International Journal of Remote Sensing
Vol. 34, No. 1, 10 January 2013, 190–210

Taylor & Francis
Taylor & Francis Group

Partially supervised hierarchical classification for urban features from lidar data with aerial imagery

Haiyan Guan[a], Zheng Ji[b]*, Liang Zhong[c], Jonathan Li[a], and Que Ren[a]

[a]*Department of Geography and Environmental Management, University of Waterloo, Waterloo, Ontario, Canada N2L 3G1;* [b]*School of Remote Sensing Information and Engineering, Wuhan University, Wuhan 430079, China;* [c]*Changjiang Spatial Information Technology Engineering Company, Changjiang Institute of Survey Planning Design and Research, Wuhan 430010, China*

(*Received 10 August 2011; accepted 17 February 2012*)

Although spatial and spectral resolutions of remotely sensed data have been improved, the usage of multispectral imagery is not sufficient for urban feature classification. This article addresses the problem of automated classification by integrating airborne lidar range data and aerial imagery. In this study, the classification procedure is divided into three phases. We first use the lidar range data to obtain the coarse lidar-based classification results, by which a lidar-driven labelled image and a lidar-driven high-rise object mask are acquired in this phase. Then, at the image-based classification level, we train samples based on the lidar-driven labelled image and conduct maximum likelihood classification experience with the lidar-driven normalized digital surface model as a high-rise object mask. Finally, we propose a knowledge-based cross-validation (KBCV) for misclassification between the lidar-based classification results and the image-based classification results. Experimental results are presented to demonstrate the benefits of the training sample selection of the lidar-driven labelled image, using the lidar-driven high-rise object mask, and the greater classification accuracy of the KBCV.

1. Introduction

Land-use classification is one of the fundamental tasks in the analysis and modelling of spatial data acquired from lidar systems. Classified ground points are used to generate digital terrain models (DTMs) that provide basic geographical data for a wide range of applications, including topographical mapping, engineering surveying, and environmental planning, whereas other lidar points, classified as building or vegetation, assist in DTM accuracy enhancement, three-dimensional (3D) digital city reconstruction, building model update in the geographical information system (GIS) database, and urban green space study. Therefore, many previous studies have shown the potential of lidar data for urban feature identification, classification, and reconstruction (Miliaresis and Kokkas 2007; Huang et al. 2008; Wang et al. 2009).

However, point data are unstructured, irregular and short of spectral information, leading to classification confusion between man-made and natural objects. On the other hand, it is difficult to directly obtain land-use information only from remotely sensed data, due to the complexity of landscapes, spectrally identical objects, and abundance of spatial and

*Corresponding author. Email: jz07@whu.edu.cn

ISSN 0143-1161 print/ISSN 1366-5901 online

http://dx.doi.org/10.1080/01431161.2012.712228
http://www.tandfonline.com

spectral information. Being aware of the limitations of a single data source, the majority of researchers have focused on fusing multi-source data, such as lidar data with QuickBird imagery (Ke, Quackenbush, and Im 2010), lidar data with IKONOS satellite imagery (Shan and Lee 2003), and lidar data with aerial imagery (Huang et al. 2008; Carlberg et al. 2009; Goodwin et al. 2009; Khoshelham et al. 2010). Because the latest lidar systems carry both laser scanners and high-performance charge-coupled device cameras, integrating lidar data with the aerial imagery is a preferred means for object identification and classification.

Researchers use many classification algorithms to convert 3D lidar data into two-dimensional (2D) range images for integration with aerial images in classification tasks. Once the data are in 2D form, conventional image-based classification algorithms can be applied to achieve urban feature classification. A host of classification methods are broadly divided into two categories: unsupervised and supervised classifiers. Unsupervised classification, an analytical procedure, is based on clustering, which separates the image data into a series of spectral classes. Then, all pixels of interest are labelled as one of these spectral classes. But the labels are unrelated to ground types and required to explicitly attach to particular ground types. Unlike prior knowledge from training samples in the supervised method, the unsupervised classification is a posterior identification; in other words, it is a segmentation of data without any known information, and then feature information is used to assign ground types to the segments established by a variety of clustering algorithms. Due to time demands for the clustering, unsupervised classification is a computationally time-consuming procedure. Thus, it is often carried out with small imagery data.

In contrast to the unsupervised classification algorithm, the supervised method requires training samples to obtain class attributes. The supervised classifiers range from traditional maximum likelihood classification (MLC) (Mesev, Gorte, and Longley 2001; Khoshelham et al. 2010) to machine learning algorithms (e.g. support vector machine (Dalponte, Bruzzone, and Gianelle 2008; Koetz et al. 2008; Jones, Coops, and Sharma 2010; Guan et al. 2011; Trinder and Salah 2011) and random forest (Pal 2005; Gislason, Benediktsson, and Sveinsson 2006; Smith 2010; Guo et al. 2011)). Both of them rely on discrimination functions using representative image attributes acquired from training samples. Such discrimination functions are then implemented in the whole image to obtain the desired ground types by one or more particular classifiers (Richards and Jia 2006). In the supervised manner, a prior knowledge about the ground types to be classified is required through a combination of field work, image interpretation, map analysis, and personal experience (Hodgson et al. 2003). However, supervised algorithms might fail if training samples are not representative of the true distribution of classes to be labelled, owing to a lack of available information and acquisition costs of ground-truth information. Meanwhile, with an increase in sensor performance such as the spatial extent, training a large amount of extensive remotely sensed data is hard and time-consuming (Tuia, Pasolli, and Emery 2011). A hybrid supervised/unsupervised classification algorithm that applies the MLC after the implementation of clustering has been studied to solve the multi-modality problem; however, researchers have seldom used this method due to the complex and time-consuming procedure (South, Oi, and Lusch 2004).

To overcome classification noises caused by individual classification of each pixel as a certain group, object-based classification (OBC) has been widely studied for high-resolution multispectral images in urban environments (Lu, Hetrick, and Moran 2010). According to Blaschke (2010), over 800 articles are related to object-based image analysis. However, critical problems have emerged along with the development of a variety of OBC techniques. An analyst must have sufficient prior knowledge of objects of interest,

and then select optimum segmentation parameters by building up a hierarchical image classification network for classification. Thus, classification accuracy is influenced by the quality of segmentation (Repake et al. 2004). Currently, the selection of segmentation parameters is based on analysts' empirical studies, which lead to uncertain classification results. In addition, the complexity of urban landscapes and the abundance of intensity and spectral information can encumber the interpretation and segmentation of aerial imagery. However, the statistical MLC, one of the most widely used supervised classification algorithms, is easy to implement and analyse with its effective computation (He and Wang 1992). It is recognized as a stable and robust classification method, and the most accurate classification scheme in standard digital image processing. In MLC, two parameters, the mean vector and the covariance matrix, are used to characterize each class of interest; therefore, an unbiased and efficient estimate of these parameters would largely depend on the determination of the sample sizes and schemes.

According to Jeon and Landgrebe (2002), sampling or training all of the classes in a given data set is labour- and time-intensive by collecting ground truth. In general, three methods are applied to collect of training data, including the collection of *in situ* information, on-screen selection of polygon training data, and on-screen seeding of training data (Jensen 2005). In most cases, the analyst uses a 'rubber band' tool to select polygon areas of interest (AOIs) by viewing imagery on the computer screen. However, this selection of AOIs requires the analyst's empirical expertise and proficiency. Volpi, Tuia, and Kanevski (2010) also mentioned that the difficulty of manually selecting samples was to keep the number of labelled training patterns small for covering intra-class variance of the image. Unlike on-screen selection of polygon training data, the on-screen seeding of training data is similar to a region-growing method. First, the analyst seeds a specific point in the image using a cursor, and then the seed expands to find more pixels with similar spectral characteristics by a given criteria. However, for high-resolution aerial imagery, spectral variation in one ground object is dramatic, which leads to generation of a number of homogenous regions. According to Tso and Mather (2001), the use of large region samples is not recommended since pixels or points in a group are not mutually independent. In other words, each region describes a part of the spectral characteristics of the objects so that statistical information of the objects from the training data might not be comprehensively representative for the following classification procedure. Congalton, Oderwald, and Mead (1983) suggested that the number of pixels in such a group should not be larger than 10.

