IMPROVED BUILDING DETECTION USING TEXTURE INFORMATION

Mohammad Awrangjeb, Chunsun Zhang and Clive S. Fraser

Cooperative Research Centre for Spatial Information,University of Melbourne
Level 5, 204 Lygon Street, Carlton Vic 3053, Australia
Phone: +61 3 8344 9182, Fax: +61 3 9349 5185
Email: {mawr, chunsunz, c.fraser}@unimelb.edu.au

Working Groups III/4

KEY WORDS: Building, Detection, LIDAR, Orthoimage, Fusion, Texture, Classification.

ABSTRACT:

The performance of automatic building detection techniques can be significantly impeded due to the presence of same-height objects, for example, trees. Consequently, if a building detection technique cannot distinguish between trees and buildings, both its false positive and false negative rates rise significantly. This paper presents an improved automatic building detection technique that achieves more effective separation of buildings from trees. In addition to using traditional cues such as height, width and colour, the proposed improved detector uses texture information from both LIDAR and orthoimagery. Firstly, image entropy and colour information are jointly applied to remove easily distinguishable trees. Secondly, a voting procedure based on the neighbourhood information from both the image and LIDAR data is employed for further exclusion of trees. Finally, a rule-based procedure using the edge orientation histogram from the image is followed to eliminate false positive candidates. The improved detector has been tested on a number of scenes from three different test areas and it is shown that the algorithm performs well in complex scenes.

1 INTRODUCTION

Building detection from remotely sensed data has a number of practical applications including city planning, homeland security and disaster management. Consequently, a large number of building detection techniques have been reported over the last few decades. Since photogrammetric imagery and LIDAR (LIght Detection And Ranging) data have their own merits and demerits, the recent trend is to integrate data from both of these sources as a means of advancing building detection by compensating the disadvantages of one with the advantages of the other.

The success of automatic building detection is still largely impeded by scene complexity, incomplete cue extraction and sensor dependency of data (Sohn and Dowman, 2007). Vegetation, and especially trees, can be the prime cause of scene complexity and incomplete cue extraction. Image quality may vary for the same scene even if images are captured by the same sensor, but at different times. The situation also becomes complex in hilly and densely vegetated areas where only a few buildings are present, these being surrounded by trees. Important building cues can be completely or partially missed due to occlusions and shadowing from trees. Therefore, many existing building detection techniques that depend largely on colour information exhibit poor detection performance.

Application of a recently developed building detection algorithm (Awrangjeb et al., 2010a) has shown it to be capable of detecting buildings in cases where cues are only partially extracted. For example, if a section of the side of a roof (at least 3m long) is correctly detected, the algorithm can also detect all or part of the entire building. However, this detector does not necessarily work well in complex scenes when buildings are surrounded by dense vegetation and when they have the same colour as trees, or where trees are other than green.

This paper presents an improved detection algorithm that uses both LIDAR and imagery. In addition to exploiting height, width and colour information, it uses different texture information in order to differentiate between buildings and trees. Firstly, image entropy and colour information are employed together to remove the trees that are easily distinguishable. Secondly, a voting procedure that considers neighbourhood information is proposed for the further exclusion of trees. Finally, false positive detections are eliminated using a rule-based procedure based on the edge orientation histogram. The improved detector has been tested on a number of scenes covering three different test areas [1].

2 CUES TO DISTINGUISH TREES AND BUILDINGS

Cues employed to help distinguish trees from buildings include the following:

- **Height**: A height threshold (2.5m above ground level) is often used to remove low vegetation and other objects of limited height, such as cars and street furniture (Awrangjeb et al., 2010a). The height difference between first and last pulse DSMs (digital surface models) have also been used (Khoshelham et al., 2008).

- **Width, area and shape**: If the width or area of a detected object is smaller than a threshold, then it is removed as a tree (Awrangjeb et al., 2010a). A number of shape attributes can be found in (Matikainen et al., 2007).

