

Multispectral airborne lidar intensity data: from correction to classification

Wai Yeung Yan

Department of Land Surveying and Geo-informatics, Hong Kong Polytechnic University, Hong Kong

*Accepted by the 40th Asian Conference on Remote Sensing, Daejeon, Korea

Abstract

Over the years, airborne LiDAR has demonstrated its merits over traditional aerial photogrammetry and remote sensing for fine scale mapping due to its elimination of shadowing and relief displacement effects. In spite of that, conventional monochromatic LiDAR system (i.e. single wavelength) has a limited analytical capability since it has a lack of fruitful spectral information. Until very recently, the world's first multispectral airborne LiDAR system manufactured by Teledyne Optech, i.e. Optech Titan, represents a significant breakthrough in LiDAR remote sensing. The multispectral LiDAR system is capable of collecting high resolution dense 3D point cloud with three laser channels (i.e. 532 nm, 1064 nm and 1550 nm). Such an enhanced capability can facilitate the use of these multi-wavelength laser intensity data for retrieving land surface properties and improving Earth's knowledge discovery. Since LiDAR intensity data unavoidably suffers different levels of noise due to the varying system and environmental conditions, LiDAR intensity data should undergo certain radiometric pre-process prior to any analysis. This paper presents a comprehensive solution, including 1) a LiDAR Scan Line Correction (LSLC) to remove the intensity banding effect found in individual LiDAR data strip, and 2) a radiometric correction model to reduce the system- and environmental induced distortions between overlapping data strips, to facilitate multispectral LiDAR data classification. With the improved radiometric quality, the intensity homogeneity was enhanced significantly by 0.2% to 52.5%, depending on the land cover classes. By using multispectral airborne LiDAR intensity data for classification, additional laser channels yielded an improvement of overall accuracy by 13% to 21%, comparing to monochromatic LiDAR intensity data classification in most of the scenarios. This work provides, the first of its kind, a practical and universal solution for radiometric pre-processing the multispectral airborne LiDAR data to support enhanced data analysis.

1. Introduction

Monochromatic (i.e. single wavelength) airborne LiDAR system has been widely used for various topographic applications due to its capability of collecting dense 3D point cloud (Yan et al., 2015). Most of the existing applications reap the benefits of the collected data's 3D geometry, such as ground filtering (Vosselman, 2000), 3D building extraction (Zhang et al.,

2006), and tree canopy modeling (Kato et al., 2009), which are hard to be achieved using traditional remote sensing images. To compensate the lack of multispectral information, airborne LiDAR can be fused with aerial photos or high resolution satellite images to support thematic analysis. Nevertheless, the associated problems, including geometric misalignment between the point cloud and image data, appearance of shadowing and effects of tilted objects, all potentially degrade the data quality and analytical accuracy. Therefore, the need of having a LiDAR system that is capable of collecting multi-wavelength laser intensity data has been sought and scientifically justified.

Recently, Teledyne Optech has announced the world’s first multispectral airborne LiDAR system, named Optech Titan, which is capable of collecting the backscattered laser signal strength with wavelengths of 532 nm (green), 1064 nm (near-infrared, NIR), and 1550 nm (infrared, IR), see Fig. 1a. Such a ground breaking development further facilitates scientists and surveyors to improve the capability of Earth surface’s analysis, and opens many new doors for various urban and environmental applications. Figs. 1b and 1c show an example of the collected multispectral LiDAR intensity data and the aerial photo. One can easily observe that the effects of shadowing dominate nearby the tree canopies and houses on the aerial photo. Also, the adjacent regions located nearby the titled objects are occluded. All these drawbacks are overcome on the multispectral airborne LiDAR intensity data due to the direct geo-referencing technique.

