Geocarto International

ISSN: 1010-6049 (Print) 1752-0762 (Online) Journal homepage: http://www.tandfonline.com/loi/tgei20

Evaluation of LiDAR and image segmentation based classification techniques for automatic building footprint extraction for a segment of Atlantic County, New Jersey

R. Prerna & Chander Kumar Singh

To cite this article: R. Prerna & Chander Kumar Singh (2015): Evaluation of LiDAR and image segmentation based classification techniques for automatic building footprint extraction for a segment of Atlantic County, New Jersey, Geocarto International, DOI: 10.1080/10106049.2015.1076060

To link to this article: http://dx.doi.org/10.1080/10106049.2015.1076060

Accepted online: 29 Jul 2015.Published online: 02 Sep 2015.

Geocarto International, 2015
http://dx.doi.org/10.1080/10106049.2015.1076060

Taylor & Francis
Taylor & Francis Group

Evaluation of LiDAR and image segmentation based classification techniques for automatic building footprint extraction for a segment of Atlantic County, New Jersey

R. Prerna[†] and Chander Kumar Singh[*]

Department of Natural Resource, TERI University, New Delhi, India

(*Received 30 January 2015; accepted 13 July 2015*)

Extracting high-quality building footprints is a basic requirement in multiple sectors of town planning, disaster management, 3D visualization, etc. In the current study, we compare three different techniques for acquiring building footprints using (i) LiDAR, (ii) object-oriented classification (OOC) applied on high-resolution aerial photographs and (iii) digital surface models generated from interpolated LiDAR point cloud data. The three outputs were compared with a digitized sample of building polygons quantitatively by computing the errors of commission and omission, and qualitatively using statistical operations. These findings showed that building footprints derived from OOC gave highest regression and correlation values with least commission error. The R^2 and R values (0.86 and 0.92, respectively) imply that the footprint areas derived by OOC matched more closely with the actual area of buildings, while a low commission error of 24.7% represented a higher number of footprints as correctly classified.

Keywords: LiDAR; building footprints; object-oriented classification; image segmentation; DSM

1. Introduction

The need to extract good-quality building footprints has grown due to their widespread use in various domains of spatial planning. Since planning is not possible without data inventories, automatic techniques become important in order to save time and resources. By definition, a building footprint is the outer boundary of a building encompassing the exterior walls without including courtyards/gardens. It basically helps in creating an exterior reconstruction of the walls and the roof especially useful for 3D visualization. Building footprint extraction is also an important segment for the creation of virtual navigation models as they not only localize buildings within a point cloud, but also reduce the search space for vertical walls and step edges, which are generally believed to be one of the most difficult components of roof modelling (Suveg & Vosselman 2002; Uden & Zipf 2013).

Lee et al. (2008) have characterized building detection process into three groups. First method uses 2D or 3D information from photogrammetric imagery (Hinz & Baumgartner 2003) which depends upon resolution of images; however, the complexity of such methods increases with 3D information derivation, occlusion and shadows and

*Corresponding author. Email: chander.singh@teriuniversity.ac.in
[†]National Centre for Antarctic and Ocean Research, Headland Sada, Vasco-da-Gama, Goa, India.

nearby objects or trees of similar heights (Lee et al. 2008; Vu et al. 2009; Maan et al., 2014). Second method is using LiDAR data to detect building footprints; however, using LiDAR also raises concern on poor accuracy on building edges and geometrically precise boundary using only LiDAR point cloud (Rottensteiner 2003; Yong & Huayi 2008). The third category of methods uses both LIDAR data and photogrammetric imagery. More specifically, intensity and height information in LIDAR data can be used with texture and region boundary information in aerial imagery to improve accuracy (Lee et al. 2008). Over the past few decades, the multifaceted advantages of LiDAR data have been realized and put to proper use in various sectors like urban planning (Priestnall et al. 2000), hydraulic and hydrologic modelling (French 2003), ecosystem studies (Lefsky et al. 2002; Hyde et al. 2005; Gwenzi & Lefsky 2014; Tang & Dai 2014), coastal management (Pe'eri & Long 2011; Collin et al. 2012; Yousef et al. 2013) and many more. Within the domain of urban planning, LiDAR technology has been effectively employed for mapping building footprints, identifying building heights, determination of lines of sight, site planning and design (Priestnall et al. 2000; Alexander et al. 2009; Vu et al. 2009; Patino & Duque 2013; Xiao et al. 2012). The applications of LiDAR have become furthermore seamless due to its powerful integration with other data-sets like orthophotos, multispectral, hyperspectral and panchromatic imagery (Vu et al. 2009; Kabolizade et al. 2010; Awrangjeb et al. 2013; Tang & Dai 2014). The use of LiDAR data with high density of points (at least 4–6 points per square meter) jointly with 2D vector digital maps of building footprints, allows the construction of a 2.5D urban surface model (Carneiro et al. 2009). Alongside, image segmentation – a process of dividing or disintegrating the image into objects or regions with homogenous spatial and spectral characteristics for the purpose of feature extraction (Haralick & Shapiro 1985; Pesaresi & Benediktsson 2001; Ouma et al. 2008; Sridharan & Qiu 2013; Bhandari et al. 2014) – is also being valued as an alternative method. Object-oriented classification (OOC) is driven by an understanding of the image object rather than its pixels (Wegner, 1987; Wang et al. 2004; Bhaskaran et al. 2010), where each image object is a homogenous group of pixels/regions that have similar spectral and/or spatial characteristics. Liu et al. (2005) performed OOC followed by a fuzzy rule-based logical classification of the segments to distinguish between building and non-building objects, giving the opportunity to include size, shape and texture parameters along with the usual spatial and spectral homogeneity, making the process more valuable. Apart from LiDAR data classification and OOC-derived footprint extraction techniques, another approach discussed in this paper is that of deriving footprints from digital surface models (DSMs). Brédif et al. (2013) have exhibited a valid approach for deriving footprints using DSMs, which can be directly obtained from the 3D point cloud provided by LiDAR sensors, or with surface reconstruction techniques using multiple images (Pierrot-Deseilligny & Paparoditis 2006; Hirschmuller 2008). Automated extraction is also difficult due to obstructions caused by surrounding objects or shadows; heterogeneous structures in terms of shape, size, pattern, geometry, etc. causing difficulty in precise edge detection (Sahar et al. 2010). One more common difficulty that has plagued such algorithms is that there is generally a lot of noise in the data points that lie around the edges of buildings which can be caused by overhanging trees and sensor noise.

With the above background, the current study compares three different approaches of building footprint extraction, which are (1) LiDAR point cloud-based classification, (2) OOC-based classification applied on aerial photographs and (3) LiDAR point cloud-derived DSM classification. Further, the authors have aimed to validate the

methods by comparing the resultant footprints – based on quantitative and qualitative statistical analysis discussed in the subsequent sections.

