INTENSITY CORRECTION OF MULTISPECTRAL AIRBORNE LASER SCANNING DATA

Wai Yeung Yan
Department of Land Surveying and Geo-informatics, The Hong Kong Polytechnic University, Hong Kong
Department of Civil Engineering, Ryerson University, Toronto, Ontario, Canada
(email: waiyeung.yan@polyu.edu.hk, waiyeung.yan@ryerson.ca)

***Abstract*—Correction of the range and angle effects has been proven to be viable to improve the radiometric quality of laser scanning data, regardless of the data collection platform. Existing range normalization (RN) method usually adopts the range to the power of two as a correction factor. Nevertheless, such a setting cannot always guarantee the best improvement of intensity homogeneity, particularly on the recently invented multispectral airborne laser scanner. Therefore, this study aims to propose and intensity correction model for the world's first multispectral airborne laser scanner, Optech Titan. The correction model is built upon the iteratively reweighted least squares, and is able to estimate the correction parameters for the range, scan angle and atmospheric attenuation through using a number of paired closest points from the overlapping data strips. Experimental results show that the correction parameters estimated yielded better than the traditional RN approach with a reduction of coefficient of variation ranging from 2.03-18.35% in channel 1 (1550 nm), 15.5658.42% in channel 2 (1064 nm), and 2.53-29.8% in channel 3 (532 nm) in six land cover classes. Therefore, intensity correction parameters should be estimated from the overlapping laser scanning data strips, regardless of the laser channel being used.**

***Index Terms*—Intensity Correction, Land Cover, Multispectral Airborne Laser Scanning, Range Normalization (RN)**

I. Introduction

Various radiometric calibration and correction models have been developed for multimodal laser scanning point cloud and full-waveform data [1]. Different applications, including land cover classification [2], maize plant detection [3], and estimation of leaf area index and leaf water content [4], have been benefited from the improved radiometric quality. Most of these applications have demonstrated significant improvement in terms of classification accuracy or retrieval of biophysical properties after the radiometric pre-process. Indeed, all these calibration or correction models are mainly built upon a physical model, i.e. radar (range) equation, which describes the relationship between the emitted and recorded laser energy with respect to the factors of range, incidence angle, surface condition, atmospheric attenuation and other system factors [5]. Thus, range normalization (RN) method has been proposed to correct the intensity data (I) [6]–[8], where it simply considers the effect of range but not other paramters. This results in the corrected intensity data (I_c) being directly proportional to the range (R) to the power of a factor f, i.e.

$$I_c = I \cdot \left(\frac{R}{R_r}\right)^f \tag{1}$$

where R_r refers to a reference range value that is commonly set to be the minimum or median range value within the collected data strips, and f value is commonly known to be equivalent to two according to the radar (range) equation [5]. Nevertheless, such a setting only handles simple scenario collected by monochromatic laser scanning system with stable settings. Different research studies have attempted to look for an optimal f value through using a cross-validation approach with a range of pre-defined values, e.g. f equals to 0 to 4.

Korpela [7] studied the use of range normalized intensity data to detect the *Cladina* lichens for understory vegetation. The study concluded that the optimal value of f varied from 2.3 to 2.5 considering the surface types with low shrub–moss–lichen, grass, or barley vegetation or gravel. The RN affected the intensity with up to $\pm$25%, and classification accuracy improvement was noted with 2 to 4%. In the subsequent study [9], similar findings were recorded where $f \approx 2$ was recommended for trees with large or clumped leaves (i.e. birch) and $f \approx 3$ was recommended for diffuse coniferous crowns (i.e. pine or spruce). Tree species classification accuracy can be improved by 2 to 3% after implementing the RN, while 6 to 9% improvement can be achieved if the effect of automatic gain control is considered (for Leica ALS50-II).

Gatziolis [8] investigated the optimal f value through cross validation of the f value from 2 to 3, and subsequently examined the coefficient of variation (cv) of three types of forest canopies (i.e. conifers, mixed and hardwoods). The study found out that by using $f = 2.042$, the cv of forest canopies was reduced up to 55.3% for the single returns, resulting in an improvement of overall classification accuracy from 44.4 to 75.6%. Recently, Kukkonen *et al.* [10] evaluated the optimal f to serve the purpose of tree species prediction in boreal forests using the latest multispectral airborne laser scanner. The study concluded the optimal f values were 1.1, 0.4 and 0 for channel 1 (1550 nm), channel 2 (1064 nm) and channel 3 (532 nm), respectively, which can minimize the intensity difference of the first canopy echoes between flight lines with a percentage of reduction yielded less than 7%.

