ROAD EXTRACTION FROM HIGH RESOLUTION SATELLITE IMAGES USING ADAPTIVE BOOSTING WITH MULTI-RESOLUTION ANALYSIS

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF INFORMATICS
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

UMUT ÇİNAR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF INFORMATION SYSTEMS

SEPTEMBER 2012

ROAD EXTRACTION FROM HIGH RESOLUTION SATELLITE IMAGES USING ADAPTIVE BOOSTING WITH MULTI-RESOLUTION ANALYSIS

Submitted by **Umut Çinar** in partial fulfillment of the requirements for the degree of **Master of Science in Information Systems, Middle East Technical University** by,

Prof. Dr. Nazife Baykal ______________________
Director, **Informatics Institute**

Prof. Dr. Yasemin Yardımcı Çetin ______________________
Head of Department, **Information Systems**

Prof. Dr. Yasemin Yardımcı Çetin ______________________
Supervisor, **Information Systems, METU**

Examining Committee Members:

Prof. Dr. Uğur Halıcı ______________________
Electrical and Electronics Engineering, METU

Prof. Dr. Yasemin Yardımcı Çetin ______________________
Information Systems, METU

Asst. Prof. Dr. Erhan Eren ______________________
Information Systems, METU

Asst. Prof. Dr. Banu Günel ______________________
Information Systems, METU

Asst. Prof. Dr. Alptekin Temizel ______________________
Work Based Learning, METU

Date: ____________________

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Umut Çinar

Signature : __________

ABSTRACT

ROAD EXTRACTION FROM HIGH RESOLUTION SATELLITE IMAGES USING ADAPTIVE BOOSTING WITH MULTI-RESOLUTION ANALYSIS

Çinar, Umut

MS., Department of Information Systems

Supervisor: Prof. Dr. Yasemin Yardımcı Çetin

September 2012, 126 pages

Road extraction from satellite or aerial imagery is a popular topic in remote sensing, and there are many road extraction algorithms suggested by various researches. However, the need of reliable remotely sensed road information still persists as there is no sufficiently robust road extraction algorithm yet. In this study, we explore the road extraction problem taking advantage of the multi-resolution analysis and adaptive boosting based classifiers. That is, we propose a new road extraction algorithm exploiting both spectral and structural features of the high resolution multi-spectral satellite images. The proposed model is composed of three major components; feature extraction, classification and road detection. Well-known spectral band ratios are utilized to represent reflectance properties of the data whereas a segmentation operation followed by an elongatedness scoring technique renders structural evaluation of the road parts within the multi-resolution analysis framework. The extracted features are fed into Adaptive Boosting (Adaboost) learning procedure, and the learning method iteratively combines decision trees to acquire a classifier with a high accuracy. The road network is identified from the probability map constructed by the classifier suggested by Adaboost. The algorithm is designed to be modular in the sense of its extensibility, that is; new road descriptor features can be easily integrated into the existing model. The empirical evaluation of the proposed algorithm suggests that the algorithm is capable of extracting majority of the road network, and it poses promising performance results.

Keywords: road extraction, high resolution multi-spectral satellite images, multi-resolution analysis, Adaptive boosting, Mean-shift image segmentation

ÖZ

ADAPTİF DESTEKLEME İLE ÇOKLU ÇÖZÜNÜRLÜK ANALİZİ KULLANARAK YÜKSEK ÇÖZÜNÜRLÜKLÜ UYDU İMGELERİNDEN YOL TESPİTİ

Çinar, Umut

Yüksek Lisans, Bilişim Sistemleri Bölümü

Tez Yöneticis: Prof. Dr. Yasemin Yardımcı Çetin

September 2012, 126 sayfa

Uzaktan algılamada uydu ve hava fotoğraflarında yol tespiti popüler bir konudur. Çeşitli araştırmacılar tarafından önerilmiş yol çıkarımı algoritmaları mevcuttur. Ancak güvenilir yol bilgisine olan ihtiyaç hala devam etmektedir çünkü yeterli ölçüde gürbüz bir yol tespiti algoritması henüz bulunmamaktadır. Bu çalışmada çoklu çözünürlük analizi ve adaptif destekleme tabanlı sınıflandırıcılardan yararlanarak yol tespiti problemi çalışılmıştır. Bir başka deyişle bu çalışmada yüksek çözünürlüklü ve çok bantlı uydu imgelerinde bulunan yolun hem tayfsal hem de yapısal özniteliklerini kullanan yeni bir yol tespiti algoritması önerilmiştir. Önerilen model üç ana bileşenden oluşmaktadır; öznitelik çıkarımı, sınıflandırma ve yol tespiti. Çokça kullanılan tayfsal bant oranlarından görüntülerin yansıtma özelliklerini temsil etmesi için yararlanılmıştır. Öte yandan imge bölütleme ve ardından uygulanan bir uzatılmışlık puanlama yöntemi sayesinde yol parçalarının çoklu çözünürlük analizi altyapısı içerisinde değerlendirilmesi mümkün olmuştur. Çıkarılan öznitelikler adaptif destekleme öğrenme işlemine tabii tutulmuştur ve söz konusu öğrenme metodu karar ağaçlarını tekrarlı olarak kaynaştırarak yüksek doğruluk ile çalışan bir sınıflandırıcı oluşturmuştur. Adaptif destekleme yöntemi tarafından verilen olasılık haritasından yol ağı tespit edilmiştir. Bu algoritma genişletilebilirlik anlamında modüler bir yapıya sahiptir yani yeni yol belirteçleri kolaylıkla varolan model entegre edilebilmektedir. Önerilen algoritmanın deneysel değerlendirmesi göstermiştir ki algoritma yol ağının büyük bir bölümünü tespit edebilmektedir ve gelecek vaadeden performans sonuçları sunmaktadır.

Keywords: yol çıkarımı, yüksek çözünürlüklü çoklu bantlı uydu imgeleri, çoklu çözünürlük analizi, adaptif destekleme, kayan ortalama imge bölütleme

To the humankind…

ACKNOWLEDGEMENTS

I am grateful to express my sincere gratitude to my supervisor Prof. Dr. Yasemin Yardımcı Çetin for her encouragement, priceless guidance in making this research possible. I also appreciate her consistent support in fostering my scientific intellectuality and commitment to my professional development.

I am also grateful to Prof. Dr. Uğur Halıcı for inspiring me to utilize adaptive boosting in the road extraction problem. I also thank my thesis jury members, Alptekin Temizel, Banu Günel and Erhan Eren for their suggestions.

Special thanks to my colleagues Ersin Karaman and Ekin Gedik for their invaluable supports in resolving implementation specific issues.

Weekly meetings with research groups of ASELSAN and HAVELSAN have inspired me with new ideas regarding my research. I also thank these institutios for supporting my graduate studies at METU.

I also thank to my mother and father for their endless love, moral support and trust.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Motivation

With developing space imagery technologies, automatic object detection from satellite images became an essential component for many applications. Likewise, automatic detection of road networks from the satellite or aerial images is a popular research subject. The automatic road extraction is primarily needed by Geographic Information Systems (GIS) applications, since manual data acquisition for GIS database update is a costly and time consuming task. GIS applications have a wide area of usage including defense industry, transportation systems, emergency management, urban planning and navigation applications. Despite the existence of high demand on the automatically acquired road information for GIS applications, a robust road extraction algorithm outperforming all other alternative extraction techniques does not exist. Therefore, the topic is needed to be studied further, and new road extraction strategies should be devised.

1.2 Scope and Goal

This study is devoted to explore the automatic road extraction problem and to propose a new road extraction algorithm based on Adaboost with multi-resolution analysis. In fact, this dissertation combines the insights gained from two years of research conducted on the subject of road extraction. In the scope of the study, the strengths of multi-resolution analysis and Adaboost learning for discriminating roads from non-road regions are analyzed and reported. 4-bands IKONOS images are utilized for development and experimentation stages. The main purpose of this study is to establish a new road detection scheme exploiting supervised learning and optimization routines in order to construct an accurate classification model.

1.3 Contribution of Thesis

The power of the proposed algorithm comes from its multi-resolution feature extraction scheme and supervised classification routine. Since the structural features extracted from various resolution levels favors in the roads with different width properties, it is convenient to boost the elongatedness features detected in various resolution levels with a basic spectral descriptor of the multi-spectral image. Moreover, Adaboost learning algorithm is capable of generating better estimates for road and non-road class discrimination compared to unsupervised alternatives. This study also includes a new mean-shift image segmentation parameter optimization strategy designed for multi-resolution structural analysis. Firstly, the individual inferences of the mean-shift parameters into the elongatedness distributions of road and non-road classes are examined. Then, we designate separate parameter sets for the segmentation of each resolution levels by the pattern search algorithm, a direct search optimization method. By this way, the classification component is utilized with the features having maximal discrimination capabilities. On the other hand, two alternative unsupervised probability map construction strategies are proposed and discussed. In other words, the proposed road extraction model can be converted to an unsupervised algorithm provided that the optimal segmentation parameters are ensured. Furthermore, three different road extraction methods are compared in order to supply application dependent design alternatives. The region growing based road extraction technique is designed to validate topological connectivity and elongatedness properties of the extracted road parts, whereas the thresholding based extraction methods are empirically proved to be promising in detecting the entire road network.

1.4 Outline of Thesis

The dissertation document is composed of five chapters including introduction, background, methodology, experiments and conclusion. The literature survey on the road extraction is presented in Chapter-1. Then, the theoretical background regarding to the proposed algorithm is elaborated in Chapter-2. In Chapter-3, the proposed road extraction model is detailed and the design principles of individual components are explained. Experimental work conducted on the algorithm is unveiled in Chapter-4, and the empirical deductions regarding to the relevance of the proposed scheme are obtained. Moreover, the parameter tuning operations for individual components as well as entire model are taken place in the experimentation chapter. Finally, the research is concluded in Chapter-5 by summarizing the major concepts and providing key elements for the direction of future work.

CHAPTER 2

BACKGROUND

In this chapter of the thesis, theoretical background behind the study is elaborated. Firstly, data and sensor types utilized in remote sensing imagery are introduced in section 2.1. The road extraction literature is investigated in section 2.2. Spectral band ratios investigated as spectral non-road features are explained in section 2.3, and the multi-resolution concept is introduced in the following section 2.4. Elongatedness analysis techniques used for road segment evaluation are detailed in section 2.5. After introducing the mean-shift theory and its application to image segmentation in section 2.6, Adaboost learning algorithm is described including its major theoretical properties. On the other hand, we employ decision trees as weak classifiers in this study, so they are elaborated in section 2.8. Histogram equalization and region growing procedures used in the road extraction component are also explained in sections 2.9 and 2.10, respectively. Finally, the pattern search algorithm, an optimization method, is presented in section 2.11.

2.1 Data and Sensor Types in Remote Sensing Imagery

Remote sensing is conceptualized as the acquisition of information about a remote object or phenomenon without establishing any physical contact with the object (Schowengerdt, 2007). The sensing term indicates the procedure of acquiring visual data by aerial sensor devices by means of propagated signals. Sensors employed for remote sensing imagery can be categorized into two classes by means of detected radiation’s source. Passive sensors rely on the natural radiation supplied by the Sun to collect spectral information of target objects emitting or reflecting the sunlight. Photographic, infrared and charge-coupled devices are examples of passive remote sensors. On the other hand, electromagnetic radiation of a specific wavelength can be conveyed by using special purpose illumination devices, and the radiation created by such devices can be used to measure the scattered radiation from the targets. Such sensors are called as active sensors. RADAR and LIDAR are examples of active remote sensors which are capable of measuring the location, height, speed and direction of an object by considering the time delay between emission and return.

Another discriminating factor among the remote sensors is the range of electromagnetic regions in which they operate such as optical and microwave (Figure 1).

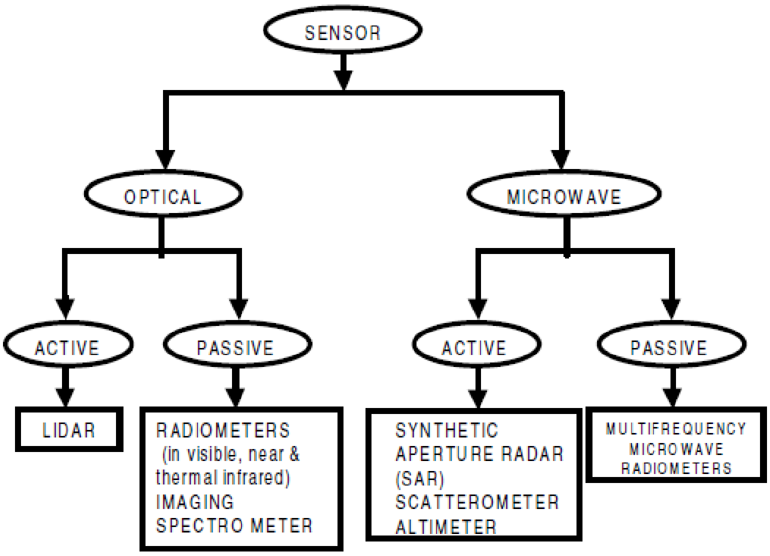


Figure 1: Classification of sensors, adapted from (Rangnath , 2001)

In microwave remote sensing, electromagnetic radiation in the microwave wavelength region is used for sensing relevant information about the Earth's atmosphere, land and ocean. Microwave sensing devices enable the usage unique features of objects such as frequency characteristics, Doppler Effect, polarization and back scattering which cannot be provided by optical remote sensing devices. The main disadvantage of microwave based remote sensing is that microwave data requires sophisticated methods to extract useful information (Murai, 1991). Synthetic aperture radar (SAR), microwave scatterometers and radar altimeters are examples of active microwave sensors whereas microwave radiometer is one of the examples of passive microwave sensing.

Optical remote sensing devices employ visible, near infrared and short wave-infrared sensors to sense the radiation reflected from objects on the Earth. Basically, different wavelengths are reflected and absorbed differently by different types of materials, therefore; reflectance properties of objects vary depending on object type. Optical remote sensors can be grouped in four main categories according to the number of spectral bands provided by device (Heidner & Straus, 2005);

i. **Panchromatic imagery devices:** Panchromatic images are single band images containing the radiation characteristics of an object in a broad wavelength range. IKONOS-PAN, SPOT and HRV-PAN are major examples of such devices.

ii. **Multi-spectral imagery devices:** Multi-spectral imagery provides both panchromatic brightness and multi-layer spectral information of the observed targets. In these devices, each sensing channel is sensitive to radiation belonging to a narrow wavelength band such as red, green, blue and near infrared. The multi-layer data typically contains less than 10 bands. LANSAT MSS, LANDSAT TM, HRV-XS and IKONOS MS are some examples of multi-spectral remote sensing devices.

iii. **Super-spectral imagery devices:** Like multi-spectral imagery devices, super-spectral devices also provide a multi-layer data in which each layer represents the radiation in a narrow wavelength band. But super-spectral devices are capable of sensing more than 10 bands, in fact; the resulting bands have narrower bandwidths enabling more precise spectral analysis for target objects. MODIS and MERIS are examples of super-spectral sensors.

iv. **Hyper-spectral imagery devices:** Hyper-spectral sensors, namely imaging spectrometers, are capable of acquiring images having a hundred or more contiguous spectral bands. Hyper-spectral sensing is the most accurate technology for detecting the unique spectral characteristics of any material. Hyperion is an example of hyper-spectral imaging devices.

2.2 Literature Survey

Significant research effort is spent for the road extraction problem from aerial or satellite imagery. The common aim is to extract roads represented by binary masks in which true pixels represent road regions and false pixels indicate non-road regions. The first research attempts for this topic have begun in the mid-70s by the study of (Bajcsy & Tavakoli, 1976). For over two decades, automatic road extraction from remote sensing imagery has been an active research area in computer vision, photogrammetry, remote sensing and geographic information system (GIS) researchers. In a detailed literature survey study, approximately 250 state-of-art techniques consisting of several road extraction studies are reviewed by (Mena J. , 2003). In this study, methods developed for road extraction from satellite images between the years of 1987-2002 are classified into several categories according to their preset objective, extraction technique applied and type of sensor utilized. In a more recent study, (Hauptfleisch, 2010) reviews the major contributions to the topic by grouping them into categories such as automatic seeding, classification, edge detection, Hough transform, multi-resolution analysis and road tracking. Both of these articles introduce the essential methods used by many researchers for years in the perspective of their studies. In this section, the studies conducted by former researchers concerned with the road extraction problem are introduced. This literature survey groups the related papers according to their core theory. These groups are consisted of

- Basic Image Processing Methods
- Road Tracking
- Multi-resolution Analysis
- Knowledge Based Methods
- Supervised Classification Methods
- Segmentation Based Methods

2.2.1 Basic Image Processing Methods

In the literature, edge detection (Mirnalinee, Das, & Varghese , 2009) Hough transformation (Guan , Wang, & Yao , 2010), snakes (Zhao & Wang , 2010) , template matching (Lin , Zhang, Liu, & Shen , 2008) and mathematical morphology (Gaetano , Zerubia, Scarpa, & Poggi, 2011) are commonly used basic image processing techniques for extraction of roads from remote sensing imagery.

A new object based road extraction strategy suitable for large scale imagemaps is elaborated in the paper (Amini, Saradjiana, Blaisc, Lucas, & Azizi, 2002). The road extraction model is composed of two parallel processes; first one aims to detect straight line segments and the second process is responsible for finding road skeleton. At the straight line extraction stage, quadratic picture tree method is utilized to obtain initial road segmentation based on series of homogeneity tests. The raw straight lines are obtained by raster-to-vector conversion, and the resulting fragmented line segments are restored to longer straight line segments by using a perceptual grouping approach containing several steps based on geometric and structural evaluation of the line segments. At the road skeleton detection stage, the input image is reduced to a lower resolution by using direct wavelet transform so that roads can be considered as rough regions having two or three pixels width. Road skeleton extraction is accomplished by the help of Split and Merge segmentation algorithm applied on the low resolution wavelet transform image. By combining straight lines and road skeleton by road-side tracing algorithm, contours identifying road regions are detected. The empirical evaluation of the algorithm suggests that influence of the segmentation method on overall accuracy is an important design issue.

An automatic road detection approach is studied by (Jinxin , Qixin, & Liguang , 2006) based on an edge detection and tracing strategy. The proposed algorithm firstly preprocesses the input image to reduce noise and adjust contrast in order to ease the separation of road objects from the background. The authors presume the intensity distributions of road regions as being distinguishable from image histogram, and raw road mask is attained by using Otsu thresholding applied on the gray scale image. Subsequently, morphological reconstruction and boundary tracing algorithms are utilized to obtain edges of the binary mask. Shape characteristics of the edges are analyzed by using pre-defined shape indices, and the extracted shape features are fed into k-means clustering algorithm to identify road regions in an unsupervised manner.

A more specific road extraction algorithm proposed in (Jin , Feng, & Li , 2008) focuses on the high resolution satellite images, and roads are considered as the regions having continuity and homogeneity properties. Therefore, the authors aim to detect road surfaces rather than road lines. For this purpose, the input image is roughly segmented with a technique called as homogram segmentation exploiting both of the spatial and spectral characteristics of the data. Before homogram construction step, the input image is exposed to contrast stretching and Gaussian smoothing operations to attain histograms having proper shapes for homogram segmentation. Then, several morphological operations are applied on the result of homogram segmentation to obtain a more accurate road mask. The final road extraction result is represented by pruned skeletons of the road mask generated by series of thinning and vectorization operations.

The method suggested in the study (Zhao & Xili, 2010) employs mathematical morphology together with active contour model (snakes) to detect roads from high resolution satellite images. The input gray-scale image is firstly converted to binary image by a threshold value which is determined by automatically examining the distribution of image’s histogram. In the second step, a connected component analysis using area-to-perimeter ratio to discriminate the road regions from non-road regions is carried out. The remaining connected component’s contour lines are fed into Snake model to identify actual road region. The extraction results are visually compared with the results obtained by manually drawn contour tracking procedure, and both of the results look similar.

2.2.2 Road Tracking

Road tracking methods utilizing statistical estimation methods devised by various researchers provide robust extraction strategies. Since roads can be generalized as elongated, continuous and relatively homogeneous regions, road tracking algorithms are considered as plausible methods for road extraction.

The study (Ye, Su, & Tang, 2006) uses a statistical model supported by basic image processing methods to detect roads by tracking them. The statistical model introduces a particle filter, a recursive and dynamic procedure. Main assumption of this study is that roads are composed of similar gray valued pixels but roads can be occluded by several objects such as vehicles and trees. Therefore, to overcome obstacle inference problem, they use both gray values and edge point distributions in a particular observation window as features. Also, the algorithm is allowed to update the road model in order to adapt to changes occurred in road profiles. It is designed to estimate the parameters describing the position and the shape of the road. The authors initiate the tracking by finding seed points with the help of Hough transform. Parallel straight lines are found in the input image, and then the points in the middle of these lines are selected as seed points. In general, this method provides less vulnerable tracking to obstacles caused by environmental effects, however, each road part to be extracted still needs separate seed points.

A practical application of road extraction is proposed in the paper (Zhou, Bischof, & Caelli, 2006). In this study, the authors combined the input taken from human operator with well-known Bayesian filters such as Kalman and particle filters. An initial road profile is identified by human operator. Then, the next state of the tracker is estimated by Bayesian filter. The state model consists of current coordinates of the tracker, direction of the road and change in road direction. The observation model is obtained by matching the reference model with the observed profile. The proposed approach is superior to many automatic road extraction algorithms existing in the literature.

In the research conducted by (Ma, Qin, Du, Wang, & Jin, 2007), an automatic road extraction algorithm which operates on 15 meters LANDSAT Enhanced Thematic Mapper (ETM) panchromatic images is proposed. This study firstly defines the basic characteristics of roads according to their geometric, radiometric, topologic and contextual features. Dual road edge following method is applied to the edge map of the image. False positives are eliminated by validating the roads' spectral homogeneity and structural properties. The authors indicate that this method needs to be extended by using different information or techniques to show good performance on different types of images.

In the paper (Hu J. , Razdan, John, Cui, & Wonka, 2007), another automatic road extraction algorithm using aerial images is proposed. They consider roads as long and thin structures depicting high contrast with respect to their neighborhoods. Based on roads' directional rectangularity, bounded width and contrast properties, the algorithm firstly extracts candidate road segments called footprints. The footprints are identified by utilizing the spoke wheel operator which is type of a directional angular operator. Then, footprints are classified as road and non-road regions by using a Bayes model in order to eliminate false positives. The features for this model are extracted by a rule based approach called toe-finding algorithm which mainly analyzes the directional variety of road or non-road footprints. In the Bayesian decision model a lognormal model distribution is constructed by using area-to-perimeter ratios of road footprints. Unlike most road tracker algorithms available in the literature such as (McKeown & Denlinger, 1988), (Baumgartner, Hinz, & Wiedemann, 2002) and (Geman & Jedynak, 1996), the authors also propose an automatic road seeding method, which makes use of structural properties of road footprints. Although the proposed method suffers from the problem of over-extraction of roads due to junction complexity, the quantitative performance of the algorithm is competitive.

In a more recent study (Li, An, & Chen, 2011), the authors propose a new method considering two main difficulties faced in the road extraction problem; identifying initial seeds and a robust tracking strategy. An adaptive Canny based method detects edges of the input image. This edge information is fed to the Harris corner detection algorithm to determine interest points. Beginning from these points, seeds are selected by the help of context features such as adjacency, parallelism, perpendicularity and intersection of line segments. This automatic seed selection procedure is based on a newly proposed algorithm called as bat algorithm. Seed points induced by context objects other than roads are eliminated by firstly creating a circular search region at the given seed and then validating its presence with pre-defined rules. Then, initial road tracking position and direction are determined, and a rectangle reference template is also constructed by using initial seeds. A variation of mutual information matching method (Maes, Collignon, Vandermeulen, Marchal, & Suetens, 1997) is used to track roads along the main road axis.

2.2.3 Multi-resolution Analysis

Multi-resolution analysis is the process of generating images with different scales by applying repeated smoothing and down-sampling operations. Multi-resolution analysis stems from the micro-local analysis and the theory of differential equations. Pyramid methods and discrete wavelet transformation are two common representation techniques for multi-resolution analysis. Scale space representation of an input image provides good approximation according to human vision, in fact; some studies signify the strong relationship between scale space theory and biological vision (Young, 1987) (DeAngelis, Ohzawa, & Freeman, 1995).

In the paper (Gong & Wang, 1997), it is stated that high resolution imagery proposes new opportunities for classification based road extraction algorithms. In the study, several experiments on both fine and coarse scale images are conducted. These experiments mainly use direction profile matching, linear analysis and classification based methods. In this comparative study, the best extraction result is obtained by firstly morphologically filtering the output of the ratio NIR/(R+NIR), and then running the linear extraction algorithm on the binary image.

Likewise, the study presented in (Mayer, Laptev, & Baumgartner, 1998) takes advantage of multi-resolution analysis by ribbon snake approach. Initial seed lines are extracted at the coarse scale which provides less information about the scene but has been less occluded by environmental factors such as cars, trees and shadows. The initial line seeds are needed to start a ribbon snake operation on fine scale which treats the roads as bright and elongated regions. Road and non-road discrimination after snake operation is done by analyzing road segment’s width homogeneity. Non-salient roads and crossings are handled separately. Non-salient roads are detected by a top-down approach exploiting the roads’ connectivity constraint. Crossing regions are also extracted by the coarse scale analysis reinforced with fine scale. Multi resolution snakes appeared on a more recent study (Péteri & Ranchin , 2003) for street extraction from high resolution imagery. The proposed method consists of two main components; road segment extraction and road junction detection. A new energy variable, namely parallelism energy, is introduced to validate parallelism of the road sides. At the last stage, intersections of extracted road segments are detected by considering their direction and position.

2.2.4 Knowledge Based Methods

The paper (Lee , Park, Lee, & Kim , 2000) proposes a knowledge based road extraction strategy which treats the roads as separate regions. A segmentation result is obtained by using hierarchical multi-scale gradient watershed transformation. The candidate road regions are identified by using prior knowledge about the road areas such as gray level value, elongatedness and connectedness. Lastly, directional cone search operation applied on the candidate road segments to satisfy their connection constraint by expanding suitable road segments. The method is tested with 1-m resolution IKONOS images, and it is observed that there are both false negative and positive detections caused by deficiencies in the segmentation stage.

The study carried in (Hinz & Baumgartner, 2003) benefits from multi-view imagery to extract roads in urban areas from aerial images. The detailed knowledge about the scene’s context is provided by explicitly defined scale-dependent rules. The road extraction strategy is based on three features incorporating contextual, self-diagnosis and multiple-view components. The actual extraction procedure exploits these features by evaluating their relative positions in the space before validation step. Lane segments belonging to road regions are detected by analyzing candidate segments with respect to their orientation difference, gap length, and width and height properties. The final road network mask is completed by an iterative topology based post-processing algorithm proposed in the study (Wiedemann & Ebner, 2000). The proposed approach is capable of resolving clutter caused by environmental factors. The model is aware of context objects such as buildings, trees and vehicles with their multi-view appearance. This is the main strength of the method and it makes road extraction procedure more reliable. As discussed in the paper, the proposed method has deficiencies extracting roads in the regions with complex junctions where it is hard to prove connection hypotheses. Furthermore, the eventual algorithm is constructed on a model with many parameters to be tuned.

The study (Mena & Malpica, 2005) concentrates on road extraction from rural and semi-urban areas. They propose a method for GIS update as well as a new road extraction technique. The proposed framework consists of four major components; data processing, segmentation, vectorization and validation. Firstly, input RGB image is rectified and geo-referenced. At this stage, existing GIS data is prepared as a prior knowledge to be used in the segmentation stage. Then, Texture Progressive Analysis is carried to get binary segmented image. For this procedure, existing GIS data is used to decide on the training regions. In the third module, morphological processing is applied on the binary road mask to obtain vector data of the road network. Finally, geometrical and topological corrections are applied on this vector data to get final result. The quantitative results pose a quality percentage of 70. This method is suitable to be generalized into a semi-automatic technique which can reliably extract roads from high resolution electro-optic satellite images.

2.2.5 Supervised Classification Methods

Genetic algorithms which are stochastic methods in machine learning can be used to extract road regions as in the study (Xiao, Bao, & Jiang, 2004). This study uses fuzzy C-means algorithm to extract features from filtered SAR image. The road model for genetic algorithm's optimization procedure is obtained by using membership values of the clusters. Continuous curves passing thought seed points are considered as individuals.

The study (Zheng, Liu, Shi, & Zhu, 2006) proposes a central road contour extraction method. Firstly, a Support Vector machine classifier is used to divide the input image into two categories; road and non-road. The training samples are provided through the user interface. To get a visually appealing and continuous road contour map, the binary mask is encoded into a tensor field. The tensor voting method is used to extract road contours from the point cloud.

In the study (Zhu, Wen, & Ling, 2011), a remotely sensed image is considered as having three classes; roads, housing and outdoor terrain. Based on this assumption Fuzz C-Means algorithm is used to cluster the input image into three groups. The membership values are given to a Support Vector Machine classifier as features. The final mask is obtained by further post-processing step based on morphological operations.

A neural network based classifier is employed in the study (Kirthika & Mookambiga, 2011). At first step, a neural network is trained with spectral features using ground truth. Then, the result of classification is considered as an indication of road regions. Further features mainly based on texture are used to cluster the result of neural network classification. By this way the final road mask is obtained.

The study (Yanqing, Yuan, Tai, & Shu, 2006) employs Adaboost for discriminating road and non-road regions. The ratio of bright regions, direction consistency of edges and local binary patterns are considered as training features. The connection constraint of the road network is validated by a sliding window operation applied on confidence mask of learning procedure.

In paper (Trinder, 2009) , a cluster analysis with inductive learning technique is generalized to form a road extraction framework, called RAIL (Road Extraction by Inductive Learning). In this hierarchical model, the road extraction work is accomplished by constructing complex objects like road networks from primitive components like road edge pairs. The proposed feature extraction model consists of four hierarchical levels; road edge pairs, linked road edge pairs, intersections and road network. The inductive learning procedure determines the best clustering algorithm with an optimized parameter set. Since the proposed inductive learning scheme validates the learning results with test set, the algorithm shows good generalization properties.

2.2.6 Segmentation Based Methods

Graph Cuts segmentation algorithm cooperating with a Markov Random Model is used in the paper (Del-Toro-Almenares, Mihai, Vanhamel, & Sahli, 2007) to automatically detect linear features. The authors defined an energy function for the model which describes spatial properties of road segments.

A segmentation based method is studied in the paper (Grote & Heipke, 2008). This study makes use of 10-cm resolution CIR aerial images to detect roads. N-cut segmentation algorithm is used to generate an over-segmented representation of the given image. Then, these initial segments are combined to get larger segments by analyzing their edge and color properties. Also, irregularly shaped road parts are split from their skeleton branches. By this way a high level of segmentation result is obtained. The authors report that after getting road parts by this method, there exist gaps between some of these road parts. To satisfy connectivity constraint of the road network, these gaps are filled if the neighbor road parts have similar direction and pose continuation smoothness. This method is capable of extracting the road network showing color and width homogeneity. The same researcher team expands this work in another paper (Grote A. , Heipke, Rottensteiner, & Meyer, 2009) by applying a more robust gap filling method to the existing results. In this method, road parts are grouped into a higher level category, namely road sub-graphs. When forming a road sub-graph, straight continuation of the given road part is the main constraint. A linear programming evaluation technique, which is restricted by pre-defined rules, is used to find the best partition for the road sub-graph. In this methodology, each gap edge is scored by considering some features including distance, road part quality, color, width, smoothness and direction. For the purpose of determination of the gap weights, the authors also introduce the context objects which are vehicles, trees, vegetated areas and non-road asphalt areas. As indicated in the paper sub-graph extraction and evaluation method clearly improves the former results.

In the study (Maurya, Gupta, & Shukla, 2011), a simple road extraction algorithm consisting of k-means clustering followed by variety of morphological operations is described. At the first stage, spectral features, Red, Green and Blue bands, of the input image are clustered into three classes by K-means algorithm. One cluster having road regions is manually selected as road cluster. Then, non-road regions are eliminated by considering connected components' structural properties. Having disregarded the possible different classes can be found in a satellite image, this study shows good performance in their dataset.

In paper (Zhang & Couloigner, 2006) , an unsupervised road extraction strategy is used to extract road regions from high resolution satellite image, and the proposed approach for road and parking lot discrimination problem is elaborated. Road extraction procedure is based on k-means clustering and fuzzy logic classifier which is capable of automatically identifying road cluster. After road detection, a series of shape descriptors for excluding parking lot regions from road mask is experimented. These descriptors are based on angular texture signatures using several segment features such as mean, compactness, eccentricity and direction. The proposed approach is successful at road class refinement by the help of angular texture signatures.

