METSÄNTUTKIMUSLAITOKSEN TIEDONANTOJA 926, 2004
FINNISH FOREST RESEARCH INSTITUTE, RESEARCH PAPERS 926, 2004

Image segmentation in multi-source forest inventory

Anssi Pekkarinen

Academic Dissertation

VANTAAN TUTKIMUSKESKUS – VANTAA RESEARCH CENTRE

Pekkarinen, Anssi. 2004. Image segmentation in multi-source forest inventory. Finnish Forest Research Institute. Research papers 926. ISBN-951-40-1929-6, ISBN 952-10-2003-2 (pdf). 35 p. + 4 original papers.

Supervisor:	Professor Erkki Tomppo Finnish Forest Research Institute National Forest Inventory
Pre-examiners:	Professor Alan Ek Department of Forest Resources University of Minnesota, USA
	Professor Håkan Olsson Department of Forest Resource Managements and Geomatics SLU, Umeå, Sweden
Opponent:	Ph.D. Tuomas Häme Chief Research Scientist VTT Information technology Remote sensing group Espoo, Finnland
Publisher:	Finnish Forest Research Institute, Vantaa Research Centre, P.O. Box 18, FI-01301 Vantaa, Finland. Accepted by Kari Mielikäinen, Research Director, June, 2004.
Author :	Anssi Pekkarinen Finnish Forest Research Institute National Forest Inventory Unioninkatu 40 A, FI-00170 Helsinki

Front cover: AISA image: ©METLA/NFI

Printed in Dark Oy, Vantaa 2004.

Contents

Abstract

This thesis examines whether the applicability of remote sensing aided forest inventory methods can be improved by using the image segment -based approach to feature extraction and image analysis. Several image segmentation techniques are developed, implemented, and tested in forest inventory applications employing both, satellite and aerial remote sensing material. The inventory applications in which the segment-based approach is tested include the estimation of plot-level timber volumes (I -III) and the stratification of forested areas (IV).

All the tested and developed segmentation algorithms are applicable to the determination of feature extraction and image analysis units. The incorporation of image segment and sub-segment features into the estimation procedure improved the plot level volume estimates in most of the cases (I-III). The achieved reductions in the RMSEs are smaller than expected and it is questionable whether these small improvements justify the use of the segment-level estimation approach. Better results are achieved in the stratification of forested areas, even though some of the results are controversial (IV). The segment-aided approach can be recommended for stratification purposes.

The selection of the appropriate segmentation algorithm depends on the application, but in general a two-phase approach starting with initial segmentation and proceeding with merging of the initial segments is advisable. The initial segmentation method presented in III, and a region merging algorithm that can be guided with minimum segment size and similarity parameters are recommended methods especially for forest inventory applications that employ VHR imagery.

List of papers

I Mäkelä, H. & Pekkarinen, A. 2001. Estimation of timber volume at the sample plot level by means of image segmentation and Landsat TM imagery. Remote Sensing of Environment 77(1): 66-75.

II Pekkarinen, A. 2002. A method for the segmentation of very high spatial resolution images of forested landscapes. International Journal of Remote Sensing 23(14): 2817-2836.

III Pekkarinen, A. 2002. Image segment-based spectral features in the estimation of timber volume. Remote Sensing of Environment 82(2-3): 349-359.

IV Pekkarinen, A. & Tuominen, S. 2003. Stratification of a forest area for multi-source forest inventory by means of aerial photographs and image segmentation. In: Advances in Forest Inventory for Sustainable Forest Management and Biodiversity Monitoring. Forestry Sciences. Vol. 76: 111-124. Kluwer Academic Publishers, Dordrecht, Netherlands, ISBN 1-4020-1715-4, 460 pp.

In I, Pekkarinen was responsible for developing and implementing image segmentation and estimation algorithms. The report was written together with Mäkelä. In IV, Pekkarinen was responsible for developing and implementing of image segmentation and clustering algorithms. The report was written together with Tuominen.

Acknowledgements

The majority of this research was conducted from April 1998 to June 2002, during which time I was a part-time post-graduate student within the “Forests in Geographical Information Systems” Graduate School, University of Helsinki and part-time research scientist at the Finnish Forest Research Institute (Metla). I greatly appreciate the financial support provided by the Foundation of Foresters (Metsämiesten Säätiö) through the Graduate School, without that this work could not have been realised. In addition, I would also like to thank both, Metla’s National Forest Inventory research program and the University of Helsinki, Department of Forest Resource Management for providing the research material and all necessary work facilities for this thesis.

Numerous people encouraged and helped me throughout the course of this study.

My supervisor Professor Erkki Tomppo supported my idea to study image segmentation and helped me on several occasions. Both, he and Dr. Jari Varjo also initially encouraged me to apply to the Graduate School. I greatly appreciate their support. I would also like to thank Prof. Jouko Laasasenaho and Dr. Markus Holopainen for guiding us doctorants through the Graduate School and for arranging inspiring post-graduate seminars in Helsinki and Hyytiälä. Thanks also go to Emeritus Professor Simo Poso, with whom I have had many interesting and thought-provoking debates during both, my MSc and doctoral studies. It may actually be, that my interest in remote sensing can be traced back to an enlightening discussion that we had in a mature spruce stand in Toivala, May 1992.

I have been very fortunate to work with the Finnish National Forest Inventory team. I would particularly like to thank my excellent co-authors and colleagues Lic. Sc. Helena Mäkelä and Mr. Sakari Tuominen In addition to my co-authors, I am especially grateful to Lic. Tech. Kai Mäkisara who provided me with an outstanding programming environment for implementing the concepts that I had. Kai also helped me on numerous occasions to debug my awkward ‘C’ code and patiently answered all of my questions related to the AISA data. I would also like to thank Dr. Juha Heikkinen and Dr. Helena Henttonen for helping me to understand what the Ohkola field data was about.

Beyond the scope of this thesis, I would like to express my gratitude to my former and present colleagues that I have had the pleasure to work with. In particular I would like to acknowledge Ms. Reija Haapanen with whom I wrote my first scientific article, Ms. Tarja Tuomainen with whom I have shared many of my work-related headaches over the past couple of years and my dear colleagues and ‘co-lunchers’: Ms. Sirpa Rajaniemi, Mr. Markus Haakana and Mr. Mikael Strandström. I have shared many enjoyable and cheerful moments with all of them.

I would like to express my sincere gratitude to Prof. Annika Kangas, her encouragement and comments on the synthesis part of this thesis were invaluable. I would also like to thank my pre-examiners Prof. Alan Ek and Prof. Håkan Olsson for their constructive and valuable preview comments. Thanks also go to Dr. Ashley Shelby for editing the language of sub-studies I-III and the synthesis part of this thesis and to Ms. Anna-Kaisu Korhonen and Ms. Anne Siika for their friendly assistance in the publishing phase of this work.

Finally, I would like to thank my wife Heli for initially encouraging me to start the post-graduate studies and for her continued support during this process and our children Pihla and Pietu for constantly and concretely reminding me that every day really is an adventure.

Helsinki, June 2004

Anssi Pekkarinen

1 Introduction

Remote sensing (RS) provides invaluable information concerning spatial and temporal distribution of land use and forests. The oldest earth observation (EO) program provides spaceborne data dating back to 1972 with the launch of Landsat 1 (originally ERTS 1). In addition, aerial RS archives contain data that originate from the first decades of the 20th Century. Combining the temporal and spatial dimensions of these data enables the assessment of not only the present state of the environment but also the changes that have occurred over various periods of time. The value of these data has been recognised among the foresters and the number of RS based forest inventory applications is constantly increasing.

Although RS aided forest inventories have been shown to produce valuable information for large and medium sized areas, their applicability at the forest holding and stand level has been limited. Stand- and plot level estimation errors have remained high with different RS information sources and estimation techniques. These results imply that the high estimation errors for small areas do not straightforwardly result from the spatial or spectral resolution of the employed RS material, but the fundamental reasons for partial failures are elsewhere.

This thesis examines whether the applicability of remote sensing aided forest inventory methods can be improved by using the image segment -based approach to feature extraction and image analysis. Several image segmentation techniques are developed, implemented and tested in forest inventory applications employing both, satellite and aerial RS material. In order to present the issue in a broader perspective, let us first review the common history of remote sensing and forest inventory.

2 Remote sensing in forest inventory

In Finland, the first tests of airborne RS aiming at cartographic mapping were conducted by the Topographic Service of Finland in 1926-1927 (Löfström 1946). The results were promising and the tests stimulated significant technical and methodological developments. The developed techniques and equipment were soon given wider applications. Before the end of the 1930's, both the Topographic Survey of Finland and the Finnish National Land Survey provided maps based on aerial photographs in order to *"satisfy the constantly increased map-demand of economic and social life"* (Löfström 1946).

The potential of aerial photographs was soon recognised among the foresters. According to Sarvas (1938) and Nyyssönen (1955), the first forestry related applications were reported in Central Europe at the beginning of 1920's, but the initiative was later "*transferred to U.S.A and Canada*" (Nyyssönen 1955). In these first applications, aerial photos were used as maps in fieldwork and in timber surveys (Nyyssönen 1962, Lillesand et al. 2004).