One can note that the f values are different with respect to objects and the laser scanning systems. This sheds the light that there is no specific f value to be applied to a certain system or study area. In addition, the target (object) size and the laser echoes also affect the optimal f value [9]. Indeed, the use of the cross validation approach (i.e. setting a range of f values for assessing the lowest cv) to look for an optimal f is not practical at all. Therefore, an automatic approach, which can estimate the correction parameters from the overlapping laser scanning data strips, is developed in this study.

II. Proposed Method

Recalling the radar (range) equation [5] that describes the relationship between the recorded laser power P_r with respect to a number of system and environmental factors:

$$P_r = \frac{P_t D_r^2}{4\pi R^4 \beta_t} \eta_{atm} \eta_{sys} \sigma \tag{2}$$

where P_t is the transmitted laser power, D_r is the diameter of the aperture and β_t is the laser beam width. All these three parameters can be assumed as a constant during the flight mission. The parameter η_{atm} describes the laser energy attenuation with respect to the atmospheric condition, which is commonly known as e^{-2cR} [2] (i.e. c is the sum of the molecular/aerosol scattering and absorption). η_{sys} is a transmission factor that is system-dependent. Examples such as the use of automatic gain control [9] in Leica ALS50 or intensity banding [11], all of which should be considered in the correction process. The target cross section σ describes the illuminated surface characteristics, and if Lambertian reflectance mechanism is assumed, σ can be modeled as:

$$\sigma = 4\pi\rho A \cos\theta \tag{3}$$

where ρ refers to the relative spectral (or pseudo) reflectance of the illuminated object, A is the laser footprint (which is a factor of the laser range and beam divergence), and θ refers to the incidence angle. By substituting Eq. 3 into Eq. 2 and removing the constant of the parameters, it becomes

$$I \propto \rho \cdot \frac{1}{R^2} \cdot \cos\theta \cdot \mathrm{e}^{-2cR} \tag{4}$$

When an airborne laser scanner collects a pair of partially overlapping data strips (i.e., $\mathbf{L}_i$ and $\mathbf{L}_j$), a total number of n pairwise closest points can be formed through constructing a kd-tree data structure. Assuming that a specific pair of closest point from the two data strips is backscattered from the same object, the relative spectral (or pseudo) reflectance ρ should be identical. As a result, the Eq. 4 can be re-written as:

$$\frac{I_i}{I_j} = \left(\frac{R_j}{R_i}\right)^2 \cdot \left(\frac{\cos\theta_i}{\cos\theta_j}\right) \cdot \mathrm{e}^{2c(R_j - R_i)} \tag{5}$$

In order to model the complex study scene, a higher order of correction factor is given to reduce the effects of range and incidence angle [12]. As a result, Eq. 6 becomes:

$$\frac{I_i}{I_j} = \left(\frac{R_j}{R_i}\right)^a \cdot \left(\frac{\cos\theta_i}{\cos\theta_j}\right)^b \cdot \mathrm{e}^{2c(R_j - R_i)} \tag{6}$$

The above Eq. 6 can be solved after linearization:

$$\begin{bmatrix} \ln\left(\frac{R_{j,1}}{Ri,1}\right) & \ln\left(\frac{\cos\theta_{i,1}}{\cos\theta_{j,1}}\right) & 2\left(R_{j,1} - R_{i,1}\right) \\ \ln\left(\frac{R_{j,2}}{Ri,2}\right) & \ln\left(\frac{\cos\theta_{i,2}}{\cos\theta_{j,2}}\right) & 2\left(R_{j,2} - R_{i,2}\right) \\ \vdots & \vdots & \vdots \\ \ln\left(\frac{R_{j,n}}{Ri,n}\right) & \ln\left(\frac{\cos\theta_{i,n}}{\cos\theta_{j,n}}\right) & 2\left(R_{j,n} - R_{i,n}\right) \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \ln\left(\frac{I_{i,1}}{I_{j,1}}\right) \\ \ln\left(\frac{I_{i,2}}{I_{j,2}}\right) \\ \vdots \\ \ln\left(\frac{I_{i,n}}{I_{j,n}}\right) \end{bmatrix} \tag{7}$$

