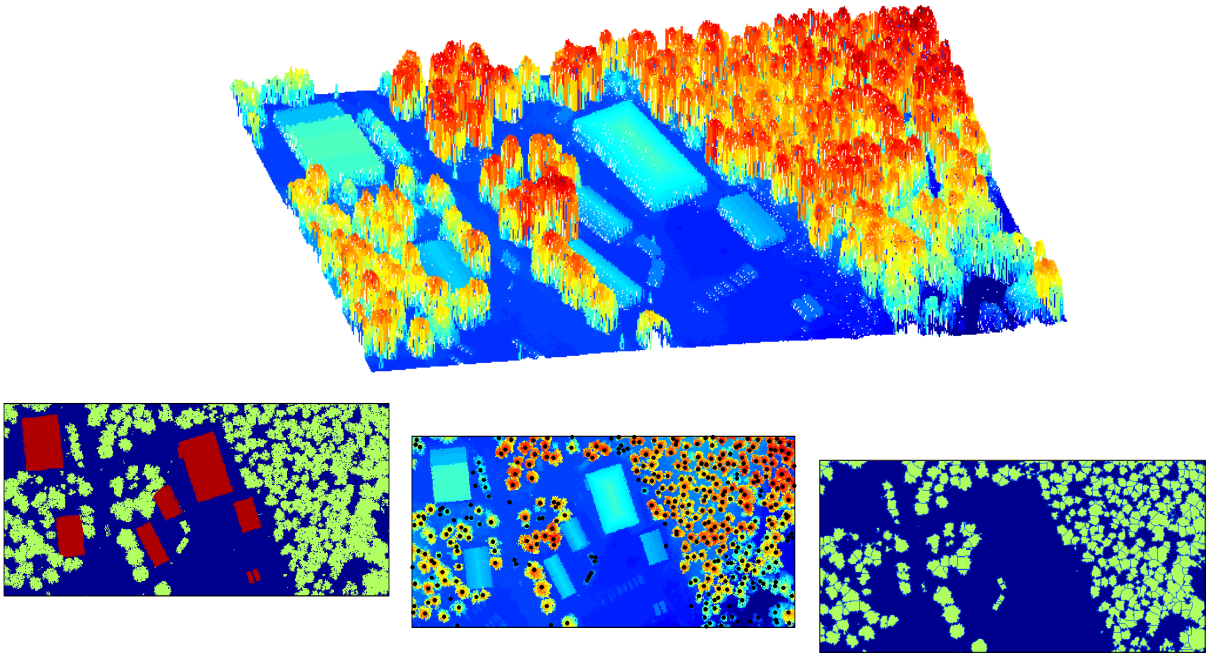
FOI
SWEDISH DEFENCE
RESEARCH AGENCY

FOI-R--0236--SE
October 2001
ISSN 1650-1942
Scientific report

Åsa Persson

Extraction of Individual Trees Using Laser Radar Data



Sensor Technology
SE-581 11 Linköping

SWEDISH DEFENCE RESEARCH AGENCY

Sensor Technology
P.O. Box 1165
SE-581 11 Linköping

FOI-R--0236--SE

October 2001

ISSN 1650-1942

Scientific report

Åsa Persson

Extraction of Individual Trees Using Laser Radar Data

Issuing organization	Report number, ISRN	Report type
FOI – Swedish Defence Research Agency Sensor Technology P.O. Box 1165 SE-581 11 Linköping	FOI-R--0236--SE	Scientific report
	Research area code 4. C4ISR	
	Month year October 2001	**Project no.** E3901
	Customers code -	
	Sub area code 42 Surveillance Sensors	

Author/s (editor/s)	Project manager
Åsa Persson	Lars Ulander
	Approved by
	Sponsoring agency
	Scientifically and technically responsible Åsa Persson

Report title

Extraction of Individual Trees Using Laser Radar Data

Abstract (not more than 200 words)

When generating 3D-landscape models using laser radar data, it is desirable to automatically extract different objects, such as buildings, trees, roads, etc, from the data set. In this study, an algorithm was implemented to extract individual trees using laser radar data. The segmentation was done in two steps. First, a segmentation based on texture measures and recording of double echoes was performed to separate man-made objects (buildings) and natural objects (vegetation). Single trees were then segmented in the identified areas of vegetation and the position, height, and crown diameter of the detected trees were estimated. The automatic recognition of the trees was compared with ground measurements and the result showed that 71 % of the trees were correctly identified. The average positional error of the detected trees was 51.4 cm. The height and the crown diameter of the detected trees could be estimated with a standard error of 0.63m and 0.61m, respectively.

Keywords

Tree extraction, laser radar, signal processing, segmentation, vegetation segmentation, feature extraction

Further bibliographic information	Language English
ISSN 1650-1942	**Pages** 38 p.
	Price acc. to pricelist
	Security classification

Utgivare Totalförsvarets Forskningsinstitut - FOI Sensorteknik Box 1165 581 11 Linköping	**Rapportnummer, ISRN** FOI-R--0236--SE	**Klassificering** Vetenskaplig rapport
	Forskningsområde 4. Spaning och ledning	
	Månad, år Oktober 2001	**Projektnummer** E3901
	Verksamhetsgren -	
	Delområde 42 Spaningssensorer	
Författare/redaktör Åsa Persson	**Projektledare** Lars Ulander	
	Godkänd av	
	Uppdragsgivare/kundbeteckning	
	Tekniskt och/eller vetenskapligt ansvarig Åsa Persson	

Rapportens titel (i översättning)

Extrahering av individuella träd med användning av laserradardata

Sammanfattning (högst 200 ord)

Vid framställning av 3D-modeller av landskap från laserradardata är det önskvärt att automatiskt kunna identifiera olika objekt som till exempel byggnader, träd och vägar från mätdata. I detta arbete implementerades en algoritm för att identifiera enskilda träd från laserradardata. Segmenteringen gjordes i två steg. Först gjordes en segmentering baserad på texturmått och dubbelekon för att separera konstgjorda (byggnader) och naturliga (vegetation) objekt. Enskilda träd segmenterades sedan i de identifierade områden av vegetation där positionen, höjden och krondiametern av träden estimerades. De automatiskt detekterade träden jämfördes med markmätningar och resultatet visade att 71 % av träden var korrekt identifierade. Medelpositionsfelet av de detekterade träden var 51,4 cm. Höjden och krondiametern av de detekterade träden kunde estimeras med ett standardfel på 0,63m respektive 0,61m.

Nyckelord

Trädextrahering, laserradar, signalbehandling, segmentering, vegetationssegmentering, egenskapsextrahering

Övriga bibliografiska uppgifter	**Språk** Engelska
ISSN 1650-1942	**Antal sidor:** 38 s.
Distribution enligt missiv	**Pris: Enligt prislista** **Sekretess**

Table of Contents

1 Introduction

At the Department of Laser Systems, 3D-models of landscapes have been developed using laser radar data. The method of using elevation data from airborne laser scanners has been shown to be very suitable for the generation of 3D-models. 3D-models of landscapes, cities, forests, etc, are used in many applications and the need for more complex models is increasing. While the requirements of accuracy and details in the models are increasing, the time to construct the models needs to be reduced.

Today, the generation of complex 3D-models is time-consuming with many manual adjustments. Therefore, it is of interest to study how different types of data processing methods could ease the work. Ideally, it is desirable to automatically generate the 3D-models using laser radar data. Different types of objects, such as buildings, trees, roads, etc, can automatically be extracted and modeled by analyzing the data. In a current project, a new landscape model is under development and the work presented here is a part of this project.

The purpose of the study was to automatically extract individual trees using laser scanner data. The position, height and crown diameter of the detected trees were estimated. A segmentation based on texture measures of local variations in height and recording of double echoes was first performed to separate man-made objects (buildings) and natural objects (vegetation). Single trees were then segmented in the identified areas of vegetation. When knowing the location and size of the trees, models of trees can directly be placed out in the 3D-landscape model. To evaluate the results, the algorithm was used over an area where ground measurements had been performed and the estimated position, height, and crown diameter of the detected trees were compared with the manual field measurements.

2 The Data of Laser Scanners

Laser scanners are used to measure the elevation of the terrain surface and elevation data from large areas can be collected in only a few hours. Laser scanners, which are operated from an airplane or helicopter, scan the ground in a zigzag pattern across the direction in which the plane is flying and measure the surface of the Earth in three dimensions (see Fig 2.1).

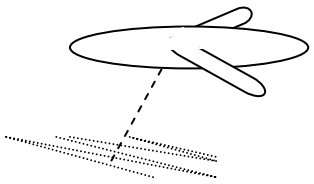
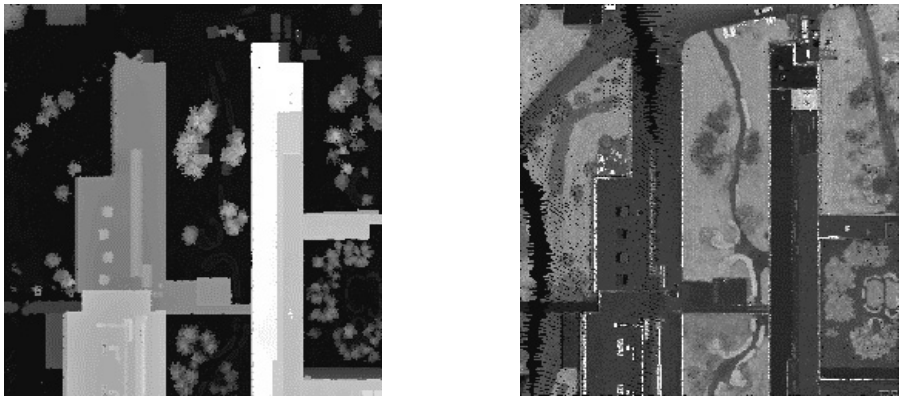


Fig 2.1. Laser scanners scan the ground in a zigzag pattern.

The range to the surface is obtained by measuring the time for the laser pulse to return. The exact geographic position of the surface points (latitude, longitude, and elevation) can be obtained with decimeter accuracy if the position of the airplane and the direction where the scanner is pointing are known. The position of the airplane is determined using

a GPS-receiver (Global Positioning System) and the direction where the scanner is pointing is obtained by an INS (Inertial Navigator System) onboard the airplane [4]. As a result, for every surface point that was hit by a laser pulse, the x-, y-, and z-coordinate can be calculated where x and y are the position and z the elevation. The laser scanner used in this thesis was the TopEye System. The data set was acquired over a 1000x2000 m area in Linköping. The point density was approximately 5 points/meter (yielding a density of about 25 points/m^2), and the accuracy of the surface points was one decimeter in the x-, y-, and z-direction. In some areas a lower density of approximately 15 points/m^2 was used. Fig 2.2a shows a part of the elevation data set.

In addition to the elevation value, the reflectance of the returning pulse is recorded (see Fig 2.2b). Since the reflectance differs for many materials, these values have been shown to be useful when segmenting the data. Most artificial materials like buildings and roads have relatively low reflectance while vegetation in general has higher values [3].



(a) (b)

Fig 2.2. Data from the TopEye System. (a) Elevation data. (b) Reflectance data.

Another useful source of information from the laser scanners is that more than one echo can be recorded from the same pulse, so called double echoes. Due to the size of the area that is covered when a beam reaches the surface, double echoes occur when two different ranges have been detected in the returning pulse. In Fig 2.3, the locations of the points giving double-echoes are shown. As shown in Fig 2.3, double echoes occur mainly in vegetation and at edges of buildings where discontinuities exist.

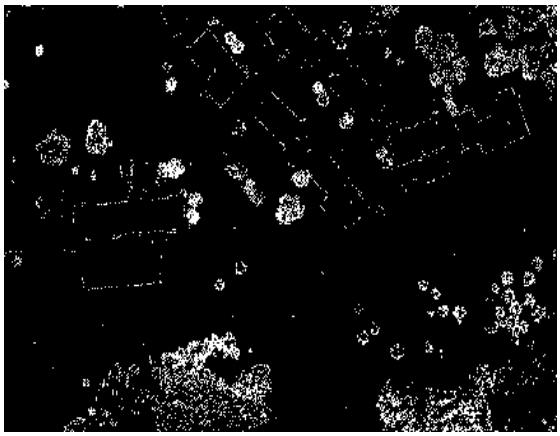


Fig 2.3. Positions where double echoes have occurred.

2.1 Pre-processing of Data

Since the beam of laser altimeters scans the ground in a zigzag pattern and the pulses are recorded in a temporal order, two subsequent pulses are not necessarily positioned next to each other depending on the height of the surface. Furthermore, the raw data points are usually not evenly spaced but have an irregular grid (see Fig 2.4a). Before processing, the data points were sorted into a rectangular array of cells and a model of the surface elevation is created which describes the shape of the surface where elevation is a function of position (see Fig 2.4b). This type of model is referred to as a digital surface model (DSM). A grid size of 0.33 meter was used. When sorting the data into a grid, some cells may contain more than one value.