After World War II, the increasing use of the infrared films stimulated the use of aerial material, and the number of applications increased. Probably the first significant Finnish effort to estimate the amount of growing stock from aerial photographs was presented in 1955 (Nyyssönen 1957). Nyyssönen concluded that the sensible way to accomplish aerial photography aided forest inventory would be the combination of fieldwork and aerial photo interpretation. The interpretation of aerial photographs could not satisfy all information needs of the forest management but, on the other hand, they provided possibilities that should not be ignored (Nyyssönen 1957). However, aerial photos were still used mainly as "*maps facilitating the work*" at the beginning of 1960's (Nyyssönen 1962).

In the late 1960's, foresters' expanded their attention from aerial photos to satellite imagery (e.g. Kuusela and Poso 1970), but a significant increase in forest inventory related satellite RS application was seen only after the launch of first satellite of Landsat program in 1972. Landsat 1 carried a multispectral scanner (MSS) instrument and provided multispectral EO data in digital format (Lillesand et al. 2004). This new data source and the rapid development of computer technology resulted in major developments in image analysis techniques and in an increasing number of RS applications.

So far, most of the environmental RS research has concentrated on examining the present and monitoring recent changes. The focus has been in land use and forest type classification (Bauer et al. 1994, Holopainen 1998, Haapanen et al. 2004), forest change (Goldberg et al. 1982, Saukkola 1982, Varjo 1996, Häme et al. 1998, Woodcock et al. 2001) and prediction of values of forest variables, such as timber volume and woody biomass, for areas of various sizes (Hagner 1987, Tomppo 1992a, Tokola et al. 1996, Mäkisara et al. 1997, Trotter et al. 1997, Kilpeläinen and Tokola 1999, Tomppo et al. 1999, Hyyppä et al. 2000, Halme and

Tomppo 2001, Holmström et al. 2001, Tomppo et al. 2001, Tuominen and Poso 2001, Anttila 2002, Dong et al. 2002, Tomppo et al. 2002).

Since Landsat 1, the spatial resolution of the RS material available in digital format has continuously improved. Currently, several satellite imaging systems provide data that possess spatial resolutions higher than 5 meters (Nieke et al. 1997). Furthermore, the availability of aerial imagery has significantly increased. New image compression algorithms and the development of web-based services provide ready access to digitized aerial photographs (METRIA 2003, ILMARI 2003) and the introduction of digital aerial cameras such as Z/I DMC (Z/I Imaging 2004) and Leica ADS40 (Leica Geosystems 2004) to operative aerial imaging services is approaching. In addition to aerial photography, aerial very high spatial resolution (VHR) data is available from airborne imaging spectrometers (e.g. AISA, CASI, DAIS 7915, ROSIS and HyMap) and active sensors such as airborne laser scanners (e.g. TopoSys, Optech ALTM and Leica ALS50) and airborne radars (e.g. CARABAS and GEOSAR). Many of these new data sources have shown to provide interesting data for forest inventory applications (e.g. Fransson et al. 2000, Holopainen 1998, Hyyppä et al. 1999, Mäkisara et al. 1997, Næsset 1997).

The challenge that the very high spatial resolution image material introduces to image analysis can be explained using the H- and L-resolution concepts (Strahler et al. 1986). In the case of L-resolution, the resolution of RS material is coarse in relation to the size of object of interest, i.e. a single L-resolution pixel contains information on several objects of analysis. In such a case, the image analysis can be approached using fuzzy or sub-pixel level image processing techniques. With these methods, the user can model the informational content of the pixel despite the fact that the objects of interest cannot be directly observed in the spatial domain. In the case of H-resolution, the problem is reversed and reasonable analysis of the object properties requires that the contextual information present in the image is taken into account. One, and probably the most simple, way to accomplish this is to use the square or rectangular neighbourhood of each pixel and to utilize spectral information from that area in the analysis instead of the information from a single pixel. The problem with this kind of approach is: how to determine the appropriate size and shape of the neighbourhood to be employed in the analysis of the phenomena of interest? Even though an adaptive solution to the definition of the size of the geographic window may be possible (Franklin et al. 1996), the determination of the shape remains a problem. It can be solved with methods that are able to compose spatial entities i.e. regions that can be used as basic units in image analysis instead of single pixels (Blaschke and Strobl 2001). These regions can be determined with help of image segmentation.

3 Image segmentation

3.1 General

Image segmentation is the division of an image into spatially continuous, disjointed and homogeneous regions. More formally, following the notation presented by Pal and Pal (1993), if a digital image is presented as

$$F_{PxQ} = \left[f(x,y) \right]_{PxQ}$$

where PxQ is the size (columns x rows) of the image and $f(x,y) \in G_L = \{0,1,...,L-1\}$ is the set of possible grey level values, image segmentation is partitioning of the set F into a set of homogeneous regions S_i in such a manner that

$$\bigcup_{i=1}^{n} S_i = F \text{ with } S_i \bigcap S_j = \varnothing, \quad i \neq j$$

The homogeneity of the regions is controlled with a homogeneity criterion, denoted by $P(S_i)$. The criterion has to be true for each region and false for adjacent regions. This ensures that every region is distinct from every other region. More formally: $P(S_i \bigcup S_j)$ is false when S_i is adjacent to S_j. $P(S_i)$ can be determined as convenient. It can, for example, be set in such a way that a segment may include only pixels that carry the same grey level value. In real-world applications, however, the criteria are usually much more complicated and may consist of, for example, a set of spectral and geometrical rules.

As pointed out by Haralick and Shapiro (1985), there is no theory of image segmentation. In addition, even though image segmentation is precisely defined, the word "segment" may be sometimes confusing. The online Merriam Webster (2004) dictionary gives, among many others, following meanings to the word "segment":

"One of the constituent parts into which a body, entity, or quantity is divided or marked off by or as if by natural boundaries" and *"Portion cut off from a geometric figure by one or more points, lines, or planes"*.

Thus, the word "segment" does not explicitly involve the requirement of spatial continuity that is characteristics to image segments. Therefore RS students and even specialists often misunderstand the meaning of an image segment by confusing it with a cluster. This is understandable when one considers the fact that the segmentation or clustering of an image may sometimes yield an identical result. This is an exception, however, and is found only in cases where the image consists of a background and a single, spatially continuous and separable, object.

3.2 Classification of image segmentation techniques

Image segmentation techniques can be classified in many different ways depending on the level of details included. Fu and Mui (1981) use three relatively coarse classes: 1) characteristic feature thresholding or clustering, 2) edge detection and 3) region extraction. Haralick and Shapiro (1985) use a more detailed classification and divide the methods into 1) measurement space guided clustering, 2) region growing schemes that include single, hybrid and centroid linkage region growing methods, 3) hybrid linkage combination techniques, 4) spatial clustering schemes and 5) split and merge schemes. Also Pal and Pal (1993) use a quite detailed classification and even separate colour image segmentation to its own class: 1) grey level thresholding, 2) iterative pixel classification, 3) surface based segmentation, 4) segmentation of colour images and 5) edge detection and 6) methods based on fuzzy set theory.

The latter two classifications are unduly complicated and technically oriented for the brief introduction to segmentation of remotely sensed images. The division into pixel-, edge and region based methods is sufficient for that.

Pixel based image segmentation methods include image thresholding, clustering in the feature space and other methods that rely on pixel-level information and employ it in the global feature space. The pixel-based methods may also include spatial components. For example, the input image may be a smoothed version of the original image. In such a case, the original pixel value has been replaced with a weighted average of its neighbouring pixels. Another example is a case in which the pixel has been assigned texture information describing its neighbourhood. These examples can, however, be classified as pixel-based segmentation approaches if the actual segmentation is conducted in the global feature space.

Image thresholding is a pixel-based technique in which an image is turned into a binary image in such a way that the objects of interest are separated from the background. The selection of an appropriate threshold value is usually based on *a priori* known properties of the object and background. Even though image thresholding may, in many cases, seem trivial that is not usually the case. The contrast between the objects may be poor and the illumination conditions may vary and cause artefacts (such as shadows) that make the determination of appropriate threshold difficult. In addition, many applications require that the appropriate thresholds can be determined automatically. However, from a forester's point of view, image thresholding, as such, is usually applicable only to relatively simple segmentation problems such as binarization an image into bright and dark areas for local maxima detection (e.g. Pitkänen 2001), and extraction of water bodies or clouds from remotely sensed images. Therefore, the discussion about different thresholding techniques is beyond the scope of this thesis. Examples of these techniques can be found in, for example, Weszka (1978) and Jain et al. (1995).

Image clustering can be seen as a multi-dimensional extension of thresholding (Fu and Mui 1981). A typical image clustering algorithm, such as ISODATA, is an iterative process that seeks to find natural classes within the feature space with help of user provided parameters. Depending on the implementation, the required parameters may include the number of clusters, the maximum number of iterations, the convergence threshold etc. A common solution is to start the clustering with a set of initial cluster centres that have been located in the multi-dimensional feature space in such a way that the distance between the centres is maximized. During the first iteration, each image pixel is assigned to that cluster centre that is closest to it in the given feature space. The locations of the cluster centres are subsequently re-determined with help of the pixels that fell to each cluster. The process is iterated until all the pixels remain in same clusters during two sequential iterations or until the proportion of the pixels changing clusters is smaller than the given convergence threshold (e.g. ERDAS 1994).

