國立交通大學

運輸科技與管理學系

博 士 論 文

路側雷達偵測器之自動化學習演算法研發

The Study of Automated Learning Algorithms for Road-Side Radar Traffic Detector

研 究 生：陳昱光

指導教授：卓訓榮　教授

周幼珍　副教授

中 華 民 國 九 十 八 年 七 月

Abstract

Traffic information is essential to effectively perform numerous traffic operations, including travel time estimation or prediction, congestion control policies, and traffic signal control strategies. Traditionally, gathering traffic information is extremely labor and cost intensive, meaning the assistance of advanced information technology equipment is crucial. For overcoming the defects of traditional detectors, the road-side radar detector is adopted in this study, and more advanced learning algorithms are developed to achieve the automation and enhance the accuracy.

The basic traffic information includes traffic count, vehicle classification and vehicle speed estimation. Counting traffic in a single lane is a basic task that can be achieved by using traffic detectors to detect passing vehicles, but it is difficult for road-side radar detectors to simultaneously detect different vehicle types in multi-lane environments, because the signals reflected from passing vehicles in a single lane influence neighboring lanes. The spread of reflected signals created difficulty in accurately identifying lanes. Hence, this study first develops a learning procedure for road-side radar detectors to form an on-line traffic lane estimator. An on-line traffic lane estimator is modelled by Gaussian mixture model (GMM) based on span and conflict information and trained by using the proposed variant of expectation maximization (EM) algorithm. The numerical results demonstrate on-line traffic lane estimator can work well in real-world scenarios, and the accuracy of traffic lane estimator is verified by counting traffic in different lanes.

Besides, the real-time vehicle classifier for road-side radar detectors in multi-lane environments is for the first time presented. A two-dimensional Gaussian Mixed

Model is employed to develop the learning model based on FMCW radar data. An EM algorithm is thus implemented to maximize the likelihood of the formulated learning model; consequently, the model could be used for classifying small and large vehicles in multi-lane environments simultaneously, so that traffic information can be obtained at a relatively lower cost. In the suburban field test, the accuracy of real-time vehicle classifier in multi-lane environments can achieve more than 88%.

Acknowledgements

I would like to thank Dr. Hsun-Jung Cho and Yow-Jen Jou, my supervisors, for their many suggestions and constant support during this research. I am also thankful to committee members and Dr. Da-Yin Liao for his guidance through the early years of chaos and confusion.

I should also mention that my graduate studies in Taiwan, R.O.C. were supported by MOE ATU Program of the Ministry of Education, Taiwan, R.O.C. and the National Science Council.

Of course, I am grateful to my parents for their patience and *love*. Without them this work would never have come into existence (literally). In addition, my aunties also do me a great of help in my study life.

Finally, I wish to thank the following: Pei-Jen Yeh (for her generousness and love); Dr. Shu-Cherng Fang and Dr. Yi-ming Li (for their knowledge of optimization), Dr Chia-Tung Lee (for his philosophy in life), Dr. Gau-Rong Liang (for his globalized economic view), Dr. Shie-Yuan Wang (for his knowledge of simulation) and many NCTU faculty (for sharing with their knowledge); Heng Huang, Akin, Banana, Urnus, Amie, Tesseric; (for their help in developing radar detectors); Isoguava, Hogan, ... (for all the good times in basketball court and gym); all the Cho-lab members (for all the good and bad times we had together); *and* my brother (because he asked me to).

National Chiao Tung University, Hsinchu Yu-Kuang Chen
June, 2009

Table of Contents

Bibliography **79**

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Traffic information is essential to effectively perform numerous traffic operations, including travel time estimation or prediction, congestion control policies, and traffic signal control strategies. Traditionally, gathering traffic information is extremely labor and cost intensive, meaning that the assistance of advanced information technology equipment is crucial. As IT develops, numbers of detectors have increased. All of this IT equipment has been developed to replace traditional detectors owning to its inexpensiveness, installation flexibility.

Generally speaking, traffic count technologies can be classified into two categories: the intrusive and non-intrusive methods. The former consist of a data recorder and a sensor placing on or in the road. These detectors have been employed for many years and are briefly introduced as follows [1][2][3][4]:

- Pneumatic road tubes: rubber tubes are placed on road lanes to detect passed vehicles from pressure changes. The pulse of air is detected and recorded by a

traffic counter. The main drawback is that its efficiency is subject to temperature, and low speed flows.

- Piezoelectric sensors: the principle is to convert the mechanical energy to electrical signal. The amplitude of the signal is proportional to the degree of deformation.

- Magnetic loops: it is the most common technology applied to collect traffic data. The loops are embedded in roadways in a square formation that generates a magnetic field. This has a generally short life expectancy because it is usually damaged by heavy vehicles, but is not affected by weather conditions. The main drawback are the implementation and maintenance costs can be expensive.

Not-intrusive detectors are remote observations. Even if manual counting is the most common way, some new technologies have recently emerged and should be noted.

- Manual counts: it is the most labor-intensive way. The trained observers collect traffic flows, pedestrians and vehicle types by using tally sheet, mechanical count boards to record.

- Passive and active infra-red: the presence, speed and type of vehicles are detected and recorded based on infrared equipments. The main drawbacks include bad weather, and limited lane coverage.

- Passive magnetic: it is also embedded in roadways, and it can count the number of vehicles, type, and speed.

- Ultrasonic and passive acoustic: these detectors transmit sound waves to detect vehicles by measuring the round trip time. The ultrasonic sensors are placed

over the lane, and the passive acoustic sensors are placed alongside the road, respectively. The main drawback of both detectors is that the performance is affected by temperature and bad weather (e.g. low temperatures, snow).

- Microwave radar: it can detect moving vehicles, speed and type, and the main advantage is the performance is not affected by weather conditions, lights.

- Video image detection: video cameras record vehicle numbers, type and speed by different image processing technologies (e.g. detecting and tracking). The main drawback is that video detectors are very sensitive to light and weather conditions.

Different detectors have different advantages and limitations. In sum, ultrasonic or acoustic sensors in noisy environments are not suitable for use. For inductive loop or magnetic sensors, the installment usually spoils the road pavement and tends to be damaged by passed vehicles. Video and laser sensors are deeply affected by the different conditions of weather and lights. Though these sensors already have high accuracy in traffic count and vehicle type classification [5][6][7], most of sensors (except for video ones) are limited to apply to a single lane environment, such that installation and maintenance costs in cities become very high.

Regarding capability of being applied to multi-lane environments, only video-based and radar-based detectors have made great progress. Notably, the video-based and radar-based detectors are capable of providing traffic counts, real-time vehicle speed, and vehicle classification simultaneously in multi-lane environments. In other words, detectors are not designed for a specific function. The development of video-based detectors are earlier than radar ones, because video equipments are easier to

be acquired, and image processing techniques are much mature. For video-based detectors, except for the effects arose from lights and weather, several environmental variations will affect the accuracy of vehicle analysis. First, vehicle shadows degrade the accuracy of video-based detectors, and lead to erroneous traffic counts. Second, the camera perspective makes vehicles geometry features inconsistent, hence, a calibration process is required in advance to fit in with video algorithms [8][9].

1.2 Characteristics of Radar-Based Detectors

Actually, the characteristics of radar-based detectors are better than that of video-based ones. For examples, radar detectors are not influenced by weather effects (such as, heavy fog, rain) and light effects (such as, shadow, dark); hence radar-based detectors eliminate the need for complex algorithms for handling these scenarios in video-based ones. Additionally, radar-based detectors do not need rigorous fine adjustments in setup, and the installation process does not spoil the road pavement. Owing to its high adaptability and inexpensiveness (in multi-lane environments), radar detectors become one of the popular equipments used in traffic surveillance system gradually.

The installation methods of radar-based detectors are classified into two major categories, including forward-looking and side-fire (or road-side) ways. The forward-looking detectors have the illuminative direction parallel to the direction of traffic, and such an installation method is only applied to a single lane. On the other hand, the side-fire approach means that the illuminative direction of radar detectors is perpendicular to the traffic direction shown in Figure 1.1. The main differences in function between video-based and radar-based detectors include (i) the forward-looking

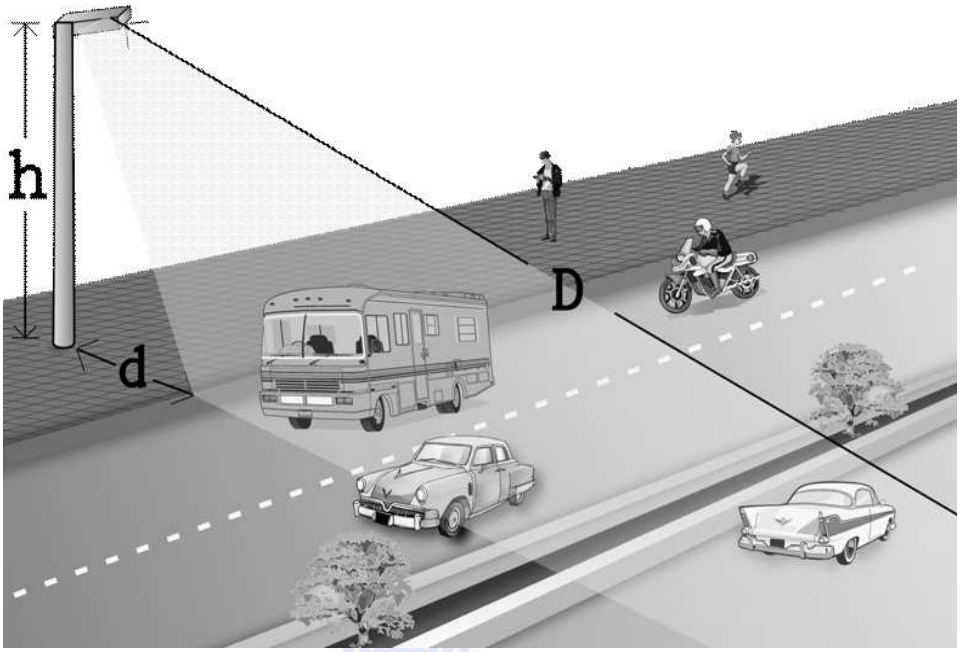


Figure 1.1: Sketch map of a multi-lane scenario. The road-side radar detector is mounted at height h, and its illuminative direction is perpendicular to the traffic direction. The least h must be above the height of passed vehicles in the nearest lane for preventing the far reflected signals are entirely blocked by the near vehicles.

approach is the most common installation method for video-based detectors to apply to multi-lane environments. (ii) Video-based detectors usually adopt background subtraction techniques to sense vehicles from a video sequence without knowing the locations of traffic lanes in advance [8][10].

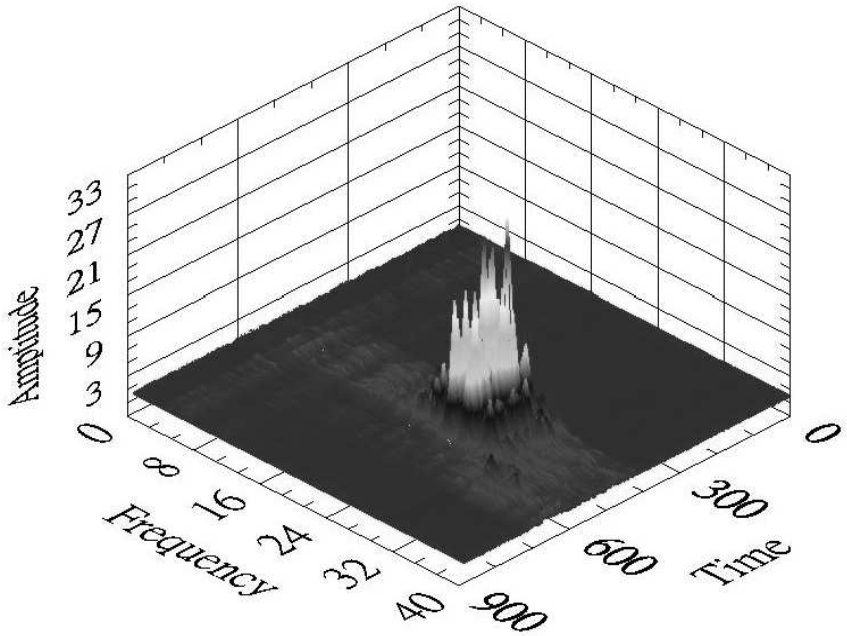
Figure 1.1 is a sketch of a multi-lane environment, and indicates that the road-side radar detector is mounted at height h and at distance d from the nearest lane, while the maximal distance for detecting passing vehicles is D. A road-side traffic detector transmits and receives electromagnetic signal across its illuminative field. The relative distance and the corresponding magnitude information among targets and a detector can be obtained by using a Fast Fourier Transform (FFT) to convert

the received time-domain voltage values into the frequency-domain information.

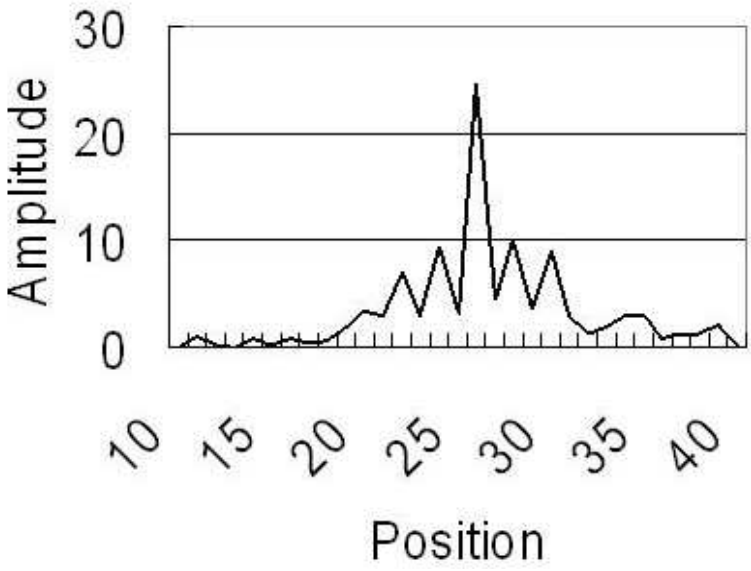
Figure 1.2(a) shows a 3D frequency-domain information of a passed vehicle, and Figure 1.2(b) shows a frequency-domain frame which is a snapshot of Figure 1.2(a). The magnitude of each FFT bin shows the amount of energy of the received signal at a particular frequency, and which depends on the geometry, angles, relative distances and material characteristics of a target vehicle. The frequency shown in Figure 1.2(a) represents the relative distance; hence the vehicle's position can be directly estimated from the frequency-domain information. For convenience, all figures mentioned later will use position rather than frequency to represent the relative distance. The panel data shown in Figure 1.4 represent a large vehicle in lane 2 and a small vehicle in lane 6, and depict the energy distributions of large and small vehicles. Figure 1.4 also shows that the spread of reflected signals of a large vehicle is so large that neighbor lanes are influenced.

1.3 Problem Description

Road-side radar-based systems have so much advantageous features; nevertheless, road-side radar-based systems do not have very reliable and robust algorithms in the vehicle detection application. Restate, road-side radar-based systems still encounter various different difficulties in attaining the following functions, including traffic count, vehicle type classification, and speed estimation in multi-lane situations, owing to the spread of the reflected signals resulting from limitations in bandwidth and equipment. It is known that the accurate lane boundary positions are required in multi-lane situations because the boundary positions are fundamental information to carry out all those functions such as a vehicle classifier and speed estimator. Since



(a)



(b)

Figure 1.2: Frequency-domain information: (a) 3D pattern of a passed vehicle; (b) Frequency-domain frame, which is a snapshot of (a), and the scale of x-axis is approximately 78 cm thick.

Figure 1.3: The top figure shows the snapshot of the received voltage signals when the bus shown in the image was within the illuminative field of the road-side radar, and the bottom figure shows the frequency-domain frame that is obtained by converting received voltage signals via FFT.

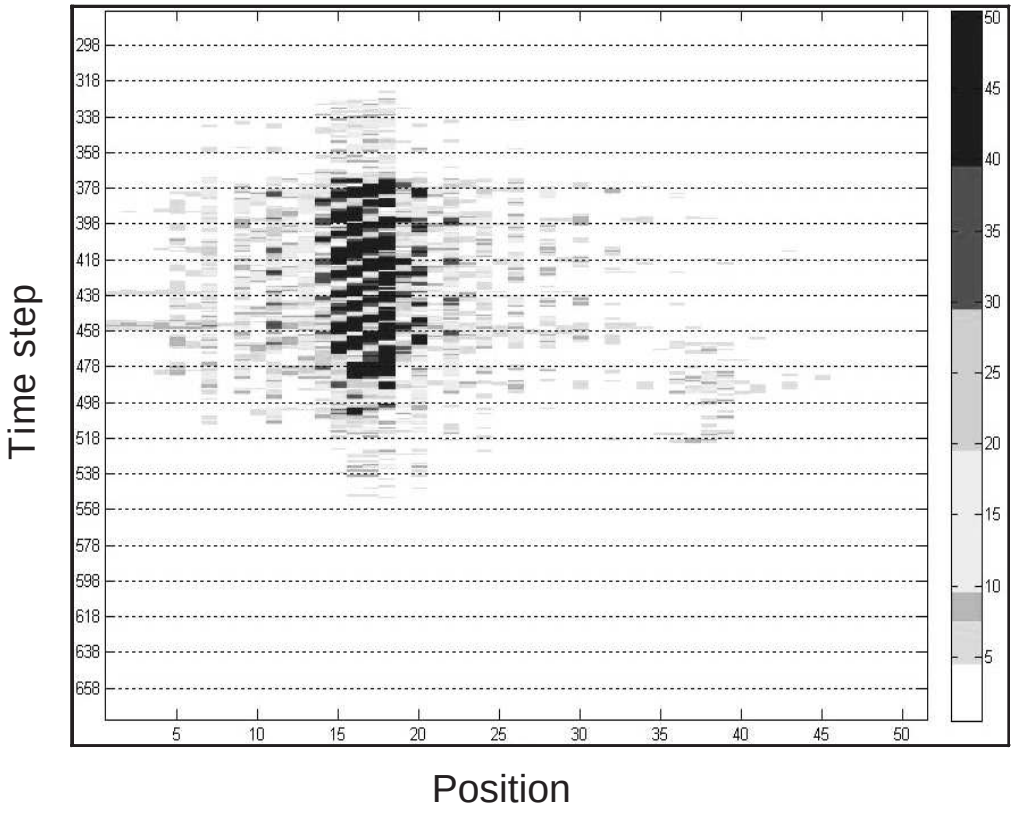


Figure 1.4: Energy comparisons between large and small vehicles. This panel data represent that these exists a large vehicle in lane 2 and a small vehicle in lane 6.

current products do not have robust algorithms to depict the lane boundary positions in various multi-lane situations, such that follow-up functions, such as vehicle type classification, and speed estimation, do not have good performances.

Additionally, the problem of the classification of vehicles arises in many areas such as truck volume estimating, traffic planning, roadway tolling, etc. Thus, how to design a vehicle classifier for distinguishing a large-sized vehicle from a small-sized vehicle in multi-lane situations is very important in software design for road-side radar-based systems.

1.4 Research Scope

The algorithms are proposed for the road-side radar detector to enhance the accuracy of basic functions in multi-lane situations, and meet the following goals.

