

SYNERGY BETWEEN AERIAL IMAGERY AND LOW DENSITY POINT CLOUD FOR AUTOMATED IMAGE CLASSIFICATION AND POINT CLOUD DENSIFICATION

Hani Mohammed Badawy^{a,*}, Adel Moussa^{a,b}, Naser El-Sheimy^a

^a Dept. of Geomatics Engineering, University of Calgary, Calgary, Canada (hmmohamm, amelsaye, elsheimy)@ucalgary.ca

^b Dept. of Electrical Engineering, Port-Said University, Port-Said, Egypt

Commission VI, WG VI/4

KEY WORDS: Classification, Automated, LIDAR, Aerial, image, Point cloud

ABSTRACT:

In this paper a synergy scheme between aerial imagery and sparse LIDAR point clouds is proposed for an automated aerial image classification. In this scheme, a point cloud and an image are chosen for a certain urban area. The point cloud is automatically classified into buildings, vegetation and roads using PCA and intensity variation. Afterwards, a projection of the point cloud into an image is obtained, such that it is registered with the aerial image. The aerial image classifier is trained with the LIDAR classification result to generate an automated classifier for aerial images. The classifier is tested with another image to demonstrate its accuracy. Another benefit of the synergy proposed is to densify the planar patches of the low density point cloud using the segmented aerial image to help modelling applications achieve more precise boundaries.

1. INTRODUCTION

The vast number of available aerial images that cover most of the world to date is a great tool for object classification in different environments, such as urban areas, forests, etc. Most of the supervised classifiers require a training session with a large number of datasets to enhance the classification accuracy. Therefore, supervised training of datasets is a tedious task that requires human intervention. On the other hand, unsupervised classification could be achieved using 3D point clouds. However; these data are relatively expensive and are rarely updated on a regular basis, such that for a certain area, available high resolution point cloud data are fewer than their corresponding aerial imagery. Therefore; the problem of laborious supervised training of image classifier along with the problem of low density of rarely available point clouds, can be tackled simultaneously if both data are fused together.

In this paper a synergy scheme between aerial images and airborne laser data, of urban areas, is proposed to handle the limitations of both types of data. The proposed scheme comprises the following steps. The first step is to classify a low density point cloud while, at the same time, segmentation of an airborne image covering the same scene is being performed. Point cloud is automatically classified using Principal Component Analysis (PCA), Digital Terrain Model (DTM) generation and intensity (LIDAR reflectance) analysis. Image segmentation is achieved using region growing segmentation algorithm.

The second step is to register the airborne image with a projected image of the point cloud. Consequently, each segment in the image can be assigned the classification label of the corresponding point cloud. This process is equivalent to an automated training process of the image classifier. By repeating this process with different scenes, the image classifier is being enhanced.

The third step is the boundary regularized densification of the point cloud. Densification of the point cloud can be accomplished by adding 3D points, to the point cloud, that follow the geometrical constraints of the corresponding image segment boundary.

Ultimately, such synergy helps to classify image scenes that are not associated with point clouds using the gained information from the synergy. At the same time, the low density point cloud can be densified

The paper is organized as follows: the rest of section (1) summarizes that related work and the paper contribution. Section (2) describes the detailed synergy scheme. Section (3) demonstrates the data set and the result obtained. The final section is a conclusion followed by the proposed future work.

1.1 Related work and paper contribution

The fusion of LIDAR and aerial images has been discussed in several literatures for different purposes. For example, 3D modelling as in (Nakagawa & Shibasaki, 2003), Building extraction as (Mao, Liu, & Zeng, 2009) and (Zhou & Zhou, 2014). (Mao et al., 2009) integrated LIDAR with CCD images on the same board and used LIDAR return to filter vegetation, and then the remaining points were clustered to determine the buildings' roofs and walls. (Zhou & Zhou, 2014) utilized geometrical features and structures of houses to extract the building and create Digital Building Models (DBM).

In (Li, Li, Wang, Wang, & Li, 2014) the aerial images with the normalized Digital Surface Model (nDSM) were used to filter the different objects in the urban areas.

Most of the authors targeting the integration of LIDAR with aerial images require both devices to be on the same board and to be utilized together for all images used in the classification process.

