COMBINED MULTIPLE CLASSIFIED DATASETS CLASSIFICATION APPROACH FOR POINT CLOUD LIDAR DATA

N. El-Ashmawy^{1*}, A. Shaker²

¹ Survey Research Institute, National Water Research Center, Cairo, Egypt - nagwa.elashmawy@(gmail.com, ryerson.ca) ² Civil Engineering Department, Ryerson University, Toronto, Canada-ahmed.shaker@ryerson.ca

Commission III, WG III/5

KEY WORDS: LiDAR, Intensity Data, Classification, Land Cover, Combined Classifier, Point Cloud

ABSTRACT:

Airborne Laser scanners using the Light Detection And Ranging (LiDAR) technology is a powerful tool for 3D data acquisition that records the backscattered energy as well. LiDAR has been successfully used in various applications including 3D modelling, feature extraction, and land cover information extraction. Airborne LiDAR data are usually acquired from different flight trajectories producing data in different strips with significant overlapped areas. Combining these data is required to get benefit of the multiple strips' data that acquired from different trajectories. This paper introduces an approach called CMCD "Combined Multiple Classified Datasets" to maximize the benefits of the multiple LiDAR strips' data in land cover information. This approach relies on classifying each strip data then combining the results based on the *a posteriori* probability of each class of the classified data and the position of the classified points.

Two datasets from different overlapped areas are selected to test the proposed CMCD approach; both are captured from different flight trajectories. A comparison has been conducted between the CMCD results and the results of the common merging data approaches. The results indicated that the classification accuracy of the proposed CMCD approach has improved the classification accuracy of the merged data-layers by 6% and 10% for the two datasets.

1. INTRODUCTION

Airborne laser scanning systems, with Light Detection and Ranging (LiDAR) technology, have been successfully developed and rapidly used in various practical applications since the 1990s. LiDAR data have been intensively used in 3D city modelling, building extraction and recognition, and digital terrain/surface models (DTM/DSM) (Ackermann, 1996, Baltsavias, 1999, Haala & Brenner, 1999, Song et al., 2002; Yan et al., 2015).

LiDAR systems can record the intensity of backscattered energy from the illuminated targets along with the range data. Most of the commercial LiDAR sensors, that acquire the ground surface, utilize laser signals that operate with near infrared (NIR) signals. At this region of the electromagnetic spectrum, high separability of spectral reflectance of various land cover materials can be observed. As a result, distinguishing different ground materials based on the values of LiDAR intensity data can be achieved (Charaniya et al., 2004, Hu et al., 2004, Bartels & Wei, 2006, Brennan & Webster, 2006, Antonarakis et al., 2008, Hui et al., 2008, Blaschke, 2010, El-Ashmawy et al., 2011, Yan et al., 2012; Yan et al., 2015).

Airborne LiDAR data are usually acquired from different flight trajectories in several strips. Sometimes same area is acquired several time within the same mission in different directions or even in different missions. Dataset of any area that is scanned several times with different acquisition characteristics (from different flight trajectories, different scanned angles, different heights, or different sensors) usually contains denser data, than single acquisition dataset. Combining data of different strips is required to get benefit of these multiple sources of data. Yet, these data have various characteristics and discrepancies in their intensity values due to changes in flying altitude, attitude, and sensor scanning angles (El-Ashmawy & Shaker, 2014 (a)). Radiometric correction of the intensity data helps to homogenize the recorded intensity data of the different datasets (Yan et al., 2012). In this research we call each group of data that has same acquisition characteristics (acquired from same flight line by same sensor) a data-layer. In LiDAR acquisition missions, the flight trajectories are planned with side lapped areas between the adjacent strips. The areas that are acquired several times are usually have denser data than those acquired once. To get benefit of these dense data in land cover mapping, including all available data-layers in the classification process is required.

The main goal of this research is to maximize the benefit of using LiDAR data, which are collected by different acquisition characteristics, in land cover information extraction. This goal is achieved by proposing the "Combined Multiple Classified Datasets" (CMCD) approach to combine data of several classified data-layers based on specific factors. These factors are the *a posteriori* probability of each classified point in each class and the proximity of the grid points to the original ones. The introduced approach, is a modified version of a previously proposed approach, which depends on the combined multiple classifiers that is introduced in the pattern recognition field (El-Ashmawy & Shaker, 2014 (b)). The main idea of the combined multiple classifiers is to rely on several decision-making schemas to improve the confidence of the decision made, by weighing various opinions and combining them through some thought process (Polikar, 2006; Yan & Shaker, 2011).

