3D Highway Curve Reconstruction From Mobile Laser Scanning Point Clouds

Zongliang Zhang, Jonathan Li, *Senior Member, IEEE*, Yulan Guo, *Member, IEEE*, Chenhui Yang, and Cheng Wang, *Senior Member, IEEE*

***Abstract*— The point clouds acquired by a vehicle-borne mobile laser scanning (MLS) system have shown great potential for many applications such as intelligent transportation systems, road infrastructure inventories, and high-definition (HD) maps to support the advanced driver-assistance systems (ADAS) and autonomous vehicles (AVs). This paper presents a novel two-step approach to automated detection and reconstruction of three-dimensional (3D) highway curves from MLS point clouds. However, when dealing with noisy, unstructured, dense point clouds, we often face some challenges, most notably in handling of the outliers introduced during road marking detection and in recognition of curve types during 3D curve reconstruction. Our approach is formed by two main algorithms: a detector based on intensity variance and a robust model fitting estimator. The experimental results obtained using both a virtual scan dataset and a real MLS dataset demonstrated that our approach is very promising in handling of the outliers and reconstruction of 3D road curves. Specifically, a relative accuracy of 0.6% has been achieved in estimation of circle radii based on the virtual scan dataset. A comparative study also showed that our road marking detection approach is more effective and more stable than state-of-the-art approaches.**

***Index Terms*— Highway geometric design, mobile laser scanning, 3D point cloud, curve parameter estimation, road marking detection.**

I. Introduction

THE geometry of highways has crucial impacts on highway construction, road safety, and driving comfort. Highway geometry is mainly determined by highway curves. The goal of highway curve reconstruction is to recover the as-built geometric information of the underlying curves of a highway. Consequently, highway curve reconstruction is useful for many applications such as digital mapping, autonomous driving, highway geometric design, traffic accident analysis, and highway as-built quality assessment [1].

There are two major types of data that can be used to reconstruct highway curves. The first type is ground surveys, mainly using real-time kinematic (RTK) Global Navigation Satellite Systems (GNSS) technique [2]. The second type is off-ground surveys using digital images acquired by the remote sensors mounted on aircrafts or satellites as well as computer vision and digital photogrammetric algorithms [3] [4]. Compared with the ground surveys, off-ground surveys can cover a larger geographic area in a cost-effective way, but its spatial accuracy of off-ground surveys is usually lower than that of ground surveys [5], [6].

A mobile mapping vehicle with onboard optical cameras and integrated GNSS and IMU (inertial measuring unit) system becomes a better solution to provide both higher accuracy and higher efficiency [7]. The trajectory data of the mobile mapping vehicle along with the images collected can be used to detect and reconstruct highway centerlines. In recent years, vehicle-borne mobile laser scanning (MLS) systems are able to rapidly acquire high-accuracy and high-density three-dimensional (3D) point clouds. Compared with its camera-based counterparts, MLS systems have considerable advantages for acquiring 3D point clouds covering roadways including centerlines [8]. The remarkable technical progress has been achieved in the research and development efforts in machine learning and deep learning algorithms for automated detection and extraction both road objects (e.g., road edges, lane lines, road markings, pavement cracks, manhole covers) and off-road 3D objects (e.g., traffic signs, street light poles, roadside tress, cars parking along roadsides) from noisy, unstructured, highly dense 3D MLS point clouds [9], [10]. Moreover, autonomous vehicles under research and development not only require road centerlines, but much more road features (e.g., edge lines, lane lines, road markings), which are key elements for future high-definition (HD) maps.

MLS point clouds have been used for a number of applications, such as road information extraction [11]–[13]. Several methods have also been proposed to use MLS data to reconstruct the geometry of highways. The rectangular regions between pavement markings were used in [14] to estimate super-elevations. The method proposed in [15] uses

Manuscript received March 18, 2018; revised September 29, 2018 and April 14, 2019; accepted September 30, 2019. This work was supported by the National Natural Science Foundation of China under Grant 41471379, Grant 61602499, Grant 61972435, and Grant 61471371. The Associate Editor for this article was D. Fernandez-Llorca. *(Corresponding author: Jonathan Li.)*

Z. Zhang is with the Fujian Key Laboratory of Sensing and Computing for Smart Cities and the Xiamen Key Laboratory of Geospatial Sensing and Computing, School of Information Science and Engineering, Xiamen University, Xiamen 361005, China, and also with the School of Basic Medical Sciences, Fujian Medical University, Fuzhou 350122, China (e-mail: zhangzongliang@stu.xmu.edu.cn).

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

Y. Guo is with the College of Electronic Science and Technology, National University of Defense Technology, Changsha 410073, China (e-mail: yulan.guo@nudt.edu.cn).

C. Yang and C. Wang are with the Fujian Key Laboratory of Sensing and Computing for Smart Cities and the Xiamen Key Laboratory of Geospatial Sensing and Computing, School of Information Science and Engineering, Xiamen University, Xiamen 361005, China (e-mail: chyang@xmu.edu.cn; cwang@xmu.edu.cn).

Digital Object Identifier 10.1109/TITS.2019.2946259

both the trajectory and scan angle to segment MLS data into blocks, and then applies a principal component analysis (PCA) algorithm to each block to estimate its slope and super-elevation. Both the intensity and scan angle were used in [16] to detect road markings, and the detected marking points was then used to calculate azimuth and curvature diagrams to estimate horizontal curve variables. The trajectory data was used to hierarchically segment MLS data into multiple blocks to estimate their slopes and super-elevations [17].

Highway curves are intrinsically 3D, as highways are 3D objects. Different from the aforementioned methods that separately reconstruct horizontal or vertical curves from MLS data, we reconstruct highway curves in a full 3D manner. Consequently, the proposed method is more suitable for the applications that prefer comprehensive 3D geometric information of highways (e.g., highway geometric design, traffic accident analysis, and highway as-built quality assessment).

The proposed method has two major steps. The first step is to detect road marking points from MLS data. A number of sophisticated methods have been proposed in the literature to detect road markings. For example, Yang et al. recently use a least squares algorithm to compute intensity gradient to detect road markings [18]. Since road marking detection is a preprocessing step in our method, the road marking detection method is expected to be as simple as possible. We propose to solely use the variance of intensity as feature to directly detect road markings from MLS data.

The second step is to reconstruct curves from road marking points. Several curve reconstruction approaches estimate the values of curve variables by assuming that the type of the curve is known. However, it is non-trivial to recognize the types of 3D highway curves. Note that, there are three types of elemental horizontal curves (i.e., line, circle, spiral), and two types of elemental vertical curves (i.e., line, parabola). 3D highway curves are combined by horizontal and vertical curves. Therefore, apart from composition curves, there are six types of elemental 3D highway curve types. We adopt the probabilistic program induction strategy [19] to perform curve reconstruction, as the strategy can automatically recognize the types of curves and estimate the values of the curve variables at the same time. Specifically, curve reconstruction is achieved by conducting model fitting for each type of curve.

Apart from curve type recognition, another challenge in curve reconstruction is brought by outliers detected in the road marking points. Road markings of interest are defined as lane-line markings in this paper, while outliers are points belonging to the other markings (e.g., arrows or stop lines), which are difficult to be removed. Outliers can severely undermine the performance of a least-squares estimation algorithm, which is used by many highway geometric design methods such as [2], [17], [15]. To robustly handle outliers, we adopted the approach proposed in [20] to perform model fitting. A geometric similarity estimator called mean measure (MM) plays the key role in that approach. The MM-estimator has very few parameters to tune, making it as convenient as a least-squares estimation algorithm.

The new main contribution of this study is automation in reconstruction of highway curves in 3D from MLS point clouds. Specifically, our method differs from existing methods in the following three aspects.

First, the curves reconstructed by our method are confined by highway geometric design standard, making our method useful not only for digital mapping but also for highway as-built quality assessment. In contrast, many existing curve reconstruction methods are not suitable for highway as-built quality assessment as they are proposed to reconstruct generic curves (e.g., B-splines [21]).

Second, our method uses MLS data to achieve curve reconstruction. In contrast to vehicle trajectory data, MLS data have richer and more reliable information for highways.

