

4-29-2015

# Automatic Extraction of Number of Lanes from Aerial Images for Transportation Applications

LI TANG

Florida International University, ltang001@fiu.edu

**DOI:** 10.25148/etd.FIDC000099

Follow this and additional works at: <https://digitalcommons.fiu.edu/etd>



Part of the [Civil Engineering Commons](#)

---

## Recommended Citation

TANG, LI, "Automatic Extraction of Number of Lanes from Aerial Images for Transportation Applications" (2015). *FIU Electronic Theses and Dissertations*. 2200.

<https://digitalcommons.fiu.edu/etd/2200>

This work is brought to you for free and open access by the University Graduate School at FIU Digital Commons. It has been accepted for inclusion in FIU Electronic Theses and Dissertations by an authorized administrator of FIU Digital Commons. For more information, please contact [dcc@fiu.edu](mailto:dcc@fiu.edu).

FLORIDA INTERNATIONAL UNIVERSITY

Miami, Florida

AUTOMATIC EXTRACTION OF NUMBER OF LANES FROM AERIAL IMAGES  
FOR TRANSPORTATION APPLICATIONS

A dissertation submitted in partial fulfillment of

the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

CIVIL ENGINEERING

by

Li Tang

2015

To: Dean Amir Mirmiran  
College of Engineering and Computing

This dissertation, written by Li Tang, and entitled Automatic Extraction of Number of Lanes from Aerial Images for Transportation Applications, having been approved in respect to style and intellectual content, is referred to you for judgment.

We have read this dissertation and recommend that it be approved.

---

L. David Shen

---

Mohammed Hadi

---

Zhenmin Chen

---

Tao Li

---

Priyanka Alluri

---

Albert Gan, Major Professor

Date of Defense: April 29, 2015

The dissertation of Li Tang is approved.

---

Dean Amir Mirmiran  
College of Engineering and Computing

---

Dean Lakshmi N. Reddi  
University Graduate School

Florida International University, 2015

© Copyright 2015 by Li Tang

All rights reserved.

## DEDICATION

I dedicate this dissertation to my parents, Huilin Tang and Xinmin Qi, and to my wife, Jingyu Huang, for their love and support.

## ACKNOWLEDGMENTS

I would like to take this opportunity to acknowledge everyone who helped me complete my doctoral degree in Civil Engineering. First and foremost, I would like to express my deepest appreciation to my major advisor, Dr. Albert Gan. Without his inspiration, guidance, and support, this dissertation would not have come to completion. Dr. Gan taught me how to think critically about research problems. He also taught me valuable lessons on professional attitude. His invaluable advice and suggestions helped me greatly during my doctoral studies, and will certainly benefit my future professional and personal life

I would also like to thank Dr. L. David Shen, Dr. Mohammed Hadi, Dr. Zhenmin Chen, Dr. Tao Li, and Dr. Priyanka Alluri, for serving on my dissertation committee. I sincerely appreciate the time they took to read my dissertation, and for providing comments and suggestions to improve it. I would like to give special thanks to Dr. Priyanka Alluri who provided precious advice and corrections for two papers I published.

I am very grateful to the following colleagues for their help throughout the years of my doctoral studies: Dr. Kaiyu Liu, Dr. Tao Wang, Dr. Jinyan Lu, Dr. Wanyang Wu, Dr. Dibakar Saha, and Dr. Feng Gui. I would also like to thank Ms. Natalie Defraene for her editing of this dissertation.

Finally, I would like to thank my parents and my wife for their support and love. Without their patience and understanding, the completion of this dissertation would not have been possible.

ABSTRACT OF THE DISSERTATION

AUTOMATIC EXTRACTION OF NUMBER OF LANES FROM AERIAL IMAGES

FOR TRANSPORTATION APPLICATIONS

by

Li Tang

Florida International University, 2015

Miami, Florida

Professor Albert Gan, Major Professor

Number of lanes is a basic roadway attribute that is widely used in many transportation applications. Traditionally, number of lanes is collected and updated through field surveys, which is expensive especially for large coverage areas with a high volume of road segments. One alternative is through manual data extraction from high-resolution aerial images. However, this is feasible only for smaller areas. For large areas that may involve tens of thousands of aerial images and millions of road segments, an automatic extraction is a more feasible approach. This dissertation aims to improve the existing process of extracting number of lanes from aerial images automatically by making improvements in three specific areas: (1) performance of lane model, (2) automatic acquisition of external knowledge, and (3) automatic lane location identification and reliability estimation.

In this dissertation, a framework was developed to automatically recognize and extract number of lanes from geo-rectified aerial images. In order to address the external knowledge acquisition problem in this framework, a mapping technique was developed to automatically estimate the approximate pixel locations of road segments and the travel

direction of the target roads in aerial images. A lane model was developed based on the typical appearance features of travel lanes in color aerial images. It provides more resistance to “noise” such as presence of vehicle occlusions and sidewalks. Multi-class classification test results based on the K-nearest neighbor, logistic regression, and Support Vector Machine (SVM) classification algorithms showed that the new model provides a high level of prediction accuracy.

Two optimization algorithms based on fixed and flexible lane widths, respectively, were then developed to extract number of lanes from the lane model output. The flexible lane-width approach was recommended because it solved the problems of error-tolerant pixel mapping and reliability estimation. The approach was tested using a lane model with two SVM classifiers, i.e., the Polynomial kernel and the Radial Basis Function (RBF) kernel. The results showed that the framework yielded good performance in a general test scenario with mixed types of road segments and another test scenario with heavy plant occlusions.



## TABLE OF CONTENTS

CHAPTER	PAGE
1. INTRODUCTION.....	1
1.1 Research Background .....	1
1.2 Problem Statement .....	2
1.3 Goals and Objectives .....	7
1.4 Dissertation Organization .....	8
2. LITERATURE REVIEW.....	10
2.1 Introduction.....	10
2.2 Road Models .....	11
2.3 Vehicle Models.....	15
2.4 Representations of Plants .....	19
2.5 Building Detection .....	22
2.6 Shadow Detection and Removal.....	24
2.7 Automatic Road Segment Recognition .....	27
2.8 Summary .....	29
3. METHODOLOGY.....	32
3.1 Introduction.....	32
3.2 Methodology Framework.....	32
3.3 Identification of Road Locations in Geo-Rectified Aerial Images .....	37
3.4 Lane-Profile Model and Profile Extraction from Aerial Images .....	40
3.5 Estimation of Number of Lanes .....	46
3.6 Summary .....	47
4. PROFILE-BASED LANE MODEL .....	49
4.1 Introduction.....	49
4.2 Similarity-Score Based Approach.....	49
4.3 Classification Based Approach .....	53
4.4 Support Vector Machine (SVM) .....	57
4.5 General Two-Class Classification Based on Profile Lane Model.....	59
4.6 Summary .....	63
5. ESTIMATION OF NUMBER OF LANES .....	65
5.1 Introduction.....	65
5.2 Saturation and Vegetation Index Thresholding.....	66
5.3 Fixed lane-length Optimization Algorithm for Lane Location Identification .....	67
5.4 Implementation of Fixed-Width Optimization and Evaluation .....	71
5.5 Flexible lane-Width Optimization Algorithm .....	76
5.5.1 Local Maximum Filter .....	79
5.5.2 Optimization Process and Objectives .....	81
5.5.3 Inner Optimization Loop.....	84
5.5.4 Outer Optimization Loop.....	90

5.6 Summary .....	94
6. EVALUATION OF FRAMEWORK.....	96
6.1 Introduction.....	96
6.2 Data Preparation, Lane Model Training, and Implementation .....	96
6.3 Review of Test Road Network .....	103
6.4 Summary .....	109
7. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS.....	111
7.1 Summary and Conclusions .....	111
7.2 Research Contribution .....	114
7.3 Recommendations for Future Research .....	115
REFERENCES .....	117
VITA .....	125

## LIST OF TABLES

TABLE	PAGE
Table 4-1 Summary of Training Data .....	56
Table 4-2 Summary of Classification Results.....	56
Table 4-3 Summary of SVM Classification Results for Feasibility Test.....	61
Table 4-4 Detailed Accuracy by Class .....	62
Table 5-1 Optimization Test Results .....	75
Table 6-1 Summary of SVM Classification Results for Test Scenarios .....	100
Table 6-2 Detailed Accuracy of Classification Results.....	101
Table 6-3 Parameter Setting of Test Program .....	102
Table 6-4 Summary of Framework Performance in Test Scenario One .....	105
Table 6-5 Summary of Framework Performance in Test Scenario Two .....	105

## LIST OF FIGURES

FIGURE	PAGE
Figure 1-1. A Two-lane Local Street with No Lane Markings.....	4
Figure 2-1. Road Model – Different Solutions (Source: Baumgartner et al. 1997).....	13
Figure 2-2. Ruskoné’s Vehicle Structure Hierarchical Model .....	16
Figure 2-3. Spectral Reflectances of Dry Soil, Wet Soil and Asymptotic Green Reflectance.....	20
Figure 2-4. Clusters in ( $a^*$ , $b^*$ ) Components of the CIELAB Color Space.....	23
Figure 2-5. Result of Edge Detection and Locations of Actual Streets .....	29
Figure 3-1. Methodology Framework.....	33
Figure 3-2. Profiles of Paved Area from Gaps between Vehicles .....	34
Figure 3-3. An Illustration of Lane Locations for a Scanned Road Region .....	346
Figure 3-4. An Example of Information File of Image.....	39
Figure 3-5. Determining the Pixel Locations of a Road Segment .....	40
Figure 3-6. Three Lane-profile Patterns and Their Brightness Fluctuations .....	42
Figure 3-7. An Example of Area Automatically Identified from Aerial Image .....	45
Figure 3-8. An Illustration of Classification Result for a Scanned Road Region .....	47
Figure 4-1. Studied Road Region (in Rectangle).....	50
Figure 4-2. Brightness Distributions of 30 Profiles .....	51
Figure 4-3. Studied Road Region with Plant Occlusions (in Rectangle).....	55
Figure 5-1. An Illustration of Classification Result for a Scanned Road Region .....	65
Figure 5-2. Saturation Distributions of Lane and Non-Lane Pixels .....	67
Figure 5-3. Aerial Images of Test Roadway Segments.....	74
Figure 5-4. Flexible Lane-Width Optimization for Lane Identification .....	78
Figure 5-5. Candidate Lane Locations Before and After Local Maximum Filtering .....	80

Figure 5-7. Workflow of the Location Optimization Module.....	91
Figure 6-1. Test Road Segments Identified by Automatic Extraction Program.....	98
Figure 6-2. Road Network with Heavy Plant Occlusion .....	99
Figure 6-3. Four Types of Road Segments Included in Test Networks .....	104
Figure 6-4. An Example of Road Segment with Heavy Occlusions.....	105

## ABBREVIATIONS AND ACRONYMS

AASHTO	American Association of State Highway and Transportation Officials
ASPRS	American Society for Photogrammetry and Remote Sensing
FP	False Positive
GIS	geographic information system
HSV	Hue-Saturation-Value
KNN	K-nearest neighbor
LSM	Least Square Matching
MrSID	Multiresolution Seamless Image Database
MSNN	morphological shared-weight neural network
NDVI	Normalized Difference Vegetation Index
RBF	Radial Basis Function
RCI	Roadway Characteristics Inventory
RGB	red, green, and blue
SVM	Support Vector Machine
TP	True Positive
Weka	Waikato Environment for Knowledge Analysis

# **CHAPTER 1**

## **INTRODUCTION**

### **1.1 Research Background**

Number of lanes is a basic roadway attribute that is widely used in many transportation applications. A good example is found in highway safety analysis, where it is classified as a required input variable for several tasks, including roadway division, roadway classification, and evaluation of traffic safety. In these tasks, number of lanes is used to reclassify the road network as site subtypes for further analysis. In addition, having the number of lanes on record also improves traffic volume estimation and aids in the calibration of regression models for the purpose of predicting crash frequency (Gan et al. 2005).

Roadway attribute databases with different detail levels are available from various sources. For Federal-aid or state roads, these attribute data are usually complete and are well maintained by transportation agencies. For example, the Florida Department of Transportation (FDOT) collects and maintains roadway characteristics data for state roads in the Roadway Characteristics Inventory (RCI) database, which includes the number of lanes. However, this information is unavailable for non-state roads which cover a majority of the road network. While it is possible to make a reasonable assumption that minor local roadways consist mostly of undivided two-lane roads, the same cannot be assumed for local collectors and arterials, which may vary from two to eight lanes, and may either be divided or undivided. In Florida, there are a total of about 1.5 million such road segments in which the number of lanes is yet to be acquired.

Traditionally, roadway characteristics data are collected through expensive field surveys. A team can collect data over a few hundred miles a day with vehicles that are equipped with a data collection system. This type of data collection can cost anywhere from \$8/km to \$20/km (*NCHRP Report*, 2000). Statewide data-collection tasks requiring field data collection for initial data acquisition or regular updates are almost impossible due to the high volume of non-state roads. Instead, a more efficient and inexpensive approach to number of lanes extraction is required.

One potential alternative is to extract these data from aerial images through automated image recognition techniques. Recent development in remote sensing technology has made it possible to use high-resolution ( $\leq 1$  meter/pixel) aerial images for large areas inexpensively. Therefore, data acquisition for an aerial image-based approach can potentially be relatively cheaper compared to field data collection. For smaller areas, the manual extraction of number of lanes could be feasible (Xiong and Floyd, 2004). However, for large areas (e.g., statewide or nationwide) involving tens of thousands of aerial images, an automatic extraction is a more feasible approach for both data collection and data updates.

## **1.2 Problem Statement**

The first step for the automatic extraction is to develop a lane model based on which travel lanes can be automatically identified in aerial images. Previous studies suggest that this task can be accomplished by recognizing several objects from aerial images, including lane markings, vehicle lines, paved areas with parallel edges, fluctuations of brightness, etc. In these studies, a lane model is usually defined using the patterns of visible objects (Baumgartner et al. 1997) or statistical features (Chi et al.,



2009). Lane models based on patterns of visible objects are typically used in the studies on lane or road extraction. For example, researchers from remote sensing and artificial intelligence communities explored methodologies to recognize paved areas with parallel edges (which can either be a lane or a road segment) from aerial images and were encouraged by the outcome of the progress (Mena, 2003). In fact, many road extraction algorithms include models that can be applied for recognition of lanes (Baumgartner et al., 1999; Hinz and Baumgartner, 2003). General approaches devised from these types of algorithms usually model a lane as the elongated homogeneous region bounded by parallel edges in aerial images. Therefore, recognition of pavement markings (i.e., parallel edges) and paved areas (elongated homogeneous regions) are two main areas of focus.

Existence of pavement markings and paved areas do not always guarantee the existence of a lane. For example, sidewalks are a type of inference that provides similar features as travel lanes. In color aerial images, sidewalks usually appear as elongated homogeneous regions bounded by clear edges that parallel the actual travel lanes. Consequently, if the aforementioned features are referred to as the “decision rules,” sidewalks are likely recognized as travel lanes in an automated recognition process. Similar to sidewalks, driveways also cause problems in automated recognition. Since driveways provide exactly the same image features as paved roadways and are usually connected to actual lanes, an automatic recognition process may define driveways as part of roadways, resulting in extra lanes.

Another complication of the marking/edge-based lane model is that, when applied to local collectors, it is unable to recognize the travel lanes because local collectors do not

have markings, edges, or lines within paved areas. Figure 1-1 shows an example of a local collector. The two-lane highway area within the rectangle has no lane markings, and it also does not provide any abrupt change that may lead to the detection of an edge within the paved area. Based on the aforementioned lane model, features extracted from this area could suggest the existence of only one lane.



Figure 1-1. A Two-lane Local Street with No Lane Markings

For the reasons discussed, lane models that rely on the identifications of pavement and markings could fail to identify travel lanes in many cases. Therefore, more information, such as the width, direction of homogeneity, and the color attributes of lane candidates are needed to increase the reliability of the automatic recognition process.

Other features or objects, e.g., abrasion on pavement (Chi et al., 2009) and lines of vehicles (Ruskoné et al., 1996) are also useful for lane recognition. In previous research on vehicle detection, Ruskoné et al. (1996) extracted vehicles and their orientations from aerial images and formed vehicle queues, which suggested the

existence of lanes. This approach turned the problem of extracting number of lanes into a vehicle detection problem. For roadways with heavy traffic volumes where pavement is likely to be occluded by vehicles, extracting number of lanes by vehicle-line detection could be a better choice. However, one major limitation with this approach is roadside parking, which suggests an additional lane in each travel direction. In addition, smaller parking lots with  $N$  lines of vehicles may be recognized as a segment of  $N$ -lane highway. Chi et al. (2009) suggested that the existence of a lane could also be identified by capturing the fluctuation pattern of brightness values on pavement. Since no specific object is required to be extracted, this technique significantly reduces the amount of information needed to detect a lane. However, to use this technique for automatic extraction of number of lanes, two preliminary tasks must be accomplished in advance:

1. Approximate location of the road segment must be specified for extraction.
2. Vehicles on the road segment must be removed from the image.

These two tasks are more difficult to implement in the automatic extraction procedure due to the need for further investigative research on road and vehicle detection.

A more common problem of the existing techniques is the inability to ensure a reliable performance. In simple scenarios where the impacts of noises (i.e., shadow or building occlusions) are insignificant, it is possible to extract lanes by recognizing the aforementioned features or objects. However, real-world aerial images, especially for urban areas, usually include various occlusions from buildings, vehicles, and shadows, all of which could seriously impede recognition efforts. Images of rural or suburban areas typically include roadways with a lesser density of vehicles and buildings. In these areas, existing techniques also do not perform well due to the presence of roadside vegetations,

which generate continuous and irregular occlusions along roadways and replace clear edges provided by curbs or lane markings. Consequently, lane models relying on the recognition of these linear features tend to yield poor results. In the highly vegetated areas of northern and central Florida, this limitation would exclude a significant percentage of roads.

The existing literature lacks in-depth discussion on reliability variation. Reliability of road or lane recognition approaches can typically be evaluated by two indicators: the percentage of roads/lanes recognized, and the success rate (Auclair-Fortier et al., 1999). In areas with numerous obstacles (e.g., heavy traffic or many buildings), an automatic road/lane recognition approach typically results in poor performance under either indicator. Current literature usually labels test results with area descriptions (e.g., rural, suburban, or urban areas) to indicate features of studied areas. This type of simple connection built between reliability and discrete labels is straightforward and has been widely used in previous work (Hinz and Baumgartner, 2003; Grote et al., 2007). However, this is not feasible in this research because the data collection that covers large areas may involve tens of thousands of images to be labeled.

The disadvantages of manual area labeling aforementioned are two-fold. First, it may need extensive manual work to classify the target images. For a task with thousands of aerial images, this classification procedure could be burdensome. Second, choosing a label for an image is an arbitrary decision due to the inaccurate definitions of labels. For example, some may understand the concept behind the label “suburban” as something entirely different from the designer of a technique. As a result of this misunderstanding, a

“suburban” area that is labeled as “rural”, for example, would lead to the expectation for a higher recognition rate.

In summary, existing models and techniques for automatic extraction of number of lanes have the following problems:

1. Difficulty in distinguishing a lane from its surroundings in aerial images.
2. Feature extraction may require external knowledge (location or direction of roadways in aerial images), which is difficult to obtain in advance for an automated procedure.
3. Reliability is difficult to predict.

### **1.3 Goal and Objectives**

The goal of this dissertation is to develop a more reliable and robust framework for automatic extraction of number of lanes from geo-rectified aerial images. The main objectives of this dissertation are:

1. Develop a reliable lane model based on the typical appearance features of travel lanes in color aerial images, which is more robust against noises such as vehicle occlusions and sidewalks.
2. Develop an algorithm to locate travel lanes in aerial images based on the improved lane model and the algorithm to distinguish number of lanes from identified lane locations.
3. Implement the framework and evaluate its effectiveness.

## **1.4 Dissertation Organization**

This dissertation consists of seven chapters. Chapter 1 introduces the background of this dissertation research, presents the problem to be solved, and sets the goal and objectives to be achieved.

Chapter 2 presents an extensive review of modeling issues for road segments and noise factors, as well as research on automatic road recognition from aerial images. The purpose of this review is to explore the existing models of the objects likely to present around paved road segments, identify the pending problems of the current research on automatic lane extraction.

Chapter 3 introduces the methodology developed for this research. This chapter first provides a brief description of the methodology framework. A technique of automatic road spatial information extraction is then presented, followed by the description of a profile-based lane model, which is made feasible with the help of spatial information. After specifying the workflow of applying the new lane model to aerial images, the rest of the chapter presents the idea of extracting number of lanes from preliminary recognition results based on their statistical features.

Chapter 4 discusses two different approaches to building a profile-based lane model, namely a similarity-score-based approach and a classification-based approach. After a general comparison, two simple test scenarios are provided to validate the classification-based approach. In the first test scenario, lane profiles and three noise factors are included. Three classification techniques, namely, K-nearest neighbor (KNN), logistic regression, and Support Vector Machine (SVM), are used to train the classifiers to recognize lane profiles from samples. In the second test scenario, test samples are

divided into lane profiles and non-lane profiles in a more general manner. SVM is selected to validate a profile-based lane model in this more complex case.

In Chapter 5, two types of optimization algorithms for automatic extraction of number of lanes from classification results are presented: fixed lane-width optimization and flexible lane-width optimization. The objectives of these optimization algorithms are to maximize the support of lane locations from classification results generated from a profile-based lane model and the reliability. A simple test case, which includes six multi-lane roadways, is provided to test this fixed lane-width optimization algorithm.

Chapter 6 describes the application and the evaluation of the framework. A MATLAB program that implements the framework using a two-class lane model and flexible lane-width optimization is presented. Road networks included in the two test scenarios are then reviewed. Lastly, the performance data in the two test scenarios are evaluated.

Chapter 7 first summarizes the major research findings from this research. It then discusses the contributions of this research and provides recommendations for future research.

## **CHAPTER 2**

### **LITERATURE REVIEW**

#### **2.1 Introduction**

As is evident from the current literature, automated extraction of number of lanes from aerial images has yet to be systematically studied as a primary research objective. There are three main research areas involved:

1. The models for lane and related objects in color aerial images.
2. The techniques used to extract features included in the model.
3. The algorithm to decide number of lanes from extracted features with reliability estimation.

Among these areas, the establishment of models for travel lane and related objects is the most critical, since extraction techniques and decision algorithms are usually adopted or designed based on the nature of features desired for extraction. Although existing road models may help identify features useful for travel lane recognition, a full review of the techniques and decision algorithms used is not feasible, nor needed, since these models have goals that are very different from those of this research. Therefore, this review is mainly focused on lane models and related objects in color images.

To achieve the objectives of this research, it is critical to target the following types of objects in aerial images:

- Pavement and markings
- Vehicles
- Buildings
- Plants



- Shadows

Among these objects, shadows are the only type of objects that cannot provide useful information for lane identification. Therefore, several shadow removal approaches are reviewed below. For each of the remaining objects, various modeling approaches will be reviewed in the following sections. Literature on road/profile extraction methodologies are reviewed in the last section to introduce several typical approaches in road segment extraction, which should also be useful for this particular research.

## **2.2 Road Models**

Appearances of roads vary depending on the resolution of images and the scenery depicted therein. In low-resolution images, roads are usually recognized as linear objects. In this case, road extraction is usually referred to as line extraction. In high resolution images (resolution  $\leq 1\text{m/pixel}$ ), roads are projected as elongated homogeneous regions (Baumgartner et al., 1999). Thus, the process of mapping the physical characteristics of road regions to visual characteristics is critical in advancing efforts in road detection and extraction.

In earlier literature, Bajcsy and Tavakoli (1976) classified the physical characteristics of roads into four groups: spectral properties, geometric properties, topological properties, and contextual properties. Although this classification method is helpful in understanding road models, it only provides general descriptive knowledge rather than specific guidance in image processing language. Later, Gruen and Li (1995) proposed a generic road model that is more practical for image feature extraction. They suggested five properties of roads that could be useful in high resolution images, as follows:

1. Road is the region with high intensity itself and lower intensity on both sides.
2. Gray level along road direction does not change within a short distance.
3. Road is a smooth object without wiggles.
4. Curvature of a road is bounded by a certain threshold.
5. Width of road does not change significantly.

These properties can be formulated mathematically as constraints in image processing. Mathematical representations of properties (1) through (4) are defined in Equations 2-1 through 2-4, respectively:

$$E_{p1} = \int \{ G[f(s)] \}^2 ds = \text{Maximum} \quad (2-1)$$

$$E_{p2} = \sum \int \Delta s_i \{ G[f(s)] - G_m(\Delta s_i) \}^2 ds = \text{Minimum} \quad (2-2)$$

$$E_g = \int | f''(s) |^2 ds = \text{Minimum} \quad (2-3)$$

$$C_g = | f''(s) | < T_1 \quad (2-4)$$

where

$f(s)$  = a vector function mapping arc-length  $s$  to point  $(x, y)$  in the image,

$G(x, y)$  = a function representing the preprocessed digital image,

$\Delta s_i$  = the  $i$ th short segment of curve  $s$ , and

$G_m(\Delta s_i)$  = the average value of  $G$  on  $\Delta s_i$ .

Baumgartner et al. (1997, 1999) proposed a road model which, for large-scale images, conceptualized road segments as a combination of pavement and markings. The structure of the model is shown in Figure 2-1.

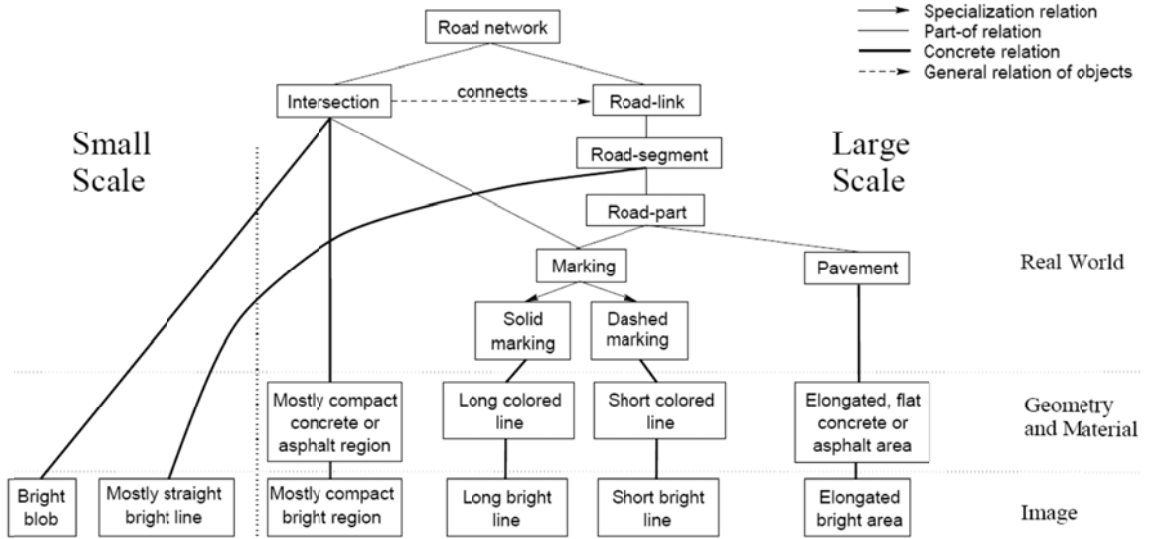


Figure 2-1. Road Model – Different Solutions (Source: Baumgartner et al. 1997)

At the image level, the physical characteristics included in Figure 2-1 are represented by an elongated bright area bounded by bright lines. More specifically, markings can be modeled as a bright line with symmetric contrast (Steger, 1998), and pavement can be modeled as a region with homogeneity along road direction (Gruen and Li, 1995). The texture feature designed to measure the degree of homogeneity,  $f$ , is defined as follows (Haralick et al., 1973):

$$f = \sum_i \sum_j \{p(i, j), d, \theta\}^2 \quad (2-5)$$

where

$p(i, j)$  = the  $(i, j)$ th normalized gray-tone spatial-dependence matrix,

$d$  = the distance between two pixels, and

$\theta$  = the degree selected from set  $\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ .

In this research, the texture feature of pavement implies that the homogeneity function value along the road direction  $(\sum_i \sum_j \{p(i, j), 1, 0^\circ\}^2)$  is no smaller than the values along any of the other three directions. This pavement texture feature was further

expanded by Hinz and Baumgartner (2003) based on the observation that roads usually show good homogeneity in driving direction but moderate homogeneity perpendicular to it. The new feature is critical in improving efforts to erase noise from driveways.

Trinder and Wang (1998) defined road segment in a similar form. Furthermore, road pavement was modeled as a region with visible boundaries (edge lines). In computer vision, an edge is detected by searching for the borders between two regions, each of which has approximately uniform brightness (Horn, 1986). Regions separated by road edges typically feature one-sided homogeneity (Lanser and Eckstein, 1992). This is a pivotal characteristic due to the fact that buildings, plants, and other objects can also generate regular parallel edges similar to those of roads.

In summary, in high-resolution aerial images, a road segment can be physically constituted by several primitives that include pavement, marking, road edge and constrained smoothness, width, and curvature constraints. Preliminary assessment of the complexity of extracting these primitives is better understood by the following three-level definition (Ballard and Brown, 1982):

1. Low-level: Pixel-level characteristics such as color and texture.
2. Medium-level: Symbolic representations such as lines and regions.
3. High-level: Semantic, cognitive knowledge of images.

Theoretically, primitives of road segments can be dealt with by low-level and medium-level processing. However, in real-world high-resolution images, noises such as occlusions or shadows will significantly impede automatic extraction if no high-level processing is involved. For example, trees and vehicles may occlude road surface and may strongly influence road appearance (Zhou et al., 2005). Therefore, it is critical to

determine which other components are included in real-world road regions, and how these components are subsequently represented in multiband aerial images.

### **2.3 Vehicle Models**

Vehicles are one of the most common objects that occlude road segments. Although it is not a physical component of the road segment, vehicles can actually provide critical information for lane detection. For example, Hinz and Baumgartner (2003) developed an automatic road extraction framework in which successfully detected vehicles were captured as part of a lane. Another study by Ruskoné et al. (1996) extracted the orientation of vehicle-region candidates from images and formed lines for vehicle queues. Intuitively, parallel vehicle-lines under certain constraints can be relied upon for the inference of number of lanes. As such, these typical correlations may raise concerns for the viability of vehicle models in aerial images. Based on the underlying models, this study has adopted the classification of Hinz (2004) by grouping the related work into two types: implicit and explicit.

Models established following the implicit approach can be found in Ruskoné et al. (1996), Rajagopalan et al. (1999), Schneiderman and Kanade (2000), Papageorgiou and Poggio (2000), Zheng et al. (2006), Jin and Davis (2007), and Leitloff et al. (2010). This group of models typically characterizes vehicles by using statistics of intensity values or textures. With classifiers built on positive (with vehicle) and negative (without vehicle) sample pictures, pixels belonging to vehicles were retrieved from target pictures. The detection procedure is then accomplished by matching the statistics of candidate regions in a target image with that of sample images.

Ruskoné et al. (1996) proposed a vehicle-structured hierarchical model with higher semantic levels (see Figure 2-2). In this model, as is similar to other literature, vehicles are represented as clusters of pixels showing common geometric (rectangular shape) and radiometric (gray level thresholds) properties. In addition, the presence of vehicles can also be further confirmed if the detected rectangles' orientations follow a line (on road) or form an angle of 45 or 90 degrees (on parking lots). Furthermore, Rajagopalan et al. (1999) modeled vehicles using high-order moments achieved from sample pictures. Moments from second and upward to order  $m$  were first generated from sample pictures to model. Joint moments of these training samples were then established as a model for the vehicle. In a recent paper (Leitloff et al., 2010), the authors tested a co-occurrence matrix, the histogram of oriented gradients, and Haar-like features proposed by Papageorgiou et al. (1998), and concluded that only the last group of features captured appearance fairly well.

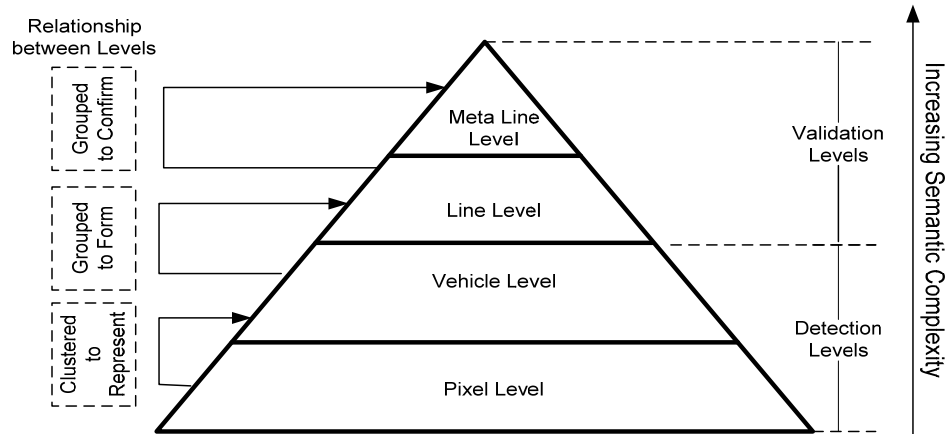


Figure 2-2. Ruskoné's Vehicle Structure Hierarchical Model

Schneiderman and Kanade (2000) decomposed images into ten sub-bands using wavelet transformation. A set of 17 attributes derived from these sub-bands were evaluated to represent pace, frequency, and orientation information, all of which modeled

the vehicle region in an image. The attribute definitions in this paper followed the four categories proposed by Cosman et al. (1996), as follows:

1. Intra-sub-band: All of the coefficients originate from the same band.
2. Inter-frequency: Coefficients originate from the same orientation, but multiple frequency bands.
3. Inter-orientation: Coefficients originate from the same frequency band, but multiple orientation bands.
4. Inter-frequency/inter-orientation: A combination of coefficients that represent cues that span a range of frequencies and orientations.

Papageorgiou and Poggio (2000) also used wavelet transformation to model vehicles. For each training pattern in their research, a quadruple density Haar transformation in each color channel was computed. Then, at each specific location and orientation, the wavelet was assigned the largest value from RGB channels. The model was finally represented by feature vectors (3,030 coefficients) from both positive (with vehicle present) and negative (without vehicle) image regions.

Jin and Davis (2007) established an interesting model that was actually a trained morphological shared-weight neural network (MSNN). Since morphological structures in MSNN could work as both a classifier and a feature detector, the authors claimed that the trained MSNN incorporated both the spatial and spectral characteristics of vehicles in images.

The explicit approach, on the other hand, creates a general 2D or 3D model of the vehicle, which does not rely on specific training images (Tan et al., 1998; Haag and Nagel, 1999; Liu et al., 1999; Michaelson and Stilla, 2000; Zhao and Nevatia, 2001; Hinz

and Baumgartner, 2001; and Kozempel and Reulke, 2009). Chellappa et al. (1996) proposed a vehicle model that described a vehicle as a two-dimensional rectangular shape in an aerial image. Since this model ignores the specific shape and color of a vehicle, the algorithm based on this model depends primarily on the gray level difference along four boundaries (Moon et al., 2002). Similar 2D models can be found in Burlina et al. (1997) and Moon et al. (2002). Inspired by psychological tests, Zhao and Nevatia (2003) created a model that integrated the boundary of the vehicular body, the boundary of the front windshield, and the vehicle's shadow as features of a vehicle in an aerial image.

Hinz (2004) proposed an adaptive model that consists mainly of geometric features, as well as radiometric properties. Compared to a model with only geometric features, the new model adjusts the expected saliency of edges based on color, vehicle orientation, viewpoint, and sun direction. This model is more specific and was claimed to be a robust approach with its disadvantages. For example, in comparison to a vehicle with good contrast, considering a vehicle with a similar color to pavement would present a weak gray level edge; therefore, the detection procedure in Hinz's model is less likely to reject the former candidate region as the result of a lowered expectation for saliency. However, the adaptive model is actually a set of models, each of which covers one type of vehicle, which is more expensive to build than a single, standalone model.

Given the above results, it is clear that both approaches have their pros and cons. Nevertheless, some of the literature (Suetens et al., 1992; Hinz, 2003; and Schlosser et al., 2003) suggests that the explicit model is more robust when applied to urban areas. One major reason is that close distances lead to the overlapping of objects in an image. Consequently, the extracted radiometric measures will often be mixed with heavy noises



and are thus more difficult to match to models generated from sample images. Another reason cited in literature is that sample images may not represent necessary information, such as illumination and viewpoint. Therefore, even if the trained classifier or learning systems yield suitable results in analogous scenes, there is still a concern for performance reliability when applied to different image sets.

## **2.4 Representations of Plants**

Another type of object that may impede road extraction efforts are plants. There are two types of plant-related problems that should be considered in road models:

1. Grass on the median or roadsides, which can provide similar texture features as pavement.
2. Trees may occlude primitives of road segment, e.g., pavement, markings, and edges.

Research efforts in remote sensing have long been using the Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1973) to distinguish vegetation from other materials, such as soil and water, in multiband images. NDVI is defined as follows:

$$NDVI = \frac{(NR - R)}{(NR + R)} \quad (2-6)$$

where  $NR$  is near-infrared band and  $R$  is the red band. The index is created based on the observation that reflectance of vegetation is high in the near-infrared band and low in the red band. Figure 2-3 shows the difference between soil and vegetation in terms of the reflectance rate.

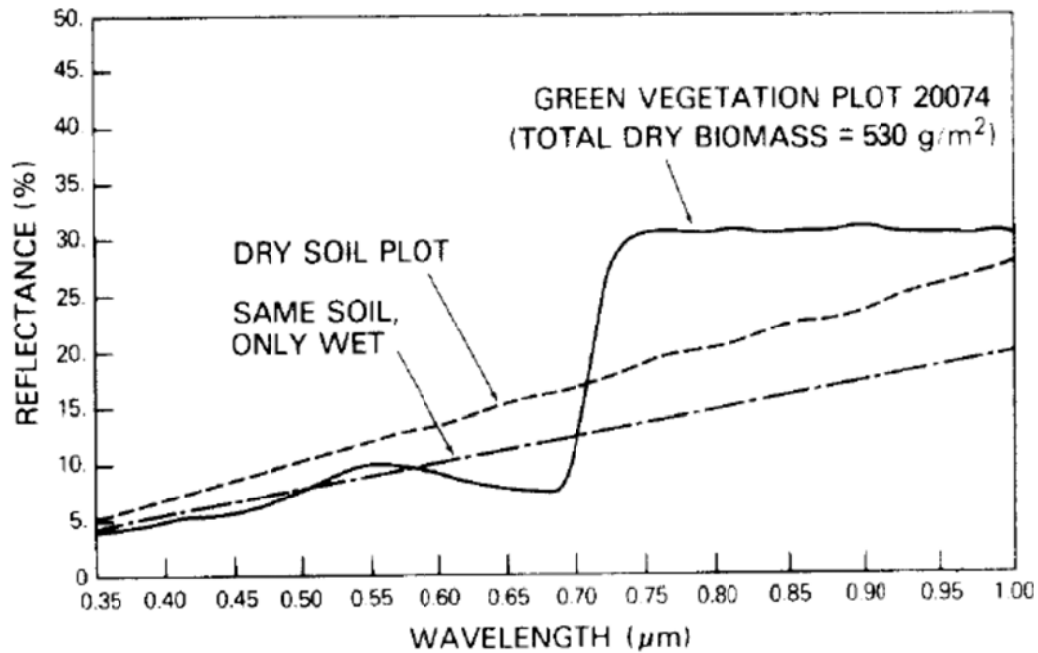


Figure 2-3. Spectral Reflectances of Dry Soil, Wet Soil and Asymptotic Green Reflectance  
(Source: Tucker 1979)

Since NDVI is a continuous value between  $[-1, 1]$ , in an aerial image, plants can be modeled as pixel clusters, each of which has an NDVI index higher than a given threshold. Therefore, it is critical to choose a reasonable threshold to model plants by NDVI. Based on theoretical and experimental results (Myneni et al., 1990; Baret and Guyot, 1991; Asrar et al., 1992; and Nemani and Running, 1997), NDVI of 0.4 can be referred to as the general guideline, even though soil brightness can actually produce significant variations (Liu and Huete, 1995). Generally speaking, NDVI values for road segment or soil are much lower. For bare soil, NDVI varies from -0.1 to 0.2 (Carlson and Ripley, 1997). For paved roads, Grote et al. (2007) suggested that 0 can be used as the threshold, and that any pixel exceeding this should be excluded from the road region.

A monochromatic image can be easily obtained from the original multiband image by evenly distributing the index into 256 gray levels. Accordingly, thresholds of

NDVI can be interpreted as gray levels to model plants in images. However, in a typical geo-rectified aerial image, the infrared band is not always available. In the digital images used in this research, only three bands are available: red, green, and blue (RGB). Tucker (1979) compared different linear combinations of two pairings of these bands, red-infrared and green-red, and concluded that the first pair, on the basis of which NDVI was designed, was superior to green-red. In his research, another greenness measurement, denoted as  $VI_{g/r}$ , as expressed below, was compared to NDVI:

$$VI_{g/r} = \frac{(G - R)}{(G + R)} \quad (2-7)$$

where  $G$  is the green band and  $R$  is the red band.

Although the significance levels of  $VI_{g/r}$  were lower than those of NDVI, most of its coefficients of determination varied from 0.73 to 0.84. The research showed that the  $VI_{g/r}$  index was a promising indicator for vegetation groups in the summer seasons, with the exception of the dry brown biomass group. Thus, the index was deemed suitable as an alternative to NDVI in images with only RGB bands. For practical reasons, researchers have actually adopted  $VI_{g/r}$  to distinguish vegetation in road extraction from digital images (Baltasavias and Zhang, 2005; Jin et al., 2009). The impact of seasons, however, was not discussed in the road extraction applications mentioned above. As can be expected, aerial images generated in the summer will provide well-separated index values for different objects better than other seasons. Statistical evidence of the seasonal impacts can be found in Tucker (1979). In practice, operators can minimize the impacts by selecting the aerial images acquired in seasons with least precipitation.

## 2.5 Building Detection

Depending on the context (rural, suburban, or urban) of an image, the impact of building occlusions on the appearance of road segments varies significantly. Empirical knowledge suggests that the demand for building detection along road segments will rise as the region studied becomes more urban. The objective of building detection in this dissertation research is to solve the occlusion problem it poses for the road extraction procedure described herein.

Classical building extraction approaches usually model the object as a 3D structure, although many of the studied data were 2D images. In one of the early literatures on building extraction, Herman and Kanade (1984) proposed a model that conceptualized buildings as rectangle prisms. Later, Shufelt (1996) proposed two 3D wireframes, namely, rectangular volume and triangular prism, to describe model buildings. A comprehensive survey of early literature on the automatic extraction of buildings from aerial images can be found in Mayer (1999). The abovementioned models are mainly designed to reconstruct buildings from an image. In this dissertation research, however, the primary objective was to distinguish the building region from road segment. Henricsson (1998) suggested an efficient strategy in which the roofs of buildings were found through the grouping of color features. The test results are given in Figure 2-4. The new variables  $a^*$  and  $b^*$  are two components of CIELAB color space transformed from RGB. The transformation is defined as follows:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2.7690 & 1.7518 & 1.1300 \\ 1.0000 & 4.5907 & 0.0601 \\ 0.0000 & 0.0565 & 5.5943 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2-8)$$

$$L^* = 116 \times \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} - 16 \quad (2-9)$$

$$a^* = 500 \times \left[ \left( \frac{X}{X_n} \right)^{\frac{1}{3}} - \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} \right] \quad (2-10)$$

$$b^* = 200 \times \left[ \left( \frac{Y}{Y_n} \right)^{\frac{1}{3}} - \left( \frac{Z}{Z_n} \right)^{\frac{1}{3}} \right] \quad (2-11)$$

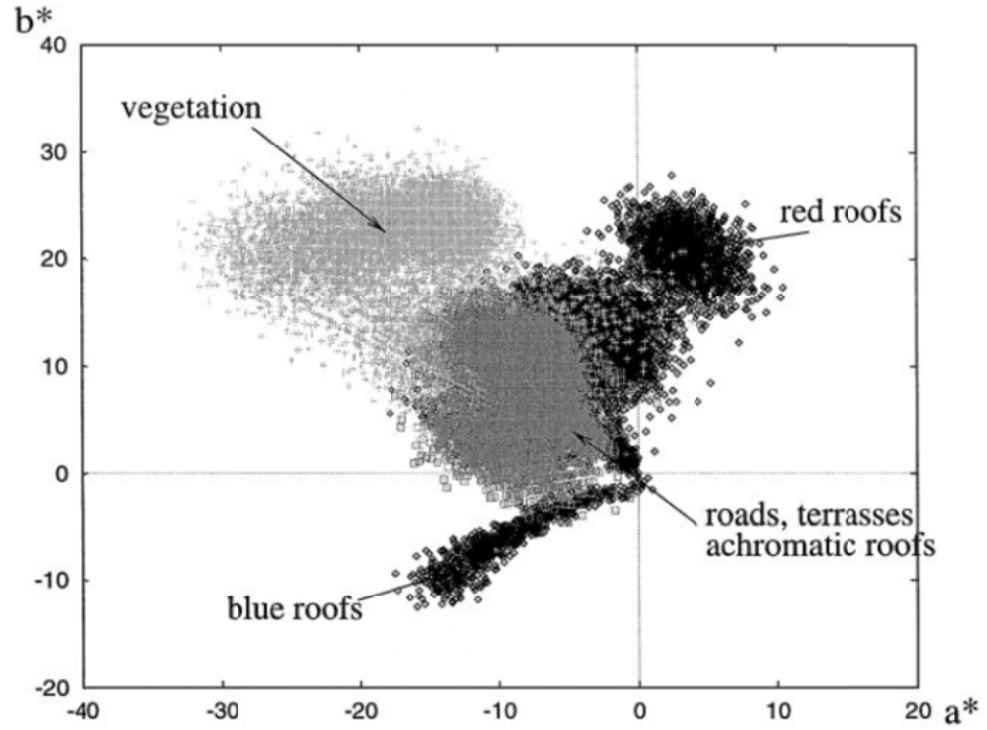


Figure 2-4. Clusters in ( $a^*$ ,  $b^*$ ) Components of the CIELAB Color Space

(Source: Henricsson 1998)

Not surprisingly, objects such as roads, grey roofs, and shadows cannot be well-separated based on color clustering. To address this issue, Fischer et al. (1998) presented a generic model that actually described buildings as a hierarchical structure consisting of building parts. According to this model, the first step of building detection

was to find the corners of a house by searching for its vertices. Although our primary interest is not the reconstruction of a building, this model suggests that the existence of building's vertices imply that the surrounding region will be pavement-free.

## 2.6 Shadow Detection and Removal

Another major negative factor that arises from buildings in images is shadow.

There are mainly two reasons that motivate interest in shadow detection and removal:

1. Unlike that of plants, the shadows of buildings can occupy the majority of road areas and thereby occlude the region of interest.
2. Even if the objects under a shadow are recognizable, the identification of the real boundaries of the desired objects (e.g., road segments) is likely to fail due to the false hue information generated by the shadow (ASPRS, 1996).

Shadow detection has been studied based on two different motives. One group of researchers is interested in detecting shadows to facilitate the detection of objects in shadow areas (Huertas and Navatia 1988, Irvin and McKeown 1989, Liow and Pavlidis 1990, Lin et al. 1994, Watanabe et al. 1998, Noronha and Nevatia 2001). The other group is mainly interested in shadow detection and removal. Scanlan et al. (1990) proposed an intensity-based algorithm to remove shadow from gray scale images. The algorithm partitioned the original  $X \times Y$  image matrix into  $l \times l$  blocks and calculated the mean intensity value  $A_{ij}$  ( $i \leq X/l, j \leq Y/l$ ) for each block. The median of all  $A_{ij}$  ( $M$ ) was then selected as the threshold. All pixels in blocks with  $A_{ij} \leq M$  were then multiplied by the scaling factor  $\alpha_{ij}$ , which is defined as:

$$\alpha_{ij} = \frac{M}{A_{ij}} \quad (2-12)$$

$$M = \text{Median} \{ A_{ij} | i \in [1, \frac{X}{L}], j \in [1, \frac{Y}{L}] \} \quad (2-13)$$

$$A_{ij} = \text{Mean} \{ G_{ij} | i \in [1, L], j \in [1, L], G_{ij} \in [0, 255] \} \quad (2-14)$$

This algorithm actually identified a block as a shadow region if its average intensity was less than the median of all blocks. Any suspected shadow block was then removed by scaling factors. When testing the algorithm, Scanlan et al. (1990) suggested a small block size ( $l=2$ ). Since the algorithm inevitably distorts the intensity values of the original image, it is important to assign a small  $l$  value to keep the texture feature.

Tsai (2003) followed the same principle and recursively partitioned the images into quad-tree sub-images until each of the sub-images provided a bimodal distribution histogram. Unlike the global threshold used in Schalan's algorithm, the new approach adopted different thresholds for different sub-images. For each sub-image, the threshold value  $T$  was then assigned for shadow recognition according to the following cost function:

$$\eta(T) = \frac{\sigma_w^2}{\sigma^2} \quad (2-15)$$

where

$\sigma_w$  = among-cluster variance of the two clusters separated by  $T$ , and

$\sigma$  = total variance of sub-image.

The detected shadow region is then removed by rescaling the shadow pixels using the histogram matching method (Richards and Jia, 2006). However, the authors did not provide any specific decision rule for the termination of partitioning. As such, a bimodal distribution requirement was not formulated.

Shadow detection is more complicated in multiband images. However, certain transformations of color space may actually minimize the impact of shadow. These special types of transformations were found to be stable, or invariant, regardless of illumination conditions and were thus named color invariants (Gevers and Smeulders, 1999). Assuming that  $F$  is a color invariant,  $F_s$  and  $F_l$  are the values of a point in the shadow and light, respectively, the nature of color invariants implies the following:

$$F_s = F_l \quad (2-16)$$

Examples of color invariants include normalized RGB, hue (H), saturation (S), and C1C2C3 (Gevers and Smeulders, 1999). The definitions of  $C_1$ ,  $C_2$  and  $C_3$  are specified as:

$$C_2(x, y) = \arctan \frac{R(x, y)}{\max(G(x, y), B(X, Y))} \quad (2-17)$$

$$C_2(x, y) = \arctan \frac{G(x, y)}{\max(R(x, y), B(X, Y))} \quad (2-18)$$

$$C_3(x, y) = \arctan \frac{B(x, y)}{\max(G(x, y), R(X, Y))} \quad (2-19)$$

where for a pixel at location  $(x, y)$ ,

$R(x, y)$  = red component,

$G(x, y)$  = green component, and

$B(x, y)$  = blue component.

Since Kender (1976) indicated that normalized RGB was unstable near the black vertex of the RGB space, and that hue was unstable near its singularities at the entire achromatic axis, Salvador et al. (2001) suggested that  $C_1C_2C_3$  color space be used to reduce shadow impacts.



Although an improvement upon previous shadow reduction efforts, color-based approaches fail to exploit image geometric information and, as such, could perform unsatisfactorily when applied to complex images (Yao and Zhang, 2005). Salvador et al. (2004) presented a new approach that included both color and geometric information. Using this approach, each test region was verified by hypotheses based on color invariance and geometric properties of shadows. Later, Yao and Zhang (2005) proposed a two-level hierarchical algorithm which was claimed to be more robust and reliable under different brightness and illumination conditions. According to their algorithm, when an image is mapped into a 2D undirected graph at the pixel level, each node represents three properties: color of pixel, shadow probability, and node reliability. The procedure of shadow detection is then converted into the problem of achieving maximum graph reliability by iteratively updating the graph.

## **2.7 Automatic Road Segment Recognition**

Different approaches to automatic road extraction have been widely studied by researchers since the 1980s. Based on the different road models discussed in the previous section, road segment detection techniques were proposed. For a typical automatic detection procedure, there are two issues to be resolved, i.e., how to detect a candidate road segment around a seed point and how to determine the next detection area. This typically requires techniques such as edge detection or clustering. After being validated under various constraints, segment candidates are connected for the reconstruction of road networks. For an extensive survey on this topic, see Mena (2003).

In this dissertation, one of the goals is a successful means of road region detection, which is similar to part of the first step of automatic road extraction. Unlike previous

studies on road extraction, this dissertation is more focused on finding robust and reliable approaches to detect details of objects in and around road segments in high-resolution images. Road profile and edges are two typical analysis units in such detection (McKeown and Denlinger, 1988; Vosselman and Knecht, 1995; Baumgartner et al., 1999; Hinz and Baumgartner, 2003; Zhou et al., 2006; Movaghati et al., 2010).

Edge detection algorithms capture abrupt changes of intensity, and are thus useful for the retrieval of road segment outlines. Common edge detectors such as Sobel, LoG (Laplacian of Gaussian), and Canny (Canny, 1986) have been widely used in the automatic road extraction. When applied to a digital image, edge detectors usually generate a large amount of discrete edges, points, or short lines. It is difficult to retrieve the actual outlines of road regions unless they involved simple scenarios, such as rural highways or freeways with low traffic volume. Figure 2-5 provides an image of detected edges from a suburban area in Miami, Florida. The algorithm used the Canny method with a mask size of 7x7. Two lines are drawn in the figure to provide the approximate locations and directions of two roads.

Another type of approaches for road segment recognition is based on profiles. A profile is defined as a slice of road segment perpendicular to the direction of traffic. Since a road typically provides low homogeneity in this direction (Hinz and Baumgartner, 2003), it can be expected that the variance of pixels within a profile is high, and rich information can thus be extracted from this sample. However, for a random image file, finding the direction of traffic is not much easier than that of recognizing a road segment. Thus, early literature usually included profile-based road recognition as a method of semi-automatic extraction. This means that either an initial direction of profile or the

traffic direction of a link must be assigned by an operator (McKeown and Denlinger, 1988; Vosselman and Knecht, 1995; Zhou et al., 2006; Movaghati et al., 2010).



Figure 2-5. Result of Edge Detection and Locations of Actual Streets

## 2.8 Summary

In this chapter, models of road segment and related objects, as well as typical features for road segment recognition, are reviewed. The multi-fold purpose of this review includes understanding the image objects included in automatic extraction of number of lanes from color images, identifying the current models of these objects, and investigating image features that need to either be extracted or controlled for lane recognition.

Literature on road segment recognition typically models it as a structure combining primitives (e.g., pavement, marking, and road edge) and constraints (e.g., smoothness, width, and curvature). As part of a road segment, a lane includes similar features. However, these features, as discussed in existing literature, lack the necessary information that must be provided to distinguish a lane from parallel lane(s), turn-bays, and other objects within the paved area, such as medians and shoulders. This is a critical issue that needs to be addressed in this dissertation research.

Another common problem with existing lane models is robustness. These models were usually developed for road extraction and are typically suitable for suburban or rural areas, in which case occlusions or noises from shadow are minimized. Hence, recognizable lane segments must be close enough to be connected for road reconstruction. When a significant part of the target road segment is occluded by vehicles or buildings, any algorithm attempting to connect lane segments will fail because the regions may be too small, making it too far away from each other to be connected under certain constraints, such as curvature and width.

