteria were used to perform the accuracy evaluation: Recall, precision, and F1-score. The expressions are given as follows: Recall $= t_p/(t_p + f_n)$, precision $= t_p/(t_p + f_p)$, and $F1 - \text{score} = 2 * \text{recall} * \text{precision}/(\text{recall} + \text{precision})$, where t_p denotes true positive, f_p indicates false positive, and f_n presents false negative classification. The recall shows how complete the extracted road markings are, while the precision describes what percentage of the extracted road markings are valid. In addition, F1-score represents an overall score with the integration between recall and precision. In this paper, t_p indicates the number of road marking points that are correctly classified, f_p indicates the number of noise points is misclassified as road marking points, and f_n is the number of road marking points is misclassified as noise points.

Table IV presents the quantitative accuracy assessment of the road marking extraction results. Consequently, an average recall, precision, and F1-score of 90.79%, 92.94%, and 91.85%, respectively, were achieved for all of six test datasets. The low performance occurs in Dataset IV, because the SOR filter cannot eliminate all outliers in the processing of road marking extraction. Besides, some MLS points belonging to road surfaces were misclassified into road markings due to the improper threshold selection, as illustrated in black circles in the right zoom-in view of Fig. 9. Moreover, the proposed equalization method at

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE IV
ACCURACY ASSESSMENT OF ROAD MARKING EXTRACTION

Dataset	Recall(%)	Precision(%)	F1-score(%)
I	90.06	92.53	91.28
II	92.81	94.76	93.77
III	90.33	93.41	91.84
IV	88.02	89.59	88.80
V	92.71	94.97	93.83
VI	90.82	92.36	91.58
Average	90.79	92.94	91.85

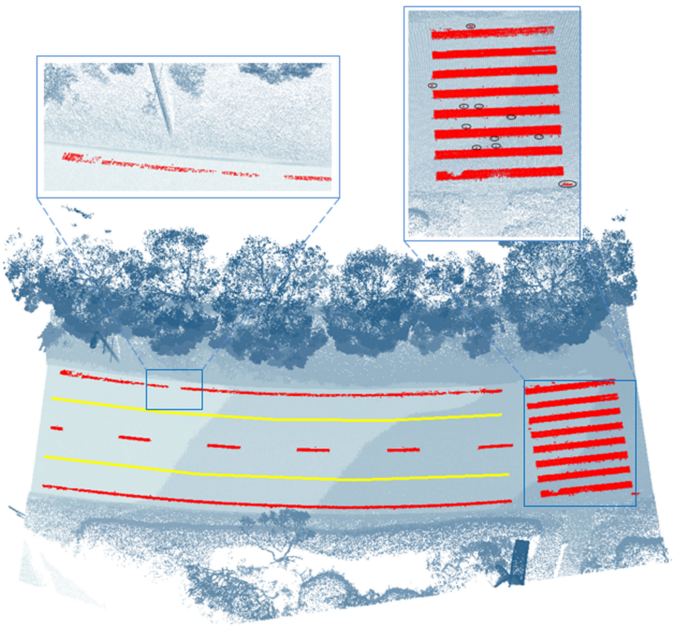


Fig. 9. Two close views of road marking extraction results from Dataset IV.

the stage of GRF intensity imagery generation, lessens the influence of intensity imbalance that might result in the decline of recall. The value of precision is larger than that of recall for each dataset, which demonstrates that certain road marking points were misclassified as road surface points, as shown in the left zoom-in view in Fig. 9. The inevitable errors evolving in the process of creating manually interpreted reference data also result in accuracy reduction for road marking extraction. Additionally, the sizes of manually interpreted reference data are larger than the road markings due to the decay of them. Thus, the overall performance of the proposed road marking extraction algorithms is underestimated in the final results.