- **Surface**: A plane-fitting technique has been applied to non-ground LIDAR points to separate buildings and trees (Zhang et al., 2006), and a polymorphic feature extraction algorithm applied to the first derivatives of the DSM in order to estimate the surface roughness has also been employed (Rottensteiner et al., 2007).

- **Colours**: While a high NDVI (normalised difference vegetation index estimated using multispectral images) value represents a vegetation pixel, a low NDVI value indicates

---

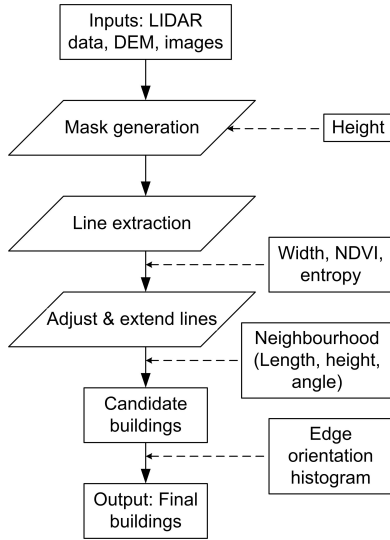[1] This paper is a condensed version of (Awrangjeb et al., 2011).

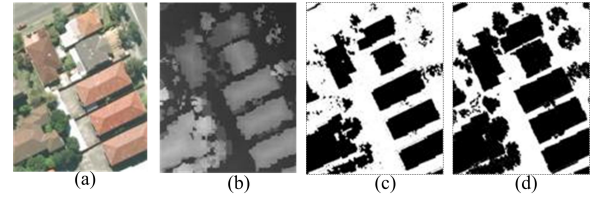Figure 1: Flow diagram of the improved building detection technique.

Figure 2: (a) Image of a test scene, (b) corresponding LIDAR data (in gray-scale), (c) primary mask and (d) secondary mask.
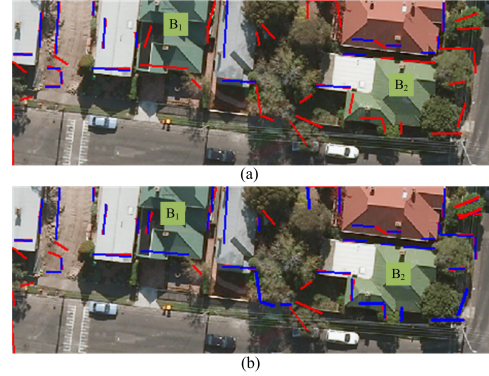

Figure 3: Detection of green buildings: (a) the NDVI information alone missed green buildings whereas (b) combined NDVI and entropy information detects green buildings. 'Blue' lines are accepted, 'red' represents rejected.

a non-vegetation pixel. This cue, although frequently used, has been found unreliable even in normal scenes where trees and buildings have distinct colours (Awrangjeb et al., 2010a). K-means clustering was applied on multispectral images to obtain spectral indices for clusters like trees, water and buildings (Vu et al., 2009). Colour invariants have also been used (Shorter and Kasparis, 2009). A number of other cues generated from colour image and height data can be found in (Matikainen et al., 2007, Salah et al., 2009).

- **Texture**: When objects have similar spectral responses, the grey level co-occurrence matrix (GLCM) can be estimated from the image to quantify the co-occurrence probability (Chen et al., 2006). Some GLCM indices, eg mean, standard deviation, entropy and homogeneity, have been applied to both height and image data in order to classify buildings and trees (Salah et al., 2009, Matikainen et al., 2007).

- **Training pixels**: Training pixels of different colours from roofs, roads, water, grass, trees and soil have been used for classification (Lee et al., 2003).

- **Filtering**: Morphological opening filters have been employed to remove trees attached to buildings (Yong and Huayi, 2008).

- **Others**: Segmentation of LIDAR intensity data can also be used to distinguish between buildings and trees (Maas, 2001). The density of raw LIDAR data has also been employed (Demir et al., 2009).