or

$$\mathbf{X}\boldsymbol{\beta} = \mathbf{y} \tag{8}$$

To solve the above linearized equation, the below objective function should be minimized:

$$\arg\min \sum_{k=1}^{n} \left(\mathbf{y}_k - \mathbf{X}_k\boldsymbol{\beta}\right)^2 \tag{9}$$

With the appearance of outliers (caused by the mis-match of closest points), a robust regression approach, i.e. iteratively re-weighted least squares together with M-estimator, is adopted. The model can first estimate the parameter $\hat{\boldsymbol{\beta}}$ using traditional least squares estimator:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^\top\mathbf{X})^{-1}\mathbf{X}^\top\mathbf{y} \tag{10}$$

The estimated parameters $\hat{\boldsymbol{\beta}}$ can be re-applied to Eq. 8 in order to calculate the residual $\boldsymbol{e}$, i.e.

$$\mathbf{e} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} \tag{11}$$

The median value of the residual $\hat{\boldsymbol{\beta}}$ is retrieved in order to aid in the determination of a weight matrix for the subsequent iteration.

$$\hat{e} = \mathrm{med}|\mathbf{e}| \tag{12}$$

A new weight $\boldsymbol{\Omega}$ is assigned to the objective function as shown in Eq. 9. In this way, Eq. 9 is re-rewritten as:

$$\arg\min \sum_{k=1}^{n} \boldsymbol{\Omega}\left(\mathbf{y}_k - \mathbf{X}_k\hat{\boldsymbol{\beta}}\right)^2 \tag{13}$$

where

$$\boldsymbol{\Omega} = \begin{bmatrix} \omega_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \omega_n \end{bmatrix} \tag{14}$$

and

$$\omega_k = \begin{cases} 1 & \text{if } e_k \le \psi, \\ \psi/e_k & \text{if } e_k > \psi. \end{cases} \tag{15}$$

where ψ is a turning constant that restricts the outliers. The Huber estimator recommends the $\psi = 1.345 \times \hat{e}/0.6745$ [13]. A new iteration is conducted based on the weighted least squares:

$$\hat{\boldsymbol{\beta}}' = (\mathbf{X}^\top\boldsymbol{\Omega}\,\mathbf{X})^{-1}\mathbf{X}^\top\boldsymbol{\Omega}\,\mathbf{y} \tag{16}$$

The above iteration terminates until the following fulfills:

$$\left|\frac{\hat{\boldsymbol{\beta}}' - \hat{\boldsymbol{\beta}}}{\hat{\boldsymbol{\beta}}}\right| < 0.0001 \tag{17}$$

Once the parameters ($a, b,$ and c) are solved, they can be assigned to the following equation so as to retrieve the relative spectral reflectance as follow:

$$\rho = I \cdot \left(\frac{R}{R_r}\right)^a \cdot \left(\frac{1}{\cos\theta}\right)^b \cdot \mathrm{e}^{2cR} \tag{18}$$

Fig. 1 summarizes the workflow of the proposed correction.

