Methods for improving the quality of a true orthomosaic of Vexcel UltraCam images created using a lidar digital surface model

Benoît St-Onge[1]

[1]University of Quebec at Montreal, Quebec, Canada st-onge.benoit@uqam.ca

Abstract

The combined analysis of lidar and image datasets for information extraction of forest structural attributes and composition requires that the image-to-lidar geometric correspondence be known accurately. We propose a series of methods for producing a 10 cm high quality true orthomosaic of Vexcel UltraCam images perfectly adjusted to the lidar digital surface model (DSM). First, we introduce a technique for filling the small cavities visible on lidar raster DSMs. We then assess the image-to-lidar registration using visualization and quantitative approaches. The small geometric discrepancy measured between the two datasets is then corrected. In the image overlap areas, the true orthomosaic is created by choosing the contributing image that has the smallest distance to the corresponding DSM pixel. Occluded pixels that can not be seen from any centre of perspective are then filled with synthetic values calculated according to their sunlit/shadowed state at the time the images were taken. The resulting true orthomosaic is perfectly registered to the lidar dataset, is complete (considering occluded pixels receive synthetic values), is not radiometrically altered, and shows no visible cut lines. The proposed process should greatly help the simultaneous analysis of lidar and image datasets.

Keywords: orthorectification, co-registration, mosaicking, UltraCam, filtering

1. Introduction

There is currently a clear trend towards the combined use of lidar and digital aerial images to produce fine scale forest maps useful in resource, biodiversity, and carbon inventories. Typically, structural attributes (height, density, etc.) are derived from the lidar canopy height model (CHM) while species composition and health are extracted from the images. A basic approach consists of processing both data types separately and integrating the results later in the process. This way of doing is however far from exploiting all the synergies between the lidar and image datasets. A more sophisticated approach would be to analyse all data simultaneously by associating image reflectance values to lidar height measurements. This requires that the images be "perfectly" registered to the lidar surface model and that true orthorectification (i.e. including visibility calculation, as per Mayr 2002) be used. Moreover, in the case of low altitude digital aerial photography, the image resolution is so high that the texture of individual crowns could be used as an additional criterion for species identification. This brings the requirement that the image texture should be modified as little as possible by the orthorectification and mosaicking processes. What is more, radiometric analysis of an orthomosaic must not be affected by such things as image-to-image histogram balancing and feathering along cut lines. All these requirements are currently not met by commercial orthorectification software. In addition, small misregistration problems cause the image to drape imperfectly over the 3D surface, which produces geometric warping in the resulting orthoimage, thus modifying crown texture and shape. Therefore, this study's aim is to:

1. Achieve the best possible registration between high resolution aerial images and a lidar dataset so that each tree crown image is mapped exactly on the corresponding 3D lidar crown shape.

2. Create a seamless and complete true orthomosaic with the best possible preservation of radiometric and textural properties of the original aerial images.

This brings us to solve the following problems: the removal of small artefact cavities in the lidar digital surface model (DSM) in order to create a better reference 3D surface for the orthorectification, the detection and correction of subtle misregistrations between the lidar and the images caused by small direct georeferencing (GPS and IMU) errors, to find a way to take advantage of the high aerial image overlap in order to fill as much orthoimage pixels as possible with no radiometric modification, and to fill the orthoimage's empty pixels (resulting from occlusion) with plausible values.

2. Study region and materials

The study site falls within the Training and Research Forest of Lake Duparquet (TRFLD, 79°22'W, 48°30'N), in the Province of Quebec, Canada. It is characterized by small hills with elevations comprised between 227 m and 335 m. The mixed vegetation is composed of common boreal species, and dominated by balsam firs (*Abies balsamea L.* [Mill.]), paper birch (*Betula papyfifera* [Marsh.]), and trembling aspen (*Populus tremuloides* [Michx]). Most stands are mature or over-mature and reach heights of 25-30 m.

