Evaluation of the Self-Organizing Map Classifier for Building Detection from Lidar Data and Multispectral Aerial Images

M. Salah
J. Trinder
A. Shaker

Integration of aerial images and lidar data compensate for the individual weaknesses of each data set when used alone, thus providing more accurate classification of terrain cover, such as buildings, roads and green areas, and advancing the potential for automation of large scale digital mapping and GIS database compilation. This paper presents work on the development of automatic feature extraction from multispectral aerial images and lidar data. A total of 22 feature attributes have been generated from the aerial image and the lidar data which contribute to the detection of the features. The attributes include those derived from the Grey Level Co-occurrence Matrix (GLCM), Normalized Difference Vegetation Indices (NDVI), and standard deviation of elevations and slope. A Self-Organizing Map (SOM) was used for fusing the aerial image, lidar data and the generated attributes for building detection. The classified images were then processed through a series of image processing techniques to separate the detected buildings. Results show that the proposed method can extract buildings accurately. Compared with a building reference map, 95.5 percent of the buildings were detected with a completeness and correctness of 83 percent and 80 percent respectively for buildings around 100m^2 in area; these measures increased to 96 percent and 99 percent respectively for buildings around 1100m^2 in area. Further, the contributions of lidar and the individual attributes to the quality of the classification results were evaluated.

KEY WORDS: *Aerial Images, Lidar, GLCM, Attributes, Building Detection, Classification, Self-Organizing Map.*

M. Salah
J. Trinder
School of Surveying and Spatial Information Systems,
The University of New South Wales, Sydney NSW 2052, Australia
m.gomah@unsw.edu.au
j.trinder@unsw.edu.au

A. Shaker
Department of Surveying Engineering,
Shoubra Faculty of Engineering,
Benha University
Egypt
ahmshaker@link.net

INTRODUCTION

Research on automated feature extraction from aerial images and lidar data has been fuelled in recent years by the increasing use of geographic information systems (GIS), and the need for data acquisition

and update for GIS. Feature extraction has been approached in many different ways by photogrammetrists, remote sensing experts and computer vision experts. Some of the methods are quite complex and require the fusion of several data sources or different scale space images (Hinz and Baumgartner, 2003).

Many different methodologies have been used for feature extraction, but none has so far proved to operate in all conditions and for all types of data. Typical methods used include image fusion (Pigeon *et al.*, 2001), fuzzy-based approaches (Agouris *et al.*, 1998), mathematical morphology (Zhang *et al.*, 1999), model based approaches (Buckner, 1998), dynamic programming (Gruen and Li, 1995), and multi-scale grouping and context (Mayer *et al.,* 1997).

The high dimensionality of aerial and satellite imagery presents a challenge for traditional classification methods based on statistical assumptions. Artificial Neural Networks (ANNs) on the other hand may represent a valuable alternative approach for land cover mapping for such highly dimensional imagery. ANNs require no assumption regarding the statistical distribution of the input pattern classes (Hugo *et al.*, 2007) and they have two important properties: the ability to *learn* from input data; and to generalize and predict unseen patterns based on the data source, rather than on any particular *a priori* model (Seto and Liu, 2003). The Self-Organizing Map is one of the most commonly used neural network classifiers. It can be adjusted to adapt to the probability distribution of the inputs (Zheng and Greenleaf, 1996).

In this paper we applied the SOM algorithm for combining multispectral aerial imagery and lidar data so that the individual strengths of each data source can compensate for the weakness of the other. The low contrast, occlusions and shadow effects in the image were compensated by the accurately detected planes in the lidar data. However, edges of features are not located accurately in lidar point clouds because of the lidar's system discrete sampling interval of 0.5m to 1m, (Li and Wu, 2008). An edge will only be located accurately if a lidar pulse is reflected exactly from that edge. Therefore, we have derived 22 attributes from both aerial image and lidar data by a number of algorithms to overcome this problem. To achieve this, we commenced by filtering the lidar point clouds into on-terrain and off-terrain points to generate a Digital Terrain Model (DTM), a Digital Surface Model (DSM) and the difference between the DTM and the DSM, the Normalised Digital Surface Model (nDSM) using a novel filtering technique based on 2D simple polynomials. Then we generated the attributes from the aerial image and the DTM, DSM and nDSM. Finally, the SOM was used to detect buildings from aerial image, lidar data and the generated attributes.

To evaluate the contribution of the lidar data and the generated attributes in the detection process, three separate SOM classification tests were carried out using different input data:

- The aerial image, the lidar data and the derived attributes,
- The aerial image and the lidar data,
- The aerial image only as input data for the SOM.

Finally, the contributions of the individual attributes to the quality of the classification results were evaluated.

RELATED WORK

There have been many research efforts on the application of aerial images and lidar data for building extraction. Chen *et al.* (2004) initially extracted 3D planes from point clouds. Then the initial building edges were detected from raster form lidar data by the Canny Edge Detector. Based on the rough edges, the precise building edges were extracted in image space through a Hough transform.

Rottensteiner *et al.,* (2005) evaluated a method for building detection by the Dempster-Shafer fusion of lidar data and multispectral images. The heuristic model

for the probability mass assignments of the method was validated, and rules for tuning the parameters of this model were discussed. Further, they evaluated the contributions of the individual cues used in the classification process to the quality of the classification results, which showed that, the overall correctness of the results can be improved by fusing lidar data with multispectral images.

Matikainen *et al.,* (2007) used a classification tree approach for building detection. A digital surface model (DSM) derived from last pulse laser scanner data was first segmented into classes *ground* and *building or tree*. Different combinations of 44 input attributes were used. The attributes were derived from the last pulse DSM, first pulse DSM and an aerial colour ortho image. In addition, shape attributes calculated for the segments were used. Compared with a building reference map, a mean accuracy of almost 90 percent was achieved for extracting buildings.

Novacheva (2008) discussed a strategy for outlining building roofs from airborne laser scanning with density of about 5.3 points per m^2 and colour aerial images of 10cm ground resolution. The main purpose of image integration was the refinement of the building outlines. The lidar data was used in guiding the edge detection process from aerial photographs. It was shown that colour images provided important information for the building extraction.

Li and Wu (2008) firstly separated the objects and ground by a filter based on morphological gradients. The non-building objects were removed by mathematical morphology and region growing. Secondly, they smoothed the aerial image by Gaussian convolution, and the gradients of the image were calculated. Finally, they created the edge buffer areas in image space by the edge points of the individual roof patches. The pixels with local maximum gradient in the buffer area were judged as the candidate edges. The final edges were determined through fusing the edges in the image and roof patches by morphological operation. Also, various other classification techniques have been applied for building detection, e.g., unsupervised classification (Haala and Brenner, 1999), rule-based classification (Rottensteiner and Briese, 2002), Bayesian networks (Brunn and Weidner, 1997), and fuzzy logic (Vögtle and Steinle, 2003).

Although ANNs are routinely applied to multispectral imagery classification, the number of studies that have utilized ANNs for highly spectrally dimensional image analysis are limited (Filippi and Jensen, 2006). Some researchers have shown that the unsupervised Self-Organizing Maps (SOM) technique possesses a speedier training rate and higher classification accuracy when classifying remotely sensed data, than the supervised backpropagation neural network (Ji, 2000). Jen-Hon and Din-Chang (2000) applied the SOM classification method for SPOT scene land cover classification using a spatial filter for multispectral feature extraction, following which they employed a supervised non-parametric self-organizing feature map (SOFM) for the classification process. The experiments showed that the method produced 93.2 percent overall classification accuracy compared to 90 percent produced by the Maximum Likelihood classifier. Hugo et al. (2007) assessed the potential of the SOM neural network to extract complex land cover information from medium resolution satellite imagery using MERIS Full Resolution data. The results show that SOM is more effective than k-Nearest Neighbour (kNN) for land cover classification with MERIS data. The results also show that, as expected, larger networks will produce better results.

