Radiometric Correction and Normalization of Airborne LiDAR Intensity Data for Improving Land-Cover Classification

Wai Yeung Yan, Member, IEEE, and Ahmed Shaker, Member, IEEE

Abstract—Radiometric correction of airborne LiDAR intensity data has been proposed based on the use of the radar (range) equation for removing the effects of attenuation due to system- and environmental-induced distortions. Although radiometric correction of airborne LiDAR intensity data has been recently investigated with results revealing improved accuracy of surface classification, there exist a few voids requiring further research effort. First, the variation of object surface characteristics (slope and aspect) plays a crucial role in modeling the recorded intensity data, and thus, the laser incidence angle is usually considered in the correction process. Nevertheless, the use of incidence angle would lead to the effects of overcorrection, particularly on those features located in steep slope. Second, line-stripping problems are usually appeared in the overlapping region of LiDAR data strips acquired by sensors configured with automatic gain control (AGC). Currently, the effects of AGC cannot be perfectly modeled due to the nondisclosure of information by the sensor manufacturers. In this paper, we attempt to fill these voids by: 1) proposing a correction mechanism using the surface slope as a threshold to select either using scan angle or incidence angle in the radar (range) equation; and 2) proposing a subhistogram matching technique to radiometrically normalize the overlapping intensity data. The proposed approaches were applied to three real airborne LiDAR data strips for experimental testing. The results showed that the coefficient of variation reached to the lowest value for most of the land-cover features with a slope threshold between $30^\circ$ and $40^\circ$. The variance-to-mean ratio of five land-cover features was significantly reduced by 70%–82% after applying the proposed correction mechanism. In addition, the systematic noises appeared in the overlapping region were significantly reduced after radiometric correction and normalization, where the overall accuracies in the overlapping region were significantly reduced after radiometric correction mechanism. Nevertheless, the use of incidence angle would lead to the effects of overcorrection, particularly on those features located in steep slope.

Index Terms—Airborne LiDAR, Gaussian mixture model (GMM), incidence angle, intensity, land-cover classification, land-cover homogeneity, radiometric correction, radiometric normalization, subhistogram matching.

I. INTRODUCTION

AERBORNE LiDAR has been extensively used for digital elevation/surface modeling [1], [2], topographic mapping [3], [4], building/road features recognition [5], [6], forestry mapping and assessment [7], [8], power line extraction [9], and 3-D city modeling [10]–[12]. When an airborne LiDAR system flies over a survey area, the LiDAR sensor emits laser pulses at a specific pulse repetition frequency, illuminates the surface objects, and records the returned laser pulse signals after backscattered from the surface objects. Single return of laser pulses can always be found in impenetrable surfaces such as roofs and roads, whereas multiple returns of laser pulses are usually obtained from tree canopies, and a portion of the laser footprint is illuminated on and backscattered from the leaves, stems, and branches before encountering the ground surface [13]. For the discrete-return LiDAR sensor, the intensity data represent the peak amplitudes recorded in the laser backscattering beam return from the illuminated object, where the intensity is usually linearized into an 8- or 11-bit data scale. For the latest full-waveform LiDAR, the sensor not only records a discrete number of echoes but also digitizes the entire waveform of the emitted pulse and the backscattered echoes [14]. As commercial topographic LiDAR sensors usually utilize the Nd:YAG laser, which operates at 1.064-μm wavelength, high separability of spectral reflectance can always be found among different land-cover materials in the near-infrared spectrum [15], [16].

Airborne LiDAR data acquisition is usually carried out with several overlapping scans in order to serve large-scale seamless mapping. Such scanning configuration can apparently increase the density of acquired data point cloud and compensate the lack of scanning surface from different angles of view. Nevertheless, combining overlapping strips of airborne LiDAR intensity data is not just simply mosaicking the data based on the geometric information $xyz$. Direct mosaicking the multiple strips without proper treatment would result in visually detrimental to the LiDAR intensity data, particularly those data acquired by LiDAR sensor configured with automatic gain control (AGC). The result of combining intensity data includes a significant line-stripping problem, which represents a source of systematic noise (see Fig. 1). The radiometric heterogeneities, which appear both within the single strip and among different strips, can be ascribed mainly due to system- and environmental-induced distortions. These thus degrade the intensity image quality and the performance of any thematic applications. As such, radiometric correction and normalization are required to resolve the effects of the line-stripping problem.

Manuscript received June 19, 2013; revised December 22, 2013, February 15, 2014, and March 26, 2014; accepted April 4, 2014. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Discovery Grant and in part by the GEOIDE Canadian Network of Excellence, Strategic Investment Initiative project SII P-IV #72.

The authors are with the Department of Civil Engineering, Ryerson University, Toronto, ON M5B 2K3, Canada (e-mail: waiyeung.yan@ryerson.ca; ahmed.shaker@ryerson.ca).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2014.2316195
A physical approach of radiometric correction has been recently proposed and addressed by [14] and [17]–[19] based on the use of radar (range) equation [20]. The correction model considers system parameters (sensor-dependent) and environmental parameters (location-dependent) in order to convert the recorded intensity data into its corresponding spectral reflectance value. Absolute calibration may be required if physical spectral reflectance is desired for the study surface by using natural or man-made reference targets as reported in [21]–[23]. The significance of radiometric correction on airborne LiDAR intensity data has been proven recently. Korpela et al. [24] conducted tree classification using discrete-return LiDAR intensity data. Accuracy improvement with 6%–9% was found after normalizing the range and AGC effects on the intensity data. Habib et al. [25] and Yan et al. [16] compared different land-cover classification scenarios using the airborne LiDAR intensity data in an urban area. Although the overall accuracy varied from 30% to 63%, an accuracy improvement was found by 7%–12% using the radiometrically corrected LiDAR intensity data. Despite these attempts, there are a few current voids requiring further research efforts to consolidate the correction approach.

One of the crucial factors in the correction model is the condition of surface topography, which influences the effects of incidence angle. The variation of topography (slope and aspect) affects the size of projected laser footprint [26] and backscattered laser pulse [27]. Therefore, consideration of surface normal is needed to determine the laser incidence angle regardless of the reflection models [28], [29]. Kuukko et al. [27] and Kaasalainen et al. [30] investigated the effects of incidence angle by conducting intensity measurement with different reference targets in the laboratory and field. It was found that the backscattered intensity data vary and mainly depend on specific target surface properties, instead of entirely following Lambert’s cosine law. This phenomenon has been also justified in our previous experiments in radiometric correction of airborne LiDAR intensity data [16], [31], where most of the tree canopies and building boundaries receive excessive correction. The excessive correction is most pronounced with targets located in steep slope, leading to the incidence angle approaching 90°. View shadow occurs if the incidence angle exceeds 90°, and thus, these slopes are occluded and produce no return [32].

Gaulton et al. [33] further mentioned that correcting the effects of incidence angle is extremely difficult for vegetation canopies using single-wavelength LiDAR intensity data.

Apart from the incidence angle, the effects of AGC have not been fully studied and modeled. The AGC aims to control the range of the recorded intensity within the radiometric resolution, for instance, 0 to 255 for an 8-bit intensity data. Vain et al. [34] discussed the principle of AGC by describing the process as: a) The LiDAR sensor emits laser pulse on a low reflective object (e.g., water bodies), a null return may be recorded, and thus, the AGC would be automatically increased; b) when the LiDAR sensor flies over highly reflective objects, the recorded backscattered laser energy may go beyond the range of the 8-bit-intensity-data capacity, and thus, the AGC would be increased. Although few proprietary software packages bundled with LiDAR sensors are used to normalize the intensity value from AGC-on to AGC-off, the method of AGC is not disclosed by the sensor manufacturers, where the technique to remove the effects of AGC remains questionable. Recently, Korpela et al. [24] and Vain et al. [34] attempted to model the effects of AGC by using two different linear empirical models, which shed the light on the black box of AGC. Based on the empirical formulas, the effects of AGC are controlled by a scale factor and a shift (offset). As a result, the transformation equation between the intensity data with AGC-on and AGC-off should be in a form of piecewise linear function. Although fitting a piecewise linear function in a joint intensity histogram (or comparison) between two identical images has been demonstrated for color mapping in computer vision studies [35], [36], such an approach is not viable in airborne LiDAR intensity data since the LiDAR footprints between two strips are not projected at the exact location.

