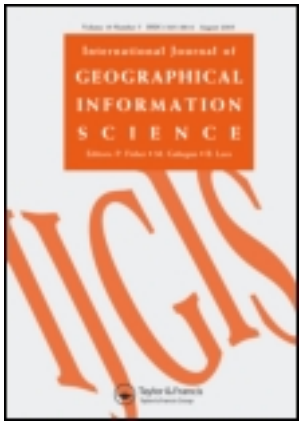

International Journal of Geographical Information Science

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/tgis20

A multiscale geographic object-based image analysis to estimate lidar-measured forest canopy height using Quickbird imagery

Gang Chen [a], Geoffrey J. Hay [a], Guillermo Castilla [a], Benoît St-Onge [b] & Ryan Powers [a]

[a] Foothills Facility for Remote Sensing and GIScience, Department of Geography, University of Calgary, Calgary, Canada

[b] Department of Geography, Université du Québec à Montréal, Montreal, Canada

Available online: 07 Jul 2011

To cite this article: Gang Chen, Geoffrey J. Hay, Guillermo Castilla, Benoît St-Onge & Ryan Powers (2011): A multiscale geographic object-based image analysis to estimate lidar-measured forest canopy height using Quickbird imagery, International Journal of Geographical Information Science, 25:6, 877-893

To link to this article: http://dx.doi.org/10.1080/13658816.2010.496729

International Journal of Geographical Information Science
Vol. 25, No. 6, June 2011, 877–893

Taylor & Francis
Taylor & Francis Group

A multiscale geographic object-based image analysis to estimate lidar-measured forest canopy height using Quickbird imagery

Gang Chen[a]*, Geoffrey J. Hay[a], Guillermo Castilla[a], Benoît St-Onge[b] and Ryan Powers[a]

[a]*Foothills Facility for Remote Sensing and GIScience, Department of Geography, University of Calgary, Calgary, Canada;* [b]*Department of Geography, Université du Québec à Montréal, Montreal, Canada*

(*Received 27 July 2009; final version received 18 May 2010*)

Lidar (light detection and ranging) has demonstrated the ability to provide highly accurate information on forest vertical structure; however, lidar data collection and processing are still expensive. Very high spatial resolution optical remotely sensed data have also shown promising results to delineate various forest biophysical properties. In this study, our main objective is to examine the potential of Quickbird (QB) imagery to accurately estimate forest canopy heights measured from small-footprint lidar data. To achieve this, we have developed multiscale geographic object-based image analysis (GEOBIA) models from QB data for both deciduous and conifer stands. In addition to the spectral information, these models also included (1) image-texture [i.e., an internal-object variability measure and a new dynamic geographic object-based texture (GEOTEX) measure that quantifies forest variability within neighboring objects] and (2) a canopy shadow fraction measure that acts as a proxy of vertical forest structure. A novel object area-weighted error calculation approach was used to evaluate model performance by considering the importance of object size. To determine the best object scale [i.e., mean object size (MOS)] for defining the most accurate canopy height estimates, we introduce a new perspective, which considers height variability both between- and within-objects at all scales. To better evaluate the improvements resulting from our GEOBIA models, we compared their performance with a traditional pixel-based approach. Our results show that (1) the addition of image-texture and shadow fraction variables increases the model performance versus using spectral information only, especially for deciduous trees, where the average increase of R^2 is approximately 23% with a further 1.47 m decrease of Root Mean Squared Error (RMSE) at all scales using the GEOBIA approach; (2) the best object scale for our study site corresponds to an MOS of 4.00 ha; (3) at most scales, GEOBIA models achieve more accurate results than pixel-based models; however, we note that inappropriately selected object scales may result in poorer height accuracies than those derived from the applied pixel-based approach.

Keywords: geographic object-based image analysis; multiscale; geographic object-based texture; shadow fraction; Quickbird; lidar; forests

1. Introduction

Lidar (light detection and ranging), a relatively recent remote sensing technique, has demonstrated the ability to provide highly accurate information on forest vertical structure (Means *et al.* 1999, Lim *et al.* 2003); however, compared with optical satellite data of similar spatial resolution and extent, the costs associated with lidar forest data acquisition are much higher.

*Corresponding author. Email: gangchen@ucalgary.ca

ISSN 1365-8816 print/ISSN 1362-3087 online

DOI: 10.1080/13658816.2010.496729
http://www.informaworld.com

The availability of very high resolution (VHR) optical imagery (<5.0 m) creates great opportunities for forest inventory. These VHR data have shown promising results to estimate various forest biophysical properties, such as individual tree species (Wulder *et al.* 2004a), stem density (Wang *et al.* 2004), and tree crown size and position (Culvenor 2003). Although VHR optical remote sensing data are typically used to acquire horizontal forest canopy structure, they can also be used to describe forest vertical structure (e.g., canopy height) when additional image-derived variables such as canopy texture and shadow information are utilized. To date, only a few studies have evaluated the accuracy of estimating forest height information exclusively from VHR optical imagery. For example, Franklin and McDermid (1993) found a significant correlation between VHR (pixel size of 1.0 by 1.3 m) CASI (Compact Airborne Spectrographic Imager) red band data and tree-top mean height at the stand level ($R = 0.75$). Hyde *et al.* (2006) compared different types of remotely sensed data to map forest structure for wildlife habitat analysis and found an R^2 of 0.566 using Quickbird (QB) imagery to estimate mean canopy height. Similarly, Donoghue and Watt (2006) used IKONOS data to estimate lidar-measured forest height in spruce plantation forests in northern England. They reported that lidar and IKONOS data appear to show a good agreement in densely stocked plantation areas with tree height less than 10 m. A small number of studies have also investigated the potential of using medium resolution optical data (e.g., 30 m resolution Landsat imagery) to update forest stand height information (Hudak *et al.* 2002, Wulder and Seemann 2003).

Traditional pixel-based approaches for extracting geographic information have been applied to remote sensing for over three decades. However, in recent years, remote sensing has undergone an evolution in data acquisition technologies (e.g., higher resolution sensors and more affordable data) and more sophisticated user requirements. At high spatial resolution, individual pixels are typically smaller than the geographic objects of interest. Although this provides more details for visual interpretation, it also creates higher spectral variance within the classes of interest that can decrease classification accuracy when using pixel-based approaches (Hay *et al.* 1996). To minimize the effect of high spectral variance, new (nonpixel-based) approaches are needed. In this respect, object-based image analysis (OBIA), which combines spatial and spectral information within image analysis using image objects (i.e., groups of connected pixels that are relatively homogeneous and different from their surroundings) instead of individual pixels as basic study units to capture geographic objects, provides a viable alternative to the traditional pixel-based paradigm. To differentiate OBIA in the geographic domain (vs. computer vision and biomedical imaging), the name geographic object-based image analysis (GEOBIA) was recently proposed by Hay and Castilla (2008). Advantages of GEOBIA over pixel-based approaches have been illustrated in several forest studies (Wulder *et al.* 2004b, Yu *et al.* 2006, Addink *et al.* 2007). However, there remain numerous challenges to be addressed (Hay and Castilla 2008), including the investigation of the best object scales [(i.e., mean object sizes (MOS)] to achieve the desired model accuracy.

