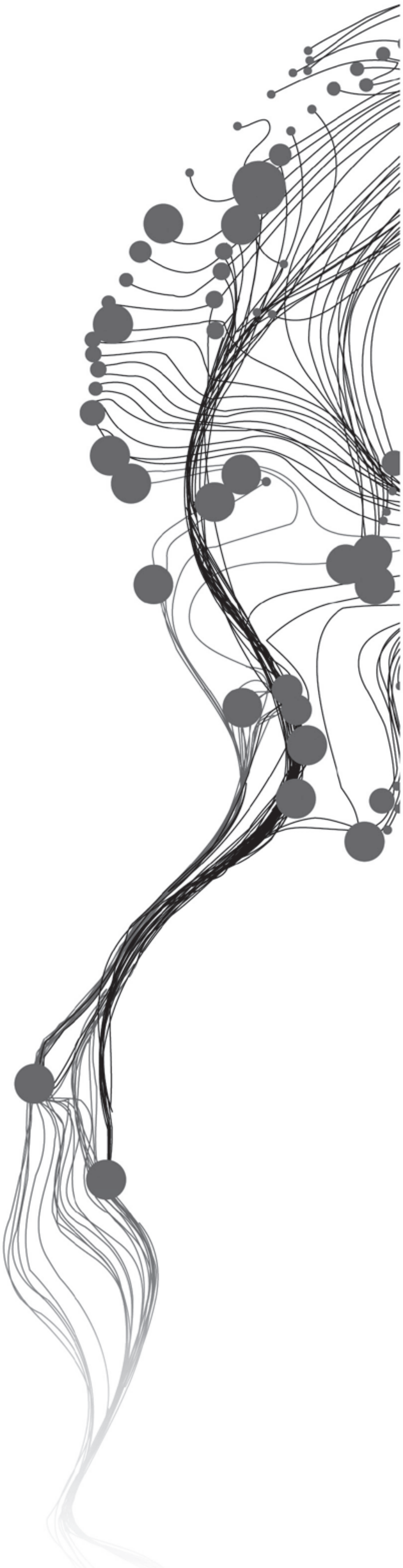
DETECTING CHANGES IN TREES USING MULTI-TEMPORAL AIRBORNE LIDAR POINT CLOUDS

WEN XIAO
February, 2012

SUPERVISORS:
Prof. Dr. Ir. M.G. Vosselman
Dr. Ir. S.J. Oude Elberink

ADVISOR:
Sudan Xu



DETECTING CHANGES IN TREES USING MULTI-TEMPORAL AIRBORNE LIDAR POINT CLOUDS

WEN XIAO
Enschede, The Netherlands, February, 2012

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: Geoinformatics

SUPERVISORS:
Prof. Dr. Ir. M.G. Vosselman
Dr. Ir. S.J. Oude Elberink

ADVISOR:
Sudan Xu

THESIS ASSESSMENT BOARD:
Prof.Dr. M.J. Kraak (Chair)
Dr. R.C. Lindenbergh (External Examiner, Delft University of Technology)
Prof. Dr. Ir. M.G. Vosselman
Dr. Ir. S.J. Oude Elberink
Sudan Xu MSc

ABSTRACT

Changes in vegetation are of great interest since they play crucial roles in ecosystem monitoring where remotely sensed data has been proven extremely profitable. Digital change detection approach has been widely utilised in conventional remote sensing technologies. As a relatively new technology, light detection and ranging (lidar) provides a promising way of change detection of vegetation in three dimensions (3D) because the laser beam will penetrate through the foliage generating point clouds with highly accurate 3D coordinates.

This research proposes a method for vegetation change detection in 3D with high level of automation. Three epoch datasets are classified into several predefined classes including high vegetation (trees). A connected components algorithm is applied to group the points of a tree together because the point clouds are unstructured. The attributes of components are used to discriminate tree components from other since a few non-tree points are misclassified. Points from neighbour trees might be clustered together, so a local maxima algorithm is implemented to distinguish single tree components with multiple tree components. After that, the parameters of trees are derived through two independent ways: point based method which refers to 3D alpha shapes and convex hull; model based method which utilises the Pollock tree model for single trees. Then the changes can be detected by comparing the parameters of corresponding tree components which are found by a tree to tree matching algorithm.

The comparison of these two methods illustrates the consistency and stability of the parameters. The detected changes clearly show the growth and pruning of trees. The results are visualized by point cloud mapper (PCM) and statistically analysed.

Keywords: change detection, high vegetation, 3D modelling, airborne lidar, point cloud

ACKNOWLEDGEMENTS

I would love to express my deepest gratitude and appreciation to my supervisors, Prof. Dr. Ir. M.G. Vosselman and Dr. Ir. S.J. Oude Elberink, for their continuous supports and brilliant suggestions. Their positive comments helped me to build my confidence and encouraged me for the challenges. This thesis could not have been written without their elaborate supervisions. It was my great pleasure and honour to work with them.

My special thanks go to my advisor, Sudan Xu who really helped me a lot during the entire process. She is kind, intelligent, thoughtful and hardworking and she has set a great example as a good researcher and thinker to me.

I sincerely thank all the GFM lecturers, who helped me to upgrade my knowledge and improve my skills of learning. I would like to specially thank Dr. Khoshelham, for his constructive criticism during proposal and mid-term defences and Dr. Rossiter, for his concern about my work and his helpful English courses.

I am grateful for the scholarship from Erasmus Mundus. And genuine thanks go to EMECW project manager, Dr. Tiejun Wang who selected me for the scholarship so that I have this precious opportunity to study at ITC.

I would like to thank Biao Xiong, Liang Zhou, Qifei Han, Ding Ma, Fangning He, Mostafa and Daniel for the advice and supports for my research. Together with the latter three, we four laser guys have had really great times together. Also I deeply appreciate the great friendships and good or bad times I spent with all GFM colleagues and fellows from other departments.

I also want to thank many many people that helped me during my study and life. Especially the Chinese fellows: Fengfan Yang, Yanbin Jiang, Peng Wang, Yang Chen, Bingbing Song, Fan Shen, Zhi Wang, Chenxiao Tang, Shaoqing Lu, Zheng Yang, Dong Yang, Chao Yang, Xiao Jing, Qiuju Zhang, Xiaojing Wu, Fangfang Cheng, Donghai Zheng, Xinping Ye, Xuelong Chen, Xuanmei Fan, Zeng Yijian, Xiaogang Ma, Xi Zhao, Xiang Zhang, etc. I love this big family abroad.

At last, I cannot thank my dear parents and lovely sister enough for their supports, encouragements and endless love.

This thesis is dedicated to my dear parents XIAO Qichun and WU Zengqun.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1. Motivation and problem statement

Change detection is a process that observes the differences of an object or phenomenon at different times (Singh, 1989). The change detection technique is of significant importance in many environmental and ecological studies such as land use change monitoring, deforestation analysis, damage assessment, study of changes in vegetation phenology, disaster monitoring and so on. Accurate change detection of the earth surface (e.g. urban areas, vegetation areas) provides essential information for better planning and management. Planners should be clear about the changes to identify the situation and the state of the environment, and then make sustainable decisions. Change detection has become a major application of remote sensing techniques which provide viable data of repetitive coverage at short time and consistent quality.

Especially, changes in vegetation covered areas are of great interest because they are crucial for ecosystem monitoring where digital change detection method is widely used (Coppin et al., 2004). Change detection can be applied to monitor the speed of deforestation and analyse the effects of measures taken by the authorities. Besides forests, vegetation in urban areas is a vital part of the living conditions. The proportion of vegetation covered areas is an essential factor for urban planning. Urbanization and industrialization will severely affect the growth of vegetation. Therefore the changes of vegetation should be monitored and estimated. In practice, land cover changes in urban area have drawn many attentions, and in some cases vegetation are specially considered (Chen et al., 2005).

Many technologies have been used to detect and monitor changes; especially the application of remotely sensed data has been proved extremely successful. Digital images are most commonly utilised, not only remote sensing satellite images (Yasuoka et al., 1990) but also digital aerial photographs. Airborne visible/infrared imaging spectrometer (AVIRIS) data was successfully used for detecting changes in vegetation in two-dimension (2D) (Elvidge and Portigal, 1990). Besides, Hoffer and Lee (1990) detected the change of forest cover using satellite radar data. Taking advantage of remotely sensed data, changes can be detected more efficiently and automatically with less labour intensity and time cost.

As a relatively new remote sensing technology, airborne laser scanners (ALS) provide a promising way of change detection of vegetation in three-dimensional (3D) perspective. With traditional remotely sensed data, change detection of vegetation is described as coverage rate change of forests or woods, thus is usually carried out in 2D. By measuring the shadow in aerial images, changes in the vegetation height can be estimated. Nonetheless, this method can be quite inaccurate. Thanks to the light detection and ranging (lidar) technology, change detection research is stepping from 2D to 3D and becoming more reliable and automatic. The main advantage of ALS compared with traditional techniques is that it generates point clouds with accurate 3D coordinates. Using high density point clouds, the changes in both coverage and height can be detected (Houldcroft et al., 2005; Yu et al., 2006). Moreover, not only forests but also small bushes and even single trees can be detected from dense point clouds, which can be used to estimate the carbon storage (Kim et al., 2010) and generate 3D tree models (Pratihast, 2010; Rutzinger et al., 2011; Vosselman, 2003).

In terms of the importance of monitoring changes of vegetation in urban areas and the development of high density laser scanners, lidar shows its great potential to detect changes in high vegetation (Murakami et al., 1998; Rutzinger et al., 2007). Vegetation changes in forestry at plot level, such as biomass or average height, have been studied (Yu et al., 2004). Lidar data processing, e.g. feature extraction, classification, has been discussed a lot and some researches have taken vegetation into consideration. However, no research has focused on tree changes in 3D in urban areas. For a certain urban area, multi-temporal data sets can be obtained by ALS. Then the majority of the vegetation can be identified after classification. Nevertheless, the results of classification have to be improved. In addition, 3D modelling of trees, automatically detecting and quantifying the changes are still problems.

1.2. Research identifications

1.2.1. Research objectives

The general objective of this research is to detect the changes of high vegetation, which refers to trees, in two epoch's lidar point clouds in urban area. To accomplish that, four sub-objectives should be fulfilled sequentially.

1.2.1.1. Classification results verification

Lidar data are points with coordinates in 3D and other attached information (e.g. intensity, pulse count). Therefore in order to detect changes of vegetation, points that belong to vegetation should be identified, which is quite easy for human visual interpolation but not the case for computers. The datasets used in this research have already been classified and more than 96% of the vegetation points were obtained after classification. But commission and omission errors remain in the data. So the result of the classified vegetation should be verified and improved.

1.2.1.2. Tree parameters derivation

In urban area, trees are the majority of vegetation and treated as important index for urban planning, not like low vegetation, e.g. bushes, grass land, which are more likely considered as land cover. Therefore, only trees are considered as the research objects in this research. Single tree is most common such as a line of trees along a road. A couple of trees or woods usually appear in parks or outer side of the city. In order to detect the changes of trees in 3D, parameters like height, crown area and crown volume of every single tree are essential. After identified as a tree, the points should be grouped together so the parameters can be derived from the components.

1.2.1.3. Change detection in two data sets

After all the previous works, the datasets are to be compared. This is the main objective of this research. The changes should be categorized and quantified. Trees that are cut or newly planted should also be treated as changes.

1.2.1.4. Quality assessment

The results must be verified. The quality of the change detection results should be assessed so that the methods and algorithms in the whole process can be evaluated and revised.

1.2.2. Research questions

For each of the sub-objective mentioned above, there are several corresponding research questions.

1) Classification verification
   How to eliminate the misclassified points or segments?
   Which features can be used?
2) Parameter derivation
   How to cluster the points of a tree together as an object?
   What parameters can be used to do change detection for single trees?
   How to derive parameters from grouped trees?
3) Change detection
   How to match the corresponding trees in two data sets?

What are the change categories?

4) Quality assessment
   How to verify the change detection results?
   How to assess the performance of the algorithms?

1.2.3. Innovation

Airborne lidar technique provides high accurate height information and the laser beam can penetrate through the vegetation thus points on and inside the canopy and even ground information can be obtained. So it has been widely used to estimate the biomass in forest area, including the change detection of tree growth. However, barely work of change detection of vegetation in urban area has been done before using airborne lidar data. Besides, trees are never exactly the same even in two strips because of scanning geometry and wind effect, so stable parameters are essential for change detection. The innovation of this research is mainly to detect the changes of high vegetation in urban area using airborne lidar data, including:

- Method and procedure for vegetation change detection with a high level of automation.
- Stable parameter derivation of trees for change detection in 3D.

1.3. Thesis structure

This thesis is organized into six chapters.
First of all, the background is introduced and the problems are stated. And the objectives, questions and innovation of this research are identified.
Chapter two describes the review of relevant literatures. The state of the art of laser scanning is explained first. Then the research about change detection of vegetation using remote sensing data sources is discussed.
Chapter three introduces the research methodology. The framework was depicted first and then each step is described in detail corresponding to the research objectives.
Chapter four illustrates the details of data processing including the value of each threshold. The flowchart of the programme is displayed and the results of each step are visualized.
Chapter five statistically analyses each step and also the final change detection results. Discussions based on the analysis are followed.
The last chapter describes the conclusions drawn on the study, and makes some recommendations for future research.

2. LITERATURE REVIEW

2.1. State of the art of airborne lidar

2.1.1. Principle of lidar

Lidar is a relatively new remote sensing technology using laser for measuring purposes. A laser scanner (see Figure 2-1) is the core of the scanning system, thus the term laser scanning is also widely used. The mechanism of airborne laser scanning is time-of-flight meaning the time of pulse flying in the medium is recorded. Then the range can be obtained because light travels with a known constant velocity.
Beside the laser scanner, an airborne laser scanning system also combines global positioning system (GPS) and inertial navigation system (INS) and all of them are mounted on an aircraft. A GPS ground base station is also necessary for differential GPS (DGPS) to improve the accuracy of the system. The whole system looks very similar to photogrammetry. The accuracy of height is only few centimetres. When the flying height is low, few centimetres accuracy can also be achieved horizontally.

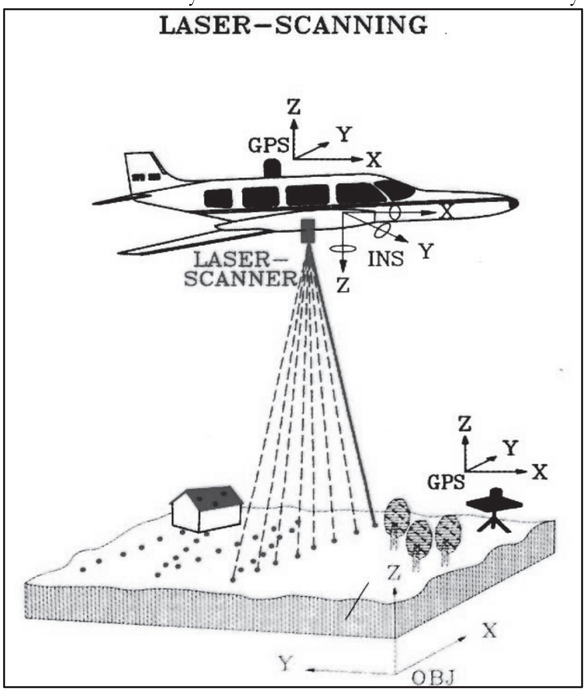


Figure 2-1: Airborne laser scanning (Pang, 2012)

Instead of imagery obtained by conventional remote sensing techniques, the data of laser scanning are unstructured points in 3D, thus also are called point cloud. The standard lidar data format is .LAS. Multiple echoes and full-waveform laser scanner has been developed in recent years. Along with 3D coordinates, other information e.g. reflectance strength, pulse count and even true colours can also be recorded. These kinds of information play crucial roles in data processing and application.
Compared with conventional data acquisition techniques, e.g. aerial photogrammetry, lidar has its own advantages and disadvantages. As an active system, laser scanning is applicable even at night. The points acquired by laser scanner have very high accuracy especially in vertical direction and the point density can be higher than 30 points per square meter (pts/m^2) nowadays. Most importantly, the beam of laser may penetrate through the foliage of vegetation thus not only the points on the surface or inside the vegetation

but also the information on the ground can be captured. So the lower stories and the structure of the vegetation can be studied and also the elevation model under the vegetation can be generated. This is why lidar has been widely utilised in forestry.

