

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/311630470>

# A Bayesian–Network–Based Classification Method Integrating Airborne LiDAR Data With Optical Images

Article in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing · December 2016

DOI: 10.1109/JSTARS.2016.2628775

CITATIONS

28

READS

729

3 authors:



**Zhizhong Kang**

China University of Geosciences (Beijing)

99 PUBLICATIONS 776 CITATIONS

[SEE PROFILE](#)



**Juntao Yang**

Shandong University of Science and Technology

22 PUBLICATIONS 180 CITATIONS

[SEE PROFILE](#)



**Ruofei Zhong**

Capital Normal University

96 PUBLICATIONS 777 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



The ISPRS Benchmark on Indoor Modelling [View project](#)



Planetary Structure Mapping and Analysis [View project](#)

# A Bayesian-Network-Based Classification Method Integrating Airborne LiDAR Data With Optical Images

Zhizhong Kang, Juntao Yang, and Ruofei Zhong

**Abstract**—Point cloud classification is of great importance to applications of airborne Light Detection And Ranging (LiDAR) data. In recent years, airborne LiDAR has been integrated with various other sensors, e.g., optical imaging sensors, and thus, the fusion of multiple data types for scene classification has become a hot topic. Therefore, this paper proposes a Bayesian network (BN) model that is suitable for airborne point cloud classification fusing multiple data types. Based on an analysis of the characteristics of LiDAR point clouds and aerial images, we first extract the geometric features from the point clouds and the spectral features from the optical images. The optimal BN structure is then trained using an improved mutual-information-based K2 algorithm to obtain the optimal BN classifier for point cloud classification. Experiments demonstrate that the BN classifier can effectively distinguish four types of basic ground objects, including ground, vegetation, trees, and buildings, with a high accuracy of over 90%. Moreover, compared with other classifiers, the proposed BN classifier can achieve the highest overall accuracies, and in particular, the classifier demonstrates its advantage in the classification of ground and low vegetation points.

**Index Terms**—Airborne light detection and ranging (LiDAR), Bayesian network (BN), classification, optical imagery, point cloud.

## I. INTRODUCTION

THE development of airborne Light Detection And Ranging (LiDAR) provides a new technology for obtaining geospatial information with high spatial resolution, which can accurately capture ground objects. LiDAR has been widely used in such areas as ground data measurement, 3-D city model reconstruction, and vegetation coverage and forest volume estimation.

Point cloud classification is an essential step for practical application of airborne laser scanning (ALS) data, and thus, its accuracy and efficiency are two important factors on which relevant studies focus. The presently proposed methods can be classified into two categories: supervised and unsupervised classification. Both approaches have their own advantages. Unsupervised classification can realize automation (e.g., [1], [2]),

while supervised classification is more adaptable and data driven (e.g., [3], [4]). In this paper, we focus on supervised classification methods integrating airborne LiDAR data with optical images. Common classifiers, including support vector machine (SVM) (e.g., [4]–[6]), random forests (RFs) [7]–[9], AdaBoost [10], Bayesian maximum likelihood [11], [12], and conditional random field [13], [14], have been used to classify point clouds. However, it is very difficult to perform automatic and accurate ground object classification and identification using only airborne LiDAR point cloud data. Chehata *et al.* [8] integrated full waveform features into point cloud classification using an RF classifier. Based on Chehata *et al.*'s work, Guo *et al.* [4] incorporated the spectral information of images into point cloud classification. Khoshelham *et al.* [3] compared different supervised techniques, namely Bayesian, Dempster Shafer and AdaBoost, for fused ALS and image data. However, concerning the integration of spectral information and point cloud features, the classifiers, such as decision tree (Dtree) and its variations, RF, and AdaBoost, may not be able to fully utilize all features (e.g., spectral information) to classify an object due to the characteristics of a Dtree structure.

Instead of directly assigning an object to a class, Bayesian-based methods compute a probability that indicates how likely the object belongs to a class. The object is then assigned to the class from which the highest probability is calculated. Because the probability is calculated in terms of all features, each feature is able to play a role in the classification process. Methods based on Bayesian maximum likelihood have been proposed [11], [12]. However, the Bayesian maximum likelihood classifier assumes that the classes have Gaussian probability distribution functions, which is not necessary for a Bayesian network (BN). A BN is a directed acyclic graph model that is used to express the probabilistic dependencies among variables [15]. Brunn and Weidner [16] proposed an approach to extract buildings from dense digital surface models (DSMs) using hierarchical Bayesian nets. However, to the best of our knowledge, no research introducing BNs into point cloud classification fusing multiple data types has been reported. Therefore, this paper proposes a BN model that is suitable for point cloud classification fusing airborne LiDAR data and optical images. Concerning the classification of point clouds, we focus on the detection of the following four major object classes in urban environments: ground, low vegetation, high vegetation, and buildings.

This paper analyzes the characteristics of two types of data (LiDAR point clouds and aerial images), extends the classification strategy that integrates multisource data, and proposes a

Manuscript received September 16, 2016; revised November 3, 2016; accepted November 8, 2016. This work was supported by the Natural Science Foundation of China under Grant No. 41471360 and the Fundamental Research Funds for the Central Universities under Grant No. 2652015176. (Corresponding author: Zhizhong Kang.)

Z. Kang and J. Yang are with the School of Land Science and Technology, China University of Geosciences, Beijing 100083, China (e-mail: zzkang@cugb.edu.cn; jtyang@cugb.edu.cn).

R. Zhong is with the State Key Laboratory Incubation Base of Urban Environmental Processes and Digital Simulation, Capital Normal University, Beijing 100048, China (e-mail: zrfss@163.com).

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

Digital Object Identifier 10.1109/JSTARS.2016.2628775

BN model that is suitable for the classification of airborne point clouds. First, the geometric features of a point cloud and the spectral features of an image object are extracted as the feature nodes of the BN. The mutual information between all the feature nodes and between the feature nodes and class nodes is then calculated. The mutual information is used as the weights between the nodes. A maximum-weight spanning tree is obtained through calculations. The BN is initialized to generate the initial edge set of the BN. A breadth-first search is then performed on the generated maximum-weight spanning tree using the class nodes as the starting points. In addition, when performing the breadth-first search, the parent-child relationship between adjacent nodes is also determined using a Bayesian-based statistical scoring method to obtain the optimal sequence of nodes. Finally, the obtained sequence of nodes is used as the input of the K2 algorithm [17] to train the optimal BN structure, and the optimal BN classifier is then obtained through parameter learning.

We begin by describing the algorithms used to extract features from both point clouds and images in Section II. Section III introduces the BN classifier for point cloud classification fusing multiple data types. Section IV discusses the test results. We then offer conclusions and suggestions for further research in Section V.

## II. FEATURE EXTRACTION FROM POINT CLOUDS AND OPTICAL IMAGES

An airborne LiDAR system can simultaneously obtain point cloud data and charge-coupled device image data. Therefore, we focus on the use of the geometric features of the point clouds and the spectral features of the image objects.

### A. Feature Extraction From Point Clouds

1) *Definition of Neighborhoods*: The calculation of the geometric features of a point cloud often requires the statistical analysis of other points in its 3-D neighborhoods. Therefore, the shape and size of the neighborhoods must be considered when calculating the geometric features of a point cloud [18]. The following three types of neighborhood definition methods are commonly used [19]–[21]: 1) a spherical neighborhood definition  $N_s$  with a fixed radius of  $r_s \in R$  around the point that is being classified; 2) a vertical cylindrical neighborhood definition  $N_c$  with a bottom surface fixed radius of  $r_c \in R$  around the point that is being classified; and 3) a neighborhood definition  $N_k$  consisting of the  $k$  points that are closest to the point that is being classified. Because of the variations in the airborne point cloud density, this paper uses definition 3 to define the neighborhood of a point that is being classified. After the neighborhood is determined, the geometric features of a point cloud can be extracted.

2) *Definition of the Geometric Features of a Point Cloud*: There are three main types of extracted geometric features of a point cloud [4], [10], [18], [19]: height-based features, plane-based features, and eigenvalue-based features.

