Segmentation and Classification of Individual Tree Crowns

in High Spatial Resolution Aerial Images

Mats Erikson

Centre for Image Analysis
Uppsala

Doctoral Thesis
Swedish University of Agricultural Sciences
Uppsala   2004

Acta Universitatis Agriculturae Sueciae
Silvestria 320

Abstract

Mats Erikson. *Segmentation and Classification of Individual Tree Crowns*. Doctoral Thesis

ISSN 1401-6230, ISBN 91-576-6704-7

By segmentation and classification of individual tree crowns in high spatial resolution aerial images, information about the forest can be automatically extracted. Segmentation is about finding the individual tree crowns and giving each of them a unique label. Classification, on the other hand, is about recognising the species of the tree. The information of each individual tree in the forest increases the knowledge about the forest which can be useful for managements, biodiversity assessment, etc.

Different algorithms for segmenting individual tree crowns are presented and also compared to each other in order to find their strengths and weaknesses. All segmentation algorithms developed in this thesis focus on preserving the shape of the tree crown. Regions, representing the segmented tree crowns, grow according to certain rules from seed points. One method starts from many regions for each tree crown and searches for the region that fits the tree crown best. The other methods start from a set of seed points, representing the locations of the tree crowns, to create the regions. The segmentation result varies from 73 to 95 % correctly segmented visual tree crowns depending on the type of forest and the method. The former value is for a naturally generated mixed forest and the latter for a non-mixed forest.

The classification method presented uses shape information of the segments and colour information of the corresponding tree crown in order to decide the species. The classification method classifies 77 % of the visual trees correctly in a naturally generated mixed forest, but on a forest stand level the classification is over 90 %.

Key words: aerial photography, high spatial resolution, image analysis, individual tree crown, remote sensing, segmentation, species classification

Author's address: Centre for Image Analysis, Lägerhyddsvägen 3, 752 37 Uppsala, Sweden

Till mamma och pappa

Contents

Segmentation and Classification of Individual Tree Crowns

Papers I – V

Papers appended to the thesis

The thesis is based on the following articles. The included non-published papers are slight modifications of earlier submitted versions.

I   Erikson, M., Segmentation of individual tree crowns in colour aerial photographs using region growing supported by fuzzy rules, *Canadian Journal of Forest Research*, Vol. 33, No. 8, 2003, pp. 1557-1563. Reprinted with permission from NRC Research Press.

II   Erikson, M., Structure-preserving segmentation of individual tree crowns by brownian motion, in Proc. of the 13th Scandinavian Conference on Image Analysis (eds. J. Bigun, T. Gustavsson), 29 June - 2 July, 2003, Gothenburg, Sweden, *Lecture Notes in Computer Science 2749*, Springer-Verlag, Berlin, Germany, 2003, pp. 283-289. Reprinted with permission from Springer-Verlag.

III   Erikson, M., Two preprocessing techniques based on grey level and geometric thickness to improve segmentation results, submitted for journal publication.

IV   Erikson, M., Olofsson, K., Comparison of three individual tree crown detection methods, submitted for journal publication.

V   Erikson, M., Species classification of individually segmented tree crowns in high-resolution aerial images using radiometric and morphologic image measures, *Remote Sensing of Environment*, Vol. 91, No. 3-4, 2004, pp. 469-477. Reprinted with permission from Elsevier.

1 Introduction

Analysis of individual tree crowns is about finding individual tree crowns in aerial images and to perform analysis, such as tree species classification, for each of these tree crowns. Such analysis of individual tree crowns in digital images requires high spatial resolution. In the community of tree crown research, high spatial resolution often means a pixel size corresponding to at most one meter on the ground. Therefore, airborne sensors are frequently used when capturing the images. In this thesis, a pixel size corresponding to at most a few decimeters is used. This means that the shape of the tree crown starts to be visible in the images if the crown diameter is at least about two meters. Since more details are included in the image with better resolution, not only the crowns in the overstory are more visible but also the understory vegetation. This can create problems when separating the interesting parts of the image from the non-interesting parts. This separation is called segmentation and is something that has to be done independently of the resolution and of the subsequential analysis.

In this thesis, segmentation means that each individual tree crown must be separated from the other tree crowns. Different segmentation techniques can be used depending on the subsequent analysis. For instance, if the objective is to count the number of tree crowns in the image, it is sufficient to find some part of each tree crown. On the other hand, if the objective is to find the species of the tree crown, the above result is not good enough. Instead of just finding a part of the tree crown, the complete tree crown must be found, otherwise, information used in the decision of the species may be lost resulting in erroneous classification. To find the precise tree crown is a task, however, that can never be solved completely. The solution can, at most, become satisfactory. One reason is that a tree crown in an image does not have a distinct border. Even for a single isolated tree, it is not obvious where the tree crown ends, see Fig. 1. This makes it impossible to know, at pixel level, if the full tree crown is found or not. Of course, it is even harder to separate two closely standing trees, where the crowns are touching each other, since the distinction between the two tree crowns is even fuzzier than that between a tree crown and the background. One of the objectives in this work is to find the tree species. Therefore, great effort is spent on the segmentation of the individual tree crowns, in the sense that the shape of the tree crown, as seen in the images, is preserved. All developed methods in the thesis belongs to the class of region growing methods. A region growing method needs seed points and rules on how the regions can grow.

The interest of finding individual tree crowns in digital images has its roots in 1920, which is the first time aerial photo-interpretation was used in forest inventories (Howard, 1991). Since the breakthrough, after World War II, the technique of using remote sensing in forestry has been frequently used. Today, almost all forest companies use a method based on manual interpretation of remotely sensed data in combination with field measurements, in order to find the forest parameters (Walter, 1998). A list of forest parameters is given in Walter (1998) in order of relevance, according to forest owners with more than 5000 ha forest. The four most important parameters are stem volume, age, tree species composition, and ecological values such as key biotopes and habitats. All these parameters can be estimated by interpre-
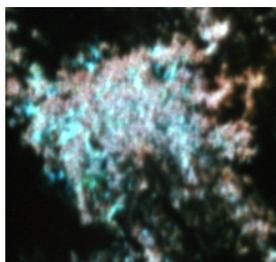
Figure 1: Even for a human, the distinction between background and object can be difficult for a single tree crown.

tation of aerial photographs using a stereoscope (Anon., 1993). Pancromatic images with a scale of 1:20,000 or 1:30,000 are usually used for this kind of interpretation. In these images, individual tree crowns are very small and not easily seen. Therefore, the parameters are estimated stand wise with probably lower accuracy in the estimates than for estimates based on individual tree crowns. Different measures are used for estimation, for example, texture, grey levels, the shapes of the crowns (from shadows), crown closure, the size of the crowns, etc. By using colour infrared (CIR) images, the distinction between tree species is easier, especially the distinction between conifers and deciduous. Deciduous will have a white to yellow-white colour while conifers will have bluish or blue-green colour. Of course, the colour depends a lot of the day of the year the image is captured and also of the age of the trees in the forest. The colour given above corresponds to an image captured in the early autumn above a mature forest. Although CIR images can distinguish between some tree species, it is often the case that colour differences are not enough. It can also be that the colour is not representative for the species due to illumination variations from clouds, etc. In order to overcome these problems as well as getting better accuracy, shape and texture of the illuminated individual tree crowns should be used, together with colour information. The measures extracted from each individual tree crown, such as tree species, tree crown area, etc., can be useful for silviculture treatments, selective cuts, and biodiversity assessments (Gougeon, 1998). From the tree crown diameter it is possible to estimate the stem dimensions and thus the stem volume (Minor, 1951; Jakobsons, 1970). By finding the tree species, the composition is also found.

1.1 Individual tree crown research

Research in the field of identifying single tree crowns in high spatial resolution images have been going on for several years. Various approaches and variations among these has been presented during the years (Pinz, 1989; Gougeon, 1995; Pollock, 1996; Brandtberg & Walter, 1998; Larsen & Rudemo, 1998; Culvenor, 2002; Olofsson, 2002; Pouliot et al., 2002). A short summary of some of these methods is given in the following subsections. It is important to remember that the methods are developed with different subsequent analysis in mind, therefore, results cannot be immediately compared.

1.1.1 The valley-following approach

The valley-following approach (Gougeon, 1995) uses an assumption that there exist dark pixels between the tree crowns, as in Fig. 2. To describe the method, an analogy where the image is seen as a landscape is used. Bright pixels (trees) represents mountains or high altitude while dark pixels (background) represents valleys or low altitude, see Fig 3. To find these valleys in the analogy, first all non-forest areas are found by using a threshold. Then, all local minima are found by using a 3×3 scanning filter. The centre pixel of the filter is a local minimum if all other pixels in the filter have greater grey-level value. The neighbourhood of these valley pixels are further examined by using four repeatedly directional filter scans to include more valley pixels. In the analogy, this is equivalent to walking in the valley towards the saddle point defined by two valleys meeting between two tops. The four scans are repeated until no further pixels are considered to belong to a valley. Although this results in an image where almost all of the crowns are delineated, some of them can still be connected. Therefore, certain rules are used to perform further splitting from the boundary of each delineated tree crown in order to get the final segmentation. With the resolution used in this thesis, too many valleys would be found with this technique and thus the tree crowns would be divided into subcrowns.
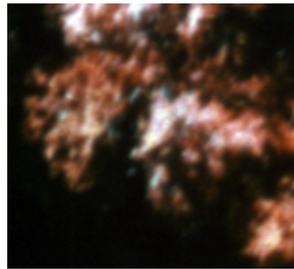

Figure 2: A usual occurrence is the dark boundary around a tree crown, even though tree crowns are connected to each other.
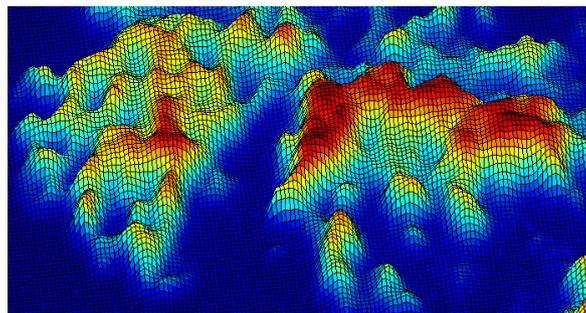

Figure 3: An image consisting of two tree crowns seen as a landscape. Compare with Fig. 2 which shows the original image.