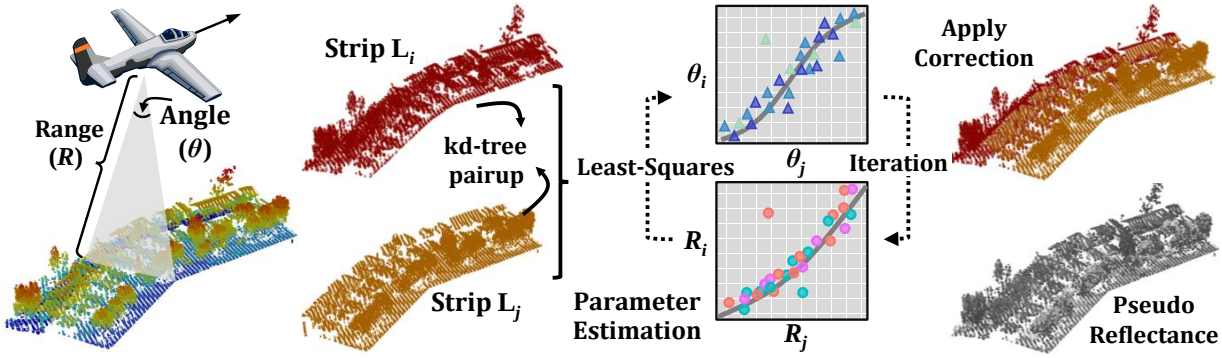

Fig. 1: Overall workflow of the proposed intensity correction.

III. Experiments and Results

A. Multispectral Airborne Laser Scanning Data

The above correction model was applied to a multispectral airborne laser scanning dataset collected by Optech Titan. Optech Titan is operated based on an oscillating mirror mechanism, and is capable of collecting laser scanning data using three laser channels: channel 1 (1550 nm), channel 2 (1064 nm) and channel 3 (532 nm), see Fig. 2a. A total of 33 partially overlapping data strips was collected with the following settings: pulse repetition frequency = 375kHz, scan frequency = 40 Hz, scan angle = $\pm 20°$, and flying height $\approx 1,100$ m. As a result, the mean point density of channel 1 to 3 is 11.9 points/m^2, 12.4 points/m^2, and 4.8 points/m^2, leading to an approximate 0.5 m of mean point spacing. No modulation of dynamic gain was applied. Since there exist a notable banding pattern on the first two data channels, a LiDAR scan line correction [11] was implemented in order to remove the banding effect. Subsequently, a number of study areas were selected to evaluate the effect of intensity correction, and these areas cover shrubland, bare ground, trees, grass, and rooftops.

B. Improvement of Intensity Homogeneity After Correction

The proposed correction reduced the cv of all the six land cover classes regardless of the data channels (see table I). The unpaved road recorded a reduction of cv after correction. The respective cv of the three channels decreased from 0.161 to 0.144 ($\downarrow 10.42\%$) in channel 1, 0.241 to 0.121 ($\downarrow 49.92\%$) in channel 2, and 0.137 to 0.01 ($\downarrow 27.14\%$) in channel 3. The house rooftop also recorded a notable reduction of cv, where the values were 0.141, 0.244 and 0.101 in the original intensity (OI). These became 0.127 ($\downarrow 9.78\%$), 0.124 ($\downarrow 49.34\%$) and 0.071 ($\downarrow 29.8\%$) after correction.

Features with multiple returns, such as tree canopies or shrubland, received the least cv reduction comparing to the rest of the land cover classes. For instance, the values of cv computed in tree canopies were 0.415, 0.356 and 0.186, respectively, for the three channels in the OI. After implementing the correction, the cv was reduced by 2.03% in channel 1, 15.56% in channel 2, and 2.53% in channel 3, resulting in the respective cv values being 0.407, 0.401 and 0.181. Regarding the shrubland, a decrease of cv was noted from 0.27 to 0.254 after implementing the correction in channel 1, leading to a reduction rate of 5.88%. In channels 2 and 3, the cv dropped from 0.356 to 0.291 ($\downarrow 18.28\%$) and from 0.152 to 0.13 ($\downarrow 14.26\%$), respectively.

On the other hand, features with high relative spectral reflectance in the corresponding laser wavelength recorded a significant reduction of cv. For instance, bare ground and grass cover all received a large reduction of cv in the three channels after correction. Sample data points of grass cover recorded a decrease of cv by 11.94% in channel 1, 58.42% in channel 2, and 25.37% in channel 3. Similarly, bare ground demonstrated a decrease of cv after correction by 18.35% in channel 1, 55.65% in channel 2, and 21.98% in channel 3. With the improved radiometric quality, the notable stripe artifacts found in the overlapping swaths were significantly removed after the intensity correction (Fig. 2b).

