

UC Berkeley

Dissertations

Title

Mobile and Stationary Computer Vision based Traffic Surveillance Techniques for Advanced ITS Applications

Permalink

<https://escholarship.org/uc/item/3kj790ws>

Author

Cao, Meng

Publication Date

2009-09-01

UNIVERSITY OF CALIFORNIA
RIVERSIDE

Mobile and Stationary Computer Vision based Traffic Surveillance
Techniques for Advanced ITS Applications

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

In

Electrical Engineering

By

Meng Cao

March 2009

Dissertation Committee:

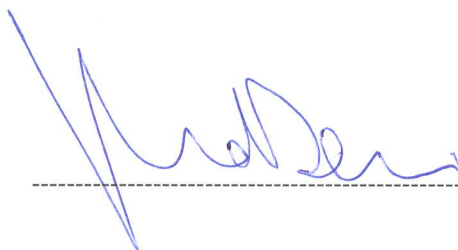
Professor Matthew Barth, Chairperson

Professor Gerardo Beni

Dr. ZuWhan Kim

Copyright by
Meng Cao
2009

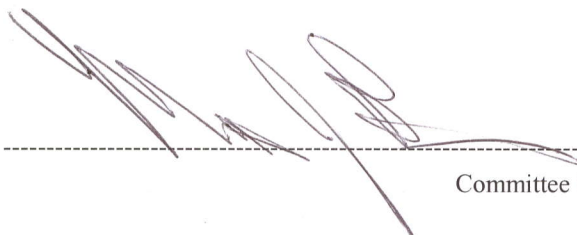
The Dissertation of Meng Cao is approved by



A handwritten signature in blue ink, appearing to be 'M. De...', written above a horizontal dashed line.



A handwritten signature in blue ink, appearing to be 'John Lin', written above a horizontal dashed line.



A handwritten signature in black ink, appearing to be 'M. De...', written above a horizontal dashed line.

Committee Chairperson

Acknowledgments

First of all, I would like to express my deep gratitude to my advisor Professor Matthew J. Barth, for the source of knowledge, guidance and support that he has been to me over the past years. His encouragement of my work and commitment to whatever project I would undertake was inspiring, and has made the time I have spent here at UC Riverside enjoyable. His sense of humor, the optimistic view toward life, as well as the talent of seeing the big picture, will benefit me in all my life. Without his help, I would not have been here.

I would also give my appreciation to Dr. ZuWhan Kim, my supervisor while I was staying at California PATH, for his deep insight, extensive knowledge and his ongoing effort of extending and improving his great software package, Cvzulib. He has been a tremendous mentor, collaborator, friend, and even a brother for me. I owe him a great many heartfelt thanks.

I am also grateful to my committee member Professor Gerardo Beni, for the helpful discussion, patient teaching and the encouragement since the first year in my graduate school. I had the luck to find with Dr. Xiaoyun Lu and Dr. Steven Shladover two extraordinary personalities as my co-supervisors, in the time I was staying at California PATH. With their motivation, dedication, and love for their work and life, they certainly did not teach me only “engineering”. Some of the work I present here is based on team

work with Dr. ZuWhan Kim, Dr. Xiaoyun Lu and Dr. Steve Shladover. Thanks for everything.

I am happy to have had some very excellent colleagues, the first of whom that I would like to thank being Anh Vu, for his effort and help in designing and constructing the vision sensor system used in this thesis. Thanks to Mike Todd, for the kind help whenever I needed. His big smile is always the sign of the TSR group. My thanks also go to Dr. Jie Du, Dr. Weihua Zhu, Dr. Kanok Boriboonsomin, George Scora, Alex Vu, Lili Huang, Anning Chen, Ashkan Sharafsaleh and Susan Dickey. I could not have afforded this Ph.D program without financial support, and I therefore would like to acknowledge the sponsors: the Department of Electrical Engineering of UC Riverside, CE-CERT at UC Riverside, UC Transportation Center and California PATH.

I have had a very wonderful time in the past 5 year, largely due to my great friends here. You know who you are-- Xiaotao Zou, who picked me up when I first landed at United States and been my roommate for three years. Zijun Fang, Gang Liu, Weihua Zhu, Bi Song, Peilin Fu, Grace Gao, Xiaoyu Wu, Shania Huang, Daniel Kuang, Yilei Xu from UC Riverside, as well as Allen Yao, Yiguang Xuan, Jing Xiong and Xiaomeng Zhong from UC Berkeley. I cherish you all.

I would also appreciate my childhood-mates: Yan Xu, Zishi Feng, Chenxin Zhang and college-mates: Chao Liu and Jiafeng Guo, for our friendships and all the experience

we have had together.

Lei, I am honored that you have stood steadfastly with me through all the ups and downs of life. May our love continue to grow.

Thanks, my dear Dad and Mom. So less words, and so much to add. I will never forget the words you put in my heart on that cold night when I was in middle school. I am strong when I am on your shoulder. I love you, for ever.

Dedicated
to my parents
and Lei

ABSTRACT OF THE DISSERTATION

Mobile and Stationary Computer Vision Based Traffic Surveillance Techniques for
Advanced ITS Applications

By
Meng Cao

Doctor of Philosophy, Graduate Program in Electrical Engineering
University of California, Riverside, March 2009
Professor Matthew Barth, Chairperson

During the past decade, new sensing technologies, such as inductive loops, laser range scanners, radar detectors and computer vision sensors have been greatly enhanced and applied to the Intelligent Transportation System (ITS) area. Among all these sensor systems, computer vision-based approaches are one of the most popular and promising techniques used in ITS for traffic evaluation and management, driver assistance, as well as other safety related research. This is primarily due to the advantages of easy maintenance, high flexibility, and low cost for traffic scene monitoring and analysis. Many stationary vision sensors have been already installed near the roadway, particularly at intersections. In addition, more and more vision sensors are now being installed on mobile vehicles, in order to have real time surrounding traffic information. This dissertation focuses on both mobile and stationary computer vision based traffic surveillance techniques, including the development of a new vision sensor, a survey and

development of vision algorithms, as well as their applications in three different aspects of ITS areas with high quantitative requirements. These areas are outlined in further detail below.

Portable Loop Fault Detection

For many years, it has been difficult to quantitatively measure real-time freeway traffic conditions. Numerous research projects have been carried out in traffic surveillance; for example, the Freeway Performance Measurement System (PeMS) operated by Caltrans and UC Berkeley collects, processes, aggregates, and examines traffic data primarily through loop detectors buried beneath the freeway. This type of stationary embedded loop sensor system provides a point measurement for the traffic flow, roadway occupancy and average speed. By aggregating these directly measured traffic data, they can be used to estimate and provide a larger picture of the traffic conditions in certain area. However, the results obtained from embedded loop sensors are not entirely reliable. The embedded loop data delivered to a Traffic Management Center (TMC) may contain errors at one or more sensors, and between the loop detector and the TMC database. As a result, loop fault detection is important. In this dissertation, a stationary-vision based technique has been developed as part of a Portable Loop Fault Detection Tool (PLFDT). This work is complementary to recent research focusing on aggregated faulty loop data at a macroscopic level (the macroscopic level generally

considers a large roadway network as consisting of links (roadways) and nodes (e.g., intersections)). The objectives of the PLFDT is to develop a real time, multi-lane, multi-vehicle tracking system for freeways using video cameras as the baseline measurement technique to compare the loop detection signal for direct fault detection for inductive loop system.

Localized Traffic Density Measurement

The embedded loop sensor system provides a direct measurement to traffic flow, roadway occupancy and average speed (only for double-loop detector). This type of sensor network does not directly measure traffic density; instead it can only be estimated. In this dissertation, we have developed systematic techniques to measure traffic conditions by utilizing both on- and off-board computer vision systems. A unique development technique is a combined computer vision and Global Positioning System (GPS) equipped mobile traffic surveillance system to measure localized traffic density. In addition, we correlate the localized density measurement with estimates from embedded loop sensor system using a space-time diagram. Experiments have shown the complementary nature of these sensing techniques. Further, most traffic surveillance computer vision algorithms and techniques are typically based on observing vehicles from stationary rectilinear cameras mounted near roadways. For many applications, some of the key tasks include extracting traffic information such as average traffic speed, flow,

and density. However, less research has been carried out in observing and generating the localized traffic map around specific vehicles in the traffic stream. In this dissertation, we have developed a vision-equipped vehicle test bed for traffic surveillance purposes and have experimentally demonstrated the generation of localized traffic density from video processing and synthesizing. In contrast to the off-board surveillance systems (e.g. embedded loop sensor networks and stationary vision monitoring system), this type of on-board surveillance system provides a temporal- and spatial- continuous measurement of the localized traffic density.

One of the key components developed is an Orthogonal Omni-directional Vision (OODV) System that has been developed to observe lane-level activity surrounding a vehicle, as well as the ability to observe the surrounding roadway geometry. This vision system uses a special catadioptric mirror providing a 360 degree orthogonal view of the environment. It is different from other catadioptric mirror-based Omni-directional vision systems in that it directly provides an orthogonal image without the need of warping a polar-coordinate based image to a perspective view. Based on this unique OODV, a roadway traffic surveillance system was designed and implemented. It consists of three major components:

- ◆ A GPS stamped roadway traffic data collection technique;
- ◆ Post video processing that includes automatic vehicle detection/tracking with the

ability to correct using interactive tools;

- ◆ A traffic parameter (localized density) estimation process.

Combined with a GPS receiver that provides approximately 2 - 3 meters spatial resolution, this traffic surveillance system can be applied not only in several traffic applications which require localized traffic density/flow/average speed measurements, but also in some other applications that require detailed roadway geometry acquisition, and vehicle activity analysis. Based on experimentation, it has been shown that the designed mobile surveillance system reports a high detection rate under the dynamic freeway environment in the experiments, with assistance from a human interactive detection module.

In order to have a better understanding of dynamic traffic conditions, we have incorporated this localized traffic density measurements into a Dynamic Roadway Network Database (DRND), which has been developed to fuse the roadway traffic data and the probe vehicle data. We believe that with the increasing use of on-board vision sensors, more and more localized traffic information samples can be reported to this type of database. The combinational analysis of temporal-spatial variable density and the embedded loop sensor data will provide a better and more reliable method for traffic condition estimation and prediction.

Bicycle Safety Support System

In addition to these new traffic data collection/analysis techniques and verification process, computer vision techniques are being applied in safety studies as well. In the third part of this dissertation, stationary vision based observations have been made of the timing of bicyclists' intersection crossing maneuvers, to support of efforts in improving traffic signal timing to accommodate the needs of bicyclists. Video recordings were made of bicyclists' crossings and the video images were processed to extract the bicyclists' trajectories. These were synchronized with video images of the traffic signals so that the timing of the bicyclists' maneuvers could be determined relative to the signal phases. The processed data have yielded cumulative distributions of the crossing speeds of bicyclists who did not have to stop at the intersection and the start-up times and final crossing speeds of the bicyclists who had to cross from a standing start. This study provides a foundation in recommendation of minimal green signal time in terms of safety purpose.

The key contributions of this dissertation are:

- ◆ A unique OODV System has been designed and developed;
- ◆ A vision based systematic approach for developing a mobile traffic surveillance system has been proposed and implemented;
- ◆ A time-space diagram based flow calculation method from localized density has been proposed and experimentally verified;

- ◆ A stationary-vision based technique has been developed as part of a Portable Loop Fault Detection Tool (PLFDT) to provide baseline data; and
- ◆ A stationary vision-based intersection monitoring system has been developed for the quantitative study of bicyclist crossings at signalized intersections.

Contents

List of Figures	xvi
List of Table	xviii
1. Introduction	1
1.1 Computer Vision as A Powerful Tool in ITS	1
1.2 Problem Statement	3
1.3 Contributions of the Dissertation	9
1.4 Organization of the Dissertation.....	10
2. Background and Literature Review	12
2.1 Roadway Traffic Surveillance Systems.....	12
2.1.1 Three Levels of Traffic Data Representation.....	12
2.1.2 Embedded Loop Data Collection Infrastructure.....	13
2.1.3 PeMS System	15
2.2 Review of Vision Based Vehicle Detection Algorithms	17
2.2.1 Knowledge Based Methods.....	17
2.2.2 Motion Based Methods	18
2.2.3 Some Other Algorithms.....	18
2.3 Review of Vision Based Vehicle Tracking Algorithms	20
2.3.1 Region Based Tracking Algorithms.....	21
2.3.2 Contour Based Tracking Algorithms	21
2.3.3 Feature Based Tracking Algorithms	21
2.3.4 3D Model Based Tracking Algorithms.....	21
2.4 Review on Loop Fault Detection Systems	22
2.5 Review on the Studies of Bicyclist Intersection Crossing Maneuvers	25
3. OODV System and Video Processing Algorithm.....	26
3.1 The Design of OODV System.....	26
3.1.1 The Structure of the OODVs	26
3.1.2 Characterization of the OODV Projection Model	28
3.1.3 Some Discussion on the OODV Projection Model	31
3.1.4 Highlights of the OODV Projection Model.....	33
3.2 OODV-Based Vehicle Detection Algorithm.....	33
3.3 Rectilinear Camera-Based Vehicle Detection and Tracking Algorithm	42
4. Stationary Vision Based Portable Loop Fault Detection Tool (PLFDT) Development.....	45
4.1 Vision Sensor Provides Baseline Data	45
4.2 Development of Portable Loop Fault Detection Tool (PLFDT).....	47
4.2.1 Overall System Structure of the PLFDT	48
4.2.2 Mobile Pole for Roadside Video Camera Mounting	49

4.2.3 Interface with Control Cabinet	50
4.3 Vision Based Multi-lane Vehicle Tracking Software Development	51
4.3.1 Synchronization of the two Computers with Wireless Communication	51
4.3.2 Real-Time Multi-lane Vehicle Tracking Algorithm	53
4.4 Comparison of Physical Loop and Virtual Loop	55
4.5 Experiment Data Analysis	56
4.6 Conclusions	60
5. Mobile Traffic Surveillance System Using a Unique OODV Sensor	62
5.1 On-Board Vision Sensor to “Sense” Soundings	62
5.2 Vision Based Mobile Traffic Surveillance System	65
5.2.1 Localized Traffic Map Captured with On-vehicle Vision Sensors	65
5.2.2 Calibration of the Vision Sensors	69
5.2.3 Video and GPS Data Acquisition	74
5.2.4 Localized Traffic Map Generation from Video Post Processing	76
5.3 Dynamic Roadway and Vehicle Traffic Data Integration	78
5.3.1 Dynamic Roadway and Vehicle Database Architecture	78
5.3.2 Relationship between Loop Data and Mobile Traffic Data	80
5.3.3 Experiment Verification	83
5.4 Conclusions	91
6. Vision Based Quantitative Measurements of Bicyclist Intersection Crossing Times	93
6.1 Computer Vision as a Reliable Method for Quantitative Measurements	93
6.2 Data Collection and Processing Approach	94
6.2.1 Bicyclist Traffic Data Collection	94
6.2.2 Data Processing	96
6.3 Analysis of Bicyclist Crossing Data	99
6.4 Summary	109
7. Conclusions and Future Work	111
7.1 Conclusions	111
7.2 Publications Resulting From This Research	113
7.3 Future Work	114
Bibliography	115
Appendix A	121
Derivation of Eqn.3-1	121
Appendix B	122
Least Square Error Line Fitting	122

List of Figures

Figure 1-1 Embedded loop sensors displayed on Google map	4
Figure 1-2 Mobile traffic data is complementary to the loop sensor data	7
Figure 2-1 Loop data collection process [3].....	14
Figure 2-2 Overview of the data process of PeMS system [3].....	16
Figure 3-1 (a) Layout of the OODV; (b) OODV used in the project.....	27
Figure 3-2 Comparison of OODV image with hyperbolic ODV image.....	27
Figure 3-3 Geometry of the OODV	28
Figure 3-4 3-D Reconstruction of the OODV mirror shape	30
Figure 3-5 Cross-section simulation of the OODV mirror shape.....	30
Figure 3-6 Rendered images of checker board using the OODV projection model.....	30
Figure 3-7 Graphical display of the imaging distortion map (a and b have different depth)	32
Figure 3-8 OODV based roadway experiment setup.....	34
Figure 3-9 Flow chart of the OODV based vehicle detection algorithm.....	35
Figure 3-10 (a).Background image and (b) Foreground image.....	39
Figure 3-11 Occlusion analysis in case (1).....	39
Figure 3-12 Occlusion analysis in case (2)	39
Figure 3-13 Lane structure detection process.....	41
Figure 3-14 An example of vehicle tracking under the presence of shadows	44
Figure 4-1 Overall system structure	48
Figure 4-2 Mobile retractable pole setup	49
Figure 4-3 Interfacing with control cabinet and smart card	50
Figure 4-4: An example feature tracking.....	54
Figure 4-5: Virtual loop is triggered when a vehicle reaches it.....	54
Figure 4-6 Software structure and interaction between the two laptops	56
Figure 4-7 Laptop RS232 serial interface with C922 3M Canoga Card	57
Figure 4-8 Visual loop in the video stream	58
Figure 4-9 Comparison of the virtual loops and physical loops data	59
Figure 5-1 View angle of on-vehicle cameras	63
Figure 5-2 Different applications of on-vehicle cameras	64
Figure 5-3 Vision sensor and GPS-equipped vehicle	67
Figure 5-4 Captured roadway image data	67
Figure 5-5 Top-down and lateral display about the field of view.....	68
Figure 5-6 Coordinate systems used in the mobile traffic surveillance system.....	68
Figure 5-7 (a) Linear fitting for the scaling parameter for a particular depth (b) Relationship between the scaling parameters and depth	72

Figure 5-8 (a) OODV coordinates re-projection in the OODV image plane (b) Auxiliary coordinates re-projection in the rectilinear image plane.....	73
Figure 5-9 GPS trajectories of the test trip on Google Earth.....	75
Figure 5-10 Duplicated detection in multiple sensors	77
Figure 5-11 Dynamic roadway and vehicle data architecture	79
Figure 5-12 Time-space diagram.....	81
Figure 5-13 (Part A) Mobile and loop data comparison results for Westbound VDS	86
Figure 5-13 (Part B) Mobile and loop data comparison results for Westbound VDS	87
Figure 5-13 (Part C) Mobile and loop data comparison results for Westbound VDS	88
Figure 5-14 (Part A) Mobile and loop data comparison results for Eastbound VDS	89
Figure 5-14 (Part B) Mobile and loop data comparison results for Eastbound VDS	90
Figure 5-14 (Part C) Mobile and loop data comparison results for Eastbound VDS	91
Figure 6-1 Video data collection sites at Telegraph and Russell in Berkeley (left) and Park Blvd. and El Camino Real in Palo Alto (right)	96
Figure 6-2 Bicyclists tracking results.....	97
Figure 6-3 Example of extraction of needed values for standing start from bicyclist trajectory plot.....	98
Figure 6-4 Cumulative distributions and histograms of speed of rolling start bicycles (mph)....	101
Figure 6-5 Cumulative distributions and histograms of offset times for standing start bicyclist crossings	103
Figure 6-6 Cumulative distributions and histograms of final crossing speed for standing start bicyclist crossings (mph).....	104
Figure 6-7 Distributions of total time for bicyclists to cross El Camino Real from a standing start.....	106
Figure 6-8 Distribution of duration of green phase (s) when bicyclists were crossing El Camino at Park Blvd.....	107
Figure 6-9 Distribution of duration of green phase (s) when bicyclists were crossing Telegraph at Russell St.....	108
Figure 6-10 Completion times of bicyclist crossings from standing start, relative to yellow onset.....	108

List of Table

Table 3-1 Comparison rate with and without lane-level occlusion analysis.....	39
Table 5-1 Traffic Parameters definition and equations.....	78
Table 5-2 List of VDS that is used for comparison	84

1. Introduction

1.1 Computer Vision as A Powerful Tool in ITS

In recent years, computer vision sensors have become increasingly important for many Intelligent Transportation System applications, mainly due to their fast response, easy installation, simple operation and maintenance, high flexibility, low cost, and their ability to monitor wide areas. Object detection and tracking is one of the major research areas in computer vision. One significant application of object detection and tracking is traffic video analysis, which is part of a variety of Intelligent Transportation System applications. Two of the most widely researched transportation applications relate to traffic surveillance systems and driver assistant systems [55].

In general, systems developed for traffic surveillance purposes aim at having a understanding of the real time traffic conditions, which require the analysis of the static environment (from a stationary vision sensor to monitor an area of interest) and the detection of static or moving obstacles, such as vehicles, bicycles, and pedestrians. Image processing and computer vision techniques are being applied to the analysis of video sequences of traffic, in order to extract traffic information, such as vehicle counting, vehicle classification, speed measurement, traffic flow, and so on. Most of these systems are based on fixed position installations, such as at urban intersections and along freeways. They offer considerable improvements over existing infrastructure-based traffic surveillance systems, such as embedded loop sensor systems and microwave detector systems.

In addition to these stationary monitoring systems, there has been increased use of “mobile” sensors. With the improvements in on-board vision sensor technology together with other technologies such as GPS technology and increasing mobile computational power, vehicle-based vision sensor systems are increasingly prevalent for the purpose of observing a vehicle’s immediate “sensed” surroundings, as a major component of driver assistant systems. Vision sensors can provide a great amount of information about the surrounding environment, but the scenes can be complex and highly demanding. Typically with computer vision techniques, different physical quantities need to be measured and pattern recognition needs to be applied before useful information can be delivered to the driver. Example subtasks include [6]:

- 1) Determining the vehicle’s position relative to the lane and checking for obstacles on the path or known road signs (for road following/keeping);
- 2) Recognizing specific vehicles’ characteristics and computing the time-to-impact (for vehicle platooning);
- 3) Sensing multiple lanes and detecting obstacles (for passing vehicles and collision avoidance);
- 4) Measuring the distance among parked vehicles (for automatic parking); and
- 5) Determining the position and following the movements of the driver’s eye and head (for driver-status monitoring).

Thus, computer vision techniques are becoming powerful tools in ITS applications to better understanding traffic conditions so that roadway travel can be made as efficient as possible (i.e., traffic surveillance systems) and in providing information surrounding vehicles so that driving can be made as safe as possible (i.e., driver assistant systems).

1.2 Problem Statement

Although large investments have been made and are continuing to be made in transportation infrastructure, inefficient (and often unscientific) transportation management and traffic control fail to keep pace with increases in population and car use (i.e., increased demand typically measured in Vehicle Miles Traveled). This results in roadway congestion. To minimize the impacts of congestion, it is becoming increasingly important to accurately measure real-time roadway-traffic conditions and manage it.

During the past decade, numerous research projects have been carried out in traffic surveillance in terms of measuring traffic performance (e.g. [1][2]). A traffic surveillance system will typically evaluate traffic conditions based on the following key parameters:

- ◆ Traffic flow Q : the rate at which vehicles pass a given point on the roadway, and is normally given in terms of vehicles per hour;
- ◆ Traffic density K : refers to the number of vehicles present on a given length of roadway and given in terms of vehicles per mile or vehicles per kilometer.
- ◆ Average speed V : also called the space mean speed, and can be found by averaging the individual speeds of all of the vehicles in the study area.

These parameters are defined in the 2D space of time and position (i.e., locations

along the roadways). Currently, most of the traffic surveillance systems obtain data at specific locations at different times of the day. An example is the California statewide loop sensor system (e.g., PeMS). The system collects real-time traffic flow, and occupancy data directly from loop detectors embedded in California's freeways and makes it available for transportation management, research, and commercial use. Figure 1.1 illustrates the locations of embedded loop sensors (the small dots) embedded in the freeways around Sacramento, California, United States, which is displayed on Google Map.

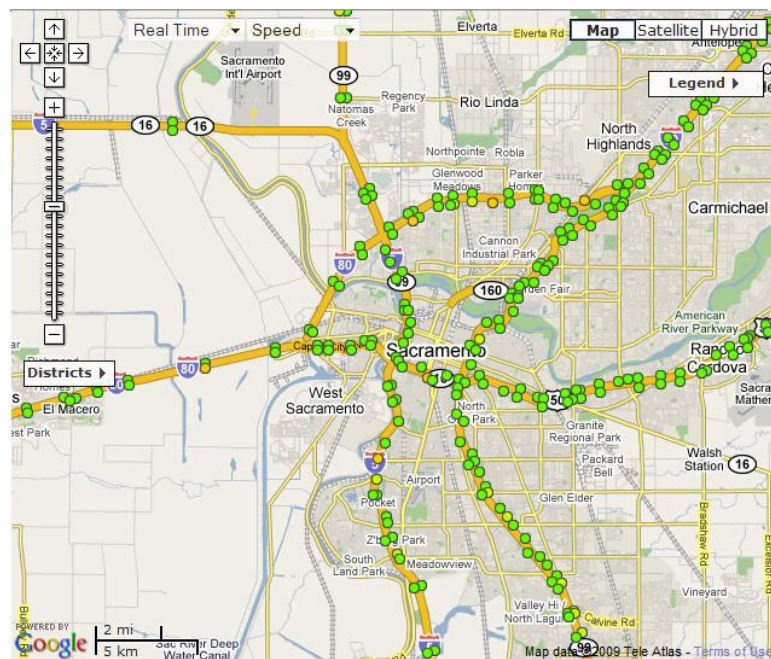


Figure 1-1 Embedded loop sensors displayed on Google map

This statewide sensor system consists of over 28,000 sensors located on major freeways. Most of these traffic sensors are inductive loops, embedded in the pavement. The collection, aggregation and analysis of these embedded loop data provide a systematic way to estimate the macroscopic level traffic conditions (see details in Chapter

2). However, these loop detector data are not entirely reliable. The loop detector data delivered to a Traffic Management Center (TMC) may contain errors created at one or several sensors, as well as in the communication between the loop detector and the TMC database. Therefore, loop fault detection is critical. Loop fault detection at a high level is done usually through an analysis of aggregated loop data [42]. Such an approach is indirect, and the faults that can be detected are usually large scale problems such as electric or communication system faults. This is actually one of the limitations to the macroscopic and mesoscopic approaches.

Only the detection at the control cabinet level can directly detect the fault(s) in hardware and software, isolate them from the communication fault, and correct the faults permanently. Data at this level can either be the ones processed by loop detector cards, which will be loop ON/OFF time instances or occupancies; or the raw loop pulse signal before the loop card. More importantly, an independent data source or baseline data needs to be obtained so that the loop detector readings can be validated against the baseline data. Obviously, computer vision sensors and stationary computer vision techniques are one of the best choices for providing such baseline measurements.

In this dissertation, a stationary-vision based technique has been developed as part of a Portable Loop Fault Detection Tool (PLFDT). This work is complementary to recent research focusing on aggregated faulty loop data at a macroscopic level. The objectives of the PLFDT is to develop a real time, multi-lane, multi-vehicle tracking system for freeways using video cameras as the baseline measurement technique to compare the loop detection signal for direct fault detection for inductive loop system.

As mentioned above, the traffic parameters estimated from embedded loop sensor system are space discontinuous: they are only spatial samples along the roadway. The cumulated traffic parameters (traffic flow and occupancy) and their derivatives (e.g., traffic density) from the embedded loop based surveillance system only have meanings from the view of statistics. By aggregating these space discontinuous traffic data over space and time, we can describe a larger picture of the traffic conditions over a range of roadways during a certain time period. However, there are missing details in between the embedded loop sensors along the roadway. Thus, it will be useful and meaningful to develop a mobile traffic surveillance system to obtain the spatial and temporal continuous traffic condition. Obviously, vision sensors and computer vision techniques are the best choices to equip such a mobile traffic surveillance system because of its low costs, high flexibility, and the ability to monitor a wide range of areas.