1.1.2 Finding edge contours at multiple scales

Scale-space theory is a mathematical tool for analysing structures in images at different scales. Different scales are achieved by smoothing the image with different Gaussian kernels (Lindeberg, 1998) or with discrete versions of the Gaussian kernel (Lindeberg, 1994), see Fig. 4. Brandtberg & Walter (1998) considered the latter case to round off the contour of each tree crown at a certain scale, when the algorithm was constructed. The discrete scale-space representation, $L(x, y, t)$, is given by,

$$L(x, y, t) = \sum_{m=-\infty}^{\infty} T(m, t) \sum_{n=-\infty}^{\infty} T(n, t) f(x - m, y - n), \qquad (1)$$

where $f(x, y)$ is the image under consideration and $T(n, t) = e^{-t} I_n(t)$, where t is the smoothing factor and $I_n(t)$ are the modified Bessel functions of integer order.

Figure 4: Examples of discrete scale-space smoothing at different scales. Left: Original image. Middle: Smoothing factor $t = 10$. Right: Smoothing factor $t = 300$.

Edge pixels of the smoothed image, $L(x, y, t)$, are pixels where

$$\frac{L_x^2 L_{xx} + 2L_x L_y L_{xy} + L_y^2 L_{yy}}{L_x^2 + L_y^2} = 0, \qquad (2)$$

and

$$\frac{L_x^3 L_{xxx} + 3L_x^2 L_y L_{xxy} + 3L_x L_y^2 L_{xyy} + L_y^3 L_{yyy}}{(L_x^2 + L_y^2)^{3/2}} < 0, \qquad (3)$$

for a certain scale t (Lindeberg, 1994). Here, $L_x, L_y, L_{xx}, L_{yy}, L_{xy}, L_{xxx}, L_{yyy}, L_{xyy}$, and L_{xxy} are partial derivatives. These pixels are local maxima of the gradient magnitude in the gradient direction of the grey levels. Thus, Eq. 2 finds the location of the edge and Eq. 3 checks whether it is a maximum or not. The 2D-curvature of the level curve defined by the edge pixels is calculated as,

$$\kappa = \frac{L_x^2 L_{yy} - 2L_x L_y L_{xy} + L_y^2 L_{xx}}{(L_x^2 + L_y^2)^{3/2}}, \qquad (4)$$

11

to remove pixels with positive curvature, $\kappa \geq 0$. For edge pixels with $\kappa < 0$, the centre of a circle with the radius of curvature $\rho = -1/\kappa$, is used to build an ellipse for each contour segment. This ellipse is then used to build a primal sketch, which describes significant curvature changes of the contours. The primal sketch is built for different scales and an accumulated primal sketch for these scales is also built. In the accumulated primal sketch, each tree crown forms a local maximum. From these maxima, each region is then grown in the accumulated primal sketch, to form the final segments. Note that the segments will not keep the shape of the tree crowns, as seen in the original images, since smoothing is involved.

1.1.3 Template Matching

In contrast to the other methods, template matching (Pollock, 1996; Olofsson, 2002) does not rely on clearly visible tree crown boundaries or large contrast between different tree crowns. Instead, template matching finds the best match between a synthetic template representing a crown and an equally large region in the image. The synthetic template are constructed by assuming the tree crowns are rotationally symmetric about a vertical axis. Thus, a tree crown can be modelled by a generalised ellipsoid of revolution,

$$\frac{z^n}{a^n} + \frac{\left(x^2 + y^2\right)^{n/2}}{b^n} = 1, \tag{5}$$

where a, b and n are integers. With $n = 2$, Eq. 5 represents an ellipsoid, where a and b are the semi-axes in the z-direction and the x- and y-directions, respectively. By changing n in Eq. 5, the shape of the generalised ellipsoid of revolution is changed. Different values for a, b, and n are used to generate 3-D models of different sizes of tree crowns as well as for different species, see Fig 5.
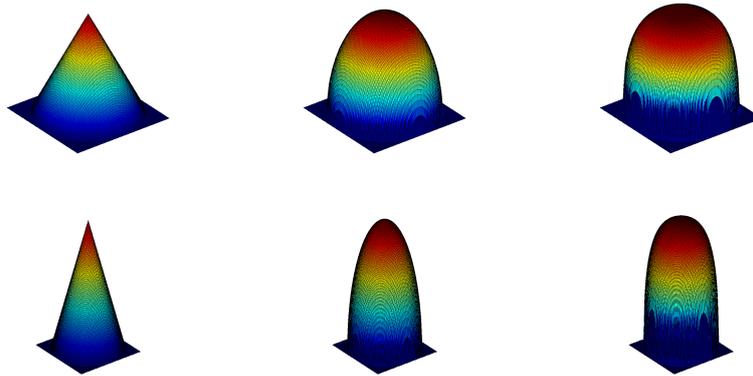

Figure 5: Examples of generalised ellipsoid of revolution for different a and n. $b = 50$ for all cases. Upper row: $a = 100$. Lower row: $a = 200$. Left column: $n = 1$. Middle column: $n = 2$. Right column: $n = 3$.

The 3D-models of the tree crowns are illuminated by a parallel light source from different viewing angles and for a specific sun angle and camera position. The illuminated model is then projected onto the image plane to create the template, see Fig. 6. With a suitable definition of the correlation between each template and each corresponding region in the image, a correlation image is calculated. Local maxima, above a certain threshold in the correlation image, represent possible tree crown positions. The template associated to a local maximum represents the tree crown. If two or more local maxima are close to each other, such that the templates overlap each other, they are candidates to the same tree crown. In this case, all except one of these local maxima are removed. In the final result, a segment is represented by a template. Since there is a finite number of templates, the actual shape of the tree crown is not possible to find, only a rough approximation.
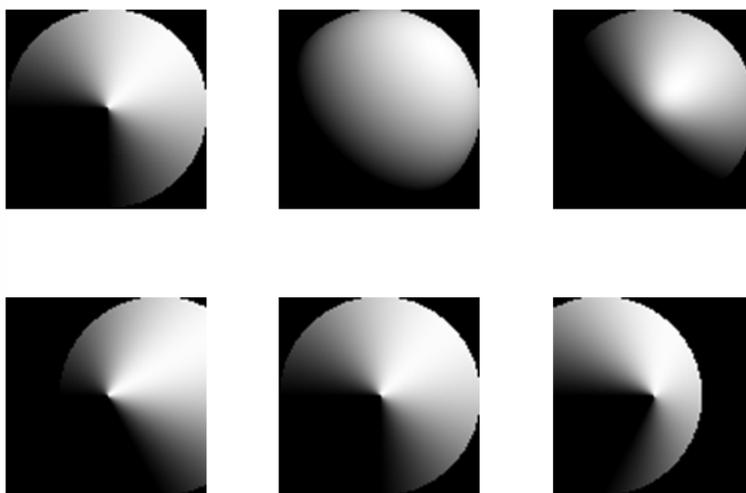

Figure 6: Examples of templates generated for different viewing angles and with different $a, b,$ and n in Eq. 5.

1.1.4 Other techniques using different sensors

This far, only aerial images have been discussed. It is worth mentioning that there exist other types of remotely sensed data used in individual tree crown research. Perhaps the most frequently used technique, beside aerial photographs, is lidar which is an abbreviation of light detection and ranging. Lidar produces a depth image by shooting laser pulses toward the ground and calculating the distances from the times the pulses take to return. Some pulses return after hitting the tree canopy while others after hitting the ground. The general strategy is to estimate the canopy surface and the ground surface from the depth image. Persson et al. (2002) estimated the surfaces using deformable surface models. Then, the separation of the tree crowns were done by seeding every pixel and letting the seeds climb up the steepest descent until all seeds reach a maximum. All seeds reaching the same maximum form a

segment. From each segment, height, crown diameter, stem diameter, and stem volume, were found. The maximum laser height above the ground surface represented the height of the tree crown. The crown diameter was calculated from the area of the segment and a linear model with the product of the crown diameter and the height was used to calculate the stem diameter. The stem diameter together with the height was used to calculate the stem volume.

The height distribution of the laser pulses of a segment can also be used for identification of the tree species. Holmgren & Persson (2004) used features like crown base height, crown shape, and various statistical measures of the height distribution to classify segments into either pine or spruce.

Radar, standing for radio detection and ranging, is also used in remotely sensed forest research. Instead of shooting electromagnetic waves with wavelength around some hundreds nanometers as in lidar, radar uses electromagnetic waves with wavelength ranging from a few centimeters to ten of meters.

2 Objectives

The main objective of this work was to develop algorithms for classifying individual tree crowns into tree species. In order to do so, the individual tree crowns must be identified, i.e. segmented. Since the classification step relies on the shape of the tree crowns, the objective in the segmentation step is to segment the tree crowns such that the shape of the segment agrees with the shape of the tree crown in the original image. The objective of the thesis is illustrated by a flowchart shown in Fig. 7.
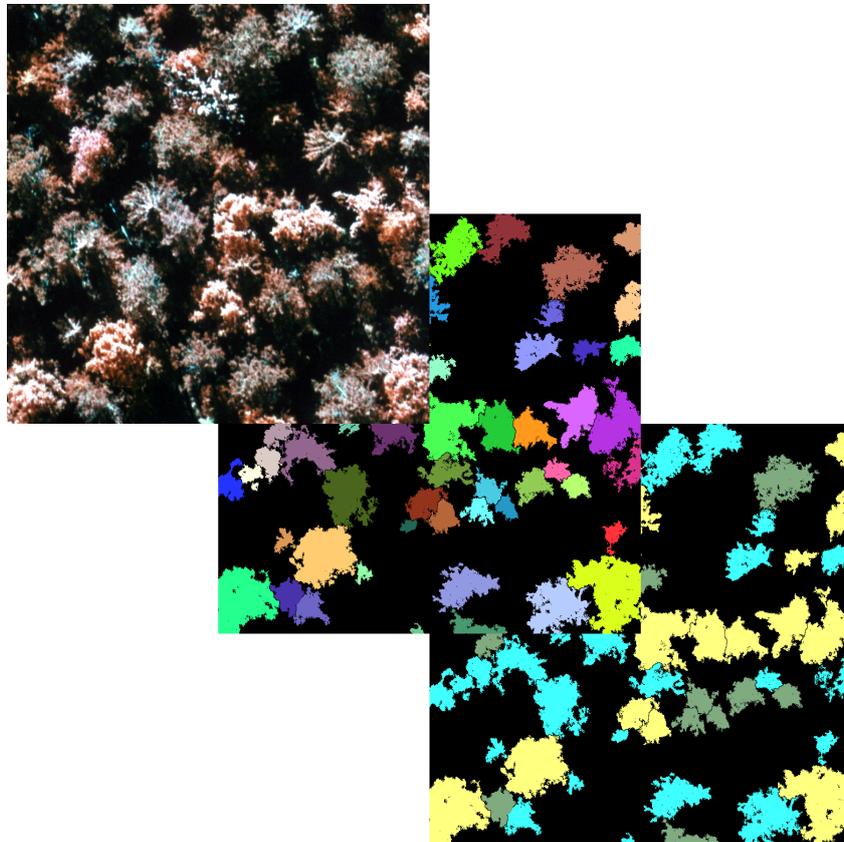

Figure 7: Original (upper) image is segmented into individual tree crowns (middle image) which are classified into tree species (lower image). The colours in the segmented image represents different tree crowns and in the classified image they represent different species.