In the research conducted by (Bong, Lai, & Joseph, 2009), a hybrid simple color space segmentation and edge detection algorithm is proposed. Before actual extraction procedure, the input image’s color and edge features are analyzed and a general lookup threshold table regarding to color values is obtained. In this analysis, it is assumed that a remotely sensed image can be divided into five categories; trees, roads, buildings, sandy regions and water regions. Along with spectral threshold values such as luminance, saturation, hue and gray value, edge information is also fused into the final model to get an improved solution. Although this method ignores the fact that different land covers result in different spectral values, it gives good performance on the test dataset.

2.3 Spectral Band Ratios

In remote sensing, spectral band rationing is a commonly used technique to differentiate certain kinds of objects from background. That is, band ratios provide essential information for the monitoring of several different land change processes such as vegetation health and status, burned area and fire severity. Illumination differences caused by shadows and surface slope directions are common in satellite imagery (Smith, 2012). Therefore, the gray-scale value of a material can vary from place to place in the image. To analyze the materials in the remotely sensed imagery, these unwanted effects should be removed. Band ratios exploit the spectral differences found in different bands of the image. Since shading and shadowing contributions to the image are approximately constant, division of two band values cancels them out. By this way, a robust indicator of interested material can be obtained.

Band ratios are common techniques to detect water, vegetation and minerals in remote sensing. According to the purpose of the application, one can choose to attenuate the occurrence of a particular material. The band in which the material is highly reflective and another band in which the material posing strong absorption properties are suitable to construct band ratios. The numerator of the ratio is occupied by the band having high reflectivity. In this study, three different band ratios are examined to provide spectral awareness to the final classifier. The spectral features are revealed by using Normalized Difference Vegetation Index (NDVI), Enhanced Vegetation Index (EVI), Soil Adjusted Vegetation Index (SAVI) and Normalized Difference Water Index (NDWI).

NDVI is defined in equation (2.1). NDVI is a simple band ratio utilized for plant growth monitoring and vegetation detection. It is known that NDVI values range between -1 to 1, and lower negative values imply water regions while the regions like bare soil and asphalt have values close to zero. Also vegetative regions in the images are represented in brighter scales.

$$NDVI = \frac{NIR - Red}{NIR + Red} \tag{2.1}$$

Like NDVI, EVI is also useful for vegetation detection. In fact, EVI is designed to provide good vegetation estimations in addition to reducing atmospheric artifacts. In the equation (2.2), the calculation of EVI is given. C1 and C2 are the aerosol resistance coefficients which are eliminating the aerosol effects from the red band by using blue band.

$$EVI = \frac{NIR - Red}{NIR + C1xRed - C2xBlue + 1} \tag{2.2}$$

SAVI is useful when the vegetative cover is relatively low (i.e. < 40%). Indeed, SAVI is an extension to the NDVI to adjust the effects caused by soil brightness. In the formulation given in (2.3), L parameter is used to control the soil brightness correction factor.

$$SAVI = \frac{NIR - Red}{NIR + Red + L} * (L + 1) \tag{2.3}$$

On the other hand, NDWI is utilized for water detection in many remote sensing applications, and it is designed to detect wet regions in the image. NDWI is defined in equation (2.4).

$$NDWI \ = \frac{Green - NIR}{Green + NIR} \tag{2.4}$$

2.4 Multi-Resolution Analysis

For multi-resolution analysis, Gaussian pyramid reduction and expansion method coined in (Burt & Adelson, 1983) is employed. Pyramid representation is a commonly used multi-resolution analysis technique that is subject to repetitive smoothing and down-sampling processes. The pyramid model consists of an arbitrary number of levels where level zero equals to the input image. By increasing the number of levels, images get smaller in size and smoothed. Formally, let $I(x, y)$ be the original image, the reduction process of the Gaussian pyramid can be formulized in equation (2.6), and the input image corresponding to level-0 is defined in (2.5).

$$G_o(x, y) = I(x, y) \tag{2.5}$$

$$G_l(x,y) = \sum_{m=-2}^{2} \sum\nolimits_{n=-2}^{2} w(m,n)\, G_{l-1}(2x+m, 2y+n) \tag{2.6}$$

where $w(m,n)$ is the weighting function which is applied to all levels. The weighting function, called the generating kernel, is a 5-by-5 window having construction weights for level-to-level averaging process. The generating kernel holds the properties (2.7), (2.8) and (2.9) to satisfy certain constraints (Burt, Fast filter transform for image processing, 1981);

$$Separability; \quad w(m,n) = w_1(m)w_2(n) \tag{2.7}$$

$$Normalized; \sum\nolimits_{m=2}^{2} w(m) = 1 \tag{2.8}$$

$$Symmetric; \ w(i) = w(-i) \tag{2.9}$$

Expansion process is the reverse of reduction: this process consists of a smoothing followed by an up-sampling operation. Expansion operation is defined in equation (2.10).

$$G_{l,n}(x,y) = \sum_{m=-2}^{2} \sum\nolimits_{n=-2}^{2} w(m,n)\, G_{l,n-1}(\frac{x-m}{2} \frac{y-n}{2}) \tag{2.10}$$

where $G_{l,n}$ is the result of expanding G_l n times.

2.5 Elongatedness Analysis

The model proposed in this study is constructed on the elongatedness assumption of roads. Therefore, a routine for elongatedness calculation is needed. Even though, elongatedness is defined as the ratio of long and short side lengths of a rectangle, almost all segments are far away from being a perfect rectangle in practice, that is; the segments obtained by particular segmentation algorithms have complicated shapes due to scenes' complexity. In this thesis, we use the elongatedness index firstly used in the study (Karaman, Cinar, Gedik, Yardimci, & Halici, 2012).

The elongatedness is based on the ratio of Major Axis Length and Minor Axis Length and can be defined as in equation (2.11).

$$Elongatedness = \frac{Ideal\ MajorAxisLength}{Ideal\ MinorAxisLength} \tag{2.11}$$

Major axis of a segment is defined as the major axis of the ellipse having the same normalized second central moments as the segment (Haralick & Shapiro, 1992). Minor axis is also considered as the fitting ellipse's minor axis. In Figure 2, an illustration of a segment and its corresponding ellipse can be seen. The solid blue lines indicate the ellipse axes and the red dots are the foci of ellipse.

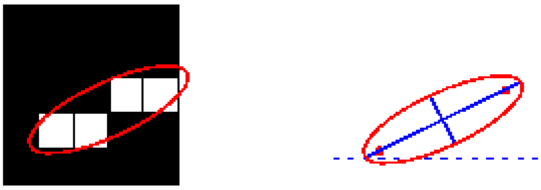


Figure 2 : An example segment and its ellipse, adapted from (MathWorks, 2012)

Approximation for major axis length can be made by using extent property of the segment. Extent is a scalar specifying the ratio of area of the segment to the area of bounding box, defined in (2.12).

$$Extent \ = \frac{Area\ of\ segment}{Area\ of\ bounding\ box} \tag{2.12}$$

Extent indicates the rate of vacancy in the bounding box; hence it describes how far the segment is from an ideal rectangle structure. Thus, we can combine extent with major axis length to obtain an approximation to its ideal major axis length as in (2.13).

$$Ideal\ MajorAxisLength = \ MajorAxisLength\ x\ (2 - Extent) \tag{2.13}$$

Then, ideal major axis length can be rewritten as in (2.14).

$$Ideal\ MinorAxisLength = \frac{Area}{Ideal\ MajorAxisLength} \tag{2.14}$$

Thus, we can define the elongatedness property with the equation (2.15).

$$Elongatedness \ = \frac{[MajorAxisLength\ x\ (2 - Extent)]^2}{Area} \tag{2.15}$$

2.6 Mean-shift

Mean-shift is a non-parametric feature space analysis algorithm firstly proposed by (Fukunaga & Hostetler, 1975). Afterwards, (Cheng , 1995) has formulated the mean-shift procedure and further investigated the versatilities of this algorithm. Convergence property for discrete data of mean-shift is proved in the study (Comaniciu & Meer, 2002), and a robust image segmentation algorithm for low-level vision tasks based on mean-shift theory is presented. The main objective of mean-shift is to find local maxima points (modes) of a density function assuming that given discrete data is sampled from that function. To estimate the density function, mean-shift makes use of kernel functions which is responsible for determining the weights of nearby points for next iteration.

The Parzen window technique presented in (Parzen, 1962) is the kernel estimation method used in mean-shift to estimate the probability density function for a given set of random variables. The multivariate kernel density estimation function with the kernel $K(x)$ and the dxd bandwidth matrix H can be formulized for given n data points $x_i, i = 1, \dots, n \; in \; the \; d - dimensional \; space \; R^d$ defined in (2.16);

$$f(x) = \frac{1}{n}\sum_{i=1}^{n} K_H(x - x_i) \tag{2.16}$$

where

$$K_H(x) = |H|^{-\frac{1}{2}} K(H^{-1/2}x) \tag{2.17}$$

The kernel function $K(x)$ is d-variate and satisfies the constraints given in (2.18), (2.19), (2.20) and (2.21);

$$\int_{R^d} K(x)dx = 1 \tag{2.18}$$

$$\int_{R^d} xK(x)dx = 0 \tag{2.19}$$

$$\lim_{\|x\| \to \infty} \|x\|^d K(x) = 0 \tag{2.20}$$

$$\int_{R^d} xx^T K(x)dx = c_k I \tag{2.21}$$

where c_k is a constant. Generalizing these constraints with the need of radially symmetric kernels, the final density estimation function is defined as follows; (Comaniciu & Meer, 2002)

$$f_{h,K}(x) = \frac{c_{k,d}}{nh^d}\sum_{i=1}^{n} k(\left\|\frac{x-x_i}{h}\right\|^2) \tag{2.22}$$

where h is the bandwidth parameter specifying the size of search window. Having defined the density estimator, we also need to define density gradient estimation procedure since the modes are the points having zero gradient. By using linearity of (2.22), density gradient estimator can be written as in (2.23)

$$\nabla f_{h,K}(x) = \frac{2c_{k,d}}{nh^{d+2}}\left[\sum_{i=1}^{n} g(\left\|\frac{x-x_i}{h}\right\|^2)\right]\left[\frac{\sum_{i=1}^{n} x_i g(\left\|\frac{x-x_i}{h}\right\|^2)}{\sum_{i=1}^{n} g(\left\|\frac{x-x_i}{h}\right\|^2)} - x\right] \tag{2.23}$$

Where $g(x)$ is the kernel profile. From this definition, we can infer that first term implies density estimate at x computed with the kernel G defined in (2.24)

$$f_{h,G}(x) = \frac{c_{k,d}}{nh^{d+2}}\left[\sum_{i=1}^{n} g(\left\|\frac{x-x_i}{h}\right\|^2)\right] \tag{2.24}$$

The second term indicates the mean shift which is defined as the difference between weighted mean. The mean shift formulation is given in (2.25)

$$m_{h,G}(x) = \frac{\sum_{i=1}^{n} x_i g(\left\|\frac{x-x_i}{h}\right\|^2)}{\sum_{i=1}^{n} g(\left\|\frac{x-x_i}{h}\right\|^2)} - x \tag{2.25}$$

Arranging these formulas into more simple form yields (2.26);

$$m_{h,G}(x) = \frac{1}{2}h^2 c\,\frac{\nabla f_{h,K}(x)}{\nabla f_{h,G}(x)} \tag{2.26}$$

This final formula states that the direction of maximum increase in the density is indicated by the mean shift vector. (Comaniciu & Meer, 2002) In other words, the local mean is always moved towards to the majority of the data points. The path to stationary points of the estimated density is guaranteed by the mean shift vector's alignment with the local gradient estimate. Indeed, these stationary points are the modes we are seeking. At each iteration of the mean-shift algorithm, the mean shift vector $m_{h,G}(x)$ is computed and the kernel $G(x)$ is moved by $m_{h,G}(x)$. Moreover, in the regions having low density of data points, the mean shift steps are relatively large. Likewise, when the search window is close to the mode, the steps are small. By considering these properties, we can declare mean-shift algorithm as an adaptive gradient ascent method.

2.6.1 Mean-shift Image Segmentation

In the context of this study, the mean-shift image segmentation implementation suggested by (Comaniciu & Meer, 2002) is employed. In this method, it is stated that an image is composed of two dimensional lattice of d-dimensional vectors (pixels). The space of the lattice represents spatial domain while the spectral information about the pixels form the range domain. Therefore, for image segmentation, we must gather spatial and range features to form a more general feature space representation. Indeed, this composition is named as joint spatial-range domain. Fortunately, for each domain, separate bandwidth parameters can be defined in the multivariate kernel such as (2.27).

$$K_{h_s,h_r}(x) = \frac{C}{{h_s}^2 {h_r}^d}\, k\left(\left\|\frac{x^s}{h_s}\right\|^2\right) k\left(\left\|\frac{x^r}{h_r}\right\|^2\right) \tag{2.27}$$

Where x^s is the vector containing spatial features and x^r is the vector containing range features, k(x) is the common kernel profile used in both of the domains, C is the normalization constant, h_s and h_r are spectral and range parameters respectively. The resolution of the mode detection is controlled by user defined bandwidth parameter $h = (h_{s,} h_r)$. In this implementation, there is also an optional parameter named as “minimum region” indicating minimum area of output segments. For given n data points $x_i, i = 1, \dots, n$ in joint spatial-range domain, the overall flow of the image segmentation algorithm is as follows;

- Initialize iteration number, $j = 1$. The point of convergence is given as, $y_{i,1} = x_i$
- Find the point of convergence $y_{i,j+1}$, $y = y_{i,c}$
- Assign $z_i = ({x_i}^s, {y_{i,c}}^r)$
- Construct image segments by grouping z_i , concatenate the basins of attraction of the convergence points.
- For each $i = 1, \dots, n$ assign label numbers
- Optionally, regions with less than M pixels are joined to closest neighbors.

The superscripts indicate the spatial and range components of feature vector, respectively. M is the user defined parameter donating minimum region of segments.

Additionally, we have used L*u*v color space for segmentation in this study. As studied in (Comaniciu & Meer, 1999) L*u*v color space is the optimally fitting feature space for mean-shift color image segmentation as it provides robust approximation to Euclidean metrics.

2.7 Adaboost

Adaptive Boosting (Adaboost), is an ensemble based machine learning method, proposed by (Freund & Schapire , 1997). The concept of adaptive boosting stems from PAC (Probably Approximately Correct) learning model proposed by (Valiant, 1984). The PAC learning model defines a strong classifier as a classifier produces an arbitrarily small test error for arbitrarily sampled random examples from any distribution with a high probability. On the other hand, weak classifiers hold same properties except having an error rate of $\varepsilon_t \geq 0.5 - \gamma$, where γ is an arbirarly small scalar. Indeed, weak learners are the classifiers having slightly better error rate than random classifiers.

The concept of boosting implies a weighted majority (Littlestone & Warmuth , 1994) based voting scheme for the problem of dynamically apportioning resources among a set of options. The boosting algorithms aim to construct a very accurate prediction rule by using combination of rough and relatively inaccurately rule-of-thumbs. The booster iteratively creates a distribution D_t over the training examples and requests a weak hypothesis (or rule-of-thumb) h_t with low error ε_t with respect to D_t from an unspecified oracle. By this way, the learning mechanism concentrates on incorrectly predicted examples at each round. The final problem of boosting is to combine weak hypothesis to devise a single strong prediction rule.

Freund et al. show that weak learning algorithms that perform just slightly better than random guessing can be combined to obtain a strong classifier with arbitrarily high accuracy. Adaboost tries to build a strong classifier by iteratively adding a weak classifier into the ensemble. Adaboost can be used in conjunction with several types of weak classifiers to improve their individual performances. Adaboost does not require prior knowledge of error rates of the weak classifiers; instead, it uses posterior error rates to generate a highly accurate weighted majority hypothesis. In this model, weight of the each weak classifier is defined as a function of its accuracy. Another important outcome of this model is that unlike previous boosting algorithms' performance bounds (Schapire, 1990) which depend on the accuracy of the least accurate weak hypothesis, the Adaboost algorithm's performance can be improved by any improvement on the combined weak hypothesis.

In the context of this thesis, Adaboost is used to classify road and non-road regions. Detailed information about theoretical background of Adaboost algorithm is presented in this section.

2.7.1 Formulation of Adaboost

```
Input:
    • Sequence of N examples $(x_1,y_1)$…… $(x_N,y_N)$ with class labels $y_i \in Y = \{1, \dots, k\}$
    • Distribution D over the examples
    • Weak classifier to be trained, WeakLearn
    • Number of iterations, T
Initialize:
    • The weight vector $w_i^1 = D(i)\ for\ i = 1 \dots N$
Do for t=1 to T
    1. Set
           $\boldsymbol{p}^t = \frac{\boldsymbol{w}^t}{\sum_{i=1}^{N} w_i^t}$
    2. Construct a WeakLearn based on $\boldsymbol{p}^t$; get back a hypothesis $h_t : X \rightarrow Y$
    3. Calculate the error of
           $h_t: \varepsilon_t = \sum_{i=1}^{N} p_i{}^t [h_t(x_i) - y_i]$
    4. Set$\beta_t = \frac{\varepsilon_t}{1 - \varepsilon_t}$.
    5. Set the new weights vector to be
           $w_i^{t+1} = w_i^t \beta_t^{1-[h_t(x_i) - y_i]}$
Output the predicted classification score function
    $h_f(x) = \sum_{t=1}^{T} a_t\, h_t(x_i)$  where  $a_t = \frac{1}{2}\left(\log \frac{1}{\beta_t}\right)$
Output the predicted class labels
    $hl_f(x) = \begin{cases} 1 & if\ h_f(x) \geq 0.5 \\ 0 & otherwise \end{cases}$
```

Figure 3: Adaboost formulation

The general formulation of the Adaboost algorithm is given in Figure 3. The algorithm is provided with a set of labeled training examples (x_1,y_1)…… (x_N,y_N), where y_i is the label associated with the instance x_i. Distribution D is created over the example set, which specifies a weight value for each example in the training set. The algorithm maintains a weight vector whose value at time t is shown $w_t = \langle w_1^t, \dots w_N^t \rangle$ during the execution.

Initially, each example has the same importance and sum of the weight vector equals to one, so each value in the weight vector w_1^t equals to 1/N, where N is the number of examples. Beginning of the each iteration $\boldsymbol{p}^t$ is computed by normalizing $\boldsymbol{w}^t$. Then, weight distribution $\boldsymbol{p}^t$ is given to the WeakLearn as an input. WeakLearn is responsible for generating a hypothesis $\mathbf{h}_t$ having a small error with respect to distribution. Having constructed weak classifier whose performance is better than a random classifier, weight vector is updated according to classification results. For this work, firstly, error rate of the current weak classifier $\mathbf{h}_t$ is calculated. According to this error value, weights of the misclassified examples are increased whereas weights of the correctly classified examples are decreased. By this way, at the next iteration, the algorithm concentrates on the hard examples, in other words, examples having large weight values.

2.7.2 Theoretical Properties of Adaboost

Freund et al. proves two important properties for our study in their article.

2.7.2.1 Bound on Training Error

First property is that the training error of the Adaboost is bounded by the equation (2.28).

$$\prod_t \mathrm{Z}_t = \prod_t 2\sqrt{\varepsilon_t(1-\varepsilon_t)} = \prod_t \sqrt{1-4{\gamma_t}^2} \leq e^{-2\sum_t {\gamma_t}^2} \qquad (2.28)$$

This property implies that if each weak classifier is at least slightly better than random guessing by $\gamma_t \geq \gamma$ for some $\gamma > 0$, then the training error drops exponentially fast in T iterations. According to the theory the training error upper bound can be at most $e^{-2T\gamma^2}$.

2.7.2.2 Bound on Generalization Error

The second property proved by Freund et al. defines a bound for the generalization error. This property implies that the VC-dimension of $\Theta_T(H)$ is at most $2(d+1)(T+1)log_2(e(T+1))$, where H is the class of binary functions of VC-dimension $d \geq 2$, T is the iteration number, and H is the class of binary functions over some domain X. Thus, if classifiers constructed by WeakLearn are selected from a class of VC-dimension $d \geq 2$, then Adaboost generates the final hypothesis after T iterations belong to a class of VC-dimension at most $2(d+1)(T+1)log_2(e(T+1))$. Therefore, the proved bound suggest that Adaboost algorithm will overfit if a large T is chosen.

With a high number of T, the learning model will begin to memorize the set of training rather than learning it. This problem is called as “over-fitting” in the machine learning community, and results in hypothesis posing good performance on the training set but not on new data points from test set. Nevertheless, the experiments conducted on real world problems by Freund et al. show that Adaboost often does not seem to overfit. On the other hand, an empirical research (Hernández-Lobato, Martínez-Muñoz, & Suárez, 2011) conducted on evaluation of several ensemble methods shows that Adaboost is not likely to be overfitted but it can be sensitive to overfitting for some problems (see Figure 4).

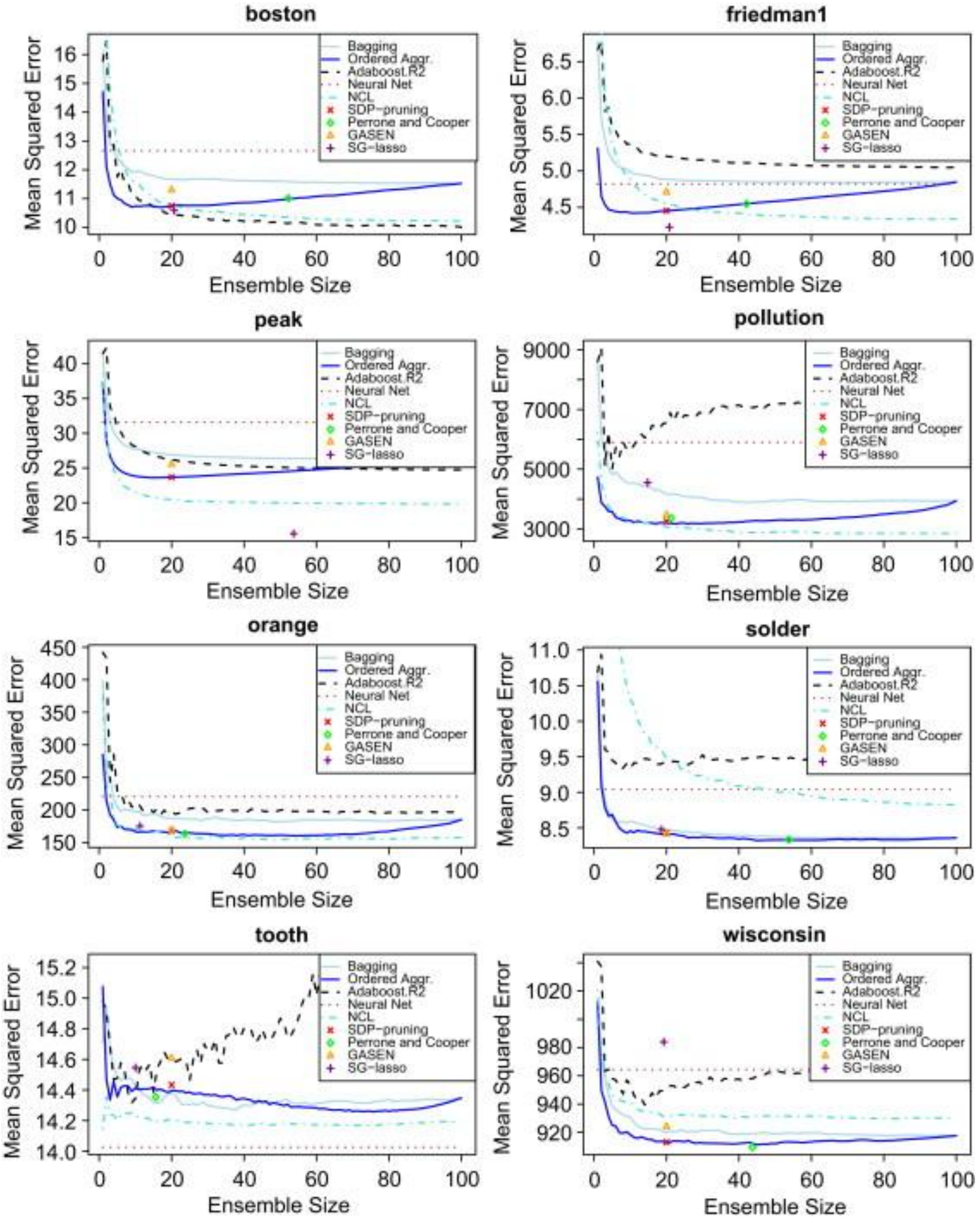


Figure 4:Comparision of some ensamble learning strategies including Adaboost

Figure 4 contains the behaviors of the ensemble methods including bagging, ordered aggregation, negative correlation learning (NCL) and Adaboost.R2. It can be seen that for a few datasets, Adaboost does not suffer from overfitting while other algorithms do. Also for some datasets, Adaboost shows extreme vulnerability to overfitting.

2.7.2.3 Margin Theory of Adaboost

The third important property of Adaboost is proved in the study (Schapire, Freund, Bartlett, & Lee, 1998). The authors show another upper bound based on margins defined as the difference between the number of correct votes and the maximum number of votes given to any incorrect label. They also study the phenomena of not overfitting by increasing ensemble size. They relate these phenomena with the distribution of margins of the training examples.

2.7.2.4 Bayesian Interpretation of Adaboost

There is an important relationship between Bayesian analysis and the final classifier indicated by Adaboost. For a given set of example distribution P on X x $\{0,1\}$ and a set of weak hypothesis $h_{1,} \dots h_T$ generated in T iterations, we want to combine them in an optimal way. According to Bayes decision principle, the label with the highest likelihood must be predicted as true class, yielding equation (2.29)

$$P_r\big(y = 1 \big| h_1(x), \dots h_T(x)\big) > P_r\big(y = 0 \big| h_1(x), \dots h_T(x)\big) \tag{2.29}$$

Assuming the different hypotheses are independent of one another, we can rewrite this decision rule in a simpler form;

$$P_r[y = 1] \prod_{t:\, h_t(x)=0} \varepsilon_t \prod_{t:\, h_t(x)=1} (1 - \varepsilon_t) > P_r[y = 0] \prod_{t:\, h_t(x)=0} (1 - \varepsilon_t) \prod_{t:\, h_t(x)=1} \varepsilon_t \tag{2.30}$$

Note that we can replace $P_r[y = 0]$ with ε_0. Taking the logarithm of both sides and rearranging the terms result in the same decision rule generated by Adaboost. This demonstrates that Adaboost is compatible with Bayesian analysis.

2.7.2.5 Improvement on Error Bound

Freund et al. suggest a method to improve error bound of the Adaboost in their study. They replace the "hard" threshold $\{0,1\}$ used in the final decision rule h_f by a soft threshold. More precisely, let us define the final strong hypothesis by $h_f(x_i) = F(r(x_i))$ and a weighted average of weak hypotheses h_t as in (2.31);

$$r(x_i) = \frac{\sum_{t=1}^{T} \left(log \frac{1}{\beta_t}\right) h_t(x_i)}{\sum_{t=1}^{T} \left(log \frac{1}{\beta_t}\right)} \tag{2.31}$$

F is defined as $F: [0,1] \rightarrow [0,1]$, so our soft threshold values can be selected from the range $[0,1]$. In this interpretation, the final hypothesis h_f is considered a randomized hypothesis and $h_f(x_i)$ corresponds the probability of predicting 1. Thus the probability of an incorrect classification contributes on the error bound defined as $E_{i \sim D}\left[h_f(x_i) - y_i\right]$.

2.8 Decision Trees

Decision tree learning, firstly formulated by (Quinlan , 1986), is a predictive model mapping observations to the classification labels. In decision trees, leaves represent category labels and branches imply conjunctions of feature components leading to resulting prediction. An illustrated sketch of a decision tree can be seen in Figure 5.

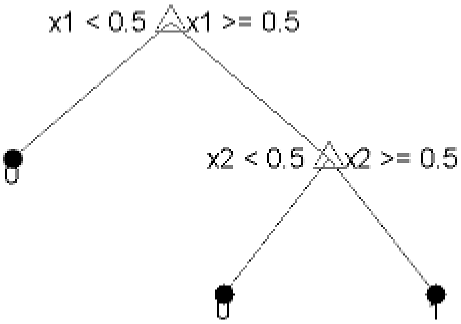


Figure 5: An illustration of desicion tree

In this example, branch nodes have decision statements, and the tree is capable of predicting based on two predictors x1 and x2. The prediction routine begins from the top node and interprets the decision statement at that node. If the decision statement is true, right branch is followed; otherwise, left branch is followed. The procedure continues in this way evaluating the decision statements found in the nodes until reaching a leaf node. The leaf node presents the decision for classification.

In this study, Classification and Regression Trees (CART), proposed by (Breiman, Friedman, Stone, & Olshen, 1984), are employed as the weak classifiers. Although CART implies two distinct groups of weak classifiers, classification trees producing class label as a result of evaluation is used.

Decision tree construction procedure relies on the splitting rule. Training samples are divided into smaller partitions at each split of the decision tree, and this process is controlled by the splitting rule. Maximum homogeneity of a tree node is the metric used to determine whether to continue with separation on the data.

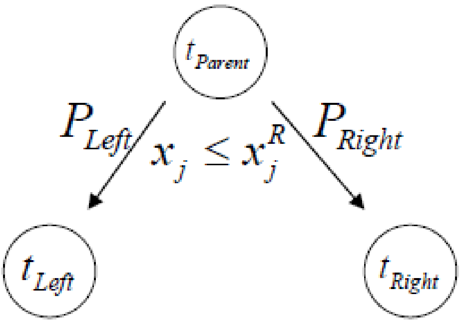


Figure 6:Decision procedure of a desicion tree

In an example sketch of a decision tree in Figure 6, t_p, t_l, t_r, are the parent, left and right children nodes, respectively. x_j is the predictor, and x_j^R is the best splitting value of the predictor x_j . The homogeneity of child nodes is calculated by using an impurity function which is the indicator of a node's purity degree. As impurity of the parent is not affected by the decision of splitting rule searching for the best split $x_j \leq x_j^R, j = 1, \dots, M$, the homogeneity of left and right nodes can be achieved by maximizing difference of impurity function $\Delta i(t)$ defined as (2.32)

$$\Delta i(t) = i(t_p) - E[i(t_c)] \tag{2.32}$$

where t_c represents children nodes, and t_p is the parent node. Let P_l and P_r be the probabilities of following the search routes to the left and right, respectively, yield (2.33)

$$\Delta i(t) = i(t_p) - P_l i(t_l) - P_r i(t_r) \tag{2.33}$$

Thus, the maximation problem given in (2.34) is subjected to be solved by the each node of CART classifier.

$$\text{argmax}_{x_j \leq x_j^R}[i(t_p) - P_l i(t_l) - P_r i(t_r)] \tag{2.34}$$

In fact, this equation indicates the training samples are scanned by the algorithm to find the best split satisfying inequality $x_j \leq x_j^R$ from the pool of all alternative splits.

2.9 Histogram Equalization

Histogram equalization is a common contrast enhancement tool used in image processing (Gonzalez & Woods, 2008). Histogram equalization is suitable for increasing global contrast of an image. That is, it stretches the histogram of an image providing a more uniform distribution.

Histogram of an image can be viewed as a probability density function assuming its values are random variables in the interval $[0, L-1]$. Thus, letting $T(r)$ is the transformation function which is continuous and differentiable in the range $[0, L-1]$, r and s denote input and output images, $p_r(r)$ and $p_s(s)$ to show the probability density functions of r and s, respectively, we obtain equation (2.35).

$$p_s(s) = p_r(r)\left|\frac{dr}{ds}\right| \tag{2.35}$$

Since the derivative of a definite integral with respect to its upper limit is the integrand computed at the limit, we can write as in equation (2.36).

$$\left|\frac{dr}{ds}\right| = \frac{dT(r)}{dr} = (L-1)p_r(r) \tag{2.36}$$

and by substitution;

$$p_s(s) = \frac{1}{L-1} \tag{2.37}$$

This result shows that the histogram equalization operation flattens the input histogram by converging to the rectangle function. The discrete representation of histogram equalization proved in (Gonzalez & Woods, 2008), as in (2.38).

$$s_k = \frac{(L-1)}{MN}\sum_{j=0}^{k} n_j \ , k = 0,1,2, \dots L-1 \tag{2.38}$$

where MN is the total number of pixels exist in the image, n_j is the number of pixels having the value of r_{kj}, and L is the number of possible intensity levels.

2.10 Region growing

Region growing is a pixel based semi-automatic image segmentation algorithm (Gonzalez & Woods, 2008) which is generally used to generate segments having spatially continuous pixels. Region growing algorithm starts with a seed pixel which is an arbitrarily relevant representation of target segment on account of its intensity. The segmentation process goes iteratively by aggregating spectrally similar neighbor pixels in all directions.