A comparison between the results of the new introduced approach and the results of the common combining approaches is another aim of this study. The common combining approaches used in this research are; one is merging the multiple data-layers before conducting the classification, and the other after classifying the separate data-layers.

2. PROPOSED CLASSIFICATION APPROACH

The workflow of the introduced classification approach is illustrated in Figure 1. It consists of three stages: data preparation, classification, and evaluation. The classification

stage has two sub-stages; classification of grid points stage and CMCD stage. Data preparation, classification of grid points, and evaluation are applicable for both, single and multiple datalayers, while the CMCD stage is applicable only for the multiple data-layers.



Figure 1: Workflow of the introduced approach for classification

2.1 Stage 1: Data Preparation

The data preparation stage includes; defining the distinguished classes within the area of interest, defining the grid space, and collecting the ground truth data to be used for validating the classification results.

- 1. **Defining Classes:** According to land cover types of the study area, M distinguished classes are defined as $\{v_{-}1, v_{-}2, ..., v_{-}M\}$.
- 2. **Defining Grid Points:** Since LiDAR footprints usually have irregular spatial distribution (i.e., not gridded) and gaps appear within the LiDAR points, to achieve land cover information for the entire study area, a grid space *S* filled with regular distributed grid points s_i with *h* distance apart is defined. s_i represents the *i*th point on the grid space *S*, where i = 1, 2, ..., N and *N* is the total number of points in the grid space *S*.
- 3. **Collecting Ground Validation Data:** For accuracy assessment, ground validation data are required. Therefore, well distributed reference points are randomly selected.

The outputs of this stage are ASCII files; a file for each datalayer, containing the data of grid points (point ID, x, y, and v_r) where, v_r is the classes of the reference points, and r = 1, 2, ..., M.

2.2 Stage 2: Classification

The classification stage consists of classification of grid points stage (Stage 2A) and the CMCD stage (Stage 2B).

2.2.1 Stage 2A: Classification of Grid Points

(1) Classifying the Original Points

In this step a classification algorithm is applied on the original LiDAR points of each data-layer D_d without resampling the points into a grid space to avoid any losses of details associated with resampling points into 2D grids as described in Bao et al., 2007; El-Ashmawy & Shaker, 2014 (b). Consequently, the original points of each data-layer are classified into the distinguished classes. For that a MATLAB code is developed to apply the selected classification algorithm on LiDAR point cloud data^{*}, based on specific attribute values of the LiDAR

^{*} The classification in this work depends only on the LiDAR data without camera images to investigate the applicability of using this approach without any external data (RGB images). That is because sometimes it is required to reduce the wait of the payloads, especially if drones are in demands. Since increasing the payloads either requires using larger drones, or reducing the acquisition time per mission.

data. In this research the considered attribute values for classification are the elevation values z and the intensity values I. The input of this code is an ASCII file containing (ID, x, y, z, I). The outputs are an ASCII file for each data-layer that contains for each point, (ID, x, y, z, I, v), where v is the assigned classes.

(2) Assigning Classes to the Grid Points

Depending on the classification results of the original LiDAR point cloud data, the grid points are assigned to the appropriate land cover classes. Figure 2 illustrates an example of assigning a grid point to the appropriate. First, the original classified point data are resampled to the generated grid space. The most frequent class of the classified point data within a square area of $h \times h$ is assigned to the resampled point that represented this area, where *h* is the spacing distance between grid points. Then, the grid points are assigned to the same classes as the resampled points they are coincide with. The rest of the grid points remain unclassified, as they are located within the gaps (between the footprints), Figure 2 (b).

To assign the unclassified grid points to the appropriate land cover, a method called Iterative Majority Moving Window (IMMW) is introduced. In this method, each unclassified point is assigned to the major class of the surrounding neighbour grid points in an iterative process starting from points that are adjacent to the classified ones. Where, at any iteration, the unclassified grid point that is adjacent to classified ones is assigned to the most frequent class of the eight neighbouring points (3x3 pixels window with the centre that coincides with the grid point to be classified). The iterative process continued until all grid points are assigned to one of the predefined classes, Figure 2(c).



