INVESTIGATING SEMI-AUTOMATED CADASTRAL BOUNDARY EXTRACTION FROM AIRBORNE LASER SCANNED DATA

XIANGHUAN LUO
March, 2016

SUPERVISORS:
Dr. R.M. Bennett
Dr. M.N. Koeva

ADVISORS:
Dr.ir. C.H.J. Lemmen
Dr. N. Quadros

INVESTIGATING SEMI-AUTOMATED CADASTRAL BOUNDARY EXTRACTION FROM AIRBORNE LASER SCANNED DATA

XIANGHUAN LUO
Enschede, The Netherlands, March, 2016

Thesis submitted to the Faculty of Geo-Information Science and Earth Observation of the University of Twente in partial fulfilment of the requirements for the degree of Master of Science in Geo-information Science and Earth Observation.
Specialization: Land Administration

SUPERVISORS:
Dr. R.M. Bennett
Dr. M.N. Koeva

ADVISORS:
Dr.ir. C.H.J. Lemmen
Dr. N. Quadros

THESIS ASSESSMENT BOARD:
Prof.mr.dr. J.A. Zevenbergen
Prof.dr.ir. P.J.M. van Oosterom, Delft University of Technology

ABSTRACT

Cadastres are commonly argued as a key part in guaranteeing land tenure security. A complete cadastral map provides confidence for the society about the range and location of land rights. However, in many developing contexts there still exists a dearth of credible land rights records. Innovative methods are needed to speed up the process of cadastral mapping. Airborne laser scanning (ALS) techniques can produce highly accurate three-dimensional data, and the technique is gaining increased popularity in the land-surveying field. This study focused on applying airborne laser scanning techniques to the land administration challenge of general boundary mapping. A semi-automated workflow is developed to extract cadastral boundaries from the airborne laser scanning data.

First, the study investigated the relationship between topographic objects and cadastral boundaries in the case context of Port Vila, Vanuatu. Overlays of cadastral parcel data on satellite images revealed that over eighty percentages of cadastral boundaries coincide with topographic objects. Specifically, in dense urban areas, road edges and building walls coincide with the majority of cadastral boundaries, while in suburban areas, the fence instead of buildings plays an important role in marking a parcel border. Therefore, constructing a map that depicts these features can contribute a lot to expedited cadastral mapping.

Second, a two-phased workflow was developed that focused on extracting digital representations of the physical objects. Point clouds were first classified into semantic components. Points of roads were identified according to their reflectance intensity, and then connected component analysis was applied to reconstruct the road surface. The outline of planar objects such as building roofs and road surfaces were generated by α-shape algorithm. Points of fence were projected into the raster, and centrelines were fitted into connected pixels to generate vector fences. The extracted lines together constructed a rough parcel map. Afterwards, the extracted vector lines were edited and completed during the post-refinement phase.

Third, the workflow result was compared with the exiting cadastral map as reference, in order to quantitatively evaluate the performance of the developed process. It was found that two thirds of the extracted lines coincided with the true cadastral boundaries, and that roughly one quarter of the cadastral boundaries can be reconstructed by the developed workflow. Therefore, it is argued that the semi-automated extraction workflow could effectively speed up cadastral surveying: both human resources and equipment costs could be reduced. A key point of advantage over image-based technique is that LiDAR is able to penetrate tree canopies: in images they are invisible. However, at this point in time the spatial accuracy of the workflow still cannot meet the requirement of many conventional cadastral mapping standards.

Finally, key lessons and possible improvements were observed and compiled. Extracting small objects such as fences occurring in the test site requires a higher point density in the scanned data. Moreover, the semi-automated extraction workflow performed better in more regular suburban areas. By contrast more researches are needed on parcel morphology of dense urban areas.

ACKNOWLEDGEMENTS

Land Administration is an essential component for a country's development, and I am so grateful that I had such an opportunity to devote much time to it. Though the research result might be modest, I have experienced a fascinating time wandering in the word of LiDAR's application for cadastral mapping. It was such an interesting and meaningful topic.

Firstly I would like to give thanks to my supervisors at ITC, Dr. Bennett and Dr. Koeva. This thesis was completed under their guidance. It was their innovative ideas and continuous encouragements that guided my work through all the bottlenecks. My sincerely thanks go to Dr. Lemmen, I have learnt a lot from his rich experiences related to cadastre. Many thanks also go to Dr. Quadros, for sharing his expansive knowledge on feature extraction. Also thanks to Land Equity and Land Administration Department of Vanuatu for supplying me with the research data.

I also would like to express my thanks to all the ITC professors and tutors who ever taught me during my stay at the Faculty ITC. The knowledge I learned from their courses not only contributes to the thesis, but also helped me to establish a solid foundation for my future career.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

ALS	Airborne Laser Scanning
DEM	Digital Elevation Model
FIG	International Federation of Surveyor
GIS	Geographic Information System
GNSS	Global Navigation Satellite System
GPS	Global Positing System
LiDAR	Light Detection And Ranging
TIN	Triangulated Irregular Network
TLS	Terrestrial Laser Scanner
RS	Remote Sensing
UN	United Nations
α-Shape	Alpha Shape Algorithm
PCs	Point Clouds

1. CHAPTER 1 INTRODUCTION

1.1. Motivation and Background

According to a recent FIG report, about 75 percent of land in the world is still not registered in a statutory cadastral system, and most of these are located in developing countries (Enemark, Bell, Lemmen, & McLaren, 2014). Low coverage of land registration hinders social development. A formal register serves as legal evidence of humans' right to land. Without confirmation from a register, people tend to find it more difficult to access and use land. In this situation, land disputes are likely to be aroused and may result in land grabbing and disorder occurs in the land market. Hence it is an urgent issue to accelerate the procedure of land registration.

A cadastre generally is considered as a comprehensive register that records the ownership, tenure as well as real property's location and dimension (Roger, Kain, & Elizabeth, 1992). Parcel boundaries are observed by means of cadastral survey. The cadastral boundary describes the legal boundary of a real property. Therefore, unambiguous cadastral boundaries are considered necessary for the land tenure security. Besides legal aspects, the characteristic of instantaneity make cadastral mapping different to basic surveying and mapping. Cadastres should be able to reflect the current status of land tenure. They must be updated in a timely manner, because people and land relations are complex and dynamic.

Conventional cadastral survey methods are often argued as time consuming and labour intensive. Primary methods in use, for example, total station and GNSS survey usage take, on average 2.5 hours[1] for one land parcel to be mapped. As such, nationwide cadastral survey projects, which are responsible for mapping millions of parcels, would take decades, if not centuries (Zevenbergen & Augustinus, 2011). Sub-Sahara Africa area provides real illustrations (Simbizi, Bennett, & Zevenbergen, 2014). Researchers have found that land mapping in countries in this area will still require some 3 hundred years to complete (Enemark et al., 2014). Therefore, to accelerate the process of land survey and registration, innovative and automated methods are in high demand.

In addition, the continually increasing urban population leads to high-rises in land prices and shortage of urban land resources. Urban areas are expanding vertically, with complex construction structures. The right to a land parcel is not necessarily coherent with the right to space (de Vries, 2013). Overlapping and interlocking constructions in urban areas illustrate the complex modern relationship between humans and land. Traditional in-situ survey methods hardly handle mapping of these overhanging constructions, but they can be identified using observations from above.

After decades' development, cadastral surveys based on remote sensing data have gained recognition and increased popularity. High-resolution satellite imagery and laser scanning point clouds are two emerging approaches, although the latter is far less developed. In contrast to optical images, laser scanning is less affected by cloud and weather conditions. Consequently, time efficiency could be achieved through conducting such laser scanning surveys on a large scale. Canopy penetration is another apparent major advantage. Point clouds can detect features covered by vegetation, which cannot be seen from optical imagery. Thus, airborne laser scanning has become an increasingly popular tool for collecting vast amounts of accurate spatial data within a short period of time (Jochem, Höfle, Wichmann, Rutzinger, &

[1] Lecture Notes of Module 5: Cadastral Data Acquisition and 3D Cadastre, Land Administration Domain, ITC.

Zipf, 2012). Airborne Laser Scanner (ALS) can produce highly accurate 3D positioning information. The height information can be applied to distinguish vertically distributed constructions. Currently, existing methods for parcel boundaries generations are mainly manual. If cadastral objects can be extracted semi-automatically from light detection and ranging (LiDAR) data, much less manpower would be needed, and cost as well as operation time would be largely reduced.

Much research focuses on feature extraction from point cloud data. Examples include reconstruction of buildings (Vosselman & Dijkman, 2001), traffic furniture (Yu et al., 2015), trees (Monnier, Vallet, & Soheilian, 2012), amongst others. Perhaps most prominently, Van Beek (2015) designed a workflow to extract general boundaries from airborne laser scanning data of the Netherlands, and the results were satisfactory. However, in general, there is but limited research focused on semi-automated extraction of cadastral boundaries. In different areas, parcel boundaries might be marked by different kinds of physical physical objects. Roads, building walls, or water bodies are all may coincide with cadastral boundaries - particularly general boundaries. Aspects of features should therefore be clearly defined. A tailored strategy should be created to take comprehensive criteria into consideration. However, cadastral boundaries are fundamentally a human construct, and not all boundaries are visible. Likewise, not all detectable features coincide with cadastral boundaries. As a consequence, manual completion is also likely to be needed. The above challenges and opportunities provide the overarching motivation for the research.

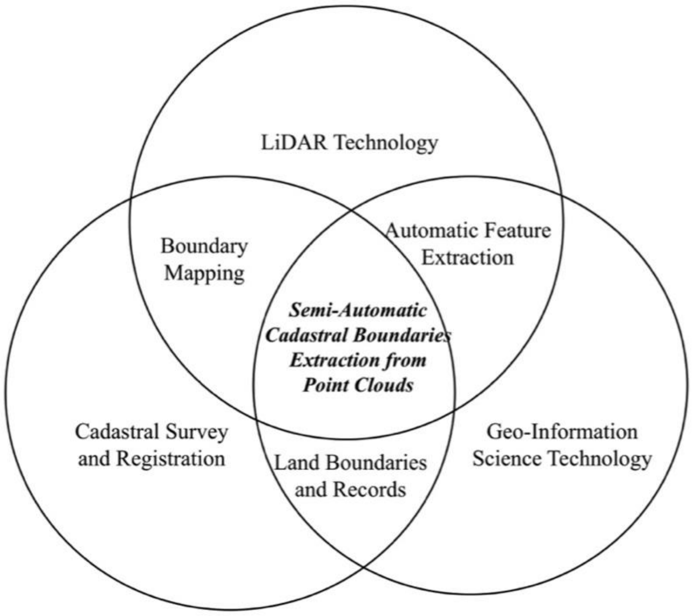
1.2. Research Problem

For some developing countries, to accomplish the full coverage of cadastre still requires decades, or even centuries (Zevenbergen, 2012). Although great progress on cadastral survey methods have been witnessed, due to the diversity morphology of cadastral boundaries, existing extraction methods are not efficient and effective for different contexts. There is a need for suitable methods, to speed up the procedure of cadastral survey. ALS data provides one potential alternative, however, to date there is only limited understanding on whether, how, where, and when the approach can be appropriately applied.

1.3. Research Objective

Derived from the research problem, the objective of this study is to develop a strategy for semi-automated extract cadastral boundaries from ALS point cloud data. Figure 1.1 which extended from the illustration of Battles (2016), describes the relationship between different research fields. Semi-automated extraction for cadastral boundaries from point clouds is an interdisciplinary that crosses over LiDAR techniques and geographic information system (GIS) as well as cadastral survey.

Figure 1.1 Conceptualizing the general objectives-(adapted from Battles, 2016)



In view of the general objective of developing a strategy for semi-automated extraction of cadastral objects from ALS point clouds data, considerations were taken into its feasibility and quality. Thus specific sub-objectives and relevant questions were designed.

1) **Objective 1: To study the cadastral boundary morphology in the study area**
   - What physical features also double as cadastral boundaries in the study area?
   - What is the probability of each kind of feature being a cadastral boundary?
   - What are the useful characteristics of these features?
   - What is the suitable point density?

2) **Objective 2: To design workflow for semi-automated extraction of cadastral features using laser scanning data**
   - What are targeted features of the workflow?
   - Which parameters are suitable for the classification in the workflow?
   - Which outline generation algorithms are suitable for planar objects?
   - How can linear objects be extracted?
   - What kind of post-refinement is needed and how should it be completed?

3) **Objective 3: To quantitatively evaluate the extraction workflow**
   - What is the error tolerance of quality?
   - What is the correctness and precision of the extracted boundaries?
   - What is the percentage of completeness?
   - What is the degree of automation?
   - What are the strengths and weaknesses of semi-automated extraction?
   - What are the possible recommendations for improvement?

1.4. Overarching Research Design

Base on the sub-objectives and research questions, the methodology of this research follows the general approach of: hypothesis-operate-evaluate procedure (Figure 1.2). The morphology of parcels in the study area was investigated in advance, to define the hypothesized target objects, and then a tailored workflow was developed to extract them to reconstruct a parcel map. The extracted result was accessed with reference data in the evaluation stage.

Figure 1.2 Overarching Research Design

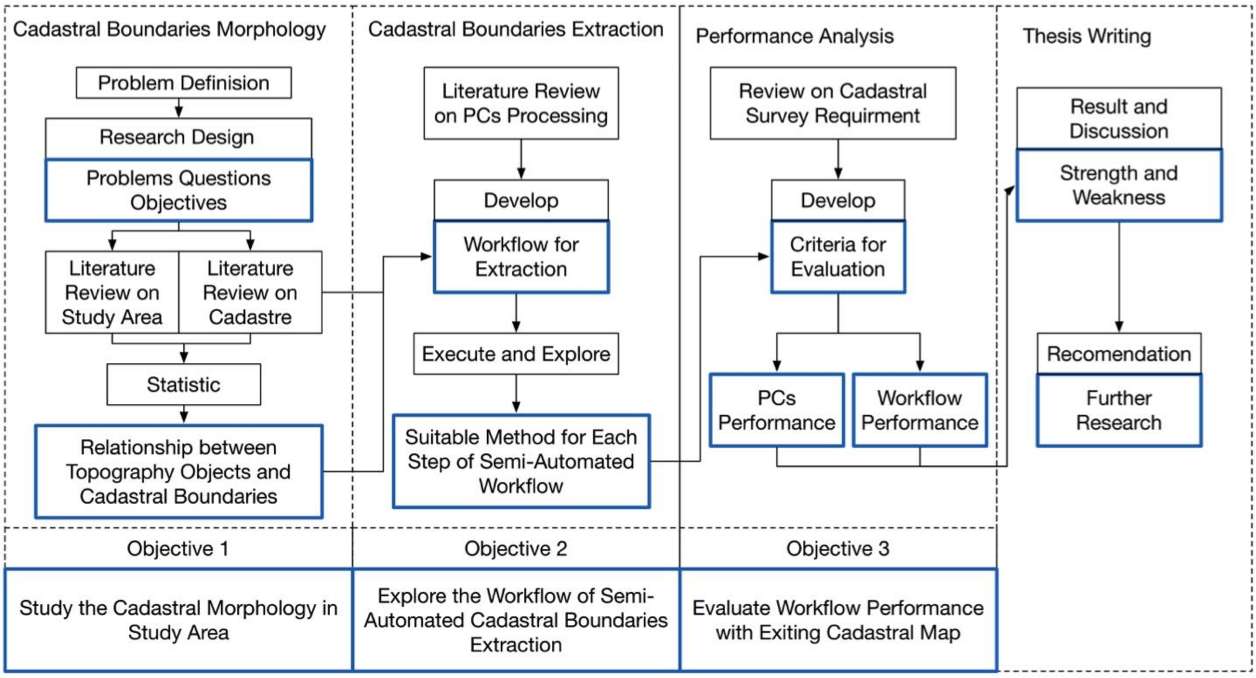


Figure 1.2 describes the research trajectory, with methods to achieve each sub-objective, as well as expected results. The details of the specific steps area provided in subsequent chapters, however, broadly speaking the included following activities:

Objective 1: Study the Morphology of Cadastral Boundaries in Vanuatu
The task of this section is to obtain specific knowledge of cadastral boundaries in the study area, Vanuatu. The morphology of land parcels is studied from the combination of the National cadastral map and an available orthophoto mosaic. A statistical analysis process aims at studying the probability of parcel boundaries coinciding with specific topographic features, they are considered as target objects. Expliciting targets can benefit the extraction process. Expected computation results consist of:

- Study regions selections and subsets;
- Percentage of extractable cadastral boundary segments;
- Priority of each object coinciding with cadastral boundaries; and
- Key parameters of target features.

Objective 2: Explore the Workflow of Semi-Automated Cadastral Boundary Extraction
According to the targets features and characteristic of them, a workflow is designed to extract these features, in order to support cadastral boundary reconstruction. Because of the diverse morphology of parcel boundaries, a further classification strategy is employed in this research. Several topographic objects are extracted in certain hierarchy from already classified point clouds. The outlines of these features are then generated to construct a rough map. Then post-refinement with visual interpretation improves the extracted result. The workflow mainly consists of the following steps:

- Further classify target features;
- Planar objects outline algorithms;
- Linear objects centerline fitting; and

- Refine the rough parcel map.

Objective 3: Evaluate the Workflow Performance with Existing Cadastral Map
The performance of the designed workflow is evaluated with the existing cadastral map. By comparing these two entities, a qualitative analysis is conducted to investigate the completeness, correctness and spatial accuracy of the extracted features. More specifically with regards to:

- Accuracy of the workflow – it is described by the tolerance of errors of the workflow;
- Correctness and precision of the workflow - they are illustrated by proportion of correct extracted lines segments.
- Capability of designed workflow - it is described by completeness of the workflow: the proportion of extracted full parcels and cadastral boundary segments.
- The degree of automation - it is described by comparing the total length of edited line segments with that of automated extracted lines.

1.5. Study Area and Data

For the purpose of the study, the island of Efate of Vanuatu, a Pacific island was utilized[2]. The country Vanuatu lies between latitudes 13° and 21°S and longitudes 166° and 171°E. Vanuatu is a Y-shaped archipelago consisting of about 82 relatively small, volcanic origin islands. There is a distance of about 1,300 kilometres between the most northern and southern islands (ThomasReuter, 2006). Vanuatu's increasing population (estimated in 2008 as growing 2.4% annually) is placing growing pressure on land and resources for agriculture, grazing, hunting, and fishing (Peace Corps, 2007). Some 90% of Vanuatu households fish and consume fish, which has caused intense fishing pressure near villages and the depletion of near-shore fish species. Figure 1.3 gives an overall view of the territory of Vanuatu.

Figure 1.3 Map of Vanuatu-(Ezilon, 2009)

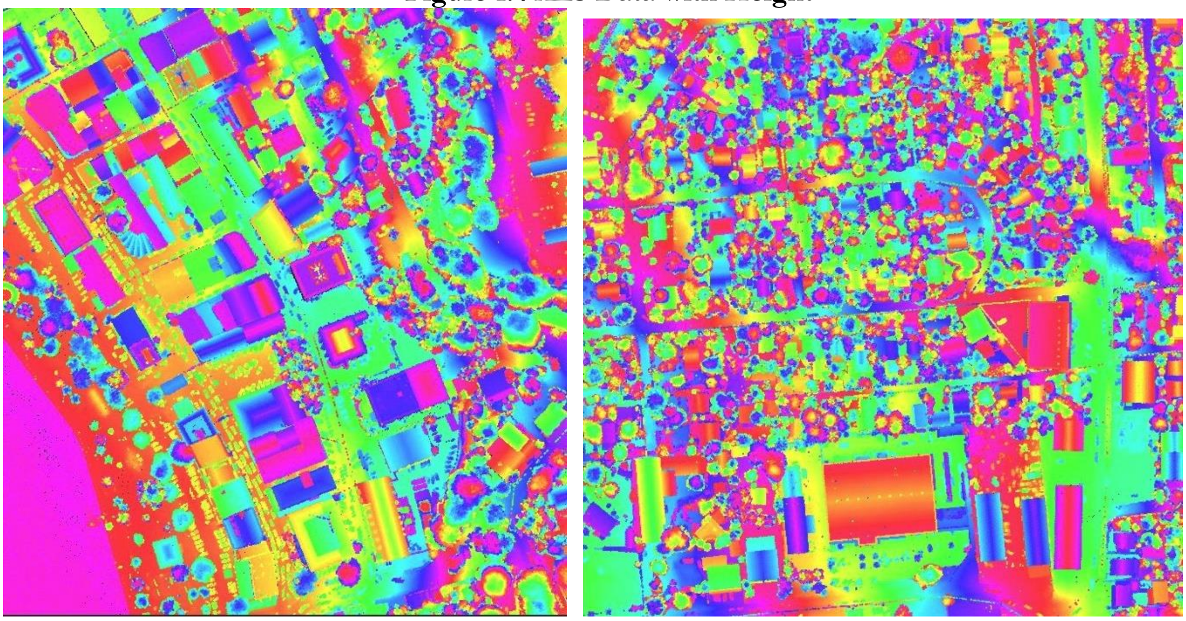
Vanuatu is an ideal case that has both cadastral and LiDAR data available. Moreover, the morphology in Vanuatu is generally different to the plain and regular NL land parcels; it arguable provides an ideal representation of a developing, if not urbanizing landscape, which is the target context of this study.

[2] Reasoning and justification is provided in Chapter 1.1.

The Cooperative Research Centre (CRCSI) of Australia provided the tested ALS LiDAR data, and Land Equity of Australia made the detailed cadastral map of Vanuatu available. However, access to both was dependent on the good will of the Government of Vanuatu Land Sector. The capital, Port Vila in Efata is covered by both data and is selected as the study area. The provided orthophoto was used as additional ground proving reference data. The coordinate system used in this study is UTM 59S, WGS 84.

Overall point density of the LiDAR data in Efate is 9.47 p/m². According to the data quality report (Quadros & Keysers, 2014), the topography LiDAR points are classified into 9 classes: unclassified, default, ground, low vegetation, medium vegetation, high vegetation, building, I/H point, and water. The classes label are coherent with LAS standards (ASPRS, 2009). Figure 1.4 shows the ALS data visualized according to the height of points. Roughly roof planes and road planes are visible because their height is constant. The changing background color illustrates the undulating topography.