Because image thresholding and clustering methods produce results that may have several spatially discontinuous units that carry the same label, the result does not fulfil the definition of segmentation until the spatially continuous regions have been identified and re-labelled. This can for example be done using connected component labelling (CCL) -algorithm (Jain et al. 1995).

The nature of edge-based image segmentation methods differs significantly from that of pixel-based methods. The first phase in all edge-based segmentation algorithms is, of course, the detection of edges. An edge point (pixel) in an image can be defined as:

"... a point in an image with coordinates [i,j] at the location of a significant intensity change in the image." (Jain et al. 1995).

Given this definition, to decide whether a pixel is an edge pixel or not one needs to analyse it and its neighbourhood. In general, edge detection consists of the following steps: a) filtering, b) enhancement and c) detection (Jain et al. 1995). The filtering step is required because most of the edge enhancement methods are relatively sensitive to image noise and therefore they perform better when using a smoothed input. The edge enhancement phase is usually carried out using specific edge operators that emphasize pixels having significantly different values than their neighbours. Most of these operators, such as Roberts, Sobel and Prewitt operators, are based on discrete approximation of the gradient that in the case of images is a two-dimensional equivalent of the first derivative (Jain et al. 1995). They usually produce sufficient results for most applications, even though they typically result in relatively thick edges. In case a more precise location of the edges is required, second derivative operators, such as Laplacian and Second Directional Derivative (Jain et al. 1995), can be used.

After image filtering and edge enhancement, the remaining step in edge detection is the recognition of edge points (pixels) among the edge candidates. This is usually carried out with help of thresholding. In the simplest case, all pixels having an edge magnitude above a threshold T are considered as edge pixels. In many real-world cases that deal with noisy images, it may be very difficult to find a threshold that keeps the probability of detecting false edges low while finding all the relevant edges. It may therefore often be necessary to use several thresholds. For example, Canny (1986) suggests the use of two thresholds T_1 and T_2. The idea is to detect an edge contour using the higher threshold T_2 and to also mark as edges all connected edge pixels of that contour that have edge magnitude higher than T_1. The suggested relation for the thresholds is: $2T_1 < T_2 < 3T_1$.

Despite the method with which the edge pixels are detected, the final phase of the edge-based methods is to link the detected edges and to compose meaningful boundaries. The simplest way to represent a boundary is to use an ordered list of its points, but more compact representations are usually preferred because they provide more efficient basis for subsequent operations. Examples of different means to represent boundaries can be found in, for example, Jain et al. (1995).

Region-based image segmentation techniques differ from pixel- and edge-based methods in the way they deal with spatial relationships. Region-based techniques can be all seen as region growing techniques (e.g. Zucker 1976) or further divided into region growing, merging and splitting techniques, and their combinations. Here, the latter classification is used.

There are several approaches to region growing. The algorithm may require a set of seed pixels or regions with which the process is started, or it may simply start with the initial image and process it pixel-by-pixel. If seeding is required, the seed pixels or areas may be shown interactively on screen or selected automatically. Where seeding is not required, the processing usually begins from the top left corner of the image and proceeds from left to right and top to bottom. Despite the processing details, the region growing techniques usually join neighbouring pixels to a same region if their spectral properties are similar enough. The similarity can be determined in terms of a homogeneity criterion or a combination of homogeneity, size or some other characteristics criteria (Zucker 1976). Following the definition of segmentation, the region growing process terminates after every pixel has been assigned to a segment.

In region merging and splitting techniques, the image is divided into sub-regions and these regions are merged or divided according to their properties. The basic idea is to start with initial regions and merge similar adjacent regions. These initial regions may be single pixels, or areas determined with help of any low-level segmentation technique. Region splitting methods operate in the opposite fashion; the input usually consists of large segments that are divided into smaller sub-segments with help of a simple geometric rules. If the sub-segments are not homogeneous enough, they are further divided and the process is continued. A common way to implement the region-splitting technique is to use a quad tree structure (Figure 1). The basis on which the splitting or merging is done may be, for example, the spectral similarity of the segments or the magnitude and length of their common edge (Zucker 1976).

The above presented classification of image segmentation techniques is not comprehensive. Many segmentation methods fuse properties of algorithms of several classes. An example of a segmentation method that combines features of edge- and region-based approaches is the “Image segmentation with directed trees” -algorithm (Narendra and Goldberg 1980).

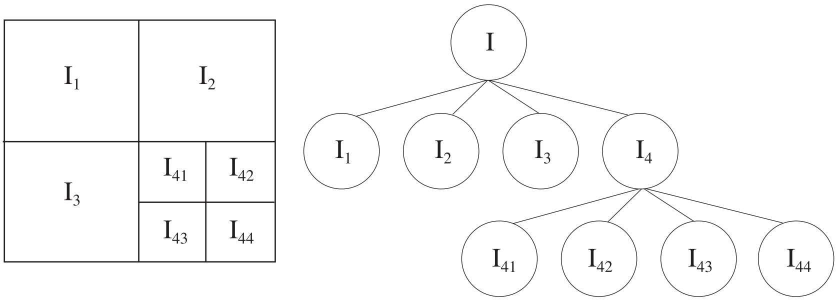


Figure 1. Segmented image and corresponding quadtree (modified from Gonzales and Woods, 1993).

The selection of the appropriate image segmentation approach and algorithm for a specific task on the basis of the algorithm description may be difficult. The segmentation approaches may have advantages and disadvantages that cannot be recognised prior to testing the algorithms with actual imagery. In addition, several different algorithms may result in similar or not dissimilar segmentation output. In practice, the decision is usually made between the algorithms that are commercially available. Unfortunately, there are few such algorithms.

Practically, the first commercially available segmentation software packages that were designed for the analysis of remote sensing data were released in 2000 (Schieve et al. 2001). Since then the interest in the development of such segmentation packages has increased and currently segmentation tools are available for many leading image processing software packages. One software package that deserves explicit mentioning in forestry context is *eCognition*. It is currently the leading commercial image segmentation and object-oriented image analysis software designed for analysis of RS data and has recently been strengthened with a new tool designed for automated tree crown delineation (Definiens 2003).

eCognition's multivariate segmentation is based on a region merging technique that starts with regions of one pixel in size. The region-merging algorithm is iterative and it merges adjacent regions based on their spectral and spatial properties. The main parameters controlling the algorithm are "scale" and "homogeneity criteria". The "scale" restricts the allowed heterogeneity of the resulting segments with help of the homogeneity criteria that can be controlled by weighting the "colour" and "shape" parameters. "Colour" refers to spectral and "shape" to geometric properties of the segments. Furthermore, the shape parameter is a combination of segments "smoothness" and "compactness" that can be weighed by the user (Baatz et al. 2002).

4 Segmentation in forest inventory

The use of image segmentation as a tool for analysing earth observation imagery is, of course, not new. In their segmentation review, Haralick and Shapiro (1985) cite EO image segmentation studies that have been conducted as early as the mid-1970's. The actual need for image segmentation tools was, however, recognised later, soon after the launch of the Landsat 4 (1982) and SPOT 1 (1986) satellites. These satellites introduced new sensors that provided images with considerably improved spatial resolution. In addition to early satellite image segmentation pioneers, many other scientist among the RS community were convinced that these sensors would motivate an increased use of contextual methods in the analysis of satellite remote sensing images.

Despite the fact that the need for contextual image analysis methods was recognised, only few forestry applications for segmentation of EO images were presented. In Scandinavia, many of the forestry related approaches were developed on the basis of directed-trees -algorithm presented by Narendra and Goldberg (1980). Tomppo (1987) tested the algorithm in stand delineation for the estimation of several stand attributes, namely total volume, mean diameter at breast height, mean age and volume proportion by species for pine, spruce and deciduous trees. Later, a similar or not dissimilar method has been applied to the segmentation of Landsat and SPOT imagery by Parmes (1992), and in the spatial generalization of pixel-level change-detection (Häme 1991) and forest site fertility classification results (Tomppo 1992b).

Algorithms designed for more specific purposes were also developed. Hagner (1990) presented a method he calls *"t-ratio segmentation"* that is used for the automatic delineation of stands. Hagner describes the method as *"a type of region growing algorithm",* but it can also be classified as a region merging method. The same segmentation method has later been used for change detection (Olsson 1994). Another example of an algorithm that has been designed for a specific purpose is a method that was aimed at the delineation of stands for the construction of forest canopy reflectance models. The method has been presented by Woodcock and Harward (1992) and was later employed in the generalization of change detection results (Woodcock and Macomber 2001).

In general, the segmentation methods produced promising results in stand delineation, in the estimation of forest parameters and in post-processing of the results of pixel-based analysis. For example, it has been concluded that stand delineation *"seems to work quite well"* and that *"it is possible to develop a stand-wise forest inventory method based on the satellite images"* (Tomppo 1987), and that segmentation based stand delineation with SPOT imagery followed by manual editing is comparable to results achieved with visual interpretation of aerial images (Hagner 1990). Further, the precision of stand-level estimates of stand volume and mean diameter was found to be comparable to the results of subjective field inventory (Hagner 1990).