STUDY AREA AND DATA SOURCES

Study area

The study area is a part of the University of New South Wales campus (Sydney, Australia) covering approximately 500m x 500m. It is a largely urban area that contains residential buildings, large Campus buildings, a network of main roads as well as minor roads, trees, open areas and green areas. The terrain

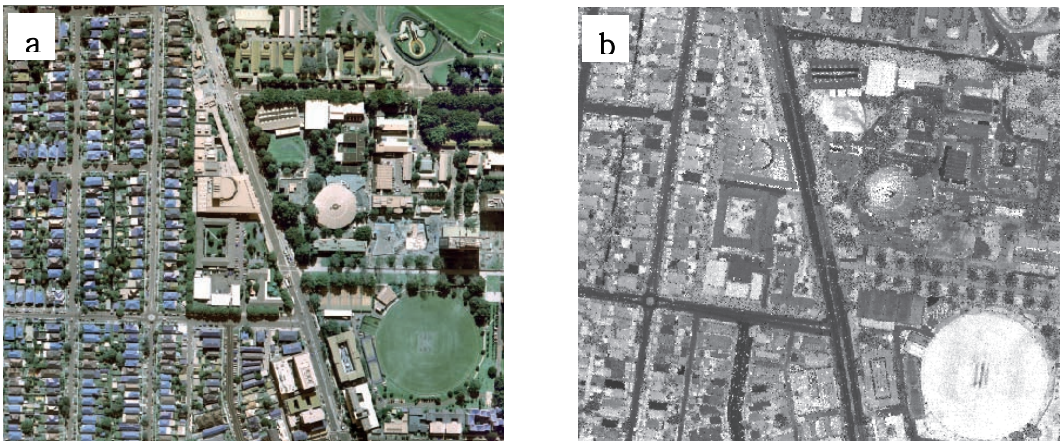

Figure 1. *(a) Multispectral aerial imagery acquired over the study area, (b) Greyscale image created from laser intensity returns*

varies from 45m to 115m above MSL (mean sea level) throughout the area.

Lidar Data

Lidar data were acquired over the study area between April 13th and 14th 2005, and provided in ASCII format (easting, northing, heights and laser intensity returns for first and last pulses). The lidar system used was the Optech ALTM 1225, which operated with a pulse repetition frequency of 25kHz at a wavelength of 1.047µm. The approximate flying height of this sensor above ground was 1100m and the laser swath width was 800m. The vertical accuracy of the lidar posts was 0.15m (1), the internal precision was 0.05m, and the original laser footprint was 22cm in diameter. The original lidar data had point spacing of the order of 16 points per m^2. Figure 1(b) shows the produced image from the lidar intensity values.

Multispectral imagery

The multispectral imagery was captured by film camera over the study area by AAMHatch on June 2005 at 1:6000 scale. The film was scanned in three colour bands (red, green and blue) in TIFF format, with 15µm pixel size (GSD of 0.09m) and radiometric resolution of 16-bit as shown in Figure 1(a).

METHODOLOGY

Building detection of the study area was implemented in several stages as shown in Figure 2.

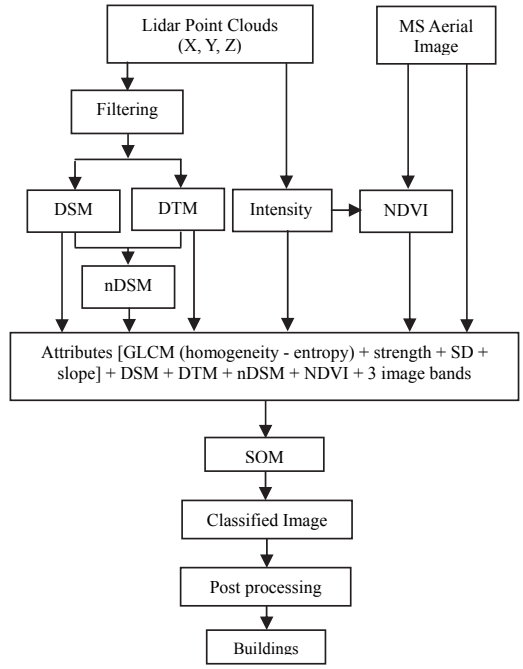

Figure 2. *The work flow for building detection by data fusion based on Self-Organizing Map.*

Filtering of Lidar point clouds

Filtering is the process of separating on-terrain points (DTM) from points falling onto objects such as buildings, cars, trees, and other natural and human made objects. Filtering approaches can be classified in terms of the source of data used into two main approaches: grid based filtering and raw data based filtering. A considerable number of grid based filtering algorithms have been created to filter lidar data. Firstly a DSM is generated from the lidar point cloud, after which filtering is conducted. Okagawa (2001) used a cluster analysis technique to separate ground and non-ground lidar measurements. Kraus and Pfeifer (1998) filtered lidar data in forest areas using an iterative, linear least squares interpolation method. Pfeifer *et al.,* (2001) expanded this method for lidar measurements in urban areas. Passini and Jacobsen (2002) developed a filter based on linear prediction of stationary random function. Haugerud and Harding (2001) employed an algorithm to remove trees in forest areas by comparing local curvatures of point measurements. Vosselman (2000) proposed a filter to remove non-ground measurements by comparing slopes between a lidar point and its neighbours. Zhang *et al.,* (2003) used mathematical morphology to identify non-ground measurements. Mathematical morphology methods suffer various problems such as ineffective removal of various sized non-ground objects due to the requirement of a fixed window size. Zhang and Whitman (2003) proposed an iterative progressive morphological filter to overcome these drawbacks. By gradually increasing the window size and using elevation difference thresholds, the progressive morphological filter removed measurements for different sized non-ground objects while preserving ground data. Elmqvist (2002) classified ground and non-ground measurements based on active contours.

Alternatively, lidar data can be filtered by selecting ground measurements iteratively from the original data set. Axelsson (2000) developed an adaptive Triangulated Irregular Network (TIN) method to find ground points based on selected seed ground measurements. Whitman *et al.,* (2003) used an elevation threshold and an expanding search window to identify and remove non-ground points. Abo Akel *et al.* (2004) used a robust method with orthogonal polynomials and road network for automatic terrain surface extraction from lidar data in urban areas. The DTM cannot always be generated in areas that do not contain road networks by this method. Also, elevated bridges will be classified as part of the DTM.

The approach adopted in this paper is to divide the whole area into small square areas, for which a tilted plane surface is constructed using a first-order polynomial. All points greater than threshold t_1 (=2.5m) above this plane were classified as off-terrain points, while all points below, on or above this plane within the threshold t_2 (=15cm), were classified as on-terrain points. Threshold t_1 was defined by the minimum building height, while t_2 was equal to the lidar system accuracy. The process used for DTM generation is summarized in the following steps:

Dividing the Area into Small Square Patches. The basic assumption of this method is that the height of a ground point is lower than the heights of neighbouring above ground object points and the terrain can be described using a simple tilted plane within small areas. The method started by dividing the data into small 30m x 30m square patches. In principle, the patch should be larger than the largest building within the test area in such a way that no object within the study area can totally cover the patch. Otherwise, points falling over buildings will be classified as on-terrain points (commission error). On the other hand, a patch that is too large will cause inaccurate polynomial fitting and hence some of the ground points will be classified as off-terrain points (omission error). In our test, grids of different sizes and locations were laid over the raster DSM for selecting the smallest sized patches which were not totally covered by buildings. Using this method, and although there are some buildings around 1100m^2, 30m square

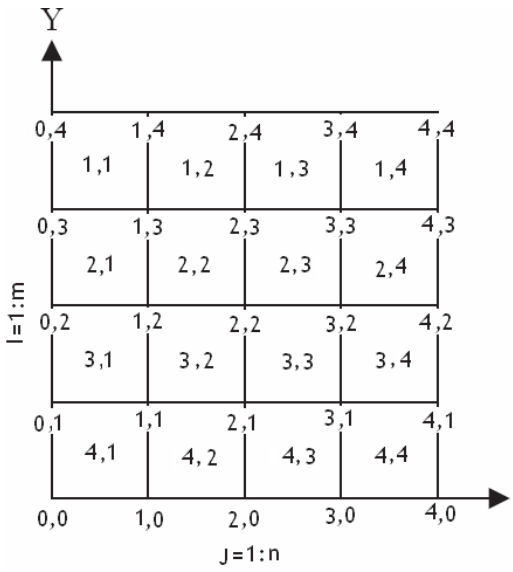

Figure 3. *Dividing the lidar data into small patches.*

patches were considered appropriate. After that, the algorithm constructed a matrix, A (m, n), where m and n are the number of patches in both X and Y directions respectively. Each patch was defined by its row (m) and column (n) numbers within the matrix (Figure 3). Then, the lower left and the upper right coordinates for each patch were determined and stored. Data from both the first and the last pulse echoes were used in order to obtain denser terrain data and hence a more accurate filtering process.

Construction of the Tilted Surface. For each patch and based on its lower left and upper right coordinates tilted plane surfaces were fitted to the terrain points using equation (1):

$$Z = a + b*X + c*Y \quad (1)$$

Where X, Y and Z are the ground coordinates of lidar point clouds. The process of plane surface construction started with the detection of two points, one on each patch border, in the X direction, which represent the minimum elevations on these borders. The two points were then shifted in Y directions by a reasonable value, for example 1000m, while Z values remained constant.

The reason behind the shifting process is to create a new set of two points to construct a comparison plane which includes the four detected points (two old and two new) and represents the general slope of the patch. The main assumption here was that the surface varies slowly from region to region over the patch of interest. The four points were then used to determine the best estimates of the polynomial coefficients of the plane by a least squares solution.