Figure 2: Example of assigning classes to the unclassified grid points using the Iterative Majority Moving Window (IMMW) method a) the classified original points and the defined grid space, b) the resampled points to the grid space, c) the classified grid points after applying the IMMW

2.2.2 Stage 2B: Combined Multiple Classified Datasets

(1) Determination of the "*a posteriori*" Probabilities

The confusion matrix (*CM*) assessing method indicates the degree of support given, by the applied classifier, to each class (Xu et al., 1992). These values of the confusion matrix can be accepted as an estimation of the *a posteriori* probability for that class after scaling these values to the [0, 1] interval, where the summation of any row/column elements has to be 1 (Kuncheva et al., 2001; Poliker, 2006). The predefined reference points are

used to form the confusion matrices. There is a confusion matrix corresponding to each data-layer. The output of scaling the elements of the confusion matrix (CM) to the [0,1] interval is called a normalized confusion matrix (NCM).

(2) Determination of the Inverse Distance Weights

The Inverse Distance Weighting (IDW) method is used, as another factor affecting the results of the CMCD approach. IDW method defines the weight factors of each data-layer at each grid point as illustrated in Equation (1). This factor are used to weigh the effect of each data-layer on the class of the grid points.

$$w_1 = \frac{dis_2}{dis_1 + dis_2} \& w_2 = \frac{dis_1}{dis_1 + dis_2}$$
(1)

Where w_1 and w_2 are the weighting factors for the first and second data-layers, respectively, and dis_1 and dis_2 are the distances between the grid point into consideration and the nearest point of the first and second data-layers, respectively.

(3) Combining the Classified Data

With the knowledge of the normalized confusion matrix NCM_d for a classified data-layer d, the uncertainty in the class v_i assigned to the point p into consideration can be described by the conditional probabilities that $p \in v_i$, i = 1, 2, ..., M are true under the occurrence of the assigning event $l_k(p) = v_i$ (Xu et al., 1992), as shown in Equation (2).

$$P(p \in v_i | l_k(p) = v_j) = \frac{n_{ij}^{(d)}}{\sum_{i=1}^M n_{ij}^{(d)}},$$

$$i = 1, \dots, M\& j = 1, \dots, M$$
(2)

Where $n_{ij}^{(d)}$ is the number of reference points of class *j* that are assigned to class *i* in the classified data-layer *d*, and $\sum_{i=1}^{M} n_{ij}^{(d)}$ is the total number of reference points of class *j* (summation of each column of the *NCM*).

The normalized confusion matrix can be considered as prior knowledge of an expert. This expert has a belief value with uncertainty that the point p, in consideration, belongs to the class v_i (for all classes), which can be expressed in the form of the conditional probability as (Xu et al., 1992):

$$bel(p \in v_i | l_k(p), EN) = P(p \in v_i | l_k(p) = v_{j_k}), \quad (3)$$

i = 1, ..., M

Where, *EN* is the environment of the common classification environments that consist of independent events. Based on the Bayesian formula, when the data-layers are independent of each other, then the events $l_1(p) = v_{j_1} \dots l_k(p) = v_{j_k}$ will be independent of each other under either the condition of $p \in$ v_i or the environment *EN*, which leads to (as adjusted from Xu et al., 1992):

$$bel(v_i) = \eta \prod_{k=1}^{K} (p \in v_i | l_k(p) = v_{j_k})$$
(4)
Where η is a constant that ensures that $\sum_{i=1}^{M} bel(v_i) = 1$

With the data of two overlapped strips, each grid point s_i is assigned to two classes with assigning class events $l_d(s_i)$, where $l_1(s_i) = v_{k_1}$ and $l_2(s_i) = v_{k_2}$. To decide which class is finally assigned to the grid point after combining the classification results, the belief of each class has to be calculated. The grid point, then, is assigned to the class with maximum belief. The belief of each class can be calculated using the formula in Equation (4). This belief is based on the *a posteriori* probability of each class.

Based on the *a posteriori* probabilities of any class v_j , the belief in that class can be defined by the conditional joint probability that a point s_i belongs to that class and is true under the occurrence of the two assigning class events, l_1 and l_2 in the environment *EN*. This relation can be described by the Equation (5) (as adjusted from Xu et al., 1992):

 $bel pp(s_i \in v_j | l_1(s_i), l_2(s_i), EN) = P(s_i \in v_j | l_1(s_i) = vk1, l_2(s_i) = vk2), j = 1, ..., M$ (5)
Where,

 $bel_{pp}\left(s_i \in v_j \middle| l_1(s_i), l_2(s_i)\right)$

is the belief, based on the *a posteriori* probability, in class v_j assigned to point s_i is true with the occurrence of, l_1 and l_2 .

