AUTOMATED TREETOP DETECTION AND TREE CROWN IDENTIFICATION USING DISCRETE-RETURN LiDAR DATA

Haijian Liu

Thesis Prepared for the Degree of

MASTER OF SCIENCE

UNIVERSITY OF NORTH TEXAS

May 2013

APPROVED:

Pinliang Dong, Major Professor
Alexandra Ponette-González, Committee Member
Chetan Tiwari, Committee Member
Paul Hudak, Chair of the Department of Geography
Mark Wardell, Dean of the Toulouse Graduate School

Liu, Haijian. Automated Treetop Detection and Tree Crown Identification Using Discrete-Return LiDAR Data. Master of Science (Applied Geography), May 2013, 51 pp, 2 tables, 31 figures, reference, 27 titles.

Accurate estimates of tree and forest biomass are essential for a wide range of applications. Automated treetop detection and tree crown discrimination using LiDAR data can greatly facilitate forest biomass estimation. Previous work has focused on homogenous or single-species forests, while few studies have focused on mixed forests. In this study, a new method for treetop detection is proposed in which the treetop is the cluster center of selected points rather than the highest point. Based on treetop detection, tree crowns are discriminated through comparison of three-dimensional shape signatures. The methods are first tested using simulated LiDAR point clouds for trees, and then applied to real LiDAR data from the Soquel Demonstration State Forest, California, USA. Results from both simulated and real LiDAR data show that the proposed method has great potential for effective detection of treetops and discrimination of tree crowns.

ACKNOWLEDGEMENTS

I would like to express my deep appreciation to my major professor and supervisor Dr. Pinliang Dong, who has helped me a lot in selecting the research topic, collecting the data, and developing the research methods. His suggestions and encouragement are crucial for me to enlarge my horizon. I learned a lot from him.

My gratitude also goes to Dr. Alexandra Ponette-González and Dr. Chetan Tiwari. They spent a lot of time reading my thesis and providing some useful suggestions. I would also like to give my special thanks to Dr. Alexandra Ponette-González who edited earlier versions of my thesis. She is a very nice and patient professor.

I also want to say thanks to my friend Dr. Wei Liu. During the time he visited our department, we discussed my thesis many times. He not only taught me some computer programming skills, but also gave me useful suggestions in my research.

During the three years of my study, many professors and friends have helped me in my research. Dr. Bruce Hunter provided me some hardware for fieldwork. Dr. George Maxey gave me a lot of encouragement. My best friend Zongpei Tang and I shared many interesting topics in GIS and geography.

Last but not least, I want to thanks my wife Shan Gao for her unconditional support and love. Without her support, it is impossible for me to finish my thesis.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS iii

LIST OF TABLES v

LIST OF FIGURES vi

INTRODUCTION 1

Statement of Problem 3

Previous Studies 5

STUDY AREA AND DATA 9

Study Area 9

Dataset 9

METHODOLOGY 12

Simulation of LiDAR Point Clouds for a Mixed Forest 12

LiDAR Data Filtering 17

Creating Raster Images from Filtered LiDAR Points 19

Detecting Treetops 20

Discriminating 3D Crown Shapes 28

RESULTS AND DISCUSSION 30

Results from Simulated Data 30

Number of Trees 30

Location of Treetop 32

Tree Types 35

Results from Real Forest 39

CONCLUSION 46

REFERENCES 48

LIST OF TABLES

Page

1. Basic information of simulated points 17

2. Number of trees detected using different methods 31

LIST OF FIGURES

Page

1. General flowchart for forest biomass estimation using LiDAR data 4
2. Study area (300m x 200m) 9
3. Location of a random point in a 2D circle 13
4. Cone crown profile 14
5. Hemisphere crown profile 15
6. Half-ellipsoid crown profile 15
7. Simulated forest in 2D by random 3D points 16
8. Raster image created from initial LiDAR points (cell size = 0.2m x 0.2m) 18
9. Filtered LiDAR points 19
10. Raster created from filtered LiDAR points (cell size = 0.2m x 0.2m) 20
11. Local maxima detected by static and variable moving windows 24
12. Calculation of cluster center 25
13. Calculation of variable windows 26
14. Treetops detected by CCHP 27
15. Delineated crown space by lowest point between two adjacent trees 28
16. Absolute and relative offset 33
17. Comparison of X-coordinate between real and detected trees location 34
18. Comparison of Y-coordinate between real and detected trees location 34
19. Angle shape signature of cone 36
20. Angle shape signature of half ellipsoid 36
21. Angle shape signature of hemisphere 37
22. Angle shape signature of all simulated trees 37
23. Weighted average value of each curve 38

24. Combination of angle signature and normalized height signature 39
25. Raster image created from filtered LiDAR data 40
26. Treetops detected by the new method 41
27. Delineated crown space by lowest point between two adjacent trees 41
28. Large crown trees (radius >= 2m).. 42
29. Combination of angle distribution and normalized height distribution 43
30. Selected testing trees 43
31. Tree clusters. Trees with similar shape crown gather together........................... 44

INTRODUCTION

Accurate estimates of tree and forest biomass are essential for a wide range of applications, including climate change assessments, bio-energy production, natural resource management, biodiversity monitoring, and conservation. Forests cover 31 percent of Earth's total land area and provide shelter for 80 percent of biodiversity (Fao 2012). Therefore, much research has been conducted on forest ecosystems, especially canopy structure, a key factor in forest ecosystem processes (Sexton 2009).

Trees not only provide timber, food and shelter for people and wildlife, but also they provide diverse ecosystem services (Carle et al. 2002). They can clean water and help to stabilize the soil by absorbing and filtering rainwater to reduce runoff (Calder 2007). Trees as well as the grass can help reduce the urban heat island effect (Armson et al. 2012). Forests are natural "carbon sinks" that absorb large amounts of carbon dioxide from the atmosphere (Kartha 2001). Based on Global Forest Resource Assessment (Fao 2010), forests store 289 gigatonnes (Gt) carbon in their biomass. Through photosynthesis, plants absorb carbon to make carbohydrates and through metabolism, plants release oxygen. This natural exchange can help promote the balance of the atmosphere, thereby stabilizing climate.

Since electricity can be created through combustion of biomass directly with less pollution (van den Broek et al. 1996), bio-energy is considered to be an alternative energy in the future. Some species of trees have high oil content and are considered suitable fuel for electricity generation. For these and other applications, it is necessary to distinguish tree species quickly and to analyze their spatial distribution.

Forest canopy structure is “the organization in space and time, including the position, extent, quantity, type and connectivity, of the aboveground components of vegetation” (Parker 1995). It strongly influences the understory light environment and plants (Pringle et al. 2003). Canopy structure includes not only leaves, branches, trunks and roots, but also the airspace around trees. Therefore, three-dimensional (3-D) canopy structure contains size (e.g., tree height) and species information (Parker 1995). The basic unit of canopy structure is the tree crown, thus 3-D crown structure contains much information about trees (e.g., location, tree height, and crown width) (Lefsky et al. 1999).

Data on forest ecosystems can be obtained in several ways, including through in situ (field) measurements, light detection and ranging (LiDAR), and high spatial resolution imagery. These three data sources were compared by Sexton (2009). For individual trees in small study areas, field measurement is a good option to obtain data on tree parameters at low cost (Sexton 2009). However, in situ data collection is time consuming and costly over large areas or remote regions. Remotely sensed data is economical and efficient over large areas. In particular, LiDAR data provides ranging information on the canopy and ground, and optical images provide horizontal forest patterns (Popescu 2002). There are two kinds of high spatial resolution optical images: air-borne aerial photos and space-borne images. These two types of images provide a two-dimensional (2-D) perspective of tree crowns and can be used in the delineation of individual crown boundaries based on different levels of brightness around the crown center. However, automated processing is difficult due to many limitations (Culvenor 2002), and it is difficult to get the vertical variation of tree canopies (e.g., tree height). In contrast, LiDAR provides three-dimensional (3-D) products. Therefore, it has been widely used in the study of forestry and for estimating biomass and tree volume (Estornel et al. 2011).

LiDAR systems are active remote sensing devices that measure the return time of pulses sent from an airborne system and then convert time to distance. This information is used to derive the characteristics of the object surface and ground surface (Popescu 2007). Because the laser pulses can penetrate the vegetation canopy in various degrees, they can precisely show 3-D crown structure. Based on LiDAR data, some basic parameters of trees can be measured, such as tree location, height, volume, and diameter at breast height (DBH) (Popescu 2007). Furthermore, tree types can be distinguished and tree biomass can be estimated based on these characteristics. Biomass is the total number of living organisms in a given area, expressed in dry weight per unit area, which represents the ability of a forest to store carbon (Costello and Chum 1998). In sum, use of LiDAR data to study forests holds several advantages over other existing methods.

Statement of Problem

Previous forest studies have focused on individual trees or stands of a discrete area of single-species forest (Culvenor 2002). These studies have investigated how to extract forest parameters and how to obtain forest biomass based on these parameters (Popescu 2007). As shown in Figure 1, the height and the boundary of individual trees can be detected from LiDAR data, and other tree parameters (e.g., crown width, DBH, and canopy closure) can be extracted from geometric models created from LiDAR data

(Lee 2004; Popescu and Wynne 2004). Further, biomass distribution can be estimated using algebraic models created from these parameters (Popescu 2007).

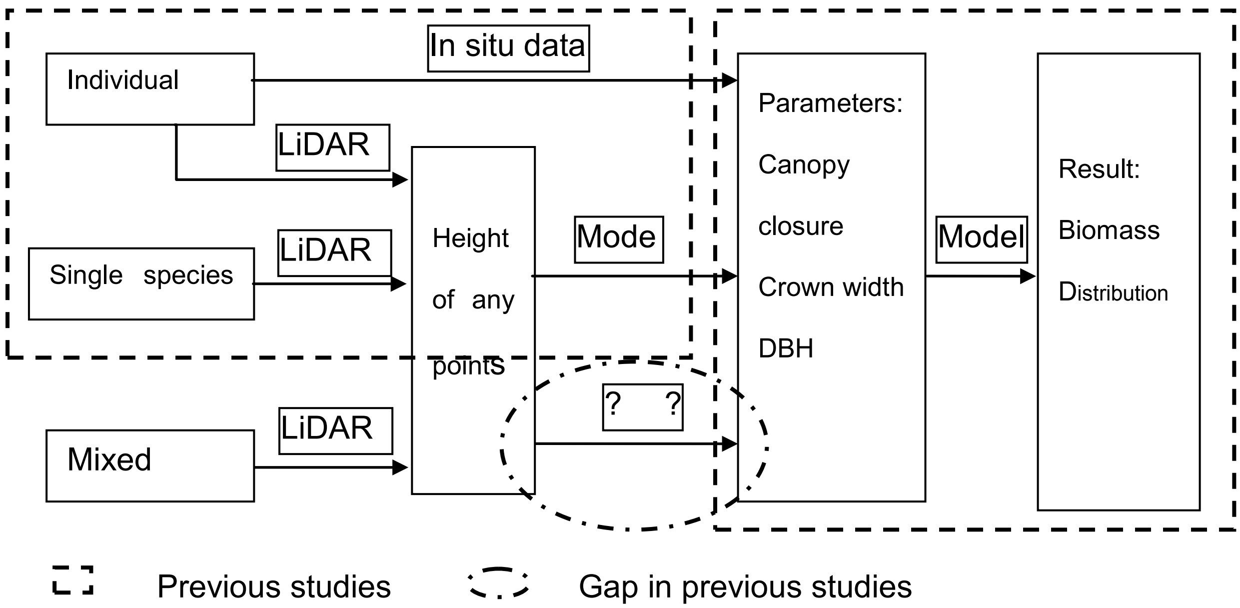


Figure 1 General flowchart for forest biomass estimation using LiDAR data

Mixed forests are more structurally complex than individual trees and single-species forests (Culvenor 2002), and thus methods for extracting tree properties in mixed forests have been relatively limited (Dong 2009). Since tree crowns overlap with one another, it is difficult to automatically detect crown boundaries, tree locations, and tree types. Currently, there is no efficient method for extracting tree parameters in a mixed forest using LiDAR data.

Therefore, the primary objective of this thesis research is to develop an automated method for treetop detection and crown shape discrimination using discrete-return LiDAR data. This study will address the following research questions:

(1) Can simulated LiDAR point clouds be used as test data of a mixed forest so that methods for treetop detection and crown discrimination can be evaluated before being applied to real LiDAR data?

(2) What methods can be developed for effective detection of treetop locations using LiDAR data?

(3) How can we discriminate three-dimensional tree crown shapes using LiDAR data?

Previous Studies

Tree location can be extracted from high spatial resolution imagery (Wulder et al. 2000; Culvenor 2002). Wulder et al. (2000) developed a local maximum (LM) filtering method for extracting tree locations from images. For conifer structures, pixels near the center of the tree crown are brighter than other pixels (Wulder et al. 2000). When a moving window passes through the image, the pixel with the highest reflectance can be determined to represent the tree location (Dralle and Rudemo 1997). The success of this method depends on window size. Comparison of tree location from different LM filtering indicates that a smaller window can detect more trees with higher commission error (falsely indicated trees) but lower omission error (missed trees). However, large windows detect fewer trees with lower commission error but higher omission error. Among these different static windows, the 3X3 LM filtering correctly located maximum trees. In order to reduce the commission error, variable window sizes were proposed. Window size can be calculated based on the relationship between tree height and crown width (Popescu and Wynne 2004) or on the average semivariance range value, which measures spatial dependency (Wulder et al. 2002).

To delineate tree crowns, a new algorithm, tree identification and delineation (TIDA), was proposed by Culvenor (2002). This method uses a “top-down” approach to delineate tree crowns, which includes local maxima identification, local minima identification, and crown pixels cluster. Local maxima are defined as the highest probability in four–way linear search direction (horizontally, vertically, and in both 45 degree planes) and local minima are defined as the lowest value in any of the four directions. Clustering is the process used to identify all pixels, which are adjacent to the local maxima pixel.

Additional information can be derived from LiDAR data, since LiDAR points include first returns from the forest canopy and last returns from the ground (Lim et al., 2003). A digital surface model (DSM) can be created by interpolation from first returns and a digital elevation model (DEM) can be created by interpolation from last returns. Further, a canopy height model (CHM) characterizing canopy height can be created by subtracting the DEM from the DSM (Popescu 2007). Since some first returns come from the interior of the crown, the CHM exhibits many holes or pits, which leads to errors in estimating forest biomass (Ben-Arie et al. 2009). These pits can be filled by low pass filtering based on focal functions to improve the quality of the CHM. Low pass filtering calculates median value to replace the high or low value within a neighborhood. It can make the canopy smooth but also changes the canopy shape to some degree (Verma and Kumar 2011).

