RAPID ROAD INVENTORY USING HIGH RESOLUTION SATELLITE IMAGERY

SITANSU BHUSHAN PATTNAIK

NATIONAL UNIVERSITY OF SINGAPORE
2003

RAPID ROAD INVENTORY USING HIGH RESOLUTION SATELLITE IMAGERY

SITANSU BHUSHAN PATTNAIK
(B.E. GeoInformatics, Anna University, Chennai, INDIA)

A THESIS SUBMITTED
FOR THE DEGREE OF MASTER OF ENGINEERING
DEPARTMENT OF CIVIL ENGINEERING
NATIONAL UNIVERSITY OF SINGAPORE
2003

Dedicated to

Rajkumar Arumugam (coke) and **Google.com** for always being there.

Acknowledgements

I am extremely grateful to Prof. Chan Weng Tat my main supervisor for believing in me and being the guiding force for my research. I am thankful to my co-supervisor Dr. Cheu Ruey Long for his support during the research.

Goh Yeow Meng and Poonam Saxena from MITL lab were very helpful in allowing me to start my research and get used to the new research environment. I am also thankful to TMSI for providing the infrastructure at the outset for my research.

I am also indebted to the National University of Singapore, for providing me an opportunity to pursue independent research towards my Master's degree. Without the help of the staff in the department of civil engineering, data collection and funding would not have been possible.

My parents, brother and bunny were my main source of motivation to complete my thesis. Last but not the least; I am grateful to my friends for their constant support and encouragement.

SITANSU BHUSHAN PATTNAIK

Table of Contents

Abstract

Human Operator has been the most important player in feature extraction procedures due to the variability and complexity of the scenarios under consideration. In this thesis an attempt has been made to utilize the power of Artificial Intelligence by the application of a self-organizing map algorithm along with traditional image processing algorithms to build a road inventory system. Satellite imagery from IKONOS was used to detect and extract roadway features in the central business district of Singapore. Edge detection from the high-resolution image has been coupled with the multi-spectral information for better classification results.

Key words: Road Inventory, Self-organizing maps, Neural Network, Edge Detection

List of Figures

List of Tables

CHAPTER 1
Introduction

1.1 Introduction

Singapore has about 3000 kilometers of roads of which the expressway network after completion would amount to 155 kilometers [1]. Major road expansions and improvements are being undertaken continuously. These make it imperative to maintain an accurate and comprehensive dataset describing the road network.

The proposed technique in this thesis is useful for situations where: (a) the road network is experiencing rapid change; (b) it is necessary to obtain a quick update of the status of the network, certainly within a time-frame that is shorter than that possible with traditional surveying techniques and (c) the network is embedded in an urban or semi-urban environment. Singapore is a good example of (a) and (c) although the size of the city-state is such that (b) does not necessarily apply. Cities experiencing urban sprawl and where there has not been a consistent effort in maintaining current records of the road network are prime targets for the application of the technique developed in the thesis. Singapore happened to be chosen as the subject of study in this thesis because (a) it is an urban city-state with a fairly extensive road network that traverses both highly urbanized and semi-urban landscapes, (b) the ready availability of high resolution satellite imagery, and (c) ease of verification through field studies and comparing against other data sets.

Road networks are always in a continuous state of being up-graded or repair; hence, the accuracy and timeliness of the data depicting the network are critical for analysis. Road inventory data for the networks include the alignment, length, condition, width and placement of signs etc. Much of the inventory data collection by transportation agencies is time consuming and labor intensive, requiring on-site surveys using traditional survey equipment. More recently, surveys using video logging and GPS measurements have been introduced to improve the productivity of the whole process. Inventorying road networks over large areas poses challenges in ensuring the currency and completeness of the datasets if traditional surveying techniques are employed exclusively. The use of image processing techniques on remotely sensed images presents an attractive alternative for acquiring road network inventory datasets; this area has also been an active area of research.

Remote sensing is the science and art of obtaining information about an object, area or phenomena through the analysis of data acquired by a device that is not in contact with the object, area or phenomena under investigation [2]. Remote sensing has been useful for obtaining a synoptic view of a widespread study area in a short time.

Traditional image processing algorithms for feature extraction, mainly based on edge detection and distribution of radiometric values, are on high resolution satellite imagery from advanced imaging systems due to the increase in spatial and spectral resolution. Context information is necessary for understanding the datasets for spatial filtering.

Road features are essentially linearly elongated having varying surface characteristics due to the pavement material, age, markings and repair patches. These diverse spectral signatures coupled with the various configurations of the road networks make it difficult for rule-based algorithms to identification and extract road features. Algorithms to decrease the level of human effort have not entirely succeeded in realizing Automatic Feature Extraction (AFE) for integration with a Geographical Information System (GIS) database. A level of human intelligence needs to be incorporated in the algorithm to cope with the wide variability of possible scenarios presented in high-resolution satellite imagery.

Artificial intelligence attempts to mimic a human being's capability to reason and generalize rules based on experience and previous knowledge. Artificial neural networks are adaptive algorithms that attempt to emulate the information processing and associative retrieval characteristics of the human brain. The hypothesis of this thesis is that the use of these algorithms, in conjunction with traditional classifiers, will improve feature extraction for integration with the current GIS databases.

1.2 Scope of the Thesis

The main objective of this research is to evaluate the functionality of the existing classifiers for feature extraction from high-resolution imagery and then to utilize domain knowledge from photo-interpretation and feature extraction for potential improvements in the feature extraction procedure. The traditional classifiers have to be modified to deal with the increased spectral and spatial resolution data available from advanced sensors collecting multi-spectral imagery. Research into data fusion and artificial intelligence for potential improvements in the classification accuracy was conducted. The methodology put forth utilizes the multi-spectral imagery with its spectra information and panchromatic imagery with its high spatial resolution in tandem for potential improvements in feature extraction. Post-processing techniques are applied to integrate the database in the GIS.

These proposed methodology to extract road features from remotely sensed imagery can be very useful in areas where continuous upgrading is undertaken like in Singapore and in many developing countries in Asia and Africa for developing an inventory due to the widespread areas and lack of information The major hurdle in the implementation of the procedures is the cost of acquiring the datasets. This can be overcome by cost sharing between different user groups. This would help in the creation of a unified database for change detection studies and for modeling. Remote sensing data collection has the niche in terms of the rapid data collection and the widespread area of coverage.

1.3 Layout of the Thesis

This thesis includes 5 chapters. Chapter 1 gives a brief introduction for the need and scope of the research. Chapter 2 considers the development of road infrastructure systems and its place in the transportation planning scenario and then compares the different techniques of data collection. Chapter 3 provides an exhaustive review of the feature extraction procedure by describing the different aspects of feature extraction and the development of techniques with the availability of new datasets. Modeling issues in road extraction is also discussed in this chapter. Chapter 4 deals with the methodology and logic used to develop a hybrid road feature extraction algorithm. The results obtained are discussed at the end of chapter 4. Conclusions and utility of the methodology developed are in Chapter 5. Recommendation and future scope have also been included in Chapter 5.

CHAPTER 2
Road inventory Systems

2.1 Introduction

The provision and operation of a transportation system requires a continuous planning function [3]. This planning function aims at assessing the performance of the present transportation system, to forecast future demand by identifying various solutions to the perceived problems in system performance.

The inventory of the facilities is the first step in transportation planning. Much of the transportation investment that occurs in a state or urban area is aimed at upgrading the physical condition of a facility or improving its performance [4].

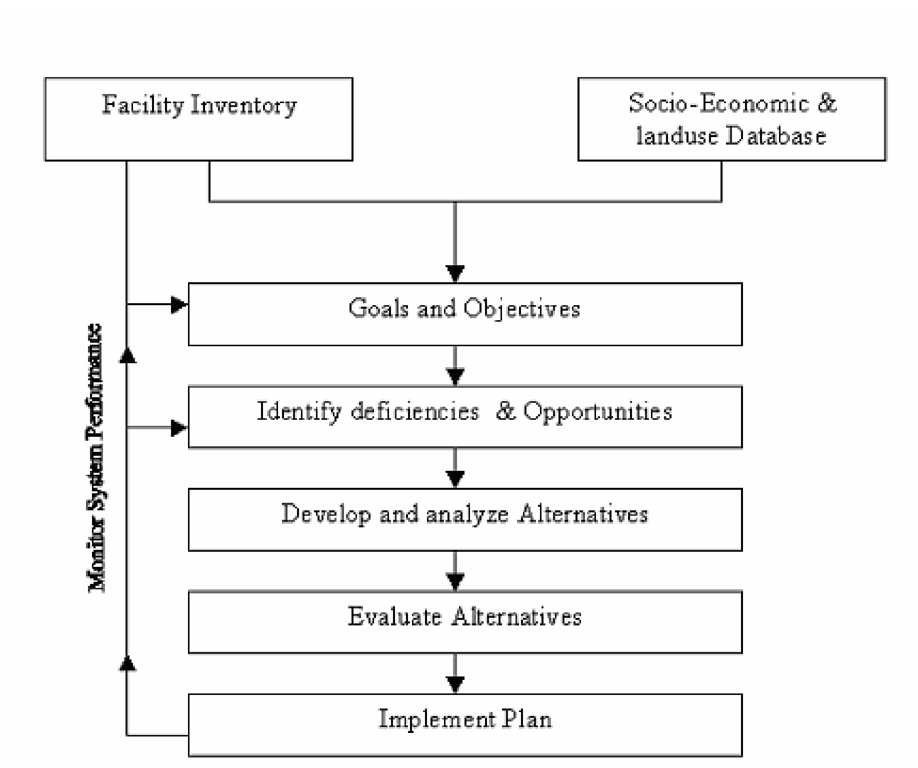

Figure 1: Basic elements of Transportation Planning [4]

Hence, inventory data is an important aspect of planning leading to judicious decision-making. The importance of inventory data in transportation planning as a whole is illustrated in Figure 1.This chapter discusses various inventorying procedures and evaluates the work undertaken in developing inventory systems.

2.2 Road Inventory Systems

A road inventory typically includes the alignment, length, condition, width and placement of signs among others. Typical highway infrastructure elements are described in Table 1.

Table 1: Typical Highway Infrastructure Elements

Category	Elements
Roadway Features	Pavement width/type/condition
	Shoulder width/type/condition
	Laneage/channelization
	Curb location/condition
	Curb cut location/condition
	Pavement marking location/type/condition/size
	Railroad crossing location/condition/control
Roadside features	Bridge location/length/width/condition
	Sidewalk location/width/condition
	Fixed object location/type/rigidity/size
	Sight distance or obstructions
	Driveway location
	Underground utility location/type/depth
Appurtenances	Sign type/location/condition/size
	Signal type/location/condition/size
	Guard rail type/location/condition/size
	Crash cushion type/location/condition/size
	Median barrier type/location/condition/size

The highway infrastructure inventories can be generally categorized [5] as:

- Traffic control device inventories that include information on the location, type, condition, operation of traffic signs and pavement markings.
- Roadway feature inventories that include information on the design characteristics of a roadway network including vertical and horizontal alignments and geometric design.
- Pavement condition inventories that describe the road surface characteristics by roughness, skid resistance and structural integrity.
- Roadside feature inventories that have the information on the roadside obstacles, safety devices and provisions of underground utilities and services.

These inventories are used primarily for maintenance, upgrading and also as the basis in planning process. Data collection methods for inventorying are influenced by the study area, financial considerations, time frame and other factors unique to the study. These data collection procedures can be broadly classified by the need to be physically present at the site as on-site and off-site surveys.

On-site survey methods include utilizing Photo-logging, video-logging, Global positioning systems, Electronic distance measurement devices, Optical measurement devices, Trundle wheel and Chains.

Off-site survey methods include aerial photography and satellite remote sensing in the optical, ultraviolet, infrared and thermal bands.

2.2.1 On-site survey procedures

Photo-logging: This technique is a photographic process whereby photographs of the highway and its environment are taken from a moving vehicle at equal increments of distance [6] resulting in a static pictorial record of the highway cross section, known as a photolog. These pictorial records are used to identify the roadside features

Video-logging: Video-logging provides a continuous visual perspective of the highway recorded on a magnetic medium [7]. The videotapes are recorded as the vehicle transverses through the road segments. Data concerning the inventory elements is entered simultaneously by using a microprocessor-based keyboard, which include the longitudinal information and the type of element on the roadway.

Global Positioning Systems (GPS): GPS is a constellation of satellites which are in specific orbits allowing a user with a receiver to fix his/her position based on the time information from the satellites by converting time units to distance units. The co-ordinates are calculated by using the time information from four satellites. The atmosphere between the satellites and the receiver affects the propagation of the signals and hence the measurements of location are influenced. These correction factors are also sent by the ground command and control centers to the satellites, which are, received by the ground receiver.

The accuracy of the GPS measurements depend on these correction factors, the capability of the receiver to calculate position from more than four satellites and the antenna configuration. Using a GPS or a Differential GPS, the road networks can be automatically mapped by traversing through the entire network.

Electronic/Optical Distance Measurements: Theodelites and Electronic Distance Measurement (EDM) devices are used to measure distances between points.

The equipment has to be located on the road while the measurements are taken.

Trundle wheel / Chaining/ Measuring tapes:

These are the two primal methods of measuring distances and these are still used to measure small sections accurately.