In this article, we explore the potential of GEOBIA and pixel-based models applied to QB imagery to best estimate forest canopy height over a range of object scales. The results of these canopy height estimates are then compared with a lidar-measured canopy height model (CHM), and errors are defined. To achieve this, the objectives of this research are (1) to examine the potential using QB imagery to estimate mean canopy height measured from small-footprint lidar data; (2) to evaluate the contribution of spectral, image-texture, and shadow information derived from QB imagery to the accuracy of canopy height estimation for GEOBIA and pixel-based approaches; (3) to explore a range of object scales (i.e., MOS) to find the most accurate canopy height estimates; and (4) to compare the performance (defined by canopy height error) of GEOBIA and pixel-based approaches in canopy height estimation.

2. Data collection

2.1. Study area

The study site (49°52′N, 125°20′W) is located approximately 10 km southwest of Campbell River on the east coast of Vancouver Island, British Columbia, Canada (Figure 1). This study was performed over a 5.1 × 5.1 km (2601 ha) area, characterized by regenerating conifer and deciduous forests and clearcuts. Conifer forest types compose 65% of the study area, with a mean canopy height of 19.7 m, and are dominated by approximately 80% Douglas-fir [*Pseudotsuga menziesii* (Mirb.) Franco], with small proportions of Western Red Cedar

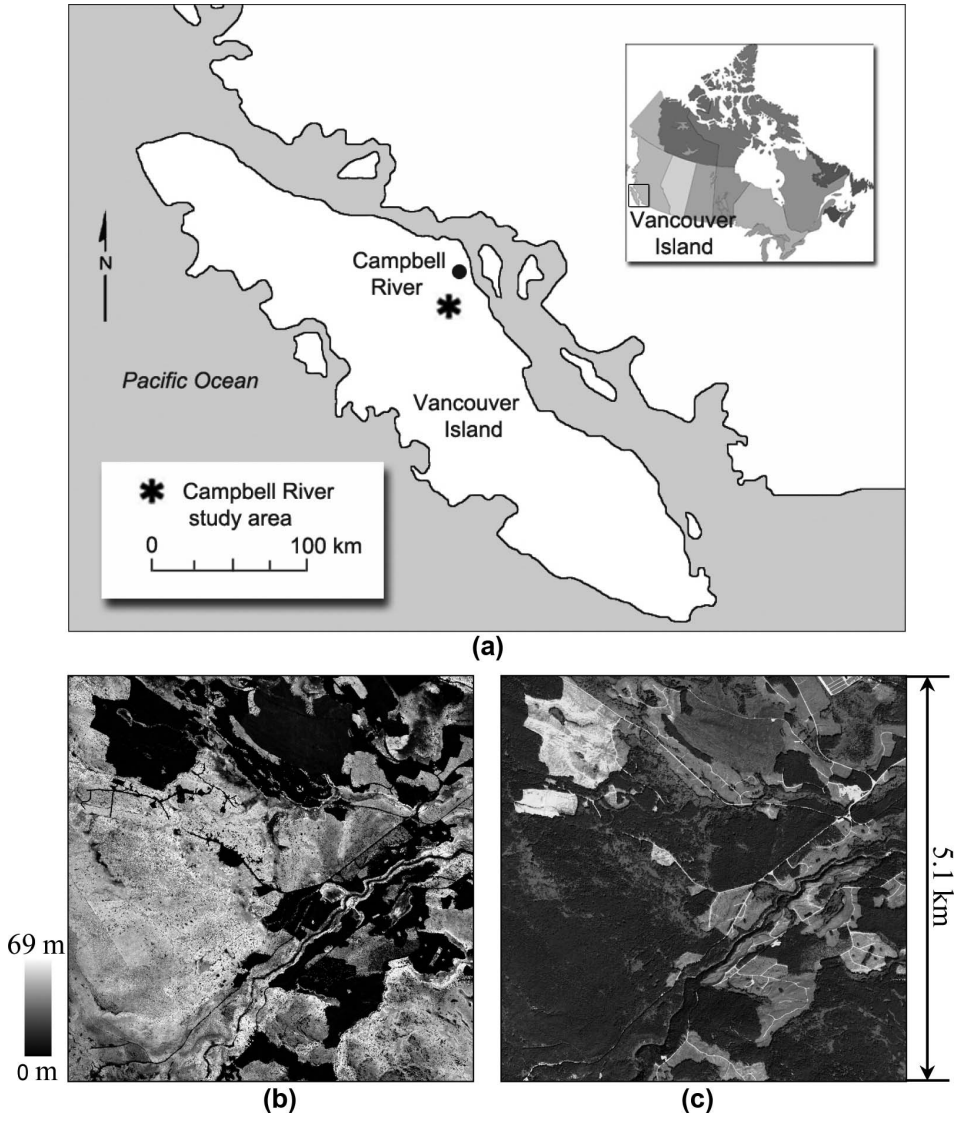

Figure 1. (a) Study area located southwest of Campbell River, Vancouver Island, Canada. (b) Lidar canopy height model (CHM). (c) Quickbird grayscale image converted from a false color composite using near-infrared (NIR), red, and green bands.

[*Thuja plicata* (Donn.)], and Western Hemlock [*Tsuga heterophylla* (Raf.) Sarg.]. Deciduous forest of Red Alder (*Alnus rubra* Bong.) comprises 16% of the study area, with a mean canopy height of 17.4 m. Most forest stands in this study area are composed of regenerated forest from harvest and are between 20 and 60 years of age. Topographically, the average elevation is approximately 300 m above sea level, ranging from 180 (southwest) to 440 m (northeast), with a gentle slope of 5°–10°.

2.2. *Lidar data*

Twelve overlapping flight lines of lidar data were acquired on 8 June 2004, by a Terrain Scanning Lidar system (Terra Remote Sensing Inc., Sidney, Canada) on a Bell 206 Jet Ranger helicopter. The positioning systems, a Litton LTN-92 inertial navigation system (INS) and an Ashtech Z-surveyor Dual Frequency P-code differential global positioning system (DGPS), were installed to record the aircraft's altitude and position within 5–10 cm. Terrain scanning lidar is a discrete return lidar system (Lightwave Model 110) with a pulse repetition frequency of 10 kHz, a wavelength of 1047 nm, a swath width of 56°, and a beam divergence of 3.5 mrad. This mission used a continuous scanning mode in the typical zigzag pattern, yielding point densities of 0.7/m^2 and a footprint size of 0.19 m.

The raw lidar point cloud data were collected containing both ground and nonground returns. Nonground points were assumed equivalent to the returns from the vegetation/tree canopy, as no artificial objects existed in the study site. Classifying point cloud data into ground and tree canopy returns was implemented with Terrascan software (v4.006 – Terrasolid, Helsinki, Finland), which was developed using iterative algorithms that combine filtering and thresholding methods (Kraus and Pfeifer 1998, Axelsson 1999). Ground and tree canopy returns were then separately interpolated (Hutchinson 1989) to form a digital elevation model (DEM) and a digital surface model (DSM) with 1-m grid cell size. The final step was to obtain the forest CHM by subtracting the DEM from the DSM. The lidar CHM represents a canopy height range from 0.0 to 72.9 m, with an average height of 19.3 m and a standard deviation of 8.0 m over the entire forest canopy. Over 50% of the forest area is covered by canopies with heights ranging from 20 to 30 m. Young canopies (lower than 10 m) account for approximately 8%, with only 1% very tall canopies (larger than 35 m).