(Adams & Bischof, 1994) have formalized the mechanisms of the region growing algorithm by considering set theory. For a given set of seeds $S_1, \dots S_q$, homogeneous regions $R_1, \dots R_q$ are constructed iteratively. As the procedure is continuous, pixels assigned to one of these regions are called as allocated pixels and the remaining ones are named as unallocated pixels. Let H be the set of all unallocated pixels which are in the neighborhood of at least one of the homogeneous regions, we define an equation (2.39).

$$H = \left\{ (x,y) \notin \bigcup_{i=1}^{q} R_i \mid N(x,y) \cap \bigcup_{i=1}^{q} R_i \neq \varnothing \right\} \tag{2.39}$$

where $N(x,y)$ is the second order neighborhood of the pixel location (x,y).

If the pixel (x,y) is currently tested point from the set of unallocated pixels, then the difference between the testing pixel and its adjacent region R_i can be stated as in (2.40).

$$\delta(x,y,R_i) = |g(x,y) - g(X_i^c, Y_i^c)| \tag{2.40}$$

Where $g(x,y)$ is the gray value of the pixel (x,y), and $g(X_i^c, Y_i^c)$ is the average gray value of region R_i centered at (X_i^c, Y_i^c). Thus, the pixel location (x,y) which is minimizing the objective function $\delta(x,y,R_i)$ allocated to the region R_i, in other words, $\varphi(x,y) \in N(x,y)$ is the formulation of whole process defined as (2.41);

$$\varphi(x,y) = min\{\delta(x,y,R_i) | j \in \{1, \dots q\}\} \tag{2.41}$$

According to the equation (3.40), the input image is segmented into homogenous regions consisted of integrated set of neighbor pixels.

2.11 Pattern search

The term, pattern search, is firstly coined by (Hooke & Jeeves, 1961), and the pattern search algorithm is generalized in the study (Torczon V. , 1997) for derivative-free unconstrained optimization on continuously differentiable functions with positive spanning directions. Later on, a simpler analysis on pattern search optimization is presented in (Audet, Charles , & Dennis, 2003) providing theoretical background on the relationship between the optimization process and the search directions.

The optimization problem can be formulated as follows;

$$\min_{x \in \Omega} f(x) \ , where\ f: R^n \to R \ \cup \{\infty\} \tag{2.42}$$

$$\text{and } \Omega = \{x \in R^n : \ell \ \leq \text{Ax} \leq \text{u}\} \text{, where A} \in \ Q^{mxn} \tag{2.43}$$

$$and\ \ell, u \in \{R \ \cup \{-\infty, +\infty\}\}^m \ provided\ that\ \ell < \text{u} \tag{2.44}$$

In this formulation, $f(x)$ is the given objective function, Ω represents any feasible solution, ℓ, u are lower and upper bounds provided by user.

(Torczon V. , 1997) proposes a barrier approach where the barrier function defined as $f_\Omega(x) = f + \psi_\Omega$ is optimized instead of $f(x)$, where ψ_Ω is the indicator function for Ω, that is; ψ_Ω has zero value on Ω and equal to ∞ elsewhere.

$$f_\Omega(x) = \begin{cases} f(x) & if\ x \in \Omega \\ \infty & otherwise \end{cases} \tag{2.45}$$

The pattern search algorithm generates a sequence of iterates $\{x_n\} \in R^n$, and it achieves non-increasing objective function values at each optimization step divided in two stages including optional SEARCH and a local POLL routines.

At the SEARCH step, the algorithm begins to search a better space point further minimizing the barrier objective function. A finite number of points on a mesh defined as a discrete subset of R^n is used to evaluate the objective function. If a better objective function value is found during the SEARCH procedure, i.e. $f_\Omega(x_{k+1}) < \ f_\Omega(x_k)$, then x_{k+1} is said to be an improved mesh point. Otherwise, the optimization procedure continues by invoking the POLL routine which seeks an optimum solution in the neighborhood of the mesh. If POLL routine cannot also provide an improved mesh point, then the current best solution is said to be a mesh local optimizer, in other words, the current mesh does not provide solutions better than its neighbor meshes. Then, the mesh size parameter Δ_k at the iteration k is updated by using below equation (3.45).

$$\Delta_{k+1} = \tau^{wk} \Delta_k \tag{2.46}$$

Where τ is a user-defined constant greater than one, and $wk \leq -1$.

Therefore, the mesh centered around the current iterate $\{x_k\} \in R^n$ at iteration k can be defined as (3.46);

$$M_k = \{x_k + \Delta_k D_z : z \in Z^+\} \tag{2.47}$$

where Z^+ is the set of non-negative integers, and D is a set of positive spanning directions in R^n.

At each iteration, POLL routine is given a distinct positive spanning matrix set D_k formed by columns of D, that is; $D_k \subseteq D$. The directions of the columns of D_k point the neighbor meshes of current iterate x_k.

$$POLL\ set : \{x_k + \Delta_k d : d \in D_k\}$$

For non-linear optimization problems, the global convergence property of the pattern search algorithm is proved in (Torczon V. , 1995), and the proof implies that the gradient of the problem is not needed to reach the global optimum.

CHAPTER 3

METHODOLOGY

In this chapter, specific components guiding our system for automatically extracting roads from remotely sensed images are explained. We have proposed a new road extraction method, and the code is written in MATLAB and C++. This chapter is devoted to fully describe the practical work behind the study and it is arranged accordingly. Qualitative research conducted on the problem is elaborated in six distinct sub-sections; data, introduction, feature extraction, mean-shift segmentation parameter optimization, classification with Adaboost and road extraction. Firstly, the data used in this research is stated, then we introduce the algorithm in introduction part. After the technique used to extract features required for classification is detailed in section 3.3, we described the parameter optimization method for mean-shift segmentation. Finally, classification with Adaboost and road extraction sections show the main principles employed for road extraction from satellite imagery.

3.1 Data

In this study, high resolution multi-spectral electro-optic satellite images taken from IKONOS-2 device are employed for algorithm development and experimentation purposes. IKONOS-2 satellite is one of the world's leading fine resolution remote sensing image provider craft operated by Geo Eye (Space Imaging, 2000). IKONOS-2 provides both urban and rural mappings for a broad purpose of study such as natural disasters, tax mapping, agriculture and forestry analysis, mining, engineering and change detection. In the context of this study, we have made use of IKONOS imagery for the purpose of road network detection. Properties of these images are summarized in Table 1.

Table 1:Properties of the image data

Spatial Resolution	• **0.8 m panchromatic (1-m PAN)** • **4-meter multispectral (4-m MS)** • **1-meter pan-sharpened (1-m PS)**	
Spectral Resolution	Band	4-m MS & 1-m PS
	1 (Blue)	0.445–0.516 µm
	2 (Green)	0.506–0.595 µm
	3 (Red)	0.632–0.698 µm
	4 (Near IR)	0.757–0.853 µm
Temporal Resolution	Three to five days off-nadir and 144 days for true nadir	
Radiometric Resolution	11-bit (values in range 0-2047)	

Satellite imagery devices can detect certain regions of the electromagnetic spectrum, and this property is referred as spectral resolution which describes the interval of the electromagnetic spectrum an instrument is capable of sensing. In the study (Arledge & Hatcher, 2008), spectral response curves for IKONOS data is illustrated as in the figure.

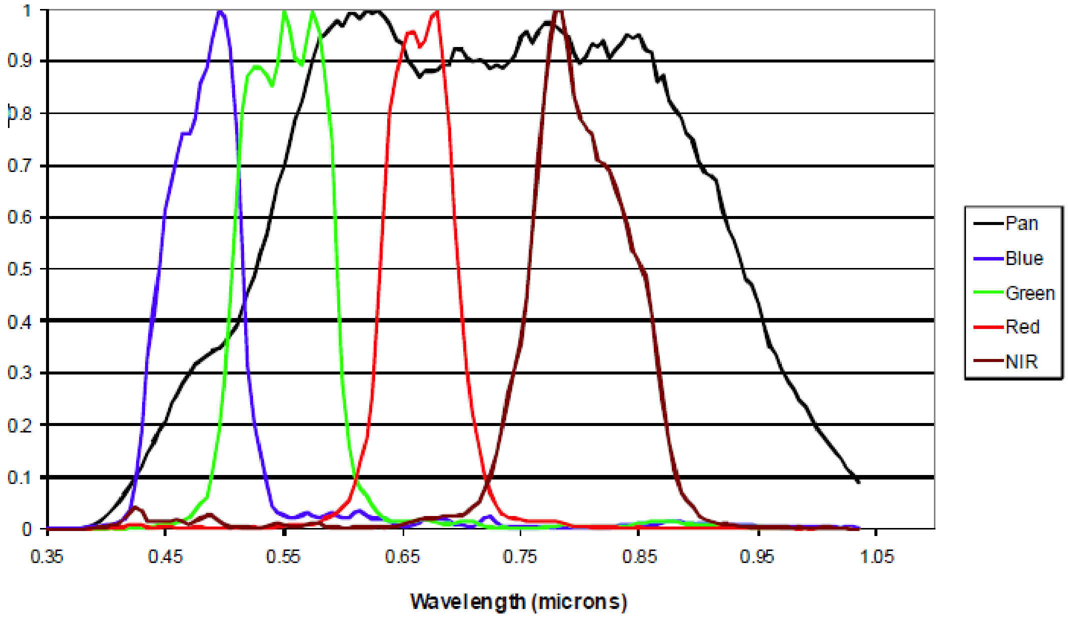


Figure 7:Spectral response curves of IKONOS data

We have 12 images in the train set and 25 images in the test set, all of the images are pan-sharpened in order to increase spatial resolution by using ERDAS®, commercial GIS software. Image sizes vary between 1024x1024 and 2048x2048. The pan-sharpening operates on 4-m Red, Green, Blue and Near IR bands in accordance with 1-m panchromatic image. By this way, lower resolution images are merged with the 1-m data existing in the panchromatic image, so higher resolution images can be obtained. Ground truths of these images are obtained by manual annotation on the images. An example of IKONOS satellite image and the corresponding ground truth image can be seen in Figure 8.


(a) (b)

Figure 8:An example of IKONOS data (a) input image in gray scale (b) Ground Truth of

3.2 Introduction

The main objective of this research is to investigate the road extraction problem as well as to develop a robust road detection model. We propose a multi-resolution feature extraction based supervised classification method for road discrimination from high resolution multi-spectral satellite imagery.

Following (Bacher & Mayer, 2005), we group the common properties of roads into two categories; local and global properties. Locally, roads are elongated structures having homogeneity in color and width. Nevertheless, a road does not have to be formed with a single fragment. They can be composed of different fragments with different orientation properties. Road regions may also be exposed to heavy noise due to unpredictable environmental factors such as shadows, vegetation and vehicles driving on the road, or they can be just undetectable because of lack of proper resolution. See Figure 9 for illustrative examples.

(a)

(b)

(c)

(d)

Figure 9:Illustration of local difficulties faced in road extraction problem. (a) Shadow (b) vegetation (c) vehicles driving on road (d) lack of proper resolution

For example, in Figure 9-a, the road is partially occluded by shadow caused by the trees lying near the road. This is a major problem for segmentation algorithm, since there are two uniform color patterns on our target; roads are responsible for gray colors, and dark colors are induced by shadows. In Figure 9-b, we see another occlusion induced by vegetation. Indeed there are some trees whose branches overlay the road. This situation generally causes to sacrifice such regions as the regions' spectral responses certainly imply that there exists vegetation. In Figure 9-c, heavy traffic congestion causes non-uniform noise on the road surface. This is the main problem that is being tried to be solved in urban road extraction studies. The reaction of segmentation algorithm for such a situation will be a local over-segmentation, that is, small segments will be generated leading to difficulties for the extraction algorithm. The most difficult problem can be observed in Figure 9-d. There is a narrow road, which cannot be properly displayed in 1-meter resolution imagery.

The second group of road properties can be investigated from the global perspective. We can assume that roads are topologically connected regions; indeed, they are expected to form a globally connected road network. Since the intention to road construction for people is to fulfill the need of transportation; that is, roads do not interrupt abnormally.

On the other hand, parking lots and buildings having the same spectral features with roads are the major environmental factors that need to be dealt with by road extraction algorithms. Moreover, some of the buildings can have elongated structures same as roads. See Figure 10 for illustrative examples.

(a)

(b)

Figure 10:Illustration of global difficulties faced in road extraction problem. (a) Parking lots (b) elongated buildings

For instance, in Figure 10-a, there is a parking lot which have the same color with the road crossing near. This is the most challenging problem in road extraction, since these types of arrangements of road and non-road regions deteriorate the elongatedness property of corresponding road. On the other hand, in Figure 10-b, the buildings have very similar structural properties with roads. This issue can be resolved by color validation among the road extraction; however, this approach can cause elimination of true positives in some cases. That is, the road network can contain road parts having different spectral properties like in Figure 11.

Figure 11:An example of road topology with different road parts having different spectral features

In Figure 11, some of the roads have the color of soil, on the other hand, some of them have the color of asphalt. Nevertheless, these road parts together forms a single road network.

Up to this point, we have explained the major problems that must be coped with in road extraction problem. The complexity of the problem is narrated with visual examples. We propose an approach designed to deal with such problems by exploiting both local and global properties of the roads. The flow chart of the algorithm is pictured in Figure 12.

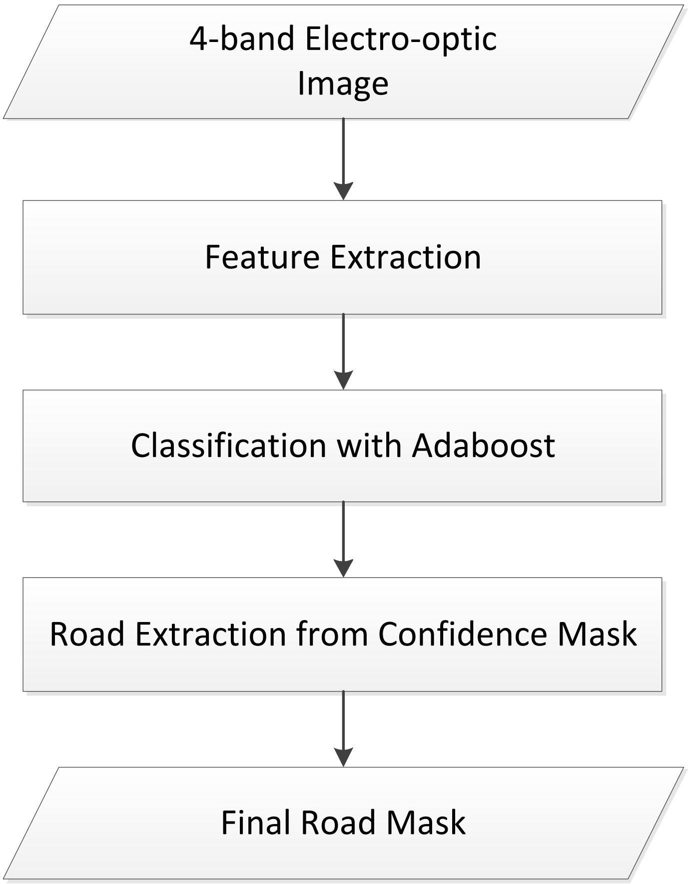


Figure 12:Flow chart of the proposed algorithm

The proposed algorithm consists of three main components including feature extraction, classification with Adaboost and road extraction. Firstly, features obtained by multi-resolution analysis and segmentation are combined with spectral features to form final feature vectors. Then, by using these features an Adaboost classifier is trained. Finally, the road regions are extracted from the confidence map provided by Adaboost classifiers.

In this method, the input image is divided into multi-scale levels, and at each level an image segmentation procedure is carried out. Multi-scale division step includes smoothing and down-sampling operations at each level, that is; gradual noise reduction and topological aggregation of road segments are achieved in the input image. By this way, at each level, the input image is analyzed into a different scale, that is, the roads hard to detect in a particular level are expected to be get easy to identify in another level. This approach is like looking and trying to understand what is in the image from near and then doing the same inspection again from a little farther distance at each step. We aim to detect as much road as possible at each level with no prior assumption about the width properties of the roads expected to be seen. Spectral features are also included into feature vectors to provide indicators for non-road regions such as water and vegetation. Another important aspect of this research is that for each level of segmentation, we provide optimal parameter sets obtained by pattern search algorithm for mean-shift segmentation. By this way, we try to provide features as discriminative as possible to Adaboost classifier. Moreover, Adaboost based learning method is adapted, as Adaboost poses good generalization properties among other learning alternatives. Finally, the road mask is extracted with a region growing based approach validating the connectivity constraint of road network.

3.3 Feature Extraction

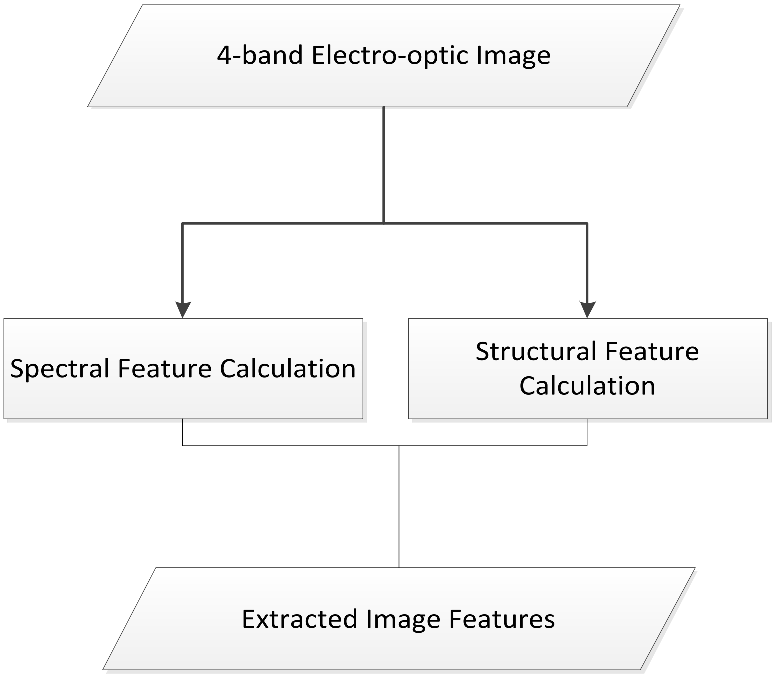


Figure 13:Flow chart of the feature extraction component

Feature extraction component is responsible for extracting required features from the input image to deliver them to the classification component. In this study, both spectral and structural properties of both road and non-road regions are examined for discrimination of roads from non-road regions. At the final feature space, each image pixel is represented by a d-dimensional vector. The internal structure of these vectors are such that;

$$Feature\_vector = [NDVI, SAVI, NDWI, structural\_score\_level\_0, \ldots\ structural\_score\_level\_L]$$

3.3.1 Spectral Features

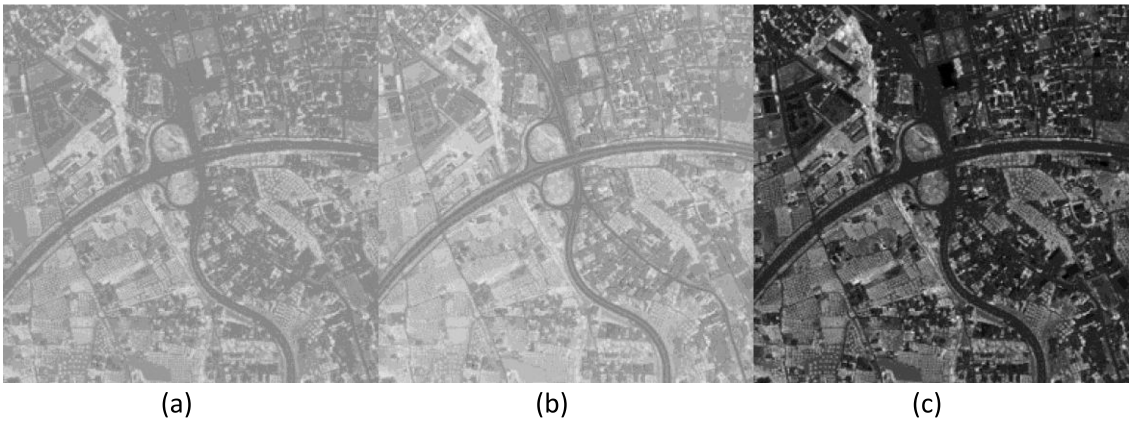
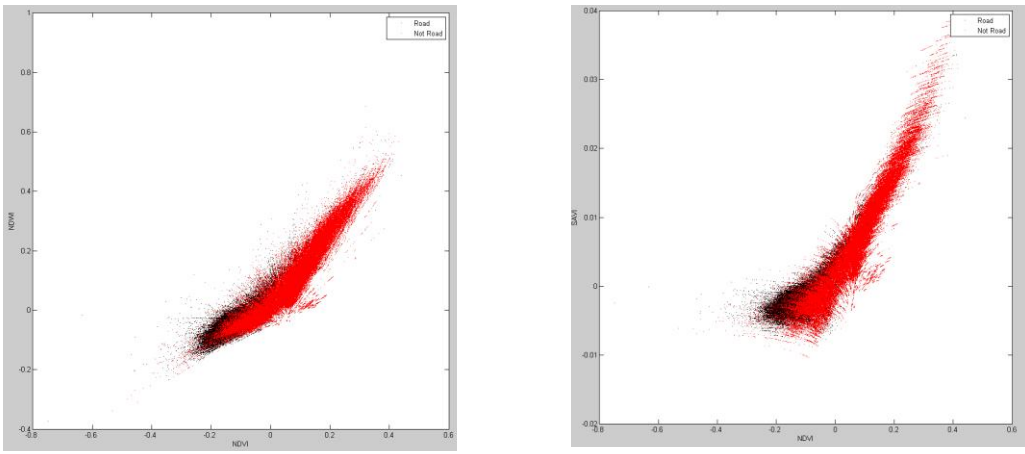


Figure 14:Illustration of spectral band ratios, for visualization purposes images are normalized to range [0,1]; (a) NDVI (b) NDWI (c) SAVI

Spectral features are not unique descriptors for road regions, since there are a variety of road types in the World having different color properties. However, band ratios can provide good indications for non-road regions such as vegetation and water. For instance, in Figure 14, NDVI, NDWI and SAVI images of the same scene is displayed. The bright regions correspond to vegetation areas; on the other hand, other regions have lower responses for such spectral band ratios. Band ratios are obtained by rationing two spectral bands, so the values of these types of spectral features can be generalized for all images in the dataset. In this study, three different band ratios are considered as spectral features; NDVI, NDWI and SAVI but not all of them are employed in the final solution. In fact, these spectral band ratios are highly correlated with each other as it can be seen in Figure 15.



(a) (b)

Figure 15:Intra-relations of spectral band ratios; black dot indicates road pixels and red dots represent non-road regions (a) NDVI vs. NDWI (b) NDVI vs. SAVI

Inclusion of highly correlated features into the same feature vector will increase complexity yet may provide no improvement. From Figure 15, the correlation of band ratios can be inferred. NDVI, NDWI and SAVI are providing similar values in the same spatial locations. Detailed analyses regarding spectral feature selection can be found in Chapter 4.

3.3.2 Structural features

Structural features are extracted by multi-resolution analysis of mean-shift segmentation results. Typically, roads are elongated regions, that is, their length sizes are much larger than their widths. We have already proposed a robust elongatedness index in another paper (Karaman, Cinar, Gedik, Yardimci, & Halici, 2012). For this study, we use this index for the purpose of structural scoring. Pseudo code of the structural feature extraction is given in Figure 16.

```
Input: Given input image I and number of reduction levels $L$
for $l$=0 to $L$
    • $I2 = pyramidial\_reduction(I, l)$, reduce the input image to level $l$
    • $S = meanshift\_segmentation(I2)$, segment the decomposed image by mean-shift
    • $C = structural\_scoring(S)$, assign structural scores to each segment
    • $O = pyramidial\_expansion(C, l)$, expand the structural score mask to its original size
    • Give output mask, $O$
endfor
Output: $O$, mask containing structural scores for each pixel in the image
```

Figure 16:Pseudo code of the structural feature extraction

Firstly, input image is reduced to desired level by using pyramidal reduction method. The reduction procedure includes a Gaussian smoothing operation followed by down-sampling. As a result, a less detailed version of the image is obtained. As the low scale image contains less noise and detail, it reacts differently to segmentation algorithm. So mean-shift segmentation algorithm is applied to reduced image. Each resolution level has its own optimum mean-shift segmentation parameter set obtained by pattern search optimization. These parameter sets are optimized by the pattern search algorithm in order to differentiate the road and non-road pixels at each level as much as possible. Then, each segment in the segmentation result is evaluated with the elongatedness index and structural scores are assigned to the corresponding pixels belonging to inspected segment. Before yielding the structural features, the structural score masks are expanded to the size of input image by using pyramidal expansion operator, since, we need to construct d-dimensional feature vectors for each pixel.

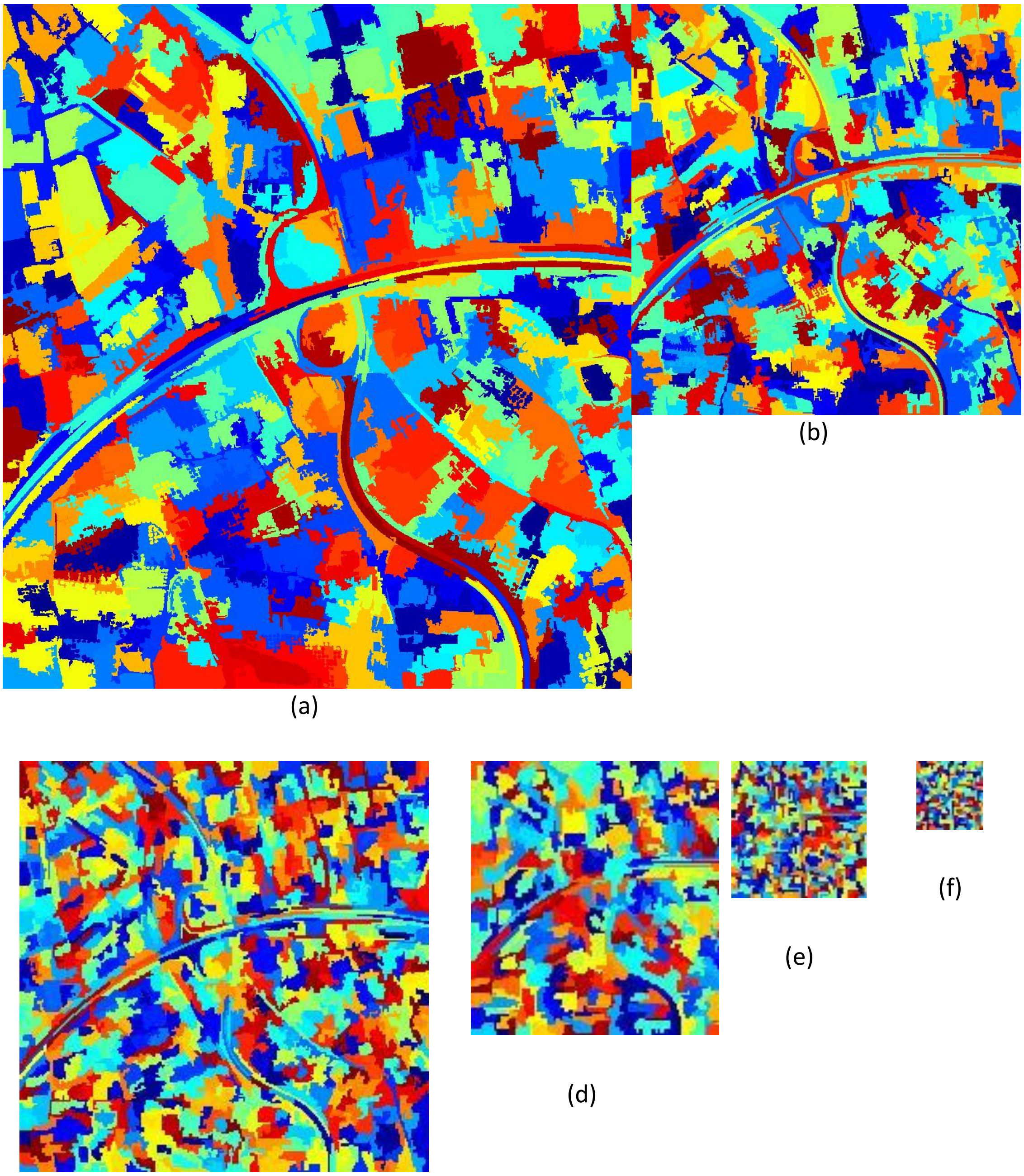
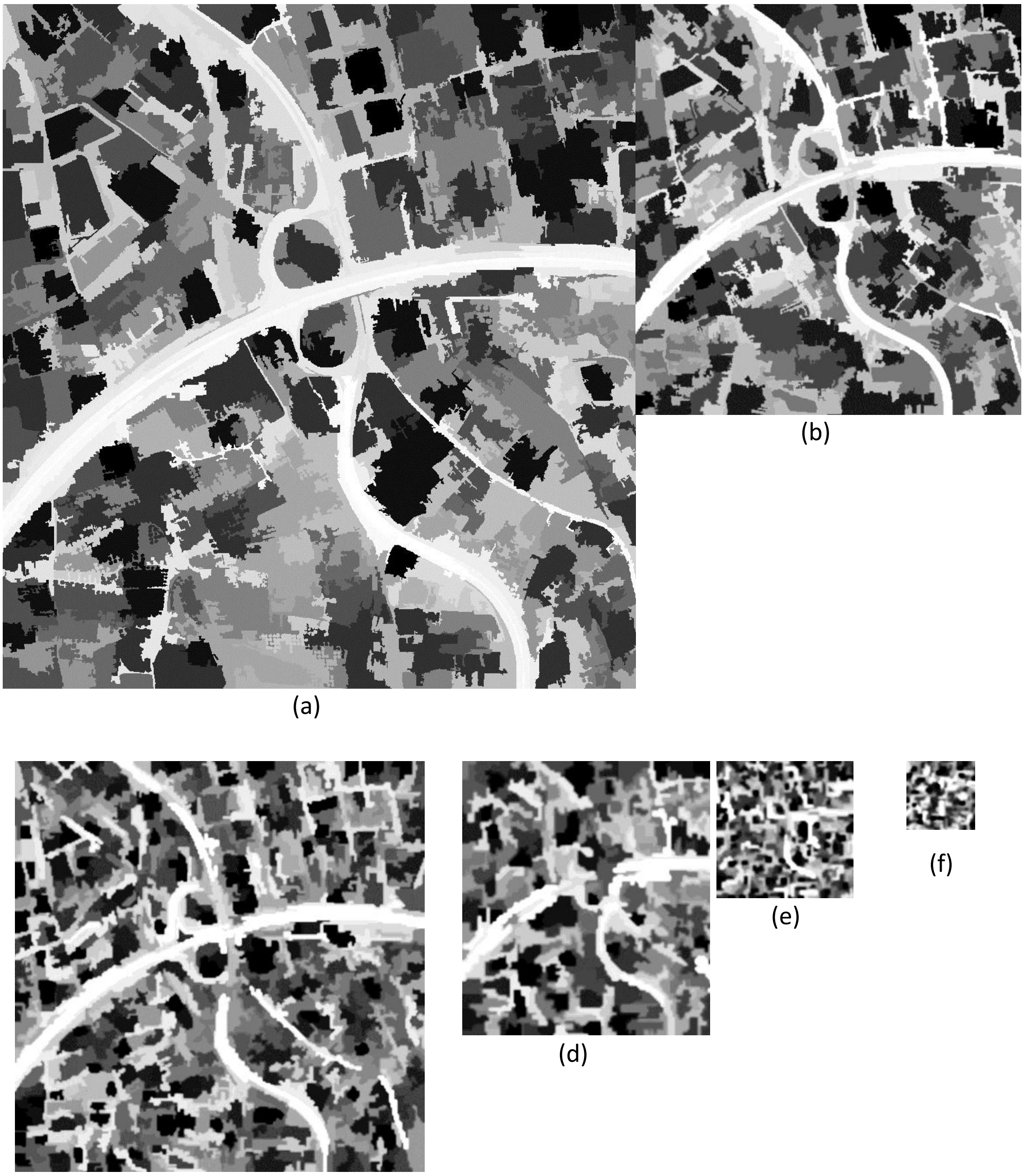

(a)
(b)
(c)
(d)
(e)
(f)

Figure 17:Showing results of multi-resolution segmentation for input images from levels of; (a) level-0 (b) level-1 (c) level-2 (d) level-3 (e) level-4 (f) level-5

In Figure 17, segmentation results for different resolution levels are illustrated. Further reductions of the input images result in segmentation results that are not relevant hence useful for road extraction after a certain number of iterations.