1. Detect signal variations caused from the environment and update the background information properly.

2. Analyze the defects of the current lane boundary estimation algorithm, and re-design a much accurate one.

3. Analyze and retrieve the characteristics of the passed vehicles, and design a real-time vehicle classifier based on the selected features.

The first problem usually results from intruding objects and the stability of equipment. If we have statistical model of the scene, an external intruding object usually spot the parts of signals that don't fit the model, such as a parked vehicle. Another reason for background variation is the equipment itself, since generated signals would

be deeply affected by working temperature and the stability of power supply. Consequently, for minimizing these uncertainties, a background adapter is proposed, and this process is usually known as "background subtraction".

The second problem is how to design a real-time and highly accurate traffic counter in multi-lane situations. For counting vehicles in each lane, the road-side radar should realize where the lane position is, consequently, the development of a lane boundary estimator is required. It is well known that numbers of equipments are designed for a single-lane traffic count, because such a design carried out in a much easier way. However, the quality of current equipments which are able to provide a multi-lane traffic count function still has a great room to improve, because the signals reflected from passing vehicles in a lane influence neighboring lanes.

Lastly, vehicle classification is a very important function for road planning, traffic signal control, and road surveillance applications, because these applications provide useful information for intelligent transportation systems (ITS) to enhance the efficiency of traffic operations. Hence, a real-time vehicle classifier is one of the necessary functions to be carried out.

1.5 Research Framework and Organization

The research framework and organization are shown in Figure 1.5, which describes the required functions of a road-side radar system. Chapter 2 will introduce the related reviews and Chapter 3 introduces an overall system architecture including hardware and software design. After that, three main algorithms will be revealed in Chapter 4, 5, 6, respectively. The aim of Chapter 4 is to develop a background adapter to accommodate environmental changes, such as an intruding parked car or

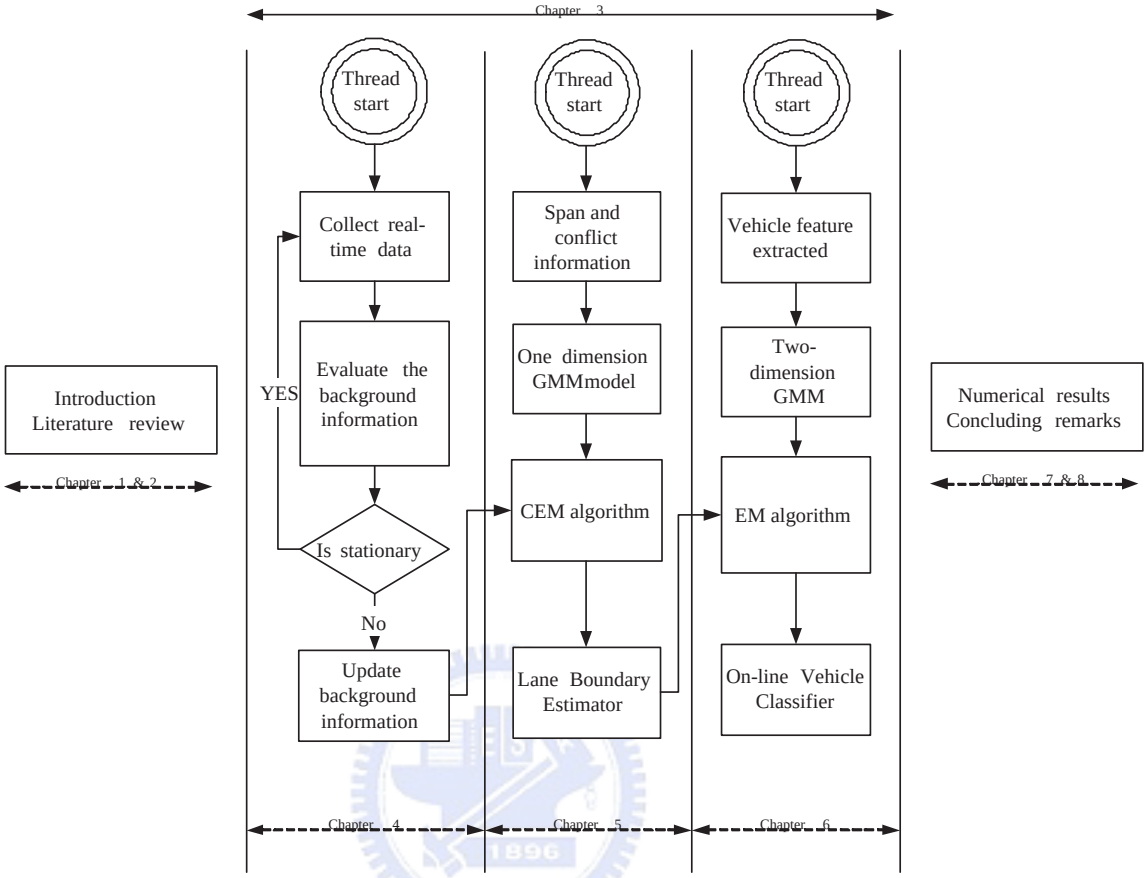


Figure 1.5: The organization of this study

temperature variation. The goal of Chapter 5 is to propose two novel features, span and conflict information, to work with one-dimensional Gaussian Mixed Model, and a classification EM algorithm is also introduced to develop a lane boundary estimator. A two-dimensional Gaussian Mixed Model is proposed in Chapter 6 and employed to develop the learning model based on FMCW radar data, and the model could be used for classifying small and large vehicles in multi-lane environments simultaneously. Finally, Chapter 7 and 8 summarize numerical results in the field test and concluding remarks, respectively.

Chapter 2

Literature Review

In this section, the related literature reviews are introduced. The first part is to review the design of lane boundary estimator for multi-lane environment. However, there is few literatures associated with this topic, since this function may not have widespread applications in real-world. So far there are only two patents that are claimed by the current marketing road-side radar manufacturers.

The second part is to review the design of radar-based vehicle classifiers no matter forward or road-side installations. In addition, the feature selected for classifying vehicles are also reviewed and discussed in detail.

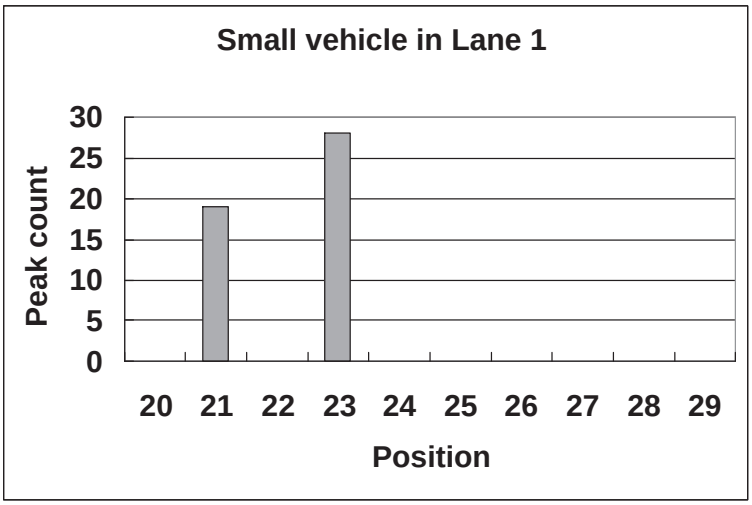
2.1 Literature Review for Lane Boundary Estimation

Essentially, developing an on-line automatic traffic lane estimator can be considered as a clustering problem, as is demonstrated in the patent [11]. This patent uses the same frequency-domain information mentioned above to decide the lane position. After obtaining the similar frequency-domain domain information shown in Figure

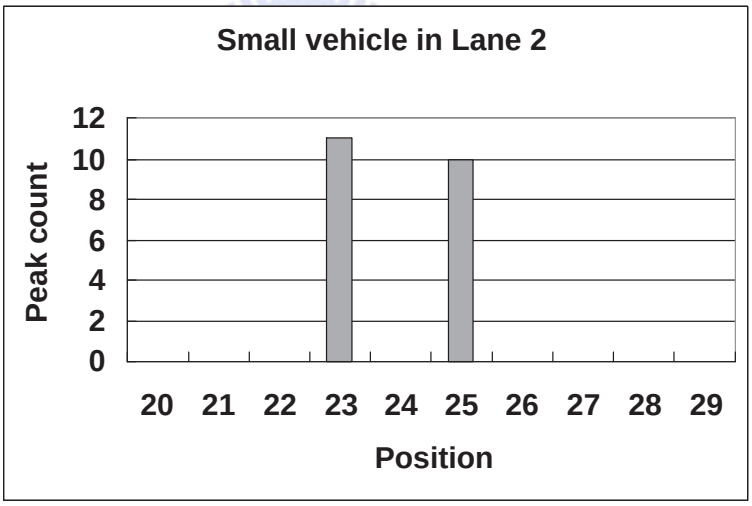
1.2, the patent utilizes the peak information to form the vehicle feature. The bin with the maximal amount of energy is termed the peak, and the bin with most peak count is used to represent the corresponding vehicle position. In other words, a single passing vehicle contributes just one vehicle count to the corresponding bin. When the sufficient vehicle data are gathered, Gaussian mixture model (GMM) is introduced to model the probability density of lane position. The peaks of probability function describe the lane center, while the low density parts of GMM distribution represent lane boundaries. The description of lane boundary of radar is just to verify that the passed vehicle belonged to which lane, in fact, the description of lane position may not correspond to the real-world lane position.

However, representing a vehicle position by adding just one vehicle count to a bin may involve bias risk, and such an expression may easily result in misleading and incorrect lane outcomes. The examples shown in Figure 2.1 are selected from real-world data, where Figure 2.1(a) and (b) show the peak counts when small vehicles passed through the detection zone in lanes 1 and 2, respectively. According to the vehicle feature proposed in the patent [11], both Figure 2.1(a) and (b) added one vehicle unit to bin position 23, respectively. In other words, even if Figure 2.1(a) and (b) had different bin spectrums arising from the different vehicles in the different lanes, both cases provided the same feature expression that possibly caused ambiguity in separating the two lanes. The detailed outcomes of the patent [11] will also be exhibited in the latter section.

The patent [12] adopts another approach to describe the lane position based on a set of lane center variables, and utilizes displacement information to adjust the lane center. Restated, every reflected signal when a vehicle passed through the detectable



(a)



(b)

Figure 2.1: Examples illustrating the accumulated peak count scenario arose from two vehicles passed through the detection zones in two lanes. (a) indicates the accumulated peak count when a small vehicle passed through the detection zone in lane 1, it reveals that position 23 has the maximum bin count. (b) indicates the scenario where a small vehicle passes through the detection zone in lane 2, but position 23 still has the maximum bin count.

area could comprise zone center information. If the newly formed information is sufficiently close to the existing zones, the nearest zone is updated; otherwise, a new zone is defined based on this new formed information. However, the patent [12] easily result in misleading and incorrect lane number outcomes because of the fixed lane width, consequently, follow-up human interventions are usually required. Although this patent has the advantage of not requiring lane number input in advance, it has the weakness that incorrect estimates of lane number make it necessary to spend a long time adding/removing corresponding lanes and adjusting lane positions.

Both of the above patents transformed reflected signals arising from a passing vehicle into single-value information (such as bin position, lane center), as may commit bias in estimating the lane position. The main difference between the two patents is that the former is a batch algorithm, while the latter is a real-time algorithm.

This study focuses on designing an online traffic lane estimator for a road-side radar detector in multi-lane environments. New features (such as span information and conflict information) are introduced to describe lane positions such that the ambiguity in feature expression can be eliminated. Additionally, the variant of EM algorithm is proposed to achieve classification and reduce the variance of each Gaussian component.

2.2 Literature Review for Radar-Based Vehicle Classification

Roe et al. (1992) [13] used FMCW radar (10.525GHz) to classify vehicle types, including bicycle, car, light goods, medium goods, heavy goods, and buses based on

Table 2.1: Literature review for vehicle classification based on radar-based systems

Authors	Radar Type	Features	Installation
Roe and Hobson	FMCW	length & height	over-head
Sang et al.	FMCW	BW & received power	side-fire
M. Cherniakov et al.	FSR	received power	side-fire
Ildar Urazghildiiev	down-looking spread	height profiles	over-head
Jianxin et al.	FMCW	doppler signature	over-head

length and height of cars, and accuracies of 75% in single lane are achieved. Park et al. (2003) [14] developed a K-band (around 24GHz) side-fire FMCW radar for a single lane application, and vehicles are classified as large, medium or small size based on received power and spectrum pattern. Cherniakov (2005) [6] use forward scattering radars (FSR) for classifying ground vehicles, but the transitive and receiving radar antennas are located at different sides of road. Principle component analysis is utilized to reduce dimensionality in frequency-domain feature vector, and the K-nearest is employed to form a classifier. The experimental cars are distinguished into three vehicle categories, including small, medium, and large. The experimental results show that recognition accuracy is only 71% when using the main lobe, while it keeps at 86% without the main lobe. Urazghildiiev et al. (2007) [5] classified vehicles based on vehicle height and length in a single lane, and this exhibits that the signals have the rough geometric pattern similar to the side view of vehicle. The key point of this study is to develop a real-time vehicle classifier for "road-side" equipments in "multi-lane environments". (1) The road-side equipment is perpendicular to the traffic direction, which makes Doppler effect unapparent to be utilized. (2) The road-side equipment receives the signals of all lanes which requires more sophisticated model and algorithm to describe and analyze our data.

After reviewing the developments of vehicle classification, it is known that, for

real-world multi-lane applications, there are few significant results and field tests in vehicle classification based on road-side radar detectors. Moreover, using road-side radar detectors for classifying small and large vehicles in multi-lane environments simultaneously will lower the cost of acquiring the traffic information, consequently, this study develops mathematical modelling as a working tool based on FMCW radar data to analyze the vehicle types in multi-lane environments.

Chapter 3

Road-Side Radar System Design

Real-time systems comprises all devices with different constraints and limitations. In general, we can classify these constraints into two categories, that is,"Hard deadlines" and "Soft deadlines". Hard deadlines are constraints that absolutely must be met. A missed deadline constitutes an erroneous computation and a system failure. In these systems, late data are bad data. On the contrary, if late data are still good data, such a system is referred to a "soft" real-time system. In this study, when the term real-time is used alone, we specifically referring to hard real-time systems [15].

In addition, this study only focuses on the details of software design, thus the hardware design is out of our research scope. However, our proposed softwares are working on the certain requirements of hardware, so these requirements of hardware design will also be briefly introduced in the following section.

3.1 Hardware Architecture and Design Specifications

An X-band CMOS-based frequency-modulation continuous-wave (FMCW) radar system is proposed by Tzuang et al. [16] and adopts two antennas to transmit and receive signal separately. The radio frequency (RF) transceiver is formed by standard 0.18 μm one-poly six-metal (1P6M) complementary metal-oxide semiconductor (CMOS) technology and is embedded in a chip area of 1.68 mm $\times$ 1.6 mm. The several significant properties of the FMCW radar are described as follows. First, the linearity of VCO is 3 % while the output frequency is ranged from 10.3 GHz to 10.8 GHz. Second, the on-chip isolation between the transmitting and receiving paths is 55.0 dB at 10.5 GHz, additionally; two planar leaky-mode antenna arrays with a gain of 18 dB are designed. Experiments indicate the isolation between two antenna arrays with a spacing of 5.0 mm is higher than 42.0 dB at 10.5 GHz.

Figure 3.1 illustrates the block diagram of the X-band FMCW radar system, which comprise dual planar antenna arrays at the transmitter output and the receiver input; a 0.18 μm 1P6M CMOS transceiver is responsible for the RF signal processing as shown inside the dashed lines, and a baseband digital signal processing unit is designated for instantaneous and simultaneous assessment of range measurements. Notably, using two antennas rather than only one eliminates the circulator, as is expensive and does not provide sufficient isolation between the transmitter and the receiver.

In the range measurement, the FMCW radar transmits the signals, and receives the reflecting waves from an intruding vehicle. The amplitude of the input triangular

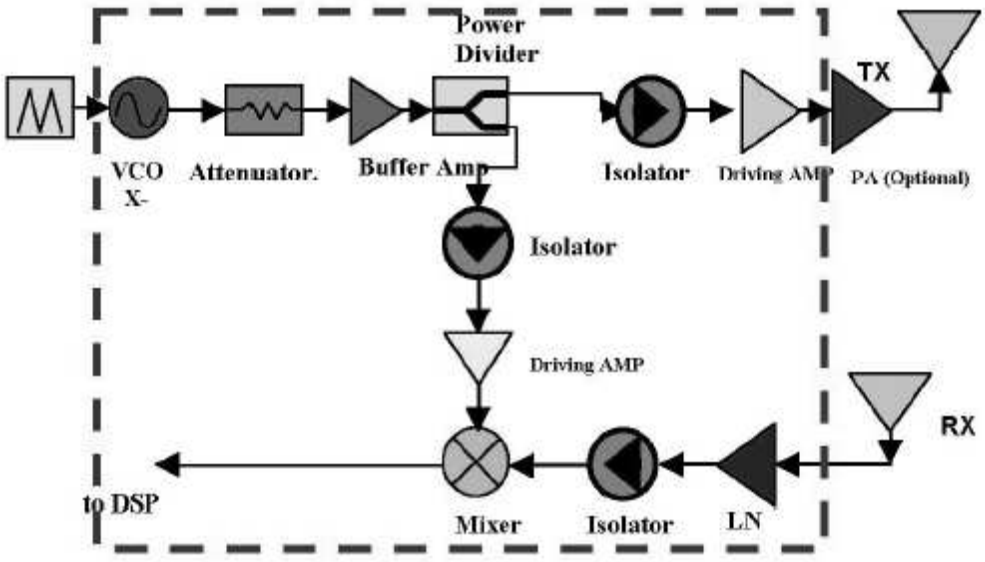


Figure 3.1: Block diagram of the X-band FMCW radar system which is composed of two antenna arrays, a single-chip CMOS transceiver (enclosed by the dashed line) and an external digital signal processing unit. Power amplifier is added to increase the output power level [16]

wave directly controls the output frequency of the transceiver. The bandwidth of the modulated signals at the transmitting path, denoted by BW, is equivalent to the frequency difference between the maximum and minimum output frequencies of the voltage controlled oscillator (VCO). Moreover, the modulated-bandwidth over the half-period of the triangular-wave, denoted by S_b, can be calculated by

$$S_b = \frac{BW}{t_m/2} \tag{3.1.1}$$

where tm denotes the period of the triangular wave, and the parameters are set when performing the signal processes. The pulse-repeated-frequency (PRF) is 500.0 Hz; the swept time (Ts) is set to 2.0ms and the period of triangle-wave t_m is 0.5 ms.