Based on the computed coefficient values of a, b and c, equation (1) was applied for each individual point i with coordinates X_i, Y_i in the lidar point clouds to find the Z value of its corresponding point on the plane. From a comparison of the elevation of each data point with its corresponding elevation on the generated plane surface, all points which were t_1 or more above the plane were classified as off-terrain points (building or trees), while all points on or within t_2 above this surface were classified as on-terrain points.

The plane surface was generated twice in both X and Y directions, so each point was classified twice. The point was assigned to the DTM surface or the DSM, according to whether it was classified as on-ground point or off-ground point respectively from the analysis in both X and Y directions. Figure 4 shows a side view of the fitting process, where the blue colour refers to the original

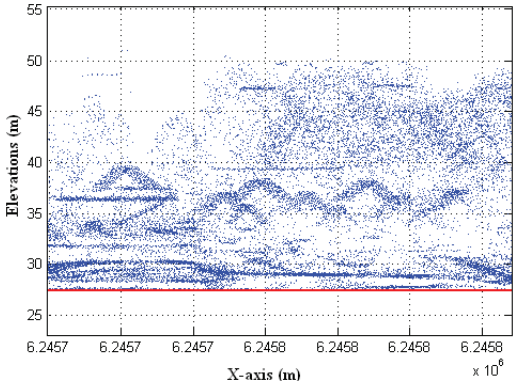

Figure 4. *Side view of the lidar point clouds (blue) and the fitted polynomial (red) in one patch.*

lidar point clouds while the red colour refers to the generated plane surface for the same patch.

DTM, DSM and nDSM generation

The filtered lidar point clouds are still distributed according to the scanning pattern of the lidar sensor and it is not possible to extract attributes or buildings directly from these points. Therefore, the lidar point clouds were converted into an image DTM and DSM by equations 2 and 3:

$$r = \text{Integer}\left(\frac{X - X_{min}}{s}\right) \quad (2)$$

$$c = \text{Integer}\left(\frac{Y_{max} - Y}{s}\right) \quad (3)$$

Where:

X, Y: Ground Coordinates.

r, c: Image Coordinates.

s: Pixel spacing or resolution of the produced image (= $1/\sqrt{n}$ =30cm).

n: the average density of lidar points.

For each pixel (i, j) in the generated images, its digital number (DN) is equal to the Z value of the lidar point that falls into the calculated pixel. For pixels with no corresponding lidar point, an interpolation process from neighbouring pixels was applied to avoid introducing new height values. If more than one point fell within the same pixel, the minimum value was assigned to that pixel. Figures 5(a) and 5(b) show the produced DSM from the original lidar point clouds and the generated DTM after the filtering process. These are grey scale images where tones range from dark for low elevations to bright for high elevations.

The nDSM was generated by subtracting the DTM from the DSM. Non-ground objects, such as buildings and trees are observed on a level plane, so the nDSM represents the absolute heights above the ground and is used in both the segmentation and classification processes. Figure 5c, shows the generated nDSM.

Image to Lidar geographic registration

The aerial photograph and DSM, DTM or nDSM of the same area can only be compared directly if they are in the same coordinate system and have the same pixel size. Otherwise image registration is required to transform one image into the coordinate system of the other image. Unfortunately, it is nearly impossible to identify conjugate points in imagery and lidar data, due to the irregular distribution of the lidar points cloud. To overcome this problem, data from both the first and the last pulse echoes were used again to generate a lidar intensity image in order to obtain a sharper image. Since the lidar intensity image has already been registered to the ground, the aerial photograph (already orthorectified) was registered to the lidar intensity image, based on nine control points visible in both images. A projective transformation was used to find the parameters that best fit the control points (from intensity image) and the corresponding image points

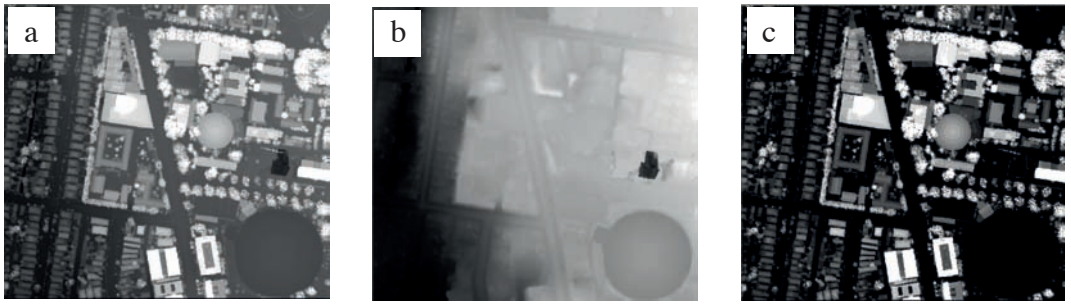

Figure 5. *(a) DSM generated from the original lidar point cloud, (b) DTM generated using the simple tilted plane filtering method, (c) nDSM generated by subtracting the DTM from the DSM.*

(from aerial photograph). The Root Mean Square (RMS) errors from the modelling process were 0.01m and 0.01m in X and Y respectively and the total RMS error was 0.02m, indicating an accurate registration between image and lidar data. The small error values in the process of image to lidar registration demonstrate that most of the geometric distortions had already been removed by the orthorectification process. Following the transformation, the image was resampled to 30cm x 30cm cell size to match the resolution of the lidar data. A bilinear interpolation was used for resampling, which results in a better quality image than nearest neighbourhood resampling and requires less processing than cubic convolution.

Generation of attributes

Features or attributes commonly used for building detection from aerial images and lidar data include height texture (Maas and Vosselman, 1999) or surface roughness (Brunn and Weidner, 1998) of the lidar data, reflectance information from aerial images (Vögtle and Steinle, 2000) or lidar data (Hug, 1997), the difference between first and last pulses of the lidar data (Alharthy and Bethel, 2002), and shape and size of objects (Tóvári and Vögtle, 2004). The attributes calculated for predefined segments or single pixels are presented as input data for a classification method.

In our test, a set of 78 attributes were selected as shown in Table 1. Because of the way the texture equations are constructed, many of them are strongly correlated with one another (Hall-Beyer, 2008). Also Clausi (2002) analysed the correlations among the texture measures to determine the best subset of measures and showed that contrast, correlation and entropy used together outperformed any one of them alone. If only one can be used, he recommended choosing from amongst contrast, dissimilarity or homogeneity. Based on these experiments, only 22 of the 78 possible attributes were available for the classification process as shown in

Attributes	Attribute	R	G	B	I	DSM	NDSM
Spectral	Mean	•	•	•	•	•	•
	St. Deviation	•	•	•	•	•	•
	Strength	•	•	•	•	•	•
GLCM	Contrast	•	•	•	•	•	•
	Dissimilarity	•	•	•	•	•	•
	Homogeneity	•	•	•	•	•	•
	A.S.M	•	•	•	•	•	•
	Entropy	•	•	•	•	•	•
	Mean	•	•	•	•	•	•
	Variance	•	•	•	•	•	•
	Correlation	•	•	•	•	•	•
Height	SD	•	•	•	•	•	•
	Slope	•	•	•	•	•	•

Table 1. *The full set of the possible attributes from aerial images and lidar data; attributes available for the classification are shown by shading.*

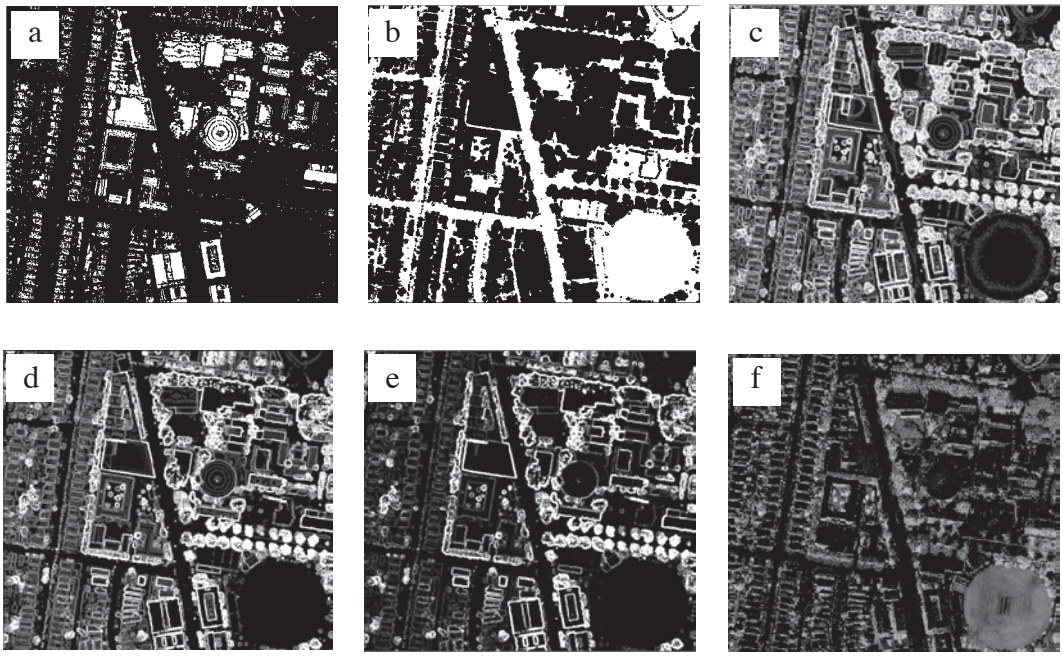

Figure 6. *(a) GLCM-homogeneity of the nDSM, (b) GLCM-entropy of the nDSM, (c) Slope percent of the nDSM, (d) SD of the nDSM, (e) Texture strength of the nDSM, (f) The NDVI.*

