

ROAD EXTRACTION FROM SATELLITE IMAGES BY  
SELF-SUPERVISED CLASSIFICATION AND PERCEPTUAL GROUPING

A THESIS SUBMITTED TO  
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES  
OF  
MIDDLE EAST TECHNICAL UNIVERSITY

BY

EDA ŞAHİN

IN PARTIAL FULLFILLMENT OF THE REQUIREMENTS  
FOR  
THE DEGREE OF MASTER OF SCIENCE  
IN  
ELECTRICAL AND ELECTRONICS ENGINEERING

JANUARY 2013



Approval of the Thesis:

**ROAD EXTRACTION FROM SATELLITE IMAGES BY SELF-SUPERVISED  
CLASSIFICATION AND PERCEPTUAL GROUPING**

submitted by **EDA ŞAHİN** in partial fulfillment of the requirements for the degree of **Master of Science in Electrical and Electronics Engineering Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen  
Dean, Graduate School of **Natural and Applied Sciences**

\_\_\_\_\_

Prof. Dr. İsmet Erkmén  
Head of Department, **Electrical and Electronics Engineering**

\_\_\_\_\_

Assoc. Prof. Dr. İlkey Ulusoy  
Supervisor, **Electrical and Electronics Engineering Dept., METU**

\_\_\_\_\_

**Examining Committee Members:**

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

\_\_\_\_\_

Assoc. Prof. Dr. İlkey Ulusoy  
Electrical and Electronics Engineering Dept., METU

\_\_\_\_\_

Prof. Dr. Gözde Bozdağı Akar  
Electrical and Electronics Engineering Dept., METU

\_\_\_\_\_

Prof. Dr. Aydın Alatan  
Electrical and Electronics Engineering Dept., METU

\_\_\_\_\_

Assist. Prof. Dr. Alptekin Temizel  
Work Based Learning Studies Dept., METU

\_\_\_\_\_

**Date: 23.01.2013**

**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 : Eda ŞAHİN

Signature :

## ABSTRACT

### ROAD EXTRACTION FROM SATELLITE IMAGES BY SELF-SUPERVISED CLASSIFICATION AND PERCEPTUAL GROUPING

Şahin, Eda

M.S., Department of Electrical and Electronics Engineering

Supervisor: Assoc. Prof. Dr. İlkay Ulusoy

January 2013, 61 pages

Road network extraction from high resolution satellite imagery is the most frequently utilized technique for updating and correcting geographic information system (GIS) databases, registering multi-temporal images for change detection and automatically aligning spatial datasets. This advance method is widely employed due to the improvements in satellite technology such as development of new sensors for high resolution imagery. To avoid the cost of the human interaction, various automatic and semi-automatic road extraction methods are developed and proposed in the literature. The aim of this study is to develop a fully automatized method which can extract road networks by using the spectral and structural features of the roads. In order to achieve this goal we set various objectives and work them out one by one. First objective is to obtain reliable road seeds, since they are crucial for determining road regions correctly in the classification step. Second objective is finding most convenient features and classification method for the road extraction. The third objective is to locate road centerlines which are defines the road topology. A number of algorithms are developed and tested throughout the thesis to achieve these objectives and the advantages of the proposed ones are explained. The final version of the proposed algorithm is tested by three band (RGB) satellite images and the results are compared with other studies in the literature to illustrate the benefits of the proposed algorithm.

Keywords: Anti-parallel Centerline, Gaussian Mixture, Perceptual Grouping, Classification

## ÖZ

### UYDU GÖRÜNTÜLERİNDEN ÖZ GÖZETİMLİ SINIFLANDIRMA VE ALGISAL GRUPLAMA YÖNTEMLERİ İLE OTOMATİK YOL ÇIKARIMI

Şahin, Eda

Yüksek Lisans, Elektrik Elektronik Mühendisliği Bölümü  
Tez Yöneticisi: Doç. Dr. İlkay Ulusoy

Ocak 2013, 61 sayfa

Yüksek çözünürlüklü uydu görüntüsünden yol ağı tespiti yöntemi, Coğrafi Bilgi Sistemi (GIS) veritabanı düzeltilmesi ve güncellenmesi, farklı zamanlı görüntülerde değişiklik tespiti için görüntülerin hizalanması, uzaysal veritabanlarının otomatik olarak hizalanması gibi işlemlerde kullanılan temel bir konudur. Gelişen alman kabiliyetleri sayesinde yüksek çözünürlüklü uydu görüntülerinin elde edilmesi ile bu konu üzerinde geniş ölçüde çalışmalar başlamıştır. İnsan etkileşiminden kaynaklı maliyet artışını önlemek için, literatürde çeşitli otomatik veya yarı otomatik yol çıkarımı yöntemleri önerilmiştir. Bu tezin amacı, yolların spektral ve yapısal özellikleri kullanılarak otomatik olarak yol ağı tespitinin sağlanmasıdır. Bu amaç doğrultusunda çeşitli hedefler belirlenerek, bu hedefler üzerinde çalışmalar gerçekleştirilmiştir. Öncelikli hedef yollara ait örnek noktaların güvenilir bir şekilde belirlenmesidir, çünkü elde edilen örnek noktaların doğruluğu sınıflandırma adımda doğru yol bölgelerinin belirlenmesindeki en önemli etkidir. İkinci hedef, sınıflandırma yönteminin ve sınıflandırma için en uygun yol özelliklerinin belirlenmesidir. Üçüncü hedef, yolun merkez hatlarının oluşturulması ve yol topolojisinin tanımlanmasıdır. Bu hedefleri gerçekleştirmek için çeşitli algoritmalar incelenmiş ve önerilen algoritmaların avantajları anlatılmıştır. Önerilen algoritma birçok üç bantlı (RGB) uydu görüntüsü ile test edilmiş, literatürdeki diğer çalışmalarla karşılaştırılarak avantajları belirtilmiştir.

Anahtar Kelimeler: Anti-paralel Merkez Hattı, Gauss Karışımı, Algısal Gruplama, Sınıflandırma

To My Dear Husband

## ACKNOWLEDGEMENTS

I would like to express my sincere gratitude and appreciation to my supervisor Assoc. Prof. Dr. İlkay Ulusoy for her supervision, guidance and encouragement throughout this study.

I would like to thank Prof Dr. Aydın Alatan, Prof Dr. Uğur Halıcı, Prof Dr. Güzde Akar and Assist. Prof. Dr. Alptekin Temizel for serving in my committee and sharing their opinions.

I would also like to thank Tübitak for their financial support during my graduate education.

I would like to also express my thanks to my friends, my colleagues and Aselsan A.Ş. for their precious support and fellowship.

I would like to express my sincere gratitude to Adem Ergül for proofreading this thesis and Mustafa Dursun for obtaining ground truth of the data.

I would like to thank my family and Kerem for their love, support and patience over the years. This thesis is dedicated to them.

## TABLE OF CONTENTS

ABSTRACT .....	v
ACKNOWLEDGEMENTS .....	viii
TABLE OF CONTENTS .....	ix
LIST OF TABLES .....	xi
LIST OF FIGURES.....	xii
CHAPTERS	
1 INTRODUCTION.....	1
2 LITERATURE REVIEW.....	3
2.1 According to Objectives .....	3
2.1.1 Road Mask Extraction Methods .....	3
2.1.2 Road Centerline Extraction Methods.....	4
2.2 According to Extraction Techniques.....	4
2.2.1 Road Seed Acquisition .....	4
2.2.2 Segmentation Based .....	5
2.2.3 Classification .....	5
2.2.4 Road Modeling .....	5
2.2.5 Road Tracking .....	6
2.2.6 Morphology and Filtrate.....	6
2.2.7 Multi-scale Analysis.....	7
2.2.8 Grouping and Linking .....	7
3 BACKGROUND METHODS .....	9
3.1 Canny Edge Detection Algorithm.....	9
3.2 Principle Component Analysis .....	9
3.3 Anti-Parallel Centerline Extraction.....	9
3.4 Hough Transform.....	11
3.5 Bilateral Filter.....	12
3.6 Gaussian Mixture Model.....	13
3.7 Features Used for Classification .....	14
3.7.1 Mean and Standard Deviation .....	14
3.7.2 HSV Color Space .....	14
3.7.3 Gabor Filters.....	15
3.8 Relative Neighborhood Graph .....	15
3.9 Minimal Spanning Tree .....	16
3.10 Self Organized Road Map.....	17
3.11 Perceptual Grouping .....	18
3.12 Fuzzy Organization of Elongated Regions (FOrgER) .....	18
4 PROPOSED ROAD EXTRACTION ALGORITHM.....	21
4.1 Seed Point Extraction.....	22
4.1.1 Edge Detection .....	22
4.1.2 ACE (Anti-parallel Centerline Extraction).....	22
4.1.3 Linking Seed Points.....	23
4.1.3.1 Hough Transform .....	23
4.1.3.2 Perceptual Grouping.....	24
Labeling Line Pixels (LLP) .....	25
Grouping Based on Gestalt Principles .....	25
Bridging Gaps.....	27
4.1.3.3 Fuzzy Organization of Elongated Regions (FOrgER).....	28
4.1.4 Removing Shadow and Water from Seed Points.....	29
4.2 Classification .....	29
4.2.1 Smoothing .....	29
4.2.2 Modeling and Classifying.....	30
4.3 Morphological Operations .....	33
4.4 Topology Construction .....	33

5 EXPERIMENTAL RESULTS .....	37
5.1 The Performance Evaluation Metrics .....	38
5.2 Analysis of ACE Inputs .....	39
5.3 Analysis of the Seed Points Linking Algorithms .....	41
5.4 Analysis of Classification Features .....	43
5.5 Analysis of Classification Algorithms .....	45
5.6 Analysis of Smoothing Algorithms .....	46
5.7 Analysis of Morphological Operations .....	48
5.8 Analysis of Topology Construction Algorithms .....	49
5.9 Comparison with the Baseline Algorithm .....	51
5.10 Comparison with Previous Studies .....	52
CHAPTER 6 CONCLUSION .....	55
6.1 Summary and Conclusions .....	55
6.2 Future Work .....	55
REFERENCES .....	57
APPENDICES	
A.EXAMPLE RESULTS ON GROUND TRUTH .....	59

## LIST OF TABLES

### TABLES

Table 2-1 The classification of road extraction methods .....	3
Table 4-1 The evaluation results of PBA and OBA [13] .....	30
Table 5-1 The performance comparison of the alternative ACE methods .....	41
Table 5-2 The performance comparison of the alternative linking methods .....	42
Table 5-3 The parameters of the linking algorithm.....	43
Table 5-4 The performance comparison of the alternative features .....	45
Table 5-5 The performance comparison of the alternative classification methods .....	46
Table 5-6 The performance improvement with smoothing filter .....	47
Table 5-7 The performance improvement with morphological operations .....	49
Table 5-8 The performance comparison of the topology construction methods .....	51
Table 5-9 The performance comparison with [1].....	52
Table 5-10 The performance comparison of classification results .....	53
Table 5-11 The performance comparison of final results.....	53

## LIST OF FIGURES

### FIGURES

Figure 3-1 ACE Horizontal Scan Line [1] .....	10
Figure 3-2 ACE Processing (a) Road Profile and Gradient Orientations of Road on input image (b) Canny Edge Result (c) Sobel Gradient Orientations of Canny Edges(Units are degrees) (d) Gradient Orientation Graph Along Selected Profile (e) The candidate road seeds from ACE...	11
Figure 3-3 The normal parameters of a line .....	12
Figure 3-4 Bilateral filter (a) input (b) weight $r \times d$ for centre pixel (c) output [36].....	12
Figure 3-5 Gaussian distribution modeling .....	13
Figure 3-6 Gaussian mixture modeling .....	14
Figure 3-7 Gabor Filter Decomposition a) The sinusoidal carrier b) The Gaussian envelope c) The corresponding Gabor Filter [29] .....	15
Figure 3-8 Relative Neighborhood Graph a) The lune of the points $p$ and $q$ b) The constructed neighborhood graph.....	16
Figure 3-9 Minimal Spanning Tree .....	16
Figure 3-10 SORM Process a) Cluster Initialization b) Cluster Centers on Voronoi Graph (c) The result of SORM Algorithm [8] .....	18
Figure 3-11 The orientation and link angle of SORM centers .....	19
Figure 4-1 Flow diagram of road centerline extraction algorithm .....	21
Figure 4-2 Seed Point Linking a) The output of ACE b) Lines found with Hough Transform.....	23
Figure 4-3 Grouping and linking line segments a) labeled line segments b) orientations of local curvature around endpoints (c) result of grouping and linking .....	24
Figure 4-4 Labeling line pixels a) Labeling connected braches b) The result of LLP algorithm.....	25
Figure 4-5 Grouping Parameters .....	26
Figure 4-6 Determining the directions of line segments .....	26
Figure 4-7 Bridging gaps between two lines.....	27
Figure 4-8 ACE Image and Perceptual Grouping Result .....	27
Figure 4-9 Selected seed points on the edge image.....	28
Figure 4-10 FOrgER Algorithm (a) Link strengths on Voronoi Diagram b) Linked centers.....	28
Figure 4-11 Shadow removal from the seed points a) Before shadow removal b) After shadow removal.....	29
Figure 4-12 Water region removal from the seed points a) Before water region removal b) After water region removal.....	29
Figure 4-13 Bilateral filter (a) input image (b) output image.....	30
Figure 4-14 Gaussian membership values (a) membership of the road class (b) membership of the non-road class (c) posterior probabilities for road class (d) result of thresholding .....	31
Figure 4-15 Classification based on Gaussian mixture model (a) membership of road class (b) the classification result .....	32
Figure 4-16 Classification based on Gaussian mixture model (a) original image (b) First cluster (c) Second cluster (d) The final result.....	33
Figure 4-17 Linking SORM Centers (a) SORM Centers (b)with MST (c) with RNG (c) with FOrgER .....	34
Figure 4-18 Morphology and Perceptual Grouping (a) Classified image after closing and hole filling (b) Subtraction of edges (c) Morphological Thinning (d) Result on the input image.....	35
Figure 5-1 Base algorithm and alternative methods.....	38
Figure 5-2 The matching stage of performance evaluation algorithm (a) Buffering around the reference data (b) Buffering around the extracted data .....	39
Figure 5-3 The Correctness of the ACE results .....	40
Figure 5-4 The completeness of the ACE results.....	40
Figure 5-5 The quality of the ACE results .....	40

Figure 5-6 The variance of the ACE result in terms of road width parameters (a) In terms of the minimum road width (b) In terms of the maximum road width .....	41
Figure 5-7 The correctness of the linking results .....	42
Figure 5-8 Completeness of the linking results .....	42
Figure 5-9 Quality of the linking results .....	42
Figure 5-10 The quality variance of linking results in terms of parameters (a) In terms of the maximum gap length (b) In terms of the minimum line length (c) In terms of the maximum angle difference .....	43
Figure 5-11 The correctness of the classification features .....	44
Figure 5-12 The completeness of the classification features .....	44
Figure 5-13 The quality of the classification features .....	44
Figure 5-14 The quality variance in terms of minimum membership values (a) In GMM (b) In normal distribution.....	45
Figure 5-15 The correctness of the classification methods .....	45
Figure 5-16 The completeness of the classification methods .....	46
Figure 5-17 The quality of the classification methods .....	46
Figure 5-18 The quality variance in terms of the parameters of bilateral filter (a) In terms of range filter variance value (b) In terms of domain filter variance value .....	46
Figure 5-19 The correctness comparison for smoothing filter .....	47
Figure 5-20 The completeness comparison for smoothing filter .....	47
Figure 5-21 The quality comparison for smoothing filter .....	47
Figure 5-22 The quality variance in terms of parameters used in morphological operations (a) Area threshold used for gap filling (b) Area threshold used for CC removing (c) The size of structuring element used for closing.....	48
Figure 5-23 The correctness comparison for morphological operations .....	48
Figure 5-24 The completeness comparison for morphological operations.....	49
Figure 5-25 The quality comparison for morphological operations .....	49
Figure 5-26 The grid size selection for SORM algorithm (a) MST (b) FOrger .....	50
Figure 5-27 The linking result of SORM algorithm (a) $\beta_{grid} = 20$ (b) $\beta_{grid} = 60$ .....	50
Figure 5-28 The correctness of the topology construction methods.....	51
Figure 5-29 The completeness of the topology construction methods .....	51
Figure 5-30 The quality of the topology construction methods .....	51
Figure 5-31 The correctness comparison .....	52
Figure 5-32 The completeness comparison .....	52
Figure 5-33 The quality comparison .....	52
Figure A-1 The result of test image 1 (green : TP, blue: FN, red : FP).....	59
Figure A-2 The result of test image 3 (green : TP, blue: FN, red : FP).....	59
Figure A-3 The result of test image 4 (green : TP, blue: FN, red : FP).....	60
Figure A-4 The result of test image 5 (green : TP, blue: FN, red : FP).....	60
Figure A-5 The result of test image 6 (green : TP, blue: FN, red : FP).....	61
Figure A-6 The result of test image 12 (green : TP, blue: FN, red : FP).....	61



## CHAPTER 1

### INTRODUCTION

Automatic road extraction from satellite imagery is a fundamental computer vision operation and this vision attracted wide attention with the availability of the high resolution data within last few decades. It is now possible produce road maps for any given region of the earth by using satellite images. Road extraction is used for many different purposes. In geospatial applications it is used for completing and updating GIS databases with accurate information while for the military purposes, it is useful for transportation, tactical missions, etc.

Although many road extraction methods are proposed, extraction from images is performed mostly manually and this is not effective in terms of cost and time. Therefore automatized procedures are required in order to overcome these drawbacks. Due to this necessity we have worked on algorithms which are capable of extracting road maps automatically from satellite imagery.

Several automatic road extraction methods and algorithms can be found on the literature. Most of these methods show superiority in certain fields while none of them outperforms all the others in every extent. In this study, we intend to select and combine the eminent parts of these methods in order to create a fast and effective road extraction method.

Therefore we have survey the literature and an automatic seeding method proposed in [1] is selected as the initial step. The proposed algorithm is performed step by step. First, the methods for improving seed points are examined and perceptual grouping is selected as the seed point grouping algorithm. Then, the need of linking seeds in the same group is emerged, so the linking algorithm proposed in [6] is used since it was appropriate for our purpose. In the classification step, we have a two-class problem such as road and non-road, but we have only the samples of the road class. The comparisons of the classification methods [34, 35] are examined and it is seen that the neural network and machine learning are not appropriate, but the statistical approach is appropriate for our purpose. After the classification step, roads must be expressed as networks and this step requires road topology construction. Therefore graph based methods and grouping methods are examined and the perceptual grouping is preferred based on the conducted experiments for topology construction.

The aim of this thesis is to extract road network automatically from three bands (RGB) satellite images and for this purpose the 4-band high resolution images obtained from the IKONOS observation satellite are used as test objects for the algorithm.

The output of the proposed algorithm is a structure that consists of straight and curvature line segments that constitute the road network of an input image. All the pixels belong to the same line segments together with the start points and the end points are specified in the same structure. Due to the construction of the structures, the proposed algorithm is applicable for not only the GIS update, but also for the navigation and registration purposes.

The contributions of the thesis are summarized as follows:

- In order to achieve the best performance, we have examined methods proposed on the literature and the self-supervised road classification algorithm proposed by Doucette et al. is preferred as a baseline algorithm. Some steps of this algorithm are modified and a remarkable enhancement at the performance is achieved.
- In seed point extraction step, we examined the Hough transform and perceptual grouping for linking the ACE responses. Shadow and water removal is performed for eliminating non-road seed points.

- In classification step, we have seen some spectral and structural features such as mean, standard deviation, HSV components, Gabor features and we also use Gaussian mixture model as an alternative of normal distribution.
- In topology construction step, morphological operations are applied to improve the classification result. Before applying thinning algorithm, edges are removed from road class image for smoothing the borders of the roads. Finally, thinning operation is performed and road centerlines are constructed with perceptual grouping. The centerlines of dual carriageways can be detected separately by the help of the edge removal.

The thesis is organized as follows, in Chapter 2, the related studies on road extractions and brief introductions about these studies are presented. In Chapter 3, the techniques used in the proposed algorithm are described in details. The formulations, graphical representations and some visualization about these techniques are given in this chapter in order to provide a good understanding on the proposed algorithm. In Chapter 4, the alternative methods examined in each step of the proposed algorithm are presented. The customization of the methods given in chapter 3 is described in this chapter. In Chapter 5, the details of the performance analysis of the developed system are presented. The quality metrics, the comparison between alternative methods, the parameter selection criteria and their effects to the overall system are illustrated. In Chapter 6, the summary of the thesis is presented and future work is discussed.

## CHAPTER 2

### LITERATURE REVIEW

In the field of road extraction, there are a large amount of works performed on satellite imagery. These works are classified as semi-automatic and fully automatic methods. Semi-automatic approaches rely on user-provided cues. These algorithms assist the human analyst to execute road extraction tasks more efficiently under human supervision. The advantage of these approaches is convenience of on-the-fly editing for the user. On the other hand, fully automatic methods are intended to extract initial clues automatically, without interactive operation of a human operator. Although the efficiency from fully automated methods is potentially enormous, in practice it is difficult to achieve a reliable performance.

General steps of the road extraction methods are examined in the basis of the survey by Mena[11] and the literature review of Hauptfleisch thesis[13] which provide an overview of the various extraction methods. The categorization methods proposed in these reviews are followed to a large extent, but they have been restructured in this study.

Mena [11] classifies the road extraction techniques according to the preset objectives and extraction techniques. We also compose our review according to this structure.

Table 2-1 The classification of road extraction methods

<b>Road Extraction Methods</b>	
<b>According to Objectives:</b>	<b>According to Extraction Techniques:</b>
<ul style="list-style-type: none"><li>• Road Mask Extraction</li></ul>	<ul style="list-style-type: none"><li>• Road Seed Acquisition</li></ul>
<ul style="list-style-type: none"><li>• Road Centerline Extraction</li></ul>	<ul style="list-style-type: none"><li>• Segmentation Based</li></ul>
	<ul style="list-style-type: none"><li>• Classification</li></ul>
	<ul style="list-style-type: none"><li>• Road Modeling</li></ul>
	<ul style="list-style-type: none"><li>• Road Tracking</li></ul>
	<ul style="list-style-type: none"><li>• Morphology and Filtrate</li></ul>
	<ul style="list-style-type: none"><li>• Multi-spectral Analysis</li></ul>
	<ul style="list-style-type: none"><li>• Grouping and Linking</li></ul>

#### 2.1 According to Objectives

Many road extraction methods focus on detecting the pixels which belong to the roads, besides many methods aims to obtain the numerical and topological definition of the roads. According to the aims, the road extraction methods are examined in two different classes. In 2.1.1, the road mask extraction methods which aim to find a road mask containing the road areas are introduced. In 1.1.1, the road centerline extraction methods which detect the centerline points are presented.

##### 2.1.1 Road Mask Extraction Methods

The segmentation and classification based algorithms generally aim to detect road network. In [28], the road class image is obtained with k-means segmentation and morphological dilation and non-road area removal. In [20], classification is performed by artificial neural network by using color information of 3x3 neighborhoods of the pixels. In [7], a fuzzy classification is performed and then the road network is constructed by means of ziplock line snakes.