The lidar data was acquired on July 12th 2007 using an Optech ALTM3100 flown at approx. 650 m AGL. Strip overlap was sufficient to avoid data gaps. The density of the first returns was approx. 3.2 hits m^{-2} (single density, i.e. outside strip overlaps). The lidar *Z* data was delivered as ellipsoidal heights. The GPS antenna/receiver was a Novatel and the IMU an Applanix AV510. The GPS and IMU were integrated in a tightly coupled solution using the POS AV 1.6 and POS Pac 4.3 software by Applanix. The vertical datum used was the GRS80 ellipsoid. The reference height of the base antenna was obtained from a geodetic point which Z data was expressed in CGVD28 orthometric height, which was converted to GRS80 using the CGG00E geoid undulation value.

The Vexcel UltraCam images were taken in full blue sky conditions on June 9th 2007 at approx. 1000 m AGL resulting in a 10 cm ground pixel size. Image overlap was approximately 80%. Only the panchromatic images were used in this study. The calibrated internal orientation parameter values where obtained from the aerial survey provider. Base data was logged using NovAtel DL GPS receivers with NovAtel Model 600-LB GPS antennas while the airborne platform used the Applanix POSAV 510 system. One base antenna was placed on a monument distinct from the one used for the lidar survey. Another one was set up as a spike and operated simultaneously. The two ground survey points were processed as a network with the geodetic point held fixed. Data was delivered as CGVD28 orthometric heights. We have converted those to GRS80 ellipsoidal heights using the GPS-H software from Natural Resources Canada and the CGG00E undulation table.

3. Methods and results

3.1 DSM creation

The design of the raster DSM creation process aims at producing a canopy surface grid that corresponds as closely as possible to reality in order to achieve a high quality of orthorectification. Because trees are not solid objects with a well defined surface but rather a hierarchical network of branches, twigs and leaves, the raster canopy surface is an abstraction of the true interface, along a vertical column, between the tree material and the medium in which the laser energy propagates. We here define the canopy surface, at the geoposition of a given pixel, as the elevation at which all the tree material (within the spatial extension of the pixel) is found immediately below, and the open medium immediately above. The creation of the DSM involves gridding the first return elevations and then filling some of the surface cavities.

3.1.1 Gridding

The gridding procedure follows Vepakomma et al. (2008). First the pixels of an empty grid having a 10 cm resolution (equivalent to that of the images) are given the corresponding elevation values of the first returns. If more then one return falls into a given pixel, only the maximum value is retained. The remaining empty pixels are then filled by interpolated values obtained using the inverse distance weighed (IDW) method of ArcGIS 9.2 applied to the first return point data. A small portion of the resulting DSM in shown in figure 1a.

3.1.2 Cavity filling

Lidar DSMs of a tree canopy surface will normally show numerous small cavities (also reported in other studies, such as Leckie et al. 2003) of two types. “Drill holes” are caused by a near vertical laser shots penetrating in small openings of a crown and generating returns well below the generalized crown hull. They represent the true canopy surface elevation at a particular geoposition but cause a deep hole in the surface which may affect the orthoimage geometric quality. A more important problem is the presence of “overhang holes”, artefacts caused by oblique laser shots travelling close to the side of a crown and intercepting a first surface under the crown, close to ground level. The normal gridding process integrates these points to the canopy surface, generating deep cavities (easily visible in figure 1a) that must be removed. We propose a process that first detects cavities and then fills them with interpolated values.