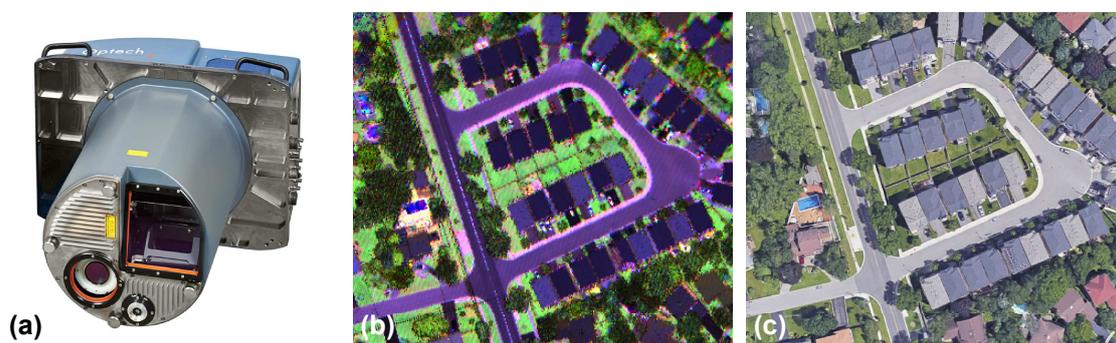


Figure 1: (a) Optech Titan, (b) Multispectral airborne LiDAR intensity data, and (c) Aerial photo.

With the invention of multispectral airborne LiDAR intensity data, it is foreseeable that traditional thematic mapping capability can further excel, particularly in fine scale, large-area land cover mapping. The current version of global land cover map, GlobeLand30, was generated using Landsat satellite images with 30-m resolution (Chen et al., 2015). Satellite remote sensing techniques are no longer capable of generating the next generation global land cover map with foreseeable spatial resolution down to 1 - 5 m. As a result, multispectral airborne LiDAR deems to be a viable alternative to fulfill such a goal. In spite of that, the intensity data noise, i.e. the stripe artifacts appeared in individual LiDAR data strip and overlapping LiDAR data strips, degrade the data quality. Currently, there is a lack of universal solution to remove these striping noises. In this paper, a comprehensive solution of improving the radiometric quality of multispectral airborne LiDAR intensity data, including

a LiDAR scan line correction and an overlap-driven intensity correction, is proposed so as to maximize the benefits of using the LiDAR intensity data for land cover classification.

2. Radiometric Correction

2.1. LiDAR Scan Line Correction

The LiDAR scan line correction (LSLC), which is previously proposed by the author (Yan and Shaker, 2018), was acquired to reduce the striping noise found in the individual LiDAR data strip. The cause of the striping noise can be ascribed by the intensity banding effect, which is mainly due to the backscattered laser signal partially falling outside the receiver’s field of view in a particular scanning direction. The recorded laser signal strength is attenuated in this specific scanning direction, resulting in a striping noise pattern found in the LiDAR intensity data. The mechanism of the LSLC is described in the Fig. 2.

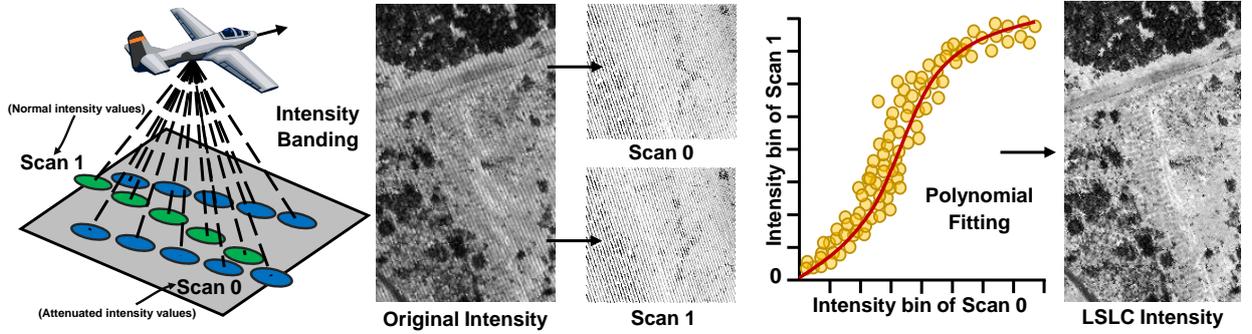


Figure 2: Overall workflow of LiDAR scan line correction for striping noise removal.