Vehicles, plants, buildings, and shadows are four types of objects that may impede the process of lane recognition. Among these four objects, vehicles detected using either implicit or explicit models may be helpful if vehicle lines are recognized at the same time. Otherwise, occluded regions should be excluded from pavement-based lane models. Areas covered by plants can be distinguished using indices developed from the color components of pixels. Therefore, these indices can be introduced into a desired robust lane model to remove the area covered by plants. Similar transformations of color components may also help reduce shadow impact. Among various techniques for shadow

removal, some transformed color systems were found to be invariant regardless of illumination condition. This implies that, even if a lane model is trained by lane segments with no shadow, it is possible that shadowed lanes can still be recognized by the model if color invariants are adopted to represent features. Lastly, buildings that occlude roadways can be recognized through shape-based models (2D/3D or vertices) or color of roofs. Although vehicles may provide similar shapes or color features, these models are still useful when trying to distinguish a paved lane area from surrounding objects.

Due to the similarities between lane and road segment, two analysis units widely used in research on road extraction, namely, profile and edge, are also investigated. Existing literature shows that profile is the more applicable analysis unit for extraction. However, initial manual guidance is always required to provide traffic direction in investigated literature. Therefore, existing profile-based extractions are usually semi-automatic and not suitable for this dissertation research.

## **CHAPTER 3**

### **METHODOLOGY**

#### **3.1 Introduction**

This chapter describes the methodology applied in this dissertation research. A framework consisting of several modules for robust automatic extraction of number of lanes is first presented. Details of each module in the framework are then discussed.

#### **3.2 Methodology Framework**

The framework of the methodology is shown in Figure 3-1. It consists of four modules: data and model preparation, lane-profile recognition, optimization, and post-processing. The inputs required for the framework is a special type of aerial images of the target area, the geo-coordinates of the target road segments, and the lowest accuracy rate the operator can accept. After the processing of the four fully automated modules, the framework will either identify the number of lanes for the target road segment or leave it for manual check.

In the first module, namely, the data and model preparation module, the applied methodology has two major functionalities: training a lane model based on lane profile patterns and extracting candidate profiles for target roadways. As the basic extraction unit used in the framework, a profile is defined as a slice of road segment (one-pixel wide) perpendicular to the direction of travel. In most existing lane models, the basic extraction units are usually more complex objects, such as marking lines, pavement areas, etc.

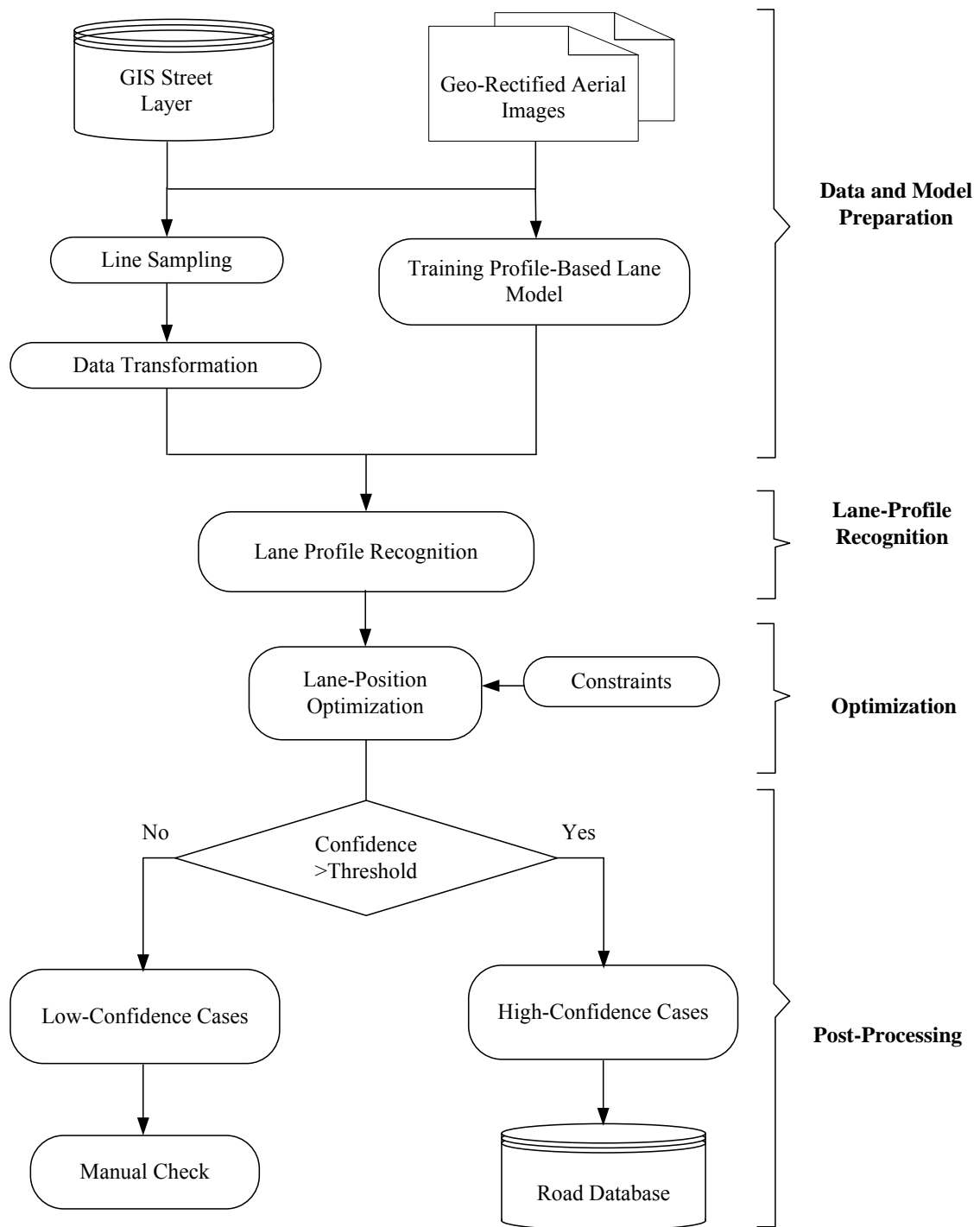


Figure 3-1. Methodology Framework

The reasons for adopting profile as the extraction unit are two-fold: (1) a profile of a paved road can provide sufficient pattern information to identify lane(s) and consistency of conclusions from multiple profiles of identical paved road can be used to validate the results (Steger, 1998; Baumgartner, 1999; Chi et al., 2009); and (2) vehicles, buildings, or plants will leave gaps between one another in paved areas, even if they occlude the majority of a road segment (see Figure 3-2).



Figure 3-2. Profiles of Paved Area from Gaps between Vehicles

The second reason is critical to the reliability of the new approach. It implies that, under the framework of the methodology, only a few successfully extracted road profiles



of paved area are needed to infer the number of lanes for a road segment. However, it must be noted that more successfully extracted profiles of a paved area provide a higher confidence in extracting the number of lanes for a particular road segment.

Using a set of training data that includes lane profiles and profiles of other noise factors, profile-based lane models can be established based on the special patterns discovered. Details of these patterns are discussed in Section 3.4. Different approaches to develop lane models are developed, tested, and compared in Chapter 4.

The other functionality of the first module is to extract candidate profiles from aerial images. Since profiles are one-pixel wide road segments perpendicular to the travel direction, it implies that the automated process needs to identify the travel direction from aerial images. To meet this need, a special type of aerial images, namely, geo-rectified images, are used in the developed framework. In Section 3.3, an algorithm is introduced to automatically identify the direction of travel for the target road. Other benefits of the algorithm are also discussed in detail in this section.

In the second module, lane profiles are identified from candidate profiles using the lane model built in the first module. After this step, the lane profiles will indicate the possible locations of travel lanes. However, the lane profiles identified are inevitably mixed with inaccurately predictions by the lane model. In order to extract more reliable lane locations, an optimization algorithm is developed in the third module. In Section 3.5, the feasibility of the algorithm is discussed. Besides the candidate lane profiles, the optimization algorithm also includes various constraints. Reliability threshold, width range and road pixel saturation are typical constraints for a reasonable optimization algorithm to follow. Different algorithms to implement the optimization are introduced

and compared in Chapter 5. In general, the multi-objective optimization algorithm applied in the third module aims to maximize the number of lane profiles included in predicted lane locations and their reliability. In the rest of this dissertation, two levels of definitions are used to describe lane locations, as illustrated in Figure 3-3:

1. In a one-pixel wide road profile, a lane location at location  $i$  indicates that there is a lane profile from the  $i$ th pixel to the  $(i+K)$ th pixel of the road profile, where  $K$  is the length of the lane.
2. In a  $W$ -pixel wide by  $M$ -pixel long road region, a lane location at location  $i$  indicates that there is a lane in the rectangular area within  $(0, i)$ ,  $(0, i+K-1)$ ,  $(M, i+K-1)$  and  $(M, i)$ , where  $K$  is the length of the lane.

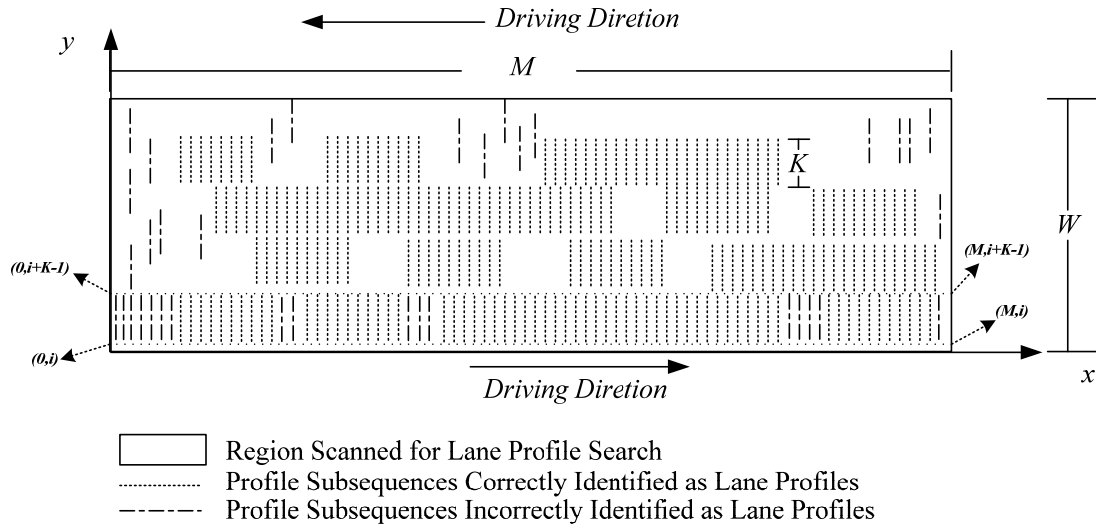


Figure 3-3. An Illustration of Lane Locations for a Scanned Road Region

In the last module, as the output the optimization, number of lanes estimated can be labeled with two confidence levels after comparing to a predetermined confidence threshold. Only the inferences with confidence levels higher than the threshold value will be accepted to update the database. In cases when confidence level is lower than the threshold, the targeted road segments will be recorded and manually checked thereafter. A

higher threshold means the accepted results in the automatic extraction process is more reliable. A lower threshold allows more extracted results to be accepted and reduces the amount of manual work required later. Therefore, the adoption of the threshold value can be considered as a tradeoff between effectiveness and reliability. The impacts of the confidence threshold and the guideline of threshold selection are explained in detail in Chapter 5.

The following sections of this chapter introduce the technique of mapping pixel data to spatial-coordinates for the data preparation module, the approaches to build lane model for lane-profile recognition, and statistical features of lane profiles on the basis of which an optimization algorithm is designed.

### **3.3 Identification of Road Locations in Geo-Rectified Aerial Images**

The coordinates of road segments are critical references in geo-rectified images. To match image pixels and road segments, the projection system of the GIS street layer and the projection system of the images must be consistent. Since the projection of the GIS street layer can be easily manipulated in application software (e.g., ArcGIS), the original projection system of geo-rectified images is adopted for both resources as the standard criterion in this research. Figure 3-4 presents an example of the excerpted information file for an image.

In this research, geo-rectified images of Florida are used. These images are available upon request through the Florida Department of Transportation's (FDOT) Surveying and Mapping Office. Color aerial images of Florida's counties in Multiresolution Seamless Image Database (MrSID) format are available for the year 2006 and later. The resolution for all of these images is one foot, and each of the images is a

5000 by 5000 square feet. With the top-left reference point specified for each image, the pixel locations of each link's start and end points can be easily inferred, as shown in Figure 3-5.

In the step of mapping geo-coordinates to pixel-coordinates, assuming that the resolution of all processing images is identical, a mapping function,  $f$ , is defined as follows:

$$f : A \rightarrow B \quad (3-1)$$

$$A = \left\{ \alpha_i = \begin{bmatrix} Sx_i \\ Sy_i \\ Ex_i \\ Ey_i \end{bmatrix} \mid i \in [1, L], \alpha_i \in R^4 \right\} \quad (3-2)$$

$$B = \left\{ \beta_i = \begin{bmatrix} Sx'_i \\ Sy'_i \\ Ex'_i \\ Ey'_i \\ Seq \end{bmatrix} \mid i \in [1, L], Sx'_i, Ex'_i \in [1, c], Sy'_i, Ey'_i \in [1, r], Seq \in [1, M] \right\} \quad (3-3)$$

where

$r$  = row index of pixels in an image,

$c$  = column index of pixels in an image,

$L$  = number of road segments in GIS street layer,

$M$  = number of images,

$Sx_i, Sy_i, Ex_i, Ey_i$  = geographic coordinates of the start and end points of the  $i$ th road,

$Sx'_i, Sy'_i, Ex'_i, Ey'_i$  = pixel coordinates of the start and end points of the  $i$ th road, and

$Seq$  = index of image in aerial image dataset.

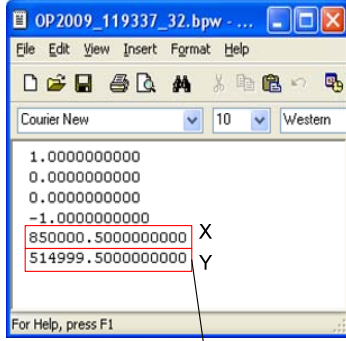
```

<PAMDataset>
<SRS>PROJCS[;NAD83(HARN) / Florida East
(ftUS);,GEOGCS[;NAD83(HARN);,DATUM[;NAD83_High_Accuracy_Regional_Network;,SPHEROID[;GRS_
1980;,6378137.0,298.257222101]],PRIMEM[;Greenwich;,0.0],UNIT[;Degree;,0.0174532925199433]],PROJECT
ION[;Transverse_Mercator;,PARAMETER[;false_easting;,656166.667],PARAMETER[;false_northing;,0.0],PA
RAMETER[;central_meridian;,-81.0],PARAMETER[;scale_factor;,0.999941177],PARAMETER[;latitude_of_ori
gin;,24.33333333333333],UNIT[;Foot_US;,0.3048006096012192]]</SRS>
<Metadata>
<MDI key="PyramidResamplingType">NEAREST</MDI>
</Metadata>
<PAMRasterBand band="1">
<Histograms>
<HistItem>
<HistMin>-0.5</HistMin>
<HistMax>255.5</HistMax>
<BucketCount>256</BucketCount>
<IncludeOutOfRange>0</IncludeOutOfRange>
<Approximate>0</Approximate>
<HistCounts>57194|5868|7254|6560|...</HistCounts>
</HistItem>
</Histograms>
<Metadata>
<MDI key="STATISTICS_MINIMUM">0</MDI>
<MDI key="STATISTICS_MAXIMUM">255</MDI>
<MDI key="STATISTICS_MEAN">86.74719916</MDI>
<MDI key="STATISTICS_STDDEV">32.376305149527</MDI>
</Metadata>
</PAMRasterBand>
<PAMRasterBand band="2">
<Histograms>
<HistItem>
<HistMin>-0.5</HistMin>
<HistMax>255.5</HistMax>
<BucketCount>256</BucketCount>
<IncludeOutOfRange>0</IncludeOutOfRange>
<Approximate>0</Approximate>
<HistCounts>17266|2898|3651|3123|...</HistCounts>
</HistItem>
</Histograms>
<Metadata>
<MDI key="STATISTICS_MINIMUM">0</MDI>
<MDI key="STATISTICS_MAXIMUM">255</MDI>
<MDI key="STATISTICS_MEAN">113.50571316</MDI>
<MDI key="STATISTICS_STDDEV">29.465471728106</MDI>
</Metadata>
</PAMRasterBand>
/PAMDataset>

```

Figure 3-4. An Example of Information File of Image

Coordinates of The Top-Left Reference Point



Coordinates of Road Segment

	FNODE	TNODE	FROM_TO_TO_FROM SN	xstart	ystart	xend	yend	xcentroid	ycentroid
1	2134	2136	210.379 210.379 N AUGUSTA DR	881174.8755	590330.5003	880964.5005	590329.2503	881069.688	210.378713
2	4905	4811	209.292 209.292 NW 79TH CT	876153.063	584995.6253	876066.938	585186.3753	876110.0005	209.291849
3	5084	4905	470.944 470.944 NW 79TH CT	876240.438	584538.6253	876153.063	584995.6253	876222.6799	470.944195
4	5640	5632	241.349 241.349 NW 181ST ST	875429.4379	583333.6878	875188.5629	583348.8128	875309.0004	241.349396
5	5758	5498	611.908 611.908 NW 28TH AVE	877261.2505	583073.9378	877232.8755	583685.1878	877247.063	611.908247
6	5405	6195	2524.49 2524.49 FLORIDA TPKE EXT	865256.5629	583866.6253	863399.3754	582156.6878	864327.9692	2524.486417
7	6364	6366	255.54 255.54 NW 125TH ST	877573.7505	581800.1878	877317.8755	581794.4378	877445.813	255.939598
8	6478	6366	259.331 259.331 NW 28TH AVE	877324.3755	581535.1878	877317.8755	581794.4378	877321.1255	259.331472
9	6831	6738	249.623 249.623 NW 79TH AVE	876215.563	580796.3128	876208.438	581045.6878	876214.0005	249.623028
10	7101	6195	2541.66 2541.66 FLORIDA TPKE EXT	861657.8754	580319.5628	863399.3754	582156.6878	862462.1903	2541.627017
11	7733	7779	1052.76 1052.76 SR 826 EXT	888180.2505	579068.8128	887132.5005	578966.1878	887656.3755	1052.763959
12	7779	7770	2071.42 2071.42 SR 826 EXT	887132.5005	578966.1878	885061.1255	578980.1878	886096.813	2071.422312
13	7791	7795	2021.37 2021.37 SR 826 EXT	881916.6255	578941.5628	879895.3755	578919.8128	880906.0005	2021.367019
14	7795	8289	2309.6 2309.6 SR 826 EXT	879895.3755	578919.8128	878091.438	577688.1878	878812.5883	2309.578729
15	8516	8467	400.692 400.692 TORPHIN PL	881465.7505	576960.5628	881113.3755	577151.3128	881289.563	400.691531
16	10597	10373	546.149 546.149 NW 67TH AVE	883524.3755	570798.6253	883511.6255	571344.6253	883518.0005	546.148846
17	10785	10597	462.319 462.319 NW 67TH AVE	883551.8795	570337.1253	883524.3755	570798.6253	883538.1255	462.318613
18	10942	10854	227.647 227.647 NW 67TH AVE	883574.5005	569875.5003	883563.3755	570102.8753	883568.938	227.646999
19	13656	13618	291.357 291.357 W 66TH ST	884108.6255	563695.9378	883857.5005	563769.4378	884003.178	291.357214

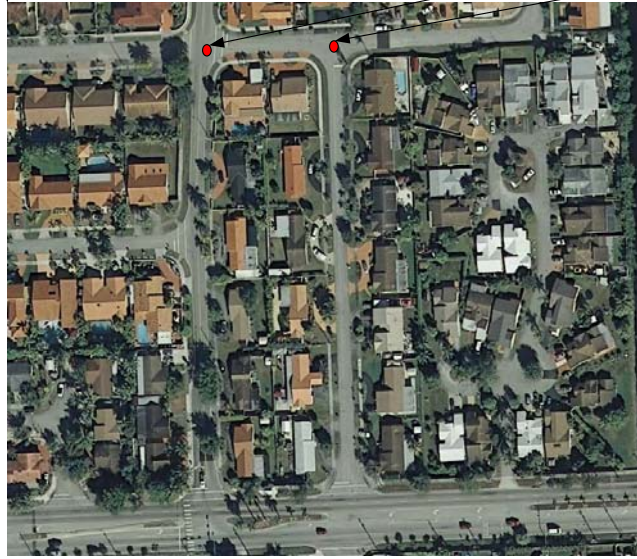


Figure 3-5. Determining the Pixel Locations of a Road Segment

The  $f$  function is a mapping function between  $\alpha_i$  and  $\beta_i$ , which specify the coordinates of a start point and end point in GIS coordinates and in image coordinates, respectively.

### 3.4 Lane-Profile Model and Profile Extraction from Aerial Images

This section presents a profile-based lane model that includes different brightness fluctuation patterns and homogeneity measurement, as well as a technique to extract target profiles from aerial images. Using  $\bar{A}^T$  to represent the transpose of the matrix  $\bar{A}$ , a profile  $\bar{v}$ , defined as a slice of road segment (one-pixel wide) perpendicular to the direction of traffic, is defined as follows:

$$\bar{v} = [e_1, e_2, \dots, e_W]^T \quad (3-4)$$

$$\bar{e}_i = [c_{i1}, c_{i2}, c_{i3}, x, y]^T \quad \text{where } 1 \leq i \leq W \quad (3-5)$$

where

$W$  = width of the road profile,

$c_{i1}, c_{i2}, c_{i3}$  = feature components of the  $i$ th pixel values (e.g., brightness value in Hue-Saturation-Value (HSV) color space), and

$x, y$  = 2-dimension coordinates of pixel in the image.

Generally, a group of adjacent profiles is collected for regional analysis. A region is defined as  $\bar{V} = [\bar{v}_1, \bar{v}_2, \dots, \bar{v}_M]^T$ , where  $M$  is the number of profiles within the region.

Three brightness-based patterns of lane profiles are proposed in this research:

1. Lane profile with light abrasion and lane markings (“U” shape).
2. Lane profile with heavy abrasion and lane markings (“V” shape).
3. Lane profile with light abrasion and no lane markings.

Figure 3-6 shows the image and brightness fluctuations of the three patterns. Pattern 1 represents newly paved and marked lanes, and is the ideal pattern. Patterns 2 and 3 represent marked multi-lane highway and unmarked two-lane highway, respectively. These three patterns are assumed to cover most of the roadways. Based on these patterns, a  $K$ -pixel lane profile can be generally modeled with  $K$  attributes, and each of these attributes represents the brightness value of a lane profile.

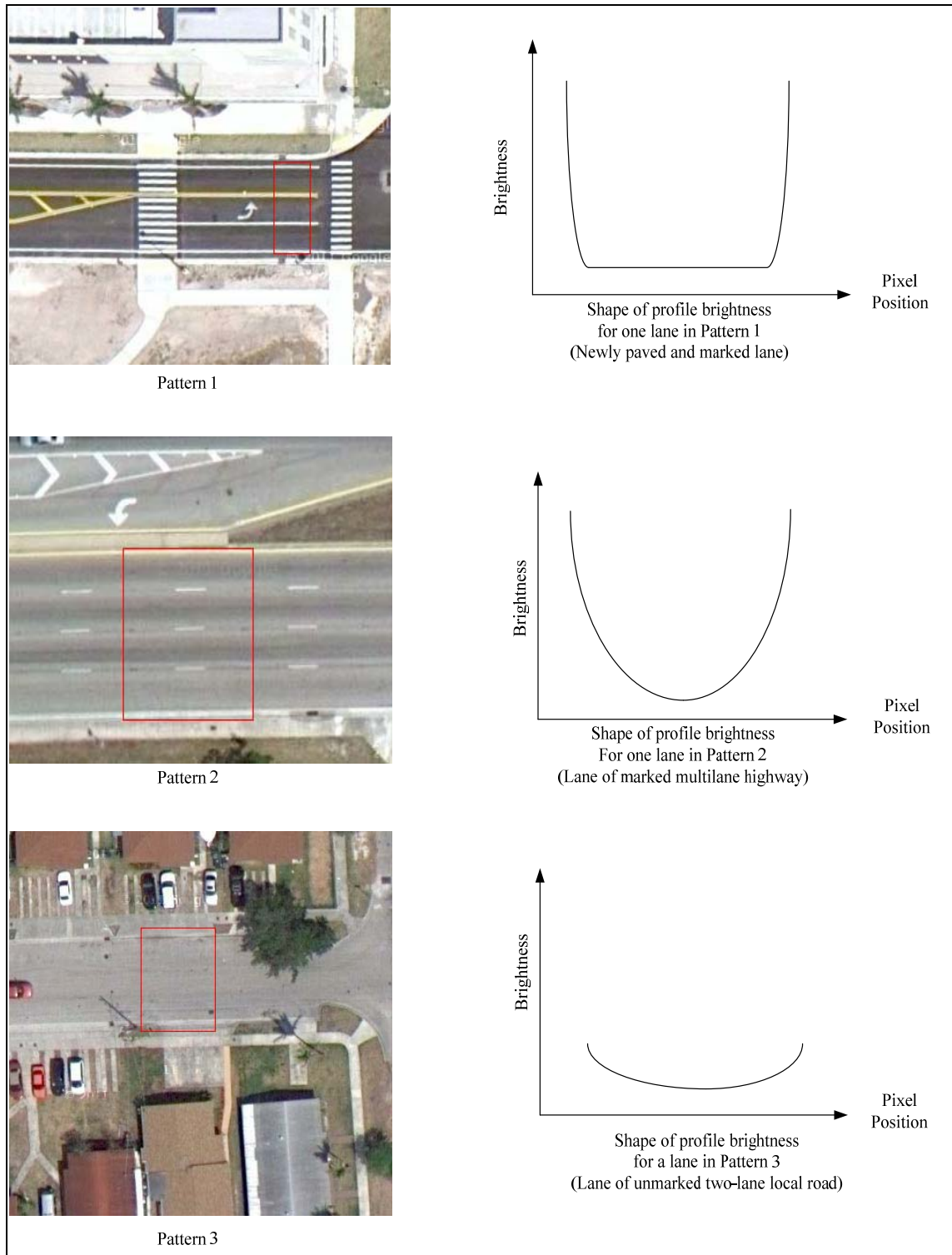


Figure 3-6. Three Lane-profile Patterns and Their Brightness Fluctuations



Homogeneity is another attribute to be considered for use in the lane model. This attribute is designed to measure the degree of homogeneity, and is generally defined as follows:

$$f = \sum_i \sum_j \{p(i, j), d, \theta\}^2 \quad (3-6)$$

where

$f$  = the degree of homogeneity,

$p(i, j)$  = the  $(i, j)$ th normalized gray-tone spatial-dependence matrix,

$d$  = distance between two pixels, and

$\theta$  = degree selected from set  $\{0, 45, 90, 135\}$ .

It has been observed that paved roadways usually show good homogeneity in driving direction. This feature of roadways is included in the proposed lane model. In this research, the homogeneity of a profile is specifically defined as the number of similar pixels along the driving direction. Given a profile subsequence  $\bar{v}_i$ , if  $\bar{v}_{i-1}$  is the profile next to  $\bar{v}_i$ , they can be represented as follows

$$\bar{v}_i = [e_{i1}, e_{i2}, \dots, e_{iw}]^T \quad (3-7)$$

$$\bar{v}_{i-1} = [e_{(i-1)1}, e_{(i-1)2}, \dots, e_{(i-1)w}]^T \quad (3-8)$$

where  $e$  is the brightness value of a pixel.

The homogeneity of  $\bar{v}_i$  is defined as:

$$H = \sum_{j=1}^K S(e_{ij}, e_{(i-1)j}) \quad (3-9)$$

$$S(x, y) = \begin{cases} 0 & \text{if } |x - y| \geq \sigma \\ 1 & \text{if } |x - y| < \sigma \end{cases} \quad (3-10)$$

where  $\sigma$  is a threshold value that describes the maximum distance allowed in determining the similarity of the two pixels.