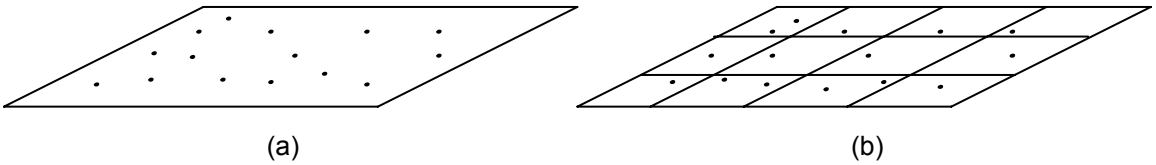


Fig 2.4. (a) Raw data. (b) Horizontal 0.33 m grid.

Furthermore, some pulses with different elevation may have the same x- and y-coordinate. For example, at edges of buildings, points with different elevation can have the same coordinates (see Fig 2.5a). Also, since the laser beam may penetrate the canopy of trees, some pulses may hit the top of the trees and some the ground below (see Fig 2.5b).

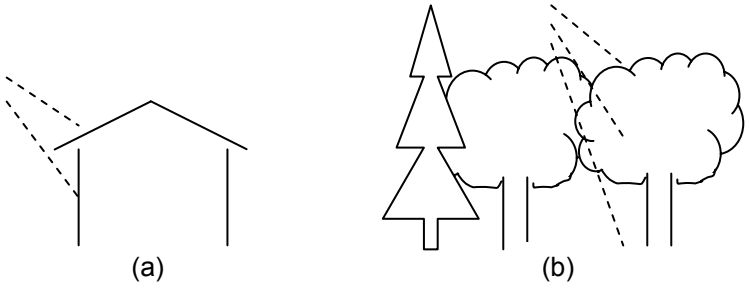


Fig 2.5. (a) Pulses may have the same x- and y-coordinate. (b) Pulses may penetrate vegetation.

Two grids were saved, one containing the largest value in each cell and one grid containing the smallest value. In addition, some cells may not be covered by the laser radar (see black dots in Fig 2.2b). These cells were filled by averaging the height values of the 8 neighboring cells where values existed.

In comparison to aerial photographs that need illumination, laser scanners work both day and night and the data will not contain shadows that make object recognition more difficult [4]. The elevation values and the recordings of double echoes were used to separate vegetation and buildings.

3 Object Extraction

To segment the objects as artificial or natural, the objects need first to be extracted from the ground. Thus, a first step is to identify the ground. In the thesis by Elmqvist [1], a segmentation of the ground was performed where objects are removed from the terrain. Fig 3.1a shows the elevation data over an area of 100x100 m when the maximum value in each cell was used. In Fig 3.1b, the ground segmentation of the area is shown. This type of model is referred to as a digital terrain model (DTM). The ground segmentation was based on the theories of active contours where the contour can be seen as a net being pushed upward from underneath the surface and sticks to the ground points. By subtracting the elevation of the ground from the measured elevation data, all objects above the terrain are left on a zero-elevation plane and a normalized digital surface model is obtained (see Fig 3.1c).

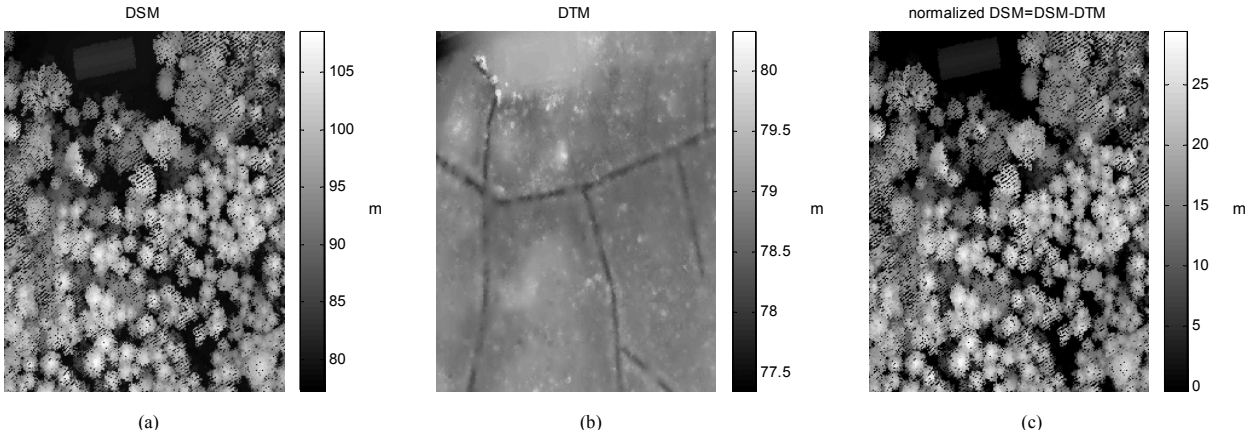


Fig 3.1. Extraction of objects above the terrain. (a) Elevation data. (b) Ground segmentation. (c) Ground elevation subtracted from surface elevation.

By thresholding the image in Fig 3.1c at 2 meters, all objects with a height of 2 meters above the ground level can be found (see Fig 3.2) [3].

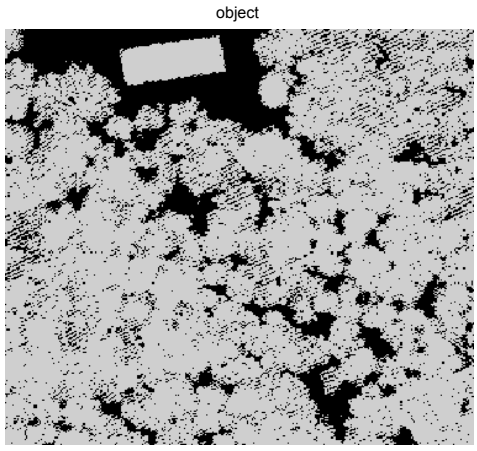


Fig 3.2. Object mask containing objects above 2 m.

Ideally, one would like to find the boundary of each individual object and classify each object separately. However, using the boundaries of the thresholded image will not perform well since trees are very close and connected to buildings in some regions. Only the pixels of the objects are known but not to which object the pixels belong. Thus, a pixelwise segmentation of the pixels above the threshold was performed.

4 Segmentation of Vegetation and Buildings

Segmentation of the data consists of dividing the data until the objects of interest have been found. First, the areas of vegetation needed to be identified; thus, the data set was separated into natural and artificial objects. Based on texture measures of local differences in height, artificial surfaces can be distinguished from the natural shape of natural objects [2]. While artificial objects such as buildings consist of continuous, compact surfaces that are bounded by discontinuous edges, natural objects such as vegetation have large vertical variations throughout the objects since the beam can penetrate the canopy of the trees and some pulses may hit the top of the trees and some pulses within the trees or the ground. In addition, the information acquired from double echoes was used for the separation.

4.1 Initial Segmentation Using Texture Measures

A pixelwise segmentation of the laser radar data was performed using measurements of local differences in elevation. First, a set of elementary properties that described the characteristics of the object (pixel) was chosen (see Fig 4.1). Several texture measures can be calculated directly from the laser scanner elevation data. These measurements are referred to as features. Finally, a classifier was used to recognize the objects from the chosen set of object descriptors.

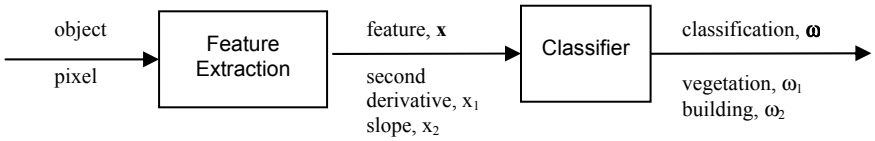


Fig 4.1. Segmentation process to distinguish vegetation and buildings.

4.1.1 Feature Extraction

To describe the characteristics of the object, measurements of local differences in height were used. Many different features can be extracted from the height data such as the slope, the second derivative, the Sobel operator, the variance in height in a window around a pixel, etc. However, most of these features show a large correlation with each other and the segmentation can therefore be based on a subset of these measurements [2]. The measurements used in this study were the second derivative and the slope. Adding a new feature such as the variation in height in a window around a pixel did not improve the segmentation.

Laplace operator
The Laplacian of a 2D function *f(x,y)* is a second derivative and is defined as

$$\nabla^2 f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{4.1}$$

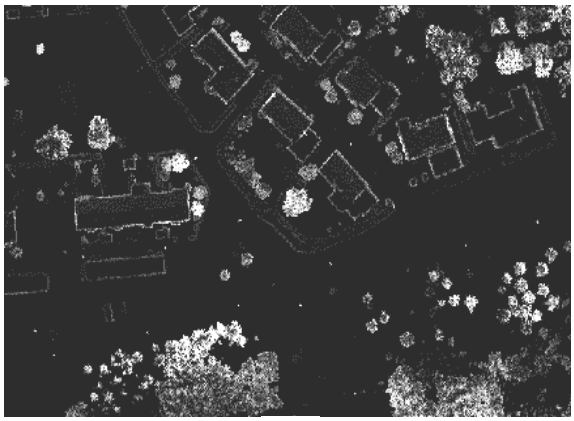
One way to implement the Laplacian operator in digital form for a 3x3 region is [5]

	-1	
-1	4	-1
	-1	

(4.2)

When the data was sorted into a grid, some cells contained more than one value. The maximum value was used for the value in the center of the mask and the minimum value for the neighbors to enhance the measures of the variation in height. In vegetation, where the height between neighboring pixels varies, the second derivative is larger than within buildings where the change in height of a flat or tiled roof is constant and the second derivative is close to zero (see Fig 4.2b). However, at the edges of buildings the Laplacian will be large.

(a)



(b)

Fig 4.2. (a) Elevation data over an area of 130x200 m. (b) Laplacian filter output for the elevation image.

Slope

The second feature used was the slope. The local slope in height was calculated for each pixel and its eight neighboring pixels and the maximum value was saved for each pixel (see Fig 4.3).

	h_{min1}	h_{min2}	h_{min3}
	h_{min4}	h_{max}	h_{min5}
dx	h_{min6}	h_{min7}	h_{min8}

dy

$$slope = \max\left(\left| h_{max} - h_{min\ i} \right| / distance \right) \quad i = 1,2,..,8$$

$$distance = \begin{cases} dx, dy & or \\ \sqrt{dx^2 + dy^2} \end{cases} \qquad (4.3)$$

Similarly to when calculating the second derivative, the maximum value in the cell was used for the center and the minimum value for the surrounding pixels. As shown in Fig 4.3, the slope over vegetation is in general larger than over flat or tiled roofs.

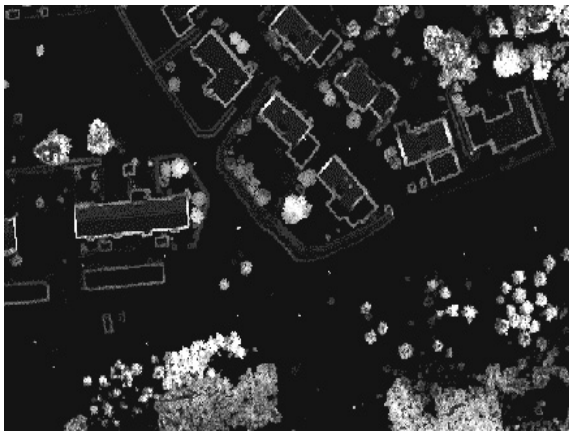
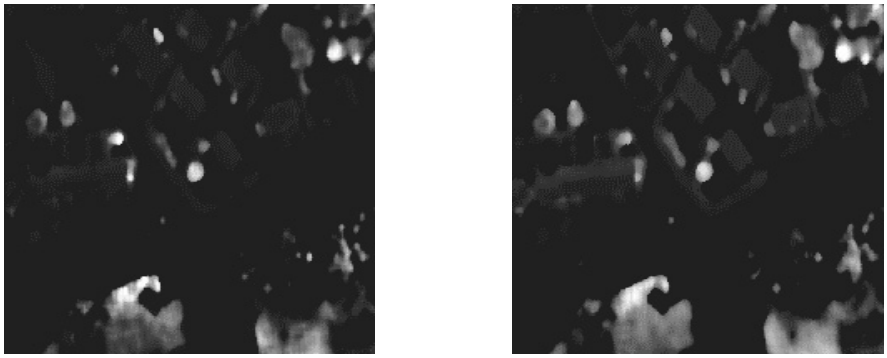


Fig 4.3. Maximum slope of the elevation image.