1) *Height-based features*. The mean normalized height ( $N_H$ ) is expressed using the difference between the DSM and the digital elevation model (DEM) [22]. Fig. 1 shows

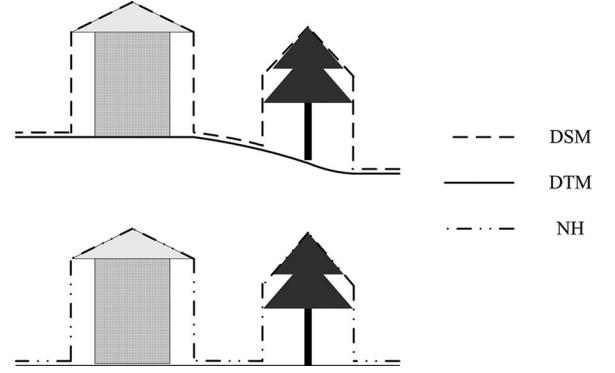


Fig. 1. Relationship among  $N_H$ , the DSM, and the DEM.

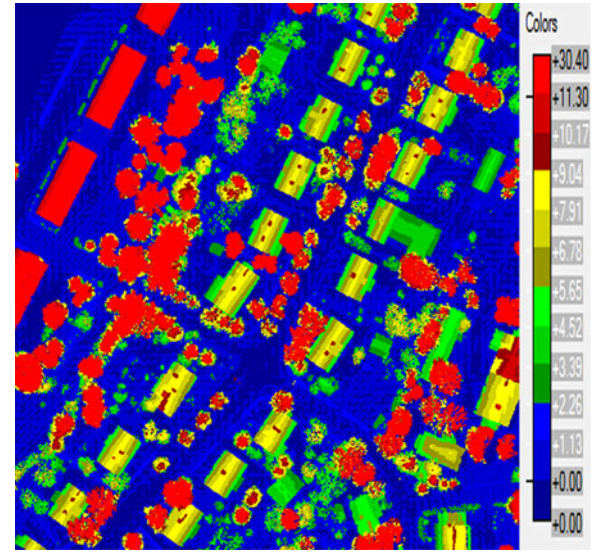


Fig. 2. Normalized height  $N_H$ .

the relationships among  $N_H$  (see Fig. 2), the DSM, and the DEM.  $N_H$  can eliminate the effect of topographic relief (see the bottom part of Fig. 1). The value of the height variance ( $\sigma^2$ ) is relatively high at the locations of vegetation (see Fig. 3).

2) *Plane-based features*: residual ( $R_z$ ) of plane fitting. A plane is fitted to the point that is being classified and all the points in its neighborhoods using the least-squares method. The distances between all the points and the fitted plane are calculated. Then, we have

$$R_z = \sum_{i \in V_p} \frac{(d_i)^l}{l} \quad (1)$$

where  $d_i$  represents the distance between point  $i$  and the plane. Here,  $l=1,2$ , as adopted in [8] and [23]. The values of  $d_i$  at the locations of vegetation are relatively high.

3) *Eigenvalue-based features*. Assuming that  $C_p$  represents the neighborhood covariance matrix of point  $p$ ,  $C_p$  can be calculated using the following equation:

$$C_p = \frac{1}{|N|} \sum_{P_i \in N} (P_i - \bar{P}) (P_i - \bar{P})^T \quad (2)$$



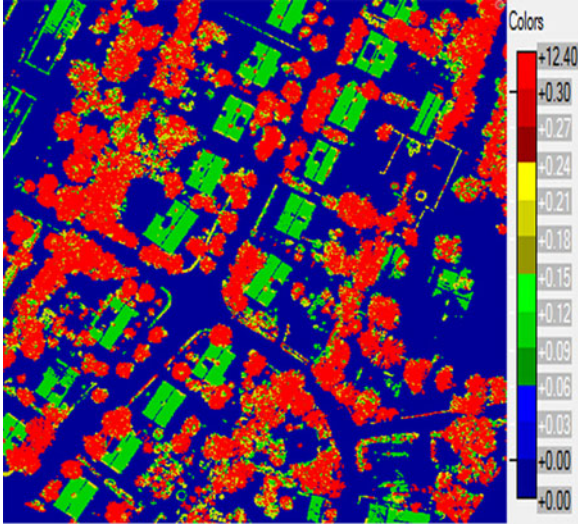
Fig. 3. Variance of height  $\sigma^2$ .

TABLE I  
VARIATION OF THE VALUES OF THE FEATURES OVER PLANAR  
AND NONPLANAR AREAS

Feature	Nonplanar area	Planar area	Linear edge
$\lambda_3$	High	Low	–
$A_\lambda$	Low	High	–
$P_\lambda$	Low	High	–
$S_\lambda$	High	Low	–
$L_\lambda$	–	–	High

where  $\bar{P} = \frac{1}{|N|} \sum_{P_i \in N} P_i$ . The covariance matrix  $C_p$  is subjected to eigenvalue decomposition, and  $\lambda_1$ ,  $\lambda_2$ , and  $\lambda_3$  are the eigenvalues of  $L_\lambda$ , where  $\lambda_1 > \lambda_2 > \lambda_3$ . Because the ranges of the eigenvalues calculated from the covariance matrices of different points are different, it is necessary to normalize the eigenvalues using

$$\lambda_i = \frac{\lambda_i}{\sum_i \lambda_i}, i = 1, 2, 3 \quad (3)$$

$$\left. \begin{aligned} A_\lambda &= (\lambda_1 - \lambda_3) / \lambda_1 \\ P_\lambda &= (\lambda_2 - \lambda_3) / \lambda_1 \\ S_\lambda &= \lambda_3 / \lambda_1 \\ L_\lambda &= (\lambda_1 - \lambda_2) / \lambda_1 \end{aligned} \right\}. \quad (4)$$

Additional features, i.e., the anisotropy  $A_\lambda$ , the planarity  $P_\lambda$ , the sphericity  $S_\lambda$ , and the linearity  $L_\lambda$ , are then obtained using (4) to distinguish planar, edge, and linear characteristics. Table I lists the variations of the values of the features over planar and nonplanar areas in terms of which ground, low vegetation, high vegetation, and buildings are classified [see Fig. 4(b)–(e)].

As illustrated in Fig. 4(a), the values of  $\lambda_3$  are relatively low in planar areas (e.g., building roofs) and relatively high in nonplanar areas (e.g., vegetation). The values of anisotropy  $A_\lambda$  are relatively low at the locations of vegetation, especially high vegetation [see Fig. 4(b)]. The values of the planarity  $P_\lambda$

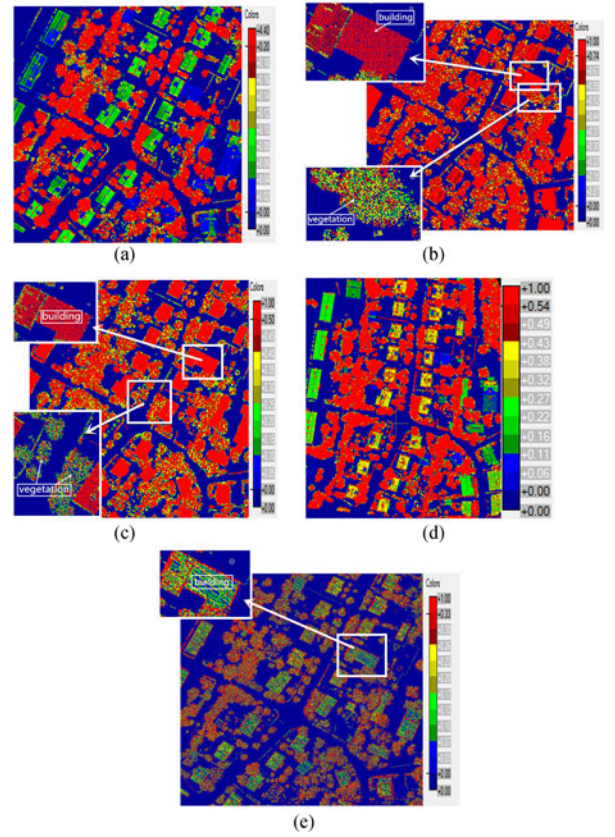


Fig. 4. Eigenvalue-based features. (a) Eigenvalue  $\lambda_3$ . (b) Anisotropy  $A_\lambda$ . (c) Planarity  $P_\lambda$ . (d) Sphericity  $S_\lambda$ . (e) Linearity  $L_\lambda$ .

are relatively high at planar areas, especially building roofs and roads [see Fig. 4(c)], whereas those of the sphericity  $S_\lambda$  are relatively high at the locations of vegetation and are relatively low in planar areas, especially building roofs [see Fig. 4(d)]. The values of the linearity  $L_\lambda$  are relatively high at the edges of building roofs [see Fig. 4(e)].