The beat frequency, denoted by f_b, is the frequency difference between two input

signals applied to the mixer. If the round trip time for signal propagation between the sensor and the target is less than that of a half-period of a triangular wave, then S_b can be maintained at a constant value. Suppose the maximum distance R between the radar and the target is set 60 meter, and the round trip time τ is described as

$$\tau = \frac{2R}{c}, \tag{3.1.2}$$

where c represents the speed of light in air. Consequently, the maximum round trip time τ is $0.4\mu m$. The beat frequency f_b shown in (3.1.3) also can be obtained if any reasonable round trip time is given.

$$f_b = \frac{2 \times BW}{t_m}\tau. \tag{3.1.3}$$

On the contrary, the quantity of R also can be calculated by (3.1.4) if f_b is given.

$$R = \frac{c \times t_m \times f_b}{4 \times BW}. \tag{3.1.4}$$

The ideal range resolution R_0 of the FMCW radar can be estimated as

$$R_0 = \frac{c}{2 \times BW}, \tag{3.1.5}$$

thus, range resolution R_0 is 1.0 meter. Equations (3.1.1) to (3.1.5) illustrate the principles of FMCW radar for range measurements. These operations are based on the assumption that the frequency modulation is linear. In other words, the slope of modulation-bandwidth, S_b, must be maintained at a constant value during the half-period of the input triangular wave. If the output frequency of VCO is not linearly proportional to the amplitude of input triangular wave, then the linear transfer function between f_b and τ shown in (3.1.3), is not valid. Consequently, the range

resolution of the FMCW radar is degraded. Moreover, the signals at the transmitting and receiving paths are assumed to be independent of each other in the FMCW radar. The coupling effects between the transmitting and receiving paths are not included in the above derivations, hence, such leakages could significantly reduce the performance of the FMCW sensor. The design issues, which include linearizing the generating frequency modulating signal and eliminating undesired couplings in the radar system, are significant challenges for the proposed single-chip CMOS transceiver.

Figure 3.2 illustrates the illuminative field of the designed antenna [16] from top and side views. The angle ψ shown in Figure 3.2(a) is called elevational plane beamwidth, and is asked to be large, because the larger angle ψ represents that more lanes can be covered. The angle θ shown in Figure 3.2(b) is called horizontal plane beamwidth, and is expected to be small. The design reasons for the horizontal plane beamwidth as θ can be described as follows. First, if the horizontal plane beamwidth displayed as a large angle $\theta^{'}$ shown in Figure 3.2(b), the information about vehicles in which lane can not be caught. Take vehicles A and B for example, vehicles A and B will form the similar frequency-domain information theoretically because both vehicles have the same relative distance to a detector. Therefore, the large horizontal plane beamwidth $\theta^{'}$ makes different vehicles' reflected signals overlap in the same position such that vehicles in which lane can not be recognized. Second, considering the farthest lane (relative to a road-side radar detector), if the horizontal plane beamwidth is $\theta^{'}$, the illuminative field in the farthest lane may contain more than one vehicle, such that road-side radar detectors can not separate vehicles one by one. Take vehicles B and C for example; vehicles B and C will be misled as an only vehicle if the horizontal plane beamwidth is $\theta^{'}$, then the erroneous traffic count occurs in that

Table 3.1: Comparison of the current commercial products

	RTMS	Smart Sensor Model 105	The designed product
Elevation	45°	80°	55°(35° ∼ 90°)
Azimuth	15°	12°	20°
Detection range	3 ∼ 60 m	3 ∼ 60 m	3 ∼ 60 m
Range resolution	2m	3m	3m
Time accuracy	10.0 ms	2.5 ms	2.0 ms
Protocol	RS-232 or RS-485	RS-232 or RS-485	RS-232
Power consumption	4.5W	7.5W	5W
Temperature	−37° ∼ 74°C	−40° ∼ 75°C	
Size	16 x 24 x 12 cm	32 x 23 x 7.6 cm	29.3 x 12.2 x 25.5 cm
Weight	2.2 kg	< 2.27kg	3.5 kg

lane. In a word, the illuminative field of the antenna determines the position of the farthest lane.

Table 3.1 shows the comparison of the current commercial products with the proposed radar sensor in this research. There is still room for improvement to the same grade products. Besides, the quality of materials and the shell over the detector may affect antenna radiation patterns such that the farthest distance is unable to be detectable. Additionally, antenna radiation patterns of dual planar antenna arrays are shown in Fig. 3.3 and 3.4. The 3dB beam width of the antenna array was 20° in the E-plane, and the main beam with a gain of 18.0 dB is at 56° in the H-plane.

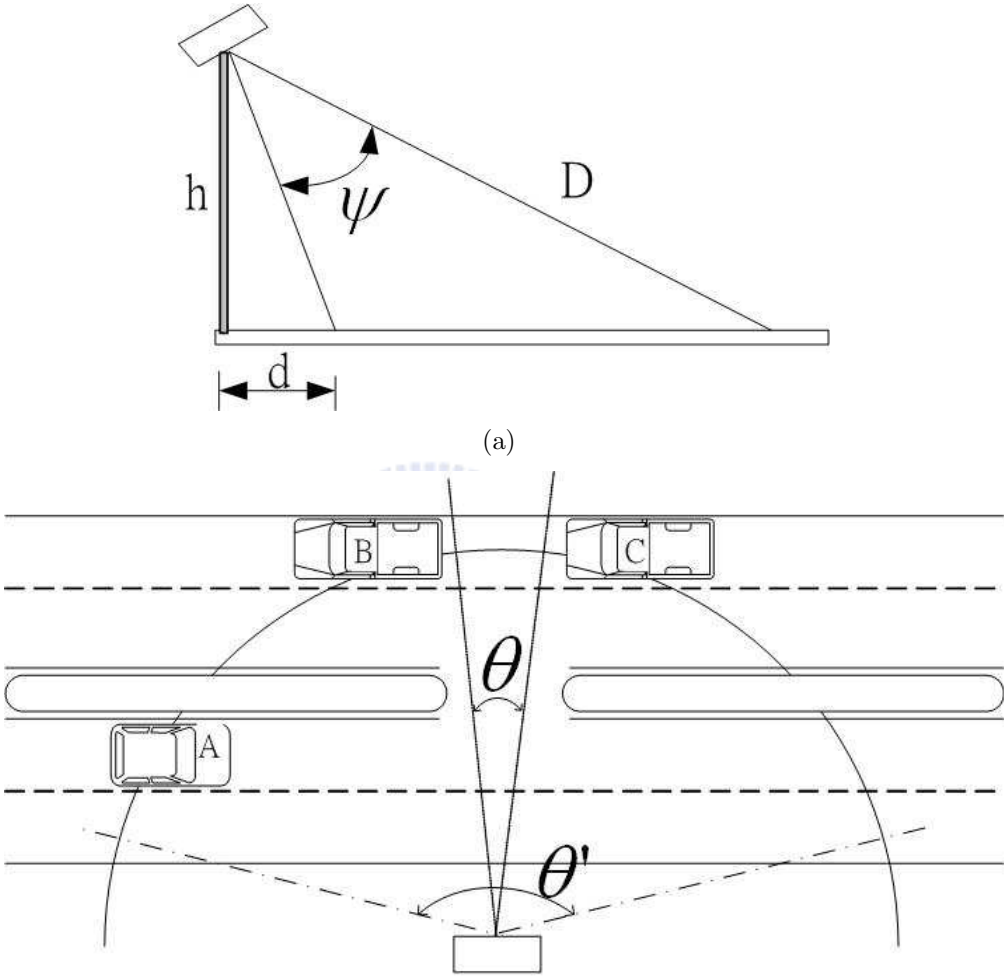


Figure 3.2: (a) A side view of road-side radar detector in multi-lane circumstances. (b) A top view of road-side radar detector in multi-lane circumstances. There are a small angle θ and a large angle $\theta^{'}$.

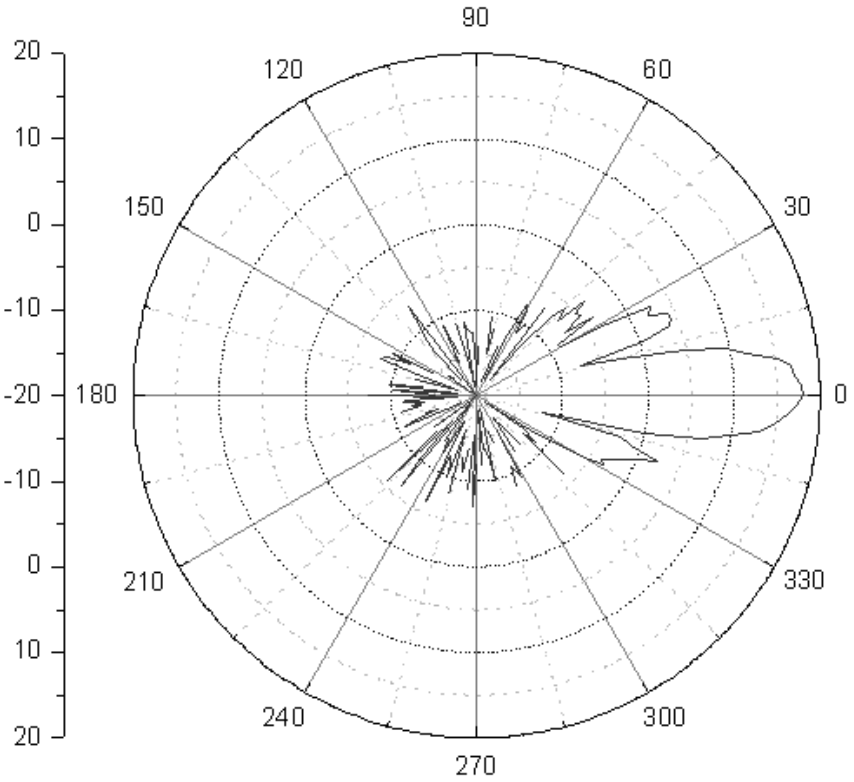


Figure 3.3: E-Plane (Microstrip Leaky-EH1-Mode Antenna Array) [16]

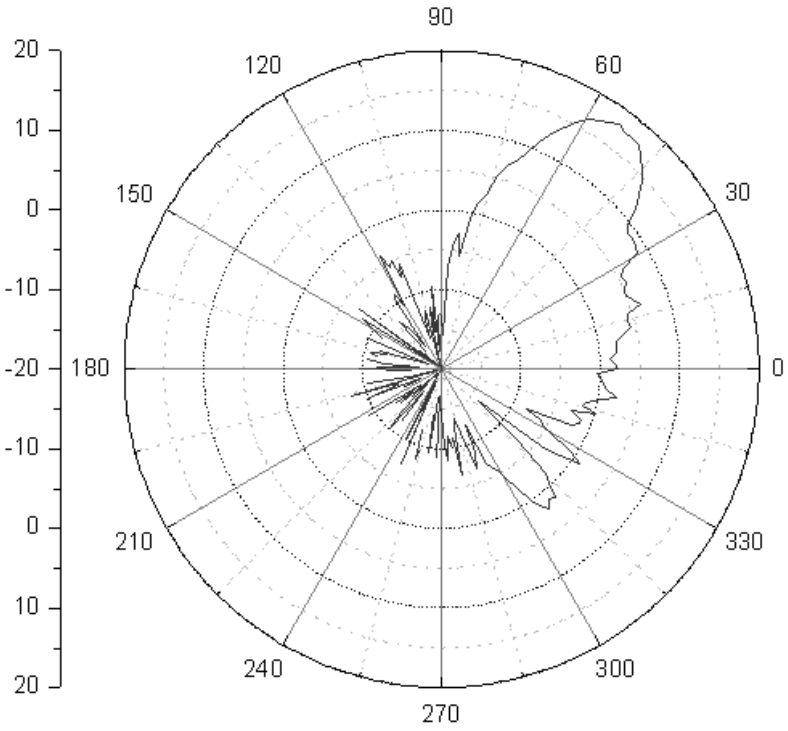


Figure 3.4: H-Plane (Microstrip Leaky-EH1-Mode Antenna Array) [16]

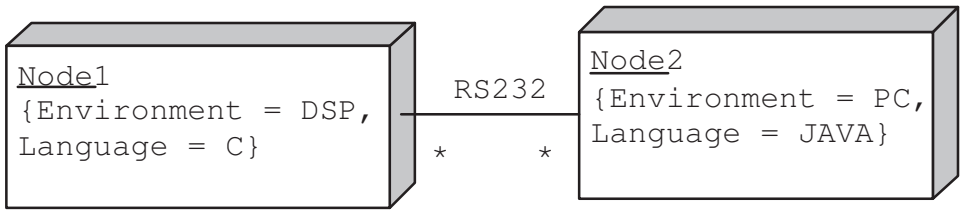


Figure 3.5: The C code is deployed on a DSP device, while the java code is deployed on a PC. The devices are connected by RS232, and communicated by the designed interface.

3.2 Software Architecture

For simplifying the developing complexity and analyzing the signals, real-time modules, video and signal retrieval interface cards, are equipped and integrated into an industry PC for simultaneously recording real-world signal and video data. The collected data can be replayed and analyzed through some integrated development environment (IDE), and also can be an off-line input of alpha testing.

3.2.1 Developing Tools and Environment

Our real-time developers use GNU plug-ins embedded on eclipse IDE, and then develops the operational C codes that are further executed in the target platforms. For enhancing the developing efficiency, the Concurrent Versions System (CVS) is introduced to help us integrate different versions of functions. For example, bugs sometimes creep in when software is modified, and you might not detect the bug until a long time after you make the modification. With CVS , you can easily retrieve old versions to see exactly which change caused the bug.

In addition, analysis tools and graphic interfaces are developed based on Java language (jfreechart-1.0.8), and the Java programs can be executed at any machine

with Java virtual machine (JVM). The communication protocol between the target platforms and analysis tools is defined, and is connected by RS-232 physically. The physical and software deployments is shown in Figure 3.5.

3.2.2 Requirements of Software Design

Based on the sampling capability of equipment ($200k \sim 250k$) and the designed range resolution (in our experimental design, the bin range is approximately 78 cm thick), there are approximate $400 \sim 500$ FFT frames generated per second. The sampling rate is limited to the chosen equipment, thus the performance of software is the most important key to be noticed. According to the equation (6.1.2) shown in Chapter 6, the minimal active duration can be obtained, and statistic characteristics should be able to represent a passed vehicle. Consequently, how to reduce the cycle time of our designated algorithms is the most part to decide whether the software is workable.

3.2.3 Software Framework

The software framework is shown in Figure 3.6. First, the parameter, the number of lane, is given, because the number of lane is always fixed while the detector was installed, and it is easy for operators to decide by the eyes. After the number of lane is determined, the learning processes are launched.

The first learning process is to form a background adapter which will be detailed in Chapter 4. The main objectives of background adapter are to find out the background information and further detect the background variation. The background variation may be arose from a parked vehicle or an accident, hence, any environmental variation should be detected and further filtered out for enhancing the accuracy

of our interesting signals.

The second learning process is to train a lane boundary estimator, the input is reflected FMCW signals and the objective is to determine that reflected signals belong to which lane. The lane boundaries are the necessary conditions for enhancing the accuracy of traffic count and vehicle classification. Consequently, the lane boundary estimator is the core part of road-side radar systems.

In addition, the spread signal should be detected and be further filtered out, because the spread signal influences the proposed vehicle classification algorithm. The spread signal can be treated as some kind of vehicle shadow, and which is proportional to the RCS of detected targets. So far, there is no vehicle classification algorithm designated for the road-side radar detector, because there are too many parameters to be considered and determined in advanced, for example, background information, lane boundaries, and sampling characteristics. In this study, the spread of reflected signals will also be filtered out by integrating certain background information for enhancing the accuracy of vehicle classification, moreover, the boundaries of different vehicle are also be estimated.

Opposite to the current marketing products, the vehicle type will be determined first in our study, the vehicle speed is derived subsequently. Theoretically, the vehicle speed can be estimated directly through the vehicle speed formula of single loop detector if the loop length is given. The estimation of loop length in each lane is out of our research scope, hence, the speed estimation is only proposed theoretically. The functional parts of software are roughly described as above, and the total execution time is asked to be the less the better, because the less execution time means the more sampling of vehicle can be obtained.

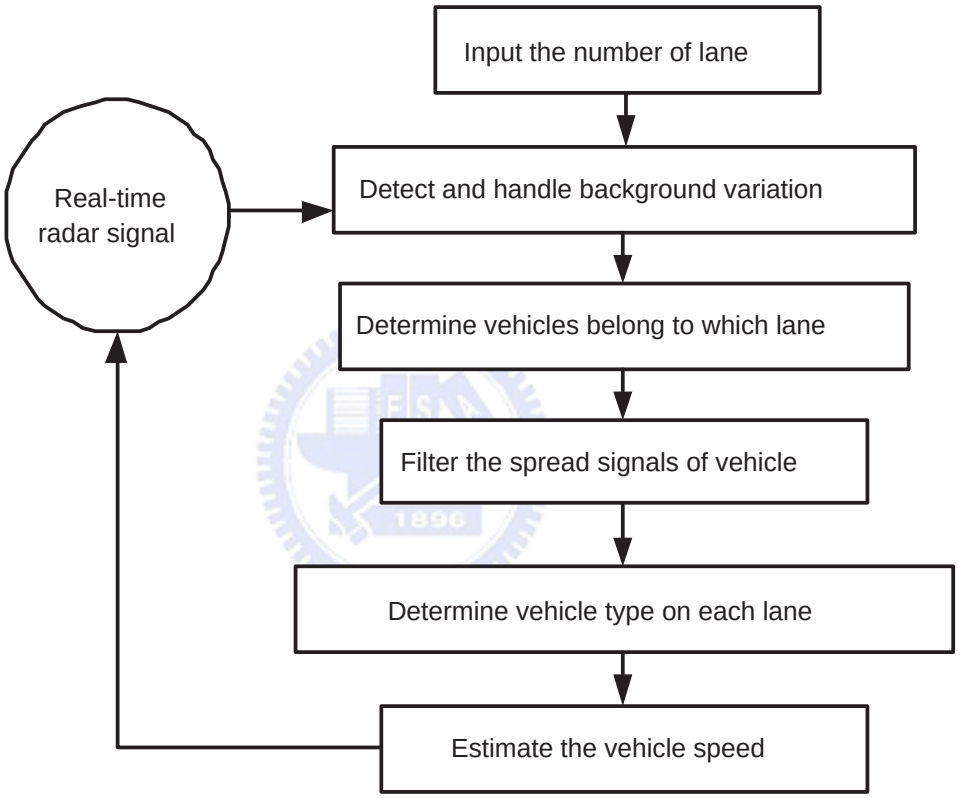


Figure 3.6: The software framework

Chapter 4

Background Adapter

The aim of background adapter is to overcome possible changes resulted from circumstances. For example, one of cases is that an intruding car parks in the detection zone after background information retrieved by radar systems. Usually background information is the scenario that the detection zone does not contain any moving objects, in other words, vehicles may park within the detection zone, but not in the moving state. The goal of designing a background adapter is to allow the radar systems to adapt different background patterns for capturing the better signals of passed vehicles, and then further enhances the accuracy of road-side radar systems.

Background subtraction is a common technique to remove non-moving objects from a video sequence for obtaining background information. The traditional approach is to select several video sequences containing no moving objects and then take a weighted average of the current background and the current frame of the video sequence [17][18][19][20].

A false alarm is an erroneous radar target detection decision caused by noise exceeding the detection threshold. In general, it is an indication of the presence of a radar target when there is no valid target. False alarms are generated while

thermal noise exceeds a pre-set detection threshold, by the presence of spurious signals (either internal to the radar receiver or from sources external to the radar), or by the malfunction of equipment. If the detection threshold is set too high, there will be very few false alarms, on the contrary, if the threshold is set too low, the large number of false alarms will restrain detection of valid targets. Constant false alarm rate (CFAR) detection refers to a common form of adaptive algorithm that vary the detection threshold as a function of the sensed environment [21][22].

After extracting background information, intruding vehicles which suddenly parked within the detection zone are still valid targets (neither noise or spurious information), consequently, the idea of FAR or CFAR is not a suitable definition for describing the scenario that vehicle parked after background information retrieved by radar systems.

