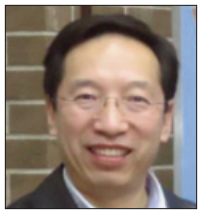
URBAN THEMATIC MAPPING BY INTEGRATING LIDAR POINT CLOUD WITH COLOUR IMAGERY

Haiyan Guan and Jonathan Li, University of Waterloo, Waterloo, Ontario
Michael A. Chapman, Ryerson University, Toronto, Ontario

This paper presents an effective approach to integrating airborne lidar data and colour imagery acquired simultaneously for urban mapping. Texture and height information extracted from lidar point cloud is integrated with spectral channels of aerial imagery into an image segmentation process. Then, the segmented polygons are integrated with the extracted geometric features (height information between first- and last-return, eigenvalue-based local variation and filtered height data) and spectral features (line segments) into a supervised classifier. The results for two different urban areas in Toronto, Canada, demonstrated that a satisfactory overall accuracy of 84.96% and Kappa of 0.76 were achieved in Scene I, while a building detection rate of 92.11%, comission error of 2.10% and omission error of 9.25% were obtained in Scene II.

Cet article présente une approche efficace à l'intégration des données lidar aéroporté et de l'imagerie en couleur acquises simultanément pour la cartographie des zones urbaines. L'information sur la texture et la hauteur extraite à partir d'un nuage de points lidar est intégrée aux canaux spectraux de l'imagerie aérienne lors d'un processus de segmentation de l'image. Ensuite, les polygones segmentés sont intégrés aux éléments géométriques extraits (information sur la hauteur entre le premier et le dernier retour de signal, la variation locale basée sur la valeur propre et les données filtrées sur la hauteur) et les éléments spectraux (segments linéaires) dans un classificateur supervisé. Les résultats pour deux zones urbaines différentes de la ville de Toronto au Canada démontrent une précision globale satisfaisante de 84,96 % et un coefficient Kappa de 0,76 dans la scène I, ainsi qu'un taux de détection des bâtiments de 92,11 %, un taux d'erreurs de commission de 2,10 % et un taux d'erreurs d'omission de 9,25 % dans la scène II.

Haiyan Guan
h6guan@uwaterloo.ca



Jonathan Li
junli@uwaterloo.ca

Michael Chapman
mchapman@ryerson.ca

1. Introduction

Today, most airborne laser scanning or light detection and ranging (lidar) systems can collect point cloud data by a laser scanner and image data by a digital camera, simultaneously. Higher thematic mapping accuracy of complex urban areas becomes achievable when both types of data are used. Airborne lidar can directly collect a digital surface model (DSM) of an urban area. Unlike a digital terrain model (DTM), the DSM is a geometric description of both terrain surface and objects located on and above this surface like buildings and trees. Lidar-derived dense DSMs have been shown to be useful in building detection, which is a classification task that separates buildings from other objects such as natural and man-made surfaces (lawn, roads) and vegetation (trees). Since a common standard property of most buildings is that they are off-terrain points, standard filtering algorithms first can be used to identify off-terrain lidar points. Existing methods for urban scene classification using lidar point clouds data alone include hierarchical Bayesian nets [*Brunn and Weidner* 1998], morphological filtering using sloped kernels [*Vosselman* 2000], and using specified features (e.g., height data, Laplace filtered height data, and maximum slope) [*Mass* 1999]. Fusing lidar point cloud with the digital imagery is promising to improve urban scene classification accuracy [*Haala* et al. 1998; *Haala and Walter* 1999; *Zeng et al*. 2002; *Rottensteiner et al*. 2003; *Collins et al.* 2004; *Charaniya et al*. 2004; *Hu and Tao* 2005; *Walter* 2005; *Rottensteiner et al*. 2005; *Brattberg and Tolt* 2008; *Huang et al*. 2008; *Chehata et al.* 2009; *Awrangjeb et al*. 2010; *Germaine and Hung* 2010]. Besides multispectral imagery, colour infrared (CIR) imagery has also been used to perform a pixel-based land-use classification [*Haala and Brenner* 1999; *Bartels and Wei* 2006].

Traditional pixel-based classification approaches only use spectral information which is inadequate for classifying high-resolution multispectral images in urban environments, because each pixel is individually classified as a certain group, which results in more noise, and classified urban land-cover coars-

er than the size of objects of interest [*Lu et al*. 2010]. Spatial dimensions such as distance, neighborhood, and texture are crucial to enhance the classification performance, which attracts a great attention on the studies of object-oriented classification methods. It has been conclusively proven that the object-oriented classification method is better than pixel-based classification approaches [*Thomas et al.* 2003; *Lu et al.* 2010; *Blaschke* 2010].

Traditional aerial imagery can provide an abundant amount of structure, intensity, colors, and texture information. However, it is difficult to recognize objects from aerial imagery due to the complexity of image interpretation. Thus, the complementary informational content of lidar point clouds and aerial imagery contribute to urban object classification. Based on a combination of lidar point cloud and aerial image data, several researchers proposed a number of methods for classification. *Haala and Walter* [1999] integrated the height information of an additional channel, together with the spectral channels into a pixel-based classification scheme. *Gamba and Houshmand* [2002] even combined lidar data, aerial imagery with SAR for land cover extraction, DTM and 3D shapes of buildings. *Hodgson et al*. [2003] pointed out that the rule-based segment classifier produced a slightly superior product than pixel-based classifier for mapping urban parcel imperviousness using lidar data and color aerial photography. *Charaniya et al*. [2004] described a supervised classification technique that classify lidar data into four classes-road, meadow, building and tree-by combining height texture, multi-return information and spectral feature of aerial images. *Brennan and Webster* [2006] presented a rule-based object-oriented classification approach to classifying surfaces derived from DSM, intensity, multiple returns, and normalized height [*Tiede et al*. 2008]. They applied to one-level-representation (OLR) concept for automated tree crown delineation in high-resolution digital camera imagery. A segmentation algorithm was used to create an initial set of image regions that are characterized by homogenous spectral information and height information. Then, these regions were treated independently to perform domain-specific class modeling using Cognition Network Language (CNL). *Zhou and Troy* [2008] presented an object-oriented approach for analyzing and characterizing the urban landscape structure at the parcel level for the Baltimore area in the United States. Additional spatial datasets including property parcel boundaries and building footprints were used to both facilitate object segmentation and obtain greater classification accuracy. *Aubrecht et al.* [2008] analyzed land-cover and urban function types on the basis of their relative heights and integrated socioeconomic data. By means of spatial disaggregation, raster population data is distributed to potential residential buildings. *Germaine and Hung* [2010] proposed a two-step classification method to delineate impervious surface in an urban area using lidar data to refine a base classification result of multispectral imagery. Their experiment showed that the use of lidar data could improve the overall accuracy by 3%. Therefore, the implementation of lidar significantly enhances the classification of optical imagery both in terms of accuracy as well as automation.