C. Comparison between RN and Proposed Correction

The proposed correction results were also compared with the same set of multispectral laser scanning data processed with the traditional RN with values reported in the literature [7]–[10]. The range correction factor (f) being examined was 1.1 for channel 1 (1550 nm), where channel 2 (1064 nm) was examined with a number of f values, including 0.4, 2, 2.042, 2.4, 2.5 and 3. Although the cv was significantly reduced after implementing the RN, none of the results outperformed the proposed correction as shown in Table I. The proposed correction consistently recorded the least cv values on all land cover types regardless of the RN parameters.

IV. Conclusions

This paper presents an intensity correction model for the latest multispectral airborne laser scanning data with three laser channels. The proposed correction was able to improve the intensity homogeneity and outperformed the traditional RN method by comparing the cv of six land cover classes, i.e., shrubland, bare ground, trees, grass, and rooftops. The proposed correction was able to reduce the cv by 2.03–18.35% in channel 1 (1550 nm), 15.56–58.42% in channel 2 (1064 nm), and 2.53–29.8% in channel 3 (532 nm). In addition, the proposed correction, embedded with the capability of outliers resistance, can robustly estimate the correction parameters.

TABLE I: Coefficient of variation (cv) of six land cover types after implementing RN and the proposed correction.

Data	Wavelength (nm)	Channel	Ground	Rooftop	Road	Grass	Shrub	Tree
Original Intensity	1550	1	0.130	0.141	0.161	0.255	0.270	0.415
Range normalization (f = 1.1) [10]	1550	1	0.112	**0.127**	0.149	0.236	0.255	**0.407**
Proposed intensity correction	1550	1	**0.106**	**0.127**	**0.144**	**0.224**	**0.254**	**0.407**
Original Intensity	1064	2	0.206	0.244	0.241	0.240	0.356	0.475
Range normalization (f = 0.4) [10]	1064	2	0.140	0.152	0.151	0.135	0.315	0.411
Range normalization (f = 2.0) [9]	1064	2	0.112	0.138	0.134	0.113	0.302	0.405
Range normalization (f = 2.042) [8]	1064	2	0.111	0.137	0.134	0.113	0.302	0.405
Range normalization (f = 2.4) [7]	1064	2	0.106	0.135	0.131	0.109	0.299	0.404
Range normalization (f = 2.5) [7]	1064	2	0.105	0.134	0.131	0.108	0.299	0.404
Range normalization (f = 3.0) [9]	1064	2	0.099	0.131	0.127	0.105	0.296	0.403
Proposed intensity correction	1064	2	**0.091**	**0.124**	**0.121**	**0.100**	**0.291**	**0.401**
Original Intensity	532	3	0.180	0.101	0.137	0.189	0.152	0.186
Range normalization (f = 0) [10]	532	3	0.180	0.101	0.137	0.189	0.152	0.186
Proposed intensity correction	532	3	**0.141**	**0.071**	**0.100**	**0.141**	**0.130**	**0.181**

*Bold items refer to the largest reduction of cv, which also indicate the best improvement of intensity homogeneity.

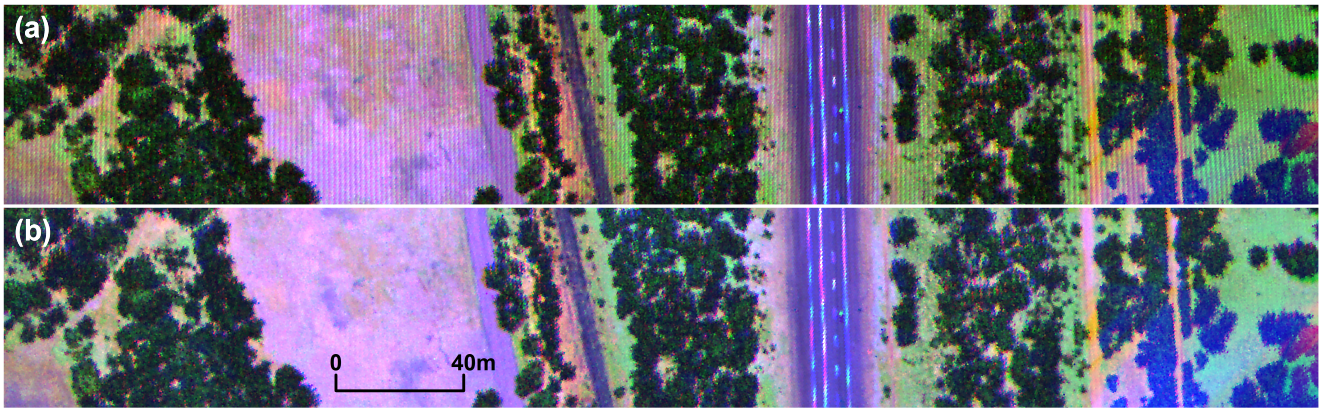

Fig. 2: An example of two overlapping swaths with a large scan angle at top and the other with a small scan angle (near nadir region) at bottom: (a) original intensity, and (b) the proposed correction.