In past decade, a lesser amount of research in ITS applications has been carried out in observing traffic conditions around specific moving vehicles in the traffic stream. Trivedi et al. [56] introduced a Dynamic Panoramic Surround Map (DPSM) generation technique based on two omni-directional vision sensors on the vehicle. However this DPSM is generated as part of a driver assistant system, not mainly for traffic surveillance purposes. Monitoring the roadway traffic based around a moving vehicle can be considered as a good complement to the existing loop sensor surveillance systems since it provides both time and space varying traffic data. As illustrated in Figure 1-2, a mobile traffic surveillance system has the ability of direct measurement to the region in between embedded loop sensors. Those parameters estimated from a mobile traffic surveillance

system are very useful to model and predict the traffic conditions at a range of roadway locations rather than at specific points. Further, relationships can be established between the traffic parameters obtained from vision sensor equipped mobile traffic surveillance system with the embedded loop sensor based surveillance system for complementary data fusion. This can be viewed as a two-level problem. First, we have to verify that mobile traffic data can be related to the loop sensor data. Second, with the increasing mobile traffic data collection, it is important to setup a model to fuse, estimate, and analyze the traffic data obtained from mobile surveillance system and loop sensor system. In this dissertation, we focus on the first level.

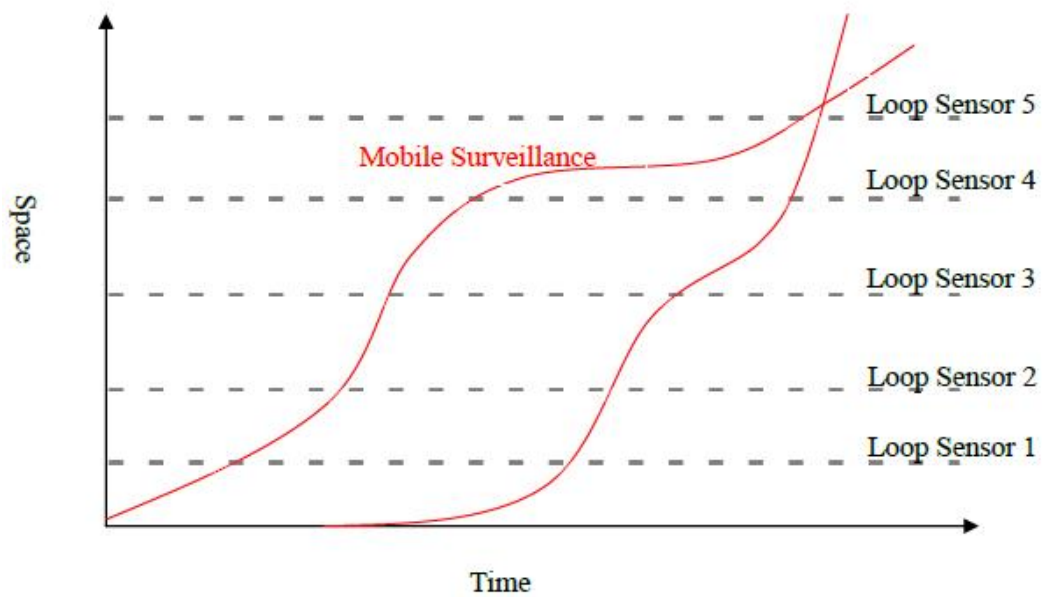


Figure 1-2 Mobile traffic data is complementary to the loop sensor data

Today's computer vision systems used for traffic surveillance can be grouped into two primary categories: those employing conventional rectilinear cameras (see, e.g. [3]) and those employing omni-directional cameras that make use of hyperbolic or spherical

mirrors (see, e.g. [5]). The rectilinear camera uses a simple geometric projection relation and most of the algorithms are designed on the basis of this perspective projection. However, the field of view is typically quite narrow (approximately 30 ~ 45 degrees). Therefore it is difficult to monitor the entire surrounding area with a single camera. On the other hand, an omni-directional camera based on hyperbolic or spherical mirrors can observe the 360 degree surroundings and thus provide a much wider angle of view than rectilinear cameras. For example, this property makes it possible to monitor an entire intersection with a single camera. However, the images resulting from these vision systems often require a preprocessing step to convert these images from a polar-coordinate to a perspective view prior to analysis [5].

In this dissertation, we introduce a novel design for a unique Omni-Directional Vision Sensor (ODVS), which can be used for the mobile vision based traffic surveillance. Prototype systems that employ these mobile ODVS provide another view point on roadway traffic conditions which can readily be integrated with embedded loop sensor based stationary traffic surveillance systems.

In addition to traffic surveillance systems, computer vision techniques can also be applied to detect and track vehicle/bicycle/pedestrians for safety applications. To support efforts to improve traffic signal timing to accommodate the needs of bicyclists, stationary vision based observations have been made of the timing of bicyclists' intersection crossing maneuvers. Video recordings were made of bicyclists' crossings and the video images were processed to extract the bicyclists' trajectories. These were synchronized with video images of the traffic signals so that the timing of the bicyclists' maneuvers

could be determined relative to the signal phases. The processed data have yielded cumulative distributions of the crossing speeds of bicyclists who did not have to stop at the intersection and the start-up times and final crossing speeds of the bicyclists who had to cross from a standing start. This study provides a foundation in recommendation of minimal green signal time in terms of safety purpose.

1.3 Contributions of the Dissertation

The primary objective of this dissertation research is to develop methodologies and systems based on both mobile and stationary vision sensors, in order to accommodate the requirement of quantitative measurements in several advanced ITS applications. The dissertation has several major contributions as listed below:

- ◆ A unique OODV System has been designed and developed to observe lane-level activity surrounding a vehicle, as well as the surrounding roadway geometry. This vision system uses a special catadioptric mirror providing a 360 degree orthogonal view of the environment without the need of warping a polar-coordinate based image to a perspective view.
- ◆ The mathematical description of the OODV projection model has been developed, which has provided a foundation for utilizing the unique properties of the OODV system.
- ◆ A systematic approach for developing a mobile traffic surveillance system has been proposed. Both OODV and rectilinear vision sensors, as well as GPS devices (providing relative and absolute position information) have been

integrated to generate a surrounding traffic density map.

- ◆ A novel method using a time-space diagram approach has been proposed to relate the localized traffic density obtained from the mobile traffic surveillance system to the aggregated traffic flow obtained from an embedded loop sensor system.
- ◆ A stationary-vision based technique has been developed as part of a Portable Loop Fault Detection Tool (PLFDT). Vision sensor is utilized as an independent source to provide baseline data to compare against the loop detection system output.
- ◆ A stationary vision-based intersection monitoring system has been developed for the quantitative study of bicycle crossings at signalized intersections. The detailed, accurate measurements of bicycle crossing times represent an unprecedented quantification of bicycle behavior and bicyclists' needs, providing the foundation for defining signal timing.

1.4 Organization of the Dissertation

The Dissertation is organized as follows: Chapter 2 reviews background and related work, including: embedded loop sensor network systems, vision-based vehicle detection algorithms and vehicle tracking algorithms, loop fault detection algorithms, as well as the studies of bicyclist intersection crossing maneuvers. In Chapter 3, we focus on the vision system and object detection algorithms developed for surveillance and safety purposes. First the OODV system is introduced, followed by the mathematical derivation of the

OODV projection model. Then the OODV based vehicle detection, tracking algorithm are presented. Lastly, a rectilinear vision based vehicle detection and tracking algorithm are introduced as well. Chapter 4 describes the developed mobile traffic surveillance system, the localized density and speed estimation, as well as the integration with embedded loop sensor traffic flow data. Some of the analytical results of the data fusion process are also given. In Chapter 5, both hardware and video processing software of the loop fault detection toolkits are introduced. The comparison results from video and physical loop signal are also presented. Chapter 6 describes the studies of the bicyclists' intersection crossing maneuvers from video analysis. We conclude the dissertation and describe potential future work in Chapter 7.

2. Background and Literature Review

2.1 Roadway Traffic Surveillance Systems

2.1.1 Three Levels of Traffic Data Representation

Roadway traffic data can be broken into three scales:

1). *Macroscale:*

The macroscale level generally considers a large roadway network as consisting of links (roadways) and nodes (e.g., intersections). Link-level traffic data such as average speed, flow, occupation are estimated by the infrastructure based roadway sensors (such as embedded loop sensors). Macroscale traffic data [1] provides important cue for path planning, which is usually based on some optimality such as shortest distance or shortest traverse time and least energy consumption.

2). *Microscale:*

The microscale level typically considers things at the vehicle level. A variety of techniques [2] are proposed to locate, detect, and track a target vehicle from different sensors. Actually, microscale traffic data can be combined and processed to form the macroscale traffic data.

3). *Mesoscale:*

The mesoscale level is a level in-between the micro- and macro-scales. A particular freeway may have a variety of features: multiple lanes, off-ramps, etc. In the driving assistance problem, these mesoscale traffic data greatly help vehicle action such as

passing other vehicles, pulling off to the side of the roadway, moving out of the way of emergency vehicles, getting in and out of specialty lanes (e.g., high occupancy toll lanes), and being in the correct lane to exit.

A typical traffic surveillance system usually consists of two components: a sensor based traffic data collection system, and a raw data process and evaluation system. An example is the statewide inductive loop sensor system embedded on the California freeways, and the Freeway Performance Measurement System (PeMS), a systematic processing and analysis tool which is based on the raw data from the loop sensor network. In the following sections, we present a brief introduction of these systems.

2.1.2 Embedded Loop Data Collection Infrastructure

A loop sensor detector [3] is an inductive coil of wires buried in the freeway. Whenever a vehicle passes over the loop, the inductance changes and the detector can sense this. The length of time that the inductance changes, or the time that the detector is on, is a function of the speed and length of the vehicle. The detector can count the total number of vehicles which have passed in a unit of time as well as the average time which the detector is on. These are referred to as the flow and the occupancy. A loop sensor detector provides the direct measurements of flow and occupancy every 30 seconds. Usually each lane has an individual loop sensor detector and several loop sensor detectors distributed at different lanes at same position along the road are logically collected by a Vehicle Detection Station (VDS), which measures the roadway traffic in a single direction.

Figure 2-1 below gives a very simple overview of the general scheme to collect loop data starts at the individual detectors.

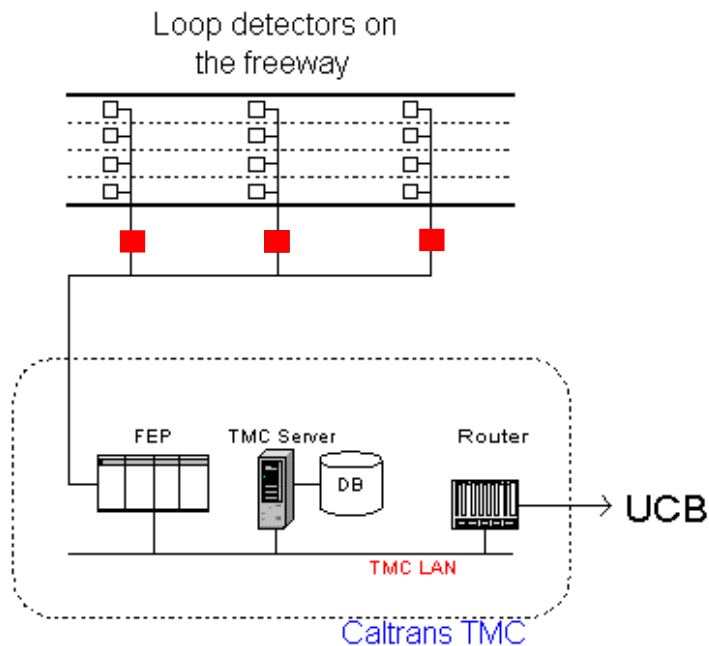


Figure 2-1 Loop data collection process [3]

The steps below [3] describe the stages that the data go through when being passed through a typical collection infrastructure.

- ◆ The loop detectors are connected to a cabinet on the side of the road and the cabinet is connected to the district office via some communications equipment.
- ◆ The data are collected in the district from the loop detectors by a device called a Front End Processor (FEP). The FEP is just a collection mechanism for the data (corresponding to a physical machine).

- ◆ The TMC Server that is in charge of the data collection gets the data from the FEP, parses it, filters it, and decides whether the data is valid or not. In addition, the processed data is written into a database in the District TMC.
- ◆ The processed data at the District TMC are also sent via the network to the centralized PeMS database.

2.1.3 PeMS System

The primary function of the PeMS system [3] is to process detector data. The PeMS system receives measurements (flow and occupancy) every 30 seconds. Figure 2-2 illustrates an overview of the data process of PeMS system starting from the smallest details, the raw data, and moving towards the higher layers. The green boxes represent data that is either collected or stored. The pink boxes represent offline analysis and various factors and parameters. The gray boxes are the real-time grinding steps.

The details of each step of the entire process are discussed as below [3]:

- ◆ In the first step, PeMS system receives the raw, 30-second data from the district TMC servers and stores it in the local database.
- ◆ The 30-second data are aggregated into 5 minute samples. This is done per lane. This data are referred to as the 5-min data with holes. A hole is any sample that is missing. A hole can occur when a detector doesn't report data.

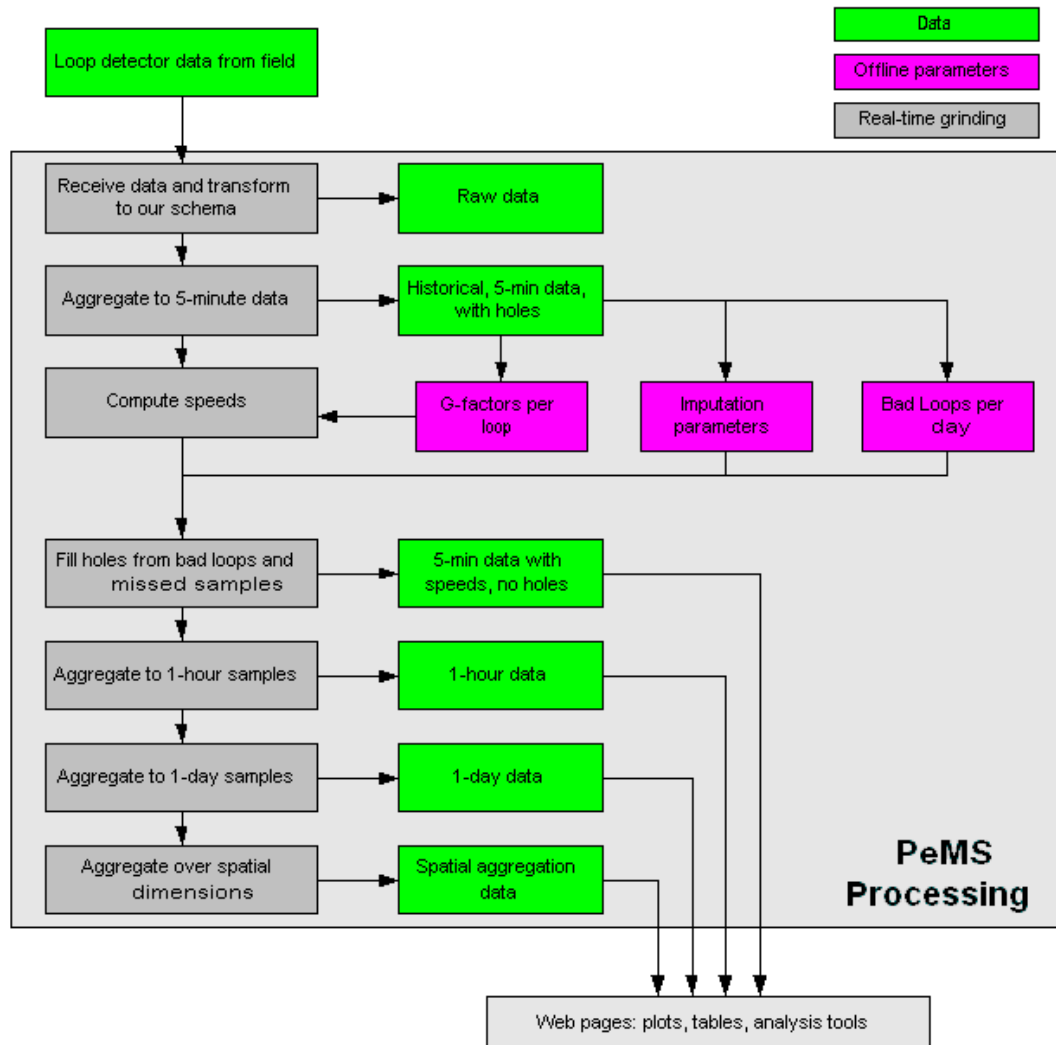


Figure 2-2 Overview of the data process of PeMS system [3]

- ◆ A number of parameters are calculated offline from this 5-minute data set. For example, g-factors used to compute speed.
- ◆ The next step is to fill in all of the holes in the data. After this step, a complete set of 5-minute data with all of the holes filled in is obtained.

- ◆ The 5-minute lane-by-lane data are then aggregated across lanes to form a logical station for the location (e.g., VDS).
- ◆ The 5-minute aggregate data values are then aggregated to hourly aggregate data samples. The hour aggregate data samples are then aggregated to daily data samples. The daily loop data samples are then aggregated across spatial dimensions.

2.2 Review of Vision Based Vehicle Detection Algorithms

Vision-base vehicle detection is actually very challenging due to significant variance within class variables. For example, vehicles may vary in shape, size, and color. Vehicle appearance depends on its pose and is affected by nearby objects. Illumination changes, complex outdoor environments (e.g. shadows), unpredictable interactions between traffic participants, and cluttered background are difficult to control. With the purpose of monitoring the traffic, vehicle detection algorithms need to process the frames real-time or at least close to real-time.

These algorithms and techniques are often based on:

2.2.1 Knowledge Based Methods

Knowledge-based methods employ a priori knowledge to hypothesize vehicle locations in an image. The useful information can be used includes symmetry [7], color [8], shadow [9], corners [10], horizontal/vertical edges [11], texture [12] and vehicle lights [13].

2.2.2 Motion Based Methods

The Knowledge-based methods discussed above use spatial features to distinguish between vehicles and background. Another important cue that can be used is the relative motion obtained via the calculation of optical flow. Approaching vehicles at an opposite direction produce a diverging flow, which can be quantitatively distinguished from the flow caused by the car ego-motion [14]. On the other hand, departing or passing vehicles produce a converging flow. In contrast to dense optical flow, “sparse optical flow” utilizes image features, such as corners [15], local minima and maxima [16], or “Color Blob” [17]. Although it can only produce a sparse flow, feature based method can provide sufficient information for vehicle detection

2.2.3 Some Other Algorithms

(1). Template-based methods apply predefined patterns of the vehicle class and perform correlation between the image and the template. A rather loose template was also used in [18], where the hypothesis was generated on the basis of road position and perspective constraints. The template contained a priori knowledge about vehicles: “a vehicle is characterized by a rectangular bounding box which satisfies specific aspect ratio constraints”.

(2). Statistic model-based methods learn the characteristics of the vehicle class from a set of training images which capture the variability in vehicle appearance. Usually, the variability of the non-vehicle class (background) is also modeled to improve performance [19].

(3). Neural network classifier methods; in [20], Principal Component Analysis (PCA) was used for feature extraction and Neural Networks (NNs) for classification. All the vehicle candidates were scaled to 20x20, then this 20x20 scaled image was divided into 25 4x4 small windows. PCA was applied on every sub window and the output of the “local PCA” was provided to a NN to verify the hypothesis.

(4). Background subtraction methods [21][22][23][24], are a much more common technique for vehicle detection. Most of the background subtraction methods are used with stationary cameras which rely on a fixed background or prior knowledge of the type of action taking place. The main goal is to subtract the road background from the current image, thereby easily detecting any moving object. One of the key problems is how to extract the background adaptively under different lighting conditions. For a stationary camera, since the background is relatively static, moving objects can be segmented from the background just by frame differencing followed by some type of thresholding. Simple thresholding can result in partial extraction of moving objects. If the threshold is too low, camera noise and shadows will produce spurious objects; whereas, if the threshold is too high, some portions of the objects in the scene will fail to be separated from the background. Hence, morphological operations are used to reconstruct incomplete targets and to remove extraneous noise. If the video streams are acquired by a camera mounted on a moving platform, then they have to utilize a stabilization algorithm in order to compensate the camera motion. Thresholding and morphological operations are carried out on the stabilized frames to detect the motion regions.

2.3 Review of Vision Based Vehicle Tracking Algorithms

The trajectory of a moving vehicle is also very important because the vehicle trajectories are collected over a length of roadway and segmentation of time, rather than at a single point and a fixed time. The additional information from the vehicle trajectories could lead to improved incident detection, both by detecting stopped vehicles within the camera's field of view and by identifying lane change maneuvers or acceleration/deceleration patterns that are indicative of incidents beyond the camera's field of view. The vehicle tracking system can also produce individual vehicle data (e.g., spacing, velocity, acceleration), which could lead to better traffic flow modeling and an improved understanding of driver behavior. Another important reason to do vehicle tracking is to bypass the relatively time consuming operation of vehicle detection on a certain image frames because a traffic surveillance and management system requires to process the traffic video in real-time.

Thus, after segmenting the traffic image frame and localizing the vehicle, the segmented vehicle with their bounding boxes and centroids are expected to be tracked in the next several frames.

The evaluation of any vehicle tracking algorithm depends on the degree of those features representing the target vehicle can be extracted from the surroundings. The most common features used in the object tracking algorithm are color and shape. On the other hand, the distortions of features caused by the projection model of vision sensor also need to be considered.

These techniques can be grouped into the following categories:

2.3.1 Region Based Tracking Algorithms

In the approach described in [57][58], the vehicles are detected and represented as a bounding box or a “blob” and then are tracked over a sequence of frame using different criterion. Typically, this method initialized by the background subtraction or inter-frame subtraction algorithm to extract the foreground objects.

2.3.2 Contour Based Tracking Algorithms

The basic idea of this approach [59] is to have a representation of the active contour of the target vehicle. This approach is based on some active contour models and will have less computational complexity comparing with the region-base methods.

2.3.3 Feature Based Tracking Algorithms

In above sections, the vehicle is represented and tracked as a whole while in this approach, several features such as color statistic, distinguishable corners, lines, etc. are chosen to represent a vehicle. In [60], the vehicle tracker builds statistical models for the target vehicle in color and shape feature spaces and continuously evaluates each of the feature spaces by computing the similarity score between the probabilistic distributions of the vehicle and model. Based on the score, the final location of the target is determined by fusing the potential locations found in different feature spaces together.

2.3.4 3D Model Based Tracking Algorithms.

In this approach, the 3D geometric descriptions of vehicle are assumed to be

established. The correspondence between the image features and the 3D model features are also determined. One of the advantages of the model-based methods is that the vehicle pose, which is a critical parameter in vehicle activity analysis, is examined. In [61], an efficient pose refinement methods based on a point-to-line segment distance was proposed to refine the vehicle's pose parameters.

2.4 Review on Loop Fault Detection Systems

A systematic document review on loop fault detection through faulty data at different aggregation level was conducted in [42]. The characteristic of this approach is to apply various statistical analysis methods to the aggregated loop data to figure out possible faults in the loop detection system. Since this approach is indirect, the faults that can be detected are usually large scale problems such as electric or communication system faults. These cannot tell exactly what the fault is, appearing at which point of which loop detection system. This is one of the limitations to the macroscopic and mesoscopic approaches. Besides, communication faults are co-mingled with the loop detection system fault. For the remainder of this review, the low level loop fault detection is the focus of attention.

Many operating agencies use specialized loop testers to assess the quality of the wiring [43][44]. However, these tools bypass the controller and the loop sensors. Therefore, they do not analyze the entire detector circuit, nor do they analyze the circuit in operation. To this end, most operating agencies employ simple heuristics, for example, manually checking the loop sensor indicator light when a vehicle passes over a loop.

Such tests are typically employed when the loop detectors are installed close to the control cabinet. Many practitioners and some researchers [44][45][46] formalized this heuristic by examining if the time series of 30 second average flow and occupancy is within a statistical tolerance.

In [48], Chen and May addressed the fault detection problem for a single loop. They used the number pulses to replace the vehicle counts to verify the loop data. Therefore, if a pulse is broken, it would cause a data problem. They developed an automated loop fault detection system using aggregated data. Their method is to examine the detector *ON/OFF* time instant. Unlike conventional aggregate measures, their approach was sensitive to errors such as "pulse breakups", where a single vehicle registered multiple detections because the sensor output flickered off and back on. This was the main disadvantage to use vehicle count from a single loop for fault detection: one could not isolate other loop fault from the pulse flickering problem.

Low level loop data correction traces back to the Freeway Service Patrol study in 1990s [47][49]. It looked at the transition time measures in sub-second of dual loop stations with the 20ft distance between upstream and downstream loops. It observed some faults in low level data including:

- ◆ missing data;
- ◆ a matching error which results in unreasonable occupancy and speed;
- ◆ on-time and off-time not always being related;
- ◆ no-flow and no-speed but with positive occupancies;
- ◆ existing pulses in both up and down streams;

The author mentioned that some causes of the abnormal phenomenon could be explained as a vehicle's changing lane. However, there was no systematic diagnosis in [48] for those faults, nor systematic methods for data correction.

In [50], Skabardonis et al use dual loop information to detect loop fault. This paper focuses on the evaluation of loop sensors and the detection of cross-talk. It was developed for off-line data analysis but could possibly be used for on-line applications in the future. The operation procedure is summarized in three steps:

- ◆ Record a large number of vehicle actuations during free flow traffic;
- ◆ For each vehicle, match actuations between the upstream and downstream loops in the given lane;
- ◆ Take a difference between the matched upstream and downstream on-times and examine the distribution on a lane-by-lane basis. Assuming that the loops are functioning properly, only a small percentage of the differences should be over 1/30 seconds. Otherwise, the “cross-talk” fault is announced.

The idea of using dual loop speed traps to identify the detector errors is also used in [50]. At free flow, the on-time difference and the off-time difference should be the same if there is no hardware problem. So if they are not the same, there may be a hardware and/or software problem. However, such an assumption is not true for non-free-flow.

Developing a portable tool for systematic and direct loop fault detection and correction from the control cabinet level has not been reported in the literature to our knowledge.

2.5 Review on the Studies of Bicyclist Intersection Crossing Maneuvers

It is not easy to collect quantitative microscopic data about bicycling, and only limited resources have been available to support efforts to collect such data. This has tended to limit the breadth and depth of the information available in the literature.

The most directly relevant prior work is the 1995 review paper by Wachtel et. al., which summarized a variety of preceding studies to point toward signal timing recommendations [52]. The data cited directly in the paper are relatively limited and are described only in terms of mean crossing times and speeds. In the absence of more detailed data it is difficult to know how to apply the findings more widely. The second relevant work is the 2005 TRB paper by Rubins and Handy [53], which summarizes an extensive set of bicyclist crossing observations in Davis, CA. These data, which represent total crossing times at intersections of varying widths, show significant scatter, making it challenging to use them as the basis for specific signal timing recommendations. Regressions showed the mean value of crossing time as a function of crossing width, and the authors derived estimates of fixed start-up time and crossing speed from these ensemble statistics. They focused on 2%ile and 15%ile statistics for formulating their recommendations, but the full distributions of crossing times were not reported, which makes it impossible to apply other percentile criteria based on their data.

3. OODV System and Video Processing Algorithm

As mentioned in Chapter 1, computer vision techniques have become a powerful tool for a variety of ITS applications, such as traffic surveillance systems and other safety applications. In this chapter, we describe a unique Orthogonal Omni-Directional Vision (OODV) system that has been developed and utilized in our mobile traffic surveillance system. A unique motion compensated background subtraction algorithm for vehicle detection in OODV videos is then introduced, followed by a description of a dynamic feature grouping based vehicle tracking algorithm [40] that is applied in the following three chapters.