(b)

(a)

(f)

(e)

(d)

(c)

Figure 18: Showing results of structural scoring function followed by multi-resolution segmentation results for input images from levels of; (a) level-0 (b) level-1 (c) level-2 (d) level-3 (e) level-4 (f) level-5 (For visualization, images are normalized into [0,1] and applied histogram equalization)

The score masks for different levels are shown in Figure 18. Realize that different road regions appear bright or dark in different levels. These visual results prove our multi-resolution segmentation approach is capable of establishing different view of evaluations for the same road regions.

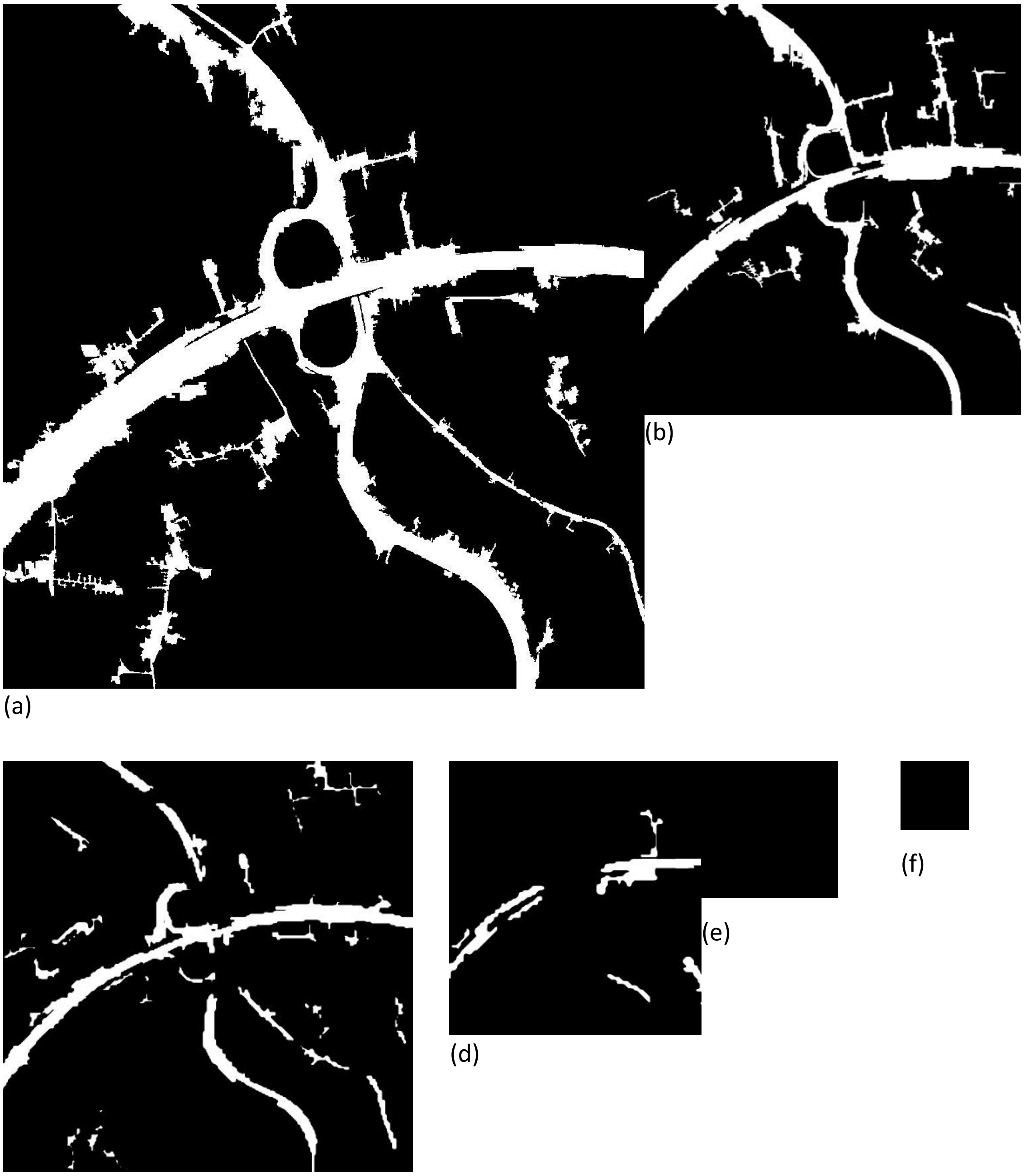

(a) (b) (c) (d) (e) (f)

Figure 19:Showing the regions having elongatedness value greater than 30(a) level-0 (b) level-1 (c) level-2 (d) level-3 (e) level-4 (f) level-5

In Figure 19, score mask pixels having elongatedness score greater than 30 are pictured in white. Note that different road parts are extracted in different levels. There is no precise rule for the properties of roads extracted at different levels; indeed, we cannot argue that certain kinds of roads are detectable only at a particular level. Furthermore, after a certain level of reduction, the road identification ability of structural features is lost. At the levels 4 and 5, no road can be extracted because the contrast between roads and non-roads is nearly lost at these levels.

The visualization of structural feature extraction procedure implies that strength of the proposed approach relies on multi-resolution analysis. As we have shown at the previous section, spectral features are not useful for discriminating roads and non-roads. On the other hand, structural features obtained by multi-resolution analysis are complementary to each other. This situation centers our proposed feature extraction algorithm into a new reinforcement modality depending on the structural analysis among different scales.

3.4 Mean-shift Segmentation Parameter Optimization

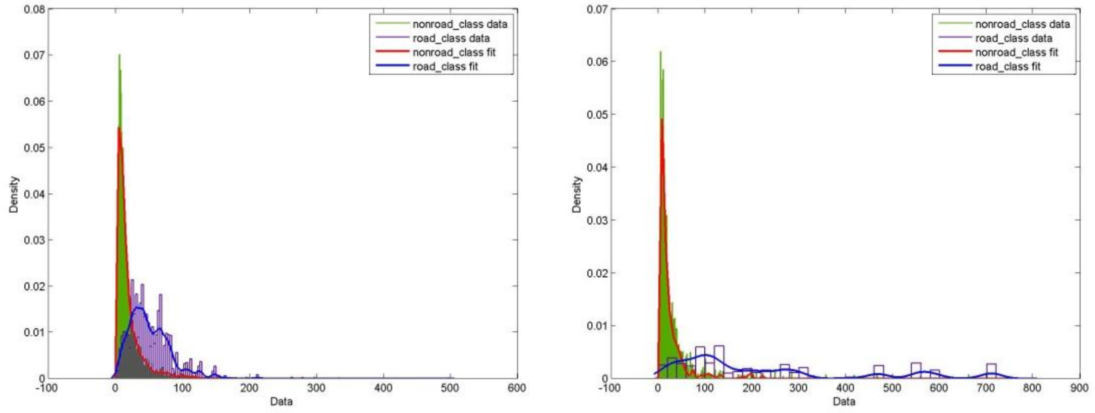
The problem of finding optimal parameter set for mean-shift segmentation is crucial issue for accurate road extraction. That is, mean-shift segmentation results should favor in structural separation of road and non-road regions. In order to obtain a good classification result with Adaboost, providing discriminative features to the training process is the key factor in our proposed model. For this purpose, proper parameters must be set in mean-shift segmentation, as mean-shift produces highly parameter dependent segmentation results. For our application, segments belonging to the road regions are desired to pose good elongatedness properties; on the other hand, non-road segments should depict different structural features from roads.

In the context of this study, we propose a mean-shift parameter optimization technique for structural road and non-road discrimination. The proposed method is based on pattern search algorithm which is a direct search optimization procedure. Pattern search algorithm does not require knowing the gradient of the problem to be optimized, and it can provide promising results for the functions that are not continuous or differentiable. Mean-shift has three distinct parameters (minimum segment area, spatial sigma, range sigma) to be set, and any combination of these parameters can produce different structural distributions of road and non-road pixels. Therefore, we need to investigate this 3-dimensional parameter space and make an optimum estimation about the parameter set that we will use. Pattern search algorithm is employed to find an optimum parameter set for each resolution level. Each level provides different interpretation of the input image, so segmentation parameters for each level must be optimized individually.

We want to obtain more separable distributions using the same elongatedness scoring strategy. In fact, the desired road distribution can be defined as having a maximal mean and minimal standard deviation. Similarly, non-road regions having minimal mean and minimal standard deviation are beneficial for creating separable distributions. Following these requirement, we devise an objective function to be maximized by pattern search algorithm, see function (3.1);

$$obj\_func = \frac{\mu_r - \mu_n}{\sigma_r + \sigma_n} \tag{3.1}$$

Where μ_r and μ_n are means, and σ_r and σ_n are standard deviations of the road and non-road region's distributions respectively. In this objective function, we want to maximize the distance of means while minimizing the in-class standard deviations.



(a) (b)

Figure 20:Illustration of the class distributions of road and non-road region's elongatedness scores. The results of evaluating segments obtained by using (a) an arbitrary mean shift parameter set (b) the parameter set decided by pattern search optimization

The pattern search algorithm is provided lower and upper bounds of the search space, and we obtain different set of parameters for each resolution levels. In Figure 20, the effect of using optimized parameter set for elongatedness evaluation is pictured on the class probability density distributions. The optimization results and investigations can be found in Chapter-4.

3.5 Classification with Adaboost

After feature extraction, these features are fed into the classification component to construct a strong classifier which is responsible for road and non-road pixel discrimination.

The training is accomplished by using Adaboost algorithm with decision trees as weak classifiers. In this iterative procedure, at each round, a weak classifier having least classification error is added up into the final strong classifier with its weight. After obtaining the strong classifier, features extracted from test images are given to this classifier to get road confidence mask. The pseudo code of the learning process can be seen in Figure 21.

```
input
    • $N$ Training examples vector of d-dimensional feature vectors, $X$
    • Label vector indicating the classes of the training data, $Y$
    • Distribution defined on training examples, $D_t(i) = \frac{1}{m}$
    • Number of iterations, $T$
for t = 1 to T
    • $construct\_weak\_classifier(X, Y, D_t)$, call WeakLearn algorithm using distribution $D_t$and get back weak classifier $h_t: X \rightarrow Y$
    • $calculate\_error(X, Y, h_t)$, calculate the error of $h_t$, $\varepsilon_t = \sum_{i=1}^{N} D_t[h_t(x_i) \neq y_i]$.
      If , then set $T = t - 1$ and abort loop
    • $set\_classifier\_weight(h_t, \varepsilon_t)$, set $a_t = \frac{1}{2}\left(\log(\frac{1-\varepsilon_t}{\varepsilon_t})\right)$
    • $update\_distribution(X, Y, a_t, D_t)$, update $D_{t+1} = \frac{D_t(i)}{\sum_{j=1}^{N} D_t(j)} e^{-a_t y_i h_t(x_i)}$
endfor
output
    • the final strong classifier, $h(x) = \sum_{t=1}^{T} a_t \, h_t(\mathrm{x})$
```

Figure 21: Pseudo code of the learning process

At the beginning of the algorithm, we define the number of iterations indicating the number of weak classifiers to be trained during the learning process. Initially, all of the training samples have the same weight values, and the iterative process adjusts this weight vector according to classification results obtained from weak classifiers. At the first step of iterations, a weak classifier, a decision tree in our case, is trained by the given samples. Naturally, the trained classifier is different from previous classifiers because at each step weights of examples are updated. The classification error is calculated by multiplying misclassified examples with their own weights. There is also an assessment statement which monitors the quality of weak classifier at subsequent iterations. As it is a requirement for weak classifiers that they must be at least slightly better than random guessing, their classification error is expected to be greater than 0.5. When the learning procedure begins to generate hypotheses having accuracies lower than 0.5, the loop is halted. After calculation of classification error, classifier weights are assigned accordingly. The computation of classifier weights depends on the classification error, so the new classifier's contribution to the final hypothesis depends on its classification success. Like classifier weights, the observation weights are updated with a function of classification error. At this update operation, former weights are normalized so as to generate a valid distribution. By the end of loop, the final strong classifier is constructed by calculating a weighted sum of weak classifiers. Indeed, the final hypothesis provides confidence values for each pixel about being belong to road. In theory, confidence values vary in the range $[-\infty, +\infty]$, and, the higher confidence value a pixel owns the more likely it is a road pixel. Each confidence value also implies a data margin which is the distance between decision boundary and the corresponding pixel.

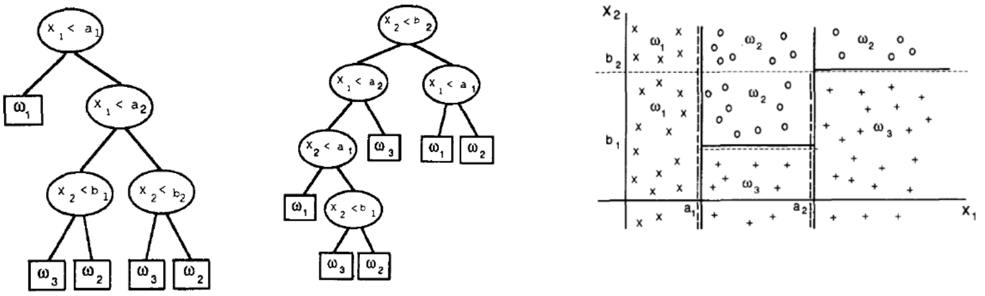
The learning process by Adaboost can be considered as the stepwise minimization of exponential loss function defined as;

$$loss = \sum_{n=1}^{N} \mathrm{w}_n \exp(-y_n \, \mathrm{h}(\mathrm{x}_n)) \tag{3.2}$$

where $y_t \in \{-1, +1\}$ the true class is label and $\mathrm{h}(\mathrm{x}_n)$ is the predicted classification score. w_n are the observation weights so that their sum is 1.

3.5.1 Decision trees

Classification and Regression Trees (CART) are used as weak classifiers which are trained in the Adaboost learning process. CART can be considered as a weak classifier because it divides the input space into axis parallel rectangles and assigns class numbers to each rectangle which is an insufficient discrimination method for complex problems. CART is a popular weak classifier for its robustness to the outliers (Timofeev, 2004); that is, it is very likely that CART classifier can isolate the outlier region in a single tree node. In the study (Safavian & Landgrebe, 1991), the authors provide a comprehensive sketch for decision boundaries generated by decision trees, which can be seen in Figure 22.



(a) (b) (c)

Figure 22:illustration of decision boundaries generated by CART. (a) Decision tree-1 (b) decision tree-2 (c) two dimensional feature space showing two possible partitionings, solid lines represent the boundaries for (a) and dashed lines represent the boundaries f

As it can be inferred from the Figure 22, there are many possible partitionings that can be generated by decision trees, although decision trees are simple classifiers. The pattern of the decision boundaries depends on the number of nodes in the tree and the threshold values assigned to these nodes. The main concern in the decision tree training process is to keep the resulting classifier simple and predictively powerful.

We can describe the mechanism of decision tree construction procedure in three sub-sections; tree construction, termination criterion definition and split optimization.

3.5.1.1 Tree Construction

Pseudo code of tree construction routine can be seen in Figure 23;

```
input
    • $N$ Training examples vector of d-dimensional feature vectors, $X$
    • Min number of observation that one node can represent, $MinObservations$
construct_tree function
for termination criterion is not met
    • $examine\_data$, on each dimension of data examine all possibilities of binary splits
    • $select\_a\_split$, a split which satisfies the best optimization criterion is computed
    • $create\_split$ , generated split is applied in the data structure
    • $construct\_tree(left\_child)$
    • $construct\_tree(right\_child)$, for both of the children same procedure is repeated
endfor
output
    • the final classifier WeakLearn
```

Figure 23:Pseudo code of tree construction routine

The tree construction function is initialized with training set and a parameter specifying minimum number of observations that a node may contain. Firstly, input vector is examined in all directions, and possible split thresholds are identified. Then, a split having the best optimization score is chosen to be applied. After constructing the split on the data structure representing the tree, the procedure is let to recursively follow child nodes. The whole tree is yielded when a termination criterion is met.

3.5.1.2 Termination Criterion

There are two possible termination criteria for decision tree construction. First possibility is valid when we reach a completely pure node which only contains observations from a single class. Since there is no need for further splits, the construction process is terminated. Another way of ending the construction process is the time when the current node represents fewer than $MinObservation$ data points. $MinObservation$ is a user defined parameter that is used to control the depth of the tree.

3.5.1.3 Split Optimization

We measure the split quality with Gini's index defined as;

$$1 - \sum_{i} p^2(i)$$

where i identifies class and $p(i)$ is the observed fraction of class i among all observations. This optimization function yields zero for pure nodes, otherwise, it will return a positive value. Indeed, Gini's index is a metric to measure impurity of node. Therefore, the tree construction method endeavors to generate splits having maximized purity.

A decision tree generated by our implementation can be seen in Figure 24.

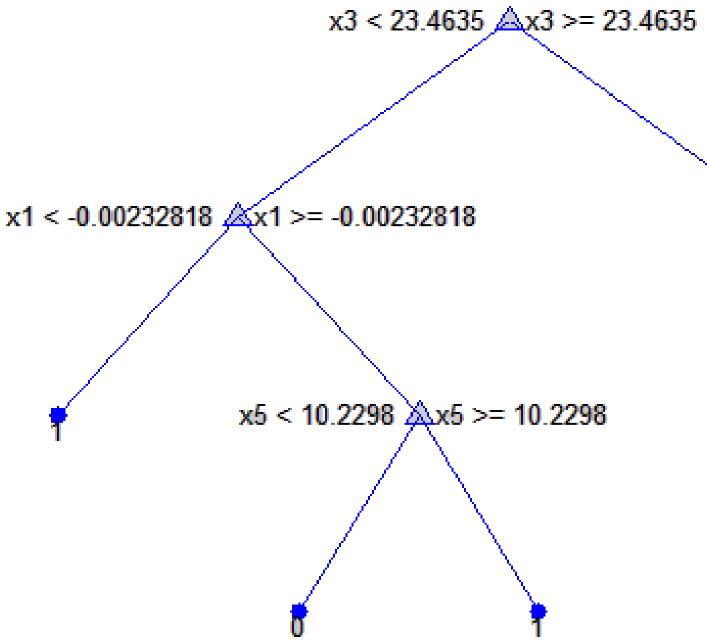


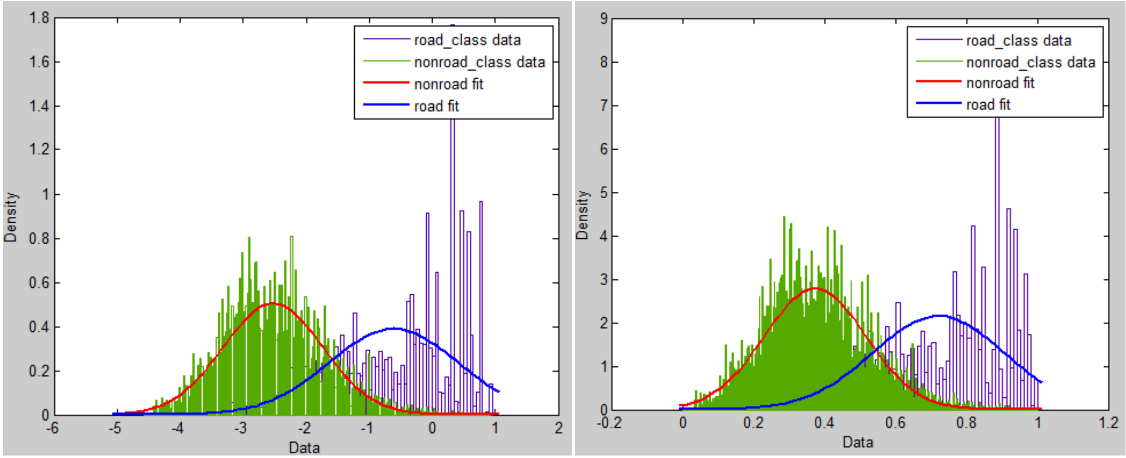
Figure 24:An illistration of a decision tree

In this tree, each node makes its decision by using a minimum of N/5 observations, where N is the total number of observations.

3.6 Road Extraction

In the original paper of Adaboost (Freund & Schapire , 1997), the authors suggest to use soft-thresholds in order to improve the error bound as the position of the error bound is application dependent. For complex problems that are hard to generate a discrimination rule, the soft-threshold is expected to be below zero. The confidence mask obtained from the strong classifier is used to extract roads; indeed, road pixels have larger values in this mask. So, it is relatively easy to extract roads from confidence mask. There are three different methods applied for the purpose of road extraction; hard thresholding, hysteresis thresholding and road detection by region growing.

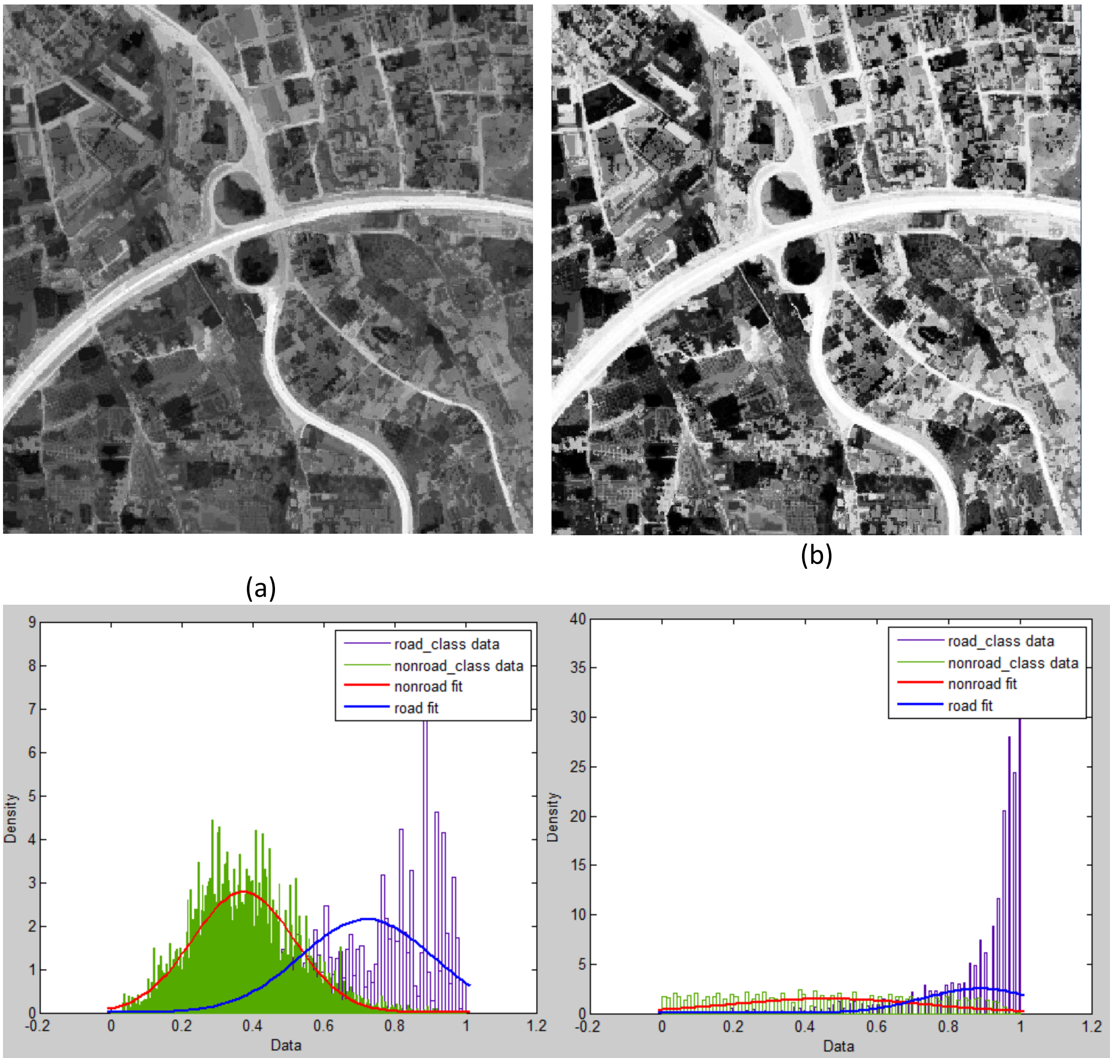
Before applying road extraction methods, confidence mask is pre-processed by a normalization operation followed by a histogram equalization operation. Normalizing the confidence mask into a pre-defined range is essential for our road extraction approach since our methods operate on a pre-defined range of [0, 1]. Additionally, histogram equalization operation increases the global contrast of the confidence mask; that is, the road regions, which are actually surrounded by non-road regions, become more separable.



(a) (b)

Figure 25:Showing the effect of normalization on data distributions (a) class distributions before normalization (b) class distributions after normalization

The normalization operation applied on the confidence mask does not reshape the distributions, it only transforms the values from an unpredictable range to the range [0,1]. By this operation, we can use same threshold values for all images; that is, the normalization operation provides generality for our road extraction methods. More significantly, we convert our confidence mask to a probability mask defined in Bayesian context where each image pixel represents the probability of being part of a road. This is an outcome of Adaboost theory which we have already explained in the Chapter-2.



(a)

(b)

(c)

(d)

Figure 26:Illustration of histogram equalization on probability mask (a) original probability mask (b) histogram equalized probability mask (c) class distributions before histogram equalization (d) class distributions after histogram equalization

	Road Data	Non Road Data
Mean (M_1)	0.7187	0.3711
Difference of Means	0.3476	
In-class Standard Deviation	0.1856	0.1437
Separation Score	1.0557	

Figure 27:Statistical information of distributions before histogram equalization

	Road Data	Non Road Data
Mean (M_1)	0.8843	0.4598
Difference of Means	0.4245	
In-class Standard Deviation	0.1598	0.2043
Separation Score	1.1658	

Figure 28:Statistical information of distributions after histogram equalization

In Figure 26, the effect of histogram equalization on the probability mask and the in-class data distributions can be seen. An important aspect of histogram equalization operation is that it increases the distance between two classes' mean; road and non-road. Also, looking to the in-class standard deviation of the road class we can infer that histogram equalization induces the narrowing in road class getting the road data closer in distribution. The distribution enhancement achieved by using histogram equalization is a distribution dependent procedure, that is; it may not always favor in class separation. However, we have observed increase in the class separation scores for most of the images. The experimental inferences are presented in Chapter-4.

3.6.1 Road Detection by Hard Thresholding

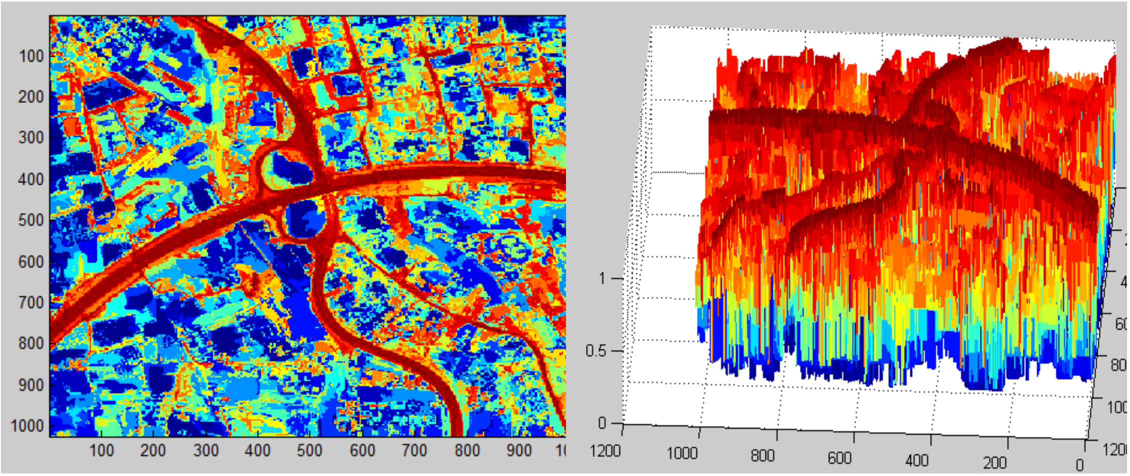
In the first method, pixels having probability values greater than a pre-defined constant are identified as road pixels. This method evaluates the individual pixel probability, and image's spatial domain information is disregarded. Although, this technique does not consider the roads as topologically connected regions, the outputs of this method seem visually appealing.

3.6.2 Road Detection by Hysteresis Thresholding

In the hysteresis method, some of the pixels are marked as roads via hard thresholding and any pixel neighbor to the road pixels are identified as roads if it has a value similar to its neighbor road pixel. This similarity is also determined by the user providing a second threshold. Therefore, there are two different threshold parameter defined by the user in this method. Hysteresis method uses the spatial domain information of the image; in fact, it operates on a 4-neighborhood range. This method results in visually satisfactory road extraction provided that a suitable parameter set is supplied.

3.6.3 Road Detection by Region Growing

Apart from elongatedness property, roads are also known as topologically connected.



(a) (b)

Figure 29:Illustration of probability mask, (a) heat map representation of probability mask, road regions have higher values. (b) 3-d mesh plotting of probability mask, realize that how road regions form continuous hills.

In Figure 29, probability mask is illustrated with a heat map and 3-d mesh representations. We can see that the road regions constitute continuous hills on 3-d graph; therefore, it is convenient to apply region growing algorithm seeded from the highest point on the probability mask. Nevertheless, a single seed for extraction of whole road network will be insufficient, so we developed an iterative method based on region growing to obtain whole road mask. The proposed method starts from the global maximum and tries to obtain an elongated segment. By this way, we can validate another important property of roads; connectivity. After completion of one region growing segmentation, another one is triggered from next global maxima. The flow chart of this approach can be seen in Figure 30, T_{sc} and T_{st} are user-defined parameters.

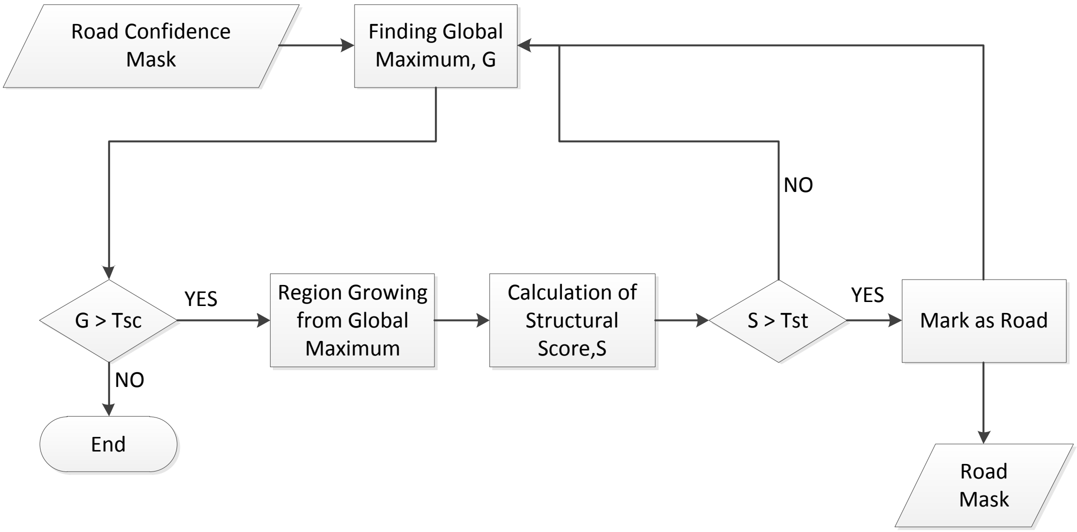


Figure 30:Flow chart of the region growing based road extraction strategy

The road extraction process begins with identifying the global maximum of the probability map. Then, a region growing operation is launched from this seed resulting in a road part. To ensure that we have obtained a segment belonging to the road network, its elongatedness property is validated by comparing its elongatedness score with a pre-defined constant. This validation step is useful when the classification algorithm gives high probability values to non-road regions. After this decision, the extracted segment is removed from the probability mask; that is, its pixel locations are set to zero in order to discard them at the next iteration. The procedure is repeated until the seed selected for region growing operation has a probability value below a pre-defined constant.

(a) (b) (c)

Figure 31:Results of road extraction methods by (a) hard thresholding, threshold value is 0.9 (b) hysteresis thresholding, threshold values 0.8 and 0.9 (c) region growing, global maximum limit is 0.9, minimum elongatedness value 20 and growing similarity is 0.1.

In Figure 31, a visual comparison between the results of the road extraction methods can be seen. Hard thresholding detects most of the roads but it is not capable of validating connectivity of the detected roads. Hysteresis method extracts more roads than hard thresholding; however, the increase in the false positive rate is noticeable. Region growing based technique seems to eliminate the regions having low elongatedness values as well as validating connectivity constraint.