2.2.2 Off-site procedures

Aerial Photography: Aerial photography is taken by placing cameras in the aircraft and then capturing photographs of the study area. The camera and the film loaded are sensitive to a wide spectrum, which includes the visual spectrum, ultraviolet region, infrared region and the thermal region in the electromagnetic spectrum. The wide area of coverage enables the analyst to visualize the surface features in their spatial conformation. The photograph gives a permanent record of the existing conditions allowing change detection studies from an archived image collection.

Satellite Remote Sensing: Remote sensing from the satellite platform is undertaken by using special sensors, which focus on the earth surface and capture data. These satellites are placed in sun-synchronous orbits and depending on the altitude, the repeat period is determined. The sensors on these satellites are classified in terms of the spatial resolution and spectral sensitivity. Imagery from the sensors gives a synoptic view of the study area and the data collection can be completed rapidly. The inventory data are then extracted from the imagery by using post-processing algorithms.

2.2.3 Comparison of On-site and Off-site Procedures

The advantages of **On-site** procedures:

- Accurate measurements of inventory elements by using theodelite, EDM.
- Workforce of medium skill required for data collection.
- Ground truth verification possible while data collection using trundle wheel and chains.
- Photo-logging/video-logging reduce the dangers to the workers on-site.
- The time taken by photo-logging/video-logging is lesser than Theodelite and EDM surveys.
- The equipment requirements for onsite survey procedures are not expensive and most the equipments are already with the surveying agencies.

The disadvantages of **On-site** procedures:

- Theodelite, EDM surveys are time consuming processes.
- Labor intensive
- Danger to the workers during data collection using Theodelites and EDM equipments.

The advantages of **Off-site** procedures:

- Rapid data collection is possible due to the enhanced viewing opportunity of the sensor based on the satellite platform and with aerial photography the possibility of flying the plane for data collection by over-flying the study area in good weather exists.
- The contextual information gathered from the imagery is useful for analysis.
- Minimal labor costs for data collection.

The disadvantages of **Off-site** procedures:

- High cost of the equipment and the imagery
- Highly skilled workers required for data extraction

The synoptic view from the aerial and satellite platform allow the analyst to asses the entire environment for planning procedures. Information extraction from the raw data is a complex and time-consuming process, however, the utility of the information extracted offsets the cost procedures. On-site procedures are traditional methods and hence are widely used presently but have inherent flaws when the volume of data collection is very high. With the development of new technology, the computing power of the machines have increased rapidly and hence there is a shift from analog to digital form of data collection and processing, which is more suitable for computer analysis. Researchers have devised procedures to upgrade from the traditional methods of surveying to utilize the capabilities of the modern sensors and computing power. The issues of feature extraction, spawned from the technological advancements are discussed in the subsequent Chapter.

2.3 Feature Extraction for Road Inventory

As discussed in the previous section, off-site procedures are useful in rapid collection and analysis of inventory elements. The data collected from off-site(remotely sensed) procedures have to be processed prior to being useful. The conversion from grayscale values in the imagery to usable information is made possible by feature extraction.

One of the earliest work in data collection for road inventory by using photologs was undertaken in the State of Iowa for generating a photolog of the highway system. Work undertaken by Office of Traffic Operations, Bureau of Public roads in the Washington DC area in the 70's aimed to have a complete "static pictorial record" of the highway system [12]. Camera equipment was tested and cost estimates were derived to generate photolog at a per mile basis. One of the main recommendations was to utilize color film for better interpretation which highlighted the need for better data quality.

Research at Vehicle Intelligence & Transportation Analysis Laboratory, University of California, Santa Barbara for developing spectral signatures of different road pavements is very useful for developing classification algorithms based on the spectral information. The work is especially significant with the availability of hyper-spectral data for urban areas which exhibit a mixture of spectral signature due to the landuse pattern where different landuse classes co-exist in a very close proximity. [14]

Dan Lee for ESRI developed an interactive procedure in ArcINFO environment to simplify building outlines and derive road centerlines from vectorized outlines [15]. The paper discuses the strategies for deriving generalization solutions useful for automating feature extraction procedures for decreasing the production time. The need for automation is put forth by the way of recommendations for reducing the human intervention to a minimum.

Doucette et al (2001) indicate the extraction procedure to include two steps. Firstly, the identification of the feature and then the delineation of the feature. The procedure to utilize an existing GIS layer to facilitate the implementation of a self-organizing map has been developed. This process uses the GIS layer to define initial weights in the network which enable the neural network to identify the road network. [16]

Tsay(2000) from National Cheng Kung University proposes a set of new symmetric gradient operators for edge extraction. Automatic edge detection is illustrated in Chinese character recognition.[17] The results for centerline extraction from traditional edge detection algorithms is seen to be similar but with very fine edges the new approach using wavelets results in better representation.

Cumani(1991) in his paper relating to edge extraction for multi-spectral images utilizes second order differential operators to derive external edges at an sub pixel resolution which could be useful in image reconstruction procedures.[18]

Trier (1996) in his paper dealing with feature extraction methods highlights the variability of the methodology used with respect to the dataset under study. The experiments performed with different algorithms compares the results for selecting the most suitable method for different datasets.[19]

Kaski(1997) in his dissertation titled " Data Exploration using Self-organizing maps" again highlights that every feature extraction process is tailored to the application and the data attributes. He also deals with visualization issues and how self-organizing maps can be useful to cluster data elements by finding hidden relationships. [20]

Campbell et. Al (1996) in their work with segmentation using SOM's utilized the color and texture of the images to get better results than by using only the color information. This establishes a case for data fusion to improve classification results. Issues with similarities in the spectral signatures due to scene conditions were also identified. The paper proposes using segmentation before classification to reduce the machine time. [21]

Brown et.al. (2001) in the work with image processing using SOM describes the different applications. Interesting finding with the application of SOM for segmentation while preserving edges while processing natural sceneries and in feature location for identifying eye movements are discussed. This paper gives an encouraging viewpoint for the utility of SOMs.[22]

Research undertaken by Hui et.al(2000) at Centre for Remote Imaging, Sensing and Processing, National University of Singapore dealt with extraction and utilization of geometry and contextual information from IKONOS imagery using color and active contours.[13] They are presently working on semi-automated methods to improve the discontinuities in the road network extracted and to address issues of vegetation overlap on the road segments which is a common phenomena in Singapore.

CHAPTER 3
Feature Extraction

3.1 Introduction

In image processing, a feature represents an element in the image, which is a manifestation of an object from the real world. Feature extraction is made possible by identifying the element based on its characteristic attributes through image interpretation. Features from remotely sensed images are extracted by considering the shape, size, pattern, tone, texture, location and association of the elements in the scene.

Road inventorying systems include features, which describe the roadway, pavement width and roadside furniture among others. "Feature databases built by utilizing extraction techniques store the identity, location and characteristics of the natural and man-made features in the imagery." [8] These databases facilitate the creation of maps and GIS databases for planning, modeling and visualization.

3.2 Background

The performance of image processing systems is dictated by the image attributes, processing algorithms and the representation of data. The goal of image processing system is to convert the raw data to usable information by feature extraction. Feature extraction is divided into low-level, mid-level and high level processes.

Low-level processes are undertaken on the discrete image domain wherein the individual attributes of the pixel are utilized for extraction. E.g. the range of radiometric values for a feature when used for extraction constitutes a low-level process.

Mid-level processes aim to extract the feature based on the geometrical characteristic of the feature. The shape of the feature in the image constitutes a mid-level feature extraction. E.g. the elongated shape of the road features when used for extraction of the linear features from the image.

High-level processes consider the interrelationships between the features spatially and by association. E.g. the presence of vegetation along the road edges as a linear features improves the knowledge that a man made feature can be nearby and can be a road segment.

These levels of feature extraction are dependent on the type of feature to be extracted. Point features in images are characterized by their location on the image and their radiometric values across different spectral bands and hence a low-level feature extraction process would be suited. Line features in the images are elongated objects, which differ from neighborhood pixels in terms of their radiometric values; these include edges, which are linear features, corresponding to a significant change in the radiometric value. Mid-level extraction procedures are suited for their extraction. Polygonal features, which are regions of homogeneous intensity, representing a single feature class, can also be extracted by mid-level extraction procedures.

High-level feature extraction procedures are implemented when the feature characteristics make it difficult to be extracted by the low-level and mid-level processes due to the variation in terms of the spectral responses and overlap in the spectral values with other feature classes. Hence, knowledge regarding the environment is necessary for successful feature extraction.

3.3 Data Availability

The choice of feature extraction technique is dependent on the application area, the sensor technology available and the cost of the datasets. The start of satellite based data acquisition spawned numerous application areas mainly in resource management and in military intelligence gathering. The resource management included the ability to estimate/model natural and human interactions for developing decision making on the whole. Specific examples are in natural disaster assessments (flood plain mapping, earthquake vulnerability assessment), urban sprawl (detecting landuse change, mapping the transportation network) and in agricultural produce estimation (crop yield, forest cover). In military intelligence gathering, remote sensing from the satellite platform, allowed the experts to look far inside the enemy territory for determining the locations and extent of the enemy military structures.

The main factors in the usability of a remote sensed image after the acquisition cost is the spatial resolution, spectral resolution, temporal resolution expressed in terms of revisit period and the data collection and processing time as shown in Figure 2.

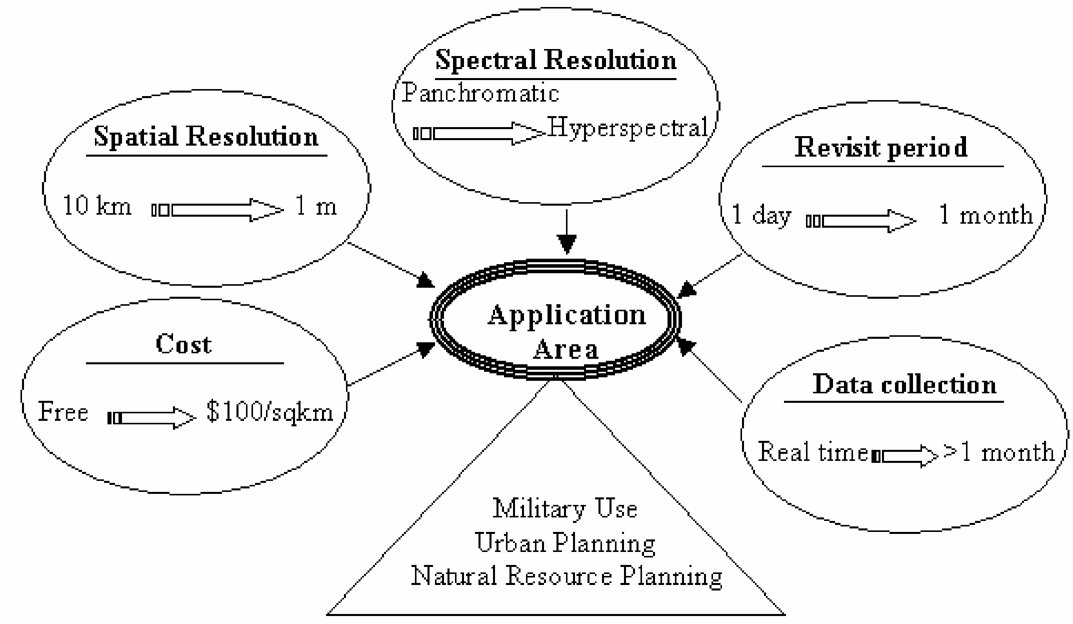

Figure 2: The factors determining the utility of Satellite Imagery.

The synoptic observation feasible due to the application of remote sensing for an elevated platform is useful in visualizing and modeling natural phenomena. Even minute changes to the environment can be detected and imaged from the satellite platform generating immense possibilities to researchers.

IKONOS satellite launched in 1999 was the first commercially available 1-meter resolution panchromatic imagery commercially available from 1999. The multi-spectral imagery from IKONOS has a spatial resolution of 4meters with Blue, Green, Red and Near Infra-red bands. [23] A detailed description of IKONOS data is given in the subsequent chapter under data description.

The availability of these high resolution datasets has opened up various avenues for research and development of application areas with unique needs for feature extraction. Prior to the availability of high spectral/spatial resolution data classification algorithms were concentrated on delineating large areas with significantly dissimilar attributes (spectral signatures), which could be detected in the available imagery. The algorithms were based on clearly demarcated boundaries between the landuse classes. Hence, the datasets were not overly influenced by minor changes in the spectral response (due to the coarse resolution) of the scene under study. But with increased sensitivity which provided an opportunity to detect finer elements (road segments, single buildings), it also degraded the discriminating capabilities of the algorithms which were essentially based on marked differences in their spectral signatures. This gave rise to an unique problem of dealing with high-resolution data as in one hand they came with the promise of better feature identification/delineation but on the other they complicated the identification process.

An example to explain this problem would be to identify the issues with road extraction from Indian Remote Sensing satellite (IRS 1C) with a horizontal resolution of 5.8 meters with the 1 meter horizontal resolution data from IKONOS 1. Road extraction from IRS 1C would be possible only along multi-lane facilities as the minimum spatial resolution limits it from detecting single lane segments. But with IKONOS 1 data, there is no problem with detecting single lane segments. The major issue would be the vehicles on the road during satellite imaging. Due to the increased spatial resolution, the pixels on the road which are occupied by vehicles do not get classified as roads. This problem is aggravated when thinning algorithms are used to extract the road network as these “holes” in the network bring in discontinuities.

Hence, the need for intelligent algorithms to disregard these “holes” is established.