 $l_1(s_i), l_2(s_i)$

are the assigning class events, for data-layer 1 and data-layer 2.

For simplicity the bel_{pp} $(s_i \in v_j | l_1(s_i), l_2(s_i))$ will be denoted as bel_{pp} $(v_j | l_1, l_2)$, and $P(s_i \in v_j | l_1(s_i) = v_{k_1}, l_2(s_i) = v_{k_2})$ as $P(v_j | l_1, l_2)$. Therefore, Equation (3) can be expressed as:

$$bel_{pp}(v_j|l_1, l_2, EN) = P(v_j|l_1, l_2)$$
 (6)

Based on the Bayesian formula, the right-hand side term of Equation (6) can be described as follows (Xu et al., 1992):

$$P(v_j | l_1, l_2) = \frac{P(l_1, l_2 | v_j) P(v_j)}{P(l_1, l_2)}, \ j = 1, \dots, M$$
(7)

Where,

 $P(v_j)$ is the *a priori* probability of the classifier for each data-layer, however, the *a priori* probabilities are considered constant for all classes.

 $P(l_1, l_2)$ is the unconditional joint probability density.

Since the classifier is trained by independent training sets for each data-layer, the two-classified data-layers are considered independent; thus, the product rule can be applied for the joint probability case.

$$P(l_1, l_2 | v_j) = P(l_1 | v_j) . P(l_2 | v_j)$$
(8)

$$P(l_1, l_2 | v_j) = P(l_1 | v_j) \cdot P(l_2 | v_j) = \prod_{d=1}^{D} P(l_d | v_j)$$
(9)

Where $P(l_1 | v_j)$ and $P(l_2 | v_j)$ are the conditional probabilities of class v_j for the two data-layers, which can be estimated by the *a posteriori* probability. The *a posteriori* probability can be calculated by evaluating the classification result of the base classifier using the confusion matrix (Yan & Shaker, 2011).

The unconditional probability can be expressed in terms of the conditional probability as (Kittler *et al.*, 1998):

$$P(l_1, l_2) = \sum_{m=1}^{M} P(l_1, l_2 | v_m) P(v_m)$$
(10)

From Equations 5–to-10, the belief of class v_j based on the *a posteriori* probabilities of the classification results of the two data-layers can be expressed as:

$$bel_{pp}(v_j) = P(v_j|l_1, l_2) = P(v_j) \frac{(P(l_1|v_j).P(l_2|v_j))}{\sum_{m=1}^{M} P(v_m).P(l_1|v_m).P(l_2|v_m)}$$
(11)

$$bel_{pp}(v_j) = P(v_j) \frac{\prod_{d=1}^{D} P(l_d | v_j)}{\sum_{m=1}^{M} P(v_m) \prod_{d=1}^{D} P(l_d | v_m)}$$
(12)

To include the effect of the distance on the final classification decision, the weighting factors w_1 and w_2 can be multiplied by the *a posteriori* probabilities of the classification results in the calculation of the belief of each class.

$$bel(v_j) = P(v_j) \frac{w_1 P(l_1 | v_j) . w_2 P(l_2 | v_j)}{\sum_{m=1}^{M} P(v_m) . (P(l_1 | v_m) . P(l_2 | v_m))}$$
(13)

For each of the grid points, the believes of the available classes are compared to the point of interest, and the class with the maximum belief is assigned to that point. Since the *a priori* probabilities are considered constant for all classes, comparing the belief for each class is not affected by the *a priori* probabilities. Hence, the *a priori* probabilities can be omitted from Equation (13). Furthermore, the denominator in this formula is constant for all classes; so it will not affect the final combining decision. Therefore, only the numerator can be considered in the belief comparison as illustrated in Equation (14).

$$assign s_{i} \to v_{j} \quad if w_{1}P(l_{1}|v_{j}) . w_{2}P(l_{2}|v_{j}) = max^{M}_{m=1} w_{1}P(l_{1}|v_{m}) . w_{2}P(l_{2}|v_{m})$$
(14)

The belief value of each class based on the *a posteriori* probability of that class and the proximity of the original points is calculated for each grid point. The grid points may have different values for each layer based on the classification results of that layer and its accuracy assessment. Finally, for each grid point, the class that has the maximum belief value is assigned to that point.