1.1. Study area and data specifications

The study area consists of four cities of the Atlantic County, New Jersey state of the USA. These four cities are Atlantic (44.125 km^2), Ventnor city (9 km^2), Margate city and Longport (4 km^2 each) totalling up to approximately 60 km^2 in area under study (Figure 1). The latitudinal extent of the study area is 39°18′ N to 39°23′ N while its longitudinal extent is 74°32′ W to 74°24′ W located at the south eastern coast of Atlantic county in the state of New Jersey. Out of the four states, the northern two i.e. Atlantic and Ventnor have a greater density of high-rise buildings (15 stories and above) with complex architectures while Margate and Longport comprise more of mid rise (8–15 stories) and low-rise buildings (8 stories and below) due to a larger area covered by residential units. This prevalent heterogenity of high-/mid-/low-rise building structures was a major reason for choosing this area for study. Another reason for selection was the availability of both – LiDAR data-sets and high-resolution aerial photographs (used for OOC) for this area. Details and specifications of the data-sets that have been used in the study are mentioned in Table 1.

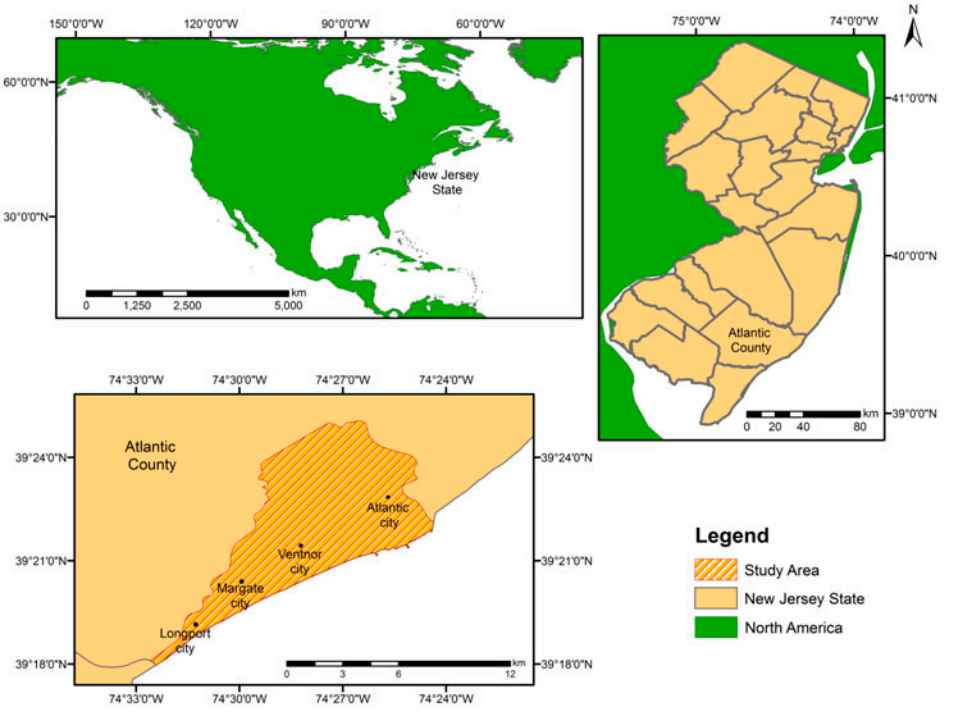


Figure 1. Study area.

Table 1. Data specifications of the study.

Data format	Resolution	DOA	Source	Projection
LiDAR points (LAS files)	Point spacing <1 m	April 2010	US geological survey	NAD83/2007 New Jersey state plane coordinate system in US survey feet
Aerial photographs (MrSID files)	1 × 1 ft	2007	New Jersey geographic information network	NAD83/2007 New Jersey state plane coordinate system in US survey feet

2. Methodology

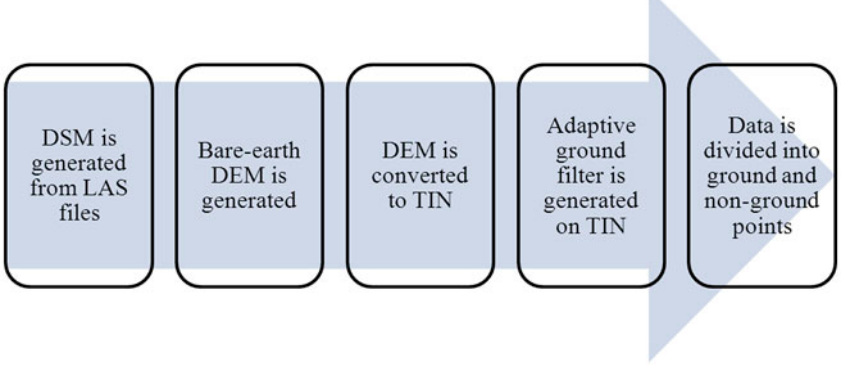
2.1. Methods for footprint extraction

2.1.1. Capturing building footprints using LiDAR point clouds

Extraction of building footprints for the four cities of the study area was performed on 22 tiles of LiDAR point clouds. The LiDAR data-sets were by default 'unclassified', i.e. no distinction existed between the points returned from the ground and those from a building/structure. These point clouds had to be first filtered and cleansed to classify the data appropriately. For this, QCoherent's LP360 (Advanced Version) was employed, which is a widely used software package meant for managing huge amounts of LiDAR data, also having a multitude of point cloud tasks for the purpose of filtering and classifying LiDAR Archive Standard (LAS) files. These tasks are specially designed to address several stages of processing LAS files; for instance, tasks for filtering data, masking data, classification, exporting subsets from the data, etc. It has an advantage over other LiDAR processing software which has its direct linkage with ArcGIS environment, allowing interoperability between LAS format files and other GIS data. In the current study, several *objective-based tasks* have been applied, where each task is dedicated for processing a particular objective e.g. differentiating the non-ground points from the unclassified point cloud; filtering of data as per height, etc. An in-depth description of the tasks used in the study is given below:

- Step 1: Filtering of ground points:

The 'unclassified' point cloud (Figure 2(a)) is segregated into ground and non-ground points using *adaptive TIN ground filtering point cloud task* (Figure 2(b)). The performance of this task is briefly summarized below:



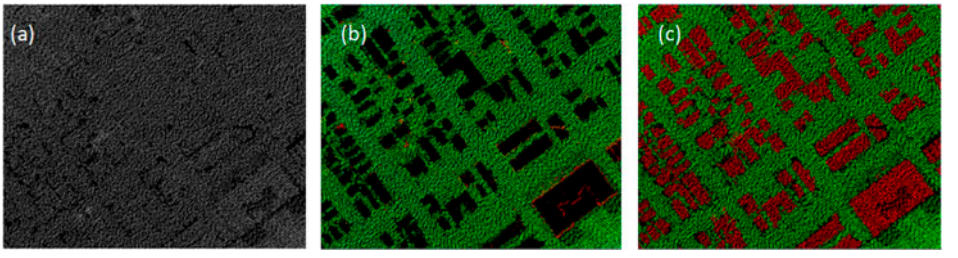


Figure 2. (a) Point clouds as unclassified points; (b) ground and non-ground points; (c) building points and ground points.

In this operation, all points with elevation greater than the ground elevation (user defined) are classified as non-ground points.