Therefore, advancement in the sensor technology providing better data, matching advancements in feature extraction procedures are necessary to suitably utilize the datasets.

3.4 Feature Extraction Procedures

Feature extraction is made possible by using image-processing techniques. These techniques can be categorized into broad types as discussed below based on the sensor model, atmospheric model and surface interaction model. [2]

As described in the earlier section, feature extraction is dependent on the data characteristics (spectral/spatial resolution, cost) and application area.

3.4.1 Image rectification and restoration

Image rectification and restoration deal with removing noise and inherent distortion due to the sensor characteristics. Noise in the imagery can be defined as locally incorrect pixel values generated due to malfunctioning sensors. Noise pixels are identifiable by examining the image histogram where they would show up as kinks.

The geometric distortions are due to earth curvature, sensor type, atmospheric refraction and relief displacement. A satellite image of the earth is a flat representation of a curved surface and hence the acquired imagery is not an exact representation. Also, the presence of atmosphere in between the earth’s surface and the satellite platform; electromagnetic radiations (reflected light) get distorted. The affect of earth’s curvature is removed by image transformation using a set of ground control points

(points with known map co-ordinates). In the transformation process the co-ordinates of the ground control points (GCP) are measured from the imagery. The transformation functions are derived by least square regression.

$X_i = f_a(X_m, Y_m) \quad Y_i = f_b(X_m, Y_m)$

Where,

(X_i, Y_i) are image co-ordinates, (X_m, Y_m) are map co-ordinates and (f_a, f_b) are transformation functions obtained from regression analysis. [2]

The transformation function therefore re-samples the image pixels to achieve a better representation of the earth's surface. Re-sampling techniques include nearest neighborhood technique which preserves the original DN of the pixels and the cubic convolution method which calculates the pixel values in the corrected image by calculating the average of the surrounding pixels to give a smooth image rather than the disjointed image from nearest neighborhood re-sampling. In this research, nearest neighborhood technique was used so that the spectral information could be preserved.

Noise removal is the next step and is necessary before applying radiometric correction. As the presence of noise pixels would affect any algorithm based on the image histogram (univariate statistics would be sensitive), these kinks have to be removed prior to implementation of any radiometric correction scheme. Fourier transformation, thresholding techniques are used to identify the noise pixels and then a synthetic value is substituted computed from the average of the neighboring pixels.

Radiometric correction is then applied to counter the effect of illumination and the atmosphere. As shown in the Figure 3 the clouds in the north western corner of the acquired imagery cover some of the regions.

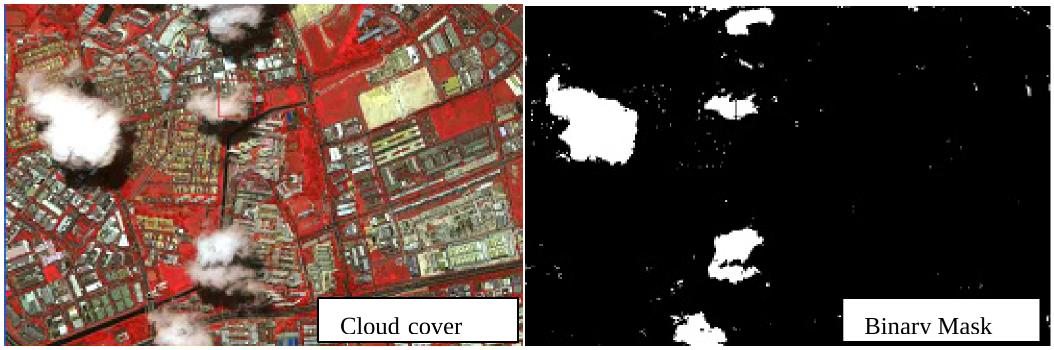

Figure 3: Binary mask used to remove the areas covered with clouds

The areas below the clouds are not visible as the clouds are very dense. Hence, a binary mask was used to remove the areas with cloud cover. If haze/smoke would have been present across the imagery then their removal is possible by applying a threshold or by using proprietary algorithms for the satellite dataset acquired. The threshold value is obtained by measuring the reflectance values over areas of nearly zero reflectance (deep clear water has no reflectance in the NIR band). The satellite data acquired for this research had been pre-processed for any presence of haze/smoke by the satellite data vendor.

3.4.2 Image enhancement

Enhancement is undertaken to improve the display and interpretation of the image. Enhancement thus , increases the visual interpretability for better selection of algorithms for analysis.

The first step in image enhancement is to compute the univariate and multivariate image statistics for understanding the distribution of DN values and the relationship between the different spectral bands. Univariate statistics are measures of central tendency (mean, variance, standard deviation) and do not give any information about relationships between pixels in different bands. Multivariate statistics on the other hand give insight into data quality and redundancy (covariance, band correlation) and are useful in classification and feature extraction.

Stretching of the DN values is done by applying a linear transform or based on the histogram, or by using the parametric distribution like the Gaussian stretch. The main aim of stretching algorithms is to use the entire range of the display for better interpretation due to increase in the contrast in the scene.

Spectral rationing is a technique used to eliminate the effect of non-homogeneous scene illumination, which causes spectral response variation within the bands. DN values from one spectral band are divided with the DN values in another band, pixel by pixel. The imagery used in this research did not require spectral rationing as the area imaged was not very large and hence the scene illumination was not a problem. Spectral rationing is very useful while dealing with large mosaics as the illumination factors change during data collection of large areas.

Normalized Difference Vegetation Index (NDVI) is a non-linear transformation of the visible (red) and near-infrared bands of satellite imagery [2]. NDVI is defined as

$$NDVI = \left[\frac{(NIR - RED)}{(NIR + RED)}\right]$$

NDVI can be used to identify areas covered with vegetation based on the general spectral signature of vegetation and is useful in image segmentation (preliminary classification).

Principal and canonical components in Multi-spectral (MSS) datasets can be derived to minimize the redundancy due to inter-band correlation. This technique reduces the dimensions and size of the data leading to faster processing but there is an inherent loss of information. A compromise between the processing time and the level of information is necessary for efficient image processing.

Intensity-Hue-Saturation (IHS) transformation is undertaken for utilizing the high spatial resolution of Panchromatic (PAN) imagery with the high spectral resolution of MSS image. Intensity (total brightness of a color), Hue (average wavelength of color) and Saturation (purity of the color) is another way to represent an image.

$$I = \frac{(R+G+B)}{3}$$

$$H = Cos^{-1}\left\{ \frac{\frac{1}{2}[(R-G)+(R-B)]}{\left[(R-G)^2+(R-B)(G-B)\right]^{\sqrt{2}}} \right\}$$

$$S = 1 - \frac{3}{(R+G+B)}\min(R,G,B)$$

where,

I = Intensity, H = Hue, S = Saturation, R = Red, G = Green and B = Blue components of the image. [24]

This is an improved strategy for visual enhancement wherein the image is separated into intensity, hue and saturation components by using the IHS transformation. IHS transformation converts the image into components, which are perceived by the analyst (human) and hence while analyzing lower resolution MSS data, the level of detail (intensity) can be separately analyzed and enhanced for improving visual interpretation.

The three primary colors red, green and blue (RGB) are used in additive color scheme to display satellite images. False Color composites (FCC) are created by assigning Red color for NIR band, Green color for Red band and Blue color for the Green band. Implementation of image enhancement algorithms on RGB images has to be applied to each of the color components and hence specific issues like improving the level of detail cannot be addressed singularly.

IHS transformation is thus useful in multi-scale image fusion wherein the intensity from a higher resolution image data (PAN image) can be used with the lower resolution image data (MSS data with high values in Hue and saturation due to the increased spectral resolution). The resulting image can be transformed to a RGB image for interpretation from the FCC format.

Edge enhancement is also performed for improving visual interpretation. Augmenting the high spatial frequency elements in the image improves the visual interpretation process which now has a better stimulus along changes in spectral classes. Examples of high and low spatial frequency regions in the satellite data are shown in Figure 4.

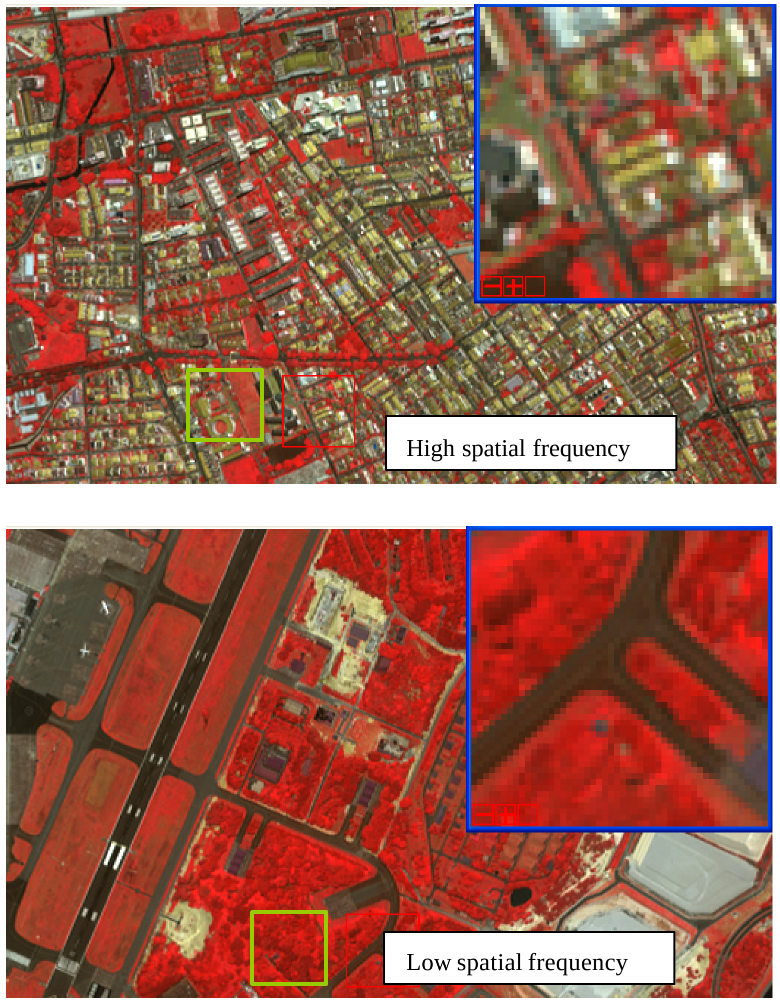

Figure 4: Regions with high and low spatial frequency in the satellite image

An edge in the context of an image can be defined as a contour of the highest gradient in the change of the radiometric values. Edge enhancement/sharpening increases the contrast between regions of different grayscale values. An edge enhancement algorithm, derives the high spatial frequency information from the image and combines with the original image, making the edges better defined.

Computation of the high frequency spatial information is undertaken by applying mid-level feature extraction procedures. These include laplacian, Sobel's and

Canny operators among various other algorithms.

Laplacian derivative is the second order derivative of the brightness (B) in the directions x and y

$$\nabla^2 B \equiv \frac{\partial^2 B}{\partial x^2} + \frac{\partial^2 B}{\partial y^2}$$

which is invariant to rotation. [25]

Sobel's operator measures the spatial gradient in the image [26]. The edges in the vertical (V) and horizontal (H) direction are computed by using roving windows with a set of weights as shown in Figure 5.

-1	0	+1
-2	0	+2
-1	0	+1

H

+1	+2	+1
0	0	0
-1	-2	-1

V

Figure 5: Sobel's convolution kernels

The magnitude and the direction of edge is computed by

$$Edge = \sqrt{H^2 + V^2}$$

$$Direction = \arctan\{V/H\}$$

One of the most widely used edge detection algorithm with successful application in noisy images is the Canny's edge detector, which is a multi-stage process wherein the image is first smoothened by a Gaussian convolution and then the first order derivatives are computed [27]. Thinning is then done to eliminate multiple edges from the same feature creating a thinned edge map. Tracking of the edge pixels is done to join the edges and suppress the responses from noise. Matlab implementation of Canny's requires the threshold for edge detection. This threshold value can be automatically computed by the algorithm by finding the natural breaks in the image histogram. In summary, Canny's edge detection algorithm aims to maximize the efficiency of edge detection by initially detection of only the most important edges followed by localization criteria which selects the located edges which are closest to the actual edge. The final step is to minimize multiple responses from edge features.

This algorithm encounters problems with junctions and corners in the image scene as illustrated in Figure 16 in chapter 4.

3.4.3 Image classification

This process aims to group the different feature classes in the image by spectral and spatial pattern recognition. Temporal pattern recognition is also a part of image classification used in datasets acquired at different epochs. Image datasets contain representations of multiple features in terms of DN values across the different spectral bands. Supervised and Unsupervised classification techniques are the strategies employed for classification of the DN values associated with different features in the imagery.

Supervised Classification is based on the postulate that similar features have similar spectral/spatial characteristics and hence by identifying representative regions for each spectral class, the image can be classified based on the statistics derived from the representative regions.

The analyst determines the separability of feature classes by defining training sets representing each category. Algorithms utilize the statistical descriptors calculated from the training set for classification. In the training stage in supervised classification the representative regions in the image are identified and the statistics of the sample are computed. Ground truth data is used to correctly characterize the representative regions. These regions must include all the intra-class variation of the radiometric values associated with the feature and need to have enough pixels for statistically significant analysis in the classification stage.