3   IMPROVED BUILDING DETECTION

The proposed improved detector employs a combination of height, width, angle, colour and texture information with the aim of more comprehensively separating buildings from trees. Although cues other than texture were used in the earlier version of the detector, the improved formulation makes use of additional texture cues such as entropy and the edge orientation histogram at four stages of the process, as shown in Fig. 1. Different steps of the detection algorithm have been presented in (Awrangjeb et al., 2010a). This paper focuses on how texture, dimensional and colour information can be applied jointly in order to better distinguish between buildings from trees. The setup of different threshold values are discussed in (Awrangjeb et al., 2011).

3.1   Application of Height Threshold

A height threshold $T_h = H_g + 2.5$m, where H_g represents the ground height, is applied to the raw LIDAR data and two building masks are created – the primary M_p and secondary M_s masks (Awrangjeb et al., 2010a). This threshold removes low height objects (grounds, grass, roads, cars etc.) and preserves non-ground points (trees and buildings). The corresponding DEM height for a given LIDAR point is used as the ground height. If there is no corresponding DEM height for a given LIDAR point, the average DEM height in the neighbourhood is used. Fig. 2 shows the two extracted masks for a scene.

3.2   Use of Width, NDVI and Entropy

The black areas in M_p are either buildings, trees or other elevated objects. Line segments around these black shapes in M_p are formed, and in order to avoid detected tree-edges, extracted lines shorter than the minimum building width $L_{min} = 3$m are removed. Trees having small horizontal area are thus removed.

The mean of the NDVI value is then applied, as described in (Awrangjeb et al., 2010a), to eliminate trees having large horizontal area. However, the NDVI has been found to be an unreliable cue even in normal scenes where trees and buildings have distinct colours (Rottensteiner et al., 2007, Awrangjeb et al., 2010a). In addition, it cannot differentiate between trees and green buildings. Fig. 3(a) shows an example where a green building B_1 cannot be detected at all since all lines around it are rejected. However, green building B_2 can be partially detected because it has a white coloured roof section. In some areas there may be non-green buildings having the same colour as trees, especially when leaves change colour in different seasons. In such cases, the removal of trees based on the NDVI will result in many buildings also being removed. Detection of these same buildings will likely also lead to detection of trees.

If the mean NDVI is above the NDVI threshold at any side of a line segment, a further test is performed before removing this
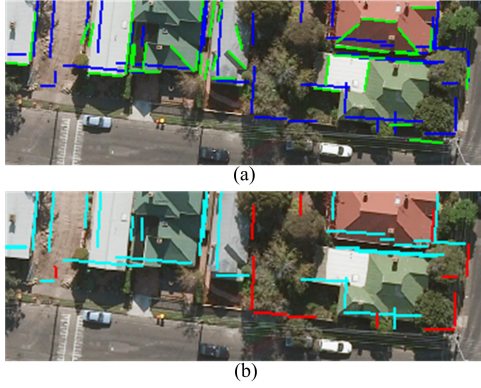
(a)

(b)

Figure 4: Use of neighbourhood information to remove tree-edges: (a) before voting: 'blue' represents lines from the primary mask after the extending procedure and 'green' represents lines from the image and (b) after voting: 'cyan' represents accepted lines after the voting procedure and 'red' represents rejected lines.

line segment as a tree-edge. This test checks whether the average entropy is more than the entropy threshold $T_{ent} = 30\%$. If the test holds, the line segment is removed as a tree edge, otherwise it is selected as a building edge. Fig. 3(b) shows that the green buildings B_1 and B_2 can be fully detected using this approach. In addition, some of the trees subject to shadowing and self-occlusion are also detected.

3.3 Voting on the Neighbourhood Information

The joint application of NDVI and entropy can remove some large trees; however, in the case when there are shadows and self-occlusions within trees, difficulties with the approach can be expected. Therefore, for each of the extended lines a voting procedure based on the information within the neighbourhood of that line is followed.