A review of current literature indicates that spectral information acquired by a remote sensor may not be sufficient to derive accurate information. Lidar point clouds may assist in achieving more accurate land-cover classification of remotely sensed imagery. *Rottensteiner et al.* [2007] demonstrated that the classification accuracy of a small residential area can be improved by 20% when fusing airborne lidar point cloud with multispectral imagery.

The advantage of the object-oriented approach is the ability to use multiple segmentation methods that are based on spatial and spectral properties of lidar point clouds and color image data, respectively. In a typical data fusion and classification method, a set of features is extracted from the two data sources and a number of class hypotheses are defined. Then, classification features, according to certain classification rules, are combined to form a feature space to decide to which of classes object belongs [*Li and Guan* 2010]. In the following section, we detail the proposed method of urban land-use classification using the spectral information from aerial imagery and spatial information from airborne lidar point clouds. Then results obtained from two urban scenes and classification accuracy assessment are reported and discussed. Finally, conclusions are drawn in the last section.

2. Methods

Figure 1 illustrates a workflow of the major components and their subsequent processes in the object-oriented classification system proposed in this study. Lidar point clouds with first- and last-return and a colour ortho-image covering the same area are used as the input. Spatial registration of lidar point clouds and aerial imagery is performed as data preprocessing, which is not discussed in this paper.

The output of a lidar mapping system is a cloud of irregularly spaced 3D points which include not only the bare ground, but also all kinds of objects (buildings, cars, trees, etc.). Separation of individual buildings from lidar point clouds is the key to

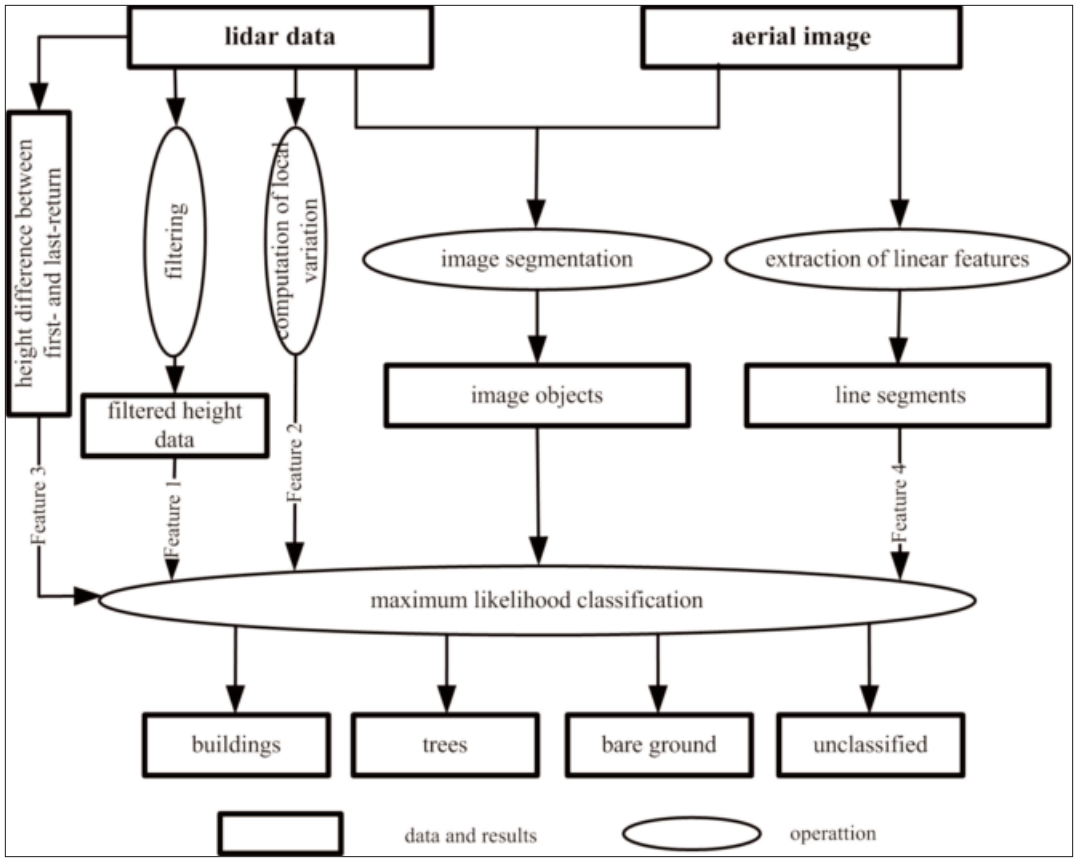


Figure 1: A workflow of the proposed object-oriented classification method

an accurate urban land-use classification. Like buildings, trees are also one of the dominant features in urban areas. Thus, a new object-oriented supervised classification method is proposed to detect individual buildings and differentiate trees from lidar point clouds, which is mainly built on image segmentation. The result of this image segmentation is the creation of a set of image objects defined as individual areas with shape and spectral homogeneity using additional height and texture information from lidar point clouds. These segmented homogeneous regions are—recognized as meaningful objects (e.g., building, trees and bare ground) that are then analyzed using the Bayesian classifier. In this study, four types of features that were used as additional channels in the classification are (1) the filtered height data obtained from lidar point clouds using the adopted progressive triangulated irregular network (TIN); (2) height difference between first- and last-returns of the laser pulse; (3) eigenvalue-based local variation of lidar points by computing dispersion matrix; (4) line segments extracted from the color aerial image. These four types of features are comprehensively utilized to determine each homogeneous region by the Bayesian decision rule as to which class it belongs.

2.1 Region-Based Segmentation

Segmentation, a process of partitioning an image space into some non-overlapping meaningful homogeneous regions (polygons) is crucial to the classification result. Segmentation of color aerial imagery contributes to the quality of classification because they can provide more additional information than gray level images. On the other hand, although the detailed spectrum information of color aerial imagery may lead to an increase of segmentation difficulty to some extent, it can be lessened based on the height and textural information from lidar data as additional channels. The region-based approach, comparing with edge-based segmentation, works generally better on noisy images, where borders are extremely difficult to detect. Watersheds transform, a region based algorithm, is a powerful tool for image segmentation. The concept of Watershed and catchment basin are well-known in topography. Image data can be interpreted as a topographic surface where the gradient image grey-levels represent the altitudes [*Sonka et al*. 2002]. Thus, Watershed transform algorithm can be used to denote a labeling of the image, such that all points of a given catchment basin have the same unique label *[Roerdink and Meijster*

2000]. In this paper, watershed segmentation is used to generate closed contours for each region in the original image due to it effectively dividing individual catchment basins in a gradient image.