In this paper, we present an approach to correct the effects of incidence angle by imitating the technique of [37], which incorporates a threshold slope value to select either using scan angle or incidence angle in the radiometric correction process, as presented in Section II. To deal with the line-stripping problem in the overlapping LiDAR intensity data, we propose a radiometric normalization model to radiometrically align the overlapping intensity data among different LiDAR data strips by using a Gaussian mixture model (GMM)-based subhistogram matching technique in Section III. We implemented the proposed approaches with three LiDAR data strips, as described in Section IV and assessed the land-cover homogeneity using different statistical evaluation methods, as presented in Section V. Finally, we demonstrate the effects of the proposed radiometric correction and normalization on different LiDAR intensity data classification scenarios in Section VI. Our ultimate goal aims to maximize the benefits of using airborne LiDAR intensity data for object recognition and surface classification and pave the way for the development of future LiDAR data processing systems.

II. RADIOMETRIC CORRECTION

A. Overall Workflow

Fig. 2 illustrates the overall workflow for radiometric correction of LiDAR intensity data. The time-tagged 3-D data...
point cloud is usually delivered in LAS file format together with a GPS trajectory file (optional) after an airborne LiDAR survey. The LAS file contains the 3-D coordinates $xyz$ and the backscattered intensity $I$ of each laser pulse. With the GPS trajectory data and the time-tagged 3-D data point cloud, instantaneous GPS coordinates were interpolated for each of the laser pulses. After this, system parameters (i.e., range, horizontal angle, and scan angle), which describe the geometric relationship between the instantaneous position of the aircraft and the illuminated object, were computed. The incidence angle, which is the angle between the incidence laser pulse and the surface normal from the topography, was computed to consider the topographic induced distortion. Atmospheric attenuation was determined from the weather information (i.e., temperature, pressure, and meteorological visibility) before radiometric correction. Finally, the intensity, range, atmospheric attenuation, and incidence angle for each laser pulse were imported into the radar (range) equation to determine the spectral reflectance of illuminated object, which was regarded as the radiometrically corrected intensity data.

### B. Radar (Range) Equation

The radar (range) equation [20] describes the physical properties of the laser beam energy with respect to the sensor configuration and different environmental parameters. Such equation has been adopted as the radiometric correction/calibration model for LiDAR intensity data, as reported in [14] and [17]–[19]. The equation is used to convert the LiDAR intensity data (which is directly proportional to the amount of received laser energy $P_t$) into the spectral reflectance $\rho$, as shown in (1) and (2) as follows:

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

\[
\sigma = 4\pi \rho A \cos \theta
\]

where $P_t$ is the transmitted laser pulse energy, $D_r$ is the diameter of the aperture, $R$ is the range, $\beta_t$ is the laser beamwidth, $\eta_{sys}$ is the system factor, and $\eta_{atm}$ is the atmospheric attenuation factor, which is assumed as a constant in some of the previous studies [38], [39]. In this paper, we modeled the atmospheric attenuation based on the Beer–Lambert law

\[
\eta_{atm} = e^{-2\tau R}
\]

where $\tau$ is the summation of aerosol scattering, molecular scattering, aerosol absorption, and molecular absorption. Details of atmospheric correction can be found in [16]. The target cross section $\sigma$ consists of the illuminated surface characteristics, including the spectral reflectance $\rho$, the projected target area to the direction of the laser beam $A$, and the direction of reflection, which is determined by the angle $\theta$ between the LiDAR sensor and the target. In case of inclined surface, the cosine $\theta$ should be the cosine of the angle of reflection $\theta_r$, which is the angle between the surface normal and the incidence laser pulse. Nevertheless, only few of the previous research works emphasized the importance of the computation of the reflection angle in the correction process [29]. The following section describes the mathematical procedures to calculate the laser incidence angle.

### C. Computation of Laser Incidence Angle

Fig. 3 presents the geometric relationship between the instantaneous position of the LiDAR sensor $L$ and the object on the ground $P$ in an $XYZ$ Cartesian coordinate system. The scan angle is denoted with $\theta$, and the distance between the sensor and the ground object is represented by the range $R$. In case of flat terrain, the incidence angle is described between two vectors, which are the vertical vector from the ground object $PV$ and the range vector $PL$, where the incidence angle is equivalent to the scan angle $\theta$. In case of rugged terrain, the ground object $P$ is located on a surface with slope $\alpha$ and aspect $\beta$, where the incidence angle should be described with the range vector $PL$ and the surface normal vector $PN$ to the ground object.

In common practice, the range data are usually not included in the LAS file; the range vector $PL$ should be calculated by the instantaneous 3-D coordinates of the aircraft and the 3-D coordinates of the illuminated object for each laser pulse. Since the laser pulse is recorded in nanoseconds (ns), which is not synchronized with the GPS measurement (measured in seconds), the instantaneous 3-D coordinates of the aircraft $(X_L, Y_L, Z_L)$ can be projected on the LiDAR 3-D data point.
cloud \((X_P, Y_P, Z_P)\) by interpolating the GPS time into the corresponding time of LiDAR data. Finally, the instantaneous range \(R\) and scan angle \(\theta\) can be determined using the 3-D coordinates of the aircraft and point clouds. In order to compute the incidence angle \((\theta_r\) or \(\angle LPN)\), three vectors \((\vec{PL}, \vec{PN}, \text{and} \vec{LN})\) in the triangle \(\triangle LPN\) should be used. The slope \(\alpha\) and aspect \(\beta\) can be derived directly from a triangulated irregular network (TIN) model generated from the first return of the LiDAR data using Delaunay triangulation [40]. Then, the vector \(\vec{PN}\) can be calculated using the following equations.

1) In \(\triangle LV\) \(P\):

\[
V P = R \cdot \cos \theta, \quad (4)
\]

2) In \(\triangle NV P\):

\[
PN = \frac{VP}{\cos \alpha}, \quad (5)
\]

3) Substituting (4) into (5)

\[
PN = \frac{R \cdot \cos \theta}{\cos \alpha}. \quad (6)
\]

On the other hand, the vector \(\vec{LN}\) can be calculated using the following equations:

1) In \(\triangle LV P\):

\[
LV = R \cdot \sin \theta. \quad (7)
\]

2) In \(\triangle NV P\):

\[
NV = VP \tan \alpha. \quad (8)
\]

3) Substituting (4) into (8)

\[
NV = R \cdot \cos \theta \tan \alpha. \quad (9)
\]

In \(\triangle NLV\) (see Fig. 3, bottom right):

\[
\angle NVL = \angle Y'VL - \angle Y'VN. \quad (10)
\]

where angle \(\angle Y'VN\) is an obtuse angle that is equal to the aspect \(\beta\) of the ground object \(P\) on the terrain. Angles \(\angle Y'VL\) and \(\angle Y'LV\) are interior angles, where \(\angle YLV\) is the projected horizontal angle between the \(Y\)-axis and the laser pulse. The projected horizontal angle \(\angle YLV\) (or \(\theta_h\)) can be computed using the plane coordinates of the laser pulse and the instantaneous position of the aircraft as follows:

\[
\theta_h = \angle YLV = \tan^{-1} \left[ \frac{X_P - X_L}{Y_P - Y_L} \right]. \quad (11)
\]

According to the cosine law

\[
LN^2 = NV^2 + LV^2 - 2(NV)(LV)(\cos \angle NVL). \quad (12)
\]

Substituting (7) and (9) into (12)

\[
LN^2 = (R \cdot \cos \theta \tan \alpha)^2 + (R \cdot \sin \theta)^2
- 2(R \cdot \cos \theta \tan \alpha)(R \cdot \sin \theta)(\cos(\beta - \theta_h)). \quad (13)
\]

Finally, the incidence angle \(\angle LPN\) can be calculated using the three vectors \((\vec{PL}, \vec{PN}, \text{and} \vec{LN})\) in \(\triangle LPN\) in accordance to the cosine law.

\[
\angle LPN = \theta_r = \cos^{-1} \frac{PN^2 + PL^2 - LN^2}{2(PN)(PL)} \quad (14)
\]

where \(\vec{PL}\) is the range vector, \(\vec{PN}\) can be obtained from (6), and \(\vec{LN}\) can be obtained from (13). By combining all the equations, (14) becomes

\[
\theta_r = \cos^{-1} \left[ \frac{(R \cdot \cos \theta)^2 + R^2 - (R \cdot \cos \theta \tan \alpha)^2}{2 \left( \frac{R \cdot \cos \theta}{\cos \alpha} \right) (R)} \right] + (-R \cdot \sin \theta)^2 + 2(R \cdot \cos \theta \tan \alpha)(R \cdot \sin \theta)(\cos(\beta - \theta_h)) \right] \quad (15)
\]