3 Material

Three different test sites are used in this work. The main material, from the Huljen test site, is located in Medelpad, Sweden, and is the same material as used in the thesis by Brandtberg (1999). The second test site is the Remningstorp estate in Västergötland, Sweden. The third test site is the Tiputini Biodiversity Station, located in Ecuador.

3.1 Huljen test site

The material consists of a data set of 50 colour infrared aerial photographs. The wavelengths of the three bands, denoted v_1, v_2, v_3, are $v_1 : 700 - 900$ nm (infrared), $v_2 : 600 - 700$ nm (red), and $v_3 : 500 - 600$ nm (green), respectively. The photographs were captured on a sunny day, 10th August 1995, between 13:00 and 14:30 local time, from an aircraft at 600 m above ground level, in a square with side length 5 km at Huljen outside Sundsvall in Central Sweden $(62°27'N, 16°55'E)$. The camera was a Wild RC30/4 NAT-S 17111 with focal length of 302.97 mm, which gives a scale of 1:2000, and the film was Kodak Aerochrome Infrared Film 2443. During the acquisition the solar zenith angle, the angle between the local zenith and the line of sight to the sun, was between $46.9°$ and $49.7°$ and the azimuth angle, the angle between the line of sight to the sun and the line of sight to the camera in the image plane, between $180.7°$ and $209.6°$. Each photograph covered an area of 450×450 m and the forest covered by the photographs is naturally regenerated and mature (approx. 80 years old). It consists of 23 % Scots pine (*Pinus Sylvestris* L.), 61 % Norway spruce (*Picea abies* Karst.), 14 % birch (*Betula pubescens* Ehrh.), and 2 % aspen (*Populus tremula* L.). No other species exists in the image material. The terrain is hilly and has an altitude around 300 m above sea level.

A square with side length 150 m on the ground was scanned in the centre of each photograph. This corresponds to a maximum view angle of $10.0°$, see Fig 8. The photographs were scanned to a size of 5000×5000 pixels which corresponds to 3 cm on the ground, 67 pixels/mm in the image, or a pixel size of 15μm. The central square was used, since the tree crowns are then seen from straight above and not "leaning" towards the border of the image, see Fig. 9. The leaning trees must be processed in different ways depending on which direction they lean towards and how much they lean while all trees seen straight from above can be treated in the same manner.

3.2 Remningstorp estate

In a comparison between different methods another image material is used, located at the Remningstorp estate in the south west of Sweden $(58°30'N, 13°40'E)$. The aerial photographs were captured on the 25th September 2001 around 12:00 from a flight height of 600 m above the ground. The camera was a Hasselblad SWCE 61085, lens Biogon 38 mm f/4.5, with a digital back. This corresponds to a scale of 1:15,800, a resolution corresponding to 15 cm on ground, 105 pixels/mm in the image, or a pixel size of 9.5μm. The images are True color RGB images, see

Figure 8: Only the photograph part inside the white square was scanned and used.

Fig. 10. The images were rectified into the geo-referenced system RT90, using known objects in the photo. Four subimages with were used, two images containing pine stands and two images containing spruce stands. The maximum view angle was $12.4°$. Each image contained one field site where all trees with larger stem diameter than $0.05\,\text{m}$ (at $1.3\,\text{m}$ height) within the field site were measured and the positions were registered. No other information of the test site is available.

3.3 Tiputini Biodiversity Station

The Tiputini Biodiversity Station is located in northern Ecuador in a tropical forest. Three different flight heights, $600\,\text{m}$, $300\,\text{m}$, and $200\,\text{m}$, were used when the images were captured as RGB images. The camera, a Kodak DC420, was mounted on an aircraft for flight heights $600\,\text{m}$ and $300\,\text{m}$, while a Nikon Coolpix 990 was mounted
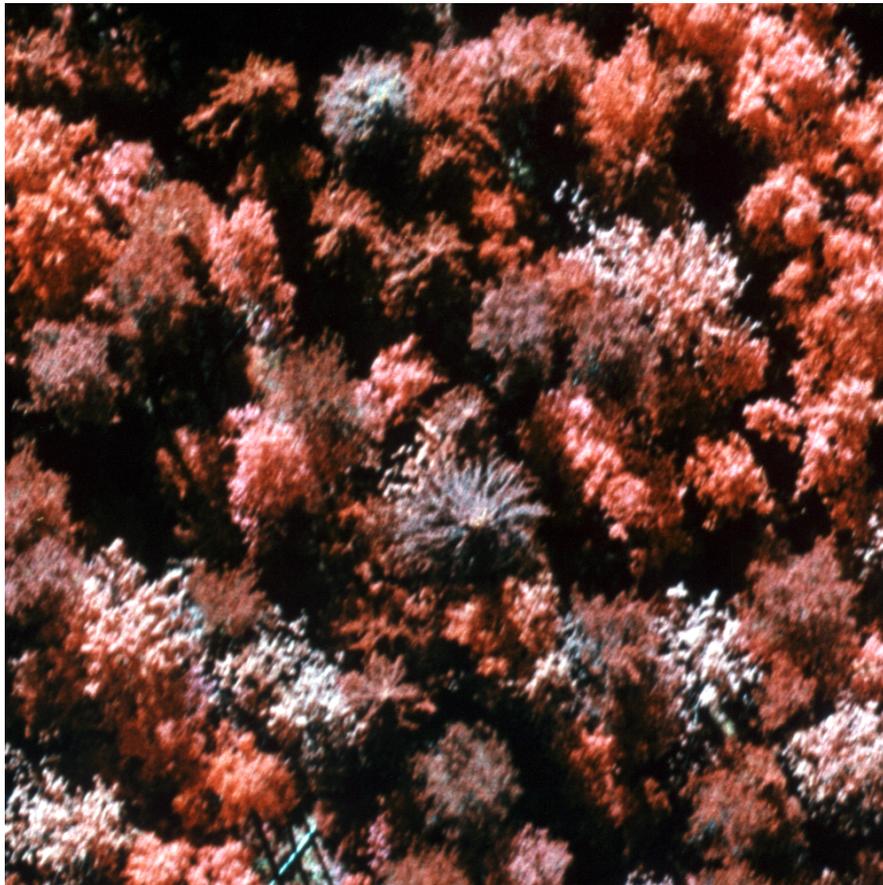
Figure 9: A CIR image from Huljen, Medelpad.

on a balloon for flight height 200 m. The number of images captured for each flight height was 24, 14, and 11, respectively. The corresponding pixel size on ground is 21.4 cm, 8.4 cm, and 7.6 cm, respectively. Example images of the three heights are shown in Figs. 11, 12, and 13 for 600 m, 300 m, and 200 m, respectively. No other information of the test site is available.

3.4   Colour images versus colour infrared images

When capturing a photograph with an ordinary colour film (RGB), the sensitivity curves of the film has peaks in the wavelengths representing red (600-700 nm), green (500-600 nm), and blue (400-500 nm) light. Such an image is written (r, g, b), to describe the bands. If colour infrared film (CIR) is used instead, the sensitivity curves has peaks in the wavelengths representing infrared (700-900 nm), red (600-700 nm), and green (500-600 nm) light. Such an image is written (i, r, g), to describe the bands. This means that when a CIR image is displayed on standard monitors, red

Figure 10: Two RGB images from Remningstorp, Västergötland. Left: A pine stand. Right: A spruce stand.

objects are shown as green and green objects are shown as blue. Objects shown as red are objects reflecting a lot of infrared light.

The reflection of healthy vegetation has, typically, a peak in the green wavelength band, a minimum in the red and the blue bands, and a very large reflectance in the infrared wavelength band. Unhealthy vegetation, though, reflects little in-
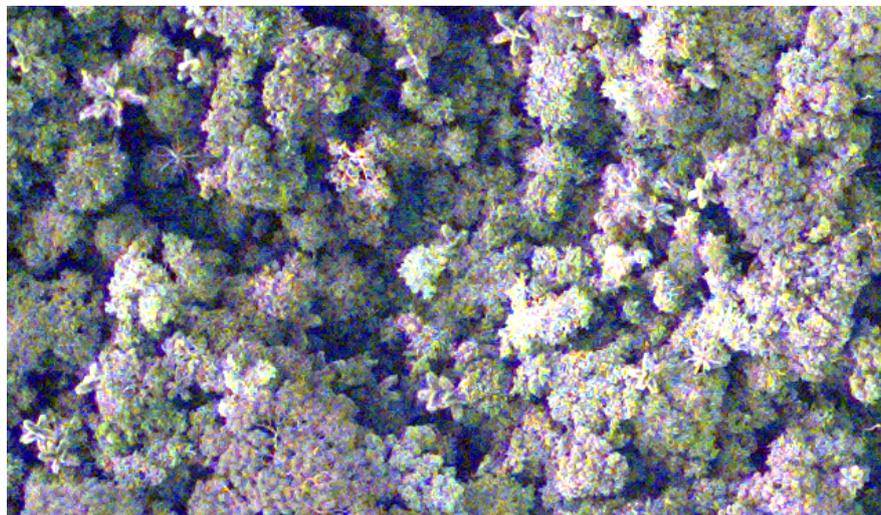

Figure 11: An image from the Tiputini Biodiversity Station, Ecuador. Flight height 600 m.
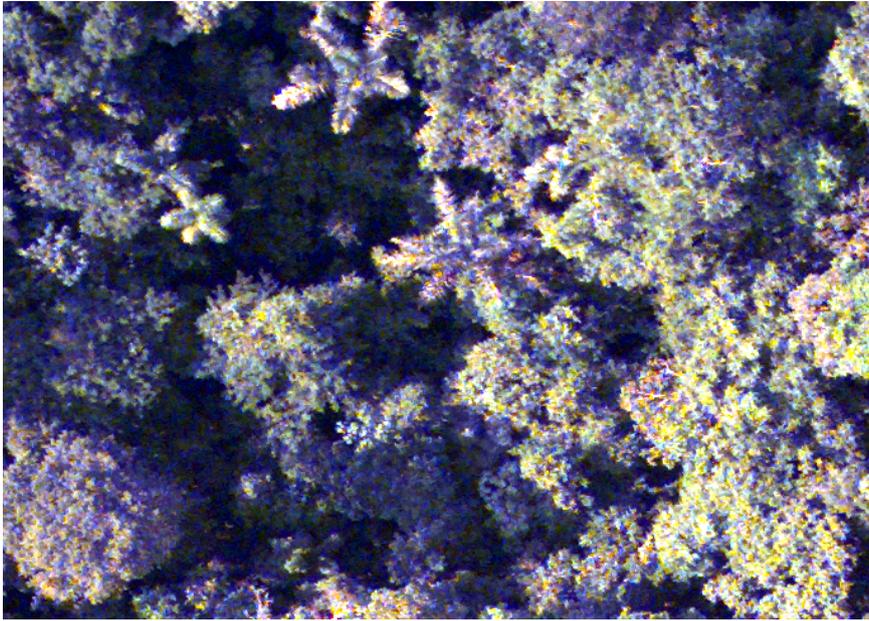
Figure 12: An image from the Tiputini Biodiversity Station, Ecuador. Flight height 300 m.