Because measuring basic parameters including tree height, crown width, and crown volume is important in the study of forest ecosystems, models that can represent the forest canopy are garnering attention. Typical canopy height models based on LiDAR data are used to detect individual trees and delineate their crowns (Leeuwen 2010), but the models cannot retrieve individual trees and crown shapes accurately, and they underestimate tree height (Holmgren 2003). The new parametric height model (PHM) was developed to simulate pine canopies using a series of cones which fit the raw LiDAR cloud (Leeuwen 2010). This simple geometric model clearly shows crown boundaries and treetops, so the crown can easily be delineated. Although this model has the potential for representing coniferous species, the cones cannot describe the canopy of trees with different crown shapes.

The traditional technique for estimating forest biomass is to cut down sample trees, dry them, and weigh the dry mass (Costello and Chum 1998). This destructive sampling method is difficult to use in the study of large areas. An algebraic model known as FORCARB was therefore developed to estimate the volume-based biomass for individual trees (Popescu 2007). The FORCARB model is a regression equation, in which the unique variable DBH can be predicted by measurements of height and crown diameter. This model can be used from local to regional scales to estimate biomass without using destructive sampling.

Three dimensional shape is a basic requirement for characterizing a tree crown. The basic idea of a 3D-shape signature is to transform the 3D object into a 2D parameterized function and the 3D-shape signature represents the 3D object, so it can be easily compared with others (Dong 2009). The 3D-shape signature can be created from DSM or point clouds, but the process based on DSM is more efficient than that based on point clouds. By comparing the 3D-shape signature of individual trees, the type of tree can be distinguished. This method is dependent on high resolution LiDAR data and it can only be used for individual trees or delineated tree crowns.

Most of the previous research has focused on individual trees or on small areas and it has been helpful to understand the structure of individual trees and to develop unique models to simulate certain types of trees. Few studies, however, have focused on deriving parameters for larger forest areas although these parameters are important for climate and ecosystem studies (Culvenor 2002).

STUDY AREA AND DATA

Study Area

The study area (300m x 200m) is located in the Soquel Demonstration State Forest (SDSF) near Santa Cruz, California (Figure 2).

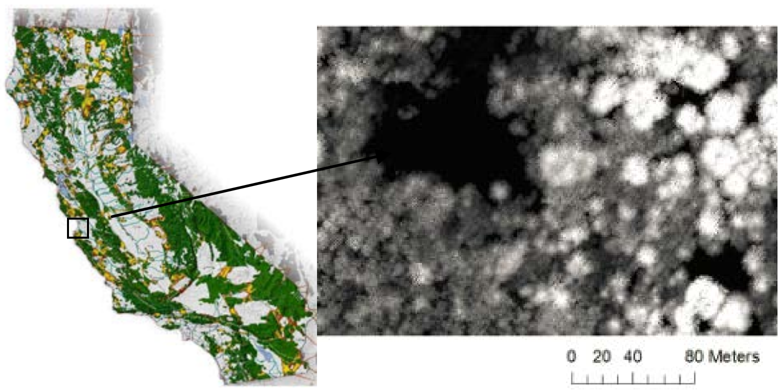


Figure 2 Study area (300m x 200m): Soquel Demonstration State Forest in California

Dominant tree species in this area include Douglas fir (*Pseudotsuga menziesii*), coast redwood (*Sequoia sempervirens*), and various oak species (*Quercus* spp.) among others. Species vary in age, shape, and size. Tree height ranges from 5-55 meters depending on the species. Crown radii also vary substantially in length, from 2-10 meters.

Dataset

The LiDAR dataset was collected by the NSF EarthScope Northern California LiDAR Project (Prentice et al. 2009) and downloaded from the NSF OpenTopography Facility at the San Diego Supercomputer Center. The Airborne LiDAR data were acquired on March 2007 with an Optech Gemini Airborne Laser Terrain Mapper (ALTM) serial number 06SEN195 mounted in a twin-engine Cessna Skymaster (Tail Number

N337P). The point density was 5.5-7/m^2. The dataset is a comprehensive point cloud (ASCII file), which includes gpstimestamp, x, y, z coordinates, intensity, class, and flightline.

1. Gpstimestamp: gps timestamp of the week (Sunday = day 0). For absolute time referencing
2. X, Y, Z: Easting, Northing and Elevation
3. Intensity: laser intensity index
4. Class: classification id number. Class1 represents Default (includes vegetation and above the ground artificial structures), class2 represents Ground
5. Flight line: flight ID number

In order to analyze the characteristics of trees and check the effectiveness of the results, dozens of trees in the study area were measured in field. The parameters include:

1. Tree crown spread (i.e., crown radius): Crown spread, the distance from the bole of the tree to the drip line, was measured with a meter tape. The widest spread of the crown and the narrowest spread of the crown were measured. These measurements were summed and averaged to obtain average crown spread.
2. Tree height: Tree height was measured with a Laser Technology TruPulse 200 Laser Rangefinder, which has an accuracy of +/- 0.3m.
3. Diameter at breast height (DBH): Tree diameter was measured 1.3 meters aboveground using a DBH tape.

4. Tree species: Trees were identified to species with the help of the park management.

Forest simulation and data analysis were conducted using ArcGIS 10, ArcGIS Software Development Kit (SDK), and Microsoft C# for methodology development and implementation.

METHODOLOGY

To evaluate the accuracy and effectiveness of the methods for treetop detection and tree crown discrimination using LiDAR data, simulated LiDAR point clouds for a forest were first generated and then analyzed. The virtual forest included three tree crown shapes: cone, hemisphere, and half ellipsoid. When this forest was generated, the basic information of each tree, such as location and crown type, was readily available. Therefore, methods for treetop detection and tree crown discrimination can be tested using the simulated LiDAR data before being applied to real LiDAR data.

The methodology is described in the following six sections: (1) Simulation of LiDAR point clouds for a mixed forest; (2) LiDAR data filtering; (3) creating a raster image from filtered LiDAR data; (4) treetop detection; (5) 3D crown shape discrimination; and (6) accuracy assessment.

Simulation of LiDAR Point Clouds for a Mixed Forest

The first component of this methodology involves simulating LiDAR point clouds for a mixed forest. To simulate a forest, three simple models of tree crowns are used: cone, hemisphere, and half ellipsoid. In order to reflect the characteristics of the forest, the location of each tree is randomly generated in 2D space with some tree crowns partially covering other crowns. The random location of each tree is determined by equation 1:

$$D > R + r * (1 - ratio) \tag{1}$$

where

D = distance between two random trees.

R = radius of the larger crown of two neighboring trees.

r = radius of the smaller crown of two neighboring trees.

ratio = overlapped radius of small crown/r.

After these random points are determined to represent tree locations, the corresponding trees can be generated around each point. The crown shape type can be randomly selected from one of the three types: cone, hemisphere, and half ellipsoid. In order to simulate these three crown shapes, random 3D points are created to simulate the crowns. The point clouds are limited in a circle in 2D surface and their locations are marked in a Cartesian coordinate system. Therefore point location (x, y) in 2D surface can be randomly selected based on random radius (R) of the crown and random angle (A). The horizontal coordinate (x) is determined by equation 2, vertical coordinate (y) is determined by equation 3, and the location of a random point is displayed in Figure 3:

$$x = r * \sin(A) \qquad (0<r<R \text{ and } 0<A\leq 360) \tag{2}$$

$$y = r * \cos(A) \qquad (0<r<R \text{ and } 0<A\leq 360) \tag{3}$$

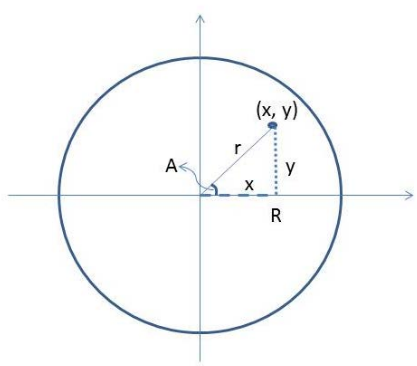


Figure 3 Location of a random point in a 2D circle.

The third dimension z represents the height of each LiDAR point which is located between two surfaces of the crown: outer surface and inner surface, so the height of any point will be randomly generated between the two surfaces $f_1(x, y)$ and $f_2(x, y)$. It is determined by equation 4:

$$\mathrm{z(x,y)} = \mathrm{f1(x,y)} + (\mathrm{f2(x,y)} - \mathrm{f(1(x,y))} * \mathrm{t} \quad (0 < t < 1) \tag{4}$$

where

$f_1(x, y)$ = inner surface.

$f_2(x, y)$ = outer surface.

t = random number between 0 and 1.

For a cone, $\mathrm{f1}\ (\mathrm{x,y})$ and $\mathrm{f2}\ (\mathrm{x,y})$ are determined by equation 5; the z value is demonstrated in Figure 4. For a hemisphere, they are determined by equation 6; the z value is demonstrated in Figure 5. For a half-ellipsoid, they are determined by equation 7; the z value is demonstrated by Figure 6.

$$\text{Cone:} \quad \mathrm{f1}\ (\mathrm{x,y}) = \frac{r}{\tan(\mathrm{B})};\ \mathrm{f2}\ (\mathrm{x,y}) = \frac{R}{\tan(\mathrm{B})} \quad (0 < B < 90) \tag{5}$$

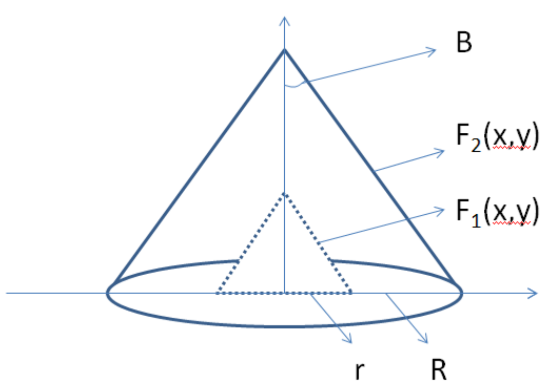


Figure 4 Cone crown profile

$$\text{Hemisphere:}\ \ f1\ (x,y)\ =\ \sqrt{r^2-(x^2+y^2)};\ \ f2\ (x,y)\ =\ \sqrt{R^2-(x^2+y^2)} \tag{6}$$

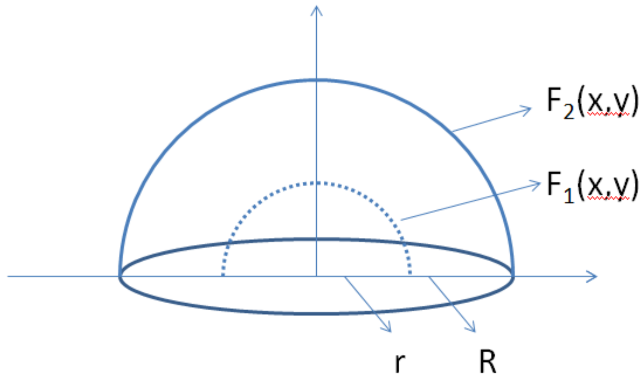


Figure 5 Hemisphere crown profile

$$\text{Halfellipsoid:}\ \ f1\ (x,y)\ =\ e*\sqrt{r^2-(x^2+y^2)};\ \ f2\ (x,y)\ =\ \sqrt{R^2-(x^2+y^2)} \tag{7}$$

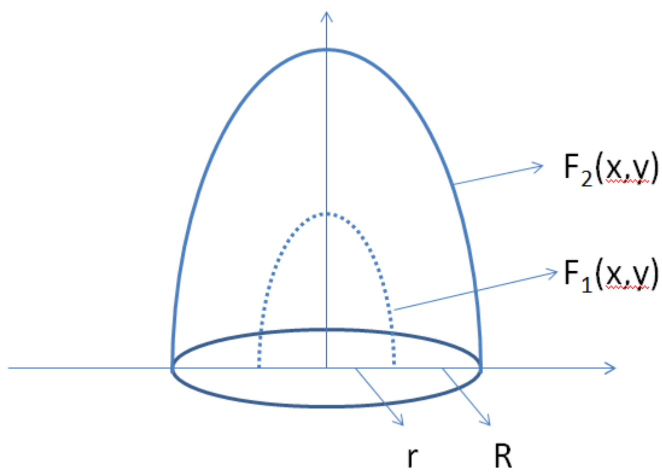


Figure 6 Half-ellipsoid crown profile

where

B = angle between the crown surface and vertical direction.

R = outer radius of the crown.

r = the inner radius of the crown.

e = the ratio between the long axis and the short axis.

Based on these conditions, random 3D points with respective coordinates (x, y, z) in a Cartesian coordinate system can be generated for each tree crown, and eventually simulated LiDAR points clouds are generated for a forest. In this research, 100 trees were simulated over a 100m x 100m area, with some ground points(Figure 7).

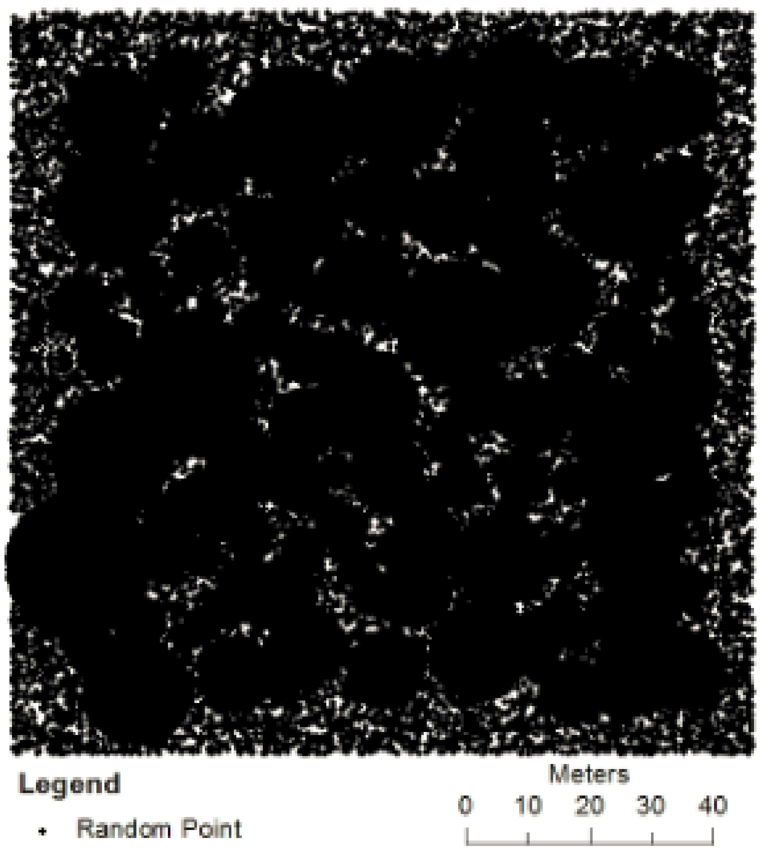


Figure 7 Simulated forest in 2D by random 3D points

When these simulated trees are generated, the main parameters of each tree and each point are recorded, such as point location, tree location, crown radius, and tree type. These original data will be used to test the result obtained from analyzing LiDAR data and estimate the effectiveness of the new methods. The initial data is recorded in Table 1.