2.3 Evaluation Stage

The third, and last, stage of the proposed classification approach workflow is the evaluation of the final classification results using the confusion matrix method. The ground validation points are used to form a confusion matrix. Two accuracy assessment values are calculated; the overall accuracy and the overall Kappa statistics.

3. STUDY AREA AND DATASETS

The study area covers part of British Columbia institute of technology, Canada. This area is scanned by a Leica ALS50 sensor, operating at 1.064 μ m wavelength, 0.33 mrad beam divergence and 83 kHz pulse repetition frequency. The acquired area covers a variety of land cover types including buildings, parking areas, trees, roads, and open spaces with and without grassy coverage, Figure 3. The LiDAR data were captured from different flight lines forming different strips with flying altitude of about 540 m (Habib et al., 2011)

The point density of the data is 4-5 points/square meter^{*}. Aerial images were captured during the same flight mission of the LiDAR data acquisition. The aerial images were geometrically

^{*} This point density is calculated based on the total number of points divided by the total area. Yet, because of the characteristics of the acquired surfaces (roads and roof materials), no reflectance occurred producing gaps between the point clouds data. In the areas without gaps the average point density is around 15 point/m².

corrected and ortho-rectified. Two different datasets from the same study area are selected to conduct the proposed classification approach. The clipped datasets are selected from overlapped areas between pairs of adjacent strips and cover around 270 m x 80 m, where the overlapped distance between the adjacent strips is around 80 m (width of the selected datasets) and the length of the datasets are selected to be with

reasonable number of points. The first dataset contains around 125,000 and 85,000 points of the two data-layers, and the second one contains 100,000 and 80,000 points in the two data-layers. Both datasets have complex land cover types relative to their sizes, where different types of roof surfaces and different ground elevations are found within the small areas.



Figure 3: Study areas (clipped from google map and google Earth)

4. RESULTS AND DISCUSSION

All the obtained LiDAR data are geometrically calibrated as described in Habib et al. (2011), and radiometrically corrected by eliminating the effects of the system characteristics, the objects geometry, and the atmospheric attenuation as described in Shaker et al. (2011). For improving the homogeneity of the intensity data, a histogram matching approach is applied to the overlapped areas between each adjacent pair of strips (Yan & Shaker, 2016).

According to land cover types of the study areas, four distinguished land cover classes are defined as; Buildings, Open areas (with and without grass), Roads, and Trees. A MATLAB code for point cloud classification using the Maximum Likelihood algorithm^{*} is developed to classify the point cloud LiDAR data based on the elevation and intensity values. Two grid spaces with 0.2 m spacing distance covering the selected areas are defined to overcome the irregularity of the spatial distribution (not gridded) of the LiDAR data footprints, and to fill the gaps within the LiDAR points. Two sets of well distributed reference points are randomly selected from the

regular grid points, where around 2000 reference points are selected in each dataset. The validation data of the reference points is collected from the ortho-rectified aerial images.

For each dataset, the developed MATLAB code is applied on the merged data-layers after applying the histogram matching process, and on each data-layer separately to merge the datalayers after classification. After that, the classified points are resampled into the defined grid space and the remaining unclassified grid points are assigned to the appropriate classes by following the iterative majority moving window (IMMW) method.

To test the CMCD approach for combining the multiple datalayers, a MATLAB code is developed to perform the CMCD approach on the classified data-layers. To determine the *a posteriori* probability the classification results of each datalayer is assessed based on the defined reference points using confusion matrices. Then, the normalized confusion matrices are calculated for each data-layer using a developed MATLAB code. An example of the confusion matrix and the normalized confusion matrix is illustrated in Figure 4. To calculate the IDW factors, the distances between each grid point and the nearest original points, of the two data-layers, are determined. The IDW for each point is calculated using the formula in Equation (1).

^{*} The Maximum Likelihood classifier is selected in this research work. However, any other classification algorithm can be used.

Based on the calculated factors, the *a posteriori* probability and the IDW, the belief of each grid point belonging to each class is calculated Then, the grid points are assigned to the classes that have maximum believes, Equation (14).

Figures 5 and 6 for the first and second datasets respectively illustrate the classification results of the grid points for the described three cases (merging approaches). Where, Figures 5a and 6a show the results of the merged data before classification. Figures 5b and 6b show the classification results of the merged classified data, where each data-layer is classified then the classified results are merged. Finally, Figures 5c and 6c show the classification results of the CMCD introduced approach.