Finally, the incidence angle is represented as

\[
\theta_r = \cos^{-1} \left[ \cos \theta \cos \alpha + \sin \theta \sin \alpha \cos(\beta - \theta_h) \right]. \quad (16)
\]

D. Combining the Use of Laser Incidence and Scan Angles in Radiometric Correction

Effects of overcorrection have been reported in [41] and [42] while using the incidence angle in topographic correction of passive remote sensing images. Since the cosine of incidence angle is commonly assumed to be indirectly proportional to the corrected intensity (or the spectral reflectance) in the correction process, excessive correction is most pronounced at the incidence angle approaching 90° [42]. This phenomenon has been justified in our previous experiment in radiometric correction of airborne LiDAR intensity data [16], [31], where most of the trees and building boundaries receive excessive correction. To resolve this problem, we imitated the technique proposed in [37], which incorporates threshold angle values to decide the use of different parameters in the correction model for satellite remote sensing images. In this paper, we used the surface slope as a control for the selection of angle in the correction process.

When the slope was less than or equivalent to \(T\), then the incidence angle was used in the radar (range) equation; while the slope exceeded \(T\), the scan angle was used. As such, the threshold value was adopted in the correction process, and (2) is rewritten as

\[
\sigma = \begin{cases} 4\pi \rho A \cos(\theta_r), & \alpha \leq T \\ 4\pi \rho A \cos(\theta), & \alpha > T. \end{cases} \quad (17)
\]

III. Radiometric Normalization

A. Overall Workflow

After applying the proposed radiometric correction for individual LiDAR data strip, the intensity heterogeneity may still appear in the overlapping region of LiDAR data strips due to the effects of AGC. This section presents a radiometric normalization model to adjust the radiometric misalignment by matching the intensity subhistogram of a data strip with reference to the intensity subhistogram generated from a reference data strip. First, overlapping areas within the first and the second LiDAR data strips were identified in those regions, which are
preferable to have a variety of land-cover features covering a wide range of intensity values. Histograms \((H_1 \text{ and } H_2)\) of the overlapping LiDAR data strips were individually generated from the intensity data \((I_1 \text{ and } I_2)\). GMM technique was applied to the histogram in order to fit a Gaussian component for each individual subhistogram. The intersection points, which were used to partition the histogram into subhistograms, were derived by finding the intersection points of the adjacent Gaussian components. The process was then repeated for all the histograms generated from the overlapping LiDAR intensity data strips. Finally, subhistogram matching was carried out based on the histogram equalization techniques to normalize the intensity of the second LiDAR data strip with reference to the intensity of the first LiDAR data strip. In the following sections, we follow the notation of GMM, as defined in [43] for presenting the mathematical model of radiometric normalization.

### B. Gaussian Mixture Model

Consider a LiDAR data set \(X\) with \(N\) number of points; \(X = \{x_1, x_2, \ldots, x_n, \ldots, x_N\}\), where \(1 \leq n \leq N\). For an 8-bit LiDAR data, intensity value \(I\) lies between 0 and 255. The intensity of LiDAR data point \(x_n\) is denoted as \(I(x_n)\). Let \(n_i\) be the number of LiDAR data points with intensity \(I\). The probability density function of \(n_i\) over \(N\) is defined by

\[
P_I = \frac{n_i}{N}. \quad (18)
\]

Given a histogram of the LiDAR data set \(X\), we need to partition the histogram into \(K\) subhistograms. Let \(i_1, i_2, \ldots, i_{K-1}\) and \(i_{K-1}\) be the \(K-1\) intersection points that partition the histogram with \(2 \leq K \leq 255\). The cumulative density function of each subhistogram is calculated by

\[
c_1 = \sum_{l=0}^{i_1} P_I, \quad c_2 = \sum_{l=i_1+1}^{i_2} P_I, \ldots, c_K = \sum_{l=i_{K-1}+1}^{255} P_I. \quad (19)
\]

As such, the sum of the cumulative density functions would be

\[
\sum_{k=1}^{K} c_k = 1 \quad \text{and} \quad 0 \leq c_k \leq 1. \quad (20)
\]

Based on (19), the mean of each subhistogram can be expressed as

\[
\mu_1 = \sum_{l=0}^{i_1} \frac{l}{c_1} P_I, \quad \mu_2 = \sum_{l=i_1+1}^{i_2} \frac{l}{c_2} P_I, \ldots, \mu_K = \sum_{l=i_{K-1}+1}^{255} \frac{l}{c_K}. \quad (21)
\]

The variance for each of the \(K\) subhistograms can be calculated by

\[
\sigma_1^2 = \sum_{l=0}^{i_1} \frac{(l - \mu_1)^2}{c_1} P_I, \quad \sigma_2^2 = \sum_{l=i_1+1}^{i_2} \frac{(l - \mu_2)^2}{c_2} P_I, \ldots, \sigma_K^2 = \sum_{l=i_{K-1}+1}^{255} \frac{(l - \mu_K)^2}{c_K} P_I. \quad (22)
\]

Commonly, a histogram is in a form of multimodal distribution, which can be regarded as a GMM. A GMM is a parametric statistical model that assumes that the data originate from a weighted sum of several Gaussian components. The probability of the LiDAR data point \(x_n\), with respect to the \(k\)th Gaussian component, is defined as

\[
G(I(x_n), \mu_k, \sigma_k^2) = \frac{1}{\sqrt{2\pi\sigma_k^2}} \exp \left[ -\frac{(I(x_n) - \mu_k)^2}{2\sigma_k^2} \right]. \quad (23)
\]

The GMM for the intensity histogram of data point \(x_n\) is a weighted sum of the individual Gaussian components, and GMM is defined as follows:

\[
P(I(x_n)) = \sum_{k=1}^{K} \alpha_k G(I(x_n), \mu_k, \sigma_k^2) \quad (24)
\]

where \(\alpha_k\) denotes the weight of the \(k\)th Gaussian function under the condition \(\alpha_1 + \alpha_2 + \cdots + \alpha_k + \cdots + \alpha_K = 1\) and \(0 \leq \alpha_k \leq 1\). The expectation–maximization (EM) algorithm is commonly used to estimate the set of parameters, i.e., mean \(\{\mu_1 \ldots \mu_K\}\), variance \(\{\sigma_1^2 \ldots \sigma_K^2\}\), and weight \(\{\alpha_1 \ldots \alpha_K\}\) in (24) through an iterative process.

#### Step 1) Partition the entire histogram into \(K\) subhistograms with equal range. Compute the mean \(\mu_1 \ldots \mu_K\) and variance \(\sigma_1^2 \ldots \sigma_K^2\) for each subhistogram using (21) and (22). Assume the weight \(\alpha_K\) as 1/K.

#### Step 2) Compute the new weight \(\alpha_{k,\text{new}}\) for each subhistogram using the following equation:

\[
\alpha_{k,\text{new}} = \frac{1}{N} \sum_{n=1}^{N} \frac{\alpha_k G(I(x_n), \mu_k, \sigma_k^2)}{\sum_{k=1}^{K} \alpha_k G(I(x_n), \mu_k, \sigma_k^2)}. \quad (25)
\]

#### Step 3) Compute the new mean \(\mu_{k,\text{new}}\) for each subhistogram using the following equation:

\[
\mu_{k,\text{new}} = \frac{\sum_{n=1}^{N} \frac{\alpha_k G(I(x_n), \mu_k, \sigma_k^2)}{\sum_{k=1}^{K} \alpha_k G(I(x_n), \mu_k, \sigma_k^2)} I(x_n)}{\sum_{n=1}^{N} \frac{\alpha_k G(I(x_n), \mu_k, \sigma_k^2)}{\sum_{k=1}^{K} \alpha_k G(I(x_n), \mu_k, \sigma_k^2)}}. \quad (26)
\]