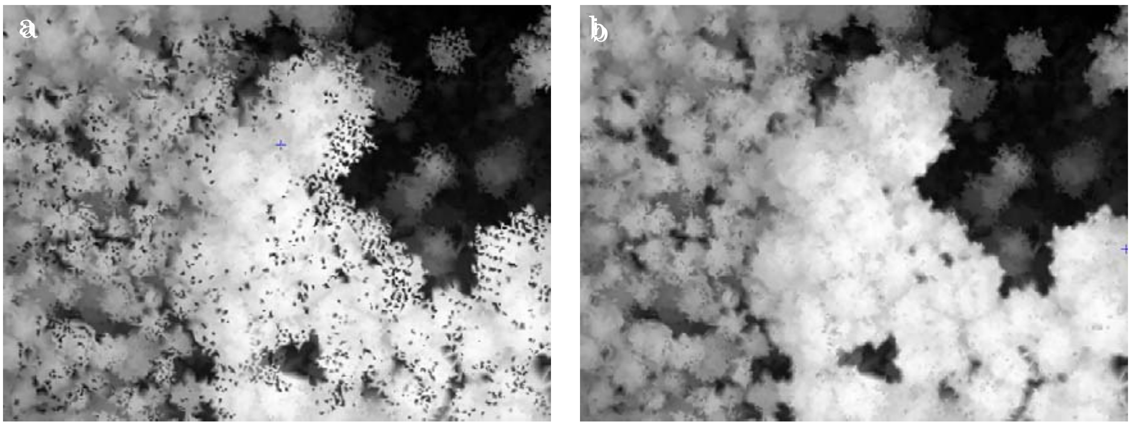

Figure 1 - a) Initial raster DSM showing cavities, and b) DSM after cavity filling. Image width is 64.4 m.

A circular Laplacian filter with negative values near the centre and positive values on the periphery is used to produce scores that reflect the likeliness that a cavity is present. These scores are then thresholded to produce a binary map showing the cavity locations. The filter radius determines the size of the detected cavity while the threshold value controls the depth of detected cavities relative to the surrounding pixels. We have empirically chosen these parameters’ values such that the sharp elevation drop at the crowns’ edges is not considered a cavity. Once the cavities are mapped, they are slightly dilated to ensure that the full extent of the cavity is captured. Cavity pixels are then given an interpolated value (IDW interpolation using the closest valid DSM pixels). In this study, we processed the DSM in two passes with the following parameter values for the Laplacian filter and dilatation radii respectively, pass 1: 3 and 2 pixels, pass 2: 1 and 0 pixel (no dilatation). The first pass removed most of the overhang holes while the second pass removed many small drill holes. These algorithms were coded in C language (as all the other special purpose functions presented in this paper). The resulting DSM is presented in figure 1b. It can be seen that most cavities are filled, that the crown edges are not modified and that natural gaps between crowns are not filled. Note also that the rest of the canopy surface, including tree apex elevations, is not affected by the cavity filling process.

3.2 Misregistration assessment and correction

To ensure accurate orthorectification, the exact image to 3D scene correspondence must be established. Our goal is to first assess if systematic errors affect this correspondence, and then to apply the appropriate corrections. Various approaches have been proposed in the literature using area or feature based matching between the lidar surface and the images (e.g. Mitishita et al. 2008). Many of these rely on the presence of buildings or man-made structures. In this study, the landscape is almost entirely covered by vegetation and devoid of buildings. We therefore propose alternatives for working in these conditions, i.e. two visualization techniques and one quantitative method. Visualization is used to detect and understand the nature of the misregistration problems. Various system or manipulation errors in the independent direct georeferencing of the respective datasets can translate into a systematic misfit taking the form of an *XY* translation (e.g. caused by a horizontal datum error), a *Z* translation (e.g. caused by a vertical datum error), rotations (e.g. caused by errors in the boresight matrix or IMU drift. The first visualization technique consisted of draping the CHM image onto the cavity-filled DSM and to project it onto the UltraCam image plane according to the uncorrected camera orientation parameter values. If this synthetic rendering is well adjusted to the corresponding real image, we conclude that the initial uncorrected orientation is correct. Otherwise, we study the amount and directions of the discrepancy to comprehend the nature and origin of the misfit.