Then the reference points are used to assess the classification results. Table 1 lists the overall accuracy and the overall Kappa statistics values for each case, for the two datasets.

		Open		
	Build	Areas	Road	Tree
Build	182	8	30	18
Open Areas	25	690	66	71
Road	61	134	537	138
Tree	42	33	49	440

Confusion matrix			
		Open	
	DI.d	A	D

	Build	Areas	Road	Tree
Build	0.7892	0.0263	0.1273	0.0571
Open Areas	0.0376	0.7873	0.0971	0.0781
Road	0.0773	0.1289	0.6659	0.1280
Tree	0.0961	0.0573	0.1097	0.7368
		-		

Normalized confusion matrix Figure 4: Example of confusion matrix and normalized confusion matrix for one of the data-layers





Figure 6: Classification results of the second dataset a) classification results of the merged data before classification, b) classification results of the merged classified data, c) classification results of the CMCD approach

Table 1: Accuracy assessment	of the classification results
------------------------------	-------------------------------

Casa	Dataset 1		Dataset2	
Case	Accuracy	Kappa	Accuracy	Kappa
Merged Original data	49%	0.37	62%	0.49
Merged Classified	49%	0.37	68%	0.56
CMCD Approach	55%	0.43	73%	0.62

By observing the classification results of the multiple datalayers of the overlapped areas, it can be noticed that merging the multiple data-layers after classifying each data-layer slightly improves the classification results. Moreover, combining the classified data by following the CMCD approach further improves the classification results of the overlapped areas. That is because the CMCD approach considers the *a posteriori* probabilities of each class and the distance between the original LiDAR points and the classified grid points. The misclassification points in the selected areas occur because of the large gaps between the original footprints and the similarity between the roof materials and the road covers.

5. SUMMARY AND CONCLUSIONS

This research aims at maximizing the benefits of the LiDAR data acquired by different characteristics for extracting land cover information. A new approach, called CMCD "Combined Multiple Classified Datasets", is introduced to combine the multiple LiDAR data-layers after classification. The CMCD combines the classified data-layers based on the a posteriori probability of each class of the classified data and the proximity of the grid points to the original acquired points. CMCD is three-stage classification approach; data preparation, classification, and evaluation. Where in the classification stage, the original point cloud data are classified, then the regular generated grid points are assigned to the appropriate class using a method called Iterative Majority Moving Window (IMMW), where the appropriate class is assigned to the unclassified grid point based on the majority classification of the surrounding points. After that the "*a posteriori* probability" is calculated for each class based on the assessment of the classification results, and the inverse distance weighting factor is calculated for each grid point based on the distance between the point and the surrounding original points. After that, the grid points are assigned to the classes that have maximum factors.

Two datasets in an urban study area are investigated; both are in the overlapped areas of pairs of adjacent strips, and contain four distinguished classes; buildings, open areas, roads, and trees. The proposed approach is compared to two other common combining data methods; merging the points into one layer and then classifying the merged data, and the other is merging the data of the classified data-layers.

For the first dataset, the accuracy achieved of the merged datalayers, is 49% for both common combining methods, and the achieved classification accuracy is 55% by applying the proposed CMCD approach.

For the second dataset, the merged classified data of the two data-layers is 62%, improved to 68% when the data-layers are merged after classification. By applying the CMCD approach the classification accuracy improved to 73%.

As a conclusion, following the proposed CMCD approach improves the classification results of the multiple data-layers by more than 5% without more significant expenses in time, nor money.

ACKNOWLEDGEMENTS

This research work is supported by the Discovery Grant from NSERC (the Natural Sciences and Engineering Research Council of Canada), and the GEOIDE Canadian Network of Excellence, Strategic Investment Initiative (SII) project SII P-IV # 72. The authors would like to thank McElhanney Consulting Services Ltd, BC, Canada for providing the LiDAR data and the aerial photos. Special thanks go to Prof. Tarek Abdel-Aziz, director of Survey Research Institute (SRI) for activating a fruitful cooperation between the researchers of SRI and the researchers of Ryerson University, Canada.

REFERENCES

Ackermann, F. (1996). Airborne laser scanning for elevation models. *Geomatics Info Magazine*, *10* (10), 24-25.