### B. Features of Images

1) *Image Segmentation Based on the Fractal Net Evolution Approach (FNEA)-Based Algorithm*: The FNEA, which was proposed by Baatz and Schape [24] in 2000, is a widely used multiscale segmentation algorithm and is also the basis and core content of popular object-oriented image analysis techniques. Fundamentally, the FNEA is a pixel-based, bottom-to-top region-growing segmentation algorithm. The FNEA, which follows the principle of minimal heterogeneity, merges adjacent pixels with similar spectral information into a homogenous image object and gives the same meaning to all pixels that belong to the same object after segmentation. The FNEA algorithm has been applied as the core segmentation algorithm of the eCognition commercial remote sensing software [25] and produces relatively good results.

2) *Definition of the Spectral Features of Image Objects*: As a prerequisite for integrating optical images with LiDAR data, the registration of the two datasets poses great challenges and is attracting increasingly more attention, e.g., [26], in which an

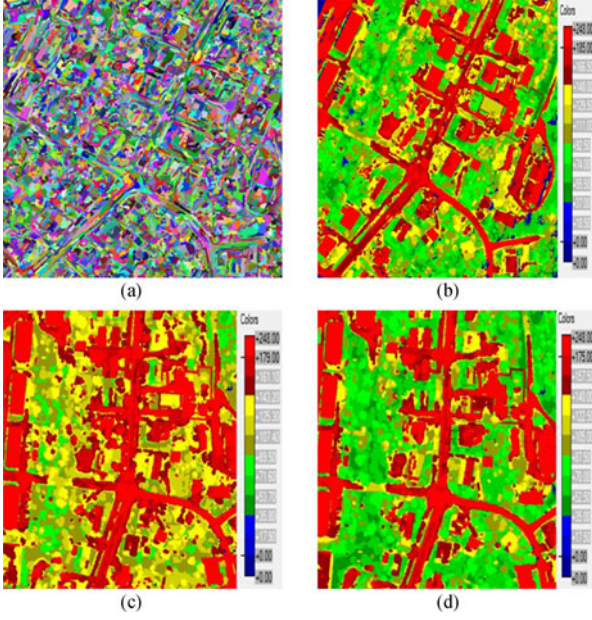


Fig. 5. Mean values of an image object in the RGB bands. (a) Image objects. (b) Mean value  $\mu_R$ . (c) Mean value  $\mu_G$ . (d) Mean value  $\mu_B$ .

automatic registration of optical images and LiDAR data captured by mini-UAVs was implemented in a novel manner via a fine registration of a LiDAR point cloud and a point cloud generated from sequent images. In this paper, direct georeferencing data contained in the orthoimage are used to register the optical images and LiDAR data because the registration accuracy satisfies the requirements of our proposed method. The extraction of the image classification features is mainly based on the image object. The rectified image is segmented using the FNEA algorithm [see Fig. 5(a)]. The features of the image object are extracted based on the segmentation results. The extracted spectral features of the image object [25] mainly include the mean value, brightness, and ratio.

*a) Mean value:* Based on the red-green-blue (RGB) values of all the pixels ( $n$  pixels) that form an image object, the mean values in the R, G and B bands are calculated by

$$\left. \begin{aligned} \mu_R &= \frac{1}{n} \sum_{i=1}^n R_i \\ \mu_G &= \frac{1}{n} \sum_{i=1}^n G_i \\ \mu_B &= \frac{1}{n} \sum_{i=1}^n B_i \end{aligned} \right\}. \quad (5)$$

In terms of the mean values in the R, G, and B bands, every sample point mapped to this image object is color-coded from blue to red [see Fig. 5(b)–(d)]. Fig. 5(b)–(d) shows that vegetation and nonvegetation areas are clearly distinguishable based on the mean values in the R, G, and B wavebands of the image objects.

*b) Brightness:* Brightness is the sum of the mean values of the layers that contain spectral information (e.g., the R, G, and B bands) divided by the number of layers of an image object, which is computed by (6). Similar to the illustration of mean values, every sample point mapped to this image object is

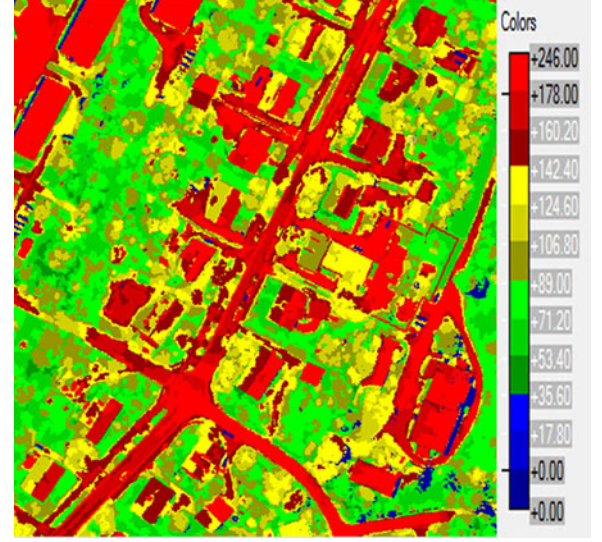


Fig. 6. Brightness of image objects.

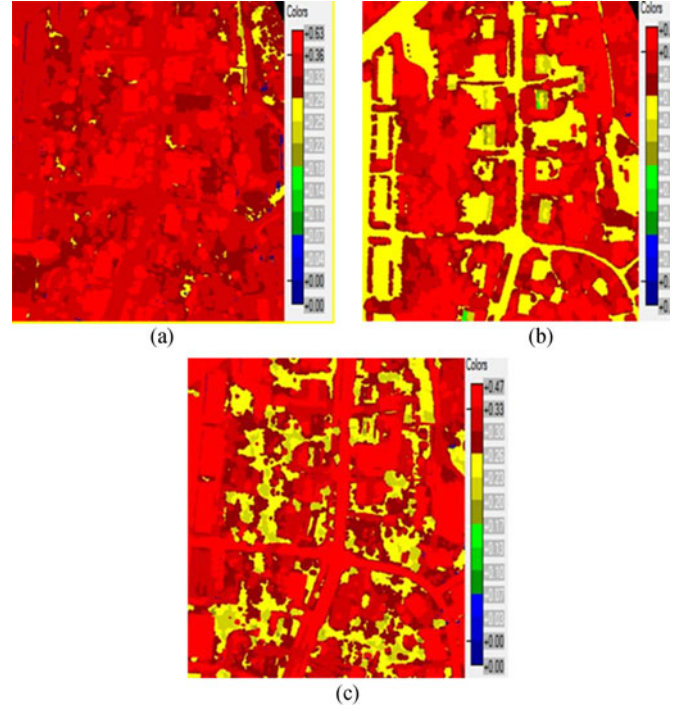


Fig. 7. Ratios of the R, G, and B bands of image objects. (a) Ratios of the R band, (b) ratios of the G band, and (c) ratios of the B band.

colored on a scale from blue to red according to the brightness (see Fig. 6). The vegetation and nonvegetation areas are also clearly distinguishable in Fig. 6:

$$B = \frac{1}{3}(\mu_R + \mu_G + \mu_B). \quad (6)$$

*c) Ratios:* The ratios of the R, G, and B bands of an image object are the mean values of the R, G, and B bands of the image object divided by the sum of the mean values of all the R, G, and B bands (7). Fig. 7(a)–(c) shows that vegetation and



nonvegetation areas are recognizable overall:

$$\left. \begin{aligned} \text{ratio}_R &= \frac{\mu_R}{\mu_R + \mu_G + \mu_B} \\ \text{ratio}_G &= \frac{\mu_G}{\mu_R + \mu_G + \mu_B} \\ \text{ratio}_B &= \frac{\mu_B}{\mu_R + \mu_G + \mu_B} \end{aligned} \right\}. \quad (7)$$

All the extracted features are combined into a feature vector. Thus, the 3-D point  $p$  is characterized by a 17-D feature vector  $F_v$ , where

$$F_v = \begin{bmatrix} N_H, \sigma^2, R_z, \lambda_1, \lambda_2, \lambda_3, A_\lambda, P_\lambda, S_\lambda, L_\lambda; \\ \mu_R, \mu_G, \mu_B, B, R_R, R_G, R_B \end{bmatrix}^T \quad (8)$$

which is used as an input of the BN network for the classification of point  $p$ .