Figure 13: An image from the Tiputini Biodiversity Station, Ecuador. Flight height 200 m.

20

frared light. The reflectance differences between different vegetation are largest for the infrared wavelength, but there are also reflectance differences depending on the time of the year. The largest colour differences between different species are in the spring during the leafing and in the autumn during the defoliation, since the time of the year for these events are different for different species (Anon., 1993). The reflection difference between two closely standing tree crowns can also be useful when segmenting them. Thus, for classification of different tree species and sometimes for identification of the individual tree crowns, CIR images are preferable.

4 Methods and Results

In the following sections, the methods developed for segmentation and classification in the appended papers are described.

4.1 Segmentation

Segmentation is often the first step the image undergoes in computerised image analysis. In this thesis, segmentation of an image means separation of the objects in the image from the background as well as separation between individual object. This is done by letting the background pixels be represented by zeros and by letting all pixels belonging to an individual object have the same nonzero number, which is unique for the object. Each number can then be represented by a colour, not necessarily unique, when the image is displayed on the monitor. Sometimes it is very easy to segment an image, for example, if the background is very dark compared to a single object in the image. A single function can then easily be found, that maps all intensity values below or equal to a given value (threshold) to zero and all intensity values above the same value to one. For most cases though, the above described procedure is not applicable, for various reasons. The image in Fig. 9, for example, has a very dark background compared to the tree crowns. However, the thresholding technique described above, will only separate the tree canopy from the ground and not the tree crowns from each other. Since this is also required, more sophisticated methods for segmentation of the individual tree crowns must be found.

Region growing methods is a class of segmentation methods where the segments grow, according to some similarity rules, from a number of seed points. Generally, if no postprocessing is made, the number of seed points decides the number of segments. Thus, in order to use a region growing method, basically two problems must be solved. The first problem concerns the seed points and the second problem concerns the similarity rules, or how the regions grow. In the following sections, some techniques for solving these questions are discussed in the application of segmentation of individual tree crowns.

4.1.1 Segmentation supported by fuzzy rules

Fuzzy set theory was first introduced by Zadeh (1965) to generalise the conventional set theory. In conventional set theory, a member can either belong to or not belong to the set. However, in fuzzy set theory, a member belongs to all sets to a certain degree. The degree is decided by a membership function which takes a member and returns a value between zero and one. With full membership of the set, the value is one and then the grade of membership decreases with decreasing value. A zero value means that the member is not compatible with the set at all. In other words, the grade of membership depends on how compatible the member is to the properties of the set.

This segmentation method is inspired by the fuzzy set theory, in the way that a pixel must have a certain degree of membership, in order to belong to a region. The assumptions for the method are that the variation of colour within the tree crown is

22

less than the variation between the crowns and that there exists a dark valley between the tree crowns, as used in the valley-following approach, see Section 1.1.1.

The first question, concerning the seed points, is solved by finding one or several seed points in each tree. In order not to end up with too many tree crowns, superfluous seed points are removed from the final segment. First, the original image is thresholded at 0.3 (a rough threshold only used for approximate results) and a distance transform (Borgefors, 1986) is performed on the thresholded image to create a distance image. The distance image is found by a two-pass scan on the image with the value 3 for the two 4-neighbours and the value 4 for the two 8-neighbours in the scanning mask which consists of the already visited neighbours. In the distance image, the grey-level values represent the shortest distance from the pixel to the background. The seed points, defined as local maxima in the distance image, are found by using a 3×3 scanning filter. The centre pixel of the scanning filter is a local maximum if all other pixels in the filter have lower grey-level value (distance value). The corresponding distance value of each seed point is used to determine the processing order of the seeds. The larger value it has, the earlier it is processed. Fig. 14 illustrates an image and its seed points.
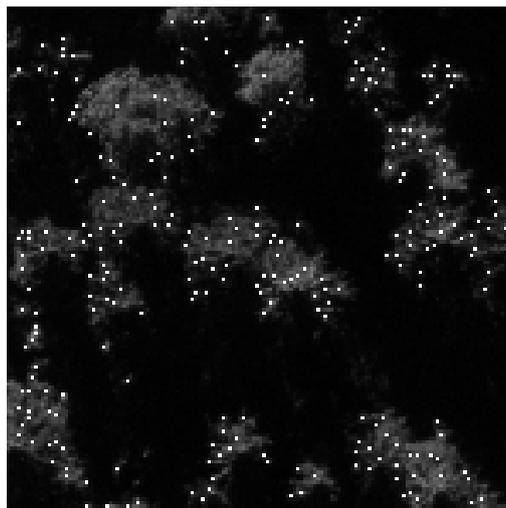

Figure 14: Seed points overlaid the original image.

For a given seed point, the segment is found by performing a number of steps. The first step is to extract the region, in a thresholded image, in which the seed is included. The threshold is calculated by a fuzzy thresholding algorithm (Chi et al., 1996), where different thresholds are tried. For each threshold, the image is divided into two sets, background and objects. A membership value for each pixel in the image is calculated, depending on which set the pixel belongs to. The optimal threshold is found by calculating an entropy measure for all possible thresholds and selecting the threshold attaining the minimum of that measure. Since the threshold only separates the tree crowns from the background, the extracted region often

consists of a number of real tree crowns. To locate one of the tree crowns, the boundary of the region is first smoothed by convolution with a Gaussian function. From the smoothed boundary, a curvature value for each pixel can be calculated. Boundary segments consisting of pixels with convex curvature are extracted and the one consisting of most pixels among these segments is used as the most significant tree crown. Fig. 15 illustrates the region in which the first processed seed point is included as well as the smoothed boundary.

Figure 15: Left: The region extracted from the thresholded image. Right: The smoothed boundary shown as convexities (white) and concavities (grey).

Since the boundary segment of the most significant tree crown can belong to one of the other tree crowns, and not the one the seed point represents, the seed point must be moved. The location of the new seed point, called starting point, is calculated as the median of the curvature values of the pixels belonging to the current boundary segment. The median is also used to calculate the σ_2-value in the exponential function,

$$\mu = e^{-\frac{1}{2}((i_d^2+r_d^2+g_d^2)/\sigma_1^2+(x_d^2+y_d^2)/\sigma_2^2)}, \tag{6}$$

which is the similarity measure used when the regions grow. Here, (x_d, y_d) is the difference in location between the starting point and the pixel under consideration, (i_d, r_d, g_d) is the colour difference between the starting point and the pixel, and σ_1 is a constant. The pixel under consideration is included in the region if μ is larger than a fixed value T and the maximum size of the region is determined by σ_1 and σ_2. If the resulting region consists of only a few pixels it is not considered a tree crown at all. Otherwise, the resulting region is only used as an approximation of the tree crown, since the agreement with the tree crown in the original image may not be the best possible. To find regions with better agreement, all pixels included in the approximation are used as starting points in Eq. 6. This creates many candidate regions. To select the region among all candidates that represents the tree crown best, the mean value of the grey levels along the border of the region is calculated.

The smaller the mean value is, the better the candidate. Recall the valley assumption and Fig. 2. Thus, if the border of a segment is located at the boundary of the tree crown, the border of the segment will contain many dark pixels and thus have a small mean value.

Finally, seed points located inside the segment are removed, see Fig. 16, which also shows the segment removed from the original image in order to show how the segment fits the tree crown. Since only the seed points are updated when a segment is found and not the original image, it is possible to have overlapping segments which is shown in the final result in Fig. 17.
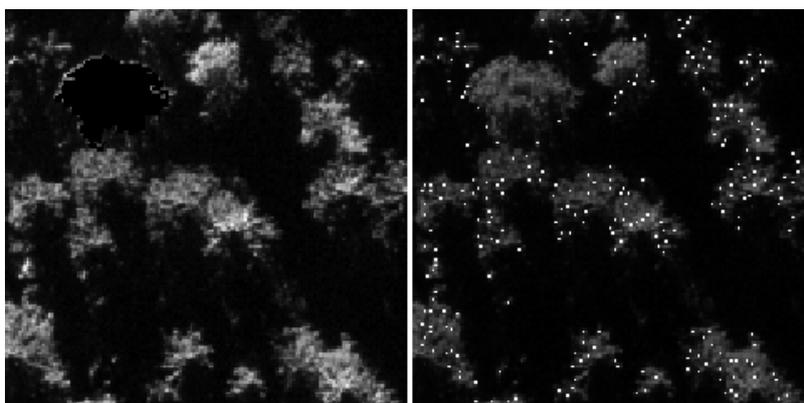

Figure 16: Left: Segment (upper left) removed in order to show the fitting between the segment and the tree crown. Right: Seed points removed.
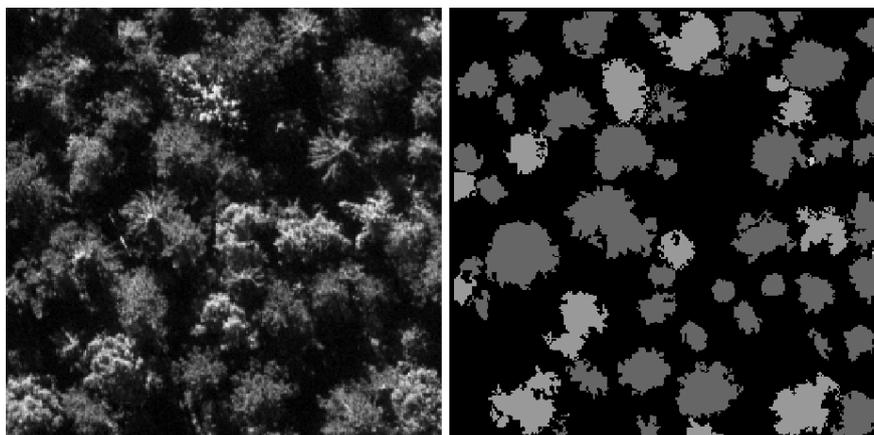

Figure 17: Result from segmentation by fuzzy rules. Left: Original image with 10 cm resolution. Right: Segmentation result in different grey levels with overlapping parts shown as small bright regions.