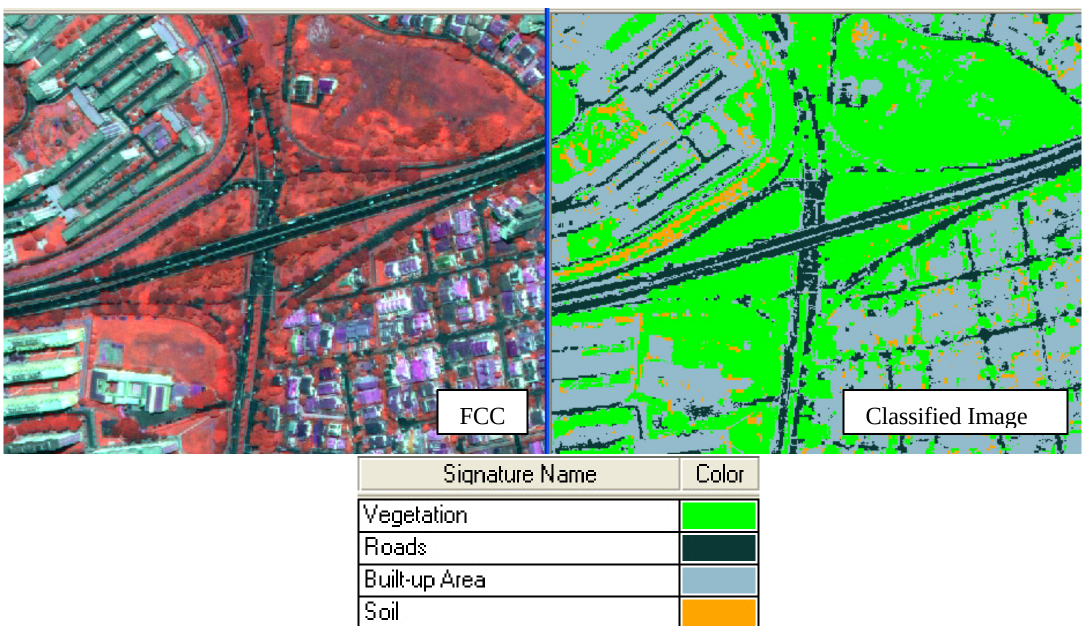

Figure 6: Supervised classification of the merged image into four classes

The nature of the information classes which are sought and the spectral and spatial complexity of the area under analysis dictates the number of training sites to represent the variability in each class. Lillesand and Keifer comment that even 100 training sites per information class can be used to assimilate majority of the intra-class variability. [2, pp 545-555] The theoretical lower limit of the number of pixels necessary in a training set is (n + 1), where n is the number of spectral bands in the dataset. Lillesand and Keifer again state that a minimum of 10n to 100n pixels are suited for describing each information class. Figure 6 shows the results from supervised classification performed on the IKONOS imagery. Sections of the road segment are missing due to the overhang of the tree canopies as shown in Figure 7 below.

Figure 7: Low oblique photograph of Singapore Central Business District. (Superstock.Inc)
Inset: Outgrowth of trees over the freeways

Minimum distance classifier computes Euclidean distance between each of the pixel vectors and each of the feature vectors, which are computed from the training sets defined by the analyst. The pixel is assigned to the feature class with which the computed Euclidean distance is minimized in the multi-feature space. This type of classification is computationally economical and is a useful tool for preliminary investigation as the accuracy of classification decreases when the features are not well separated.

In *Parallelepiped classifiers*, the DN values of the unknown pixel are compared to upper and lower bounds of the spectral signature of each feature as derived from the training sets and then the unknown pixel is assigned to the feature class within which it lies. Overlapping of the range of the DN values of the feature classes causes the unknown pixel to be assigned to more than one class leading to classification errors.

The *Maximum Likelihood* (MXL) classifier quantitatively evaluates the variance and covariance of the class when classifying an unknown pixel. The statistical probability of a given pixel belonging to a particular class is computed based on the mean vector and the covariance matrix generated from the training data. [2] The algorithm assumes the features class to have a normal distribution of DN values while computing the class statistics.

The next higher level of classification is undertaken by *expert systems*, which are computer based inference engines operating on a set of rules derived from domain knowledge.

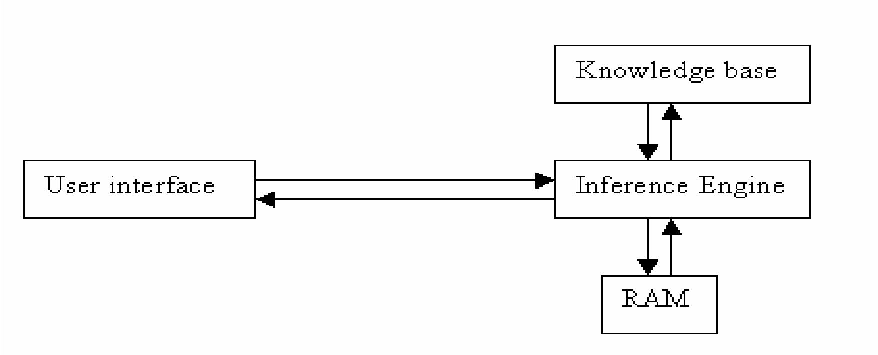

Figure 8: Schematic diagram of an Expert System

Expert systems attempt to emulate the human decision making process by applying sound inference / reasoning techniques on collections of domain knowledge expressed in the form of *if – then* rules implementing chains of logical deduction by human experts.[29] Figure 8 shows the typical architecture of an expert system.

Unsupervised classification does not use *a priori* knowledge associated with the dataset for categorization. Instead, the output classes from the implementation of the algorithm give an insight into the separable clusters in the image dataset. These clusters do not represent any feature class defined by the spectral response patterns but are used as a guide for applying the domain knowledge during characterization.

K-means and IsoDATA (Iterative Self-Organizing Data Analysis Technique algorithm) are widely used unsupervised algorithms in remote sensing based on a simple iterative procedure which clusters data into a predefined number of clusters (K) of the data. The selection of the centriods of the dataset is an iterative process, with new centriods being calculated for the clusters until the clusters (and their centroids) stabilize. This method of unsupervised classification is important to understand the natural clustering in the dataset. The shortcomings of this approach are the sensitivity of the unsupervised classification algorithm to initial cluster starting points and the identification of a stopping rule; these lead to difficulties in reproducing the same classification results between different attempts leading to inconsistencies in classification. [9] These limitations can be eliminated by using artificial intelligence techniques.

Artificial Intelligence is the ability of machines to mimic human decision-making in varied scenarios. Artificial intelligence can be achieved by using Neural Networks which are simple computational units connected through a dense network of connections.

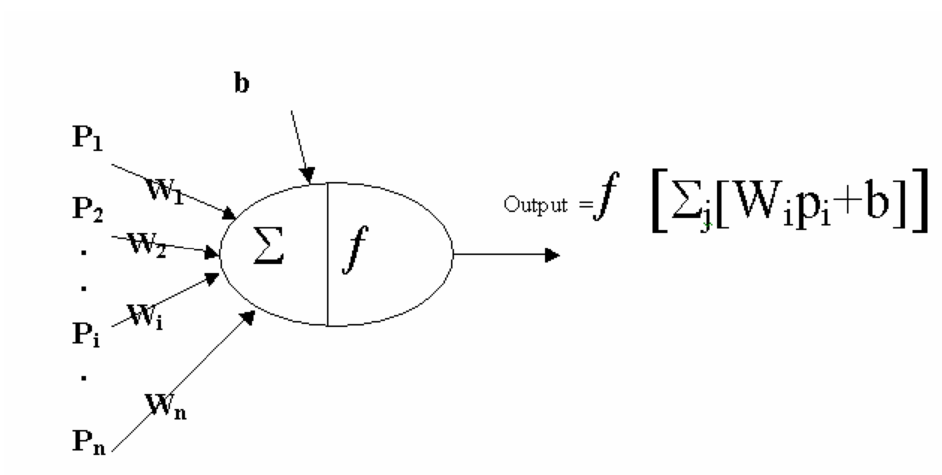

Figure 9: Model of an artificial neuron

Figure 9 shows a typical model of a neuron which is performs the summation of the input vectors(P_i) with their initial weights (W_i) and the bias (b) to give an output which is defined by the output function (f).

The first stage in developing a neural network is its architecture, which is defined by the connection between the input and the output. Thereafter, the network is initialized by defining initial weight values for the weightages. Learning or the training stage is next step when the weights of the connections in the neural network are updated based on a given set of rules by using a subset of the dataset. After the weight vectors have stabilized, vectors unseen by the network can be input to the network for analysis.

Neural networks need a large number of training sets and time to stabilize the decision rules but once the networks are trained the processing time is minimal when compared to other non-parametric classifiers. Table 2 shows the different applications of neural networks in various fields on image processing.

"Kohonen's self organizing maps (SOM) are simple analogs of human brain's way of organizing information in a logical manner". [9] SOM is a unsupervised-competitive learning process for forming a topology-preserving map. Topology preservation/mapping preserves the relative distance between the points(input vectors). Points that were close to each other in the input space are mapped to nearby map units in the resulting SOM. This paradigm is useful to utilize the contextual information(E.g. vegetation along road networks) in the road environment for classification. The high dimensional input space is mapped into lower dimensional mapping space(usually a 2-D lattice) for better data visualization/interpretation. Hence while dealing with high dimensional datasets, SOM can be used to determine clusters without having any a-priori knowledge of the input dataset. The generalization capability of SOM can then be used to classify unknown input data.

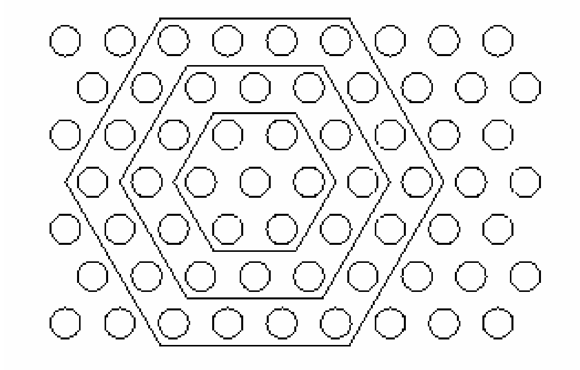

Figure 10: Illustration of Hexagonal topology of SOM

Figure 10 shows a hexagonal topology which is commonly used in the mapping space. This topology is useful in topology preservation as each neuron has six immediate neighbors and hence relationships between input vectors can be better presented.

The architecture of a typical Self-organizing feature map is shown in Figure 11.

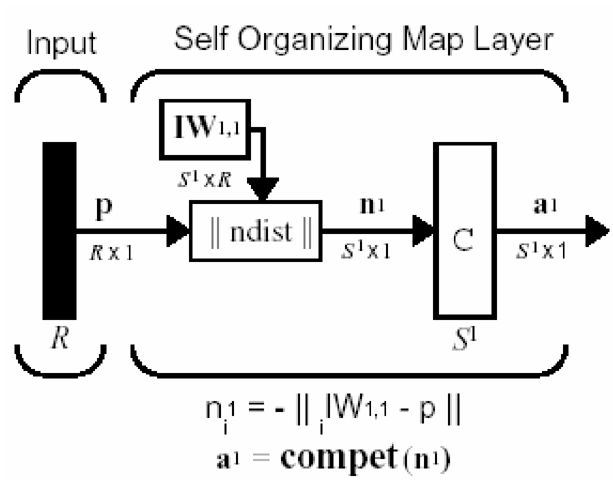

Figure 11: Architecture of a Self-organizing Map (Matlab Neural network toolbox)

An implementation of SOM is illustrated in Figure 12, which shows the randomly initialized input vector, and the codebook vector in the mapping space after the network is trained.

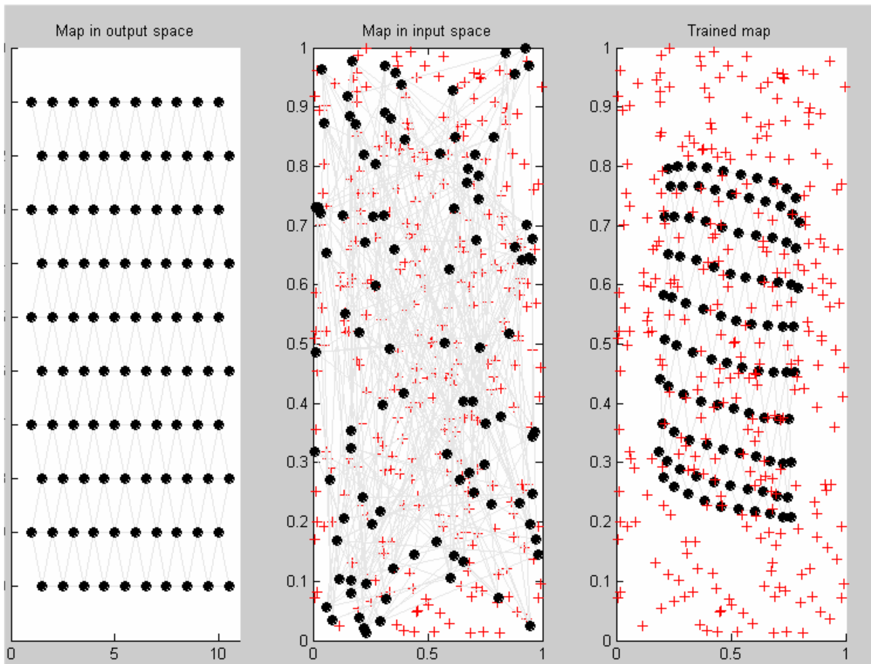

Figure 12: An example of SOM implementation (SOM toolbox)

The input vector "**p**" is a matrix whose rows represent input vectors and the columns represent the attributes associated with each input vector. The initial weights matrix IW contains random initial values, which are used with the results from the distance measurement of each input vector to find the best matching units (BMU), which are computed by determining the Euclidean distance between the input vector "p", and the codebook vectors in the mapping space.

