Automated Detection of Urban Road Manhole Covers Using Mobile Laser Scanning Data

Article *in* IEEE Transactions on Intelligent Transportation Systems · April 2015

DOI: 10.1109/TITS.2015.2413812

CITATIONS
10

READS
98

3 authors, including:

Yongtao Yu
Huaiyin Institute of Technology
41 PUBLICATIONS **369** CITATIONS
SEE PROFILE

Haiyan Guan
Nanjing University of Information Science & …
52 PUBLICATIONS **427** CITATIONS
SEE PROFILE

Automated Detection of Urban Road Manhole Covers Using Mobile Laser Scanning Data

Yongtao Yu, Haiyan Guan, *Member, IEEE*, and Zheng Ji

***Abstract*—This paper proposes a novel framework for automated detection of urban road manhole covers using mobile laser scanning (MLS) data. First, to narrow searching regions and reduce the computational complexity, road surface points are segmented from a raw point cloud via a curb-based road surface segmentation approach and rasterized into a georeferenced intensity image through inverse distance weighted interpolation. Then, a supervised deep learning model is developed to construct a multilayer feature generation model for depicting high-order features of local image patches. Next, a random forest model is trained to learn mappings from high-order patch features to the probabilities of the existence of urban road manhole covers centered at specific locations. Finally, urban road manhole covers are detected from georeferenced intensity images based on the multilayer feature generation model and random forest model. Quantitative evaluations show that the proposed algorithm achieves an average completeness, correctness, quality, and F_1-measure of 0.955, 0.959, 0.917, and 0.957, respectively, in detecting urban road manhole covers from georeferenced intensity images. Comparative studies demonstrate the advantageous performance of the proposed algorithm over other existing methods for rapid and automated detection of urban road manhole covers using MLS data.**

***Index Terms*—Deep learning, manhole cover, mobile laser scanning (MLS), random forest, road distress, road safety.**

I. Introduction

FOR many intelligent transportation systems (ITS), the accurate information about current road conditions, road surface marks, and road surface fixtures is greatly important to conduct correct behaviors and improve safety. Accurate and real-time detection and recognition of these characters form significant inputs to many intelligent transportation related applications, such as driver assistance and safety warning systems [1], [2] and autonomous driving [3], [4]. However, the absence of accurate feedbacks from the road to the ITS might lead to terrible traffic accidents in urban areas. Therefore, rapidly and effectively detecting potential road surface distresses can reduce the occurrence of severe casualties and improve traffic safety. Among the road surface fixtures, road manholes are a common character on urban roads. They usually function to conduct rainwater, drainage, power cables, telecommunication cables, and other things. Generally, they are covered with a metal-made or concrete-made cover to keep vehicles, pedestrians, and other things from dropping into the wells. However, if the manhole cover is removed or stolen by someone, or broken caused by some uncertain factors, it is very dangerous to the moving vehicles and pedestrians, especially at night. Thus, rapid and automated detection of road manhole covers not only assists the transportation infrastructure agencies to conduct monitoring and repairs for driving safety, but also provides useful information to the ITS for the warning of potential road distresses and safety hazards.

Traditionally, the monitoring and repairing of road surface distresses were mainly accomplished by on-site inspections and maintenances. Such field measurements were time consuming, labor intensive, and costly for maintaining complicated urban road networks. Sometimes, it was even greatly dangerous to work on highways or in tunnels. With the development of optical imaging techniques, mobile mapping systems (MMS) integrated with digital camera(s) or video(s) [5], [6] were widely used in a variety of transportation related applications. However, due to the passive mapping ways, optical imaging based MMS relied greatly on the illumination conditions of the environment. Thus, the mapping mission was only limited to the daytime. In addition, the MMS suffered greatly from distortions, blurs, occlusions caused by nearby moving objects, shadows cast by roadside high buildings and trees, and lack of accurate geospatial information of the measured targets.

In the past two decades, laser scanning and navigation technologies have rapidly developed and been used in a variety of applications, such as military, transportation, mining, forestry, map drawing, heritage documentation, and basic surveying and mapping. Among the laser scanning products, mobile laser scanning (MLS) systems [7] have shown outstanding advantages in transportation related activities. Compared to its airborne counterparts, an MLS system has the following advantages: 1) more direct views of road surfaces and roadside vertical objects, 2) higher point density and higher accuracy of 3-D point clouds, and 3) less costly mapping missions. Moreover, compared to optical imaging based MMS, MLS

Manuscript received September 15, 2014; revised January 9, 2015; accepted March 12, 2015. Date of current version November 23, 2015. This work was supported by the National Natural Science Foundation of China under Grant 41471379. The Associate Editor for this paper was H. Wang. *(Corresponding author: Yongtao Yu)*.

Y. Yu is with the Fujian Key Laboratory of Sensing and Computing for Smart Cities, Xiamen University, Xiamen 361005, China (e-mail: allennessy.yu@gmail.com).

H. Guan is with the College of Geography and Remote Sensing, Nanjing University of Information Science and Technology, Nanjing 210044, China (e-mail: guanhy.nj@nuist.edu.cn; guanhy.nj@gmail.com).

Z. Ji is with the School of Remote Sensing Information and Engineering, Wuhan University, Wuhan 430079, China (e-mail: jz07@whu.edu.cn).

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/TITS.2015.2413812

systems can collect high-density and high-accuracy 3-D point clouds of real-world coordinates over a large area within a short time period. Due to the active mapping ways of MLS systems, the mapping mission can be carried out day and night. Therefore, MLS systems are a promising and cost-effective alternative for rapidly monitoring road surface distresses.

In this paper, we propose an automated algorithm for detecting urban road manhole covers using MLS data. First, to narrow searching regions and reduce the computational complexity, road surface points are segmented from a raw point cloud using a curb-based road surface segmentation approach and rasterized into a georeferenced intensity image through inverse distance weighted (IDW) interpolation. Then, a supervised deep learning model is developed to construct a multi-layer feature generation model for depicting high-order features of local image patches. Next, a random forest model is trained to learn mappings from high-order patch features to the probabilities of the existence of road manhole covers centered at specific locations. Finally, road manhole covers are detected based on the combination of these models. The contributions of this paper are as follows: 1) a multi-layer feature generation model is proposed to effectively generate high-order patch features; and 2) a random forest model is developed to cast weighted votes for detecting road manhole covers.

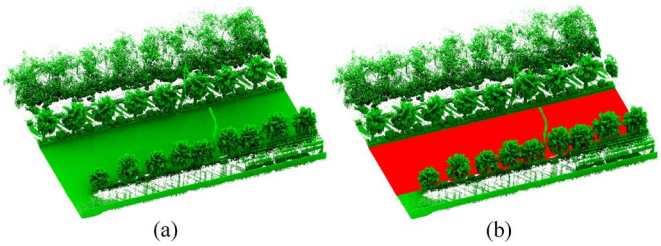
The rest of this paper is organized as follows: Section II reviews some related work on manhole cover detection. Section III describes the proposed road manhole cover detection framework. Section IV reports and discusses the experimental results. Finally, Section V gives the concluding remarks.

II. Related Work

Most of existing methods for manhole cover detection are based on digital imagery. A morphological method was developed in [8] for detecting round-shaped manhole covers from road surface images. First, a black top-hat operation with disk-shaped structure elements was performed to extract round-shaped components. Then, a masking operation was applied to the round-shaped components with a thresholded input image. Finally, manhole covers were obtained by eliminating the black regions with small areas or the regions without any holes. In [9], inspired by separability filters, a novel method was proposed to detect obscure and textured manhole covers from noisy and inhomogeneous contrast road surface images. Rather than analyzing the intensity difference between the manhole covers and their surroundings, the separability and uniformity of the image intensity distributions were studied using the Bhattacharyya coefficient. Finally, three indicators (circular object indicator, oriented separability indicator, and uniformity indicator) were defined and used to achieve the detection of manhole covers. A model-based method was introduced in [10] for detecting and localizing circular-shaped manhole covers in aerial images. This method was based on a parametric intensity model, which defined a circular-shaped manhole cover model with five parameters for respectively depicting the radii and intensity distributions. Localization of manhole covers was achieved by directly fitting this model to the observed image intensities.

