

Segmentation of Airborne Point Cloud Data for Automatic Building Roof Extraction

Syed Ali Naqi Gilani^{a*}, Mohammad Awrangjeb^b and Guojun Lu^c

^a*Faculty of Information Technology, Monash University, Clayton Vic 3800, Australia;* ^b*School of Info. and Comm. Tech. Griffith Sciences, Griffith University, Nathan QLD 4111, Australia;*

^c*School of Eng. and Info. Tech., Federation University Australia, Churchill Vic 3842, Australia*

(Received on XX/XX/XXXX)

Roof plane segmentation is a complex task since point cloud data carry no connection information and do not provide any semantic characteristics of the underlying scanned surfaces. Point cloud density, complex roof profiles, and occlusion add another layer of complexity which often encounter in practice. In this paper, we present a new technique that provides a better interpolation of roof regions where multiple surfaces intersect creating non-manifold points. As a result, these geometric features are preserved to achieve automated identification and segmentation of the roof planes from unstructured laser data. The proposed technique has been tested using the ISPRS benchmark and three Australian datasets, which differ in terrain, point density, building sizes, and vegetation. The qualitative and quantitative results show the robustness of the methodology and indicate that the proposed technique can eliminate vegetation and extract buildings as well as their non-occluding parts from the complex scenes at a high success rate for building detection (between 83.9% to 100% per-object completeness) and roof plane extraction (between 73.9% to 96% per-object completeness). The proposed method works more robustly than some existing methods in the presence of occlusion and low point sampling as indicated by the correctness of above 95% for all the datasets.

Keywords: Airborne LiDAR, roof detection, segmentation, reconstruction

*Corresponding author. Email: alinaqi.gilani@monash.edu

1. Introduction

Buildings in urban environment are indispensable components on the terrain and objects of most interest for various urban information and management systems. The emergence of spatial data acquisition technologies has sped up 3D characterisation of the earth's surface and objects on it (Song, Wu, and Jiang 2015). Among these sources, airborne Light Detection And Ranging (LiDAR) has become the most attractive choice for spatial data researchers due to its ability to capture 3D georeferenced spatial points of the buildings, roads, vegetation, and other objects expediently at a high point density and accuracy. These characteristics make it feasible for automatic extraction and reconstruction of the buildings and their distinct features.

Regarding the detection of the building regions, most of the methods (Awrangjeb and Fraser 2014a, Chen et al. 2012, 2014, Gilani, Awrangjeb, and Lu 2016, Gilani, Awrangjeb, and Lu 2015) initiate a filtering process using a digital surface model (DSM) or digital terrain model (DTM) to classify the 3D point cloud into terrain and off-terrain points, most often, using an elevation threshold. Nevertheless, segmentation and labelling of the coplanar LiDAR points into homogeneous surfaces for roof plane extraction is the most significant process. Two most common strategies reported for roof extraction are model-driven (parametric or bottom-up) and data-driven (non-parametric or top-down) (Awrangjeb and Fraser 2014a, Chen et al. 2014, Maas and Vosselman 1999). In model-driven approach, a predefined catalogue of parameterised roof models is prescribed and the model that best fits the corresponding LiDAR points is chosen. Therefore, the final roof shape is always topographically correct, but the determination of a specific roof model remains a key issue (Song, Wu, and Jiang 2015). Moreover, a complex roof structure cannot be realised if such a parametric roof model is missing in the library (Awrangjeb and Fraser 2014a). In data-driven methods, on the other hand, the building's roof is reconstructed by aggregating the individually derived roof planes from a segmentation algorithm. Such methods preserve the original structure of an underlying surface. However, a challenge here is to determine a roofs' topology and the relationship among the plane surfaces. The main advantage is that such methods are not restricted to any primitive shape or orientation and can reconstruct any polyhedral building in more detail (Awrangjeb and Fraser 2014a, Chen et al. 2014). However, due to the low resolution of the point cloud data, curved surfaces and roof features, for instance, chimneys, dormers, and vents cannot be well approximated, albeit, possible in the presence of high point density.

1.1. *Research challenges and contribution*

Roof plane extraction has been a challenge to the research community and found restricted by the following quality issues: LiDAR data (1) have systematic and stochastic measurement inaccuracy; (2) points are spatially unorganised and have variable point density; (3) have sparsity and gaps due to the occlusion by neighbouring objects, e.g., vegetation clusters (Elberink and Vosselman 2011); (4) have no connection information among the 3D points; (5) have a presence of noisy laser pulses due to physical limitation of data acquisition sensors and multiple reflectance (multipath effect); and (6) have no statistical distributional pattern especially for the points around anisotropic surfaces (Sotoodeh 2006); Also, the absorption of laser pulses by water and the reflection from vents and transparent roof structures are few added issues that make roof plane detection more stimulating.

Urban scenes are characterised by the existence of diverse objects such as buildings, trees, bridges, and road infrastructure, offering a high degree of complexity. In many cases, vegetation is found very close to the buildings that often occludes parts of the roofs. These buildings are generally ignored and therefore, removed during elimination of the false objects. To overcome these limitations, a technique aiming at roof plane extraction and building detection using LiDAR data is proposed. The initial idea discussing building detection was briefly presented in (Gilani, Awrangjeb, and Lu 2016). Howbeit, this paper extends the initial work in the following aspects.

- Point cloud density is used at different stages of execution for handling sparsely sampled point sets and making the proposed technique robust for various data acquisition sources, e.g., airborne and terrestrial (mobile) platforms.
- A roof plane extraction method and a comprehensive objective assessment using several datasets have been included in this study. These datasets differ in scene complexity, topographical conditions, and point density (1.6 to 35 points/m²).
- A new LiDAR-based boundary tracing technique is also included that seamlessly extracts inner and the outer boundaries of an object without any limitation.
- Two new algorithms: anisotropic point selection and saliency feature estimation, are introduced in this research. The first algorithm identifies points on intersecting surfaces using a local rather a global threshold while latter estimates saliency features accurately that help in extracting occluded roof planes and is robust to noise.

The proposed technique has a light computing burden since it uses geographic location and height information of a point cloud for roof plane segmentation and boundary extraction. On the contrary,

other techniques incorporate more features of LiDAR data that include the timestamp, the strength of backscatter (intensity data), colour or scan angle. Note that, the employed point cloud segmentation method prefers buildings with planar surfaces which widely exist in urban environments and herein, the focus of the proposed study.

The remainder of the paper is structured as follows: Section 2 presents a review of the segmentation paradigms and prominent data-driven techniques for building roof extraction. Further, this section describes various saliency feature estimators and the relevant principles. Section 3 discusses the methodology for robust roof extraction and building detection technique including a detailed description of each step of the workflow. Section 4 provides a comprehensive performance evaluation and comparative analysis of the proposed method using four benchmark datasets. Concluding remarks followed by future outlook are finally provided in Section 5.

2. Literature review

Buildings are of particular importance amongst the numerous features contained in an airborne LiDAR data, which are subsequently processed for the extraction of roof primitives. For this objective, the RANdom SAMple Consensus (RANSAC), Hough Transform, and region growing algorithms, which are three major contenders, are often utilised for point set segmentation (Boulch and Marlet 2012, Huang and Brenner 2011, Schnabel, Wahl, and Klein 2007, Vosselman and Klein 2010). In principle, RANSAC is a randomised procedure that iteratively fits an accurate model to a set of observed data which may contain outliers. Hough Transform, however, describes the primitives in which each data point casts its vote for candidate planes in a parameter space. On the other hand, the region growing algorithm finds the primitive shapes from unorganised point cloud by accumulating the neighbouring points into a region satisfying some conditions.

To apprehend the theoretical and practical feasibility of these approaches, Deschaud and Goulette (2010) conducted a comparative study. They showed that RANSAC is very efficient for detecting large planes in noisy point cloud but very slow for the small planes in large datasets. They, further, demonstrated that Hough Transform is computationally expensive and time-consuming for the plane fitting and extraction. Moreover, they argued that region growing, on the other hand, is quite robust and fast segmentation approach and offers a strong resilience against noise (Liu and Xiong 2008). But, it is not highly accurate due to sensitivity and location of initial seed (Deschaud and Goulette 2010). However, this issue can better be addressed when global information is used (Liu and Xiong 2008). Since urban areas are a blend of variable sized roof primitives where large and small planes are equally important,

115 therefore, in this study, we consider all the raised observations for finding a reliable and robust region
116 growing based segmentation algorithm.

117 Zhou and Neumann (2008) proposed a region growing technique for segmenting building rooftops.
118 They used contouring algorithm to extract the building boundaries. Later, histogram statistics was
119 used to determine the principal direction of the buildings and roof planes for footprint generation.
120 However, this method works only for flat roofs and needs manual interaction for the identification of
121 non-flat surfaces. Awrangjeb and Fraser (2014a) presented a roof plane and building extraction method.
122 They first divided the LiDAR data into the ground and non-ground points and then used non-ground
123 points for extraction of the buildings and roof planes. This method, although, achieved a reasonably
124 high pixel-based accuracy but failed to extract small and occluded roof planes.

125 For point cloud filtering, Vosselman (2000) proposed a slope-based classification algorithm that first
126 computes the slope between any two adjacent LiDAR points and classifies them as non-ground points
127 if the slope value is larger than the threshold. However, selection of the threshold value is critical that
128 directly affects the filtering process and can easily flatten terrain details. Rather, a method proposed by
129 Arastounia and Lichti (2013) used the points' height histogram mechanism for dynamic selection of the
130 threshold value. A peak in histogram corresponds to the ground surface. The threshold value was chosen
131 where the bin entries had decreased dramatically and then remained small. However, this automatic
132 selection criterion works fine for the flat terrain and does not accurately produce a threshold value for
133 hilly and sloppy surfaces. Yang et al. (Yang, Xu, and Yao 2014) proposed the Gibbs energy model for
134 the building objects and fits them into the corresponding LiDAR points using reversible jump Markov
135 chain Monte Carlo algorithm. It then refines to get rid of non-ground and outlier points and extract the
136 primitive's outline. There are few other techniques for automatic or semi-automatic object recognition
137 such as the ones presented in the following works (Guo and Du 2017, Teimouri, Mokhtarzade, and
138 Valadan Zoej 2016).

139 An area wide point cloud segmentation is presented by Jochem et al. (2012), where data were processed
140 in the form of several overlapping tiles. The candidate building regions of all the tiles were detected
141 first and then merged into a single polygon layer for roof plane segmentation. However, the roof planes
142 below a certain height (< 1.5 m) and smaller than < 6 m² area were removed. This technique works
143 seamlessly for larger planes and non-occluding building parts. In addition, roof artefacts (e.g. dormers,
144 chimneys) cannot be extracted due to the merger of adjacent candidate regions. Sampath and Shan
145 (2010) presented a framework to extract and reconstruct polyhedral building roofs from LiDAR point
146 cloud data. First, planar and non-planar points were determined through eigenvalue analysis followed

by the extraction of the planar surfaces. Then, the roof vertices, ridges, and edges were used using adjacency matrix for building reconstruction.

Building roofs in urban environments vary from flat to steeply pitched surfaces and often, have a complex arrangement of slopes, gables, and hips. These distinct parts define sharp features at their intersections like edges, ridges, and corners. So, LiDAR points on these features describe fundamental characteristics of the underlying geometry and considering them in advance, improves the performance of a segmentation process (Park, Lee, and Lee 2012) and thereof, an integral component of our segmentation method. Similarly, estimation of saliency features such as, normal and curvature, is an essential step in surface approximation process because the quality of the extracted surfaces heavily depends on the quality of the estimated point normals (Li et al. 2010, Zhang et al. 2013). Often, Principal Component Analysis (PCA) is used but, it overly smoothes the normals at the points near/on the sharp features due to non-robust location and scatter analysis (Li et al. 2010, Nurunnabi, Belton, and West 2012). Therefore, any segmentation based on these erroneous descriptors results in unreliable and inaccurate surfaces (Mitra, Nguyen, and Guibas 2004).