However, people usually are interested in one or a small number of classes (Foody et al. 2006). For example, urban land-use classification focuses on four main urban features of interest: grass, high vegetation, building, and ground (Secord and Zakhor 2007; Huang et al. 2008; Guo et al. 2011). Each class can be distinguished from a given data set based on its available prior information derived from known training samples. Jeon and Landgrebe (2002) call this method partially supervised classification. Therefore, we design a partially supervised classifier as the main classifier followed by a knowledge-based cross-validation (KBCV). Unlike conventional classification methods of fusing lidar data with imagery, we first extract four coarse classes of interest from lidar data to form a lidar-driven labelled image and a lidar-driven high-rise object mask. According to the locations of each class in the lidar-driven labelled image, corresponding training samples in the calibrated aerial images can be randomly and automatically selected to obtain statistic information for performing statistical MLC. Meanwhile, the use of a lidar-driven high-rise object mask simplifies the image-based classification processing. Next, a KBCV strategy is developed

to analyse the two classification results – the lidar-based and image-based classification results – and to adjust misclassified points by constructing rules.

In the following section, we detail the proposed partially supervised hierarchical classification method for urban feature classification tasks in three steps: the lidar-based coarse classification for the selection of training samples and lidar-driven high-rise object mask generation, the image-based MLC, and the KBCV for final classification results. Then, two urban scenes are used to test the proposed algorithm, and the classification accuracy assessment is reported and discussed. Finally, conclusions are drawn in the last section.

2. Methodology

Figure 1 illustrates the workflow of the major components and their subsequent processes in the partially supervised hierarchical classification (PSHC) system proposed in this study. Lidar point clouds with intensity information and a colour aerial image covering the same area are used as the input. Spatial registration of lidar point clouds and aerial imagery is performed as data pre-processing, but is not discussed in this article.

The proposed PSHC actually comprises three stages, which classify unstructured lidar points into four classes (grass, bare ground, high vegetation, and building) by combining lidar point data and aerial imagery. The lidar-based and image-based classification stages provide their classification results mainly based on their single data sources. However, there are some interactions between these two classification stages. For example, in the lidar-based stage, the Red band of the aerial image is used to combine with the intensity information for the separation of grass and ground. In the image-based stage, the lidar-based classification results are applied to the training sample selection and a high-rise object mask generation. Finally, the knowledge-based rules constructed from lidar data are implemented to obtain final classification results by comprehensively analysing the discrepancy of the image-based and lidar-based classification results.

2.1. Lidar-based classification

In the lidar-based classification stage, we deal with only lidar data to obtain coarse classification results not only for acquiring training samples of the four classes of interest, but also for calculating the normalized digital surface model (nDSM) as a high-rise object mask

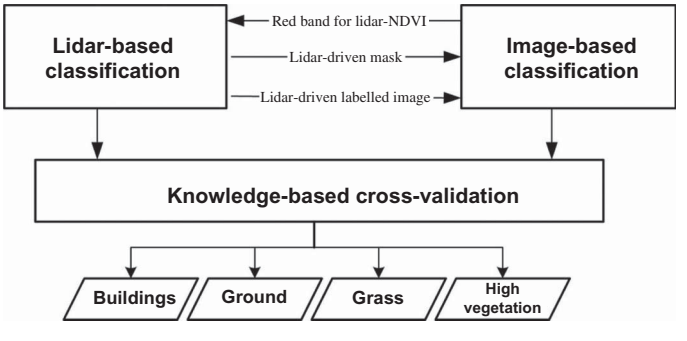


Figure 1. Flow chart of the proposed method.

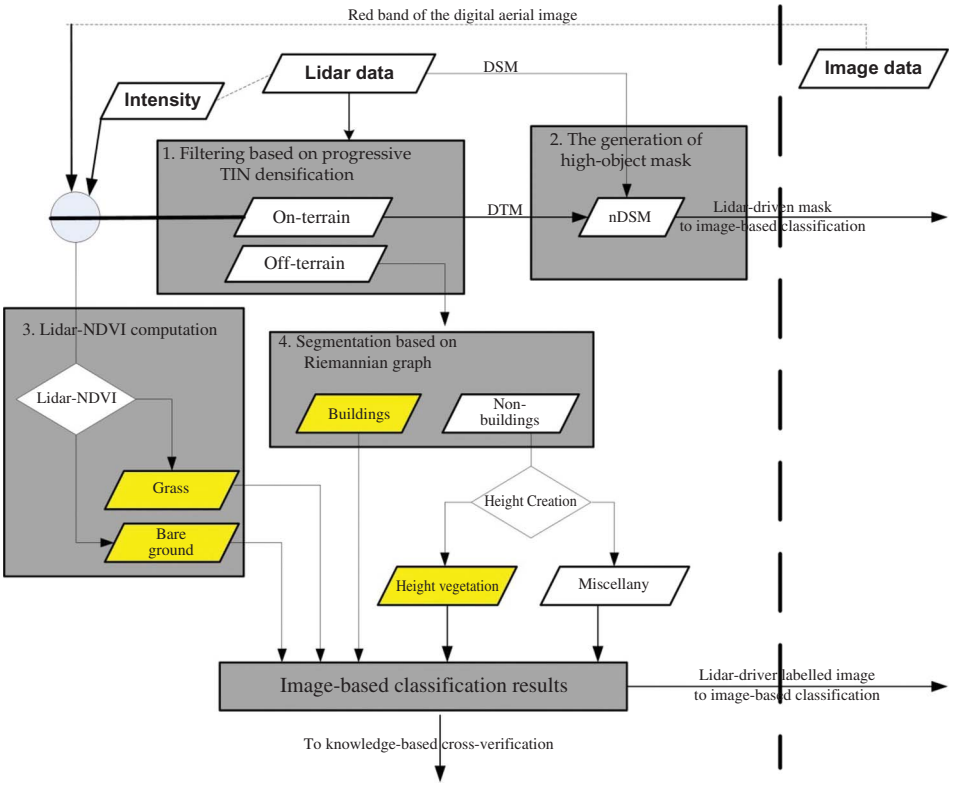


Figure 2. Sub-flow chart of the lidar-based classification.

for the next image-based MLC. Figure 2 shows a detailed flow chart of the lidar-based classification. The processing steps are divided into four parts.

(1) *The Separation of on-terrain and off-terrain Points*
Lidar point cloud data are filtered into off-terrain and on-terrain points based on the algorithm of progressive triangulated irregular network (TIN) densification that is detailed in Axelsson (2000). This filtering algorithm, one version of which has been used in the commercial software Terrasolid®, is considered to be robust and steady for modelling surfaces with discontinuities such as urban areas. The filtering method assumes that objects on the ground, such as trees, cars, and buildings, etc., are usually higher than those of on-terrain points. First, the lowest points in lidar data, such as seed points, within a user-defined grid of a size greater than the largest type of features are selected to join an on-terrain data set. In order to minimize grid size sensitivity to the selection of seed points, from which a parametric plane is derived, some seed points will be removed if their perpendicular distances to the fitting plane do not satisfy the median value estimated from the histograms. Then, the rest of the seed points generate a sparse TIN as an initial digital elevation model. For each iteration, points in each TIN facet are added to the on-terrain data set if they meet the criteria based on the calculated parameters, distances to the TIN facets, and angles to the nodes. At the end of each iteration, the TIN and the

thresholds are recomputed. The iterative process continues until no more points meet the threshold values.

(2) *The generation of a high-rise object mask*
An nDSM, a representation of elevated objects on a flat surface, is generated by subtracting DTM from the digital surface model (DSM). Both DSM and DTM are in the raster format with the same spatial resolution. A high-rise object mask is made from the nDSM that includes possible points of buildings and high vegetation. Generally, the heights of buildings are more than 2.0 m in urban areas; therefore, the high-rise object mask is derived with a given height threshold of 2.0 m for the separation of the majority of buildings in the image-based classification.