Despite the advantages, there are also some drawbacks of using lidar. Water surface or objects with water on top may cause data missing when the scan angle is not very small. Even the point density becomes higher and higher with the development of laser scanners, texture of the data is still not as informative as imagery. Break lines can be estimated by fitting and tracking intersecting planes (Vosselman and Maas, 2010). Roof edges can be detected with the help of aerial images which are often obtained simultaneously with the laser scanning data.

2.1.2. Point cloud processing

2.1.2.1. Pre-processing

To identify the vegetation in lidar point clouds, filtering, segmentation and classification are the prerequisite steps.

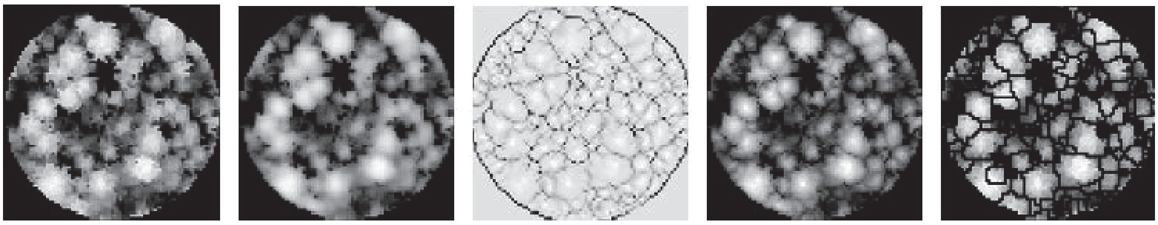
Several mathematical morphology filtering algorithms can be found, among which slope based filtering (Vosselman, 2000) can filter the ground out in more complex terrain situation. In order to avoid reconstruct objects based on massive unstructured points, segmentation is usually applied beforehand. 3D Hough transform or random sample consensus (RANSAC) can be performed to detect a local set of co-planar points as a seed segment, then a bigger segment can be found through surface growing.

Information extraction from lidar data, e.g. buildings, roads and bridges (Maas and Vosselman, 1999; Oude Elberink and Vosselman, 2009; Sithole and Vosselman, 2006), shows a great potential of detecting vegetation in urban areas. Feature-based classification and object-based classification methods both have been utilised in airborne lidar data where feature-based can divide into point-wise and segment-wise classification. A number of attributes of segments are commonly utilised, e.g. the number of points in a segment (segment size), the normal distribution, the average height of the points inside, the percentage of last echo if available, etc.. Moreover, full-waveform and spectral information have been proven profitable for classification of airborne lidar data.

2.1.2.2. Vegetation extraction and delineation

Vegetation has very irregular and unique point distribution in laser scanning data. Rutzinger, et al. (2007) detected high vegetation in urban areas using airborne lidar data. First of all, the original lidar data was segmented then the attributes of segments were calculated to discriminate vegetation segments from others so that vegetation was classified and labelled.

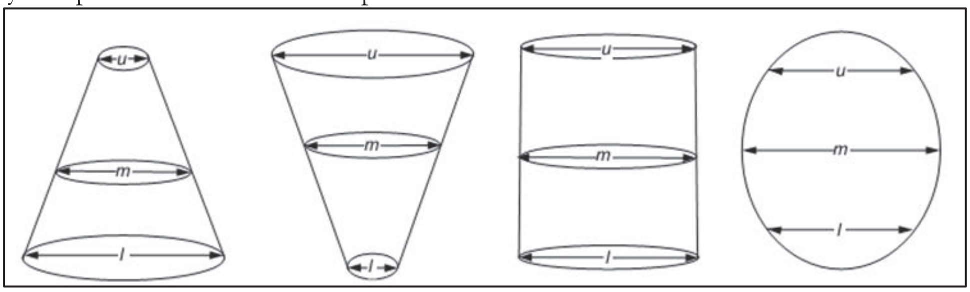
Parameters of individual tree can be generated directly from the point clouds. Yu et al. (2011) developed an approach for extracting individual tree attributes, i.e. height, diameter at breast height (DBH) and stem volume, based on 26 geometrical and statistical features derived from airborne lidar data. A height image based individual tree detection method (see Figure 2-2) was demonstrated and random forests technique was utilised for estimation. A canopy height model (CHM) was smoothed first, and then the minimum curvatures were computed. The image was scaled then the local maxima were found as single tree crown. The method was tested using 1476 trees in Finland. Correlation coefficients between the observed and predicted values were 0.93, 0.79 and 0.87 for the three attributes mentioned above respectively.



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

Figure 2-2: Individual tree detection, (a) original CHM, (b) smoothed image, (c) minimum curvature image, (d) stretched image, (e) watershed segmentation (Yu, et al., 2011)

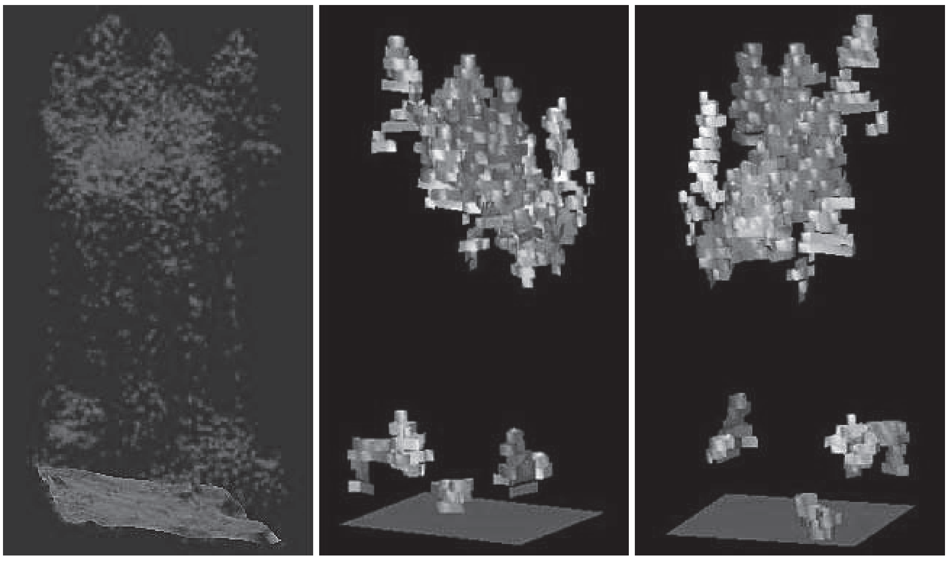
What is more, 3D tree modelling in lidar data becomes a hot topic recently. Models used in traditional remote sensing techniques and computers science are commonly utilised in lidar data. A fixed shape model or individual tree-wise models were applied in both mobile laser scanning (MLS) (Pratihast, 2010) and airborne laser scanning (ALS) data. Rutzinger et al. (2011) utilised four different crown shapes (see Figure 2-3) having different diameters at three height levels using 2D enclosing circles. Then trees were modelled by an open source framework OpenAlea.



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

Figure 2-3: Crown shape types, (a) conical, (b) inverse conical, (c) cylindrical, (d) spherical (Rutzinger, et al., 2011)

Wang et al. (2008) analysed the vertical canopy structure of forest and also modelled trees in 3D. A voxel based method for individual trees delineation was implemented at different height levels. Then tree crowns were modelled as Figure 2-4, and several crown parameters, e.g. tree height, crown height, crown diameter and volume can be derived.



(a) (b)

Figure 2-4: Virtual Reality Modelling Language (VRML) prismatic model, (a) raw lidar data, (b) VRML models from two views (Wang, et al., 2008)

Vosselman (2003) detected trees in airborne lidar data by computing the local maxima with a detection rate of 97%. The tree crown was modelled using a fixed shape (see Figure 2-5) whose diameter was adaptive to the height of the local maximum.

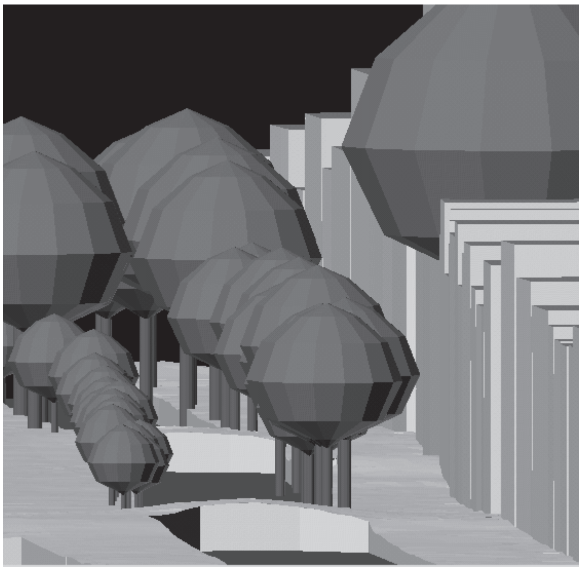


Figure 2-5: Fixed shape tree models (Vosselman, 2003)

Instead of regression models, Kato et al. (2009) developed a 'wrapped surface reconstruction' method. Tree parameters, e.g. tree height, crown diameter, crown base and volume were derived by the wrapped surfaces (see Figure 2-6). And the results were validated by comparing with total station surveyed field measurements.

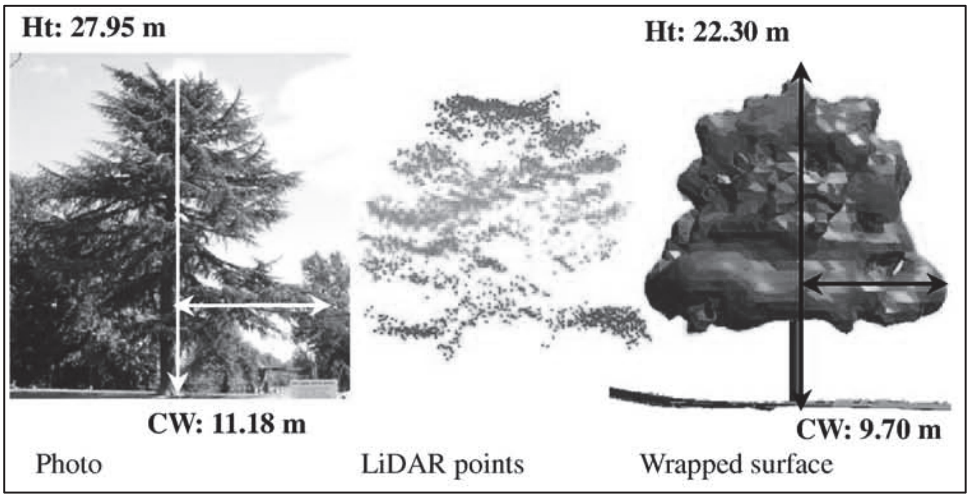


Figure 2-6: An example of wrapped surface comparison (Kato, et al., 2009)

2.2. Change detection of remote sensing data

2.2.1. Change detection approaches

Change detection techniques have been developing for decades. Several reviews have categorized change detection approaches from different perspectives.

Singh (1989) generalized the basic approaches into two, i.e. classification based comparison and comparison directly on raw multi-temporal data. Lu et al. (2004) summarized ten different applications of change detection and grouped the change detection approaches into seven categories, i.e. algebra, transformation, classification, advanced modelling, geographic information system (GIS) integration, visualization and others. Both the advantages and disadvantages of the approaches were discussed and the degree of complexity was ranked. Based on previous categorizations, Gong et al. (Gong et al., 2008) classified change detection approaches into two general groups, i.e. bi-temporal change detection and temporal trajectory analysis regarding the data sources that used. Furthermore, they also organized the algorithms into seven categories, i.e. direct comparison, classification, object-orientated method, modelling, time-series analysis, visualization and hybrid method (see Figure 2-7).

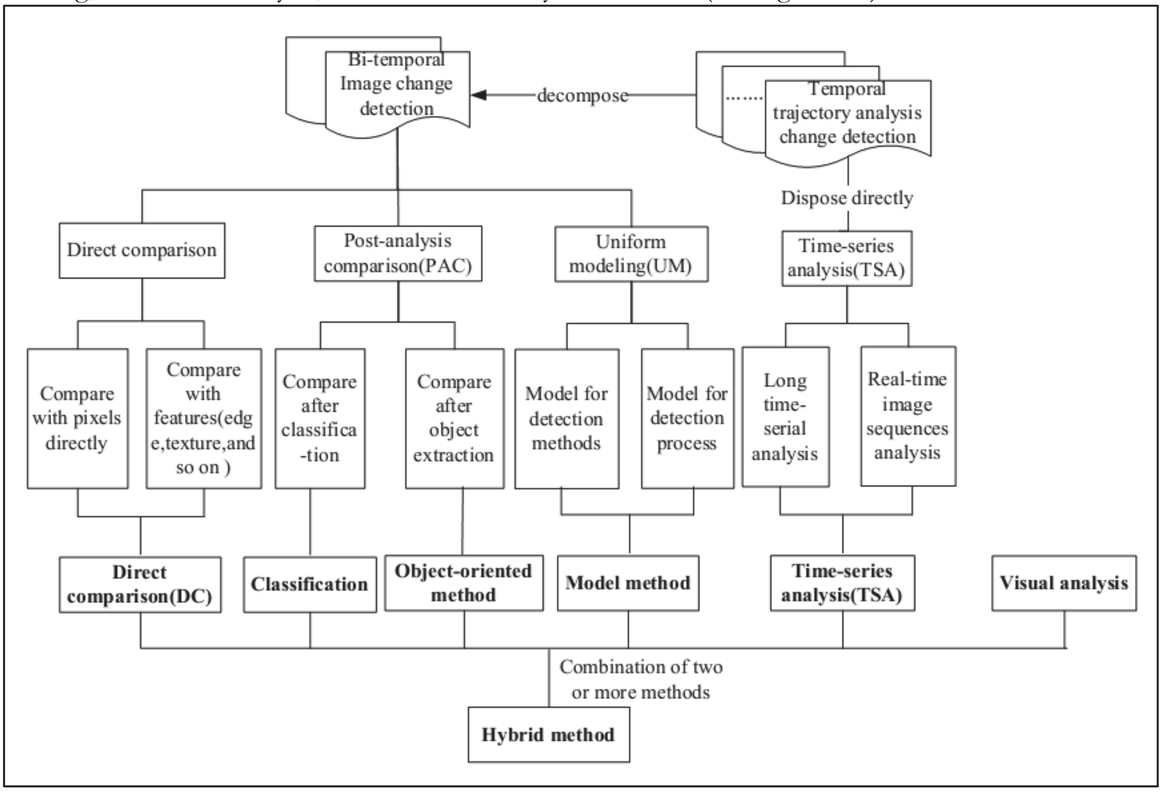


Figure 2-7: Change detection categories (Gong, et al., 2008)

2.2.2. Change detection of lidar data

Change detection using conventional remote sensing techniques is still popular nowadays. Some works also focused on changes of vegetation in urban areas utilising Landsat imageries (Chen, et al., 2005). The potential of ALS system for change detection in urban area was discussed early on by Murakami (1998), who also detected the changes of buildings using ALS data (Murakami et al., 1999).

