# Building Detection and Reconstruction using Airborne Imagery and LiDAR Data

**Syed Ali Naqi Gilani**

A thesis submitted for the degree of *Doctor of Philosophy* at

Faculty of Information Technology,

Monash University, Australia

February 06, 2018

**Main Supervisor:** Professor Guojun Lu

**External Supervisor:** Dr. Mohammad Awrangjeb

# Copyright Notices

# Declarations

This thesis contains no material which has been accepted for the award of any other degree or diploma at any university or equivalent institution and that, to the best of my knowledge and belief, this thesis contains no material previously published or written by another person, except where due reference is made in the text of the thesis.

The ideas, development, and writing up of all the papers in the thesis were the principal responsibility of myself, the candidate, working within the *Faculty of Information Technology, Monash University, Australia* under the supervision of *Prof. Guojun Lu* and *Mohammad Awrangjeb*.

[redacted] 06/02/2018

Student's Signature Date

[redacted] 6/2/2018

Main Supervisor's Signature Date

*To my beloved mother…!*

[**1956 - 2012**]

# Acknowledgments

Alhamdulillah, all the best praises to the Almighty Allah for giving me the ability to complete this thesis successfully. The journey through Ph.D. research has been a wonderful lifetime experience and many people and events have affected me. I would never have been able to finish my dissertation without the help and encouragement of numerous people. First and most importantly, I express my heartfelt gratitude to my supervisor, Mohammad Awrangjeb, for his mentoring, advice, encouragement, spontaneous guidance, and cordial support. He has been an excellent support to me in both academic and social life.

I am grateful to Guojun Lu, my main supervisor, who provided positive thoughts and assistance with fruitful discussion to make my research smooth. The research reported here was supported by a Discovery Early Career Researcher Award by the Australian Research Council (project number DE120101778). I would like to extend my gratitude to all associated with the project. My appreciation goes to the Faculty of Information Technology and the management staff of both Monash University and Federation University Australia for providing me with the best possible facilities to conduct the research. I am also grateful to Monash University for providing scholarships and to the Australian people for their welcome.

My boundless gratitude goes to my parents and grandparents for their untiring support and love and their prayers for my success and happiness in life. Special thanks are due to my brother, sister and their families, and other relatives for their constant support and love. At the same time, I am also indebted to my wife Dr AB whose companionship and continuous untiring support have been a source of strength, inspiration, and encouragement. I would like to express my deepest gratitude to my friends Aamir Malik, Ahmad Shareef, Fasahat Siddiqui, Mahbub Khoda, Mohammad Manzur, Najam Uddin, Priyabrata Karmakar, Rajkumar Das, Shahriyar Kasiar, Zafar Iqbal and all the other friends and colleagues especially from Pakistan. In the end, an important acknowledgment goes to Dr Alex McKnight for proofreading the final draft and to the reviewers of the thesis.

# Abstract

Buildings are key components of geographic information systems, which offer support and analysis in various fields. Examples include disaster prevention and control, city and transportation planning, cadastral maps, installation of solar devices, cultural heritage documentation, and insurance. Therefore, studies on the rapid detection and reconstruction of buildings have increased over time.

Of the various remote sensing data sources, Light Detection and Ranging (LiDAR) and airborne imagery have received a good deal of attention, as they provide information about the terrain, natural and other built-up features with high accuracy and resolution. The algorithms mostly designed for building detection and three-dimensional (3-D) modelling usually fall short of producing comprehensive results, mainly because they often impose constraints on the size, height, area, and orientation of objects. Buildings and roofs small in size, in shadow or partly occluded are removed during the filtering process, and this adversely affects the detection and modelling performance. In addition, the scene heterogeneity, innumerable possibilities of building structures, and unavoidable noise due to the hardware sensors and the environment make robust modelling of the objects extremely challenging. Therefore, there is great research potential in this field to address these shortcomings of building detection and reconstruction techniques.

The research in this thesis mainly concerns the development of robust techniques in a hierarchical framework for the data-driven detection of buildings, the extraction of roof planes, and the reconstruction of building models through the use of LiDAR and aerial images. Firstly, the thesis begins with a review of the literature that allows us to identify potential research directions in the field. In the remainder of the thesis, we address several of the identified limitations in the development of the framework. Secondly, we propose a methodology to detect and regularise buildings integrating the spectral and spatial features extracted from LiDAR data and aerial images respectively.

This enables the technique to extract small, partially occluded, and shadowed buildings. These features also allow the detection technique to generate building footprints using the detected building boundaries. Thirdly, we propose a segmentation technique that provides a better interpolation of roof regions and does not apply any constraint to the shape or size parameters; therefore, it is more effective in extracting roofs and their non-occluding parts. Fourthly, the thesis also reports an automatic technique for the 3-D roof reconstruction and modelling of polyhedral buildings. The proposed technique processes datasets in data-driven fashion and reconstructs the buildings represented at lower levels with coarse boundaries (3-D roof planes) to the higher levels (3-D building models). Finally, we present an industrial application for solar potential assessment utilising information on building rooftops.

# Abbreviations

| | |
|---|---|
| LiDAR | Light Detection And Ranging |
| ALS | Airborne Laser Scanner |
| RO | Research Objective |
| RADAR | RAdio Detection And Ranging |
| SAR | Synthetic Aperture Radar |
| NIR | Near Infrared |
| NDVI | Normalized Difference Vegetation Index |
| DTM | Digital Terrain Model |
| DEM | Digital Elevation Model |
| DSM | Digital Surface Model |
| nDSM | Normalized Digital Surface Modelling |
| RMSE | Root Mean Square Error |
| DT | Delaunay Triangulation |
| PCA | Principal Component Analysis |
| RANSAC | RANdom SAmple Consensus |
| KNN | K Nearest Neighbourhood |
| MDL | Minimum Description Length |
| BSP | Binary Tree Partitioning |
| GMM | Gaussian Mixture Model |
| LoD | Level of Detail |
| SPA | Solar Positional Algorithm |
| AV | Aitkenvale |
| HB | Hervey Bay |
| EL | Eltham |
| HT | Hobart |
| VH | Vaihingen |

# Notations

| | |
|---|---|
| $C_m$ | Object completeness |
| $C_{mp}$ | Pixel completeness |
| $C_r$ | Object correctness |
| $C_{rp}$ | Pixel correctness |
| $Q_l$ | Object quality |
| $Q_{lp}$ | Pixel quality |
| $h_t$ | Height threshold |
| $h_g$ | Ground height |
| $h_{rf}$ | Height relief factor |
| $\Delta H$ | Height difference |
| $\mathcal{G}$ | Disconnected graph |
| $\mathbb{g}$ | Edge weighted graph |
| $k$ | Local neighbours of a LiDAR point |
| $D_{pt}$ | Determinant point |
| $P_{in}$ | Inside point |
| $P_{mid}$ | Mid point |
| $\mathbb{R}^3$ | Tri-dimensional space |
| $\mathbb{R}^2$ | bi-dimensional space |
| $d_{max}$ | Maximum point spacing of LiDAR data |
| $\lambda$ | Eigenvalue |
| $\vec{v}$ | Eigenvector |
| $\hat{n}$ | Direction vector |
| $w_i$ | Weight /curvature |
| $w_t$ | Weight threshold |

# Notations (continued)

| | |
|---|---|
| $r_c$ | Concentration of feature points |
| $\theta$ | Angle between two point normals |
| $\varepsilon_t$ | Plane fitting to LiDAR points error threshold |
| $\xi_t$ | Point to plane tolerance threshold |
| $\theta_t$ | Angle threshold |
| $r_u$ | Segmented to unsegmented points ratio |
| $r_p$ | Roof plane |
| $dist_p$ | Euclidean distance between the neighbouring planes |
| $I_{pnt}$ | 3D intersection point between the roof planes |
| $G$ | Gnomon point |
| $g$ | Length of the gnomon |

# Publications

- Journal Publication

  1. **Syed Ali Naqi Gilani**, M. Awrangjeb, and G. Lu, "*Segmentation of Airborne Point Cloud Data for Automatic Building Roof Extraction*", GIScience & Remote Sensing, 2018, 55 (1), 63-89. http://dx.doi.org/10.1080/15481603.2017.1361509 (**impact factor**: 3.049). *Corresponds to Chapter 5*.
  2. **Syed Ali Naqi Gilani**, M. Awrangjeb, and G. Lu, "*An Automatic Building Extraction and Regularisation Technique using LiDAR Point Cloud Data and Orthoimage*" Remote Sensing 2016, 8, 258. http://dx.doi.org/10.3390/rs8030258 (**impact factor**: 3.244). *Corresponds to Chapter 4*.

- Conference Publication

  1. **Syed Ali Naqi Gilani**, M. Awrangjeb, and G. Lu, "*Robust Building Roof Segmentation using Airborne Point Cloud Data*" IEEE International Conference on Image Processing (2016): 859–863. https://doi.org/10.1109/ICIP.2016.7532479. *Corresponds to Chapter 5*.
  2. **Syed Ali Naqi Gilani**, M. Awrangjeb and G. Lu, "*Fusion of LiDAR data and Multispectral Imagery for Effective Building Detection Based on Graph and Connected Component Analysis*," International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 1 (2015): 65–72. http://dx.doi.org/10.5194/isprsarchives-XL-3-W2-65-2015. *Corresponds to Chapter 4*.

- Planned Journal Publication

  1. **Syed Ali Naqi Gilani**, M. Awrangjeb, and G. Lu, "*A Robust Data-Driven approach for Polyhedral Scene Reconstruction*". *Corresponds to Chapter 6*.
  2. **Syed Ali Naqi Gilani**, M. Awrangjeb, G. Lu, N. Quadros, and P. Delaney, "*Automatic Assessment and Visualisation of Rooftop Solar Potential Using Remote Sensing data*". *Corresponds to Chapter 7*.

# Contents

# List of Figures

# List of Tables

Chapter 1

# Introduction

> *"The beginning is the most important part of the work."*
>
> Plato

In this chapter we provide an introduction to the work presented in this thesis. We describe the research problem in Section 1.1, the research challenges in Section 1.2, the research objectives in Section 1.3, and finally the contribution of this research in Section 1.4.

## 1.1 Problem Statement

Building being a key urban object is indispensable to various Geographic Information System (GIS) applications. At first glance, building appears to be a simple object on the earth's surface that can easily be identified. However, in reality, detection of the buildings, roof planes, and their 3-D city form are quite challenging, due to differences in the viewpoint, the surrounding environment, and structural complexities. The aim of the detection of building and roof planes is to identify the area of a building region and its constituent roof surfaces, respectively. Modelling subsequently uses these geometrical parameters for the development of an equivalent computer-based representation of the buildings. According to a survey by the European Organization for Experimental Photogrammetric Research [1], more than 90% of participants expressed their interest in the construction of realistic 3-D city models that include information about the buildings, the traffic network, and the vegetation.

Geospatial technologies have long been proved to be reliable and efficient tools for understanding the urbanisation process and generating building models [2, 3]. However, traditional methods, such as ground surveys, intrinsically degrade the automation and processing speed [4]. Meanwhile, the emergence of spatial data acquisition technologies has made various types of data, such as aerial and oblique images, LiDAR data and terrestrial laser scanning data, available for 3-D characterisation of the earth's surface and the objects on it [5].

Beginning with a single data type, either aerial imagery or LiDAR data, modelling of buildings is performed in two interdependent tasks i.e., detection and reconstruction [6–9], in which the accuracy of modelling is arguably subject to the reliability of the detection process. The development of high spectral imagery, which captures image data within several wavelength ranges across the electromagnetic spectrum, makes it possible to sense individual buildings in an urban scenario [10]. However, the increasing spectral band ranges and textural information do not warrant a proportional increase in building detection and modelling accuracy [11, 12], but rather adversely increase classification ambiguities [13]. Consequently, similar objects (e.g., green and red trees) may appear with different spectral signatures, whereas different objects (e.g., a green tree and a green roof) may appear to have similar spectral signatures under various background conditions [14]. These factors, together with sensory noises, reduce the building detection rate and increase the modelling error. Therefore, the use of spectral information alone to differentiate these objects eventually results in poor performance [10].

LiDAR data, on the other hand, provide height information for objects, such as buildings, trees, bushes, terrain, and other 3-D objects. Therefore, adopting height variation to distinguish these urban objects is a more suitable cue than spectral and texture changes. However, the accuracy of the detected boundaries is often compromised due to point cloud sparsity, and this reduces planimetric accuracy [12, 14]. In addition, trees and buildings in complex urban scenes sometimes appear to have similar height variations [15], and height information alone is not perceived to produce a finer classification [11]. Therefore, several researchers have developed the consensus to use multisource data in designing better strategies to increase the building detection rate and the accuracy of the reconstructed models [12, 14–17].

The acquisition and modelling of buildings include not only detection but also several non-trivial processes such as segmentation, classification, structuring, hypothesis generation, and geometric modelling. In fact, the seamless integration of these processes to acquire the required 3-D information in a conventional way would be not only unrealistic but also laborious. We, therefore, present a solution for 3-D building modelling as a framework that performs building detection, roof plane extraction, and 3-D model generation of variably-shaped complex structures separately. Moreover, we have developed an automated system for industrial use for large-scale assessment of solar potential by leveraging the proposed framework. To achieve our goals, photogrammetric imagery and LiDAR data are used addressing the particular challenges and research questions presented in the following section.

## 1.2 Research Challenges

Urban areas are continuously changing as a result of construction and extension with complex building structures such as dormers, hedges, and ventilation systems. According to the United Nations, the world's population living in urban regions is projected to increase from 54% to 66% by the year 2050 [18]. Another study reports that the building modelling and mapping markets are expected to expand from $1.1 billion to $7.7 billion in a short period of 5 years beginning from 2013 [19]. A more refreshed construction of 3-D building models is therefore imperative in order to contend with the issues of growing urban agglomeration and city management.

With the advent of advanced technology and fast computational power, several efforts have been made over the last two decades to achieve automatic building modelling, showing various degrees of success [20–23], but usually failing to produce promising results, mostly because of the intrinsic characteristics and the large size of the input data. These systems fall short in extracting buildings with unusual geometrical structures and occlusion by dense vegetation or shadows. Therefore, there is great research potential for robust detection and modelling tasks because of difficulties created by scene heterogeneity in appearance, unavoidable noise due to the environment, terrain complexities, and the indefinite possibilities of different structures. The research challenges addressed in this thesis are:

- **Occlusions** and **shadows** by surrounding objects are major issues, often causing object recognition processes to fail or at least, increasing the misclassification rate. They usually reduce the accuracy of the process in hypothesising the major structures of the buildings. Figure 1.1 shows complexity scenarios where buildings are partially occluded and/or shadowed by surrounding trees.

**Figure 1.1** Occluded and shadowed building samples.

- Most strategies either neglect or fall short in the detection of buildings of **small size** and **low height**. The problem is aggravated when several small structures form a block of buildings and identifying a discrimination element automatically becomes challenging for modelling purpose. In addition, **noise** and **outlier** data introduce the issue of over-segmentation, which at the later stages causes ambiguity in building geometric modelling. Example scenarios are sketched in Figure 1.2.

- Building **shape heterogeneity** and **complex structures** prohibit the use of specific object models for the 3-D representation of all possible buildings in an area. Therefore, many researchers, as alternatives, use different constraints on the geometric regularity of the building structures that prevent the detection and modelling of several structures which do not satisfy the conditions perfectly. Figure 1.3 shows sketches of possible structures and the heterogeneity of buildings that make the modelling of buildings and their constituent roof planes a challenging task.

- The problem of **misregistration** for the integration of LiDAR and aerial imagery has a profound impact on the robustness of a detection and modelling procedure. It is challenging to identify and extract the correct features and combine

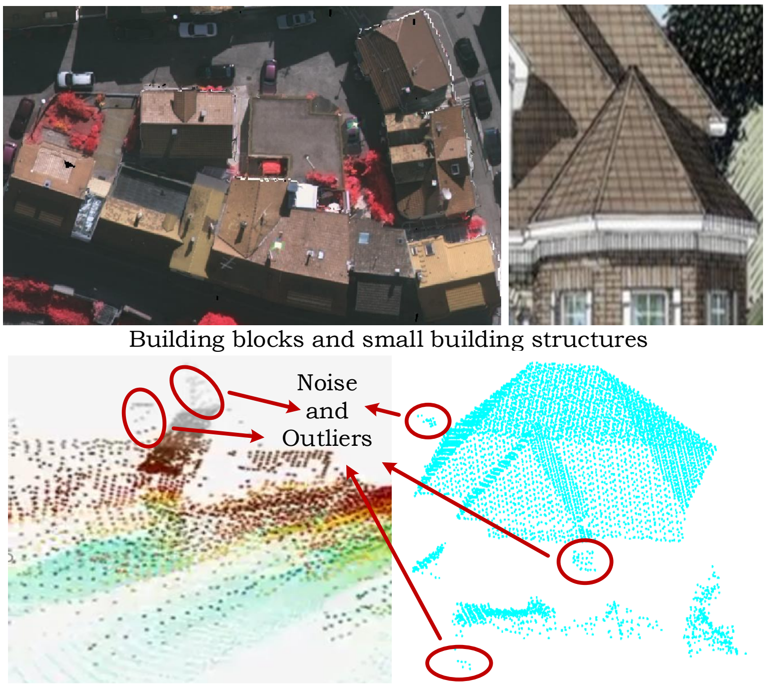



**Figure 1.2** Sample small structures, and noise and outlier points in LiDAR data.

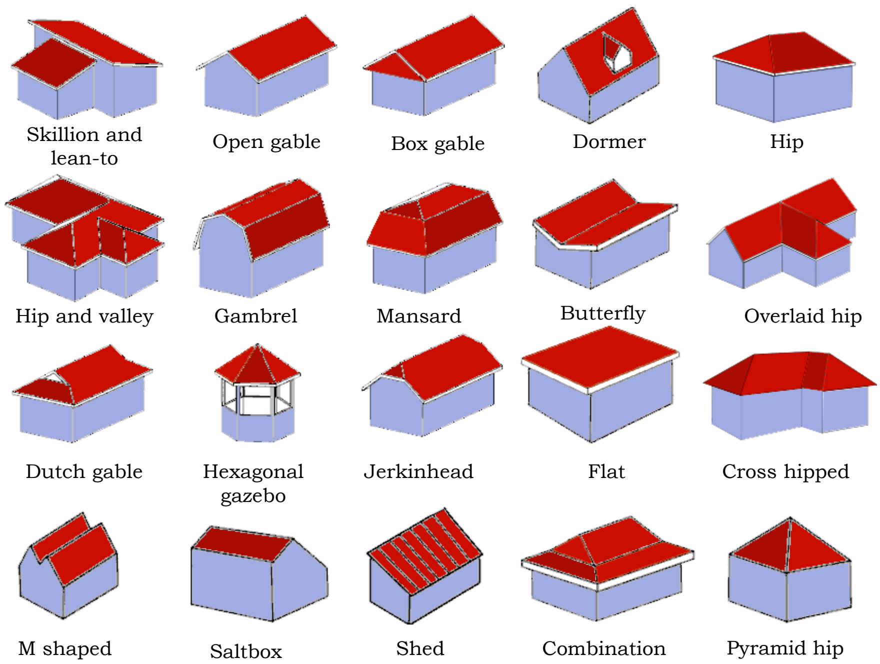



**Figure 1.3** Some possible different arrangements of building rooftops.

them such that their strengths can be exploited to achieve a high detection rate and reduce modelling errors. Figure 1.4 shows misregistration/misalignment between the LiDAR points and the corresponding building in the aerial images.

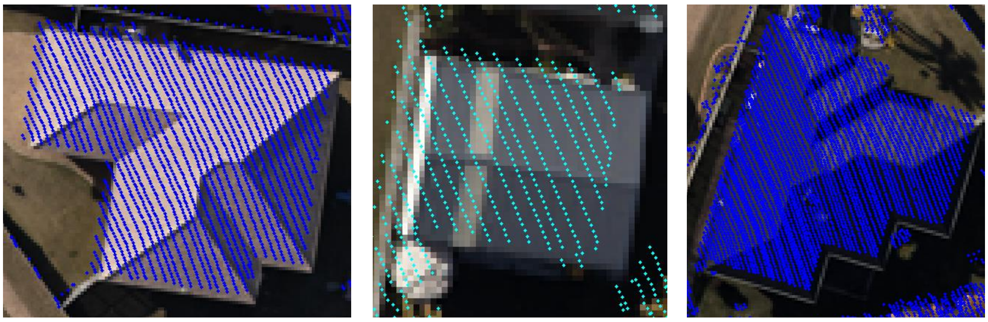


**Figure 1.4** Misregistration between LiDAR and the corresponding aerial images when the LiDAR points are overlaid on the corresponding buildings showing a moderate offset from their actual location in the aerial image.

## 1.3 Research Objectives

On the basis of the discussion in Sections 1.1 and 1.2, the main question this research aims to answer is: **How can the effectiveness of building detection and 3-D modelling using LiDAR and aerial imagery be improved?** To help answer this question, this thesis addresses the following list of research objectives (ROs):

1. **RO1**: To identify the research gaps and limitations, and conduct a review of existing approaches to building detection, roof plane extraction, and 3-D building modelling.

2. **RO2**: To develop a technique which can handle moderate misregistration errors for the automatic detection of buildings which have variable shapes or sizes or can be partially occluded or in shadow.

3. **RO3**: To develop a roof detection technique which can extract roof planes and small structures, and can handle occlusion, noise and moderately corrupt data.

4. **RO4**: To develop a modelling technique that can use the extracted roof planes to develop 3-D models for heterogeneous polyhedral buildings.

5. **RO5**: To develop an application using roof geometric information for the large-scale assessment of solar potential.

## 1.4 Thesis Contribution and Structure

To achieve the objectives, the thesis is organised into eight chapters. Each chapter starts with an introduction and highlights the main contributions it makes. The structure of the thesis is sketched in Figure 1.5.

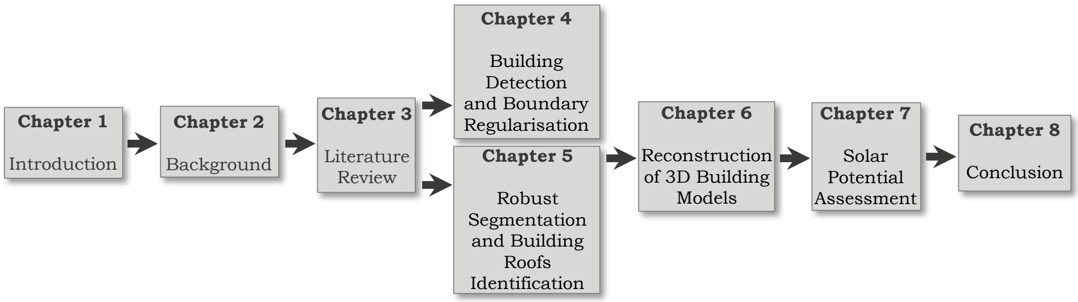



**Figure 1.5** Thesis organisation.

The remainder of this thesis is structured as follows:

- **Chapter 2** provides background information on key remote sensing technologies and data sources, the mathematical concepts of the performance evaluation systems, and benchmark datasets.
- **Chapter 3** addresses **RO1** by presenting an extensive literature review of building detection, roof plane extraction, and 3-D building modelling. The review process allows us to identify potential research gaps and directions in the field of detection and reconstruction of urban objects.
- **Chapter 4** focuses on **RO2** and introduces a technique for building area detection and footprint generation. The proposed technique integrates LiDAR data and aerial imagery to eliminate vegetation, extract partly-occluded buildings, and generate building outlines. The synthesising approach to the use of features from both data sources enables the proposed technique to extract not only small and shadowed buildings but also building regions where LiDAR data have large sparsity. The proposed method offers high detection rates, even in the presence of a moderate registration error between the input data sources.

- **Chapter 5** addresses **RO3** and proposes an effective system for the recognition of roof planes. Since the availability of aerial imagery for particular LiDAR data can never be guaranteed, instead of using the output of the previous chapter, the proposed system utilises only LiDAR data for the extraction of the roof planes and their constituent primitives e.g., dormers, chimneys, vents, and the building regions. The use of robust saliency features enables the technique to handle noise and eliminate vegetation to extract roofs as well as their partially occluded parts from complex scenes with a high success rate. A new LiDAR-based boundary tracing algorithm is also proposed, which seamlessly extracts the inner and the outer boundaries of an object without imposing any geometric constraint.

- **Chapter 6** provides an automatic technique for 3-D modelling of polyhedral buildings in order to address **RO4**. The proposed technique uses the extracted roof planes (from Chapter 5) as the only information for the construction of building models. This chapter also introduces a robust procedure to approximate the missing roof planes that were not extracted from LiDAR previously due to low point density or missing data. The proposed technique develops an interrelation among the building roof planes and identifies their interconnections, which are later used for the reconstruction of 3-D building models.

- **Chapter 7** focuses on **RO5**. This chapter aims to develop an application for the installation of photovoltaic (PV) systems for a solar potential assessment project at the Collaborative Research Centre for Spatial Information (CRCSI)[1]. For project development, building roofs and other geometrical information produced in Chapters 5 and 6 are leveraged for the effective assessment of solar PV deployment for both local and state governments and solar energy companies in Australia.

- **Chapter 8** concludes the thesis with a summary of the achievements and suggestions for future research directions.

[1] http://www.crcsi.com.au/

Chapter 2

# Background

*"Success is the result of perfection, hardwork, learning from failure, loyalty, and persistence."*

Collin Powell

## 2.1 Introduction

The aim of this chapter is to present background information for the research work described in this thesis. We provide descriptions of key technologies, data sources, mathematical concepts of the performance evaluation systems, and benchmark datasets and their characteristics. Section 2.2 provides an overview of different input data sources for building detection, roof extraction, and 3-D building reconstruction. Section 2.3 presents a brief discussion of different basic principles that are used in this thesis. A summary of different evaluation systems is provided in Section 2.4, and the metrics to measure the performance of the proposed methodologies are discussed in Section 2.5. The benchmark datasets are introduced in Section 2.6 and Section 2.7 concludes the chapter.

## 2.2 Remote Sensing Data Sources

Remote sensing is a technique for acquiring information about objects without physical contact. The basic principle is to detect and measure the electromagnetic radiations of different wavelengths reflected or emitted from features on the surface, in the atmosphere or in the ocean by which they may be identified and categorised. This technique has been applied to numerous disciplines, including geography, land surveying, and the earth sciences. Remote sensing, in its current usage, is performed through satellite- or aircraft-based platforms equipped with various sensors operating at specific wavelengths. It can be an active or passive remote, depending on the sensors. In active remote sensing, the signals are emitted by a satellite or aircraft and the energy reflected from objects is detected by the sensors. Radio detection and ranging (RADAR) and LiDAR are examples where the time difference between emission and return is measured, and the location and direction of an object are determined. In contrast, a sensor in passive remote sensing records the energy emitted from different objects [24–26]. Examples include film photography, infrared, charge-coupled devices, and radiometers [27]. The difference between active and passive remote sensing is illustrated in Figure 2.1.

The most common remote sensing data types used for building detection, roof extraction, and 3-D reconstruction are image and LiDAR data [28, 29]. Many researchers have tried using image information [30–33], while others have simply used the LiDAR data in different applications for urban monitoring and planning [34–37]. Recently some researchers have tried integrating both types of data to perform detection and reconstruction activities [32, 38–44]. In this thesis, Chapter 3 provides an extensive literature review on building detection, roof plane extraction, and 3-D building modelling strategies. However, this chapter briefly discusses the most commonly used technologies involved in object extraction and its representation.

### 2.2.1 Airborne Imagery

The American Society for Photogrammetry and Remote Sensing (ASPRS) defines photogrammetry as “the art, science, and technology of obtaining reliable information about physical objects and the environment through the process of recording, measuring, and interpreting photographic images and patterns of recorded radiant electromagnetic energy and other phenomena” [45]. Photogrammetry has been the traditional way of generating 3-D information and still is a major data source for GIS building detection and reconstruction applications.

In mainstream usage, two types of remotely-sensed images frequently used are satellite imagery and aerial imagery. A satellite imagery is captured by a satellite sys-

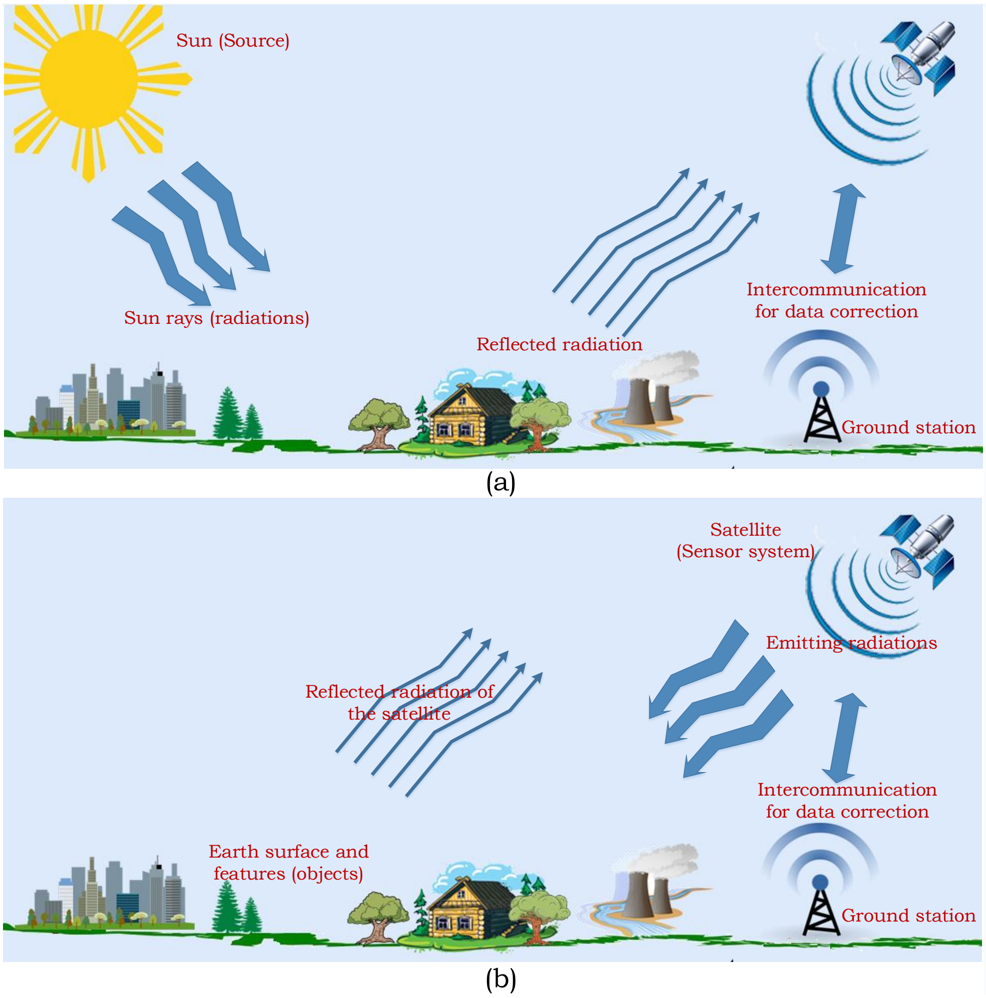



**Figure 2.1** Illustration of remote sensing: (a) Passive remote sensing and (b) Active remote sensing.

tem that is distant from the real objects [46]. In contrast, an aerial vehicle flying at low height is utilised to capture the real objects in aerial imagery [47]. There are a number of imagery sources, and the choice of which imagery to be used is entirely dependent upon the context of the problem. To provide an in-depth survey based on all the available sensors and their data is nearly impossible, due to the sheer number of sensors available. Instead, we explore those most frequently used in the literature for building identification and its related GIS applications.

Most often, aerial orthoimagery, which is simply a remotely-sensed image that has been geometrically corrected, is used to measure the true distance of objects on the earth in various geomorphological and earth science applications. Of the satellite imagery systems, *Synthetic Aperture Radar* (SAR), QuickBird, WorldView, and GeoEye-1 have been used by researchers for building detection, roof extraction, and 3-D building reconstruction purposes [32, 48]. A well-known aerial image used extensively in the literature is *multispectral* image [15, 49]. The following section provides details of multispectral imagery and its associated advantages and disadvantages.

#### 2.2.1.1 Multispectral Imagery

Multispectral sensors capture spectral information from multiple discrete bands of the electromagnetic spectrum. They sense radiations from multiple wavelength regions of the visible, near infrared, middle infrared and thermal infrared bands and transform them into a digital image known as a multispectral image [50]. A multispectral image, or simply a multi-band image, is a colour image consisting of a red, a green and a blue band, each taken with a sensor sensitive to a different wavelength. Multispectral images are most commonly-used for remote sensing applications such as building detection, building change detection, flood monitoring, and forest mapping. The technique to capture multispectral imagery is shown graphically in Figure 2.2.

Multispectral imagery has several advantages over satellite imagery. For example, since it has high spatial resolution (number of pixels), processing the images and handling objects become quite easy. It also offers high location accuracy [52] and describes various urban attributes with different colour bands e.g., healthy vegetation appears in shades of red; water shows up almost black; and concrete and gravel appear in shades of grey. Therefore, multispectral images are considered a strong candidate for a variety of observation and monitoring applications e.g., urban objects, forestry, and plant health. More details on discriminating these objects is provided later in the thesis, specifically in the context of vegetation elimination as it is one of the key subjects being addressed here.

There are also some limitations with the use of multispectral imagery. It has low area coverage and high cost per unit area compared with satellite imagery. Often, weather conditions and sun-light cause delays in data acquisition. In addition, the accuracy of an object extraction procedure is often affected by the variability and visibility of the object's structure when multispectral images are used [53]. The variability

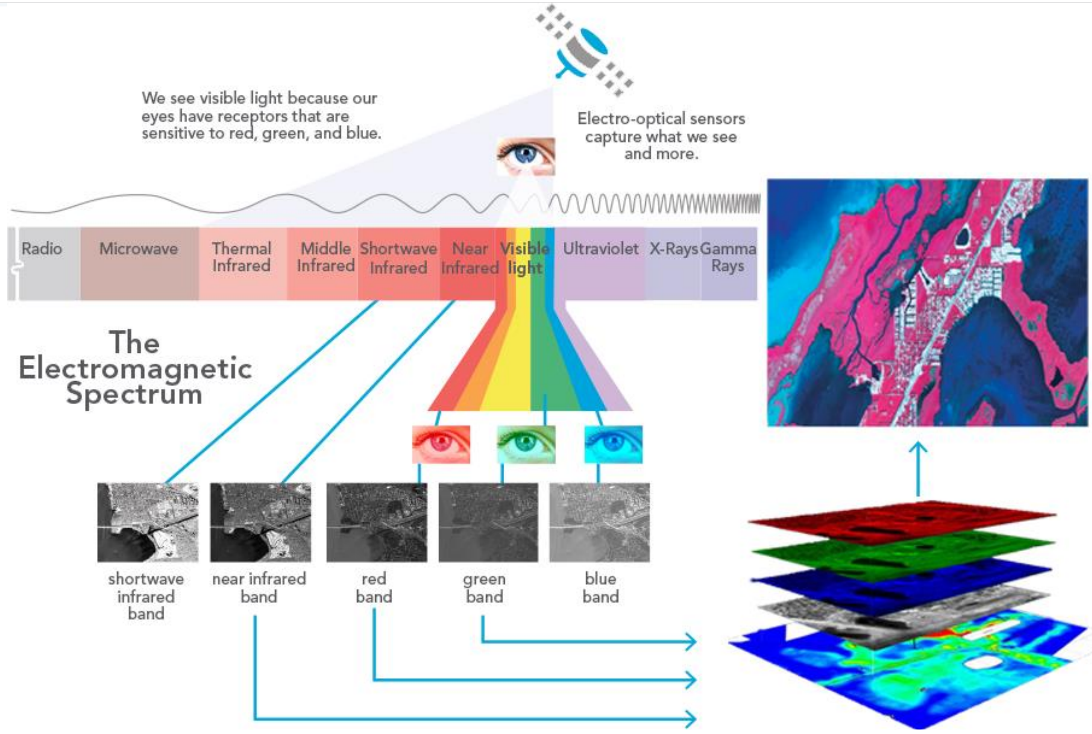



**Figure 2.2** Production of multispectral imagery from the electromagnetic spectrum [51] (best seen in colour).

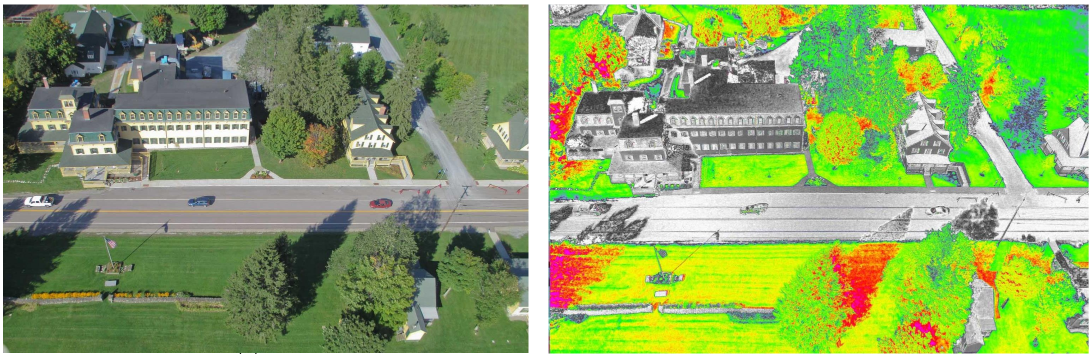
aspect of an object deals with the nature of the content e.g., density (rural or urban territories), architectural complexity (residential, industrial or commercial areas), terrain (flat or hilly), and vegetation (low, moderate or heavy), whereas object structure visibility addresses the resolution, quality of the contents (blurriness or sharpness) and view of an object (visible, shadowed, or occluded).

#### 2.2.1.2 Derived data - Normalized Difference Vegetation Index

Multispectral images provide very interesting analysis through the use of different indices. A multispectral image comprises of green (555–580 nm), red (665–700 nm) and near infrared (NIR, 740–900 nm) bands, which assist the observation of vegetation mapping and plant health. The chlorophyll in plants basically absorbs the NIR wavelength of light that is more for a healthier plant than a dried one. The NIR camera detects this difference and calculates different vegetation indices. The *Normalized Difference Vegetation Index* (NDVI) is an index which is extensively used in the research literature for the elimination of vegetation and the classification of different objects [14, 17, 42, 49, 54, 55]. It is calculated using Equation 2.1 [42, 55]:

$$NDVI = \frac{(NIR - Red)}{(NIR + Red)} \tag{2.1}$$

A high value of the NDVI for a pixel indicates the presence of plants. For vegetation analysis, different NDVI thresholds are used simultaneously for vigorous plants with green leaves and plants with red leaves. Figure 2.3 shows the original image and the corresponding NDVI image. The images have been slightly modified for the purposes of the thesis based on [56].



(a) (b)

**Figure 2.3** An oblique photo pair: (a) Original image and (b) The corresponding NDVI image (best seen in colour).

### 2.2.2 Airborne Laser Scanning Data

Airborne LiDAR has gained popularity among spatial researchers and mapping professionals due to its speed, precision, and accuracy in capturing 3-D geo-referenced spatial information about buildings, roads, vegetation, and other objects expediently at a high point density. These characteristics make it feasible to examine natural and built environments across a wide range of scales for the automatic extraction and reconstruction of buildings and their distinct features. Figure 2.4(a) and Figure 2.4(b) show the LiDAR data and the corresponding aerial imagery of a site in Australia, respectively.

A complete LiDAR system is made up of several components which work together to generate, record, and geo-reference the data. The components include a *Light amplification by stimulated emission radiation* (Laser) unit, a *Global Positioning System* (GPS),

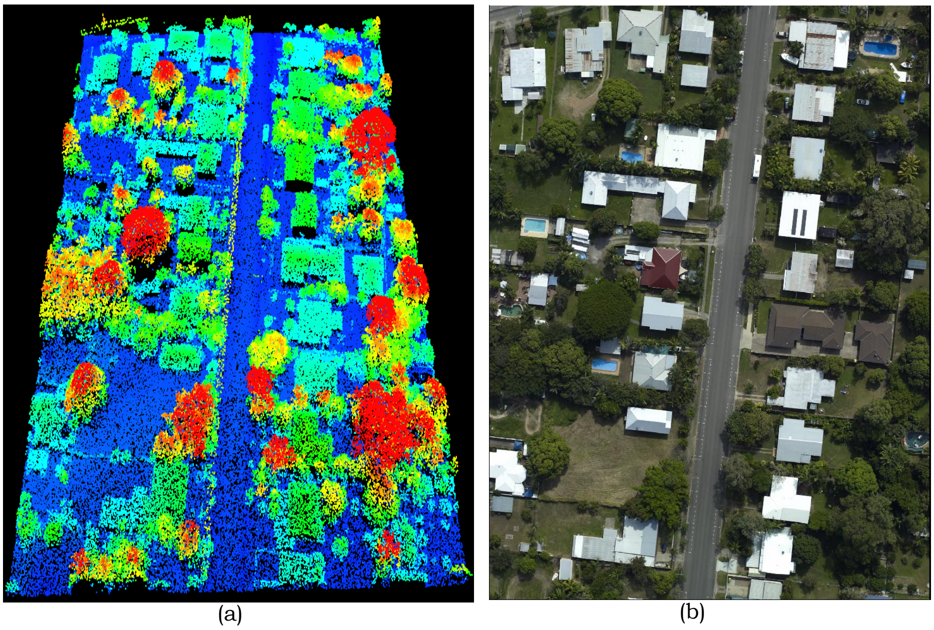



**Figure 2.4** (a) LiDAR data over an area where different colours correspond to different height values and (b) The corresponding aerial imagery.

an *Inertial Measurement Unit* (IMU), and a general-purpose computer. The laser unit generates energy in discrete pulses and captures the returned energy, the GPS unit records the precise location of the scanner, and the IMU sensor measures the velocity, orientation and gravitational forces. A LiDAR system also utilises a fixed ground base station to correct and improve the data collected by the sensors. Moreover, a LiDAR system employs a computer unit for the integration of data received from the laser system, the GPS and the IMU to produce LiDAR point cloud data. Of the cost-effective platforms, airplanes and helicopters are the most commonly-used carriers for acquiring LiDAR data over large areas. A particular arrangement of a LiDAR system which operates in airborne vehicles is shown in Figure 2.5.

LiDAR uses electromagnetic energy to detect various GIS objects and determine their height values. Each time the laser hits a feature on the earth, it generates a return. Each return carries the 3-D coordinates (x, y, and z or latitude, longitude, and elevation) of the objects which are measured from 1) the time difference between the emitted and returned laser pulses, 2) the angle at which the pulses were “fired”, and 3) the absolute location of the sensor on or above the surface of the earth. In the case of buildings and the ground, we have a single return, but in the case of trees, which

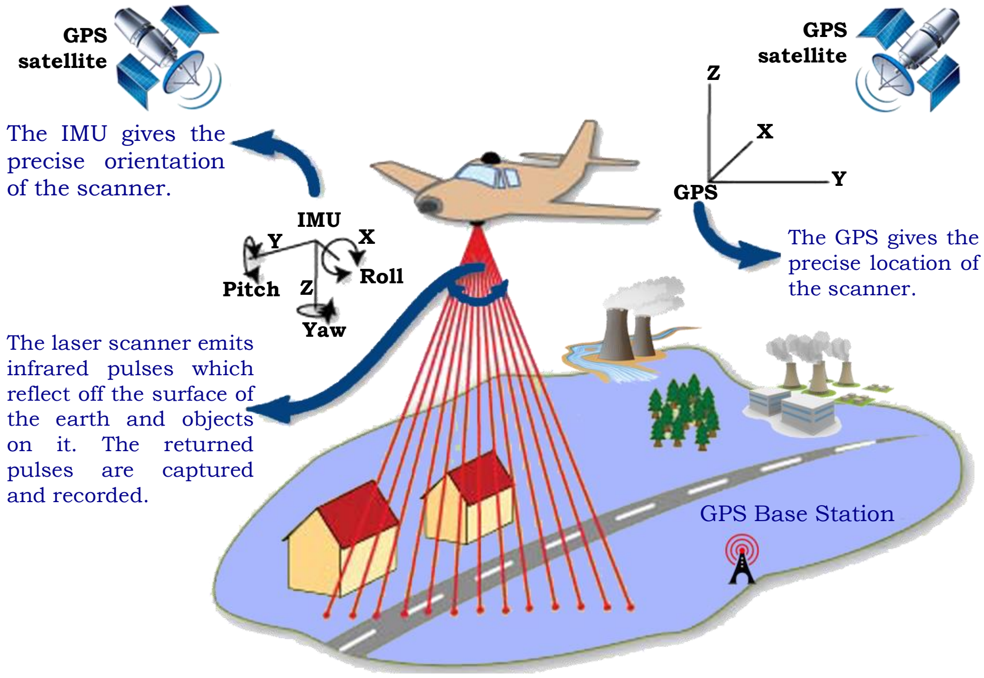



**Figure 2.5** Schematic diagram of an airborne LiDAR data collection.

are not dense features, we may have multiple laser returns from different heights. However, LiDAR systems have some limitations. For example, the laser pulses cannot penetrate through clouds and dense haze. It is therefore recommended to record the data during fair weather conditions and this is the reason why flight missions are often carried out at night.

#### 2.2.2.1 Derived data - DEM, DSM, and DTM

LiDAR data allow us to generate models of the earth's surface. The general term for a model of the elevation of an area is a *Digital Elevation Model* (DEM). DEMs come in two common types: *Digital Surface Models* (DSMs) and *Digital Terrain Models* (DTMs). High quality DSMs are the direct products of laser pulses reflected off the earth's surface or the tops of man-made or natural objects. More simply, DSMs represent the earth's surface and all the objects on it. With minute or almost no post-processing, a DSM can provide 3-D landscape models for fly- or walk-through visualisation, landscape design, and computer game production. In contrast, the DTM describes only the bare ground surface without any objects such as trees and man-made objects. Figure 2.6 illustrates the difference between a DSM and DTM using a red solid line with blue circles to represent LiDAR discrete points.

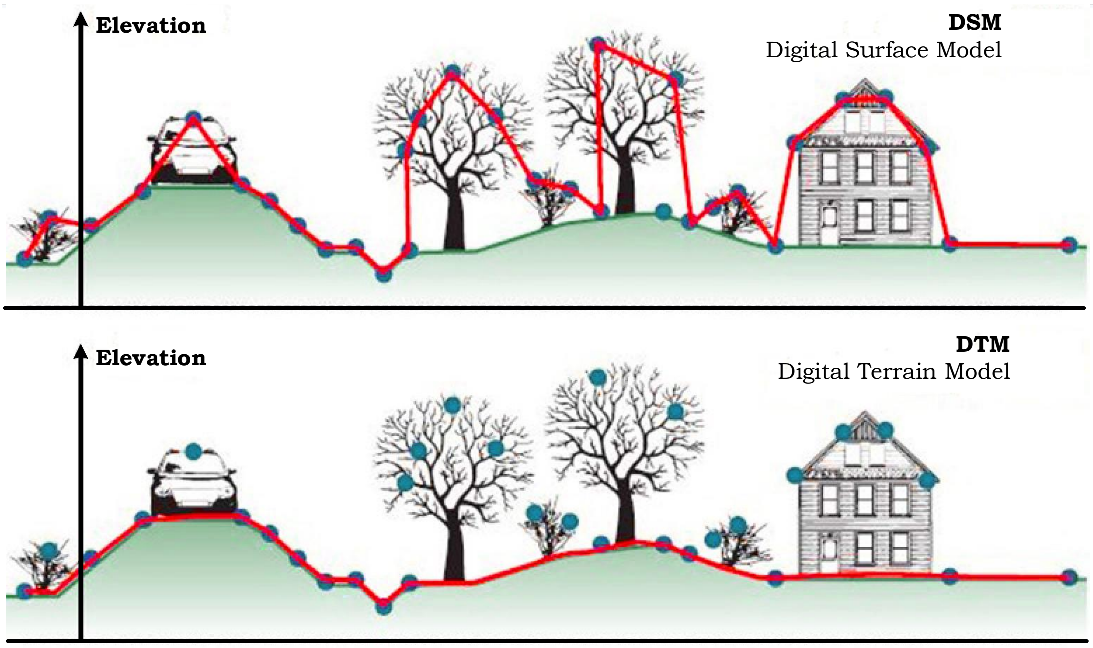



**Figure 2.6** Surfaces represented by DSM and DTM [57].

The term DEM often loosely refers to both DSM and DTM in the literature. DEM is a bare earth (topology) model of the earth's surface generated by removing the LiDAR points reflected off non-ground objects. When using a DEM, the reference surface must always be stated. Alternatively, it can be generated using stereo photogrammetry or land survey techniques [58]. The usage of DEMs as principal data is widespread for supporting applications in both the engineering and scientific fields, such as geomorphological analysis, flood-risk mapping, oil and gas exploration, and real-estate development. To a certain extent, city modelling and urban reconstruction could be considered as a subset problem of DEM production.

### 2.2.3 Characteristics of LiDAR

This section presents the characteristics of LiDAR data that are important to several applications which involve LiDAR point cloud processing. Of its various properties, point spacing, average point density, and accuracy are of common interest; therefore, these factors are briefly discussed in the following sections.

#### 2.2.3.1 Point Distribution and LiDAR Density

LiDAR point distribution or point spacing and LiDAR density are amongst the most frequently used metrics to describe the quality of the point cloud data in the literature [59, 60]. LiDAR point distribution refers to the point-to-point distance of the data and describes how close the laser points are to each other. It is considered analogous to the pixel size of an aerial image [61]; however, the spacing value is not uniform, unlike image pixels, as the points of the LiDAR data are sparse and highly dispersed. Alternatively, the nominal spacing value is generally regarded as the best representation of point distribution, which is the average spacing of the unorganised points in both x and y directions [60, 62].

LiDAR point density, on the other hand, refers to the number of points per unit area [60]. This is typically measured by first using a square grid on the LiDAR data and then determining the average number of points per cell, which is taken as the point density of the data. The greater the number of points per grid cell, the denser the dataset is, and finer the quality of the reconstructed objects which can be achieved. In contrast, several features, like gullies, mounds, chimneys, slope changes, and depressions are often obscured in a low-resolution dataset, and the resultant building models therefore have low quality and high reconstruction errors [61]. The point densities of the datasets we use in the present thesis vary between 1 point/$m^2$ to 40 points/$m^2$. In terms of point distribution, these datasets have point spacings varying between 0.2 m to 1.2 m. More details on the benchmark datasets are provided in the following section.

#### 2.2.3.2 Accuracy

Photogrammetric elevation generation is a time-consuming and labour-intensive process, especially for high-accuracy products. According to Carter et al. [61], determining the required level of data accuracy and the level achieved are important parts of a data collection procedure and its subsequent usage. Often, the regions hidden by or below trees have low elevation accuracy and this requires such locations to be visible from other images by a photogrammetric elevation generation process. In contrast, LiDAR has not only a similar cost but also captures the topographic mappings for large areas, including hidden locations rapidly and with a higher level of accuracy [61, 63].

The vertical accuracy of LiDAR data has improved over time with the increasing sophistication of the technology. Now the LiDAR datasets have *vertical Root Mean Square Error* (*RMSE*) values of less than 20 centimeters (8 inches) [61], which adds great value to applications for building detection, roof extraction, and GIS object modelling. However, the horizontal resolution refers to the point spacing concept of a dataset and directly affects the planimetric accuracy of extracted objects.

### 2.2.4 Which is better?

To date, there has not been a single authoritative study to indicate which technology is better between imagery- and LiDAR-based systems. Over the last decade, research in the field of photogrammetry and remote sensing clearly shows that the usage of LiDAR has increased considerably for practical and industrial applications. However, this does not guarantee the advantages of LiDAR over the competing technologies. In this regard, Leberl et al. [64] carried out two separate studies comparing the performance of point clouds produced from airborne- and ground-based LiDAR systems with those generated using optical images.

Their research findings demonstrate that the accuracy achieved by using the photogrammetric technique is comparable with that of the LiDAR-based technique. The authors identified fifteen additional advantages in order to support photogrammetric methods despite the fact of tremendous increase in the use of LiDAR data. Other advantages and disadvantages and the selection of a particular data type are discussed in the following sections while addressing the challenges and describing the proposed methodologies.

## 2.3 Basic Principles of the Proposed Techniques

This section provides a brief review of some basic principles used in the areas of building detection, roof extraction, and 3-D building reconstruction.

### 2.3.1 Delaunay Triangulation

LiDAR data has a collection of laser points which are spatially unorganised and have variable point density. LiDAR point clouds often have gaps due to occlusion by neighbouring objects, e.g., vegetation clusters [65]. These points do not have any connection information. In fact, the main reason for using the *Delaunay triangulation* (DT) is to solve the problem of the spatial adjacency of disconnected points. According to [66], "a DT for a given set $P$ of discrete points in a plane is a triangulation DT($P$) such that no point in $P$ is inside the circumcircle of any triangle in DT($P$). DTs maximise the minimum angle of all the angles of the triangles in the triangulation; they tend to avoid long/thin shapes".

The authors [66] further state that "for a set of points on the same line there is no DT (the notion of triangulation is degenerate for this case). For four or more points on the same circle (e.g., the vertices of a rectangle) the DT is not unique: each of the two possible triangulations that split the quadrangle into two triangles satisfies the Delaunay condition, i.e., the requirement that the circumcircles of all triangles have empty interiors. The circumcircle of a triangle actually is the unique circle passing through the three vertices of the triangle".

For a set of points in 2-dimensions (2-D) or 3-D, a DT of these points ensures the circumcircle associated with each triangle contains no other point in its interior. Once a DT is created, a variety of topological and geometric queries can be performed, such as surface interpolation, determining the density or intensity of point sampling, locating a facet that contains a specific point, find the nearest neighbours of a particular point, and generating a convex hull. Figure 2.7 shows 2-D and 3-D DTs using 30 random points. In this thesis, we utilise a variant of 2-D DT called constrained Delaunay triangulation for the development of a building region detection technique and a boundary outline tracing algorithm, which are described in Section 5.2.1.

### 2.3.2 Principal Component Analysis

Principal Component Analysis (PCA) is one of the most widely-used multidimensional statistical technique for dimension reduction and data visualisation [67]. PCA begins

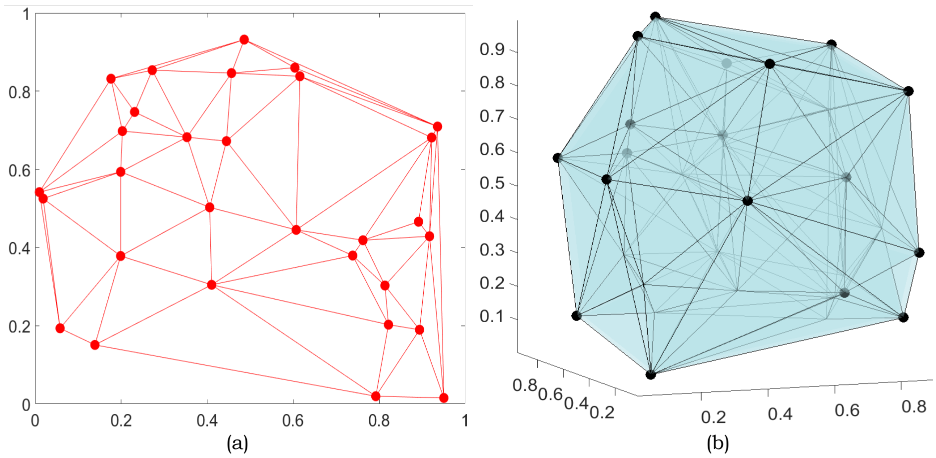



**Figure 2.7** Illustration of Delaunay Triangulation with 30 points in (a) 2-D and (b) 3-D spaces.

to find directions of maximum or minimum variability in the data space, and tries to represent the data across orthonormal axes with maximum de-correlation. Through the transformation process, PCA generates new set of uncorrelated and orthogonal variables that can explain the underlying covariance structure of the data. The new set of variables, the Principal Components (PCs), are the linear combinations of the mean-centred original variables that rank the variability in the data through the variances and produces the corresponding directions using the eigenvectors of the covariance matrix [68, 69]. The PCs are usually ranked in descending order, explaining the underlying data variability of the corresponding eigenvalues.

In LiDAR data processing, the first two PCs serve the orthogonal basis of the 3-D plane, whilst the third PC, being orthogonal to the first two, corresponds to a normal of the fitted plane and is used to compute saliency features (e.g., surface normal and slope). Figure 2.8 shows a plane-fitting illustration in a sample point cloud. LiDAR points are shown using blue and red points with the plane-fitting residuals shown graphically using lines of the same colours, respectively. The variations in the data are shown with arrows where the arrow towards to the least variation corresponding to plane normal.

### 2.3.3 Region Growing

Urban scenes are characterised by the existence of diverse objects, such as buildings, trees, bridges, and road infrastructures, offering a high degree of complexity. In order

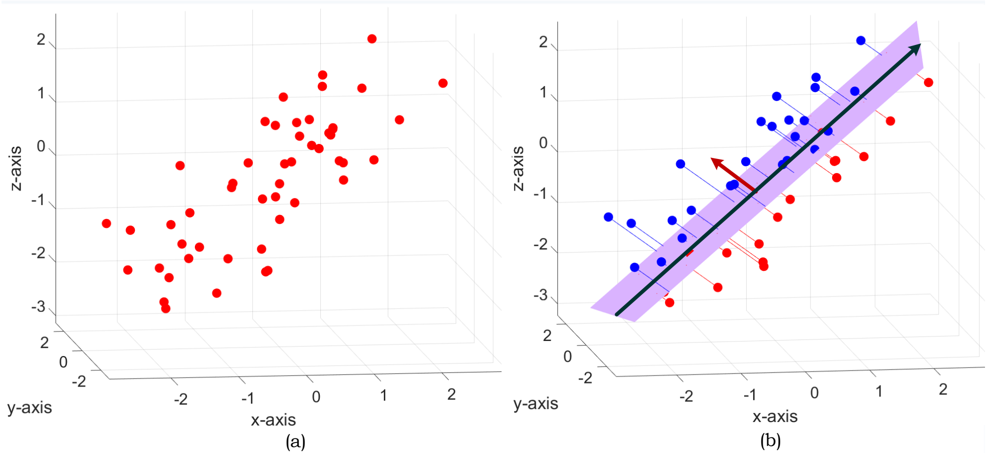



**Figure 2.8** Plane-fitting demonstration using PCA: (a) Sample point cloud and (b) Plane-fitting, residuals, and PCs showing directions of data variations in their respective directions.

to identify these underlying objects from any input data source, segmentation is performed, which is a process of separating and labelling the most similar features into a number of separate surfaces. Region growing in reality is an approach for performing the segmentation task.

Region growing was originally introduced in the context of the region-based image segmentation method. The process examines the neighbouring pixels of initial seed points and determines whether the pixel neighbours should be added to the region based on some similarity e.g., intensity, colour, or texture information. The process continues in the same manner and segments the image into different regions. Due to its simplicity, this approach has been adopted for data clustering algorithms in other disciplines, including LiDAR systems.

We explain the region-growing approach in terms of segregating the LiDAR points based on some similarity criteria. Since LiDAR data points do not have any statistical distributional pattern in the data and provide no connection information, we approximate saliency features, namely point normal and slope, using the PCA. We then use them as the similarity criteria for segmentation of LiDAR points into different underlying regions. Figure 2.9 shows the region growing-based segmented LiDAR points. The region-growing process generally has the following stages:

- Start by choosing a seed point (often the one with the least variation)

- A region is grown from a seed point by adding the neighbouring LiDAR points based on the similarity criteria. This results in the region expanding in size.
- The growth of one region stops when no more points can be added to the region. Another seed point is then chosen from the unsegmented points.
- For the remaining unsegmented points, repeat the previous steps.

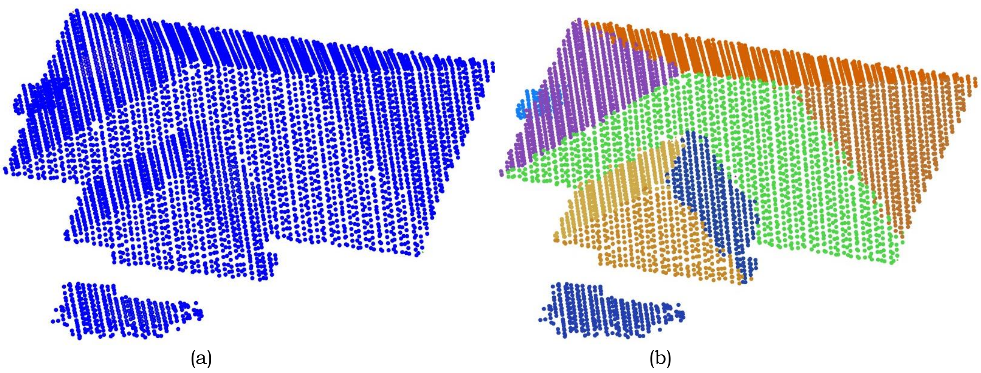



**Figure 2.9** Point cloud segmentation using region growing: (a) Sample building LiDAR data and (b) Segmented LiDAR data.

The selection of a seed point depends entirely on the nature of the problem. Often, the random selection of a point as the seed can segment the data but it generally requires some prior knowledge. However, in the presence of noise, outliers and the diverse features in urban areas, a particular criterion needs to be defined for the selection of a seed point. In contrast with images where the neighbouring pixels can be selected in four-connected neighbourhoods or eight-connected neighbourhoods, neighbouring points in LiDAR are selected using a fixed-distance neighbourhood and *k-nearest neighbourhood* (Knn) methods [69]. Figure 2.10 shows the difference between the two neighbourhood selection approaches. $P_i$ shows a random LiDAR point and $N_{P_i}$ shows the neighbours of $P_i$. In the case of a fixed-distance neighbourhood, the radius $r$ of a concentric-circle needs to be defined and $N_{P_i}$ corresponds to the points which lie within the circular region defined by $r$. In contrast, the Knn method uses Euclidean distance for the construction of a *kd-tree* from the LiDAR point cloud. The value of parameter $k$ signifies the number of neighbours of point $P_i$.