Assume LiDAR dataset L having the intensity banding, the process is described as below:

1. The LSLC starts with splitting the dataset L into L_0 and L_1 , which represent the data with scan flag 0 and 1, respectively.
2. The subset serves as a reference set if its mean intensity is higher than the other one.
3. The data points from L_1 (assuming it to be the reference set) should be matched with data points from L_0 through a kd -tree search, resulting in a n number of paired points.
4. A high order polynomial model (e.g. 3rd order) is constructed with parameters (i.e. a to j) including intensity I and scan angle θ . The matrix form is listed as below:

$$\begin{bmatrix} I_{1,1} \\ I_{1,2} \\ \vdots \\ I_{1,n} \end{bmatrix} = \begin{bmatrix} I_{0,1}^3 & \theta_{0,1}^3 & I_{0,1}^2 \theta_{0,1} & I_{0,1} \theta_{0,1}^2 & I_{0,1}^2 & \theta_{0,1}^2 & I_{0,1} \theta_{0,1} & I_{0,1} & \theta_{0,1} & 1 \\ I_{0,2}^3 & \theta_{0,2}^3 & I_{0,2}^2 \theta_{0,2} & I_{0,2} \theta_{0,2}^2 & I_{0,2}^2 & \theta_{0,2}^2 & I_{0,2} \theta_{0,2} & I_{0,2} & \theta_{0,2} & 1 \\ \vdots & \vdots \\ I_{0,n}^3 & \theta_{0,n}^3 & I_{0,n}^2 \theta_{0,n} & I_{0,n} \theta_{0,n}^2 & I_{0,n}^2 & \theta_{0,n}^2 & I_{0,n} \theta_{0,n} & I_{0,n} & \theta_{0,n} & 1 \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ \vdots \\ j \end{bmatrix} \quad (1)$$

5. By using iteratively re-weighted least-squares together with M-estimator, the above equation can be solved, resulting in the correction parameters a to j .
6. Finally, the subset L_0 is combined with L_1 after the intensity values of L_0 is corrected using the parameters. The scan line corrected L should be free from the striping noise.

2.2. Overlap-driven Intensity Correction

Although the LSLC can significantly remove the striping noise in individual LiDAR data strip, a notable level of striping noise can be observed at the swath edge when combining multiple overlapping LiDAR data strips due to the system- and environmental-induced distortions. As a result, an overlap-driven intensity correction is proposed with an enhancement to the author's previous models (Yan and Shaker, 2014, 2016) and (Ding et al., 2013). Recalling the radar (range) equation that is commonly used for radiometric correction of LiDAR intensity data:

$$P_r = \frac{P_t D^2}{4\pi R^2} \cdot \frac{4\pi \rho A \cos\theta}{4\pi R^2} \cdot \eta_{sys} \cdot \eta_{atm} \quad (2)$$

where P_r is the recorded laser power, P_t is the transmitted laser power, D is the aperture diameter, R is the range, ρ is the surface reflectance, A is the laser footprint (which is dependent on R), θ is the incidence angle, η_{sys} is the system transmission factor, and $\eta_{atm} = e^{-2cR}$ is the atmospheric attenuation factor (Yan and Shaker, 2014). P_r is commonly regarded as the intensity data I after linearization and ρ is the desired (corrected) intensity data. Since most of the parameters, including the P_t, D, η_{sys} , are constant, the above Eq. 2 becomes

$$\rho = I \cdot \left[\frac{R}{R_m} \right]^2 \cdot \frac{1}{\cos\theta} \cdot e^{2cR} \quad (3)$$

where R_m is a constant and is usually regarded as the minimum range value. The above equation adopts the Lambertian reflectance mechanism, where a number of studies have reported its limitations in LiDAR intensity correction (Carrea et al., 2016; Tan et al., 2016). To overcome such a drawback, a higher degree of correction can be applied toward the range and angle:

$$\rho = I \cdot \left[\frac{R}{R_m} \right]^a \cdot \left[\frac{1}{\cos\theta} \right]^b \cdot e^{2cR} \quad (4)$$

Assume there exist two LiDAR data strips (L_i and L_j) that are partially overlapped, data points from the two LiDAR data strips are matched by searching closest distance based on building a kd -tree, resulting in a total n number of paired points. Assume the paired points located on the identical surfaces, as a result the ρ of the two data points should be identical. Substituting the two data strips into Eq. 4 and dividing them becomes:

$$\frac{I_i}{I_j} = \left[\frac{R_j}{R_i} \right]^a \cdot \left[\frac{\cos\theta_i}{\cos\theta_j} \right]^b \cdot e^{2(R_j - R_i)c} \quad (5)$$