Caused by frequent occlusions, regular changes in illumination conditions, substantial viewpoint variances, and varying textured appearances and designed shapes, accurate detection, recognition, and 3-D localization of manhole covers from images captured by moving vans is greatly challenging. In [11], a multi-view method integrated with 2-D and 3-D techniques was presented to detect road manhole covers from road surface images. The detection process was separated into single-view and multi-view processing stages. At single-view processing stage, the position and inclination of the ground plane were first estimated in order to projectively warp the images to generate fronto-parallel views. Then, single-view processing was applied to such fronto-parallel views using either a single part-based detector or a cascaded framework composed of segmentation, area, aspect ratio, variance, symmetry, and texture-based filters. At multi-view processing stage, the detections from the single views were fused and grouped into more reliable 3-D hypotheses, which were fed into a graph-cut segmentation filter. Finally, these hypotheses were used for accurate 3-D localization of manhole covers. Considering the complex background of road surface images, an improved Hough transform was proposed in [12] for the detection of manhole covers. First, on binary edge images, all the contours were ascertained through contour tracking. Then, false alarms were eliminated using contour filters. Finally, the improved Hough transform was applied to detect circular-shaped manhole covers.

Due to the superior properties of MLS systems over the optical imaging based systems, MLS data have also been investigated by some researchers to detect road manhole covers. In [13], marked point processes were proposed to detect circular-shaped and rectangular-shaped manhole covers from MLS point clouds. First, road surface points were segmented using a curb-based approach and then rasterized into 2-D georeferenced intensity images through IDW interpolation. Next, two types of marked points were defined to model the geometric features of manhole covers. Finally, circular-shaped and rectangular-shaped road manhole covers were detected and optimized based on the reversible jump Markov chain Monte Carlo (RJMCMC) algorithm. A multi-scale tensor voting approach was developed in [14] for detecting road manhole covers from MLS point clouds. Similarly, road surface points were first segmented and rasterized into 2-D georeferenced intensity images. Then, manhole cover candidates were ascertained through distance-dependent intensity thresholding. Next, a multi-scale tensor voting framework was performed to suppress noise and preserve manhole cover pixels. Finally, manhole cover regions were extracted via distance-based clustering and morphological operations. A new method combined with multi-view matching and feature extraction techniques was developed in [15] to detect road manhole covers using close-range images and MLS data. To achieve accurate road manhole cover detection, a novel edge detection and feature extraction method was introduced to overcome the difficulties of viewpoint variances, varying ground materials, and complex road scenes with shadows and vehicles. The MLS data were used to segment road scenes so that roadside features and off-road objects, such as vehicles and pedestrians, were excluded from the road. Then, the arcs detected by Canny detectors were extracted and fitted to form



(a) (b)

Fig. 1. (a) Raw point cloud, and (b) segmented road surface points (red).

ellipses, which were finally resampled and matched to adjacent images for checking the existence of road manhole covers.

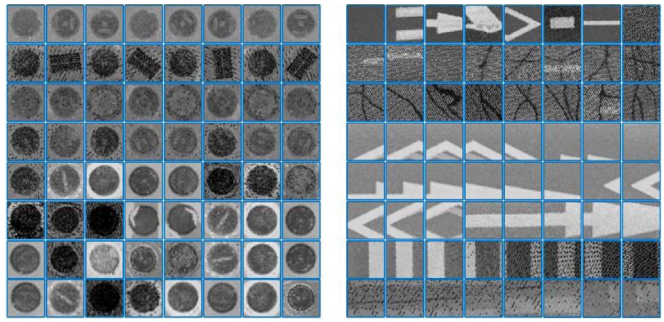
III. Method

A. Data Preprocessing

In a data acquisition mission, the 3-D point cloud data of both road surfaces and roadside features can be simultaneously collected by an MLS system. However, the detection of road manhole covers only focuses on road surfaces instead of the entire scene. Therefore, in order to reduce the quantity of the data to be processed and narrow the searching regions, road surface points should be segmented from the raw point cloud accurately and rapidly. In our previous study [16], we developed a curb-based road surface segmentation approach. This approach investigates the geometric properties of the curbs that are designed to separate the road from the sidewalks on urban roads. First, with the assistance of the trajectory data that is acquired by the onboard navigation system, a raw point cloud is vertically partitioned into a set of data blocks with a certain width (e.g., 3 m) along the direction of the trajectory. Rather than handling the entire data block, a vertical profile is generated perpendicularly to the trajectory within each data block. Then, through profile analysis, curb points are located within each profile by selecting the points on opposite sides of the trajectory with specific elevation gradients constrained by an elevation filter. Finally, the curb points ascertained from all profiles are fitted into curb-lines. Based on the knowledge that curbs indicate the boundaries of the road, road surface points are easily segmented from the raw point cloud by using the fitted curb-lines. This curb-based road surface segmentation approach operates rapidly and effectively and achieves high accuracy in road surface segmentation. Therefore, in this study, we use this curb-based road surface segmentation approach [16] to prepare road surface points for the detection of road manhole covers. Fig. 1 illustrates a visual example of the segmented road surface points using the curb-based approach.

Instead of handling the 3-D discrete road surface points, we rasterize them into a 2-D georeferenced intensity image based on the intensity information of the laser points. In our previous study [13], [16], we propose an inverse distance weighted (IDW) interpolation method to generate georeferenced intensity images from 3-D MLS point clouds. According to the IDW interpolation method, the road surface points are first vertically divided into a grid structure with a certain spacing (e.g., 2.5 cm) in the XY plane. These grids correspond to the pixels in the resultant georeferenced intensity image. Next, the laser points within each grid are weightedly interpolated to generate a single pixel based on the following rules [16]: 1) a point with higher intensity obtains a greater weight; and 2) a point with shorter distance away from the grid's geometric center obtains a greater weight. In this way, the generated georeferenced intensity image well reflects the retro-reflectivity properties of the road surface. Therefore, in this study, we use the IDW interpolation method [16] to convert the 3-D discrete road surface points into 2-D orthographic road surface images. These orthoimages are used for road manhole cover detection. Fig. 2 shows a visual example of the generated 2-D georeferenced intensity image of the segmented road surface points using the IDW interpolation method.

Fig. 2. Illustration of the generated 2-D georeferenced intensity image.



(a) (b)

Fig. 3. Illustration of (a) a subset of positive training samples, and (b) a subset of negative training samples.

B. Patch Feature Learning

Recently, deep learning models [17]–[19] have attracted great attention for their superior capabilities in learning hierarchical deep features from large amounts of unlabeled data. Among these deep learning models, deep Boltzmann machines (DBM) [19] have been proven to be a promising and powerful tool. A DBM model is actually a layer-wise structure composed of a stack of restricted Boltzmann machines (RBM) [20] with feedbacks from upper layers. In this paper, we propose a supervised learning strategy to jointly train a DBM model from a set of manually labeled training samples. The trained joint DBM model is used to construct a multi-layer feature generation model for describing the high-order feature representation of a local image patch.