All the extracted and extended lines that reside around the same black shape in the primary mask M_p fall into the same neighbourhood. Let $\Omega = \{l_i\}$, $0 \leq i \leq n_t$ be such a neighbourhood obtained after the application of the width threshold L_{min} in the previous section, where l_i indicates an extracted line, its length $L_{l_i} \geq 3$m, and there are a total of n_t extracted lines. Furthermore, let n_e lines, out of n_t extracted lines in Ω, survive after the extending procedure discussed above, with the average length of these being $L_{\Omega,avg}$. We also consider the longest image line, extracted from the grey-scale orthoimage, which resides around l_i. The longest local image line ℓ_i for l_i within a rectangular area of width 3m around l_i is obtained. Let the length of ℓ_i be L_{ℓ_i}. In some cases, no ℓ_i may be found due to poor image contrast or if l_i is a tree edge. Fig. 4(a) shows the extended lines from M_p and the accepted lines from the orthoimage.

For each line l_i in the proposed voting procedure, four votes v_k, $0.0 \leq v_k \leq 1.0$ are cast by exploiting its neighbourhood information as follows:

- $v_1 = 1.0$, if $L_{l_i} \geq L_{\Omega,avg}$; else $v_1 = \frac{L_{\Omega,avg} - L_{l_i}}{L_{\Omega,avg}}$.

- $v_2 = \frac{\Theta - \theta_i}{\Theta}$, where θ_i is the adjustment angle between l_i and the longest line in Ω, which was used as the base line in the adjustment procedure, and $\Theta = \frac{\pi}{8}$ is the angle threshold used in the adjustment procedure (Awrangjeb et al., 2010a).

- $v_3 = \frac{n_e}{n_t}$. This is based on the observation that line segments around a building are more likely to be adjusted, which
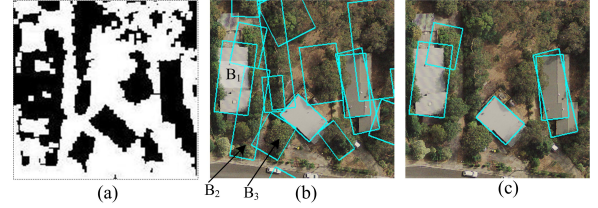

(a)   (b)   (c)

Figure 5: A complex scene: (a) primary mask, (b) detected candidate buildings with a large number of false detections and (c) detected final buildings after removing false positives.

means that they are either parallel or perpendicular to the base line around the same black shape in M_p.

- $v_4 = 1.0$ if $L_{\ell_i} \geq 2L_{min}$; else $v_4 = \frac{L_{\ell_i} - L_{min}}{L_{min}}$. If there is no image line found around l_i, then $v_4 = 0.0$.

The voting procedure is executed for n_e lines in Ω. A line l_i is designated a building edge if it obtains a majority vote. This means that the mean of v_k, $1 \leq k \leq 4$, is greater than 0.50. Fig. 4(b) shows that the majority of tree edges can be removed by applying the voting procedure. A candidate building set is then obtained using the extended lines that survive the voting procedure (Awrangjeb et al., 2010a).

In areas with dense vegetation, the black shapes of buildings and nearby trees are not separable and consequently a building may be connected with another building a few metres away (see Fig. 5). If the connected buildings are not parallel to each other, then the improved adjustment procedure will likely still fail. This is why in the improved detection algorithm, the adjustment and voting procedure is available as an optional step, the choice of which will depend upon vegetation density. In either case, there may be some false buildings present in the candidate building set, as shown in Fig. 5(b). A procedure utilising the edge orientation histogram from the orthoimage is then applied in order to remove false positives.

3.4 Application of Edge Orientation Histogram

Following the detection of candidate buildings, a gradient histogram is formed using the edge points within each candidate building rectangle. Edges are first extracted from the orthophoto using an edge detector and short edges (less than 3m in length) are removed. Each edge is then smoothed and the gradient (tangent angle) is calculated on each point using the first order derivatives. The gradient will be in the range $[-90°, +90°]$. A histogram with a successive bin distance of $D_{bin} = 5°$ is formed using the gradient values of all edge points lying inside the candidate rectangle.