There are few literature reviews of background extraction applied for radars, but the idea of background extraction can be also borrowed from video field for overcoming the mentioned scenario. The next subsection will introduce the operational flow chart of background adapter and explain in detail.

4.1 Operational Flow Chart of Background Adapter

The input of background adapter is real-time voltage data, and could be converted to frequency-domain information by Fast Fourier Transform. Based on the capability of equipment, the averaged sample rate is $200k \sim 250k$. According to the designed range resolution 0.78 meters, 512 sampling points are required to form a frame of frequency-domain information. A frame of frequency-domain information at time t is denoted as $\overrightarrow{F}_t$. Consequently, there are $400 \sim 500$ frames generated per second. However, in actually, the number of frequency-domain frames could be less than that,

because algorithms, such as Fast Fourier Transform, background information update, lane boundary estimation, vehicle type classification and speed estimation, occupy certain "short" time to be executed. Consequently, the less computational burden, the more number of frequency-domain frames can be analyzed.

After obtaining a frequency-domain frame $\overrightarrow{F}_t$, a common background subtraction technique applied for video-based equipments is employed to extract background information. Background subtraction information $\overrightarrow{R}_t$ can be derived based on (4.1.1), and the magnitude of $\overrightarrow{R}_t$ is expressed as $\left\|\overrightarrow{R}_t\right\|$.

$$\overrightarrow{R}_t = \overrightarrow{F}_t - \overrightarrow{F}_{t-1}. \tag{4.1.1}$$

Moreover, such background subtraction information is updated all the time for detecting any moving objects in the detection area. A pre-set threshold level δ is a small value, and is applied to look for any feasible background information. If $\left\|\overrightarrow{R}_t\right\|$ is less than threshold level δ, it implies there is no moving objects within the detection area, so information will be kept and further be analyzed to detect whether background information changes. If $\left\|\overrightarrow{R}_t\right\|$ is larger than threshold level δ, a variation, that maybe caused by moving objects, occurs certainly in the detection area. δ can be regarded as a sensitivity measure to detect background changes.

The array size for storing background information is N * P, where P is the number of bin position of frequency-domain domain information, and N is the number of frame. Usually, the number of N is set to 500, and the significant level α is set to 0.01. The null hypothesis is that the sample mean of background information is equal to the previously decided background information on the corresponding bin position.

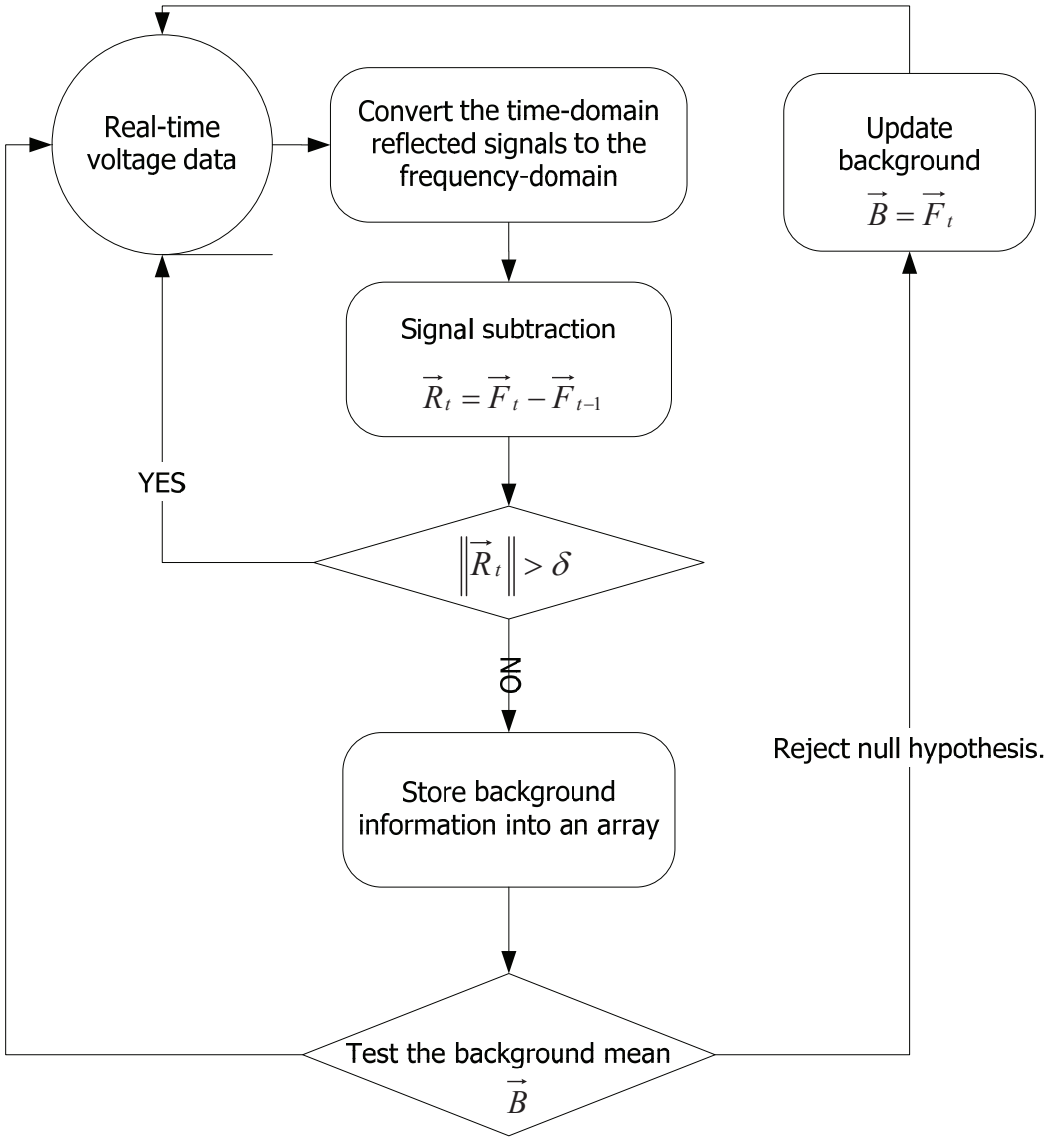


Figure 4.1: Flow chart of a background adapter

If the null hypothesis is rejected, background information $\overrightarrow{B}$ is updated by the latest background information $\overrightarrow{F}_t$, moreover, if the changed background information is within the certain lane position, the vehicle classifier on that corresponding lane will be asked to train again. If the null hypothesis is not rejected, the array for storing background information keeps to update if $\left\|\overrightarrow{R}_t\right\|$ is less than threshold level δ.

Chapter 5

Lane Boundary Estimator

The traffic surveillance function, the most basic and fundamental of all ITS functions, is responsible for gathering continuous, real-time information on the state of the transportation system and driving environment. The information gathered forms the basis for the different traffic management and control actions that can be taken to optimize performance and also serves as the basis for the information to be disseminated to travellers [23].

Traffic surveillance has relied upon measuring devices, which have traditionally been embedded into the road, for both measuring traffic conditions and providing control to signaling mechanisms that regulate traffic flow. However, many of sensors have been "intrusive" types. Several "no-intrusive" sensor technologies have been developed for traffic monitoring, however, none of these devices can be reconfigured or adapted without the assistance of certified technicians. Such an on-site modification to the sensors may take several hours to several days for a single reconfiguration.

This section develops a learning algorithm of road-side radar sensors to monitor traffic conditions when vehicles pass through a field of view and process the data into an estimation of the position of each of the detected vehicles. The positions

are defined and recorded for use in a probability density function estimation without requiring manual set-up and definition of the traffic lane boundaries. The update process is done by continuously updating the parameters of the GMM [11].

5.1 Extraction of Span and Conflict Information

In supervised learning, category labels or training patterns are provided to estimate the unknown model parameters to reduce errors or costs. In unsupervised learning or clustering, no category labels or human interventions exist to guide error measure reduction [24]. Besides, if the instances in the learning problems are characterized by the presence of a small percentage of labelled data or side-information only, such learning algorithms are the ones dubbed as semi-supervised learning. The side-information that could be useful in clustering problems is primarily obtained from background knowledge about the domain or instances [25][26].

Constraint information is one of side-information to form constraints to assist in clustering. Basically, the constraints primarily comprise two types, equivalence constraints (positive constraints, must-link constraints) and not-equivalent constraints (negative constraints, cannot-link constraints). Equivalence constraints specify two un-labelled instances must belong to the same cluster, while not-equivalent constraints refer to the situation where the two instances come from different sources [27][28][29][30].

5.1.1 Span Information

To overcome the defects in the patents mentioned above, this study proposes a new feature, span information, for building a learning feature rather than adopting only a single sample point to describe lane positions when a vehicle passed through. The span information denotes an idea of a signal range of a single lane width formed by the reflected signal from a passing car.

Conceptually, the idea of span information is analogous to instance-level equivalence constraint, because the selected bins come from the same source (same vehicle). In other words, The span information modifies the shape based on background knowledge to form learning samples. The span information is carried out by applying the concept of a convex hull to establish a closed set, and by grouping the selected bins that belong to a lane into a closed interval. The closed set comprises the bin positions where peaks had occurred, and should pass through the following two steps to form reasonable information. Notably, vehicles that are applied to form the span information don't include the motorcycles, because it is meaningless for merging two motorcycles detected on adjacent lanes. Therefore, the motorcycles will be filtered out through the vehicle detection algorithm.

1. "Maximum separation" indicates the maximal distance between the adjacent bin positions with peaks of a passed vehicle. The common distance between the adjacent bin positions is same to the illustrated examples shown in Figure 2.1 that the bin range is approximately 78 cm thick, hence, the reasonable inference that there seldom exists an empty subinterval (for example, a subinterval with a width exceeding 1 meter) in which no peak occurs as a vehicle passes by. Consequently, if the distance between any adjacent bin positions with peaks

is less than 1 meter, the two bins should be grouped into a closed interval. The iterative merging procedure will not stop until the distance between any adjacent bin with peaks is larger than 1 meter. On the contrary, if the distance between any adjacent bin with peaks is larger than 1 meter, the reflected signals should be regarded as two different vehicles.

2. "Maximum extent" describes the maximum reach of the span information. According to the regular lane width design, the lane width is about $4 \sim 5$ meter. This design aims to catch reflected signals from individual vehicles, while avoiding those from pairs of vehicles in adjacent lanes that are detected and gathered simultaneously. Because such information is inappropriate for addressing a single lane's positions, hence, the maximum extent is employed to filter out the vehicles in adjacent lanes that are detected simultaneously.

Restated, the span information is formed by using the above two steps, and which accords with the idea of instance-level equivalent constraint that group the select bins belonging to a vehicle into a closed interval. Take Figure 2.1 for example, the span information in (a) ranges from 21 to 23, while that in (b) ranges from 23 to 25. The primary advantage of span information is that the ambiguous results derived from the mentioned patents can be avoided.

5.1.2 Conflict Information

Besides, for developing the more accurate on-line traffic lane estimator, another feature, conflict information, is extracted to work together with span information. Conflict information is an idea to represent the lane boundary, and can be directly derived

from the property of span information without introducing other external information. Notably, introducing the external intervention is unrealistic and inappropriate when dealing with a general road environment, because the attention devoted to installation and calibration leads to inconvenience and operational costs.

From the collected data, span information obtained from vehicles in different lanes do not overlap except for boundary positions, therefore, the overlap implies a higher possibility of lane boundaries. For obtaining overlapping positions, this study assumes if the left boundary of span information is same to the right boundary of another, a conflict pair is formed at that position. Conceptually, conflict information is an idea analogous to an instance-level not-equivalent constraint, because different span information arose from vehicles in different lanes can be considered as different sources.

To identify the possible conflict positions on the arrival of span information, an intuitive and heuristic matching mechanism is designed to count the conflict pairs at every position and expressed as follows:

$$C_i = \min\{L[i], R[i]\}, \tag{5.1.1}$$

where the conflict count at position i, denoted as C_i, and $L[i]$ and $R[i]$, respectively, denoted the accumulated left and right boundary counts of span information on position i. Taking Figure 2.1 as an example, the span information of a small vehicle in lane 1 is 21-23, thus, the left boundary is 21 and the right boundary is 23 ($L[21] = 1$ and $R[23] = 1$). Additionally, the span information of a small vehicle in lane 2 is 23-25, hence, the left boundary is 23 and the right boundary is 25 ($L[23] = 1$ and $R[25] = 1$). Suppose there are only two vehicles are taken into consideration, then

only a conflict pair is derived at position 23, and the conflict count is one; that is $C_{23} = 1$.

Furthermore, if the histogram of accumulated span information can take the conflict information provided by Eqn. (5.1.1) into account, the resultant accumulated histogram more closely reflects vehicle distribution. Hence, the accumulated histogram of span information is further adjusted by incorporating the conflict information according to the following Eqn. (5.1.2)

$$A_i^{'} = \min\{A_i - C_i\}, \tag{5.1.2}$$

where the accumulated counts on position i denoted as A_i and the refined accumulated counts on position i denoted as $A_i^{'}$. In sum, conflict information is an outcome generated by the different span information without knowing that span information belongs to which lane, and is generated by the on-line panel data without the guidelines provided by external experts or teachers, hence, the whole learning process can be executed automatically.

Additionally, even if vehicles rarely pass through one of lanes, the subtraction of the accumulated span information by using conflict information still should be implemented. Since conflict information usually appears at the boundary of span information, and just occupies a small part of span information. Hence, the subtraction of conflict counts will not make that the lane with rare vehicles becomes less apparent. On the contrary, the adjustment for the accumulated span information will sharpen the distribution so that lane positions can be described more precisely.

5.2 Incorporating a Variant of EM Algorithm with a GMM

5.2.1 Gaussian Mixture Model

Gaussian Mixture Model is a very popular technique for model selection [31][32] and clustering [33], and its purpose is to express the probability density function of the chosen features as a weighted sum of several Gaussian distributions. Owing to the flexibility of GMM, it also has been successfully applied to other applications in relation to image processes [34][35][36], speech recognition [37][38], and other fields [39][40][41]. Besides, the Gaussian component also can be replaced by another probability density function to apply to different applications [42].

Generally, large variance results from real-world imperfect or noisy data, and leads to the GMM being ineffective in the clustering problem. Lee and Cho [43] attempted to overcome the weakness of large variance by applying the committee approach, which combines several single models to reduce variance. Banfield and Raftery [44] suggested to use a uniform noisy component to handle noise, and it would also reduce the variance of Gaussian components. The component with small variance is extremely important, particularly in relation to boundary determination, because the large variance may easily result in misleading and incorrect lane position outcomes. As a consequence, the variant of EM algorithm [45] is proposed to reduce the variance of Gaussian components and achieve classification, and is detailed in the following subsection.

5.2.2 A Variant of EM Algorithm

This study develops a variant of EM algorithm to achieve classification and reduce the variance of each Gaussian component; that is, data point x_i is assigned to the component which has the highest likelihood at the start of each iteration, and each component updates its sample mean and variance by only calculating the instances assigned to it rather than using all instances. The proposed variant of EM algorithm will not stop until such time as difference between consecutive iterations is below some specific tolerance level and none of the instances changed its component. In other words, the variant of EM algorithm can be regarded as a classification version of the EM algorithm. The proposed variant of EM algorithm incorporating the above GMM could form an on-line lane estimator. More formally, a GMM $p(x|\theta)$ with M components can be described as:

$$p(x|\theta) = \sum_{m=1}^{M} \alpha_m g(x; \mu_m, \sigma_m^2), \tag{5.2.1}$$

where $\alpha_1, ..., \alpha_M$ denote the non-negative weights and add up to one, while μ_m and σ_m^2, respectively, denote the estimate of the mean and variance of the m^{th} Gaussian distribution $g(x; \mu_m, \sigma_m^2)$ (or $g_m(x)$). Data point x_i is assumed i.i.d. where $i \in \{1,\ 2, \ldots,\ n\}$ and represents the relative distance.

All of GMM parameters $\theta = [\alpha_1, \alpha_2, \ldots, \alpha_M, \mu_1, \mu_2, \ldots, \mu_M, \sigma_1^2, \sigma_2^2, \ldots, \sigma_M^2]$ can be estimated and optimized by using the variant of EM algorithm, and the expectation of log-likelihood is detailed as follows:

$$E - step : \tag{5.2.2}$$

$$\begin{aligned} & J(\theta|\theta^{(k)}) \\ & = E\left[\ln\left[\prod_{i=1}^{n} p(x_i)\right]\middle|\,\theta^{(k)}\right] \\ & = E\left[\sum_{i=1}^{n} \ln p(x_i)\middle|\,\theta^{(k)}\right] \end{aligned} \tag{5.2.3}$$

where $\theta^{(k)}$ is the estimate at iteration k.

Additionally, let $\left|X_j^{(k)}\right|$ denote the set size for data point x_i being assigned to component j which has the highest likelihood at iteration k, and can be expressed as follows:

$$X_j^{(k)} : \{i : \ x_i \ is \ assigned \ to \ the \ component \ j \ at \ iteration \ k\},$$

where $\left|X_j^{(k)}\right| \neq 0$, and $\bigcup_{j=1}^{M} X_j^{(k)} = \{1, 2, 3, ..., n\}$.

For obtaining the optimal θ, we follow the method of Maximum Likelihood. First, Gaussian density function is differentiated with respect to μ and σ.

$$\begin{aligned} & \nabla_\mu g(x; \mu, \sigma^2) \\ & = g(x; \mu, \sigma^2)\left(-\tfrac{1}{2\sigma^2}\right)\nabla\left[(x-\mu)^T(x-\mu)\right] \\ & = g(x; \mu, \sigma^2)\left(\tfrac{(x-\mu)}{\sigma^2}\right) \end{aligned}$$

$$\begin{aligned} & \nabla_\sigma g(x; \mu, \sigma^2) \\ & = (2\pi)^{-d/2}\left\{\begin{array}{l} (-d)\sigma^{-d-1}\exp\left[-\frac{(x-\mu)^T(x-\mu)}{2\sigma^2}\right] + \\ \sigma^{-d}\left[\frac{(x-\mu)^T(x-\mu)}{\sigma^3}\right]\exp\left[-\frac{(x-\mu)^T(x-\mu)}{2\sigma^2}\right] \end{array}\right\} \\ & = g(x; \mu, \sigma^2)\left(\tfrac{(x-\mu)^T(x-\mu)}{\sigma^3} - \tfrac{d}{\sigma}\right) \end{aligned}$$

For simplifying the expression, the $\beta_j^{(k)}(x_i) = \frac{\alpha_j^{(k)} g_j^{(k)}(x_i)\delta_{ij}^{(k)}}{\sum_{j=1}^{M} \alpha_j^{(k)} g_j^{(k)}(x_i)\delta_{ij}^{(k)}}$ is introduced. The update rules are obtained by differentiating Eqn. (5.2.3) with respect to μ_j, and σ_j, and the update formula for the parameters for each component are addressed as follows.

$$\begin{aligned}
&\nabla_{\mu_j} J(\theta) = \sum_{i=1}^{n} \frac{\alpha_j g(x_i;\mu_j,\sigma_j^2)\delta_{ij}}{\alpha_1 g(x_i;\mu_1,\sigma_1^2)\delta_{i1}+\alpha_2 g(x_i;\mu_2,\sigma_2^2)\delta_{i2}+\ldots+\alpha_M g(x_i;\mu_M,\sigma_M^2)\delta_{iM}} \frac{(x_i-\mu_j)}{\sigma_j^2} \\
&= \sum_{i=1}^{N} \beta_j(x_i) \left(\frac{(x_i-\mu_j)}{\sigma_j^2}\right)
\end{aligned}$$

$$\begin{aligned}
&\nabla_{\sigma_j} J(\theta) \\
&= \sum_{i=1}^{N} \frac{\alpha_j g(x_i;\mu_j,\sigma_j^2)\delta_{ij}}{\alpha_1 g(x_i;\mu_1,\sigma_1^2)\delta_{i1}+\alpha_2 g(x_i;\mu_2,\sigma_2^2)\delta_{i2}+\ldots+\alpha_M g(x_i;\mu_M,\sigma_M^2)\delta_{iM}} \left(\frac{(x-\mu)^T(x-\mu)}{\sigma_j^3} - \frac{d}{\sigma_j}\right) \\
&= \sum_{i=1}^{N} \beta_j(x_i) \left(\frac{(x-\mu)^T(x-\mu)}{\sigma_j^3} - \frac{d}{\sigma_j}\right)
\end{aligned}$$

where $\delta_{ij}^{(k)}$ is a binary indicator. $\delta_{ij}^{(k)} = 1$, if data point x_i is assigned to the component j; otherwise $\delta_{ij}^{(k)} = 0$; that is $\delta_{ij}^{(k)} = 1$.