The main contribution of this paper is the synergy of LIDAR data with any aerial image that does not have to be on the same board, and generate a classifier that is trained by the classified LIDAR data. Consequently, the classifier can classify other aerial images without the need of further LIDAR data for the new image.

2. SYNERGY SCHEME

The Synergy scheme is illustrated in figure (1) below.

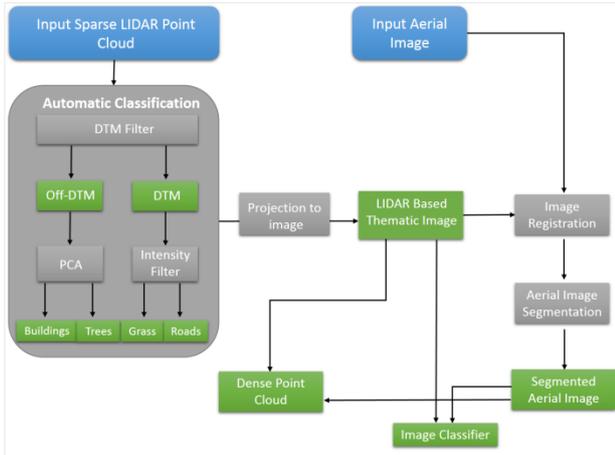


Figure (1): Synergy scheme for building unsupervised image classifier and densifying point cloud.

In the stage of generating an image classifier the input comprises an aerial image and airborne LIDAR point cloud. The point cloud is being classified (automatically) using the Digital Terrain Model (DTM) filter which discriminates the ground (DTM data) from the trees and buildings (off-DTM data). Principal Component Analysis (PCA) is then applied to the off-DTM data to discriminate between the trees and buildings. Grass and roads are extracted from the land surface based on the LIDAR return intensity (reflectance).

After classification, the point cloud is projected to an image by eliminating the z-coordinate. A LIDAR based thematic map is obtained from this projection.

By default, the aerial image and the airborne point cloud are georeferenced. That simplifies the image registration between the thematic map and the aerial image.

The aerial image is segmented and information about each segment is fed from the LIDAR based thematic map to the generated image classifier. Simultaneously, the segmented image constrains the point cloud densification to certain shape boundaries.

In the following subsection each of these processes is discussed in detail.

2.1 LIDAR data classification

As discussed in the previous subsection, LIDAR data is classified via DTM filter, PCA and the intensity filter, in this paper we follow the work done in (Badawy, Moussa, & El-Sheimy, 2014) with a slight modification to the DTM filter.

2.1.1 DTM filter

DTM generation was discussed in many literature, such as (Badawy et al., 2014), (Sithole, 2001) and (Vosselman, 2000). The main idea in those papers was that the height gradient from one point to another is computed and tested against a threshold. In this paper, the same process is done but with a different approach. First, the data are divided by a plane passing through the z-axis. The point with minimum height h_0 is selected and the plane, with z-component equals to $h_0 + \epsilon$, is used to divide the data. ϵ is a parameter that equals to few meters and is chosen according to the data characteristics.

If the height gradient is used in a local area of few meters it usually identifies points with sudden change in height as off-DTM points. In case of a roof of a building with an area of a few meters; the height on such roof will have a zero gradient. Consequently; a point on the roof of a building could be falsely

identified as DTM point while it actually belongs to the set of off-DTM points. Hence, the plane cut will group all the buildings points and their roofs in the off-DTM set of points except for few points below the cutting plane.

Now, these off-DTM points that are embedded in the DTM points can be filter with the following DTM filter. First the area is divided into grids of an equal size. This allows for the identification of local minima of the height in each grid as shown in figure (2).

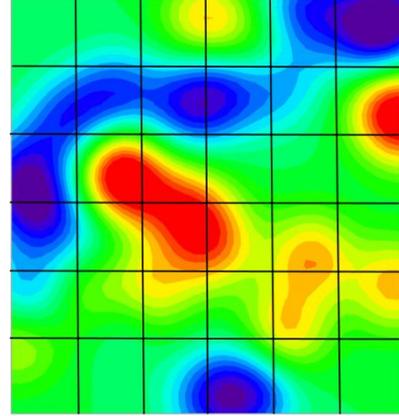


Figure (2): Equally divided grids with different height minima.