2.2 Four Classification Features

In this section, four classification features, which are extracted from both the lidar point cloud and the aerial image data, are discussed respectively. These four features are filtered height data, height information, local variation calculated from lidar data, and line segment extracted from the aerial image.

Filtered Height Data

Before lidar point clouds can be further used, it has to be preprocessed which filtering is commonly used to separate on-terrain points and off-terrain points [*Kraus and Pfeife* 1998]. Commonly used filtering algorithms include morphological filtering, progressive TIN densification, and robust filtering. An extensive overview of different filtering approaches can be found in [*Sithole and Vosselman* 2004]. In this study, we adopted the progressive TIN densification that was firstly proposed by *Axelsson* [2000]. This algorithm works well when modeling surfaces with discontinuities, which is common in urban areas.

Eigenvalue-Based Local Variation

An eigenvalue-based local variation of lidar data is defined as the spatial feature of each point by calculating a dispersion matrix of its neighbors. It is another indicator for distinguishing tree points from the other features. Eigenvalues of a point's dispersion matrix can reflect the spatial information of this point. For each point v_j under consideration, its neighborhood points can be found by the KD-tree. A 3×3 dispersion matrix S_j of point v_j is given by:

$$S_j = \sum_{i=0}^{n} (P_i - M)^T (P_i - M) \ \forall (j = 0,1,2 \ldots N-1) \quad (1)$$

where S_j is the dispersion matrix of point v_j, n is the number of neighborhood points of point v_j, P_i is the coordinate (x_j, y_j, z_j) of point v_j, N is the total number of the lidar points, M is the mean matrix of its neighborhood points.

In this 3×3 dispersion matrix, each point v_j has three eigenvalues. An eigenvalue can be used to represent the spatial information of a lidar point because it is a scalar value associated with an eigenvector which reflects spatial distribution of a lidar point. Three possibilities are considered: (1) If one of the three eigenvalues is much larger than the others, its lidar point is labeled as an "edge" point. (2) If two of the three eigenvalues are much larger than the other, its lidar point is labeled as a "plane" point. (3) If the three eigenvalues are larger than a given threshold, the lidar point is labeled as a "discrete" point. In general, the trees show the property of divergence, while the ground and buildings exhibit the local planarity.

Height Difference Between First- and Last-Return

Height difference between first- and last- return usually differentiates tree features from lidar data. One of lidar system's characteristics is the capability of laser beam to penetrate the trees canopy through a small opening. The number of returns counts on the object within the travel path of the laser pulse. Many commercial lidar systems can measure multiple returns. Although the laser beam can penetrate the trees canopy to the ground, the height difference between first- and last-return alone is unreliable to distinguish the tree from the lidar point clouds. This is because the laser beam hitting on the edge of building also generates two returns. Secondly, if the density of trees is high, the small-footprint lidar cannot penetrate the tree's canopy. It is necessary to combine other features from lidar data and imagery to distinguish trees and buildings.

Line Segments

Generally speaking, a typical building consists of the regular geometric primitives (e.g., lines and corners of rooftops). Given the fact that geometric features (e.g., lines of building rooftops) in color aerial imagery are easier to be interpreted than those in lidar point clouds because the aerial images possess sharp buildings edges and have a generally higher horizontal accuracy, the Canny edge detector [*Canny* 1986] and 2D Hough transform [*Sonka et al.* 2002] were used to extract the line segments from the color aerial image in this study. These line segments were used as one of the additional channels in an object-oriented maximum likelihood classifier.

2.3 Supervised Classification

Compared with an unsupervised classification approach, a supervised classification approach is preferred by most researchers because it generally gives more accurate class definitions and higher accuracy. Given the class w_i $(i = 1,2 \cdots n)$, where n is the number of the classes. Classification features mentioned in previous section, yield a D dimensions feature space F. So, the probability of a region represented by its feature vector $X(X \in F)$ belongs to class w_i, is defined by the Bayesian decision rule:

$$P(w_i / X) = \frac{P(X / w_i) \bullet P(w_i)}{P(X)} \quad (2)$$

where $P(w_i)$ is the prior probability of class w_i. $P(X / w_i)$ is the conditional probability of class w_i has data X. $P(w_i / X)$ is the posterior probability of data X belonging to class w_i. $P(X)$ can be considered as constant value for class w_i. Therefore, Equation 2 can be reduced to:

$$P(w_i / X) = P(X / w_i) \bullet P(w_i) \quad (3)$$

where, i =1, 2 and 3, defining three distinctive classes, named building w_1, bare ground w_2, tree w_3. The objects that do not belong to these three classes are labeled as unclassified ones w_4. The prior probabilities of building $P(w_1)$, bare ground $P(w_2)$, tree $P(w_3)$ and unclassified objects $P(w_4)$ are obtained in accordance with specific training data set that can represent the typical features of urban areas, and meet the following equation :

$$P(w_1) + P(w_2) + P(w_3) + P(w_4) = 1 \quad (4)$$

The next step is to quantify features before determining conditional probability of each class.

(1) Filtered height data by adopted progressive TIN $X_1(f)$ is the ratio of on-terrain points to points in the homogenous region. $1 - X_1(f)$ implies the percent of off-terrain points in region.
(2) Bayesian of a lidar point using eigenanalysis $X_2(f)$ is the ratio of "scatter" and "edge" points to points in the homogenous region. $1 - X_2(f)$ represents the percent of 'plane' points in the region.
(3) Height difference between first- and last-return $X_3(f)$ is the ratio of points where the height information is over the given threshold, to the points in the homogenous region.
(4) Line segments extracted from aerial image $X_4(f)$ are the ratio of the length of line segments near to the region to the length of region boundary.