Multi-spectral analyzing methods use edge information to find road segments and group these segments to complete the road network. [15, 19]

A road tracking method is proposed in [12] which bases on the intensity profile matching of road cross sections. The road pixels are detected in terms of the geometric and radiometric parameters.

### **2.1.2 Road Centerline Extraction Methods**

Some of the centerline extraction methods extract the road topology from the road network. SORM algorithm in [1,8] are proposed for this purpose. The binary road class image is used to obtain road centerlines by applying a clustering algorithm and graph-based grouping method. Mena at al.[17] also propose a system consists of four different modules. Data preprocessing and binary segmentation is performed on the input image by considering textural information, and then automatic vectorization by means of skeleton extraction and system evaluation is carried out.

Most of road tracking and modeling algorithms also detect road topologies. Koutaki at al. [26] uses the position and direction information of road intersections to generate a contour for road tracking. Then, the ribbon snakes are used to extract the road topology. In [14, 27], dynamic programming and active contour models are implemented to extract linear features.

Hu at al. [4] propose a ribbon line detector to obtain main road centerlines and a hierarchical perceptual grouping for linking fragmented centerlines.

## **2.2 According to Extraction Techniques**

In section 2.2.1, the road seed acquisition methods generally used in fully automatic approaches are discussed. In section 2.2.2 and 2.2.3, as important parts of road network construction, segmentation and classification algorithms are introduced. In section 2.2.4, the road modeling methods generally used in the road network updating algorithms are explained. The road tracking methods are introduced in section 2.2.5. In section 2.2.6, the morphological and filtering operations which are preferred for smoothing the road network, eliminating false regions and extracting the road centerlines are illustrated. Multi-scale approaches are defined in section 2.2.7. And finally in section 2.2.8, the grouping and linking methods used for constructing a complete road network are introduced.

### **2.2.1 Road Seed Acquisition**

Road seeds are the markers on the roads that are used to identify approximate feature locations and the topology of the road class. Most of road extraction methods need to know some certain points on the road network such as the intersection points of the roads, the high curvature points, the points along the centers of the roads.

Road following approaches are usually initialized by the seed points on the roads to determine the starting points of the algorithms and to define the road directions.[12] Active contour models and dynamic programming approaches are initialized by a few seed points describing coarsely the roads.[14] In classification based methods, road seeds are used as training data of the road class.[19]

Some of the available road extraction methods select the seed points manually and some of them find automatically. Parallel edge detection is the most popular approach to automatic road seeding. These approaches are based on the fact that the edges of the roads are parallel to each other and have opposing gradient orientations. The seed points are formed by locating the road centerlines between these parallel edges or creating rectangles by connecting the end points of these parallel edges. In [1], anti-parallel edge centerline extraction (ACE) is used to detect the road seed points. In [21], the minimal oriented bounding box (MOBB) method which detects the homogeneous and long rectangular regions between parallel edges is used for automatic seeding.

Geometric model matching is the other popular method used for road seed detection. The road candidates are selected based on the shape and the spectral properties and they are compared with the pre-defined road models. The correctly matched profiles are selected as seed points. In [26], spectral classification is applied to find road candidates and they are compared with the road intersection models. The overlapped regions are selected as seed points.

### 2.2.2 Segmentation Based

Supervised or unsupervised segmentation is used to extract the road network which is used as an input of higher level processes of the road extraction. In the road extraction methods, segmentation is conducted to simplify the classification step of the algorithms. Segmentation forms meaningful partitions from the image and these partitions are classified in terms of spectral, textural and geometric properties of the roads.

Segmentation is a method of grouping similar pixels together to form some subsets on the image. The distances between some selected features of the pixels are used for grouping these pixels. Various segmentation algorithms are used in road extraction. In [26], Iterative Self-Organizing Data Analysis Technique (ISODATA) is used to participate the image into connected regions of similar colors and features. In [17], the color texture information which is constructed by fusing proximity, color and the texture information is used in segmentation process. In [18], a region growing technique based on similarity, smoothness and compactness criteria is applied on the classified image to extract the shape information. In [22], Gaussian mixture modeling is used for segmentation, where the maximum likelihood estimation is used to model the pixel intensities of the image. In [28], k-means clustering is used for image segmentation and the road class is determined by selecting the segments that have elongated structure.

### 2.2.3 Classification

Classification is the method that categorizes the image pixels into different labeled groups on the basis of a training set of pixels whose category is known. Spectral, textural and geometric properties of the image pixels are used for classification. The training data is obtained by the road seed acquisition or obtained from some training database. Classification methods used in literature are probabilistic classification, fuzzy classification and artificial neural network (ANN). In probabilistic methods, road pattern is modeled with some specified parameters by using the training samples. These models are used to assign road membership values to each image pixels and the road network is constructed by thresholding these membership values. In Artificial Neural Network, a set of input parameters is related to an output set by a transformation encoding by the network weights. First, the network is constructed by calculating the network weights by using the road samples and then the image pixels are classified with the constructed neural network. In fuzzy classification, the image pixels have memberships in many different classes with different degrees, such that the membership values are constrained so that all of the membership values for a particular pixel sum to 1.

Gaussian membership function is used in [19, 7], the mean and the standard deviation of gray values are calculated by using the training areas. In [19], calculation is performed for each training area and the results are combined to form a final membership value for every pixel. In [7], parameters are calculated for each channel and the final membership values of the pixels are found by combining the membership values of each channel with an AND operator. A rank filter is performed to construct the road class.

Support Vector Machine (SVM) is designed for binary classification in [18], which obtains good generalization performance through maximizing the margin between two classes.

In [20], a neural network is constructed to classify image to road and non-road classes. The inputs of the network are a single pixel, single pixel with normalized distance value, 3x3 window and a 3x3 window with normalized distance values. 10 neuron in 1 hidden layer is used and the network trained with the road seeds. The road class is generated by calling the network for each pixel.

### 2.2.4 Road Modeling

In the road extraction algorithms with dynamic programming, generic road models are formulated taking into account photometric and geometric road properties. A polygon representing the road is generated from few seed points describing coarsely the road by using dynamic programming algorithm.

Gruen and Li [27] uses dynamic programming approach in low resolution images that roads manifest as lines. They formulate the road model according to some knowledge about the road object, such as a road pixel in the image is lighter than its neighbors on both road sides, grey values along a road usually do not change very much within a short distance, a road is usually smooth and does not have small wiggles, the local curvature of a road has an upper bound, the width of a road does not change significantly.

Dal Poz and Vale [16] are modified Gruen and Li's approach to extract the road centerlines from medium and high resolution images. With the help of road edge properties, curve fitting and energy minimization are used to model the road with a merit and a constraint function.

Active contour modeling (snake) is the other method used for modeling roads. The main concept of snakes is curve fitting through an optimization process. This optimization process is an energy minimization process that completes the snake contour which is initialized by some seed points around the curve. The energy function of a snake contains descriptions of internal and external forces, as well as external constraints. While internal forces emphasize the smoothness of the snake curve, external forces tends to move curve to edge locations in the image.

Agouris et al. [14] presents a different variation of the traditional snakes' solution. To better describe the complexity of the extracted line, additional nodes are inserted to snake contour, and redundant nodes are deleted. Along the high curvature areas, the number of nodes is increased, so the geometry of the curve is captured better. On linear segments the number of nodes are decreased, so the complexity of the snake algorithm.

Baumgartner et al.[15] use ribbon snake in high resolution image to fill gaps between road primitives that are supplied by line extraction from low resolution image. Ribbon snake differs from the conventional snake with an additional parameter for the road width.

### **2.2.5 Road Tracking**

Road tracking is an iterative line growing process that starts with a seed point and adds new segments to the road network based on some constraints defined for the road class. Road trackers typically consist of the prediction and the matching states. The possible feature positions are predicted by considering the current and the preceding steps, these possible positions are checked to decide if the prediction has been accurate and also to determine the orientation of the road.

McKeown and Denlinger [12] proposed a road tracking method based on the intensity profile matching of road cross sections. The road candidate pixels are compared with a reference profile and differences between the two profiles are measured in terms of the geometric and radiometric parameters.

In [21], Hu et al. span circular regions around the initial seed points and observe the absolute intensity changes around the seed points. Sharp intensity changes are called as cutting points. By combining cutting points, footprints around the seeds are generated. The footprints are compared with normal, L-shaped, T-shaped, X-shaped templates and the direction and the size of road is determined from the toes of the footprints.

### **2.2.6 Morphology and Filtrate**

In road extraction methods, mathematical morphology is often conducted in binary road class image to gain a useful representation to the road regions. In the mathematical framework, morphology provides an approach to the processing of digital images based on geometrical shapes. Union, intersection, complementation, dilation, erosion, thinning are generally used operators.

In [28], the road network is constructed by applying morphological operations to the road classification result. Dilation operation is used to fill the gaps on the boundaries of the road class. Connected component labeling, area calculation and removing small areas are the operations that are used to remove non-road regions such as parking lots, buildings.

In [18], morphology is used to extract road centerlines from road-class image. Morphological thinning which is implemented in a two pass erosion process is used to eliminate boundary pixels. In the first pass, the candidates of erosion are determined and in the second pass, the ones that don't break the connection are removed.

In [15], morphology is used for smoothing road boundaries. Closing operation is performed by applying dilation and erosion operations sequentially.

Doucette et al[1] are also applied morphological operators to the road class image for filling gaps and eliminating noise by closing and minimum component size filtering.

### **2.2.7 Multi-scale Analysis**

Multi-scale or multi-resolution based road extraction algorithms are stent on the approach that roads show different behaviors in different scales of the images. In low resolution images roads mainly appear as lines, while in high resolution images roads appear as elongated regions.

Baumgartner et al [15] proposed a multi-scale based approach which models roads with homogenous regions and intersections linking these regions. They extract lines in the reduced resolution image and extract edges in the original resolution images. While lines represent possible road positions, edges represent possible road sides. If a region enclosed by parallel edges is homogenous and a corresponding line exists in the reduced resolution image, the region is marked as a road segment.

Gruen and Li [27] built a particular wavelet to sharpen roads. By using a special scaling function, they derive the scale that linear features are enhanced. The maximum gradients in the corresponding scale represent the ridges and the valleys of the image which are used as candidate seed points of the road network.

### **2.2.8 Grouping and Linking**

Grouping and linking is the process that is used to build a complete road network. After extracting the primitives of the roads, these primitives are grouped and the gaps on the road network are bridged based on some features of the roads, namely similarity, continuity, proximity.

Hu and Tao [4] proposed a strategy based on a fundamental question in computer vision: "How do we (i.e., the human visual system) perceive objects (i.e., roads) in an image?". They developed a hierarchical approach for grouping and linking fragmented road segments by using Gestalt principles. A segment connectivity matrix is constructed which contains linking probabilities of the road segments. Based on this matrix, collinear, adjacent and similar lines are detected and the gaps between these lines are bridged.

In [22], grouping and linking is used to complete grid-structured road segments. The radius rotating intersection detection method is proposed to detect the intersections of the roads and the linear discontinuous road segments are extracted before applying the least squares method for detecting road centerline positions. The extracted road centerlines are grouped in terms of colinearity and the centerlines that have random directions are removed from road network. The gaps on remaining centerline network are bridged based on gap length, curvature, proximity and colinearity features.



## CHAPTER 3

### BACKGROUND METHODS

In this chapter, the fundamental subjects used in this study are explained.

#### 3.1 Canny Edge Detection Algorithm

Canny edge detection algorithm is the most popular approach used for detecting edge points [3]. The algorithm is based on the performance criterions defined to extract edges precisely which are good detection, good localization and only one response to a single edge. Good detection means finding all true edges while minimizing the number of false edges. Good localization means finding edge points as close to the actual positions of the edges. And the third criterion, one response to a single edge means finding only a single point corresponding to each true edge point. The algorithm can be summarized as follows:

**Noise reduction:** The image is smoothed with a Gaussian filter by using a specified standard deviation.

**Finding the gradients:** Local gradients and the gradient directions are computed for each pixel. The horizontal derivative and the vertical derivative are computed by using one of the edge detection operators such as Roberts, Prewitt or Sobel. The magnitude and the direction of the gradients are determined by using the horizontal and vertical derivatives.

**Non-Maximum Suppression:** The gradient directions are grouped with 45 degrees separation and the local maximums in the same direction of gradient are marked as candidate edge pixels.

**Edge Tracing and Thresholding:** The candidate edge pixels are thresholded with two different threshold values,  $T_1$  and  $T_2$  where  $T_1 < T_2$ . If the gradient of the candidate edge pixel is greater than  $T_2$ , it is said to be a strong edge pixel and if the gradient value is between  $T_1$  and  $T_2$ , it is said to be a weak edge pixel. Beside the strong edges, and the weak edges in the 8-pixel neighborhood of the strong edges are marked as true edges.

#### 3.2 Principle Component Analysis

Principle component analysis (PCA) is a variable reduction procedure that analyzes the variances and redundancies of the variables and reduces the number of variables into a smaller number of principal components.

First, the mean vector of the data set is calculated and subtracted from the data to form a zero mean data set. Then, the covariance matrix of the corresponding data set is calculated and the eigenvalues and the eigenvectors of this matrix are determined. The eigenvector with the high eigenvalues are called the principle components of the data set. Finally, the data set is multiplied by the principal components to generate a new data set in the new coordinates defined by the principle components (eigenvalues).

#### 3.3 Anti-Parallel Centerline Extraction

The ACE is designed to identify anti-parallel edges and extract road center points by finding the midpoints of these edges[1]. Input of this algorithm is the monochrome image which is obtained by directly averaging the RGB components of the input image.

The gradient signature for roads is demonstrated in Figure 3-2(a). A brighter road segment is elongated on a darker background. The gradient orientations of the either side of the road (shown with

arrows) opposes to each other. To find all of the potential road sides, edge detection algorithm is used and then 3x3 Sobel kernels are convolved with the original image to find the gradient orientation image. For an image A, horizontal and vertical derivative approximations are computed with the following operations:

$$G_X = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix} * A, \quad G_Y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix} * A \quad (3-1)$$

The gradient orientation is calculated as follows:

$$\phi = \tan^{-1}(G_Y/G_X) \quad (3-2)$$

After finding gradient orientation, the edge image and the gradient orientation image are scanned horizontally and vertically to obtain successive edge pixels p and q that satisfy the following distance and orientation conditions. Figure 3-1 illustrates the horizontal scan. The distance between p and q is w and  $w_{pd}$  is the vector projection in the gradient direction. Since  $w_{pd}$  represents the road width, it should be in the range of possible road widths.

$$w_{pq} = w \cdot \cos \phi_p \quad (\text{for horizontal scan}) \quad (3-3)$$

$$w_{pq} = w \cdot \sin \phi_p \quad (\text{for vertical scan}) \quad (3-4)$$

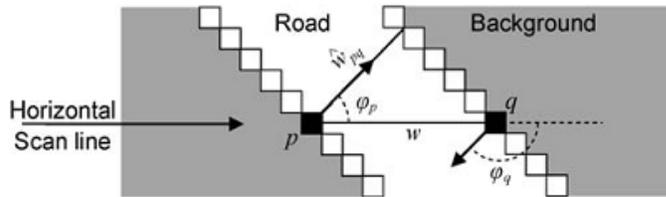


Figure 3-1 ACE Horizontal Scan Line [1]

For horizontal scan,  $w_{pd}$  is computed by the equation (3-3) and in vertical scan  $w_{pd}$  is computed by the equation (3-4). The other condition is about the angle difference between the orientations of the gradients where the angle difference should be approximately equal to  $180^\circ$ . The graph in Figure 3-2(d) illustrates the orientation graph of the selected road profile and the distance between edge pixels.

When the above conditions are assured, the mid-point of the pixels p and q is marked as the centerline, which will be used as the seed point. After scanning the image completely, noise filtering is performed by eliminating small sized connected components.

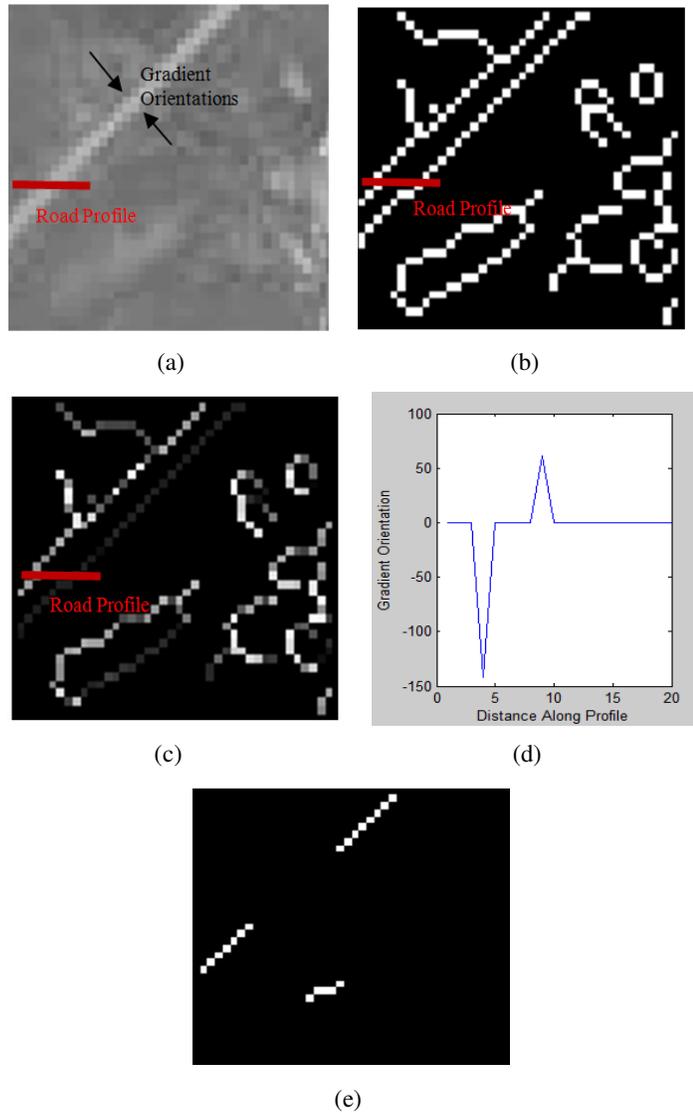


Figure 3-2 ACE Processing (a) Road Profile and Gradient Orientations of Road on input image (b) Canny Edge Result (c) Sobel Gradient Orientations of Canny Edges(Units are degrees) (d) Gradient Orientation Graph Along Selected Profile (e) The candidate road seeds from ACE

### 3.4 Hough Transform

The Hough Transform is a technique used to detect an arbitrary object (in an image), which is described with a model. The parameters of the model are the components of the space that the image is transformed to. As long as the values of the transformation parameters that map the model into the image are known, the position of the model in the image can be determined.

Duda[2] proposed a method that uses normal parameterization for lines that specifies a straight line by the angle  $\theta$  to its normal and its algebraic distance  $\rho$  from the origin. The line equation is given in (3-5) and the normal parameters of a line are shown in Figure 3-3.

$$p = x \cos \theta + y \sin \theta \quad (3-5)$$

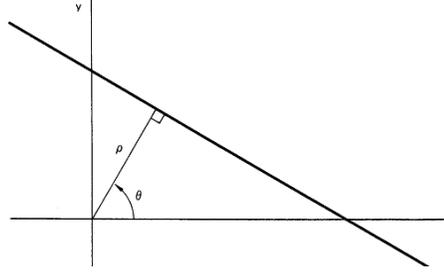


Figure 3-3 The normal parameters of a line

To compute the Hough Transform of the line pixels in a binary image, first the ranges of the parameters  $\rho$  and  $\theta$  are specified. The range of  $\theta$  is restricted to the interval  $[-90, 90)$  and the range of  $\rho$  is restricted to the interval  $[-D, +D]$ , where  $D$  is the distance of the corners of the image. With these restrictions, every  $(\rho, \theta)$  pair corresponds to a unique line in the image. The resolution of the parameters (the difference between two successive  $\rho$  and  $\theta$  values) is also specified in terms of the scale of the image. If the scale of the image is low then the difference between two successive parameter values is selected lower, else if the scale of image is high then the difference between two successive parameter values is selected higher. The binary image is scanned through the lines each corresponds a point  $(\rho_i, \theta_i)$  in parameter space. For each line, the number of pixels lying on that line is the value of  $(\rho_i, \theta_i)$  index of Hough Transform matrix.

### 3.5 Bilateral Filter

Bilateral filtering is a non-linear edge preserved smoothing technique which proposed by Tomasi at al. [36]. The filter combines the pixel values in terms of the photometric similarity and geometric closeness. The weights of the filter coefficients are determined with two Gaussian filters, namely domain filter and range filter. While domain filter measures the geometric closeness between the center pixel of the filter window and the neighboring pixels, the range filter measures the photometric similarity. Equation (3-6) represents the coefficients of the neighboring pixels assigned by range filter, where  $x_0$  is the center pixel and  $f$  is the function of intensity. Equation (3-7) represents the coefficients assigned by domain filter, where  $d$  is the distance function. Equation (3-8) shows the bilateral filter generated by combining the range and domain filters. This equation contains normalization constant which is used to maintain zero-gain.

$$r(x_i) = e^{-\frac{(f(x_i) - f(x_0))^2}{2\sigma_r^2}} \quad (3-6)$$

$$d(x_i) = e^{-\frac{(d(x_i, x_0))^2}{2\sigma_d^2}} \quad (3-7)$$

$$h(x_0) = k^{-1} \sum_{i=0}^{n-1} f(x_i) \cdot r(x_i) \cdot d(x_i) \quad (3-8)$$

$$k = \sum_{i=0}^{n-1} r(x_i) \cdot d(x_i) \quad (3-9)$$

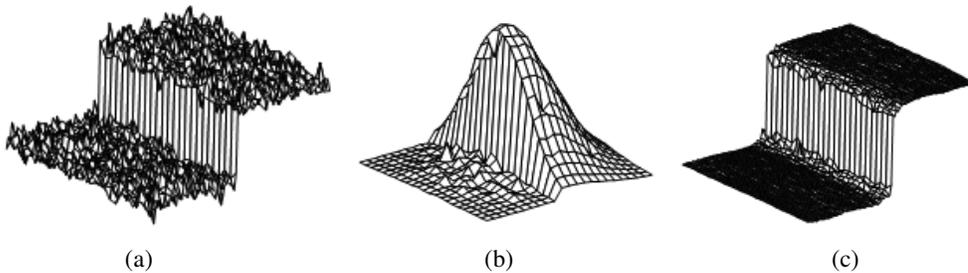


Figure 3-4 Bilateral filter (a) input (b) weight  $r \times d$  for centre pixel (c) output [36]

As shown in the figure above, the filter smooths weak intensity differences while preserving strong differences. The coefficients illustrated in Figure 3-4 (b) are formed for the case that the centre pixel lies in the bright side of the edge.

The standard edge-preserved filters are applied to the each band of the color image, which causes corrupted colors close to the edges. However, bilateral filter is applied to the three bands of the image at once by using the intensity vectors in range filter, so the perceptually similar colors are bring together and the artifacts disappear.