Rectangles containing the whole or major part of a building should have one or more significant peaks in the histogram, since edges detected on building roofs are formed from straight line segments. All points on an apparent straight line segment will have a similar gradient value and hence will be assigned to the same histogram bin, resulting in a significant peak. A significant peak means the corresponding bin height is well above the mean bin height of the histogram. Since edge points whose gradient falls into the first (at $-90°$ to $-85°$) and last (at $85°$ to $90°$) bins have almost the same orientation, located peaks in these two bins are added to form a single peak.

Fig. 6 illustrates three gradient histogram functions and mean heights for candidate buildings B_1, B_2 and B_3 in Fig. 5(b). Two bins at $\pm 90°$ basically form one bin, because lines in these two
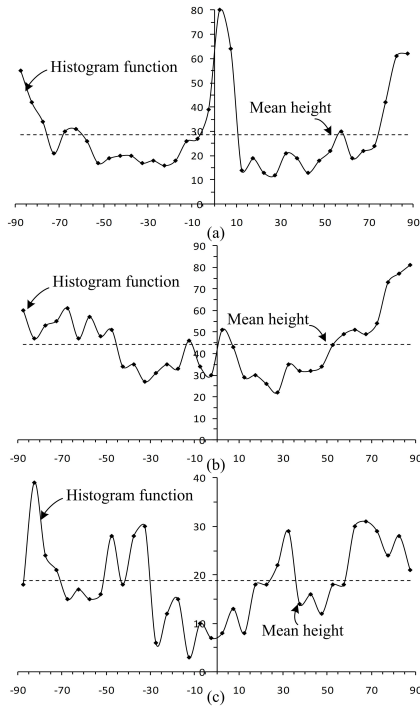
Figure 6: Gradient histogram functions and means for rectangles (a) B_1, (b) B_2 and (c) B_3 in Fig. 5(b): x-axis is in degrees and y-axis is in pixels (bin heights).

bins are perpendicular to the x-axis and reside above & below this axis. Therefore, these can be a peak at either of these bins and their heights can be accumulated to form a single peak. Fig. 6(a) shows that B_1 has two significant peaks: 80 pixels at $0°$ and 117 $(55+62)$ pixels at $\pm 90°$, these being well above the mean height of 28.6 pixels. The two significant peaks separated by $90°$ strongly suggest that this is a building. From Fig. 6(b) it can be seen that B_2 has one significant peak at $\pm 90°$ but a number of insignificant peaks. This points to B_2 being partly building but mostly vegetation, which is also supported by the high mean height value. With the absence of any significant peak, but a number of insignificant peaks close to the mean height, Fig. 6(c) indicates that B_3 is comprised of vegetation. Although there may be some significant peaks in heavily vegetated areas, a high average height of bins between two significant peaks can be expected. Note that the orthophoto resolution in this case was 10cm, so a bin height of 80 pixels indicates a total length of 8m from the contributing edges.

The observations above support the theoretical inferences. In practice, however, detected vegetation clusters can show the edge characteristics of a building, and a small building having a flat roof may not have enough edges to show the required peak properties. As a result, some true buildings can be missed, while some false buildings may be detected. A number of precautions can be formulated in order to minimize the occurrence of false detections.

Two types of histograms are formed using edges within each detected rectangle. In the first type, one histogram considers all the edges collectively, and in the second type histograms for individual edges whose length is at least L_{min} are formed. Let the collective histogram be symbolized as H_{col}, with an individual histogram being indicated by H_{ind}. Tests on H_{col} and H_{ind} can be carried out to identify true buildings and remove trees. If a detected rectangle passes at least one of the following tests it is selected as a building, otherwise it is removed as vegetation.

1. *Test 1*: H_{col} has at least two peaks with heights of at least $3L_{min}$ and the average height of bins between those peaks is less than $2L_{min}$. This test ensures the selection of a large building, where at least two of its long perpendicular sides are detected. It also removes vegetation where the average height of bins between peaks is high.

2. *Test 2*: The highest bin in H_{col} is at least $3L_{min}$ in height and the aggregated height of all bins in H_{col} is at most 90m. This test ensures the selection of a large building where at least one of its long sides is detected. It also removes vegetation where the aggregated height of all bins is high.