(3) *The separation of ground and grass*
According to Jutzi and Gross (2009), the intensity information is the physical power of incoming echoes, considered as a synonym for the amplitude, reflectance, or energy in the terminology of laser scanning. Different ground objects have different reflecting characteristics within the range of the near-infrared wavelength, which is useful for recognizing different types of features. For example, road surfaces paved with asphalt usually appear darker than building roofs with concrete, and water often appears the darkest because it absorbs infrared laser light. Thus, the intensity image is similar to the low-resolution grey image. Based on this feature, grass points are roughly separated from filtered on-terrain points using the lidar-normalized difference vegetation index (lidar-NDVI). The lidar-NDVI can be calculated by the following equation:

$$\text{Lidar-NDVI} = \frac{\text{NIR} - \text{R}}{\text{NIR} + \text{R}}, \tag{1}$$

where R is the Red band from the ancillary aerial image and NIR is the grey value of the intensity image. Compared with tree intensity, which varies with type, species, and height, grass intensity information is relatively stable. Then, the results of lidar-NDVI is stretched into the range from 0 to 255 as a vegetation-index image, from which the median threshold method is used to separate grass from ground points based on the histogram statistics.

(4) *The separation of buildings and non-buildings*
We adopt a Riemannian graph to segment classified off-terrain lidar points into two classes: building and high vegetation. Because lidar points are acquired at an altitude of around 2000 m, the point density of a discontinuous feature, such as a building breakline, is lower than that of the object inside; thus, the distances of points in an intra-object are shorter than those of between-object points. According to this theory, different objects can be distinguished if we label points close to oneanother as one object by using the Riemannian graph.

The Riemannian is a graph that consists of the edges to the *k*-nearest neighbours for every data point. According to Gumhold, Wang, and Mcleod (2001), the Riemannian graph is robust under conditions of very noisy data and is able to construct the connectivity information. The *Kd*-tree data structure contributes to the finding of the neighbourhood of a lidar point by using the approximate nearest-neighbours (ANN algorithm, open source developed by David M. Mount, University of Maryland). During the tests, a neighbourhood *k* size of 14–20 is suitable for delivering good results. In theory, each object is a sub-graph of interconnected points so that objects are obtained by iterating the Riemannian graph of

lidar points to find sub-graphs based on the graph distances. However, in practice, due to the presence of noise and the complexity of objects, k-neighbours of each point could not be inside objects of interest. Thus, a height difference criterion is introduced into the clustering process. When the graph distances and height differences of points meet the given thresholds, the neighbours are considered to be connective and belong to the same objects. Starting from a known building point, all points having the same characteristics with the start point are obtained to form a building segment by iterating the Riemannian graph.

After building points are separated from off-terrain points, the remaining off-terrain points, according to the height criterion, are classified into high vegetation and miscellany (e.g. cars). Generally at the end, a whole 3D lidar point cloud is divided roughly into five classes: building, high vegetation, ground, grass, and miscellany. By using the bi-linear interpretation method, a lidar-driven labelled image is generated from the first four urban features with the same spatial resolution as aerial imagery for the selection of training samples.

2.2. Image-based classification

In this stage, we focus on processing mainly the aerial image simultaneously acquired with lidar data. With the assistance of the lidar-based classification results, the statistical MLC is applied to obtain five classes (i.e. grass, building, ground, high vegetation, and miscellany). Figure 3 shows the flow chart of the image-based classification.

The aerial image needs to be pre-processed before the classification is performed, because during the image acquisition and the electronic signal transmission process,

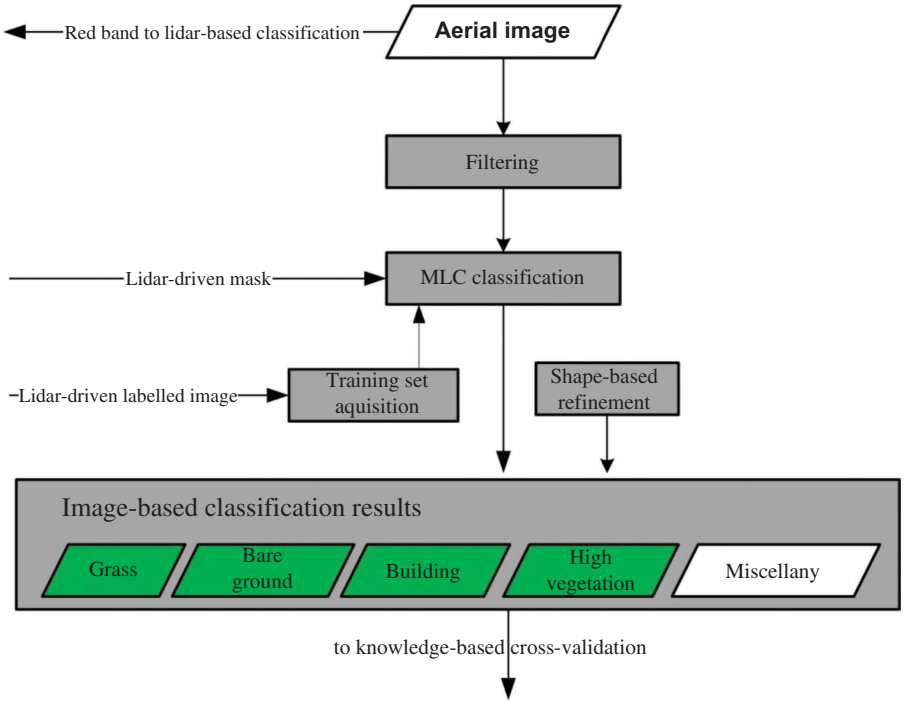
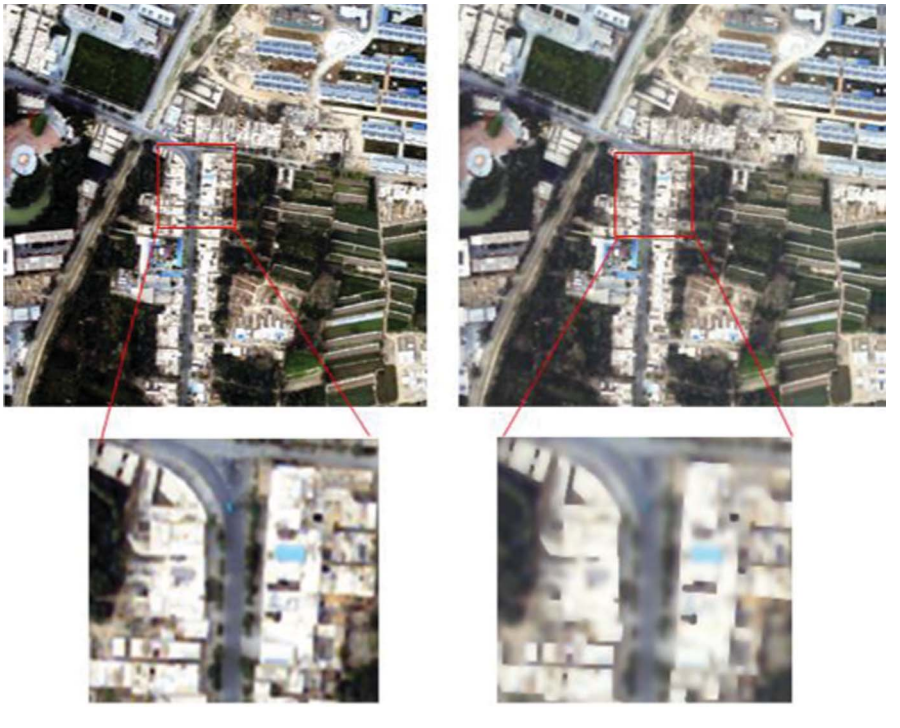


Figure 3. Sub-flow chart of the image-based classification.

external factors such as the atmospheric effect and instrumentation errors cause random noises superimposed on the pixel brightness. This noise can be removed by the process of low-pass filtering or smoothing. We use anisotropic diffusion, also called Perona–Malik diffusion. It is a technique aimed at reducing image noise without removing significant parts of the image content, typically edges, lines, or other details that are important for the interpretation of the image (Perona and Malik 1990). Figure 4 shows an example of the Perona–Malik diffusion. Figure 4(*a*) is the image without using the Perona–Malik diffusion and Figure 4(*b*) shows the filtering results using the Perona–Malik diffusion. This comparison demonstrates that building boundaries and edges are sharpened and enhanced, and buildings' inner detail information is smoothed to reduce speckle noises, which is beneficial to classification implementation.