Table 1 Basic information of simulated points

FID	Shape	Point-X	Point-Y	Point-Z	Crown radius	Tree No.	Tree type	Tree-location-X	Tree-location-Y
0	Point	45.3	82.6	9.15	5.11	1	0	43.52	78.46
1	Point	42.9	82.4	6.66	5.11	1	0	43.52	78.46
2	Point	47.4	75.4	7.18	5.11	1	0	43.52	78.46
…	…	…	…	…		…	…	…	…
20000	Point	79.4	73.2	8.85	9.72	14	1	72.04	69.97
20001	Point	69.6	71.5	8.61	9.72	14	1	72.04	69.97
….	…	…	…	…		…	…	…	…

LiDAR Data Filtering

Laser pulses generated by LiDAR systems may penetrate the crown and can be reflected from interior crown surfaces, including leaves, the trunk, or the understory. When all points are used to create a raster image to show crown shape, many pits appear on the surface (Figure 8).

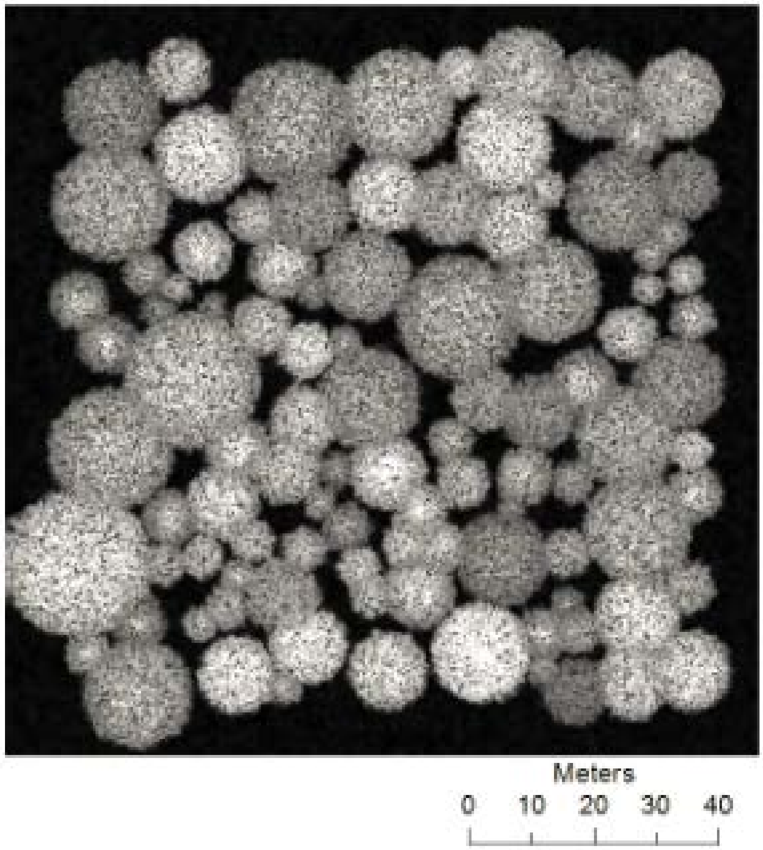


Figure 8 Raster image created from initial LiDAR points (cell size = 0.2m * 0.2m)

In order to create smooth tree crowns, data points must be filtered. A fixed size window of 1 m x1 m can be used to move through the entire region and only higher values are retained while all other points are deleted. The relatively higher points make up a thin layer which can represent the actual tree crown surface. In this research, the higher value is set as $(\frac{4}{5} * \text{local maxima})$, so all the points which are greater than 4/5*local maxima in 1mx1m are selected (Figure 9).

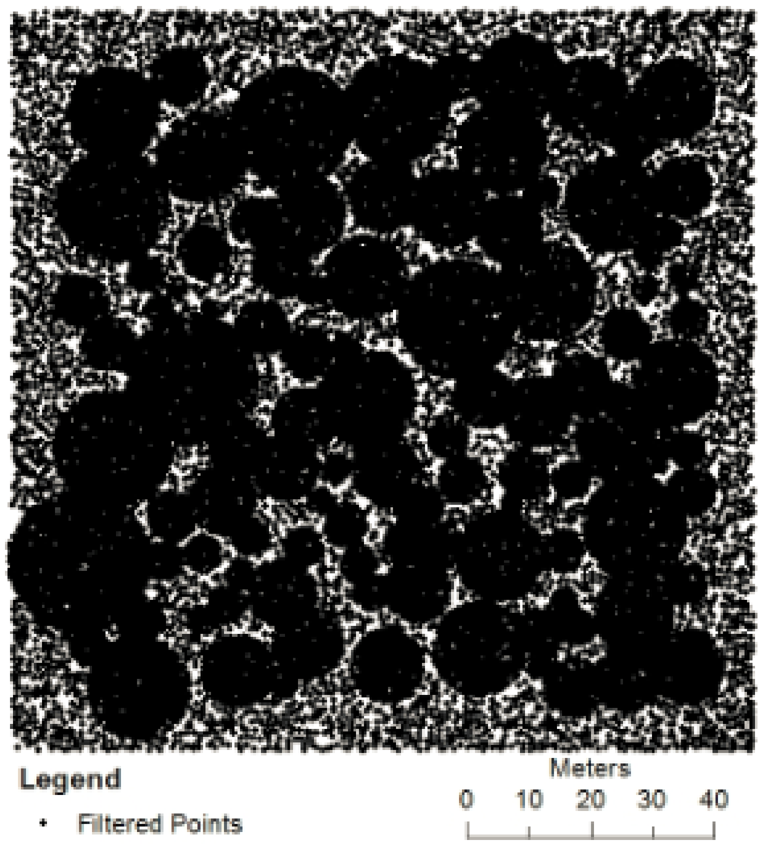


Figure 9 Filtered LiDAR points

Creating Raster Images from Filtered LiDAR Points

LiDAR data points are presented as a series of discrete point clouds. As a result, calculation of 3D shape signatures from original point clouds is much slower than calculation from raster data (Dong 2009). Therefore, creating a continuous surface is considered useful for handling large amounts of data for large areas. Gridded-based image created from LiDAR data can be used to represent the continuous surface.

The simulated forest includes 100 variable-sized trees as is typical of an actual forest ecosystem. If the resolution of the raster image is low, it is difficult to discern the crown shape properties of small trees. If the resolution is too high, it is hard to employ the raster over a large area. For this reason, the raster image in this research is created by interpolation from filtered LiDAR points whose resolution is set as 0.2 meters. Figure 10 shows a smooth raster image.

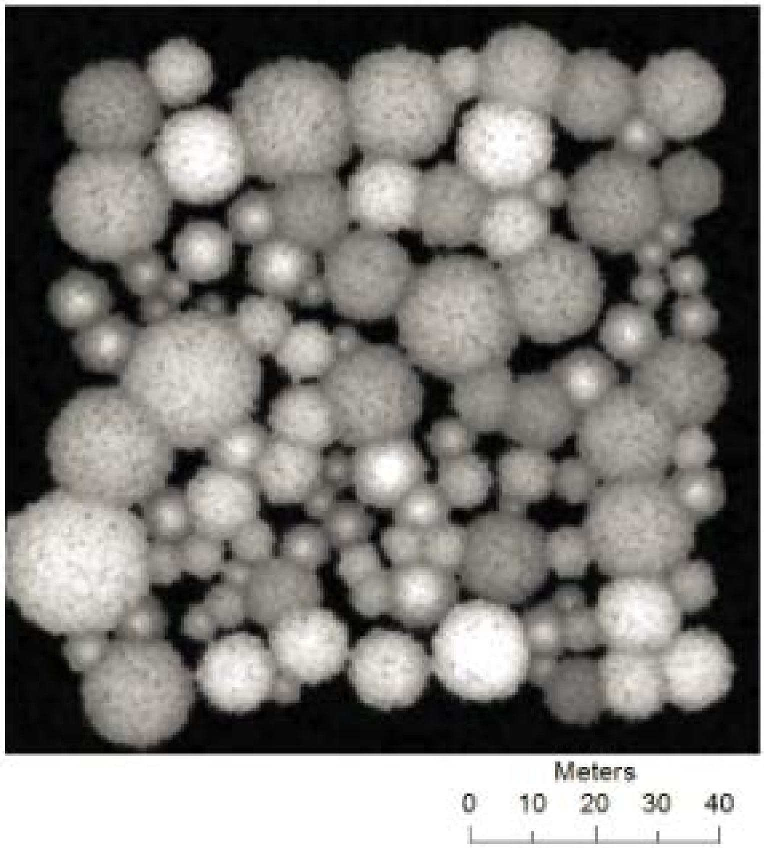
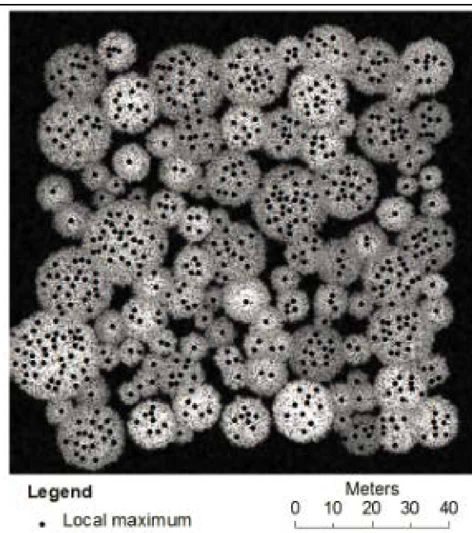
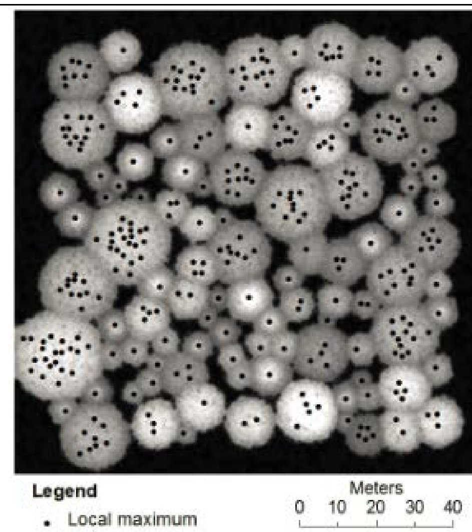
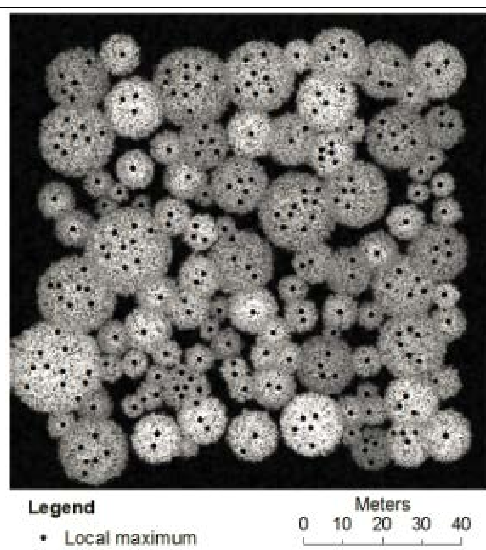
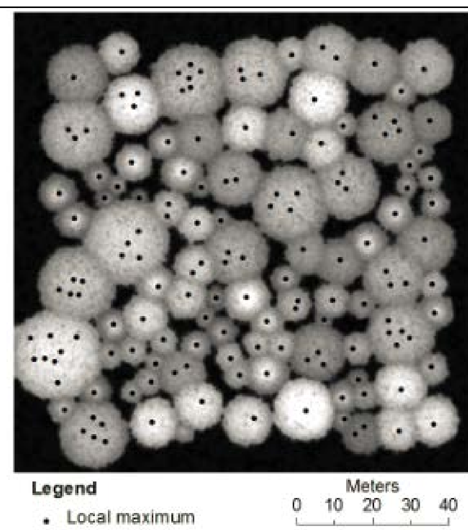
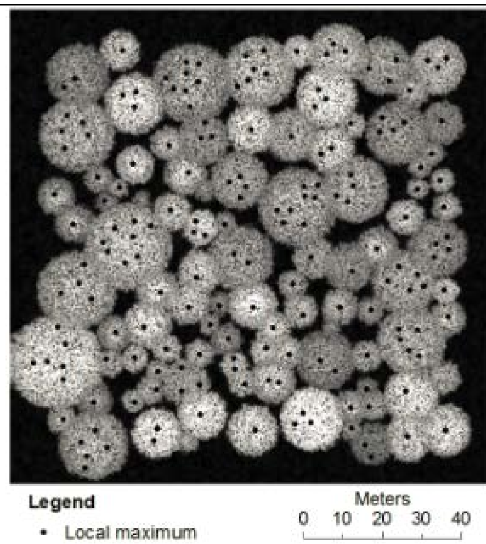
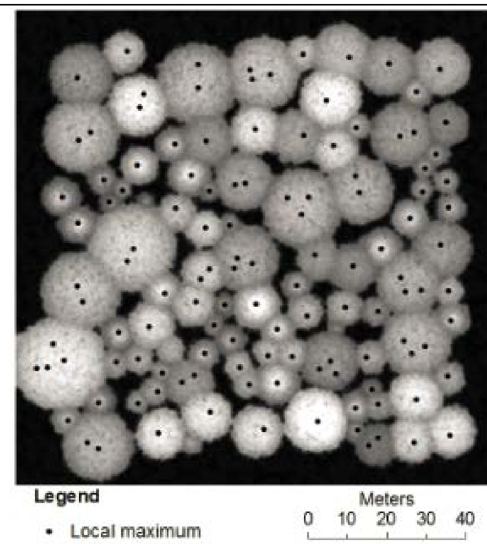
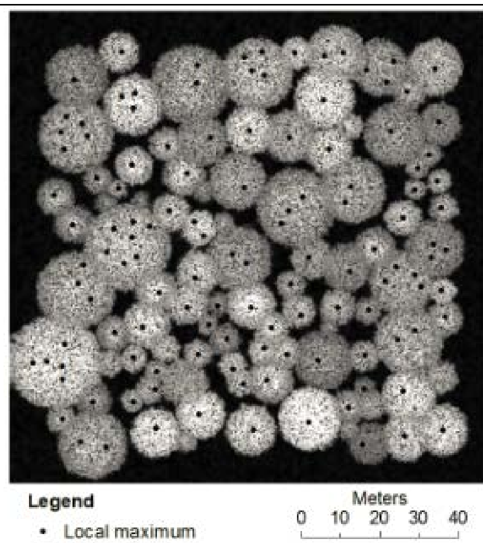
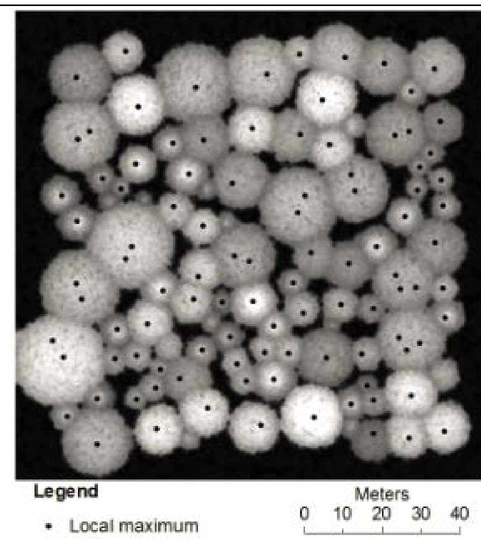