3.1 The Design of OODV System

3.1.1 The Structure of the OODVs

The OODVs consists of a special catadioptric mirror, a CCD camera and a computer, as shown in Figure 3-1(a) and (b). Digital cameras are used, which use the IEEE-1394 standard to transmit both data and power. The image data are purely digital, eliminating the need for analog frame grabbers typically used for NTSC video capture. The catadioptric mirror and digital camera are mounted together inside an acrylic-cast plastic tube with a black needle coming out at the center of the mirror, pointing along the optical axis of the camera to eliminate any internal reflection in the tube. Image data are stored and processed using a 2.0GHz Core2 Duo, 2.0GB RAM computer with 1TB of disk space. This allows for parallel processing of threads and a robust way to flush the live video sequence onto hard drive for post processing when needed.

The catadioptric mirror allows for a 360 degree view with 124 degree downward-looking field of view. It was designed using ray tracing techniques, allowing that the distance between any 3D points in any plane perpendicular to the camera's optical axis to be a linear function of the corresponding pixel distance in the image. This unique perspective projection model provides a lot of properties similar to the conventional rectilinear cameras using perspective projection. The orthogonal images captured from this OODVs can be used for high-speed (e.g real-time) applications requiring 360 degree view. Figure 3-2 shows the comparison of the images captured from the OODVs and hyperbolic ODV.

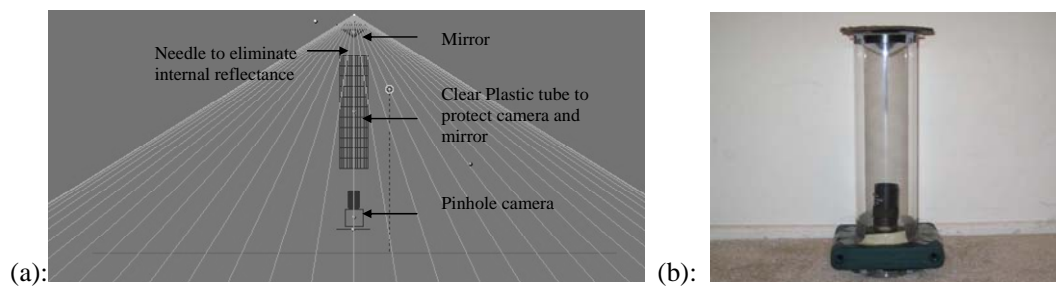


Figure 3-1 (a) Layout of the OODV; (b) OODV used in the project

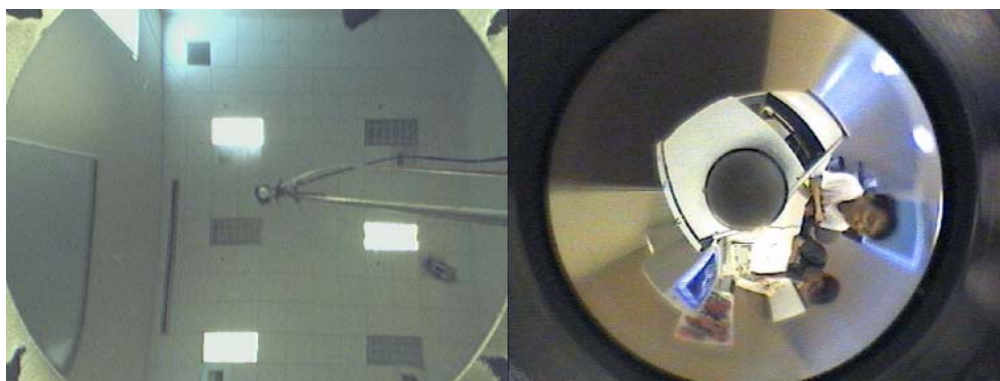


Figure 3-2 Comparison of OODV image with hyperbolic ODV image

3.1.2 Characterization of the OODV Projection Model

In optics, catadioptrics refers to the devices consisting of curved mirrors and lenses. These systems have many advantages compared with conventional rectilinear vision sensors, for instance, they can obtain a much wider angle of view. Ishiguro, et al. pioneered the design of these sensors [32]. In [33], Baker and Nayar derived the complete class of single-lens single-mirror catadioptric sensors, in which some general mirrors were discussed such as spherical, hyperboloid mirrors. All these mirrors preserve the fixed viewpoint constraint. In some other cases, this constraint is broken to satisfy some special projection relationships. Andrew Hicks [34][35] presents two families of mirrors that preserve the geometry of a plane perpendicular to their axis of symmetry, using the pinhole and orthographic models respectively. In this research, the mirror used in the OODVs is similar to those described by Hicks in his work.

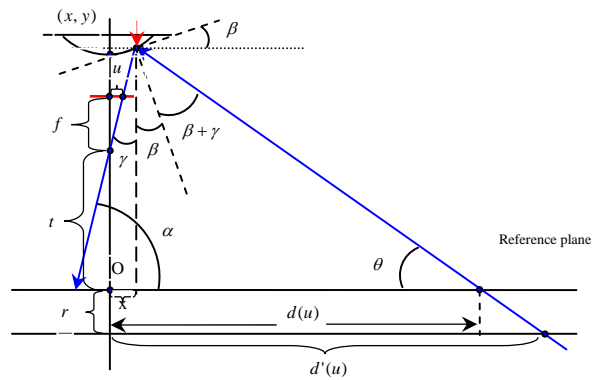


Figure 3-3 Geometry of the OODV

Figure 3-3 illustrates the geometry of the OODV. Using a pin-hole modeling approach, a ray starts from the view point of the camera, goes through one point in the

CCD array, hits the point (x, y) on the mirror, and then intersects the reference plane. We define u as the CCD distance from the point to the center of CCD, and $d(u)$ as the corresponding real world distance between the center of the reference plane and the point of the reflected ray on the mirror. Eqn.3-1 has been developed (see) as the mathematical description for the mirror surface so that u and $d(u)$ satisfy $d(u) = a \times u$, where a is a constant

$$\frac{2y'}{1-y'^2} = \frac{-2xy^2 + (afx + 3xt)y - (aft + t^2)x}{y^3 - 2ty^2 + (t^2 - x^2)y + (af + t)x^2} \quad (\text{Eqn.3-1})$$

Eqn.3-1 mathematically describes the surface shape of the mirror. Based on this initial model, the surface of the mirror $y(x)$, the parameter t , the camera focal length f , and the designed scale parameter a are related. Thus the image view can be designed based on the parameters in Eqn.3-1. However for such a differential equation; there is no closed-form solution. Eqn.3-2 shows the numerical version of the Eqn.3-1 (see Appendix A for the derivation), which can be solved and simulated.

$$\Delta y = \frac{-2\Delta x + \sqrt{4(\Delta x)^2 + 4g^2(\Delta x)^2}}{2g} = \frac{(\sqrt{1+g^2}-1)\Delta x}{g} \quad (\text{Eqn.3-2})$$

Figure 3-4 and Figure 3-5 shows the simulation of the 3D display and cross-section of the OODV mirror shape, created through the numerical method described in Eqn.3-2. The initial conditions are: $x = 0, y(0) = 34, a = 54, f = 6, t = 15$. The step $\Delta x = 0.001$. Figure 3-6 illustrates two images created using the ray tracing rendering technique based on the OODV system, whose mirror is created by the numerical method as shown in Figure 3-4 and Figure 3-5. The physical model of the checker board is placed on a perpendicular plane to the optics axis in the OODV coordinate system. This rendering process shows the OODV generates the desirable imaging properties, hence, proves

Eqn.3-1 & Eqn.3-2 mathematically describe the mirror shape well.

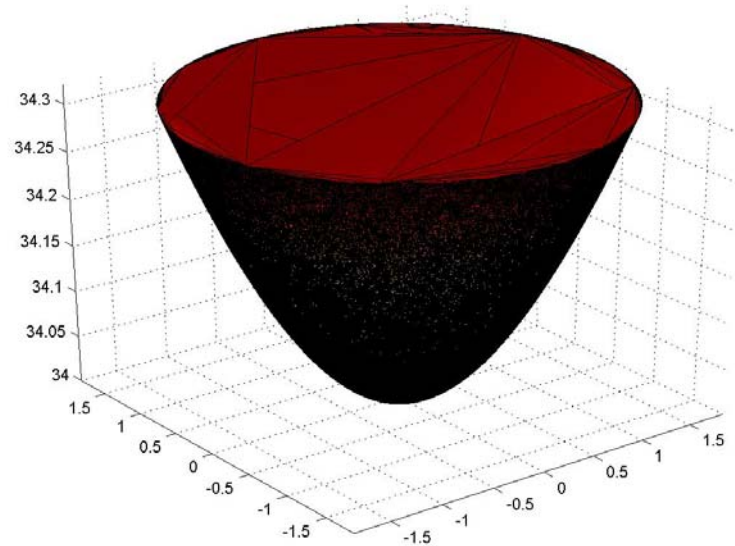


Figure 3-4 3-D Reconstruction of the OODV mirror shape

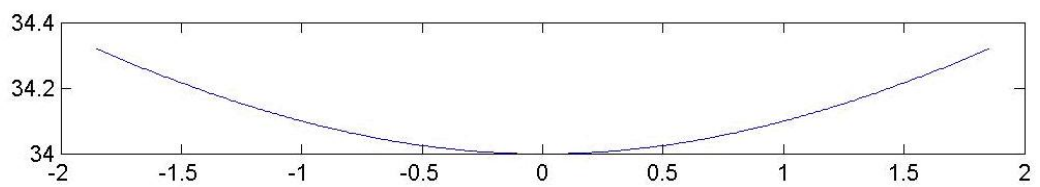


Figure 3-5 Cross-section simulation of the OODV mirror shape

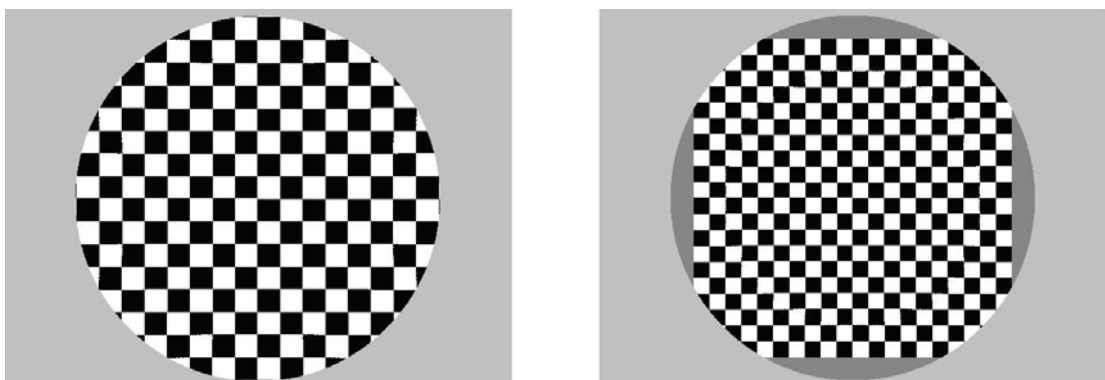


Figure 3-6 Rendered images of checker board using the OODV projection model

3.1.3 Some Discussion on the OODV Projection Model

For some planes other than the reference plane, as in Figure 3-3, we consider a plane with a different depth $y(0) + r$. $d(u)$ in the reference plane and $d'(u)$ in the plane $(y(0) + r)$ correspond to the same line u in the CCD array. From the geometrical relationship, we have

$$\begin{aligned} d'(u) &= \left(a + \frac{r}{y(x)} \cdot a - \frac{x \cdot r}{y(x) \cdot u} \right) \cdot u = \frac{y(x) + r}{y(x)} \cdot a \cdot u - \frac{x \cdot r}{y(x)} \\ &= \frac{y(0) + r}{y(0)} \cdot a \cdot u + \left[\frac{y(0) - y(x)}{y(0) \cdot y(x)} \cdot r \cdot a \cdot u - \frac{x \cdot r}{y(x)} \right] \end{aligned} \quad (\text{Eqn.3-3})$$

In Eqn.3-3, we can see that $d'(u)$ is the summation of the approximate perspective part and the distortion part. From the expression, it makes sense that:

- 1) when u, x, y are fixed, the depth Z increases (equal to r increases), the distortion also increases;
- 2) if depth Z is fixed, the increasing u makes x, y increase, which finally results in increased distortion.

This can also be verified by the simulated distortion map shown in Figure 3-7(a) & (b). The distortion of each pixel in the rendered image is created based on the distortion part in Eqn.3-3. All the parameters settings remain the same except the depth of imaging plane. As we see in Figure 3-7, the distortion is distributed like a special shape, getting rid of a conic from a column, with small distortion in the middle and larger one at the edge. Since the mirror is 360 degree symmetric, the distortion is determined by the distance D between the pixel and the image center, rather than the position of the pixel. In addition, the distortions in (b) are generally larger than those in (a) because of a larger

depth Z .

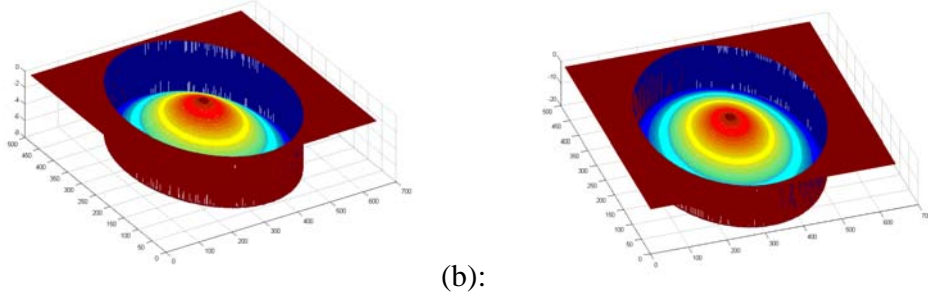


Figure 3-7 Graphical display of the imaging distortion map (a and b have different depth)

For a particular depth Z , the distortion can be modeled only by several polynomial coefficients. Using 2nd order polynomial curve fitting as an example, the distortion can be represented as $Distortion = a'D^2 + b'D + c'$, where D is the distance between the pixel and the center of the image. Usually we consider D as a function of CCD distance u : $D = D(u)$. This is less complicated comparing with the theoretical distortion part in Eqn.3-3 and the coefficients a', b', c' can be found easily by curve fitting techniques. By exploring more than 100 different planes paralleling to the reference plane, the coefficients a', b', c' can be determined by a linear relationship with depth Z .

Thus the OODV projection relationship can be modeled as a sum of a perspective projection and a polynomial residue:

$$d(u) = \frac{Z}{A} \cdot u + [a' \cdot D(u)^2 + b' \cdot D(u) + c'] \quad (\text{Eqn.3-4})$$

where $a' = w \cdot Z + v$, $b' = p \cdot Z + q$, $c' = m \cdot Z + n$,

w, v, p, q, m, n are constants and can be obtained from the line fitting. Z is the depth and $D(u)$ is the pixel distance. u is the corresponding CCD distance and A is a constant.

3.1.4 Highlights of the OODV Projection Model

The Eqn.3-3 can be simplified as below:

$$\begin{aligned}d'(u) &= \frac{y(0)+r}{y(0)} \cdot a \cdot u + \left[\frac{y(0)-y(x)}{y(0) \cdot y(x)} \cdot r \cdot a \cdot u - \frac{x \cdot r}{y(x)} \right] \\ &\approx \frac{y(0)+r}{y(0)} \cdot d(u) = \frac{depth}{y(0)} \cdot d(u)\end{aligned}\quad (\text{Eqn.3.5})$$

Eqn.3-5 shows the mirror makes the OODV approximate a perspective projection with the view point at $A(0, y(0))$, even though it is not a true perspective projection.

3.2 OODV-Based Vehicle Detection Algorithm

The wide field of view (FOV) of the OODV, as well as its orthogonal projection property, makes it suitable for observing not only surrounding vehicles and other objects, but also to observe roadway infrastructure as the OODV equipped vehicle travels down the roadway. In this dissertation, we have carried out a variety of experiments with this type of sensor. For example, a truck-trailer has been mounted with four of these unique OODVs on the top of each corner, as shown in Figure 3-8 (a), driving on the freeway to collect the traffic data. Figure 3-8 (b) illustrates the recorded roadway traffic video data, and the 3D display about the field of view is given in Figure 3-8 (c), which is generated by 3D rendering.

To accommodate this type of downward-looking OODV videos, a motion compensated background subtraction based vehicle detection algorithm has been proposed. Figure 3-9 illustrates the flow chart of this method. The basic idea is making use of the unique geometric projection model of the OODV to compensate for the spatial

motion of the moving platform. In addition, luminance compensation is applied by computing the mean of brightness in the background region to make the background stable in consecutive frames.

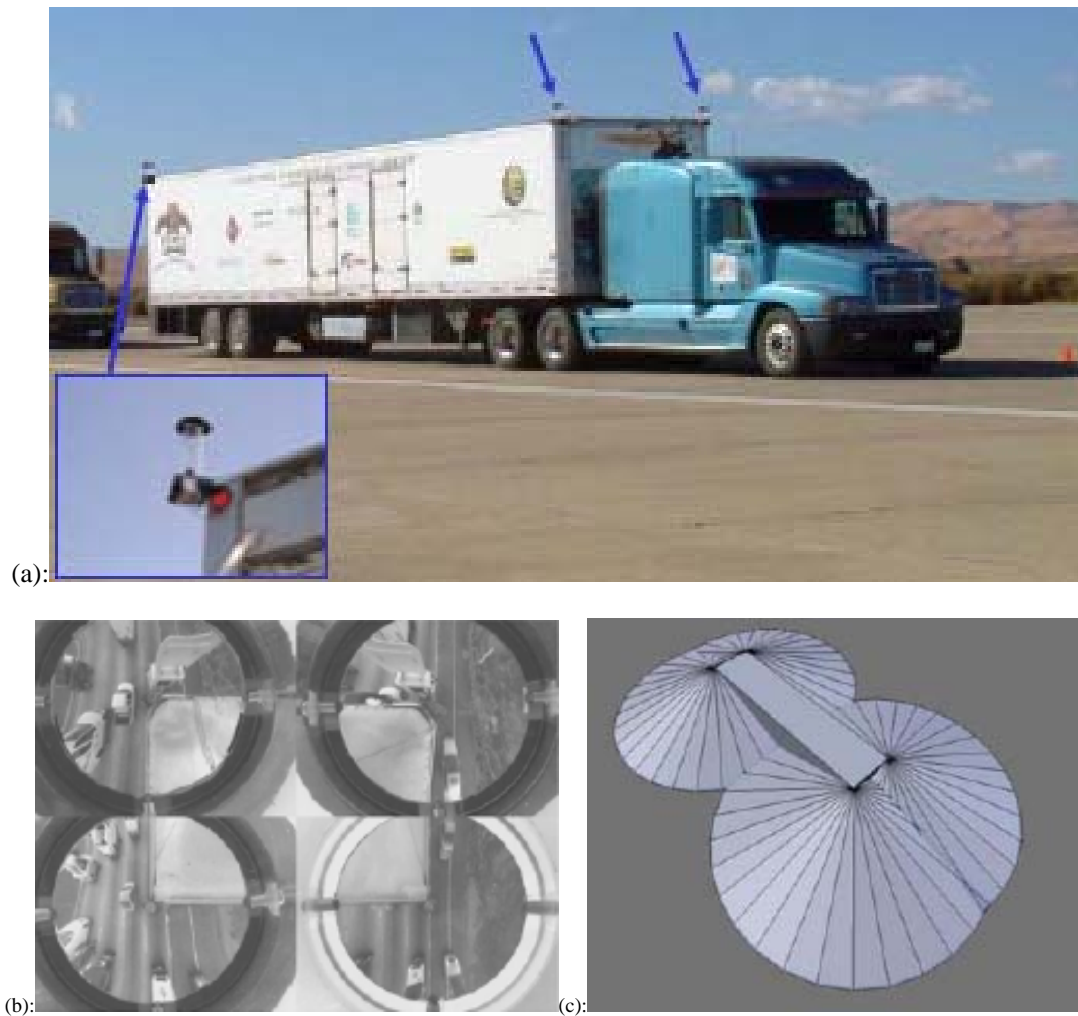


Figure 3-8 OODV based roadway experiment setup

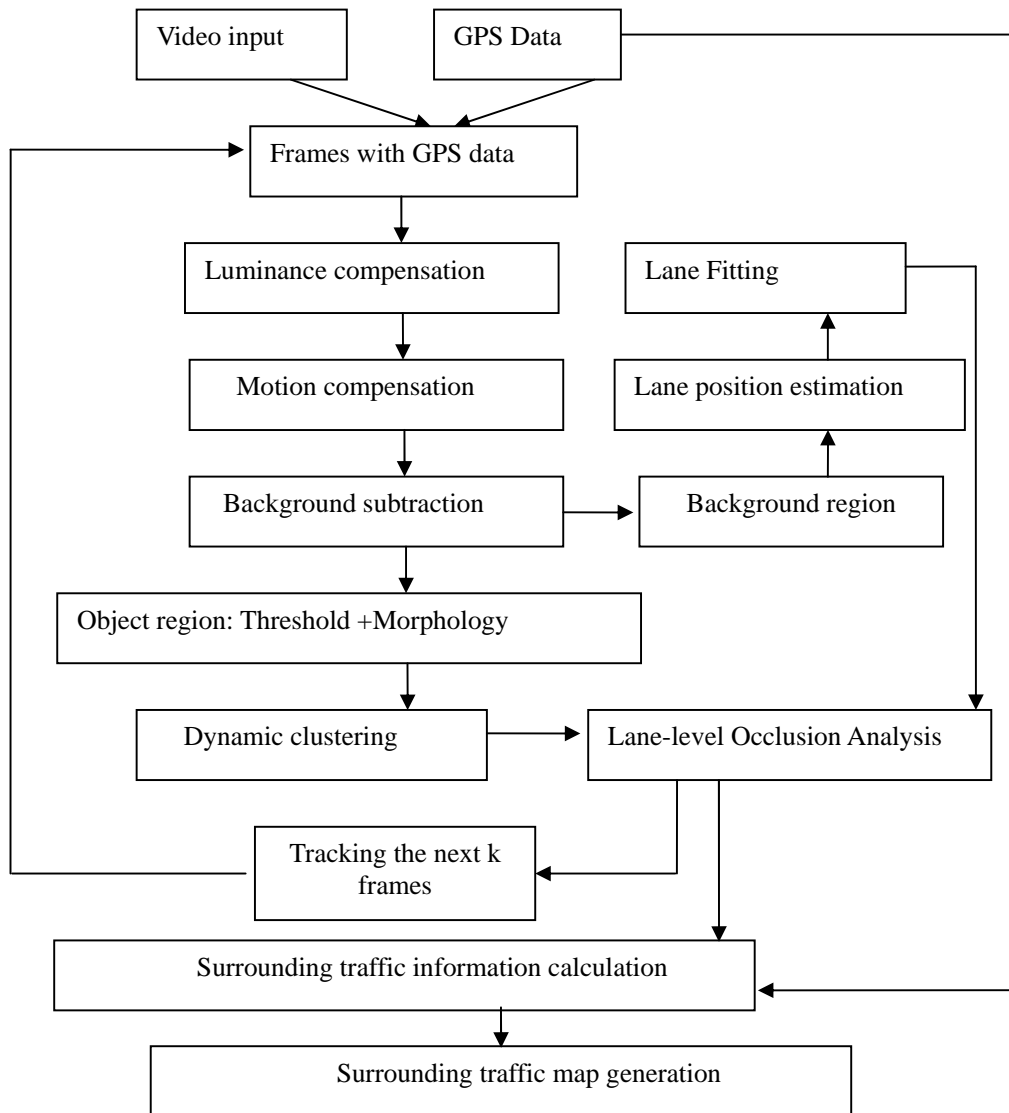


Figure 3-9 Flow chart of the OODV based vehicle detection algorithm

To stabilize the ground plane background, a ground plane motion compensation algorithm was proposed in [39], but this method is applied to the applications where the conventional rectilinear camera is mounted in the front or rear of the car such that the points on the ground plane have different scene depth. However in our application, the optical axis of the OODV is perpendicular to the ground and the view angle is

downward-looking. These properties ensure that all the points on the ground plane have the same depth. According to the special geometric projection relationship of the OODV mentioned, the spatial motion compensation of the frame becomes the pixel shifting in the frame which is much easier to be implemented. The shift distance in the background can be obtained directly from the vehicle moving distance in the real world using Eqn.3-6:

$$PIX_dist = \frac{GND_dist}{AZ} \quad (\text{Eqn.3-6})$$

where A is the projection parameter for the OODV, which has been obtained in the calibration process. GND_Dist is the driving distance of the truck in a certain time-slot. It can be computed directly from the GPS data as in Eqn.3-7.

$$GND_Dist = (t_{curr} - t_{prev}) \times V_{GPS} \quad (\text{Eqn.3-7})$$

The sudden change of brightness has been considered while doing gray-scale compensation. Let cur_u be the mean brightness of the background region in the current frame and $prev_u$ be the mean brightness of the background region in the previous frame. Then the brightness change in the region without moving objects equals to $cur_u - prev_u$. Then after the spatial motion compensation and luminance compensation, we have a temp frame as given in Eqn.3-8.

$$I_{curr}'(i, j) = I_{prev}(i + PIX_dist, j) + (cur_u - prev_u) \quad (\text{Eqn.3-8})$$

The background in $I_{curr}(i, j)$ and $I_{curr}'(i, j)$ is supposed to be the same and the difference image can be obtained using Eqn.3-9. The image is then segmented into the background region where the gray values in the difference image are less than a threshold, as shown

in Figure 3-10(a), and the foreground region where the gray values are large than the threshold, as shown in Figure 3-10(b).

$$Diff(i, j) = abs(I_{curr}(i, j) - I_{curr}'(i, j)) \quad (\text{Eqn.3-9})$$

A simple threshold operation on the difference image can result in partial extraction of surrounding vehicles. If the threshold is too low, noise and shadows could be detected as vehicles; whereas, if the threshold is too high, some portions of the vehicles will be classified as the background. Hence, morphological operations are used to reconstruct incomplete objects and to remove extraneous noise.

To detect objects from the foreground image, the classic K-mean clustering algorithm has been modified so that it can dynamically determine the number of clusters to be grouped. The grouping criterion is the pixel distance from each foreground point to the centroid of each cluster.

Vehicle occlusions analysis is another important issue in the surveillance system. Occlusions will result in the failure of vehicle detection and further degrade the accuracy of vehicle tracking, classification, and counting. Generally, shadows and optical projection effects are the major factors to cause vehicle occlusion. These vehicle occlusions can be simply grouped into three classes: 1) One vehicle is classified as two different vehicles; 2) Two or more vehicles are classified as one vehicle; and 3) The vehicle is missing.

In this research, lane structure is used to eliminate the occlusions. First, each detected vehicle is marked with the lane index. If the centroids of two detected vehicle are close enough in the same lane, then we can consider them as the same vehicle. Figure 3-11

illustrates how the lane information helps to eliminate the occlusions in case 1). In case 2), two vehicles in the consecutive lanes may probably be detected as one vehicle if they are overlapping with each other. Thus if the bounding box of a detected vehicle is larger than two times the lane width, it is reasonable to consider it as a potential occlusion. Further, color information in the HSV color space is used to confirm whether it is an occlusion. Figure 3-12 illustrates such an example of the case 2) occlusions. As shown in Table 3-1, we have processed a 2000-frame sample video which is captured by the instrumented truck going straight down in the middle of a 6-lane freeway without lane changing and turning. Through the lane structure occlusion analysis, the false detection in case 1) and 2) are greatly reduced.