- Step 2: Classification on non-ground points:

All points classified as non-ground points from a LiDAR data-set may not necessarily be buildings. Hence, the components of shape, area, height and pattern come into play. The *planar filter* and *Height* or *Basic* filters are used in conjunction to deduce the building points from the point clouds. The *planar filter* point cloud task is a planar-based point classifier, which searches for points that fit user-defined planes and classifies those points to a user-specified class. The software supports classification and export of LiDAR points (into other formats) belonging to different user-defined classes like buildings, vegetation, ground, water, etc. In this case, the user-specified class was buildings. Along with the execution of the planar filter, the data also need to be classified using the *Height* or *Basic* filters merely on the basis of elevation. The Height filter divides the points in the building class as per a user-specified maximum and minimum height constraint. For this, a calculation of the distance between (a) ground and the lowest planar surface; and (b) ground and the highest planar surface must be done for getting minimum and maximum height values, respectively. The essential parameters involved in these tasks are the following:

(1) Units: These are the units of measurement for all parameters fed into the system for filtering, which can be changed as per requirement of the user, not necessarily matching with the units of the data. In the current study, units of measurement were feet.
(2) Point spacing: This distance is used to construct a moving window where the moving window will have a length and width of twice the point spacing value. A window sized with the dimensions as previously described, is fit around each point and an equation of a plane will be fit to these points. If the area of the planar surfaces is less than the dimensions of the moving window, then they are not classified. Several trials are required before arriving at the optimum value, which can be achieved by visual inspection across the data-set.
(3) Maximum area: The maximum object area parameter is used as a limit to the surface growth. An increasing maximum area would cause more points to be classified as the limit of region growing is expanded. A value 5000 ft was

considered to be appropriate for this function as the largest buildings in the study area had dimensions of ≤5000 ft.

(4) Z threshold: The Z threshold parameter represents the maximum orthogonal distance to the plane under consideration required for a point to be classified. It depends on the accuracy of the point cloud as it further affects the vertical accuracy at which the points would be classified.
(5) Plane fit: This is used to determine the 'goodness of fit' test of the particular plane i.e. it shows the variance of the residuals of the fitted plane.
(6) Minimum and maximum slope: The object slope is in degrees. Planes with slopes less than or greater than the minimum and maximum slopes, respectively, are not classified.
(7) Minimum and maximum height: Points with a height above the ground surface and below the minimum/maximum object height value will not be evaluated and is applicable only when the 'Use Height Filter' option is checked.
(8) Clean up percent: This value affects the degree of missed points to be cleaned up thereby removing the noisy points.

However, it must be noted that this task does not create vectors (footprints in this case) and only 're-classifies' the data into a defined class. Figure 2(a) and (b) shows the unclassified point cloud and ground (green) and non-ground points (black), respectively, while Figure 2(c) represents the result of the *Height filters* with building points shown in red and ground points shown in green.

- Step 3: Extraction of building footprints:

Once the non-ground points are categorized into a 'Building' class, the *point group tracing* task was executed to trace boundaries along the class. Figure 3 illustrates a preview window showing the pattern of footprints identified using the building extractor parameters. This task completes the vectorization process of the classified points derived by Step 2. The traced outlines are then exported as shapefiles along with attributes.

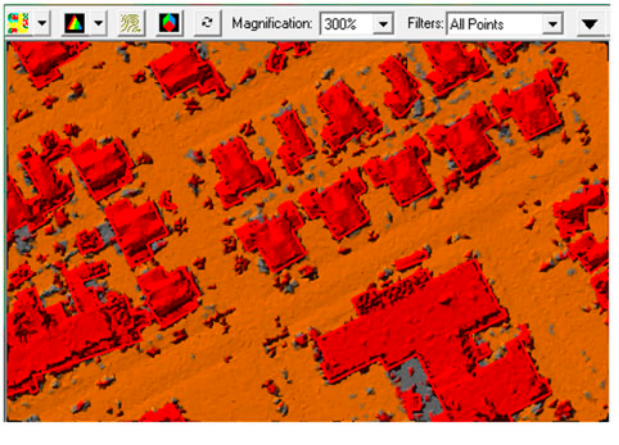


Figure 3. Example of preview window showing the polygons of building footprints growing along the structures (using the building extractor parameters explained in Step 3 under Section 2.1.1).

Figure 4. A sample data-set from Longport Borough showing buildings before and after applying *squaring* function.

Results so obtained had to be smoothed because of the occurrence of noisy points near the building edges. This was performed using a *squaring* function that drastically helped in producing smoother and neater polygons. The variation in building polygons before squaring (in red) and those obtained after squaring (in yellow) is clearly seen in Figure 4.

For successfully employing the various tasks, the analyst needs to be acquainted with the study area and the data-set to be able to provide correct parametric values. Units of measurement, point spacing, trace window size, Z threshold, min–max slope, height, etc. are some parameters for which accurate values must be given. These have been described below:

(1) Grow window: The Grow Window parameter is a moving window size used to group points based on the Boundary Trace Class. Its application is same as the point spacing parameter in the previous task. Grouping of adjacent objects is possible if the grow window is large. A window of 04 ft was chosen as optimum value as it seemed to capture even the minute building edges/corners.
(2) Trace window: The Trace Window parameter is the moving window size used to trace a grouped set of points. This parameter is also dependent on the ground sample distance. This adds smoothness in the data-set as pointed building edges get smoothened and squared. It is mostly maintained as twice the grow window size, hence 08 ft was considered appropriate.
(3) Minimum area: It allows the user to filter small objects from the extracted output. A value of 100 ft implied that a structure with dimensions <100 was not included in the classification.
(4) Minimum points: This option allows the user to place the points below a certain threshold into another user-defined class which was kept unchecked because only building class was needed for this study.

(5) Perform squaring: This action is done to produce an approximation of the roof outlines of buildings by squaring the traced outlines. An angle of 20° was maintained in this study.

2.1.2. *Capturing building footprints using image segmentation and OOC*

The second aim of the study was to extract building footprints from OOC for which image segmentation was a pre-processing step. For the purpose of OOC and segmentation, it was necessary to identify features that were homogeneous not only in terms of spectral response but also as per shape, size and texture. It was observed that where building structures displayed similar geometry – shape, size and rooftop tone; objects were more easily identified as opposed to regions where heterogeneity existed in the building shapes.

The first part of segmenting the image (Figure 5(A)) was done using the *multi-resolution segmentation* option in *eCognition software* which executes segmentation based on input parameters. This software offers the capability of applying different scale parameters and colour and/or shape combinations decided by the user in order to build a hierarchical network of image objects whose results are then used to outline different materials within the image (Ngcofe & Minnaar 2012). The scale and colour parameters are the most important inputs as they effect segmentation the most. A scale parameter is used to control the average image object size (Baatz & Schape 2000). The multi-resolution segmentation option is a bottom-up approach which moves from the smallest entity of the image (a pixel) to a larger homogenous object (group of pixels meeting the criteria). This way, it starts to include other adjacent pixels growing into larger 'objects' making it a local optimization procedure (Karakis et al. 2006). In eCognition, the scale parameter range is from 5 to 250, which results in variation in the size of the segments created. A higher scale parameter defined by the user would result in larger sized objects because the scale value is basically a measure of the number of pixels to be coalesced as *one* object. Hence, the concept of image segmentation based on *scale* is more of a region merging technique, where at each step – more and more image objects are merged into one larger object (Baatz & Schape 2000).