Table 2 shows the utility of various neural network paradigms in the realm of image processing. As it can be seen SOM's are well suited for a variety of tasks in image processing and analysis.

3.4.4 Biophysical modeling (ground truth)

The dataset obtained by remote sensing is correlated to the phenomena measured on the ground. Ground truth data is necessary to quantitatively associate the datasets. This step allows the analyst to utilize the data in the modeling process.

3.5 Automated feature Extraction (AFE)

Feature extraction for building databases is time and labor intensive; introducing computer automation will increase the productivity. Productivity improvement in feature extraction comes not from a single dramatic change but from incremental improvements in a number of areas. [8] These incremental improvements include increase in the computing power, newer sensor technology and new application areas.

The success of an AFE operation is dependent upon user interfaces, which can streamline user interaction and provide the collateral information used in every step and the logic behind the analysis. AFE methods can be integrated with the existing algorithms for possible improvements.

Combining the different classification methods and taking advantage of improved sensor capabilities, hybrid classification scheme can be implemented. Interactive digitizing of features cannot be productive due to the time taken for large datasets and would have lower quality as it is dependent on the skills levels and fatigue of the analyst.

Table 2: The different types of neural networks applied to the various tasks in the image processing chain. [10]

Image processing task / Neural network type	Image reconstruction	Image restoration	Image enhancement	Image compression	Feature extraction	Segmentation, pixel-based	Segmentation, feature-based	Object recognition, pixel-based	Object recognition, feature-based	Image understanding	Optimization
Feed-forward, regression	•	•	•	•	•		•				•
Feed-forward, auto-association				•	•			•			
Feed-forward, classification			•			•	•	•	•	•	
Feed-forward, shared weights								•			
Feed-forward, recursive							•	•			
Perceptron					•					•	
Radial basis function network (RBF)				•		•	•				
Self-organising feature map (SOM)				•	•	•	•	•	•	•	
(Fuzzy) Learning vector quantization (LVQ)				•			•				
Hopfield	•	•	•		•	•	•	•	•		•
Cellular (CNN)		•				•	•				
Generalized adaptive neural filters (GANF)		•									
Adaptive resonance theory (ART)			•					•			
Associative memories (and RAM)						•		•	•	•	
ADALINE	•										
Neocognitron								•			
Probabilistic						•					
Neural decision tree										•	
Neural belief network										•	
Higher order network								•			
Counterpropagation network							•				
Fuzzy neural / Neuro-fuzzy system		•			•	•		•	•		
Other		•	•	•	•	•	•	•	•		•

The consistency and time savings afforded by automated methods make it worthwhile to invest in powerful computer systems and allow rapid analysis of data. The next chapter describes a hybrid feature extraction procedure developed to extract road segments from urban and peri-urban scenes (E.g. the central-business-district of Singapore).

3.6 Accuracy Indices

The National Standard for Spatial Data Accuracy (NSSDA) implements a well-defined statistic and testing methodology for positional accuracy of maps and geospatial data derived from sources such as aerial photographs, satellite imagery, or maps.[11] Accuracy is reported in ground units.

3.6.1 Types of Accuracy estimates

The main elements of data accuracy are :

- Positional accuracy
- Attribute Accuracy
- Logical consistency
- Completeness
- Lineage

While evaluating the accuracy of an inventorying program the most important element is attribute accuracy which is followed by completeness and then positional accuracy. The positional accuracy of the inventory process is dependent on the baselayer which is used for the analysis. In this research, 1 m resolution PAN imagery from IKONOS was used and hence according to specifications of the product the horizontal accuracy is within 12 meters which is reduced to within 2 meters by registration with existing datasets [Table 3]. This accuracy is improved by ground survey which is used to rectify the imagery. Ortho-photos can be generated by investing further to obtain higher positional accuracies. However the major thrust of an inventorying process is the attribute accuracy. The results of classification have to be compared with ground truth to determine the accuracy. As the spatial resolution was 1 meter any object below this threshold cannot be detected, hence elements like utility poles, signals, road marking; elements of a road inventory cannot be identified. The possible alternatives are discussed in the section dealing with future recommendations wherein higher spatial resolution data can be obtained by using aerial photographs or by on-site procedures.

The accuracy indices are mainly concerned with positional accuracy. RMSE NSSDA statistic (1.7308*RMSE) which is another form of RMSE is mainly utilized to evaluate how closely the locations match with the actual co-ordinates [10].

In an inventorying procedure, the accuracy indices are mainly the percentage of positive identification for each type of facility. The total number of road pixels that would typically constitute a road segment is compared with the number of road pixels identified after the application of the algorithm.

Accuracy of the SOM network can be determined by using pure pixels as input to the network and locating the part of the network which houses the BMU. The location of the BMU with respect to the input vectors actual cluster would determine the accuracy of the network in classifying unknown input.

CHAPTER 4
Hybrid Feature Extraction

4.1 Introduction

Any algorithm composed of simpler algorithms can be labeled as a hybrid algorithm. In this research, a hybrid algorithm is built using a combination of established image processing techniques and artificial neural networks (specifically, an unsupervised neural network called a Self-Organizing Map). The hybrid algorithm is intended to be used for feature extraction in the complex scenes experienced in urban areas. It is postulated that the inherent limitations of traditional classifiers for clustering and feature extraction can be overcome using the ability of artificial neural networks to discern contextual patterns in the dataset, thus improving the overall clustering ability of the hybrid algorithm.

Road features in an urban area are highly dense, with a very close configuration with vegetation and buildings (similar spectral signature). A typical road intersection in an urban area would have numerous lane markings (for pedestrian and turn movements), traffic islands with landscaping and various billboards. The end result of such complex configuration leads to the pixels being imaged with a value which is a sum total of all the elements contributing to the reflectance pattern being captured by the satellite based sensor. Hence, spectral information alone would not be adequate to correctly classify the pixels.

Multi-spectral information along with edges derived from panchromatic image was used to characterize the road features for potential improvement in their detection.

The imagery was preprocessed to be suitable for input into the neural network package. Edge information along with intensity values from the PAN image were used to characterize each pixel representing the scene. These input vectors were used by the SOM network for clustering. The overview of the entire process is shown in Figure 13.

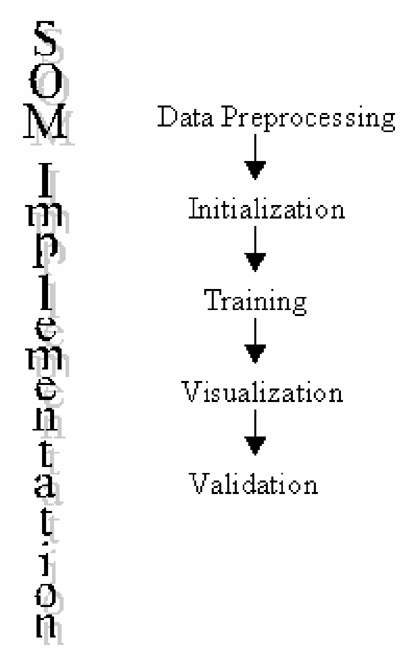

Figure 13: Schematic diagram of SOM based clustering

These clusters were then associated with information classes based on a set of pure pixels defining each class. It was hypothesized that the data would be clustered into roads, vegetation, urban features, mixed pixels (describing the road/vegetation /urban boundary) and unclassified. The cluster of mixed pixels would be of interest as these would be generally unclassified by traditional classifiers, but with the edge information and the ability of the SOM to preserve the topology, these boundary pixels would mostly likely be classified into one of the closest landuse classes in terms of its Euclidean distance from the landuse class vector. Moreover, the ability of neural networks to generalize these class attributes by compensating for different data collection conditions and different conformations of the road features would allow these boundary pixels to be successfully classified.

4.2 Data Description

IKONOS satellite was used to acquire the imagery over Singapore. The area included the Central Business district of Singapore. There is a wide diversity in the scenes in the photograph, with highly urbanized areas, water transportation facilities, peri-urban areas and highly vegetated areas. This dataset would allow for proper development and design of the algorithm to suit varying environments. The specification for the data is listed in Table 3.

Table 3: Satellite Data description [23]

Spatial Resolution	PAN = 1m MSS = 4m
Spectral bands (microns)	PAN = 0.45 to 0.9 MSS Blue = 0.45-0.52 Green = 0.52-0.60 Red= 0.63-0.69 NIR= 0.76-0.90
Revisit Cycle	3 days using multiple satellites
Sensor type	Push broom
Pixels	11 bit
Swath Width	11 km X 11 km
Cloud cover	Less than 10 %
Format	Multi-band Geo-tiff
Area of coverage	
Processing Level:	Standard Geometrically Corrected
Interpolator Method:	Bicubic
Map Projection:	Universal Transverse Mercator UTM Specific Parameters Hemisphere: N Zone Number: 48 Datum: WGS84
Bits per Pixel per Band:	11 bits per pixel
Sensor	IKONOS-2
Acquisition Date/Time	: 2000-05-15 03:00
Metric Accuracy	No ground control : 12-meter horizontal 10-meter vertical. With Ground control : 2-meter horizontal 3-meter vertical

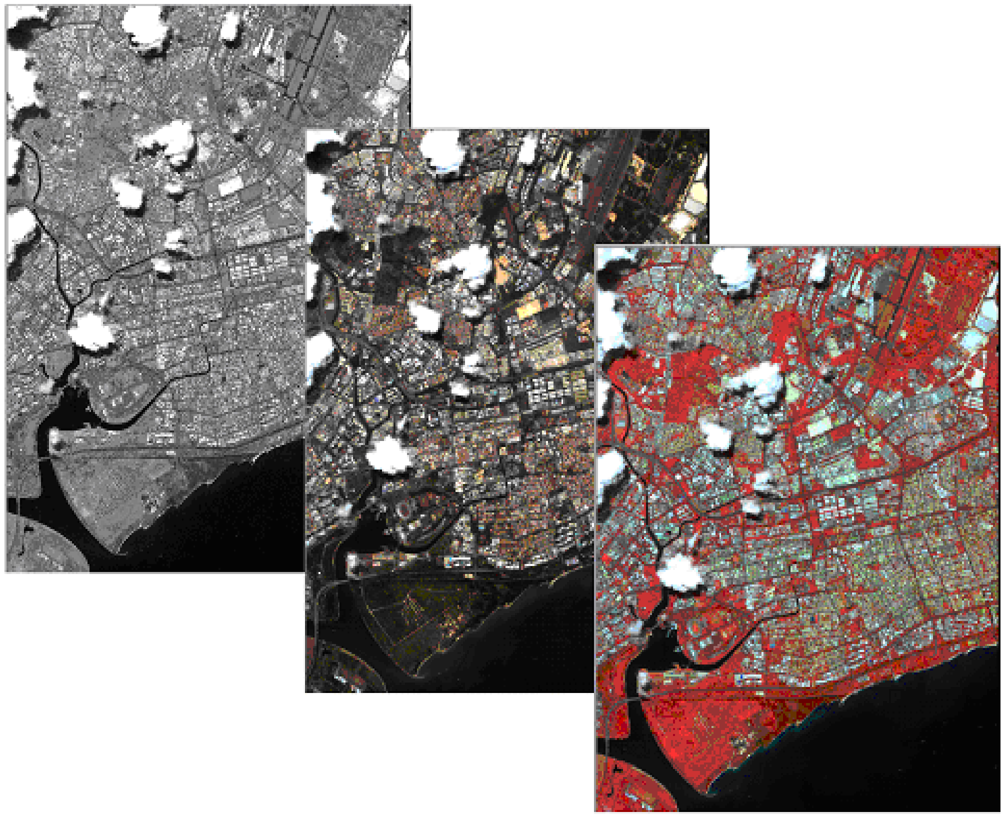

Figure 14: Satellite imagery (PAN, Natural Color, False Color)

The data was provided in the Geotiff format for each spectral band namely: blue, green, red and near-infra red. The panchromatic imagery and the output of band combinations are shown in Figure 14.

The file size of the PAN image was nearly 110 megabytes and each band of the MSS image was nearly 7 megabytes. The data provided had cloud cover of less than 10 percent as specified in the contract with Space Imaging regional center in Singapore.

4.3 Preprocessing

The available imagery was preprocessed in order to be used as an input into the neural network. Preprocessing is necessary to reduce the dimensionality of the input and extract the information necessary for developing the feature extraction methodology. The four bands (PAN, Green, Red, NIR) available as four different images were converted into matrices for input into the neural network. The PAN data with its 1 meter resolution was merged with MSS data of 4 meter resolution. NIR, Red, Green components of the MSS imagery were used in the merge procedure which resulted in an image with 1 meter spatial resolution. Blue band was not utilized as it is most affected by the atmosphere and hence would not aid in the clustering process. The edges were also extracted and were input into the neural network. Hence, the input vector "**p**" would consist of the spectral responses in the Green, Red, NIR band along with extracted edges. The steps involved in pre-processing are discussed in the following sub-sections.