In spite of the promising early results, the number of segment-based forestry applications remained low. There are two probable reasons for this. First, it was soon observed that the improved spatial resolution of the satellite imagery did not necessarily require contextual image analysis. Despite the preliminary doubts, the pixel-level analysis worked reasonably well with these imageries. Another reason for few reported applications was the lack of commercial image segmentation software.

The need for segmentation was perceived again in the 1990's. Increasing availability of digitized aerial photographs and the approach of a new generation of satellites providing VHR data re-stimulated the discussion concerning contextual image analysis. As a consequence of that the interest in object oriented image analysis has steadily increased, the pixel-by-pixel approach to image analysis has been increasingly criticized (e.g. Blaschke and Strobl 2001) and image segmentation has been tested in numerous VHR applications. Examples of these applications are delineation of habitats for biodiversity assessment (Holopainen 1998), delineation of individual tree crowns from aerial and other high spatial resolution imagery (Pitkänen 2001, Gougeon, 1995, Burnett 2003) and change detection (Pekkarinen and Sarvi 2002, Saksa et al. 2003). Even commercial services that are based on segmentation technology and aim at stand-level inventories are already available (FACT 2004). In addition to the optical imagery, image segmentation has been increasingly used to analyse airborne laser scanning (ALS) data. Examples of segment-aided ALS applications include the delineation of trees for change detection and growth estimation (Yu et al. 2003) and for the extraction of forest inventory parameters (Diedershagen et al. 2003). It seems, that image segmentation of RS imagery is experiencing its second renaissance.

5 Objectives

The main objective of this thesis has been to study whether the applicability of multi-source forest inventory (MSFI) methods can be improved using the image segment based approach. More specifically, the objective has been to study the applicability of segment-level analysis to the estimation of timber volume and stratification of forested areas. The particular objectives of sub-studies I-IV have been:

I To study the effect of segment-level feature extraction and image analysis on the accuracy of plot-level multi-source forest inventory estimates with help of Finnish National Forest Inventory field data and Landsat TM imagery. In addition, the selection of appropriate segmentation method has been addressed.

II To develop a method for the segmentation of very high spatial resolution imagery of forested landscapes and to evaluate its applicability to MSFI, specifically with respect to the estimation of timber volume.

III To introduce and test multi-scale segment based features in MSFI, specifically in the estimation of plot-level timber volume with help of field data and VHR imagery.

IV To evaluate an image-segment based approach to the stratification of forested area with help of aerial photographs.

6 Material

6.1 Field data

The sub-studies were carried out at four different test sites. The field data included four data sets two of which were gathered by the personnel of the National Forest Inventory (NFI) of Finland. The first of those sets included a subset of plots from 9th NFI (I) and the second set consisted of a dense grid of systematically sampled NFI-like field plots (II and III). The third and fourth field data sets (IV) consisted of circular and relascope field sample plots located in two study areas in Southern-Finland and measured by the Department of Forest Resource Management of the University of Helsinki.

6.2 Image material

The sub-studies of this thesis employ images from several different RS data sources. Sub-study I was based on an analysis of spaceborne imagery, namely two Landsat TM images. The thematic mapper (TM) is mounted on a satellite platform that orbits the earth at a nominal altitude of 705 kilometres. It sweeps the earth from west-to-east and east-to-west and collects data during both sweeps. It has seven bands, a quantization range of 8 bits and a spatial resolution of 30 (bands 1-5 and 7) and 120 meters (band 6) (Lillesand et al. 2004).

In sub-studies II and III, the imagery employed was acquired with a pre-series version of Airborne Imaging Spectrometer for Applications (AISA). AISA is a pushbroom type scanner recording radiation in the range 450 to 900 nm. The pre-series version of AISA has 286 spectral channels and the number of pixels per line is 384. The instrument is programmable and has four operating modes. The selectable parameters of AISA include the number of channels, wavelength and bandwidth of each channel, operating mode and integration time. The instantaneous field of view (IFOV) of the instrument is 1 milliradian and its dynamic range 2500 digital numbers. Across track pixel size of the instrument depends on the IFOV and flight height and along track pixel size on the velocity of the aeroplane and the integration time. For example, one meter pixel size is achieved with flight height of 1000 m, speed of 50 m/s and integration time of 20 ms (Mäkisara et al. 1993). The details of the AISA data employed here can be found in II and III. The first prototype of AISA was developed in the early 1990's, and currently the AISA family consists of three different systems: AISA+, AISA Eagle and AISA Hawk (SPECIM 2004).

In sub-study IV, the analysis was carried out using CIR aerial imagery. The images were obtained with a Wild RC30 camera, UAGA-F 13158 optics and

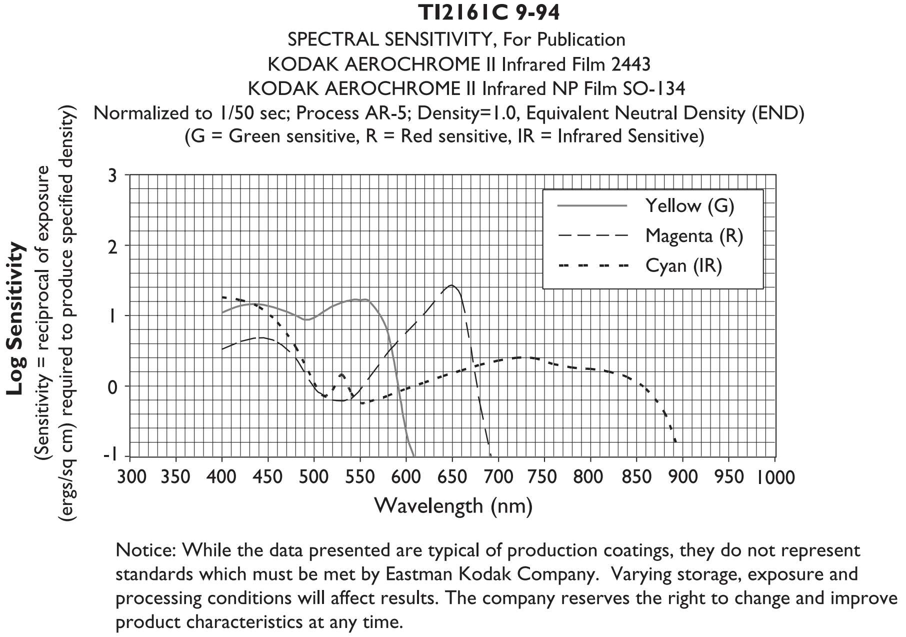
Kodak Aerochrome II Infrared Film 2243. The film characteristics curve is presented in figure 2. The antivignetting AV520 nm, and IR80% filters were used. The images were scanned using Zeiss Scai -scanner and 14 µm resolution and resampled to the pixel size of 0.5 metres.

All these data sources have different resolution characteristics that affect to their applicability in forest inventory applications. Note, that the term "resolution" refers to spatial, spectral or radiometric resolution (Lillesand et al. 2004). Spatial resolution describes the sensors capability to record spatial details, whereas spectral resolution determines the wavelength area to which the sensor is sensitive. The sensor's radiometric resolution determines the magnitude of the differences in the radiation that can be observed. In the case of aerial films, the radiometric resolution is usually described with help of the film characteristics curves (Lillesand et al. 2004). In real imaging systems, there is always a trade-off between these different types of resolution, and the choice of the appropriate sensor depends on the task to be conducted. In the following, only the differences in spectral and spatial resolution characteristics of the employed imagery are discussed.

The wavelength areas of the imagery employed in the sub-studies are presented in Table 1. From the standpoint of multi-source forest inventory the best performing sensors are TM and AISA. TM covers the widest range of spectrum and AISA is capable of dividing the spectrum into very narrow bands. This may be useful in the analysis of a phenomenon that can be observed only in a narrow range of the spectrum. Note, that in II, the estimation was carried out using the original 30 spectral AISA channels whereas in III these channels were generalised to four channels imitating the spectral characteristics of new generation VHR satellites (e.g., IKONOS). Detailed spectral characteristics of the employed AISA imagery are presented in II and III.

The drawback in both aerial AISA and CIR imagery is that the spectral sensitivity of the sensor (or film) is limited to the range of about 400 nm to about 900 nm. However, both of these data sources provide superior spatial resolution when compared to that of the TM sensor. The spatial resolution of channels of Landsat TM imagery employed is 30 meters. The corresponding figures with AISA and aerial imagery were 1.6 and 0.5 meters, respectively. In IV, however, the aerial imagery was resampled to a pixel size of 1.5 meters prior to the analysis

In addition to resolution characteristics, there are other factors that affect the applicability of remote sensing imagery to multi-source forest inventory. The radiance that a given remote sensing sensor observes is affected by the sun-object-sensor geometry, atmospheric attenuation, bidirectional reflectance and for non-lambertian surfaces also factors such as land cover or vegetation type (Leckie 1987). The magnitude with which each of these factors affects the observed radiance depends mainly on the view angle and imaging altitude. In satellite imagery, the main disturbing factors are usually scattering and absorption in the atmosphere (Song et al. 2001) and in low altitude imaging systems the sun-object-sensor geometry and bidirectional reflectance of the land surface (Pellikka et al. 2000).