D. Comparative Study

Furthermore, a comparative study was carried out concentrating on the extracted road marking results by using the proposed algorithms and other methods, i.e., Chen *et al.* [48] and Yu *et al.* [12]. Since Test Datasets I, VI, and V contain many types of road markings, they were selected to evaluate the performance of such methods. MLS point clouds were directly used in the process of road marking extraction in both Chen's [48] and Yu's [12] methods. Chen's [48] method mainly focuses on the lane marking extraction along the moving direction of the vehicle, resulting in limitations at the stage of complex and semantic road marking extraction (e.g., arrows, words, and curved road markings). Meanwhile, based on deep learning and PCA methods, Yu's [12] approach can be applied in the extraction of any types of road markings, but it has limitations in the process of curved road marking extraction, and it requires rich prior knowledge.

TABLE V
ROAD MARKING EXTRACTION RESULTS OBTAINED BY USING DIFFERENT METHODS

Method	Dataset	Recall(%)	Precision(%)	F1-score(%)
Yu et al. [12]	I	81.83	90.47	85.93
	IV	76.27	91.88	83.35
	V	85.34	91.17	88.16
Chen et al. [48]	I	72.95	90.89	80.94
	IV	74.95	92.02	82.61
	V	82.84	90.95	86.71
In this paper	I	90.06	92.53	91.28
	IV	88.02	89.59	88.80
	V	92.71	94.97	93.83

The overall performance of the proposed method in this paper and other two methods were evaluated, based on the quantitative assessment (i.e., recall, precision, and F1-score). As shown in Table V, Chen's [48] method achieved an average recall, precision, and F1-score of 76.91%, 91.27%, and 83.42%, respectively; Yu's [12] method achieved an average recall, precision, and F1-score of 81.15%, 91.17%, and 85.81%, respectively; while the proposed method in this paper achieved an average recall, precision, and F1-score of 90.26%, 92.36%, and 91.30%, respectively. Since Chen's [48] method cannot completely extract complex road markings (e.g., words) especially for horizontally curved road sections, it had a low recall, precision, and F1-score in Test Datasets I, IV, and V. Meanwhile, Yu's [12] method had a relatively low F1-score in these three test datasets, since its road surface extraction and road marking extraction methods were both depending on the fixed-interval straight trajectory. Compared to both Chen's [48] and Yu's [12] methods, the proposed method can achieve a better performance than their methods, in terms of both recall and precision at horizontally curved road sections. Nevertheless, road marking points were not effectively extracted due to the loss of reflectance, intensity deduction, and the decay of road markings. Road surface points with high reflectance might be incorrectly identified as road marking points. Additionally, outliers around some road markings were not completely eliminated by using the SOR filter, and the extraction accuracy was therefore affected.

E. Accuracy Assessment of Driving Line Generation

According to [60], the generated driving lines were evaluated using the buffer overlay statistics method. This method specifies a different buffer size, based on reference lines and reference orthoimagery through overlaying and statistics. According to the georeferenced MLS point clouds, UAV orthoimagery with 4 cm resolution was calibrated first as the reference

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

TABLE VI
ACCURACY ASSESSMENT OF GENERATED DRIVING LINES FROM TEST DATASET I

Buffers \ Driving lines	1		2	
	Recall(%)	Precision (%)	Recall(%)	Precision (%)
5 cm	74.82	24.48	70.98	28.23
10 cm	92.33	7.24	91.27	8.32
15 cm	100.00	0	100.00	0

data. Then, reference driving lines were manually created on the UAV orthoimagery using ArcGIS Desktop v10.2.2. In order to conduct both visual inspection and quantitative evaluation of driving lines, buffering and overlaying is performed iteratively. Given a number (n) of buffer sizes $S_i (1 \leq i \leq n)$, the following three procedures were conducted. First, creating S_i buffers in the reference data R based on the manually created driving lines, and defining the generated buffer zones S_iR. Next, overlay the generated driving lines L with S_iR to produce mixed datasets LS_iR. Third, compute the total length of the generated driving lines in L, the sum of the length of the generated driving lines from L inside S_iR in LS_iR, and the sum of the length of the generated driving lines from L outside S_iR in LS_iR, respectively.