Figure 2 shows UltraCam image subsets and corresponding synthetic images in which brightness values are proportional to the CHM heights. The red contours correspond to the silhouettes of certain trees digitized according to the real aerial image. When transferred to the synthetic image, we see that for a subset located at the image centre (nadir view), the image contours (2a) correspond well to the projected CHM morphology (2c). However, for the subset taken near the corner of the image (2b), there is a clear displacement in the radial direction (2d). A similar misfit was visible at all image corners, for all inspected images. This strongly suggests that there is a scale problem, i.e., a *Z* offset between the image and lidar vertical datums. The other visualization technique consisted of orthorectifying a few overlapping images according to the uncorrected orientation and looking at the fit quality between the orthoimages over well defined and smooth surfaces (road segments). This analysis revealed that discrepancies existed (planimetric offsets in the corresponding patches of the orthoimages, not shown in this paper), corroborating the fact that a slight misregistration between the lidar and the images existed.

The quantitative analysis was based on computing the *XYZ* position of image conjugate points and comparing the *Z* value to the corresponding lidar elevation. Thirty points were measured on a set of four consecutive images. Precise features (inflections in the shadow silhouettes visible on road segments, e.g. figure 6) were used. Individual points fell on two to four images. Spatial intersection was computed for all intersecting rays for each possible pair of rays. For example, a conjugate point falling on all four test images yielded six different estimates of *XYZ* (1-2, 1-3, 1-4, 2-3, 2-4, 3-4). For each pair, the *Y* estimate was calculated based on Y_{01}, Z_{01} (*Y* and *Z* of the perspective centre for first image) and Y_{02}, Z_{02} according to Kraus and Waldhaüsl (1994), giving two different *Y* estimates. Conjugate point coordinates leading to *Y* discrepancies of more than 20 cm were rejected, leaving only very accurate points. The average of the two *Y* values for each valid point was used as the *Y* estimate. The lidar elevation (Z_L) was read from the DSM at the ray intersection position *XY* for all valid combination of rays for all 30 points and subtracted from the photogrammetric Z_P, yielding 43 observations of elevation difference (Z_P - Z_L). The average difference revealed a 1.41 m bias. The Z_0 value of the centre of projection of all images was corrected for that bias. After this correction, new synthetic images were produced (figure 3). No significant change is visible on the centre sub-image (3a), but the red contours on the corner subset now fit exactly with the visible CHM morphology (3b). Using the second visualization technique, we observed that the individual orthoimages now overlap almost perfectly. This signifies that the image to lidar correspondence is now solved with great accuracy.