Notice: While the data presented are typical of production coatings, they do not represent standards which must be met by Eastman Kodak Company. Varying storage, exposure and processing conditions will affect results. The company reserves the right to change and improve product characteristics at any time.

Figure 2. Spectral sensitivity curve of KODAK AEROCHROME II CIR film.

In spite of the fact, that the correction of atmospheric attenuation is often necessary prior to the classification and analysis of satellite EO imagery, that is not always the case. If the training data and the imagery are on the same relative scale, as in I, atmospheric correction has only minor effect on the image analysis results and is therefore unnecessary (Song et al. 2001).

The radiometric distortions in aerial imagery are often larger than in spaceborne material. The sun-object-sensor geometry and bidirectional reflectance effects cause radiometric distortions that may complicate the analysis of the imagery. The most important factors affecting the bidirectional reflectance of the forests include the hotspot effect and effects caused by mutual shadowing between trees, branches and leaves. The hotspot effect is observed when the viewing and illumination positions coincide, because the shadows are hidden behind illuminated objects and only bright object are registered on the scene (Li and Strahler 1992). In addition, the objects in the sun-side of the image appear darker, because only the shadowed part of the crowns and trunks are visible to the sensor. The phenomenon is more obvious at large viewing angles.

The AISA imagery employed in II and III consisted of data collected from seven flight lines. The analysis of the radiometric differences between the lines revealed, that adjustment of the pixel values was necessary. The adjustment was carried out using overlap areas of adjacent flight lines and cumulative histogram matching. In IV, the radiometric quality of the image mosaic was generally good. Similar objects had similar or close-to-similar properties in different parts of the images and therefore the radiometric correction was considered unnecessary after the exclusion of one sub-area that had significantly different spectral properties than the rest of the area.

Table 1. Characteristics of the image material employed.

Image type	**Channel**	**Sensitivity, nm**
Satellite Landsat TM)	1	0.450 - 0.520
	2	0.520 - 0.600
	3	0.630 - 0.690
	4	0.760 - 0.900
	5	1.550 - 1.750
	6	10.400 - 12.500
	7	2.080 - 2.350
Imaging spectrometer (AISA)	1	0.470 - 0.477
	2	0.488 - 0.495
	3	0.505 - 0.512
	4	0.523 - 0.530
	5	0.542 - 5.488
	6	0.550 - 0.556
	7	0.565 - 0.572
	8	0.579 - 0.587
	9	0.600 - 0.607
	10	0.623 - 0.630
	11	0.647 - 0.654
	12	0.668 - 0.676
	13	0.676 - 0.683
	14	0.686 - 0.691
	15	0.697 - 0.701
	16	0.701 - 0.706
	17	0.711 - 0.715
	18	0.726 - 0.733
	19	0.733 - 0.738
	20	0.742 - 0.747
	21	0.750 - 0.755
	22	0.776 - 0.780
	23	0.785 - 0.790
	24	0.794 - 0.799
	25	0.803 - 0.808
	26	0.818 - 0.823
	27	0.844 - 0.849
	28	0.853 - 0.858
	29	0.861 - 0.866
	30	0.865 - 0.870
Aerial CIR photographs	See figure 2	

7 Methods

7.1 Image segmentation

In I, the Landsat TM imagery was segmented using two different approaches: a) measurement space guided clustering followed by connected component labelling (ISOCCL) and b) modified implementation (NG) of the “Image segmentation with directed trees” -algorithm (Narendra and Goldberg 1980). The resulting initial segmentations were fine-tuned using two different region merging (RM) algorithms that differ in the way they compute the similarity of adjacent segments. The similarity was determined using either the Euclidean distance between the segments or their t-ratio (see e.g. III, page 353). The RM algorithms were guided by a minimum segment size parameter that was set to 0.5 hectares, and in the case of t-ratio RM by a spectral similarity threshold.

In II, a new two-phase segmentation algorithm was developed. The method is based on the assumption that a VHR image of a forested area is a composition of certain spectral classes that have spatial relationships that can be modelled with their co-occurrence statistics. The developed method forms initial segments via clustering and CCL (as in I) and merges the initial regions to applicable segments with the help of a novel RM algorithm, namely co-occurrence region merging (CRM). The CRM is an iterative algorithm that combines each segment smaller than a given minimum size to a neighbouring segment with which it has the largest content co-occurrence. During the first iteration, the co-occurrence is determined using the original cluster labels of the segments. In subsequent iterations, the mode of the cluster labels of possibly merged segments is employed; see II for a more detailed description of the algorithm.

The applicability of the developed algorithm was tested in the segmentation of AISA imagery. Two different initial segmentations were derived, one based on clustering of the first three principal components (PCs) and row and column coordinates (S1), and one based on the use of the three first PCs only (S2). In the subsequential region merging phase the CRM -algorithm was guided by a minimum segment size parameter that was set to 0.03 hectares.

Paper III describes the implementation of a segmentation algorithm (INISEG) that is based on the ideas of Narendra and Goldberg (1980). The algorithm is used as an initial segmentation method in a two-phase segmentation process that aims to delineation of feature extraction and image analysis units for VHR imagery. The initial segments produced by the algorithm were further processed using a region-merging algorithm that compares the spectral similarity of adjacent segments with the aid of their mutual t-ratio. The two-phase segmentation algorithm was tested in the segmentation of AISA imagery that was spectrally generalized to correspond to the spectral properties of new generation VHR satellite images. The minimum segment size was set to 10 pixels and three

Table 2. Summary of applied image segmentation and region-merging algorithms.

Initial segmentation algorithms	
Algorithm	Employed in sub-studies
Modified implementations of "Image segmentation with directed trees": NG and INISEG	I, III, IV
Measurement space guided clustering followed by connected component labelling (CCL)	I, II
Region merging (RM) algorithms	
Algorithm	Employed in sub-studies
t-ratio RM	I, III
Co-occurrence RM	II
Euclidean distance RM	I, IV

different segmentations (SA, SB and SC) were derived using t-ratio threshold values 0, 24 and 40.

In IV, the initial segmentation of the aerial imagery employed was derived using the INISEG algorithm described in III. The initial segments were merged to larger entities using RM algorithm that was guided by a minimum segment size parameter. All segments smaller than the given minimum size were merged to the most similar neighbouring segment. The similarity of the segments was determined by means of their Euclidean distance in the spectral feature space. Two different minimum sizes, 380.25 and 675 m^2, were tested.

A summary of the developed, implemented and tested segmentation algorithms is presented in Table 2.

7.2 Feature extraction and selection

In I, the Landsat TM spectral average features were extracted in two different ways: from square shaped windows surrounding the plots (reference features) and from those pixels within the windows that belonged to the same segments as the plot pixel (segment-restricted features). Window sizes from 1 x 1 (0.06 ha) to 11 x 11 pixels (7.56 ha) were tested. A basically similar approach was applied in II, but due to the better spatial resolution of the AISA imagery, a larger number of different window sizes was tested. Both, reference and segment-restricted features were extracted from windows of sizes from 3 x 3 pixels (about 0.002 hectares) to 121 x 121 pixels (about 3.75 hectares). In II, the extracted spectral feature set consisted of sub-optimal features found in an earlier AISA study (Mäkisara et al. 1997).

The objective in III was to test the performance of multi-scale segment based features. The analysis was therefore extended to include tests of segment-level features that were derived using all pixels within the segments in addition to reference and segment-restricted features. Segment-restricted and reference features were extracted from square shaped windows of 31 x 31 pixels (about 0.25 hectares). The features extracted included spectral averages and standard deviations computed from the spectrally generalized AISA imagery. Similar segment-level and reference features were extracted and employed in IV, but in the case of reference features, the size of extraction window was set to about 20 x 20 m^2, which corresponds to the size recommended in Holopainen and Wang (1998).

In III, the different number of reference and segment-based spectral features complicated the comparison of their performance. A subset of five best-performing features from both datasets was therefore selected for estimation tests. The feature selection was carried out using a sequential forward selection algorithm. It started by selecting the feature giving the lowest RMSE and proceeded by adding the feature that gave the best performance with the already selected features. The five best features were chosen from both reference and segment-based datasets for the evaluation tests.

7.3 Evaluation of segment-based approaches to estimation and stratification

7.3.1 General

The applicability of the segment-aided approach to feature extraction and image analysis was evaluated by employing segment-based spectral features, namely spectral averages (I-IV) and standard deviations (II-IV) and their sub-optimal combination (II), in the estimation of plot-level timber volume (I-III) and stratification of an inventory area (IV). The performance of segment-based spectral features was compared to that of reference features that were extracted in a more straightforward manner. These methods were selected for evaluation of segment-based approaches because they provide a more objective basis for comparisons than, for example, visual analysis of segmentation result. The evaluation of different segmentation results (I and II) could have been based on GIS analysis of the location of the segments borders, shape of segments and other segment properties, but because there is no analytical way to determine the "correct" segmentation, indirect analysis was needed. An alternative way to compare the segmentation results would have been the analysis of the within and between segment variances of stand characteristics with help of dense grid of field plots (Hagner 1990).