### 3.6 Gaussian Mixture Model

#### Gaussian Distribution:

Gaussian distribution is a continuous probability distribution that models data in terms of mean and standard deviation parameters. The distribution equation is specified in (3-10) where  $k$  is the length of the feature vector  $x$ ,  $\mu$  is the mean vector,  $\Sigma$  is the covariance matrix of the Gaussian and  $N$  is the number of data. The parameters  $\mu$  and  $\Sigma$  are estimated by the maximum likelihood estimation (MLE) which is formulized with the equations given in (3-11). Figure 3-5 shows an example of a data histogram and corresponding Gaussian distribution.

$$g(x|\mu, \Sigma) = \frac{1}{(2\pi)^{k/2} |\Sigma|^{1/2}} \exp\left(-\frac{(x-\mu)\Sigma^{-1}(x-\mu)^t}{2}\right) \quad (3-10)$$

$$\mu = \frac{1}{N} \sum_{n=1}^N x_n, \quad \Sigma = \frac{1}{N} \sum_{n=1}^N (x_n - \mu)(x_n - \mu)^T \quad (3-11)$$

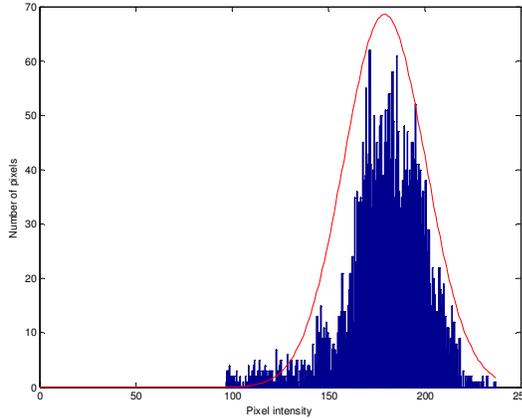


Figure 3-5 Gaussian distribution modeling

#### Gaussian Mixture Model:

Gaussian Mixture Model (GMM) is a probability density function constructed by combining Gaussian components with some weights. This model is proposed for modeling different distributions accurately. A Gaussian mixture model consists of  $K$  components is given by the equation (3-12).

$$p(x) = \sum_{i=1}^K w_i g(x|\mu_i, \Sigma_i) \quad (3-12)$$

The parameters  $w_i$ ,  $\mu_i$  and  $\Sigma_i$  are estimated by the expectation maximization (EM) method in an iterative way. The parameters are initiated with the values  $w_{i0}$ ,  $\mu_{i0}$  and  $\Sigma_{i0}$  and iteratively converged until the likelihood function reaches its maximum value. In the expectation step, probability of belonging to component  $i$  is computed for all  $x_n$  using the available parameter values. In the maximization step, new values of the parameters are computed by using  $p(i|x_n)$ . Following equations are the values of parameters at the  $m^{\text{th}}$  iteration. Figure 3-6 shows an example of the data histogram and corresponding Gaussian mixture model.

$$p^m(i|x_n) = \frac{p_i^m g(x_n|\mu_i, \Sigma_i)}{\sum_{n=1}^N p_i^m g(x_n|\mu_i, \Sigma_i)} \quad (3-13)$$

$$\mu_i^{m+1} = \frac{\sum_{n=1}^N p^m(i|x_n) x_n}{\sum_{n=1}^N p^m(i|x_n)} \quad (3-14)$$

$$\Sigma_i^{m+1} = \sqrt{\frac{1}{D} \frac{\sum_{n=1}^N p^m(i|x_n) \|x_n - \mu_i^{m+1}\|^2}{\sum_{n=1}^N p^m(i|x_n)}} \quad (3-15)$$

$$w_i^{m+1} = \frac{1}{N} \sum_{n=1}^N p^m(i|x_n) \quad (3-16)$$

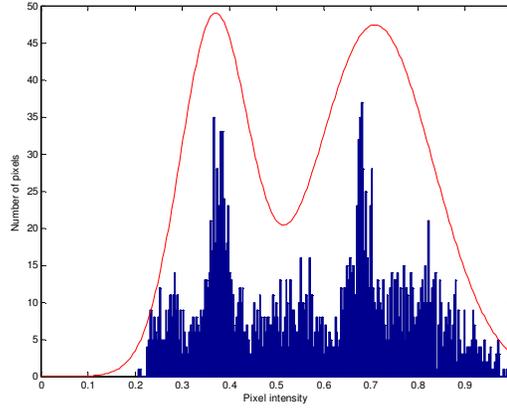


Figure 3-6 Gaussian mixture modeling

### 3.7 Features Used for Classification

#### 3.7.1 Mean and Standard Deviation

Mean and standard deviation are the basic statistical features of the images. In road extraction, these features are calculated in an image window to determine local textures. Mean represents the average of the pixel intensities in the window and the standard deviation indicates how much variation exists from the mean value. Mean and standard deviation in an image window are calculated by the equations (3-17), (3-18) and where  $N$  is the number of pixels in the window,  $I(x_i, y_i)$  is the intensity value of pixel  $i$  of image  $I$ .

$$\text{mean} = \frac{1}{N} \sum_{i=1}^N I(x_i, y_i) \quad (3-17)$$

$$\text{std} = \sqrt{\frac{1}{N} \sum_{i=1}^N [I(x_i, y_i) - \text{mean}]^2} \quad (3-18)$$

#### 3.7.2 HSV Color Space

HSV is a color space which uses cylindrical representation of colors with the components Hue, Saturation and Value. Hue is the dominant wavelength in a mixture of light, simply the color attribute that describes pure color like red, purple, yellow. Saturation is a measurement of the purity, the amount of white light diluting the pure color. Value represents the brightness relative to the white light. The translation of the HSV color components from RGB color components is given in (3-19).

$$H = \begin{cases} \theta & \text{if } B \leq G \\ 360 - \theta & \text{if } B > G \end{cases}$$

$$\text{where } \theta = \cos^{-1} \left\{ \frac{1/2 [(R - G) + (R - B)]}{[(R - G)^2 + (R - B)(G - B)]^{1/2}} \right\}$$

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

$$V = \frac{1}{3}(R + G + B) \quad (3-19)$$

### 3.7.3 Gabor Filters

Gabor filtering is one of the textural feature extraction methods which has similar characteristic with human perception. Gabor filter is composed of a complex sinusoid which is known as the carrier function and a Gaussian function which is known as the envelope function.

**Carrier Function:** The carrier function serves as sinusoidal filter centered at the origin with the spatial frequency  $\theta$  and the orientation  $\varphi$  as shown in (3-20).

$$c(x, y) = e^{2\pi\theta i(x \cos \varphi + y \sin \varphi)} \quad (3-20)$$

**Gaussian Envelope:** The Gaussian envelope maintains the spatial localization of the carrier function, which is shown in (3-21).

$$e(x, y) = e^{-\frac{(x^2 + y^2)}{\sigma^2}} \quad (3-21)$$

The spatial deviation of the Gaussian envelope ( $\sigma$ ) depends on the spatial frequency of the carrier function ( $\theta$ ). The relationship is given in the equation (3-22) where  $b$  is the half response spatial frequency bandwidth.

$$\sigma = \frac{1}{\pi\theta} \sqrt{\frac{\ln 2}{2} \frac{2^b + 1}{2^b - 1}} \quad (3-22)$$

Figure 3-7 illustrates decomposition of Gabor filter where (a) shows the 2D sinusoidal carrier function, (b) shows the 2D Gaussian envelope and (c) is the Gabor filter supplied by multiplying (a) and (b).

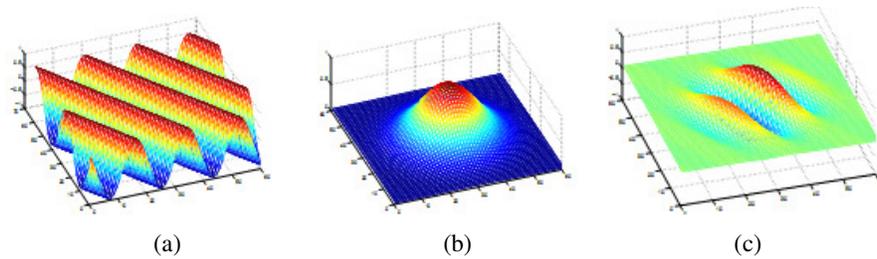


Figure 3-7 Gabor Filter Decomposition a) The sinusoidal carrier b) The Gaussian envelope c) The corresponding Gabor Filter [29]

### 3.8 Relative Neighborhood Graph

Relative neighborhood graph (RNG) is the method of extracting the shape and structure of the point sets. The structure is described in terms of the edges that connect the points. The two points in the point set are connected by an edge whenever there does not exist a third point that is closer to both than they are to each other.

For points  $p, q$ , connectivity condition is defined in (3-23) where  $d$  represents the dimension of feature space that the points are defined in and  $V$  is the set of the points that RNG will be constructed. The lune of the points  $p$  and  $q$  means the interior region formed by the intersection of two  $d$ -dimensional spheres that centered at  $p$  and  $q$  and the radius of the spheres are equal to the distance between  $p$  and  $q$  as shown in Figure 3-8 (a). Figure 3-8 (b) shows an example of RNG that is constructed on a set of planar points. The integers show the distances between corresponding points and the bold lines are the edges that provide connectivity condition of RNG.

$$[pq] \in RNG(V) \text{ iff } \text{lune}(p, q) \cap V = \emptyset, \text{ where } p, q \in R^d \quad (3-23)$$

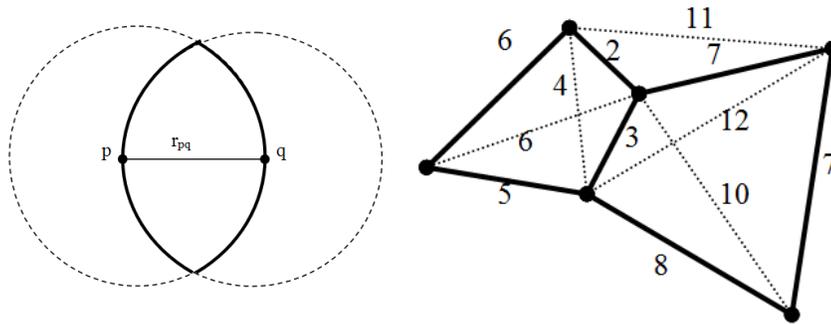


Figure 3-8 Relative Neighborhood Graph a) The lune of the points  $p$  and  $q$  b) The constructed neighborhood graph

### 3.9 Minimal Spanning Tree

Spanning Tree is a graph-based approach that is a structure composed of edges between vertices and forms a tree that includes every vertex. Minimum spanning tree (MST) is a spanning tree that the total length (or weight) of the edges is less than other spanning trees. Figure 3-9 shows a MST on a group of planar nodes. The integers show the distances between corresponding points and the bold lines are the edges that provide connectivity condition of MST.

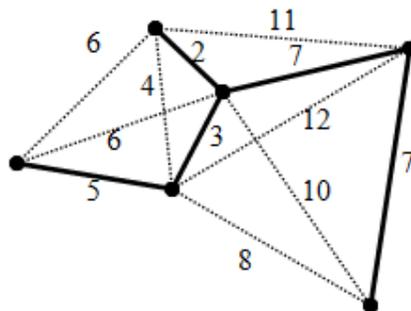


Figure 3-9 Minimal Spanning Tree

There are two algorithms to find MST. One of them is Prim's algorithm [25] and the other is Kruskal's algorithm [24].

Kruskal's algorithm starts with a tree consisting of a single vertex, continuously adds the shortest edge which does not form any loops with the edges already present, until there is no edges to add.

Prim's algorithm has two principles that any isolated vertex is connected to the nearest vertex, any isolated sub-tree is connected to the neighbor (vertex or another sub-tree) that has the shortest available link.

### 3.10 Self Organized Road Map

SORM technique is inspired from Kohonen's self-organizing map [9] and provides an approach to road centerline delineation from binary road class image [1, 8]. It is an adaptation of the K-means algorithms and differs from K-means in some respects. The method works as follows:

**Step 1:** Initialize K cluster centers ( $c_1, c_2, \dots, c_k$ ) in the road class image such that the nodes construct a grid structure on the image with a grid spacing  $\beta_{\text{grid}}$ .

**Step 2:** Generate the clusters  $S_j$  by finding the closest center for each road pixel  $x$ . The cluster generation is formulated in (3-24) where  $t$  is the number of iterations.

$$x \in S_j(t) \quad \text{if} \quad \|x - c_j(t)\| < \|x - c_i(t)\|$$

$$i = 1, 2, \dots, K \text{ and } i \neq j \quad (3-24)$$

**Step 3:** Detect the clusters whose number of samples is less than a threshold value and remove these clusters by translating their members to the nearest clusters.

**Step 4:** Update the cluster centers by finding the median of samples of each cluster. The median of the samples is found according to (3-25), where  $N_j$  is the number of samples of  $S_j$ . For determining the median of the cluster, the samples of the cluster are sorted in terms of coordinate values and if  $N_j$  is odd, the middle element of the sorted samples is selected as the cluster center; if  $N_j$  is even the average of the two middle elements is selected as the cluster center.

$$c_j(t+1) = \begin{cases} \text{sort}(x)_{(N_j+1)/2}, & \text{if } N_j \text{ is odd} \\ \frac{1}{2}(\text{sort}(x)_{N_j/2} + \text{sort}(x)_{(N_j/2+1)}), & \text{if } N_j \text{ is even} \end{cases}$$

$$j = 1, 2, \dots, K \text{ and } x \in S_j(t) \quad (3-25)$$

**Step 5:** There is also a minimum distance threshold ( $\beta_{\text{min}}$ ) between cluster centers. If the distance between two clusters is less than  $\beta_{\text{min}}$ , merge the clusters. If the cluster merging occurs, return to Step2.

**Step 6:** Iterate to Step2 until cluster centers are unchanged, such that  $c_j(t+1) = c_j(t)$ .

**Step 7:** Using the cluster center locations as input, apply a linking algorithm to construct a consistent road topology. While a MST-based linking algorithm is used in [8], a gestalt grouping algorithm is used in [1].

The process defined above is illustrated in Figure 3-10.

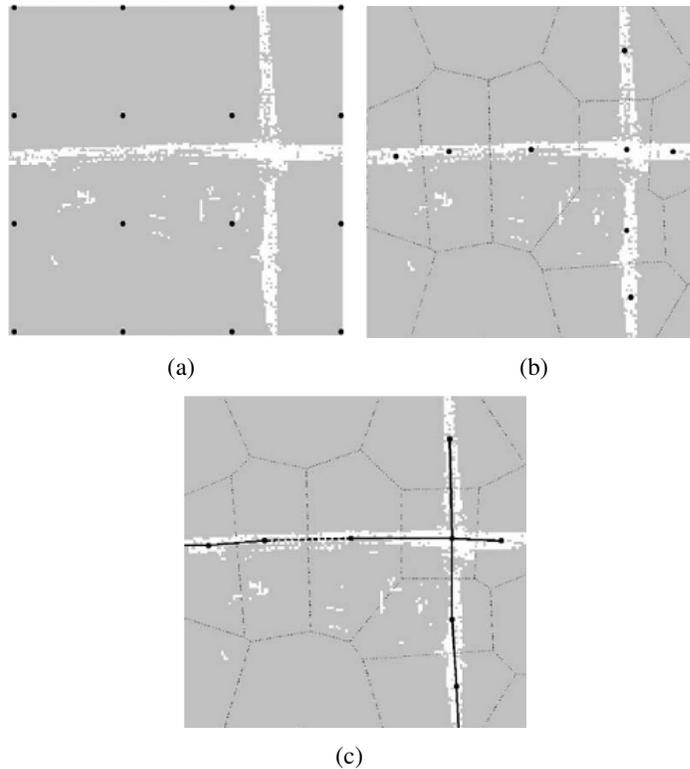


Figure 3-10 SORM Process a) Cluster Initialization b) Cluster Centers on Voronoi Graph (c) The result of SORM Algorithm [8]

### 3.11 Perceptual Grouping

Perceptual grouping is based on human visual ability to extract significant image relations from primitive image features and to group these relations for supplying meaningful higher level structures. The Gestalt psychologists have proposed a set of important principles in perceptual organization, namely, proximity, continuity, similarity, closure and symmetry [5].

For linking fragmented road segments proximity, continuity and similarity principles are meaningful. When it is thought in terms of road segment linking, definitions of these principles can be expressed as follows:

1. Continuity: Main roads are continuous, so the road segments tend to continue their shapes beyond their ending points.
2. Proximity: Road segments are close to each other and they are likely to be bridged.
3. Similarity: Segments that have similar directions are more likely to be linked.

These constraints comprise both geometric and radiometric similarities of road segments. By defining continuity, proximity and similarity constraints, the road segments are grouped and continuous road network is constructed by linking the road segments in the same groups.

### 3.12 Fuzzy Organization of Elongated Regions (FOrgER)

Fuzzy organization of elongated region (FOrgER) is proposed by Doucette at al. in order to generate larger curvilinear networks from binary road regions and construct the road network topology. FOrgER motivated from the Gestalt grouping principles (e.g. similarity, proximity and continuity) for grouping and linking elongated nodes.

The inputs of the algorithm are SORM centers and the binary road class image. The metrics used to guide topology construction are proximity, link angle and orientation angle. The proximity is the

Euclidian distance between the two given SORM centers. The link angle is the angle between the corresponding SORM centers.

The orientation of each center is the direction of the distributed nodes in its Voronoi region. In Figure 3-11,  $\theta_i$  is the orientation angle of Voronoi region  $i$  and  $\theta_l$  is the link between nodes  $i$  and  $j$ . The difference between link angle and orientation angle is called deflection angle ( $\theta_d$ ), which stands for the similarity between Voronoi regions.

The calculation of orientation angle and deflection angle is given in equation (3-26), where  $\sigma_x$ ,  $\sigma_y$  are standard deviations in x and y directions and  $\sigma_{xy}$  is the covariance.

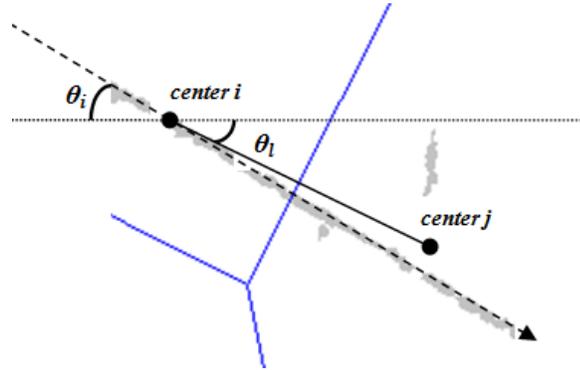


Figure 3-11 The orientation and link angle of SORM centers

$$\theta_i = \frac{1}{2} \tan^{-1}\left(\frac{2\sigma_{xy}}{\sigma_x^2 - \sigma_y^2}\right), \quad \theta_d^i = |\theta_l - \theta_i| \quad (3-26)$$

The topology construction is organized in term the proximity, link and orientation metrics. The grouping is defined in three levels.

The rules of low level grouping:

- If  $\theta_d^i$  and  $\theta_d^j$  are small, proximity is close then link\_strength  $i,j$  is high.
- If  $\theta_d^i$  or  $\theta_d^j$  are small, proximity is close then link\_strength  $i,j$  is possible.
- If  $\theta_d^i$  and  $\theta_d^j$  are large, proximity is close then link\_strength  $i,j$  is low.
- If proximity is far then link\_strength  $ij$  is impossible.

The rules of middle level grouping:

- If link\_strength  $h,i$  and link\_strength  $j,k$  are high then link\_strength  $i,j$  is high.
- If link\_strength  $h,i$  or link\_strength  $j,k$  are high then link\_strength  $i,j$  is possible.
- If link\_strength  $h,i$  or link\_strength  $j,k$  are low then link\_strength  $i,j$  is low.

In high level grouping, the linking is performed in term of linking possibilities to construct a network closure.



## CHAPTER 4

### PROPOSED ROAD EXTRACTION ALGORITHM

The proposed algorithm is a fully automated road extraction method from high resolution multispectral satellite imagery. First step of the algorithm is to analyze the image gradient and find some road centerlines based on this analysis. Then, extracted centerlines are used for training a road model based on various features of the multispectral image. Finally, classification is performed to separate road regions from non-road regions and the road topology is constructed based on the classification result. The flow diagram of the whole method is given in Figure 4-1.

When the methods in the literature are considered, our proposed algorithm is a road centerline extraction method which bases on classification. We use seed point extraction methods to make it a self-supervised algorithm and grouping and linking methods to determine the topology of the road network.

Section 4.1 introduces the seed point extraction method which consists of edge detection, ACE (Anti-parallel Centerline Extraction) and centerline linking algorithms. Section 4.2 presents the classification algorithms which use extracted seed points as the training data. Section 4.3 explains the morphological operations performed for improving classification results and determining the candidates of the road centerlines. In section 4.4 topology construction algorithms are discussed based on grouping and morphological operations.

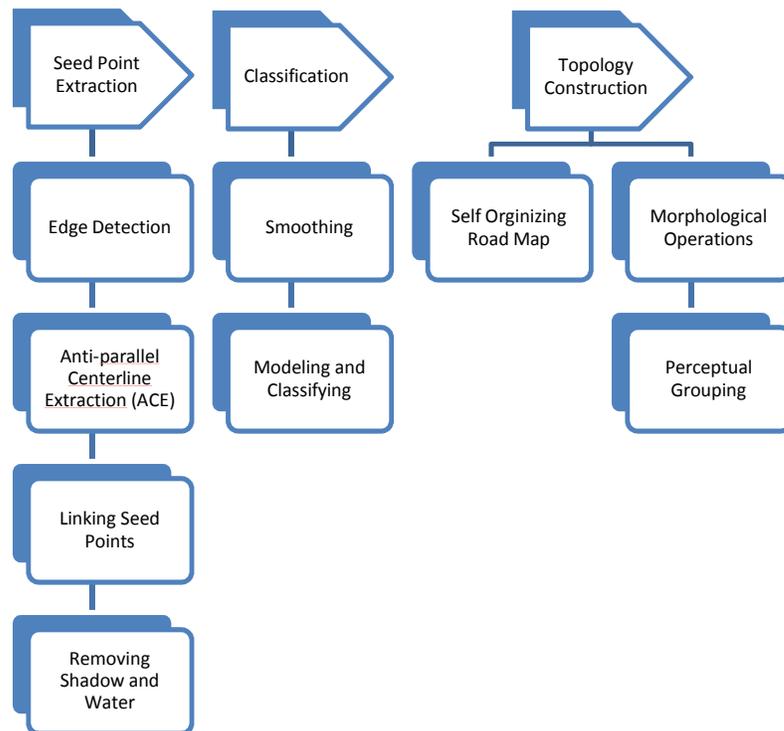


Figure 4-1 Flow diagram of road centerline extraction algorithm

## 4.1 Seed Point Extraction

Most of the available road extraction methods need a human operator to benefit from the interactive user-provided information to have some prior information about the road seeds, directions, width, road intersections and to allow manual editing. Depending on the need of the human operator, the road extraction strategies are categorized as semi and fully automated. The ACE (Anti-parallel Centerline Extraction) algorithm is proposed by Doucette et al. [1] to minimize the need of human operator by automatically finding the seed points of the roads. The main idea is detecting the road center points by considering geometric properties of the road structures.

When looking at high-resolution imagery whose ground sample distance is greater than 1.0 m/pixel, roads are elongated regions bounded by parallel edges. These parallel edges with opposing gradients are called Anti-Parallel edges. ACE is designed to extract road center points by finding the midpoints of these edges.