As Fig 4.2 and Fig 4.3 show, the second derivative and the slope will be large at edges of buildings and where antennas, chimneys, etc exist. To reduce this noise caused by the discontinuous parts of the objects on the roofs and the edges of buildings, the texture measures were median filtered (see Fig 4.4) [2]. Without the median filter, pixels within objects and edges of buildings will often be misclassified.



(a) (b)

Fig 4.4. Median filtering. (a) Second derivative. (b) Maximum slope.

The sizes of the masks were chosen so that in general the noise was completely removed. The median filtered texture measurements were the inputs to the classifier.

4.1.2 Classifier

When having obtained the features representing the characteristics of the objects, a classifier was used to recognize the objects. The feature vector is denoted

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

x_1 – second derivative
x_2 - slope

Classification is based on automatically assigning the feature vectors to the two pattern classes ω_1 and ω_2 where ω_1 and ω_2 represent vegetation and buildings, respectively. Thus, each pixel was classified as vegetation or building depending on its value of the feature vector. Fig 4.5a shows x_1 and x_2 from training regions of vegetation and buildings.

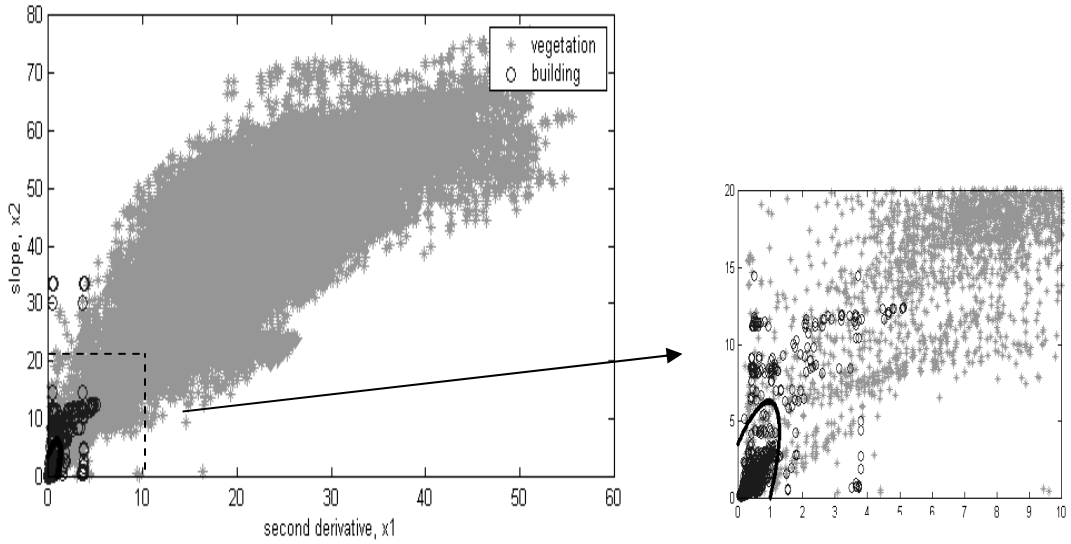


Fig 4.5. Feature vectors. (a) Median filtered slope, x_2 vs. median filtered second-order derivative, x_1. (b) Decision boundary.

Distance measures

Classification deals with the problem of finding decision functions $d_1(\mathbf{x})$ and $d_2(\mathbf{x})$ for the two pattern classes ω_1 and ω_2. The decision functions should have the property that, if the pattern belongs to vegetation, $d_1(\mathbf{x}) > d_2(\mathbf{x})$, and similarly, $\mathbf{x}$ belongs to buildings if $d_2(\mathbf{x}) > d_1(\mathbf{x})$. Therefore, the unknown feature $\mathbf{x}$ belongs to the decision function yielding the largest value. The decision boundary separating ω_1 and ω_2 is given by the values of $\mathbf{x}$ when $d_1(\mathbf{x}) = d_2(\mathbf{x})$ or $d_1(\mathbf{x}) - d_2(\mathbf{x}) = 0$.

One way to find the decision functions is to use a minimum distance classifier where each class is represented by its mean vector and the feature vector is assigned to the class with the closest mean vector. This classifier performs well when the spread of each class around its mean is small compared to the distance between the means. However, to have large mean separation and small class spread occur seldom in practice. The

measurements will vary to some degree within the class. For example, the Laplacian and slope are in general small for pixels inside the edges of a building. However, at the edges, the slope and Laplacian may be as large as over vegetation. As Fig 4.5 shows, the feature space is non-separable (see lower left corner).

On the other hand, a statistical classifier is useful in the present case because of the randomness under which pattern classes normally occur. A classifier that is optimal in the sense that, on the average, it yields the lowest probability of committing classification errors is called the Bayes classifier. If a classifier decides that **x** belongs to ω_1 when it actually belongs to ω_2, a loss has occurred. With Bayes classifier, the total average loss with respect to all decisions will be minimized. The decision function has the form

$$d_i(\mathbf{x}) = p(\mathbf{x}|\omega_i)P(\omega_i) \qquad i = 1, 2 \tag{4.4}$$

where $P(\omega_i)$ is the a priori probability of occurrence of class ω_i and $p(\mathbf{x}/\omega_i)$ is the probability density function of the patterns from class ω_i. It was assumed that a pixel is equally likely to belong to vegetation as to buildings, thus $P(\omega_i) = 0.5$. Assuming that $p(\mathbf{x}/\omega_i)$ is a multivariate Gaussian probability density function, the following quadratic decision function can be derived

$$d_i(\mathbf{x}) = \ln P(\omega_i) - 0.5 \ln|\mathbf{C}_i| - 0.5\left[(\mathbf{x} - \mathbf{m}_i)^T \mathbf{C}_i^{-1}(\mathbf{x} - \mathbf{m}_i)\right] \tag{4.5}$$

where $\mathbf{m}_i$ is the mean vector and $\mathbf{C}_i$ the covariance matrix. With a maximum likelihood classifier, the samples from the training regions are used to estimate the parameters $\mathbf{m}_i$ and $\mathbf{C}_i$

$$\mathbf{m}_i = \frac{1}{N_i}\sum_{\mathbf{x}\in\omega_i}\mathbf{x} \qquad \mathbf{C}_i = \frac{1}{N_i}\sum_{\mathbf{x}\in\omega_i}\mathbf{x}\mathbf{x}^T - \mathbf{m}_i\mathbf{m}_i^T \qquad i=1,2 \tag{4.6-7}$$

where $\mathbf{m}_i$ is the sample mean vector, $\mathbf{C}_i$ the sample covariance matrix, and N_i the number of sample vectors. A second-order surface is placed between the pattern classes. If the pattern classes are truly Gaussian, no other surface would yield a smaller average loss in classification [5].

The mean vector and covariance matrix were calculated for the two pattern classes using the training regions of vegetation and buildings in Fig 4.5. Substituting $\mathbf{m}_i$ and $\mathbf{C}_i$ into Eq (4.5), the decision functions $d_1(\mathbf{x})$ and $d_2(\mathbf{x})$ were evaluated for every pixel and the feature vector was assigned to the class yielding the largest value. The decision boundary, where $d_1(\mathbf{x}) - d_2(\mathbf{x}) = 0$, that separate ω_1 and ω_2 is shown in Fig 4.5b. A second-order curve is placed between the two pattern classes.

Fig 4.6 shows the classification result. Since the second derivative and the slope were median filtered, the edges of the buildings are in general correctly classified. However, since the masks of the filters were large and edges of trees are not straight, edges of trees surrounded by ground are removed and hence misclassified as buildings. To remove misclassified pixels, the information acquired from double echoes was used to improve this initial segmentation using texture measures.

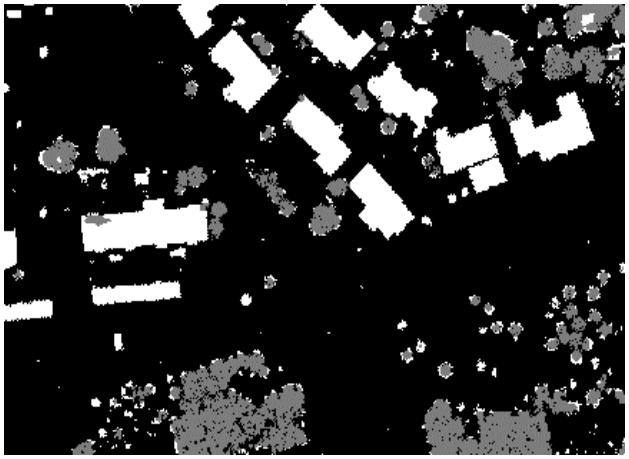


Fig 4.6. Initial classification results.

4.2 Improved Classification of Vegetation Using Double Echoes

To improve the above segmentation result, an additional segmentation of vegetation was performed using double echoes. As shown in Fig 4.7a, the recordings of double echoes, using the area of the upper right corner of Fig 4.2a, occur mainly in vegetation since trees have many vertical discontinuities. Double echoes also occur at the edges of buildings but not very often within the compact surfaces of roofs. By performing a dilation, double echoes will connect to others where they occur frequently, which are mainly in vegetation (see Fig 4.7b). At dense trees, no double echoes will occur at the center of the trees but only at the edge around the trees. As a result, after the dilation dense trees will often have an edge of double echoes surrounding the trees. The result after filling relatively small holes created by double echoes is shown in Fig 4.7c.

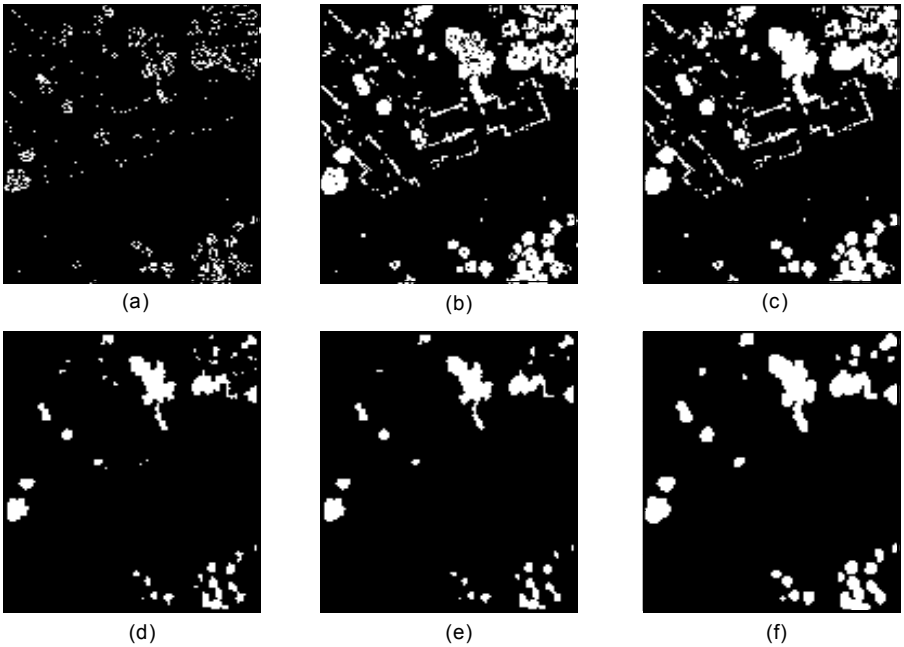


Fig 4.7. Classification of vegetation. (a) Double echoes. (b) Dilation. (c) Fill holes. (d) Erosion. (e) Median filtering. (f) Dilation.

Since double echoes also occur at edges of buildings, an erosion was performed to reduce this noise (see Fig 4.7d). As Fig 4.7d shows, some noise from edges of buildings might still be left. To remove remaining noise, a 5x5 median filter was used (see Fig 4.7e). Finally, a dilation was performed on the remaining areas (see Fig 4.7f) (see Appendix A for morphological operations: dilation, erosion, closing, opening). This result was combined with the previous segmentation. All pixels classified as vegetation were set to vegetation in the previous segmentation and some of the misclassified pixels of vegetation were removed (see Fig 4.8). By looking at the second derivative and the double echoes of the smaller uncertain areas classified as buildings, the result was further improved.

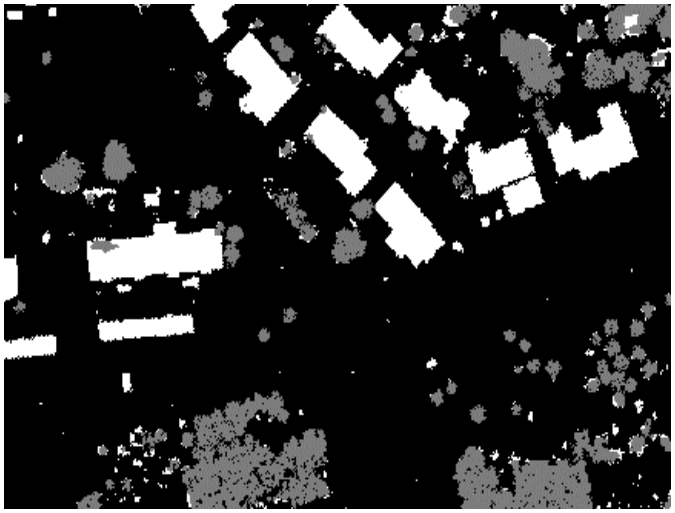


Fig 4.8. Improved classification results.