### III. BN CLASSIFIER

A BN can be written as a two-tuples  $(\xi, \theta)$ , where  $\xi$  represents the network structure between the variables and  $\theta$  represents the probability distribution between the variables. The construction of a BN classifier consists of the discretization of extracted features, structure learning, parameter learning and classification.

#### A. Discretization of Extracted Features

To establish the BN network, the values of the extracted features, including geometric features of the point clouds and spectral features of the image objects as the evidence variables, need to be discretized based on information entropy [27]. The process is based on finding an “optimal” discretization by minimizing both the loss of entropy and the number of partitions. Here, the intervals are of varying lengths and frequencies. The entropy of a discrete random variable  $X$  is defined as follows:

$$H(X) = - \sum_{x \in X} p(x) \log_2 p(x). \quad (9)$$

Let each distinct value of a continuous random variable  $X$ , which denotes the geometric features of the point clouds or the spectral features of the corresponding image objects, be represented by a separate interval. Let the number of intervals be  $k$ ,  $1 \leq k \leq m$ , for  $m$  cases. Let the probability of  $X$  in each interval be  $p(i)$ ,  $1 \leq i \leq k$ . Let the entropy of the distribution of  $k$  discrete intervals be  $H(p_k)$ .

$H(p_k)$  is a concave function over  $k$ , where each decrease in the number of intervals is chosen to minimize the change in  $H(p_k)$ . Starting at the point of maximum entropy and maximum number of  $k$  intervals ( $2 \leq k \leq m$ ), the two adjacent intervals are merged, which results in the smallest change in  $H(p_k)$  that gives  $H(p_{k-1})$ . The smallest change in  $H(p_k)$  results from merging the two adjacent intervals with the smallest difference between  $p(i)$  and  $p(i+1)$ . This procedure is iterated until the stopping criteria, i.e., when the change in  $p_k$  becomes greater than the change in the entropy of  $p_k$  are satisfied.

After the discretization process, the values of the feature vector  $F_v$  are replaced with the sequence number of their corresponding intervals. For example, the range of the ratio of the B band  $\text{ratio}_B$  is divided into the following 18 intervals of

varying lengths:  $(-\infty, 0.175]$ ,  $(0.175, 0.185]$ ,  $(0.185, 0.205]$ ,  $(0.205, 0.215]$ ,  $(0.215, 0.245]$ ,  $(0.245, 0.255]$ ,  $(0.255, 0.265]$ ,  $(0.265, 0.285]$ ,  $(0.285, 0.295]$ ,  $(0.295, 0.305]$ ,  $(0.305, 0.315]$ ,  $(0.315, 0.335]$ ,  $(0.335, 0.345]$ ,  $(0.345, 0.355]$ ,  $(0.355, 0.365]$ ,  $(0.365, 0.375]$ ,  $(0.375, 0.4]$ ,  $(0.4, 0.425]$ . Then, each distinct value of the extracted feature  $\text{ratio}_B$  is represented by the sequence number (i.e., 1–18) of its corresponding intervals. The discretized feature vector is used to train the optimal BN structure.

#### B. Structure Learning

The structure learning of BNs is generally classified into the following two types: score-search-based algorithms and dependence-analysis-based algorithms. The K2 algorithm is a classic score-search-based algorithm but requires the determination of the sequence of nodes in advance. Based on the analysis of the BN structure, this paper adopts an improved mutual information-based BN structure learning algorithm.

Mutual information is a concept in mutual information theory [28] and describes the impact of the selection of the value of a certain variable on the selection of the value of another variable. As the value of the mutual information increases, the impact of the selection of the value of one variable on the selection of the value of the other variable increases. Therefore, the degree of dependence between random variables can be evaluated based on the mutual information between them. Specifically, the degree of dependence between the class variable, geometric feature variable, and spectral feature variable is represented by the mutual information in this paper.

The mutual information between random variables  $X$  and  $Y$  ( $X$  and  $Y$  can be class variables, geometric feature variables or spectral feature variables) is denoted as  $I(X; Y)$ , which is defined as follows:

$$I(X; Y) = H(X) - H(X|Y) \quad (10)$$

where  $H(X)$  and  $H(Y)$  represent the information entropy of  $X$  and  $Y$ , respectively, and  $H(X|Y)$  represents the information entropy of  $X$  when the conditions of  $Y$  are given.  $H(X)$  and  $H(X|Y)$  are defined as follows:

$$H(X) = - \sum_{i=1}^n P(X_i) \log(P(X_i)) \quad (11)$$

$$H(X|Y) = - \sum_{i=1}^n \sum_{j=1}^m P(X = x_i, Y = y_j) \cdot \log(P(X = x_i|Y = y_j)) \quad (12)$$

where  $n$  and  $m$  represent the number of states of  $X$  and  $Y$ , respectively. The mutual information between  $X$  and  $Y$  is symmetrical, i.e.,  $I(X; Y) = I(Y; X)$ .

The algorithm in this paper that is used for the structure learning calculates the mutual information between the nodes, obtains a maximum-weight spanning tree through calculations using the mutual information as the weights between the nodes, initializes the BN, and generates the initial edge set of the BN. The algorithm then performs a breadth-first search on the

generated maximum-weight spanning tree using the class nodes as the starting points and, while searching, determines the parent-child relationship between adjacent nodes using a Bayesian-based statistical scoring method to obtain the optimal sequence of nodes. Finally, the algorithm uses the obtained sequence of nodes as the input of the K2 algorithm to train the optimal BN structure. The detailed process for calculating the optimal sequence of nodes is as follows.

- 1) Calculate the maximum-weight spanning tree and initialize the edge set of the BN.
- 2) Define the linked list order =  $\{\}$ .
- 3) Add the class nodes into the linked list order and perform a breadth-first search on the maximum-weight spanning tree using the class nodes as the starting points (the current node ( $V_i$ ) has a sequence neighbor node ( $V_j$ ) in the linked list order).
- 4) Calculate the scores (if  $V_i$  is the parent node,  $score_i = CH(< v_i, v_j > | D)$ ; if  $V_j$  is the parent node,  $score_j = CH(< v_j, v_i > | D)$ ).
- 5) If  $score_i > score_j$ , assign the node after  $V_j$  in the linked list to  $V_j$  and return to Step 4 until  $score_i < score_j$  or the last node in the linked list order is reached, and then, insert  $V_i$  before  $V_j$ . If  $score_i < score_j$ , assign the node before  $V_j$  in the linked list to  $V_j$ , return to Step 4 until  $score_i > score_j$  or the first node in the linked list order is reached, and then, insert  $V_i$  after  $V_j$ .
- 6) Repeat Step 3 until all the nodes have been traversed.

### C. Parameter Learning

Once the BN structure has been established, the network parameter learning can be performed. To estimate the parameters of the BN, with  $\theta$  as a random variable, Bayesian estimation is used to determine the conditional probability distribution of each node variable. The prior knowledge of  $\theta$  is expressed as a prior probability distribution  $p(\theta)$  so that parameter learning is used to calculate the posterior probability distribution ( $p(\theta|D)$ ) of  $\theta$  given the observed data  $D$ .

Consider a set of variables  $X = \{X_1, X_2, \dots, X_n\}$  ( $n$  is the number of variables or nodes, which is 18 in this paper, including one class variable and 17 feature variables), and the variable  $X_j$  takes on values in the set  $S_j := \{x_j^{(1)}, x_j^{(2)}, \dots, x_j^{(k_j)}\}$ , i.e., the sequence number of the intervals after discretization in this paper. The probability distribution between the variables  $\theta$  is computed by the following equation:

$$\theta_{jil} = p\left(\{X_j = x_j^{(i)}\} | \{\Pi_j = \pi_j^{(l)}\}\right) \quad (13)$$

where  $\Pi_j$  denotes the set of parent variables of  $X_j$  ( $1 \leq j \leq n$ ). Suppose that the parent configurations  $\Pi_j$  take on values in the set  $\{\pi_j^{(q_1)}, \pi_j^{(q_2)}, \dots, \pi_j^{(q_j)}\}$  for  $j = 1, 2, \dots, n$ .  $k_j$  is the number of possible states that the variable  $X_j$  can take,  $q_j$  is the number of possible parent configurations for variable  $X_j$ , and  $\theta_{jil}$  is the conditional probability that variable  $X_j$  is in state  $i$ , given that the parent configuration is configuration  $l$ . The notation  $\pi_j^{(l)}$  denotes that the parents of variable  $X_j$  are in state  $l$ .

