Development of a rapid assessment technique to identify human disturbance features in a forested landscape

August 2003

UNIVERSITY OF CALGARY

Development of a Rapid Assessment Technique to Identify Human Disturbance Features in a Forested Landscape

by

Joan Louise Pereverzoff

A DOCUMENT

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF GEOGRAPHIC INFORMATION SYSTEMS

DEPARTMENT OF GEOGRAPHY

CALGARY, ALBERTA

AUGUST 2003

UNIVERSITY OF CALGARY

FACULTY OF GRADUATE STUDIES

Approval Page from Exam

Supervisor, (Dr. Darren Bender, Department of Geography)

(Dr. Mryka Hall-Beyer, Department of Geography)

(Dr. Isabelle Couloigner, Department of Geomatics Engineering)

Date

Abstract

Accurate and up to date inventories of natural and anthropogenic features in northern Canadian landscapes are critical for effective resource and wildlife management. Images derived from remote sensing offer the advantage of providing wide coverage of inaccessible areas at a variety of spatial resolutions. Panchromatic data from satellite images or aerial photographs are abundant, cover a large time period and are available at a range of prices. However, challenges exist in the efficient processing of high spatial resolution panchromatic data. The large volume of data and limited spectral characteristics contribute to lower accuracies in computerized image processing.

This paper examines object-oriented image classification techniques as an alternative to the traditional pixel-based classification techniques. Object-oriented image classification works by first segmenting an image into meaningful image objects that can then be classified with a set of spectral and shape characteristics as well as logical statements (e.g., adjacency rules).

Object-oriented image processing techniques were applied to two types of panchromatic data: a 0.75 m spatial resolution orthophotograph and a 5 m IRS image of a forested landscape in northern Alberta. The orthophotograph was segmented and the natural and anthropogenic features such as clearcuts, wellsites, and roads were classified with a high overall accuracy (88%, $\hat{k}$ = 0.77). The larger IRS scene was more difficult to segment and detect the features of interest (82%, $\hat{k}$ = 0.49). Results indicate that while expert input is required and the learning curve is steep, object-oriented image analysis is an improvement over manual aerial photograph interpretation and pixel-based image classification.

Acknowledgements

Successful completion of this project would not have been possible without the help of:

Dr. Darren Bender
Rebecca Ritson-Bennett
Gordon Stenhouse and the Foothills Model Forest
Dr. Mryka Hall-Beyer
Stefan Lang
Dr. Stephen Franklin
Julia Linke
Dr. Yvonne Martin
Greg McDermid

Dedication

I would like to dedicate this project to Mekefor (Dad), Della (Mom) and Susan Pereverzoff. Thank you for your support. I would never have attempted something so big if you did not encourage me to achieve to my full potential.

TABLE OF CONTENTS

List of Tables

List of Figures

CHAPTER ONE: INTRODUCTION

Resource and wildlife management in Canada depends on accurate inventories of natural and anthropogenic features in the landscape. Natural features such as forests, bogs, grasslands, lakes, rivers, riverbanks, areas of bare rock, and anthropogenic features such as clearcuts, roads, seismic lines, well sites, mines, town sites and trails all contribute to baseline data for further GIS analysis. In Canadian forests, the landscape does not stay static. Natural disturbances such as fire, pest, and forest succession, as well as human disturbances such as logging, mining and oil and gas exploration create a dynamic environment.

The management of landscapes is enhanced when these features of interest are accurately mapped. However, challenges in inventorying the northern Canadian landscape exist because of the remoteness and vastness of the area. The need for current, comprehensive and cost efficient data sets of natural and man-made features in a landscape has driven research into improving feature extraction techniques.

Remotely sensed images, in the form of orthophotographs and satellite images offer an ideal solution because of the wide coverage of difficult to monitor areas and their potentially high spatial resolution. However, limitations in the interpretation of the images has occurred as a result of difficulties in processing large volumes of high resolution imagery, as well as limited spectral separability in orthophotographs and other panchromatic imagery.

Object-oriented image classification is an image processing technique. The technique involves grouping pixels with similar spectral and spatial characteristics to image objects through the process of segmentation. Next, the image objects are classified using a series of logical statements and fuzzy classifiers. Both spectral and spatial characteristics of the image objects are incorporated in the classification rule set. The final product can be a classified raster image or a vector data set that is ready for further GIS analysis.

The first goal of this paper is to fully explore the technique and theories supporting object oriented image classification. This is accomplished by a detailed literature review of the development and current research involving object oriented image classification, with a special focus on the applicability to forested landscapes.

The second goal of this paper is to apply the object oriented technique to a set of panchromatic images in a forested landscape in west central Alberta, specifically in the Foothills Model Forest. A 0.75m spatial resolution panchromatic orthophotograph and a resampled 5m spatial resolution, panchromatic band of an IRS image were classified using the eCognition software.

CHAPTER TWO: LITERATURE REVIEW

2.1. Why a better method for forest classification is needed

Historically, the classification of forest and features within forests was realized through the interpretation of aerial photographs and field observations. Aerial photographs have been used to classify the landscape into homogeneous forest stands (Gillis & Leckie 1996), creating a cartographic base. Descriptive attributes are added to the forest stands through field measurements and the data derived from each aerial photograph is combined to form a regional forest inventory.

More recently, satellite imagery has been added as a source of data for forest classification and forest inventory map updates. Yatabe and Leckie (1995) recognize that satellite remote sensing has gained prominence as a forestry applications tool. It can be used to determine individual forest stand characteristics as well as natural and anthropogenic changes over time. Satellite imagery and aerial photographs are also used for the mapping of corridors and other linear features within forested areas (Innes & Koch, 1998). Innes and Koch (1998) feel that remotely sensed data are ideally suited for determining linear feature characteristics such as length and width of edges, orientation, connectivity and the nature of adjacent areas.

In Canada, over 340 million hectares of forested land have an inventory system (Gillis & Leckie, 1996). But landscapes are complicated systems (Blaschke & Hay, 2001) that require frequent updates for harvests, burns, insect and disease, silviculture and roads (Gillis & Leckie, 1996). Due to the increased use of the forest inventory and the demand for accurate and up-to-date forest inventory information (Gillis & Leckie, 1996), better methods for extracting landscape information are needed.

While aerial photographs contain a wealth of information, manually interpreting them is a labour intensive and expensive process (Hall, Dams, and Lyseng, 1991; Blaschke & Hay, 2001). Manual interpretation can also lead to inconsistent results when generated by different interpreters (Blaschke & Hay, 2001). Holopainen and Wang (1998) identify the need for simpler and more automated methods for forestry classification. Gillis and Leckie (1996) also recognize that organizations are searching for more efficient and effective inventory methods. They see the implementation of forest management inventory into Geographic Information Systems (GIS) as an opportunity to develop and use new technologies for forest change detection. Hall *et al.*

(1989) also associate the increasing use of GIS and remote sensing technology with an inexpensive and less labour-intensive method for updating forest cover maps.

The newly developed technique of object-oriented image classification has excellent potential to radically improve forest inventory updates. Object-oriented image classification can be defined as the processing of spatially contiguous homogeneous image segments (Hofmann, 2001c). Object-oriented image classification deviates from the traditional pixel based classification because pixels are first grouped into meaningful objects through the process of image segmentation, and are classified not only by spectral information but on neighbourhood relationships, orientation, texture and shape properties as well. Fifteen years ago, Tomppo (1987) concluded that it was possible to develop an object-oriented classification methodology for stand-level forest inventories based on the segmentation of satellite images. Since then, the object-oriented software eCognition (Definiens Imaging, 2001) has been successfully applied to forest and biotype monitoring, monitoring of conservation areas, maps and tree cadastres (Neubert, 2001; Flanders *et al.*, 2003). Object-oriented software has also been used to generate forest fire damage maps (Willhauck, Benz & Siegert, 2000) and update land use datasets including mixed terrain landscapes (Fuller, Groom & Jones, 1994). Object-oriented classification makes computer aided aerial photograph interpretation more accurate due to the incorporation of the spatial attributes of the image (Willhauck *et al.*, 2000). It also offers the potential for automation and trouble-free modification (Willhauck *et al.*, 2000), thus providing an inexpensive and less labour-intensive option for forest inventory updates.

This paper will present the main techniques for object-oriented classification, including an introduction to image segmentation, fuzzy logic classification rules, and automation. Next, this paper will discuss when object-oriented image classification is preferable to per-pixel classification, as well as the problems and limitations with object-oriented classification.

2.2. Techniques for object-oriented image classification

Research into the development of object-oriented image classification programs began as early as the 1970's (Blaschke & Strobl, 2001). A segmentation program called Machineseg was developed in 1984 that used shapes, sizes and spectral data in an image to grow regions of similar characteristics (McKeown, 1998). In the late 1980's, the object-based ARF (A Road Finder) program was capable of using multi-level, object-based algorithms to identify roads from digital, high-resolution aerial photographs. Other programs used segmentation as a means to detect roads, rivers, and field boundaries in Synthetic Aperture Radar images (Quegan *et al.*, 1988).

The early models of object-oriented image classification were limited because they were not able to fuse data for multi-level analysis, produced conflicting results, and were not able to perform updates or become automated (Pereverzoff, 2001). There were also obstacles such as inefficiency in processing. Hardware limitations, unintelligent software, poor resolution of images and huge numbers of possible solutions (Blaschke & Strobl, 2001) also prevented object-oriented image classification from replacing the pixel-based classification standard.

By the late 1990's, hardware capabilities increased dramatically and several commercial and government organizations launched satellite sensors with high spatial resolution (de Kok *et al.*, 1999). For example, SPOT was released in 1986 with 10m panchromatic bands, and 20m multispectral, the Indian Remote Sensing (IRS) with 5.8m panchromatic band and 23.5m multispectral and the Ikonos with 1m panchromatic, 4m multispectral were launched in 1999. Most importantly, in depth research and development of image segmentation algorithms, fuzzy logic and other intelligent programs played a key role in advancing object-based analysis. By 2000, a user-friendly, multi-scaled, fully functional, object-based processing software program called eCognition was available from Definiens (Blaschke & Strobl, 2001).

Current methods of object-oriented image analysis are based upon the classification of homogeneous image regions (known as objects) created through image segmentation (Willhauck *et al.*, 2000). With eCognition software, the procedure for object-oriented image classification can be divided into several key steps: image segmentation at multiple scales, creation of classification rules, classification-based segmentation, accuracy assessment and automation.

2.2.1. Image segmentation

Segmentation is the division of an image into regions (Schiewe *et al.* 2001). The goal of image segmentation is to divide an image into distinct, spatially continuous, homogeneous regions. This step marks a definite separation from pixel-based image classification in that the regions derived from the segmentation become the basis for classification. Correctly segmenting an image is very important so that the regions can accurately represent the real world objects. For example, a very high resolution image could be segmented so individual buildings or roads are separated correctly, creating one image segment per building or road.

An image can be segmented using different segmentation algorithms. The choice of the segmentation method is crucial to reduce the large number of possible solutions. Image segmentation can be divided into two main categories: data-driven and knowledge-driven approaches (Definiens Imaging, 2001).

Data-driven methods are referred to as a bottom-up approach because the methods are applied to the entire image based on a set of statistical methods and parameters (Definiens Imaging, 2001). Because *a priori* knowledge of the real world objects is not included in the segmentation process, the bottom-up approach is considered to be a means of data abstraction or data compression. The bottom-up approach has been criticized because at the end of the segmentation, the image segments have no meaning, *i.e.* they do not represent real-world objects (Definiens Imaging, 2001).

Data-driven segmentation methods include edge-based and region growing (Blaschke & Strobl, 2001) as well as texture-based (Willhauck *et al.*, 2000) algorithms. The edge-based method segments an image by determining where the boundaries are between groups of relatively homogeneous pixels. The edge-based method finds edges between regions using filters, enhancements, edge detection and localization (Pekkarinen, 2002). The region-growing method of image segmentation finds groups of homogeneous pixels and grows the region until pixels with significantly different DN's are located adjacent to the region. Region- growing segmentation involves growing, merging, splitting and combining until a threshold of dissimilarity is reached (Pekkarinen, 2002).

Texture-based segmentation algorithms derive characteristic features from texture measures like spatial frequencies, co-occurrence matrices, fractal indices and others, then group them together using clustering functions.

Knowledge-driven approaches to image segmentation received their name because they incorporate information derived from training areas into the segmentation approach. The knowledge-driven approach can be as simple as merging all adjacent pixels of the same class together into an image object (Definiens Imaging, 2001). Typically, this method uses only spectral information or filtered images, using processes like sharpening and edge enhancement to initially segment the image.

With eCognition, segmentation is largely knowledge-free (Definiens Imaging, 2001). The segmentation is multi-scale, and can incorporate data from several input channels at once. It works well for textured or low contrast data including radar and very high spatial resolution imagery. Segmentation algorithms in eCognition seek to minimize the average heterogeneity of the image segments, weighted by their size. Image segments with more pixels can tolerate a higher heterogeneity than smaller sized image segments. The procedure begins with single pixel objects that use region-merging algorithms. During each iteration the smaller objects are merged together until the size-weighted heterogeneity threshold is exceeded.

To customize the segmentation process in eCognition to produce an appropriate set of image segments, the layer weights, image object level, scale parameter, segmentation mode, homogeneity criterion and type of neighbourhood must be set by the user with the segmentation graphical user interface in eCognition (Definiens Imaging, 2001).

The weight of the image channels is based on the number of input channels containing spectral or other information. The stronger the object of interest can be separated in a channel, the stronger the channel should be weighted (Hofmann, 2001).

The scale parameter is used to determine the object size (Hofmann, 2001). It has been criticized as being an abstract value with no direct correlation to the object size (Willhauck *et al.*, 2000). In fact, the scale parameter is based on the heterogeneity limit for objects. For heterogeneous data, the objects created at a certain scale will be smaller than for homogeneous data. In general, the larger the scale factor, the larger the average size of image segments.

The homogeneity criterion is an indication of the minimum heterogeneity allowable in the image segments. It can be adjusted according to the proportion of influence that the colour versus the shape have in determining the homogeneity of each segment. If the colour proportion is less than the shape proportion, than the spectral value homogeneity is less influential than the shape homogeneity in determining the image objects. The shape proportion can be further subdivided into smoothness and compactness proportions. The smoothness parameter is used to segment the image into objects with smooth borders as opposed to frayed borders. The compactness parameter indicates how compact or spread out the image object shapes can be. The compactness and the smoothness parameters combine to determine the shape parameter for the homogeneity criteria.

When the homogeneity criteria is calculated based on a single band of image data, the criteria uses only that band's spectral information. In the case of multiband imagery, the weighted mean of the image bands is used. The weight of influence of each band is assigned in the layer weight section of the segmentation criteria.

Specifying the type of neighbourhood (diagonal, also known as "8 corner" or planar, also known as "4 corner") determines whether pixels can be included in an object segment or not. The 4 corner pixel neighbour looks for homogeneity in the pixels that share a common plane border (4 pixels) whereas the diagonal pixel neighbourhood looks for homogeneity in the plane border or the corner points (4 plane plus 4 corner pixels). Figure 1 is a diagram of the eligible pixels for segmentation.

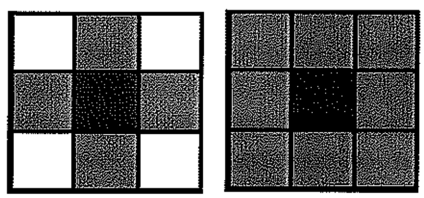

Figure 1. 4 corner pixel and 8 corner pixel neighbourhood (eCognition Users' Manual).

The grey squares indicate pixels that are considered to be joined together in creating image segments.

It is recommended that to achieve the best segmentation for an image, a series of trial segmentations must be performed on a subset of the image, varying the above parameters, until a suitable segmentation is achieved (Neubert, 2001). Through *a priori* knowledge of the characteristics of the real world objects (*i.e.* clearcuts are a certain size, with a certain shape) as well as evaluation of previous segmentation results, each segmentation parameter value can be adjusted independently, or in coordination with other parameters, until the image objects are satisfactory. The ease with which the correct segmentation parameters can be determined is dependent on the operator's knowledge of the software as well as knowledge of the attributes of the objects of interest. The segmentation procedure is repeated at different segmentation levels to produce a hierarchical network of image objects of different sizes (Willhauck *et al.*, 2000a), (see Figure 2 in Section 2.2.2.). Each object "knows" its relationship to neighbour sub- and super-objects (Willhauck *et al.*, 2000a and Willhauck, *et al.*, 2000b). The objects' "knowledge" of its neighbours is due to the creation of topology in the segmentation process. The spatial extent of each object is calculated and references are recorded, indicating that the adjacent objects on the same segmentation level are neighbours. The objects that are contained within the spatial extent of the super-object at the above segmentation level are referenced as its sub-objects. In eCognition, all image objects are linked to the network, providing valuable context information for classification purposes (eCognition White Paper).

2.2.2. Multiple scale/resolution analysis

Object-oriented image classification can be made more accurate by segmenting the image several times, each time adjusting the scale factor to produce objects at multiple spatial resolutions. A single solution of image segments will never represent meaningful objects at all scales and for all applications (Blaschke & Strobl, 2001), therefore it is necessary to design a multi-scale image segmentation approach to insure that each object of interest can be segmented correctly. For forestry applications, the typical spatial object can range from the scale for large forest stands to individual tree crowns (DeKok *et al.*, 1999). Research by Innes and Koch (1998) support the idea that remote sensing is useful for evaluating ecosystem diversity at the landscape level as well as at a number of smaller scales. Community characteristics such as patterns, extents, disturbances and species can be classified using remotely sensed data.

When an image is segmented at varying values for the scale parameter, a network of image objects is created within each resolution scale (Definiens Imaging, 2002). Each layer of objects is linked to the layer above and below, creating a hierarchy (Figure 2).

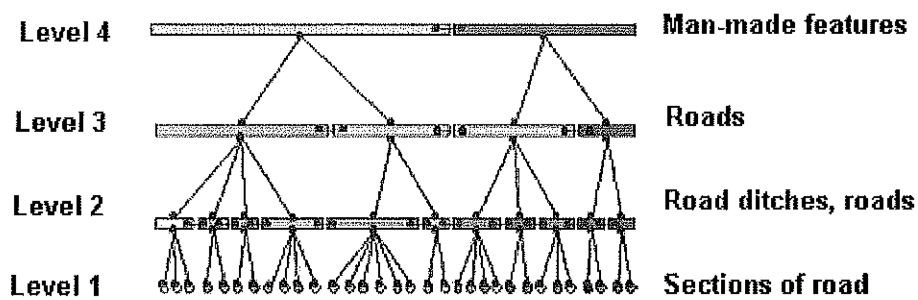

Figure 2. Hierarchical network of image objects (taken from page 3-28 of eCognition User Guide)

This hierarchy allows for the evaluation of context and dependency among levels (GISCafe, 2001). At a fine scale, small segments could represent trees and small objects while at a coarser scale, larger segments could represent large forest and water polygons. The network of objects will create a dependency between objects. The tree image segments will belong to the larger forest polygon.

The different levels of segmentation can be created following a top-down or bottom-up approach. In the case of a top-down created multi-level segmentation, the first level created contains larger objects, which are sub-divided in lower levels of the hierarchy. In a bottom-up segmentation, the smaller objects are created first, then are aggregated into larger objects in the higher levels of the network. Geneletti and Gorte (2002) suggest high resolution imagery can be used to segment image objects at a fine scale while poorer resolution satellite imagery can be used as a reference at a coarse level in the segmentation process.

2.2.3. Classification rules

Whereas pixel-based classification assigns each pixel to a class independently, object-oriented image classification classifies image segments that are composed of several to many pixels. Summary data are extracted from the image segments including mean spectral value, standard deviation, and other information. The summarized data, and not the individual pixel values are used to classify each image segment.

Classification methods for remotely sensed imagery could be either supervised or unsupervised (Lillesand and Kiefer, 1994). In supervised classifications, the user must input the properties of the desired classes. In contrast, an unsupervised classification uses statistical measures of the feature space (multi-band) or image histogram (single band) to separate objects into classes, requiring less user input.

Classification in eCognition is supervised and occurs after the image is segmented. The methods can be divided into sample-based and rule-based supervised classification (Definiens Imaging, 2002). Sample-based classification uses training samples in nearest neighbour clustering. The user selects the sample objects for each class and nearest neighbour clustering is used to assign the image objects to their appropriate classes (Hofmann, 2001).