Furthering to smoothness issue, PCA is also found inefficient in estimating normals in the presence of noise and considered a direct and/or indirect reason for the failure of region growing processes (Nurunnabi, Belton, and West 2012). To overcome these limitations, several methods (Boulch and Marlet 2012, Fleishman, Cohen-Or, and Silva 2005, Li et al. 2010, Mitra, Nguyen, and Guibas 2004, Nurunnabi, Belton, and West 2012, Öztireli, Guennebaud, and Gross 2009, Zhang et al. 2013) have been proposed which can handle data inconsistencies to a certain extent. To obtain statistically robust normals along the sharp features, we propose PCA mollification method that uses pre-computed normals to generate consistent point normals. PCA mollification is based on Low-Rank Subspace Clustering framework with Prior Knowledge (LRSPCK) technique (Zhang et al. 2013) that uses PCA's precomputed point normals as prior knowledge and employs an unsupervised learning process to gracefully compute robust normals around anisotropic regions regardless of data inconsistencies. Another advantage of low-rankness is that it better captures the global structure of the data that make it robust to noise (Luo et al. 2014) and enables to handle corrupt data (Zhang et al. 2013). In the literature, mollification has been used in computer graphics for the representation of geometric models (Öztireli, Guennebaud, and Gross 2009). However, to the best of our knowledge, this is a first study for the detection of roof planes and buildings from LiDAR data using PCA mollification.

2.1. PCA mollification

PCA internally reduces dimensions of the data and finds a matrix V (representing eigenvectors $\vec{v}_1, \vec{v}_2, \vec{v}_3$) and scalar λ (representing eigenvalues $\lambda_1 \leq \lambda_2 \leq \lambda_3$), where λ_i denotes a spatial variation along the corresponding \vec{v}_i . In case of 3 dimensional point cloud data, first two eigenvectors can describe a planar surface where the smallest eigenvalue λ_1 corresponding to \vec{v}_1 defines a point normal \hat{n}_i (Nurunnabi, Belton, and West 2012). However, PCA approximates inaccurate normals at the feature points, the points not lying on planar surfaces, such as edges. Therefore, we combine LRSPCK at this stage, which segments the neighbours of each feature point into several isotropic subspaces and re-estimate these normals. An affinity matrix, which is dense amongst the same classes and sparse otherwise, is generated by seeking the lowest rank representation on the PCA normals. A plane is then fit to the feature point and each of its subneighbourhoods to estimate a fitting residual. A subneighbourhood that exhibits a minimum residual is identified as a consistent subneighbourhood and used to approximate an accurate normal.

A careful analysis of LiDAR data shows that building roof planes have diverse geometry and predominantly, three invariants generally exist when a dihedral angle at the intersection of two or more planes makes; 1) an acute/obtuse angle; 2) a right angle; or 3) a jump edge (from mutually superposing surfaces). We demonstrate the robustness of the PCA mollification using three real-world point cloud samples and visually assess the quality of the estimated normals. Figure 1 shows that the point normals approximated using PCA mollification can accurately approximate the underlying surfaces as compared to the PCA.

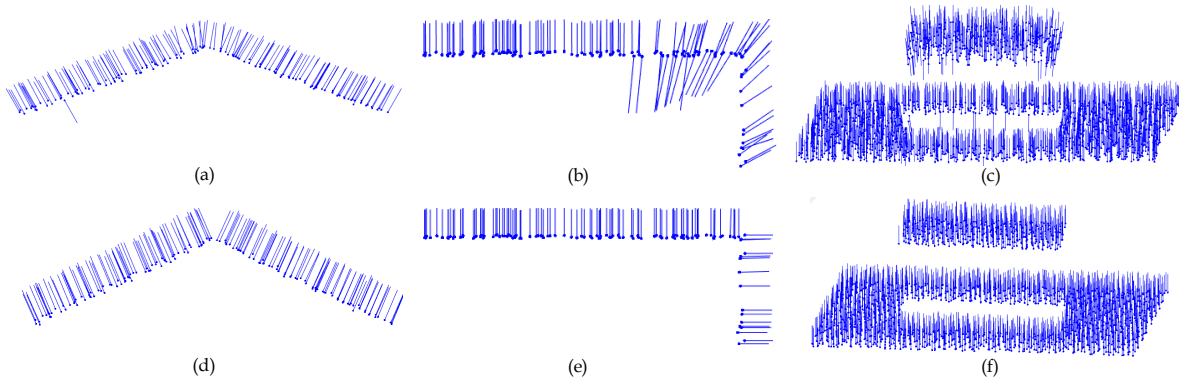


Figure 1. Point normals for three real-world roof plane samples using PCA (a-c) and (d-f) PCA mollification using LRSPCK. Planes forming an acute/obtuse angle (left), right angle (middle) and jump edges (right).

3. Methodology

Figure 2 shows the workflow of the proposed roof extraction and building detection technique. The input data consist of raw LiDAR point set and the corresponding DTM. For this study, DTM with 1 m horizontal resolution was available for each benchmark dataset. Otherwise, it can be generated using any commercial software, such as MARS[®] explorer (Merrick 2015). The proposed technique comprises of three major stages. First, we separate the LiDAR data into the ground and non-ground point sets and use them later to identify the building regions. Second, the proposed segmentation method extracts the planar surfaces from the point cloud of each identified building region by using saliency features. Finally, a refinement procedure eliminates the non-building planes and then, approximate the boundaries of the roof planes and buildings using a proposed boundary tracing algorithm. We used Matlab[®] 2016a for all the experiments and utilised the built-in functions, where applicable, to exploit parallel processing and gain high performance. The detailed explanation of all the intermediate stages is provided in the following sections.

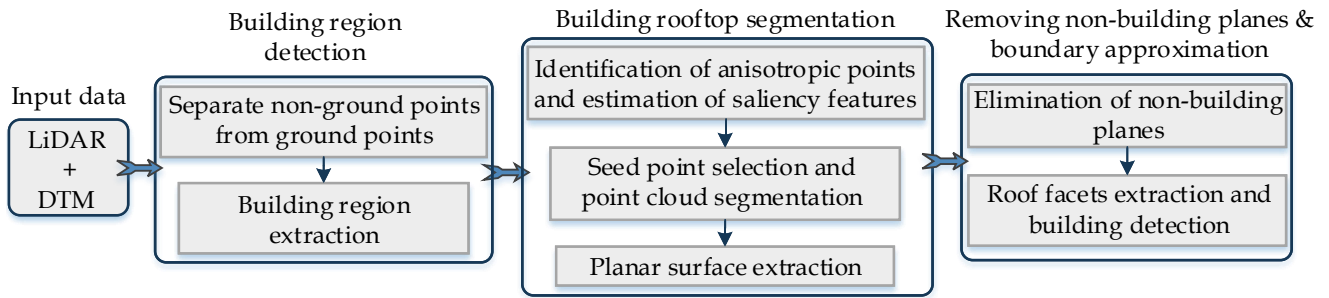


Figure 2. Workflow of the proposed technique.

3.1. Building region detection

Figure 3a presents the test dataset which is an urban area named Aitkenvale (AV) situated in Queensland, Australia. The aerial image will be used for the demonstration of different stages of the proposed methodology and to show the planimetric accuracy of the extracted roof planes and building boundaries. Even though the test dataset is small but it better offers the challenges of vegetation and occlusion as shown in Figure 3a and the magnified rectangle. It covers an area of 66 m \times 52 m containing moderate vegetation and six buildings comprising 24 roof planes. LiDAR coverage of AV comprises of the first pulse returns with a point density of 35 points/m² and a spacing of 0.17 m in both in- and across-flight directions.

The proposed method takes the LiDAR point cloud $C \in \mathbb{R}^3$ (tri-dimensional space) and its corre-

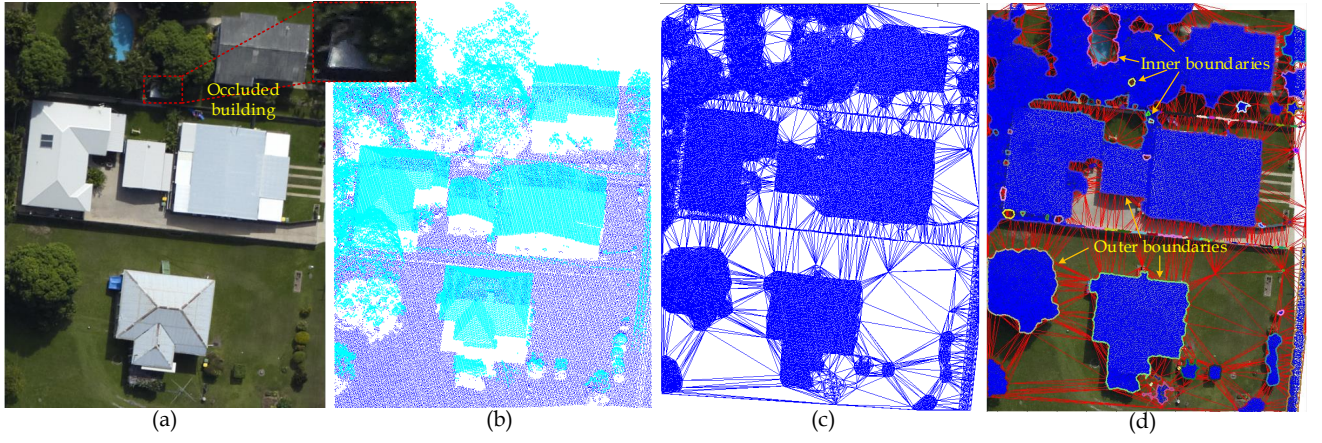


Figure 3. (a) Aerial image of test dataset, (b) LiDAR point set, (c) Delaunay triangularisation of non-ground points, and (d) Building region identification. (For interpretation of the references to color in this figure legend, the reader is referred to the digital version of this article.)

sponding DTM as input. Generally, airborne LiDAR data contain points returned from different features such as ground surface, trees, buildings, and other 3D objects. Therefore, we first separate the area of interest from other ground objects for the detection of building roof planes. For the reason to separate the non-ground points, a height threshold $h_t = h_g + h_{rf}$ for each LiDAR point is computed using its ground-height h_g from DTM and a relief factor h_{rf} which is 1 m in this study. This process eliminates all the low height objects below h_{rf} including bare-earth, roadside furniture, cars, and bushes while preserving the objects above the threshold including buildings and trees. Notably, many points on low height trees and bushes may be classified as the non-ground points provided they are above the h_{rf} . Figure 3b shows the LiDAR points separated into the ground and non-ground points sketched with blue and cyan colours, respectively.

From this point onwards, we use only the non-ground LiDAR points $P \subseteq C$ for extracting the building regions which will subsequently be processed for the detection of roof planes. To identify the building regions, a neighbourhood connection among \mathbb{R}^2 (bi-dimensional space) representation of P is established using the Delaunay triangulation as shown in Figure 3c. The edges of any triangle having a length $\geq 2d_{max}$ are determined as anomalous connections (see red lines in Figure 3d), where d_{max} corresponds to the maximum point spacing of the data. These edges are then removed. The resultant triangles form contiguous regions which do not have any connection with the others and named herein *building regions*. We used Matlab's `DelaunayTri` function for the construction of the Delaunay triangulation and `edges` method to identify the unwanted constrained triangles.

The proposed boundary tracing algorithm further takes the building regions and approximate their boundaries. It can be observed from Figure 3d (digital copy) that each side of an inner triangle of the connected region is associated with exactly two neighbouring triangles. However, one of the sides

of a triangle along the periphery of the region or inscribed hole is associated with only one triangle. It is computationally inefficient to search such triangles sequentially. Therefore, Matlab’s built-in method `freeBoundary` is used to get the edges of triangles along a region’s periphery and inside the holes/concavities. This method returns an unorganised connectivity list of triangle edges where each record has start and end vertices. The boundary tracing algorithm pops the top-most edge from the list and chooses its start vertex as a beginning point of the object boundary. An edge is iteratively selected from the list whose start vertex is the end vertex of the previous edge and adds it into a boundary segment. The boundary approximation stops if the process meets an edge whose end vertex is the beginning point of the boundary. The proposed boundary tracing algorithm continues extracting boundaries of the building regions until the connectivity list has no further edges left.