First, a group of training samples with a size of $n \times n$ pixels are manually selected from the generated georeferenced intensity images. These training samples are manually labeled into two categories: positive training samples consisting of manhole cover patches [see Fig. 3(a)] and negative training

samples consisting of the background road surface patches [see Fig. 3(b)]. Then, we jointly train a two-layer DBM model [see Fig. 4(a)] using the labeled training samples. Let $\boldsymbol{v} \in [0,1]^{n^2}$ be a vector of real-valued visible units that represent a linear arrangement of an image patch (a positive or negative training sample). Let $\mathbf{L}$ be a binary label vector with a "1-of-K" encoding pattern [19]. That is, $\mathbf{L} = [1,0]^{\mathrm{T}}$ and $\mathbf{L} = [0,1]^{\mathrm{T}}$ respectively encode a positive training sample and a negative training sample. Let $\mathbf{h}^1 \in \{0,1\}^{D_1}$ and $\mathbf{h}^2 \in \{0,1\}^{D_2}$ represent the lower and higher layer binary hidden variables, respectively. Here, D_1 and D_2 are the number of hidden units in the lower and higher hidden layers, respectively. Then, for this joint DBM model, the energy of the joint configuration $\{\boldsymbol{v}, \mathbf{L}, \mathbf{h}^1, \mathbf{h}^2\}$ is defined as [19]:

$$E(\boldsymbol{v}, \mathbf{L}, \mathbf{h}^1, \mathbf{h}^2; \theta) = \frac{1}{2}\sum_{i=1}^{n^2}\frac{v_i^2}{\sigma_i^2} - \sum_{i=1}^{n^2}\sum_{j=1}^{D_1}\frac{v_i}{\sigma_i}w_{ij}^1 h_j^1 - \sum_{j=1}^{D_1}\sum_{m=1}^{D_2} h_j^1 w_{jm}^2 h_m^2 - \sum_{k=1}^{2}\sum_{m=1}^{D_2} l_k w_{km}^{\mathrm{L}} h_m^2 \tag{1}$$

where $\theta = \{\mathbf{W}^1, \mathbf{W}^2, \mathbf{W}^{\mathrm{L}}, \boldsymbol{\sigma}\}$ are the model parameters. $\mathbf{W}^1$, $\mathbf{W}^2$, and $\mathbf{W}^{\mathrm{L}}$ represent the visible-to-hidden, hidden-to-hidden, and label-to-hidden symmetric interaction terms, respectively; $\boldsymbol{\sigma}$ represents the standard deviations of the visible units. The marginal distribution over the visible vector $\boldsymbol{v}$ with a label $\mathbf{L}$ takes the following form:

$$P(\boldsymbol{v}, \mathbf{L}; \theta) = \frac{\sum_{\mathbf{h}^1, \mathbf{h}^2} \exp\left[-E(\boldsymbol{v}, \mathbf{L}, \mathbf{h}^1, \mathbf{h}^2; \theta)\right]}{\int_{\boldsymbol{v}'} \sum_{\mathbf{h}^1, \mathbf{h}^2, L} \exp\left[-E(\boldsymbol{v}', \mathbf{L}, \mathbf{h}^1, \mathbf{h}^2; \theta)\right] d\boldsymbol{v}'}. \tag{2}$$

The conditional distributions over the visible, label, and two sets of hidden units are defined as:

$$p\left(h_j^1 = 1 | \boldsymbol{v}, \mathbf{h}^2\right) = g\left(\sum_{i=1}^{n^2} w_{ij}^1 \frac{v_i}{\sigma_i} + \sum_{m=1}^{D_2} w_{jm}^2 h_m^2\right) \tag{3}$$

$$p\left(h_m^2 = 1 | \mathbf{h}^1, \mathbf{L}\right) = g\left(\sum_{j=1}^{D_1} w_{jm}^2 h_j^1 + \sum_{k=1}^{2} w_{km}^{\mathrm{L}} l_k\right) \tag{4}$$

$$p(v_i = x | \mathbf{h}^1) = \frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{\left(x - \sigma_i \sum_{j=1}^{D_1} w_{ij}^1 h_j^1\right)}{2\sigma_i^2}\right) \tag{5}$$

$$p(l_k | \mathbf{h}^2) = \frac{\exp\left(\sum_{m=1}^{D_2} w_{km}^{\mathrm{L}} h_m^2\right)}{\sum_{s=1}^{2} \exp\left(\sum_{m=1}^{D_2} w_{sm}^{\mathrm{L}} h_m^2\right)} \tag{6}$$

where $g(x) = 1/(1+\exp(-x))$ is the logistic function [19].

Exact maximum likelihood learning in this joint DBM model is intractable. In order to rapidly and effectively train this model, a greedy layer-wise pre-training [21] is first applied to initialize the model parameters θ. Then, a joint training algorithm that benefits from variational and stochastic approximation approaches [19] is adopted to jointly fine-train the model parameters.

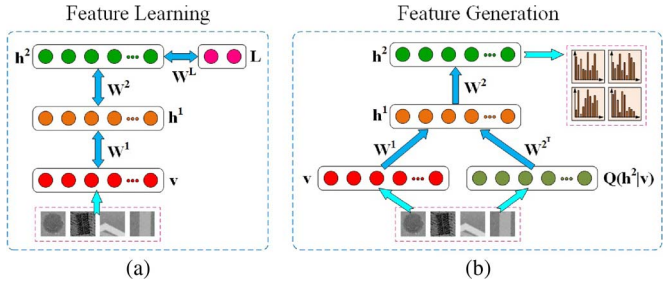


Fig. 4. (a) Jointly trained DBM model, and (b) feature generation model.

After the joint DBM model is trained, the stochastic activities of binary features in each hidden layer are replaced by deterministic real-valued probabilities to construct a multi-layer feature generation model [see Fig. 4(b)]. Considering the feedbacks from the hidden layers, for each input visible vector $\boldsymbol{v}$, the mean-field inference [22] is applied to obtain an approximate posterior distribution $Q(\mathbf{h}^2|\boldsymbol{v})$. Then, the marginal $q(\mathbf{h}^2|\boldsymbol{v})$ of this approximate posterior is input as an augment to this multi-layer feature generation model. Finally, the output of this multi-layer feature generation model produces a high-order feature representation (I) for vector $\boldsymbol{v}$ associated with a 2-D image patch:

$$I^{\mathrm{T}} = g\left(g\left(\frac{\boldsymbol{v}^T}{\sigma^{\mathrm{T}}}\mathbf{W}^1 + q(\mathbf{h}^2|\boldsymbol{v})^T (\mathbf{W}^2)^{\mathrm{T}}\right)\mathbf{W}^2\right) \in [0,1]^{D_2}. \tag{7}$$

C. Random Forest Training

Random forests [23], [24] are a promising and powerful model for learning mappings from local image patch features to the probabilities of the existence of an object centered at specific locations. They have been successfully used for spatial context modeling [23], object detection [24], classification [25], segmentation [26], etc. In this paper, we train a random forest model to cast weighted votes about the certainty of the existence of road manhole covers based on the high-order features of local image patches.

First, the training samples (positive and negative training samples) (see Fig. 5) are normalized and linearized to form real-valued visible inputs to the multi-layer feature generation model. Then, the output of the multi-layer feature generation model produces the high-order features of these training samples. Finally, the obtained high-order features along with the class labels of the training samples form the training data for constructing a random forest (see Fig. 5).

For a random forest, each tree in the forest is constructed separately based on a set of local patches $\{p_i = (I_i, c_i)\}$ (see Fig. 5), where I_i is the high-order feature representation of the patch; c_i is the class label of the patch (0 means a negative sample and 1 indicates a positive sample). As shown in Fig. 5, each internal node of a constructed tree in the random forest represents a binary test function, which bipartitions the patches

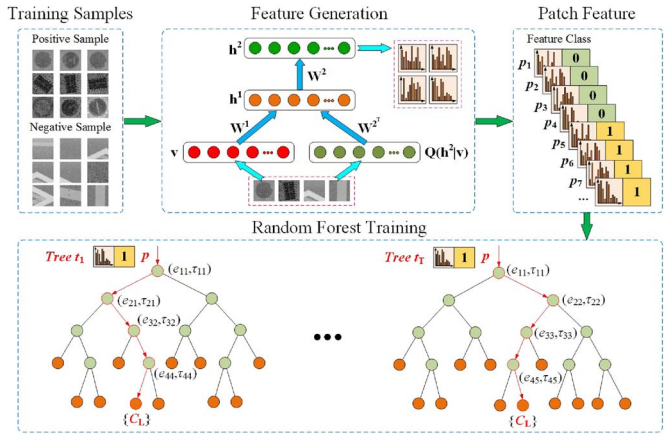


Fig. 5. Illustration of random forest training framework.