2.3. *Quickbird data*

A cloud-free QB image was acquired on 11 August 2004 over the same study area. The image used in this study consists of four multispectral bands [i.e., blue, green, red, and near-infrared (NIR)] and one panchromatic band, with an off-nadir view angle of 11.1°.

The different spatial resolutions between the QB image (2.4 m multispectral bands and 0.6 m panchromatic band) and lidar data (1.0 m) make comparison difficult. Therefore, a principal components spectral sharpening technique (Welch and Ahlers 1987) was used to rescale the QB image to a 1.0 m spatial resolution multispectral image by fusing the blue, green, red, and NIR channels with the panchromatic band. This method was selected as both spectral values and high spatial resolution information content were well retained. The QB image was then geometrically coregistered to the lidar data using 118 ground control points. A second-order polynomial warping method and the nearest neighbor resampling were selected for the coregistration, yielding a RMSE of 0.85 m. Because the study area was covered by relatively dense forests, the coregistration was performed using tree-tops only.

3. Data analysis

The flowchart shown in Figure 2 summarizes the steps completed in this study, whereas the following subsections provide greater detail and explanation.

3.1. GEOBIA approach

3.1.1. Multiscale image segmentation

The basic study units in the GEOBIA approach are segmentation-derived regions. These are typically referred to as *segments*, whose purpose is to spatially model geographically referenced image-objects or *geo-objects*. Geo-objects represent delineated areas, entities, or objects in the image that are meaningful to the image analyst – within a geographic context. In our case, well-generated geo-objects represent forest patches of varying size and shape that resulted from applying the *size-constrained region merging* (SCRM) algorithm to the pan-sharpened multispectral QB imagery. SCRM is a freeware image segmentation tool, containing six main processing steps in the algorithm, that is, image resampling, image smoothing, gradient magnitude image generation, watershed partition, region merging, and vectorization (Castilla *et al*. 2008). Key advantages of SCRM over currently existing segmentation algorithms are (1) the size of the objects of interest can easily and explicitly be controlled through input parameters, such as MOS, minimum object size, and maximum object size; and (2) the well-defined polygons contain smooth pixel boundaries, which are similar to human-made delineation.

In this study, 15 different object scales (i.e., MOS) were evaluated with the SCRM input parameters listed in Table 1. Figure 3 illustrates examples of the segmentation results derived at five scales for a small sample area. A MOS of 0.04 ha was chosen as the smallest scale. The main reason for this size was that lidar and optical remote sensing systems have different

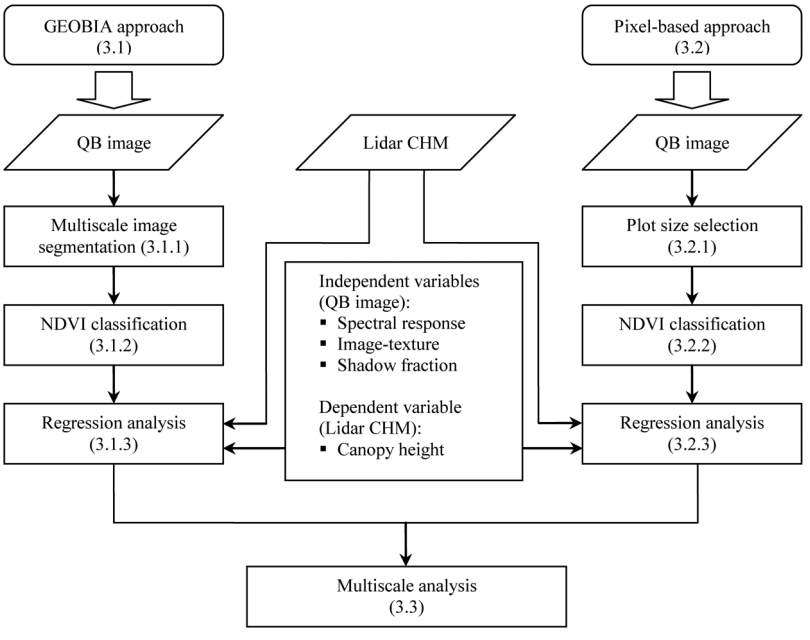

Figure 2. Flowchart of the research process with reference to Data Analysis sections.

Table 1. Segmentation parameters used in SCRM.

Scale	Minimum object size (ha)	Mean object size (ha)	Maximum object size (ha)
1	0.008	0.040	0.048
2	0.030	0.150	0.180
3	0.070	0.350	0.420
4	0.130	0.650	0.780
5	0.200	1.000	1.200
6	0.300	1.500	1.800
7	0.400	2.000	2.400
8	0.500	2.500	3.000
9	0.600	3.000	3.600
10	0.700	3.500	4.200
11	0.800	4.000	4.800
12	0.900	4.500	5.400
13	1.000	5.000	6.000
14	1.100	5.500	6.600
15	1.200	6.000	7.200

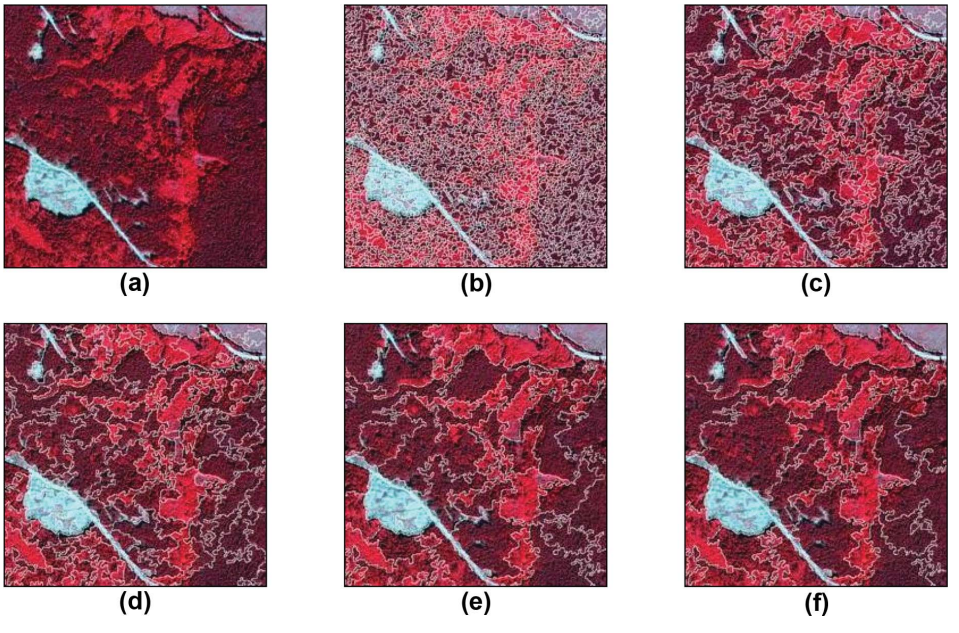

Figure 3. (a) A sample area in the study site, and the SCRM-derived segmentation boundaries overlaid on the corresponding area with the mean object size (MOS) of (b) 0.04 ha, (c) 0.36 ha, (d) 1.00 ha, (e) 4.00 ha, and (f) 6.00 ha.

data acquisition geometry, which introduce relatively large coregistration errors – especially in areas where large canopy height variation exists. Therefore, model performance using a MOS smaller than 0.04 ha was in some places highly affected by the coregistration error. The MOS of 6.00 ha was selected as the largest scale, because, in preliminary tests, the QB image tended to be undersegmented, meaning that the resulting segments contained various land cover types (rather than a single cover-type) when using a MOS larger than 6.00 ha.