#### Step 4) Compute the new variance \(\sigma_{k,\text{new}}^2\) for each subhistogram using the following equation:

\[
\sigma_{k,\text{new}}^2 = \frac{\sum_{n=1}^{N} \frac{\alpha_k G(I(x_n), \mu_k, \sigma_k^2)}{\sum_{k=1}^{K} \alpha_k G(I(x_n), \mu_k, \sigma_k^2)} (I(x_n) - \mu_k)^2}{\sum_{n=1}^{N} \frac{\alpha_k G(I(x_n), \mu_k, \sigma_k^2)}{\sum_{k=1}^{K} \alpha_k G(I(x_n), \mu_k, \sigma_k^2)}}. \quad (27)
\]

#### Step 5) Check the difference of the new and previous values. If \(|\alpha_{k,\text{new}} - \alpha_k| < \text{threshold}\) and \(||\mu_{k,\text{new}} - \mu_k|| < \text{threshold}\) and \(||\sigma_{k,\text{new}}^2 - \sigma_k^2|| < \text{threshold}\), then the process stops; else, go to step 2. The threshold used in this paper is \(10^{-3}\).
\section{C. Histogram Partition}

After fitting the Gaussian component for each subhistogram, the next step is to partition the entire histogram based on the intersections of all the pairwise adjacent Gaussian components. The intersection point can be found by equaling the function of any adjacent Gaussian components. Mathematically, it can be solved by

\[ \alpha_k G \left( I, \mu_k, \sigma_k^2 \right) = \alpha_{k+1} G \left( I, \mu_{k+1}, \sigma_{k+1}^2 \right) \]

or equivalently

\[
\frac{\alpha_k}{\sqrt{2\pi \sigma_k^2}} \exp \left[ \frac{-(I - \mu_k)^2}{2\sigma_k^2} \right] = \frac{\alpha_{k+1}}{\sqrt{2\pi \sigma_{k+1}^2}} \exp \left[ \frac{-(I - \mu_{k+1})^2}{2\sigma_{k+1}^2} \right] 
\]

\[
\frac{\alpha_k \sigma_k^2}{\alpha_{k+1} \sigma_{k+1}} \ln \left[ \frac{\alpha_k \sigma_{k+1}}{\alpha_{k+1} \sigma_k} \right] = \frac{-(I - \mu_{k+1})^2}{2\sigma_{k+1}^2} - \frac{-(I - \mu_k)^2}{2\sigma_k^2} \]

\[
2\sigma_k^2 \sigma_{k+1}^2 \ln \left[ \frac{\alpha_k \sigma_{k+1}}{\alpha_{k+1} \sigma_k} \right] = -\sigma_k^2 (I - \mu_k)^2 + \sigma_{k+1}^2 (I - \mu_{k+1})^2
\]

\[
\sigma_k^2 (I - \mu_k)^2 + 2 \left( \mu_{k+1} \sigma_k^2 - \mu_k \sigma_{k+1}^2 \right) I + \sigma_{k+1}^2 \mu_k^2
\]

\[
-\sigma_k^2 \mu_k^2 - 2\sigma_k^2 \sigma_{k+1}^2 \ln \left[ \frac{\alpha_k \sigma_{k+1}}{\alpha_{k+1} \sigma_k} \right] = 0.
\]

Equation (33) can be represented in the form of

\[ aI^2 + bI + c = 0 \]

where

\[ a = \sigma_{k+1}^2 - \sigma_k^2 \]

\[ b = 2 \left( \mu_{k+1} \sigma_k^2 - \mu_k \sigma_{k+1}^2 \right) \]

\[ c = \sigma_{k+1}^2 \mu_k^2 - \sigma_k^2 \mu_{k+1}^2 - 2\sigma_k^2 \sigma_{k+1}^2 \ln \left[ \frac{\mu_k \sigma_{k+1}}{\mu_{k+1} \sigma_k} \right]. \]

The solution of (34) would be

\[ I = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \]

\section{D. Subhistogram Matching}

Finally, the intersection points of all the pairwise adjacent Gaussian components are computed. Recalling the notation as defined in Section III-B, the entire histogram is partitioned into \( K \) histogram based on the intersection points \( \{0, i_1, i_2, \ldots, i_{K-1}, 255\} \). Assuming that the intensity data \( X_A \) of LiDAR data strip A are normalized with reference to the intensity data \( X_B \) of LiDAR data strip B, the intersection points of both strips’ histogram are denoted as \( \{0, i_1^A, i_2^A, \ldots, i_{K-1}^A, 255\} \) and \( \{0, i_1^B, i_2^B, \ldots, i_{K-1}^B, 255\} \). The subhistogram matching process first computes the cumulative probability density function for each subhistogram of \( X_A \) based on the intersection points

\[ c_1^A = \sum_{i=0}^{i_1^A} P_{i}, c_2^A = \sum_{i=i_1^A+1}^{i_2^A} P_{i}, \ldots, c_K^A = \sum_{i=i_{K-1}^A+1}^{255} P_{i}. \]

The intensity value in each subhistogram of \( X_A \) is transformed to the intensity of the corresponding subhistogram of \( X_B \) by using the histogram equalization technique

\[ f_1[I(x_n)] = 0 + (i_1^B - 0) c_1^A, \]

\[ f_2[I(x_n)] = i_1^B + (i_2^B - i_1^B) c_2^A, \ldots, \]

\[ f_K[I(x_n)] = i_{K-1}^B + (255 - i_{K-1}^B) c_K^A. \]

Combining all the transformation functions, the entire histogram equalization model is represented as

\[ F[I(x_n)] = f_1[I(x_n)] \cup f_2[I(x_n)] \cup \ldots \cup f_K[I(x_n)]. \]

Based on the transformation function \( F \), the intensity data in \( X_A \) are normalized with reference to the intensity data in \( X_B \) where the normalized intensity of \( X_A \) is computed by \( F[I(x_n)] \) \( \forall x_n \in X_A \). Fig. 4 demonstrates a pictogram example for the entire radiometric normalization process.

\section{IV. Experimental Work}

\subsection{A. Study Area and Data Set}

Three LiDAR data strips covering the British Columbia Institute of Technology (BCIT) located in Burnaby, BC, Canada (122°59’ W, 49°15’ N) were acquired for experimental testing. The LiDAR survey was carried out on July 17, 2009, using the Leica ALS50 sensor, which operates in 1.064-\( \mu \)m wavelength, 0.33-mrad beam divergence, 4 maximum returns, and 83-kHz pulse repetition frequency. Details of the LiDAR system settings and the data specification are listed in Table I, and the data specification of the three LiDAR data strips is shown in Table II. The LiDAR flight was carried out in the afternoon (14:00 to 15:30 local time) with a sunny weather. The temperature, pressure, and visibility for both dates were 29.8°C, 101.81 kPa, and 48.3 km, respectively, as delivered by the National Climate Data and Information Archive from Environment Canada. Since the airborne LiDAR survey did not acquire any information regarding the atmospheric conditions, these archived climate data were used for atmospheric correction.

As shown in Fig. 5, the acquired LiDAR data consisted of three data strips. The first and the third strips were scanned from east to west, and the middle strip (second strip) was scanned in reverse direction (west to east). The percentage of overlap in the first two scans (strips 1 and 2) and the last two scans (strips 2 and 3) was about 30\% and 25\%, respectively. The average flying altitude of all scans was approximate 600 m,
resulting in a point density of 4–5 points/m². Digital aerial photos were also collected during the same flight mission. An orthorectified aerial imagery was produced for the three image bands (i.e., red, green, and blue) with 0.5-m spatial resolution. The LiDAR data were delivered with the 3-D data point cloud in LAS format together with the time-tagged GPS trajectory. Since the geometry of the overlapping data strips is critical in order to produce a viable result in radiometric modeling, the airborne LiDAR data were geometrically calibrated with the quasi-rigorous method [44] with consideration of overlapping primitives. The horizontal accuracy and vertical accuracy were \( \pm 8 \) and \( \pm 20 \) mm, respectively [44]. The outcome of the calibration produced a more accurate 3-D point cloud and the derived data product (the range and the scan angle). In addition, the incidence angle of each laser pulse was computed by using the slope and aspect derived from the 3-D surface, as described in Section II.

B. Design of Experiment and Evaluation

Following the method presented in Sections II and III, radiometric correction was first applied to the airborne LiDAR intensity data based on the radar (range) equation, and radiometric normalization was performed on the corrected LiDAR intensity data. The LiDAR data set was displayed in the ESRI ArcGIS 9.3 platform, where the core of the correction model was built by using ArcObjects and Visual Basic applications. To evaluate the effects of the proposed approach in the correction model, we performed radiometric correction on the intensity data using different slope threshold values ranging from 0° to 90°. By interpolating the first-return LiDAR data points in ArcGIS using a “void-filling” method, intensity images were generated and exported in TIFF format with 0.4-m resolution for classification. The “void-filling” method interpolates the void area by computing the intensity values in the void pixels based on the local average of the neighboring intensity pixels with a 3 \( \times \) 3 moving window [45]. Fig. 6(a)–(c) shows the original intensity (OI) data of the three LiDAR data strips, respectively.