CHAPTER 4

EXPERIMENTS

In this chapter, empirical inferences are gained by observing the proposed algorithm with different experiment configurations. This chapter is organized as follows. Firstly, the measurement metrics utilized in the experiments are introduced. Then, the experimentation model consisting of component and entire model based experiments are elaborated.

4.1 Measurement Metrics

Our road extraction strategy is composed of different types of components having theoretical interpretations in different field of information systems, so they must be evaluated with their own performance criteria. In this study, we group the performance measurement metrics in two groups; classification and road detection metrics. Following sub-sections are intended to give information about these metrics.

4.1.1 Classification Metrics

There are several measurement tools for classification assessment and validation. In the experiments, fundamental classification metrics are employed including classification error, and margin.

4.1.1.1 Classification Error

In its most basic form, classification error is the fraction of the data points misclassified by the classifier. Classification error indicates the accuracy of classifier, and it is used when considering different classifier alternatives. It is computed as in equation (4.1).

$$Classification\ Error\ = \frac{Number\ of\ Misclassified\ Data}{Number\ of\ Available\ Data} \quad (4.1)$$

However, we use exponential loss function to calculate classification error, as the Adaboost algorithm is designed to minimize the exponential loss function. By this way, we can monitor the success of the Adaboost algorithm in minimizing the loss function. For given N number of examples, loss function is defined as in equation (4.2).

$$Classification\ Error = \ Loss = \sum_{n=1}^{N} \mathrm{w}_n \exp(-y_n\, \mathrm{h}(\mathrm{x}_n)) \quad (4.2)$$

Where $y_t\ \in \{-1, +1\}$ the true class is label and $\mathrm{h}(\mathrm{x}_n)$ is the predicted classification score. w_n is the observation weights normalized to added up to 1. For the experiments, the weight vector w_n is set equal weights for each data point; that is, $\mathrm{w}_n = 1/N$ for any n value.

4.1.1.2 Classification Margin

In machine learning, classification margin is defined as the data point's distance to the decision boundary. In Adaboost, the margin is calculated by subtracting the maximal classification score for the false classes from the observation's score for the true class.

$$Classification\ Margin\ = \ Score\ for\ true\ class - Maximal\ score\ for\ false\ classes$$

By using classification margin, one can decide the quality of the classification; that is, high margin values imply better generalization achieved by classifier.

4.1.2 Road Detection Metrics

Our road extraction component outputs a binary mask in which true pixels represent road and false pixels represent non-road regions. The ground truth images generated for the training and testing purposes are also in the same format. To compare the detected road mask with the ground truth, we use precision and recall metrics.

The indicators below are calculated to assess the performance;

True positives (TP): the number of data correctly classified as belonging to the positive class is called as true positives

False positives (FP): the number of data incorrectly classified as belonging to the positive class is called as false positives

True negatives (TN): the number of data correctly classified as belonging to the negative class is called as true negatives

False negatives (FN): the number of data incorrectly classified as belonging to the negative class is called as false negatives

The definitions of TP, FP, TN and FN is presented in Table 2.

Table 2:The indicators used for performance evaluation

	Actual class (expectation)	
Predicted class (observation)	True Positive (TP) Correct result for true class	False Positive (FP) False alarm
	False Negative (FN) Missed target	True Negative (TN) Correct result for false class

4.1.2.1 Precision

Precision is the fraction of detected points that are indeed part of the target; that is, precision indicates the probability that a randomly selected sample from the pool of retrieved samples is detected correctly. This metric is calculated as in equation (4.3).

$$Precision = \frac{TP}{TP + FP} \tag{4.3}$$

In our case, precision show the fraction of actual road pixels among the detected road class.

4.1.2.2 Recall

Recall is the fraction of target points that are retrieved; that is, recall indicates the probability that a randomly selected sample from the pool of relevant samples is retrieved. This metric is calculated as in equation (4.4).

$$Recall = \frac{TP}{TP + FN} \tag{4.4}$$

In our case, recall show the fraction of road pixels that our algorithm detects among the ground truth pixels.

4.1.2.3 F-measure

F-measure is the harmonic mean of precision and recall. F_β measure is computed as in equation (4.5).

$$F_\beta = (1+\beta)\frac{Precision\ x\ Recall}{\beta^2 xPrecision + Recall} \qquad (4.5)$$

The β parameter controls the importance weight given to precision or recall. $F_{0.5}$, F_1 and F_2 measures are three commonly used variations of F-measure. $F_{0.5}$ measure gives more importance to precision while F_2 puts more emphasis on recall. F_1 measure provides the balanced importance given to both precision and recall. F-measure combines precision and recall into one metric, and we can examine the aggregated performance of precision and recall.

4.2 Experimentation Model

We partition the experiments carried out into a hierarchical model. The flow chart of the experimentation model can be seen in Figure 32.

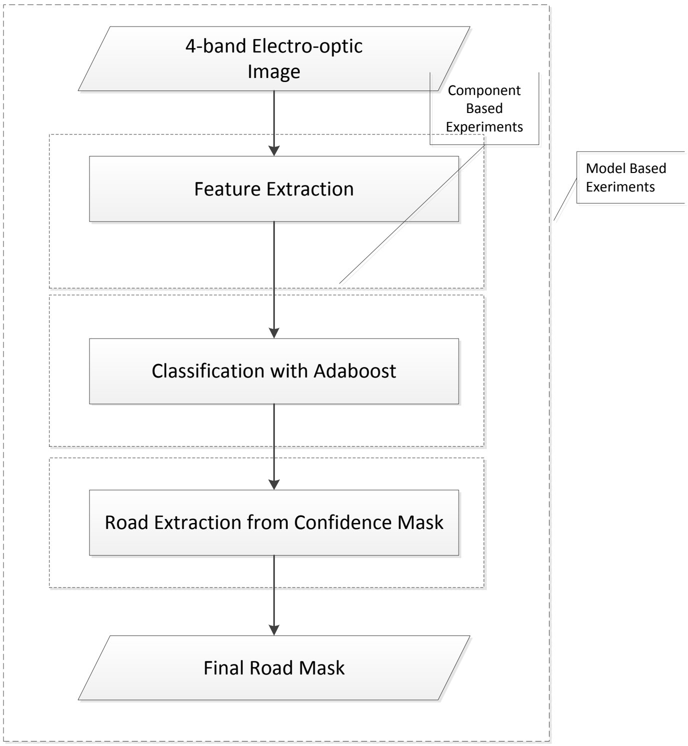


Figure 32:Experimental flow chart

In this model, firstly, the behaviors of each component with respect to its parameters are examined. Component based experiments aim to investigate and understand the internal and external dependencies of the components individually. Additionally, validation of the proposed methods and performance optimization of components are accomplished in component based experimentation analyses. The component based experimentations progress sequentially. After completing the optimizations and assessments on the former component, the next component is examined with the optimal input provided by the former component.

On the other hand, model based experiments operate on the entire algorithm; indeed, model based experiments aim to provide higher level descriptions and investigations regarding to the proposed road extraction strategy. In these series of experiments, interactions of components and validity of proposed techniques are analyzed.

For consistency, each experiment is described in a common format which is composed of five four headlines including purpose, procedure, observations and remarks/interpretations.

4.3 Component Based Experiments

4.3.1 Experiments on Feature Extraction Component

The main objective of the feature extraction component is to extract and generate the classification feature vectors in an optimal way. As discussed in Chapter-3, the objective function defined in equation (3.1) is the essential indicator of road and non-road class separation. The objective function requires the construction of two class vectors consisting of their densities and corresponding elongatedness scores; also we sketch the probability density distributions of the classes for visualization purposes. The discrete data used for the construction of class vectors is depicted on the grid, which can be seen in Figure 33.

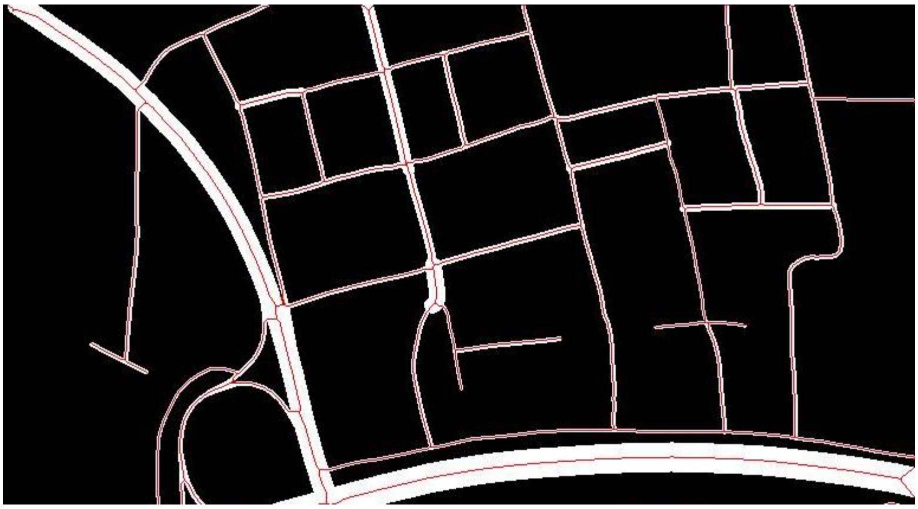


Figure 33:Illustration of the class data formation.

In Figure 33, white pixels indicate the road mask, red pixels represent the data used in road class distribution analysis, and black pixels are evaluated as non-road class data in distribution analysis. In other words, the pixels belonging to the skeleton of the road mask are taken into account in class distribution analysis. By this way, each road component is treated equally. That is, roads have a variety of width values, so we employ the skeleton pixels to represent the road class in probability distribution in order to suppress the domination effect caused by wide roads.

The parameter analysis and tuning experiments are conducted on two distinct image sets; optimization and validation sets. These image sets are composed of six images, and the images of the optimization set is a subset of the training set employed in the classification component whereas the validation set includes the images from the test set of the classification component.

In this series of the experiments, we examine the dynamics affecting feature extraction component, and component based parameter tuning is another function of the experiments conducted in this section.

4.3.1.1 Observing Individual Behaviors of Mean-shift Parameters

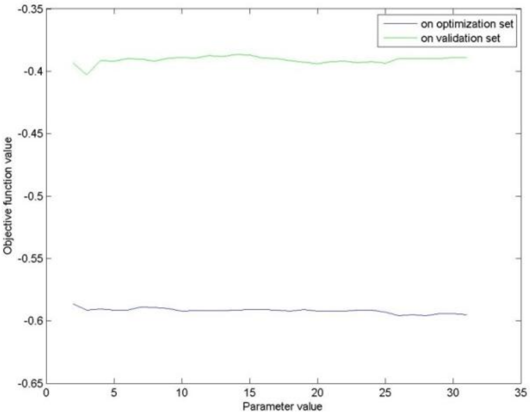
Purpose:

Mean-shift image segmentation algorithm has three distinct parameters directly influencing segmentation results which are spatial domain bandwidth, range domain bandwidth and minimum segment area. In this experiment, the main purpose is to observe individual behaviors of each of three parameters, that is; we need to define effect of any segmentation parameter change on the elongatedness distributions of road and non-road classes.

Procedure:

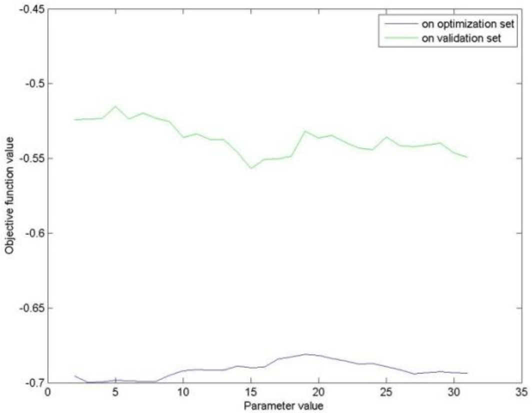
Since this experiment aims to investigate individual parameter effects on elongatedness distributions, one parameter is set varying and the other two are introduced as constants. The variable parameter is increased at each step with a pre-defined step size, and objective function is evaluated at each step. Output of the objective function is used to compare the parameter tradeoffs. The experiment is conducted on both optimization and validation sets so that we can verify our observations.

Observations:



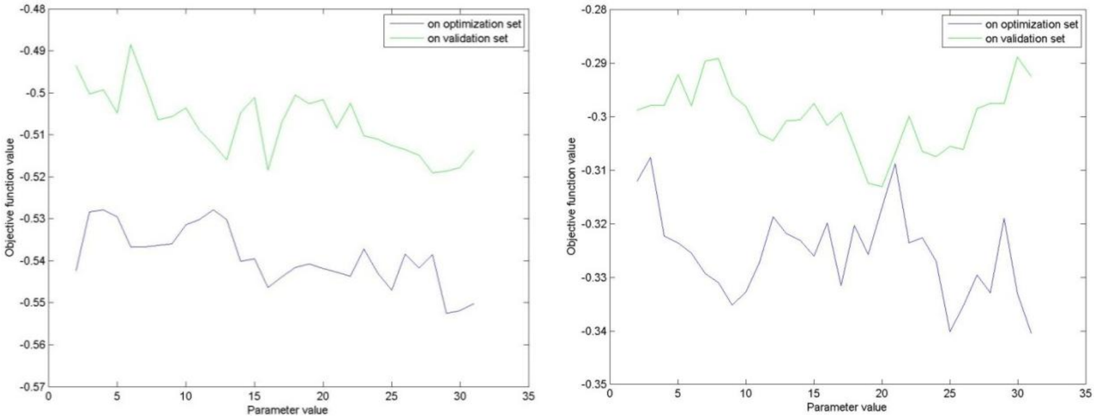
	Optimization set	Validation set
Maximum	0.5958	0.4027
Minimum	0.5866	0.3864
Average	0.5919	0.3909
Std	0.0020	0.0030
Optimum Value	28	3

(a)



	Optimization set	Validation set
Maximum	0.6997	0.5569
Minimum	0.6811	0.5152
Average	0.6914	0.5367
Std	0.0055	0.0105
Optimum Value	3	15

(b)



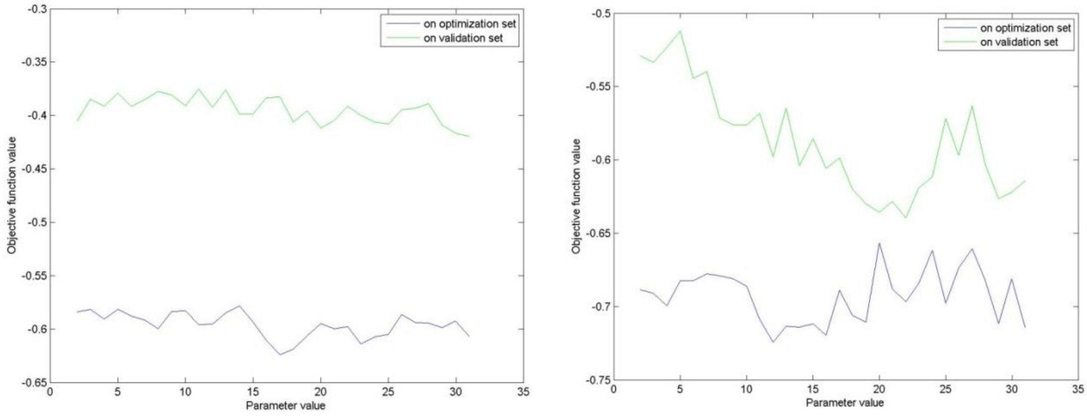
	Optimization set	Validation set
Maximum	0.5525	0.5191
Minimum	0.5278	0.4885
Average	0.5392	0.5072
Std	0.0069	0.0078
Optimum Value	29	28

(c)

	Optimization set	Validation set
Maximum	0.3404	0.3131
Minimum	0.3076	0.2888
Average	0.3254	0.3001
Std	0.0082	0.0062
Optimum Value	31	20

(d)

Figure 34:Influence of spatial bandwidth parameter on objective function for (a) level-0 (b) level-1 (c) level-2 (d) level-3

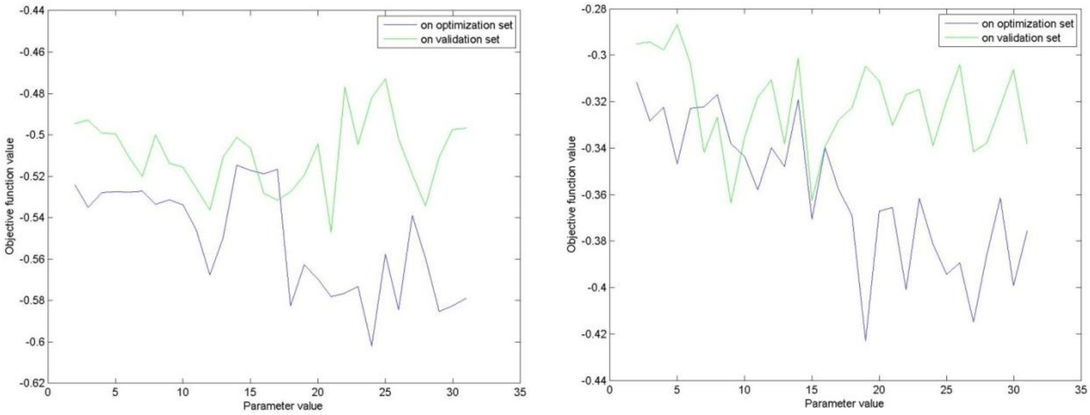


	Optimization set	Validation set
Maximum	0.6240	0.4193
Minimum	0.5780	0.3750
Average	0.5960	0.3946
Std	0.0116	0.0121
Optimum Value	17	31

(a)

	Optimization set	Validation set
Maximum	0.7243	0.6395
Minimum	0.6564	0.5120
Average	0.6924	0.5870
Std	0.0181	0.0363
Optimum Value	12	22

(b)



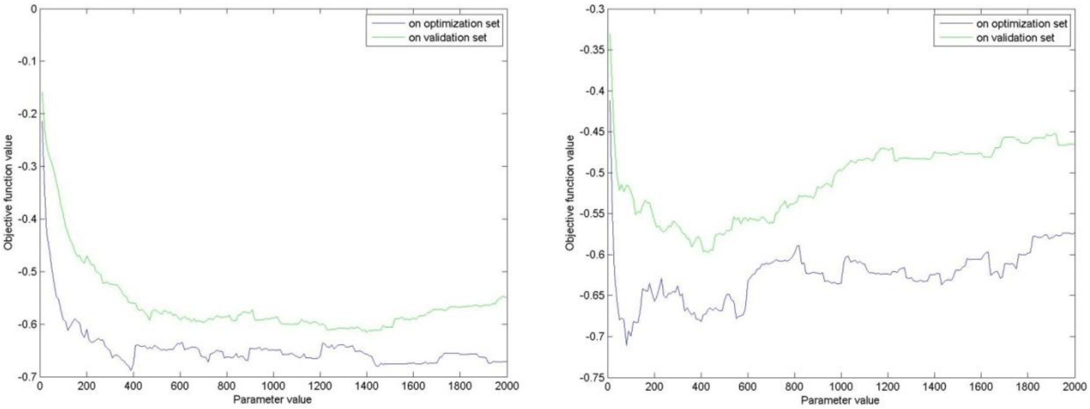
	Optimization set	Validation set
Maximum	0.6020	0.5471
Minimum	0.5147	0.4730
Average	0.5510	0.5094
Std	0.0259	0.0176
Optimum Value	24	21

(c)

	Optimization set	Validation set
Maximum	0.4230	0.3637
Minimum	0.3117	0.2869
Average	0.3592	0.3217
Std	0.0304	0.0196
Optimum Value	19	9

(d)

Figure 35:Influence of range bandwidth parameter on objective function for (a) level-0 (b) level-1 (c) level-2 (d) level-3

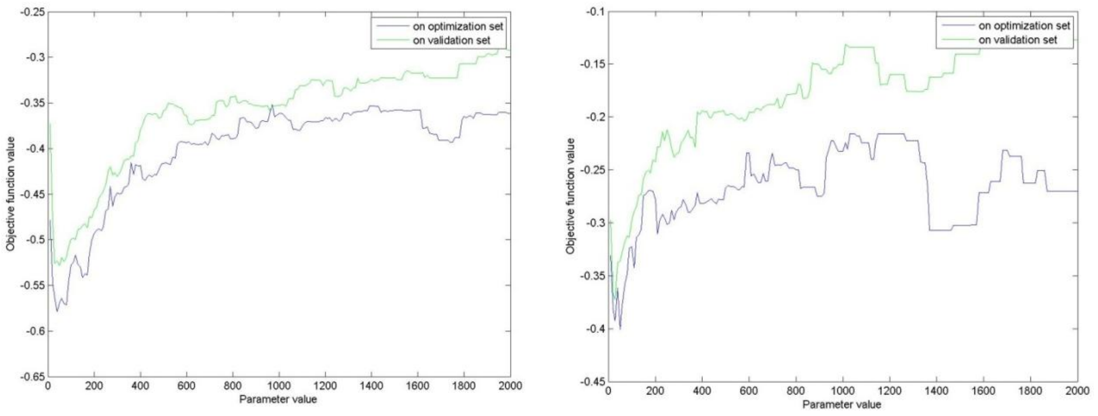


	Optimization set	Validation set
Maximum	0.6879	0.6157
Minimum	0.2142	0.1586
Average	0.6452	0.5596
Std	0.0500	0.0724
Optimum Value	390	1400

(a)

	Optimization set	Validation set
Maximum	0.7108	0.5978
Minimum	0.4121	0.3301
Average	0.6251	0.5094
Std	0.0330	0.0452
Optimum Value	80	430

(b)



	Optimization set	Validation set
Maximum	0.5789	0.5280
Minimum	0.3515	0.2903
Average	0.4001	0.3592
Std	0.0550	0.0565
Optimum Value	40	50

(c)

	Optimization set	Validation set
Maximum	0.4008	0.3723
Minimum	0.2155	0.1257
Average	0.2665	0.1774
Std	0.0355	0.0507
Optimum Value	50	30

(d)

Figure 36:Influence of minimum region parameter on objective function for (a) level-0 (b) level-1 (c) level-2 (d) level-3

Remarks and Interpretations:

In this experiment, the optimization set achieves better separation scores in all resolution levels compared to validation set. Therefore, we can conclude that an optimum parameter set suitable for a particular set of images might not produce same objective function value when applied on a different set of images. Variety of images in each set is the main reason for observing different objective values, indeed; elongatedness property among the images belonging to the optimization set is more prominent than the images in the validation set. However, parameter behaviors of two sets are almost the same.

In first level, spatial bandwidth parameter poses nearly no effect on the objective function value, but after first level the fluctuations on graph begins to be notable. This level by level increase in the behavior fluctuations implies that effect of spatial bandwidth parameter becomes unstable in lower resolution images. For example, after a certain level number it is obvious that there is no correspondence between the graphs of optimization and validation sets. Indeed, an arbitrary parameter change in the optimization set can cause a score decrease while the same parameter alteration is beneficial for validation set resulting in a score increase. Likewise, range bandwidth parameter is subject to the same behavior in different resolution levels. Another important observation for this experiment is that the instability of the bandwidth parameter behaviors begins to increase after a certain parameter value. Therefore, we should consider defining a rough upper bound on these parameters to avoid image specific optimums. In relation to this investigation, (Comaniciu & Meer, 2002) state that the mean-shift image segmentation is not very sensitive to the choice of resolution parameters, and the parameter selection is a problem dependent issue.

On the other hand, the minimum segment area parameter shows a straightforward behavior on different resolution levels. By reducing the resolution of the input image, we are actually reducing the area of targeted road objects, and we can observe this phenomenon by comparing the optimum values obtained in each level. The parameter behaviors on the optimization and validation sets are also similar which implies that the experiment is valid. Nevertheless, in the first two levels, the optimum value for optimization set is much smaller than the value found four validation set, but the behaviors on two distinct sets still coincide. That is, a global optimum for one set is a local optimum for another set providing a score value slightly worse than its global optimum. A stable optimum value can be found by adding extra images to optimization data set and letting the process to converge. In this experiment, we conclude that an optimum parameter value for minimum segment area can provide a significant score gain.

4.3.1.2 Mean-shift Parameter Optimization by Pattern Search Algorithm (Uncensored)

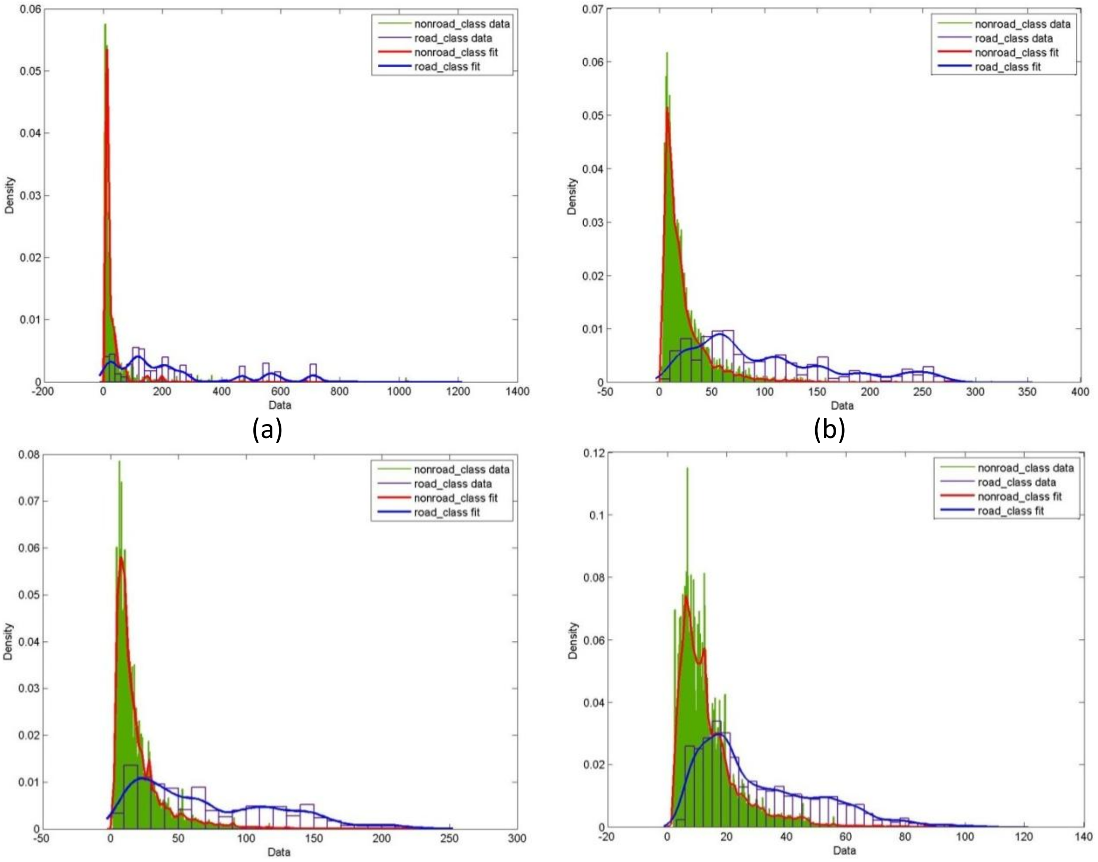
Purpose:

In the former experiment, we have investigated the individual behaviors of mean-shift parameters. The purpose of this experiment is to find out and validate the optimal mean-shift parameter sets for each resolution level. In multi-resolution image interpretation, each reduction level may favor in different parts of the road network eliminating the noise and redundant detail from the image. Another aim of this experiment is to explain the dependencies between different parameters of the mean-shift segmentation from the perspective of elongated structure generation.

Procedure:

The pattern search algorithm is executed for each resolution level on both of the optimization and validation sets. Pattern search algorithm explores the three dimensional parameter space for the purpose of minimizing the objective function. The optimization procedure terminates and outputs the best parameter set if there is no increase in the best objective function value after a pre-defined number of iterations.

__Observations:__



(a) (b)

(c) (d)

Figure 37:Probability density distributions of road and non-road class data after optimization, the sketches are generated for image-1, (a) level-0 (b) level-1 (c) level-2 (d) level-3

Table 3: Optimum distribution statistics for optimization and validation sets, (a) level-0 (b) level-1 (c) level-2 (d) level-3

	Optimization set	**Validation set**
Road class Min	3.4584	3.5359
Road class Max	447.5522	426.7034
Road class	117.2570	82.4156

	Optimization set	**Validation set**
Road class Min	3.7715	3.4683
Road class Max	254.0864	292.6097
Road class	63.4460	86.4723

Mean		
Road class Std	111.5840	77.9947
Nonroad class Min	2.4823	2.3888
Nonroad class Max	492.7288	426.7034
Non-road class Mean	17.0980	18.8462
Non-road class Std	24.5660	27.7980
Optimum Obj value	0.6728	0.6157

(a)

Mean		
Road class Std	50.8308	76.7928
Nonroad class Min	1.3985	0.3755
Nonroad class Max	246.8408	298.2379
Non-road class Mean	15.0162	19.1456
Non-road class Std	14.6371	26.4106
Optimum Obj value	0.7153	0.6035

(b)

	Optimization set	Validation set
Road class Min	4.0074	3.0965
Road class Max	169.1651	161.3542
Road class Mean	44.7980	36.8128
Road class Std	35.2206	27.8226
Nonroad class Min	2.0098	2.0068
Nonroad class Max	176.9158	162.2808
Non-road class Mean	14.0775	14.5259
Non-road class Std	13.1126	13.4124
Optimum Obj value	0.6071	0.5369

(c)

	Optimization set	Validation set
Road class Min	3.2194	2.5424
Road class Max	85.5774	81.8495
Road class Mean	21.6144	22.5641
Road class Std	15.4091	16.4753
Nonroad class Min	2.0006	1.9411
Nonroad class Max	86.2532	94.9094
Non-road class Mean	11.4405	12.3379
Non-road class Std	8.6100	9.5288
Optimum Obj value	0.4060	0.3782

(d)

Table 4:Optimum parameter sets determined by the pattern search algorithm for optimization and validation sets, (a) level-0 (b) level-1 (c) level-2 (d) level-3

	Optimization set	Validation set
Spatial Bandwidth	4	4
Range Bandwidth	4	4
Minimum Region Area	717	1401
Optimum	0.6728	0.6157

	Optimization set	Validation set
Spatial Bandwidth	4	4
Range Bandwidth	8	6
Minimum Region Area	81	417
Optimum	0.7153	0.6035

Obj value		

(a)

Obj value		

(b)

	Optimization set	Validation set
Spatial Bandwidth	8	4
Range Bandwidth	8	4
Minimum Region Area	57	47
Optimum Obj value	0.6071	0.5369

(c)

	Optimization set	Validation set
Spatial Bandwidth	8	4
Range Bandwidth	4	6
Minimum Region Area	23	29
Optimum Obj value	0.4060	0.3782

(d)

<u>Remarks and Interpretations:</u>

In this experiment, combinational behaviors of the mean-shift parameters are examined in a three dimensional search space by the pattern search algorithm. The optimization process is expected to find optimum segmentation parameter sets for each resolution level minimizing the objective function. In Figure 37, illustration of probability density distributions of two class obtained by optimum parameter segmentation can be seen. In early resolution levels, the distributions seem to be separated well although the disjunction between two classes is lost by decreasing resolution of the image. This observation can be also verified by examining the in-class statistics of the distributions presented in *Table 3*. On the other hand, there is a huge difference between the minimum and maximum values of road class data which implies that a complete separation of two classes at this stage of the road extraction model is not possible.

Another important deduction can be obtained by examining the in-class statistics of the distributions of first two levels. Although level-0 road class distribution has much greater class mean than the level-1, the objective function value indicates better class separability for level-1. The underlying reason is that segmentation with higher minimum region area parameters causes to spatially combine the fragmented road segments resulting in a better elongatedness score. At first glance it seems favorable in road extraction process, yet the thin road parts also tend to be unified with the segments in vicinity which causes a deliberative loss of targeted road regions. Moreover, the standard deviation of the road class in level-0 is very high compared to level-1, which also a good indicator for the wrong tendency of the optimization procedure. This challenge is dealt with at the next experiment by censoring the elongatedness scores.