In essence, the proposed lane models turned the task of recognizing  $K$ -pixel wide lane profile into a typical sequence matching problem for  $(K+1)$  dimension records  $\overline{v_i} = [e_{i1}, e_{i2}, \dots, e_{iK}, H_i]^T$ . The goal of lane profile recognition is to find all  $K$ -pixel wide profile subsequences that match a lane profile model derived from the three patterns specified in Figure 3-6. In order to prepare data for automated recognition using the aforementioned model, candidate profiles must first be extracted from aerial images.

Using the mapping technique described in Equations 3-1 through 3-3, any target road segment is recorded as a line from pixel  $(x_1, y_1)$  to  $(x_2, y_2)$  in an image. Along this line, the  $M$  centers of profiles are determined using the Bresenham line algorithm (Bresenham, 1965). Then, a profile with  $W$  pixels is extracted along the direction perpendicular to the roadway link based on the first point in  $M$ . Data in the profile are then transformed to acquire the brightness and the homogeneity values. After the transformation procedure, all  $K$ -pixel subsequences of the profile are extracted for the lane identification process. When the extraction for this profile is completed, the next point in  $M$  is used to find the next  $W$ -pixel profile. The procedure is repeated until all points in set  $M$  are processed. For an area with  $M \times W$  pixels, the total number of  $K$ -pixel subsequences extracted is  $M \times (W - K + 1)$ . However, depending on the length of the target roadway, the first and the last 25-100 pixels are typically excluded to avoid the pixels that cover the turn-bay areas.

One of the advantages of using geo-rectified images is that the number of profiles to be classified is greatly reduced. Without the information on roadway coordinates,

profile extraction needs to be applied to the entire pixel matrix of the image with  $r \times c$  pixels. Since aerial images usually have several thousand pixels in each dimension, the matrix to be processed will include tens of millions of pixels. In this approach, the area of interest is limited to a matrix with  $M \times W$  pixels. For a 1000-ft road link, with  $W$  set to 120, the size of the matrix to be classified can be less than 1% of the original size.

Figure 3-7 shows an example of road area automatically identified in an aerial image. The rectangular area marked in the image is where the profiles are extracted for lane recognition. Compared to the original size of the image, a very small target area is identified and processed using the aforementioned mapping technique. This technique of precisely identifying the target area also helps to exclude misleading objects such as parking lots (Hu et al., 2004). In Figure 3-6, if the entire parking lot that is located beside the target road is included within the extracted area, it will require significant additional effort to separate it from the actual travel lanes (Zhang and Couloigner 2006).

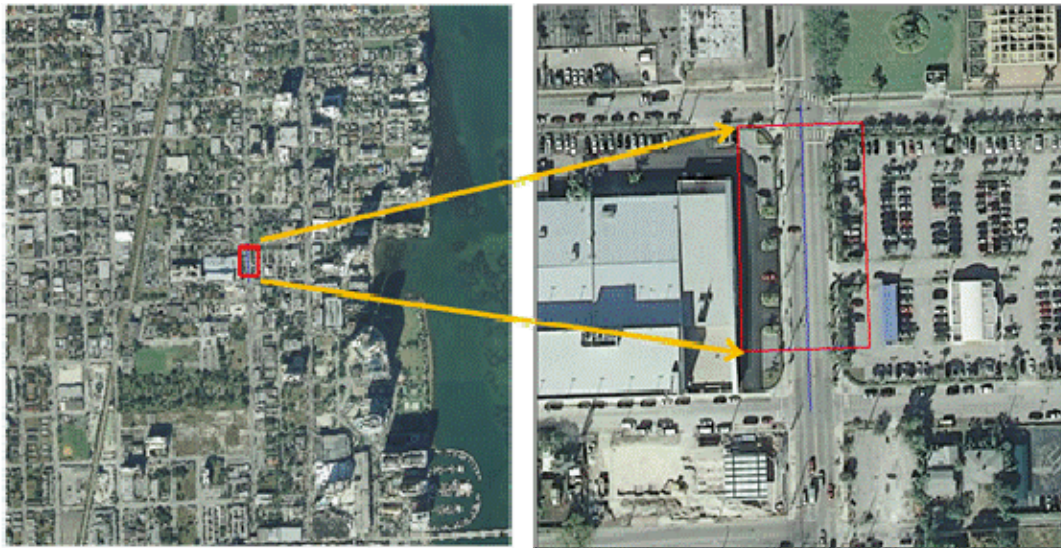


Figure 3-7. An Example of Area Automatically Identified from Aerial Image

In a traditional aerial image pixel matrix with  $r$  rows and  $c$  columns, since the direction of road is unknown,  $W$ -pixel wide profiles need to be extracted along all

directions to find the lane profiles. In theory, there are infinite options for the roadway direction ( $\theta \in \left(0, \frac{\pi}{2}\right)$ ). Although the number of the possible directions for pixel matrix is limited, as only integer coordinates of points are meaningful, the entire  $r \times c$  pixel matrix still needs to be processed for each possible direction. In practice, this approach will greatly increase the number of instances that are incorrectly classified. With geo-rectified images, the probable direction of a target roadway can be estimated. Therefore, profiles need to be extracted and classified in a  $M \times W$  matrix only once. Once the complete lane profiles are extracted and identified, the areas covered by lanes can be estimated. Two types of lane-profile modeling approaches and a test result are provided in Chapter 4.

### 3.5 Estimation of Number of Lanes

The previous two sections mainly discussed the issue of determining possible lane positions in a single profile. Figure 3-8 provides an illustration of possible recognition results for a scanned road region with  $M \times W$  pixels. It shows that, if the accuracy of lane profile recognition is high, large groups of lane profiles that are identified will share the same vertical coordinates and cover the majority of the area included in an actual lane. With different groups of lane profiles, areas of different lanes are expected to be identified. Although subsequences could be mistakenly identified as lane profiles, with high-accuracy lane models, their variance of vertical locations will be high and different from that of true lane profiles.

To estimate number of lanes, image data are reviewed in another direction, namely, the direction of traffic. Since pixels along this direction should provide high homogeneity (Gruen and Li, 1995), statistical features are expected to be captured along

the assumed direction of traffic. Although homogeneity is included in profile features as described in Section 3.3, it is used to measure the similarity between two neighbor profiles. In the lane-location estimation procedure, it is scrutinized in the entire roadway section in order to examine the lane-profile recognition results with the aforementioned statistical feature. In this sense, the algorithm designed for lane-location estimation is actually seeking locations that include the maximum number of lane profiles under certain constraints. Details of the optimization algorithm and the test results are discussed in Chapter 5.

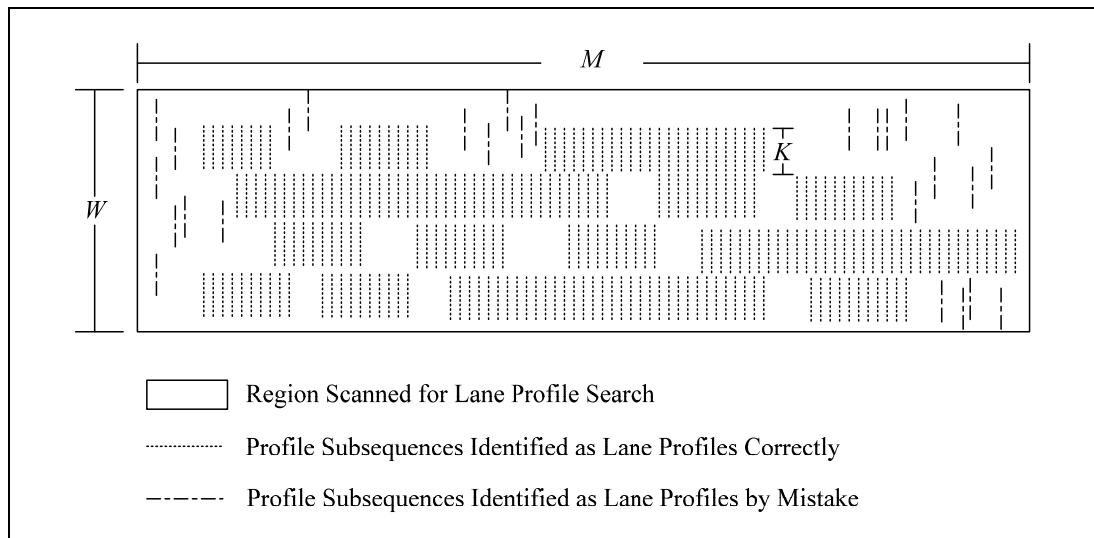


Figure 3-8. An Illustration of Classification Result for a Scanned Road Region

### 3.6 Summary

In this chapter, the methodology framework of this dissertation is presented. Given the geographic coordinates of the target road segments and the geo-rectified aerial images of the target area, the framework automatically extracts the number of lanes for each of the target road segment if the reliability of prediction is higher than a thresholding value. The methodology develops solutions to three critical steps in the automatic

extraction process: the estimation of road position, the identification of travel lanes, and the estimation of number lanes based on identification results.

First, in order to identify approximate road position, the framework adopts geo-rectified aerial images for automatic extraction. A mapping algorithm is developed to precisely estimate the pixel coordinates of road segments based on the geographic information included in aerial images. The algorithm greatly reduces the required number of pixels scanned for lane recognition and increases the accuracy of lane recognition. Second, a profile-based lane model is proposed. Compared to the existing lane models that require the extraction of complete lane area, this model is expected to be more reliable and robust when applied to road segments that are heavily occluded. Different approaches to build the profile-based lane model are presented in Chapter 4. Lastly, the statistical features of the lane recognition results are applied in the estimation of number of lanes. The algorithms for the estimation procedures are presented and discussed in detail in Chapter 5.

## **CHAPTER 4**

### **PROFILE-BASED LANE MODEL**

#### **4.1 Introduction**

The lane-profile models described in Chapter 3 turned the problem of finding lane-profile candidates into one of finding qualified  $(K+1)$ -pixel wide sequences in road profiles. The qualification of these sequences is justified by statistical lane models built on sample data. In this chapter, two types of modeling approaches are introduced, namely the similarity-score based approach and the classifier based approach. The classification-based approach is then applied in two test cases to validate the profile-based lane model presented in this research.

#### **4.2 Similarity-Score Based Approach**

Figure 4-1 illustrates a relatively ideal scenario. The pixel information of each road profile for the area within the rectangle was collected using the Bresenham line drawing algorithm (Bresenham, 1965). This algorithm helps to determine the pixels that should be selected in order to form a close approximation to a straight line between two end points. The entire study area consists of 80 profiles, each of which includes 120 vertical pixels.

According to previous studies on road models (Gruen and Li 1995, Steger 1998, Chi et. al 2009), a profile's brightness information can provide evidence of lane presence. As such, the brightness information of 30 profiles in the middle of the target area is given in Figure 4-2. It can be seen from the figure that the fluctuations in the brightness of all profiles are similar. Each of these seven lanes visible from Figure 4-2 can be mapped to a "V" region. This reflects the fact that lighter abrasion can be discovered in the center of a

lane (Chi et. al, 2009). Due to the presence of lane marking, it is even easier to find the dashed marking lines that bound a particular lane. For pavement with less serious abrasion, a “U” style intensity distribution of the lane profiles can be expected.



Figure 4-1. Studied Road Region (in Rectangle)

Figure 4-2 also shows a visual display of homogeneity along driving direction. The width of lanes 1 through 3 and lanes 5 through 7 can be determined by the pixel distance between peaks, which is 12 in most cases. Since the resolution of an image is 1 ft/pixel, the lane areas concluded also provide reasonable lane width data. Within each lane, for a given profile point, the difference between various profiles is small. Regions outside of the paved area generally have a much greater diversity of intensity distributions. However, bounded homogenous regions are also evident due to the presence of sidewalks or shoulders. Compared to paved lanes, these are more likely to have a bright and flat profile intensity distribution. In this simple scenario, the goal is to



find a model to distinguish a moderately-used and/or marked lane from noise factors such as sidewalks.

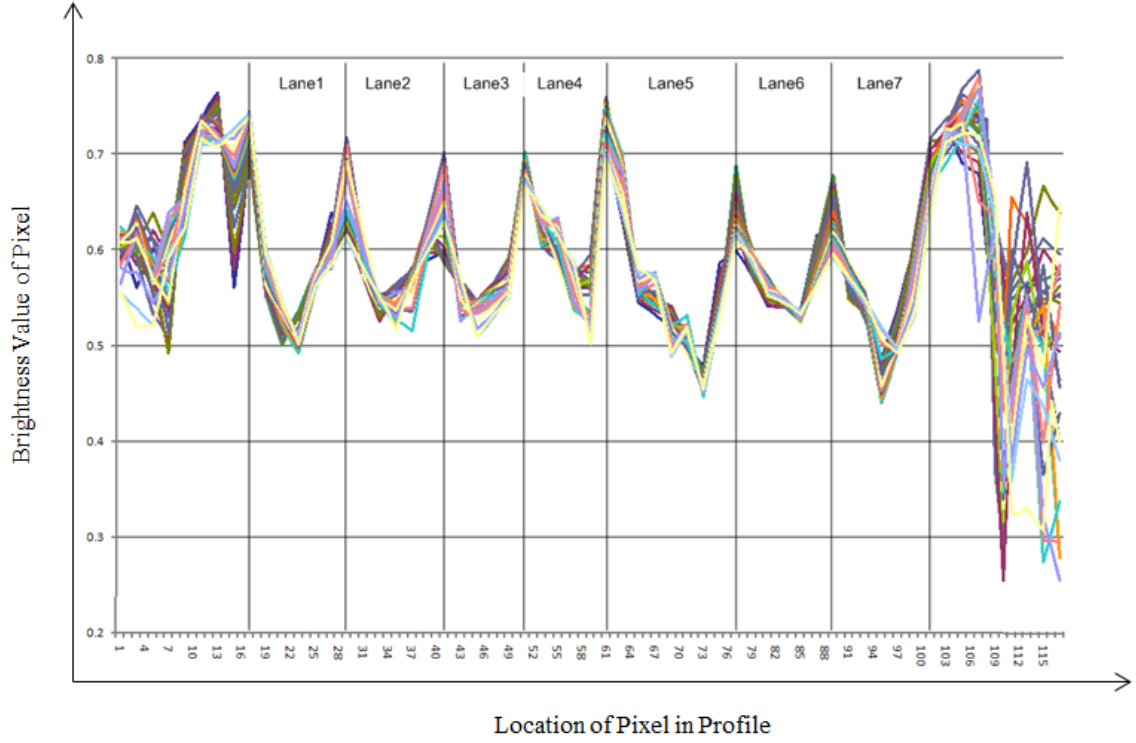


Figure 4-2. Brightness Distributions of 30 Profiles

One of the simplest models would be a reference profile data vector. Using Equations 3-4 and 3-5, a subsequence profile  $\bar{v}_i^{mn}$  of profile data vector  $\bar{v}_i$  can be defined as follows:

$$\bar{v}_i^{mn} = [e_{im}, e_{i(m+1)}, \dots, e_{in}]^T, \quad 1 \leq m \leq n \leq K \quad (4-1)$$

where  $K$  is the width of the subsequence profile.

Assuming that  $\bar{v}_i^{mn}$  always represents a subsequence that is mapped to a lane, by taking the average of  $H$  subsequences (Vosselman and Knecht, 1995; Zhou et al., 2006; Chi et al., 2009), a simple model for lane recognition can be defined as follows:

$$\bar{\gamma}_{mn} = \left[ \sum_{i=1}^H \frac{e_{im}}{H}, \sum_{i=1}^H \frac{e_{i(m+1)}}{H}, \dots, \sum_{i=1}^H \frac{e_{in}}{H} \right]^T, \quad 1 \leq m \leq n \leq K \quad (4-2)$$

For any new profile segment  $\bar{v}_{new}^{mn}$ , the  $(m-n+1)$ -dimensional Euclidean distance can be applied to measure its similarity to reference model  $\bar{\gamma}_{mn}$ . However, the simple average model will not be able to deal with the following two cases where the Euclidean distance is applied:

1. when  $\bar{v}_{new}^{mn}$  and  $\bar{\gamma}_{mn}$  fluctuate in different scales, which means that the degrees of abrasion on the pavement are different, and
2. when  $\bar{v}_{new}^{mn}$  and  $\bar{\gamma}_{mn}$  fluctuate between different maximums and minimums, which means that the brightness of markings and pavements are inconsistent for different road segments.

An alternative definition of  $\bar{\gamma}$  is suggested in this research, which actually normalizes the vector, as follows:

$$\bar{\gamma}_{mn} = \left[ \sum_{i=1}^H \frac{e'_{im}}{H}, \sum_{i=1}^H \frac{e'_{i(m+1)}}{H}, \dots, \sum_{i=1}^H \frac{e'_{in}}{H} \right]^T, \quad 1 \leq m \leq n \leq K \quad (4-3)$$

$$e'_{ij} = \frac{(e_{ij} - \mu(\bar{v}_i^{mn}))}{\sigma(\bar{v}_i^{mn})}, \quad 1 \leq m \leq n \leq K, \quad 1 \leq i \leq H \quad (4-4)$$

where

$\mu(\bar{X})$  = mean of all elements in vector  $\bar{X}$ , and

$\sigma(\bar{X})$  = standard deviation of all elements in vector  $\bar{X}$ .

More generally, denote  $\Phi$  as the similarity measurement function between two vectors:

$$S = \Phi(\bar{X}, \bar{Y}) \quad (4-5)$$

For a new sequence  $\bar{v}_{new}^{mn}$ , the distance between  $\bar{v}_{new}^{mn}$  and  $\bar{\gamma}_{mn}$  is defined as follows:

$$S_i = \Phi(\bar{v}_{new}^{mn}, \bar{\gamma}_{mn}) \quad (4-6)$$

This distance should be bounded by some thresholds to be considered as a lane segment profile. Define the minimum distance of all training sequences to  $\bar{\gamma}_{mn}$ , as follows:

$$S_{\min} = \text{Min} \left( \Phi(\bar{v}_i^{mn}, \bar{\gamma}_{mn}) \right) \quad (4-7)$$

Given proper selection of training sequences, this minimum distance can be considered as a reasonable candidate of threshold distance. In this similarity-score based approach, reference choice and the design of similarity measurements are two key issues to be explored. For example, Least Square Matching (LSM) (Ackermann 1983) is a widely adopted similarity measurement approach. Similarity score-based methods such as LSM are straightforward and easy to understand. Moreover, the processing of candidate profile sequences is computationally efficient. However, this method only selects one reference profile to validate the candidate sequence. In practice, the predictions based on only one reference profile tend to be biased. Furthermore, if a more complex (longer and noisier) profile data sequence model is desired, it is difficult to make the exact match between the subsequences and the reference sequence.

### 4.3 Classification Based Approach

The classification-based approach seeks out rules or functions that will distinguish different categories of data. These rules or functions, known as classifiers, are trained based on a set of data, each of which has a “class” label. Using the training data set, classification algorithms generate classifiers that decide the “class” label of any new data

record. The classification method of sequential data, or time series, has attracted considerable attention from researchers.

Most typical classification techniques can be applied in this research. To prove the feasibility of the classification based approach, three types of classification algorithms were tested in this section, namely, K-nearest neighbor (KNN) (Altman, 1992), logistic regression (Cox, 1958), and Support Vector Machine (SVM) (Vapnik, 1998). The KNN algorithm calculates the distances between the target record and every record in the training data set. The K nearest training records will then vote to decide the “class” label of the target record. Logistic regression is a special type of regression in which the dependent variable is categorical. For a data record requesting a "class" label, the logistic regression model calculates its probabilities of belonging to each class using a logistic function. The label with the highest probability is then assigned to the new data record. The SVM classification uses a more complex methodology called kernel to build classifiers. Details of this type of classification algorithm are provided in the next section.

To prepare the training data set for classifications, the profile definition of Equation 4-1 needs to be extended. In addition to the feature variables defined in Equation 4-1, a class label is added to the feature vector. If the profiles of the lane areas are labeled as “A”, a subsequence mapped to the lane areas can be defined as follows:

$$\bar{v} = [e_1, e_2, \dots, e_K, H, 'A']^T \quad (4-8)$$

where

$K$  = width of subsequence,

$H$  = homogeneity of the subsequence, and

$e_i$  = brightness value of the  $i$ th pixel of the subsequence.

In the test performed in this section, the  $K$  is set to 12.

The training of classifiers also requires the exclusion of specific datasets, which in this research are noises. For the purpose of demonstration, this test case also included sequence data on the left side of Lane 1 and on the right side of Lane 7 in Figure 4-2. These two sets of sequences, labeled as “B” and “C”, respectively, have the same width. Another type of noise caused by plants was also considered in this case (see Figure 4-3). Sequence data with a length of 12 pixels were also collected from this noisy region and marked as “D”.



Figure 4-3. Studied Road Region with Plant Occlusions (in Rectangle)

A summary of the training data sets is provided in Table 4-1. Along with two-fold cross-validation, the modeling results for SVM, logistic regression, and KNN when  $K=1$  are provided in Table 4-2. The algorithm with highest accuracy rate was KNN which had

91.80% of the test records correctly labeled. Logistic regression algorithm correctly predicted 85.79% of the test records which is the lowest among the three algorithms tested. All three classification algorithms yielded satisfying results using the test set with 183 test records. One possible problem of the aforementioned tests is generality. Although classes “B”, “C”, and “D” represent three types of common, unwelcome objects in lane recognition, not all noise factors are included in these three classes. The models for noise factors as reviewed in Chapter 2, provide possible approaches to excluding noises in automatic extraction processes.

Table 4-1 Summary of Training Data

Type Label	Number of Records	Type of Area	Sequence Length
A	46	Lane	12
B	28	Sidewalks with Shadow	
C	30	Sidewalks with Driveways	
D	79	Plants	

Table 4-2 Summary of Classification Results

Algorithm	Correctly Classified Instances	Incorrectly Classified Instances	Accuracy Rate	Notes
One Nearest Neighbor (KNN)	168	15	91.80%	
Logistic Regression	157	26	85.79%	
Support Vector Machine (SVM)	161	22	87.98%	With Linear Kernel

However, in practice, it is not feasible to enumerate all possible noise factors that may appear in aerial images. An ideal lane model should be able to provide features that are able to stand out from the rest of the image contents. In the next test case, a “non-lane” class is chosen as a more general class against the “lane” class classification to show the effectiveness of the profile-based lane model. The SVM classification is applied in the test case due to two reasons: (1) it tends to provide “natural boundaries” between classes

and is less likely to be biased, and (2) the moderate accuracy rate makes better approximation to normal distribution. The second reason is critical in the prediction reliability estimation which will be explained in detail in Chapter 5. Since SVM is used in the following two-class classification test, a brief introduction to SVM is provided below.

#### 4.4 Support Vector Machine (SVM)

In the following test, SVM is used to find complete lane profiles. SVM is a statistical learning technique designed to solve classification and regression problems. Detailed reviews of the technique can be found in (Vapnik, 1998; Burges, C.J.C, 1998; Schölkopf and Smola, 2002). In this section, linear SVM is explained using a binary class problem.

The goal of building a classifier for a training dataset is to find a hyperplane that separates the positive cases from the negative cases. A training set with  $l$  records and two class labels is defined as follows:

$$\{(\bar{x}_1, y_1), (\bar{x}_2, y_2), \dots, (\bar{x}_l, y_l) | \bar{x}_i \in R^d, y_i \in \{-1, 1\}\} \quad (4-9)$$

where

$\bar{x}_i$  = the  $i$ th record in the training data set,

$y_i$  = the class label of the  $i$ th record in the training data set,

$R$  = real number space, and

$R^d$  =  $d$ -dimensional real number space.

In linear SVM separable cases, the desired classifier of the SVM algorithm is the one that maximizes the margins between the hyperplane and positive cases, as well as the margins between the hyperplane and negative cases. The target hyperplane has the following general functional form:

$$f(x) = \langle \omega, x \rangle + b, \omega \in R^d, b \in R \quad (4-10)$$

where  $\langle ., . \rangle$  represents the dot product in  $R^d$ . Using  $\|\omega\|$  to represent the Euclidean norm of  $\omega$ . The problem of finding the target hyperplane can then be formulated as an optimization problem, as follows:

$$\text{Minimize: } \frac{\|\omega\|^2}{2} \quad (4-11)$$

$$\text{Subject to: } y_i (\omega \cdot \bar{x}_i + b) \geq 1, \quad i = 1, 2, \dots, l \quad (4-12)$$

The constraint of Inequality (4-12) is not always feasible for the optimization problem. A general form that introduces a “soft margin” can be found in Vapnik (1995).

SVM applications have increased significantly due to some of the unique features of this technique. Unlike other classification methods, such as the rule-based classifier and artificial neural network (which tend to find local optimum), SVM formulates a classifier searching task as a convex optimization problem that has only one unique minimum. Moreover, SVM also makes a breakthrough in the effort of solving a well-known problem: the curse of dimensionality. In practice, the direct impact of this phenomenon is that the required number of training data increases significantly as the dimension of dataset increases. In nonlinear boundary cases, the vector  $\bar{x}_i$  in the constraint of Inequality (4-12) is simply replaced with  $\Phi(\bar{x}_i)$  which represents a map of  $\bar{x}_i$  to a higher dimension in order to make the boundary linear (e.g., with map function  $f: \Phi(x_1, x_2) \rightarrow (x_1^2, x_2^2, \sqrt{2}x_1, \sqrt{2}x_2, 1)$ ). Since the similarity can be computed in transformed space using the original data, this mapping function actually converts the



dimension of the classification problem from  $l$  into the dimension of  $\Phi$ . In SVM, the mapping function,  $\Phi$ , is usually referred to as a kernel function.

The dimension of a profile dataset is related to the resolution of aerial images. Typically, for aerial images with a 1 ft/pixel, the dimension of the profile dataset must be more than 10 to cover one lane. Using SVM with the polynomial kernel function, the dimension of dataset is then decreased to the length of the mapped vector (e.g., if  $\Phi:(x_1, x_2) \rightarrow (x_1^2, x_2^2, \sqrt{2}x_1, \sqrt{2}x_2, 1)$ , and the similarity computation is based on 5-dimension vectors). Subsequently, lesser amounts of data are required to build a classifier.

#### 4.5 General Two-Class Classification Based on Profile Lane Model

In this section, 490 instances of 12-pixel wide profile subsequences are extracted to validate a lane-profile model using SVM. Among these instances, 299 are classified as different types of complete lane profiles (labeled “A”), and the rest are random profiles collected from pixels belonging to objects around roadways or the pixels crossing two lanes (labeled “B”). Each data record consists of 13 attributes representing the 12 brightness values and the homogeneity value (H) of the corresponding subsequence, respectively. Since the 12-pixel width is used, the H value varies between 0 and 12 based on the definition in Equations 3-9 and 3-10. A higher H value generally means that the profile and its adjacent profile are more closely matched along the direction of the roadway. The format of instances is defined as  $\bar{v}_i = [e_{i1}, e_{i2}, \dots, e_{i13}]^T$ , where  $e_{i1}, e_{i2}, \dots, e_{i12}$  are the brightness values of a profile subsequence with 12 consecutive pixels, and  $e_{i13}$  is the homogeneity of the subsequence. All 13 features are normalized

using the following equation to reduce the impact of outliers before applying SVM for classifier training:

$$e'_{ij} = \frac{[e_{ij} - \mu(\bar{v}_i^m)]}{\sigma(\bar{v}_i^m)}, \quad 1 \leq m \leq 13, 1 \leq i \leq 490 \quad (4-12)$$

where  $\mu(x)$  and  $\sigma(x)$  are the mean and standard deviation of vector  $x$ , respectively.

The polynomial function kernel and Radial Basis Function (RBF) kernel are tested for the test instances. The polynomial kernel was tested five times, with the exponent set to be 1 to 5, and the RBF kernel was tested three times, with  $\gamma=0.001$ , 0.01, and 1.