4.3 Final Classification

Since the texture measurements were median filtered, most large buildings are correctly classified. Instead some edges of trees have been misclassified. Thus, the problem is to determine if the smaller areas classified as buildings are small buildings or sheds or vegetation. Each of these areas was checked to see if the area was correctly classified as a smaller building or a part of a tree that had been misclassified. The texture measure of the second derivative without the median filter and the double echoes were used to distinguish buildings from vegetation of the smaller objects.

Since buildings consist of planar segments, the second derivative will be close to zero within the borders of the roofs of buildings. Only at the edges large values will occur. Therefore, the mean value of the second derivative using only the values inside the borders will be small for buildings compared to vegetation where the elevation will vary [3]. In addition, since double echoes occur mainly at the edges of buildings, and in general not within the compact surfaces of the roofs, the mean value of the number of double echoes inside the boundaries of the objects was also used. For the small areas classified as buildings, all connected pixels above 2 meters were used to represent the object (see Appendix A for algorithm for extraction of connected components). However, if small buildings are connected to trees, one of the classes will be misclassified.

In Fig 4.9a, the pixels classified as buildings are shown. To remove noise a 5x5 median filter was used (see Fig 4.9b). Each set of pixels smaller than 250 pixels was checked. For each of these sets, the corresponding object (all connected pixels above 2 m) and its boundary were found (see Fig 4.9c).

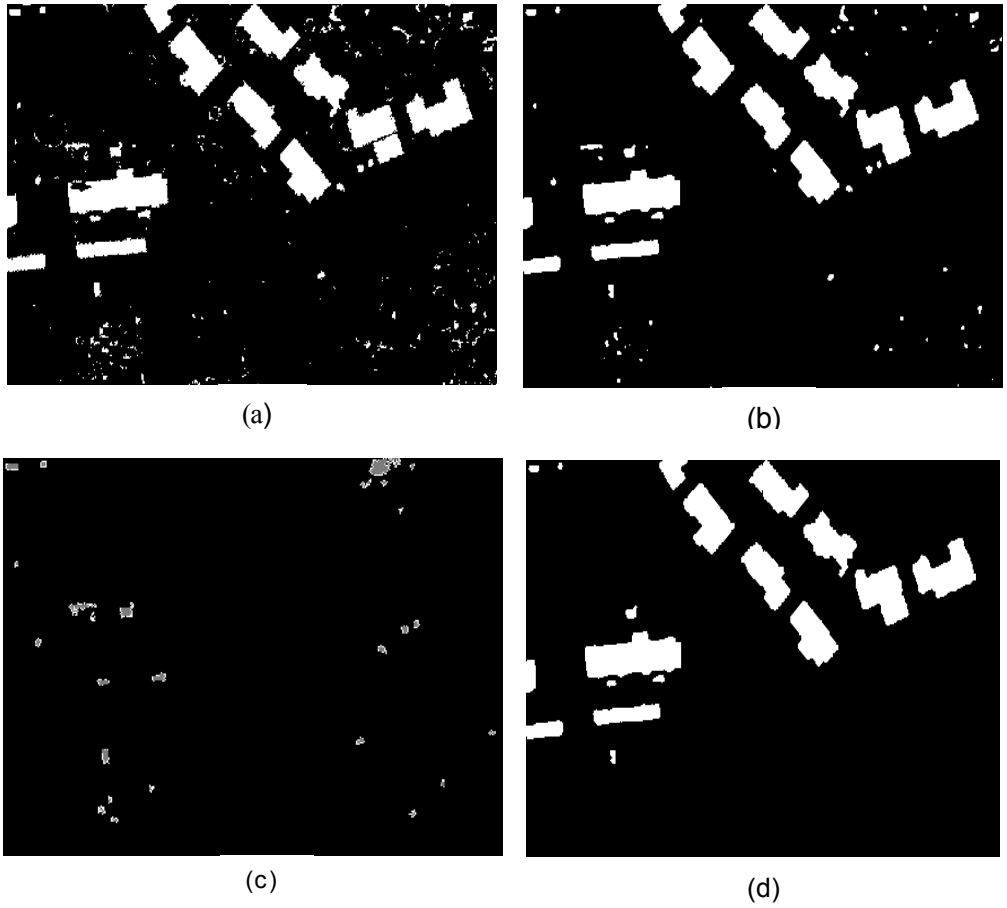
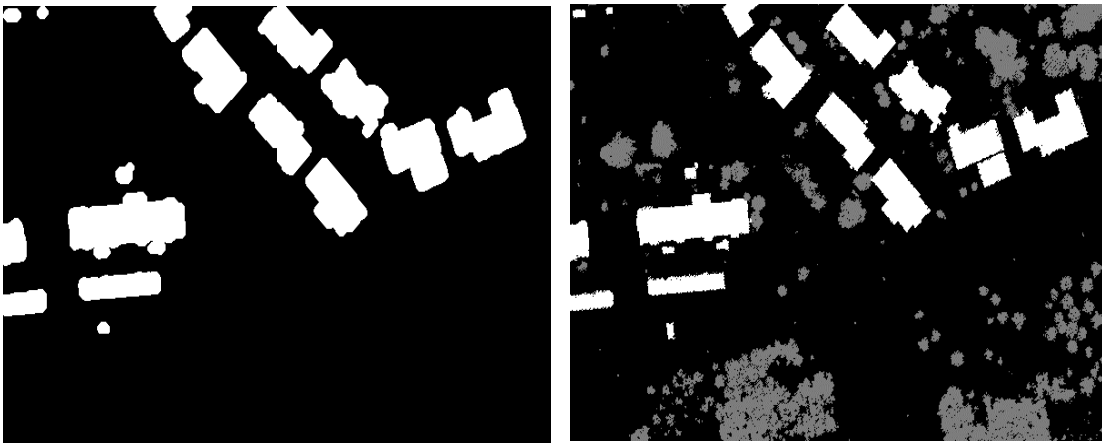


Fig 4.9. Final classification. (a) Pixels classified as buildings. (b) 5x5 median filter. (c) Objects. (d) Pixels in (a) remaining. The area is 130x200 m.

The mean value of the second derivative and the number of double echoes was calculated for the values inside the boundaries. Usually these values are smaller for buildings compared to vegetation; however, since only small areas of pixel are examined, the characteristics of the height measurements that usually make vegetation and buildings distinguishable are not as obvious. Some trees do not have a large variation in height but instead the characteristics of buildings. In addition, no double echoes might have occurred within the trees if the trees are dense.

The two mean values were thresholded, and if any of these values were above the threshold, the object was classified as vegetation. If both values were below the threshold, the object was taken as a correctly classified smaller building. In Fig 4.9d, the remaining pixels classified as buildings are shown after checking these two mean values of the smaller uncertain objects.

Finally, a closing followed by a dilation was performed to fuse narrow gaps of the pixels classified as buildings (see Fig 4.10a). Then, by letting the remaining pixels in the object mask be classified as vegetation and the pixels below 2 m as ground, the final segmentation result was obtained (see Fig 4.10b).



(a) (b)

Fig 4.10. Classification results. (a) Pixel classified as buildings. (b) Final results (white areas: buildings, gray areas: vegetation, black areas: ground).

As shown in Fig 4.10b, the misclassified pixels have been reduced and small buildings are detected. The result of the segmentation of the height data over an area of 900x400 m is shown in Table 4.1. In Appendix B.1, the segmentation of the area is shown.

	Vegetation	Buildings	Sum	Error of omission
Vegetation	294928	880	295809	0.30 %
Buildings	500	183663	184162	0.27 %
Sum	295428	184543		
Correctly classified pixels	99.8 %	99.5 %		

Table 4.1. Evaluation of the classification results.

Out of the 295428 pixels classified as vegetation, approximately 500 pixels belonged to buildings. Thus, 99.8 % of the pixels of vegetation were correctly classified. About 880 of the pixels classified as buildings belonged to vegetation. As a result, the error of omission of vegetation was 0.30 %. Out of the 184543 pixels classified as buildings, 99.5 % were correctly classified. The number of misclassified pixels was estimated by visual interpretation.

5 Locating and Analyzing Individual Trees

After the data set was separated into vegetation and buildings, a segmentation of single trees was performed in the identified areas of vegetation where the position, height, and crown diameter of the detected trees were estimated. The laser beam has the characteristic of being able to penetrate the canopy of trees. The ability to obtain ground hits even in dense areas of vegetation makes it possible to estimate the ground surface. Using the ground surface, the height of trees can be determined by subtracting the ground level from the measured height data. However, the laser beam's ability to penetrate the canopy of trees may result in large variations in height within single trees making it difficult to separate tree crowns from each other. Thus, one problem is to determine and remove the pulses that have penetrated the canopy and create a model of the outer part of the crowns. To remove the penetrations, the same active contour surface that was used to estimate the ground level was applied from above so that the surface followed the outer part of the crowns.

The process to detect individual trees was based on smoothing the image and the location of the trees was estimated by identifying local height maxima. To remove the height variations caused by branches within individual tree crowns so each tree has a single height maximum, a certain scale of smoothing should be used depending on the size of the trees. Thus, another problem when automatically extracting individual trees is the large variation in sizes of trees. Trees positioned close to each other with intersecting tree crowns may have no clear valley between the trees. On the other hand, an individual tree crown can have more than one peak. Three different scales were used to smooth the image. A parabolic surface was fitted to the elevation data to select the appropriate scale in different parts of the image. Finally, the height and crown diameter of the detected trees were estimated.

In previous studies, individual trees have been identified using airborne laser scanner data. Pyysalo detected individual trees and obtained a standard error of the tree height of 1.5 m [8]. Hyyppä also identified single trees and obtained a standard error of 1.8 m for the mean tree height [9]. In [10], tree height measurements from a high-resolution airborne laser scanning system are validated. A standard error of 0.97 m was obtained for the laser derived tree height.

5.1 Active Contour Modeling of the Canopy of Trees

Fig 5.1 shows the elevation data from one row in the image. As shown in the figure, many pulses have penetrated the trees. When lowpass filtering the vegetation, one would like to use the canopy of the trees and not include pixels where pulses have penetrated the foliage and hit the ground or within the trees.

To remove the pulses that had penetrated the vegetation, the active contour surface that was used to find the ground was applied from above (see Fig 5.1) [1]. The net sticks to the top of the trees and by adjusting the elasticity of the net an appropriate stiffness can be obtained so the net will not stretch down where the laser has penetrated the foliage.

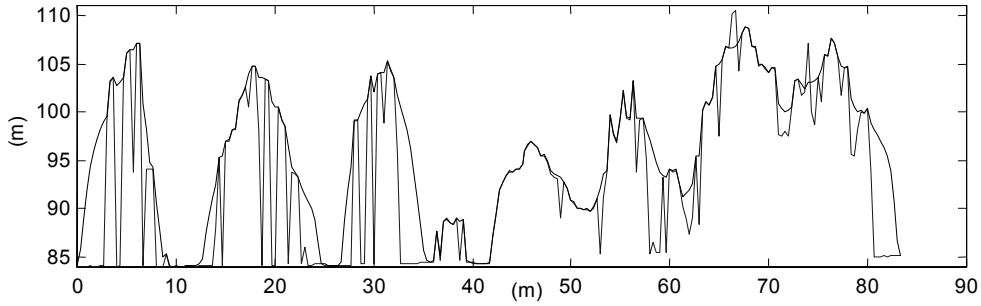


Fig 5.1. Elevation data and active contour surface applied from above.

To avoid that pixels that were close to a tree were assigned values of the net that were interpolated between the tree crown and the ground, the object mask was median filtered. The pixels remaining as ground after applying the median filter can be seen as the ground with no tree coverage and the net was set to zero at these pixels (see Fig 5.2a).