Accordingly, the accuracy of the generated driving lines was assessed, based on recall and miscoding. The miscoding represents what percentage of the generated driving lines are located outside of the reference buffers. It is calculated by

$$\text{Miscoding} = \frac{\text{Length}\,(L \text{ outside } S_iR \text{ in } LS_iR)}{\text{Length}\,(L)} \times 100. \quad (11)$$

Reference buffer zones with different ranges established by the manually interpreted driving lines, were generated and overlapped with the generated driving lines on the georeferenced high-resolution UAV orthoimagery.

Table VI indicates the quantitative assessment in both recall and miscoding of the generated driving lines from Test Dataset I. Reference buffers with the width of 5, 10, and 15 cm were established to evaluate the performance of proposed driving line generation algorithms. As a consequence, the proposed algorithms are capable of achieving an average recall of 72.90% in 5 cm level reference buffers, 91.80% in 10 cm level reference buffers, and 100.00% in 15 cm level reference buffers for two generated driving lines. The values of miscoding decrease with the increased width of reference buffers, which demonstrates that the majority of generated driving lines are located within the precision permissible reference buffers. Experimentally, the proposed algorithms provide a 15 cm level localization accuracy in order to meet the quality requirements of HD maps, and improve the safety of autonomous diving [8].

Additionally, Fig. 10 shows the final results of generated driving lines within 5, 10, and 15 cm width of reference buffers. It is identified that black rectangles in Fig. 10 indicate miscoding parts, while setting the width of reference buffers to be 5 and