Unlike the algorithms in Chen et al. (2014) and Awrangjeb (2016), where outer and inner building boundaries are identified and processed separately, the proposed algorithm traces a primitive’s boundary irrespective of its alpha-shape and location. The proposed technique does not struggle like (Chen et al. 2014) and (Weber, Hahmann, and Hagen 2011), where the value of α is carefully chosen in order to avoid producing exclusively convex hull. Our boundary tracing technique, on the other hand, relies only upon a single parameter, i.e., d_{max} to remove the unwanted edges. Furthering the robustness, our method does not degenerate a convex hull because all of the unwanted long edges are removed upfront before the boundary extraction process begins. Another advantage of the proposed method is that it is easy to implement and exploits the underlying hardware for parallel execution. The time complexity of the boundary tracing algorithm using big O notation is approximated as $O(n)$ i.e., linear. Figure 3d presents a snapshot of the proposed boundary tracing technique and also shows the extracted building regions and their boundaries.

3.2. *Building rooftop segmentation*

Estimation of a point normal largely depends on appropriate selection of a local neighbourhood size and the method to search a point’s neighbours. Commonly, three neighbourhood selection methods are widely practised which are k-nearest neighbours (Knn), Delaunay triangulation, and fixed-distance neighbours. However, Knn has an advantage over others due to its adaptiveness towards the sparsity of an unstructured point cloud that makes it suitable for airborne LiDAR processing. This research, therefore, adopts Knn method to determine local neighbourhood for point cloud segmentation and saliency feature estimation. Matlab’s `KDTreeSearcher` and its relevant functions were used to find a point’s local neighbours in an optimised fashion.

273 3.2.1. Feature points selection

274 PCA discussed in Section 2.1 is used for the Eigen value analysis to estimate normal for each LiDAR
 275 point. These normals are then used for the identification of sharp features so that points on/around these
 276 intersecting surfaces can be preserved to avoid degenerating the segmentation process. It is achieved by
 277 computing weight (curvature) w_i of each point $p_i \in P$ that measures the likelihood of p_i belonging to
 278 a sharp feature. w_i is computed using equation 1 as defined by Pauly et al. (Pauly, Keiser, and Gross
 279 2003).

$$w_i = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3} \quad (1)$$

280 The use of a global weight threshold to determine feature points will be irrational since the underlying
 281 point set of each building region can have entirely a different geometry. Therefore, we extended the
 282 original principle of Zhang et al. (2013) that automatically calculates a single threshold for the entire
 283 point set. We incorporated an adaptive threshold estimation mechanism that first computes a histogram
 284 capturing the distribution of the weights for each building region separately. Then, a threshold w_t is
 285 defined as the horizontal ordinate where the plane fitting residual begins to have a slow decrease. We
 286 used points within $2d_{max}$ of each LiDAR point as its k local neighbours for plane fitting. As Zhang et al.
 287 (2013) defines the distribution of $\{w_i\}_{i=1}^N$ as f_w and smooth it by the following function.

$$\min_{\hat{f}_w} \|\hat{f}_w - f_w\|_F + \|Df_w\|_1 \quad (2)$$

288 where D is the second difference matrix, $\|\cdot\|_F$ and $\|\cdot\|_1$ represent l_2 norm and l_1 norm, respectively.
 289 A threshold w_t value for a building region is chosen after the first peak of the smoothed distribution, as
 290 indicated by the red dotted line in Figure 4a. So, the LiDAR points of a building region having weights
 291 below w_t are classified as feature points (red) as shown in Figure 4b. Similarly, all the building regions
 292 are processed concurrently to determine a local region-specific w_t for the selection of feature points. The
 293 results are shown in Figure 5a. The feature points of each building region are further used to estimate
 294 robust normals using the proposed PCA mollification method (see Section 2.1). The estimated normals
 295 in Figure 5b show that the point normals near connected vegetation and across sharp edges are quite
 296 accurate and robust.

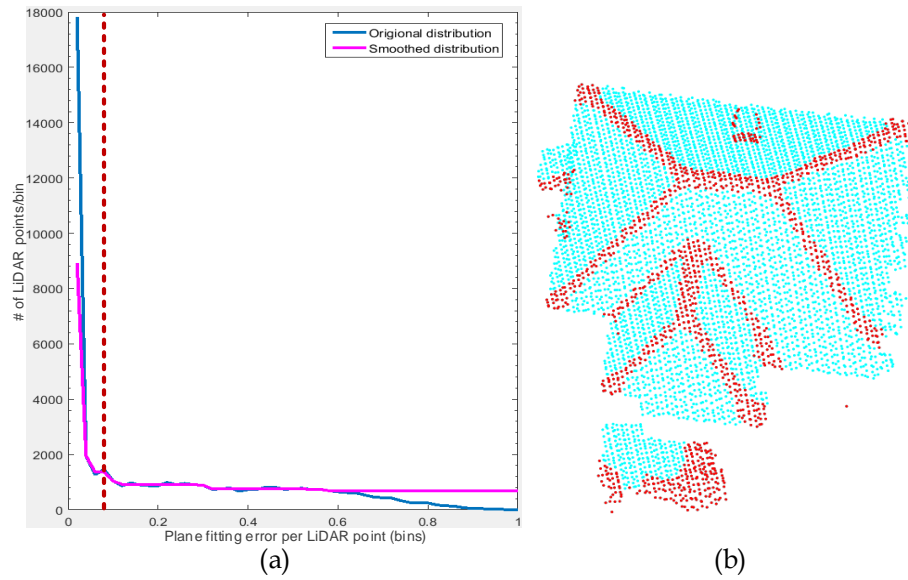


Figure 4. (a) Adaptive selection of weight threshold w_t , and (b) Feature points (red) of a building region. The red dotted line represents the position of a selected threshold w_t .

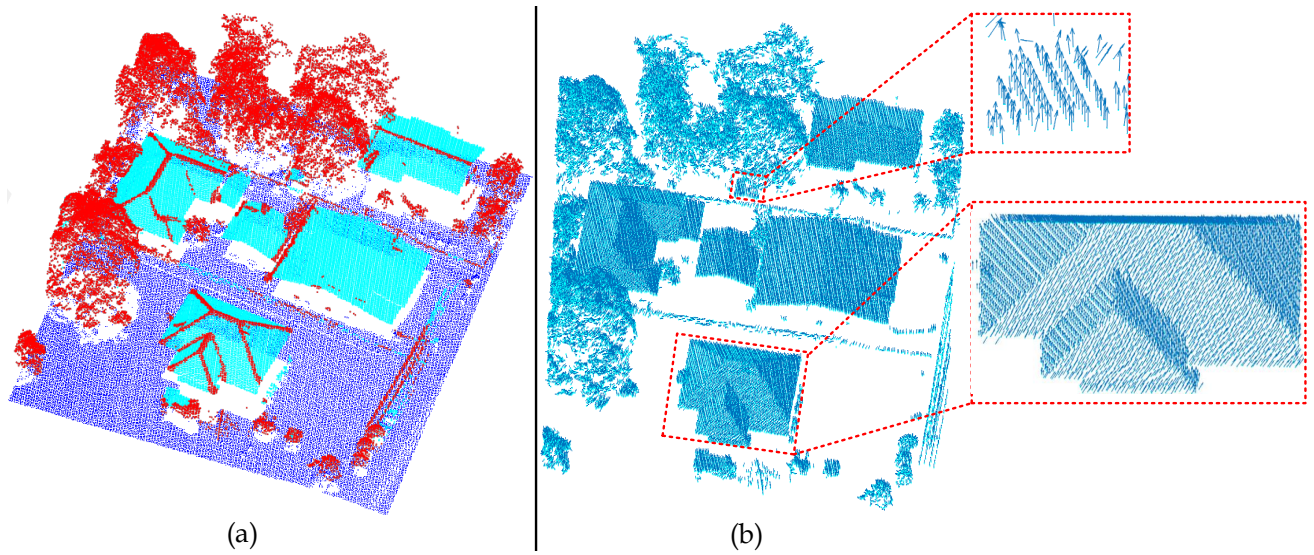


Figure 5. Feature point estimation: (a) Ground points (blue), non-ground points (cyan), feature points (red) and (b) Estimated point normals using proposed PCA mollification. Insets show magnified snapshots of two buildings and their estimated point normals.

297 3.2.2. Segmentation

298 A region-growing technique commonly uses several parameters to determine the coplanarity of the
 299 LiDAR points. However, the proposed segmentation method utilises only the critical parameters to
 300 achieve robustness and efficiency. Therefore, surface curvature and normal orientation (Chen et al.
 301 2014, Nurunnabi, Belton, and West 2012) of a point are chosen as a proximity criterion, whereas, to
 302 distinguish superstructures, a point-to-plane tolerance and plane fitting error thresholds (Awrangjeb
 303 and Fraser 2014a) are used as a coherence criterion. Curvature σ , which measures the rate of change
 304 of surface normal, is estimated using equation 1 that earlier referred as weight. Generally, real-world

305 datasets have inherent noise and even the points reflected from a smooth surface in a local vicinity have
 306 some height variations. Therefore, coherence criterion is found inevitable for better convergence of the
 307 plane surfaces.

308 A region growing technique begins with a selection of a seed point which is sensitive to the segmenta-
 309 tion process. However, the proposed segmentation process defines a robust seed point selection criterion
 310 where a point is chosen from non-feature point set that has the least curvature value. It is believed
 311 that a region growing will be more successful for areas where spatial variation is the least. If point
 312 normals closely approximate the true normals, a usual case is that the angle θ between two normals
 313 across neighbouring surfaces will be larger than a minimum angle threshold. However, when the normals
 314 belong to the same surface, θ will be smaller than the threshold. Note that θ between two normals \hat{n}_i
 315 and \hat{n}_j can be estimated as:

$$\theta_{i,j} = \cos^{-1}|\hat{n}_i \cdot \hat{n}_j| \quad (3)$$

316 Following the principle that a smooth (non-feature) point in a sufficiently small local neighbourhood
 317 always lies on a planar surface, we, therefore, choose a relatively small neighbourhood k i.e., LiDAR
 318 points within $2d_{max}$ for the coplanarity check and the extraction of planar surfaces. Trees, in a point
 319 cloud, if assumed to be composed of several randomly oriented surfaces which are intersecting arbitrarily,
 320 then, they would have a high concentration of feature points as compared to the building roofs. This cue
 321 can be exploited in two meaningful ways— to reduce the amount of data and to eliminate the vegetation.
 322 Therefore, for the point cloud segmentation, only those building regions, where the concentration of
 323 feature points is less than 95%, are processed to extract planar surfaces. Besides, this condition helps
 324 in early removal of several small bushes and trees. Figure 5a shows that tree regions, which can be
 325 spotted from the image in Figure 3a, have a high concentration of feature points (red) in contrast to
 326 the buildings.

327 Segmentation process begins with the selection of a seed point and is chosen from the set of non-
 328 feature points that has the least curvature value. Next, we take k local neighbourhood points N_p of
 329 the seed point. Then, the points having angular difference between the normal of the seed and N_p
 330 within a predefined threshold θ_t are used for plane fitting using PCA. If the plane fitting error and
 331 a difference between point's height and fitted plane's height are smaller than their respective fitting
 332 error ε_t and point-to-plane tolerance ξ_t thresholds, then these LiDAR points are added into a planar
 333 region. This region further grows as long as new neighbouring points meet the criteria; otherwise, a new

region is instantiated. The proposed segmentation process considers all the unallocated LiDAR points including feature and non-feature points to determine a local neighbourhood while expanding a region. This procedure of region growing adds most of the anisotropic points into a planar surface. At this stage, the extracted planar primitives can be seen in Figure 6a.