3.1.2. NDVI classification

Forests with different tree species tend to possess very different spectral reflectance characteristics in optical remote sensing imagery. In our study area, deciduous canopies tend to show higher digital numbers (DNs) than conifers in the NIR band, whereas conifers contain more spectral variability than deciduous trees because of their steeper conical crowns, resulting in more visible forest gaps and shadows. Therefore, to improve canopy height estimation using QB imagery, the study area was classified into three classes: conifers, deciduous, and clearcuts (i.e., non- or sparse-forest areas) using an NDVI (normalized difference vegetation index) thresholding method. Based on visual inspection, deciduous canopies also consistently show higher NDVI values than conifers. Thus, two thresholds (0.55 and 0.70) were applied to classify the NDVI image. The first threshold was used to separate clearcuts from forests, whereas the latter was applied to differentiate conifers and deciduous vegetation. This schema was applied to the segmented images at all 15 scales.

3.1.3. Regression analysis – variables and models

Regression models were developed to estimate canopy height with the predicting variables derived from QB data, and the canopy height as dependent variable, derived from the lidar CHM. The independent (i.e., predicting) variables used in the analysis were (1) the *spectral response* – (i.e., DNs) of four optical bands; (2) the *image texture* derived from the standard deviation of the DNs *within* objects, as well as *between* neighboring objects for each of the four optical bands; and (3) the *shadow fraction* – a quotient of the shaded areas and the entire corresponding forest unit area – calculated from the NIR band (Table 2). More details of these variables are provided below.

3.1.3.1. Spectral response. There are four spectral response variables, one per each band of the QB image (i.e., NIR, red, green, and blue). The values of these variables in a given segment are the mean DNs of pixels within the segment.

3.1.3.2. Image-texture. Remote sensing image-texture refers to the local variation of brightness in an image, which is typically measured using the Grey-Level-Co-occurrence Matrix method. Essentially, this is based on the co-occurrence of DNs located within an arbitrarily sized square moving window, and evaluated at a choice of different directions and lags (Hay and Niemann 1994). In this study, image-texture is considered a surrogate spatial

Table 2. Independent variables used in the regression analysis.

Variables	GEOBIA approach	Pixel-based approach	Description
Spectral response	DN*i*_Obj	DN*i*_Pix	Average of DNs for the *i*th[a] band within study area units[b]
Image-texture	TXIt*i*_Obj	TXIt*i*_Pix	Internal standard deviation for the *i*th band within study area units
	GEOTEX*i*	TXNb*i*_Pix	Neighboring standard deviation for the *i*th band within neighboring study area units
Shadow fraction	SF_Obj	SF_Pix	A quotient of the size of shadow areas and the size of corresponding entire study area units

[a]*i* is the band number (i.e., 0 – blue band, 1 – green band, 2 – red band, and 3 – NIR band).
[b]For the GEOBIA approach, the study area units are objects; whereas for the pixel-based approach, the study area units are fixed-size windows.

information source that represents both vertical and horizontal forest structure. Instead of a traditional square window we used the geo-object boundary as an adaptive window. Thus, image-texture is calculated in two ways: (1) internal-object texture was calculated as the standard deviation of all DNs within the geo-object extent; and (2) *geographic object-based texture* (GEOTEX) was calculated as the standard deviation of the averaged DNs of the center object and its neighboring objects.

3.1.3.3. Shadow fraction. Shadows become more visually recognizable in VHR imagery. Shadow fraction provides clues about tree vertical structure and forest species type because of the geometric relationship between tree shapes, sun rays, and shadows (Figure 4). The shadow fraction of a given areal unit (segment or window) is the ratio of shaded area to total area within that unit. For this research, shadow fraction was derived using the QB NIR band because it contained the highest contrast between shaded and nonshaded forest areas among the available spectral bands. The shadow fraction was calculated using two steps: (1) a heuristically defined DN threshold of 420 was used to separate the NIR band pixels into shaded and nonshaded groups; and (2) the quotient of the number of shaded pixels to the number of all pixels within each object was calculated to obtain the shadow fraction value. At all scales, the independent variables (i.e., spectral response, image-texture, and shadow fraction) were computed for the individual geo-objects. Similarly, the dependent variable (i.e., canopy height) was calculated by averaging all height values from the lidar CHM within the extent of the corresponding multiscale geo-objects.

3.1.3.4. Regression models. Nonlinear multiple regression models were developed to estimate canopy height for the GEOBIA approach at all 15 scales. After initial tests, a model formulated using a combination of exponential and quadratic forms was found suitable for canopy height estimation:

$$\mathrm{CH} = \exp\left(\sum_{i=0}^{n} (a_i X_i^2 + b_i X_i + c_i)\right) \quad (1)$$

where CH is canopy height; X_i is the ith independent variable; a_i, b_i, and c_i are coefficients for the ith variable; and n is the number of independent variables.

To avoid the overfitting problem common to regression models, correlation coefficients were first calculated between all independent and dependent variables. Each independent variable was evaluated and retained under two rules: (1) its correlation value with any other independent variable is lower than 0.7; and (2) it shows a correlation value greater than 0.7,

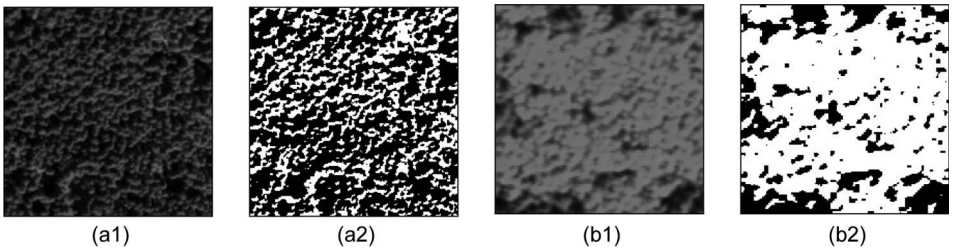

Figure 4. A grayscale Quickbird image converted from a false color image (using NIR, red, and green bands) illustrating forest structure of (a1) conical-shaped conifers, and (b1) irregular-shaped deciduous trees, with their corresponding binary shadow images (a2 and b2). Black tones in shadow images represent shaded forest areas.

but it has the highest correlation value with the canopy height of all other variables with which its correlation value exceeds 0.7. After discarding redundant variables, a stepwise method was used in the regression analysis to develop models at a 0.05 significance level.