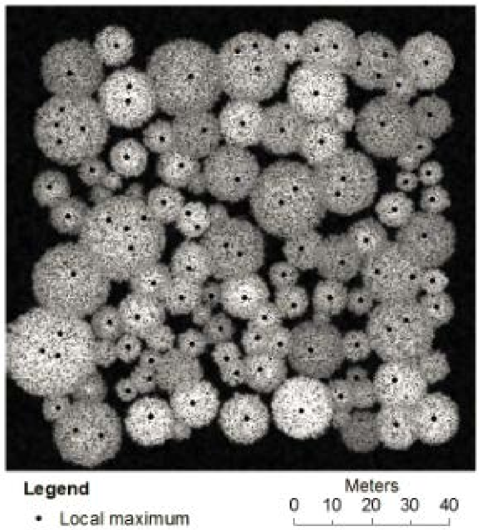
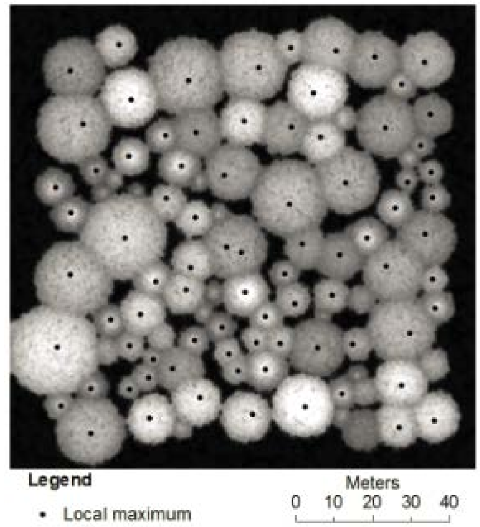
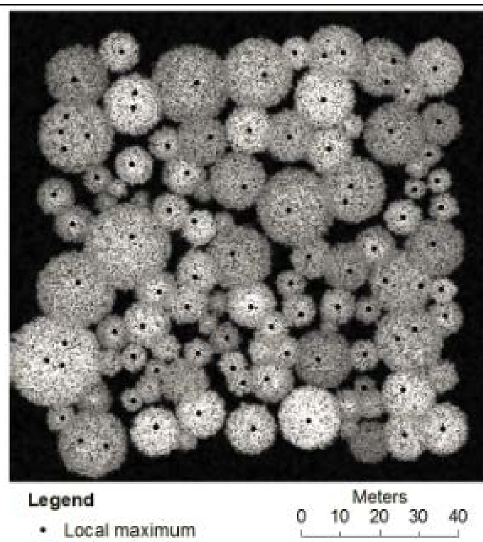
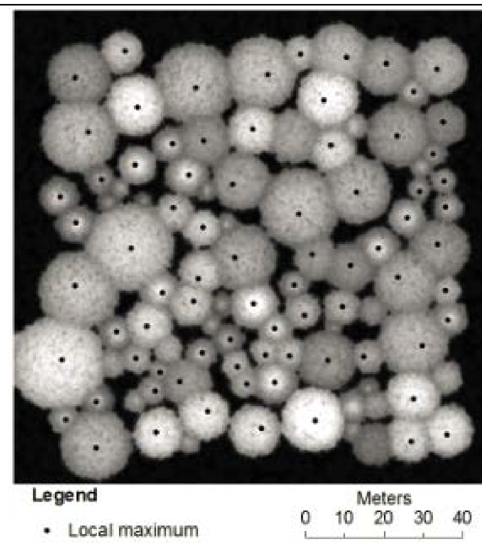
Figure 10 Raster created from filtered LiDAR points (cell size = 0.2m * 0.2m)

Detecting Treetops

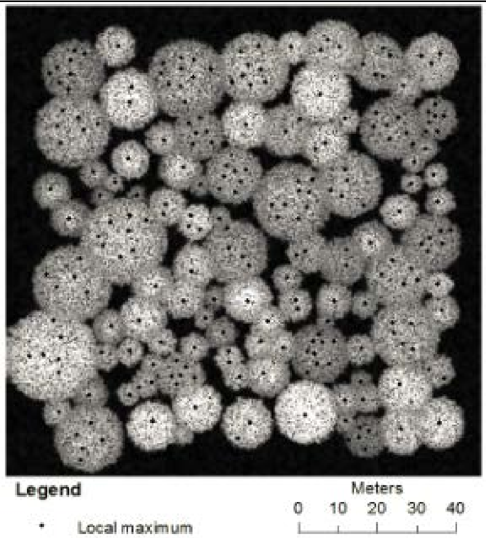
The location of the treetop is the basis for extracting tree height, partitioning the crown space, and further distinguishing tree species. Therefore, determining the treetop location is critical. In previous work, the treetop usually was determined by the brightest point of the crown in the image (Wulder et al. 2000; Culvenor 2002) or the highest LiDAR point (local maximum) of a single tree in LiDAR data (Popescu 2007; Popescu and Wynne 2004). However, these assumptions have some limitations and can only be applicable to homogeneous crowns. Especially when the crown is large, several local maxima can be found on a single crown surface. Both static window filtering and variable window filtering display higher errors in detecting tree locations (Figure 11).

window	Based on initial raster image	Based on filtered raster image
3m * 3m		
4m * 4m		

5m * 5m		
6m * 6m		

<table>
<tr><td>7m
*
7m</td><td>
</td><td>
</td></tr>
<tr><td>8m
*
8m</td><td>
</td><td>
</td></tr>
</table>

Variable window



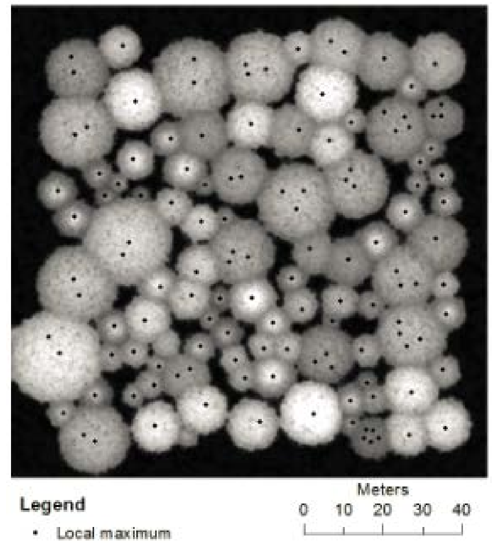


Figure 11 Local maxima detected by static and variable moving windows.

In order to improve the accuracy of treetop detection, a new assumption is proposed. Treetop here refers to the cluster center of higher points (CCHP) in an individual tree. The cluster center is the point, which has the shortest horizontal distance from all other points on a crown surface. It is difficult to find all the points because of overlapping crowns in the forest. Therefore, only higher points (such as the points that are greater than the median height) in the crown area are selected. The distance from the first point (point1) to all other points is calculated and added together to obtain the first distance (D1). Then, the other distances (D2, D3…) can be calculated in the same manner. Finally, the point that has the shortest accumulative distance is the cluster center.

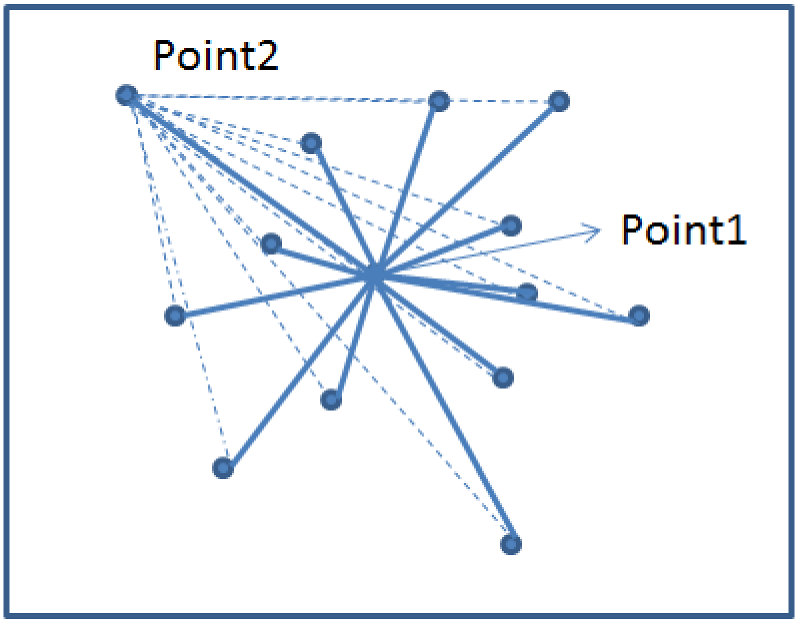


Figure 12 Calculation of cluster center

Crown sizes vary in a forest. Therefore, it is difficult to use a fixed-size window to detect treetops. For example, while a large window covers a large area, small trees cannot be detected. In contrast, a small window can be used to detect small trees, but it will detect more than one treetop for a large crown. In order to detect treetops for tree crowns with different sizes, windows with variable sizes were used. Because LiDAR provides a 3-D product, the slope in a setting window around local maxima can be calculated. The different slopes show different crown shapes and sizes. Thus, the slope can be used to calculate the crown size. When the slope in the setting window is calculated, the crown radius can be calculated as well. In theory, windows of any size can be used, but fewer points will be included in a small window, and the slope calculated using a small window may not be accurate, making it difficult to analyze the crown properties. Therefore, here the radius of the window is set at 2 meters and the crown radius can be determined by equation 8 as demonstrated in Figure 18.

$$R = \frac{\mathrm{r}}{\sin(180 - 2 * A)} \tag{8}$$

where

r = preset window radius.

A = angle between the line from a point on the crown surface to treetop and the vertical direction.

R = radius of calculated tree crown.

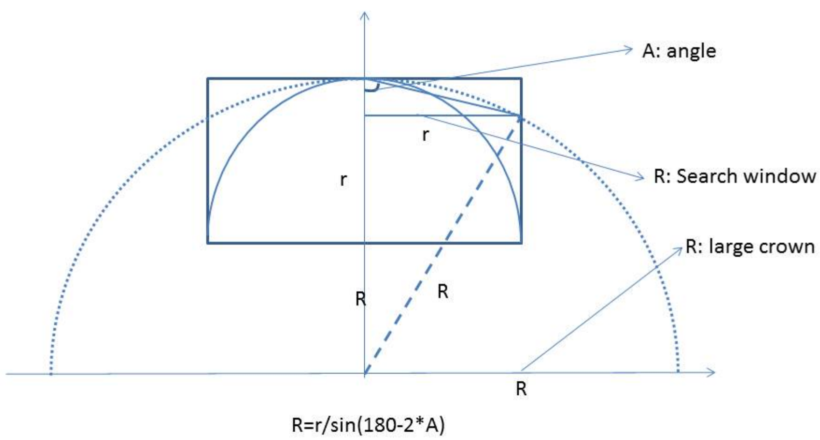


Figure 13 Calculation of variable windows

Even though crown slope centered in each pixel can be calculated, the process consists of a large amount of computation and the running speed will be very slow. In order to avoid calculating all pixels' slopes and improve computing speed, local maxima detected by a small window (3m x 3m) can be used as basic points to further detect tree location. Crown slope centered in each basic point in a setting window (r = 2m) can be calculated, and then the radius of crown can be calculated by equation 8. Since some trees are overlapped by others, only part of the crown whose radius equals to the half of whole radius is used to represent its own and uncovered crown. The cluster center of higher points (CCHP) in this small crown space can be calculated. If the calculated cluster center is same as the basic point, it is considered a tree location. Otherwise, the cluster center is set as new basic point and a new cluster center is calculated; this process can be iterated until the calculated cluster center is the same as the basic point. Following this iteration, the final tree location can be detected (Figure 14).