Figure 1.4 ALS Data with Height



1.6. Thesis Structure

The thesis consists of 6 chapters.

Chapter 1: Introduces the background and objectives. It describes the research design, and defines the research problems, according to specific research objectives and related research questions. A general introduction to the methodology and study area is also presented in this chapter.

Chapter 2: Literature review on relevant fields, including: Cadastre - content of cadastral survey and cadastral map, cadastral boundary system types (Fixed and general boundary); Automated feature extraction - diverse approaches for extraction, both single-entity and multi-entity extraction, and introduction of relevant algorithms; and Geography and institutional context of the study area.

Chapter 3: Objective 1: Learning the morphology of cadastral boundaries in the study area, including statistics on topographic objects coinciding with cadastral boundaries; priority target objects; the parameter of these objects, and an investigation of suitable point density for cadastral mapping in the case area.

Chapter 4: Objective 2: Develop a workflow for semi-automated extraction of cadastral boundaries, including describe the steps and the methods used; results from executing processing and illustrating the performance of each step; and comparing the different methods applied in each step, in order to determine suitable methods.

Chapter 5: Objective 3: Quantitatively evaluate the workflow performance including analyzing errors and determining tolerances; and calculating the statistic on the correctness and the completeness of extracted line segments as well as parcels; and determines the degree of automation of the workflow.

Chapter 6: Discussion and conclusion, includes a reflection on results; a synthesis of answers relating to each research question.

2. CHAPTER 2 LITERATURE REVIEW

As described in Chapter 1, the semi-automated extraction for cadastral boundaries from point clouds is an emerging interdisciplinary research field. Accordingly, this review reflects upon several disciplinary areas including cadastral studies, feature extraction techniques and LiDAR data, and also a focus is given to the specific study area. The concept of cadastre is essential in this study, as the ultimate goal is to produce a cadastral map. Specifically, a cadastre is a comprehensive register, whilst cadastral survey is the way to acquire the cadastre data. Feature extraction is a means of deriving informative values from measured data. Observed from literature reviews, some extraction algorithms are considered useful for cadastral purpose. Each concept is now covered in more depth.

2.1. Advances in Cadastral Concepts

1) Defining Cadastre

A cadastre is a comprehensive record of the real property's boundary and ownership (Roger et al., 1992). It is an official record of a real property's ownership, value as well as its location and dimension (Dale & McLaughlin, 2000). Zevenbergen (2002) describes cadastres as systematically arranged public registers of property data, which depend on the survey of property boundaries (Zevenbergen, 2002). Promoted by FIG, the description of Cadastres emphases its parcel-based feature (FIG, 1995). Cadastres play important roles in different land registration systems worldwide, for either juridical or fiscal purposes. In some countries that implement title register, cadastres act like a conjunction with other records (Baigent, 1992). In a cadastral system, every parcel is constrained by a unique parcel identifier. It acts as an important role in the descriptive part of the cadastre, in that it enables national-wide, statewide, or jurisdiction-wide systematically management. A "Fit for purpose" concept is raised in view of the urgent need for a cost-effective and sustainable management system in less developed countries (Enemark et al., 2014). In a "fit for purpose" system, a cadastre record is flexible that focus more on tenure security rather than spatial accuracy (Bennett & Alemie, 2015). An alliance of "fit for purpose" land tools is developed to form the Global Land Tool Network (GLTN), in order to establish a full coverage cadastre and improve the land management as well as security of land tenure globally ("GLTN - Global Land Tool Network," 2016)

2) Cadastral Surveying and Cadastral Mapping

A cadastre usually consists of two parts: a geographic part that is often represented as maps or plans, and a descriptive part, called register or indexes (Zevenbergen, 2002). It is produced by a cadastral survey that results in a cadastral index map, and potentially other more detailed parcel descriptions. Cadastral Surveying is the sub-field of surveying that specializes in the establishment and re-establishment of the real property size, shape, and location (Robillard, Brown, & Wilson, 2003). It refers to the cadastre, or collective record of lands that many nations have established. Parcels and their owners that are presented on cadastral maps can be used as the basis for a land tax (Bruce, 1998), which illustrates the cadastre's can have both juridical and fiscal natures. Specifically, cadastral maps are graphic representation of the unit of lands that supply additional assurance. These depictions for configurations are made at large scale, in order to give a persuasive overview of the unit of land (Baigent, 1992). Furthermore, Dale (1988) stated that each parcel survey should be related to the adjoining properties. However, along with transaction, parcels will be split and amalgamated from time to time. Such changes in the boundaries of parcels have to be surveyed accordingly. Therefore, normally the content of cadastral map consist of parcels boundaries and land improvements, which are mainly composed of constructions.

The fundamental principle of cadastral surveying is to measure each property boundary point. Along with the development of the land survey technique, distance measurement and positioning were introduced into cadastral surveying. Over decades, if not centuries, many countries have established national coordinate systems, and this reference network is introduced to all surveying activities including cadastral survey. Absolute positioning offers confidence on parcels' location and area, contributing to tenure security.

In Vanuatu, the Ministry of Land and Natural Resource is responsible for cadastral information collecting and managing. However, since 2010, cadastral information has been digitized in a GIS environment, as part of the automating process of cadastral information management ("Vanuatu Ministry of Lands and Natural Resources - Land Survey," 2016).

3) Cadastral Boundaries

Theoretically, according to subject-object-right relationship in a cadastre, a boundary involves two parties related. As Dale & McLaughlin (1988) stated, "In a legal sense, a boundary is a vertical surface that defines where one land owner's territory ends and the next begins". Ideally this vertical surface extends from the deepest underground into the endlessness of space, but customarily it is recognized as how the boundary line lies on ground. From a traditional perspective, the cadastral boundary demarcates rights from ground surface and extends to infinity. In more complex urban environments, constructions are built on top of each other that increase the difficulties of boundary demarcation (Stoter, Ploeger, & van Oosterom, 2013). Thus in fact, the boundary that is visible in the terrain may not be coherent with the legal boundary (Zevenbergen, 2002). In the context of automatic feature extraction, there appears the opportunity to develop an innovative method to effectively and efficiently identify cadastral boundaries – in an automated or semi-automated fashion. Cadastral boundaries are either fixed or general. Arguments on fixed or general boundary never rest. Dale & McLaughlin (1988) gave each concept three descriptions.

A general boundary line between adjoining parcels is left undetermined (Tuladhar, 1996). It mainly consist of manmade or natural objects, with less emphasis placed on precision. Giving priority to land registration procedures, a gap with unspecified width or uncertain ownership could be left behind. Therefore fuzziness on cadastre are allowed to exist between parcels while still guaranteeing the title of each (Zevenbergen, 2002). The major advantages of such a general boundary include that less standard surveys are required, as well as the maintenance of the cadastral register is generally cheaper and convenient. It is particular useful in sporadic adjudication, because when demarcating consulting the owners of the adjoining properties is not compulsory (Zevenbergen, 2002). As general boundaries are often visible, they are more likely to be extracted from remote sensing data. In the context of this study, visible general boundaries are key focus.

A fixed boundary is a specific boundary that has been accurately surveyed (Dale & McLaughlin, 1988). Agreements are made at the time of adjudication of the land. Thus lines and corner points of fixed boundaries are fixed in space, their location cannot change without some documentation (Dale & McLaughlin, 1988). Fixed boundaries provide sufficient confidence for register (Bruce, 1998). In spite surveyors' measurements serves as importance evidence, a boundary is fixed or not depends on whether there has been a survey (Zevenbergen, 2002). In contrast to general boundaries, fixed boundaries supply parties with confidence on spatial extend of properties. In addition, in some areas most boundary lines are invisible so that precise measurements need to be done for retracement (Zevenbergen, 2002). In this context of this work, automatic extraction results can serve as base map for in situ adjudication, and further simplify survey procedures.

2.2. Developments in Feature Extraction

When analysing complex data, the major obstacle is the number of variables involved. Feature extraction is a general term used to broadly define methods of establishing combination of variables to address this problem while still describing the data with sufficient accuracy.

Kern (2006) describes automation of the feature extraction process as the "Holy Grail" of the photogrammetric data collection industry. Thanks to the rapid development of laser scanning, techniques for object classification and detection from point clouds have been applied to many user applications. An object-based approach would be more straightforward for cadastral boundaries extraction: by this means useful knowledge of objects could be applied to achieve the best results of detection.

1) Physical Objects Extraction from ALS

Much research has been done on automatic extraction of physical objects, some of which inspires thinking on application in the land administration domain. Overly, Bodum, & Kjems (2004) used the Hough Transform to reconstruct 3D buildings. The Hough Transform was applied to detected rough roof planes. Erroneous planes were filtered out using a Projected-Cluster-Test that determines true 3D planes. After merging detected planes, building models were reconstructed by generating meshes from planes. The created buildings had a success rate of 84%, but solutions for high jump lines and aligning ridges were still absent. For 3D cadastral purposes, more detailed results are expected.

A comprehensive approach for building extraction from ALS was proposed by Dorninger & Pfeifer (2008). They initialized segmentation by hierarchical clustering, with each point represented by its local regression plane. Then mean shift algorithm and region growing were applied in turn to define the roof plane. Polygon roof outlines generation was achieved by a 2D $\alpha - \text{shape}$ computation from building points. It forces line segments to become either orthogonal or parallel to each other, which conforms to the ground truth. The portions of complete and properly modeled buildings were about 75%. However heavy computation of the whole approach makes it unsuitable for wide-area cadastral survey.

Elberink, J, & Vosselman (2009) explored extracting complex road junctions information from point clouds. The points-in-polygon algorithm was a process applied for fusing point clouds with a topographic map. The surface-growing algorithm is the trigger in this process for determining road elements. After connecting road parts, they used the Hough Transform to determined local direction of roads. In this way, interlocking roads were separated in three dimensions. This method reflects thoughts on extraction of roads and urban overlapping constructions.

Much work that has been done in this domain can be linked to the application of cadastral boundary extraction, such as the hierarchical learning process for road marking extraction developed by Yu et al. (2015), and the method of classifying powerline scenes with Markov Random Field (MRF) proposed by Sohn, Jwa, & Kim (2012). In addition, Maltamo & Packalén (2007) investigated the capability of ALS for forest modelling; and Vosselman & Dijkman (2001) integrated ground plan to improve the quality for building reconstruction from point clouds.

2) Multi-Entity Extraction from ALS

The approach to classify various physical objects is a much focused research topic. Some methods make use of training set knowledge. The system for recognizing urban physical objects developed by Golovinskiy, Kim, & Funkhouser (2009) provides a good an illustration. Initially they extracted a set of features describing the shape and context of the objects. Afterwards they were used as a training set to classify objects. In this way, potential location of the objects can be obtained, and further segmentation accuracy could be increased. Location predictions were conducted before labelling objects, and its performance validated large-scale recognition. Therefore this strategy is inspiring for wide-area cadastral coverage. However, though the system was able to recognize 65% of small objects, many errors occur in spatial positioning of large objects such as buildings.

Other researchers divide the task into planar and non-planar objects. Vosselman (2013) explored a combinative process of different segmentation and post-processing methods, aiming at improving detection of various urban objects. He used the 3D Hough transform to determine the surface, and subsequently applied a segment-growing algorithm to address non-planar components. Finally the planar segments were merged and outliners were filtered. This method requires less manual pre-processing while still employing specific knowledge of object classes to enhance classification quality. Connected component analysis that applied to merge fragmental surfaces contributes to large plane reconstruction and vegetation elements grouping.

More specifically, Xu, Vosselman, & Oude Elberink (2014) proposed a workflow consist of four steps on multi-entity classification. They firstly obtained rough classification of planar segments (ground, water, vegetation, roof and unclassified objects) by surface growing algorithms. Then they further labelled unclassified points with point-wise classifiers. Finally, mean shift segmentation was applied to reclassify roof elements. The final results achieved an overall classification accuracy of 97%, 90% completeness and a correctness of 85%; the results are highly encouraging for the cadastral domain. The surface-growing algorithm is more suitable for cadastral plane detection, because it computes locally and tolerates certain amount of roughness, which makes it adaptive for complex cadastral contexts.

In summary, diverse approaches are proposed to achieve automated feature extraction. Algorithms for each step also vary. Due to the complex shapes of cadastral boundaries, it appears the object-based strategy can achieve higher accuracy. Specifically, more attention should be paid to planar surface extraction because of the nature of the parcels. Therefore, planar detection and objects classification are considered two essential in the steps to determination of cadastral objects.

2.3. A Primer on Algorithms

From the literature review on feature extraction, some algorithms and methods are considered suitable for cadastral purposes.

1) Segment Points with Connected Component Analysis

In the planned research, when extracting road points, the segmentation step is to be achieved by using the connected component algorithm. It was selected because it is widely used and easy to operates function (Rubio, Lenskiy, & Ryu, 2013). Hopcroft and Tarjan (1973) illustrate essentially this algorithm, with the statement that at that point it was "well known". Originally developed from the image-processing field, the algorithm computes the connected components in linear time, in the search manner of either breadth-first or depth-first. No matter the selected search approach, the search begins at some particular vertex γ, loop until it found all connected component that containing γ (Hopcroft & Tarjan, 1973), according to the numbers of the vertices and edges of the graph. Then it begins another a new loop.

The algorithm has been extended for segmentation of LiDAR point clouds, though calculating distances between consecutive points. Kd-tree is used to structure the returned LiDAR Data. The whole process can be summarized into five steps. Firstly, the points are converted to a Cartesian coordinate system, before being projected onto an occupancy grid. The grid imitates a binary image and is then used for segmentation (Rubio et al., 2013). Finally, the algorithm searches for connected components over the grid and return segments. The whole process is robust and fast, which makes it suitable for the large amounts of data associated with cadastral surveys.

2) Centreline Detection by Skeleton

In the context of this work, the skeleton algorithm specifically refers to topological skeleton. It has been widely used in computer vision and image processing. The algorithm computes a thin version of a shape, presented as a skeleton that is equidistant to its boundaries. It emphasizes geometrical and topological properties of the shape, for instance its length, direction, and width (Bai, Latecki, & Liu, 2007). Points of a skeleton and its distance to the boundary serve as representation of the shape.

A medial axis is the major component to define a skeleton. It use an intuitive model of fire breed on a grass field that possesses the given shape to compute the skeleton (Abeysinghe & Baker, 2008). Once a "set fire" at all points along the boundary in the grass field is triggered, the skeleton appears as a connection of these points. Planar sets are determined in advance to assume a connected bounded subset. The subset is composed of a finite number of disjoint curves. These curves should be closed and simple, and each consists of a certain amount of analytic curves (Bai et al., 2007). The skeleton algorithm does not introduce any restriction on object contours, but regards the images as polygonal curves with vertices. These vertices make up the skeleton pixels.

3) Straight Line Detection by Hough Transform

The Hough Transform is a famous feature extraction technique. The principle of this algorithm is to vote procedure over parameterized objects, to perform edge points grouping into object candidates (Hough, 1962).

Straight lines detection is the simplest case of Hough transforming. A straight line can be represented by general equation $y = mx + b$, with a point (b, m) falling in the parameter space. During computing, slope parameter is given a rise, thus a Hessen normal form is applied to describe it.

$$r = x \cos\theta + y \sin\theta$$

Therefore, each line is associated with an image pair (r, θ), which referred to as Hough space for a set of straight lines. An accumulator, a two-dimensional array, detects the line described by this normal form. The accumulator's dimension is two, equals to the number of unknown parameters (r, θ). Then the Hough transform algorithm searches for the highest value that is the local maxima, then the most likely lines can be determined (Ballard, 1981). A two-dimensional matrix stores the final result of a linear Hough transform. The number of curves through a point is represented as a cell value. In Hough space, extreme bright points are points that have the most votes. They form the Hough parameters of lines.

4) Outline Generation by Alpha Shape Algorithm

In the context of this work, after some form of initial classification, clusters of points represented detected building roofs and different land parcels. In order to obtain the line – or vector based - based cadastral map, the outlines of each cluster could be generated with Alpha shape (α-shapes) algorithm. α-shape algorithm was originally introduced by Edelsbrunner, Kirkpatrick, & Seidel (1983). It begins with the approach to formalize the intuitive notion of "shape" for spatial point sets. After that, the convex hull of a point set is generalized, followed by deriving a family of shapes from the Delaunay triangulation parameterized by α. Indeed, a α-shape is demarcated by a frontier, which is a linear approximation of the original shape (Wei, 2008).

Given a dense unorganized set of data points S, the set S of points has a α-shape in polygon. The parameter α controls the precision of the boundary. Imagine that a circle with a radius α is rolling around the point-set S. If there is no other point falling within the circle, the point is defined as boundary points automatically, and these points are connected to obtain one boundary line. When the α value is approaching infinity ($\alpha \rightarrow \infty$), α-shape would be the convex hull (Amenta, 2015). On the other hand, when the α value is very small ($\alpha \rightarrow 0$), every point might be the boundary. For any α, the α-shape is a sub-graph of the Delaunay triangulation.

Though the α-shape concept has been expanded into 3D space lately, for cadastral purpose, 2D is satisfactory. In this case, point-set S containing evenly distributed points, they are projected onto ground plane. An optimal value was determined for α approaching, and then detected outline points are connected with α-shape. Therefore the α-shape can extract the inner and outer outlines of the polygon at the same time

5) Outline Generation by Canny Edge Detector

Originally developed by Canny (1986), the Canny algorithm in principle is an edge detector that works in images. Being widely applied in the computer vision field, Canny can detect edges of an image with much less error. An edge in an image would be recognized only once, which make it a noise robust operator.

Firstly, a Gaussian filter is used to remove the noise of an image to prevent false detection. This step significantly smooths the image, and the kernel size of Gaussian filter will affect the performance of detector. Subsequently, four filters are applied to detect horizontal, vertical and diagonal edges respectively in the blurred image. The result of this step is returned as a number. Thirdly, an edge thinning technique called non-maximum suppression is applied to sharpen the detected edges while excluding local maxima. After doubling the threshold, the strong edge is tracked by hysteresis, in which way edges extracted from noises are removed. During this final step, 8-connected neighbourhood pixels of an edge are searched for by a Binary Large Object (BLOB) analysis (Lindeberg, 1998).

The adjustable parameters of Canny algorithm are the kernel size of Gaussian filter and the threshold of hysteresis (Deriche, 1987). The process of Canny detection is straightforward and computes in a short, fixed amount of time, which makes it a flexible detector that can adapt to diverse environment, and it therefore appears suitable for a wide range of cadastral purposes.

2.4. Contextualizing Vanuatu

Nowadays two land tenure systems run in Vanuatu: a formal deeds system inherited from the colonial period, and an uncodified customary system. Different cultural backgrounds relating to the two tenure systems cultivate conflicts in perception and conceptions of tenure rights (Nari, 2000). Approximately 97 percent of land in Vanuatu is customary land, and this customary tenure system plays a dominate role is all activities relating to land (de Burlo, 1989). Vague to outsiders, the behaviour of the customary system is significantly diverse among communities over how land is transferred, subdivide or inherited (Interaction, Modern, Tenure, & In, 2011). Customary rules are given the privilege on ownership and use of land. Land transactions are being approved or opposed by members of the communities living adjacent to, or even on a top of customary land. Arguably, the lack of security provided via written regulation and a concrete register means, land conflicts happen frequently on custom land. Up until recently, registering land in a formal system was still a voluntary action (Nari, 2000). Additionally, diverse rights to land are not equally distributed amongst members of a group (Holt, Sullivan, & Weaver, 2004). Therefore, individual ownership is always unknown. During the independence period, the land issue was the political cornerstone amongst participating parties. And the Vanuatu Land Program, a long-term commitment by the Government of Vanuatu Land Sector Framework is carrying out to implement the land sector reforms (Land Equity International ty Ltd, 2016).

Additionally, Vanuatu is facing increasing urban migration, especially to Port Vila. The past twenty years have witnessed an increasing tendency in land disputes. People leave outer islands, either temporarily or permanently, to live and work on Efate (Farran, 2007). The city can hardly accommodate the flooded population, and consequently in widespread urban areas squatter settlements and informal license arrangement emerge. Disordered land use contributes to vague tenure rights. In Vanuatu, the cadastral information is used to support public registering, to facilitate customary land and state land subdivision and lease.

Under the formal system, the Vanuatu Ministry of Land and Natural Resource is responsible for all land related activities. In view of the high proportion of customary land, the ministry pays particular attention to customary land leasing, which encourages landowners to liberate their land for activating the land market. Since customary land cannot be alienated, the most common way to circulate land is leasing, which is allowed for up to 75 years. Leasing is common in urban areas of Port Vila, and is on the increase in the urban and coastal areas. Real time records of customary land are not only remained important but are becoming increasingly significant (de Burlo, 1989).

Historically, claims to land and rights to land are confirmed by oral records and the statement of kinship links. Small portions of land are accurately surveyed, and boundary markers tend to be physical features of the land such as trees, stones and streams (Regenvanu, 2008). These markers are unstable: this also impacts on the affect the security of the register. Hence a faster cadastral survey method to monitor the changes could help in the maintenance of land records.

In view of the urgent demand for a full coverage cadastral map, Vanuatu has accelerated its cadastral survey procedure (Regenvanu, 2008). However, in urban areas, the undulating topography leads to irregular landscape. Innovative survey methods are therefore needed to speed up the procedure.

2.5. Summary of Literature Review

Summarized from the literature review, accelerating the cadastral survey procedure is of essential importance for approaching a full coverage cadastre, and further secure land tenure. The feature extraction technique of LiDAR data could be used to develop an effective and efficient cadastral survey method. Therefore, a semi-automated workflow: a combination of automatic feature extraction technique from LiDAR data and manual completions may be suitable for the cadastral purpose and at least demand exploration.

3. CHAPTER 3 THE MORPHOLOGY OF CADASTRAL BOUNDARIES IN VANUATU

Until this point, the planned research, including methodology, and necessary background justification have been provided. This chapter presents results from the first objectives – that relating to the morphology of cadastral boundaries in Vanuatu. Specifically, the results of the investigation of the relationship between parcel boundaries and topographic objects are provided. Going further, this Chapter then describes how the feature extraction approach can contribute to cadastral mapping.