The gradient is computed for each point in the grid as compared to the grid's local minima. This can be written mathematically as (Badawy et al., 2014):

$$p_{min} = \{p_i \in N_i \subset C : \forall p_j \in N_i : h_{p_i} \leq h_{p_j}\} \quad (1)$$

$$DTM = \{p_j \in N_i \subset C : h_{p_j} - h_{p_m} \leq \Delta h(d(p_j, p_{min}))\} \quad (2)$$

where p_{min} are the local minima (point with minimum height value) in the neighbourhood which is a subset of the point cloud C . The height function Δh is a function of distance from the minima point $d(p_j, p_{min})$ and is defined to be $0.25d$.

The DTM filter results in two data sets, the DTM data defining the ground of the test area, and the off-DTM. Off-DTM data are further classified using the PCA and DTM data are classified via the LIDAR intensity.

2.1.2 PCA

The principal component analysis (PCA) is well documented in the literature. In this paper, it is used to classify the off-DTM data into trees and buildings.

The data are organized via the kD-tree data structure. The covariance matrix is computed for each subset of points in the kD-tree search radius R . Covariance matrix determines the geometry at which the LIDAR points are spread. This can be inferred from the eigenvalues of the covariance matrix.

If the covariance matrix is denoted C_v the eigenvalue problem is expressed as:

$$(C_v - \lambda I)X = 0 \quad (3)$$

where I is the identity matrix, λ is the eigenvalue corresponding to the eigenvector X .

Since the covariance matrix is 3×3 , there will be three eigenvalues corresponding to three eigenvectors. If the largest eigenvalue is much larger than the two other eigenvalues, then the points are spread along a line (representing edges). Similarly, if the points are spread over a plane (roofs and walls), two eigenvalues will be close in value and larger than the third.

If the three eigenvalues are close in values, the points are scattered in the 3D space (representing trees). These three conditions are summarized here:

- 1) Planar points, if $\lambda_1 \approx \lambda_2 \gg \lambda_3$.
- 2) Scattered points if $\lambda_1 \approx \lambda_2 \approx \lambda_3$.
- 3) Linear points, if $\lambda_1 \gg \lambda_2 \approx \lambda_3$.

Further details about PCA can be found in (Shi & Zakhor, 2011) and (Carlberg, Gao, Chen, & Zakhor, 2009).

2.1.3 Intensity Filter

LIDAR data are usually combined with the intensity (reflectance) of the reflected laser beam. The reflectance is dependent on the object's material. Typically, vegetation has larger reflectance than roads. Therefore, based on the reflectance, one can classify the DTM data into roads and grass. Since the DTM is either roads or grass and their return will be relatively constant, the histogram of the reflectance will normally show two peaks. The peak at larger reflectance corresponds to the vegetation and the other corresponds to the roads. There might be other small peaks which will be ignored and classified as ambiguous objects. These objects do not contribute to the training process of the classifier.

2.2 Projection of the LIDAR data to an image

The LIDAR data are projected to an image by eliminating the z-coordinate and using the (x, y) coordinates of the LIDAR data and projected to an image, where a transformation is performed to map from the (x, y) ground coordinate to (x_p, y_p) image coordinates. The equations of projection used here are discussed in (Mohammed, 2015) and are listed as:

$$(x_k, y_k) = P_k(X, Y) * S \quad (4)$$

$$(x', y') = (x_k - x_m, y_k - y_m) \quad (5)$$

$$i = -x_k + \frac{r}{2} \quad (6)$$

$$j = y_k + \frac{c}{2} \quad (7)$$

$$f(i, j) = RGB_{P_k} \quad (8)$$

The point cloud is first scaled when multiplied by the scale factor. Then the points are shifted to the centre by subtracting the centre of mass of the point cloud from each point. Afterwards, the indices of the pixel representing the point in the point cloud are obtained from equations (6) and (7) where c is the number of columns and r is the number of rows in the generated image. The RGB values associated with the point cloud are assigned to the image at the corresponding pixels as in equation (8).

The registration of the two images (The projected LIDAR based image and the Aerial image) is based on their georeferencing information.