The conditional probability of classis w_i ascertained by the choices of the weighs (m_1, m_2, m_3, m_4) to features $(X_1(f), X_2(f), X_3(f), X_4(f))$:

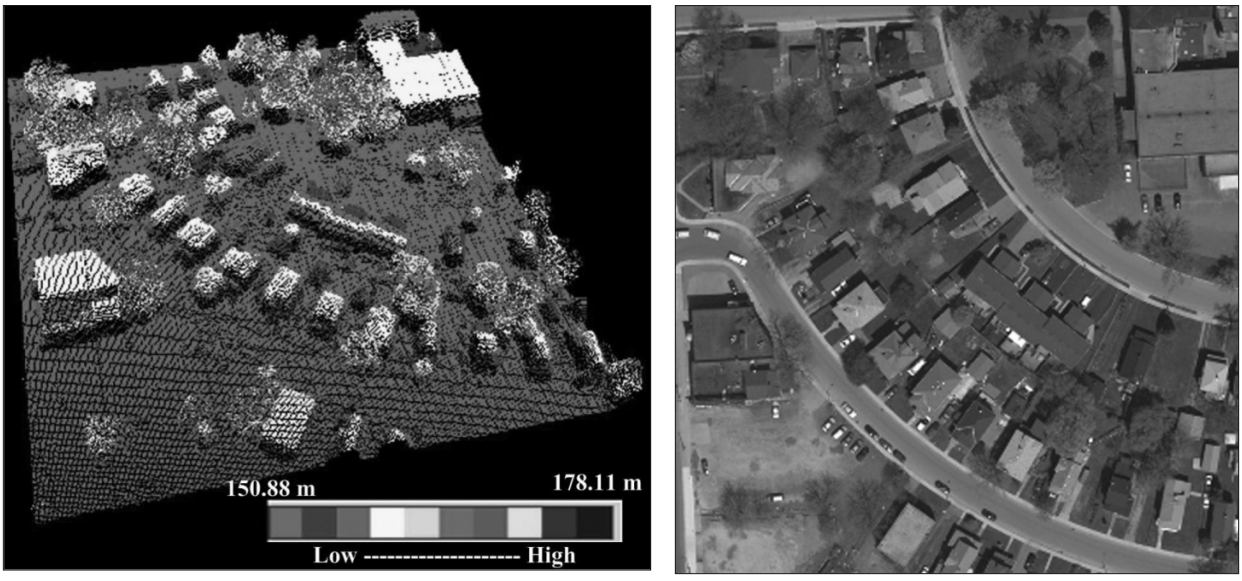
$$P(X / W_i) = m_1 \text{ x } X_1(f) + m \text{ x } X_2(f) + m_3 \text{ x } X_3(f) + m_4 \text{ x } X_4(f) \quad (5)$$

For each region, we can get three class values $\{d_1(X), d_2(X), d_3(X)\}$ according to Equation 3. The maximum of three results is $d_i(X) = \max\{d_1(X), d_2(X), d_3(X)\}$, which labels the region to the class it belongs. If $d_1(X) \approx d_2(X) \approx d_3(X)$, the region is then temporarily labeled as an unclassified class.

3. Results and Discussion

The lidar point cloud data covering two different urban areas (Scenes I and II) in the City of Toronto, Ontario, acquired by Optech ALTM 3100 system were used in this study. The lidar dataset consists of the first and last returns of the laser beam. The true colour image data used were taken by an onboard 4k × 4k digital camera simultaneously. Two different scenes covered by the same dataset were selected to evaluate the performance of the proposed object-oriented classification approach.

Figure 2 shows Scene I, in which (a) shows a raster DSM, containing a total of 105,298 points, which was interpolated with both the first and the



(a) (b)

Figure 2 Scene I: (a) DSM, (b) Colour aerial image.

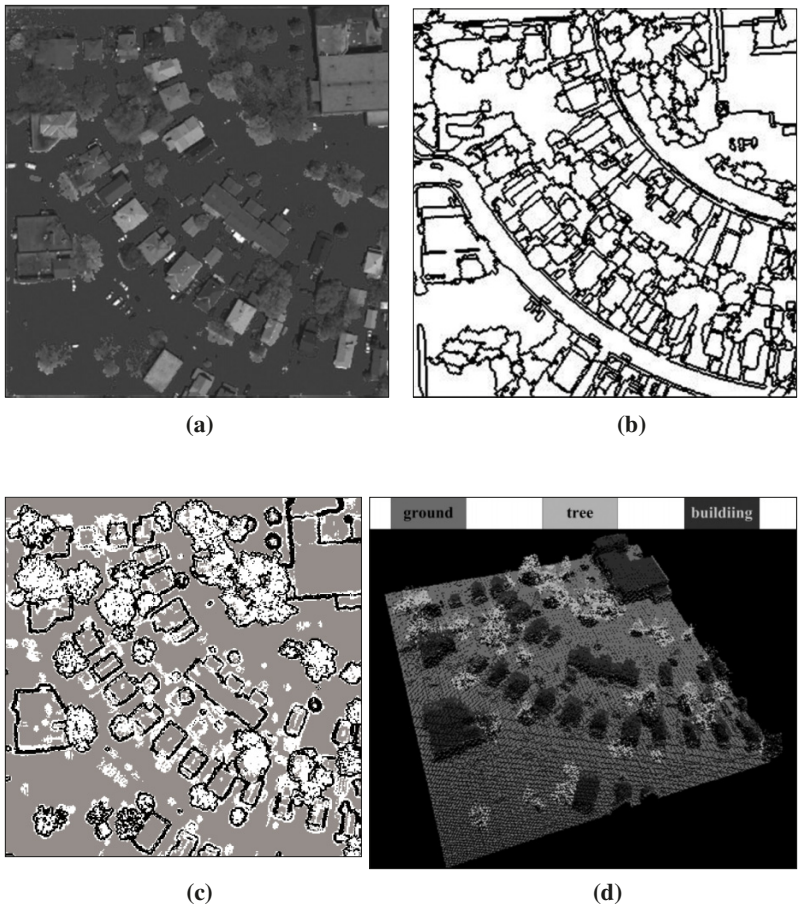


Figure 3: Results of Scene I: (a) filtered height data, (b) segmentation result by fusing aerial image with height and texture information from lidar point clouds, (c) local variation by eigenanalysis, and (d) results of object-oriented classification.

last pulse returns by the bi-linear interpolation method. The width and height of the grid equals to the ground sample distance (GSD) of the aerial image (0.5 m). The elevation of the study area ranges from 150.00 m to 178.11 m. Besides buildings, several clusters of trees located along the street; (b) shows a true color aerial image that was re-sampled to 0.5 m ground pixel. The majority of buildings that appear in the colour image are with gable roofs or hip roofs.