Lane markings provide the most important roadway geometry information for lane-level traffic analysis and navigation tasks. Lane detection is performed based on the following assumption: on the flat ground, the intensity of the lane marker area is much brighter than that in the lane areas.

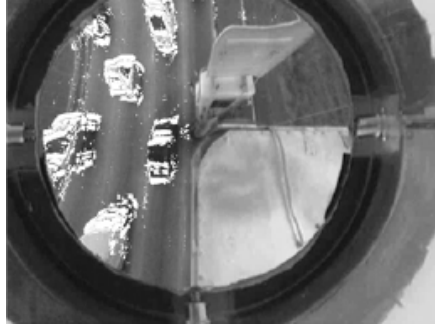
In [38], these bright sections are extracted by comparing the pixel intensities with their neighboring pixels at a distance m , where m represents the width of lane markers. The value of m could be obtained in the calibration process. Only those pixels whose intensities are brighter than both of their left- and right- m neighbor pixel over a threshold are classified as lane markers. This classifying criterion can be represented as:

$$u(i, j) = \begin{cases} d_{m+}(i, j) + d_{m-}(i, j), & d_{m+}(i, j) > I(i, j)/k, d_{m-}(i, j) > I(i, j)/k \\ 0, & \text{otherwise} \end{cases} \quad (\text{Eqn.3-10})$$

where

$$d_{m+}(i, j) = I(i, j) - I(i, j + m)$$

$$d_{m-}(i, j) = I(i, j) - I(i, j - m)$$



(a): Figure 3-10 (a).Background image and (b) Foreground image

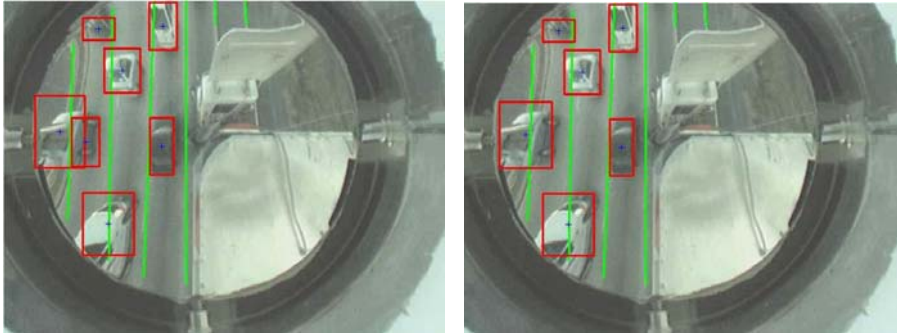


Figure 3-11 Occlusion analysis in case (1)

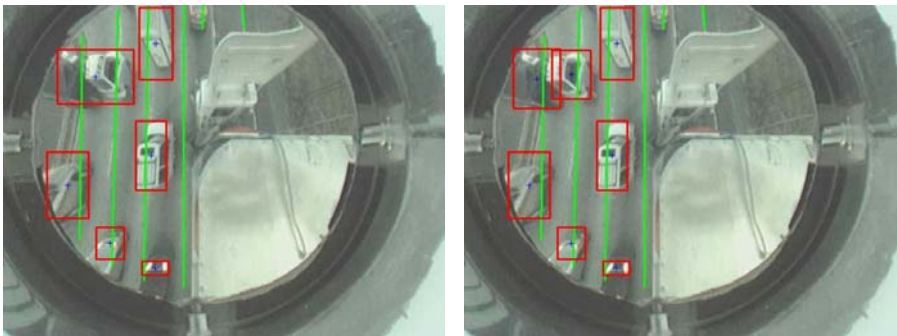


Figure 3-12 Occlusion analysis in case (2)

	Frame	Total Vehicle	Detected Vehicle	False detection (case 1,2)	Missing (Case3)	Detection rate
W/O Lane level Occlusion Analysis	2000	15253	11074	2473	1706	72.6%
W/ Lane level Occlusion Analysis	2000	15253	12749	718	1786	83.7%

Table 3-1 Comparison rate with and without lane-level occlusion analysis

Since sometimes part of the lane markers are occluded by vehicles, the above process is performed in the background image obtained in the previous sub-section. Then, morphological operations are used to enhance the lane markers and to remove the noise. Binarization is performed on the enhanced image:

$$r(i, j) = \begin{cases} 1, & u(i, j) > 0 \\ 0, & otherwise \end{cases} \quad (\text{Eqn.3-11})$$

The lane maker areas are usually space continues in the background image except that part of it is occluded by vehicle. Thus a recursive connected components analysis and labeling process is performed.

In the ideal case, all the pixels in a lane are connected and are labeled by the same number. However, usually, one lane is cut by the existence of the on-road vehicles; thus, centroid information is used to explore the vertical correlation. Finally, polynomial curve fitting is used to extract the sketch of lane information. Figure 3-13 illustrates the entire process of the lane detection; (a) shows the original frame; (b) presents the rough lane marker detection results; (c) indicates the enhanced image after noise removing; (d) illustrates the finally result using polynomial model to fitting the lane curve.

When a vehicle is detected, 96.8% of the vehicles can be classified into the correct lane it is driving in, based on the same 2000-frame video test with occlusion analysis.

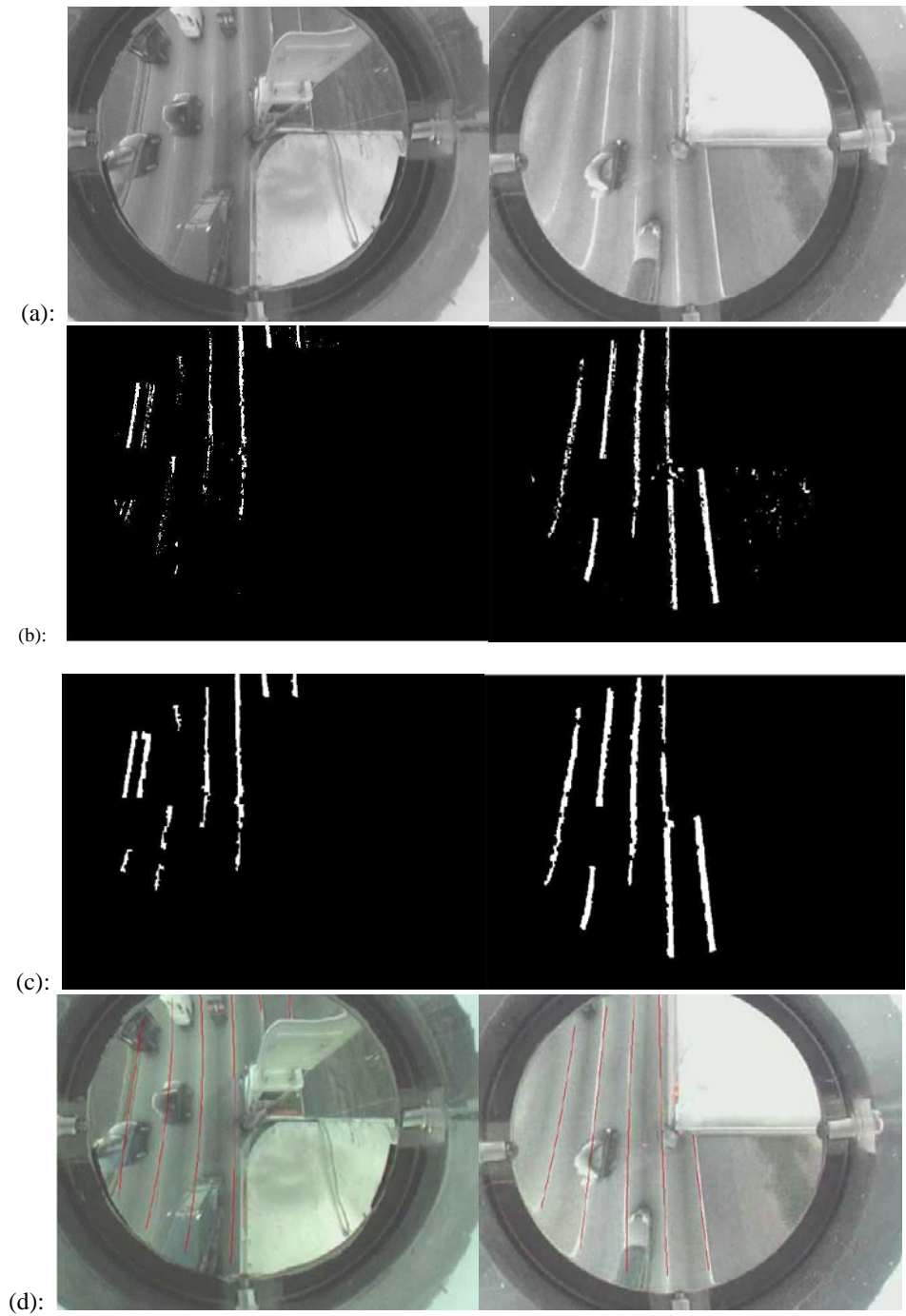


Figure 3-13 Lane structure detection process

3.3 Rectilinear Camera-Based Vehicle Detection and Tracking Algorithm

In recent years, various vehicle detection and tracking algorithms have been introduced based on the rectilinear camera model. The majority of approaches are either based on background subtraction or feature tracking and grouping. Background subtraction methods are straight-forward, but have high variability based on illumination conditions. On the other hand, feature based approach usually obtain feature's trajectories by tracking the detected corner feature and grouping them based on the analysis of the trajectories. However, it is difficult to maintain robust feature trajectories for a long time.

In [40], Kim proposed an object detection and tracking approach that combines the background subtraction algorithm and the feature tracking and grouping algorithm. The highlights of this approach are:

- ◆ background subtraction is used as an additional cue for robust feature detection and tracking, and vice versa;
- ◆ It proposes a robust multi-level feature grouping algorithm in order to handle different sizes of objects, such as vehicles and bicycles in the traffic scene.
- ◆ the dynamic feature grouping algorithm produces a high quality of object trajectory, in compare with the previous approaches that produce trajectories simply by linking the center of the object.

This algorithm has been experimentally proven its robustness under different environment, such as at a congested intersection, freeway, etc. This method is adopted in the video processing part of all the three applications described in the following three

chapters, including virtual loop vehicle detection in the loop fault detection application (see Chapter X), forward vehicle detection in mobile traffic surveillance system (see Chapter Y), as well as high quality trajectory production in the intersection bicycle detection application (see Chapter Z). All of these applications have had reliable results.

A summary of this algorithm is given below:

- ◆ **Background Subtraction:** A temporal median approach is combined with alpha-blending to update the pixel intensities to find the background. To compensate the sudden illumination change, an illumination correction procedure is provided as well. Eqn.3-12 describes the background update equation, where $Ic()$ is an illumination-correction function and N_t is the temporal median of the recent frame:

$$B_{t+1}(i, j) = \begin{cases} Ic(B_t(i, j)) & M_t = 1 \\ Ic((1 - \alpha)B_t(i, j) + \alpha N_t) & M_t = 0 \end{cases} \quad (\text{Eqn.3-12})$$

- ◆ **Feature Tracking and Grouping:** the Harris corner feature are detected in the foreground region and is tracked by a cross-correlation template matching method. Then a multi-level grouping is applied, where low-level corner features are first grouped into small clusters and then clusters are grouped into object. A Bayesian reasoning is applied to determine the corner feature's membership by considering the following parameters: p – the proximity between the feature and cluster; r – history of falling in the same background subtraction blob region; m – motion history; t – trajectory's maximum disparity; c – previous cluster membership. Eqn.3-13 describes the Bayesian reasoning equation:

$$P(\text{member} / p, r, m, t, c) = \frac{P(p, r, m, t, c | \text{member})}{P(\text{member})P(p, r, m, t, c | \text{member}) + P(\text{--member})P(p, r, m, t, c | \text{--member})}$$

(Eqn.3-13)

- ◆ **Trajectory generation:** the position and shape of the object is determined by the positions of all its member corner feature and all trajectories are computed in the 3D space (the image position is projected into the real-world coordinates, assuming a predetermined height).

Figure 3-14 shows a nice example of vehicle detection and trajectory generation results under the presence of shadows, using the methodology [40] described in this section.

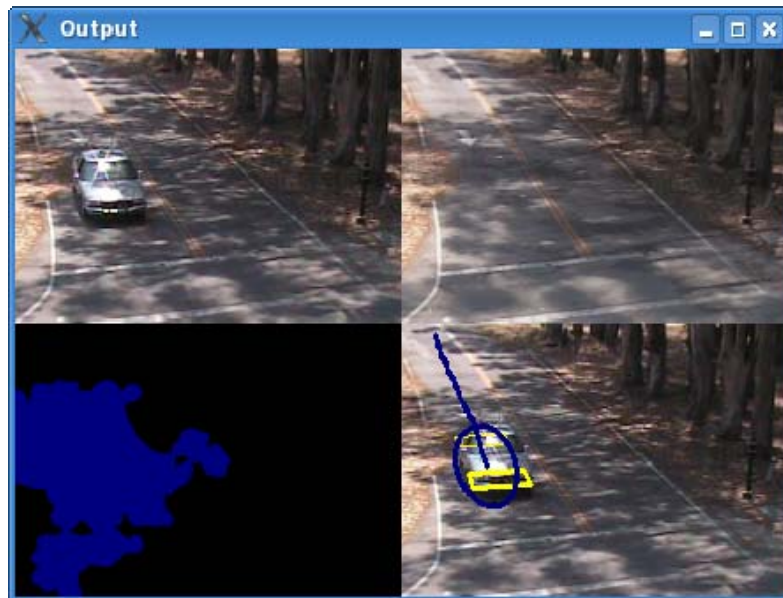


Figure 3-14 An example of vehicle tracking under the presence of shadows

4. Stationary Vision Based Portable Loop Fault Detection Tool (PLFDT) Development¹

4.1 Vision Sensor Provides Baseline Data

As described in Chapter 1, a systematic approach to detecting faulty traffic detector loops is crucial to traffic operation, traveler information, and corridor management. Traffic detection systems are widely used for traffic management and control in California. The statewide loop sensor system consists of over 28,000 sensors located on the mainline and ramps, and grouped into 8,000 vehicle detector stations (VDS). However, loop data are not reliable. The loop data delivered to a TMC may contain error created at a point or several points between the loop detector and the TMC database, which presents a great challenge. To solve this problem, it is necessary to take a systematic approach. This approach is composed of three complementary tasks:

- ◆ loop fault detection;
- ◆ faulty loop data correction/imputation; and
- ◆ loop detection system maintenance.

Loop fault detection at the high level is done usually through an analysis of aggregated loop data. Such approach is indirect, and shortcomings are obvious: (a) data aggregation in time and space would smear the faulty problem; and (b) a communication fault caused by data error/loss make it impossible to isolate the loop fault detection problem. Only the detection at the control cabinet level can directly detect the fault(s) in

¹ This work was conducted at California PATH as a joint research with Lu, Kim, Varaiya, Horowitz and Palen, and described in the Proc. 88th TRB Annual Meeting, 2009 [63]

hardware and software, isolating them from the communication fault, and correct the faults permanently.

As listed in [42], the loop faults to be diagnosed at the microscopic level include hardware, software and installation problems, and loop card faults. Those faults appear as: mis-assignment, temporary data missing, crosstalk, absence of data or constant for a period of time, broken cable, chattering, broken card, card sensitivity being too high or too low, broken pulse, mismatch of ON/OFF time instances between upstream and downstream loops for dual loop stations.

This is the only level that one could conduct direct loop fault detection and isolate the loop faults from possible other system faults. Data in this level can either be the ones processed by loop detector card, which will be loop ON/OFF time instances or occupancies; or the raw loop pulse signal before the loop card. The main characteristics of those data are that (a) they do not pass any communication media and thus there is no possibility of communication fault which usually pollutes or loses the data stream; (b) all the raw information is available with a proper interface with the control cabinet; (c) real-time data are available; and (d) most importantly, an independent data source or baseline data could be obtained at this level thus loop fault detection could be conducted by comparing the loop detector reading with these data. Obviously, vision sensor is the best choice to provide the baseline data, due to its easy installation, operation and maintenance, high flexibility, low costs and their ability to monitor wide areas.

4.2 Development of Portable Loop Fault Detection Tool (PLFDT)

The PLFDT, as shown in Figure 4-1, is designed for systematic loop fault detection at the control cabinet level. A loop detector(s) could be identified as being *suspicious* from a higher level data analysis in TMC/PeMS. The suspicious loops will then be diagnosed further using the portable tool. The tool will enable the operator to use independent video stream of traffic measurements for comparing with the suspicious loop detector data. This portable tool will be designed to achieve the following objectives:

- ◆ determination of the exact fault type and causes in the detection system
- ◆ on-site diagnosis of faults on:
 - mis-assignment
 - malfunctioning such as misfiring
 - inappropriate card sensitivity settings
 - inductance variation due to temperature and humidity
 - broken loop circuit due to
 - ◆ improper installation
 - ◆ road surface maintenance
 - ◆ fatigue
- ◆ Facilitate on-site detector precision evaluation and calibration.

4.2.1 Overall System Structure of the PLFDT

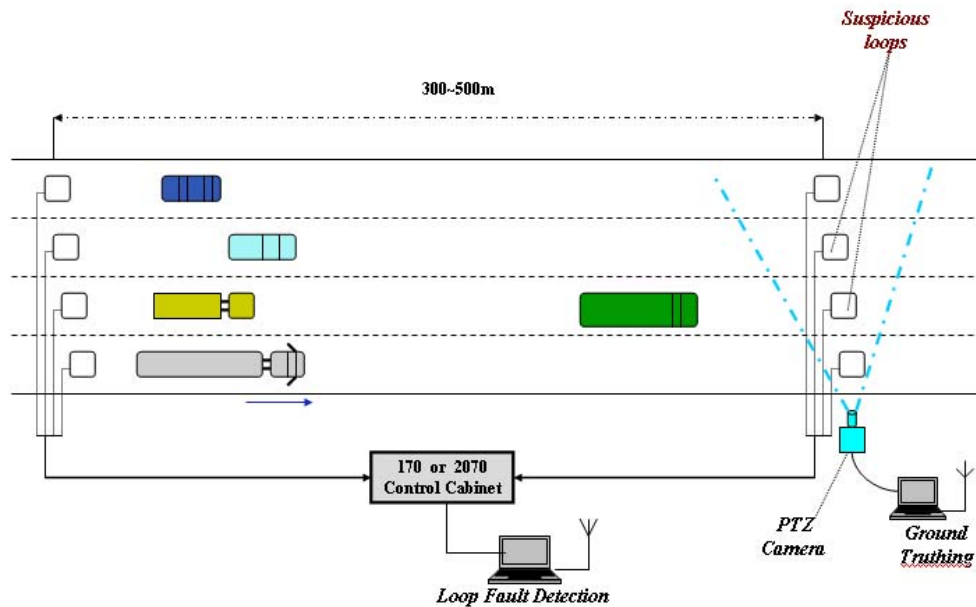


Figure 4-1 Overall system structure

The system development includes hardware, software and algorithm development.

Our hardware setup consists of the following components:

- ◆ Mobile trailer which can be towed to the site near the suspicious loop location
- ◆ Retractable pole with a PTZ (Pan-Tilt-Zoom) camera mounted on it
- ◆ Two laptop computers with the Linux operating system
- ◆ *Computer A* to capture video images, process them in realtime and send out the result via IEEE 802.11b wireless
- ◆ *Computer B* to interface with a loop detector card, receive the video processing data from Computer A through IEEE 802.11b wireless, and to compare the synchronized signals for loop fault detection;

4.2.2 Mobile Pole for Roadside Video Camera Mounting



Figure 4-2 Mobile retractable pole setup

A mobile pole for the roadside camera setup has been developed as in Figure 4-2. The mobile trailer has four retractable folding legs for supporting the platform for leveling and robustness. The Pan-Tilt-Zoom camera is mounted on top of the mobile pole and looks at the loop for vehicle tracking to obtain ground truth. The PTZ parameters can be controlled using the remote controller or using control software running under Microsoft Window System through the RS-232 serial port interface. This setup process is necessary for the camera to view the loops on the ground and to display on the computer

screen so that a virtual loop can be overlaid on the actual loop.

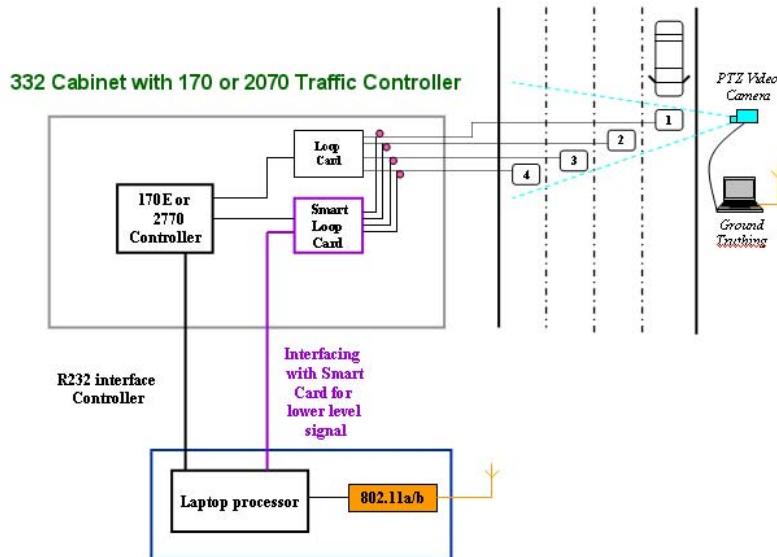


Figure 4-3 Interfacing with control cabinet and smart card

4.2.3 Interface with Control Cabinet

A loop card receives raw analog signals from each loop circuit, processes them with a physical oscillator and amplifier, and outputs traffic signals. Loop cards can be divided into two types: single-layer and multi-layer output cards. Single-layer output cards have only two outputs (the vehicle count and occupancy), that results from processing the input signal from the loop circuit. For example, Sarasota GP5 and Reno 222 cards are single-layer output ones that are widely used. There is no direct interface port with those cards. Instead, their signals are directly fed into the controller. The output from the card to the controller is either 1 or 0 without the lower level signal available. The low level signal is more attractive than the binary data for several reasons: (a) it can be used for extracting vehicle signatures for re-identification; (b) it tells if the sensitivity of the card

is properly adjusted; (c) it tells if any algorithm in the card has a flaw; and, (d) most importantly, we can remove time delay incurred in the traffic controller. Multi-layer output or *Smart Cards*, such as the 3M Canoga and IST cards, have multi-level output information including the start detection times, the occupancy, the vehicle count, fault status, and even the inductance intensity signals calculated from the frequency. A smart card has a built-in RS232 interface port, and thus lower level signal can be obtained. We chose a smart card, 3M Canoga C922, which is compatible with 332 Traffic Control Cabinet and both 170 and 2070n controller, for our current development. The update rate for the 3M Canoga C922 card is 13Hz.

4.3 Vision Based Multi-lane Vehicle Tracking Software Development

The software has the following two components:

- ◆ High precision synchronization of the timers on the two computers through wireless communication;
- ◆ Real-time multi-lane and multi-vehicle tracking using the video camera, and matching signals from the two data streams, which are described respectively in this section.

4.3.1 Synchronization of the two Computers with Wireless Communication

The two data streams (from the loop and the video camera) include timestamps for matching. Potential faults are diagnosed by comparing the matched data pairs. The video processing data and the loop data are collected with time stamps in two different

computers. Therefore, computer system time synchronization is critical.

We use wireless-based (UDP) synchronization tool developed by the California PATH to synchronize the two computers within 1 millisecond difference. The procedure is described as follows:

Step 1: *Computer A* send a signal packet, *MSG1*, to *Computer B* containing its current system time, say *Start_TIME*.

Step 2: When *Computer B* receives the signal packet *MSG1*, it immediately sends back the acknowledge signal packet, *MSG2*, to *Computer A*, which contains:

a) the *Start_TIME* from the packet it received

b) the current system time on *Computer B* at the time of receiving the packet from *Computer A*, say *Rcv_TIME*

Step 3: *Computer A* gets the acknowledge message, *MSG2*, and marks the current system time when it received *MSG2* from *Computer B*. Then, the round trip time *R1* of data passing is calculated by subtracting the *Start_TIME* from the current system time. The clock skew between the two computers are then estimated by comparing *Rcv_TIME* with $(Start_TIME+0.5*R1)$.

Step 4: *Computer A* sends a time setting packet to *Computer B* with the clock skew and *Computer B* adjusts its system time accordingly.

The above process is iterated for 100 times and the average round trip time is used to estimate the clock skew. According to our experiment, the resulting clock skew is far less than 1 millisecond. It is a much more accurate and reliable way to synchronize the

two computers than other affordable methods, such as using GPS units.

4.3.2 Real-Time Multi-lane Vehicle Tracking Algorithm

We have designed a computer vision system to obtain baseline measurements to compare. The whole system consists of three parts:

- ◆ a camera (Canon VC50i pan-tilt-zoom communication camera);
- ◆ a moving platform; and
- ◆ Linux-based video processing software

The Canon VC50i camera provides a wide range of view by panning through a broad reach of 200 degree, tilting of 120 degree, as well as a 26x optical zooming. It provides superior camera optics that a good quality of images can be obtained even with challenging illumination conditions, such as strong shadow cast that causes too high image contrast. An Intel Pentium Core laptop computer equipped with a USB frame grabber was used for video data processing.

There are many commercial vision-based vehicle detection systems (“virtual loop detectors”) also available. However, most virtual loop detectors are based on the background subtraction algorithm. They normally use frontal-view video images to avoid difficulties caused by occlusions and to get better lane positioning. Since our application requires the camera on the roadside, it is difficult to adopt those systems.

We applied the rectilinear camera based vehicle tracking algorithm described in Section 3.3 to detect and track vehicles in multiple lanes at the same time. The algorithm combines the background subtraction algorithm with the feature tracking and grouping

algorithm to better handle the occlusion problem. The developed algorithm is more robust to shadow and occlusions than conventional virtual loop detectors and, thus, better separates between lanes. An example detection result is shown in Figure 4-4. We see that the upper-left background subtraction result cannot separate the vehicles in multiple lanes but the newly developed algorithm can correctly localize them.

An example placement of a “virtual loop” over the loop mark on the ground is shown in Figure 4-5. The trajectories of all the vehicles in the image are estimated, and the virtual loop is triggered by analyzing the trajectories.

The software was developed under Linux environment using the OpenCV library. The algorithm runs on a Pentium Core processor (1.83GHz) in real-time at 10 fps. The details of the image processing algorithm are described in [40].

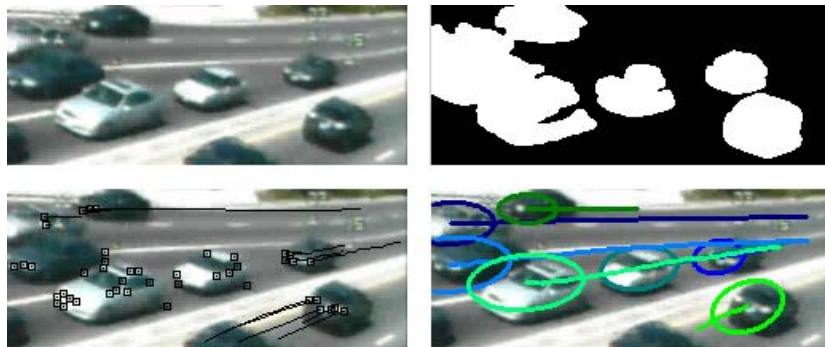


Figure 4-4: An example feature tracking²



Figure 4-5: Virtual loop is triggered when a vehicle reaches it

² The video tracking algorithm part was developed by Dr. Kim

The 10Hz update rate is frequent enough to avoid any missing vehicle count due to high vehicle speed. For example, when a vehicle moves at 70 mph or 31.3 m/s, and loop length 2m and vehicle length 4m, the total crossing length for a vehicle starting on upstream edge and leave at downstream edge is: $2+2*4 = 10m$. Thus the expected duration is $10.0/31.3 = 0.32s$. If the video camera update rate is 10 Hz, there are at least several frames of the video where the vehicle is on the virtual loop. In addition, even if a fast vehicle passes through the virtual loop between the frames, we can still infer vehicle's passage by analyzing the continuous trajectory that the vision algorithm provides.