2.3 Migrating the LiDAR based classification into spectral attributes of aerial images

For training the classifier, spectral attributes of the image segments are fed as input features while the corresponding classes are obtained using the previous LiDAR based classified image. To build such input features/corresponding classes, the aerial image of the training scene is segmented using a region growing segmentation approach conducted over the L*a*b* representation of the image. Since the spectral components of each segment are used directly as input features, the employed segmentation has been intentionally forced to have high number

of small segments (over segmentation) to increase the spectral purity of the segments and to avoid generating segments of more than one dominant spectral tone.

After the segmentation of the aerial image, the class that has the highest count of pixels in each segment is assigned to this segment. The segments that has significantly mixed classes (percentage of pixels belongs to dominant class < 75%) are excluded from the training data set. This exclusion helps to avoid the mixed areas typically found between trees and other classes of the scene. The mean values of the RGB and the L*a*b* components of the accepted segments are computed to serve as input features.

2.4 Aerial image classification algorithm

The training features and classes obtained in the previous step are employed for classification of aerial images of nearby scenes in an object based image analysis fashion.

For a new aerial image to be classified, the image has to be segmented first and the spectral attributes are computed for each segment (mean RGB and L*a*b* components). These attributes are used as an input to K Nearest Neighbour (KNN) classifier using the training features and classes obtained in the previous step. Over segmentation of test scene has also been emphasized to avoid generating segments of many different spectral tones.

2.5 Point cloud densification

Densification of the point cloud is another advantage of the synergy between point clouds and images. As the aerial image is segmented, each segment can be used to fill the gaps in the projected LIDAR based image. Those gaps are a result of the projection of LIDAR points without interpolation. In fact filling the gaps from the information from the segmented image can be considered as an interpolation process. The interpolated points in the image space need to be transformed to 3D space to obtain a densified point cloud. Equations (4) through (8) are used to derive the inverse mapping equation to obtain the (x_k, y_k) from the indices (i, j) . Obtaining the z-coordinate can be done via fitting a plane using the neighbourhood points. The segmented image is used to regularize the densified point cloud.

3. RESULT ANALYSIS AND DISCUSSION

3.1 Dataset

The data used for the test are obtained for parts of the state of Indiana, USA. The aerial images are obtained from Google Earth. The Airborne LIDAR data (IndianaMap Framework Data) are obtained from OpenTopography source with links : <http://www.indianamap.org> and <http://dx.doi.org/10.5069/G9959FHZ>

The aerial image which is used for the training process of the classifier is shown in figure (3) below.



Figure (3): Aerial image of the original scene used for training the classifier.

3.2 Result analysis

3.2.1 Lidar DTM extraction

Figure (4) shows the result obtained for the DTM (ground) and off-DTM extraction. The points in red represent the off-DTM data, while the grey points represent the DTM data.

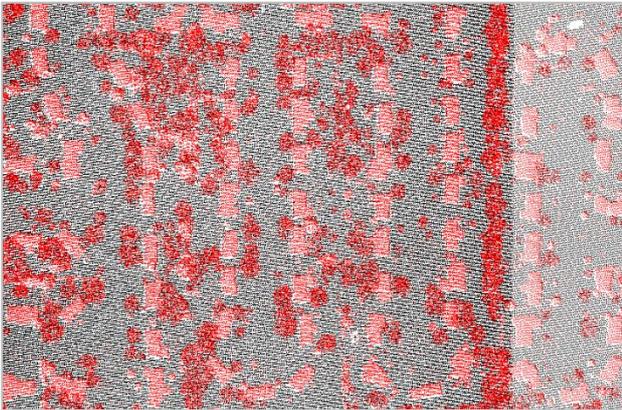


Figure (4): DTM and off-DTM data extracted using the DTM filter.

3.2.2 Intensity filter

The DTM data are classified into roads, grass and ambiguous data based on the histogram of the LIDAR reflectance. In figure (5) the histogram is plotted with two main peaks one around the reflectance value of 20 (representing roads) and the other around the reflectance value of 222 (representing grass). A window is taken around both values as a threshold for the ground data classification. Between the two peaks there are few other peaks corresponding to data that are left unclassified. The reason for excluding those data points from the classification and training process is that they will be most probably misclassified. Therefore, ignoring those points is preferable.