Third, our method reconstructs highway curves in 3D, whereas most existing methods separately reconstruct 2D highway curves. Note that, in recent years, the highway geometric design community tends to design new highway curves in 3D [22]–[24] [25]. In contrast to 2D methods, our method can be painlessly extended to handle such new 3D highway curves.

The rest of this paper is organized as follows. Section II reviews related work. Section III describes the proposed method. Section IV presents and discusses the experimental results. Section V concludes the paper.

II. Related Work

In this section, we review some existing studies related to our work.

A. Road Marking Detection

Road markings, especially lane-line markings, can well represent the geometry of highways. Therefore, it is natural to use road marking points to reconstruct highway curves. However, it is challenging to detect markings from MLS data.

As painted on pavement with high reflective materials, the intensities of road markings are larger than that of the surrounding road surfaces. However, the intensities of different road marking points are usually very different, because their intensity values also depend on other factors such as incident angles and ranges [26]. Consequently, a single intensity threshold was used in [27]–[29] [30], but showed insufficiency in detection of all road markings points [18]. A number of advanced methods were thus proposed to address the problem. In [31], [32], [13], an inverse distance weighted interpolation algorithm was used to overcome the range effect of intensity to detect road markings. In [33] [26], the input point cloud was segmented into several segments, road markings were then detected from each segment by a thresholding algorithm.

It can be concluded that, existing road marking detection methods are more or less not straightforward to implement, making them inconvenient to use. In contrast, our proposed detection method is considerably easier to implement and is able to handle MLS data with sufficiently high quality.

B. Robust Curve Fitting

Many methods have been proposed to perform curve fitting on relatively clean data, such as [21] [34]. However, it is

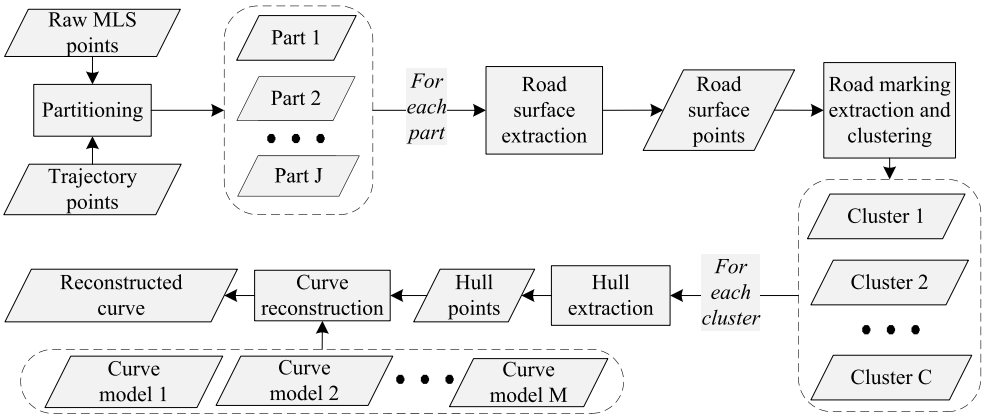


Fig. 1. Pipeline of the proposed method.

more challenging to perform robust curve fitting on corrupted data. One of the most popular robust curve fitting paradigms is random sample consensus (RANSAC) [18], [35] [36], [37]. The RANSAC-based methods perform model fitting by maximizing the inlier number. An extra parameter called inlier threshold is thus required to tune in RANSAC-based methods, making them less convenient than a least-squares estimation algorithm or the MM-estimator [20].

III. Method

The input to our method consists of a raw MLS point cloud dataset and its corresponding trajectory point set. A raw MLS point has four-dimensional (4D) properties, including three location coordinates (x, y, z), and one intensity (i) value. The GNSS-IMU position and orientation system (POS) mounted on the MLS vehicle can provide accurate vehicle's trajectory data (x, y, z) in real time. As shown in Fig. 1, our method is composed of five steps: raw data partitioning, road surface detection, road marking detection and clustering, hull detection, and curve reconstruction.

First, the vehicle's trajectory data were used to partition raw MLS point clouds into several parts. Second, road surface points were then detected from each part using a smoothness-based region growing algorithm. Third, road marking points were detected from the road surface points by a thresholding algorithm according to the variance of intensity, and then the road marking points were grouped into several clusters by computing the Euclidean distance between a point and the cluster centroids. Fourth, hull points were detected from each cluster by the alpha-shape algorithm. Finally, curves were reconstructed in 3D from the hull points according to predefined curve models. These five steps are sequentially detailed in the following sections.

A. Raw Data Partitioning

A raw MLS point cloud dataset usually contains a large number of unstructured, uneven distributed, highly dense points, it is therefore difficult to process the whole point set simultaneously. A common way to handle this problem is to partition the raw MLS point set into several sub-point sets. In this paper, we utilized the vehicle's trajectory points to perform partitioning, as the trajectory points are organized in the order of acquisition time. Given the raw MLS point set D and the trajectory point set T, we first split T into f sub-sets $T_1, T_2, \cdots, T_f$. After splitting, each sub-set has β points except for the last sub-set T_f. We then partition D into f parts $D_1, D_2, \cdots, D_f$ as follows.

Given a point $\mathbf{p} = (x_\mathbf{p}, y_\mathbf{p}, z_\mathbf{p})$, its set of neighboring points in D is defined as:

$$N_\mathbf{p}^D = \{\mathbf{q} \in D \mid d(\mathbf{q}, \mathbf{p}) < r\}, \tag{1}$$

where $d(\cdot, \cdot)$ is the Euclidean distance between two points, and r is a distance threshold to determine neighbors and is fixed to 10 m in this step. Then the first part is defined as $D_1 = \bigcup_{\mathbf{p} \in T_1} N_\mathbf{p}^D$, and each part with $j > 1$ is defined as:

$$D_j = \bigcup\nolimits_{\mathbf{p} \in T_1} N_\mathbf{p}^D - \bigcup\nolimits_{\mathbf{q} \in T_{j-1}} N_\mathbf{q}^D. \tag{2}$$

Fig. 2 shows the consistent partition results obtained using the proposed approach, with which those points far away from roadways have been filtered out.

B. Road Surface Detection

Since the objects of interest in this paper are road markings on road surfaces, road surface points are useful for the detection of road markings. In general, most highway surfaces are smooth. Therefore, we use the smoothness-based region growing algorithm [38] to detect road surface points. For the highway surfaces, the closer to the laser scanners on the top of the vehicle, the higher density of the MLS point clouds. Consequently, with region growing, the road surface usually corresponds to the region with the largest number of points. This step requires two parameters: a distance threshold to determine neighbors to calculate the surface normal vectors, and a smoothness threshold to compare the surface normal vectors. In this paper, the smoothness threshold is set to 2°. Fig. 3 shows the results of detected road surfaces.

C. Road Marking Detection and Clustering

Road markings are the symbols and text painted on roadway surfaces. A road marking is usually brighter than its surroundings. Consequently, the intensity of a road marking

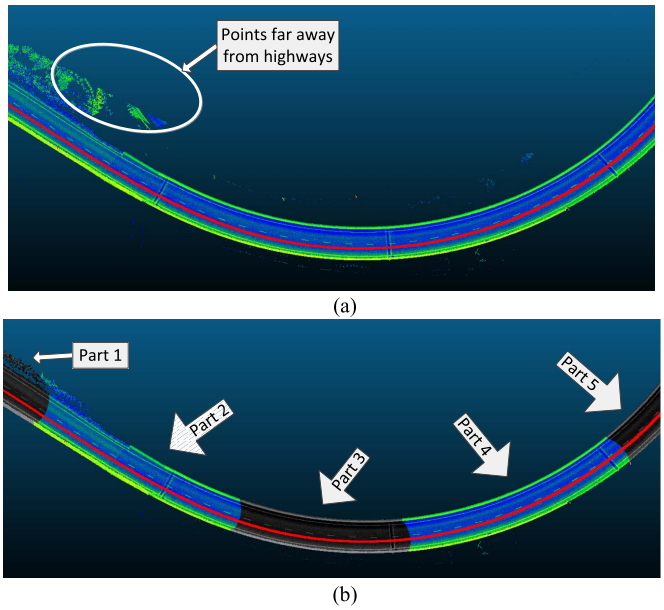


Fig. 2. Partition results (trajectory highlighted by red curve). (a) A raw MLS point set rendered with intensities in color, and (b) partitioning result, where Parts 1, 3, and 5 are rendered with intensities in gray.