Though applying logarithm on both sides of equation, Eq. 5 can be linearized. The matrix form of the equation would be:

$$\begin{bmatrix} \ln \left[\frac{R_{j,1}}{R_{i,1}} \right] & \ln \left[\frac{\cos\theta_{i,1}}{\cos\theta_{j,1}} \right] & 2(R_{j,1} - R_{i,1}) \\ \ln \left[\frac{R_{j,2}}{R_{i,2}} \right] & \ln \left[\frac{\cos\theta_{i,2}}{\cos\theta_{j,2}} \right] & 2(R_{j,2} - R_{i,2}) \\ \vdots & \vdots & \vdots \\ \ln \left[\frac{R_{j,n}}{R_{i,n}} \right] & \ln \left[\frac{\cos\theta_{i,n}}{\cos\theta_{j,n}} \right] & 2(R_{j,n} - R_{i,n}) \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \ln \left[\frac{I_{i,1}}{I_{j,1}} \right] \\ \ln \left[\frac{I_{i,2}}{I_{j,2}} \right] \\ \vdots \\ \ln \left[\frac{I_{i,n}}{I_{j,n}} \right] \end{bmatrix} \quad (6)$$

Similar to the LSLC, the above linearized equation can be solved using iteratively re-weighted least squares together with M-estimation. The solved parameters (a , b and c) can be re-applied to Eq. 4 for both L_i and L_j in order to retrieve the ρ , which is treated as the corrected intensity data. After implementing the overlap-driven intensity correction, the striping noise found in the swath edge should be significantly reduced after combining the corrected L_i and L_j .

3. Experimental Work

3.1. Multispectral Airborne LiDAR Data

Four multispectral LiDAR data subsets collected by Optech Titan were used to demonstrate the effects of the proposed radiometric correction and the subsequent land cover classification. The datasets were collected on different land cover scenarios, where all of them were located in Ontario, Canada. The pulse repetition frequency (PRF) were set to be 600 kHz for the three laser channels, i.e. channel 1 (1550 nm), channel 2 (1064 nm) and channel 3 (532 nm), with scan frequency of 40 Hz and flying height ranging from 400 m to 1 km. With these settings, the mean point density yielded higher than 10 to 15 points/m², resulting in a mean point spacing better than 0.5 m. The multispectral LiDAR datasets cover both rural and sub-urban areas.

3.2. Assessment of Intensity Homogeneity

To examine the effects of radiometric correction, the intensity homogeneity was assessed on the multispectral LiDAR intensity data before and after implementing the correction. The coefficient of variation (cv) was acquired to quantitatively measure the intensity homogeneity. Within each of the studied land covers, the cv can be derived as

$$cv = \frac{\sigma_i}{\mu_i}, \quad \forall i \subseteq L \quad (7)$$

where σ_i represents the standard deviation of a set of LiDAR data points i within a certain land cover, and μ_i represents the mean of a set of LiDAR data points i within a certain land cover. A reduction of cv after radiometric correction implies an improvement of intensity homogeneity, which also indicates a successful removal or reduction of striping noise.

3.3. Land Cover Classification

Finally, one of the multispectral airborne LiDAR datasets was used to perform land cover classification. Training data of five land cover classes, including water bodies, grass cover, tree canopies, houses and paved ground, were collected on the LiDAR data. They were used to train a multivariate Gaussian classifier with different combinations of feature sets, including the intensity and elevation of the three laser channels. Once the classification results based on different feature sets were generated, accuracy assessment was conducted based on the random check points generated on the reference aerial photos. The overall accuracy was computed based on the derived confusion matrix. The overall aim of this task is to demonstrate the effects of additional laser channels toward the classification.

4. Results and Analysis

4.1. Effects of Radiometric Correction

Fig. 3 shows the multispectral airborne LiDAR intensity data before and after implementing the radiometric correction. It shows five land cover classes, including harvest farmland, grassland, paved road, tree canopy, and wetland. One can obviously note that a significant striping noise appeared on the multispectral airborne LiDAR intensity data (see Figs. 3a to 3d). This can be explained by the intensity data collected for scan flag 1 (i.e. a positive scan direction which is a scan moving from the left side of the in-track direction to the right side and negative the opposite) suffered from significant energy loss. Such intensity banding effect only occurred in the channel 1 (1550 nm) and channel 2 (1064 nm). As shown in Figs. 3e to 3h, the proposed correction successfully provided a robust matching of the intensity data bin of scan flag 1 with respect to the scan flag 0, leading to a significant striping noise removal in the resulted intensity images.