Edge detection is the first step of the seed point extraction algorithm, which is followed by the ACE algorithm. Third step is linking the extracted seed points. Finally, the seed points belonging to shadow regions are removed from the seed points set.

### 4.1.1 Edge Detection

Edges are the sharp intensity changes in the imagery. These intensity changes correspond to the boundaries of the objects, surface markings, depth changes, illuminations changes and so on.

In road extraction, edges are so important that they give clues on structural properties of the images. In low resolution images, roads appear as lines that correspond to single edges. When looking at high-resolution imagery, roads are elongated regions bounded by parallel edges. These parallel edges with opposing gradients are called as Anti-Parallel edges.

Various edge detectors exist in the literature, such as Sobel, Laplace, Canny, Nalwa–Binford, Sarkar–Boyer. In [31], Heath et al. represent a comparison between four edge detector algorithms. They demonstrate that if the images to be analyzed are similar in content Canny edge detector is the best edge detection algorithm amongst the others. They also mention that most of the papers that describe new edge detectors compare the results with the results supplied by Canny. This means that Canny algorithm is regarded as the best among its counterparts. Since we used satellite images that have similar contents, Canny edge detector is chosen as the edge detection algorithm.

### 4.1.2 ACE (Anti-parallel Centerline Extraction)

As mentioned in section 3.3, the ACE algorithm which is designed to detect road center points by considering the geometrical properties of the road structures is preferred as the road seed point detection algorithm. First, the edges on the image are extracted with the Canny edge detector. Second, the gradient orientations of the extracted edge points are calculated by applying the Sobel operators in the neighborhood of these points. Third, the edge image is scanned horizontally and vertically to identify anti-parallel edges and extract road center points by finding the midpoints of these edges.

The ACE Algorithm uses two parameters namely the range of road width and the angle difference between edge gradient orientations. The range of road width is determined in terms of the scale of the image. Since a single road lane width is constant for each country (3.6 m in Turkey), the minimum road width is selected as a single road lane width and the maximum road width is selected as the multiplication of a single road lane width and the maximum number of lanes in the image. The angle difference between edge gradient orientations is selected in the range of  $[\frac{3\pi}{2} \ 5 \ \frac{\pi}{2}]$  which shows good performance for all images.

In [1], multispectral image content is also used to improve the output of ACE. The ACE algorithm is applied to each image layer and the result is merged to find the final result, which is called Discrete Layer Gradient Analysis (DLGA). Merging is performed by finding union of the results of the layers and morphological thinning is applied to find final result. Another approach for using multi-spectral content is performing Principal Component Analysis (PCA). PCA analysis is applied to supply monochrome image from RGB layers.

DLGA increases the number of samples extracted by ACE but the correctness of the samples decreases. PCA gives better results than using the gray level which is the average of the RGB layers.

To improve DLGA results, merging is performed by merging the individual layer ACE results by intersection. This approach improves the correctness of ACE results; the disadvantage of this approach is the decrease in the number of samples.

Considering the previously mentioned observations, principal component image is preferred as the input of the ACE algorithm.

#### 4.1.3 Linking Seed Points

Instead of directly using the ACE results as the training data, only the straight lines longer than a predefined threshold on the ACE image are selected as the training data. The minimum line length threshold is selected by considering the ACE response of buildings. Selecting the threshold as longer than the length of the ACE response of buildings eliminates the seed points belonging to buildings and some other non-road regions.

To find the straight lines from the ACE output, two methods are applied. One of them is the Hough transform and the other one is the perceptual grouping.

##### 4.1.3.1 Hough Transform

As introduced in section 3.4, Hough transform is a method of modeling the image pixels using some parameters. For straight line modeling, Hough transform is formed in the normal parameter space, which specifies a straight line by the angle  $\theta$  to its normal and its algebraic distance  $\rho$  from the origin.

For linking, the seed points are obtained by the ACE algorithm, the ACE output image is transformed to normal parameter space and the following steps are performed to extract the lines in the image: First, the peak points of the Hough Transform matrix are found. The maximum number of peaks can be limited. The locations of the pixels which have contributed to a peak (associated with a line) are determined. These are the pixels belonging to a line. The neighboring pixels are merged to construct the corresponding line segment. In this step, small gaps are ignored. The line segments longer than the predefined threshold are labeled as the lines in the image. 3x3 neighborhoods of Hough lines are selected as road seed points [1].

In Figure 4-2 (a), the edges are shown with red color and the ACE results are shown with blue. Figure 4-2(b) shows the extracted Hough lines with blue color on ACE output image.

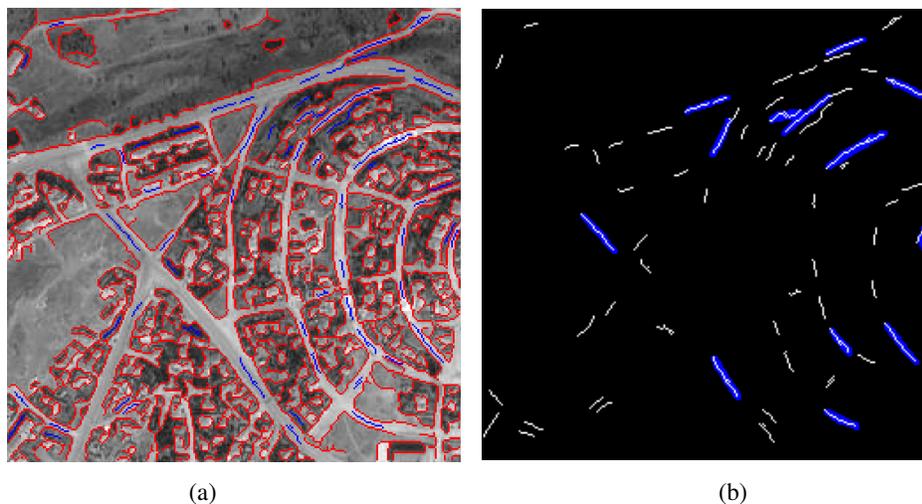


Figure 4-2 Seed Point Linking a) The output of ACE b) Lines found with Hough Transform

#### 4.1.3.2 Perceptual Grouping

As introduced in section 3.11, the perceptual grouping method is used to link fragmented line segments by using proximity, continuity and similarity principles. For linking seed points, the grouping is performed on ACE output image and for topology construction which is introduced in section 4.4, the grouping is performed on the road centerline image. Although, the ACE output image generally consists of straight line segments, the road centerline image consists of straight and curvilinear line segments. Therefore, the grouping algorithm is proposed as it is used for both the ACE output and the road centerline image.

In [22], only the straight line segments are labeled and grouping is performed on these straight line segments. In [4], straight collinear segments, polylines, and the curvilinear segments are grouped separately. In this thesis, any connected line segment is labeled and grouping is performed on these line segments if the local curvatures around the endpoints are similar to each other, which is one of the contributions of this thesis. With this approach, different type of road networks can be extracted. Figure 4-3 shows some examples about the proposed grouping and linking method, in (a) unconnected line segments are illustrated, in (b) the orientations of the local curvatures around the endpoints of the lines are shown. Since the orientations are similar and the line segments are close to each other, the segments are linked and a connected line is formed, shown in (c).

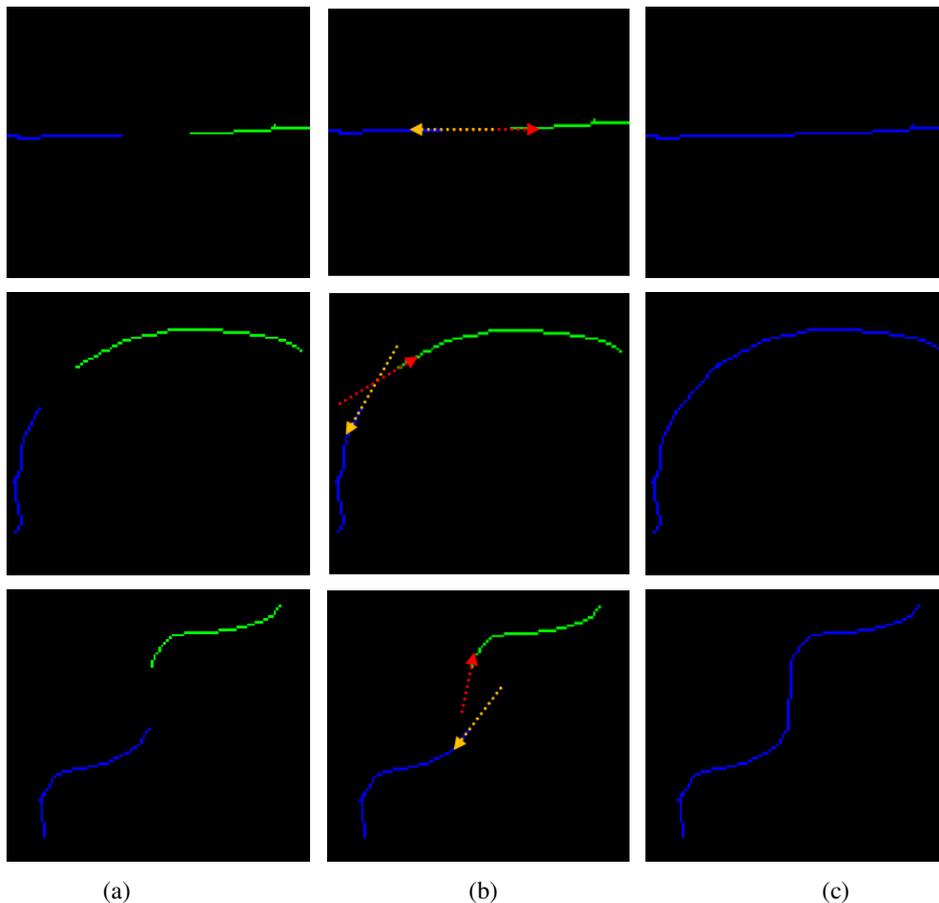


Figure 4-3 Grouping and linking line segments a) labeled line segments b) orientations of local curvature around endpoints (c) result of grouping and linking

The grouping method is proposed for binary images and it assumes that segments are one pixel wide. First, labeling line segments algorithm is performed to determine pixel locations of the road segments. After labeling, the Gestalt grouping principles are used to find the segments that belong to the same road. Then, gaps between grouped segments are filled to construct a higher level road structure.

### Labeling Line Pixels (LLP)

This algorithm produces line segments by labeling the pixels that constitutes a connected line. The branches of the line segments are labeled with different labels from the main branch, by this way small line branches may be discriminated from longer line segments.

LLP algorithm assumes that the lines are one pixel wide, so morphological thinning on ACE output is required before applying LLP algorithm to obtain one pixel wide lines.

The algorithm is as follows:

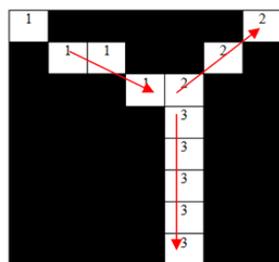
**Step 0.** n is the label number which is initialized with 1 at the beginning.

**Step 1.** If there is no line pixel in the line image, go to Step 3. Else, select a pixel in the line image, label it with label n and delete this pixel from the line image.

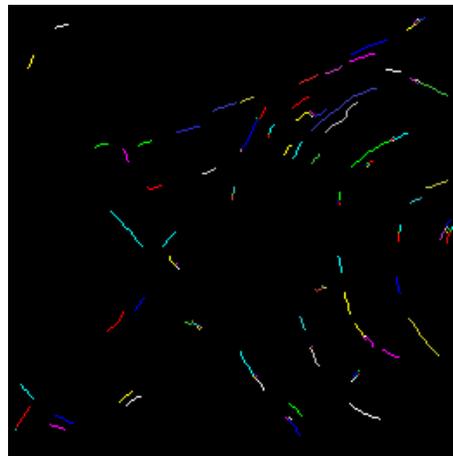
**Step 2.** Examine 3x3 neighborhood of the reference pixel.

- If there is one line pixel in the neighborhood, label it with label n, delete this pixel from the line image and repeat Step 2.
- If there is more than one pixel in the neighborhood, for each neighboring pixel, increase n, label the pixel with label n, and delete it from the line image. Repeat Step 2 for each neighboring pixel.
- If there are no pixels in the neighborhood, increase n and return to Step 1.

**Step 3.** Each label represents a line in the image and the pixels having the same label are the pixels that belong to the same line.



(a)



(b)

Figure 4-4 Labeling line pixels a) Labeling connected braches b) The result of LLP algorithm

### Grouping Based on Gestalt Principles

Gestalt principles namely continuity, proximity and similarity are used as primary constraints for linking labeled line segments. Line segments that are closer to each other and have similar angles are concatenated. Considering that line pixel labeling algorithm finds not only the straight lines but also the curvilinear lines, the slopes of the segments are calculated only near the endpoints because the local curvilinearity of the road segments is small.

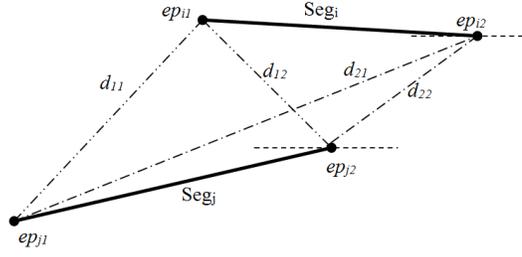


Figure 4-5 Grouping Parameters

Figure 4-5 illustrates the parameters that are used for grouping.  $ep_{i1}$ ,  $ep_{i2}$ ,  $ep_{j1}$  and  $ep_{j2}$  are the endpoints of segments of  $seg_i$  and  $seg_j$ . The gap between the segments is the minimum distance between endpoints as shown in the equations (4-1) and (4-2).

The line directions are calculated around the corresponding endpoints. When it is assumed that  $ep_{i2}$  and  $ep_{j2}$  are the nearest end points, the direction of  $seg_i$  ( $\alpha_i$ ) is calculated near  $ep_{i2}$  and the direction of  $seg_j$  ( $\alpha_j$ ) is calculated near  $ep_{j2}$ . As shown in Figure 4-6, the direction of  $seg_i$  is calculated using the line from  $(x_{i1}, y_{i1})$  to  $(x_{i2}, y_{i2})$  and the direction of  $seg_j$  is calculated using the line from  $(x_{j1}, y_{j1})$  to  $(x_{j2}, y_{j2})$  with the equation (4-3).

$$d_{gap} \leq \max\_gap\_length$$

$$\|\alpha_i - \alpha_j - \pi\| \leq \pi/6$$

$$\|\alpha_i - \alpha_{gap} - \pi\| \leq \pi/6 \quad (4-1)$$

$$d_{gap} = \min(d_{11}, d_{12}, d_{21}, d_{22})$$

$$\text{where } d_{11} = \|ep_{i1} - ep_{j1}\|, d_{12} = \|ep_{i1} - ep_{j2}\|$$

$$d_{21} = \|ep_{i2} - ep_{j1}\|, d_{22} = \|ep_{i2} - ep_{j2}\| \quad (4-2)$$

$$G_{yi} = y_{i2} - y_{i1}, G_{xi} = x_{i2} - x_{i1}$$

$$\alpha_i = \begin{cases} \tan^{-1} G_{yi}/G_{xi}, & G_{yi} \geq 0 \text{ and } G_{xi} \geq 0 \\ \tan^{-1} G_{yi}/G_{xi} + \pi, & G_{xi} < 0 \\ \tan^{-1} G_{yi}/G_{xi} + 2\pi, & \text{otherwise} \end{cases} \quad (4-3)$$

The direction of the gap filling segment is also important. Because, short parallel line segments ensure similarity and proximity principles but they do not satisfy continuity principle. The direction of the gap filling segment ( $\alpha_{gap}$ ) is calculated using the line from  $(x_{i1}, y_{i1})$  to  $(x_{j1}, y_{j1})$ .

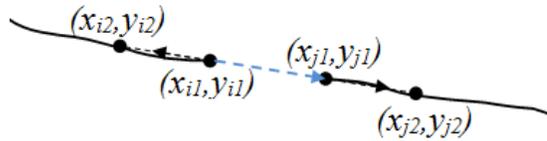


Figure 4-6 Determining the directions of line segments

All of the labeled lines are compared with all other labeled lines. If two lines ensure the grouping constraints shown in (4-1), they are merged and the gap between merged lines is filled. This process is repeated until there are no remaining un-compared line pairs left.

### Bridging Gaps

Bridging the gaps between the merged lines is necessary to construct the road structure as a whole. To bridge a gap between two lines, one of the nearest end points of the lines are selected as the starting point of the bridge and the other is selected as the end point of the bridge. The bridge direction is calculated from the starting point to the end point and three neighboring pixels of the starting point that are compatible with this direction are examined. The directions from these neighbors to the end point of the bridge are also calculated and the neighbor that has a direction nearest to the bridge direction is added to the bridge [6].

For instance in Figure 4-7, the starting point is  $(x_1, y_1)$  and the orientation of the gap filling segment is in the interval  $[-3\pi/8, -\pi/8]$ . Then, only the points  $(x_{i+1}, y_{i1})$ ,  $(x_{i1}, y_{i1+1})$ , and  $(x_{i1+1}, y_{i1+1})$  are examined. The point that is appropriate for the orientation is  $(x_{i1+1}, y_{i1+1})$  so it is added to the bridge. This process is continued until the endpoint of the gap, namely the point  $(x_2, y_2)$  is reached.

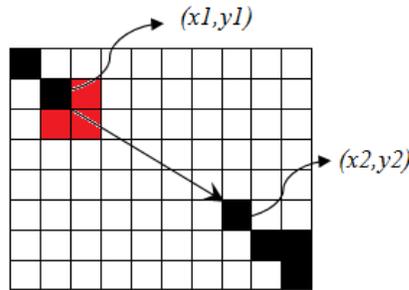


Figure 4-7 Bridging gaps between two lines

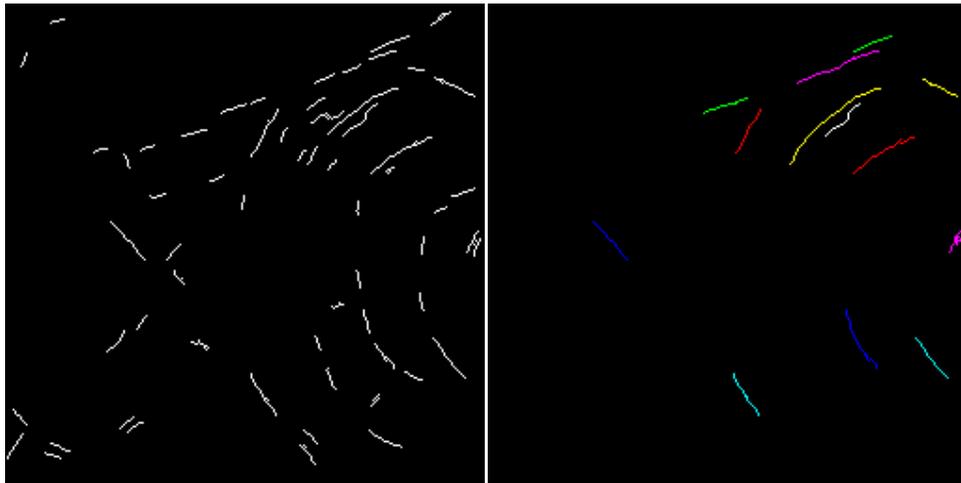


Figure 4-8 ACE Image and Perceptual Grouping Result

After the bridging step, for eliminating non-road responses of the ACE algorithm, the grouping result is thresholded and the lines smaller than a predefined line length, are erased from the seed point image.

When the results of Hough transform are compared with the results of perceptual grouping in terms of the correctness of the seed points (belonging to the road), perceptual grouping is better than the Hough transform. Therefore, perceptual grouping method is the preferred method for linking the seed points. The result of perceptual grouping is used to form the training set for the classification. Since the neighbors of the road center pixel also belong to the road, the pixels between road centers and the corresponding edges are also added to the training set. Figure 4-9 shows the training seeds on the edge image.

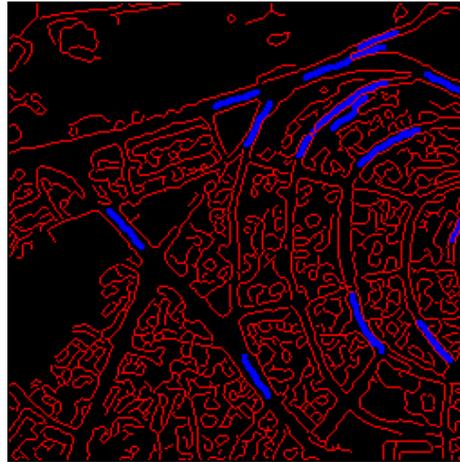


Figure 4-9 Selected seed points on the edge image

#### 4.1.3.3 Fuzzy Organization of Elongated Regions (FOrgER)

As explained in section 3.12, FOrgER is a graph-based method for constructing a road network from distributed road regions. The algorithm bases on Gestalt principles like perceptual grouping. In [1], Doucette et al. use FOrgER to link ACE results to enlarge the training set.

In Figure 4-10 (a), the red line are highly probable links, greens are possible ones and the yellows are lowly probable links. When these links are organized, the final network is constructed, Figure 4-10 (b).

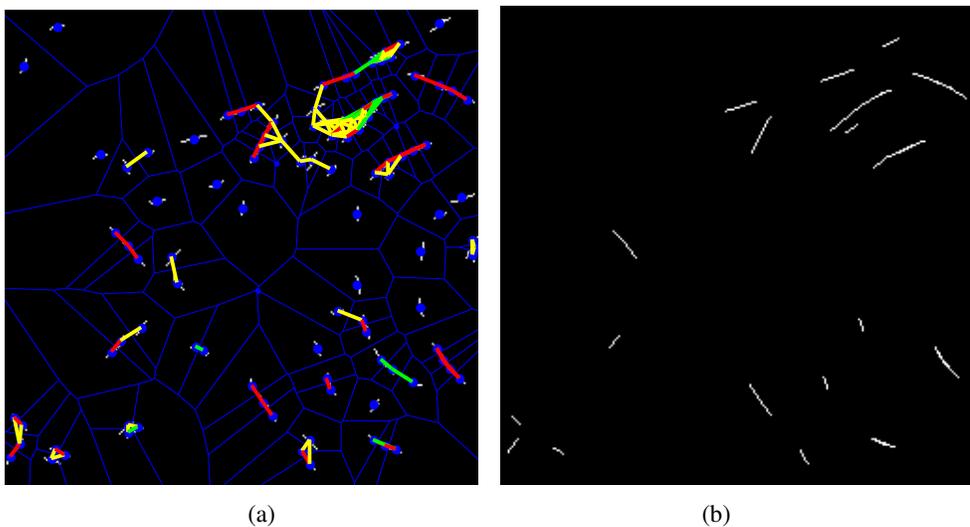


Figure 4-10 FOrgER Algorithm (a) Link strengths on Voronoi Diagram (b) Linked centers

#### 4.1.4 Removing Shadow and Water from Seed Points

With the previous steps of the seed point extraction algorithm, the elongated long regions are detected. Although these elongated regions generally belong to the road regions, in some satellite images, shadow and water also exist as elongated regions. To decide whether the seed points belong to the road region or shadow and water regions, RGB information is used.