To validate our models, a 10-fold cross-validation was performed to estimate the model accuracy. Theory and extensive tests have shown that the fold (i.e., subset) number of 10 can provide a better estimate of error than other fold numbers (Witten and Frank 2005). The 10-fold cross-validation was applied by (1) randomly splitting all object groups into 10 subsets of approximately equal size; (2) training the model using 9 subsets and calculating the errors (i.e., RMSE) for the remaining subset; and (3) repeating the second step 10 times with each of the subsets used only once as validation dataset. For the model validation, it is important to be aware that in the GEOBIA approach the objects are of varying sizes – even at the same scale (i.e., MOS). Consequently, large objects tend to have more influence than small objects in the canopy height error calculation. Therefore, an object area-weighted RMSE was applied using the following equation:

$$\mathrm{RMSE}_{\text{Area-weighted}} = \sqrt{\frac{1}{A_N}\sum_{i=1}^{N}\left[A_i(\mathrm{CH}_{\mathrm{QB\_i}} - \mathrm{CH}_{\mathrm{Lidar\_i}})^2\right]} \quad (2)$$

where $\mathrm{RMSE}_{\text{Area-weighted}}$ is the area-weighted RMSE; $\mathrm{CH}_{\mathrm{QB\_i}}$ is the canopy height calculated from the regression model using QB imagery for the *i*th object; $\mathrm{CH}_{\mathrm{Lidar\_i}}$ is the canopy height measured from lidar data for the *i*th object; A_i is the area size for the *i*th object; A_N is the forest-covered area; and N is the number of objects.

3.2. Pixel-based approach

3.2.1. Plot size selection

We evaluated the plausibility of using geo-objects as areal units to estimate canopy height by comparing the results with those obtained using areal units of fixed arbitrary size and shape, namely, square plots of 20 m side. This size and shape is similar to that used in conventional forest inventories, with the particularity that in our case the plots cover the study area exhaustively, that is, the entire study site was divided into 65,025 adjacent 20 × 20 m plots. Because each pixel in VHR imagery only covers a small portion of a tree area, each plot encompasses at least several trees. Compared with individual pixels, the plots also reduced the errors caused by the co-registration between two different data types (i.e., QB and lidar). Therefore, the plots are akin to pixels of 20 m size and will hereafter be referred to as *pixel-based*.

3.2.2. NDVI classification

Similar to the methods used in the previously described GEOBIA approach, we assigned each plot to one of the three classes (i.e., conifers, deciduous trees, and clearcuts) using the majority rule, that is, each plot was assigned to the most frequent class among the pixels of the classified image within it.

3.2.3. Regression analysis

To make a direct comparison with the GEOBIA approach, the modeling procedure in the pixel-based approach followed the rules of the previous GEOBIA approach by using similar variables and methods (Table 2), including a nonlinear multiple regression model

(Equation 1) and a 10-fold cross-validation method. It should be noted that there are two differences between these two approaches: (1) the areal units for the pixel-based approach are 20 × 20 m square plots instead of geo-objects with varying sizes and shapes; and (2) no weights were considered in the pixel-based approach when calculating RMSE in the validation step, because all plots were of the same size for both canopy types (i.e., conifer and deciduous).

3.3. Multiscale analysis

We also evaluated the object scale (i.e., MOS) that achieves the most accurate height estimates. In forest studies, there are typically two types of methods applied to select the best scale(s): (1) expert judgment with consideration of the specific requirements (Desclée *et al*. 2006, Yu *et al*. 2006, Wulder *et al*. 2007); and (2) the comparison of modeling results derived from various scales (van Aardt *et al*. 2006, Addink *et al*. 2007). Method (1) requires *a priori* knowledge of the study area, which is not always available or known. By applying method (2), regression model-derived R^2 and RMSE are typically plotted at each scale, when continuous biophysical properties (e.g., canopy height and volume) are the research subjects. Typically, the 'best' scale is the one where the highest R^2 and the lowest RMSE are located. This is based on the models that only consider *between-object* canopy height variability, regardless of object size and internal variance. However, objects are the basic areal units for GEOBIA, and their internal variability changes with scale. Therefore, we propose that *within-object canopy height variability* should also be considered when conducting multiscale geo-object-based regression.

To better understand whether within-object canopy height variability affects the model performance and the decision of best scale(s), a new canopy height map was created as the *reference* dataset in the study. All regression model-derived results (from 15 scales) were compared with this dataset. Because individual trees represent meaningful basic objects in a real forest height scene, the reference dataset was generated at the individual tree level by averaging canopy height within each tree crown extent from the lidar CHM. The dataset was created using four steps: (1) a median filter with a 3 × 3 pixel window was used to smooth the lidar CHM; (2) a watershed algorithm (Dougherty and Lotufo 2003) was applied to the smoothed lidar CHM using the tree-top pixel as a local minimum (i.e., seed) for flooding the individual tree areas; (3) a height threshold (i.e., 2.0 m) was used to derive tree crowns by deleting nontree areas (e.g., tree gaps and shrubs); and (4) each tree crown area was filled with the average lidar height data. The average height is used based on previous studies that have demonstrated a high correlation between the average height of lidar returns and forest aboveground biomass and volume (Lefsky *et al*. 2002, Lim *et al*. 2003). To compare the model performance from both GEOBIA- and pixel-based approaches, both sets of results were also evaluated using this reference dataset.

4. Results and discussion

4.1. Performance of spectral response, image-texture, and shadow fraction in canopy height estimation

4.1.1. GEOBIA approach

The performance (i.e., adjusted R^2 and RMSE) of the GEOBIA models at all 15 scales with MOS from 0.04 to 6.00 ha, and that of the pixel-based models at 20 m are shown in Figure 5. The models for deciduous and conifers gave similar performance trends, showing that larger MOS results in higher adjusted R^2 values and a lower RMSE. Specifically, there are two

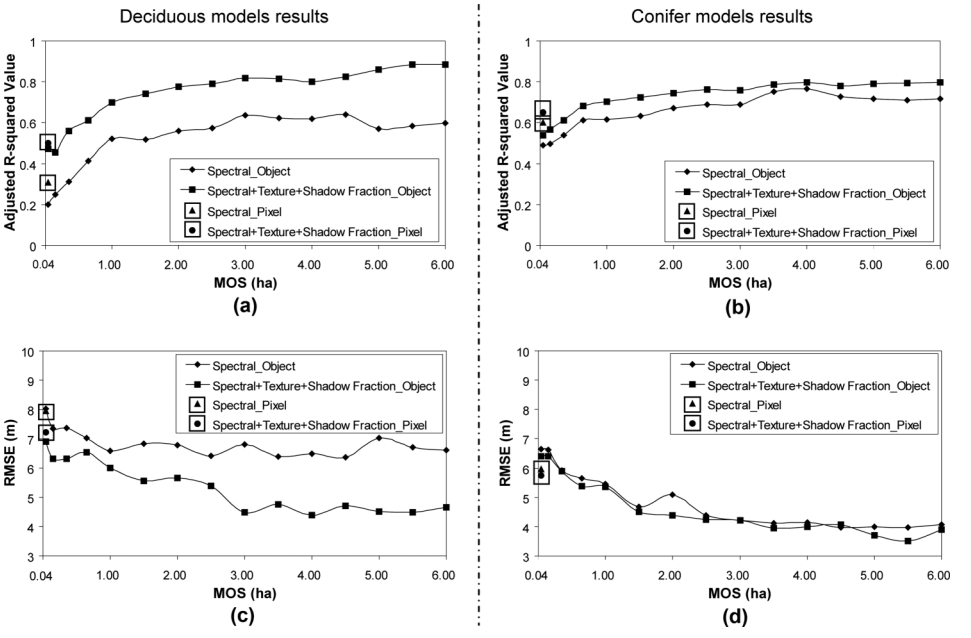

Figure 5. The change of adjusted R^2 values at various mean object sizes (MOS) (i.e., MOS) for (a) deciduous trees and (b) conifers; and the change of RMSE at various MOS for (c) deciduous trees and (d) conifers, using GEOBIA and pixel-based approaches.