4.3.1 Resolution merging

The first step was to merge the high resolution PAN imagery with the MSS imagery by using the IHS transformation as discussed in section 3.3.2. The results are shown in Figure 15. The merged image has higher intensity from the PAN image and hence visual interpretation is improved. The spatial and spectral resolutions compete with each other during sensor design. High spatial resolution means that large energy levels are required for completing the sensing task which is only possible with a large sensitivity window as in PAN images which senses the energy in the entire visual spectrum.

Figure 15: IHS transformation for merging 1 meter resolution PAN with 4 meter MSS imagery

However, MSS imagery has lower spatial resolution as with increased spectral resolution the sensitivity window is reduced for each band, making it difficult to absorb higher energy levels required for imaging at higher spatial resolution. Hence, the compromise with spatial resolution for improved spectral resolution is almost annulled by the merging procedure.

4.3.2 Edge detection

The next step was to derive the edges from the PAN imagery. Canny's edge detection method was used as it was necessary to maximize the edge detection along road/vegetation boundary while minimizing responses from other features in the image.

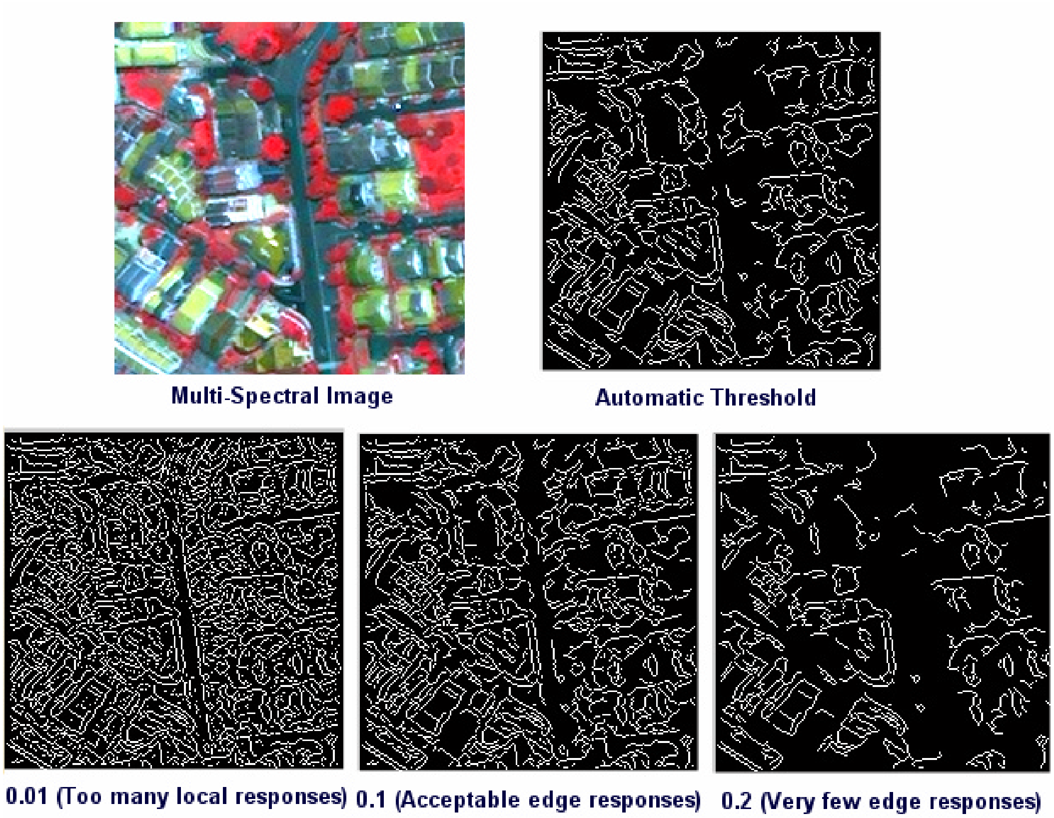

Figure 16: Implementation of Canny's edge detection algorithm.

Thresholding and thinning operations applied in the Canny's algorithm maximized the location of important edges while minimizing local responses. The resulting matrix after edge detection is shown in Figure 16. Threshold value of 0.1 was utilized for deriving the results after comparing the results as shown in Figure 16. As it can be seen, most of the edges in the image were detected by the application of Canny's algorithm with a threshold of 0.1.

4.3.3 Unsupervised classification

Unsupervised classification into ten classes was performed by using ISO-data procedure on the four meter resolution MSS imagery and the 1 meter resolution merged image. The results are shown in Figure 17. The main purpose of unsupervised classification was to identify natural clusters in the dataset so that the

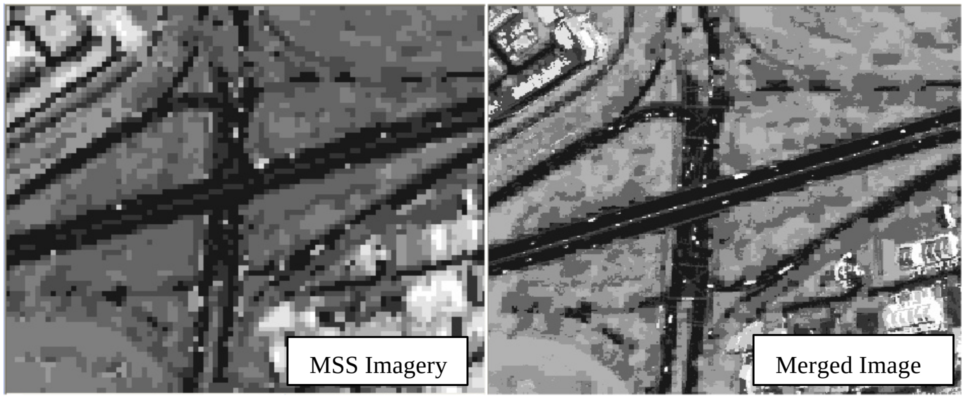

Figure 17: Results of unsupervised classification into 10 classes.

training sets

could be chosen training. The vehicles and the roads marking contribute to the noise in the merged dataset.

Hence, even though the enhanced resolution due to the intensity values from the PAN image in the merged dataset improved visual interpretation, spectral classification alone was not beneficial as shown in Figure 17. The classes thus generated were only useful in selecting the pure pixels which would be used to check the accuracy of the final clustering.

4.3.4 Data scaling

The attrributes of the input vector "**p**" as described earlier are then scaled to have unit variance. This assures that for each attribute, the difference between two samples(sample variance) have approximately an equal contribution to the summed distance measure between an input dataset and codebook vector in the mapping space.Unit variance of the input vector is calculated by dividing the attributes by the square roots of their corresponding variances.

4.4 SOM Implementation

SOM toolbox for Matlab, developed by Laboratory of Information and Computer Science in the Helsinki University of Technology was utilized to design the SOM. The first step in the implementation of the SOM network was to select training samples. The selected input vectors were then formatted to be suitable for the neural network algorithm which mainly uses "struct arrays" (arrays capable of storing different variable types) for analysis. Use of "struct arrays" makes simplified data analysis while increasing the time required for data formatting. The network was then initialized before the actual training procedure.

4.4.1 Random selection of training samples

Random samples numbering 100 were chosen for input into the neural network. The training sets were of 200 meters by 200 meters in size and hence being large enough to accommodate and model variations encountered in the vicinity of road pixels. Road alignments are typically 15 meters in width in urban scenarios and hence by choosing a square block with sides measuring at least 20 meters, landuse classes interacting with the road pixels can be assessed and modeled in the analysis. A larger area chosen would allow road intersections to be selected which would facilitate in training the network with all possible combinations of road configurations available in urban scenarios. The total number of pixels in the merged image (1 meter resolution) is approximately equal to 60 million (8845(rows)*6677(cols)) of which nearly 10 million were not suitable in the analysis as they represented the ocean.

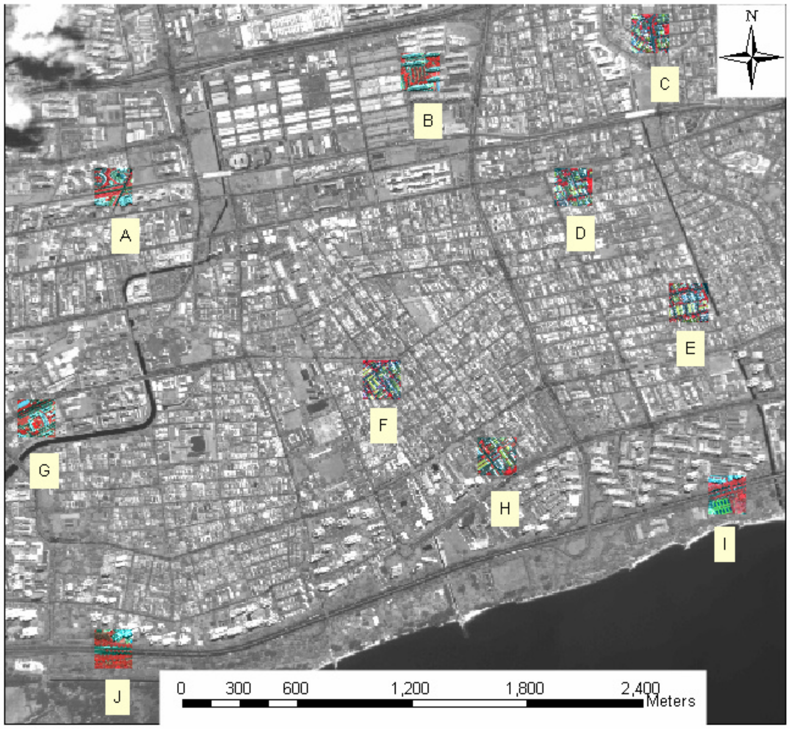

Figure 18: Location of training sets in the Southeast region in the satellite image.

Of the remaining 50 million pixels, 4 millions pixels were used in training and validating the results. The number of pixels used was sufficient considering that the recommended amount of training pixels is 10 times the number of weights in the network.[28] The training sets randomly selected in the southeast region of the image are shown in Figure 18.

4.4.2 Data formatting for SOM

The input to the neural network was in the form of a matrix wherein the rows represented unique inputs and the columns described the input vector. The DN numbers from the merged image and the edges derived by applying the Canny's edge detector were used to populate the input vector.

The data format:

Input Vector p = f (Near Infra-Red, Red, Green, Edges)

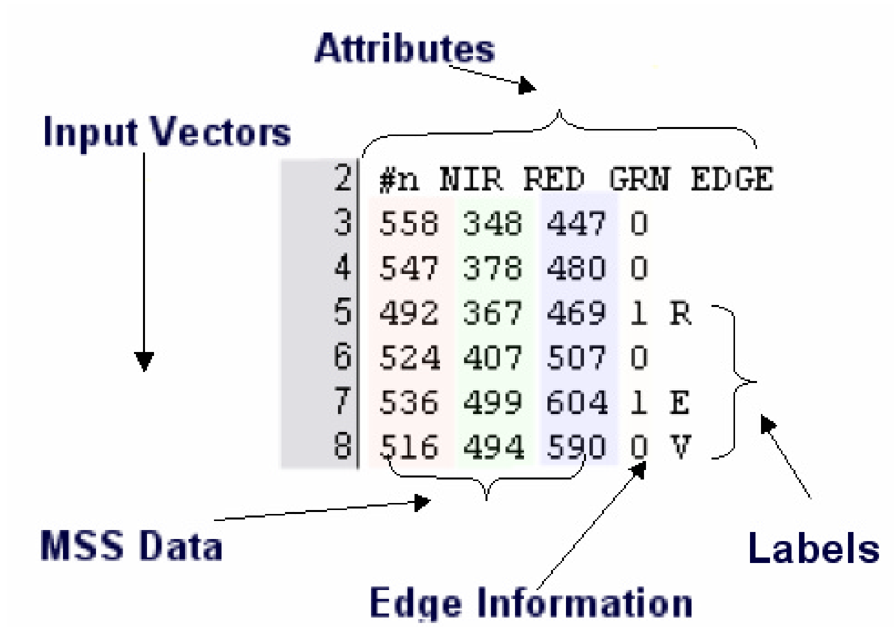

Figure 19: Formatted input for the training.

Figure 19 shows the input file used for training. Other dimensions that could be analyzed include the DN values from the blue band, the results of supervised classification and NDVI derived from imagery.

Data scaling as described in section 4.3.4 was performed to achieve an unit variance. The GUI tool called by function “preprocess” was used to perform preprocessing tasks like data normalization and histogram inspection. The screenshot of the tool is shown in Figure 20.