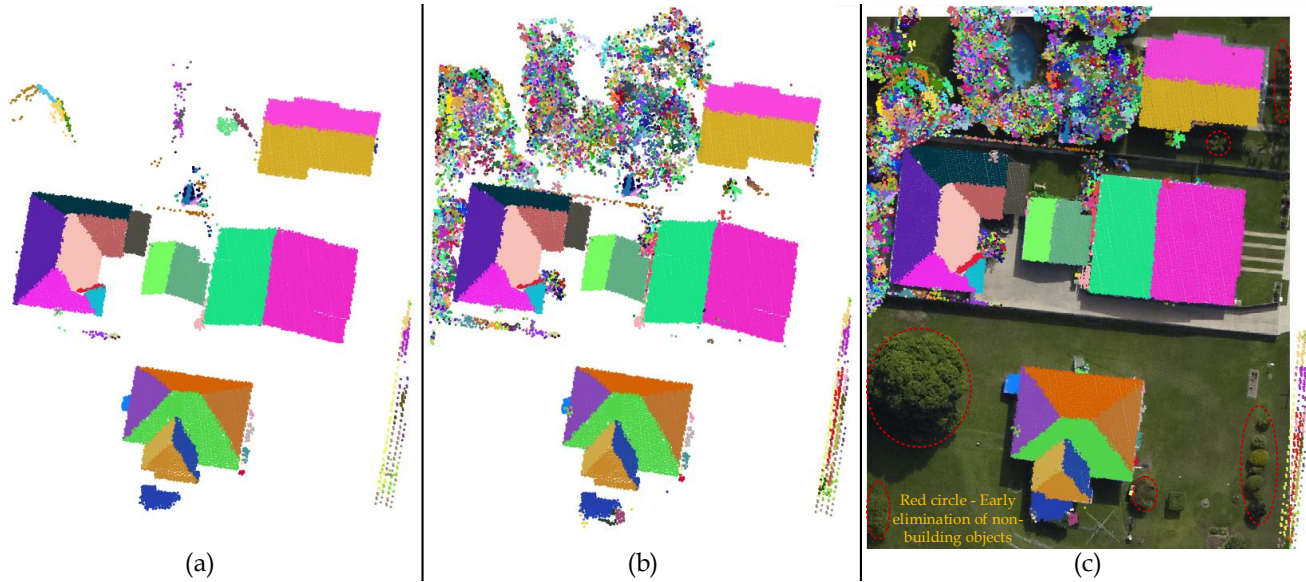


Figure 6. Plane extraction from 3D point cloud using: (a) non-anisotropic points, (b) anisotropic points, and (c) overlaying extracted planes on the test image for the demonstration.

Often, a building roof can have small facets such as dormers, chimneys and vents. The LiDAR returns from these objects might get classified as feature points. Therefore, the feature points that have not yet been classified into any planar surface are segmented following the proposed segmentation technique. Since tree canopies can better be approximated using a non-planar structure, therefore, points reflected from such regions generate a significant number of small planar surfaces. For instance, Figure 6b shows that vegetation breaks up into many smaller planar surfaces. These planar surfaces do not render any parts of the trees rather sets of points reflected from branches that are nearly coplanar. These planes are further shown in Figure 6c using the aerial image, where red-dotted ovals indicate trees which were eliminated due to a high concentration of feature points ($>95\%$) even before the segmentation process started. Ideally, any two points on a truly flat plane should have the similar heights and normals. But, due to inherent noise and surface roughness, there are some random errors in their estimated LiDAR-determined heights and normal directions. Therefore, for better convergence of the segmentation process, the proposed region growing technique adopts the threshold values for $\theta_t = 10^\circ$ from Chen et al. (2014) and $\varepsilon_t = 0.1$ m and $\xi_t = 0.15$ m from Awrangjeb and Fraser (2014a).

3.3. *Removing non-building planes and boundary approximation*

We formulated the problem of removing non-building planes in an unsupervised non-parametric fashion by identifying the underlying patterns of their LiDAR data points. The concentration of feature points r_c has been observed as a good indicator to identify non-planar regions and is proposed to be a useful cue to distinguish vegetation from roof planes. The sizes of tree segments are typically minuscule which can be a useful feature to differentiate between vegetation and roofs (Elberink and Vosselman 2012) and thereof, used herein. Also, a low number of the segmented to unsegmented points ratio r_u in a plane segment, which typically is high for vegetation but low for roof surfaces, is employed. This attribute serves as a reliable cue for the classification of vegetation segments. To avoid removal of small roof planes like chimneys and vents, the plane refinement procedure adopts a criterion from Awrangjeb et al. (Awrangjeb, Lu, and Fraser 2014). This test marks a misclassified plane as a roof plane if it resides within the boundary of an accepted plane.

Typically, urban buildings have a complex arrangement of dormers which do not extend their footprints rather provide a great architectural detail. These dormers are generally built with several intersecting surfaces which are small in size and in proximity, for instance, hexagonal gazebo dormer. As a result, sparsely distributed LiDAR points reflecting from these planar surfaces are often classified as feature points that alone are insufficient for the discrimination of a roof plane in contrast to a tree plane. So, another test is performed to detect if any potential roof plane exists in segmented anisotropic points. This test ensures a plane that has not yet been classified exists in the local neighbourhood of a correct plane and has a straight line segment of at least 2 m in length along its boundary.

The proposed boundary tracing technique is capable of deriving the boundary of a segment in \mathbb{R}^3 because of each transformed point in \mathbb{R}^2 has the same reference index as of the input LiDAR data. Therefore, 3D boundaries of the segmented point cloud can be determined as shown in Figure 7a. In this procedure, a plane segmented from a tree or any non-building region is usually small in size and has both a high concentration of feature points (r_c) and low point usage (r_u) while compared to a roof plane. Therefore, such a plane, known as a false plane, is removed regardless of the other parameters. Also, a plane that exists in a local neighbourhood of a false plane is considered as a false alarm, and thus, is removed. However, the presence of long straight lines along the boundary of a plane and its occurrence near an accepted plane tests are utilised to detect roof planes connected to vegetation and identify small roof facets like dormers and sheds. These line segments are extracted using Canny edge detector from a plane's binary mask. Figure 7b shows a snapshot of false plane removal using image where roof planes (plotted with blue and black-dotted on cyan colour) and non-building segments (represented with

red, yellow, cyan and magenta colours) are sketched. The planar surfaces plotted with black-dotted on cyan colour boundaries correspond to the roof planes which are identified using the neighbourhood criterion and line segment along the boundary test. Figures 7c and 7d show all the roof planes with their boundaries obtained after the elimination of non-building structures in \mathbb{R}^3 and \mathbb{R}^2 , respectively.

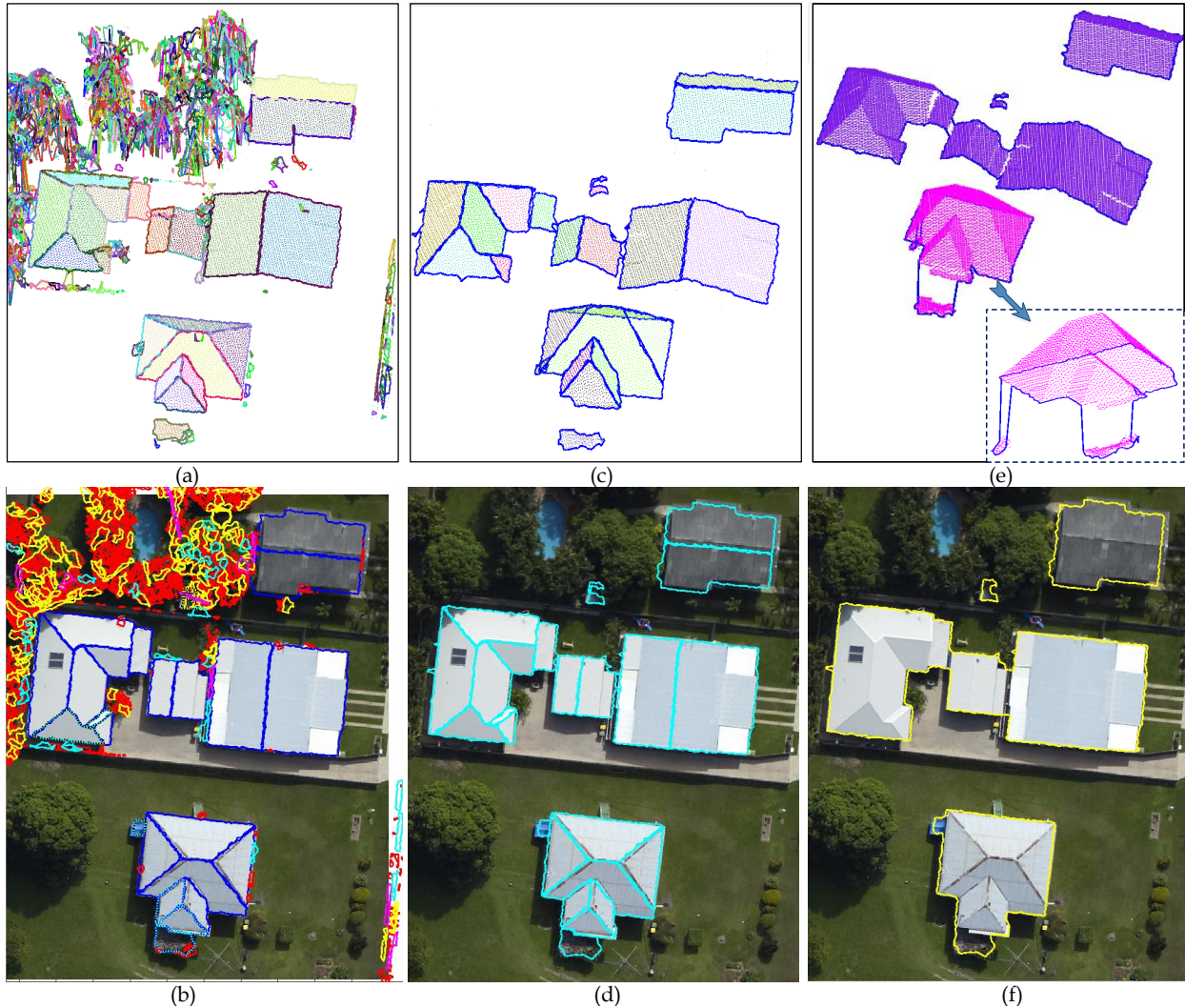


Figure 7. Demonstration of removing non-building planes and building detection using LiDAR and imagery respectively - (a - b) All extracted plane segments and their boundaries, (c-d) Detected building roof planes, and (e-f) Detected buildings.

3.4. Building outline generation

All the buildings in the test area can now be extracted as we have detected all the roof planes. For this purpose, we process all the building regions sequentially. We first collect LiDAR points of all the detected roof planes in a building region. Then, the proposed boundary tracing technique is used to determine any connected region(s)/building(s) and approximate their outlines following the procedure in Section 3.1. Figures 7e and 7f show the extracted buildings with their corresponding boundaries in \mathbb{R}^3 and \mathbb{R}^2 , respectively.

4. Performance study

To validate the performance of the proposed technique, we provide a comprehensive evaluation using four benchmark datasets, which have different LiDAR resolutions, topography, and surrounding conditions. The International Society for Photogrammetry and Remote Sensing (ISPRS) dataset, Vaihingen, Germany has three areas, whereas the other three datasets have one area each captured over different geographic locations in Australia.

4.1. Datasets

Figure 8 shows the datasets used herein for the evaluation where reference boundaries of the roof planes and buildings are sketched using blue and cyan colours respectively. The first dataset is Vaihingen (VH), from the ISPRS benchmark and has three test sites as shown in Figures 8a–c. Area 1 is situated in the centre of the city and characterised by dense construction consisting of historic buildings having complex shapes. Area 2 is characterised by a few high-rise residential buildings surrounded by trees. Finally, Area 3 is purely residential, with detached houses and many surrounding trees. Each area has a point density of 3.5, 3.9 and 3.5 points/m². The number of buildings (larger than 2.5 m²) in these areas is 37, 14, and 56, and the corresponding number of planes is 288, 69 and 235, respectively.

The other three Australian datasets, AV (introduced earlier), Hervey Bay (HB), and Hobart (HT) have point densities of 35, 12, and 1.6 points/m², respectively. Topographically, the AV and HB areas are flat while HT is hilly containing mostly residential buildings with different level of vegetation. The AV and HT datasets have dense vegetation, where the surrounding trees severely occlude many of the buildings. These test areas, shown in Figures 8d–f, cover 66×52 m², 108×104 m², and 303×302 m² respectively. The HB dataset contains 24 buildings (three are between 4 to 5 m² and six are between 5 to 10 m²), consisting of 166 roof planes. The HT dataset has 69 buildings (thirteen are less than 10 m², including four within 1 to 5 m²) containing 257 planes (24 less than 5 m²). The methods in (Gilani, Awrangjeb, and Lu 2016) and (Awrangjeb and Fraser 2014a) are among the other studies which used the employed benchmark datasets for performance analysis.