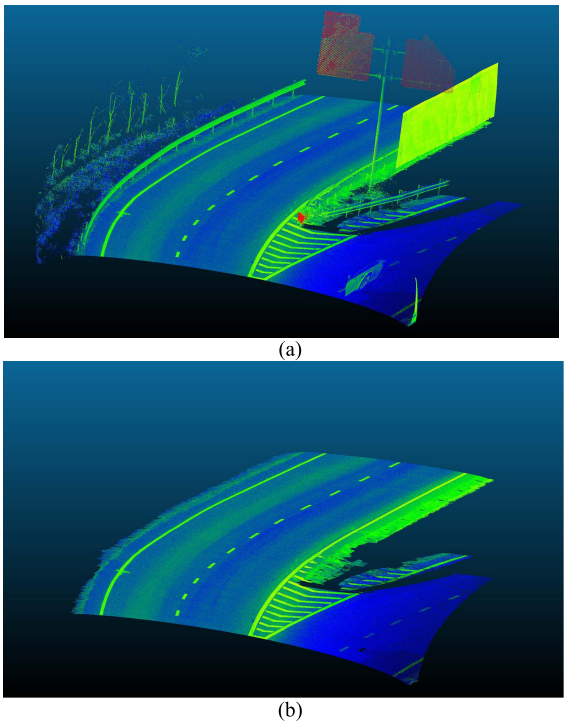


Fig. 3. Extracted road surfaces. (a) A section of a raw MLS point cloud dataset, and (b) Road surface points extracted from (a).

point is usually higher than that of its surrounding points. However, the intensities of different road marking points can vary significantly (Fig. 5b) due to many factors such as laser scanner incident angles and ranges [26]. Therefore, a simple approach such as intensity thresholding is usually insufficient to detect all road marking points.

Note that, the goal of this paper is to reconstruct highway curves, while the edge points of road markings are sufficient to achieve this goal. We therefore, propose a simple approach

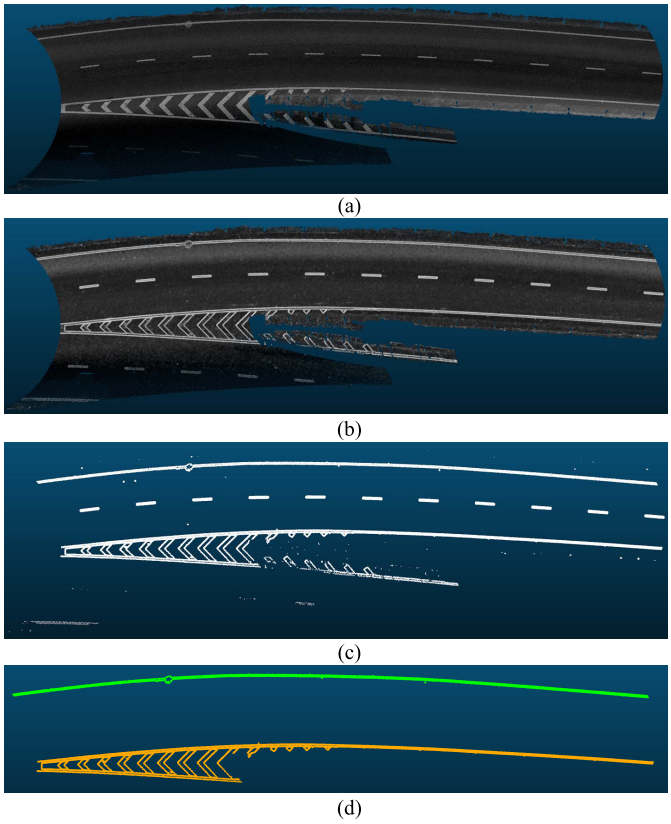


Fig. 4. Results of road marking detection and clustering. (a) Road surface rendered with intensities. (b) Road surface rendered with intensity variances. (c) Extracted road marking points. (d) Two largest clusters.

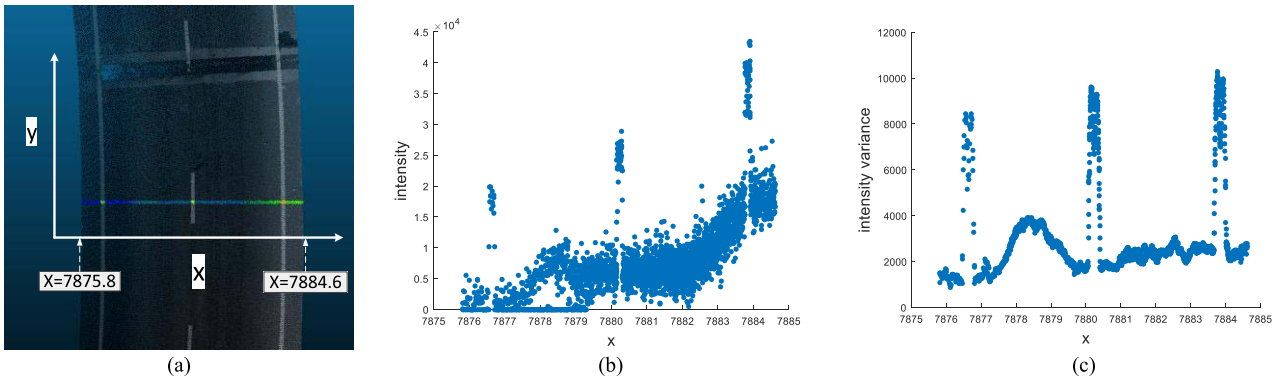
to detect edge points of road markings. Our insight is that, the intensity variance of neighboring points of an edge point is usually larger than that of a non-edge point. More importantly, the intensity variances are quite stable among different edge points (Fig. 5c). Therefore, the edge points can be detected by thresholding the intensity variances using a single threshold.

The intensity variance of a point $\mathbf{p} \in D$ is calculated as follows:

$$v_{\mathbf{p}} = \sqrt{\frac{1}{\left|N_{\mathbf{p}}^{D}\right|} \sum\nolimits_{\mathbf{q} \in N_{\mathbf{p}}^{D}} \left(i_{\mathbf{q}} - i_{\mathbf{p}}\right)^{2}}, \tag{3}$$

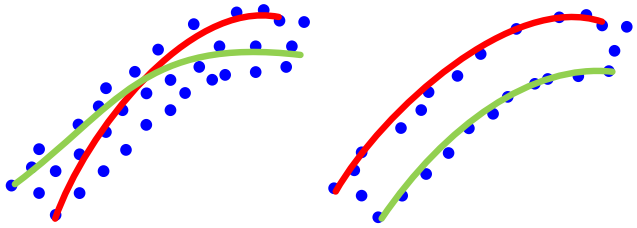
where D is the set of detected road surface points, i is the intensity of a point, $|\cdot|$ denotes the number of points in a point set, and $N_{\mathbf{p}}^{D}$ is the neighbors in D of $\mathbf{p}$ as defined in Eq. (1).

With the intensity variances, we use the Otsu's algorithm [39], [26] to compute a threshold to distinguish edge points from other points. That is, a point with variance higher than the threshold is detected as an edge point. Considering that the detected points usually do not belong to a single road marking, as a road surface usually contains several road markings, we further use the Euclidean clustering algorithm to group the detected points into several clusters. This step requires two parameters: a distance threshold to determine neighbors to calculate intensity variance, and a distance threshold for clustering. Fig. 4 shows a result of road marking detection and clustering.



(a) (b) (c)

Fig. 5. A comparison of intensity and intensity variance. (a) A slice (color) of a road surface (gray). (b) The intensities of the points in the slice. (c) The intensity variances of the slice.



(a) (b)

Fig. 6. Illustration of hull point extraction. (a) Road marking points (blue dots), and two curves (red and green) which fit well to road markings. (b) Hull points (blue dots) extracted from road markings, and two curves (red and green) which fit well to hull points.

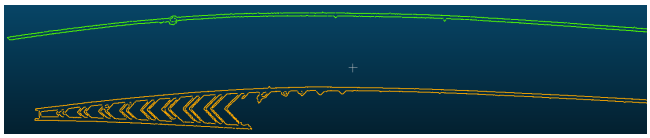


Fig. 7. Hull points extracted from Fig. 4d.