clear trends: (1) a relatively strong increase of adjusted R^2 values and a decrease of RMSE occur below the scale of 1.00 ha; and (2) relatively steady trend lines are found above the scale of 1.00 ha, with a slightly decrease in model performance at some scales. However, the object scale plays a more important role for the deciduous trees than conifers. When only spectral response were used in modeling, 20% of the canopy height variance was explained (with an RMSE of 8.03 m) for deciduous trees at the scale of 0.04 ha; whereas for the conifer model, spectral response explained approximately 49% of the canopy height variance (with an RMSE of 6.66 m) at the same scale. With the increase of the MOS, the highest R^2 values reached 64% and 76%, occurring at the MOS of 4.50 ha and 4.00 ha for deciduous trees and conifers, respectively (Figure 5a and b). Compared with conifers (27%), a remarkable change of the R^2 values for deciduous trees (44%) indicates that the two forest types exhibit different sensitivities to the scales applied in our research. A similar condition of tree-type-dependent scale sensitivity is also found in the modeling results when the variables of spectral response, image-texture, and shadow fraction are used in a combined fashion for the canopy height prediction (Figure 5).

The addition of the variables image-texture and shadow fraction increased the model performance for both canopy types at all 15 scales. For deciduous trees, the average increase of R^2 values is approximately 23%, with a 1.47 m decrease in RMSE at all scales (Figure 5a and c) compared with the model performance using spectral response only. For conifers, the values are only 7% and 0.20 m (Figure 5b and d). Overall, our results show that modeling the canopy height of deciduous trees can reach a similar performance as modeling the height of conifers using the three combined types of variables and a GEOBIA approach (Figure 5), even though the biophysical properties for deciduous forests are normally more difficult to model than conical-shaped conifers because of their irregular shaped canopy.

Table 3. Relationships (adjusted R^2) between Quickbird-derived variables and lidar-measured canopy height for deciduous and conifer forests at three sample scales (i.e., mean object sizes).

Input independent variables	Deciduous model			Conifer model		
	Scale 1[a]	Scale 2[b]	Scale 3[c]	Scale 1	Scale 2	Scale 3
Spectral[d]	0.20	0.62	0.60	0.49	0.76	0.72
Spectral + image-texture_1[e]	0.35	0.78	0.84	0.53	0.80	0.78
Spectral + image-texture_2[f]	0.47	0.70	0.69	0.55	0.79	0.77
Spectral + shadow fraction	0.22	0.51	0.49	0.56	0.81	0.80

[a]0.04 ha.
[b]4.00 ha.
[c]6.00 ha.
[d]Spectral response.
[e]Internal-object texture.
[f]GEOTEX.

Table 3 represents an example of the relationships (i.e., adjusted R^2) between various types of QB-derived variables and lidar-measured canopy height for two types of canopies at three sample scales (i.e., 0.04 ha, 4.00 ha, and 6.00 ha). For both canopy types, the addition of image-texture variables (i.e., both internal-object texture and GEOTEX) proved to be useful to increase model performance. It should be noted that, when small scales are used, GEOTEX is more significant than the internal-object texture. This makes sense, as Tobler's first law of geography (Tobler 1970) states that spatially near entities tend to be similar. When the MOS is small, close objects tend to be of similar classes or types; however, this relationship changes at larger MOS. In fact, as MOS gets larger, internal-object texture is better than GEOTEX to estimate canopy height.

The shadow fraction variable plays different roles for the two canopy types. Shadow fraction performs better than image-texture variables when estimating conical-shaped conifers, whereas it decreases the model performance for deciduous forests. One possible explanation for this may be due to the difficulty in creating the deciduous shadow threshold (Section 3.1.3.3 and Figure 4). Because of their complex canopy structure, a single deciduous threshold may not fully capture its shadow content, whereas conifer shadows are simpler, even though their tree heights vary greatly.

4.1.2. Pixel-based image analysis approach

The performance of the pixel-based approach was evaluated using the plot size of 0.04 ha (i.e., a 20 × 20 m window, Figure 5). Similar to the GEOBIA approach, a better result was obtained for conifers (adjusted R^2 = 0.60 and RMSE = 5.97 m) than for deciduous trees (adjusted R^2 = 0.31 and RMSE = 7.94 m) when we used the spectral response only. The inclusion of image-texture and shadow fraction also resulted in a better canopy height estimation for conifers (adjusted R^2 = 0.65 and RMSE = 5.74 m) than for deciduous trees (adjusted R^2 = 0.50 and RMSE = 7.21 m), although a greater improvement (i.e., 20% increase of R^2 with 0.74 m decrease of RMSE) of model performance was achieved for deciduous trees (Figure 5).

4.2. Multiscale analysis

Figure 6 compares the RMSE directly derived from the regression models with the RMSE derived by differentiating the regression model estimated canopy height and the reference

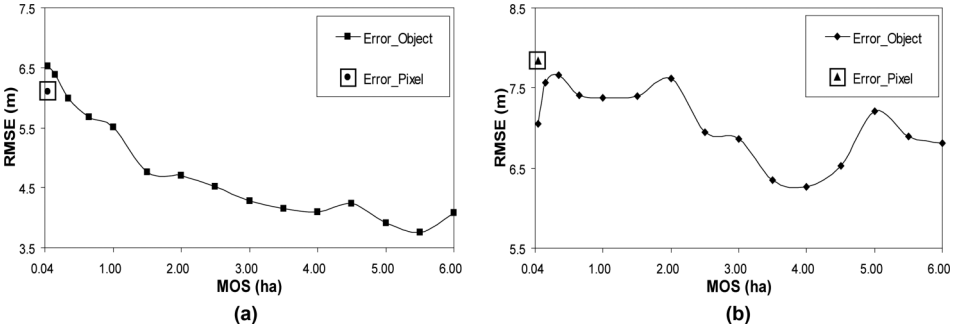

Figure 6. The errors (a) directly derived from the regression models (only considering height variablity between-objects), and (b) derived by differentiating the regression model estimated canopy height and the reference dataset (considering height variability both between- and within-objects) for all trees in the study area at 15 scales.

dataset for all forest canopies in our study area. The former RMSE represents the error resulting from *between-object* canopy height variability, whereas the latter represents the error of both *between-* and *within-object* height variability. The error trend of regression models clearly shows that the regression models developed at larger scales have better performance (Figure 6a). For example, the lowest error of 3.75 m is located at the scale of 5.50 ha. By considering the height variability within objects, Figure 6b shows that the lowest error (i.e., RMSE = 6.27 m) is found at the scale of 4.00 ha. The error trend is also different, as low errors exist between small and large scales. There are two main reasons which might cause the error differences in this study:

(1) Large scales (i.e., large MOS) tend to decrease the canopy height variability between objects and increase the height variability within individual objects. Because typical GEOBIA models are developed only based on the height variability between objects, the results are prone to be of high R^2 and low RMSE. It is therefore expected that a theoretically *best* result may still be obtained, even though the forest patches are undersegmented;
(2) When small MOS are selected, the height variability within each object tends to be low, whereas the height variability between objects is relatively high. This normally causes a low-performance (i.e., large estimation error) GEOBIA model. However, the details (i.e., forest structure) within objects are retained.