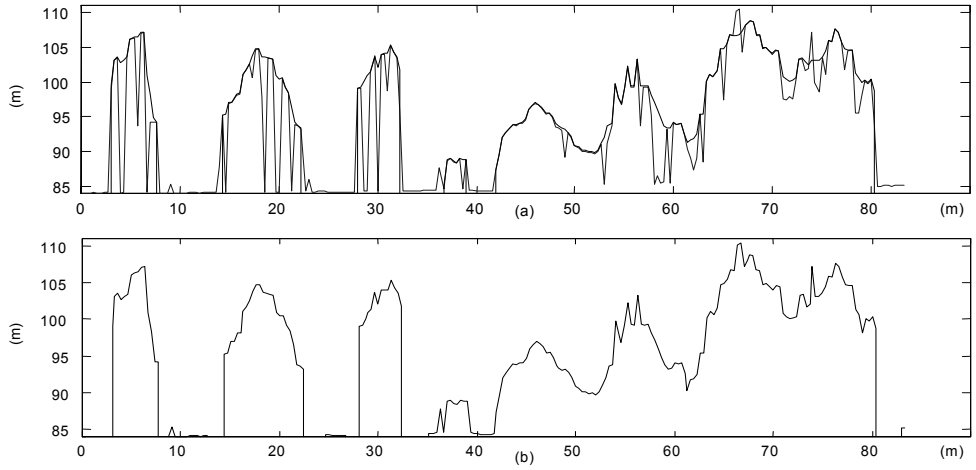


Fig 5.2. (a) Net set to zero at the ground. (b) Penetrations removed.

The pixel values where the original laser height was more than 2 meters below the net were replaced by the value of the net (see Fig 5.2b). Thus, pulses that hit below the canopy were eliminated and resulted in fewer variations in height within single trees.

5.2 Smoothing of the Canopy of Trees

To remove the variations in height within individual tree crowns, the image was smoothed. Image smoothing is based on averaging the height values in some neighborhood. A 2D Gaussian filter was used

$$G(x,y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{5.1}$$

The only parameter of the Gaussian filter is the standard deviation, σ. Pixels farther away than σ will have small influence and pixels with a distance of more than 3σ away will have negligible influence [6]. Depending on the size of the trees, different scales of

smoothing should be used. Three different σ-settings were used: 4/π, 6/π, and 8/π. The finest and the coarsest scale are shown in Fig 5.3b-c below.

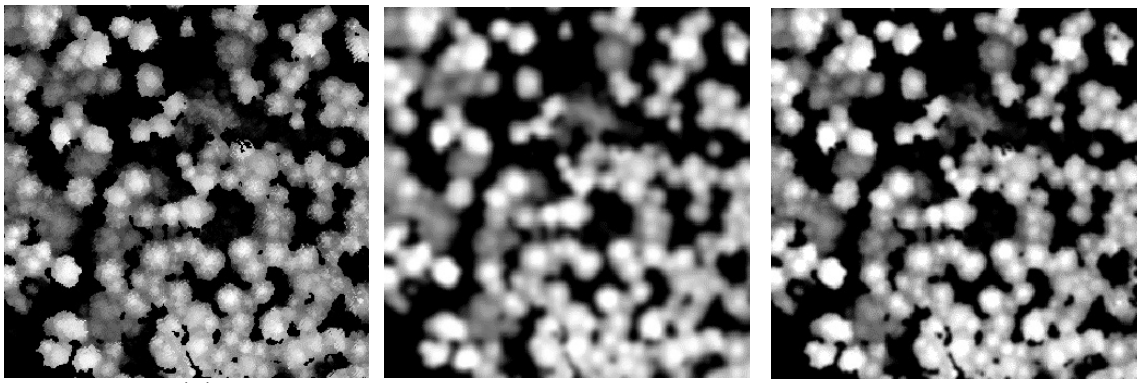


Fig 5.3. Gaussian filtering. (a) Elevation data with penetrations removed. (b) σ = 8/π. (c) σ = 4/π.

The location of the trees was estimated by searching for local height maxima in the smoothed images. Seeds were placed out in every pixel classified as vegetation and let to climb in the direction having the largest slope. When a seed reached a position where all neighboring pixels had lower values, a local maximum was found. Fig 5.4 shows the locations of the height maxima using the coarsest and the finest scale respectively. The crown coverage was estimated by grouping those pixels that climbed to the same maximum (see Fig 5.4).

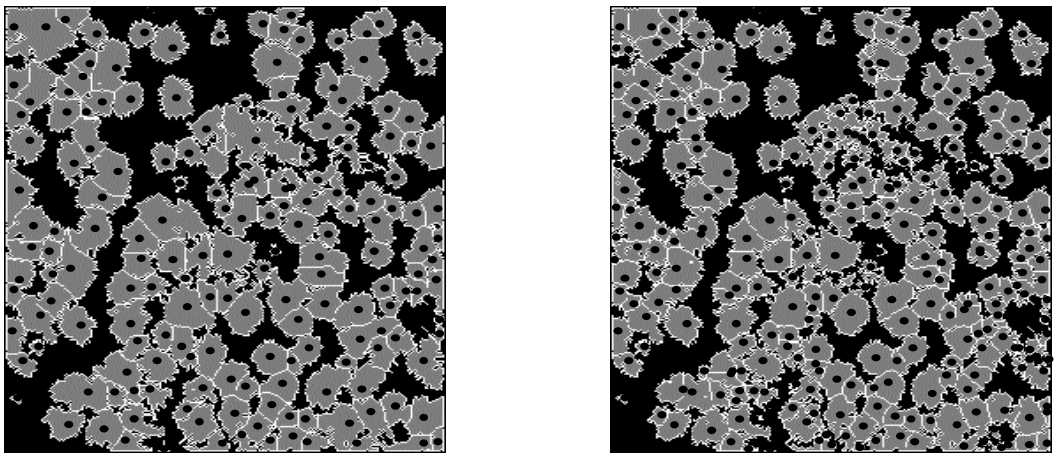


Fig 5.4. Local height maxima and crown coverage. (a) Coarsest scale. (b) Finest scale.

The smoothing of the coarsest scale was chosen so that in general no tree had more than one maximum. The finest scale was chosen so that small trees in compact groups were detected. Fig 5.5 shows the estimated positions of the trees marked on the original height values.

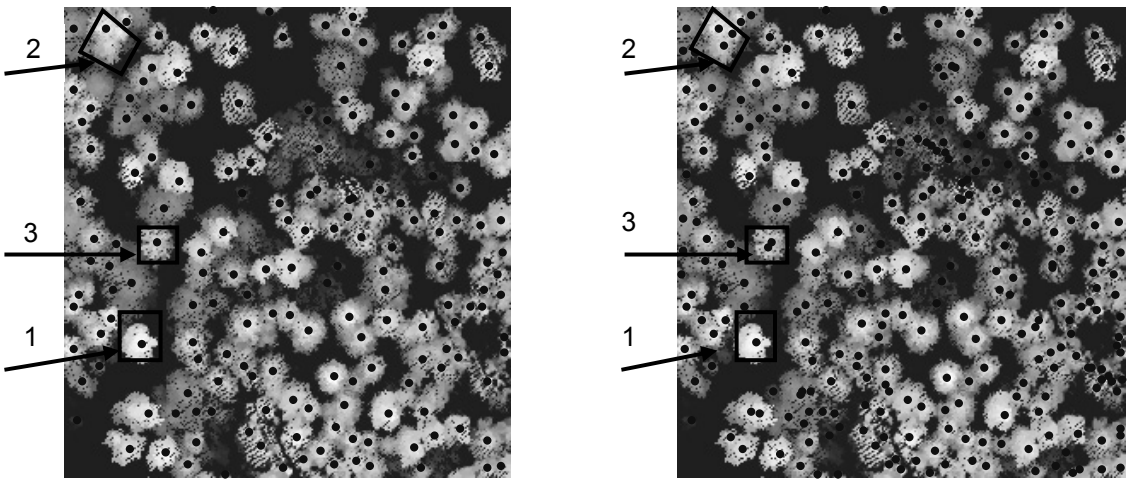


Fig 5.5. Estimated locations of trees. (a) Coarsest scale. (b) Finest scale.

At the coarser scale, different tree crowns may have been merged, and on the contrary, at the finer scale, variations within a single tree may be detected as several trees. Thus, a combination of the scales is desired.

5.3 Combination of Scales

To determine which scale to choose for different parts of the image, a parabolic surface was fitted to the top 30% of the segments of the elevation data in Fig 5.3a. Three different cases could occur when comparing the segmented areas of trees from the coarser scale with the corresponding area at the finer scale (see Fig 5.6).

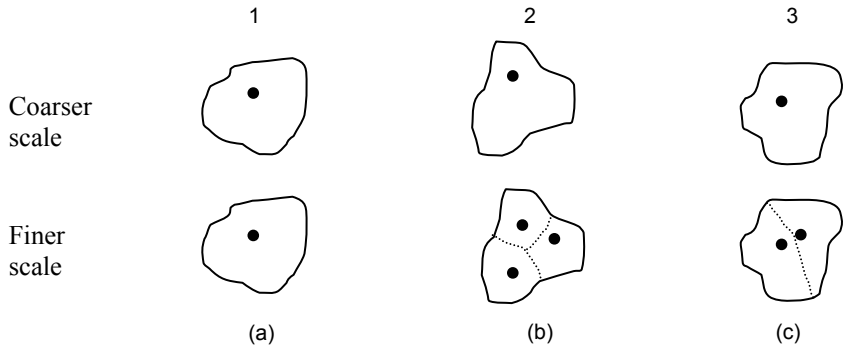


Fig 5.6. Possibilities. (a) One tree. (b) Trees merged. (c) Variation within a tree.

One possibility is that the finer scale could also have only one maximum within the crown coverage of the coarser scale (see Fig 5.6a). In this case, the top was taken as correctly determined. The two other possibilities are when the finer scale has detected more than one maximum. In case two, the coarser scale has merged tops of trees. The third case is when a variation within a tree has been detected.

For case two and three, the problem was to determine if the additional maxima at the finer scale should be judged as separate trees or belong to the treetop detected at the coarser scale. To determine which scale to use a second-order parabolic surface

$$z = a(x - x_0)^2 + b(y - y_0)^2 + c \tag{5.2}$$

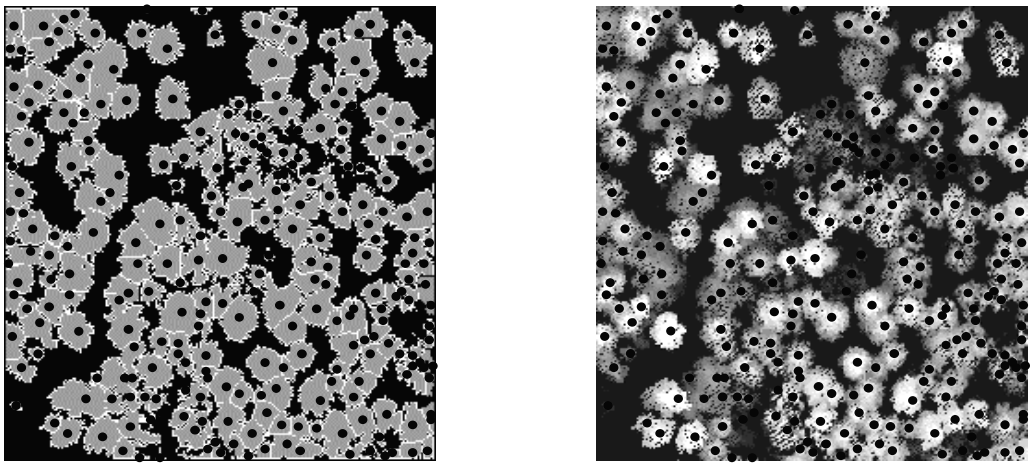
was used and the center of the surface, x_0 and y_0, was placed at the maximum found at the coarser scale. The three unknown parameters a, b and c were chosen so the square error was minimized

$$min\ \delta = \sum_i (z_i - z)^2 \tag{5.3}$$

Taking the derivative of Eq (5.3) with respect to a, b and c and setting the equations to zero resulted in three equations and three unknowns. The equations were solved for a, b and c.

Each additional maximum at the finer scale was tested starting with the one closest to (x_0, y_0). First, the surface was fitted to the segment at the finer scale that belonged to the maximum (x_0, y_0) and the segment of the additional treetop being tested at the finer scale. Then, the surface was fitted with the center placed at the same position but now using only the segment from the finer scale that belonged to the maximum that was detected at the coarser scale. If the sum of residuals decreased by more than 8% at the pixels where the parabolic surfaces overlapped, the treetop only found at the finer scale was judged as a separate treetop. This procedure was first performed with the coarsest and the middle scale. The resulting segments were then compared with the finest scale.

Fig 5.7 shows the result when a combination of the scales was used. Now, the merged trees in example 2 in Fig 5.5 have been replaced by the trees detected at the finest scale, and the coarsest scale has been chosen in example 3 so the variation within the tree is not detected. The result is similar to the visual interpretation. In Appendix B.3 and B.5, the estimated positions of trees over some regions are shown.



(a) (b)

Fig 5.7. Combination of the scales. (a) Crown coverage. (b) Estimated positions of trees.