The five-folder cross-validation method was used to evaluate the performance of the two kernel functions. In other words, instances included in this test were randomly divided into five partitions. For each run, four partitions were used to train support vectors, and the fifth partition is used to test. This process was repeated five times until each partition was tested. Cross-validation is a widely used method for prediction error estimation. As the partition number increases, the validation bias decreases, while the variance increases (Hastie et. al, 2009). In practice, it is generally recommended that either a ten-folder or five-folder be used as a compromising choice (Breiman and Spector, 1992; Kohavi, 1995).

Table 4-3 provides a summary of the classification statistics. The polynomial kernel with exponent  $d=1$  to 5 and RBF kernel with  $\gamma=0.001$ , 0.01 and 1 provide accuracy data ranging from 76.73% to 95.92%. The performance of the polynomial kernel improves with higher order exponents. With  $d=5$ , the polynomial kernel achieved the highest accuracy rate by correctly classifying 95.92% of the instances. The Kappa

statistic included in Table 4-3 shows how close the prediction of classifier is to the actual class. The closer it is to 1, the more likely the prediction will be consistent with the true class. According to this indicator, the RBF kernel with  $\gamma=0.01$  yields the lowest probability of accurate prediction, while the polynomial kernel with  $d=5$  yields the highest probability.

Table 4-3 Summary of SVM Classification Results for Feasibility Test

	Parameter Setting	Correctly Classified Instances	Kappa Statistic
Polynomial Kernel* $(x_i, x_j) = (\langle x_i, x_j \rangle)^d$	d=1	392 (80.00%)	0.5998
	d=2	422 (86.12%)	0.7216
	d=3	456 (93.06%)	0.8607
	d=4	465 (94.90%)	0.8977
	d=5	470 (95.92%)	0.918
RBF Kernel $(x_i, x_j) = \exp(-\gamma \ x_i - x_j\ ^2)$	$\gamma=0.01$	376 (76.73%)	0.5418
	$\gamma=0.10$	401 (81.84%)	0.6366
	$\gamma=1.00$	447 (91.23%)	0.8241

\* Low order terms are not used

Table 4-4 provides detailed information on the accuracy by class. In this table, the True Positive (TP) rate represents the proportion of instances classified as class x among all instances that are truly in class x. In other words, it gives the proportion of instances in class x that is successfully recognized by the classifier. The False Positive (FP) rate represents the proportion of instances that are classified as class x but are actually not among all instances that are not in class x. It indicates the percentage of instances in other classes that are labeled as x by error. Finally, the Precision represents the rate of instances that are classified as x and are truly in class x.

The statistics for class “A” recognition are a substantial part of this research, as the main target is to recognize class “A” (i.e., complete lane profile). Given that 91.3% of instances in class “A” were successfully recognized, the RBF kernel with  $\gamma=0.01$  actually yielded a very good performance based on the TP rate. However, only 69.0% of the

instances that are classified as “A” can be trusted in this test case. This means that when  $\gamma$  is small, although the RBF kernel can capture most of the complete lane profiles, a great deal of noise is also included. When  $\gamma=0.1$ , the percentage of complete lane profiles being captured dropped, but the precision of the classifier increased from 0.69 (when  $\gamma=0.01$ ) to 0.787 (when  $\gamma=0.1$ ).

Table 4-4 Detailed Accuracy by Class

	Parameter Setting	Class Set	True Positive (TP) Rate	False Positive (FP) Rate	Precision
Polynomial Kernel* $(x_i, x_j) = (\langle x_i, x_j \rangle)^d$	d=1	A	0.817	0.215	0.770
		B	0.785	0.183	0.830
	d=2	A	0.860	0.138	0.845
		B	0.862	0.140	0.875
	d=3	A	0.930	0.069	0.922
		B	0.931	0.070	0.938
	d=4	A	0.956	0.057	0.936
		B	0.943	0.044	0.961
	d=5	A	0.956	0.038	0.956
		B	0.962	0.044	0.962
RBF Kernel $(x_i, x_j) = \exp(-\gamma \ x_i - x_j\ ^2)$	$\gamma=0.01$	A	0.913	0.360	0.690
		B	0.640	0.087	0.893
	$\gamma=0.10$	A	0.838	0.199	0.787
		B	0.801	0.162	0.850
	$\gamma=1.00$	A	0.921	0.096	0.894
		B	0.904	0.079	0.929

\* Low order terms are not used

All three indicators show that the performance of the polynomial kernel improves when  $d$  increased. The Precision of this kernel in the worst case scenario is 77.0%, which is obviously better than the lowest Precision achieved by RBF kernel. It must also be noted that, when  $d$  reaches 3, the performance only improves slightly with the exponent. The test statistics show that when  $d=8$  (which is not provided in the table), the TP rate dropped to 88.6%, although the precision increased slightly to 96.7%. This implies that

the polynomial kernel with  $d=3$  had already set a reasonable boundary between these two classes, with the 13-feature lane-profile model.

A comparison of precision data between class “A” and class “B” shows that the TP rate of class “B” is either roughly equal to or lower than the TP rate of class “A.” However, the precision indicators of class “B” are always higher than those of class “A.” This means that although a lower percentage of instances in class “B” is recognized, the classification results of class “B” are more reliable.

#### **4.6 Summary**

This chapter presented two approaches to building a profile-based lane model. It was concluded that the classification-based approach should be adopted in this research. Two test cases are provided in this chapter to validate the classifier-based lane model. In the first test case, three types of noise and lane profiles were included. KNN, logistic regression, and SVM classification techniques all resulted in promising accuracy rates. In the second test case, the SVM classification method was tested to recognize lane profiles from non-lane profiles based on a 13-feature lane profile model and a dataset with 490 instances. It showed that using the polynomial kernel and the RBF kernel techniques, a prediction precision of about 90% can be achieved. This test was validated by the five-folder cross-validation method. It was found that with  $\text{exponent}=3$ , the polynomial kernel yielded good precision, and further increase of exponents only provided a slight improvement in performance. The RBF kernel with  $\gamma=1.0$  also achieved a precision of about 90%.

In conclusion, the framework in the classification-based modeling approach for identifying profile-based lane profiles from aerial images generated promising results. In

the second test dataset used in this chapter, candidate profiles were divided into only two categories, namely, complete lane profile and non-lane profile. The category of non-lane profile was inclusive of all noises from vegetation, vehicles, or buildings. Compared to the four-class classification used in the first test, a lower accuracy rate was achieved when a classifier was applied to the two-class test data. However, predictions based on natural boundaries between lane-profiles and all other profiles were found to be more reliable and less likely to be biased.

## CHAPTER 5

### ESTIMATION OF NUMBER OF LANES

#### 5.1 Introduction

In the previous chapter, profile-based lane models were developed in order to identify candidate lane profiles from aerial images. The predictions of lane models indicate the possible locations of lanes in a road segment. To identify the actual lane locations from the predictions and eventually estimate the number of lanes, incorrect candidate lane locations need to be excluded. Figure 5-1, reproduced here from Figure 3-3, provides an illustration of the recognition results after applying the profile-based lane model in a road region with  $M \times W$  pixels.

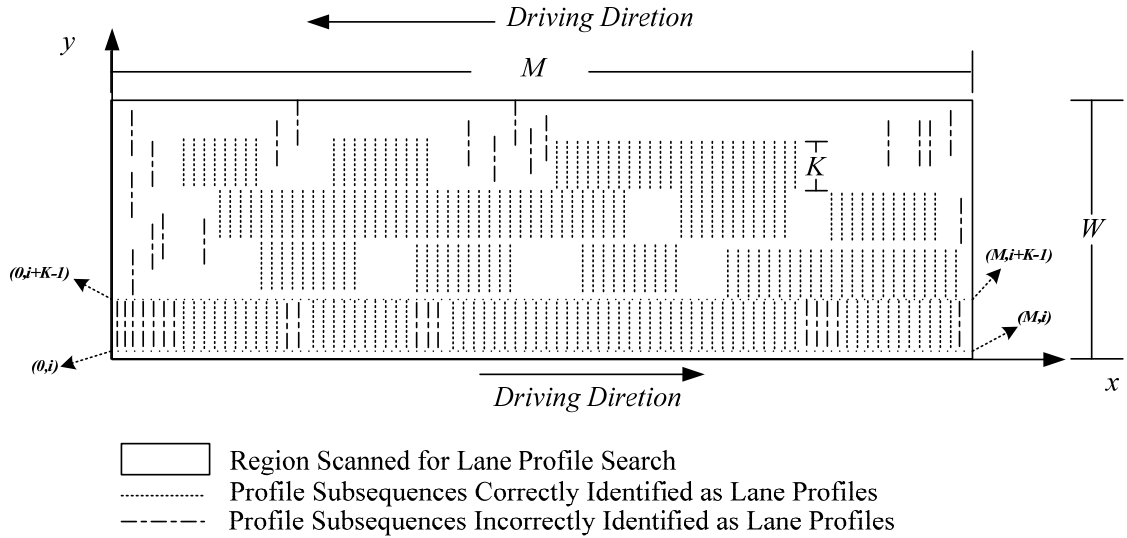


Figure 5-1. An Illustration of Classification Result for a Scanned Road Region

For the convenience of discussion, the bottom-left point of the roadway is referred to as the origin. Assuming the target lane profile length is  $K$ , the figure shows that if a lane exists in the rectangular area within  $(0, i)$ ,  $(0, i+K-1)$ ,  $(M, i+K-1)$  and  $(M, i)$ , the majority of the profiles included in this area should be identified as lane profiles. This

pattern makes it possible to optimize the locations of lanes and eventually estimate the number of lanes. This chapter presents two optimization algorithms for identifying number of lanes based on lane-profile recognition results.

## **5.2 Saturation and Vegetation Index Thresholding**

The optimization module gathers all of the classified  $K$ -pixel profiles as inputs to determining the number of lanes for a road segment. Theoretically, all  $K$ -pixel profiles can be classified using the profile-based lane model directly. However, it was found that the saturation and vegetation threshold could efficiently exclude profiles of plants, vehicles, and building roofs.

Saturation is a measurement used in color space to describe the dominance of hue. Generally, pixels of paved lane areas do not appear to be colorful. Therefore, they typically have a very low saturation degree. Figure 5-2 provides the saturation distributions of 200 lane pixels and non-lane pixels extracted from aerial images. As shown in the figure, the difference between these two types of pixels is evident. In this research, saturation thresholding was used as a reinforcement of classification procedure in order to boost the performance. The saturation threshold for all profiles is 0.1. This means that if the maximum saturation of a profile's 12 pixels is greater than 0.1, the profile is excluded from the classification procedure.

In the current literature, there are two types of formulas that are widely used for vegetation index measurements, namely, the Normalized Difference Vegetation Index (NDVI) and the Greenness Index (VIg/r). These indices have been defined in Equations 2-6 and 2-7, respectively. Since the aerial images used in this research are in RGB color space and no infra-red component is available, the VIg/r index was applied to the



thresholding process, as was done in some road extraction studies (Baltsavias and Zhang, 2005 and Jin et al., 2009). The threshold value of  $V_{lg/r}$  was set to 0, as suggested by Grote et al. (2007).

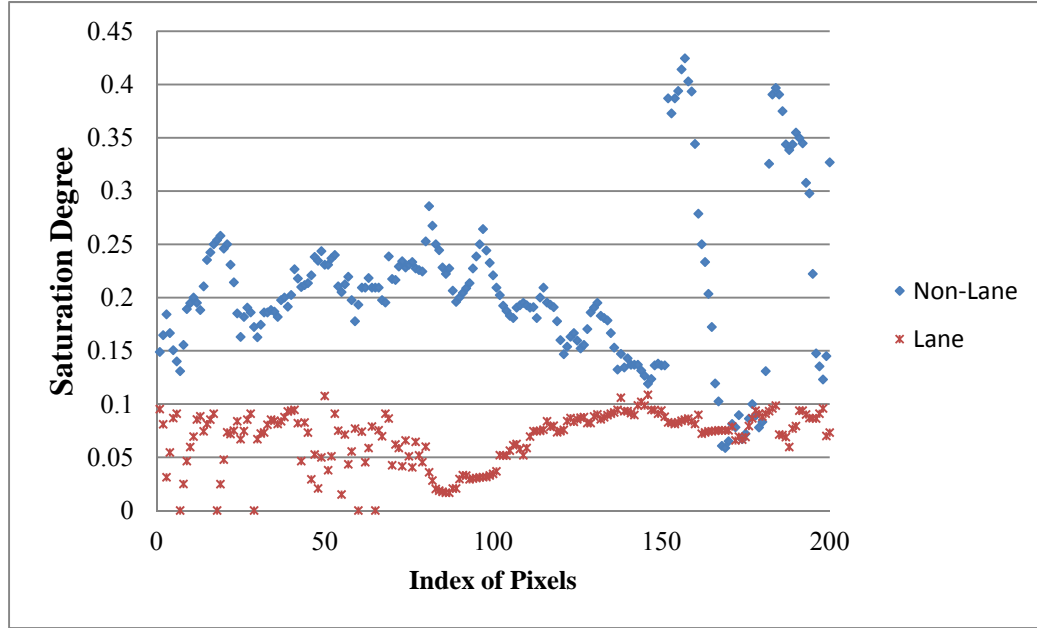


Figure 5-2. Saturation Distributions of Lane and Non-Lane Pixels

The purpose of bringing the two types of thresholds is twofold. First, the thresholds exclude the amount of candidate profiles input into classifiers to improve the identification accuracy. Second, excluding non-lane profiles that appear in lane areas lessens the discrepancy between the expected number of lane profiles and the lane profiles that are actually identified. The latter is critical to increasing the credibility of the lane candidates detected. The benefits will be explained in more detail in Section 5.4.

### 5.3 Fixed lane-length Optimization Algorithm for Lane Location Identification

After excluding the profiles that exceed the saturation threshold, the recognition of lane profile candidates for the remaining profiles becomes a task of typical binary classification, when the general two-class classification model specified in Chapter 4 is

adopted. These  $K$ -pixel profiles are either labeled “lane profile” or “non-lane profile.” The goal of this process is to recognize and mark all the possible lane profiles along the  $W$ -pixel road profiles. The width of a profile ( $W$ ) was determined in advance based on the resolution of aerial images. This profile width represents the reasonable width of a roadway link. Although the width is always unknown,  $W$  should generally be greater than the width of two lanes, but less than that of 15 lanes.

To better understand the above process, in the following description of the optimization process, a lane profile that covers the  $i$ th pixel through the  $(i+K-1)$ th pixel in a one-pixel wide image profile is referred as a lane profile at location  $i$ . Furthermore, a lane included in the rectangular area defined by  $(0, i)$ ,  $(0, i+K-1)$ ,  $(M, i+K-1)$  and  $(M, i)$  is referred to as a lane at location  $i$ .

Along the direction of a road profile, for each pixel location  $i$ , define:

$$C_i : \text{number of lane profiles identified by classifier} \quad (5-1)$$

$$w_i : \text{number of profiles input into classifier} \quad (5-2)$$

where  $1 \leq i \leq (W - K + 1)$ ,  $0 \leq C_i \leq M$ ,  $0 \leq w_i \leq M$ .

Different locations typically have different  $w_i$  values because some of the profiles are excluded through saturation thresholding (i.e., when the maximum saturation of pixels is greater than 0.1). In the case of  $w_i = M$ , the minimum  $C_i$  needed to support the existence of a lane can be estimated based on the accuracy rate of the classifier. In a general two-class classification, assuming that the probability that a classifier identifies a non-lane profile correctly is a constant  $p$  (for a location  $i$  where there is no lane), which is neither close to 1 nor to 0, the probability that the random variable  $C_i$  equals or less than

$X$  can be estimated using the normal distribution density function, defined as follows, when  $w_i$  is large enough:

$$f(X) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(X - \mu)^2}{2\sigma^2}\right) \quad (5-3)$$

where  $\mu = w_i \times (1 - p)$ ,  $\sigma = \sqrt{w_i \times p \times (1 - p)}$ . In practice,  $w_i$  generally increases with the length of the target road segment. In high resolution aerial images, the lengths of road segments in pixels are typically long enough for the normal distribution approximation.

If there is not a lane at location  $i$ , the probability that  $C_i \leq C_{im}$  is  $P\{X \leq C_{im}\}$ . In other words, if more than  $C_{im}$  profiles are identified as non-lane profiles, in confidence level  $P\{X \leq C_{im}\}$ , there should not be a lane at location  $i$ . Otherwise, the location is considered a candidate lane location. In the following steps of the lane identification module, only the locations with more than  $(w_i - C_{im})$  identified lane profiles are used to estimate the lane locations.

The true locations of travel lanes should include most of the lane profiles while keeping a reasonable lane width. In this framework, finding the locations of lanes based on the identified lane profiles is formulated as an optimization problem, defined as follows:

$$\text{Maximize: } f(L) = \sum_{i=L_1}^{i=L_r} u_{p_i} \quad (5-4)$$

$$\text{Subject to: } \exists \theta_n \in \theta \quad (5-5)$$

$$\forall p_i, p_j, i \neq j \Leftrightarrow |p_i - p_j| \geq \theta_n$$

$$\forall p_i, 1 < i < L_r \Leftrightarrow |p_i - p_{i-1}| \leq 3\theta_n, |p_i - p_{i+1}| \leq 3\theta_n$$

$$\forall p_i, \exists p_j, i \neq j, |p_i - p_j| = \theta_n$$

$$\forall p_i \text{ and } p_j, i \neq j \Leftrightarrow p_i \neq p_j$$

where

$$u_{p_i} \in \{\alpha_i C_i \mid 1 \leq i \leq (W - K + 1)\},$$

$$1 \leq p_i \leq (W - K + 1),$$

$$1 \leq r \leq \left\lfloor \frac{W}{K} \right\rfloor,$$

$W$  = number of pixels in the road profile,

$K$  = width of the target lane profile to be identified by lane model,

$C_i$  = number of lane profiles identified at location  $i$  of road profile,

$u_i$  = adjusted number of lane profiles detected at location  $i$ ,

$L_i$  = the location of the  $i$ th lane identified in the optimization process,

$r$  = number of lane locations identified in the optimization process,

$p_i$  = index of the  $i$ th pixel in a  $W$ -pixel wide road profile,

$\theta_n$  = the  $n$ th lane width value in the possible lane width set  $\theta$ , and

$\alpha_i$  = adjustment factor based on lane width and difference of  $C_i$  to  $C_{i\pm1}$  and  $C_{i\pm2}$ .

The set of possible lane width values based on general guidance for geometric design includes eight integers between 9 and 16 (AASHTO, 2011). The constraints included in Equation 5-5 specify that if a lane width  $\theta_n$  is adopted in the optimized scenario, then:

- For any two predicted lane locations, the distance between them must be no less than  $\theta_n$  ft;

- For any predicted lane location, its distance to any adjacent lane should be less than  $3\theta_n$  ft; and
- For any predicted lane location, there must be another lane  $\theta_n$  ft away from it.

These three constraints were included in order to exclude the cases of turn bays and limit the searching space for optimization. When  $L_r = L_0$  that maximizes  $f(L_r)$  in Equation 5-4 is found, the corresponding  $L_0$  is used as the number of lanes for the target road segment. It is noted that this algorithm assumes that the lanes to be detected share the identical width.

#### **5.4 Implementation of Fixed-Width Optimization and Evaluation**

To evaluate the performance of the developed algorithm, the number of lanes is extracted from real-world aerial images using the MATLAB and Weka (Waikato Environment for Knowledge Analysis) software.

First, two-class SVM classifiers were trained using the lane model. The training data was presented in Chapter 4. As discussed earlier, a polynomial kernel of SVM with  $d=2$  was used to avoid the over-fitting problem.

To prepare the data for the road extraction module, geo-rectified aerial images of Miami-Dade County, Florida were downloaded from the FDOT Surveying and Mapping Office. The resolution of these images is 1 ft/pixel and is identical to the resolution of images from which training data are extracted. The length of the profile to be extracted ( $K$  value) was set to 12. The projection system of the Miami-Dade County street layer was then converted in ArcGIS to make it consistent with the projection that is used in aerial images, i.e., Florida State Plane East NAD83.

Together with GIS street layer and geo-rectified aerial images, a MATLAB program was developed to accomplish the tasks specified in the road profile extraction and lane identification modules. A total of six two-lane and four-lane roadways were tested. Figure 5-3 provides the locations of each test roadway link. All of the test roadway links are located in suburban or urban areas with moderate traffic volumes. Occlusion from shadows and buildings existed, but did not affect the extraction of lane profiles in the majority of the paved areas. The link number assigned to each test road is a unique ID assigned to the corresponding road in the Miami-Dade County GIS street layer. When the number of lanes is automatically identified for a test roadway, the link ID is used to manually locate the roadway in the geo-rectified image in the GIS system. The identified information can then be validated quickly by checking the corresponding area in the aerial image.

For each of these roadways, 12-pixel profiles were first recursively retrieved from images for each point of road link indicated and identified in the GIS street layer. The maximum profile width ( $W$  value) extracted for classification was set to be 150 pixels (i.e., 150 ft), which was wide enough to cover a ten-lane roadway section. Depending on the length of the target road, the first and the last 25-100 pixels were excluded in order to avoid the turn-bay areas.

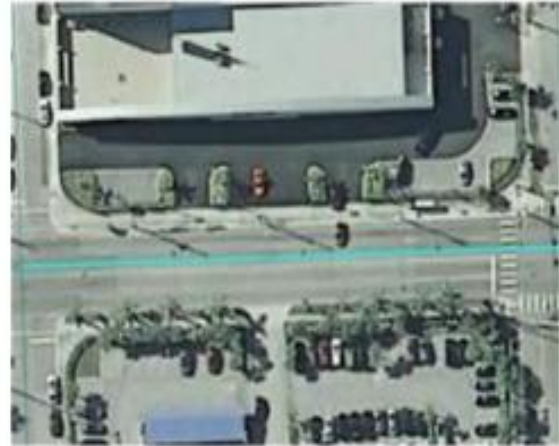
The extracted fixed-width profiles were tested using a 0.1 minimum saturation threshold. Profiles exceeding the saturation threshold were removed. The remaining profiles are then classified by SVM and labeled “lane” or “not a lane” in Weka. The lane profile candidates successfully identified by the SVM were optimized using the MATLAB program developed for the test. Specifically,  $P\{X \leq C_{im}\}$  was set to 99.7% to

estimate the required minimum number of identified lane profiles for each location  $i$  in the program.

Link 54081 Four-Lane



Link 41997 Four-Lane



Link 54369 Four-Lane



Link 58514 Two-Lane



Link 41990 Four-Lane



Link 58510 Two-Lane



Figure 5-3. Aerial Images of Test Roadway Segments



Table 5-1 provides the optimization results of the number of lanes and lane information. Based on the coordinates provided in Figure 5-1, a lane identified at location  $i$  means that a lane is identified in the rectangular area  $(0, i), (0, i+11), (M, i+11)$  and  $(M, i)$ . In the table, the column entitled “Adjusted Support Lane Profiles” specifies the adjusted number of lane profiles identified by the SVM classifier at location  $i$ .

Table 5-1 Optimization Test Results

Link ID	Number of Actual Lanes	Identified Number of Lanes	Identified Lane Width (ft)	Lane Location (i = 1 to 150)	Adjusted Support Lane Profiles ( $a_i \times C_i$ )
Link 54081	4	4	12	49	195
				61	74
				79	58
				91	312
Link 54369	4	4	12	49	125
				61	315
				81	205
				93	489
Link 41990	4	4	12	49	68
				61	269
				75	198
				87	53
Link 41997	4	4	12	47	91
				59	335
				71	172
				83	68
Link 58514	2	2	12	61	117
				73	87
Link 58510	2	2	13	58	122
				71	114

In all six test cases, the automated program yielded the correct number of lanes. The test results also show that the automatic identification method provides approximate locations of the lane, as well as a reasonable lane width. The automated extraction method successfully handled the presence of noises, such as sidewalks, shadows, and plants in the test images. For segments with turn bays in the center (e.g., Link 54081 and

Link 54369), the method successfully rejected the existence of extra lane in the center. Instead, an empty space was left in between the two predicted center lanes.

### **5.5 Flexible Lane-Width Optimization Algorithm**

Tests performed with the fixed lane-width optimization in the previous section showed that it was feasible to extract number of lanes by maximizing the number of lane profiles included at lane locations with an appropriate fixed lane width. This algorithm was designed to find the optimal lane width in order to identify lane locations that maximize the number of lane profiles. The problem with fixed lane-width optimization is threefold:

1. Mapping between geographic coordinates to pixel locations may not be precise enough to locate the actual lane locations.
2. The algorithm is likely to include more lanes of a lower reliability.
3. The overall reliability of predicted lane locations has no impact on the optimization results.

Mapping a continuous geographic coordinate pair to a discrete pixel coordinate pair was an approximation approach. The actual locations of lanes could be shifted or rotated using the mapping technique described in Chapter 3. Given that the Bresenham algorithm, which was applied to the profile extraction process, was also an approximation approach, some discrepancy between the actual and calculated lane locations is inevitable in the automated process. Consequently, the fixed lane-width algorithm could fail to locate an adjacent lane if its distance to the existing predicted lane is not equal to  $\theta_{max}$ , which maximizes the objective function in Equation 5-4.

Second, the minimum number of lane profiles used in the fixed lane-width optimization was calculated based on the expected error lane profiles predicted at the non-lane locations using the formula defined in Equation 5-3. In practice, this threshold is too small. The algorithm is likely to include sidewalks or driveways, both of which may provide a smaller  $C_i$ .

Lastly, this algorithm maximizes the number of lane profiles included in the optimal solution; however, it does not include the reliability of the prediction. For example, for two road segments with different lengths that have identical amounts of lane profiles detected, the reliability of prediction for the shorter segment is evidently higher. Furthermore, an objective function that only aims to maximize lane profiles gathered in the final results tends to overestimate the number of lanes.

To address the above issues, a flexible lane-width optimization algorithm was developed. The workflow of the algorithm is described in Figure 5-4. In this algorithm, the distance between predicted lanes is flexible within a reasonable width. A function was also included in the algorithm to measure the reliability of predicted lanes. The objectives of the new optimization algorithm are to maximize the support from lane profiles and maximize the prediction reliability. The rest of this section introduces each step of the algorithm in detail.