To select the *best* object size, both *between-* and *within-object* height variability should be considered, where the errors are derived by differentiating the regression model estimated canopy height and the reference dataset, even though a larger error (6.27 m vs. 3.75 m) might be obtained. Based on these concepts, we chose the MOS of 4.00 ha as the best scale in this study.

Figure 7 illustrates an error map when subtracting regression model estimated canopy height (at the scale of 4.00 ha) from the reference dataset. Approximately 86% of the errors in forests are below 9 m. Some large errors tend to exist in (1) river bank areas, where spectral information were highly affected by topography and heavy tree shadows (Figure 7a1 and a2); and (2) mixed canopy types including both deciduous and conifer forests (Figure 7b1 and b2). Figure 8 represents the change of estimation error and canopy cover at various canopy heights. As canopy cover increases, the estimation accuracy

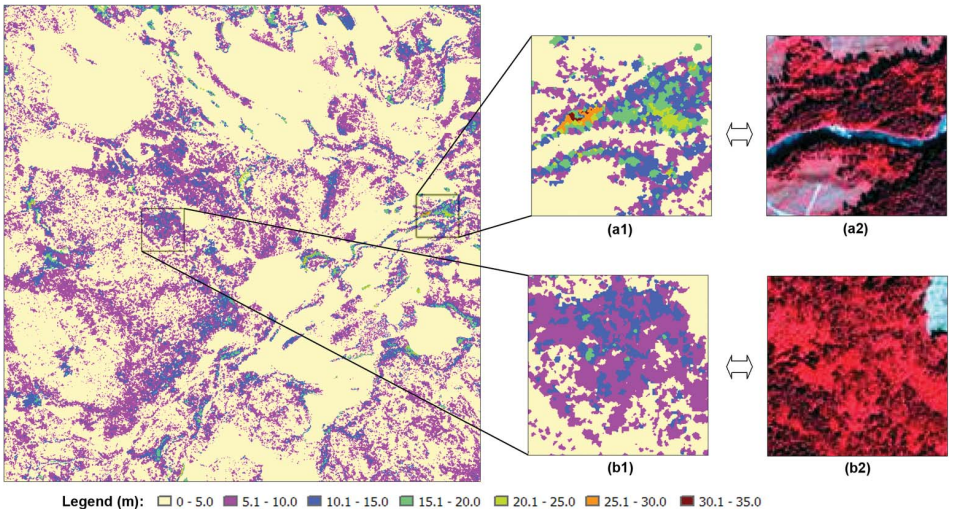

Figure 7. Error map derived at the scale of 4.00 ha with samples representing large errors for (a1) river bank areas and (b1) mixed canopy types, and their corresponding Quickbird images (a2) and (b2).

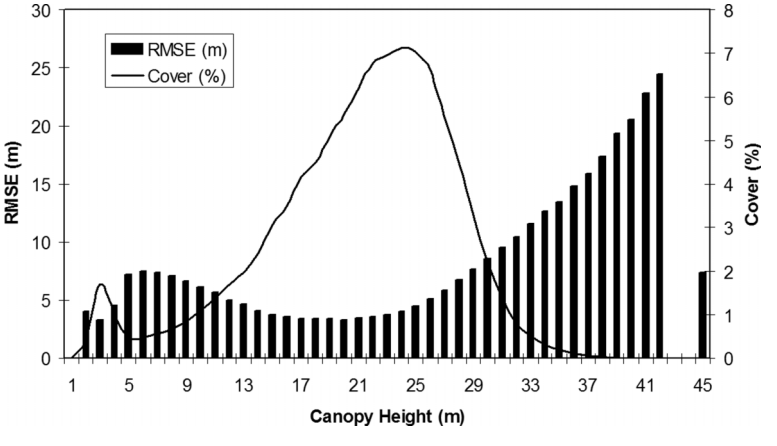

Figure 8. The change of estimation error and canopy cover at various canopy heights.

increases. High accuracy (RMSE <5 m) occurs in two ranges: (1) young trees with canopy heights lower than 5 m; and (2) canopies between 12 and 25 m. When two stands have a similar cover, low canopies (i.e., lower than 20 m) tend to have lower errors than those of tall canopies (i.e., higher than 20 m).

4.3. GEOBIA versus pixel-based approach

The basic study units in the pixel-based approach are arbitrarily determined pixels or plots (e.g., square plots of 0.04 ha) compared with the varying sized and shaped geo-objects. For

both tree types, the pixel-based approach results in a lower accuracy for the estimated canopy height than for most of the GEOBIA models over the 15 tested scales, that is, 15 different MOS (Figures 5 and 6). However, it should be noted that the pixel-based approach still has comparable or better performance than the GEOBIA approach at some very small scales. The potential reason is that segmentation in GEOBIA is highly affected by tree shadows in VHR imagery when small scales are performed. This indicates that selecting an appropriate object scale is crucial for the GEOBIA approach, because the selection of arbitrary scales may result in poorer canopy height accuracy than for a canopy height derived with a pixel-based approach.

5. Conclusions

In this study, we investigated the potential of VHR QB imagery to estimate canopy height and compared the results with a CHM derived from small-footprint lidar data. GEOBIA and pixel-based models were developed for two types of trees: deciduous and conifers. A novel object area-weighted error calculation approach (Section 3.1.3.4), which considered the importance of object size, was also developed to evaluate the RMSE of the CHM derived from GEOBIA. Results show that VHR QB imagery can be used to accurately estimate canopy height using a GEOBIA approach with approximately 86% of canopy height errors less than one British Columbia forest inventory height class (i.e., ~9 m).

Specific results show that the addition of image-texture [both internal-object texture and GEOTEX, a dynamic neighboring image-texture measure (Section 3.1.3.2)] increased GEOBIA model performance over those using only spectral response. Shadow fraction (Section 3.1.3.3) was also found to more significantly improve the model performance for conifers than other variables. However, it increased error for deciduous forests. Overall, the addition of image-texture and shadow fraction resulted in an R^2 increase of approximately 23% with a 1.47 m decrease in RMSE for deciduous trees, compared with the model performance using spectral response only. For conifers, the values were 7% and 0.20 m, respectively.