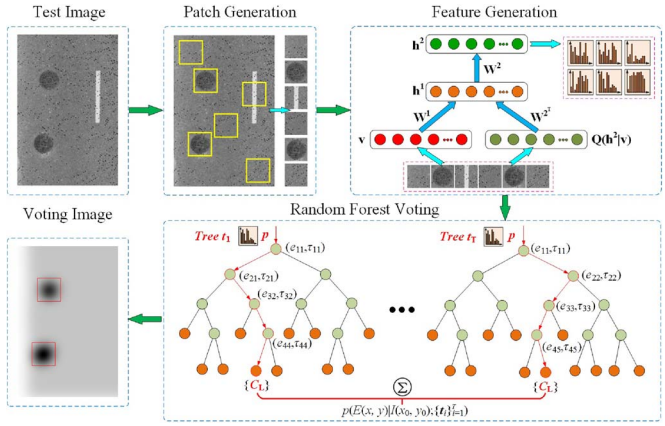


Fig. 6. Illustration of road manhole cover detection framework.

arriving at this node into two subsets according to their feature representations and respectively redistributes these two subsets to its left and right children nodes for further processing. Here, we define the binary test function as follows:

$$BT(I) = \begin{cases} 0, & \text{if } I(e) < \tau \\ 1, & \text{otherwise} \end{cases} \tag{8}$$

where $I(e)$ represents the e'th feature channel of feature I; $\tau \in (0, 1)$ is a real handicap value for determining the distribution direction of a patch. Thus, a patch with a test value of 0 is distributed to the left child node, whereas a patch with a test value of 1 is distributed to the right child node. For each leaf node in a tree, the statistical information about the patches reaching this node is analyzed and stored. Here, the label proportion information $C_{\mathrm{L}} \in [0, 1]$, which represents the proportion of positive patches, is computed and stored at the leaf node. In this way, the leaf nodes form a discriminative codebook with information about the probabilities of the distributions of the manhole covers' centers. At detection stage, this information is used to cast weighted votes about the existence of the manhole covers centered at specific locations in a voting space.

At training stage, each tree in the random forest is recursively constructed starting at the root. During construction, each newly constructed tree node receives a set of training patches. If the depth of the node reaches the pre-defined maximal depth $d_{\max}$ or the number of patches lies below a threshold $N_{\min}$, the constructed node is labeled as a leaf node. Then, the label proportion information C_{L} is computed and stored at this leaf node. Otherwise, an internal node is constructed and an optimal binary test function is determined to bipartition and redistribute the training patches. Finally, the split two subsets are respectively distributed to two newly created children nodes. In order to suppress the uncertainty of class labels, we adopt the same way, which uses class-label uncertainty measures, to design the binary test function as in [27].

D. Road Manhole Cover Detection

Fig. 6 illustrates a detailed workflow of the proposed algorithm for automated detection of road manhole covers from the georeferenced intensity images that are generated from road surface point clouds. At detection stage, we adopt a sliding window strategy to partition a test image into a group of local image patches with a size of $n \times n$ pixels. In order to generate a full and redundant coverage of the test image, two adjacent patches are designed to have an overlapping size of n_{o} pixels. Then, these patches are input into the multi-layer feature generation model to obtain high-order feature representations.

Next, each of the generated patches is distributed into the constructed random forest, where each tree receives a copy of this patch. When the patch arrives at an internal node, the binary test function stored at this node is used to conduct a correct route for this patch according to its high-order feature representation. Once the patch arrives at a leaf node, the label proportion information C_{L} stored at this leaf node is used to cast weighted votes.

Consider a patch $p(x, y) = (I(x, y), c(x, y))$, where $I(x, y)$ is the high-order feature representation of the patch centered at position (x, y) in a test image; $c(x, y)$ is the hidden class label of the patch. Let $E(x, y)$ be the random event corresponding to the existence of a road manhole cover centered at position (x, y) in a test image. By this definition, the existence of a road manhole cover centered at position (x, y) inevitably indicates $c(x, y) = 1$. Then, the conditional probability $p(E(x, y)|I(x_0, y_0))$ that the patch feature $I(x_0, y_0)$ of a patch $p(x_0, y_0)$ centered at position (x_0, y_0) estimates the certainty of the existence of a road manhole cover centered at position (x, y) is expressed as:

$$\begin{aligned} &p(E(x, y)|I(x_0, y_0)) \\ &\quad = p(E(x, y), c(x, y) = 1|I(x_0, y_0)) \\ &\quad = p((E(x, y)|c(x, y) = 1, I(x_0, y_0))p(c(x, y) = 1|I(x_0, y_0)). \end{aligned} \tag{9}$$

Both terms in (9) can be estimated by passing the patch feature $I(x_0, y_0)$ through the trees in the constructed random forest. The information stored at the leaf nodes that the patch finally

Fig. 7. Illustration of RIEGL VMX-450 MLS system and its components.

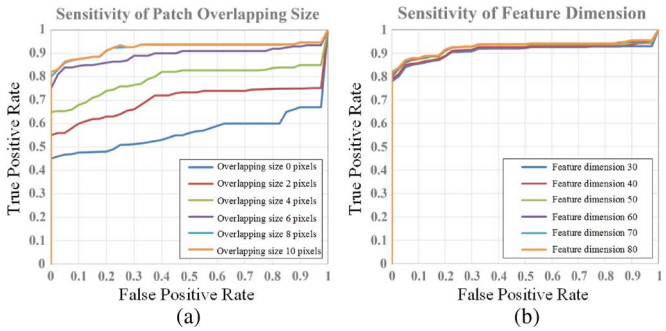


Fig. 8. Illustration of the ROC curves for different parameter configurations of (a) patch overlapping size, and (b) patch feature dimension.

arrives at is used to compute the conditional probability. For a single tree $\boldsymbol{t}$, the probability estimation is given by:

$$p\left(E(x,y)|I(x_0,y_0);\boldsymbol{t}\right) = \frac{C_{\mathrm{L}}}{2\pi\sigma^2} \times \exp\left(-\frac{(x-x_0)^2+(y-y_0)^2}{2\sigma^2}\right) \tag{10}$$

where σ^2 is the variance of the Gaussian Parzen window [27]. By integrating the entire random forest $\{\boldsymbol{t}_i\}_{i=1}^{T}$, we average the probabilities coming from all trees to generate a forest-based probability estimation:

$$p\left(E(x,y)|I(x_0,y_0);\{\boldsymbol{t}_i\}_{i=1}^{T}\right) = \frac{1}{T}\sum_{i=1}^{T} p\left(E(x,y)|I(x_0,y_0);\boldsymbol{t}_i\right). \tag{11}$$

This forest-based probability estimation casts a weighted vote by a single patch $p(x_0,y_0)$ about the existence of a road manhole cover in its vicinity. Then, by aggregating the votes cast from all the patches, we construct a 2-D voting space $VS(x,y)$, where each position (x,y) contains the accumulated votes about the existence of a road manhole cover centered at this position:

$$VS(x,y) = \sum_{p(x_0,y_0)} p\left(E(x,y)|I(x_0,y_0);\{\boldsymbol{t}_i\}_{i=1}^{T}\right). \tag{12}$$

The operations suggested by (10)–(12) might be inefficient when constructing the voting space. As an efficient alternative, first, each patch $p(x_0,y_0)$ casts a vote $(1/T)\sum_{i=1}^{T} C_{\mathrm{L}}^{i}$, where C_{L}^{i} is the label proportion information stored at the leaf node that the patch ends up in the i'th tree, to the position (x_0,y_0). Then, the voting space is obtained by Gaussian-filtering the votes accumulated at each position.

Finally, based on the constructed voting space, the centers of road manhole covers are estimated by ascertaining the positions with local maxima values via a traditional non-maximum suppression process (see Fig. 6).