3. *Test 3*: H_{col} has at least two peaks with heights of at least $2L_{min}$, and the highest bin to mean height ratio R_{Mm1} is at least 3. This test ensures the selection of a medium size building, where at least two of its perpendicular sides are detected. It also removes vegetation where the highest bin to mean height ratio is low.

4. *Test 4*: The highest bin in H_{col} has a height of at least L_{min} and the highest bin to mean height ratio R_{Mm2} is at least 4. This test ensures the selection of a small or medium size building where at least one of its sides is at least partially detected. It also removes small to moderate sized vegetation areas where the highest bin to mean height ratio is low.

5. *Test 5*: The highest bin in H_{ind} has a height of at least L_{min} and the aggregated height of all bins in H_{col} is at most 90m. This test ensures the selection of buildings which are occluded on at most three sides.

6. *Test 6*: The ratio R_{aTp} of the detected rectangular area to the number of texture pixels (N_{Tp}, the aggregated height of all bins in H_{col}) is at least 45. This test ensures the selection of all buildings which are at least partially detected but the roof sides are missed.

The application of these tests on the complex scene in Fig. 5(b) produces the result shown in Fig. 5(c). Note that for simple scenes with small amounts of vegetation, the NDVI and entropy together can successfully remove most trees so subsequent application of the voting procedure and edge orientation histogram can be considered as optional, leading to a saving of computation time.

4 RESULTS AND DISCUSSIONS

The threshold-free evaluation system involved in the performance study conducted makes one-to-one correspondences using nearest centre distances between detected and reference buildings. The descriptor 'threshold-free' means the evaluation system does not involve any thresholds based on human choice. Some 15 evaluation indices in three categories, namely object-based, pixel-based and geometric, have been employed. Whereas pixel-based evaluation considers only spectral properties in the imagery, object-based evaluation takes into account spatial and contextual properties in both the imagery and LIDAR data. The root mean square positional discrepancy value (RMSE) is employed to quantify the geometric accuracy. The detailed procedure of the threshold-free evaluation system and the evaluation indices can be found in (Awrangjeb et al., 2010b).

The test data sets employed cover three suburban areas in Australia, Fairfield, NSW; Moonee Ponds, Victoria and Knox, Victoria. The Fairfield data set covers an area of 588m $\times$ 417m and contains 370 buildings, Moonee Ponds covers 447m $\times$ 447m and

has 250 buildings and Knox covers 400m × 400m and contains 130 buildings. Fairfield contains many large industrial buildings and in Mooney Ponds there were some green buildings. Knox can be characterized as outer suburban with lower housing density and extensive tree coverage that partially covers buildings. In terms of topography, Fairfield and Mooney Ponds are relatively flat while Knox is quite hilly.

LIDAR coverage comprised last-pulse returns with a point spacing of 0.5m for Fairfield, and first-pulse returns with a point spacing of 1m for Moonee Ponds and Knox. For Fairfield and Knox, RGB colour orthoimagery was available, with resolutions of 0.15m and 0.1m, respectively. Moonee Ponds image data comprised RGBI colour orthoimagery with a resolution of 0.1m. Bare-earth DEMs of 1m horizontal resolution covered all three areas.

Reference data sets were created by monoscopic image measurement using the Barista software [2]. All rectangular structures, recognizable as buildings and above the height threshold T_h, were digitized. The reference data included garden sheds, garages, etc. These were sometimes as small as 10m^2 in area.

Tables 1 to 3 show results of the object-based, pixel-based and geometric accuracy evaluations of the improved building detection algorithm in the three test areas. A visual illustration of sample building detection results are shown in Fig. 7. The improved algorithm produced moderately better performance than the original in all three evaluation categories within both Fairfield and Mooney Ponds. The better performance was mainly due to proper detection of large industrial buildings in Fairfield, detection of some green buildings in Mooney Ponds, and elimination of trees in both Fairfield and Mooney Ponds.