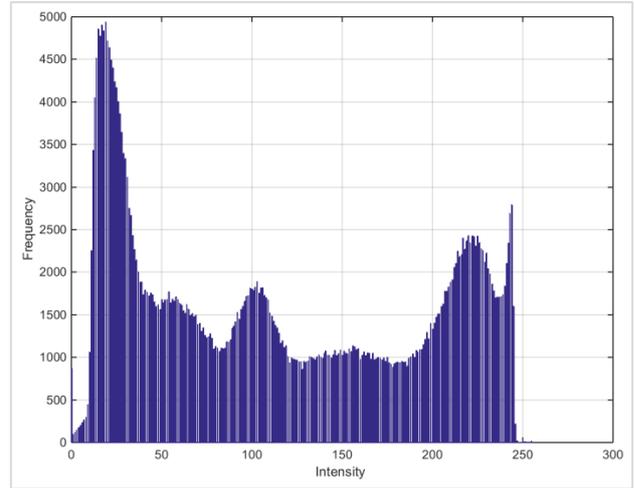


Figure (5): Histogram of the reflectance (intensity) of the LIDAR data.

3.2.3 PCA

PCA was used to determine the geometry of the objects in the point cloud. Then they are given the brown label if they are planar (representing buildings) or given a green label to represent the trees.

Figure (6) is a projection of the completely classified LIDAR data where the green colour represents trees, the brown colour represents buildings, the roads are represented with black colour, and the grass is represented with the yellow colour. The blue colour represents unclassified areas. It can be clearly noticed that the classification result agrees with the aerial image in figure (3).

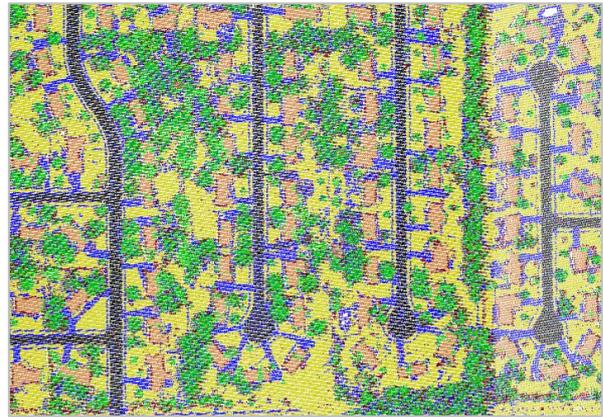


Figure (6): Projection of the LIDAR data after complete classification.

3.2.4 Aerial image segmentation

Figure (7) shows the segmented aerial image. The image is over-segmented so that when it is registered with the LIDAR based thematic map each segment lies inside the classified object in the point cloud. This ensures that segmented regions will contain spectral attributes from only one object.



Figure (7): Aerial image over-segmented to avoid generating segments of many different spectral tones.

Each segment is recognized automatically using the LIDAR data classification. Those segments are then used to train the classifier and the classifier is tested with a test image.

3.2.5 Test scene

The test image is acquired for the same state of Indiana in USA to ensure that the spectral characteristics of the original and the test images are similar. Otherwise; there will be misclassification especially between roads and buildings.

Figure (9) shows the segmented aerial image for the test scene. The result of classification is shown in figures (10), (11) and (12) for the buildings, the vegetation and the roads.

The results exhibit very good classification accuracy especially in the classification of the vegetation area. Those results are also supported with the confusion matrix in table (1).



Figure (9): segmentation of the test scene that is used with the trained classifier.



Figure (8): Test image at the same state of Indiana, USA.



Figure (10): Buildings extracted from the test image using the trained classifier.