Figure 3.1 An urban area of Port Vila in Vanuatu

As shown in Figure 3.1, in Port Vila, intense constructions splits the land into small plots of parcels. This circumstance not only increases the workload for ground-based cadastral surveying, but also land boundaries extracted from remote sensing data. Figure 3.2 illustrates that in suburban areas parcels boundaries are marked by thin fences but they are mixed in vegetation and are hardly seen from images.

Figure 3.2 Illustration of Parcel

3.1. Study Region Subsets

In order to investigate different parcel boundary morphologies in the dense urban area and the suburban area, two regions in Port Vila were selected as the subset for further statistic analysis (Figure 3.3). The subset splitting was based on the hypothesis that major difference of two regions lies in the fact that there is less vegetation and small fences in the dense urban area. From the orthophoto, these two regions are clear, with diverse land covers clearly distinguishable. They represent typical dense urban and suburban areas respectively. Additionally, the research focus on general boundaries in the urban area, and the selected test regions were both covered by the reference data. Separating these two regions helped in exploring the performance of the workflow on the diverse landscape of the urban area.

Upon overlaying the cadastral map with the orthophoto, it was noticed that in the urban area, the parcel boundaries mainly coincide with roads and building outlines. Other small segments of parcel boundaries falls in low vegetation thus they are hardly visible. In the urban region, buildings are close to each other. Constructions and roads mainly occupy the land. It is difficult to define the parcel range from images. By contrast, in the suburban region, buildings are sporadically distributed, with grassland and vegetation surrounding them. Within a block, parcel boundaries cut through the middle of two buildings. In many cases one parcel contains one building. As an exploratory study, the decision to split the region into these two subsets (urban and suburban) also helps in shortening the processing time, enabling the research to focus on the relationship of physical objects and cadastral boundaries.

Figure 3.3 Overview of Efate in Vanuatu

3.2. Physical Objects that Coincide with Cadastral Boundaries

To study the relationship between physical objects and cadastral boundaries, the number of parcel boundary segments coinciding with each kind of objects were counted by visual interpretation. This step focused on defining the visible boundaries that can be detected from remote sensing data. Table 3.1 illustrates the statistical result of the target objects. Table 3.1 shows that in the two study regions (region 1: dense urban; region 2: suburban), cadastral boundary segments coincide with similar kinds of physical objects but in different priorities. By counting the parcel boundary segments, it was studied that in these two regions, a total of 797 boundary segments were identified, of these, 82.4 % of them coincide with visible physical objects. The remaining segments (17.6%) are invisible from the satellite images. Specifically, road edges coincide with the most percentage of parcel boundary segments in both regions, with 46.2% and 49.1% respectively. They clearly separate blocks of construction clusters. In region 1, building outlines play the second most significant role (35.1%) while it drops to only 4.2% in region 2. In region 2, fences serve as monuments and indicate one third of visible parcel boundaries. The computation results as well as the priority of objects are as followed Table 3.1.

Table 3.1 Statistics of Cadastral Boundaries with Physical Objects

	Region 1		Region 2		Total	
Total number of cadastral boundaries	396		401		**797**	
Roads Edges	183	46,21%	197	49,13%	380	**47,68%**
Fences	2	0,51%	102	25,44%	104	**13,05%**
Building Outlines	139	35,10%	17	4,24%	156	**19,57%**
Vegetation	12	3,03%	5	1,25%	17	**2,13%**
Invisible	60	15,15%	80	19,95%	140	**17,57%**

Characteristics of these physical objects were studied, in order to derive a suitable extraction approach from the LiDAR point clouds. On-land objects such as construction and vegetation are certainly above ground, while roads normally lie on the ground. Therefore, height from ground was given the priority for rough classification. In terms of buildings and vegetation, local roughness and numbers of return were two major constrains. From the nadir view, building roofs present as regular planes, whilst vegetation represents as a group of messy points. As for small objects like parapets and fences, they were mixed into low vegetation, but the intensity of constructions is different with vegetation. On the other hand, from ground points, it was comparatively difficult to distinguish roads, bare lands and car parks. Since height difference is no longer effective, intensity and smoothness were selected to identify them. Roads and carparks are made from cement or pitch, while bare lands are soil and grass. Table 3.2 illustrates the observed characteristics for describing each target objects.

Table 3.2 Characteristic of Target Objects

Object Types	Criteria	Output
Road	On ground; plain; uniform material; linear shape;	Boundary
Building	Off ground; plain; height steady; minimum 2m above; rectangular shape; local smooth	Boundary
Vegetation	Off ground; rugged; height jump; cluster shape;	Remove
Parapet & Fence	Off ground; rugged; height steady; narrow; minimum 1m above; linear shape;	Centreline

3.3. Suitable Point Density for Target Objects

After defining the target objects, the Equation 3.3 determined a suitable point cloud, to inspect the capability of study data. The smallest dimension of a particular object in a point cloud was used to determine the minimum point cloud density for detecting that particular object. When determining the smallest dimension of a target object. From the nadir view, only width and length of object are considered. The point spacing should be smaller than the smallest object dimension.

$$p = \frac{1}{(smallest\ object\ dimension/2)^2}, \qquad 3.3$$

In this study, the thinnest target object is the fence. Normally the width of the fence is 0.5m, thus the smallest point cloud density for detecting the fence is 16 points per square meter (p/m²). However, the point density of the study data is only 9 p/m²: the smallest detectable dimension is 1.3 m, smaller objects like the fence may be undetectable in this study, while the others are larger than the minima.

3.4. Summary of Objective 1

The result of objective 1 shows that in the dense urban area and the suburban area of Port Vila, parcel boundary segments coincide with different objects. The extraction workflow should give privilege to road edges, building outlines and fences: they have the most prominent to double with parcel boundary segments. Some object characteristics could be applied when classifying points: uniform material may be useful to recognize points of roads, whilst the height of fences may differ them from other objects. The capability of study data should be taken into consideration: the smallest detectable dimension is 6 meter, which cannot meet the accuracy requirement of cadastral survey.

4. CHAPTER 4 THE SEMI-AUTOMATED WORKFLOW OF CADASTRAL BOUNDARIES EXTRACTION

Based on the cadastral boundary morphology of Port Vila, it was determined that the workflow would focus on identifying and extraction roads, buildings and fences; they have the most relevance with regards to coincidence with parcel boundaries. This Chapter reports results on the procedure used to develop the workflow and the outputs stemming from it.

4.1. Overview of Semi-Automated Workflow

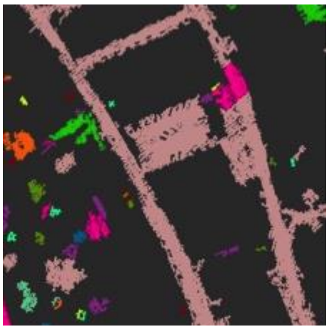
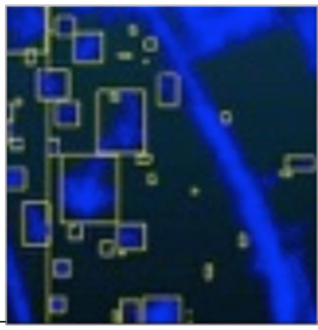
The developed workflow consists of two phases: automated extraction and post-refinement. Three steps made up the automatic extraction phase:

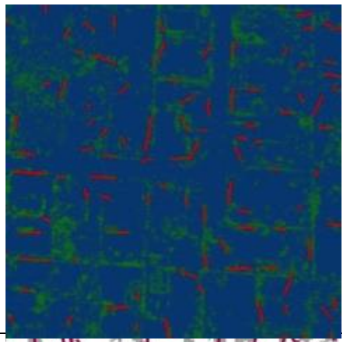
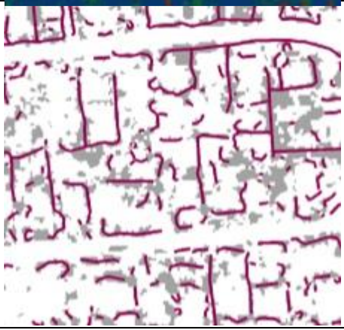
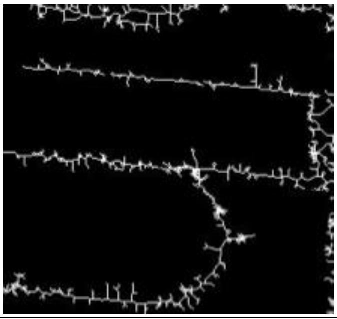
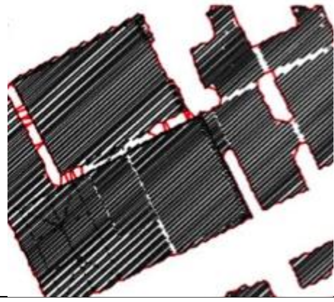
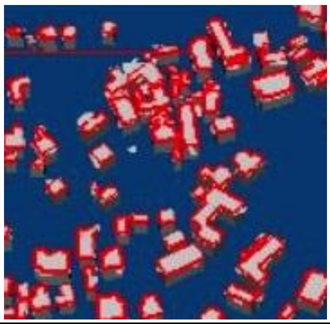
1) further classify points to target objects; 2) generate planar object outlines; 3) fit centreline to linear objects.

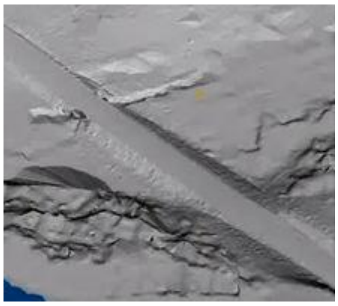
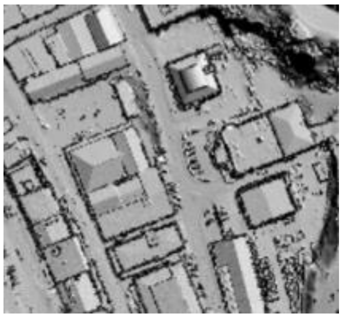
Specific approaches were selected to conduct each step. In view of the complex morphology of cadastral boundaries, these approaches were able to deal with the large amount of datasets as well as with particular targets of the cadastral objects. Afterwards, extracted line segments were edited and completed in the post-refinement phase.

Diverse types of software were tested in different steps. Diverse software was tested in order to fine an effective and efficient approach. Comparing the performance and efficient of each piece of software, with time and budget considered (Table 4.1), the most suitable software was determined for each step, and they are highlighted with underline.

Table 4.1 Comparison of Software

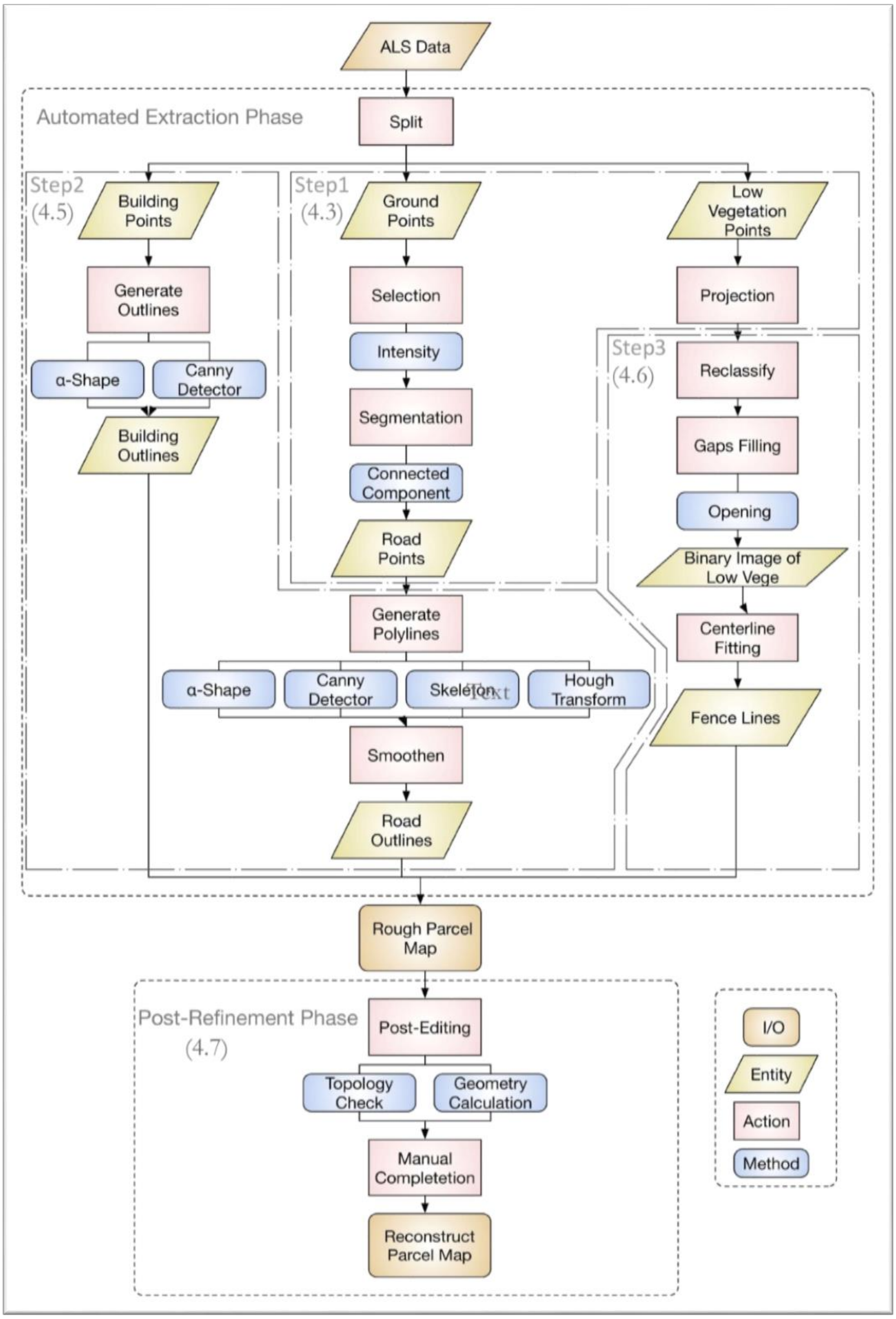
Purpose	Software		Sample Output
Segmentation	*PointCloudMapper*		
	Cost	Campus License	
	Function	Segmentation	
	Process Time	Medium	
	Data Type	.las	
	CloudCompare		
	Cost	Open source	
	Function	Segmentation	
	Process Time	Fast	
	Data Type	.las	

Line Generation	*VRmesh*		
	Cost	Commercial	
	Function	Point-tracing	
	Process Time	Medium	
	Data Type	.shp	
	ArcScan		
	Cost	Campus License	
	Function	Vectorize	
	Process Time	Medium	
	Data Type	.shp	
	Matlab		
	Cost	Campus License	
	Function	Skeleton	
	Process Time	Fast	
	Data Type	.svg	
Outline Delineation	*Matlab*		
	Cost	Campus License	
	Function	Outline	
	Process Time	Medium	
	Data Type	.svg	
	VRmesh		
	Cost	Commercial	
	Function	Outline	
	Process Time	Fast	
	Data Type	.shp	

<table>
<tr><td rowspan="10">Terrain Visualization</td><td colspan="2">VRmesh</td><td rowspan="5"></td></tr>
<tr><td>Cost</td><td>Commercial</td></tr>
<tr><td>Function</td><td>Meshed Surface</td></tr>
<tr><td>Process Time</td><td>Medium</td></tr>
<tr><td>Data Type</td><td>.tif</td></tr>
<tr><td colspan="2"><u>LasTools</u></td><td rowspan="5"></td></tr>
<tr><td>Cost</td><td>Free License</td></tr>
<tr><td>Function</td><td>Hillshade</td></tr>
<tr><td>Process Time</td><td>Fast</td></tr>
<tr><td>Data Type</td><td>.tif</td></tr>
</table>

Specifically, Matlab executed outline generation algorithms because the tested algorithms were implemented in it; ArcScan was selected to achieved centreline fitting for its outstanding vectorizing function; LasTools was used to produced the hillshade images because it is the fastest solution for LiDAR data processing; CloudCompare, an open source and efficient software, were selected to conducted the segmentation and points filtering; and the well-known ArcGIS was used for output visualization and post-refinement. An overview of the extraction framework is given in Figure 4.1. It illustrates the designed steps and algorithms applied.

Figure 4.1 Extraction Frameworks



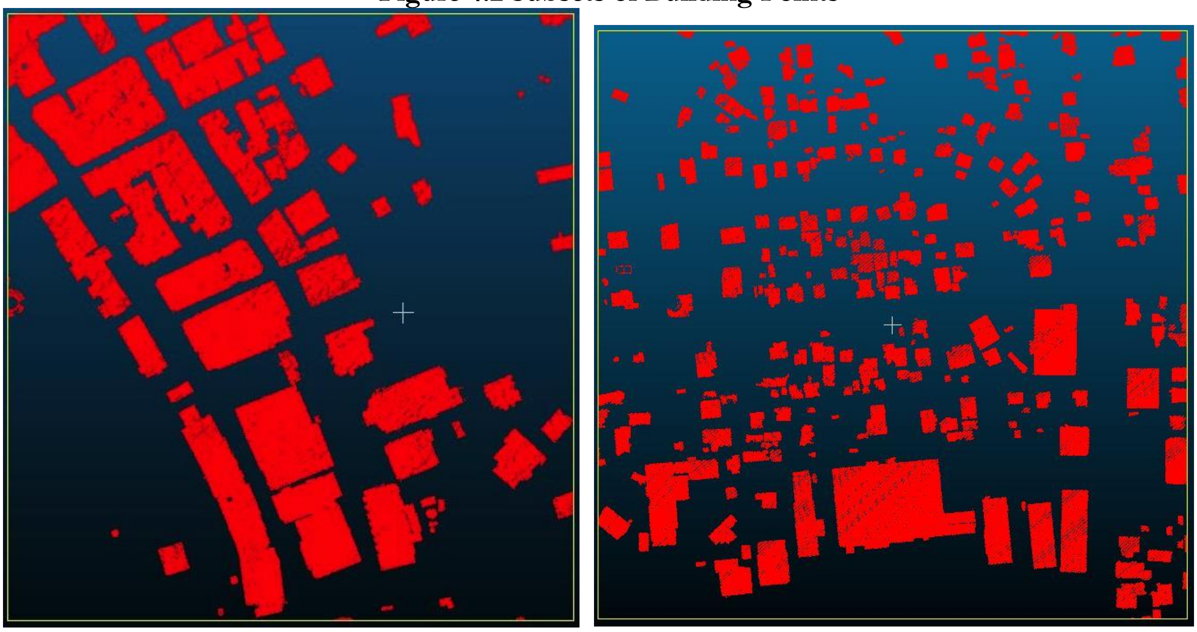
4.2. Preparation of Classified Point Clouds

Before conducting each step, these labelled points were separated into semantic subsets. The study data has been classified into 9 classes, which including the building class and the vegetation class[3], whereas

[3] Metadata of the study data is introduced in Chapter 1.5.

roads and fences have to be recognized from these original classes by the further classification step. One subset of building points is displayed as Figure 4.2. The following sub-Chapters describe the process and result of each step.

Figure 4.2 Subsets of Building Points



4.3. Further Classification for Expected Objects

The very first step of the workflow is to further classify points into target components, which are the road and the fence in this study. Difference methods were applied to recognize these two objects.

4.3.1. Road Detection from Ground Points

Illustrated in the cadastral situation analysis of Vanuatu, roads are very likely coincides with cadastral boundaries. As described in section 3.1, roads lie at ground level; that they cannot be separated from height difference. Inspired by work of Clode, Kootsookos, and Rottensteiner (2004), in spite of the noisy value of intensity returned by the scanning unit, road material is usually uniform along a road section. Therefore, points were then selected when their last pulse intensity values fell in the acceptable range for this type of road material. By searching for a particular intensity range (defined be equation 4.3), it is possible to extract most LiDAR points that were on roads, even though there were also some other on-road detections that were also produced. Equation 4.2 illustrates how the LiDAR points were filtered based on their intensity, in order to create a new subset of points (Clode et al., 2004).

$$S_1 = \{P_i \in S: i_{min} < P_i < i_{max}\}, \qquad 4.2$$

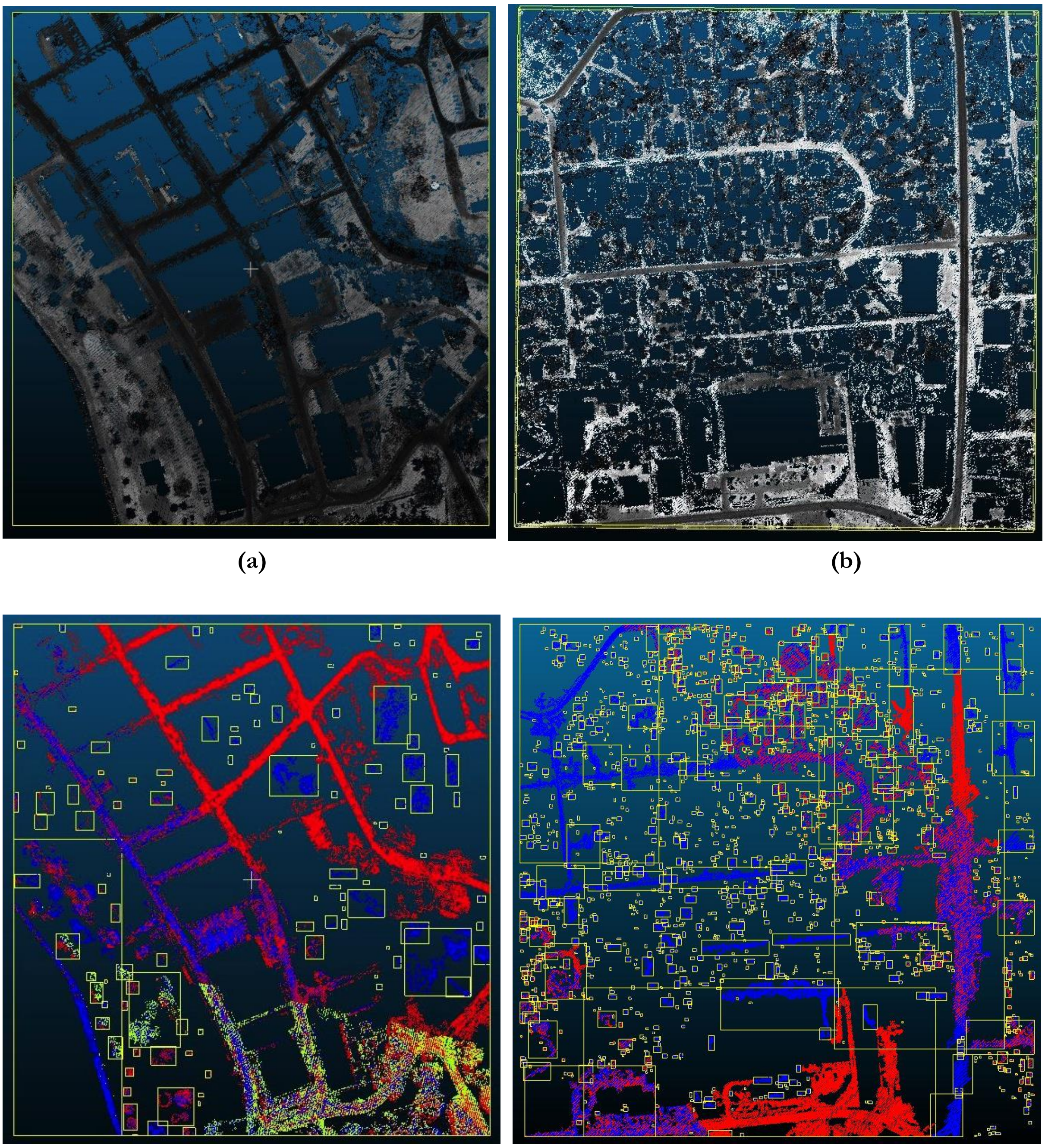
where and i_{min} and i_{max} are the minimum and maximum acceptable LiDAR intensities at point P_i.