Once the neighbours of a seed point are identified by the region-growing process,

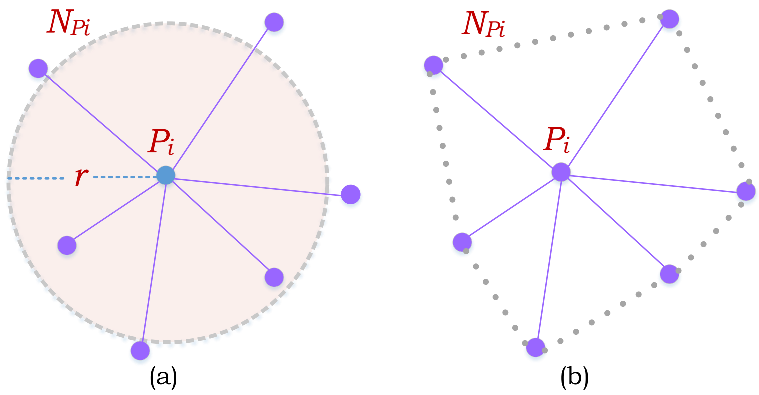



**Figure 2.10** (a) Fixed-distance neighbourhood and (b) K nearest neighbourhood.

they are included in a region only after examining all the given criteria. The entire procedure is iterative and stops when a significant change is observed between successive iterations or no new neighbouring points are added. Throughout the present thesis, we employ the Knn method for neighbourhood selection, since it quickly adapts to the limitation of point density variation and sparseness. It is a well-established fact that point density variation occurs because of the variations in both data acquisition sensors and the orientation of a surface with respect to the scanner position. More specifically, we utilise the Knn method to find the local neighbours of a laser point in point cloud segmentation and approximate their saliency features using the PCA in Section 5.2.2.

## 2.4 Performance Evaluation Systems

In light of the end goal to gauge the quality of building recognition and rooftop extraction techniques, and to evaluate their pertinence for functional applications, it is important to assess them by contrasting the detected and the reference information. An evaluation framework measures the robustness of a technique by different performance measurements. Such systems perform a coordinated correspondence between the identified polygons and the reference polygons. The overlapping region of the polygons is determined through pixel-by-pixel comparison, in order to establish the accuracy of a technique. The overlapping region is categorised into four types as follows:

1. TP (True Positive): A detected polygon is classified correctly in the reference data as a building;

2. TN (True Negative): An entity is classified as other object in both the detected and reference data;

3. FP (False Positive): A detected polygon (a building) does not have a corresponding object in the reference data;

4. FN (False Negative): A reference polygon (a building) is identified as other object (not a building);

Several possible quality metrics can be derived using these quantities in order to assess the performance of building detection and rooftop extraction techniques. Irrespective of the general understanding, to date no single uniform evaluation system is available which can run performance evaluation for any benchmark dataset in the realm of airborne laser scanning [70]. Therefore, various studies use different evaluation methods for measuring the quality of their methods. This selection is heavily dependent upon the dataset used for the analysis. In the present study, we use the International Society for Photogrammetry and Remote Sensing (ISPRS) benchmark datasets [71] and the Australian benchmark datasets. Therefore, the threshold-based system [70] has been used for evaluation of the ISPRS datasets while for the Australian benchmark datasets, the present study uses an automatic and threshold-free evaluation system [72].

Both the evaluation systems perform three categories of evaluation: object-based, pixel-based, and geometric, and each category uses several metrics. The object-based metrics evaluate the performance by counting the number of buildings, while the pixel-based metrics measure the detection accuracy by counting the number of pixels. The pixel-based metrics also correspond to the horizontal accuracy of the detected polygon and indicate the area which is classified accurately. In addition, the geometric metric indicates the accuracy of the extracted boundaries with respect to the reference entities. It also measures the height accuracy of a detected polygon with respect to a reference polygon.

### 2.4.1 Threshold-based Evaluation System

The ISPRS working group II[1] quantitatively evaluates the performance of a technique using a threshold-based evaluation system [70]. This system applies a number of thresholds for establishing the correspondence between the detected and the reference buildings for performance evaluation. Here, we provide a summary of the evaluation system. The pixel-based evaluation of the system compares the detected and the reference polygons pixel by pixel in order to compute the TP, FP, FN, and TN pixels. This evaluation further measures the pixel-based completeness, correctness, and quality of the detected buildings. In contrast, in object-based evaluation, the overlapping region between the detected and reference polygons is determined. This approach actually counts the pixels assigned to a building ($B_r$) in the reference label image with a building ($B_a$) in the detected label image.

In the case of object-based TP, the reference polygon is considered to be detected if more than 10% of its pixels overlap those of the detected polygon. In the case of pixel-based TP, the overlapping pixels of the detected polygon are considered positively identified if 50% of its pixels overlap those of the reference polygon. In reality, a detected building might have zero, one, or multiple overlaps with other buildings present in the reference. Therefore, the authors used a split and merge technique to accurately establish the correspondence, which is demonstrated graphically in Figure 2.11.

Figure 2.11(a) shows the scenario of a merge operation when a detected building ($B_a$) is overlapping $N$ reference buildings ($B_{r1}$ and $B_{r2}$). Therefore, the detected building is split into $N$ new buildings. In contrast, a merge operation is invoked if the correspondence of $B_a$ is greater than that of $B_r$ as shown in Figure 2.11(b). Similarly, both the split-and-merge operations are performed for a scenario presented in Figure 2.11(c). For more technical detail, readers are encouraged to study the following article [70].

[1] ISPRS Working Group II/4 aims to make progress in the automatic recognition and 3-D reconstruction of objects in complex scenes from images, point clouds, and other sensor data. More detail can be found at `http://www2.isprs.org/commissions/comm2/wg4.html`

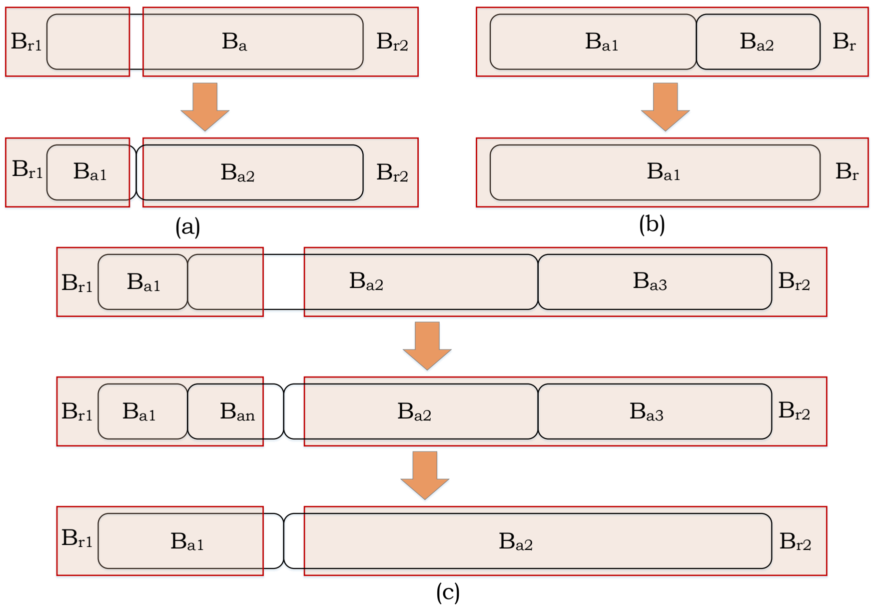



**Figure 2.11** (a) Split of a detected building into two new buildings; (b) Two detected buildings are merged into one building; and (c) The detected buildings are split and merged to match the reference buildings [70].

### 2.4.2 Threshold-free Evaluation System

In contrast with the previous system, the threshold-free evaluation system [72] uses pseudo one-to-one similarity matching to avoid the use of thresholds. In this system, each building of one dataset has at most one correspondence in the other dataset. In the case where a detected building overlaps more than one reference building, the nearest reference building is considered as a true matched reference building for the detected building. The pseudo one-to-one correspondence is illustrated in Figure 2.12.

The detected building is labelled FP if it does not overlap any reference building. Similarly, if a reference building overlaps any detected building, the reference building is labelled FN. In the case where the reference buildings have one-to-one correspondence with the detected buildings, both are labelled TP, as shown in Figure 2.12(a). If several reference buildings overlap a detected building, the reference building (e.g., 1 in Figure 2.12(b)) which is nearest to the centre of the detected building is labelled TP, and others (e.g., 2 in 2.12(b)) are labelled FP. If a reference building overlaps multiple detected buildings (as in Figure 2.12(c)), both the detected buildings are labelled multiple detections (MDs). Otherwise, the detected buildings are not labelled as MDs (see Figure 2.12(d)).

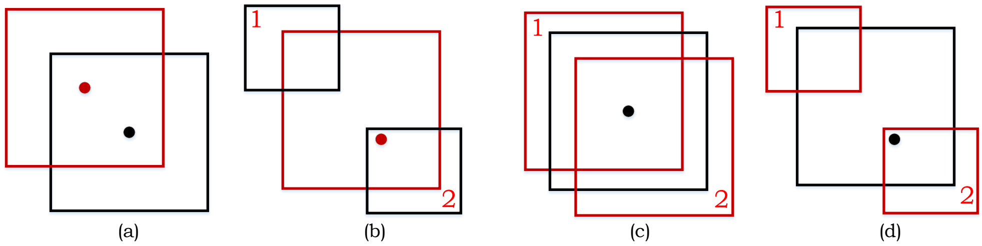



**Figure 2.12** (a) Different scenarios for establishing a pseudo one-to-one correspondence, where a red rectangle denotes a detected building and a black rectangle denotes the reference building [72].

## 2.5 Quality Assessment Criteria

As mentioned previously, both the evaluation systems run three categories of evaluation, i.e., object-based, pixel-based, and geometric. Each category includes a number of metrics to measure the performance of a proposed method. The object-based metrics (i.e., completeness ($C_m$ ), correctness ($C_r$), quality ($Q_l$), over-segmentation (one-to-many - $1:M$), under-segmentation (many-to-one - $N:1$) and over-and-under segmentation (man-to-many - $N:M$) errors) measure the performance quantitatively by counting the number of building objects, whereas the pixel-based metrics (i.e., pixel completeness ($C_{mp}$), pixel correctness ($C_{rp}$), and pixel quality ($Q_{lp}$)) evaluate the performance by counting the number of pixels of the detected objects.

$$Completeness = \frac{TP}{TP+FN} \tag{2.2}$$

$$Correctness = \frac{TP}{TP+FP} \tag{2.3}$$

$$Quality = \frac{TP}{TP+FN+FP} \tag{2.4}$$

In contrast to the system adopted by the ISPRS [70] which measures one-to-many, many-to-one, and many-to-many segmentation errors, the threshold-free evaluation system measures detection cross-lap ($D_{cl}$) and reference overlap $R_{cl}$. To compute $D_{cl}$, the formula defined in [72] uses $C_{ld}$ which is the number of detected buildings with more than one correspondence with the reference buildings.

$$D_{cl} = \frac{C_{ld}}{TP+FP+MD} \tag{2.5}$$

Similarly, for the calculation of $R_{cl}$, the formula uses $C_{lr}$ which gives the number of reference buildings overlapping more than one detected building.

$$R_{cl} = \frac{C_{lr}}{TP+FP+MD} \tag{2.6}$$

In addition, the geometric accuracy in terms of *RMSE* is determined, which measures the average distance between a pair of detected and reference buildings [73]. This can be represented as:

$$RMSE = \sqrt{\frac{\sum d^2}{N}} \tag{2.7}$$

Here, $d$ corresponds to the Euclidean distance between the corresponding boundary points of the detected and reference objects, whereas $N$ refers to the number of points for which a correspondence has been found. Moreover, the geometric evaluation process also measures the height accuracy ($RMS_z$), which is the root mean square height distance of the points within the corresponding planes. The height values of all points corresponding to both the planes are used to calculate $RMS_z$ as:

$$RMS_z = \sqrt{\frac{\sum (Z_{ref_i - Z_{det_j}})^2}{N}} \tag{2.8}$$

Here, $Z_{ref_i}$ and $Z_{det_j}$ are the height values of the points from both the reference and the detected objects, respectively.

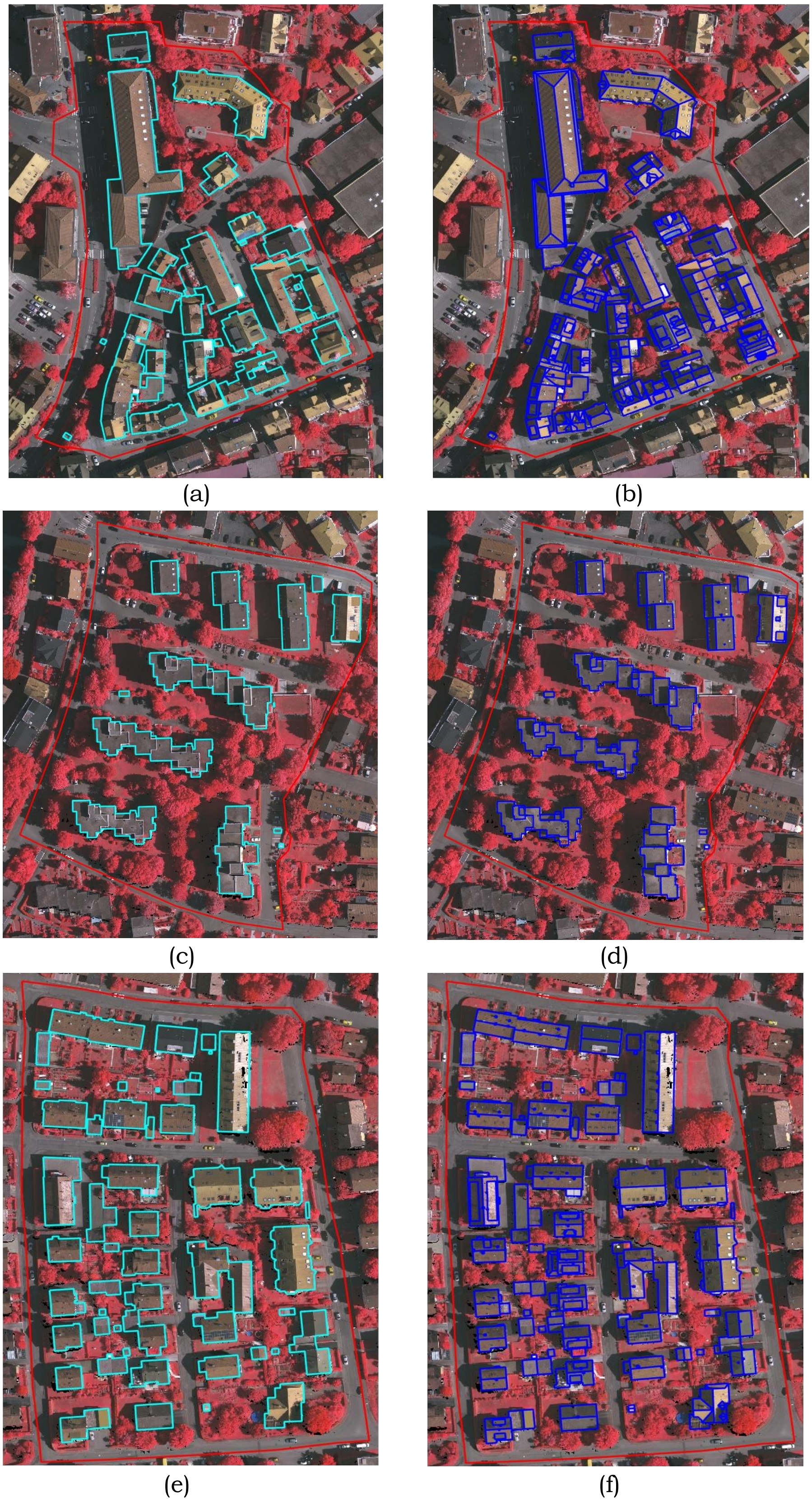
## 2.6 Benchmark Datasets

This section provides information about the data used throughout the thesis. The information about the characteristics of the data is important for the robust interpretation of the proposed methods and the results achieved.

### 2.6.1 The ISPRS Benchmark Dataset

The first dataset is Vaihingen (VH) from the ISPRS benchmark and provided by the German Society for Photogrammetry, Remote Sensing and Geo-information (DGPF) [71]. It has three test areas, as shown in Figure 2.13. Each area has a point density of 3.5, 3.9, and 3.5 points/m$^2$, respectively. The VH1 area is situated in the centre of the city and characterised by dense construction consisting of historic buildings. These buildings have complex shapes, such as gables and hip roofs, with small chimneys and dormers. In most cases, the buildings are located beside vegetation that often partially occludes the buildings. The VH2 area is characterised by a few high-rise residential buildings surrounded by dense trees. Finally, VH3 area is purely residential, with detached houses and many surrounding trees. The numbers of buildings larger than 2.5 m$^2$ in these three areas are 37, 14, and 56, and the corresponding numbers of planes are 288, 69, and 235, respectively. The red polygons in all the areas of the ISPRS benchmark dataset show the region the evaluation system uses for quantitative performance analysis. For independent evaluation, building detection and roof extraction results are submitted to the ISPRS, which quantitatively evaluates the performance of a technique and publishes the results online[2].

[2] http://www2.isprs.org/commissions/comm3/wg4/results.html



(a) (b) (c) (d) (e) (f)

**Figure 2.13** The ISPRS Vaihingen dataset: (a)–(b) VH1 ; (c)–(d) VH2; and (e)–(f) VH3. Column 1 shows polygons for the reference buildings and Column 2 shows polygons for the reference roof planes.

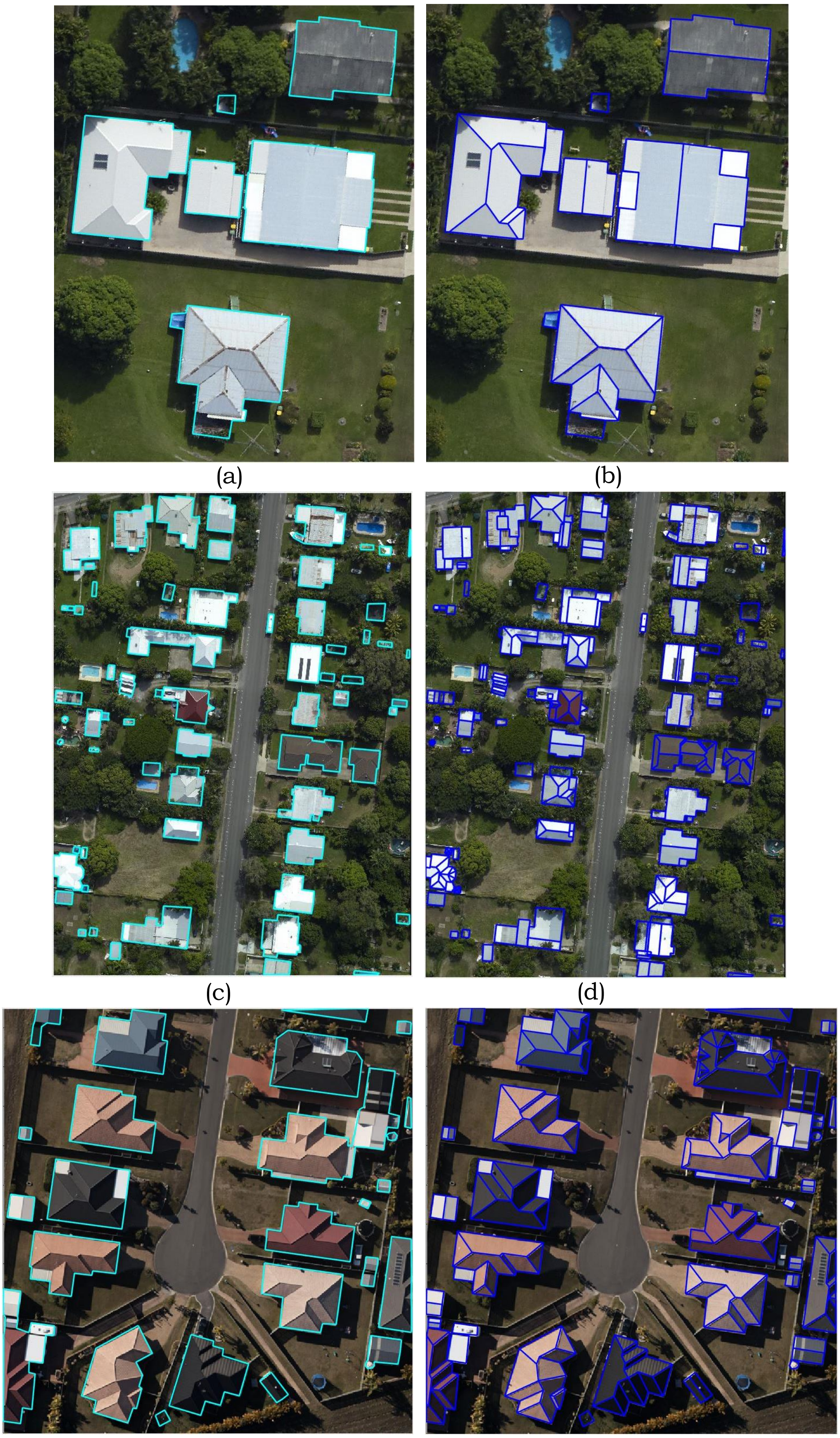
### 2.6.2 The Australian Benchmark Dataset

Four Australian datasets, Aitkenvale (AV), Hervey Bay (HB), Eltham (EL), and Hobart (HT) were used for testing the applicability of the proposed solutions on diverse areas. The first two datasets were provided by Ergon Energy[3] in Queensland, Australia while the latter two were supplied by Photomapping Services and the Department of Environment and Primary Industries of Victoria, Australia, respectively.

The AV dataset has two areas for which we used two different acronyms, AV1 and AV2. The AV1, AV2, HB, EL, and HT datasets have point densities of 35, 29.3, 12, 4.8, and 1.6 points/$m^2$, respectively. Topographically, the AV and HB areas are flat, while EL and HT are hilly, containing mostly residential buildings with different levels of vegetation. The AV, EL, and HT datasets have dense vegetation, and many of the buildings are severely occluded by the surrounding trees. The test areas AV1, AV2, HB, EL, and HT cover $66 \times 52$ $m^2$, $214 \times 159$ $m^2$, $108 \times 104$ $m^2$, $393 \times 224$ $m^2$, and $303 \times 302$ $m^2$ areas, respectively, as shown in Figure 2.14(a)–(j).

AV1 contains 6 buildings comprising 24 roof planes, while AV2 has 63 buildings (four are between 4 to 5 $m^2$ and 10 are between 5 to 10 $m^2$) comprising 211 roof planes. The EL dataset contains 75 buildings (nine are less than 10 $m^2$, including five within 3 to 5 $m^2$) consisting of 441 planes. The HT dataset has 69 buildings (thirteen are less than 10 $m^2$, including four within 1 to 5 $m^2$) containing 257 planes. The HB dataset contains 28 buildings (three are between 4 to 5 $m^2$ and six are between 5 to 10 $m^2$), consisting of 166 roof planes. The reference datasets included parking shades, huts, and shelter umbrellas as buildings in the evaluation.

[3] http://www.ergon.com.au


(a) (b) (c) (d) (e) (f)

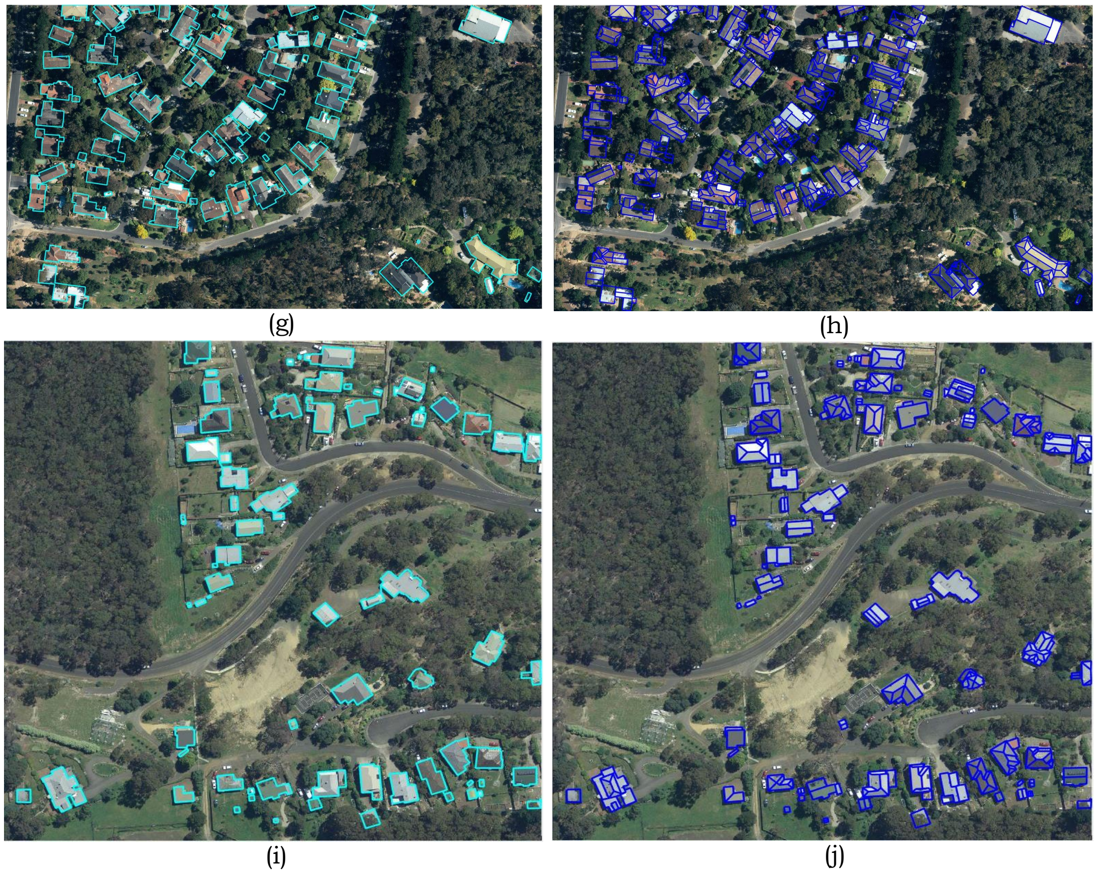



**Figure 2.14** Australian benchmark datasets: (a)–(b) AV1; (c)–(d) AV2; (e)–(f) HB; (g)–(h) EL; and (i)–(j) HT. Column 1 shows polygons for reference buildings and Column 2 shows polygons for the reference roof planes.

## 2.7 Summary

The background material relevant to understanding the thesis has been discussed in this chapter. We summarised the characteristics of remote sensing systems, both satellite- and airborne-based, and provided a discussion on which data source is better for GIS object detection and classification. We briefly discussed the technical principles to establish the foundation for a better understanding of the work presented in the following chapters. We also briefly presented the performance evaluation systems and the metrics to measure different quality aspects in terms of pixels and objects. Finally, the datasets used in the entire thesis were introduced in the last section. The next chapter presents the relevant literature to provide the motivation for the research on building detection, roof plane extraction, and 3-D building modelling.

Chapter 3

# Literature Review

> *"If you can't explain it simply, you don't understand it well enough."*
>
> Albert Einstein

## 3.1 Introduction

In this chapter, we address the first research objective **RO1**, i.e., "**To identify the research gaps and limitations, and conduct a review of existing approaches to building detection, roof plane extraction, and 3-D building modelling**". Recent developments in the fields of photogrammetry and remote sensing for automating the measurement and scene interpretation tasks have resulted in sophisticated methods with promising results for data acquisition, and the identification and reconstruction of urban objects. Many scientists have developed building recognition and reconstruction techniques utilising image information only, others have utilised LiDAR point cloud, and some have attempted to integrate LiDAR and aerial images for several GIS applications.

In recent years, studies on the detection and reconstruction of buildings have made significant advances. These methods can broadly be classified into three categories on the basis of their processing strategy: model-driven, data-driven, and hybrid approaches [12, 17, 74, 75]. A model-driven method uses a predefined building model (shape) and fits into the input data for the extraction purposes, in contrast to a data-driven method that uses the input data and extracts one or more features (corners, lines, and planes) for the detection and reconstruction of buildings. A hybrid method, on the other hand, exhibits the characteristics of both model- and data-driven approaches. All these strategies differ significantly in terms of degree of automation, the data sources used, methodologies, and the processing strategies. Therefore, they can be categorised according to the following different criteria [22, 76]:

- Degree of automation: automatic or semi-automatic
- Input data source: photogrammetry image and/or laser scanning data
- Processing strategy: supervised (probabilistic or non-probabilistic classifiers), unsupervised (rule-based) or statistical (heuristic model as energy function)
- Data features: corners, lines, illumination, contrast (from image) and/or geographical location, point density, height, intensity, flight scan-line information (from LiDAR data)
- Pre-existing knowledge: none; building geographical position, cadastral map or off-line pre-computed information
- Output: buildings, roof planes, training parameters; geometric information; building model; or scene description

There are numerous possibilities for categorising the existing techniques on the basis of given criteria. However, the input data source parameter is adopted for the classification in this chapter to simplify the systematic review of the techniques. Further, the scope of the discussion is limited to the context of the proposed study. Therefore, only the relevant techniques are reviewed. A schematic categorisation of the literature review is shown in Figure 3.1.

In Section 3.2, we provide the definitions of some basic concepts, followed by a review of building detection techniques in Section 3.3. The techniques for roof plane extraction and 3-D building modelling are discussed in Sections 3.4 and 3.5, respectively. The research challenges addressed in the thesis are presented in Section 3.6. The chapter concludes with Section 3.7.

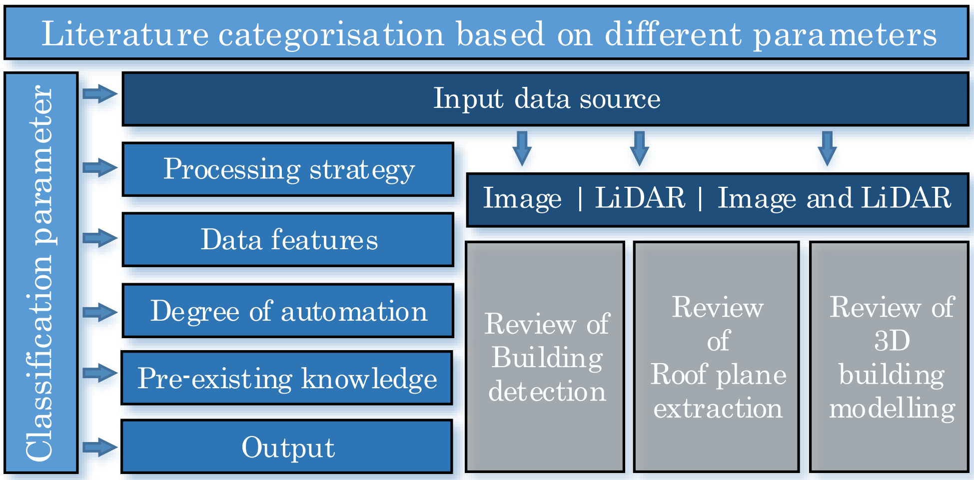



**Figure 3.1** Schematic of literature review classification.

## 3.2 Preliminaries

This section provides the definitions of some basic terms in order to assist understanding of the topics discussed. These terms will frequently be used throughout the dissertation. The definitions are as follows:

- An **automatic method** refers to a technique which does not require any user interaction in the course of the classification/segmentation/detection steps, apart from the selection of model parameters or determining the training data off-line, and the output of the method is not subject to any post-editing [22].
- The **building detection** procedure aims to identify the location, area, and the boundary of a building region. The extracted boundary is generally represented as 2-D LiDAR points/image pixels.
- The **building regularisation** technique generates a 2-D regular outline of a building periphery by replacing the sequence of boundary points with straight lines.
- **Roof plane extraction** techniques classify the LiDAR point cloud into homogeneous and non-homogeneous points in order to detect the roof plane surfaces. These techniques also estimate the area and boundary of the extracted surfaces.

- **Roof reconstruction** is a procedure for generating a 3-D reconstructed view of building roofs with the minimum number of geometric primitives. This procedure is also referred to as the roof plane regularisation procedure.

- **Building reconstruction** is the determination of the geometrical parameters of a building in 3-D space located in a given region of interest [77]. In addition to the roof model, the wall models are generated and added to obtain a complete building model.

- A **building delineation error** denotes when the boundaries of buildings are not properly approximated.

- A **building detection error** corresponds to an error when buildings are not detected by a building detection process.

## 3.3 Building Detection

Of the countless objects on the earth, buildings are a fundamental component and quite diverse for several applications in the field of urban planning, civilian and military emergency responses, cartographic mapping, and crisis management [17, 78]. Accurate and current information on buildings is imperative to keep these applications operational. In the past, some promising solutions used an interactive initialisation set-up, followed by an automatic extraction procedure [79]. However, this approach is practical only if a trained human operator makes the initialisation and supervises the extraction process. In contrast, automatic building extraction has proven to be a non-trivial task [9]. The following sections provide a review of techniques using single and multiple data sources.

### 3.3.1 Review of Image-based Methods

The development of *very high resolution* (VHR) space-borne images makes it possible to sense individual buildings in an urban scenario [10], which is imperative in various reconstruction, cartographic, and crisis management applications. However, the increasing information in high-resolution imagery does not guarantee a proportional increase in building detection performance [11, 12]. Rather, it adds to the spectral ambiguities [13]. Consequently, similar urban objects may have different spectral signatures, whereas different objects may have similar spectral signatures under various background conditions [14]. These factors, together with hardware errors, reduce the building detection rate. Therefore, using spectral information alone to discern buildings from other urban objects eventually results in poor performance [10]. A practical solution is to combine the information of both LiDAR and images for building region detection. Following a major principle of scientific writing that the readers must understand the concept, we, therefore, provide a brief discussion of the methods utilising images as the only source for building detection. The following Section 3.3.3 primarily focuses on a literature review of studies using the integration of multiple data types.

Song et al. [80] proposed a building detection technique based on single high-resolution satellite imagery. The detection was performed in two stages: hypothesis generation and hypothesis verification. Firstly, the input image was decomposed into small regions, called *Candidate Building Regions* (CBRs). Next, image lines were extracted from each CBR and classified into nine sub-sets. Then, four random lines from each sub-set were used in developing a rectangular building hypothesis. Finally, the authors applied an edge verification approach to remove all the unreasonable hypotheses for further processing. The method was tested using satellite imagery that did not have any vegetation. According to the performance evaluation results, the method detected more than 90% of buildings. However, buildings which were small or constructed with dark rooftops were not identified.

Wei and Prinet [81] applied a probability function for the detection of building regions. They first performed image segmentation that decomposed the image into several regions. Next, they computed many features, including shadow ratios, region entropy, contour edges, and shape features for each region. The authors then applied a probability function to calculate a confidence value which defines whether or not a region corresponds to a building. The probability function utilises the following features for calculating the confidence value: distance to straight line segments, contour region entropy, edges, grey level, standard deviation, shape, and shadow ratio. The values of many parameters were identified by manual interaction, making this method semi-automatic. The authors identified the presence of shadows as a key factor in poor performance.

Izadi and Saeedi [82] proposed a building detection technique using aerial images. The images were decomposed using a hierarchical segmentation algorithm. Unlike a generic segmentation algorithm that uses colour or intensity information for image segmentation, a hierarchical segmentation algorithm segments the given images based on adaptive colour range resolution. Using this algorithm, a segmented image along with a set of regions for each range were produced. The authors then organised the information on the image segments and the set of regions in a tree structure which was later used to obtain the best colour range resolution for various image regions. After applying some building rooftop features, trees, roads, and other non-building objects were removed. Finally, the regions which were left were the potential buildings. The authors tested their method using seven aerial images, and the results showed high accuracy. However, since the test datasets had little vegetation, the robustness of the method remains unclear for datasets with dense vegetation.

Sirmacek and Unsalan [83] employed colour-invariant features and shadow information in a feature- and area-based technique. The shadows were detected using the information of the blue colour channel, while roofs constructed with red coloured material were detected using the red colour channel. The authors detected the shadow regions first and the regions opposite to them were then selected using the illumination angle as candidate regions. They next applied the canny edge detector and then the box-fitting algorithm for candidate region extraction. This technique assumes that all the buildings have rectangular shapes and contain only a single texture on the rooftops. Therefore, the method is not applicable for other types of buildings and multi-textured rooftops.

### 3.3.2 Review of LiDAR-based Methods

Airborne LiDAR data provide height information on salient ground features, such as buildings, trees, bushes, terrain, and other 3-D objects. Therefore, the adoption of height variation to distinguish various urban objects is a more suitable cue than spectral and texture changes. However, the detected boundaries have low planimetric accuracy due to point cloud sparsity [12, 14]. In addition, the appearance of trees and buildings is sometimes similar in complex urban scenes [15], when height information alone is not perceived to produce a finer classification [11]. Therefore, the fusion of LiDAR with aerial images is regarded as a promising strategy to increase the building detection rate [12, 14–17]. The following Section 3.3.3 focuses on integration techniques for building detection. However, this section discusses detection techniques using only LiDAR data.

Cheng et al. [84] presented a reverse iterative morphological algorithm for building extraction. They used a large window size at the beginning for the morphological operations to separate the ground points and the non-ground points (on buildings, sparse and dense trees). After each iteration, the window size was decreased and only the non-ground points were filtered based on the height difference between two successive morphological operations. The authors used some other parameters the values of which were adjusted manually, based on the minimum and maximum building size and the minimum building height. Finally, the tree regions were removed at the post-processing stage with the help of surface roughness measurement and the building regions were detected. This technique was tested using 10 building samples and the minimum building size was over 500 $m^2$. Therefore, the algorithm's ability to detect buildings as small as 10 $m^2$ is unclear. Moreover, this method failed to detect one large building and some of the building parts, and this shows low detection accuracy.

Yang et al. [85] optimised the Gibbs energy model for building detection. Initially, a LiDAR point on each building that served as prior knowledge of the building position was marked in terms of energy. The authors used *Reversible-Jump Markov Chain Monte Carlo* (RJMCMC) to sample the global energy to fit the LiDAR data to the marked building points for the generation of the building cuboid. Finally, the false buildings were removed and the boundaries of the potential buildings were refined. The method was tested on three datasets. Several low height buildings were not detected, resulting in low object-based evaluation. However, the technique achieved high pixel-based accuracy. The energy optimisation process needs an exact number of iterations to achieve a building's global optimal location, which is a computationally expensive process and a major limitation of this method.

Mongus et al. [86] proposed a LiDAR-based building detection technique. The authors introduced a differential morphological profile algorithm which makes the process of segmentation independent of building size and shape. The buildings were detected in a three-stage process. Firstly, the LiDAR points were arranged into a grid in which the grid resolution was defined by the density of the data. Next, the outliers were removed using a data de-noising process. At this stage, bushes and small trees were removed for further processing. The authors finally applied a grid segmentation process to detect the building regions. The performance was evaluated using three datasets and the results showed high accuracy in terms of objects and pixels. However, the user-defined values were not consistent in the different datasets, indicating that the method is quite subjective.

Awrangjeb and Fraser [87] presented a region-growing-based building extraction method using LiDAR data. They first separated the LiDAR data and generated a building mask. The authors then decomposed the mask into equal-sized cells and used a region-growing technique for the extraction of planar regions. The segmentation technique used the coplanarity of the laser points and their locality for segmenting the non-ground LiDAR points into planar surfaces. The planar surfaces on trees and other non-building objects were then removed using area, point height difference, and used-to-unused point ratio. The technique establishes a neighbourhood relationship among the adjacent planes which is used finally to detect the building regions. Although this method achieves reasonably high pixel-based accuracy, it is unable to extract small and occluded buildings.

### 3.3.3 Review of Methods using Multisource Data

Despite the agreement to use multiple data sources, how to extract and integrate the distinct features so that their weaknesses can be compensated effectively is a hot area of investigation. Haala [88] classified the integrated methods into two categories. The first group of techniques [12, 14, 17, 89] primarily uses the image features, for instance NDVI, entropy, and illumination, for the elimination of unwanted objects and the extraction of buildings. Therefore, the detection rate and planimetric accuracy of the detected buildings is fairly high.

Chen et al. [14] proposed an integration method for building detection using LiDAR data and QuickBird imagery. A *normalised DSM* (nDSM) was first generated from the LiDAR data. Then, it was decomposed into equi-sized grids and the cells representing unwanted objects were eliminated using prior knowledge of minimum building height. The authors next utilised the NDVI and spatial relationship in order to eliminate trees and to extract the building roofs. This technique performs poorly when buildings have small roof planes or complex roof structures. Moreover, several buildings with green roofs were also removed due to inappropriate usage of features from LiDAR and images. Since the method was tested using single dataset, the robustness of this technique is unclear.

Grigillo et al. [55] follows the mask generation process in [14] and then eliminated the vegetation under shadows by truncating the areas with low homogeneity. However, this technique does not address the occlusion issue and produces inaccurate building boundaries when they are surrounded by trees but works well when trees are isolated. Another technique [49] generates a DSM from the first and last pulse return of the LiDAR data, and uses the NDVI and spatial relationship between buildings and trees to complement the detection process. However, it has the similar limitations to the previous technique.

A technique by Awrangjeb et al. [17] initially separated the LiDAR data into ground and non-ground points using a height threshold value. The non-ground points were then used to generate a building mask. The line segments were then extracted from the image and used for the segmentation of non-ground LiDAR points. Finally, the authors, used image features like NDVI and entropy for the removal of vegetation and identification of the building regions. Due to a large height threshold (2.5 m) to avoid roadside furniture, bushes, and several low height objects, many small buildings (with areas $< 10$ m$^2$) were not detected. In addition, buildings, which were partially occluded and in shadow were also not detected.

The techniques in the second group utilise LiDAR as the primary cue for building detection and employ image features only to remove vegetation [15, 90]. Consequently, such approaches show poor horizontal accuracy of detected buildings due to LiDAR point discontinuity. A method based on the Dempster–Shafer theory [15] classified the LiDAR data into several groups representing buildings, trees, grassland, and bare soil. A morphology operation was later performed to eliminate the small segments and detect the building regions. However, this procedure requires tuning of several parameters with an estimation of wooded areas. This technique results in poor detection rates in the case of small buildings because it uses a large threshold (area $\geq 30\ m^2$) and an untrained Dempster–Shafer model [91]. Xiao et al. [9] used edge and height information in a dense image-matching technique to detect the façades in oblique images. The authors considered these façades as a representation of vertical planes to define the building hypotheses. A major disadvantage is that the method fails to extract small buildings and simply ignores building attachments and small structures in backyards and open areas.

Qin and Fang [13] first obtained an initial building mask hierarchically by considering the shadow and off-terrain objects. A graph cut optimisation technique based on spectral and height similarity was then used to refine the mask by exploring the connectivity between the building and non-building pixels. This method can handle shadows and small buildings to a good extent, but building patches on steep slopes, roof parts in shadow, and roofs with vegetation cannot be extracted. Zhang et al. [10] proposed a dual morphology top-hat profile to overcome spectral ambiguity using a DSM and ultra-high-resolution image for feature classification. However, the accuracy of the DSM remains a critical factor in building extraction, particularly for small objects.

Awrangjeb et al. [41] proposed a residential building detection technique using LiDAR data and orthoimagery. They generated two masks from the LiDAR data to represent void and filled areas, and then used the prior mask for extracting the line segments. Trees were removed next using the NDVI from the input orthoimage. Using the line segments, this method further identified the initial positions of the buildings. Finally, the building footprints were extracted from their initial positions using the generated masks and orthoimagery. The evaluation results showed that the method has good object-based accuracy than pixel-based accuracy. Moreover, the technique cannot represent the position of a complex building using single polygon. Instead, different parts of some buildings were detected separately.

Since the LiDAR-based methods generally produce ragged building boundaries, some researchers have tried to outline the ragged boundaries with minimum possible lines, which is called boundary regularisation or generalisation in literature. The following techniques give an overview of boundary regularisation studies. The polygon extraction method proposed in [92] determined the dominant direction of a building using cross-correlation mapping and later, used a rotating template and angle histogram to obtain a regularised boundary. Fu and Shan [93] used three primitive models based on locating 2-D rectangles to construct 3-D polyhedral primitives followed by assembling the final buildings with right-angled corners.

Ma [94] categorised the lines into two perpendicular classes and performed a weighted adjustment to calculate the azimuth of these classes using the Gauss-Markov model. Finally, the adjacent parallel lines were combined together to construct a regularised boundary. Sampath and Shan [74] utilised a conditional hierarchical least squares method ensuring the lines participating in the regularisation process have the slopes of parallel lines equal or a product of the slopes of perpendicular lines to be $-1$ (orthogonal). Another similar technique [95] forced boundary line segments to be either parallel or perpendicular to the dominant building orientation when appropriate and fitted the data without a constraint elsewhere.

A compass line filter [96], on the other hand, extracted straight lines from irregularly-distributed 3-D LiDAR points and constructed a boundary by the topological relations between adjacent planar and local edge orientation. For building boundary extraction, Wei [97] used the alpha-shape algorithm and applied a circumcircle regularisation approach to outline rectangular buildings.

Rottensteiner [15] found through the comparative analysis of several techniques that a major reason for any building detection technique producing low detection rates is incompetence in the extraction of partially-occluded buildings. These techniques also produce a high delineation error because trees standing close to buildings are wrongly merged into the building region.

## 3.4 Building Roof Detection

Building roof extraction techniques aim to identify individual planes of buildings and their constituent components e.g., dormers, chimneys, and vents. To the best of our knowledge, no attempt has been undertaken to date to recognise roof planes using only image data. We, therefore, do not provide any discussion of this category. Our statement is further supported by another study [98] that also claims the unavailability of image-based roof extraction techniques. Therefore, the following sections provide a review of different roof plane extraction techniques using LiDAR data and those integrating multiple data sources.

### 3.4.1 Review of LiDAR-based Methods

For building roof detection using LiDAR data, *RANdom SAmple Consensus* (RANSAC), *Hough Transform*, and *region-growing algorithms* are three major contenders often utilised for point set segmentation [99–102]. 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 the unorganised point cloud by accumulating points into regions satisfying certain conditions.

To comprehend the theoretical and practical feasibility of these approaches, Deschaud and Goulette [103] conducted a comparative study. They showed that RANSAC is very efficient for detecting large planes in a noisy point cloud but very slow for small planes in large datasets. They also showed that the Hough Transform is computationally expensive and time-consuming for plane fitting and extraction. In contrast, they argued that region growing is a quite robust and fast segmentation approach which offers strong resilience to noise [104]. However, it is not highly accurate due the sensitivity and location of the initial seed [103]. But this issue can better be addressed when global information is used by the segmentation process [104].

Awrangjeb and Fraser [87] proposed a roof plane and building extraction method using LiDAR data. They first separated the LiDAR data into the ground and non-ground points, and generated a primary building mask. Points from the earth, roadside furniture, and small bushes were represented in white in the mask whereas non-ground points, including buildings and trees denoting the elevated objects, were represented in black. The authors then divided the mask into equal-sized grids and applied a region-growing technique to extract the planar regions. They then used a rule-based method to remove the false planes and extract the roof planes. Although, this method achieves a reasonably high pixel-based accuracy, it is unable to extract small and occluded roof planes.

For point cloud filtering, Vosselman [105] proposed a slope-based classification algorithm that first computed the slope between any two adjacent LiDAR points and classified them as non-ground if the slope value was larger than a threshold value. However, the selection of a threshold value is quite critical to the method which directly affects the filtering process and can flatten the terrain details. In contrast, a method proposed by Arastounia and Lichti [106] used the point height histogram mechanism for the dynamic selection of the threshold value. A peak in the histogram represents the ground surface and a threshold was chosen where the bin entries were decreasing dramatically. Although, this automatic selection criterion works well for the flat terrain, it does not accurately produce a threshold value for hilly and slopping surfaces.

Sampath and Shan [35] presented a framework to extract and reconstruct polyhedral building roofs from LiDAR point cloud data. First, planar and non-planar points were determined through eigenvalue analysis. Next, plane segments were extracted using planar points by a modified *fuzzy k-means clustering algorithm* [107]. The non-planar points were then assigned to the appropriate planar clusters, where finding a cluster centre was computationally expensive. The authors showed results using a small number of building samples and the effect of noise was completely ignored although a surface becomes noisy in dense datasets [52]. In addition, the results were not convincing for small roof planes.

Dorninger and Pfeifer [34] assumed that buildings comprise of planar structures and the LiDAR points of these surfaces have similar local normals. They began with the identification of the seed points through a histogram analysis. A predominant bin in the histogram was chosen as the feature space parameter. They later detected the planar patches utilising seed points and coplanarity analysis in a region growing technique. Then, the alpha-shape algorithm was applied to approximate the boundary of the building regions. The authors proposed a boundary regularisation technique that use the mean angular directions of the approximated boundary to regularise the building outline. This study shows the applicability of the technique using two datasets, but does not provide any discussion of partial occlusion and the elimination of vegetation. Further, the authors advised initialising the coarse selection of building regions interactively to achieve better accuracy. In addition, this method requires some of the algorithmic parameters to be tuned through manual interaction, which causes this technique to be categorised as semi-automatic.

Zhou and Neumann [108] proposed a region growing technique for segmenting building rooftops. They used a contouring algorithm to extract the building boundaries. Later, histogram statistics were used to determine the principal direction of the buildings and roof planes for footprint generation. This method works only for flat roofs and needs manual interaction for the identification of non-flat surfaces.

In another technique proposed by Awrangjeb and Fraser [109], a rule-based segmentation technique was presented using LiDAR point cloud. They initially divided the LiDAR data into two groups based on a height threshold and generated a building mask. The non-ground points were further classified based on the coplanarity of points. Using planarity and neighbouring plane information, a rule-based method was applied to remove the tree segments and identified the roof planes. Results show that this method missed some small buildings and several roof planes. Moreover, the detected boundaries were ragged and had low planimetric accuracy.

An area-wide point cloud segmentation was proposed by Jochem et al. [110], where data were processed in the form of several overlapping tiles. The candidate building regions of all the tiles were detected first and then merged into a single polygon layer for plane extraction. The normal vector to all the points in the polygonal layer were approximated by fitting an orthogonal regression plane and using the 3-D k-nearest neighbourhood method. Further, surface roughness was determined as the standard deviation of the orthogonal fitting residuals for local planarity analysis. The authors next used the similarity of the normal vector and the 3-D distance to the seed point in a region-growing algorithm for plane surfaces. Finally, using the plane's slope, aspect, and area attributes, the plane segments were classified to identify the building roofs. The technique showed fast handling of the data but roof planes below a certain height ($< 1.5$ m) and area ($< 6$ m$^2$) were not detected. This technique works seamlessly for larger planes and non-occluding building parts. However, roof artefacts (e.g., dormers and chimneys) were not extracted due to the merger of adjacent candidate regions.

### 3.4.2 Review of Methods using Multisource data

Various integrated methods have been introduced in the literature in which multiple data sources are used for roof plane extraction. Typically, two main approaches are taken to the segmentation of roof surfaces: the top-down approach first identifies the building periphery and then detects the roofs and other primitives, while the bottom-up approach detects planar patches first and then determines the building regions. Irrespective of the approach taken, the ultimate aim is to extract the building roof surfaces. This section provides an overview of method of building roof detection.

Awrangjeb et al. [17] proposed an automatic data-driven method for roof plane extraction using LiDAR data and orthoimagery. The method extracted the line segments from a grey-scaled version of the image using the canny edge detector. It then classified all the extracted lines as edge-, ridge-, ground-, and tree-lines. The authors used the ridge lines to determine the seed point and applied the region growing technique for segmenting the point cloud into planar segments. They used the NDVI and entropy of the image in a rule-based refinement procedure to remove the planes on trees and non-building objects. The technique applied a height threshold of 2.5 m to remove LiDAR points on the earth and smaller objects, including roadside furniture and bushes. This technique offered high accuracy but is unable to detect small-sized primitives due to the large threshold values for area and width parameters. Moreover, roofs which were partially occluded in complex areas and in shadow were also not detected.

In the technique proposed by Khoshelham et al. [91], split and merge technique was adopted for roof extraction using both aerial imagery and height information.

They addressed over- and under-segmentation issues in image segmentation by integrating the height values from the DSM with the image segmentation algorithm. They exploited a simple observation to resolve the segmentation issues. In the case of under-segmentation, multiple regions were detected as single over-grown region, but in reality these regions had different height values and the opposite is the case for over-segmented regions when multiple segments had the similar height values. The authors evaluated the performance of the method through the *RMS* error using four simple gable roofs. Only one metric for the evaluation is not found to be convincing as the avoidance of segmentation issues could have led to other errors.

Since the detection of roof planes is considered a sub-task in 3-D building reconstruction, a survey of many other techniques in the context of roof extraction is provided in the following Section 3.5.3. Regarding the integration approach, Sun [111] argues that although combining information from multisource data has several advantages, it becomes problematic to determine the correspondences between different types of data for the detection of roof planes. It also becomes difficult to assess which particular information or what level of amalgamation should be achieved for the development of a generic approach.

## 3.5 3-D Building Reconstruction

The fundamental task of building modelling is the transformation of low-level building primitives to a high-level model description. In the research literature, a wide range of methods for reconstruction of buildings have been proposed [21, 23, 30, 31, 35, 52, 65, 112–121]. Theoretically, these techniques differ from each other in the generality and degree of automation, the data sources, the geometric modelling methodologies and the strategies to achieve GIS building models. This section provides a review and discussion of recent developments in the reconstruction of 3-D buildings based on the input data sources used, as shown in Figure 3.1.

### 3.5.1 Review of LiDAR-based Methods

Vosselman [23] proposed a building reconstruction technique using 3-D Hough transformation for the extraction of roof surfaces from *Airborne Laser Scanning* (ALS) data.

The edges were recognised through cross-intersection of the extracted surfaces and analysis of the discontinuities in Delaunay triangulation of the original height points. A roof topology was built assuming buildings were polyhedral and their edges corresponding to height discontinuities were either parallel or perpendicular. The large gaps in the reconstruction were resolved by building regularity constraints and manual intervention. This strategy has the disadvantage that adjacent vegetation is considered as part of the building and partial occlusion by nearby trees also hampers the reconstruction process.

In a subsequent approach, Vosselman et al. [115] used ground plans and the 3-D Hough transformation for building reconstruction. The ground plans were used to obtain the accurate location of the outer roof face edges and information about the structure of a building. The first strategy explored the Hough Transform for the detection of intersection lines and planar faces. In the event that building outlines were unavailable, they were drawn manually in a display of the laser points with grey-value-coded heights. The second strategy adopted some predefined simple roof models (flat, shed, gable, hip, spherical, or cylindrical) that were refined on the basis of fitting them into the input point cloud data. The first strategy is often deficient in finding the intersection lines or the height jump edges and therefore fails to refine the initial ground plan. However, the second strategy is unsatisfactory while working with small details of buildings. Moreover, some unnecessary extensions in the building regions are observed, mainly due to misalignment between the ground plan and the corresponding laser data.

Kim and Shan [117] proposed a novel roof plane segmentation and building reconstruction technique using airborne LiDAR data. A segmentation process based on a multiphase level set was applied to extract the roof planes that used point normal for local planarity analysis and exclude the non-planar points lying on roof ridges or step edges. The reconstruction of buildings was then performed in two stages. The roof structure points were first determined through the intersection of adjacent roof planes or line segments of the building boundary and repositioning the structure points based on their topological relations inferred from the segmentation results. Although, this technique shows good results, it suffers from over-segmentation and neglects the effect of vegetation in the segmentation process.

Sohn et al. [122] proposed a generative modelling approach to reconstruct 3-D polyhedral building models using LiDAR point cloud. The coarse boundaries of the buildings were detected and the tree segments were eliminated at the beginning. The building points were then partitioned into homogeneous rooftop regions based on height and plane similarity criteria using the RANSAC algorithm. The step edges and the plane intersecting lines were identified using a *Compass Line Filter* (CLF) and the intersection of the adjacent roof planes, respectively. A *Binary Tree Partitioning* (BSP) tree was used to produce initial rooftop vectors using the topological relationships between the adjacent planar surfaces. Finally, *Minimum Description Length* (MDL)-based regularisation was performed by rectifying the geometrical distortions between adjacent roofs and the intersection lines, and then an optimal rooftop model was reconstructed. The study showed good evaluation results and is suitable for updating cadastral maps, but the extracted models tend to shrink in comparison with the reference vectors which are digitised using the multispectral image. This method also shows low geometric accuracy due to over-segmentation and the applied error tolerance. In addition, the occurrence of deformation in rooftop models is observed when the curved segments are extracted as linear segments. Some initial work and a similar study have been published in [123] and [124], respectively.

Kong et al. [36] proposed a classification method based on the k-plane clustering algorithm. The point cloud data were used to obtain the clustering objects and real intersecting lines which were mapped directly onto the *xy*-plane. Using the location of intersecting lines on the *XY*-plane, the point cloud data of the obtained clusters were segmented to identify the building roof planes. By employing the largest polygons composed of the intersecting lines and boundary lines, 3-D models of the buildings were reconstructed. Since the method since adopts a plane as the clustering model, it is unable to classify the laser points for curved building roofs. Moreover, information about the initial clusters and the normal vectors of the neighbouring planes need to be generated in advance for successful clustering of the objects.

Poullis and Suya [125] presented a reconstruction pipeline method for building detection and the production of high-fidelity geometric models using LiDAR data. The input LiDAR was first interpolated and subdivided into a grid space. Next, the neighbouring points which have similar geometrical properties were grouped into homogeneous regions describing the roof planes. These regions from laser points were extracted through the segmentation of maps following a plane-fitting technique. They used the *Gaussian Mixture Model* (GMM) to classify the boundary points into different orientations. Through experiments, the authors show that their method is capable of delineating linear as well as non-linear boundaries. However, no detail is provided on building roof topology in their results.

Jung et al. [126] presented a data-driven reconstruction technique to develop 3-D rooftop models at city-scale from ALS data. This technique does not differ much from their earlier work [122]. Using height- and plane-similarity, building-labelled laser points were clustered into homogeneous regions. In the next stage, external outlines, intersection lines, and step lines of the planar segments were extracted as part of the linear modelling cues for reconstruction purposes. The topology relationship among the modelling cues were recovered through the use of the BSP technique. The building rooftops were modelled using an implicit regularisation process based on *Hypothesise and Test* (HAT) optimisation in the MDL framework. The parameters governing the MDL optimisation were approximated using min-max optimisation and entropy-based weighting methods which were used for selecting an optimal model from the possible hypotheses. Although the experimental results show good performance for large objects, some small roof planes were not detected, and were therefore, not reconstructed. This method suffers from under-segmentation issues since many roof planes were merged into their adjacent clusters. In addition, this study lacks a discussion of the reconstruction of partially occluded buildings and their detection.

Wu et al. [127] offer a graph-based technique to reconstruct urban building models from airborne LiDAR data. In their paper, they represented buildings as topological structures using a graph theory-based localised contour tree method. The contour tree was then split into different individual parts for surface modelling by analysing their topological relationships. The technique further constructed a weighted bipartite graph between any two adjacent contours and solved the correspondence problem for surface modelling using bipartite graph matching. Finally, building models were reconstructed by gluing all individual parts of the building models obtained from bipartite graph matching into a complete model. Although this technique provides descriptive models, it fails to capture the sides of buildings and produces high geometric distortion. Similarly, multi-storey buildings and buildings constructed with transparent materials were unable to be reconstructed, and therefore, the technique produces low completeness and high modelling errors.

Rottensteiner et al. [112] proposed a 3-D building reconstruction technique using airborne LiDAR data. They detected roof planes using surface normal vectors and subsequently determined the intersection lines and step edges among the adjacent roof planes. Finally, the intersection lines and the step edges were combined to construct 3-D building models. More recent contributions to 3-D building reconstruction by Rottensteiner can also be found in [22, 113, 114]. In addition, Rottensteiner et al. [22] have reported comparative research results for urban object detection and 3-D building reconstruction using the ISPRS benchmark datasets. They selected fourteen different building reconstruction methods, of which ten methods were based on LiDAR point cloud, two methods used images, one method employed a raster DSM from ALS data, and one method used both image and LiDAR data. The performances of these techniques were evaluated on the basis of different quality matrices. Readers interested in more in-depth knowledge are encouraged to read the case study [22].

### 3.5.2 Review of Image-based Methods

For image-based roof reconstruction, researchers have proposed a variety of techniques. One goal common to all the techniques is the identification of the 3-D intersection points and this is where different approaches and their variations have been developed. This section provides a brief description of some salient building reconstruction methods.

Noronha and Nevatia [120] proposed a 3-D wireframe roof modelling technique using multiple aerial images. The authors chose a dual-phase modelling procedure for hypothesis generation and verification of flat and symmetric gable roof structures. Initially, hypotheses for rectangular rooftops were generated by grouping the extracted lines in the images hierarchically. Next, junction points were determined through matching and grouping them into different classes of parallel or U-shaped line segments. These lines and junctions were utilised to generate roof hypotheses and were verified by searching for the presence of predicted walls and shadows. The verified hypotheses were finally combined to reconstruct the 3-D roof model. The authors show that occlusion and shadows limit the performance because of overlap mismatch when different viewpoint images were used.

Sportouche et al. [31] presented a semi-automatic building reconstruction technique using DTM, high resolution optical, and SAR images. The authors argued that the amalgamation of optical images with SAR is a difficult task and therefore favoured the intervention of an operator in a restricted manner. First, individual rectangular buildings were distinguished using a region-based segmentation technique, followed by a boundary refinement procedure utilising a contour-based approach. Next, the optical footprints were projected and registered in the SAR data in order to obtain a fine superposition between the optical and SAR features. The last stage was to retrieve height and validation of the identified buildings based on the optimisation of two SAR criteria. The authors chose a confidence score as the evaluation parameter which was higher for most of the reconstructed buildings. However, they have been used many other user-defined parameters for different datasets that make the robustness of this approach questionable. The technique was tested using only rectangular and flat rooftops. Therefore, reconstruction of complex buildings is not found to be convincing using the proposed arrangements.