Vosselman et al. (2004) applied change detection of lidar data for updating medium scale map. First, the laser data was segmented and classified. Then the segments of buildings were matched against the building objects of a topographical database. In the change detection experiment, all newly constructed buildings were detected reliably.

Choi et al. (2009) presented a feature based approach to automatically detect changes in urban areas. The main processes are first detecting changed areas through the subtraction of two DSMs, then segmenting and classifying the changed patches to predefined classes, ground, vegetation and building, and last determining the types of changes based on the classes and properties. This method was able to detect the changes in a sufficient degree of accuracy semi-automatically. However, there was no quantitative evaluation.

The multi-temporal lidar data analysis and change detection studies in forest area are also studied in many researches (Hyyppä et al., 2004). Yu et al. (2004) detected harvested trees and forest growth using airborne lidar data. The estimation of height growth was accomplished by individual tree delineation and a tree to tree matching algorithm. First, trees were located by detecting the local maxima in the CHM, and then the crown shape was determined by watershed segmentation. Last, the trees were matched by the location with a threshold distance at 0.5m. The precision was about 5cm and 15cm at stand level and plot level respectively based on field and statistical analysis. Individual tree height growth was detected again by Yu et al. (2006), who presented three change detection manners, i.e. differentiation between DSMs and CHMs, canopy profile comparison and analysis of height histograms. Hausdorff distance, the maximum distance from a point in one set to the closest point in a different set, was applied to improve the tree to tree matching result.

2.3. Summary

Airborne lidar point clouds have obvious advantage for vegetation analysis. Due to the penetration of laser beams, the information upon, inside and even below vegetation areas can be captured.
Several classification approaches have been applied to lidar data. Knowledge based method is quite popular since the attributes of objects differ from each other. Especially, vegetation points have their own patterns and features.
Vegetation structures and parameters have been extracted utilising both point based method and model based method. Before that, individual tree delineation results still need to be improved. Different models of trees have been studied. Nonetheless, a stable model for change detection is still necessary.
Different change detection approaches have been applied to detect changes in lidar data. The combination of several methods is mostly adopted for vegetation analysis.

3. RESEARCH METHODOLOGY

3.1. Introduction

This research is an experimental research, in which methods and algorithms are developed to process the datasets. For each phase, the result is assessed and the methods and algorithms are revised. As for the verification of classification results, in order to eliminate points that belong to other classes and noises, certain features are selected and several try-outs are carried out. Then proper parameters are chosen to represent the trees. The results of parameters, which are verified through two independent methods, rely on the procedures above. Algorithms are modified and retested if the verification does not perform well. In brief, this research is composed of repeated procedures of problem analysis, algorithm developing, experimental test, result assessment. However, perfect results are not necessarily to be achieved in the end.

3.2. Framework of the methodology

The procedure of the methodology is introduced with respect to the four sub-objectives, including four steps, i.e. vegetation verification, tree parameters derivation, tree oriented change detection and quality assessment. The overall framework is depicted as Figure 3-1.

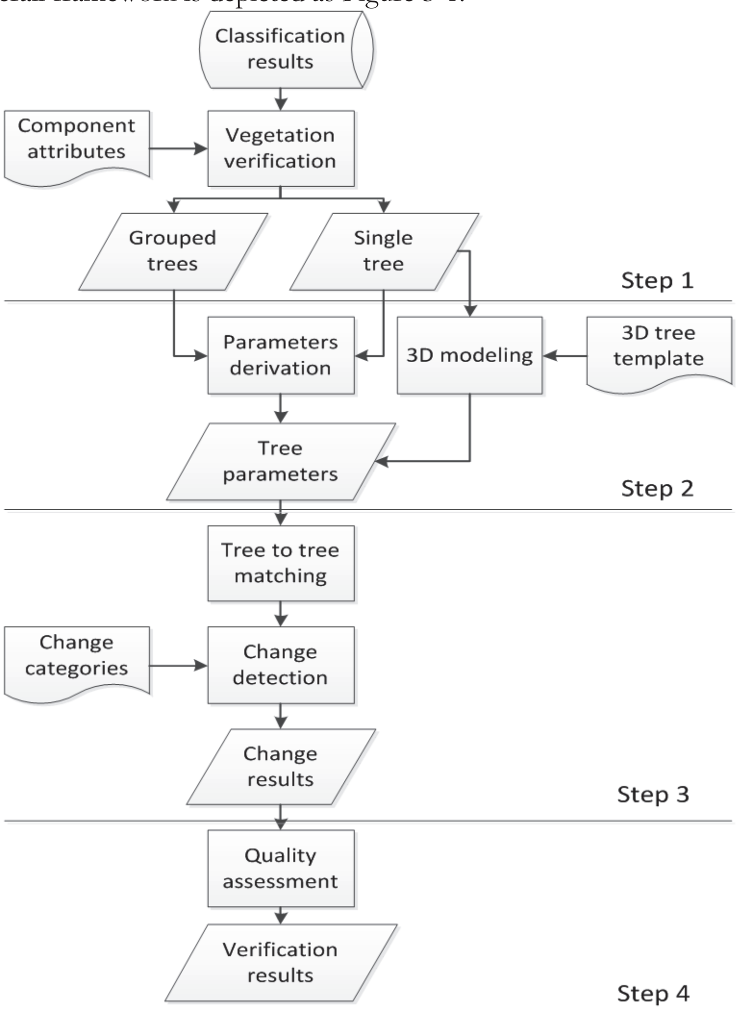


Figure 3-1: Framework of the methodology

3.3. Vegetation verification

The datasets used in this research have been classified into several predefined classes in which vegetation is one of them. But the classification results have to be improved and several kinds of misclassifications are to be corrected. Also noises in the scene should be removed.

Normally, misclassification is categorized as commission and omission errors. By visual inspection, several kinds of commission errors can be recognized in the data set. Small segments like walls, roofs of complex shape on buildings as well as cars, poles and even some ground points are classified as vegetation. On the other hand, some vegetation points are classified as other classes like buildings and ground, which is omission error. Points that were not classified should remain in the data set because most of them are vegetation points.

The overall accuracy of classification reached to 98.1%, while the completeness was 96.9% and the correctness was 97.8% (Xu et al., 2012). The omission errors are a few misclassified points which are minority with respect to the point amount of a tree. So the error will hardly affect the parameter derivation result. This part of the research focused on commission error.

3.3.1. Connected components

To eliminate the misclassified vegetation points and to group the points of a tree together as a component, connected components algorithm was implemented. It can not only cluster the points of a tree together, which is fundamental for the following work, but more importantly, it provides a way to verify the classification results using the attributes of the components. Segment based method was used for the data classification beforehand in the way that the attributes of the segments vary from different classes. For instance, the normal of wall segment is often parallel to the ground but not for other objects like tree segments. Same as segment based classification, the attributes of component can also be used to distinguish vegetation components from others.

After classification, most of the misclassified components were very small fragments. Therefore, the size of a component (number of points inside a component) was quite an efficient way to differentiate trees from other fragments.

The second attribute used was the height span of a component (the distance from the lowest point to the highest point). As we know, the points of a tree are normally from the ground or from the bottom of the crown to the top of the crown if the trunk was not scanned. So the height span of a tree component should be no more than the real height of that tree. A certain threshold can be determined by the base knowledge of the height of trees, and then components that have greater height span than an upper threshold or smaller height span than a lower threshold can be removed from the dataset.

In order to remove components from high buildings or points hanging in the air, the minimum height of a component can be utilised. If the lowest point of a component is higher than a normal tree, then this component can be eliminated. Except the above attributes, normal distribution of the components might also be helpful because some of the fragments have regular shapes.

In addition to the geometry attributes, spectral information, e.g. reflectance, intensity and true colour are also useful. For example, the reflectance of trees is usually different from other objects like building roofs, so the average reflectance of a tree component differs from a roof component.

3.3.2. Local maxima

Trees are normally planted with a certain distance, but as trees grow, they will become connected with each other. After implementing connected components, these kinds of trees will be connected together. So the dataset will have both single trees and grouped trees.

For single trees, the parameters can be derived directly and also can be fitted by a mathematical model. But for grouped trees, further process might be necessary. So it is important to find out the number of trees in each component, or at least label the component as single tree component or multiple tree component, namely the component contains one or more than one tree.

To accomplish that, a point based local maxima algorithm was implemented. For each component, the highest point within a certain range, which was adaptive to the component height, was found. So the number of highest points of a component was regarded as the number of trees in that component. For each point in a component, the distances from other points were calculated, and the height was compared with its neighbour points that were within a threshold range. If there was no higher neighbour point found, this point was the local maximum.

3.4. Tree parameter derivation

After the first step, the dataset was composed of a group of components which contained single trees or grouped trees. In order to simplify the procedure of parameter derivation, the trunks were filtered first and only the crowns were left. For multiple tree components, the areas and volumes of the crowns can be calculated by 3D alpha shapes algorithm. For single tree components, the areas and volumes can also be computed by 3D alpha shapes or convex hull (same as 3D alpha shapes when $\alpha \rightarrow \infty$). Also they can be fitted by Pollock model. Then the parameters will be generated by the fitting.

3.4.1. Trunk removing

Because of the occlusion by the leaves and the scanning geometry, some of the trees have sufficient points on the trunks but others do not. This inconsistency will affect the process of parameter derivation and the result of change detection when in one epoch dataset a tree has trunk but in the second epoch dataset the same tree has no trunk. So it is better to remove the points on the trunk before parameter derivation.
For single tree components, an interval of height was moved upwards from the bottom of the component, so the component was cut into slices. For each slice, together with the previous slice below, a 2D bounding box was computed (see Figure 3-2). If the hypotenuse of the bounding box is smaller than a predefined threshold, the points within the slice were considered as trunk points. The bounding box will keep moving upwards until the hypotenuse is greater than the threshold. In the end, a reference height will be obtained. The points above the height belong to the crown and the points below belong to the trunk.

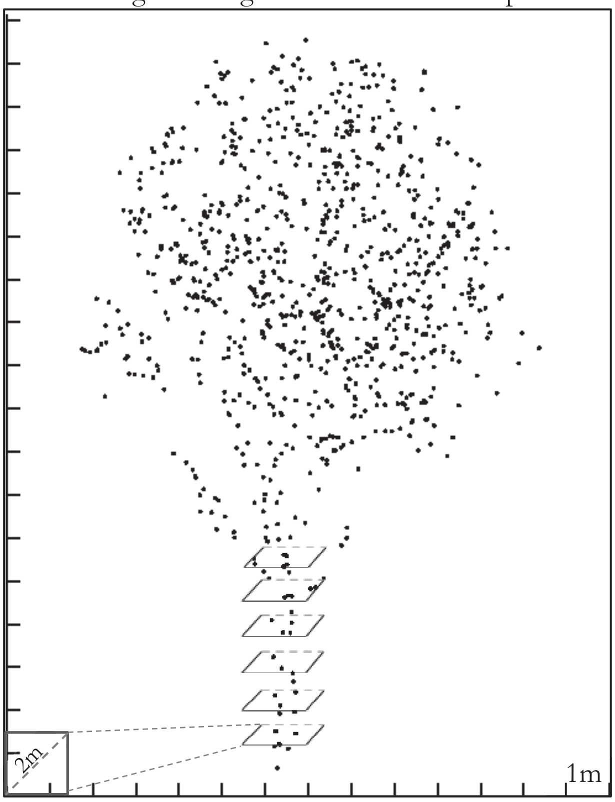


Figure 3-2: Sketch of trunk removing

The general idea for removing trunks of multiple tree components was quite the same as above, the only difference was that the multiple trunks within the slice were separated using connected components algorithm again. Then the bounding box was generated for each trunk component. If the trunk component was smaller than the threshold, the slice would keep moving upwards until each bounding box reached the crown.

3.4.2. 3D alpha shapes

Alpha shapes algorithm is famous of shape reconstruction from a dense unorganized set of points. Indeed, an alpha shape is a linear approximation of the original shape (Da et al., 2011). The definition of alpha shapes is based on an underlying triangulation (usually a Delaunay triangulation). Then for 2D alpha shapes, circles with a certain radius (alpha) will try to pass through the data points until they touch points on the edges of the triangles. As shown in Figure 3-3, the edges touched with circles describe an approximate shape of the original points. For 3D alpha shapes, a triangulation is calculated first, and then spheres, instead of circles in 2D, with a certain radius (alpha) will be generated based on the outer side triangles that are on the boundary. So the triangles that are touching spheres will represent the original shape of the data points.

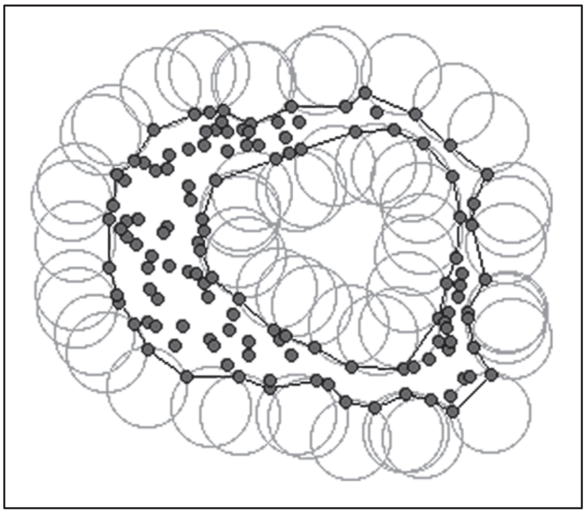


Figure 3-3: 2D alpha shapes (Da, et al., 2011)

The points belonging to the vertices of the triangles will be sufficient to describe the shape of a tree, so the points inside the alpha shape can be eliminated. The area and volume of points can be calculated through the alpha shape; moreover the data amount will be reduced significantly.
For multiple tree components, to reduce the effect of the connection of trees, the alpha values were optimized according to the shape of each component. The optimized alpha value is defined as a smallest alpha value such that the alpha complex satisfies the following properties: the number of solid component of the alpha complex is equal to or smaller than a limited given number which should be one in this research. Then the areas and volumes of grouped treed components can be extracted from the alpha shapes.
For single tree components, a bigger alpha value is necessary because most of these components are small trees and optimized alpha shapes of small trees still have points inside the crown. In this case, the volume of the alpha shapes is actually much smaller than the real crown volume. Also if the alpha values are quite different from each other of the same tree from two epoch datasets, the change detection result will also be affected. To avoid the problems stated above, a consistent and big enough alpha value should be set for single tree components. If the alpha value is positive infinity ($\alpha \rightarrow \infty$), the alpha shapes are the same as convex hull. So the areas and volumes of single tree components are derived by convex hull.

3.4.3. 3D tree modelling

3D alpha shapes and convex hull are both point based parameter derivation method which has both advantages and disadvantages. Point based method is straight forward, simple to process but sensitive to noises and outliers. On the contrary, model based method is more stable. These two methods can be compared so as to assess the quality of the results. In this research, single tree components were modelled and the results were compared with point based method.

3.4.3.1. Adjusted Pollock model

A 3D model template should be able to represent the shape of trees which may be diverse from individual to individual. There are many models and approaches can choose from. A fixed shape model, e.g. elliptic paraboloid, ellipsoid or sphere, can be used to simulate the shape of crown in a simple way (Vosselman, 2003). But to be more precise, a parametric model developed by Pollock (1994) was used to model trees of different shapes. The shape of the crown can be adjusted by a parameter from an ellipsoid to a cone:

$$\frac{z^n}{a^n} + \frac{(x^2 + y^2)^{\frac{n}{2}}}{b^n} = 1 \tag{3-1}$$

The origin is the centre of the circle of the crown, and z-axis points vertically upwards. In the equation a is the radius of the intersection of surface with the z-axis, b is the radius of the circle of the crown and n is a positive real number that determines the shape of the crown surface. When n=2 the surface is an ellipsoid, and as n decreases to 1 the surface becomes a cone (see Figure 3-4). As n decreases from 1 to 0, the surface becomes increasingly concave.