Considering that the Dirichlet distribution is the conjugate prior probability distribution of a multinomial distribution and that its conjugate properties can reduce the computational complexity and improve the efficiency, it is generally assumed that the prior distribution is a Dirichlet distribution. Hence

$$p(\theta|D) \propto \prod_{j=1}^n \prod_{i=1}^{q_j} \prod_{l=1}^{k_j} \theta_{jil}^{m_{jil} + \alpha_{jil} - 1} \quad (14)$$

which means that the posterior probability distribution  $p(\theta|D)$  is a Dirichlet distribution  $D[\alpha_{j11} + m_{j11}, \alpha_{j12} + m_{j12}, \dots, \alpha_{jk_j} + m_{jk_j}]$ , where  $m_{jil}$  ( $l = 1, 2, \dots, k_j$ ) represents the number of cases satisfying  $X_j = i$  and  $\Pi_j = \pi_j^{(l)}$  in  $D$  and  $\alpha_{j11}, \alpha_{j12}, \dots, \alpha_{jk_j}$  represent the hyperparameters of the prior distribution. For additional details, refer to [33].

### D. Classification

For classification problems, the nodes of a BN are generally classified into feature nodes and class nodes. The established BN model provides a complete joint probability distribution of all the variables. The classification process based on the established BN model calculates the posterior information of the class nodes in the BN when the exact values of a set of feature nodes are given. The classifier performs the classification based on the maximum posterior probability criterion. The classification processes for BNs are mainly classified into exact reasoning and approximate reasoning. The commonly used reasoning algorithms include the junction-tree algorithm [34], message transmission algorithm [35], and random sampling algorithm [36], [37]. Among these algorithms, the junction-tree algorithm has become the most widely used exact reasoning algorithm for BNs because it is efficient and can provide accurate reasoning results.

The junction-tree algorithm uses a graphical expression of a junction tree to express the joint probability, and its main idea is to transform a BN into a junction tree and then to transfer messages by defining message delivery on it to calculate the probability. The main process includes the following steps.

- 1) The original BN is converted into a junction tree.
- 2) The junction tree is initialized using the joint probability function of the BNs.
- 3) The junction tree satisfies global consistency; that is, steady state is reached through message delivery among the group nodes (each node in the junction tree is described as the group node, which consists of a set of random variables in the previous BNs and is also the largest fully connected subgraph of an undirected graph). In this case, we can calculate the probability distribution of any random variable of the original BN.

The feature vector  $F_v$  of a sample point is used as the evidence variables of the BN so that the posterior probability of the sample point belonging to each class (i.e., ground, vegetation, trees, or buildings) can be computed. The sample point is classified to the class that holds the highest posterior probability.

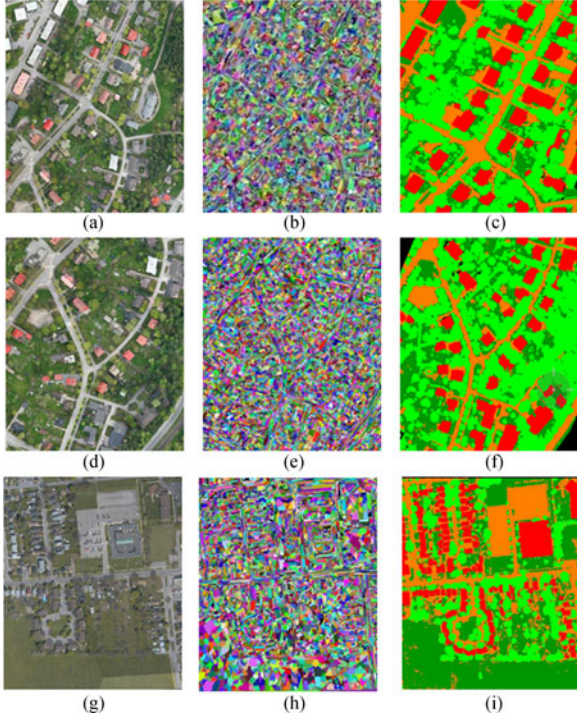


Fig. 8. Experimental data. Point types of ground, low vegetation, high vegetation, and buildings are colored in orange, dark green, light green, and red, respectively. (a) 0.2 m orthoimage (test area 1 in scene I), (b) segmented image (test area 1 in scene I, scale=20), (c) manual classification (test area 1 in scene I), (d) 0.2 m orthoimage (test area 2 in scene I), (e) segmented image (test area 2 in scene I, scale=20), (f) manual classification (test area 2 in scene I), (g) 0.2 m orthoimage (test area 3 in scene II), (h) segmented image (test area 3 in scene II, scale=20), and (i) manual classification (test area 3 in scene II).

#### IV. EXPERIMENTATION AND ANALYSIS

##### A. Experimental Data

To verify the proposed method, we performed both qualitative and quantitative evaluations on two scanned urban scenes. In scene I, airborne point cloud data of Jyväskylä, Finland<sup>1</sup> from 2011 with a point density of approximately 17 points/m<sup>2</sup> were used as experimental data. The orthoimages of the corresponding area have a resolution of 0.2 m. In scene II, airborne point cloud data of Niagara, Canada, with a point cloud density of approximately 3 points/m<sup>2</sup> were used as experimental data. The corresponding orthoimages have a resolution of 0.2 m. Fig. 8 shows the orthoimages, the segmented images, and the manual classification of the point clouds of test areas 1 and 2 in scene I and test area 3 in scene II. The original point clouds were classified into the following four types: ground points, low vegetation points, high vegetation points, and building points. Table II lists the statistics of the point clouds of all the training and test sets.

This paper used the recall, precision, accuracy, and F-score as the classifier evaluation indexes.

- 1) Recall represents a measure of completeness or quantity.
- 2) Precision denotes a measure of exactness or quality.

<sup>1</sup> Airborne point cloud data of Jyväskylä, Finland, are publicly available at [http://www.terrasolid.com/training/training\\_animations.php](http://www.terrasolid.com/training/training_animations.php) (accessed 11 June 2016).

TABLE II  
EXPERIMENTAL DATA

Class	Scene I			Scene II	
	Train set	Test area 1	Test area 2	Train set	Test area 3
Ground	127 643	265 785	183 031	86 131	266 561
Low vegetation	137 307	390 184	444 078	138 638	297 386
High vegetation	114 580	399 825	338 198	49 005	105 083
Buildings	72 690	138 834	106 202	62 422	97 141
Sum	452 220	1 194 628	1 071 513	396 196	766 171

TABLE III  
PERFORMANCE COMPARISON OF TWO IMAGE FEATURES

		Ground	Low vegetation	High vegetation	Buildings	Overall accuracy
Test area 1 in scene I	Pixel	72.21%	78.85%	89.84%	88.92%	87.58%
	Object	96.86%	93.28%	97.26%	90.10%	95.04%
Test area 3 in scene II	Pixel	87.99%	86.87%	80.36%	91.17%	86.95%
	Object	96.61%	90.91%	89.58%	91.56%	92.80%

- 3) Overall accuracy reflects the overall performance of the classifier on the test set.
- 4) F-score combines precision and recall with equal weights.
- 5) Let us assume that the original samples include positive samples and negative samples. Then, we have

$$\text{Precision} = \frac{TP}{TP + FP} \quad (15)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (16)$$

$$\text{Overall Accuracy} = \frac{TP + FN}{TP + FP + TN + FN} \quad (17)$$

$$F_{\text{-score}} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}} \quad (18)$$

where TP denotes the number of positive samples that are correctly determined as positive, FN represents the number of positive samples that are incorrectly classified as negative, FP denotes the number of negative samples that are incorrectly determined as positive, and TN is the number of negative samples that are correctly identified as negative.