The rule-based supervised classification uses a very different approach for image classification. For each feature a set of classification rules can be designed. The spectral value, form, size, texture as well as neighbourhood and hierarchy relations can be used in the classification (Meinel *et al.*, 2001; Neubert, 2001; Schwarz *et al.*, 2001). Classification rules can integrate linguistic expert knowledge or can incorporate fuzzy membership functions of an object to a class (Definiens Imaging, 2002). Expert knowledge rules include statements using the object's context, such as "roads do not cross houses" or "riversides border on ditches" (Hofmann, 2001a). These classification rules are able to utilize relationships between objects. Quegan *et al.* (1998) describes the incorporation of contextual information, especially size, shape, texture of regions and the direction or adjacency of objects, as a valuable tool for producing land-use classification from remotely sensed images.

For example, Hofmann (2001b) used many classification rules to improve his object-oriented image classification. Shape properties such as the mean size as well as spectral characteristics like the variability of colour of small objects and the average difference to neighbouring objects were incorporated. Contextual information was also used, including higher relative frequency of small shadows and objects with smaller relative size, to identify vegetation patches.

The linguistic rule that "shadows are long, thin and dark" (Quegan *et al.*, 1988) can be formulated into a classification rule in eCognition by defining the shadow class by its shape and size (long), its high length to width ratio (thin), and its spectral signature (dark). This definition of shadows was appropriate for the image Quegan used, however it cannot be applied universally for all shadow classes.

The fuzzy membership functions describe intervals of the characteristic to which an object would be likely or unlikely to belong (Hofmann, 2001). These membership functions can be assigned values of zero, low or high to describe an object that is not typical, less typical, or typical for a class. Logical operators such as "or", "and" and "not" can be included in classification rules. For example, a fuzzy membership function for the spetral range of bright roads would be a value of 1 for very bright spectral values and decrease to 0 membership for very dark spectral values.

DeKok *et al.* (1999) state that fuzzy membership rules provide the proper tool to group spatial objects into meaningful classes as well as reduce the complexity of classification to a condensed, clear set of end-membership functions. They also promote the fact that the rule-based classification produces transparency of the decision rules (Willhauck *et al.*, 2000a).

The classification rules based on expert knowledge and fuzzy membership functions are combined to build a substantial knowledge base (Definiens Imaging, 2002) that will create a meaningful classification.

2.2.4. Classification-based segmentation

Recalling that the initial segmentation of an image in eCognition is based on colour and form characteristics, it is possible to refine the shape and size of the image objects with classification-based segmentation. The segments created in the classification-based segmentation procedure are classified using knowledge-based rules. This segmentation procedure is suitable for the initial classification because the goal is to reduce the information in the pixels to image objects. The initial objects created may not represent meaningful real world objects; therefore a second segmentation is needed that is based on knowledge of the real world objects.

eCognition has devised a classification-based segmentation that incorporates the information in the knowledge base. The classification-based segmentation refines the objects shapes and borders. It also splits or aggregates image objects (Definiens Imaging, 2002).

eCognition has two tasks to enable classification-based segmentation. The first task is classification-based fusion. It merges adjacent objects that are in the same class. The second task is shape correction, where sub-objects that do not belong to a super-object are regrouped to their correct super-object. Shape correction can be done by border optimization. This procedure regroups sub-objects under the wrong super-object to become a part of an adjacent super-object if the super-object is correct for the sub-object. Shape correction can also be performed by the extraction of sub-objects. In this technique, all sub-objects are resegmented to their correct super-object, regardless of the sub-object's adjacency to super-objects.

2.2.5. Accuracy assessment

The quality of the image classification can be assessed in eCognition. It is possible to compute statistics for image objects as well as relational features not used in the classification themselves, such as embedding or distances to other classes. Typical statistics for accuracy of the classification such as confusion matrices, as well as fuzzy classification can be evaluated with eCognition. The output can be tabular or graphical. Tables can be exported in comma-separated text files and the graphical results can be exported as raster files.

Basic statistical operations such as the sum, mean, standard deviation and range (minimum, maximum) of areas for selected classes and features can be computed. Histograms of frequency distributions of digital numbers of individual image channels can be plotted, enabling display of the mean value and standard deviation of the image layer.

Advanced accuracy assessments available in eCognition are measures of best classification result, classification stability, error matrices based on a training or test area (TTA) mask, and error matrices based on samples. The first two advanced measures indicate the spatial distribution of the fuzzy classification stability, also known as the membership value. In fuzzy classification, objects can belong to several classes but with different degrees of membership. The first advanced accuracy assessment tool is the best classification result, that is a display of the membership value assigned to each image object. The class with the highest assignment value is displayed in the colour range of dark green for a perfect assignment (1.0) to a red for a poor assignment (0.0). The best classification result tool also allows tabular output with the number of objects per class, the mean, standard deviation and range of best classification assignments.

The second accuracy assessment tool is the classification stability measure. It is a measure of the difference of the degree of membership of an object between the best and the second best class assignments. This measure gives an indication of the ambiguity of classification. An object can meet the criteria for more than one class. When it is a high degree of membership to several classes, the classification is ambiguous. A useful accuracy assessment is to calculate the difference between the best and second best memberships (μ values). The higher the difference, the less ambiguously the object belongs to a class. The display of the classification stability can be visual, ranging from dark green for an unambiguous classification ($\mu = 1.0$) to red for an absolutely ambiguous classification ($\mu = 0.0$). The tabular output is a summary per class of the number of objects, the mean, standard deviation, and range of the difference between membership values.

The information gained from the best classification result and the classification stability measures is extremely useful because one wants to define classes as unambiguously as possible. This information can be used to modify membership functions so there is less overlap in value ranges, or define other features that are capable of separating objects.

Two types of error matrices can be computed as accuracy assessment measures, the error matrix based on a TTA mask and the error matrix based on samples. The error matrix based on a TTA mask uses test areas as a reference for classification quality. The TTA mask is created from the individual pixels in the classified image. It is also possible to import an externally created TTA mask. The statistical output is a confusion matrix with user's and producer's accuracy, overall accuracy, Hellden, Short and KIA (Kappa Index of Agreement) statistics.

The overall accuracy is the proportion of all reference pixels that are classified correctly. It is the sum of the diagonal entries of the confusion matrix divided by the number of reference pixels. The producer's and user's accuracies are used to determine the classification accuracy of individual classes. The producer's accuracy, the complement of the errors of omission, gives a measure of how well the classification agrees with the true classes, while the user's accuracy, the compliment of the error of commission, gives a measure of how well the classification predicts a class. Short's accuracy is the product of the producer's and user's accuracy. Hellden's accuracy measure is the harmonic mean of the producer's and user's accuracy. Short's accuracy measure is more pessimistic than Hellden's accuracy measure. Finally, Cohen's KIA assumes that both the image classification and reference classification are independent class assignments of equal reliability, and measures how well they agree (Jensen, 1996). The KIA can be used to evaluate the statistical significance of the differences between two classification attempts.

The error matrix based on samples gives a measure of the quality of classification based on sample object, not individual pixels as with the TTA mask. It is not recommended to use the same objects for the error matrix as ones that were used in the nearest neighbour classification. The error matrix tool produces a table indicating the number of samples per class and the proportion of the samples that are assigned to the correct class and to each of the other classes.

2.2.6. Automation

Bird, Taylor and Brewer (2000) used aerial photography as their main data source in a land use classification project, but lamented that its interpretation was very labour intensive and expensive. Other researchers agree (DeKok *et al.*, 1999) that the cost of visual interpretation is the bottleneck in image interpretation. They believe that considerable improvements in the automation of image analysis are necessary to reduce the costs of image interpretation. And because the amount of images taken has increased, automation is the only practical solution to extract important information from the images.

In an assessment of Canadian forest inventory update techniques, Gillis and Leckie (1996) acknowledged that automatic classification of remotely sensed digital images works reasonably well, saving time in delineation and digitizing. Visual confirmation is still required when using current methods of automated image interpretation.

When a stable sequence of processing steps for image interpretation has been arranged in eCognition, the protocol automation tool may be used to classify a series of images. All major steps of the classification process can be included in a protocol file (Willhauck *et al.*, 2000b) and can be automatically applied to an image, assuming the number of thematic layers and image objects levels is consistent throughout the images (Definiens Imaging, 2001). Operations such as multiresolution segmentation, classification-based segmentation, customized features, class hierarchies, classification rule base, polygon creation, accuracy assessment results and image object exports can be scripted to run with the protocol tool (Definiens Imaging, 2002).

2.3. Potential limitations in the software

Most evaluations of the object-oriented image analysis eCognition software have been positive, but some researchers have expressed concerns over limitations in the type and amount of data that can be analysed as well as processing errors.

Meinel *et al.* (2001) describe a major disadvantage of eCognition: the long computing time involved in processing images. For example, an IRS scene with 40 million pixels could take upwards of 24 hours to segment and build polygons. It is necessary to subset larger scenes for initial classification then apply the classification scheme to the whole scene. The disadvantage in this approach is that features and properties of the whole scene may not be represented adequately in the subset image. If the whole scene or a larger subset is used to improve the initial classification then the processing time increases dramatically. Therefore, handling of huge data sets is not practical at this moment (Neubert, 2001).

These researchers also found that in general, heterogeneous areas in images take more time and are less securely classified than large, homogeneous area. Neubert (2001) noted that a limitation of eCognition is that only raster data of the same scale can be used, so images must be resampled to the same spatial resolution in order to be incorporated in the image analysis. The resampling could potentially reduce accuracy and introduce errors.

It has been noted that classification errors, namely misclassifications, can be introduced in the processing of images with eCognition. Schwarz *et al.* (2001) found that errors appeared in the classification of the borders of damaged areas in a scene, as well as ofsteep sloped and other shadowed areas.

Schiewe, Tufte and Ehlers (2001) found it difficult to automate the segmentation process because many repetitions of segmentation were necessary to determine the correct scale parameters and the suitable weights for the input data. The task of determining these parameters is described as being non-trivial because there are virtually infinite segmentation solutions (Blaschke& Hay, 2001). Schiewe *et al.* (2001) also found that determining the classification rules with fuzzy membership was a costly process that could not be easily determined.

DeKok *et al.* (1999) raise concerns that eCognition requires the operator to be an expert on the objects of interest. A forester needs to know the ecology of the area to determine the proper influence of the spatial and spectral characteristics in classifying the objects of interest. They conclude that this form of object-oriented classification expects a high input of the field expert to correctly define the decision functions with fuzzy logic.

The lack of transparency in the segmentation process is a clear disadvantage to this technique. The lack of information on how the segmentation algorithm works adds a considerable amount of time to the process of segmentation because the user must infer the algorithms behaviour based on the results of each trial of segmentation.

2.4. When Object-oriented is desirable over per-pixel classification

Object-oriented image processing techniques are more intuitive than per-pixel classification techniques, incorporating spatial theories and identifying landscape patterns and features in complex scenes. Object-oriented classification incorporates more than spectral values to improve classification, mimics pattern recognition by humans, can handle large volumes of data, reduces the speckle effect in SAR imagery, and can produce raster or vector GIS-ready data.

2.4.1. Pixel-based classification uses only spectral values

Classification based on image objects is a superior method when compared to pixel-based techniques when spectral values cannot be classified unambiguously (Meinel *et al.*, 2001). eCognitions claim to have a higher classification accuracy and better class differentiation versus conventional pixel-based methods because in comparison to pixels, image objects carry more useful information (Definiens Imaging, 2002) in addition to spectral values, or information derived from spectral values, form, texture, neighbourhood and context information. Willhauck *et al.* (2000b) assert that important semantic information, often necessary to accurately classify an image, is found in image objects and their mutual relations, not in single, independent pixels.

Neubert (2001) was able to differentiate more land use classes with object-oriented classification by including spectral as well as other types of information in the classification. In his case and other classification situations, per-pixel classification is at a disadvantage because it does not make use of spatial characteristics and spatial relationships in the image (Blaschke & Hay, 2001). As well, pixel-based classification results in more isolated, disconnect pixels that cannot be clearly assigned to any class (Meinel *et al.*, 2001). When Neubert (2001) classified images for land use class, eCognition yielded classifications that were more accurate. The classification yielded more homogenous regions and had a smaller portion of unclassified areas compared to pixel-based classification methods.

When spectral values are relied on as the main source of information for image classification, misclassification can occur for several reasons. First, boundaries that cross pixels can be misclassified because the pixel representing the boundary contains the reflectance values from two or more land cover types (Smith & Fuller, 2001). With object-oriented classification, the boundary pixel could be classified based on contextual information, shape or connectivity characteristics. Second, groups of pixels that represent the same land cover type may not have the same spectral value due to noise, atmospheric effects and natural variation within the land class (Smith & Fuller, 2001). In comparison, object-oriented classification uses the averaged spectral values and the standard deviation of the spectral values of the group of pixels to compensate for the natural variation in the spectral data. Third, pixels do not represent the actual landscape structure that is supposed to be classified. Pixels are an arbitrary grid system whereas the image objects have shape and form that mimic landscape structures. Using the shape information in classification rules adds distinguishing data to improve classification. Finally, pixel-based classification that relies heavily on the spectral value of individual pixels does not incorporate the spatial context of the pixel, thus depriving the classification of the use of neighbourhood information.

2.4.2. Similar to how humans identify patterns

"Human perception does not observe, nor do we actually think in pixels" is a strong statement made by Blaschke and Strobl (2001:16) in their criticism of pixel-based image classification. Aerial photographs are not classified manually on a per-pixel basis. Instead, interpreters delineate ecological boundaries or create polygons of similar land use classes. While manual interpretation of ecological patches is common, monitoring needs are increasing, thus more results from manual interpretation will be expensive and more time consuming than an automated procedure that does not have such a high human interpretation input (Blaschke & Hay, 2001). Object-based image classification claims to simulate manual interpretation, but at lower costs by incorporating several characteristics of human perception. Human perception is capable of discerning features in natural scenes that are not limited to the square shape of pixels. Second, human perception is able to generalize one area and see details in the data of another area simultaneously. Third, human perception can detect patterns and texture readily.

In sciences like landscape ecology and forestry, the ecological phenomenon being studied transcends the boundaries of pixels. The statistical analysis of pixels does not use the correct areas for analysis based on fundamental ecological concepts. Geneletti and Gorte (2002) describes the pixel-based system as a model of the terrain that does not match with geographical objects such as land parcels and water bodies. This is because a land class can rarely fit within the constraints of a single pixel or a single scale of analysis. It is more likely that the land use classes will be discernable at a range of scales, extending over several pixels in an organic shape not in the rigid grid shape of square pixel filters.

The shape of an image can be characterized by several properties, extending its boundaries to create more meaningful image objects. The size of the object, direction, asymmetry and length to width ratio are shape properties that can be used in eCognition to produce more representative features in an image than analysis of each pixel independently. Pixels and box shaped filters only have their size as a meaningful shape characteristic.

Human perception is also capable of generalizing an image into homogeneous areas and can detect discontinuities to determine where one ecological region should be distinguished from another (Blaschke & Hay, 2001). When an image is first separated into homogenous areas, then details within the areas are discerned, the result is more intuitive features (Blaschke & Strobl, 2001). Object-oriented image analysis attempts to include these human perception tools to improve image classification. Smith and Fuller

(2001) used object-oriented image analysis techniques to smooth unwanted variation in the data. Pedley and Curran (1991) found that object-based analysis reduced the internal variability of land parcels by averaging the spectral values within an image segment, much like the way human perception smoothes variability to delineate regions of similar characteristics. They found that the classification accuracy increased significantly with more spatially variable land cover classes. The classification time decreased and mixed classes within fields were eliminated (Pedley & Curran, 1991).

Schwarz *et al.* (2001) maintain that the pixel-based approach has reached its operational limit for forestry because it is not easily able to deal with internal variance and the high level of texture in images. Object-oriented image analysis produces a less variable classification and more intuitive classifications while requiring less user input, time and expenses.

2.4.3. Can handle the large volume of data in high resolution images

The use of very high resolution (VHR) imagery as the primary data source in classifications can improve the identification of features in an image, hence improve the overall classification accuracy. As a result, VHR images have many more pixels than lower spatial resolution imagery covering the same area. The disadvantage of using VHR imagery is that this enormous volume of data that needs to be processed. The large volume of data creates a strong need for methods that can exploit the data effectively (Blaschke & Strobl, 2001; Pekkarinen, 2002). Pekkarinen (2002) asserted that pixel-based analysis methods are not applicable for VHR images in multi-source forest inventory applications. He offers the solution of reducing the volume of data by analysing image units larger than the single pixels through image segmentation.

The segmentation and classification of image objects greatly reduces the number of image elements for classification and significantly reduces the volume of data that needs to be processed in each step (Meinel *et al.*, 2001). Schwarz suggests that eCognition software handles the high level of detail and highly variable data better than pixel-based classification algorithms. He found the pixel-based approach was limited for processing VHR images.

Data reduction through image segmentation enables faster processing time for image classification once the segmentation parameters have been finalized. Reducing the volume of data through image segmentation is a valid option based on the understanding that it is very likely that a pixel will contain the same spectral value as its neighbours (DeKok,*et al.*, 1999). Based on this concept, it is possible to reduce the large volume of pixels into representative image objects without reducing the classification accuracy.

2.4.4. Reduces the salt/pepper effect

Quegan *et al.* (1998) describes the image classification of images with speckle as the first obstacle to overcome in any attempt to interpret SAR images automatically. Willhauck, *et al.* (2000b) suggest that speckle effects can be overcome by segmenting the image before it is classified. This procedure ensures that pixels near one another will most likely be classified into the same class because the pixels are grouped into objects in the segmentation process.

In an experiment conducted on SAR imagery, Willhauck *et al.* (2000b) found that there was the same overall accuracy for the pixel and the object-based classifications, but their appearances of the resulting maps were different. The pixel-based classification had a salt and pepper effect whereas the object-based classification had a smoother classification because the image was segmented before classification.

When an image is first segmented into image objects, Meinel *et al.* (2001) found a lower proportion of pixels were left unclassified and a far more homogeneous product was the result.

2.4.5. GIS ready

To simplify the export of the classified remotely sensed data, the image objects are converted to a polygon vector layer (DeKok *et al.*, 1999) with an identifier and a series of attributes including coordinates, area and degree of membership to a class. The advantages of having image objects exported as polygons means that the geo-information is immediately usable (Geneletti & Gorte, 2002) in a GIS program. Blaschke and Strobl (2001) praise object-oriented image analysis techniques as a critical bridge between GIS and remote sensing. Many projects use results gained from analysing remotely sensed images for further analysis in GIS programs. Smith and Fuller (2001) found that the results from object-oriented image analysis are suitable for incorporation with existing data sets, easing the transition from raster- to vector-based data.

In fact, the relationship between remotely sensing and GIS can be the inverse as well. GIS objects can be used directly in remote sensing software packages as a contributing layer in the analysis of an image. In eCognition, vector layers can be added to assist in delineating image object boundaries and in their classification.

A further advantage of the strong link between remote sensing and GIS in object-oriented image analysis is that the GIS objects are synthesized without ignoring the geographical concepts of relationships and scales (Blaschke & Strobl, 2001).

2.5. Conclusion

Object-based image classification of forested scenes offers several clear advantages over pixel-based classification and manual interpretation of aerial photographs. Through image segmentation, fuzzy logic classification rules and automation of procedures, object-based classification is able to produce highly accurate delineations of forest disturbance features. Object-oriented software programs, such as eCognition, have the potential to overcome difficulties with large data sets in VHR imagery, reduce the speckle effect in SAR imagery, produce meaningful image objects by simulating human perception abilities, and produce GIS ready outputs.

In many forestry application projects, object-oriented image classification is preferable to per-pixel classification because data can be analysed more efficiently, saving both time and financial resources. At present, there are limitations to the object-based classification software, eCognition, including the need to subset larger images while determining segmentation parameters, resample data of different spatial resolutions and difficulties in determining fuzzy classification rules. However, these are practical rather than theoretical limitations. Once these limitations are overcome, it is very possible that object-oriented image classification will become the preferred choice for forest disturbance inventories and updates.

CHAPTER THREE: METHODS

The goal of the technical part of this study was to apply object oriented image classification to two types of panchromatic data to produce highly accurate maps of major natural and anthropogenic features in a forested landscape in northern Alberta. A digital, panchromatic orthophotograph with very high spatial resolution (0.75m) and a resampled IRS image (5.0m) were used for this study. The analysis was conducted on a subset of the orthophotograph and the IRS image as well as a full IRS scene. The purpose of subsetting the IRS scene was to examine the segmentation of a smaller scene and to compare the results with the subset orthophotograph of the same area.