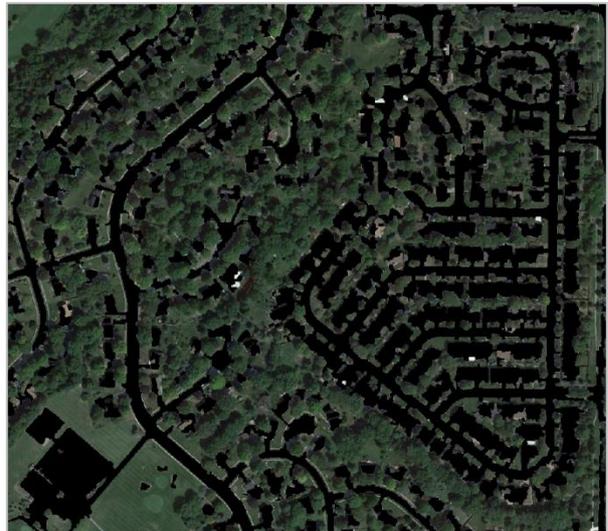


Figure (11): Vegetation extracted from the test image using the trained classifier.



Figure 12: Roads extracted from the test image using the trained classifier.

Parts of the roads are misclassified as buildings due to the similarity of their spectral attributes in parts of the image.

3.3 Classification quality assessment

To assess the classification result, a manual classification was performed and used as a ground truth. The confusion matrix is listed in table (1). This matrix describes the amount of agreement between the classified objects in the test image and the ground truth.

Associated with the confusion matrix is the Cohen kappa coefficient κ (Cohen, 1960) and (Monserud & Leemans, 1992) given by:

$$\kappa = \frac{P_0 - P_e}{1 - P_e} \quad (9)$$

where

$$P_0 = \sum p_{ii} \quad (10)$$

$$P_e = \sum_{i \neq j} p_i p_j \quad (11)$$

And p_i are the elements of the confusion matrix.

Table (1): Confusion matrix of the classification result.

Actual (%)		Classified		
		Buildings	vegetation	roads
Buildings	100	72.7977	26.5766	0.6257
Vegetation	100	0	100	0
Roads	100	19.7390	11.5642	68.6968

The computed value of κ for this test is $\kappa = 0.7809$ which indicates a very good to excellent agreement.

Both of the confusion matrix and the value of κ are in agreement with the results of figures (10) to (11). For example; it can be clearly seen from figure (11) that the vegetation is perfectly classified without misclassifications with roads or buildings due to sharp difference in spectral attributes between vegetation and the other objects. When looking at the confusion matrix, it is seen that 100 % of the vegetation are classified as vegetation.

3.4 Point cloud densification result

The densification of the sparse point cloud was performed based on the segmented aerial image. As discussed earlier, segmented image offer a smoother boundaries for the densified point

cloud. Moreover; the point cloud obtained can be correctly labelled using the aerial image.

Figure (13) demonstrates part of the point cloud for a building roof before and after densification. Distortion in the image can be seen clearly in the (a). While in the right image (b) the roof of the building is denser and more regular.

In figure (14) the result of densification is demonstrated with the surrounding environment of the building. The geometry of the building is more accurate with increased data points. The increase of regularized data points is useful if a surface mesh is to be computed, since gaps in the data points lead to inaccurate surface reconstruction. It can be seen also that the ambiguous points can be removed using the segmented image.

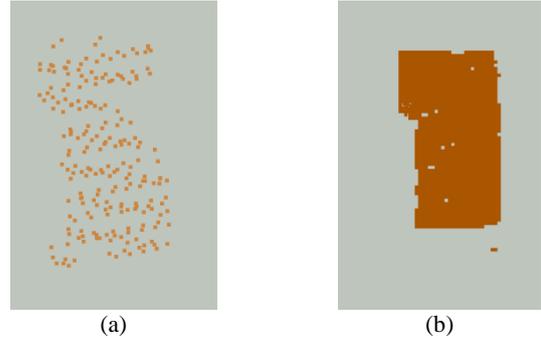


Figure (13): Part of the sparse point cloud of a building roof before densification in (a) and after densification (b).

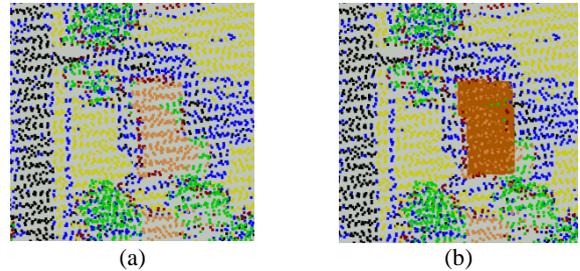


Figure (14): Part of the building with its surroundings before densification in (a) and after densification (b).