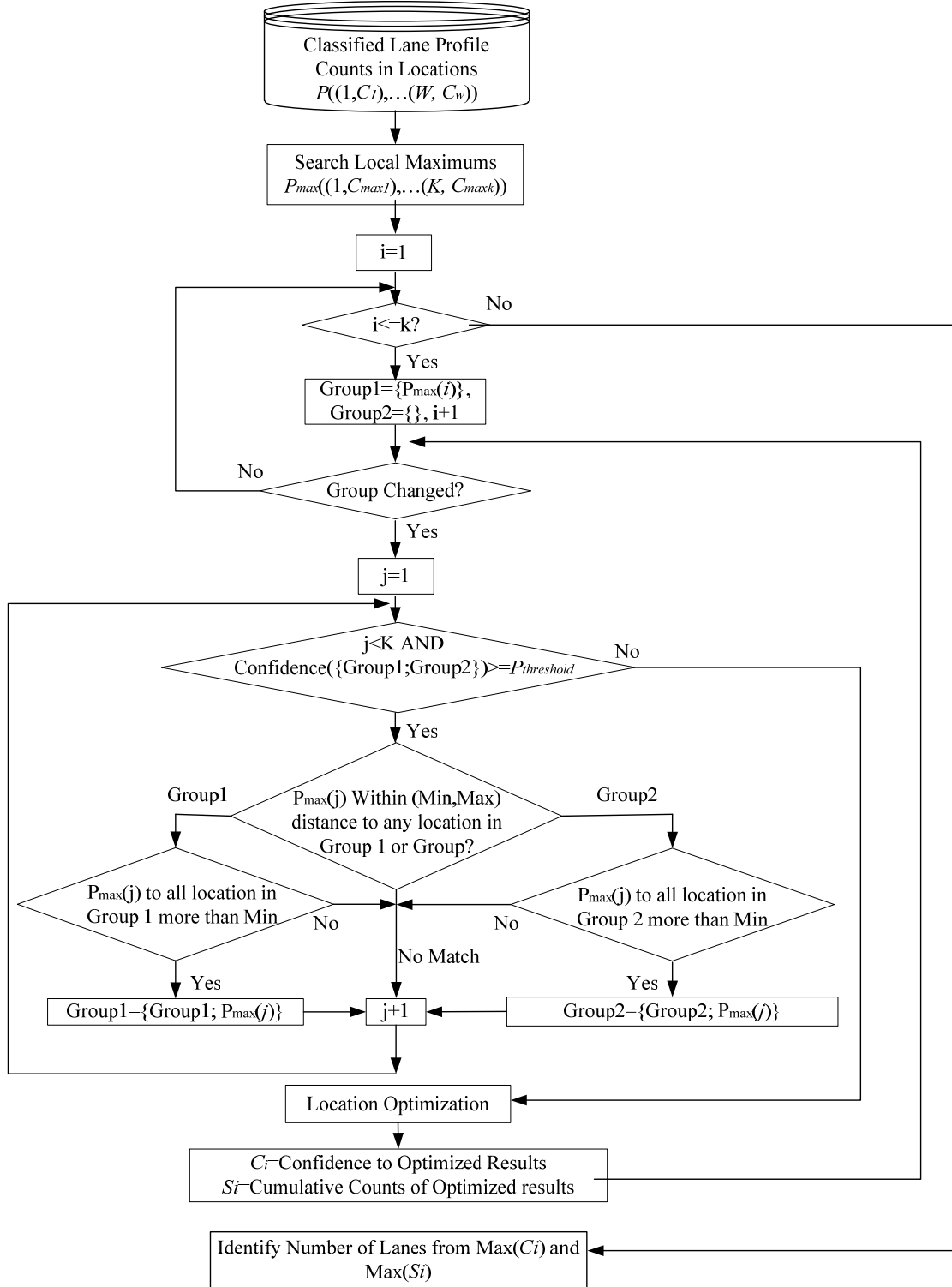


Figure 5-4. Flexible Lane-Width Optimization for Lane Identification

### 5.5.1 Local Maximum Filter

In the first step of the optimization, only locations with a local maximum number of candidate lane profiles are kept. Assuming that the algorithm is scanning a road segment that includes  $W$  pixels in a road profile, the local-maximum location and the corresponding counts can be defined as follows:

$$C_{\max i} = \begin{cases} C_i, & \text{if } C_i \geq C_{i\pm 1}, C_i \geq C_{i\pm 2}, i \in [3, W-2] \\ 0, & \text{otherwise} \end{cases} \quad (5-6)$$

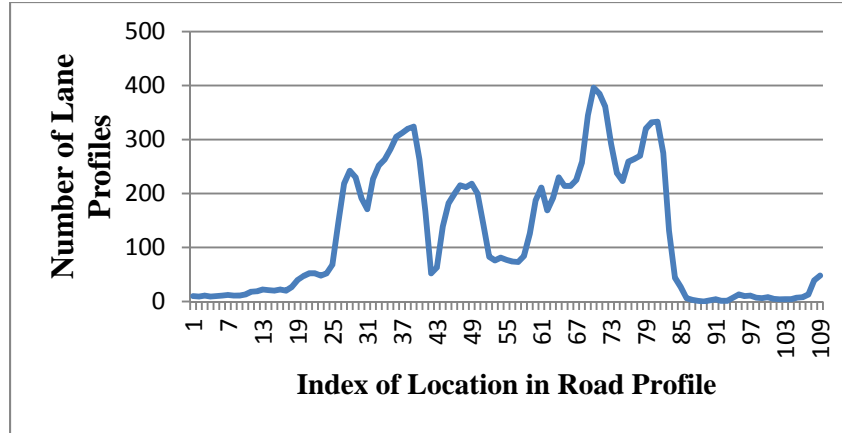
$$L_{\max} = \left\{ L_{\max 1}, L_{\max 2}, \dots, L_{\max k} \mid C_{\max L_{\max i}} > 0, L_{\max i} \in [1, W] \right\} \quad (5-7)$$

where

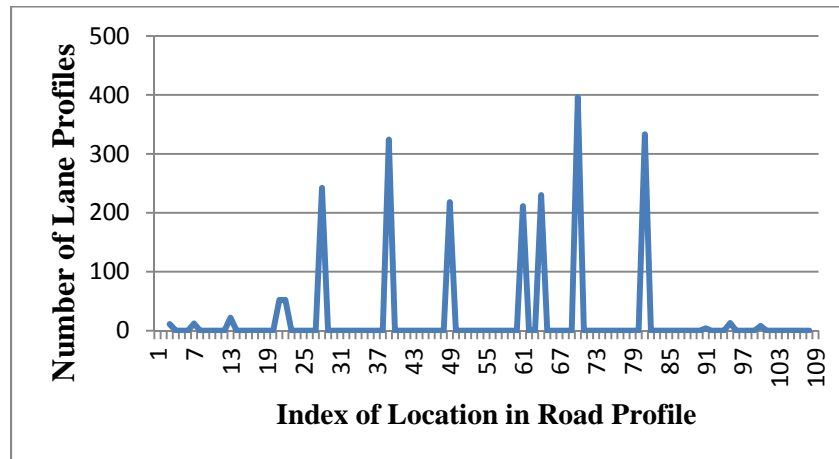
$C_i$  = number lane profiles identified at location  $i$ , and

$L_{\max i}$  = index of  $C_{\max i}$  in road profile.

The purpose of applying the local-maximum filter is to keep the locations that are most likely to be lane locations. Since the optimization process is trying to predict the lane locations in the combinations of all candidate locations, a smaller number of candidate locations will significantly reduce the size of the search space. An example of the candidate lane locations before and after local maximum filtering in a 500×120 pixel road segment is provided in Figure 5-5. Assuming that the length of lane profile is 12 pixels, in the first case, almost all 109 locations along the road profile qualified as candidate lane locations. After local maximum filtering, the number of candidate locations dropped to 15, or approximately 12% of the original size.



(a) Before Local Maximum Filtering



(b) After Local Maximum Filtering

Figure 5-5. Candidate Lane Locations Before and After Local Maximum Filtering

In Chi et. al (2009), local maximums/minimums were used to find the troughs and peaks in the brightness fluctuations of road profiles, as described in Figure 4-2, in order to identify the lane locations. The trough-peak pattern described in the research can be considered a special case that is included in the profile-based lane model developed in this research. The presence of the trough-peak pattern was used as a sufficient evidence to indicate the existence of a lane in Chi et. al. In practice, noises from turn bays, sidewalks or driveways may provide similar patterns. Therefore, the local maximums/minimums

found should be inclusive of all lane locations, but not all of them are true lane locations. The flexible lane-width optimization algorithm developed in this research provides an error-tolerant approach to estimate the actual lane locations based on the results of local-maximum filtering.

### 5.5.2 Optimization Process and Objectives

The flexible lane-width optimization algorithm as depicted in Figure 5-4 includes two major loops. For all  $k$  local maximum locations found, the outer loop repeatedly chooses the “seed” points from  $L_{max}$  to expand the search for candidate lane locations in the inner loop. Feedback from the inner loop then returns the candidate locations found in two groups, which represent two parts of lanes, as follows:

$$\text{Group1} = \{p_{11}, p_{12}, \dots, p_{1m} \mid p_{1i} < p_{1j} \quad \text{if } i < j, m \in [1, k]\} \quad (5-8)$$

$$\text{Group2} = \{p_{21}, p_{22}, \dots, p_{2n} \mid p_{2i} < p_{2j} \quad \text{if } i < j, n \in [1, k]\} \quad (5-9)$$

where  $P_{ij}$  is the location of the  $j$ th candidate lane in group  $i$ . In the following discussion, it is assumed that  $P_{1i} < P_{2i}$ . The locations provided in the two groups are then processed further in the location optimization module to predict the lane locations in the outer loop.

In the inner loop and the lane location optimization module depicted in Figure 5-4, the constraints defined in Equation 5-5 were applied and extended to locate the lane locations based on Group1 and Group2. The target is to find location set  $G$ , defined as follows:

$$G = \{p_1, p_2, \dots, p_r \mid p_i < p_j, \text{ if } i < j, r \in [0, k], p_i \in \{\text{Group1}, \text{Group2}\}\} \quad (5-10)$$

Subject to

$$\forall p_i, p_j, i \neq j \Leftrightarrow p_i \neq p_j \quad (5-11)$$

$\forall p_i, p_j$ , if  $\exists p_{i'}, p_{j'}$ ,

$$|p_i - p_{i'}| \in [D_{\max}, D_{\max g}] \text{ and } |p_j - p_{j'}| \in [D_{\max}, D_{\max g}] \Leftrightarrow p_i = p_j \quad (5-12)$$

$$\forall p_i, \exists p_j, i \neq j, |p_i - p_j| \geq D_{\min} \text{ and } |p_i - p_j| \leq D_{\max} \quad (5-13)$$

$$\forall p_i \text{ and } p_j, i \neq j \Leftrightarrow |p_i - p_j| \geq D_{\min} \quad (5-14)$$

$$\forall p_i \exists p_j, i \neq j, \left| \frac{p_i + p_j}{2} - \frac{p_{\text{floor}(\frac{r}{2})} + p_{\text{floor}(\frac{r}{2})+1}}{2} \right| < D_{\text{center}}$$

Unless  $\exists p_m < p_n, p_i \in [p_m, p_n]$

$$\left| \frac{p_m + p_n}{2} - \frac{p_{\text{floor}(\frac{r}{2})} + p_{\text{floor}(\frac{r}{2})+1}}{2} \right| < D_{\text{center}} \quad (5-15)$$

$$\forall p_i \text{ and } p_j, \text{ if } |p_i - p_j| \in [D_{\max}, D_{\max g}] \Leftrightarrow$$

$$\forall p \in L_{\max}, \text{ if } p_i \neq p, p_j \neq p, |p_i - p| > D_{\text{clear}} \text{ and } |p_j - p| > D_{\text{clear}} \quad (5-16)$$

where

$\text{floor}(r)$  = largest previous integer of real number  $r$ ,

$p_i$  = index of the  $i$ th pixel in a  $W$ -pixel wide road profile,

$D_{\max}$  = maximum distance between two lanes,

$D_{\max g}$  = maximum distance between two lane groups,

$D_{\min}$  = minimum distance between two lanes,

$D_{\text{center}}$  = maximum distance between centers of “mirror” lanes, and

$D_{\text{clear}}$  = minimum distance between middle lane and local maximum locations.

For each location included in  $L_{\max}$ , the optimization algorithm will take it as the seed point and repeat the above procedures to search for  $G$ . If the lane locations found by



expanding seed point  $L_{max i}$  are defined as  $G_{L_{max i}}$ , each  $G_{L_{max i}}$  can be evaluated by two objective functions, namely, the number of lane profiles included and the corresponding confidence. These two objectives are an extension of the objective function in Equation 5-4. Assuming that the accuracy rate of a classifier model is  $\omega$  and the presence of a lane profile is a random variable, for a two-class classification with a large enough sample size, the expected number of lane profiles identified from the lane area can be estimated using a normal distribution, as described in Equation 5-3. In the fixed lane-width optimization, this feature was used only for thresholding. It could also be applied to estimate the reliability for lane prediction. For a predicted lane location  $i$  in the road segment profile, the confidence in this prediction is calculated as:

$$\Phi\left(\frac{C_{\max i} - \mu}{\sigma}\right) \quad (5-17)$$

where

$$\mu = w_i \times \omega,$$

$$\sigma = \sqrt{w_i \times \omega \times (1 - \omega)},$$

$\omega$  = prediction accuracy of lane model applied in classification, and

$w_i$  = number of profiles classified at profile location  $i$ .

For multiple lane locations predicted by a search starting from local maximum location set  $L_{\max i}$ , define:

$$f(G_{L_{\max i}}) = \sum_{i=p_1}^{i=P_r} C_{\max i} \quad (5-18)$$

$$g(G_{L_{\max i}}) = \Phi\left(\frac{\sum_{i=p_1}^{i=P_r} C_{\max i} - \mu_G}{\sigma_G}\right) \quad (5-19)$$

where

$$\mu_G = \sum_{i=p_1}^{i=p_r} w_i \times \omega$$

$$\sigma_G = \sqrt{\sum_{i=p_1}^{i=p_r} w_i \times \omega \times (1 - \omega)}$$

$\omega$  = prediction accuracy of lane model applied in classification, and

$p_i$  = index of the  $i$ th pixel in road profile.

Note that although the length for each location is always  $M$ , the actual number of profiles  $w_i$  classified can be different at each location  $i$ . This is because many profiles are excluded by the saturation and vegetation index threshold presented in Section 5.2. Without the process of thresholding, at every profile location  $i$ , it was assumed that  $w_i = M$ . In cases where all lane areas are extracted for classification, this assumption works well. However, when a lane area is occluded, the number of lane profiles that can be detected with a 100% accuracy rate will be significantly lower than the expected number that is estimated based on a sample size of  $M$ . Consequently, the reliability of lane prediction based on Equation 5-19 tends to be very low. In order to lessen the discrepancy between the estimated reliability and the actual reliability, the saturation and vegetation thresholds for pre-processing are both necessary and effective.

### 5.5.3 Inner Optimization Loop

In a search expanded from a seed point, the inner loop visits every local maximum location in  $L_{max}$  to find the possible lane locations. These locations are then divided into Group1 and Group2 as defined in Equations 5-8 and 5-9. At a candidate location, the probability that there is not a lane is calculated based on Equation 5-17. If

the probability is higher than a threshold value,  $T_{seach-lane}$ , the location is skipped. Otherwise, the inner loop checks if the location is within a reasonable distance to the locations in Group1 or Group2. If the location is merged into Group1 or Group2, the inner loop revisits every location left in  $L_{max}$  again until no changes have been made to the two groups after traversing all the remaining locations in  $L_{max}$ .

In the implementation of the inner loop, if the elements of  $L_{max}$  are visited in their paired  $C_{maxi}$ 's descending order, the execution of the inner loop can usually stop earlier. Since this feature benefits every inner loop search expanded from every seed location, the speed of optimization execution could be significantly improved, compared to visiting in the  $L_{max}$ -based order. The pseudo code for the inner loop is given in Figure 5-6.

In the cases when all the elements in  $L_{max}$  are visited in their paired  $C_{maxi}$ 's descending order, the inner loop of the search stops when one of the following two conditions is met:

1. All locations in the local maximum set  $L_{max}$  have been visited.
2. Group1 and Group2 already included more than two candidate locations, and the probability estimated using Equation 5-19 is lower than the threshold reliability level  $T_{seach-lane}$ .

The second condition specifies when the optimization should stop searching for more candidate lane locations for Group1 and Group2.

```

% CntSort: Number of lane profiles Ci and their location index in descending order of Ci
% cdf_cnt: Array of probability that is no lane at location index
% CntMax: Number of profiles Wi classified and their location index
Posg1=CntSort(Seed_Location);
Posg2=[];

GroupChanged=1;
while(GroupChanged==1)
GroupChanged=0;
fori=1:Maximum_Location_Index
    % Reach the location with 0 lane-profiles
    ifCntSort(i,1)==0
break;
    % High Confidence there is NO lane in index i
    elseifcdf_cnt(CntSort(i,2))>=cdf_threshold_nolane
continue;
    % Wi not large enough for Normal Distribution Estimation
    %or Number of lane profiles less than minimum threshold
    elseif (CntMaxCntSort(i,2)<Mininum_Sample_Size )
        ||(CntSort(i,2)<Minimum_LaneProfile_Counts)
continue;
    else
    if (Distance between CntSort(i) AND
        ANY element in Posg1 is between [min_lane_distance, max_lane_distance])
        AND
        (Distance between CntSort(i) AND
        ALL element in Posg1 are no less than min_lane_distance)
        CntSort(i) merge to Posg1;
    else
        if size(Posg2)==0
        CntSort(i) merge to Posg2;
        GroupChanged=1;
        else
        if (Distance between CntSort(i) AND
            ANY element in Posg2 is between [min_lane_distance, max_lane_distance])
            AND
            (Distance between CntSort(i) AND
            ALL element in Posg2 are no less than min_lane_distance)
            CntSort(i) merge to Posg2;
            GroupChanged=1;

    if (size(Posg1)+size(Posg2))>2
        % Calculate reliability of current lane locations based on
        % Total_Cnt: Total number of lane profiles included
        % Total_MaxCnt: Total number of profiles classified
        % p:Accuracy Rate of Classifier

        cdf_sum=cdf('norm',Total_Cnt, Total_MaxCnt*p, sqrt(Total_MaxCnt*p*(1-p)));
        ifcdf_sum<cdf_threshold_lane
            GroupChanged=0;

break;

```

Figure 5-6. Pseudo Code for Inner Loop Optimization

In summary, two types of reliability thresholds are applied in the inner loop, namely,  $T_{seach-lane}$  and  $T_{seach-nolane}$ . The threshold  $T_{seach-lane}$  specifies the minimum overall reliability of all lane candidates found in Group1 and Group2, and the threshold  $T_{seach-nolane}$  specifies the maximum amount of non-lane profiles that can be ignored at a candidate lane location.

The application of  $T_{seach-nolane}$  is straightforward because all profiles at non-lane locations should be non-lane profiles. The probability that there is not a lane can be estimated with the total number of profiles classified,  $n$ , the model accuracy rate,  $p$ , and the detected non-lane profiles using Equation 5-17. The application of  $T_{seach-lane}$  is more complex due to occlusion in the lane areas. Unlike the situation of non-lane areas, not all profiles in lane areas are lane profiles. Profiles from vehicles, buildings or plants can also be included in the input of the lane model. Therefore, the number of actual lane profiles detected by the lane model can be significantly lower than the expected number calculated using Equation 5-17 when all the profiles in a candidate lane location are assumed as lane profiles. In the previous section, saturation thresholding and vegetation index thresholding were introduced to exclude the non-lane profiles for lane areas. However, unhandled non-lane profiles will still present in practice.

Table 5-2 provides the impacts of unhandled non-lane profiles to the reliability of predicted lane locations. Define  $n$  as the total number of profiles included in a lane location,  $\omega$  as the accuracy rate of lane model, and  $q$  as the percentage of actual lane profiles included in  $w_i$  of Equation 5-17, the four columns on the right side show how the reliabilities of the predictions change along with  $q$ . For instance, if  $n=100$ ,  $q=100\%$ , and the number of lane profiles identified by the lane model equals the expected value  $n \times \omega$ ,

the reliability of the prediction is 50%. With the same  $n$  and  $q$ , if the number of identified lane profiles is one or two standard deviations more than  $n \times \omega$ , the reliabilities increase to 84.13% and 97.72%, respectively. If  $q=90\%$ , given a predicted lane location with  $n \times \omega \times q$  predicted lane profiles, the reliability should still be 50%. However, since the 10% non-lane profiles are not handled, the reliability calculated by assuming  $w_i = n \times \omega$  drops quickly.

The lane models with a higher  $\omega$  are affected more easily when applied to longer road segments with a higher  $n$ . In the best case included in Table 5-2, when  $n=100$ ,  $\omega=0.6$ , the predicted lane profiles are two standard deviations more than the actual expected value. If 90% of  $w_i$  are actual lane profiles in this case, the reliability of the prediction drops from 97.72% to 78.09%. When 70% of  $w_i$  are actual lane profiles, the reliability of the prediction drops to the level of 4.70%.

In other cases with higher  $\omega$  and  $n$ , the calculated reliabilities are close to 0% with 90% or 80% lane profiles in  $w_i$ . Since it is extremely difficult, if not impossible, to exclude all the non-lane profiles presented at the lane locations before the lane model recognition, the threshold  $T_{seach-lane}$  should be set to a very low level. The accuracy rate  $\omega$  of the lane model should also be reduced in order to lessen the discrepancy between the actual and calculated reliabilities.

Table 5-2 Impacts of Unhandled Occlusions to Prediction Reliability

	$\omega$	Reliability of Lane Locations with Unhandled Occlusion			
$n=100$		Predicted Lane Profiles = $n \times \omega \times q$			
		$q=100\%$	$q=90\%$	$q=80\%$	$q=70\%$
	0.9	50.00%	0.13%	0.00%	0.00%
	0.8	50.00%	2.28%	0.00%	0.00%
	0.7	50.00%	6.33%	0.11%	0.00%
	0.6	50.00%	11.03%	0.72%	0.01%
		Predicted Lane Profiles = $n \times \omega \times q + 1 \times \text{stddev}$			
		$q=100\%$	$q=90\%$	$q=80\%$	$q=70\%$
	0.9	84.13%	2.28%	0.00%	0.00%
	0.8	84.13%	15.87%	0.13%	0.00%
	0.7	84.13%	29.89%	1.99%	0.02%
	0.6	84.13%	41.11%	7.36%	0.37%
		Predicted Lane Profiles = $n \times \omega \times q + 2 \times \text{stddev}$			
		$q=100\%$	$q=90\%$	$q=80\%$	$q=70\%$
	0.9	97.72%	15.87%	0.00%	0.00%
	0.8	97.72%	50.00%	2.28%	0.00%
	0.7	97.72%	68.17%	14.57%	0.49%
	0.6	97.72%	78.09%	32.65%	4.70%
$n=200$		Predicted Lane Profiles = $n \times \omega \times q$			
		$q=100\%$	$q=90\%$	$q=80\%$	$q=70\%$
	0.9	50.00%	0.00%	0.00%	0.00%
	0.8	50.00%	0.23%	0.00%	0.00%
	0.7	50.00%	1.54%	0.00%	0.00%
	0.6	50.00%	4.16%	0.03%	0.00%
		Predicted Lane Profiles = $n \times \omega \times q + 1 \times \text{stddev}$			
		$q=100\%$	$q=90\%$	$q=80\%$	$q=70\%$
	0.9	84.13%	0.06%	0.00%	0.00%
	0.8	84.13%	3.37%	0.00%	0.00%
	0.7	84.13%	12.30%	0.04%	0.00%
	0.6	84.13%	23.21%	0.69%	0.00%
		Predicted Lane Profiles = $n \times \omega \times q + 2 \times \text{stddev}$			
		$q=100\%$	$q=90\%$	$q=80\%$	$q=70\%$
	0.9	97.72%	1.25%	0.00%	0.00%
	0.8	97.72%	20.37%	0.01%	0.00%
	0.7	97.72%	43.63%	1.02%	0.00%
	0.6	97.72%	60.56%	7.16%	0.07%

\*stddev = Standard Deviation

The majority of the optimization work that follows the constraints specified in Equations 5-11 through 5-16 is accomplished by the inner loop of the algorithm. The purpose of the inner loop is to find the lane locations for Group1 and Group2, as defined in Equations 5-8 and 5-9, respectively. In the developed optimization algorithm, the target lane locations are modeled as two groups of lanes that are divided by medians, plants, or turn-bays. If travel lanes are not divided, all candidate lanes would be included in Group1 when they are output to the outer loop.

#### 5.5.4 Outer Optimization Loop

An iteration of the outer loop picks a “seed” location from  $L_{max}$  and executes the following tasks:

- Search for the lane candidates Group1 and Group2 in the inner loop.
- Optimize the lane locations included in Group1 and Group2.
- Compare the optimization results expanded from the “seed” location  $i$  with the current global optimal to keep the better one as the new global optimal.

Figure 5-7 shows the flowchart of the lane location optimization module in the outer loop. The first three steps in the location optimization module check the results of Group1 and Group2 against constraints defined in Equations 5-12, 5-14, and 5-16 respectively. After that, the rest of the candidate lane locations are processed by an important “Mirror Match” algorithm developed to fulfill the requirement of the constraint specified in Equation 5-15. The description of the algorithm is provided in Figure 5-8. This algorithm first calculates the center  $M$  of the two center candidate lane locations. The center is then used as the reference point for candidate lane locations to match their mirror lane locations. To qualify as two mirror lane locations, the distance of the center  $M'$



of the two candidate lane locations to  $M$  must be smaller than the  $D_{center}$  threshold specified in the constraint in Equation 5-15. If the mirror locations can be matched, both lane locations are treated as the true lane locations. This search continues until no mirror lane locations can be found for the rest of the lane locations specified in Group1 and Group2.