After this, we assessed the homogeneity of the intensity data before and after radiometric correction. Although ground truth is a desirable approach for the evaluation, \textit{in situ} measurements by spectroradiometer were not available in this paper. Aerial photos were acquired during the same flight of the LiDAR survey; nevertheless, the photos were not captured in the same wavelength as the LiDAR sensor did. Since identical reference measurements cannot be achieved, we imitated the statistical

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>AIRBORNE LiDAR SYSTEM SETTINGS AND DATA SPECIFICATION</th>
</tr>
</thead>
<tbody>
<tr>
<td>System Parameters</td>
<td>Settings</td>
</tr>
<tr>
<td>LiDAR Sensor</td>
<td>Leica ALS50</td>
</tr>
<tr>
<td>Pulse Repetition Frequency</td>
<td>83 kHz</td>
</tr>
<tr>
<td>Beam Divergence</td>
<td>0.33 mrad</td>
</tr>
<tr>
<td>Wavelength</td>
<td>1.064 ( \mu )m</td>
</tr>
<tr>
<td>Flying Height</td>
<td>( \approx 600 ) m</td>
</tr>
<tr>
<td>Scan Angle</td>
<td>( \pm 30^\circ )</td>
</tr>
<tr>
<td>Data Specification</td>
<td>Settings</td>
</tr>
<tr>
<td>Number of Data Strips</td>
<td>3</td>
</tr>
<tr>
<td>Mean Point Density</td>
<td>( \approx 4.3 ) points/m²</td>
</tr>
<tr>
<td>Mean Point Spacing</td>
<td>( \approx 0.48 ) m</td>
</tr>
<tr>
<td>Footprint Diameter</td>
<td>( \approx 15 ) cm at nadir</td>
</tr>
<tr>
<td></td>
<td>( \approx 20 ) cm at max. extent</td>
</tr>
<tr>
<td>Number of Returns</td>
<td>4</td>
</tr>
<tr>
<td>Intensity</td>
<td>( 0 - 255 ) (8 bit)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE II</th>
<th>DATA SPECIFICATION OF THE THREE LiDAR DATA STRIPS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Strips</td>
<td>Data Strips 1</td>
</tr>
<tr>
<td>Direction</td>
<td>East to West</td>
</tr>
<tr>
<td>Number of Points</td>
<td>1,705,016</td>
</tr>
</tbody>
</table>
evaluation methods adopted in the magnetic resonance image (MRI) correction, as summarized in [46] and [47].

A number of approaches were proposed to evaluate the MRI correction. The first approach is to assess the variance of the entire or partial data set, which is supposed to be reduced after correction. Nevertheless, the results could be misleading as the units of the intensity are different before and after correction (i.e., the OI is equivalent to the received laser power, where the corrected intensity data are directly proportional to the spectral reflectance of the illuminated object).

The second approach is to assess the coefficient of variation $cv$, which is computed by dividing the standard derivation of a class $\omega_i$ by its mean. This approach is scale invariant, which can overcome the limitation of the first method

$$cv(\omega_i) = \frac{\sigma(\omega_i)}{\mu(\omega_i)}. \tag{40}$$

In addition, the variance-to-mean $vmr$ ratio can be used to assess homogeneity, where it can be used to describe the degree of discrepancy of the intensity values within the same type of land-cover feature. Both $cv$ and $vmr$ can serve as an indicator for assessing the within-class variation of land-cover samples

$$vmr(\omega_i) = \frac{\sigma^2(\omega_i)}{\mu(\omega_i)}. \tag{41}$$

Similarly, the coefficient of joint variation $cjv$ between two classes is commonly used to assess the variation between two different classes (i.e., $\omega_i$ and $\omega_j$) by using (42), and the relative change of the coefficient of joint variation can be determined by (43) as follows:

$$cjv(\omega_i, \omega_j) = \frac{\sigma^2(\omega_i) + \sigma^2(\omega_j)}{|\mu(\omega_i) - \mu(\omega_j)|} \tag{42}$$

$$cjv\Delta(\omega_i, \omega_j) = \frac{cjv_{rc}(\omega_i, \omega_j) - cjv_{oi}(\omega_i, \omega_j)}{cjv_{oi}(\omega_i, \omega_j)} \times 100\%. \tag{43}$$

In this context, a smaller $cjv$ corresponds to a less between-class confusion of intensity within a pair of land-cover classes. In this paper, ten evenly distributed target areas were identified for each of the land-cover features (i.e., building, grass, road, soil, and tree) in each LiDAR data strip. These target areas were carefully selected in homogeneous surface with reference to the orthorectified aerial imagery, and these target areas were evenly distributed within each LiDAR data strip, covering from

---

**Fig. 5.** Study area in the BCIT, Burnaby, BC, Canada.

**Fig. 6.** OI data of airborne LiDAR data: (a) strip 1, (b) strip 2, and (c) strip 3. The yellow zone indicates the overlapping region between LiDAR data strips.

**TABLE III**

<table>
<thead>
<tr>
<th>Number of LiDAR Data Points Acquired for Statistical Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Strip 1</td>
</tr>
<tr>
<td>Building</td>
</tr>
<tr>
<td>Grass</td>
</tr>
<tr>
<td>Road</td>
</tr>
<tr>
<td>Soil</td>
</tr>
<tr>
<td>Tree</td>
</tr>
</tbody>
</table>

---
nadir to its maximum coverage. The total size of the target areas in each land-cover type was in proportional to the area of land cover found in the data strip. These target areas were reselected if the land-cover type was mixed or questionable. Finally, the LiDAR data points that fall within the target areas were used to compute \(cv\), \(vmr\), and \(cjv\) for each land-cover type. Table III shows the number of LiDAR data points used for statistical analysis. Fig. 7(a)–(c) shows the distribution of the target sample areas for the five land-cover features in data strips 1 to 3, respectively.

V. RESULTS AND ANALYSIS

A. Determination of Slope Threshold Value for Radiometric Correction

Fig. 8 shows the coefficient of variation \(cv\) of the five land-cover samples, which are collages of corresponding target sample areas extracted from the three LiDAR data strips. Since the purpose of radiometric correction aims to reduce the intensity heterogeneity within-class and between different data strips, low \(cv\) value is desired for all the land-cover samples. By analyzing these figures, an optimal slope threshold value can be determined by searching a specific range of low \(cv\) value. There are three general patterns of \(cv\) observed within these figures. First, \(cv\) increases in proportional to the slope threshold value. For instance, the tree sample had low \(cv\) values with slope threshold ranging from 0° to 25° [for data strip 1, see Fig. 8(e)], 35° [for data strip 2, see Fig. 8(j)], and up to 40° [for data strip 3, see Fig. 8(o)]. Additional examples can be found in the building sample of data strips 1 to 3 [see Fig. 8(a), (f), and (k)], grass sample of data strip 2 [see Fig. 8(g)], and road sample of data strip 3 [see Fig. 8(m)].

The second observed pattern is similar to the first scenario; however, \(cv\) has a cutting increase at small value of slope threshold and then reaches to a steady stage. For instance, the road sample of data strip 2 and the soil sample of data strips 2 and 3 were in a low \(cv\) value, and it gradually increased to a maximum value at slope threshold between 15° and 20° until it reached to a steady stage afterward. The last observed pattern is a V-shaped form of \(cv\) across the slope threshold range. For instance, the \(cv\) of grass sample of data strips 1 and 3 [see Fig. 8(b) and (l)] dropped sharply at the beginning, reached to a minimum at the slope threshold between 30° and 35°, and climbed up after 40°. Fig. 8(c) and (d) also demonstrates such
large fluctuations in \( \text{vmr} \) also appeared in the grass sample (from \( \uparrow 8\% \) to \( \downarrow 75\% \)), road sample (from \( \uparrow 16\% \) to \( \downarrow 72\% \)), and soil sample (from \( \uparrow 149\% \) to \( \downarrow 78\% \)). It is worth noting that the elevation in the south of the study area is higher than the elevation in the north of the study area (see Fig. 5); the change in slope would cause an increase in incidence angle, resulting in the overcorrection effects, as reported in the previous literature works for satellite remote sensing images [41], [42]. In the tree sample, the \( \text{vmr} \) even spiked from \( \sim 10 \) (in the OI data) to \( \sim 80 \) in the three radiometrically corrected intensity data using incidence angle. Therefore, the ultimate solution of radiometric correction should combine the use of both laser incidence and scan angles.