IV. Results and Discussion

A. MLS System and Point Cloud Data Sets

In this paper, the 3-D MLS point cloud data were acquired using a state-of-the-art RIEGL VMX-450 MLS system (see Fig. 7) on the urban roads in Xiamen City, a port city in southeast China. The VMX-450 system was mounted on the roof of a Buick minivan with a driving speed of approximately 30–50 km/h when collecting the point cloud data. As shown in Fig. 7, the VMX-450 system is integrated with two full-view RIEGL VQ-450 laser scanners, four high-resolution digital cameras, and a set of position and orientation systems, which include two dual-frequency global navigation satellite system (GNSS) antennas, an inertial measurement unit (IMU), and a wheel mounted distance measurement indicator (DMI). The two laser scanners are respectively installed on the left and right sides of the main body with an "X" configuration pattern. The two laser scanners rotate to emit laser beams with a maximum effective measurement rate of 1.1 million measurements per second and a line scan speed of up to 400 scans per second. The accuracy and precision of the resultant 3-D point cloud data are within 8 mm and 5 mm, respectively.

TABLE I
Parameter Configurations in Road Manhole Cover Detection

$n\times n$ (pixels)	D_1	D_2	$d_{\max}$	$N_{\min}$	n_0 (pixels)	T
36×36	500	50	15	20	8	15

From the collected point cloud data, three data sets covering the areas of Siming Road South (SRS), Software Park Phase II (SPP), and International Conference and Exhibition Center (ICEC) were selected for evaluating the performance of the proposed road manhole cover detection algorithm. These three selected road sections cover urban roads paved with different materials (cement and asphalt) and of different geometric and road surface conditions. The SRS data set contains about 406 million points and has a road section of about 2998 m. This is an old asphalt-paved two-directional-four-lane road near the seaside with many road markings and some cracks caused by high traffic loads and moisture. The SPP data set contains about 714 million points and has a length of about 3105 m in the road direction. This is a newly constructed cement-paved two-directional-two-lane road with little road makings and cracks. The ICEC data set contains 568 million points and has a distance of about 2947 m along the road. This is an old asphalt-paved two-directional-two-lane road with many road markings and serious cracks caused by high traffic loads.

TABLE II
ROAD MANHOLE COVER DETECTION RESULTS, QUANTITATIVE EVALUATIONS, AND COMPUTING TIME

Dataset	Ground Truth	Detection Result		Quantitative Evaluation				Computing Time (seconds)	
	Manhole Cover	Manhole Cover	False Positive	Completeness	Correctness	Quality	F_1-measure	Preprocessing	Detection
SRS	186	179	7	0.962	0.962	0.927	0.962	50.95	167.27
SPP	201	194	9	0.965	0.956	0.924	0.960	53.52	174.13
ICEC	147	137	6	0.932	0.958	0.895	0.945	49.73	163.98
Total	534	510	22	0.955	0.959	0.917	0.957	154.20	505.38

B. Parameter Sensitivity Analysis

At both training and detection stages, the configurations of the following two parameters have a significant impact on the performance of the proposed algorithm: patch overlapping size (n_o) between two adjacent patches and dimension of the high-order patch feature (D_2). In order to obtain an optimal configuration for these two parameters, we conducted a group of experiments to test the performance of each parameter configuration on the road manhole cover detection results. The testing results were presented and analyzed using receiver operating characteristic (ROC) curves (see Fig. 8). As shown in Fig. 8(a), when the patch overlapping size increases, the performance improves accordingly. However, when the overlapping size is greater than 8 pixels, the performance almost stays unchanged. In fact, with the increase of the overlapping size, more redundant patches will be generated and a road manhole cover will be more likely to be segmented into a single patch. However, when the overlapping size increases, the number of the generated patches will increase dramatically at detection stage. This will surely bring more computational burdens and slow down the detection speed. Therefore, considering both performance and time complexity, we set the overlapping size at 8 pixels. As shown in Fig. 8(b), the dimension of the high-order patch feature has a very slight influence on the detection performance. This actually benefits from the implementation of the joint DBM model for constructing the multi-layer feature generation model. The multi-layer feature generation model produces a high-order feature representation for a local image patch. The high-order feature representation is actually an integration of a set of low-order features rather than a single feature. Therefore, it is more powerful and distinctive for depicting a local image patch. In this paper, we set the patch feature dimension at 50.

C. Road Manhole Cover Detection

We applied the proposed road manhole cover detection algorithm to the three selected data sets (SRS, SPP, and ICEC data sets) to evaluate its performance. The parameters and their configurations used at training and detection stages are listed in Table I. To narrow the searching regions and reduce the time complexity, these data sets were first preprocessed to segment road surface points, which were subsequently rasterized into georeferenced intensity images. Then, the proposed road manhole cover detection algorithm was carried out to detect road manhole covers based on the georeferenced intensity images. To quantitatively evaluate the accuracy and correctness of the road manhole cover detection results on the three data sets, we used the following four measures: *completeness*, *correctness*, *quality* [28], and F_1*-measure* [16]. Completeness measures the proportion of true positives in the ground truth; correctness evaluates the proportion of true positives in the detection result; quality and F_1-measure are two overall measures for depicting the overall performance of the detection result. These four quantitative measures are defined as follows:

$$\text{completeness} = \frac{TP}{TP + FN} \tag{13}$$

$$\text{correctness} = \frac{TP}{TP + FP} \tag{14}$$

$$\text{quality} = \frac{TP}{TP + FP + FN} \tag{15}$$

$$F_1\text{-measure} = \frac{2 \cdot \text{completeness} \cdot \text{correctness}}{\text{completeness} + \text{correctness}} \tag{16}$$

where TP, FN, and FP are the number of true positives, false negatives, and false positives, respectively.

The road manhole cover detection results, as well as their quantitative evaluations, are detailed in Table II. As reflected in Table II, compared to the ground truths, the majority of road manhole covers were correctly detected in each data set. In addition, the number of false positives is quite low. Fig. 9(a) shows some typical examples of georeferenced intensity images containing different types of road manhole covers under different road surface conditions. Their associated voting images and detected road manhole covers are respectively presented in Fig. 9(b) and (c). The voting images are a visual representation of the voting space constructed after random forest voting. As shown by the red boxes labeled #1 and #2 in Fig. 9(c), these two road manhole covers are partially painted with white road markings. Benefiting from the use of the multi-layer feature generation model to construct high-order feature representations for local image patches, our proposed algorithm can well handle such conditions and correctly detect these road manhole covers. However, as shown by the blue boxes labeled #3, #4, #5, and #6, these four road manhole covers show quite low contrasts with the background road surface and they are of higher intensities than those of the normal road manhole covers. Because of heavy traffic flows, these road manhole covers are covered with thick dusts, thereby resulting in different intensities and low contrasts with the background road surface. Such conditions usually happen to the road manhole covers in the middle of the road rather than near the road boundary. Therefore, our proposed algorithm failed to detect such road manhole covers. In addition, as shown by the yellow boxes labeled #7 and #8, these two patches show very similar appearance to the road manhole covers and they obtain very high votes in the voting space. Such similarities are usually