Another important parameter to be considered for image segmentation is colour/spectra. The degree of spectral heterogeneity of image objects can be derived from the variance *or* standard deviation (s.d.) of spectral mean values (Baatz & Schape 2000). In other words, the colour parameter determines the weighted use of spectral value and/or shape and texture value of an object being studied (Ngcofe & Minnaar 2012). Based on variance/s.d. values, images get broken down into *spectrally homogenous objects* which are identifiable due to spectral contrast existing between pixels. This heterogeneity in the image has to be understood before deciding the values for the colour parameter. In eCognition, values range from 0.1 to 0.9 with high weighted values giving more emphasis on spectra and less on shape, while low values emphasize more on shape and less on spectra.

In this study, the scale parameter mostly ranged from 45 to 60 while colour ranged from 0.45 to 0.65. These two input variables were decided after repeated attempts based on the desirable degree of segmentation. Secondly, in post image segmentation, classification is performed using the *optimal box classifier* wherein the user has to specify sample objects belonging to each class. These sample objects have to be carefully picked from amongst all the segments and is a manual operation. Here only two classes – positive (red) and negative (green) can be created for which appropriate samples need

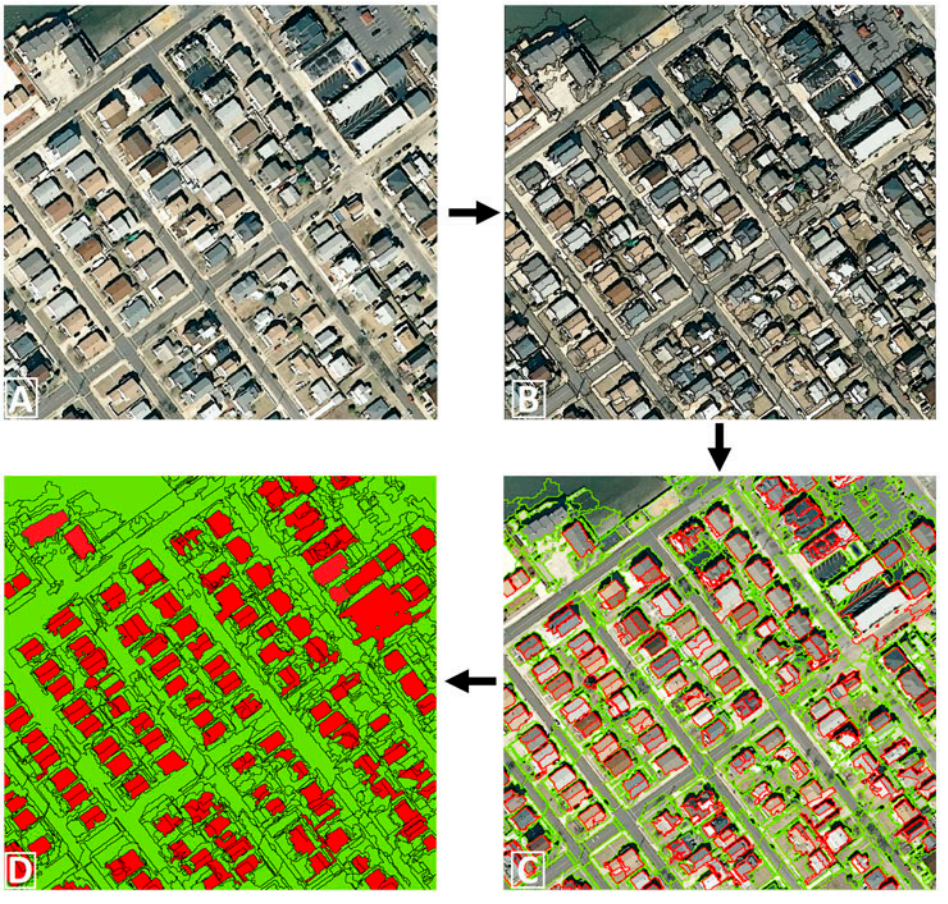


Figure 5. (A) Starting from top left, the aerial photograph from a section of the study area is shown; (B) the image is segmented into several homogenous image objects after feeding scale and colour parameters; (C) The segments are divided into two classes – building (red) and non-building (green) based upon user defined samples; (D) exported polygons showing building and non-building class.

to be collected. Using these samples, the image is then classified into two classes – building (red) and non-building (green), which can also be merged so that all objects belonging to one class could appear as a singular entity.

Finally, the classified vectors were exported as shapefiles (.shp) into the ArcGIS environment. For ease of handling data, attributes were added to the polygons, in this case, a simple binary code of 0 and 1 – 0 being non-building and 1 for building. Running a query on such attributes help in categorizing data so that they can be grouped to retain only building polygons while the remaining can be discarded. The entire flow of the image segmentation operation can be understood in Figure 5(A)–(D).

2.1.3. Capturing building footprint using DSM derived from LiDAR point cloud

The third approach for footprint capturing was from DSMs generated from LiDAR point clouds. This technique has been suggested as an alternative to footprint extraction

directly from LiDAR data. The LiDAR data points inclusive of elevation/height values were exported to ESRI point shapefiles using ArcGIS 10.1. These were interpolated to a raster surface representing elevation values (using IDW interpolation technique). It is important to preserve the resolution of the source data and hence the DSMs generated must conform to that; for instance – if the data points are spaced at every 1 m, the corresponding DSM must also be in a 1 × 1 m grid. Figure 6(A) and (B) shows DSM derived for one portion of the study area with its respective aerial photograph. This DSM was thereafter classified in ERDAS Imagine 10 using unsupervised tool for classification and several classes were obtained. The classes were reduced to two – building and non-building class. Since only the *buildings* class was important for the study, the associated polygons were exported as shapefiles. The DSMs derived were also utilized further to generate triangulated irregular network (TIN) grids for the study area. Raster to TIN conversion available in ArcGIS 10.1 was applied. Capturing footprints using LiDAR data is often performed using high-end devoted software designed especially for the purpose. However, this technique would allow analysts to use good-quality data without following the preordained methodologies of LiDAR data processing.

Validation of the results attained by the three techniques is explained in the section below.