4.2. Performance evaluation systems

This section discusses the various performance results obtained through qualitative and quantitative evaluation procedures. Performance evaluation results of building and roof plane detection are presented and analysed in separate sections. Due to a non-uniform evaluation system (Rutzinger, Rotten-



Figure 8. ISPRS Vaihingen dataset (a) Area 1, (b) Area 2, and (c) Area 3; Australian dataset, (d) Aitkenvale (AV), (e) Hervey Bay (HB), and (f) Hobart (HT).

steiner, and Pfeifer 2009), the objective evaluation in this study follows both the threshold-based system (Rutzinger, Rottensteiner, and Pfeifer 2009) adopted by the ISPRS and an automatic and threshold-free evaluation system (Awrangjeb and Fraser 2014b) for the German and the Australian datasets, respectively. Both the evaluation systems perform three categories of evaluations: object-based, pixel-based, and geometric, where each category uses several metrics. For example, the object-based metrics (completeness, correctness, quality, under- and over-segmentation errors, and reference cross-lap rates) evaluate the performance by counting the number of buildings, while the pixel-based metrics (completeness, correctness, quality, area-omission, and area-commission errors) measure the detection accuracy by counting the number of pixels. In addition, the geometric metric (root mean square error *i.e.*, RMSE) indicates the accuracy of the extracted boundaries to the reference polygons. For all the datasets, the minimum areas for large and small buildings have been set to 50 m² and 10 m², respectively. The methods in (Gilani, Awrangjeb, and Lu 2016) and (Awrangjeb and Fraser 2014a) are among the other studies which adopted the employed evaluation systems for performance assessment.

4.3. Building detection

Tables 1 and 2 show the per-object and per-area level quantified evaluation for the ISPRS and the Australian datasets. Figures 9 and 10 show the respective detection results and few detection examples from the benchmark test areas. For the ISPRS dataset, the detailed quantitative and qualitative measures for building detection can be found on the ISPRS portal (ISPRS 2016) under acronym Mon5.

Test-case	Per-object			Segmentation			Per-object $\geq 50 \text{ m}^2$			Per-area			<i>RMSE</i>
	C_m	C_r	Q_l	M	N	B	$C_{m,50}$	$C_{r,50}$	$Q_{l,50}$	C_m	C_r	Q_l	
Area 1	91.9	97.1	89.5	0	8	0	100.0	100.0	100.0	89.0	90.2	81.1	0.93
Area 2	92.9	100.0	92.9	0	2	0	100.0	100.0	100.0	91.1	92.8	85.1	0.82
Area 3	83.9	91.7	78.0	0	8	0	97.4	100.0	97.4	89.7	87.9	79.9	0.75
Average	89.6	96.3	86.8	0	6	0	99.1	100.0	99.1	89.9	90.3	82.0	0.83

Table 1. Building detection evaluation results using ISPRS reference classification of Vaihingen, Germany. (C_m = completeness, C_r = correctness, and Q_l = quality in percentage; M = over-segmentation and N = under-segmentation, B = both over- and under-segmentation in number of buildings; *RMSE* = planimetric accuracy in metres).

Test-case	Per-object			Segmentation		Per-object ≥ 50 and 10 m^2				Per-area			<i>RMSE</i>
	C_m	C_r	Q_l	C_{rd}	C_{rr}	$C_{m,50}$	$C_{r,50}$	$C_{m,10}$	$C_{r,10}$	C_m	C_r	Q_l	
AV	100.0	100.0	100.0	0	0	100.0	100.0	100.0	100.0	96.2	96.5	92.9	0.42
HB	100.0	100.0	100.0	0	0	100.0	100.0	100.0	100.0	95.2	92.2	88.1	0.54
HT	86.4	98.1	85.3	3	5	100.0	100.0	96.2	100.0	85.6	93.6	80.9	1.33
Average	95.5	99.4	95.1	1	1.6	100.0	100.0	98.7	100.0	92.3	94.1	87.3	0.8

Table 2. Building detection evaluation results using threshold-free reference classification of Australian datasets. (C_m = completeness, C_r = correctness, and Q_l = quality in percentage; C_{rd} = detection cross-lap (under-segmentation) and C_{rr} = reference cross-lap (over-segmentation) rates; *RMSE* = planimetric accuracy in metres).

First, we will proceed with a qualitative evaluation. The buildings in Figures 9d-e and 10g,i show classic examples of small huts, which were accurately segmented and successfully detected (yellow polygons). At the same time, Figures 9f and 10d-f,i show few complex scenarios, where non-occluded building parts were accurately separated from the nearby vegetation, representing robustness of the proposed technique towards the noise and non-homogeneous surface points. The buildings in Figure 9g have a combination of several flat and hipped rooftops, which collectively form inner boundaries. The proposed technique traced both the inner and outer building boundaries even though their boundary points have no topological relationship. In addition, benefits of the robust normal estimator are found twofold; extract the planar surfaces and eliminate curved surfaces, e.g., trees and domes, as shown in Figure 9h where roof-mounted umbrella was accurately separated from the neighbouring planar surface.

A visual inspection tells that our proposed method obtains good results on all the datasets. However, there are still some point cloud segmentation errors. The proposed algorithm missed few building attachments (green rectangle) and small buildings (cyan rectangle) as shown in Figure 9. This is because VH has sloped terrain and building regions on stilts have LiDAR points below 1 m (h_{rf}) which were ousted during the separation of ground and non-ground points. Whereas, small buildings were missed

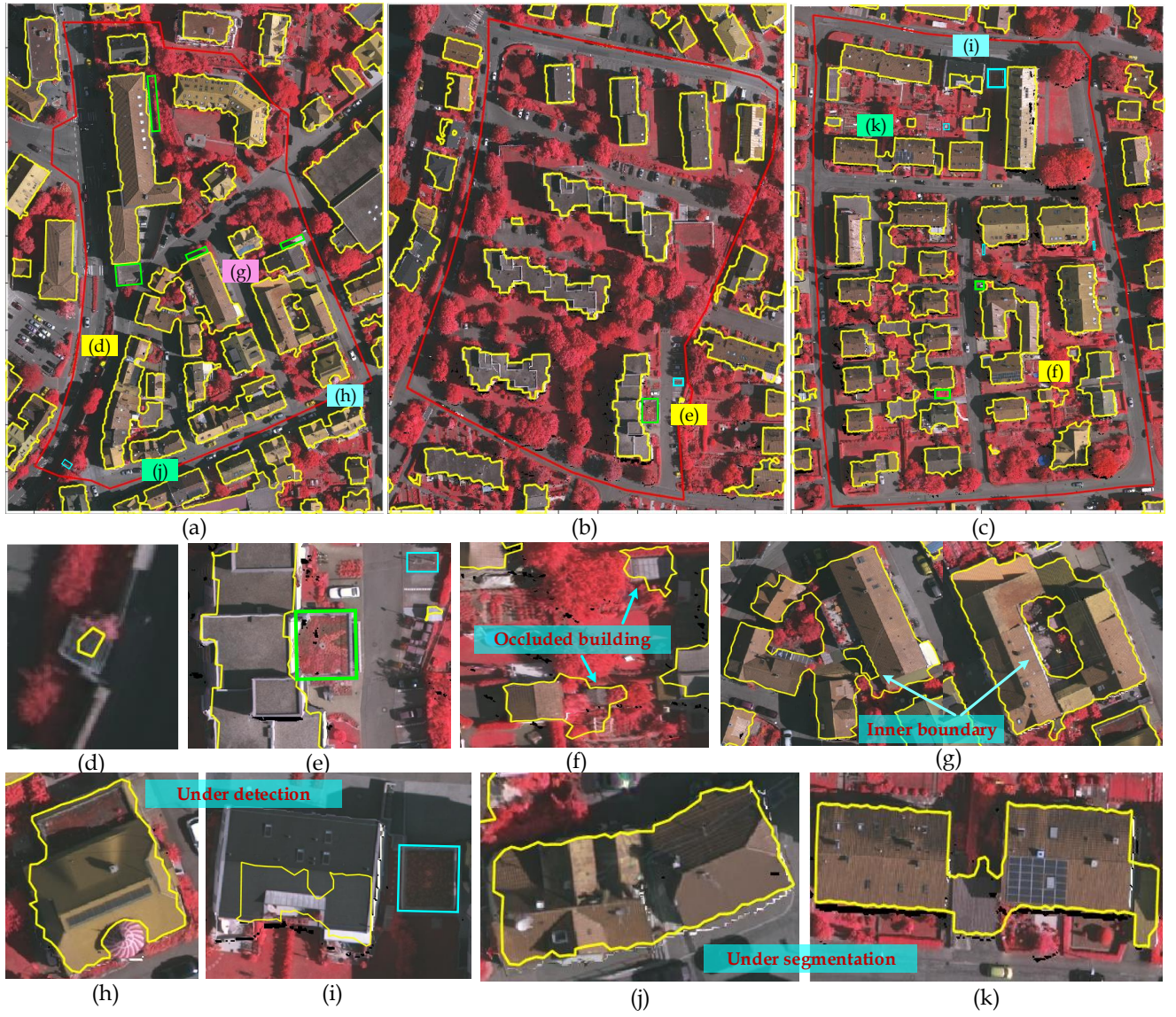


Figure 9. Building detection on VH dataset: (a) Area 1, (b) Area 2, and (c) Area 3. Snapshots (d)-(h) show examples of small, occluded and under detection cases.

from HT dataset as shown in Figure 10c and magnified Figures 10h-i. It was due to the low point density (≈ 1 point/m²) and severe occlusion offered by the neighbouring vegetation.

To quantitatively evaluate the detection results, these detected buildings are further analysed. Considering all the buildings in VH, Table 1 shows that the overall object-level completeness for Area 1 and Area 2 are 91.9% and 92.9% with the corresponding correctness of 97.1% and 100.0%. For the buildings larger than 50 m², the proposed technique achieved 100% object-based completeness, correctness, and quality in Areas 1 and 2. However, Area 3 had a lower object-level accuracy because of missing point cloud data of a partially detected large building as shown in Figure 9i. Some under-segmentation cases occurred when nearby buildings were found close to one another. As shown in Figure 9j, a carport between two buildings and two carports in Figure 9k were merged with their neighbouring buildings.