The main methodology was to first identify image enhancement and edge detection filters to assist in the segmentation and classification process. A DEM was linked to the larger IRS scene to facilitate classification. The images were segmented using eCognition software at a range of levels to delineate objects of interest. The images were then classified using a range of spectral and shape characteristics, as well as logical statements. Accuracy assessment was performed on a random sample of points in the image. Field data collected near the time of year that the images were acquired as well as photo-interpretation were used in the accuracy assessment. The classified images were exported as raster and vector data sets.

PCI software was used for image filter calculation and the overall accuracy assessment of the classified images.

3.1. Study Area: Foothills Model Forest

The study area is situated within the Foothills Model Forest, near Hinton, Alberta. The spatial extent of the study area is UTM Zone 11 (Upper Left) 465466.75m E, 5900873.03m N, and (Lower Right) 500736.46m E, 5871468.30m N. It is characterized by boreal, montane and sub-alpine forest types (Foothills Model Forest Network, 2003). Figure 2 depicts the location of the Foothills Model Forest in Alberta.

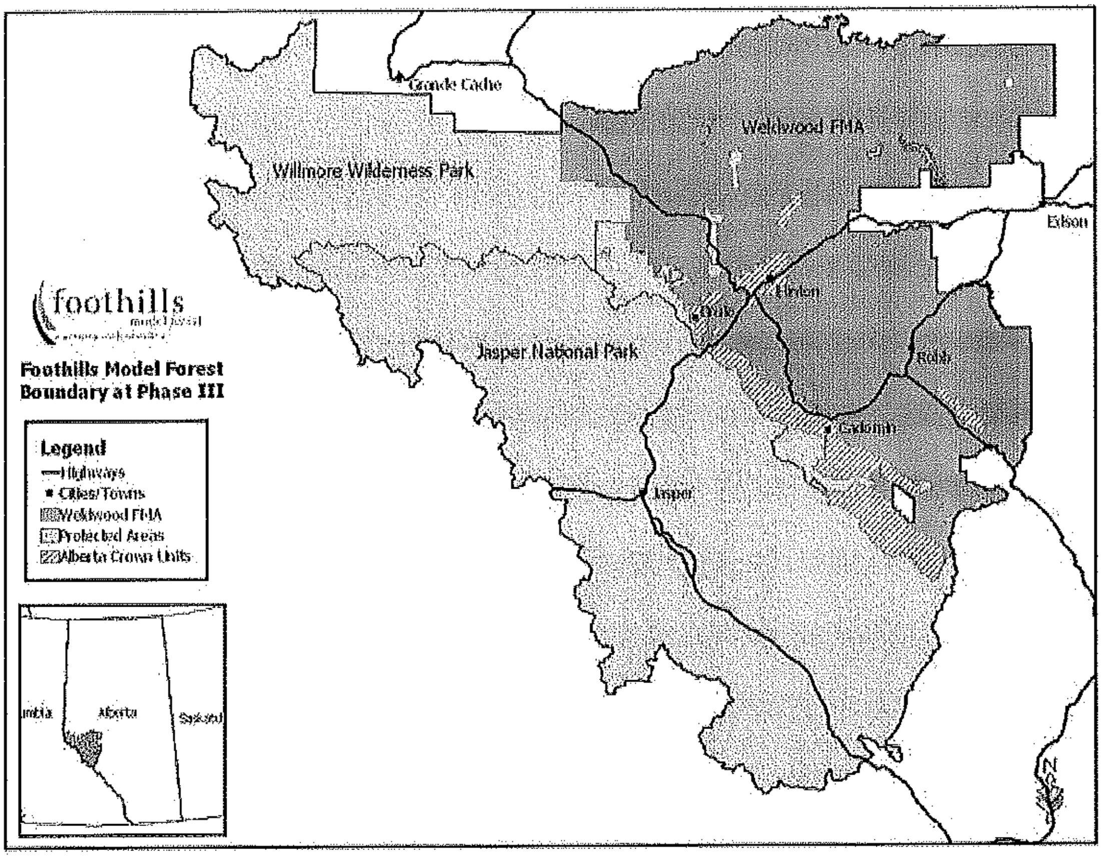

Figure 2. Study Area in the Foothills Model Forest

Typical tree species include pine, spruce, fir, and poplar. The Foothills Model Forest includes Jasper National Park, Willmore Wilderness Park, Weldwood of Canada's Forest Management Agreement area, as well as various oil and gas projects. The majority of the landscape is hilly and forested with meandering rivers and small water bodies. A dense pattern of seismic lines, clearcuts, roads, trails, and well sites has emerged in recent decades due to human activity in the area. Images of the study area are displayed in Figure 3.

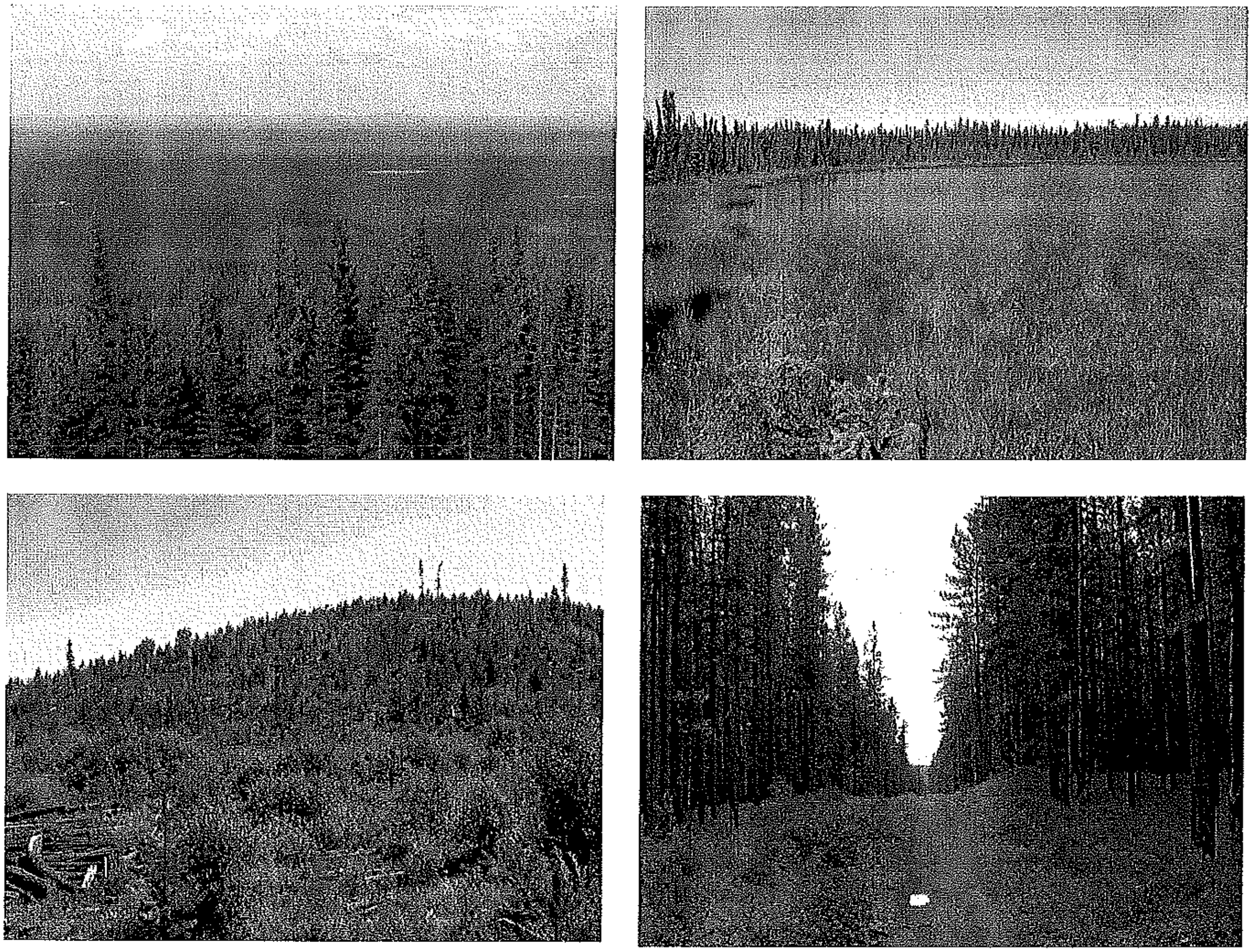

Figure 3. Images of the study area in the summer of 2002, (clockwise, forested landscape, pond, trail, clearcut)

The features of interest for this study are at the landscape level. The natural features are northern forests, primarily coniferous, sparser forests such as bogs, and other natural areas without a lot of trees, rivers and their sand banks, and bare rock and snow at high elevation. The anthropogenic features are older, regrown clearcuts, newer clearcuts, roads, the ditches along roads, well sites, and coal mines.

3.1.1 Field Validation

Field validation was conducted in August 2002 in the Foothills Model Forest. Dozens of clearcuts, well sites, roads and seismic lines were evaluated and observations were recorded to assist in characterizing the features of interest. Observations of clearcuts, well sites and sparsely vegetated areas included the GPS location, the elevation, age, cover type, tree height, and shrub height, percent of different cover types (ex Pine, Spruce, shrub, grass, rock, etc.), the understory plants and additional comments. Observations for linear features included the GPS location, elevation, line width, effective line width, orientation, slope, age of line, soil moisture, ATV use, percent tree, shrub, grass and other cover types, and if the line was visible on the ground. For the area surrounding the seismic line or linear feature, the crown closure, cover type, tree height, shrub height, percent cover type, understory plants and other comments were recorded. The data was entered into a database and shapefiles were created to overlay on the panchromatic data.

Many seismic lines were not easily discernable in the images or could not be segmented in the object oriented classification software, because of the narrow width of the seismic line and the disconnectedness of the line segments, thus the field validation data was used primarily to identify roads, clearcuts and well sites as well as bogs and sparsely vegetated areas. The field validation data and general observations made in the field were used to perform the accuracy assessment. Aerial photograph interpretation, based on field observations, was also used.

3.2. Data Sources: Orthophotographs

3.2.1. Description of Image

Orthophotos were obtained from Weldwood of Canada Limited from their survey in the summer of 2000 as part of a data sharing agreement with Foothills Model Forest. The data was recorded in the panchromatic band (0.4-0.7 μm) of the electromagnetic spectrum at a spatial resolution of 0.75 m. The image was subset to 3000 rows by 3000 columns covering a 5 km^2 area (2.25 km x 2.25 km) on the ground. The geographical extent is UTM coordinates Zone 11: (Upper Left) 489238.25m E, 5885348.00m N and (Lower Right) 491488.25m E, 5882098.00m N. The study area is depicted in Figure 4.

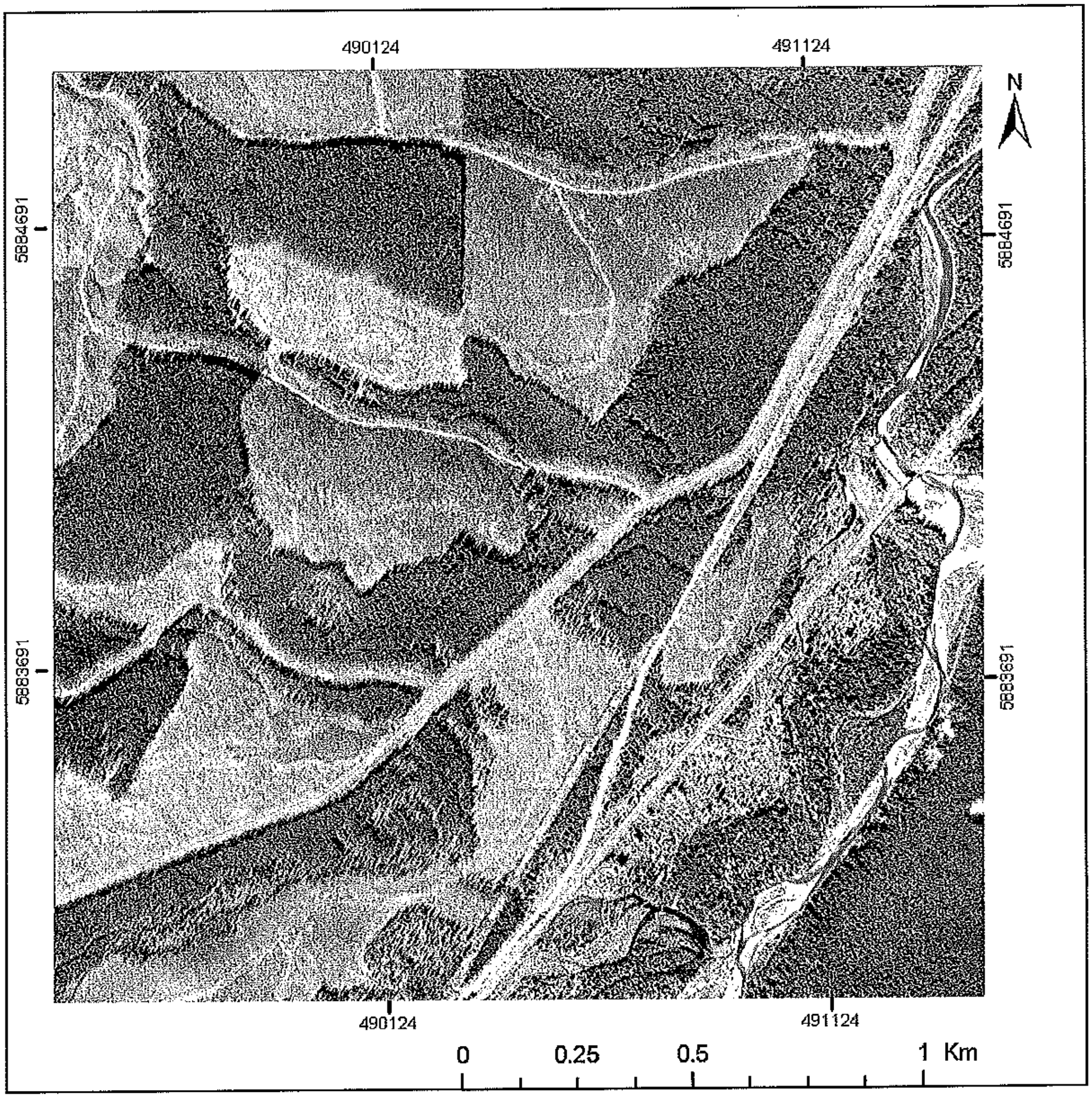

Figure 4. Subset of orthophotograph in study area.

The features of interest in the subset orthophotograph are 10 clearcuts, 6 main roads, areas of sparse vegetation, and a meandering river within the forest matrix. Smaller roads and trails are visible in some clearcuts and the forest.

A flowchart of the methodology to produce the classification of the orthophotograph can be viewed in Figure 5.

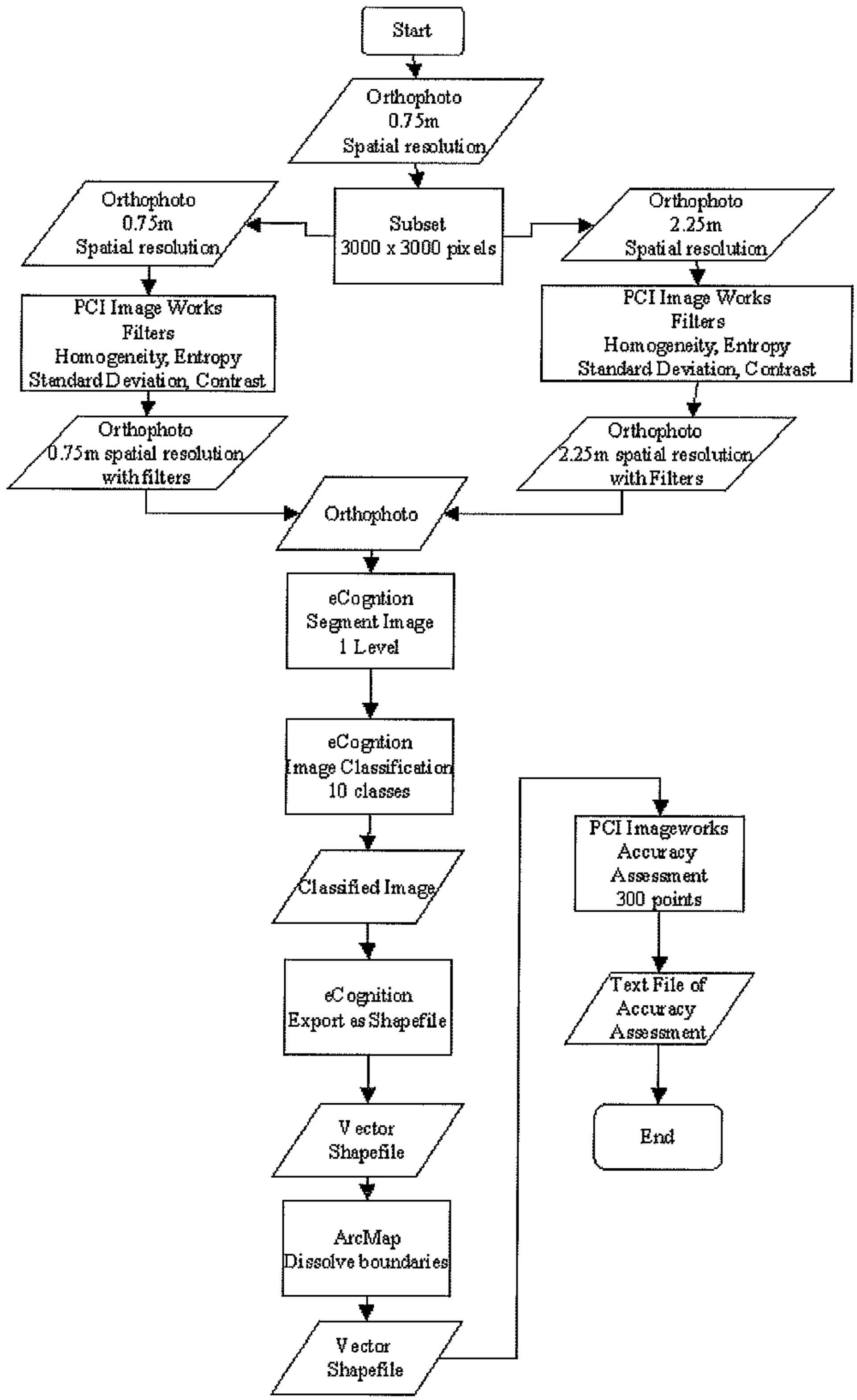

Figure 5. Flowchart of methodology to classify the orthophotograph

3.2.2. Image Pre-processing

In several trials to determine the segmentation parameters, the image was resampled to 2.25 m x 2.25 m, reducing the number of pixels from 90,000 to 10,000. The image was resampled to decrease the heterogeneity of the data due to the high spatial resolution of the orthophotograph.

Various texture measures were applied to the orthophotograph. The filters that enhanced edges to improve the ability to detect edges of features or differences in texture of forest versus clearcuts and other features were selected for both the segmentation and classification process. The window size was varied until an appropriate resolution was obtained to differentiate the characteristics of the desired features in the landscape such as texture of the overall forest or clearcut. Entropy, standard deviation, homogeneity, and contrast measures were calculated using PCI Xpace using moving windows ranging from 3x3 to 9x9 pixels.

A 30m DEM was available for the orthophotograph but was not used, as the change in elevation was so small that the elevation data would not improve the segmentation or classification of the subset image.

3.2.3. Image Segmentation

Image segmentation was achieved using eCognition version 2.0. The segmentation was conducted using the 4 corner pixel neighbourhood segmentation method. In order to perform segmentation the following parameters needed to be input: the scale, colour (or grey level, in the case of panchromatic data) and shape parameters of the heterogeneity criterion and the smoothness and compactness components of the shape parameter. A range of scales was used to create very small image segments (10 pixels) up to very large segments (hundreds of pixels).

Different ratios of colour and shape parameters of the heterogeneity criterion were used to determine how to segment the image, ranging from 100% colour, no shape to equal value to no colour and 100% shape. A higher colour ratio forces the segmentation algorithm to examine heterogeneity of the colour values of the input bands whereas a higher shape ratio forces the examination of heterogeneity in the shape properties of the image segments.

Different smoothness and compactness ratios were used to determine the shape parameter. Smoothness of edges of segments versus compactness of objects were proportioned ranging from greater influence of smoothness to greater influence of compactness.

3.2.4. Classification Rules

Several classes were created to assist the classification process. There were classes for the features of interest: clearcuts, roads, ditches, rivers, sand banks, forests and sparse forests, as well as classes to initially differentiate between groups: forested versus not forested, and old clearcuts versus new clearcuts.