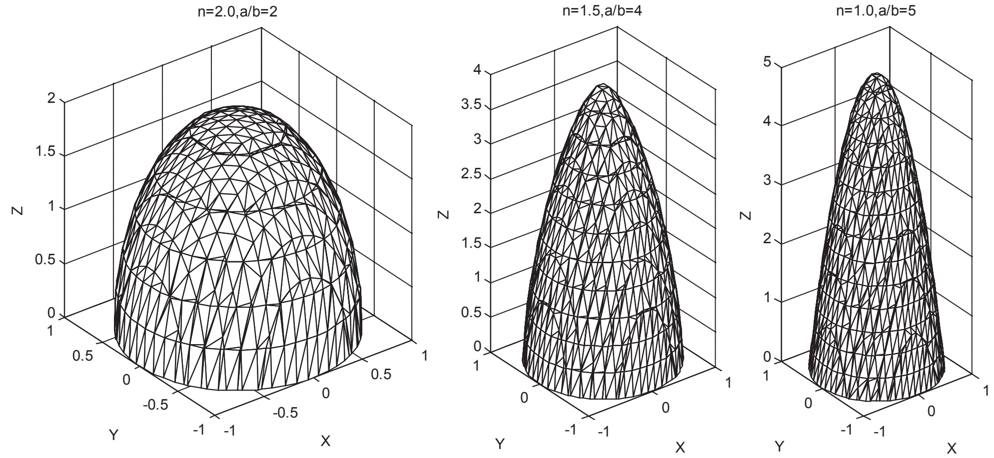


Figure 3-4: Crown shapes with different parameters (Pollock, 1994)

To make the model more realistic, the base was changed to a ellipse instead of a circle, and then a rotation in x-y plane and the shift of coordinates were added.

$$\frac{z^{\,n}}{c^n} + \left(\frac{x^2}{a^2} + \frac{y^2}{b^2}\right)^{\frac{n}{2}} = 1 \tag{3-2}$$

$$x = (X - X_0)\cos\beta - (Y - Y_0)\sin\beta \tag{3-3}$$

$$y = (X - X_0)\sin\beta + (Y - Y_0)\cos\beta \tag{3-4}$$

In which, a, b are the two semi axes of the crown base ellipse, c is the semi axis in z direction, n is still the real number that determine the crown shape, X_0, Y_0, Z_0 are coordinates of the origin in the global coordinate system and β is the rotation angle.

3.4.3.2. 3D model fitting

Point density and distribution will affect the modelling results on the same tree, so, as mentioned above, in order to eliminate the effects of different alpha values, alpha value was set to 10m (fixed)(Rutzinger, et al., 2011) instead of optimized value for each component. After that, the point density became consistent and most of the points lied on the boundary of the tree crown but not inside.

The adjusted Pollock model was implemented using nonlinear least square fitting in three steps: 2D crown base fitting, upper crown fitting and lower crown fitting. The crown was divided into upper crown and lower crown because they have different shapes. For bigger trees that have sufficient points on the surfaces of the crowns, the results of directly 3D crown nonlinear least square fitting was not bad, but for the small tree that have only a few points on the crown, the shape of the model was quite far from the reality. For this reason, the crown base (a, b, β) was estimated first in 2D, then a and b were used as constraints for the 3D Pollock model fitting.

To be more specific, principle component was used to transfer the points from the global coordinate system to a local coordinate system. Next a convex hull was implemented to derive the outer points which were then fitted by an ellipse. The initial values of the two axes were the half length of the point range in x-y plane. So after the fitting, a, b, β were obtained. The height used to distinguish the upper crown and lower crown was determined by the average height of the points on the convex hull, so Z_0 was known now. For the upper crown, initial c was the distance from the highest point to the crown base (Z_0). Initial X_0, Y_0 were the interquartile values of all the points on the upper crown. And initial n was set to 1.5 because it represented a shape in the middle of ellipsoid and cone. For the lower crown, the position of the origin was determined by the upper crown fitting so X_0, Y_0 were constant. Initial c was the distance from the lowest point to the crown base and the initial n was also 1.5.

After the 3D model fitting, the height of the tree was the height of the crown origin plus the vertical axis ($Z0 + c$). The area of the crown was the area of the fitted ellipse (πab) and the volume of the crown was the sum of the volumes of the upper crown and lower crown.

3.5. Tree oriented change detection

3.5.1. Tree to tree matching

After deriving the parameters of the trees in each data set, the parameters can be compared. But before the comparison, the corresponding trees should be identified based on the locations of trees in both epochs. The datasets are in the same local coordinate system, so the corresponding trees should be at the same location. Yu et al. (2004) applied tree to tree matching method with the threshold distance at 0.5m. In this research, the tree to tree matching was done by finding the corresponding points of the same tree in each epoch dataset.

This point based matching was accomplished by calculating the overlapping of bounding boxes and point to point distances. First of all, for each of the components in dataset 1, a bounding box was derived. Then the overlapping bounding box in dataset 2 was searched. To further check whether these two bounding boxes were the corresponding components, the distances from points in dataset 1 to points in dataset 2 were calculated. If the number of distances that were smaller than 1m was greater than a certain percentage of the smaller size of the two components in each dataset, it meant some of the points of the two components were quite near from each other thus these two components were matched.

Because it happened that some of the components were grouped trees components but in the other epoch dataset the trees were not connected, this led to the situation that several components in dataset 1 corresponded to one component in dataset 2. And that was why the number of nearby points should be compared with the smaller one between the two components.

After processing all the components in the two dataset, the corresponding components were given the same label in both dataset. If there were no corresponding components found in the other dataset, these kinds of components were not labelled.

3.5.2. Change detection

Since the corresponding trees were found, the changes can be detected just by subtract the parameters of dataset 1 from dataset 2. Four categories were introduced to analyse the changes: cut, newly plant, area change and volume change. The latter two are further divided into increase and decrease. Considering the errors of pre-processing and modelling, every tree pair will be slightly different even if no changes were happened. So a very small proportion of change should be taken care of. The categories are shown as Table 3-1.

Table 3-1: Change detection categories

Categories	Cut	Planted	Area change		Volume change	
Change	Only in data1	Only in data2	Area ↑	Area ↓	Volume ↑	Volume ↓

As explained in the above section, the corresponding relation between the components in two datasets might be many-to-one. In this case, the parameters (area, volume) of components in dataset one will be added up and then compared with the component in the other dataset. Since grouped trees components have more than one tree inside, it is meaningless to compare the change of height as a whole component. So the tree height was not compared.

3.6. Quality assessment

The classification verification results can be verified by visual inspection because human eyes can discriminate trees from others in point clouds. Because of lack of ground truth data, completeness and correctness can be assessed based on the original datasets.

The results of local maxima algorithm might not be very accurate because of the irregular combination of trees. But what has to be assessed is that whether the component is a single tree component or a grouped tree component because this will affect the following computation and comparison, especially when a two trees component has only one local maximum. One way to check the quality is also visual inspection.

Two independent methods, point based and model based, were utilised for parameters derivation, thus the results were assessed by comparing these two methods. For multiple tree components, the comparison with single tree components which have the same size as both of the grouped trees will somehow suggest the consistency and accuracy.

Generic knowledge helps to assess the quality of previous methods and the change detection results. For example, trees of the same size nearby will grow quite the same as each other. If the changes of same size trees were quite different, the accuracy of the parameters might be doubtful. Also younger trees grow faster than older trees, so the percentages of the change should show this tendency.

3.7. Summary

In this chapter, a framework of the methodology corresponding to the objectives was presented. The four steps of the methodology were then discussed in detail sequentially. Several algorithms were introduced in each of the first three steps and, quantitative and qualitative methods were both given for quality assessment.

The datasets were already classified, so the original data for this research were unstructured points classified as vegetation which actually contained many other non-tree points. Connected components algorithm was applied first to group the points so the points of an object will be clustered together. Also the attributes of the components were utilised to eliminate the non-tree components. Then the local maxima of each component were found in order to distinguish the single tree and multiple tree components. To derive the parameters of trees, the trunks were removed because they might affect the results. Then two independent methods, point based and model based, were implemented. Point based method used 3D alpha shapes and convex hull algorithms and the model based method fitted an adjusted Pollock model by nonlinear least square fitting. The parameters were assigned to each component for comparison. After that, corresponding trees were matched by the bounding boxes of components and point to point distances. So the changes can be detected directly by comparing the parameters of the components. In the end visual inspection, comparison of methods and generic knowledge were proposed to assess the quality of the results.

4. POINT CLOUDS PROCESSING AND VISUALIZATION

4.1. Introduction

Point clouds were processed according to the methodology using DEV C++, MATLAB and a mapping library. A programme dealing with all the algorithms was developed. The flowchart of the programme is depicted as Figure 4-1.

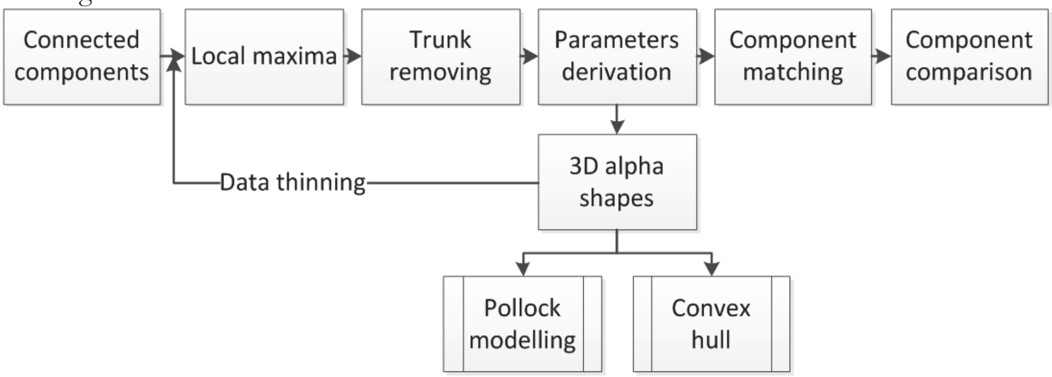
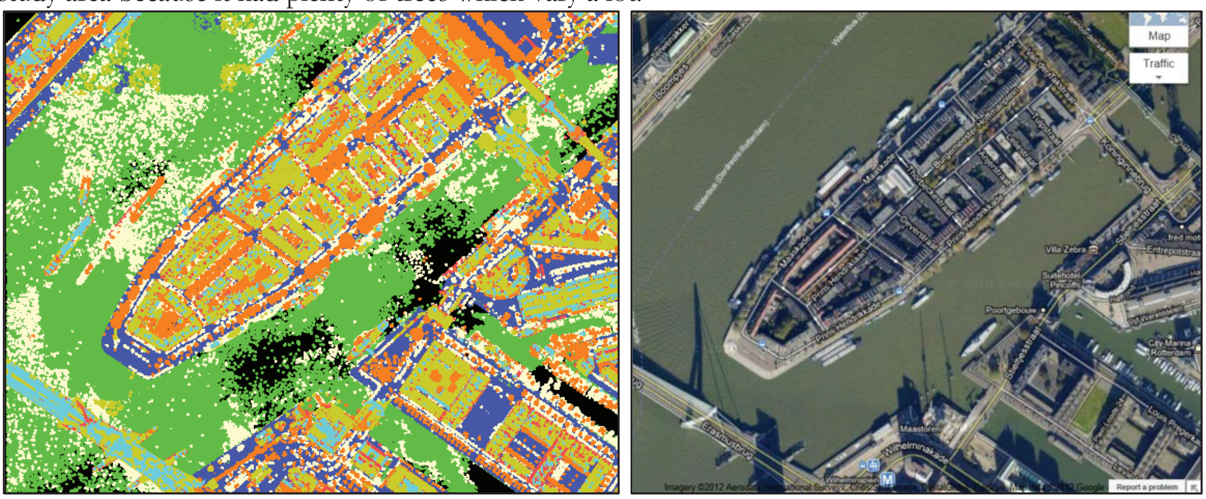


Figure 4-1: Flowchart of the data process programme

4.2. Three epoch datasets

Three datasets were obtained on behalf of municipality of Rotterdam, two of them were obtained in 2008 (March and November) and the third one in April 2010. They are both under the Dutch coordinate system. Point density of the data in March of 2008 is around 10-15 pts/m^2, while the other two are about 30-50 pts/m^2.

A part of the small island (Noordereiland) along the river in Rotterdam (see Figure 4-2) was selected as study area because it had plenty of trees which vary a lot.



(a) (b)

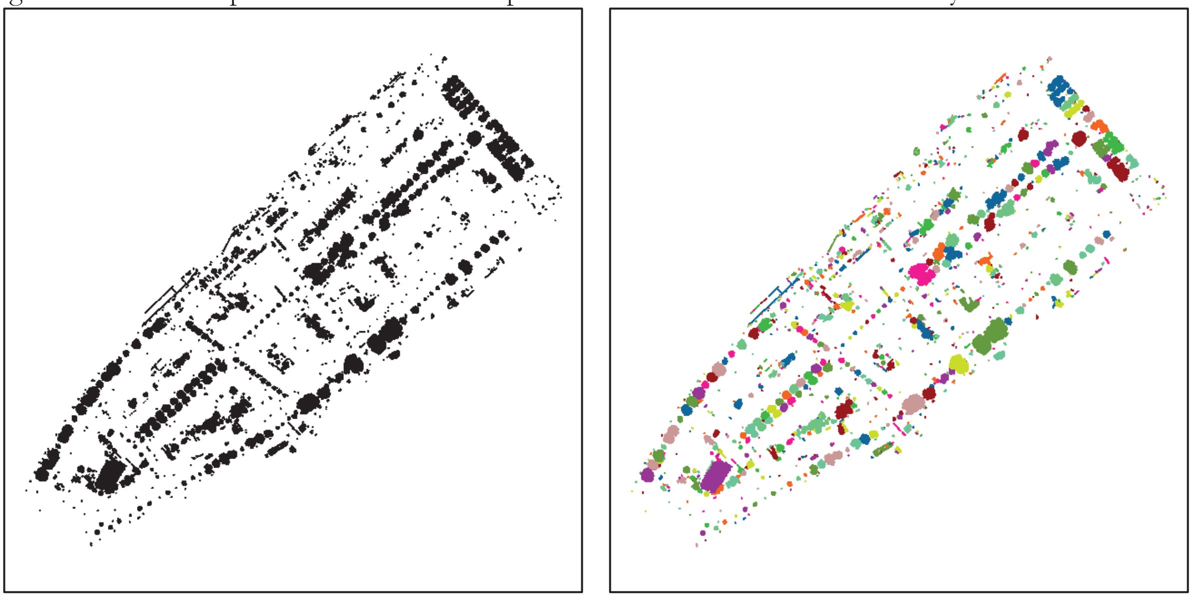
Figure 4-2: Study area in Rotterdam, (a) Lidar points, (b) Google Map image

4.3. Vegetation verification

4.3.1. Connected components

4.3.1.1. Points clustering

First, all the vegetation points were extracted from the original datasets. And then the connected components algorithm was implemented. A Kd-tree was generated on the points and edges between points and their neighbours greater than 1m were removed, so the points still connected were treated as components. Then each of the components was labelled by a component number. Figure 4-3 (a) shows the points classified as vegetation which were unstructured, (b) shows the result of connected components algorithm. Each component had its own component number which was illustrated by colour.



(a) (b)

Figure 4-3: Connected components results, (a) before connected, (b) after connected

4.3.1.2. Non-vegetation components removing

As shown in Figure 4-3, many non-tree components were remaining in the dataset. To remove them, some of the geometric and spectral attributes were utilised such as component size, height span, minimum height, colour, reflectance, normal distribution, plane residuals and so on. Some of the thresholds were assigned based on generic knowledge and others were determined by experimental tests.