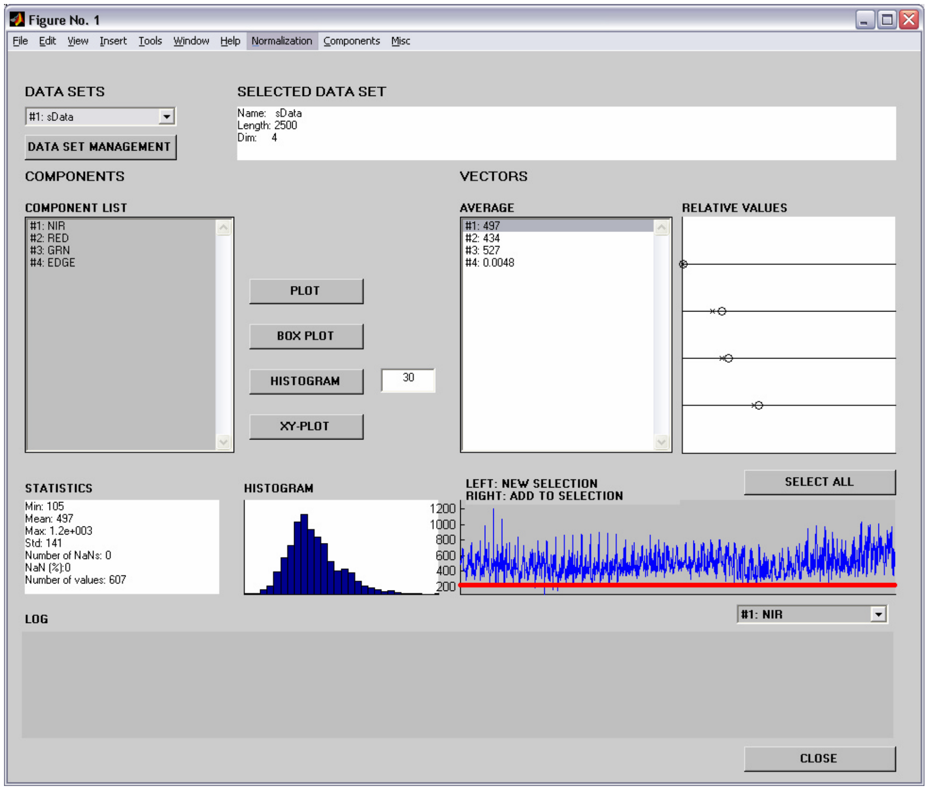

Figure 20: Preprocessing toolbox used to format input vectors.

The tools allowed the user to store the image matrices as "struct" elements for use with the toolbox.

4.4.3 Initialization

Linear initialization of network weights was performed so that the weight vectors are initialized in a systematic manner along the linear subspace spanned by the two principal eigenvectors of the input dataset. A hexagonal grid topology was used to characterize the mapping space. A hexagonal topology would allow for maximizing the influence of each neuron on the neighborhood. The size of the mapping space was calculated to be composed of 45 rows and 22 columns. The map size calculated is a function of the dimensions of the input data (number of input

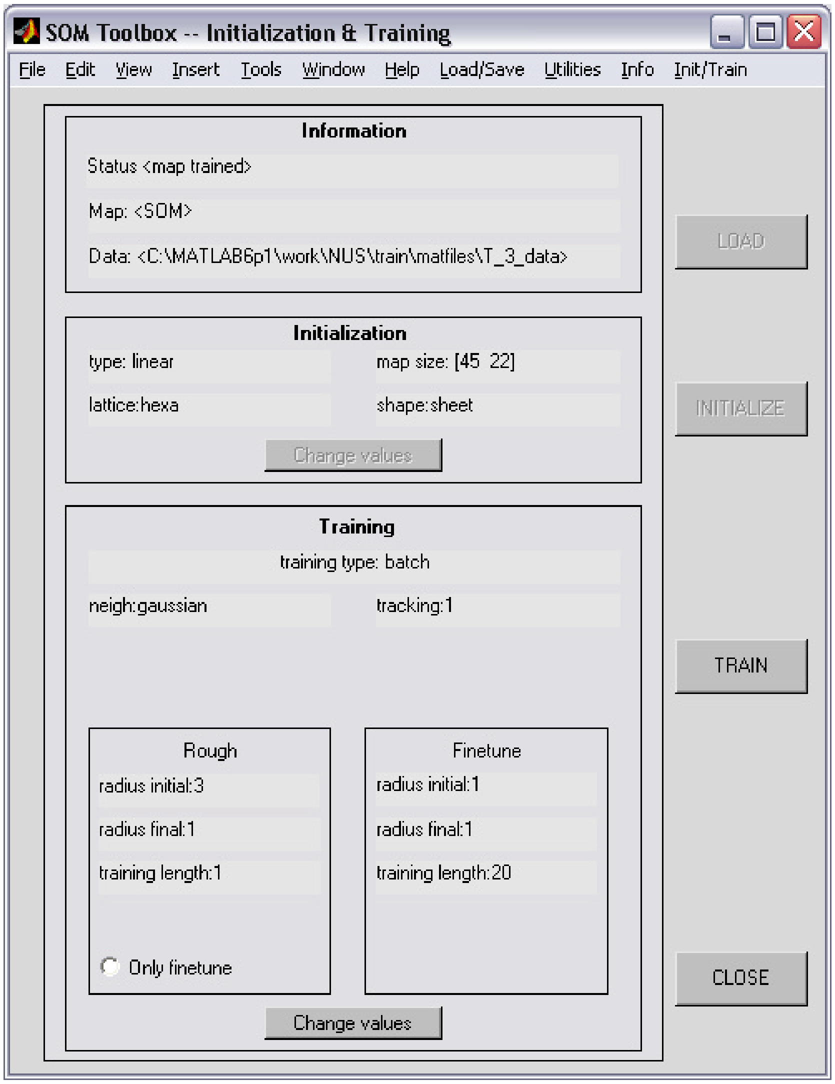

Figure 21: GUI for initialization and training of SOM.

vectors and dimension of each input vector) and the two biggest Eigen values of the training data.[30, pp33-34] Figure 21 shows the GUI tool for initializing and training the network.

4.4.4 Training

Training was initiated with the hypothesis that the contextual information between input vectors would facilitate the network in the clustering process. Hence, the set of input vectors presented, should characterize the pixels in any single analysis scene (200 meters by 200 meters) as selected while determining the size of training sets. Batch mode of training was therefore chosen for training the network, so that the network would perform clustering, one scene at a time.

The batch training algorithm processes the training sets one by one and in every training step, each input vector is assigned to the map unit which it closest to it. The weight vectors are then recalculated as:

$$W_i(t+1) = \frac{\sum_{j=1}^{n} h_{ic}(t) p_j}{\sum_{j=1}^{n} h_{ic}(t)},$$

where

$W_i(t+1)$ is the weight vector for input vector p_j at time t, $h_{ic}(t)$ is the neighborhood kernel around the winner $c = \arg\min_n \{\| p_j - W_n \|\}$, n is the number of input vectors.

Gaussian neighborhood function was used during the training process for determining the zone of influence of the winning neuron in the mapping space.

4.4.5 Visualization

Visualization of the trained network by representing the mapping space as a grid and then in 3-Dimension to display the configuration of the network. Figure 22 shows the Unified distance matrix (U-matrix) which is a visualization of the neighborhood distances between the map units. High values (darker pixels) of U matrix indicate a cluster border and uniform areas represent the clusters formed. The clustering in the component (attribute) planes is also shown in Figure 22.

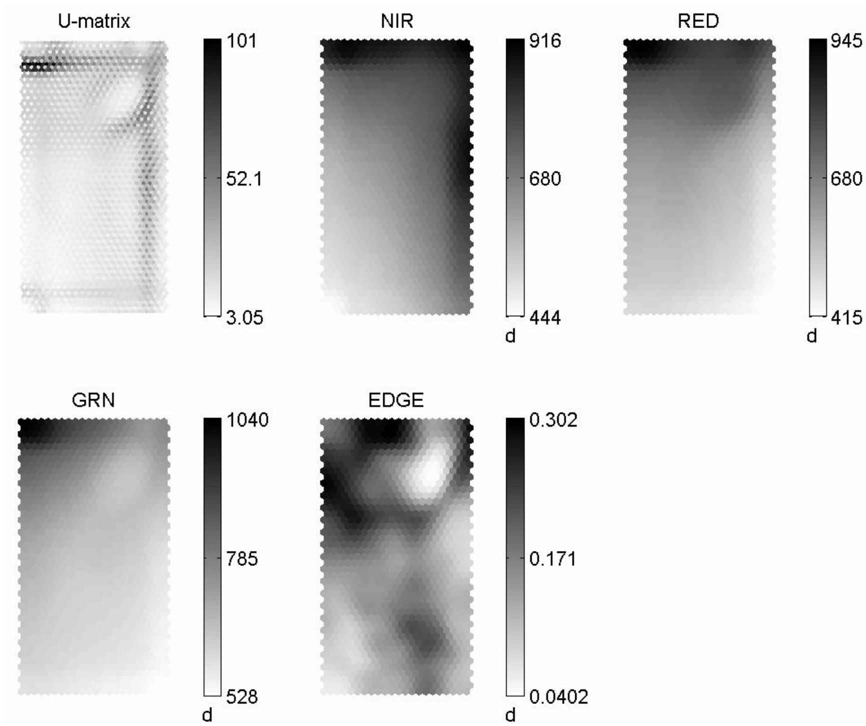

Figure 22: The clusters structure after training

The topology formed in the results is representative of the actual ground situation where the road pixels and the mixed pixels are adjacent to each other. Hence, the neurons adjacent to neurons mapping the pure road pixels are located close to those mapping the edges.

The location of the neurons in the mapping space is shown in Figure 23. The connections are directly proportional to the distance between the vectors describing them. From the 3D rendition it is seen that the density of neuron in the central part is highest.

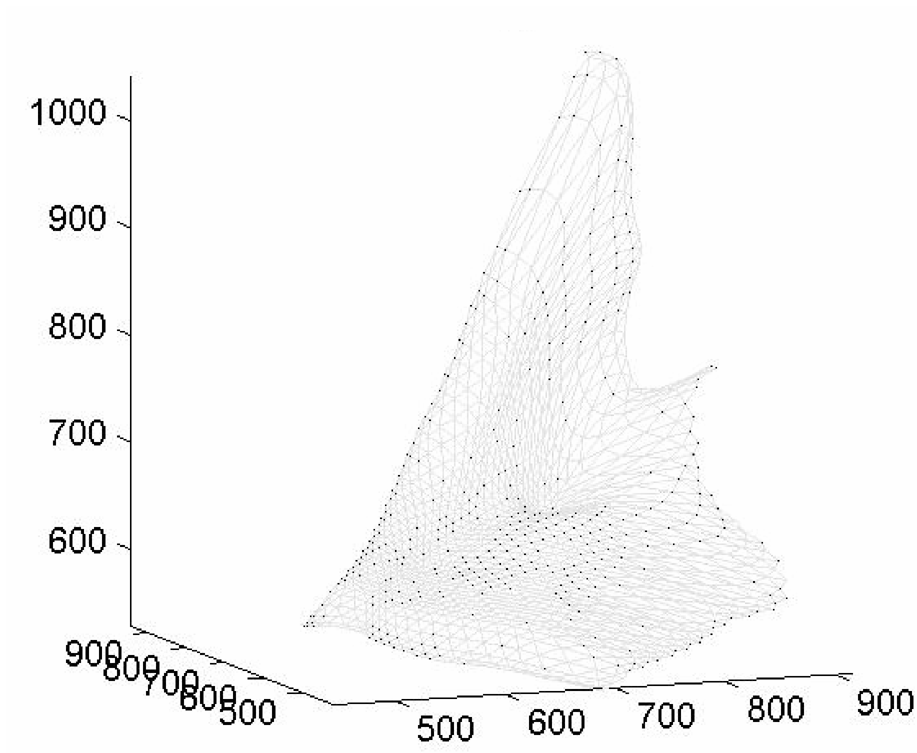

Figure 23: Location of the neurons in the final mapping space

4.4.6 Validation

The clustering capability of the network is assessed by using a set of vectors with known class attributes. This would show the regions in the map which are sensitive to specific landuse classes. Figure 24 shows how the different class vectors are mapped by the network.

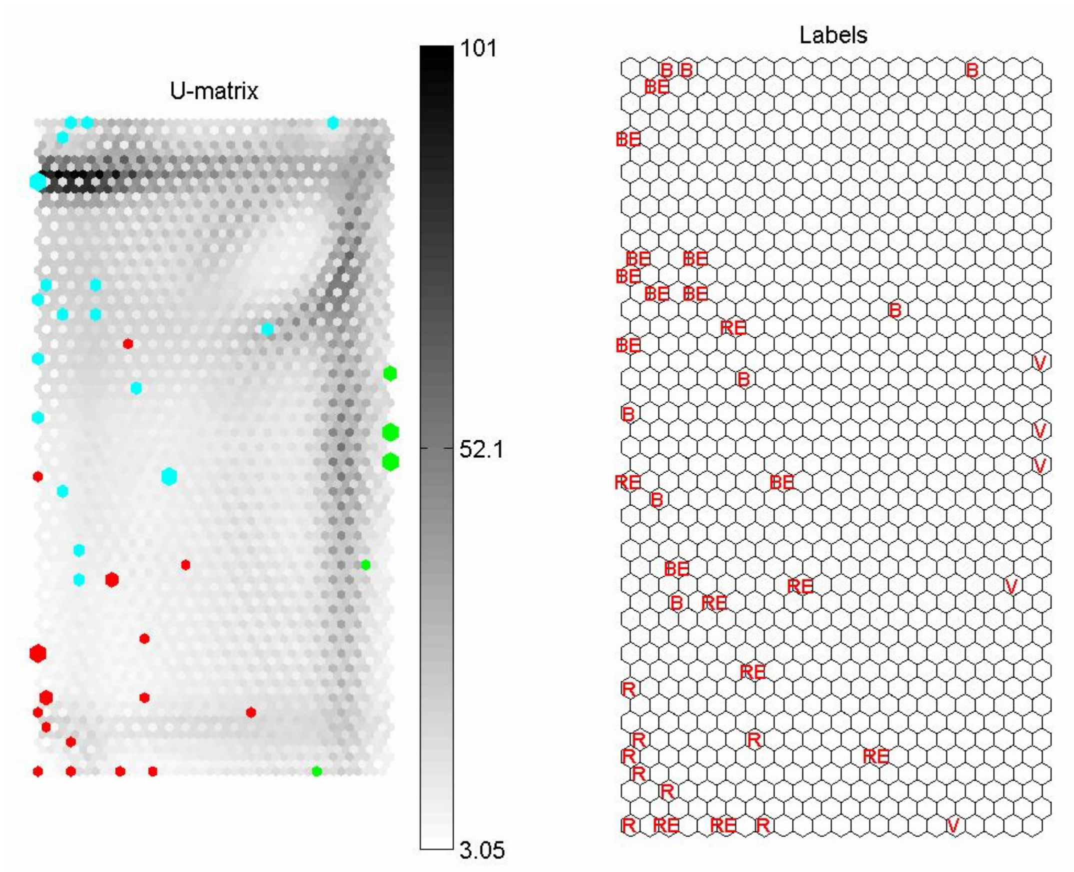

Figure 24: Location of the input vectors with known landuse

Red color has been assigned to road pixels, blue to buildings and green to the vegetation class. Roads and buildings each have a subclass "edge" (labeled as RE, BE), which defines the edges associated with roads and buildings. As seen in the mapping a few vectors describing the "road-edge" class have been mapped inside the building class. This is because those edges could have been located at road/building junctions. The final mapping also shows that SOM is able to successfully discriminate the vegetation class from the other (built classes). The boundary between the building class and the road class is not extensively distinct. Even there the road pixels have been clearly been clustered away from the building class.