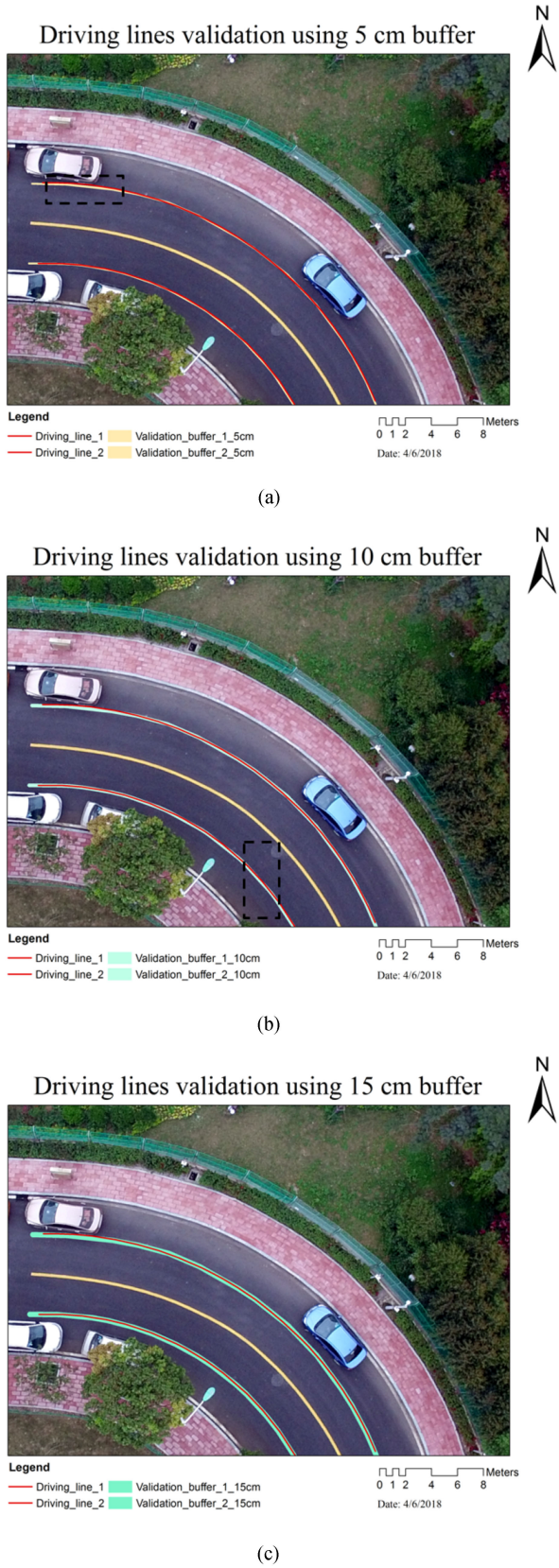


Fig. 10. Driving line generation results from Test Dataset I within reference buffers. (a) Generated driving line validation using 5-cm reference buffer. (b) Generated driving line validation using 10-cm reference buffer. (c) Generated driving line validation using 15-cm reference buffer.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

10 cm, respectively. Moreover, it is worth noting that the generated driving lines are completely located within the reference buffers with the width of 15 cm.

V. Conclusion

The generated driving lines are vital components in HD maps. Such maps are integrated and preloaded on AVs to cooperate with onboard sensors for precise localization and navigation services. In this paper, we have presented a method for driving line generation using MLS point clouds at horizontally curved road sections. Based on the curvature analysis, the curb-based road surface extraction and multithreshold road marking extraction algorithms were revised to apply to horizontally curved road sections. Furthermore, by using a cubic spline curve fitting approach and considering road construction regulations, we proposed a novel method for driving line generation from MLS point clouds.

The feasibility and validity of the proposed algorithms were evaluated on six MLS point cloud datasets. The quantitative evaluation indicated that the revised road marking extraction algorithm achieved an average recall, precision, and F1-score of 90.79%, 92.94%, and 91.85%, respectively. Through the comparative studies, it demonstrated that the revised method outperformed the other two existing methods in completely and correctly extracting road markings from various horizontal curves. Based on the buffer overlay statistics method, the proposed algorithm was capable of achieving an average recall of 72.90% within 5 cm level reference buffers, 91.80% within 10 cm level reference buffers, and 100.00% within 15 cm level reference buffers in generating driving lines, with the assistance of 4-cm resolution UAV orthoimagery. Thus, the proposed algorithms can successfully and effectively generate driving lines at horizontally curved road sections from six MLS point cloud datasets with 15-cm localization accuracy.

In conclusion, the proposed method is capable of efficiently generating driving lines at horizontally curved road sections from MLS point clouds to provide highly accurate localization services. We have provided a reliable solution to demonstrate that MLS point clouds can be effectively utilized to extract on-road information (i.e., road markings) and generate driving lines to support the development of HD maps and autonomous driving. However, the proposed driving line generation method mainly concentrate on the circular horizontal curves. It cannot deal with complex and spiral road curves, such as intersections, which need to be further studied.

Acknowledgment

The authors would like to express deep gratitude and appreciation to the anonymous reviewers for their valuable comments.

References

[1] J.-F. Bonnefon, A. Shariff, and I. Rahwan, "The social dilemma of autonomous vehicles," *Science*, vol. 352, no. 6293, pp. 1573–1576, 2016.

[2] D. J. Fagnant and K. Kockelman, "Preparing a nation for autonomous vehicles: Opportunities, barriers and policy recommendations," *Transp. Res. Part. A—Policy Pract.*, vol. 77, pp. 167–181, 2015.

[3] N. H. T. S. Administration, "National highway traffic safety administration," in *Automated Driving Systems 2.0: A Vision for Safety*, Washington, DC: US Department of Transportation, DOT HS, 812, 2017, p. 442.

[4] E. Guizzo, "How Googles' self-driving car works," 2012. [Online]. Available: https://spectrum.ieee.org/automaton/robotics/artificial-intelligence/how-google-self-driving-car-works

[5] J. Dokic, B. Müller, and G. Meyer, "European roadmap smart systems for automated driving," *EPoSS*, p. 39, 2015.