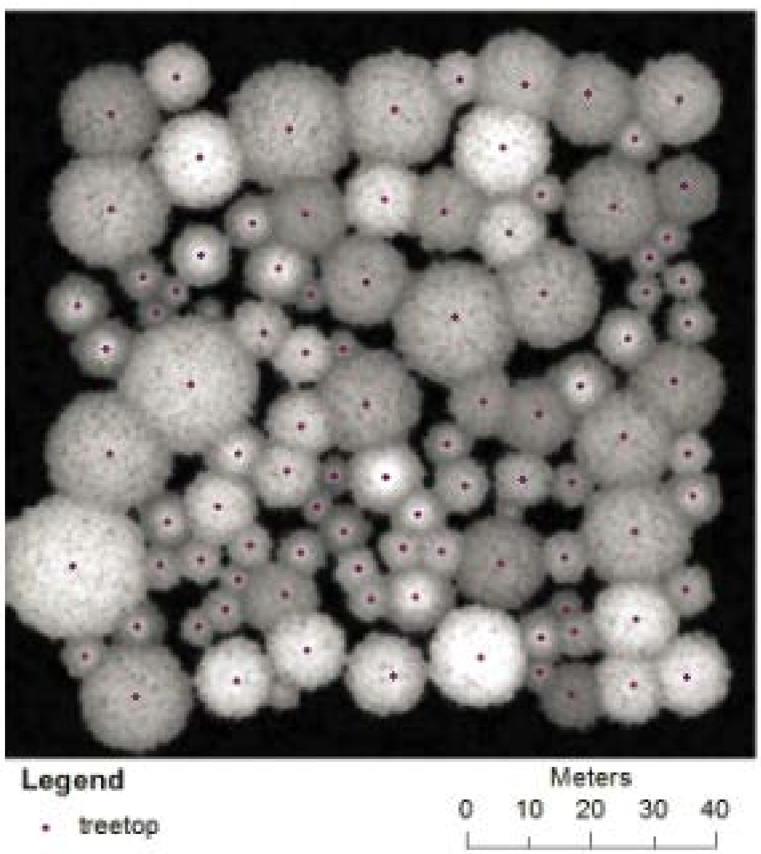


Figure 14 Treetops detected by CCHP

Thus far, treetops have been detected and a relatively smooth raster has been created from filtered LiDAR points. Therefore, the crown space can be delineated.

Usually, the crown edge represents the lowest part of the crown. The lowest points between any two adjacent trees can be considered as the boundary separating the two trees. Several low points between each tree and adjacent trees can be used to determine the edge of the crown. The circles around treetops in this area can be drawn to represent the crown space (Figure 15).

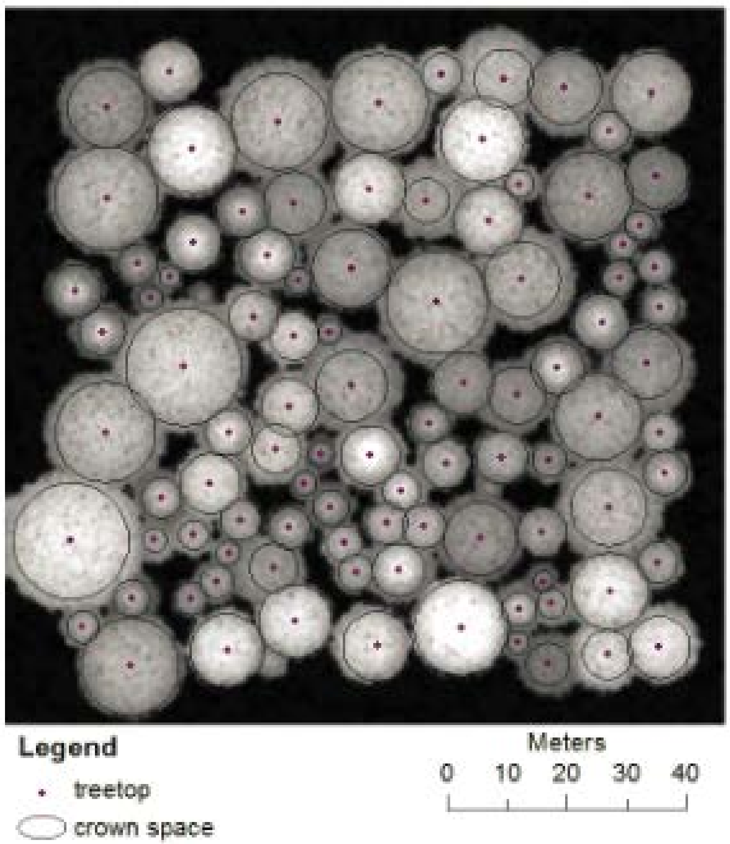


Figure 15 Delineated crown space by lowest point between two adjacent trees

Discriminating 3D Crown Shapes

It is difficult to directly compare two 3D objects based on their shapes (Osada 2002). The most common method is to convert the 3D shape to 2D functions (Osada 2002). These 2D functions are called shape signatures of 3D objects, which is the probability distribution representing the geometric properties of the 3D objects. There are several kinds of shape signatures based on distance, aspect, or slope. We can test these different signatures to determine the best one and use it to analyze the types of tree crowns. The similar shape signatures represent the same type of trees. The shape signatures can be evaluated quickly and are not sensitive to scale, rotation or translation (movement) of objects.

RESULTS AND DISCUSSION

Results from Simulated Data

After detecting treetops, partitioning tree spaces and determining tree species, the effectiveness of this process can be analyzed by comparing the results with the initial data collected when the forest was simulated. The variables compared include: number of trees, location of treetops, and types of trees.

Number of Trees

When the virtual forest was simulated, 100 trees were generated in an area of 100m x 100m. In order to analyze the effectiveness of the proposed method in detecting the number of treetops, results were compared to those obtained using traditional methods. Traditional methods usually assume the local maximum as the treetop; this local maximum can be detected using a fixed moving window or a variable-size window. The new method detected a cluster center of higher points as the treetop. In this case, the range of fixed window size is set from 3 meter to 10 meter. Variable window size is calculated by the regression equation: crownwidth = 2.51503 + 0.00901 * height * height (Popescu and Wynne 2004). The numbers of trees detected by the different methods are listed in Table 2.

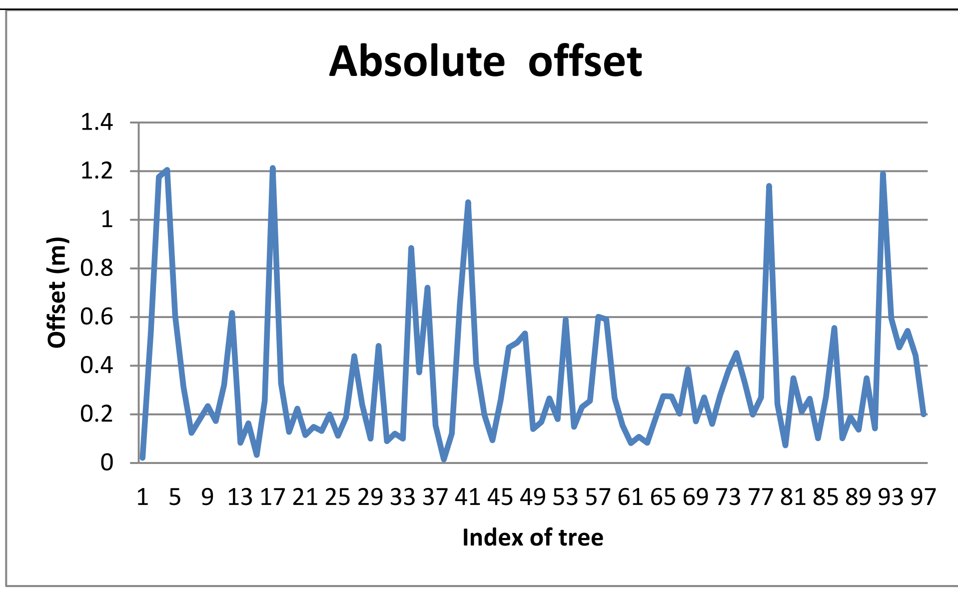
Table 2 Number of trees detected using different methods

Method \ Result (Raster)		Based on original raster			Based on filtered raster		
		Commiss-ion error	Omiss-ion error	Correct tree	Commiss-ion error	Omiss-ion error	Correct tree
Fixed wind-ow	3m*3m	63	0	37	40	0	60
	4m*4m	44	1	55	21	1	78
	5m*5m	30	5	65	19	4	77
	6m*6m	17	11	72	10	11	79
	7m*7m	14	18	68	1	19	80
	8m*8m	9	23	69	0	22	78
	9m*9m	5	28	67	0	29	71
	10m*10m	3	33	64	0	35	65
Variable window		33	1	66	20	2	78
New method		---	---	---	0	3	97

Comparison of the results shows the variation in precision among the three methods. Results from a fixed moving window approach depend on window size. Small windows detected fewer treetops with higher commission error but large windows detected more treetops with higher omission error. In this case, the maximum number of treetops detected by the fixed window approach is 80 and 19 small trees are missed. The second largest number of treetops is 79 and 10 large trees are falsely detected, which are very important in estimating biomass. The number of treetops detected using a variable window is 78, but two small trees are missed and 20 large trees are falsely detected. This is similar to the result obtained using a 4m*4m or 5m*5m window in the fixed window method. The total number of treetops detected using the new method is 97. By this method, all the large trees are successfully detected and only three small trees which are much closed to some large tree crowns are missed.

Location of Treetop

When the points representing the location of trees are detected using the new method, the coordinates of these points are determined at the same time. Since the real coordinate of each tree and the radius of each crown was recorded when these trees were simulated, the absolute offset and relative offset between the detected treetop and real treetop can be calculated. Therefore the precision of each tree location can be analyzed by comparing the result and the initial coordinate. The absolute offset is the distance between detected treetop and the real treetop, while the relative offset equals the absolute offset divided by radius of tree crown. Absolute offset and relative offset are shown on Figure 16.

Absolute offset	

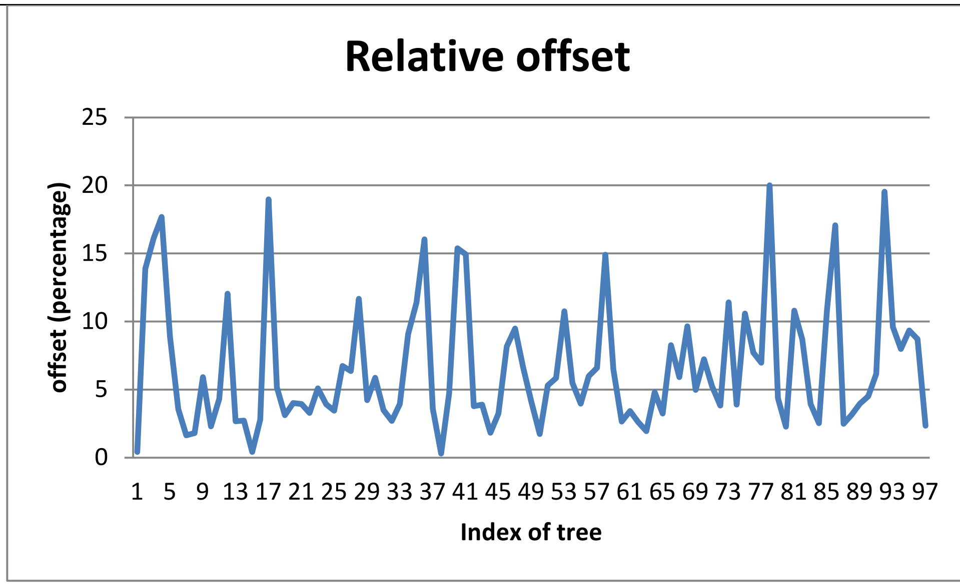
Relative offset	

Figure 16 Absolute and relative offset.

Absolute offset indicates that the greatest distance between the initial and detected location of all the trees is less than 1.2 meters; most of the offsets are range from 0-0.6 meter. In relative terms, the maximum offset between the initial and detected location is less than 20 percent of crown width but most detected trees offset from the initial trees are less than 10 percent of crown width. Based on research by Dong (2010), an offset value of less than 20 percent of crown width is not likely to affect determination of crown shape. In sum, results of the new method match the real location well, meaning this technique is useful to further delineate crown space based on the detected treetops. Both real and detected x-coordinates are shown in Figure17; both real and detected y-coordinates are shown in Figure 18.

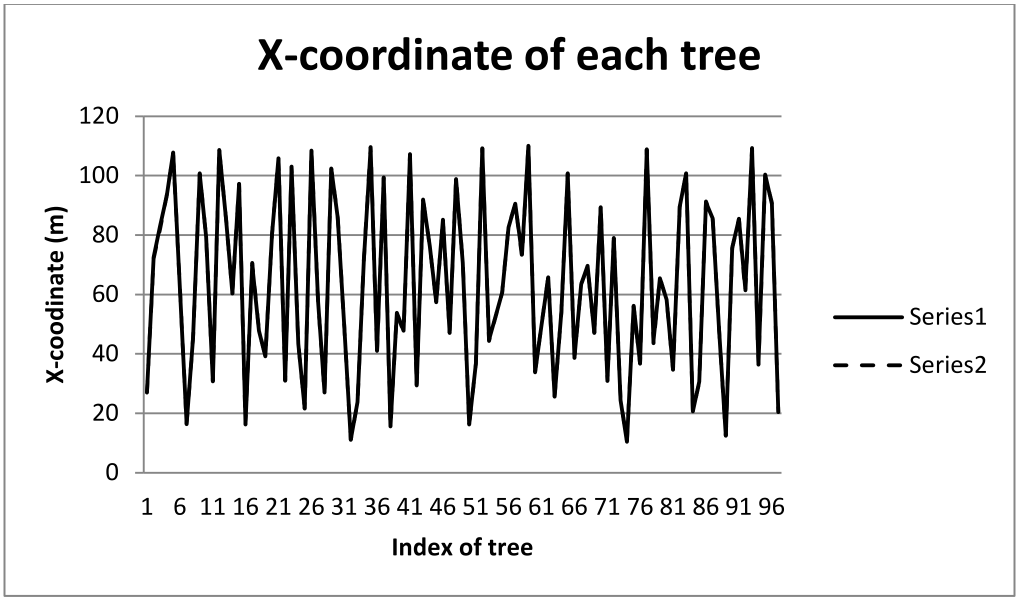


Figure 17 Comparison of X-coordinate between real and detected trees location

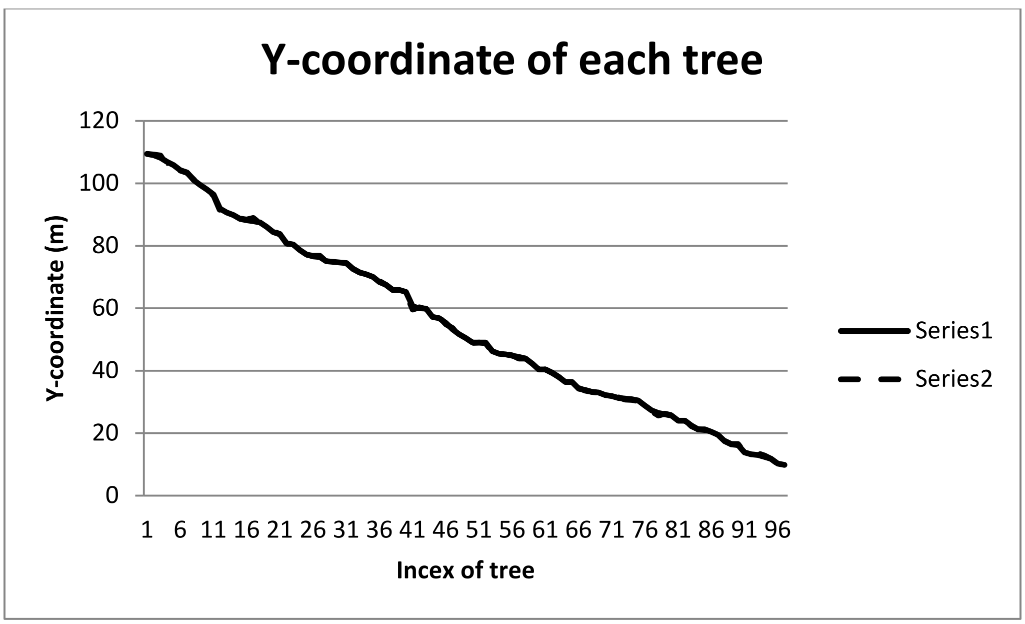


Figure18 Comparison of Y-coordinate between real and detected trees location

Tree Types

In order to distinguish the different crown shapes, two shape signatures were selected. The first signature is the angle signature, which represents the distribution of angles between the vertical axis in treetop points and the lines between treetops and random points on crown surface. The second signature is the normalized height signature, which represents the distribution of normalized height of random point on the tree crown. Normalized height is calculated by dividing the height of random point by the radius of the detected tree crown.