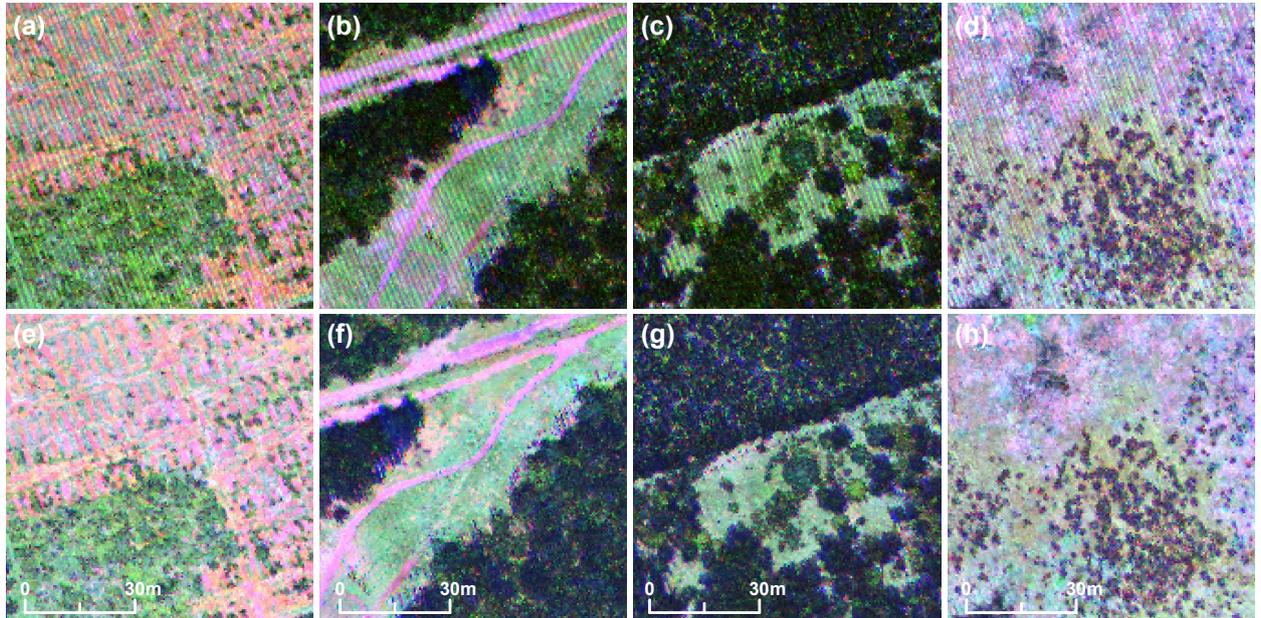


Figure 3: (a) to (d) Multispectral airborne LiDAR intensity data, and (e) to (h) corrected intensity data.

Table 1 shows the computed cv of e channels 1 and 2 on the original intensity data and the corrected intensity data. Since channel 1 did not have significant striping noise, therefore, the cv reduction was mild among the three land cover classes. Both vegetation features (i.e. grass cover and tree canopies) were recorded with a cv reduction by less than 1% in channel 1, while the farmland and wetland were recorded with a cv reduction by 6%. Since the absolute cv value was low on the paved road, the cv on the original intensity was 0.167 and it decreased to 0.148 after implementing the correction, resulting in a drop of cv by 11.5% . In channel 2, since the intensity banding effect was obvious, a notable reduction of cv can be found after running the correction. Similar to channel 1, the tree canopies

were recorded with the less cv reduction, which was only 6.2%. The cv of the grass cover was decreased from 0.397 to 0.323 (\downarrow 18.6%) after applying the correction. A significant reduction of cv was observed in the rest of the three land cover classes, where the cv of farmland, wetland and paved road was reduced from 0.285 to 0.184 (\downarrow 35.5%), 0.239 to 0.141 (\downarrow 40.9%), and 0.242 to 0.115 (\downarrow 52.5%), respectively.

Table 1: A summary of cv of five land cover classes before and after applying radiometric correction.