Usually, in the detection of shadow and water, IR band is used. However, in this study, we only used RGB bands to detect shadow and water pixels. Since using RGB representation of the image for shadow detection gives erroneous results, HSV representation is preferred. Shadow and water have similar spectral characteristics in HSV space. They both have higher values of saturation and lower values of value components.

The image is firstly converted to HSV representation from the RGB. Then, a mask which filters the seed points that have higher values of saturation and lower values of value components on the HSV space. After applying this mask, shadow and water pixels are removed from the set of seed points. Figure 4-11 shows an example image that contains elongated shadow regions, with the shadow detection step these regions are removed from the set of seed points. Figure 4-12 illustrates the removal of the elongated water regions from the set of the seed points.



Figure 4-11 Shadow removal from the seed points a) Before shadow removal b) After shadow removal



Figure 4-12 Water region removal from the seed points a) Before water region removal b) After water region removal

## 4.2 Classification

### 4.2.1 Smoothing

In order to eliminate noises in the image and obtaining a smooth classification result, a smoothing algorithm is decided to be experienced. In order to maintain edge information, an edge preserving algorithm (bilateral filter) is used. First, RGB image is transformed to Lab color space, which is suggested in [36] and filter is applied to the transformed image. The variance of the range filter is selected as a small value in order to preserve comparatively weak edges.

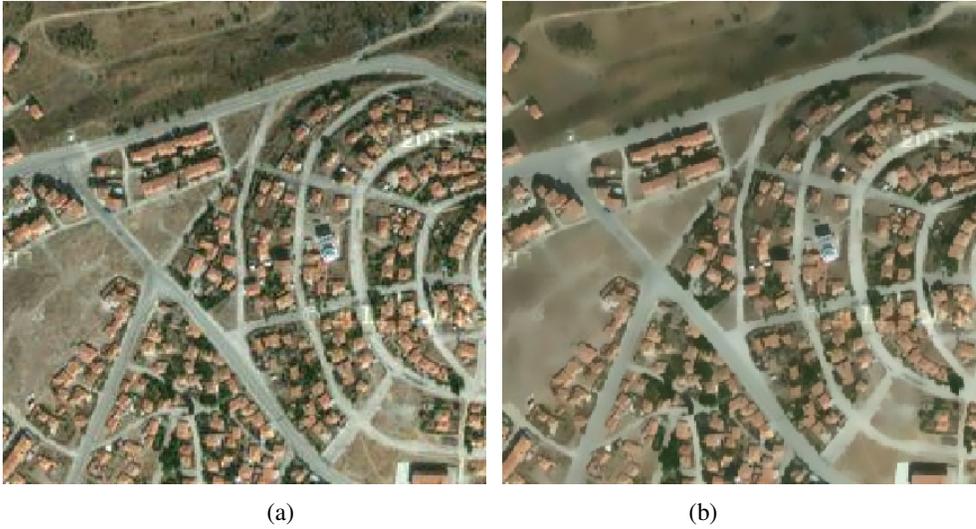


Figure 4-13 Bilateral filter (a) input image (b) output image

#### 4.2.2 Modeling and Classifying

There are two different approaches that are used in image classification such as pixel-based and object based approaches. In pixel-based approaches (PBA), the features of each pixel are used as the inputs of the classification algorithm. In object-based approaches (OBA), firstly the image is segmented by using a segmentation algorithm and the features of each segment are used as the inputs of classification. There are many papers comparing these two approaches for different data sets, but we need to know the effects of these approaches in the aspect of the road extraction. In [13], Hauptfleisch conducts an experiment about this subject. He performs pixel-based and object-based classification in 50 different scenes by fusing different classifiers with a fusion algorithm and then he applies topology construction algorithm on the classification results. The evaluation which is given in Table 4-1 is calculated by using the results of topology construction and the manually constructed ground truth images.

Table 4-1 The evaluation results of PBA and OBA [13]

	Correctness	Completeness	Quality
<b>PBA</b>	83.8	73.5	64.4
<b>OBA</b>	83.4	73.4	64

The evaluation results show that using pixel-based or object based methods for road extraction purposes gives similar results. Therefore, considering the computational complexity, the pixel-based approach is selected in the classification step.

Many road extraction methods use spectral statistics (mean and variance) of the roads for classification [1, 7, 19]. Textural features are also used to improve the classification results [17]. In this thesis, four different features and their combinations are experienced, such as the mean of 3x3 neighborhood, the variance of 3x3 neighborhood of gray image, HSV components and Gabor filters. Experiments show that HSV and Gabor features are stronger classifiers than the gray scale mean and the variance. HSV features generally produce good results in different images, however if the color spectrum of the road pixels is similar to other objects in the image, it produces noisy results. In order to avoid this artifact, Gabor filters are included to the classification step to take into account the textural features of the road pixels.

Gabor filter is a directional filter that consists of a sinusoidal wave with a frequency and orientation. Therefore, using a single Gabor Filter gives the texture along the corresponding orientation and the frequency. A circular symmetric version of Gabor filter proposed in [32] is used to find a rotation

invariant texture. Gabor filter responses are calculated for 30° angle separations and the frequency values recommended in [32]. The frequency values are calculated by (4-4) where M is the size of the image and  $i = 1, 2, \dots, \log_2(M/8)$ .

$$F = (0.25 - 2^{i-0.5}/M)(0.25 + 2^{i-0.5}/M) \quad (4-4)$$

Two alternative methods that are used for classification are defined below:

**Normal Distribution:**

Based on the Gaussian distribution models of road and non-road training sets, the road and non-road membership values of all image pixels are calculated. The pixels that have high road membership value and low non-road membership value are marked as road class pixels, and the others are marked as non-road class pixels. This rule bases on Bayesian decision rule, but there is a little difference such that we subtracted non-road membership from road membership value and compare this result with a probability threshold. Figure 4-14 shows the classification result obtained by normal distribution.

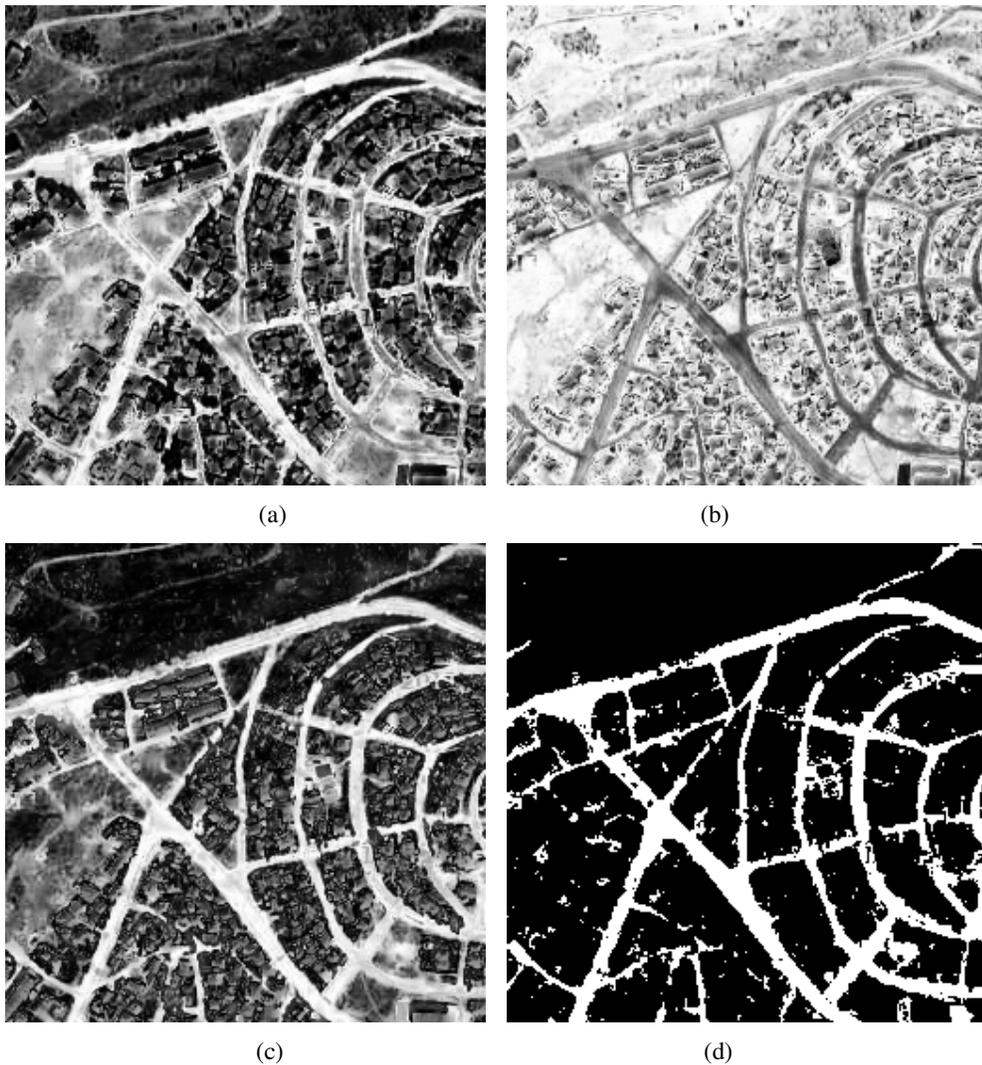


Figure 4-14 Gaussian membership values (a) membership of the road class (b) membership of the non-road class (c) posterior probabilities for road class (d) result of thresholding

### Gaussian Mixture Modeling:

In some images, different road types may exist, so training samples which have different characteristics are obtained. Modeling these training samples with normal distribution causes false responses in classification result. To overcome this drawback, we proposed to use Gaussian mixture model in classification step.

Figure 4-15 shows the classification result of the reference image. Since the roads in the image have similar characteristics and means of the Gaussians are near to each other, a similar result with normal distribution is obtained.

Figure 4-16 illustrates the classification of the image that consists of two different road types. Since different characteristics can be modeled with GMM, the classification result contains whole road network.

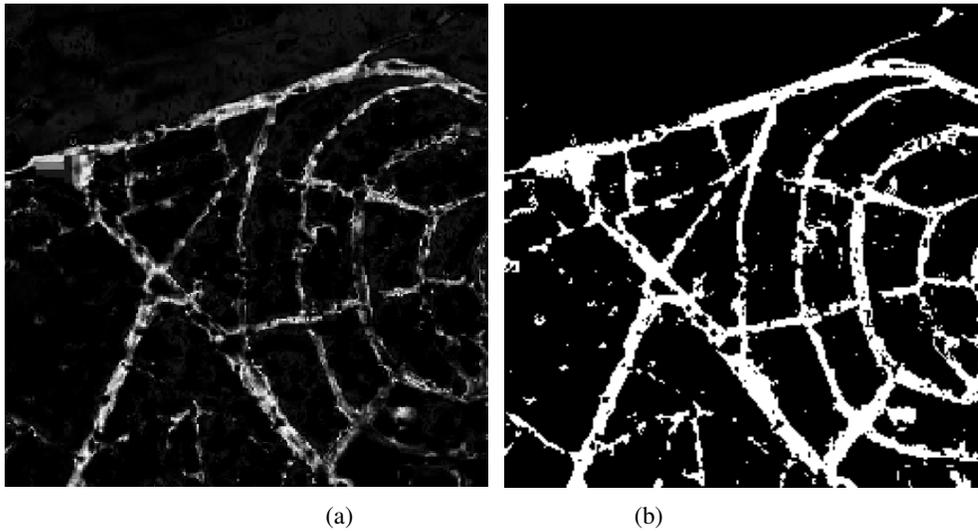


Figure 4-15 Classification based on Gaussian mixture model (a) membership of road class (b) the classification result

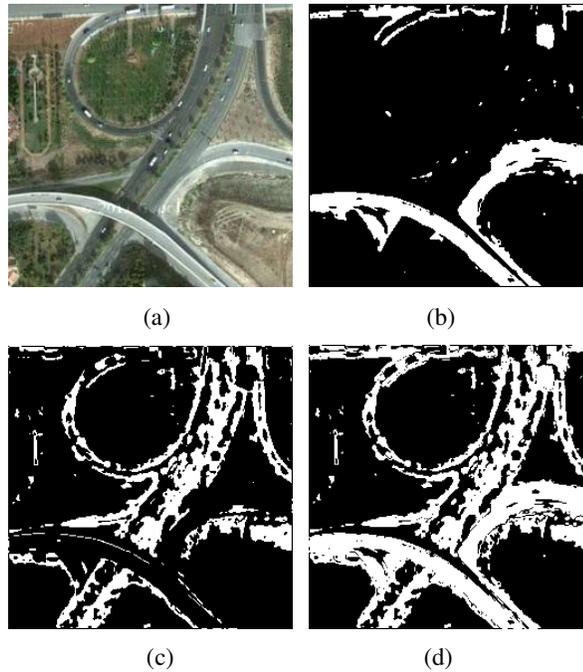


Figure 4-16 Classification based on Gaussian mixture model (a) original image (b) First cluster (c) Second cluster (d) The final result

### 4.3 Morphological Operations

Morphological operations are used to obtain a compact and noise free classification result before topology construction algorithms.

In order to avoid noises distributed on the classification result, the regions that have smaller area are removed from image and to fill small gaps in the road regions, hole filling is performed.

Closing operation is used for smoothing the road class image. A square shaped structuring element is used for this purpose since the orientation of the roads may be any direction. Finding the optimum size of the structuring element is hard such that large structuring element integrates noisy regions to the road regions and small structuring element is not enough for smoothing the road regions and the boundaries. By considering that edges separate roads from surroundings, a solution to this problem is proposed. A large structuring element is used for closing operation and then edge pixels are subtracted from the resulting image for separating roads from surroundings. Thinning is performed before perceptual grouping. With the thinning operation, the skeleton of the road network is generated. The result of the morphological operations is the rough topology where perceptual grouping is used for eliminating noises.

### 4.4 Topology Construction

Topology construction is used to extract the road network from the road candidate pixels. One of the methods for topology construction is finding regional road clusters with Self Organizing Road Map (SORM) algorithm (introduced in section 3.10) and linking the centers of these regional clusters [1]. Three different methods which are proposed in [1, 8, 13] are performed for linking the SORM centers. These are graph-based methods of generating progressively larger curvilinear networks, namely Minimal Spanning Tree (MST), Relative Neighborhood Graph (RNG) and Fuzzy Organization of Elongated Regions (FOrgER).

Figure 4-17 shows the MST, RNG and FOrGER edges of the SORM centers in the voronoi diagram. RNG algorithm produces better results than the MST algorithm since it allows closed loops within the network. FOrGER is good in linear regions but not good in curvatures.

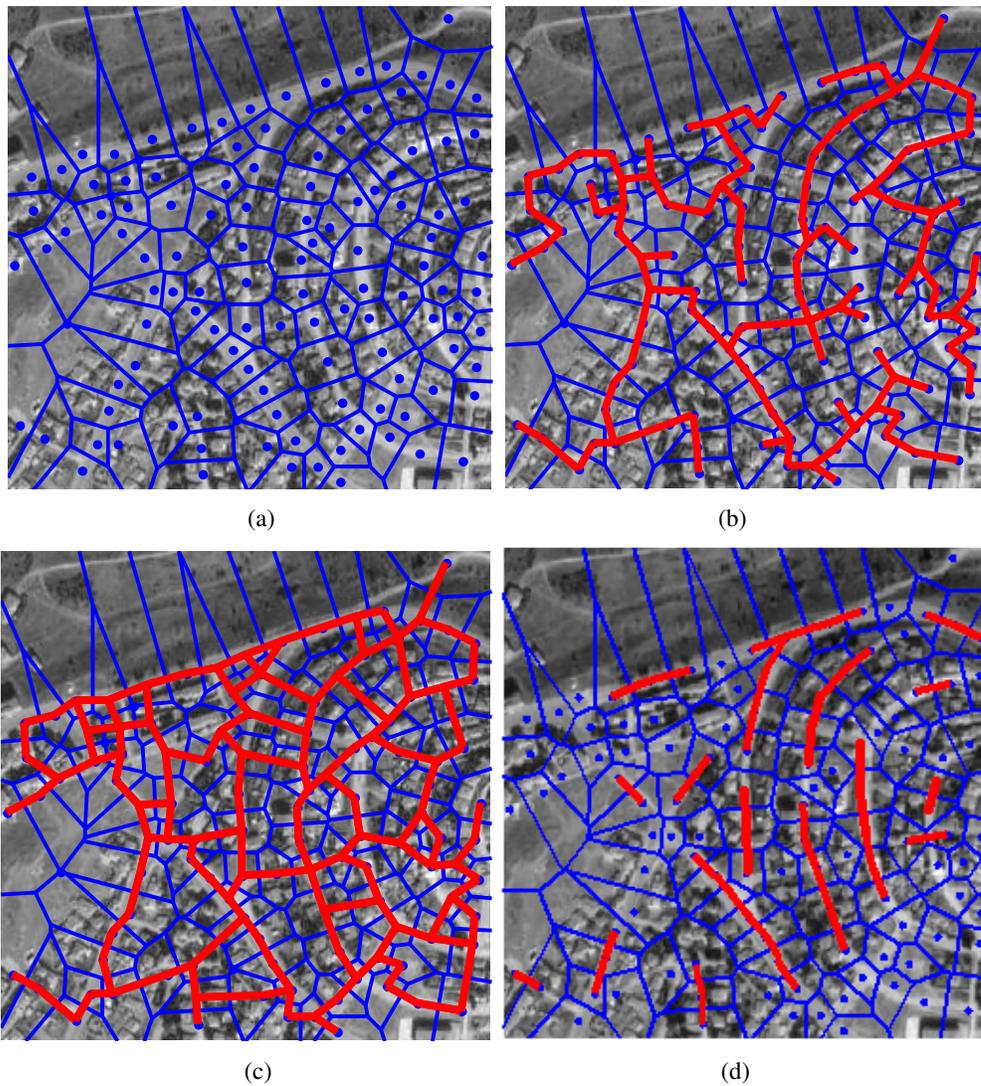


Figure 4-17 Linking SORM Centers (a) SORM Centers (b)with MST (c) with RNG (c) with FOrGER

The other method used for defining the road topology is applying morphology and perceptual grouping. Grouping and gap filling are performed with the methods described in section 4.1.3.2. After grouping the lines, the ones shorter than a minimum connected line threshold are removed from the image.

In Figure 4-18, (a) is the classified image after noise removal, closing and gap filling operations, (b) is the image after removing edge pixels, (c) is the result of thinning and (d) is the result of perceptual grouping.

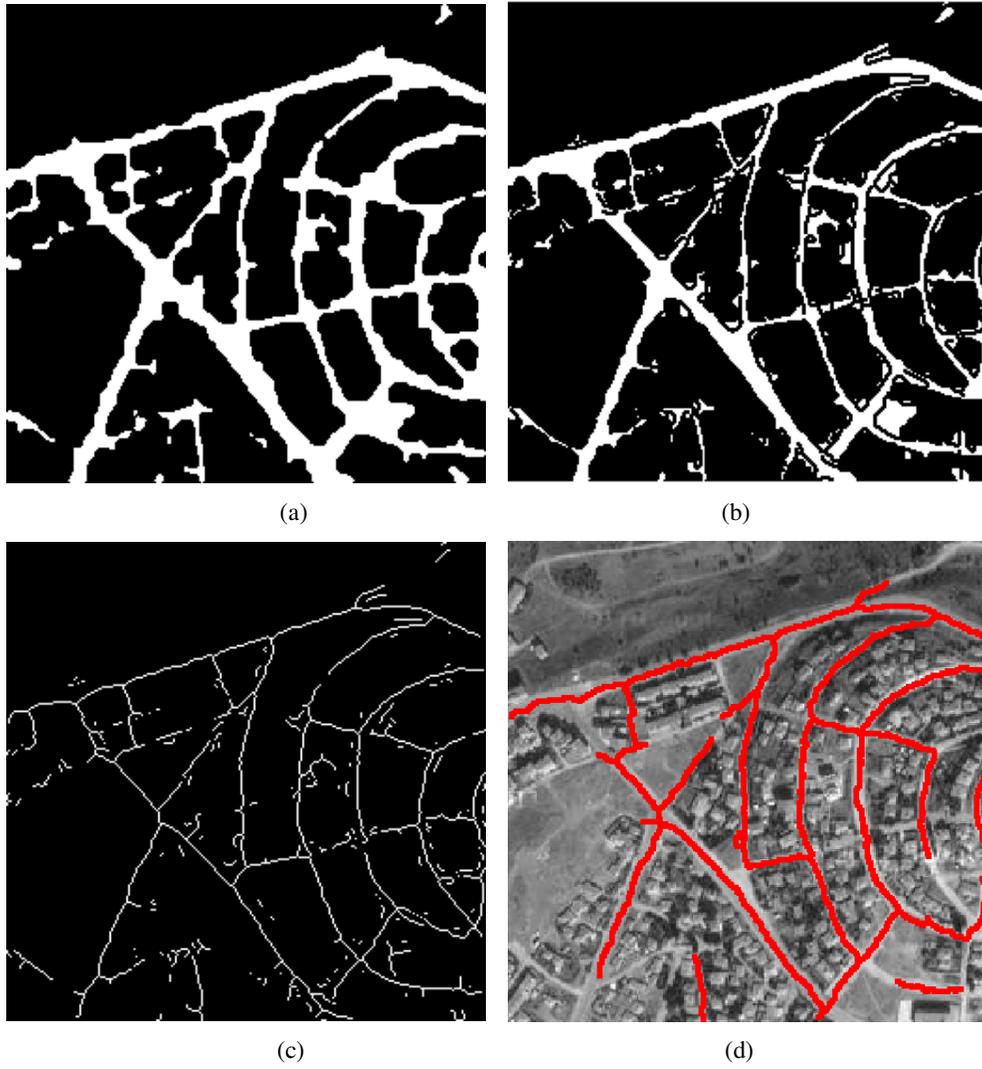


Figure 4-18 Morphology and Perceptual Grouping (a) Classified image after closing and hole filling (b) Subtraction of edges (c) Morphological Thinning (d) Result on the input image



## CHAPTER 5

### EXPERIMENTAL RESULTS

In this chapter, the experimental results of the implemented algorithms are given. The experiments are conducted on an algorithm which is proposed by Doucette et al. [1]. Figure 5-1 shows the base algorithm and the alternatives experienced to find the optimum solution. For each experiment, previously selected methods are used for previous steps and the methods of baseline algorithm are used for following steps. After generating proposed algorithm, the baseline algorithm and the proposed algorithm are compared with another experiment.

To evaluate the performances of alternative methods, 20 high resolution satellite images with 1m resolution and 1024 x 1024 sizes are used. Ground truth data are constructed manually by marking all main and side road pixels. During the evaluation of seed point extraction and classification methods, the classification results are considered, since the computational complexities of topology construction algorithms are high for some test images. The performance metrics used in these experiments are obtained by pixel by pixel comparison.

In the evaluation of topology construction algorithms, performance metrics are calculated by defining a buffer width of 5 pixels around the reference and extracted data. The method is explained in Section 5.1.

Most of the algorithms examined in this thesis need some parameters. In order to obtain most convenient ones, parameter tuning is performed by using three different images such as urban, suburban and rural.