2.2. *Comparison of automated building footprints with manually digitized footprints*

One of the simplest, logical and quantitative measures for assessing the accuracy of building footprints is to compare the area obtained by an automated technique with the area of manually digitized building polygons. Such an approach has been previously followed where simplification techniques of building footprints were suggested and the

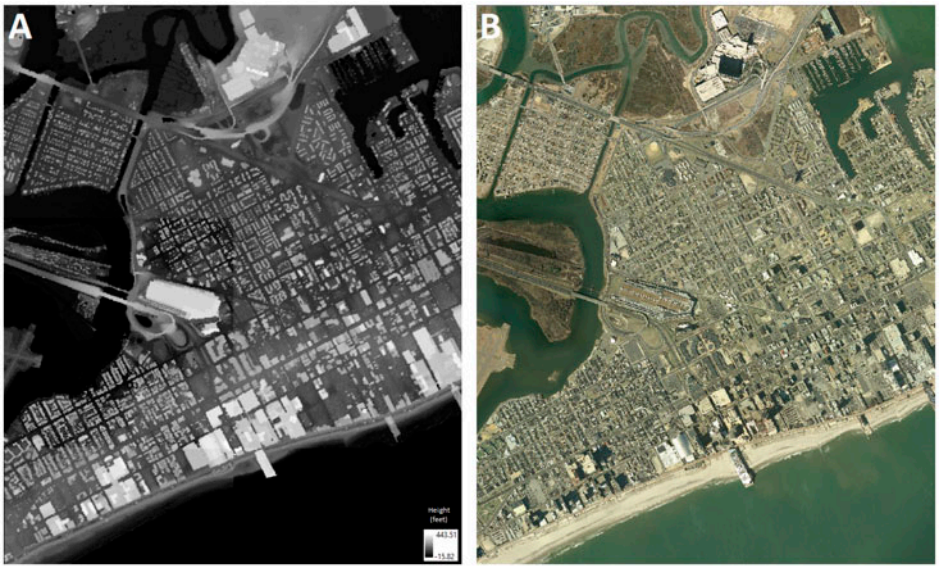


Figure 6. (A) DSM generated from LiDAR point cloud showing height (feet); (B) aerial photograph of the same area.

accuracy was assessed using area comparisons with the original polygons (Chen et al. 2004). For this, a small segment of buildings (roughly 200) were manually digitized using the high-resolution aerial photographs in ArcMap 10.1 for creating a data-set of actual building footprints with area calculation.

Another estimation of accuracy was carried out using commission and omission percentages for the three data-sets. In simple terms, percentage of commission denotes the 'additional' area included in the footprint results which does not actually exist on ground, while percentage of omission depicts the 'left out' area that is there on ground but not included in the derived footprints. So in a way, commission error tells how the model has overestimated footprint area while omission error is a measure of the model's underestimation. This method is an area-based metric method adopted by Rüther et al. (2002). The commission and omission errors, attained by comparing final building footprints with corresponding known building footprints (Zhang et al. 2006) has hence been established as a useful quantitative method of footprint validation. To facilitate understanding, it can be said that if an actual ground footprint measures 100 ft^2, and if the model predicts its value as 110 ft^2, a commission error of 10% would prevail. On the other hand, if the model calculates the area to be 90 ft^2, 10% of omission error would be committed.

In this study, commission/omission errors were calculated for the 200+ polygons manually digitized with their corresponding derived footprints – from all three methods. An average of the values helped in assessing the level of accuracy achieved by the different techniques. Qualitative evaluation of the results was also done by regression and correlation analysis to check the degree of induced errors between the actual and resultant data. The methodology followed in the study – footprint extraction from LiDAR point data directly, image segmentation and DSMs has been shown as a flow chart in Figure 7.

3. Results

3.1. *Building footprints derived from LiDAR data*

LAS files downloaded from the open source domain were inclusive of only elevation values devoid of any categorization. Classification of the points into different features like water, ground, vegetation, building, etc. needed the point cloud to be filtered beforehand because removal of ground points and other noise points is essential to carefully extract the returns coming from buildings i.e. the LiDAR signals received from buildings only. The parameters were judged after previewing the results multiple times and finally those values were fed as inputs which exhibited finer results. The building footprints acquired using the methodology discussed in Section 2.1.1, were further classified into three groups (Figure 8) based on their heights – high-rise building with heights greater than 180 ft (15 stories and above); mid-rise building with heights from 96 to 180 ft (9–15 stories) and low-rise buildings that were lesser than 96 ft (8 stories and below).

For 3D visualization, LAS files were draped upon high-resolution aerial photographs for Atlantic City shown in Figure 9. Also, footprint shapefiles exported from LiDAR data-set were viewed and extruded in ArcScene showing clearly defined building edges (Figure 10).

3.2. *Building footprints extracted from OOC and image segmentation*

Segmentation of the images on the basis of *scale* and *colour* was performed. Numerous objects were created appearing as broken segments (Figure 5(B)), which had to be subsequently grouped together and merged to form one class. The use of *optimal box*

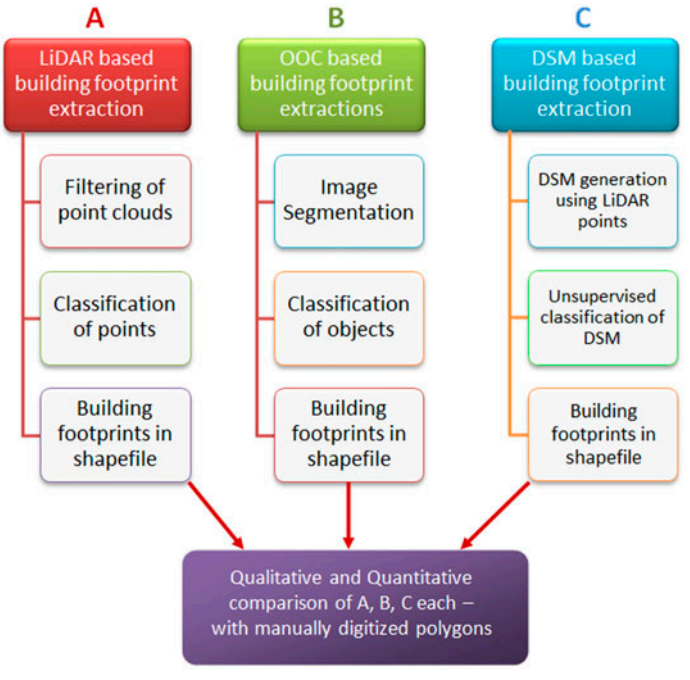


Figure 7. Methodology for the study.

Figure 8. Map showing building heights in different categories derived from LiDAR point clouds.

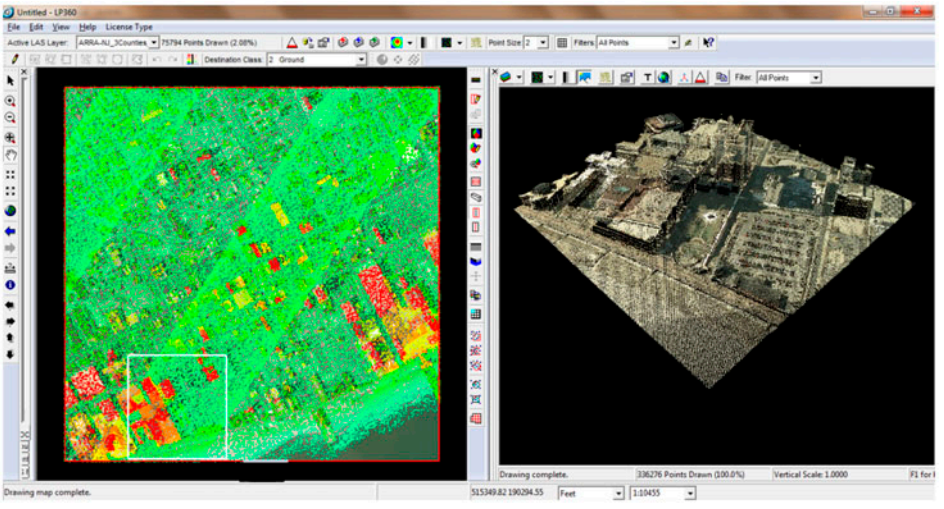


Figure 9. A preview window showing aerial photograph draped on LiDAR data.
Note: Area in white box on the left image is represented on the right.