1) Component size
   Component size was the number of points within a component. Small segments can be removed by this attribute. The threshold was proportional to the point density. Components with less than 100 points were assumed as fragments and removed from the dataset when the point density was 30-50 pts/m^2.
2) Height span
   Height span was calculated as the distance from the highest point to the lowest point in a component. Components that have height span less than 3m or more than 25m were treated as non-vegetation components. Since this research was focus on high vegetation, bushes that lower than 3m were removed. Big segments parallel to the ground were also filtered out.
3) Minimum height
   Components with lowest point higher than 15m are not trees. Because some of the components were fragments from high buildings, so this attribute helped to remove big segments hanging in the point cloud.

4) Colour
   True colour of the components can be used to discriminate trees from others, especially the green channel. If the average colour value of a component was greater than a threshold, it will be treated as vegetation component.
5) Reflectance
   Vegetation has smaller reflectance than other urban objects like buildings or roads in the datasets. So the components with extremely big average reflectance values were removed. Based on several experimental tests, the threshold was set to 200.
6) Normal
   Some of the non-tree components were fragments from man-made constructions like building walls having certain normal directions, so a plane was fitted to each component. If the normal was horizontal or vertical, the component was most likely a non-vegetation component. However, a vegetation component might also have a horizontal or vertical normal depending on the point distribution, so the results were further checked.
7) Plane residual
   Residuals of vegetation components were normally quite big, so the average residual of each component was calculated. Components were removed if their average residuals were smaller than 0.3.

Geometric attributes showed the efficiency for removing fragments that dislike the shape of trees. Spectral information facilitated the filtering of vegetation-liked components. Figure 4-4 shows the result of non-vegetation components removing.

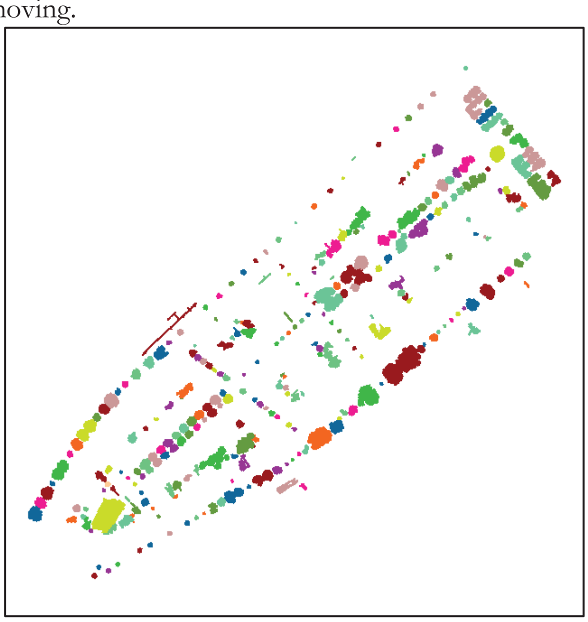
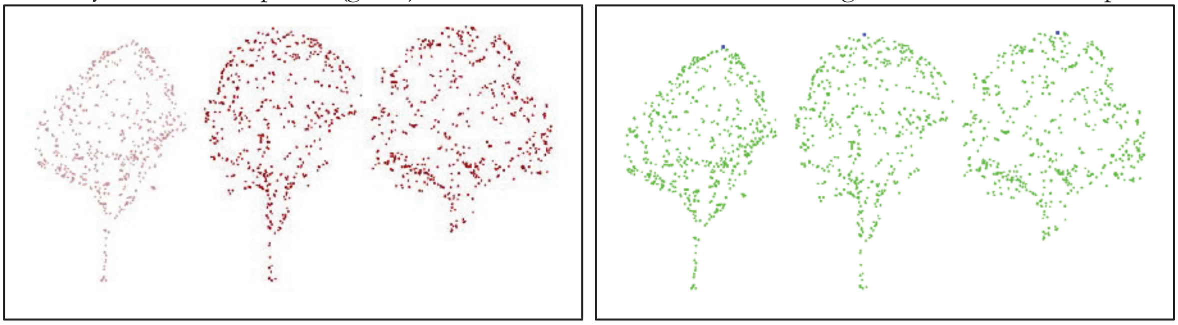


Figure 4-4: Components removing result

4.3.2. Local maxima

As shown in Figure 4-4, some components contain more than one tree inside. In order to find the number of trees in each component so that single tree components and multiple tree components can be treated differently in the following process, the local maxima of each component were found. The highest point of each component within a certain range was considered as a local maximum. The threshold of the range was set to a quarter of the absolute height of each component according to experimental tests. To reduce the cost of computation, the 3D alpha shapes algorithm was applied first to thin the data. Then a subset of the data was selected because even after components filtering some non-vegetation or bush components, which will affect the change detection results, still remained in the data. Thus the tree components along the roads were selected manually.

In the result shown as Figure 4-6, two trees on the right were in the same orange colour meaning they belonged to a same multi-tree component. A local maximum (blue) was detected on each tree and labelled differently with the rest points (green). The number of the maxima was assigned back to each component.



(a) (b)

Figure 4-5: Local maxima results, (a) single tree and multi-tree components, (b) local maxima (blue)

Figure 4-6 shows the number of local maxima in each component. Different colours represent different number of local maxima. Reddish yellow colour means one local maximum in the components and the other colours means multiple local maxima within the components.

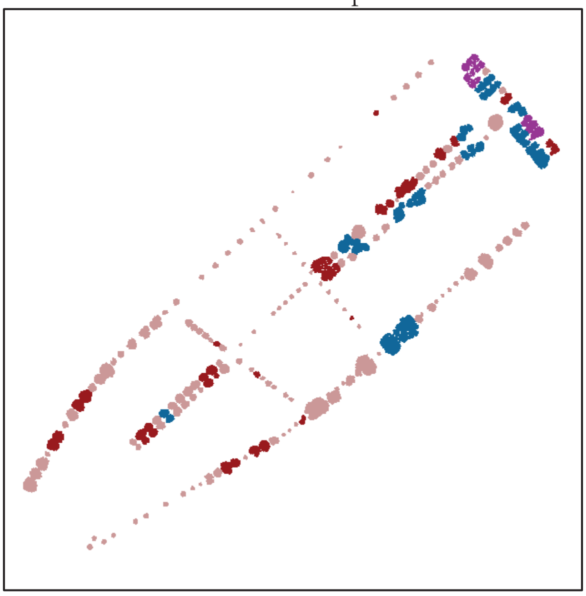


Figure 4-6: Local maxima label results (reddish yellow-1; dark red-2; blue-3; purple-5 components)

4.4. Tree parameters derivation

4.4.1. Trunk removing

To remove the trunks, components were cut into slices from bottom upwards. The height span of the slice was defined as 0.2m for single tree components. For each interval, points in it and in the previous interval below (0.4m in total) were considered. Based on experimental tests, the threshold of the bounding box was set to 2m which was quite big because small branches at the bottom of the crown should be removed otherwise they will significantly affect the volume. Figure 4-7(a) and (b) show the result of trunk removing for single tree components.

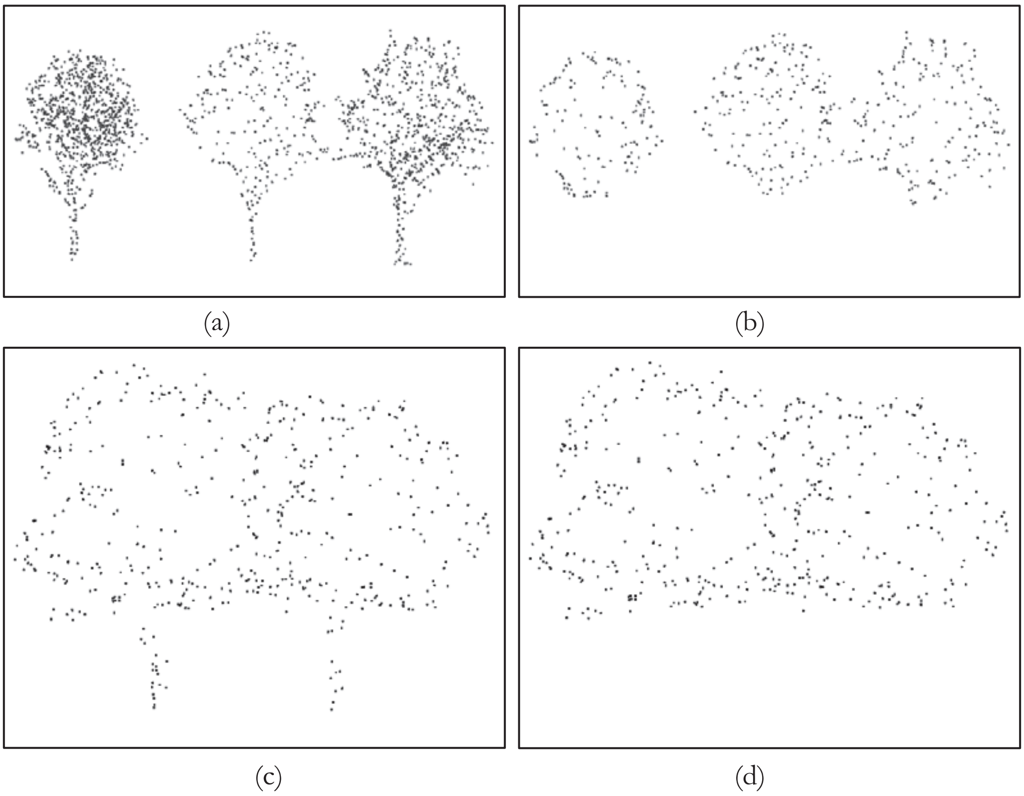


Figure 4-7: Trunk removing, (a) single tree components before trunk removing, (b) after trunk removing, (c) multiple tree component before trunk removing, (d) after trunk removing

As for multiple tree components, connected components algorithm was used again however the maximum edges in the algorithm was set to 3m because the data were sparse after data thinning. The threshold of the bounding box was also 2m. Figure 4-7 (c) and (d) show the result of trunk removing for a multiple tree component.

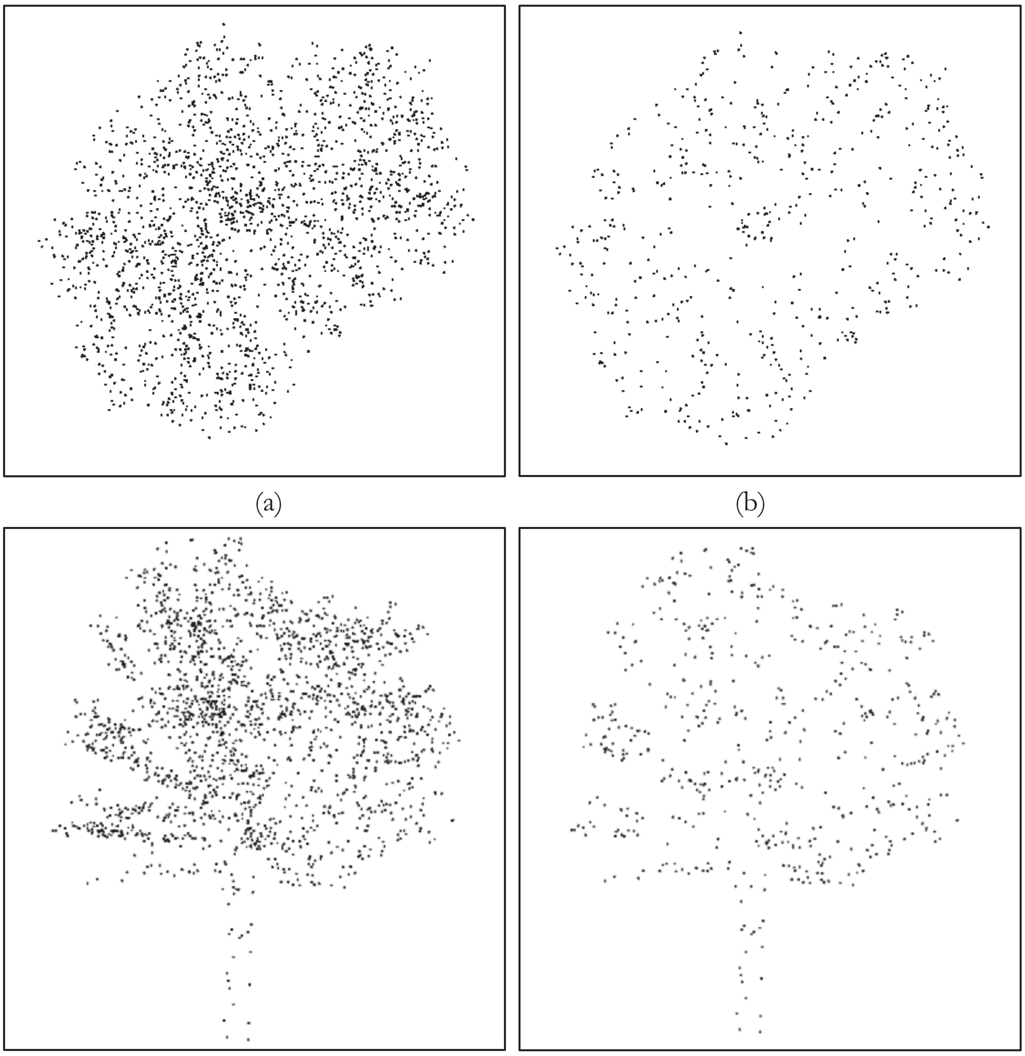
4.4.2. 3D alpha shapes

3D alpha shapes algorithm was applied just after connected components to thin the datasets. So the cost of computation for the following steps was much lower because the data amount was significantly reduced (see Figure 4-8).

After removing the trunks, the parameters of trees were to be derived. But before that, 3D alpha shapes algorithm was implemented again for both single and multiple tree components.

Single tree components were to be modelled, but the point density and distribution will affect the modelling results on the same tree. So in order to eliminate the effects of different alpha values, alpha value was set to 10m (fixed) instead of optimized value for each component. After that, the point density became consistent and most of the points lied on the boundary of the tree crown but not inside. The remaining points were used directly for 3D tree modelling.

As for multiple tree components, the alpha values were optimized for each component. The optimized alpha value is the smallest one such that the remaining points belong to one solid component. This method minimized the gaps between connected trees when calculating the area and volume of the crown.



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

Figure 4-8: Single tree thinning, (a) before thinning from top view, (b) after thinning from top view, (c) before thinning from horizontal view, (d) after thinning from horizontal view

4.4.3. 3D tree modelling

The areas and volumes of single trees had been calculated by convex hull, which was a point based method. Another way to derive the parameters is fitting models to the points. The Pollock model was adjusted to be more suitable for trees. The whole procedure was as following:

1) the rotation angle β was found first using principle component and the points were translated to the local coordinate system;
2) a convex hull was generated for each component and the crown base ellipse (a, b) was fitted with the points on the convex hull in 2D;
3) the position of the crown X_0, Y_0 were initialized by the median of all points and Z_0 was fixed as the average height of the points on the convex hull;
4) points above Z_0 (upper crown) were modelled by nonlinear least square fitting, so X_0 and Y_0 were fitted;
5) points below Z_0 (lower crown) were modelled the same way, where X_0, Y_0, Z_0, a, b, c, β were treated as constant so only shape parameter (n) was fitted;
6) the area and volume, which was the sum of the upper and lower crowns, were calculated by the model.

Figure 4-9 illustrates the process of 2D crown fitting. The points were transferred to a local coordinate system first. The dash line is the convex hull and the real line is the ellipse fitted by the vertices.