4.4 Comparison of Physical Loop and Virtual Loop

Figure 4-6 illustrates the system structure developed to graphically monitor and compare the loop information. The instantaneous physical loop information and virtual loop information packets are processed and formatted as follows:

<pre> Loop Information Package typedef struct{ double timestamp; double Inductance[Max_Loops]; }Loops_TYPE </pre>	<pre> Virtual Loop information package typedef struct { double timestamp; double On[Max_Loops]; }Virtual_Loops_TYPE </pre>
---	--

Thus the physical loop packet and virtual loop packet are matched based on the time stamp. However, the packet update rate from the vision system and that from Canoga card are different. The update of information packet for a vehicle over the virtual loop in the vision system is about 10 fps. The maximum update rate from the Canoga card is around 13 Hz. So this is not a one-to-one matching. On the other hand, the messages from both sides could possibly have some delay due to wireless communication or some

other unknown reasons. To solve this problem, two First-In-First-Out (FIFO) buffers were built on the cabinet computer. One is used to store virtual loop packets from vision system and the other is used to store the physical loop packets from the Canoga card. With those two buffers, two initially synchronized computers can work independently as long as the data is time stamped. Each packet from the video computer (which has a lower update rate) is matched with the physical loop packet which has the closest time stamp by looking up their buffer. This approach significantly increases the reliability of the system.

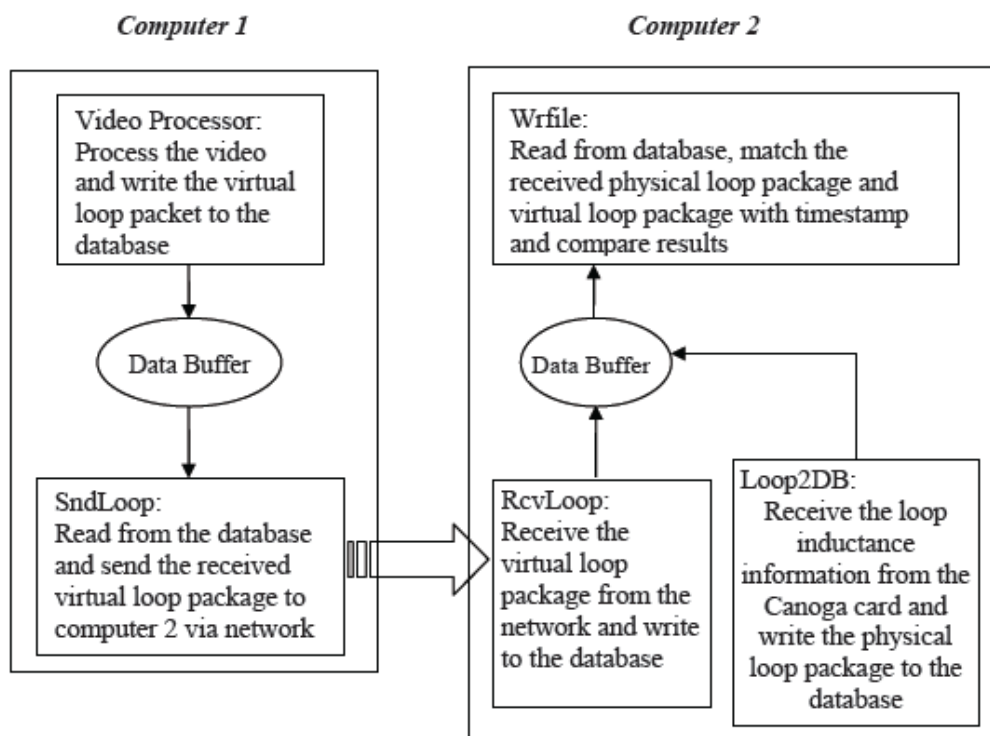


Figure 4-6 Software structure and interaction between the two laptops

4.5 Experiment Data Analysis

Tests have been conducted at an experimental intersection at PATH Headquarters,

RFS, U.C. Berkeley. A 3M Canoga 922 card was connected to a 322 traffic control cabinet to read the raw loop inductance data directly from the physical loop, as shown in Figure 4-7(a). It also transfers these loop information to the laptop through the RS-232 serial cable as shown in Figure 4-7 (b). The 3M Canoga 922 card could read at most two physical loops at the same time.

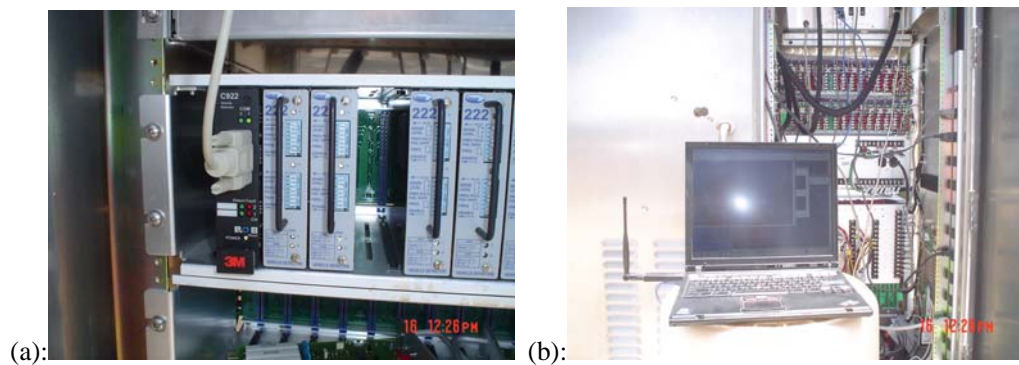


Figure 4-7 Laptop RS232 serial interface with C922 3M Canoga Card

On the other side, the vision system was set up as illustrated in Figure 4-2. The camera was mounted on the top of the trailer pole to look downward towards the loop detectors at the RFS test intersection. The camera's intrinsic parameters were estimated by using the Camera Calibration Toolbox for Matlab [65]. The extrinsic parameters are estimated by a simple external calibration algorithm which uses a single rectangle [41]. USB 300mW WiFi adapter with 9dBi and 5dBi antennas were used for reliable wireless communication with about 800m distance coverage.

Figure 4-8 shows the vehicle detection and tracking process. A virtual loop is turned on (as highlighted) when the vehicle ellipse hits the loop rectangle in the world coordinates in the vision system.

We extensively tested the system at the Richmond Field Station at UC Berkeley and

all of the vehicles passing through the intersection have been detected from both virtual loop system and physical loop system based on the observation. Note that it is a particularly difficult environment for video processing due to heavy moving shadow of trees movements by a strong wind. The packet buffer described in the previous section makes the synchronization only have an average error of 0.0436 seconds, which means an error of 1.16 meters in space if the vehicle runs at 60 mps. Figure 4-9 (a-e) shows the comparison of detections from virtual loops and physical loops when the vehicle is running at different speeds. The red bars represent the virtual loops' on/off information and the blue bars indicate the inductance changes of the physical loops. Each column shows a pair of results for the corresponding virtual and physical loop. The reference origin time in the four sub-figures is exactly the same. The x axis represents the time domain with the number of packets as a unit. Since the vision algorithm works at around 10Hz, each packet is about 0.1s long. In our experiment, two loops were monitored. The exact physical loop size was 2 meters in width by 1.8 meters in length. Considering the vehicle's physical length, the efficient length of a regular sedan is around 6 meters.

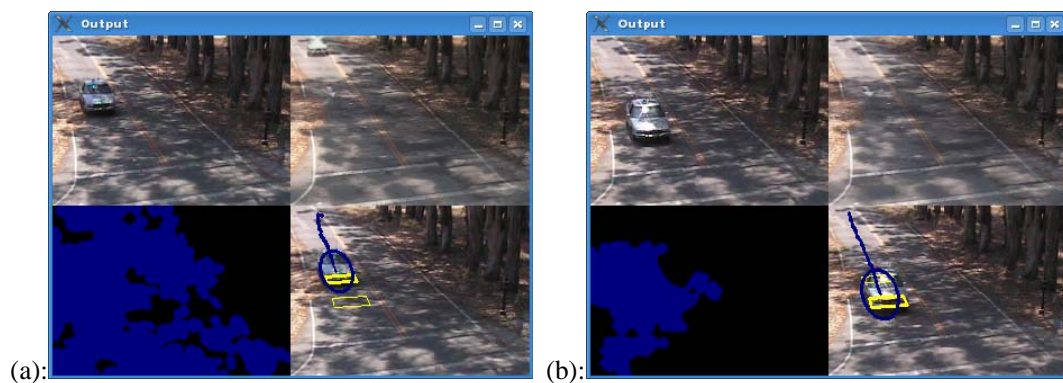
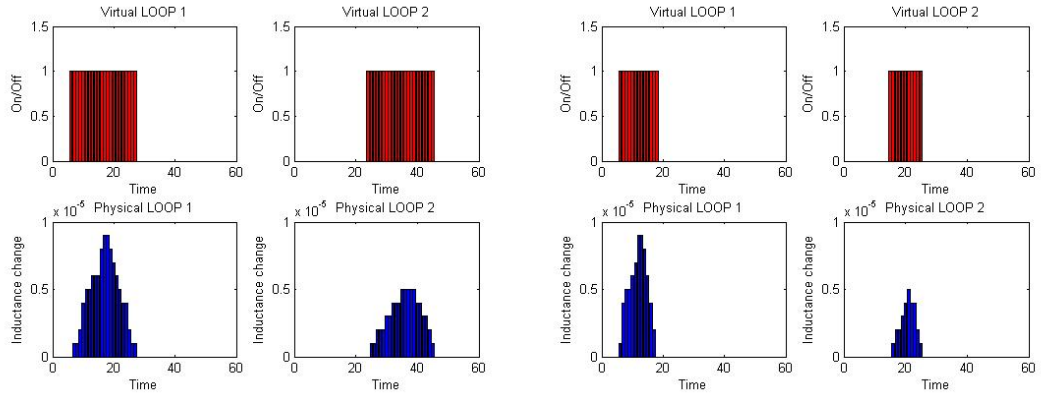
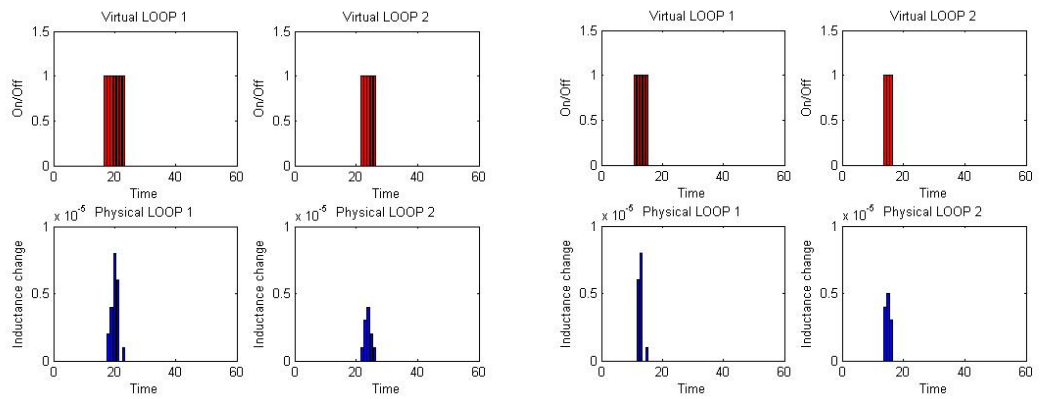


Figure 4-8 Visual loop in the video stream



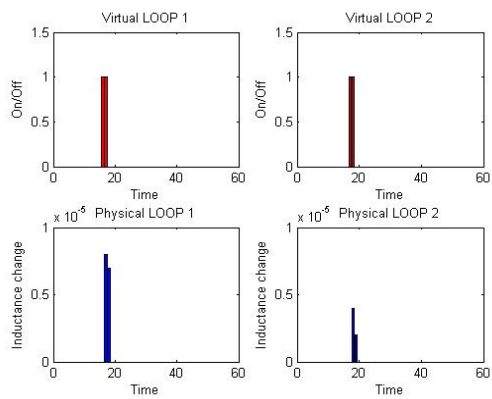
(a) vehicle speed at 5mph

(b) vehicle speed at 15mph



(c) vehicle speed at 25 mph

(d) vehicle speed at 45 mph



(e) vehicle speed at 50mph

Figure 4-9 Comparison of the virtual loops and physical loops data

In Figure 4-9, we see that at a low speed, such as 5 mph in (a), 15 mph in (b) and 25 mph in (c), the physical loop data is bell-shaped, while at high speed, such as 45 mph in (d) and 50 mph in (e), it is shown as a signal pulse. The vertical is the inductive intensity variation calculated based on variation of the pulse frequency of the loop as a vehicle passing over relative to the inductance of absence of vehicles. It can be observed that as the vehicle speed increases, the occupancy time decreases. At the speed of 50mph, the duration only lasted for two time steps. Even in higher speed, it is expected that, the vehicle over loop can still be caught due to internal reset frequency of the loop. For the vehicle tracking over virtual loop with video, the time instant for virtual loop ON can still be guaranteed due to continuous tracking in advance. Note that the inductance intensities of the two nearby loops are different even for the same vehicle passing at the same speed. This implies a practical challenge in directly using the inductance data as the only vehicle signature for re-identification over different loops.

4.6 Conclusions

In this chapter, we presented the research and development of a portable tool for systematic loop fault detection at the control cabinet level, in which vision sensor and vision based vehicle detection and tracking technique plays an important role, serving as an independent source to provide baseline data for comparison. Experimental tests up to 50mph of a vehicle speed demonstrated that this concept is feasible in operating in real-time. Continuous tracking of the vehicles (using the unique vehicle tracking algorithm described in Section x) from further the upstream of the loop guaranteed that

the vehicle is reliably caught over a loop even at a high speed. An effective and reliable synchronization and data communication scheme was also presented.

5. Mobile Traffic Surveillance System Using a Unique OODV Sensor

5.1 On-Board Vision Sensor to “Sense” Soundings

Vehicle-based vision sensor systems are developed for the purpose of determining the vehicle’s position relative to its immediate “sensed” surroundings. The original drivers of developing the video-based object detection and tracking system are the autonomous driving and driver assistance systems. These different applications require cameras to be mounted at different places on the vehicle. Figure 5-1 shows the view angle coverage of cameras at different mounting positions. Figure 5-2 shows different applications with cameras at different mounting positions:

- ◆ Automatic Parking Systems typically need forward and rear looking cameras;
- ◆ Lane-Departure Warning/ Lane Changing Assistance System need side or forward-looking cameras;
- ◆ Forward Collision Avoidance Systems need forward-looking cameras; and Blind Spot Detection Systems need side-looking cameras.

These systems are mainly designed for providing more information about the driving surroundings to the driver for purposes of safety. For example, Bertozzi et al. in [54] provides a comprehensive survey of the use of computer vision techniques in intelligent vehicles. Approaches for obstacle and pedestrian detections are described and analyzed. Recently, a number of researchers have considered lane detection and tracking as the

main purpose for on-vehicle vision system [62]. This is really an important and challenging vision-based application, because:

- ◆ The system process should operate in real-time;
- ◆ The system must robustly identify lane markers, even though they may have a variety of appearances, even across small regions.
- ◆ The system must minimize as many assumptions as possible in the processing.

Most of the research and development deal with obstacles in front, side vehicles and lanes in isolation, and less has provided full surrounding scenarios together. On the other hand, almost all of the previous work focuses on the driver assistant system and there is a less research in developing a system that is for the purpose of observing and evaluating the localized traffic conditions.



Figure 5-1 View angle of on-vehicle cameras

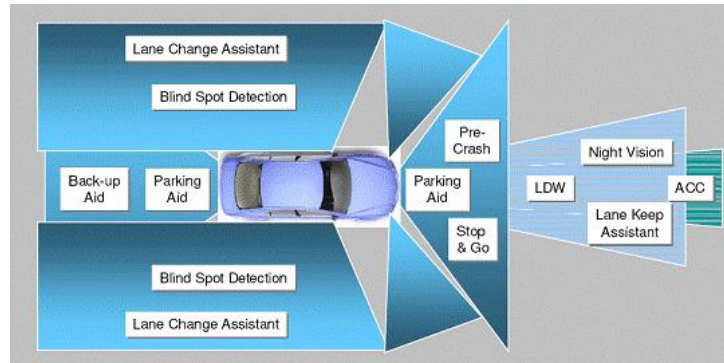


Figure 5-2 Different applications of on-vehicle cameras

In recent years, there has been an increasing interest in omni-directional vision sensing in the roadway environment. Traditionally, forward and backward looking rectilinear vision sensors are used to detect the surrounding vehicles for potential collision avoidance as well as the lane information for navigational purposes. However, the narrow field of view limits the applications of rectilinear cameras, since it is often important to observe 360 degree of surrounding environment as well as the ground plane. In [36], Gandhi and Trivedi give a good review on the capture of vehicle surroundings using video inputs and describe a novel method to create a Dynamic Panoramic Surround (DPS) map using stereo and motion analysis of videos from a pair of hyperboloid mirror-based omni-directional cameras. This system can be applied in real-time systems for driver assistance or off-line for studying vehicle behavior. Most of these omni-directional vision sensors obtain a 360 degree of panoramic surrounding views and a transformation process is required to unwarp the polar-coordinate image to the perspective view.

In this research, aiming at measuring the spatial and temporal varying traffic data (mainly focus in local density and average speed), a novel OODV system based roadway

traffic surveillance system has been designed and implemented. It consists of three major components:

- ◆ A GPS stamped roadway traffic data collection technique;
- ◆ Post video processing that includes automatic vehicle detection/tracking with the ability to correct using interactive tools;
- ◆ A traffic parameter (localized density) estimation process.

5.2 Vision Based Mobile Traffic Surveillance System

5.2.1 Localized Traffic Map Captured with On-vehicle Vision Sensors

By mounting two OODV sensors (for details, refer Chapter 3) on the back corners of a box truck and a single forward-looking rectilinear camera (Point Grey Dragonfly 2 digital camera) on the top of truck box, as shown in Figure 5-3, the surveillance system could be used to observe not only surrounding vehicles and other objects, but also to observe roadway infrastructure as the vehicle travels down the roadway. This roadway infrastructure includes lane markings, barriers, pavement defects (e.g., potholes), and other items lying on the ground plane. These observations can also be combined with high-accuracy position and speed information provided by differential GPS receivers.

In this experiment, the optical axis of each OODV is carefully adjusted to be perpendicular to the ground plane, providing a top-down view of freeway ground map. Figure 5-4 illustrates a snapshot of the recorded roadway traffic video data; Each OODV gives a view of its side of the road, as well as a rear side view. OODV calibration is performed offline in the laboratory to determine the system parameters (details refer to

next section). Combined with the accurate height measurement, the relationship between the OODV pixel coordinate and OODV camera coordinate can be determined. The motion compensated background subtraction based method (described in Section 3.2) has been applied to detect the moving vehicles in the OODV image domain. The forward rectilinear camera adopted a different detection algorithm, which is based on the dynamic feature detection and grouping algorithm as described in Section 3.3. Once the vehicle is detected, its center point is transformed from its image domain to the camera coordinate system, and finally transformed to the vehicle coordinate, with the assumption that the vehicle is at a fixed height. The three frames captured simultaneously by OODVs and rectilinear camera are processed independently and the results can be fused to provide a full description about the localized traffic map around the truck. By carefully calibration, the three cameras in the mobile traffic surveillance system provide a 130 (feet) by 85 (feet) monitoring area on the ground plane. The top-down and lateral display about the field of view is given in Figure 5-5.

The following coordinate systems are used to create the full description of the surrounding traffic map, as illustrated in Figure 5-6:

A) Vehicle Coordinate System: this is the main reference system for the surrounding traffic map. All the vehicle positions calculated from each vision sensor in their camera coordinate systems are transformed into the vehicle coordinate system to remove dual detection and create the surrounding traffic map. The origin O is placed in the center of the vehicle projection on the ground. The X axis points to the right side of the vehicle and Y axis points to the front of the vehicle. Z axis points to the ground upwardly.



Figure 5-3 Vision sensor and GPS-equipped vehicle



Figure 5-4 Captured roadway image data

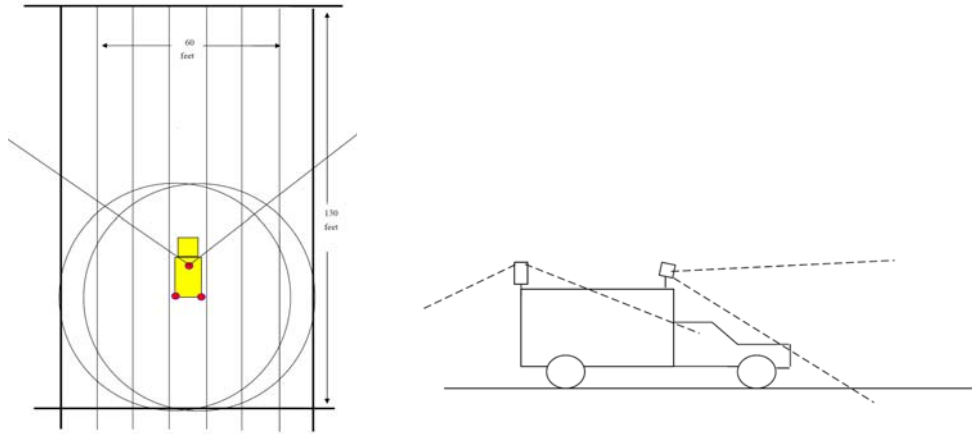


Figure 5-5 Top-down and lateral display about the field of view

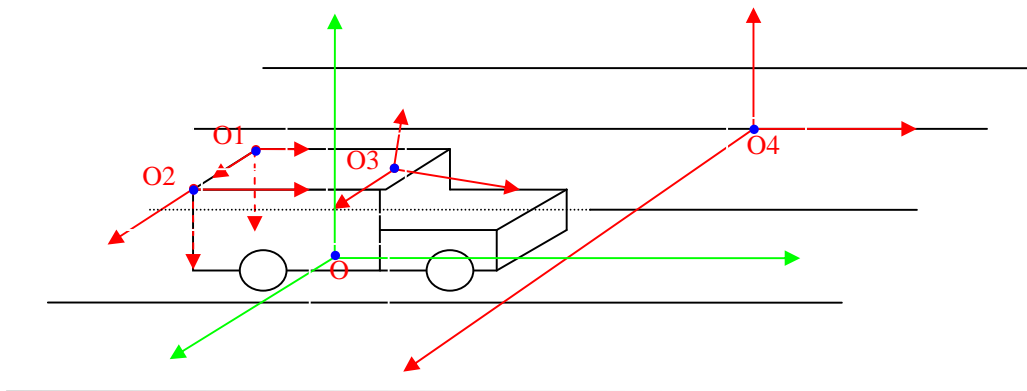


Figure 5-6 Coordinate systems used in the mobile traffic surveillance system

B) OODV Camera Coordinate System: The OODVs are well aligned to be perpendicular to the ground plane, so the origin of the each OODV coordinates, denoted as O1 and O2 in Figure 5-6, is placed in the center of the OODV mirror. The X axis and Y axis are along the edges of the truck box (cube): X axis points to the right of the truck box and Y axis points to the front of the truck box. Z axis is up-to-down perpendicularly pointing to ground.

C) Rectilinear Camera Coordinate System: the rectilinear camera is looking forward and slightly pointing downward. The origin O3 is illustrated in Figure 5-6. The Z axis is

along the optical axis of the camera, pointing to the forwarding area of the vehicle. X axis and Y axis form a perpendicular plane to the optical axis, with X axis points to the right side of the vehicle.

D) Rectilinear Camera Auxiliary Coordinate System: by extrinsic calibration, we can build a relationship between the rectilinear camera coordinate system and a real-world coordinate system, with the origin placed on the ground plane. However, the origin of the vehicle coordinate system O is on the back of the rectilinear camera and thus is invisible in the image. As a result, it is very difficulty to collinear the real-world coordinate system with the vehicle coordinate system in the calibration process. Thus we introduce an auxiliary coordinate system, which serves as a bridge between the camera coordinate system and the vehicle coordinate system. Its origin O4 is placed on the ground in front of the vehicle. The X axis and Y axis are in parallel with the X axis and Y axis of the vehicle coordinate system, pointing to the right side and front side respectively. Z axis points upwardly on the ground plane.

E) Image Plane Pixel Coordinates: defined in the image domain and represented by pixel positions.

5.2.2 Calibration of the Vision Sensors

A). OODV Calibration

Section 3.1 provides a theoretical analysis for the projection model of OODV. We conclude that it approximates a pin-hole based perspective projection with the view point at the mirror center. The distortion component is eliminated in this application because

the processing is performed in the image domain. However, we do need to know the relationship between the OODV image plane coordinate and the OODV camera coordinate, because the position of the vehicle center will be projected to the vehicle coordinate to generate the surrounding traffic map. As described in Section 3.1, the OODV system has a linear scaling property that in a perpendicular plane of the OODV optical axis, the length of an arbitrary line is actually the product of the pixel length of its image and a scaling parameter. A different perpendicular plane has a different scaling parameter. Thus the main purpose of OODV calibration is to find the relationship between this scaling parameter and the depth of the perpendicular plane.

This process is performed offline in the laboratory. In our experiments, characterization is completed by capturing images of OODV with varying heights from a gridded plane perpendicular to the sensor's optical axis, see Figure 3-2, Left. The relationship is then determined by extracting the grid's pixel coordinates from the images and comparing it to real world's coordinates. The size of the grid is 2 (feet) \times 2 (feet). 12 images were taken by the ODVS with different depths between the gridded planes to the OODV. For a fixed distance, 28 points were chosen in the gridded plane to find their corresponding points in the images.

When the focal length f and the distance from the gridded plane to camera d is fixed 28 points were selected for each in the image. We used the ray-tracing technique to find their corresponding 3D real-world point. In Figure 5-7(a), the blue line shows the relationship between the pixel distance and the real-word distance and the red line is the linear fitting of this relationship using Least Square (LS) error criterion (see Appendix B),

which means the slope of the fit line actually equals to the scaling parameter for the particular depth.

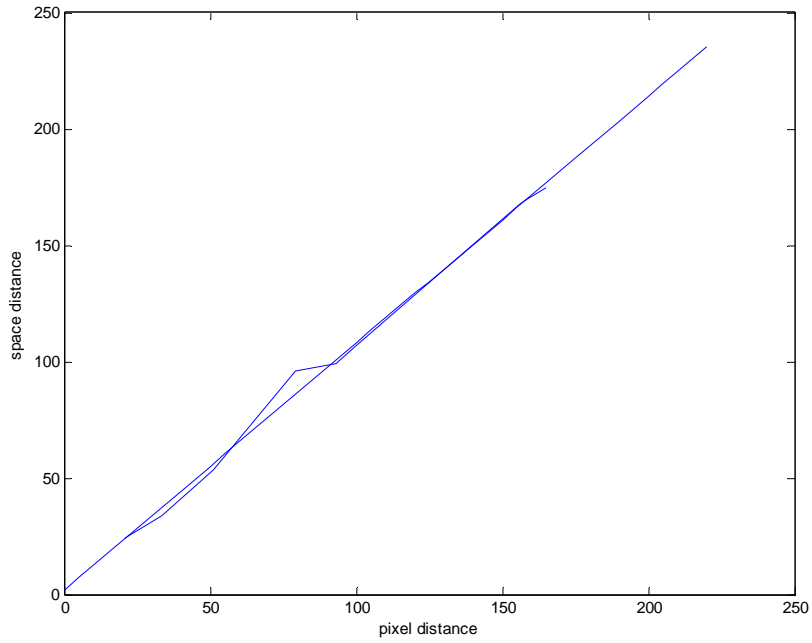
For 12 different depths d from the gridded plane to the camera, we repeat the above process 12 times and thus we can get 12 different scaling parameters. Figure 5-7 (b) illustrates the relationship between scaling parameters and depth.