As described in Section II, a new processing scheme was proposed by using the surface slope as a threshold value in selecting using either the incidence angle (when slope \( \leq 40^\circ \)) or the scan angle (when slope \( > 40^\circ \)) in the radiometric correction process. The \( \text{vmr} \) obtained from the results derived by this approach showed a consistent reduction and reached the smallest values, as shown in Table IV. Except the tree sample, all the \( \text{vmr} \) recorded for the corrected intensity of building, grass, road, and soil samples were mostly with values less than 1. The percentage change in \( \text{vmr} \) was reduced by 70\%–82\%. Although the \( \text{vmr} \) of tree sample data points were recorded with 2.076–2.525 in the radiometrically corrected intensity data using incidence and scan angles, the \( \text{vmr} \) values are relatively small when compared with their corresponding values 7.781–9.737 in the OI data. With respect to the significant reduction in \( \text{vmr} \), the proposed approach can counteract the overcorrection effects that occur when the surface slope exceeds 40° and thus produce improvement of land-cover homogeneity for all the land-cover types after radiometric correction.

2) Coefficient of Joint Variation (\( \text{cjv} \)): Fig. 9 shows the \( \text{cjv} \) derived for all combinations of pairwise land-cover features for the three LiDAR data strips. At a first glance of the figures, it is easy to notice that the \( \text{cjv} \) of the radiometrically corrected intensity data using incidence angle always produced higher \( \text{cjv} \) values than that of the OI data and the radiometrically corrected intensity data using incidence and scan angles. Such phenomenon can be ascribed by the overcorrection effects leading to an increase in \( \text{cv} \), which may intensify the between-class confusion, as reflected in the high \( \text{cjv} \) values. However, the radiometrically corrected intensity data using incidence angle still improved the \( \text{cjv} \) compared with those derived from the OI data in certain circumstances. Examples can be found in the \( \text{cjv} \) of building and road derived from data strip 3 [see Fig. 9(b)], the \( \text{cjv} \) of grass and soil derived from data strip 1 [see Fig. 9(f)], and the \( \text{cjv} \) of soil and tree derived from data strip 2 [see Fig. 9(j)]. The \( \text{cjv} \) was reduced by 26\%–75\% within these radiometrically corrected intensity data using incidence angle while comparing with those derived from the OI data. Such results also echo our previous findings in [31], where the use of laser incidence angle can help improve the variance-to-mean ratio of land-cover features. Among most of the figures, it is easy to recognize that the proposed approach, i.e., radiometrically corrected intensity data using incidence and scan angles, generated lower \( \text{cjv} \) values than that of the OI data, with an exception of two combinations in Fig. 9(b) and (d),

### Table IV

**Coefficient of Variation of Five Land Cover Features Generated From the Original and Radiometrically Corrected Intensity Data for the Three LiDAR Data Strips**

<table>
<thead>
<tr>
<th></th>
<th>Building</th>
<th>Grass</th>
<th>Road</th>
<th>Soil</th>
<th>Tree</th>
</tr>
</thead>
<tbody>
<tr>
<td>OI</td>
<td>1.192</td>
<td>3.415</td>
<td>3.992</td>
<td>3.566</td>
<td>9.377</td>
</tr>
<tr>
<td>RCI_IA</td>
<td>2.965</td>
<td>3.691</td>
<td>4.634</td>
<td>0.799</td>
<td>82.624</td>
</tr>
<tr>
<td>RCI_IA_SA</td>
<td>0.807</td>
<td>0.626</td>
<td>1.029</td>
<td>0.783</td>
<td>2.525</td>
</tr>
<tr>
<td></td>
<td>(74%)</td>
<td>(74%)</td>
<td>(74%)</td>
<td>(74%)</td>
<td>(74%)</td>
</tr>
<tr>
<td></td>
<td>(74%)</td>
<td>(74%)</td>
<td>(74%)</td>
<td>(74%)</td>
<td>(74%)</td>
</tr>
<tr>
<td>OI</td>
<td>0.962</td>
<td>1.659</td>
<td>1.538</td>
<td>3.925</td>
<td>8.990</td>
</tr>
<tr>
<td>RCI_IA</td>
<td>0.265</td>
<td>0.677</td>
<td>0.429</td>
<td>9.756</td>
<td>84.957</td>
</tr>
<tr>
<td>RCI_IA_SA</td>
<td>0.265</td>
<td>0.371</td>
<td>0.429</td>
<td>0.963</td>
<td>2.437</td>
</tr>
<tr>
<td></td>
<td>(72%)</td>
<td>(72%)</td>
<td>(72%)</td>
<td>(72%)</td>
<td>(72%)</td>
</tr>
<tr>
<td></td>
<td>(72%)</td>
<td>(72%)</td>
<td>(72%)</td>
<td>(72%)</td>
<td>(72%)</td>
</tr>
<tr>
<td>OI</td>
<td>2.144</td>
<td>8.120</td>
<td>2.395</td>
<td>2.815</td>
<td>7.781</td>
</tr>
<tr>
<td>RCI_IA</td>
<td>2.904</td>
<td>2.026</td>
<td>1.893</td>
<td>6.482</td>
<td>74.344</td>
</tr>
<tr>
<td>RCI_IA_SA</td>
<td>0.654</td>
<td>1.999</td>
<td>0.688</td>
<td>0.719</td>
<td>2.076</td>
</tr>
<tr>
<td></td>
<td>(70%)</td>
<td>(75%)</td>
<td>(72%)</td>
<td>(75%)</td>
<td>(75%)</td>
</tr>
<tr>
<td></td>
<td>(70%)</td>
<td>(75%)</td>
<td>(72%)</td>
<td>(75%)</td>
<td>(75%)</td>
</tr>
</tbody>
</table>

1. Original intensity data
2. Radiometrically corrected intensity data using incidence angle
3. Radiometrically corrected intensity data using incidence and scan angles

For a tradeoff among all experiments, a slope threshold ranging from 30° to 40° seems to be a good choice for ensuring low \( cv \) leading to an improvement in land-cover homogeneity. The choice of such threshold value can be also interpreted by the nature of inclined surfaces in the study area, which were found to be mostly less than this value. In addition, the laboratory measurements reported in [27] found that the decrease in brightness with the incidence angle is negligible up to 30°–40° for a set of gravel samples with high reflectance. As a result, the slope threshold value with 40° was adopted for radiometric correction using the proposed approach.

### B. Assessment of the Land-Cover Homogeneity

1) Variance-to-Mean Ratio (\( \text{vmr} \)): Table IV shows \( \text{vmr} \) for the five land-cover features (i.e., building, grass, road, soil, and tree) on the OI, radiometrically corrected intensity data using incidence angle, and the proposed approach using slope threshold for selecting either incidence angle or scan angle for radiometric correction. In the radiometrically corrected intensity data using incidence angle, more than half of the results were recorded with an increase in \( \text{vmr} \) in data strips 1 to 3. In the building sample, the \( \text{vmr} \) was raised by 36\%–149\% in data strips 1 and 3, whereas contradictorily, a decrease in \( \text{vmr} \) (72\%) was recorded in the building sample obtained from data strip 2. This can be explained by considering that the sample data points acquired in data strips 1 and 3 are mostly located on the inclined rooftops of small houses, whereas those of data strip 2 are mainly located on the rooftops of BCIT campus buildings. Such pattern for road and soil samples of data strip 1; however, \( cv \) did not vary after reaching the minima.
Fig. 9. Coefficient of joint variation $c_{jv}$ of land-cover features generated from the OI data, radiometrically corrected intensity data using scan angle (RCI_IA), and radiometrically corrected intensity data using incidence and scan angles (RCI_IA_SA). (a) Building versus grass. (b) Building versus road. (c) Building versus soil. (d) Building versus tree. (e) Grass versus road. (f) Grass versus soil. (g) Grass versus tree. (h) Road versus soil. (i) Road versus tree. (j) Soil versus tree.