Several seed points are randomly selected for each class of interest in the lidar-driven labelled image. These seeds then grow into a number of regions in eight directions. According to the location of training samples in the lidar-driven labelled image, corresponding statistical information of training samples can be gathered in the aerial image.

Training samples from a known data set such as lidar data will provide a large number of high-precision samples, which means that larger sample sizes generally lead to increased precision when estimating unknown parameters. It is necessary to identify training samples for all classes of interest in the image to be analysed. However, it is impractical to expect the availability of large quantities of training samples in the traditional supervised classification methods because labelling data requires much laborious human effort and takes much time. According to Richards and Jia (2006), if pixels in the training samples are not representative of the signature of a given class, the desired class will not be classified correctly. The



(*a*) (*b*)

Figure 4. An example of the Perona–Malik diffusion: (*a*) before and (*b*) after.

proposed partially supervised analysis employs training samples provided by lidar data, a different data set, to bias the process of discovering structure in the imagery data. Liu, Shi, and Zhang (2011) mention that if the pixels or points are adequately trained by a large number of samples, the traditional statistic MLC can achieve the same classification accuracy as other machine learning classifiers. Thus, prior information of particular classes of interest, extracted from lidar data, can provide sufficient training samples and assist in calculating relatively highly accurate statistical information.

Meanwhile, an nDSM is adopted as a lidar-driven high-rise object mask when the MLC is carried out for the aerial image. Huang et al. (2008) stated that a lidar-driven nDSM has great potential for classification tasks. According to the lidar-driven high-rise object mask, the ancillary aerial image can be divided into high-height and low-height regions. In the high-height regions, the ancillary aerial image is classified into two classes (high vegetation and buildings) depending on the training samples discussed above. In the low-height regions, the image is labelled as the other two classes of interest (grass and bare ground). Usage of the lidar-driven high-rise object mask can effectively lessen the confusion between spectrally identical or similar objects; for example, sometimes it is hard to distinguish alpha roads from building roofs because of their similar material spectral characteristics. Thus, the ancillary aerial image is finally classified into the four classes of interest and class miscellany by the partially supervised MLC by means of the lidar-driven high-rise object mask.

Shape information of ground objects, the basis of information extraction from ground objects, can be calculated for distinguishing buildings from other lidar points because the majority of buildings are square or rectangular in the real world (Zhang and Zhu 2011). Considerable research has been dedicated to improving land-use/land-cover classification accuracy by integrating ground object shape information (Werff and Meer 2008). Segl et al. (2003) point out that the majority of buildings have the characteristics of a rectangular or quadratic shape with a length of 10–50 m and a width of 5–20 m, by which industrial or residential buildings are separated from roads. The shape index can be represented in three ways: the perimeter and area, the area alone, and the ratio of the perimeter to the area (compactness). The calculation of the object area (S_{area}) is a summation of all pixels in an object; the perimeter ($S_{\text{perimeter}}$) of an object is a summation of pixels belonging to its boundary. Compactness is defined as (Burghardt and Steriniger 2005):

$$\text{Comtactness} = 2\frac{\sqrt{\pi S_{\text{area}}}}{S_{\text{perimeter}}}. \tag{2}$$

The compactness is a measure of the regularity of the object contour (Heinzle, Anders, and Sester 2006). Its value is between 0 and 1; the maximum value of 1.0 is for a circle. The compactness shows significant difference between regular, man-made objects like buildings and irregular, narrow-shaped natural objects like trees. The greater the value of compactness is, the higher the possibility of buildings. Thus, using compactness, misclassified building pixels can be removed from the building class. Our experiments demonstrate that a threshold of 0.55 can remove most narrow, irregular, and jagged connected objects from the building class.

2.3. Knowledge-based cross-validation

After the image-based classification, the final classification results are achieved by comprehensive cross-validation of the image-based and lidar-based classification results. The

lidar-based classification mainly uses lidar data, whereas the image-based classification is mainly based on spectral information. Thus, their classification results are complementary. In this stage, a KBCV method is applied to combine two-type preliminary classification results. Although the lidar-driven labelled image and the lidar-driven high-rise object mask are acquired from the lidar-based classification results, the MLC still causes lots of classification errors relying on the spectral and shape information. Due to the limitations of any single classifier, it has been experimentally observed that obtaining satisfaction classification results from a single classifier is impossible for any given task. There is a tendency to combine multiple classification methods or multiple classification results to improve classification accuracy (Liu, Skidmore, and Oosten 2002). There are two classification results given in this article: the lidar-based and the image-based results. Any discrepant points in two-classification results are cross-verified. In other words, if a point in the lidar-based classification results is classified into the same class with its corresponding pixel in the image-based classification results, the labelled point is considered as right. The developed KBCV method is based on three characteristics of lidar data, including intensity, height, and eigenvalue. Thus, we create validation rules to adjust those misclassified points.

2.3.1. *Height-based rule*

The height-based rule is created to adjust the misclassified building and bare ground points. The building pixels are often mixed up with the bare ground pixels in the spectral images because of their similar material characteristics; however, they can be distinguished from each other in lidar data in terms of the height criterion. Thus, we iterate lidar points (P_i, where $i = 0, 1, 2 \ldots N$) assigned as the building class, interpolate them into the DEM generated from ground points, and then compare interpolated elevation values with their real elevation values. This rule is represented as:

$$\text{if } H_i < H_{\text{threshold}} \text{ then } P_i \in \text{G}, \tag{3}$$

where H_i is the height value of point P_i. The given threshold $H_{\text{threshold}}$ is determined by the classified building results in the lidar-based classification stage. The absolute height of each classified building point is counted to form a histogram, a useful graphic representation of the information content of buildings.

2.3.2. *Intensity-based rule*

As mentioned in Section 2.1, intensity information is useful for identifying vegetation from other features in the data sets. Thus, the intensity-based rule is used to correct misclassified points between high vegetation and buildings. Local intensity mean values are calculated in building (BI) and high vegetation (TI) regions, respectively. Owing to this type of misclassification occurring along buildings' boundaries, buffer zones along the building boundaries are first required to build up, and then building points ($P_i, i = 0, 1, 2 \ldots N$) are iterated inside these buffer zones. This rule is represented as:

$$\begin{aligned} &\text{if} \quad \begin{array}{l} \text{Int}_i < (\text{BI} - \text{BI}/4) \text{ and } \text{Int}_i > (\text{BI} + \text{BI}/4) \\ \text{and } (\text{TI} - \text{TI}/4) < \text{Int}_i < (\text{TI} + \text{TI}/4) \end{array} \quad \text{then } P_i \in \text{High vegetation} \\ &\text{else } P_i \in \text{Miscellany}, \end{aligned} \tag{4}$$

where Int_i is the intensity value of point P_i.

2.3.3. Eigenvalue-based rule

The last rule is eigenvalue based. The eigenvalue-based rule is defined as the spatial feature of each point by calculating a variance–covariance matrix of its neighbours. It is another auxiliary indicator for distinguishing planes, edges, corners, and volumes (Gross and Thoennessen 2006). For each point v_j under consideration, its neighbourhood points can be found by the *Kd*-tree (*K*-dimensional tree), a space-partitioning useful data structure for organizing points in k dimensional space. A 3×3 variance–covariance matrix $\mathbf{S_j}$ of point v_j is given by

$$\mathbf{S_j} = \sum_{j=0}^{n} (P_j - M)^T \left(P_j - M\right) \forall \, (j = 0, 1, 2 \ldots N - 1)\,, \tag{5}$$

where $\mathbf{S_j}$ is the variance–covariance matrix of point v_j, n is the number of neighbourhood points of point v_j, P_j is the coordinate $\left(x_j, y_j, z_j\right)$ of point v_j, N is the total number of lidar points, and M is a 1×3 mean vector of its neighbourhood points.