In this case, 1000 random points on each crown surface were first selected, and then the angles of these points were calculated. Usually the angles of random points were less than 90 degrees. Therefore when 90 degrees were separated into 30 bins, each bin contained 3 degrees. Every angle can be put into one bin. Percentage distribution can be calculated using the number in each bin divided by the total numbers (1000). The following figures show the angle percentage distribution of cone, ellipsoid, and hemisphere.

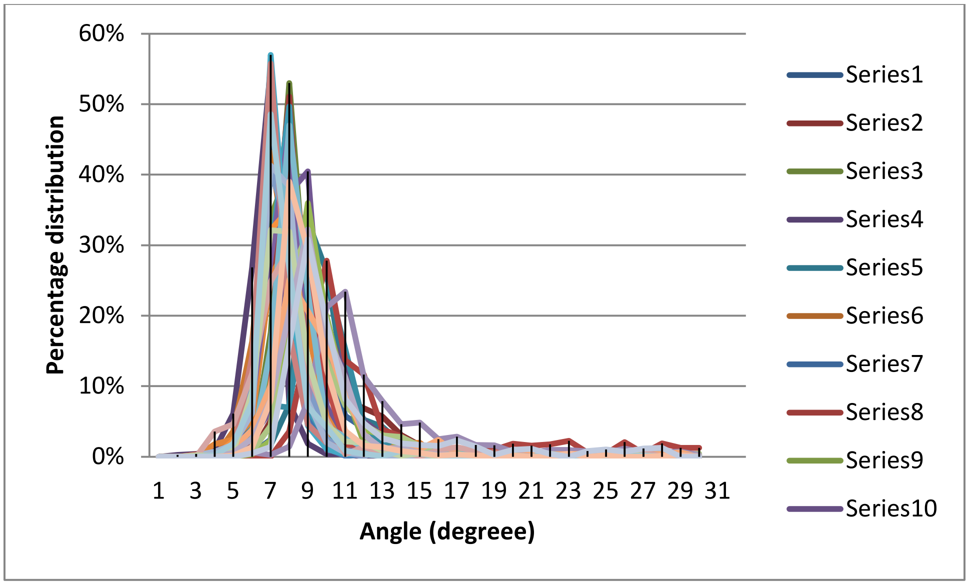


Figure 19 Angle shape signature of cone

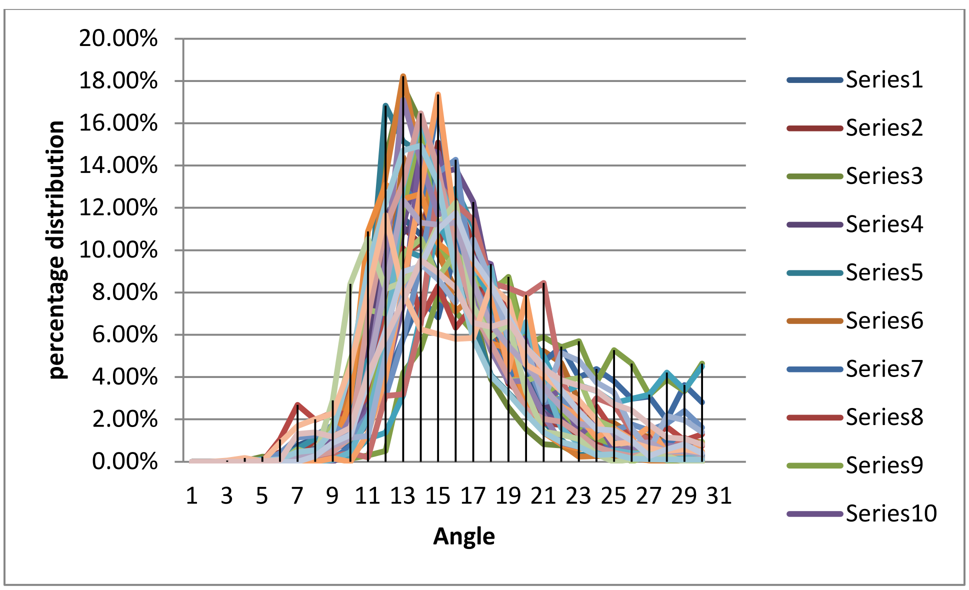


Figure 20 Angle shape signature of half ellipsoid

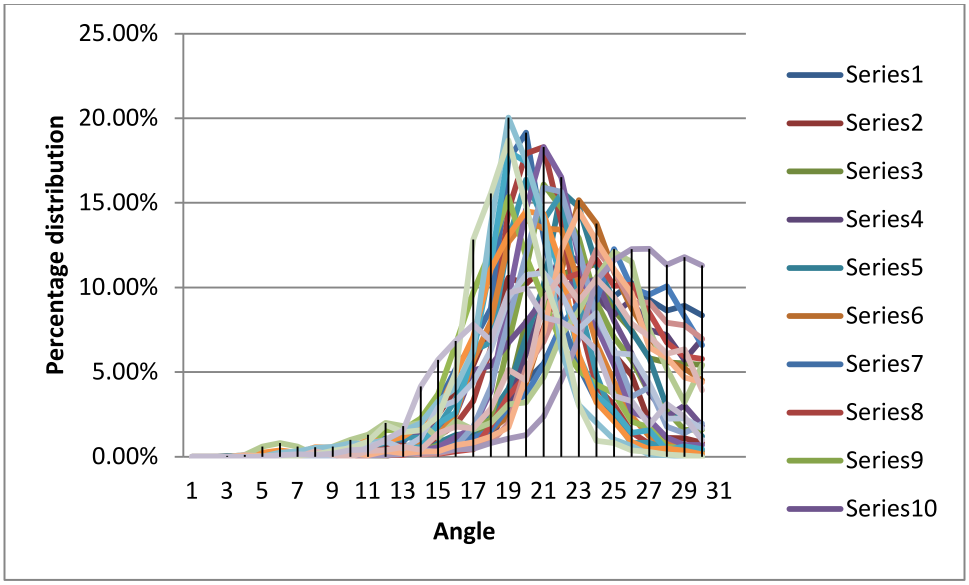


Figure 21 Angle shape signature of hemisphere

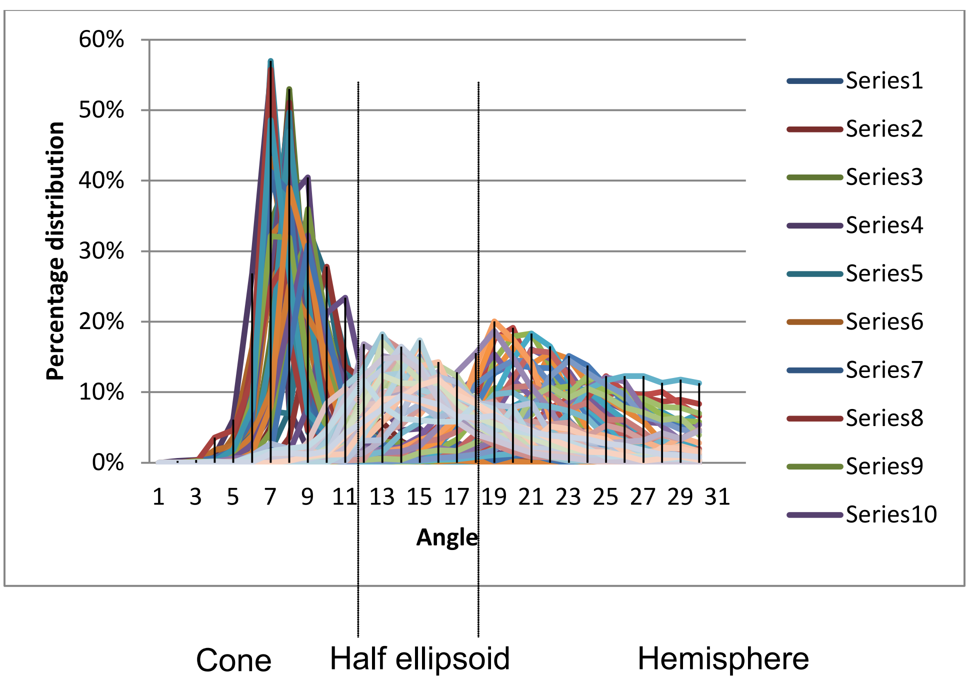


Figure 22 Angle shape signature of all simulated trees

The shape signatures based on angles of random points display the characteristics of the three different shapes in 2D. Cones gather along the left, half ellipsoids gather in middle, and hemispheres gather along the right side hand of the graph. Even though the signatures of the same kinds of shapes cluster together, the peaks are different and it is difficult to find the boundary between any two clusters. Therefore, weighted average value rather other the peak of curve can be used to represent the location of each curve (Figure23).

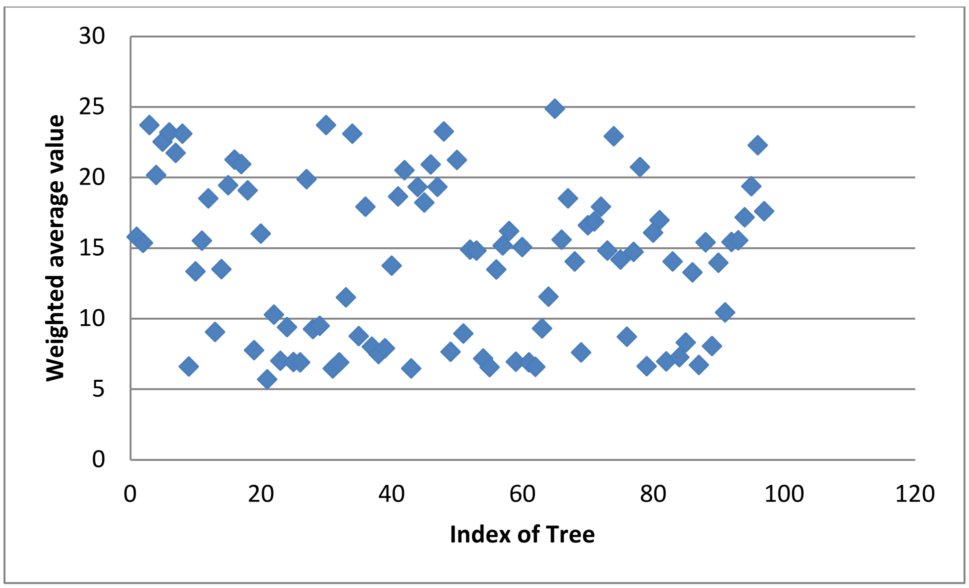


Figure 23 Weighted average value of each Curve. Each of them is located by value on the curve as weights

Similar to the angle signature, the normalized height signature can also represent crown shapes in 2D. The normalized height signature can be automated drawn and the location of each curve can be calculated as angle signature which will show the similar result as angle signature. When these two kinds of signature are calculated and combined together, three clusters with very clear boundaries can be seen. Each cluster represents one crown shape category (Figure 24).

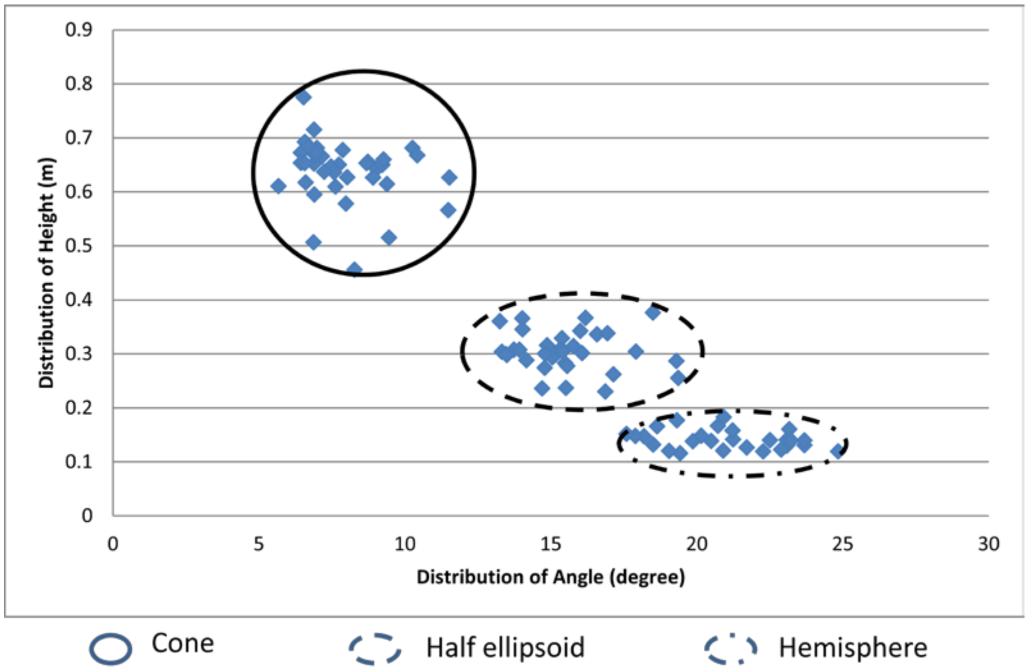


Figure 24 Combination of angle signature and normalized height signature. The horizontal axis is angle distribution and the vertical axis is normalized height distribution.