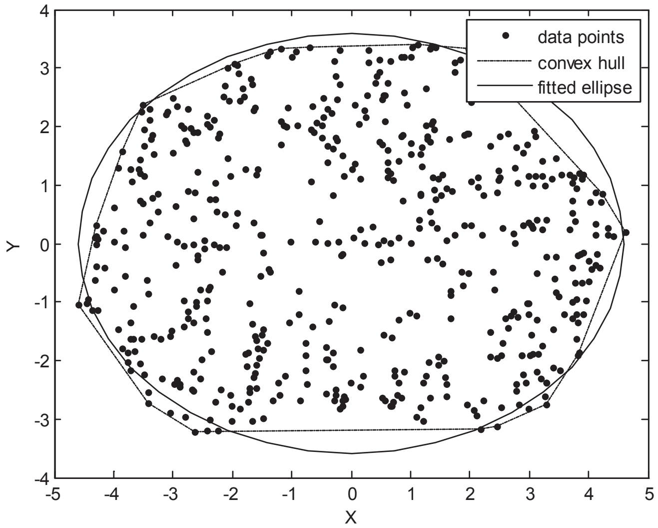
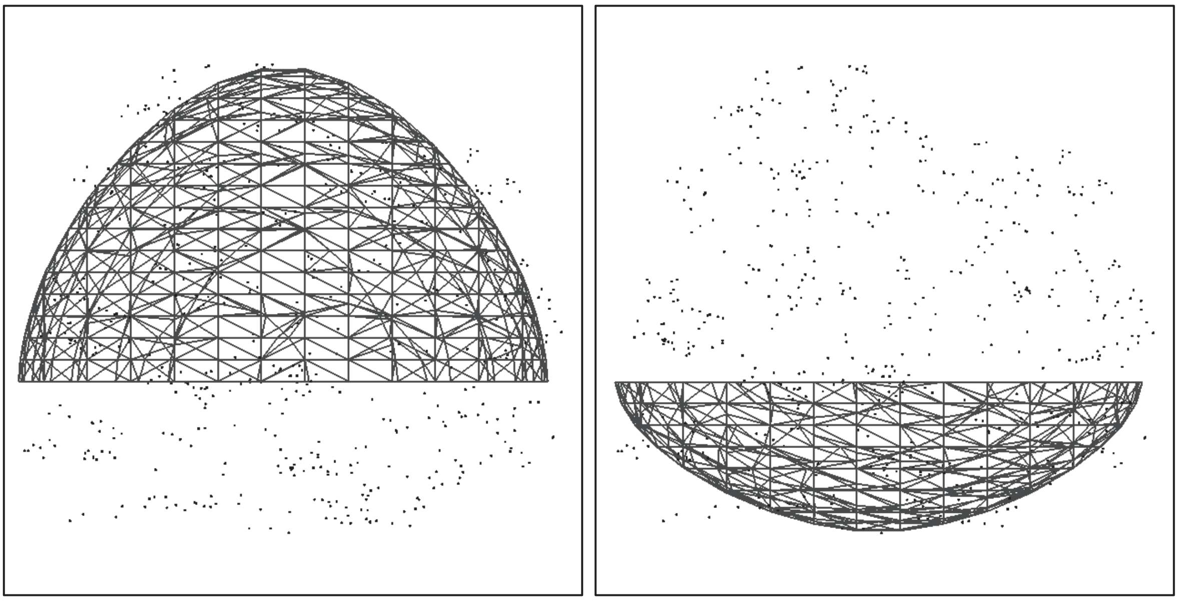


Figure 4-9: 2D crown base fitting



(a) (b)

Figure 4-10: 3D crown fitting, (a) upper crown fitting, (b) lower crown fitting

Figure 4-10 shows the process of 3D crown fitting. The height of the base was determined by the average height of the vertices and points above the base were fitted. After that, the crown position was fixed. If the upper crown and the lower crown were modelled individually, the crown positions were most likely different with each other. Because the upper crown normally has more points and bigger than the lower crown, the crown position was fixed using the upper crown so that the whole model has a continuous smooth surface which can be directly used for 3D virtual city visualization. Figure 4-11 shows the modelling results of different shapes and sizes.

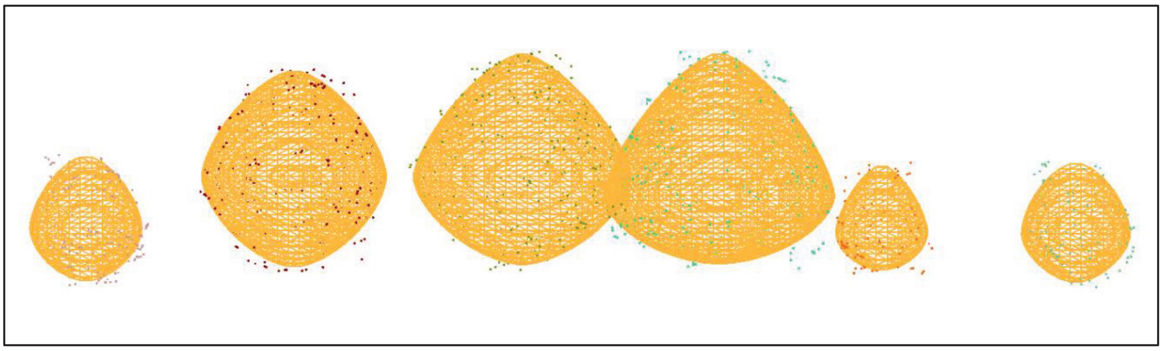


Figure 4-11: 3D modelling results

4.5. Tree oriented change detection

4.5.1. Tree to tree matching

The parameters extracted were then assigned back to each component as its features (area, volume) so that the components can be compared and visualized. Before the comparison, corresponding components were matched by checking the overlay of bounding boxes and the number of matched points. If the number of matched points was greater than a quarter of the size of the smaller component, the components were labelled a same number. If a component in the earlier epoch dataset had no counterpart in the other dataset, it meant the component has been cut and was labelled number 0. On the contrast, if a component in the later epoch dataset had no counterpart, it was newly planted and was not labelled.

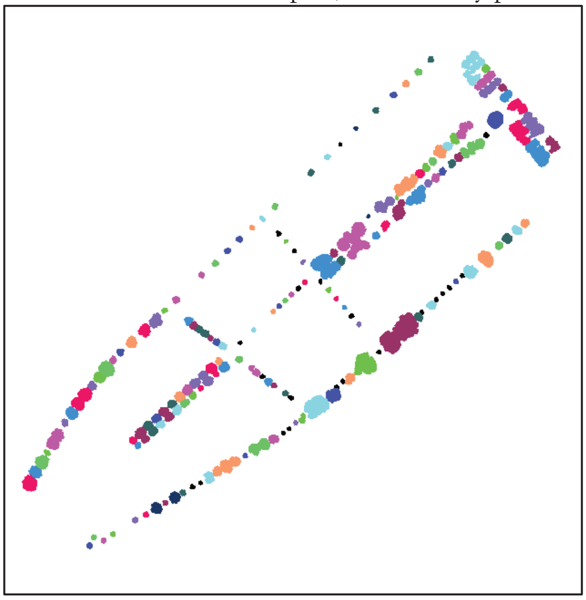


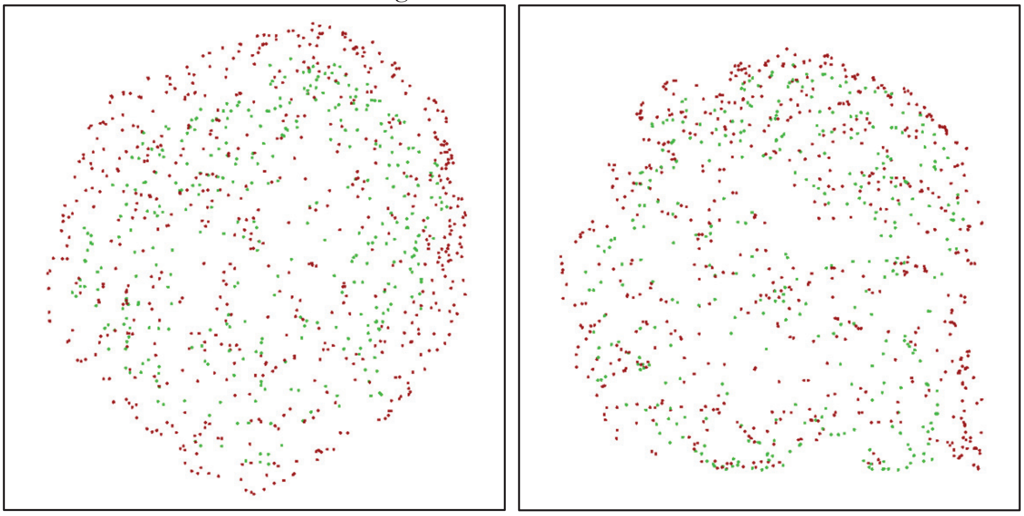
Figure 4-12: Corresponding components matching result (cut trees in dark blue; planted trees in black)

Figure 4-12 shows the result of matching. Two epoch datasets were merged together, and corresponding components had the same label while different matches had various labels shown by colours. Components that have been cut were in dark blue colour whereas newly planted components are in black.

4.5.2. Change detection

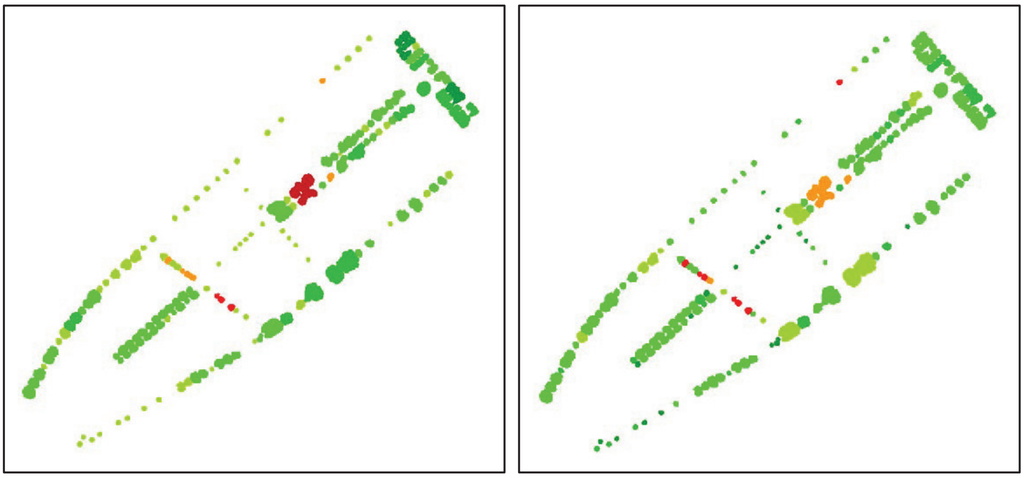
Since corresponding components have the same label, the changes can be found by comparing the parameters of each component. The area and volume were compared by subtracting them of the earlier dataset from the later one, for instance volumes of 2010 minus volumes of 2008, and then the results were assigned back to both corresponding components in order to be visualized.

Figure 4-13 shows a tree in March 2008 and in April 2010, green points are from 2008 and red ones are from 2010. It is clear that the tree has changed in 3D.

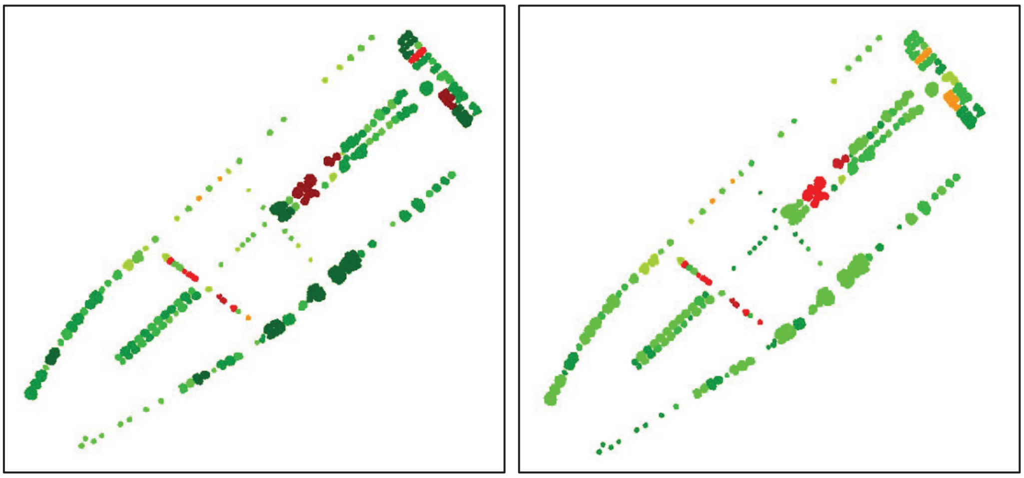


(a) (b)

Figure 4-13: Two epoch datasets (green for 2008; red for 2010), (a) top view, (b) horizontal view



(a) (b)



(c) (d)

Figure 4-14: Change detection results (2008.03-2010.04), (a) area change, (b) area change percentage, (c) volume change, (d) volume change percentage

Figure 4-14 (a) and (c) show the changes of areas and volumes between March 2008 and April 2010, in which a sequence of colours illustrates different changes. Figure 4-14 (b) and (d) show the change percentages for area and volume. So not only the changes were detected and also the changes of different size trees can be found. Most of the components increased in area and volume. Similar size components had more or less the same changes. Smaller components had equal or even bigger change percentages. The colour legend is shown as Table 4-1. Area, volume and change percentage have the same legend, so the unit of values are m², m³ and % respectively.

Table 4-1: Colour legend

Value	-1000	-100	-50	-10	0	10	50	100	1000	
Colour										
R G	100 0	150 0	200 0	255 0	255 150	150 255	0 255	0 200	0 150	0 100

Dataset of November 2008 was also compared with the other two. Figure 4-15 shows the change detection results between March and November 2008. And the comparison results between November 2008 and April 2010 are shown as Figure 4-16.