4.3.1.3 Mean-shift Parameter Optimization by Pattern Search Algorithm (Censored)

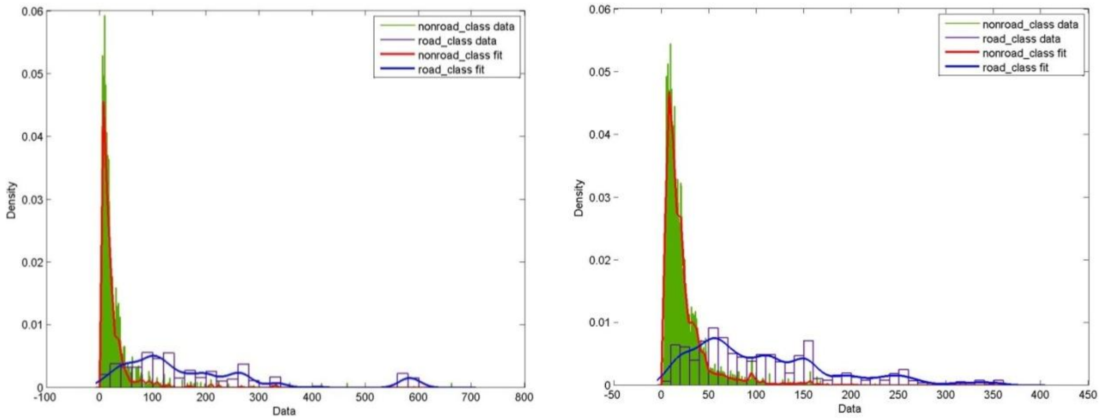
<u>Purpose:</u>

The purpose of this experiment is to establish a better optimization model for mean-shift segmentation parameters. From the former experiment, we observe the presence of elongatedness scores about 117 at the first resolution level which means that an average road segment's long side length is greater than its short side length in multiple of 117. For the road extraction procedure, the segments having such elongatedness scores are over-qualified, and we integrate a censoring mechanism into optimization model in order to avoid over-qualified road segments.

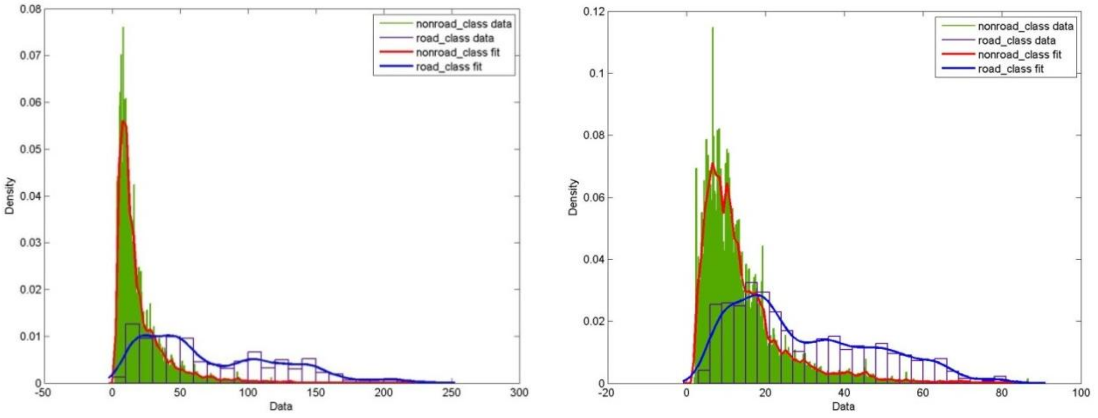
<u>Procedure:</u>

Different from the former experiment, the scoring routine is adjusted so that none of the data points can receive a score higher than *score_limit*, which is a pre-defined constant. By this way, the class distributions are left censored, and the pattern search algorithm has a four dimensional search space to find an optimum parameter set including the censoring variable.

Observations:



(a) (b)



(c) (d)

Figure 38: Probability density distributions of road and non-road class data after optimization, the sketches are generated for image-1, and for proper visualization the data points are assigned to their actual values rather than censored values (a) level-0 (b) level-1 (c) level-2 (d) level-3

	Optimization set	Validation set
Road class Min	3.6190	3.6159
Road class Max	64.0000	48.0000
Road class Mean	45.3169	36.0496
Road class Std	19.2270	14.5101
Nonroad class Min	2.4223	2.3714
Nonroad class Max	64.0000	48.0000
Non-road class Mean	14.8025	15.1121
Non-road class Std	11.8485	11.7662
Optimum Obj value	0.9878	0.8045

(a)

	Optimization set	Validation set
Road class Min	3.6348	3.3681
Road class Max	56.4290	48.6083
Road class Mean	38.1667	34.6795
Road class Std	16.1771	13.9329
Nonroad class Min	1.9342	1.8189
Nonroad class Max	56.5311	48.7668
Non-road class Mean	14.2480	14.7452
Non-road class Std	10.3722	11.1954
Optimum Obj value	0.9039	0.8053

(b)

	Optimization set	Validation set
Road class Min	4.0982	3.1443
Road class Max	54.0502	38.0440
Road class Mean	29.6826	24.9885
Road class	14.9524	10.4963

	Optimization set	Validation set
Road class Min	2.9678	2.5443
Road class Max	25.9960	26.0029
Road class Mean	15.5147	15.9168
Road class	6.3078	6.7144

Std		
Nonroad class Min	2.0105	1.9310
Nonroad class Max	54.0851	38.0700
Non-road class Mean	13.2044	12.9491
Non-road class Std	9.5597	8.4565
Optimum Obj value	0.6630	0.6369

(c)

Std		
Nonroad class Min	1.9726	1.7004
Nonroad class Max	26.0039	26.0048
Non-road class Mean	10.2628	10.7588
Non-road class Std	5.5516	5.9501
Optimum Obj value	0.4398	0.4081

(d)

Figure 39: Optimum distribution statistics for optimization and validation sets, (a) level-0 (b) level-1 (c) level-2

	Optimization set	Validation set
Spatial Bandwidth	4	4
Range Bandwidth	6	4
Minimum Region Area	383	697
Score Limit	64	48
Optimum Obj value	0.9878	0.8045

(a)

	Optimization set	Validation set
Spatial Bandwidth	4	10
Range Bandwidth	4	4
Minimum Region Area	101	125
Score Limit	56	48
Optimum Obj value	0.9039	0.8053

(b)

	Optimization set	Validation set
Spatial Bandwidth	8	4
Range Bandwidth	8	4
Minimum Region Area	55	43
Score Limit	54	38
Optimum Obj value	0.6630	0.6369

(c)

	Optimization set	Validation set
Spatial Bandwidth	8	4
Range Bandwidth	4	4
Minimum Region Area	21	21
Score Limit	26	26
Optimum Obj value	0.4398	0.4081

(d)

Figure 40: Optimum parameter sets found by the pattern search algorithm for optimization and validation sets, (a) level-0 (b) level-1 (c) level-2 (d) level-3

Table 5: Comparison of optimum objective values achieved in the optimization experiments.

	On optimization set			On validation set		
	Optimum Objective Value (Uncensored)	Optimum Objective Value (Censored)	Gain Percentage	Optimum Objective Value (Uncensored)	Optimum Objective Value (Censored)	Gain Percentage
Level-0	0,6728	0,9878	46,82%	0,6157	0,8045	30,66%
Level-1	0,7153	0,9039	26,37%	0,6035	0,8053	33,44%
Level-2	0,6071	0,663	9,21%	0,5369	0,6369	18,63%
Level-3	0,406	0,4398	8,33%	0,3782	0,4081	7,91%

Remarks and Interpretations:

The Illustrations of probability density distributions of this experiment found at Figure 38 resemble the ones at the former experiment. When we examine the distribution statistics presented in Figure 39, the difference looks more distinct that both road class mean and standard deviation are decreased. Obviously, the reason of this difference is the introduction of censoring variable in the optimization process. The decrease observed in the road class mean is not favorable to road and non-road class discrimination, nevertheless; the standard deviation of the road class is also decreased at a remarkable rate which may provide better separation. The most accurate analysis can be conducted by comparing the gains in the objective function values from

Table 5. The censoring based optimization is clearly superior to the uncensored experiment according to the objective function value gain.

4.3.1.4 The Relationship between Road Width and Multi-Resolution Analysis

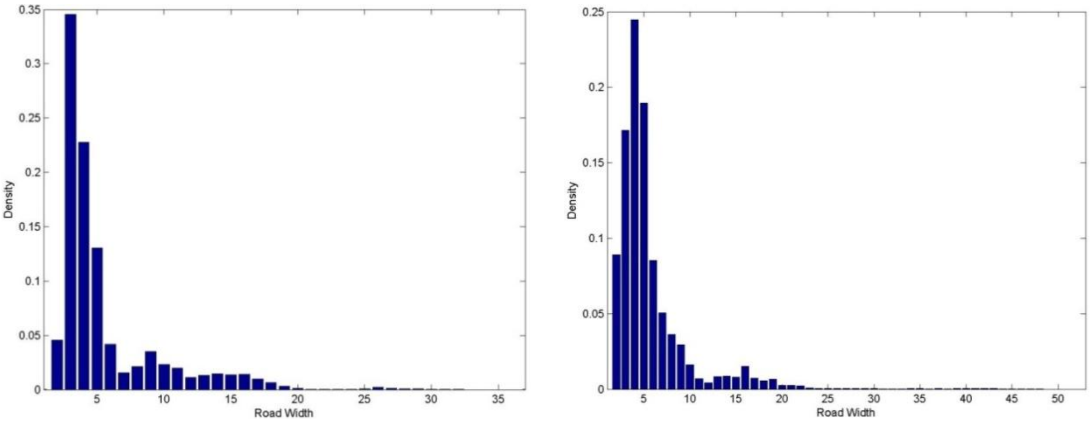
Purpose:

In this experiment, the relationship between road width and resolution levels is investigated. In 1-m satellite images, there may be various road width values, which are expected to be in range 1 to 50. Multi-resolution approach applying gradual resolution reduction employed in this study is expected to provide conformity for roads having particular width in each level. Therefore, extent of multi-resolution influence exposed to roads is needed to be clarified.

Procedure:

Firstly, we observe the road width distributions for both of the image sets. Then, the mean-shift segmentation algorithm is applied on both optimization and validation datasets with the optimum parameters obtained from the former experiment. For each possible value of road width, the average elongatedness score of the road pixels having a certain width value is calculated in all resolution levels. That is, the elongatedness values are grouped by averaging operator according to the road width.

Observations:

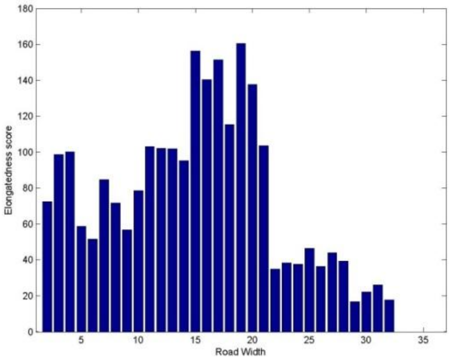


(a) (b)

Figure 41: Probability density distributions of widths of roads on the (a) optimization and (b) validation datasets.

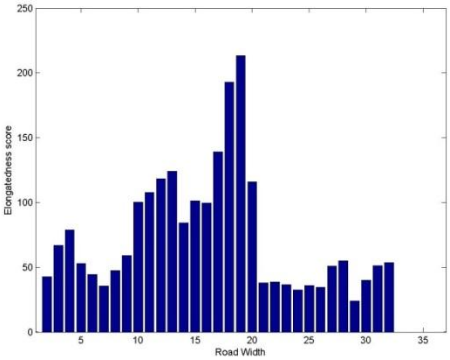
Table 6: Table of road width densities reported in a fixed interval.

Road Width	Density on Optimization Set	Density on Validation Set
1-5 m	0.7493	0.6947
5-10 m	0.1364	0.2178
10-15 m	0.0721	0.0362
15-20 m	0.0343	0.0369
20-25 m	0.0022	0.0068
25-30 m	0.0051	0.0023
30-35 m	0.0004	0.0014
Greater than 35 m	N/A	0.0037



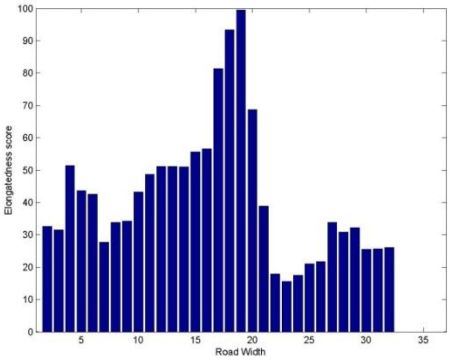
Road Width	Average Elongatedness Score on Optimization Set	Average Elongatedness Score on Validation Set
1-5 m	82.5787	72.7323
5-10 m	68.7495	63.3961
10-15 m	111.726	71.8873
15-20 m	141.086	131.9201
20-25 m	52.203	66.5068
25-30 m	31.801	31.6684
30-35 m	21.851	23.8819
Greater than 35 m	N/A	15.7976

(a)

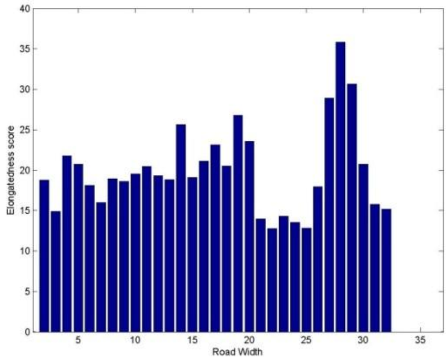


Road Width	Average Elongatedness Score on Optimization Set	Average Elongatedness Score on Validation Set
1-5 m	60.4015	59.5434
5-10 m	57.3289	48.0894
10-15 m	107.2169	53.1197
15-20 m	152.2776	86.7555
20-25 m	36.4292	69.1810
25-30 m	40.9465	45.8065
30-35 m	52.5571	34.8420
Greater than 35 m	N/A	12.5570

(b)



Road Width	Average Elongatedness Score on Optimization Set	Average Elongatedness Score on Validation Set
1-5 m	39.8022	33.1173
5-10 m	36.3392	28.7587
10-15 m	51.5815	27.0464
15-20 m	79.9040	58.8991
20-25 m	22.1719	41.9091
25-30 m	28.8629	18.4284
30-35 m	25.8644	28.0213



Road Width	Average Elongatedness Score on Optimization Set	Average Elongatedness Score on Validation Set
1-5 m	19.0445	18.7807
5-10 m	18.2381	17.7985
10-15 m	20.6738	24.2749
15-20 m	23.0227	36.8910
20-25 m	13.4930	27.6103
25-30 m	26.8267	24.2754
30-35 m	15.4772	27.7071

Greater than 35 m	N/A	20.5138

(c)

Greater than 35 m	N/A	15.5171

(d)

Figure 42: road width vs. elongatedness graphs for different resolution levels calculated on the optimization set (a) level-0 (b) level-1 (c) level-2 (d) level-3

Table 7: Comparison of score losses according to level-0 on optimization set

Road Width	Level-0 Average Elongatedness Score	Level-1 Average Elongatedness Score	Level-2 Average Elongatedness Score	Level-3 Average Elongatedness Score	Level-1 Elongatedness Score Loss	Level-2 ElongatednessScore Loss
1-5 m	82,5787	60,4015	39,8022	18,2381	26,86%	51,80%
5-10 m	68,7495	57,3289	36,3392	20,6738	16,61%	47,14%
10-15 m	111,7260	107,2169	51,5815	23,0227	4,04%	53,83%
15-20 m	141,0860	152,2776	79,9040	13,4930	-7,93%	43,37%
20-25 m	52,2030	36,4292	22,1719	26,8267	30,22%	57,53%
25-30 m	31,8010	40,9465	28,8629	15,4772	-28,76%	9,24%
35-40 m	21,8510	52,5571	25,8644	18,2381	-140,52%	-18,37%
Greater than 40 m	N/A	N/A	N/A	N/A	N/A	N/A

Table 8 : Comparison of score losses according to level-0 on validation set

Road Width	Level-0 Average Elongatedness Score	Level-1 Average Elongatedness Score	Level-2 Average Elongatedness Score	Level-3 Average Elongatedness Score	Level-1 Elongatedness Score Loss	Level-2 Elongatedness Score Loss	Level-3 Elongatedness Score Loss
1-5 m	72,7323	59,5434	33,1173	18,7807	18,13%	54,47%	74,18%
5-10 m	63,3961	48,0894	28,7587	17,7985	24,14%	54,64%	71,92%
10-15 m	71,8873	53,1197	27,0464	24,2749	26,11%	62,38%	66,23%
15-20 m	131,9201	86,7555	58,8991	36,891	34,24%	55,35%	72,04%
20-25 m	66,5068	69,181	41,9091	27,6103	-4,02%	36,99%	58,48%
25-30 m	31,6684	45,8065	18,4284	24,2754	-44,64%	41,81%	23,35%
35-40 m	23,8819	34,842	28,0213	27,7071	-45,89%	-17,33%	-16,02%
Greater than 40 m	15,7976	12,557	20,5138	15,5171	20,51%	-29,85%	1,78%

Remarks and Interpretations:

In Figure 41, the road width distributions for optimization and validation datasets is illustrated, and the density values of this distributions are presented in Table 6. From these it is inferred that majority of the roads are thinner than 5m as the percentage of the roads in interval 1-5m is about 75% for optimization set and %70 for validation set. Moreover, 1-10m roads have the percentage of 89% for optimization set and 91% for validation set.

The objective scores gained by different width groups during the feature extraction process are summarized in the Figure 42 for optimization set. The road width vs. elongatedness graphs indicates the variation in objective scores among different resolution levels. In high resolution thinner roads have greater score values, and by decreasing resolution the elongatedness scores achieved by thinner roads also decrease. A more comprehensible analysis can be accomplished by considering the percentages of score losses observed in the resolution reduction process. Table 7 and Table 8 show the score loss percentages calculated based on level-0 for optimization and validation sets, respectively. In optimization set, the loss of level-0 to level-1 resolution reduction for 1-5m roads is 27% and, and the loss of level-0 to level-2 resolution reduction for the same roads is about 52%. The loss of elongatedness scores for 1-5m roads continues to increase in subsequent levels, so we can infer that the elongatedness scores achieved by thinner roads tend to decrease in lesser resolutions. Nevertheless, the loss encountered for thin roads is not valid for wider roads in the context of multi-resolution analysis. For example, in optimization set, the loss of level-0 to level-1 resolution reduction for 30-35m roads is about -140% implying an elongatedness score gain, and this gain also increases in the level-0 to level-2 reduction.

This experiment proves the suitability of multi-resolution approach for the road extraction problem. In different resolution levels, different parts of the roads become more identifiable. That is, thin roads pose more distinguishable form in high resolution, whereas, the lesser resolutions favor in the detection of wider roads.

4.3.1.5 Analysis of Multi-spectral Band Number for Segmentation

Purpose:

The purpose of this experiment is to investigate the multi-spectral data combinations used in mean-shift segmentation algorithm, and the optimal band number is needed to be determined. Although using as much as data possible seems to generate better segmentation results, (Comaniciu & Meer, 2002) state that high-dimensionality may cause the empty space phenomenon as the kernel density estimation function tends to concentrate on a small region of the space. They also suggest employing a projection pursuit to reduce data dimension if the feature space has more than six dimensions. Therefore, the ideal number and combination of spectral features to use in the mean-shift segmentation routine is needed to be addressed.

Procedure:

In our dataset, the multispectral images have four bands including R (red), G (green), B (blue) and NIR (near infrared). By using the optimal parameter sets for both optimization and validation sets, we execute the mean-shift segmentation algorithm with different combinations of the four bands. After each run, objective scores for each resolution level is computed and reported in Table 9.

Execution Environment Configuration:

Observations:

Table 9: Spectral band combinations and their corresponding objective scores based on the resolution levels.

	On optimization set				On validation set			
	Level-0	Level-1	Level-2	Level-3	Level-0	Level-1	Level-2	Level-3
R	0.6929	0.5916	0.3824	0.2609	0.6208	0.5601	0.4128	0.2639
G	0.7424	0.5528	0.3772	0.2616	0.6040	0.5428	0.3883	0.2409
B	0.7103	0.5551	0.4185	0.2434	0.5868	0.5319	0.3734	0.2300
NIR	0.6288	0.5384	0.3767	0.2128	0.5426	0.5056	0.3653	0.2536
R-G	0.8506	0.7762	0.6284	0.3880	0.7043	0.7869	0.5558	0.3770
R-B	0.9127	0.8295	0.6233	0.4053	0.7504	0.7380	0.5379	0.3652
R-NIR	0.9331	0.8936	0.7050	0.4248	0.7469	0.7647	0.5715	0.3905
G-B	0.8689	0.8007	0.6182	0.3935	0.6760	0.7144	0.5309	0.3399
G-NIR	0.9660	0.8785	0.6671	0.3784	0.7503	0.8267	0.6137	0.4088
B-NIR	0.9112	0.8381	0.6070	0.3571	0.7241	0.7712	0.5837	0.3692
R-G-B	0.9113	0.7530	0.5149	0.2903	0.7044	0.6920	0.5012	0.3027
R-G-NIR	0.9438	0.8273	0.6290	0.3860	0.7330	0.7674	0.5984	0.3392
R-B-NIR	0.9419	0.8502	0.6134	0.3861	0.7207	0.7234	0.5463	0.3691

G-B-NIR	0.9193	0.8597	0.6041	0.3781	0.7441	0.7173	0.6004	0.3618
R-G-B-NIR	0.9878	0.9039	0.6630	0.4398	0.8045	0.8053	0.6369	0.4081

__Remarks and Interpretations:__

From Table 9, the R-G-B-NIR combination outperforms all other band combinations regarding to achieved objective score in all levels. Therefore, we conclude that the empty space phenomenon and curse of dimensionality are not valid for the road segmentation problem using four dimensions of data. The optimal band combination for mean-shift image segmentation is determined to be R-G-B-NIR.

4.3.1.6 Multi-resolution Level Number Analysis

__Purpose:__

In this experiment, we aim to investigate contributions of multi-resolution levels in data discrimination as well as deciding optimum number of resolution levels to be employed in the final configuration. The reduction process in multi-resolution analysis gradually eliminates the details from the image implying that after a certain extent of reduction, the roads become indistinguishable from the environment. Therefore, we need to decide an optimum level number to utilize.

__Procedure:__

Multi-resolution reduction is applied to the input image up to level six, and the mean-shift algorithm is employed to segment the resolution levels. Then, objective scores are computed for each level.

__Execution Environment Configuration:__

Segmentation parameter sets: optimum sets

Number of spectral bands for segmentation: 4 (R,G,B,NIR)

Feature vector for classification: structural features are about to be optimized

Number of decision trees for classification: 100

Minimum number of observations for decision tree: equals to number of training data

<u>Observations:</u>

Table 10: Objective values achieved in each resolution level.

	On optimization set	On validation set
Optimum Objective Value in Level-0	0.9878	0.8045
Optimum Objective Value in Level-1	0.9039	0.8053
Optimum Objective Value in Level-2	0.6630	0.6369
Optimum Objective Value in Level-3	0.4398	0.4081
Optimum Objective Value in Level-4	N/A	0.2544
Optimum Objective Value in Level-5	N/A	N/A
Optimum Objective Value in Level-6	N/A	N/A

<u>Remarks and Interpretations:</u>

In Table 10, objective scores implied by each resolution level are summarized. The N/A indicates that the non-road class mean exceeded the road class mean, so the objective value calculation procedure stopped. This phenomenon occurred due to lack of road and non-road segment discrimination, in other words; the resolution is so low to generate a promoting segmentation result. Up to level-3, there is class separation suitable for the classification; therefore, we decide to employ structural features level-0 to level-3.

4.3.1.7 Generalizing the Parameter Selection Procedure for Mean-shift

<u>Purpose:</u>

In this experiment, we aim to combine insights gained from the former experiments conducted on mean-shift parameter optimization in order to obtain a global segmentation parameter set valid for 1-m multispectral satellite images. The concept of generalization can be accounted as an extension to available investigations on the mean-shift parameter optimization with the objective of reducing them into a less-specific criterion. Taking our optimization and validation sets into consideration, they are actually composed of sample sets of images, indeed; we cannot expect a specific image set to represent whole population of satellite data. In ideal case, all available IKONOS satellite images must be included in the optimization set to achieve ideal parameterization, yet it is not a feasible solution due to time and cost constraints. The optimization and validation sets allow particular parameter sets to exhibit their behavior but a globally optimum parameter set valid for any arbitrarily compromised image set can be determined by empirical evaluation. Firstly, the observations must be extended by introducing new image sets which is expected to provide other local optimum locations. In fact, the observations attained from the parameter optimization experiments carried out up to this section imply the local optimum locations of the parameter space. Based on the available local optimum points, we aim to arbitrate the construction of an empirically determined global optimum parameter set.

Procedure:

Experiment results deduced in the former experiments are gathered. In addition to the optimization and validation sets, two extra validation image sets are introduced to generate new observations. Like the experiment 3.1.4, the pattern search algorithm is utilized to find optimal parameter sets for two new validation sets. Therefore, four local optimum locations pointed by the experiments conducted on four different image sets are obtained. As the local optimums are actually basins of attraction, they indicate the lower and upper bounds of the global optimum. In this manner, we can enforce the pattern search algorithm to explore the space bounded by local optimums resulting in an empirically determined global optimum. Although the global optimum found by this approach produces sub-optimal solutions for the individual image sets, it describes an optimal and general solution for the parameter determination problem.

Observations:

Table 11: Parameter sets and their corresponding optimum scores obtained by optimization procedure for different image sets, the right-most column indicates the generalized parameters

		Optimization Set	**Validation Set-1**	**Validation Set-2**	**Validation Set-3**	**After Generalization (Ø)**
Level-0	Spatial Bandwidth	*4*	*4*	*4*	*8*	8
	Range Bandwidth	*6*	*4*	*4*	*10*	4
	Minimum Region Area	*383*	*697*	*675*	*481*	481
	Score Limit	*64*	*48*	*84*	*46*	56
	Optimum Obj Value	*0.9878*	*0.8045*	*0.8739*	*0.6413*	0.8313
	Obj Value	*0.9649*	*0.8020*	*0.8698*	*0.6175*	N/A

	for set Ø					
Level-1	Spatial Bandwidth	*4*	*10*	*10*	*8*	8
	Range Bandwidth	*4*	*4*	*4*	*4*	4
	Minimum Region Area	*101*	*125*	*179*	*101*	171
	Score Limit	*56*	*48*	*56*	*48*	55
	Optimum Obj Value	*0.9039*	*0.8053*	*0.8311*	*0.6027*	0.7660
	Obj Value for set Ø	*0.8528*	*0.7857*	*0.7903*	*0.5808*	N/A
Level-2	Spatial Bandwidth	*8*	*4*	*10*	*4*	4
	Range Bandwidth	*8*	*4*	*10*	*8*	8
	Minimum Region Area	*55*	*43*	*59*	*97*	85
	Score Limit	*54*	*38*	*44*	*48*	53
	Optimum Obj Value	*0.6630*	*0.6369*	*0.6421*	*0.4775*	0.5881
	Obj Value for set Ø	*0.6502*	*0.6004*	*0.5893*	*0.4772*	N/A
Level-3	Spatial Bandwidth	*8*	*4*	*8*	*10*	10
	Range Bandwidth	*4*	*4*	*8*	*6*	6
	Minimum Region Area	*21*	*21*	*21*	*21*	21
	Score Limit	*26*	*26*	*30*	*28*	27
	Optimum Obj Value	*0.4398*	*0.4081*	*0.3716*	*0.3444*	0.3983
	Obj Value for set Ø	*0.4314*	*0.4060*	*0.3665*	*0.3446*	N/A

Remarks and Interpretations:

In this experiment, we adopt an inductive reasoning approach to reach a global optimal solution from the given local optimum parameter patterns. That is, the final optimization procedure conducted by the pattern search algorithm is deployed to explore the search region bounded by the local optimums. The ultimate results of the mean-shift parameter optimization efforts are summarized in Table 11, and the global optimum parameter set is presented at the right-most column of this table. Note that the optimum objective score achieved by the global optimum is lower than some of the optimum values suggested by former experiments, and likewise some of the local optimum scores are lower than the global optimum score. Moreover, the generalized parameter set, $\emptyset$, is experimented on each of the image sets to observe performance differences. As the objective class separation scores are decreased slightly, the generalized parameter set provides satisfactory generalization for the mean-shift image segmentation procedure. This phenomenon indicates that each set of images has its own ideal parameter set, nevertheless; the variances of the objective scores achieved by each ideal parameter set are considerably low.

4.3.1.8 Band Ratios Combination Analysis

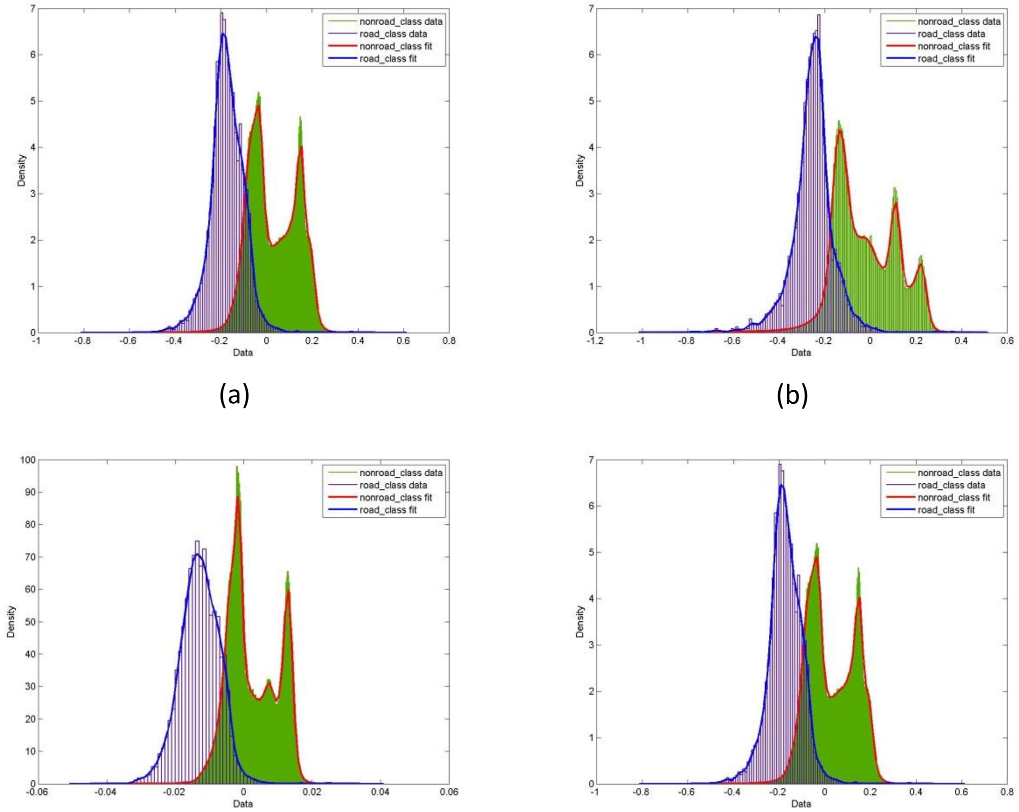
Purpose:

The multi-spectral imagery containing NIR bands provides the capability of object identification based on their common spectral characteristics. In this study, well-known band ratios are investigated to find the optimal spectral ratio describing the non-road regions effectively. The purpose of this experiment is to decide on a set of spectral band ratios to use them as classification features. Three popular band ratios are determined as candidate features including NDVI, NDWI, SAVI and EVI.

Procedure:

Firstly, band ratios are computed by using their corresponding rationing formulas defined in equations (2.1), (2.2), (2.3) and (2.4). Then, their Pearson product-moment correlation coefficients (PPMCC) are determined to measure their inter-similarities. Their objective scores are also calculated to find out their road and non-road class separation capabilities. All of the observations are generated both on the optimization and validation sets. SAVI's L value is set to 0.5, and EVI has the following configuration C1 = 6, C2 = 7.5 and G = 2.5.

Observations:



(a) (b)

(c) (d)

Figure 43:Probability density distributions of road and nonrodad classes for (a)NDVI (b) NDWI (c)SAVI and (d) EVI (Note: these distributions are obtained from the image-2)

Table 12: Correlation matrix of the spectral band ratios calculated on optimization set

	NDVI	NDWI	SAVI	EVI
NDVI	1.0000	0.9292	0.9645	0.8924
NDWI	0.9292	1.0000	0.8817	0.8567
SAVI	0.9645	0.8817	1.0000	0.9726
EVI	0.8924	0.8567	0.9726	1.0000

Table 13: Correlation matrix of the spectral band ratios calculated on validation set

	NDVI	NDWI	SAVI	EVI
NDVI	1.0000	0.9573	0.9606	0.8967
NDWI	0.9573	1.0000	0.8959	0.8685
SAVI	0.9606	0.8959	1.0000	0.9764
EVI	0.8967	0.8685	0.9764	1.0000

Table 14: Objective score achieved by the spectral band ratios.