In this 3×3 variance–covariance matrix, each point v_j has three eigenvalues $(\lambda_1, \lambda_2, \lambda_3)$. An eigenvalue $\lambda_i \, (i = 1, 2, 3)$ represents the spatial information of a lidar point because it is a scalar associated with an eigenvector $\vec{e}_i \, (i = 1, 2, 3)$, which reflects the spatial distribution of a lidar point.

$$\text{Anisotropy} = \frac{\lambda_1 - \lambda_3}{\lambda_1}, \tag{6}$$

$$\text{Planarity} = \frac{\lambda_2 - \lambda_3}{\lambda_1}, \tag{7}$$

$$\text{Linearity} = \frac{\lambda_1 - \lambda_2}{\lambda_1}. \tag{8}$$

The planar objects, such as buildings and ground, show high values of planarity. On the other hand, high vegetation, in general, obtains a high value of anisotropic distribution (Chehata, Guo, and Mallet 2009).

3. Results and discussion

Two data sets are used in this study to test the performance and applicability of the proposed classification method. Test site I data were collected in the multi-return mode by a Leica ALS50-II scanner (airborne laser scanner; Leica Geosystems Inc., Norcross, GA, USA) and test site II data were collected by an Optech ALTM 3100 system (airborne laser scanner; Optech Inc., Toronto, Canada). Both systems were mounted on helicopters.

3.1. Dunhuang test site in Gansu

Test site I data were acquired in Dunhuang, Gansu, western China, in October 2009. The sensor, operated at a fixed wavelength (1.064 μm) and flown at 2750 m above sea level, generated points at nadir for a 45° field of view. Figure 5(*a*) shows a lidar point cloud of 628,085 points. The lidar data have average point spacing of 1.4 m in and along the cross-track directions, with a horizontal accuracy of 27 cm and a vertical accuracy of 15 cm. Lidar points of different height values are shown as a different shade of grey. Figure 5(*b*) is a colour aerial image acquired from an RCD105 digital frame camera with a ground

(*a*) (*b*)

Figure 5. Study area of Dunhuang, China. (*a*) Lidar data and (*b*) aerial image.

sample distance (GSD) of 0.5 m. The orientation parameters of the aerial image are known. The experimental area is a typical urban region that contains variously sized buildings with different orientations, as well as trees and grass interspersed among buildings. Meanwhile, the study area and its vicinity are relatively flat, with elevations ranging from approximately 1132 to 1170 m.

3.1.1. Lidar-driven training samples and high-rise object mask experiments

In this experiment, the lidar-driven training sample strategy and high-rise object mask are investigated by quantitative assessment, compared to a conventional photo-interpretation method. In general, the primary step in the supervised classification is the prior identification of training samples. Although sometimes map data in the GIS database can be used to assist in selection of training samples by superimposing them over the image data, training samples in most cases are chosen by the photo-interpretation method from images formed from the multispectral data to be classified (Richards and Jia 2006).

For the sake of comparison, the sample size of the proposed partially supervised method was identical/close to the photo-interpretation method using commercial software ENVI®. Table 1 lists the number of training samples. As a proportion of the full image to be analysed, the number of training samples would represent less than 1–5%. For accuracy assessment, an adequate number of testing data are required per class of interest. Congalton

Table 1. The training samples and test data for Dunhuang study site.

	Training samples (manual)		Training samples (lidar-driven)		Test data
Categories	Pixels	ROIs	Pixels	ROIs	Pixels
Buildings	1465	53	1453	57	415
High vegetation	709	87	733	83	197
Ground	1404	58	1431	49	317
Grass	1334	45	1306	51	193

and Green (2009) pointed out that it is necessary to have sufficient testing data for building a valid error matrix to represent classification accuracy. Thus, the sample size N was determined by Equation (9) for the binomial probability theory (Fitzpatrick-Lines 1981):

$$N = \frac{Z^2 p\,(100 - p)}{E^2}, \tag{9}$$

where p is the expected percentage accuracy, E is the allowable error, and $Z = 1.96$ from the standard normal deviant for the 95% two-sided confidence level. An expected accuracy of 95% was selected because the land-use classification system specifies that each class category should be mapped to at least 85% accuracy, and then the allowable error of 5% is chosen. For this study area, the sample size (N) of 996 meets the demand of Congalton and Green's (2009) rule-of-thumb of a minimum of 50 samples per class.

Three experiments are performed for comparison. First, the aerial image alone is used as the standard case for comparison (case MLC1 in Table 2). The training samples are required by manual photo-interpretation. Second, the selection of training samples is changed into the lidar-driven strategy (case MLC2). Third, the lidar-driven high-rise object mask is added (case MLC3).

Figure 6(*a*) shows that the lidar-driven labelled image, interpolated from the coarse lidar-based classification results, is used to extract training samples by a region-growing algorithm for the image-based classification in the second stage. Starting from randomly selecting several seeds of each class of interest, training sample regions are acquired. As shown in Figures 6(*b*)–(*d*), building, tree, and ground training samples are overlapped on the aerial image. Visual inspection shows that these training samples completely cover the desired classes. Unlike traditional sample selection methods, the lidar-driven training sample selection is automated and not only guarantees sample accuracy and capacity but also saves training time.

One of the most common methods of expressing classification accuracy is the preparation of a classification error matrix (confusion matrix). An error matrix is an effective way to assess accuracy in that it compares the relationship between known reference data and the corresponding results of the classification (Congalton 1991). It is a square matrix E of $N \times N$ elements, where N is number of classes. The element E_{ij} is the number of points known to belong to class i and classified as belonging to class j. Thus, the elements on the leading diagonal E_{ii} correspond to correctly classified points, whereas the off-diagonal elements correspond to erroneous classifications (i.e. the commission and omission errors). From the confusion matrix, the user's (UA), producer's (PA), overall accuracy (OA) (Story and Congalton 1986), and kappa coefficient (Congalton, Oderwald, and Mead 1983; Congalton and Green 2009) can be calculated. Table 2 lists the accuracy

Table 2. Accuracy assessment of training sample experiments.

			Producer's accuracy (%)				User's accuracy (%)			
Experiment	OA (%)	Kappa	Bs	HV	Gd	Gs	Bs	HV	Gd	Gs
MLC1	70	0.57	66	60.8	71.9	74	81.1	43.1	65.1	81.1
MLC2	78.1	0.66	84.1	62.1	70.0	84.3	75.3	56.1	81.7	85.1
MLC3	85.6	0.79	87.4	80.6	85.9	75	88.1	69.1	84.8	84.8

Notes: MLC 1, photo-interpretation training sampling; MLC2, lidar-driven training sampling; MLC3, lidar-driven training sampling and lidar-driven nDSM mask; Bs, buildings; HV, high vegetation; Gd, ground; and Gs, grass.

Figure 6. Lidar-based classification results: (*a*) lidar-driven labelled image, (*b*) lidar-driven building training samples overlapping on the aerial image, (*c*) lidar-driven high vegetation training samples overlapping on the aerial image, and (*d*) lidar-driven ground training samples overlapping on the aerial image.