In Knox, the improved algorithm exhibited significantly better performance over the original, due to two main reasons. Firstly, the improved algorithm better accommodated the dense tree cover and randomly oriented buildings that characterized the Knox data. Fairfield and Mooney Ponds on the other hand are low in vegetation cover and buildings are generally well separated and more or less parallel or perpendicular to each other. Secondly, the improved algorithm showed its merits in better handling varying topography. Knox is a hilly area (maximum height $H_M = 270$m and minimum height $H_m = 110$m), whereas Fairfield ($H_M = 23$m and $H_m = 1$m) and Mooney Ponds ($H_M = 43$m and $H_m = 23$m) are moderately flat.

The original algorithm detected a large number of false buildings in Knox, as illustrated in Figs. 7 (a) and (c). Moreover, many buildings detected with the original algorithm were not properly aligned. Consequently, in object-based evaluation, 56% quality was observed with 77% completeness and 67% correctness. The reference cross-lap rate was above 85%, with 39% detection overlap rate. In pixel-based evaluation, 27% quality was found with 44% completeness and 42% correctness. The area omission error was more than 50% and both branching and miss factors were above 120%. The geometric accuracy was no better than 33 pixels.

In contrast, as shown for Knox in Figs. 7 (b) and (d), the improved detector removed a large number of false buildings using its orientation histogram. In object-based evaluation, when compared to the original algorithm, the quality increased to 82%, a 26% rise. The detection overlap rate decreased to 13% and the reference cross-lap rate reduced to 62%. In pixel-based evaluation, again when compared to the original algorithm, the quality went up to 39%, a 12% growth, while the branching factor

[2]The Barista Software, www.baristasoftware.com.au, May 2011.

Table 1: Object-based evaluation results in percentages (C_m = completeness, C_r = correctness, Q_l = quality, M_d = multiple detection rate, D_o = Detection overlap rate, C_{rd} = detection cross-lap rate and C_{rr} = reference cross-lap rate).

Scenes	C_m	C_r	Q_l	M_d	D_o	C_{rd}	C_{rr}
Fairfield	95.1	95.4	92.2	2.7	8.6	3.5	9.7
MPonds	94.5	95.3	89.2	6.2	13.1	7.3	17.5
Knox	93.2	87.2	82.0	9.3	12.8	23.3	61.6
Average	**94.0**	**91.3**	**86.4**	**6.9**	**11.9**	**14.4**	**37.6**

Table 2: Pixel-based evaluation results in percentages (C_{mp} = completeness, C_{rp} = correctness, Q_{lp} = quality, A_{oe} = area omission error, A_{ce} = area commission error, B_f = branching factor and M_f = miss factor).

Scenes	C_{mp}	C_{rp}	Q_{lp}	A_{oe}	A_{ce}	B_f	M_f
Fairfield	83.2	84.5	72.4	15.3	12.5	13.5	20.3
MPonds	87.2	85.4	75.3	12.7	13.2	16.7	17.3
Knox	49.0	65.9	39.1	51.0	30.9	51.8	104.0
Average	**73.1**	**78.6**	**62.3**	**26.3**	**18.9**	**27.3**	**47.2**

declined dramatically to 52% and the miss factor was also moderately improved to 104%. Geometric accuracy improved to 20 pixels, or by approximately 50%.

In object-based evaluation, the improved algorithm offered on average across the three data sets a more than 10% increase in completeness and correctness and a 15% increase in quality. Multiple detection and detection overlap rates were also low. In pixel-based evaluation, there was also a reasonable rise in completeness (4%), correctness (10%) and quality (7%). Area omission and commission errors were less than those obtained with the original algorithm. In addition, there was a 5 pixel improvement in geometric accuracy.