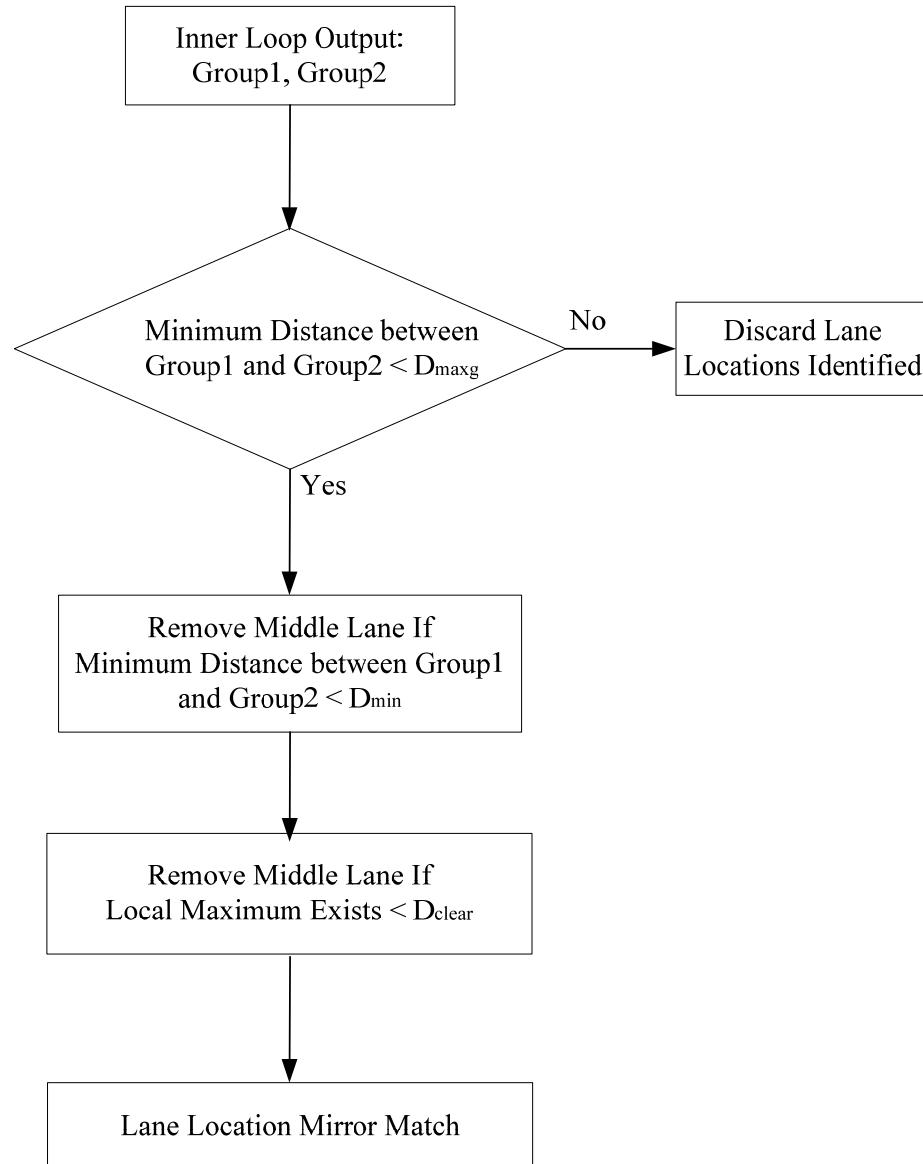


Figure 5-7. Workflow of the Location Optimization Module

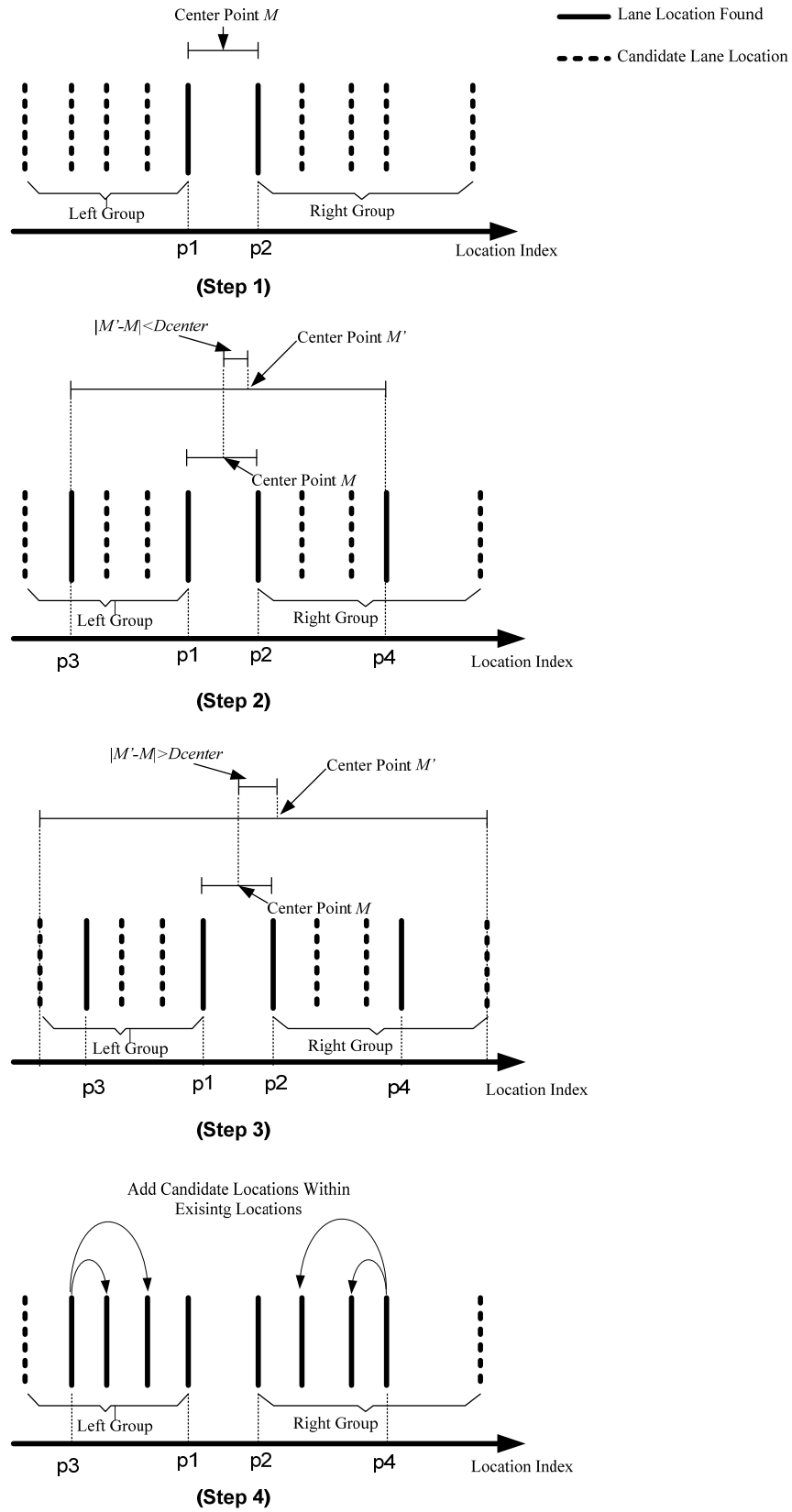


Figure 5-8. Mirror-Lane Match in the Outer Loop

At the end of each outer loop execution,  $G_{L_{max}i}$  is provided by the location optimization module. After finishing the searches expanded from each point of  $L_{max}$  in the outer loop, the  $G_{L_{max}i}$  with the highest  $f$  and  $g$  are adopted. When two predictions generate maximum  $f$  and  $g$ , respectively, the one generated by the seed point with a larger  $C_{maxi}$  is adopted. For the rest of this dissertation, the final optimized locations are noted as  $G_{max}$ . The optimization algorithm fails to provide  $G_{L_{max}i}$  when no groups with more than one lane that satisfies the constraints in Equations 5-11 through 5-16 exist. There could be two reasons behind this failure:

1. The inner loop cannot find enough amounts of lane candidates before  $g(G_{L_{max}i})$  drops to a threshold level  $T_{search-lane}$ .
2. The optimal results,  $G_{max}$ , generated by this algorithm have the highest/lowest lane reliability or group reliability that is lower than a threshold level  $T_{result}$ .

The two reliability thresholds apply to different situations in the optimization process. The value of  $T_{search-lane}$  is applied to each inner loop where the algorithm is searching for candidate lane locations. It decides when the loop should stop searching for more lane locations for Group1 and Group2 to avoid the problem of overestimating number of lanes. The second threshold  $T_{result}$  is applied to the optimization results  $G_{max}$ . It decides whether the results should be adopted. When  $g(G_{max}) < T_{result}$ , the  $G_{max}$  is discarded, and the target road segment is labeled as “unable to identify.”

Both failure reasons are a critical improvement over the current automatic road/lane extraction studies. In existing literature, the accuracy rate is assumed by applying the accuracy rate summarized in similar test scenarios. However, the reliability of the automatic extraction algorithm is different when applied to different road segments

because of the varying noise levels. This difference is not convincing enough to be summarized in descriptive categories that may divide scenarios of aerial images as suburban vs. urban or high-volume vs. low-volume traffic. In the automatic extraction process feedback, the  $g(G_{Lmaxi})$  values of the target road segments provide a practical approach to reducing the amount of manual work required in data collection. Due to occlusions from non-pavement objects, sometimes it is not possible to manually recognize number of lanes from aerial images. Therefore, it is extremely difficult develop a model that can correctly predict number of lanes for road segments in all circumstances. The purpose of automatic extraction is to reduce the manual work needed by extracting data from as many road segments as possible, while providing a satisfying accuracy rate. In the framework developed in this research, the reliability drops when the discrepancy between the expected and identified number of lane profiles grows. For road segments with a very low  $g(G_{max})$ , manual identification is required. In the automatic extraction process, it means that the higher the accuracy required, the lower percentage of target road segments can be automatically processed.

## **5.6 Summary**

In this chapter, two optimization algorithms were developed in the lane identification module to estimate the number of lanes. The objective of these algorithms is to maximize both the number of lane profiles previously identified by the profile-based lane model and the reliability of lane locations.

A fixed lane-width optimization algorithm was developed and tested using a SVM classifier. It was found that the method was very accurate when applied to suburban or urban areas where the majority of paved areas were not occluded. Although the noise

from shadows, vehicles, plants, and buildings existed in the aerial images, the number of lanes was correctly extracted for all six test roadway links. The automated program also suggested approximated locations and reasonable lane widths. However, the fixed lane-width optimization left the pixel-mapping precision and reliability estimation problems unresolved. The flexible lane-width optimization addressed both issues. The algorithm formulated the problem of estimating number of lanes as a multi-objective constrained optimization problem that aims to maximize the number lane profiles and their reliability. It is concluded that precision and coverage are two conflicting objectives in the automatic extraction framework developed in this dissertation.

In the next chapter, a test is performed using the flexible lane-width optimization algorithm in a roadway network to evaluate the automatic extraction framework.

## **CHAPTER 6**

### **EVALUATION OF FRAMEWORK**

#### **6.1 Introduction**

In the previous chapter, different profile-based lane models and lane location optimization algorithms were presented and compared. For the test cases provided in Chapters 4 and 5, the multi-class classification-based lane model, the two-class classification-based lane model, and the fixed lane-width optimization algorithm were found to provide good prediction performance. However, based on the conclusions in Section 4.4 and 5.5, the two-class classification lane model and flexible lane-width optimization are theoretically superior to their alternatives. In this chapter, the automatic extraction framework is implemented using the two-class lane model and the flexible lane-width optimization. To evaluate the framework, two test road networks that include a total of 490 road segments are used.

#### **6.2 Data Preparation, Lane Model Training, and Implementation**

In this chapter, two cases are included to evaluate the general performance of the developed framework and its performance in areas with heavy plant occlusions, respectively. The first test case includes two areas that cover approximately 1.8 square miles and more than 500 road segments. The second test case covers two areas with a high percentage of plant coverage in the first test.

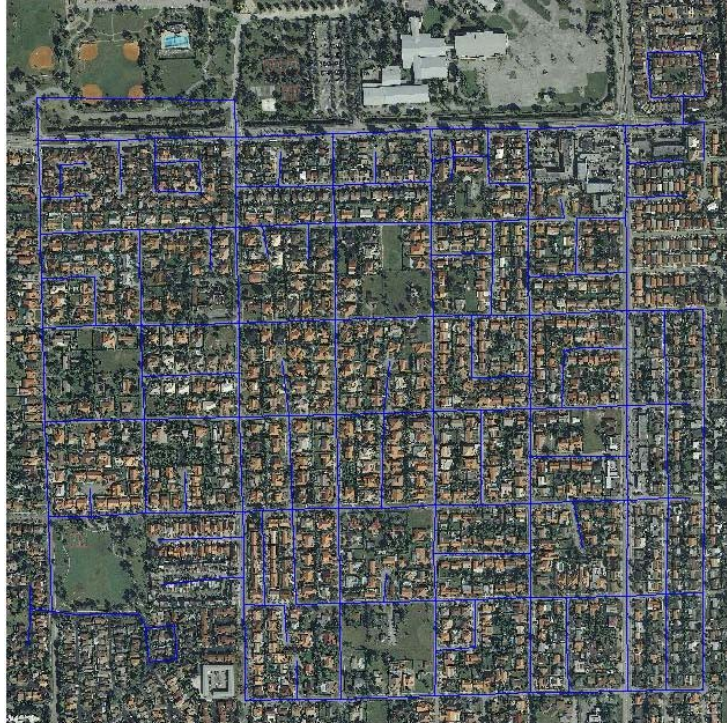
The areas were randomly selected from geo-rectified aerial images of Miami-Dade County, Florida. The resolution and format of the image used in this test is identical to the images used in the fixed lane-width optimization test in Chapter 5. A GIS layer of roadway networks for Miami-Dade County is used in the automatic extraction

program. A MATLAB program was developed to identify all roadways located within the selected aerial image. Identified locations of road segments in the pixel matrix are shown in Figure 6-1. The implemented program detected 490 road segments within the two target areas using the mapping technique presented in Chapter 3. To evaluate the performance of the network when heavy occlusions from plants are presented in road areas, road networks that cover this special type of area are also selected for testing. In a special test case, 180 road segments from the areas, shown in Figure 6-2 were included. As seen in special cases of the road networks presented in Figure 6-1, most test road segments in Figure 6-2 are typically covered by plants and their shadows.

The training profiles used in this test consist of 389 “lane profiles” and 493 “non-lane profiles”. Profile features were extracted and calculated using the same method applied in two-class classification tests in Chapter 4. Each of the  $k$ -pixel wide test profiles are represented by a  $(k+1)$ -dimension vector  $\overline{v}_i = [e_{i1}, e_{i2}, \dots, e_{ik}]^T$  following the definitions described in Equations 3-4 through 3-8. While the similarities between pixels are still calculated using Equation 3-10, the sum of similarity data were normalized to achieve the homogeneity measure  $H_i$ . The following equation is applied in this test in place of Equation 3-9:

$$H = \frac{\sum_{j=1}^K S(e_{ij}, e_{(i-1)j})}{k} \quad (6-1)$$

After normalization of the similarity measurement, all  $(k+1)$  features will fall within  $[0,1]$ . The training data of 822 instances were classified using SVM, with both the RBF kernel and the Polynomial kernel. The five-folder cross-validation method was used to evaluate the performances.



(a) Test Road Network 1



(b) Test Road Network 2

Figure 6-1. Test Road Segments Identified by Automatic Extraction Program





(a) Test Region 1



(b) Test Region 2

Figure 6-2. Road Network with Heavy Plant Occlusion

Table 6-1 gives a summary of the classification statistics when individual classifiers were applied to the training data set. To measure the prediction accuracy for different classes, Table 6-2 gives the results for the following three accuracy rates:

$$TP\ Rate = \frac{TP}{P} \quad (6-2)$$

$$FP\ Rate = \frac{FP}{P} \quad (6-3)$$

$$precision = \frac{TP}{(TP + FP)} \quad (6-4)$$

where

$P$  = number of positive instances,

$N$  = number of negative instances,

$TP$  = true positive instances predicted, and

$FP$  = false positive instances predicted.

Table 6-1 Summary of SVM Classification Results for Test Scenarios

	Parameter Setting	Correctly Classified Instances	Kappa Statistic
Polynomial Kernel* $(x_i, x_j) = (\langle x_i, x_j \rangle)^d$	d=2	681 (82.85%)	0.6580
	d=3	721 (87.71%)	0.7550
	d=4	730 (88.81%)	0.7769
RBF Kernel $(x_i, x_j) = \exp(-\gamma \ x_i - x_j\ ^2)$	$\gamma=0.01$	586 (71.29%)	0.4338
	$\gamma=0.10$	660 (80.29%)	0.6093
	$\gamma=1.00$	708 (86.13%)	0.7242

In the four test cases, the  $TP$  of lane profiles is always higher than the  $TP$  of non-lane profiles. This means that a higher percentage of lane profiles is likely to be successfully identified compared to the percentage of non-lane profiles identified. However, a lower  $FP$  rate and a higher  $precision$  value of non-lane profiles indicate that the predictions of non-lane profiles are more reliable.

Table 6-2 Detailed Accuracy of Classification Results

	Parameter	Class	TP Rate	FP Rate	Precision
Polynomial Kernel	d=2	Lane Profiles	0.879	0.217	0.784
		Non-Lane Profiles	0.783	0.121	0.878
	d=3	Lane Profiles	0.928	0.169	0.832
		Non-Lane Profiles	0.831	0.072	0.928
RBF Kernel	gamma=0.1	Lane Profiles	0.907	0.291	0.737
		Non-Lane Profiles	0.709	0.093	0.895
	gamma=1	Lane Profiles	0.938	0.208	0.802
		Non-Lane Profiles	0.792	0.062	0.935

To build the lane-model for the two test scenarios, the ensemble method is applied to include two SVMs with both the RBF kernel ( $\gamma=1$ ) and the Polynomial kernel ( $d=3$ ). The lane model labels a candidate profile as a “lane profile” when both classifiers provide the identical positive predictions. The purpose of adopting the ensemble method is to improve the precision of “lane profile” class predictions. Compared to the lane models applied in the test in Chapter 5, this new model is likely to miss more lane profiles; however, the reliability of predictions for lane profiles can be expected to improve.

A MATLAB program was developed to implement the following functions of the automatic extraction framework:

- Two-class lane model training.
- Data extraction from aerial image.
- Lane profile classification.
- Estimation of number of lanes using the flexible lane-width optimization algorithm.

The program first built a lane model based on a training data set and parameter settings, as mentioned above. Afterwards, the data preparation process read the geographical coordinates of all road segments included in the target aerial images. The

geographical information was then converted to pixel coordinates based on the candidate profiles that were extracted and classified. Finally, number of lanes for each road segment was estimated by the optimization module. The parameter settings of the program are given in Table 6-3.

Table 6-3 Parameter Setting of Test Program

Pre-processing	Saturation Threshold	0.1
	Vegetation Index Threshold	0
Profile Scan	Profile Width $W$	120
	Lane Model Accuracy Rate $\omega$	0.76
	Minimum $W_i$	$10/\omega$
Optimization Setting	Minimum $C_i$ of Lane Location $i$	=Minimum $W_i$
	$D_{max}$	14
	$D_{maxg}$	36
	$D_{min}$	7
	$D_{center}$	3
	$D_{clear}$	3
	$T_{search}$ : No Lane	0.95
	$T_{search}$ : Have Lane	0.01
	$T_{result}$	0

As indicated in Chapter 5, 0.1 and 0 were applied for the saturation thresholding and the vegetation index thresholding, respectively. Since the resolution of the aerial images used is one foot per pixel, the maximum profile width extracted is set to 120 pixels, which is large enough to cover a ten-lane road segment. The lane model accuracy is set to 0.76. The accuracy rate used here is lower than the actual accuracy rate of SVM with the polynomial kernel or the RBF kernel. As discussed in Chapter 5, a lower accuracy rate adopted in the optimization process can help to handle the occlusion of lane area by narrowing the discrepancy between the true and predicted lane location reliabilities. As a result, the lane area with heavy occlusions is less likely to be excluded incorrectly in the optimization process.

### 6.3 Review of Test Road Network

Road segments in the test networks include typical two-lane undivided local roads, two-lane divided highways, four-lane highways, and six-lane highways. Sample images for each type of roads that were cut from the test aerial images are shown in Figure 6-3. Among the 490 road segments in the images, typical noise factors, such as vehicles, turn bays, road medians, shadows, and plants were present. Some road segments also have turn bay in the middle, which is likely to be recognized as travel lanes using a traditional marking or pavement-based lane models.

Test areas with heavy plant occlusion focused on road segments, where a small part of the pavement is exposed. Using the traditional lane models, travel lanes cannot be identified unless the objects that occlude the area can first be modeled and identified. Since the new lane model developed in this research no longer requires a reconstruction of the whole paved area, impacts of occluded parts are minimized. Therefore, areas that are similar to that described in Figure 6-4 were included in the second test to evaluate the performance of the framework in this special scenario.

The performance results of the implemented framework for the two test scenarios are given in Tables 6-4 and 6-5, respectively. In Chapter 5, two types of reliability threshold were introduced in the flexible lane-width optimization, namely  $T_{search}$  and  $T_{result}$ . In the program settings, it was specified that  $T_{result} = 0$ , which means that no threshold was applied in the optimization results. The accuracy rate and the coverage rate in this case are 85.09% and 88.98%, respectively.





(a) Two-Lane Undivided Local Collectors



(b) Two-Lane Divided Highway with Median



(c) Four-Lane Highway with Turn-Bay



(d) Six-Lane Highway with Median

Figure 6-3. Four Types of Road Segments Included in Test Networks



Figure 6-4. An Example of Road Segment with Heavy Occlusions

Table 6-4 Summary of Framework Performance in Test Scenario One

	$T_{search}$ Only	Highest Lane $g>0.01$	Highest Lane $g>0.1$	Lowest Lane $g>0.001$	Lowest Lane $g>0.01$	All-Lane $g>0.01$	All-Lane $g>0.1$
Correct	371	326	293	196	158	262	227
Incorrect	65	59	59	30	22	55	41
Excluded	54	105	138	264	310	173	222
Accuracy	85.09%	84.68%	83.24%	86.73%	87.78%	82.65%	84.70%
Coverage	88.98%	78.57%	71.84%	46.12%	36.73%	64.69%	54.69%

Table 6-5 Summary of Framework Performance in Test Scenario Two

	$T_{search}$ Only	Highest Lane $g>0.01$	Highest Lane $g>0.1$	Lowest Lane $g>0.001$	Lowest Lane $g>0.01$	All-Lane $g>0.01$	All-Lane $g>0.1$
Correct	146	102	92	60	44	85	74
Incorrect	17	9	9	2	1	8	7
Excluded	17	18	28	67	84	36	48
Accuracy	89.57%	91.89%	91.09%	96.77%	97.78%	91.40%	91.36%
Coverage	90.56%	86.05%	78.29%	48.06%	34.88%	72.09%	62.79%

The rest of the six columns in Table 6-4 and 6-5 specified the performance of framework in three types of situations when  $T_{result}$  was applied to different reliability indicators of optimization results, as follows:

1.  $T_{result}$  was applied to the highest reliability of the lane included in optimization results,
2.  $T_{result}$  was applied to the lowest reliability of the lane included in optimization results, and
3.  $T_{result}$  was applied to the reliability of all lanes included in optimization results.

It is evident that the threshold with the highest lane reliability covered more road segments than the other two approaches, while the threshold with the lowest reliability yielded the highest accuracy rate. For each situation, two different levels of  $T_{result}$  were applied. Theoretically, a higher  $T_{result}$  should increase the accuracy rate, but reduce the coverage rate. However, while the coverage rate dropped, the accuracy rate did not improve significantly in the test cases. For thresholding with the highest lane reliability approach, the accuracy rate actually dropped when  $T_{result}$  increased.

Compared to the performance data when no  $T_{result}$  was applied, the accuracy rate only increased when it was applied to the lowest-lane-reliability approach. However, the coverage rates are too low to be accepted in automatic extraction practice. Since the process of applying thresholds to the optimization results produced a very slight improvement to the prediction accuracy rate, as compared to the amounts of road segments excluded when  $T_{result}$  was applied, the benefit of this accuracy improvement was deemed limited. Therefore, in this test network, it can be concluded that only  $T_{search}$  should be applied, in which case both the accuracy rate and the coverage rate are considered promising.

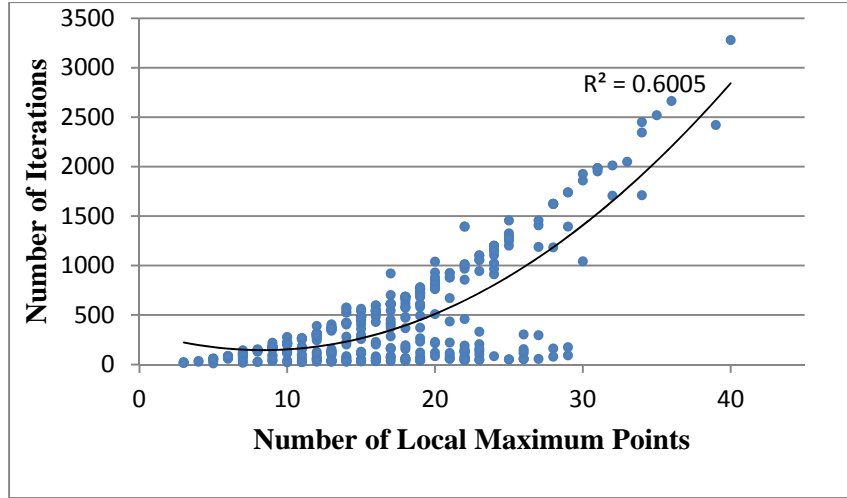
The performance of the test framework in test scenario two is given in Table 6-5. Overall, it shows that the accuracy and the coverage measurements in test scenario two



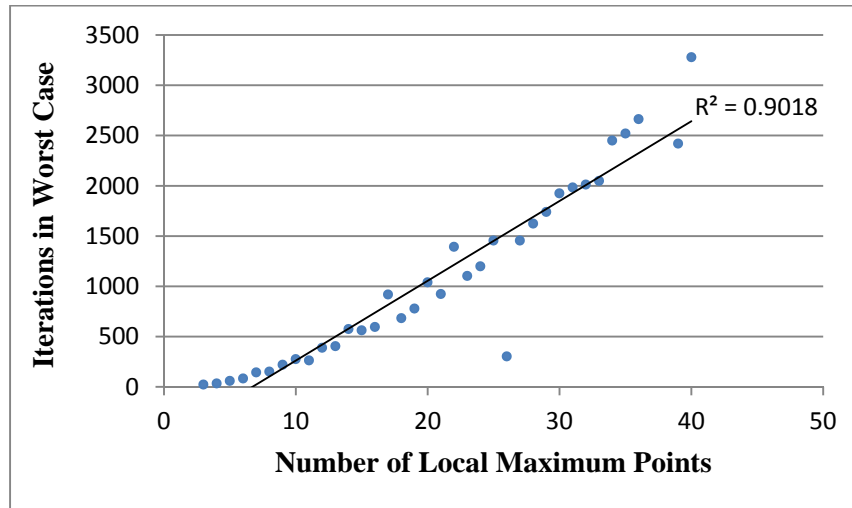
are superior to the measurements in test scenario one. The only exception is the coverage rate measurements when the lowest lane reliability was used for thresholding. Since test scenario two was created specifically to test the performance of the framework in areas with heavy plant occlusion, it suggested that this framework could handle the occlusion problem caused by plants more effectively than by other types of occlusion or noise factors.

The efficiency of the optimization process is another indicator to be measured in evaluating the automatic extraction framework developed in this research. Depending on the configuration of the test machine, a large variation of the optimization execution time can be expected. Therefore, in this research, the efficiency of the automatic extraction framework is measured by the number of inner loop executions. Intuitively, the number of iterations should increase along with the number of local maximum points in  $L_{maxi}$  as defined in Equation 5-7. Figure 6-5 shows the number of iterations executed for all test cases and the maximum number of iterations for each  $L_{maxi}$  size.

It can be concluded from Figure 6-5 that the variation between numbers of iterations for the similar number of local maximum points is high. Although a polynomial trend is observed, the R-Squared is relatively low ( $R^2=0.6005$ ). Figure 6-5(b) shows the linear trend of worst cases, i.e., the highest number of executions for identical number of local maximum, is evident, as described in Figure 6-5(b). Therefore, the execution time of the optimization algorithm should approximate the expected linear increase, along with the increase of the number of local maximum points in  $L_{maxi}$ .



(a) Iterations of All Test Cases



(b) Iterations in Worst Cases

Figure 6-5. Impact of Local Maximum Number to Inner Loop Iterations

The reliability of lane group  $g$  and the number of lane profiles included in lane group  $f$  are two objectives that the flexible lane-width optimization algorithm aims to maximize. In Chapter 5, it was concluded that if classified profiles  $w_i$  from location  $i$  is always equal to the length of road segment  $w$ , the reliability of prediction always increased with the number of lane profiles included in the optimal results. In this test, as specified in Table 6-2, the saturation and vegetation index thresholds were applied in the pre-processing

module to provide a closer estimation of lane areas exposed in aerial images. The optimization results for the test road segments show that the correlation coefficient between  $g$  and  $f$  is 0.2, which is similar to the random pattern. Figure 6-6 depicts the number of lane profiles and their reliability for each test road segment.

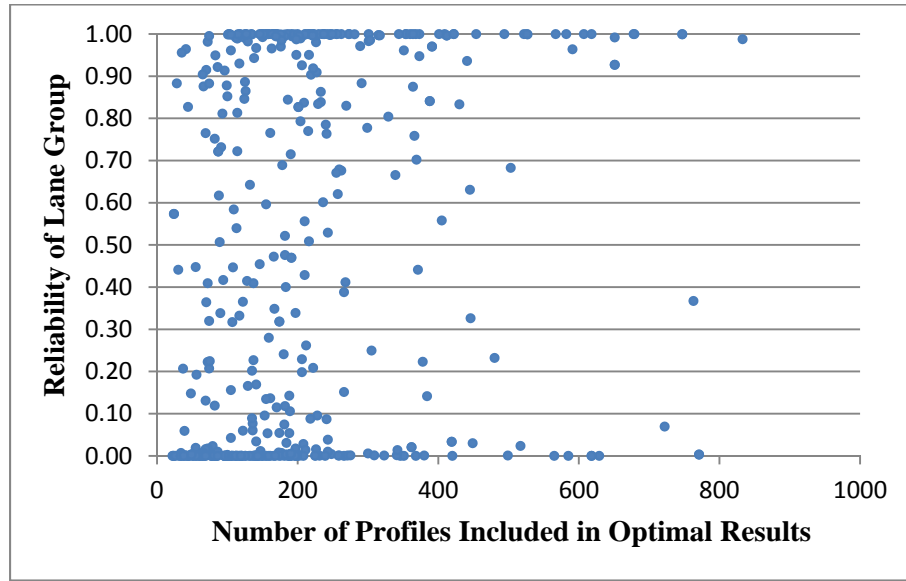


Figure 6-6. Number of Lane Profiles in Optimal Lane Locations vs. Reliability

#### 6.4 Summary

In this chapter, the proposed automatic extraction framework was implemented in MATLAB. The MATLAB program trained SVMs with the Polynomial kernel and the RBF kernel. These two classifiers were adopted in the profile-based lane model for ensemble classification. The flexible lane-width optimization developed in Chapter 5 was implemented in the program to automatically identify the number of lanes from the classification results.