5.4 Results

To evaluate the results, the algorithm was used over an area where ground measurements of forest parameters had been performed. In September 2000, the Swedish University of Agricultural Sciences (SLU) in Umeå performed field measurements at a test site located in southern Sweden (lat. 58° 30' N, long. 13° 40' E). The main tree species were Norway spruce (*Picea abies L. Karst.)*, Scots pine (*Pinus sylvestris L.)* and birch (*Betula spp.)*. The area was essentially flat with a variation in altitude ranging from 120–145 meters above the sea level. The laser data acquisition was performed using the TopEye system. The flight altitude was approximately 130 m above the ground, the speed 16 m/s, the scan frequency 16.67 Hz, the scan width ±10°, and the beam divergence 1 mrad. The point density was approximately 5 points/m^2.

Along the flight lines, twelve rectangular plots (50x20m) were selected for the field measurements. The main tree species were spruce for six of the plots and pine for six plots. The stem position and stem diameter were measured for all trees within the plots and for a random sample of trees (approximately 15 trees/plot), the height and crown diameter were measured.

5.4.1 Number of Detected Trees

To validate how single trees can be detected, each detected tree was linked to the corresponding field-measured tree. For each segment, three different cases could occur: *i*) no field tree was within the segment, *ii*) one field tree was within the segment, and *iii*) more than one field tree were within the segment. For case (*i*), the segment was judged as a segment that had no field tree. For case (*ii*), the field tree was linked to the laser-detected tree. For case (*iii*), the field tree that was closest to the position of the laser-detected tree was linked to the tree. When the laser trees and the field trees had been linked with the rules above, each field tree that had not been linked was examined. For each of these trees, a search was done at a maximum distance of 2 pixels in all directions. If a segment was found that had not been linked and it was within the plot, the field tree was linked to this segment.

Fig 5.8 shows the elevation data for four of the plots and the automatically detected trees (black marks). The white marks show the positions of the measured field trees where a white dot indicates that the tree has been detected and linked, and a white x-mark that the tree has not been identified.

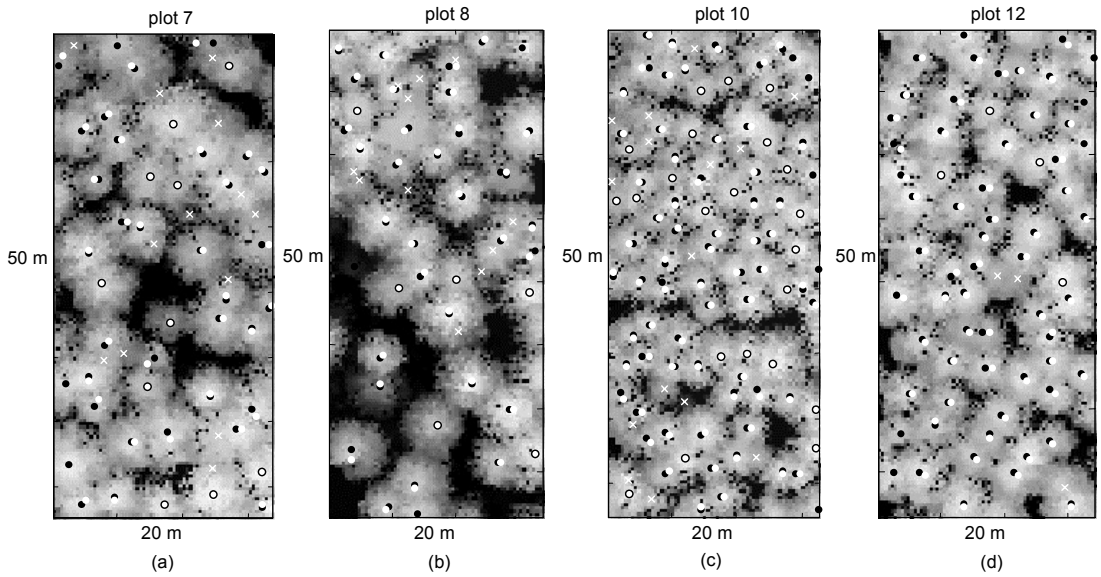


Fig 5.8. Estimated positions of trees and field measurements. (a) 42/55 trees detected. (b) 37/48 trees detected. (c) 86/101 trees detected. (d) 59/62 trees detected. ● = automatically estimated tree positions, ○ = field measured stem positions, 𝕏 = non-detected trees, and ⚬ = black and white mark at the same pixel.

All trees within the plots having a stem diameter ≥5.0 cm were measured. Thus, small and hidden trees that cannot be seen from above are included. Out of the 795 trees, 562 (71%) were detected. Two false detections not corresponding to trees were found. Table 5.1 shows the number of identified trees for each of the twelve plots.

Plot	Number of detected trees	Total number of trees	Percent
1	27	28	96%
2	31	63	49%
3	27	31	87%
4	45	73	62%
5	43	64	67%
6	58	143	41%
7	42	55	76%
8	37	48	77%
9	52	61	85%
10	86	101	85%
11	55	66	83%
12	59	62	95%
Total	562	795	71%

Table 5.1. Number of detected trees.

The result showed that small trees standing close to large trees were in general not detected. Also, trees standing close to each other with intersecting tree crowns and no clear valley between the trees were not detected. For example, plot 6 consisted of young spruces where many trees had less than a meter in between the stems and were thus not visible from the height data.

When evaluating the number of identified trees for different tree sizes, the result showed that most of the large trees were detected. In Fig 5.9, the number of detected trees for different stem diameters is shown. A large portion of the undetected trees has a small stem diameter.

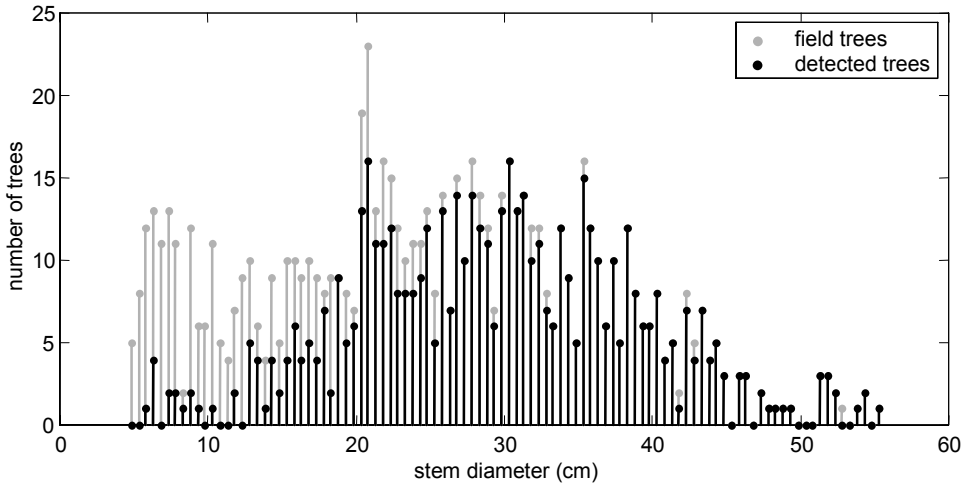


Fig 5.9. Number of detected trees for different stem diameters.

In Table 5.2, the number of detected trees for different stem diameters is shown. For example, 90 % of the trees with a stem diameter >20 cm were detected.

Plot	Detected trees with a stem diameter			
	≥ 5.0 cm	>10.0 cm	>15.0 cm	>20.0 cm
1	96%	96%	96%	96%
2	49%	50%	61%	69%
3	87%	96%	96%	96%
4	62%	93%	93%	93%
5	67%	89%	93%	93%
6	41%	54%	70%	89%
7	76%	78%	82%	89%
8	77%	77%	78%	85%
9	85%	85%	87%	86%
10	85%	86%	89%	98%
11	83%	83%	85%	87%
12	95%	95%	95%	97%
Total	562/795 (71%)	549/694 (79%)	531/621 (86%)	471/522 (90%)

Table 5.2. Number of detected trees for different stem diameters.

5.4.2 Estimating the Stem Positions

The estimated positions of the identified trees were compared with the measured stem positions. The center of the pixel was defined as the position of the detected trees. Fig 5.10 shows the positional errors relative to the field measurements for the detected trees. The average positional difference of the stem positions was 51.4 cm.

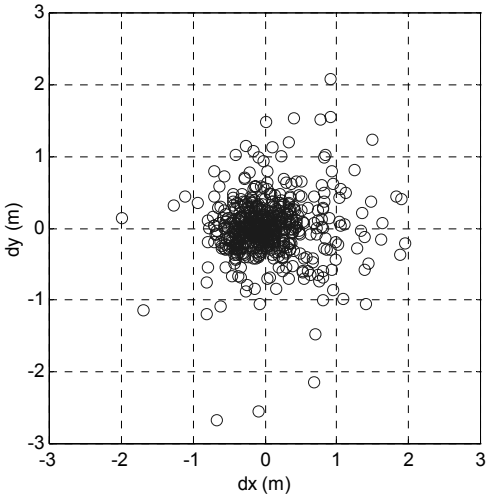


Fig 5.10. Positional errors of estimated stem positions relative to field measured stem positions.

5.4.3 Estimating the Height of Trees

In Fig 5.11, the field measured tree height is plotted against the laser estimated tree height. The height of the field trees was measured using an ultrasound distance measurement and an electronic angle decoder. For each segment in the laser data set, the maximum elevation value above the ground surface was used as the estimate of the tree height. The correlation coefficient showing the linear relationship of the heights was 0.99 and the standard error of the laser estimated tree height was 0.63 m.

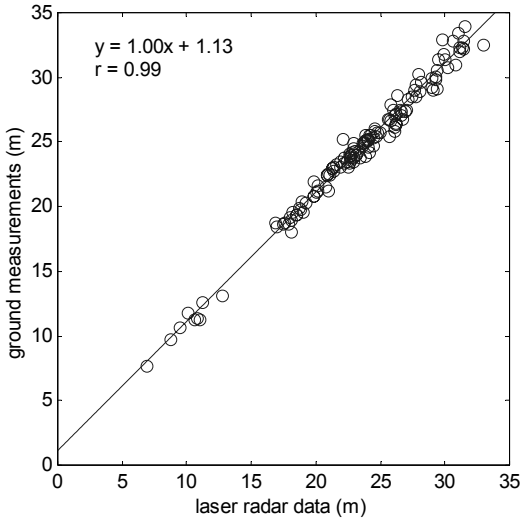


Fig 5.11. Estimated height of trees vs. ground measurements, 135 trees.

If a tree was not detected in the laser data, typically a small tree next to a large tree, this smaller field tree could sometimes be linked to the detected tree if positioned closer than the correct higher field tree. Therefore, measurements where the laser measured tree height was twice the height of the linked field measured tree (seven trees) were excluded from the data set.

The standard error for *Suunto* hypsometers varies from 0.4 to 0.8 m [11]. If assuming that the electronic hypsometer in this study had the same standard error, a significant portion of the mean error for the laser estimated tree height could have been caused by error in the field data.

5.4.4 Estimating the Crown Diameter of Trees

The estimated crown diameter of the trees was compared with the field-measured crown diameter (see Fig 5.12). For the field measurements, the projected crown diameter on the ground was measured using a plumb to determine where on the ground the outermost part of the crown was projected. The procedure was performed on both sides of the crown and the distance between the two points was measured. For the laser data, the area of the segments was used to calculate the diameter of the trees as if the tree crowns had the shape of a circle.

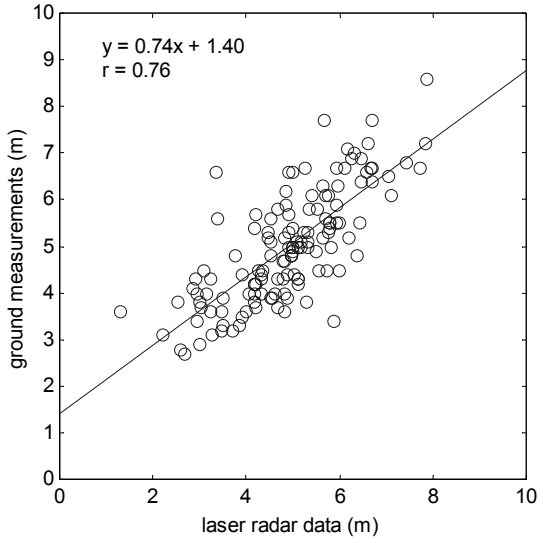


Fig 5.12. Estimated crown diameter of trees vs. ground measurements, 135 trees.

Since the segmented trees do not overlap, trees with intersecting tree crowns will get too small values. On the other hand, the area of non-detected trees will be added to surrounding trees resulting in a too large diameter for these trees. In addition, since trees are not exactly circular, the diameter of the field-measured trees will vary depending on where the measurement is made. A correlation coefficient of 0.76 was obtained and the standard error of the estimated crown diameter was 0.61 m.