By visual interpretation, the intensity of road points in two study regions was similar. It illustrates that in Port Vila, the material used in road constructions is uniform. The selected range was 0-25 and 30-55. The followed equation describes the road points subset S_2.

$$S_2 = \{P_i \in S: 0 < P_i < 25 || 30 < P_i < 55\}, \qquad 4.3$$

Since roads are flat linear networks, they were assumed to be connected planes. After the points were selected, they were segmented by connected component analysis, based on planar distance among points. During the connected component segmentation, octree level was determined as 10 in both regions[4]. Segments size were computed, then small segments were defined as unvalued physical objects and removed. The remaining points were points on roads. Figure 4.3 (**a** – **f**) shows how a rough road network was extracted in two regions respectively. Specifically, **a** and **b** present the intensity filtering result, while **c** and **d** illustrate the segmentation process, and **e** as well as **f** show the road extraction results of two region.

Figure 4.3 Road Points Extraction



(a) **(b)**

(c) **(d)**

[4] Comparison of different level of octree is illustrated in Appendix 2.

(e)

(f)

The result (**e**, **f**) shows that some portion of roads were wrongly deducted, while some roadside bare lands still remained. Figure **f** (region 1) contains more irregular physical objects, while Figure **e** (region 2) presents a clearer linear road structure. This might be caused by more car parks and bare land in developed region1, compared to region 2 where was covered by vegetation. Gaps exist in both regions, because of the wrong removal of small segments. Uneven points distribute on road surface consequent in this incorrect segmentation.

4.3.2. Fence Detection for Low Vegetation

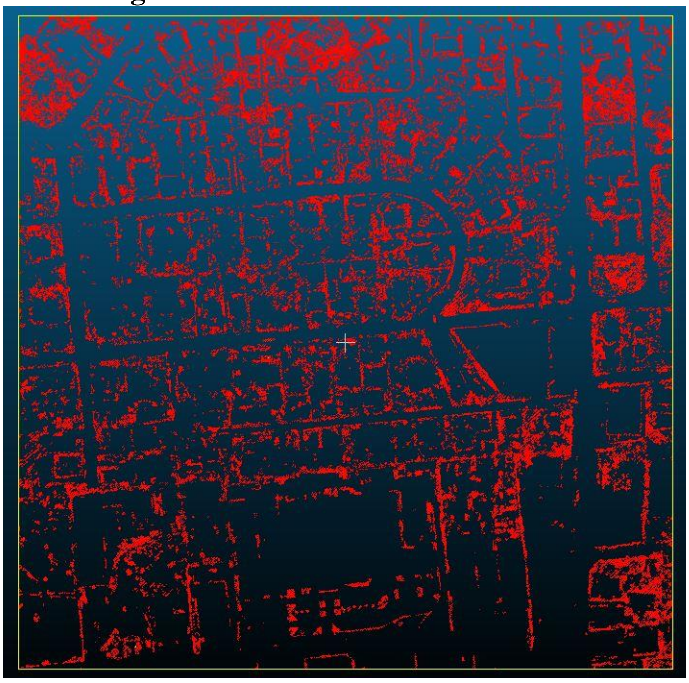
In region 2, fences coincide the most with cadastral boundaries, however, they are hard to distinguished from low vegetation: their heights were very similar. Furthermore, they cannot be computed from local smoothness, as they are too narrow to form planes. The average thickness of fences in the study area is about 0.5m, which is smaller than the smallest detectable object (1.3 m)[5] of the study data point density.

Besides, the material of the fence was uncertain: they may be made up by bush, concrete or wood. In addition, the LiDAR data only provided measurements in near inferred band: the spectral information provided by study data was insufficient for recognizing fences. Except from reflectance intensity, an extra criteria - local point distance was computed, and the threshold was set to 1 pixel and the standard deviation was computed of 6 neighbourhood points, afterwards sporadic points were removed. However, as shown in Figure 4.4, the detection performance was modest still.

From the very general view on region 2, a rough impression of parcels can be identified. Some building edge points remain as they were misclassified. And in denser vegetation cover areas, parcels cannot be distinguished.

[5] Calculated and Discussed in Chapter 3.3.

Figure 4.4 Fence Points Extraction



4.4. Complemented Knowledge from Height Jump

A complementary procedure was integrated to delineate the outline of objects from height information. This supplementary process provided more knowledge on parcels boundaries, in cases where the topography relief result in incorrect detection. Hillshade images were applied to visualize the height difference in this step. LAStools was applied to conduct this process. The Blast2dem function of LAStools triangulated ground points into a seamless Triangulated Irregular Network (TIN), and then rasterize the TIN on a digital elevation model (DEM).

Figure 4.5 Hillshade Images of Two Regions

As illustrated as Figure 4.5, in region 1, building outlines were highlighted. Specifically, very closed building roofs were distinguished as they were at different heights. The topography relief was also visible, which aids in separating road plane with roadside slopes. On the contrary, in region 2, whose topography is a flatter, hillshade visualization did not add much value. Even though building roofs were also highlighted, dense low vegetation was also mixed inside and that ended up as noise.

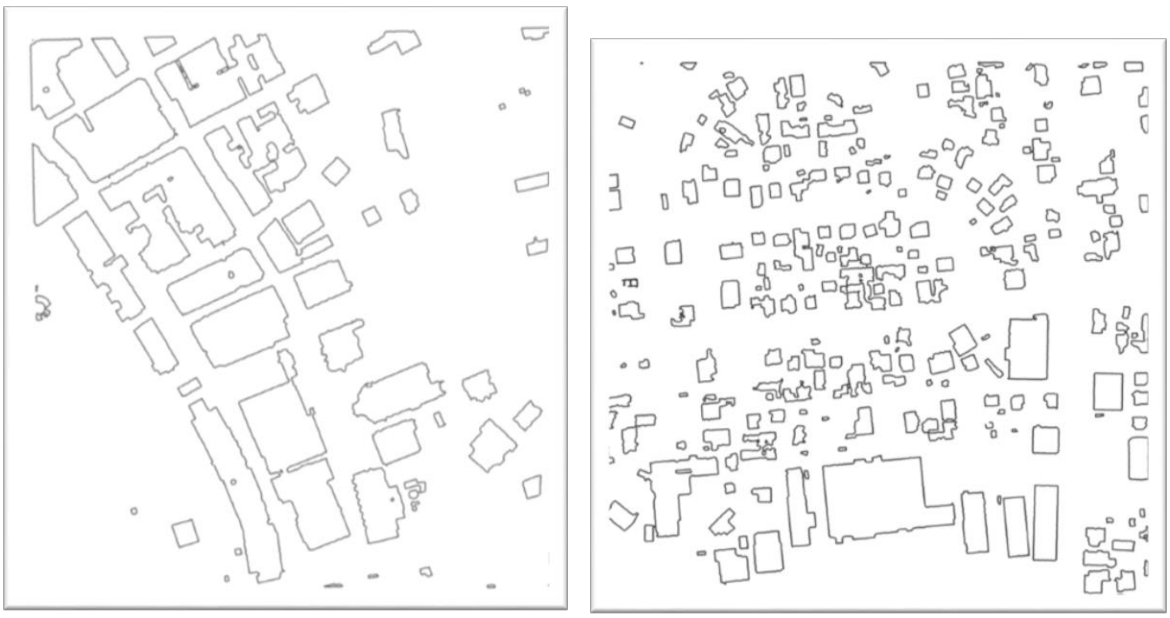
4.5. Outline Generation of Detected Planar Objects

After the recognition of roads and fences, points are classified into target objects; the second step of automatic extraction phase is to derive outlines from planar objects. According to the nature of objects, these physical objects are either planar or linear. The boundary approach was applied in planar objects such as building roofs, while the line fitting approach was more suitable for linear objects such as fences. However, roads lie between these two types of physical objects, and therefore both approaches were tested on roads. Following sub-chapters describe the process and results of each algorithms tested in the outline delineation of planar objects, and it is the preliminary step for vectorization of boundaries from objects. A number of approaches were tested, including α-shape, Canny detector, Hough transform, and Skeleton. Though some of them worked on very regular context, most of them did not perform adequately.

4.5.1. Building Outlines Extraction

A subset of building roof points was separated. α-shape and Canny detector were tested for building outline generation (the results are shown in Figure 4.6 and Figure 4.7). When testing α-shape, the radius was set to 1 in both regions, in order to acquire the best-fit and appropriately detailed outlines. In region 1, less vegetation mixed with roof edges, resulted in the shape of points being comparatively regular, the outlines were straight. However region 1 also has higher intense constructions, with buildings close to each other, therefore, they were difficult to be separated.

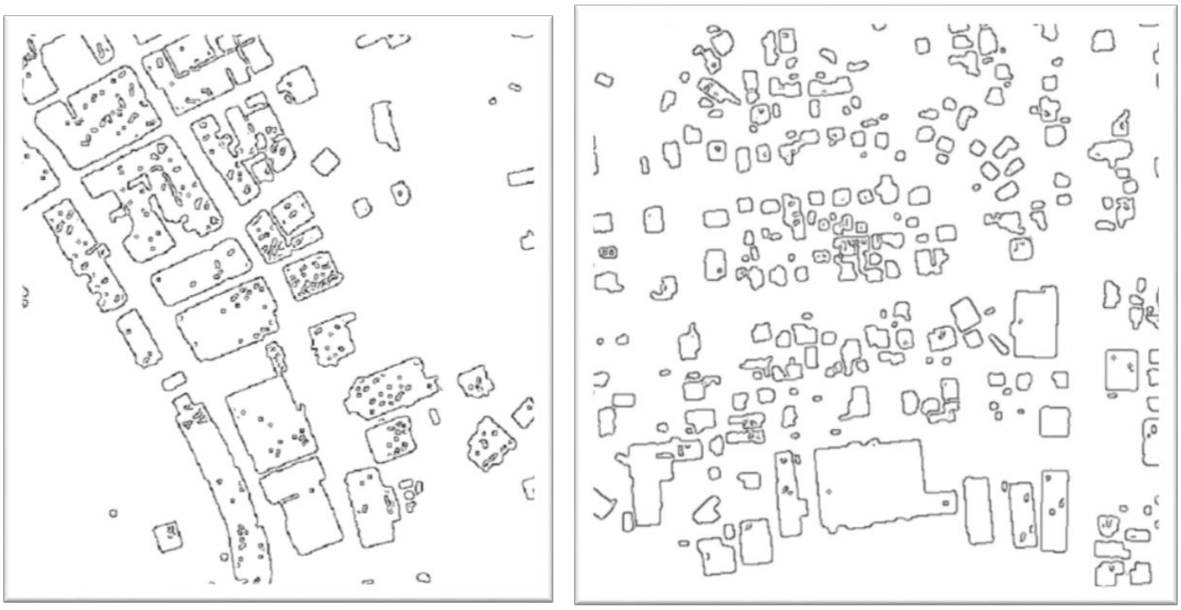
Figure 4.6 Building Outlines by α-Shape



Buildings in region 2 were sporadically mixed with low vegetation, with their borders covered. As a consequence, though building roof outlines were not as regular as in area 1, they were better separated.

The Canny algorithm can be used on remotely sensed images. Building points were projected on raster images, with pixel size equal to point space. It provided similar result to α-shape. However, the extra rasterizing step introduced a decrease in the resolution. The edges were not as sharp as the result of the α-shape. Especially in region 2, small constructions lost their original shape. The most obvious error produced by the Canny algorithm was the small dots inside buildings, and it may be originated from uneven point density. Therefore, in this study, α-shape was adopted.

Figure 4.7 Building Outlines by Canny Detector



4.5.2. Road Outlines Extraction

Roads lie between planar or linear constructs. α-shape and Canny were applied to generate road outlines, while Skeleton, Hough transform were used to detect road trend. Due to the coarse road point classification, the shapes of roads were irregular. When generating road outlines, noisy objects were also included, and result in a zigzag outline. α-shape produced a detail road outline map that described the roadside objects (Figure 4.8). For both regions, the radius was set to 1.6, in order find a balance between straightening lines and maintaining details[6]. Specifically, in region 1, the shape of roads on the slope was odd. On the other hand, since the width of roads in region 2 was smaller than in region 1, disconnects occurred with smaller radius - but enlarging the radius setting would have resulted in a loss of detail.

[6] Comparison on performance of different radiuses is provided in Appendix 3.

Figure 4.8 Road Outlines by α-Shape Algorithm

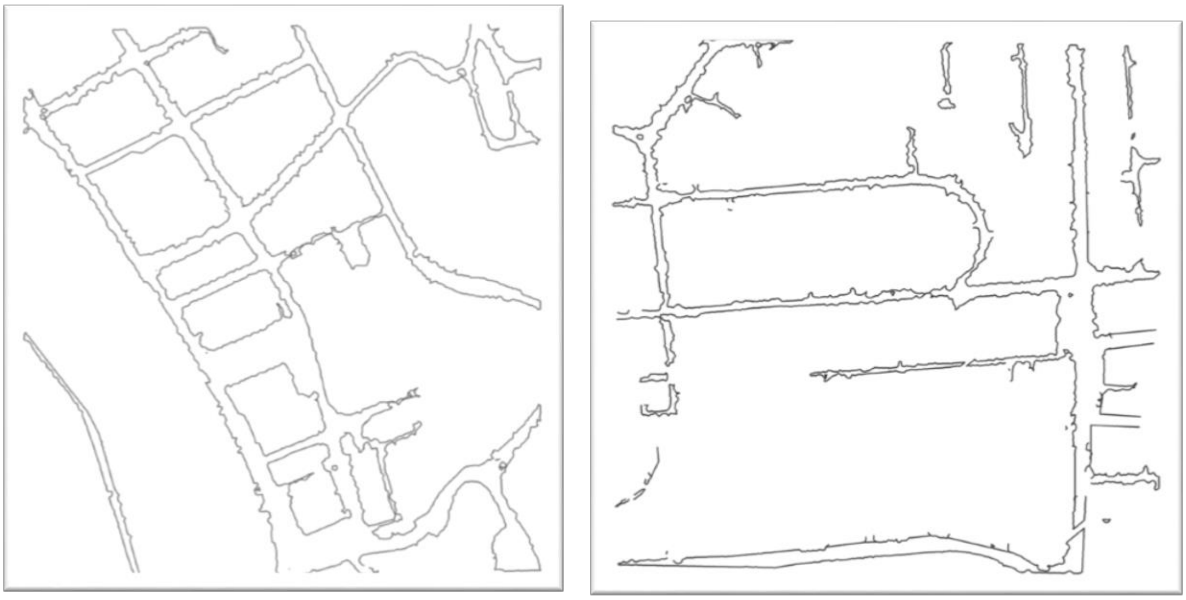
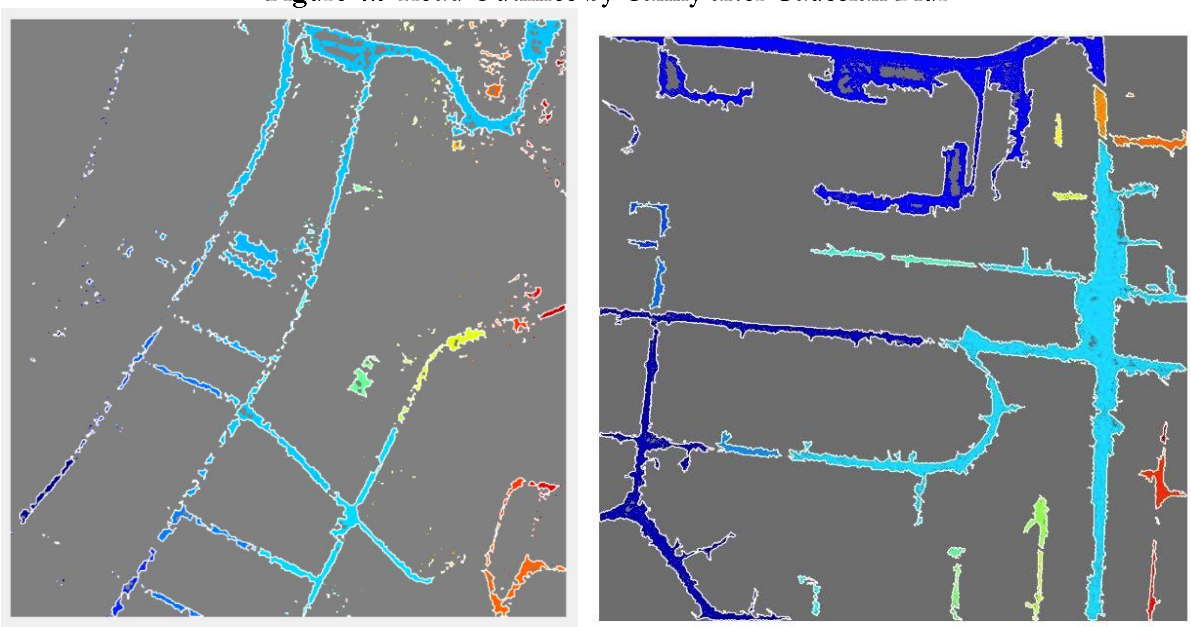


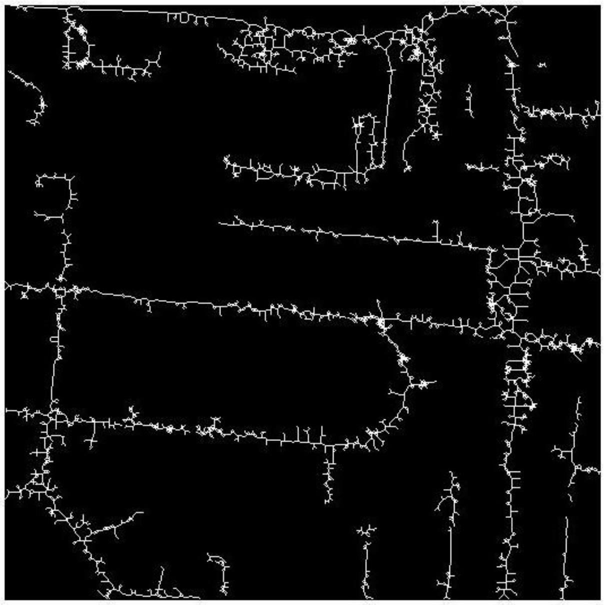
Figure 4.9 displays the result of the outline generated by the Canny algorithm. A Gaussian blur was integrated before the computation of Canny, in order to decrease the thickness of roads. The process helped in removing the noisy roadside objects, meanwhile introducing disconnection. The Canny outline computed the concave hull of road segments, excluding small gaps inside road surface. In both two regions, the performance of the Canny algorithm highly depended on the quality of the classification of road points.

Figure 4.9 Road Outlines by Canny after Gaussian Blur



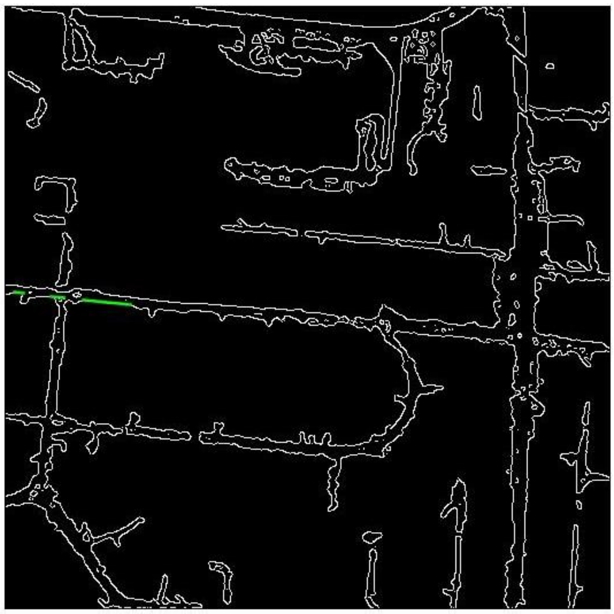
Skeleton drew the rough centerline of roads (Figure 4.10). Thought trends of the road were clear visible, the width of roads was lost. However, parcel boundaries usually coincide with road edges – the role of the road width cannot be overlook. In addition, line segments generated by skeleton were unconnected, further simplification is required.