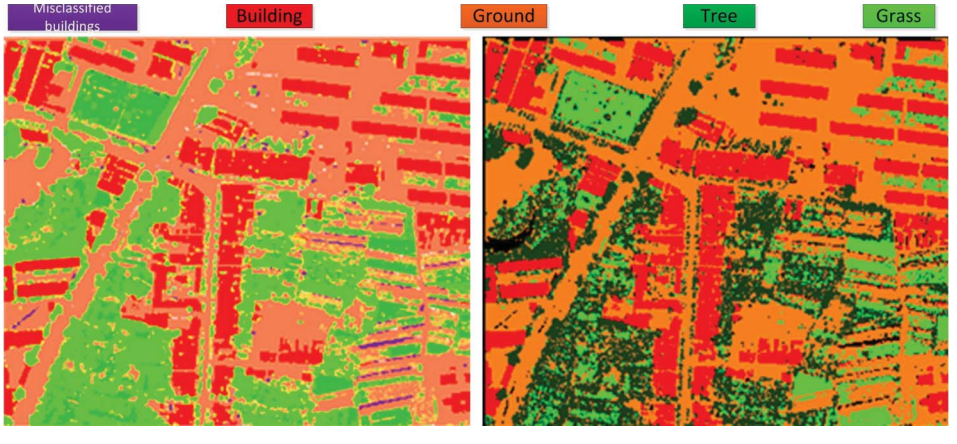
assessment results for the comparative experiment on training sample selection (cases MLC1 and MLC2). The overall accuracy of only the standard aerial image experiment (case MLC1) is 70.0%, which is unsatisfactory for urban classification applications. The overall accuracy is improved by about 8% when using the lidar-driven labelled image (case MLC2). Generally, the overall accuracy and kappa coefficient of the lidar-driven training sampling strategy are higher than those of photo-interpretation. In particular, the producer's and user's accuracy have more significant improvement than those based on an interpreter's empirical expertise. Because high vegetation is rare and sparsely scattered, ground or partial building boundary pixels may be included in the high vegetation training samples under the photo-interpretation method and might lead to decreased classification accuracy. Adding the lidar-driven high-rise object mask improves the overall accuracy by about 8% (case MLC3), compared with case MLC2. Spectral confusion between buildings and roads and between high vegetation and grass might be solved, because the lidar-driven high-rise object mask leads classification to be implemented in the high-height and low-height levels, respectively. Thus, the contribution of the lidar-driven labelled image and high-rise object mask has been confirmed by accuracy assessment in this study.

3.1.2. The PSHC and OBC experiments

The OBC and PSHC are performed to investigate the proposed PSHC performance. The feature vector used for this experiment contains shape index, spectral information, lidar intensity, lidar-driven nDSM, height variance, and eigen-based values. The first two are derived from aerial imagery, whereas the other four are lidar based. These features are calculated using C++ programming language.

The OBC experiment was implemented by the segmentation of an image into a network of homogeneous image regions and the use of a supervised nearest-neighbour classifier to train and build up a knowledge base for the classification of image objects, which was performed using the Definiens Professional Earth® commercial software (also called eCognition®). As discussed above, the nDSM mask has a great influence on the accuracy of classification so that a high weight value is assigned. Other features are assigned identical weight values. Visual inspection determines a set of segmentation parameters to create a number of regions by repeating experiments. Afterwards, we use the same lidar-driven training samples listed in the Table 1. After a number of sample objects have been declared as initial information for a nearest-neighbour classification, the four categories are classified. Similarly, the classification accuracy is evaluated by using the testing samples listed in Table 1.

The proposed PSHC is performed using C++ programming language in Microsoft Studio 2005®. The prior knowledge of desired classes is derived from training samples of the lidar-driven labelled image. The lidar-driven high-rise object mask divides the aerial image into high-height and low-height levels. At each level, the statistical MLC is performed. Because a building, in general, is regular, the value of shape measure calculated by Equation (2) would be closer to 1.0 for a regular boundary. Therefore, the shape measure removes a partial misclassification and improves classification accuracy. Figure 7(*a*) is the image-based classification results refined by the shape-based information. In the figure, the purple pixels represent the misclassified buildings that should belong to the high vegetation class. It should be noted that the 'salt and pepper' classification effect occurs owing to spectral confusion between class and spectral variation within class; for example, there are a few discrete and discontinuous regions in the building or high vege-



(*a*) (*b*)

Figure 7. The results of classification: (*a*) the image-based classification results refined by the shape-based information and (*b*) the final classification results after the KBCV stage.

Table 3. Accuracy assessment of KBCV and OBC experiments.

			Producer's accuracy (%)				User's accuracy (%)			
Experiment	OA (%)	Kappa	Bs	HV	Gd	Gs	Bs	HV	Gd	Gs
OBC (Definiens)	86.5	0.82	86.4	72.7	87.1	77.6	91.1	79.6	90	84.3
PSHC	90.4	0.88	92.4	73.3	95.7	83.1	90.5	81.4	93.2	87.8

tation classes. In the KBCV stage, according to lidar-based and image-based classification results, the discrepancy between these two classification results is cross-validated by the analysis of three rules: height-based, intensity-based, and eigen-based rules. Thresholds are set from expert experience, including height, intensity, and spatial distribution parameters. Figure 7(*b*) shows the final classification results after the KBCV stage.

Table 3 shows the accuracy assessment results for the OBC and PSHC. Compared with case MLC3 in Table 2, 5% improvement of the accuracy indicates that the KBCV is the necessary step to improve the classification performance after the implementation of MLC. As shown in the Table 3, the overall accuracies of the OBC and PSHC are higher than those of the traditional statistical MLC (cases MLC1 to MLC3 in Table 2). However, because the OBC segmentation parameters are chosen based on visual interpretation of the segmentation results, it is difficult to determine the segmentation parameters. For example, small objects may not be identified if the segmentation parameters are large, while a number of much smaller objects may be so small that their characteristic cannot be calculated if the segmentation parameters are very small. Thus, the overall accuracy and kappa coefficient values of the PSHC are greater than those of the OBC.

3.2. Niagara Falls test site in Ontario

The second lidar point cloud data were acquired by the Optech ALTM 3100 system, covering a typical residential area in the region of Niagara Falls, Ontario, Canada. The lidar data set consists of the first and last returns of the laser beam. The true colour image data were simultaneously taken by an onboard 4k × 4k digital camera. Figure 8(*a*) shows a raster DSM, containing a total of 190,944 points, which was interpolated with both the first and last pulse returns by the bi-linear interpolation method. The width and height of the grid are equal to the GSD of the aerial image (0.5 m). The elevation of the study area ranges from 150.00 to 178.11 m. Figure 8(*a*) shows that several clusters of trees are located around buildings along the streets and Figure 8(*b*) shows a true colour aerial image that was re-sampled to a 0.5 m ground pixel. The majority of buildings appearing in the colour image have gable or hip roofs. From the aerial image, it is difficult to classify high vegetation because the colour of vegetation ranges from green to brown and the canopy density is relatively low.

Figure 8(*c*) shows the resultant classification image of the image-based classification process. A number of classification errors involve the missing of building boundaries, and can be compensated for by spatial information of lidar data in the cross-validation stage. Figure 8(*d*) shows the final classification resultant image, from which a majority of speckles and noises which resulted from the image-based classification are removed by the KBCV. Table 4 shows the accuracy assessment results for three classification stages in the study. The overall accuracy is increasing from the lidar-based classification to the KBCV stage. The overall accuracy of the image-based classification does improve by about 9%, and