Fig. 10. Subarea 1: (a) Aerial photo, (b) DSM, (c) OI data, and (d) radiometrically corrected and normalized intensity data of an area located in the overlapping region of LiDAR data strips 1 and 2.

C. LiDAR Intensity Images Before and After Radiometric Correction and Normalization

Radiometric normalization was applied to the radiometrically corrected intensity data for all the three LiDAR data strips. Four subareas in the two overlapping regions were selected for further examination. Fig. 10(a)–(d) shows the aerial photo, digital surface model (DSM), OI data, and radiometrically corrected and normalized intensity data, respectively, of an area including different land-cover types in the overlapping region of LiDAR data strips 1 and 2. Systematic stripping noises were found on the building rooftop in the OI image [see Fig. 10(c)], where such noises were reduced in the radiometrically corrected and normalized intensity image [see Fig. 10(d)].

Fig. 11(a)–(d) shows the aerial photo, DSM, OI image, and radiometrically corrected and normalized intensity image, respectively, in another area located in another overlapping region of LiDAR data strips 1 and 2. In this example, the line-stripping noises were manifest on the grass cover and the building rooftop. The radiometrically corrected and normalized intensity...
image seemed to have slight reduction on the variability of intensity. One should note that the building rooftops in these two examples are inclined; radiometric correction demonstrated a positive influence in reducing the intensity difference between the inclined surfaces. Fig. 12 demonstrates an example of inclined building rooftops in the same study area as Fig. 11. As shown in the aerial photo, a building roof with opposite inclined orientation is located in the north of the flight line. In the OI data, the intensity value on surface B (close to the aircraft) was much higher than that on surface A, where the variance-to-mean ratio of the entire rooftop was 1.308. Difference of intensity was observed in surfaces A and B on the building rooftop. After applying radiometric correction, the \( \text{vmr} \) reduced 27% and reached to 0.948. Therefore, the results showed that radiometric correction improves the normalized intensity results.

In the example of Figs. 13 and 14, the study areas are located in the overlapping region of LiDAR data strips 2 and 3; the stripping noise problem was extreme. Unlike the previous examples, the noises appeared on the ground, grass cover, and building rooftops in a vertical direction [see Figs. 13(c) and 14(c)]. After radiometric normalization, such discrepancy was significantly reduced in the radiometrically corrected and normalized intensity image [see Figs. 13(d) and 14(d)], leaving a few low-level noises in the result images.
VI. LAND-COVER CLASSIFICATION

Several land-cover classification scenarios were carried out using the airborne LiDAR intensity data in order to evaluate the effects of the proposed radiometric correction and normalization on the classification accuracy. The reason for testing different classification scenarios is to avoid any biases of the experimental testing toward a specific problem domain. Due to the significant noise reduction in subareas 3 and 4, we compared the classification results between the OI image and the radiometrically corrected and normalized intensity image to justify the impact of the radiometric correction and normalization. Similar to our previous experiment in [16], the design of land-cover classes followed the standardized national United States Geological Survey (USGS) Land-Cover Classification Scheme (LCCS) for remote sensing image classification at different scales and resolutions [48]. The first scenario classified the study area into two land-cover classes: 1) “Built-up Land” (class 1 in level I of the USGS LCCS); and 2) “Natural Land”. The second scenario subdivided the “Natural Land” into “High Rangeland” (or “Tree”) and “Low Rangeland” and remained the “Built-up Land” as unchanged. The third scenario is composed of four classes: 1) “Built-up Land”; 2) “Tree”; 3) “Grass Land”; and 4) “Barren Land” (class 7 in level I of the USGS LCCS), which is described as an area of thin soil, sand, or rocks where less than one-third of the area has vegetation or other cover. The last scenario included five classes after subdividing the “Built-up Land” class into “Building” and “Road” (level II of the USGS LCCS). The five land-cover classes were as follows: 1) “Road”; 2) “Building”; 3) “Tree”; 4) “Barren Land” (or “Soil”); and 5) “Grass Land”. Training sites were selected for each of the land-cover classes with reference to the high-resolution aerial photo. The maximum-likelihood classifier was applied to the LiDAR data sets for each classification scenario. Finally, 510 random sample checkpoints were generated as a reference data for accuracy assessment in each of the subareas with the overall accuracy and the Kappa statistics (KS).

A. Subarea 3

In subarea 3, all four land-cover classification scenarios were applied to the two LiDAR data sets (i.e., OI image and radiometrically corrected and normalized intensity image). Table V summarizes the overall accuracy and KS in this case study. In the two-class scenario, the overall accuracy produced by the OI image was close to 83%, where the radiometrically corrected and normalized intensity image produced a slightly improved accuracy value of 86.5%. Such increase was mainly due to the improvement of distinguishing “Built-up Land” after radiometric correction and normalization, where the radiometrically corrected and normalized intensity image recorded a KS value of 0.758 when compared with 0.594 recorded in the result of the OI image. In the three-class scenario, the overall accuracy of the OI image was 71.8%; an 8% accuracy improvement was achieved in the classification result of radiometrically corrected and normalized intensity image. The KS value of the radiometrically corrected and normalized intensity image was found increased by ~0.1 in both “High Rangeland” and “Low Rangeland”. In the four- and five-class scenarios, the overall accuracy dropped to 63.5% and 45.1%, respectively. Nevertheless, an increase in the overall accuracy was found in the result of radiometrically corrected and normalized intensity image classification. Respective increase of 13.4% and 16.5% was observed in the four- and five-class scenarios, respectively. The KS value of “Tree”, “Grass”, “Soil”, and “Built-up Land” were found with an increase from 0.08 to 0.22 in the result of radiometrically corrected and normalized intensity image classification, when compared to the corresponding KS values generated from the OI image.

As shown in Fig. 13(c), the OI image in subarea 3 had the most serious line-stripping noise. Owing to the successful removal of noise in the intensity data, accuracy improvement was found to be significant after radiometric correction and normalization. The result in the first three classification scenarios achieved more than 75% overall accuracy. The over 60% classification accuracy generated using the radiometrically corrected and normalized intensity image in the five-class scenario was even much higher than all the results of radiometrically corrected intensity image classification in our previous experiment [16] (maximum overall accuracy was 45% only). One should note that the more line-stripping noise on the land-cover features in the OI image, the more improvement of KS in the radiometrically corrected and normalized intensity image classification can be achieved. For instance, the KS of “Tree” in the five-class scenario was 0.077 using the OI image, and it increased to 0.298 using the radiometrically corrected and normalized intensity image. The KS of class “Grass” in the same scenario was 0.174 in the OI image classification result.

<table>
<thead>
<tr>
<th>TABLE V</th>
<th>OVERALL ACCURACY AND KS OF LAND-COVER CLASSIFICATION IN SUBAREA 3</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Original Intensity</td>
</tr>
<tr>
<td>2-Class Scenario</td>
<td></td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td>82.7%</td>
</tr>
<tr>
<td>KS of Natural Land</td>
<td>0.699</td>
</tr>
<tr>
<td>KS of Built-up Land</td>
<td>0.594</td>
</tr>
<tr>
<td>3-Class Scenario</td>
<td></td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td>71.8%</td>
</tr>
<tr>
<td>KS of Built-up Land</td>
<td>0.612</td>
</tr>
<tr>
<td>KS of High Rangeland</td>
<td>0.064</td>
</tr>
<tr>
<td>KS of Low Rangeland</td>
<td>0.743</td>
</tr>
<tr>
<td>4-Class Scenario</td>
<td></td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td>63.5%</td>
</tr>
<tr>
<td>KS of Built-up Land</td>
<td>0.612</td>
</tr>
<tr>
<td>KS of Grass</td>
<td>0.174</td>
</tr>
<tr>
<td>KS of Soil</td>
<td>0.618</td>
</tr>
<tr>
<td>KS of Tree</td>
<td>0.077</td>
</tr>
<tr>
<td>5-Class Scenario</td>
<td></td>
</tr>
<tr>
<td>Overall Accuracy</td>
<td>45.1%</td>
</tr>
<tr>
<td>KS of Building</td>
<td>0.347</td>
</tr>
<tr>
<td>KS of Grass</td>
<td>0.174</td>
</tr>
<tr>
<td>KS of Soil</td>
<td>0.618</td>
</tr>
<tr>
<td>KS of Tree</td>
<td>0.077</td>
</tr>
<tr>
<td>KS of Road</td>
<td>0.206</td>
</tr>
</tbody>
</table>
and 0.384 in the radiometrically corrected and normalized intensity image classification result. Up to this point, subarea 3 successfully proves the effectiveness of the proposed radiometric correction and normalization on land-cover classification.