Figure 4.10 Roads Skeleton



Houghline used the Hough Transform to detect lines from binary images (Figure 4.11). In more regular region 1, a certain percentage of straight lines can be generated, when gaps smaller than 10 pixels were filled and lines shorter than 3 pixels were removed. However, in the denser region 2, small bulges along the roadside caused shortcut of lines. Straight lines segments were shorter than the detectable range. Only very limited lines were generated from region 2.

Figure 4.11 Roads Detection by Hough Transform



Comparing different approaches, the α-shape algorithm was the most suitable for road generation. The zigzag outlines were simplified and straightened with ArcGIS (Figure 4.12), the maximum acceptable distance was set to 2 meters[7] in both regions.

[7] Comparison on different degrees of simplification is provided in Appendix 4.

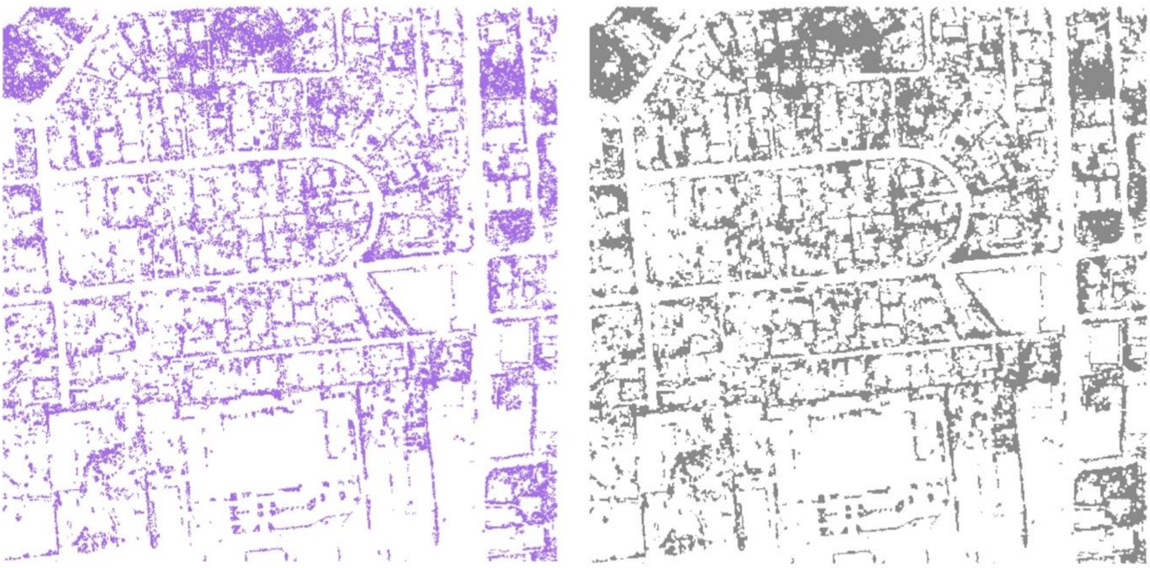
Figure 4.12 Simplified Road Outlines

4.6. Line Fitting From Linear Fences

The last step for automatic extraction phase is delineating lines from linear objects. ArcScan has a centreline fitting function: it can vectorize lines by tracing pixels. The roughly classified fence points were projected on to raster images. Pixels were reclassified into two classes, foreground and ground level. Line fitting was conducted on foreground pixels.

Before operating the line-fitting progress, an opening pre-processing was applied to filled in the small gaps (Figure 4.13). Afterwards raster clearance removed single points by calculating the local stand deviation of point distances.

Figure 4.13 Projected Fences Images before and after Opening



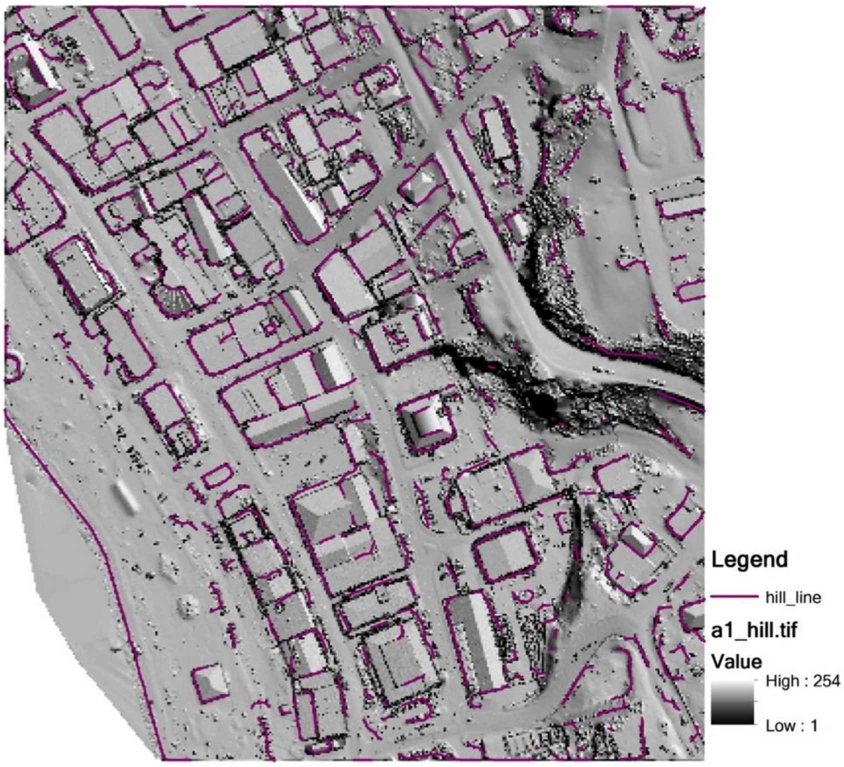
Line fitting results are shown in Figure 4.14. Both centreline and outline were generated. Roughly certain percentages of fence centrelines were drawn. The overview of parcels was vectorized into a polyline.

Disconnecting pixels produced many very short line segments (shorter than 2 meters), and they were unvalued noise.

Figure 4.14 Fences Line Fitting of Region 2

A similar approach was conducted in hillshade images of region 1 (Figure 4.15), to generate topography relief and building outlines, because building outline generation from points cannot be separated from close buildings, this results in more detail building interpretation. In region 1, building walls coincide with a large portion of cadastral boundaries. Therefore, it was necessary to strike a more detailed building outline.

Figure 4.15 Line Fitting from Hillshade Image of Region 1



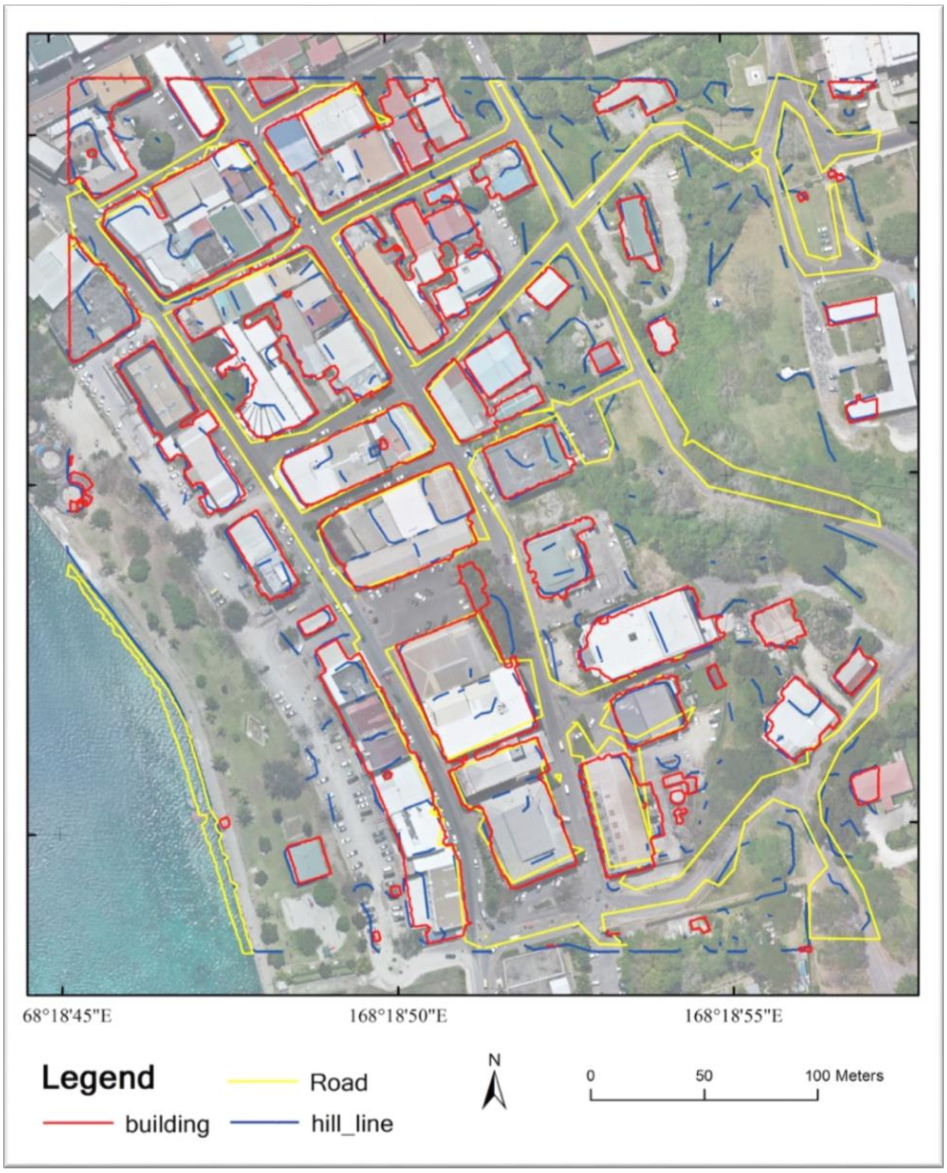
4.7. Reconstruct Parcel Map by Post-refinement

When rough lines have been generated by automatic extraction phase, the results were further refined by the second phase of the workflow: post-refinement phase.

4.7.1. Overview on Automated Extraction Results

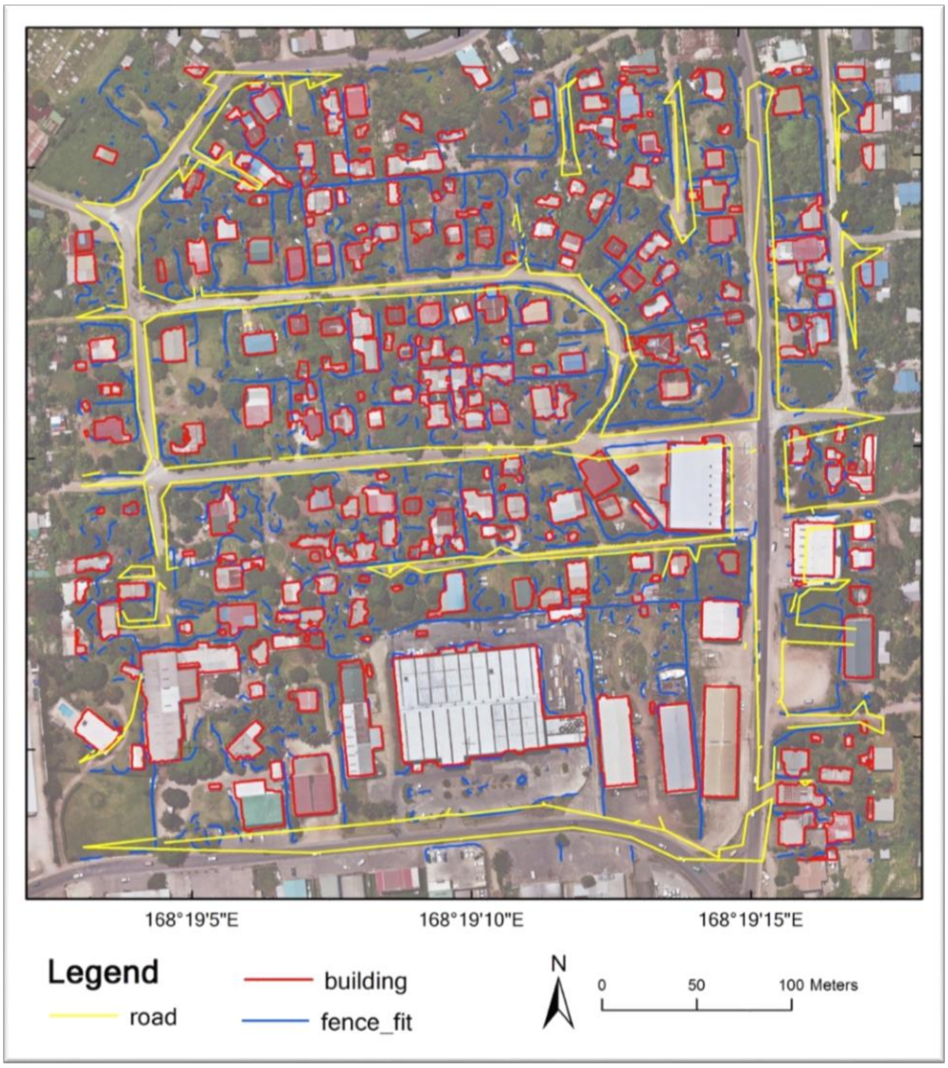
An overview of the study regions cadastral situation was obtained by overlaying all the automated extracted results with orthophoto mosaic. Obvious mismatches between road boundaries and fitted lines emerged in both regions, on account of the road thickness deduction along extraction. Furthermore, fitting lines and building outlines provides much more detail than roads boundaries. Generally, building outlines coincided with fitted lines, illustrating that building extraction process approximately maintains their original shapes. However, the centreline-fitting approach produced lots of short line segments. Either shape corner of road outlines or sporadic short lines contribute to difficulties of visual adjudication.

Figure 4.16 Overview of Automated Extraction in Region 1



In region 1, the fitted line of the hillshade visualization serves as complementary for distinguishing building blocks. However, as a highly developed area, morphology of parcels in region 1 is so irregular that it was hard to define the lines that are likely to coincide with cadastral boundaries. In region 2, where the majority of cadastral boundaries were composed of fence and road, building outlines serve as a supplement for defining parcel locations. In particular, when determining useful lines from line clusters, parcel boundaries normally surround buildings rather than cutting through them.

Figure 4.17 Overview of Automated Extraction in Region 2



4.7.2. Reconstruction with Post-Refinement

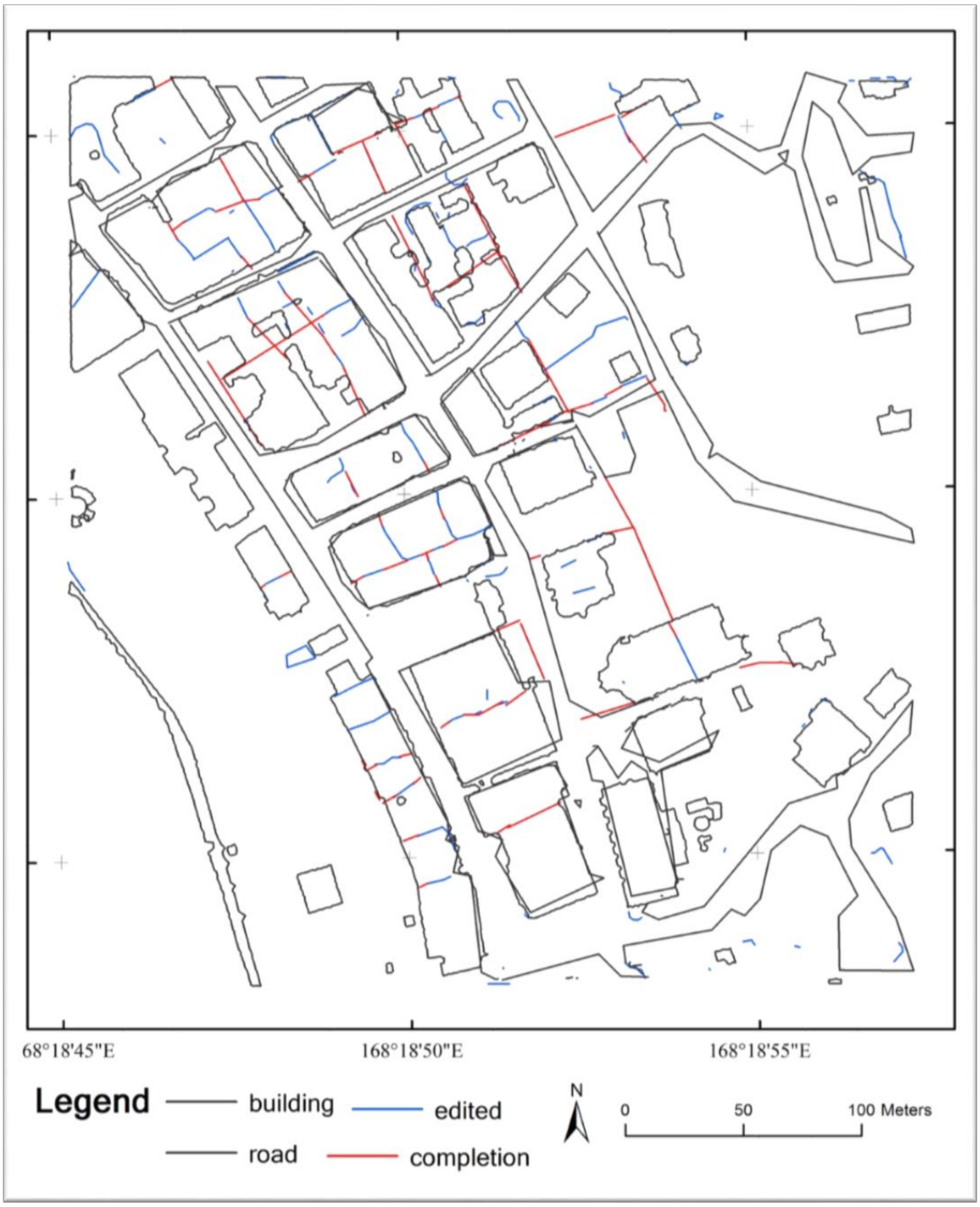
After acquiring general knowledge from the automated extraction results, manual editing was conducted to reconstruct a rough parcel map. Different strategies were designed for these two regions.

An overall reorganization of the detected lines was preliminarily conducted on both regions. With a topology check, close and intersecting lines were merged. In the post refinement phase, less editing was executed on automated extracted building outlines and road outlines because they had higher positional accuracy.

The editing was mainly conducted on fitted lines. In region 1, a topology check was used to search for fitted lines that coincide with building boundaries; they were removed because of redundancy. On the other hand, the left fitted lines created from gaps between buildings or topography reliefs were kept. Comparatively, in region 2, in spite of fence centreline, portions of the fitted lines were generated from roadsides. Topology was also applied to remove this redundancy. Afterwards, the length of the fitted lines was computed as a line attribute: line segments shorter than 2 meters, and were then removed. Based on the target feature strategy, two regions applied a different combination of object lines. A rough parcel map of region 1 consisting of automatically extracted road outlines and building outlines and edited fitted lines from hillshade visualization, while for region 2, it was grouped by automated extracted road outlines and edited fence fitted lines.

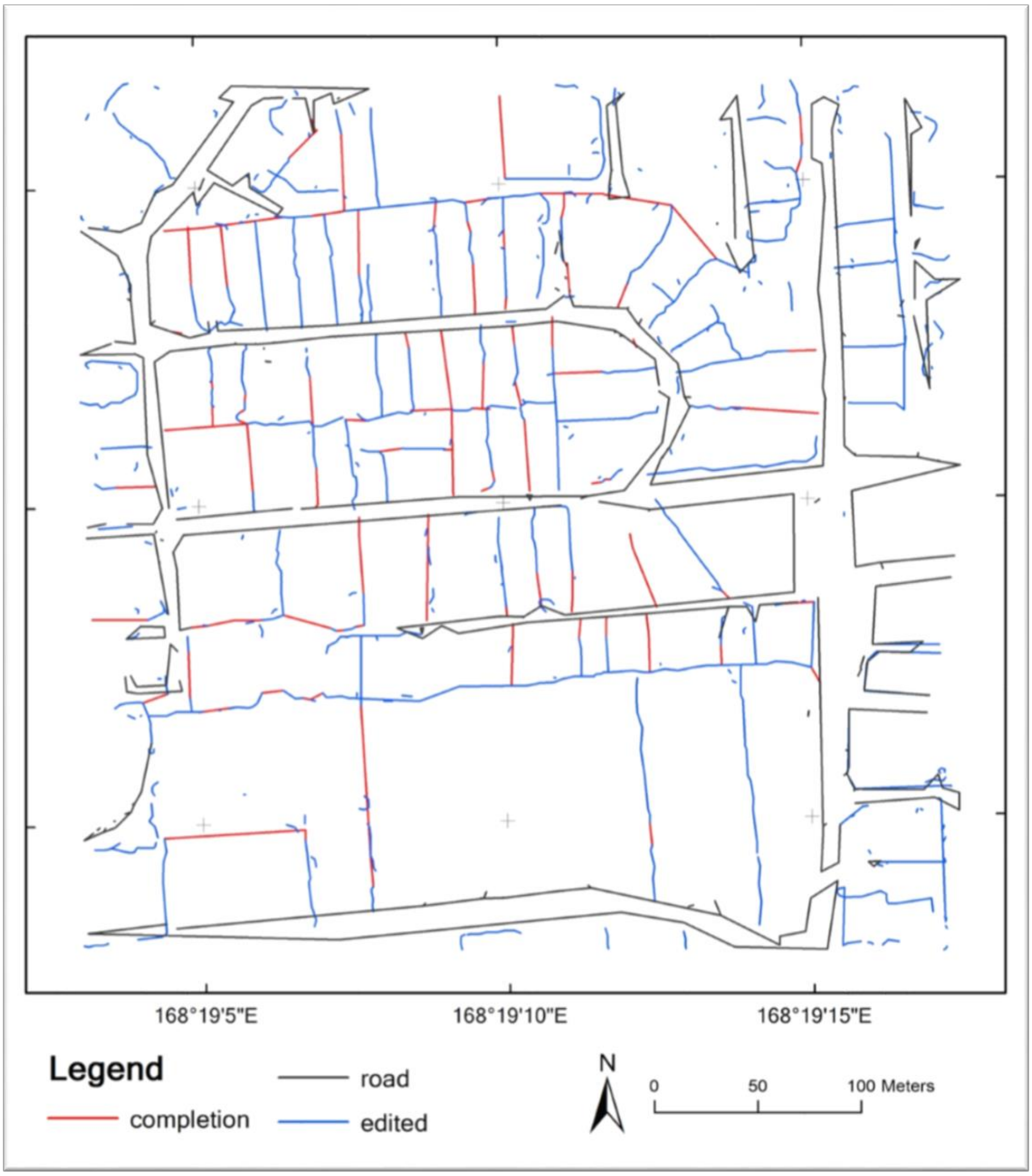
Manual completion was the final step for scene clearance. After determining the useful line segments with the help of visual interpretation, gaps among these line segments were manually filled in and short lines were connected. Partition lines were created on the place where were thought to be parcel boundaries. The final vector draft parcel maps of two regions are shown as Figure 4.18 and Figure 4.19, respectively, edited lines are automated extracted lines but have been corrected or changed, while completion refers to lines that manually added.