Classification of the orthophotograph was conducted in eCognition version 2.0 using the membership function classification rule set. Experimentation with different combinations of image layers, shape metrics, and logical rules was performed. The primary image layer attributes were the image segments' mean spectral value and standard deviation for the panchromatic band and the enhancement filters. Examination of the shape characteristics of the image segments included the metrics of: area, length, width, border length, length to width ratio, shape index, density and asymmetry. Logical rules using AND and OR were also included in the membership functions. Finally, membership functions detecting borders and proximity with other features were designed to help correctly classify the image segments.

3.2.5. Data Export

Data was exported as a vector shapefile, compatible with ESRI GIS software. Each image segment was exported with any combination of image layer attributes and shape descriptors such as area. The classified image was also exported as a raster PCI .pix file.

3.2.6. Accuracy Assessment

The accuracy of the classified image was assessed in PCI Image Works using the exported PCI .pix file. The classified image layer was added to the original 0.75m spatial resolution, panchromatic image. A random sample of 300 pixels was used to perform the accuracy assessment. Users and Producers accuracy, overall accuracy and Kappa Hat statistics were computed for the image (Jensen, 1996). The random sample of points to perform the accuracy assessment was selected as it follows typical image classification accuracy assessment techniques (Jensen, 1996). A stratified random sample of points based on class was considered for the accuracy assessment, however the intent of the accuracy assessment was to determine how accurately any pixel was classified. All the image classes were incorporated in the accuracy assessment, including both natural and anthropogenic features and were considered in the measure of the overall accuracy.

3.3. Data Sources: IRS Imagery

3.3.1. Description of Data

IRS imagery was obtained through the Foothills Model Forest taken in the summer of 2000. An image was selected that encompassed the spatial extent of the subset orthophotograph. The spatial extent of the image was UTM Zone 11 (Upper Left) 465466.75m E, 5900873.03m N, and (Lower Right) 500736.46m E, 5871468.30m N. The spatial extent of the image can be seen in Figure 6 while the quality of the spatial resolution can be assessed in Figure 7.

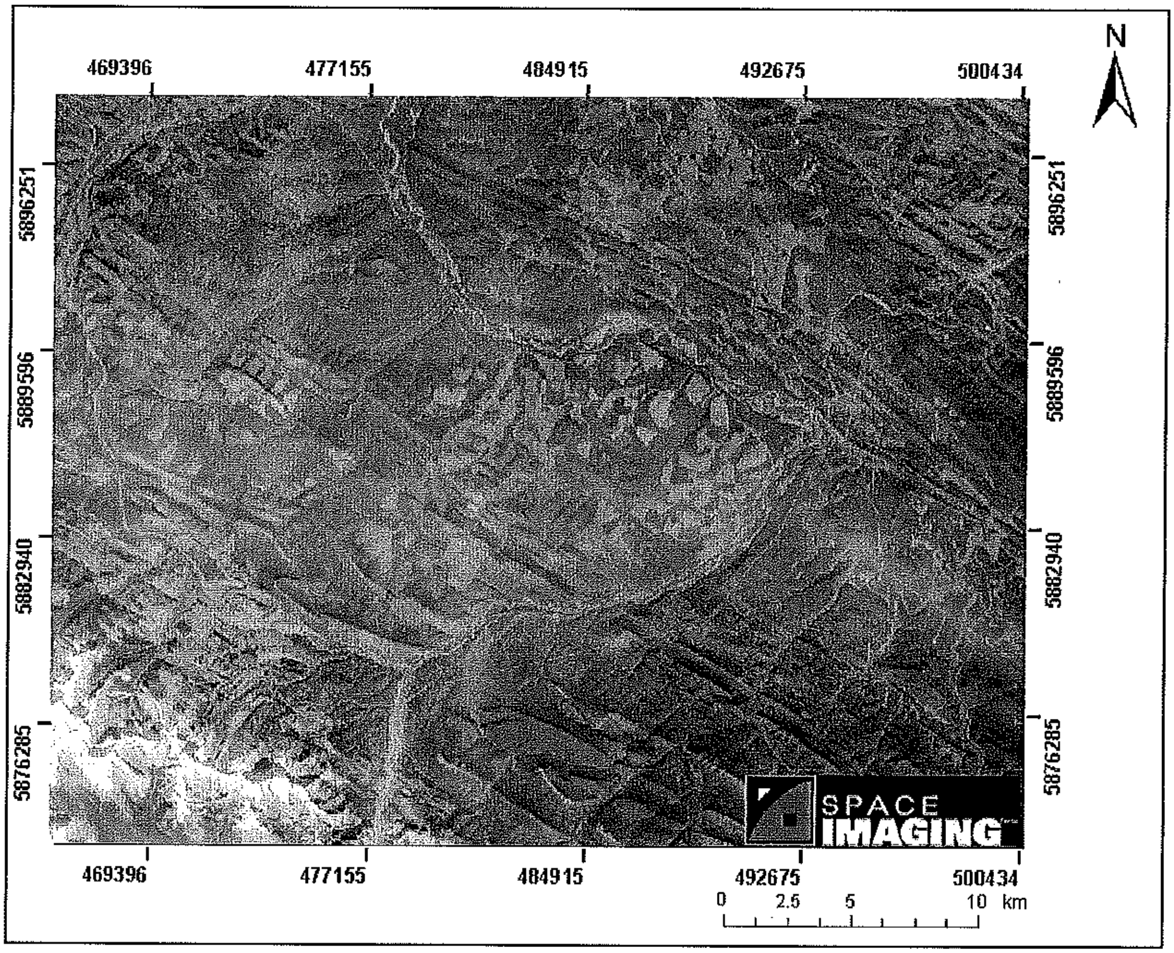

Figure 6. Panchromatic Band of IRS image

Figure 7. Detailed view of IRS imagery

The IRS imagery is a panchromatic image with a spatial resolution of 5.6m. The Foothills Model Forest researchers resampled the images to 5.0m. The 5.0m resampled image used was 7054 x 6088 pixels in size covering an area of 1072 km^2 (35km x 30km).

Features discernable in the full IRS scene include more than 70 clearcuts, several major roads including the two-lane Highway 40, the logging roads Pembina Road and Tri Creeks Road as well as several smaller roads, power lines, seismic lines and trails. There is a coal mine situated at the SW corner of the image. This image has an elevation range of 1326 m –1497 m. There are several large rivers and several smaller rivers in the image. Shadowed forest, sparsely vegetated areas, and forests cover the remainder of the full IRS scene.

The methodology used to classify the IRS imagery is outlined in a flowchart in Figure 8.

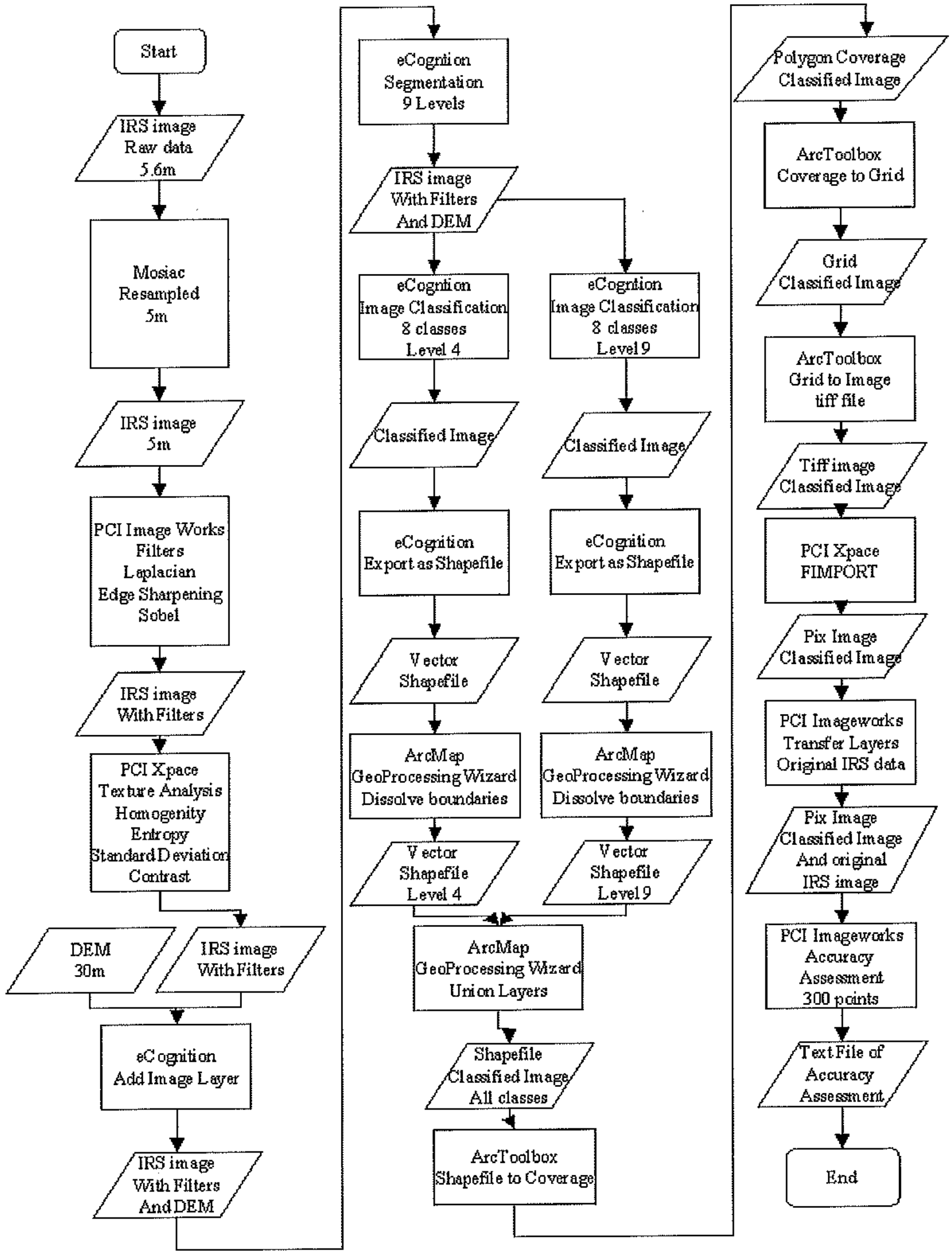

Figure 8. Flowchart of methodology to classify the IRS image

3.3.2. Pre-processing of Image

The IRS image was pre-processed in PCI Image Works. The texture measures Standard Deviation, Contrast, Entropy and Homogeneity were calculated for the image using a moving window with the dimensions 3 x 3 to 9 x 9 pixels.

The edge detection filters: Sobel, Edge Sharpening, Type 1 Laplacian and Type 2 Laplacian filters were created for the IRS image with dimensions also ranging from 3 x 3 to 9 x 9 pixels. Type 1 Laplacian filter has the following weight matrix (0,1,0 1, -4,1 0,1,0), while the Type 2 Laplacian filter has (-1,-1,-1 –1,8,-1 –1,-1,-1). Advantages of each operation are outlined in Table 1.

Table 1. Operations computed for the IRS image

Operation	Enhancement for panchromatic data
Standard Deviation (9x9)	Separating roads from other areas
Contrast (9x9)	Identifying river valleys and other relief
Entropy (9x9)	Separating clearcuts and bare rock from forest
Homogeneity (3x3)	Identifying bare rock and clearcuts to some extent
Sobel	Identifying linear features and edges
Edge Sharpening (5x5)	Increases contrast between features of interest
Type 1 Laplacian (3x3, 5x5)	Identifying linear features and edges
Type 2 Laplacian (3x3, 5x5)	Identifying linear features and edges

A 30 m digital elevation model was added to the IRS image to assist in classifying the panchromatic image.

The large IRS scene was subset to cover an area approximately the same as the subset orthophotograph. The spatial extent is slightly different because the pixel size was 5m and did not coincide exactly with the orthophotograph subset. The subset IRS scene was subjected to the same treatments as the subset orthophotograph except the filters computed for the entire IRS scene were subset for the subset IRS image. The DEM was not included in the analysis of the subset IRS scene.

3.3.3. Image Segmentation

Image segmentation was performed in eCognition version 2.0. The image was segmented several different ways with the goal of achieving image segments that represented real world objects such as clearcuts and roads.

The 4 corner pixel neighbourhood and the 8 corner pixel neighbourhood option of the segmentation were used to segment the image. The image was segmented at a variety of levels from image objects representing a few pixels to large segments representing greater areas. The scale parameter, colour, shape, smoothness and compactness criteria were adjusted in an attempt to optimize the segmentation of the image.

The subset IRS image was segmented using the same segmentation parameters as the subset orthophotograph.

3.3.4. Classification Rules

Several image classes were created to assist in the classification of the IRS image, either as intermediate classes or final classes and are summarized in Table 2.

Table 2. Classes created to classify IRS imagery

Natural Feature	Anthropogenic Feature
Forest	Roads
Sparse Forest	Clearcuts
Forest Shadows	Wellsites
Rivers	Mines
Sand Banks	
Bare Rock	

A "Road Ditches" class was not created, as ditches were not distinguishable from the roads as they were in the orthophotographs.

Two types of classification rules were applied to the IRS imagery: membership functions and nearest neighbour classifier, both available in eCognition version 2.0. Membership functions use fuzzy classifiers based on a single image attribute. Each image object is evaluated for fuzzy membership to each membership function. The image object is assigned to the image class with the highest membership value assigned to it. The combination of several membership functions produces the classification rules for each image class.

For the IRS imagery, the shapes of the membership functions were created manually or were generated by eCognition based on samples of image objects belonging to an image class. Membership functions were created based on image layer spectral properties, shape attributes and logical expressions. Image layer spectral values included the image objects' mean, standard deviation, and ratio of image layers of panchromatic data and filters. The logical expressions included statements such as "similar to bare rock" and "not classified as roads".

The nearest neighbour classifiers are based on a multi-dimensional separation in feature space of spectral and shape attributes (Definiens Imaging, 2001a). The separation of image classes is computed in eCognition based on the user-selected attributes. All the image layer attributes and shape characteristics were chosen for nearest neighbour analysis.

The subset IRS scene was classified using similar classification rules to those that were applied to the subset orthophotograph.

3.3.5. Data Export

The classifications created based on the 4 corner pixel neighbourhood segmentation parameter were exported as vector shapefiles and as raster PCI .pix files. The classifications based on the 8 corner pixel neighbourhood could only be exported as raster files because the software algorithm can only build polygons based on the 4 corner pixel neighbourhood.

When the best classification required extracting image objects from two segmentation levels, the image objects from both levels were exported from eCognition as vector shapefiles with smoothed geometry. Using the GeoProcessing wizard in ArcGIS, the boundaries were dissolved between objects of the same class and the two layers were reduced to one layer using the UNION command. The shapefile was converted to a polygon coverage, an Arc grid and a .tiff image using the ArcToolbox conversion tools. In PCI XPace, the .tiff image was converted to a PCI .pix file and was transferred as a new layer added to the original IRS imagery file.

3.3.6. Accuracy Assessment

Accuracy assessment was conducted in PCI Image Works. A random sample of 300 points was used to assess the classification of the IRS imagery. Producer's, user's and overall accuracy as well as k^ were computed for the classified images.

CHAPTER FOUR: RESULTS

A trial and error approach was used to determine the best technique to segment and classify the images. Due to the large number of trials (exceeding 50), only the best results will be presented and discussed.

4.1. Orthophotographs

4.1.1. Segmentation

One level of segmentation was used to achieve the best delineation of features of interest. The panchromatic band, the entropy, standard deviation and contrast measurements layers were all assigned equal weight for the segmentation process. The scale parameter was set to 100, the colour and shape parameters of the heterogeneity criterion were set to 0.5 each, and the smoothness and compactness components of the shape parameter were set to 0.9 and 0.1 respectively. The results of the segmentation can be viewed in Figure 10.

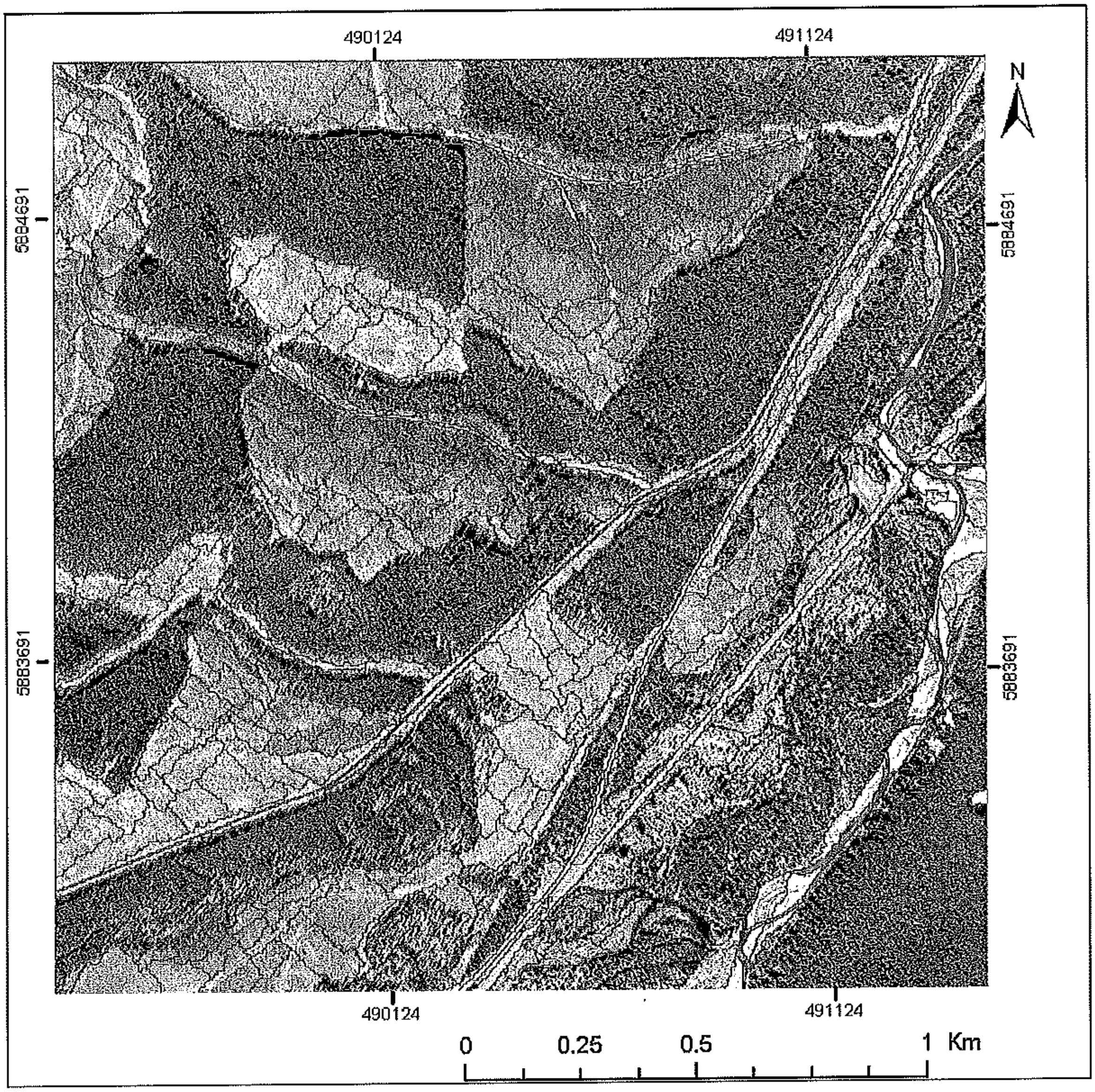

Figure 10. Segmentation of orthophotograph using eCognition.

The segmentation resulted in a very accurate delineation of features of interest. Sections of the river as well as adjacent sand banks were delineated accurately. Roads and smaller roads within clearcuts were also detected well. At the scale of segmentation of 100, clearcuts and large sections of the forest were not grouped as one segment per clearcut or forest. Instead, the clearcut or section of forest was divided into several smaller segments.

4.1.2. Classification

The best classification of the orthophotograph is presented in Figure 11.