Figure 10. 3D depiction of the building structures derived from LiDAR data in ArcScene 10.1. Buildings are shown as high-rise (red – 180 ft and above), mid-rise (yellow – 180–96 ft) and low-rise (green – 96 ft and below).

classifier was done wherein samples of positive (buildings) and negative class (non-buildings) had to be fed, as explained in Section 2.1.2. The result of this classification can be clearly seen in Figure 5(C). Here the positive class i.e. buildings are depicted in red and negative class i.e. non-buildings are shown in green. This is an automated process and works well if an accurate sample is fed. Since the focus was only on extracting building objects, all segments of buildings class were grouped together. These segments were eventually exported and area calculations were done for validating results.

3.3. Building footprints extracted from DSMs

Rasterization of points is a conversion of vector point shapefiles into raster surfaces on the basis of a given input value. In this case, elevation values i.e. Z value of the LAS points were given as an input and raster surfaces were generated with the cell resolution of 5 ft (1.5 m). The LAS points were roughly spaced at 1–10 ft (0.3–3 m) (horizontal spacing between points) and thereby 5 ft pixel resolution was chosen as an average value of the input file spacing.

Since these raster surfaces depict values in terms of elevation not only for the terrain of the area but also for the overlying structures, the output so attained was a DSM and not a DTM/DEM, which are surfaces representing only the earth surface elevation. The classes belonging to buildings were subsequently converted to polygons and calculation of area was done. The accuracy of the DSM derived was inspected on a 3D platform using ArcScene 10.1. In Figure 11, an area of Atlantic City has been represented as a TIN. This approach was quite successful in capturing the outer structure of the buildings and exhibits a good methodology to derive TIN surfaces from LiDAR point clouds.

4. Discussion

4.1. Comparison of automated building footprints and manually digitized footprints

4.1.1. Commission and omission error assessment

Every footprint in the sample data-set of 200+ polygons was compared with the corresponding calculated building area by measuring the percentage of commission and omission (Table 2). From the results, it was concluded that OOC derived footprints exhibited the least commission error, while the results of DSM showed lowest omission error. The commission errors can be attributed to the inclusion of certain other objects under buildings' class, for instance large size vehicles and sometimes closely spaced buildings get identified as one singular block, thereby increasing the total area. Omission errors may have been induced due to shadows created by high-rise buildings, omitting the adjacent buildings and trees covering certain parts of the roofs in the case of single/double-storied buildings. Since a single technique did not exhibit the lowest

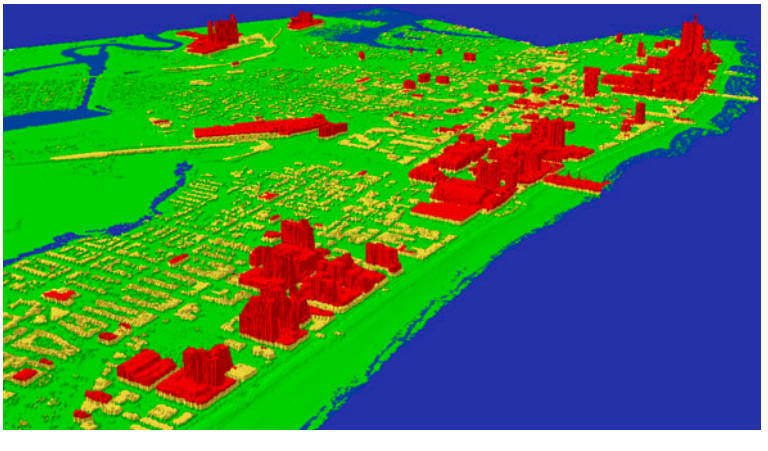


Figure 11. TIN surface generated from DSM for Atlantic City.

Table 2. Commission and omission percentages of the three data-sets compared with actual buildings.

Technique	Commission percentage (%)	Omission percentage (%)
LiDAR	33.18	18.37
OOC	24.70	17.64
DSM	34.32	14.19

Table 3. REG and COR values between different techniques and digitized polygons shown from high to low.

Variables compared	R^2 value	R^2 (adj) value	COR value
Digitized and LiDAR-based polygons	0.74	0.74	0.86
Digitized and OOC-based polygons	0.86	0.86	0.92
Digitized and DSM-based polygons	0.53	0.52	0.72

commission and omission error values together, it was sought best to also check the statistical dependence of the derived data-sets on the actual footprints.

4.1.2. Regression and correlation assessment

All the three approaches showed high regression values indicating that the areas derived by each method fitted closely to the ground reality (Table 3). As can be seen graphically in Figure 12(a)–(c), results of OOC method show strongest positive correlation between derived area and actual area as compared to the other methods.

The evaluation of methods was done based on error of omission and commission as well as a correlation was also studied between manually digitized polygons and footprints derived using the above mentioned three methods. In LiDAR-based footprint detection, while 82% of building areas were completely detected, resulting in an 18% omission error, 67% of detected areas were correct, offering a 33.18% of commission error. Similarly for DSM-based footprint extraction we observed that 86% of the buildings were correctly detected with a commission error of 34% i.e. only 66% footprints detected were correct. However, OOC-based method was almost close to LiDAR derived footprints with 82% of buildings detected however 75% of the buildings were being classified correctly i.e. with a commission error of 25%. Since the commission error was larger than the omission error, the false positive rate of the proposed technique is greater than its false negative rate. This large error is mainly due to the fact that majority of residential buildings are small and LiDAR data alone is not able to capture the boundaries of the buildings properly. In OOC-based method, 92% of the buildings were correctly classified whereas the accuracy in LiDAR and DSM-based method ranged from 86 to 72% respectively as per its correlation values and it was also observed that the adjusted R^2 values are much higher for OOC-based method i.e. 0.86 whereas the adjusted R^2 values for LiDAR and DSM derived method were only 0.74 and 0.52 (Figure 13(a)–(c)).