The implemented program was evaluated in a general test road network with 490 road segments and a special network with 180 road segments heavily occluded by plants. The test results show that the framework yielded promising results, which were supported

by both the accuracy rate and the coverage rate. Execution efficiency of the developed framework was also estimated during the test. Measured by the number of inner loop iterations, it was found that, in general, the number of iterations and the number of local maximum points approximately followed a polynomial trend for the test cases. The test results also show that the correlation between the two optimization objectives was sparse. Therefore, the two objectives acted independently in the flexible lane-width optimization when the number of profiles classified is pre-processed through the saturation and vegetation index thresholds.

## **CHAPTER 7**

### **SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS**

#### **7.1 Summary and Conclusions**

The acquisition and update of number of lanes for transportation applications are traditionally in need of heavy labor work. High-definition aerial images make it possible to manually extract this road characteristic especially for smaller areas. However, for larger areas it is more efficient and economically feasible using an automatic extraction approach. In this dissertation research, a comprehensive review was conducted to investigate lane models, as well as models of typical noise factors around roadways. Although existing lane models were reported to work well in road extraction studies, they appear to apply only when the target lanes are not heavily occluded. This is due to the fact that automatic extractions based on existing models require reconstruction of the whole target road segment. In practice, road segments presented in aerial images are usually occluded by vehicles, buildings or plants. On the other hand, reliability of existing techniques is hard to be concluded during an extraction process. Accuracy rates presented in test scenarios are assumed in applications in similar scenarios. Since similarities between test scenarios and application scenarios (e.g., urban or suburban area images) are subjectively decided, actual prediction accuracy of existing techniques can be significantly different when compared to the claimed accuracy summarized from the test data. As such, the goal of this research was to develop a fully automated extraction framework that is better against noises and is able to provide measures on the prediction reliability. As stated, the objectives of this dissertation include the following:

- Develop a reliable lane model based on the typical appearance features of travel lanes in color aerial images, which is more robust against noises such as vehicle occlusion or sidewalks.
- Develop an algorithm to locate travel lanes in aerial images based on the improved lane model and the algorithm to decide number of lanes from identified lane locations with the reliability measurement provided.
- Implement the framework and validate its effectiveness using geo-rectified images.

The feasibility of this framework relies on the automatic acquisition of particular pixel-level information of target road segments, namely, road orientation and location estimation. A mapping technique was presented to estimate the pixel coordinates of road in the pixel matrix by using its geographic coordinates. This technique requires aerial images to be geo-rectified in order to retrieve resolution and geographic information. Adoption of this specific type of aerial image in the automatic extraction process makes it unnecessary to include manual work typically required in road extraction studies for initial road location and orientation input.

To overcome the limitations of existing lane models, a profile-based lane model was developed in this research. The new model identifies one-pixel wide profiles of lanes instead of the whole lane segment, making it possible to identify lane locations, even with heavy occlusion of roads in aerial images. The multi-class classification and two-class classification were tested to build a classifier-based profile lane model. Both approaches yielded promising results. The two-class classification approach was recommended due to the reason that there are in theory infinite types of objects to be modeled. Therefore, it

cannot be claimed that a high-performance classifier against three noise factors (classes) will yield similar promising results in other scenarios.

Along with the identification of lane profiles, two types of optimization algorithms were presented, namely, the fixed lane-width optimization and the flexible lane-width optimization. The objectives of optimization include maximizing the number of lane profiles included in determined lane locations and the probability of making the correct decision. Since the latter is estimated based on the accuracy rate of classifier(s), the non-lane objects should be excluded from lane areas before classification. This is due to the fact that the expected number of lane profiles to be identified will be significantly higher than the actual lane profiles identified in lane areas exposed in aerial images. Saturation thresholding and vegetation index thresholding were developed to lower the impact of the occluded lane area. In a test that included six divided roads, fixed lane-width optimization was tested in order to validate the general principle used in the location optimization algorithm presented in the research. Although the performance of fixed lane-width optimization was shown to be promising, the flexible lane-width optimization was recommended. The flexible lane-width algorithm was developed to address two fixed lane-width optimization issues:

1. Mapping between geographic coordinates to pixel locations may not be accurate enough to locate the actual locations of lanes.
2. Locations are not optimized to maximize the probability of lane presence.

The accuracy rate of the lane model classifier was introduced in the reliability estimation of the optimization algorithm. With large enough input profiles, the probability of two-class classification results were estimated using a normal distribution. As such, the

reliability of maximized lane locations could be measured by the cumulative density function. The purpose of providing this probability is not only to provide automatic extraction process feedback, but also to provide the operators an opportunity to compromise between extraction accuracy and extraction coverage.

The automatic extraction framework was implemented in MATLAB with a two-class classification lane model and the flexible lane-width optimization algorithm developed in this dissertation. Two test scenarios were used to evaluate the performance of the implemented program. The first test scenario included two road networks and a total number of 490 road segments. The second scenario picked out the areas with heavy plant occlusions from the first scenario to evaluate how effectively this type of occlusion could be handled in the automatic extraction framework. The test results showed that the framework provided a promising accuracy rate, a good coverage rate, and with high execution speed. Furthermore, the difference between performance data in the two test scenarios showed that the framework could yield better results when applied to areas with only heavy occlusion from plants.

## **7.2 Research Contribution**

Number of lanes is a basic roadway attribute that is needed in many transportation applications. To automatically collect or update the data for large road networks, methodologies developed in road extraction studies are usually applied. The major problem found in the lane models adopted in existing studies is that it must identify complete lane objects, such as lane markings or paved areas. However, the aerial images used for automatic road/lane extraction usually have occlusions that cover part of the target objects. This dissertation developed a profile-based lane model, which requires the



recognition of part of a lane area exposed in aerial images. A mapping technique was also developed to locate the approximate pixel coordinates of road segments in geo-rectified aerial images in order to assist the application of the new model. Furthermore, this dissertation provides an optimization algorithm that estimates the reliability of each road segment and each lane location that the framework predicts.

The framework developed in this dissertation provides a more robust lane model that can be applied to heavily occluded road segments without modeling each non-lane object. It also reduces the travel lane search area to exclude similar objects, such as driveways and parking lots. The reliability estimation function helps reduce the amount of work required for data collection or updates.

### **7.3 Recommendations for Future Research**

Future research to further improve the automatic extraction framework could consider the following three areas:

- For the profile-based lane model, future research could expand lane profile feature studies. Features adopted in this research reflected the fluctuation of brightness for the complete lane, and their brightness similarity to adjacent profiles. It is likely that brightness values in various lane profile locations contribute to the production of different lane models. There may be fewer features needed to reflect fluctuation patterns. Additional features, such as saturation and hue values, could also contribute to lane profile identification. The impacts these features require further study in order to improve the lane model developed in this research.

- The reliability estimation function developed for the optimization algorithm in this dissertation provided valuable extraction process feedback. In reality, since the reliability is calculated based on the assumption that only pavement area profiles are input into the classifier, this measurement level is very likely to be much lower than the actual reliability level. This dissertation implemented the saturation threshold and vegetation threshold to exclude the profiles of noise factors. In future studies, the methodology for pavement area detection could be further explored to more precisely estimate the reliability of extraction results.
- Reliability estimations based on scenarios or types of roadways are valuable statistics for operations. In future studies, the specific reliabilities of the framework for road segments with different importance level or in areas with different traffic volumes could be further explored. The studies on the variations will give valuable guidelines to determining the part of the processed area that may need further manual check.

## REFERENCES

- AASHTO, A. Policy on Geometric Design of Highways and Streets. American Association of State Highway and Transportation Officials, Washington, DC. 2011.
- Ackermann, F., "High Precision Digital Image Correlation", *Proceedings 39<sup>th</sup> Photogrammetric Week*, Stuttgart, pp. 231-243, 1983.
- Altman, Naomi S. "An introduction to kernel and nearest-neighbor nonparametric regression", *The American Statistician*, Vol. 46, No. 3, pp. 175-185, 1992.
- American Society for Photogrammetry and Remote Sensing (ASPRS), *Digital Photogrammetry: An Addendum to the Manual of Photogrammetry*, Bethesda, Maryland, USA, 1996.
- Asrar, G., R. B. Myneni, and B. J. Choudhury, "Spatial Heterogeneity in Vegetation Canopies and Remote Sensing of Absorbed Photosynthetically Active Radiation: A Modeling Study", *Remote Sensing of Environment* 41(2), pp. 85-101, 1992.
- Auclair-Fortier, M. F., D. Ziou, C. Armenakis, and S. Wang, "Survey of Work on Road Extraction in Aerial and Satellite Images", *Technical Report No. 241*. Département de mathématiques et informatique, Université de Sherbrooke, Ontario, Canada, 1999.
- Baltsavias, E. P. and C. Zhang, "Automated Updating of Road Databases from Aerial Images", *International Journal of Applied Earth Observation and Geoinformation*, Vol.6, pp. 199-213, 2005.
- Baret, F. and G. Guyot, "Potentials and Limits of Vegetation Indices for LAI and APAR Assessment", *Remote Sensing of Environment*, No. 35, pp. 161-173, 1991.
- Baumgartner, A., C. Steger, H. Mayer, and W. Eckstein, "Semantic Objects and Context for Finding Roads", *Integrating Photogrammetric Techniques with Scene Analysis and Machine Vision III*, Proc. SPIE 3072, pp. 98-109, 1997.
- Baumgartner, A., C. Steger, H. Mayer, W. Eckstein, and E. Heinrich, "Automatic Road Extraction Based on Multi-scale, Grouping, and Context", *Photogrammetric Engineering and Remote Sensing* 65, pp. 777-786, 1999.
- Breiman, L. and P. Spector, "Submodel Selection and Evaluation in Regression: The X-random Case", *International Statistical Review*, Vol. 60, pp. 291-319, 1992.
- Bresenham, J.E, "Algorithm for Computer Control of a Digital Plotter", *IBM Systems Journal* Vol. 4, No.1, pp. 25-30, January 1965.
- Burges, C.J.C., "A Tutorial on Support Vector Machines for Pattern Recognition", *Data Mining and Knowledge Discovery* Vol. 2, Issue 2, pp. 121-167, 1998.

Burlina, P., R. Chellappa, C. Lin, and X. Zhang, "Context-based Exploitation of Aerial Imagery", *IEEE Workshop on Context-Based Vision*, pp. 38-49, 1995.

Burlina, P., V. Parameswaran, and R. Chellappa, "Sensitivity Analysis and Learning Strategies for Context-based Vehicle Detection Algorithms", *DARPA Image Understanding Workshop*, pp. 577-584, 1997.

Canny, J., "A Computational Approach to Edge Detection", *IEEE Transaction on Pattern Analysis and Machine Intelligence*, Vol. 8, No. 6, pp. 679-698, 1986.

Carlson, T. N. and D. A. Ripley, "On the Relation between NDVI, Fractional Vegetation Cover, and Leaf Area Index", *Remote Sensing of Environment*, Vol. 62, Issue 3, pp. 241-252, 1997.

Chellappa, R., X. Zhang, P. Burlina, C. Lin, Q. Zheng, L. Davis, and A. Rosenfeld, "An Integrated System for Site Model Supported Monitoring of Transportation Activities in Aerial Images", *Proceedings of the 1996 ARPA Image Understanding Workshop*, pp. 275-304, 1996.

Chi, H., C. Lu, F. Zhao, and D. Shen, "Vehicle Detection from Satellite Images: Intensity-, Chromaticity-, and Lane-Based Method", *Journal of the Transportation Research Board*, Vol. 2105, pp. 109-117, 2009.

Cosman, P. C., R. M. Gray, and M. Vetterli. "Vector Quantization of Image Subbands: A Survey", *IEEE Transactions on Image Processing*, Vol. 5, Issue 2, pp. 202-225, 1996.

Cox, D. R. "The regression analysis of binary sequences". *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 215-242, 1958.

Fischer, A., T. H. Kolbe, F. Lang, A. B. Cremers, W. Forstner, L. Plumer, and V. Steinhage, "Extracting Buildings from Aerial Images Using Hierarchical Aggregation in 2D and 3D", *Computer Vision and Image Understanding* 72(2), pp. 185-203, 1998.

Gan, A., J. Shen, and A. Rodriguez, "Update of Florida Crash Reduction Factors and Countermeasures to Improve the Development of District Safety Improvement Projects", Final Report Submitted to FDOT Research Center, Lehman Center for Transportation Research, Florida International University, 2005.

Gevers, T. and A. W. M. Smeulders, "Color-based Object Recognition", *Pattern Recognition* 32, pp. 453-464, 1999.

Grote, A., M. Butenuth, and C. Heipke, "Road Extraction in Suburban Areas Based on Normalized Cuts", *International Archives of Photogrammetry and Remote Sensing*, Vol. 36, Part 3/W49A, pp. 51-56, 2007.

Gruen, A. and H. Li, "Road Extraction from Aerial and Satellite Images by Dynamic Programming", *ISPRS Journal of Photogrammetry and Remote Sensing*, 50(4), pp. 11-20, 1995.

Haag, M. and H.-H.Nagel, "Combination of Edge Element and Optical Flow Estimates for 3D-Model-Based Vehicle Tracking in Traffic Sequences", *International Journal of Computer Vision* 35(3), pp. 295-319, 1999.

Haralick, R. M., K. Shanmugam, and I. Dinstein, "Textural Features for Image Classification", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC-3, pp. 610-621, 1973.

Hastie, T., R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Springer, New York, 2009.

Herman, M. and T. Kanade , "The 3D MOSAIC Scene Understanding System: Incremental Reconstruction of 3D Scenes from Complex Images", *Proceedings of Image Understanding Workshop*, pp. 137-148, 1984

Henricsson, O., "The Role of Color Attributes and Similarity Grouping in 3-D Building Reconstruction", *Computer Vision and Image Understanding* 72(2), pp. 163-184, 1998.

Highway Capacity Manual, Transportation Research Board, National Research Council, Exhibit 23-5, pp.23-6, 2000.

Hinz, S., "Detection and Counting of Cars in Aerial Images", *2003 Proceedings of International Conference on Image Processing*, Vol.2 III, pp. 997-1000, 2003.

Hinz, S. and A. Baumgartner, "Automatic Extraction of Urban Road Networks from Multi-view Aerial Imagery", *ISPRS Journal of Photogrammetry and Remote Sensing*, 58(1-2), pp. 83-98, 2003.

Hinz, S., "Detection of Vehicles and Vehicle Queues for Road Monitoring Using High Resolution Aerial Images", *Photogrammetrie Fernerkundung Geoinformation* 2004 (3) pp. 201-213, 2004.

Hinz, S. and A. Baumgartner, "Vehicle Detection in Aerial Images Using Generic Features, Grouping, and Context", *Pattern Recognition* 2001 (DAGM Symposium 2001), Lecture Notes on Computer Science 2191, Springer Verlag, 2001.

Horn, B.K.P. , *Robot Vision*, MIT Press, pp. 161-163, 1986.

Hu, X., C. Tao, and Y. Hu., "Automatic Road Extraction from Dense Urban Area by Integrated Processing of High Resolution Imagery and LIDAR Data", *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Istanbul, Turkey 35, B3, 2004.

Huertas, A. and R. Navatia, "Detecting Buildings in Aerial Images", *Computer Vision, Graphics, and Image Processing*, Vol. 41, pp. 131-152, 1988.

Irvin, R. B. and D. M. McKeown, "Methods for Exploiting the Relationship between Buildings and Their Shadows in Aerial Imagery", *IEEE Transaction on Systems, Man, and Cybernetics*, Vol.19, No.6, pp. 1564-1575, 1989.

Jin, H., Y. Feng, and Z. Li, "Extraction of Road Lanes from High-Resolution Stereo Aerial Imagery Based on Maximum Likelihood Segmentation and Texture Enhancement", *Digital Image Computing: Techniques and Applications*, pp. 271-276, 2009.

Jin, X. and C. H. Davis, "Vehicle Detection from High-Resolution Satellite Imagery Using Morphological Shared-Weight Neural Networks", *Image Vision Computing*, Vol. 25, No. 9, pp. 1422-1431, 2007.

Kender, J, "Saturation, Hue, and Normalized Colors: Calculation, Digitization Effects, and Use", Tech. Report, Carnegie-Mellon University, 1976.

Kohavi, R., "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection", International Joint Conference on Artificial Intelligence (IJCAI), Morgan Kaufmann, pp. 1137-1143, 1995.

Kozempel, K. and R. Reulke, "Fast Vehicle Detection and Tracking in Aerial Image Bursts", *Object Extraction for 3D City Models, Road Databases and Traffic Monitoring - Concepts, Algorithms and Evaluation (CMRT)*, Paris, France, Vol. 38 (3/W4), pp. 175-180, 2009.

Lanser, S. and W. Eckstein, "A Modification of Deriche's Approach to Edge Detection", *Proceedings of 11th International Conference on Pattern Recognition*, Vol.III, pp. 633-637, 1992.

Leitloff, J., S. Hinz, and U. Stilla, "Vehicle Detection in Very High Resolution Satellite Image of City Areas", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 48, No. 7, pp. 2795-2804, 2010.

Lin, C., A. Huertas, and R. Nevatia, "Detection of Buildings Using Perceptual Grouping and Shadows", *Proceedings CVPR '94, 1994 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pp. 62-69, 1994.

Liow, Y. and T. Pavlidis, "Use of Shadows for Extracting Buildings in Aerial Images", *Computer Vision, Graphics and Image Processing*, Vol. 49, pp. 242-277, 1990.

Liu, G., L. Gong, and R. Haralick, "Vehicle Detection in Aerial Imagery and Performance Evaluation", ISL-Technical Report, Intelligent Systems Laboratory, University of Washington, Seattle, WA, 1999.

Liu, Q. and A. Huete, "A Feedback Based Modification of the NDVI to Minimize Canopy Background and Atmospheric Noise", *Geoscience and Remote Sensing, IEEE Transactions on* 33(2), pp. 457-465, 1995.

Mayer, H. and C. Steger, "Scale-Events and Their Link to Abstraction for Road Extraction", *ISPRS Journal of Photogrammetry and Remote Sensing*, 53(2), pp. 62-75, 1998.

Mayer, H., "Automatic Object Extraction from Aerial Imagery-A Survey Focusing on Buildings", *Computer Vision and Image Understanding*, Vol. 74, Issue 2, pp. 138-149, 1998.

Michaelsen, E. and U. Stilla, "Estimating Urban Activity on High-Resolution Thermal Image Sequences Aided by Large Scale Vector Maps", *IEEE/ISPRS joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, pp. 25-29, 2001.

Moon, H., R. Chellappa, and A. Rosenfeld, "Performance Analysis of a Simple Vehicle Detection Algorithm", *Image and Vision Computing* 20(1), pp. 1-13, 2002.

Movaghati, S., A. Moghaddamjoo, and A. Tavakoli, "Road Extraction from Satellite Images Using Particle Filtering and Extended Kalman Filtering", *IEEE Transactions on Geoscience and Remote Sensing* Vol. 48, Issue 7, pp. 2807 - 2817, 2010.

Myneni, R., G. Asrar, and S. A. W. Gerstl, "Radiative Transfer in Three Dimensional Leaf Canopies", *Transport Theory and Statistical Physics* 19, pp. 1-54, 1990.

National Highway Cooperative Research Program (NHCRP) Report, "Collection and Presentation of Roadway Inventory Data", Transportation Research Board, National Research Council, pp. 7, 2000.

Nemani, R. and S. Running, "Land Cover Characterization Using Multitemporal Red, Near-IR, and Thermal-IR Data from NOAA/AVHRR", *Ecological Applications* 7, pp. 79-90, 1997.

Noronha, S. and R. Nevatia, "Detection and Modeling of Buildings from Multiple Aerial Images", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, Issue: 5, pp. 501-518, 2001.

Papageorgiou, C., M. Oren, and T. Poggio, "A General Framework for Object Detection", *Proceeding of 6th International Conference on Computer Vision*, Bombay, India, pp. 555-562, 1998.

Papageorgiou, C. and T. Poggio, "A Trainable System for Object Detection", *International Journal of Computer Vision* 38(1), pp. 15-33, 2000.

Rajagopalan, A., P. Burlina, and R. Chellappa, "High-Order Statistical Learning for Vehicle Detection in Images", *13th International Conference on Pattern Recognition*, Vol. 3, pp. 900-903, 1999.

Richards, J. A. and X. Jia, *Remote Sensing Digital Image Analysis: An Introduction*, Springer-Verlag, Berlin, pp. 97-99, 2006.

Rouse, J. W., R. H. Haas, J. A. Schell, and D. W. Deering, "Monitoring vegetation systems in the Great Plains with ERTS", *Third ERTS Symposium*, NASA SP-351 I, pp. 309-317, 1973

Ruskoné, R., L. Guiges, S. Airault, and O. Jamet, "Vehicle Detection on Aerial Images: A Structural Approach", *13th International Conference on Pattern Recognition*, Vol. 3, pp. 900-903, 1996.

Salvador, E., A. Cavallaro, and T. Ebrahimi, "Shadow Identification and Classification Using Invariant Color Models", *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, pp. 1545-1548, 2001.

Salvador, E., A. Cavallaro, and T. Ebrahimi, "Cast Shadow Segmentation Using Invariant Color Features", *Computer Vision and Image Understanding*, Vol. 95, Issue 2, pp. 238-259, 2004.

Scanlan, J. M., D. M. Chabries, and R. W. Christiansen, "A Shadow Detection and Removal Algorithm for 2-D Images", *International Conference on Acoustics, Speech, and Signal Processing 1990 (ICASSP-90)*, Vol. 4, pp. 2057-2060, 1990.

Schneiderman, H. and T. Kanade, "A Statistical Method for 3D Object Detection Applied to Faces and Cars", *2000 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'00)*, Vol. 1, pp. 1746, 2000.

Schlosser, C., J. Reitberger, and S. Hinz, "Automatic Car Detection in High Resolution Urban Scenes Based on an Adaptive 3D-model", *Proceedings of the 2nd RSS/ISPRS Joint Workshop on Data Fusion and Remote Sensing over Urban Area*, Berlin, Germany, pp. 167-170, 2003.

Schölkopf, B. and A. Smola., *Learning with Kernels*. MIT Press, Cambridge, MA, 2002.

Shufelt, J. A., "Exploiting Photogrammetric Methods for Building Extraction in Aerial Images", *Proceedings, International Archives of Photogrammetry and Remote Sensing*, 1996, B6/VI, pp. 74- 79, 1996.

Suetens, P., P. Fua, and A. Hanson, "Computational Strategies for Object Recognition". *ACM Computing Surveys* 24(1), pp. 5-60, 1992.

Steger, C., "An Unbiased Detector of Curvilinear Structures", *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (2), pp. 113-125, 1998.



- Tan, T., G. Sullivan, and K. Baker, "Model-Based Localisation and Recognition of Road Vehicles", *International Journal of Computer Vision* 27(1), pp. 5-25, 1998.
- Trinder, J. C. and Y. Wang, "Knowledge-based Road Interpretation in Aerial Images", *International Archives of Photogrammetry and Remote Sensing*, 32(4), pp. 635-640, 1998.
- Tsai, V. J. D., "Automatic Shadow Detection and Radiometric Restoration on Digital Aerial Images", *Proceedings of 2003 IEEE International Geoscience and Remote Sensing Symposium*, Vol. 2, pp. 732 - 733, 2003.
- Tucker, C. J., "Red and photographic infrared linear combinations for monitoring vegetation", *Remote Sensing of Environment*, pp. 127-150, 1979.
- Univeristy of Florida, "Development and Calibration of Highway Safety Manual Equations for Florida Conditions", University of Florida Research Proposal, FDOT Research Center, 2009.
- Vapnik V., *The Nature of Statistical Learning Theory*, Springer, New York, 1995.
- Vapnik, V., *Statistical Learning Theory*, John Wiley and Sons, New York, 1998.
- Vosselman, G. and J. Knecht, "Road Tracing by Profile Matching and Kalman Filtering", *Proceedings of the Workshop on Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pp. 265–274. Birkhaeuser, Germany, 1995.
- Watanabe, S., K. Miyajima, and N. Mukawa, "Detecting Changes of Buildings from Aerial Images Using Shadow and Shading Model", *Proceedings of Fourteenth International Conference on Pattern Recognition*, pp. 1408, 1998.
- Xiong, D. and R. Floyd, "Highway Feature and Characteristics Database Development Using Commercial Remote Sensing Technologies", *Combined with Mobile Mapping, Gis and Gps*. Oak Ridge National Laboratory, pp. 37-43, 2004.
- Yao, J. and Z. Zhang, "Hierarchical Shadow Detection for Color Aerial Images," *Computer Vision and Image Understanding*, Vol. 102, Issue 1, pp. 60-69, 2005.
- Zhang, Q. and I. Couloigner, "Benefit of the Angular Texture Signature for the Separation of Parking Lots and Roads on High Resolution Multi-spectral Imagery", *Pattern recognition letters*, 27(9), pp. 937-946, 2006.
- Zhao, T. and R. Nevatia, "Car Detection in Low Resolution Aerial Image", *Image and Vision Computing*, Vol. 21 (8), pp. 693-703, 2003.
- Zheng, H., L. Pan, and L. Li, "A Morphological Neural Network Approach for Vehicle Detection from High Resolution Satellite Imagery", *Neural Information Processing*, pp. 99-106. Springer Berlin Heidelberg, 2006.

Zhou, J., W. F. Bischof, and T. Caelliz, "Robust and Efficient Road Tracking in Aerial Images", *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (CMRT05), Vol. XXXVI, Part 3/W24, pp. 35-40, 2005.

Zhou, J., W. F. Bischof, and T. Caelli, "Road Tracking in Aerial Images Based on Human-Computer Interaction and Bayesian Filtering", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 61, pp. 108, 2006.

## VITA

### LI TANG

1999 – 2003	B.S., Computer Science Northeastern University Shenyang, China
2003– 2006	M.S., Computer Science Northeastern University Shenyang, China
2006 – 2008	Assistant Engineer Guangdong Planning and Design Institute of Telecommunication Guangzhou, China
2008– 2015	Doctoral Candidate in Civil Engineering Florida International University Miami, Florida
2008 - Present	Student member, Institute of Transportation Engineers
2008 – 2009	Graduate Teaching Assistant Department of Civil and Environmental Engineering Florida International University Miami, Florida
2009 – 2013	Graduate Research Assistant Lehman Center for Transportation Research Florida International University Miami, Florida
2013 – 2015	System Engineer HHT International Consulting, Inc. Coral Gables, Florida

## PUBLICATIONS

1. Tang, L, A. Gan and P. Alluri, “Automatic Extraction of Number of Lanes from Geo-Rectified Aerial Images”, *Transportation Research Record No. 2460, Journal of the Transportation Research Board*, pp. 86-96, 2014.
2. Tang, L, A. Gan, and P. Alluri, “A Framework of Automatic Identification and Extraction of Travel Lane Information from Georectified Aerial Images Using

Support Vector Machine”, in *Proceedings (CD-ROM) of the 92nd Annual Meeting of the Transportation Research Board*, January 2013.

3. Tang, L, A. Gan, and F. Cevallos, “Characteristics of Bus Transit Vehicles in the United States: How They Have Changed Over a Quarter Century”, in *Proceedings (Online) of the 53rd Annual Transportation Research Forum*, March 2012.
4. Gan A., F. Gui, and L. Tang, “System for Transit Performance Analysis Using the National Transit Database”, *Journal of Public Transportation*, Vol. 14. No.3, pp. 87-108, 2011.