Figure 4.18 Reconstructed Parcel Map of Region 1



Generally speaking, for maintaining the research objective, in the post-refinement step, uncertain errors were left maintained, and only obvious and identifiable errors were edited. Certain rules were designed for this post-refinement. For instance, when editing road outlines, the over simplified sharp corners were left, but when the shape of roads were irregular rather than linear, these part were removed. In terms of centreline fitting, either building gaps fitting in region 1 or fence fitting in region 2, two conditions were considered worth editing: One was that the angle between line segments were larger than certain degrees, because parcel corners are not likely to be too acute. The other condition was when the endpoint of a line segment was close to an intersection. If line segments were randomly orientated or sporadically distributed, they were removed.

Figure 4.19 Reconstructed Parcel Map of Region 2



4.8. Summary of Objective 2

The developed workflow targeted at generates parcel boundaries from road edges, building outlines and fences. The whole process consists of two phases: automatic extraction phase and post refinement phase. Specifically, three steps make up the automatic extraction phase: 1) further classify points to target objects; 2) generate planar object outlines; 3) fit centreline to linear objects. Road points were recognized from their intensity, whilst points of fences were mixed in low vegetation. After exploring diverse approaches, it was found that α-shape is more suitable for derive planar object outlines, and centreline generation may be appropriate for delineating linear objects. In post-refinement phase, automated generated line segments were edited and corrected, through topology check and geometry calculation, to solve the problem of redundancy and sporadic short line segments. Finally, manual completion was conducted to fill vacancy leaks and reconstruct a rough parcel map.

5. CHAPTER 5 WORKFLOW PERFORMANCE EVALUATION

5.1. Comparison with Exiting Cadastral Map

The evaluation stage compared the extracted results of both phase (automated extraction phase and post-refinement phase) with reference data: the cadastral map, and made assessment in terms of the correctness and completeness of the developed workflow. Figure 5.1 and 5.2 provides an overview of the workflow performance. There is a small portion of reconstructed lines that completely coincide with ground truth, whilst the others possess errors like offsets and wrong detections. Section 5.2 will describes the error sources as well as tolerances.

Figure 5.1 Workflow Results vs Reference Data in Region 1

Specifically, the workflow performs better in region 2 than region. In region 2, the scene is clearer with less disruptive lines, while on the contrary in region 1, redundant and single stand lines exist. Cadastral morphology is indeterminate in dense urban areas. For instance, when a block consists of a couple of parcels, the outer parcel boundaries normally coincide with roads surrounding the block. But the inter divisional lines normally cut through building gaps. However, in dense urban areas, buildings often adjoin to roads: more efficient land use caused by higher land price. Thus, when the workflow extracts both buildings and roads outlines, a certain redundancy occurs on block outlines. On the other hand, the parcel shape and size are not uniform - such as some parcels possess more than one building, some parcels are larger and the other are smaller. This phenomenon contributes difficulties to the boundary location prediction. Therefore, diverse noise made the overall result of region 1 looks coarser.

Figure 5.2 Workflow Result vs Reference Data in Region 2

Comparatively, the scene of region 2 is cleaner. As illustrated by Figure 5.2, there are more recognizable parcels in the scene. Similar with region 1, blocks are separated by roads outlines. Within blocks, parcels are regularly arranged with evenly distributed areas. Most parcels are rectangle and north orientated. Though fences were hardly seen from images when they were covered by vegetation, the workflow filtered out the high vegetation and fences mixed with the low vegetation were then left. They serve as markers describing the range of land parcels. However, in real-situation, boundaries to not always cut through right in middle of thick brush, therefore there are obvious offsets existing between extracted fence lines and the reference data. In general, the fitted centreline partly draws the division boundaries.

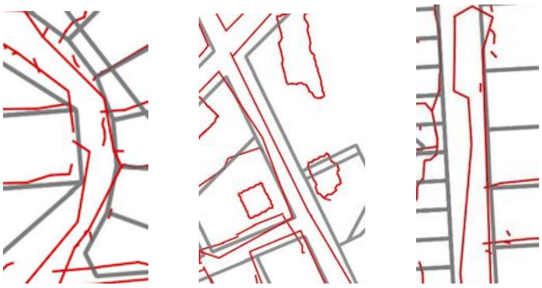
5.2. Error and Tolerance

As described in Chapter 5.1, extracted lines are not totally aligned with the ground truth. Errors can be summarized into three categories, by their sources and appearance. One is raised from outlines generation; the second present as offset, and the third is caused by misalignment between features and cadastral boundaries. By studying the characteristic of errors, standards for the result evaluation were defined based on whether the error can be improved. Tolerance was determined according to maximum distance between extracted lines and the relative reference data. If errors can be eliminated through further improvement, they were recognized as acceptable. From observations, the approximate average distance between the workflow result and reference data is 3 meters.

1) Over simplified

The coarse outline extraction introduced errors when connecting boundary points. Especially in road outline generation, the reducing approach wrongly subtracted points on road as well as decreased the road's thickness. Though a simplification step was integrated, but it also brought errors of over simplification. As shown in Figure 5.3, the roads width was decreased as well as the shape of some corners were lost. A better threshold for straighter and simplified road outlines should have been investigated.

Figure 5.3 Illustrations of Road Outline Error



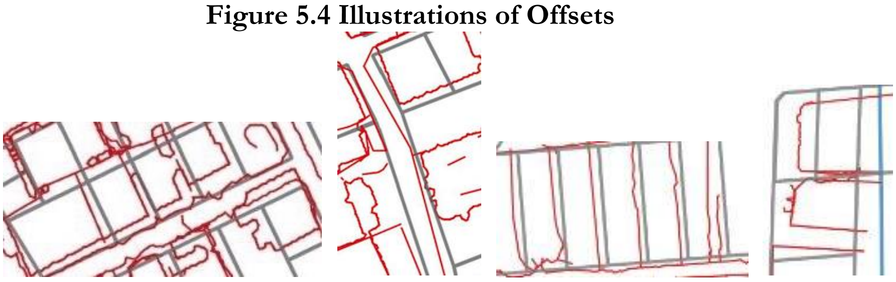
(a) **(b)** **(c)**

If the maximum distance to reference data is smaller than 4 meters[8], which is one thirds of road width, such as Figure 5.3.a and 5.3.b, the extraction were recognized as correct, because the On the contrary, if the maximum distance is too large, though they were correctly extracted from features, they were determined as errors, as Figure 5.3.c.

2) Offset

The most common error is offset between the extraction result and the reference data. Offsets occur in both building outlines and the fence fitting lines, but the causes of them were different. No systematic shift was observed between these two datasets, and the causes of the offset might be dynamic. In region 1, building outlines were shifted at certain angle, and it might be caused by the vague edge of building points. As described in Chapter 3, the cadastral map is not totally coincidental with the orthophoto. In region 2, probably because of the fact that not all cadastral boundaries cut through the centre of fences, offsets exist between the extracted result and the reference data. Fortunately, the thickness of brushwood and fences are small enough to control the maximum shift distance.

Figure 5.4 Illustrations of Offsets



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

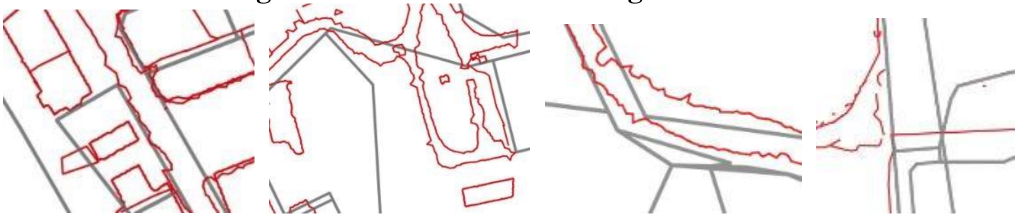
Similar to the previous type of error, when the average distance between the extraction result and the reference data was smaller than 4 meter, like Figure 5.4, errors were considered within tolerance. Otherwise, they were recognized as unvalued, like Figure 5.4.b and Figure 5.4.d.

[8] Justification is provided in Chapter 6.1.3.

3) Misalignment

The statistic analysis in Chapter 3 has already illustrated that over 70% of cadastral boundaries are visible, and therefore extractable. The remaining 30% of lines are neither coincidental with any features nor visible by human eyes. Vice versa, these feature outlines representing no parcel boundaries. Simply improving extraction technique cannot decrease this type of error. More intelligent approaches should be integrated to mitigate human definition of cadastral boundaries. So this type of errors falls outside tolerance.

Figure 5.5 Illustrations of Misalignment



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

5.3. Workflow Correctness and Completeness

Since errors and tolerance are determined, the performance was quantitative evaluated with statistic. The percentage of correctly extracted lines and the proportion of detected parcels were selected criteria to describe the completeness of the designed workflow. Correctness is illustrated by the percentage of correct extraction from total number of extracted lines.

5.3.1. Proportion of Detected Lines from Each Kind of Features

The table 5.1 describes the number of both true and false line segments from total extracted lines. Road extraction was conducted on both regions: building outlines and fence fitting were executed in region 1 and region 2 respectively. The total number of extracted line segments was counted. True lines are extracted lines that coincide with relative cadastral boundaries, while false lines are either wrongly detected or their error are larger than tolerance.

Table 5.1 Number of Detected Lines Coincide with Features

	Feature	Extracted	True	False	Lost
Region 1	Road	111	90	21	13
	Building	97	23	74	62
Region 2	Road	177	148	29	49
	Fence_fit	106	91	15	11

The correctness was computed as percentage of true line segments from total extraction. In region 1, in total 54.32% of extraction are true, while this number raise to 84.1% in region 2. Completeness was computed by backing to Table 3.1, the percentage of extracted line segments from detectable boundaries. In region 1, in total 71.76 % of cadastral boundaries are detectable, and 60.11% of them were successfully extracted by the workflow. In region 2, 79.93% of lines segments from 74.56 % detectable lines were successfully extracted, approximately 20 % more than that of region 1.

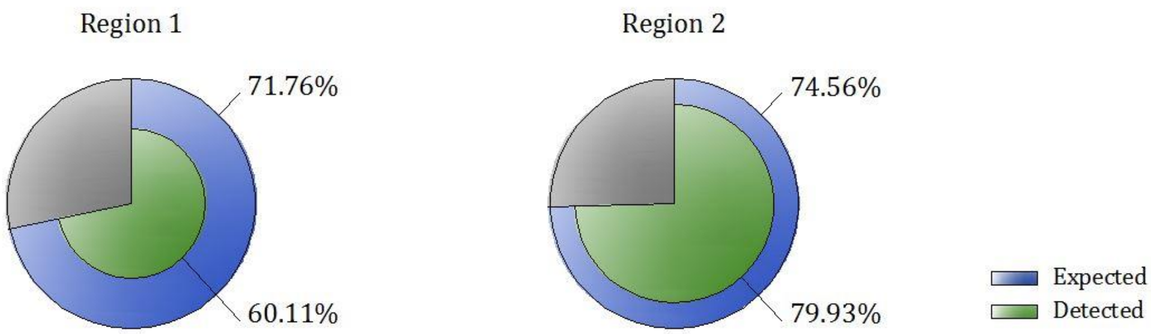
The road extraction achieved more than 80% correctness in both regions (Table 5.2). However, only 23.7% of extracted building outlines coincide with parcel boundaries, which also reveals the complicated cadastral boundary morphology in urban area. The fence fitting realized the highest correctness in this study.

Table 5.2 Overall Performance of Workflow with Features

	Feature	Correctness	Completeness	Overall	Expected	Completeness
Region 1	Road	81,08%	87,38%	**43,13%**	**71,76%**	**60,11%**
	Building	23,71%	27,06%			
Region 2	Road	83,62%	75,13%	**59,60%**	**74,56%**	**79,93%**
	Fence_fit	85,85%	89,22%			

Figure 5.6 illustrates the overall performance of workflow in two regions. The blue area describes the proportion of detectable lines, while the green area stands for overall completeness by relative ratio of radius.

Figure 5.6 Illustrations of Completeness of Workflow



5.3.2. Performance of Full Parcel Identification

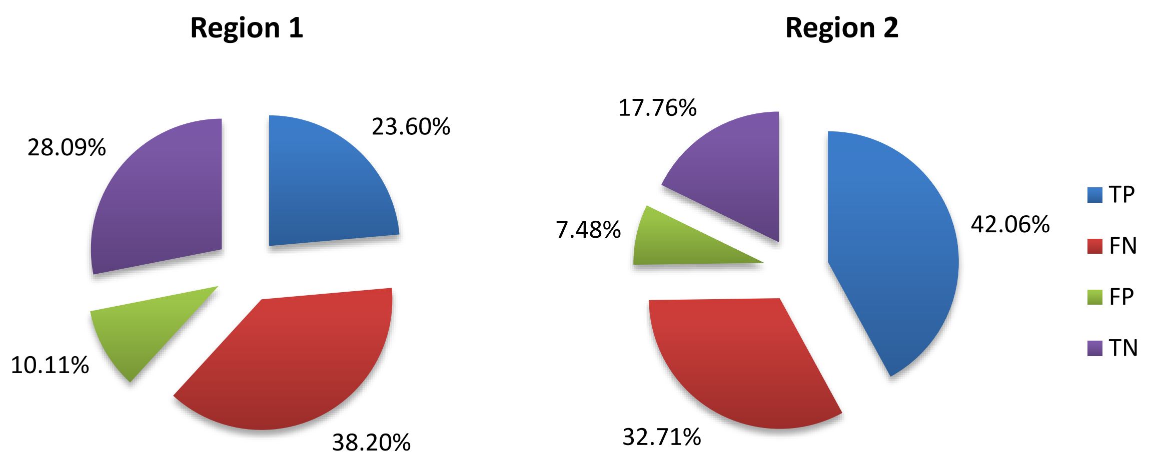
The table 5.3 illustrates the fully detected parcel. Identifiable parcels are parcels where all their boundary segments have been extracted. It is illustrated that the ratio of identified parcels of region 2 is almost double than that of region 1. Shifted parcels stands for extracted parcels but are significantly shifted from right position. The workflow extracted less shifted parcels from region 2 than from region 1, reflecting that the workflow performed higher positional precision in more regular area.

Table 5.3 Parcels Identification Correctness

	Tested	True	False	Shifted	Lost	Total	Correctness
Region 1	55	21	9	25	34	89	**38,18%**
Region 2	72	45	8	19	35	107	**62,50%**

Figure 5.7 displays a straightforward illustration on parcels identification. It uses true, false, positive and negative to summary the overall performance on the parcel extraction. TP is parcels appear in both extracted result and reference data. FP is where parcels were in reference data but were not extracted. FN is where parcels appear in result but do not exist in reference data. And TN is parcels in reference data but that is only partly extracted or have visible shift. TP describes the completeness of the parcel identification, and both regions achieved less than half, region 1 even have only 23.60% completeness.

Figure 5.7 Illustrations of Parcels Identification Completeness



5.4. Degree of Automation

Two phases of the semi-automated workflow were separately evaluated, in order to assess the workflow's degree of automation. Firstly, the total length of the line segments generated in both automatic extraction and post-refinement phase were computed. Then the correctness of each phase was assessed respectively, and it indicates that the workflow performed better in region 2 in either phase. Table 5.4 illustrates the different performance of the workflow in two regions.

Table 5.4 Comparison of Automated Extraction and Post-Refinement

	Feature	Length	True	Correctness
Region 1	Manual	982,05	667	67,93%
	Edited	1111,54	727	65,46%
	Road	5354,7	3945	73,69%
	Building	12133	9214	75,94%
Region 2	Edited	4787,57	4376	91,41%
	Manual	1651,8	1374	83,22%
	Road	5428,36	5011	92,33%
	Building	11392,59	11252	98.80%
	Fence_fit	10099,8	--	--

In Region 1, the workflow was conducted on the two most probably features: roads and buildings, and achieved both approximately two-thirds correctness. Without ground reference, it is hard to judge whether extracted lines from road points are correct or not. Therefore, for maintaining this experiment objectivity, in post-refinement step, uncertain errors were left maintained, only obvious and identifiable errors were edited.

Equation 5.4 calculates the degree of automation D, where the total length is the total length of workflow result that consist of automated extraction result and post-refinement result; automated extraction is the total length of the automated extraction phase result that have been used for reconstruction. In this study, in region 1, automated extraction results of road and building outlines were used, whereas in region 2, only road outlines were used, and all fence fitted lines were edited.

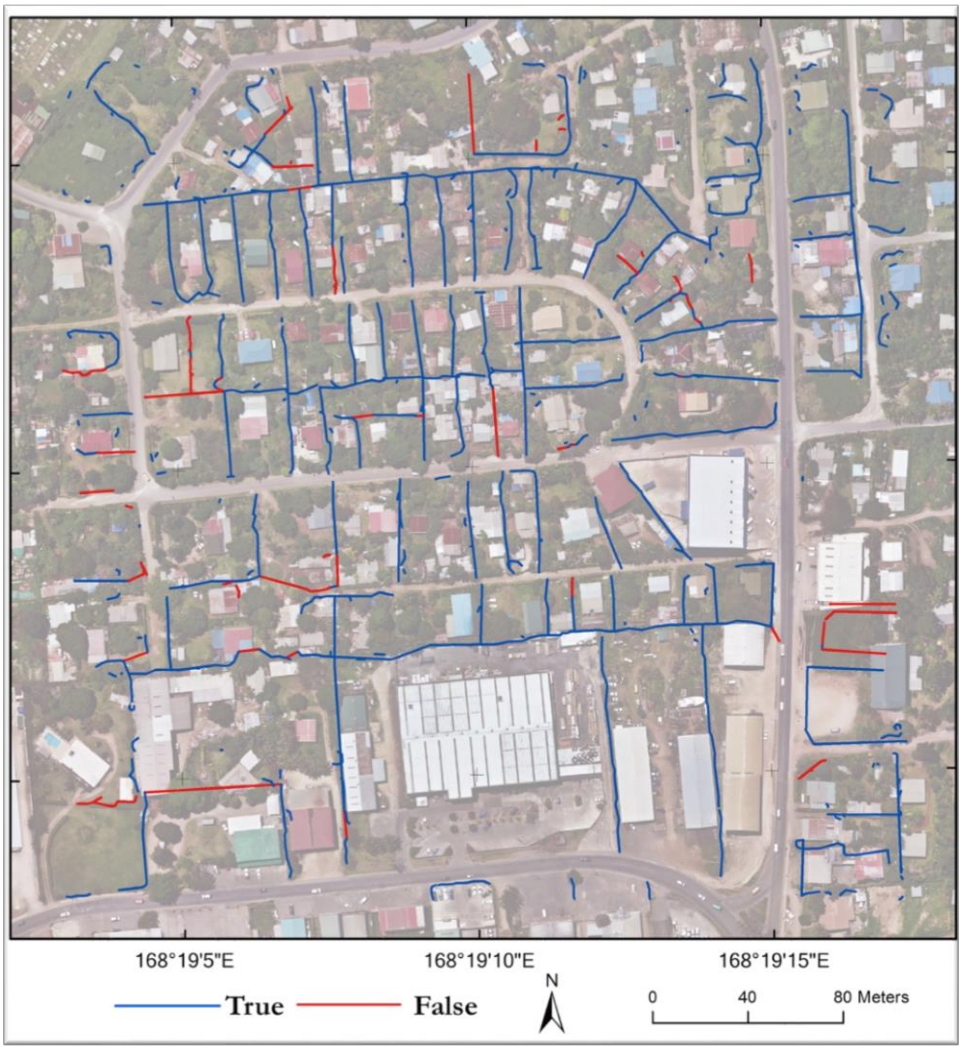
$$D = \frac{Automated\ Extracted}{Total\ Length} * 100\%, \qquad 5.4$$

The automation degree of workflow in region 1 is 89.3%, and the number drop to 45.75% in region 2. Furthermore, only 10.13% of true line segments in region 1 are created by the post-refinement phase, the corresponding number in region 2 jumps to 34.24%. This slight difference is caused by more unpredictable parcel morphology in the dense urban area, which is region 1 in this study. It was difficult to judge whether an extracted line in region 1 is valuable, as a consequence in many cases, errors were maintained with less manual editing. Comparatively, the fence-fitting result was coarse and all of them have been manual inspected.

Figure 5.8 Correctness of Post-Refinement in Region 1

From another perspective, the post refinement phase achieved over 85% correctness in region 2, compared to 65% in region 1. It further proves that without reference, regular parcel shape and uniform parcel size would contribute to correctness of the post-refinement activities, in that human interpretation highly depends on these regular patterns. Besides, the correctness of automated of extraction of different feature were all around 70%, indicating that capabilities of the developed workflow is still less than cadastral mapping requirements. Figure 5.8 and Figure 5.9 visualize both the true and false edited line segments as well as manually created line segments.