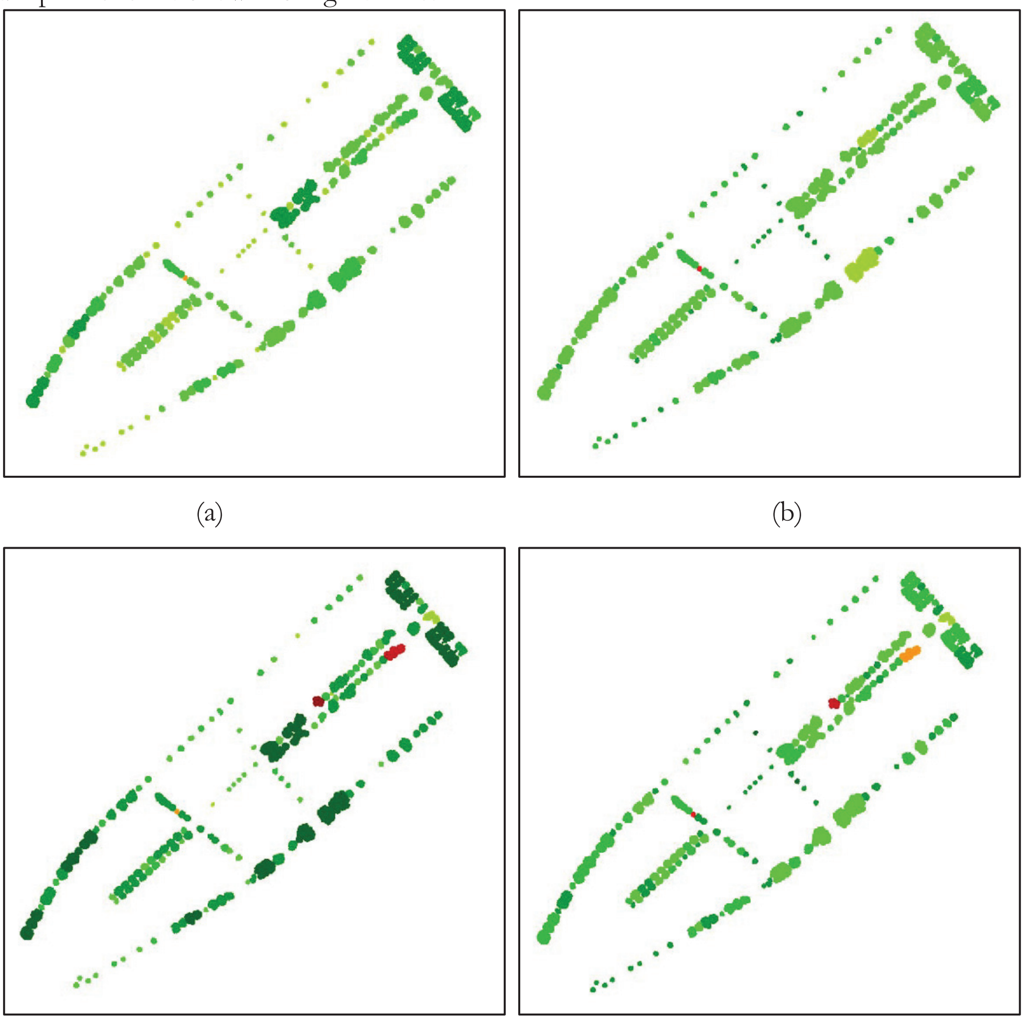
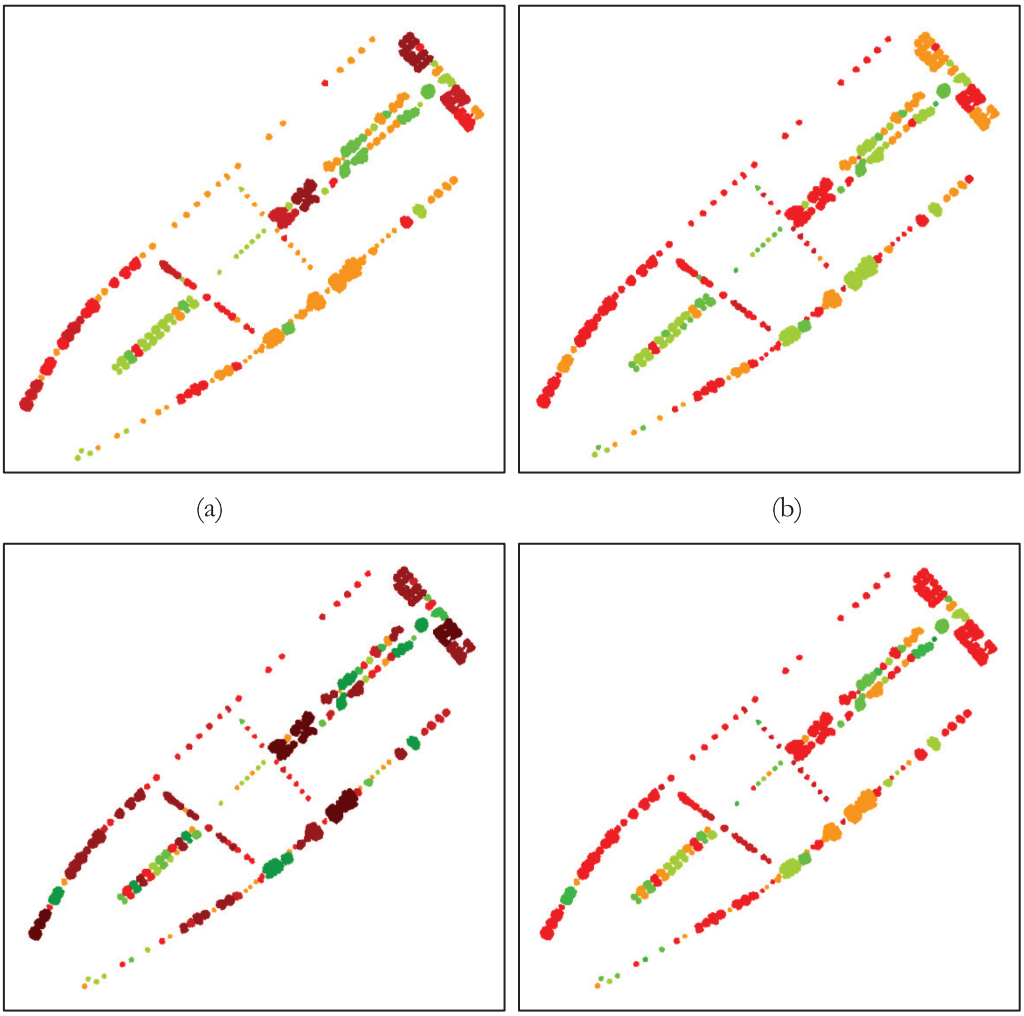


Figure 4-15: Change detection results (2008.03-2008.11), (a) area change, (b) area change percentage, (c) volume change, (d) volume change percentage



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

Figure 4-16: Change detection results (2008.11-2010.04), (a) area change, (b) area change percentage, (c) volume change, (d) volume change percentage

Almost all the components have grown from March to November in 2008. Only one component decreased in area and several in volume.

The majority of the components from November 2008 have decreased in both area and volume compared with April 2010. However, no component has decreased more than 100%; most of them are from 10% to 50%.

Components that have no counterparts in the other dataset are not displayed because no mathematical changes could be detected. Through the results of components matching, trees that have been newly planted and been cut can be found. Because of errors brought from the former steps, trees that changed smaller than 10% have great possibility of no change. Statistical analysis is presented in next chapter.

4.6. Summary

In the beginning of this chapter, a flowchart of programme processing the datasets was introduced. Connected components algorithm was applied first and datasets were thinned by 3D alpha shapes. Then local maxima were found to discriminate single tree components and multiple tree components. For single tree components, 3D alpha shapes algorithm was implemented again to make the point density consistent. Then the parameters were derived using adjusted Pollock model and convex hull. For multiple tree

components, parameters (area and volume) were computed through 3D alpha shapes directly. Then the corresponding components were matched and the changes were detected.
After the introduction, the datasets were presented and processed. Then the procedures were described in detail and the results were visualized.

5. STATISTICAL ASSESSMENT AND DISCUSSION

5.1. Analysis of vegetation verification

5.1.1. Connected components analysis

Non-vegetation components and small bush components were to be removed because only high vegetation (trees) components were the objects of this research. However some small trees were removed in the meantime and some non-tree components were not. So completeness and correctness (Heipke et al., 1997) were used to assess the quality of this step. True positive (TP) is the real tree component after removing; false positive (FP) is the remaining incorrectly classified tree component and false negative (FN) is the real tree component that have been removed. The completeness and correctness of the result were assessed by visual inspection.

$$Completeness = \frac{TP}{TP + FN} \tag{5-1}$$

$$Correctness = \frac{TP}{TP + FP} \tag{5-2}$$

$$Quality = \frac{TP}{TP + FN + FP} \tag{5-3}$$

The completeness, correctness and overall quality were calculated as the above formulas. The number of component before and after removing and the accuracy results are shown as Table 5-1.

Table 5-1: Number of component and the accuracy analysis results

Dataset	2008.03	2008.11	2010.04
Before removing	1451	1169	2118
After removing	306	229	275
Completeness	93%	97%	98%
Correctness	86%	97%	93%
Overall Quality	81%	94%	91%

The reference number of missing tree components was observed visually, thus some mistakes might have occurred, especially when a component had a very small size or the shape was not clearly a tree-like shape. Similarly, the number of remaining tree components was also determined by visual inspection.

The first dataset had the lowest overall quality; especially the correctness was quite low. That was mainly affected by classification because visually it was obvious that the classification result of the first dataset was worse than the other two.

Seven component attributes were discussed in the section 4.3.1.2, but due to the diversity of tree component's attribute values and the risk of removing tree components, the thresholds were set not strict at all. Some of the attributes could only remove one or two non-tree components or even fail to remove any one. It also happened that some non-tree components were removed but in the meantime several tree components were also removed. In this case the threshold was adjusted to avoid missing any trees.

The overall accuracy was high in terms of classification and feature extraction, however, as for change detection the error will be directly considered in the following steps. So to avoid the effects of non-tree components, a subset in which the trees were mostly along roads was selected.

5.1.2. Local maxima analysis

Before Local maxima, 3D alpha shapes algorithm was applied to thin the sub datasets. Table 5-2 shows the number of component in each subset and the number of points before and after thinning.

Table 5-2: Results of data thinning

Dataset	2008.03	2008.11	2010.04
No. of component	170	153	168
Before thinning	100174	418459	416040
After thinning	54731	121918	138282
Thinning ratio	45%	71%	67%

Even though the range threshold was adaptive to the component height, some local maxima were incorrectly computed. If a single tree was labelled as a multiple tree component, the parameters would be derived through 3D alpha shapes algorithm, which was a point based method, similar to convex hull. The parameters would also be quite the same. So this scenario would not severely affect the change results. However, if a multiple tree component was mistaken as a single tree, the parameters might be far from the real. So the threshold was set in favour of minimizing the error of single tree components.
The confusion matrix (error matrix) of dataset April 2010 was illustrated as Table 5-3.

Table 5-3: Confusion matrix of Local maxima

Local maxima	Single	Multi		User
Single	132	6	138	0.96
Multi	2	28	30	0.93
	134	34	168	
Producer	0.99	0.82		0.95

Two components out of 168 were mislabelled. Figure 5-1 shows an example component which contained two trees but only one local maximum was detected.

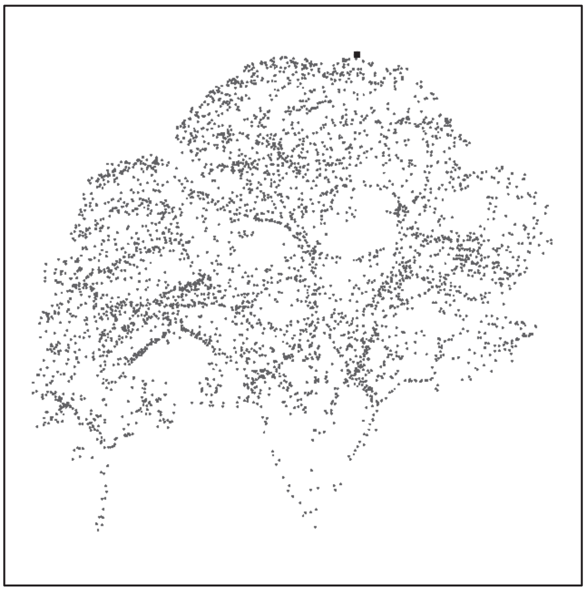


Figure 5-1: Mislabel of Local maxima

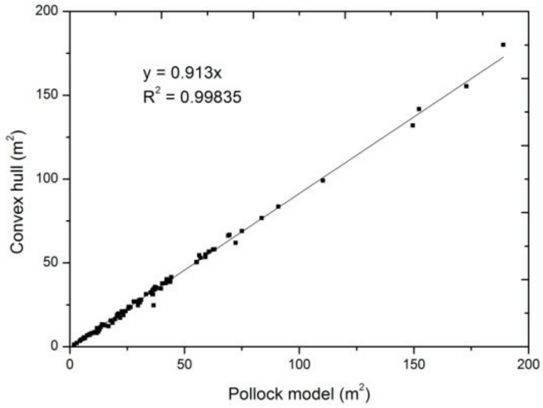
As shown in Figure 5-1, the left small tree was attached to the bigger one. It was failed to detect the local maximum for the small tree. Detecting the trunks might be helpful in this case, but not all components have sufficient points on the trunks. Models and parameters were derived considering them as a single tree, so the result of changes would be not accurate.

5.2. Tree parameters assessment

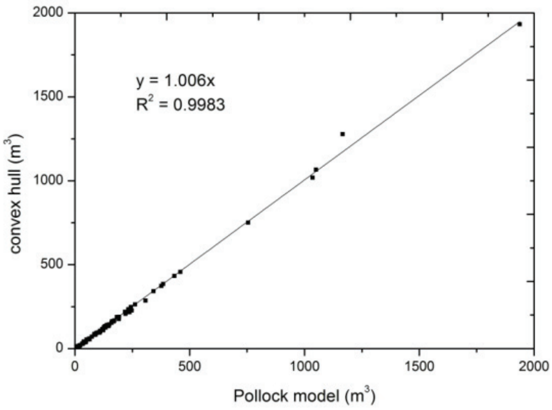
5.2.1. Parameters comparison

Adjusted Pollock model and convex hull were both utilised to derive tree parameters, so the results of these two can be compared which will show the consistency of these two independent methods. Under the circumstance that no true data were available, the comparison results could to some extent indicate the stability of parameters.

The dataset of March 2008 was selected in which 87 single tree components existed. The results of Pollock model and convex hull were compared as Figure 5-2.



(a)



(b)

Figure 5-2: Linear correlation between Pollock model and convex hull, (a) area fitting, (b) volume fitting

These two methods showed very strong linear correlation even one was based on model but the other was based on points. Especially the slope of volume was 1.006 which meant they had almost the same volume results. The area of convex hull was little bit smaller than Pollock model since the slope was 0.913.

The ratio of the difference between the two results with respect to the Pollock model result was calculated as following:

$$Diff\ ratio = \frac{Pollck\ model - Convex\ hull}{Pollck\ model} \tag{5-4}$$

The standard deviation of the difference of area was 0.0721, and standard deviation of the difference of volume was 0.0764 which were both under 10%. The standard deviation of the difference ratio also indicated that the differences between these two methods were very small.

5.2.2. Change percentage comparison

To find out the consistency of the changes through the two methods, the differences of change with respect to the Pollock model results were calculated.

$$Change = Parameter2 - Parameter1 \tag{5-5}$$

$$Diff = Model\ change - Convex\ hull\ change \tag{5-6}$$

$$Normalized\ diff = \frac{Diff}{Parameter1} \tag{5-7}$$

The normalized differences of changes between March 2008 and April 2010 were depicted as Figure 5-3, in which 78 corresponding single tree components were found. The standard deviation of the differences in area was 0.0607 and 1.5 times the half of the interquartile was 0.0518. Moreover, the standard deviation of the differences in volume was 0.0985 and 1.5 times the half of the interquartile was 0.0789. Both of the values were under 10% which further improved the consistency of these two methods.

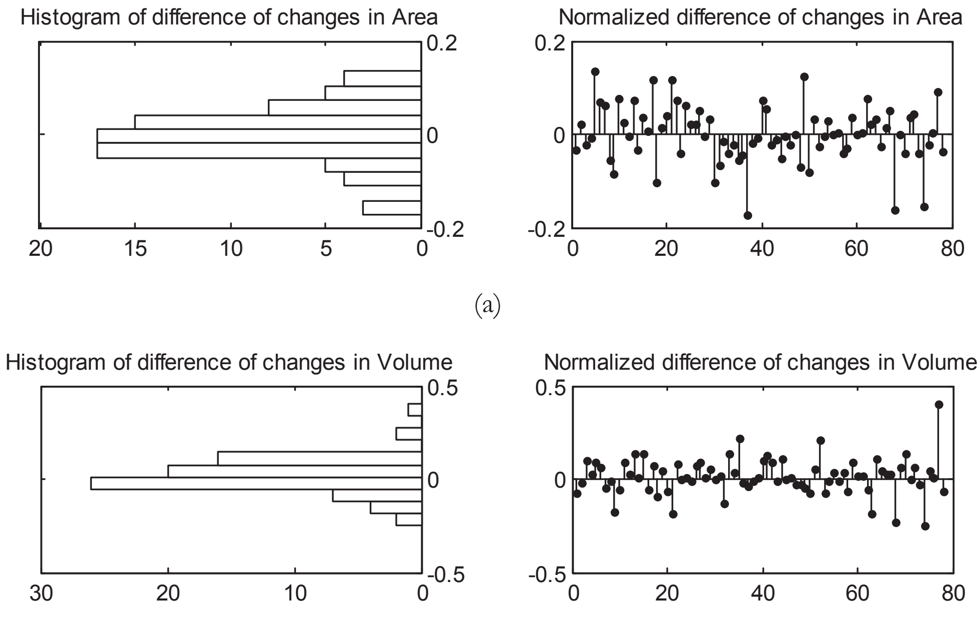


Figure 5-3: Differences of changes between two methods, (a) difference of changes in area, (b) difference of changes in volume

Since the values of the two methods were quite near each other, the differences of changes were also very small. Without ground truth, the results of the adjusted Pollock model cannot be validated and no further conclusion can be drawn. But the comparison of these two methods showed the stability of the model.