4.1.2 Segmentation via Brownian motion

As pointed out earlier, the shape of the tree crown should be kept in the segments. However, visually, the result from the segmentation method supported by fuzzy rules is not completely satisfactory in this respect, thus a new segmentation method was developed. Furthermore, this method was developed to preserve, or even enhance, the large structure in the tree crowns. The structure was supposed to be used in the classification step. The assumption of the method is, again, that there exist dark pixels between the tree crowns. The method is based on a phenomenon called Brownian motion. Brownian motion is the irregular continuous path of a particle moving in a liquid. To simulate a rough estimate of the motion in 2D, normally distributed random 2D-vectors are summed. The location of the particle after n steps is the sum of the first n random vectors.

If there is no restriction on where the particle can move, except that the particle must stay in the image, the particle will eventually reach all positions in the image. In order to force the particle to stay inside the tree crowns in the image, the elements of the normally distributed vectors are multiplied by a weighting factor, before the vector is added to the old position. The weighting factor is proportional to the grey level at the position it should enter if the weighting factor is one. Thus, if p_{i-1} is the location of the particle at step $i-1$, the location at step i is given by

$$p_i = p_{i-1} + g\left(f\left(p_{i-1} + v\right)\right) v, \tag{7}$$

where v is the normally distributed random vector and g is a function mapping the grey level values of the image f onto the real interval $[0, 1]$. This means that if the particle is heading towards a dark pixel, the step will be smaller than if it is heading towards a bright pixel. Fig. 18 illustrates the use of the weighting factor for two cases. Note that the positions will not necessarily be integers.
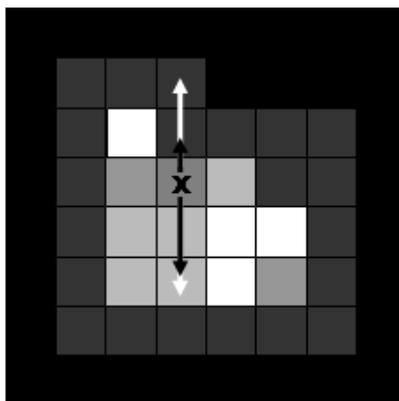

Figure 18: Illustration of the weighting factor. **x** is the current position of the particle. The two white arrows are possible moves while the black arrows are the actual moves, respectively.

Since the particle must start somewhere in the image, at least one seed point is needed. By counting the number of times the particle visits a pixel, a new grey level image, *numPass*, is created. The pixels are found by truncating the positions. The *numPass*-image can be seen as a semi-thresholded image, where the background pixels are put to zero (the pixels the particle has not visited) and the object pixels have grey levels (in this case according to the visiting number rather than from the original image). This new image is used instead of the original image in the rest of the segmentation.

To be sure that the particle visits all tree crowns, one seed for each tree crown is preferable. Thus, in order to find the starting points, the same technique as for the method supported by fuzzy rules is used. But in this case, the distance image is smoothed with a Gaussian filter before local maxima are detected. If the size of the filter is equal to the size of a mean crown, this reduces the number of seed points to one in almost every tree crown.

The next step is to initialize the final segmented image, *seg*, by letting each seed point have a unique label. In order to ensure that the segment will have an area, each seed point is expanded to a circle with the same label. The radius of the circle is proportional to the distance value at the location of the seed point, where the distance image is found from the *numPass*-image. Then, all pixels with nonzero values in the *numPass*-image, which is also a border pixel to only *one* region in *seg*, is added to a list of pixels. Pixels that are border pixels to more regions creates dividing lines between the regions. The pixel with largest value in the list is, added to *seg* with the label equal to the label of the corresponding border pixel. The list of border pixels is updated by removing the pixel and by including the new neighbouring pixels. The largest value in the new list is added to *seg* and so on, until the list of pixels is empty. The result from the method is shown in Fig. 19.
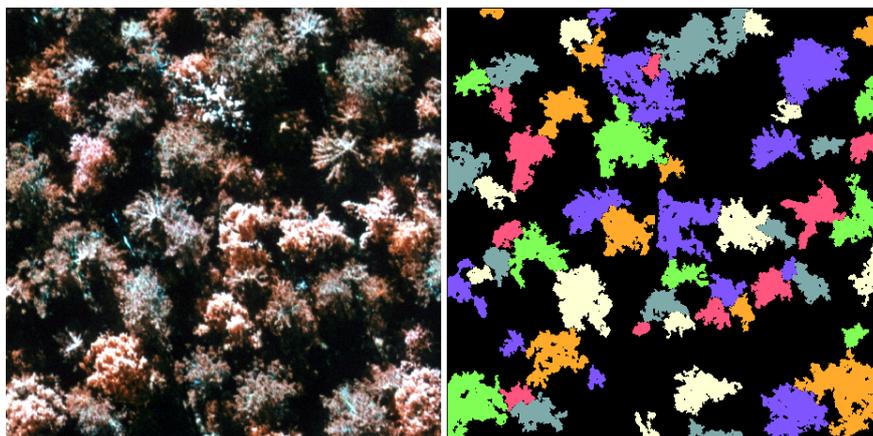

Figure 19: Result from segmentation via Brownian motion. Left: Original image with a pixel size corresponding to 3 cm on the ground. Right: Segmentation result.

4.1.3 Segmentation via random walk

Although the result from the segmentation method via Brownian motion finds the shape of the tree crown much better than the method supported by fuzzy rules, there are still some errors to be fixed. Again, it turns out that segmentation via Brownian motion does not visually find a good dividing line between all connected objects, see for instance the spruce in Fig. 20 which is located at the upper border of the image in Fig. 19.
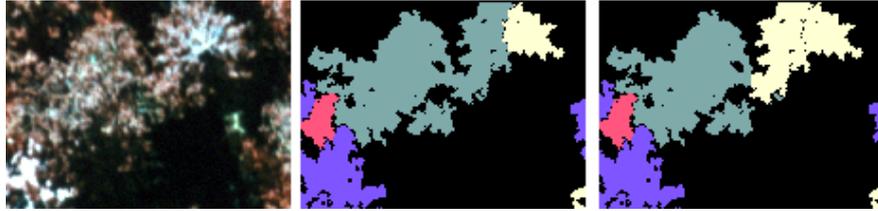

Figure 20: The middle image shows the separation between the spruce and the pine from segmentation via Brownian motion. However, the right image shows a more natural separation between the two crowns.

In order to find better dividing lines between the tree crowns, the segmentation method via Brownian motion is slightly modified to create the segmentation method via random walk. The same assumption with dark pixels, or valleys, between different crowns is used. However, this method tries to enhance the contrast between these valleys and the tree crowns. To explain the differences, the segmentation method via Brownian motion is divided into a preprocessing step, the creation of *numPass*, and an expansion step. This method uses the same expansion step but modifies the preprocessing step. The first modification is to use uniformly distributed random vectors instead of normally distributed random vectors and only take integer steps. This is why the name is changed from Brownian motion to random walk. The second change is to find a neighbourhood representing the positions the particle can move to in a single step. The intended new position is found by adding a uniformly distributed random vector, randomized inside the neighbourhood, to the current position of the particle. If the step is taken or not, depends on a new random number, z, randomized uniformly between zero and the maximum grey level in the neighbourhood. If z is less than the grey level at the position the particle is trying to reach, the step is taken. Otherwise, a new position must be found in the same manner. This is similar to the multiplicative factor used in Brownian motion in order to prevent the particle to enter the background. The largest change, though, is to vary the size of the neighbourhood depending on the sum of the included grey levels in the neighbourhood, thus taking care of the geometrical thickness of the object. The size, ν, of the neighbourhood, N_ν, is found by

$$\min_{\nu} \sum_{q \in N_\nu} f(q) > T, \tag{8}$$

for some threshold T and where $f(q)$ is the grey level at position q. The threshold

T is a number between one and the sum of all included grey levels in the image. If T is small, the resulting image, which counts the number of times the particle visit a pixel, is similar to the original image (but with different dynamic range). On the other hand, if T is large, a pixel becomes darker the closer a border of an object it is, see Fig. 21. The effect of the darkness is proportional to the size of the neighbourhood.
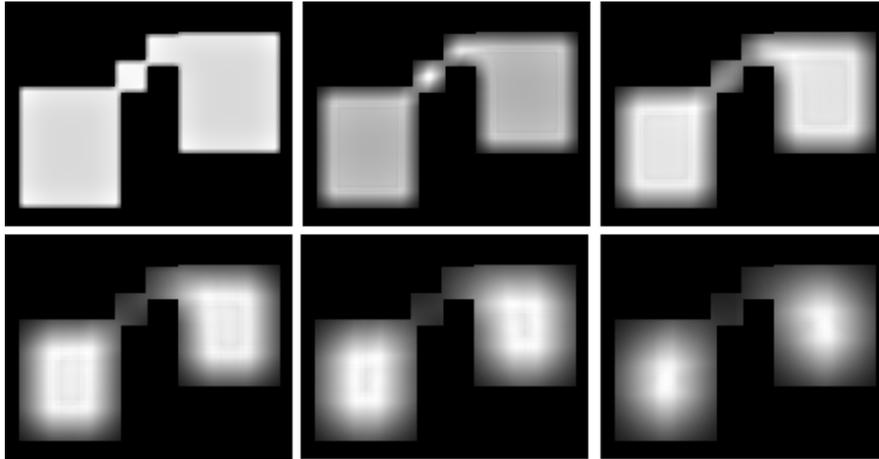

Figure 21: A binary test image with different T. Upper left: $T = 9$. Upper middle: $T = 100$. Upper right: $T = 225$. Lower left: $T = 400$. Lower middle: $T = 625$. Lower right: $T = 900$. Each image is divided by its maximum value in order to be displayed.

At the first sight, the resulting images are very similar to the result using a distance transform. However, there are three major differences. The first is that the grey levels in these images are not distance values as in the distance transform. The second is that this method is insensible to noise, which is not true for the distance transform. The third difference is that the method is not limited to binary images. As a matter of fact, the method is developed for grey level images.