	Channel 1 (Original Intensity)	Channel 1 (After Correction)	Channel 2 (Original Intensity)	Channel 2 (After Correction)
Farm	0.273	0.258 (\downarrow 5.5%)	0.285	0.184 (\downarrow 35.5%)
Grass	0.374	0.373 (\downarrow 0.2%)	0.397	0.323 (\downarrow 18.6%)
Road	0.167	0.148 (\downarrow 11.5%)	0.242	0.115 (\downarrow 52.5%)
Tree	0.629	0.626 (\downarrow 0.6%)	0.642	0.602 (\downarrow 6.2%)
Wetland	0.212	0.200 (\downarrow 5.7%)	0.239	0.141 (\downarrow 40.9%)

4.2. Multispectral LiDAR Data Classification

The multispectral LiDAR intensity data was further used to explore land cover classification. In the author’s previous experiments as reported in Shaker et al. (2019), incorporation of additional laser channels did not significantly improve the land-water classification, since these two classes (i.e. land and water) already had significant separability based on the elevation. However, in this experiment, the land cover classification scenario was further broken down into five land cover classes, including grassland, house, paved land, tree canopies and water bodies. As shown in Fig. 4, individual laser intensity data, regardless of channel 1, 2 or 3, does not seem to provide a good separability among the five classes. Nevertheless, combining the three laser channels obviously provides an ideal feature space to delineate these five classes (see the middle three sub-figures in Fig. 4).

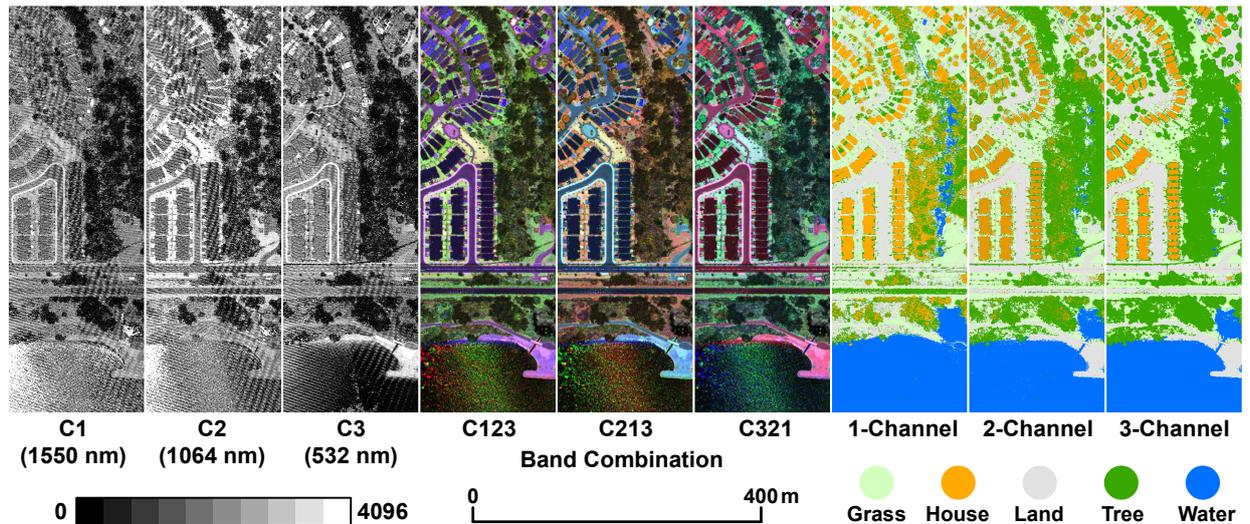


Figure 4: Individual LiDAR intensity data, multispectral LiDAR intensity data displayed with different band combinations and land cover classification results using different combinations of laser channels.

Table 2 shows the overall accuracy of land cover classification using different combinations of laser channels. It is obvious that the use additional laser channels together with the elevation can yield a high classification accuracy. Without incorporating the elevation, an overall accuracy ranging from 30% to 37% was achieved purely based on the individual LiDAR intensity data. By using two laser channels, the overall accuracy was increased to 44% to 45%, and it finally reached to 51% when all the three LiDAR intensity data were used as feature sets for the five-class classification. The merit of airborne LiDAR data over remote sensing image for classification can be attributed to the LiDAR-derived height features. With incorporation of LiDAR elevation, the overall accuracy of single-channel classification ranged from 72% to 81%, resulting in an over 40% improvement to those without using elevation. Combing two laser channels together with the elevation, the overall accuracy yielded more than 83%. An overall accuracy of 85% was achieved when the three laser channels were used. The sub-figures in the right hand side of Fig. 4 show the classification results using one to three laser channels. Fig. 5 shows the 3D classification result using the three laser channels together with the elevation feature.