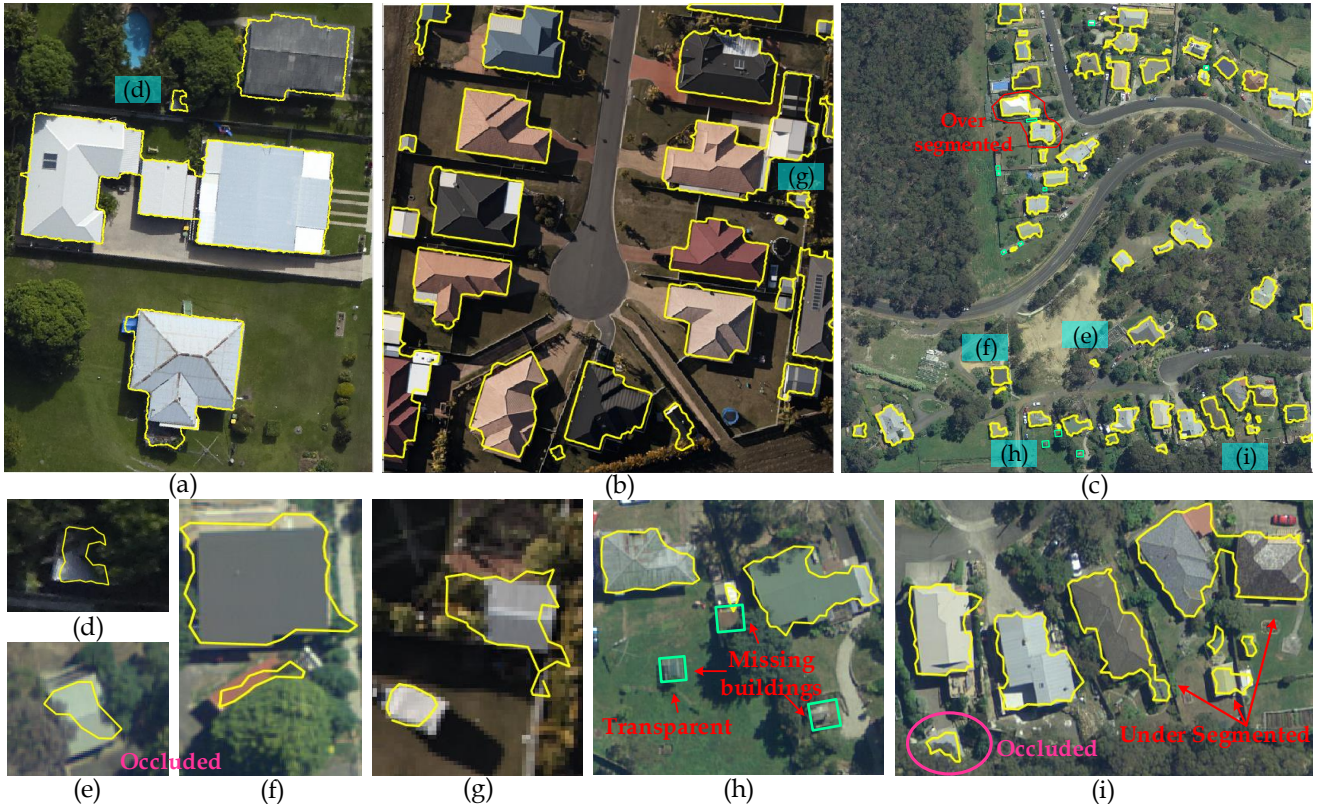


Figure 10. Building detection on the Australian datasets: (a) AV, (b) HB, and (c) HT. Snapshots (d)-(i) show small, occluded and under detection cases.

This unexpected merging was due to the low density of the input LiDAR data that can be avoided by analysing the height spikes among the neighbouring planes at the time of delineating the building periphery. As far as per-area accuracy is concerned, statistics in Table 1 show both the average completeness and correctness were around 90%, indicating accurate detection of correct pixel points. Results in Table 1 further indicate that the proposed technique is absolutely free from over- and many-to-many segmentation errors.

A similar detection trend was observed in the AV and HB datasets, as presented in Table 2. The statistics represent a high detection rate particularly the overall accuracy, quantified in terms of completeness, correctness, and quality, was 100.0%. Additionally, there were no under- and over-segmentation cases because the buildings were well separated from each other and both AV and HB datasets had high point densities in contrast to the VH dataset. Results in Table 2 shows that the proposed method extracted all the small buildings (see Figures 10d,g), which were as small as 10 m² apart from the buildings larger than 50 m². However, while considering all the buildings in HT dataset, object-based completeness was comparatively lower than the other two datasets. The reason was the missing buildings (marked cyan) caused by severe occlusion and transparent roof material (Figures 10c and h). LiDAR pulses generally pass through transparent roof materials and return from the ground. Consequently, such building

points were removed as ground measurements and were not used in point cloud segmentation. The results further show that the buildings over 50 m^2 were extracted with 100% object-based completeness and correctness while nearly equal completeness (96.2%) and correctness of 100.0% were achieved on the buildings over 10 m^2 . However, three very close buildings in HT as shown in Figure 10i were unexpectedly merged since the buildings were situated in a proximity having a distance less than $2d_{max}$. Such complex cases increased the detection cross-lap rate (under-segmentation) in HT. Moreover, a missing transparent building caused reference cross-lap (over-segmentation) as can be seen in Figure 10c. In terms of per-pixel accuracy, the performance of HT was lower (85.6%) than AV (96.2%) and HB (95.2%) but had almost similar correctness rate.

The statistics in Tables 1 and 2 tell us that the achieved planimetric accuracies were close to one to two times the horizontal point spacing of the point cloud data. Overall, these experiments suggest that proposed method obtains a high detection performance and extracts buildings of variable sizes and occluded buildings from flat to hilly terrain under different surrounding complexities. Moreover, the statistics in both the Tables also show that the per- object and pixel accuracies are promoted with the increase of the point cloud density. A constantly higher ($> 97\%$) correctness further tells that our technique is robust to scene complexity. In fact, the qualitative and quantitative results dictate that the proposed detection method can eliminate vegetation and extract buildings as well as their non-occluded parts from the complex scenes at a high object- and pixel-based accuracy.

4.4. Building roof detection

Tables 3 and 4 show the object- and pixel-based evaluation results for roof plane extraction for the ISPRS and Australian datasets, respectively. Figures 11 and 12 show the roof plane extraction results for the ISPRS and Australian dataset, respectively. These figures also present some samples for plane extraction results from the corresponding datasets.

Test-case	Per-object			Segmentation			Per-object $\geq 10 \text{ m}^2$			Per-area			RMSE
	C_m	C_r	Q_l	M	N	B	$C_{m,10}$	$C_{r,10}$	$Q_{l,10}$	C_m	C_r	Q_l	
Area 1	74.3	98.7	73.6	15	32	10	89.8	100.0	89.8	80.7	98.7	79.8	0.76
Area 2	73.9	94.8	71.0	27	1	2	89.6	100.0	89.6	81.4	99.5	81.1	1.06
Area 3	80.9	99.3	80.4	7	36	4	91.1	99.1	90.4	83.8	97.6	82.1	0.79
Average	76.4	97.6	75.0	16.3	23	5.3	90.2	99.7	89.9	82.0	98.6	81.0	0.87

Table 3. Roof planes evaluation results using ISPRS reference classification of Vaihingen, Germany. (C_m = completeness, C_r = correctness, and Q_l = quality in percentage; M = over-segmentation, N = under-segmentation, and B = both over- and under-segmentation in number of buildings; $RMSE$ = planimetric accuracy in metres).

In the VH dataset, the proposed roof extraction algorithm performed better on Area 3, which is purely residential and has detached houses. Table 3 shows that the planes larger than 10 m^2 were detected

Test-case	Per-object			Segmentation		Per-object $\geq 10 \text{ m}^2$			Per-area			<i>RMSE</i>
	C_m	C_r	Q_l	C_{rd}	C_{rr}	$C_{m,10}$	$C_{r,10}$	$Q_{l,10}$	C_m	C_r	Q_l	
AV	96.0	100.0	96.0	4.1	0	100.0	100.0	100.0	89.1	89.2	81.0	0.3
HB	94.6	96.2	92.1	7.8	4.3	100.0	100.0	100.0	89.7	92.5	83.6	0.56
HT	82.49	92.17	77.10	11.20	3.0	93.39	92.17	86.53	72.22	93.86	68.97	1.21
Average	91.0	96.1	88.4	11.5	2.4	97.8	97.4	95.5	83.7	91.9	77.9	0.69

Table 4. Roof planes evaluation results using threshold-free reference classification of Australian datasets. (C_m = completeness, C_r = correctness, and Q_l = quality in percentage; C_{rd} = detection cross-lap (under-segmentation) and C_{rr} = reference cross-lap (over-segmentation) rates; $RMSE$ = planimetric accuracy in metres).

at per-object completeness and correctness of 90.2% and 99.7%. Few examples are shown in Figures 11d-f (yellow ovals). However, there were many under-segmentation cases, where small roof planes were not extracted separately and they merged with the neighbouring large planes as shown in Figure 11g (purple oval). In addition, there were few over-segmentation cases when a roof plane was detected into two or more split facets as shown in Figures 11e,h,i (aqua ovals). Some roof structures were also missed because either they were situated below the height threshold or smaller than 1 m^2 as sketched in Figures 11g-i (red ovals). Consequently, per-object completeness was a bit low for all the areas. However, the proposed technique achieved per-area completeness and correctness of 82.0% and 98.6%, respectively.

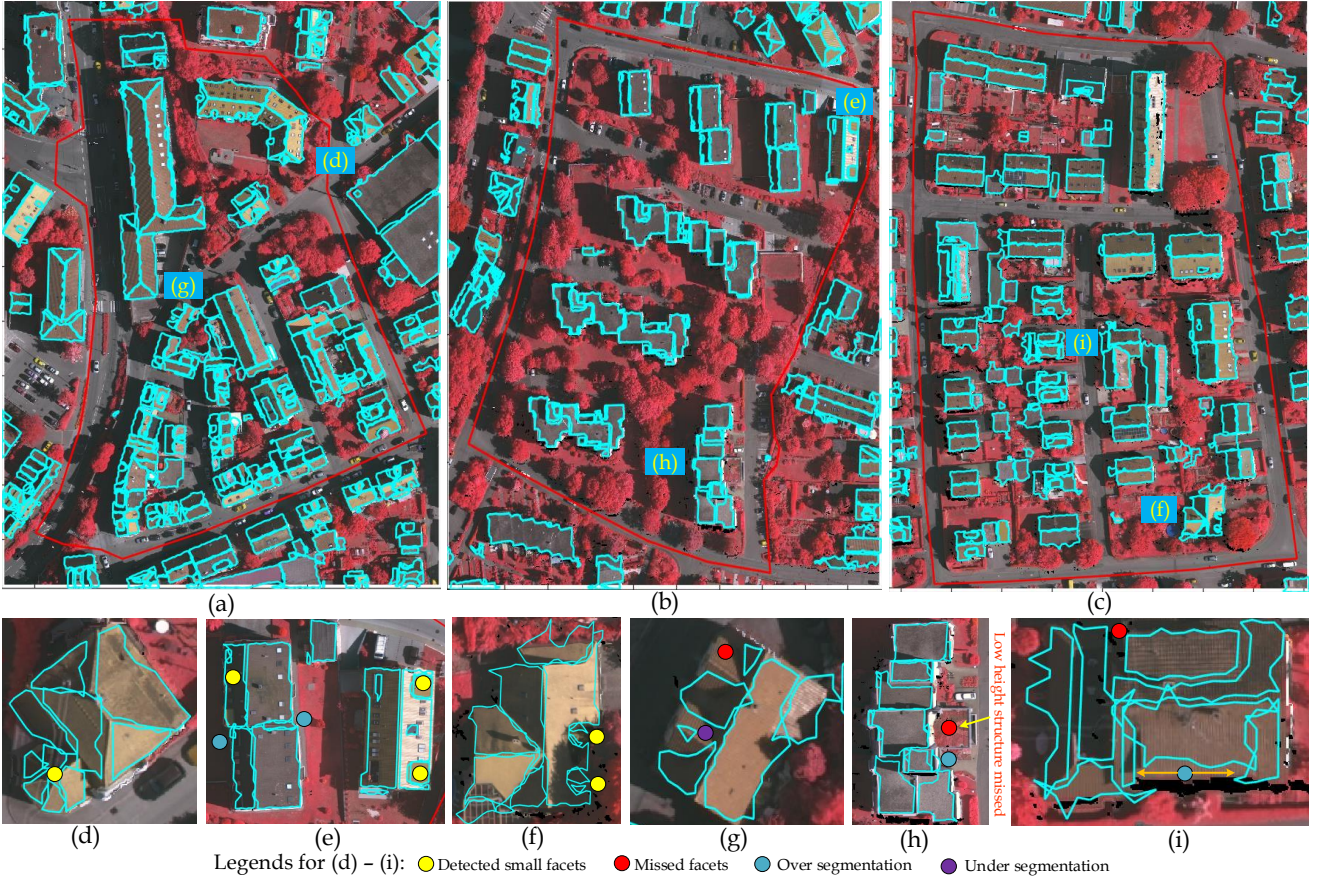


Figure 11. Roof plane extraction on the ISPRS dataset (a) Area 1 (b) Area 2, and (c) Area 3. Areas marked by letters in (a), (b) and (c) are magnified in (d-i).