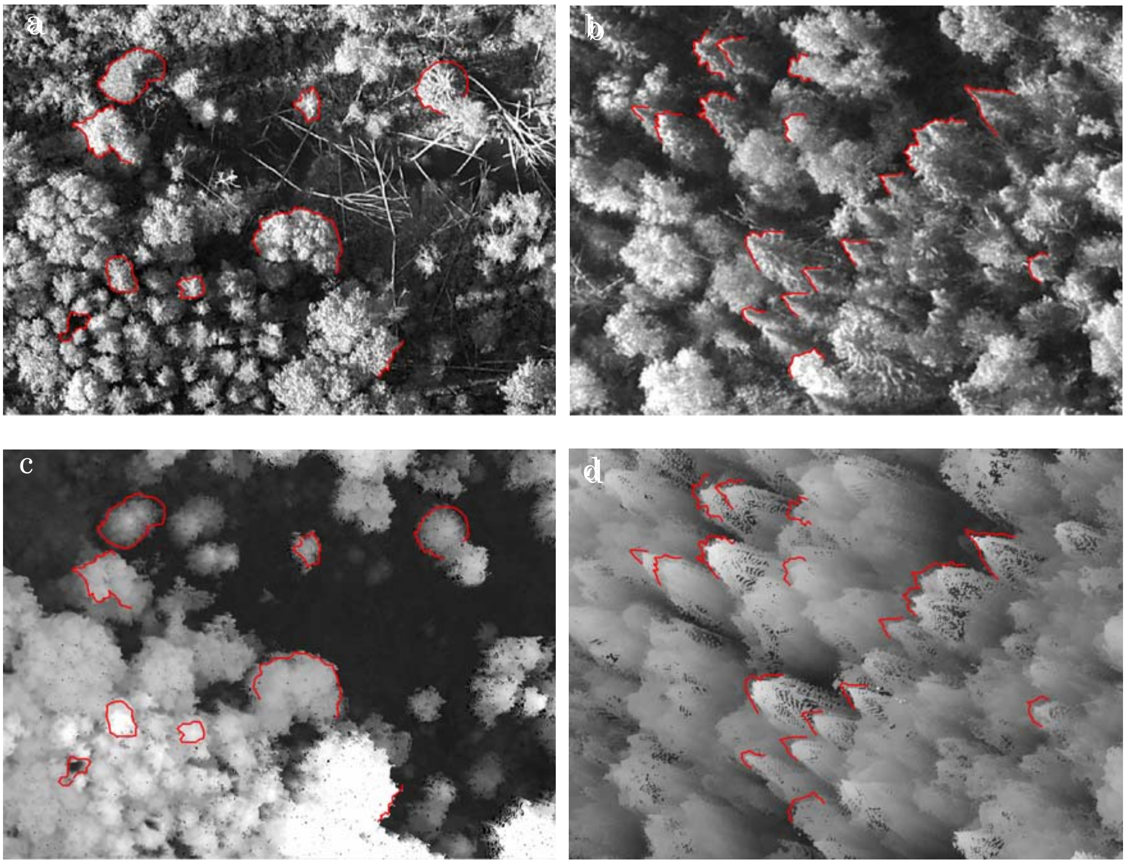

Figure 2 - a & b: two Vexcel UltraCam subimages, respectively extracted at the centre and near the top left corner of the sample image; c & d: CHM image draped over the DSM and projected into the Vexcel image plane according to the uncorrected orientation. The red contours were manually digitized on the real images and overlaid afterwards on the synthetic ones.

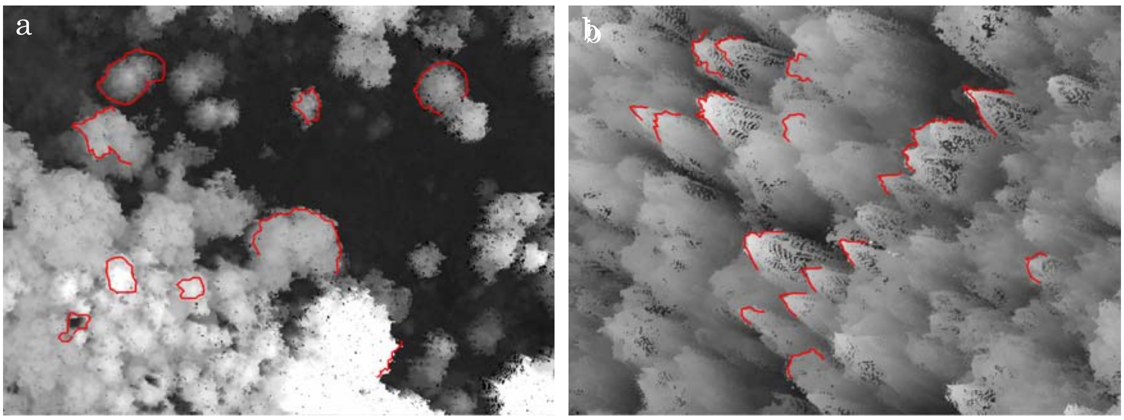

Figure 3 – Same as 2 c & d, after the absolute orientation was corrected.

3.3 True orthorectification and mosaicking

The goal of the orthorectification was to produce an orthomosaic with the best possible geometric and radiometric quality. True orthorectification calculates pixel visibility based on the 3D scene model (i.e. the lidar DSM) and leaves occluded pixels empty (Mayr 2002). In our proposed approach, a non-occluded ortho-pixel receives the value of the closest image. Proximity is established by retaining the image with the shortest distance between the DSM pixel and the centre of perspective, among all images encompassing this pixel. Visibility is computed using the classical *z*-buffer technique. This proximity approach has the following advantages: 1) the image with the most vertical view angle is always used (helping preserve the quality of the image texture and reducing BRDF effects), 2) a DSM pixel occluded in the closest image still receives a value if it can be seen from an other image, and 3) cut lines between images are optimal and automatically calculated. Moreover, we avoided inter-image histogram balancing as well as feathering along the cut lines (which are actually very complex due to the interspersed image contributions, e.g. figure 5b) between images to preserve the original radiometry.