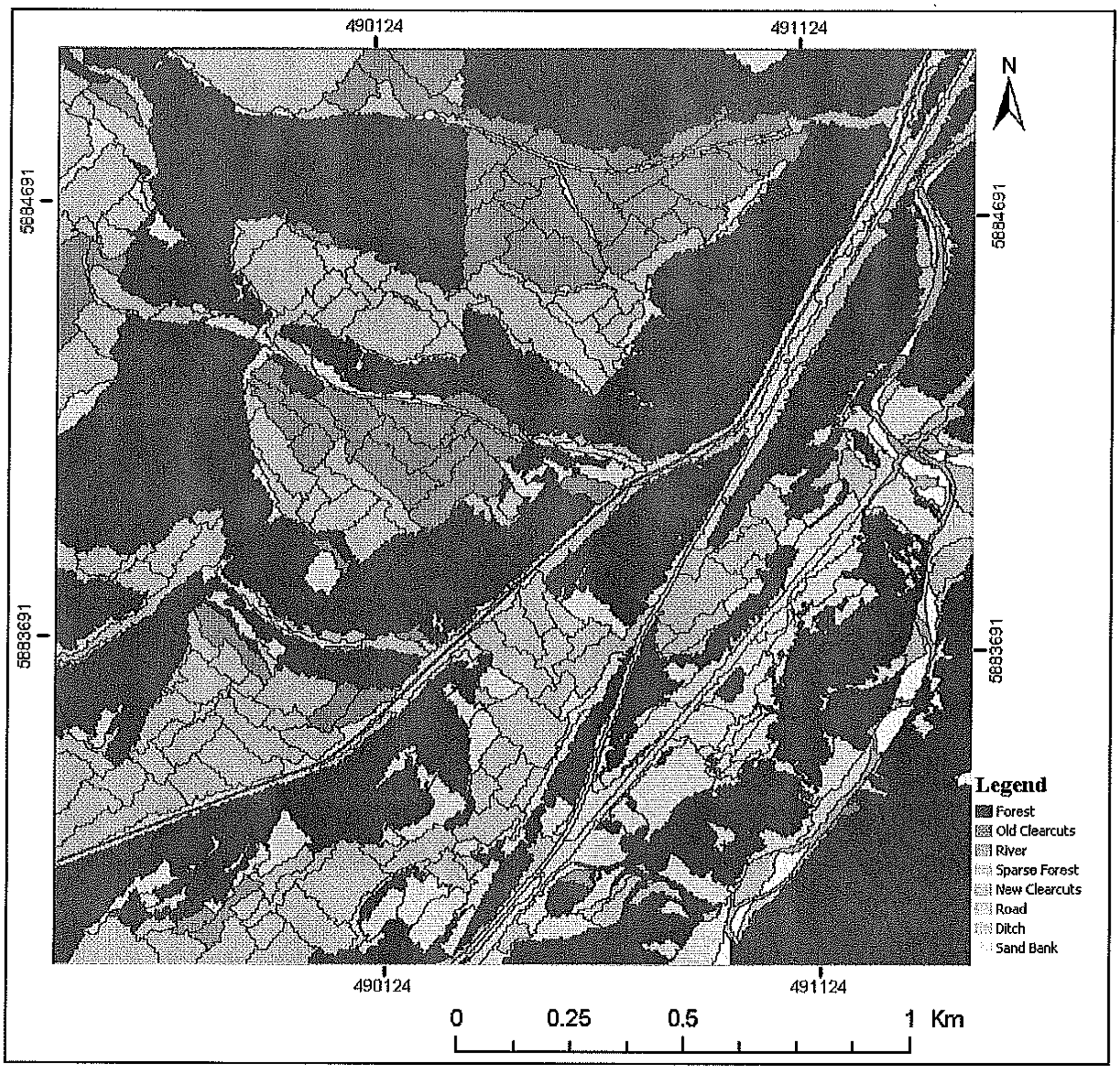

Figure 11. Classification of orthophotograph

The classification rules were all hard classifiers, meaning that the rule assigned 100% membership or 0% membership to a class, and no values in between. The classification rules are presented in Table 3.

Table 3. Classification rules for the orthophotograph

Forest		Non-Forest					
Mean contrast filter 17-43		Mean contrast filter 0-22					
Sparse Forest	**Forested**	**Road**	**Sandbank**	**Ditches**	**River**	**Old Clearcut**	**New Clearcut**
Mean spectral value 140-231	Mean spectral value 0-160	Length/Width 2-3182	Mean entropy 0-3.8	Length/Width 10-3000	Density 0-1.55	Brightness 35-255	Length/Width 0-22
	Mean contrast filter 0-43	Mean spectral value 220-255	Mean spectral value 220-255	Mean spectral value 210-252	Mean spectral value 44-181	Density 1.2-5	Mean spectral value 180-234
		St. dev. of mean spectral value 0.5-4.2	St. dev. of mean spectral value 5-45	St. dev. of mean spectral value 9-38	St. dev. of mean spectral value 0-53	Mean spectral value 58-183	St. dev. of mean spectral value 16-47
					Border to Ditches 5-2250	St. dev. of mean spectral value 20-66	
					Border to Sandbanks 5-2250		

The initial classification rules were to separate forested areas (forest and sparse forest) from non-forested areas (old clearcuts, new clearcuts, roads, ditches, sandbanks and rivers). Forested areas were best separated from non-forested areas using the spectral feature: contrast measurement layer. Forested areas were divided into sparsely vegetated areas and forested areas by the mean spectral value of the image segments and the mean of the contrast filter layer. A flow chart of the procedure is displayed in Figure 12.

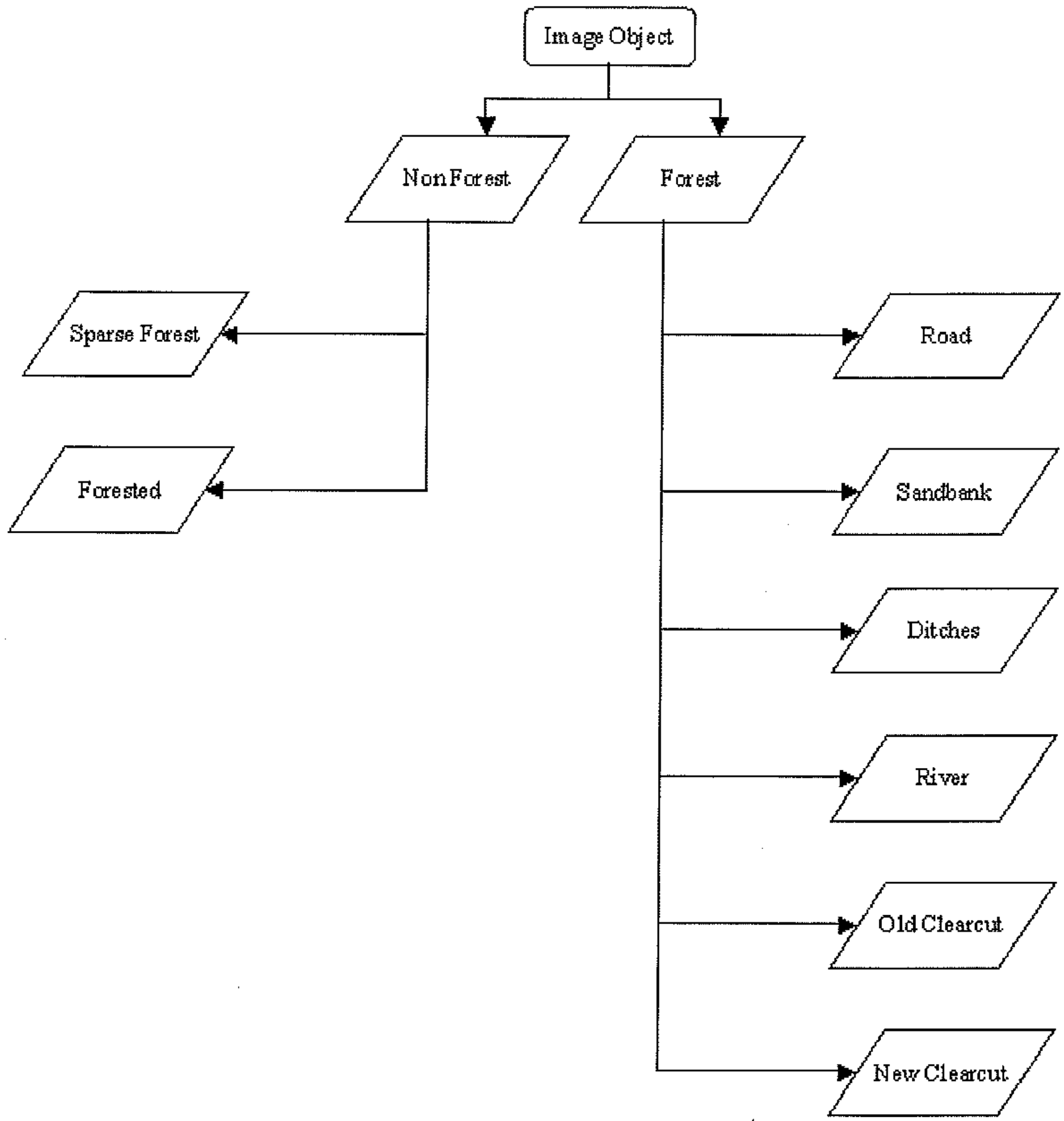

Figure 12. Flowchart of classification rules for the orthophotograph

Non-forested areas were separated using the spectral features of mean spectral value of the panchromatic band, standard deviation of the panchromatic band, mean contrast and entropy filters as well as the mean brightness value of the combined image and filter layers. The shape features used to separate non-forest classes were: length to width ratio and density of image objects. The class-related feature of border to neighbours was used to classify rivers based on the distance to the ditches and sandbank classes.

The boundaries between image objects with the same classification were dissolved in ArcMap to produce meaningful objects (Figure 13).

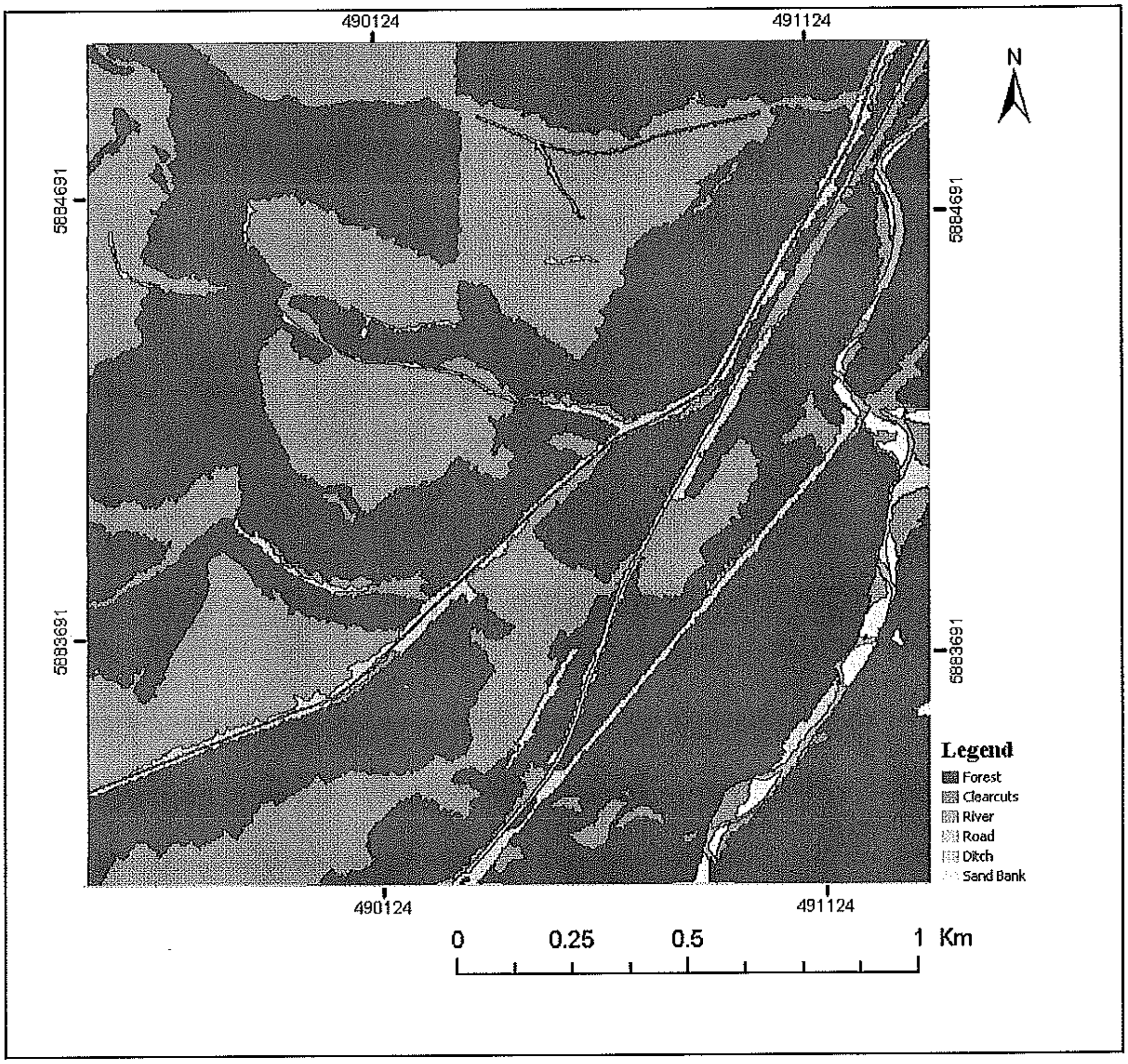

Figure 13. Final classification of the orthophotograph

4.1.3. Accuracy Assessment

The results of the accuracy assessment are summarized in Tables 4 and 5.

Table 4. Number of samples per class used for the accuracy assessment.

	Reference						
Classified	**Road**	**Clearcuts**	**River**	**Sand**	**Ditch**	**Forest**	**Totals**
Road	1	0	0	0	0	0	1
Clearcut	0	70	0	5	6	4	85
River	0	0	2	0	0	0	2
Sand	0	0	0	1	0	0	1
Ditch	0	1	0	1	11	0	13
Forest	0	10	0	2	7	179	198
Unknown	0	0	0	0	0	0	0
Total	1	81	2	9	24	183	300

Table 5. Producer's and User's accuracy for the orthophotograph

Class Name	**Producer's Accuracy**	**User's Accuracy**
Road	100%	100%
Clearcut	86%	83%
River	100%	100%
Sand	11%	100%
Ditch	46%	85%
Forest	98%	90%

The overall accuracy was 88% with a Kappa hat statistic of 0.768. The highest classification accuracies were for roads, clearcuts, rivers, and forest, each resulting in user's and producer's accuracies over 80%. There was confusion in classifying sand, with a low producer's accuracy of 11%. Sand banks were misclassified as clearcuts, ditches and forest. There was also confusion in classifying ditches along roads with a producer's accuracy of 46%. Ditches were misclassified as clearcuts and forest.

4.2. IRS Imagery

4.2.1. Segmentation

The segmentation of the large IRS image was a challenge. Several levels of segmentation were needed to achieve objects that would be feasible to classify. The best segmentation achieved in this experiment is summarized in Table 6.

Table 6. Segmentation parameters for IRS imagery

Level	Scale	Colour	Shape	Smooth-ness	Compact-ness	Image Layers
1	5	0.8	0.2	0.9	0.1	All textures and filters
2	10	0.8	0.2	0.9	0.1	All textures and filters
3	20	0.5	0.5	0.9	0.1	Only IRS panchromatic
4	50	0.8	0.2	0.9	0.1	Only IRS panchromatic
5	70	0.8	0.2	0.9	0.1	Only IRS panchromatic
6	10	0.8	0.2	0.9	0.1	Only IRS panchromatic
7	120	0.8	0.2	0.9	0.1	Only IRS panchromatic
8	150	0.8	0.2	0.9	0.1	Only IRS panchromatic
9	180	0.8	0.2	0.9	0.1	Only IRS panchromatic

Initially, the IRS image contained 43 million pixels. The first level of segmentation created 669,642 small image objects. Level 4 of the segmentation merged the lower level image objects to 16,967 image objects. Level 9 merged the smaller image objects into 2375 larger objects. The Level 4 and Level 9 segmentations are displayed in Figures 14 and 15.

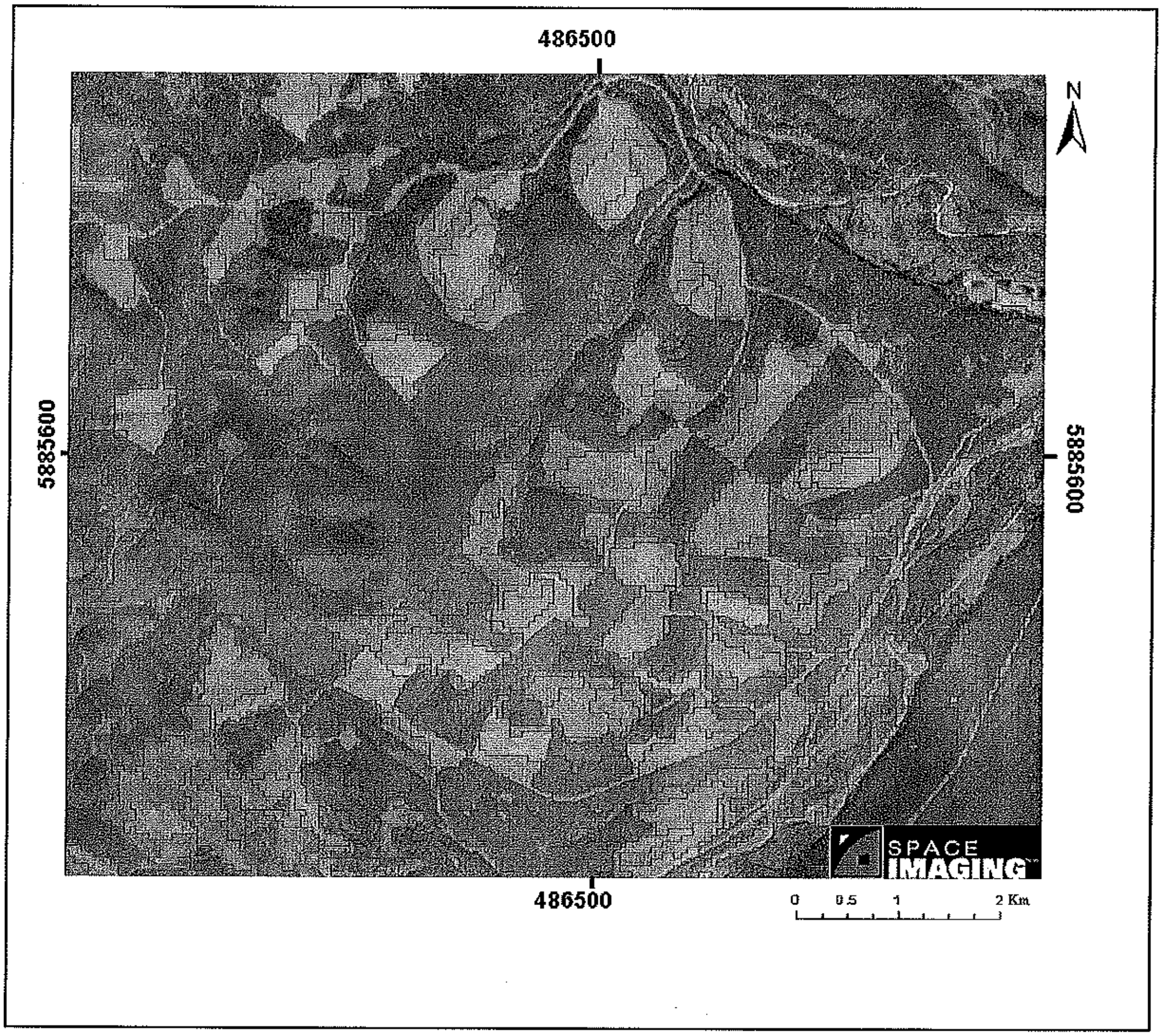

Figure 14. Subset of Level 4 segmentation of IRS imagery using 4 corner pixel neighbourhood segmentation