Arefi and Reinartz [128] introduced a building reconstruction method integrating DSMs and orthoimagery. Firstly, building ridge lines were extracted using orthoimagery and height information in a ridge-based decomposition process. Next, parametric rooftops were reconstructed using a projection-based algorithm for each ridge line by projecting the 3-D points onto the 2-D plane which was defined by the orientation of the processed ridge line. The authors then fit a predefined 2-D model into the data and back-projected to the 3-D space. Further, these parametric and prismatic models were merged and the coinciding nodes and corners were refined in a post-processing stage to form a final 3-D model of the building. This method is quite sensitive to the location and accuracy of the extracted ridge lines and a model cannot be reconstructed if there is no ridge line available. Occlusion with nearby trees is another limitation of the technique, which causes incorrect detection of the building walls.

Haala [30] used a DSM to detect building areas for building reconstruction. The author utilised height isolines to decompose the DSM. For each region, size and compactness parameters were computed and the regions with building-like features were selected for building reconstruction. The straight lines were extracted from the image and a stereo matching was performed using the height information. Based on the assumption that the ridge-line of the roof represents the dominant direction of the building and has the maximum elongation, a prismatic building model was fitted to the candidate 3-D edges to obtain the reconstructed object.

Yu et al. [33] reported an automated data-driven reconstruction technique using *Terrestrial Laser Scanning* (TLS) images. They aimed to describe an entire 3-D building by using predefined grammar e.g., Lindenmayer systems (L-systems) and the use of the *Maximum A Posteriori* (MAP) principle. First, the geometry handling process converted the unstructured geometry data into structured geometry in terms of polygons. Then, the mapping-relationship between the structured geometry objects and semantic building objects were determined by the grammar-based reconstruction procedure. Finally, the MAP estimator was used as part of the rule selection strategy to assess the goodness of fit between the observed data and the selected model, as well as the complexity of the fitted model, in order to determine the best hypothesis.

### 3.5.3 Review of Methods using Multisource data

In the context of integrating more than one data source for roof reconstruction, researchers have proposed a variety of techniques. These techniques integrate the features obtained from different sources and attempt to extract building primitives, locate 3-D intersection points, and approximate the roof topology for the construction of building models. This section presents a summary of some building reconstruction techniques.

The approach of Brenner and Haala [28] integrated both DSMs and 2-D ground-plans data sources in a semi-automatic reconstruction process. They used a heuristic algorithm to decompose the ground plans into rectangular regions. For each region, several 3-D parametric primitives were instantiated and their optimal values were approximated. Then, a primitive with the least fitting error was selected using the area and slope threshold parameters. The selected 3-D primitives and their parameters were later refined using aerial images by a human operator during the semi-automatic post-processing stage. The authors provided an interactive mode to modify or add the ground plan rectangles and estimate a best matching 3-D primitive. Finally, all the 3-D primitives were merged to obtain the reconstructed object.

The semi-automatic technique of Wang et al. [129] offered a building reconstruction from LiDAR data and aerial images by introducing floating models. In this technique, the operator made a model choice and estimated its alignments to the building boundary on the aerial photographs. Next, the model's optimal fit parameters were assessed using aerial photographs and the model's vertical parameters were determined by LiDAR data using an iterative least-squares model-data fitting algorithm. Finally, using the model parameters and standard deviations, the wire-frame-model was reprojected onto all the overlapping aerial photographs to enable the operator to further modify the results if necessary. The operator actually selects a model, which is rectilinear and approximately fits it. This study is limited to the construction of simple buildings, and for complex buildings, the design of more model types is required.

Seo [130] used LiDAR data as a combination of wing models for generating building model hypotheses. The author used surface patches as the fundamental primitives and stored their adjacent relationships in an adjacent relationship graph. A wing model was initiated with an antisymmetric plane and the remaining patches were next examined to determine if they can be grouped into any existing wing model. Building models/hypotheses were then generated by combining the wing models. These hypotheses were finally verified by evaluating the consistency between the model hypothesis and the corresponding aerial imagery. However, the parameter optimisation procedure was not explained in detail by the author.

Suveg and Vosselman [131] utilised existing ground-plans with LiDAR information for building reconstruction. They proposed both data-driven and model-drive methods. In their data-driven method, initially, the building plans were decomposed into polygon segments. Next, the LiDAR points for each polygon were extracted and supplied to a least square plane fitting method to compute the parameters of the planes. The authors investigated the topological relationship between the plane surfaces by determining their mutual intersections. Their model-driven technique begins with the generation of a building hypothesis. 3-D lines were then extracted to determine the building edges and the were projected back onto the images. Finally, the building hypotheses were verified with the gradient analysis of these projected edges. The main limitation of the proposed methods is their restriction to the use of an external ground-plan.

Zhang et al. [43] introduced a model-driven 3-D building roof modelling technique combining both LiDAR point cloud and aerial imagery. The fundamental concept was to characterise the building rooftops by parametric primitives and develop cost functions utilising the information from both the input data. The cost function was then minimised by the use of shape parameters present in the library. The purpose of the minimisation function was to transform the 3-D modelling problem into a optimisation problem. Since the authors choose a rectangular shape for the building primitives, buildings with complex structures or non-rectilinear shapes cannot not be detected and hence reconstructed.

Kwak [44] proposed a hybrid building modelling technique integrating both LiDAR point cloud and aerial images. The method decomposed a complex building into rectangular primitives using the *Recursive Minimum Bounding Rectangle algorithm* (RMBR). Next, the parameters associated with the extracted primitives were approximated from LiDAR data. Building edges were identified from the aerial images. Finally, the model primitives were adjusted using edges through the least-squares adjustment procedure, i.e., model-based image fitting. However, this method cannot reconstruct buildings with hipped roof structures, because the fundamental primitive used for the reconstruction was a rectangular primitive.

## 3.6 Research Challenges

Since the beginning of the development of techniques that aim to provide solutions for building detection, roof plane extraction, and 3-D building modelling, there has been a clear path of progress, starting from early techniques that solve the problem of occlusion, shadows, vegetation elimination, and urban object differences using naive approaches, moving to 3-D building modelling that allows the reconstruction of urban

GIS objects. However, some research gaps can be focused on to improve the effectiveness of such techniques. This thesis addresses some of these identified shortcomings as follows:

- **Building detection**: There is a class of building detection techniques which either neglect the effect of vegetation altogether or consider only the constrained presence of vegetation. Another class of techniques either applies large height threshold or uses large windows in morphological filters to remove unwanted GIS objects, while others utilise ground-plans to address only the region of interest. Due to spectral ambiguities, many image-based techniques have difficulty distinguishing buildings with coloured rooftops from the connecting vegetation. As described in Section 2.2.3, since LiDAR data have discrete laser pulses and sparsity, building outlines normally have low planimetric accuracy. In contrast, buildings extracted from imagery data have high planimetric accuracy but limited vertical accuracy due to the unavailability of accurate height information. To overcome these issues, a robust technique for automatic building detection and regularisation is introduced in Chapter 4. We consider the complementary advantages of LiDAR and imagery data, and choose the fusion of two sources as a promising strategy to increase the building detection rate and the planimetric accuracy of the building regions.

- **Roof plane detection**: Most often, roof plane extraction techniques choose to apply constraints on plane size, area, and orientation to address forefront challenges because of point cloud sparsity, urban object differences, surrounding complexity, and high spectral variability, which adversely affect their detection performance. Some methods use colour features for segmentation and therefore lack in detecting the rooftops constructed using unspecified colours. Furthermore, techniques based on clustering and region-growing approaches have difficulty in determining an initial cluster count and a seed region respectively, resulting in false segmentation and high computational expense. Another group of techniques focus on simple rooftops e.g., flat surfaces, and are therefore unsuitable for complex roof structures. On the other hand, integrated techniques have not been proven successful in providing promising results as finding a correct correspondence among the different features becomes problematic [111]. By using LiDAR data in a data-driven fashion, the technique proposed in Chapter 5 not only detects polyhedral buildings with diverse roof types but also roofs which are partially occluded, situated in shadow or have colourful surfaces. The proposed technique provides a better interpolation of roof regions where multiple surfaces intersect creating non-manifold points. As a result, these geometric features are preserved to provide automated identification and segmentation of roof planes from unstructured laser data.

- **Building reconstruction**: The challenges of building reconstruction have been addressed in previous studies. However, the quality of reconstructed models appears to be limited by the accuracy and quality of the input data, resulting in several low quality measures such as completeness, correctness or planimetric accuracy. Further, providing the specifications of different input models does not warrant that the reconstructed model has high geometric accuracy [65]. Many general parameters, including minimum footprint size and positional accuracy values, are used by several techniques to develop building models with certain level of detail [132]. Many existing approaches have demonstrated promising results in building reconstruction, but there are still a number of issues to be improved. For instance, in segmentation, segmented roofs are mostly disconnected, causing difficulty in determining the neighbourhood relationships among the roof planes. Furthermore, locating step edges only from LiDAR data is also hard and often requires additional information or constraints. In addition, the approximation of roof patches, which are generally missed because of the low resolution of the LiDAR data, requires the operator to make assumptions and often produces high reconstruction errors. To resolve these issues, Chapter 6 introduces a data-driven 3-D reconstruction technique that constructs buildings represented at lower levels with coarse boundaries (3-D roof-planes) to the higher levels (3-D building models).

## 3.7 Summary

This chapter has reviewed techniques of building detection, roof plane extraction, and 3-D building reconstruction. These techniques were categorised based on the type of input data sources, including LiDAR, imagery or both. In the review, the execution stages of different techniques were probed and their performances analysed to identify their limitations in performing the designated tasks. These limitations include: (1) inability to detect partially occluded buildings and roof surfaces, (2) limited robustness to detect small-sized and low height buildings, (3) difficulty in recognising buildings in shadows, (4) unavailability of effective theories to eliminate the vegetation from regions of different complexities, (5) inadequacy in overcoming noise and outlier effects, and (6) lack of a sound foundation for the integration of multiple data sources.

In the following chapters, we propose solutions for automatic building detection and footprint generation, roof plane extraction, and 3-D building modelling and scene reconstruction to address the highlighted limitations. Furthermore, a detailed performance and comparative analysis is also provided using a variety of datasets in each chapter.

Chapter 4

# Building Detection and Boundary Regularisation

> "Science knows no country, because knowledge belongs to humanity, and is the torch which illuminates the world. Science is the highest personification of the nation because that nation will remain the first which carries the furthest the works of thought and intelligence."
>
> Louis Pasteur

## 4.1 Introduction

The automated extraction and localisation of urban objects is an active field of research in photogrammetry with the focus on detailed representation. The literature survey indicates that the success of most existing detection methods relies on the quality of the DEM and the accuracy of co-registration between multisource data. They often impose constraints on several features such as height, area, and orientation to distinguish different urban objects and remove vegetation. These methods have been observed to be unable to address buildings which are small in size, in shadow or partly occluded. Furthermore, the building outlines produced by these data-driven approaches are generally ragged. Therefore, the research objectives of this chapter research objective (aligned to **RO2**) are to develop a strategy able to:

- deal with moderate misregistration between the LiDAR point cloud and the corresponding orthoimagery,

- identify buildings which are partially occluded and in shadow,
- extract small buildings without affecting larger ones, and
- generate regularised and well-delineated building boundaries.

This chapter concentrates on building detection and boundary regularisation using multisource data. It includes a comprehensive evaluation and analysis of a wide range of test datasets. These datasets differ in scene complexity, vegetation, topography, building sizes, and LiDAR resolution (1 to 29 points/$m^2$). We further incorporate LiDAR's point density feature in different processes to make the proposed technique flexible and robust in relation to multiple data acquisition sources, e.g., airborne and mobile laser scanning systems. An adaptive local height threshold is utilised in detection for fine delineation of building boundaries. Moreover, a new boundary regularisation technique is also introduced which generates 2D building footprints using spectral information (image lines) assuming buildings are rectilinear.

To meet the set objectives and evaluate the proposed technique, it is tested using the ISPRS (German) and the Australian benchmark datasets. Compared to the ISPRS benchmark, the Australian datasets are far more complex and challenging due to the hilly terrain (Eltham and Hobart), dense vegetation, shadows, occlusion, and low point density (1 point/$m^2$). Often, the buildings are covered by nearby trees or shadows, as indicated by some real scenarios in Figure 4.1. The evaluation study, particularly on the Australian datasets, demonstrates the robustness of the proposed technique in regularising boundaries and separating partly occluded buildings from connected vegetation and detecting small buildings as well as larger ones.

**Figure 4.1** Complex scenarios in Australian datasets: Partly occluded and shadowed buildings.

The rest of the chapter is organised as follows: Section 4.2 details the proposed building detection and boundary regularisation techniques. Section 4.3 presents the performance study and discusses the experimental results using five test datasets, followed by a comparative analysis. Section 4.4 concludes the chapter.

## 4.2 Proposed Method

The proposed technique identifies the candidate building regions and subsequently segments the regions into grids. Next, vegetation elimination, building detection and extraction of their partially occluded parts are achieved by synthesising the point cloud and image data. Finally, the detected buildings are regularised by exploiting the image lines in the building regularisation process. The workflow of the proposed technique has three stages, as sketched in Figure 4.2. In the data pre-processing stage, several data metrics are generated to be used at later stages, namely the building mask and height difference data from the ALS, and entropy, NDVI, and image lines from the orthoimagery. In the building detection stage, the candidate building regions are identified following a proposed *graph-based line clustering process* to remove the superfluous objects. Next, the buildings and their parts occluded or in shadow are identified and vegetation is eliminated using a proposed *cell-based clustering process*. Subsequently, the detected area is enlarged to reduce misalignment between the aggregated data. Finally, the building footprints are generated using a new proposed *building regularisation process*.

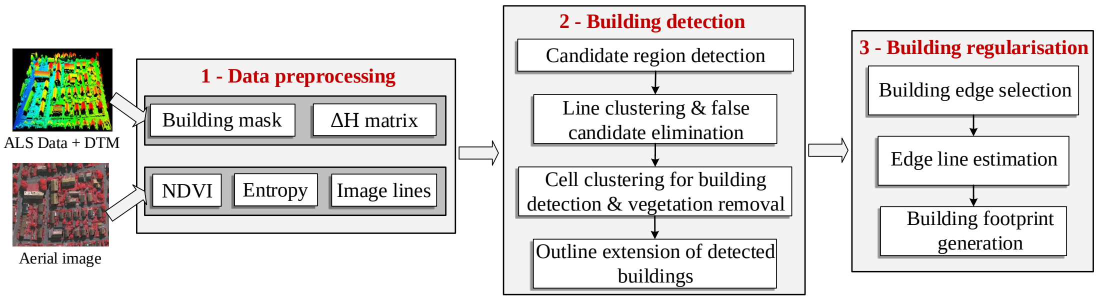



**Figure 4.2** Workflow of the proposed building detection and regularisation technique.

### 4.2.1 Data Pre-processing

The proposed technique takes ALS data, orthoimagery, and a DTM as inputs. 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 [133].

#### 4.2.1.1 Test Data

Figure 4.3(a) presents a test dataset, AV2, which was introduced in Section 2.6. The AV2 dataset can be regarded as a moderately vegetated area containing complex neighbourhood structures, small and occluded buildings, shadowed, impervious and transparent roofs, and flat and gabled structures. It is, therefore, selected to evaluate the performance and illustrate the different processes of the proposed technique.

The test ALS dataset has a point spacing of 0.17 m with 29.3 points/m$^2$ and the corresponding RGB colour orthoimage has a resolution of 0.05 m. The available orthoimages for the Australian sites were created using DTM. These datasets were registered using the mutual information-based method [134], which exploits the statistical dependency between same- and multi-modal datasets to produce a similarity measure and uses LiDAR intensity in simultaneous registration. Nevertheless, the building roofs and tree tops are considerably displaced with respect to the LiDAR, owing to the absence of true orthophotos.

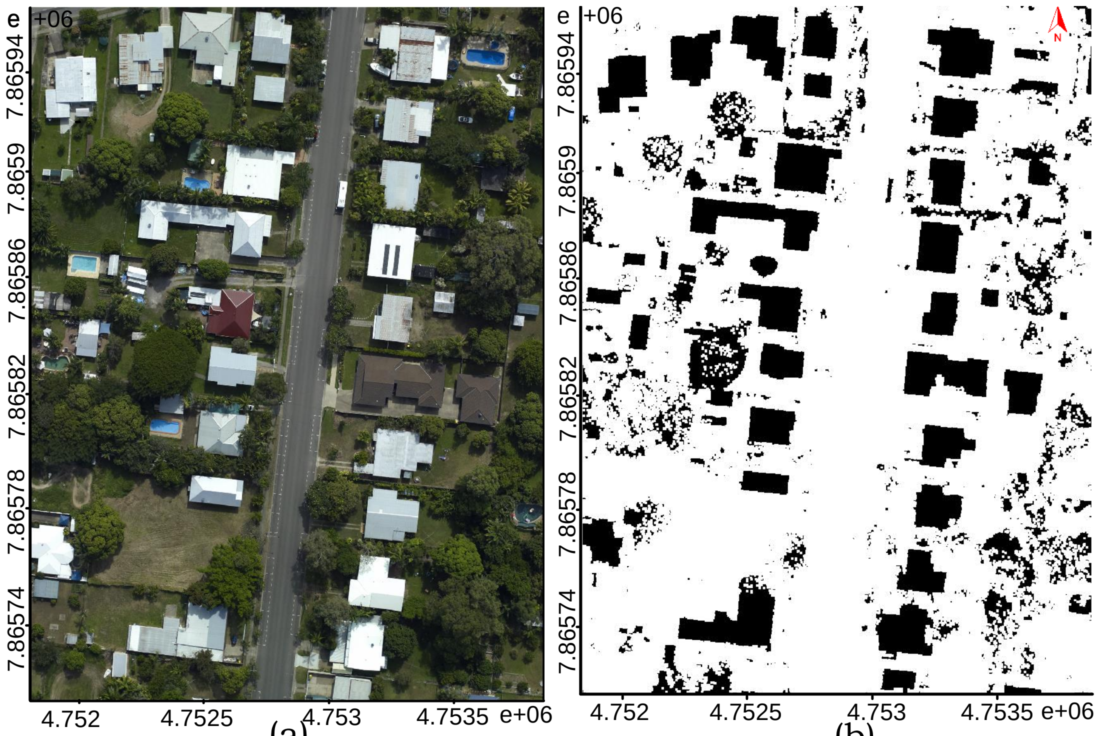



**Figure 4.3** Sample data: (a) RGB orthoimage and (b) Building mask.

#### 4.2.1.2 ALS-Generated Data

A *building mask* is generated using the ALS data and ground height from DTM to divide the point cloud into the ground and non-ground points. For each point, a height threshold is calculated as $h_t = h_g + h_{rf}$, where $h_g$ is the ground-height taken from DTM, while $h_{rf}$ is a relief factor that separates low height objects from large height objects. For this study, $h_{rf}$ was set to 1 m to keep the low height objects [37], which classifies many points on bushes and low height trees as non-ground points. Figure 4.3(b) shows the building mask, where black regions correspond to elevated objects (buildings and trees) while white regions represent bare earth, including roadside furniture, cars, and bushes.

Moreover, height difference data are generated by dividing the ALS data into a uniform grid twice the ALS data spacing in magnitude. The method is explained with a sample image taken from the test dataset, as shown in Figure 4.4(a). If a cell contains only one laser point, the elevation of the point is assigned to the cell. Otherwise, the elevation of the laser point which is closest to the centre of the cell is assigned to the grid (see Figure 4.4(b)). If a cell has no laser point, the elevation of the grid is assigned zero height. Finally, the *average height difference* $\Delta H$ of each cell is computed by averaging the elevation differences of 8 connected neighbouring cells. $\Delta H$ is used by the *cell clustering process* to delineate buildings and identify vegetation.

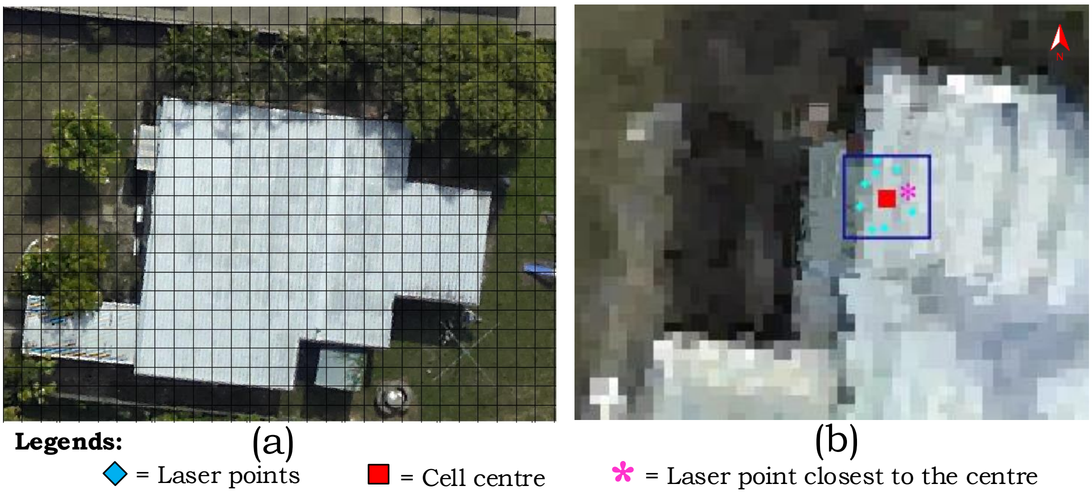



**Figure 4.4** (a) Grids overlaid on a sample image and (b) LiDAR points within a grid cell.

#### 4.2.1.3 Aerial Image Generated Data

NDVI has been used extensively in the research literature to eliminate vegetation and classify scenes [14, 42, 49, 55]. Nevertheless, many authors [14, 17] emphasise combining it with entropy, since NDVI alone is not a promising feature to handle shadows and coloured buildings. Therefore, texture information in the form of entropy [135] is utilised on the basis of the observation that trees are rich in texture and have higher surface roughness than building roofs [14]. If multispectral orthoimagery (RGBI) is not available, a pseudo-NDVI is then calculated from a colour orthoimage (RGB) following the process explained in [15, 17, 54], which assumes the three colour channels in the order of I-R-G to use the standard NDVI formula. Henceforth, the term NDVI is used to refer to both NDVI and pseudo-NDVI.

Furthermore, using the method explained in [17], the image lines are extracted from the test input image and categorised into *edge* (building border), *ridge* (intersection of two roof planes), and *ground* classes sketched with *blue*, *cyan*, and *red* colours, respectively, in Figure 4.5. The structural lines define the shape and direction of a building. These can be used as a cue to differentiate vegetation and buildings and later, to generate building footprints in the building regularisation stage.

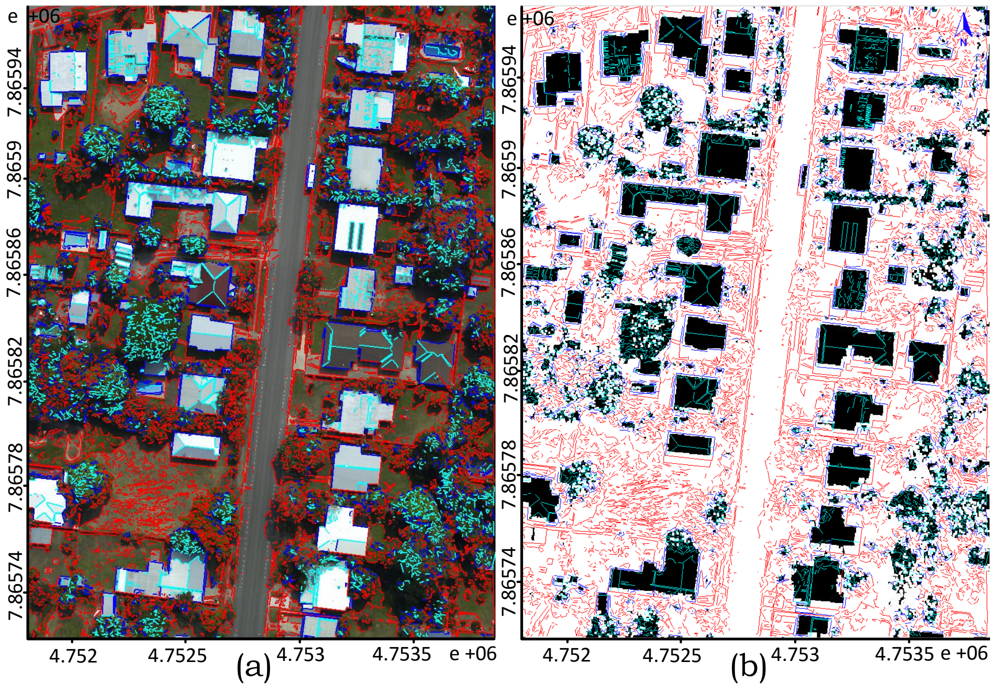



**Figure 4.5** Image line extraction: (a) All classified lines on RGB orthoimage and (b) All classified lines on the building mask.

### 4.2.2 Building Detection

In this section, the building detection stage shown in Figure 4.2 is presented in more detail using the test dataset.

#### 4.2.2.1 Candidate Region Detection

Candidate building regions are identified from the building mask using *connected component analysis* and their boundaries are estimated using the *Moore-Neighborhood tracing algorithm* [135], which provides a list of connected pixels of an object in clockwise order. Figure 4.6(a) shows the detected candidate regions, and each region is labelled and sketched in a different colour from its neighbour with the boundary black in colour. It is shown in Figure 4.6(b) that the extracted boundaries (from the LiDAR-based building mask) are misaligned with the image (often over 1 m) due to misregistration between the aggregated data.

Similarly, Figure 4.6(c) shows the inaccuracy of the extracted objects, as although the large buildings are detected, their boundaries are inaccurately delineated. It also shows the presence of several false objects on trees and the inaccurate inclusion of nearby vegetation in the building region. For instance, two buildings (labelled (vi)) have been wrongly detected as a single object due to dense vegetation between them. Some complex situations are shown in Figure 4.6(d) with their corresponding locations marked with labels (i)–(vi) in Figure 4.6(c).

#### 4.2.2.2 Line Clustering and False Candidate Elimination

To associate the lines (from the image) and the candidate building regions (from the LiDAR-based building mask), a graph is constructed as part of this process. Each pixel of the building mask corresponds to a node. An edge exists between a pair of nodes if a pixel and its $3 \times 3$ neighbouring pixels are *black*, denoting a non-ground object. All these edges carry a weight equal to 1. However, if a pixel or the neighbouring pixels are *white* (ground object), the corresponding nodes do not have any edge. Therefore, the resultant graph $\mathcal{G}$ is disconnected where each candidate region of the building mask corresponds to a strongly connected edge-weighted graph $\mathbb{g}$, such that $\mathcal{G} = \{\mathbb{g}_1, \mathbb{g}_2, ..., \mathbb{g}_{\mathbb{n}}\}$ with $\mathbb{n}$ describing the number of candidate regions.

Figure 4.5(b) shows the misalignment issue, as we can see that edge lines (blue) are substantially away from the borders of their respective candidate regions. This issue is overcome by using a determinant point $D_{pt}$ that can be either a line's midpoint

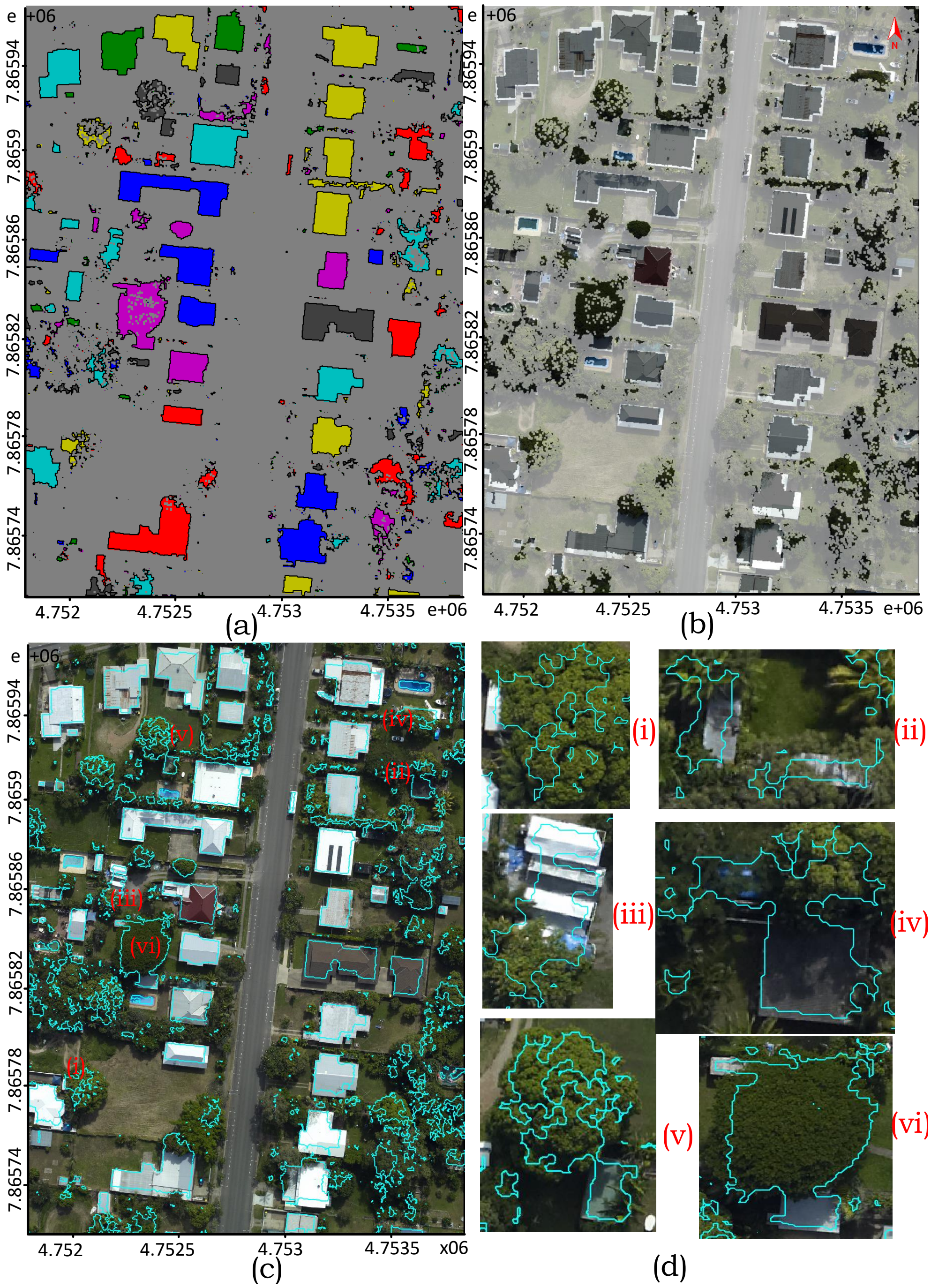



**Figure 4.6** (a) Connected components extracted from the building mask of test data; (b) Overlaying building mask and orthoimage (misalignment and squeezed boundaries); (c) Boundaries of the candidate regions sketched on input orthoimage; and (d) Complex scenes representing false boundary delineation from (i) to (vi).

$P_{mid}$, vertex, or inside-point $P_{in}$ for the clustering of image lines around the candidate building regions. An inside-point of a line is actually a perpendicular offset point (i.e., 1.5 m) from a line's $P_{mid}$ that is given by the *line extraction process* in [17]. The *line clustering process*, with reference to Figure 4.7(a), works as follows: A pool of candidate lines is created by defining a buffer region (in cyan) using uniform matrix scaling transformation. The longest line is then chosen as a *seed*, such that its $D_{pt}$ resides within the region's boundary. If no seed is found, the procedure continues to process the next building region. Finally, the candidate lines with finite distances from the seed are added to the cluster. The *Bellman-Ford algorithm*, which computes the shortest path between two nodes, is used to calculate the distances between $D_{pt}$'s of the lines. The procedure continues until the pool is empty and all the candidate regions are processed iteratively. A candidate region which fails to cluster any line is removed for further investigation. Figure 4.7(b) shows the remaining candidate building regions with their clustered lines (it displays 307 candidate regions from an initial count of 936 from the building mask for brevity).

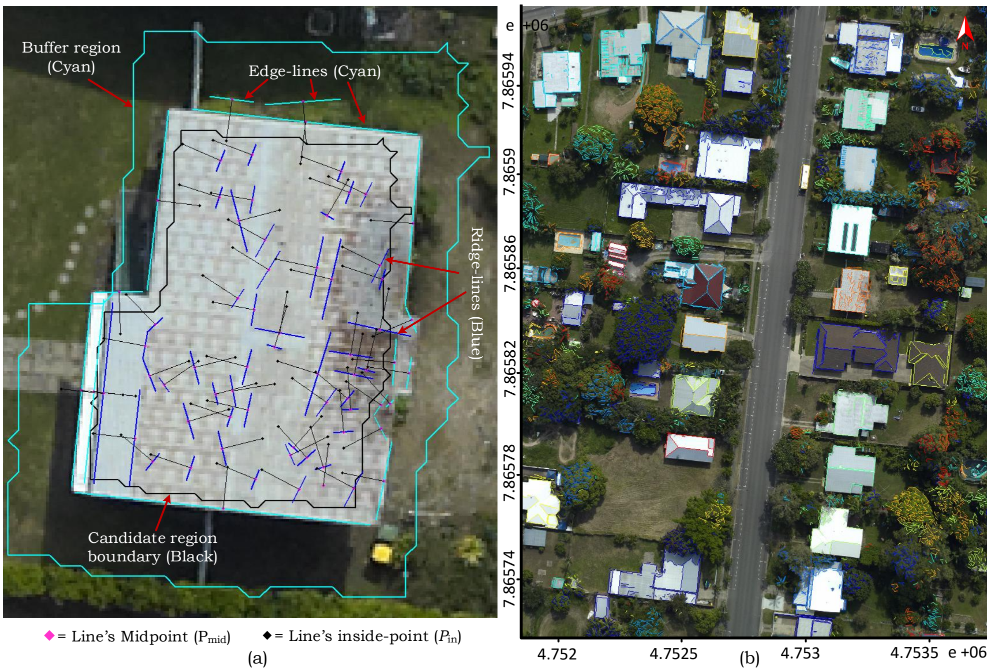



**Figure 4.7** Line clustering process: (a) Pool of candidate lines to a sample building and (b) Clusters found with their associated lines in different clustering colours.

#### 4.2.2.3 Cell Clustering for Building Detection and Vegetation Removal

It is a core process to distinguish urban objects, eliminate vegetation, and identify buildings partly occluded or in shadow in addition to larger buildings and those small in size. Features such as $\Delta H$, NDVI, and entropy are used to detect objects using a *cell-based region growing process*. To make the segmentation flexible and robust in relation to multiple data acquisition sources, e.g., airborne and mobile laser scanning systems [136], the point cloud density per grid cell $P_d$ is combined with the preceding features. The test area is divided into a uniform grid structure and $P_d$ is computed in relation to LiDAR resolution and cell size. For example, $P_d$ for the AV2 dataset is 7 ($floor(29 \times 0.25)$) using a grid of size 0.25 m with 29 laser points/m$^2$, while for the German benchmark $P_d$ is 3 with a grid size of 1 m. The process is explained as follows with reference to Figure 4.8 (sample occluded building labelled (vi) in Figure 4.6):

The cells of a candidate building region are marked as unused (blue circles in Figure 4.8(a)) and a *seed* cell is chosen i.e., an unused cell with the lowest $\Delta H$. Its neighbours are found iteratively in a region-growing fashion. Neighbouring cells are added to the cluster if their values for $\Delta H$, $P_d$, NDVI, and entropy are smaller than the user-defined thresholds (see Table 4.1). If region-growing stops, a cell with the smallest $\Delta H$ is then chosen. This becomes the new seed cell, and the process continues. Any seed that fails to grow is removed. Figure 4.8(b) shows the different clusters (shown in red) within the region's boundary conceived during the cell clustering process.

| | **Parameters** | **Values** | **Sources** |
|---|---|---|---|
| | Ground height $h_g$ | DTM height | input ALS data |
| | Height threshold $h_t$ | $h_g + 1$ m | [37] |
| | $Entropy_t$ | 8.0 | [17] |
| RGBI | $NDVI_t$ | 10 | [17] |
| RGBI | $NDVI_{t-max}$ | 15 | this chapter |
| RGB | $NDVI_t$ | 48 | [17] |
| RGB | $NDVI_{t-max}$ | 75 | this chapter |
| $\Delta H_t$ | Flat terrain | 0.4 | this chapter |
| $\Delta H_t$ | Hilly terrain | 0.8 | this chapter |

**Table 4.1** Parameters used by the proposed building detection technique.

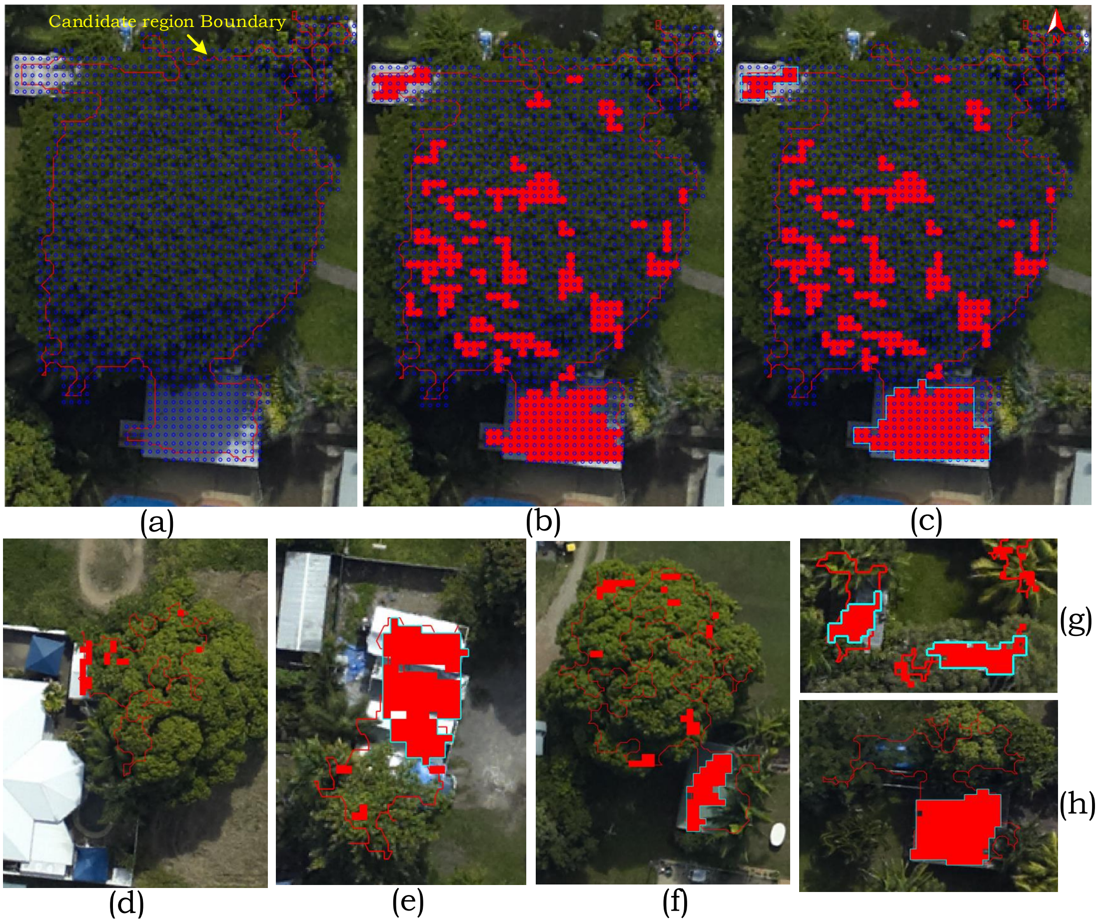



**Figure 4.8** Cell clustering process: (a) Cell collection for a candidate region (4 blue holes correspond to a cell); (b) Clustered cells (red); (c) Clusters marked as buildings (cyan boundary); and (d)–(h) Boundary delineation samples.

Due to errors in hardware sensors, a data acquisition system fails to capture the laser returns from various parts of a single object. Therefore, a segmentation process based on ALS data alone extracts these parts as different buildings and therefore, exposes the strategy to over-segmentation. To circumvent this limitation, the cell clustering process entirely uses NDVI and entropy if a maximum 25% of the total cells of a particular candidate region does not have any laser point. The seed cell is chosen with the lowest NDVI whereas a cell's neighbours are chosen based on both NDVI and entropy rather than the ALS-based features.

Finally, a rule-based procedure employing NDVI, entropy, and boundary intersection is used to identify clusters within a region's boundary as buildings or vegetation. It was found that buildings in shadow or with coloured roofs have a higher NDVI value, which can be removed if a smaller threshold is employed. Therefore, $NDVI_{t-max}$ is introduced which works with image entropy to remove vegetation without eliminating any potential building. A cluster $c$ is marked as a tree if its boundary does not intersect with the boundary of the candidate region and $\overline{NDVI_c} > NDVI_{t-max}$ OR $\overline{NDVI_c} < NDVI_{t-max}$ AND $\overline{entropy_c} > entropy_t$. Similarly, $c$ is identified as a building if its boundary intersects with the boundary of the candidate region and $\overline{NDVI_c} < NDVI_t$ OR $\overline{NDVI_c} < NDVI_{t-max}$ AND $\overline{entropy_c} < entropy_t$.

Figure 4.8(c) shows two partially occluded buildings that are identified and separated from the connected tree whereas the clusters within the region's boundary are eliminated as vegetation. Similarly, some more detected buildings either occluded or under shadows, marked earlier in Figure 4.6, are shown in Figures 4.8(d)–(h) with their boundaries in cyan while the candidate regions' boundaries are shown in red. The final building positions are obtained by the cell clustering process after the elimination of vegetation.

#### 4.2.2.4 Building Area Enlargement

Owing to misalignment and a large height difference on the building perimeter, building edges are usually under-detected (Figures 4.9(a) and 4.10(a)). This is compensated by claiming individual pixels based on the NDVI, entropy, and LiDAR height difference. An adaptive local height threshold is employed instead of a global threshold in the *area enlargement process* by taking into account buildings with gables, flat, and complex roofs. The process with reference to Figure 4.9 is described as follows:

A boundary pixel $P_b$ is first selected and its neighbouring pixels $P_n$ which are outside the current boundary (black) are determined. Subsequently, a concentric square of a grid-size (magenta) around $P_n$ is considered. If the NDVI and entropy values of $P_n$ are less than their thresholds presented in Table 4.1, and the mean height inside the square remains within $\pm 1$ standard deviation of the building height (within the building boundary), it is included as a new boundary pixel and $P_b$ is added in the building region. The building outline process continues until all the boundary pixels including the new ones are processed as described in Figures 4.9(b) and (c). The building boundary before and after the enlargement process is presented in Figure 4.9(d).

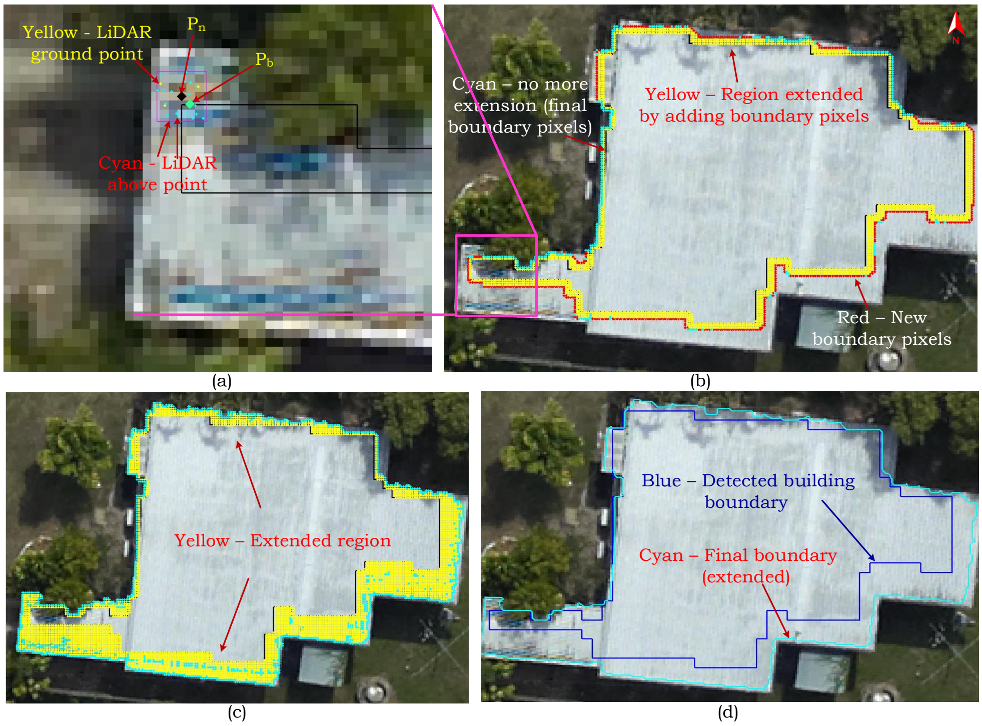



**Figure 4.9** Pixel-based enlargement process: (a) Candidate boundary pixel ($P_n$) selection; (b) Snapshot of the accumulated pixels; (c) Extended region; and (d) Detected outline and final building boundary.

Figures 4.10(a) and (b) show the building detection results before and after the pixel-based enlargement process. It is discernible that the final detected boundaries in Figure 4.10(b) cover rooftops close to the building edges and the building boundaries are detected accurately. The same boundaries are used eventually for evaluation purposes. Since the building boundaries are independent, the clustering (line and cell-based) and area enlargement processes are executed in parallel to improve their performance.

### 4.2.3 Building Regularisation

Due to point cloud sparsity, the extracted boundaries are jagged, as shown in Figure 4.10(b), which can be regularised by obtaining structural lines using the MDL

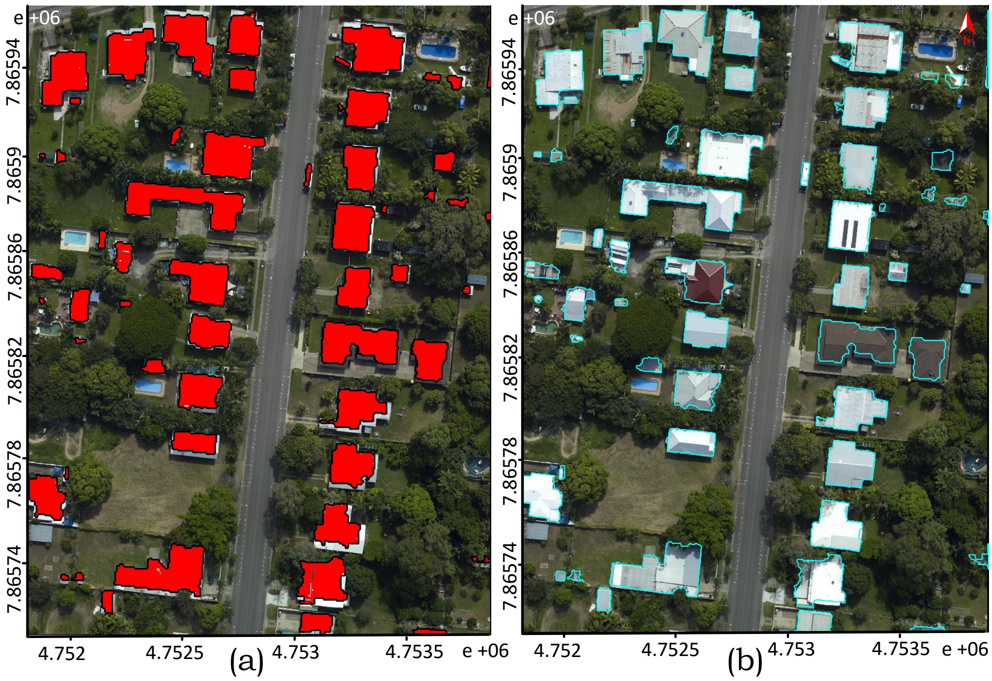



**Figure 4.10** (a) Building boundaries before enlargement (under-detected edges with black boundary and red region); (b) Final detected buildings after enlargement (cyan boundary).

method [7], although it is computationally intensive. However, the proposed regularisation method employs image lines to regularise the boundary, assuming buildings are rectilinear and adjacent edges are either parallel or perpendicular. Since all possible image lines could not be extracted due to shadows and low contrast, this makes the German dataset (Vaihingen) suitable to demonstrate the robustness of the regularisation technique. Therefore, we chose a sample scene from VH3, to explain the building footprint generation process.

The workflow of the regularisation process is shown in Figure 4.2. Figure 4.11(a) shows the boundary of the candidate building region (from the building mask) and classified image lines before the application of the building detection steps described in Section 4.2.2. Figure 4.11(b) shows the clustered lines to the candidate building region and the refined building boundary after the application of the building detection procedure. The proposed building regularisation method works in three steps: line selection, line estimation, and footprint generation.

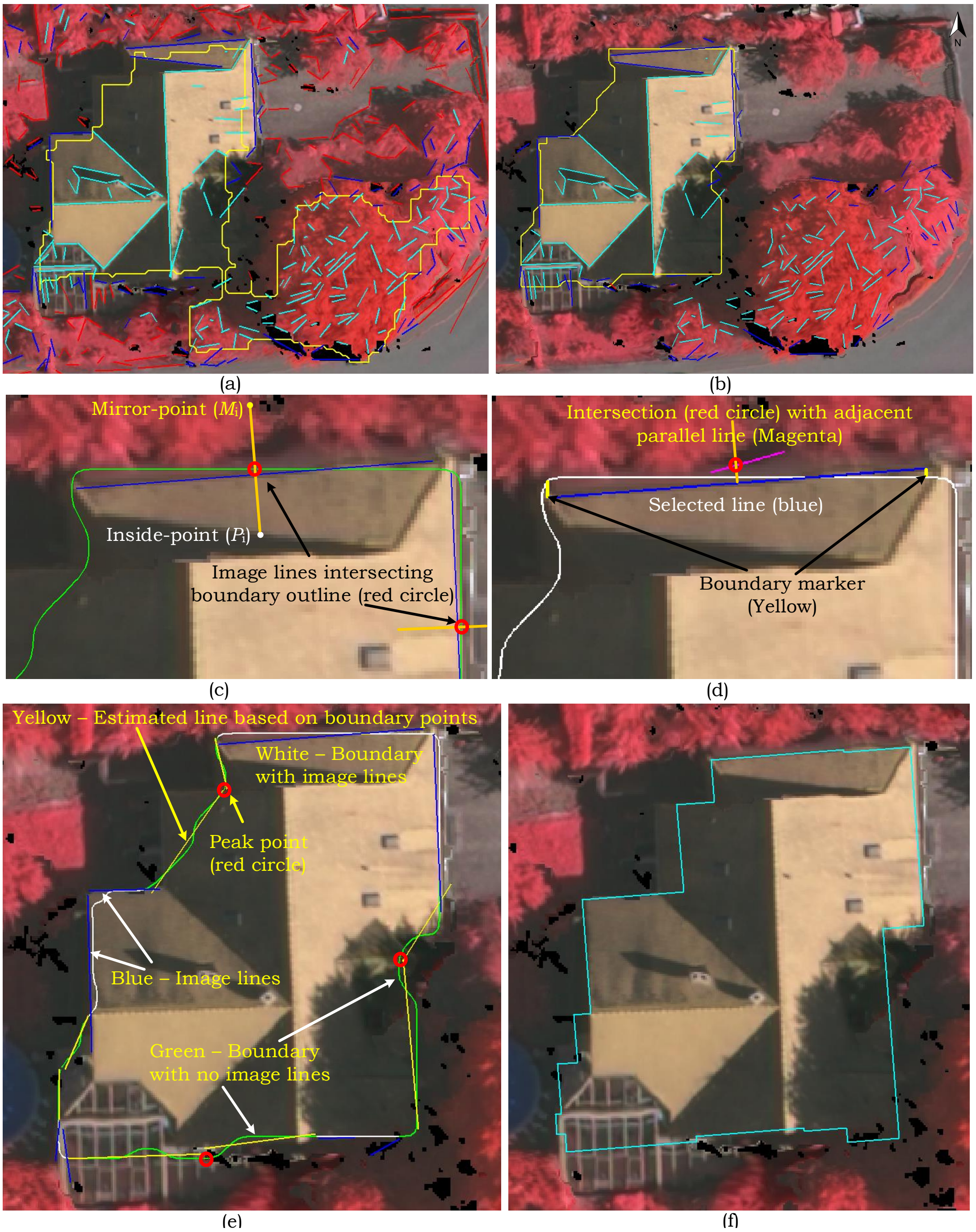



**Figure 4.11** Building regularisation process: (a) Candidate boundary region (yellow) from building mask and image extracted lines; (b) Clustered lines to candidate region and final detected building; (c) Boundary lines selection; (d) Edge selection and boundary marking; (e) Edge line estimation for unmarked boundary outline; and (f) 2-D building footprint/regularised boundary.

#### 4.2.3.1 Building Edge Selection

For each clustered line with its midpoint $P_{mid}$, we use its inside-point $P_{in}$ and mirror-point $P_{mir}$ to determine whether the line is a candidate line for the building boundary. For instance, with reference to Figure 4.11(c), the image line (blue) exists near the building perimeter, the perpendicular line from $P_{mir}$ to $P_{in}$ through $P_{mid}$ intersects the extracted boundary (green), making this line a candidate boundary line. The corresponding boundary part of the selected line is also marked. Figure 4.11(d) shows the corresponding boundary parts within two yellow perpendicular lines through the end points of the selected line. This selected line can be from any class of *edge*, *ridge*, or *ground* to overcome misalignment between the LiDAR-derived building boundary and the image-derived lines. All the candidate lines are determined iteratively. If a line fails to intersect the building boundary, it is removed. If two or more candidate lines are found for a part of the boundary, the line that has the lowest mean perpendicular distance of the boundary points to the line is selected for the boundary part.

#### 4.2.3.2 Edge Line Estimation

The edge lines are estimated only if the boundary points are left unmarked, as illustrated by the green colour boundary in Figure 4.11(e). These boundary points are first smoothed using the Gaussian function and then corner detection is carried out using the technique described in [137]. We determine the curvature peaks (red circles in Figure 4.11(e)) that specify the locations where a considerable change in curve direction occurs. A straight line is next fitted to the boundary segments using the least-square technique such that the line is rotated by a half degree clockwise/anti-clockwise around its centre to minimise the mean perpendicular distance to boundary points. The estimated edge lines and the chosen image lines (yellow and blue respectively) can be seen in Figure 4.11(e).

#### 4.2.3.3 Building Footprint Generation

Assuming buildings mainly have two principal directions along their length and width, lines at least 6 m long are considered to be long lines [138]. The long lines from the image are kept fixed. If an image line does not exist, the estimated long lines are kept fixed. In the case of single fixed line, all the other line segments are made parallel or perpendicular to it. The two closest small lines on both sides of a fixed line are adjusted first, and the next two nearest small lines are tuned with reference to the last fixed lines. This process continues until all the lines are adjusted to a parallel and perpendicular relationship.

If there exist two or more fixed lines, the small lines between them are gradually made parallel or perpendicular to their nearest fixed lines. In the case where a small line is at an equal distance from the fixed lines, it is adjusted according to the fixed line with which it makes the smaller angle. Perpendicular lines are then introduced between the successive parallel lines. Finally, the regularised building boundary is generated by intersecting the consecutive lines. The regularised building footprint for the sample scene is sketched on the input image in Figure 4.11(f).

## 4.3 Performance Evaluation

The performance of the proposed technique was tested on five datasets (introduced in Section 2.6) with different LiDAR point densities, topographies, and surrounding conditions. The ISPRS benchmark dataset, Vaihingen, Germany has three areas, whereas the other four datasets have one area captured over different geographic locations in Australia.

### 4.3.1 ISPRS Benchmark Results and Discussion

Tables 4.2 and 4.3 show the official per-object and per-area level evaluation results for the three test areas of the benchmark dataset. Figure 4.12 shows the per-pixel level visual evaluation of all the test areas (column 1) for the building delineation technique (column 2) and the corresponding regularisation outcome (column 3). Detailed quality measures for the building delineation technique before and after regularisation can be found on the ISPRS portal [139] under detection with the acronyms `Fed_1` and `Fed_2`[1], respectively.

[1] http://www2.isprs.org/commissions/comm3/wg4/results/a1_detect.html

| **Proposed Detection** | **Areas** | $C_m$ | $C_r$ | $Q_l$ | $C_{m,50}$ | $C_{r,50}$ | $Q_{1,50}$ | $1{:}M$ | $N{:}1$ | $N{:}M$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Before Regularisation | VH 1 | 83.8 | 100 | 83.8 | 100 | 100 | 100 | 0 | 6 | 0 |
| | VH 2 | 85.7 | 91.7 | 79.5 | 100 | 100 | 100 | 0 | 2 | 0 |
| | VH 3 | 82.1 | 95.7 | 79.2 | 100 | 100 | 100 | 0 | 7 | 0 |
| | **Average** | **83.87** | **95.80** | **80.83** | **100** | **100** | **100** | **0** | **5** | **0** |
| After Regularisation | VH 1 | 83.8 | 100 | 83.8 | 100 | 100 | 100 | 0 | 6 | 0 |
| | VH 2 | 85.7 | 100 | 85.7 | 100 | 100 | 100 | 0 | 2 | 0 |
| | VH 3 | 82.1 | 95.7 | 79.2 | 100 | 100 | 100 | 0 | 5 | 0 |
| | **Average** | **83.87** | **98.57** | **82.90** | **100** | **100** | **100** | **0** | **4.3** | **0** |

**Table 4.2** Object-based building detection results for Vaihingen (VH) dataset before and after regularisation stage. ($C_m$ = completeness, $C_r$ = correctness and $Q_l$ = quality, are for all buildings and over 50 $m^2$ in percentage; 1:M = over-segmentation and N:1 = under-segmentation, N:M = both over- and under-segmentation in the number of buildings).

| **Proposed Detection** | **Areas** | $C_{mp}$ | $C_{rp}$ | $Q_{lp}$ | *RMSE* |
|---|---|---|---|---|---|
| Before Regularisation | VH 1 | 84.9 | 86.5 | 74.9 | 1.2 |
| | VH 2 | 87.9 | 84.4 | 75.6 | 1.28 |
| | VH 3 | 88.7 | 85.0 | 76.7 | 1.1 |
| | **Average** | **87.17** | **85.30** | **75.73** | **1.19** |
| After Regularisation | VH 1 | 85.4 | 86.4 | 75.4 | 1.06 |
| | VH 2 | 88.8 | 84.5 | 76.4 | 1.21 |
| | VH 3 | 89.9 | 84.7 | 77.4 | 1.06 |
| | **Average** | **88.03** | **85.20** | **76.40** | **1.11** |

**Table 4.3** Pixel-based building detection results for Vaihingen (VH) dataset before and after regularisation stage. ($C_{mp}$ = completeness, $C_{rp}$ = correctness and $Q_{lp}$ = quality are in percentage, *RMSE* = planimetric accuracy in metres).

Considering all the buildings, Table 4.2 shows that the overall object-based completeness and correctness before regularisation are 83.87% and 95.80%, respectively. However, the buildings eliminated during the mask generation process, shown with green arrows in Figures 4.12(a),(d), and (g) became an issue with a relatively reduced performance, although 100% object-based accuracy was achieved on the large buildings. Since the overlap areas of the resultant and reference polygons after the regularisation process were not greatly changed, a substantial increase in per-object evaluation accuracy was not achieved. Table 4.3 shows the per-area evaluation (completeness and correctness $\simeq$ 88% and 85%, respectively) of the proposed detection before regularisation in VH2 and VH3. However, per-area completeness in VH1 was lower (84.9%) because some carports below the height threshold, marked with yellow dashed circles in Figure 4.12(b) and one in VH3 (Figure 4.12(h)), were eliminated by

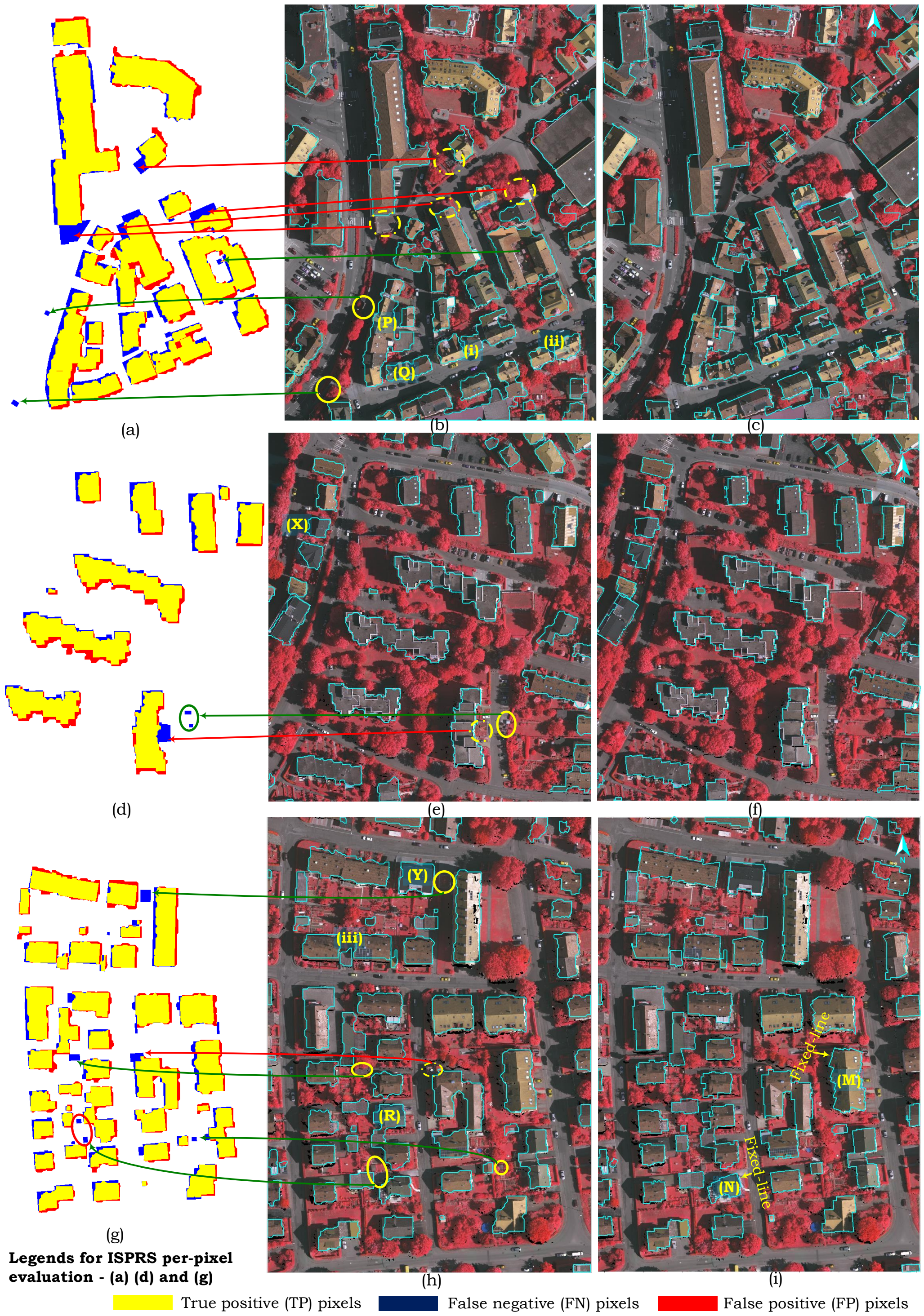



**Figure 4.12** Building detection on the ISPRS German dataset: (a)–(c) VH1, (d)–(f) VH2, and (g)–(i) VH3. Column 1: pixel-based evaluation, Column 2: boundary before regularisation, and Column 3: regularised boundary.

the building mask generation process.