In this paper, classification accuracy relies heavily on the quality of segmentation. Although the abundant spectral information from aerial imagery is beneficial to classification procedure, it increases the difficulty of segmentation, because there is much spectral confusion between-class and spectral variation within-class. For example, the roof material of a building along a road sometimes is similar to the roads; dead trees and cars adjacent to a building have similar spectral characteristics to buildings. If the materials used on buildings and grounds were similar, they might be partitioned into a homogenous region. The quality of segmentation cannot be guaranteed from the range image when it is used alone. For instance, if a tree is close to a building, it is very difficult to segment one from another. The texture and height information obtained from the lidar data as additional channels can improve the quality of colour image segmentation.

Figure 3(a) shows filtered height data overlaid on the color aerial image. A close visual inspection shows that off-terrain points have been removed very well. The threshold of the iterative distance is usually below the height of car, and it gradually decreases with the increase of iterative times. Figure 3(b) shows the result of segmentation by fusing lidar point clouds and the aerial image. Figure 3(c) represents the local variation results of lidar point clouds by eigenanalysis. As shown in Figure 3(c) regions with abundant vegetation are described to have 'discrete' and 'edge' characteristics whereas regions with many buildings are said to have as 'plane' characteristics.

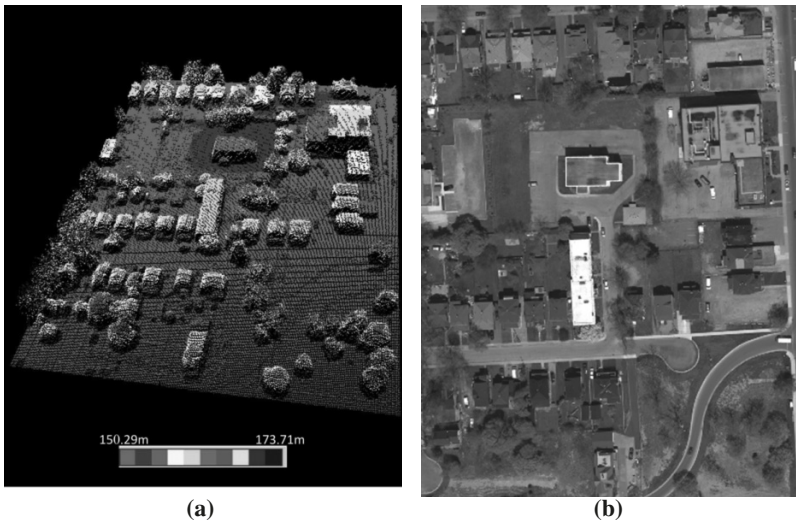
Some parameters for supervised object-oriented classification need to be pre-determined. First, the parameters of prior probabilities can be derived from the typical training set. The most important parameters in the study are the choices of weights to features. The best way to choose weights is self-adaptive to features of each homogenous region. Since features ($X_1(f)$, $X_2(f)$, $X_3(f)$, $X_4(f)$) have already been normalized, the weights used for each feature is its value:

$$m_1 = X_i(f) / (X_1(f) + X_2(f) + X_3(f) + X_4(f)) \; i = 1,2,3,4 \quad (6)$$

For each homogeneous region, the weights can change depending on the features in the region. For example, if the values of four features are {0.8, 0.8, 0.2, 0.1} in a region, the correspondent weights are {0.42, 0.42, 0.1, 0.06} respectively. Figure 3(d) shows the results of classification. By visual inspection with the aerial image, a few buildings cannot be differentiated from the trees, owing to trees with a large crown covering the major part of buildings. Spectral characteristics of some trees are similar to the material of buildings covered, and whose heights almost are near to that of these buildings. Therefore, these trees and buildings cannot be segmented correctly, which leads to the occurrence of classification errors.

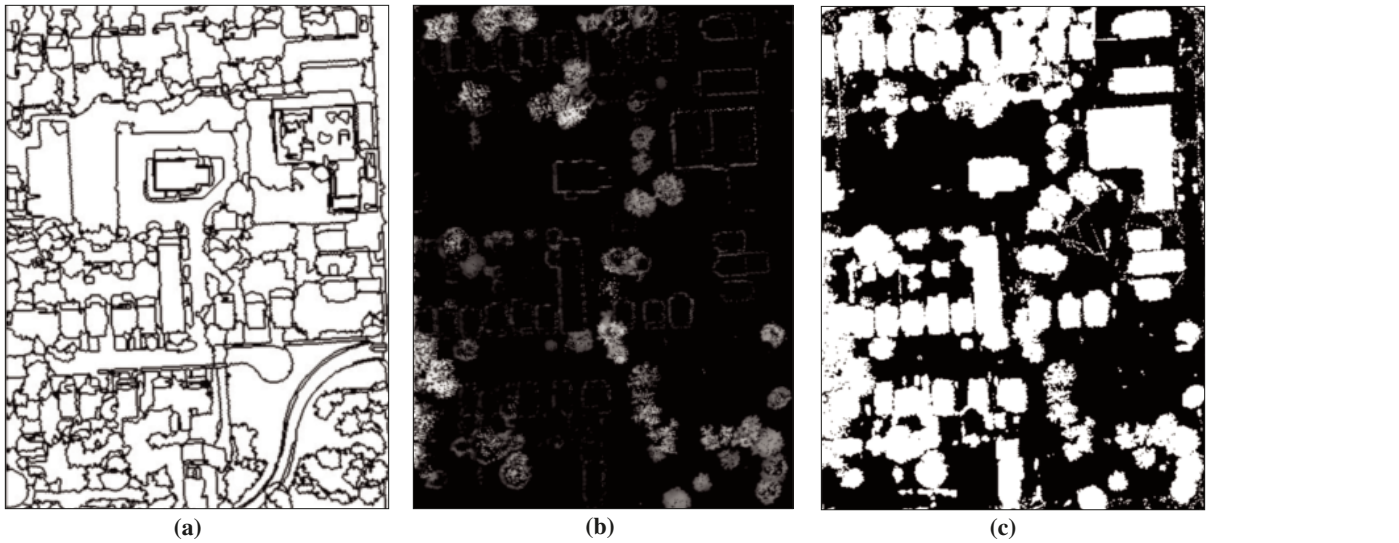
Figure 4 shows Scene II, in which (a) shows a raster DSM, containing a total of 166,495 points, which was interpolated using the same method as that covering the first study site; (b) shows the true color aerial image that was re-sampled to 0.5 m ground pixel.