D. Hull Detection

As shown in Fig. 4c, the markings detected by the proposed approach are usually wider than the real markings. Consequently, the detected markings may be not accurate enough for precise curve reconstruction. Actually, it can be observed that even the real markings (Fig. 4a) may be too wide for precise curve reconstruction. As shown in Fig. 6a, the curves directly reconstructed from all points of a marking may not able to correctly represent the underlying shape of the marking. Therefore, we propose to use the hull points to address this problem (Fig. 6b). The hull points are detected by the alpha-shape algorithm [40] which requires only one parameter (i.e., the alpha). In this paper, we set the alpha to 0.1. Fig. 7 shows the hull points detected from Fig. 4d.

E. Curve Reconstruction

According to highway geometric design standard [41], there are three elemental types of horizontal curves and two elemental types of vertical curves (Table I). An elemental 3D highway curve is combined by a horizontal curve and a vertical curve. Consequently, there are six elemental types of 3D highway curves. For example, the mathematical model of a 3D curve combined by horizontal circle and vertical parabola is as follows:

TABLE I
MATHEMATICAL MODELS OF HIGHWAY CURVES

Horizontal Line	$\begin{cases} x(s) = x_0 + s\cos\mu_0 \\ y(s) = y_0 + s\sin\mu_0 \end{cases}$
Horizontal Circle	$\begin{cases} x(s) = x_0 + (\sin(\mu_0 + \kappa_0 s) - \sin\mu_0)/\kappa_0 \\ y(s) = y_0 - (\cos(\mu_0 + \kappa_0 s) - \cos\mu_0)/\kappa_0 \end{cases}$
Horizontal Spiral	$\begin{cases} x(s) = x_0 + \int_0^s \cos(\mu_0 + \kappa_0 t + \frac{1}{2}\psi t^2)dt \\ y(s) = y_0 + \int_0^s \sin(\mu_0 + \kappa_0 t + \frac{1}{2}\psi t^2)dt \end{cases}$
Vertical Line	$z(s) = z_0 + s\xi_0$
Vertical Parabola	$z(s) = z_0 + s\xi_0 + \frac{1}{2}\eta s^2$

$$\begin{cases} x(s) = x_0 + (\sin(\mu_0 + \kappa_0 s) - \sin\mu_0)/\kappa_0 \\ y(s) = y_0 - (\cos(\mu_0 + \kappa_0 s) - \cos\mu_0)/\kappa_0 \\ z(s) = z_0 + s\xi_0 + \frac{1}{2}\eta s^2 \end{cases} \quad (4)$$

This model is expressed in a parametric form. The parameters $\in [0, h]$, where h is the arc-length of the curve. Including h, this model has 9 variables. The variables x_0, y_0 and z_0 define the start location, μ_0 is the start horizontal azimuth, ξ_0 is the start slope, η is the vertical curvature, κ_0 corresponds to the start horizontal curvature, and ψ is the horizontal curvature change rate. It is worth noting that curvature is defined as reciprocal of radius.

The goal of curve reconstruction is to find a curve that best fits a given data point set. In this paper, the curve is defined by one of the six curve models. To achieve this goal, we first conduct model fitting for each curve model to find a candidate curve, the best curve is then selected from the candidate curves. The pipeline of our curve reconstruction method is shown in Fig. 8. The model fitting process is crucial for the curve reconstruction. As the hull points may contain many outliers (Fig. 7), a robust model fitting method [20] is

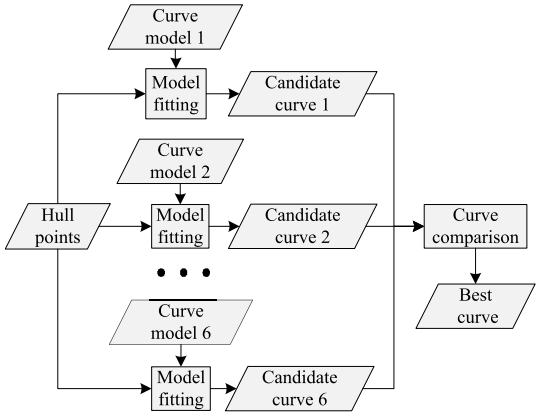


Fig. 8. The pipeline of curve reconstruction.

TABLE II
THE RANGES OF MODEL VARIABLES

Variable	μ_0	ξ_0	h	κ_0	ψ	η
Min value	0	-0.2	1.0	-0.05	-0.0025	-0.01
Max value	2π	0.2	200	0.05	0.0025	0.01

used to handle outliers. The fitting method is based on the error from model to data, and thus is more robust than many other fitting methods that are based on the error from data to model.

For each curve model $n \in \{1, 2, \cdots, 6\}$, the model fitting method aims to find the candidate curve:

$$C_n^* = \arg\max_{C \in \mathcal{C}_n} g(C, D), \tag{5}$$

where $\mathcal{C}_n$ is the set of curves defined by the model n, D is the data point set, and $g(\cdot, \cdot)$ is the fitness function called mean measure:

$$g(C, D) = h(C)/d^2(C, D), \tag{6}$$

where $h(C)$ is the arc-length of the curve C, $d(C, D)$ is the modified Hausdorff distance from C to the data D [42]:

$$d(C, D) = \frac{1}{|C^\delta|} \sum\nolimits_{\mathbf{p} \in C^\delta} \min_{\mathbf{q} \in D} d(\mathbf{p}, \mathbf{q}) \tag{7}$$

where C^δ is a point set uniformly sampled from C with a small resolution. Then the best curve $C^{**} = C_b^*$, where

$$b = \arg\max_{1 \le n \le 6} g\left(C_n^*, D\right) \tag{8}$$

We use the cuckoo search algorithm [43] as described in [20] to solve the maximization problem (see Eq. (5)). The algorithm requires each variable of a model to be within a range. The location variables (x_0, y_0, z_0) are within the bounding box of the data. The ranges of the other variables are presented in Table II. These ranges are set with regard to highway geometric design standard. Similar to [20], we set the resolution for sampling points from curves to 0.3 times data resolution. In this paper, the data resolution is 20 cm, as the data (i.e., the hull points) are down-sampled to a resolution of 20 cm to remove redundancy. Therefore, the resolution for sampling points from curves is set to 6 cm. Then, the convergence tolerance for model fitting is the only parameter remained in this step.

IV. EXPERIMENTS

We implemented our method in C++ (for road marking extraction), Python (for curve reconstruction), and Java (for generating virtual scanning data). We conducted experiments on a machine running Windows 10 with an Intel Core i7-8700 3.20 GHz CPU and 16 GB RAM. The experiments include comparative study for road marking detection (Section IV-A), curve reconstruction on the virtual scan dataset (Section IV-B), and curve reconstruction on the real MLS dataset (Section IV-C).

A. Comparison of Intensity Variance and Other Features

We compared the intensity variance (Section III-C) against three other features (i.e., the intensity (Section III-C), the intensity median [44], and the intensity gradient [18]) by detecting road markings from the data shown in Fig. 4a. The data consists of 2,056,202 points. The way of using the intensity median or intensity gradient is similar to that of using the intensity variance as described in Section III-C. We compute the magnitude of the intensity median or intensity gradient for each point in the data. Then a magnitude threshold is calculated to detect road marking edges.

The results of road marking extraction by using the different features are shown in Fig. 9, which demonstrates that intensity variance has two superiorities upon the other features. First, intensity variance is more effective for extracting edges of road markings than intensity and intensity gradient. As shown in Figs. 9b, 9e, and 9f, the road markings extracted by using the intensity or intensity gradient contain many outliers. In contrast, as shown in Figs. 4c and 9a, the edges extracted by using the intensity variance are cleaner. Second, the performance of intensity variance is more stable under different distance threshold values than intensity median. When distance threshold is set to 10 cm, intensity median performs well (Fig. 9c). However, when distance threshold is set to 20 cm, many road marking points have not been extracted by intensity median (Fig. 9d). In contrast, intensity variance performs well both for threshold values 10 cm (Fig. 4c) and 20 cm (Fig. 9a).