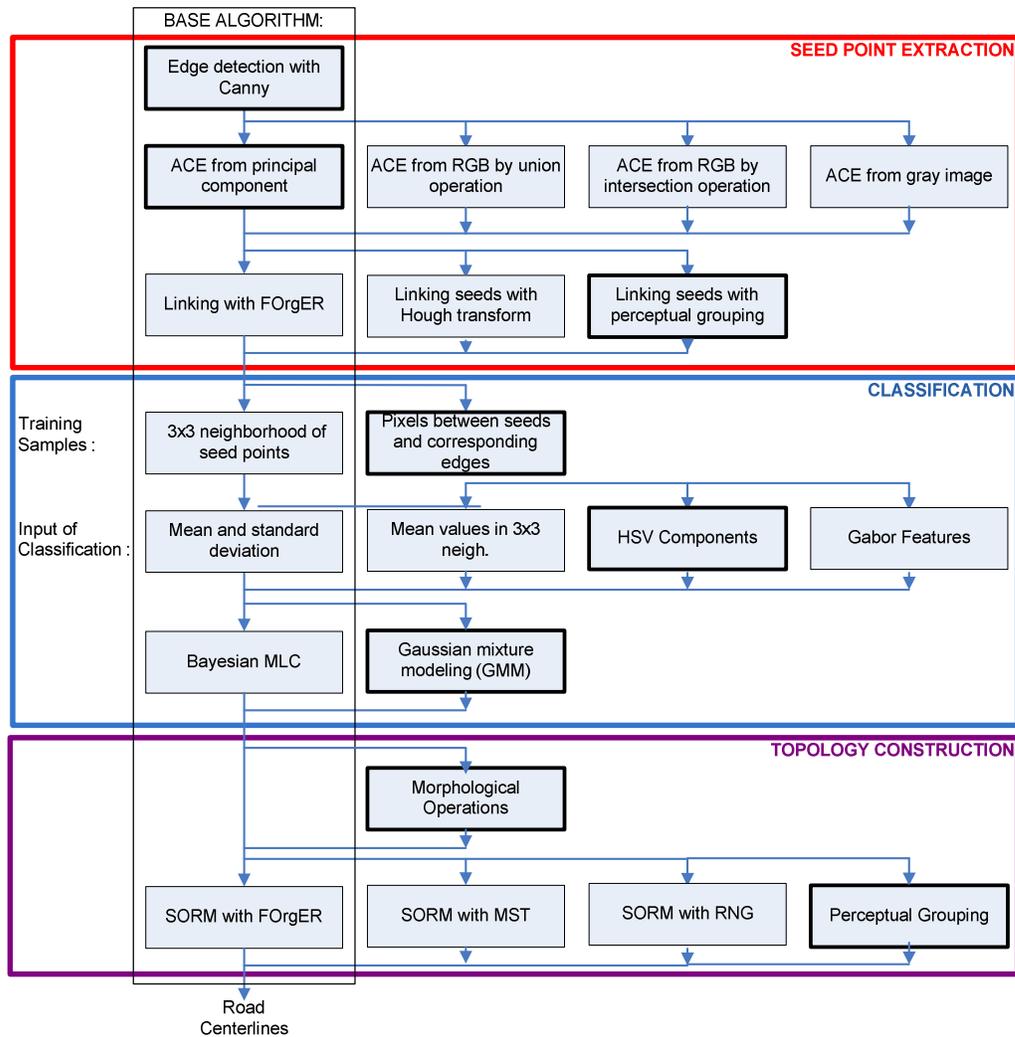


Figure 5-1 Base algorithm and alternative methods

### 5.1 The Performance Evaluation Metrics

The performance evaluation of the obtained results is essential for determining the reliability of the proposed algorithm. Heipke et al. [33] proposed a method of performance evaluation of automatic road extraction algorithms by comparing the extracted road pixels with the manually plotted linear road axis which is used as the reference data. This method consists of matching and calculation stages.

The matching stage is performed in two steps; that are matching the extracted data and matching the reference data. The buffering method is used in matching steps. The buffering means that constructing a constant predefined width around the data. In the first step, the buffer is constructed around the reference data and the region where the extracted data and the buffer of the reference data intersect is called as the matched region. In the second step, the buffer is constructed around the extracted data and the region where the reference data and the buffer of the extracted data intersect is called as the matched region.

With the matching stage, the following measuring elements are determined:

**True Positive (TP):** The number of matched pixels supplied by one of the matching steps.

**False Positive (FP):** The number of extracted pixels that don't belong to the buffered region around the reference data.

**False Negative (FN):** The number of the reference pixels that don't belong to the buffered region around the extracted data.

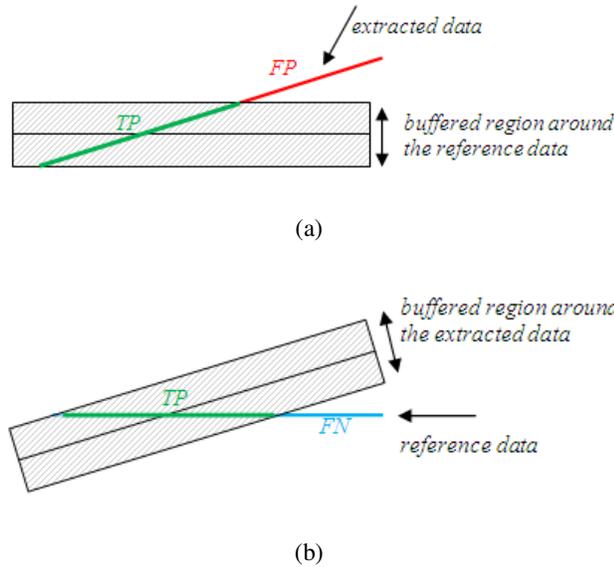


Figure 5-2 The matching stage of performance evaluation algorithm (a) Buffering around the reference data (b) Buffering around the extracted data

In the calculation stage, a number of quality measures are defined to interpret the matching results in an absolute way. The quality measures are based on two criteria: the correctness of the extracted road pixels and the completeness of the extracted road network. The completeness, correctness and the quality measures are defined in the following.

**Correctness:** The percentage of the correctly extracted road pixels, i.e, the ratio of the matched pixels to the extracted road pixels.

$$\text{Correctness} = \frac{TP}{TP+FP} \quad (5-1)$$

**Completeness:** The percentage of the reference pixels that overlap the extracted road pixels, i.e, the ratio of the matched pixels to the reference pixels.

$$\text{Completeness} = \frac{TP}{TP+FN} \quad (5-2)$$

**Quality:** The measure of the goodness of the road extraction algorithm, i.e, the ratio of the matched pixels to the union of the extracted and the reference road pixels.

$$\text{Quality} = \frac{TP}{TP+FP+FN} \quad (5-3)$$

## 5.2 Analysis of ACE Inputs

Finding anti-parallel centerline points is the first step of the road seed point extraction algorithm. The performance of this stage is quite important since all of the following stages of the proposed algorithm are based on the outputs of the ACE algorithm. There are four alternative methods of extracting anti-parallel centerline points: using gray level image formed by averaging RGB layers as the input of the ACE, conducting discrete layer gradient analysis (DLGA) by the intersection or the union operations, or using PCA (Principal Component Analysis) image as an input. Discrete layer gradient analysis (DLGA) means conducting the ACE algorithm for each layer of the RGB image and forming the final result by the intersection or the union of all layers.

The performance evaluation of the ACE methods is performed by comparing classification results obtained with each alternative. Figure 5-3, Figure 5-4, Figure 5-5 show the correctness, completeness and qualities of the results. The performances of alternative methods are close to each other, but the one that has highest quality measure is selected. Using principal component as the input of the ACE algorithm is preferred with respect to the Table 5-1.

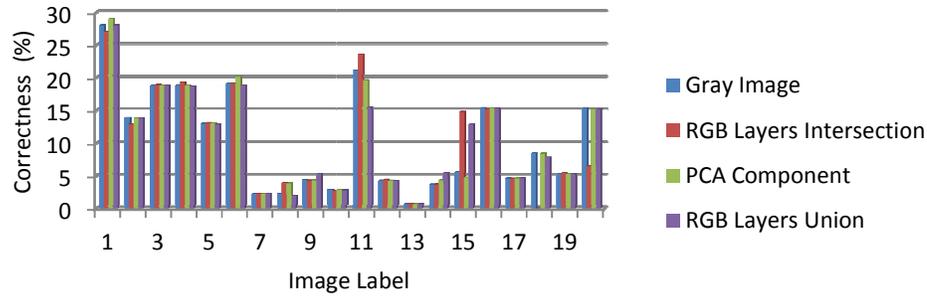


Figure 5-3 The Correctness of the ACE results

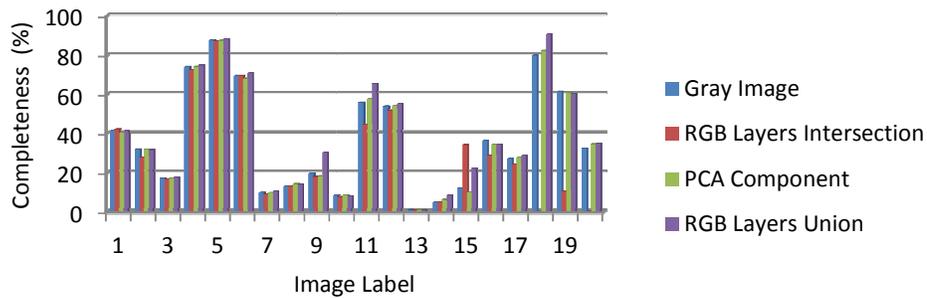


Figure 5-4 The completeness of the ACE results

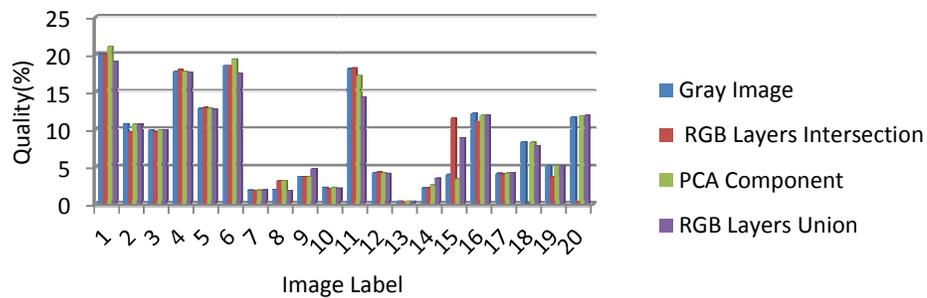


Figure 5-5 The quality of the ACE results

Table 5-1 The performance comparison of the alternative ACE methods

	Correctness	Completeness	Quality
<b>Gray Image</b>	10.38	36.71	8.44
<b>RGB Layers Intersection</b>	10.10	28.01	7.72
<b>PCA Image</b>	10.47	36.75	8.53
<b>RGB Layers Union</b>	10.05	39.22	8.45

The parameters of the ACE algorithm are determined based on the resolution of the input image. Since a single road layer has a width about 3 to 4 meters, the width of a single layer is 3-4 pixels in 1 m/pixel resolution image. By considering the edge tolerance and assuming that the number of the road lanes is between 1 and 3, the minimum road width is selected as 2 pixels and the maximum road width is selected as 15 pixels for 1 m/pixel resolution images.

To evaluate the parameter selection criteria of the ACE algorithm, the variation of the ACE result in terms of the road width parameters are examined in three 1 m/pixel resolution images. The images are selected as rural, urban and sub-urban. First, the minimum road width parameter is changed while the maximum road width parameter is hold as constant. Second, the maximum road width parameter is changed while the minimum road width parameter is hold as constant. Figure 5-6 illustrates the quality measures in terms of the road width parameters. When these graphs are examined, it is seen that the parameter selection criteria is fair enough.

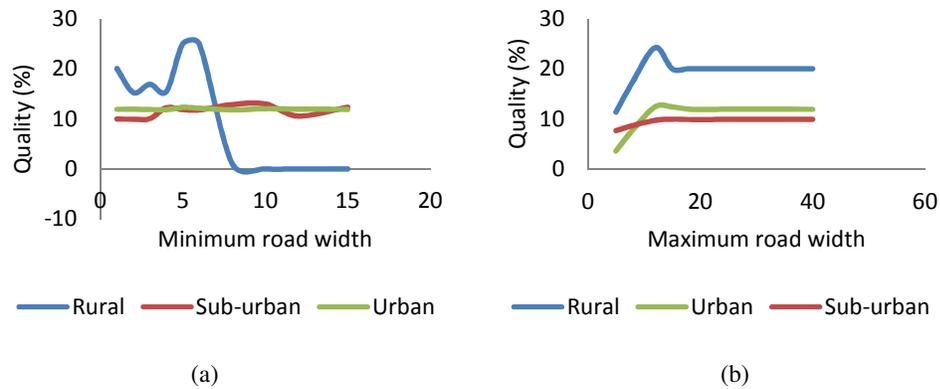


Figure 5-6 The variance of the ACE result in terms of road width parameters (a) In terms of the minimum road width (b) In terms of the maximum road width

### 5.3 Analysis of the Seed Points Linking Algorithms

The purpose of seed point linking is eliminating non-road ACE responses and increasing the number of seed points. Man-made objects generally have straight boundaries, so the edge detection algorithm finds these edges too. Since the roads are elongated long regions, by taking only the long ACE responses as seeds, the non-road responses can be eliminated.

As the seed point linking algorithm, we suggest two different methods such as Hough transform and perceptual grouping. The SORM algorithm with FORgER method which is proposed by Doucette at al. and the proposed algorithms are compared. As mentioned in section 4.1.3.2, to increase the number of samples the neighbors of the road centers lying between the corresponding edges are also added to the training set. "Perceptual Enlarged" represents the enlarged training set after perceptual grouping.

In terms of the quality measures, perceptual grouping is selected as the seed point linking algorithm.

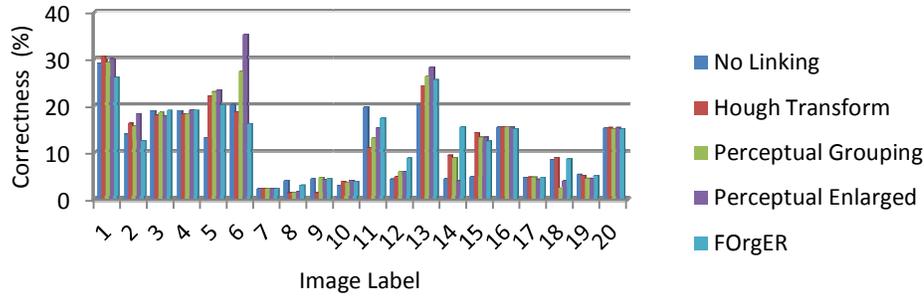


Figure 5-7 The correctness of the linking results

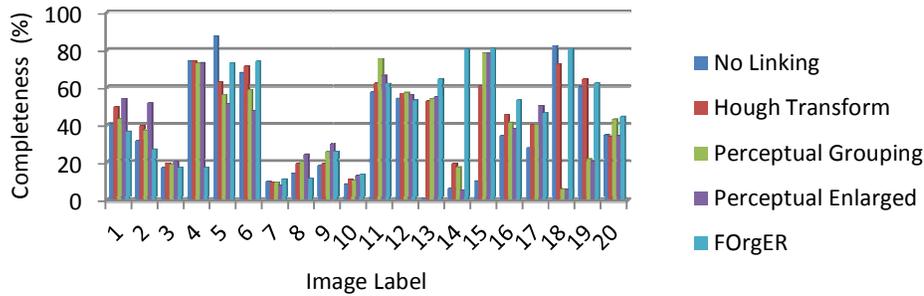


Figure 5-8 Completeness of the linking results

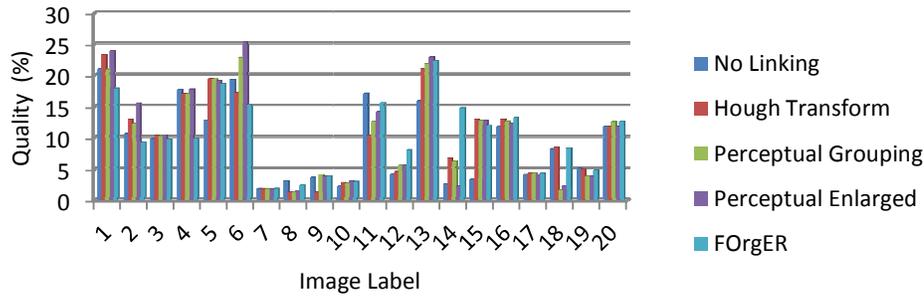


Figure 5-9 Quality of the linking results

Table 5-2 The performance comparison of the alternative linking methods

	Correctness	Completeness	Quality
<b>No linking</b>	9.42	30.23	7.63
<b>Hough Transform</b>	11.03	41.56	9.23
<b>Perceptual Grouping</b>	11.32	36.57	9.24
<b>Perceptual Enlarged</b>	11.85	36.26	9.56
<b>FOrgER</b>	10.47	36.75	8.53

The parameters of the perceptual grouping algorithm are the maximum angle between the line segments, the maximum length of the gap between the line segments, the minimum centerline length that is used for thresholding the grouping result.

The maximum length of the gap is selected as smaller than the average road width, in order to avoid bridging the ACE responses of the buildings that appear at each side of the roads. The maximum angle between the line segments is selected as  $30^\circ$ . The minimum centerline length is selected as greater than the average road width by considering the directional rectangularity feature of the roads. These parameters are evaluated with the images which are used for the parameter evaluation of the ACE algorithm. Figure 5-10 illustrates the quality measures of the results in terms of the grouping and linking parameters. By examining these graphs and considering the criteria mentioned above, the parameter selection is done as shown in Table 5-3. The SORM parameters are stated by Doucette et al. in [1].

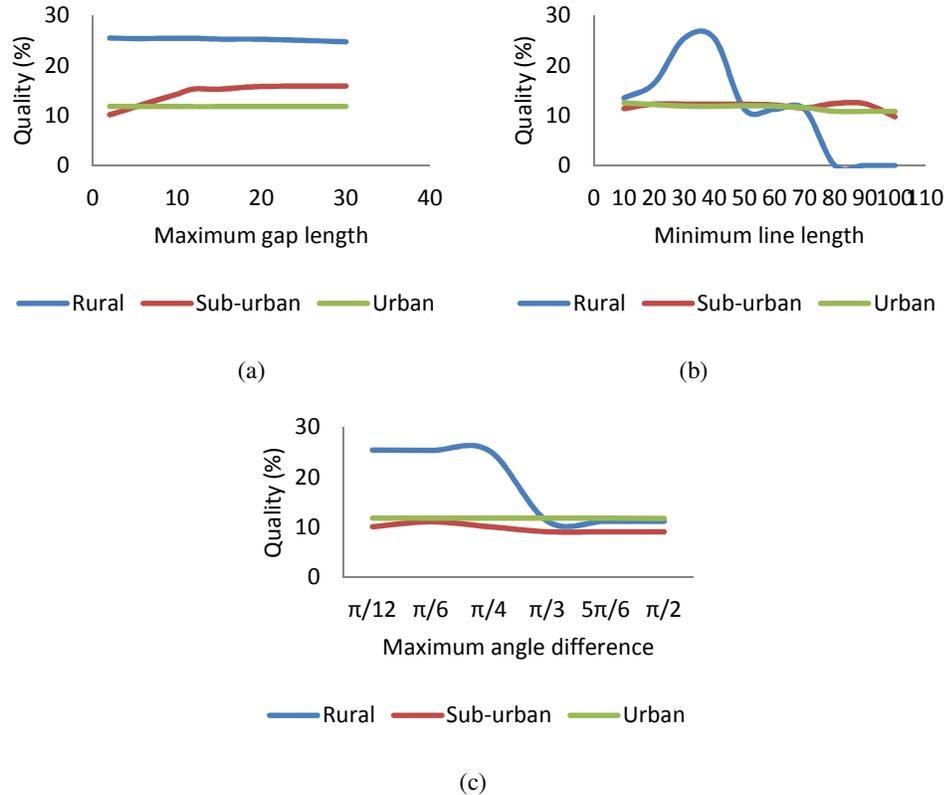


Figure 5-10 The quality variance of linking results in terms of parameters (a) In terms of the maximum gap length (b) In terms of the minimum line length (c) In terms of the maximum angle difference

Table 5-3 The parameters of the linking algorithm

	1 m/pixel
The maximum gap length	15
The maximum angle difference	$\pi/6$
The minimum line length	40

#### 5.4 Analysis of Classification Features

For selecting the best features for the road extraction, spectral and textural features are evaluated. In literature, some papers use only the mean filtered intensity values as a feature; some papers use features contain mean and variance values. In order to use color information, we experienced HSV features and in order to use textural features, we experienced Gabor filters.

Therefore, four different features are tested, such as mean of 3x3 neighborhood of gray image, mean and variance of 3x3 neighborhood of gray image, HSV components of RGB image, Gabor filters of RGB image. The intersection of classification results obtained by using HSV and Gabor filters is also compared with the other classification results. During the experiments, previously selected algorithms and parameters are used. As the classification algorithm, normal distribution is used.

As illustrated in the following figures and Table 5-4, using HSV color space shows best performance, therefore HSV color space is selected as the feature for classification.

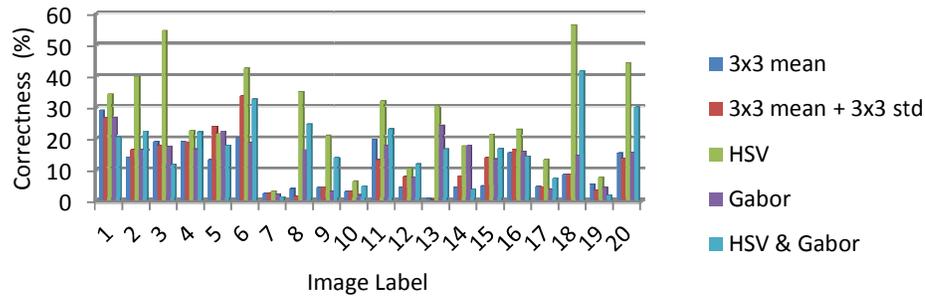


Figure 5-11 The correctness of the classification features

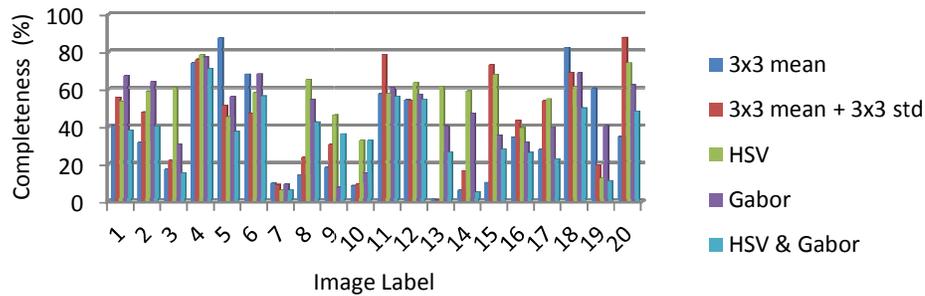


Figure 5-12 The completeness of the classification features

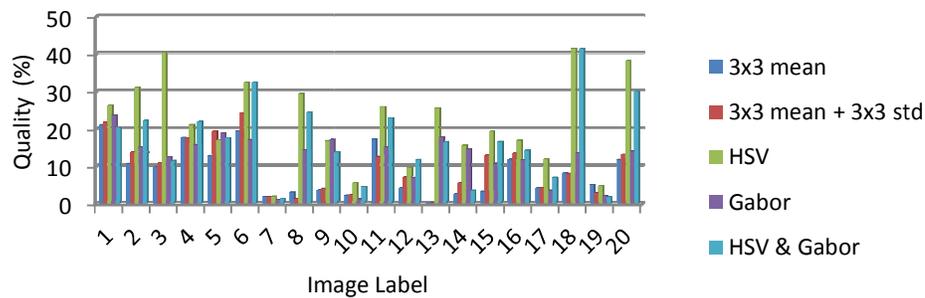


Figure 5-13 The quality of the classification features

Table 5-4 The performance comparison of the alternative features

	Correctness	Completeness	Quality
<b>Mean</b>	10.57	35.12	9.06
<b>Mean + Standard Deviation</b>	11.85	36.26	9.56
<b>HSV Color Space</b>	26.73	52.62	21.54
<b>Gabor Filter</b>	13.72	46.45	12.32
<b>HSV + Gabor</b>	24.89	34.93	16.80

### 5.5 Analysis of Classification Algorithms

Two different methods are evaluated in the classification step. First method is modeling the features with the normal distribution function for road and non-road class and using Bayesian decision rule to determine road class. Second method is modeling the road class with Gaussian mixture, and directly using this model for calculating the membership values of the road class.