In the current study, the primary aim was to derive footprints with least error for which OOC was seen to be a more appropriate technique. Most of the buildings in the study area displayed regular patterns and thereby, this technique followed suit and matched best with actual data. Problem of shadows and trees can also be dealt with because shadows are often extremely contrasting in terms of colour/tone as compared

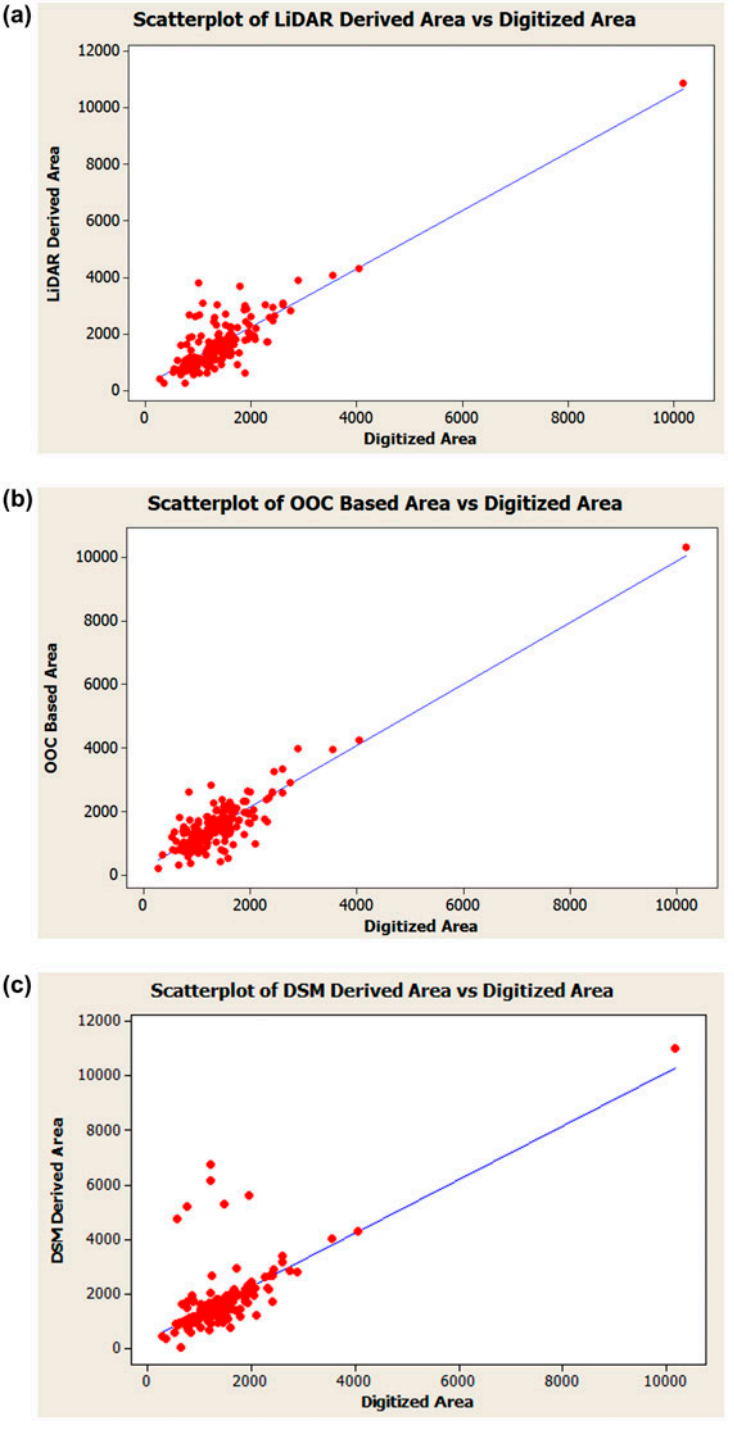


Figure 12. (a) Scatterplot of LiDAR derived area vs. digitized area (square feet); (b) scatterplot of OOC derived area vs. digitized area (square feet); (c) scatterplot of DSM derived area vs. digitized area (square feet).

Figure 13. (a) Correlation between LiDAR derived area and digitized area (square feet); (b) correlation between OOC derived area and digitized area (square feet); (c) correlation between DSM derived area and digitized area (square feet).

to adjacent pixels thereby making them easy to classify, and trees because of their inherent shape/canopy characteristics can be differentiated from others. Even so, the limitation with OOC is that appropriate values of *shape* and *colour* have to be used in order to optimally classify any region. No fixed values can promise global application and hence, have to be adjusted from area to area. Also, only high-resolution aerial images can promise good quality segmentation, which are not commonly available.

The third technique – extraction of footprints using DSMs derived from LiDAR points was a simpler technique to create a 3D surface without dependence on dedicated LiDAR software. A straightforward interpolation of LiDAR points to create a raster surface and its subsequent classification can be executed in several GIS software. Its methodology was simplistic and easier to apply as equated to others however; its final outputs were lacking in accuracy. For creating surface generalizations for any area, this method may be adopted, but for precise footprint measurements, corrections like marked point access discussed by Brédif et al. (2013) would have to be essentially applied.

5. Conclusion

The study shows that building footprint extraction using – classification of LiDAR point cloud data; OOC on aerial photographs; and classification of LiDAR data-based

DSMs. Footprint extraction was possible with all the methods, but qualitative and quantitative results proved that OOC is a better method. It can be concluded that all three approaches are efficient for capturing building footprint and can be applied to any region, although, results may vary as per heterogeneity in shape and geometry. With slight modifications to the parameters used – for filtering, classification or segmentation, these methods would perform well for footprint capturing. Observing the comparative results, it was concluded that OOC performed better than others. It must be however also understood that all techniques pose certain advantages over others. With a prior knowledge of the parameters involved for classification, building footprint capturing from LiDAR data demonstrates to be an efficient technique. Also, for 3D depiction of structures, LiDAR has several advantages over other data sources. However, acquiring accurate footprints for complicated building structures sometimes pose an issue because every building may not be evenly covered by LiDAR beams and tall buildings may overshadow the coverage of lower structures.

Acknowledgement

The authors acknowledge the USGS for providing LiDAR data-sets, duly acquired under the Accessibility Policy (Section 508). We would also like to thank the New Jersey Geographic Information Network developed and maintained by the NJ Office of Information Technology (Office of GIS) for providing high-resolution aerial photographs of the study area. Further, we would like to extend our sincere gratitude to Ms Paige Parker from the GeoCue Corporation for equipping us with LP360 QCoherent Software (Advanced version) which immensely aided in putting this work in its current form.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Alexander C, Smith-Voysey S, Jarvis C, Tansey K. 2009. Integrating building footprints and LiDAR elevation data to classify roof structures and visualise buildings. Comput Environ Urban Syst. 33:285–292.

Awrangjeb M, Zhang C, Fraser CS. 2013. Automatic extraction of building roofs using LIDAR data and multispectral imagery. ISPRS J Photogramm Remote Sens. 83:1–18.

Baatz M, Schape A. 2000. Multi-resolution segmentation: an optimization approach for high quality multi-scale image segmentation. AGIT Symposium, Salzburg.

Bhandari AK, Singh VK, Kumar A, Singh GK. 2014. Cuckoo search algorithm and wind driven optimization based study of satellite image segmentation for multilevel thresholding using Kapur's entropy. Expert Syst Appl. 41:3538–3560.

Bhaskaran S, Paramananda S, Ramnarayan M. 2010. Per-pixel and object-oriented classification methods for mapping urban features using Ikonos satellite data. Appl Geogr. 30:650–665.