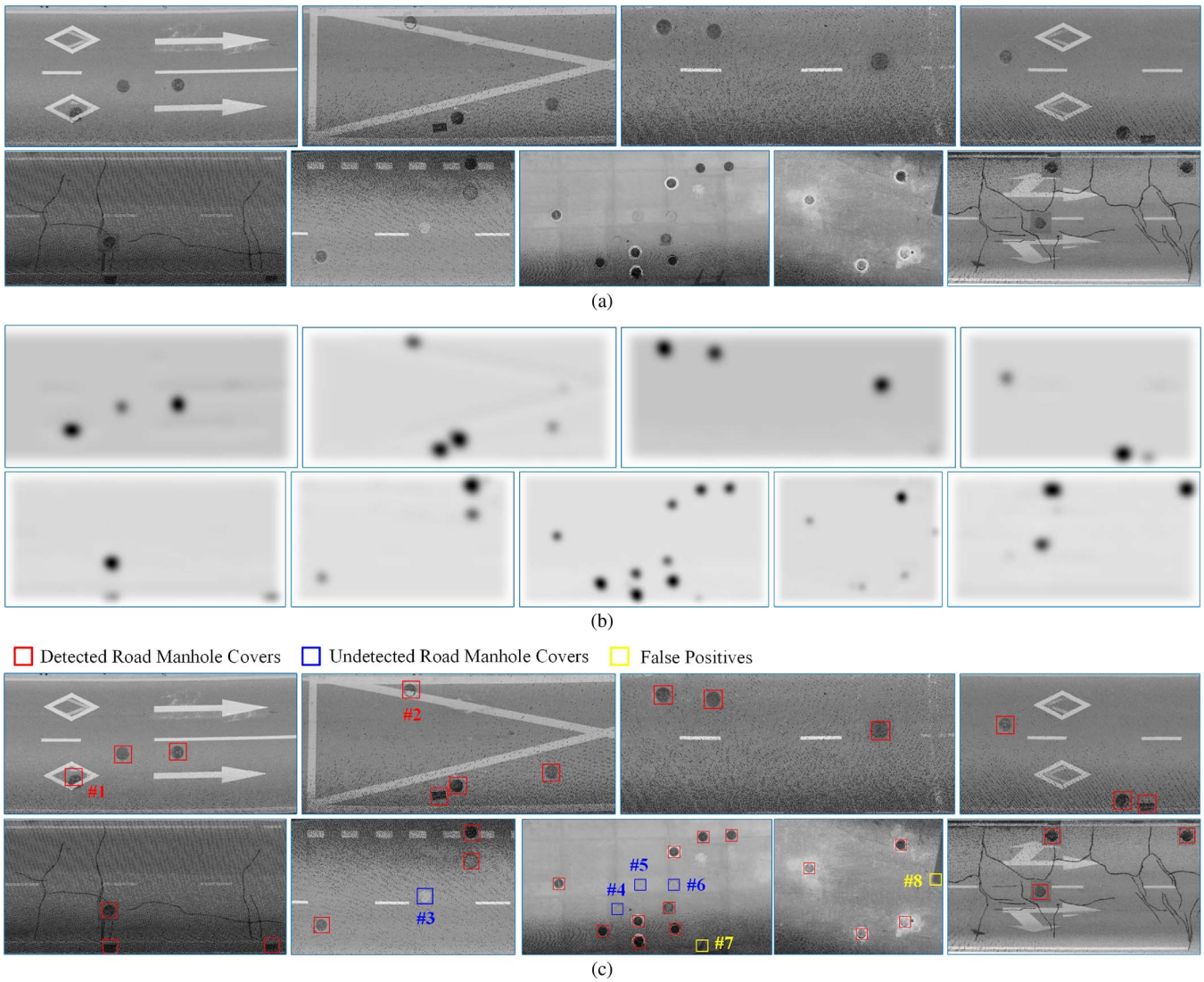


Fig. 9. Illustrations of (a) some typical georeferenced intensity images containing different types of road manhole covers under different road surface conditions, (b) voting images which are visual representations of the voting spaces constructed after random forest voting, and (c) road manhole cover detection results.

caused by the shadows of mobile obstacles on the road and the large-area asphalt blocks used for repairing road surface cracks. Thus, our proposed algorithm falsely detected them as road manhole covers. Through quantitative evaluations, the proposed algorithm achieves an average completeness, correctness, quality, and F_1-measure of 0.955, 0.959, 0.917, and 0.957, respectively, on the three selected data sets. On the whole, the proposed algorithm achieves very promising performance and obtains very good accuracy in detecting varying types of road manhole covers under different road surface conditions from MLS point clouds.

The proposed algorithm was coded using C++ on the platform of Visual Studio 2010 and run on an HP Z820 8-core-16-thread workstation. To evaluate the computational performance of the proposed algorithm, the model training time, preprocessing time, and road manhole cover detection time were also recorded. In this paper, the number of positive and negative training samples for training the multi-layer feature generation model and the random forest model was 7820 and 7820, respectively. The time cost for training the multi-layer feature generation model and the random forest model was approximately 5.3 and 0.6 hours, respectively. At detection stage, each data set was first partitioned into a group of segments with a length of about 50 m along the road based on the vehicle's trajectory. Then, all the segments were distributed to a multi-thread computing environment with 16 parallel threads. Hence, 16 segments were simultaneously being processed. Such a strategy can dramatically improve the performance and reduce the time complexity of the proposed algorithm. After statistics, the preprocessing time for road surface segmentation and georeferenced intensity image rasterization was 50.95, 53.52, and 49.73 seconds, respectively, for the SRS, SPP, and ICEC data sets. The computing time for road manhole cover detection was 167.27, 174.13, and 163.98 seconds, respectively, for the SRS, SPP, and ICEC data sets (see Table II). In conclusion, the proposed algorithm is very feasible and promising for rapidly and automatically detecting road manhole covers using MLS point cloud data.

TABLE III
ROAD MANHOLE COVER DETECTION RESULTS AND QUANTITATIVE EVALUATIONS ON THE SYNTHETIC DATA OF DIFFERENT POINT DENSITIES

Dataset	Synthetic Data	Detection Result		Quantitative Evaluation			
	Point Interval (points)	Manhole Cover	False Positive	Completeness	Correctness	Quality	F_1-measure
SRS	1	179	7	0.962	0.962	0.927	0.962
	2	171	9	0.919	0.950	0.877	0.934
	3	165	15	0.887	0.917	0.821	0.902
	4	137	26	0.737	0.840	0.646	0.785
	5	124	32	0.667	0.795	0.569	0.725
SPP	1	194	9	0.965	0.956	0.924	0.960
	2	188	10	0.935	0.949	0.891	0.942
	3	174	19	0.866	0.902	0.791	0.884
	4	155	23	0.771	0.871	0.692	0.818
	5	138	29	0.687	0.826	0.600	0.750
ICEC	1	136	6	0.925	0.958	0.889	0.941
	2	134	8	0.912	0.944	0.865	0.928
	3	126	13	0.857	0.906	0.788	0.881
	4	113	19	0.769	0.856	0.681	0.810
	5	97	25	0.660	0.795	0.564	0.721

However, due to the scanning mechanism of MLS systems, the mobile obstacles (e.g., moving vehicles and pedestrians) on the road might generate occlusions for the measurement of road surfaces. Thus, the road manhole covers occluded by the mobile obstacles cannot be completely scanned by using MLS systems. Such occlusions cause challenges for accurately and completely measuring road manhole covers. In practice, in order to reduce the influence of mobile obstacles on the measurement of road surfaces, the mapping missions should be carried out at night or during low-traffic-flow time period. In our mapping mission, the occlusions of road manhole covers also exist in the three selected data sets. Therefore, the ground truths in Table II are actually the road manhole covers completely and correctly scanned in the point cloud data. In addition, the materials and colors of road manhole covers and the unevenness and materials of the road surface also bring challenges for the accurate measurement of road manhole covers. In this paper, to improve the performance of the proposed algorithm, we selected different shapes, colors, and materials of road manhole covers as positive training samples and different materials and conditions of road surface patches as negative training samples for training the high-order feature generation model and the random forest model (see Fig. 3). Moreover, three selected road sections of different road surface materials and road surface conditions were used for evaluating the performance of the proposed algorithm. The experimental results demonstrate the feasibility and accuracy of the proposed algorithm in detecting different types of road manhole covers under different road surface conditions.

D. Performance Evaluation

Different MLS systems are configured with different laser scanners and navigation systems. Due to different scanning rates of laser scanners and accuracy of navigation systems, the point density and data quality of the collected point clouds vary greatly among different MLS systems. Thus, it is necessary to evaluate the performance of the proposed algorithm in handling lower-grade MLS point clouds toward road manhole cover detection. However, due to the lack of publicly available point clouds of other MLS systems, we generated a group of synthetic point cloud data of different point densities and data qualities using the point clouds collected by our RIEGL VMX-450 system. Then, we conducted a group of simulated experiments to evaluate the performance of the proposed algorithm in detecting road manhole covers from the synthetic data. In this simulated study, the segmented road surface point clouds in the SRS, SPP, and ICEC data sets were used for performance evaluation.