the shaded cells of Table 1. The attributes include those derived from the Grey Level Co-occurrence Matrix (GLCM) (Haralick, 1979), Normalized Difference Vegetation Indices (NDVI) (Myneni *et al.*, 1995), standard deviation of elevations (Rutzinger *et al.*, 2008), slope (Escobar *et al.*, 2001) and the polymorphic texture strength (Rottensteiner, 2001) based on the Förstner operator (Förstner and Gülch, 1987).

For GLCM construction and texture calculation, a window of 3x3 pixels was used. The window was placed over the top left 3 x 3 block in the image, then the GLCM was constructed and texture was calculated for the 9 pixel values within that window. The texture value was then written to the location in the image of the central pixel of that window to form a new raster layer. Then the 3 x 3 window was moved one pixel at a time, and the process was repeated until the whole image was scanned by the 3x3 block. For the data lines outside the area covered by the window, no values were derived and therefore data for these locations were interpolated from the computed values. Figure 6 (a - f) shows the generated attributes from the nDSM plus the generated NDVI.

Another way to automatically select the most useful attributes for the classification process is the classification tree analysis method, which automatically constructs a classification tree (Breiman *et al.*, 1984). The main defect of this method is that the user has to control the size (number of splits and hierarchical levels) of the tree by defining a complexity parameter (CP). The CP effectively indicates the *cost* of the process, as it increases for every split, thus influencing the total number of splits included. A high CP could lead to a smaller tree and hence many attributes will be cancelled, even if they are useful for the classification process (Breiman *et al.*, 1984; Therneau and Atkinson, 1997), while a low CP could produce more branches and a more complex tree, which would have a negative effect on stability and separability of the classes (Oczipka, 2007).

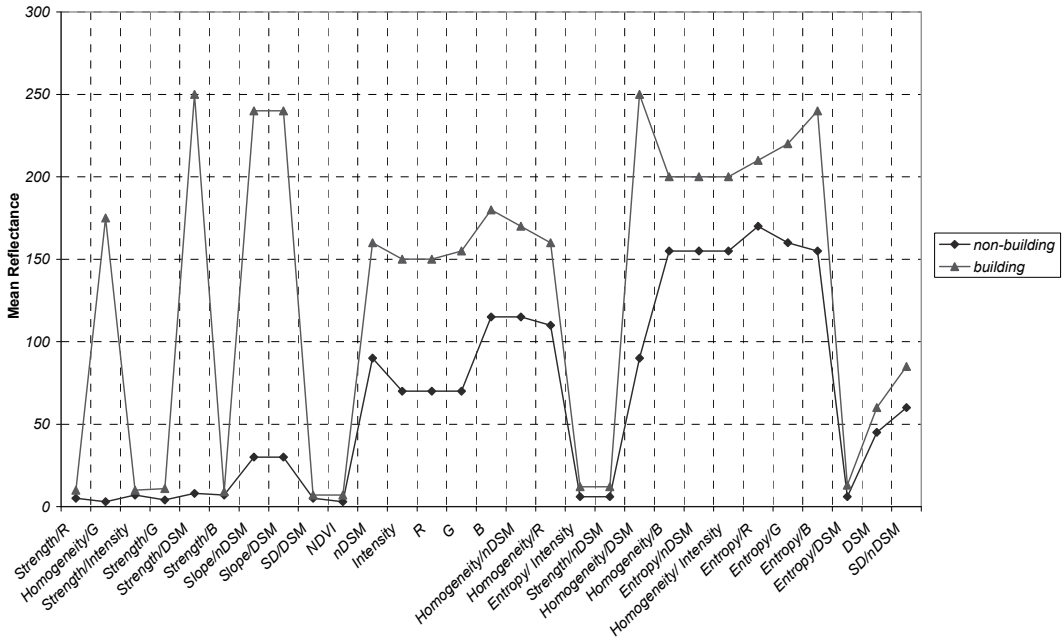

Figure 7. *Mean reflectances for all attributes for buildings and non-buildings.*

Creation of signatures

The overall objective of the creation of signatures is to assemble a set of statistics that describe the spectral response patterns for each land cover type to be classified in the image (Lillesand and Kiefer, 2004). The minimum number of pixels required for a signature is the number of bands plus one (N+1), which is the necessary condition for the covariance matrix to be positive definite (Schowengerdt, 2001).

Twenty samples evenly distributed through the image were selected for buildings and non-buildings (roads and green areas) and the mean reflectance computed (Figure 7) of the created signatures from the 29 bands: 22 generated attributes, 3 image bands (R, G and B), intensity image, DTM, DSM and nDSM. Figure 7 displays a wide range of reflectances (colour image, black and white elevation data, and black and white attributes), but there are only a few bands which appear to have similar reflectances. The clear separation for most reflectances indicates that the selected signatures represent a distinct set of pixels in the bands, which is desirable for any classification model to give a good classification result.

Self-Organizing Map Classifier (SOM)

The SOM undertakes both unsupervised and supervised classification of imagery using Kohonen's Self-Organizing Map (SOM) neural network (IDRISI, 2006). Figure 8 illustrates the basic architecture of an SOM. The input layer represents the input feature vector and thus has neurons for each measurement dimension. In our study, we applied a separate neuron for each band. Therefore, the SOM has 29 input neurons which are: 22 generated attributes, 3 image bands (R, G and B), intensity image, DTM, DSM and nDSM. The output layer of an SOM was organized as a two-dimensional (square) array of neurons. We used a 15 x 15 array of neurons as an output for the SOM (255 neurons). This number was selected because, as recommended by Hugo *et al.* (2007), small networks result in some unrepresented classes in the final labelled network, while large networks lead to an improvement in the overall classification accuracy. Each

output layer neuron is connected to all neurons in the input layer by synaptic weights. To classify the image, four consecutive steps were applied which are: a coarse tuning phase that is effectively a form of unsupervised classification; the assignment of neurons to training classes (a process known as labelling); a fine tuning stage which refines intra-class decision boundaries using a Learning Vector Quantization (LVQ) procedure; and finally display of the feature map to view the classes coded by colour and classification of the unlabelled neurons.

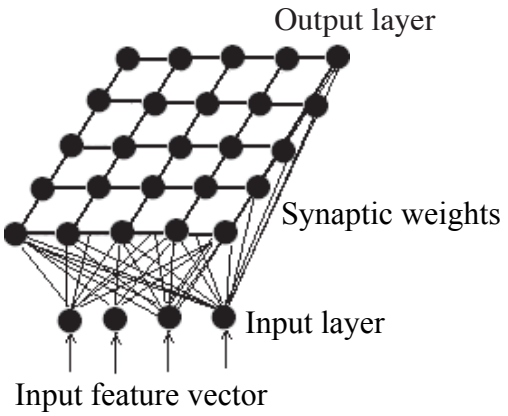

Figure 8. *Example of the architecture of a SOM with a 4 neurons input layer and an equally spaced 5x5 neurons output layer.*

Coarse Tuning. Coarse tuning is an unsupervised classification stage in which competitive learning leads to a fundamental regional organization of neuron weights that represent the underlying clusters and sub-clusters in the input data. Let $X= (x_1, x_2, x_3..., x_n)$ be a vector of reflectances for a single pixel input to the SOM. We took the previously mentioned 29 values (22 generated attributes, 3 image bands, R, G and B, intensity image, DTM, DSM and nDSM) as the vector of reflectances of each pixel. The weight vector could be initialized either to small random values or sampled evenly from the subspace spanned by the two largest principal component eigenvectors (Kohonen, 2001). Initially, synaptic weights between the output and input neurons were randomly assigned (0-1). The Euclidean distances between a reference vector (weight vector) and an input feature vector were then calculated, and the neuron in the output layer with the minimum distance to the input feature vector (known as the winner) was then determined as in equation (4):

$$\text{Winner} = \arg\min\left(\sum_{i=1}^{n}(x_i^t - w_{ji}^t)^2\right) \quad (4)$$