5 CONCLUSIONS

This paper has presented an improved automatic building detection technique that exhibits better performance in separating buildings from trees. In addition to employing height and width thresholds and colour information, it uses texture information from both LIDAR and colour orthoimagery. The joint application of measures of entropy and NDVI helps in the removal of vegetation by making trees more easily distinguishable. The voting procedure incorporates neighbourhood information from the image and LIDAR data for further removal of trees. Finally, a rule-based procedure based on the edge orientation histogram from the image edges assists in eliminating false positive building candidates. The experimental results reported showed that while the improved algorithm offered moderately enhanced performance in Fairfield and Mooney Ponds, it yielded a very significant improvement in performance in Knox across all three evaluation categories.

ACKNOWLEDGEMENTS

REFERENCES

Awrangjeb, M., Ravanbakhsh, M. and Fraser, C. S., 2010a. Automatic detection of residential buildings using lidar data and multispectral imagery. ISPRS Journal of Photogrammetry and Remote Sensing 65(5), pp. 457–467.

Figure 7: Building detection by the previous (left) and the improved (right) algorithms on two samples from Knox.

Table 3: Geometric accuracy.

Scenes	metres	pixels
Fairfield	2.4	16.0
Mooney Ponds	1.6	16.0
Knox	2.0	20.0
Average	**2.0**	**17.3**

Awrangjeb, M., Ravanbakhsh, M. and Fraser, C. S., 2010b. Building detection from multispectral imagery and lidar data employing a threshold-free evaluation system. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38(part 3A), pp. 49–55.

Awrangjeb, M., Zhang, C. and Fraser, C. S., 2011. Infomration removed to facilitate the blind review system. Submitted to IS-PRS Journal of Photogrammetry and Remote Sensing.

Chen, L., Teo, T., Hsieh, C. and Rau, J., 2006. Reconstruction of building models with curvilinear boundaries from laser scanner and aerial imagery. Lecture Notes in Computer Science 4319, pp. 24–33.

Demir, N., Poli, D. and Baltsavias, E., 2009. Extraction of buildings using images & lidar data and a combination of various methods. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 38(part 3/W4), pp. 71–76.

Khoshelham, K., Nedkov, S. and Nardinocchi, C., 2008. A comparison of bayesian and evidence-based fusion methods for automated building detection in aerial data. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37(part B7), pp. 1183–1188.

Lee, D., Shan, J. and Bethel, J., 2003. Class-guided building extraction from ikonos imagery. Photogrammetric Engineering and Remote Sensing 69(2), pp. 143–150.

Maas, H. G., 2001. The suitability of airborne laser scanner data for automatic 3d object reconstruction. In: Proc. 3rd International Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images, Ascona, Switzerland, pp. 345–356.

Matikainen, L., Kaartinen, H. and Hyypp, J., 2007. Classification tree based building detection from laser scanner and aerial image data. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 36(part 3/W52), pp. 280–287.

Rottensteiner, F., Trinder, J., Clode, S. and Kubik, K., 2007. Building detection by fusion of airborne laser scanner data and multi-spectral images : Performance evaluation and sensitivity analysis. ISPRS Journal of Photogrammetry and Remote Sensing 62(2), pp. 135–149.

Salah, M., Trinder, J. and Shaker, A., 2009. Evaluation of the self-organizing map classifier for building detection from lidar data and multispectral aerial images. Journal of Spatial Science 54(2), pp. 1–20.

Shorter, N. and Kasparis, T., 2009. Automatic vegetation identification and building detection from a single nadir aerial image. Remote Sensing 1(4), pp. 731–757.

Sohn, G. and Dowman, I., 2007. Data fusion of high-resolution satellite imagery and lidar data for automatic building extraction. ISPRS Journal of Photogrammetry and Remote Sensing 62(1), pp. 43–63.

Vu, T., Yamazaki, F. and Matsuoka, M., 2009. Multi-scale solution for building extraction from lidar and image data. International Journal of Applied Earth Observation and Geoinformation 11(4), pp. 281–289.

Yong, L. and Huayi, W., 2008. Adaptive building edge detection by combining lidar data and aerial images. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37(part B1), pp. 197–202.

Zhang, K., Yan, J. and Chen, S. C., 2006. Automatic construction of building footprints from airborne lidar data. IEEE Trans. on Geoscience and Remote Sensing 44(9), pp. 2523–2533.