6 Further Work

Laser radar data are usually combined with additional information such as aerial photographs in commercial systems. Although aerial images contain shadows, the images often have higher resolution; thus, aerial photographs can be integrated with the laser radar data to improve the result. A continuation of the segmentation will be to identify the trees as deciduous or coniferous trees. In [7], methods to classify tree species of individual trees in aerial images are presented. In addition, the reflectance data may be used to extract structures on the ground, such as roads, from surrounding grass using the ground pixels.

7 Conclusion

An automatic extraction of individual trees was performed using laser scanner data. First, the data set was segmented to separate man-made and natural objects using texture measurements and the recordings of double echoes. The segmentation result showed that 99 % of the pixels of vegetation and buildings were correctly classified.

In the identified areas of vegetation, a further analysis of single trees was performed where the position, height and crown diameter of the trees were estimated. By applying an active contour surface from above, the canopy of the trees was modeled and the pulses that had penetrated the vegetation could be removed. The process to detect individual trees was based on Gaussian smoothing and the location of the trees was estimated by identifying local height maxima. A combination of three different scales was used for the smoothing and a parabolic surface was fitted to the elevation data to determine which scale to choose for different parts of the image. The automatic recognition of trees was compared with manual ground measurements. The result showed that 71 % of the trees were correctly identified. Most of the large trees were detected and a large portion of the undetected trees had a small stem diameter. The average positional error of the detected trees was 0.51 m. The height of the segmented trees had a correlation of 0.99 with the reference data while the crown diameters of the trees had a correlation coefficient of 0.76. The standard error of the estimated tree height and crown diameter was 0.63 m and 0.61 m, respectively.

8 References

[1] Elmqvist M. *Automatic Ground Modelling using Laser Radar Data.* Master thesis LiTH-ISY-EX-3061, Department of Electrical Engineering, Linköping University, March 2000.

[2] Maas H-G. *Segmentation of Airborne Laserscanner Data*. Delft University of Technology, 1999.

[3] Hug C. *Extracting Artificial Surface Objects from Airborne Laser Scanner Data*. In Gruen A, E. P Baltsavias and O. N. Henricsson. *Automatic Extraction of Man-Made objects from Aerial Space Images (II).* Birkhauser, 1997.

[4] http://www.geolas.com/Pages/laser.html p1-5

[5] Gonzalez R and R Woods. *Digital Image Processing.* Addison–Wesley, 1993.

[6] Sonka M, V Hlavac and R Boyle. *Image Processing, Analysis, and Machine Vision, 2nd ed.* Brooks/Cole, 1999.

[7] Brandtberg T. *Automatic Individual Tree-Based Analysis of High Spatial Resolution Remotely Sensed Data*. Doctoral thesis, Acta Universitatis Agriculturae Sueciae, Silvestria 118. Swedish University of Agricultural Science, Uppsala, 1999.

[8] Pyysalo U. *A Method to Create a Three-dimensional Forest Model From Laser Scanner Data.* The Photogrammetric Journal of Finland. Vol. 17, No. 1, pp34-42, 2000.

[9] Hyppää J, O Kelle, M Lehikoinen and M Inkinen. *A Segmentation Based Method to Retrieve Stem Volume Estimates from 3-D Tree Height Models Produced by Laser Scanners.* IEEE Transactions on Geoscience and Remote Sensing, Vol. 39, No. 5, pp969-975, May 2001.

[10] Hyyppä J., U. Pyysalo, H. Hyyppä, and A. Samberg. *Elevation accuracy of laser scanning-derived digital terrain and target models in forest environment.* In Proceedings of the 4th EARSeL workshop on LIDAR Remote Sensing of Land and Sea, Dresden, Germany, 2000.

[11] Lindgren, O. *A study of circular plot sampling of Swedish forest compartments.* Section of Forest Ministration and Management, Swedish University of Agricultural Sciences, Report 11, 153 p, 1984.

9 Appendix A

Binary morphological operations
Mathematical morphology is based on set theory. The sets are members of the 2D integer space Z^2 in binary images [5].

Dilation
Let A and B be sets in Z^2. The dilation of A by B is defined as

$$A \oplus B = \left\{ x \left| \left(\hat{B} \right)_x \cap A \neq \phi \right. \right\} \tag{A.1}$$

where $\emptyset$ is the empty set, $\hat{B}$ the reflection of B, and $(B)_x$ the translation of B by x. The dilation of A by B consists of obtaining the reflection of B about its origin and then shifting $\hat{B}$ by x. All displacements of x, where $\hat{B}$ and A overlap by at least one nonzero element, constitute to the dilation (see Fig A.1).

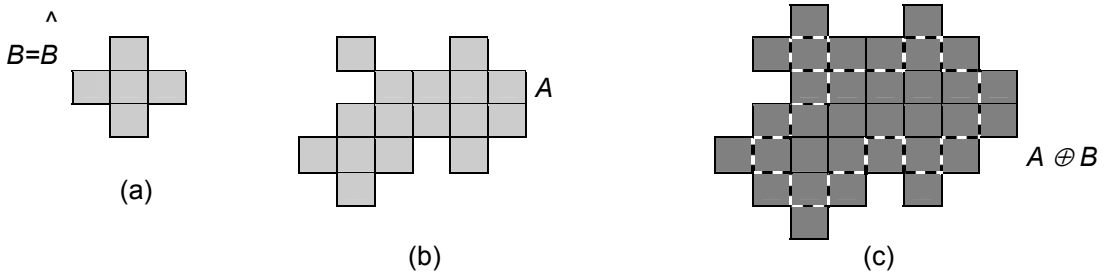


Fig A.1. Dilation. (a) Structuring element B and its reflection. (b) Set A. (c) Dilation of A by B.

Set B is referred to as the structuring element and can be viewed as a convolution mask where B is reflected about its origin and successively displaced so it slides over A. Dilation results in an expansion of the image.

Erosion
With A and B in Z^2, the erosion of A by B is defined as

$$A \ominus B = \left\{ x \left| \left(B \right)_x \subseteq A \right. \right\} \tag{A.2}$$

As shown in Fig A.2, the erosion of A by B is the set of all points where B, displaced by x, is contained in A. Erosion shrinks the image.

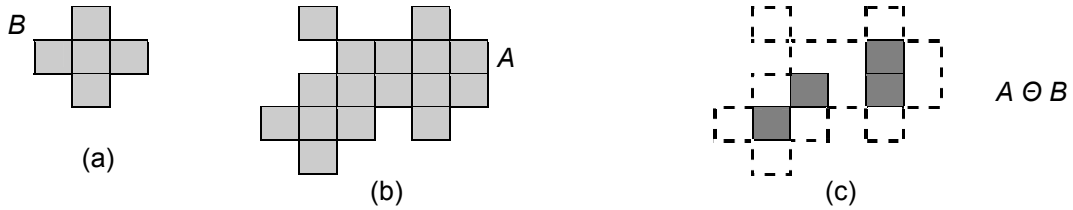


Fig A.2. Erosion. (a) Structuring element B. (b) Set A. (c) Erosion of A by B.

Opening
The opening of set A by B is denoted $A \circ B$ and is defined as

$$A \circ B = (A \ominus B) \oplus B \tag{A.3}$$

Thus, the opening of set A by B is the erosion of A by B followed by a dilation of the result by B. Opening tends to remove thin protrusions and small islands and smooth the contour of an image.

Closing
The closing of set A by B is denoted $A \bullet B$ and is defined as

$$A \bullet B = (A \oplus B) \ominus B \tag{A.4}$$

The closing of set A by B is the dilation of A by B. An erosion of the result by B is then performed. Closing generally eliminates small holes, fuses narrow breaks and fills gaps in the contour.

Extraction of connected components
If Y represents a component contained in a set A and a point p of Y is known, the following iterative expression extracts all the elements of Y

$$X_k = (X_{k-1} \oplus B) \cap A \qquad k = 1, 2, 3, \ldots \tag{A.5}$$

where $X_0 = p$ and B is a structuring element. When $X_k = X_{k-1}$, the algorithm has converged and $Y = X_k$ (see Fig A.3) [5].

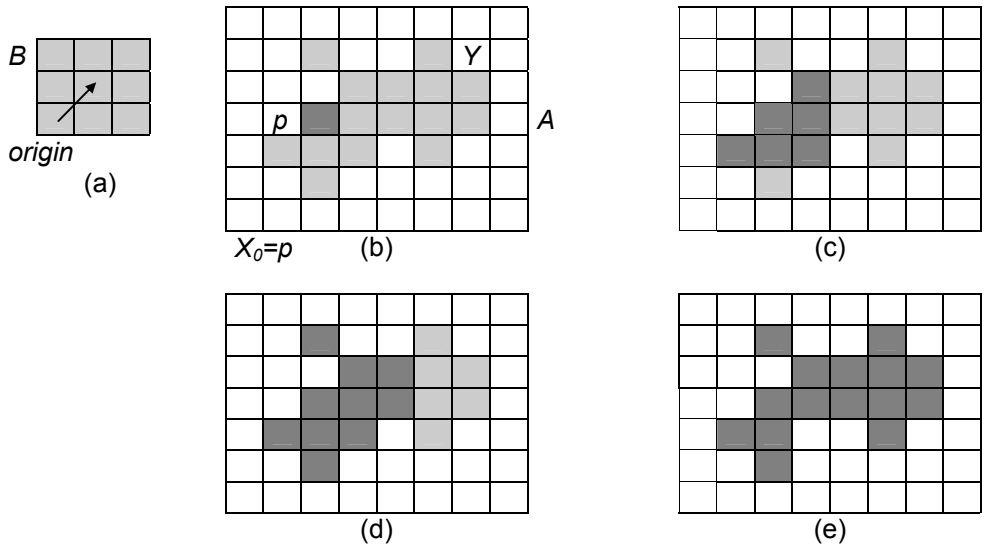
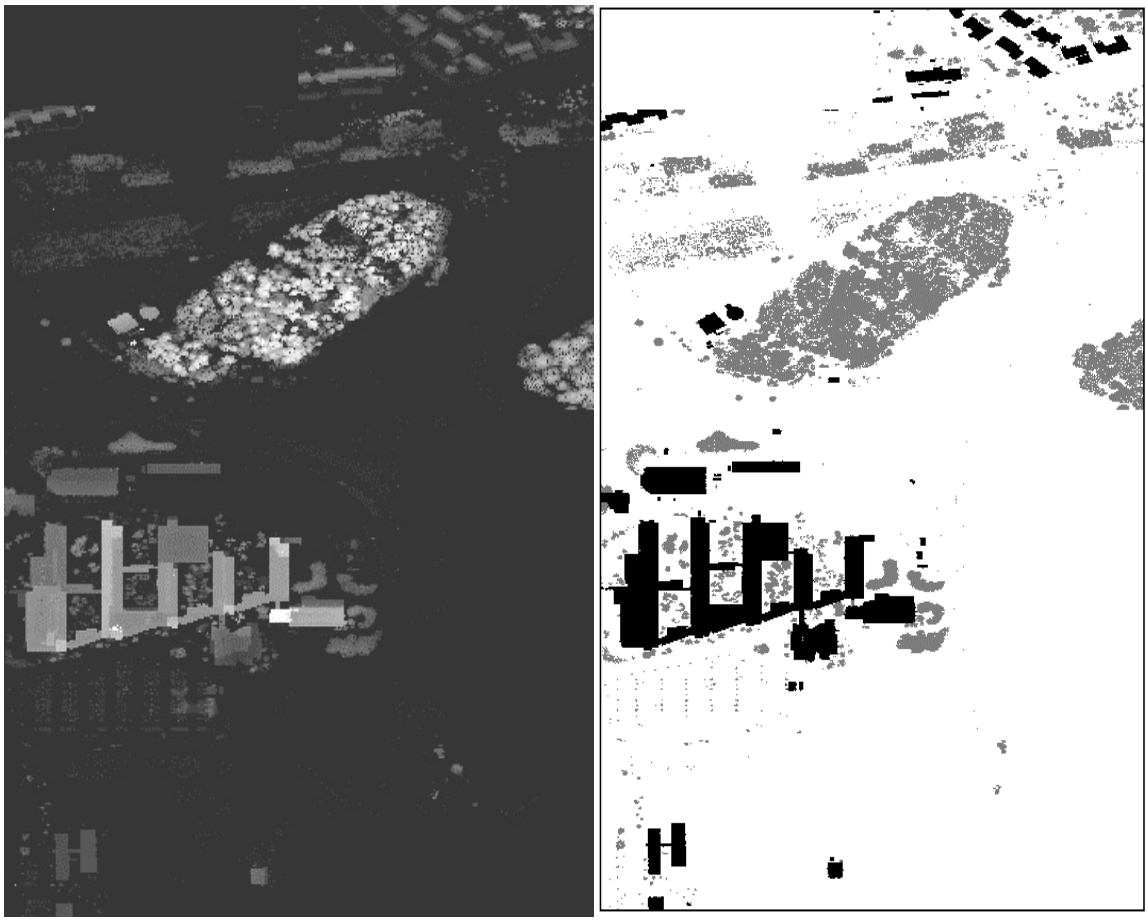