According to the first order necessary condition for optimality, let $\nabla_{\mu_j} J(\theta) = 0$ and $\nabla_{\sigma_j} J(\theta) = 0$. Consequently, $M - step$ can be addressed as follows:

$$\alpha_j^{(k+1)} = \frac{\sum_{i \in X_j^{(k)}} \beta_j^{(k)}(x_i)}{\sum_{j} \sum_{i \in X_j^{(k)}} \beta_j^{(k)}(x_i)} \tag{5.2.4}$$

$$\mu_j^{(k+1)} = \frac{\sum_{i \in X_j^{(k)}} \beta_j^{(k)}(x_i) x_i}{\sum_{i \in X_j^{(k)}} \beta_j^{(k)}(x_i)} \tag{5.2.5}$$

$$\sigma_j^{2^{(k+1)}} = \frac{\sum_{i \in X_j^{(k)}} \beta_j^{(k)}(x_i)(x_i - \mu_j^{(k)})^2}{\sum_{i \in X_j^{(k)}} \beta_j^{(k)}(x_i)} \tag{5.2.6}$$

As is clearly visible, each component only calculates the data points that are assigned to it, and the computational burden is relatively less then the EM algorithm. The variant of EM algorithm can be expressed as follows:

1. Set an initial values ($k = 0$) of parameters $\theta^{(k)}$. Basically, the initial weights can be set to $\alpha_1^{(0)} = \alpha_2^{(0)} = \ldots = \alpha_M^{(0)} = \frac{1}{M}$, and component center $\mu_j^{(0)}$ can be set by a K-mean initialization; the initial $\sigma_j^{(0)}$ is set to one third of lane width.

2. All the instances x_i are assigned to their nearest component j with highest likelihood.

3. Use $\theta^{(k)}$ to calculate a posterior probability $\beta_j^{(k)}(x_i)$.

4. Calculate a new weight $\alpha_j^{(k+1)}$.

5. Calculate a new mean $\mu_j^{(k+1)}$.

6. Calculate a new variance $\sigma_j^{2(k+1)}$.

7. If $\left\|\theta^{(k+1)} - \theta^{(k)}\right\|$ is less than a tolerant value, and none of the instances changed its component, stop; otherwise, set $k = k + 1$, and go back to 2.

Definition 5.2.1. For component j with mean μ_j at iteration k, define a displacement $d_i^{(k)} = \left|x_i - \mu_j^{(k)}\right|,\ i \in \{1,\ 2, \ldots,\ n\}$ such that $d_1^{(k)} \leq d_2^{(k)} \ldots \leq d_n^{(k)}$ is satisfied.

Definition 5.2.2. Define a set $H_l^{(k)} = \{x_i :\ d_i^{(k)}\ \leq\ d_l^{(k)}$ at iteration $k\}$.

Proposition 5.2.1. *For component j with mean μ_j at iteration k, given nature numbers $l,\ l+1 \in \{1,\ 2, \ldots,\ n\}$, then $\sigma_j^2(H_l^{(k)}) \leq \sigma_j^2(H_{l+1}^{(k)})$ is true.*

Proof. By definition, $\sigma_j^2(H_l^{(k)}) = \frac{\sum_{i=1}^{l} \beta_j^{(k)}(x_i) d_i^{(k)2}}{\sum_{i=1}^{l} \beta_j^{(k)}(x_i)}$, hence, we can find a $\overline{d}^{(k)}$, which satisfies $d_1^{(k)2} \leq \overline{d}^{(k)2} = \sigma_j^2(H_l^{(k)}) \leq d_l^{(k)2}$ so that $\sigma_j^2(H_l^{(k)}) = \frac{\sum_{i=1}^{l} \beta_j^{(k)}(x_i) d_i^{(k)2} + \beta_j^{(k)}(x_{l+1}) \overline{d}^{(k)2}}{\sum_{i=1}^{l+1} \beta_j^{(k)}(x_i)}$. Because $\overline{d}^{(k)} \leq d_l^{(k)} \leq d_{l+1}^{(k)}$, hence,

$$\sigma_j^2(H_l^{(k)}) = \frac{\sum_{i=1}^{l} \beta_j^{(k)}(x_i) d_i^{(k)2} + \beta_j^{(k)}(x_{l+1}) \overline{d}^{(k)2}}{\sum_{i=1}^{l+1} \beta_j^{(k)}(x_i)} \leq \frac{\sum_{i=1}^{l} \beta_j^{(k)}(x_i) d_i^{(k)2} + \beta_j^{(k)}(x_{l+1}) d_{l+1}^{(k)2}}{\sum_{i=1}^{l+1} \beta_j^{(k)}(x_i)} = \sigma_j^2(H_{l+1}^{(k)}) \qquad \square$$

Proposition 5.2.2. *For component j at iteration k, suppose the component mean estimated by the EM algorithm and that estimated by the variant of EM algorithm are the same, then the variance estimated by the variant of EM algorithm is not larger then the variance estimated by the EM algorithm.*

Proof. Let $x_1,\ x_2, \ldots,\ x_n$ be with displacements $d_1^{(k)},\ d_2^{(k)}, \ldots,\ d_n^{(k)}$, respectively and $d_1^{(k)} \leq d_2^{(k)} \leq \ldots \leq d_n^{(k)}$. Given nature numbers l and $l+1$ satisfy $\alpha_j^{(k)} g_j^{(k)}(x_l) > \alpha_{j'}^{(k)} g_{j'}^{(k)}(x_l)$ and $\alpha_j^{(k)} g_j^{(k)}(x_{l+1}) < \alpha_{j'}^{(k)} g_{j'}^{(k)}(x_{l+1})$, where $j, j' \in M$ and $j \neq j'$ at iteration k. Then based on the classification rule of the variant of EM algorithm, instances with displacements not larger than $d_l^{(k)}$ can be gathered into a set $H_l^{(k)}$, and are assigned to the component j to update the parameters. Hence based on Proposition 5.2.1, $\sigma_j^2(H_l^{(k)})$ is not larger than $\sigma_j^2(H_p^{(k)})$ where $p \in \{l+1,\ l+2, \ldots,\ n\}$. Furthermore, it is known that $\sigma_j^2(H_l^{(k)})$ is the estimated variance obtained by the variant of EM algorithm, while $\sigma_j^2(H_n^{(k)})$ is the estimated variance obtained by the EM algorithm; that is, the variance estimated by the variant of EM algorithm is not larger than that estimated by the EM algorithm. $\square$

Hence, according the Proposition 5.2.2, the lower variance by using the variant of EM algorithm can also be expected provided that the component mean estimated by

the EM algorithm and that estimated by variant of EM algorithm are the same.

5.2.3 Learning Procedure

Figure 5.1 shows the entire flow chart of the learning procedure of the on-line traffic lane estimator, and Figure 5.2 shows the corresponding illustrative example.

Initially, the reflected real-time signals regarded as an input were converted to the frequency-domain information via a Fast Fourier transform algorithm. When a passed vehicle is captured with a 3D pattern like that shown in Figure 1.2(a), the peak count of a vehicle shown in Figure 5.2(a) can be obtained. There are four illustrative vehicles shown in Figure 5.2, and the information of peak count shown in Figure 5.2(a) can be converted into span information shown in Figure 5.2(b). When enough vehicles are gathered (about 30 vehicles), the histogram of accumulated span information can be marshaled as Figure 5.2(c). Afterward, the conflict counts on each bin can be calculated using Eqn. (5.1.1) based on span information. In this case, only one conflict occurs at the bin position 24 shown in Figure 5.2(d). The histogram of the accumulated span information shown in Figure 5.2(e) can be refined by using Eqn. (5.1.2). Subsequently, a GMM is introduced to model the refined histogram and the parameters can be estimated by using the variant of EM algorithm. Finally, the outcome similar to several Gaussian functions shown in Figure 5.2(f) can be drawn geometrically. Notably, this learning procedure only requires the input of lane number. The other information (such as span and conflict information) can be automatically derived from the passed vehicles without any external intervention.

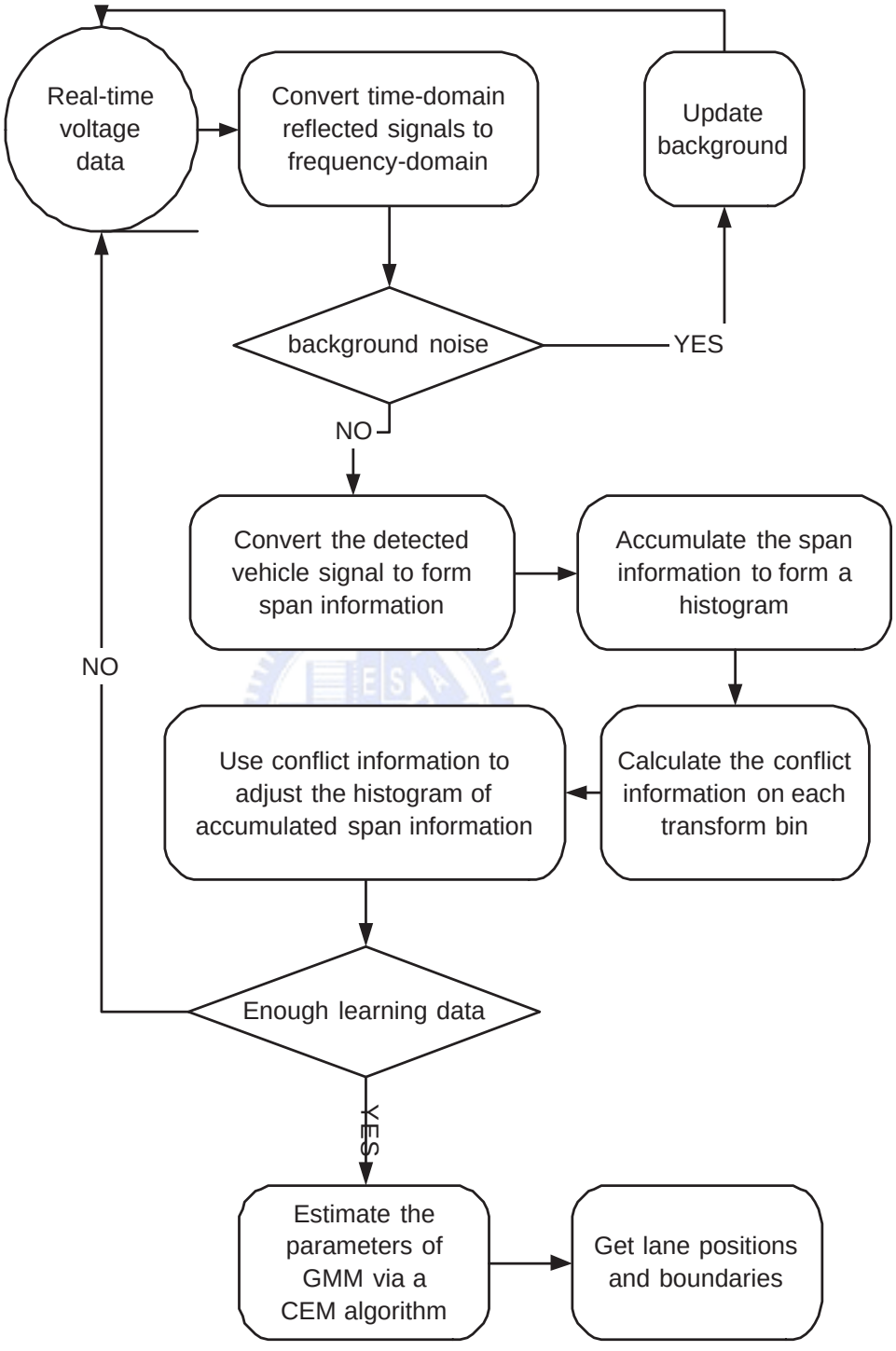


Figure 5.1: Learning procedure of the on-line traffic lane estimator

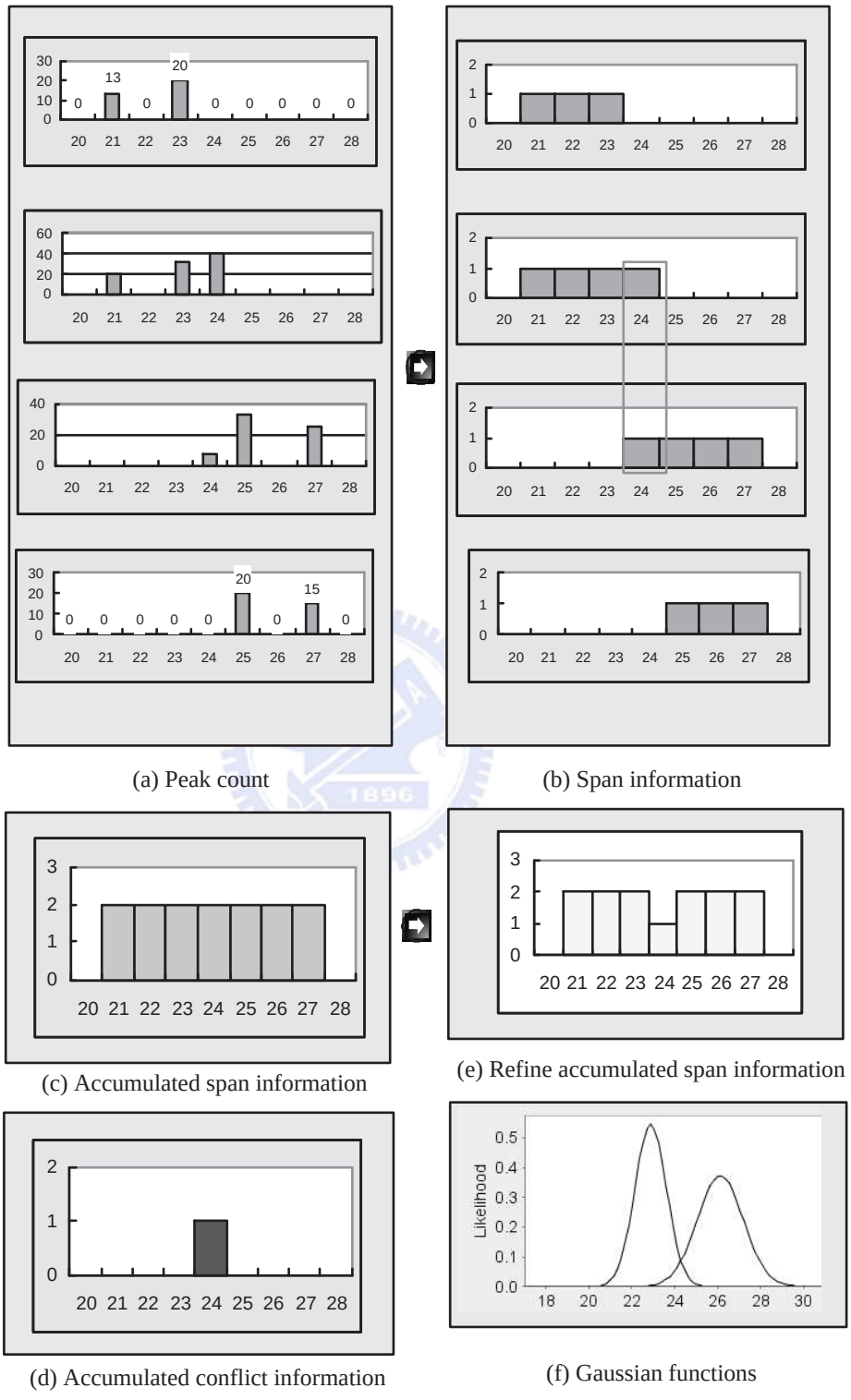


Figure 5.2: Illustrative example of the learning procedure

Chapter 6

Vehicle Classifier

6.1 Automatic Learning Framework

The panel data shown in Figure 1.4 represent a large vehicle in lane 2 and a small vehicle in lane 6, and depict the energy distributions and magnitudes of vehicles. Figure 1.4 shows that the spread-spectrum of a large vehicle is so wide that the neighbor lanes are influenced, and reveals that the large vehicle tends to have the higher reflected energy and the wider spread-spectrum than a small one.

Conceptually the energy magnitude and spread-spectrum could be selected as features for classifying vehicles. However, in multi-lane environments, the spread-spectrum of a large vehicle may simultaneously cause a virtual pattern which is similar to a small vehicle or motorcycle in the adjacent lanes, such that we may confuse whether a vehicle or motorcycle exists in the adjacent lanes.

In this work, an automatic learning framework shown in Figure 6.1 is proposed for developing a real-time vehicle classifier, and the flow chart can be described as follows. First, the received voltage data will be converted into frequency-domain information shown in Figure 1.2 via Fast Fourier Transform. Then an algorithm is

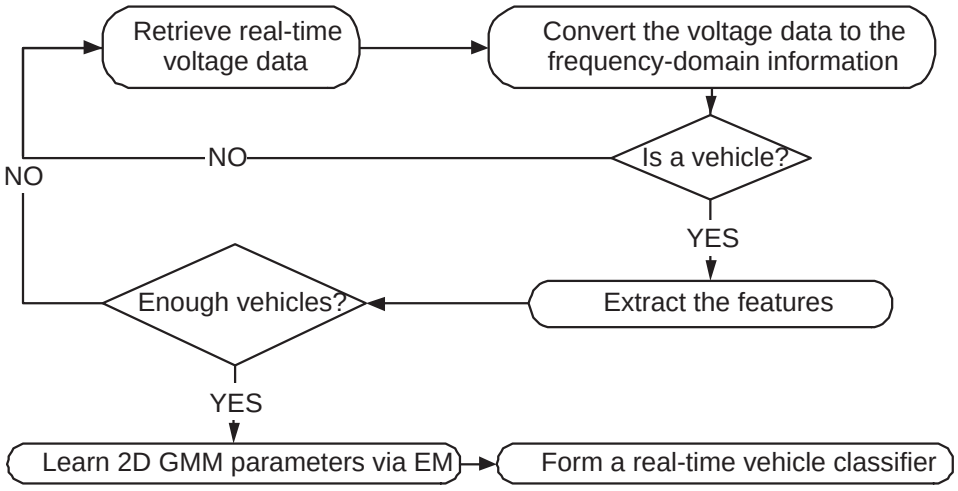


Figure 6.1: Automatic learning framework of real-time vehicle classifier. The first step is to retrieve the real-time voltage data, and the voltage date are large enough to generate 200 frequency-domain frames per second uniformly.

applied to determine if the transformed information forms vehicles. If reflected signals of a vehicle is obtained, the features of a vehicle will be extracted from a series of frequency-domain frames. While sufficient training vehicles are gathered (about 30 vehicles), a two-dimensional Gaussian mixed model is introduced to describe these extracted features, and combines an expectation maximization algorithm (1977) [45] to estimate the GMM parameters.