It can be seen from Figure 24 that the cones gather in upper left corner with higher height distribution and lower angle distribution, the hemispheres gather in the lower right corner with lower height distribution and higher angle distribution, and the half ellipsoids gather in the middle with middle height distribution and middle angle distribution.

Results from Real Forest

Following the steps used to detect treetops and discriminate tree crowns in the simulated forest, the real forest can be analyzed. The real LiDAR data includes several categories: ground, default, 3rd stop, low point, aerial points, and isolated points. Ground category is the points reflected from ground which can be used to create digital elevation model (DEM). Default points are those reflected from vegetation and above the ground artificial structures, which can be used to create digital surface model (DSM). During the process of creating the raster image, the raw LiDAR points were filtered first by 1mx1m moving window, only the points greater than $(\frac{4}{5} * \text{local maxima})$ in each window were selected. Therefore the raster created by these filtered LiDAR points shows smooth canopy. The canopy height model or normalized digital surface model (nDSM) can be created by subtracting DEM from DSM (Figure 25). Based on nDSM, treetops were detected using the new method (Figure 26), and the tree crown space was discriminated (Figure 27).

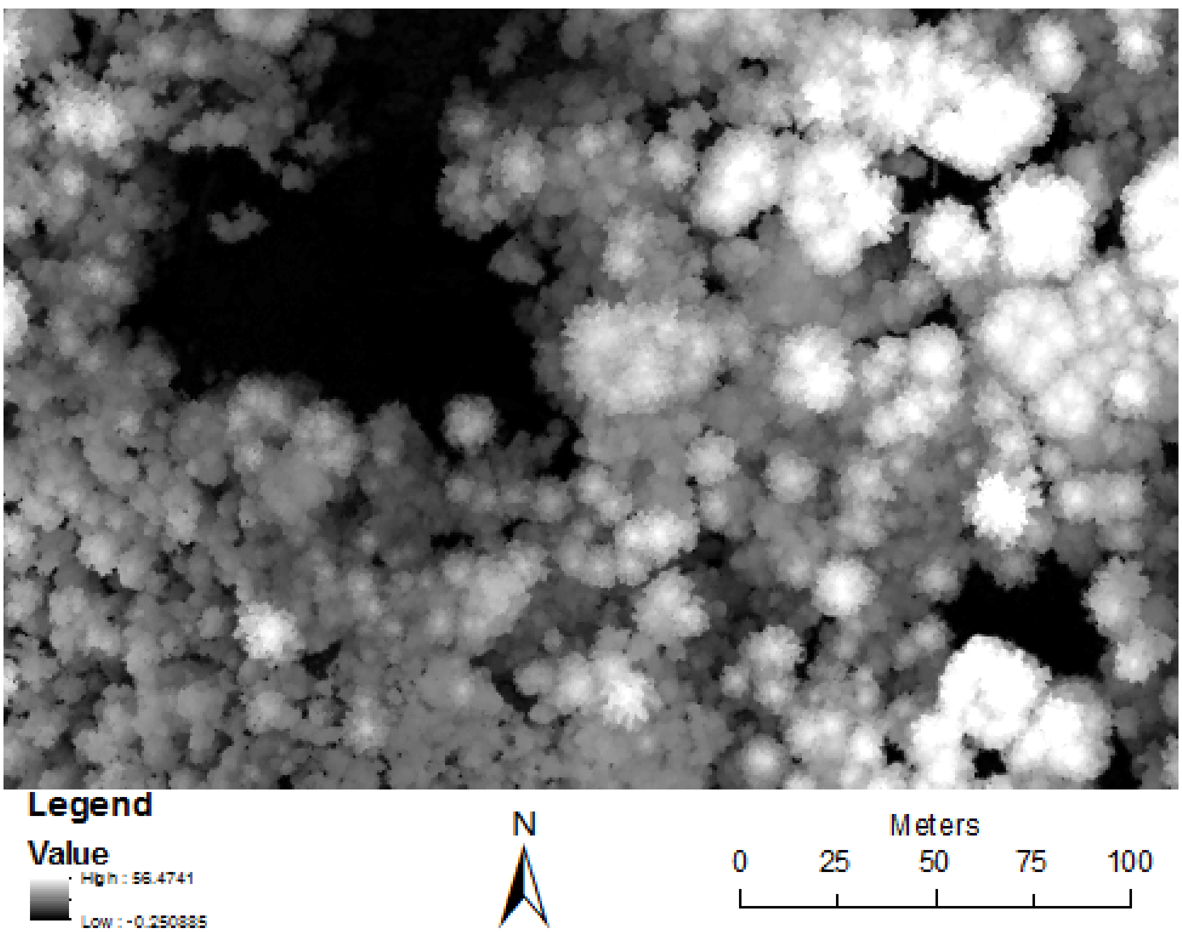


Figure 25 Raster image created from filtered LiDAR data

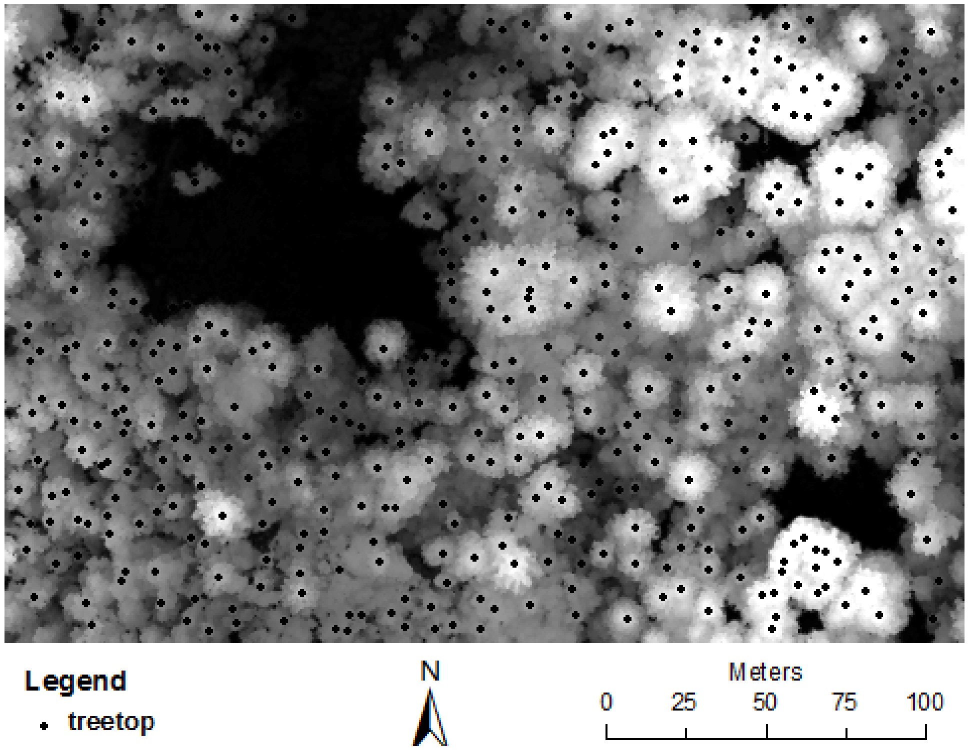


Figure 26 Treetops detected by the new method

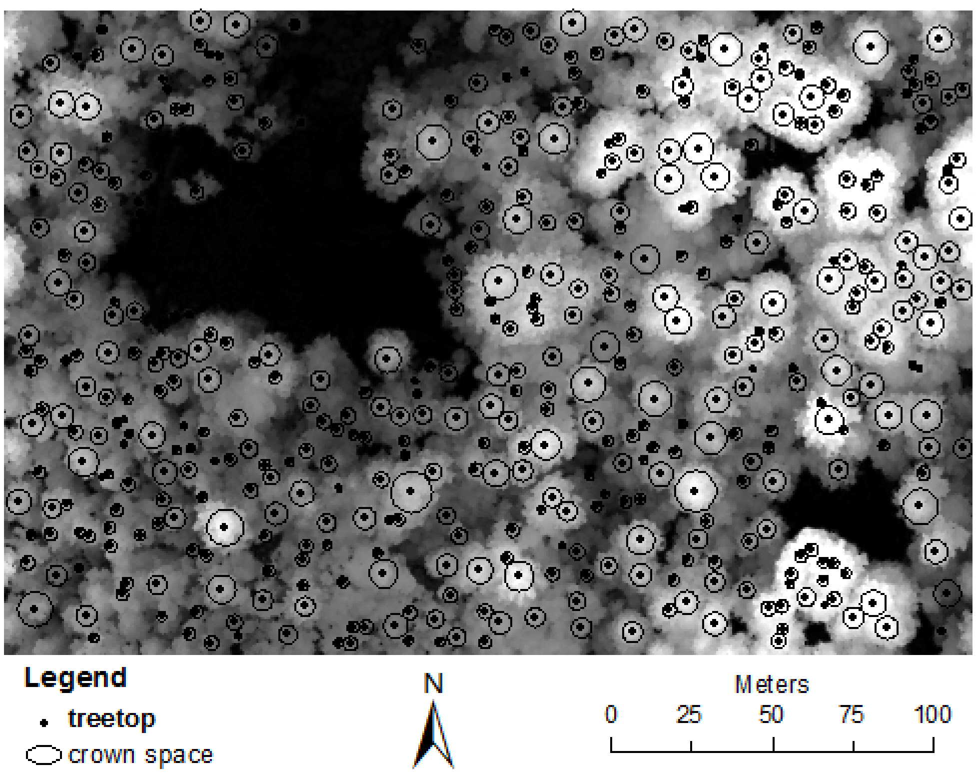


Figure 27 Delineated crown space by lowest point between two adjacent trees

Since different trees have different crown sizes and crowns that are too small do not contain enough information to display the crown shape, only the large crowns (radius > 2 meters) were selected and analyzed in order to accurately study the crown shape and determine tree species. The following map shows the large crown spaces (Figure 28).

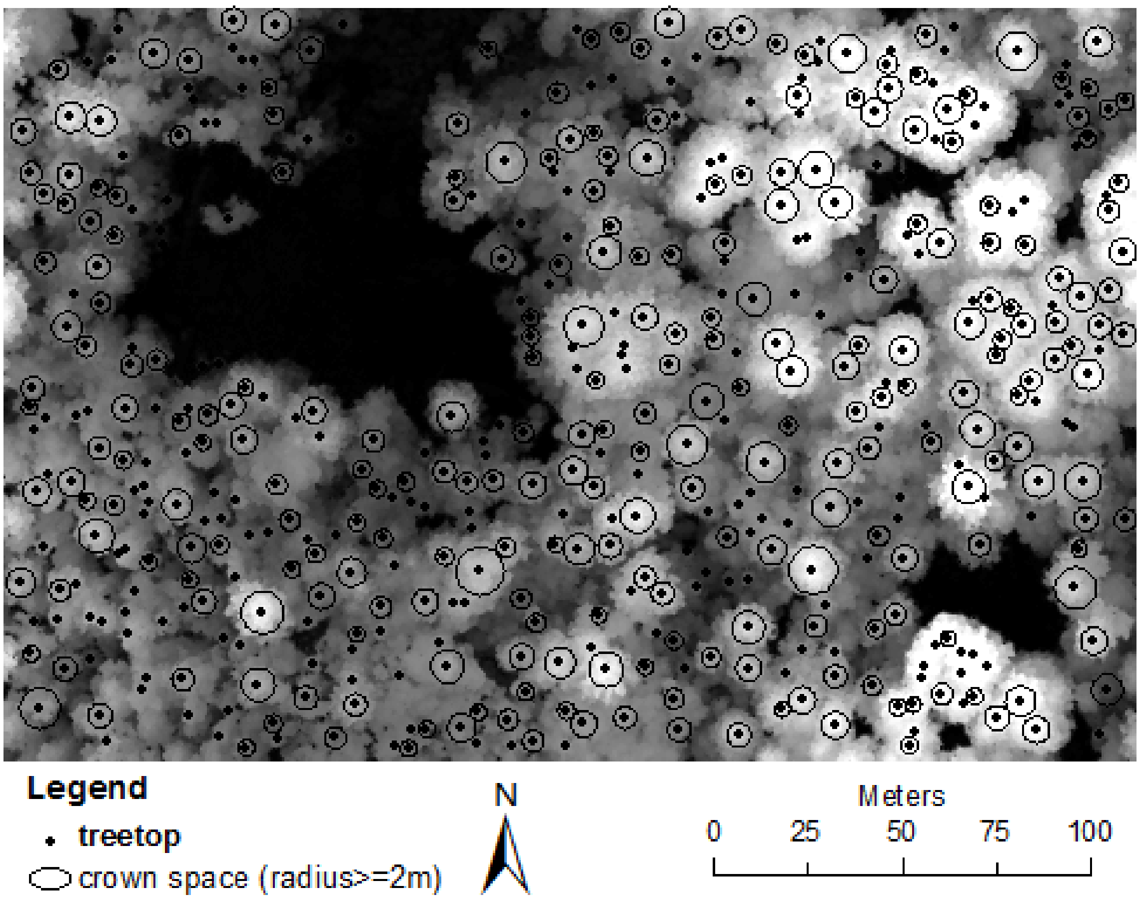


Figure 28 Large crown trees (radius >= 2m). These trees have large crown for precisely analyzing characteristics of crown shape.

Angle signature and normalized height signature were calculated and combined. The location of each tree is displayed in the Figure 29.