[6] J. Levinson, M. Montemerlo, and S. Thrun, "Map-based precision vehicle localization in urban environments," vol. 4, pp. 344–352, 2007.

[7] D. Bétaille and R. Toledo-Moreo, "Creating enhanced maps for lane-level vehicle navigation," *IEEE Trans. Intell. Transp. Syst.*, vol. 11, no. 4, pp. 786–798, Dec. 2010.

[8] J. Du, "Enhanced digital mapping project-final report," U.S. Department of Transportation, Washington, DC, USA, Tech. Rep. DTFH61-01-X-00014, 2004.

[9] I. Puente, H. González-Jorge, J. Martínez-Sánchez, and P. Arias, "Review of mobile mapping and surveying technologies," *Measurement*, vol. 46, no. 7, pp. 2127–2145, 2013.

[10] N. Haala, M. Peter, J. Kremer, and G. Hunter, "Mobile LiDAR mapping for 3D point cloud collection in urban areas—A performance test," *ISPRS. Arch.*, vol. 37, pp. 1119–1127, 2008.

[11] G. Vosselman and H. Maas, *Airborne and Terrestrial Laser Scanning*. Scotland, U.K.: Whittles, 2010.

[12] Y. Yu, J. Li, H. Guan, F. Jia, and C. Wang, "Learning hierarchical features for automated extraction of road markings from 3-d mobile lidar point clouds," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 2, pp. 709–726, Feb. 2015.

[13] H. G. Seif and X. Hu, "Autonomous driving in the iCity—HD maps as a key challenge of the automotive industry," *Engineering*, vol. 2, no. 2, pp. 159–162, 2016.

[14] H. Guan, J. Li, S. Cao, and Y. Yu, "Use of mobile LiDAR in road information inventory: A review," *Int. J. Image Data Fusion*, vol. 7, no. 3, pp. 219–242, 2016.

[15] D. Zai *et al.*, "3-d road boundary extraction from mobile laser scanning data via supervoxels and graph cuts," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 3, pp. 802–813, Mar. 2018.

[16] L. Ma, Y. Li, J. Li, C. Wang, R. Wang, and M. Chapman, "Mobile laser scanned point-clouds for road object detection and extraction: A review," *Remote Sens.*, vol. 10, no. 10, p. 1531, 2018.

[17] C. Guo, K. Kidono, J. Meguro, Y. Kojima, M. Ogawa, and T. Naito, "A low-cost solution for automatic lane-level map generation using conventional in-car sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 8, pp. 2355–2366, Aug. 2016.

[18] M. Javanmardi, E. Javanmardi, Y. Gu, and S. Kamijo, "Towards high-definition 3d urban mapping: Road feature-based registration of mobile mapping systems and aerial imagery," *Remote Sens.*, vol. 9, no. 10, pp. 975–1005, 2017.

[19] G. Khan, A. Bill, M. Chitturi, and D. Noyce, "Horizontal curves, signs, and safety," *Transport. Res. Board.*, no. 2279, pp. 124–131, 2012.

[20] O. R. S. A. Reports, "Preliminary 2016 Ontario road safety annual report selected statistics," Toronto, Canada: Elsevier, 2016, p. 4.

[21] K. Jo and M. Sunwoo, "Generation of a precise roadway map for autonomous cars," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 3, pp. 925–937, Jun. 2014.

[22] B. Riveiro, H. González-Jorge, J. Martínez-Sánchez, L. Díaz-Vilariño, and P. Arias, "Automatic detection of zebra crossings from mobile lidar data," *Opt. Laser Technol.*, vol. 70, pp. 63–70, 2015.

[23] P. Kumar, P. Lewis, C. P. McElhinney, P. Boguslawski, and T. McCarthy, "Snake energy analysis and result validation for a mobile laser scanning data-based automated road edge extraction algorithm," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 2, pp. 763–773, Feb. 2017.