With increasing point density in Australian datasets, improved roof detection results were obtained in the AV and HB as statistics show in Table 4. The proposed technique extracted the planes larger than

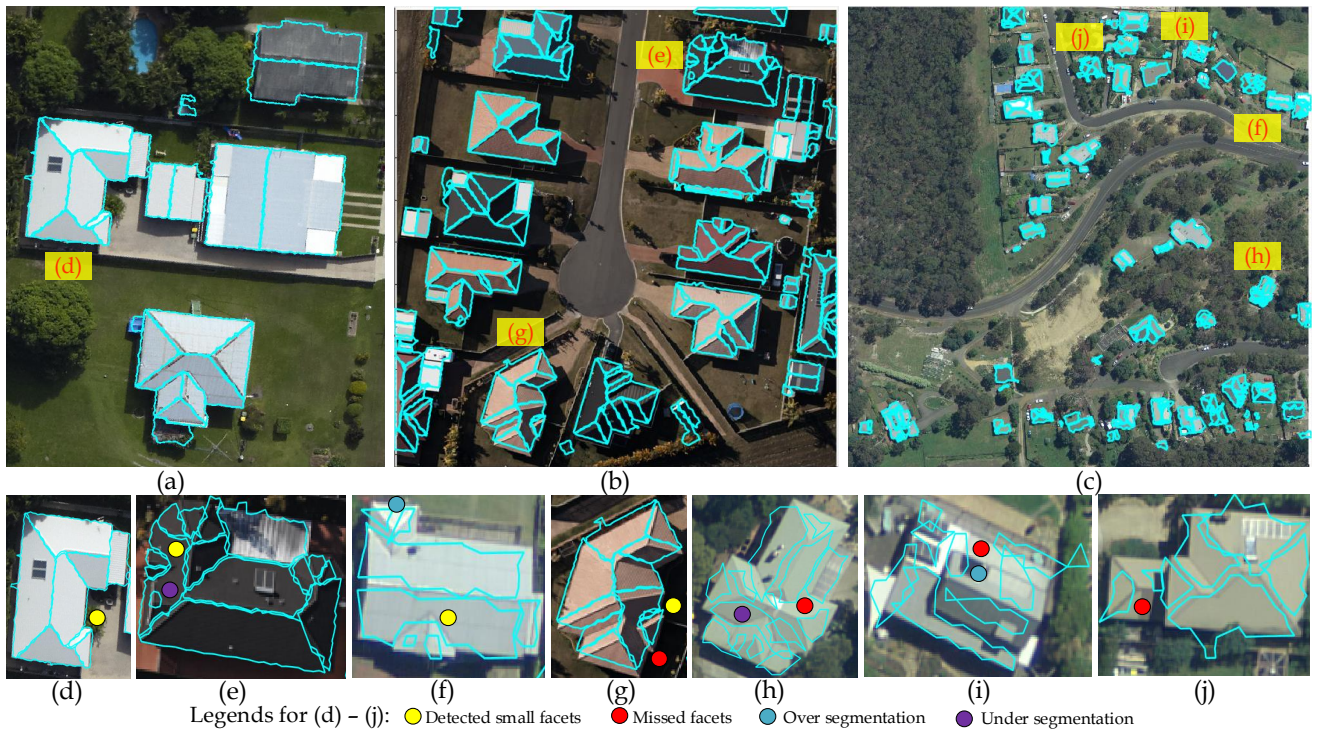


Figure 12. Roof plane extraction on the Australian datasets (a) AV (b) HB, and (c) HT. Areas marked by letters in (a), (b) and (c) are magnified in (d-j).

10 m² with 100% per-object completeness and correctness in AV and HB datasets and correspondingly achieved 96.0% and 94.6% per-object completeness when all the planes were considered (see yellow ovals in Figures 12d-e,g). Figures 12g-j show that many small planes (< 10 m²) were missed (red ovals) and merged into neighbouring planes and thus, increased the under-segmentation and over-segmentation rates. These segmentation errors were more in the HT dataset since it had a low point density (≈ 1 point/m²) and severe occlusion. The statistics further show that the planimetric accuracy in roof plane extraction was within one to two times of the horizontal point spacing for all the datasets. Despite the registration error between detected roofs (from LiDAR data) and reference roofs (from orthoimage), the proposed method achieved nearly 90% per-area completeness in AV and HB datasets and around 72% completeness in HT dataset. A constantly higher correctness of above 90% demonstrates the robustness of our method in roof plane extraction under different complex scenes. It is also observed that the detection performance of the proposed method degrades gracefully with the decrease in point density and does not severely impact the planimetry accuracy of the extracted polygons.

4.5. Comparative analysis and limitation

The proposed automatic technique is data-driven and works with airborne LiDAR data alone. Therefore, for comparative study, the methods which are (1) automatic and data-driven; (2) use only LiDAR data;

and (3) unsupervised, are chosen from ISPRS portal (ISPRS 2016) and a research survey (Rottensteiner et al. 2014). The evaluation results of the HKP, VSK, and TUD on the VH dataset are available on ISPRS portal. However, Yang (Yang, Xu, and Dong 2013) and MA (Awrangjeb and Fraser 2014a) results are taken from their corresponding papers.

Methods	C_m	C_r	$C_{m,50}$	$C_{r,50}$	C_{mp}	C_{rp}	RMSE
Area 1: 3.5 laser points/m ²							
HKP	83.8	100.0	100.0	100.0	92.0	97.4	0.9
Yang	81.1	96.8	100.0	96.6	87.9	91.2	0.9
MA	83.8	96.9	100.0	100.0	92.7	88.7	1.11
VSK	78.4	100.0	96.4	100.0	85.7	98.1	0.8
Proposed	91.9	97.1	100.0	100.0	89.0	90.2	0.9
Area 2: 3.9 laser points/m ²							
HKP	78.6	91.7	100.0	100.0	93.0	98.4	0.6
Yang	78.6	100.0	100.0	100.0	88.8	94.0	0.8
MA	85.7	84.6	100.0	100.0	91.5	91	0.83
VSK	85.7	100.0	100.0	100.0	85.4	98.4	0.9
Proposed	92.9	100.0	100.0	100.0	91.1	92.8	0.8
Area 3: 3.5 laser points/m ²							
HKP	76.8	97.8	97.4	100.0	89.2	97.7	0.7
Yang	73.2	97.6	97.6	92.1	85.2	89.5	0.8
MA	78.6	97.8	97.4	100.0	93.9	86.3	0.89
VSK	75.0	100.0	97.4	100.0	86.3	98.7	1.0
Proposed	83.9	91.7	97.4	100.0	89.7	87.9	0.7

Table 5. Comparing building detection results for the Vaihingen (VH) data set. Object-based C_m = completeness, C_r = correctness ($C_{m,50}$ and $C_{r,50}$ are for buildings over 50 m²) and pixel-based C_{mp} = completeness and C_{rp} = correctness are in percentage. RMSE = planimetric accuracy in metres.

Methods	C_m	C_r	$C_{m,10}$	$C_{r,10}$	$M/N/B$	RMSE	RMSZ
Area 1: 3.5 laser points/m ²							
MA	76.4	83.3	84.4	84.9	6/42/7	1.05	0.41
VSK	72.2	96.7	80.3	95.9	7/42/6	0.9	0.3
TUD	67.4	96.2	68.0	97.8	1/33/1	0.8	0.2
Proposed	74.3	98.7	89.8	100.0	15/32/10	0.8	0.3
Area 2: 3.9 laser points/m ²							
MA	73.9	91.9	93.8	92.6	7/3/1	0.74	0.37
VSK	73.9	100.0	91.7	100.0	3/5/1	0.7	0.3
TUD	68.1	98.1	85.4	100.0	5/3/0	0.6	0.3
Proposed	73.9	94.8	89.6	100.0	27/1/2	1.06	1.40
Area 3: 3.5 laser points/m ²							
MA	82.1	93.9	92.7	96.7	5/45/0	0.89	0.27
VSK	76.6	99.1	86.3	100.0	3/50/0	0.8	0.1
TUD	74.5	93.0	83.1	98.0	0/42/1	0.7	0.1
Proposed	80.9	99.3	91.1	99.1	7/36/4	0.8	0.2

Table 6. Comparing plane results for the Vaihingen (VH) data set. Object-based C_m = completeness, C_r = correctness ($C_{m,10}$ and $C_{r,10}$ are for buildings over 10 m²). M = over-segmentation, N = under-segmentation, and B = both over- and under-segmentation in number of buildings; RMSE = planimetric accuracy; RMSZ = height accuracy; in metres.

In all three VH benchmark areas, Table 5 shows that the proposed building detection technique offered significantly better per-object level completeness and similar correctness than the alternative methods. For the buildings larger than 50 m², our technique achieved 100% accuracy in terms of completeness and correctness, whereas, VSK was unsuccessful in detecting a large building in Area 1 as shown by low $C_{m,50}$ value in Table 5. In terms of per-area accuracy, HKP and MA, however, obtained slightly

more per-area completeness in Areas 1 and 3 since our method missed few carports below the height threshold. In terms of planimetric accuracy of the extracted polygons, the proposed technique obtained better or slightly low performance than the counterparts in all three areas.

The proposed roof plane extraction method, in all areas of the VH dataset, offered better per-object completeness and correctness than VSK and TUD as Table 6 shows. However, MA remained slightly better than the proposed technique when all the planes are considered in Areas 1 and 2. Concerning the planes larger than 10 m^2 , Table 6 shows that our technique achieved better correctness but performs slightly poorer than MA on Areas 2 and 3 in terms of completeness. Nevertheless, it performed better than VSK and TUD on Areas 1 and 3 when performance is quantified as completeness ($C_{m,10}$). The planimetric accuracy (RMSE) and height error (RMSZ) of the proposed technique are not much different from the counterparts, but RMSZ difference in Area 2 is larger. It is due to the reason that we do not clean the LiDAR data and treat each point as information.

Table 6 shows that the proposed plane extraction technique encountered more segmentation errors than the other alternative methods, which is explained using Figures 13a,b taken from Area 1 of the VH dataset. The left roof of building in Figure 13a has two surfaces: roof's base plane and superpose plane, i.e., a dormer. LiDAR points of the base primitive had a hole due to the points of the superimposed primitive. Therefore, inner and outer boundaries were approximated using LiDAR points of the base surface as sketched with red colour in Figure 13a while cyan colour boundary corresponds to the outline of the overhung dormer. The evaluation system at the ISPRS considers that the proposed technique detected the inner primitive twice and such cases badly affected the overall performance of the proposed plane extraction technique. Regarding over-segmentation, roof plane coloured red in Figure 13b was although identified accurately but two boundary segments were determined because the corresponding LiDAR points were separated more than twice of the point density. In addition, irregular surfaces like a dome and curvy structures were eliminated as the employed point cloud segmentation method prefers objects with planar surfaces. Therefore, a missing roof-mounted umbrella can be seen in Figure 13c.

5. Conclusion and outlook

Point cloud segmentation is meaningful for various applications, but this task is challenging due to inherent characteristics of the input data. This paper particularly focuses on the automatic detection of buildings and their roof planes, which define the general three steps: feature preservation, surface growing, and false plane elimination. The proposed technique is data-driven and introduces a feature preservation based segmentation algorithm that effectively uses robust saliency features for roof extrac-

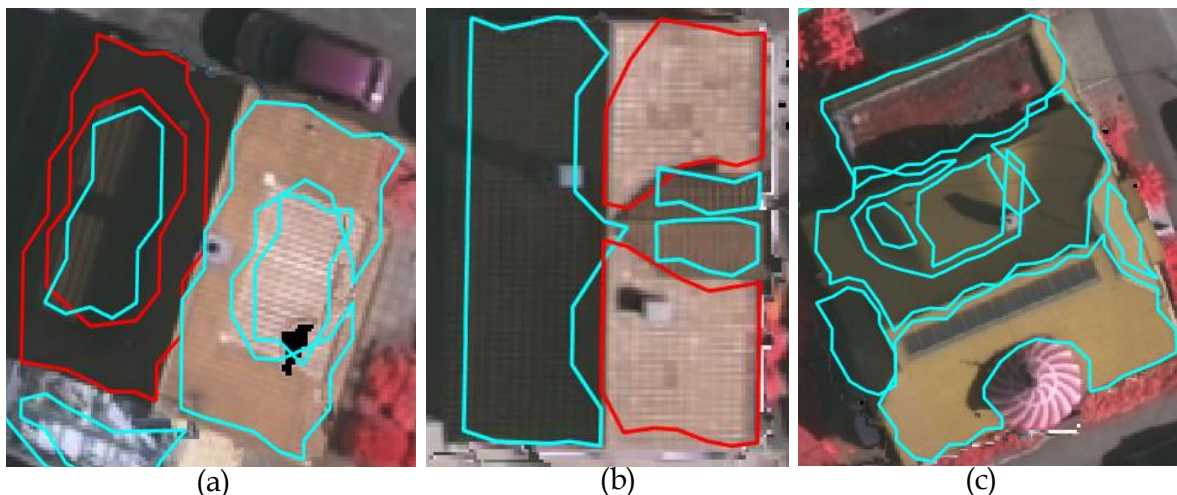


Figure 13. Roof plane samples: (a) Multi-plane extraction (b) Over-segmentation, and (c) Missing roof mounted umbrella.

tion, which is less sensitive to noise and avoids over- and under-segmentation. We have also proposed a boundary extraction technique that seamlessly extracts a primitive's boundary and approximate the outline of any inner hole using LiDAR points. It has been demonstrated experimentally that the proposed technique achieves a high building detection rate and good roof plane extraction performance on several datasets of variable point densities, terrain and surrounding complexity. It is equally capable of detecting small buildings and extracting small roof planes. Moreover, in most cases, our method can separate buildings and non-occluded parts from the connected vegetation.