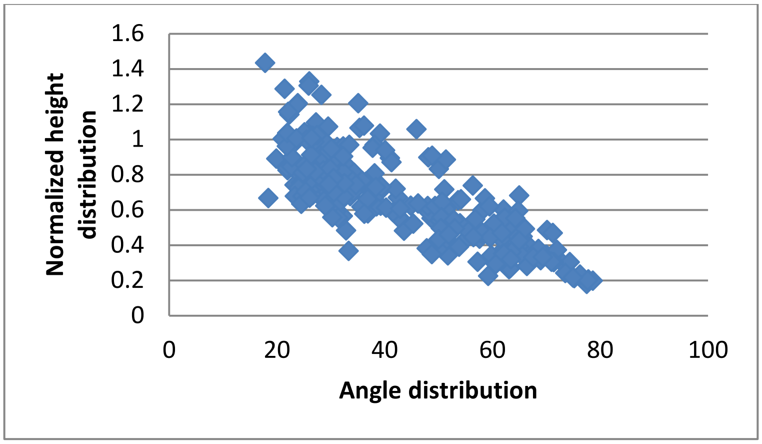


Figure 29 Combination of angle distribution and normalized height distribution

In general, the points are divided into two groups. In order to further analyze what kinds of trees these groups of points represent, the field data can be compared with this distribution. The field data includes 14 oak trees, 7 Douglas fir and 17 coast redwood trees. These trees are located around two open spaces in the forest. These trees are shown in the following image (Figure 30).

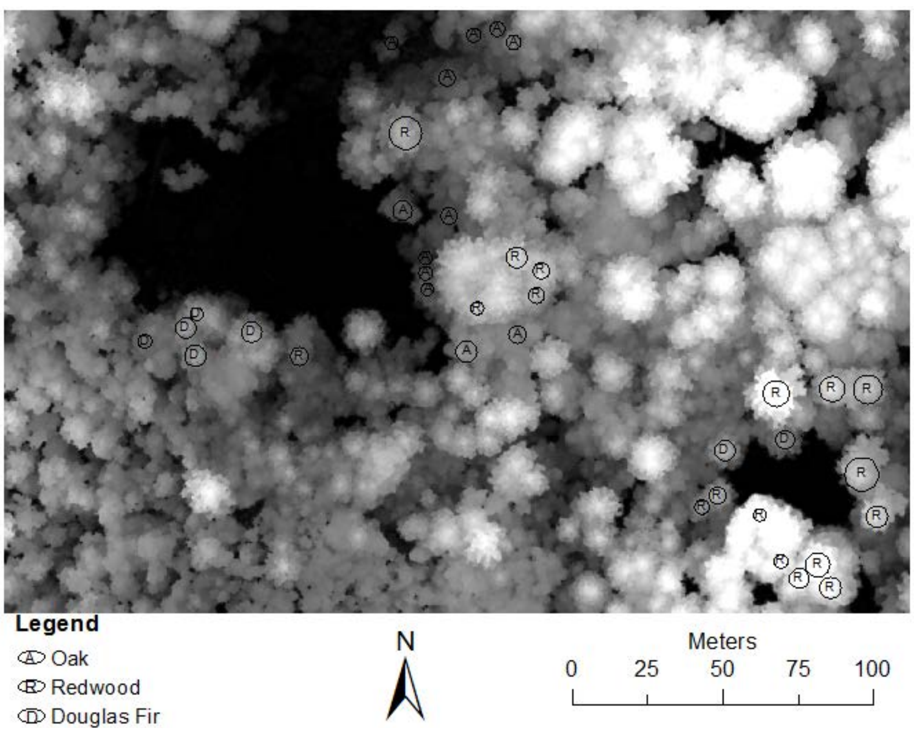


Figure 30 Selected testing trees. These trees had been identified in field and were selected to test the result obtained from LiDAR data.

All these trees were identified in the field and their signatures including angle distribution and normalized distribution were calculated. Therefore the location of each tree can be found in the graph of signature combination. When all of them are labeled on the figure, the cluster of each kind of tree can be distinguished (Figure 31).

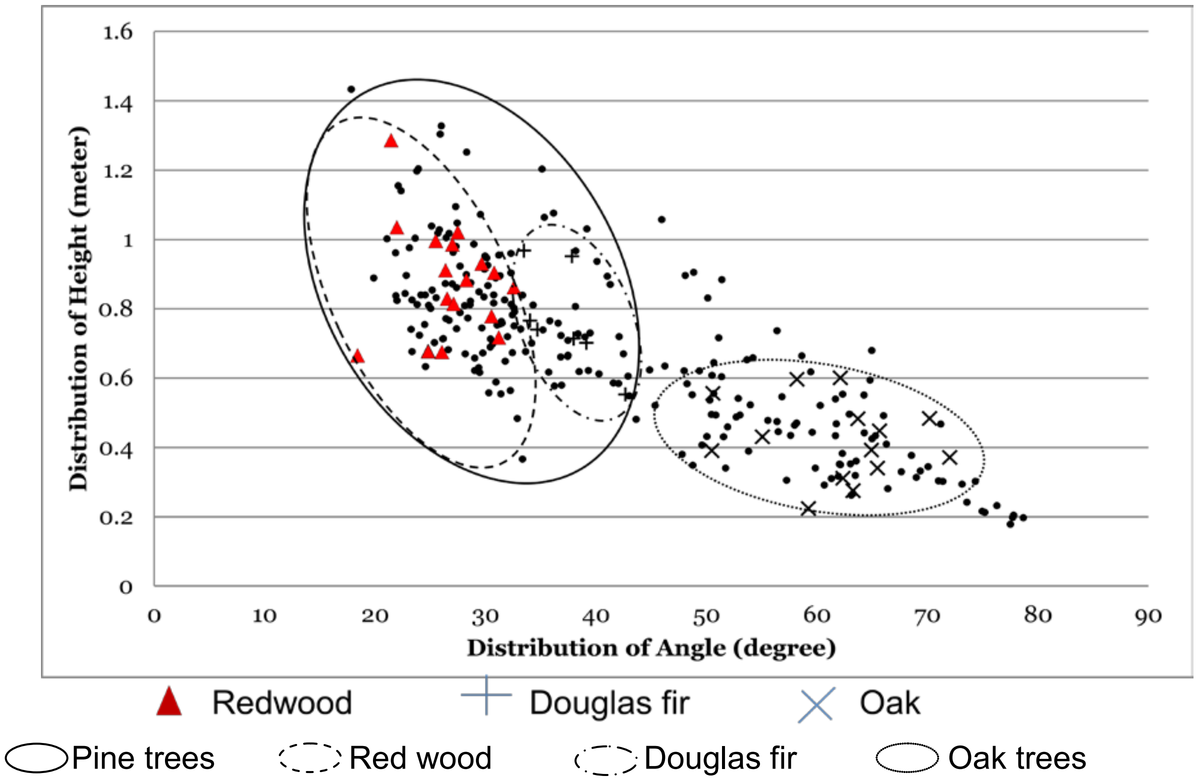


Figure 31 Tree clusters. Trees with similar shape crown gather together.

Oak trees can be seen along the right side and the group is characterized by a horizontal ellipse, which implies normalized height signature of the tree crown with gentle slope change less than angle signature. Redwoods gather along the left right side, the group is characterized by a vertical ellipse which implies normalized height signature of the tree crown with gentle slope change greater than angle signature. Douglas firs gather in the middle part and share similar characteristics as redwoods. Since redwood and Douglas fir groups are close and partly overlap with each other, it is difficult to differentiate between them. Therefore, the two groups can be combined into one pine group, which is consistent with reality: redwood and Douglas firs have similar crow shapes.

CONCLUSION

In this study, a new method of treetop detection is proposed in which the treetop is determined by the cluster center of selected points rather than the highest point. During this process, the possible radius of the crown is calculated based on the slope in a window centered on the local maxima, and then the cluster center in half radius window is calculated. If the cluster center is the same as the local maxima, it is considered to be the treetop. Otherwise, a new window is set centered in the cluster center, and then new slope and new radius are calculated. Furthermore, new cluster center is calculated again. This process is repeated until the last two cluster centers are the same point which represents treetop.

The new method was first applied to a simulated forest. Result from this virtual forest show that the new method works well in a mixed forest. A total of 97 percent trees were successfully detected, and all the large trees were detected. The detected location of each tree is close to the real location. The greatest error was less than 20 percent of the crown width. Therefore the treetop detected by the new method was deemed accurate enough to be used for further crown space delineation. Moreover, the angle signature and the normalized height signature successfully distinguished different crown shapes from one another. The combination of these two signatures obviously separates 97 out of 100 trees into three groups with clear boundaries.

The above method was also applied to real LiDAR data. Because of the limitation of parameters collected in the field, only 38 trees were available to test the results. However, all these trees were detected by the new method. Furthermore, the high resolution images available on the Internet (such as Bing Maps and Google Maps) can be used to check the results. Visual comparison showed most of the trees were successfully detected. The combination of the angle signature and the normalized height signature can separate trees into two main groups: oak trees and pine trees, where the pine trees include coast redwood and Douglas fir.

The new method (CCHP) has high accuracy in detecting treetops in both simulated and real forests. The two key processes of this method are calculating cluster centers and calculating possible radius of crown. The advantage of using cluster centers for representing treetops is to reduce or eliminate the influence of extreme values. Due to the limitation of real data collected in field, the effectiveness of new method cannot be completely tested. Future research is recommended to improve the method by collecting more real data for test.

REFERENCES

Armson, D., Stringer. P., Ennos, A.R. 2012. The effect of tree shade and grass on surface and globe temperatures in an urban area. *Urban Forestry & Urban Greening* 11: 245-255.

Ben-Arie, J.R., Hay, G.J., Powers, R.P., Castilla, G., St-Onge, B. 2009. Development of a pit filling algorithm for LiDAR canopy height models, *Computers & Geosciences* 35: 1940-1949.

Calder, I.R. 2007. Forests and water---Ensuring forest benefits outweigh water costs. *Forest Ecology and Management* 251: 110-120.

Carle, J., Vuorinen, P., Del, L.A. 2002. Status and trends in global forest plantation development. *Forest Products* 52: 7.

Costello, R., Chum, H.L. 1998. Biomass, bioenergy, and carbon management. *Expanding Bioenergy Partnerships*. http://boc.mtholyoke.edu/courses/tmillett/course/geog_304B/25695.pdf

Culvenor, D.S. 2002. TIDA: an algorithm for the delineation of tree crowns in high spatial resolution remotely sensed imagery. *Computers & Geoscience* 28: 33-44.

Dong, P.L. 2009. Characterization of individual tree crowns using three-dimensional shape signatures derived from LiDAR data. *International Journal of Remote Sensing* 30: 6621-6628.

Dong, P.L. 2010. Sensitivity of LiDAR-derived three-dimensional shape signatures for individual tree crowns: a simulation study, *Remote Sensing Letters* 1: 159-167.

Dralle, K. and Rudemo, M. 1997. Stem number estimation by kernel smoothing of aerial photos. *Canadian Journal of Forest Research* 26: 1228-1236.

Estornell, J., Ruiz, L.A., Velazquez_Marti. B., Fernandez_Sarra. A. 2011. Estimation of shrub biomass by airborne LiDAR data in small forest stands. *Forest Ecology and Management* 262: 1697-1703.

Fao 2010. Global forest resource assessment. Fao Forestry Paper 163.

http://www.fao.org/docrep/013/i1757e/i1757e.pdf.

Fao (2012). Forests, jungles, woods & their trees.

http://wwf.panda.org/about_our_earth/about_forests/.

Holmgren, J., Nilsson, M., Olsson, H. 2003. Estimation of tree height and stem volume on plots using airborne laser scanning. *Forest Science* 49: 419-428.

Kartha, S. 2001. Biomass sinks and biomass energy: key issues in using biomass to protect the global climate. *Energy for Sustainable Development* 5: 10-14.

Lee, W.K., Gadow, K.V., Chung, D.J., and Lee J.L. 2004. DBH growth model for Pinus densiflora and quercus variablilis mixed forests in central Korea. *Ecological Modelling* 176: 187-200.

Leeuwen, M.V., Coops, N.C., and Wulder, M.A. 2010. Canapy surface reconstruction from a LiDAR point cloud using Hough transform. *Romote Sensing Letters* 1: 125-132.

Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G., Spies, T.A., and Harding, D. 1999. Lidar remote sensing of the canopy structure and biophysical properties of douglas-fir western hemlock forests, *Remote Sens. Environ* 70: 339-361.

Lim, K., Treitz, P., Wulder, M., St-Onge, B., Flood, M. 2003. LiDAR remote sensing of forest structure. *Progress in Physical Geography* 27: 88-106.

Osada, P., Funkhouser, T., Chazelle, B., Dobkin, D. 2002. Shape distributions. *ACM Transaction on Graphics* 21: 807-832.

Parker, G.G. 1995. Structure and microclimate of forest canopies. *Forest Canopies*, P. 73-106.

Prentice, C. S., Crosby, C. J., Whitehill, C. S., Arrowsmith, J. R., Furlong, K. P., and Phillips, D. A. 2009. Illuminating northern california's active faults. *American Geophysical Union* 90: 55-56.

Popescu, S.C. 2007. Estimation biomass of individual pine trees using airborne LiDAR. *Biomass and Bioenergy* 31: 646-655.

Popescu, S.C., Wynne, R. 2004. Seeing the trees in forest: using LiDAR and multispectral data fusion with local filtering and variable window size for estimating tree height. *Photogrammetric Engineering & Remote Sensing* 70: 589-604.

Popescu, S.C. 2002. Estimating plot-level tree heights with LiDAR: local filtering with a canopy-height based variable window size. *Computers and Electronics in Agriculture* 37: 71-95.

Pringle, R.M., Webbb, H.K., Shine, R. 2003. Canopy structure, Microclimate, and Hatitat selection by a nocturnal snake, hoplocephalus bungaroides. *The Ecological Society of America* 84: 2668-2679.

Sexton, A.J.O., Bax, T., Siqueira, P., Swenson, J.J., Hensley, S. 2009. A comparison of LiDAR, radar, and field measurements of canopy height in pine and hardwood forests of southeastern North America. *Forest Ecology and Management* 257: 1136–1147.

van den Broek, R., Faaij, A., van Wijk, A. 1996. Biomass combustion for power generation, *Biomass and Bioenergy* 11: 271-281.

Verma, A.K., Kumar, A. 2011. Design of low-pass filters using some defected ground structures. *International Journal or Electronidx and communications* 65: 864-872.

Wulder, M., Nienann, K.O., Goodenough, D.G. 2000. Local maximum filtering for the extraction of tree locations and basal area from hight spatial resolution imagery. *Remote Sensing Environment* 73: 103-114.