During the cell clustering process, the cells lying partially on the building perimeter were included in the region. Consequently, some very close buildings were combined unexpectedly (marked as P, Q, and R) apart from the false positive pixels, as shown in Figure 4.12. This can be avoided by analysing the white pixels from the building mask during region growing. Nevertheless, the proposed technique extracted several close buildings with greater accuracy, as shown in Figures 4.12(b) and (h) with Labels (i)–(iii). Moreover, two buildings labelled X and Y, as shown in Figures 4.12(e) and (h), respectively were completely extracted using the proposed cell clustering process, although major parts of the buildings did not have any LiDAR data present in the dataset. The results in Table 4.2 further indicate that the proposed technique is completely free of over- (1:M) and many-to-many (M:N) segmentation errors, although there are some under-segmentation cases.

Table 4.3 shows that the regularisation process has improved the per-area completeness, correctness, and quality. This increase might be substantial if the image lines could come in place of all the boundary segments, which are a more accurate representation of the building outline than the ALS-derived boundary. However, a slight decrease in VH3 correctness after regularisation was observed, because the lines and their intersections (corners) shifted from their actual positions during regularisation. Notably, the regularisation process obtained higher planimetric accuracy, reaching up to two times the horizontal point spacing of the ALS data. Overall, it can be established from the evaluation performance that the proposed delineation together with the regularisation process can not only eliminate vegetation and extract buildings and their parts from the connected trees but also regularise the boundary with 98.57% average objective correctness.

### 4.3.2 Australian Benchmark Results and Discussion

Tables 4.4 and 4.5 present object- and pixel-based evaluations of the detection technique before and after regularisation using the threshold-free evaluation system. Figures 4.13 and 4.14 show the extracted buildings and their corresponding building footprints for the AV2, EL, HT, and HB datasets. The proposed detection extracted 55, 68, 62, and 22 buildings out of 65, 75, 69, and 22 reference buildings in the AV2, EL, HT, and HB datasets, respectively.

Table 4.4 shows that the object-based completeness of HT for all the buildings was comparatively lower than that of AV2 and EL. The reason was the missing buildings (marked red) caused by severe occlusion (Figures 4.13(c) and (f) and 4.14(e)) and transparent roof material (Figure 4.14(g)). The results further show that buildings over 50 $m^2$ were extracted with 100% objective completeness and correctness while nearly equal average completeness ($\simeq$ 95%) was achieved for buildings over 10 $m^2$.

Figures 4.13(c) and (f)–(h) show scenarios where non-occluded building parts were extracted after eliminating the connected vegetation. However, two very close buildings in EL, labelled (i) in Figure 4.13(f), were unexpectedly merged due to a small connecting region. Such complex cases increased the cross-lap rate (under-segmentation) in EL compared with AV2 and HT, as given in Table 4.4. This could be prevented by considering white pixels between the connected regions using the building mask. Since the overlap areas of the resultant and reference polygons before and after regularisation were not greatly changed, the objective evaluation was nearly the same as that shown in Table 4.4.

Table 4.5 shows the pixel-based evaluation (completeness and correctness $\simeq$ 87% and 95%) before regularisation for AV2, HT, and HB. However, there was relatively lower completeness in EL (80.57%) than others caused by missing fewer buildings which were under severe occlusion and evident from the high area omission errors. For buildings larger than 50 $m^2$, average completeness and correctness of about 87% and 95% were achieved, while similar results were obtained for buildings over 10 $m^2$. More than a 2% increase in average completeness was recorded after the regularisation process when all the buildings were considered. However, the planimetric accuracy was compromised, since the regularised lines and their intersections were repositioned to generate a regularised boundary. Nevertheless, Table 4.5 shows a significant increase in average pixel-based completeness from 86.95% to 91.80% in the HB dataset after regularising all the buildings. The reason is that most parts of the delineated boundary were replaced with image lines, which increased the accuracy.

The performance of the proposed technique is demonstrated using several datasets which have transparent to complex structure buildings with varying sizes, flat to hilly

| Proposed Detection | Areas | $C_m$ | $C_r$ | $Q_l$ | $C_{m,50}$ | $C_{r,50}$ | $C_{m,10}$ | $C_{r,10}$ | $C_{rd}$ | $C_{rr}$ |
|---|---|---|---|---|---|---|---|---|---|---|
| Before Regularisation | AV2 | 84.62 | 100 | 84.62 | 100 | 100 | 91.67 | 100 | 1.81 | 1.53 |
| | EL | 84.0 | 100 | 85.0 | 100 | 100 | 94.03 | 100 | 2.9 | 9.3 |
| | HT | 82.61 | 98.39 | 82.43 | 100 | 100 | 95.0 | 100 | 1.6 | 7.2 |
| | HB | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 0 | 13.6 |
| | **Average** | **87.81** | **99.6** | **88.01** | **100** | **100** | **95.18** | **100** | **1.57** | **7.91** |
| After Regularisation | AV2 | 84.62 | 100 | 84.62 | 100 | 100 | 91.67 | 100 | 1.80 | 1.5 |
| | EL | 84.0 | 100 | 85.0 | 100 | 100 | 94.03 | 100 | 2.9 | 9.3 |
| | HT | 82.61 | 98.39 | 82.43 | 100 | 100 | 95.0 | 100 | 1.6 | 7.2 |
| | HB | 100 | 100 | 100 | 100 | 100 | 100 | 100 | 0 | 13.6 |
| | **Average** | **87.81** | **99.6** | **88.01** | **100** | **100** | **95.18** | **100** | **1.57** | **7.91** |

**Table 4.4** Object-based building detection results for Australian datasets before and after regularisation. ($C_m$ = completeness, $C_r$ = correctness, are for all buildings 50 $m^2$ and over 10 $m^2$ in percentage; $C_{rd}$ = detection cross-lap (under-segmentation) and $C_{rr}$ = reference cross-lap (over-segmentation) rates).

| Proposed Detection | Areas | $C_{mp}$ | $C_{rp}$ | $Q_{lp}$ | $A_{oe}$ | $A_{ce}$ | $RMSE$ | $C_{mp,50}$ | $C_{rp,50}$ | $C_{mp,10}$ | $C_{rp,10}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Before Regularisation | AV2 | 88.85 | 96.71 | 86.25 | 11.14 | 3.28 | 0.58 | 91.0 | 96.71 | 89.32 | 96.71 |
| | EL | 80.57 | 95.13 | 77.38 | 19.4 | 4.86 | 2.08 | 81.24 | 95.13 | 80.79 | 95.13 |
| | HT | 87.9 | 92.04 | 81.7 | 12.10 | 7.9 | 1.03 | 89.31 | 92.15 | 88.61 | 92.15 |
| | HB | 86.95 | 96.72 | 84.45 | 13.0 | 3.28 | 1.18 | 86.95 | 96.71 | 86.95 | 96.71 |
| | **Average** | **86.06** | **95.15** | **82.45** | **13.91** | **4.83** | **1.22** | **87.13** | **95.18** | **86.42** | **95.18** |
| After Regularisation | AV2 | 90.11 | 97.50 | 88.08 | 9.88 | 2.49 | 1.30 | 92.43 | 97.50 | 90.71 | 97.50 |
| | EL | 82.25 | 95.43 | 79.13 | 17.75 | 4.56 | 2.08 | 82.93 | 95.43 | 82.47 | 95.43 |
| | HT | 89.4 | 93.67 | 84.28 | 10.62 | 6.33 | 1.42 | 91.0 | 93.74 | 90.11 | 93.74 |
| | HB | 91.80 | 96.08 | 88.49 | 8.19 | 3.91 | 1.78 | 91.80 | 96.08 | 91.80 | 96.08 |
| | **Average** | **88.39** | **95.67** | **84.99** | **11.61** | **4.32** | **1.65** | **89.54** | **95.76** | **91.54** | **95.76** |

**Table 4.5** Pixel-based building detection results for Australian datasets before and after the regularisation stage. ($C_{mp}$ = completeness, $C_{rp}$ = correctness and $Q_{lp}$ = quality are for all buildings 50 $m^2$ and over 10 $m^2$; $A_{oe}$ = area omission error and $A_{ce}$ = area commission error; in percentage, $RMSE$ = planimetric accuracy in metres).


**Figure 4.13** (a),(b) Building detection and regularisation on AV2 dataset; (c) Building regularisation example in AV2; (d),(e) Building detection and regularisation on EL dataset; and (f)–(h) Building regularisation examples in EL. Areas marked in (b) and (e) are magnified in (c) and (f)–(h), respectively.

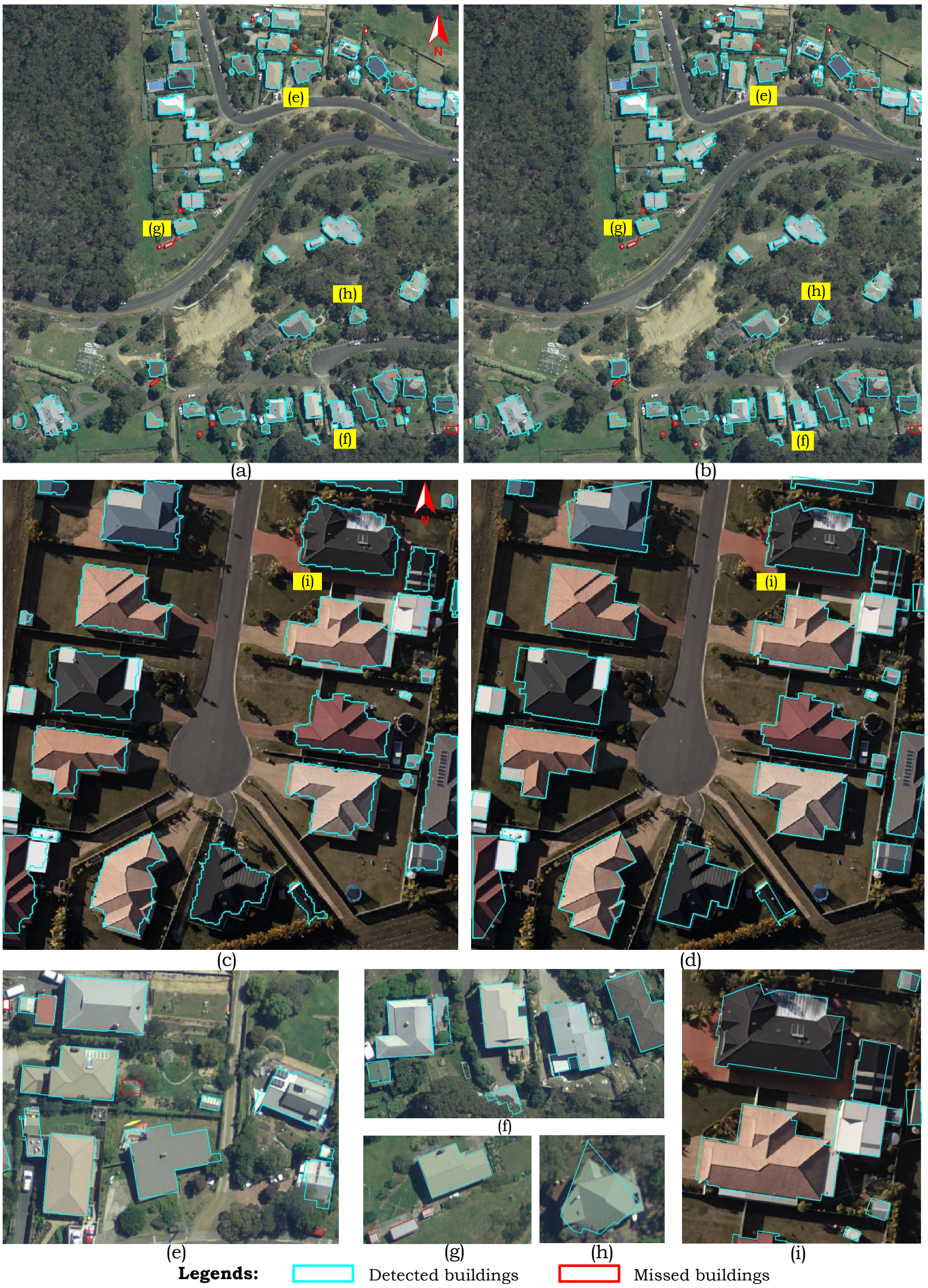



**Figure 4.14** Building detection and regularisation on Hobart (a),(b) and Hervey Bay (c),(d). Building detection examples after regularisation in: (e)–(h) HT dataset (1.6 points/$m^2$) ; and (i) HB dataset (12 points/$m^2$). Areas marked in (b) and (d) are magnified in (e)–(i).

terrains, low to dense vegetation, and different point cloud resolutions. The qualitative (figures) and quantitative (Tables 4.2–4.5) results show that the proposed detection method can eliminate vegetation and extract buildings as well as their non-occluded parts from complex scenes with high object- and pixel-based accuracies. A constantly higher ($>$ 95%) correctness suggests that the proposed detection technique is robust. Moreover, the applicability of the regularisation process after detection is effective in the generation of building footprints.

### 4.3.3 Comparative Analysis

The proposed method is automatic and data-driven, and integrates LiDAR and orthoimagery. It predominantly uses LiDAR to create the building mask and building extraction. From the ISPRS portal and the methods classified in [22], we selected those which (1) use both LiDAR and images; (2) in which pixel- and point-based processing is equally important; (3) are automatic; and (4) are unsupervised and data-driven [22]. Due to a lack of integrated methods meeting the above criteria, two supervised integrated (criteria (1)–(3)) and two LiDAR only (criteria (2)–(4)) methods were also chosen, as shown in Table 4.6. The quantitative evaluation of the German dataset (VH1, VH2, and VH3) for KNTU_mod, Whuz, IIST, Mon2, and the proposed detection and regularisation technique (FED_2) are available on the ISPRS' website. However, Yang's [85] results are taken from the paper. To the best of our knowledge, the only method evaluated using the evaluation system in [72] on the Australian datasets is MA [37]. Therefore, it was chosen for the comparative study in order to conduct a fair evaluation.

| Benchmark dataset | Method's Name | Data Types | Processing Strategy | Reference |
|---|---|---|---|---|
| ISPRS, Germany | KNTU_mod | LiDAR + image | supervised | [140] |
| | Whuz | LiDAR + image | supervised | [141] |
| | IIST | LiDAR + image | Data-driven | [142] |
| | Mon2 | LiDAR | Data-driven | [138] |
| | Yang | LiDAR | Data-driven | [85] |
| Australian | MA | LiDAR | Data-driven | [37] |

**Table 4.6** Existing methods compared with the proposed technique (FED_2).

A comparison of FED_2 and other methods is presented in Tables 4.7 and 4.8, where bold numbers show better or equal performance of our technique. FED_2 performed

| **Methods** | $C_m$ | $C_r$ | $C_{m,50}$ | $C_{r,50}$ | $C_{mp}$ | $C_{rp}$ | *RMSE* | **1:M** | **N:1** | **N:M** |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | VH1 - 3.5 laser points/m$^2$ | | | | | | |
| KNTU_mod | 83.8 | 100.0 | 100.0 | 100.0 | 91.4 | 94.3 | 0.8 | - | - | - |
| Whuz | 78.4 | 43.5 | 89.3 | 96.3 | 84.4 | 83.9 | 1.2 | - | - | - |
| IIST | 78.4 | 83.3 | 82.1 | 95.8 | 75.8 | 74.3 | 1.4 | - | - | - |
| Mon2 | 89.2 | 91.4 | 100.0 | 100.0 | 88.1 | 90.0 | 1.0 | 0 | 6 | 0 |
| Yang | 81.1 | 96.8 | 100.0 | 96.6 | 87.9 | 91.2 | 0.9 | 0 | 3 | 2 |
| **FED_2** | 83.8 | **100.0** | **100.0** | **100.0** | 85.4 | 86.4 | 1.0 | **0** | 6 | **0** |
| | | | | VH2 - 3.9 laser points/m$^2$ | | | | | | |
| KNTU_mod | 83.8 | 100.0 | 100.0 | 100.0 | 86.5 | 93.6 | 0.8 | - | - | - |
| Whuz | 57.1 | 42.3 | 80.0 | 90.9 | 79.6 | 91.9 | 0 | - | - | - |
| IIST | 71.4 | 62.5 | 100.0 | 90.9 | 78.8 | 92.6 | 0.9 | - | - | - |
| Mon2 | 85.7 | 92.3 | 100.0 | 100.0 | 87.1 | 94.0 | 0.8 | 0 | 2 | 0 |
| Yang | 78.6 | 100.0 | 100.0 | 100.0 | 88.8 | 94.0 | 0.8 | 0 | 2 | 0 |
| **FED_2** | **85.7** | **100.0** | **100.0** | **100.0** | **88.8** | 84.5 | 1.2 | **0** | **2** | **0** |
| | | | | VH3 - 3.5 laser points/m$^2$ | | | | | | |
| KNTU_mod | 85.7 | 98.0 | 100.0 | 100.0 | 88.3 | 99.0 | 0.7 | - | - | - |
| Whuz | 64.3 | 79.2 | 81.6 | 100.0 | 76.9 | 92.6 | 1.1 | - | - | - |
| IIST | 67.9 | 58.3 | 86.8 | 94.3 | 86.2 | 78.4 | 1.2 | - | - | - |
| Mon2 | 83.9 | 97.9 | 97.4 | 100.0 | 87.7 | 89.0 | 1.0 | 0 | 8 | 0 |
| Yang | 73.2 | 97.6 | 97.6 | 92.1 | 85.2 | 89.5 | 0.8 | 0 | 6 | 0 |
| **FED_2** | 82.1 | 95.7 | **100.0** | **100.0** | **89.9** | 84.7 | 1.1 | **0** | **5** | **0** |

- means results not available.

**Table 4.7** Comparison of building detection results for Vaihingen (VH) dataset. Object-based $C_m$ = completeness, $C_r$ = correctness ($C_{m,50}$ and $C_{r,50}$ are for buildings over 50 m$^2$) and pixel-based $C_{mp}$ = completeness and $C_{rp}$ = correctness are in percentages. *RMSE* = planimetric accuracy in metres. 1:M = over-segmentation and N:1 = under-segmentation, N:M = both over- and under-segmentation in the number of buildings. **Bold values** show better or equal performance of the proposed technique.

significantly better on the Australian datasets which are more complex than the ISPRS benchmark due to dense vegetation, shadows, and topography. Likewise, FED_2 achieved similar or better object- and area-level completeness and correctness than the counterparts in all the German areas. In VH1, Mon2 offered better object completeness, since it captured the carports at ground height due to more accurate DTM, however, high per-object correctness was obtained by FED_2 (see Table 4.7), which is an indicator of a method's robustness. Moreover, the proposed method had better objective completeness than Yang, which had more per-area completeness on VH1. In addition, FED_2 performed better in VH2 than the other methods, and Mon2 achieved slightly better per-object completeness in VH3. Nevertheless, FED_2 achieved higher per-area and better per-object completeness on the large buildings in VH3.

Table 4.7 further shows that large buildings (50 m$^2$) were extracted with 100% completeness compared with Yang and Mon2, which missed large buildings from VH3 and IIST from VH1 and VH3. IIST also performed poorly on all the areas in the German dataset due to the inappropriate integration of features extracted from LiDAR and images. Notably, FED_2 was mainly free of over- and many-to-many segmentation errors, but had some under-segmentation cases, unlike Yang which suffered from many-to-many and under-segmentation. Compared with supervised methods, which are trained to perform a better classification, the proposed method achieved better per-object completeness and correctness in all the areas except VH3. Likewise, it performed better in per-area level except in VH1 where KNTU_mod had better completeness and correctness. However, Whuz performed consistently poorly on all the areas in the German dataset and even missed buildings larger than 50 m$^2$ in VH1 and VH3. Since the segmentation error results of the IIST, KNTU_mod, and Whuz are not available, a comparative discussion is not provided.

| **Methods** | $C_m$ | $C_r$ | $C_{m,10}$ | $C_{r,10}$ | $C_{mp}$ | $C_{rp}$ | $RMSE$ |
|---|---|---|---|---|---|---|---|
| Aitkenvale (AV2) - 29.3 laser points/m$^2$ | | | | | | | |
| MA | 67.2 | 100 | 81.1 | 100 | 87.2 | 94.9 | 0.66 |
| **FED_2** | **84.62** | **100** | **91.67** | **100** | **90.11** | **97.50** | 1.3 |
| Eltham (EL) - 4.8 laser points/m$^2$ | | | | | | | |
| MA | 77.6 | 88.2 | 77.6 | 88.2 | 85.6 | 90.1 | 1.31 |
| **FED_2** | **84.0** | **100** | **94.03** | **100** | 82.25 | **95.43** | 2.08 |
| Hobart (HT) - 1.6 laser points/m$^2$ | | | | | | | |
| MA | 71.2 | 80.8 | 80.8 | 79.3 | 80 | 80.2 | 1.33 |
| **FED_2** | **82.61** | **98.39** | **95** | **100** | **89.4** | **93.67** | 1.42 |
| Hervey Bay (HB) - 12 laser points/m$^2$ | | | | | | | |
| MA | 73.2 | 97.6 | 97.6 | 92.1 | 80 | 80.2 | 0.68 |
| **FED_2** | **100** | **100** | **100** | **100** | **91.80** | **96.08** | 1.78 |

**Table 4.8** Comparison of building detection results for the Australian datasets. Object-based $C_m$ = completeness, $C_r$ = correctness ($C_{m,10}$ and $C_{r,10}$ are for buildings over 10 m$^2$) and pixel-based $C_{mp}$ = completeness and $C_{rp}$ = correctness are in percentage. $RMSE$ = planimetric accuracy in metres. **Bold values** show better or equal performance of the proposed technique.

FED_2 achieved better object and area level accuracy than MA on the Australian datasets, as shown in Table 4.8. However, in EL, the per-area completeness of MA was marginally better but performed poorly (per-object) in AV2 due to missing small buildings. Nevertheless, FED_2 offered high completeness and correctness on all the datasets but showed lower planimetric accuracy than MA. It is concluded that FED_2 is quite robust for complex scenarios in the Australian datasets and achieves higher completeness and correctness. However, it achieved better or equal performance with the exception of a few cases on the ISPRS benchmark datasets with constantly high correctness.

## 4.4 Summary

A building detection and footprint generation technique is presented in this chapter, which is fully data-driven and automatic. The candidate building regions are first identified using connected component analysis. Next, the buildings are extracted including those partly occluded and shadowed after vegetation removal through the grid index structure and multisource data. Finally, the building footprints are generated using the image lines and the extracted building boundaries.

The performance of the proposed technique was tested on several datasets with different point densities (1 to 29 points/m$^2$), topographies, and vegetation conditions. The results showed that the technique can not only extract small, partially occluded and shadowed buildings, but it can generate footprints irrespective of the surrounding complexity. The proposed method offers high detection rates, even in the presence of moderate registration errors between the ALS data and the orthoimagery. The experimental results further demonstrated that the proposed method is completely free of many-to-many and over-segmentation errors, which is imperative to obtain high objective accuracy. Compared with six existing methods, the proposed technique performs better with correctness of above 95%. In addition, the building outlines produced are regularised, in contrast with the recent methods which generate only ragged boundaries. In the next chapter, we present a roof plane extraction technique, which is a subsequent task in 3-D building modelling.

Chapter 5

# Robust Segmentation and Building Roof Identification

*"Stay positive and happy. Work hard and don't give up hope. Be open to criticism and keep learning. Surround yourself with happy, warm and genuine people."*

Tena Desae

## 5.1 Introduction

The aim of the research in this thesis is the development of a 3-D building framework that includes not only building detection but also extraction of the roof planes. The previous chapter introduced the main aspects of the building recognition task, encompassing building detection, vegetation elimination, and generation of building footprints. Based on examples and performance evaluation using several datasets, it was demonstrated that the proposed integration technique has the ability to eliminate vegetation and extract buildings as well as their non-occluded parts with high object- and pixel-based accuracy. We also observed that image lines were utilised with LiDAR-based building boundaries for the useful generation of building footprints that have high planimetric accuracies. These building footprints can be used subsequently in the task of building roof segmentation for the extraction of roof planes, and this addresses the third research objective **(RO3)**. However, there are some salient reasons for devising a new technique and declining the use of detection results (the output of the previous chapter) for the extraction of building roofs and building regions. The reasons are as follows:

- The detection technique proposed in the previous chapter integrates both LiDAR point cloud and the corresponding aerial imagery, which are not necessarily available for all possible areas. Therefore, there is an urgent need for the development of a technique that can use LiDAR point cloud for the detection of buildings and their constituent roof surfaces.

- A large misalignment between LiDAR and its corresponding imagery generally results in under-detected roof regions, because both LiDAR and image features are generally used together in the extraction process. As a result, the detected building regions result in areas which are compressed compared with their actual areas from images. Consequently, LiDAR point segmentation using these building regions for the extraction of building roofs may worsen the misalignment effect and produce under-detected roof surfaces.

- The building footprints sometimes deviate from their principal direction during the regularisation process and wrongly include nearby non-building regions in their detected areas. Therefore, the use of such building regions for the extraction of roof planes result in anomalous planes which may belong to non-building surfaces.

Evaluation of the existing techniques (in Section 3.4) shows that building roof extraction has been challenging for the research community and is 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 occlusion by neighbouring objects, e.g., vegetation clusters [65]; (4) have no connection information among 3-D laser points; (5) show the presence of noisy laser pulses due to the physical limitations of data acquisition sensors and multiple reflectance (multipath effects); and (6) have no statistical distributional pattern, especially for points around anisotropic surfaces (where multiple surfaces intersect) [143]. The absorption of laser pulses by water and reflection from vents and transparent roof structures are additional issues which make roof plane detection based on LiDAR data alone more challenging.

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 very close to buildings and often occludes parts of roofs. These buildings are generally ignored and therefore removed during the elimination of false objects by some existing methods. To overcome these limitations and address the highlighted challenges, a technique of roof plane extraction and building detection using LiDAR data is proposed here. The main contributions of the chapter are as follows:

- Point cloud density is used at different stages of execution for handling sparsely sampled point sets and making the proposed technique robust for multiple data acquisition sources, e.g., airborne and terrestrial (mobile) platforms.
- A new LiDAR-based boundary-tracing technique is included, which seamlessly extracts the inner and outer boundaries of an object without any limitation.
- Two new algorithms, anisotropic point selection and saliency feature (e.g., surface normal and slope) estimation, are introduced in this research. The first algorithm identifies points on intersecting surfaces using a local rather a global threshold, while the latter estimates saliency features accurately to help in extracting occluded roof planes and is robust to noise.
- A roof plane extraction method and a comprehensive objective assessment using several datasets are included in this study. These datasets differ in scene complexity, topographical conditions, and point density (1.6 to 35 points/m$^2$).

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

The remainder of the chapter is structured as follows: Section 5.2 discusses the methodology for robust roof extraction and building detection, including a detailed description of each stage of the workflow. Section 5.3 provides a comprehensive performance evaluation and comparative analysis of the proposed method using four benchmark datasets. Concluding remarks are provided in Section 5.4.

## 5.2 Methodology

The proposed technique preserves sharp surfaces and combines PCA and the *Low-Rank Subspace Clustering framework with Prior Knowledge* (LRSCPK) technique [144] for the estimation of robust saliency features. The use of these robust features makes the segmentation of LiDAR data more resistive to noise. Figure 5.1 shows the workflow of the proposed roof extraction and building detection technique. The input data consist of raw LiDAR points 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 [133].

The proposed technique comprises three major stages. First, we separate the LiDAR data into 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 using saliency features. Finally, a refinement procedure eliminates the non-building planes and then approximates 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.

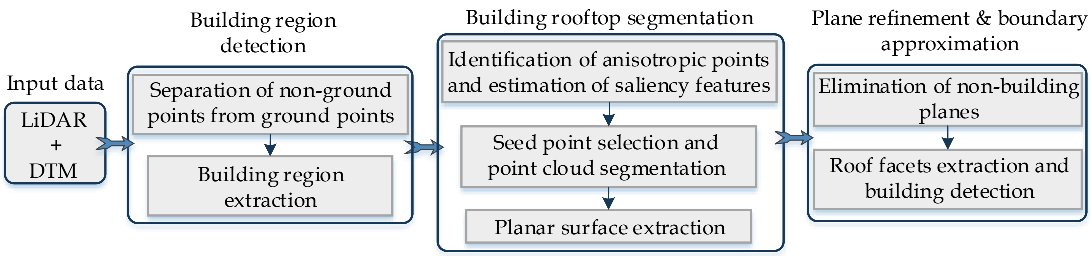



**Figure 5.1** Workflow of the proposed technique.

### 5.2.1 Building Region Detection

Figure 5.2(a) shows the test dataset AV1 introduced in Section 2.6. The aerial image is 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. Although the test dataset is small, it offers the challenges of vegetation and occlusion, as shown in Figure 5.2(a) 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 AV1 comprises the first pulse returns with a point density of 35 points/m$^2$ 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 corresponding DTM as inputs. Generally, airborne LiDAR data contain points returned from different features such as the ground surface, trees, buildings, and other 3-D objects. Therefore, we first separate the area of interest from other ground objects for the detection of building roof planes. In order 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 non-ground points provided they are above the $h_{rf}$. Figure 5.2(b) shows the LiDAR points separated into ground and non-ground points sketched in 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 Delaunay triangulation (see Section 2.3) as shown in Figure 5.2(c). The edges of any triangle having a length $\geq 2d_{max}$ are determined as anomalous connections (see red lines in Figure 5.2(d)), 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 are named herein *building regions*. We used Matlab's `DelaunayTri` function for the construction of the Delaunay triangulation and the `edges` method to identify the unwanted constrained triangles.

The proposed *boundary tracing algorithm* then takes the building regions and approximates their boundaries. It can be observed in Figure 5.2(d) (digital copy) that

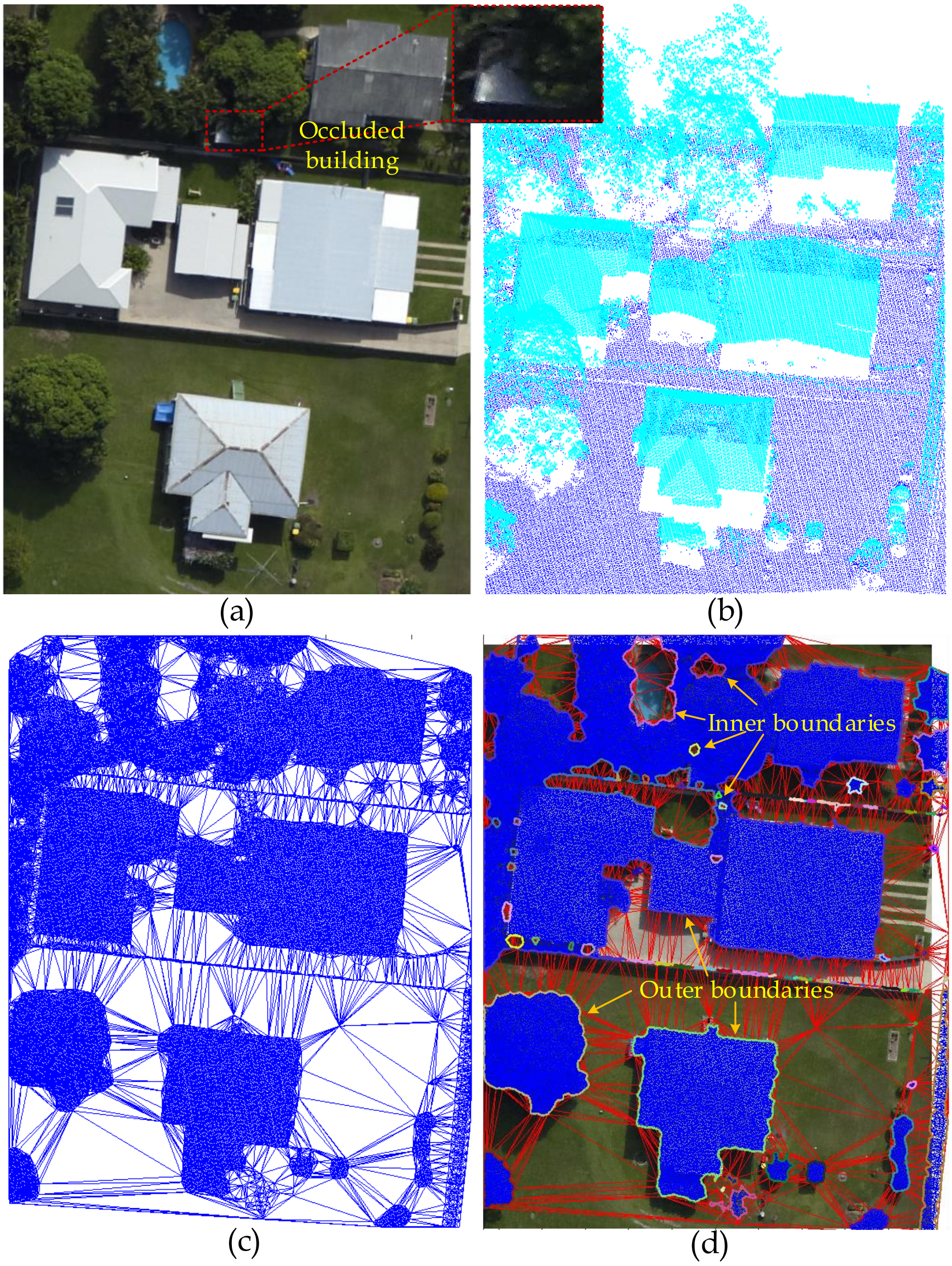



**Figure 5.2** (a) Aerial image of test dataset; (b) LiDAR point set; (c) Delaunay triangularisation of the non-ground points; and (d) Building region identification.

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 obtain the edges of triangles along a region's periphery and inside 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 the start vertex of which is the end vertex of the previous edge and adds it in a boundary segment. The boundary approximation stops if the process meets an edge the end vertex of which 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 [145] and [146], 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 those in [145] and [147], where the value of $\alpha$ is carefully chosen in order to avoid producing exclusively convex hulls. The boundary tracing technique, on the other hand, relies only on 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 5.2(d) presents a snapshot of the proposed boundary tracing technique and also shows the extracted building regions and their boundaries.

### 5.2.2 Building Rooftop Segmentation

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. Therefore, LiDAR points on these features describe fundamental characteristics of the underlying geometry and considering them in advance improves the performance of a segmentation process [148] and therefore an integral component of our segmentation method.

Similarly, estimation of *saliency features* such as normal and curvature, is an essential step in the surface approximation process, because the quality of the extracted surfaces depends heavily on the quality of the estimated point normals [144, 149]. Often, Principal Component Analysis (described in Section 2.3.2) is used but, it overly smooths the normals at the points near/on the sharp features due to non-robust location and scatter analysis [69, 149]. Therefore, any segmentation based on these erroneous descriptors results in unreliable and inaccurate surfaces [150].

Further to the smoothness issue, PCA is also inefficient in estimating normals in the presence of noise and a direct and/or indirect reason for the failure of region-growing processes [69]. To overcome these limitations, several methods [69, 99, 144, 149–152] have been proposed which can handle data inconsistencies to a certain extent. To obtain statistically robust normals along sharp features, we propose a *PCA mollification* method that uses pre-computed normals to generate consistent point normals. PCA mollification is based on the LRSCPK technique [144], which uses the PCA's pre-computed point normals as prior knowledge and employs an unsupervised learning process to 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, making it robust to noise [153] and enables the handling of corrupt data [144]. In the research literature, mollification has been used in computer graphics for the representation of geometric models [151]. However, to the best of our knowledge, this is the first study of the detection of roof planes and buildings from LiDAR data using PCA mollification.

In addition, the estimation of a point normal largely depends on the appropriate selection of a local neighbourhood size and the method to search a point's neighbours. Commonly, two neighbourhood selection methods are widely practised (as discussed previously in Section 2.3): Knn and fixed-distance neighbours. However, Knn has the advantage over others due to its adaptiveness towards the sparsity of an unstructured point cloud, which makes it suitable for airborne LiDAR processing. The present research, therefore, adopted the Knn method to determine local neighbourhoods 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.

#### 5.2.2.1 PCA Mollification

PCA internally reduces the dimensions of the data and finds a matrix $V$ (representing eigenvectors $\vec{v_1}, \vec{v_2}, \vec{v_3}$) and scalar $\lambda$ (representing eigenvalues $\lambda_1 \leq \lambda_2 \leq \lambda_3$), where $\lambda_i$ describes a spatial variation along the corresponding $\vec{v_i}$. In the case of 3-D point cloud data, the first two eigenvectors can describe a planar surface where the smallest eigenvalue $\lambda_1$ corresponding to $\vec{v_1}$ defines a point normal $\hat{n}_i$ [69]. However, PCA approximates inaccurate normals at the feature points, those not lying on planar surfaces, such as edges. Therefore, we combine LRSCPK at this stage, which segments the neighbours of each feature point into several isotropic subspaces and re-estimates 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 fitted 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 geometries 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 5.3 shows that the point normals approximated using PCA mollification can accurately approximate the underlying surfaces as compared to the PCA.

#### 5.2.2.2 Feature Point Selection

PCA performs *Eigen value analysis* to estimate the normal for each LiDAR point. These normals are then used for the identification of sharp features so that points on/around these intersecting surfaces can be preserved to avoid degenerating the segmentation process. It is achieved by computing the weight (curvature) $w_i$ of each point $p_i \in P$ that measures the likelihood of $p_i$ belonging to a sharp feature. $w_i$ is computed using Equation 5.1 as defined by Pauly et al. [154]:

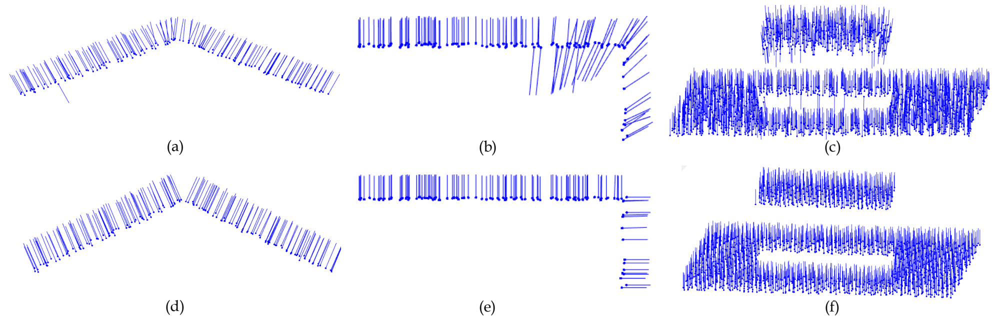


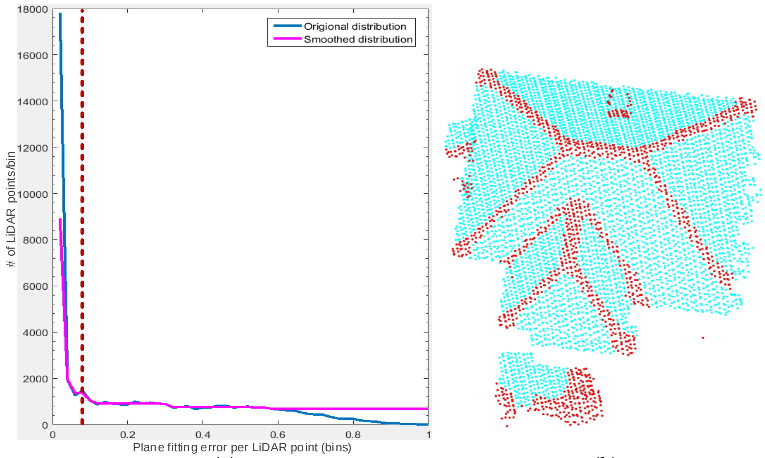

**Figure 5.3** Point normals for three real-world roof plane samples using PCA (a)–(c) and (d)–(f) PCA mollification using the LRSCPK. Planes forming an acute/obtuse angle (left), right angle (middle), and jump edges (right).

$$w_i = \frac{\lambda_1}{\lambda_1 + \lambda_2 + \lambda_3} \tag{5.1}$$

The use of a global weight threshold to determine feature points will be irrational since the underlying point set of each building region can have an entirely different geometry. Therefore, we extended the original principle proposed by [144] that automatically calculates a single threshold for the entire point set. We incorporated an adaptive threshold estimation mechanism that first computes a histogram capturing the distribution of the weights for each building region separately. A threshold $w_t$ is then defined as the horizontal ordinate where the plane fitting residual begins to show a slow decrease. We used points within $2d_{max}$ of each LiDAR point as its $k$ local neighbours for plane fitting. Zhang et al. [144] define the distribution of $\{w_i\}_{i=1}^{N}$ as $f_w$ and smooth it by the following function:

$$min_{\hat{f}_w} \quad \|\hat{f}_w - f_w\|_F + \|Df_w\|_1 \tag{5.2}$$

where, $D$ is the second difference matrix, and $\|\cdot\|_F$ and $\|\cdot\|_1$ represent $l_2$ norm and $l_1$ norm, respectively. A threshold $w_t$ value for a building region is chosen after the first peak of the smoothed distribution, as indicated by the red dotted line in Figure 5.4(a). So, the LiDAR points of a building region having weights below $w_t$ are classified as feature points (red), as shown in Figure 5.4(b). Similarly, all the building regions are processed concurrently to determine a local region-specific $w_t$ for the selection of feature points. The results are shown in Figure 5.5(a). The feature points of each building region are further used to estimate robust normals using the proposed PCA mollification method (see Section 5.2.2.1). The estimated normals in Figure 5.5(b) show that the point normals near connected vegetation and across sharp edges are quite accurate and robust.




(a) (b)

**Figure 5.4** (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 $w_t$.

#### 5.2.2.3 Segmentation

A region-growing technique commonly uses several parameters to determine the coplanarity of the LiDAR points. However, the proposed segmentation method utilises only the critical parameters to achieve robustness and efficiency. Therefore, the surface curvature and normal orientation [69, 145] of points are chosen as proximity criteria, whereas, to distinguish superstructures, point-to-plane tolerance and plane fitting error thresholds [37] are used as coherence criteria. Curvature $\sigma$, which measures the rate of change of surface normal, is estimated using Equation 5.1 earlier referred to as weight. Generally, real-world datasets have inherent noise and even the points reflected from a smooth surface in a local vicinity have some height variations. Therefore, the coherence criterion is inevitable for better convergence of the plane surfaces.

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

$$\theta_{i,j} = cos^{-1}|\hat{n}_i \cdot \hat{n}_j| \tag{5.3}$$

Following the principle that a smooth (non-feature) point in a sufficiently small local neighbourhood always lies on a planar surface, we chose a relatively small neighbourhood $k$ i.e., LiDAR points within $2d_{max}$ for coplanarity check and the extraction of planar surfaces. Trees, in point cloud data, are assumed to be composed of several randomly oriented surfaces which are intersecting arbitrarily, would have a high concentration of feature points compared to building roofs. This cue is exploited in two meaningful ways: to reduce the amount of data and to eliminate the vegetation. Therefore, for point cloud segmentation, those building regions, where concentration of the feature points is less than 95%, are processed to extract planar surfaces. This condition also helps with removal of several small bushes and trees at the beginning of the segmentation process. Figure 5.5(a) shows tree regions, which can be seen in the image in Figure 5.2(a), have a high concentration of feature points (red) in contrast with buildings.

The segmentation process begins with the selection of a seed point from the set of non-feature points that has the least curvature value. Next, we take $k$ local neighbourhood points $N_p$ of the seed. Then, the points with angular differences between the normal of the seed and $N_p$ within a predefined threshold $\theta_t$ are used for the plane fitting using PCA. If plane fitting error and difference between a points' height and fitted plane's height are smaller than the corresponding *fitting error* $\varepsilon_t$ and *point-to-plane tolerance* $\xi_t$ thresholds, these LiDAR points are added in a planar region. This region continues growing as long as new neighbouring points meet the criteria. Otherwise, a new region is instantiated. The proposed segmentation process considers all the unsegmented LiDAR points, including feature and non-feature points for determining a local

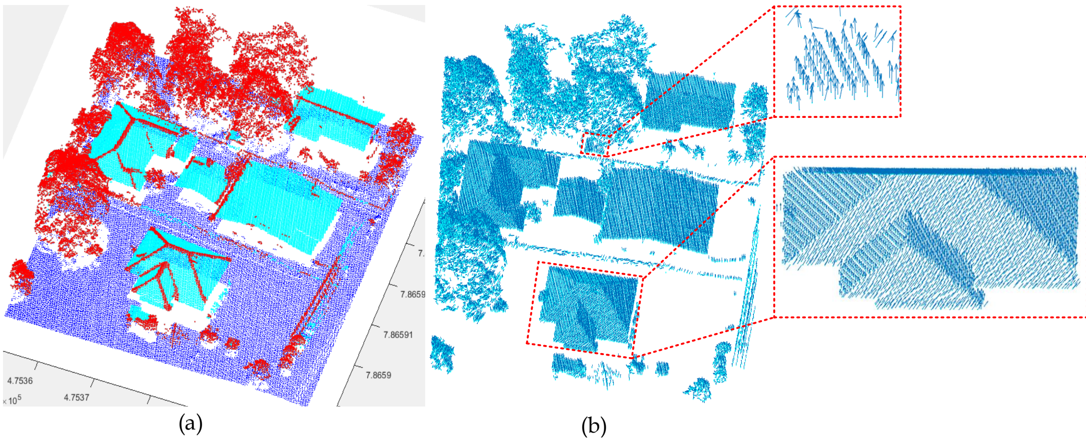



**Figure 5.5** Feature point estimation: (a) Ground points (blue), non-ground points (cyan), feature points (red) and (b) Estimated point normals using proposed PCA mollification. Insets show magnified snapshots of two buildings and their estimated point normals.

neighbourhood, while expanding any region. This procedure of region growing adds most of the anisotropic points to some planar surfaces. The extracted planar primitives at this stage are shown in Figure 5.6(a).

Often, building roofs have small facets such as dormers, chimneys, and vents, and the LiDAR returns of them may be classified as feature points. Therefore, 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, points reflected from such regions generate a significant number of small planar surfaces. For instance, Figure 5.6(b) shows vegetation that breaks up into many smaller planar surfaces. These planar surfaces do not render any parts of trees rather sets of points reflected of tree-branches that are nearly coplanar. These planes are shown in Figure 5.6(c) using the aerial image, where red-dotted ovals indicate trees which were eliminated due to high concentration of feature points ($> 95\%$) at the beginning of segmentation process. Ideally, any two points on a truly flat plane have similar heights and normals. However, due to noise and surface roughness, there are some random errors in the 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 $\theta_t = 10^o$ from [145] and $\varepsilon_t = 0.1$ m and $\xi_t = 0.15$ m given in [37].

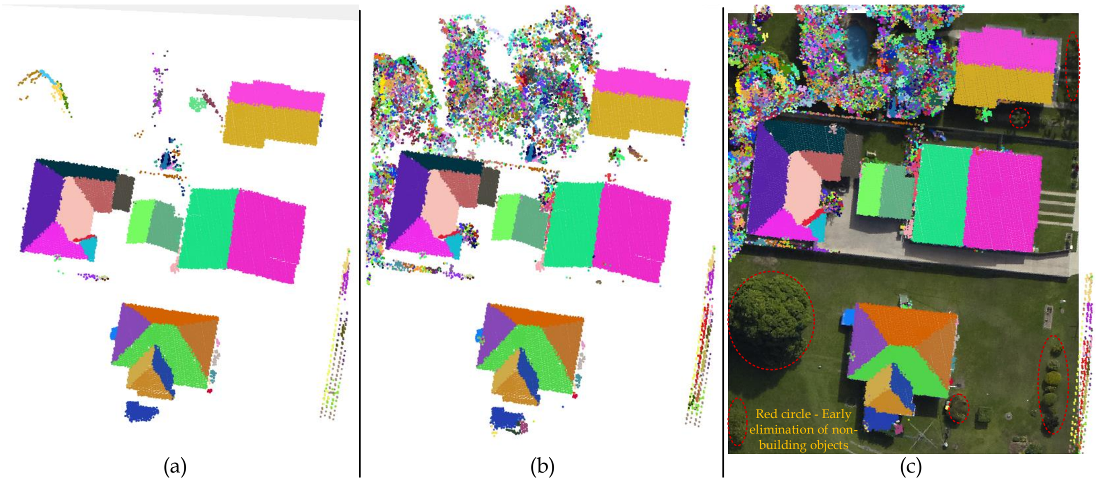



**Figure 5.6** Plane extraction from 3-D point cloud using: (a) non-anisotropic points; (b) anisotropic points; and (c) overlaying the extracted planes on the test image for the demonstration.

### 5.2.3 Removing non-building Planes and Boundary Approximation

We formulated the problem of removing non-building planes in unsupervised non-parametric fashion by identifying the underlying patterns of LiDAR data points. The *concentration of feature points* $r_c$ is observed a key indicator to identify non-planar regions and is proposed as a useful cue to distinguish vegetation compared to roof planes. The sizes of tree segments are typically minuscule, which is a useful feature to differentiate between vegetation and roofs [155], and is therefore, used here. In addition, low number of the *segmented to unsegmented points ratio* $r_u$ in plane segments, which typically is high for vegetation but low for roof surfaces, is employed. This attribute serves a reliable cue for the classification of vegetation segments. To avoid the removal of small roof planes, like chimneys and vents, the plane refinement procedure adopts the following criterion from Awrangjeb et al. [156]. This test marks any misclassified plane as a roof plane if it resides within the boundary of an accepted plane.

Typically, urban buildings have complex arrangement of dormers, which do not extend their footprints but rather provide architectural detail. These dormers, as shown in Figure 5.7, are generally constructed using several intersecting surfaces, which are small in size and in proximity e.g., hexagonal gazebo dormers. As a result, LiDAR points reflecting from these planar surfaces are often classified as feature points, because of insufficient cues to differentiate them from tree planes. Therefore, another test is performed to detect valid roof planes if any exists in anisotropic points. This test ensures a plane, that has not yet been classified, exists in local neighbourhood of a valid roof plane and has a straight line at least 2 m along its boundary. All the user-defined parameters used in the refinement process are listed in Table 5.1.


**Figure 5.7** Various types of dormers.

| Parameters | Values | Sources |
|---|---|---|
| Feature points Concentration $r_c$ | 95 (%) | This chapter |
| Straight line length | 2 m | This chapter |
| Minimum plane area | 1 $m^2$ | [37] |
| Used point ratio $r_u$ | 60 (%) | [37] |

**Table 5.1** Parameters for plane refinement.

The proposed boundary tracing technique is capable of deriving boundary of a segment in $\mathbb{R}^3$, because each transformed point in $\mathbb{R}^2$ has the same reference index as the input LiDAR data. Therefore, the 3-D boundaries of the segmented point cloud are determined and shown in Figure 5.8(a). In this procedure, a plane segmented from tree or any non-building region is usually small in size and has both high concentration of feature points ($r_c$) and low point usage ($r_u$) compared to a roof plane. Therefore, such planes, known as false planes, are removed regardless of other parameters. Further, planes which exist in local neighbourhood of false planes are considered false alarms, and are removed. However, two tests: presence of long straight lines along the boundary of a plane and its occurrence near an accepted plane, are performed to detect roof planes connected to vegetation and identify small roof facets like dormers and sheds. These line segments are extracted using the canny edge detector from the plane's mask.

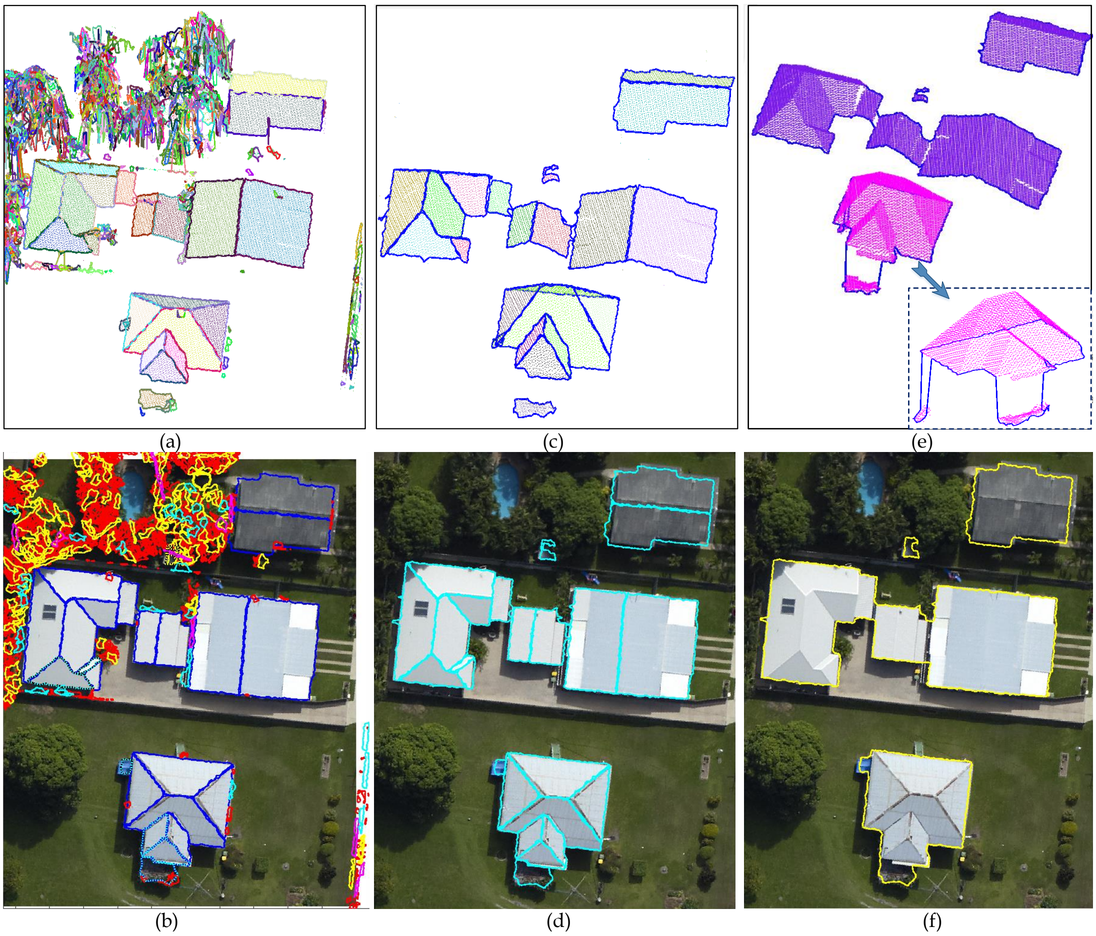



**Figure 5.8** Demonstration of removing non-building planes and building detection using LiDAR and imagery respectively: (a)–(b) All the extracted plane segments and their boundaries; (c)–(d) The detected building roof planes; and (e)–(f) The detected buildings.

Figure 5.8(b) shows a snapshot of false plane removal, where roof planes (plotted in blue and black-dotted on cyan colour) and non-building segments (represented in red, yellow, cyan and magenta colours) are sketched. The planar surfaces plotted in black dots on cyan boundaries correspond to roof planes which are identified using the neighbourhood criterion and the line presence along plane's boundary test. Figures 5.8(c) and 5.8(d) show all the roof planes and their boundaries after the elimination of non-building structures in $\mathbb{R}^3$ and $\mathbb{R}^2$, respectively.

### 5.2.4 Building Outline Generation

The buildings present in test area are now extracted using all the detected roof planes. We process the building regions sequentially and collect the LiDAR points of all the detected roof planes. Next, the proposed boundary tracing technique is employed to determine connected region(s)/building(s) and approximate their outlines using the procedure in Section 5.2.1. Figures 5.8(e) and 5.8(f) show the extracted buildings and their corresponding boundaries in $\mathbb{R}^3$ and $\mathbb{R}^2$, respectively.

## 5.3 Performance Evaluation

To validate the performance of proposed technique, we provide comprehensive evaluation using four benchmark datasets (introduced in Section 2.6), which have different LiDAR resolutions, topographies, and surrounding conditions. The ISPRS dataset has three test areas and other three datasets have one area each captured from different geographic locations in Australia.

### 5.3.1 Building Detection

Tables 5.2 and 5.3 show per-object and per-area level quantified evaluation for the ISPRS and the Australian datasets. Figures 5.9 and 5.10 show the respective detection results and some detection examples from the benchmark test areas. For the ISPRS dataset, detailed quantitative and qualitative measures of building detection can be found on the ISPRS portal [139] under the acronym Mon5[1].

First, we proceed with qualitative evaluation. The buildings in Figures 5.9(d)–(e) and 5.10(g) and (i) show classic examples of small huts, which were accurately segmented and successfully detected (yellow polygons). Figures 5.9(f) and 5.10(d)–(f) and (i) show few complex scenarios, where partially occluded buildings were accurately separated from nearby vegetation, demonstrating the robustness of the proposed technique to noise and non-homogeneous surface points. The buildings in Figure 5.9(g) have a combination of several flat and hipped rooftops, which collectively form

[1] http://www2.isprs.org/commissions/comm3/wg4/results/a1_recon.html

| Test-case | Per-object | | | Segmentation | | | Per-object $\geq$ 50 m$^2$ | | | Per-area | | | *RMSE* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $C_m$ | $C_r$ | $Q_l$ | $M$ | $N$ | $B$ | $C_{m,50}$ | $C_{r,50}$ | $Q_{l,50}$ | $C_m$ | $C_r$ | $Q_l$ | |
| VH1 | 91.9 | 97.1 | 89.5 | 0 | 8 | 0 | 100.0 | 100.0 | 100.0 | 89.0 | 90.2 | 81.1 | 0.93 |
| VH2 | 92.9 | 100.0 | 92.9 | 0 | 2 | 0 | 100.0 | 100.0 | 100.0 | 91.1 | 92.8 | 85.1 | 0.82 |
| VH3 | 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 5.2** Building detection evaluation results using the 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 $\geq$ 50 and 10 m$^2$ | | | | Per-area | | | *RMSE* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $C_m$ | $C_r$ | $Q_l$ | $C_{rd}$ | $C_{rr}$ | $C_{m,50}$ | $C_{r,50}$ | $C_{m,10}$ | $C_{r,10}$ | $C_m$ | $C_r$ | $Q_l$ | |
| AV1 | 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 5.3** Building detection evaluation results using threshold-free reference classification of the 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).

inner boundaries. The proposed technique traces both the inner and outer building boundaries although their boundary points have no topological relationship. The additional benefits of the robust normal estimator are twofold; it extracts planar surfaces and eliminates curved surfaces. For instance, see trees and domes in Figure 5.9(h), where a roof-mounted umbrella was accurately separated from neighbouring planar surface.