The parameters of classification methods are the minimum membership values.

The minimum membership thresholds are selected based on the graphics shown in Figure 5-14 which are obtained from three different image types. The selected thresholds are: 0.5 for GMM, 0.7 for normal distribution.

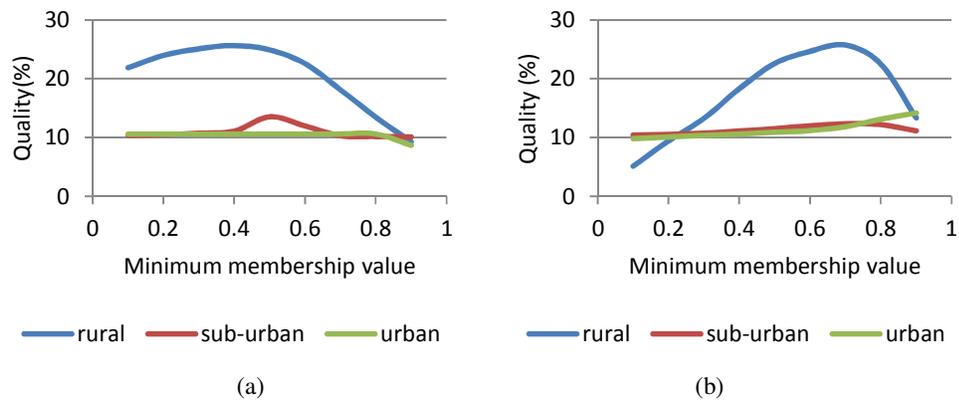


Figure 5-14 The quality variance in terms of minimum membership values (a) In GMM (b) In normal distribution

The comparison result of the classification methods are shown below. Based on these results GMM is selected as the classification method.

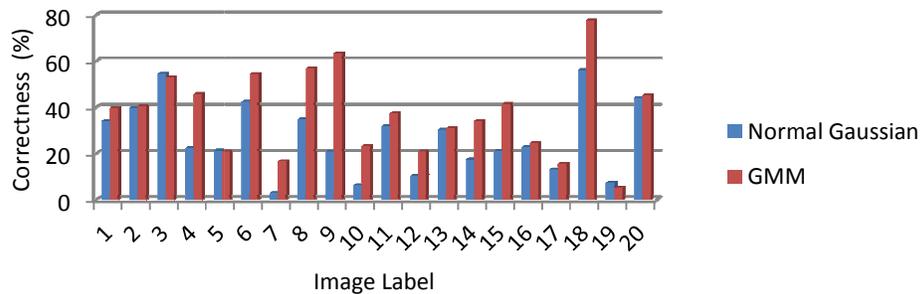


Figure 5-15 The correctness of the classification methods

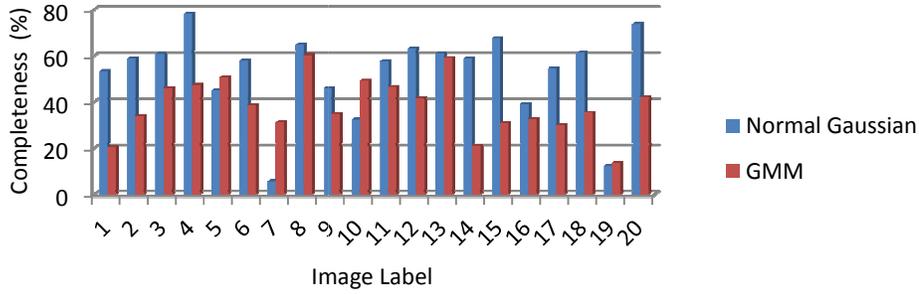


Figure 5-16 The completeness of the classification methods

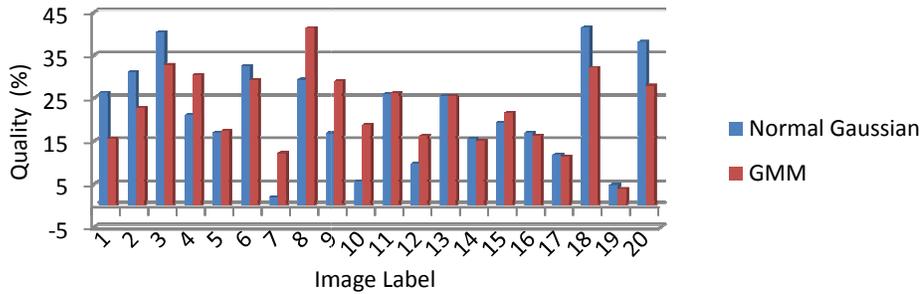


Figure 5-17 The quality of the classification methods

Table 5-5 The performance comparison of the alternative classification methods

	Correctness	Completeness	Quality
<b>Normal distribution</b>	26.73	52.62	21.54
<b>GMM</b>	37.37	38.31	22.26

### 5.6 Analysis of Smoothing Algorithms

In this step, the smoothing filter is examined if it improves the classification results or not. The parameters of bilateral filter are selected by testing this filter in three different images. Based on the graphics given in Figure 5-18, the variance of range filter is selected as 0.3 and the variance of domain filter is selected as 3.

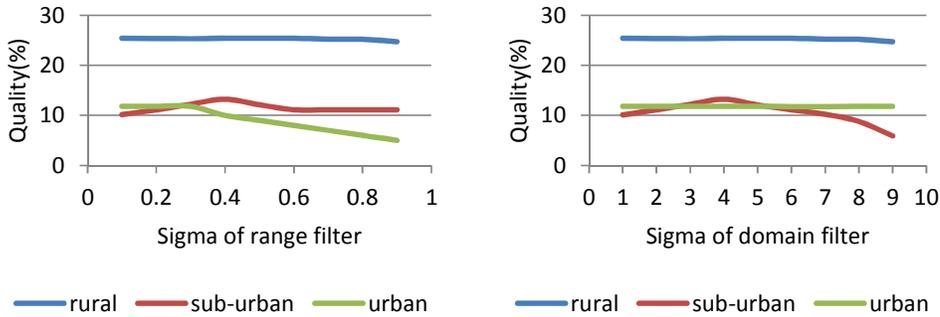


Figure 5-18 The quality variance in terms of the parameters of bilateral filter (a) In terms of range filter variance value (b) In terms of domain filter variance value

The following graphs show that bilateral filter give better result than no filtering. However, the computational complexity of filtering is high for big images and when the resulting images are observed, it is seen that the results of filtered image can be supplied by noise filtering and closing operations. Therefore, the smoothing filter is not included in the proposed algorithm.

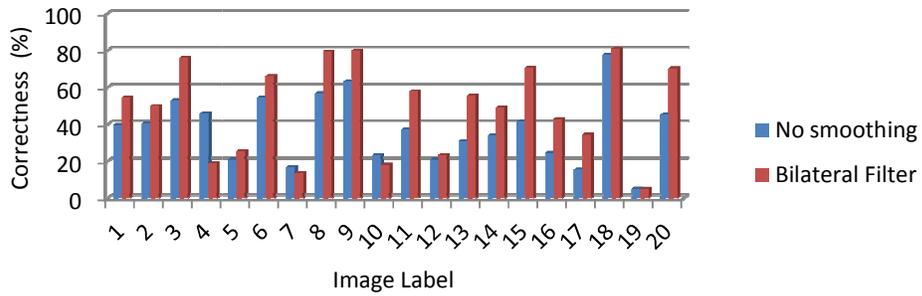


Figure 5-19 The correctness comparison for smoothing filter

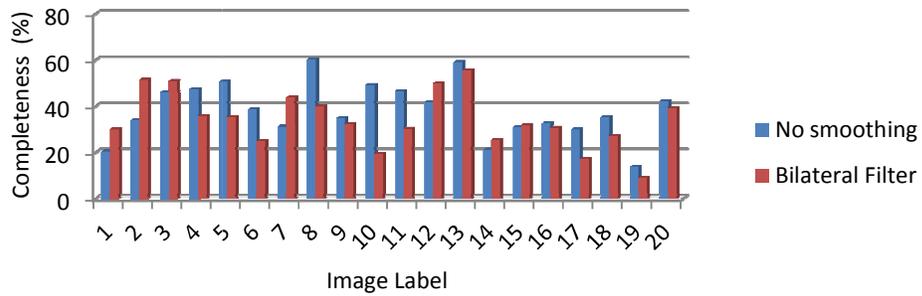


Figure 5-20 The completeness comparison for smoothing filter

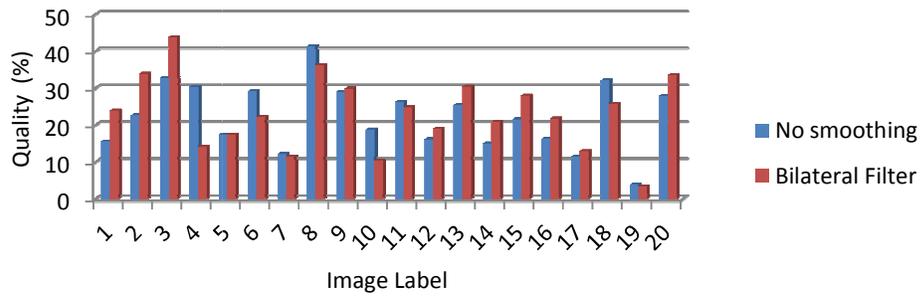


Figure 5-21 The quality comparison for smoothing filter

Table 5-6 The performance improvement with smoothing filter

	Correctness	Completeness	Quality
<b>No smoothing</b>	37.37	38.31	22.26
<b>Bilateral filter</b>	48.56	33.98	23.17

### 5.7 Analysis of Morphological Operations

Morphological operations are applied to filter noises in classification result and smoothing the road network. These operations are small area removal, closing and hole filling as mention in Section 4.3. The morphological operations improve the classification results remarkably.

First, morphological closing is performed and the connected components (CC) whose areas are smaller than a predefined area are removed. Then, the gaps on the image are filled. When the effect of morphological parameters to the final result is examined, it is seen that the area thresholds used for gap filling and connected component removing do not affect the result so much. However, the size of structuring element used for closing affects the quality. In terms of the result given Figure 5-22 (c), the structuring element is selected as 11x11square.

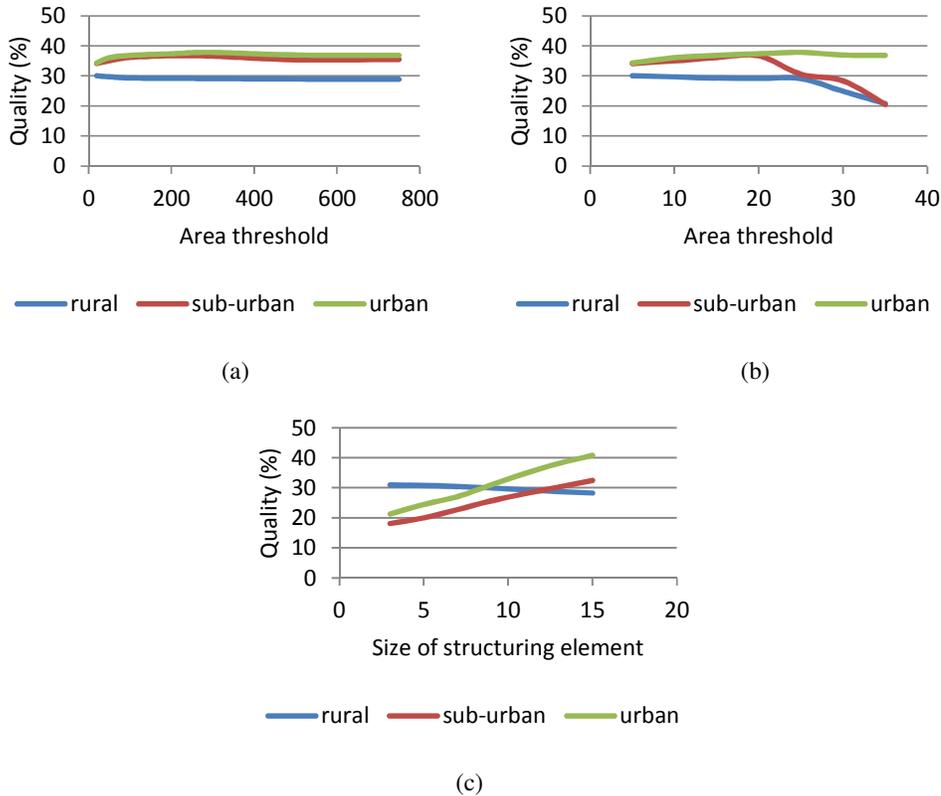


Figure 5-22 The quality variance in terms of parameters used in morphological operations (a) Area threshold used for gap filling (b) Area threshold used for CC removing (c) The size of structuring element used for closing

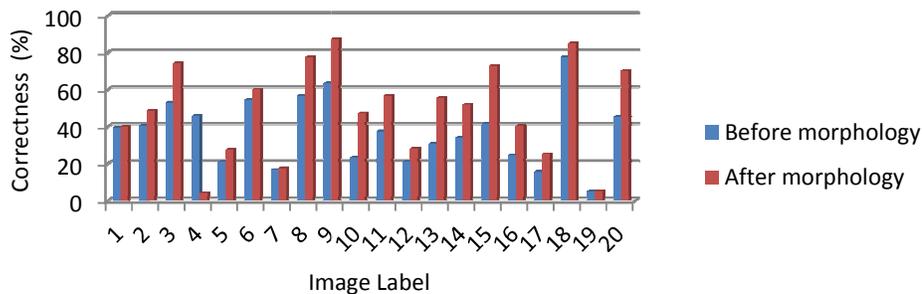


Figure 5-23 The correctness comparison for morphological operations

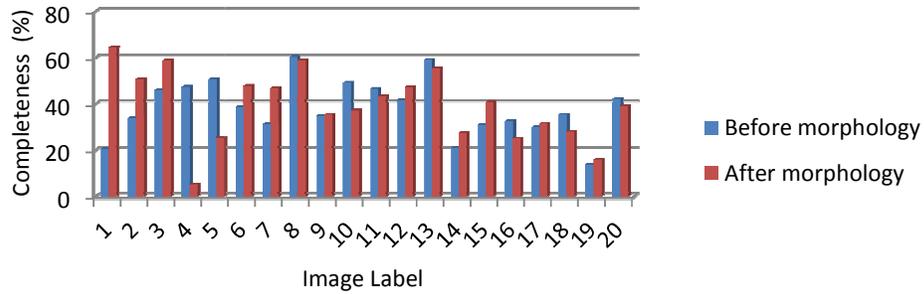


Figure 5-24 The completeness comparison for morphological operations

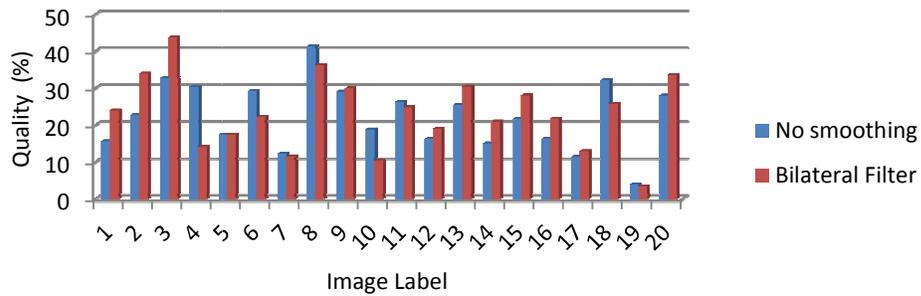


Figure 5-25 The quality comparison for morphological operations

Table 5-7 The performance improvement with morphological operations

	Correctness	Completeness	Quality
<b>Before morphology</b>	37.37	38.31	22.26
<b>After morphology</b>	48.77	39.22	27.09

### 5.8 Analysis of Topology Construction Algorithms

As introduced in Section 4.4, extracting road topology from road class image four different methods are evaluated, such as SORM with MST, RNG, FOrGER algorithms and perceptual grouping. For perceptual grouping, the parameters used in the seed point linking step are also used in topology construction step. For SORM algorithm, the parameters are determined by performing parameter tuning, as shown in Figure 5-26. The reason of not using parameters in section 5.3 is that the classification result includes road regions but the ACE response includes centerlines.

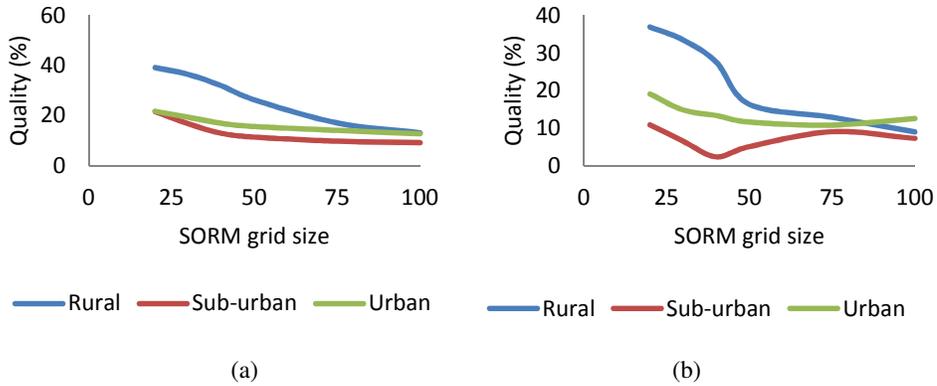


Figure 5-26 The grid size selection for SORM algorithm (a) MST (b) FOrGER

MST and RNG algorithms link the SORM centers in terms of the distance between centers. When the grid size of the SORM algorithm is selected as small, the linking result contains zigzags; when it is selected as a high value, different road regions can be included in a Voronoi region, besides false responses occur at curvatures and intersections. By using FOrGER algorithm, zigzags are avoided, but the orientation of the region cannot be detected when the grid size is small. Therefore, finding an optimum grid size for different images is very hard. The MST responses in terms of grid size are illustrated in Figure 5-27.

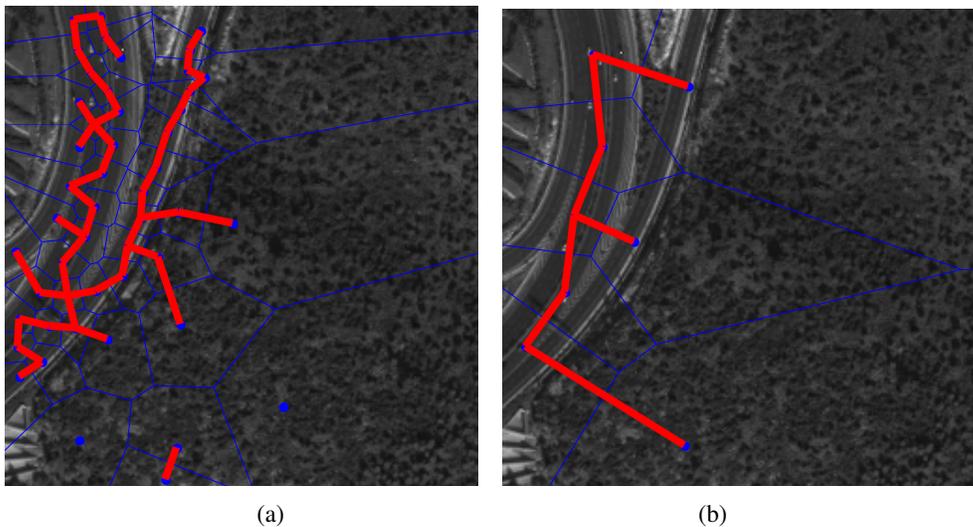


Figure 5-27 The linking result of SORM algorithm (a)  $\beta_{grid} = 20$  (b)  $\beta_{grid} = 60$

Before applying perceptual grouping, the edge pixels are removed from classification result and morphological thinning operation is applied to find the rough topology of the road network.

The experimental result of topology construction algorithms are given below. Based on these metrics, perceptual grouping is selected as the topology construction algorithm.

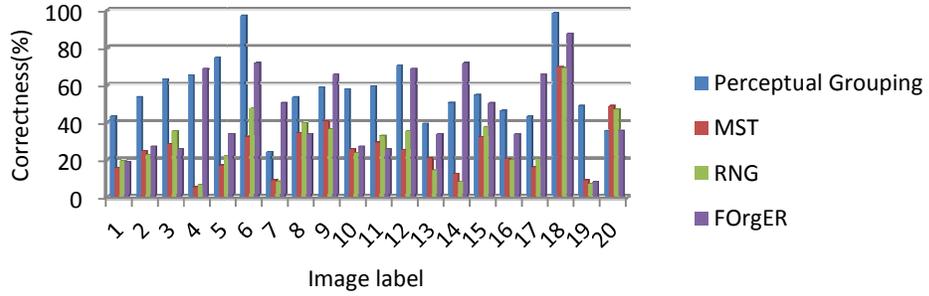


Figure 5-28 The correctness of the topology construction methods

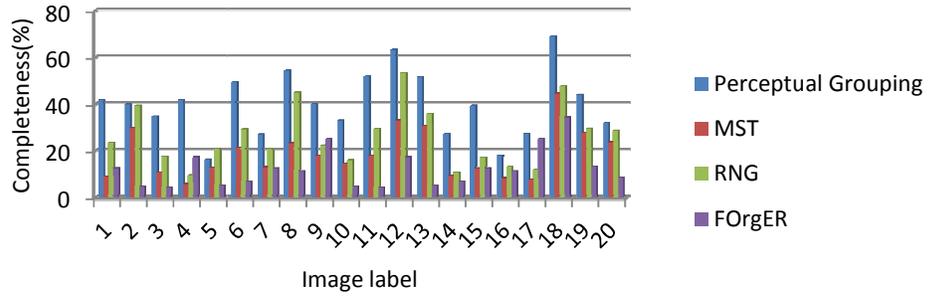


Figure 5-29 The completeness of the topology construction methods

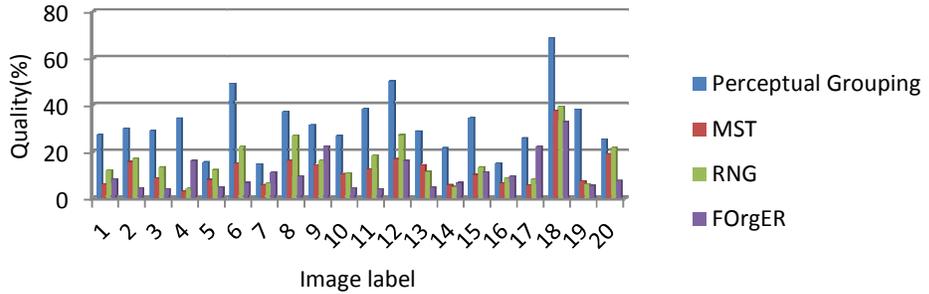


Figure 5-30 The quality of the topology construction methods

Table 5-8 The performance comparison of the topology construction methods

	Correctness	Completeness	Quality
<b>Perceptual Grouping</b>	57.13	40.00	32.29
<b>MST</b>	25.74	18.62	11.75
<b>RNG</b>	27.58	26.05	14.88
<b>FOrgER</b>	45.07	12.10	10.42

### 5.9 Comparison with the Baseline Algorithm

The analysis of alternative methods are conducted on baseline algorithm but using the optimum solutions for previous steps. Therefore, these analyses don't include the performance of baseline algorithm. In this part, the performance of baseline algorithm and the proposed algorithm is compared.