Based on the experiment, it is concluded that the correction parameters are data-driven and system-dependent. Therefore, it is recommended not to label a specific value of correction parameter for a certain type of study site/land cover or a specific laser scanning system, similar to the way of traditional RN. Instead, the correction parameters should be estimated in each flight mission based on the use of the overlapping data strips. In the future, different BRDFs [12], [14] should be examined in the proposed framework. Similar to satellite remote sensing, a stratified strategy, such as the use of spectral indices, number of returns, etc., can be further explored.

REFERENCES

[1] A. Kashani, M. Olsen, C. Parrish, and N. Wilson, "A review of LiDAR radiometric processing: from ad hoc intensity correction to rigorous radiometric calibration," *Sensors*, vol. 15, pp. 28099–28128.

[2] W.Y. Yan, A. Shaker, A. Habib, and A. Kersting, "Improving classification accuracy of airborne LiDAR intensity data by geometric calibration and radiometric correction,". *ISPRS J. of Photogram. and Remote Sens.*, vol. 67, 2012, pp. 35–44.

[3] B. Höfle, "Radiometric correction of terrestrial LiDAR point cloud data for individual maize plant detection". *IEEE Geosci. and Remote Sens. Lett.*, vol. 11, 2014, pp. 94–98.

[4] H. You, T. Wang, A. Skidmore, and Y. Xing, "Quantifying the effects of normalisation of airborne LiDAR intensity on coniferous forest leaf area index estimations". *Remote Sens.*, vol. 9, 2017, pp. 163.

[5] A. Jelalian, *Laser Radar Systems*. Boston: Artech House, 1992.

[6] C. Hopkinson, "The influence of flying altitude, beam divergence, and pulse repetition frequency on laser pulse return intensity and canopy frequency distribution," *Canadian J. of Remote Sens.*, vol. 33, 2007, pp. 312–324.

[7] I. S. Korpela, "Mapping of understory lichens with airborne discrete-return LiDAR data," *Remote Sens. Environ.*, vol. 112, 2008, pp. 3891–3897.

[8] D. Gatziolis, "Dynamic range-based intensity normalization for airborne, discrete return lidar data of forest canopies," *Photogramm. Eng. & Remote Sens.*, vol. 77, 2011, pp. 251–259.

[9] I. Korpela, *et al.*, "Range and AGC normalization in airborne discrete-return LiDAR intensity data for forest canopies,"*ISPRS J. of Photogramm. and Remote Sens.*, vol. 65, 2010, pp. 369–379.

[10] M. Kukkonen, M. Maltamo, L. Korhonen, and P. Packalen, "Multi-spectral airborne LiDAR data in the prediction of boreal tree species composition," *IEEE Trans. on Geosci. and Remote Sens.*, vol. 57, 2019, pp. 3462–3471.

[11] W.Y. Yan and A. Shaker, "Airborne LiDAR intensity banding: Cause and solution," *ISPRS J. of Photogramm. and Remote Sens.*, vol. 142, 2018, pp. 301–310.

[12] Q. Ding, *et al.*, "Combination of overlap-driven adjustment and Phong model for LiDAR intensity correction," *ISPRS J. of Photogramm. and Remote Sens.*, vol. 75, 2013, pp. 40–47.

[13] P. J. Huber, "Robust estimation of a location parameter," *The Annals of Mathematical Statistics*, 1964, vol. 35, pp. 73-101.

[14] D. Bolkas, "Terrestrial laser scanner intensity correction for the incidence angle effect on surfaces with different colours and sheens," *Int. J. of Remote Sens.*, vol. 40, 2019, pp.7169–7189.