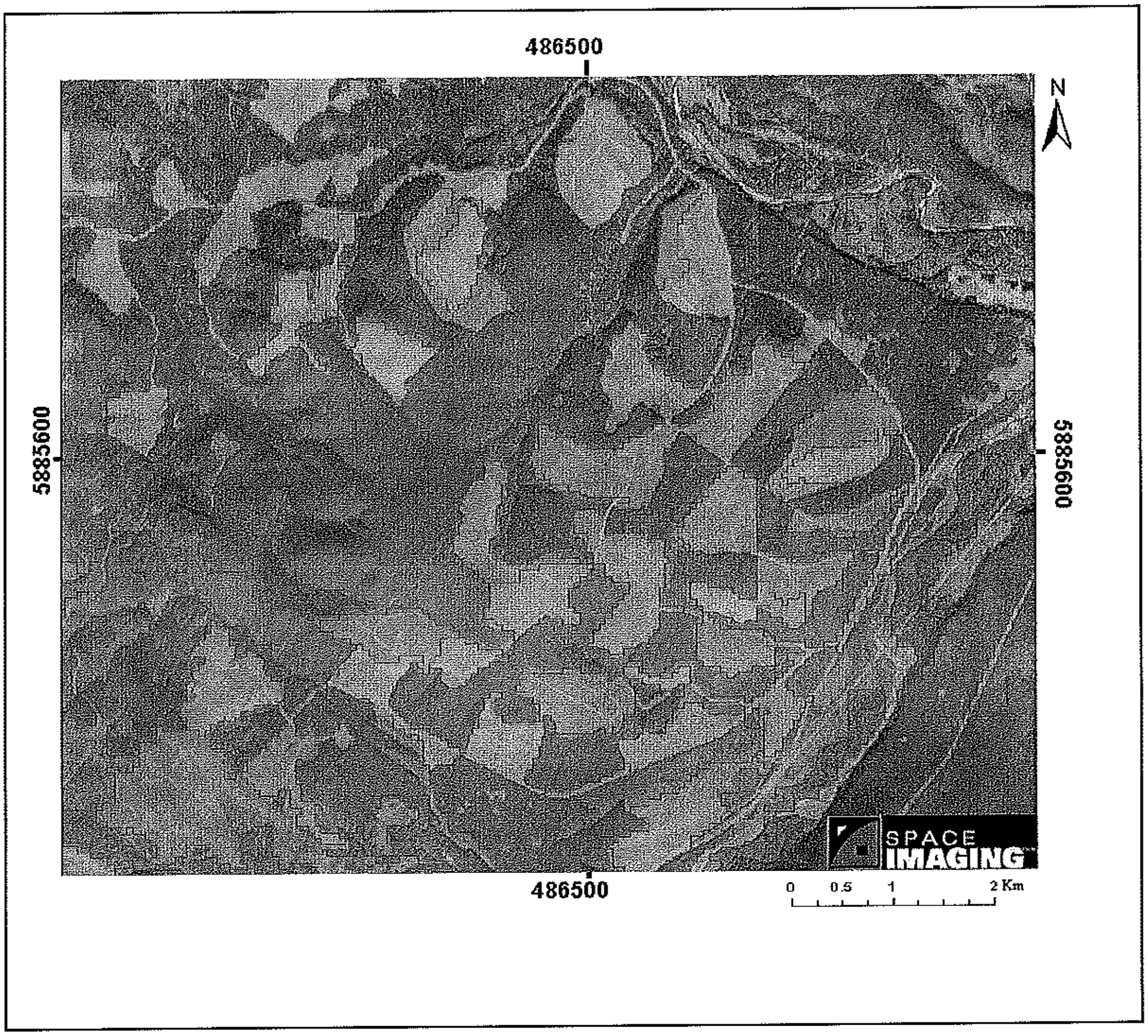

Figure 15. Subset of segmentation of IRS image at Level 9, using 4 corner pixel neighbourhood segmentation

While none of the levels of segmentation captures perfectly the objects of interest, Level 4 is useful for classifying narrow roads and small wellsites. Level 9 is useful for classifying clearcuts, forested areas, and mining areas as the spectral and shape values of the representative image segments can be used to compute membership functions.

The subset IRS image was segmented using similar segmentation parameters as the subset orthophotograph. The results of the segmentation can be viewed in Figure 16.

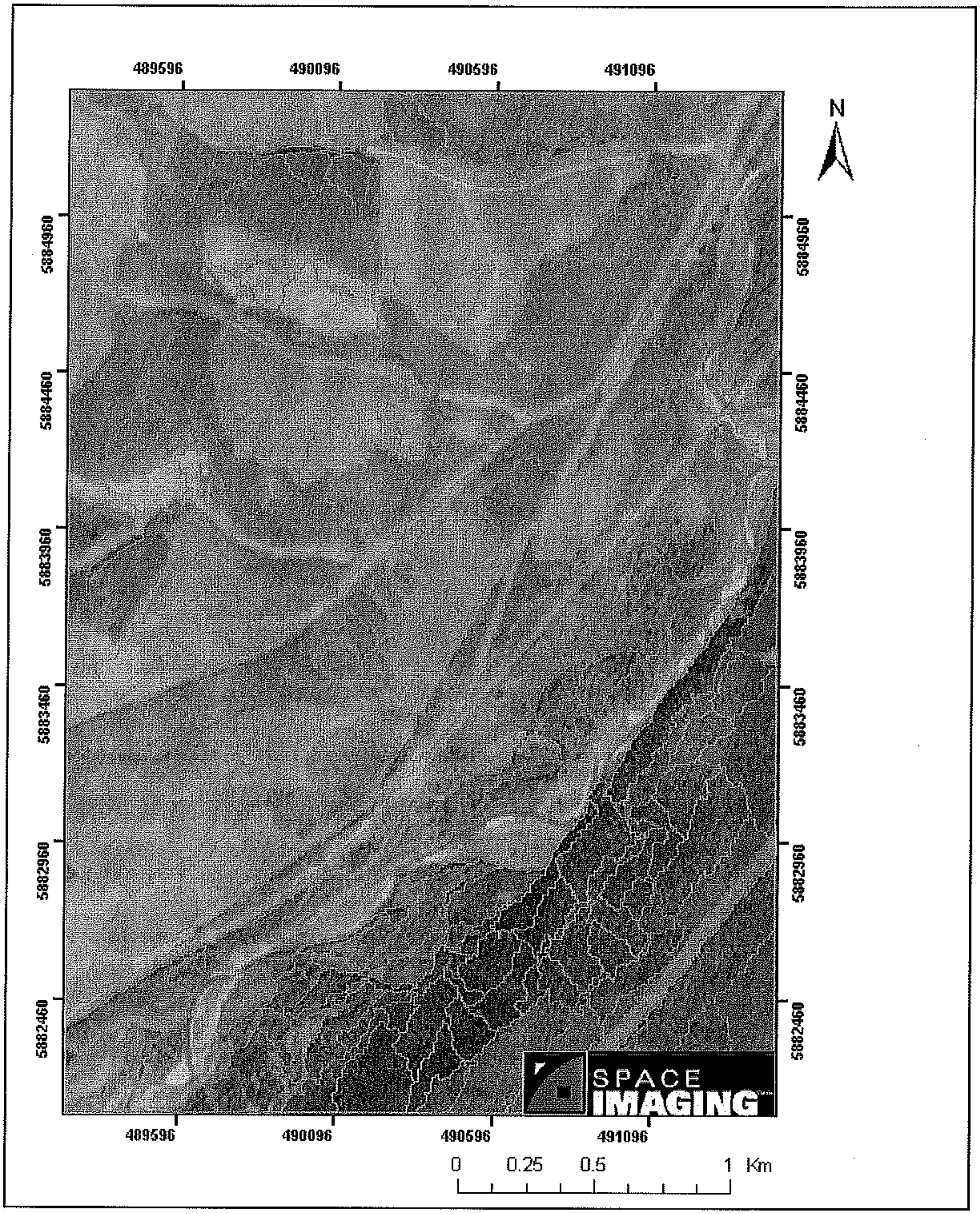

Figure 16. Segmentation of subset IRS image

4.2.2. Classification

Classification of the IRS imagery was done using membership functions. The classification rules for Level 4 and Level 9 classification are summarized in Tables 7 and 8.

Table 7. Level 4 classification rules for the IRS imagery

Roads	**Wellsites**	**Not Roads**
Level 4	Level 4	Level 4
Length/Width 3-200	Area 0-40,000	Membership to Roads 0-0.8
Mean spectral value 120-190	Mean DEM 0-2200	Not bare rock
Not bare rock	Mean spectral value 138-197	
Ratio DEM 0.55 – 0.71	Shape Index 0-2.11	
Ratio spectral values 0.04-0.1	Width 8-50	
Ratio Laplacian filter 0.01-0.12		
Width 0-103		

Table 8. Classification rules for Level 9 of the IRS imagery

Bare Rock	**Forest**	**Sparse Veg near Bare Rock**	**Sparse Forest**	**Forest Shadows**	**Clearcut**	**Mine**	**Roads**
Level 9	Level 9	Level 9	Level 9	Level 9	Level 9	Level 9	Level 9
Mean DEM 2200-3500	Mean spectral value 27-170	Distance to Bare Rock 0-830	Mean spectral value 44-255	Mean spectral value 0-99	Distance to Bare Rock 47-300 +	Mean spectral value 100-220	Length/Width 3-200
	Shape index 0.7-4.5	Mean spectral value 42-255	Mean Laplacian filter 0-255		Length/Width 5-50	X center 467100-477000	Mean spectral value 120-190
	St. dev. of mean spectral value 5-30	Not Bare Rock	Not Mine		Mean spectral value 120-220	Y center 5874000-5884200	Not bare rock
					Not Bare Rock		Ratio DEM 0.55 – 0.71
					Not Mine		Ratio spectral values 0.04-0.1
					St. dev. Edge Sharpening filter 0-100		Ratio Laplacian filter 0.01-0.12
							Width 0-103

To achieve the best classification of roads given the limitations of the segmentation at Level 4, the shape features of length to width ratio and width as well as the spectral features of mean IRS panchromatic band, the ratio of the mean Laplacian filter and mean DEM values to the image and the logical statement that roads could not be on high elevation Bare Rock were used.

Well sites were classified using the shape features: width, area and shape index as well as the mean spectral value of the IRS panchromatic band and the mean elevation of the image objects.

Level 9 classification was achieved by first separating the bare rock at higher elevations using the mean elevation value. The areas affected by coal mining activities were delineated based on their proximity to the coalmine and the mean spectral value. Roads in Level 9 were classified using the same membership rules as in Level 4 (length/width ratio, width, mean panchromatic band, ratio Laplacian filter, ratio of DEM).

Clearcuts were classified by not being Bare Rock or Mine, being 47m distance from Bare Rock (as clearcuts do not occur at very high elevations), as well as by their spectral characteristics of mean IRS panchromatic band and standard deviation of the edge sharpening filter.

Forest image objects were classified as Forest, Sparse Vegetation, Sparse Vegetation near Bare Rocks (high elevation) or Forest Shadows. The Forest objects were classified based on the mean and standard deviation of the IRS panchromatic band and the shape index. The Sparse Vegetation class was based on the mean IRS band and the mean Laplacian filter values of the image objects as well as not being classified as a Mine. Sparse Vegetation near Bare Rock was classified based on proximity to Bare Rock while not being Bare Rock, and on the mean spectral value of the IRS image. Finally, Forest Shadows were classified based on the mean IRS panchromatic band.

Level 4 and Level 9 classified images were exported and combined in ArcMap. The separate and combined images can be viewed in Figures 17, 18 and 19.

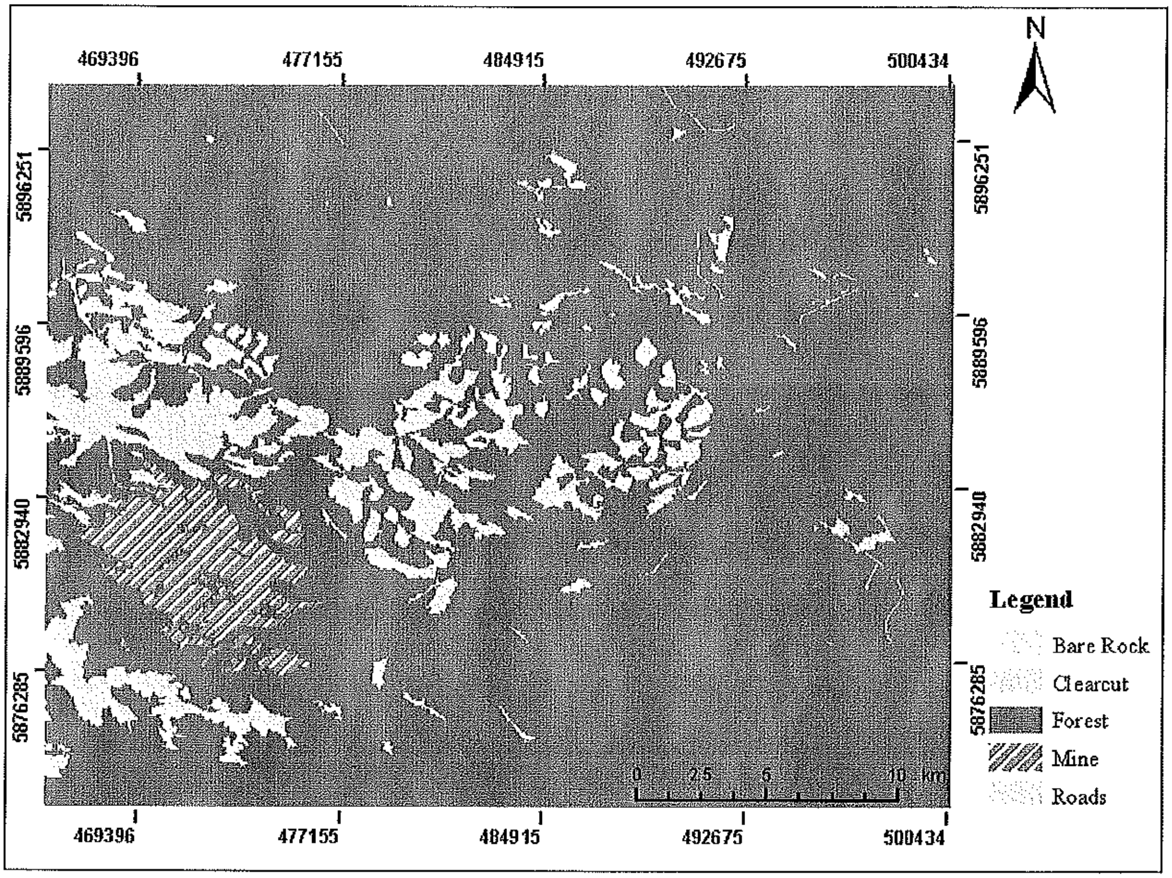

Figure 17. Classification of IRS imagery using Level 9 segmentation.

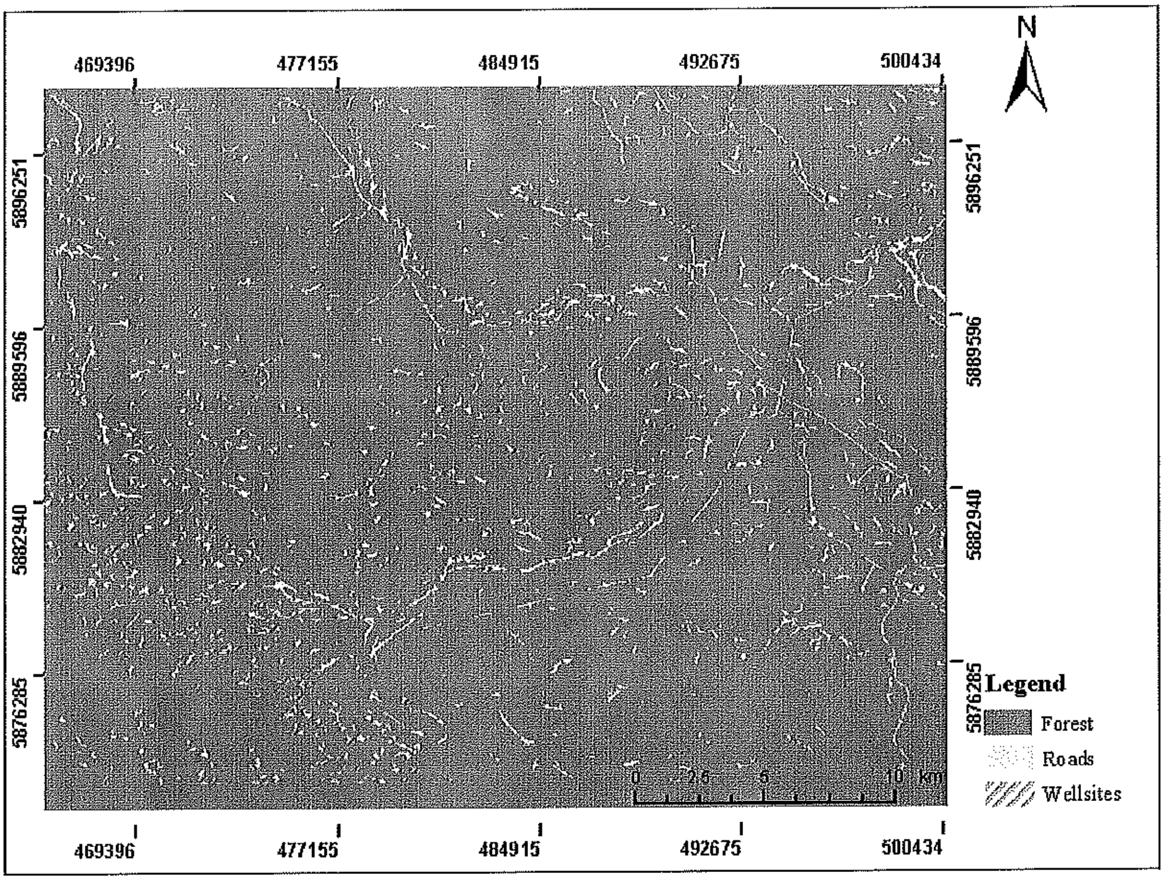

Figure 18. Classification of IRS imagery using Level 4 segmentation

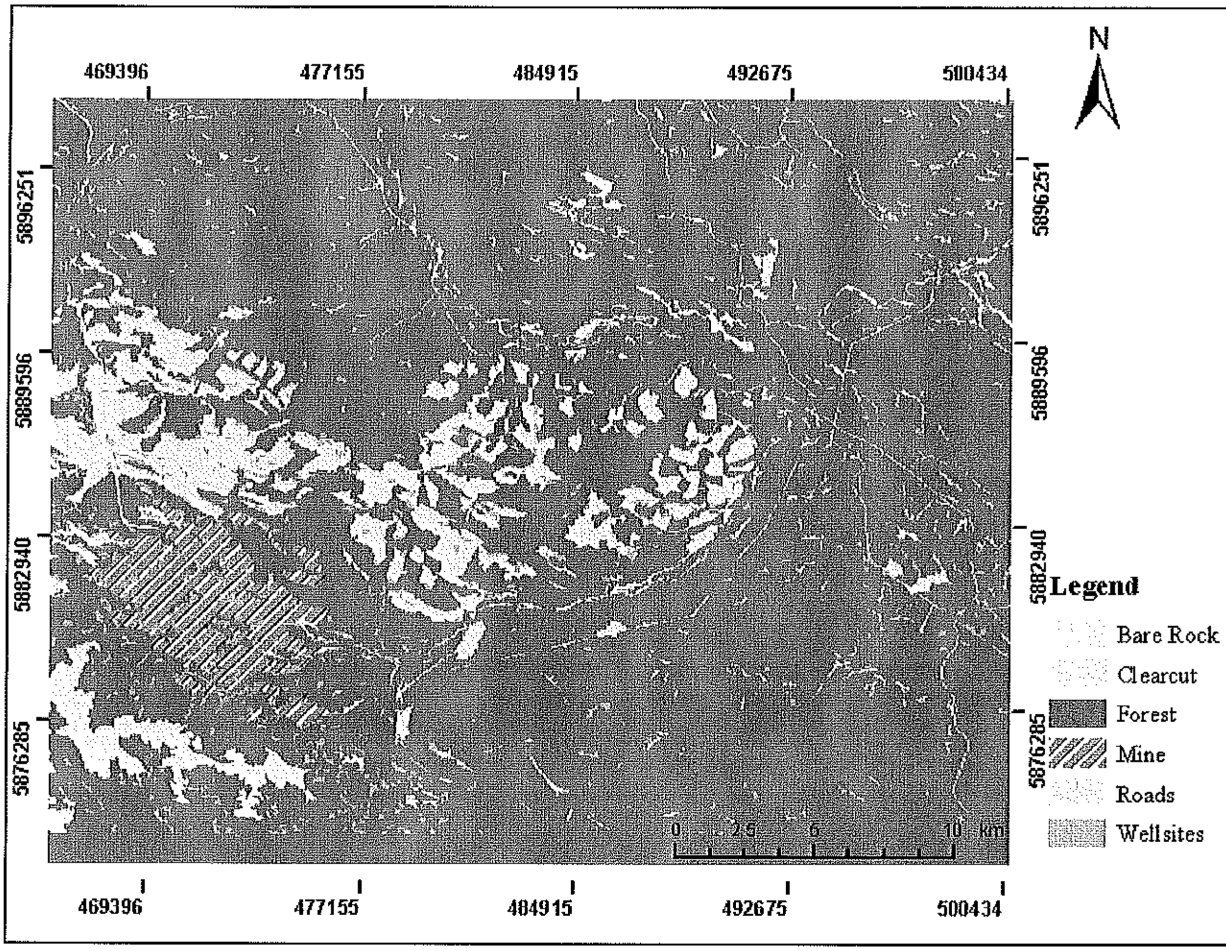

Figure 19. Final classification of IRS imagery, combination of Level 4 and 9 segmentation

The results of the classification of the subset IRS image can be viewed in Figure 20.