Where x_i^t is the input to neuron i at iteration t, and is the synaptic weight from input neuron i to output neuron j at iteration t. The weight of winner and its neighbours within a radius γ were then altered (while those outside were left unaltered) according to a learning rate α^t as shown in equations (5,6):

$$w_{ji}^{t+1} = w_{ji}^t + \alpha^t(x_i^t - w_{ji}^t), \quad \forall d_{\text{winner } j} \in \gamma^t \quad (5)$$

$$w_{ji}^{t+1} = w_{ji}^t, \quad \forall d_{\text{winner } j} \notin \gamma^t \quad (6)$$

where α^t is the learning rate at iteration t and $d_{\text{winner } j}$ is the distance between the winner and other neurons in the output layer. α^t was calculated from equation (7)

$$\alpha^t = \alpha_{\max}\left(\frac{\alpha_{\min}}{\alpha_{\max}}\right)^{\frac{1}{t}} \quad (7)$$

An identical time-decay function was also used to reduce the radius from an initial size that can include all of the neurons in a single neighbourhood to a final one which includes only the winner.

The above process is controlled by three primary parameter groups: the number of neurons in the output layer, the number of input samples and the maximum and minimum learning rates (Tso and Mather, 2001; Villmann *et al.*, 2003; Ji, 2000). The coarse tuning parameter values for our test were selected according to the suggestions proposed by Vesanto *et al.* (2000) to improve the classification accuracy without giving any computer memory constraints. Table 2 shows the coarse tuning parameters used for the test.

Input Neurons	22
Out put neurons	225 (15*15)
Initial γ	25
Min. α	0.5
Max. α	1

Table 2. *Coarse tuning parameters used for the test.*

Labelling. Before an SOM can perform a classification, the class to which each output neuron belongs must be determined by inputting training data into the coarsely tuned network. The labelling stage is performed fully automatically by the SOM without any user interaction and based on the training data (CAD signatures file) which were previously created. So the SOM method appears to be a highly automatic approach for extracting buildings compared to other methods, especially when its operations for building detection are based on a graphical user interface (GUI). The training site class that was assigned most frequently to a neuron then became its label (Tso and Mather, 2001). The labelling process resulted in four groups of neurons, namely: one for buildings, two for non-buildings (roads and green areas), and one for unclassified pixels.

Fine Tuning. The goal of fine tuning is to refine the decision boundaries between classes based on the training site data. In the procedure outlined by Kohonen (1990, 2001), fine tuning was achieved through the use of Learning Vector Quantization (LVQ). For LVQ, known patterns (from training sites) were fed again to the SOM. The neuron with the most similar weights to the input vector of reflectances was determined and its weights were adjusted as in equations (8), (9) and (10):

If x is correctly classified

$$w_c^{t+1} = w_c^t + \delta^t (x_i - w_c^t) \quad (8)$$

If x is incorrectly classified

$$w_c^{t+1} = w_c^t - \delta^t (x_i - w_c^t) \quad (9)$$

If x is not classified

$$w_c^{t+1} = w_i^t, \text{ if } i \neq c \quad (10)$$

Where w_c is the weight vector of the winner, and δ^t is a gain term that will decrease with time.

The primary control for the fine tuning is thus the gain factor δ_t setting (Tso and Mather, 2001; Ji, 2000). Again, the fine tuning parameter values for the test were also selected according to the suggestions proposed by Vesanto *et al.* (2000). Table 3 shows the fine tuning parameters used for the test.

Min. δ^t	0.0001	Fine tuning rule	LVQ
Max. δ^t	0.0005	Fine tuning epoch	50

Table 3. *Fine tuning parameters used for the test.*

Display of the Feature Map and Classification of the Unlabeled Neurons. After the coarse tuning, labelling and fine tuning stages, the feature map was generated, which is colour coded by information classes (Tso and Mather, 2001). Thus all neurons associated with a particular class share the same colour as shown in Figure 9(a). Classes with substantial variability, such as trees, had many neurons associated with them, while those with little variability, such as roads, were associated with only a few neurons. The feature map gave an indication that the performance of the SOM was good. From the feature map we notice that both buildings and roads occupied close neighbour neurons since they have similar spectral reflectance on the three bands of the aerial image. On the other hand, green areas occupied neurons which were spatially separated from those of buildings or roads since they are different from roads and buildings in terms of spectral reflectance in our case.

From the produced feature map, black colour refers to unlabelled neurons by the SOM. All pixels within the unlabelled neurons were assigned a class that has a minimum D_i as in equation (11). Figure 9(b) shows the SOM classified image.

$$D_{\min i} = \min_j \{\|x - w\|\} ---- \forall \text{Label}(j) = i \quad (11)$$

where x is the input vector and w the

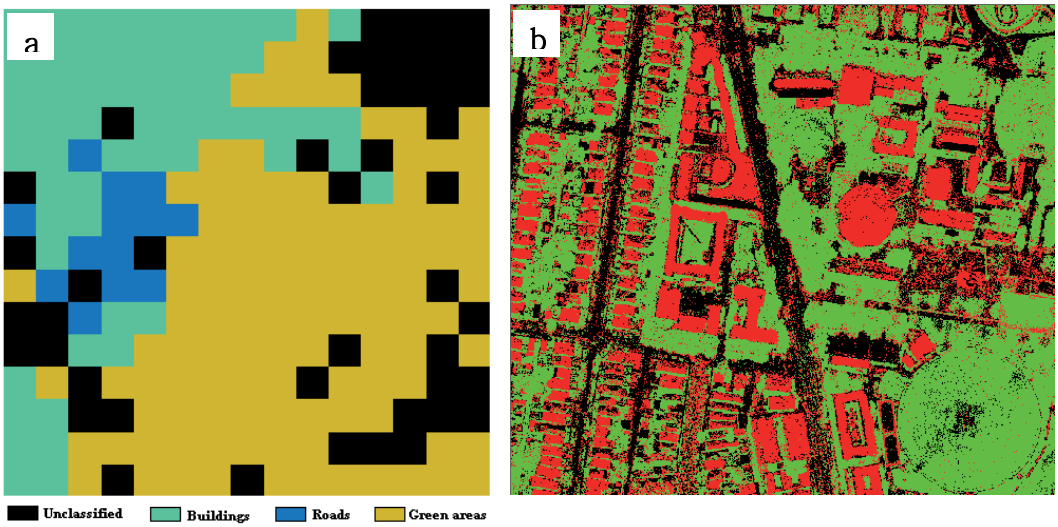

Figure 9. *(a) The produced feature map using training data, (b) The SOM classification results.*

organized reference vector (weight vector) (Li and Eastman, 2001).

Post classification smoothing

Since this study was only interested in building detection, buildings were separated by converting the digital number of buildings to one, while the digital numbers of non-building pixels (roads and trees) were converted to zeros. Then, the smaller raster homogeneous building regions were merged into larger neighbouring homogeneous regions or deleted according to an arbitrary 1m distance and 50m^2 area threshold, respectively. The area threshold represents the expected minimum building area, while the distance threshold was set to 1m to fill in any gaps produced through the classification process. Regions were retained if they were larger than the given area threshold and/or were adjacent to a larger homogeneous region by a distance less than 1m.

Finally, building borders were cleaned by removing structures that were smaller than 8 pixels and that were connected to the image border. There was a compromise between

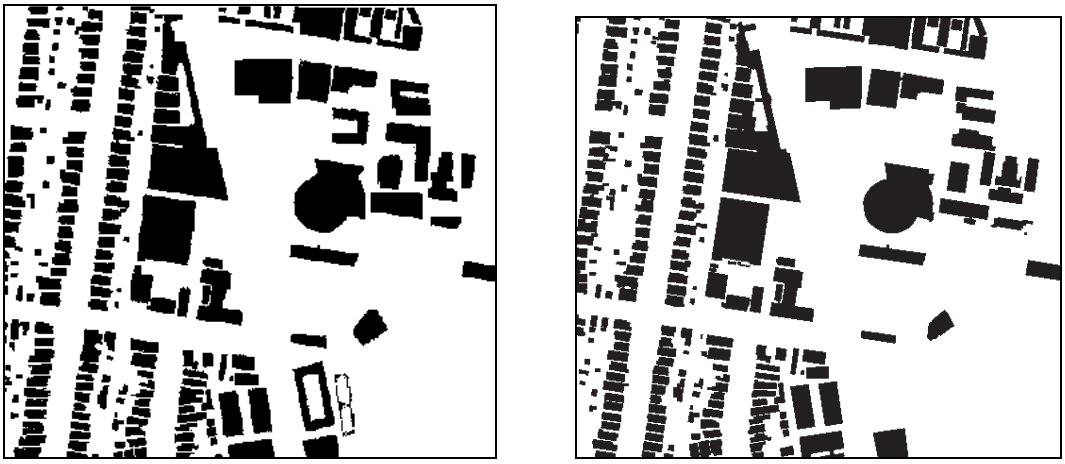

Figure 10. *(a) The final detected buildings, (b) The manually digitized buildings.*

cleaning thresholds less than 8 pixels, which may leave the original buildings uncleaned, and thresholds greater than 8 pixels which may remove parts of the buildings. The result was a black and white image that represents the detected buildings without noisy features and also without holes, as shown in Figure 10(a).