Figure 5(a) shows the results of segmentation by fusing lidar point clouds and an aerial image. Figure 5(b) represents the height information of first- and last-return. Figure 5(c) illustrates the filtered height data of the adopted progressive TIN method, in where white pixels represent off-terrain objects, while black ones represent on-terrain objects. A visual inspection indicates that all off-terrain points are separated from lidar point clouds and on-terrain features are retained quite well. Figure 5(d) indicates the spatial local variation by eigenanalysis. As shown in Figure 5(d), building regions are basically exhibiting "plane" features, and tree regions are exhibiting "discrete" and "edge" features. However, a few tree regions consist of "plane" features because of its high density; Figure 5(e) shows the line segments extracted by using the Canny edge detector and 2D Hough transform.

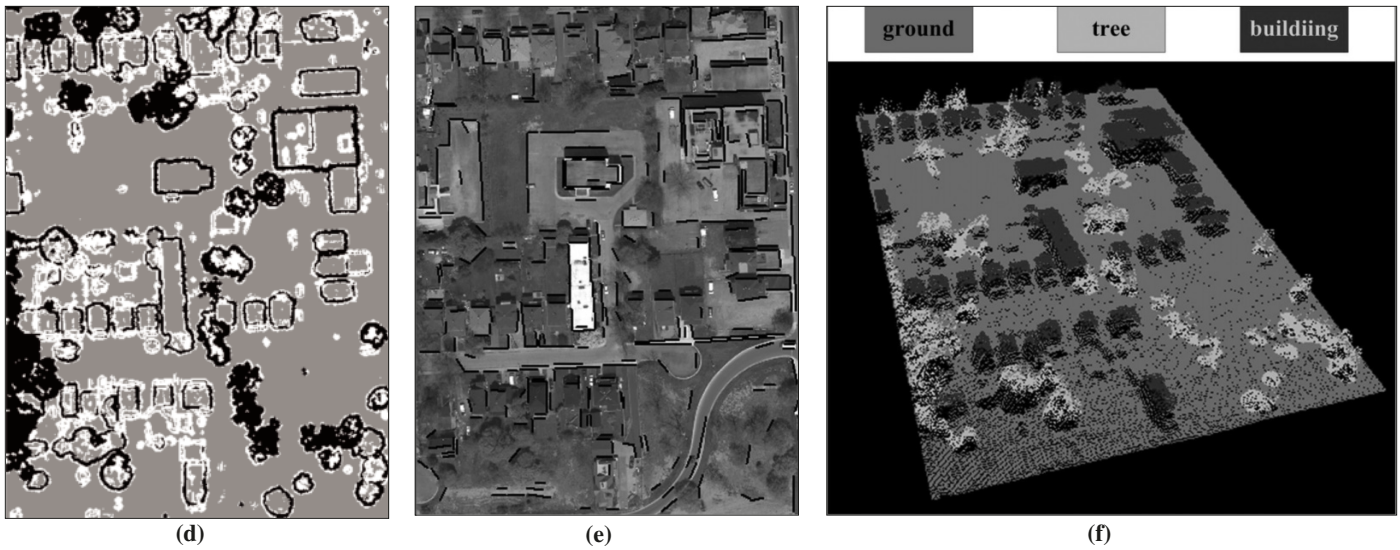


(a) (b)

Figure 4: Scene II: (a) DSM; and (b) Colour aerial image.



(a) (b) (c)



(d) (e) (f)

Figure 5: Results of Scene II: (a) segmentation result by fusing aerial image with height and texture information from lidar point clouds; (b) height information between first- and last-return; (c) filtered lidar data by adopted progressive TIN; (d) local variation by eigenanalysis; (e) line segments extracted by the Canny edge detector and 2D Hough Transform; and (f) the results of our classification method.

Parameters of prior probabilities for supervised object-oriented classification are equal to those of Scene I. That is, the prior probabilities of building $P(w_1)$, bare ground $P(w_2)$, tree, $P(w_3)$ and unclassified objects $P(w_4)$ are 0.4, 0.35, 0.2, and 0.05, respectively. Through self-adaptive weight computation, the weights used for buildings, grounds and trees are (0.12, 0.33, 0.16, 0.39), (0.34, 0.39, 0.23, 0.06), and (0.29, 0.32, 0.30, 0.09), respectively. As shown in Figure 5(f), the classification result of Scene II is satisfactory.

To evaluate the overall performance of our classification method, we utilized Definiens Professional Earth commercial software (also called eCognition) to produce the reference data for comparing with the classified results pixel-by-pixel. 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_{ij} 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), and overall accuracy (OA) [*Story and Congalton* 1986] and kappa coefficient [*Congalton et al.* 1983; *Congalton and Green* 2009] can be calculated.

Table 1 summarizes the confusion matrix of Scene I. The overall accuracy obtained is 84.96%. The classes "bare ground" and "building" demonstrate higher accuracy. Although "tree" has specific spectral characteristics (height and spatial informa-

Table 1: Confusion matrix of Scene I.

		Reference					
Segmentation-based classification	Class	Building	Bare ground	Tree	Unclassified	Σ	UA
	Building	***21169***	2016	2318	37	25540	82.88%
	Bare ground	4015	***53016***	3083	672	60786	87.21%
	Tree	1097	1894	***14052***	91	17134	82.01%
	Unclassified	56	477	79	***1226***	1838	66.71%
	Σ	26337	57403	19532	2026	105298	
	PA	80.37%	92.35%	71.94%	60.51%		
	Overall accuracy = (21169+53016+14052+1226)/105298 = 84.96%						
	Kappa coefficient = 75.68						

Table 2: Classification error of buildings in Scene II.

Building reference points(#)	Objected-oriented classification					
	Building points(#)	Real building points(#)	Omission points(#)	Omission error (%)	Commission points(#)	Commission error (%)
24137	22410	21904	2233	9.25	506	2.10

tion) that makes it easily distinguishable from "building" and "bare ground", many trees are still misclassified as "bare ground" due to their smaller height values, or as "building" due to their higher canopy density and their surroundings which consists of mainly buildings. Thus, it is inevitable that there are some classification errors in reference data; even these errors were carefully removed manually from the classification results obtained using eCognition. Besides the errors of the reference data, the use of high-difference, as one of classification features, is another factor leading to the classification errors between buildings and trees. Full waveform digitalization techniques have been developed to detect trees, including tree types and structures. Also, it is useful for delivering the information on low vegetation that is hard to detect in the discrete pulse system owing to the finite length of laser pulses.

In addition, the partial classification errors are caused by shadow-effects. Although shadow-effect usually is used to extract geospatial information, it also brings problems on the extraction of geometric and semantic information from remotely sensed data, particularly in the urban areas. Due to the presence of shadow-effect, it is difficult to accurately judge shadowing areas, especially near to high-rise buildings, which leads to the increase of the classification errors. However, lidar, on the other hand, can be used as one of solutions to shadow-effect correction by considering building heights, the sun elevation angle and azimuth [Sohn and Yun 2008].