The center of the mirror is defined to be the origin of the OODV coordinate system. The Z axis is the optical axis of the camera, pointing down to the ground. The X axis and Y axis are lying in the tangent plane of the mirror, but the direction of X- and Y- axis can be arbitrary, because it is 360 degree symmetrical. Suppose (X,Y,Z) is a point in the OODV camera system, (x,y) will be its image point. From the geometry relationship and linear relationship of the mirror, we have the correspondence between the (X,Y,Z) and (x,y) as follow:

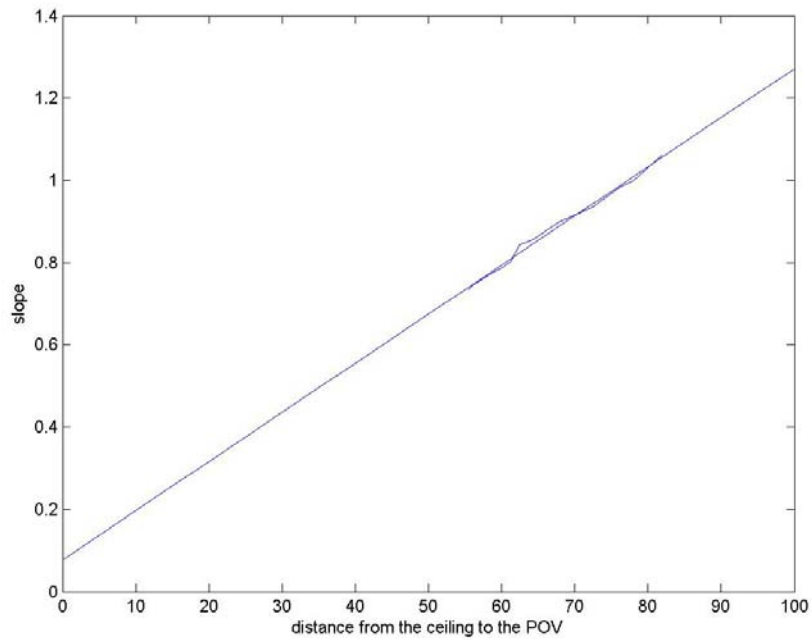
$$x = \frac{X}{AZ} , \quad y = \frac{Y}{AZ} \quad (\text{Eqn.5.1})$$

where Z is the scene depth and A is the system parameter implying the OODV structure information and intrinsic transformation information.

The scaling parameters relationship of OODV is determined in the laboratory and the height of the OODV mounting position is accurately measured before the experiment. The calculated scaling parameter can also be verified by measuring the length and width of the parallel lines in the parking lot. The OODV is 360 degree symmetrical, thus we can align the X-axis and Y-axis to be along with the edges of the box truck (considered as a cube). Figure 5-8 (a) shows the projection of the coordinates in the OODV image plane.



(a):



(b):

Figure 5-7 (a) Linear fitting for the scaling parameter for a particular depth (b) Relationship between the scaling parameters and depth

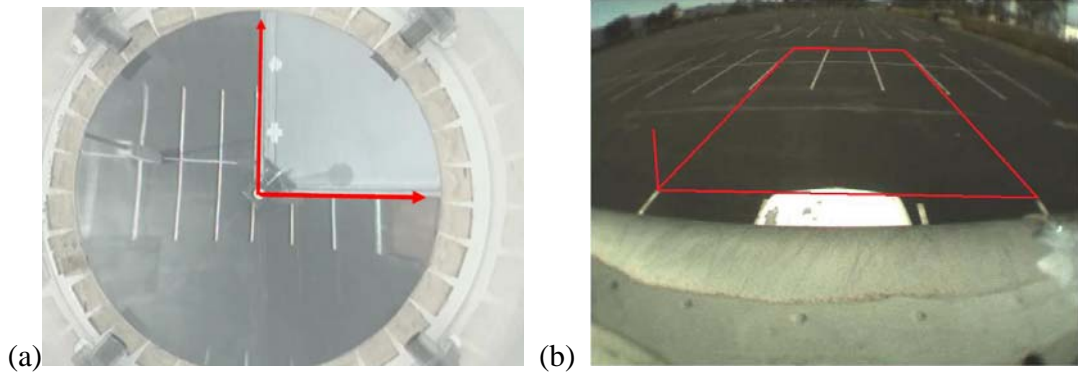


Figure 5-8 (a) OODV coordinates re-projection in the OODV image plane (b) Auxiliary coordinates re-projection in the rectilinear image plane

B) Rectilinear Camera Calibration

A variety of calibration algorithms have been proposed in the past decade to find the intrinsic and extrinsic parameters for the rectilinear camera [41]. The intrinsic parameters determined by the internal properties of the camera (such as principle point, focal length, etc.) describe the transformation between image plane and camera plane. Extrinsic parameters then determine the transformation between the camera plane and the real world coordinates. Vanishing points of an image usually contain important information for the camera calibration. For example, in [41], Kim revisits the geometry of the vanishing point and proposed an extrinsic parameter estimation algorithm based on a single rectangle. This simple algorithm has been proven to provide a competitive accuracy compared with the complex and state-of-the-art calibration algorithm. This method is suitable for the requirement of extrinsic calibration for the roadway mobile video applications. To apply this method, the only thing required is manually drawing two sets of parallel lines in the image plane which forms a rectangle in the real world.

Once you measure the width or length of the rectangle in the real world, you can determine the transformation matrix from camera coordinates to the real world coordinates. In the roadway mobile application, this is easy to achieve, because lane markers provide a natural way to define parallel lines in the scene and the width of the lane is typically 12 feet. On the other hand, this method also provides a potential efficient way to automatically estimate the vehicle pose by real time vanishing point extraction through a lane detection technique.

Figure 5-8 (b) shows the re-projected auxiliary coordinates in the image plane. The auxiliary coordinate system is defined in the forward looking area of the vehicle and the Y axis is always collinear with the heading direction of the vehicle. Thus there is only translation between the auxiliary coordinate system and vehicle coordinate system.

In this experiment, the intrinsic parameters are estimated offline using the Matlab tool box [65]. The calibration of extrinsic parameters is performed by taking the vehicle into a scene with a number of parallel lines in the direction of the vehicle axes. A rectangle is manually created in the image and the translation vector between auxiliary coordinate system and vehicle coordinate system is carefully measured. Thus a position of a detection vehicle in the forward view video can be transformed through auxiliary coordinate to the vehicle coordinate, in order to create the surrounding traffic map.

5.2.3 Video and GPS Data Acquisition

As illustrated in Figure 5-3, a portable GPS receiver (E-TEK EB85A) has been used to capture the GPS signal, and two OODVs and one rectilinear camera have been used to

capture video data simultaneously. The capture rate is 3-4 records per second for GPS data and 8-10 fps for video data. Both of the datasets are indexed by a timestamp each in millisecond. A post data fusion process is applied, so that each frame in the video stream is tagged with the timestamp, instantaneous speed, longitude, and latitude data.

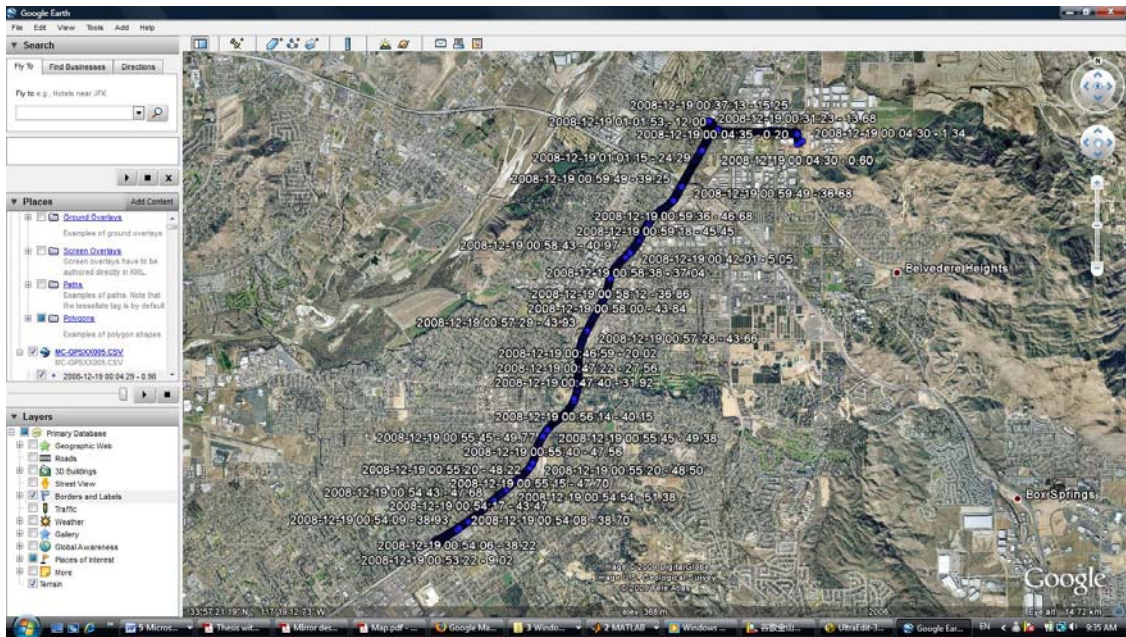


Figure 5-9 GPS trajectories of the test trip on Google Earth

A freeway segments between Riverside, California, United States and Corona, California, United States along State Route-91 is selected as the test region. The length of the test segment is approximately 20 miles one way, which covers 21 effective VDS on westbound SR-91 and 25 VDS on eastbound SR-91. 12 round trips have been taken to collect the mobile traffic data in a four-day test period. The trip times were well distributed from early morning to late afternoon, in order to cover all different traffic conditions including morning rush hour, afternoon rush hour, and free flow traffic.

5.2.4 Localized Traffic Map Generation from Video Post Processing

Traffic parameter estimation is the key goal for this traffic surveillance system. This requires the mobile surveillance system report two outputs for each processed frame: 1) the number of vehicles detected from the three vision sensors; 2). the average relative speed of all the detected vehicles in the scene. To simply the problem, we assume that the instrumented vehicle is always following the traffic and thus assume the average relative speed of surrounding vehicles is zero. This also means the instrumented vehicle speed measured from the on-board GPS is used to represent the average absolute speed of all vehicles in the scene.

To ensure error-free detection, we applied an interactive image processing tool to manually confirm the detected vehicles, removing false positives and adding a negative truth of each vehicle. The detection and tracking was then done automatically based on the methods described in Section 3.2 and Section 3.3 for OODV video and rectilinear video respectively, but any visible detection errors were fixed manually using the interactive tool. We manually go through all the video data we collected in the experiment.

The video data is processed at full rate, but only report a localized density measurement every a second.

Each vision sensor is processed by individual vehicle detector. The OODV based vehicle detector maintains a list of pixel locations for the centers and bounding box of the detected vehicles. As illustrated in Figure 5-10, there are overlapped region for OODV stream in the rear side of the instrumented vehicle, so the same vehicle can be possibly

detected by both of the OODV detector as shown in Figure 5-10. Both the center and bounding box of the detected vehicles in OODV streams are projected to the vehicle coordinate system (assuming a fixed height, say 1 meter in this application). An automatic counting process is performed by eliminating the dual detected vehicle if one re-projected center locates inside of another bounding box and the overlapped region of bounding box is larger than half area of either bounding box. The front vehicle detector maintains an object trajectory in the Auxiliary Coordinate System, thus the position of the object is transformed to the Vehicle Coordinate System for every reported frame. An automatic overlapping checking process is performed by determining whether the object falls into a bounding box of the OODV stream. Through the above procedure, the localized density was reported every second.



Figure 5-10 Duplicated detection in multiple sensors

5.3 Dynamic Roadway and Vehicle Traffic Data Integration

5.3.1 Dynamic Roadway and Vehicle Database Architecture

As stated in Chapter 1, general traffic conditions can be represented by the following parameters: traffic flow rate (volume), average speed, and traffic density (concentration). These parameters are related for uninterrupted flow: flow rate equals to the product of density and speed. Their definitions are as below:

Parameter	Definition	Equation
Density K	the number of vehicles N in a section with length L from an instantaneous photograph	$K = N / L$
Flow rate Q	the number of vehicles M observed from a fixed location during time T.	$Q = M / T$
Space Mean Speed V	average speed of all vehicles	$\frac{1}{n} \sum_{i=1}^n v_i$

Table 5-1 Traffic Parameters definition and equations

The embedded loop sensor system described in Section 2.1 provides a way to measure the flow rate at different locations along roadways. The mobile traffic surveillance system equipped with GPS and vision sensors, described in this dissertation, provides a way to measure the localized density.

In order to incorporate the above two quantitative measurements of roadway traffic conditions from embedded loop sensor system and mobile surveillance system, a dynamic roadway and vehicle data architecture has been designed at CE-CERT, as shown in Figure 5-11. The embedded loop data from the PeMS database is reported to the Roadway Database. This database provides a great support for many applications, such as

eco-friendly navigation, navigational mobility index, and dynamic eco-driving. On the other hand, probe data from other on-board sensors (e.g., GPS) is reported to the Vehicle Database as well, which can be used in applications such as vehicle activity studies, active vehicle tracking, emission analysis, etc. The data from Roadway Database and Vehicle Database can also be integrated and analyze together, for some other applications, such as shared-use vehicle system and city dynamic traffic management. This architecture provides a system level integration of roadway data and vehicle data from a variety of sensors to benefit a variety of applications.

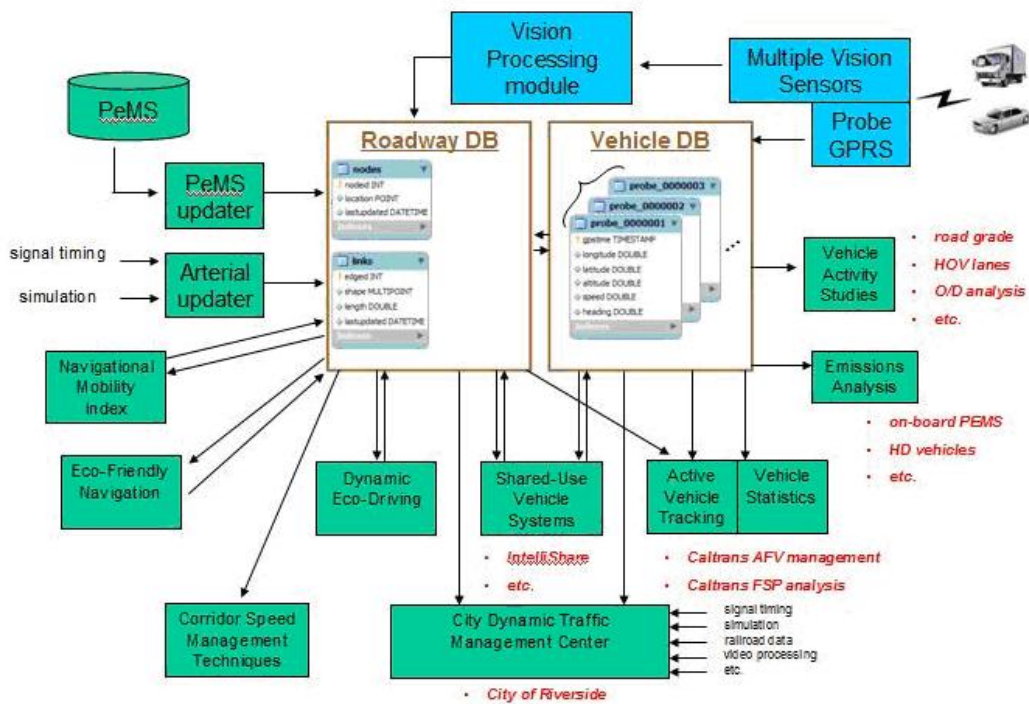


Figure 5-11 Dynamic roadway and vehicle data architecture

The temporal-spatial variable traffic parameters (localized density / average speed, number of lanes) obtained from the vision processing module of the instrumented vehicle

will be filled into the roadway database as well for future post-processing and analysis, which is explained and discussed in the following section.

5.3.2 Relationship between Loop Data and Mobile Traffic Data

To better understand the relationship between loop sensor data and mobile traffic data, we examine the density K and flow Q from another point of view [66]. Figure 5-12 illustrates a time-space diagram including the trajectories of vehicles traveled between locations $[0, L]$ during time period $[0, T]$. Section 5.3.1 introduced the concept of density over a section of road for a specific time as the number of vehicles observed in a photograph of the section with length L , as shown in Figure 5-12 the green bar with time difference infinitesimal $dt \rightarrow 0$. The concept of flow at a fixed position as the number of vehicles observed over a period of time T , as show in Figure 5-12 the yellow bar with length infinitesimal $dl \rightarrow 0$.

By multiplying the numerator and denominator of the equation of density by a small differential of time, we obtain:

$$k = \frac{N}{L} = \frac{Ndt}{Ldt} \quad (\text{Eqn.5-2})$$

Referring to Figure 5-12, the denominator of Eqn.5-2 is the area of the green bar as shown. If we ignore the possibility that the vehicle trajectories enter and leave the green bar from bottom and top side, which is logically impossible, the numerator of Eqn.5-2 represents the total vehicle time (vehicle-hour) spent inside the green bar. Thus we can view the density K as total time in time-space region (veh-hrs) divided by “area” of time-space region (mile-hrs).

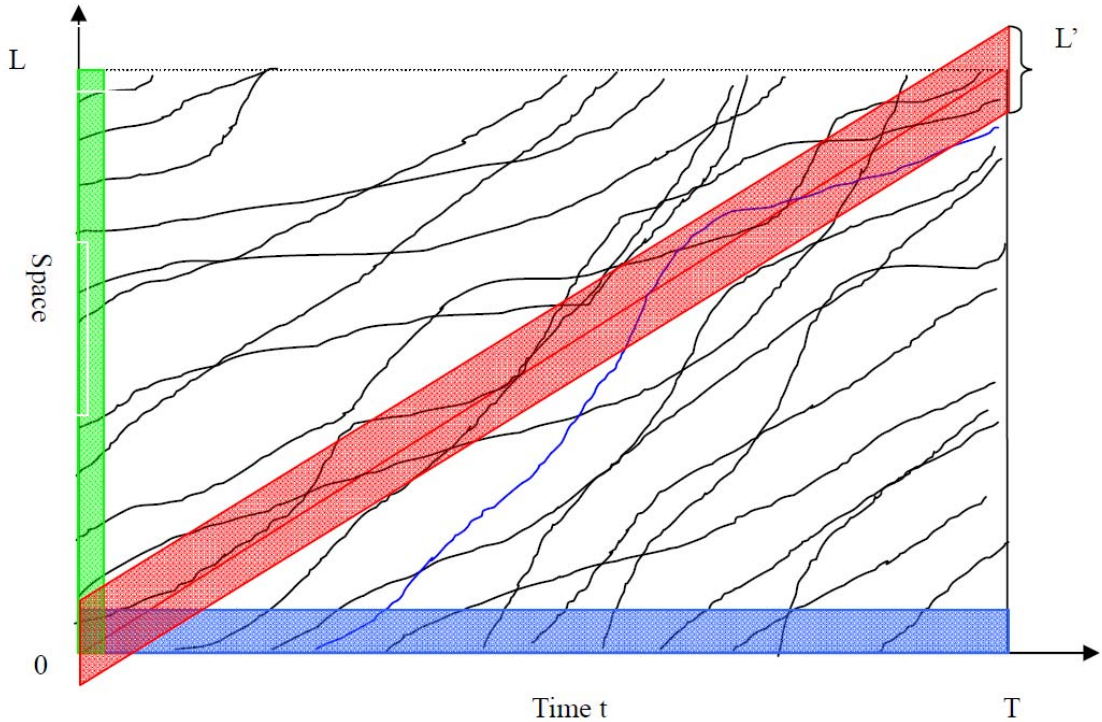


Figure 5-12 Time-space diagram

Using the same analysis process, we multiply the numerator and denominator of the equation of flow by a small differential of space, we get:

$$Q = \frac{M}{T} = \frac{Mdl}{Tdl} \quad (\text{Eqn.5-3})$$

Referring to the blue bar in Figure 5-12, flow rate Q can be explained as total distance in time-space region (veh-mile) divided by “area” of time-space region (mile-hrs).

Eqn.5-2 and Eqn.5-3 provide another view toward density and flow. More importantly, the advantage of this explanation is that it can be applied to arbitrary time-space region over L and T (even non-rectangular).

If we define the total time in time-space region R as $t(R)$ and the total distance in

time-space region R as $d(R)$, then the ratio of flow over density of region R is:

$$\frac{Q(R)}{K(R)} = \frac{d(R)}{t(R)} = \frac{\sum_{i=1}^n v_i dt}{ndt} = \frac{1}{n} \sum_{i=1}^n v_i = V(R) \quad (\text{Eqn.5-4})$$

Eqn.5-4 implies that flow equals to the product of density and average speed, for an arbitrary region in the time-space diagram.

The above explanation is also useful in aggregating the localized density over time and space. The red line in the time-space diagram represents a trajectory of our vision/GPS equipped mobile vehicle. Each snapshot in the red bar at a fixed time t contains information about the local density; we can then define the density over the red bar region as the total vehicle time spent in the red bar divided by the “area” of the red bar, as

$$K = \frac{\int k dt}{L \times T} = \frac{\sum k_i t_i}{L \times T} \quad (\text{Eqn.5-5})$$

where L represents the length of the local observation area and k_i represents the local density at time t_i .

As described above, the best way to characterize the density and flow over a length $[0, L]$ during time $[0, T]$, is by observing the full trajectories of all the vehicles. However this is difficult to accomplish. Loop sensor systems observe traffic at a fixed positions over continuous time to record samples of flow rate, as shown in Figure 5-12 (blue bar) and the mobile traffic surveillance system captures samples of localized density, as shown in Figure 5-12 (red bar). Thus if we view the traffic time-space diagram a random field, then loop sensor data are samples over time at fixed locations and the mobile data are

samples over variable space and time.

Traffic conditions over a certain region are very dynamic. Of course, the more mobile data we have, the better they describe the traffic conditions in the region. The more loop detectors we have, the better they describe the traffic. It does not make sense to compare one sample from a loop sensor system and one from mobile traffic surveillance system, but it does make sense to compare them statistically as a whole. In other words, theoretically, one traffic parameter described for a time-space region, should satisfy the same distribution, no matter how it is measured, i.e, from a loop detector sensor system or from a mobile surveillance system. Based on experiments described in the next section, this has been verified by comparing the 5-mins flow rate Q , which is measured directly from a loop detector sensor system and the calculated and scaled value through localized density measurement by the mobile surveillance system.

5.3.3 Experiment Verification

As described in Section 3.2, we took 12 test trips between Riverside and Corona, in Southern California, United States along the State Route-91, which is approximately 20 miles one way and covers 21 VDS on westbound SR-91 and 25 effective VDS on eastbound SR-91. The trips have been taken in a consecutive four-day period and trip times were well selected to distribute from early morning to late afternoon, in order to cover all different traffic conditions including morning rush hour, afternoon rush hour, and free flows.

We selected 8 VDS each on westbound and on eastbound for data comparison. The

selected VDS are roughly uniformly distributed along SR-91 and the distance between adjacent VDS is between 2.2 miles and 2.7 miles. Table 5-2 shows the details of the VDS information:

Direction	vds_id	Name	Abs PM
Westbound	801559	14 th Ave	57.33
Westbound	801543	Arlington	55.15
Westbound	801527	Adams	52.99
Westbound	801511	Tyler	50.34
Westbound	801496	Pierce	48.15
Westbound	811332	E.Grand	44.23
Westbound	810854	E.Lincoln	41.93
Westbound	811314	Green River	40.12
Eastbound	811309	Green River	40.03
Eastbound	801464	Smith	42.09
Eastbound	811325	E Grand	44.10
Eastbound	811389	Buchanan	47.34
Eastbound	811400	Tyler	49.84
Eastbound	801523	Jackson	52.06
Eastbound	801541	Arlington	54.62
Eastbound	810110	Pachappa	56.73

Table 5-2 List of VDS that is used for comparison

The primary traffic parameter we examine is flow Q . From the historical data of a VDS in the loop detector sensor system, we can plot the distribution of flow Q over speed for a certain space (position of the loop detector) and time (5-mins aggregated flow data over the 4-day test period) region. On the other hand, we can calculate a sample flow from the measured localized density and speed by the mobile surveillance system. What is examined is whether the calculated flow-speed pair satisfies the distribution given by historical data from the loop sensor system.

Here is the method we calculate flow from mobile traffic data: suppose that the surveillance vehicle passes VDS A,B,C at time t_A , t_B , t_C respectively, we wish to calculate the density between A and C during time period $[t_A, t_C]$ by using Eqn.5-5 from the localized density reported by the vision processing module. The instrumented vehicle follows the traffic all the time, thus the average speed of all vehicles is estimated by the time mean of the instrumented vehicle's GPS-derived speed. The flow rate is calculated by the speed-density-flow relationship:

$$q = k \times v$$

The VDS are selected so that the distance between A and C is around 5 miles, so it will take approximately 5 minutes in travelling from A to C at speed 60 mph. The latitude, longitude information for each VDS is compared with the trajectory of the instrumented vehicle in order to extract the exact timestamp that the instrumented vehicle passes point A and point B, as well as the duration time $t_C - t_A$. Then the calculated flow in $t_C - t_A$ is scaled to 5 minutes by linear scaling, noted as Q' .

We plot the 12 calculated flow-speed pairs (one for each trip) into the historical

distribution (from the 4-day test period) of 5-mins aggregated flow-speed diagram at each VDS. We repeat the same process over all 12 samples (test trip) at 16 VDS sites, with the results shown in Figure 5-13 and Figure 5-14. It can be seen that most calculated flow matches the flow distribution very well, except for the VDS 810854. In PeMS, VDS 810854 is claimed as a mainline type in the VDS index page, however in its description page, it is announced that VDS810854 only receives loop data from LDS 813863 (HOV type), which made a great confusion. However, when we go back to the videos, it is a 4 lane freeway segments with one HOV lane. This error may result from a system error from the embedded loop sensor system.