Antonarakis, A. S., Richards, K. S., & Brasington, J. (2008). Object-based land cover classification using airborne LiDAR. *Remote Sensing of Environment*,112(6), 2988-2998.

Baltsavias, E. P. (1999). Airborne laser scanning: existing systems and firms and other resources. *ISPRS Journal of Photogrammetry and Remote sensing*, 54(2), 164-198.

Bao, Y., Cao, C., Chang, C., Li, X., Chen, E., & Li, Z. (2007, September). Segmentation to the point clouds of LiDAR data based on change of kurtosis. In *International Symposium on Photoelectronic Detection and Imaging: Technology and Applications 2007* (pp. 66231N-66231N). International Society for Optics and Photonics.

Bartels, M., & Wei, H. (2006). Segmentation of LIDAR data using measures of distribution. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *36*(7), 426-31.

Blaschke, T. (2010). Object based image analysis for remote sensing. *ISPRS journal of photogrammetry and remote sensing*, 65(1), 2-16.

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

Charaniya, A. P., Manduchi, R., & Lodha, S. K. (2004, June). Supervised parametric classification of aerial LiDAR data. In Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshop, CVPRW'04. (pp. 25-32). IEEE.

El-Ashmawy, N., Shaker, A., & Yan, W. (2011). Pixel vs object-based image classification techniques for LiDAR intensity data. *ISPRS-International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 38/12, 43-48.

El-Ashmawy, N., & Shaker, A. (2014a). Raster vs. point cloud LiDAR data classification. *ISPRS-International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*,40/7, 79-83.

El-Ashmawy, N., & Shaker, A. (2014b). Innovated Approach for LiDAR Intensity Data Classification. *In Geoscience and Remote Sensing Symposium (IGARSS), 2014 IEEE International* (pp. 169-172). IEEE.

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

Habib, A. F., Kersting, A. P., Shaker, A., & Yan, W. Y. (2011). Geometric calibration and radiometric correction of LiDAR data and their impact on the quality of derived products. *Sensors*, *11*(9), 9069-9097.

Hu, X., Tao, C. V., & Hu, Y. (2004). Automatic Road Extraction from Dense Urban Area by Integrated Processing of High Resolution Imagery and LiDAR Data. *ISPRS-International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences.* Istanbul, Turkey, 35, B3.

Hui, L., Di, L., Xianfeng, H., & Deren, L. (2008, July). Laser intensity used in classification of LiDAR point cloud data. In *International Geoscience and Remote Sensing Symposium*, 2008. *IGARSS 2008. IEEE* (Vol. 2, pp. II-1140). IEEE.

Kuncheva, L. I., Bezdek, J. C., & Duin, R. P. (2001). Decision templates for multiple classifier fusion: an experimental comparison. *Pattern Recognition*, *34*(2), 299-314.

Polikar, R. (2006). Ensemble based systems in decision making. *Circuits and Systems Magazine, IEEE*, 6(3), 21-45.

Shaker, A., Yan, W. Y., & El-Ashmawy, N. (2011). The effects of laser reflection angle on radiometric correction of airborne LiDAR intensity data. *ISPRS-International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *38/12*, 213-217.

Song, J. H., Han, S. H., Yu, K. Y., & Kim, Y. I. (2002). Assessing the possibility of land-cover classification using lidar intensity data. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *34*(3/B), 259-262.

Xu, L., Krzyzak, A., & Suen, C. Y. (1992). Methods of combining multiple classifiers and their applications to handwriting recognition. *IEEE Transactions on Systems, Man and Cybernetics*, vol. 22, no. 3, pp. 418–435.

Yan, W. Y., & Shaker, A. (2011). The effects of combining classifiers with the same training statistics using Bayesian decision rules. *International Journal of Remote Sensing*, *32*(13), 3729-3745.

Yan, W. Y., & Shaker, A. (2016). Radiometric normalization of overlapping LiDAR intensity data for reduction of striping noise. *International Journal of Digital Earth*, 9(7), 649-661.

Yan, W. Y., Shaker, A., & El-Ashmawy, N. (2015). Urban land cover classification using airborne LiDAR data. *Remote Sensing of Environment, 158, 295-310*

Yan, W. Y., Shaker, A., Habib, A., & Kersting, A. P. (2012). Improving classification accuracy of airborne LiDAR intensity data by geometric calibration and radiometric correction. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, 35-44.