Figure 5.9 Correctness of Post-Refinement in Region 2



5.5. Summary of Objective 3

The result of objective 3 shows that in region 1, half of the results of the developed workflow was correct, while in region 2, this number increased to over four fifths. Besides, around 40 percentages and 60 percentages of the parcel boundaries in region 1 and region 2, respectively, were generated by the workflow. The automation degree of the workflow was 90 percentages and 45 percentages in region 1 and region 2, respectively. The correctness and completeness of the workflow is promising, and the evaluation results indicate that the workflow performed better in the more regular suburban area. However, the complex parcel morphology in the dense urban area contributes difficulties to parcel boundaries delineation.

Though various errors exited, the tolerance was set to 4 meters, according to whether the error could be illuminated by further improvement: the modest spatial accuracy of the workflow still cannot meet the requirement of cadastral survey, and further researches deserve exploration.

6. CHAPTER 6 DISCUSSION AND CONCLUSION

6.1. Discussion on Research Results

Diverse observations from the research processes as well as the result of the research reflect on strengths and weaknesses of the workflow. The chapter is structured around the specific sub-objectives and answers to related research questions in each case.

6.1.1. Reflections on Cadastral Boundary Morphology Study

The cadastral boundaries morphology was studied from orthophoto mosaic and existing cadastral map in advance, to determine the kinds of topographic objects that might coincide with cadastral boundaries. The result indicates that in different area, cadastral boundaries may coincide with different topographic objects. The landscape plays an important role in determining the cadastral boundary morphology. Therefore, two regions were selected as typical subsets, an urban region and a suburban region respectively. The following study and the workflow were conducted on these two subsets. By statistical means, the number of cadastral boundary segments that coincide with each topography object; the target objects for feature extraction to support cadastral survey were determined. There is no standard on which kind of feature should be used as markers for cadastral boundaries. Both manmade and natural objects have the capacity for enclosing spatial rights – especially in general boundary system - on condition that they are visible and constant.

In the dense urban area and the suburban area of Port Vila, Vanuatu, parcel boundary segments coincide with different objects. Roads coincide with the most percentage of cadastral boundaries in both regions. In the urban region, the price of land in dense urban areas is higher, thus the land use there is more efficient, which results in constructions close to each other. Therefore, in most cases, parcel boundaries are close to construction inside the parcel. Besides, the shape and size of parcels in the urban area are very diverse. It is difficult to define how many constructions a parcel contains from images. Within building clusters, cadastral boundaries mainly coincide with building gaps. On the contrary, in the suburban region, buildings are sporadically distributed, with grassland and vegetation surrounding them. Within a block, parcel boundaries cut through the middle of two buildings; in many cases one parcel contains one building. One third of visible cadastral boundaries coincide with fences. They are widely accepted for private owned land demarcation, very likely because fences are low cost but unambiguous.

Though higher point density provides more details, it also costs more. In contrast, some skinny stand-alone cadastral boundaries cannot be detected from lower density point clouds. Thus, appropriate data volume should be investigated based on fit-for-purpose concept. Therefore a suitable point density should be defined in advance, according to accuracy requirement of cadastral survey. The point density of the study data is 9 p/m^2, thereby according to the Equation 3.3, the smallest detectable object should larger than 6m. Therefore, fences are not detectable. Buildings and roads are both larger than the minima. However, very thin and long objects still have possibility to be seen from point cloud. In this study, fences were partly detectable. Obviously, a higher point cloud density would have helped to detect more fences.

In summary to the specific research questions:

- What physical features also double as cadastral boundaries in the study area?

In Port Vila of Vanuatu, parcel boundaries in different areas coincide with different physical objects. Generally, road edges, building outlines, fences, and vegetation are all likely to coincide with cadastral boundaries.

- What is the probability of each kind of feature being a cadastral boundary?
  Road edges coincide with half cadastral boundaries in both the urban and the suburban area. In region 1, building outlines plays the second significant role whilst it has much less value in region 2. In region 2, fences mark one forth of parcel boundaries. In total, over eighty per cent of cadastral boundaries in both the urban area and the suburban area coincide with visible objects; these boundary segments are extractable, and the rest are invisible from remote sensing data.

- What are the useful characteristics of these features?
  Table 3.2 describes the useful parameters for each feature. Roads can be identified from reflectance intensity; Local smoothness and height characterize building roofs, and fences have uniform height, but they are so thin to be separated from vegetation.

- What is the suitable point density?
  The suitable point density for the workflow was determined according to the smallest dimension of target objects, which was the fence width in this study. Normally the width of the fence is 0.5m, thus the smallest point cloud density for detecting the fence is 16 p/m^2, based on the Equation 3.3. However, the point density of the study data is only 9 p/m^2.

6.1.2. Reflections on the Semi-Automated Extraction Workflow Exploration

The second objective was to design a tailored workflow to reconstruct a parcel map from airborne laser scanning data. The workflow consisted of two phases: automatic extraction phase and post-refinement phase. And the automatic extraction phase consists of three steps: 1) further classify points into target objects; 2) generate planar object outlines; 3) generate centreline of linear object. Then the extracted result is edited and completed in the post-refinement phase. Different algorithms and methods were tested for each step. The workflow aimed at extracting roads and buildings in urban areas as well as roads and fences in suburban areas.

LiDAR point cloud is a newly proposed data source for cadastral survey without a long history application. Despite diverse advantages of using LiDAR point clouds, its feasibility for cadastral survey has not been certified. In addition, reflected laser beams contain no information on objects but only 3D coordinates. Therefore, knowledge on target features is essential and should be confirmed in advance. In particular, priority should be given to characteristics such as shapes and spatial distribution, so that they can be computed easily from LiDAR point clouds.

After exploring the proposed workflow, the α-shape algorithm was adopted for both building outline and road outline generation. Since this study aims at the reconstruction of a 2D cadastral map, only planar coordinates of points were taken into consideration. But this manner introduces errors, in that different roof planes cannot be separated from different heights. Additionally, the gaps between buildings are thinner than the smallest detectable object from a 9 p/m^2 dataset. In many cases, close building roofs were wrongly segmented into one cluster. The accuracy of the α-shape algorithm highly depends on point clouds segmentation quality, in that it connects all boundary points of a segment to construct the boundary polyline. As a result, when generating building outlines in the urban area, close building roofs were wrongly merged into one polygon.

In terms of the road point classification, the quality was restricted by the single band of point clouds. The LiDAR data in this study works on near-inferred bands, which offer much less spectrum information than multi-spectral images. The workflow detected points of roads from its intensity, however, the reflectance from road surface and roadside constructions are too similar so wrong detections occur along roadside. A fusion of point clouds and multi spectral images might provide a better quality result.

Since a cadastral boundary is a societal construct, manual verification can be largely reduced through semi-automated extraction rather than eliminated totally. However, not all cadastral boundaries are visible, and not all objects coincide with cadastral boundaries. Therefore, it was hard to judge whether the line is valuable, and in consequence redundancy and incompleteness are inevitable during the whole workflow. Whether the accuracy of automated extraction meets the requirement of cadastral survey needs a quantifiable justification. During the centreline fitting, either building gaps fitting in region 1 or fence fitting in region 2, two conditions were considered worth editing. One was line segments that were perpendicular to true boundaries, because parcels are likely to be rectangular. The other condition was when a line segment was single stand, but its end points are close to intersection. If line segments were random orientated or sporadically distributed, they were removed. In the post-refinement phase, manual interpretation was applied on the automatically extracted result, to identify which lines might coincide with cadastral boundaries. It was difficult to make judgments in the study area, especially in the dense urban area. Because in the dense urban area, there was no basis to define which building outline segment is useful, in that a parcel might contain one or multiple buildings. This situation results in a lower correctness of post-refinement in region 1 compared to region 2.

In response to the relevant research questions relating to objective 2, the following statements can be made:

- What are targeted features of the workflow?
  For dense urban areas, the workflow targeted at roads and building outlines extractions, therefore in total about eighty percent of cadastral boundaries were detectable. In terms of suburban areas, roads and fences were selected for extraction, thus approximately seventy of cadastral boundaries coincide with these features, and were therefore extractable.

- Which parameters are suitable for the classification in the workflow?
  When carrying out further classification of points for feature extraction, diverse parameters were selected for different objects. The workflow computed the point intensity and the segment size to extract road surfaces. As for points of fences, their heights were taken into consideration. After removing high vegetation, fences are still mixed with low vegetation. Fences are difficult to be identified from ALS point clouds at the low density of points per square meter.

- Which outline generation algorithm is suitable for planar objects?
  After comparing α-shape and Canny, α-shape was more suitable for both building outlines and road outline generation: it created a concave hull of point clusters by computing boundary point of segments and connecting them, therefore sharp edges are maintain while small gaps are filled.

- How can linear objects be extracted?
  When exploring suitable ways to depict the relationship between objects and parcel boundaries, fences and roads were considered linear objects for testing centerline generation. The results obtained showed that outline generation was more suitable for roads, in that it maintain roads width. In this study, an approach for centerline fitting from raster was integrated to detect the

centerline of linear objects. Points of fences were projected onto raster images, after opening and reclassification, centerlines of connected pixels were computed by ArcScan.

- What kind of post-refinement is needed and how should it be completed?
  After rough line segments were automatically extracted, several steps can be used to improve the quality of the parcel map. Topology checks detected the disconnection of line segments as well as redundant lines. Line smoothing and line simplification straightened the lines. Subsequently, visual interpretation editing supplemented the detection to fill vacancy leaks.

6.1.3. Reflections on the Workflow Performance Evaluation

The reconstructed result was compared with the exiting cadastral map. The correctness and completeness of results were studied to quantitatively assess the workflow performance. The degree of automation illustrates the efficiency of the workflow.

During the performance evaluation, the tolerance was determined to 4 meters, which is significantly larger than cadastral survey requirement. But the point density of study data restricts the precision and accuracy of this workflow. Therefore, as an exploratory approach, the tolerance was set modestly, according to whether the error can be improved.

Correctness and completeness are two major criteria for performance evaluation. The workflow works better in the suburban area than in the dense urban area. The completeness of parcel identification in the suburban area is double than that in the urban area. Furthermore, the correctness of line segments extraction in the suburban area is 20% more than that in the urban area. This difference mainly caused by the diverse parcel morphology in the urban area. In the urban area, size and shape of parcels are diverse. However, in the suburban area parcels are usually rectangle in shape and regularly distributed with similar area. However, Therefore, the workflow performs better on regular context. Parcels morphology plays an important role in extracting their boundaries.

In response to objective 3, the following can be stated with regards to the research questions:

- What is the error tolerance of the workflow?
  The tolerance in this study was 4 meters. Two conditions were considered acceptable: one was that results were shifted from reference data, and the distance was less than 4 meters. The other emerges in the over-simplified: when the majority part of results was coincided with reference data, and the maximum distance between humps and reference data were less than 4 meters.

- What is the correctness of the extracted boundaries?
  The correctness was described by computing the ratio of the true extraction to the whole result. In the automated extraction phase, the road outline generation achieved over 80 percent of correctness in both urban and suburban areas. Only one-fifth of generated building outlines were correct, while fence fittings has three times more correctness. The overall correctness was about 50 percent and more than 80 percent in the urban and the suburban area, respectively.

- What is the percentage of completeness?
  Completeness was illustrated by the ratio of the true extraction to the total cadastral boundaries. In the urban area, the workflow achieved an overall completeness of more than 40 percent, and around 60 percent of detectable cadastral boundaries were extracted. As for the suburban area, four fifths of detectable cadastral boundaries were extracted by the developed workflow.

- What is the degree of automation?
  The automation degree of the workflow in region 1 was almost 90 percent, and the number dropped to less than 50 percent in region 2. Furthermore, only 10 percentages of true line segments in region 1 were created by the post-refinement phase, the corresponding number in region 2 jumped to two times more.

6.2. Conclusion and Recommendation

This study, which explored the feasibility of point cloud data for cadastral survey, achieved the main objective: develop a workflow for semi-automated extraction of cadastral boundaries from airborne laser scanning data. Port Vila of Vanuatu was selected as study area, in order to investigate the capability of semi-automated cadastral mapping on developing country context. The study focused on the visible general boundaries, through exploring suitable methods; an object-based workflow was developed to semi-automatically extract cadastral boundaries from ALS point clouds. The result of the developed workflow is promising, with around 50 percentages of parcel boundaries successfully extracted. A coarse parcel map can be arranged with the workflow within several hours. If one brings the parcel map to field for verification, only incorrect boundaries need to be digitized afterwards. However, the spatial accuracy of this workflow is still modest, because most steps of the workflow introduce errors. Furthermore, the workflow is context specific. It was fortunate that in the study area, a large proportion of cadastral boundaries coincided with topographic objects. It may not be suitable for irregular area such as dense slum areas. Moreover, the workflow performed better in the more regular suburban area. Due to the complexity of the cadastral boundary morphology in the dense urban area, the performance of the workflow in the dense urban area is still modest. More research on the relationship between topographic objects and parcel boundaries deserves further exploration.

LiDAR data possesses both strengths and weakness for cadastral purpose, and the workflow explore the feasibility of LiDAR in cadastral surveys. The most significant advantage is its ability to penetrate through a vegetation canopy, which contributes a lot to the fence extraction. Fences are widely used as cadastral markers but they are often invisible from aerial images. In spite that there is high vegetation covering above, laser beams can penetrate through the canopies and return footprints of objects below high vegetation, which in this study was the low vegetation mixed fences. Another strength of using LiDAR is it provides height information. Point clouds data can provide high accuracy xyz measurements. Height plays an important role in objects recognition. On the contrary, the major weakness of using point clouds is its accuracy. The level of detail that LiDAR can provide is highly dependent on the point density. Extracting small objects such as the fences occurring in the test site requires a higher point density, but, of course, a higher point density also means a larger data size as well as higher costs of data acquisition. This is a disadvantage in application in developing countries – where costs remain a primary concern and inhibitor.

In view of the strengths and weakness of the workflow, several recommendations are derived for improvement. The first one is to improve the feature extraction approach. A constraint that restricts the road width was not integrated into this workflow. By computing the distance between points, a fixed width road surface can be formed and then the zigzag outline problem may be solved. In addition, when generating building outlines, the workflow only considered the xy coordinates. If the α-shape algorithm is extended to 3D, close roof planes can be separated by their different heights. Furthermore, a combination with multi-spectral images can enrich spectral information and makes contribution to point-based classification. The third recommendation is to automate the post-refining phase. This can be achieved by integrating them into line generation. Taking a topology check and the line simplification as an example, if they can be executed during the generation of lines, not only much manpower could be saved, but also the whole workflow would be accelerated.

Another suggestion for improvement is to adapt the object-based approach to investigate the relationship between objects and parcel boundaries, such as the possible distance between constructions and parcel boundaries. In urban areas, cadastral boundaries may be close to a building outline; whilst in suburb areas, cadastral boundaries may evenly partition an area. These parameters could be taken into consideration for predicting parcel range and location by means of machine learning. Due to the time limit placed on this research, it was not possible to realize these suggestions in this study, but they may be worth exploring in the future.

LIST OF REFERENCES

Abeysinghe, S. S., & Baker, M. (2008). Segmentation-free skeletonization of grayscale volumes for shape understanding. *2008 IEEE International Conference on Shape Modeling and Applications*, 63–71. doi:10.1109/SMI.2008.4547951

Amenta, N. (2015). Directory of Computational Geometry Software. Retrieved November 16, 2015, from http://www.geom.uiuc.edu/software/cglist/

ASPRS. (2009). Las Specification 1.3, (November 2011), 17. Retrieved from http://www.asprs.org/a/society/committees/standards/LAS_1_4_r13.pdf

Bai, X., Latecki, L. J., & Liu, W. Y. (2007). Skeleton pruning by contour partitioning with discrete curve evolution. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(3), 449–462. doi:10.1109/TPAMI.2007.59

Baigent, R. J. P. K. and E. (1992). *The Cadastral Map in the Service of the State: A History of Property Mapping*. The University of Chicago Press.

Ballard, D. H. (1981). Generalizing the Hough transform to detect arbitrary shapes. *Pattern Recognition*, *13*(2), 111–122. doi:10.1016/0031-3203(81)90009-1

Battles, E. (2016). Surveying, Mapping and GIS: Epic Battles: GIS vs. Surveying? Retrieved January 26, 2016, from http://surveying-mapping-gis.blogspot.nl/2008/07/epic-battles-gis-vs-surveying.html

Bennett, R., & Alemie, B. (2015). Fit - For - Purpose Land Administration: Lessons from Urban and Rural Ethiopia. *Survey Review*, *000*(60).

Bruce, J. W. (1998). Review of tenure terminology. *Tenure Brief*, (1), 1–8.

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *8*(6), 679–698. doi:10.1109/TPAMI.1986.4767851

Clode, S., Kootsookos, P., & Rottensteiner, F. (2004). The Automatic Extraction of Roads from LIDAR data. *RealTime Imaging*, *35*, 231–237. doi:10.1.1.133.520

Dale, P., & McLaughlin, J. (2000). *Land Administration. OUP Catalogue*. Oxford University Press. Retrieved from http://ideas.repec.org/b/oxp/obooks/9780198233909.html

Dale, P., & McLaughlin, J. D. (1988). *Land information management. An introduction with special reference to cadastral problems in developing countries*. Clarendon Press. Retrieved from http://agris.fao.org/agris-search/search.do?recordID=GB8908430

de Burlo, C. (1989). Land alienation, land tenure, and tourism in Vanuatu, a Melanesian Island nation. *GeoJournal*, *19*(3), 317–321. doi:10.1007/BF00454578

Deriche, R. (1987). Using Canny's criteria to derive a recursively implemented optimal edge detector. *International Journal of Computer Vision*, *1*(2), 179. doi:10.1007/BF00123164

Dorninger, P., & Pfeifer, N. (2008). A comprehensive automated 3D approach for building extraction, reconstruction, and regularization from airborne laser scanning point clouds. *Sensors*, *8*, 7323–7343. doi:10.3390/s8117323

Edelsbrunner, H., Kirkpatrick, D., & Seidel, R. (1983). On the shape of a set of points in the plane. *IEEE Transactions on Information Theory*, *29*(4), 551–559. doi:10.1109/TIT.1983.1056714

Elberink, O., J, S., & Vosselman, G. (2009). 3D information extraction from laser point clouds covering complex road junctions. *Photogrammetric Record*, *24*(March), 23–36. doi:10.1111/j.1477-9730.2008.00516.x

Enemark, S., Bell, K., Lemmen, C., & McLaren, R. (2014). *Fit-for-Purpose Land Administration, Joint FIG/World Bank Publication*. Retrieved from http://www.fig.net/pub/figpub/pub60/Figpub60.pdf

Ezilon. (2009). Road Map of Vanuatu and Vanuatu Road Maps. Retrieved January 27, 2016, from http://www.ezilon.com/maps/oceania/vanuatu-road-maps.html

Farran, S. (2007). Land in Vanuatu:Moving Forward, Looking Backward, 213–223.

FIG. (1995). FIG Statement on the Cadastre. In *FIG PUBLICATION NO. 11*. The International Federation of Surveyors. Retrieved from http://www.fig.net/resources/publications/figpub/pub11/figpub11.asp

GLTN - Global Land Tool Network. (2016). Retrieved February 4, 2016, from http://www.gltn.net/index.php?option=com_content&view=article&id=437

Golovinskiy, a., Kim, V. G., & Funkhouser, T. (2009). Shape-based recognition of 3D point clouds in urban environments. In *Computer Vision, 2009 IEEE 12th International Conference on*. doi:10.1109/ICCV.2009.5459471

Holt, L., Sullivan, R., & Weaver, S. (2004). A Review of Land and Forestry Law in Vanuatu and Their Implications for Designing Forest Based Emission Trading Activities in Vanuatu, 1–28.

Hopcroft, J., & Tarjan, R. (1973). Algorithm 447: efficient algorithms for graph manipulation. *Communications of the ACM*, *16*(6), 372–378. doi:10.1145/362248.362272

Hough, P. V. C. (1962). Method And Means For Recognizing Complex Patterns. Retrieved from http://www.osti.gov/scitech/biblio/4746348

Interaction, T. H. E., Modern, O. F., Tenure, L., & In, S. (2011). *State Society and in Governance Melanesia*.

Jochem, A., Höfle, B., Wichmann, V., Rutzinger, M., & Zipf, A. (2012). Area-wide roof plane segmentation in airborne LiDAR point clouds. *Computers, Environment and Urban Systems*, *36*(1), 54–64. doi:10.1016/j.compenvurbsys.2011.05.001

Kern, P. (2006). Automated Feature Extration : A History. In *MAPPS/ASPRS*. San Antonio.

Land Equity International ty Ltd. (2016). Vanuatu Land Program. Retrieved February 4, 2016, from http://www.landequity.com.au/projects/mama-graon-vanuatu-land-program/

Lindeberg, T. (1998). *Edge Detection and Ridge Detection with Automatic Scale Selection*. *Computer Vision* (Vol. 30). Retrieved from http://www.nada.kth.se/cvap/abstracts/cvap191.html

Maltamo, M., & Packalén, P. (2007). Experiences and possibilities of ALS based forest inventory in Finland. In *ISPRS Workshop on …* (pp. 270–279). Retrieved from http://www.isprs.org/proceedings/XXXVI/3-W52/final_papers/Maltamo_2007_keynote.pdf

Monnier, F., Vallet, B., & Soheilian, B. (2012). Trees Detection From Laser Point Clouds Acquired in Dense Urban Areas By a Mobile Mapping System. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *I-3*(September), 245–250. doi:10.5194/isprsannals-I-3-245-2012

Nari, R. (2000). Land Tenure and Resource Management: a major challenge in Vanuatu. *Pacific Economic Bulletin*.

Overly, J., Bodum, L., & Kjems, E. (2004). Automatic 3D building reconstruction from airborne laser scanning and cadastral data using Hough transform. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Systems*, *34*, 296–301. Retrieved from http://cartesia.org/geodoc/isprs2004/comm3/papers/284.pdf

Peace Corps. (2007). *The Peace Corps welcomes you to Costa Rica*.

Quadros, N., & Keysers, J. (2014). *LiDAR Acquisition in the Vanuatu – Efate , Malekula & Espiritu Santo*.

Regenvanu, R. (2008). Issues with Land Reform in Vanuatu. *Journal of South Pacific Law*, *12*(1), 63–67.