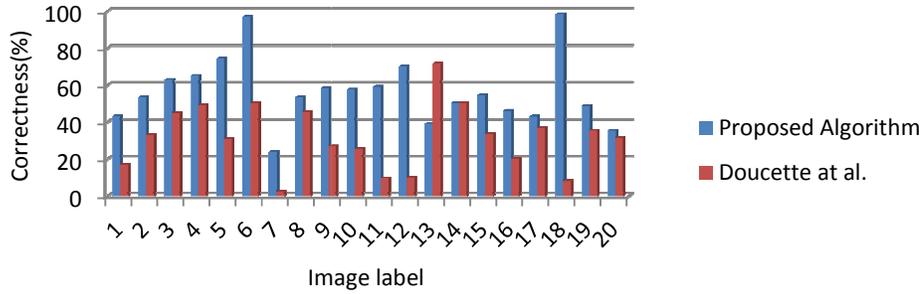


Figure 5-31 The correctness comparison

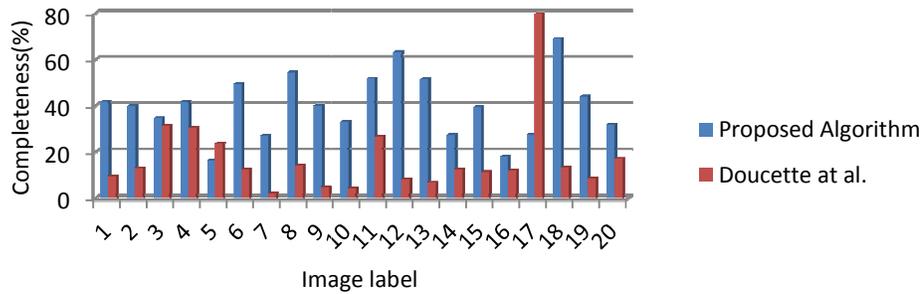


Figure 5-32 The completeness comparison

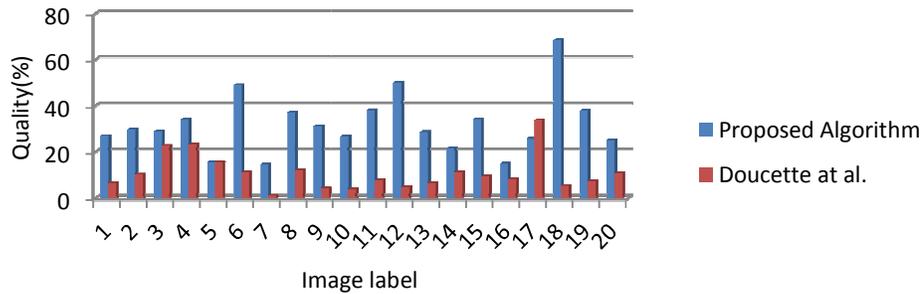


Figure 5-33 The quality comparison

Table 5-9 The performance comparison with [1]

	Correctness	Completeness	Quality
<b>Proposed Algorithm</b>	57.13	40.00	32.29
<b>Doucette at al.</b>	17.45	16.94	7.83

### 5.10 Comparison with Previous Studies

In this step, we compare the results of our algorithm with a recently proposed algorithm, (M. Dursun, 2012). In this study, the input image is segmented by mean-shift segmentation algorithm and the seed points are selected from these segments in terms of structural features. Classification is performed by modeling these samples with a modified GMM algorithm. Finally classification results are strength with structural verification and morphological operations.

The output of Dursun's proposed algorithm is the road network includes all detected road pixels, however our proposed algorithm extract the road centerlines of the input image. Therefore, we decided to compare the algorithms in two different levels. First, the classification results before post processing are compared. Then, the final results are compared by enlarging the centerlines with an average road width extracted from the input image. Since the database is common, the results are compared in terms of the average performance metrics. As shown in Table 5-10, the classification results of our proposed algorithm is better than Dursun's. It means that the performance of the seed point selection algorithm and the classification methods are better than Dursun's.

Table 5-10 The performance comparison of classification results

	Correctness		Completeness		Quality	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
<b>Dursun's method</b>	19.69	12.81	75.92	6.60	18.28	11.45
<b>Our proposed method</b>	37.37	18.17	38.31	12.33	22.26	9.03

Dursun improves his classification results with some post processing operations. Our post processing operations are based on the detection of the centerlines and eliminating false responses by perceptual grouping and linking. In order to compare the algorithms on whole, the results of our proposed algorithm are enlarged with an average road width and these enlarged images are used to calculate the performance metrics. As shown in Table 5-11, the results are similar with the algorithm proposed by Dursun [37].

Table 5-11 The performance comparison of final results

	Correctness		Completeness		Quality	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
<b>Dursun's method</b>	55.82	14.77	50.74	19.04	34.69	11.32
<b>Our proposed method</b>	58.35	17.93	47.79	13.16	33.51	12.39



## CHAPTER 6

### CONCLUSION

#### 6.1 Summary and Conclusions

In this thesis, an unsupervised approach of extracting centerlines of the road network has been presented. Our approach includes automatic seeding, classification and topology construction steps.

The ACE algorithm detects the possible seed points from the edge response of the input image based on gradient analysis. With the perceptual grouping approach, non-road seeds are eliminated and the number of seed points is extended. In the classification step, by using the Gaussian mixture model different types of road regions are determined and road class image is constructed. Topology construction extracts the road centerlines by forming the skeleton of the road network, eliminating non-road responses and linking the gaps on the network.

The perceptual grouping generally used for grouping linear features in road extraction and this makes the concept being proper only for the grid structures. However, by defining the co-linearity principle of the grouping algorithm around the nearest end-points of the road segments, it is shown that the perceptual grouping is also applicable for non-linear road segments.

As mentioned previously, the classification can be performed pixel-based and object based. In the field of remote sensing, object based classification is advantageous because it provides feature variety and gives the chance of examining pixels as a group. However, the classification method used in this thesis (modeling the features) is not appropriate for object-based classification. Therefore, pixel-based classification is preferred based on modeling the features of the roads. Gaussian modeling is preferred in most of the papers in the literature. Initially, we try to construct the road and non-road class with Gaussian model and by applying Bayesian decision rule to develop the road class. Various tests showed that Gaussian model is actually not sufficient for this purpose when different types of road structures are present in the image and therefore instead of plane Gaussian model we decided to perform Gaussian mixture modeling.

For topology construction, SORM algorithm is a good choice since it doesn't require morphological operations and isn't affected very much by the noise. Although this algorithm has advantages, it also creates false responses near the intersections of roads and doesn't perform as expected when the road regions are close to each other. The SORM algorithm is compared with the proposed topology construction method, which is the combination of morphological thinning and perceptual grouping algorithms for many different images and the results showed that the proposed topology construction method has a good performance.

Every step of the proposed algorithm depends on the certain parameters as an input values. In order to perform a fully automatic road extraction algorithm, these parameters also need to be determined automatically. By taking into consideration that the input dataset used for road extraction have a constant scale it is trivial that the most of the parameters are obtained in terms of this scale. The parameters which are not related to the scale of given images are tuned for the test data set and the most appropriate parameter set is selected.

We have showed at Chapter 4 and 5 that the proposed algorithm is capable of extracting road networks from different satellite images as long as the seed points are detected correctly.

#### 6.2 Future Work

The first step of the proposed algorithm is the automatic seeding and the small errors it this step usually causes to huge errors in the proceeding steps.

Therefore new criteria can be added to this step in order to improve the performance of the seed point extraction algorithm. In this thesis, we have used gray level image for obtaining the road seeds although RGB images can be employed as well.

Different algorithms together with the RGB images might show superior performance than proposed one for finding seed points.

In the perceptual grouping step, proximity and co-linearity principles are used for obtaining the line segments belong to the same group. Other Gestalt principles such as similarity, smoothness, closure can be used as well. In literature, some of the studies define weight to each criterion and form a segment connectivity matrix to select segments to be grouped. This type of grouping is stricter than the proposed one and may be advantageous.

The morphological closing operation is performed before the thinning operation in order to smooth the boundaries of road class image and fill small holes. Instead of the square structuring element used for closing operation, a structuring element formed in terms of the direction of road can be developed. With this adaptive structuring element, smoothing the road boundaries can be performed better than now and the road centerlines could be straighter.

In [21], a road tree pruning algorithm is proposed and this algorithm shows significant improvement at the quality of the results. In order to apply this type of pruning algorithm, a segment based approach can be developed in future. After obtaining the segments in the image, the road class image and the segments can be combined to detect road segments. The compactness of the segments can be used for eliminating non-road responses and a road tree can be constructed for bridging gaps.

## REFERENCES

- [1] P. Doucette, P. Agouris, and A. Stefanidis, "Automated Road Extraction from High Resolution Multispectral Imagery", *Photogrammetric Engineering & Remote Sensing* vol. 70, no. 12, pp. 1405–1416, December 2004.
- [2] R.O. Duda, P.E. Hart, "Use of the Hough Transformation to Detect Lines And Curves in Pictures", *Comm. ACM*, vol. 15, no. 1 pp. 11-15, January 1972
- [3] J.F. Canny, "A Computational Approach to Edge Detection", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679-698, 1986.
- [4] X. Hu, C.V. Tao, "Automatic Extraction of Main Road Centerlines from High Resolution Satellite Imagery Using Hierarchical Grouping", *Photogrammetric Engineering & Remote Sensing* vol. 73, no. 9, pp. 1049–1056, September 2007.
- [5] I. Rock, S. Pamler, "The Legacy of Gestalt Psychology", *Science American*, pp. 84–90, December 1990.
- [6] C. Steger, "An Unbiased Detector of Curvilinear Structures", Technical Report, Technical University of Munich, 1996.
- [7] U. Bacher, H. Mayer, "Fuzzy-Classification And Ziplock Snakes For Road Extraction From Ikonos Images", *ISPRS / EARSeL Joint Workshop on High Resolution Mapping from Space*, 2003
- [8] P. Doucette, P. Agouris, A. Stefanidis, M. Musavi "Self-organised clustering for road extraction in classified imagery" *ISPRS Journal of Photogrammetry & Remote Sensing* 55, 2001
- [9] T. Kohonen, "Self-Organizing Maps", 2nd edn. Springer-Verlag, 1997
- [10] J. McQueen, "Some methods for classification and analysis of multivariate observations", *Proc. Fifth Berkeley Symposium on Math. Stat. and Prob.*, vol. 1, pp. 281–296, 1967.
- [11] J. B. Mena, "State of the art on automatic road extraction for GIS update: A novel classification", *Pattern Recognition Letters* 24, pp. 3037-3058, 2003
- [12] D. M. McKeown, J. L. Denlinger, "Cooperative methods for road tracking in aerial imagery," *Proc. CVPR*, pp. 662–672, 1988
- [13] A. C. Hauptfleisch, "Automatic Road Network Extraction from High Resolution Satellite Imagery using Spectral Classification Methods", Master Thesis in University of Pretoria, 2010
- [14] P. Agouris, S. Gyftakis, A. Stefanidis, "Dynamic node distribution in adaptive snakes for road extraction", In *Proceedings of Vision Interface*, 2001.
- [15] A. Baumgartner, C. Steger, H. Mayer, W. Eckstein, H. Ebner, "Automatic Road Extraction In Rural Areas", *International Archives of the Photogrammetry. Remote Sensing*, pp. 107–112, 1999
- [16] A. P. Dal Poz, G. M. do Vale, "Dynamic Programming Approach For Semi-Automated Road Extraction From Medium- And High-Resolution Images", *ISPRS Archives*, Vol. XXXIV, Part 3/W8, 2003
- [17] J.B. Mena, J.A. Malpica, "An automatic method for road extraction in rural and semi-urban areas starting from high resolution satellite imagery", *Pattern Recognition Letters* 26, pp. 1201–1220, 2005
- [18] M. Song, D. Civco, "Road Extraction Using SVM and Image Segmentation" *Photogramm. Eng. Remote Sens.*, vol. 70, no. 12, pp. 1365–1371, 2004
- [19] U. Bacher, H. Mayer, "Automatic Road Extraction from Multi-spectral High Resolution Satellite Images" *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 36, 3/W24, pp. 29–34, 2005

- [20] M. Mokhtarzade, M. J. Valadan Zoej, "Road Detection From High Resolution Satellite Images Using Artificial Neural Networks" *International Journal of Applied Earth Observation and Geoinformation* 9(1): 32-40, 2007
- [21] J. Hu, A. Razdan, J. C. Femiani, M. Cui, P. Wonka, "Road Network Extraction and Intersection Detection From Aerial Images by Tracking Road Footprints", *IEEE Transactions On Geosciences And Remote Sensing*, vol. 45, no. 12, 2007
- [22] J. Zhao, S. You, J. Huang, "Rapid Extraction and Updating of Road Network from Airborne LiDAR Data", *IEEE Applied Imagery Pattern Recognition Workshop (AIPR)*, pp. 1-7, 2011
- [23] X. Hu, Z. Zhang, C. V. Tao, "A Robust Method for Semi-Automatic Extraction of Road Centerlines Using a Piecewise Parabolic Model and Least Square Template Matching", *Photogrammetric Engineering & Remote Sensing*, vol. 70, no. 12, pp. 1393-1398, 2004
- [24] B. Joseph, Jr. Kruskal, "On The Shortest Spanning Subtree of a Graph and the Traveling Salesman Problem", *Proceedings of the American Mathematical Society*, vol. 7, no. 1, 1956
- [25] R.C. Prim, "Shortest Connection Network and Some Generalization", *The Bell System Technical Journal*, 1957
- [26] G. Koutaki, K. Uchimura, "Automatic road extraction based on cross detection in suburb", *Proceedings of The International Society for Optical Engineering*, 2004
- [27] A. Gruen, H. Li, "Semi-automatic linear feature extraction by dynamic programming and LSB-snakes", *Photogrammetric Engineering and Remote Sensing*, 63(8), pp. 985-995, 1997.
- [28] R. Maurya, P.R. Gupta, A. S. Shukla, "Road Extraction Using K-Means Clustering and Morphological Operations", *International Conference on Image Information Processing (ICIIP 2011)*, 2011
- [29] V.S.N. Prasad, J. Domke, "Gabor Filter Visualization", *Technical Report, University of Maryland*, 2005
- [30] K. J. Supowit, "The Relative Neighborhood Graph, with an Application to Minimum Spanning Trees", *Journal of the Association for Computing Machinery*, vol 30. no 3, pp 428-448, 1983.
- [31] M. Heath, S. Sarkar, T. Sanocki, K. Bowyer, "Comparison of Edge Detectors, A Methodology and Initial Study", *Computer Vision and Image Understanding*, vol. 69, no. 1, pp. 38-54, 1998
- [32] J. Zhang, T. Tan, L. Ma, "Invariant Texture Segmentation Via Circular Gabor Filters" *Proceedings of the 16th IAPR International Conference on Pattern Recognition (ICPR)*, vol II, pp. 901-904, 2002.
- [33] C. Heipke, H. Mayer, C. Wiedemann, "Evaluation Of Automatic Road Extraction", *Proceedings of the ISPRS Conference*, 1997
- [34] D. Michie, D. J. Spiegelhalter, C.C. Taylor, "Machine Learning, Neural and Statistical Classification", 1994
- [35] J.A. Benediktsson, P.H. Swain, O.K. Ersoy, "Neural Network Approaches Versus Statistical Methods in Classification of Multisource remote Sensing data", *IEEE Transaction on Geosciences and Remote Sensing*, vol. 28: pp. 540-551, 1990
- [36] C. Tomasi, R. Manduchi, "Bilateral Filtering for Gray and Color Images", *IEEE International Conference on Computer Vision*, 1998
- [37] M. Dursun, "Road Detection By Mean Shift Segmentation And Structural Analysis", *METU*, 2012

## APPENDICIES

### APPENDIX A: EXAMPLE RESULTS ON GROUND TRUTH



Figure A-1 The result of test image 1 (green : TP, blue: FN, red : FP)

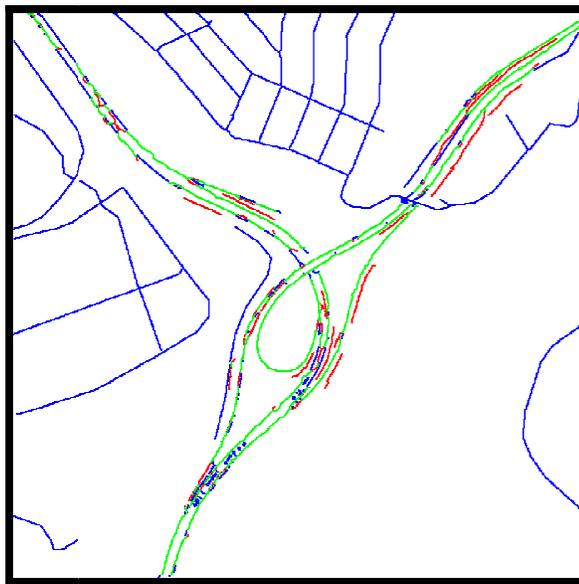


Figure A-2 The result of test image 3 (green : TP, blue: FN, red : FP)

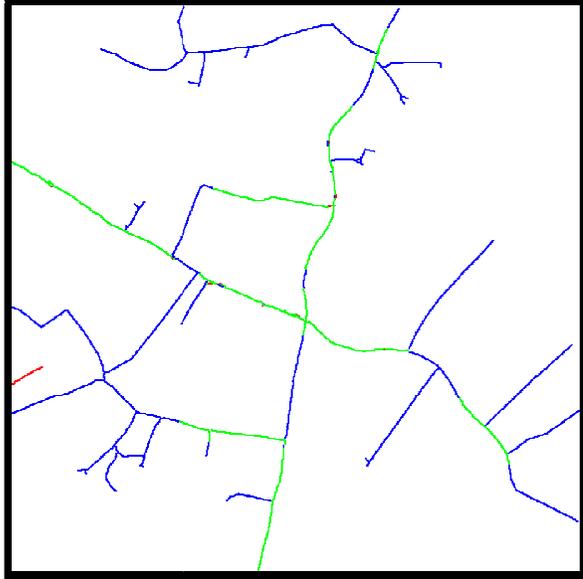


Figure A-3 The result of test image 4 (green : TP, blue: FN, red : FP)

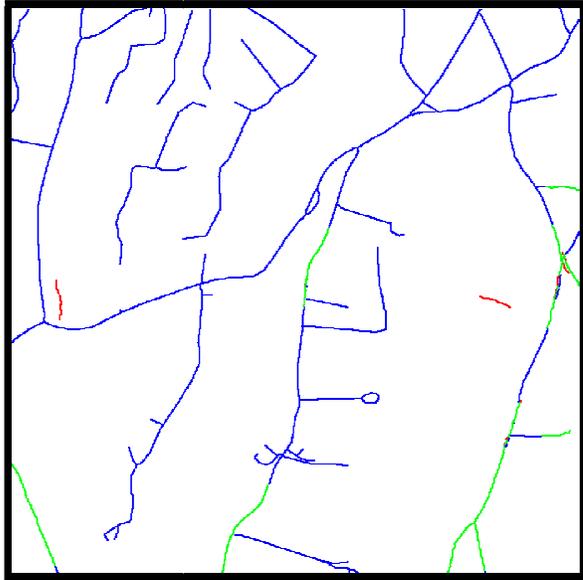


Figure A-4 The result of test image 5 (green : TP, blue: FN, red : FP)

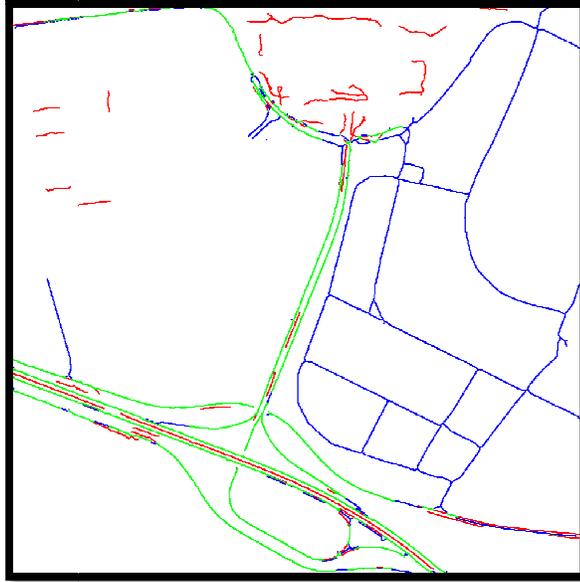


Figure A-5 The result of test image 6 (green : TP, blue: FN, red : FP)

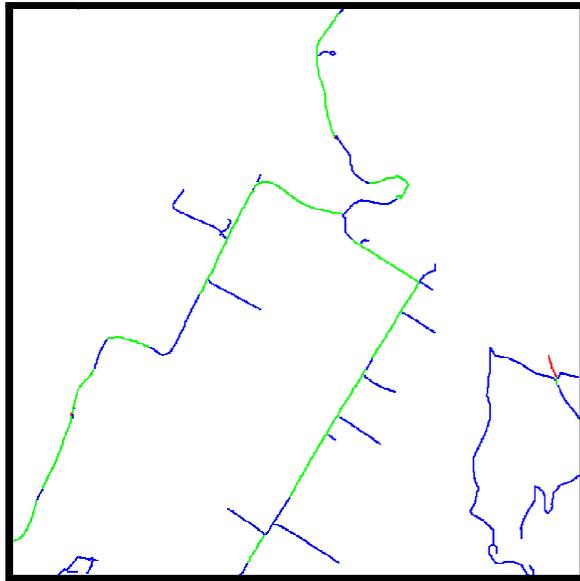


Figure A-6 The result of test image 12 (green : TP, blue: FN, red : FP)