[24] A. Hervieu and B. Soheilian, "Semi-automatic road/pavement modeling using mobile laser scanning," *ISPRS Ann. Photogrammetry, Remote Sens. Spatial Inf. Sci.*, vol. 2, pp. 31–36, 2013.

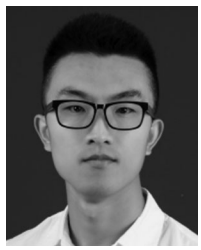
[25] M. Cheng, H. Zhang, C. Wang, and J. Li, "Extraction and classification of road markings using mobile laser scanning point clouds," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 3, pp. 1182–1196, Mar. 2017.

[26] P. Huang, M. Cheng, Y. Chen, H. Luo, C. Wang, and J. Li, "Traffic sign occlusion detection using mobile laser scanning point clouds," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 9, pp. 2364–2376, Sep. 2017.

[27] S. Pu, M. Rutzinger, G. Vosselman, and S. O. Elberink, "Recognizing basic structures from mobile laser scanning data for road inventory studies," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 6, pp. 28–39, 2011.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

[28] H. Guan, J. Li, Y. Yu, C. Wang, M. Chapman, and B. Yang, "Using mobile laser scanning data for automated extraction of road markings," *ISPRS J. Photogramm. Remote Sens.*, vol. 87, pp. 93–107, 2014.
[29] H. Wang *et al.*, "Road boundaries detection based on local normal saliency from mobile laser scanning data," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 10, pp. 2085–2089, Oct. 2015.
[30] Y. Yu, J. Li, H. Guan, and C. Wang, "Automated extraction of urban road facilities using mobile laser scanning data," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 4, pp. 2167–2181, Aug. 2015.
[31] H. Wang, C. Wang, H. Luo, P. Li, Y. Chen, and J. Li, "3-D point cloud object detection based on supervoxel neighborhood with Hough forest framework," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 4, pp. 1570–1581, Apr. 2015.
[32] F. Wu *et al.*, "Rapid localization and extraction of street light poles in mobile LiDAR point clouds: A supervoxel-based approach," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 2, pp. 292–305, Feb. 2017.
[33] Y. Zhou *et al.*, "A fast and accurate segmentation method for ordered LiDAR point cloud of large-scale scenes," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 11, pp. 1981–1985, Nov. 2014.
[34] B. Yang, L. Fang, and J. Li, "Semi-automated extraction and delineation of 3D roads of street scene from mobile laser scanning point clouds," *ISPRS J. Photogramm. Remote Sens.*, vol. 79, pp. 80–93, 2013.
[35] H. Guan, J. Li, Y. Yu, and C. Wang, "Rapid update of road surface databases using mobile LiDAR: Road-markings," in *Proc. 5th Int. Conf. Geo-Inf. Technol. Natural Disaster Manage.*, 2013, pp. 124–129.
[36] J. Hernández and B. Marcotegui, "Filtering of artifacts and pavement segmentation from mobile LiDAR data," *ISPRS. Arch.*, vol. 38, no. 3–8, pp. 329–332, 2009.
[37] H. Wang *et al.*, "Automatic road extraction from mobile laser scanning data," in *Proc. Int. Conf. Comput. Vision Remote Sens.*, 2012, pp. 136–139.
[38] C. McElhinney, P. Kumar, C. Cahalane, and T. McCarthy, "Initial results from European Road Safety Inspection (EURSI) mobile mapping project," *ISPRS. Arch.*, vol. 38, no. 5, pp. 440–445, 2010.
[39] B. Yang, L. Fang, Q. Li, and J. Li, "Automated extraction of road markings from mobile LiDAR point clouds," *Photogramm. Eng. Remote Sens.*, vol. 78, no. 4, pp. 331–338, 2012.
[40] J. Tan, J. Li, X. An, and H. He, "Robust curb detection with fusion of 3D-LiDAR and camera data," *Sensors*, vol. 14, no. 5, pp. 9046–9073, 2014.
[41] A. Boyko and T. Funkhouser, "Extracting roads from dense point clouds in large scale urban environment," *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 6, pp. 2–12, 2011.
[42] A. Jaakkola, J. Hyyppä, H. Hyyppä, and A. Kukko, "Retrieval algorithms for road surface modelling using laser-based mobile mapping," *Sensors*, vol. 8, no. 9, pp. 5238–5249, 2008.