It should be noted that, although intensity variance is suitable for extracting edge points of road markings, it may not be suitable for extracting whole points of road markings. Fortunately, road curves can be reconstructed with edge points.

B. Experiments on Virtual Scan Dataset

We use the virtual scan data VD-A (green part in Fig. 10a) to validate our method, since the ground-truth models of the virtual dataset VD-A are exactly known. We design a synthetic 3D highway surface (gray part in Fig. 10a), and then use the HELIOS virtual scanning system [45] to virtually scan the surface. Specifically, the virtual data VD-A is acquired by the virtual RIEGL VQ-450 scanners [45] with 7.2 km/h moving speed and 10 Hz scanning frequency.

The ground-truth horizontal and vertical curves representing the surface are circle and line, respectively. There are two

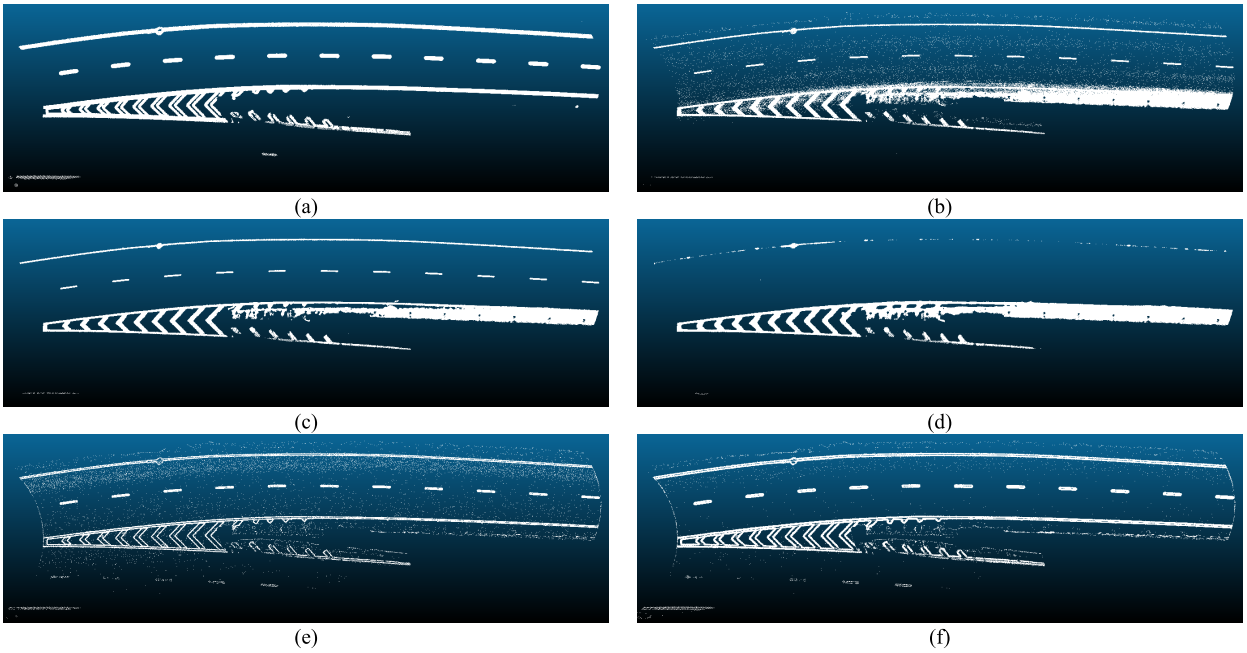


Fig. 9. Comparison of intensity variance and other features. (a) and Fig. 4c show the edge points extracted by using the intensity variance with distance thresholds 20 cm and 10 cm, respectively. (b) shows the marking points extracted by using the intensity. Note that it does not need a distance threshold to use the intensity. (c) and (d) show the marking points extracted by using the intensity median with distance thresholds 10 cm and 20 cm, respectively. (e) and (f) show the edge points extracted by using the intensity gradient with distance thresholds 10 cm and 20 cm, respectively.

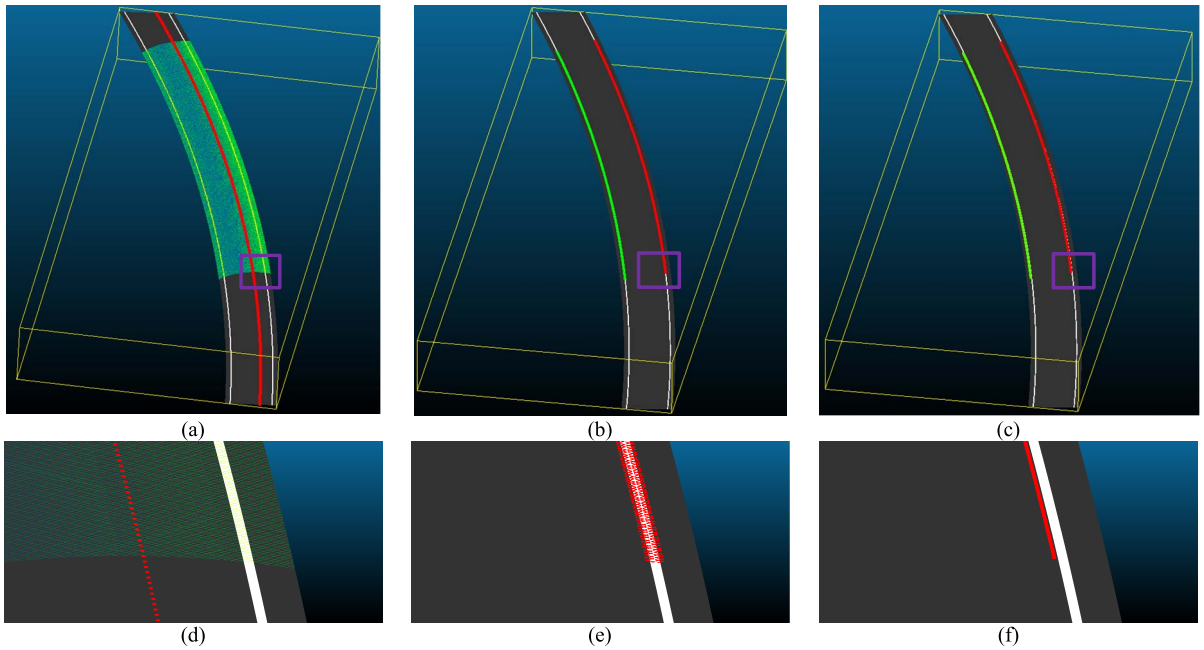


Fig. 10. Reconstructed curves from virtual scanning data VD-A. (a) The synthetic highway surface (gray), the trajectory (red) of the virtual scanner, and the virtually scanned data VD-A (green). (b) The extracted road markings (red and green) from the virtual scanned data VD-A. (c) The reconstructed curves (red and green) from the extracted road markings. The bottom row zooms in the parts within the purple boxes of the top row.

(inner and outer) white road markings on the surface. The width between the outer edge of the surface and the outer marking is 80 cm. The width of the road markings is 20 cm. The width between the two markings is 7 m. The horizontal radius of the outer edge of the surface is 350 m. The slope of the surface is 0.08.

As shown in Fig. 10c, the curves are well reconstructed by our method. Fig. 11 shows the evolution of fitness for fitting the six curve models to the detected inner and outer markings. As shown in Fig. 11a, the model best fitting the inner marking is the ground-truth model, i.e., the model combined by horizontal circle and vertical line. That is, our method correctly reconstructs the curve for the inner marking with finding the ground-truth model.