The resulting preprocessed image is used as the *numPass*-image in Section 4.1.2 and the same expansion is performed to find the segments. Thus, the final segmented image, *seg*, is initialized with the seed points having unique labels which are expanded to be circles with radii proportional to the distance value at the respective seed point location. Then, all pixels with nonzero values in the *numPass*-image which is also a border pixel to only *one* region in *seg*, is added to a list of pixels. Other pixels creates dividing lines. The pixel with largest value in the list, is added to *seg* with the label equal to the label of the corresponding border pixel. The list of border pixels is updated and the largest value in the new list is added to *seg* and so on, until the list of pixels is empty. The segmentation result from the method is shown in Fig. 22. Fig. 23 shows the close up corresponding to Fig. 20. The seed points are the same seed points as used in Fig. 19.
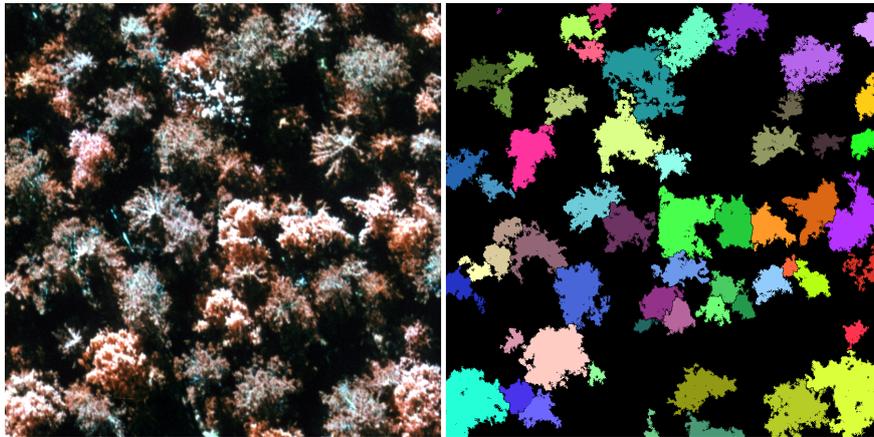
Figure 22: Result from segmentation via random walk. Left: The original image with a pixel size corresponding to 3 cm on the ground. Right: The segmentation result.

Figure 23: The same pine and spruce as in Fig. 20 but with a more natural dividing line, achieved by random walk.

4.2 Evaluation of segmentation results

To evaluate a segmentation method, a reference data set with regions representing the tree crowns are used as the truth. The reference data are manually delineated with or without field measurements as a complement to the aerial images. The evaluation can be done in many different ways and no golden standard exists. One technique is to tabulate the number of segments per delineated reference crown. Another technique is to count the number of pixels that belongs to both a segment and a region in the reference delineation. Both techniques have been used to evaluate the segmentation methods presented here. The details for each evaluation are given in the following subsections.

4.2.1 Segmentation supported by fuzzy rules

The segmentation method supported by fuzzy rules is evaluated against a manual delineated reference data set by counting the number of segments per reference crown using the following rules:

30

- A segment with less than 50 % coverage of the reference delineation is interpreted as missing. This means that the number of missing trees is increased by one.

- A segment is also missing if the segment covers two or more reference segments but none of them are covered completely. This means that the number of missing trees is increased by the number of reference segments.

- If a segment covers one reference segment with more than 50 % and less than 25 % of another reference segment, the segment is correct. Thus, the number of correct trees is increased by one.

Almost 9000 visible tree crowns from 30 randomly selected images from Huljen test site, except the images used for developing the algorithm, were used in the evaluation with more than 6500 of them correctly found, according to the chosen criteria. This means that 73 % of the visible tree crowns were correctly found. Table 1 shows the result from the evaluation, tabulated in seven categories. Note that as we have both "one crown segmented into several" and " several crowns segmented into one" errors, the total number of trees will be better determined than the single ones.

Table 1: Summary of the result from 30 images. The category A:B means the segment was divided into B parts by the program but in A parts by the interpreter. The category 1:1 is the correctly segmented crowns. The total number of visible crowns was counted from the subimages

	original:segment (manual:computer)							
	1:1	1:2	1:3	2:1	3:1	1:0	Others	Total
No. of crowns	6554	426	171	331	60	329	1083	8954

4.2.2 Comparison between different segmentation methods

In order to find strengths and weaknesses of different segmentation methods, they must be evaluated and compared using the same image material. Therefore, the segmentation method supported by fuzzy rules, segmentation via Brownian motion, and segmentation via random walk was evaluated against the template matching method in Olofsson (2002), using the image material from Remningstorp. Recall from Section 1.1.3 how the templates are generated. The material consists of four images covering two different pine stands and two different spruce stands.

The four methods were compared with reference data that was manually delineated using field data. A segment with more than 50 % covering of the delineated reference crown was considered a correct match, otherwise it was considered a false match. The number of correct matches and the number of false matches are counted and presented in Table 2, together with the percentage of both correctly segmented visible tree crowns and of all segments against the visible tree crowns.

Segmentation via Brownian motion and segmentation via random walk performed best of all methods for all stands, with most correct and least false matches. These two methods had identical results since the seed points are found in the same way. False matches with these methods were due to extra seed points and missing trees due to lack of seed point in the crown. Template matching performed better than segmentation supported by fuzzy rules for the pine stands, mostly due to less false matches but also due to a few more correct matches. For the spruce stands they were equally good. False matches for segmentation supported by fuzzy rules often came from multiple segments for a delineated crown, probably as a consequence of bad estimation of the size of the tree crown. Missing trees were also due to bad estimation of the size of the crown. For template matching, the false matches were either matches in the background or the criterion of 50 % covering was not met. The trees were the covering condition was not met were smaller than the smallest radius used when constructing the templates and can therefore not be expected to be found.

Table 2: Comparison between segmentation supported by fuzzy rules (Fuzzy), segmentation via Brownian motion (BM), segmentation via random walk (RW), and template matching (TM). Correct - correct match between segment and tree crowns. False - segment not representing a tree crown. Total - the total number of segments. The number of visible tree crowns in the image is 164 and the total number of trees from the field measurements is 202

	#Correct	#False	#Total	Segmented (%) Correct	Total
Fuzzy	131	29	160	79.9	97.6
BM	156	8	164	95.1	100.0
RW	156	8	164	95.1	100.0
TM	135	21	156	82.3	95.1

4.2.3 Tiputini Biodiversity Station

To further evaluate the segmentation method using random walk, the image material from the Tiputini Biodiversity Station is used. In this case, manually defined seed points are used since the forest is very dense. In addition, a modification in the expansion step is made. Recall the ordinary method from Section 4.1.3. The modified expansion step works in the following way: The final segmented image, *seg*, is initialized with the seed points having unique labels which are expanded to be circles with radii proportional to the distance value at the respective seed point location. Then, all pixels with nonzero values in the *numPas*-image, found from the preprocessing step, which is also a border pixel to only *one* region in *seg*, is added to a list of pixels. Other pixels are omitted and creates dividing lines between the segments.The pixel with largest value in the list *and* for which the intensity value for the red and green band, taken from the original image, differs less than T_{exp} from the respective mean intensity, is added to *seg* with the label equal to the label of the corresponding border pixel. The mean intensity for each band is calculated as the mean of the grey levels of the pixels included in the expanded circle around the seed point. The grey levels are from the respective band of the original image. The

list of border pixels is updated and the largest value in the new list is added to *seg* and so on, until the list of pixels is empty. The expansion step is repeated once but with a larger T_{exp}-value than the first time and starting from the result in *seg*. The first time the expansion step is called, $T_{exp} = 0.1$, and the second time, $T_{exp} = 0.5$. The reason for repeating the expansion step is to let all segments have a chance to grow.

No images have yet been evaluated, however, Figs. 24, 25, and 26 shows segmentation results from 200 m, 300 m, and 600 m, respectively.
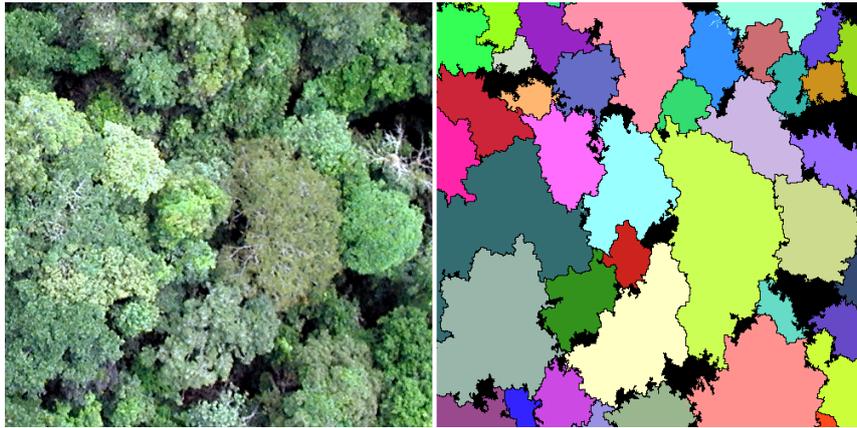

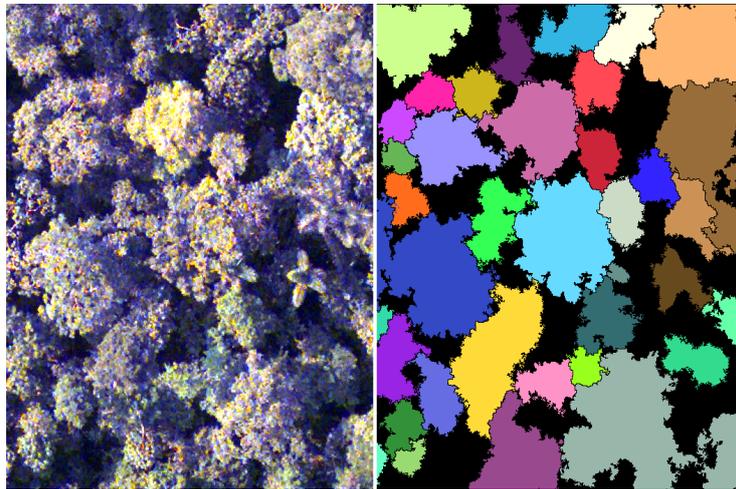Figure 24: Segmentation result from TBS, 200 m. Left: Original image. Right: Segmented image.

Figure 25: Segmentation result from TBS, 300 m. Left: Original image. Right: Segmented image.
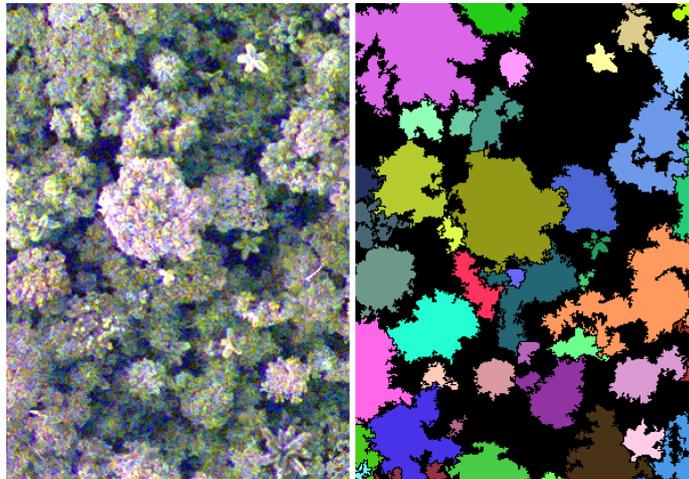
33

Figure 26: Segmentation result from TBS, 600 m. Left: Original image. Right: Segmented image.