6.1.1 Vehicle Detection Algorithm

The aim of the vehicle detection algorithm is to determine whether a vehicle exists from the panel data shown in Figure 1.4. For achieving the purpose, a multi-threshold detection algorithm is proposed for detect vehicle and described as follows.

Step 1 start point detection. Assume there are around 200 frequency-domain frames per second based on the capability of the analog digital converter, and the arrival

rate is uniform. The variance and magnitude thresholds (Th_1 and Th_2) are set to obtain the start point of a potential vehicle. If the number of successive frames that satisfy the thresholds (i.e., above Th_1 and Th_2) exceeds three times on each lane's frequency range, the start point of a potential vehicle is obtained. Otherwise, reflected signals are regards as the background information. Note that Th_1 and Th_2 are relatively low so that a potential vehicle can be detected as possible.

Step 2 end point detection. The end point can be obtained if the number of successive frames that satisfy the thresholds (i.e., below Th_1 and Th_2) exceeds Th_3 times. Th_3 is the least number of frames required to separate vehicles one by one, and Th_3 can obtained by the following equation.

$$Th_3 = \frac{frames \times (headway - length\ of\ detection\ zone)}{the\ lowest\ speed\ of\ vehicle}, \tag{6.1.1}$$

where the minimal headway has to be larger than the virtual loop, or there is no way to separate vehicles one by one.

Step 3 minimal active duration detection. According to the start and end points, the active duration can be calculated, and the minimal active duration threshold Th_4 can be calculated based on the following equation.

$$Th_4 = \frac{frames \times (vehicle\ length + length\ of\ detection\ zone)}{the\ highest\ speed\ of\ vehicle}. \tag{6.1.2}$$

If the active duration is less than Th_4, the data will be regarded as a virtual pattern or noise from the adjacent lanes.

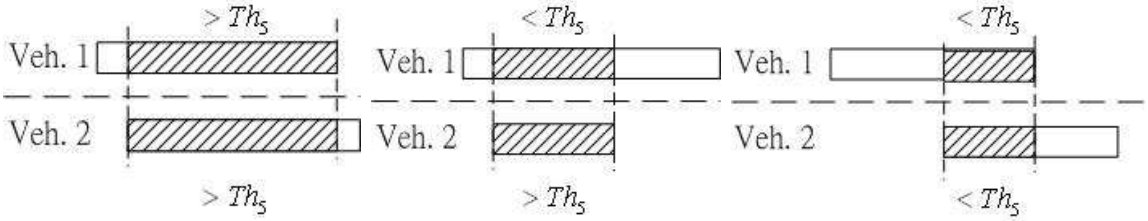


Figure 6.2: Virtual vehicle detection. (a) is a scenario that a virtual vehicle exists, and (b)(c) show there may be two vehicles in the adjacent lanes.

Step 4 virtual vehicle detection. Because the spread-spectrum of a large vehicle may simultaneously cause a virtual pattern which is similar to a small vehicle or motorcycle in the adjacent lanes, such that the appearing duration of the "real" and "virtual" vehicles have high overlap from time series perspective. Therefore, for determining whether virtual vehicles exist, an overlapping ratio threshold Th_5 is proposed to detect whether a virtual vehicle exists. An overlapping ratio in each lane is calculated by using overlapping duration to divide the active duration.

If any overlapping ratios in the neighboring lanes are above the overlapping ratio threshold (i.e., above Th_5), a virtual vehicle is detected. Figure 6.2(a) illustrates the idea of virtual vehicle that can be detected; one of vehicles could be a virtual vehicle. Figure 6.2(b)(c) shows there are two potential vehicles in the adjacent lanes.

Step 5 virtual vehicle obviation. After a virtual vehicle is detected, the energy magnitude within the overlapping duration will be utilized to compare. Since reflected signals of a real vehicle are always larger than these of virtual ones.

Hence, if a lane always has the higher odds in energy magnitude comparison, the reflected signals in that lane will be regard as a real vehicle, whereas others are virtual vehicles.

6.1.2 Vehicle Feature Extraction

The main difficulties regarding the vehicle feature extraction include: first, for providing the accurate features, the vehicle detection algorithm and accurate lane positions are required in advance. However, the spread of reflected signals of a large vehicle usually influences neighbor lanes such that the accuracy of vehicle detection degraded. Second, a good feature for describing moving vehicles is difficult to obtain, because most of features are affected by different vehicle speeds. One of our contributions is that the two proposed features can be obtained in real-time and are less sensitive to speed and the signals of vehicles in neighboring lanes. The selected magnitude and variance in the corresponding range of lane highly depend on the vehicle sizes when vehicles passed through the detection zone, consequently, a pair of features, "maximal energy peak" and "maximal energy variance", are proposed to form a learning sample $x \in R^{2+}$.

After obtaining the reflected signals of a vehicle, The maximal energy peak can be extracted as follows. First, the maximal energy value within the corresponding lane can be found from a frequency-domain frame shown in Figure 1.2(b), then a series of maximal energies within the active duration can form a frequency-domain characteristic, and the peak of this series is chosen to form the maximal energy peak.

Another selected feature, maximal energy variance is processed in the same way, except the maximal energy value is replaced by the energy variance that is the variance

of the energies within each lane.

6.1.3 Two-Dimensional Gaussian Mixed Model

In this study, large and small vehicles represent the two groups to be detected, because the effective reflected length of a motorcycle is too short and the magnitude of motorcycle is too small to distinguish. Consequently, motorcycles will be filtered out. A GMM with two components can be described as:

$$p(x) = \alpha_1 g(x; \mu_1, \Sigma_1) + \alpha_2 g(x; \mu_2, \Sigma_2), \tag{6.1.3}$$

where $g(x; \mu, \Sigma)$ is a Gaussian density function shown as below:

$$g(x; \mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^2 |\Sigma|}} \exp\left[-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right], \tag{6.1.4}$$

and α_1, α_2 denote the non-negative weights and add up to one. μ and Σ denote the mean and covariance matrix of Gaussian distribution respectively. All of these parameters can be estimated by using the EM algorithm, and the entire iterative procedure will stop until the difference between several successive iterations is below some specific tolerance level. The expectation of likelihood with two vehicles is detailed as follows:

$E(xpectation) - Step$:

$$\begin{aligned} J(\theta|\theta^{(k)}) &= E\left[\ln\left[\prod_{i=1}^{n} p(x_i)\right]\middle| x, \theta^{(k)}\right] \\ &= E\left[\sum_{i=1}^{n} \ln p(x_i)\middle| x, \theta^{(k)}\right] \\ &= E\left[\sum_{i=1}^{n} \ln\left[\alpha_1 g(x_i;\mu_1,\Sigma_1) + \alpha_2 g(x_i;\mu_2,\Sigma_2)\right]\middle| x, \theta^{(k)}\right], \end{aligned}$$

where $\theta^{(k)}$ is the estimate at step k. For simplifying the discussion, the posterior probability $\beta_j(x)$ of vehicle j is expressed as follows:

$$\beta_j(x) = \frac{\alpha_j g(x;\mu_j,\Sigma_j)}{\alpha_1 g(x;\mu_1,\Sigma_1) + \alpha_2 g(x;\mu_2,\Sigma_2)}. \tag{6.1.5}$$

For obtaining the maximum likelihood of $J(\theta)$, the update formula for μ_j, σ_j^2 of each vehicle j can be derived from the first order necessary condition and addressed as follows.

M(aximization) - Step:

$$\mu_j = \frac{\sum_{i=1}^{n} \beta_j(x_i) x_i}{\sum_{i=1}^{n} \beta_j(x_i)} \tag{6.1.6}$$

and

$$\sigma_j^2 = \frac{\sum_{i=1}^{n} \beta_j(x_i)(x_i-\mu_j)(x_i-\mu_j)^T}{\sum_{i=1}^{n} \beta_j(x_i)}. \tag{6.1.7}$$

Moreover, in order to meet the equality constraint $\sum_{j=1}^{2} \alpha_j = 1$, Lagrange multipliers

are introduced to redefine the new objective function J_{new}, and α_j also can be derived as follows.

$$\begin{aligned} J_{new} &= J + \lambda(1 - \alpha_1 - \alpha_2) \\ &= \sum_{i=1}^{n} \ln\left[\alpha_1 g(x_i; \mu_1, \Sigma_1) + \alpha_2 g(x_i; \mu_2, \Sigma_2)\right] + \lambda(1 - \alpha_1 - \alpha_2) \end{aligned}$$

and

$$\begin{aligned} \frac{\partial J_{new}}{\partial \alpha_j} &= \sum_{i=1}^{n} \frac{g(x_i; \mu_j, \Sigma_j)}{\alpha_1 g(x_i; \mu_1, \Sigma_1) + \alpha_2 g(x_i; \mu_2, \Sigma_2)} - \lambda = 0 \\ &= \frac{1}{\alpha_j} \sum_{i=1}^{n} \beta_j(x_i) - \lambda = 0, j = 1, 2 \\ &\Rightarrow \quad \alpha_1\lambda + \alpha_2\lambda = \sum_{i=1}^{n} \left[\beta_1(x_i) + \beta_2(x_i)\right] \\ &\Rightarrow \quad (\alpha_1 + \alpha_2)\lambda = \lambda = \sum_{i=1}^{n} 1 = n \\ &\Rightarrow \quad \alpha_j = \frac{1}{n} \sum_{i=1}^{n} \beta_j(x_i), j = 1,\ 2. \end{aligned}$$

Based on the above update formula, the $\theta = (\alpha_1, \mu_1, \Sigma_1, \alpha_2, \mu_2, \Sigma_2)$ can be updated iteratively, and such that the likelihood of GMM approaches the maximum.

6.2 Estimation of Vehicle Speed

Most radars use Doppler effect to obtain the relative speed of the target, however, the direction of illumination of road-side radar detectors is perpendicular to the direction of traffic, thus Doppler effect is unapparent and hard to be utilized in our study.

In this study, the principle of single loop detector for estimating vehicle speed is employed. The formula is shown in equation 6.2.1. The length of vehicle can be derived directly from the result of vehicle type shown in the previous section, but the length of the same vehicle is assumed to be equal. In addition, the length of detection zone is also assumed to the 3dB beam width of the antenna in the E-plane. The denominator is the time difference between start point and end point shown in Step 1 and 2 of the vehicle detection algorithm. The estimated value of vehicle speed is very rough, because all variables shown in equation (6.2.1) are of much uncertainty. Consequently, the estimation of vehicle speed of equation (6.2.1) is just a theoretical value.

$$vehicle\ speed = \frac{(vehicle\ length + length\ of\ detection\ zone)}{the\ occupancy\ of\ time}. \tag{6.2.1}$$

Chapter 7

Numerical results

7.1 Field Test of Lane Boundary Estimator

The real-world experimental environment shown in Figure 7.1 can be mapped to the sketch map shown in Figure 1.1, h in the real-world experimental scenario is 5 meters, while D is around 60 meters. The real-world data, including signal and video files, are simultaneously gathered via an industrial PC. The operating frequency of the road-side radar detector is at 10.5 GHz, and the bandwidth of RF is 50MHz. The hardware blocks of the road-side radar detector include the CMOS transceiver, digital signal processing (DSP) unit and two antenna arrays [16], all of which are integrated in a metallic box. The algorithm in DSP is written in C codes. The bin range can be approximately 78 cm long, although this may vary with the sensor resolution.

A total of 102,041 files involving 172 real-world vehicles are gathered, and each file comprises 512 voltage values. The frequency-domain information shown in Figure 2(b) can be obtained by using a FFT to convert the time-domain voltage values into the frequency domain. The summary of vehicle types in multi-lane environment is shown in Table 7.1, and reveals that the learning samples comprise different vehicle

(a) (b)

Figure 7.1: Real-world experimental environments in the suburbs. (a) is a sunny scenario, and (b) is a rainy scenario.

types, including motorcycle, small car, and large car. Notably, since the property of motorcycle is not suitable for identifying the lane positions, hence the motorcycle information will be filtered out in the entire learning procedure.

7.1.1 Results of Single-Value Information

According to the feature; that is, a single passing vehicle contributes just one count to the corresponding bin, proposed in [11], the accumulated vehicle count could be marshaled via a histogram shown in Figure 7.2. Then the GMM with different initials

Table 7.1: Summary of vehicle types in a multi-lane environment, which includes motorcycles, small vehicles (i.e., Honda civic, Toyota Camry, Volks wagan T4), and large vehicles (i.e., cement, gravel truck, and large bus).

	Motorcycle	Small vehicle	Large vehicle
Lane 1	8	35	1
Lane 2	1	36	7
Lane 3	7	38	10
Lane 4	2	27	0

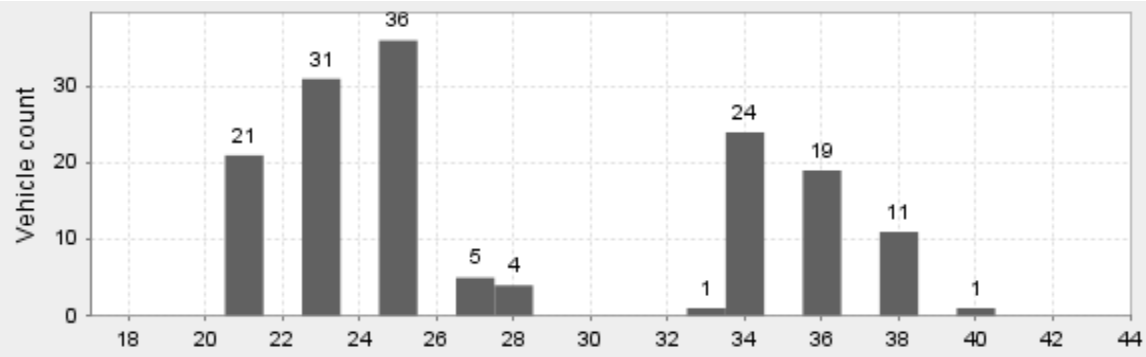
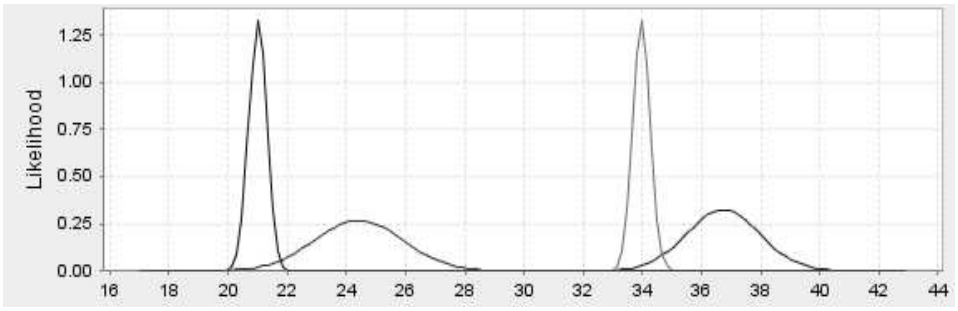
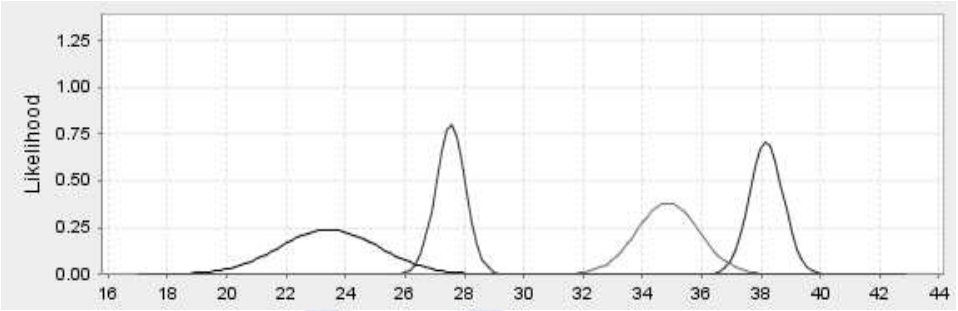


Figure 7.2: The histogram is the accumulated vehicle count mentioned in the patent [11]. That is; a single passing vehicle contributes just one count to the corresponding bin.

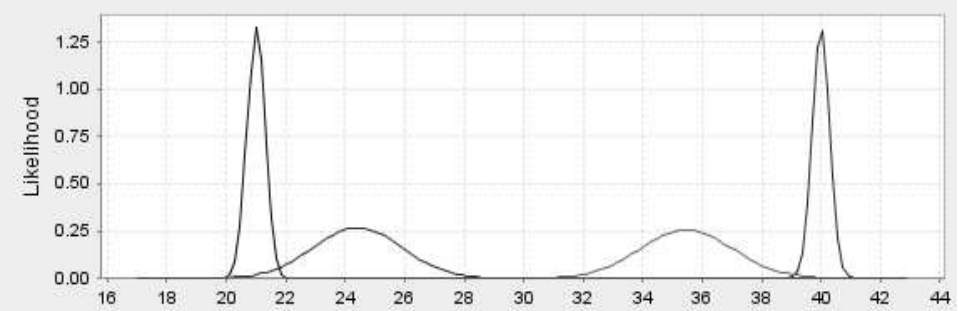
(including a K-mean initialization) are applied to work with an EM algorithm to address the lane position, and the learning result with different initials converges to the four primary outcomes shown in Figure 7.3. The peak of the probability density functions describes the center of each lane, and the valley of the probability density functions represents the lane boundary. The recorded video images are applied to verify whether the passed vehicle appeared in the corresponding lane range, but the result shows that the traffic count in each lane can not be captured accurately, since the Gaussian component associated with the larger variance mostly covers two adjacent lanes.



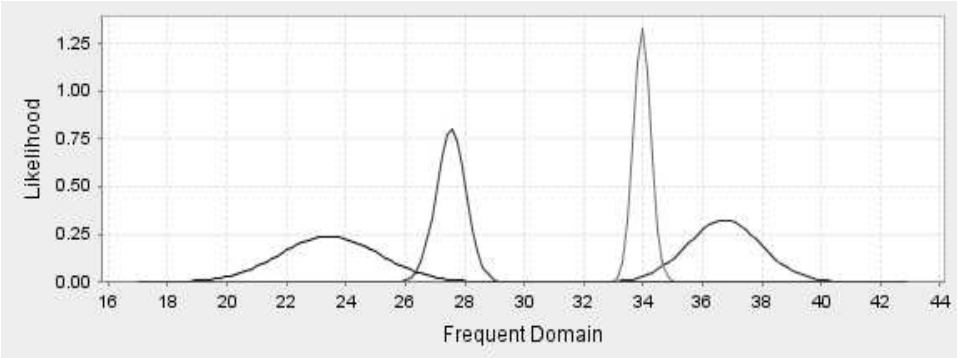
(a)



(b)



(c)



(d)

Figure 7.3: The learned Gaussian distribution density functions of GMM based on the histogram shown in Figure 7.2, (a)(b)(c)(d) are the results obtained by an EM algorithm associated with different initials.