However, as shown in Fig. 11b, for the outer marking, the best fitting model is Circle-Parabola, which is not the ground-truth model (Circle-Line). The reason is as follows. As shown in Table I, when $\eta = 0$, vertical parabola is reduced

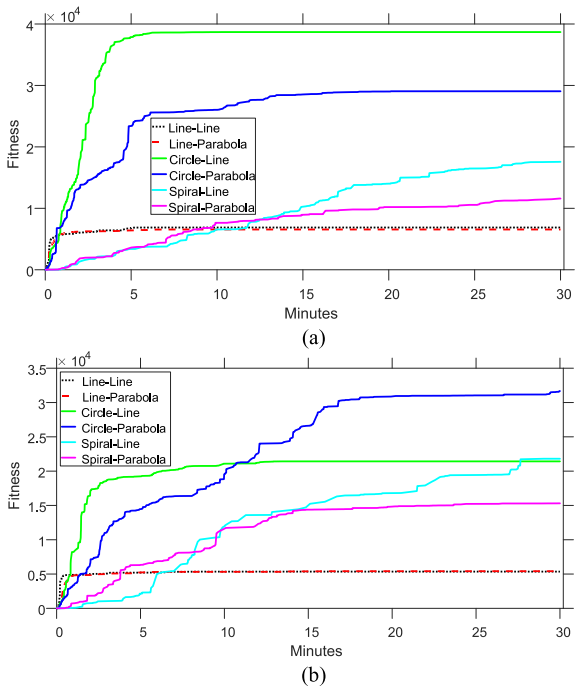


Fig. 11. Results of fitting the 6 curve models on the virtual scanned data VD-A in Fig. 10. The types of curve models are shown in the legend, where horizontal types are shown in ahead and vertical curves are shown in rear. For example, Spiral-Parabola denotes the 3D curve type combined by horizontal spiral and vertical parabola. (a) Fitting on the inner marking (green part of Fig. 10b). (b) Fitting on the outer marking (red part of Fig. 10b). The types of the ground-truth models of these data are Circle-Line.

TABLE III

GROUND-TRUTH AND RECONSTRUCTED VALUES OF SOME IMPORTANT VARIABLES OF THE INNER AND OUTER CURVES FOR DATA VD-A

	Variable	ξ_0	$1/\kappa_0$	ψ	η
Inner curve	Ground-truth	0.08	341.8	0	0
	Reconstructed	0.07999819	341.6351	0	0
Outer curve	Ground-truth	0.08	349	0	0
	Reconstructed	0.08006336	349.0036	0	-0.00000093

to vertical line. That is, line is a special case of parabola. Therefore, in practical applications, if η is approximately equal to zero, then the parabola can be seen as line. Table III shows the reconstructed value of η of Circle-Parabola (i.e., the best fitted curve) for the outer marking. It can be seen that, η is very close to zero. In other words, to a large extent, our method also finds the ground-truth model for the outer marking. Besides η, Table III also shows the reconstructed values for other important variables. All the reconstructed values are very close to the ground-truth values. Therefore, it can be concluded that, our method is able to accurately reconstruct highway curves in the virtual scan data VD-A.

We now compare the estimator used in our method against several other estimators for reconstructing curves on another virtual scan data VD-B. There are three differences between the data VD-B and the aforementioned data VD-A which is shown in Fig. 10a. First, the ground-truth model of the synthetic highway surface for VD-B is Circle-Parabola. The vertical curvature of the synthetic surface is 0.0005. Second, VD-B is acquired by the virtual RIEGL VQ-450 scanners [45] with 54 km/h moving speed and 100 Hz scanning frequency. Third, the trajectories and the scanned distances are contaminated by Gaussian noise with standard deviations 0.01 and 0.04, respectively.

TABLE IV

RECONSTRUCTED VALUES OF SOME IMPORTANT VARIABLES BY DIFFERENT ESTIMATORS FOR VD-B

Variable	$1/\kappa_0$	η
Estimator		
Ground-truth	349.2	0.0005
Mean squared error	308.16974426745486	0.000495742062
Inlier number (IT=0.05)	350.8747	0.00040696
Inlier number (IT=0.1)	335.0921	0.00062226
Inlier number (IT=0.2)	362.6496	0.00041677
M-estimator (IT=0.025)	64.7709	0.00200421
M-estimator (IT=0.05)	346.6015	0.00048958
M-estimator (IT=0.01)	347.3349	0.00050015
M-estimator (IT=0.2)	347.1531	0.00050011
M-estimator (IT=0.4)	308.1697341058159	0.000495742045
Mean measure (Our)	347.2981	0.00049981

The estimators used for comparison are mean squared error, inlier number [35], and M-estimator [36]. Inlier number and M-estimator are among the state-of-the-art robust estimators [37] and both require to tune an extra parameter called inlier threshold (IT). In the comparative experiments, we only use the ground-truth model (i.e., Circle-Parabola) to perform curve fitting on the extracted outer marking points. We use the same cuckoo search algorithm [43] to perform optimization for all the estimators. The convergence tolerance for the optimization is set to 2,000,000 iterations.

The fitting results of horizontal radius (i.e., $1/\kappa_0$) and vertical curvature (i.e., η) for VD-B are shown in Table IV. For variable η, the reconstructed value by our method (i.e., mean measure) is very close to the ground-truth value. However, for variable $1/\kappa_0$, the absolute reconstruction error (i.e., the difference between reconstructed value and ground-truth value) of our method is near 2 m. This error is about ten times larger than the error obtained on the virtual data VD-A (see Fig. 10 and Table III). This is due to that VD-B is acquired by higher moving speed and higher noise than VD-A. Let relative reconstruction error be the ratio of absolute reconstruction error to ground-truth value. The relative reconstruction error of our method for reconstructing variable $1/\kappa_0$ is only about 0.6% (i.e., 2/350). In this sense, our method is also able to accurately reconstruct highway curves for the virtual scan data VD-B.

In contrast, the reconstruction errors for horizontal radius or vertical curvature by using mean squared error and inlier number are bigger than by our method. Only when the inlier threshold (IT) is carefully chosen (in the case of Table IV, IT should be larger than 0.025 and smaller than 0.4), the performance of M-estimator is comparable with our method.

C. Experiments on Real MLS Data

The MLS data used for experiments were collected by the same RIEGL VMX-450 system as in [32]. In the experiments, the distance thresholds used for normal calculation (Section III-B), intensity variance calculation

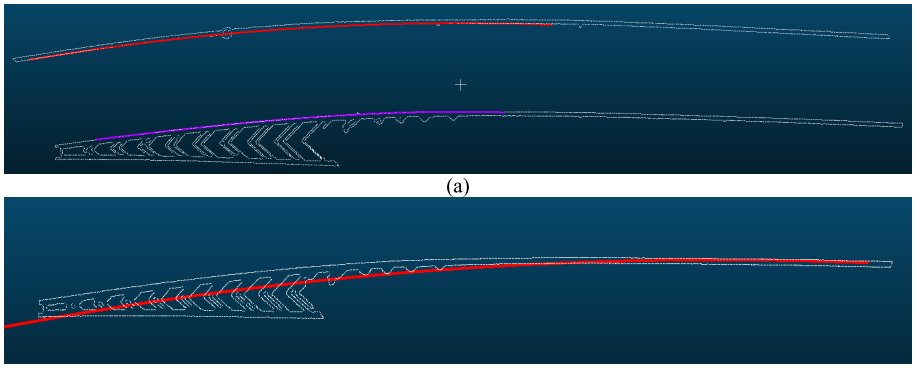


Fig. 12. Curves (color) reconstructed from the hull points (white) given in Fig. 7, using (a) our method, and (b) using the least-squares estimation method.

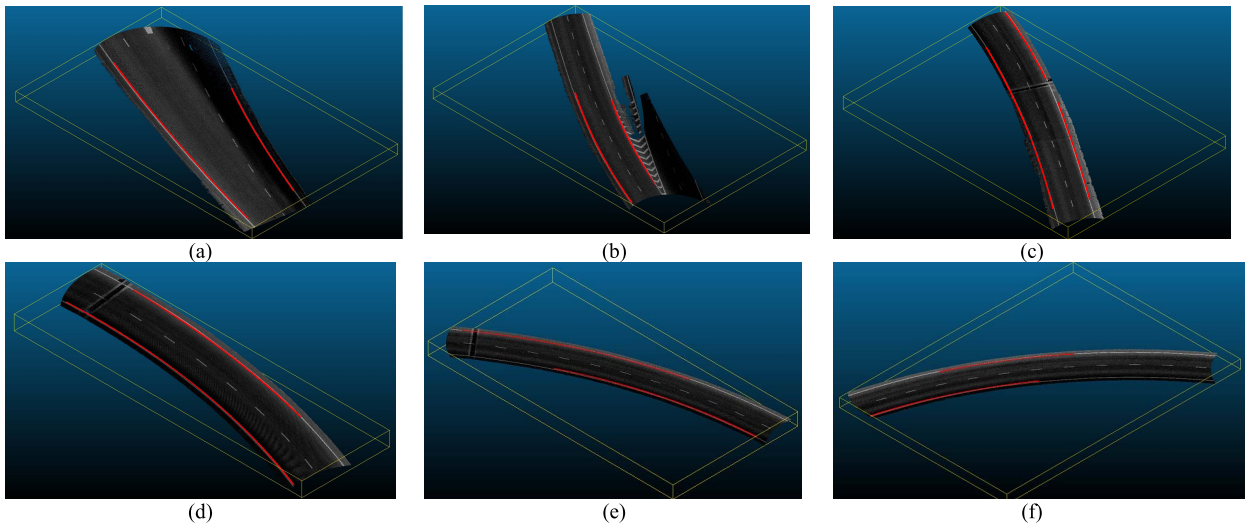


Fig. 13. Overlaps of the reconstructed curves (red) and their corresponding road surface points (gray). (a) to (f) sequential parts of MLS data of a highway. Note that, (b) corresponds to Fig. 7 and Fig. 12.