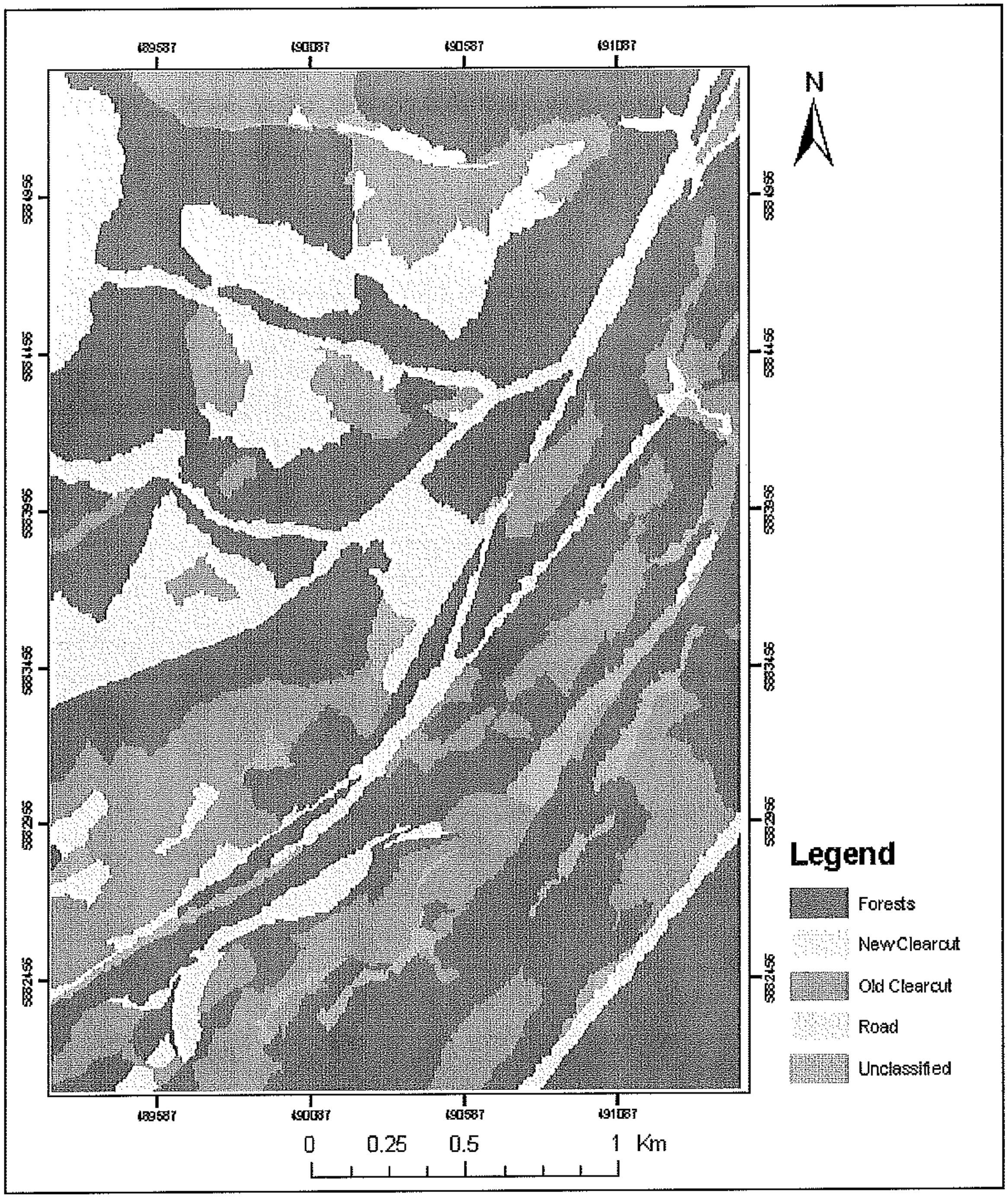

Figure 20. Classification of subset IRS image.

4.2.3. Accuracy Assessment

Accuracy assessment was conducted using PCI Imageworks. The results are summarized in Table 9 and 10.

Table 9. Number of samples per class used for the accuracy assessment.

	Reference						
	Roads	**Clearcut**	**Forest**	**Wellsites**	**Bare Rock**	**Mine**	**Total**
Classified							
Roads	2	0	3	0	0	0	5
Clearcut	0	14	11	0	0	1	26
Forest	5	15	216	0	5	5	246
Wellsites	0	1	3	1	0	2	7
Bare Rock	0	0	0	0	6	0	6
Mine	1	0	1	0	0	8	10
Total	8	30	234	1	11	16	300

Table 10. Producer's and User's accuracy for the IRS image

Class Name	**Producer's Accuracy**	**User's Accuracy**
Road	25%	40%
Clearcut	47%	54%
Forest	92%	88%
Wellsites	100%	14%
Bare Rock	55%	100%
Mine	50%	80%

The overall accuracy was 82% with a Kappa hat statistic of 0.49. The high overall accuracy can be attributed to the large proportion of samples taken from the Forest class (234/300). Forest classes attained a user's and producer's accuracy over 85%. The other classes had user's and producer's accuracies between 25 and 80%. Roads and Clearcuts were misclassified as Forest while Mines were classified as Roads and Clearcuts.

Wellsites were classified as being much more abundant than in reality. Image segments were classified as Wellsites when they were Clearcuts, Forests or Mines.

Bare Rock objects were all correctly classified as Bare Rock. Due to the range of elevation values used to classify Bare Rocks, several Forest image segments were incorrectly included in the Bare Rock class.

The accuracy assessment of the subset IRS scene is summarized in Tables 11 and 12.

Table 11. Number of samples per class used for the accuracy assessment of the subset IRS scene.

Reference	**Forest**	**Roads**	**Old Clearcuts**	**New Clearcuts**	**Unclassified**	**Total**
Classified						
Forest	129	3	4	1	0	137
Roads	3	10	3	4	0	20
Old Clearcuts	26	3	13	12	0	54
New Clearcuts	2	0	2	26	0	30
Unclassified	6	3	0	0	0	9
Total	166	19	22	43	0	250

Table 12. Producer's and User's accuracy for the IRS image

Class Name	**Producer's Accuracy**	**User's Accuracy**
Forest	78%	94%
Roads	53%	50%
Old Clearcuts	59%	24%
New Clearcuts	60%	87%
Unclassified	0%	0%

The overall accuracy of the subset IRS image was 71% with a kappa hat statistic of 0.512. In general every class was poorly classified except the clearcuts and the forest. None of the river features were classified and several road and clearcut segments were left unclassified using the same classification rules as for the subset orthophotograph.

CHAPTER FIVE: DISCUSSION

5.1. Orthophotographs

Object-oriented image processing was successful in correctly delineating and classifying the features of interest in the orthophotograph of the Foothills Model Forest. In particular, the segmentation correctly grouped pixels together to form realistic objects with the correct shape and spectral characteristics. For example, the roads were nearly perfectly delineated, each created as a long, continuous road segment. The low standard deviation, high spectral values as well as the high length-to-width ratio made it easy to distinguish the road features from the other image objects.

Acceptable segmentation of the orthophotograph was achieved using only one level. Equal weight was assigned to the colour and shape parameters of the homogeneity criterion for the image segmentation. The balance between segmenting based on similar grey level characteristics of adjacent pixels and the shape of the resulting image segments produced very realistic image objects. The delineation of the roads, rivers and sand banks image objects can be considered to be very similar to how a person would have delineated for the same study area. However, the forest polygons and clearcuts were segmented into several smaller image objects. Ideally, the image could have been segmented to have one image object represent one clearcut or one parcel of forest.

By resampling the image to 2.25m spatial resolution, the segmentation of the image was improved. This is contrary to the idea that higher spatial resolution would produce better results. One would expect that the more detail discernable in the image, the better the image would be classified.

The orthophotograph at 0.75 m spatial resolution recorded much of the heterogeneity in a forest stand, including individual trees, their shadows and the bare ground between trees. However, the purpose of this study was to gather larger scaled details of the landscape, ie. separating forests from clearcuts. Therefore, individual tree details deter from the generalized classification and resampling the data created a slightly more homogeneous landscape to classify.

The resampling of the image to a lower spatial resolution was beneficial in this study because the objects of interest were larger in size. If the intent of the study were to count individual trees, as was the case of Nughuro *et al.* (2002), resampling the image would deter from identifying individual trees. It is likely that there is a correlation between spatial resolution or pixel cell size and the size of the objects of interest. In this study, the reduction or smoothing of data through resampling enabled the software to see the forest in spite of the trees.

A good rule of thumb to remember when deciding what spatial resolution would be appropriate before beginning the analysis is to consider the size of the objects of interest. Ideally the spatial resolution would be smaller than the objects of interest, but not so small that the object of interest cannot be delineated because of all the detail in the higher spatial resolution.

The process to determine the classification rules was tedious. Each spectral layer feature (e.g., mean, standard deviation and ratio) and shape feature was examined and repeated adding and deleting features to determine the best combination of multiple spectral and shape features. Ultimately, the mean and standard deviation of the spectral values of the image segment for the panchromatic image and the contrast texture measure proved to be the most useful for distinguishing different classes. The shape features and neighbourhood statements were invaluable in correctly classifying image objects that had similar spectral characteristics. For example, the length to width ratio for the roads and the adjacency of rivers to sandbanks and ditches were instrumental in correctly classifying those features.

This could potentially be an improvement over the pixel-based classification because pixel based should theoretically have not been able to resolve the confusion between similar spectral characteristics in the panchromatic data and could have misclassified many pixels. For example, newer clearcuts and roads have very similar spectral characteristics, but the use of the shape characteristic of length to width ratio enabled the two classes to be separated.

In order to be able to correctly classify river image segments, adjacency rules needed to be applied. For example, if the river were next to a sandbank, it was likely that the polygon of darker spectral value was a river. The river was identified in the orthophotograph though the use of the statement about border to neighbours with sandbanks and ditches. In the case of the rivers, their spectral values were not sufficiently unique to classify the image objects correctly because they overlapped with other classes. But the adjacency rules enabled an image object with an ambiguous classification to be definitely assigned to the river class.

The accuracy assessment was conducted with 300 random points in the image. The results were very positive and suggest that all the features of interest can be correctly classified using object-oriented image classification. The classification rules for Sand Banks and Ditches need to be modified to improve the separation of those two classes.

5.2. IRS Imagery

The key to successful image classification using object-oriented techniques is to first be able to properly segment the image into meaningful objects. In the case of the IRS imagery, it was not possible to create an image segment level or a combination of several layers to correctly delineate the objects of interest. In particular, the segmentation, even at the lowest scale parameter, created blocky, stepped segments that misrepresented objects.

While it is not entirely certain what caused the blockiness of the segmentation, it is expected that it is due to the large image dimensions of the IRS scene. The segmentation parameter "scale" is used to tell the segmentation algorithm about how many image segments to create during the segmentation process. It is dependent on the number of pixels being segmented. A scale value of 10 in a small image will create a few image segments, while a scale value of 10 in a large image will produce many small image objects. With a very large scene, the scale value will not go below the threshold value of 1. The result is that the segmentation algorithm will only produce a blocky segmentation because there is a fixed number of image objects to create, with a minimum size. If a feature of interest contains only a very small number of pixels, it will be segmented as a larger image object because the segmentation algorithm cannot make a smaller image object.

The suggestion that the blocky segmentation is due to limitations in the scale parameter of the segmentation algorithm only that. Further investigations in altering the image dimensions and the specific behaviour of the segmentation algorithm would be necessary to fully understand the unsatisfactory segmentation results.

However, the subset IRS scene did not produce a blocky segmentation. The purpose of subsetting the IRS scene to cover a spatial extent similar to the subset orthophotograph was to determine if it was possible to achieve a less blocky segmentation than was being experienced with the large IRS scene. The segmentation of the subset IRS produced reasonably shaped objects, based on a 5m pixel size. Therefore the problems in the larger image are most likely not the result of the spatial resolution.

It is probable that the image used to test the technique on IRS imagery was too large for the segmentation algorithm in eCognition. The image contained over 40 million pixels. The lowest segmentation level needed to group large numbers of pixels together to form initial image objects created very blocky initial objects. The first level of segmentation of the IRS imagery produced blocky segments that did not delineate narrow roads, meandering rivers, seismic lines and other features very well.

When fewer pixels were used, either by using a smaller image or the same size image with lower spatial resolution, the segmentation algorithm is better able to segment the image properly.

Another possibility that would improve the segmentation, especially of linear features, would be to use the 8 corner pixel neighbourhood segmentation method. In cases where the linear feature is one or two pixels in width, the 8 corner pixel neighbourhood method would be able to connect pixels with similar spectral characteristics into thin linear features.

Seismic lines were not detected at all with the segmentation, even though several seismic lines were clearly visible in the landscape. Road pixels were also poorly grouped into image objects. Typically, they were joined with adjacent, non-road pixels. Where segments delineating roads were detected, it was only in pieces, separating small lengths of roads when segmentation occurred at the upper levels.

The addition of the DEM was helpful in the classification process. It was used to identify features at high elevation (Bare Rock) as well as features that were relatively lower in than the surroundings, *i.e.* in a depression. The ratio of the DEM value of an image object divided by the sum of all spectral layer mean values, called “Ratio to DEM” is displayed in Figure 21. It was used to highlight objects such as rivers and clearcuts that are typically lower than their surrounding areas.

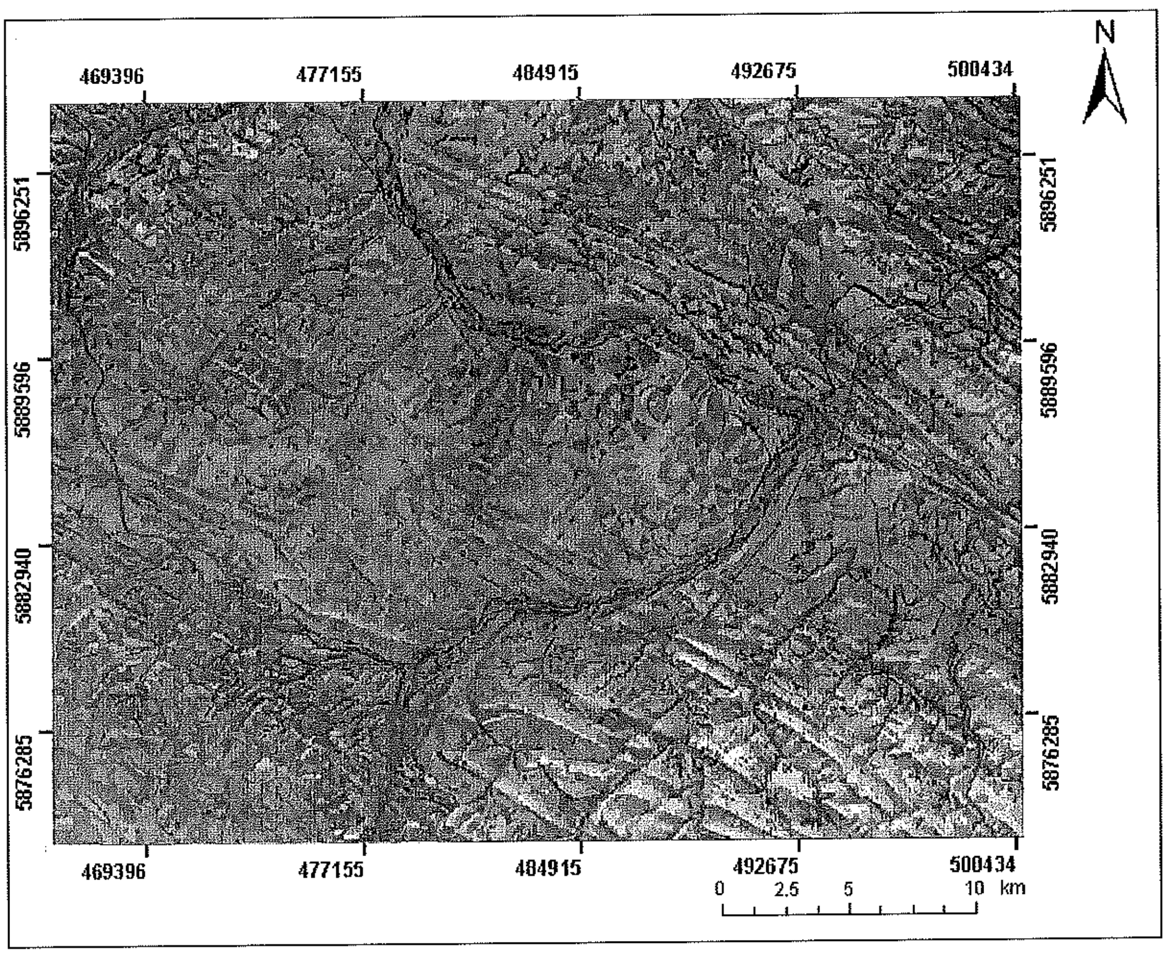

Figure 21. Ratio of DEM to other image layers. Highlights rivers and roads that are lower elevation than the surroundings.

The classification rules were made through membership functions. Overall, the classification was very successful for bare rocks and forested areas, but was not successful for all linear features and well site detection. While the overall accuracy for the classification of the IRS imagery was high (82%), the main features of interest were classified with a much lower level of accuracy. The majority of the accuracy assessment points were forest and were classified quite well. An evaluation of each feature class will be described below.

The mine was classified well, primarily due to the spatial restriction of where it could be situated. This would require *a priori* knowledge of the study area to create the membership functions of proximity to the centre point of the mine.

Roads were not segmented well, nor were they classified well. This is likely due to the fact that the image segments did not have the shape and other important characteristics that would distinguish them from other classes. Road characteristics like: width, length to width ratio, shape index and other shape characteristics should make it easy to separate roads from other classes.

Rivers and sand banks surrounding rivers were not well classified at any level of segmentation. The shape characteristics that could have been used to classify the image objects depend on the width, length to width ratio, and a shape index to take into account the meandering of the river. Sand banks or rivers could have been also classified based on a combination of their spectral value and texture but more importantly, their proximity to each other. For example, if sand banks were already properly identified in the image, the rivers could be classified based on the fact that they are nearly always adjacent to the sand bank class.

There was some confusion between classified clearcuts and sparsely vegetated areas. Recent clearcuts have a uniform shape with smooth edges. The standard deviation of the spectral value is usually low within clearcuts as well. These characteristics helped to classify the clearcuts at Level 9 of the segmentation. The older clearcuts tended to have rougher edges and less uniform shapes and a higher standard deviation because of the regrowth of vegetation. The characteristics of the older clearcuts were very similar to those of natural, sparsely vegetated areas in the panchromatic imagery. Therefore, there were mediocre results in distinguishing between old clearcuts and natural, sparsely vegetated forests.

Well sites have a very unique, square shape in a natural landscape and because of their small size and spectral characteristics (high spectral value, low standard deviation) they should have been relatively easy to classify. Due to the segmentation procedure, the shape of the well sites was altered to be less square and wellsites were often merged with the adjacent road leading to the site. Also, the segmentation process at Level 4 created many other image objects that were the same size and shape. Therefore, most wellsites were correctly identified with the classification rules but many other non-wellsite image segments were also accorded the Well site label. Correct segmentation and grouping at higher levels could solve this issue and enable more well sites to be correctly classified.

Finally, forest shadows were distinguished well using mostly the spectral characteristics of the panchromatic band as forest shadows tend to be very dark.

5.3. Image Comparison

This study examined the application of object oriented image classification to two types of panchromatic band data. The two images differed in image characteristics, segmentation parameters, classification procedures and results, and the amount of time required to determine and execute the analysis.