7.3.2 Estimation of timber volume

The estimation tests (I-III) were carried out using field sample plot data and extracted spectral features. The estimates for the plot level timber volumes, i.e. total volume and volumes by tree species, were derived using an inverse distance weighted non-parametric *k* -nearest neighbour estimator (*k*-NN, e.g. III equation 4) (e.g. Tokola et al. 1996, Tomppo 1996, Franco-Lopez et al. 2001) and a leave-one-out cross-validation technique. In cross-validation, every field plot was in turn omitted from the dataset and its characteristics were predicted with the aid of the other plots. The number of employed nearest neighbours was chosen separately for each study and was 5 in III and 10 in I and II. The performance of reference and segment-based features in the estimation was compared on the basis of root mean square error (RMSE) and relative RMSE (I-III), and empirical bias (III). In sub-study III, the analysis was extended to volume classes with the help of confusion matrices and their user's, producer's and overall accuracy (Stehman 1997).

7.3.3 Stratification of forest area

In IV, the evaluation of the segment-based approach to stratification was based on an analysis of within-strata variation in the spectral information and forest attributes. The stratification was conducted for segment-based and reference data using extracted spectral average and standard deviation features and a k-means algorithm (MacQueen 1967). Different strata numbers (20 - 50) were tested and the homogeneity of each stratum was characterized with the aid of area weighted mean standard deviations of forest attributes.

8 Results

8.1 Estimation tests

8.1.1 Landsat TM imagery (I)

The sensitivity of the traditional pixel-by-pixel image analysis methods to locational errors has been one reason for high plot- and stand-level estimation errors of MSFI applications employing satellite imagery. One possibility to diminish the effect of locational errors in the plot-based training data is to extract the spectral information for each plot not only from the plot pixel, but also from its spatial neighbourhood. This task can be conducted with the aid of image segments. The approach was tested with Finnish NFI data and Landsat TM imagery. Four different segmentations were derived using two different initial segmentation (NG and ISOCCL) and region merging (NN and TR) algorithms.

In general, segment based features gave lower estimation errors than features extracted from square-shaped windows. In the case of the estimation of total volume, the features derived using the combination of ISOCCL and TR gave the best estimation result. Best estimates for pine and spruce volumes were achieved with features derived using the combination of ISOCCL and NN, and best estimates for the volume of broadleaved species with features extracted from the image segmented using a combination of NG and TR. The differences in the performances of segment-based features and features extracted from square-shaped windows were, however, insignificant.

The clustering phase of ISOCCL is computationally demanding. In addition, it produced about 2.5 times higher number of initial segments than the NG. The NG segments also corresponded better to forest stand structure. Consequently, NG was judged to be the better choice as the initial segmentation method for forest inventory applications.

8.1.2 AISA imagery (II and III)

In II, a new two-phase algorithm that was designed for the segmentation of VHR images was developed and tested in the segmentation of AISA imagery. Two different segmentations (S1 and S2) were derived. The input of S1 consisted of the first three principal components (PCs) of the AISA imagery, and row and column coordinates. S2 was derived on the basis of the three first PC's only.

The visual analysis of the segmentation results showed clear differences between S1 and S2. The use of row and column coordinates as additional input channels resulted in improved recognition of spectrally homogeneous areas, such as recent clear cuttings, but at the same time some relevant segment borders were lost. The exclusion of the coordinate channels from the analysis resulted in a more distinct segment structure. The drawback of S2 was that even spectrally very homogeneous areas were in many cases divided into several segments.

The estimation results showed only minor differences in the performances of the segment-based and reference feature sets. Segment-based features extracted using S1 gave slightly smaller RMSEs for all estimated variables than features extracted using S2. In the case of estimates of total volume and the volume of deciduous species, S1 features gave the best results, whereas the lowest RMSEs for the estimates of pine and spruce volumes were achieved with the reference feature set. S1 features were less sensitive to the change in the size of the extraction window. The differences in the RMSEs were insignificant in most of the cases.

One means of ensuring that the extracted spectral information is representative at the field sample plot level, is to extract multi-scale features and select the best performing features for the actual analysis. That approach was tested in III. The spectrally generalized AISA imagery was segmented with a two-phase segmentation method. Initial segmentation was carried out with the INISEG algorithm and the resulting output was further processed with a t-ratio-based region merging algorithm with three different similarity thresholds.

The segmentation algorithm performed well and produced segments that were visually appealing. In addition, the results of the estimation tests showed that segment based features (SF) performed better than reference features (RF) in the estimation of all tested variables except in the estimation of pine volume. The differences in the performances of the feature sets were, however, small. Even though SF performed better than RF, the evaluation of confusion matrices and user's (UA) and producer's (PA) accuracies revealed problems especially in the estimation of high timber volumes.

8.2 Stratification test (IV)

In Finland, the data for forest management planning has been traditionally gathered by stand-level visual field inventories. The inventory method has been criticized because of its subjectivity, and the introduction of more objective methods have been suggested. One of the alternatives suggested is a two-phase sampling scheme (e.g. Poso and Waite 1996). The problem with the suggested approach is that it is difficult to determine the appropriate density for the first phase sample. Paper IV suggests that this problem could be solved by replacing the two-phase scheme with a segment-based approach. Stratification based on clustering of spectrally homogeneous segments is assumed to result in strata that are more homogeneous in their spectral and forest characteristics than strata derived from the two-phase sampling approach.

The results were controversial. In both study areas (S1 and S2) the distributions of the extracted spectral values proved that segment-level features retain better the original spectral variation of the images than features extracted from square shaped windows surrounding the first phase sample plots. This observation was evident in both study areas and holds true for both the minimum segment sizes employed. In spite of these results, the within-strata variation of forest attributes did not reveal clear difference between the approaches. In general, the results for S1 imply that when forest characteristics are considered the stratification based on a two-phase sampling scheme results in more homogeneous clusters than the segment-based approach. In S2, the results were partly reversed. The spectral features employed in the clustering phase have, however, a relatively large effect on the results. In S2, for example, the stratification based on merely spectral averages showed that segment based stratifications produce more homogeneous strata than the two-phase sampling approach. The inclusion of standard deviation features into the clustering process, however, implied that the two-phase sampling strategy would result in better strata.

9 Discussion and conclusions

The sub-studies of this thesis developed and tested several different approaches to image segmentation and their MSFI applications. The developed and implemented algorithms were tested in the estimation of plot-level timber volumes (I-III) and in the stratification of forested areas (IV). The incorporation of image segment- and sub-segment features into the estimation procedure improved the plot level volume estimates in most of the cases (I-III). The achieved reductions in the RMSEs were, however, smaller than expected and it is questionable if these small improvements can be used as basis for the recommendation of segment-level estimation approach. Better results were achieved in the stratification of forested areas, even though some of the results were controversial. In spite of that, the segment-aided approach can be recommended for stratification purposes.

The finding that the segment-based approach did not significantly improve the estimation results may be due to many reasons. First, the type of field data employed in the estimation studies (I-III) is sensitive to locational accuracy and may be unrepresentative for their neighbourhood (Koivuniemi 2003). The suitability of the employed data to segment-level analysis is therefore questionable. Furthermore, the field data employed in II and III had been pre-processed in such a way that the studies may give an over-optimistic impression concerning the performance of window-based feature extraction approaches that were used as benchmarks in the evaluation of segment-based results.

In general, the results of remote sensing based estimation and stratification applications are largely dependent on the spectral and radiometric properties of the imagery employed. Atmospheric attenuation, sun-object-sensor geometry and bidirectional reflectance effects may cause radiometric distortions that hinder the remote sensing -based analysis of phenomena that manifest themselves in slight spectral changes. The effect of these factors is generally larger in data that has been acquired from low-altitudes and using wide-angle lenses (Pellikka et al. 2000). From the viewpoint of this thesis, however, these factors are of minor importance. Even though these factors may have affected the absolute estimation errors (I-III) and within strata variations (IV), it is unlikely that they affect the mutual relationship of results of segment and non-segment based analysis on which the judgement of the segment-based approach was based.

All the tested and developed segmentation algorithms were applicable to the determination of feature extraction and image analysis units. However, bearing in mind that the selection of appropriate segmentation parameters is often based on trial-and-error, segmentation methods in which the creation of initial segments is based on computationally intensive iterative clustering algorithms may be too time consuming, at least from an operational point of view. In addition, initial segmentations created with these algorithms and VHR imagery usually produce a large number of initial segments that are complex in shape and do not necessarily correspond well with the stand structure. The complexity of the initial segments can, at least in some cases, be diminished by introducing spatially varying components (e.g. image coordinates) to the clustering process. However, that is not advisable in general, because some of the relevant borders may be lost and the use of other than spectral information complicates the interpretation of the segmentation result. Most of these problems can be avoided if initial segments are created using algorithms that are based solely on local image properties (see III).