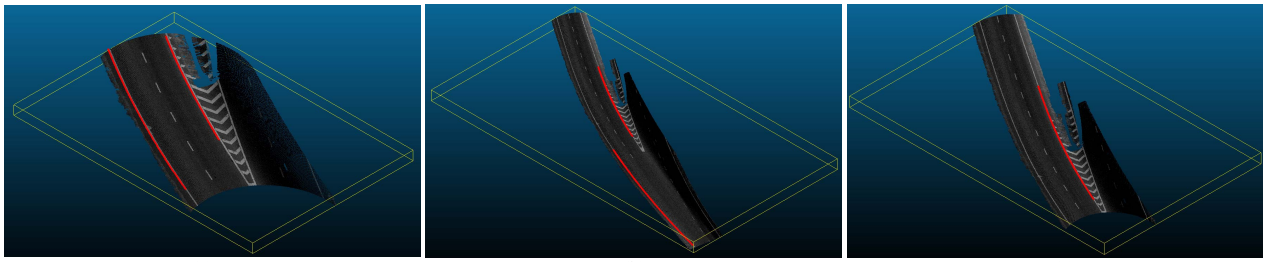


Fig. 14. Reconstructed curves (red) from a road surface partitioned by (a) $\beta = 500$, (b) $\beta = 2000$, and (c) Reconstructed curves (red) from a road surface (given in Fig. 13b) by a distance threshold 0.05.

(Section III-C), and clustering (Section III-C) are set to the same value. Therefore, the number of parameters used in our method was reduced to 3. Specifically, we set the distance threshold to 10 cm, β to 1000, and the convergence tolerance to 500,000 iterations.

Fig. 12a shows the curves reconstructed from the hull points in Fig. 7. It can be seen that, although the hull points contain a large number of outliers, our method is still able to correctly find the curve. In contrast, as shown in Fig. 12b, the curve reconstructed by a least-squares estimation method is significantly biased by the outliers. As shown in Fig. 13, the reconstructed curves can well represent the geometry of the highway.

We also conducted experiments to test the influence of several important parameters of our method. As shown in Figs. 14a and 14b, although our method can work for a large value of β, a small value of β is preferred. Our method reconstructs only one elemental curve from one road marking cluster. A large value of β results in a long road curve, which may contain multiple elemental curves. That means some

elemental curves may miss to be reconstructed. Fig. 14c shows the influence of the distance threshold. In Fig. 13b, where the distance threshold is set to 0.1, two curves (left and right) are reconstructed.

However, in Fig. 14c, where the distance threshold is set to 0.05, only one curve is reconstructed. The reason is that, our method fails to cluster the left curve into a big cluster and decides to reconstruct only the right curve.

V. Conclusion

In this paper, we have presented a new automated method to reconstruct 3D highway curves from noisy, unstructured, dense MLS point clouds. Our method consists of two major steps: detection of road markings solely using intensity variance, and curve reconstruction (i.e., recognizing the types of curves and estimating the values of curve parameters) through robust probabilistic program induction. Our method has been validated by experiments on a virtual scan dataset. The virtual experiments showed that, our method achieves the relative accuracy of less than 0.6% for estimation of horizontal circle radii and the vertical slope of 0.0001, respectively. The results obtained using a real MLS dataset also demonstrated effectiveness of our method. The final parametric curves obtained using our method is confined by the highway geometric design standard.

References

[1] G. M. Gibreel, S. M. Easa, Y. Hassan, and I. A. El-Dimeery, "State of the art of highway geometric design consistency," *J. Transp. Eng.*, vol. 125, no. 4, pp. 305–313, 1999.

[2] P. Di Mascio, M. Di Vito, G. Loprencipe, and A. Ragnoli, "Procedure to determine the geometry of road alignment using GPS data," *Procedia Social Behav. Sci.*, vol. 53, pp. 1202–1215, Oct. 2012.

[3] I. Laptev, H. Mayer, T. Lindeberg, W. Eckstein, C. Steger, and A. Baumgartner, "Automatic extraction of roads from aerial images based on scale space and snakes," *Mach. Vis. Appl.*, vol. 12, no. 1, pp. 23–31, 2000.

[4] S. M. Easa, H. Dong, and J. Li, "Use of satellite imagery for establishing road horizontal alignments," *J. Surv. Eng.*, vol. 133, no. 1, pp. 29–35, 2007.

[5] D. Chaudhuri, N. K. Kushwaha, and A. Samal, "Semi-automated road detection from high resolution satellite images by directional morphological enhancement and segmentation techniques," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 5, pp. 1538–1544, Oct. 2012.

[6] W. Shi, Z. Miao, and J. Debayle, "An integrated method for urban main-road centerline extraction from optical remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3359–3372, Jun. 2014.

[7] K. Lakakis, P. Savvaidis, and T. Wunderlich, "Evaluation of a low-cost mobile mapping and inspection system for road safety classification," *Amer. J. Geograph. Inf. Syst.*, vol. 2, no. 1, pp. 6–14, 2013.

[8] H. Cai and W. Rasdorf, "Modeling road centerlines and predicting lengths in 3-D using LIDAR point cloud and planimetric road centerline data," *Comput. Aided Civil Infrastruct. Eng.*, vol. 23, no. 3, pp. 157–173, 2008.

[9] 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.

[10] 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.

[11] 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.

[12] 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.

[13] 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.

[14] Y. Tsai, C. Ai, Z. Wang, and E. Pitts, "Mobile cross-slope measurement method using LiDAR technology," *J. Transp. Res. Board*, vol. 2367, no. 1, pp. 53–59, 2013.

[15] 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, Oct. 2014.

[16] 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 Infrastruct. Eng.*, vol. 30, no. 3, pp. 217–228, 2015.

[17] J. Wang, Z. Hu, Y. Chen, and Z. Zhang, "Automatic estimation of road slopes and superelevations using point clouds," *Photogramm. Eng. Remote Sens.*, vol. 83, no. 3, pp. 217–223, 2017.

[18] B. Yang, Y. Liu, Z. Dong, F. Liang, B. Li, and X. Peng, "3D local feature BKD to extract road information from mobile laser scanning point clouds," *ISPRS J. Photogramm. Remote Sens.*, vol. 130, pp. 329–343, Aug. 2017.

[19] B. M. Lake, R. Salakhutdinov, and J. B. Tenenbaum, "Human-level concept learning through probabilistic program induction," *Science*, vol. 350, no. 6266, pp. 1332–1338, 2015.

[20] Z. Zhang *et al.*, "Robust procedural model fitting with a new geometric similarity estimator," *Pattern Recognit.*, vol. 85, pp. 120–131, Jan. 2019.

[21] W. Wang, H. Pottmann, and Y. Liu, "Fitting B-spline curves to point clouds by curvature-based squared distance minimization," *ACM Trans. Graph.*, vol. 25, no. 2, pp. 214–238, 2006.

[22] Y. Hassan, "Three-dimensional approach for roadway alignment design incorporating driver perception," *Adv. Transp. Stud.*, vol. 2, pp. 15–32, Jul. 2004.

[23] W. Kühn, "The basics of a three-dimensional geometric design methodology," in *Proc. 3rd Int. Symp. Highway Geometric Design*, 2005, pp. 1–18.