Brédif M, Tournaire O, Vallet B, Champion N. 2013. Extracting polygonal building footprints from digital surface models: a fully-automatic global optimization framework. ISPRS J Photogramm Remote Sens. 77:57–65.

Carneiro C, Morello E, Desthieux G. 2009. Assessment of solar irradiance on the urban fabric for the production of renewable energy using Lidar data and image processing techniques. Adv GISci Lec Notes Geoinf Cartogr. 83–112.

Chen Z, Delis A, Bertoni HL. 2004. Building footprint simplification techniques and their effects on radio propagation predictions. Comput J. 47:103–133.

Collin A, Long B, Archambault P. 2012. Merging land-marine realms: spatial patterns of seamless coastal habitats using a multispectral LiDAR. Remote Sens Environ. 123:390–399.

French JR. 2003. Airborne LiDAR in support of geomorphological and hydraulic modelling. Earth Surf Processes Landforms. 28(3):321–335.
Gwenzi D, Lefsky MA. 2014. Modeling canopy height in a savanna ecosystem using spaceborne lidar waveforms. Remote Sens Environ. 154:338–344.
Haralick RM, Shapiro LG. 1985. Image segmentation techniques. Comput Vision Graphics Image Process. 29:100–132.
Hinz S, Baumgartner A. 2003. Automatic extraction of urban road networks from multi-view aerial imagery. ISPRS J Photogramm Remote Sens. 58:83–98.
Hirschmuller H. 2008. Accurate and efficient stereo processing by semi-global matching and mutual information. IEEE Trans. 30:328–341.
Hyde P, Dubayah R, Peterson B, Blair JB, Hofton M, Hunsaker C, Knox R. 2005. Mapping forest structure for wildlife habitat analysis using waveform lidar: validation of montane ecosystems. Remote Sens Environ. 96:427–437.
Kabolizade M, Ebadi H, Ahmadi S. 2010. An improved snake model for automatic extraction of buildings from urban aerial images and LiDAR data. Comput Environ Urban Syst. 34:435–441.
Karakis S, Marangoz AM, Buyuksalih G. 2006. Analysis of segmentation parameters in ecognition software using high resolution quickbird MS imagery. ISPRS Arch. XXXVI-1/W41. ISPRS Workshop on Topographic Mapping from Space, Ankara, Turkey; February 14–16th, 2006.
Lee D, Lee K, Lee S. 2008. Fusion of Lidar and imagery for reliable building extraction. Photogramm Eng Remote Sens. 74:215–225.
Lefsky MA, Cohen WB, Parker GG, Harding DJ. 2002. Lidar remote sensing for ecosystem studies. BioScience. 52:19–30.
Liu ZJ, Wang J, Liu WP. 2005. Building extraction from high resolution imagery based on multiscale object oriented classification. Appears in Geoscience and Remote Sensing Symposium, 2005, IGARSS '05 Proceedings. IEEE Int. 4:2250–2253.
Maan GS, Singh CK, Singh MK, Nagarajan B. 2014. Tree species biomass and carbon stock measurement using ground based-LiDAR. Geocarto Int (ahead-of-print). 2014:1–18.
Ngcofe L, Minnaar H. 2012. A study on automated segmentation for object-based image analysis for geological mapping in the northern Cape Province, South Africa, Proceedings of the 4th GEOBIA; 2012 May 7–9; Rio de Janeiro, Brazil; p. 129.
Ouma YO, Josaphat SS, Tateishi R. 2008. Multiscale remote sensing data segmentation and post-segmentation change detection based on logical modeling: theoretical exposition and experimental results for forestland cover change analysis. Compu Geosci. 34:715–737.
Patino JE, Duque JC. 2013. A review of regional science applications of satellite remote sensing in urban settings. Comput Environ Urban Syst. 37:1–17.
Pe'eri S, Long B. 2011. LIDAR technology applied in coastal studies and management. J Coastal Res Spec Issue. 62:1–5.
Pesaresi M, Benediktsson JA. 2001. A new approach for the morphological segmentation of high-resolution satellite imagery. IEEE Trans Geosci Remote Sens. 39:309–320.
Pierrot-Deseilligny M, Paparoditis N. 2006. A multiresolution and optimization-based image matching approach: an application to surface reconstruction from SPOT5-HRS stereo imagery. Int Arch Photogramm Remote Sens Spat Inf Sci. 36 (Part 1/W41).
Priestnall G, Jaafar J, Duncan A. 2000. Extracting urban features from LiDAR digital surface models. Comput Environ Urban Syst. 24:65–78.
Rottensteiner F. 2003. Automatic generation of high-quality building models from lidar data. IEEE Comput Graphics Appl. 23:42–50.
Rüther H, Martine HM, Mtalo EG. 2002. Application of snakes and dynamic programming optimisation technique in modeling of buildings in informal settlement areas. ISPRS J Photogramm Remote Sens. 56:269–282.
Sahar L, Muthukumar S, French SP. 2010. Using aerial imagery and GIS in automated building footprint extraction and shape recognition for earthquake risk assessment of urban inventories. IEEE Trans Geosci Remote Sens. 48:3511–3520.
Sridharan H, Qiu F. 2013. Developing an object-based hyperspatial image classifier with a case study using WorldView-2 data. Photogramm Eng Remote Sens. 79:1027–1036.
Suveg I, Vosselman G. 2002. Automatic 3D building reconstruction. SPIE. 4661:59–69.

Tang T, Dai L. 2014. Accuracy test of point-based and object-based urban building feature classification and extraction applying airborne LiDAR data. Geocarto Int. 29:710–730.
Uden M, Zipf A. 2013. Open building models: towards a platform for crowdsourcing virtual 3D cities. In: Progress and new trends in 3D geoinformation sciences. Berlin Heidelberg: Springer; p. 299–314.
Vu TT, Yamazaki F, Matsuoka M. 2009. Multi-scale solution for building extraction from LiDAR and image data. Int J Appl Earth Obs Geoinf. 11:281–289.
Wang Z, Wei W, Zhao S, Chen X. 2004. Object-oriented classification and application in land use classification using SPOT-5 PAN imagery. IGARSS proceedings: Geoscience and Remote Sensing Symposium. 2004 Sept 20–24; Anchorage, AK; 3158–3160.
Wegner P. 1987. The object-oriented classification paradigm. Cambridge: MIT Press.
Xiao J, Gerke M, Vosselman G. 2012. Building extraction from oblique airborne imagery based on robust façade detection. ISPRS J Photogramm Remote Sens. 68:56–68.
Yong L, Huayi W. 2008. Adaptive building edge detection by combining LIDAR data and aerial images. Int Arch Photogramm Remote Sens Spat Inf Sci. 37 (Part B1):197–202.
Yousef F, Kerfoot WC, Brooks CN, Shuchman R, Sabol B, Graves M. 2013. Using LiDAR to reconstruct the history of a coastal environment influenced by legacy mining. J Great Lakes Res. 39:205–216.
Zhang K, Yan J, Chen SC. 2006. Automatic construction of building footprints from airborne LIDAR data. IEEE Trans Geosci Remote Sens. 44:2523–2533.