Figure 4 presents an overview of the orthomosaic of four overlapping images (4a) with the map of the contributing image identity (4b). We see that at this scale the image appears seamless, that the contributing images each have their "principal domain" but contribute locally outside their domain. Note that black pixels indicate that the corresponding DSM pixel could not be seen from any of the images. Figure 5 shows an enlarged orthomosaic subset (5a) centred on a principal cut line between two images (visible in 5b and as a red dashed line in 5a). Even at this large scale the transition between the two images is invisible and the integrity of crown shapes and texture is preserved. Similar quality can be found along all cut lines.

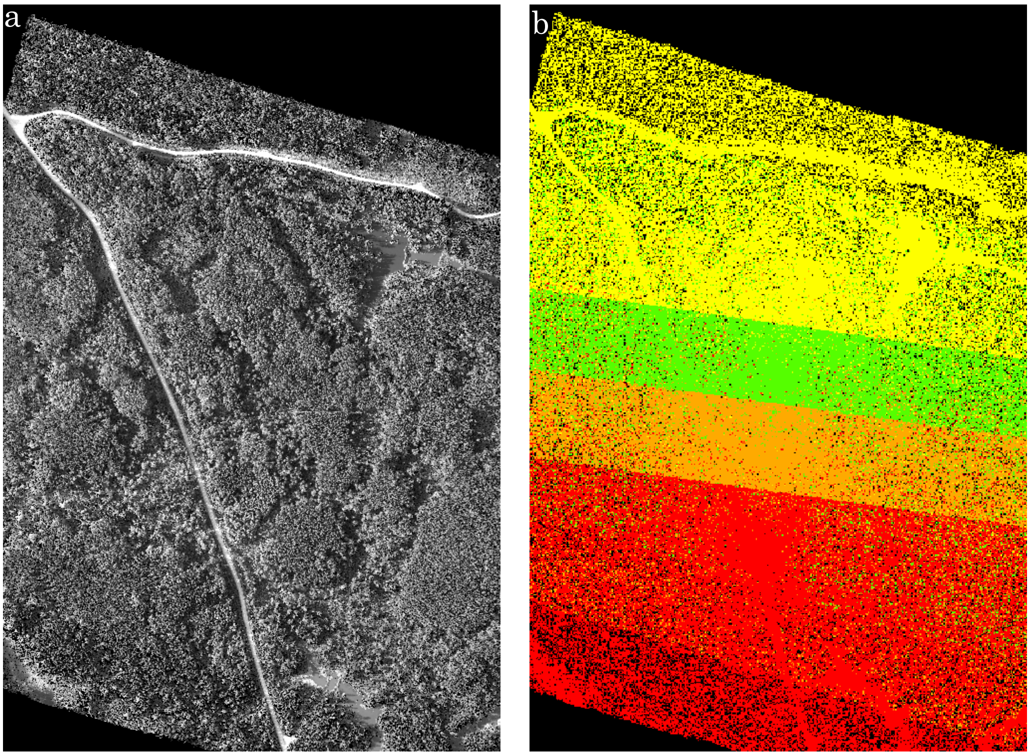

Figure 4 – a) Orthomosaic created from four Vexcel UltraCam images, b) Color-coded identification of the contributing image. No data pixels (in black) result from occlusion or absence of image data. Note that the nearest neighbour resampling used to reduce the scale of 4b for this figure degrades resolution.