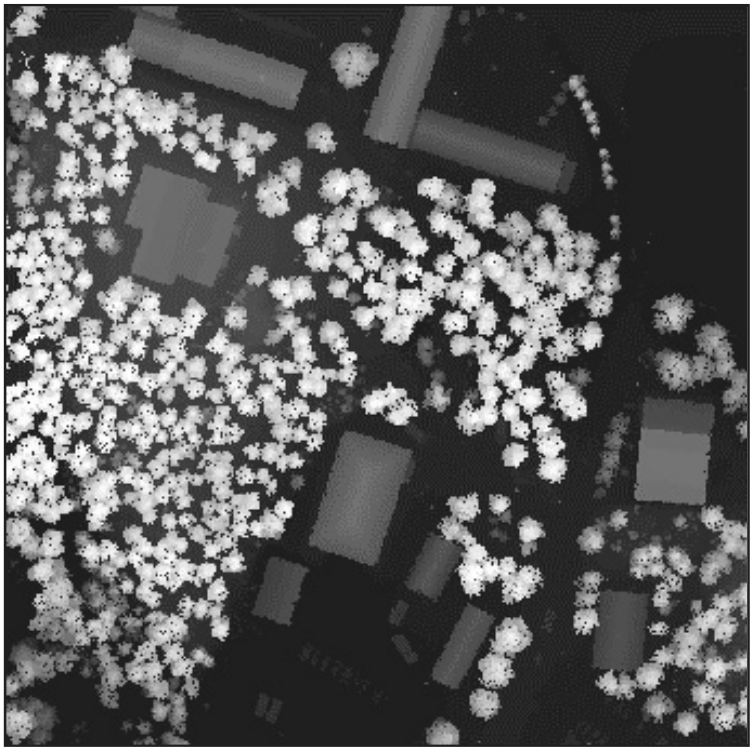
Fig A.3. (a) Structuring element B. (b) Connected component Y and initial point p contained in set A. (c) Result of first iterative step. (d) Result of second iterative step. (e) Final result. From [5].

10 Appendix B

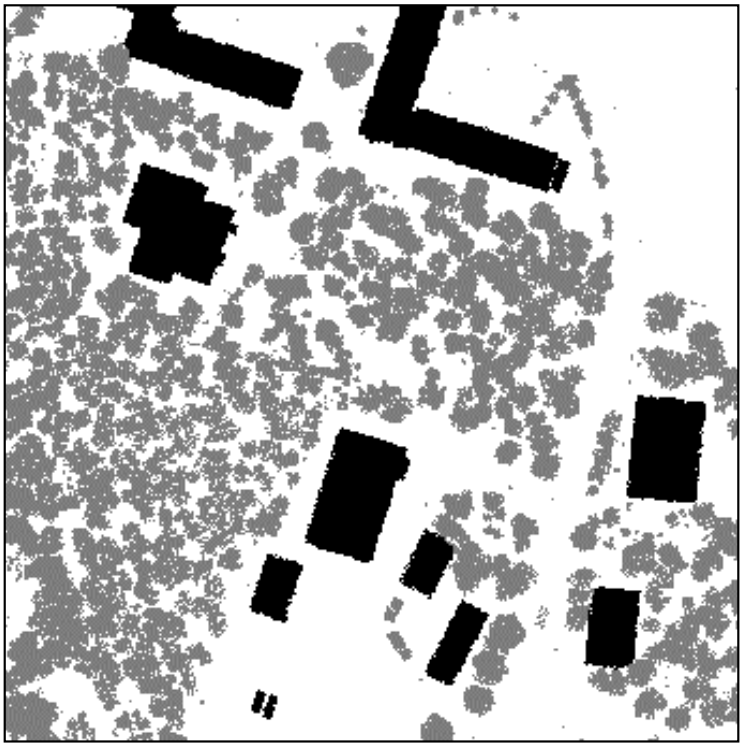


(a) (b)

Fig B.1. (a) Elevation data over an area of 900x400 m. (b) Segmentation and classification of vegetation and buildings.



(a)



(b)

Fig B.2. (a) Elevation data over an area of 200x200 m. (b) Segmentation and classification of vegetation and buildings.

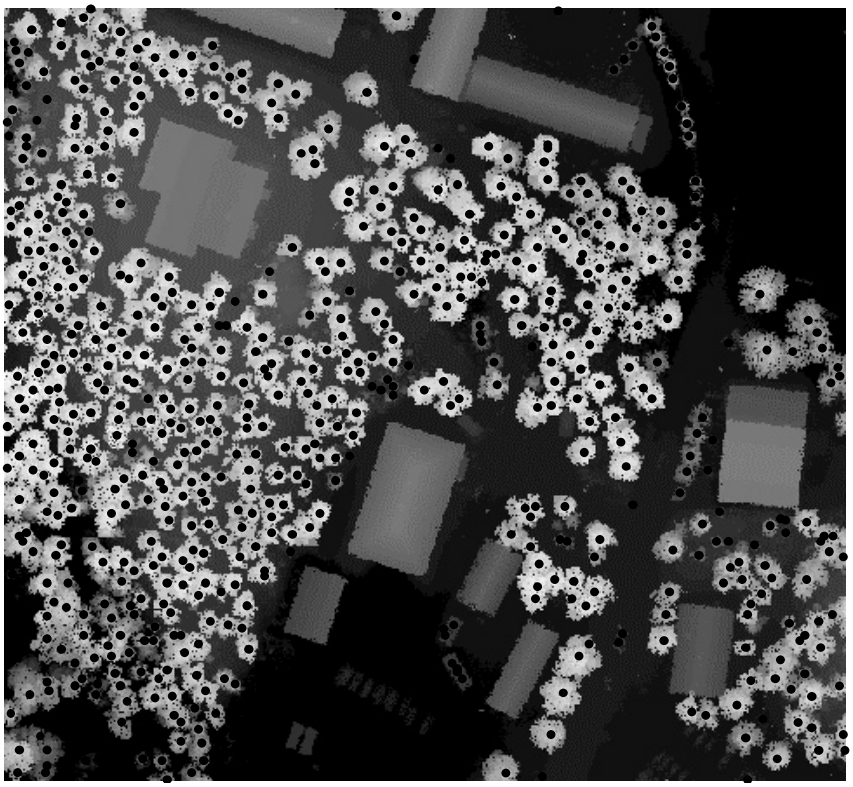


Fig B.3. Estimated positions of trees marked on the elevation data over an area of 200x200 m.

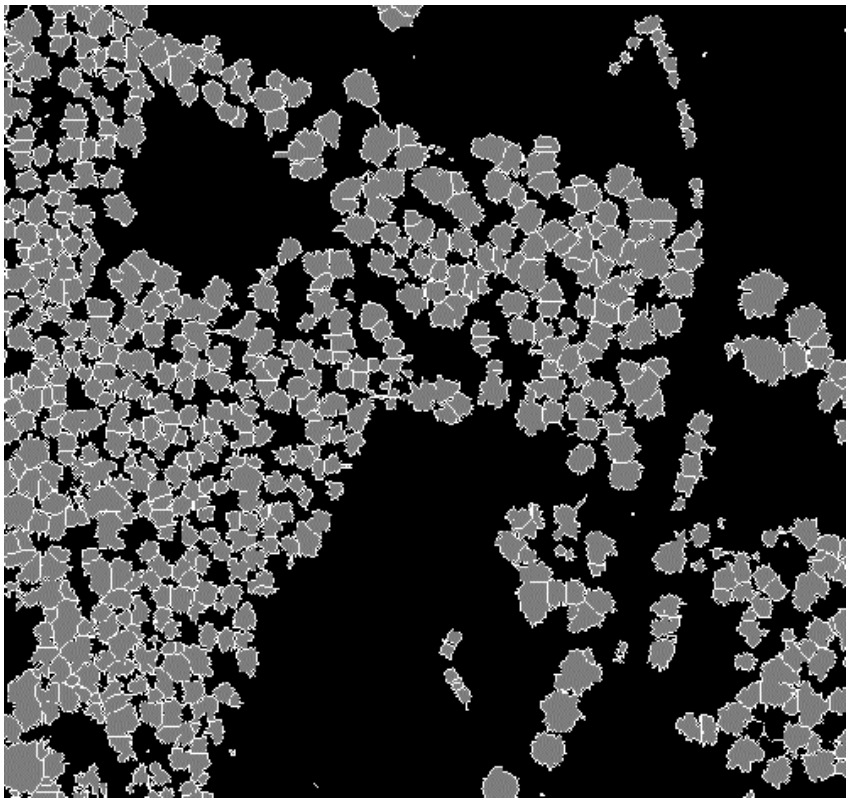


Fig B.4. Estimated crown coverage.

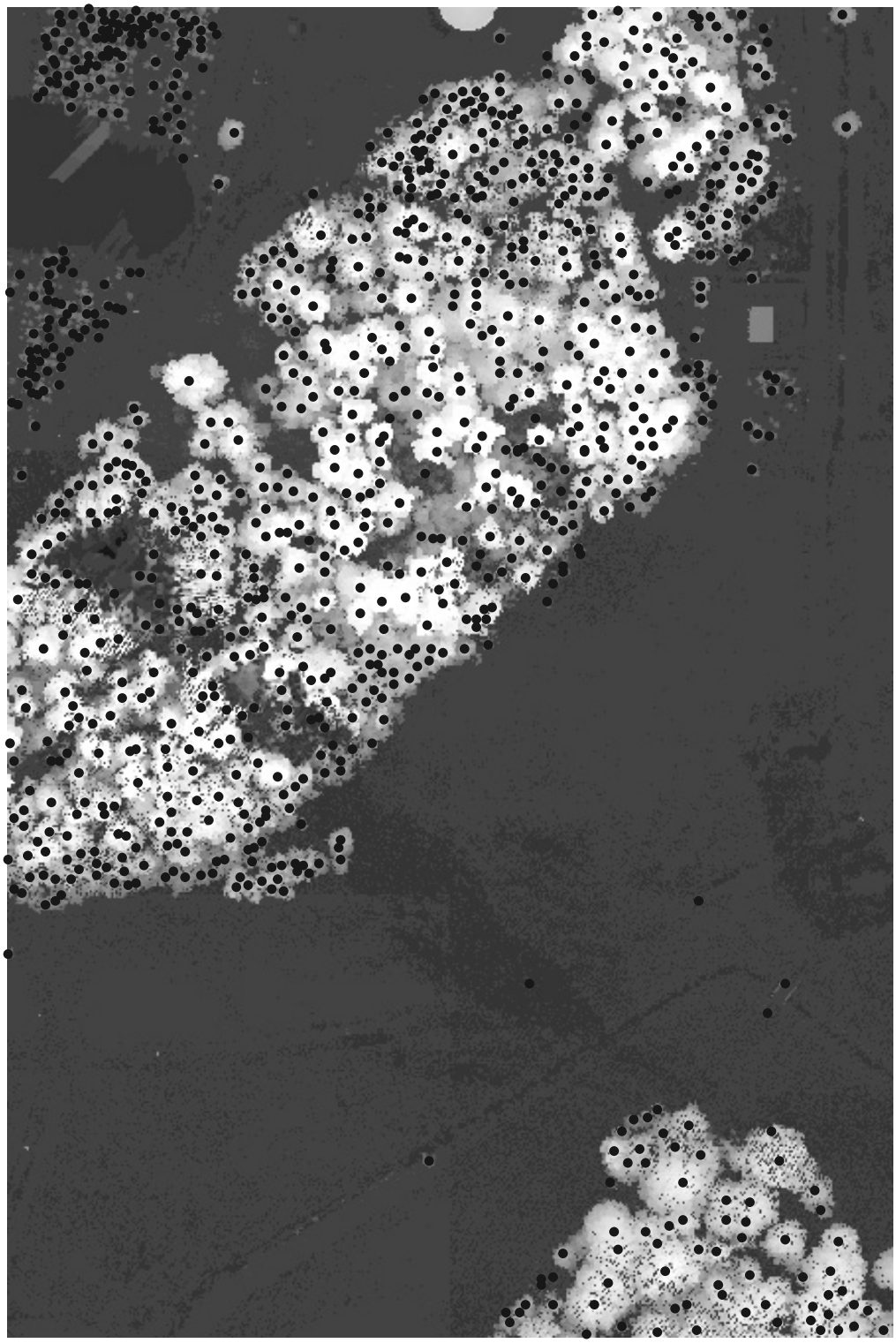


Fig B.5. Estimated positions of trees marked on the elevation data over an area of 300x200 m.