In order to evaluate the classification accuracy, buildings were manually digitized in the multispectral image as shown in Figure 10(b) to serve as the reference data. Adjacent buildings that were joined but obviously separated were digitized as individual buildings. Otherwise, they were merged as one polygon.

RESULTS AND ANALYSIS

Evaluation of the classification results

In order to evaluate the classification accuracy, manually digitized buildings were compared against the reference data. Then, the completeness and correctness of the results, described in (Rottensteiner *et al.*, 2005), were computed as shown in equations (12) and (13). Completeness is the percentage of the actual buildings that are detected by an algorithm, and the correctness is the percentage of the buildings detected by an algorithm that correspond to real buildings.

$$\text{Completeness} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{12}$$

$$\text{Correctness} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{13}$$

TP denotes the number true positives, i.e., the number of entities found to be available in both reference and experiment data sets. FN is the number of false negatives, i.e., the number of entities in the reference data set that were not detected automatically, and FP is the number of false positives, i.e., the number of entities that were detected, but do not correspond to an entity in the reference data set.

Figure 11 shows the completeness and correctness for buildings against the building size. Buildings around 100m^2 were detected with both completeness and correctness around 83 percent and 80 percent, respectively. Completeness and correctness were improved for increasing building size. Buildings around 1100m^2 were detected with both completeness and correctness around 96 percent and 99 percent, respectively. Similar accuracies have been reported in Rottensteiner *et al.* (2005). They evaluated a method for building detection by the Dempster-Shafer fusion of airborne laser scanner (ALS) data and multispectral images using different data sets. By Dempster-Shafer fusion, 95 percent of all buildings larger than 70m^2 were detected and 95 percent of all detected buildings larger than 70m^2 were correct.

The present study also compared the areas of the detected buildings against the corresponding ones from the reference data. The minimum difference in areas was -2.5m^2, the max difference was 3.3m^2, the mean difference was -0.3m^2 and the root mean square error of differences was 2.1m^2. The main reasons for the negative mean of differences in areas are the lost pixels during the cleaning step and the poor detection of edges by the lidar may have contributed to this error.

The total number of detected buildings was also compared against the digitized number of buildings and the results

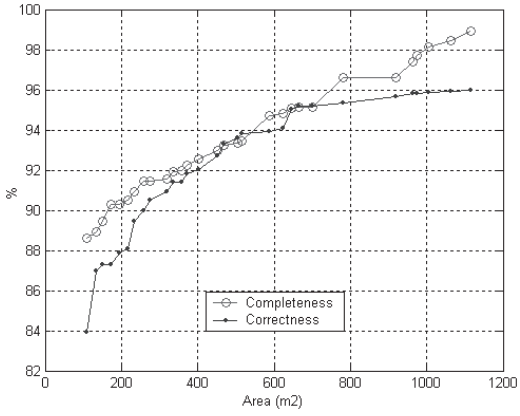

Figure 11. *Completeness and correctness against building areas.*

showed that 95.5 percent of all buildings in the reference data were detected. Only 11 buildings were not detected, since they may have been merged with larger buildings or deleted during the post classification smoothing process.

Since classification has resulted in thematic output, an error matrix was applied to express the classification accuracy in this test. The error matrix in Table 4 presents the results of comparisons of the numbers of pixels in the classified image (rows) to the number of pixels in the manually digitized image (columns). For example, row 1/ Table 4 shows that 782 092 pixels that are identified in the reference data (digitized image) as buildings were correctly classified as buildings. On the other hand, 17676 pixels were misclassified as non-buildings.

	Buildings	Non-buildings	Total
Buildings	782092	17676	799768
Non-buildings	4415	852426	856841
Total	786507	870102	1656609

Table 4. *Error matrix of the SOM classified image (rows) against digitized image (columns) using aerial image, lidar data and attributes.*

From Tables 4, the overall accuracy for the classification result was 97.8 percent and calculated by equation (14):

$$\mathrm{OCA} = \frac{\mathrm{NCP}}{\mathrm{NRP}} \tag{14}$$

Where OCA is the overall classification accuracy; NCP is the total number of correctly classified pixels (along the diagonal of the error matrix) and NRP is the total number of reference pixels.

To evaluate the contribution of lidar data and attributes to the classification accuracy, the SOM was performed twice more, firstly, by using the aerial image alone, and secondly, using the aerial image plus the lidar data. The results of these two tests were compared to the first test based on all available data and the error matrices were tabulated for each case (Tables 5 and 6). See also Figure 12 (a, b, c, d) for a comparison of results of the three tests.

	Buildings	Non-buildings	Total
Buildings	7714	654045	661759
Non-buildings	237560	757290	994850
Total	245274	1411335	1656609

Table 5. *Error matrix of the SOM classified image (rows) against digitized image (columns) using aerial image only.*

	Buildings	Non-buildings	Total
Buildings	682030	99	682129
Non-buildings	107804	866676	974480
Total	789834	866775	1656609

Table 6. *Error matrix of the SOM classified image (rows) against digitized image (columns) using aerial image and lidar data.*

The overall classification accuracies for the three different cases were: 49.7 percent (aerial image only), 87.3 percent (aerial image and lidar data) and 97.8 percent (aerial image, lidar data and attributes). For the first case, many buildings were classified as roads because they have the same spectral reflectance and the overall classification accuracy was low (Figure 12(b)). For the second case, using the lidar data along with the aerial image increased the classification accuracy to about 87.3 percent due to its ability to detect planes accurately but still some errors occurred due to the poor horizontal accuracy of edge detection in the lidar data (Li and Wu, 2008) (Figure 12(c)). As demonstrated above the use of aerial image, lidar data and extracted attributes increased the classification accuracy to about 97.8 percent since the attributes compensated for the weakness of lidar for edge detection (Figure 12(d)).

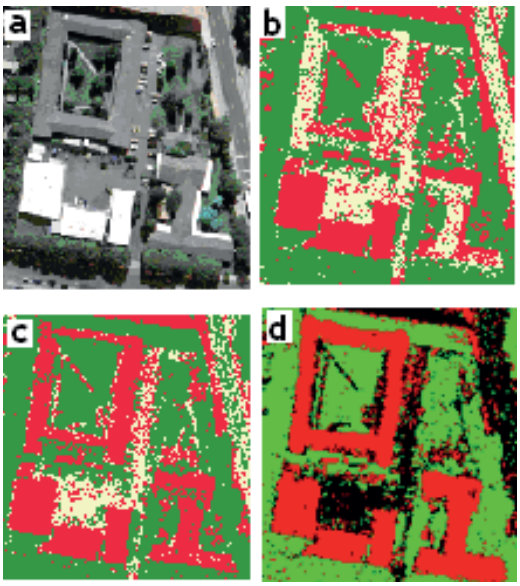

Figure 12. *(a) The multispectral aerial image, (b) The SOM classified image using aerial image only, (c) The SOM classified image using aerial image and lidar data, (d) The SOM classified image using aerial image, lidar data and attributes.*

Also the amount of improvement in the detection process which was reached by the noise removal in the classification results was evaluated. From Table 7, post-classification, smoothing and noise removal improved the detection process by 5.2 percent; the completeness by 3.9 percent, but decreased correctness by 1.8 percent, which is insignificant.