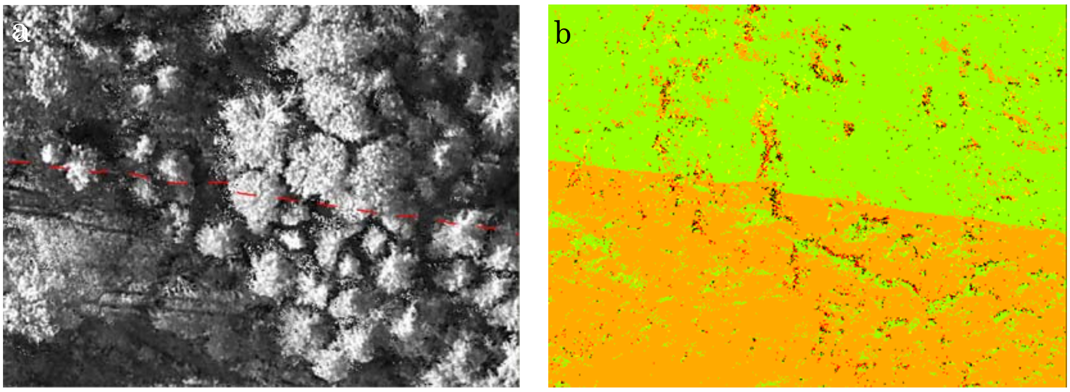

Figure 5 – Enlargement of a portion of 4 a & b. The principal cut line visible in 5b (transition between the green and orange zones) is drawn as a dashed red line in a.

3.4 Cosmetic filling

After the true orthorectification and mosaicking is complete, the DSM pixels occluded in all images have a no-data value. Although this represents the image acquisition reality, no-data pixels may be detrimental to the visual analysis of the orthomosaic, or cause problems for certain automated tree delineation algorithms. We have therefore produced a cosmetically filled version of the orthomosaic based on the following process. First, a map of projected shadows was calculated based on the lidar DSM. The average time at which the four test images were taken was extracted from the Vexcel images' GPS time stamps (images were taken at approx. 1.5 seconds intervals). The NOAA *Solar Position Calculator* was used to obtain the elevation and azimuth of the sun at the time of image capture. Shadows were computed based on the lidar DSM by modelling sun ray's as parallel lines. No-data orthomosaic pixels received a synthetic value which depended on the sunlit/shadowed state of the corresponding DSM pixel. Observations led us to conclude that, as a first approximation and pending further developments, a constant value for sunlit tree pixels would be acceptable. The overall average image value of visible sunlit trees was calculated and assigned to the sunlit no-data (occluded) pixels. A more complex solution was sought for the shadowed pixels. The high radiometric sensitivity of the Vexcel UltraCam makes individual trees discernable even in the shadowed areas. It appears that an important factor affecting the shadowed tree pixel values is the amount of diffuse sky irradiance received by each pixel. This quantity depends on height of the target pixel relative to that of its neighbours. Again as a first approximation, we have calculated the volume of tree material above the target pixel elevation in a 400 m^2 area centred on the target pixel. These volume values were regressed with all corresponding non-occluded shadow pixel image values, resulting in a model for predicting shadowed pixel brightness based on the 3D layout of its surroundings. This model was used to attribute a brightness value to occluded shadowed pixels.