Classification accuracy of the tree reference data generated using eCognition may be unreliable due to the complexity of landscape, especially characteristics of spectrum diversity and dispersion of spatial information. However, buildings are easier to identify and have more reliable classification accuracy as reference data. Therefore, in order to avoid and eliminate negative influence of reference data errors, classification accuracy of Scene II can be analyzed by comparison of building classification. In this study, accuracy analyses include two aspects: the number of the detected individual buildings and the number of the detected building rooftop points. There are 38 individual buildings in Scene II, while we detected 35 individual buildings. That is, the ratio of building detection is 92.11%. As shown in Table 2, omission error, the ratio of the building points misclassified as the non-building points to the classified building points, is (24137-21904)/24137=2.10%; commission error, the ratio of the number of the non-building points that are mislabeled as the building points to the classified building points, is (22410-21904)/24137=9.25%.

Conclusions

In this work, we develop a region-based classification method on the use of the fusing airborne lidar point cloud and colour aerial imagery. Since it is difficult to segment solely from, either aerial imagery or lidar point clouds, geometric information extracted from lidar point clouds as additional bands is incorporated with traditional spectral information to improve segmentation quality. Likewise, each segmented homogenous region is identified by comprehensively analyzing multiple features, including height difference, eigenvalue-based local variation, filtered height data and line segments.

The classification results of Scene I demonstrate that our method can successfully classify several dominant features in the urban areas. The overall accuracy and kappa statistics of Scene I classification are 84.96% and 0.76, respectively. The "building" and "bare ground" classes have 80.37% and 92.3% in the producer's accuracy, while 82.88% and 87.21% in the user's accuracy, respectively. To eliminate the influence of the reference data errors, only buildings were used as reference data to evaluate the classification accuracy in Scene II. Our classification results show that the building omission and commission errors are 2.10% and 9.25%.

In the future work, we will improve the classification automation by adaptively selecting classification parameters involved, and we will make use of the radiomtrically corrected intensity information to decrease classification errors of trees. Particularly, those trees are close to or overlapping with buildings. The development of the full waveform analysis methods provides another way to effectively highlight vegetation. Therefore, although multiple features can be extracted from laser scanning data and imagery, it is necessary to further research how to combine these features to improve the classification accuracy.

References

Aubrecht, C., K. Steinnocher, M. Hollaus, and W. Wagner. 2008. Integrating earth observation and GIScience for high resolution spatial and functional modeling of urban land use, *Computers, Environment and Urban Systems*, 33 (1): 15-25.

Awrangjeb, M., M. Ravanbakhsh and C.S. Fraser. 2010. Automatic detection of residential buildings using lidar data and multispectral imagery, *ISPRS Journal of Photogrammetry and Remote Sensing*, 65(5): 457-467.

Axelsson, P. 2000. DEM generation from laser scanner data using adaptive TIN models, *International*

Archives of the Photogrammetry and Remote Sensing, 33(B4/1): 110-117.

Bartels, M., and H. Wei. 2006. Maximum likelihood classification of lidar data incorporating multiple coregistered bands. *Proceedings of 4th International Workshop on Pattern Recognition in Remote Sensing/18th International Conference on Pattern Recognition*, 20 August, Hong Kong, China, pp.17-20.

Blaschke, T. 2010. Object based image analysis for remote sensing, *ISPRS Journal of Photogrammetry and Remote Sensing*, 65 (1): 2-16.

Brattberg, O. and G. Tolt. 2008. Terrain classification using airborne lidar data and aerial imagery. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences,* 37 (B3b): 261-266.

Brennan, R. and T. L. Webster. 2006. Object-oriented land cover classification of lidar derived surfaces. *Canadian Journal of Remote Sensing*, 32 (2): 162-172.

Brunn, A. and U. Weidner. 1998. Hierarchical Bayesian nets for building extraction using dense digital surface models, *ISPRS Journal of Photogrammetry and Remote Sensing*, 53(5):296–307.

Canny, J. 1986. A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 8(6): 679-698.

Charaniya, A.P., R. Manduchi, M. Roberto and S.K. Lodha. 2004. Supervised parametric classification of aerial lidar data, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, June 27-July 2, Baltimore, MD. Vol. 3, pp. 25-32.

Chehata, N., L. Guo and M. Clement. 2009. Contribution of airborne full-waveform lidar and image data for urban scene classification. *Proceedings of the 16th IEEE International Conference on Image Processing*, November 7-11, Cairo, Egypt, pp.1669-1672

Collins, C. A., R. C. Parker, and D.L. Evans. 2004. Using multispectral imagery and multi-return lidar to estimate trees and stand attributes in a southern bottomland Hardwood forest. *Proceedings of 2004 ASPRS Annual Conference*, May 23-28, Denver, Colorado, unpaginated CD-ROM.

Congalton, R.G. 1991. A review of assessing the accuracy of classifications of remotely sensed data, *Remote Sensing of Environment*, 37(1): 35-46.

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

Gamba, P. and B. Houshmand. 2002. Joint analysis of SAR, lidar and aerial imagery for simultaneous extraction of land cover, DTM and 3D shape of buildings, *International Journal of Remote Sensing,* 23(20): 4439–4450.

Germaine, K.A. and Ming-Chih Hung. 2011, Delineation of impervious surface from multispectral imagery and lidar incorporating knowledge based expert system rules. *Photgrammetric Engineering & Remote Sensing*, 77(1):77-87.

Haala, N., C. Brenner and K. Anders. 1998. 3D urban GIS from laser altimeter and 2D map data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 32(3): 339-346.

Haala, N. and C. Brenner. 1999. Extraction of buildings and trees in urban environments, *ISPRS Journal of Photogrammetry and Remote Sensing,* 54(2–3):130–137.

Haala, N. and V. Walter. 1999. Automatic classification of urban environments for database revision using lidar and color aerial imagery, *International Archives of Photogrammetry and Remote Sensing,* 32(7-4-3W6): 76-82.

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 & Remote Sensing*, 69(9):973-980.

Hu, Y. and C.V. Tao. 2005. Hierarchical recovery of digital terrain models from single and multiple return lidar data, *Photogrammetric Engineering & Remote Sensing,* 71 (4): 425–433.

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 & Remote Sensing*, 74(12):1473-1485.