Region merging algorithms that can be guided with minimum segment size and similarity parameters provide a meaningful way to aggregate small, potentially irrelevant regions to meaningful spatial entities that are, in most cases, more applicable in MSFI analysis than the initial areas. Where there exists areas that are small, but spectrally separable and relevant, the similarity parameter can be used in such a way that it prohibits the merging of two segments that are spectrally very dissimilar. In a forestry context, this kind of functionality might be needed in, for example, the separation of small underproductive rocky areas from the surrounding productive forest land.

Even though image segmentation is often referred as an objective way to isolate and determine spatial units for image analysis, segmentation result may be very sensitive to the user-defined parameters. The selection of various spectral or other characteristics parameters, such as input channels and minimum segments size, is often based on subjective decisions that may have a drastic influence on the final segmentation result. In addition, image segmentation is highly sensitive to the quality of image material. Special attention has to be paid to segmentation of imagery acquired from low altitudes and using wide-angle lenses because they typically include a large proportion of shadows. In some forestry applications, such as automated stand delineation, these shadows may cause serious locational errors in the final product. Even with its drawbacks, however, image segmentation is a tool that should not be bypassed when considering alternatives for VHR image analysis.

The algorithms developed here were implemented in such a way that they aim at a high level of automation. They are therefore guided by few parameters and the user has practically no other control over the segmentation procedure. Even though the implemented software has been successfully applied in this thesis and for many purposes not reported here (e.g. Pekkarinen and Sarvi 2002, Sell 2002, Saksa et al. 2003, Tuominen and Pekkarinen 2004), the highly automated approach is probably not the best alternative for all forestry applications. If segmentation is used in tasks that require a knowledge-based interpretation, such as semi-automatic stand delineation, a different approach is recommendable. For such purposes the algorithms should be implemented in such a way that they allow better interaction and thus the incorporation of superior image analysis capabilities of human beings into computer-aided image processing systems.

References

Anttila, P. 2002. Nonparametric estimation of stand volume using spectral and spatial features of aerial photographs and old inventory data. Canadian Journal of Forest Research 32: 1849-1857.

Baatz, M., Benz, U., Dehghani, S., Heynen, M., Höltje, A., Hofmann, P., Lingenfelder, I., Mimler, M., Sohlbach, M., Weber, M. & Willhauck, G. 2002. eCognition - User Guide 3.

Bauer, M.E., Burk, T.E., Ek, A.R., Coppin, P.R., Lime, S.D., Walsh, T.A. & Walters, D.K. 1994. Satellite inventory of Minnesota forest resources. Photogrammetric Engineering and Remote Sensing 60: 287-298.

Blaschke, T. & Strobl, J. 2001. What's wrong with pixels? Some recent developments interfacing remote sensing and GIS. GeoBIT/GIS 6: 12-17.

Burnett, C. and Blaschke, T. 2003. A multi-scale segmentation/object relationship modelling methodology for landscape analysis. Ecological modelling 168: 233-249.

Canny, J. 1986. Computational approach to edge detection. IEEE Transactions on pattern analysis and machine intelligence PAMI-8(6): 679-898.

Gougeon, F.A. 1995. A system for individual tree crown classification of conifer stands at high spatial resolution. pp. 635-642 In*:* Proceedings of 17th Canadian Symposium of Remote Sensing, Saskatoon, Saskatchewan, Canada, June 13-15, 1995.

Definiens 2003. http://www.definiens-imaging.com/ecognition/forester/.

Diedershagen, O., Koch, B., Weinacker, H. & Schutt, C. 2003. Combining Lidar- and GIS data for the extraction of forest inventory parameters. In: ScandLaser Scientific Workshop on Airborne Laser Scanning of Forests. Umeå, Sweden. Swedish University of Agricultural Sciences, Department of Forest Resource Management and Geomatics.Working paper 112. 273 p.

Dong, J., Kaufmann, R.K., Myneni, R.B., Tucker, C.J., Kauppi, P.E., Liski, J., Buermann, W., Alexeyev, V. & Hughes, M.K. 2002. Remote sensing estimates of boreal and temperate forest woody biomass: carbon pools, sources and sinks. Remote Sensing of Environment 84(3): 393-410.

ERDAS Inc. 1994. ERDAS FIELD GUIDE, Third Edition. ERDAS, Inc. Atlanta, GA.

FACT 2004. http://www.falconinformatics.com/

Franco-Lopez, H., Ek, A.R. & Bauer, M.E. 2001. Estimation and mapping of forest stand density, volume, and cover type using the k-nearest neighbour method. Remote Sensing of Environment 77: 251-274.

Franklin, S.E., Wulder, M.A. & Lavigne, M.B. 1996. Automated derivation of geographic window sizes for use in remote sensing digital image texture analysis. Computers & Geosciences 22: 665-673.

Fransson J.E.S., Walter F., and Ulander L.M.H., 2000. Estimation of Forest Parameters Using CARABAS-II VHF SAR Data, IEEE Transactions on Geoscience and Remote Sensing, 38: 720-727

Fu, K.S. & Mui, J.K. 1981. A survey of image segmentation. Pattern Recognition 13(1): 3-16.

Gonzales, R. C. and Woods, R. E. 1993. Digital Image Processing. Addison-Wesley Publishing Company. USA.

Goldberg, M., Schlaps, D., Alvo, M. & Karam, G. 1982. Monitoring and change detection with Landsat imagery. Proceedings of 6th International Conference on Pattern Recognition, Munich, Germany, Oct. 19-22, 1982 1: 523-526.

Haapanen, R., Ek, A.R., Bauer, M.E. & Finley, A.O. 2004. Delineation of forest/nonforest land use classes using nearest neighbor methods. Remote Sensing of Environment 89: 265-271.

Hagner, O. 1987. Remote sensing-aided forest inventory. Helsingin yliopiston metsänarvioimistieteen laitoksen tiedonantoja 19: 130-139.

—, 1990. Computer aided forest stand delineation and inventory based on satellite remote sensing. The usability of remote sensing for forest inventory and planning. Proceedings from SNS/IUFRO workshop in Umeå 26-28 February 1990. Swedish University of Agricultural Sciences, Remote Sensing Laboratory, Umeå. pp. 94-105 .

Halme, M. & Tomppo, E. 2001. Improving the accuracy of multisource forest inventory estimates by reducing plot location error - a multicriteria approach. Remote Sensing of Environment 78: 321-327.

Häme, T. 1991. Spectral interpretation of changes in forest using satellite scanner images. Acta Forestalia Fennica 222: 1-11.

—, Heiler, I. & Miguel-Ayanz, J.S. 1998. An unsupervised change detection and recognition system for forestry. International Journal of Remote Sensing 19(6): 1079-1099.

Haralick, R.M. & Shapiro, L.G. 1985. Survey: Image Segmentation Techniques. Computer Vision, Graphics, and Image Processing 29(1): 100-132.

Holmström, H., Nilsson, M. & Ståhl, G. 2001. Simultaneous estimations of forest parameters using aerial photograph interpreted data and the *k* nearest neighbour method. Scandinavian Journal of Forest Research 16: 67-78.

Holopainen, M. 1998. Forest habitat mapping by means of digitized aerial photographs and multispectral airbone measurements. University of Helsinki, Department of Forest Resource Management. Publications 18: 1-49.

— & Wang, G. 1998. The calibration of digitized aerial photographs for forest stratification. International Journal of Remote Sensing 19(4): 677-696.

Hyppänen, H. 1996. Spatial autocorrelation and optimal spatial resolution of optical remote sensing data in boreal forest environment. International Journal of Remote Sensing 17: 3441-3452.

Hyyppä, J., Hyyppä, H., Inkinen, M., Engdahl, M., Linko, S. & Zhu, Y.-H. 2000. Accuracy comparison of various remote sensing data sources in the retrieval of forest stand attributes. Forest Ecology and Management 128: 109-120.

—, Hyyppä, H., Samberg, A., 1999, Assessing Forest Stand Attributes by Laser Scanner, Laser Radar Technology and Applications IV, 3707: 57-69.

ILMARI 2004. www.ilmari.fi

Jain, R., Kasturi, R. & Schunck, B.G. 1995. Machine vision. McGraw-Hill International Edition.

Katila, M. 2004. Error variations at the pixel level in the k-nearest neighbour estimates of the Finnish multisource National Forest Inventory. In: Controlling the estimation errors in the Finnish multisource National Forest Inventory. The Finnish Forest Research Institute, Research Papers, 910.

Kilpeläinen, P. & Tokola, T. 1999. Gain to be achieved from stand delineation in Landsat TM image-based estimates of stand volume. Forest Ecology and Management 124: 105-111.

Koivuniemi, J. 2003. The accuracy of compartmentwise forest inventory based on stands and located sample plots. Faculty of Agriculture and Forestry. Department of forest resource management. University of Helsinki, Helsinki. 143 p.

Kuusela, K. & Poso, S. 1970. Satellite pictures in the estimation of the growing stock over extensive areas. Photogrammetric Journal of Finland 4(1).

Leckie, D. G. 1987. Factors affecting defoliation assessment using airborne multispectral scanner data. Photogrammetric Engineering and Remote Sensing (53)12: 1665-1674.