Table 2: Accuracy assessment of multispectral LiDAR data classification.

Feature Set	Without Elevation	With Elevation
Channel 1 (1550 nm)	30.28%	71.51%
Channel 2 (1064 nm)	33.59%	75.54%
Channel 3 (532 nm)	36.64%	80.66%
Channels 1 + 2	44.78%	83.21%
Channels 1 + 3	43.51%	80.66%
Channels 2 + 3	45.29%	83.21%
Channels 1 + 2 + 3	50.89%	84.73%

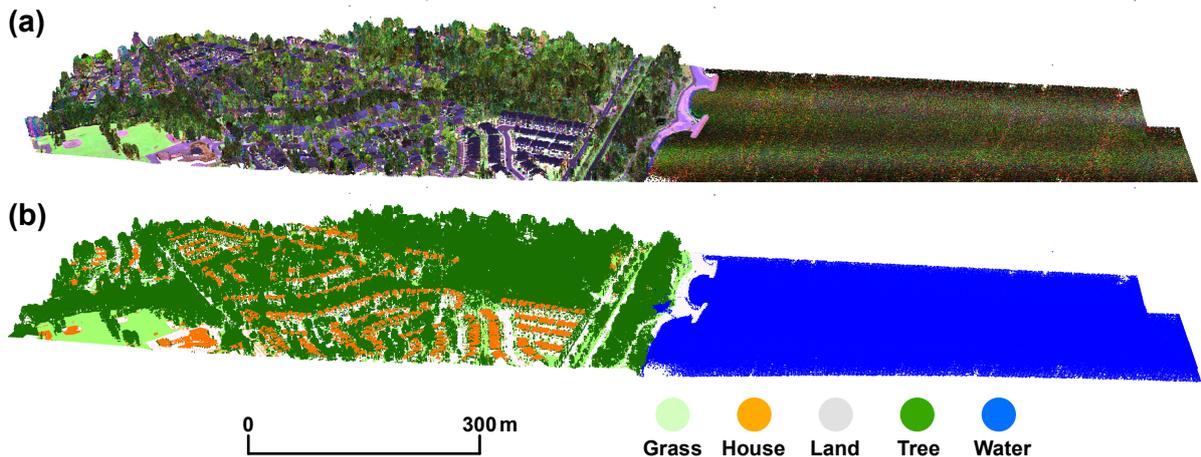


Figure 5: (a) Multispectral LiDAR intensity data, and (b) land cover classification result using tri-channels.

Based on the experimental testing, an improvement of classification accuracy ranged from 13% to 21% was recorded in the scenario with or without using the elevation feature. Since

the multispectral airborne LiDAR dataset was collected with high PRF, the point density yielded more than 10 points/m² in the testing dataset. As a result, undesired tiny objects (such as vehicles, road markings, powerlines, etc.) appeared in the study scene causing spectral mixture among the land cover classes. These undesired objects were the main cause of the mis-classification, even though multispectral LiDAR intensity data from all three channels were used. Since the experimental testing only aimed to assess the performance of using multispectral LiDAR intensity data for land cover classification, different combinations of LiDAR intensity data together with the elevation were examined. Other new feature sets, such as surface normal, texture measures, intensity/height variation, or normalized difference feature index, can be generated in order to further improve the classification accuracy, which can yield better than 90% (Shaker et al., 2019).