4.4.7 Results and accuracy analysis

The accuracy of the network in clustering the input vectors is calculated by finding the percentage of "pure" class pixels which are correctly classified into the correct cluster. The reference image used was created at the sampling locations by visual interpretation. 1m data was used along with 4 m MSS data to create reference images to be used in accuracy analysis. The classes in the reference image included urban features sans roads, roads, vegetation, water, vegetation/road edge and building edges and road edges. The vegetation/road edge class was combined with the road class to increase the likelihood of the road pixels covered by the canopy overhangs to

Figure 25: Classification results of road extraction overlaid on PAN imagery

be included in the road class. Road edges (RE) and building edges (BE) were unique classes in final classification. Road edges (RE) were combined with the road class and building edges were included in the urban class (B).

Figure 25 shows the extracted pixels belonging to the road class. Area/length based threshold was used to remove some of the apparent mis-classifications. The other mis-classified polygons were removed manually.

The error matrix for classification accuracy in low spatial frequency locations are shown in Table 4 and Table 5.

Table 4: Error matrix obtained after supervised classification in a low frequency image subset

Classified Data	Reference data		
	Road	Vegetation	Buildings
Road	91.3	1	8.6
Vegetation	2	95.4	3.1
Buildings	5.1	3.4	84.3

Table 5 shows the error matrix obtained after classifying at a low frequency location using the SOM.

Table 5: Error matrix obtained after classification using SOM in low frequency image subset

Classified Data	Reference data		
	Road	Vegetation	Buildings
Road	90.5	1.3	7.8
Vegetation	2.1	93.8	3.5
Buildings	3.7	3.2	81.5

Examination of the Table 4 and 5 shows that the classification accuracies of supervised classification is slightly better than the results obtained by using SOM. This could be because edge information was not very useful in regions of low spatial frequency.

The error matrix obtained after supervised classification at high frequency region is given in Table 6.

Table 6: Error matrix obtained after supervised classification in high spatial frequency image subset

Classified Data	Reference data		
	Road	Vegetation	Buildings
Road	82.8	1.7	16.2
Vegetation	3.3	89.2	3
Buildings	14.3	2.9	75.1

It can be seen that the classification accuracies have reduced in comparison to the results obtained in Table 4. The main reason as discussed earlier is due to the close configuration of different landuse classes leading to mixed spectral responses by the pixels.

Table 7: Error matrix obtained after application of SOM in high spatial frequency image subset

Classified Data	Reference data		
	Road	Vegetation	Building
Road	87.2	1.6	9.5
Vegetation	2	85.3	3.5
Buildings	6.4	3.7	79.8

Table 7 shows the improved classification accuracy with respect to road pixels. The edge information incorporated in the input vector has allowed the road pixels to be classified with a greater accuracy in high spatial frequency images. Moreover the classification accuracy of buildings is also improved with respect to supervised classification. Only the classification accuracy of vegetation pixels is reduced in an urban scenario. The percentage of cross-classified pixels is also reduced but at the cost of increased number of unclassified pixels.

Hence, the results stress the utility of hybrid algorithm for improving classification when applied to high resolution imagery.

CHAPTER 5
Conclusions and Future Recommendations

5.1 Motivation

Advances in the field of remote sensing have spawned numerous application areas in resource management. High resolution satellite imagery provides a quick glimpse of the region under study. Synoptic view which is synonymous with remote sensing has made it possible for high resolution imagery to establish a niche in terrestrial monitoring/management/analysis. In this thesis, the potential utility of high resolution imagery in landuse classification in regions of high spatial frequency, by the application of application of a hybrid algorithm has been investigated.

With the availability of high resolution data, existing techniques for feature extraction were unable to utilize the opportunity for improved classification accuracies. Hence, an improved procedure which utilizes the strengths of existing classifiers and achieves better feature extraction is necessary.

The primary motivation for this research was the need to accurately build and update road network database. Road networks have varying spectral characteristics and spatial characteristics. And hence provide a rich testing ground for implementing hybrid algorithms to improve classification. The results could be utilized in rapid building of road network database, in regions without accurate networks and for rapidly updating the network in regions with continuous changes in the network.

Implementation of an algorithm to automate the inventory process aims to facilitate improved identification percentage and shorter updating time, leading to increased productivity. The foremost issue is in classifying the error pixels into a suitable landuse class. While classifying into various classes there are a few doubtful classes which can be resolved by using artificial intelligence.

By the use of self-organizing maps for clustering, pixels in the neighborhood (edges) of the roads are identified as separate from the road pixels. The final clustering highlights the similarities in the spectral characteristics of the pixels surrounding the road segments. These ambiguous pixels influence the classification algorithms by skewing the discriminating statistics and hence by classifying them into separate category, classification accuracies of the other landuse classes could also improve.

5.2 Building the Transportation Network

The extraction of the main elements of a transportation network as described in Table 1 was attempted. With AFE, only the pavement width and road network were extracted which included road intersections. The pixels were converted to polygon vectors using Erdas Imagine. Parking lots were included in the road class due to the similarity in the surface materials. The resulting network had discontinuities due to tree canopies and due to the tall buildings. Vegetation road edge class was combined with the road class but the road pixels in proximity to tall buildings were lost. This problem was higher farther from the image center due to the affects of relief distortion.

However the freeways and major arterials were extracted with a few discontinuities. Only the local streets close to the tall buildings could not be extracted.

This did not affect the accuracy as the reference image was built by manually classifying the sample subsets. The road pixels shadowed by the buildings (hence not visible) were classified into building class. Therefore, the buildings did not affect the accuracy estimate.

Additional analyst interpretation is required to extract other inventory elements.

5.3 Conclusions

With close to 90% of road pixels being identified from 1-meter resolution imagery; lane width, section length were extracted. The results obtained were very encouraging and the application of SOM as a successful visualization for high dimensional data was re-established. However, much work is still needed regarding the post-processing stage and the interpretation of results. In this thesis the results from hybrid extraction were validated using manual feature delineation. This process was very time consuming as the class membership of each pixel was assigned manually for building the reference images.. Research into region growing is necessary for improved data validation process which could be used to interactively select pixels from the same information class.

5.4 Future Recommendations

Success of any inventory initiative involves rapid data collection followed by timely data analysis. Feature extraction procedures as explained in the previous chapter have to be coupled with Change detection techniques to update road inventories. With multi-user data requirement the costs of acquisition would be reduced and hence planners would be able to acquire data at frequent intervals for assessing the changes in the network and the roadside facilities.

3 D models can be analyzed with images to get edges which are accentuated in surface models. The ambiguity in determining the roadside elements will be substantially minimized when the imagery can be draped on the surface model. Suitable attributes like 3D edges/contours should be generated such that they could be used along with the other feature attributes for improved clustering.

Finally for a complete road inventory, use of satellite imagery seems unsuitable unless the sensor technology witnesses major developments to collect even higher resolution data. Aerial photography can be used to augment the synoptic satellite imagery along the road corridors to collect inventory data.

Hence, satellite imagery can be successfully utilized to extract the network characteristics like number of lanes, lane width. And with improvements in AFE procedures, complete removal of the human element in feature extraction is not far away.

References

1. Land Transport Authority of Singapore, www.lta.gov.sg.

2. *Remote Sensing and Image Interpretation*, Thomas M Lillesand, Ralph W. Keifer, 1994.

3. *Transportation Engineering: Planning and design*, Paul H. Wright, Norman J. Ashford, 3rd Edition, John Wiley and Sons.

4. *Highway Engineering with contribution*, James Lai et al., Wright Paul H, 6th Edition, John Wiley and Sons, 1996

5. Infrastructure Inventories- The starting point of effective highway Management, Opiela K.,Perkins D., Vol. 56, No. 2, Feb 1986, ITE Journal.

6. Neural Networks, Theoretical Foundations and Analysis, Edited by Clifford Lau, IEEE press.

7. *Automated Feature Extraction: The Key to future Productivity*, Firestone, L., S. Rupert, J. Olson, and W. Mueller, Photogrammetric Engineering and Remote Sensing, Vol. 62, No. 6, pp. 671-674,1996.

8. Kohonen's Self Organizing Maps and their use in Interpretation, Dr. M. Turhan Taner, Rock Solid Images,2000.

9. *Image processing with neural networks – a review*, M. Egmont-Petersen*a*,D. de Ridder*b*, H. Handels*c*, Pattern Recognition, (ã Elsevier Science), 2001.

10. *National Standard for Spatial Data Accuracy*, http://www.fgdc.gov/standards/status/sub1_3.html [accessed Jan 2000]

11. *Manual of Photographic Interpretation, Second Editio,n* Warren R. Philipson.. American Society for Photogrammetry and Remote Sensing, Maryland, 1997.

12. *Photologging*, Baker William T, Transportation Research Board, National Research Council, Washington, D.C., 1982.

13. Extraction and Utilization of Geometrical and Contextual Information in very High Resolution IKONOS Satellite Imagery, Annie HUI, et al., Asian Conference On Remote Sensing - (ACRS),Taiwan, 2000.

14. *Spectrometry and Hyperspectral Remote Sensing of Urban Road Infrastructure Martin* Herold et/ al., Online Journal of Space Communication, Issue No. 3, http://satjournal.tcom.ohiou.edu/Issue03/main.html

15. *Practical Solutions for Specific Generalization Tasks*, Dan Lee, ICA Generalization Working Group Workshop, 1999.

16. *Self-organized clustering for road extraction in classified imagery*, Doucette, P., Agouris, P., Stefanidis, A. & Musavi, M. ISPRS Photogrammetry & Remote Sensing. 55. 347-358, 2001.

17. A New Approach for Automatic Determination of Edge Center Lines in a Digital Image, Jaan-Rong Tsay, Asian Conference On Remote Sensing -

(ACRS), Taiwan, 2000.

18. *Edge detection in multispectral images*, Cumai A.,Computer Vision, Graphics and Image Processing: Image Understanding, Vol 53, pp. 40-51, Jan 1991.

19. *Feature extraction methods for character recognition*, Trier o., Jani A., Taxt T., Pattern Recognition, Vol 29, No. 4,pp. 641-662, 1996.

20. *Data Exploration using Self-organizing maps*, Kaski S., Dissertation, Helsinki University of Technology, 1997.

21. Segmentation of Natural Images using Self-organizing Feature Maps, Campbell N., Thomas B., and Troscianko T., British Machine Vision Conference, pages 223—232,1996.

22. *Towards Core Image Processing with Self-Organizing Maps*, Brown D., Craw I., Lewthwaite J.,http://citeseer.nj.nec.com/brown01towards.html.,2001.

23. *IKONOS imagery products and product guide*, Space Imaging, http://www.spaceimaging.com/whitepapers_pdfs/IKONOS_Product_Guide.pdf , 2002.

24. *Digital image processing*, Gonzalez, R. C. and R. E. Woods, Addison-Wesley, 1992.

25. *The image processing Handbook*, John C. Russ, CRC press, 1992.

26. *Digital Image Processing*, R. Gonzalez and R. Woods Addison Wesley, pp 414 – 428, 1992.

27. *A Computational Approach to Edge Detection*, J. Canny, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 8, No. 6, Nov. 1986.

28. *What size net gives valid generalization?* , Baum, E., and Haussler, D., Advances in Neural Information Processing Systems 1, Morgan Kaufmann, pp 81-90, 1989.

29. *Expert Systems Tutorial*, MAJ George Hluck, U.S. Army War College in Carlisle, PA, Nov 1996.

30. *SOM Toolbox for Matlab 5* , Juha Vesato,Report A57, Johan Himberg, Esa Alhoneimi and Juha Parhankangas, Helsinki University of Technology, 2000