Lillesand, T.M., Kiefer, R.W. & Chipman, J. W. 2004. Remote Sensing and Image Interpretation. John Wiley and Sons Inc.

Li, X. and Strahler, A. H. 1992. Geometric-optical bidirectional reflectance modeling of the discrete crown vegetation canopy: effect of crown shape and mutual shadowing. IEEE Transactions on Geoscience and Remote Sensing 30(2): 276-292.

Leica Geosystems 2004. http://gis.leica-geosystems.com/products/

Löfström, K. 1946. Ilmakuvakartoitus Suomessa. The Photogrammetric Journal of Finland 1(1): 78-109.

MacQueen, J. 1967. Some methods for classification and analysis of multivariate observations. Volume 1 of Proceedings of the Fifth Berkeley Symposium on Mathematical statistics and probability, pp 281-297. Berkeley, 1967. University of California Press.

Merriam -Webster Online Dictionary 2004. “segment”. www.merriamwebster.com

METRIA 2004. www.lantmateriet.se.

Mäkisara, K., Heikkinen, J., Henttonen, H., Tuomainen, T. & Tomppo, E. 1997. Experiment with imaging spectrometer data in large-area forest inventory context. Proceedings of the Third International Airbone Remote Sensing Conference and Exhibition. Development, Integration, Applications & Operations 7-10 July 1997, Copenhagen, Denmark II: 420-427.

—, Meinander, M., Rantasuo, M., Okkonen, J., Aikio, M., Sipola, K., Pylkkö, P. & Braam, B. (eds.). 1993. Airborne Imaging Spectrometer for Applications (AISA). 1993 International Geoscience and Remote Sensing Symposium (IGARSS’93). Tokyo, Japan. pp 479-481.

Narendra, P.M. & Goldberg, M. 1980. Image segmentation with directed trees. IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-2: 185-191.

Næsset, E. 1997. Estimating timber volume of forest stands using airborne laser scanner data. Remote Sensing of Environment, 61(2): 246-253.

Nieke, J., H., S., Neumann, A. & Zimmermann, G. 1997. Imaging Spaceborne and Airborne Sensor Systems in the Beginning of the Next Century. The European Symposium on Aerospace Remote Sensing (IEE); Conference on Sensors, Systems and Next Generation Satellites III. London, UK. SPIE3221-71p.

Nyyssönen, A. 1955. On the estimation of the growing stock from aerial photographs. Communicationes Instituti Forestalis Fenniae 46: 1-57.

—, 1962. Aerial photographs of tropical forests. Unasylva 16(1).

Olsson, H. 1994. Monitoring of Local Reflectance Changes in Boreal Forests using Satellite Data. PhD thesis. Research report 8. Department of Forest Resource Management and Geomatics. Swedish University of Agricultural Sciences, Umeå.

Pal, N.R. & Pal, S.K. 1993. A review on image segmentation techniques. Pattern recognition 26(9): 1277-1294.

Parmes, E. 1992. Segmentation of SPOT and Landsat satellite imagery. The Photogrammetric Journal of Finland 13(1): 52-58.

Pellikka, P., King, D. J. & Leblanc, S. G. 2000. Quantification and removal of bidirectional effects in aerial CIR imagery of deciduous forest using two reference land surface types. Remote Sensing Reviews, Special issue on “Multi-angle Measurements and Models”, 19: 259-291.

Pekkarinen, A. & Sarvi, V. 2002. Detection of clearcuttings with help of high-altitude panchromatic aerial photographs and image segmentation. In: Operational Tools in Forestry using Remote Sensing Techniques. ForestSAT Symposium. Edinburgh, Scotland, August 5th-9th 2002.CD-ROM.

Pitkänen, J. 2001. Individual tree detection in digital aerial images by combining locally adaptive binarization and local maxima methods. Canadian Journal of Forest Research 31: 832-844.

Poso, S. & Waite, M.-L. 1996. Sample based forest inventory and monitoring using remote sensing. Remote sensing and computer technology for natural resource assessment. Joensuu. Research notes of University of Joensuu 48.

Saksa, T., Uuttera, J., Kolström, T., Lehikoinen, M., Pekkarinen, A. & Sarvi, V. 2003. Clear cut detection in boreal forest aided by remote sensing. Scandinavian Journal of Forest Research 18(6): 537-546.

Sarvas, R. 1938. Ilmavalokuvauksen merkityksestä metsätaloudessamme. Silva Fennica 48: 2-45.

Saukkola, P. 1982. Monitoring regeneration fellings by satellite imagery. Technical Research Centre of Finland - Reports 89: 1-108.

Schieve, J., Tufte, L. & Ehlers, M. 2001. Potential and problems of multi-scale segmentation methods in remote sensing. GeoBIT/GIS 6: 34-39.

Sell, R. 2002. Segmentointimentelmien käyttökelpoisuus ennakkokuvioinnissa. Metsätieteen aikakauskirja 3: 499-507.

Song, C., Woodcock, C.E., Seto, K.C., Pax Lenney, M., and Macomber, S.A. 2001. Classification and change detection using Landsat TM data: when and how to correct atmospheric effects? Remote Sensing of Environment 75:230-244.

SPECIM 2004. www.specim.fi

Stehman, S. V. 1997. Selecting and interpreting measures of thematic classification accuracy. Remote Sensing of Environment 62: 77-89.

Strahler, A.H., Woodcock, C.E. & Smith, J.A. 1986. On the nature of models in remote sensing. Remote Sensing of Environment 20: 121-139.

Tokola, T., Pitkänen, J., Partinen, S. & Muinonen, E. 1996. Point accuracy of a non-parametric method in estimation of forest characteristics with different satellite materials. International Journal of Remote Sensing 17: 2333-2351.

Tomppo, E. 1987. Stand delineation and estimation of stand variates by means of satellite images. Remote Sensing-Aided Forest Inventory. University of Helsinki, Department of Forest Mensuration and Management Research Notes 19. 60-76 p.

—, 1992a. Multi-source national forest inventory of Finland. Metsäntutkimuslaitoksen tiedonantoja 444: 52-60.

—, 1992b. Satellite image aided forest site fertility estimation for forest income taxation. Acta Forestalia Fennica 229: 1-70.

—, 1996. Multi-source national forest inventory of Finland. New Thrusts in forest inventory, EFI proceedings 7: 27-41.

—, Goulding, C. & Katila, M. 1999. Adapting Finnish multi-source forest inventory techniques to the New Zealand preharvest inventory. Scandinavian Journal of Forest Research 14: 182-192.

—, Korhonen, K.T., Heikkinen, J. & Yli-Kojola, H. 2001. Multi-source inventory of the forests of the Hebei Forestry Bureau, Heilongjiang, China. Silva Fennica 35: 309-328.

—, Nilsson, M., Rosengren, M., Aalto, P. & Kennedy, P. 2002. Simultaneous use of Landsat-TM and IRS-1C WiFS data in estimating large area tree stem volume and aboveground biomass. Remote Sensing of Environment 82: 156-171.

Trotter, C.M., Dymond, J.R. & Goulding, C.J. 1997. Estimation of timber volume in a coniferous plantation forest using Landsat TM. International Journal of Remote Sensing 18: 2209-2223.

Tuominen, S. & Pekkarinen, A. 2004. Local radiometric correction of digital aerial photographs for multi source forest inventory. Remote Sensing of Environment 89: 72-82.

— & Poso, T. 2001. Improving multi-source forest inventory by weighting auxiliary data sources. Silva Fennica 35: 203-214.

Varjo, J. 1996. Detecting manmade forest activities and natural disasters using Landsat TM satellite data - a method presented for controlling continuously updated forest information in Finland. Presented in joint meeting of the Council on Forest Engineering and International Union of Forest Research Organizations Subject Group S3.04.00, Marquette, MI, July 29-August 1 1996: 1-9.

Weszka, J.S. 1978. A survey of threshold selection techniques. Computer Graphics and Image Processing 5: 382-399.

Woodcock, C.E. & Macomber, S. 2001. Monitoring large areas for forest change using Landsat: Generalization across space, time and Landsat sensors. Remote Sensing of Environment 78: 194-203.

—, Macomber, S.A., Pax-Lenney, M. & Cohen, W.B. 2001. Monitoring large areas for forest change using Landsat: generalization across space, time and Landsat sensors. Remote Sensing of Environment 78: 194-203.

— & Strahler, A.H. 1987. The factor of scale in remote sensing. Remote Sensing of Environment 21: 311-332.

Yu, X.W., Hyyppä, J., Rönnholm, P., Kaartinen, H., Maltamo, M. & Hyyppä, H. 2003. Detection of harvested trees and estimation of forest growth using laser scanning. In: ScandLaser Scientific Workshop on Airborne Laser Scanning of Forests. Umeå, Sweden. Swedish University of Agricultural Sciences, Department of Forest Resource Management and Geomatics. Working paper 112. 273 p.

Z/I Imaging 2004. http://www.ziimaging.com/

Zucker, S., W. 1976. Region Growing: Childhood and Adolescence. Computer Graphics and Image Processing 5: 382-399.