4.3 Classification

Classification in this thesis means to analyse each segment in the segmented image in order to find the tree species it represents. As already mentioned, the better the segmentation is, the better the analysis can become. Here, segmentation via Brownian motion is the method used to separate the tree crowns, since segmentation via random walk was not developed at the time. Fig. 27 shows typical examples of the four most commonly occurring tree species in Sweden, birch, aspen, spruce, and pine. Aspen is less common than the others, but still frequent.

Figure 27: The four most common tree species in Sweden. From left to right: birch, aspen, spruce, and pine.

As already pointed out in the introduction, separation between deciduous and conifers can in many cases be done rather easily using CIR images, simply because deciduous trees reflect a lot more infrared light than conifers. This is quite evident from Fig. 27 where the two leftmost are deciduous and the two rightmost are conifers. Another observation that can be made from the figure is that the crown of the aspen consists of a lot of small bright "blobs" with approximately the same size.

34

The last observation from the figure is that the tree top is clearly visible in the crown of a spruce and a large part below the tree top is shadowed. All these observations are used to construct the measures used in the classification.

The first measure finds birch by counting the number of bright pixels and "reddish" pixels in the segment and compare them with the total number of pixels in the segment. The second measure finds aspen and compares the bright pixels with the total number of pixels in the segment and also the number of "blobs" at two thresholds. Measure three finds spruce by locating the largest concavity in the segment below the centre of mass. This concavity must be located as in Fig. 28 due to the known position of the sun. It also measures a distance between the concavity and the tree top, defined as the largest region in the segment when thresholding the original image at the maximum intensity in the segment. The threshold must be above a certain value in order to be a tree top, otherwise the tree top is missing. The forth measure finds pine and is a complement to measure three, which means a wrongly located concavity or that the tree top is missing.

Figure 28: A correctly located spruce concavity in light grey. Segment in dark grey.

In order to perform the classification, each species corresponds to one of the above described measures. Instead of calculating all measures for each segment to create a feature vector and use a classifier, only one measure at a time is calculated. This measure compares with the criteria for the species the measure corresponds to. If the measure fulfills the criteria, the segment is supposed to represent the corresponding species. If the criteria are not fulfilled, the next measure is calculated and compared and so on. Some of the segments do not fulfil any of the criteria of the four measures and are therefore classified as the species they are closest to.

Since only one measure at a time is calculated, the order becomes important. The order is birch, aspen, spruce, and pine for two reasons. The first reason is that, as pointed out earlier, deciduous are rather easy to distinguish from conifers. Therefore, it is natural to start with one of these two classes. The second reason is that a segment representing a birch can easily be classified as a spruce if the measure corresponding to birch succeeds the measure corresponding to spruce.

The method is developed using two images from the Huljen test site and evaluated using 14 other randomly chosen images from the same test site. The overall result is given in Table 3. The rows represent the manually interpreted species and the columns represent the computer classified species. Each cell holds the number of segments that are classified as belonging to each respective species. Table 4 shows the same result as distribution of the four species.

Table 3: The number of classified segments for each species. The total number of trees is 791

Interpreted as	Classified as				
	spruce	pine	birch	aspen	% correct
spruce	**257**	61	19	6	74.9
pine	46	**148**	1	-	75.9
birch	30	17	**192**	-	80.3
aspen	2	-	2	**10**	71.4
% correct	76.7	65.5	89.7	62.5	**76.7**

Table 4: The estimated distribution from the classification and the interpreted distribution for each species

Distribution	spruce	pine	birch	aspen	Total
Interpreted (%)	43	25	30	2	100
Estimated (%)	42	29	27	2	100

5 Discussion

Segmentation is a fundamental problem in image analysis and is often very difficult. Individual tree crown segmentation is no exception. The three segmentation methods developed in this thesis focus on the agreement of the segments with the tree crowns by using region growing algorithms. However, the methods concentrate mostly on the question about how the regions grow and not very much about how to find the seed points, a problem which is still to be solved in an appropriate way. This far, a simple and quick technique using Gaussian filters is used to find the seeds, which works fairly well if the forest is not too dense. However, if the forest *is* dense, some other technique must be used. Here, manually found seed points are used in this case.

In the comparison of the three segmentation methods and the template matching method, the forest, Remningstorp estate, is almost an ideal forest with respect to density. It is neither dense nor sparse, since many of the tree crowns are standing fairly alone but not distanced so far away from each other to make the ground visible. A visible ground can create problems since these patches have similar properties as tree crowns. Also, the stands in the test site contain one species almost exclusively. These facts can explain why the segmentation method supported by fuzzy rules performs better on the images from Remningstorp than from Huljen, since it is easier to estimate the size of the crown in the former case. However, the difference in evaluation technique is also relevant. The situation in Fig. 29 is interpreted as missing for both segments by the technique used in Section 4.2.1, but as one correct segment by the technique used in Section 4.2.2. Thus, the latter technique is more generous than the former. This means that if the same technique is used for both image materials, the results will differ less.

Figure 29: Using different evaluation techniques, the result can be interpreted as both segments missing or as one correct segment. In the right image, the small bright region is overlap between the two segments.

In the comparison, segmentation via random walk and segmentation via Brownian motion gave identical results when the number of correctly found tree crowns was used as evaluation. This results from the fact that they find the seed points in the same way. Thus, the only difference between the two methods is the dividing lines between the tree crowns. However, with the evaluation methodology presented in Brandtberg et al. (2003), a larger difference between the two methods would probably appear. That evaluation technique is based on the idea that the reference delineation of the tree crown is more accurate at the centre of the crown and

becomes fuzzier the closer the border it gets. Thus, by finding the distance value for each pixel to the closest border pixel, using a distance transform, weights for the accuracy are found. A similar procedure is performed on the segment and the weights are compared to find an assessment value. The assessment value is the ratio between the sum of a maximum and a minimum image. These two images are found by overlaying the reference image and segmented image to find the maximum and minimum, respectively, for each pixel. However, if the segment is slightly shifted compared to the reference region, the assessment value for the above described technique has a larger error than the actual amount of wrong pixels. For instance, the two circles in Fig. 30 A) are slightly shifted with respect to each other. The reference circle, R, is shown in white together with the left arc segment, i.e. $R = C_1 \cup A_1$, and the segment circle, S, is shown in white together with the right arc segment, i.e. $S = C_1 \cup A_2$. The assessment value in this case is 85.8 % but the false pixels are only 9.6 %. The reason that the assessment value is not higher is that there are errors also in the intersection of the two circles and not only in the two distinct arcs, see Fig 30 D), which shows the difference between the maximum and minimum image. One way to overcome this is to use the maximum distance value in the reference region as a centre for both the reference region and the segment and fit a monotonically decreasing function around this centre, for instance, by using a Gaussian function. The problem then is to decide a proper function and how it should vary depending on the size of the reference region.
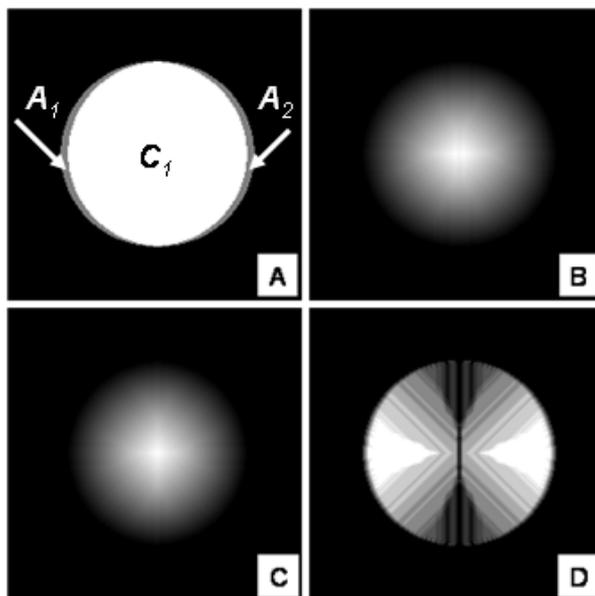

Figure 30: A) Reference circle, R, in white together with left arc segment in grey, i.e. $R = C_1 \cup A_1$ and segment circle, S, in white together with right arc segment in grey, i.e. $S = C_1 \cup A_2$. B) Maximum image. C) Minimum image. D) Difference between maximum and minimum image, scaled with the maximum difference.

The reason for getting visually better results using segmentation via random walk than using segmentation via Brownian motion is the use of different sizes of the neighbourhoods for segmentation via random walk. The change of the size of the neighbourhood, depending on the location of the pixel in the object, creates the dark effect near the border of the object as shown in Fig. 21. Thus, this effect enhances the contrast between different tree crowns, as the method was supposed to do. Since the regions grow in grey level order from the preprocessed image, this means that the valleys between the tree crowns are the last to be added to the regions. Thus, the valleys creates the dividing lines between the tree crowns.

All classification results are dependent on the segmentation result. If the segmentation is wrong, it introduces errors in the classification step since the wrongly segmented tree crowns can not be treated as they should. Thus, a perfect segmentation is needed in order to conclude how good the classification method is. However, perfectly segmented tree crowns are impossible to achieve. Manually segmented tree crowns are the closest we can get. However, using manually segmented tree crowns will only show how the classification works in theory and not how it works in practice. This is the reason for not using manually segmented tree crowns when classifying the crowns into species. However, a comparison using both manual and computer segmented tree crowns would have been preferable.

Since the classification method is developed for CIR images, the method may not work well for RGB images. The method is also dependent on the position of the sun when the images are captured due to the location of the concavity in the spruce measure. However, by rotating the image the latter problem can be overcome. In order to be able to use the method operationally, the thresholds must be calculated automatically instead of having them as fixed numbers.

In the classification step, shape and colour information of the segment is used in order to find the species. As mentioned in the introduction, texture is another information that can be useful in classification. Texture describes properties such as smoothness, coarseness and regularities of the tree crown (Gonzalez & Woods, 1992). Looking at Fig. 27, texture seems to be a good idea for separating spruce and pine. There are three principal approaches for texture. These approaches find statistical, structural, and spectral properties, respectively. The last one searches for periodicities in the texture, which apparently does not exist in this case. The structural approach deals with arrangements of image primitives such as regularly spaced lines etc, which do not exist either. Thus, the only approach that can be useful here is the statistical approach which either uses moments of the grey level histogram of the tree crown or a co-occurrence matrix of the grey levels to describe the statistical behaviour of the texture. The element c_{ij} in the co-occurrence matrix is the number of times the grey level z_i is located in a certain position with respect to the grey level z_j. Different relations in locations are used in order to make the statistics. Fig. 31 shows subimages of spruce in the upper row and of pine in the lower row. From these subimages, there is no obvious difference between the two species. Thus, in a local scale, the grey levels are more or less random and therefore neither moments of the histogram nor the co-occurrence matrix can be used. This means that if texture is to be used, large structures in the tree crowns must be found.