[43] P. Kumar, C. P. McElhinney, P. Lewis, and T. McCarthy, "Automated road markings extraction from mobile laser scanning data," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 32, pp. 125–137, 2014.
[44] C. Wen, X. Sun, J. Li, C. Wang, Y. Guo, and A. Habib, "A deep learning framework for road marking extraction, classification and completion from mobile laser scanning point clouds," *ISPRS J. Photogramm. Remote Sens.*, vol. 147, pp. 178–192, 2019.
[45] L. Smadja, J. Ninot, and T. Gavrilovic, "Road extraction and environment interpretation from lidar sensors," *ISPRS. Arch.*, vol. 38, no. 3, pp. 281–286, 2010.
[46] H. Guan, J. Li, Y. Yu, M. Chapman, and C. Wang, "Automated road information extraction from mobile laser scanning data," *IEEE Trans. Intell. Transp. Syst.*, vol. 16, no. 1, pp. 194–205, Feb. 2015.
[47] N. Ohtsu, "An automatic threshold selection method based on discriminate and least squares criteria," *Trans. Inst. Electron. Commun. Eng. Jpn.*, vol. 63, no. 4, pp. 349–356, 1980.
[48] X. Chen, B. Kohlmeyer, M. Stroila, N. Alwar, R. Wang, and J. Bach, "Next generation map making: Geo-referenced ground-level lidar point clouds for automatic retro-reflective road feature extraction," in *Proc. 17th ACM SIGSPATIAL Int. Symp. Advances Geographic Inf. Syst.*, 2009, pp. 488–491.
[49] R. Huang, H. Liang, P. Zhao, B. Yu, and X. Geng, "Intent-estimation-and motion-model-based collision avoidance method for autonomous vehicles in urban environments," *Appl. Sci.*, vol. 7, no. 5, p. 457, 2017.
[50] M. M. Atia, S. Liu, H. Nematallah, T. B. Karamat, and A. Noureldin, "Integrated indoor navigation system for ground vehicles with automatic 3D alignment and position initialization," *IEEE Trans. Veh. Technol.*, vol. 64, no. 4, pp. 1279–1292, Apr. 2015.
[51] V. Gikas and J. Stratakos, "A novel geodetic engineering method for accurate and automated road/railway centerline geometry extraction based on the bearing diagram and fractal behavior," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 1, pp. 115–126, Mar. 2012.
[52] K. Jo, M. Lee, and M. Sunwoo, "Road slope aided vehicle position estimation system based on sensor fusion of GPS and automotive onboard sensors," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 1, pp. 250–263, Jan. 2016.
[53] A. Habib, Y.-J. Lin, R. Ravi, T. Shamseldin, and M. Elbahnasawy, "LiDAR-based mobile mapping system for lane width estimation in work zones," Indiana Department of Transportation, Indiana Polis, IN, USA, Rep. FHWA/IN/JTRP-2018/10, 2018.
[54] A. Holgado-Barco, D. González-Aguilera, P. Arias-Sanchez, and J. Martinez-Sanchez, "Semiautomatic extraction of road horizontal alignment from a mobile lidar system," *Comput.-Aided Civil Infrastructure Eng.*, vol. 30, no. 3, pp. 217–228, 2015.
[55] K. Lakakis, P. Savvaidis, and T. Wunderlich, "Evaluation of a low-cost mobile mapping and inspection system for road safety classification," *Amer. J. Geogr. Inf. Syst.*, vol. 2, no. 1, pp. 6–14, 2013.
[56] A. Holgado-Barco, D. Gonzalez-Aguilera, P. Arias-Sanchez, and J. Martinez-Sanchez, "An automated approach to vertical road characterisation using mobile LiDAR systems: Longitudinal profiles and cross-sections," *ISPRS J. Photogramm. Remote Sens.*, vol. 96, pp. 28–37, 2014.
[57] A. Nurunnabi, G. West, and D. Belton, "Robust locally weighted regression techniques for ground surface points filtering in mobile laser scanning three dimensional point cloud data," *IEEE Trans. Geosci. Remote Sensing.*, vol. 54, no. 4, pp. 2181–2193, Apr. 2016.
[58] *Code for Design of Urban Road Engineering*, Standard CJJ37-2012, pp. 1–143.
[59] D. M. Powers, "Evaluation: From precision, recall and f-measure to roc, informedness, markedness and correlation," *J. Mach. Learn. Tech.*, vol. 2, no. 1, pp. 37–63, 2011.
[60] H. Tveite, "An accuracy assessment method for geographical line data sets based on buffering," *Int. J. Geogr. Inf. Sci.*, vol. 13, no. 1, pp. 27–47, 1999.