5.3. Change results analysis

There were 170, 153 and 168 components in the three datasets as shown in Table 5-2. After the matching algorithm, corresponding components which contained one or more components from each dataset were combined together. Table 5-4 shows the detected changes among three datasets.

Table 5-4: Change detection results

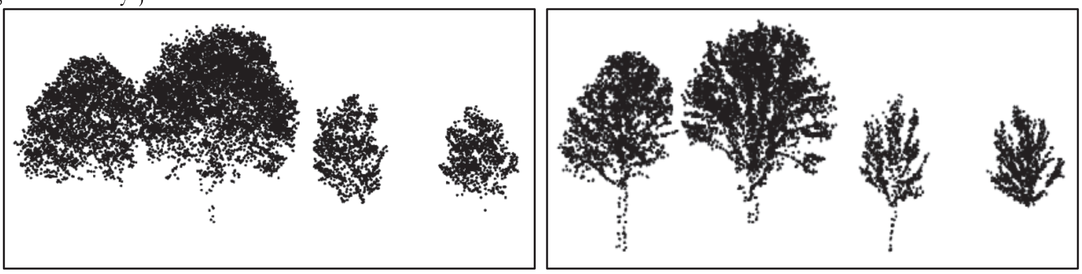
Categories	Cut	Planted	Area change		Volume change	
Change	Only in data1	Only in data2	Area↑	Area ↓	Volume ↑	Volume ↓
08.03-10.04	4	22	132	9	128	13
08.03-08.11	3	20	131	1	129	3
08.11-10.04	3	4	44	102	40	106

For the first comparison, 12 components' change ratios out the 141 were smaller than 10 per cent (positive and negative), which actually cannot be assured by the parameter derivation methods. 13 components' volumes had decreased while only 9 components' areas had decreased which might be real but also might be caused by the errors of parameters since most of the parameters of these components were near to zero.

According to the change detection results, many tree components were detected as planted in the latter two datasets compared with the first dataset. It might be true that some of the trees were indeed planted, but more likely, that was caused by omission error. Some real tree components in the first dataset were incorrectly removed because the component size and point density were small even though the threshold of component size was proportional to the density.

Completeness and correctness errors will end up with changes as cut or planted, thus these two categories were verified through the original vegetation points. The change result between March 2008 and April 2010 was visually inspected. 3 out of 4 cut trees and only 2 out of 22 planted trees were confirmed. The results of cut or planted were severely affected by commission and omission errors at tree level and also the accuracy of classification. Therefor the qualities of these two categories were low.

Seasonal difference can affect the change results. Figure 5-4 illustrates the difference in data distribution between leaf-on and leaf-off season. Branches and trunks can somehow be captured in leaf-off season whereas barely of them can be seen from leaf-on season. With dense foliage, the chance of missing the leaves for the laser beam is smaller than sparse leaf trees. Also, the penetration distance of dense foliage trees is often smaller than sparse ones. This might be the reason that most trees were decrease in area and volume from Nov. 2008 to April 2010. Other reason could be that the municipality pruned the trees. Seasonal affects were neglected between March 2008 and April 2010 because trees normally would not change obviously just after the winter.



(a) (b)

Figure 5-4: Seasonal difference, (a) leaf-on season (Nov.2008) (b) leaf-off season (April 2010)

Multiple tree components had greater change per tree than single tree components, even though 3D alpha shapes minimized the gaps between connected trees. The gap will become bigger along with the growth of trees. In some cases, the matching relations were many to one, such as two single tree components in first dataset to one big component in second dataset. Then the change results will be bigger than real because first two components were computed individually however the second one took the gap into consideration.

As we all known, small trees have smaller changes in area and volume compared with bigger trees and trees of similar sizes have the same behaviour of changes, which had been reflected by the change results shown in 4.5.2. Moreover, the change percentages of small trees are the same and even greater than bigger trees because in principle smaller trees grow faster. Thus even without ground truth, these generic knowledge somehow assess the results and suggest the feasibility of the methodology.

6. CONCLUSIONS AND RECOMMENDATIONS

6.1. Conclusions

To detect the changes between different epoch lidar point clouds, a conceptual framework was proposed and a series of algorithms were applied. Following conclusions were drawn according to the processes and results.

- Connected components algorithm can cluster the points of a same object together. Components features were feasible for non-tree components removing especially when considering both geometrical and spectral features. Component size and reflectance strength were proven to be the most efficient two attributes.
- Point based Local maxima was a very simple and fast method to distinguish single tree component from multiple tree components. The accuracy of the result was 95%.
- Trunk removing was necessary before parameter derivation since some tree trunks were scanned but others might be not. And they had no meaning for crown parameter derivation like 3D modelling. The proposed method performed well enough for the following processes.
- 3D alpha shapes first of all thinned the datasets allowing us to process much larger area or bigger datasets which is vital for the development of high point density laser scanners. Duplicated points were also removed and the computation cost of following steps was reduced. Most importantly, the shape of each component was maintained so it will not affect the parameter derivation. The points on the vertex were reserved thus they could be directly used for 3D tree modelling and other parameter derivation method. Moreover, the area and volume can also be calculated through the alpha shape. Point density of different epoch datasets were unified using a same alpha value which assured that the change results would not be affected by the differences in point density.
- The Adjusted Pollock model showed an obvious advantage for 3D tree modelling by providing the crown shape parameter n. So every single tree have its own crown shape, which is more realistic. The separation of upper crown fitting and lower crown fitting was proven feasible and the results had very high linear correlation with convex hull which was based on the points. Thus it was independent and comparable to Pollock model. The standard deviation of the difference ratio was about 8%. Also the differences of changes between these two methods were also under 10%. This suggested that adjusted Pollock model had great potential of modelling trees accurately and vividly for a 3D virtual city.
- Corresponding components were supposed to have the same location in the same coordinate system. The tree to tree matching using component bounding boxes and point to point distances were quite efficient. No mismatching had been observed through visual inspection.
- Multiple tree components were compared together as big objects instead of separating them apart into single tree components so that the error of separating could be avoided. Compared with convex hull, 3D alpha shapes minimized the gaps between trees. Nevertheless, the minimum alpha value, which would be affected by the distribution of the points, was essential for the parameter derivation.
- The proposed conceptual framework provided a guideline for change detection of trees in multi-temporal airborne lidar point clouds. The growth and pruning of trees were successfully detected. Also the programme processed the datasets semi-automatically.

6.2. Recommendations

The proposed methodology had been proven feasible to accomplish the research objectives. Nonetheless, due to the limitation of time and knowledge, following up studies could be done based on the current work. Thus some recommendations were made as following:

- Reference data could be helpful to further verify the parameters and change results. So airborne lidar data with ground truth are highly welcomed. The imagery obtained simultaneously with the point clouds can also help to identify the commission and omission errors at the level of trees.
- The study area was only a part of the overall dataset, so if needed, the proposed methodology can be applied to the entire dataset and even the whole city. Thus the result would be more meaningful for urban planning and decision making.
- Classification results will affect the change detection results to a great extent, thus better vegetation classification result is preferable. Component features were quite useful for distinguishing tree components from others, so datasets with more features are highly recommended.
- Because of existing of non-tree components, such as building fragments or bushes, a subset of the study area was selected. Alternative way is to remove non-tree components manually.
- To detect the height changes of every single tree, further efforts may be made to separate multiple tree components. Then the error brought by the separation needs to be taken into account. Normally connected trees can be separated by a straight line (in 2D, plane in 3D) in between, but actually trees intersect with each other which may affect the model and parameter results. The final choice should be made based on both methods.
- Mobile mapping system (MMS) is capable of capturing the details of the trunk and lower crown of trees. So the combination of mobile and airborne laser scanning data will facilitate the modelling of trees with accurate trunk information.

LIST OF REFERENCES

Chen, X. L., Jabbar, M. T., Cai, X. B., Wu, Z. Y., & Liang, Y. T. (2005). Detection and evaluation of vegetation change and urbanization in the central China. *IGARSS 2005: IEEE International Geoscience and Remote Sensing Symposium, Vols 1-8, Proceedings*, 230-233.

Choi, K., Lee, I., & Kim, S. (2009). A feature based approach to automatic change detection from Lidar data in urban areas. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVIII(3/W8)*, 259-264.

Coppin, P., Jonckheere, I., Nackaerts, K., Muys, B., & Lambin, E. (2004). Digital change detection methods in ecosystem monitoring: a review. *International Journal of Remote Sensing, 25*(9), 1565-1596.

Da, T. K. F., Loriot, S., & Yvinec, M. (2011). CGAL User and Reference Manual. Retrieved 18th October, 2011, from http://www.cgal.org/Manual/latest/

Elvidge, C. D., & Portigal, F. P. (1990). Change Detection in Vegetation Using 1989 Aviris Data. *Imaging Spectroscopy of the Terrestrial Environment, 1298*, 178-189.

Gong, J., Sui, H., Ma, G., & Zhou, Q. (2008). A review of multi-temporal remote sensing data change detection algorithms. *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XXXVIII-3/W*(Part B7), 757-762.

Heipke, C., Mayer, H., Wiedemann, C., & Jamet, O. (1997). Evaluation of automatic road extraction. *IAPRS, 32*(3SECT 4W2).

Hoffer, R. M., & Lee, K. S. (1990). Change Detection of Forest Cover Using Satellite Radar Data. *Protecting Natural Resources with Remote Sensing*, 349-355.

Houldcroft, C. J., Campbell, C. L., Davenport, I. J., Gurney, R. J., & Holden, N. (2005). Measurement of canopy geometry characteristics using LiDAR laser altimetry: a feasibility study. *Geoscience and Remote Sensing, IEEE Transactions on, 43*(10), 2270-2282.

Hyyppä, J., Hyyppä, H., Litkey, P., Yu, X., Haggrén, H., Rönnholm, P., et al. (2004). Algorithms and methods of airborne laser-scanning for forest measurements. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36*(8/W2), 82-89.

Kato, A., Moskal, L. M., Schiess, P., Swanson, M. E., Calhoun, D., & Stuetzle, W. (2009). Capturing tree crown formation through implicit surface reconstruction using airborne lidar data. *Remote Sensing of Environment, 113*(6), 1148-1162.

Kim, S.-R., Kwak, D.-A., oLee, W.-K., Son, Y., Bae, S.-W., Kim, C., et al. (2010). Estimation of carbon storage based on individual tree detection in Pinus densiflora stands using a fusion of aerial photography and LiDAR data. *SCIENCE CHINA Life Sciences, 53*(7), 885-897.

Lu, D., Mausel, P., Brondizio, E., & Moran, E. (2004). Change detection techniques. *International Journal of Remote Sensing, 25*(12), 2365-2401.

Maas, H. G., & Vosselman, G. (1999). Two algorithms for extracting building models from raw laser altimetry data. *Isprs Journal of Photogrammetry and Remote Sensing, 54*(2-3), 153-163.

Murakami, H., Nakagawa, K., Hasegawa, H., Shibata, T., & Iwanami, E. (1999). Change detection of buildings using an airborne laser scanner. *ISPRS Journal of Photogrammetry & Remote Sensing, 54*, 148-152.

Murakami, H., Nakagawa, K., Shibata, T., & Iwanami, E. (1998). Potential of an airborne laser scanner system for change detection of urban features and orthoimage development. *International Archives of Photogrammetry and Remote Sensing, 32*, 422-427.

Oude Elberink, S. J., & Vosselman, G. (2009). 3D information extraction from laser point clouds covering complex road junctions. *Photogrammetric Record, 24*(125), 23-36.

Pang, A. (2012). Visualizing Distributions from Multi-Return Lidar Data to Understand Forest Structure. Retrieved 20th January, 2012, from http://slvg.soe.ucsc.edu/lidar.html

Pollock, R. J. (1994). *Model-based approach to automatically locating tree crowns in high spatial resolution images*. Paper presented at the Proceedings of SPIE - The International Society for Optical Engineering.

Pratihast, A. K. (2010). *3d tree modelling using mobile laser scanning data*. Unpublished MSc Thesis, Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands.

Rutzinger, M., Höfle, B., & Pfeifer, N. (2007). Detection of high urban vegetation with airborne laser scanning data. *Proceedings forestsat 2007*, digital media.

Rutzinger, M., Pratihast, A. K., Oude Elberink, S. J., & Vosselman, G. (2011). Tree modelling from mobile laser scanning data-sets. *The Photogrammetric Record, 26*(135), 361-372.
Singh, A. (1989). Digital Change Detection Techniques Using Remotely-Sensed Data. *International Journal of Remote Sensing, 10*(6), 989-1003.
Sithole, G., & Vosselman, G. (2006). Bridge detection in airborne laser scanner data. *Isprs Journal of Photogrammetry and Remote Sensing, 61*(1), 33-46.
Vosselman, G. (2000). Slope based filtering of laser altimetry data. *International Archives of Photogrammetry and Remote Sensing, XXXIII* (B3/2), 935-942.
Vosselman, G. (2003). 3d reconstruction of roads and trees for city modelling. *International Archives of Photogrammetry and Remote Sensing, XXXIV-3/W13*, 6.
Vosselman, G., Gorte, B. G. H., & Sithole, G. (2004). Change detection for updating medium scale maps using laser altimetry. *Internat. Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 35*, 207-212.
Vosselman, G., & Maas, H.-G. (Eds.). (2010). *Airborne and terrestrial laser scanning*. GB: Whittles Publishing.
Wang, Y., Weinacker, H., & Koch, B. (2008). A lidar point cloud based procedure for vertical canopy structure analysis and 3D single tree modelling in forest. *Sensors, 8*(6), 3938-3951.
Xu, S., Oude Elberink, S. J., & Vosselman, G. (2012). Entities and features for classification of airborne laser scanning data in urban area. Unpublished PhD research. Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente.
Yasuoka, Y., Yokota, T., Miyazaki, T., & Iikura, Y. (1990). Detection of Vegetation Change from Remotely Sensed Images Using Spectral Signature Similarity - Discrimination of Long-Term Changes from Seasonal-Changes. *Remote Sensing Science for the Nineties, 1*(3), 1609-1612.
Yu, X., Hyyppä, J., Kaartinen, H., & Maltamo, M. (2004). Automatic detection of harvested trees and determination of forest growth using airborne laser scanning. *Remote Sensing of Environment, 90*(4), 451-462.
Yu, X., Hyyppä, J., Kukko, A., Maltamo, M., & Kaartinen, H. (2006). Change detection techniques for canopy height growth measurements using airborne laser scanner data. *Photogrammetric Engineering and Remote Sensing, 72*(12), 1330-1348.
Yu, X., Hyyppä, J., Vastaranta, M., Holopainen, M., & Viitala, R. (2011). Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *Isprs Journal of Photogrammetry and Remote Sensing, 66*(1), 28-37.