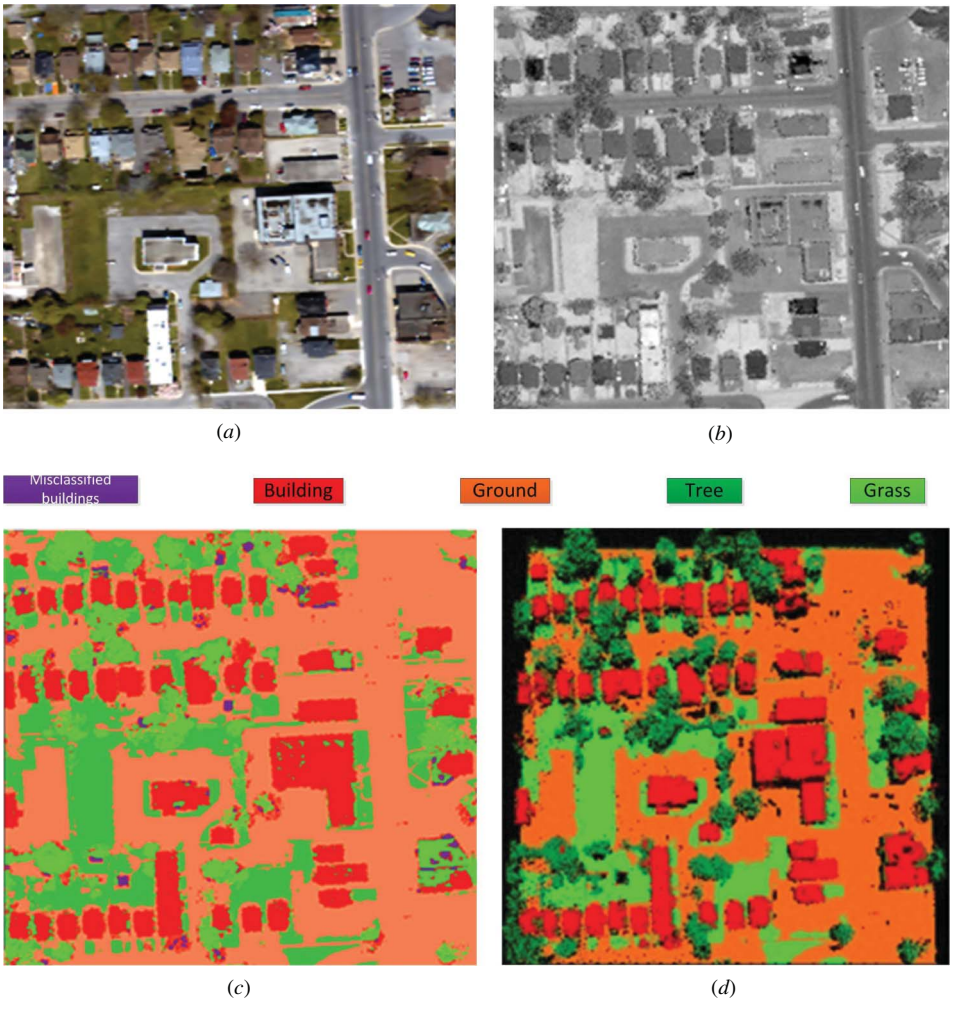


Figure 8. Niagara Falls test site in Ontario: (*a*) the aerial image, (*b*) the intensity image, (*c*) the image-based classification results, and (*d*) the final classification results.

Table 4. Accuracy assessment of KBCV and OBC experiments.

			Producer's accuracy (%)				User's accuracy (%)			
Experiment	OA (%)	Kappa	Bs	HV	Gd	Gs	Bs	HV	Gd	Gs
Lidar based	73.5	0.658	78.5	46.0	78.4	57.6	83.5	56.9	76.2	77.3
Image based	82.1	0.83	86.1	67.8	82.5	82.6	84.3	79.8	83.1	81.7
PSHC	91.7	0.9	91.8	75.7	94.9	85.3	91.2	82.4	92.6	88.4

the KBCV overall accuracy of 91.7% improves significantly from the image-based overall accuracy of 82.1%. This result indicates that the hierarchical classification strategy is feasible for the classification accuracy increase by passing down the previous classification results to the next classification stage. This simple hierarchical scheme offers a wide range of construction of different analysis strategies.

Generally, the advantage of the proposed hierarchical classification framework is that it offers a mechanism not only to improve training sample accuracy and reduce training time, but also to overcome the ambiguity between high- and low-height objects. Based on this design, the categories at the low-height level are roads and grass, while only buildings and trees need to be classified at the high-height level. The KBCV stage also improves classification abilities by adding a variety of knowledge-based rules to adjust the lidar-based and image-based classification results.

4. Conclusions

Since training samples are used to 'teach' a supervised classifier, it is essential to make sure that training samples are sufficient and representative. Generally, the manual selection of training samples often includes at least two types of land-use features within training areas. Individual classes are difficult to distinguish correctly by using incomplete statistic information extracted from training samples. This condition results in the decrease of classification accuracy to some extent. The proposed PSHC strategy first utilizes the lidar-based classification results to generate a labelled image for training sample selection and a lidar-driven high-rise object mask. Automatic training sample selection not only keeps classification results consistent, but also improves the quality of training samples. A lidar-driven high-rise object mask reduces spectral confusion between land-use feature types; for example, at the high-height level, high vegetation and buildings are easier to separate because their spectral characteristics are quite distinct. Similarly, at the low-height level, ground and grass are significantly separable. Lidar-driven experiments have shown that the overall accuracy is improved by approximately 8% when using the lidar-driven training samples compared to training samples selected by photo-interpretation; in addition, the lidar-driven high-rise object mask improves the overall accuracy by about 8%.

In terms of spatial and spectral information, the lidar-based and image-based classification results are complementary. Meanwhile, the knowledge-based rules are created to cross-validate the discrepancy between two types of classification results. The OBC and PSHC experiment results indicate that the overall accuracy of the PSHC is 3% better than that of the OBC method. Moreover, the Niagara Falls test site experiment demonstrates that the multi-step classification strategy improves the accuracy of classification by analysing previous classification results. Another advantage of the proposed PSHC approach is that the KBCV also can be extended by adding various rules from lidar data and imagery according to users' requirements.

Acknowledgements

This research was supported by the National Basic Research Programme of China (2012CB719904). We would like to acknowledge the anonymous reviewers for their constructive criticism.

References

Axelsson, P. 2000. "DEM Generation From Laser Scanner Data Using Adaptive TIN Models." *International Archives of the Photogrammetry and Remote Sensing* 33, no. B4/1: 110–17.

Blaschke, T. 2010. "Object Based Image Analysis for Remote Sensing." *ISPRS Journal of Photogrammetry and Remote Sensing* 65, no. 1: 2–16.

Burghardt, D., and S. Steiniger. 2005. "Usage of Principal Component Analysis in the Process of Automated Generalization." In *Proceedings of the Twenty-second International Cartographic Association Conference*, A Coruña, Spain, July 8–9 (CD–ROM).

Carlberg, M., P. Gao, G. Chen, and A. Zakhor. 2009. “Classifying Urban Landscape in Aerial LiDAR using 3D Shape Analysis.” IEEE International Conference on Image Processing (ICIP), Cairo, Egypt, November 7–10.

Chehata, N., L. Guo, and C. Mallet. 2009. “Airborne Lidar Feature Selection for Urban Classification Using Random Forests.” *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* 39, Part 3/W8: 207–12.

Congalton, R. G. 1991. “A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data.” *Remote Sensing of Environment* 37, no. 1: 35–46.

Congalton, R. G., and K. Green. 2009. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. 2nd ed. Boca Raton, FL: Taylor & Francis.

Congalton, R. G., R. G. Oderwald, and R. A. Mead. 1983. “Assessing Landsat Classification Accuracy Using Discrete Multivariate Statistical Techniques.” *Photogrammetric Engineering and Remote Sensing* 49, no. 12: 1671–78.

Dalponte, M., L. Bruzzone, and D. Gianelle. 2008. “Fusion of Hyperspectral and LIDAR Remote Sensing Data for Classification of Complex Forest Areas.” *IEEE Transactions on Geoscience and Remote Sensing* 46, no. 5: 1416–27.

Fitzpatrick-Lines, K. 1981. “Comparison of Sampling Procedures and Data Analysis for a Land Use and Land Cover Maps.” *Photogrammetric Engineering and Remote Sensing* 47, no. 3: 343–51.

Foody, G. M., A. Mathur, C. Sanchez-Hernandez, and D. S. Boyd. 2006. “Training Set Size Requirements for the Classification of a Specific Class.” *Remote Sensing of Environment* 104: 1–14.

Gislason, P. O., J. A. Benediktsson, and J. R. Sveinsson. 2006. “Random Forests for Land Cover Classification.” *Pattern Recognition Letters* 27, no. 4: 294–300.

Goodwin, N. R., N. C. Coops, T. R. Tooke, A. Christen, and J. A. Voogt. 2009. “Characterizing Urban Surface Cover and Structure with Airborne Lidar Technology.” *Canadian Journal of Remote Sensing* 35, no. 3: 297–309.

Gross, H., and U. Thoennessen. 2006. “Extraction of Lines from Laser Point Clouds.” In *Symposium of ISPRS Commission III: Photogrammetric Computer Vision PCV06. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 3)*, edited by W. Förstner and R. Steffen, 86–91.

Guan, H., J. Li, M. A. Chapman, L. Zhong, and R. Que. 2011. “Support Vector Machine for Urban Land-use Classification Using Lidar Point Clouds and Aerial Imagery.” International Symposium on LiDAR and RADAR Mapping: Technologies and Applications, Nanjing, China, May 26–29.

Gumhold, S., X. Wang, and R. Mcleod. 2001. “Feature Extraction from Point Clouds.” Proceedings of 10th International Meshing Roundtable, 293–305, Newport Beach, CA, October 7–10.