Lingfei Ma (S'18) received the B.Sc. and M.Sc. degrees in geomatics engineering from the University of Waterloo, Waterloo, ON, Canada, in 2015 and 2017, respectively. He is currently working toward the Ph.D. degree in photogrammetry and remote sensing with the Mobile Sensing and Geodata Science Laboratory, Department of Geography and Environmental Management, University of Waterloo.

His research interests include autonomous driving, mobile laser scanning, intelligent processing of point clouds, 3-D scene modeling, and machine learning.

Jonathan Li (M'00–SM'11) received the Ph.D. degree in geomatics engineering from the University of Cape Town, Cape Town, South Africa.

He is currently a Professor with the Departments of Geography and Environmental Management and System Design Engineering, University of Waterloo, Waterloo, ON, Canada. He has coauthored more than 300 publications, more than 150 of which were published in refereed journals, including the IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING, IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, *ISPRS Journal of Photogrammetry and Remote Sensing*, and *Remote Sensing of Environment*. His research interests include information extraction from mobile LiDAR point clouds and from earth observation images.

Dr. Li is the Chair of the ISPRS WG I/2 on LiDAR, Airborne and Spaceborne Optical Sensing (2016–2020), and the ICA Commission on Sensor-driven Mapping (2015–2019). He is an Associate Editor for the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS and IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

Ying Li received the B.Eng. degree in geomatics engineering from the Hefei University of Technology, China, in 2014, and M.Sc. degree in remote sensing from Wuhan University, China, in 2017. She is currently working toward the Ph.D. degree with the Mobile Sensing and Geodata Science Laboratory, Department of Geography and Environmental Management, University of Waterloo, ON, Canada.

Her research interests include autonomous driving, mobile laser scanning, intelligent processing of point clouds, geometric and semantic modeling, and augmented reality.

Zilong Zhong received the M.Eng. degree in electronic and communication engineering from Lanzhou University, Lanzhou, China, in 2014. He is currently working toward the Ph.D. degree in remote sensing with the University of Waterloo, Waterloo, ON, Canada.

His research interests include computer vision, deep learning, probabilistic graphical models, and their applications in the context of high-dimensional remotely sensed data.

Michael A. Chapman received the Ph.D. degree in photogrammetry from Laval University, Qubec City, QC, Canada.

He was a Professor with the Department of Geomatics Engineering, University of Calgary, Calgary, AB, Canada, for 18 years. Currently, he is a Professor of geomatics engineering with the Department of Civil Engineering, Ryerson University, Toronto, ON, Canada. He has authored or coauthored over 160 technical articles. His research interests include algorithms and processing methodologies for airborne sensors using GNSS/IMU, geometric processing of digital imagery in industrial environments, terrestrial imaging systems for transportation infrastructure mapping, and algorithms and processing strategies for biometrology applications.