##### B. Experimental Analysis

1) *Comparison of Image Features*: To determine the optimum image feature to be used in the proposed method, two image features, i.e., single pixels and image objects, were employed to classify test area 1 in scene I and test area 3 in scene II, respectively. Table III shows that the image object achieved much better classification accuracies on all types of points than did the single pixels. The incorrect classification of the ground points as low vegetation points was used as an example to visualize the performance difference. Fig. 10 shows the comparison results. The point cloud classification results using single pixels exhibit a significant “salt and pepper” phenomenon, and a large





Fig. 9. Segmentation of a rooftop. (a) Original image. (b) Segmented results.

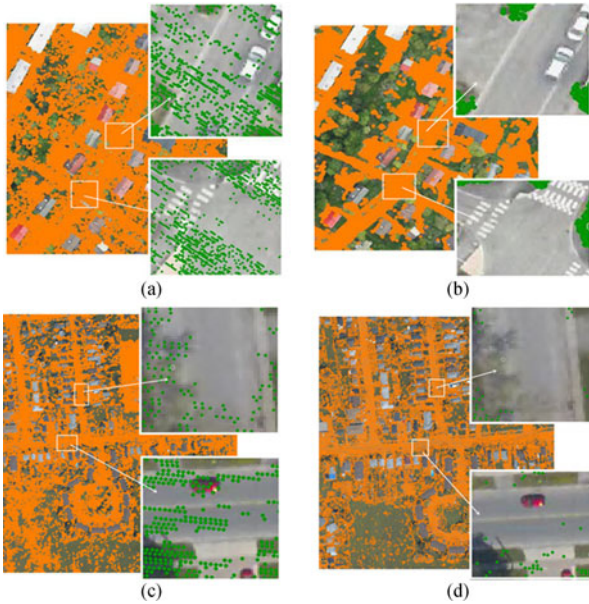


Fig. 10. Comparison of classification results using different image features. The points of ground and low vegetation are colored in orange and dark green, respectively. (a) Classification based on single pixel (test area 1 in scene I). (b) Classification based on image object (test area 1 in scene I). (c) Classification based on single pixels (test area 3 in scene II). (d) Classification based on image object (test area 3 in scene II).

number of ground points on the road were classified as low vegetation points. However, by using image objects, the classification accuracies were improved by 24.65% and 8.62%. As a result, image objects were selected by the proposed method.

The proposed method has a dependence on the image resolution and quality. If the image resolution is too low or the image quality is too poor, different objects may become mixed at the borders. As a result, deviations may be present in the image features calculated from the extracted image objects. However, the 0.2-m-resolution orthoimages used in this manuscript are commonly used products in practical applications. Moreover, the proposed approach can achieve a good result on such images, which shows good applicability.

Considering Fig. 9(a), it is normally quite difficult to ensure a uniformly colored texture of a rooftop. However, a uniform texture likely presents in small patches that can be extracted using a region-growing segmentation algorithm [see Fig. 9(b)]. Therefore, in this paper, small patches, instead of the whole rooftop, were regarded as image objects and were thus used to compute the spectral features, i.e., the mean value, brightness, and ratio.

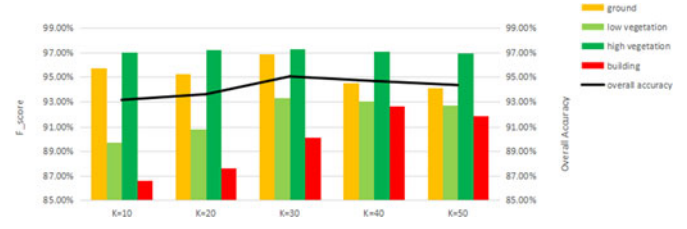


Fig. 11. F-scores of different neighborhood sizes (test area 1 in scene I).

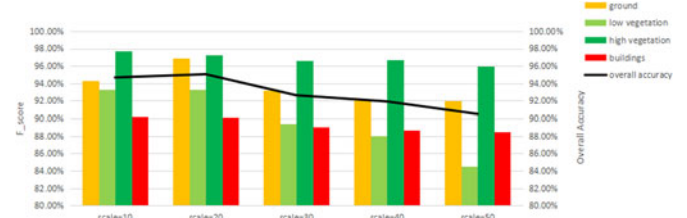


Fig. 12. F-scores of different segmentation scales (test area 1 in scene I).

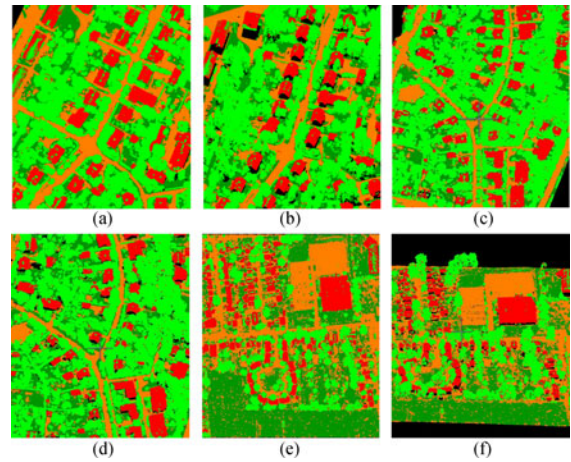


Fig. 13. Classification results ( $K=30$ ,  $\text{scale}=20$ ). The points of ground, low vegetation, high vegetation, and buildings are colored in orange, dark green, light green, and red, respectively. (a) Top view (test area 1 in scene I). (b) Oblique view (test area 1 in scene I). (c) Top view (test area 2 in scene I). (d) Oblique view (test area 2 in scene I). (e) Top view (test area 3 in scene II). (f) Oblique view (test area 3 in scene II).

As a result, even solar-panel- and water-storage-like objects can be present on the rooftop surface, and the computation of the spectral features will not be degraded by their presence.

2) *Optimization of Classification Parameters:* To optimize the selection of classification parameters, we calculated the point cloud classification features using multiple neighborhood sizes (e.g.,  $K=10, 20, 30, 40$ , and  $50$ ) and multiple segmentation scale sizes (e.g.,  $\text{scale}=10, 20, 30, 40$ , and  $50$ ).

Using test area 1 in scene I as an example, the point cloud was classified using the proposed BN classifier based on multiple neighborhood sizes and multiple segmentation scale sizes. Fig. 11 shows that the neighborhood size has a significant influence on the classification accuracies of all types, except the high vegetation. The influence on buildings is most significant.

Fig. 12 illustrates that the segmentation scale has a relatively large influence on the classification accuracies of the ground and low vegetation points.

TABLE IV  
CONFUSION MATRIX OF CLASSIFICATION (TEST AREA 1 IN SCENE I,  $K=30$ ,  
SCALE=20, OVERALL ACCURACY IS 95.04%)

	Ground	Low vegetation	High vegetation	Buildings	Precision
Ground	257 430	7918	414	23	96.86%
Low vegetation	15 160	363 945	10 901	178	93.28%
High vegetation	138	6326	388 859	4502	97.26%
Buildings	290	104	13 353	125 087	90.10%
Recall	94.29%	96.21%	94.03%	96.38%	

TABLE V  
CONFUSION MATRIX OF CLASSIFICATION (TEST AREA 2 IN SCENE I,  $K=30$ ,  
SCALE=20, OVERALL ACCURACY IS 94.48%)

	Ground	Low vegetation	High vegetation	Buildings	Precision
Ground	169 193	12 194	1485	159	92.44%
Low vegetation	10 966	413 831	18 013	268	93.40%
High vegetation	199	722	328 772	8505	97.21%
Buildings	185	116	6260	99 645	93.82%
Recall	93.71%	96.95%	92.73%	91.77%	

TABLE VI  
CONFUSION MATRIX OF CLASSIFICATION (TEST AREA 3 IN SCENE II,  $K=30$ ,  
SCALE=20, OVERALL ACCURACY IS 92.80%)

	Ground	Low vegetation	High vegetation	Buildings	Precision
Ground	257 534	8331	593	103	96.61%
Low vegetation	25 739	270 361	1108	178	90.91%
High vegetation	552	776	94 134	9621	89.58%
Buildings	53	0	8145	88 943	91.56%
Recall	90.72%	96.74%	90.53%	89.98%	

As shown in Figs. 11 and 12, when the neighborhood size and segmentation scale are set as 30 and 20, respectively, the classification achieves the highest F-scores. Therefore, these two values were selected as the optimum values of the two parameters.

With the optimum classification parameters and image objects, test areas 1 and 2 in scene I and test area 3 in scene II were classified using the proposed BN classifier (see Fig. 13). Tables IV–VI list the confusion matrices of the classifications. The presented BN classifier achieved high accuracies of over 90%.