Guo, L., N. Chehata, C. Mallet, and S. Boukir. 2011. “Relevance of Airborne Lidar and Multispectral Image Data for Urban Scene Classification Using Random Forests.” *ISPRS Journal of Photogrammetry and Remote Sensing* 66, no. 1: 56–66.

He, D., and L. Wang. 1992. “Unsupervised Texture Classification of Images Using the Texture Spectrum.” *Pattern Recognition* 25, no. 3: 247–55.

Heinzle, F., K. Anders, and M. Sester. 2006. “Pattern Recognition in Road Networks on the Example of Circular Road Detection.” In *Proceedings of Geographic Information Science: 4th International Conference, GIScience*, edited by M. Raubal, H. J. Miller, A. U. Frank, and M. F. Goodchild, 153–67. Munster, Germany, September 20–23. Berlin: Springer-Verlag.

Hodgson, M. E., J. R. Jensen, J. A. Tullis, K. D. Riordan, and C. M. Archer. 2003. “Synergistic Use of Lidar and Color Aerial Photography for Mapping Urban Parcel Imperviousness.” *Photogrammetric Engineering and Remote Sensing* 69, no. 9: 973–80.

Huang, M., S. Shyue, L. Lee, and C. Kao. 2008. “A Knowledge-Based Approach to Urban Feature Classification Using Aerial Imagery with Lidar Data.” *Photogrammetric Engineering and Remote Sensing* 74, no. 12: 1473–85.

Jensen, J. R. 2005. *Introductory Digital Image Processing: A Remote Sensing Perspective*. 3rd ed. 352, Prentice Hall Series in Geographic Information Science. Englewood Cliffs, NJ: Prentice Hall.

Jeon, B., and D. A. Landgrebe. 2002. “Partially Supervised Classification Using Weighted Unsupervised Clustering.” *IEEE Transactions on Geoscience and Remote Sensing* 37, no. 2: 1073–79.

Jones, T. G., N. C. Coops, and T. Sharma. 2010. "Assessing the Utility of Airborne Hyperspectral and LiDAR Data for Species Distribution Mapping in the Coastal Pacific Northwest, Canada." *Remote Sensing of Environment* 114, no. 12: 2841–52.

Jutzi, B., and H. Gross. 2009. "Normalization of Lidar Intensity Data Based on Range and Surface Incidence Angle." *Proceedings of ISPRS Workshop: Laserscanning 2009*, Paris, France, XXXVIII (Part 3/W8): 213–18.

Ke, Y., L. J. Quackenbush, and J. Im. 2010. "Synergistic Use of Quickbird Multispectral Imagery and Lidar Data for Object-Based Forest Species Classification." *Remote Sensing of Environment* 14, no. 6: 1141–54.

Khoshelham, K., C. Nardinocchi, E. Frontoni, A. Mancini, and P. Zingaretti. 2010. "Performance Evaluation of Automated Approaches to Building Detection in Multi-Source Aerial Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 65, no. 1: 123–33.

Koetz, B., F. Morsdorf, S. Van der Linden, T. Curt, and B. Allgower. 2008. "Multi-Source Land Cover Classification for Forest Fire Management Based on Imaging Spectrometry and Lidar Data." *Forest Ecology and Management* 256: 263–71.

Liu, K., W. Shi, and H. Zhang. 2011. "A Fuzzy Topology-Based Maximum Likelihood Classification." *ISPRS Journal of Photogrammetry and Remote Sensing* 66: 103–14.

Liu, X. H., A. K. Skidmore, and H. V. Oosten. 2002. "Integration of Classification Methods for Improvement of Land-Cover Map Accuracy." *ISPRS Journal of Photogrammetry and Remote Sensing* 56, no. 4: 257–68.

Lu, D., S. Hetrick, and E. Moran. 2010. "Land Cover Classification in a Complex Urban-Rural Landscape with Quickbird Imagery." *Photogrammetric Engineering and Remote Sensing* 76, no. 10: 1159–68.

Mesev, V., B. Gorte, and P. A. Longley. 2001. "Modified Maximum-Likelihood Classification Algorithms and their Application to Urban Remote Sensing." In *Remote Sensing and Urban Analysis*, edited by J. P. Donnay, M. Barnsley, and P. A. Longley, 71–94. London: Taylor and Francis.

Miliaresis, G., and N. Kokkas. 2007. "Segmentation and Object-Based Classification for the Extraction of Building Class from LIDAR DEMs." *Computers and Geosciences* 33, no. 8: 1076–87.

Pal, M. 2005. "Random Forest Classifier for Remote Sensing Classification." *International Journal of Remote Sensing* 26, no. 1: 217–22.

Perona, P., and J. Malik. 1990. "Scale-Space and Edge Detection Using Anisotropic Diffusion." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 12, no. 7: 629–39.

Repake, R. S., D. Truax, E. Kolstad, and C. Hara. 2004. "Comparing Spectral and Object Based Approaches for Classification and Transportation Feature Extraction from High Resolution Multispectral Imagery." ASPRS annual conference processing, 204–15, Denver, CO, May 23–28 (CD-ROM).

Richards, J. A., and X. Jia. 2006. *Remote Sensing Digital Image Analysis: An Introduction*. 4th ed. Berlin: Springer-Verlag.

Secord, J., and A. Zakhor. 2007. "Tree Detection in Urban Regions Using Aerial Lidar and Image Data." *IEEE Geoscience and Remote Sensing Letters* 4, no. 2: 196–200.

Segl, K., S. Roessner, U. Heiden, and H. Kaufmann. 2003. "Fusion of Spectral and Shape Features for Identification of Urban Surface Cover Types Using Reflective and Thermal Hyperspectral Data." *ISPRS Journal of Photogrammetry and Remote Sensing* 58, no. 1–2: 99–112.

Shan, J., and S. Lee. 2003. "Combining Lidar Elevation Data and Ikonos Multispectral Imagery for Coastal Classification Mapping." *Marine Geodesy* 26, no. 1–2: 117–27.

Smith, A. 2010. "Image Segmentation Scale Parameter Optimization and Land Cover Classification Using the Random Forest Algorithm." *Journal of Spatial Science* 55, no. 1: 69–79.

South, S., J. Qi, and D. P. Lusch. 2004. "Optimal Classification Methods for Mapping Agricultural Tillage Practices." *Remote Sensing of Environment* 91, no. 1: 90–7.

Story, M., and R. G. Congalton. 1986. "Accuracy Assessment: A User's Perspective." *Photogrammetric Engineering and Remote Sensing* 52, no. 3: 397–9.

Trinder, J. C., and M. Salah, 2011. "Support Vector Machines: Optimization and Validation for Land Cover Mapping Using Aerial Images and Lidar Data." 34th International Symposium on Remote Sensing of Environment, Sydney, April 10–15.

Tso, B., and P. M. Mather. 2001. *Classification Methods for Remotely Sensed Data*. London: Taylor & Francis.

Tuia, D., E. Pasolli, and W. J. Emery. 2011. "Using Active Learning to Adapt Remote Sensing Image Classifiers." *Remote Sensing of Environment* 115, no. 9: 2232–42.

van der Werff, H. M. A., and F. D. van der Meer. 2008. "Shape-Based Classification of Spectrally Identical Objects." *ISPRS Journal of Photogrammetry and Remote Sensing* 63, no. 2: 251–8.

Volpi, M., D. Tuia, and M. Kanevski. 2010. "Advanced Active Sampling for Remote Sensing Image Classification." Proceedings of 2010 IEEE International Geoscience and Remote Sensing Symposium (IGARSS), 1414–17, Honolulu, HI, July 25–30.

Wang, C., M. Menenti, M. Stoll, A. Feola, E. Belluco, and M. Marani. 2009. "Separation of Ground and Low Vegetation Signatures in LiDAR Measurements of Salt-Marsh Environments." *IEEE Transactions on Geoscience and Remote Sensing* 47, no. 7: 2014–23.

Zhang, R., and D. Zhu. 2011. "Study of Land Cover Classification Based on Knowledge Rules Using High-Resolution Remote Sensing Images." *Expert Systems with Applications* 38, no. 4: 3647–52.