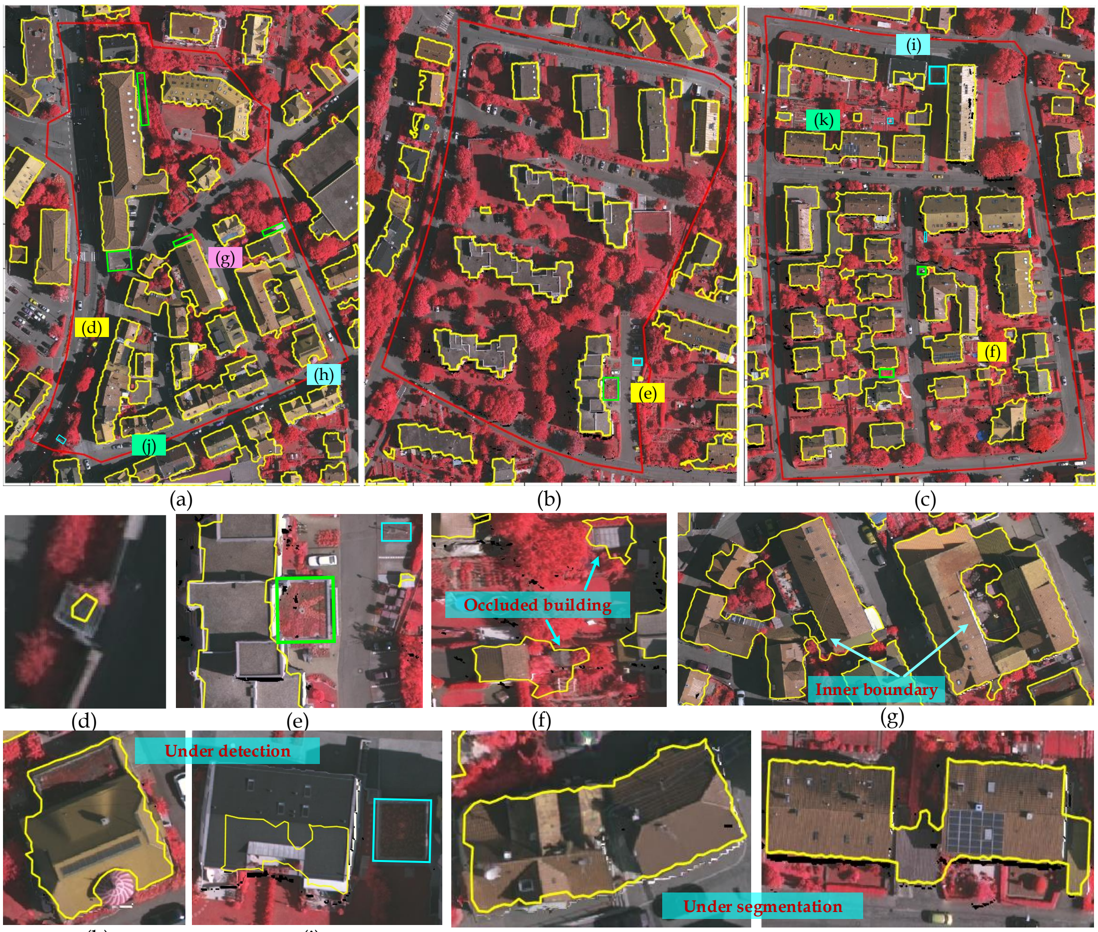



**Figure 5.9** Building detection on the VH dataset: (a) VH1; (b) VH2; and (c) VH3. Snapshots (d)–(h) show examples of small, occluded, and under-detected cases.

Visual inspection indicates that proposed method obtains good results on all the datasets. However, there are some segmentation errors. The proposed algorithm has missed some building attachments (green rectangles) and small buildings (cyan rectangles), as shown in Figure 5.9. This is because VH has slopping terrain and building regions on stilts have LiDAR points below 1 m ($h_{rf}$), which were removed during sepa-

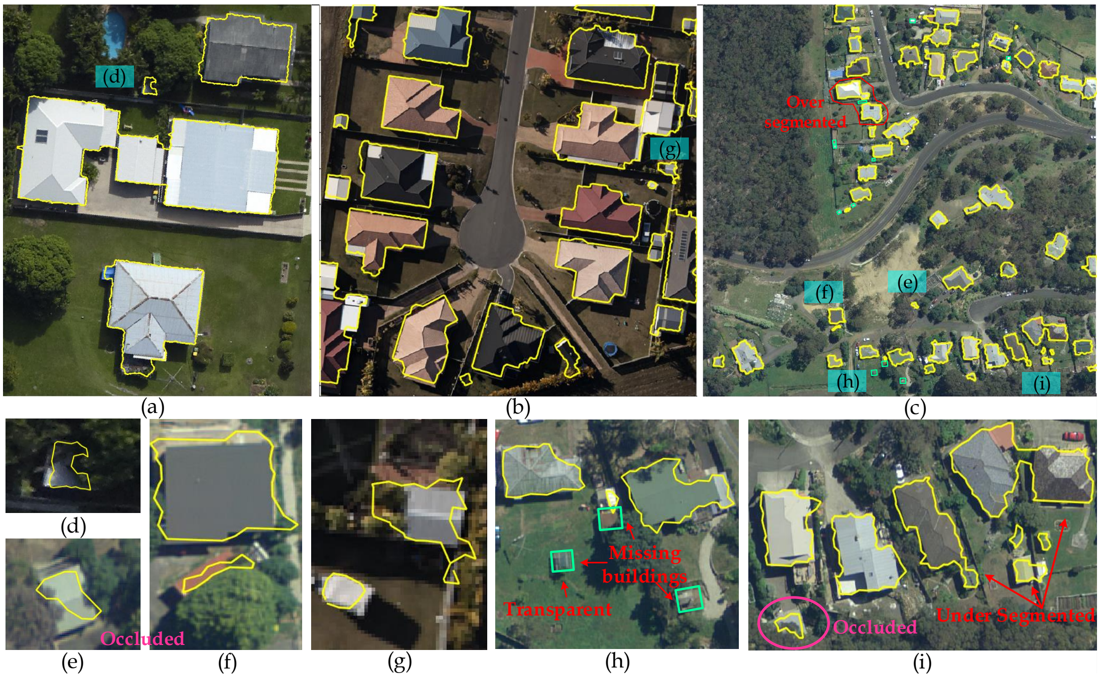



**Figure 5.10** Building detection on Australian datasets: (a) AV1; (b) HB; and (c) HT. Snapshots (d)–(i) show small, occluded, and under-detected cases.

ration of ground and non-ground points. Some small buildings were missed in the HT dataset, as shown in Figure 5.10(c) and magnified Figures 5.10(h)–(i). This was due to low point density ($\approx$ 1 point/m$^2$) and severe occlusion of neighbouring vegetation.

To quantitatively evaluate the detection results, these detected buildings are further analysed. Considering all the buildings, Table 5.2 shows that overall object-level completeness for VH1 and VH2 are 91.9% and 92.9% with corresponding correctness of 97.1% and 100.0%. For buildings larger than 50 m$^2$, the proposed technique achieved 100% object-based completeness, correctness, and quality in VH1 and VH2. However, VH3 has lower object-level accuracy because of missing point cloud data of partially detected large building, as shown in Figure 5.9(i). Some under-segmentation cases occurred when nearby buildings were close to one another. As shown in Figure 5.9(j), a carport between two buildings and two carports in Figure 5.9(k) were merged with their neighbouring buildings. This unexpected merging was due to low density of the input LiDAR data, which can be avoided by analysing height spikes among the neighbouring planes at the time of delineating the building peripheries. As far as per-area accuracy is concerned, the statistics in Table 5.2 show both average completeness and correctness were around 90%, indicating accurate detection of correct pixel points. The results in Table 5.2 also indicate that the proposed technique is entirely free of over- and many-to-many segmentation errors.

A similar detection trend was observed in the AV1 and HB datasets, as presented in Table 5.3. The statistics represent high detection rate, and in particular, the overall accuracy, quantified in terms of completeness, correctness, and quality, was 100.0%. In addition, there were no under- and over-segmentation cases because buildings were well separated and both AV1 and HB datasets had high point densities in contrast with the VH dataset. The results in Table 5.3 show that the proposed method has extracted small buildings (see Figures 5.10(d) and (g)), which were as smaller as 10 m$^2$, including buildings larger than 50 m$^2$. However, considering all the buildings in the HT dataset, object-based completeness was comparatively lower than two other datasets. The reason was some missing buildings (marked in cyan) caused by severe occlusion and transparent roof material (Figures 5.10(c) 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 buildings over 50 m$^2$ were extracted at 100% object-based completeness and correctness, while nearly equal completeness (96.2%) and correctness of 100.0% were achieved for buildings over 10 m$^2$. However, three close buildings in HT, as shown in Figure 5.10(i), were merged unexpectedly since they were situated in close proximity a distance of less than $2d_{max}$. Such complex cases have increased the detection cross-lap rate (under-segmentation) in HT. Moreover, missing transparent buildings results in reference cross-lap (over-segmentation), as shown in Figure 5.10(c). In terms of per-pixel accuracy, performance of HT was lower (85.6%) than that of AV1 (96.2%) and HB (95.2%) but it has a similar correctness.

The statistics in Tables 5.2 and 5.3 indicate that the planimetric accuracies achieved were close to one to two times of horizontal resolution of the LiDAR data. Overall, these experiments suggest that the proposed method obtains high detection performance and extracts buildings of variable sizes and partially occluded from flat to hilly terrains under different surrounding complexities. Moreover, statistics of both the tables show that per-object and per-pixel accuracies are promoted with a proportional increase in point cloud densities. A constantly higher ($> 97\%$) correctness further indicates that our technique is robust to scene complexity. In fact, both qualitative and quantitative results show that the proposed detection method can eliminate vegetation and extract buildings as well as their non-occluded parts from complex scenes with high object- and pixel-based accuracies.

### 5.3.2 Building Roof Detection

Tables 5.4 and 5.5 show object- and pixel-based evaluation results of roof plane extraction for the ISPRS and the Australian datasets, respectively, and Figures 5.11 and 5.12 show the roof plane extraction results for the ISPRS and the Australian datasets, respectively. These figures also present some samples of plane extraction results from the corresponding datasets.

In the VH dataset, the proposed roof extraction algorithm performed better on VH3, which is purely residential and has detached houses. Table 5.4 shows that planes larger than 10 $m^2$ were detected with per-object completeness and correctness of 90.2% and 99.7%, respectively. Some examples are shown in Figures 5.11(d)–(f) (yellow ovals). However, there were many under-segmentation cases, where small roof planes were not extracted separately, and they were merged in neighbouring large planes, as shown in Figure 5.11(g) (purple oval). In addition, there were some over-segmentation cases, when roof planes were detected in two or more splits, as shown in Figures 5.11(e),(h), and (i) (aqua ovals). Some roof structures were also missed, because they were either situated below the height threshold or smaller than 1 $m^2$, as shown in Figures 5.11(g)–(i) (red ovals). Consequently, per-object completeness was rather low for all the areas. However, the proposed technique achieved per-area completeness and correctness of 82.0% and 98.6%, respectively.

With increasing point densities in the Australian datasets, improved roof detection results were obtained in AV1 and HB, as the statistics show in Table 5.5. The proposed technique extracted planes larger than 10 $m^2$ with 100% per-object completeness and correctness in the AV1 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 5.12(d)–(e) and (g)). Figures 5.12(g)–(j) show that many small planes

| Test-case | Per-object | | | Segmentation | | | Per-object $\geq 10$ m$^2$ | | | Per-area | | | *RMSE* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $C_m$ | $C_r$ | $Q_l$ | $M$ | $N$ | $B$ | $C_{m,10}$ | $C_{r,10}$ | $Q_{l,10}$ | $C_m$ | $C_r$ | $Q_l$ | |
| VH1 | 74.3 | 98.7 | 73.6 | 15 | 32 | 10 | 89.8 | 100.0 | 89.8 | 80.7 | 98.7 | 79.8 | 0.76 |
| VH2 | 73.9 | 94.8 | 71.0 | 27 | 1 | 2 | 89.6 | 100.0 | 89.6 | 81.4 | 99.5 | 81.1 | 1.06 |
| VH3 | 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 5.4** Roof plane evaluation results using the 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).

| Test-case | Per-object | | | Segmentation | | Per-object $\geq 10$ m$^2$ | | | Per-area | | | *RMSE* |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | $C_m$ | $C_r$ | $Q_l$ | $C_{rd}$ | $C_{rr}$ | $C_{m,10}$ | $C_{r,10}$ | $Q_{l,10}$ | $C_m$ | $C_r$ | $Q_l$ | |
| AV1 | 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 5.5** Roof plane evaluation results using threshold-free reference classification of the 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).

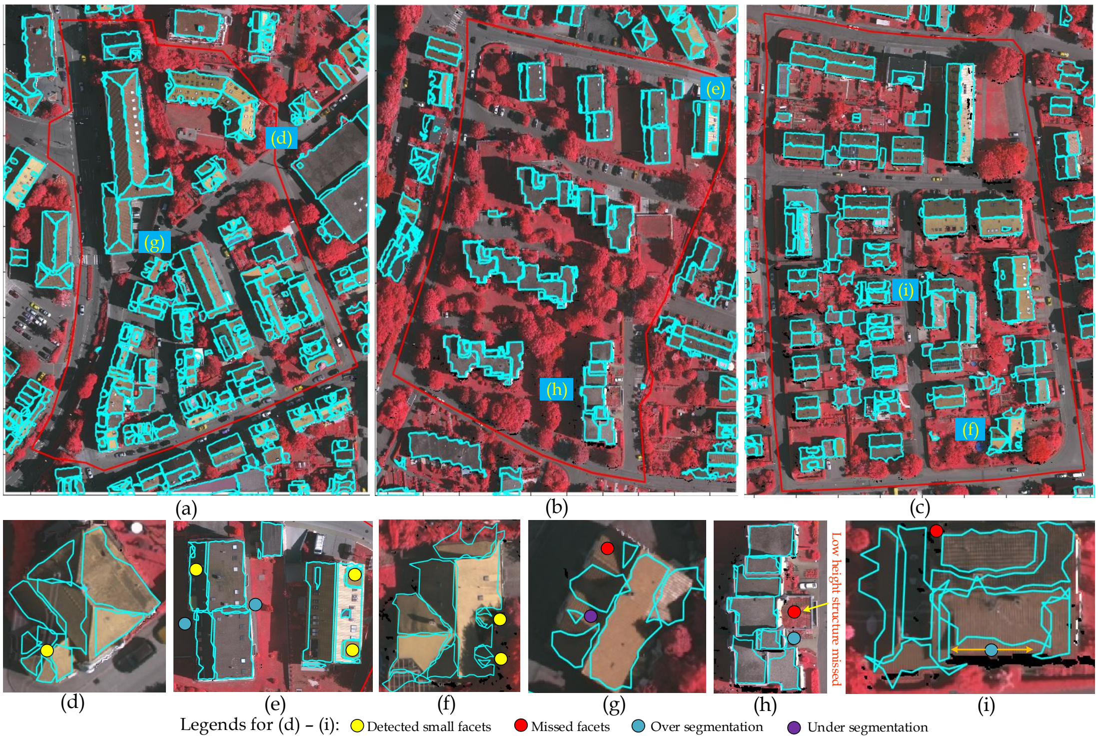



**Figure 5.11** Roof plane extraction on the ISPRS dataset: (a) VH1; (b) VH2; and (c) VH3. Areas marked by letters in (a), (b), and (c) are magnified in (d)–(i).

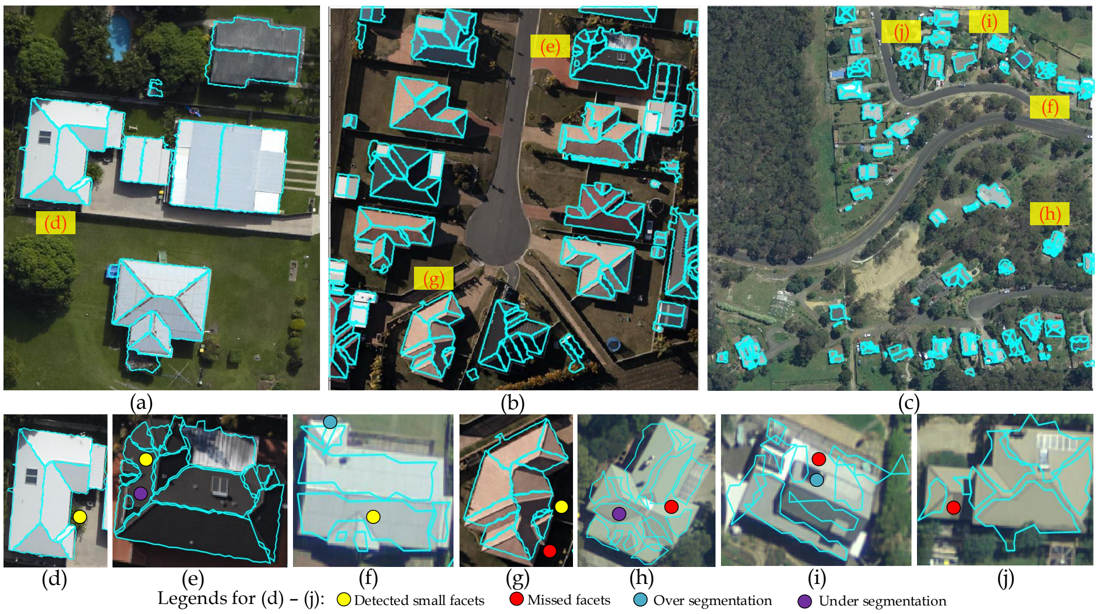



**Figure 5.12** Roof plane extraction on the Australian datasets: (a) AV1; (b) HB; and (c) HT. Areas marked by letters in (a), (b), and (c) are magnified in (d)–(j).

($< 10$ m$^2$) were missed (red ovals) and merged into neighbouring planes, and this increased the under-segmentation and over-segmentation rates. These segmentation errors were more numerous in HT dataset, since it had low point density ($\approx 1$ point/m$^2$) and severe occlusion. The statistics further show that the planimetric accuracy of roof plane extraction was within one to two times of horizontal point spacing of the input LiDAR points. Despite the registration error between detected roofs (from LiDAR data) and reference roofs (from orthoimages), the proposed method achieved nearly 90% per-area completeness for the AV1 and HB datasets, and around 72% completeness for the HT dataset. A constantly higher correctness of above 90% demonstrates robustness of the method in roof plane extraction for different complex conditions. It is observed that detection performance of the proposed method degrades gracefully with the decrease in point density, and does not severely impact planimetric accuracy of the extracted polygons.

### 5.3.3 Comparative Analysis

The proposed automatic technique is data-driven and only use airborne LiDAR data. Therefore, for comparative purposes, the methods which are (1) automatic and data-driven; (2) use only LiDAR data; and (3) unsupervised, were chosen from the ISPRS portal [139] and the methods classified in [22]. The evaluation results of the HKP, VSK, and TUD on the VH dataset are available at the ISPRS portal. However, the Yang [85] and MA [37] results are taken from their papers. Tables 5.6 and 5.7 present performance evaluation of building detection and roof plane extraction results.

For all VH areas, Table 5.6 shows that the proposed building detection technique offers significantly better per-object level completeness and similar correctness compared with the alternative methods. For buildings larger than 50 m$^2$, our technique achieved 100% accuracy in terms of completeness and correctness in contrast to VSK, which was unsuccessful in detecting large buildings in VH1, as indicated by the low $C_{m,50}$ in Table 5.6. In terms of per-area accuracy, HKP and MA obtained slightly more per-area completeness in VH1 and VH3, since our method missed some carports below the height threshold. In terms of the planimetric accuracy of extracted polygons, the proposed

| **Methods** | $C_m$ | $C_r$ | $C_{m,50}$ | $C_{r,50}$ | $C_{mp}$ | $C_{rp}$ | $RMSE$ |
|---|---|---|---|---|---|---|---|
| VH1: 3.5 laser points/m$^2$ | | | | | | | |
| HKP [157] | 83.8 | 100.0 | 100.0 | 100.0 | 92.0 | 97.4 | 0.9 |
| Yang [85] | 81.1 | 96.8 | 100.0 | 96.6 | 87.9 | 91.2 | 0.9 |
| MA [37] | 83.8 | 96.9 | 100.0 | 100.0 | 92.7 | 88.7 | 1.11 |
| VSK [158] | 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 |
| VH2: 3.9 laser points/m$^2$ | | | | | | | |
| HKP [157] | 78.6 | 91.7 | 100.0 | 100.0 | 93.0 | 98.4 | 0.6 |
| Yang [85] | 78.6 | 100.0 | 100.0 | 100.0 | 88.8 | 94.0 | 0.8 |
| MA [37] | 85.7 | 84.6 | 100.0 | 100.0 | 91.5 | 91 | 0.83 |
| VSK [158] | 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 |
| VH3: 3.5 laser points/m$^2$ | | | | | | | |
| HKP [157] | 76.8 | 97.8 | 97.4 | 100.0 | 89.2 | 97.7 | 0.7 |
| Yang [85] | 73.2 | 97.6 | 97.6 | 92.1 | 85.2 | 89.5 | 0.8 |
| MA [37] | 78.6 | 97.8 | 97.4 | 100.0 | 93.9 | 86.3 | 0.89 |
| VSK [158] | 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 5.6** Comparison of building detection results for the VH dataset. Object-based $C_m$ = completeness, $C_r$ = correctness ($C_{m,50}$ and $C_{r,50}$ are for buildings over 50 m$^2$) and pixel-based $C_{mp}$ = completeness and $C_{rp}$ = correctness are in percentages. $RMSE$ = planimetric accuracy in metres.

technique obtained better or slightly lower performance than the counterparts in all three areas.

The proposed roof plane extraction method, in all areas of VH dataset, offers better per-object completeness and correctness than the VSK and TUD, as shown in Table 5.7. However, MA remains slightly better when all the planes are considered in VH1 and VH2. Concerning planes larger than 10 m$^2$, Table 5.7 shows that our technique achieves better correctness but has performs slightly more poorly than MA on the VH2 and VH3 in terms of completeness. Nevertheless, it performs better than VSK and TUD on the VH1 and VH3 when performance is quantified as completeness ($C_{m,10}$). The planimetric accuracy ($RMSE$) and height error ($RMS_Z$) of the proposed technique do not much differ much from the counterparts but the $RMS_Z$ difference in VH2 is larger because the proposed technique do not remove noise to avoid reducing point density of the data.

| **Methods** | $C_m$ | $C_r$ | $C_{m,10}$ | $C_{r,10}$ | $M/N/B$ | $RMSE$ | $RMS_Z$ |
|---|---|---|---|---|---|---|---|
| | | | VH1: 3.5 laser points/m$^2$ | | | | |
| MA [37] | 76.4 | 83.3 | 84.4 | 84.9 | 6/42/7 | 1.05 | 0.41 |
| VSK [158] | 72.2 | 96.7 | 80.3 | 95.9 | 7/42/6 | 0.9 | 0.3 |
| TUD [159] | 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 |
| | | | VH2: 3.9 laser points/m$^2$ | | | | |
| MA [37] | 73.9 | 91.9 | 93.8 | 92.6 | 7/3/1 | 0.74 | 0.37 |
| VSK [158] | 73.9 | 100.0 | 91.7 | 100.0 | 3/5/1 | 0.7 | 0.3 |
| TUD [159] | 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 |
| | | | VH3: 3.5 laser points/m$^2$ | | | | |
| MA [37] | 82.1 | 93.9 | 92.7 | 96.7 | 5/45/0 | 0.89 | 0.27 |
| VSK [158] | 76.6 | 99.1 | 86.3 | 100.0 | 3/50/0 | 0.8 | 0.1 |
| TUD [159] | 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 5.7** Comparison of plane results for the VH dataset. Object-based $C_m$ = completeness, $C_r$ = correctness ($C_{m,10}$ and $C_{r,10}$ are for buildings over 10 m$^2$). M = over-segmentation, N = under-segmentation, and B = both over- and under-segmentation in number of buildings; $RMSE$ = planimetric accuracy; $RMS_Z$ = height accuracy; in metres.

## 5.4 Summary

This chapter focuses on the automatic detection of buildings and their roof planes, and defines the three steps: feature preservation, surface growing, and false plane elimination. The proposed technique is data-driven and introduces a feature preservation-based segmentation algorithm. This method uses robust saliency features, which are less sensitive to noise and avoids over- and under-segmentation, for the extraction of roof planes. A boundary tracing algorithm is also proposed, which approximates boundary of objects using LiDAR points.

The proposed segmentation technique achieves high building detection and roof extraction performance on several datasets of variable point densities, terrains, and surrounding complexities. The use of robust saliency features enables the proposed method to separate buildings and roofs from connected vegetation. The results show that the proposed technique is capable of detecting small roof planes as well as buildings. This technique since extracts rooftops and buildings directly using LiDAR data, the planimetric accuracy of the detected polygons is limited by its horizontal point spacing. In the next chapter, we propose a building reconstruction technique that utilises the extracted roof planes for the development of 3-D building models.

Chapter 6

# Reconstruction of 3-D Building Models

*"It's not what other people believe you can do; it's what you believe you can do."*

Gail Devers

## 6.1 Introduction

3-D building roof reconstruction i.e., the fourth research object **RO4**, has been a topic of active research for more than a decade. Most early roof reconstruction methods were semi-automatic, with the involvement of a trained human operator who performed accurate measurements. However, human intervention reduces the speed of execution in achieving high productivity and in processing large datasets. Recently, several building reconstruction systems have been developed and reviewed in Section 3.5 which show the construction of definitive models to be a non-trivial task due to variations in geometric and functional descriptions of buildings. We also noted that some methods suffer from high reconstruction errors, while others are able reconstruct only a certain type of building.

The previous chapter introduced the main aspects of point cloud segmentation and discussed the reasons and methodology used in low-level processes to recognise and extract building rooftops, which have meaningful correspondence with the buildings actually present in a scene. It is now time to move to more model-oriented representation of buildings, which is carried out here in the reconstruction part of the thesis. This chapter utilises the roof planes from Chapter 5 for the development of 3-D building models. The proposed technique is entirely data-driven and constructs buildings represented at lower levels with coarse boundaries (3-D roof-planes) to higher levels (3-D building models). The method is tested on two benchmark datasets consisting of many complex buildings, and the outcome confirms the applicability of the proposed method on more heterogeneous areas. The reconstruction strategy consists of a number of intermediate, interrelated processes to form a framework in such a way that every process provides more object-related information to its immediate higher level process.

Section 6.2 describes the complexities involved in reconstruction of building models. The proposed building modelling technique is presented in Section 6.3: identification of local neighbouring roof planes is described in Section 6.3.1, and both approximation of missing planes and extraction of 3-D intersection lines among roof planes are introduced in Section 6.3.2. Section 6.4 provides details of building rooftop adjacency relationship and roof model generation processes. Section 6.5 presents a visual assessment of the proposed method using two datasets and its applicability in reconstructing polyhedral buildings. Section 6.6 concludes the chapter.

## 6.2 Complexity of 3-D Modelling

3-D building modelling techniques differ significantly, based on the primitive shapes used and the input data sources [160]. Nonetheless, the vast majority of reported methodologies for 3-D building modelling follow a specific succession of processing steps: segmentation, building recognition, roof extraction, and 3-D geometric modelling. Irrespective of the methodology, the complexity of building reconstruction is heavily dependent on several factors, including the resolution of the input data source, the algorithms adopted and the specific requirements of the application. Therefore, it is essential to choose an appropriate model complexity, because buildings in urban regions have a diversity of shapes and levels of detail. For example, applications for flood and disaster management or city planning focus on terrain shape, while those for the assessment of solar potential are more concerned with the orientation of roof planes.

The Open Geospatial Consortium has developed the international *Level of Detail* (LoD) standard described in OpenGIS City Geographic Markup Language (CityGML), with the aim of not only providing visualisation of the geometric characteristics of city models, but also enabling the analysis of the same entity at different degrees of resolution [29, 161]. Therefore, CityGML supports five different well-defined consecutive LoDs (from $\text{LoD}_0$ to $\text{LoD}_4$) for the multiscale representation of city entities, where a higher level corresponds to more detail of the building model. Figure 6.1 illustrates the consecutive levels of detail of a real building.

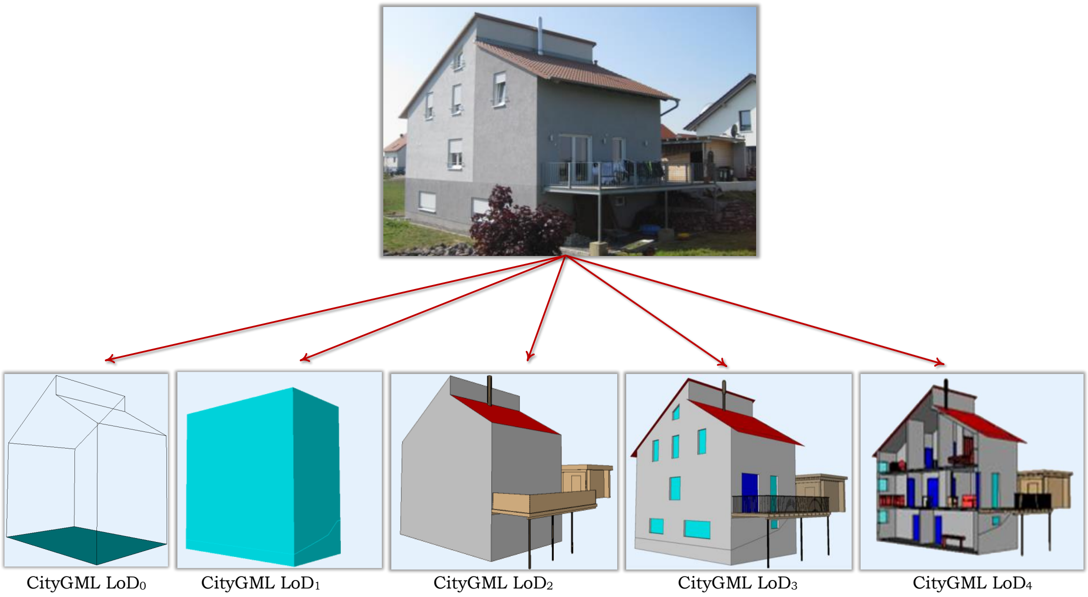



**Figure 6.1** Representation of a real detached house using $\text{LoD}_0$–$\text{LoD}_4$ [162, 163]

The coarsest level of detail ($\text{LoD}_0$) is essentially a DTM by which a building is described by its footprint or roof outline. It further shows that there is no immediate association of the LoD with different building structures. The first level of detail ($\text{LoD}_1$) represents a building as a block model i.e., a vertical extruded solid from DTM to a certain height of LiDAR points without any semantic structuring. In comparison, a building with the second level of detail ($\text{LoD}_2$) has a more differentiated rooftop and thematically-separated boundary surfaces. The region of a building is represented by a geometrically simplified exterior shell/wall. $\text{LoD}_3$ is characterised by the inclusion of architectural components with detailed wall and roof structures e.g., doors, windows, balconies, bays, and projections. In addition, image textures can be mapped onto these structures of the constructed model. Finally, in addition to $\text{LoD}_3$ representation of the building, $\text{LoD}_4$ adds the interior structures of 3-D objects such as rooms, stairs, interior doors, and even furniture.

The LoD attribute of a model is actually a sign of quality, the LoDs describe varying levels of accuracy. In other words, LoD is a measure of the consistency between real world features and modelled features, both geometrically and semantically. Furthermore, the transition of detail of a model from $LoD_0$ to $LoD_2$ indicates the 2.5 dimensional characteristics that are mainly the result of the use of airborne LiDAR or multispectral images. In contrast, $LoD_3$ and $LoD_4$ exhibit complete 3-D representations of city objects, indicating a need for more data sources to obtain the information on the façades and interiors of buildings. In relation to data sources, Elberink et al. [65] reasoned that the resolution of LiDAR data limits the detailed representation of modelling objects which cannot be detected and hence be approximated during the modelling phases. They further argued that point density of less than $1{\sim}2$ points/$m^2$ is often insufficient for the development of a detailed 3-D building model, and the only viable representation left in this scenario is to reconstruct buildings using a simple box (i.e., $LoD_1$).

In the context of data resolution, the level of detail of a 3-D model is leveraged by the level of detail achievable with LiDAR data. Therefore, only rough and generalised features can be reconstructed, since small structures cannot be determined if the data source has low resolution. We further observed (and discussed earlier in Section 2.2.3) that the use of input data sources plays a key role in attaining the desired level of detail for 3-D building reconstruction. Theoretically, no single level of detail exist that can fulfil the requirements of all the applications using different data sources and their constituent properties.

## 6.3 Proposed Methodology

In the context of model generation, it is essential to choose an appropriate model complexity because urban objects have a diversity of shapes and details. In addition, it needs to be taken into account that the LoD within a 3-D model is subject to the resolution of the input LiDAR data. When given data have large point spacing, only rough and generalised features can be extracted since tiny structures cannot be measured.

The proposed automatic 3-D building modelling technique utilises the roof planes extracted previously in Chapter 5 and the corresponding DTM as the input data. We use the initial version of the technique [164] that could not extract small roof planes instead of the final results. The reason for this is to comprehensively test the ability of the proposed modelling technique to approximate missing roof planes for the successful reconstruction of building models. As we do not use a model-driven approach, the proposed technique does not rely on pre-defined building models. Figure 6.2 shows the workflow of the proposed 3-D building modelling technique.

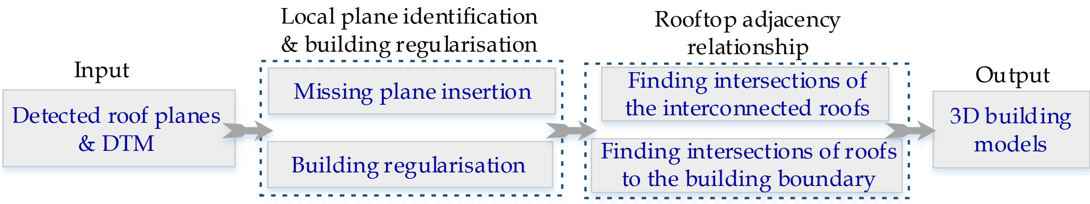



**Figure 6.2** Workflow of the proposed 3-D building modelling and scene reconstruction technique.

Our technique relies entirely on airborne LiDAR data, and the building models that can be developed while providing a good level of detail is $LoD_2$. Therefore, we first need to find any possible missing roof plane, since we know that the roof plane extraction procedure generally rejects planes smaller than a user-defined specification and often do not segment the LiDAR points which do not meet the segmentation criteria. The boundary of each building object is then approximated using the information of the extracted roof planes and the boundary is regularised to develop its building footprint. The subsequent procedure determines the interrelations between the interconnected roof planes and the building periphery. Finally, these intersection points are used to reconstruct a building model. The following sections provide more detail on these intermediate processes and illustrate the concepts with appropriate figures.

### 6.3.1 Adjacency of Roof Planes

The primary elements for the generation of building models are the roof planes which are input to the proposed technique. In the context of 3-D objects, a building can generally be perceived as a complex arrangement of different roof planes which are mutually interconnected. Information about the interconnection needs to be ascertained before we actually model a building using only the available information i.e., the roof planes. Therefore, we establish a neighbourhood relation matrix among the roof planes to determine the topological relations between these 3-D roof primitives, which is originally an adjacency matrix and stores the records of the neighbours of each roof plane. Figure 6.3(a) presents the AV1 dataset for the demonstration of different stages of the proposed methodology and the construction of a neighbourhood adjacency matrix.

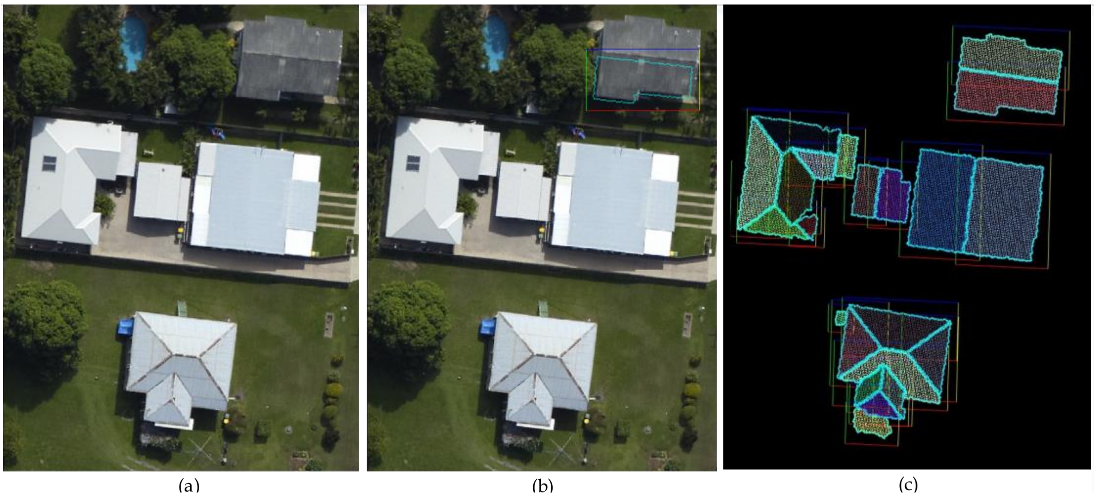


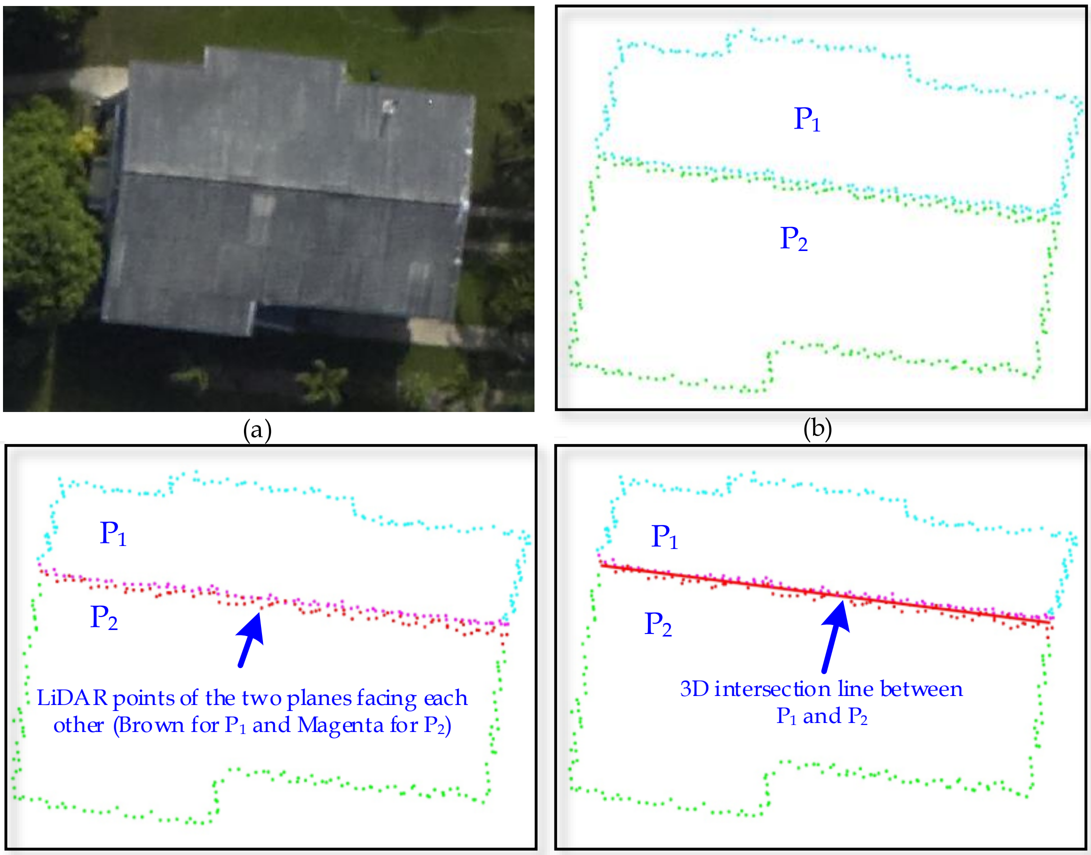

**Figure 6.3** Steps for generation of neighbourhood adjacency matrix: (a) Image of the test dataset; (b) Identification of neighbouring planes - input plane (cyan) enclosed in a rectangular region; and (c) LiDAR representation of input planes and rectangular regions for selection of neighbouring roofs for entire dataset.

Generally, the roof planes extracted using LiDAR data have low horizontal accuracy, due to the sparsity and irregularity of the laser pulses around the edges, as discussed previously in Sections 2.2.3 and 4.2.2.4. Therefore, the resolution of LiDAR data is kept in consideration for the determination of neighbouring planes and the construction of the neighbourhood matrix. Consider $R_p = \{r_{p_1}, r_{p_2}, ..., r_{p_n}\}$ is a set of $n$ input roof planes and an adjacency matrix $M$ of the same size is instantiated i.e., $M_{n\times n}$. The roof planes that remain within the Euclidean distance $dist_p$ of a source plane $r_{p_i}$, are considered its neighbours and the corresponding rows and columns of $M$ are updated accordingly with the roof plane's ID. The value of $dist_p$ is chosen as twice the maximum point-to-point distance ($d_{max}$) of the input LiDAR point cloud. To speed up the generation of $M$, we first determine the planes within the imaginary rectangular region around the source plane $r_{p_i}$ (see Figure 6.3(b)) rather than computing the distances of $r_{p_i}$ from the rest of the input roof planes $r_{p_{n-i}}$. Next, the distances between the boundary points of $r_{p_i}$ and the planes within the rectangular regions are determined. The particular records of the roof planes against the $i^{th}$ row and column of $M$ are updated which lie within the $dist_p$. The procedure continues and all the input roof planes are processed iteratively to establish the interrelations among them, as shown in Figure 6.3(c).

### 6.3.2 Identification of Intersection Lines

The procedure to determine the *roof topology* and the *intersection lines* between the adjacent planes is demonstrated in Figure 6.4. The multispectral image of the sample building and its corresponding input roof planes are shown in Figures 6.4(a) and 6.4(b), respectively. For a roof plane $P_1$, its adjacent plane(s) are determined from $M$, which in our example is single plane i.e., $P_2$. Next, the process uses the plane equations of both $P_1$ and $P_2$ to assess whether or not these planes mutually intersect in 3-D space. If they do, we obtain a point called the *intersection point* $I_{pnt}$ and a direction vector $\hat{n}$ in 3-D space. Subsequently, straight lines are approximated using the roof boundary points which face each other, as shown in Figure 6.4(c). We used Matlab's `polyfit` built-in function for the approximation of line segments. Following the concepts of 3-D coordinate geometry and using the approximated straight lines, $I_{pnt}$, and $\hat{n}$ altogether, we estimate 3-D intersection lines between the adjacent roof planes, as shown in Figure 6.4(d).

Success in extracting the intersection lines is entirely governed by the user-defined threshold of the parameter $I_{pnt}$. To understand how, imagine a building that has two multistorey (superimposed) roof planes, such that both are adjacent in 2-D space but reasonably distant because of their different heights in 3-D space. These planes, when a vertical plane joining them is not present, will definitely intersect each other (following the above mentioned principles), but the location of $I_{pnt}$ will be irrational compared to their real intersection point. Therefore, the value of $I_{pnt}$ is chosen entirely according to the resolution of the input LiDAR data, which is $2d_{max}$. The *plane insertion procedure* first assesses the value of $I_{pnt}$ of any two adjacent planes, which if it exceeds the user-defined threshold, will potentially indicate either of two general possibilities: 1)



(a) (b) (c) (d)

**Figure 6.4** (a) Test building image for demonstration; (b) Input roof planes for the corresponding building; (c) Highlighted LiDAR points of roof planes intersecting each other; and (d) Intersection line between two neighbouring planes.

a missing LiDAR-based plane or 2) a missing vertical plane, between these adjacent roof planes. Both of these two possible scenarios are described using diagrams in the following sections.

### 6.3.3 Detection of Missing Roof Planes

Recall that in the segmentation procedure describe in Section 5.2.2, point cloud data were segmented following the similarity criterion for the extraction of roof planes. Since the proposed segmentation method prefers buildings with planar surfaces, some small planes intersecting in a close proximity, e.g., hexagonal gazebo dormers and other non-building structures, were eliminated as invalid roof planes. In addition, planes with an area less than 1 $m^2$ were also removed by the refinement procedure described in Section 5.2.3. An accurate building model in cases of missing planes cannot be reconstructed because of the erroneous and inconsistent roof topological relationship. Therefore, the identification of any prospective missing building primitive is imperative before we move on to any procedure for the construction of 3-D building models.

#### 6.3.3.1 LiDAR-based Roof Plane Insertion

The proposed modelling technique iteratively takes a plane and its neighbours to approximate their intersection lines. However, if the value of $I_{pnt}$ is more than the user-defined threshold, the *plane insertion process* attempts to search for any unsegmented LiDAR points between the participating roof planes. If such points are found, the process infers the presence of a plane or at least a non-planar surface between these roof planes. LiDAR-based roof plane insertion (scenario 1) is graphically explained with reference to Figure 6.5, and the sample building, all the roof planes, the location of missing plane, and the participant planes $P_1$ and $P_2$ are sketched in 6.5(a) and (b).

The process invokes a region-growing segmentation technique to extract a planar region using the unsegmented LiDAR points (see the green points in Figure 6.5(c)). In addition to the available points, the segmentation process uses points of the neighbouring planes ($P_1$, $P_2$, and $P_3$) for the extraction of new planes assuming these points might have been added wrongly to neighbouring planes due to different settings of the algorithm. Therefore, each iteration of the region-growing technique computes a plane-fitting residual of new and neighbouring planes. If the new plane results in a height error smaller than those of the neighbouring planes' errors, the LiDAR points of the neighbouring planes are removed from their respective regions.

The segmentation process continues growing the region until it finds points complying with the height residual criterion. After the segmentation process stops, the proposed technique estimates the boundary of the new plane and updates the boundary information of the neighbouring planes, as shown in Figure 6.5(c). The process also updates the neighborhood matrix $M$ with the information on the new plane and new neighbouring relations. Subsequently, the intersection lines between the identified plane and the participating planes are estimated using their boundary points, as

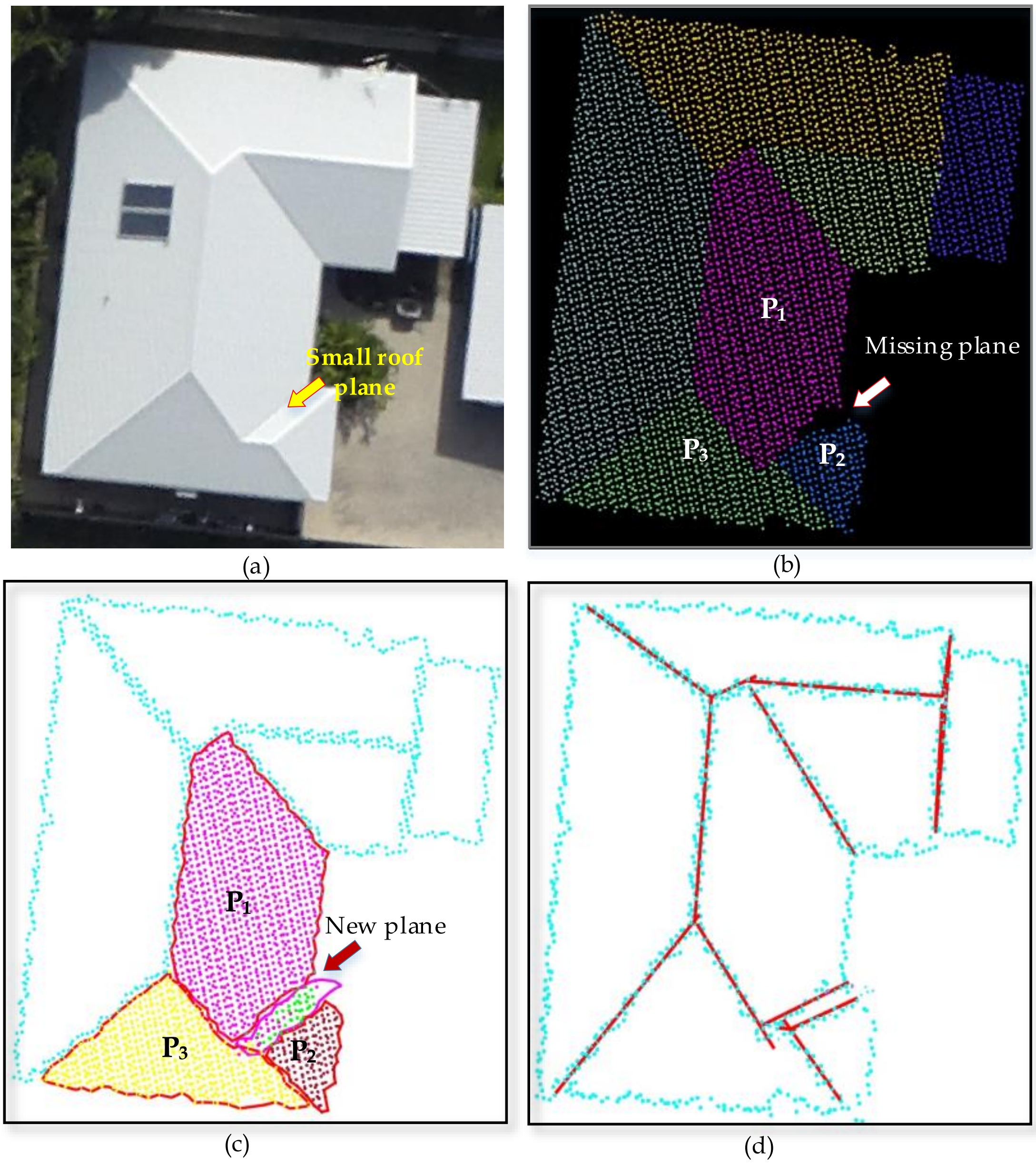



**Figure 6.5** (a) Test building image for demonstration; (b) Input roof planes for corresponding building and the location of a missing plane; (c) New roof plane using unsegmented LiDAR points (green) and points from neighbours; and (d) 3-D intersection lines between roof planes.

explained in Section 6.3.2. These intersection lines are also recorded against their adjacent roof planes. Figure 6.5(d) shows all the roof planes of the sample building and the intersection lines between them.

#### 6.3.3.2 Vertical Roof Plane Insertion

Section 2.2.2 outlines that airborne LiDAR data is collected by an aircraft emitting laser pulses that hit different features on the earth. However, these pulses do not directly hit the hidden vertical surfaces and hence no laser returns are captured for a surface joining two parallel roof planes and the building walls. As a result, the point cloud segmentation procedure cannot detect the presence of such surfaces. Therefore, a possible way to recover the roof topology that addresses scenario 2 is through the insertion of a vertical plane. This phenomenon is explained using a building model shown in Figure 6.6(a) which has two adjacent roofs ($P_1$ and $P_2$) and a vertical plane joining them.

The plane insertion process uses the boundary points of the participating planes $P_1$ and $P_2$, which face each other, to approximate a hypothetical vertical plane, as shown in Figure 6.6(b). It then determines the intersection lines between the vertical plane and both the planes $P_1$ and $P_2$, as shown in Figure 6.6(c), using the procedure explained in Section 6.3.2. The height of the approximated vertical plane is determined by these intersection lines, as shown in Figure 6.6(d). The plane insertion procedure also records the newly-inserted plane and the plane adjacency information in $M$. To further demonstrate the applicability of the entire procedure on real-world data, we take a sample building from the HB dataset that has LiDAR density of 12 points/m$^2$.

Figure 6.7(a) shows a sample building, a small slanted plane that locates between two adjacent roof planes labelled $P_1$ and $P_2$ in Figure 6.7(b). This plane was not detected by the proposed segmentation technique due to the unavailability of LiDAR points, as shown in Figure 6.7(b). Therefore, we use the plane insertion procedure that takes the boundary points of $P_1$ and $P_2$ to determine the intersection lines and approximate a new plane $P_4$, as shown in Figure 6.7(c). The procedure not only keeps track of new planes, but also maintains the correct neighbourhood information in $M$. Therefore, even after the insertion of $P_4$, we find through the neighbourhood selection method (described in Section 6.3.1) that $P_4$ has a new neighbouring plane i.e., $P_3$ as can be seen visually in Figures 6.7(b) and (d). There is a dire need to determine the intersection between $P_3$ and $P_4$ to precisely establish the topological relationships among the building roofs and reconstruct a model with a good LoD.

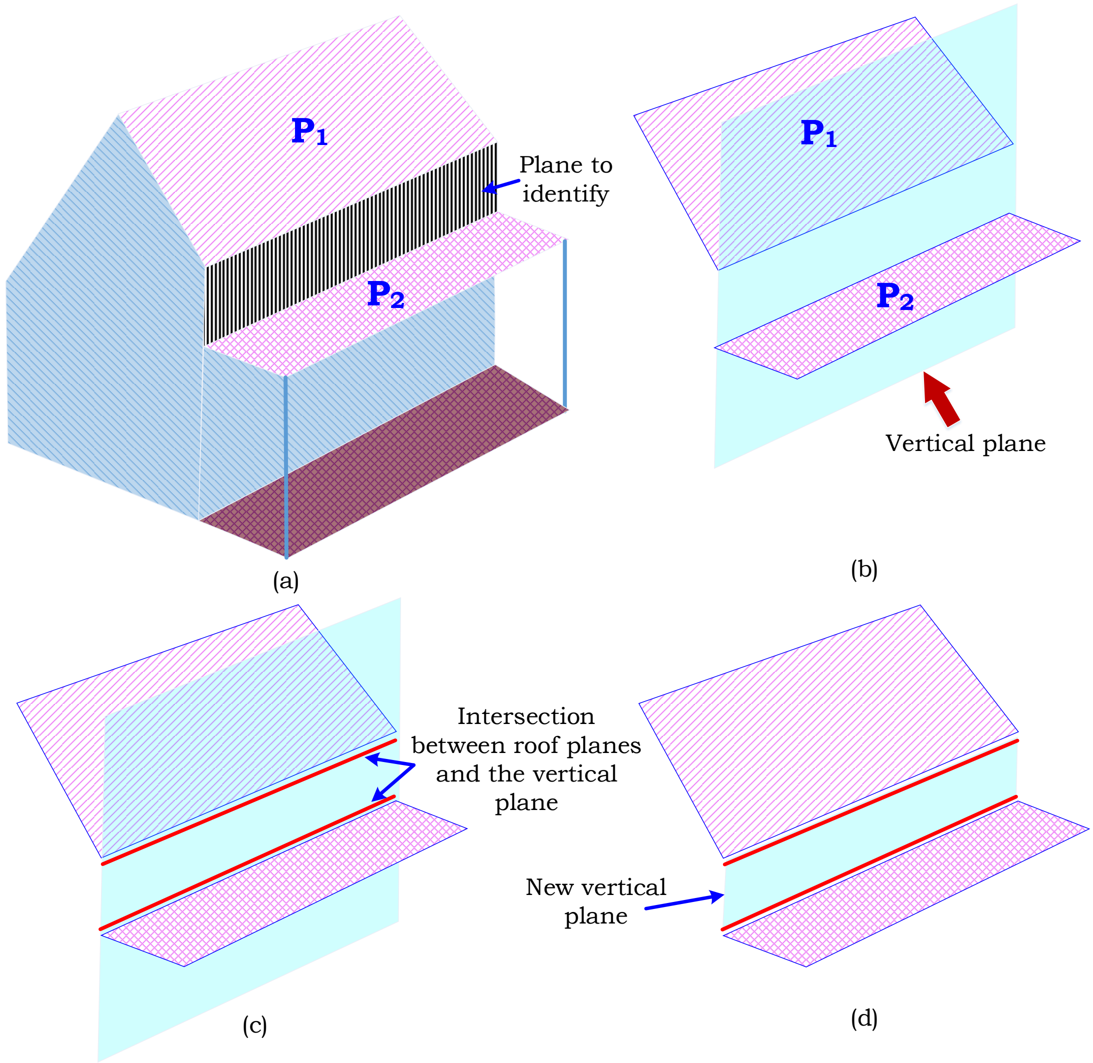



**Figure 6.6** Pipeline for inserting a hypothetical vertical plane. (a) Virtual 3-D building model; (b) Visualisation of a hypothetical vertical plane between adjacent roof planes; (c) Identification of intersection lines between vertical plane and adjacent roof planes; and (d) 3-D intersection lines and inserted vertical plane.

The plane insertion procedure is executed following the similar steps and inserts a new vertical plane between $P_3$ and $P_4$ using their boundary points. The procedure further approximates the intersection lines between the participating roof planes and the new vertical plane in 3-D space, as shown in Figure 6.7(e). The procedure stops once all the roof planes are processed. Figure 6.7(f) shows all the building roof planes and their intersection lines in 3-D space for better visualisation.

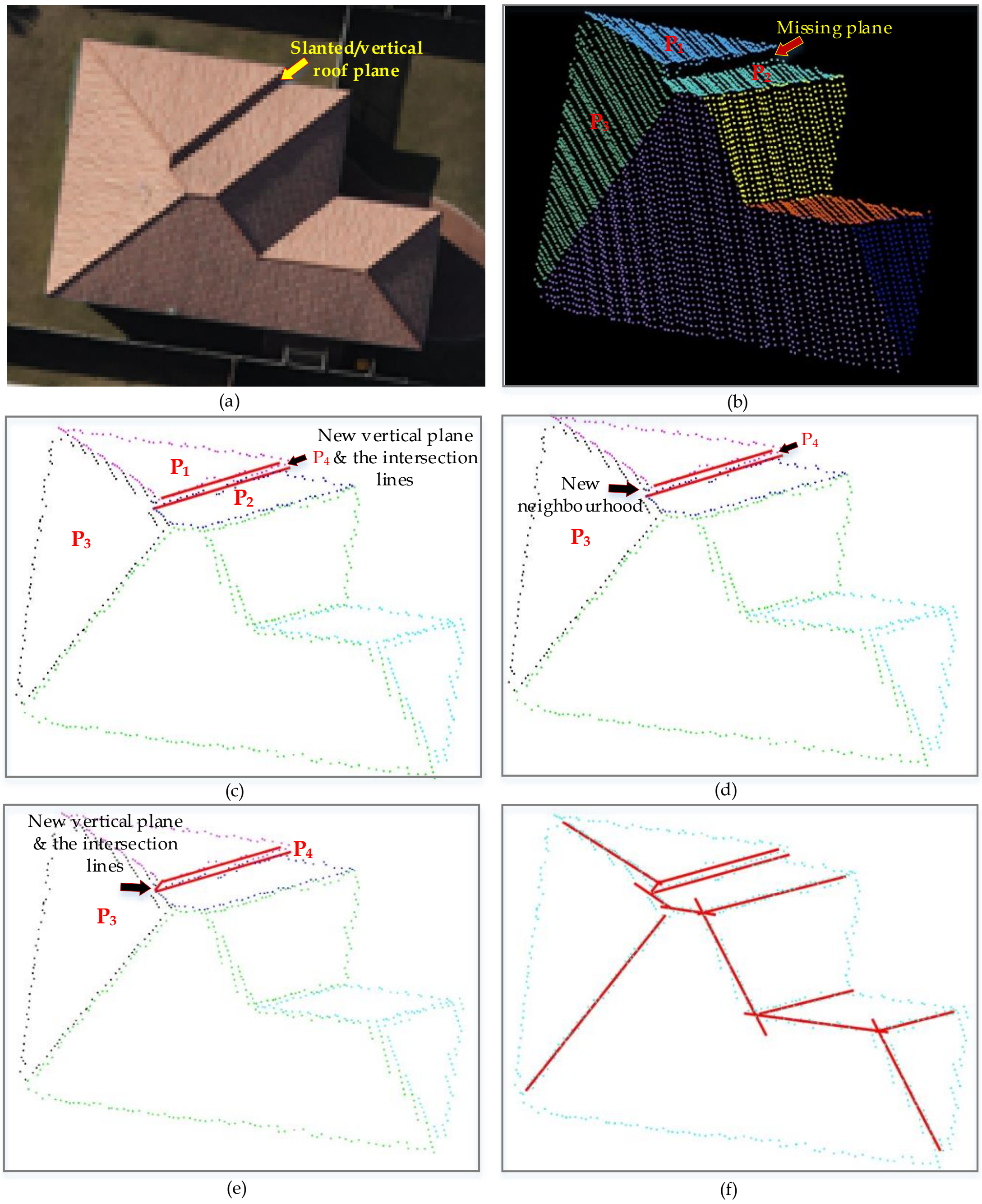



**Figure 6.7** (a) Test building image for demonstration; (b) Input roof planes for corresponding building and location of missing plane; (c) Insertion of vertical plane; (d) Adjacency of vertical plane to other roof planes; (e) Insertion of vertical plane between LiDAR-based plane and approximated vertical plane; and (f) 3-D view of building roof planes and insertion lines for construction of interrelation between roofs.

### 6.3.4 Building Regularisation

All the intersection lines can be regarded as structural lines defining the interconnection of the roof planes of a real building. We call these intersection lines the *ridge* lines. However, for the construction of a 3-D building model, information on the building periphery is as important as that on the roof planes'. By using the information on all the input roof planes and those which are approximated, we determine the boundary of the building and then regularise to develop its footprint. Recall that in the boundary tracing procedure described in Section 5.2.1, the LiDAR data corresponding to all the roof planes are collected and the number of buildings and their respective boundaries are estimated using Delaunay triangularisation. It is worth remembering that the building boundaries extracted from the LiDAR point cloud are ragged and have low horizontal accuracy. However, the regularisation of the building boundaries using the extracted buildings and the construction of their footprints are obtained using the technique proposed by Awrangjeb [165]. The entire process of footprint generation is explained using a sample image from the AV1 dataset, as shown in Figure 6.8.

Figure 6.8(a) shows the input building roof planes drawn over the corresponding building's image for demonstration purpose. The following Figure 6.8(b) shows the extracted building boundary plotted in magenta. The footprint generation process first applies the Gaussian smoothing function with a scale of $\sigma = 3$ to the extracted building boundary, which results in a boundary that is smooth and has a reduced ragged effect, as displayed in Figure 6.8(b). Next, the contour-based corner detector [166] is employed to obtain the curvature peaks along the smoothed boundary as corner points. We can see these corners in Figure 6.8(b), which actually split the smooth boundary into different curve segments. The regularisation technique then fits straight lines into each segment using a least-squares technique. These estimated lines (in green) are shown in Figure 6.8(c). The procedure next chooses a line that has the least perpendicular error from its corresponding boundary points and considers it a fixed line. All the other lines are then adjusted by making them either parallel or perpendicular to the fixed line. Finally, a building footprint is obtained after inserting perpendicular lines where needed between any two successive parallel lines. Figure 6.8(d) shows the building footprint approximated using only the LiDAR points.