B. Subarea 4

As shown in Table VI, the classification result of the OI image achieved 87.5% in the two-class scenario. The radiometrically corrected and normalized intensity image did not demonstrate significant impact on the classification result, where less than 1% of increase in overall accuracy was found. The KS values showed an opposite trend in the intensity data classification results. The “Natural Landscape” demonstrated a drop of KS 0.01 in the result of radiometrically corrected and normalized intensity image classification. Contradictorily, the KS value of the “Built-up Land” class increased from 0.697 in the result of OI image classification to 0.760 in the radiometrically corrected and normalized intensity image classification result. In the three-class scenario, the radiometrically corrected and normalized intensity image produced a 6.4% accuracy improvement, when compared to the classification result of the OI image, which was only 66.3%. “High Rangeland” and “Low Rangeland” were observed with an increase in KS with more than 0.1 in the radiometrically corrected and normalized intensity image classification result when compared with that of the OI image. Following the trend in the three-class scenario, a 9.2% overall accuracy improvement was found in the radiometrically corrected and normalized intensity image, where the accuracy of classification using OI image was 59% only. Consistently, an increase in KS values ranging from 0.04 to 0.21 was found in all the land-cover features classified using the radiometrically corrected and normalized intensity image. Since subarea 4 is located in the same overlapping region of subarea 3, the line-stripping noisy effect was manifest. After radiometric correction and normalization, the discrepancy of intensity within the same land-cover feature was significantly reduced, resulting in a large accuracy improvement in the three- and four-class scenarios.

VII. DISCUSSION AND LIMITATIONS

In our previous experiments [16], [31], we adopted the use of laser incidence angle to perform radiometric correction of airborne LiDAR intensity data. Although an improvement of classification accuracy is found using the radiometrically corrected intensity data, a large standard deviation of intensity values appeared on the land features with steep slope (i.e., building boundaries and tree canopies). In this context, the corrected intensity does not truly reflect and represent the relative spectral reflectance of land features in the operating wavelength of airborne LiDAR sensor. To relieve the overcorrection effect, a correction mechanism was proposed by using the slope as a threshold to select using either scan angle or incidence angle in the radar (range) equation. Significant improvement was found in the land-cover homogeneity by assessing the coefficient of variation, coefficient of joint variation, and variance-to-mean ratio in the radiometrically corrected intensity data. The intention of the work aims to improve the intensity homogeneity of land-cover features in order to facilitate classification and mapping applications. However, the corrected intensity data still require absolute calibration for deriving the physical spectral reflectance values, if necessary.

The main objective of this paper aims to maximize the benefits of using airborne LiDAR intensity data for land-cover classification. The use of the correction mechanism is restricted to the LiDAR data acquired for urban environment, which includes both penetrable and impenetrable objects with different land-cover types. Due to the complications in modeling the leaf orientation, the accuracy of the incidence angle derived on the tree canopies from the TIN model is not guaranteed. As such, the proposed approach may not be applicable for forestry applications, such as biomass estimation and tree species classification. Nevertheless, the use of the radiometrically calibrated full-waveform LiDAR data has been proven to be viable for these applications [49], [50]. As another limitation, implementation of radiometric normalization basically normalizes the histogram acquired in the overlapping region, which should cover a variety of land-cover features so that the histograms generated from the LiDAR data strips are representable. In this regard, a detailed flight mission should be carefully planned in order to assure this setting.

For the way forward, future research should investigate and develop a universal laser reflectance model for modeling both the peculiar and usual characteristics of laser intensity at different backscattering geometries and wavelengths. There are a number of experimental LiDAR sensors operating at different wavelengths where these sensors are currently being developed in the laboratory for retrieving the backscattered reflectance of small plant and tree canopy [33], [51]–[53]. Preliminary findings have already proven the usefulness of multil wavelength LiDAR data to detect small changes in leaf reflectance with respect to the biochemical concentration at leaf level [54]. The backscattered intensity data at different wavelengths were used to derive the normalized difference vegetation index and photochemical reflectance index for measuring plant physiology.
Improved efficiency in classification and interpretation compared with the traditional monochromatic LiDAR data was demonstrated [56], [57]. Nevertheless, most of these works focused on the short-range LiDAR sensor for terrestrial mapping without performing radiometric correction. Therefore, it is worth investigating the radiometric correction of backscattered intensity signals recorded by multispectral airborne LiDAR sensors to serve large-scale surface mapping and object recognition applications.

VIII. CONCLUSION

We have presented an improved radiometric correction model to relieve the effects of overcorrection and a radiometric normalization model to remove the line-stripping problem in the overlapping region of airborne LiDAR intensity data. The correction model utilizes the slope as a threshold to control either using the scan angle or the incidence angle in the radar (range) equation for airborne LiDAR intensity data, and the normalization model relies on the subhistogram matching technique to normalize the intensity data in the overlapping region. It was found that the coefficient of variation of most of the land-cover samples reached to the lowest value with a slope threshold between 30° and 40°. The variance-to-mean ratio was assessed in five different land-cover samples (i.e., building, grass, road, soil, and tree) generated from the original and corrected intensity data. Regardless of the land-cover features, an increase in the variance-to-mean ratio was found in some of the land-cover samples ranging from 8% to 855% in the radiometrically corrected intensity data using incidence angle, which indicates an overcorrection effect. In our proposed approach, which incorporated both laser incidence and scan angles in radiometric correction, experimental results demonstrated that the variance-to-mean ratio within the same land-cover feature was reduced by 70% to 82%. To assess the between-class variation, the coefficient of joint variation was applied to pairwise combination of land-cover features before and after correction. It was found that the coefficient of joint variation of radiometrically corrected intensity data using incidence angle had higher values than that of the OI data. However, most of the results derived from the radiometrically corrected intensity data using incidence and scan angles produced a reduction of coefficient of joint variation by 1%–35%, and the maximum reduction achieved up to 74%. We also implemented different classification scenarios on the overlapping region of OI data with serious line-stripping noise. An increase in the overall accuracy in between 5.7% and 16.5% (excluding the two-class scenario) was found in the classification results of radiometrically corrected and normalized intensity image. With the justification of different statistical assessments in this paper, the proposed radiometric correction and normalization can significantly reduce the within-class variation and between-class confusion of the land-cover features, as well as reduce the line-stripping noise in the overlapping region. Although the proposed approaches demonstrate significant improvement in land-cover homogeneity and land-cover classification, a universal laser reflection model is still desired for the upcoming multispectral airborne LiDAR data.

ACKNOWLEDGMENT

The authors would like to thank Prof. A. Habib and his research team from the University of Calgary, Calgary, AB, Canada, and McElhanney Consulting Services, Ltd., Calgary, BC, Canada, for providing the LiDAR and image data sets. The authors appreciate the comments of the three anonymous reviewers, which helped the final manuscript to be in a much better shape.

REFERENCES

[18] B. Höfler and N. Pfeifer, “Correction of laser scanning intensity data: A joint variation of radiometrically corrected intensity data using incidence and scan angles in radiometric correction, experimental results demonstrated that the variance-to-mean ratio within the same land-cover feature was reduced by 70% to 82%. To assess the between-class variation, the coefficient of joint variation was applied to pairwise combination of land-cover features before and after correction. It was found that the coefficient of joint variation of radiometrically corrected intensity data using incidence angle had higher values than that of the OI data. However, most of the results derived from the radiometrically corrected intensity data using incidence and scan angles produced a reduction of coefficient of joint variation by 1%–35%, and the maximum reduction achieved up to 74%. We also implemented different classification scenarios on the overlapping region of OI data with serious line-stripping noise. An increase in the overall accuracy in between 5.7% and 16.5% (excluding the two-class scenario) was found in the classification results of radiometrically corrected and normalized intensity image. With the justification of different statistical assessments in this paper, the proposed radiometric correction and normalization can significantly reduce the within-class variation and between-class confusion of the land-cover features, as well as reduce the line-stripping noise in the overlapping region. Although the proposed approaches demonstrate significant improvement in land-cover homogeneity and land-cover classification, a universal laser reflection model is still desired for the upcoming multispectral airborne LiDAR data.