3) *Comparison With Other Classifiers:* To further evaluate the performance of our proposed BN classifier in point cloud classification, commonly used classifiers (e.g., SVM [4]–[6], Dtree, RFs [7]–[9], and AdaBoost [10]) were compared to the BN classifier using the three datasets. Tables VII–IX list the results, which show that the BN classifier achieved the highest overall accuracies. Moreover, the BN classifier is quite advantageous for the classification of ground and low vegetation points. Fig. 14(a) and (b) shows that Dtree and RFs falsely classified the low vegetation points as ground points, whereas the BN classifier obtained correct classification results.

TABLE VII  
PERFORMANCE COMPARISON BETWEEN DIFFERENT CLASSIFIERS  
(TEST AREA 1 IN SCENE I)

Classifier	Ground	Low vegetation	High vegetation	Buildings	Overall accuracy
BN	96.86%	93.28%	97.26%	90.10%	95.04%
Dtree	87.39%	89.20%	95.50%	91.78%	91.20%
AdaBoost	94.94%	93.55%	97.25%	91.96%	94.91%
RF	92.83%	91.97%	97.53%	92.39%	94.04%
SVM	81.89%	81.71%	86.63%	74.65%	82.58%

TABLE VIII  
PERFORMANCE COMPARISON BETWEEN DIFFERENT CLASSIFIERS  
(TEST AREA 2 IN SCENE I)

Classifier	Ground	Low vegetation	High vegetation	Buildings	Overall accuracy
BN	92.44%	93.40%	97.21%	93.82%	94.48%
Dtree	84.65%	89.49%	96.52%	95.13%	91.39%
AdaBoost	92.43%	88.65%	97.94%	95.76%	92.90%
RF	88.78%	92.74%	97.95%	95.74%	93.97%
SVM	70.59%	80.34%	82.43%	65.58%	77.78%

TABLE IX  
PERFORMANCE COMPARISON BETWEEN DIFFERENT CLASSIFIERS  
(TEST AREA 3 IN SCENE II)

Classifier	Ground	Low vegetation	High vegetation	Buildings	Overall accuracy
BN	96.61%	90.91%	89.58%	91.56%	92.80%
Dtree	87.99%	86.79%	80.36%	91.17%	86.95%
AdaBoost	92.11%	90.63%	83.35%	94.66%	90.69%
RF	92.13%	91.00%	90.06%	94.49%	91.71%
SVM	74.10%	77.46%	70.55%	81.27%	75.83%



Fig. 14. Performance comparison between Dtree, RF, and the proposed method. The points of the ground are colored in orange. (a) Dtree classification. (b) RF classification. (c) BN classification corresponding to the result of Dtree. (d) BN classification corresponding to the result of RF.



## V. CONCLUSION

Using a classification strategy that integrates multisource data, this paper extracted the spectral information of aerial images using an object-oriented method addressing the “salt and pepper” phenomenon that occurs when image information is extracted based on single pixels. In addition, this paper incorporated the spectral information of image objects with the point cloud classification process using a BN model for the purpose of improving the robustness of point cloud classification.

The proposed algorithm was implemented using three airborne LiDAR datasets. The point clouds were classified using multiple neighborhood sizes and multiple segmentation scale sizes. The experimental results show that the neighborhood size has little influence on the classification accuracy of the high vegetation and that the segmentation scale has a relatively high influence on the classification accuracies of the ground and low vegetation points. The results also indicate that, when using image objects, the classification accuracies were improved by 24.65% and 8.62%, respectively, compared with the results using single pixels. The classification results of the three datasets achieved high accuracies of over 90%. Moreover, the presented BN classifier achieved the highest overall accuracies and demonstrated its advantage for the classification of ground and low vegetation points compared to commonly used classifiers (e.g., SVM, Dtree, RFs and AdaBoost).

Compared to the performances on ground points, low vegetation points and high vegetation points, the classification of building points by the proposed method achieves a less satisfactory accuracy, which can be attributed to the noise present at the edges of the roofs and ridge lines as well as the image-to-scan registration error. Thus, future research will focus on the improvement of the classification accuracy on building points.

## REFERENCES

- [1] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik, “Building detection by fusion of airborne laser scanner data and multi-spectral images: Performance evaluation and sensitivity analysis,” *ISPRS J. Photogramm. Remote Sens.*, vol. 62, no. 2, pp. 135–149, 2007.
- [2] F. Lafarge and C. Mallet, “Creating large-scale city models from 3D-point clouds: A robust approach with hybridation,” *Int. J. Comput. Vis.*, vol. 99, no. 1, pp. 69–85, 2012.
- [3] K. Khoshelham, C. Nardinocchi, E. Frontoni, A. Mancini, and P. Zingaretti, “Performance evaluation of automated approaches to building detection in multi-source aerial data,” *ISPRS J. Photogramm. Remote Sens.*, vol. 65, no. 1, pp. 123–133, 2010.
- [4] J. García-Gutiérrez, D. Mateos-García, M. Garcia, and J. C. Riquelme-Santos, “An evolutionary-weighted majority voting and support vector machines applied to contextual classification of LiDAR and imagery data fusion,” *Neurocomputing*, vol. 163, pp. 17–24, 2015.
- [5] Z. Q. Li *et al.*, “A three-step approach for TLS point cloud classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 9, pp. 5412–5424, Sep. 2016.
- [6] Y. Cao, H. Wei, H. J. Zhao, and N. Li, “An effective approach for land-cover classification from airborne lidar fused with co-registered data,” *Int. J. Remote Sens.*, vol. 33, no. 18, pp. 5927–5953, 2012.
- [7] L. Guo, N. Chehata, C. Mallet, and S. Boukir, “Relevance of airborne lidar and multispectral image data for urban scene classification using random forests,” *ISPRS J. Photogramm. Remote Sens.*, vol. 66, no. 1, pp. 56–66, 2011.
- [8] N. Chehata, G. Li, and C. Mallet, “Airborne lidar feature selection for urban classification using random forests,” *Geomatics Inf. Sci. Wuhan Univ.*, vol. 38, pp. 207–212, 2009.
- [9] M. Gerke and J. Xiao, “Fusion of airborne laser scanning point clouds and images for supervised and unsupervised scene classification,” *ISPRS J. Photogramm. Remote Sens.*, vol. 87, no. 1, pp. 78–92, 2014.
- [10] Z. Wang *et al.*, “A multiscale and hierarchical feature extraction method for terrestrial laser scanning point cloud classification,” *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2409–2425, May 2015.
- [11] V. Walter, “Object-based classification of remote sensing data for change detection,” *ISPRS J. Photogramm. Remote Sens.*, vol. 58, no. 3, pp. 225–238, 2004.
- [12] M. Bartels and H. Wei, “Maximum likelihood classification of LIDAR data incorporating multiple co-registered bands,” in *Proc. 4th Int. Workshop Pattern Recognit. Remote Sens./18th Int. Conf. Pattern Recognit.*, vol. 20, 2006, pp. 17–20.
- [13] J. Niemeyer, F. Rottensteiner, and U. Soergel, “Conditional random fields for lidar point cloud classification in complex urban areas,” *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. 1, nos. 1/3, pp. 263–268, 2012.
- [14] E. H. Lim and D. Suter, “Multi-scale conditional random fields for over-segmented irregular 3D point clouds classification,” in *Proc. Comput. Vis. Pattern Recognit. Workshop*, 2008, pp. 1–7.
- [15] J. Cheng, R. Greiner, J. Kelly, D. Bell, and W. Liu, “Learning Bayesian networks from data: An information-theory based approach,” *Artif. Intell.*, vol. 137, nos. 1/2, pp. 43–90, 2002.
- [16] A. Brunn and U. Weidner, “Hierarchical Bayesian nets for building extraction using dense digital surface models,” *ISPRS J. Photogramm. Remote Sens.*, vol. 53, no. 5, pp. 296–307, 1998.
- [17] G. F. Cooper and E. Herskovits, “A Bayesian method for the induction of probabilistic networks from data,” *Mach. Learn.*, vol. 9, no. 4, pp. 309–347, 1992.
- [18] B. Guo, X. Huang, F. Zhang, and G. Sohn, “Classification of airborne laser scanning data using JointBoost,” *ISPRS J. Photogramm. Remote Sens.*, vol. 100, pp. 71–83, 2015.
- [19] M. Weinmann, B. Jutzi, and C. Mallet, “Feature relevance assessment for the semantic interpretation of 3D point cloud data,” *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. II-5/W2, pp. 313–318, 2013.
- [20] S. Filin and N. Pfeifer, “Neighborhood systems for airborne laser data,” *Photogramm. Eng. Remote Sens.*, vol. 71, no. 6, pp. 743–755, 2005.
- [21] M. Weinmann, B. Jutzi, S. Hinz, and C. Mallet, “Semantic point cloud interpretation based on optimal neighborhoods, relevant features and efficient classifiers,” *ISPRS J. Photogramm. Remote Sens.*, vol. 105, pp. 286–304, 2015.
- [22] J. Shan and S. Aparajithan, “Urban DEM generation from raw LiDAR data,” *Photogramm. Eng. Remote Sens.*, vol. 71, no. 2, pp. 217–226, 2005.
- [23] Z. Zhang and G. Xu, *Epipolar Geometry in Stereo, Motion and Object Recognition: A Unified Approach*. Norwell, MA, USA: Kluwer, 1996.
- [24] M. Baatz and A. Schäpe, “Multiresolution segmentation: An optimization approach for high quality multi-scale image segmentation,” in *Angewandte Geographische Informationsverarbeitung XII*, J. Strobl, T. Blaschke, and G. Griesebner, Eds. Heidelberg, Germany: Wichmann, 2000, pp. 12–23.
- [25] Trimble, *eCognition Developer 8.7.2 Reference Book*. München, Germany: Trimble Germany GmbH, 2012, pp. 34–40.
- [26] B. Yang and C. Chen, “Automatic registration of UAV-borne sequent images and LiDAR data,” *ISPRS J. Photogramm. Remote Sens.*, vol. 101, pp. 262–274, 2015.
- [27] E. J. Clarke and B. A. Barton, “Entropy and MDL discretization of continuous variables for Bayesian belief networks,” *J. Intell. Inf. Syst.*, vol. 15, no. 15, pp. 61–92, 2000.
- [28] L. M. D. Campos, “A scoring function for learning Bayesian networks based on mutual information and conditional independence tests,” *J. Mach. Learn. Res.*, vol. 7, pp. 2149–2187, 2006.
- [29] T. Koski and J. Noble, *Bayesian Networks: An Introduction*. Hoboken, NJ, USA: Wiley, 2009.
- [30] J. Dash, E. Steinle, R. P. Singh, and H. P. Bahr, “Automatic building extraction from laser scanning data: An input tool for disaster management,” *Adv. Space Res.*, vol. 33, no. 3, pp. 317–322, 2004.
- [31] S. Xu, E. S. Oude, and G. Vosselman, “Entities and features for classification of airborne laser scanning data in urban area,” *Anal. Chim. Acta*, vol. 1-4, pp. 257–262, 2012.
- [32] Y. Chen, W. Su, J. Li, and Z. Sun, “Hierarchical object oriented classification using very high resolution imagery and LIDAR data over urban areas,” *Adv. Space Res.*, vol. 43, no. 7, pp. 1101–1110, 2009.
- [33] A. S. Antonarakis, K. S. Richards, and J. Brasington, “Object-based land cover classification using airborne LiDAR,” *Remote Sens. Environ.*, vol. 112, no. 6, pp. 2988–2998, 2008.