## 6.4 Rooftop Adjacency Relationships and Roof Modelling

The fundamental concept of geometric modelling is to combine adjacent building roofs to develop a building model. The approach is to stitch the adjacent roofs along their common ridges. These ridges, which are actually 3-D intersection lines, here play a major role in gluing the interrelated roofs and the construction of complex building models, which require considerable attention to their topology. Therefore, the prime focus is to establish connectivity and dimensional continuity among the adjacent building roof planes in the context of topology. At this juncture, we have the information about the buildings, all their possible roof planes, intersection lines, and the adjacency relationship. This procedure is executed in two stages where the first stage determines

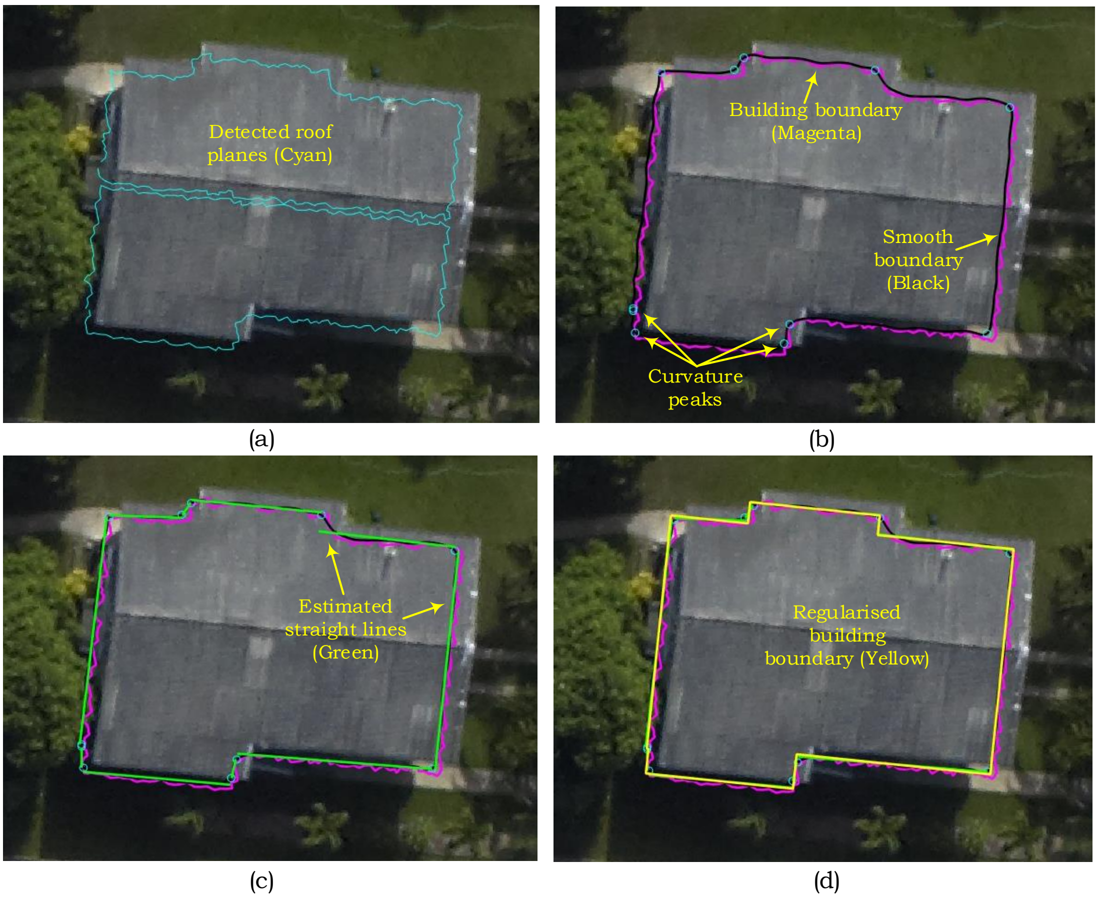



**Figure 6.8** (a) Test building image for demonstration and input roof planes; (b) Approximation of building boundary and corner points; (c) Representation of boundary segments with straight lines; and (d) Building footprint.

the intersection points among the connecting roofs using the ridge lines, while the second stage determines the intersection points of the ridges to the building periphery and the edge points.

The end vertices of the ridge lines approximate the locations of internal vertices and each successive ridge often has a small gap from the preceding vertex, as shown in Figure 6.9(a). To fill the gaps and establish connectivity among the adjacent roof planes, we use the *Roof Topology Graph* (RTG) following the principles proposed by Verma et al. [167] to identify the intersection points of the ridge lines. A topology graph is an undirected graph that is used to describe the adjacency relationships among the building rooftops. As shown in Figure 6.9(b), each rooftop is represented as a vertex in RTG and two adjacent building roofs are connected through an edge. These rooftops are labelled with their vertex numbers in Figure 6.9(a).

In the context of RTG, a basic cycle indicates an internal vertex that belongs to ridges [168]. For instance, the roof planes $P_1$, $P_4$, and $P_5$ form a basic cycle and the intersection of the corresponding ridge lines determines a ridge intersection point. We also update the corresponding vertices of the ridge lines participating in the intersection determination process. These points can also be referred as *inner* or *ridge points* and will be helpful at the later stage to approximate the model shape. An RTG can also be represented as a composition of several basic cycles, as shown in Figure 6.9(c). The building rooftop shown in Figure 6.9(a) has six basic cycles and so does the ridge intersection points. We apply the least squares approach to approximate the intersection between the ridge lines of the participating roof planes, as shown in Figure 6.9(d).

During the first stage, both the vertices of most of the ridge lines are updated and all the gaps are covered with the updated intersection points. However, in the second stage, for ridges which have one of their vertices not updated, the adjacency relationship procedure determines the intersections with the building periphery in 3-D space. These vertices are recorded as *outer* or *edge points* and are used for approximating the wall planes to develop a building model from a roof model. Figures 6.10(a) and (b) show the ridges to boundary intersections i.e., the edge points, in two different perspective views.

At this stage, we have an adjacency plane relationship, a regularised building boundary, and two groups of 3-D intersection points: inner and outer junction points.

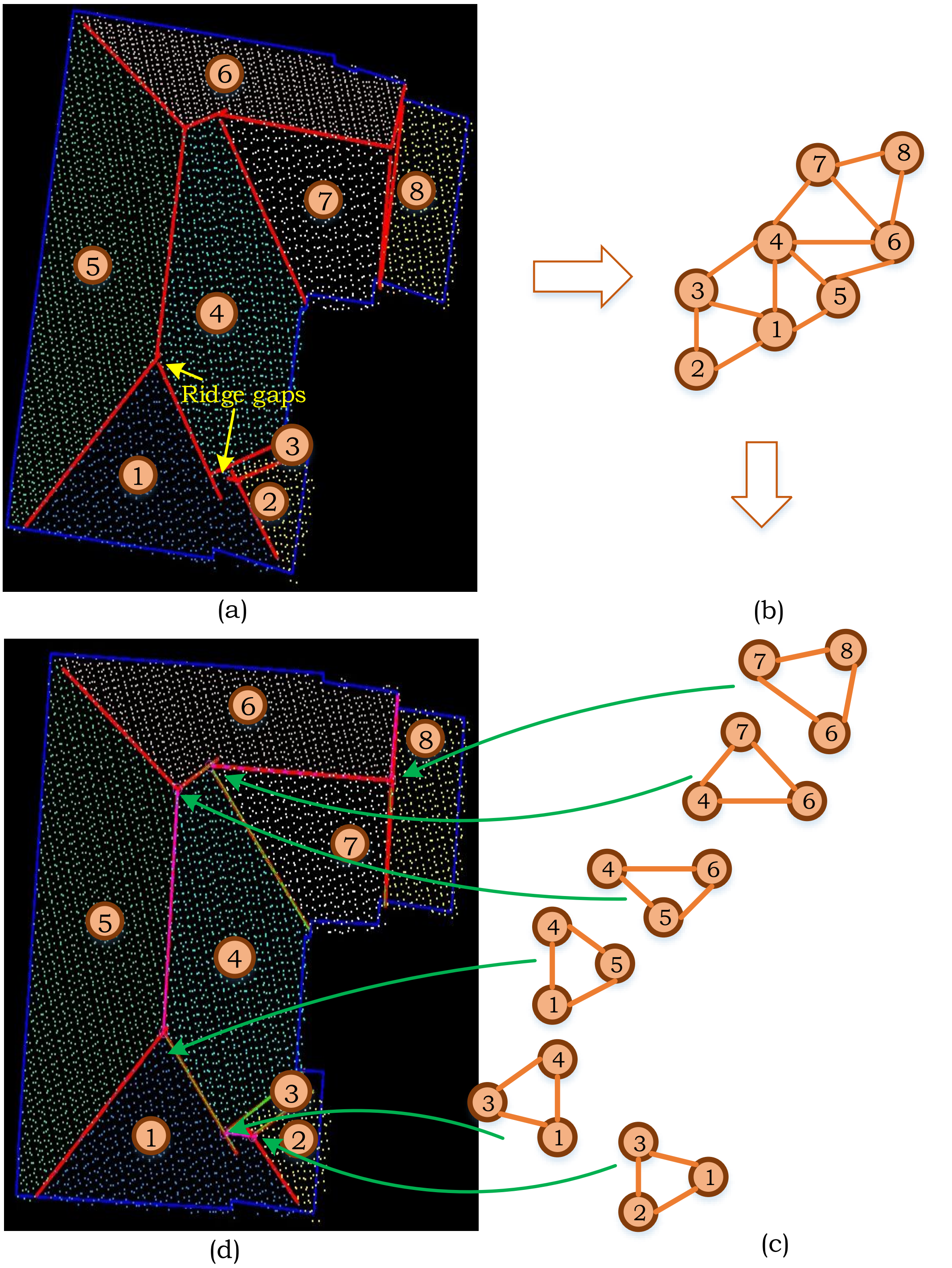
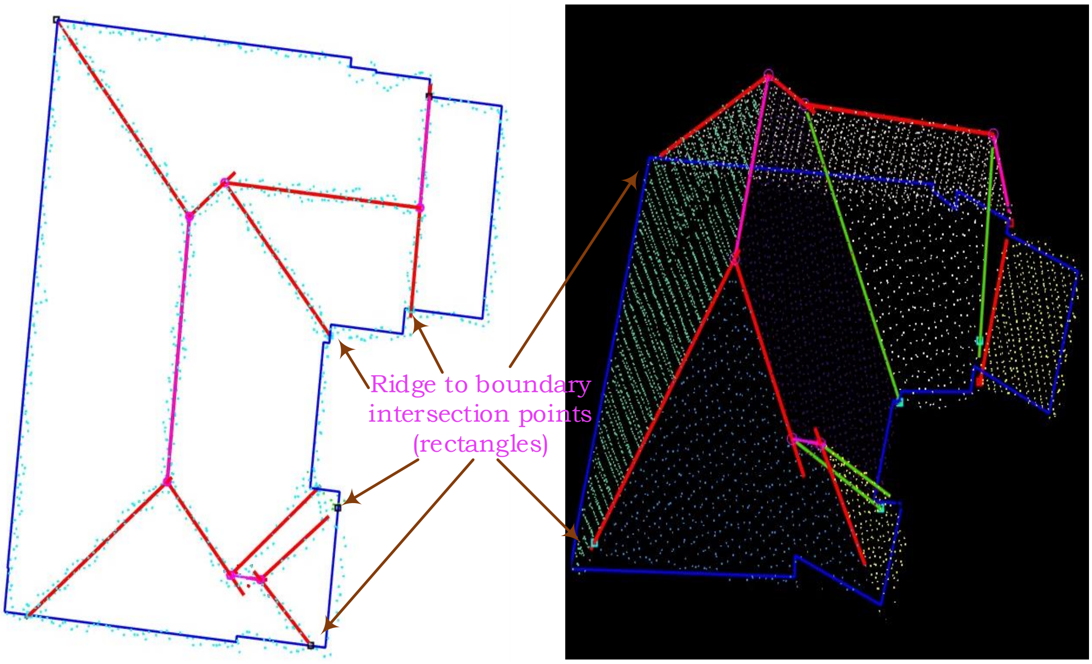



**Figure 6.9** Determination of ridge intersection points: (a) Roof planes and intersection ridges; (b) Roof topology graph; (c) Closed cycles; and (d) Corresponding ridge intersection i.e., inner/ridge points.




(a) (b)

**Figure 6.10** Determination of edge intersection points: (a) Edge points (ridge to building boundary intersection points) and (b) 3-D view of building showing edge points.

It is time to develop a building roof model before we actually develop a 3-D building model. The roof modelling mechanism is explained graphically in the following Figure 6.11. For roof modelling, each building is processed separately, and the procedure first finds the 3-D points around each plane boundary and extracts each roof segment. To do this, 3-D intersection lines whose junction points (red ovals) have been updated are recalled. Then, the junction points (edge or ridge points) of 3-D intersection lines are re-ordered in succession around the plane using the information on the corresponding LiDAR-based building boundary points. This is shown in Figure 6.11(a), where the junction points are labelled as $N_1$ to $N_5$ terminating at the starting point i.e., the start and end intersection points are the edge points. All the roof planes of each building are processed iteratively and the corresponding roof model is generated that has regularised plane boundaries, as shown in Figure 6.11(b).

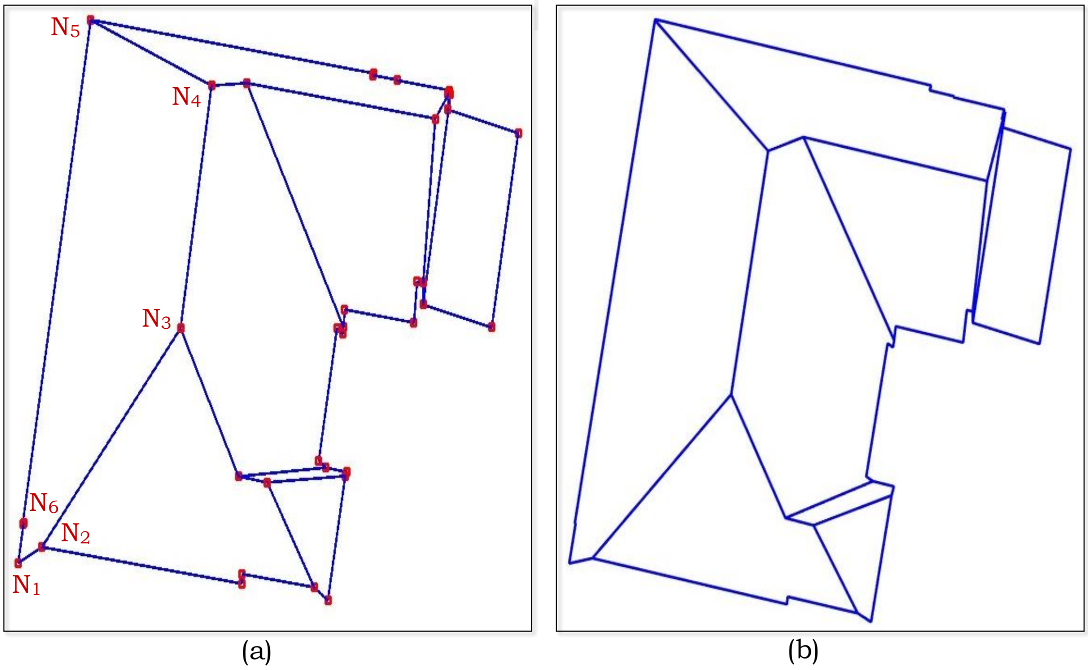
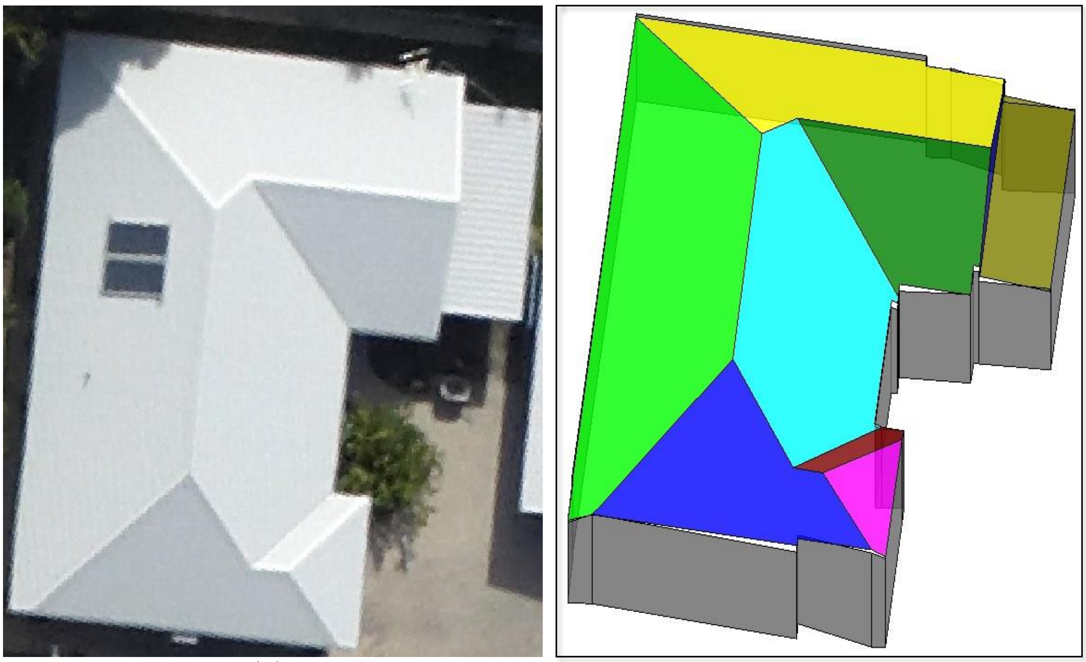



**Figure 6.11** (a) Identification of roof segments and (b) Roof model of sample building.

## 6.5 3-D Building Modelling

For building model generation, it is necessary to generate walls from the periphery of the roof model to its floor. In this regard, we use the edge points to generate the approximated building floor first. The ground height of each edge point is determined from the DTM so that the model is a replica of its respective real building. All the consecutive ground points are connected to obtain the building floor. Finally, the building walls are determined by extruding the edge points to their corresponding floor points. Figure 6.12 presents the real building and its 3-D reconstructed model, where the walls are represented in a transparent grey colour.



(a) (b)

**Figure 6.12** (a) aerial image and (b) 3-D sample building.

The proposed geometric modelling technique relies entirely on LiDAR data, while images are used throughout the chapter for the purpose of visual inspection. We tested the proposed data-driven modelling technique using two Australian benchmark datasets: AV1 and HB, and the scene reconstruction results are presented in Figures 6.13 and 6.14, respectively.

Little work has been carried out which has used the angular difference between the reference and reconstructed buildings to measure the accuracy [117, 169], which is very dependent on the dataset used. However, to date, there is no standard performance evaluation criterion for measuring the performance of building reconstruction. In addition to the absence of a standard evaluation system, 3-D reference polygons of the benchmark datasets are not available to quantitatively analyse and compare the performance of the proposed 3-D modelling technique. Therefore the performance can only be assessed qualitatively through visual inspection by inspecting solutions to the limitations addressed. The evaluation of existing techniques in Section 3.5 shows that most building modelling strategies have high reconstruction error mainly because of missing roof planes, failure to assess roof topology, variations in geometric descrip-

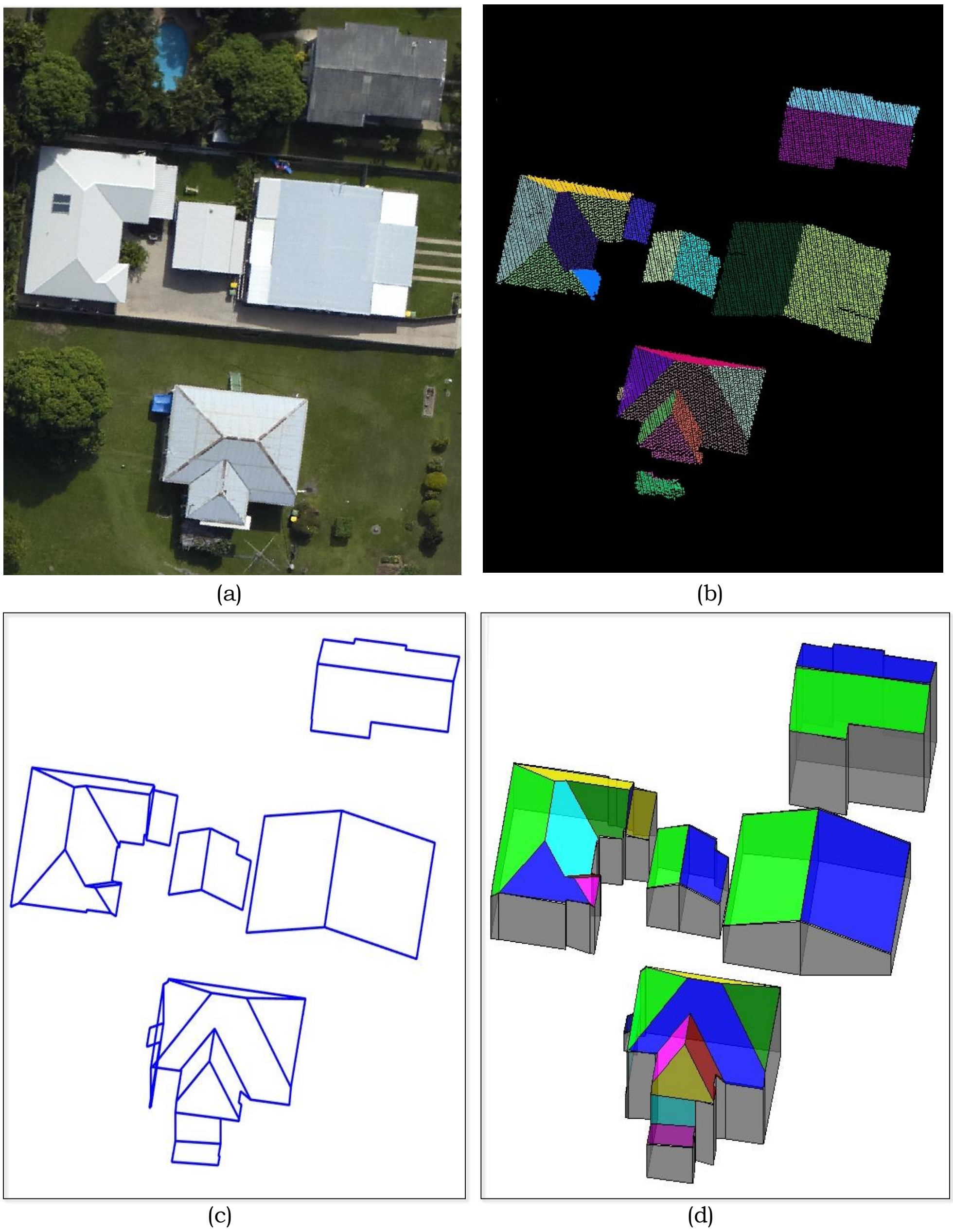



**Figure 6.13** (a) Australian benchmark site image; (b) LiDAR points of input roof planes; (c) Building roof model; and (d) 3-D modelled buildings and reconstructed scene.

tions of the buildings, and the inability to detect step edges that often limit them to the reconstruction of only certain type of buildings. In contrast, the proposed technique addresses these issues, as demonstrated using real building samples (in Section 6.3.2). The proposed modelling technique is data-driven and does not rely on any predefined shape or model, it can detect and approximate missing planes and reconstruct

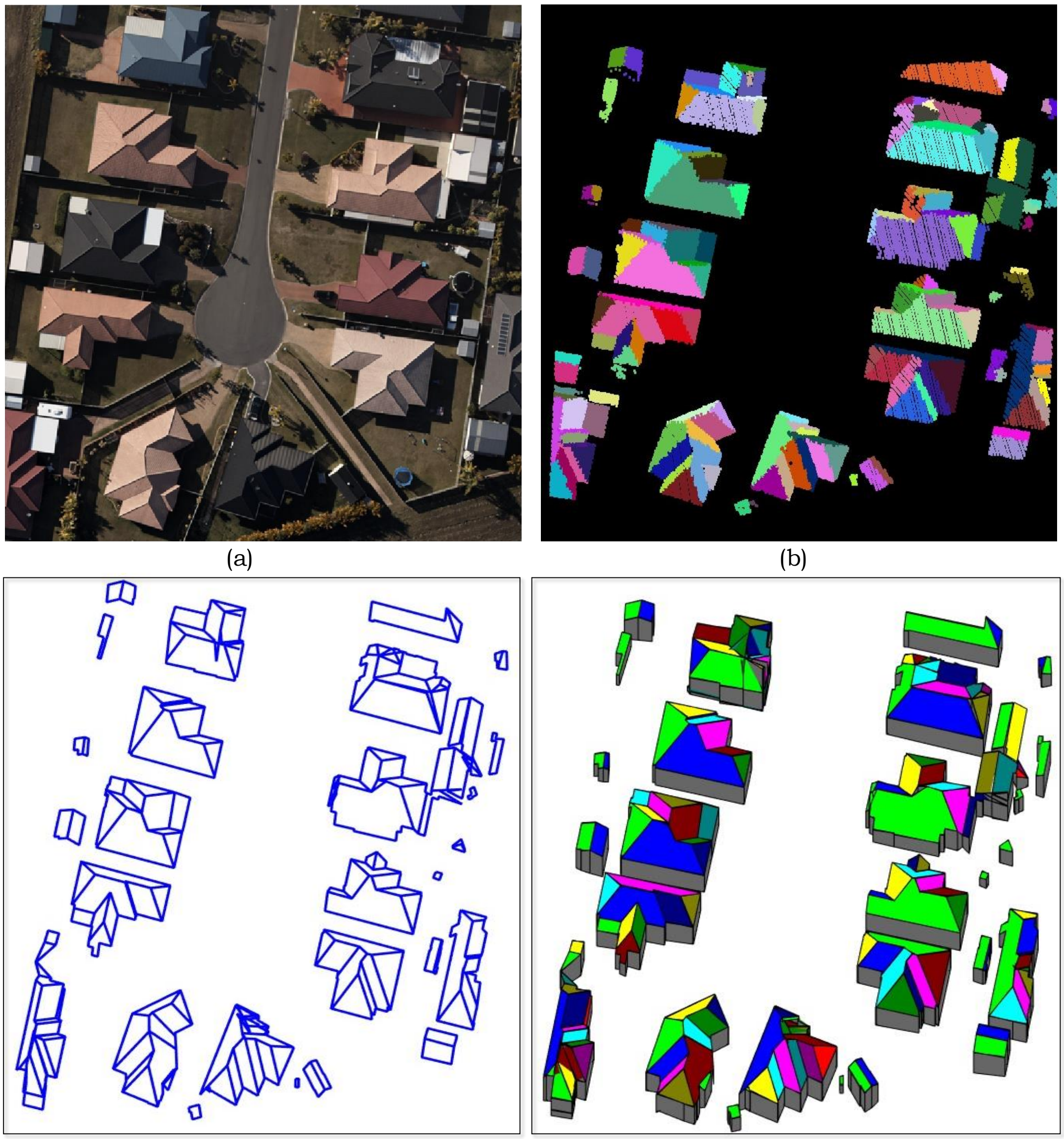



**Figure 6.14** (a) Australian benchmark site image; (b) LiDAR points of input roof planes; (c) Building roof model; and (d) 3-D modelled buildings and reconstructed scene.

models of heterogeneous building types. Figure 6.15, for instance, shows some results of slopping, gable, and pitched roofs, which validate the application of the proposed modelling technique on other datasets and its ability to reconstruct complex buildings. The first column of Figure 6.15 shows the LiDAR points of the input roof planes, and the second column presents the reconstructed building models.

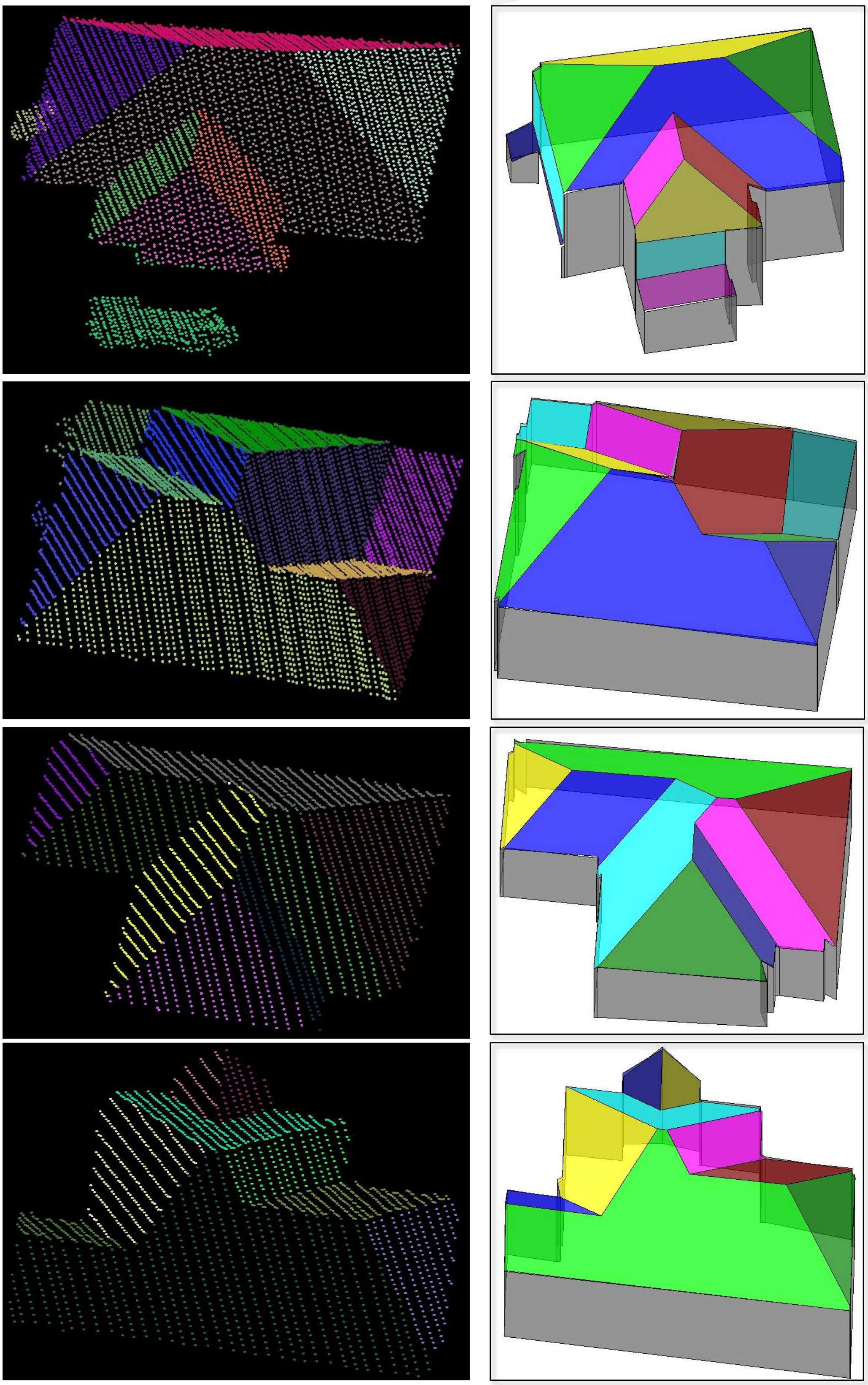


**Figure 6.15** Highlighted building models. The first column shows building roof results and the second column illustrates reconstruction building models.

## 6.6 Summary

In this chapter, reconstruction of 3-D building modelling technique has been presented. The proposed technique uses LiDAR's extracted roof planes as the only information for the construction of building models. This chapter describes the robust procedure to approximate the missing roof planes that are not extracted due to low point density, noisy data or the vertical nature of the planes. The proposed technique develops an interrelation among the building roof planes and identifies their interconnections, which are later used for the construction of building models. The method is unsupervised and data-driven, the roof types are not restricted to a pre-existing model catalogue, which can never have all possible models due to building shape variability. We have noted that since the roof boundaries used in this chapter are extracted based on alpha-shaped objects, both convex and concave building roof planes can be reconstructed.

Chapter 7

# Automatic Assessment of Solar Potential

*"At its heart, engineering is about using science to find creative, practical solutions. It is a noble profession."*

Queen Elizabeth II

## 7.1 Introduction

The previous chapters introduced the main aspects of 3-D modelling task, including building detection, vegetation elimination, building footprint generation, roof extraction, and geometric reconstruction. This chapter focuses on **RO5** and aims to develop an application for the installation of photo voltaic (PV) systems using building roofs and other geometrical information produced in Chapters 5 and 6. The study presented here was carried out by the author at the Collaborative Research Centre for Spatial Information (CRCSI)[1]. It is an international research and development centre, which conducts user-driven research in spatial information. The CRCSI provides essential services in health, energy provision, agriculture, defence, and urban planning.

According to the Australian photovoltaic Institute (APVI), the growth in the market for solar PV panels in Australia was only around 15% between 2001 and 2010, but a period of extremely rapid growth occurred from 2010 to 2013, when the reported installed PV capacity has increased from 137 megawatts (MW) to 3,897 MW [170]. Over recent years, the PV system installation price has steadily decreased, which has

[1] http://www.crcsi.com.au/

increased the average size of installations. Most PV systems in Australia are small-scale rooftop installations, most of which have been installed on residential building rooftops [170].

Since building rooftops in urban areas are interesting locations for PV system installation, an effective assessment (in both low and high level detail) for solar PV deployment on rooftops is important for both local and state governments and solar energy companies. A low level of detailed information for large areas, e.g., an administrative district, is necessary for dissemination, marketing, and sales purposes. A high level of detailed information is also crucial to energy companies, house owners and other individuals, in order to identify suitable roof surfaces and estimate the solar potential and gauge the annual return on investment for individual buildings.

Previously, the application of PV deployment assessment was limited to small areas with a limited number of buildings due to the manual or semi-automatic extraction of building rooftops. However, a high success rate in automated extraction of buildings and roof planes [40, 87, 164] has made the investigation of PV deployment in large areas possible. The proposed project investigates the application of extracted building information in PV deployment for large-scale assessment of solar potential. The project uses extracted roof planes and their boundaries (from Chapter 5) and plane equations (from Chapter 6) and calculates the area, slope (tilt), and azimuth (orientation) for all rooftops and approximate shadowing for the estimation of solar potential for individual roof planes and buildings. The development has the following aims:

1. Automatic estimation of building roof parameters, e.g., area, slope, azimuth,
2. Estimation of shadowing effects on buildings by surroundings, including buildings and trees, and
3. Estimation of annual solar potential for individual roof planes and buildings.

The following sections present a summary of the sun-earth geometry, solar angles, factors influencing solar energy, and different processes involved in shadow estimation and solar potential calculation.

## 7.2 Solar Potential Estimation and Basic Principles

The sun is the largest star in the earth's solar system and the source of energy. This energy, also known as the solar radiation, strikes the earth and heats the air, water, and soil. Before discussing solar potential, it is essential to know the relationship between the sun and the earth and the effect of incident radiation striking a surface over a specified period of time.

### 7.2.1 Sun Position

Solar energy applications need reasonably accurate predictions of where the sun is in the sky at a given time of day and year. Sun-earth geometry involves the study of the earth's rotation and revolution as well as the tilt angle of the earth's axis. As the earth rotates and revolves around the sun, there are significant seasonal and hourly positional changes of the sun and the length of day. The relative position of the sun is a major factor in the performance of PV systems. The sun's position with respect to an observer on the earth can be fully described by means of two astronomical angles, the solar altitude and the solar azimuth. Figure 7.1 describes the sun-earth geometry [171, 172], different weathers and the angles between the sun and the earth. Before providing the equations for solar altitude and azimuth angles, the solar declination and hour angle need to be defined first. These are required in all other solar angle formulations.

#### 7.2.1.1 Basic Earth-Sun Angles

The *declination angle* $\delta$ is the angle between a line connecting the centre of the sun and the earth and the projection of that line on the equatorial plane. This angle varies from $+23.45^\circ$ to $-23.45^\circ$ throughout the year. The following equation (developed from work by Spencer [173]) describes this angle, depending of the day of the year:

$$\delta = 23.45^\circ \times sin\left[\frac{360}{365} \times (n + 284)\right] \tag{7.1}$$

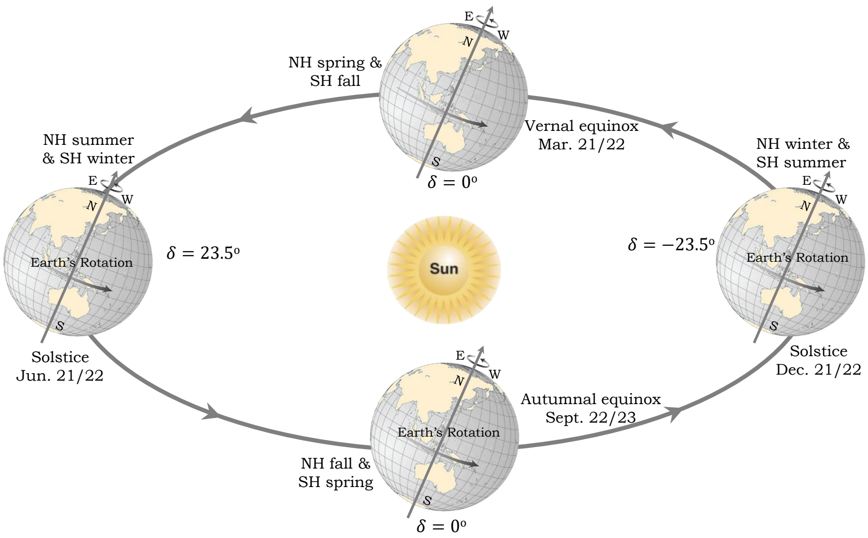



**Figure 7.1** Annual orbital motion and sun-earth geometry. NH and SH stands for northern and southern hemispheres, respectively. [171, 172]

In Equation 7.1, $n$ is the day number i.e., $1 \leq n \leq 365$ (e.g., n = 1 means January $1^{st}$). The other requisite parameter, the hour angle, depends on the longitude and the time of day. It is the angle through which the earth rotates since solar noon and is defined as the angular distance between the meridian of the observer and and the meridian whose plane is parallel to the sun's rays. Since the earth rotates at $360^\circ/24$ hours, i.e., $15^\circ$/hour, the *hour angle* $\omega$ is positive in the evening and negative in the morning, and is given by [171, 172, 174]:

$$\omega = (localtime - 12) \times 15^\circ \tag{7.2}$$

#### 7.2.1.2 Factors influencing Solar Irradiance

The quantity of solar radiation that reaches a surface is influenced by several factors, including the position of the sun in the sky and the clearness of the atmosphere, as well as the nature and the orientation of the surface. As shown in Figure 7.2, part of the incident energy is scattered and absorbed by air molecules, clouds, and other particles in the atmosphere. The radiation that is not reflected or scattered and reaches the surface directly is called *direct irradiation* $I_b$. The scattered radiation which reaches the ground is called *diffuse irradiation* $I_d$. Some of the radiation is reflected from the ground onto the receiver; this is called *reflected irradiation* $I_r$. The *total irradiation* $I$ that strikes an absorber is the summation of these components. The term *solar irradiance* is also referred to as *solar insolation* and both are used interchangeably throughout the literature and the unit of estimation is watts per square metre ($W/\mathrm{m}^2$). In this project, we do not consider $I_r$ because its impact on urban areas is negligible [172, 175, 176].

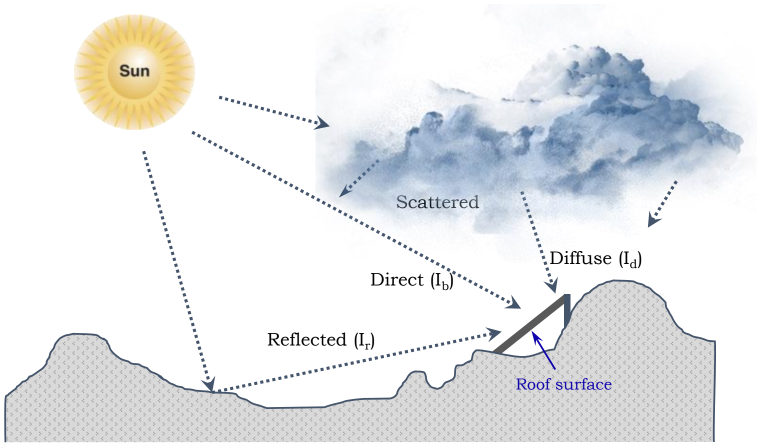



**Figure 7.2** Components of solar radiation.

### 7.2.2 Solar Potential Estimation

The solar potential estimation technique in this chapter is based on several research studies and closely follows two recent methods [175, 176] for the calculation of building roof parameters and solar irradiance approximation. The solar position as observed from a point on the earth can be defined by two angles, the *solar altitude* $\alpha$ and *azimuth* $z$ [171, 177]. The solar altitude or elevation is the angle between the horizontal plane and the central ray from the sun. The solar azimuth is the angle produced by the projection of the central ray from the sun on the horizontal plane and the south axis. The two angles can be expressed as functions of the location latitude, solar declination angle, and the hour angle [171, 172, 175].

The estimation of solar potential requires a normalised azimuth angle for each building surface/plane. Therefore, the azimuth angles calculated are normalised with respect to true north following the template shown in Figure 7.3. The azimuth angle/orientation of the roof plane increases in clockwise direction.

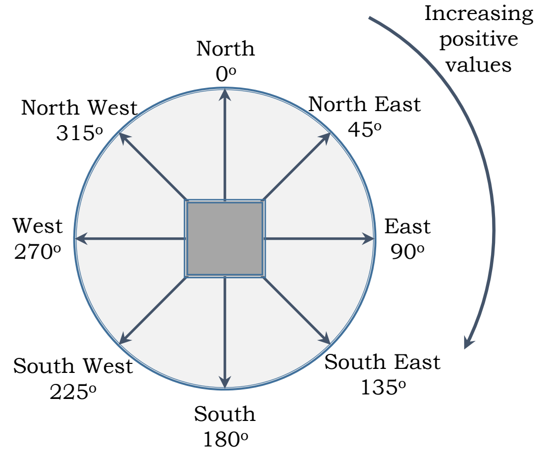



**Figure 7.3** Estimation of azimuth angle of the roof planes.

In addition to solar angles, an accurate estimation of the roof parameters is crucial for the good approximation of solar potential. These parameters are estimated following Figure 7.4 which shows a roof's normal vector $\vec{n_p}$, the normal vector for horizontal surface $\vec{n_z}$, slope $\beta_p$, aspect/orientation $\gamma_p$, angles of incidence for tilted surface $\theta_p$, the horizontal surface $\theta_{p_z}$, and solar irradiance ($I_b$ and $I_d$).

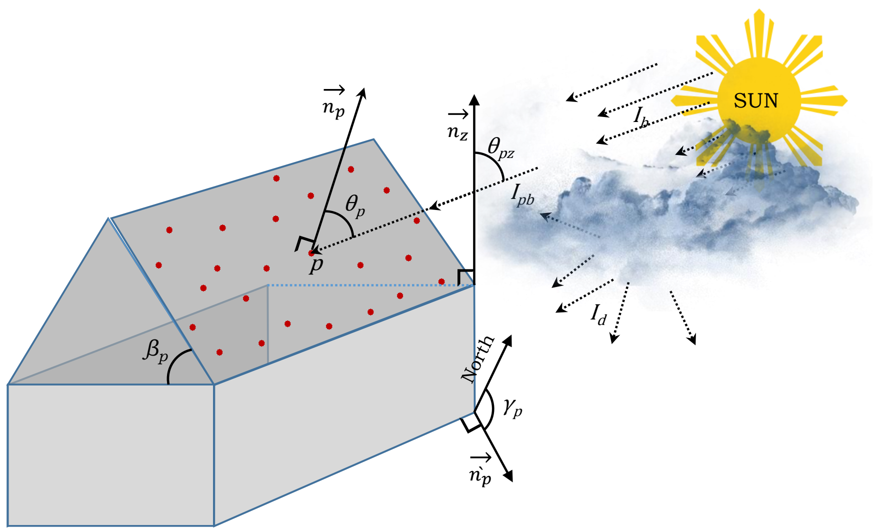



**Figure 7.4** Illustration of solar irradiance on a building plane that has LiDAR points.

Beyond the earth's atmosphere, solar irradiation is almost constant [171]. Equation 7.3 shows the value for the solar constant $s_c$ outside the atmosphere:

$$s_c = 1367 \tag{7.3}$$

There are several factors which affect the incident radiations striking buildings and other objects on the earth. Generally, a pyranometer is used for the measurement of solar radiation at a given location [175, 176]. This device collects data over several years which are then utilised for the production of solar energy applications. The measured data are usually available as hourly average values and contain the influence of atmospheric scattering, air thickness (air mass) within the atmosphere, and cloud cover.

The data for direct and diffuse radiation are not available for our study and, therefore, these data are generated empirically, for both the horizontal and the tilted planes, following the formulas in [172, 175–177]. The estimation requires to use theoretical values for transmittance ($t$ = 0.75 unit-less), air pressure, and optical air mass, which are adopted from the study of Campbell and Norman [178]. As part of this study, the sun position throughout the day at different time intervals was estimated using the *Solar Position Algorithm* (SPA) [179] by the National Renewable Energy Laboratory (NREL), which offers accuracy up to $0.0003^{\circ}$. The terrestrial direct irradiance $I_{p_b}$ for each roof plane is estimated as:

$$I_{p_b} = I_b \times R_{p_b} \tag{7.4}$$

where $R_{p_b}$ is the correction factor for $I_b$, the ratio between the angles of incidence for the point's horizontal surface $\theta_{p_z}$ (i.e., the zenith angle) and its tilted surface $\theta_p$ [180], in order to compensate for $I_b$, which is only measured for horizontal surfaces, as illustrated in Figure 7.4.

$$
\begin{aligned}
R_{p_b} &= \frac{\cos(\theta_p)}{\cos(\theta_{p_z})} \\
\cos(\theta_{p_z}) &= \cos(\delta)\cos(\varphi_p)\cos(\omega) + \sin(\delta)\sin(\varphi_p) \\
\cos(\theta_p) &= \sin(\delta)\sin(\varphi_p)\cos(\beta_p) - \sin(\delta)\cos(\varphi_p)\sin(\beta_p) \times \cos(\gamma_p) \\
&\quad + \cos(\delta)\cos(\varphi_p)\cos(\beta_p)\cos(\omega) + \cos(\delta) \\
&\quad \times \sin(\varphi_p) sin(\beta_p)\cos(\gamma_p)\cos(\omega) \\
&\quad + \cos(\delta)\sin(\varphi_p) \times \sin(\gamma_p)\sin(\omega)
\end{aligned} \tag{7.5}
$$

where, $\varphi_p$ is latitude of a plane's location (i.e., centre of the roof plane), $\beta_p$ is slope of the plane, and $\gamma_p$ is aspect/orientation of the plane. The hour angle depends on the solar time (e.g., 10:00 is equal to $-30^{\circ}$, and 14:00 to $+30^{\circ}$). These equations give us the value for $R_{p_b}$ (a ratio between horizontal and tilted planes) but $I_b$ is unavailable. Therefore, direct irradiance for the horizontal and tilted planes for an hour is calculated as [178]:

$$
\begin{aligned}
& I_b = \cos(\theta_{p_z}) \times s_c \times t^m \\
& I_d = 0.3 \times (1.0 - t^m) \times s_c \times \cos(\theta_{p_z}) \\
& \text{where airmass } m \text{ is estimated as:} \\
& m = \frac{(\text{air pressure})}{(101.3 \times \cos(\theta_{p_z}))}
\end{aligned} \tag{7.6}
$$

The terrestrial diffuse irradiance $I_{p_d}$ for a given location is:

$$
I_{p_d} = I_d \times R_{p_d} \tag{7.7}
$$

where, $R_{p_d}$ is the correction factor for $I_d$; only the slope angle $\beta_p$ is considered, and the diffuse irradiance is assumed to be isotropic [180].

$$
R_{p_d} = \cos^2 \frac{\beta_p}{2} \tag{7.8}
$$

The total irradiance $I$ on each planar surface is simply the addition of both these irradiance components i.e.,

$$I = I_{p_b} + I_{p_d} \tag{7.9}$$

Most studies reported in the literature are optimised for the northern hemisphere for the estimation of solar potential. To optimise the solar potential formulas for the southern hemisphere, the equation to approximate the incident angle is corrected following Figure 7.5 [174]. The incidence angle of the sun with respect to the tilted surface, i.e., $\cos(\theta_p)$ is modified, such that the aspect/azimuth angle of the surface $\cos(\gamma_p)$ is replaced with $\cos(\gamma_p - 180^\circ)$.

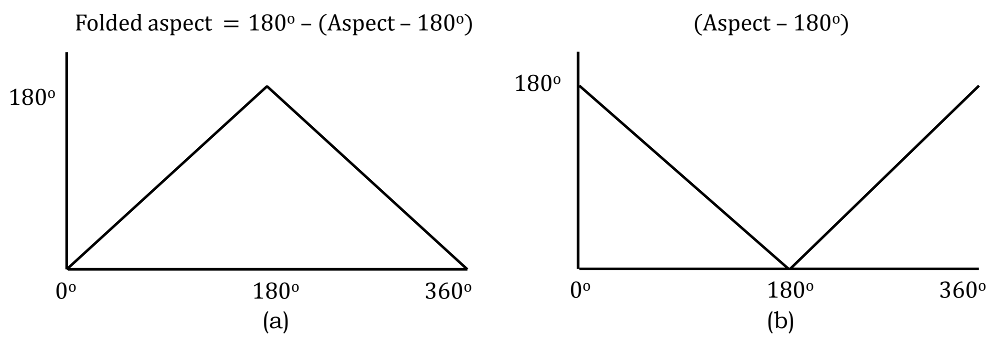



**Figure 7.5** Illustration of angle calculation for incident radiation due to folded aspect: (a) Northern hemisphere and (b) Southern hemisphere.

### 7.2.3 Solar Potential Demonstration

The test area used to demonstrate the solar potential is situated in Aitkenvale, Queensland, Australia. It is located at latitude -19.30052680 and longitude: 146.7656440. This area has 5 non-occluded buildings with 22 roof planes and a LiDAR density of 35 points/m$^2$. Figure 7.6 shows the test area with marked roof planes. The information on each plane $p$ including its boundary, centre, normal ($\vec{n_p}$), and slope ($\beta_p$) were supplied by our roof plane extraction technique [164]. However, the aspect $\gamma_p$ for each roof plane is measured as the angle between the plane's normal vector $\vec{n_p}$ on the horizontal plane and the vector to the geographical north.

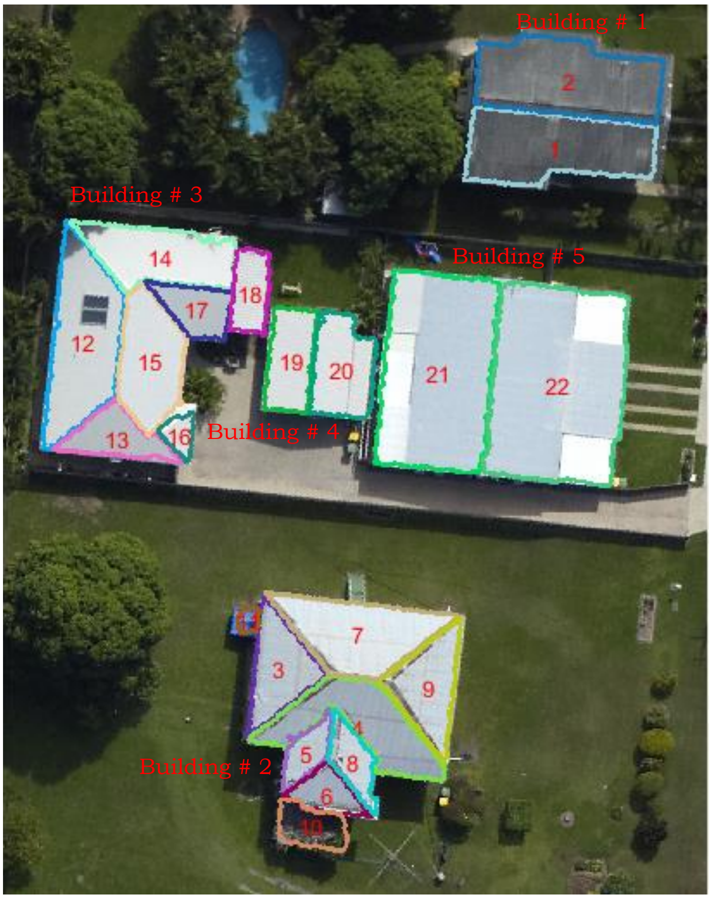



**Figure 7.6** Test area for demonstration of solar potential assessment.

The total solar potential ($I$) per day is calculated between 7:00 am – 05:00 pm with 1 hour time difference. $I_{p_b}$ and $I_{p_d}$ for each plane on the described time intervals are computed per square-metre area and per plane as well for all the buildings. We used different graphs to represent the estimated solar potential without considering shadows to determine the applicability and optimisation of the formulas for the Australian region.

#### 7.2.3.1 Annual Insolation

Figure 7.7(a) shows the total irradiance for all five buildings with their total area and number of planes in the following graph (in watts). For example, Building # 3 has 7 roof planes with an area of 193.41 m$^2$ and produces nearly $5 \times 10^8$ watts insolation annually.

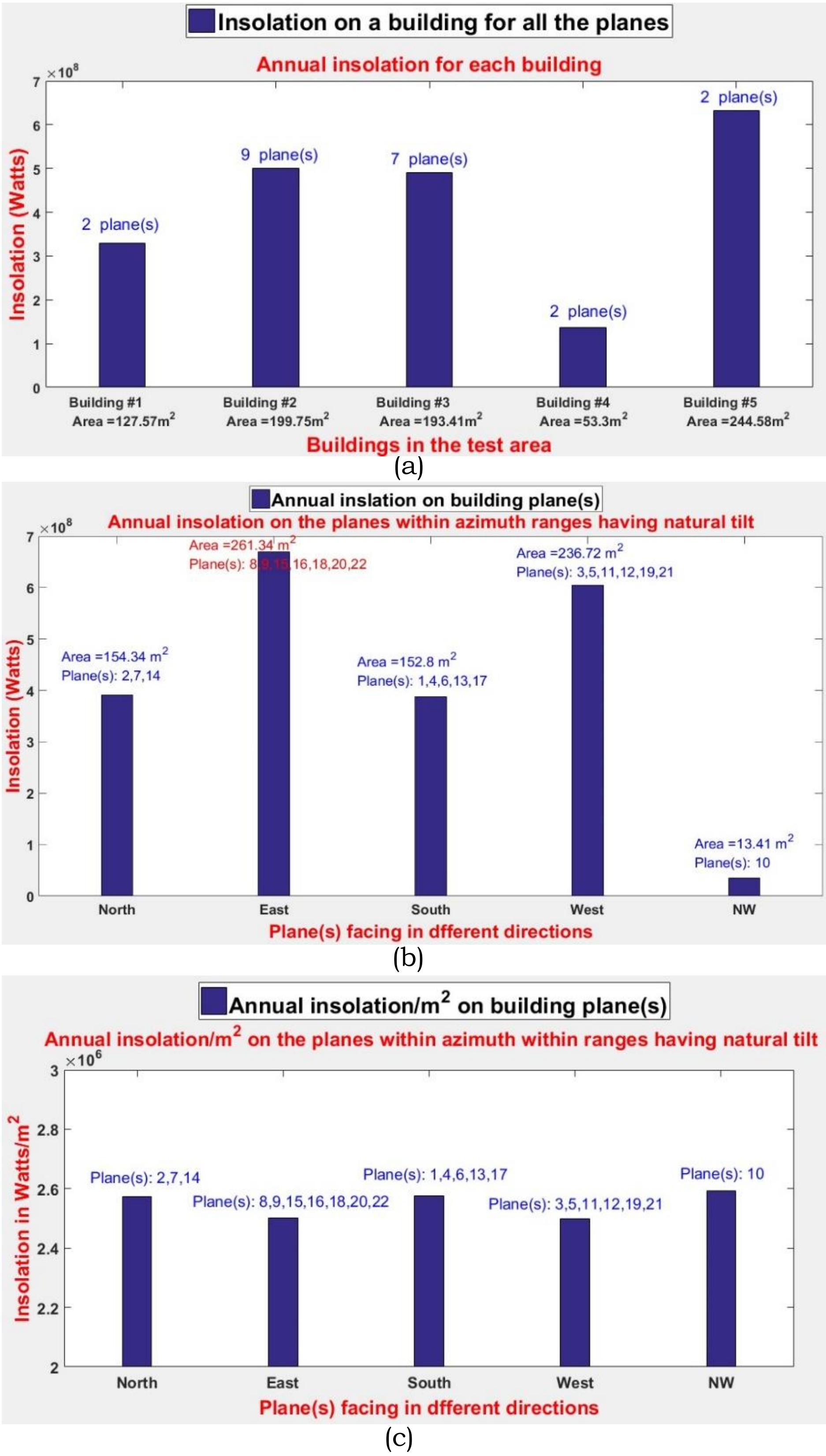



**Figure 7.7** (a) Annual Insolation per building and (b) Irradiance within different azimuth bins. (NW = north west)

#### 7.2.3.2 Irradiance within different Azimuth Bins

The roof planes of the current area face in different directions and have different orientations. The azimuth angles ($0^{\circ} - 360^{\circ}$) are separated into 16 bins, and each bin covers $22.5^{\circ}$ angles following Figure 7.3. Different azimuth bins/directions are shown on the $x$-axis while planes pointing in these directions are plotted on top of each bar. Figure 7.7(b) shows the annual solar irradiance on all the planes within a bin, whilst, for the estimation of insolation, the real slopes of all the planes are kept intact. Figure 7.7(c) shows the annual solar irradiance per $m^2$ of each plane in these bins, which indicates that the north-facing planes can produce more solar energy. For example, three planes (Planes 2, 7, and 14) face north and produce $4 \times 10^8$ watts insolation annually.

#### 7.2.3.3 Slope Variation Effect on Annual Solar Potential

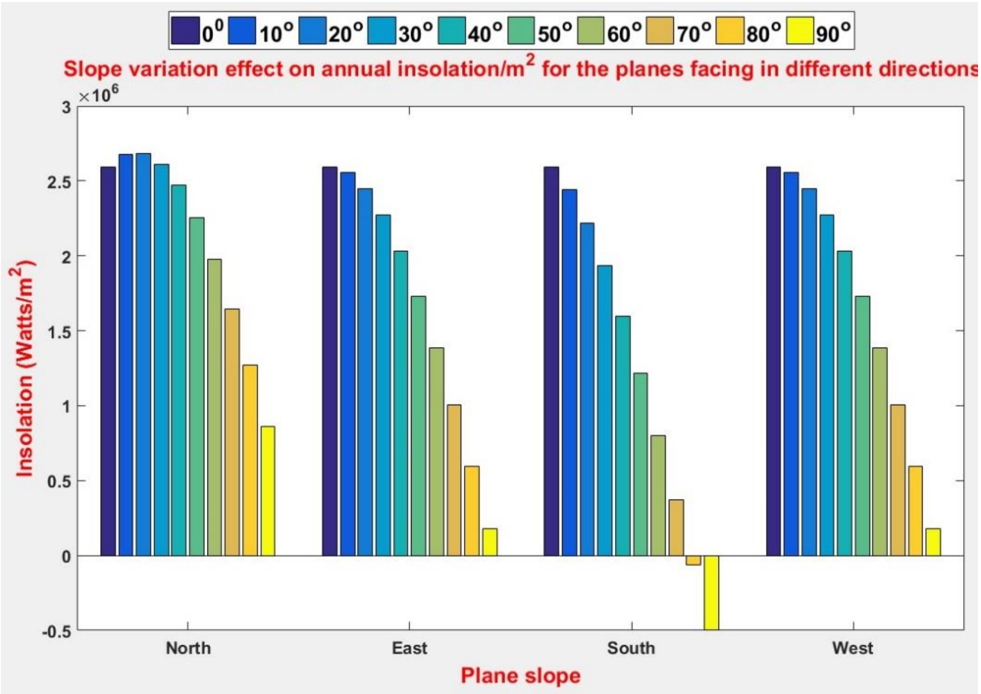
This is a simulation where the slopes of all extracted planes are changed between $0^{\circ}$ and $90^{\circ}$ and the insolation on all the planes is estimated. Figure 7.8(a) shows the effect of slope variation on annual insolation for all planes facing north, east, south, and west. For example, maximum insolation of around $2.7 \times 10^6$ watts/$m^2$ annually can be obtained from a north-facing plane with a slope of $20^{\circ}$.

#### 7.2.3.4 Effect of Orientation Variation on Annual Solar Potential

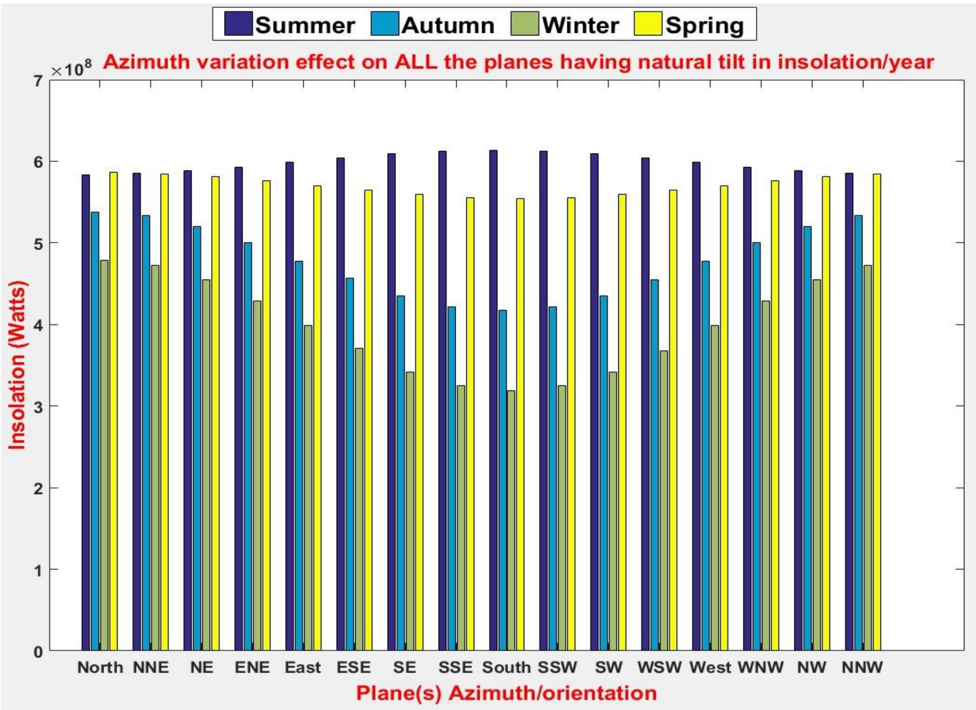
Note that the azimuth of a plane is calculated in clock-wise direction from true north, as shown previously in Figure 7.3. For this simulation, the azimuths of all the extracted planes are changed between $0^{\circ}$ – $360^{\circ}$ (North–East–West–South–North) at $22.5^{\circ}$ intervals and the annual solar insolation on all the planes is estimated while keeping their natural tilt unchanged, as shown in Figure 7.8(b). For example, the summer and winter bars in the north ($0^{\circ}$ from North) azimuth shows irradiances computed on all the planes (22) are around $5.9 \times 10^8$ and $4.6 \times 10^8$ watts for the entire year, respectively.