Robillard, W., Brown, C., & Wilson, D. (2003). *Brown's boundary control and legal principles*. Retrieved from https://books.google.nl/books?hl=en&lr=&id=vOKAggY92-oC&oi=fnd&pg=PR7&dq=Brown%27s+boundary+control+and+legal+principles&ots=kxZ5M3Rs5S&sig=1TFlOYxG0tZa6sj_3FRf6XfcitE

Roger, J., Kain, P., & Elizabeth, B. (1992). *The Cadastral Map in the Service of the State: A History of Property Mapping*. The University of Chicago Press.

Rubio, D. O., Lenskiy, A., & Ryu, J. H. (2013). Connected Components for a Fast and Robust 2D Lidar Data Segmentation. In *Asia Modelling Symposium 2013: 7th Asia International Conference on Mathematical*

Modelling and Computer Simulation, AMS 2013 (pp. 160–165). Hong Kong. doi:10.1109/AMS.2013.31

Simbizi, M. C. D., Bennett, R. M., & Zevenbergen, J. (2014). Land tenure security: Revisiting and refining the concept for Sub-Saharan Africa's rural poor. *Land Use Policy*, *36*, 231–238. doi:10.1016/j.landusepol.2013.08.006

Sohn, G., Jwa, Y., & Kim, H. B. (2012). Automatic Powerline Scene Classification and Reconstruction Using Airborne Lidar Data. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *I*(September), 167–172.

Stoter, J. (2013). 3D Cadastre. *Fig Ladm 2013*, (September), 47–48. doi:10.1201/9781420005677

Stoter, J., Ploeger, H., & van Oosterom, P. (2013). 3D cadastre in the Netherlands: Developments and international applicability. *Computers, Environment and Urban Systems*, *40*, 56–67. doi:10.1016/j.compenvurbsys.2012.08.008

ThomasReuter. (2006). *Sharing the Earth, Dividing the Land: Land and Territory in the Austronesian World.* (T. Reuter, Ed.). ANU E Press. Retrieved from https://books.google.com/books?id=Y43gLk8IyeEC&pgis=1

Tuladhar, a. (1996). Spatial Cadastral Boundary Concepts and Uncertainty in Parcel-based Information Systems. *International Archives of Photogrammetry and Remote Sensing*.

van Beek, N. (2015). *Using Airborne Laser Scanning for Deriving Topographic Features for the Purpose of General Boundary Based cadastral Mapping.* University of Delft.

Vanuatu Ministry of Lands and Natural Resources - Land Survey. (2016). Retrieved February 9, 2016, from https://mol.gov.vu/index.php/en/land-survey/282-land-survey

Vosselman, G. (2013). Point cloud segmentation for urban scene classification. *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, *XL-7/W2*(November), 257–262. doi:10.5194/isprsarchives-XL-7-W2-257-2013

Vosselman, G., & Dijkman, S. (2001). 3D Building Model Reconstruction from Point Clouds and Ground Plans. *International Archives of Photogrammetry and Remote Sensing*, *XXXIV*, 22–24.

Wei, S. (2008). Building Boundary Extraction Base on Lidar Point Clouds Data. *Proceedings of the International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences 37*, 157–161.

Xu, S., Vosselman, G., & Oude Elberink, S. (2014). Multiple-entity based classification of airborne laser scanning data in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, *88*, 1–15. doi:10.1016/j.isprsjprs.2013.11.008

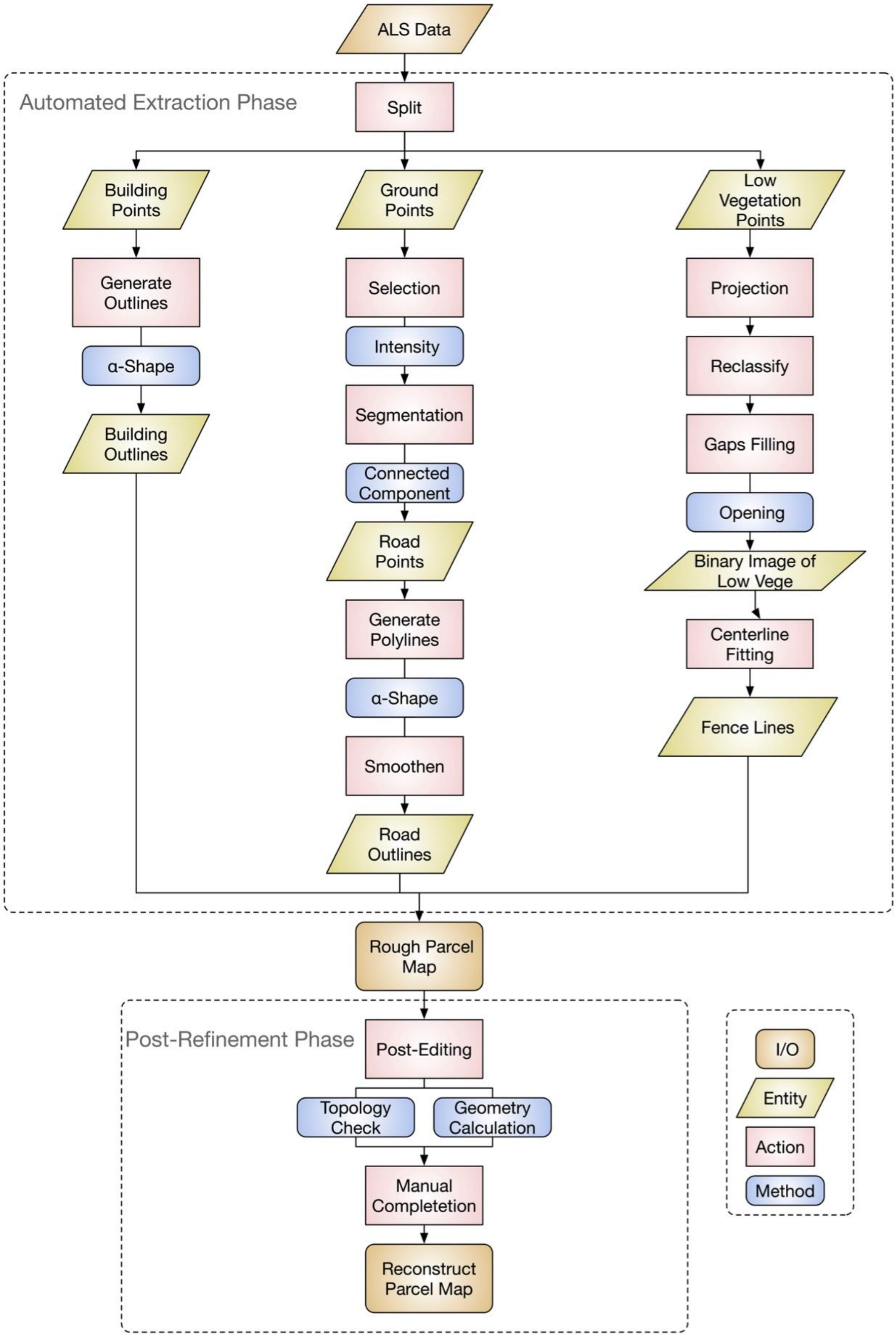
Yu, Y., Li, J., Member, S., Guan, H., Jia, F., & Wang, C. (2015). Learning Hierarchical Features for Automated Extraction of Road Markings From 3-D Mobile LiDAR Point Clouds. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(2), 709–726.

Zevenbergen, J. (2002). *Systems of Land Registration, Aspects and Effects*. *Geodesy*.

Zevenbergen, J., & Augustinus, C. (2011). Designing a pro poor land recordation system. In *FIG Working Week* (pp. 18–22).

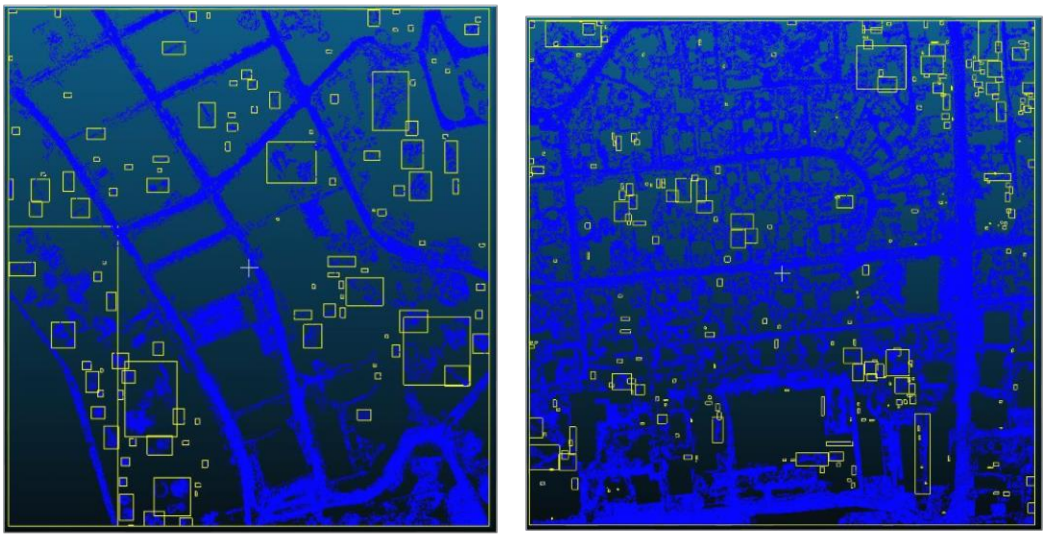
APPENDIX

1. The Semi-Automated Workflow for Cadastral Boundaries Extraction

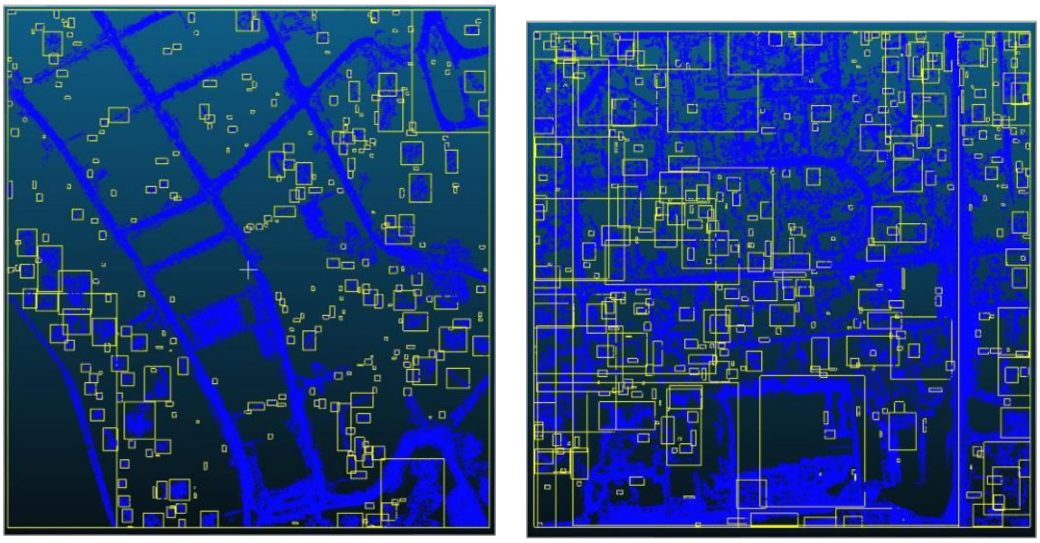


2. Performance of Different Octree Levels in Connect Component Segmentation

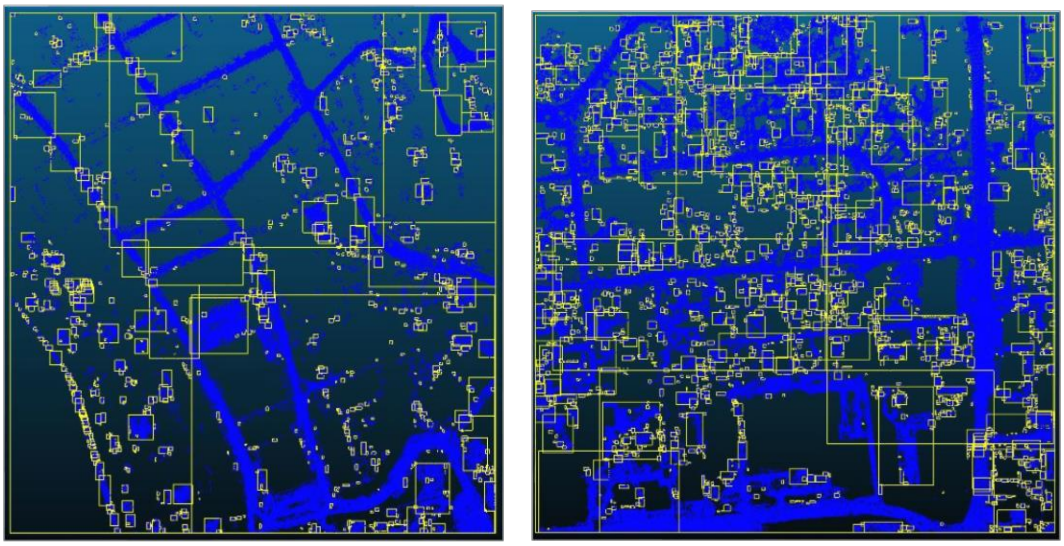
Octree=8



Octree=9



Octree=10



3. Performance of Different Radius in α-Shape Algorithm

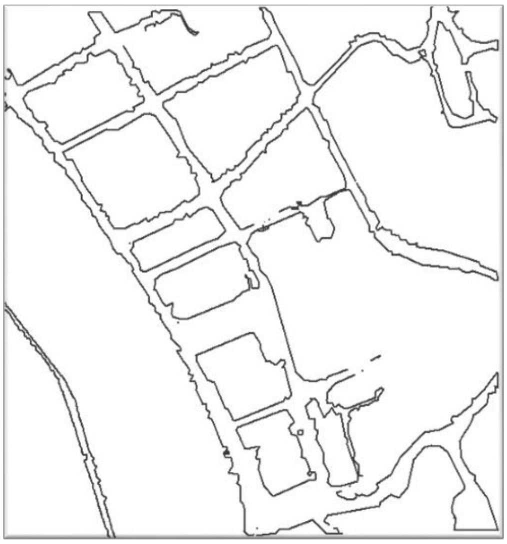
Radius=1

Radius=1.5

Radius=2

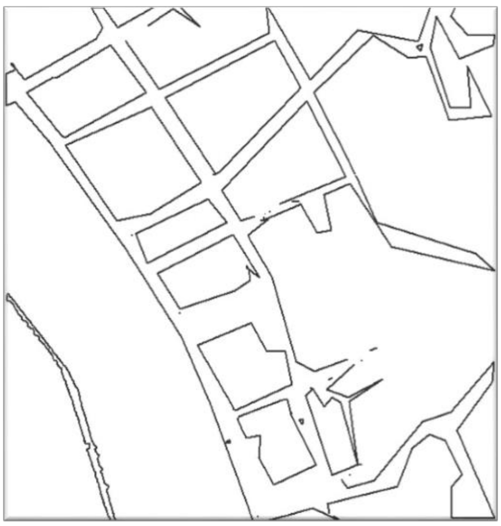
4. Performance of Different Maximum Distance in Line Simplification

MD=1

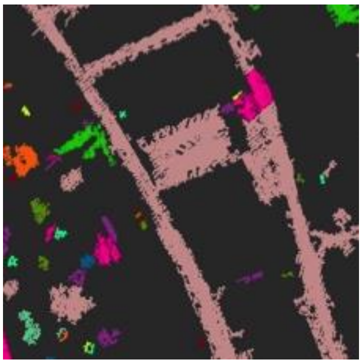
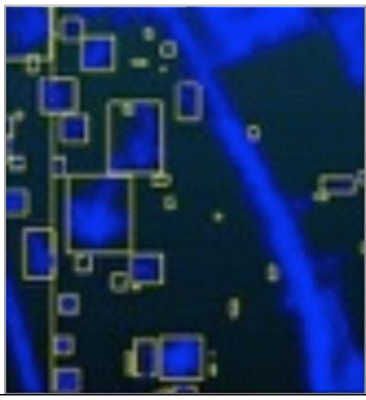
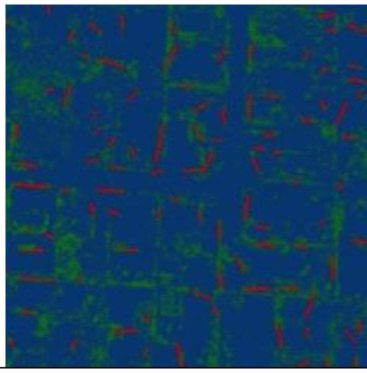


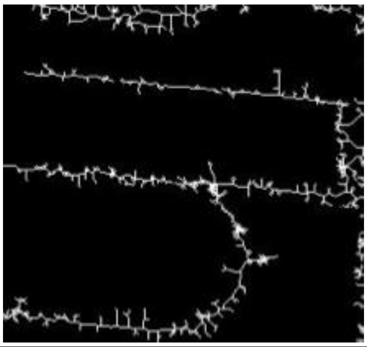
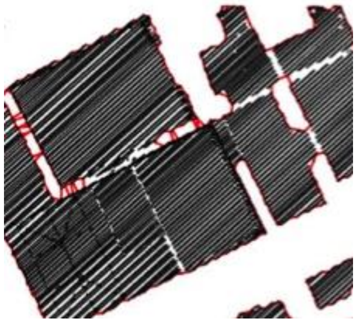
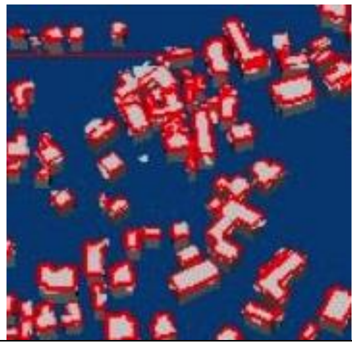
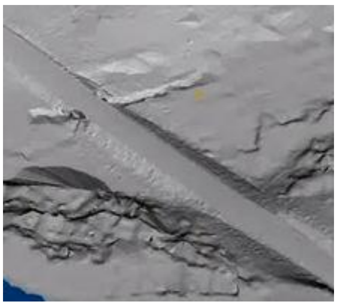
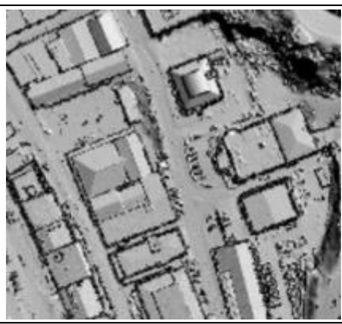
MD=2

MD=3



5. Comparison of Tested Software

<table>
<tr><th>Purpose</th><th colspan="2">Software</th><th>Sample Output</th></tr>
<tr><td rowspan="10">Segmentation</td><td colspan="2">PointCloudMapper</td><td rowspan="5"></td></tr>
<tr><td>Cost</td><td>Campus License</td></tr>
<tr><td>Function</td><td>Segmentation</td></tr>
<tr><td>Process Time</td><td>Medium</td></tr>
<tr><td>Data Type</td><td>.las</td></tr>
<tr><td colspan="2">CloudCompare</td><td rowspan="5"></td></tr>
<tr><td>Cost</td><td>Open source</td></tr>
<tr><td>Function</td><td>Segmentation</td></tr>
<tr><td>Process Time</td><td>Fast</td></tr>
<tr><td>Data Type</td><td>.las</td></tr>
<tr><td rowspan="10">Line Generation</td><td colspan="2">VRMesh</td><td rowspan="5"></td></tr>
<tr><td>Cost</td><td>Commercial</td></tr>
<tr><td>Function</td><td>Point-tracing</td></tr>
<tr><td>Process Time</td><td>Fast</td></tr>
<tr><td>Data Type</td><td>.shp</td></tr>
<tr><td colspan="2">ArcScan</td><td rowspan="5"></td></tr>
<tr><td>Cost</td><td>Campus License</td></tr>
<tr><td>Function</td><td>Vectorize</td></tr>
<tr><td>Process Time</td><td>Medium</td></tr>
<tr><td>Data Type</td><td>.shp</td></tr>
</table>

	Matlab			
		Cost	Campus License	
		Function	Skeleton	
		Process Time	Medium	
		Data Type	.svg	
Outline Delineation	Matlab			
		Cost	Campus License	
		Function	Outline	
		Process Time	Medium	
		Data Type	.svg	
	VRmesh			
		Cost	Commercial	
		Function	Outline	
		Process Time	Fast	
		Data Type	.shp	
Terrain Visualization	VRmesh			
		Cost	Commercial	
		Function	Meshed Surface	
		Process Time	Fast	
		Data Type	.tif	
	LasTools			
		Cost	Free License	
		Function	Hillshade	
		Process Time	Medium	
		Data Type	.tif	

5. Research Matrix

Objective	Questions	Method	Results
1.To identify feature types that needed to be extracted.	1. What are features being cadastral boundaries? 2. What are possibility of each kind of feature? 3.How is priority of target features? 4. What are parameters of these features?	Statistic from cadastral map and satellite images (ArcGIS, Office)	1. Number of extractable cadastral boundaries; 2. Priority of target features; 3. Parameters of target objects; 4. Suitable point density.
2.To design a proper workflow of extraction.	1.What is percentage of features can be extracted by workflow? 2.What parameters are suitable for classification of workflow? 3.Which outline generation algorithm is suitable for planar objects? 4.How to extract linear objects? 5. How to post edit?	First Classification, then generate planar feature and linear feature centerline, finally refine by Post editing (LasTools, PCM, CloudCompare, Matlab)	1. Target feature 2. Outlines of planar surface (Roofs, roads) 3. Linear features (fence) 4. Rough parcel map
3.To test and quantify the extraction result.	1. How is classification accuracy? 2. What is completeness? 3. What are correctness and precision? 4. What is degree of automation? 3. What are strength and weakness? 4. What are recommendations for improvement?	Compare with existing cadastral map (ArcGIS, Office, CloudCompare)	1. Tolerance of positional accuracy 2. Completeness 3. Correctness 4. Degree of automation