	On optimization set	On validation set
NDVI	0.5250	0.6320
NDWI	0.5229	0.6073
SAVI	0.7203	0.6935
EVI	0.6207	0.6133

Remarks and Interpretations:

In Figure 43, class distributions for the four band ratios are unveiled. From the distribution sketches, we can infer that the class distribution shapes are nearly the same. Correlation matrix of the three band ratios for the images in the optimization set is presented in Table 12. The minimum correlation coefficient value is above 0.85 implying nearly a complete uniformity among all four band ratios, and the correlation coefficients obtained from the validation set (see

Table 13) are not different than the ones obtained from the optimization set. Therefore, it is irrelevant to include more than one band ratio into the feature vector. By this way, the spectral feature selection procedure is reduced to the decision of best band ratio for the road extraction problem. In Table 14, the objective scores achieved by the spectral band ratios are given. The objective scores indicate the extent of class separation implied by the given spectral band ratios. The best score is achieved by SAVI, therefore; it is appropriate to include SAVI into feature vectors.

4.3.1.9 Optimal Feature Extraction Component Parameters

Up to this point, we have completed to optimize all parameters and decided ideal configuration of the feature extraction component. In Table 15, the final feature extraction component configuration is summarized.

Table 15:Final configuration of the feature extraction component

	Configuration
Number of bands used in mean-shift image segmentation	4 (R-G-B-NIR)
Number of multi-	4

resolution levels					
Optimum parameter sets		**Level-0**	**Level-1**	**Level-2**	**Level-3**
	Spatial Bandwidth	8	8	4	10
	Range Bandwidth	4	4	8	6
	Minimum Region Area	481	171	85	21
	Score Limit	56	55	53	27
Spectral features employed	SAVI				

4.3.2 Experiments on Adaboost Classification Component

The quality of the training procedure is controlled by two parameters; complexity of weak learners and total number of weak learners. In these series of experiments, we aim to find out the optimal parameters for the classification component, also the advantages and disadvantages of Adaboost algorithm for the road extraction problem is intended to be investigated. The final strong classifier generated by Adaboost should not be too complex resulting in an overfitting phenomenon. Similarly, an ensemble of too simple weak classifiers tends to cause underfitting, deficiency in class discrimination. In machine learning, overfitting is related to the complexity of the classifier since the increased complexity causes the training model to depend on too many parameters relative to a specific training set. Nevertheless, the classifier is desired to have high generalization capability in order to assess good performance on the test set consisting of unseen observations. In Adaboost, the generalization bounds are examined by considering the minimum margin value. Firstly, we find out the ideal weak classifier complexity for this component, and then we observe the ensemble behavior in terms of variable iteration number.

4.3.2.1 Weak Classifier Complexity Analysis

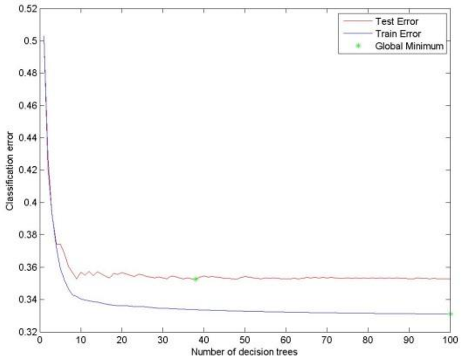
Purpose:

In this experiment, the complexity of weak learners is examined to investigate its effects on the classification performance. In our study, classification and regression trees, a type of decision trees, are employed to generate weak hypothesis in the Adaboost training procedure. In our implementation, the depths of the trees are controlled by a single parameter, “MinLeaf”, indicating the minimum numbers of observations are belonging to per leaf node. That is, the node with observations more than “MinLeaf” number is split and two new children node is constructed for the corresponding node in order to distribute the observations among the new nodes. The tree construction procedure continues in this way increasing the tree depth until all leaf nodes have observations less than “MinLeaf” parameter. Therefore, this parameter allows us to determine the complexity of the decision trees which are reported to be influential in classification performance (Freund & Schapire , 1997). On the other hand, high level of complexity in weak learners triggers the overfitting problem in ensemble methods.

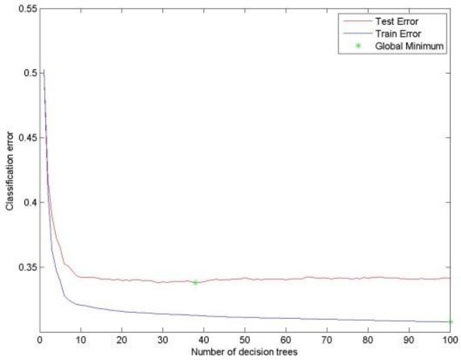
Procedure:

The ideal complexity level is determined by experimenting the several complexity degrees for the weak learners. Assuming N is the total number of observations used for training a strong classifier, and the “MinLeaf” parameter is fed into the classification routine by dividing N by a variable ranging 1 to 12 in each run. Each learning procedure takes 100 iterations to finish. After the training procedure is completed, statistical data regarding to the experiment is evaluated.

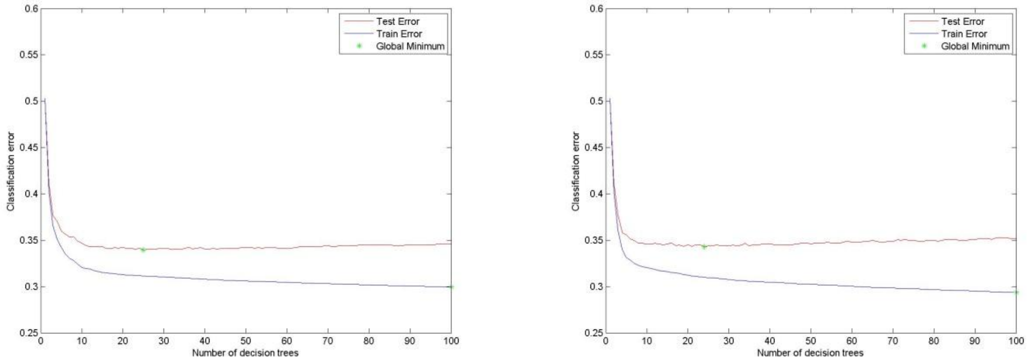
Observations:



(a)



(b)



(c) (d)

Figure 44:Classification error curves for weak learners having different complexities (a) N/3 (b) N/6 (c) N/9 (d) N/12

Table 16: Learning statistics on training set

Complexity of Decision Tree	Minimum Classification Error	Number of Decision Trees provide minimum error	Minimum Margin
N	0.403819	49	-1
N/2	0.403819	49	-1
N/3	0.330938	100	-0.8931
N/4	0.318149	100	-0.8349
N/5	0.311947	100	-0.9556
N/6	0.307633	100	-0.9166
N/7	0.304691	100	-0.8890
N/8	0.301044	100	-0.8522
N/9	0.299367	100	-0.9010
N/10	0.297081	100	-0.9394
N/11	0.295151	100	-0.8868
N/12	0.293407	100	-0.8355

Table 17:Learning statistics on test set

Complexity of Decision Tree	Minimum Classification Error	Number of Decision Trees provide minimum error	Minimum Margin
N	0.416177	5	-1
N/2	0.417603	5	-1
N/3	0.352446	38	-0.9326
N/4	0.344487	77	-0.9466
N/5	0.342926	77	-0.9707
N/6	0.338105	38	-0.9545
N/7	0.340038	26	-0.9567
N/8	0.340073	38	-0.9534
N/9	0.339623	25	-0.9655
N/10	0,343041	18	-0.9553
N/11	0,34247	12	-0.9393

N/12	0,342972	24	-0.9537

Remarks and Interpretations:

In Figure 44, the classification error curves are pictured for different weak learner complexities, and the global minimum values are painted in green with “*” sign. In all cases, the training error poses steady decrease while the behavior of the test error is observed to be dependent on the weak classifier complexity. That is, in low complexity levels, like the test error the training error is subjected to a slow decrease by the iteration number. However, high complexity levels like 11 and 12 in the weak classifiers causes the test error to rise after a pit in the error curve. This phenomenon is known to be a basic indicator of overfitting problem, and it can be observed on the complexity levels after 9. Therefore, we need to choose a complexity degree parameter which is not susceptible to cause overfitting, also this parameter is expected to provide lowest test error. In *Table 16* and Table 17, learning statistics on training and testing sets are presented corresponding to this experiment. In runs N and N/2, the minimum margin values are equal to minus one indicating the early termination of the learning due to lack of appropriate weak learners generating error lower than 0.5. The other minimum margin values show local fluctuations resulting in very close values, which mean that the Adaboost algorithm is unable to generalize the given observations by variable weak classifier complexity. The non-road regions with high elongatedness values and road like spectral features simulate a uniform class label noise situation in this training procedure. On the other hand, the minimum test error is obtained in N/6 case, therefore; the optimum weak classifier complexity parameter is determined to be N/6.

4.3.2.2 Number of Weak Classifiers Analysis

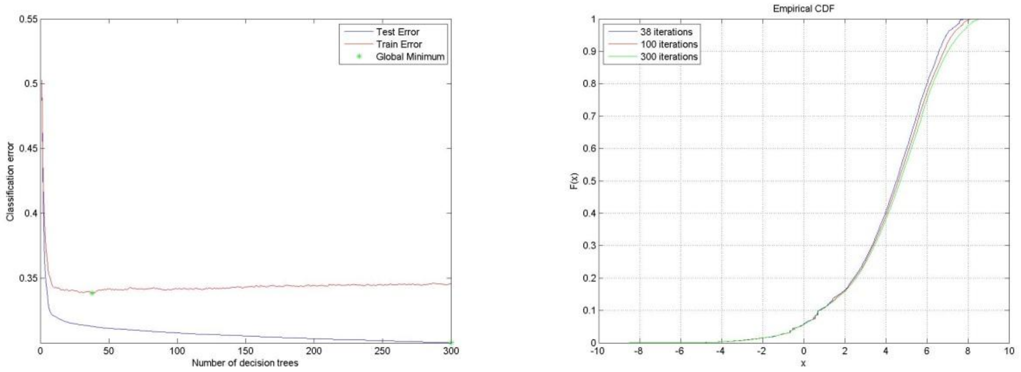
Purpose:

In the former experiment, we find out the ideal complexity level of the weak learners for the training set. In this experiment, we aim to investigate the influence of the number of weak classifiers into the strong classifier’s performance. At each learning step, Adaboost generates a weak hypothesis from the pool of weak classifiers, and the final combination of these weak learners constitutes the strong classifier. Therefore, Adaboost gradually concentrates on harder examples in a greedy learning scheme. The addition of new weak learners into the final classifier is expected to improve the prediction power, however; excessive number of weak classifiers also increases the complexity of the final strong classifier. That is, the ideal number of weak classifiers is needed to be determined.

Procedure:

The Adaboost algorithm is executed with “MinLeaf” parameter of N/6 up to iteration 300. Error curves of the learning and empirical cumulative density functions of margin distributions are extracted for each iteration.

Observations:



(a) (b)

Figure 45:Observations of the experiment (a) error curves (b)Emprical CDFs

Remarks and Interpretations:

In Figure 45-a, the sketch of the number of iterations vs. training and testing error is pictured. As it can be inferred from this graph, the training error continues to drop gradually while the iteration number counts up. However, the testing error does not improve after iteration 38, in fact; there is small noise like fluctuations in the testing error. On the other hand, achieved generalization by the final strong classifier can be examined by the empirical CDFs presented in Figure 45-b. The classifier margin distribution is improved by increasing iteration number resulting in a better generalization bound. Nevertheless, the improvements are arguably limited in a small range, and the tradeoff between the classifier complexity and the generalization error bound should be considered. Indeed, the empirical observations suggest that the best classifier for our test set is produced at 38th iteration, and the generalization error bound is improved by further iterations with the loss of classification accuracy. Therefore, it is convenient to set iteration number to 38 as the simplest final classifier with minimum error is provided at that step.

4.3.3 Experiments on Road Detection Component

Firstly, the relevance of histogram equalization operation applied on the probability map is empirically proved. Then, road extraction parameters for three distinct extraction strategies are optimized by using pattern search algorithm. In the optimization procedure, F-measure is utilized to score extraction results, that is; a pixel-wise performance evaluation is made on the road mask generated by the extraction and the ground truth of the corresponding image. In our road extraction experiments, we tune the extraction parameters for different β values of the F-measure including $F_{0.5}$, F_1 and F_2 in order to reveal different precision and recall tradeoffs among different road extraction techniques. For example, using $F_{0.5}$ measure as the objective function in pattern search forces the optimization routine to give more importance on the parameter sets providing better precision values rather than recall. Conversely, the objective function using F_2 measure emphasis on the recall values instead of precision. A balanced precision and recall tradeoff is obtained by using F_1 measure in objective function's scoring scheme. Furthermore, the objective functions calculate F-measures for the images from the training set of the Adaboost classification component, so the parameter sets are optimized only for the same training set with the classification component. Likewise, the performance comparisons are conducted on the measurements obtained from the testing set. The performances of the road extraction methods operated on the probability map are compared in terms of precision, recall and F-measure.

4.3.3.1 Histogram Equalization Analysis

Purpose:

We employ histogram equalization routine to enhance the probability map's contrast. To validate relevance of this technique, we need to observe an increase in the class separation scores. This experiment is designed to show that the histogram equalization improves the discrimination of road and non-road classes.

Procedure:

The average class separation scores are measured for both train and test sets.

Observations:

Table 18:Effect of the histogram equalization on the class separation scores.

	On Train Set	On Test Set
Before Histogram Equalization	1.0239	1.1443
After Histogram Equalization	1.0908	1.1987
Score Gain Percentage	6,53%	4,75%

Remarks and Interpretations:

In Table 18, class separation scores for train and test sets are unveiled. Obviously, applying histogram equalization operation on the probability map comes with better class separation scores, so it is convenient to use histogram equalization before extracting roads from the probability map.

4.3.3.2 Hard Thresholding Parameter Tuning

Purpose:

In this experiment, the optimal hard threshold value utilized for road extraction from the probability map is determined.

Procedure:

The pattern search algorithm is employed to tune the threshold parameter. Three alternative objective functions are considered in the optimization.

Observations:

Table 19:Result of parameter tuning on training set for hard thresholding

		$F_{0.5}$	F_1	F_2
Precision	Mean	0.7326	0.6412	0.5008
	Std	0.1821	0.2129	0.2105
Recall	Mean	0.3879	0.5251	0.6927
	Std	0.1574	0.1405	0.1011
Accuracy	Mean	0.9149	0.9150	0.8932
	Std	0.0407	0.0277	0.0154
F	Mean	0.5799	0.5376	0.6075
	Std	0.0852	0.0801	0.0725

Threshold Value	0.9656	0.9344	0.8719

Remarks and Interpretations:

In Table 19, the optimal threshold values suggested by the optimization procedure are presented with the performance measurements obtained by that threshold value. The best precision value, 0.73, appears in the parameter set optimized by the $F_{0.5}$ measure as expected, and the best recall value, 0.69, is accomplished by the optimization configuration using F_2 type objective function. The road extraction accuracy values are relatively close to each other with small standard deviation values.

The best objective score is accomplished by F_2 measure implying that false positive detection is more probable than false negative detection. In other words, the road extraction strategy tends to identify non-road regions as roads rather than labeling road regions as non-road regions.

4.3.3.3 Hysteresis Thresholding Parameter Tuning

Purpose:

The hysteresis thresholding based road extraction method employed in this study requires two different threshold values to detect road mask. The pixels with probability values greater than the second threshold are labeled as roads, and the pixels with the probability values in the range bounded by first and second thresholds are identified as road if the pixel has a neighbor already labeled as road. It is obvious that both of the threshold parameters are directly influencing the performance of road detection. Therefore, the purpose of this experiment is to decide on the optimum threshold values.

Procedure:

The pattern search algorithm is utilized to find out the optimum threshold parameters. Three types of objective functions are considered in the optimization.

Observations:

Table 20:Result of parameter tuning on training set for hysteresis thresholding

		$F_{0.5}$	F_1	F_2
Precision	Mean	0.6596	0.6364	0.5651

	Std	0.2070	0.2144	0.2088
Recall	Mean	0.5273	0.5500	0.6519
	Std	0.1532	0.1415	0.1085
Accuracy	Mean	0.9173	0.9161	0.9095
	Std	0.0279	0.0241	0.0188
F	Mean	0.5971	0.5498	0.6041
	Std	0.1343	0.0826	0.0658
Threshold Value-1		0.9187	0.9187	0.8719
Threshold Value-2		0.9875	0.9719	0.9875

Remarks and Interpretations:

In Table 20, performance measurements regarding to the experiment are presented in addition to the optimum parameters detected by the pattern search algorithm. Like the former experiment conducted on hard thresholding based road extraction method, this experiment has also the same interpretations. The precision values among different runs employing different objective functions are observed to decrease from left to right columns in the same row. Similar behavior is observed for the recall measurements except that they increase from left to right columns.

4.3.3.4 Region Growing Based Road Detection Parameter Tuning

Purpose:

The region growing based road detection strategy extracts road separate parts and validates their elongatedness properties satisfying the topological connectivity constraint. This method requires three distinct parameters to operate; region growing similarity value, minimum peak value and minimum elongatedness value. These parameters are needed to be optimized to obtain more accurate extraction results.

Procedure:

Like the former experiments conducted on the road extraction strategies, we employ the pattern search algorithm to identify optimal parameter set for the region growing based road detection method. Three different objective function alternatives are used for optimization.

Observations:

Table 21:Result of parameter tuning on training set for region growing based method

		$F_{0.5}$	F_1	F_2
Precision	Mean	0.6497	0.6433	0.6080
	Std	0.1957	0.1884	0.1838
Recall	Mean	0.5052	0.5468	0.5655
	Std	0.1602	0.1461	0.1443
Accuracy	Mean	0.9156	0.9186	0.9152
	Std	0.0300	0.0262	0.0258
F	Mean	0.5835	0.5578	0.5527
	Std	0.0820	0.0785	0.0973
Region growing similarity value		0.0500	0.0551	0.0529
Minimum peak value		0.9781	0.9781	0.9313
Minimum elongatedness value		30	30	30

Remarks and Interpretations:

Table 21 includes the performance measurements of the experiment, and the optimum parameter set suggested by the optimization procedure is also presented. The same inferences for the former road extraction experiments are also valid for this experiment.

4.3.3.5 Comparison of Road Detection Method’s Performances on Test Set

Purpose:

In the sections 4.3.3.2, 4.3.3.3 and 4.3.3.4, we determine the optimal parameter sets for different road extraction strategies, and the performance measurements are also presented for the corresponding strategies. On the other hand, the road extraction method selection tradeoffs and their relative advantages or disadvantages are needed to be discussed. In this section, we compare the road extraction methods in terms of their achieved performance on the test set. Moreover, visual observations of the road masks detected by different road extraction strategies are issued.

Procedure:

In this experiment, the observations are obtained by running three different road extraction strategies with their optimal parameter sets on the testing set.

Observations:

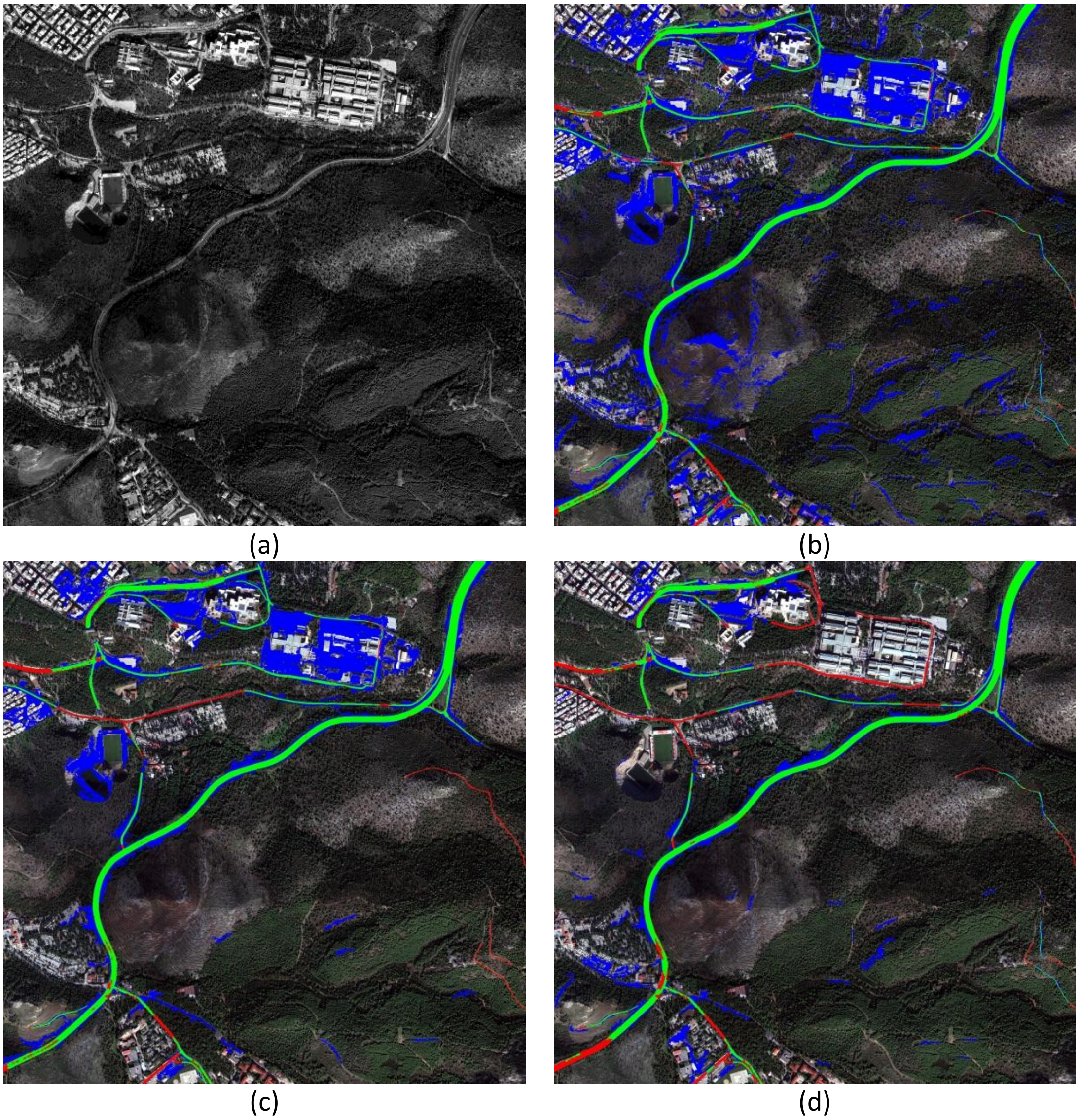


Figure 46: Visual results obtained by different road extraction strategies. (a) original image, and road extracted by (b) hard thresholding (c) hysteresis thresholding (d) region growing (Note: green regions represent true positives, blue regions are false positives and red regions indicate false negatives)

Table 22:Performance measurements with tuned parameters for three different road extraction strategies. (Note: the best objective scores achieved by optimization procedure are shaded.)

		$F_{0.5}$	F_1	F_2
		Hard Thresholding		
Precision	Mean	0.6133	0.5403	0.4188
	Std	0.2172	0.2298	0.2211
Recall	Mean	0.3734	0.4942	0.6540
	Std	0.1235	0.1268	0.1187
Accuracy	Mean	0.9178	0.9113	0.8820

	Std	0.0382	0.0288	0.0251
F	Mean	0.5136	0.4768	0.5483
	Std	0.1408	0.1103	0.1116
Hysteresis Thresholding				
Precision	Mean	0.5439	0.5268	0.4739
	Std	0.2176	0.2185	0.2220
Recall	Mean	0.4698	0.4999	0.5812
	Std	0.1653	0.1707	0.1734
Accuracy	Mean	0.9102	0.9087	0.8976
	Std	0.0360	0.0340	0.0317
F	Mean	0.4815	0.4637	0.5168
	Std	0.1439	0.1519	0.1516
Region Growing Based Road Extraction				
Precision	Mean	0.5745	0.5596	0.5211
	Std	0.2030	0.2016	0.1888
Recall	Mean	0.4826	0.5037	0.5224
	Std	0.1471	0.1474	0.1505
Accuracy	Mean	0.9217	0.9210	0.9176
	Std	0.0329	0.0315	0.0310
F	Mean	0.5315	0.5014	0.5049
	Std	0.1610	0.1334	0.1363

Remarks and Interpretations:

In Figure 46, the road masks extracted by different methods are illustrated with their true positives, false positives and false negatives. At first glance, the road mask suggested by hard thresholding takes attention for its completeness in the road network, nevertheless; this road mask has also has more false alarms than the other two methods. The smallest false detection rate is accomplished by the region growing based method, that is; it can be observed that the regions with low elongatedness properties are successfully eliminated. On the other hand, the hysteresis thresholding strategy comes with a lower false detection rate than the hard thresholding method. Visual inspection of output road masks pose only the image specific differences between the road extraction techniques. To understand the quantitative performance disparities, we need to analyze the performance measurements presented in Table 22. The best objective scores for $F_{0.5}$ and F_1 measures are achieved by the region growing based road extraction method. On the other hand, the hard thresholding method has the best objective score for F_2 measure. This difference is caused by the varying capabilities of the road extraction methods. For instance, the hard thresholding method tends to extract more road pixels than the other two methods, and this behavior results in a low precision with a high recall accomplishment. Conversely, the region growing based strategy validates the topological connectivity constraint for each extracted road segment, so it may fail in detecting roads jointed to the non-road regions with low elongatedness scores and high probability values. Furthermore, the hysteresis thresholding method does not provide better results than the other methods but it still promises a robust extraction performance.

The choice between the different road extraction strategies should be made by considering the application's needs. In fact, all of the methods seem to provide promising road detection results but they pose different behaviors for different situations. Moreover, the objective function tradeoffs should be taken into account in terms of precision and recall requirements.

4.4 Entire Model Based Experiments

Entire model based experiments aim to investigate and explain the component dependencies by evaluating them together. Our proposed road extraction algorithm is composed of three major components, and interactions of these components are needed to be revealed. In fact, the component interactions can be explained by the modifications applied on the middle component, the Adaboost classification component. Therefore, in these series of the experiments, the relevance of the Adaboost classification component is investigated by replacing it with different probability map construction approaches.

The Adaboost classification component utilizes a supervised technique to generate the probability map used in the road extraction component, so the contribution of the supervised learning routine is needed to be addressed. For this purpose, two different unsupervised probability map construction scheme is proposed. These schemes rely on the features provided by the feature extraction component to generate estimates for the probability map in an unsupervised manner. There are two different unsupervised techniques experimented in this study.

4.4.1 Unsupervised Probability Map Construction

Purpose:

The strong classifier suggested by the Adaboost algorithm can be considered as a linear combination of the given predictors, namely features as in equation (4.1)

$$f(\mathrm{x}_n) = a_n x_1 + b_n x_2 + c_n x_3 + d_n x_4 + e_n x_5 \tag{4.1}$$

where $f(\mathrm{x})$ is the strong classifier, x_n is the observation vector composed of predictors x_1, x_2, x_3, x_4 and x_5. Also, a_n, b_n, c_n, d_n and e_n are the predictor weights specifically generated for the observation x_n by Adaboost.

As it can be inferred from equation (4.1), the classification score is computed with a weighted sum of the given features, and the weights are determined by the Adaboost classifier. Since Adaboost utilizes decision trees as weak classifiers, different weight values are generated for different observation vectors.

__Procedure:__

In this experiment, we replace the predictor weights with constant weights in order to eliminate the need of supervised classification. The generalized objective separation scores, presented in Table 11, determined in the mean-shift parameter generalization experiment are employed as constant weights for the corresponding predictor. Moreover, the features need to be converted in a common distribution type in order to remove metrical differences among the feature values. That is, x_1 predictor contains the SAVI value of the corresponding pixel, in the range of [-1, 1], and x_2, x_3, x_4 and x_5 predictors contain elongatedness scores computed at different resolution levels, ranging from zero to infinity. Assuming the features form normal distribution with different mean and standard deviations, we can standardize these normal distributions by using equation (4.2).

$$s(x) = \frac{\mu - x}{\sigma} \tag{4.2}$$

where μ and σ are the mean and standard deviation of the feature distribution respectively, and $s(x)$ is the standardized normal distribution. Hence, the final form of the unsupervised probability map construction is obtained by placing the weights and standard normal distributions, see equation (4.3).

$$g(\mathrm{x}_n) = -0.7203\overline{x_1} + 0.8313\overline{x_2} + 0.7660\overline{x_3} + 0.5881\overline{x_4} + 0.3983\overline{x_5} \tag{4.3}$$

where $g(\mathrm{x})$ is the generated probability map, x_n is the observation vector composed of standardized predictors $\overline{x_1}$, $\overline{x_2}$, $\overline{x_3}$, $\overline{x_4}$ and $\overline{x_5}$. The explanation of these predictors can be seen in Table 23.

Table 23: Predictor explanations

x_1	SAVI
x_2	Elongatedness value at level-0
x_3	Elongatedness value at level-1
x_4	Elongatedness value at level-2
x_5	Elongatedness value at level-3

After construction of the probability map by this unsupervised approach, we also need to optimize road extraction parameters according to the probability maps generated by this method. The same parameter tuning techniques introduced in the section 4.3 are applied, and the corresponding optimization results are unveiled in Appendix-A.

Observations:

Table 24:Performance measurements on test set with tuned parameters for three different road extraction strategies. (Note: the best objective scores achieved by optimization procedure are shaded.)

		$F_{0.5}$	F_1	F_2
Hard Thresholding				
Precision	Mean	0.4575	0.3726	0.2943
	Std	0.2047	0.2042	0.1880
Recall	Mean	0.4049	0.5852	0.7231
	Std	0.1780	0.1681	0.1391
Accuracy	Mean	0.9297	0.9101	0.8695
	Std	0.0337	0.0208	0.0248
F	Mean	0.4141	0.4129	0.5067
	Std	0.1383	0.1268	0.1281
Hysteresis Thresholding				
Precision	Mean	0.4642	0.3747	0.3027
	Std	0.2055	0.2028	0.1937
Recall	Mean	0.3945	0.5869	0.7133
	Std	0.1767	0.1733	0.1438
Accuracy	Mean	0.9306	0.9114	0.8748
	Std	0.0344	0.0212	0.0243
F	Mean	0.4164	0.4171	0.5078
	Std	0.1498	0.1470	0.1425
Region Growing Based Road Extraction				
Precision	Mean	0.3997	0.3706	0.3372
	Std	0.1909	0.1883	0.1794
Recall	Mean	0.4871	0.5405	0.6102
	Std	0.2229	0.2068	0.1905

Accuracy	Mean	0.9219	0.9210	0.9019
	Std	0.0280	0.0269	0.0290
F	Mean	0.3725	0.4070	0.4870
	Std	0.1760	0.1701	0.1553

Remarks and Interpretations:

In Table 24, performance measurements on test set are presented regarding to this experiment. Although, there is not a remarkable achieved score difference between different road extraction strategies, the best F-measurements are suggested by the hysteresis thresholding technique.

4.4.2 Unsupervised Probability Map Construction (Refined)

Purpose:

In experiment 4.4.1, we proposed an unsupervised probability map construction method utilizing the feature separation scores. A refined version of this construction scheme can be obtained by thresholding the vegetation feature, SAVI in our case. In fact, many vegetation detection applications apply a hard threshold on SAVI in order to obtain vegetation mask, since the vegetative regions pose the same reflectance properties in the images captured by the same device. Therefore, we can use such prior information to eliminate vegetation regions and the combined structural features can be used as main descriptors for the roads.

Procedure:

Following the same theory elaborated in experiment 4.4.2, we construct a refined version of probability map by equation (4.4).

$$h(\mathrm{x}_n) = (0.8313x_2 + 0.7660x_3 + 0.5881x_4 + 0.3983x_5)\llbracket \mathrm{x}_1 > 0 \rrbracket \qquad (4.4)$$

where $h(\mathrm{x})$ is the constructed probability map, x_n is the observation vector composed of predictors x_1, x_2, x_3, x_4 and x_5. In this formulation, SAVI is thresholded with zero, that is; the pixels with a SAVI value higher than are set to zero in the final probability map. By this way, the vegetation regions are eliminated from the probability map.

The parameters of the road extraction strategies are also tuned for the probability maps generated by this unsupervised method. The result of the parameter optimization runs can be seen in Appendix-B.

Observations:

Table 25:Performance measurements on test set with tuned parameters for three different road extraction strategies(Note: the best objective scores achieved by optimization procedure are shaded.)