## 7.3 Shadowing and Shadow Path

The solar potential results presented in the previous section do not take any shadows into account. The purpose of the experiments was to estimate the solar potential accu-




(a)




(b)

**Figure 7.8** (a) Slope variation effect of planes on Irradiation and (b) Azimuth variation effect of planes on annual solar potential. N - North, S - South, E - East, and W - West.

rately and optimise the formulations for the Australian environment. The results were calculated using available real data and experts at the CRCSI verified that north-facing surfaces produce more solar potential than others in Australia.

Shadows continuously change throughout the day and year due to movement of the sun in the sky and this affects the performance of any PV system. A more realistic estimation of solar irradiance can be achieved only when shadowing is not ignored. A building's roof can be shadowed by surrounding buildings, vegetation, and other objects on the roof. In order to determine shadows quickly and efficiently, shadowing is proposed to be performed using a *sundial* and the *sun's position* in this research study. Since sunlight travels in straight lines, the projection of an obscuring point onto the ground (or any other surface) can be described in terms of simple geometry.

A good estimate of the sun's spherical position at a given time can be calculated using the SPA method [179]. It has an uncertainty of $\approx 0.0003^{\circ}$, and is several times more precise than other solar positional algorithms. A sundial is constructed following the *Gnomon principles* for the estimation of shadow direction [181]. Consider the shadow of a gnomon point (the latitude of a given plane) caste on to a horizontal plane, as shown in Figure 7.9.

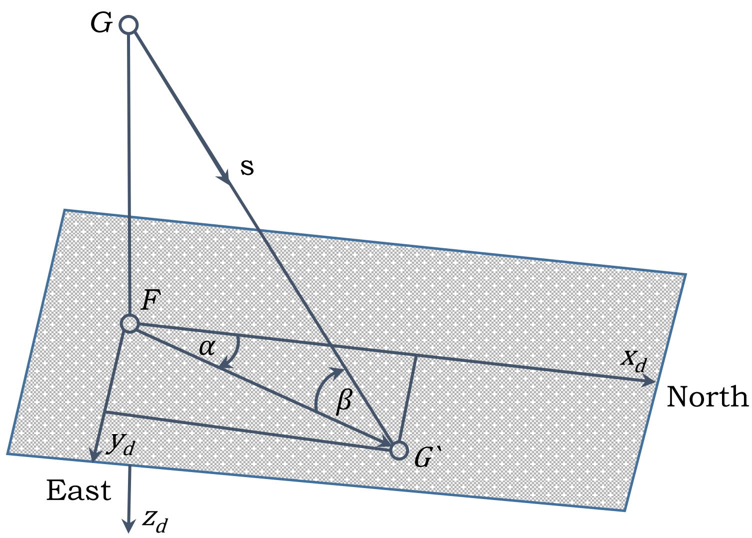



**Figure 7.9** Coordinate system of a sundial.

The direction of a gnomon point can be expressed as a vector $s$ in the following coordinate system. Its origin is at the foot $F$ of the perpendicular from the gnomon point $G$ onto the plane. The $x_d$-axis points to the north, the $y_d$-axis to the east and $z_d$-axis in the direction of the perpendicular. In this coordinate system, the vector $s$ is given as:

$$s_d = \begin{bmatrix} \cos(\alpha)\cos(z) \\ \cos(\alpha) sin(z) \\ \sin(\alpha) \end{bmatrix} . \tag{7.10}$$

where, $z$ and $\alpha$ denote azimuth and altitude, respectively. Further, the *shadow* $\acute{G}$ of the gnomon point on the plane has the coordinates:

$$g_1' = \frac{s_1}{s_3} g, \quad g_2' = \frac{s_2}{s_3} g, \quad g_3' = 0,$$

where, $g$ denotes the *length of the gnomon*, i.e., the distance $\overline{GF}$. Using the horizontal coordinate system, the vector $s$ can be expressed as:

$$s_h = \begin{bmatrix} \cos(\alpha)\cos(z) \\ \cos(\alpha) sin(z) \\ \sin(\alpha) \end{bmatrix} = -S_d \quad . \tag{7.11}$$

Consider now the task to determine the shadow of the gnomon point on the dial if the position of the sun is given in the equatorial coordinate system. The sun's position $v$ depending on the declination $\delta$ and the hour angle $\omega$ is:

$$V_q = \begin{bmatrix} \cos(\delta)\cos(\omega) \\ \cos(\delta) sin(\omega) \\ \sin(\delta) \end{bmatrix} . \tag{7.12}$$

A plane rotation is necessary to transform equatorial coordinates into horizontal coordinates. We have

$$V_h = \begin{bmatrix} \cos(\rho) & 0 & -\sin(\rho) \\ 0 & 1 & 0 \\ \sin(\rho) & 0 & \cos(\rho) \end{bmatrix} \times V_q \,, \tag{7.13}$$

where, $\rho = 90^\circ - \varphi$ and $\varphi$ is the observer's latitude. Therefore, given the $\omega$ (hour angle) and $\delta$ (declination) of the sun, $\varphi$ (the observer's latitude) and $g$ (gnomon length), we can calculate the shadow throughout the day at different hour stamps. The calculation of *shadow points* $(G')$ represents a point-wise mapping of the upper hemisphere onto the horizontal plane. This mapping is called gnomonical projection.

### 7.3.1 Shadow Length

Shadow length measurement gives the shadow cast of a vertical object according to its geographic location and the position of the sun using simple trigonometry. The length of the shadow map is normalised (changed with the magnification/zoom) and the direction is the opposite azimuth. The measurement of the length of the shadow depends on the height of the obstacle and the elevation of the sun. The following Figure 7.10 describes the shadow length of the Faisalabad clock-tower[2] and the formula to calculate shadow in Equation 7.14 is:

$$L = \frac{h}{\tan(\alpha)} \tag{7.14}$$

where, $L$ corresponds to shadow length, $h$ is the object height, and $\alpha$ denotes the angle between the sun and the horizon (sun altitude). We can use gnomon principles to not only estimate the direction but also the length of the shadow. In this regard, we need to determine the length of the gnomon's shadow at solar noon (12:00 pm) on the summer solstice when the sun is at its maximum altitude, knowing that the shadow cast by the gnomon of a horizontal sundial is the shortest. At any other time and date the shadow will be longer.

[2] One of the oldest monuments in Pakistan still standing in its original state from the period of British rule over the subcontinent. https://en.wikipedia.org/wiki/Clock_Tower,_Faisalabad

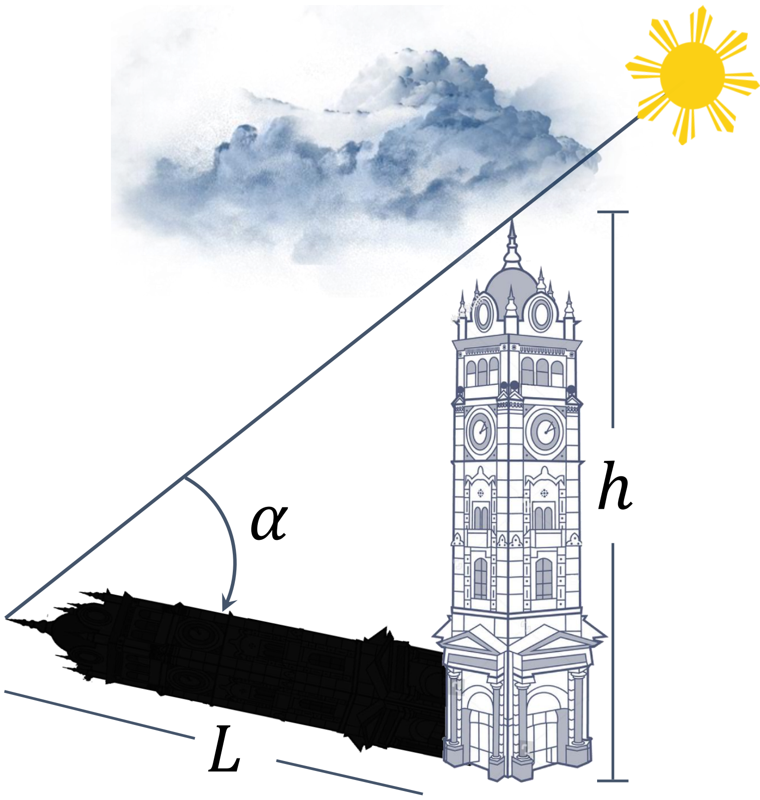



**Figure 7.10** Description of shadow length measurement for clock tower situated in Faisalabad, Pakistan.

The shadow length formula is slightly changed from the object's height with the height of the gnomon i.e., 1 metre. Later, the height of the object is multiplied by this unit height to obtain the exact shadow length.

$$L = \frac{\text{Gnomon height}}{\tan(\alpha)} \times \text{object height} \tag{7.15}$$

### 7.3.2 Shadowing Path and Demonstration of Shadow Estimation

Shadow direction and shadow length are estimated using the gnomon principle and the accuracy is verified against the available online sun positioning and shadowing tools [182–184]. The shadow of each building's plane and its length are estimated throughout the entire year and shadowing regions for each building are approximated.

Figure 7.11 shows the direction of the shadows between 7:00 am – 6:00 pm on February 23, 2016 for an observer located at latitude -19.3002 at 1 m height. Its shadow lengths are shown in grey lines and time on top of the shadow location whereas the direction of the shadow is from the observer's location to the shadow points (blue solid circles on the curvy line). The curve shows the shadow path of the object over the entire day.

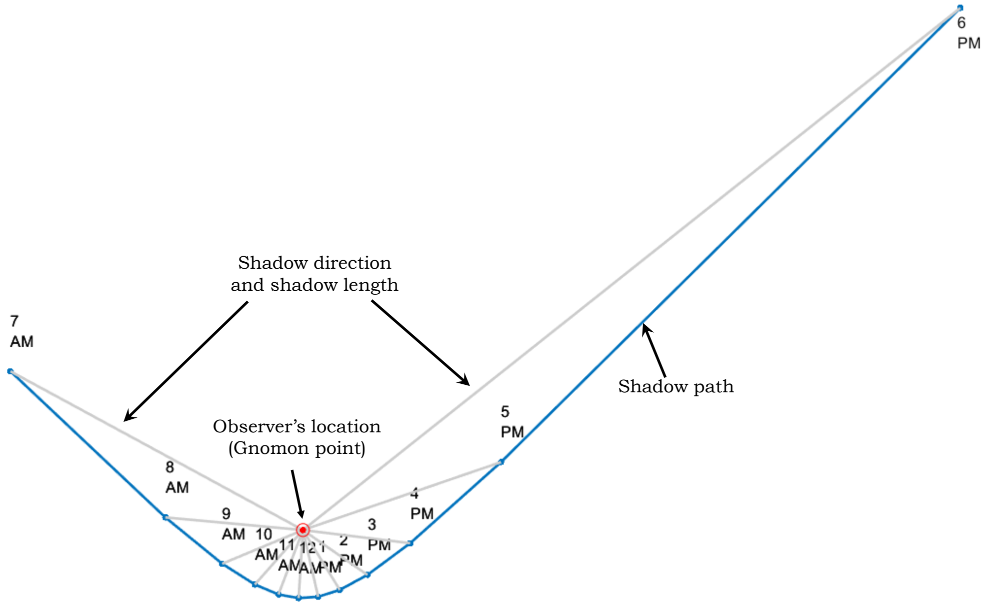



**Figure 7.11** Shadow path and its length during the day.

Throughout the year, the shadow path changes due to the movement of the sun in the sky. The sun's position is determined using the SPA method at different intervals of the day. It can be clearly seen that at solar noon on the day of the summer solstice the shadow cast by the gnomon of a horizontal sundial is the shortest. For the same location, we took a day in a month and plot the shadow path between 8:00 am – 4:00 pm, as shown in the following Figure 7.12.

Following the gnomon principles, we estimate the shadow direction and the shadow length for each plane of the building. In Figure 7.13, the building roof plane is highlighted in cyan while the estimated shadow path is shown in blue line with blue dots showing the time intervals labelled with the time of the day. Each line (grey) between the gnomon location (centre of the plane) towards each blue dot shows the length of the shadow at that particular time. This shadow length indicates the shadow of the plane on the flat surface (ground) and not on any elevated objects in the surroundings.

Since we are using the centre of the plane as the gnomon's location, a plane's

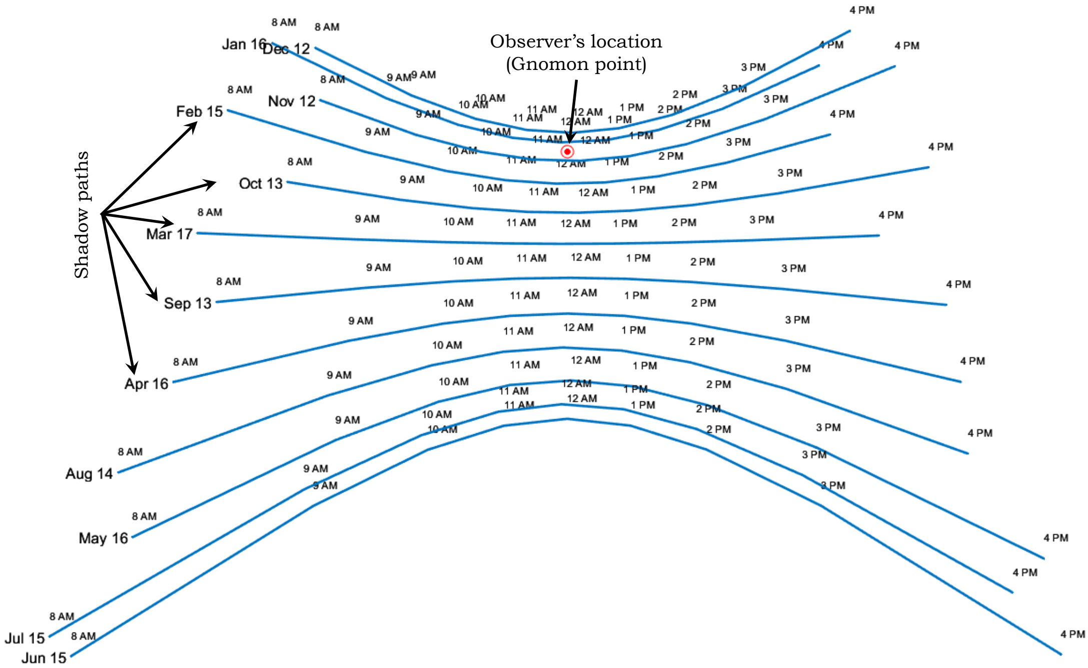



**Figure 7.12** Shadow growth throughout the year.

geometry is replicated on the shadowed position i.e., between 8:00 am – 4:00 pm for each day. Figure 7.14 shows placements of the plane's boundary at different shadow locations throughout the entire day between 8:00 am – 4:00 pm on March 18, 2016.

To find a building object under a shadow cast by a building plane, its shadowing region needs to be determined. Figure 7.15 shows the shadow location for Plane 1 (green in colour) at different times of the day while the magenta colour boundary shows the shadowing region of a plane for the whole day. In other words, the shadow cast by Plane 1 remains within the region bounded by the magenta colour boundary. Similarly, we repeat the same process for all other planes and approximate the shadowed region on March 18, 2016 between 8:00 am - 4:00 pm, as shown in Figure 7.16.

As shown in Figure 7.12, the shadow of gnomon pin grows throughout the year due to the varying position of the sun in the sky. The building planes which never appear under the shadows cast by another building throughout the whole year are the most suitable planes for harnessing solar potential. Therefore, it is necessary not only to determine the shadowing regions of a building plane in a day but also the shadow region for the entire year. Figure 7.17 shows the shadow path and shadow position of Plane 1 during the time intervals of a day each month. Rather than taking 365 days, a

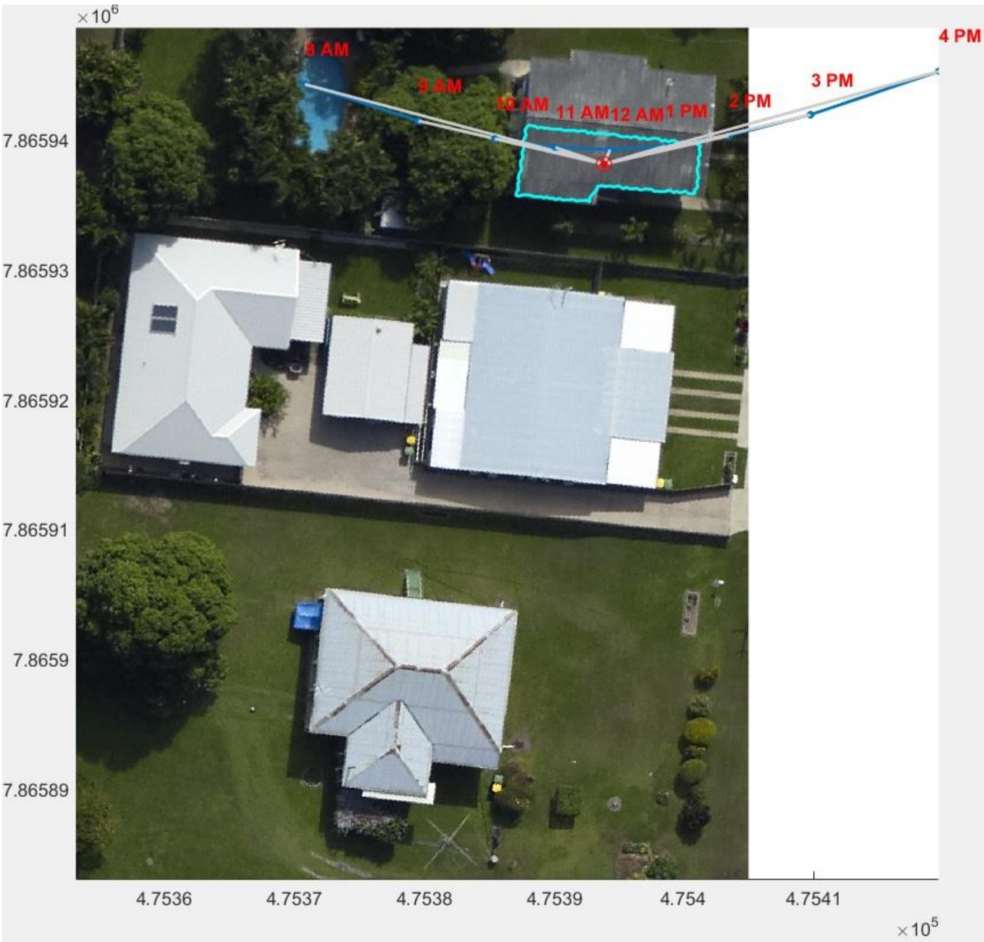



**Figure 7.13** Shadow path of Plane 1 on March 18, 2016 between 8:00 am – 4:00 pm.

single day of each month is considered at 30 days difference for the estimation of the shadowing region throughout the year. The entire process is quite fast and enables the rapid estimate of the shadow cast by all the planes of the buildings present in the given dataset.

For all the shadowing measurements, the gnomon is positioned at the centre of the plane and kept stationary for the entire year. We can see different shadow paths, based on different positions of the sun in the sky. Each red dot in Figure 7.17 indicates the time interval, the blue line represents the shadow path and the green shows the plane stacked at the position of the shadow, i.e., the red dot.

Figure 7.18 shows the shadowing region of Plane 1 over the entire year. In other words, it shows that the shadow of Plane 1 will remain within the region sketched in

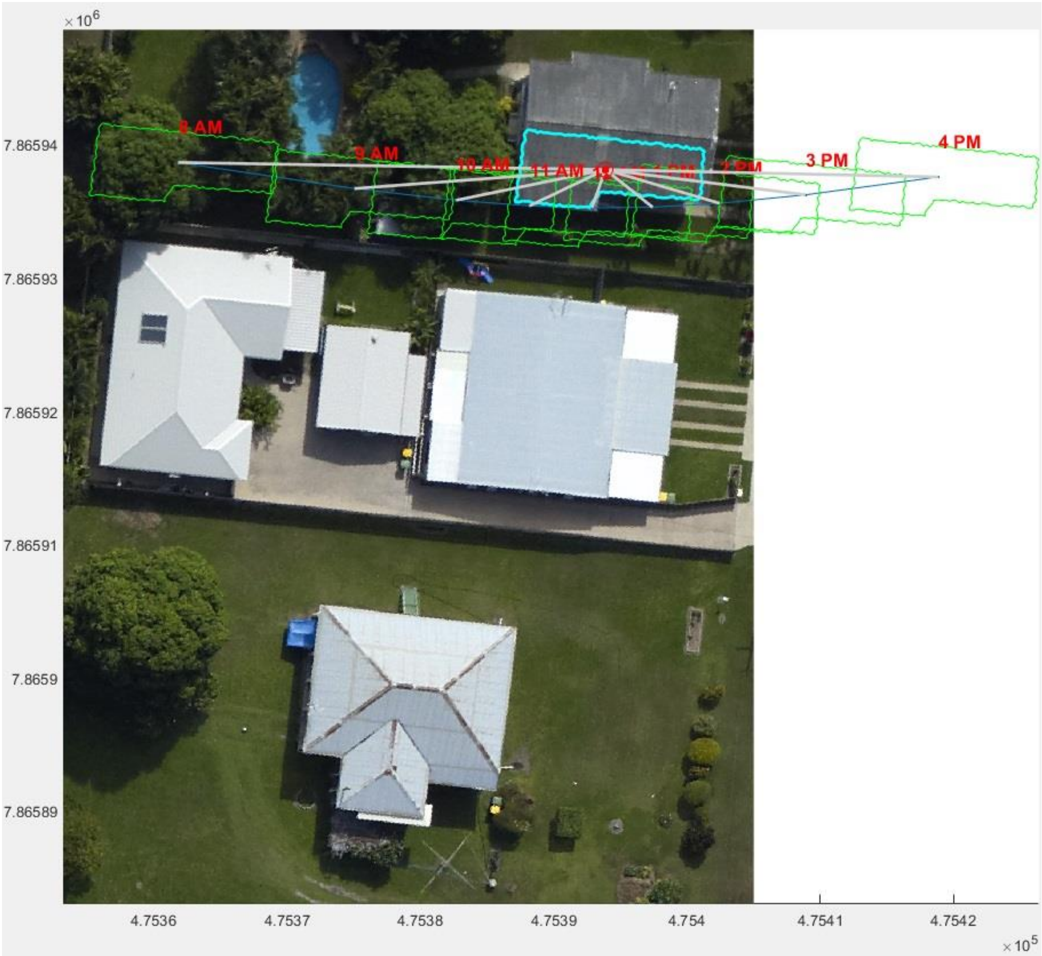



**Figure 7.14** Shadow of Plane 1 on flat ground during the entire day along the shadow path.

magenta colour during any time the year. It should be noted that this shadow region is on the flat surface. We repeated the above process and estimated the shadowing region for each building plane over the whole year. The shadowing region for a day in each month is accumulated to approximate the shadowing region for the whole year. Figure 7.19 shows the annual shadowing region of each plane sketched using different colours.

Next, the building planes under the shadowing region are determined and marked as unsuitable for production of solar potential. We, therefore, utilise the altitude difference and the distance of the planes from the source plane's shadowing region to decide a plane's suitability for generation of solar potential. This is an iterative process in which all the planes under the shadowing region of another plane are determined


**Figure 7.15** Shadowed region of FULL day for Plane 1 on March 18, 2016.

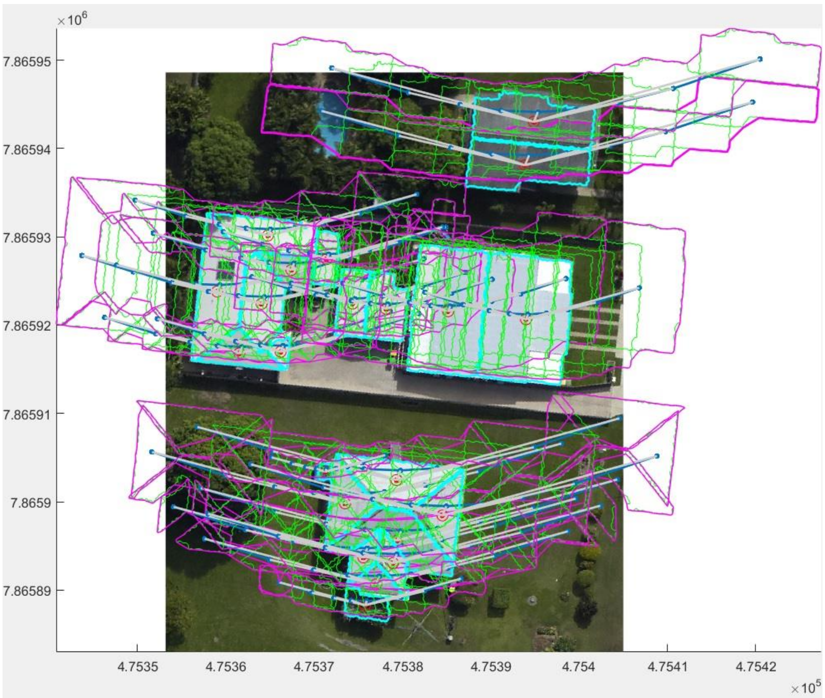



**Figure 7.16** Shadowed region of FULL day for all planes on March 18, 2016.

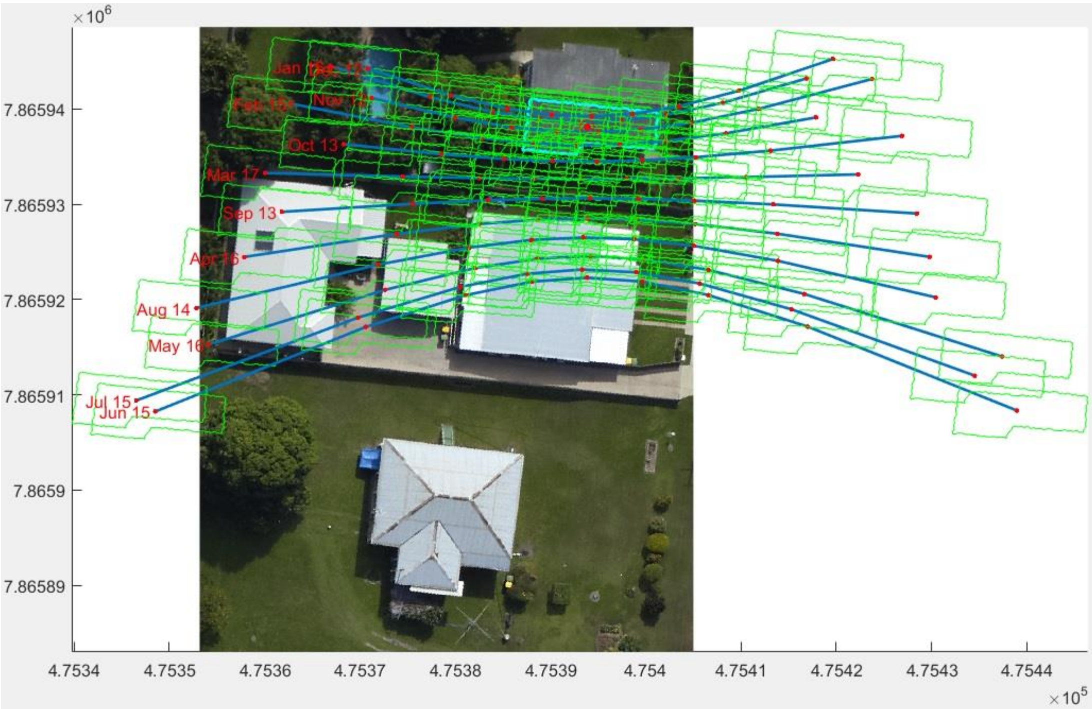



**Figure 7.17** Shadow change effect for Plane 1 throughout the year. The difference between any two successive shadowing time intervals is 30 days from January to December.

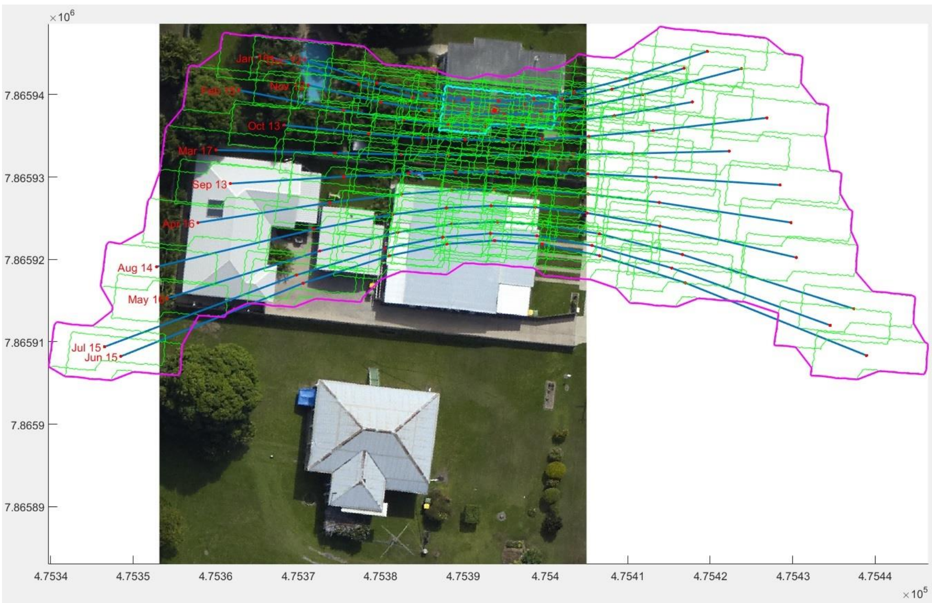



**Figure 7.18** Shadowing region estimation for Plane 1 on flat earth surface based on shadow paths between different days of entire year.

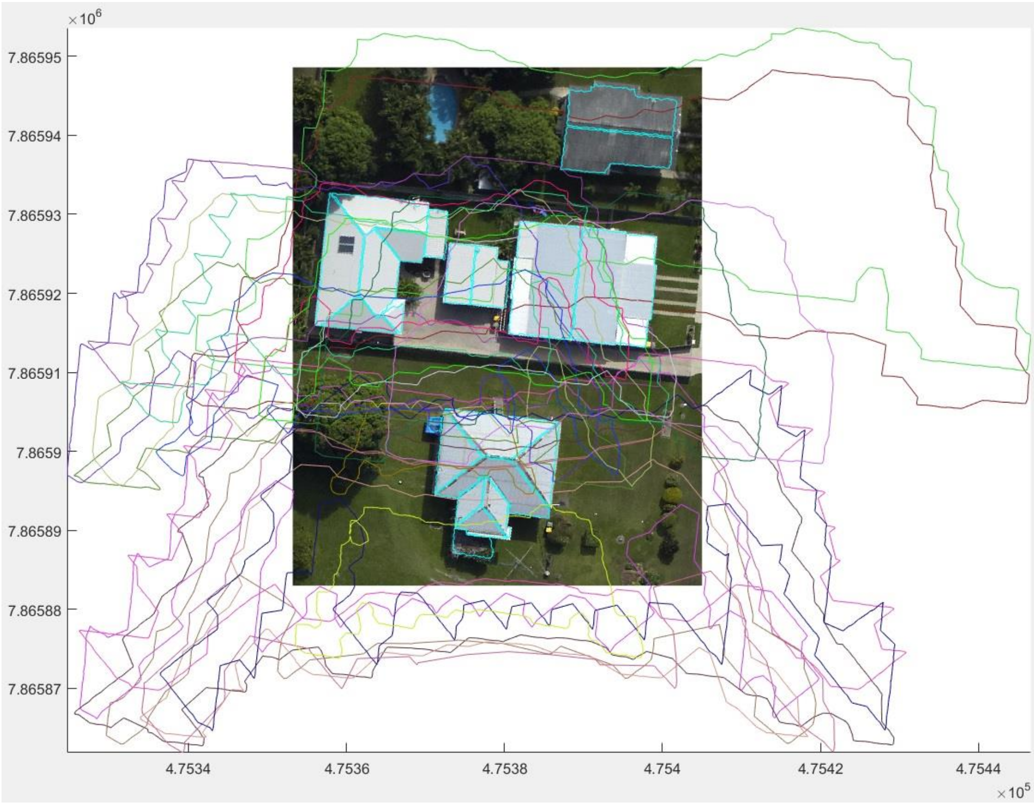



**Figure 7.19** The shadowing region for all the buildings' planes on flat ground for entire year.

using the defined criteria. Finally, the planes highlighted in magenta in Figure 7.20 show the planes which are shadowed due to self-occlusion (a roof under the shadow of another roof) and not suitable for harnessing solar energy, whereas the planes shown in cyan are unshadowed and suitable for the installation of solar systems.

The next section provides information on how to estimate the shadow cast by surrounding vegetation following the same principles and determine the building planes which are suitable for harnessing solar potential. This task requires the identification of vegetation regions and also approximates the tree canopies.

### 7.3.3 Tree Canopy and Shadowing Approximation

Tree canopy estimation is one of the most challenging activities, as trees grow leaves which drop in different seasons. The shadows cast by trees can be determined only if the tree canopy region can be identified and its corresponding boundary can be

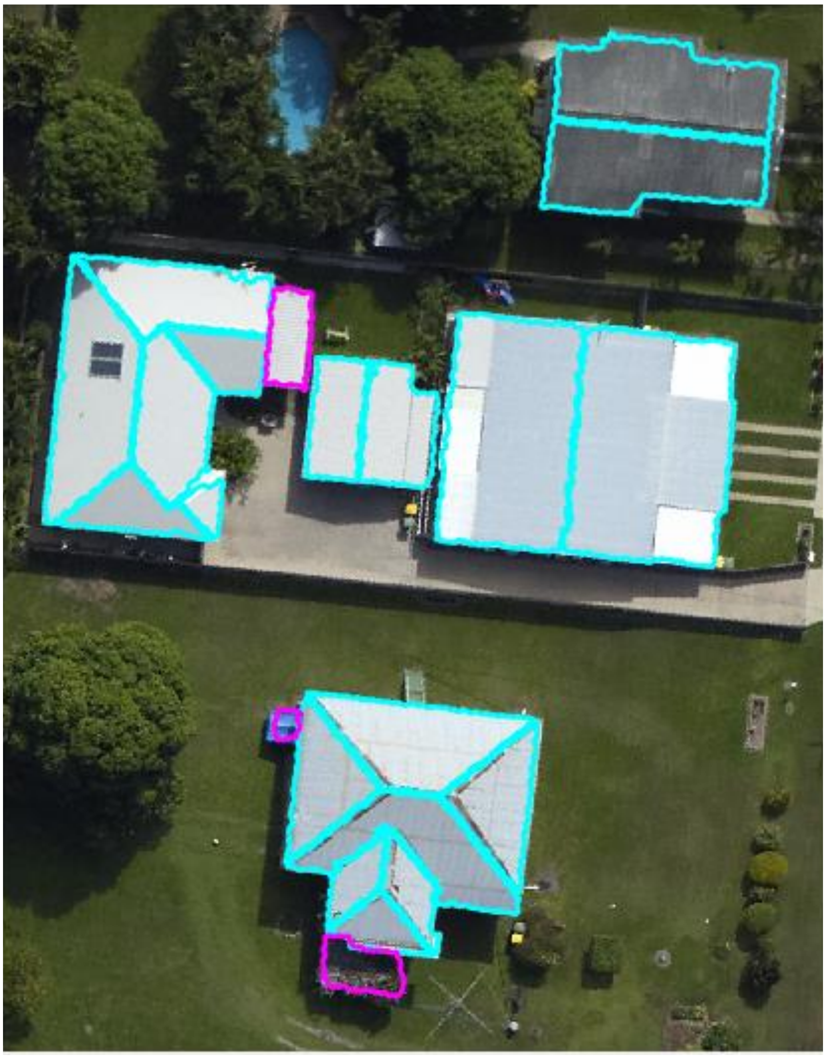


**Figure 7.20** Suitable and unsuitable planes for solar potential estimation (cyan colour shows suitable planes while magenta shows unsuitable planes due to shadowing effect).

estimated. Then, following the gnomon principles, the shadow path of the trees and the shadowing region can then be approximated. Concerning the solar potential estimation and the nature of trees, precise and accurate detection of tree regions is not considered mandatory in this study. Therefore, we use a height threshold window to segregate the tree LiDAR points on vegetation into different height segments.

We propose a region-growing technique that uses an adaptive height threshold (height window) and a fixed-neighbourhood approach for the estimation of tree canopies. The height threshold is user-defined which specifies the separation of the LiDAR points from the highest unsegmented point to the points not below the minimum limit of window size. Similarly, the LiDAR points which are left in the first iteration are considered in the following iterations, which are then aggregated in other layers at lower height levels. In our study, we chose a height threshold of 6 m. As a result, the LiDAR points on vegetation are segmented into multiple elevation groups.

Figure 7.21(a) shows the LiDAR points corresponding to the tree regions present at the test site. The proposed segmentation technique divides the entire point cloud into two separate vegetation layers (green and purple), as shown in Figure 7.21(b). Finally, contiguous LiDAR points are accumulated to identify the tree canopies. Next, the boundary tracing algorithm discussed in Section 5.2.1 is applied to the segmented data for the boundary approximation of the regions, in order to extract the individual tree regions. Figures 7.21(c) and 7.21(d) show the estimated boundaries of tree canopies extracted from the segmented LiDAR points presented in different colours.

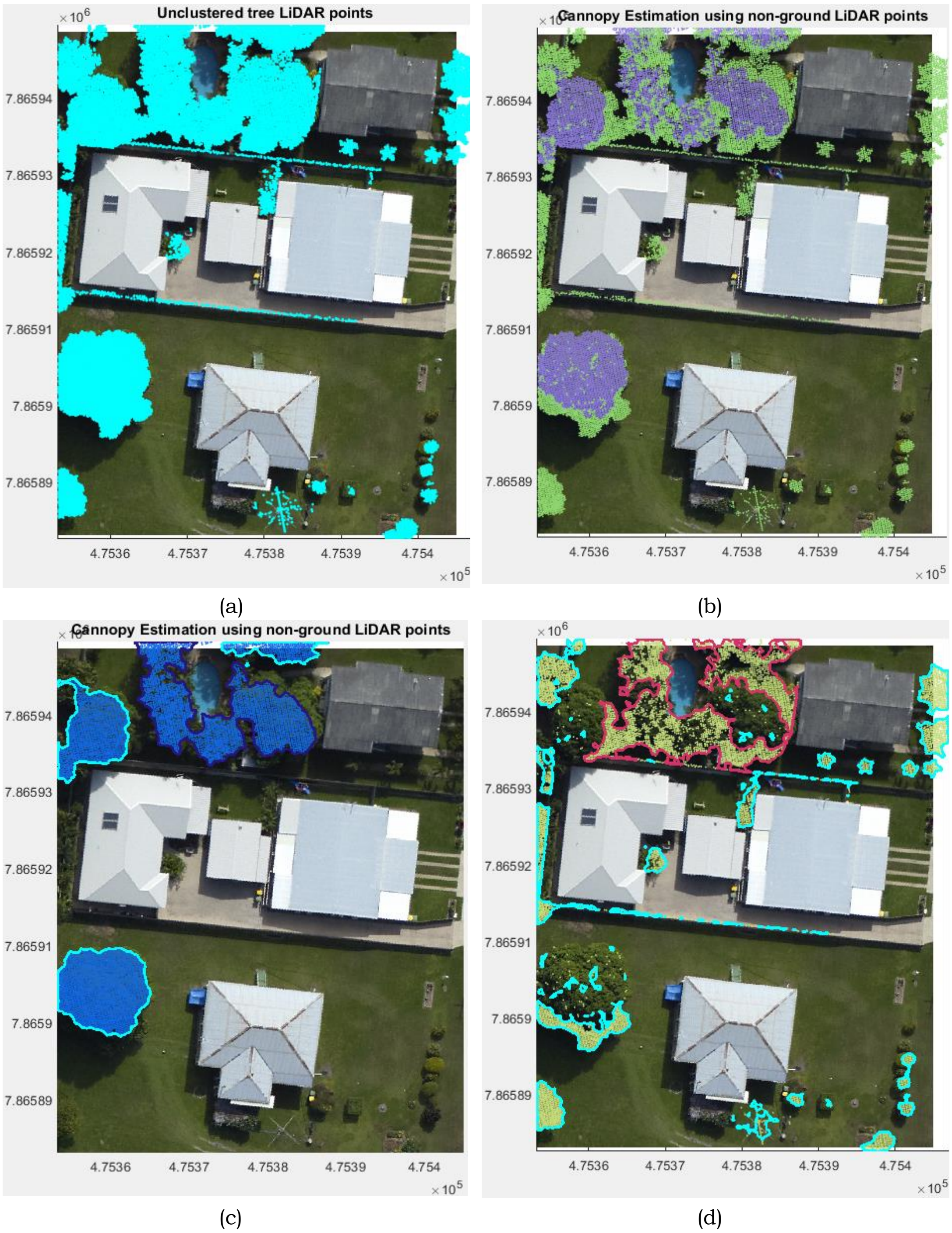
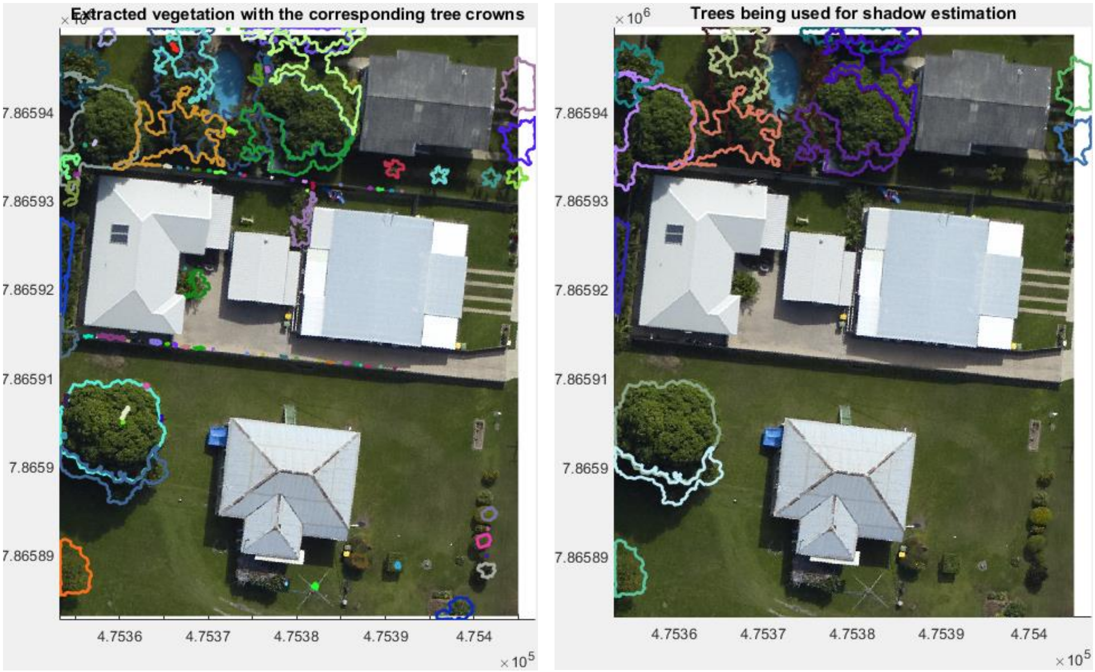



**Figure 7.21** Tree LiDAR points segmentation and canopy approximation: (a) Un-clustered tree LiDAR points; (b) Segmented LiDAR points based on height and local neighbourhood (two height groups are formed here); (c) Boundary approximation of tree canopies (using purple LiDAR points from first height window shown in sub-figure (b)); and (d) Boundary approximation of tree canopies (using light green LiDAR points of the following height window shown in sub-figure (b)).

Figure 7.22(a) shows the estimated boundaries for all the tree canopies extracted at different levels of height. Using the user-defined area threshold, which is 3 $m^2$ in our current study, all the superfluous trees with an area less than the threshold value are eliminated and the final results are shown in Figure 7.22(b).




(a) (b)

**Figure 7.22** Elimination of superfluous trees and estimation of tree boundaries: (a) All possible trees and (b) Trees after removal of small ones.

#### 7.3.3.1 Estimation of Shadowing Regions

Shadowing considerably reduces solar irradiance when building roof planes remain in shade during the day, especially at solar noon. Before shadow approximation, solar potential estimation can be performed, but it would not be a realistic estimation. Therefore, shadows cast by trees and buildings (self-occlusion) are estimated during the daytime only, since there is no solar irradiance to be considered at night. We use the gnomon principles and the sun position algorithm to determine the direction of the sun and the length of the shadow with respect to the sun's position in the sky. The technical details of the gnomon principles and the estimation of shadow have been covered in previous sections. Previously, we chose a day in each month to find the position of the sun at different time intervals and estimate the shadowing region of a building roof for the entire year. Here instead, we divide a year into four weather seasons: Summer (Dec – Feb), Autumn (Mar – May), Winter (Jun – Aug), and Spring (Sep – Nov) for the estimation of shadowing regions of vegetation and buildings. We consider a single day from each week (approximately four days a month) and determine the sun position from 10:00 am to 3:00 pm, since most energy in a day is produced around solar noon (12:00 pm). Figure 7.23 shows the shadowing region of a tree canopy. The boundary of the tree canopy is sketched in cyan colour and the shadowing regions for the Summer, Autumn, Winter, and Spring seasons are plotted in red, magenta, blue, and yellow, respectively.

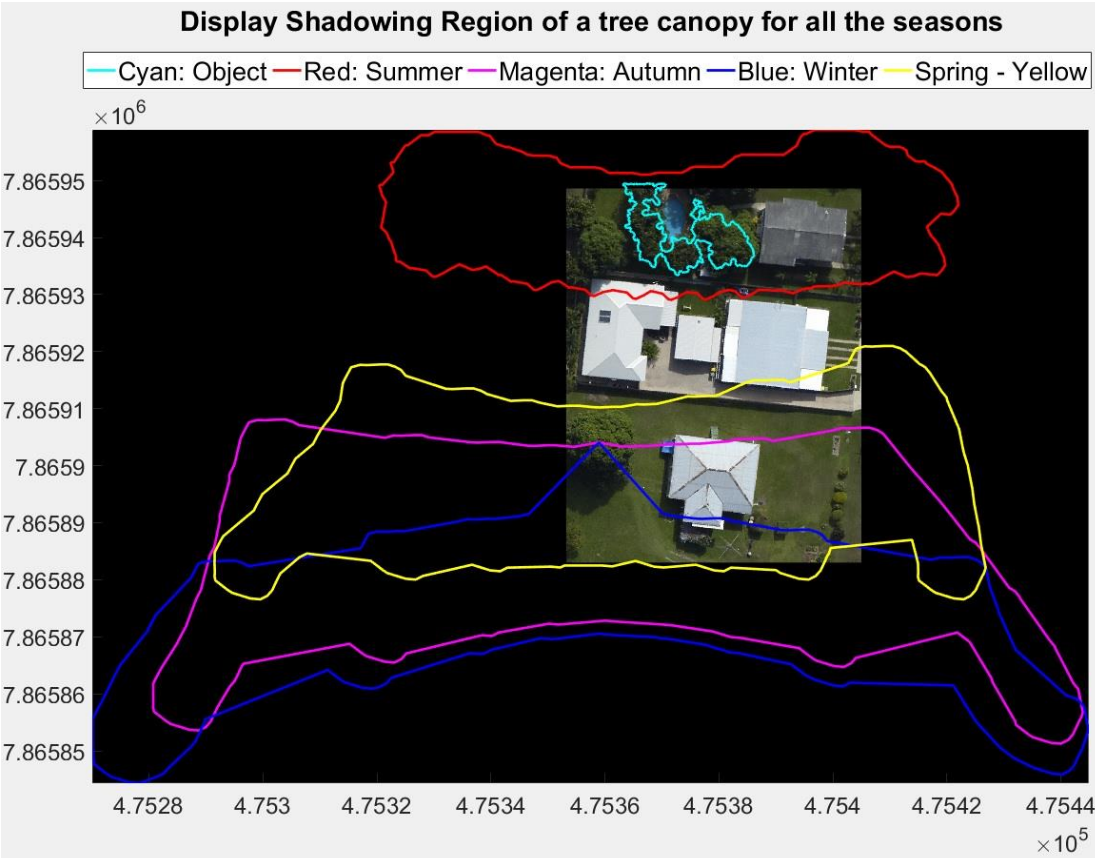



**Figure 7.23** Shadowing region of tree canopy in different seasons.

Similarly, using the gnomon principle, we estimate the shadowing regions of a building plane in four seasons, as shown in Figure 7.24. The building roof plane is shown in cyan colour whereas shadowing regions in Summer, Autumn, Winter, and

Spring are shown in red, magenta, blue, and yellow, respectively.

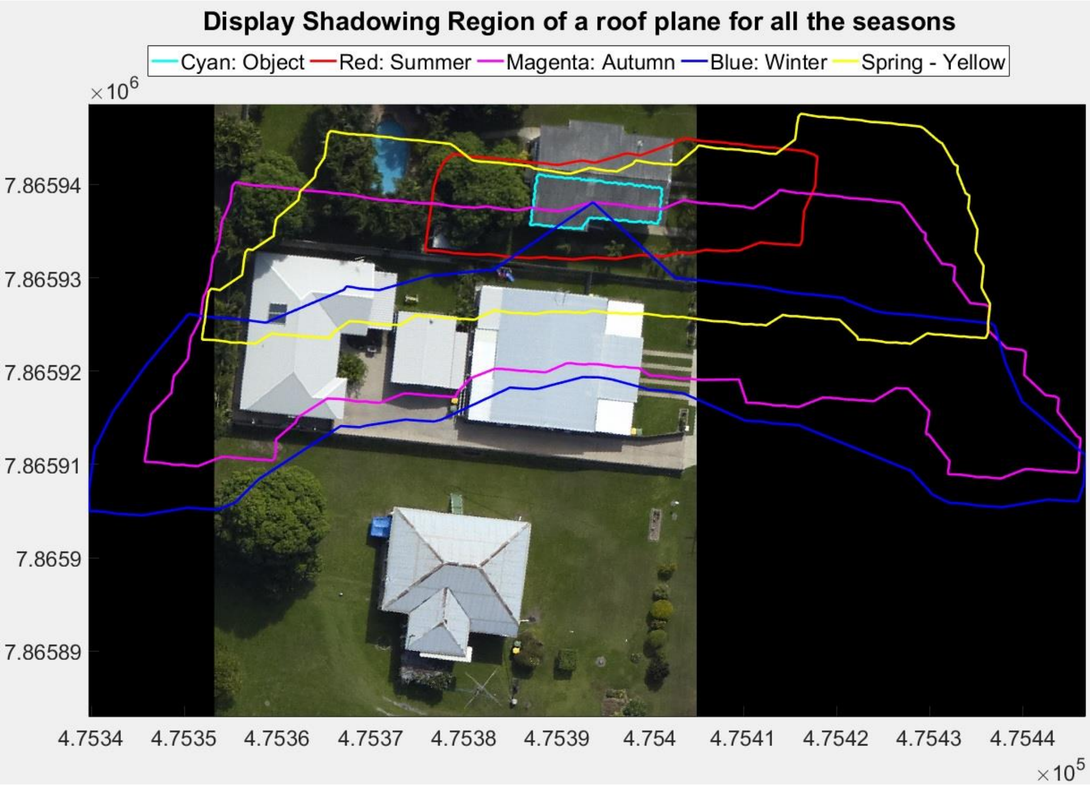



**Figure 7.24** Shadowing region of a building roof in different seasons.

#### 7.3.3.2 Estimation of Shadowed Roofs by Trees

In order to determine suitable planes for the production of solar potential, we iteratively consider the shadowing region of a tree canopy in each season and determine whether a building roof falls under the shadow. Figure 7.25 demonstrates the procedure graphically. We use a tree canopy in a particular season and determine whether the area (60% in our case) of a neighbouring building plane is covered by its shadow. The height difference between the tree and the neighbouring building plane and their mutual Euclidian distance are also considered to decide a shadowed plane. The tree canopy is plotted in red broken lines with its shadowing regions in red solid lines whereas the building roof planes are shown in cyan.

This is an iterative procedure that identifies the building roof planes, of which 60% of the area are covered by shadows in a particular season. The Figures 7.26(a)–(d)

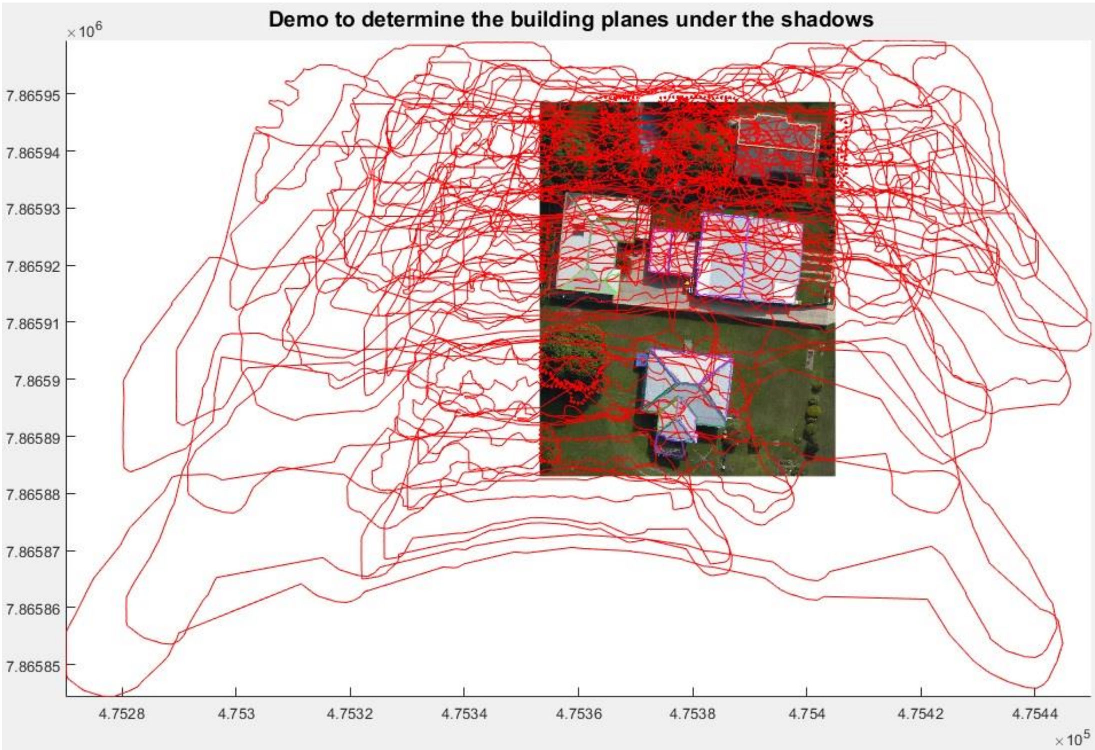



**Figure 7.25** Graphical illustration showing shadowing roof planes because of trees.

show the building planes suitable for the production of solar energy and the shadowed planes in different seasons. The Summer, Autumn, Winter, and Spring shadowed planes due to trees are shown in red, magenta, blue, and yellow, respectively. The roof planes suitable for solar potential are plotted in cyan.

#### 7.3.3.3 Estimation of Self-shadowing Roofs

We repeated the same iterative procedure explained in the last section to determine the self-shadowing planes, i.e., a building roof under a shadow cast by another neighbouring roof. The aim is to identify a building roof plane of which 60% of the area is covered by the shadow of a neighbouring building in each season. Figures 7.27(a)-(d) show the building planes suitable for the production of solar energy and the shadowed planes due to nearby buildings in each season. For Summer, Autumn, Winter, and Spring, the shadowed planes are shown in red, magenta, blue, and yellow, respectively. The roof planes suitable for solar potential are plotted in cyan.

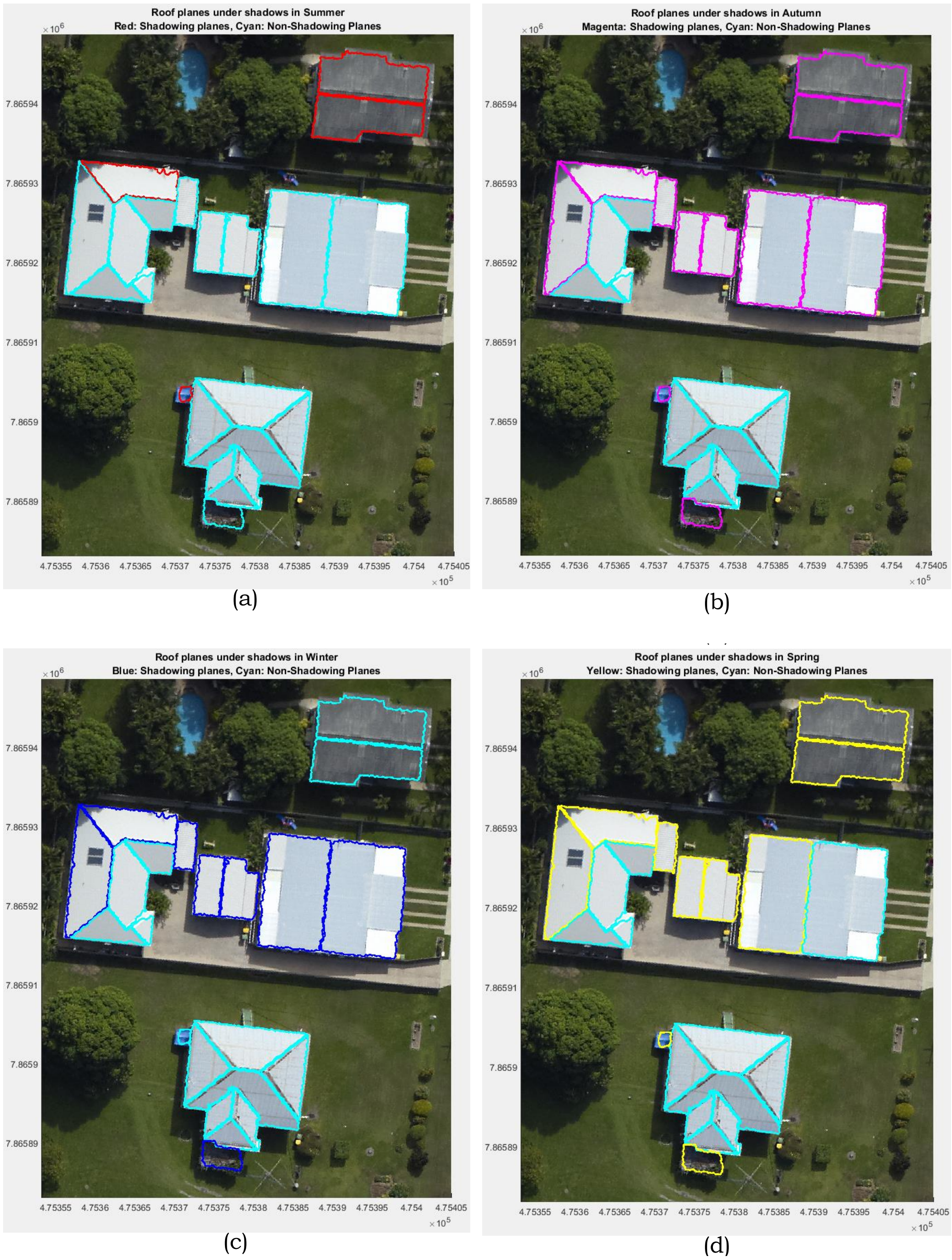



**Figure 7.26** Shadowed building roofs in different seasons due to surrounding trees: (a) Summer; (b) Autumn; (c) Winter; and (d) Spring.