- [34] S. L. Lauritzen and D. J. Spiegelhalter, "Local computations with probabilities on graphical structures and their application to expert systems," *J. Roy. Stat. Soc. Series B: Stat. Methodol.*, vol. 50, pp. 157–224, 1988.
- [35] J. Pearl, "Fusion, propagation, and structuring in belief networks," *Artif. Intell.*, vol. 29, no. 3, pp. 241–288, 1986.
- [36] P. Dagum, R. Karp, M. Luby, and S. Ross, "An optimal algorithm for Monte Carlo estimation," *SIAM J. Comput.*, vol. 29, pp. 1484–1496, 2000.
- [37] R. M. Chavez and G. F. Cooper, "A randomized approximation algorithm for probabilistic inference on Bayesian belief networks," *Networks*, vol. 20, no. 5, pp. 661–685, 1990.
- [38] H. Y. Guan, Z. Ji, L. Zhong, J. Li, and Q. Ren, "Partially supervised hierarchical classification for urban features from lidar data with aerial imagery," *Int. J. Remote Sens.*, vol. 34, no. 1, pp. 190–210, 2013.
- [39] M. Weinmann, S. Urban, S. Hinz, B. Jutzi, and C. Mallet, "Distinctive 2D and 3D features for automated large-scale scene analysis in urban areas," *Comput. Graph.*, vol. 49, pp. 47–57, 2015.
- [40] W. Y. Yan, A. Shaker, and N. El-Ashmawy, "Urban land cover classification using airborne LiDAR data: A review," *Remote Sens. Environ.*, vol. 158, pp. 295–310, 2015.
- [41] C. H. Lin, J. Y. Chen, P. L. Su, and C. H. Chen, "Eigen-feature analysis of weighted covariance matrices for LiDAR point cloud classification," *ISPRS J. Photogramm. Remote Sens.*, vol. 94, no. 94, pp. 70–79, 2014.
- [42] Z. Zhang *et al.*, "A multilevel point-cluster-based discriminative feature for ALS point cloud classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3309–3321, Jun. 2016.
- [43] A. Zarea and A. Mohammadzadeh, "A novel building and tree detection method from LiDAR data and aerial images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 5, pp. 1864–1875, May 2016.
- [44] P. Babahajiani, L. Fan, J. Kamarainen, and M. Gabbouj, "Automated super-voxel based features classification of urban environments by integrating 3D point cloud and image content," in *Proc. IEEE Int. Conf. Signal Image Process. Appl.*, 2015, pp. 372–377.
- [45] X. Zhu and T. Toutin, "Object-oriented interpretation and classification with airborne LiDAR 3D data," in *Proc. Int. Symp. Image Data Fusion*, 2011, pp. 1–4.
- [46] M. Carlberg, P. Gao, G. Chen, and A. Zakhori, "Classifying urban landscape in aerial LiDAR using 3D shape analysis," in *Proc. IEEE Int. Conf. Image Process.*, 2009, pp. 1701–1704.
- [47] K. Khoshelham, C. Nardinocchi, E. Frontoni, A. Mancini, and P. Zingaretti, "Performance evaluation of automated approaches to building detection in multi-source aerial data," *ISPRS J. Photogramm. Remote Sens.*, vol. 65, no. 1, pp. 123–133, 2010.



**Zhizhong Kang** received the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2004.

He was a Postdoctoral Researcher at Delft University of Technology, The Netherlands, from 2006 to 2008. He is currently an Associate Professor and the Head of the Department of Remote Sensing and Geo-Information Engineering, China University of Geosciences, Beijing, China. He has published more than 60 refereed journal and conference publications.

His research interests include digital photogrammetry and computer vision including indoor mapping and modeling, processing of LiDAR data, lunar structure recognition and extraction.

Dr. Kang has served as the Chair of ISPRS WG IV/5: Indoor/Outdoor seamless modelling, LBS, mobility and Member of Committee on Photogrammetry and Remote Sensing, Chinese Society for Geodesy Photogrammetry and Cartography. He received 2015 ERDAS Award for Best Scientific Paper in Remote Sensing by ASPRS.



**Juntao Yang** was born in Shouguang, China, in 1991. He received the B.S. degree in geomatics from the College of Information Science and Engineering, Shandong Agricultural University, Tai'an, China, in 2014, and is currently working toward the M.S. degree in photogrammetry and remote sensing at China University of Geosciences, Beijing, China.

Since 2014, he has been involved in the research of the classification of point cloud data. His research interests include remote sensing image processing, the classification of point cloud data, machine learning, and lunar structure recognition and extraction.



**Ruofei Zhong** received the Ph.D. degree in geoinformatics from the Chinese Academy of Science's Institute of Remote Sensing Applications, Beijing, China, in 2005.

He is currently a Professor in Beijing Advanced Innovation Center for Imaging Technology and Key Lab of 3D Information Acquisition and Application, College of Resource Environment and Tourism, Capital Normal University, Beijing. His research interests include LiDAR data processing and data collection system with laser scanning.