Table 7.2: The learned parameters of GMM are obtained by an EM algorithm

	Weight	Mean	Variance
Component 1	0.36	22.15	0.98
Component 2	0.30	25.66	1.21
Component 3	0.27	35.08	0.95
Component 4	0.06	38.43	0.98

Table 7.3: The learned parameters of GMM are obtained by the variant of EM algorithm

	Weight	Mean	Variance
Component 1	0.39	22.29	0.97
Component 2	0.27	25.92	1.00
Component 3	0.28	35.10	0.95
Component 4	0.06	38.61	0.82

7.1.2 Results of Span and Conflict Information

The histograms of the accumulated span and conflict information are shown in Figure 7.4. The learned results of the EM algorithm and its variant based on the histogram are shown in Figure 7.5(a) and (b), respectively. Tables 7.2 and 7.3 represent the parameters of GMM obtained by the EM algorithm and its variant respectively.

The fact that lanes 3 and 4 contain more small cars and large cars than lanes 1 and 2 is shown in Table 7.1, but the accumulated span and conflict information in lanes 3 and 4 are less than that in lanes 1 and 2, it is because the peak information caused by passed vehicles in lanes 3 and 4 is less apparent than that in lanes 1 and 2 when vehicles passed through the front of the detection area simultaneously.

The learned cross points between lane 1 and lane 2 shown in Figure 7.5(a)(b) are around 24, respectively, hence, the lane boundary can be set to 24. For avoiding the ambiguous areas, the range of lane 1 can be set to 21-23, and the range of lane 2 can be set to 25-27. Similarly, the range of lane 3 can be set to 34-36, while the range of

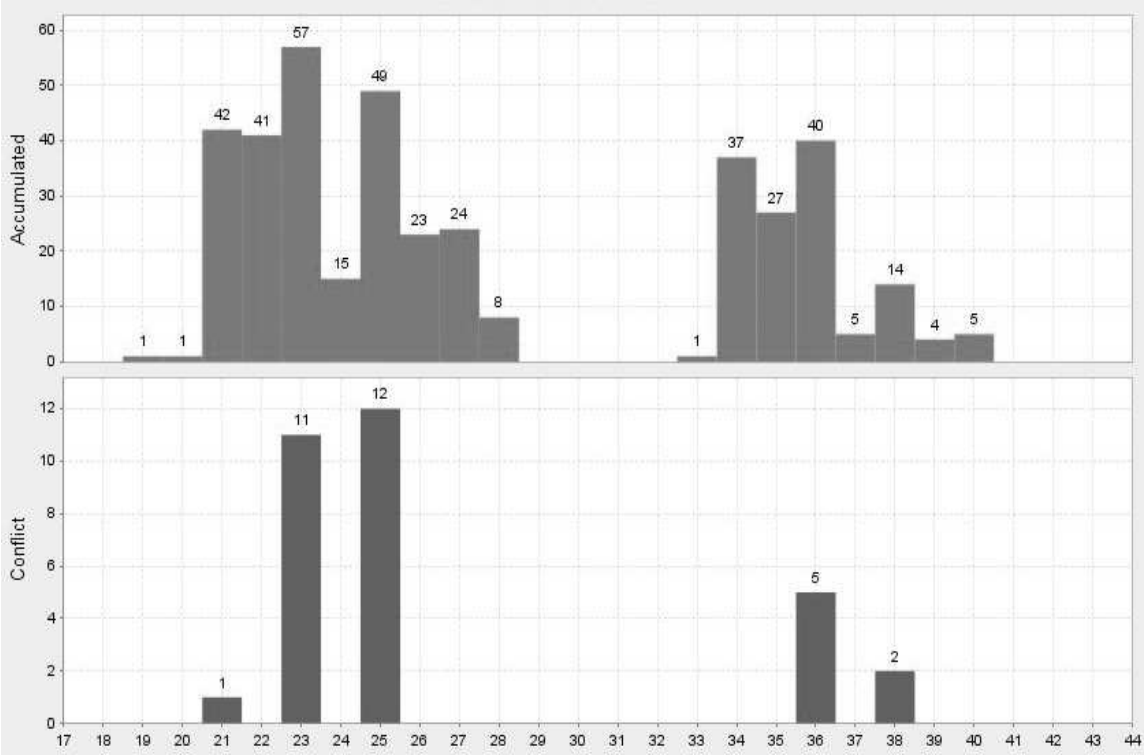
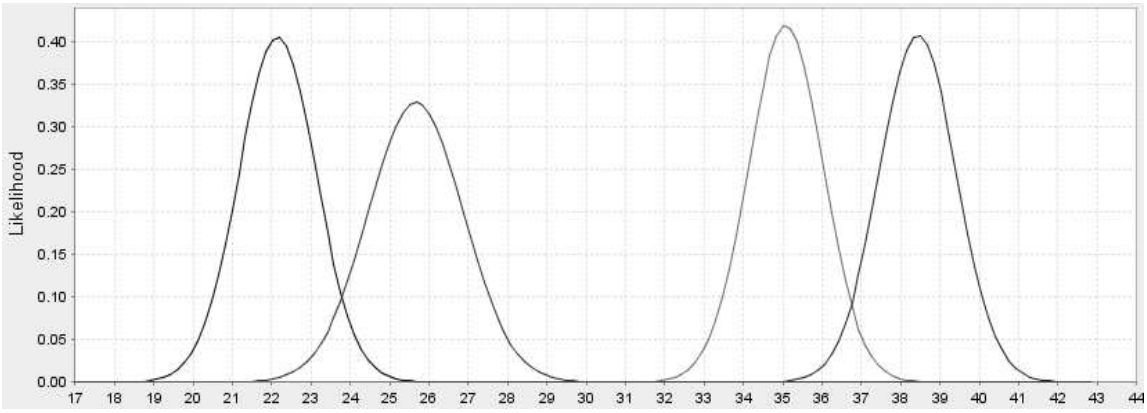
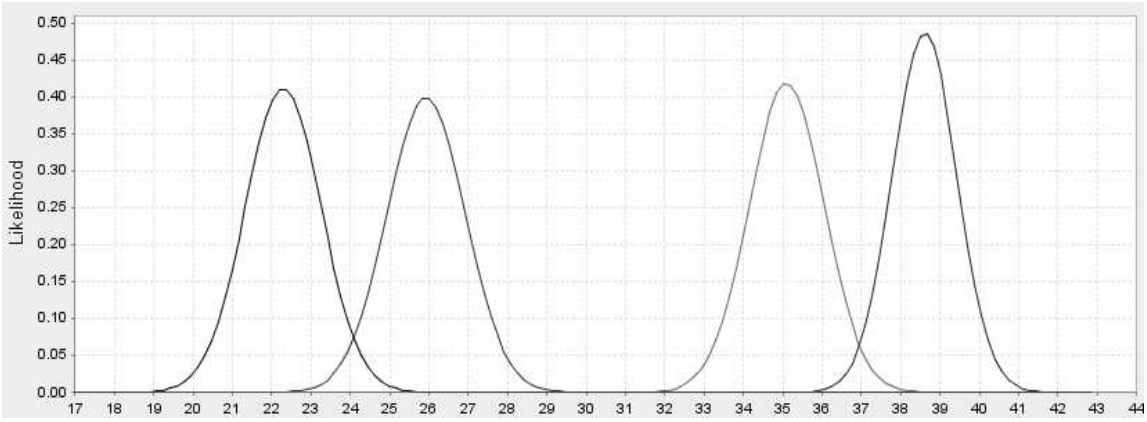


Figure 7.4: The histograms of real-world data. The top histogram is for the accumulated span information, while the lower histogram is for the conflict information.



(a)



(b)

Figure 7.5: (a) is the learned Gaussian distribution density functions obtained by using an EM algorithm, and (b) is the outcome obtained by using the variant of EM algorithm. The difference in learning results between (a) and (b) is that the variance of Gaussian components in (b) is not bigger than that in (a) underlying the same span and conflict information. Both of results are more accurate than that shown in Figure 7.3, and which indicated that the span and conflict information are superior to single-value information

lane 4 is set to 38-40. Using these rules, the classification results can be summarized in Table 7.4.

In our numerical experiments, based on the histogram generated by the proposed span and conflict information obtained from the gathered radar data, the mean of Gaussian components estimated by an EM algorithm is usually very close to the mean estimated by its variant, and hundreds of simulation results also show that the variance reduction always can be achieved by using the variant of EM algorithm. The variance of Gaussian functions shown in Table 7.3 obtained by the variant of EM algorithm is not bigger than that shown in Table 7.2 obtained by an EM algorithm. The computational burden of the variant of EM algorithm is less than that of EM algorithm, because the learning samples for each component only comprise these assigned points based on the highest likelihood.

Basically, the convergence speed of the EM variant is faster than the native EM, because the learning samples for each component only comprise these assigned points based on the component probability. The number of calculating Gaussian probability is about $N \times M$, while that of the EM variant is N, which is independent of the component number. Furthermore, the component variance shown in Figure 7.5(b) is less on average than that shown in Figure 7.5(a); that is, the proposed EM variant can be expected to reduce the variance.

The classification result shown in Table 7.4 exhibits that over 95% accuracy in terms of recording passing vehicles can be captured in the corresponding lanes. The accuracy can be further improved provided that the passed vehicles do not exhibit lane-changing behavior when passing through the detection area. Each vehicle type except for motorcycle in each lane almost reaches such a high rate of accuracy whether

Table 7.4: The classifying results (including vehicles exhibit lane-changing behavior).

	Lane 1	Lane 2	Lane 3	Lane 4	Learned results
21-23	33	1	0	0	34
25-27	3	43	0	0	46
34-36	0	0	48	0	48
38-40	0	0	0	27	27
Real-world	36	44	48	27	97.4%

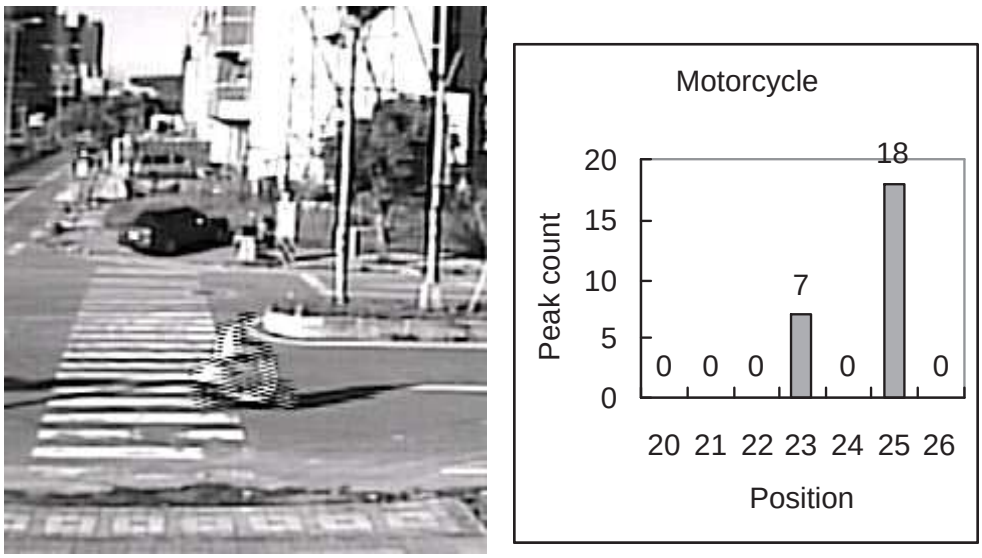
Table 7.5: The classifying results only consider small and large vehicles, and the passed vehicles do not exhibit lanechanging behavior when passing through the detection area.

	Sampling accuracy	Type I Error α
Lane 1	100%	3.9%
Lane 2	100%	2.8%
Lane 3	100%	2.3%
Lane 4	100%	2.5%

sunny or rainy days.

7.1.3 Motorcycles and Lane Change

Figure 7.6 and 7.7 show two scenarios that selected from the real-world data. Both scenarios exhibit the adverse information for identifying traffic lane boundaries. The fist case shown in Figure 7.6 indicates that a motorcycle is at the position close to the dividing line when passing through the detection area. Since motorcycles may appear at any position of roads possibly, the information of motorcycle is incapable of aiding in identifying lane positions. The second case shown in Figure 7.7 indicates the passed vehicle exhibits lane-changing behavior when passing through the detection area. Both cases generate the same span information that ranges from 23 to 25, and just fall in the conflict area. In other words, both scenarios cause difficulty in learning traffic lane positions. Hence, for achieving the higher accuracy, the installation location



(a) (b)

Figure 7.6: A motorcycle scenario and its histogram of peak count. This case shows that the track of motorcycle is not proper to describe the lane boundary information.

Table 7.6: Lane range

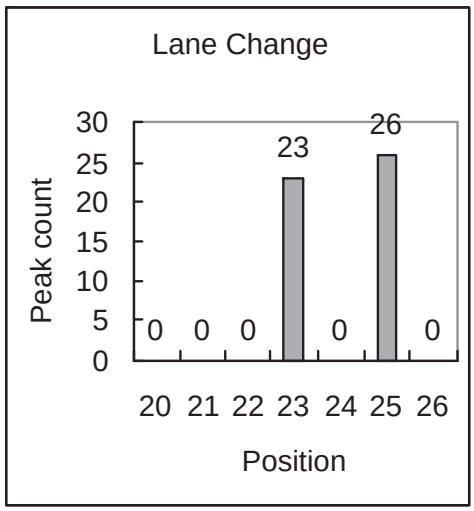
Lane	Range
1	16.38 (m) ∼ 17.94 (m)
2	19.50 (m) ∼ 21.06 (m)
3	26.52 (m) ∼ 28.08 (m)
4	29.64 (m) ∼ 31.20 (m)

of the road-side radar should avoid areas where vehicles usually exhibit lane change behavior, and the signal of motorcycles should be filtered out.

7.2 Field Test of Vehicle Classifier

Since lane 2 and 3 have more large vehicles relative to lane 1 and 4, thus, only lane 2 and 3 are selected as our targets for the vehicle classification learning. There are totally 43 and 48 vehicles obtained in lane 2 and 3 respectively by using an

(a)



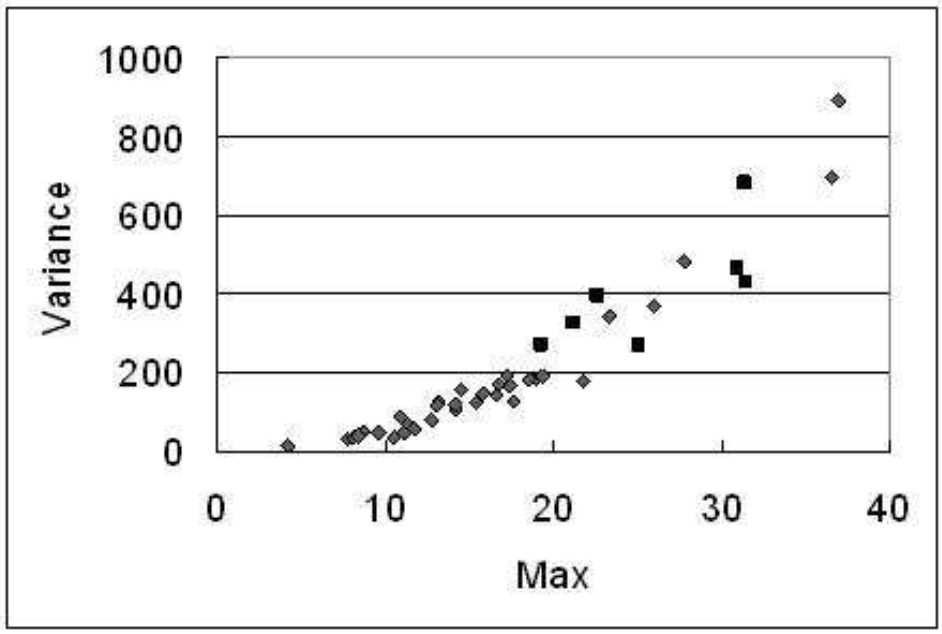
(b)

Figure 7.7: A lane change scenario of a small vehicle and its histogram of peak count. It indicates that the lane-changing behavior of vehicle is also not proper to describe the lane boundary information.

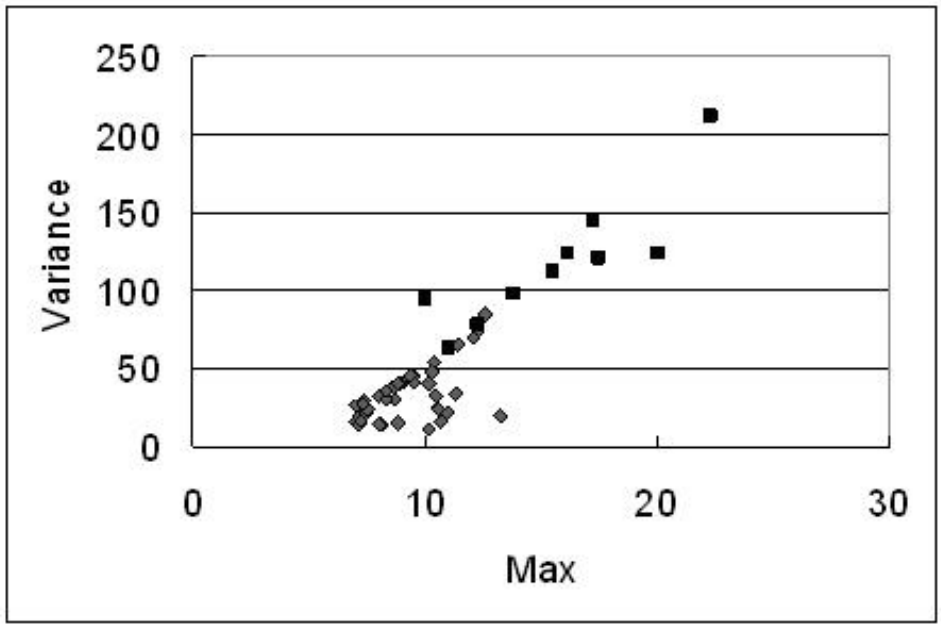
detection algorithm, if the passed vehicles do not exhibit lane change behavior when passing through the detection area, and the reflected signals of motorcycles are also filtered out. If only small and large vehicles are considered, the accuracies of vehicles appearing in the corresponding lanes can achieve 95% above.

Figure 7.8 shows vehicle feature scatter diagrams in lane 2 and 3. For distinguishing the different vehicle types' features, the large vehicles are marked as black squares, and small vehicles are represented by dark gray rhombuses. Each of vehicle sample are described by two features, average energy maximum, and average energy variance shown in x-axis and y-axis of Figure 7.8 respectively. Notably, the scales of axes are different for drawing all of the vehicle samples within a fixed screen size.

Table 7.7 describes the estimated GMM parameters in lane 2 solved by EM algorithm and the results can be geometrically plotted in Figure 7.9. The large vehicles



(a)



(b)

Figure 7.8: (a) shows the scatter diagram in lane 2. (b) shows the scatter diagram in lane 3.

Table 7.7: GMM parameters for lane 2. The component with the smaller mean represents the small vehicles, and the component with the larger mean represents the large vehicles.