5.3.1. Spatial Differences

The significant differences in the images were the spatial resolution and the spatial extent. The orthophoto had significantly higher spatial resolution than the IRS imagery at 0.75 m compared to 5 m, though less so at 2.25m. With the improved spatial resolution, many details of the forest and the features of interest can be discerned. However, the analysis in this study is coarser than the detail than the orthophotos provide. For example, this study was concerned with differentiating between forested areas and clearcuts, not in counting every tree in the forest, or the gaps between trees in the forest. With the improved spatial resolution comes more spectral heterogeneity in the image object, which can be detrimental to properly segmenting and classifying the image using object-oriented image classification.

When the orthophotograph was resampled to 2.25 m spatial resolution, the heterogeneity in the data was reduced and it was easier to segment and classify the image. All the relevant details in the image were preserved; including the length of roads and boundaries of clearcuts and all the detail about the forest was reduced to a more manageable set of spectral information.

The IRS image has a lower spatial resolution, but it is still sufficient to identify nearly all the features properly. Interestingly, a road with a width of approximately 5 m could completely cover several pixels at 0.75 m resolution, but would only be a line feature at 5 m spatial resolution. This fact will influence the type of segmentation that would need to be performed. When the width of the feature of interest is larger than the pixel size, creating an area of coverage, pixel-neighbour segmentation is preferred. Alternately, when the width of the feature of interest is larger than the pixel size, the 8 corner neighbourhood segmentation is recommended because it considers all 8 pixels surrounding the pixel of interest during the segmentation process. In the case of a road 5m wide in an image with 0.75 m spatial resolution, the road would cover an area and could be successfully segmented using only the 4 neighbours in pixel-neighbourhood segmentation.

The spatial extent of the two images is much different when comparing the subset orthophotograph and the large IRS scene. The reason the two images were evaluated was to determine the optimal set of panchromatic data to classify the objects of interest. In this study, the extent or dimensions of the image and the pixel size was varied and compared. The subset of the orthophotograph covers an area approximately 2.25km x 2.25km (5.0 km^2) while the IRS image covers an area 35km x 30km (1072km^2). The number of pixels is also different by an order of magnitude (1 million vs 43 million). This has potential implications for the segmentation of the image because the scale factor needs to be set to determine the number of image objects, hence the number of pixels per object, to create. At the same segmentation scale factor, the segmentation will force many more pixels to be assigned to each object with the IRS imagery than the orthophotograph because of the greater number of pixels in the scene.

The spatial extent also influences the number of features that can be identified in one classification process. The much smaller orthophotograph only identified a small number of image objects. If that spatial extent were used, it would take many iterations through the program to identify all the features in the Foothills Model Forest. The IRS image covered a much larger area than the orthophotograph; hence, many more features were classified making it easier to classify large tracts of land.

Finally, the spatial extent determines how many features will be identified per image, which in turn influences how many image objects will need to be matched up and merged into a final vector coverage. The large IRS scene would need less merging with other images than the many orthophotographs that are needed to create the same coverage.

5.3.2. Segmentation Differences

There were differences in the segmentation of the two types of panchromatic band imagery. The orthophotograph was segmented successfully using one segmentation level whereas the IRS imagery needed up to 9 levels of segmentation to create meaningful image segments. The number of levels needed to correctly segment the image depends on the similarity in size of the desired image classes. It was the case for the orthophotograph that the roads, clearcuts and river were all approximately equal in size. Therefore, the first and only level of segmentation created the required image objects of similar size.

In the large IRS image, areas of forest, bare rock and other features were large compared to small well sites and thin roads. Therefore, lower levels of segmentation created small image objects to represent well sites and roads while higher levels of segmentation were needed to create large image objects to represent the large features such as forest, bare rock and clearcuts. It is possible that all 9 levels of segmentation were not needed for the IRS image. Further manipulation of the segmentation parameters could reduce the number of segmentation levels needed to correctly delineate the features of interest.

The segmentation of the orthophotograph was successful in delineating all objects of interest at the 2.25m spatial resolution. The roads were delineated especially well because of the consistency (*i.e.*, low standard deviation) of their spectral values as well as the fact that the roads covered an area of pixels, not a linear chain of individual pixels. The roads were delineated much better when compared to the IRS imagery on which only a fraction of the roads were delineated correctly. The segmentation of IRS imagery produced image segments that represented discontinuous sections of roads, and several image segments consisted of a section of road and the non-road pixels surrounding it. The result was the shape characteristics could not be produced for the image segments to help differentiate the linear features from other image objects.

5.3.3. Classification Differences

The orthophotograph produced a better image classification than the IRS imagery in this study even when the resampled orthophotograph was not that different in spatial resolution. The image classes were more detailed than could be identified in the IRS imagery. For example, the Clearcut class in the IRS imagery could be separated into Older Clearcuts, with some vegetation regrowth and Newer Clearcuts, with less vegetation regrowth using the detail in the orthophotograph. The Road class in the IRS image could be divided into the actual Road and the non-vegetated Ditch surrounding the road using the orthophotograph. The advantage of the orthophotograph is a result of the higher spatial resolution of the resampled image. The intermediate resolution of 2.25m proved to give the best classification results because the 0.75m resolution contained too much information for the goals of this study and the 5m IRS image did not provide enough.

Besides the details in the orthophotograph, the other reason the classification results were superior is due to the spectral and shape characteristics of the image objects accurately representing the features of interest. Image objects such as sections of the river and sand bank, and roads could be classified using their shape and spectral characteristics, such as length to width ratio. In the IRS imagery, the segmentation did not produce meaningful objects, so the shape characteristics of the features, such as roads and rivers, could not be used to help the classification. Panchromatic imagery has the limitation of only 1 spectral band. This can cause ambiguity about which spectral values to associate with which image class. For example, the spectral range for sparsely vegetated forests and clearcuts is very similar, but the two classes have very different shape characteristics, which can be exploited to differentiate the two classes.

The final difference in classification between the two types of images is the use of a digital elevation model to assist in the classification. The DEM for the area has a spatial resolution of 30 m. At that resolution the DEM was not especially helpful in the classification of the small orthophotograph scene as there was only a slight change in elevation. For the IRS image, a large area was covered in the scene and there was a large range in elevation values. The DEM was particularly useful in separating the Bare Rock and the Vegetation near the Bare Rocks classes from the rest of the image. The ratio of the DEM to the rest of the image was helpful in identifying roads and rivers in the image.

The subset orthophotograph at 2.25m spatial resolution and the subset IRS scene at 5m spatial resolution can also be briefly compared. When the classification rules for the orthophotograph were applied to the IRS subset, the results were not the same. This is due to a number of factors. The contrast texture measure was applied to the entire IRS scene, then was subset, thus the contrast in the larger scene did not describe local differences in contrast, important for the separation of forested and non-forested areas. Also, the standard deviation of the forest polygons was diminished because of the smoothing of the data due to the larger spatial resolution. The mean spectral values of the roads were also lower because the brightness of the roads was decreased due to the combination of the surrounding vegetation in the 5m pixels.

5.3.4. Time Differences

Lastly, the difference of time between the two images will be compared. The orthophotograph was subset to a smaller area to improve the processing time. Analysis of the orthophotograph was a first attempt at object-oriented image classification, thus the learning curve took much longer with the orthophotograph than the IRS image.

The orthophotograph was segmented, built polygons and was classified quicker than the IRS image. However, the IRS image covers a greater area, thus the amount of time to classify a set of orthophotographs that cover the same geographic area would take longer to process for two reasons. First, the orthophotograph at 2.25m spatial resolution has approximately four pixels to every one pixel in the IRS image, hence a greater volume of data to segment. Second, the smaller orthophotographs would need to be merged and radiometrically correctly before processing. Alternately, the orthophotographs could be classified with object-oriented techniques, then the classified polygons could be joined and adjusted to match in post-processing.

5.4. Advantages and Disadvantages of Object-oriented Image Classification

Some clear advantages of object-oriented classification emerged as a result of this study. For the processing of panchromatic images of the study area, object-oriented image classification reduces the large volume of highly detailed data in the image to manageable image segments. The IRS scene contained 43 million pixels, by reducing the data to fewer than 2000 image objects. Ideally, those image segments should have been segmented correctly to yield image objects with meaningful spectral and spatial characteristics. When large-sized panchromatic images are typically classified using per-pixel methods, the large volume of data can increase the processing time. With object-oriented image classification, this study was able to reduce the volume of data to a fraction of that of the original data.

A second advantage of object-oriented image classification is the greater success in classifying the panchromatic images, especially the orthophotograph. The scattered, discontinuous classification was eliminated. In forest stands, with the very high spatial resolution of the images, trees and gaps between trees are represented by individual pixels. Traditional classification would have had non-forest pixels mixed in a forest stand. The intent of this study was to delineate natural and anthropogenic features in the landscape, and uniform tracts of forest needed to be identified. With object-oriented image classification, this type of classification was achieved with the orthophotograph and was attempted with the IRS image.

Third, eCognition produces vector classification data as well as raster data for export. This offered an advantage over per-pixel classification, where the results are typically exported as raster data, needing to be converted, post classification, to vector. With object-oriented classification, the benefits of vectorizing the image can be exploited even during the classification process. The topological descriptions, in terms of shape characteristics, as well as the spectral properties of the polygon, were used to properly identify features in the landscape. With per-pixel methods, confusion between the overlapping spectral range of roads, sand banks and ditches would have made it nearly impossible to separate those classes. With the shape characteristics, and the enhanced spectral values (such as standard deviation of the image object's spectral values), amounting to approximately 100 characteristics, it was possible to distinguish between these classes.

Finally, the object-oriented program eCognition has protocol tools available to automate nearly every stage in the classification process. The segmentation, class creation, classification rules, classification procedure, refinement, accuracy assessment and final vector or raster file export could be automated and saved in a file. Slight modification to any of the steps, for example the segmentation parameters, can be written into a file, so different scenarios may be run through the automation procedure.

Several major disadvantages of using object-oriented image classification for this particular set of images were also encountered. The main drawback of using object-oriented image classification was the amount of time required to learn the eCognition program and the steps to classify the image. This is because there are no clues or advice on how to determine the correct segmentation parameters for the image. A minimum of

10 attempts at determining the first level of segmentation was needed for the orthophotograph. Approximately 50 attempts were needed to determine the segmentation parameters and the number of levels of segmentation to create the IRS image objects. The extra time was needed because different levels of segmentation needed to be created to capture small and large features in the landscape.

The classification rules also took an exceedingly long time to obtain the correct combination of rules. The nearest neighbour classifier in the more recent version of eCognition (version 3.0) features an option of the program suggesting the best classification rules to separate classes. The suggestion is based on a set of samples from each class.

The second main problem of using object-oriented image classification for this project has to do with the limitation of the panchromatic data. With only one band of spectral values, it was a challenge to properly segment and classify the images. Multispectral information, and vegetation indices based on those bands, would have helped in separating water features and vegetated areas from the other features.

Lastly, this study was unsuccessful at obtaining a highly accurate classification of the IRS imagery. The failure to accurately classify the IRS scene could be due to limitations in the segmentation algorithm when confronted with a very large scene, but is mostly due to a lack of time to develop the appropriate segmentation parameters and the classification rules.

5.5. Recommendations for Future Work

Several recommendations to improve the classification of the study area are outlined below. The recommendations concern the choice of image data, the segmentation procedure and the classification rules needed to accurately classify the images.

5.5.1. Choice of image data

The addition of high resolution, multispectral data, with a maximum spatial resolution of 30 m, likely would have improved the identification of most of the features of interest. The river was not classified in the IRS image but with the multispectral data, the river would more easily be segmented and classified well. The sand banks could next be classified as being adjacent to the river. Barren Clearcuts could be distinguished from forested areas by vegetation indices. Roads could be identified by their shape and spectral characteristics. It is doubtful that it would be a disadvantage to use multispectral data, if it were available for the study area, especially if the multispectral data were high resolution.

5.5.2. Segmentation

The segmentation parameters need to be improved for the IRS imagery. As previously explained, the segmentation needs to be nearly perfect because it forms the basis for the classification. 8 corner pixel neighbourhood segmentation could be evaluated for identifying linear features whose width does not have an area, *i.e.* is a single pixel. This could help create long, slim, continuous image segments.

Determining how many and the appropriate segmentation levels to properly capture the different feature sizes is also critical. At this stage, 9 segmentation levels were needed, but it is likely that fewer levels would be needed to reduce the data to meaningful objects.

Further examination of subsetting the IRS image may give better results with the segmentation algorithm in this version of eCognition. Likewise, it would be interesting to apply the segmentation parameters to a larger orthophotograph scene. By subsetting a large image it is possible to determine the correct classification rules and manipulations of the data with a more manageable data set before applying the rule set to a large scene.

5.5.3. Classification rules

The classification rules could be improved though the use of the nearest neighbour classifier. The goal is to use as few classifier rules as possible to keep the classification simple and repeatable for different scenes. The nearest neighbour classifier indicates the improvement in classification with each classification rule added until the separability of classes' peaks.

6.0. Conclusion

The last few decades have seen an increase in the use of object-oriented image analysis for image classification. The technique offers several advantages over pixel based image classification, namely the ability to process high resolution data, reduced speckle effect, and the production of meaningful image objects that are GIS-ready. Additionally, object oriented image classification enables the classification of panchromatic data. Panchromatic data contains limited spectral information and it would not be possible to achieve highly accurate classifications based on spectral information alone. The object oriented technique was able to incorporate a range of spatial information, adjacency and logical rules to increase the amount of data to separate the classes of interest.

This study used eCognition object oriented software to determine if object oriented techniques were appropriate for classifying two types of panchromatic images. A subset of an orthophotograph produced a highly accurate classification of human disturbance as well as natural features. A large IRS scene was not classified as accurately. Difficulties in properly segmenting the scene to delineate image objects at different levels prevented an accurate classification. A subset of the IRS scene yielded better segmentation results.

This study was a preliminary investigation of the applicability of object oriented image techniques to panchromatic data. Early results suggest that more work is needed to create an accurate map of human disturbances and natural features. Future work recommended in this study includes investigating the newest version of eCognition and nearest neighbour classifiers, subsetting the IRS scene to a more manageable size for the segmentation algorithm and reducing the amount of time to determine the segmentation parameters, classes and classification rules. The addition of multispectral data would also facilitate the identification of human disturbance features.

The object oriented analysis field is still in development and its techniques continue to improve. With improvement to object oriented theories and technology, it may be possible to increase the accuracy in classifying panchromatic images.

REFERENCES

Bird, A.C., Taylor, J. C., Brewer, T. R., 2000. Mapping national park landscape from ground, air and space. *International Journal of Remote Sensing*, 21(13 & 14), 2719-2736.

Blaschke, T., Hay, G., 2001a. Object-oriented image analysis and scale-space: theory and methods for modeling and evaluating multiscale landscape structures. *International Archives of Photogrammetry and Remote Sensing*, 34(part 4/W5), 22-29.

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

De Kok, R. Schneider, T, Ammer, U., 1999. Object based classification and applications in the Alpine forest environment. *Proceedings Joint ISPRS/EARSeL Workshop "Fusion of sensor data, knowledge sources and algorithms"*, Valladolid, Spain.

Definiens Imaging, 2001a. *eCognition User Guide*, Munchen, Germany.

Definiens Imaging, 2001b. eCognition White Paper, Product overview. Definiens Imaging web site : www.definiens-imaging.com/down/whitepaper.pdf, retrieved March 27, 2002.

Flanders, D., Hall-Beyer, M., Pereverzoff, J. 2003. Preliminary evaluation of eCognition object-based software for cut block delineation and feature extraction. Canadian Journal of Remote Sensing, 29(4) pp. 441-452.

Foothills Model Forest Network, 2003. FMF web site: http://www.fmf.ab.ca/abo.html, retrieved June 23, 2003.

Fuller, R., Groom, G., Jones, A. 1994. The land cover map of Great Britain: An automated classification of Landsat Thematic Mapper data, *Photogrammetric Engineering and Remote Sensing*, 60(5), 553-562.

Geneletti, D., Gorte, B. 2002. A method for object-oriented land over classification combining Landsat TM data and aerial photographs. *International Journal of Remote Sensing*. In press.

Gillis, M., Leckie, D. 1996. Forest inventory update in Canada. *The Forestry Chronicle*, 72(2), 138-156.

GISCafe, 2001. Definiens AG presents eCognition: A completely new dimension in image analysis. Web site: www.giscafe.com/NEWS/CorpNews2/20001205_eCognition.html, retrieved March 26, 2002.

Hall, R. Kruger, A., Scheffer, J., Titus, S. Moore, W. 1989. A statistical evaluation of LANDSAT TM and MSS data for mapping forest cutovers, *The Forestry Chronicle*, 441-449.

Hall, R., Dams, R., Lyseng, L. 1991. Forest cut-over mapping with SPOT satellite data, *International Journal of Remote Sensing*, 12(11), 2193-2204.

Hoffman, P., 2001a. Detecting buildings and roads from IKONOS and DEM data with eCognition, *eCognition Application Notes* 2(5). Web site: http://www.definiens-imaging.com/index.htm, retrieved March 20, 2002.

Hoffman, P., 2001b. Using eCognition to detect informal settlements from IKONOS images, *eCognition Application Notes* 2(8). Web site http://www.definiens-imaging.com/index.htm, retieved March 20, 2002.

Hofmann, P. 2001c. Detecting urban features from IKONOS data using an object-oriented approach. *Remote Sensing and Photogrammetry Society (Editor): Proceedings of the First Annual Conference of the Remote Sensing and Photogrammetry Society*, September 12-14, 2001: 28-33.

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

Innes, J., Koch, B. 1998. Forest biodiversity and its assessment by remote sensing, *Global Ecology and Biogeography Letters*, 7, 397-419.

Jensen, J. 1996. *Introductory Digital Image Processing*; Prentice-Hall, Inc., Upper Saddle River, N. J. pp. 318.

Lillesand, T., Kiefer, R., 1994. *Remote Sensing and Interpretation*, 3rd edition, John Wiley and Sons, Inc, New York, pp. 750.

McKeown, D., 1988. Building knowledge-based systems for detecting man-made structure from remotely sensed imagery, *Philosophical Transactions of the Royal*

Society of London. Series A, Mathematical and Physical Sciences, 324(1579), 423-435.

Meinel, G., Neubert, M. Reder, J. 2001. The potential use of very high resolution satellite data for urban areas – First experiences with IKONOS data, their classification and application in urban planning and environmental monitoring. *2nd Symposium of Remote Sensing of Urban Areas, Regensburg* 35, 196-205.

Neubert, M. 2001. Segment-based analysis of high resolution satellite and laser scanning data. In: Hilty, L., Gilgen, P. *Sustainability in the Information Society, Proceedings of the 15th International Symposium Informatics for Environmental Protection, Zurich*, 379-386.

Nugroho, M., Suryokusumo, B., Hoekman, D., de Bos, R., 2002. Evaluation of Interferometric Radar Tree Mapping Parameters using Tree Grouping Approach, *Proceedings from ForestSAT Symposium, Edinburgh, August 5-9, 2002.*

Pedley, M., Curran, P. 1991. Per-field classification: An example using SPOT HRV imagery, *International Journal of Remote Sensing*, 12(11),2181-2192.

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.

Pereverzoff, J., 2002. An evaluation of linear feature detection in orthophotos using pixel and object based image analysis techniques, *Final project report Geography 795.17,* University of Calgary.

Quegan, S., Rye, A., Hendry, A., Skingley, J., Oddy, C., 1988. Automatic interpretation strategies for Synthetic Aperture Radar images, *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences,*324(1579), 409-420.

Schiewe, J., Tufte, L., Ehlers, Mr. 2001. Potential and problems of multi-scale segmentation methods in remote sensing. *GeoBIT/GIS* 6, 34-39.

Schwarz, M., Steinmeier, C. Waser, L. 2001. Detection of storm losses in alpine forest areas by different methodic approaches using high-resolution satellite data. *21st EARSeL Symposium, Paris.*

Smith, G., Fuller, R. 2001. An integrated approach to land cover classification: An example in the Island of Jersey, *International Journal of Remote Sensing*, 22(16), 3123-3142.

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*, 60-76.

Wilhauck, G., Schneider, T., De Kok, R. Ammer, U. 2000a. Comparison of object-oriented classification techniques and standard image analysis for the use of change detection between SPOT multispectral satellite images and aerial photos. *ISPRS*, 33.

Willhauck, G., Benz, U.C. Siegert, F. 2000b. Semiautomatic classification procedures for fire monitoring using multitemporal SAR images and NOAA-AVHRR hotspot data. In: *Proceedings of the 4th European Conference on Synthetic Aperture Radar*, Cologne, Germany.

Yatabe, S., Leckie, D. 1995. Clearcut and forest-type discrimination in satellite SAR imagery, *Canadian Journal of Remote Sensing*, 21(4), 455-467.