The best object scale (i.e., MOS) for defining the most accurate canopy height estimates using GEOBIA was a MOS of 4.00 ha. When estimating continuous forest biophysical properties (e.g., canopy height), typical GEOBIA modeling approaches tend to focus only on the performance of regression models (i.e., between-object variability), regardless of object size and internal variability. In this study, we evaluated both the *between-* (as typically modeled) and the *within-object* height variability as they represent two unique perspectives of forest structure. Specifically, forest-objects should exhibit (relatively) low internal variance. When this variance significantly changes (through scale), the forest class is no longer representative of its initial class conditions and thus requires reevaluation.

This study also compared GEOBIA and pixel-based approaches for canopy height estimation. The best GEOBIA models achieved better results than those developed using only the pixel-based approach. Furthermore, we confirm that the selection of an appropriate object scale plays an important role in GEOBIA (Addink *et al*. 2007); as an arbitrary scale selection can result in accuracies that may be lower than those derived from a pixel-based approach.

Future research will focus on evaluating whether forest canopy height can be accurately estimated from individual lidar transects, rather than a full lidar dataset covering the entire study area. If this is possible, we will evaluate whether narrow lidar transects and VHR broad extent optical imagery can be integrated with GEOBIA to accurately define canopy height, biomass, and volume for complex forest cover over very large areas.

Acknowledgments

This research has been supported by an Alberta Ingenuity Fund (AIF) and an Alberta Informatics Circle of Research Excellence (iCore) PhD scholarship awarded to Gang Chen. Dr. Hay also acknowledges the support from an NSERC (Natural Sciences and Engineering Research Council) discovery grant, an AIF New Faculty Award and the University of Calgary – Foothills Facility for Remote Sensing and GIScience. The authors acknowledge the valuable discussions with Dr. Stefan Steiniger at the University of Calgary, Mike Stoutley at Terra Remote Sensing Inc., and the insightful comments from two anonymous reviewers.

References

Addink, E.A., de Jong, S.M., and Pebesma, E.J., 2007. The importance of scale in object-based mapping of vegetation parameters with hyperspectral imagery. *Photogrammetric Engineering and Remote Sensing*, 73, 905–912.

Axelsson, P., 1999. Processing of laser scanner data – algorithms and applications. *ISPRS Journal of Photogrammetry & Remote Sensing*, 54, 138–147.

Castilla, G., Hay, G.J., and Ruiz, J.R., 2008. Size–constrained region merging (SCRM): an automated delineation tool for assisted photointerpretation. *Photogrammetric Engineering and Remote Sensing*, 74, 409–419.

Culvenor, D., 2003. Extracting individual tree information: a survey of techniques for high spatial resolution imagery. *In*: M. Wulder and S. Franklin, eds. *Remote sensing of forest environments: concepts and case studies*. Boston: Kluwer Academic Publishers, 255–277.

Desclée, B., Bogaert, P., and Defourny, P., 2006. Forest change detection by statistical object-based method. *Remote Sensing of Environment*, 102, 1–11.

Donoghue, D.N.M. and Watt, P.J., 2006. Using LiDAR to compare forest height estimates from IKONOS and Landsat ETM+ data in Sitka spruce plantation forests. *International Journal of Remote Sensing*, 27, 2161–2175.

Dougherty, E.R. and Lotufo, R.A., 2003. *Hands-on morphological image processing*. Bellingham: SPIE Optical Engineering Press.

Franklin, S.E. and McDermid, G.J., 1993. Empirical relations between digital SPOT HRV and CASI spectral response and lodgepole pine (*Pinus contorta*) forest stand parameters. *International Journal of Remote Sensing*, 14, 2331–2348.

Hay, G.J. and Castilla, G., 2008. Geographic object-based image analysis (GEOBIA). *In*: T. Blaschke, S. Lang, and G.J. Hay, eds. *Object-based image analysis – spatial concepts for knowledge-driven remote sensing applications*. Berlin: Springer-Verlag, 77–92.

Hay, G.J. and Niemann, K.O., 1994. Visualizing 3-D texture: a three dimensional structural approach to model forest texture. *Canadian Journal of Remote Sensing*, 20, 90–101.

Hay, G.J., Niemann, K.O., and McLean, G.F., 1996. An object-specific image-texture analysis of H-resolution forest imagery. *Remote Sensing of Environment*, 55, 108–122.

Hudak, A.T., et al., 2002. Integration of Lidor and Landsat ETM+ data for estimating and mapping forest an canopy height. *Remote Sensing of Environment*, 82, 397–416.

Hutchinson, M.F., 1989. A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *Journal of Hydrology*, 106, 211–232.

Hyde, P., *et al*., 2006. Mapping forest structure for wildlife habitat analysis using multi-sensor (LiDAR, SAR/InSAR, ETM+, Quickbird) synergy. *Remote Sensing of Environment*, 102, 63–73.

Kraus, K. and Pfeifer, N., 1998. Determination of terrain models in wooded areas with airborne scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 54, 193–203.

Lefsky, M.A., *et al*., 2002. Lidar remote sensing of above-ground biomass in three biomes. *Global Ecology & Biogeography*, 11, 393–399.

Lim, K., *et al*., 2003. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. *Canadian Journal of Remote Sensing*, 29, 658–678.

Means, J.E., *et al*., 1999. Use of large-footprint scanning airborne lidar to estimate forest stand characteristics in the Western Cascades of Oregon. *Remote Sensing of Environment*, 67, 298–308.

Tobler, W.R., 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography*, 46, 234–240.

Van Aardt, J.A.N., Wynne, R.H., and Oderwald, R.G., 2006. Forest volume and biomass estimation using small-footprint lidar-distributional parameters on a per-segment basis. *Forest Science*, 52, 636–649.

Wang, L., Gong, P., and Biging, S., 2004. Individual tree-crown delineation and treetop detection in high-resolution aerial imagery. *Photogrammetric Engineering and Remote Sensing*, 70, 351–357.

Welch, R. and Ahlers, W., 1987. Merging multiresolution SPOT HRV and landsat TM data. *Photogrammetric Engineering & Remote Sensing*, 53, 301–303.

Witten, I.H. and Frank, E., 2005. *Data mining: practical machine learning tools and techniques* (*Second* Edition). San Francisco: Morgan Kaufmann.

Wulder, M.A. and Seemann, D., 2003. Forest inventory height update through the integration of lidar data with segmented Landsat imagery. *Canadian Journal of Remote Sensing*, 29, 536–543.

Wulder, M.A., *et al*., 2004a. High spatial resolution remotely sensed data for ecosystem characterization. *BioScience*, 54, 511–521.

Wulder, M.A., *et al*., 2004b. Estimating time since forest harvest using segmented Landsat ETM+ imagery. *Remote Sensing of Environment*, 93, 179–187.

Wulder, M.A., *et al*., 2007. Integrating profiling LIDAR with Landsat data for regional boreal forest canopy attribute estimation and change characterization. *Remote Sensing of Environment*, 110, 123–137.

Yu, Q., *et al*., 2006. Object-based detailed vegetation classification with airborne high spatial resolution remote sensing imagery. *Photogrammetric Engineering and Remote Sensing*, 72, 799–811.