Before			After		
No.	Cp	Cr	No.	Cp	Cr
224	89.8	91.3	237	93.7	89.5

Table 7. *No.: Number of buildings. Cp, Cr: average completeness and correctness for building pixels.*

Contributions of the individual attributes

Furthermore, the contributions of the individual attributes to the quality of the classification results was evaluated. The red, green and blue bands of the aerial image were considered as the primary

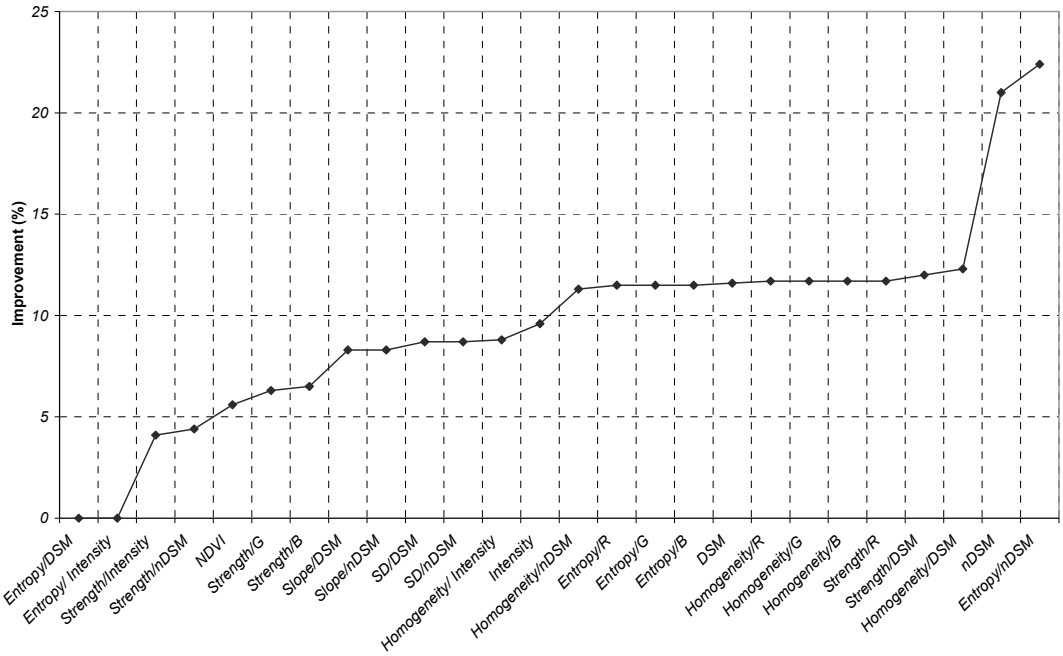

Figure 13. *Contributions of the individual attributes to the quality of the classification results.*

data source and were available in each test. The classification results for the images and each attribute were compared against those derived from the image alone. Figure 13 shows the improvement in the average completeness and correctness by each individual attribute. Polymorphic strength derived from the lidar intensity and polymorphic strength derived from the nDSM performed worse than the others and only improved the average accuracy by 4.1 and 4.4 percent, respectively. The nDSM and entropy from the nDSM performed best, improving the average accuracy by 21 and 22.4 percent, respectively. Entropy from the DSM and entropy from the intensity image made no contribution to the quality of the results. The remaining attributes improved the results by values from 5.6 up to 12.3 percent with some of these attributes performing equally as shown in Figure 13.

CONCLUSION

A method for building detection based on Self-Organizing Map fusion of lidar, multispectral aerial images and 22 auxiliary attributes was presented. The attributes that were generated from the lidar data and multispectral images include: texture strength, Grey Level Co-occurrence Matrix (GLCM) homogeneity and entropy, Normalized Difference Vegetation Indices (NDVI), standard deviation of elevations and slope. The approach significantly improves the accuracy of detection of buildings over approaches when only images and/or lidar data are used. The results show that using lidar data in the SOM improves the accuracy by 38 percent compared with using aerial photography alone, while using the generated attributes as well improve the result by a further 10 percent. An investigation into the contributions of the individual attributes showed that nDSM and entropy from the nDSM generally proved most effective and improved the average accuracy by 21 and 22.4 percent respectively, while entropy from the DSM and entropy from the intensity image made no contribution to the quality of the results. Also the results showed that the attributes of high spectral separations performed better than those with low spectral separations, and hence spectral reflectance can contribute to the selection of the most useful attributes for the classification process. In the future, the authors intend to construct a hybrid classifier based on multiple classifiers operating simultaneously to achieve a more effective and robust decision making process.

REFERENCES

Abo Akel, N., Zilberstein, O. and Doytsher, Y. (2004) A Robust Method Used with Orthogonal Polynomials and Road Network for Automatic Terrain Surface Extraction from LIDAR Data in Urban Areas. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Science*, Vol. 35, ISPRS 274-279.

Agouris, P., Gyftakis, S. and Stefanidis, A. (1998) Using a Fuzzy Supervisor for Object Extraction Within an Integrated Geospatial Environment. *In: International Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part III/1, pp. 191-195.

Alharthy, A. and Bethel, J. (2002) Heuristic Filtering and 3d Feature Extraction from LIDAR Data. *ISPRS the International Achieves of the Photogrammetry, Remote Sensing and Spatial Information Sciences* Vol. XXXIV.

Axelsson, P. (2000) DEM Generation from Laser Scanner Data Using Adaptive TIN Models. *International Archives of Photogrammetry and Remote Sensing*, XXXIII, Part B3:85–92.

Breiman, L.; Friedman, J., H.; Olshen, R., A. and Stone, C., J. (1984) Classification and Regression Trees. *Chapman and Hall*.

Brunn, A. and Weidner, U. (1997) Extracting Buildings from Digital Surface Models. In: *IAPRS XXXII* / 3-4W2, pp. 27-34.

Brunn, A. and Weidner, U. (1998) Hierarchical Bayesian Nets for Building Extraction Using Dense Digital Surface Models. *ISPRS Journal of Photogrammetry & Remote Sensing*, 53(5), pp. 296–307.

Buckner, J. (1998) Model Based Road Extraction for the Registration and Interpretation Of Remote Sensing Data. *In: International Archives of Photogrammetry and Remote Sensing*, Stuttgart, Germany, Vol. 32, Part 4/1, pp. 85-90.

Chen, L.C., Teo, T.A., Shao, Y.C., Lai, Y.C. and Rau, J.Y. (2004) Fusion of LIDAR Data and Optical Imagery for Building Modelling. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 35(B4), pp. 732-737.

Clausi, D. A. (2002) An Analysis of Co-Occurrence Texture Statistics as a Function of Grey-Level Quantization. *Canadian Journal of Remote Sensing*, vol. 28 no. 1 pp. 45-62.

Escobar, F., Hunter, G., Bishop, I. and Zerger, A. (2001) Digital Elevation Models Specific Theory. Available at: http://www.geom.unimelb.edu.au/gisweb/DEMModule/DEM_Theory. [Accessed on February 2009].

Elmqvist, M. (2002) Ground Surface Estimation from Airborne Laser Scanner Data using Active Shape Models. *ISPRS Commission III Symposium*, Photogrammetric and Computer Vision, Graz, Austria, pp. 114–109.

Filippi, A. M. and Jensen, J. R. (2006) Fuzzy Learning Vector Quantization for Hyperspectral Coastal Vegetation classification. *Remote Sensing of Environment*, 100: 512-530.

Förstner, W. and Gülch, E. (1987) A Fast Operator for Detection and Precise Location of Distinct Points, Corners and Centres of Circular Features. *In ISPRS Intercommission Workshop*, pages 281–305, Interlaken.

Gruen, A. and Li, H. (1995) Semi-Automatic Road Extraction by Dynamic Programming. *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(4), pp. 11-20.

Haala, N. and Brenner, C. (1999) Extraction of Buildings and Trees in Urban Environments. *ISPRS J. Ph & RS* 54(2-3), pp. 130-137.

Hall-Beyer, M. (2008) The GLCM Tutorial Home Page. Available at: http://www.fp.ucalgary.ca/mhallbey/tutorial.htm. [Accessed on February 2009].

Haralick, R.M. (1979) Statistical and structural approaches to texture. *Proceedings of the IEEE*, 67, pp. 786-804.

Haugerud, R.A., and D.J. Harding (2001) Some Algorithms for Virtual Deforestation (VDF) of LIDAR Topographic Survey Data. *International Archives of Photogrammetry and Remote Sensing*, XXXIV, Part 3/W4:211–218.

Hinz, S. and Baumgartner, A. (2003) Automatic Extraction of Urban Road Networks from Multi-View Aerial Imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 58/1-2, pp. 83–98.

Hug, C. (1997) Extracting Artificial Surface Objects from Airborne Laser Scanner Data. In: Gruen, A., Baltsavias, E. P., Henricsson, O. (Eds.), Automatic Extraction of Man-Made Objects from Aerial and Space Images (II), *Birkhäuser Verlag*, Basel, pp. 203-212.

Hugo, C., Capao, L., Fernando, B. and Mario, C. (2007) Meris Based Land Cover Classification with Self-Organizing Maps: preliminary results, *EARSeL SIG Remote Sensing of Land Use & Land Cover*.

IDRISI, (2006) Idrisi Andes trainning reference manual. *Clark University*, Worcester, USA.