The proposed technique extracts roof planes and buildings directly from LiDAR data and therefore, the planimetric accuracy of the detected polygons is limited by its point density. At present, it does not work with any irregular building primitives, such as domes and curved roof surfaces. Also, the extracted polygons approximated herein have ragged boundaries. The research to smooth the boundaries of building primitives and reconstruct building roof models using image data is under investigation. Furthermore, it would be interesting to incorporate geometric clues in order to detect irregular surfaces. Future work also aims to use LiDAR intensity and image gradient to extract the buildings constructed with transparent materials.

Acknowledgment

The research is supported by the Discovery Early Career Researcher Award by the Australian Research Council (project number DE120101778). The Aitkenvale and Hervey Bay datasets were provided by Ergon Energy (<http://www.ergon.com.au>) in Queensland, Australia. The Hobart dataset was supplied by Photomapping Services, Australia. The Vaihingen dataset was provided by the German Society for

Photogrammetry, Remote Sensing and Geoinformation (DGPF) Cramer (2010). Special thanks go to Uwe Breitkopf and Markus Gerke of the ISPRS Working Group III/4 for evaluating the roof and building outlines.

References

- Arastounia, M, and DD Lichti. 2013. "Automatic extraction of insulators from 3D LiDAR data of an electrical substation." *ISPRS Ann Photogramm Remote Sens Spat Inf Sci* 2 (5/W2): 19–24.
- Awrangjeb, M. 2016. "Using point cloud data to identify, trace, and regularize the outlines of buildings." *Int J Remote Sens* 37 (3): 551–579.
- Awrangjeb, M., and C. S. Fraser. 2014a. "Automatic Segmentation of Raw LiDAR Data for Extraction of Building Roofs." *Remote Sens* 6 (5): 3716–3751.
- Awrangjeb, M., and C.S. Fraser. 2014b. "An Automatic and Threshold-Free Performance Evaluation System for Building Extraction Techniques From Airborne LiDAR Data." *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7 (10): 4184–4198.
- Awrangjeb, M., G. Lu, and C. S. Fraser. 2014. "Automatic building extraction from LiDAR data covering complex urban scenes." *Int Arch Photogramm Remote Sens Spat Inf Sci* I: 25–32.
- Boulch, A., and R. Marlet. 2012. "Fast and robust normal estimation for point clouds with sharp features." In *Comput Graph Forum*, Vol. 31, 1765–1774.
- Chen, D., L. Zhang, J. Li, and R. Liu. 2012. "Urban building roof segmentation from airborne LiDAR point clouds." *Int J Remote Sens* 33 (20): 6497–6515.
- Chen, D., L. Zhang, PT Mathiopoulos, and X. Huang. 2014. "A Methodology for Automated Segmentation and Reconstruction of Urban 3-D Buildings from ALS Point Clouds." *IEEE J Sel Top Appl Earth Obs Remote Sens* 7 (10): 4199–4217.
- Cramer, M. 2010. "The DGPF-test on digital airborne camera evaluation–overview and test design." *Photogrammetrie Fernerkundung Geoinf.* 2010 (2): 73–82.
- Deschaud, Jean-Emmanuel, and François Goulette. 2010. "A fast and accurate plane detection algorithm for large noisy point clouds using filtered normals and voxel growing." In *Proc 3D Processing Vis Transm Conf*, .
- Elberink, Sander Oude, and George Vosselman. 2011. "Quality analysis on 3D building models reconstructed from airborne laser scanning data." *ISPRS J Photogramm Remote Sens* 66 (2): 157–165.
- Elberink, Sander Oude, and George Vosselman. 2012. "Entities and Features for Classification of Airborne Laser Scanning Data in Urban Area." *ISPRS Ann Photogramm Remote Sens Spat Inf Sci*

- 257–262.
- Fleishman, S., D. Cohen-Or, and C. Silva. 2005. “Robust moving least-squares fitting with sharp features.” In *ACM Trans Graph*, Vol. 24, 544–552.
- Gilani, S.A.N., M. Awrangjeb, and G. Lu. 2016. “An Automatic Building Extraction and Regularisation Technique Using LiDAR Point Cloud Data and Orthoimage.” *Remote Sens* 8 (3): 258.
- Guo, Zhou, and Shihong Du. 2017. “Mining parameter information for building extraction and change detection with very high-resolution imagery and GIS data.” *GIScience & Remote Sensing* 54 (1): 38–63.
- Huang, Hai, and Claus Brenner. 2011. “Rule-based roof plane detection and segmentation from laser point clouds.” In *Joint Urban Remote Sens Event*, 293–296.
- ISPRS. 2016. “Results: ISPRS Test Project on Urban Classification and 3D Building Reconstruction.” (accessed on Oct 28, 2016), <http://www2.isprs.org/commissions/comm3/wg4/results.html>.
- Jochem, Andreas, Bernhard Höfle, Volker Wichmann, Martin Rutzinger, and Alexander Zipf. 2012. “Area-wide roof plane segmentation in airborne LiDAR point clouds.” *Comput Environ Urban Syst* 36 (1): 54–64.
- Li, B., R. Schnabel, R. Klein, Z. Cheng, G. Dang, and S. Jin. 2010. “Robust normal estimation for point clouds with sharp features.” *Comput Graph* 34 (2): 94–106.
- Liu, Y., and Y. Xiong. 2008. “Automatic segmentation of unorganized noisy point clouds based on the Gaussian map.” *Comput Aided Des* 40 (5): 576–594.
- Luo, Pei, Zhuangzhi Wu, Chunhe Xia, Lu Feng, and Bo Jia. 2014. “Robust normal estimation of point cloud with sharp features via subspace clustering.” In *Int Conf Graphic and Image Proc*, Vol. 9069, 90691S–90691S–6.
- Maas, Hans-Gerd, and George Vosselman. 1999. “Two algorithms for extracting building models from raw laser altimetry data.” *ISPRS J Photogramm Remote Sens* 54 (2): 153–163.
- Merrick. 2015. “The Merrick Advanced Remote Sensing (MARS®).” (accessed on 29 December 2015), <http://www.merrick.com/>.
- Mitra, N. J., A. Nguyen, and L. Guibas. 2004. “Estimating surface normals in noisy point cloud data.” *Int J Comput Geom Appl* 14 (04n05): 261–276.
- Nurunnabi, A., D. Belton, and G. West. 2012. “Robust segmentation in laser scanning 3d point cloud data.” In *Int Conf Digit Image Comput Tech Appl*, 1–8.
- Öztireli, A Cengiz, Gael Guennebaud, and Markus Gross. 2009. “Feature Preserving Point Set Surfaces based on Non-Linear Kernel Regression.” In *Comput Graph Forum*, Vol. 28, 493–501.

- 657 Park, Min Ki, Seung Joo Lee, and Kwan H Lee. 2012. "Multi-scale tensor voting for feature extraction
658 from unstructured point clouds." *Graph Models* 74 (4): 197–208.
- 659 Pauly, M., R. Keiser, and M. Gross. 2003. "Multi-scale Feature Extraction on Point-Sampled Surfaces."
660 In *Comput Graph Forum*, Vol. 22, 281–289.
- 661 Rottensteiner, Franz, Gunho Sohn, Markus Gerke, Jan Dirk Wegner, Uwe Breitkopf, and Jaewook Jung.
662 2014. "Results of the ISPRS benchmark on urban object detection and 3D building reconstruction."
663 *ISPRS J Photogramm Remote Sens* 93: 256–271.
- 664 Rutzinger, M., F. Rottensteiner, and N. Pfeifer. 2009. "A comparison of evaluation techniques for
665 building extraction from airborne laser scanning." *IEEE J. Sel. Top. Appl. Earth Obs. Remote*
666 *Sens.* 2 (1): 11–20.
- 667 Sampath, Aparajithan, and Jie Shan. 2010. "Segmentation and reconstruction of polyhedral building
668 roofs from aerial LiDAR point clouds." *IEEE Trans Geosci Remote Sens* 48 (3): 1554–1567.
- 669 Schnabel, Ruwen, Roland Wahl, and Reinhard Klein. 2007. "Efficient RANSAC for point-cloud shape
670 detection." In *Comput Graph Forum*, Vol. 26, 214–226.
- 671 Song, Jingwei, Jianwei Wu, and Yongyao Jiang. 2015. "Extraction and reconstruction of curved surface
672 buildings by contour clustering using airborne LiDAR data." *Int J Light Electron Opt* 126 (5):
673 513–521.
- 674 Sotoodeh, S. 2006. "Outlier detection in laser scanner point clouds." *Int Arch Photogramm Remote Sens*
675 *Spat Inf Sci* 36 (5): 297–302.
- 676 Teimouri, Maryam, Mehdi Mokhtarzade, and Mohammad Javad Valadan Zoej. 2016. "Optimal fusion
677 of optical and SAR high-resolution images for semiautomatic building detection." *GIScience &*
678 *Remote Sensing* 53 (1): 45–62.
- 679 Vosselman, G, and R Klein. 2010. "Visualisation and structuring of point clouds." *Airborne Terr Laser*
680 *Scanning* 1: 43–79.
- 681 Vosselman, George. 2000. "Slope based filtering of laser altimetry data." *Int Arch Photogramm Remote*
682 *Sens* XXXIII: 935–942.
- 683 Gilani, SAN., M. Awrangjeb, and G. Lu. 2015. "Fusion of LiDAR Data and Multispectral Imagery
684 for Effective Building Detection Based on Graph and Connected Component Analysis." *Int Arch*
685 *Photogramm Remote Sens Spat Inf Sci* I: 65 – 72.
- 686 Gilani, SAN., M. Awrangjeb, and G. Lu. 2016. "Robust Building Roof Segmentation using Airborne
687 Point Cloud Data." *IEEE Int Conf Image Proc* 859 – 863.
- 688 Weber, Christopher, Stefanie Hahmann, and Hans Hagen. 2011. "Methods for feature detection in point

- clouds.” In *OpenAccess Series in Informatics*, Vol. 19.
- Yang, Bisheng, Wenxue Xu, and Zhen Dong. 2013. “Automated Extraction of Building Outlines From Airborne Laser Scanning Point Clouds.” *IEEE Geosci Remote Sens Lett* 10 (6): 1399 – 1403.
- Yang, Bisheng, Wenxue Xu, and Wei Yao. 2014. “Extracting buildings from airborne laser scanning point clouds using a marked point process.” *GIScience & Remote Sensing* 51 (5): 555–574.
- Zhang, J., J. Cao, X. Liu, J. Wang, J. Liu, and X. Shi. 2013. “Point cloud normal estimation via low-rank subspace clustering.” *Comput Graph* 37 (6): 697–706.
- Zhou, Qian-Yi, and Ulrich Neumann. 2008. “Fast and extensible building modeling from airborne LiDAR data.” In *Proc ACM SIGSPATIAL Int Conf Adv Geogr Inf Syst*, 7.