Kraus, K. and N. Pfeifer. 1998. Determination of terrain models in wooded areas with airborne laser scanner data, *ISPRS Journal of Photogrammetry and Remote Sensing*, 53(4):193-203.

Li, J. and H. Guan. 2010. 3D building reconstruction from lidar point clouds fused with aerial imagery, *Urban Remote Sensing* (X. Yang, editor), Wiley-Blackwell, Chichester, UK, 75-92.

Lu, D., S. Hetrick, and E. Moran. 2010. Land cover classification in a complex urban-rural landscape with QuickBird imagery. *Photgrammetric Engineering & Remote Sensing*, 76(10):1159-1168.

Mass, H.-G. 1999. The potential of height texture measures for the segmentation of airborne laser scanner data, *Proceedings of 21st Canadian Symposium on Remote Sensing*, June 21-24, Ottawa, Ontario, Canada, Vol. 1, pp.154-161.

Roerdink, J.B.T.M. and A. Meijster. 2000. The watershed transform: definitions, algorithms and parallelization strategies, *Fundamenta Informaticae*, 41: 187-228.

Rottensteiner, F., J. Trinder, S. Clode and K. Kubik. 2003. Detection of buildings and roof segments by combining lidar data and multispectral images. *Image and Vision Computing*, November 26-28, Massey University, Palmerston North. New Zealand, pp. 60-65.

Rottensteiner, F., J. Trinder, S. Clode and K. Kubik. 2005. Using the Dempster–Shafer method for the fusion of lidar data and multispectral images for building detection. *Information Fusion*, 6(4): 283-300.

Rottensteiner, F., J. Trinder, S. Clode and K. Kubik. 2007. Building detection by fusion of airborne laser scanner data and multispectral images: Performance evaluation and sensitivity analysis,

ISPRS Journal of Photogrammetry and Remote Sensing, 62(2):135-149.

Sithole, G. and G. Vosselman. 2004. Experimental comparison of filtering algorithms for bare-earth extraction from airborne laser scanning point clouds, *ISPRS Journal of Photogrammetry and Remote Sensing*, 59(1-2): 85-101.

Sohn, H. and K. Yun. 2008. Shadow-effect correction in aerial color imagery, *Photogrammetric Engineering & Remote Sensing*, 74(5):611-618.

Sonka, M., V. Hlavac and R. Boyle. 2002. *Image Processing, Analysis, and Machine Vision*, 2nd edition. CL Engineering, Ludhiana, India, pp. 134-135.

Story M. and R.G. Congalton. 1986. Accuracy assessment: A user's perspective. *Photogrammetric Engineering & Remote Sensing*, 52 (3): 397–399.

Tiede, D., S. Lang and C. Hoffmann. 2008. Domain-specific class modeling for one-level representation of single trees, *Object-Based Image Analysis* (Blaschke, T., S. Lang, and G. Hay, editors). Springer, New York, pp. 133-151.

Thomas, N., C. Hendrix, and R.G. Congalton. 2003. A comparison of urban mapping methods using high-resolution digital imagery, *Photogrammetric Engineering & Remote Sensing*, 69(9): 963–972.

Vosselman, G. 2000. Slope based filtering of laser altimetry data. International *Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 33 (B3/2): 935-942.

Walter, V. 2005. Object-based classification of integrated multispectral and lidar data for change detection and quality control in urban areas. *Proceedings of URBAN 2005/ URS 2005,* 14-16 March, Tempe, AZ, unpaginated CD-ROM.

Zeng, Y., J. Zhang, G. Wang and Z. Lin. 2002. Urban land-use classification using integrated airborne laser scanning data and high resolution multispectral satellite imagery. *Proceedings of Pecora15/Land Satellite Information IV/ISPRS Commission I/FIEOS 2002 Conference*, 10-15 November, Denver, CO, unpaginated CD-ROM.

Zhou, W., and A. Troy. 2008. An object-oriented approach for analyzing and characterizing urban landscape at the parcel level, *International Journal of Remote Sensing*, 29 (11): 3119-3135.

MS Rec'd 11/11/05
Revised MS Rec'd 12/01/04

Authors

Dr. Haiyan Guan received her M.Sc. and Ph.D. degrees, both in Photogrammetry and Remote Sensing from Wuhan University, China in 2005 and 2009, respectively. She is currently a Ph.D. candidate in Geomatics at the Department of Geography and Environmental Management, University of Waterloo. Her research interests are airborne, terrestrial and mobile laser scanning data processing algorithms and 3D spatial modeling and reconstruction of critical infrastructure and landscape. She has co-authored more than a dozen of research papers published in refereed journals, books, and proceedings.

Dr. Jonathan Li is Professor of Geomatics and head of the Geospatial Technology and Remote Sensing (GeoSTARS) group in the Department of Geography & Environmental Management, University of Waterloo. He received his Ph.D. degree in Remote Sensing and GIS from the University of Cape Town, South Africa. Prof. Li has his current research focused on use of RADARSAT images for monitoring coastal and marine environments and use of mobile laser scanning point clouds for 3D surface modeling. He has published extensively in leading remote sensing journals. Prof. Li has received several prestigious awards, including Talbert Abrams Award, ESRI Best Paper Award and MDA Best Paper Award. He is also Adjunct Professor of York University in Canada and Guest Professor of several top-ranked universities in China. Dr. Li is currently Chair of ISPRS ICWG V/I on Land-based Mobile Mapping Systems (2008-2012), Vice Chair of ICA Commission on Mapping from Remote Sensor Imagery (2011-2015), Vice Chair of FIG Commission IV Hydrography as well as Chair of FIG WG4.4 on Maritime and Marine Spatial Information Management (2011-2014). He has been the remote sensing Editor of *GEOMATICA* since 2007.

Dr. Michael Chapman is Professor of Photogrammetry and Remote Sensing with the Geomatics Program at the Department of Civil Engineering, Ryerson University in Toronto. He received his Ph.D. degree in photogrammetry from Laval University, Quebec. Prior to joining Ryerson University, Dr. Chapman was Professor in the Department of Geomatics Engineering at the University of Calgary for 18 years. His current research interests include algorithms and processing methodologies for airborne sensors using GPS/INS; geometric processing of digital imagery in industrial environments; terrestrial imaging systems for transportation infrastructure mapping; algorithms and processing strategies for bio-metrology applications. He is author and co-author of more than 160 technical articles. Dr. Chapman was the photogrammetry editor of *GEOMATICA* before 2007 and he is currently Co-Chair of ISPRS Working Group V/3 on Image Sensor Technology (2008-2012). ❑