[24] M. K. Jha, G. A. K. Karri, and W. Kuhn, "New three-dimensional highway design methodology for sight distance measurement," *J. Transp. Res. Board*, vol. 2262, no. 1, pp. 74–82, 2011.

[25] M. K. Jha, G. Karri, and W. Kuehn, "Three-dimensional highway design methodologies based on piecewise polynomials and integral Bézier splines," in *Proc. Transp. Res. Board 92nd Annu. Meeting*, 2013, pp. 53–62.

[26] 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.

[27] 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.

[28] M. Thuy and F. León, "Lane detection and tracking based on lidar data," *Metrol. Meas. Syst.*, vol. 17, no. 3, pp. 311–321, 2010.

[29] A. Mancini, E. Frontoni, and P. Zingaretti, "Automatic road object extraction from mobile mapping systems," in *Proc. IEEE/ASME Int. Conf. MESA*, Suzhou, China, Jul. 2012, pp. 281–286.

[30] M. Soilán, B. Riveiro, J. Martínez-Sánchez, and P. Arias, "Segmentation and classification of road markings using MLS data," *ISPRS J. Photogramm. Remote Sens.*, vol. 123, pp. 94–103, Jan. 2017.

[31] 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.

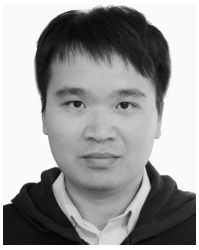
[32] H. Guan, J. Li, Y. Yu, C. Wang, M. Chapman, and B. Yang, "Using mobile laser scanning data for automated extraction of road markings," *ISPRS J. Photogramm. Remote Sens.*, vol. 87, pp. 93–107, Jan. 2014.

[33] P. Kumar, C. P. McElhinney, P. Lewis, and T. McCarthy, "An automated algorithm for extracting road edges from terrestrial mobile LiDAR data," *ISPRS J. Photog. Remote Sens.*, vol. 85, pp. 44–55, Nov. 2013.

[34] A. Iglesias, A. Gálvez, and A. Avila, "Immunological approach for full NURBS reconstruction of outline curves from noisy data points in medical imaging," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 15, no. 6, pp. 1–14, Nov./Dec. 2018. doi: 10.1109/TCBB.2017.2688444.

[35] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with apphcatlons to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[36] P. H. S. Torr and A. Zisserman, "MLESAC: A new robust estimator with application to estimating image geometry," *Comput. Vis. Image Understand.*, vol. 78, no. 1, pp. 138–156, 2000.
[37] H. Le, T.-J. Chin, and D. Suter, "An exact penalty method for locally convergent maximum consensus," in *Proc. IEEE Conf. CVPR*, Honolulu, HI, USA, Jul. 2017, pp. 379–387. doi: 10.1109/CVPR.2017.48.
[38] G. Vosselman, B. G. H. Gorte, G. Sithole, and T. Rabbani, "Recognising structure in laser scanner point clouds," *ISPRS Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. 46, no. 8, pp. 33–38, 2004.
[39] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-9, no. 1, pp. 62–66, Jan. 1979.
[40] H. Edelsbrunner, D. Kirkpatrick, and R. Seidel, "On the shape of a set of points in the plane," *IEEE Trans. Inf. Theory*, vol. IT-29, no. 4, pp. 551–559, Jul. 1983.
[41] E. Bertolazzi and M. Frego, "On the G^2 hermite interpolation problem with clothoids," *J. Comput. Appl. Math.*, vol. 341, pp. 99–116, Oct. 2018.
[42] M.-P. Dubuisson and A. K. Jain, "A modified Hausdorff distance for object matching," in *Proc. 12th ICPR*, Jerusalem, Israel, vol. 1, Oct. 1994, pp. 566–568.
[43] X. S. Yang and S. Deb, "Engineering optimisation by cuckoo search," *Int. J. Math. Model. Numer. Optim.*, vol. 1, no. 4, pp. 330–343, 2010.
[44] L. Yan, H. Liu, J. Tan, Z. Li, H. Xie, and C. Chen, "Scan line based road marking extraction from mobile lidar point clouds," *Sensors*, vol. 16, no. 6, p. 903, 2016.
[45] S. Bechtold and B. Höfle, "HELIOS: A multi-purpose LIDAR simulation framework for research, planning and training of laser scanning operations with airborne, ground-based mobile and stationary platforms," *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. III-3, pp. 161–168, Jun. 2016.



Zongliang Zhang received the B.Eng. degree from Fuzhou University in 2010, the M.Eng. degree from Northwest A&F University in 2013, and the Ph.D. degree from Xiamen University, China, in 2018, all in computer science and technology.

He is currently a Lecturer with Fujian Medical University, China. He has published an article in *Pattern Recognition* journal and coauthored several articles published in refereed conference proceedings, including IGARSS, AAAI, and ISPRS. His current research interests are 3D computer vision, 3D geometric modeling, point cloud processing, and robust model fitting.

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

He is currently a Professor and the Head of the Mobile Sensing and Geodata Science Group, Department of Geography and Environmental Management, and cross-appointed at the Department of Systems Design Engineering, University of Waterloo, Canada. He has coauthored more than 400 publications, over 170 of them were published in refereed journals, including IEEE TGRS, IEEE TITS, IEEE GRSL, IEEE JSTARS, ISPRS JPRS, IJRS, PE&RS, and RSE, and over 170 of them were refereed conference proceedings, including CVPR, AAAI, IJCAI, IGARSS, and ISPRS. His current research interests include machine/deep learning algorithms for converting image pixels and LiDAR point clouds into HD maps to support autonomous vehicles and smart cities. He is also the Chair of the ISPRS Working Group I/2 on LiDAR, Air- and Spaceborne Optical Sensing (2016–2020) and the ICA Commission on Sensor-driven Mapping (2015–2019) and an Associate Editor of the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, the IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING, and the *Canadian Journal of Remote Sensing*.

Yulan Guo received the B.Eng. and Ph.D. degrees from the National University of Defense Technology (NUDT) in 2008 and 2015, respectively. He was a Visiting Ph.D. Student with The University of Western Australia from 2011 to 2014 and a Post-Doctoral Research Fellow with the Institute of Computing Technology, Chinese Academy of Sciences from 2016 to 2018. He has authored over 80 articles in journals and conferences, such as the IEEE TPAMI and IJCV. His current research interests focus on 3D vision, particularly on 3D feature learning, 3D modeling, 3D object recognition, and 3D biometrics. He was a PC Member for several conferences, such as ACM MM, IJCAI, and AAAI, and a recipient of the CAAI Outstanding Doctoral Dissertation Award in 2016. He served as an Associate Editor for the *IET Computer Vision* and *IET Image Processing*, a Guest Editor for the IEEE TPAMI, a Reviewer for over 30 journals, and an Organizer for a tutorial in CVPR 2016 and a workshop in CVPR 2019.

Chenhui Yang received the Ph.D. degree in mechanical engineering from Zhejiang University, Zhejiang, China, in 1995.

He was a Visiting Scholar with the University of Chicago/Argonne National Laboratory from 1999 to 2000 and the University of Southern California from 2014 to 2015. He is currently a Professor with the School of Information Science and Engineering, Xiamen University, China. He has coauthored more than 80 research articles published in journals and conference proceedings. His research interests focus on computer vision, computer graphics, data mining, and their applications in transportation, security, and medical studies.

Cheng Wang (M'04–SM'16) received the Ph.D. degree in information and communication engineering from the National University of Defense Technology, Changsha, China, in 2002.

He is currently a Professor and the Associate Dean of the School of Information Science and Engineering, Xiamen University, China, where he is also the Executive Director of the Fujian Key Laboratory of Sensing and Computing for Smart Cities. He has coauthored over 80 articles in refereed journals, including IEEE-TGRS, IEEE-TITS, IEEE-GRSL, IEEE-JSTARS, IJRS, and ISPRS-JPRS. His research interests include remote sensing image processing, mobile LiDAR data analysis, and multisensory data fusion. He is also a fellow of IET and a Council Member of the China Society of Image and Graphics. He is also the Chair of the ISPRS WG I/3 on Multi-Platform Multi-Sensor System Calibration (2016–2020).