		$F_{0.5}$	F_1	F_2
Hard Thresholding				
Precision	Mean	0.5618	0.4920	0.3712
	Std	0.2159	0.2196	0.2090
Recall	Mean	0.3688	0.4774	0.6044
	Std	0.1967	0.1995	0.1987
Accuracy	Mean	0.9408	0.9370	0.9078
	Std	0.0339	0.0264	0.0421
F	Mean	0.4615	0.4424	0.4919
	Std	0.1685	0.1590	0.1682
Hysteresis Thresholding				
Precision	Mean	0.5633	0.5456	0.3945
	Std	0.2182	0.2222	0.2220
Recall	Mean	0.4089	0.4532	0.5928
	Std	0.2052	0.2063	0.2028
Accuracy	Mean	0.9418	0.9413	0.9138
	Std	0.0304	0.0276	0.0409
F	Mean	0.4780	0.4480	0.4929
	Std	0.1809	0.1733	0.1717
Region Growing Based Road Extraction				
Precision	Mean	0.6164	0.5752	0.4975
	Std	0.2112	0.2125	0.2158
Recall	Mean	0.4054	0.4483	0.4941
	Std	0.1854	0.2143	0.2113
Accuracy	Mean	0.9427	0.9391	0.9355
	Std	0.0301	0.0327	0.0311
F	Mean	0.5131	0.4505	0.4680
	Std	0.1827	0.1926	0.1768

Remarks and Interpretations:

In Table 25, the performance measurements on test set are presented. The achieved objective scores are very close to each other, however; the best objective scores for $F_{0.5}$ and F_1 functions are accomplished by the region growing based road extraction method. Also, the hysteresis thresholding detects road regions better than other two methods when the parameters tuned for F_2 function.

4.4.3 Comparison of Probability Map Construction Strategies

Purpose:

In sections 4.4.1 and 4.4.2, we proposed two unsupervised probability map construction methods. In this experiment, the supervised probability map construction method accomplished by Adaboost classification component and two unsupervised methods are compared in terms of their capabilities on class separation and influences on the road extraction performances.

Procedure:

In this experiment, two categories of observations are examined; class separation scores and road extraction performances. The probability maps generated by different construction strategies are evaluated by using the objective function defined in equation (3.1). These scores are useful for analyzing class separation capabilities of the suggested probability maps. On the other hand, the road extraction performance measurements are obtained from the former experiments conducted on the road extraction component tuning.

Observations:

Table 26:Objective separation scores achieved by different probability map construction strategies

	Class Separation Score
SAVI (x_1)	0.7203
Elongatedness at level-0 (x_2)	0.8313
Elongatedness at level-1 (x_3)	0.7660
Elongatedness at level-2 (x_4)	0.5881
Elongatedness at level-3 (x_5)	0.3983
Unsupervised-1 Probability Map	0.7460
Unsupervised-2 Probability Map	0.7759
Supervised Probability Map	0.9001

Table 27:Objective separation gain percentages achieved by different probability map construction strategies

	Gain Percentage
from Unsupervised-1 to Unsupervised-2	4,01%
from Unsupervised-1 to Supervised	20,66%
from Unsupervised-2 to Supervised	16,01%

Table 28:Average F-measures summarized for different probability map construction strategies. (Note: the best objective scores achieved by optimization procedure are shaded.)

	$F_{0.5}$	F_1	F_2
Hard Thresholding			
Unsupervised-1 F	0.4141	0.4129	0.5067
Unsupervised-2 F	0.4615	0.4424	0.4919
Supervised F	0.5136	0.4768	0.5483
Hysteresis Thresholding			
Unsupervised-1 F	0.4164	0.4171	0.5078
Unsupervised-2 F	0.4780	0.4480	0.4929
Supervised F	0.4815	0.4637	0.5168
Region Growing Based Road Extraction			
Unsupervised-1 F	0.3725	0.4070	0.4870
Unsupervised-2 F	0.5131	0.4505	0.4680
Supervised F	0.5315	0.5014	0.5049

Table 29:Gain percentages summarized for different probability map construction strategies

	$F_{0.5}$	F_1	F_2
Gain Percentages of Hard Thresholding			
from Unsupervised-1 to Unsupervised-2	11,45%	7,14%	-2,92%
from Unsupervised-1 to Supervised	24,03%	15,48%	8,21%
from Unsupervised-2 to Supervised	11,29%	7,78%	11,47%
Gain Percentages of Hysteresis Thresholding			
from Unsupervised-1 to Unsupervised-2	14,79%	7,41%	-2,93%
from Unsupervised-1 to Supervised	15,63%	11,17%	1,77%
from Unsupervised-2 to Supervised	0,73%	3,50%	4,85%
Gain Percentages of Region Growing Based Road Extraction			
from Unsupervised-1 to Unsupervised-2	37,74%	10,69%	-3,90%
from Unsupervised-1 to Supervised	42,68%	23,19%	3,68%
from Unsupervised-2 to Supervised	3,59%	11,30%	7,88%

<u>Remarks and Interpretations:</u>

The proposed road extraction scheme is composed of three main components, and the first component outputs five different predictors whose class separation scores are unveiled in

Table 26. The weighted summation of these predictors is not expected to generate a combined probability map with class separation score higher than the maximum score suggested by one of the predictors. This inference can be observed by examining the unsupervised probability maps' class separation scores with the ingredient predictors' scores. However, the supervised probability map construction technique directed by Adaboost learning algorithm is capable of generating better class separation estimates. The achievement of the supervised classification on class separation scores is highlighted in *Table 27*. In fact, employing supervised probability map instead of unsupervised estimates brings about 20% and 16% score gains. On the other hand, refined unsupervised probability map construction technique, unsupervised-2, is superior to its ancestor by 4% score gain, unsupervised-1, since we set vegetation pixels to zero in that case rather than leaving it to weighted summation operation.

It is plausible to obtain better road extraction performance for the probability maps with higher class separation scores. This consideration can be proved by examining the road extraction performances obtained from the probability maps generated by different strategies, which is presented in *Table 28*. The best f-measure scores are always achieved by the supervised method. To obtain better insights about the probability map construction strategies, we should interpret the gain percentages given in Table 29. The probability map suggested by Adaboost classification is capable of providing up to 42% performance gain on the road extraction procedure. Although, the performance gain percentage seems to be dependent on the road extraction technique utilized, the supervised method always pose better road extraction performance compared to unsupervised methods. Moreover, replacing the unsupervised method with its refined version often implies better performance results, and the exception occurs when the road extraction procedure is tuned for F_2 objective function. Therefore, we can infer that the exclusion of vegetation mask from the search area is beneficial for obtaining better precision while sacrificing from the recall.

4.4.4 Comparison with Other Road Extraction Algorithms

Purpose:

In this experiment, we compare our algorithm's road detection performance with another road extraction algorithm (Dursun, 2012) proposed recently. In that study, the algorithm suggested in (Long & Zhao, 2005) is selected as baseline, and some of the improvements are proposed for the baseline algorithm. Firstly, the input 3-band IKONOS image is filtered to remove noisy features in order to increase the image quality. Then the filtered image is segmented by using mean-shift algorithm. Spectral classification with pre-defined thresholds followed by morphological operators is used to obtain initial road mask. The final road mask is extracted after applying contour tracing and convex hull algorithms in order to eliminate false alarms.

The reason to choose this algorithm for comparison with ours is that our image dataset contains some of the images employed in theirs. In fact, their image dataset is composed of 20 IKONOS images which are also available in our dataset.

__Procedure:__

The performances of the common images which are both available in our and their datasets are evaluated. We collect the performance measurements from the corresponding article and our experiments. The hard thresholding based road extraction technique is considered for comparison.

__Observations:__

Table 30:Performance comparison

	$\boldsymbol{F_{0.5}}$	$\boldsymbol{F_1}$	$\boldsymbol{F_2}$
(Dursun, 2012)	0.5081	0.5052	0.5004
Our method	0.5245	0.5095	0.5315
Difference percentage	3,23%	0,85%	6,22%

__Remarks and Interpretations:__

In Table 30, performance comparison is presented based on various F-measures including $F_{0.5}$, F_1 and F_2. As the difference percentage in Table 30 implies our algorithm performs slightly better than the algorithm devised by (Dursun, 2012).

CHAPTER 5

CONCLUSION

5.1 Summary

In this study, a new road extraction method operating on the high-resolution multi-spectral satellite images is proposed. 4-band IKONOS satellite images having 1-m resolution are utilized. The proposed approach takes the advantage of multi-resolution analysis and supervised classification accomplished by Adaboost algorithm. The proposed road extraction model is composed of three fundamental components including feature extraction, Adaboost classification and road extraction. The feature extraction component is responsible for providing both spectral and structural features to the classification component. Spectral band ratios represent the reflectance characteristics of the input image, that is, they are beneficial for non-road region detection such as vegetation and water. The input image is exposed to Gaussian pyramid based multi-resolution analysis, and mean-shift algorithm is used to partition the 4-band image into spatially connected segments. Then the segments are scored according to their elongatedness properties. The feature vector contains five predictors; a spectral feature and four structural features which are obtained by elongatedness scoring carried out in several resolution levels. On the other hand, mean-shift image segmentation algorithm is tuned in order to generate optimum segmentation results at each resolution levels for discriminative elongatedness scoring. In this study, the pattern search algorithm is utilized for parameter optimization when needed, and it proved its robustness in parameter tuning tasks carried in this study. By the help of the pattern search algorithm, we generalize the parameter tuning procedure for mean-shift image segmentation yielding a globally valid parameter set. That is, the local hills obtained by optimizing the subsets of the image dataset can be used for acquiring a more generalized parameter set by bounding the pattern search with the local hills.

Moreover, the classification component is utilized by the features with maximal class distribution separability, and the experiments on Adaboost classification component show that Adaboost increases the class distribution separability even further by combining the given predictors. There also two distinct unsupervised probability map construction strategies proposed in this study, so the model can be converted to an unsupervised scheme if it is desired. In fact, the supervised probability map construction method outperforms all other unsupervised strategies, however; the unsupervised methods still pose promising extraction results. Three different road extraction methods are examined in order to supply design alternatives including hard thresholding, hysteresis thresholding and region growing based road extraction strategies.

5.2 Discussion

We consider roads as locally elongated structures, and from the global perspective roads constitute topologically connected regions. Following these assumptions, we utilize both structural and spectral features for classification, whereas, the region growing based road extraction method proposed in this study is capable of validating topological connectivity of the extracted road parts. Although the segments yielded by mean-shift segmentation algorithm are hard for structural scoring due their complicated shapes, the elongatedness index employed for structural scoring provides a good elongatedness evaluation scheme for road segments. It is also observed that some of the non-road objects like buildings and plowed fields may contain elongated structures as well. Topological connectivity validation achieved by region growing strategy is able to deal with some of the situations like these.

Objective function describing class distributions' separation seems to be a robust technique for feature vector quality assessment. This function generates a score expressing the magnitude of separation between two distributions; road and non-road. In parameter optimization technique proposed in this study, we exploit this capability in order to lead pattern search optimization procedure. The biggest impact on the segmentation results comes from minimum segment area parameter of the mean-shift image segmentation algorithm. Bandwidth parameters of the mean-shift describing the window size used for kernel density estimation in range and spatial spaces are little effect on the segmentation results when we compare the results according to elongatedness discrimination.

After empirical evaluations, SAVI is determined as the spectral descriptor maximizing the class distribution separation for road and non-road classes. In fact, the band ratios examined in this study pose high correlation with each other, since they are constructed by similar inferences obtained by band combinations. The choice of the spectral feature used in our road extraction framework should be dependent on the satellite type. In fact, different sensing devices provide color bands occupying slightly different frequency ranges in the electromagnetic spectrum. Therefore, spectral feature selection procedure should be designed by the sensor considerations.

It is shown that the choice of extraction strategy should be based on the application's requirements. For example, the region growing based road extraction method provides better precision by analyzing the extracted road parts in terms of their elongatedness; also it validates the topological connectivity of the road network. On the other hand, hard and hysteresis thresholding methods are two similar road extraction techniques that can provide better recall measurements than the region growing based road extraction strategy.

The proposed road extraction model is easy to be extended by introducing new modules. For instance, new features beneficial for road and non-road discrimination can be added to the feature vector. On the other hand, the model is suitable for replacing existing classification component with a different supervised or unsupervised classification strategy. Likewise, the road extraction methods operating on gray scale image can be employed to detect road regions from the probability map suggested by the classification component.

5.3 Future Work and Outlook

The best F-measure score achieved by our algorithm is 0.54 providing 41% of precision and 65% of recall values, which are obtained by the hard thresholding applied on the probability map obtained from the Adaboost classifier. At this stage, the algorithm can be said to be competitive with the other algorithms available in literature, however; there is still some future work to enhance performance further.

First of all, the algorithm seems to be dependent on the mean-shift image segmentation results, that is; any improvement accomplished in the segmentation procedure will increase the road extraction accuracy. For this purpose, one future work can be defined as employing multi-spectral satellite images having more than four bands. Although the increase data dimension may cause problems for the kernel density estimation based algorithms like mean-shift, the projection of high dimension data is still thought to provide more precise segmentation results. On the other hand, fusion of the segmentation results obtained by different image segmentation algorithms is a promising idea. The segment fusion can be carried out by combining the segmentation results with a rule based heuristic approach. Another option for segmentation fusion is the utilization of training routine with separate elongatedness scores obtained from different segmentation results as predictors.

The major limitation of the proposed algorithm is that the resolution constraint of the input image puts a lower bound for the detectable object size, and thin roads become hard to be identified. Considering that the majority of the target roads are thinner than 5-meters (see *Table 6*), the lack of proper spatial resolution is an important limitation. On the other hand, aerial photometry can provide multi-spectral images with resolution up to 10- centimeters, and these very high resolution images are expected to provide better accuracy when the multi-resolution analysis is employed. The major disadvantage of very high resolution images is that they may refer irrelevant details existing in the scene, since too much detail on the image causes noise. Nevertheless, the features extracted by using multi-resolution analysis are safe to use when a supervised learning scheme is employed, as in our case. The very high resolution images also allow using contextual features such as buildings, trees and cars, as they become visible in such resolution. The contextual features can establish a new era for road extraction, since many occlusions occurred on the road network can be resolved by the help of contextual features.

In this study, we have excluded the post-processing methods which can refine the binary road masks, since we have concentrated on feature extraction and classification components. However, an effective post-processing method employing morphologic and gradient operators is very likely to improve road detection performance.

REFERENCES

Lee , H. Y., Park, W., Lee, H.-K., & Kim , T.-g. (2000). Towards Knowledge-Based Extraction of Roads from 1m-resolution Satellite Images. *Image Analysis and Interpretation, 2000. Proceedings. 4th IEEE Southwest Symposium* (pp. 171 - 176). Austin: IEEE.

Xiao, Z.-q., Bao, G.-s., & Jiang, X.-q. (2004). Road network extraction in classified SAR images using genetic algorithm. *Journal of Central South University of Technology*, 180-184.

Adams, R., & Bischof, L. (1994). Seeded Region Growing. *Pattern Analysis and Machine Intelligence, Vol.16, No.6*, 641-647.

Amini, J., Saradjiana, M. R., Blaisc, J., Lucas, C., & Azizi, A. (2002). Automatic road-side extraction from large scale imagemaps. *International Journal of Applied Earth Observation and Geoinformation*, 95–107.

Arledge, R. K., & Hatcher, E. B. (2008). *INVESTIGATING THE EFFECTS OF HIGHER SPATIAL RESOLUTION ON BENTHIC CLASSIFICATION ACCURACY AT MIDWAY ATOLL.* MONTEREY, CALIFORNIA: NAVAL POSTGRADUATE SCHOOL.

Audet, Charles , & Dennis, J. (2003). Analysis of Generalized Pattern Searches. *SIAM Journal on Optimization, Volume 13, Number 3*, 889–903.

Bacher, U., & Mayer, H. (2005). Automatic Road Extraction from Multispectral High Resolution Satellite Images. *Proceedings of the ISPRS Workshop CMRT 2005* (pp. 29-34). Vienna : ISPRS.

Bajcsy, R., & Tavakoli, M. (1976). Computer Recognition of Roads from Satellite. *IEEE Transactions on Systems and Cybernetics, Vol. SMC-6,*, 623-637.

Baumgartner, A., Hinz, S., & Wiedemann, C. (2002). Efficient Methods and Interfaces for Road Tracking. *International Archives of Photogrammetry and Remote Sensing vol. 34*, 28-31.

Bong, D. B., Lai, K. C., & Joseph, A. (2009). Automatic Road Network Recognition and Extraction for Urban Planning. *International Journal of Applied Science, Engineering and Technology*, 54-59.

Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). *Classification and Regression Trees.* Monterey: Wadsworth & Brooks/Cole Advanced Books & Software.

Burt, P. J. (1981). Fast filter transform for image processing. *Computer Graphics and Image Processing*, 20–51.

Burt, P. J., & Adelson, E. H. (1983). The Laplacian Pyramid as a Compact Image Code. *IEEE Transactions on Communications, Vol.3 , No. 4*, 532-540.

Cheng , Y. (1995). Mean shift, mode seeking, and clustering. *Pattern Analysis and Machine Intelligence*, 790 - 799.

Comaniciu, D., & Meer, P. (1999). Mean shift analysis and applications. *The Proceedings of the Seventh IEEE International Conference on Computer Vision* (pp. 1197 - 1203 vol.2). Kerkyra: IEEE.

Comaniciu, D., & Meer, P. (2002). Mean Shift: A Robust Approach Toward Feature Space Analysis. *Pattern Analysis and Machine Intelligence*, 603 - 619 .

DeAngelis, G., Ohzawa, I., & Freeman, R. D. (1995). Receptive-field dynamics in the central visual pathways. *Trends Neuroscience*, 451–458.

Del-Toro-Almenares, A., Mihai, C., Vanhamel, I., & Sahli, H. (2007). Graph Cuts Approach to MRF Based Linear Feature Extraction in Satellite Images. *Progress in Pattern Recognition, Image Analysis and Applications*, 162-171.

Dursun, M. (2012). *IMPROVEMENT OF MEAN-SHIFT SEGMENTATION BASED ROAD DETECTION.* Ankara: METU.

Freund, Y., & Schapire , R. E. (1997). A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*, 119 - 139 .

Fukunaga, K., & Hostetler, L. (1975). The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory*, 32 - 40 .

Gaetano , R., Zerubia, J., Scarpa, G., & Poggi, G. (2011). Morphological road segmentation in urban areas from high resolution satellite images . *Digital Signal Processing (DSP)* (pp. 1 - 8). Corfu: IEEE.

Geman, D., & Jedynak, B. (1996). An Active Testing Model for Tracking Roads in Satellite Images. *IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 18*, 1-14.

Gong, P., & Wang, J. (1997). Road Network Extraction from Airborne Digital Camera Images. *Geoscience and Remote Sensing,IGARSS* (pp. 895 - 897). Singapore: IEEE.

Gonzalez, R. C., & Woods, R. (2008). *Digital Image Processing.* New Jersey: Pearson.

Grote, A., & Heipke, C. (2008). Road Extraction for The Update of Road Databases in Suburban Areas. *Archives of Photo-grammetry, Remote Sensing and Spatial Information Science*, 563-568.

Grote, A., Heipke, C., Rottensteiner, F., & Meyer, H. (2009). Road extraction in suburban areas by region-based road subgraph extraction and evaluation. *Urban Remote Sensing Joint Event* (pp. 1 - 6). Shanghai: IEEE.

Guan , J., Wang, Z., & Yao , X. (2010). A new approach for road centerlines extraction and width estimation. *Signal Processing (ICSP)* (pp. 924 - 927). Beijing: IEEE.

Haralick, & Shapiro. (1992). *Computer and Robot Vision vol I.* Addison-Wesley.

Hauptfleisch, A. C. (2010). *Automatic Road Network Extraction from High Resolution Satellite Imagery using Spectral Classification Methods.* Pretoria: University of Pretoria.

Heidner, R. F., & Straus, J. M. (2005). The Future of U.S. Commercial Remote Sensing from Space. *Global Priorities in Land Remote Sensing.* South Dakota: ASPRS.

Hernández-Lobato, D., Martínez-Muñoz, G., & Suárez, A. (2011). Empirical analysis and evaluation of approximate techniques for pruning regression bagging ensembles. *Neurocomputing*, 2250–2264.

Hinz, S., & Baumgartner, A. (2003). Automatic extraction of urban road networks from multi-view aerial imagery. *ISPRS Journal of Photogrammetry & Remote Sensing*, 83-98.

Hooke, R., & Jeeves, T. A. (1961). Direct search' solution of numerical and statistical problems. *Journal of the Association for Computing Machinery (ACM)*, 212–229.

Hu , J., Razdan , A., Femiani, J., Ming , C., & Wonka, P. (2007). Road Network Extraction and Intersection Detection From Aerial Images by Tracking Road Footprints. *Geoscience and Remote Sensing, IEEE Transactions on*, 4144 - 4157 .

Hu, J., Razdan, A., John, F. C., Cui, M., & Wonka, P. (2007). Road Network Extraction and Intersection Detection From Aerial Images by Tracking Road Footprints. *IEEE Transactions on Geoscience and Remote Sensing*, 4144 - 4157.

Jin , H., Feng, Y., & Li , B. (2008). Road network extraction with new vectorization and pruning from high-resolution RS images. *Image and Vision Computing* (pp. 1 - 6). Christchurch: IEEE.

Jinxin , C., Qixin, S., & Liguang , S. (2006). A methodology for Automatic Detection and Extraction of Road Edges from High Resolution Remote Sensing Images. *Industrial Technology ICIT* (pp. 69 - 74). Mumbai: IEEE.

Karaman, E., Cinar, U., Gedik, E., Yardimci, Y., & Halici, U. (2012). A new algorithm for automatic road network extraction in multispectral satellite images. *Geographic Object Based Image Analysis* (p. 298). Rio: ISPRS.

Kirthika, A., & Mookambiga, A. (2011). Automated Road Network Extraction Using Artificial Neural Network. *Recent Trends in Information Technology (ICRTIT)* (pp. 1061 - 1065). Chennai: IEEE.

Li, G., An, J., & Chen, C. (2011). Automatic Road Extraction from High-Resolution Remote Sensing Image Based on Bat Model and Mutual Information Matching. *Journal of Computers, Vol. 6, No. 11*, 2417-2426.

Lin , X., Zhang, J., Liu, Z., & Shen , J. (2008). Integration method of profile matching and template matching for road extraction from high resolution remotely sensed imagery. *Earth Observation and Remote Sensing Applications* (pp. 1 - 6). Beijing: IEEE.

Littlestone, N., & Warmuth , M. K. (1994). The weighted majority algorithm. *Information and Computation*, 212 - 261 .

Long, H., & Zhao, Z. (2005). Urban road extraction from high-resolution optical satellite images. *International Journal of Remote Sensing*, 4907–4921.

Ma, H., Qin, Q., Du, S., Wang, L., & Jin, C. (2007). Road Extraction from ETM Panchromatic Image. *Geoscience and Remote Sensing Symposium IGARSS 2007. IEEE International* (pp. 460-463). Barcelona: IEEE International.

Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., & Suetens, P. (1997). Multimodality image registration by Maximization of Mutual Information. *IEEE Transactions on Medical Imaging, Vol. 16, No. 2*, 187-198.

MathWorks. (2012). *MathWorks*. Retrieved from mathworks.com: http://www.mathworks.com/help/toolbox/images/ref/regionprops.html

Maurya, R., Gupta, P., & Shukla, A. S. (2011). Road Extraction Using K-Means Clustering and Morphological Operations. *Image Information Processing (ICIIP)* (pp. 1 - 6). Himachal Pradesh: IEEE.

Mayer, H., Laptev, I., & Baumgartner, A. (1998). Multi-scale and Snakes for Automatic Road Extraction. In H. Burkhardt, & B. Neumann, *Computer Vision ECCV* (pp. 720-733). Berlin: Springer.

McKeown, D. M., & Denlinger, J. L. (1988). Cooperative Methods For Road Tracking In Aerial Imagery. *Computer Vision and Pattern Recognition* (pp. 662-672). Ann Arbor: IEEE.

Mena, J. (2003). State of the art on automatic road extraction. *Pattern Recognition Letters, 24*, 3037–3058.

Mena, J., & Malpica, J. (2005). An automatic method for road extraction in rural and semi-urban areas starting from high resolution satellite imagery. *Pattern Recognition Letters*, 1201-1220.

Mirnalinee, T., Das, S., & Varghese , K. (2009). Integration of Region and Edge-based information for Efficient Road Extraction from High Resolution Satellite Imagery. *Advances in Pattern Recognition* (pp. 373 - 376). Kolkata: IEEE.

Murai, S. (1991). Microwave Remote Sensing. In Y. Akamatsu, K. Arai, & K. Inoue, *Remote Sensing Note* (p. 56). Tokyo: Japan Association on Remote Sensing.

Parzen, E. (1962). On Estimation of a Probability Density Function and Mode. *The Annals of Mathematical Statistics*, 1065-1076.

Péteri, R., & Ranchin , T. (2003). Multiresolution snakes for urban road extraction from ikonos and quickbird images. *23nd EARSeL Annual Symposium Remote Sensing in Transition* (pp. 2-5). Rotterdam: Millpress.

Quinlan , J. R. (1986). Induction of Decision Trees. *Machine Learning*, 81 - 106 .

Rangnath , N. R. (2001). Remote Sensing 2. Sensors and Platforms. *Basics and applications, Resonance, Vol.6, No.12*, 51-60.

Safavian, R. S., & Landgrebe, D. (1991). A Survey of Decision Wee Classifier Methodology. *IEEE TRANSACTIONS ON SYSTEMS, MAN, AND CYBERNETICS, VOL. 21, NO. 3*, 660-674.

Schapire, R. E. (1990). The Strength of Weak Learnability. *Machine Learning*, 197 - 227 .

Schapire, R. E., Freund, Y., Bartlett, P., & Lee, W. S. (1998). Boosting the Margin:A New Explanation for the Effectiveness of Voting Methods. *The Annals of Statistics*, 1651-1686.

Schowengerdt, R. A. (2007). *Remote sensing: models and methods for image processing.* Academic Press.

Smith, R. B. (2012). *Introduction to Remote Sensing of Environment.* Lincoln : MicroImages, Inc.

Space Imaging. (2000). *IKONOS scene po-37836.* Virginia: GeoEye.

Timofeev, R. (2004). *Classification and Regression Trees Theory and Applications.* Berlin: Master Thesis, Center of Applied Statistics and Economics.

Torczon, V. (1995). Pattern Search Methods for Nonlinear Optimization. *SIAG/OPT Views-and-News: A Forum for the SIAM Activity Group on Optimization, No. 6*, 7 - 11.

Torczon, V. (1997). On the convergence of pattern search algorithms. *SIAM Journal on Optimization*, 1–25.

Trinder, J. (2009). Towards Automation of Information Extraction from Aerial and Satellite Images. In D. Li, J. Shan, & J. Gong, *Geospatial Technology for Earth Observation* (pp. 289-327). Berlin: Springer.

Valiant, L. (1984). A theory of the learnable. *Magazine Communications of the ACM*, 1134 - 1142.

Vezhnevets, A. (2010). *GML AdaBoost Matlab Toolbox Manual.* Moscow: MSU Graphics & Media Lab.

Wiedemann , C., & Ebner, H. (2000). Automatic Completion And Evaluation Of Road Networks. *International Archives of Photogrammetry and Remote Sensing*, 979-986.

Yanqing, W., Yuan, T., Tai, X., & Shu, L. (2006). Extraction of Main Urban Roads from High Resolution Satellite Images by Machine Learning. *7th Asian conference on Computer Vision* (pp. 236-245). Heidelberg: Springer.

Ye, F.-m., Su, L., & Tang, J.-l. (2006). Automatic Road Extraction Using Particle Filters from High Resolution Images. *Journal of China University of Mining and Technology*, 490–493.

Young, R. A. (1987). The Gaussian derivative model for spatial vision: I. Retinal mechanisms. *Spatial Vision, Vol. 2, No. 4*, 273-293.

Zhang, Q., & Couloigner, I. (2006). Benefit of the angular texture signature for the separation of parking lots and roads on high resolution multi-spectral imagery. *Pattern Recognition Letters*, 937-946 .

Zhao , L., & Wang , X. (2010). Road extraction in high resolution remote sensing images based on mathematic morphology and snake model. *Image and Signal Processing (CISP)* (pp. 1436 - 1440). Yantai: IEEE.

Zhao, L., & Xili, W. (2010). Road Extraction in High Resolution Remote Sensing Images Based on Mathematic Morphology and Snake Model. *International Congress on Image and Signal Processing* (pp. 1436-1440). Yantai: IEEE.

Zheng, S., Liu, J., Shi, W.-Z., & Zhu, G.-X. (2006). Road Central Countour Extraction from High Resolution Satellite Image using Tensor Voting Framework. *Proceedings of the Fifth International Conference on Machine Learning and Cybernetics* (pp. 3248 - 3253). Dalian: IEEE.

Zhou, J., Bischof, W. F., & Caelli, T. (2006). Road tracking in aerial images based on human–computer interaction and Bayesian filtering. *ISPRS Journal of Photogrammetry & Remote Sensing*, 108–124.

Zhu, D.-M., Wen, X., & Ling, C.-L. (2011). Road Extraction Based on The Algorithms of MRF and Hybrid Model of SVM and FCM. *Image and Data Fusion (ISIDF)* (pp. 1-4). Yunnan: IEEE.

APPENDICES

Appendix-A: Road Extraction Parameter Tuning Results for The Unsupervised Probability Map Construction

Table 31:Result of parameter tuning on training set for hard thresholding

		$F_{0.5}$	F_1	F_2
Precision	Mean	0.5293	0.4121	0.3207
	Std	0.1909	0.1722	0.1455
Recall	Mean	0.4130	0.5941	0.7277
	Std	0.1813	0.1722	0.1338
Accuracy	Mean	0.9290	0.9095	0.8713
	Std	0.0325	0.0233	0.0137
F	Mean	0.4677	0.4475	0.5347
	Std	0.1385	0.1286	0.1147
Threshold Value		0.9656	0.9187	0.8562

Table 32:Result of parameter tuning on training set for hysteresis thresholding

		$F_{0.5}$	F_1	F_2
Precision	Mean	0.5352	0.4066	0.3313
	Std	0.1940	0.1657	0.1503
Recall	Mean	0.4028	0.6053	0.7177
	Std	0.1763	0.1775	0.1368
Accuracy	Mean	0.9294	0.9093	0.8769
	Std	0.0328	0.0233	0.0159
F	Mean	0.4687	0.4514	0.5361
	Std	0.1411	0.1349	0.1135
Threshold Value-1		0.9656	0.9031	0.8562
Threshold Value-2		0.9719	0.9719	0.9094

Table 33:Result of parameter tuning on training set for region growing based method

		$F_{0.5}$	F_1	F_2
Precision	Mean	0.4389	0.4318	0.3553
	Std	0.1676	0.1684	0.1493
Recall	Mean	0.5495	0.5383	0.6733
	Std	0.1791	0.1816	0.1579
Accuracy	Mean	0.9184	0.9182	0.8924
	Std	0.0267	0.0281	0.0209
F	Mean	0.4405	0.4511	0.5329
	Std	0.1496	0.1283	0.1100
Region growing similarity value		0.0502	0.0568	0.0511
Minimum peak value		0.9781	0.9781	0.9156
Minimum elongatedness value		30	30	30

Appendix-B: Road Extraction Parameter Tuning Results for The Refined Unsupervised Probability Map Construction

Table 34:Result of parameter tuning on training set for hard thresholding

		$F_{0.5}$	F_1	F_2
Precision	Mean	0.5842	0.5058	0.3774
	Std	0.2113	0.2117	0.2008
Recall	Mean	0.4098	0.5326	0.6661
	Std	0.2329	0.2602	0.2602
Accuracy	Mean	0.9404	0.9350	0.9042
	Std	0.0313	0.0260	0.0368
F	Mean	0.4812	0.4591	0.5215
	Std	0.1948	0.1913	0.2028
Threshold Value		0.9500	0.9031	0.7625

Table 35:Result of parameter tuning on training set for hysteresis thresholding

		$F_{0.5}$	F_1	F_2
Precision	Mean	0.5767	0.5427	0.3995
	Std	0.1993	0.1968	0.2040
Recall	Mean	0.4668	0.5151	0.6548
	Std	0.2544	0.2640	0.2616
Accuracy	Mean	0.9411	0.9393	0.9101
	Std	0.0291	0.0270	0.0350
F	Mean	0.4894	0.4666	0.5227
	Std	0.2067	0.2043	0.2036
Threshold Value-1		0.9187	0.8875	0.7625
Threshold Value-2		0.9719	0.9719	0.8625

Table 36:Result of parameter tuning on training set for region growing based method

		$F_{0.5}$	F_1	F_2
Precision	Mean	0.6094	0.5513	0.5064
	Std	0.1937	0.1904	0.1773
Recall	Mean	0.4946	0.5563	0.6338
	Std	0.2138	0.2269	0.2180
Accuracy	Mean	0.9391	0.9344	0.9288
	Std	0.0260	0.0258	0.0227

F	Mean	0.5398	0.5028	0.5659
	Std	0.1674	0.1642	0.1593
Region growing similarity value		0.0552	0.0793	0.1071
Minimum peak value		0.9781	0.9625	0.9000
Minimum elongatedness value		30	30	30