4. CONCLUSION

This paper investigates the possibility of using LIDAR point cloud with aerial images, mainly, to obtain unsupervised (automated) classification of aerial images and to densify the point cloud associated with the original image. Different classification steps have been employed on the LIDAR data, such as DTM filter, intensity filter and PCA to accurately classify the point cloud. This classification is used to train the image classifier. The classification result was shown to be very good. The densification algorithm proved to provide a regularized dense point cloud from a sparse point cloud. The limitation of the classification algorithm appears in the misclassification between buildings and roads due to the similarity of the spectral attributes between them. Shadow is another problem that faces the classification.

ACKNOWLEDGEMENTS

This work was supported by Dr. Naser El-Sheimy research funds from NSERC and Canada Research Chairs programs. This material is based on [data, processing] services provided by the OpenTopography Facility with support from the National Science Foundation under NSF Award Numbers 1226353 & 1225810.

The 'IndianaMap Framework Data are found at:
<http://www.indianamap.org>
<http://dx.doi.org/10.5069/G9959FHZ>

REFERENCES

- Badawy, H. M., Moussa, A., & El-Sheimy, N. (2014). Automatic Classification of Coarse Density Lidar Data in Urban Area, *XL*(June), 2500. <http://doi.org/10.5194/isprsarchives-XL-5-77-2014>
- Carlberg, M., Gao, P., Chen, G., & Zakhor, A. (2009). Classifying urban landscape in aerial lidar using 3D shape analysis. In *Proceedings - International Conference on Image Processing, ICIP* (pp. 1701–1704). <http://doi.org/10.1109/ICIP.2009.5413385>
- Cohen, J. (1960). A coefficient of agreement of nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46. <http://doi.org/10.1177/001316446002000104>
- Li, S., Li, Z., Wang, H., Wang, J., & Li, L. (2014). Urban land cover classification using aerial LiDAR and CCD images. In *2014 IEEE Geoscience and Remote Sensing Symposium* (pp. 1967–1970).
- Mao, J., Liu, X., & Zeng, Q. (2009). Building extraction by fusion of LIDAR data and aerial images. In *2009 Joint Urban Remote Sensing Event*. <http://doi.org/10.1109/URS.2009.5137631>
- Mohammed, H. (2015). *Fusion of Terrestrial and Airborne Laser Data for 3D Modeling Applications*. University of Calgary.
- Monserud, R. A., & Leemans, R. (1992). Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling*, *62*(4), 275–293. [http://doi.org/10.1016/0304-3800\(92\)90003-W](http://doi.org/10.1016/0304-3800(92)90003-W)
- Nakagawa, M., & Shibasaki, R. (2003). Integrating high resolution air-borne linear CCD (TLS) imagery and LIDAR data. In *2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas, URBAN 2003* (pp. 236–240). <http://doi.org/10.1109/DFUA.2003.1219994>
- Shi, X., & Zakhor, A. (2011). Fast approximation for geometric classification of LiDAR returns. *Proceedings - International Conference on Image Processing, ICIP*, 2925–2928. <http://doi.org/10.1109/ICIP.2011.6116272>
- Sithole, G. (2001). Filtering of laser altimetry data using a slope adaptive filter. In ... *Archives of Photogrammetry Remote Sensing and ...* (Vol. XXXIV, pp. 22–24). Retrieved from http://lr.tudelft.nl/fileadmin/Faculteit/LR/Organisatie/Afdelingen_en_Leerstoelen/Afdeling_RS/Optical_and_Laser_Remote_Sensing/Publications/Papers/018-2001/doc/sithole_annapolis.pdf
- Vosselman, G. (2000). Slope based filtering of laser altimetry data. *International Archives of Photogrammetry and Remote Sensing, Vol. 33, Part B3/2, 33*(Part B3/2), 678–684. [http://doi.org/10.1016/S0924-2716\(98\)00009-4](http://doi.org/10.1016/S0924-2716(98)00009-4)
- Zhou, G., & Zhou, X. (2014). Seamless fusion of LiDAR and aerial imagery for building extraction. *IEEE Transactions on Geoscience and Remote Sensing*, *52*(11), 7393–7407. <http://doi.org/10.1109/TGRS.2014.2311991>