To generate point clouds of different point densities, which represent the point clouds collected by the laser scanners of different scanning rates, a raw point cloud was sampled along the scan lines to select a set of laser points with certain point intervals. Such synthetic different-density point clouds were used to detect road manhole covers. To this end, each of the three data sets was sampled scan line by scan line to generate a group of synthetic data using the following point intervals: 1 point, 2 points, 3 points, 4 points, and 5 points. These synthetic data correspond to the point cloud data collected by the MLS systems with laser scanning rates of approximately 550 000, 367 000, 275 000, 220 000, and 183 000 measurements per second, respectively. Table III lists the road manhole cover detection results on the synthetic data. Quantitative evaluations using completeness, correctness, quality, and F_1-measure were also performed on the detection results, as shown in Table III. With the decrease of point density, the detail and quality of the georeferenced intensity image is degraded greatly, thereby resulting in different feature representations and lower saliencies of road manhole covers in the degraded georeferenced intensity image. Consequently, the road manhole cover detection performance decreases with the decrease of point density. As reflected in Table III, when the point interval is $\leq$ 3 points (laser scanning rate is $\geq$ 275 000 measurements per second), the proposed algorithm can still obtain promising performance on the synthetic data. However, when the point interval exceeds 3 points, the performance decreases greatly, especially for the SRS and ICEC data sets. Therefore, point density affects significantly on the detection of road manhole covers from MLS data. For the lower-grade MLS systems with lower laser scanning rates (e.g., < 275 000 measurements per second), it

TABLE IV
ROAD MANHOLE COVER DETECTION RESULTS AND QUANTITATIVE EVALUATIONS ON THE SYNTHETIC DATA OF DIFFERENT DATA QUALITIES

Dataset	Synthetic Data	Detection Result		Quantitative Evaluation			
	Variance	Manhole Cover	False Positive	Completeness	Correctness	Quality	F_1-measure
SRS	0.05	179	7	0.962	0.962	0.927	0.962
	0.10	173	12	0.930	0.935	0.874	0.932
	0.15	168	14	0.903	0.923	0.840	0.913
	0.20	161	17	0.866	0.904	0.793	0.885
	0.25	155	22	0.833	0.876	0.745	0.854
SPP	0.05	194	10	0.965	0.951	0.919	0.958
	0.10	190	13	0.945	0.936	0.888	0.940
	0.15	186	16	0.925	0.921	0.857	0.923
	0.20	183	18	0.910	0.910	0.836	0.910
	0.25	178	24	0.886	0.881	0.791	0.883
ICEC	0.05	137	8	0.932	0.945	0.884	0.938
	0.10	133	13	0.905	0.911	0.831	0.908
	0.15	129	17	0.878	0.884	0.787	0.881
	0.20	122	20	0.830	0.859	0.731	0.844
	0.25	117	23	0.796	0.836	0.688	0.816

TABLE V
ROAD MANHOLE COVER DETECTION RESULTS, QUANTITATIVE EVALUATIONS, AND COMPUTING TIME FOR DIFFERENT ALGORITHMS

Dataset	Algorithm	Detection Result		Quantitative Evaluation				Time (seconds)
		Manhole Cover	False Positive	Completeness	Correctness	Quality	F_1-measure	
SRS	Marked Point Process [13]	164	19	0.882	0.896	0.800	0.889	428.64
	Multi-Scale Tensor Voting [14]	138	21	0.742	0.868	0.667	0.800	184.31
SPP	Marked Point Process [13]	182	17	0.905	0.915	0.835	0.910	495.22
	Multi-Scale Tensor Voting [14]	177	19	0.881	0.903	0.805	0.892	198.73
ICEC	Marked Point Process [13]	129	21	0.878	0.860	0.768	0.869	399.59
	Multi-Scale Tensor Voting [14]	108	26	0.735	0.806	0.624	0.769	175.37

is not suitable to perform road manhole cover detection on the point clouds acquired by such systems.

Due to the mechanism of laser scanning, the backscattered laser pulse intensity is influenced greatly by the quality of laser scanners, quality of navigation systems, material property of the measured objects, and range between the object and the scanner. Such intensity perturbations result in speckle-like noise (multiplicative noise) in the georeferenced intensity images. To generate point clouds of different qualities, we superimposed the georeferenced intensity image with different levels of speckle noise. Such speckle-noise-contaminated georeferenced intensity images were used for road manhole cover detection. To this end, the georeferenced intensity images of the three data sets were contaminated by superimposing speckle noises with mean zero and the following different variances: 0.05, 0.10, 0.15, 0.20, and 0.25. Table IV details the road manhole cover detection results on the synthetic data and the quantitative evaluations on the detection results. Specifically, the contamination of speckle noise degraded the quality of the georeferenced intensity images. Such degradation of the georeferenced intensity images lowered the saliencies of road manhole covers, thereby leading to the decrease of detection performance. However, benefited from the use of the multi-layer feature generation model, when the level of speckle noise is relatively low (≤ 0.20), the proposed algorithm can still obtain promising detection performance on the synthetic data. When the georeferenced intensity images are contaminated by high levels of speckle noise (≥ 0.25), the detection performance decreases dramatically, especially for the ICEC data set. In conclusion, the quality of point clouds influences greatly on the detection of road manhole covers from MLS data. Thus, for the lower-grade MLS systems with lower point cloud qualities, it is not feasible to perform road manhole cover detection on the point clouds collected by such systems.

E. Comparative Studies

To further examine the performance of the proposed algorithm, we also conducted a group of comparative studies to compare our proposed algorithm with an existing marked point process based road manhole cover detection algorithm [13] and a multi-scale tensor voting based road manhole cover detection algorithm [14]. The marked point process based algorithm considered the geometric and intensity properties of road manhole covers. The detection task was achieved by simulating and optimizing the marked point process of disks and rectangles. The multi-scale tensor voting based algorithm utilized the intensity information of road manhole covers for segmenting low-intensity regions. The detection framework consisted of a chain of conventional techniques including intensity thresholding, tensor voting, and region growing. These two algorithms have the following similarities: 1) both of them are developed to detect road manhole covers based on MLS point clouds; 2) both of them utilize georeferenced intensity

images, which are generated from road surface point clouds, to detect road manhole covers; and 3) both of them have the capability of detecting both circular-shaped and rectangular-shaped road manhole covers. In this comparative study, we used the entire data of the three selected data sets with different road surface conditions to test these two algorithms for road manhole cover detection. Table V lists the road manhole cover detection results obtained by using these two algorithms. Quantitative evaluations using completeness, correctness, quality, and F_1-measure were also performed on the detection results, as shown in Table V. Comparatively, the multi-scale tensor voting based algorithm obtained the lowest performance and generated relatively less true positives and more false positives. Because of the complex road conditions and the varying and imbalanced intensities of road manhole covers in these data sets, the multi-scale tensor voting based algorithm cannot obtain promising performance by using intensity thresholding and segmentation techniques. However, our proposed algorithm achieved a better performance than the other algorithms and generated relatively less false positives.

To compare the computational efficiency of these algorithms, the computing time was also recorded for each data set, as shown in Table V. For fairness, all the algorithms were executed in a multi-thread computing environment with 16 parallel threads. Comparatively, the multi-scale tensor voting algorithm took similar time to our proposed algorithm, while the marked point process based algorithm took more time than the other two algorithms because of the complex simulation and optimization procedures of the marked point process. In conclusion, compared to the marked point process based algorithm and the multi-scale tensor voting based algorithm, our proposed algorithm achieves better performance in correctly detecting different types of road manhole covers under different road surface conditions and obtains promising computational efficiency.