	Component 1 (Small vehicles)	Component 2 (Large vehicles)
Weight	0.72	0.28
Mean	(137.13, 60.11)	(277.08, 126.72)
Variance(x)	41.16	56.54
Variance(y)	17.99	28.89
Covariance(x, y)	25.05	33.04

Table 7.8: Compare learned results and real-world date for lane 2. The diagonal counts are vehicles that are classified correctly, and non-diagonal counts are erroneous results.

	Small vehicles	Large vehicles	Learned results
Component 1	31	0	31
Component 2	5	7	12
In real-world	36	7	Accuracy : 88.37%

are marked as dark squares, and small vehicles are of gray squares. Each additional standard deviation, an oval is plotted. The last convergence value of maximum likelihood is -10.17, and the learned proportion of large vehicles is around 0.28, while the real proportion of large vehicles is 0.16. The accuracy for lane 2 is (31+7) / 43 = 88.37% shown in Table 7.8, where accuracy is equal to that the vehicles classified correctly are divided by total vehicles.

$$Accuracy = \frac{vehicle\ are\ chosen\ correctly}{total\ vehicle} \times 100\%$$

Table 7.9 is the estimated GMM parameters in lane 3 solved by EM algorithm and such results also can be geometrically shown in Figure 7.10. The last convergence value of maximum likelihood is -8.14, and the learned proportion of large vehicles is around 0.23, whereas the actual proportion is 0.21. The gap between real and statistic

Table 7.9: GMM parameters for lane 3. The component with the smaller mean represents the small vehicles, and the component with the larger mean represents the large vehicles.

	Component 1 (Small vehicles)	Component 2 (Large vehicles)
Weight	0.77	0.23
Mean	(92.3, 30.27)	(158.48, 55.26)
Variance(x)	15.97	31.48
Variance(y)	6.56	7.92
Covariance(x, y)	6.45	14.84

Table 7.10: Compare learned results and real-world date for lane 3. The diagonal counts are vehicles that are classified correctly, and non-diagonal counts are erroneous results.

	Small vehicles	Large vehicles	Learned results
Component 1	35	1	36
Component 2	3	9	12
In real-world	38	10	Accuracy : 91.67%

model is very close. The accuracy for lane 3 is (35+9) / 48 = 91.67% according to the results shown in Table 7.10.

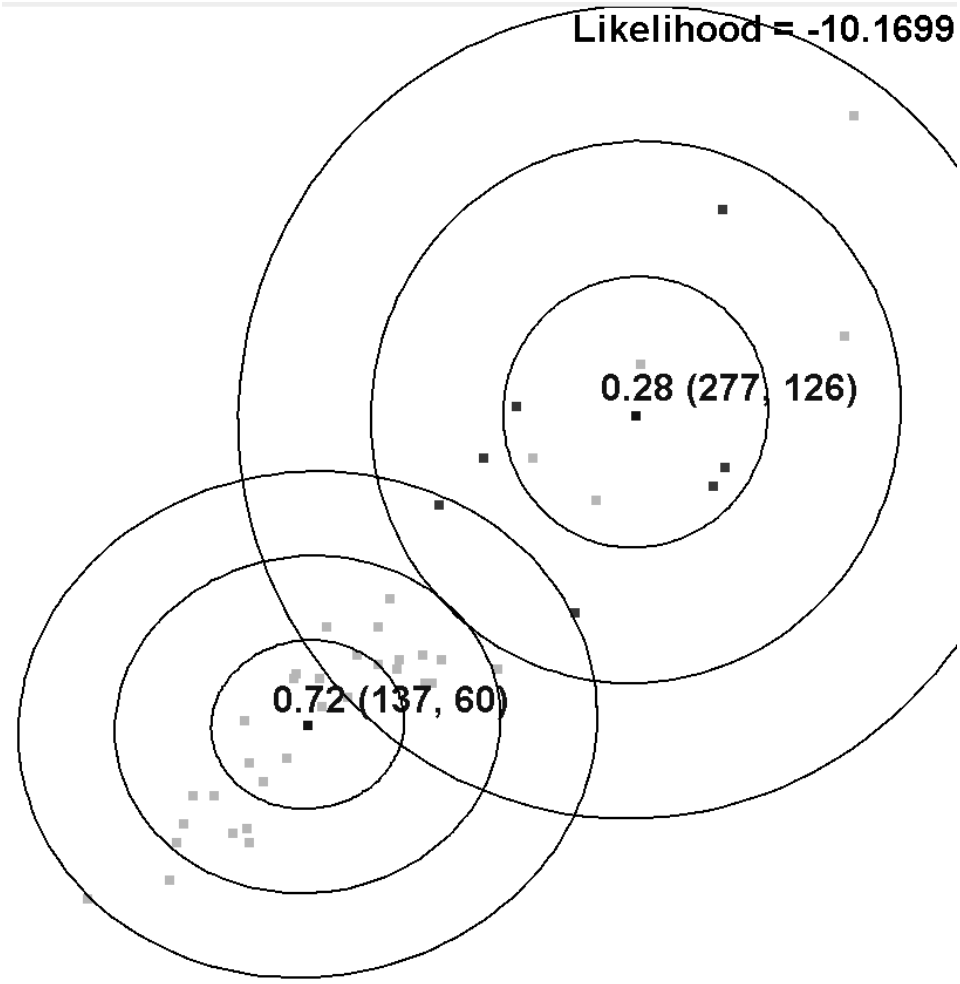


Figure 7.9: The learned result of a two-dimensional GMM for lane 2.

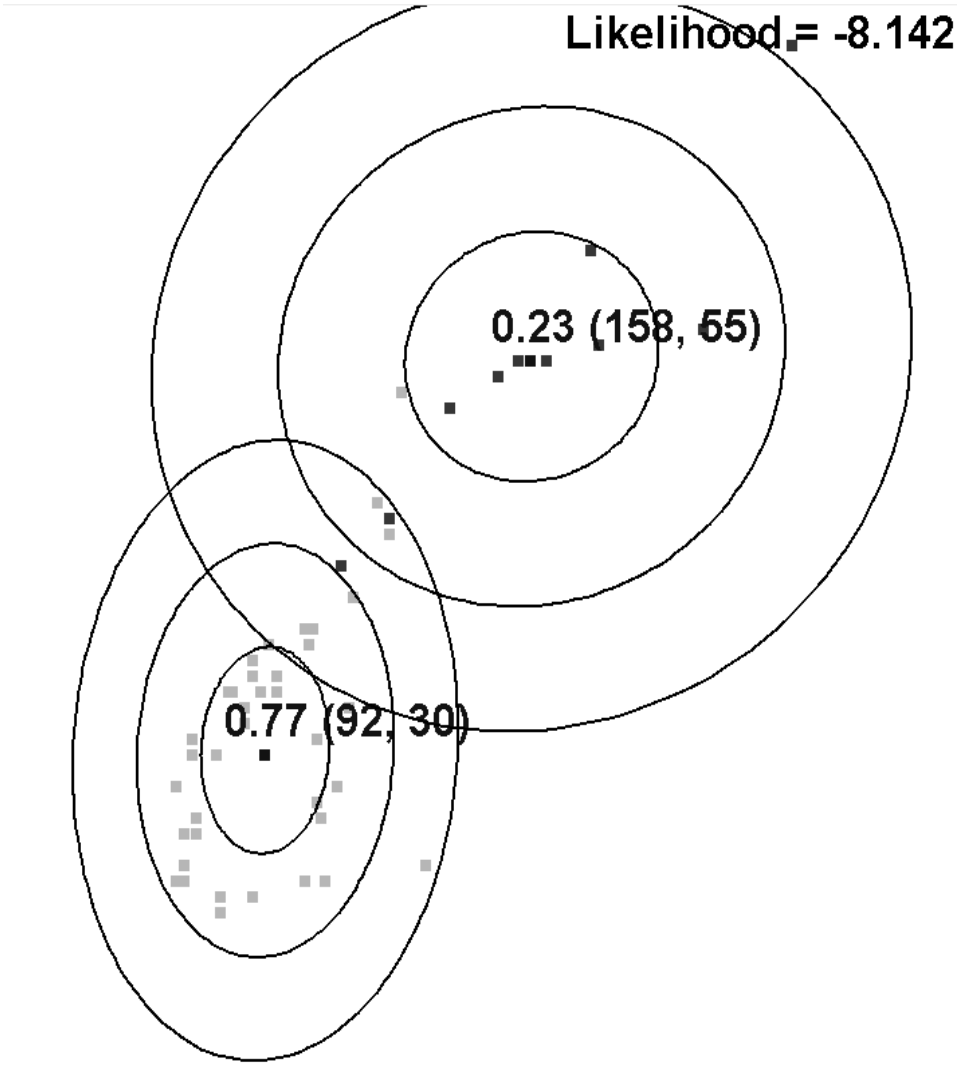


Figure 7.10: The learned result of a two-dimensional GMM for lane 3.

Chapter 8

Conclusion

The proposed learning procedure for the road-side radar detector is illustrated in detail above, and utilizes span information to form a reliable set describing the appropriate lane width. Additionally, the idea of conflict information is derived from the span information without any external interventions, and the proposed variant of EM algorithm achieves the variance of Gaussian components based on the collected data. Experimental results further indicate that passed vehicles except for motorcycle can be captured with over 95% accuracy, the accuracy can be further improved provided that the passed vehicles do not exhibit lane change behavior when passing through the detection area. This study avoids the ambiguous situation associated with the method mentioned in [11], and shows satisfying results of the proposed on-line traffic lane estimator.

Besides, this study proposes a real-time vehicle classifier applied to multi-lane environments, which introduces a two-dimensional Gaussian Mixed Model to form a learning model. A two-dimensional Gaussian Mixed Model utilizes a pair of features, maximal energy peak and maximal energy variance, extracted from the frequency-domain information of passed vehicles to form its learning samples, and integrates

an EM algorithm to enhance the accuracy. Field tests in the suburbs show that an accuracy of vehicle classification can attain above 88%, and present the results are better than all reviewable literatures on this field.

Bibliography

[1] Guillaume Leduc, "Road Traffic Data: Collection Methods and Applications," *Working Papers on Energy, Transport and Climate Change N.1*, 2008.

[2] P. T. Martin, Y. Feng, and X. Wang, "Detector Technology Evaluation," *Technical Report* , Utah Transportation Center, 2003.

[3] C. R. Bennett, A. Chamorro, C. Chen, H. de Solminihac, and G. W. Flintsch, "Detector Technology Evaluation," *Data Collection Technologies for Road Management*, Version 1.0, East Asia Pacific Transport Unit, The World Bank, Washington, D.C., April 2005.

[4] M. Schmidt, L. Giorgi, M. Chevreuil, S. Paulin, S. Turvey, and M. Hartmann, "GALILEO: Impacts on road transport," *JRC-IPTS Technical Report EUR 21865*, 2005.

[5] I. Urazghildiiev, R. Ragnarsson, P. Ridderström, A. Rydberg, E. Öjefors, K. Wallin, P. Enochsson, M. Ericson, and G. Löfqvist, "Vehicle Classification Based on the Radar Measurement of Height Profiles," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 245-253, 2007.

[6] M. Cherniakov, R.S.A Raja Abdullah, P. Jancovic, and M. Salous, "Forward Scattering Micro Sensor for Vehicle Classification," *IEEE Digital Object Identifier*, pp. 184-189, 2005.

[7] J. Gajda, R. Sroka, M. Stencel, A. Wajda, and T. Zeglen, "A Vehicle Classification Based on Inductive Loop Detectors," *IEEE Instrumentation and Measurement Technology Conference*, 2001.

[8] S. Gupte, O. Masoud, R. F. K. Martin, and N. P. Papanikolopoulos, "Detection and classification of vehicles," *IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS*, vol. 3, no. 1, pp. 37-47, Mar. 2002.

[9] D. Beymer, P. McLauchlan, B. Coifman, and J. Malik, "A real-time computer vision system for measure traffic parameters ," *In Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 496-501, 1997.

[10] G. L. Foresti, V. Murino, and C. Regazzoni, "Vehicle recognition and tracking from road image sequences," *IEEE Trans. Veh. Technol.*, vol. 48, no. 1, pp. 301-318, Jan. 1999.

[11] J. L. Waite, T. W. Karlinsey, and D. V. Arnold, Patent No. US 6556916 B2, Apr. 2003.

[12] D. M., Patent No. US2007016359 A1, Jan. 2007.

[13] H. Roe and G. S. Hobson, Improved discrimination of microwave vehicle profiles, *In Proc. IEEE MTT-S Int. Microw. Symp.* (1992), 717-720.

[14] S. J. Park, T. Y. Kim, S. M. Kang, and K. H. Koo, A Novel Signal Processing Technique for Vehicle Detection Radar, *IEEE MTT-S Digest* (2003), 607-610.

[15] B. P. Douglass, *Real-Time UML*, Addison-Wesley, ISBM: 0-201-32579-9, 1998.

[16] K. C. Tzuang, C. H. Chang, H. S. Wu, S. Wang, S. X. Lee, C. C. Chen, C. Y. Hsu, K. H. Tsai, and J. Chen, "An X-Band CMOS Multifunction-Chip FMCW Radar," *Proc. of the 2006 IEEE MTT-S Int. Microwave Symp. Dig.*, San Francisco, CA, pp. 2011 2014, Jun. 2006.

[17] S. C. Chen, M. L. Shyu, S. Peeta, and C. Zhang, "Learning-Based Spatio-Temporal Vehicle Tracking and Indexing for Transportation Multimedia Database Systems", *IEEE Transations on Intelligent Transportation Systems*, vol. 4, no. 3, pp. 154-167, Sep. 2003.

[18] Chengcui Zhang, Xin Chen, and Wei-bang Chen, "A PCA-based Vehicle Classification Framework", *Proceedings of the 22nd International Conference on Data Engineering Workshops*, vol. 4, no. 3, pp. 154-167, Sep. 2003.

[19] K. P. Karmann and A. von Brandt, "Moving object recognition using an adaptive background memory", in *Proc. Time-Varying Image Processing and Moving Object Recognition*, vol. 2, V. Capellini, Ed., 1990.

[20] Gian Luca Foresti, Vittor Murino, and Carlo Regazzoni, "Vehicle Recognition and Tracking from Road Image Sequences", *IEEE Transations on Vehicle Techinology*, vol. 48, no. 1, January 1999.

[21] H. Amirmehrabi and R. Viswanathan, "A New Distributed Constant False Alarm Rate Detector", *Digital Object Identifier*, vol. 33, no. 1, January 1997.

[22] A. De Maio, A. Farina, and G. Foglia, "Design and Experimental Validation of Knowledge-Based Constant False Alarm Rate Detectors", *Digital Object Identifier*, vol. 1, no. 4, pp. 308-316, Aug. 2007.

[23] Mashrur A. Chowdbury and Adel Sadek, *Fundamentals of Intelligent Transportation Systems Planning*, Artech House Publishers, ISBN: 1-58053-160-1, 2003.

[24] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification, 2nd ed.*, New York: Wiley, pp. 16-17, 2000.

[25] D. H. Fisher Jr, M. J. Pazzani, and P. Langley, *Concept Formulation: Knowledge and Experience in Unsupervised Learning*, Morgan Kaufmann Publishers, Inc., ISBN: 1-55860-201-1, 1991.

[26] Ethen Alpaydm, *Introduction to Machine Learning*, The MIT Press, ISBN: 0-262-01211-1, 2004.

[27] N. Shental, A. Bar-Hillel, T. Hertz, and D. Weinshell, "Computing Gaussian Mixture Models with EM using Equivalence Constraints," *Advances in neural information processing systems*, MIT Press, pp. 465-472.

[28] K. Wagstaff and C. Cardie, "Clustering with Instance-level Constraints", *Proceedings of the Seventeenth International Conference on Machine Learning*, pp. 1103-1110, 2000.

[29] K. Wagstaff, C. Cardie, S. Rogers, and S. Schroedl, "Constrained K-means Clustering with Background Knowledge," *Proceedings of the Eighteenth International Conference on Machine Learning*, pp. 577-584, 2001.

[30] D. Klein, S. D. Kamvar, and C. D. Manning, "From Instance-level Constraints to Space-level Constraints: Making the Most of Prior Knowledge in Data Clustering," *In The Nineteenth International Conference on Machine Learning*, 2002.

[31] J. Ma, T. Wang, and L. Xu, "A gradient BYY harmony learning rule on Gaussian mixture with automated model selection," *Neurocomputing*, vol. 56, no. 1, pp. 481-487, 2004.

[32] Z. Lu, "An iterative algorithm for entropy regularized likelihood learning on Gaussian mixture with automatic model selection," *Neurocomputing*, vol. 69, pp. 1674-1677, 2006.

[33] L. Li and J. Ma, "A BYY scale-incremental EM algorithm for Gaussian mixture learning," *Applied Mathematics and Computation*, vol. 205, no. 2, pp. 832-840, 2008.

[34] F. Cardinaux, C. Sanderson, and S. Bengio, "User authentication via adapted statistical models of face images," *IEEE Digital Object Identifier*, vol. 54, no. 1, pp. 361-373, 2006.

[35] Z. Zivkovic, "Improved Adaptive Gaussian Mixture Model for Background Subtraction," *In Proceedings of the 17th International Conference on Pattern Recognition*, pp. 28-31, 2004.

[36] Y. Wang, J. Yang, and Y. Zhou, "Color-texture segmentation using JSEG based on Gaussian mixture modeling," *Journal of Systems Engineering and Electronics*, vol. 17, no. 1, pp. 24-29, 2006.

[37] D. A. Reynolds, T. F. Quatieri, and R. B. Dunn, "Speaker Verification Using Adapted Gaussian Mixture Models," *Digital Signal Processing*, pp. 19-41, 2000.

[38] T. H. Dat, K. Takeda, and F. Itakura, "On-line Gaussian mixture modeling in the log-power domain for signal-to-noise ratio estimation and speech enhancement," *Speech Communication*, vol. 48, no. 11, pp. 1515-1527, 2006.

[39] M. H. Zhang and Q. S. Cheng, "Gaussian mixture modelling to detect random walks in capital markets," *Mathematical and Computer Modelling*, pp. 503-508, 2003.

[40] J. Ju, E. D. Kolaczyk, and S. Gopal, "Gaussian mixture discriminant analysis and sub-pixel land cover characterization in remote sensing," *Remote Sensing of Environment*, vol. 84, no. 4, pp. 550-560, 2003.

[41] I. Buckley, D. Saunders, and L. Seco, "Portfolio optimization when asset returns have the Gaussian mixture distribution," *European Journal of Operational Research*, vol. 185, no. 3, pp. 1434-1461, 2008.

[42] K. E. Ahmad and A. M. Abd-Elrahman, "Updating a nonlinear discriminant function estimated from a mixture of two Weibull distributions," *Mathematical and Computer Modelling*, vol. 19, no. 11, pp. 41-51, 1994.

[43] H.-j. Lee and S. Cho, "Combining Gaussian Mixture Models," *LNCS*, pp. 666-671, 2004.

[44] J. D. Banfield and A. E. Raftery,"Model-based Gaussian and non-Gaussian clustering," *Biometrics*, vol. 49, pp. 803-821, 1993.

[45] A. P. Dempster, N. M. Laird, and D. B. Rubin, "Maximum likelihood from incomplete data via the EM algorithm," *J. R. Statist. Soc.*, ser. B, no. 39, pp. 1-38, 1977.