5. Conclusions

To embrace the next generation global land cover mapping, it is believed that satellite remote sensing images are no longer able to generate reliable results due to the presence of shadowing and tilted object effects. As a result, the latest invention of multispectral airborne LiDAR system can certainly overcome these drawbacks and generates a fine-scale, large-area land cover map. Due to the presence of striping noises found in individual LiDAR data strip and overlapping LiDAR data strips, multispectral LiDAR intensity data should undergo certain pre-processing in order to improve the radiometric quality. Since there is a lack of practical solution in existing literature, this work proposes a dual-correction approach to remove the striping noises. The LiDAR scan line correction is proposed to adjust the radiometric misalignment between the two scan lines due to the intensity banding effect. An overlap-driven intensity correction is developed to adjust the system and environmental induced distortions. Based on the experimental testing, the intensity homogeneity was measured on five different land cover classes on the intensity data before and after implementing the correction, where a significant enhancement of data quality was achieved by removing these stripe artifacts. A reduction of cv was obvious on ground features (i.e. farmland, grassland, wetland and paved road) ranging from 0.2% to 11.5% in channel 1 (1550 nm) and 18.6% to 52.5% in channel 2 (1064 nm). The intensity homogeneity of tree canopies only recorded a reduction of cv by 0.6% and 6.2% in these two channels, respectively. The multispectral airborne LiDAR intensity data was explored with different classification scenarios. Comparing to only using a single laser channel, the use of tri-laser channels achieved an improvement of classification accuracy by 13% to 21%, regardless of using the elevation feature or not. To the best of the knowledge, this work is the first of its kind that provides a practical and universal solution to remove the striping noise found on the monochromatic and multispectral airborne LiDAR intensity data, where such process should be treated as an essential step prior to using the LiDAR intensity data for any thematic analysis.

Acknowledgments

The author would like to express his appreciation to the following individuals who brought in many illuminating discussions during the study and provided the testing datasets:

Dr. Ahmed Shaker (Ryerson University), Dr. Paul E. LaRocque & Dr. Ana P. Kersting (Teledyne Optech), and Dr. Karin van Ewijk & Prof. Paul Treitz (Queen's University).

References

- Carrea, D., Abellan, A., Humair, F., Matasci, B., Derron, M.H., Jaboyedoff, M., 2016. Correction of terrestrial LiDAR intensity channel using Oren–Nayar reflectance model: An application to lithological differentiation. *ISPRS Journal of Photogrammetry and Remote Sensing* 113, 17–29.
- Chen, J., Chen, J., Liao, A., Cao, X., Chen, L., Chen, X., He, C., Han, G., Peng, S., Lu, M., et al., 2015. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing* 103, 7–27.
- Ding, Q., Chen, W., King, B., Liu, Y., Liu, G., 2013. Combination of overlap-driven adjustment and Phong model for LiDAR intensity correction. *ISPRS Journal of Photogrammetry and Remote Sensing* 75, 40–47.
- Kato, A., Moskal, L.M., Schiess, P., Swanson, M.E., Calhoun, D., Stuetzle, W., 2009. Capturing tree crown formation through implicit surface reconstruction using airborne lidar data. *Remote Sensing of Environment* 113, 1148–1162.
- Shaker, A., Yan, W.Y., LaRocque, P.E., 2019. Automatic land-water classification using multispectral airborne LiDAR data for near-shore and river environments. *ISPRS Journal of Photogrammetry and Remote Sensing* 152, 94–108.
- Tan, K., Cheng, X., Cheng, X., 2016. Modeling hemispherical reflectance for natural surfaces based on terrestrial laser scanning backscattered intensity data. *Optics Express* 24, 22971–22988.
- Vosselman, G., 2000. Slope based filtering of laser altimetry data. *International Archives of Photogrammetry and Remote Sensing* 33, 935–942.
- Yan, W.Y., Shaker, A., 2014. Radiometric correction and normalization of airborne LiDAR intensity data for improving land-cover classification. *IEEE Transactions on Geoscience and Remote Sensing* 52, 7658–7673.
- Yan, W.Y., Shaker, A., 2016. Radiometric normalization of overlapping LiDAR intensity data for reduction of striping noise. *International Journal of Digital Earth* 9, 649–661.
- Yan, W.Y., Shaker, A., 2018. Airborne lidar intensity banding: cause and solution. *ISPRS Journal of Photogrammetry and Remote Sensing* 142, 301–310.
- Yan, W.Y., Shaker, A., El-Ashmawy, N., 2015. Urban land cover classification using airborne LiDAR data: a review. *Remote Sensing of Environment* 158, 295–310.
- Zhang, K., Yan, J., Chen, S., 2006. Automatic construction of building footprints from airborne LIDAR data. *IEEE Transactions on Geoscience and Remote Sensing* 44, 2523–2533.