Jen-Hon, L. and Din-Chang, T. (2000) Self-Organizing Feature Map for multi-spectral spot land cover classification. *GIS development.net*, AARS, ACRS 2000.

Ji, C. Y. (2000) Land-Use Classification of Remotely Sensed Data Using Kohonen Selforganizing Feature Map Neural Networks. *Photogrammetric Engineering and Remote Sensing*, 66: 1451-1460.

Kohonen, T. (1990) The Self-Organizing Map. *Proceedings of the IEEE*, 78: 1464-80.

Kohonen, T. (2001) Self-Organizing Maps. *Third Edition*. New York, Springer.

Kraus K. and Pfeifer N. (1998) Determination of Terrain Models in Wooded Areas with Airborne Laser Scanner Data. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 53 pp193-203.

Lillesand, T., and Kiefer, R. (2004) Remote Sensing and Image Interpretation. *Fourth Edition*, John Willey & Sons, Inc., New York.

LI Y. and WU H. (2008) Adaptive Building Edge Detection by Combining LiDAR Data and Aeria Images. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XXXVII. Part B1. Beijing 2008

Li, Z. and Eastman, J. (2001) The Nature of and Classification of Unlabelled Neurons in the Use Of Kohonen's Self-Organizing Map (SOM) For Supervised Classification. *Transactions in GIS*, upcoming.

Maas, H. and Vosselman, G. (1999) Two Algorithms for Extracting Building Model from Raw Laser Altimetry Data. *ISPRS Journal of photogrammetry and remote sensing*, 54 (2-3):153-163, 1999.

Matikainen, L., Kaartinen, H. and Hyyppä, J. (2007) Classification Tree Based Building Detection from Laser Scanner and Aerial Image Data. *ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007*, Espoo, September 12-14, 2007, Finland.

Mayer, H., Laptev, I., and Baumgartner, A. (1997) Automatic Extraction Based on Multi-Scale Modeling, Context And Snakes. *In: International Archives of Photogrammetry and Remote Sensing*, Vol. 32, Part 3-2W3, pp. 106-113.

Myneni, R. B., F. G. Hall, P.J. Sellers and A.L. Marshak (1995) The Interpretation of Spectral Vegetation Indexes. *IEEE Transactions on Geoscience and Remote Sensing*, 33, 481-486.

Novacheva, A. (2008) Building Roof Reconstruction from Lidar Data and Aerial Images through Plane Extraction and Colour Edge Detection. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. XXXVII. Part B6b, Beijing 2008.

Oczipka, M. (2007) Objektbasierte Klassifizierung hochauflösender Daten in urbanen Räumen unter besonderer Berücksichtigung von Oberflächenmodellen. *Doctoral thesis*, Freie Universität, Berlin, Germany.

Okagawa, M. (2001) Algorithm of Multiple Filter to Extract DSM from LIDAR Data. *2001 ESRI International User Conference*, Kraus, K., and N. Pfeifer, 1998. Determination of terrain models in wood areas with airborne laser scanner data, *ISPRS Journal of Photogrammetry & Remote Sensing*, 53(4):193–203.

Passini, R. and B.D. Jacobsen (2002) Filtering of Digital Elevation Models. Proceedings of the ASPRS 2002 Annual Convention, 19–26 April, Washington DC (*American Society for Photogrammetry and Remote Sensing*, Bethesda, Maryland), 9 p., unpaginated CD-ROM.

Pfeifer, N., P. Stadler and C. Briese (2001) Derivation of Digital Terrain Models in the SCOP++ Environment. *Proceedings of OEEPE Workshop on Airborne Laserscanning and Interferometric SAR for Detailed Digital Elevation Models,* 01–03 March 2001, Stockholm, Sweden, 13 p.

Pigeon, L., Solaiman, B. and Toutin, T. (2001) Linear Planimetric Feature Domains Modelling for Multi-Sensors Fusion in Remote Sensing. *In: Proceedings of SPIE AeroSense - International Symposium on Aerospace/Defence Sensing*, Simulation, and Controls, Orlando, Vol. 4051, 8 p.

Rottensteiner, F. (2001) Semi-Automatic Extraction of Buildings Based on Hybrid Adjustment Using 3d Surface Models and Management of Building Data in a TIS. *PhD thesis*, Institute of Photogrammetry and Remote Sensing, Vienna University of Technology.

Rottensteiner, F. and Briese, C. (2002) A New Method for Building Extraction in Urban Areas from High-Resolution LIDAR Data. In: *IAPRSIS, XXXIV / 3A*, Graz, Austria, pp. 295 – 301.

Rottensteiner, F., Summer, G., Trinder J., Clode S. and Kubik, K. (2005) Evaluation of a Method for Fusing Lidar Data and Multi-Spectral Images for Building Detection. *In: Stilla U, Rottensteiner F, Hinz S (Eds) CMRT05. IAPRS,* Vol. XXXVI, Part 3/W24 - Vienna, Austria, August 29-30, 2005.

Rutzinger, M. Hofle, B., Hollaus, M. and Pfeifer N. (2008) Object-Based Point

Cloud Analysis of Full-Waveform Airborne Laser Scanning Data for Urban Vegetation Classification. *Article, Sensors 2008*, 8, 4505-4528; DOI: 10.3390/s8084505, ISSN 1424-8220.

Schowengerdt, R. (2001) Models and Methods for Image Processing. *Second Edition*, ACADEMIC PRESS, San Diego.

Seto, K. and Liu, W. (2003) Comparing ARTMAP Neural Network with the Maximum-Likelihood Classifier for Detecting Urban Change. *Photogrammetric Engineering & Remote Sensing*, 69: 981-990.

Therneau, T. and Atkinson, E. (1997) An Introduction to Recursive Partitioning using the Rpart Routines. Department of Health Science Research, Mayo Clinic, Rochester, MN. *Technical report 61 edition*.

Tóvári, D. and Vögtle, T. (2004) Classification Methods for 3D Objects in Laser Scanning Data. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Istanbul, Turkey, Vol. XXXV, Part B3, pp. 408-413.

Tso, B. and Mather, P.M. (2001) Classification Methods for Remotely Sensed Data. *(New York: Taylor and Francis)*, 114-121.

Vesanto, J., Himberg, J., Alhoniemi E. and Parhankangas J. (2000) SOM Toolbox for Matlab 5. *Technical Report A57*, Helsinki University of Technology, Neural Networks Research Centre, Espoo, Finland.

Villmann, T., Merenyi, E., and Hammer, B. (2003) Neural Maps in Remote Sensing Analysis. *Neural Networks,* 16: 389-403.

Vögtle, T. and Steinle, E. (2003) On the Quality of Object Classification and Automated Building Modeling Based on Laser scanning Data. In: *IAPRSIS XXXIV / 3W13*, pp. 149-155.

Vögtle, T. and Steinle, E. (2000) 3D Modelling of Buildings Using Laser Scanning and Spectral Information. In: *International Archives of Photogrammetry and Remote Sensing,* Amsterdam, the Netherlands, Vol. XXXIII, Part B3, pp. 927-934.

Vosselman, G. (2000) Slope Based Filtering of Laser Altimetry Data. *International Archives of Photogrammetry and Remote Sensing*, XXXIII, Part B4:958–964.

Whitman, D., K. Zhang, S.P. Leatherman, and W. Robertson (2003) Airborne Laser Topographic Mapping: Application to Hurricane Storm Surge Hazards. Earth Sciences in the Cities (G. Heiken, R. Fakundiny, and J. Sutter, editors), *American Geophysical Union*, Washington DC, pp. 363–376.

Zhang, C., Murai, S. and Baltsavias, E. (1999) Road Network Detection by Mathematical Morphology. *In: Proceedings of ISPRS Workshop on 3D Geospatial Data Production*: Meeting Application Requirements, Paris, France, Pages: 185-200.

Zhang, K. and Whitman, D. (2003) Comparison of Three Algorithms for Filtering Airborne Lidar Data. *Photogrammetric Engineering & Remote Sensing*, Vol. 71, No. 3, March 2005, pp. 313–324.

Zhang, K.Q., S.C. Chen, D. Whitman, M.L. Shyu, J.H. Yan, and C.C. Zhang (2003) A Progressive Morphological Filter for Removing Non-ground Measurements From Airborne LIDAR Data. *IEEE Transactions on Geoscience and Remote Sensing*, 41 (4): 872–882.

Zheng, Yi. and Greenleaf, F. (1996) The Effect of Concave and Convex Weight Adjustments On Self-Organizing Maps. *IEEE Trans.* Neural Networks, vol. 7, no. 1, pp. 87-96.

ACKNOWLEDGEMENTS

The Lidar data and multispectral aerial imagery was provided by AAMHatch (www.aamhatch.com.au).