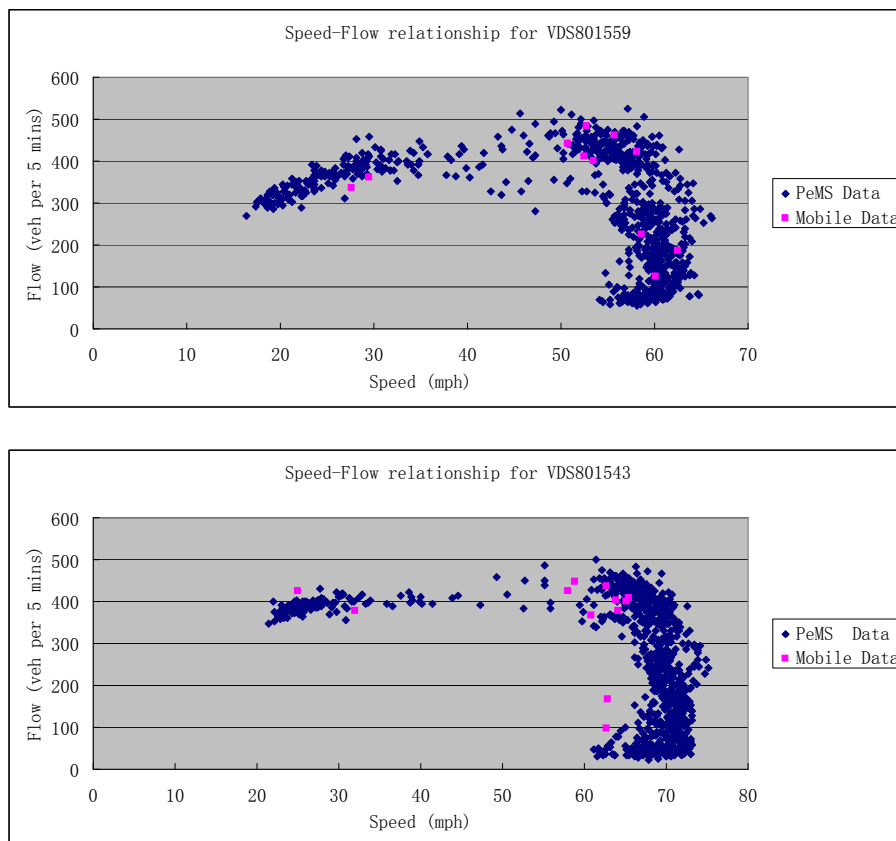


Figure 5-13 (Part A) Mobile and loop data comparison results for Westbound VDS

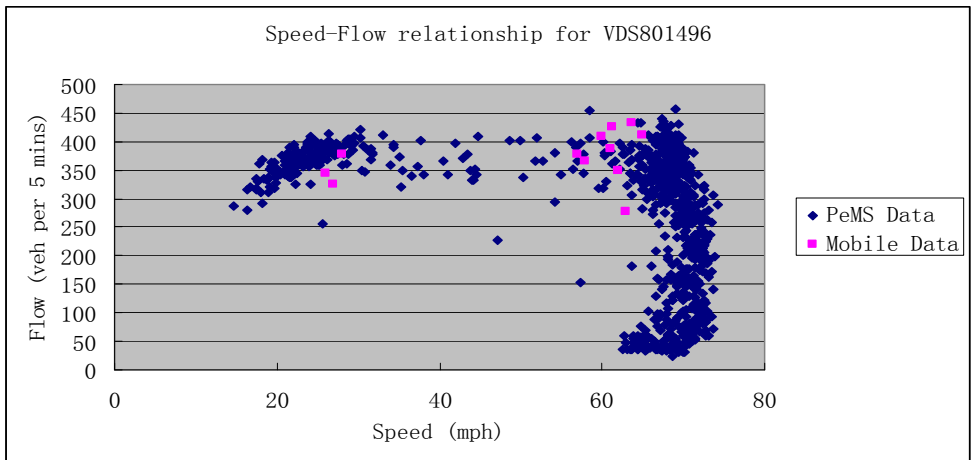
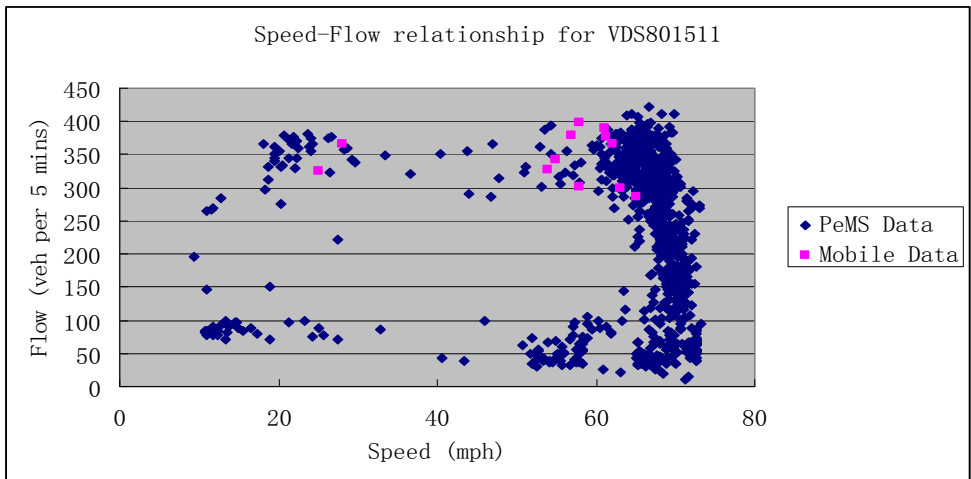
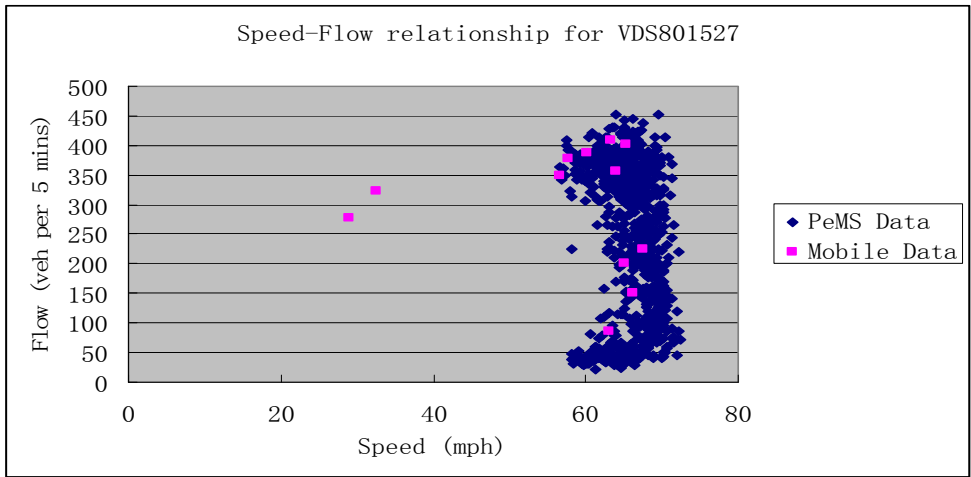


Figure 5-13 (Part B) Mobile and loop data comparison results for Westbound VDS

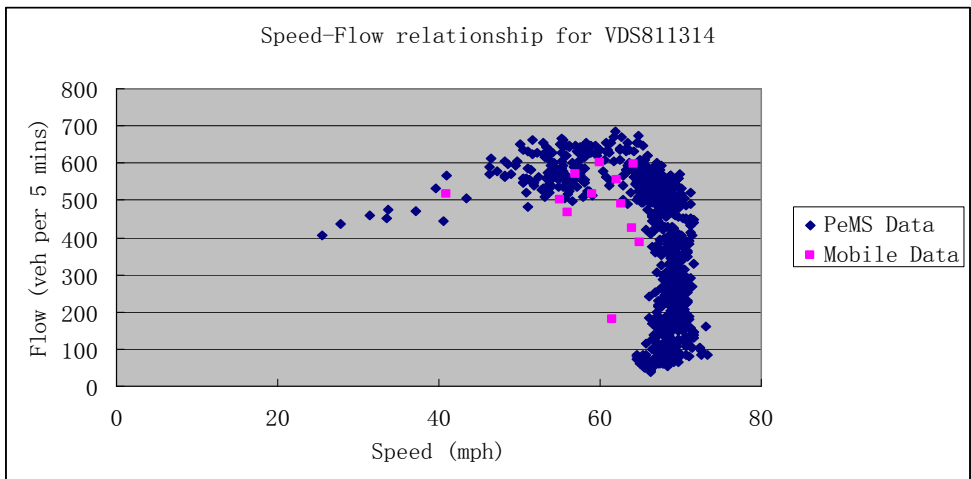
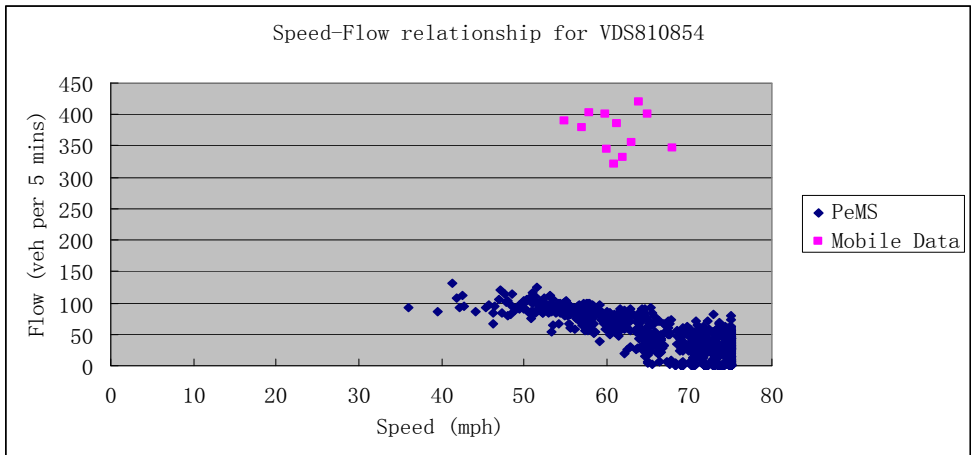
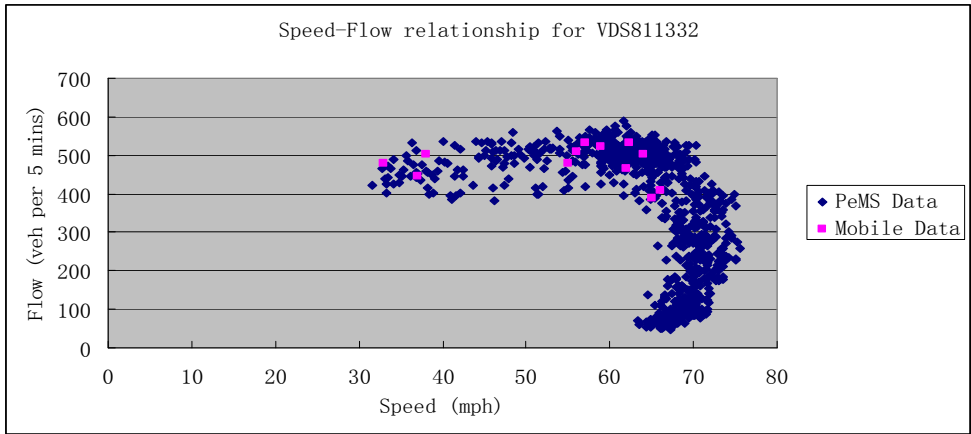


Figure 5-13 (Part C) Mobile and loop data comparison results for Westbound VDS

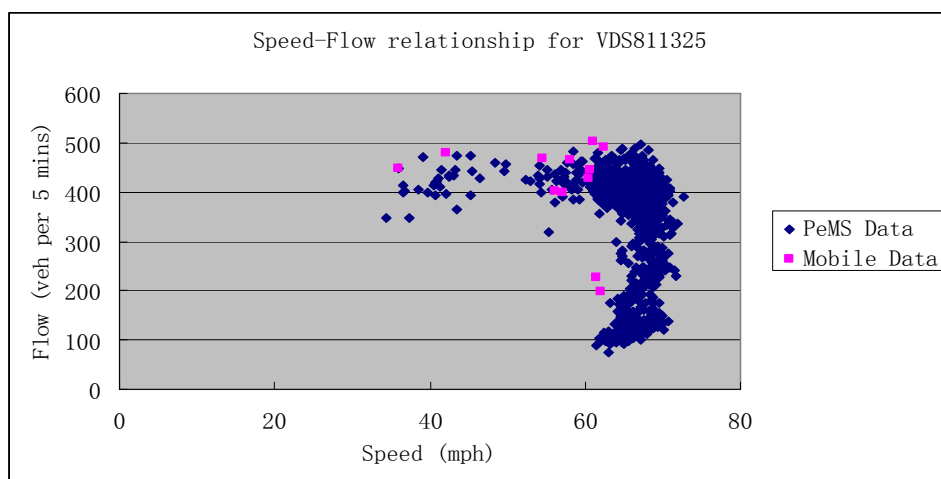
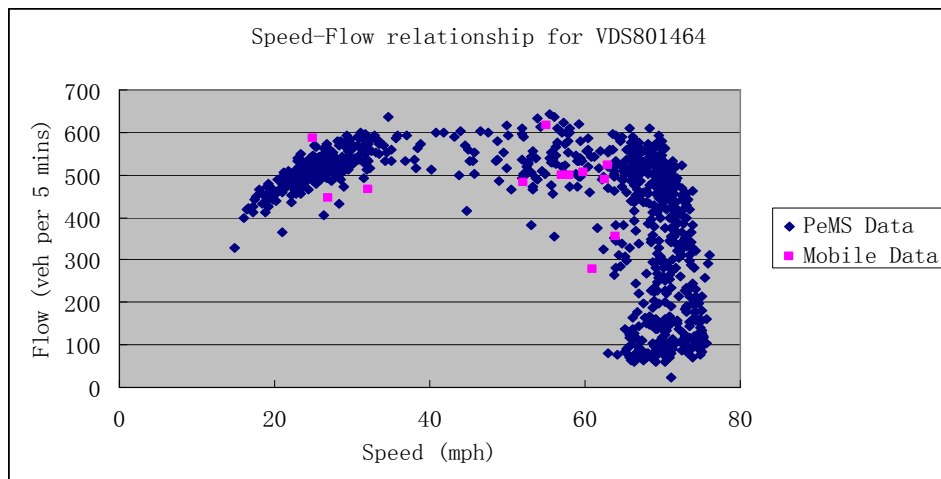
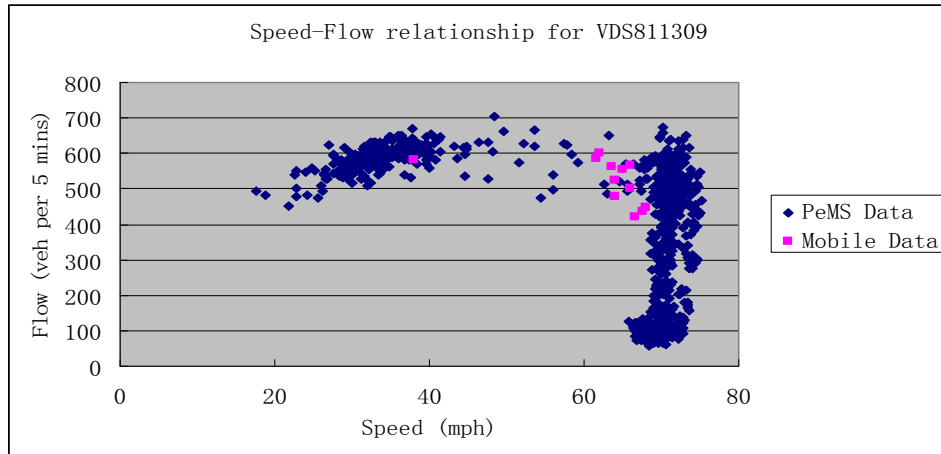


Figure 5-14 (Part A) Mobile and loop data comparison results for Eastbound VDS

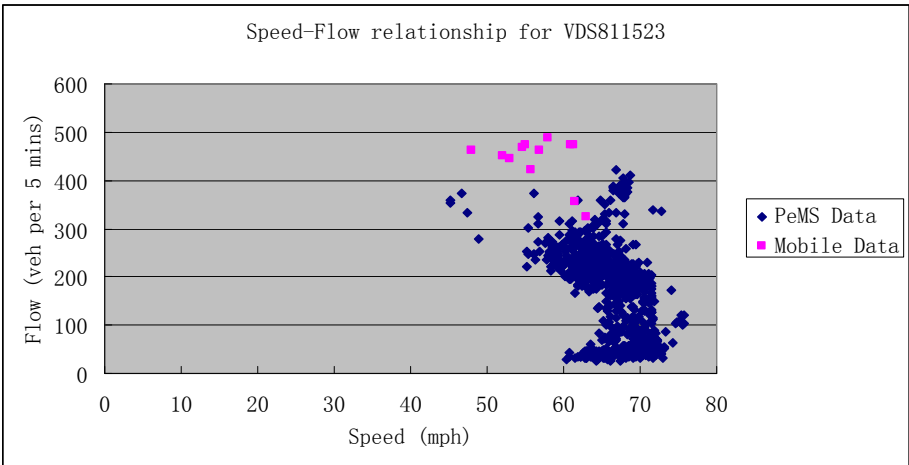
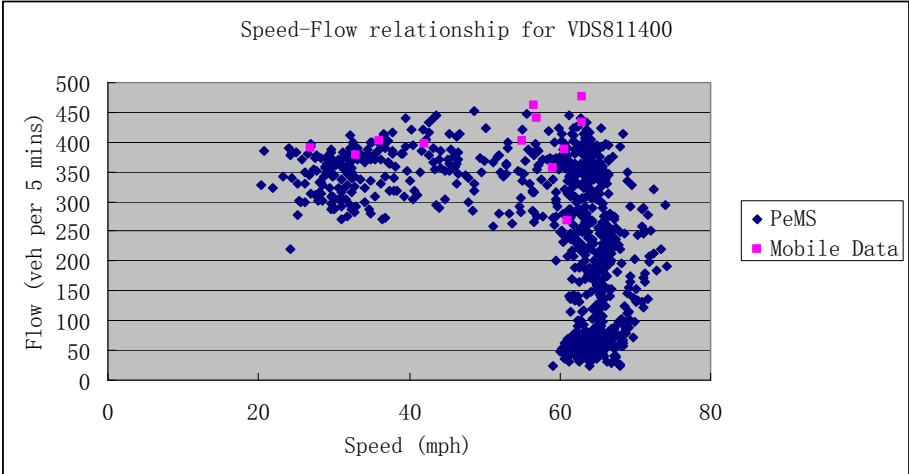
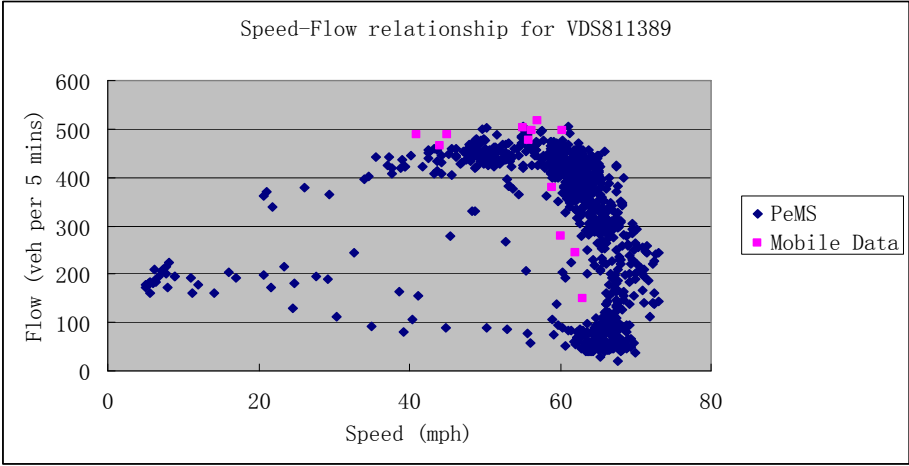


Figure 5-14 (Part B) Mobile and loop data comparison results for Eastbound VDS

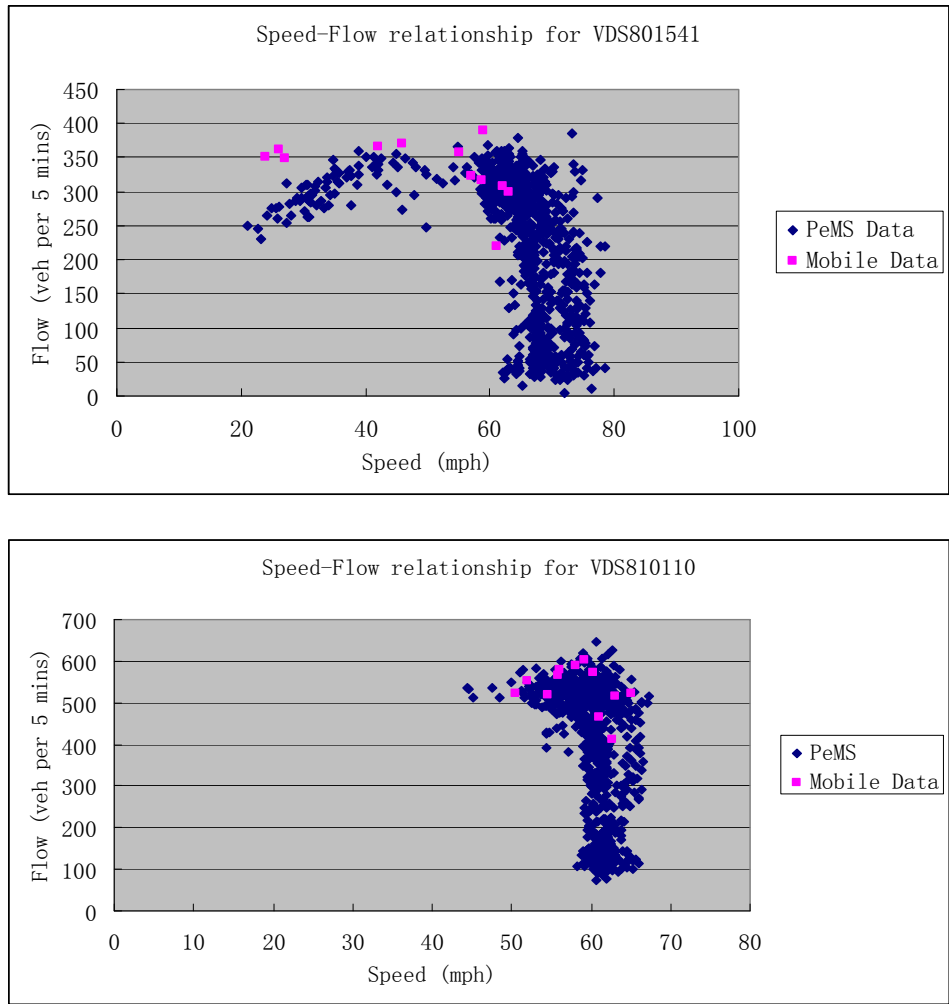


Figure 5-14 (Part C) Mobile and loop data comparison results for Eastbound VDS

5.4 Conclusions

In this chapter, we implemented a vision-based roadway traffic surveillance system installed on moving platform. A surrounding traffic map has been generated through the video analysis both in the OODV stream and rectilinear video stream. The localized traffic density estimated from the implemented mobile traffic surveillance system in this

section is very useful to model and predict the traffic conditions at a range of roadway rather than a certain point. In addition, we theoretically explain the relationship between the loop traffic data and mobile traffic data measured from our mobile traffic surveillance system and experimentally verified the relationship. The work in this section provides a solid foundation in the promising potential data fusion of the roadway and vehicle traffic data.

6. Vision Based Quantitative Measurements of Bicyclist Intersection Crossing Times³

6.1 Computer Vision as a Reliable Method for Quantitative Measurements

As we indicated in Chapter 1, some safety applications in ITS area require reliable quantitative measurements over a large set of traffic data, which is very difficult and insufficient to be collected and processed manually.

A good example is the need for Caltrans to determine how to specify the minimum green signal intervals throughout the state to give bicyclists sufficient time to cross wide arterials from a standing start.

If bicyclists could be detected reliably at intersection approaches and clearly distinguished from motor vehicles, it would be possible to modify the signal timing so that the bicyclists would receive longer green times than the motor vehicles. In this way, the longer green intervals would only be triggered when bicyclists are present, minimizing the negative effects on mainline green time. Unfortunately, this is not within the current state of the art, so bicyclists and motor vehicles will be receiving the same green signal time for the foreseeable future. This increases the urgency of determining the minimum length of green that is really needed by the bicyclists, since this green interval will have to apply all the time, with or without bicyclists being present.

In order to make intelligent decisions about the length of green intervals to accommodate bicyclists, several kinds of information are needed:

³ This work was conducted at California PATH as a joint research with Shladover, Kim, Sharafsaleh and Li, and described at *Proc. Transportation Research Board*, 2009 [64]

- ◆ time needed for bicyclists to start up from a stop after their signal turns green;
- ◆ speed of bicyclists crossing a wide arterial, once they have reached their cruising speed;
- ◆ dependence of bicyclist crossing times on street width;
- ◆ dependence of crossing time on road geometry details (grade, arterial crown);
- ◆ dependence of crossing time performance on bicyclist demographics (especially age); and
- ◆ effects of shortened green time on mainline traffic conditions, with both coordinated and independent signals.

Unfortunately, all the above information requires quantitative measurement on a large set of bicycle traffic data, either created on manual tabulation of bicyclist crossing times using stopwatches or sensor based data collection and processing. Once again, vision sensors become the best choice due to their easy installation, operation and maintenance, high flexibility, low costs and their ability to monitor wide areas. In addition, the development of computer vision techniques, especially the object detection and tracking techniques, makes it possible that efficient automatic video processing (at least, computer-aided processing) provides a reliable quantitative measurements.

6.2 Data Collection and Processing Approach

6.2.1 Bicyclist Traffic Data Collection

The quantification of bicyclist crossing behavior relied on digital video recording and offline video image processing of the bicyclist movements. A video camera was mounted

on top of a mast at a height of 6 – 7 m, atop a trailer which was parked near the subject intersection. The location of the trailer was chosen to provide a clear view of the bicyclist crossing path, with the camera height selected to be high enough that passing vehicles could not occlude the view of the bicyclists. A second video camera was mounted at a height of 2 - 3 m, with a view of the traffic signal head governing the bicyclist crossing movement.

An embedded computer (PC/104) equipped with an MPEG hardware compression board was used to convert the analog video signals into pairs of MPEG video clips. The video clips were synchronized to have less than 33 ms (one scan) error.

The two intersections that were selected for the data collection reported here were El Camino Real and Park Blvd. in Palo Alto, CA, which was recommended by the City of Palo Alto as the crossing with the highest bicyclist volume there, and Telegraph Ave. and Russell Street in Berkeley, CA where the Russell St. bicycle boulevard crosses the major arterial Telegraph. The Palo Alto site is at one of the entrances to the Stanford University campus and therefore serves primarily Stanford students and employees, generally a population of young adults. The Berkeley site provides a more diverse population of bicyclists, including school-age children and more mature adults, in a more residential area. The El Camino crossing has a total width of 125 ft, while the Telegraph crossing is 84 ft. wide, providing another source of variation between the two data sets.