V. Conclusion

In this paper, we have proposed a novel algorithm for automatically detecting urban road manhole covers using MLS point cloud data. The proposed algorithm have been successfully applied to three selected MLS point cloud data sets, which were acquired by a state-of-the-art RIEGL VMX-450 MLS system, to detect both circular-shaped and rectangular-shaped road manhole covers. Through quantitative evaluations on the detection results, the proposed algorithm achieved an average completeness, correctness, quality, and F_1-measure of 0.955, 0.959, 0.917, and 0.957, respectively, on the three selected data sets. Computational performance analysis showed that, by using a multi-thread computing strategy, the proposed algorithm can rapidly handle large-volume point cloud data toward road manhole cover detection. In addition, comparative studies have demonstrated that the proposed algorithm outperforms the other two algorithms in correctly detecting road manhole covers with varying appearances and under heterogeneous road surface conditions. Therefore, MLS point cloud data are a promising data source for rapidly and accurately measuring road surface features and distresses. They can be exploited to assist in many intelligent transportation related applications.

Acknowledgment

The authors would like to thank the anonymous reviewers for their valuable comments.

References

[1] N. N. Zheng *et al.*, "Toward intelligent driver-assistance and safety warning systems," *IEEE Intell. Syst.*, vol. 19, no. 2, pp. 8–11, Mar./Apr. 2004.

[2] H. Cheng, N. Zheng, X. Zhang, J. Qin, and H. van de Wetering, "Interactive road situation analysis for driver assistance and safety warning systems: Framework and algorithms," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 1, pp. 157–167, Mar. 2007.

[3] J. Choi *et al.*, "Environment-detection-and-mapping algorithm for autonomous driving in rural or off-road environment," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 2, pp. 974–982, Jun. 2012.

[4] A. Broggi *et al.*, "Extensive tests of autonomous technologies," *IEEE Trans. Intell. Transp. Syst.*, vol. 14, no. 3, pp. 1403–1415, Sep. 2013.

[5] S. Murray *et al.*, "Mobile mapping system for the automated detection and analysis of road delineation," *IET Intell. Transp. Syst.*, vol. 5, no. 4, pp. 221–230, Dec. 2011.

[6] M. Brogan, S. McLoughlin, and C. Deegan, "Assessment of stereo camera calibration techniques for a portable mobile mapping system," *IET Comput. Vis.*, vol. 7, no. 3, pp. 209–217, Jun. 2013.

[7] K. Williams, M. J. Olsen, G. V. Roe, and C. Glennie, "Synthesis of transportation applications of mobile LiDAR," *Remote Sens.*, vol. 5, no. 9, pp. 4652–4692, Sep. 2013.

[8] N. Tanaka and M. Mouri, "A detection method of cracks and structural objects of the road surface image," in *Proc. IAPR Workshop Mach. Vis. Appl.*, Tokyo, Japan, 2000, pp. 387–390.

[9] H. Niigaki, J. Shimamura, and M. Morimoto, "Circular object detection based on separability and uniformity of feature distributions using Bhattacharyya coefficient," in *Proc. 21st Int. Conf. Pattern Recog.*, Tsukuba, Japan, 2012, pp. 2009–2012.

[10] C. Drewniok and D. Rohr, "Model-based detection and localization of circular landmarks in aerial images," *Int. J. Comput. Vis.*, vol. 24, no. 3, pp. 187–217, Sep. 1997.

[11] R. Timofte and L. van Gool, "Multi-view manhole detection, recognition, and 3D localization," in *Proc. IEEE Int. Conf. Comput. Vis. Workshops*, Barcelona, Spain, 2011, pp. 188–195.

[12] Y. Cheng, Z. Xiong, and Y. H. Wang, "Improved classical Hough transform applied to the manhole cover's detection and location," *Opt. Tech.*, vol. 32, no. S1, pp. 504–508, 2006.

[13] Y. Yu, J. Li, H. Guan, C. Wang, and J. Yu, "Automated detection of road manhole and sewer well covers from mobile LiDAR point clouds," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 9, pp. 1549–1553, Sep. 2014.

[14] H. Guan *et al.*, "Automated extraction of manhole covers using mobile LiDAR data," *Remote Sens. Lett.*, vol. 5, no. 12, pp. 1042–1050, Dec. 2014.

[15] S. Ji, Y. Shi, and Z. Shi, "Manhole cover detection using vehicle-based multi-sensor data," in *Proc. Int. Archives Photogramm. Remote Sens. Spatial Inf. Sci.*, Melbourne, Australia, vol. 39-B3, 2012, pp. 281–284.

[16] H. Guan *et al.*, "Using mobile laser scanning data for automated extraction of road markings," *ISPRS J. Photogramm. Remote Sens.*, vol. 87, pp. 93–107, Jan. 2014.

[17] G. Carneiro and J. C. Nascimento, "Combining multiple dynamic models and deep learning architectures for tracking the left ventricle endocardium in ultrasound data," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 11, pp. 2592–2607, Nov. 2013.

[18] B. Chen *et al.*, "Deep learning with hierarchical convolutional factor analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1887–1901, Aug. 2013.

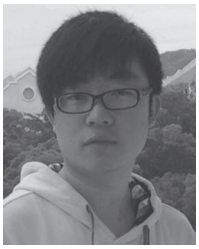
[19] R. Salakhutdinov, J. B. Tenenbaum, and A. Torralba, "Learning with hierarchical-deep models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1958–1971, Aug. 2013.

[20] T. Tieleman, "Training restricted Boltzmann machines using approximations to the likelihood gradient," in *Proc. 25th Int. Conf. Mach. Learn.*, Helsinki, Finland, 2008, pp. 1064–1071.

[21] R. Salakhutdinov and G. Hinton, "Deep Boltzmann machines," in *Proc. 12th Int. Conf. Artif. Intell. Statist.*, Clearwater Beach, FL, USA, 2009, vol. 2, pp. 448–455.

[22] R. Salakhutdinov and G. Hinton, "An efficient learning procedure for deep Boltzmann machines," *Neural Comput.*, vol. 24, no. 8, pp. 1967–2006, 2012.

[23] B. Ni, S. Yan, M. Wang, A. A. Kassim, and Q. Tian, "High-order local spatial context modeling by spatialized random forest," *IEEE Trans. Image Process.*, vol. 22, no. 2, pp. 739–751, Feb. 2013.
[24] S. Du and S. Chen, "Salient object detection via random forest," *IEEE Signal Process. Lett.*, vol. 21, no. 1, pp. 51–54, Jan. 2014.
[25] S. Havens, H. Marshall, C. Pielmeier, and K. Elder, "Automatic grain type classification of snow micro penetrometer signals with random forests," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 6, pp. 3328–3335, Jun. 2012.
[26] D. Mahapatra, "Analyzing training information from random forests for improved image segmentation," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1504–1512, Apr. 2014.
[27] J. Gall, A. Yao, N. Razavi, L. V. Gool, and V. Lempitsky, "Hough forests for object detection, tracking, and action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 11, pp. 2188–2202, Nov. 2011.
[28] M. Rutzinger, F. Rottensteiner, and N. Pfeifer, "A comparison of evaluation techniques for building extraction from airborne laser scanning," *IEEE J. Sel. Topics Earth Observ. Remote Sens.*, vol. 2, no. 1, pp. 11–20, Mar. 2009.



Yongtao Yu received the B.S. degree in computer science and technology from Xiamen University, Xiamen, China, in 2010. He is currently working toward the Ph.D. degree in computer science and technology with the Department of Computer Science, Xiamen University. He has coauthored more than 20 research papers published in refereed journals and proceedings. His current research interests include pattern recognition, computer vision, machine learning, mobile laser scanning, and information extraction from 3-D point clouds.

Haiyan Guan (M'15) received the Ph.D. degree in geomatics from the University of Waterloo, Waterloo, ON, Canada, in 2014. She is currently a Professor at the College of Geography and Remote Sensing, Nanjing University of Information Science and Technology, Nanjing, China. She has coauthored more than 30 research papers published in refereed journals, books, and proceedings. Her research interests include airborne, terrestrial, and mobile laser scanning data processing algorithms and 3-D spatial modeling and reconstruction of critical infrastructure and landscape.

Zheng Ji received the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2007. He is currently an Associate Professor at the School of Remote Sensing Information and Engineering, Wuhan University. He has coauthored over 20 research papers in refereed journals and proceedings. His research interests include terrestrial LiDAR data processing, close-range photogrammetry, and UAV data processing.