Figure 31: Subimages of twelve different tree crowns. Each subimage is 40×40 pixels or 1.2×1.2 m. Upper row: Tree crowns from spruce. Lower row: Tree crowns from pine.

An attempt to preserve or even enhance the large structures in the tree crowns was made using the *numPass*-image found by segmentation via Brownian motion. However, no useful feature that could separate between the species could be found from these structures.

6 Conclusions and future research

Automatic derivation of forest parameters from aerial photographs has a promising future. This thesis shows that it is possible to identify individual tree crowns in high spatial resolution images with good results. Table 2 shows that really good results can be achieved for certain types of forests while Table 1 shows that the result is good for more complex forests as well. The thesis also shows that further analysis of the segments, such as tree species classification, is possible by using colour information of the corresponding tree crown and shape information of the segment in the classification step. However, there is a lot of work left to be done in the future.

Perhaps the most interesting subject for future research is to combine the aerial images with laser data, not only to use laser data to measure the height of the trees but also to improve segmentation results, for example, by using laser data to find the seed points.

Of course, improvements of the methods presented here are also a possible way for future research. As already mentioned, the seed points are found in a primitive and simple way. If an automatic method is the objective, this is a step that must be improved if the forest is complex. The preprocessing step with random walk is a one way to create the image used in the expansion step. However, in some cases this image is not enough. Therefore, the expansion step can be improved to also incorporate other features such as distance to the seed point or colour differences. The colour difference is incorporated when using the images from Tiputini Biodiversity Station.

Combination of different methods is also a possible way to improve the methods. Perhaps the most interesting combination is to combine segmentation via random walk with the method presented in Section 1.1.2 which uses scale-space theory. Here, the last method can be used to locate the tree crowns, i.e. finding seed points, and the former when the regions are growing from these seed points.

One comparison of different segmentation methods is included in this work. However, this is done for only one forest and between four segmentation methods. Future research should evaluate all reasonable segmentation methods against multiple image materials in order to conclude which method is most suitable for a particular forest type. This is a huge project, which involves a lot of people. In order to be able to compare different methods, a golden standard for evaluation must be agreed on.

The classification method has one obvious weakness, the measure for pine, which does not search for a feature representing a pine. Instead, it is based on the complement of the feature representing a spruce. If it is possible to find a feature for pine which is also implementable, the classification step is most likely improved. It is also possible that the classification result becomes better if all measures are found simultaneously and then deciding which class the segment belongs to. Another interesting continuation of the classification method is to include other species as well.

References

Anon., (1993). *Flygbildsteknik och Fjärranalys*. Nämnden för Skoglig Fjärranalys, Skogsstyrelsen, SE-551 83 Jönköping, 1993, 430 p. (In Swedish.)

Borgefors, G. (1986). Distance Transformations in Digital Images, *Computer Vision, Graphics, and Image Processing*, 34, 1986, pp. 344-371.

Brandtberg, T. (1999) *Automatic Individual Tree-Based Analysis of High Spatial Resolution Remotely Sensed Data*, Doctoral Thesis, Acta Universitatis Agriculturae Sueciae, Silvestra 118, SLU Service/Repro, Uppsala, 1999, 155 p.

Brandtberg, T., Walter, F. (1998). Automatic delineation of individual tree crowns in high spatial resolution aerial images by multiple-scale analysis, *Machine Vision and Applications*, 11(2), 1998, pp. 64-73.

Brandtberg, T., Warner, T.A., Landenberger, R.E., McGraw, J.B. (2003), Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in North America, *Remote Sensing of Environment*, 85, 2003, pp. 290-303.

Chi, Z., Yan, H., and Pham, T. (1996), *Fuzzy algorithms: with applications to image processing and pattern recognition*, World scientific publishing Co. Pte Ltd., Singapore. 1996. 225 p. (Chapter 3.4)

Culvenor, D.S. (2002), TIDA: an algorithm for delineation of tree crowns in high spatial resolution remotely sensed imagery, *Computers & Geoscience*, 28(1), 2002, pp. 33-44.

Gonzalez, R.C., Woods R.E. (1992), *Digital Image Processing*, Addison-Wesley Publishing Company, Reading, Massachusetts, USA, 1992. 716 p.

Gougeon, F.A. (1995). A crown-following Approach to the Automatic Delineation of Individual Tree Crowns in High Spatial Resolution Aerial Images, *Canadian Journal of Remote Sensing* 21(3), 1995, pp. 274-284.

Gougeon, F.A. (1998). Automatic individual tree crown delineation using a valley-following algorithm and a rule-based system, *In* Proceedings of the the International Forum on Automated Interpretation of High Spatial Resolution Digital Imagery for Forestry (eds. D.A. Hill and D.G. Leckie), February 10-12, 1998. Natural Resources Canada / Canadian Forest Service, Pacific Forestry Centre. Victoria, British Columbia, Canada, pp. 11-23.

Holmgren, J., Persson, Å. (2004). Identifying species of individual trees using airborne laser scanner, *Remote Sensing of Environment*, 90, 2004, pp. 415-423.

Howard, John A. (1991). *Remote Sensing of Forest Resources, theory and application*, Chapman & Hall, London, England, 1991. 422 p.

Jakobsons, A., (1970). Sambandet mellan trädkronans diameter och andra träd-faktorer, främst brösthöjdsdiametern. (Summary: Relationship between crown diameter and diameter at breast height.) Rapport 14, Inst. f. skogstaxering, Skogshögskolan, Stockholm. (In Swedish with English summary.)

Larsen, M., Rudemo, M. (1998). Optimizing templates for finding trees in aerial photographs, *Pattern Recognition Letters*, 19, 1998, pp. 1153-1162.

Lindeberg, T. (1994). *Scale-space Theory in Computer Vision*, Kluwer Academic Publisher, Dordrecht, The Nederlands, 1994. 423 p.

Lindeberg, T. (1998) Feature Detection with Automatic Scale Selection, *International Journal of Computer Vision* 30(2), 1998. pp. 79-116.

Minor, C.O. (1951). Stem-crown diameter relations in souther pine. *Journal of Forestry*, 49(7), 1951, pp. 490-493.

Olofsson, K. (2002). Detection of Single Trees in Aerial Images using Template Matching. In ForestSat 2002, Operational tools in forestry using remote sensing techniques. Proceedings CD-ROM, talk FI6.3, session Forestry Inventory 6, Monitoring Forest Establishment and Development, Forest Research, Forest Commission, Heriot Watt University, Edinburgh, Scotland, August 5-9, 2002.

Persson, Å., Holmgren, J., Söderman, U. (2002). Detection and Measuring Individual Trees Using an Airborne Laser Scanner. *Photogrammetric Engineering & Remote Sensing*, 68(9), 2002, pp. 925-932.

Pinz, A. (1989). Final results of the vision expert system VES: Finding trees in aerial photographs. In: Pinz, A. (ed): Wissensbasierte Masterkennung (Knowledge-based pattern recognition), OCG-Schriftenreiche 49, Oldenbourg Verlag, pp. 90-111.

Pollock, R.J. (1996). *The automatic Recognition of Individual trees in Aerial Images of Forests Based on a Synthetic Tree Crown Image Model*, Doctoral Thesis, University of British Colombia, Vancouver, Canada. 1996, 172 p.

Pouliot, D.A., King, D.J., Bell, F.W., Pitt, D.G. (2002). Automated tree crown detection and delineation in high-resolution digital camera imagery of coniferous forest regeneration, *Remote Sensing of Environment*, 82, 2002, pp. 322-334.

Walter, F. (1998). Fjärranalys för skoglig plannering. (Summary: Remote sensing for forestry planning.), The Forest Research Institute of Sweden, Science Park, SE-751 83 Uppsala, Sweden, Report no. 9, 1998, 40 p. (In Swedish with English summary.)

Zadeh, L.A. (1965). Fuzzy Sets, *Inform. Control*, Vol. 8, 1965, pp. 338-353.

Other publications and conferences

The author has also been author or co-author of the following papers.

- Erikson, M. Structure-keeping Colour Segmentation of Tree Crowns in Aerial Images, in Proc. of the 12th Scandinavian Conference on Image Analysis (ed. I. Austvoll), 11-14 June, 2001, Bergen, Norway, Norwegian Society for Image Processing and Pattern Recognition, pp. 185-191.

- Erikson, M. Colour Segmentation of Individual Tree Crowns in Aerial Images, in Proc. of SSAB'02 Symposium (ed. K. Åström), 7-8 March, 2002, Lund, Sweden, pp. 177-180.

- Erikson, M., Vestlund, K., Finding tree-stems in laser images of young mixed stands to perform selective cleaning, in Proc. of the ScandLaser Scientific Workshop on Airborne Laser Scanning of Forests (eds. J. Hyyppä, E. Næsset, H. Olsson, T. Granqvist, and H. Reese), 3-4 September, Umeå, Sweden, 2003, pp. 244-250.

- Bengtsson, E., Erikson M. (eds.), Proceedings of the SSBA'04 Symposium on Image Analysis, 11-12 March, 2004, Uppsala, Sweden, 202 p.

The following international conferences were attended.

- 12th Scandinavian Conference on Image Analysis (SCIA'01), Bergen, Norway, 11-14 June, 2001 (Oral presentation)

- 13th Scandinavian Conference on Image Analysis (SCIA'03), Gothenburg, Sweden, 29 June - 2 July, 2003. (Poster presentation)

- ScandLaser Scientific Workshop on Airborne Laser Scanning of Forests, Umeå, 3-4 September, 2003. (Poster presentation)

Acknowledgments

Writing a thesis involves a lot of people. Some of them have contributed to the actual writing of the thesis or the appended papers, while others to the life aside the research. All of them are important and gratefully thanked. Especially, the following persons are thanked:

Börje and Barbro - my loving parents, who always been there for me.

Björn, Liza, Frida, Ulf, and Lonneke - my dear brothers and their families.

Ola - for being like a roommate during the years at CBA.

All biking friends and other friends

Gunilla - my supervisor.

Ola, Xavier, Gunilla, and Mattias - the proof-readers, for valuable comments.

Olle - for the help with computer problems.

Lena - for the help with all other problems.

Lunchgänget - for the nice company during the lunches.

All CBA-employees during the years - for the great atmosphere to be in.

Karin, Kenneth, and Pablo - my collaborators.

Tack alla!

Mats

Uppsala, October 2004.