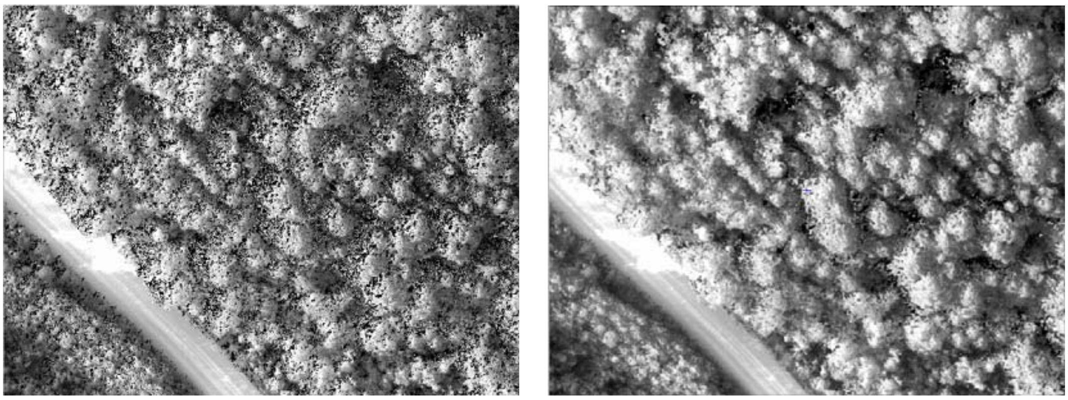

Figure 6 – a) raw orthomosaic (black pixel have no-data value), and b) orthomosaic after cosmetic filling.

As a final step, to remove small artefacts caused by the *z*-buffer technique used to calculate visibility, we have used a 3x3 Laplacian filter to detect anomalies (e.g. a single very bright pixel surrounded by dark values) and replaced their values by the 3x3 average calculated using the non-anomalous pixels. Only a very small portion of the orthomosaic is affected by this filtering. Figure 6 shows a subset of the "raw" orthomosaic (6a) and the corresponding cosmetically filled version. We observe that all pixels receive a value, and that for the most part, these values are sufficiently close to the plausible brightness levels as to improve visualization, and most probably, individual tree delineation.

4. Concluding remarks

The methods presented in this paper allow the creation of high quality complete and true orthomosaics with very few artefacts. As these orthoimages fit exactly with a lidar dataset, the simultaneous analysis of 3D and radiometric data at high resolution will likely be significantly improved. The cavity filling ensures that the image texture and crown shapes are not degraded during the rectification. The assessment of the misregistration between the datasets used in this study revealed a Z error of 1.41 m, despite all the care put in the positioning of the base GPS antennae, system mounting, direct georeferencing data collection and integration, and vertical datum transformations. Even such a small error affects the geometric quality of the orthoimages. Although the source of this error was not identified, the techniques presented here guarantee that the error itself can be detected and corrected. Moreover, note that using comparisons of Z_P and Z_L, other types of errors could be detected and empirically corrected. However, there is a limit to the registration between image and lidar data due to wind sway of the trees. In the presence of moderate or strong winds, the top of trees will sway by several decimetres such that the crown image and its 3D shape may not fit exactly even though the image and lidar datasets would be tightly registered. Improvements in the orthomosaic creation are still possible. We noted for example that the brightness values of a given crown viewed from two different consecutive images, although very similar, are significantly different. Radiometric calibration would therefore be a useful addition to the mosaicking process. What is more, rough radiometric approximations were made based on the sun and scene geometries to cosmetically fill occluded pixels. More sophisticated approaches involving rigorous radiometric modelling could be employed instead. Nevertheless, we have already achieved a level of image-lidar integration quality that should allow fruitful developments in the field of precise forest mapping.

Acknowledgements

We thank the National Science and Engineering Council of Canada for funding, and Chris Hopkinson for his much appreciated flexibility in scheduling the lidar survey.

References

Kraus, K. and Waldhäusl, P., 1994. *Photogrammetry, Fundamentals and Standard.*

Leckie, D., Gougeon, F., Hill, D., Quinn, R., Armstrong., L., Shreenan, R., 2003. Combined high-density lidar and multispectral imagery for individual tree crown analysis. *Canadian Journal for Remote Sensing 29*, 633-649.

Mayr, W., 2002, True orthoimages. *GIM International*, 16, April, pp. 37–39.

Mitishita, E., Habib, A, Centeno, J., Machado, A, Lay, J., and Wong, C, 2008. Photogrammetric and lidar data integration using the centroid of a rectangular roof as a control point. *The Photogrammetric Record*, 23, 19-35.

Vepakomma, U., St-Onge, B. and Kneeshaw, D., 2008. Spatially explicit characterization of boreal forest gap dynamics using multi-temporal lidar data. *Remote Sensing of Environment*, 112, 2326-2340.