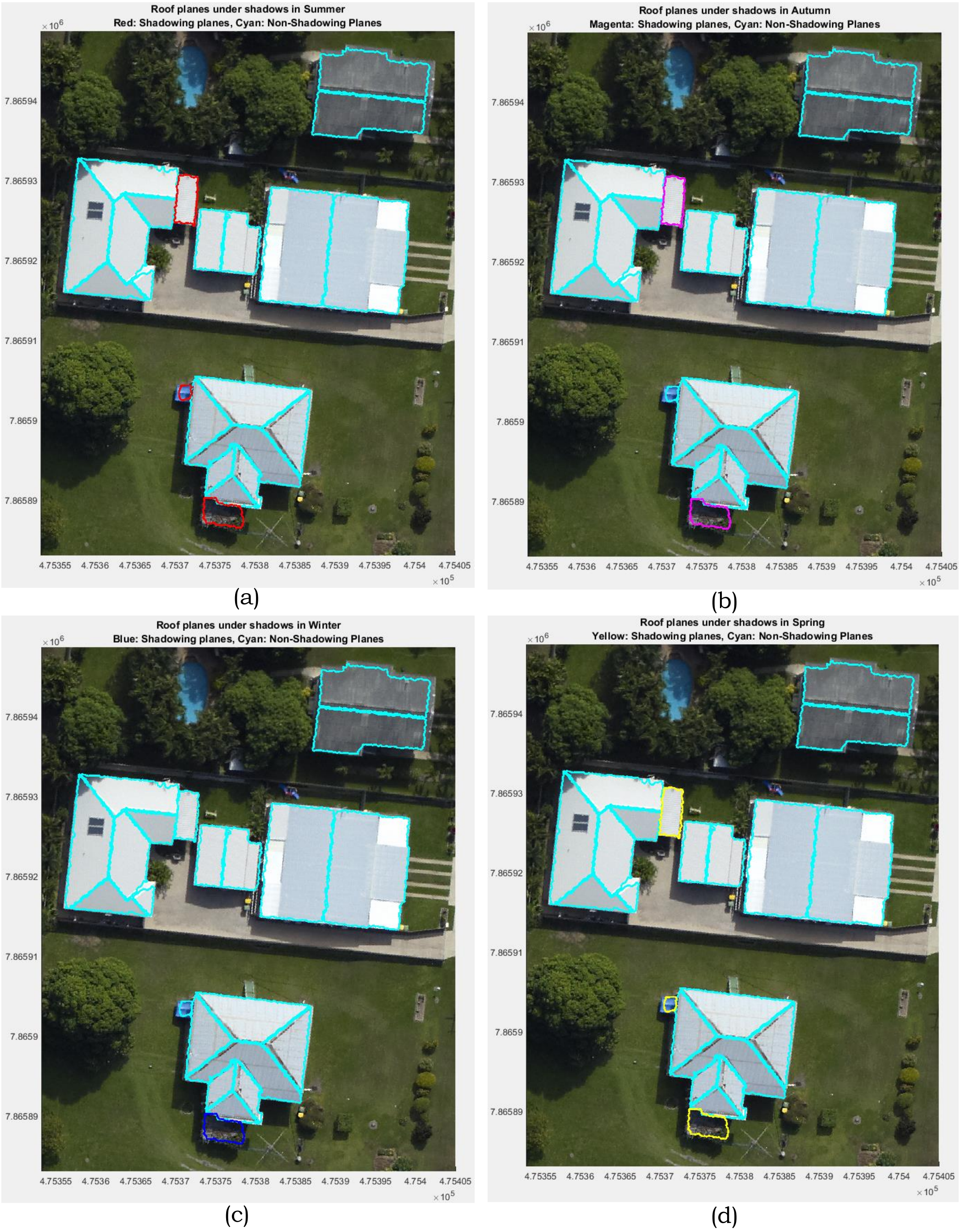
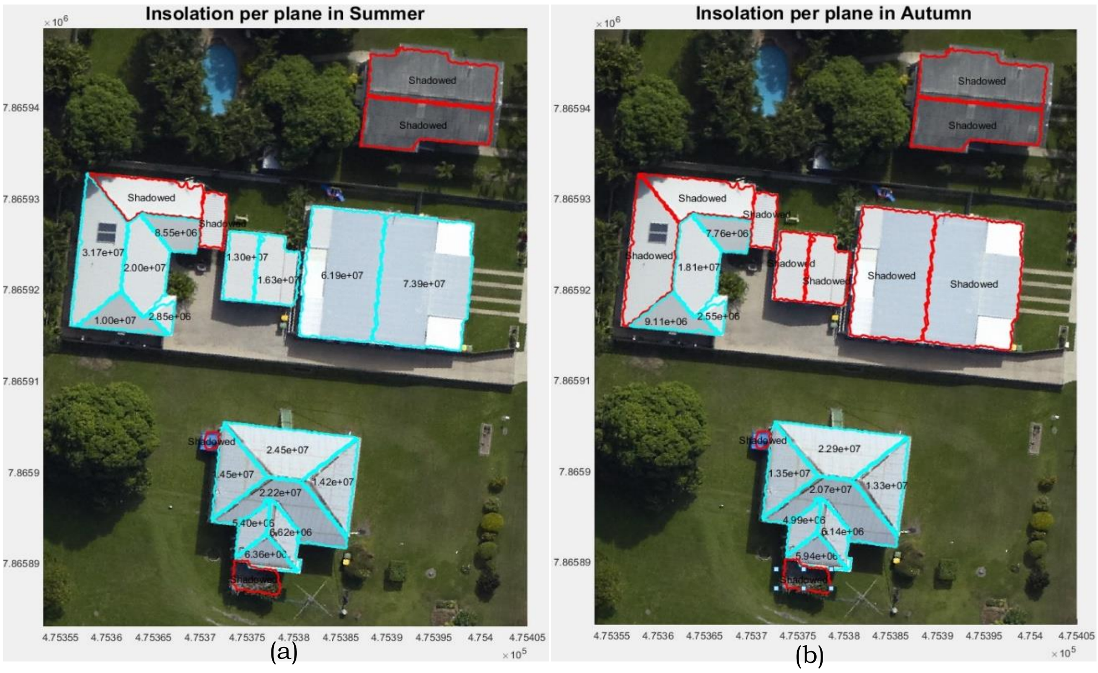



**Figure 7.27** Shadowed building roofs due to neighbouring tall roof planes (self-occlusion) in different seasons: (a) Summer; (b) Autumn; (c) Winter; and (d) Spring.

## 7.4 Solar Potential Estimation on non-shadowed Roofs

The previous sections described various forms of radiation, estimation of sun angles, and how solar potential is estimated on a particular roof plane. The solar radiation striking building roofs and other features is generally affected by a number of mechanisms. Part of the incident energy is scattered and absorbed by air molecules, clouds and other particles in the atmosphere. Part of the radiation that is not reflected or scattered reaches the surface directly, which is direct irradiation. The other type of radiation, which reaches the ground is referred to as diffuse irradiation. Some radiations reach the receiver after being reflected from the ground, which is known as reflected irradiation. Therefore, the total irradiation a roof plane can have is the sum of these three components. In this study, we do not consider reflected irradiance, because its impact on urban areas is negligible.

Figures 7.28(a)–(d) and the numerical results in Figure 7.29 show the approximate solar potential in watts for each building plane in different seasons. The roof planes which are marked as shadowed due to trees and/or neighbouring tall roofs are combined to represent roof surfaces under shadow in a particular season. The building roof surfaces which remain under shadow are shown in red. However, roof surfaces feasible for the estimation of solar potential are plotted in cyan with their respective seasonal solar potential in watts.



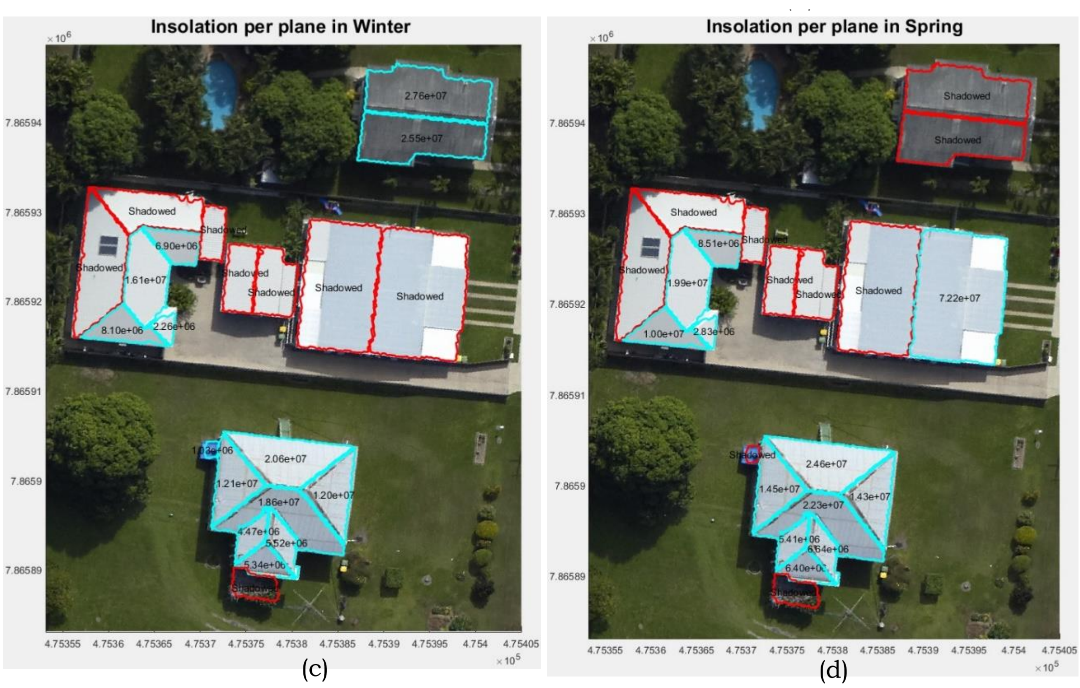
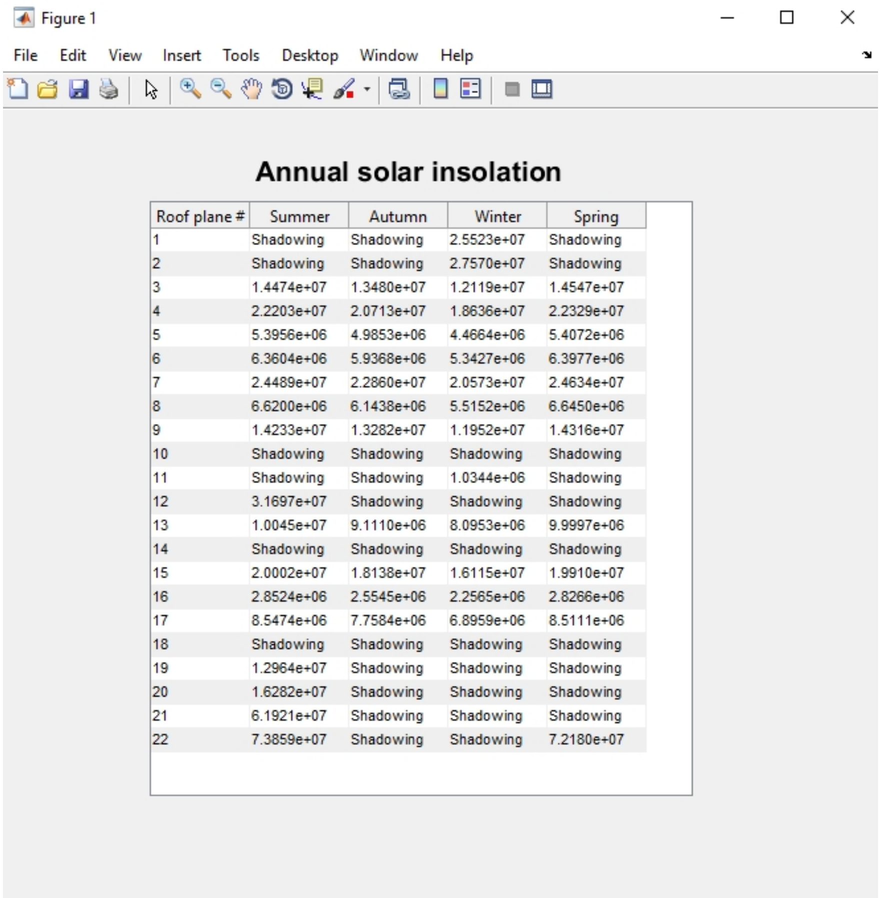



**Figure 7.28** Unshadowed roofs and total solar potential in different seasons: (a) Summer; (b) Autumn; (c) Winter; and (d) Spring.

## 7.5 Summary

This chapter discusses a practical application of the research work carried out in the thesis. The assessment of the solar potential project covered in this chapter was supported by the CRCSI, Australia, in an effort to support both local and state governments and solar energy companies. We studied several topics including a study of the automatic estimation of building roof parameters, a mathematical formulation for the estimation of shadows and their effect on solar potential estimation, different types of radiation, the sun-earth relationship, and finally, the estimation of annual solar potential for individual roof planes and buildings. This work is planned to be extended to identify the unshadowed parts of roof planes which can be used for the generation of solar energy. In addition, the creation of robust displays showing thermal affects is under consideration, to better represent the potential of solar energy for the entire area apart from the numerical measurement of energy.




Annual solar insolation

| Roof plane # | Summer | Autumn | Winter | Spring |
|---|---|---|---|---|
| 1 | Shadowing | Shadowing | 2.5523e+07 | Shadowing |
| 2 | Shadowing | Shadowing | 2.7570e+07 | Shadowing |
| 3 | 1.4474e+07 | 1.3480e+07 | 1.2119e+07 | 1.4547e+07 |
| 4 | 2.2203e+07 | 2.0713e+07 | 1.8636e+07 | 2.2329e+07 |
| 5 | 5.3956e+06 | 4.9853e+06 | 4.4664e+06 | 5.4072e+06 |
| 6 | 6.3604e+06 | 5.9368e+06 | 5.3427e+06 | 6.3977e+06 |
| 7 | 2.4489e+07 | 2.2860e+07 | 2.0573e+07 | 2.4634e+07 |
| 8 | 6.6200e+06 | 6.1438e+06 | 5.5152e+06 | 6.6450e+06 |
| 9 | 1.4233e+07 | 1.3282e+07 | 1.1952e+07 | 1.4316e+07 |
| 10 | Shadowing | Shadowing | Shadowing | Shadowing |
| 11 | Shadowing | Shadowing | 1.0344e+06 | Shadowing |
| 12 | 3.1697e+07 | Shadowing | Shadowing | Shadowing |
| 13 | 1.0045e+07 | 9.1110e+06 | 8.0953e+06 | 9.9997e+06 |
| 14 | Shadowing | Shadowing | Shadowing | Shadowing |
| 15 | 2.0002e+07 | 1.8138e+07 | 1.6115e+07 | 1.9910e+07 |
| 16 | 2.8524e+06 | 2.5545e+06 | 2.2565e+06 | 2.8266e+06 |
| 17 | 8.5474e+06 | 7.7584e+06 | 6.8959e+06 | 8.5111e+06 |
| 18 | Shadowing | Shadowing | Shadowing | Shadowing |
| 19 | 1.2964e+07 | Shadowing | Shadowing | Shadowing |
| 20 | 1.6282e+07 | Shadowing | Shadowing | Shadowing |
| 21 | 6.1921e+07 | Shadowing | Shadowing | Shadowing |
| 22 | 7.3859e+07 | Shadowing | Shadowing | 7.2180e+07 |

**Figure 7.29** Annual solar insolation in different seasons.

Chapter 8

# Conclusion and Future Directions

> "Success is no accident. It is hard work, perseverance, learning, studying, sacrifice and most of all, love of what you are doing or learning to do."
>
> Pele

In this chapter, we conclude the work presented in this thesis (Section 8.1) and recommend some possible future research directions (Section 8.2).

## 8.1 Contributions and Conclusion

The main objective of this research was to enhance the robustness of detection and reconstruction of 3-D buildings. To this end, we have proposed a framework that performs building detection, roof plane extraction, and 3-D modelling as sub-tasks. The proposed methods exploit the benefits of both LiDAR data and multispectral images. Since multispectral images are not always available, spatial and geometric features of LiDAR data were primarily used in this research.

In this thesis, the proposed framework has addressed problems of uncertainties in point cloud data, heterogeneity in appearance, the unavoidable noise due to the environment, terrain complexities, and the indefinite possibilities of different structures in building modelling. The contributions of our work to achieve the research objectives of this thesis are listed below:

- To address research objective **RO1**, i.e., identify the limitations of existing techniques for detection, roof plane recognition, and 3-D modelling of buildings, in

**Chapter 3** we conducted a comprehensive review of several techniques published in the last two decades. In addition to the spectral and spatial limitations of the data, several challenges, including shape variability, occlusion, shadows, object size, and height, were identified which were mainly due to point cloud sparsity, urban object differences, surrounding complexity, and data misalignment. We observed that the detection and reconstruction strategies for buildings and roof planes are not fully automatic and they use a set of parameters to perform their activities which make them subjective to particular datasets. These methods were further found to face difficulties in extracting and modelling buildings and their constituent roof planes which are small, partially occluded or shadowed. In addition, the effect of vegetation was either ignored altogether or resolved by eliminating the vegetation using a high height threshold value to separate the regions of interest from other objects.

- In **Chapter 4**, we addressed **RO2** and developed a methodology for the extraction and regularisation of buildings. Given the complementary advantages of LiDAR and image data, the fusion of two sources was chosen as a promising strategy to increase the building detection rate and the planimetric accuracy of building regions. The proposed technique is fully data-driven and automatic. The building delineation process is carried out by identifying the candidate building regions using connected component analysis. Next, the buildings are extracted, including partly occluded and shadowed buildings, after vegetation removal using the grid index structure and multisource data. Finally, the detected buildings are regularised by exploiting the image lines in the building regularisation process. These footprints represent the buildings at complexity level of $\text{LoD}_1$. The performance of the proposed technique was tested on the ISPRS benchmark and four Australian datasets introduced in Section 2.6. The results showed that our technique is not only able to extract small, partially occluded and shadowed buildings, but also to generate footprints irrespective of the surrounding complexity. The proposed method achieves a high detection rate even in the presence of a moderate registration error between the LiDAR data and the image.

- Since both LiDAR data and images are not always available, a robust segmentation technique using only LiDAR data, particularly focusing on the automatic detection of buildings and their roof planes, is proposed in **Chapter 5** to address **RO3**. The set objectives were met by following 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 utilises robust saliency features for roof extraction, which is less sensitive to noise and avoids over- and under-segmentation errors. Furthermore, a boundary extraction technique is presented that seamlessly extracts the boundary of an object and approximates the outline of any inner hole using only the LiDAR points. Based on experiments, it was demonstrated that the proposed technique achieves a high building detection rate and roof plane extraction performance on several datasets of variable point densities, terrain, and surrounding complexities. The technique is equally capable of detecting small buildings as well as small roof planes. Moreover, in most cases, the proposed technique is able to separate buildings and non-occluded parts from connected vegetation.

- **Chapter 6** addressed the core research objective, **RO4**, using the roof planes extracted in the previous chapter for automatic reconstruction of 3-D building models. As the modelling task is performed in unsupervised and data-driven fashion, the roof types are not restricted to a pre-existing model catalogue, as is the case in model-driven techniques. The roof planes, which are not extracted due to low point density, noise, and/or the vertical nature of the structures, are hypothesised using the roof topology assumption. As part of the modelling process, interrelations and interconnections among the building roof planes are used for the reconstruction of building models. Due to the effectiveness of the boundary tracing algorithm introduced in Chapter 5, building roofs of both the convex and concave types are reconstructed successfully. It was demonstrated that the buildings at higher levels of detail ($LoD_3$) are reconstructed by using individual roof planes and their interconnections based on their spatial adjacency.

- Chapter 7 addressed **RO5** and introduced an industrial application for large-scale assessment of solar potential. The proposed application utilised the extracted roofs and their geometrical information produced in Chapters 5 and 6 for the estimation of solar potential of individual buildings and their roof planes.

## 8.2 Directions for Further Research

As mentioned in the previous section, this research has covered several potential issues, particularly in the detection and reconstruction of building models. All the proposed techniques have been implemented and rigorously tested on a variety of datasets, and it can be concluded that the research objectives set out in this thesis have been achieved. However, from the viewpoint of the creation of a true 3-D building modelling framework, there are issues still left unattended that need to be investigated, some of which were highlighted in the literature review chapter. Below are some areas in which further research may be carried out:

- The quality of DTM plays a critical role in the accurate processing of input data, in particular, filtering the LiDAR data to separate the LiDAR points corresponding to the region of interest. Improvement in the generation of DTM and DSM will certainly increase the effectiveness of detection strategies which will subsequently improve reconstruction performance.

- Buildings made of transparent materials which allow lasers to penetrate through to the ground or below a certain threshold in height (1 m in this thesis) were generally neglected and therefore, such objects were not extracted. To overcome these limitations, the use of adaptive heights for LiDAR point cloud filtering, and both LiDAR intensity and image gradient for the extraction of low height and transparent buildings, respectively will be interesting to further improve the robustness of the detection strategies.

- The planimetric accuracy of the boundary was affected because of the registration error between LiDAR and image and the low horizontal accuracy of the LiDAR data. The resultant accuracy was also often affected by severe occlusion. Hence, further to the use of both radiometric and geometric clues, the planarity principle, which states the flatness of the planar segment while a tree does not, can be investigated to acquire high per-area and planimetric accuracy and high detection and reconstruction rates in complex areas.

- The proposed point cloud segmentation technique prefers planar surfaces which are common in urban environments. However, this assumption is not always correct. Therefore, the use of non-planar structures for point cloud segmentation should be investigated, in order to detect and model irregular surfaces e.g., spheres, cylinders, and cones.

- The proposed building modelling technique relies entirely on LiDAR data that sometimes fail to extract height discontinuities (step edges) due to the sparsity of the data points and the under-detected sides of roof planes. Therefore, recovery of intersection points between planes is difficult. These shortcomings can be overcome by using additional features e.g., the coplanarity of the neighbouring planes, the location of adjacent intersection points, or the use of more constraints. Another possible solution is to integrate the spectral features extracted from the corresponding aerial imagery. The information on lines extracted from images can also be used in conjunction with the LiDAR-based approximated intersection lines, in order to obtain accurate building models and reduce reconstruction errors.

- Some aspects e.g., buildings in close proximity, small and non-planar objects, vegetation on roofs, trees beside or overlapping roof planes, small roof surfaces, shadows, and occlusions, are covered in this research but these issues remain ongoing research challenges that need further investigation.

- The proposed techniques have been tested with a particular emphasis on complex scenes with a perfect blend of variable building structures, vegetation, shadows, and occlusions, which are common in suburban areas. Future work should test the applicability of the proposed detection and reconstruction techniques on city areas which are characterised by high-rise buildings.

This research has made progress on the creation of a robust and fully automatic modelling process for real-world complex objects. The modular approach of the proposed framework allows it to expand its functionalities at higher levels of detail for virtual reality applications, including precise mapping and monitoring of urban areas, while ensuring the consistency, accuracy, and reliability of the reconstructed models.

# Bibliography


[1] C. Fuchs, W. Forstner, E. Gülch, C. Heipke, and K. Eder, *OEEPE survey on 3D-city models*. Bundesamt für Kartographie und Geodäsie, 1998. (page 2)

[2] E. H. Wilson, J. D. Hurd, D. L. Civco, M. P. Prisloe, and C. Arnold, "Development of a geospatial model to quantify, describe and map urban growth," *Remote sensing of Environment*, vol. 86, no. 3, pp. 275–285, 2003. (page 2)

[3] S. Martinuzzi, W. A. Gould, and O. M. R. González, "Land development, land use, and urban sprawl in puerto rico integrating remote sensing and population census data," *Landscape and Urban Planning*, vol. 79, no. 3, pp. 288–297, 2007. (page 2)

[4] D. Chen, L. Zhang, J. Li, and R. Liu, "Urban building roof segmentation from airborne LiDAR point clouds," *International journal of remote sensing*, vol. 33, no. 20, pp. 6497–6515, 2012. (page 3)

[5] J. Song, J. Wu, and Y. Jiang, "Extraction and reconstruction of curved surface buildings by contour clustering using airborne LiDAR data," *International Journal for Light and Electron Optics*, vol. 126, no. 5, pp. 513–521, 2015. (page 3)

[6] C. Lin, A. Huertas, and R. Nevatia, "Detection of buildings using perceptual grouping and shadows," *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 62–69, 1994. (page 3)

[7] U. Weidner and W. Förstner, "Towards automatic building extraction from high-resolution digital elevation models," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 50, no. 4, pp. 38 – 49, 1995. (page 3, 80)

[8] M. Fradkin, H. Maìtre, and M. Roux, "Building detection from multiple aerial images in dense urban areas," *Computer Vision and Image Understanding*, vol. 82, no. 3, pp. 181 – 207, 2001. (page 3)

[9] J. Xiao, M. Gerke, and G. Vosselman, "Building extraction from oblique airborne imagery based on robust façade detection," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 68, no. 0, pp. 56 – 68, 2012. (page 3, 42, 48)

[10] Q. Zhang, R. Qin, X. Huang, Y. Fang, and L. Liu, "Classification of ultra-high resolution orthophotos combined with dsm using a dual morphological top hat profile," *Remote Sensing*, vol. 7, no. 12, pp. 16 422–16 440, 2015. (page 3, 42, 43, 48)

[11] G. Zhou, C. Song, J. Simmers, and P. Cheng, "Urban 3d gis from lidar and digital aerial images," *Computers & Geosciences*, vol. 30, no. 4, pp. 345–353, 2004. (page 3, 43, 45)

[12] Y. Li and H. Wu, "An improved building boundary extraction algorithm based on fusion of optical imagery and LiDAR data," *International Journal for Light and Electron Optics*, vol. 124, no. 22, pp. 5357 – 5362, 2013. (page 3, 39, 43, 45, 46)

[13] R. Qin and W. Fang, "A hierarchical building detection method for very high resolution remotely sensed images combined with dsm using graph cut optimization," *Photogrammetric Engineering & Remote Sensing*, vol. 80, no. 9, pp. 873–883, 2014. (page 3, 43, 48)

[14] L. Chen, S. Zhao, W. Han, and Y. Li, "Building detection in an urban area using lidar data and quickbird imagery," *International Journal of Remote Sensing*, vol. 33, no. 16, pp. 5135 – 5148, 2012. (page 3, 16, 43, 45, 46, 47, 71, 72)

[15] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik, "Using the dempstershafer method for the fusion of LiDAR data and multi-spectral images for building detection," *Information Fusion*, vol. 6, no. 4, pp. 283–300, 2005. (page 3, 14, 45, 47, 49, 72)

[16] L. Yong and W. Huayi, "Adaptive building edge detection by combining LiDAR data and aerial images," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVII, no. Part B1, pp. 197–202, 2008. (page 3, 45)

[17] M. Awrangjeb, C. Zhang, and C. S. Fraser, "Automatic extraction of building roofs using lidar data and multispectral imagery," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 83, no. 9, pp. 1 – 18, 2013. (page 3, 16, 39, 42, 45, 46, 47, 53, 71, 72, 75, 76)

[18] U. Nations, *World Urbanization Prospects 2014: Highlights*. United Nations Publications, 2014. (page 4)

[19] L. Zhu, "A pipeline of 3d scene reconstruction from point clouds," Ph.D. dissertation, Finnish Geospatial Research Institute, 2015. (page 4)

[20] O. Henricson, "Inferring homogeneous regions from rich image attributes," in *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, G. A., K. O., and A. P., Eds. Birkhuser Verlag, 1995, pp. 13–22. (page 4)

[21] C. Brenner, "Building reconstruction from images and laser scanning," *International Journal of Applied Earth Observation and Geoinformation*, vol. 6, no. 3, pp. 187–198, 2005. (page 4, 54)

[22] F. Rottensteiner, G. Sohn, M. Gerke, J. D. Wegner, U. Breitkopf, and J. Jung, "Results of the ISPRS benchmark on urban object detection and 3d building reconstruction," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 93, pp. 256–271, 2014. (page 4, 40, 41, 58, 91, 119)

[23] G. Vosselman, "Building reconstruction using planar faces in very high density height data," *International Archives of Photogrammetry and Remote Sensing*, vol. XXXII, no. 3, pp. 87–94, 1999. (page 4, 54)

[24] J. R. Jensen, *Remote sensing of the environment: An earth resource perspective 2/e*. Pearson Education India, 2009. (page 12)

[25] D. Sarala and S. Jacob, "Digital image processing-a remote sensing perspective," *International Journal of Innovative Research and Development*, vol. 3, no. 12, 2014. (page 12)

[26] L. B. Lentile, Z. A. Holden, A. M. Smith, M. J. Falkowski, A. T. Hudak, P. Morgan, S. A. Lewis, P. E. Gessler, and N. C. Benson, "Remote sensing techniques to assess

active fire characteristics and post-fire effects," *International Journal of Wildland Fire*, vol. 15, no. 3, pp. 319–345, 2006. (page 12)

[27] Remote sensing from wikipedia. https://tex.stackexchange.com/questions/3587/how-can-i-use-bibtex-to-cite-a-web-page. Last accessed: 09-08-2017. (page 12)

[28] C. Brenner and N. Haala, "Rapid acquisition of virtual reality city models from multiple data sources," *International Archives of Photogrammetry and Remote Sensing*, vol. XXXII, pp. 323–330, 1998. (page 12, 60)

[29] A. Henn, G. Gröger, V. Stroh, and L. Plümer, "Model driven reconstruction of roofs from sparse lidar point clouds," *ISPRS Journal of photogrammetry and remote sensing*, vol. 76, pp. 17–29, 2013. (page 12, 125)

[30] N. Haala, *Gebäuderekonstruktion durch Kombination von Bild-und Höhendaten*. Bayerischen Akademie der Wissenschaften, 1996, phd Thesis. (page 12, 54, 60)

[31] H. Sportouche, F. Tupin, and L. Denise, "Extraction and three-dimensional reconstruction of isolated buildings in urban scenes from high-resolution optical and sar spaceborne images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 3932–3946, 2011. (page 12, 54, 59)

[32] A. Ferro, D. Brunner, and L. Bruzzone, "Automatic detection and reconstruction of building radar footprints from single vhr sar images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 2, pp. 935–952, 2013. (page 12, 14)

[33] Q. Yu, P. Helmholz, and D. Belton, "Semantically enhanced 3d building model reconstruction from terrestrial laser-scanning data," *Journal of Surveying Engineering*, vol. 143, no. 4, p. 04017015, 2017. (page 12, 60)

[34] P. Dorninger and N. Pfeifer, "A comprehensive automated 3d approach for building extraction, reconstruction, and regularization from airborne laser scanning point clouds," *Sensors*, vol. 8, no. 11, pp. 7323–7343, 2008. (page 12, 51)

[35] A. Sampath and J. Shan, “Segmentation and reconstruction of polyhedral building roofs from aerial LiDAR point clouds,” *IEEE Transactions on geoscience and remote sensing*, vol. 48, no. 3, pp. 1554–1567, 2010. (page 12, 51, 54)

[36] D. Kong, L. Xu, X. Li, and S. Li, “K-plane-based classification of airborne lidar data for accurate building roof measurement,” *IEEE Transactions on Instrumentation and Measurement*, vol. 63, no. 5, pp. 1200–1214, 2014. (page 12, 56)

[37] M. Awrangjeb and C. S. Fraser, “Automatic segmentation of raw lidar data for extraction of building roofs,” *Remote Sensing*, vol. 6, no. 5, pp. 3716 – 3751, 2014. (page 12, 70, 76, 91, 105, 107, 109, 119, 120, 121)

[38] S. Gilani, M. Awrangjeb, and G. Lu, “Fusion of lidar data and multispectral imagery for effective building detection based on graph and connected component analysis,” *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. I, pp. 65 – 72, 2015. (page 12)

[39] S. Gilani, M. Awrangjeb, and G. Lu, “An automatic building extraction and regularisation technique using LiDAR point cloud data and orthoimage,” *Remote Sensing*, vol. 8, no. 3, p. 258, 2016. (page 12)

[40] S. A. N. Gilani, M. Awrangjeb, and G. Lu, “Segmentation of airborne point cloud data for automatic building roof extraction,” *GIScience & Remote Sensing*, vol. 55, no. 1, pp. 63–89, 2018. (page 12, 148)

[41] M. Awrangjeb, M. Ravanbakhsh, and C. S. Fraser, “Automatic detection of residential buildings using lidar data and multispectral imagery,” *ISPRS Journal of photogrammetry and remote sensing*, vol. 65, no. 5, pp. 457 – 467, 2010. (page 12, 48)

[42] M. Gerke and J. Xiao, “Fusion of airborne laserscanning point clouds and images for supervised and unsupervised scene classification,” *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 87, pp. 78 – 92, 2014. (page 12, 16, 71)

[43] W. Zhang, H. Wang, Y. Chen, K. Yan, and M. Chen, “3d building roof modeling by optimizing primitives parameters using constraints from lidar data and aerial imagery,” *Remote Sensing*, vol. 6, no. 9, pp. 8107–8133, 2014. (page 12, 62)

[44] E. Kwak, "Automatic 3d building model generation by integrating lidar and aerial images using a hybrid approach," Ph.D. dissertation, University of Calgary, 2013. (page 12, 62)

[45] P. R. Wolf and B. A. Dewitt, *Elements of Photogrammetry: with applications in GIS*. McGraw-Hill New York, 2000, vol. 3. (page 12)

[46] L. Zhao, X. Zhou, and G. Kuang, "Building detection from urban sar image using building characteristics and contextual information," *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, p. 56, 2013. (page 13)

[47] A. Huertas and R. Nevatia, "Detecting buildings in aerial images," *Computer Vision, Graphics, and Image Processing*, vol. 41, no. 2, pp. 131–152, 1988. (page 13)

[48] X. Huang and L. Zhang, "A multidirectional and multiscale morphological index for automatic building extraction from multispectral geoeye-1 imagery," *Photogrammetric Engineering & Remote Sensing*, vol. 77, no. 7, pp. 721–732, 2011. (page 14)

[49] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik, "Building detection using LiDAR data and multispectral images," *Digital Image Computing Techniques and Applications*, vol. 2, pp. 673–682, 2003. (page 14, 16, 47, 71)

[50] K. Navulur, *Multispectral image analysis using the object-oriented paradigm*. CRC press, 2006. (page 14)

[51] Expanding your point of view. https://learn.arcgis.com/en/arcgis-imagery-book/chapter4/. Last accessed: 09-08-2017. (page 15)

[52] A. Novacheva, "Building roof reconstruction from lidar data and aerial images through plane extraction and colour edge detection," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVII, pp. 53–57, 2008. (page 14, 51, 54)

[53] H. Mayer, "Automatic object extraction from aerial imagery - a survey focusing on buildings," *Computer Vision and Image Understanding*, vol. 74, no. 2, pp.

138–149, 1999. (page 14)

[54] M. Awrangjeb, M. Ravanbakhsh, and C. Fraser, "Automatic building detection using lidar data and multispectral imagery," *Digital Image Computing Techniques and Applications*, pp. 45–51, 2010. (page 16, 72)

[55] D. Grigillo and U. Kanjir, "Urban object extraction from digital surface model and digital aerial images," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. I, no. 3, pp. 215 – 220, 2012. (page 16, 47, 71)

[56] The leaffest line. https://publiclab.org/notes/cfastie/09-04-2014/the-leaffest-line. Last accessed: 11-08-2017. (page 16)

[57] Dital elevation model. http://www.blograf.com/dem-dtm-ve-dsm-nedir/. Last accessed: 11-08-2017. (page 19)

[58] Z. Li, C. Zhu, and C. Gold, *Digital terrain modeling: principles and methodology*. CRC press, 2004. (page 19)

[59] P. Gueudet, G. Wells, D. Maidment, and A. Neuenschwander, "Influence of the postspacing density of the lidar-derived dem on flood modeling," in *Geographic Information Systems and Water Resources III–AWRA Spring Specialty Conference. AWRA, Nashville, Tennessee*, 2004. (page 20)

[60] F. EarthData, "Inc., lidar mapping fact sheet, white paper published in the company web-site." (page 20)

[61] J. Carter, K. Schmid, K. Waters, L. Betzhold, B. Hadley, R. Mataosky, and J. Halleran, "Lidar 101: An introduction to lidar technology, data, and applications," *National Oceanic and Atmospheric Administration (NOAA) Coastal Services Center*, pp. 7–11, 2012. (page 20, 21)

[62] S. S. Stevens *et al.*, *On the theory of scales of measurement*. Bobbs-Merrill, College Division, 1946. (page 20)

[63] F. Robert A., A. Samberg, F. Martin, and T. Greaves, *Topographic and Terrestrial Lidar*. American Society for Photogrammetry and Remote Sensing Bethesda, 2007. (page 21)

[64] F. Leberl, A. Irschara, T. Pock, P. Meixner, M. Gruber, S. Scholz, and A. Wiechert, "Point clouds," *Photogrammetric Engineering & Remote Sensing*, vol. 76, no. 10, pp. 1123–1134, 2010. (page 21)

[65] S. O. Elberink and G. Vosselman, "Quality analysis on 3d building models reconstructed from airborne laser scanning data," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 2, pp. 157–165, 2011. (page 22, 54, 64, 96, 126)

[66] Delaunay triangulation from wikipedia. https://en.wikipedia.org/wiki/Delaunay_triangulation. Last accessed: 14-08-2017. (page 22)

[67] I. T. Jolliffe, "Principal component analysis and factor analysis," in *Principal component analysis*. Springer, 1986, pp. 115–128. (page 22)

[68] D. C. Lay, S. R. Lay, and J. J. McDonald, *Linear Algebra and Its Applications, 5th Edition*. Pearson, 2012. (page 23)

[69] A. Nurunnabi, D. Belton, and G. West, "Robust segmentation in laser scanning 3d point cloud data," in *Digital Image Computing Techniques and Applications*, 2012, pp. 1–8. (page 23, 25, 102, 103, 105)

[70] M. Rutzinger, F. Rottensteiner, and N. Pfeifer, "A comparison of evaluation techniques for building extraction from airborne laser scanning," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 2, no. 1, pp. 11 – 20, 2009. (page 27, 28, 29, 30)

[71] M. Cramer, "The DGPF-test on digital airborne camera evaluation–overview and test design," *Photogrammetrie Fernerkundung Geoinformatics*, vol. 2010, no. 2, pp. 73–82, 2010. (page 27, 32)

[72] M. Awrangjeb and C. Fraser, "An automatic and threshold-free performance evaluation system for building extraction techniques from airborne LiDAR data,"

*IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 10, pp. 4184–4198, 2014. (page 27, 29, 30, 31, 91)

[73] W. Song and T. L. Haithcoat, "Development of comprehensive accuracy assessment indexes for building footprint extraction," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 2, pp. 402–404, 2005. (page 31)

[74] A. Sampath and J. Shan, "Building boundary tracing and regularization from airborne LiDAR point clouds," *Photogrammetric Engineering & Remote Sensing*, vol. 73, no. 7, pp. 805–812, 2007. (page 39, 49)

[75] A. Habib, E. Kwak, and M. Al-Durgham, "Model-based automatic 3d building model generation by integrating lidar and aerial images," *Archives of Photogrammetry, Cartography and Remote Sensing*, vol. 22, pp. 187–200, 2011. (page 39)

[76] H. Wiman, *Least Squares Matching for Three Dimensional Building Reconstruction*. Springer, 1997, pp. 223–232. (page 40)

[77] A. Brunn, "Techniques for automatic building extraction," *Third Course in Digital Photogrammetry, Institute for Photogrammetry at Bonn University and Landesvermessungsamt Nordrhein-Westfalen, Bonn, Germany*, 1998. (page 42)

[78] D. H. Al-Khudhairy, "Geo-spatial information and technologies in support of eu crisis management," *International Journal of Digital Earth*, vol. 3, no. 1, pp. 16 – 30, 2010. (page 42)

[79] L. Sahar and A. Krupnik, "Semiautomatic extraction of building outlines from large-scale aerial images," *Photogrammetric Engineering & Remote Sensing*, vol. 65, pp. 459 – 466, 1999. (page 42)

[80] Z. Song, C. Pan, Q. Yang, F. Li, and W. Li, "Building roof detection from a single high-resolution satellite image in dense urban area," in *Proceeding of ISPRS Congress*, 2008, pp. 271–277. (page 43)

[81] W. Liu and V. Prinet, "Building detection from high-resolution satellite image using probability model," in *Proceedings of IEEE International Geoscience and Remote Sensing Symposium*, vol. 6, 2005, pp. 3888–3891. (page 43)

[82] M. Izadi and P. Saeedi, "Automatic building detection in aerial images using a hierarchical feature based image segmentation," in *International Conference on Pattern Recognition*, 2010, pp. 472–475. (page 44)

[83] B. Sirmacek and C. Unsalan, "A probabilistic framework to detect buildings in aerial and satellite images," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 1, pp. 211–221, 2011. (page 44)

[84] L. Cheng, J. Gong, M. Li, and Y. Liu, "3d building model reconstruction from multi-view aerial imagery and LiDAR data," *Photogrammetric Engineering & Remote Sensing*, vol. 77, no. 2, pp. 125–139, 2011. (page 45)

[85] B. Yang, W. Xu, and Z. Dong, "Automated extraction of building outlines from airborne laser scanning point clouds," *IEEE Geoscience and Remote Sensing Letters*, vol. 10, no. 6, pp. 1399 – 1403, 2013. (page 45, 91, 119, 120)

[86] D. Mongus, N. Lukac, D. Obrul, and B. Zalik, "Detection of planar points for building extraction from lidar data based on differential morphological and attribute profiles," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 1, pp. 21–26, 2013. (page 46)

[87] M. Awrangjeb and C. S. Fraser, "Automatic segmentation of raw LiDAR data for extraction of building roofs," *Remote Sensing*, vol. 6, no. 5, pp. 3716–3751, 2014. (page 46, 50, 148)

[88] N. Haala and V. Walter, "Automatic classification of urban environments for database revision using lidar and color aerial imagery," *International Archives of Photogrammetry and Remote Sensing*, vol. XXXII, no. Part 7, pp. 4 – 3, 1999. (page 46)

[89] N. Demir, D. Poli, and E. Baltsavias, "Extraction of buildings using images & LiDAR data and a combination of various methods," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVIII, no. part 3/W4, pp. 71 – 76, 2009. (page 46)

[90] T. Vu, F. Yamazaki, and M. Matsuoka., "Multi-scale solution for building extraction from LiDAR and image data," *International Journal of Applied Earth*

*Observation and Geoinformation*, vol. 11, no. 4, pp. 281–289, 2009. (page 47)

[91] K. Khoshelham, S. Nedkov, and C. Nardinocchi, "A comparison of bayesian and evidence-based fusion methods for automated building detection in aerial data," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVII, no. part B7, pp. 1183 – 1188, 2008. (page 48, 53)

[92] A. Alharthy and J. Bethel, "Heuristic filtering and 3d feature extraction from lidar data," *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, vol. XXXIV, no. 3/A, pp. 29 – 34, 2002. (page 49)

[93] C.-S. Fu and J. Shan, "3-d building reconstruction from unstructured distinct points," *International Archives of Photogrammetry and Remote Sensing*, vol. XXXV, pp. 553 – 558, 2004. (page 49)

[94] R. Ma, "Dem generation and building detection from lidar data," *Photogrammetric Engineering & Remote Sensing*, vol. 71, no. 7, pp. 847 – 854, 2005. (page 49)

[95] S. R. Lach and J. P. Kerekes, "Robust extraction of exterior building boundaries from topographic lidar data," *IEEE International Geoscience and Remote Sensing Symposium*, vol. 2, pp. II – 85, 2008. (page 49)

[96] G. Sohn, X. Huang, and V. Tao, "Using a binary space partitioning tree for reconstructing polyhedral building models from airborne lidar data," *Photogrammetric Engineering & Remote Sensing*, vol. 74, no. 11, pp. 1425 – 1438, 2008. (page 49)

[97] S. Wei, "Building boundary extraction based on lidar point clouds data," *International Archives of Photogrammetry Remote Sensing and Spatial Information Sciences*, vol. XXXVII, pp. 157 – 161, 2008. (page 49)

[98] S. Abdullah, "Height adaptive lidar segmentation for building extraction and roof reconstruction," Ph.D. dissertation, School of Information Technology, Monash University, 2014. (page 50)

[99] A. Boulch and R. Marlet, "Fast and robust normal estimation for point clouds with sharp features," in *Computer graphics forum*, vol. 31, no. 5, 2012, pp. 1765–1774. (page 50, 102)

[100] H. Huang and C. Brenner, "Rule-based roof plane detection and segmentation from laser point clouds," in *Joint Urban Remote Sensing Event*, 2011, pp. 293–296. (page 50)

[101] G. Vosselman and R. Klein, "Visualisation and structuring of point clouds," *Airborne and Terrestrial Laser Scanning*, vol. 1, pp. 43–79, 2010. (page 50)

[102] R. Schnabel, R. Wahl, and R. Klein, "Efficient ransac for point-cloud shape detection," in *Computer graphics forum*, vol. 26, no. 2, 2007, pp. 214–226. (page 50)

[103] J.-E. Deschaud and F. Goulette, "A fast and accurate plane detection algorithm for large noisy point clouds using filtered normals and voxel growing," in *International Symposium on 3D Data Processing, Visualization and Transformation*, 2010. (page 50)

[104] Y. Liu and Y. Xiong, "Automatic segmentation of unorganized noisy point clouds based on the gaussian map," *Computer-Aided Design*, vol. 40, no. 5, pp. 576–594, 2008. (page 50)

[105] G. Vosselman, "Slope based filtering of laser altimetry data," *International Archives of Photogrammetry and Remote Sensing*, vol. XXXIII, pp. 935–942, 2000. (page 51)

[106] M. Arastounia and D. Lichti, "Automatic extraction of insulators from 3d lidar data of an electrical substation," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 2, no. 5/W2, pp. 19–24, 2013. (page 51)

[107] A. K. Jain, M. N. Murty, and P. J. Flynn, "Data clustering: a review," *ACM computing surveys (CSUR)*, vol. 31, no. 3, pp. 264–323, 1999. (page 51)

[108] Q.-Y. Zhou and U. Neumann, "Fast and extensible building modeling from airborne LiDAR data," in *Proceedings of SIGSPATIAL international conference on Ad-*

*vances in geographic information systems*, 2008, p. 7. (page 52)

[109] M. Awrangjeb and C. S. Fraser, "Rule-based segmentation of lidar point cloud for automatic extraction of building roof planes," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 2, no. 3/W1, pp. 1–6, 2013. (page 52)

[110] A. Jochem, B. Höfle, V. Wichmann, M. Rutzinger, and A. Zipf, "Area-wide roof plane segmentation in airborne LiDAR point clouds," *Computers, Environment and Urban Systems*, vol. 36, no. 1, pp. 54–64, 2012. (page 52)

[111] S. Sun, *Automatic 3D building detection and modeling from Airborne LiDAR point clouds*. Rochester Institute of Technology, 2013. (page 54, 64)

[112] F. Rottensteiner, J. Trinder, S. Clode, and K. Kubik, "Automated delineation of roof planes from lidar data," in *ISPRS Workshop Laser scanning*, vol. XXXVI, 2005, pp. 221–226. (page 54, 58)

[113] F. Rottensteiner, G. Sohn, J. Jung, M. Gerke, C. Baillard, S. Benitez, and U. Breitkopf, "The isprs benchmark on urban object classification and 3d building reconstruction," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 1, no. 3, pp. 293–298, 2012. (page 54, 58)

[114] F. Rottensteiner, G. Sohn, M. Gerke, and J. D. Wegner, "Isprs test project on urban classification and 3d building reconstruction," *Commission III-Photogrammetric Computer Vision and Image Analysis, Working Group III/4-3D Scene Analysis*, pp. 1–17, 2013. (page 54, 58)

[115] G. Vosselman, S. Dijkman *et al.*, "3d building model reconstruction from point clouds and ground plans," *International archives of photogrammetry remote sensing and spatial information sciences*, vol. XXXIV, no. 3, pp. 37–44, 2001. (page 54, 55)

[116] G. Vosselman and I. Suveg, "Map based building reconstruction from laser data and images," *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pp. 231–239, 2001. (page 54)

[117] K. Kim and J. Shan, "Building roof modeling from airborne laser scanning data based on level set approach," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 4, pp. 484–497, 2011. (page 54, 55, 142)

[118] L. Cheng, L. Tong, Y. Chen, W. Zhang, J. Shan, Y. Liu, and M. Li, "Integration of lidar data and optical multi-view images for 3d reconstruction of building roofs," *Optics and Lasers in Engineering*, vol. 51, no. 4, pp. 493–502, 2013. (page 54)

[119] A. F. Habib, R. Zhai, and C. Kim, "Generation of complex polyhedral building models by integrating stereo-aerial imagery and lidar data," *Photogrammetric engineering & remote sensing*, vol. 76, no. 5, pp. 609–623, 2010. (page 54)

[120] S. Noronha and R. Nevatia, "Detection and modeling of buildings from multiple aerial images," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 23, no. 5, pp. 501–518, 2001. (page 54, 58)

[121] N. Haala, "Combining multiple data sources for urban data acquisition, institut für photogrammetrie," *Proceeding of Photogrammetrische Woche, Fritsch/Spiller (Eds.)*, vol. 6, 1999. (page 54)

[122] G. Sohn, Y. Jwa, J. Jung, and H. Kim, "An implicit regularization for 3d building rooftop modeling using airborne lidar data," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. I, no. 3, pp. 305–310, 2012. (page 56, 57)

[123] Y. Jwa, G. Sohn, V. Tao, and W. Cho, "An implicit geometric regularization of 3d building shape using airborne lidar data," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXVII, pp. 69–76, 2008. (page 56)

[124] G. Sohn, X. Huang, and V. Tao, "Using a binary space partitioning tree for reconstructing polyhedral building models from airborne lidar data," *Photogrammetric Engineering & Remote Sensing*, vol. 74, no. 11, pp. 1425–1438, 2008. (page 56)

[125] C. Poullis and S. You, "3d reconstruction of urban areas," in *International Conference on 3D Imaging, Modeling, Processing, Visualization and Transmission (3DIMPVT)*, 2011, pp. 33–40. (page 56)

[126] J. Jung, Y. Jwa, and G. Sohn, "Implicit regularization for reconstructing 3d building rooftop models using airborne lidar data," *Sensors*, vol. 17, no. 3, p. 621, 2017. (page 57)

[127] B. Wu, B. Yu, Q. Wu, S. Yao, F. Zhao, W. Mao, and J. Wu, "A graph-based approach for 3d building model reconstruction from airborne lidar point clouds," *Remote Sensing*, vol. 9, no. 1, p. 92, 2017. (page 57)

[128] H. Arefi and P. Reinartz, "Building reconstruction using dsm and orthorectified images," *Remote Sensing*, vol. 5, no. 4, pp. 1681–1703, 2013. (page 59)

[129] S. Wang, Y.-H. Tseng, and A. F. Habib, "Least-squares building model fitting using aerial photos and lidar data," in *Annual Conference, Proceedings of the ASPRS*, vol. 2630, 2010. (page 61)

[130] S. Seo, "Model-based automatic building extraction from lidar and aerial imagery," Ph.D. dissertation, The Ohio State University, 2003. (page 61)

[131] I. Suveg and G. Vosselman, "Reconstruction of 3d building models from aerial images and maps," *ISPRS Journal of Photogrammetry and remote sensing*, vol. 58, no. 3, pp. 202–224, 2004. (page 61)

[132] T. H. Kolbe, G. Gröger, and L. Plümer, "Citygml–interoperable access to 3d city models," *Geo-information for disaster management*, pp. 883–899, 2005. (page 64)

[133] Merrick. (2015) The merrick advanced remote sensing (mars®). (accessed on 29 December 2015). [Online]. Available: http://www.merrick.com/ (page 69, 98)

[134] E. G. Parmehr, C. S. Fraser, C. Zhang, and J. Leach, "Automatic registration of optical imagery with 3d lidar data using statistical similarity," *ISPRS Journal of photogrammetry and remote sensing*, vol. 88, pp. 28 – 40, 2014. (page 70)

[135] R. C. Gonzalez, R. E. Woods, and S. L. Eddins, *Digital Image Processing Using MATLAB*. New Jersey: Prentice Hall, 2003. (page 71, 73)

[136] Z. Lari and A. Habib, "An adaptive approach for the segmentation and extraction of planar and linear/cylindrical features from laser scanning data," *ISPRS Journal of photogrammetry and remote sensing*, vol. 93, pp. 192 – 212, 2014. (page 76)

[137] M. Awrangjeb, G. Lu, C. S. Fraser, and M. Ravanbakhsh, "A fast corner detector based on the chord-to-point distance accumulation technique," *Digital Image Computing Techniques and Applications*, pp. 519 – 525, 2009. (page 82)

[138] M. Awrangjeb, G. Lu, and C. Fraser, "Automatic building extraction from lidar data covering complex urban scenes," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. I, pp. 25 – 32, 2014. (page 82, 91)

[139] ISPRS, "Results: ISPRS test project on urban classification and 3d building reconstruction," 2016, (accessed on Oct 28, 2016). [Online]. Available: http://www2.isprs.org/commissions/comm3/wg4/results.html (page 83, 111, 119)

[140] A. Zarea and A. Mohammadzadeh, "Introducing a new approach for buildings detection using lidar data and aerial image," 2015, (accessed on 29 December 2015). [Online]. Available: https://www.itc.nl/external/ISPRS_WGIII4/ISPRSIII_4_Test_results/papers/KNTU_2.pdf (page 91)

[141] Q. Zhan, Y. Liang, C. Wei, and Y. Xiao, "Ground object recognition using combined high resolution airborne images and dsm," *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. XXXIX, p. B3, 2012. (page 91)

[142] A. M. Ramiya and R. R. Nidamanuri, "Automated building detection using lidar data and digital aerial images," 2014, (accessed on 29 December 2015). [Online]. Available: www.itc.nl/external/ISPRS_WGIII4/ISPRSIII_4_Test_results/papers/IIST_detection.pdf (page 91)

[143] S. Sotoodeh, "Outlier detection in laser scanner point clouds," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 36, no. 5, pp. 297–302, 2006. (page 96)

[144] J. Zhang, J. Cao, X. Liu, J. Wang, J. Liu, and X. Shi, "Point cloud normal estimation via low-rank subspace clustering," *Computers & Graphics*, vol. 37, no. 6, pp. 697–706, 2013. (page 98, 102, 104)

[145] D. Chen, L. Zhang, P. Mathiopoulos, and X. Huang, "A methodology for automated segmentation and reconstruction of urban 3-d buildings from als point clouds," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 10, pp. 4199–4217, 2014. (page 101, 105, 107)

[146] M. Awrangjeb, "Using point cloud data to identify, trace, and regularize the outlines of buildings," *International Journal of Remote Sensing*, vol. 37, no. 3, pp. 551–579, 2016. (page 101)

[147] C. Weber, S. Hahmann, and H. Hagen, "Methods for feature detection in point clouds," in *OpenAccess Series in Informatics*, vol. 19, 2011. (page 101)

[148] M. K. Park, S. J. Lee, and K. H. Lee, "Multi-scale tensor voting for feature extraction from unstructured point clouds," *Graphical Models*, vol. 74, no. 4, pp. 197–208, 2012. (page 101)

[149] B. Li, R. Schnabel, R. Klein, Z. Cheng, G. Dang, and S. Jin, "Robust normal estimation for point clouds with sharp features," *Computers & Graphics*, vol. 34, no. 2, pp. 94–106, 2010. (page 102)

[150] N. J. Mitra, A. Nguyen, and L. Guibas, "Estimating surface normals in noisy point cloud data," *International Journal of Computational Geometry & Applications*, vol. 14, no. 04n05, pp. 261–276, 2004. (page 102)

[151] A. C. Öztireli, G. Guennebaud, and M. Gross, "Feature preserving point set surfaces based on non-linear kernel regression," in *Computer Graphics Forum*, vol. 28, no. 2, 2009, pp. 493–501. (page 102)

[152] S. Fleishman, D. Cohen-Or, and C. Silva, "Robust moving least-squares fitting with sharp features," in *ACM transactions on graphics*, vol. 24, no. 3, 2005, pp. 544–552. (page 102)

[153] P. Luo, Z. Wu, C. Xia, L. Feng, and B. Jia, "Robust normal estimation of point cloud with sharp features via subspace clustering," in *Fifth International Conference on Graphic and Image Processing*, vol. 9069, 2014, pp. 90 691S–90 691S–6. (page 102)

[154] M. Pauly, R. Keiser, and M. Gross, "Multi-scale feature extraction on point-sampled surfaces," in *Computer graphics forum*, vol. 22, no. 3, 2003, pp. 281–289. (page 103)

[155] S. O. Elberink and G. Vosselman, "Entities and features for classifcation of airborne laser scanning data in urban area," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 257–262, 2012. (page 108)

[156] M. Awrangjeb, G. Lu, and C. S. Fraser, "Automatic building extraction from LiDAR data covering complex urban scenes," *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. I, pp. 25–32, 2014. (page 108)

[157] H. Hu. (2016) New approach for buildings detection using LiDAR supervoxels and multi-level semantic. (accessed on Oct 28, 2016). [Online]. Available: https://www.itc.nl/external/ISPRS_WGIII4/ISPRSIII_4_Test_results/papers/method_description_HanHK.pdf (page 120)

[158] P. Dorninger and N. Pfeifer, "A comprehensive automated 3d approach for building extraction, reconstruction, and regularization from airborne laser scanning point clouds," *Sensors*, vol. 8, no. 11, pp. 7323–7343, 2008. (page 120, 121)

[159] S. N. Perera, H. A. Nalani, and H.-G. Maas, "An automated method for 3d roof outline generation and regularization in airborne laser scanner data," *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 281–286, 2012. (page 121)

[160] G. Sohn, X. Huang, and V. Tao, "A data-driven method for modeling 3d building objects using a binary space partitioning tree," *Topographic Laser Ranging and Scanning: Principles and Processing*, pp. 479–509, 2009. (page 124)

[161] G. Gröger, T. Kolbe, and A. Czerwinski, "Candidate opengis citygml implementation specification (city geography markup language)," *Open Geospatial Consortium Inc.*, 2007. (page 125)

[162] M.-O. Löwner, J. Benner, G. Gröger, U. Gruber, K.-H. Häfele, and S. Schlüter, "Citygml 2.0-ein internationaler standard für 3d-stadtmodelle, teil 1: Datenmodell," *Zeitschrift für Geodäsie, Geoinformation und Landmanagement*, vol. 6, no. 2012, pp. 340–349, 2012. (page 125)

[163] S. 3D, "Modeling guide for 3d objects - part 2: Modeling of buildings (lod1, lod2, lod3)," *SIG 3D Quality Working Group*, 2014, (Last updated on 12 November 2014). [Online]. Available: http://en.wiki.quality.sig3d.org/index.php/Modeling_Guide_for_3D_Objects_-_Part_2:_Modeling_of_Buildings_(LoD1,_LoD2,_LoD3) (page 125)

[164] S. Gilani, M. Awrangjeb, and G. Lu, "Robust building roof segmentation using airborne point cloud data," *IEEE International Conference on Image Processing*, pp. 859 – 863, 2016. (page 127, 148, 155)

[165] M. Awrangjeb, "Using point cloud data to identify, trace, and regularize the outlines of buildings," *International Journal of Remote Sensing*, vol. 37, no. 3, pp. 551–579, 2016. (page 136)

[166] M. Awrangjeb, G. Lu, and C. S. Fraser, "Performance comparisons of contour-based corner detectors," *IEEE Transactions on Image Processing*, vol. 21, no. 9, pp. 4167–4179, 2012. (page 136)

[167] V. Verma, R. Kumar, and S. Hsu, "3d building detection and modeling from aerial lidar data," in *Conference on Computer Vision and Pattern Recognition*, vol. 2, 2006, pp. 2213–2220. (page 138)

[168] G. S. N. Perera and H.-G. Maas, "Cycle graph analysis for 3d roof structure

modelling: Concepts and performance," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 93, pp. 213 – 226, 2014. (page 138)

[169] M. Kabolizade, H. Ebadi, and A. Mohammadzadeh, "Design and implementation of an algorithm for automatic 3d reconstruction of building models using genetic algorithm," *International Journal of Applied Earth Observation and Geoinformation*, vol. 19, pp. 104–114, 2012. (page 142)

[170] Solar pv maps and tools. http://pv-map.apvi.org.au/. Last accessed: 21-08-2017. (page 147, 148)

[171] B. Keller and A. Costa, "A matlab gui for calculating the solar radiation and shading of surfaces on the earth," *Computer Applications in Engineering Education*, vol. 19, no. 1, pp. 161–170, 2011. (page 149, 150, 151, 152)

[172] A. A. Abood, "A comprehensive solar angles simulation and calculation using matlab," *International Journal of Energy and Environment*, vol. 6, no. 4, p. 367, 2015. (page 149, 150, 151, 153)

[173] J. Spencer, "Fourier series representation of the position of the sun," *Search*, vol. 2, no. 5, pp. 172–172, 1971. (page 149)

[174] B. McCune, "Improved estimates of incident radiation and heat load using nonparametric regression against topographic variables," *Journal of Vegetation Science*, vol. 18, no. 5, pp. 751–754, 2007. (page 150, 155)

[175] N. Lukač and B. Žalik, "Gpu-based roofs' solar potential estimation using lidar data," *Computers & Geosciences*, vol. 52, pp. 34–41, 2013. (page 151, 153)

[176] N. Lukač, D. Žlaus, S. Seme, B. Žalik, and G. Štumberger, "Rating of roofs surfaces regarding their solar potential and suitability for pv systems, based on lidar data," *Applied energy*, vol. 102, pp. 803–812, 2013. (page 151, 153)

[177] A. H. Hafeez, I. Reda, J. Shazly, and M. Eteiba, "Simulation and estimation of a daily global solar radiation in egypt," *Middle East Journal of Scientific Research*, vol. 23, no. 5, pp. 880–895, 2015. (page 151, 153)

[178] G. S. Campbell and J. M. Norman, *An introduction to environmental biophysics*. Springer Science & Business Media, 2012. (page 153, 154)

[179] I. Reda and A. Andreas, "Solar position algorithm for solar radiation applications," *Solar energy*, vol. 76, no. 5, pp. 577–589, 2004. (page 153, 160)

[180] J. A. Duffie and W. A. Beckman, *Solar engineering of thermal processes*. John Wiley and Sons, New York, NY, 1980. (page 153, 154)

[181] W. Gander and J. Hrebicek, *Solving problems in scientific computing using Maple and Matlab®*. Springer Science & Business Media, 2011. (page 160)

[182] Shadow length. http://planetcalc.com/1875/. Last accessed: 25-08-2017. (page 163)

[183] Tools for consumers and designers of solar. http://www.sunearthtools.com/dp/tools/pos_sun.php. Last accessed: 25-08-2017. (page 163)

[184] Find my shadow. http://www.findmyshadow.com/. Last accessed: 25-08-2017. (page 163)