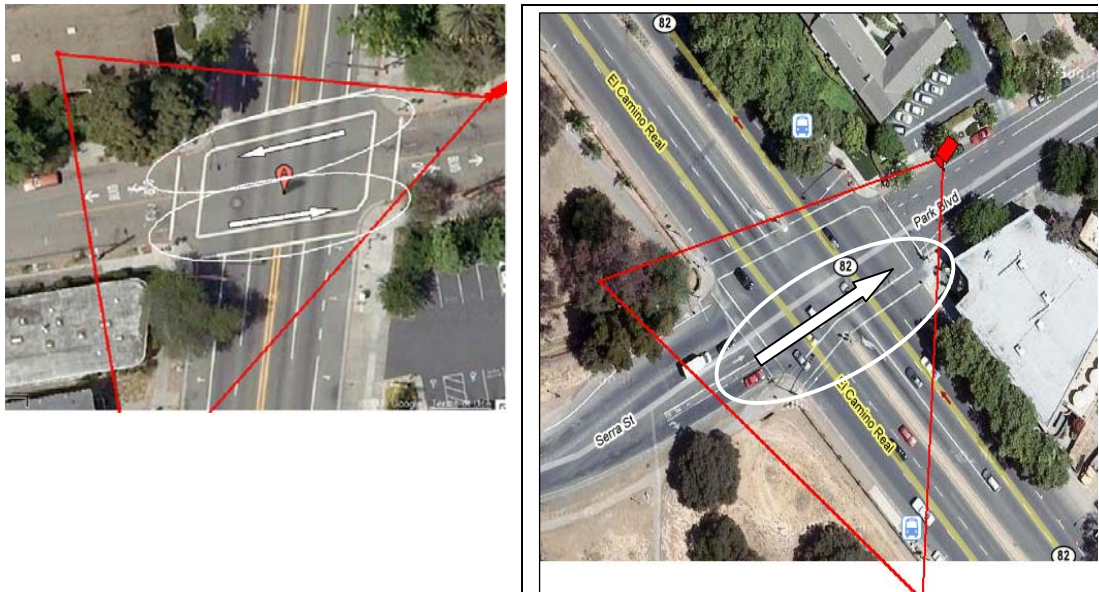


Figure 6-1 Video data collection sites at Telegraph and Russell in Berkeley (left) and Park Blvd. and El Camino Real in Palo Alto (right)

Figure 6-1 shows Google Earth views of the two intersections where data were collected and fully analyzed. The aerial views give an indication of the camera locations and fields of view relative to the intersections and the directions of bicyclist crossings that were recorded.

6.2.2 Data Processing

After the video data have been recorded in the field, they are analyzed in the laboratory to extract the information of interest. The image processing system automatically detects moving objects in the video and generates their trajectories [40]. There are many challenges for fully automated detection, such as poor image resolution on the far side of the image and difficulty in discriminating bicyclists from pedestrians and other vehicles. To ensure error-free detection, we used an interactive image

processing tool. An operator manually confirmed, repositioned or manually marked the initial position of each bicyclist. The tracking was then done automatically but any visible tracking errors were fixed manually using the interactive tool. A screen shot of the interactive tool, with the image processing system's symbols identifying the bicyclists with a thick ellipse, is shown in Figure 6-2. Note the superimposed image of the traffic signal from the second video camera, visible in the upper left corner of each image. An image processing tool was used to extract the signal phase from the image patch, and the result was shown to the user to ensure that the phase detection is correct.

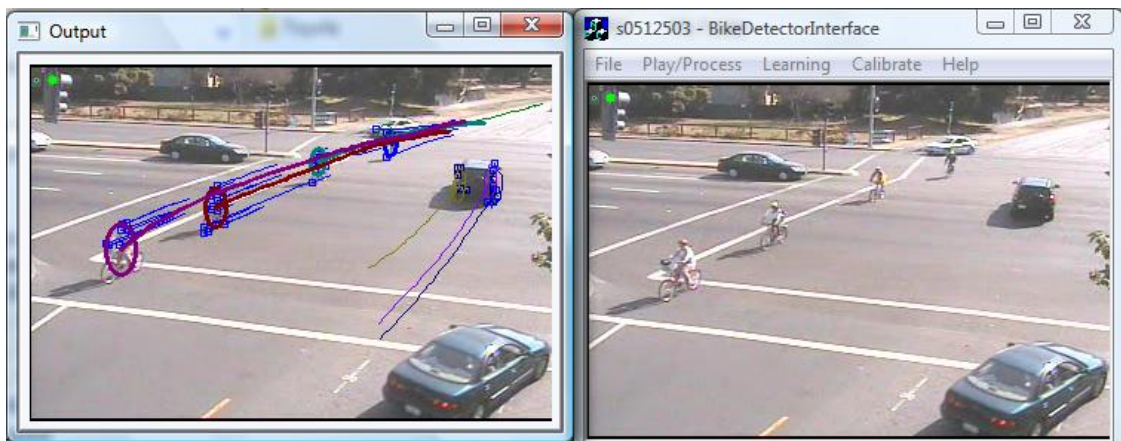


Figure 6-2 Bicyclists tracking results

The error of the crossing time since the beginning of the green signal phase is less than 100 ms. The trajectories may contain some errors from the conversion from image to world coordinates (calibration errors) and/or tracking errors due to poor image resolution. Such an error can be up to 1 m within the intersection box and even larger behind the stop bar on the far side of the intersection, where the image resolution is poor. Therefore, determining the crossing starting time from the trajectory alone may result in errors, but this is not a major concern because we are primarily interested in the timing

relative to the green signal onset.

The output of the video image processing software, as applied using the playback tool, is a plot of the location of the bicyclist crossing the intersection versus time. The traffic signal phase information is used to establish the reference time. The start of the green phase is defined as time zero for each crossing trajectory, and the times of the yellow and red onsets are marked by vertical yellow and red lines on the plots. An example of the plot that is generated for each bicyclist crossing is shown in Figure 6-3.

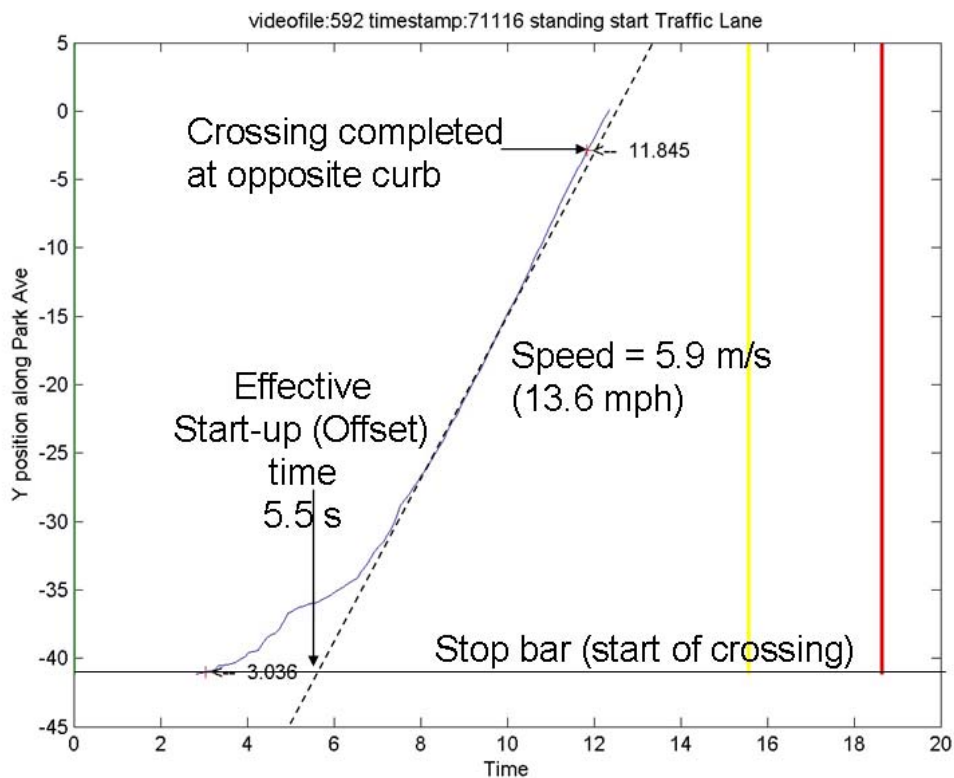


Figure 6-3 Example of extraction of needed values for standing start from bicyclist trajectory plot

Figure 6-3 shows a standing start by a bicyclist, with the trajectory indicated by the solid blue curve. The stop bar on the far side of the intersection where the bicyclist had to wait before starting to cross is at location -41 m and the curb line on the near side of

the intersection, representing completion of the crossing, is at -3 m. The vertical yellow line indicates the yellow onset after about 15.5 s of green, and the vertical red line indicates the red onset 3 s later.

This bicyclist started moving about 3 s after the signal changed to green, and reached full speed of 5.9 m/s (defined graphically by the tangent to the trajectory curve) about 7 s after the start of the green phase. In order to provide a single concise representation of standing starts that can be generalized to intersections of arbitrary width; we have defined the “offset time” as the intersection of the full crossing speed tangent with the line representing the start of the crossing. In this example, that offset time is 5.5 s, and this complete crossing maneuver is described most concisely in terms of that offset time and the final crossing speed of 5.9 m/s (13.6 mph).

6.3 Analysis of Bicyclist Crossing Data

The data collection system was used throughout the daylight hours for two days in Palo Alto and three days in Berkeley. In Palo Alto, 320 bicyclist crossings were observed, but only 265 of these produced usable data, with complete enough trajectories for analysis. Of the usable crossings, 188 were from standing starts and 77 were from rolling starts, the latter indicating that the signal was already green when the bicyclist arrived at the stop bar to begin the crossing. In Berkeley, 439 usable crossings were recorded in both directions of travel, with 279 standing starts and 160 rolling starts. The Berkeley site had more unusable crossings because of difficult lighting conditions in the late afternoon and a wider range of complicated bicyclist maneuvers that were not readily

classifiable.

The standing start and rolling start cases were analyzed separately because of their different implications for signal timing. The standing starts are important for determining the length of the minimum green interval, and are therefore the primary focus of the analysis. The rolling starts could be considered as inputs to decisions about the length of the yellow and all-red intervals, but Caltrans believes that those must be determined based on other considerations. The key attribute of the rolling start crossings is the crossing speed, the distribution of which is shown in Figure 6-4. The superimposed lines on the cumulative distribution plot show the 10%, 20% and 50%ile values, indicating potential criteria for consideration in signal timing. It was interesting to note that the values of these key percentile values were approximately 4 mph faster for the Palo Alto bicyclists (primarily young adult commuters) than for the more diverse Berkeley bicyclists. There was a further contrast between the speeds of the eastbound and westbound bicyclists at the Berkeley intersection based on the grade on Russell Street. Westbound bicyclists approached the intersection on a -3.4% grade while eastbound bicyclists approached on a +2.5% grade, leading to a difference of 2 mph in their median speeds, 4.5 mph in their 80%ile speeds and 5.5 mph in their 90%ile speeds.

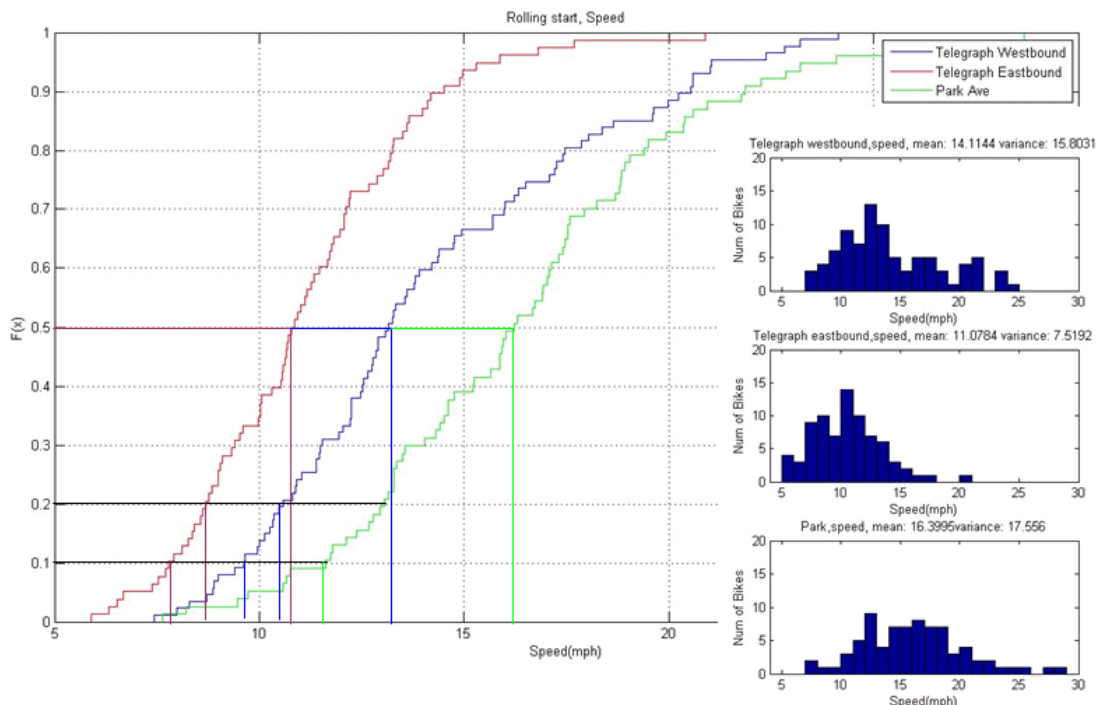


Figure 6-4 Cumulative distributions and histograms of speed of rolling start bicycles (mph)

The statistics of the standing start bicyclist crossings are most important for selecting the length of the green interval. The distributions of the offset times and of final steady speed of crossing are shown in Figure 6-5 and Figure 6-6 respectively. A scatter plot of these samples for the Palo Alto bicyclists was generated to visualize potential correlations between these parameters. Fortunately, and somewhat surprisingly, there was virtually no correlation between them so that they can be treated as independent variables, simplifying the development of recommendations for signal timing.

Figure 6-5 shows a strong concentration of offset times in the range of 4 to 8 seconds for the Palo Alto bicyclists, which is considerably longer than the start-up times reported by Wachtel, et.al. in the same city, while the Berkeley bicyclists were concentrated in the range of 2 to 5 seconds. The median value was about 6.5 s, and the 80th and 90th %iles

were about 8.3 and 9.3 s respectively in Palo Alto, while the corresponding values were about 3 seconds less in Berkeley. These values are this large for three reasons:

- ◆ they are counted from the green onset, accounting for the time the bicyclists need to recognize that the signal has changed;
- ◆ at times of heavy bicycle traffic, the bicyclists are queued behind the stop bar and the bicyclists further back in the queue cannot start until the preceding bicyclists have moved out of their way, accounting for some of the larger sampled values; and
- ◆ they represent the intersection of the constant speed tangent curve with the starting location from Figure 6-4, rather than the actual start-up time. This final point is an important distinction, which is needed when the results from one intersection have to be applied to an intersection with a significantly different width. The offset time is the component of the crossing time that is independent of intersection width.

The faster offset times in Berkeley were initially somewhat surprising, particularly considering that their rolling start crossing speeds were significantly slower. Examination of the video data indicated that the Berkeley bicyclists were much more likely to start crossing before they had a green signal (which none of the Palo Alto bicyclists did). In the end, it appears that the Palo Alto bicyclists had to be considerably more cautious about starting their intersection crossing because they were crossing a wider street with heavier and much faster traffic and they had poorer visibility of that crossing traffic from their starting position. The offset times in Berkeley were

significantly different for the two directions of travel, with the westbound bicyclists having significantly lower mean and variance in their offset times. Closer inspection of the Berkeley intersection revealed that eastbound bicyclists had clearer visibility of the approaching cross traffic on the near side of Telegraph Ave. (southbound) most of the time because of a bus stop at the corner, encouraging them to start crossing when no cars were visible on Telegraph during its yellow phase, but that visibility could be blocked when a bus was stopped there, leading to some longer offset times as well.

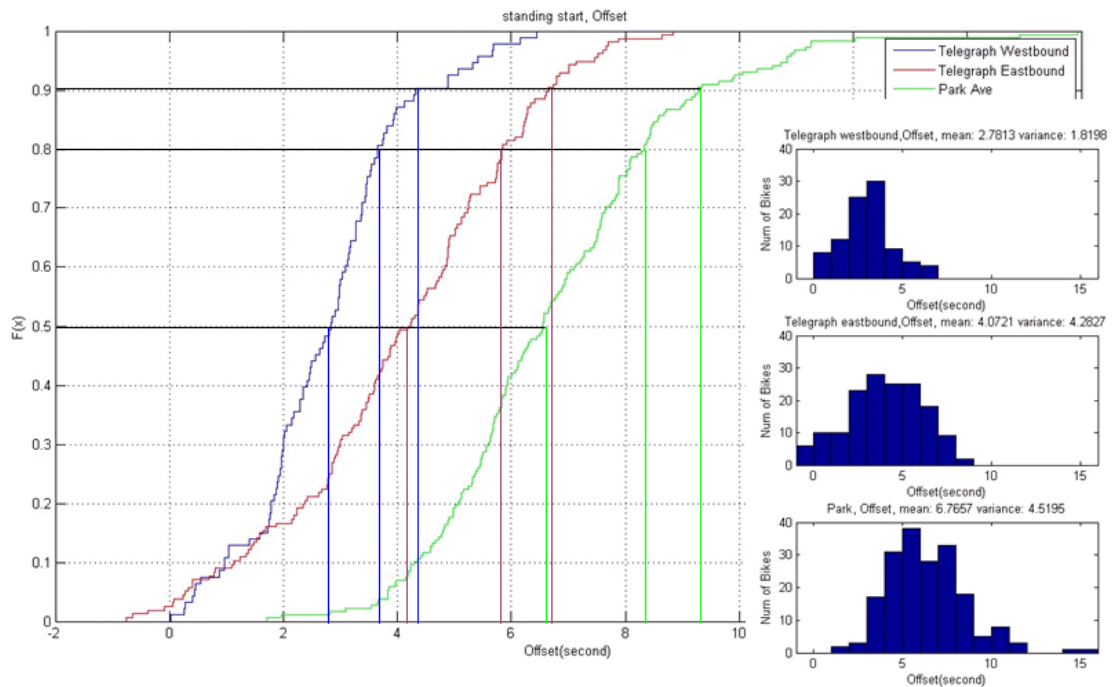


Figure 6-5 Cumulative distributions and histograms of offset times for standing start bicyclist crossings

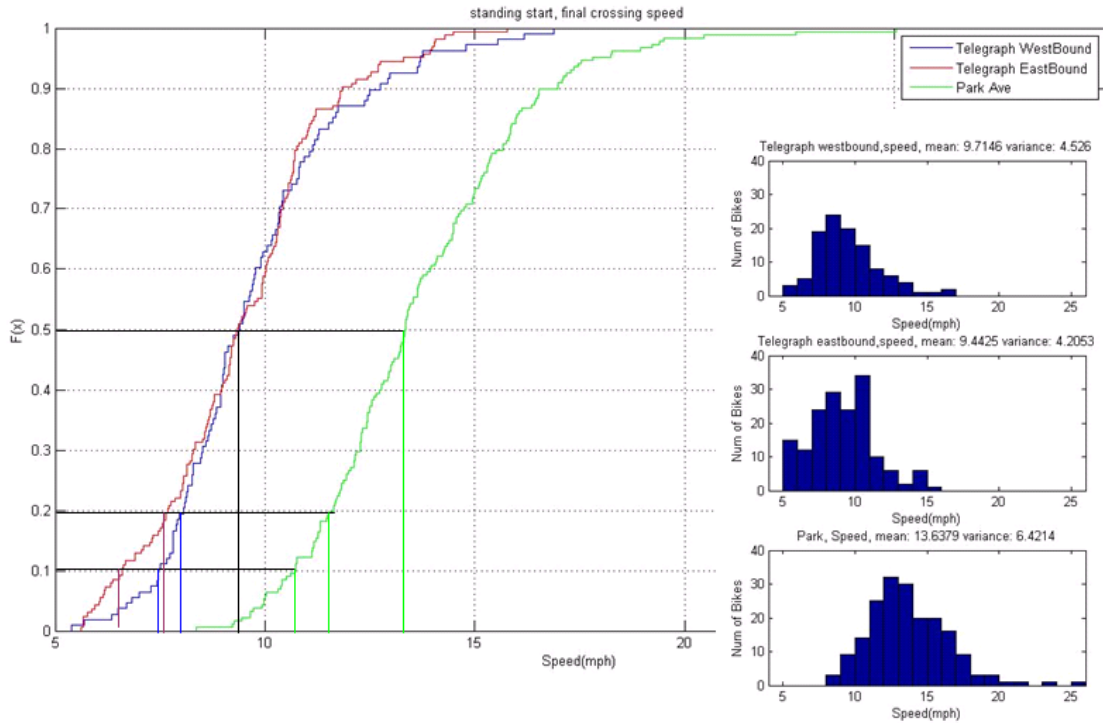


Figure 6-6 Cumulative distributions and histograms of final crossing speed for standing start bicyclist crossings (mph)

Figure 6-6 shows that the final crossing speeds of the Palo Alto bicyclists were clustered in the range of 10 to 18 mph, with only a few outliers at significantly higher speeds, while the Berkeley bicyclists were mainly clustered from 7 to 11 mph. The median speed in Palo Alto was 13.3 mph, with 10 and 20%ile values of about 10.5 and 11.5 mph respectively, which were about 4 mph faster than the corresponding values observed in Berkeley, an indication of the different demographic composition of the bicycling population at these two intersections. In [52] the 10%ile bicyclist speeds were described as 8 mph for casual adults and 6 mph for children. The speeds reported here are direct measurement of the actual final cruising speed rather than an average speed that includes some of the start-up transient. The final crossing speeds for the standing start

bicyclists in Berkeley were very similar for the two directions of travel, since the intersection is flat even though the approach blocks have significant grades.

As a cross-check on the independence of the offset time and final crossing speed measurements, we have also plotted the distributions of the total time that the bicyclists took to cross the 125 foot width of El Camino Real at Park Blvd from a standing start. These are shown in Figure 6-7, indicating a median crossing time in excess of 13 s, with the 80 and 90%iles at about 15 and 16.5 seconds respectively. These are all larger than the sum of the current minimum green plus yellow plus all-red intervals at this intersection ($7 + 3 + 1 = 11$ s). That minimum setting is only meeting the needs of the fastest 20% of the bicyclists.

Using the values of offset time and final crossing speed, the crossing time can be estimated as:

$$\text{Crossing Time} = \text{Offset time} + (\text{Width in ft./Crossing speed in mph}) * 0.68$$

Using the measured median values of offset time and crossing speed, the crossing time estimate is:

$$\text{Crossing time} = 6.5 + (125/13.3)*0.68 = 12.9 \text{ s (compared to 13.3 s measured in data)}$$

Similarly, focusing on the 20% slowest bicyclists, we use the 80%ile of offset time and 20%ile of crossing speed to estimate crossing time as:

$$\text{Crossing time} = 8.3 + (125/11.5)*0.68 = 15.7 \text{ s (compared to 15 s measured in data)}$$

Finally, for the 10% slowest bicyclists, we use the 90%ile of offset time and 10% of crossing speed to estimate crossing time as:

Crossing time = $9.3 + (125/10.5)*0.68 = 17.4$ s (compared to 16.5 s measured in data).

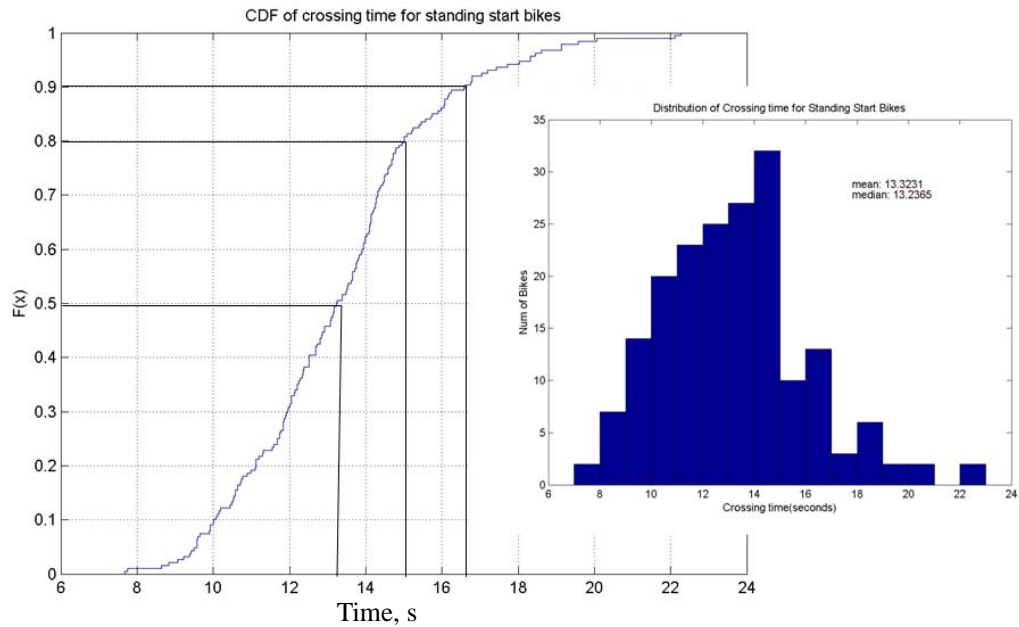


Figure 6-7 Distributions of total time for bicyclists to cross El Camino Real from a standing start

The video data include direct observations of the traffic signal cycles at the intersection at the times that bicyclists were crossing. Figure 6-8 shows the distribution of the duration of the green phase for Park Blvd., with each sample representing one bicyclist crossing. This signal is traffic actuated, rather than operating on a fixed cycle, and the green duration ranged from 10 to 37 s. Clearly most of the bicycle traffic occurs when there is also significant vehicular traffic triggering the detectors to extend the green phase beyond its minimum duration. The colors on the histogram indicate the signal phase when the bicyclist completed the crossing. Obviously, the shorter duration green cycles caused a significant proportion of the bicyclists to complete their crossing in yellow and red. This plot indicates the desirability of at least 13 s of green to enable

most bicyclists to complete their crossing in the green at this intersection. The analogous distribution for Russell Street is shown in Figure 6-9. Here, an even larger majority of the bicycle traffic was crossing when the green phase was more than adequate to meet their needs, and only a few of the bicyclists had to complete their crossing in the red when the green phase was only 7 s long.

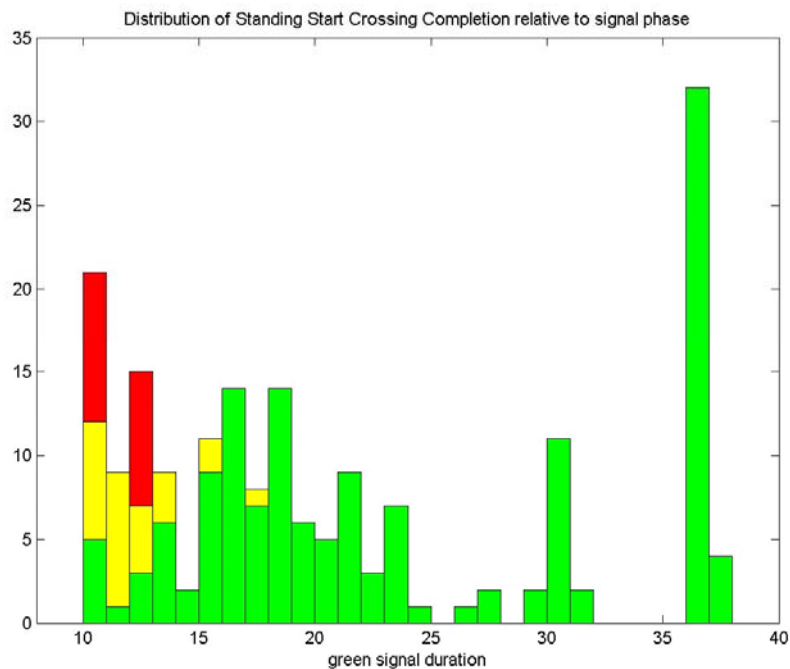


Figure 6-8 Distribution of duration of green phase (s) when bicyclists were crossing El Camino at Park Blvd.

It is also useful to understand how far into the signal phases the bicyclists are completing their crossings, which is illustrated in Figure 6-10 for Park Blvd. In this plot, time zero is the yellow onset and samples with positive times indicate crossings completed prior to the yellow onset. There is a 3 s yellow interval at this intersection, and the samples from -3 to 0 indicated a crossing completed in the yellow, while the samples for larger negative values indicate crossings completed in the red phase, as much

as 4 s after the red onset (3 seconds after El Camino traffic has received a green signal).

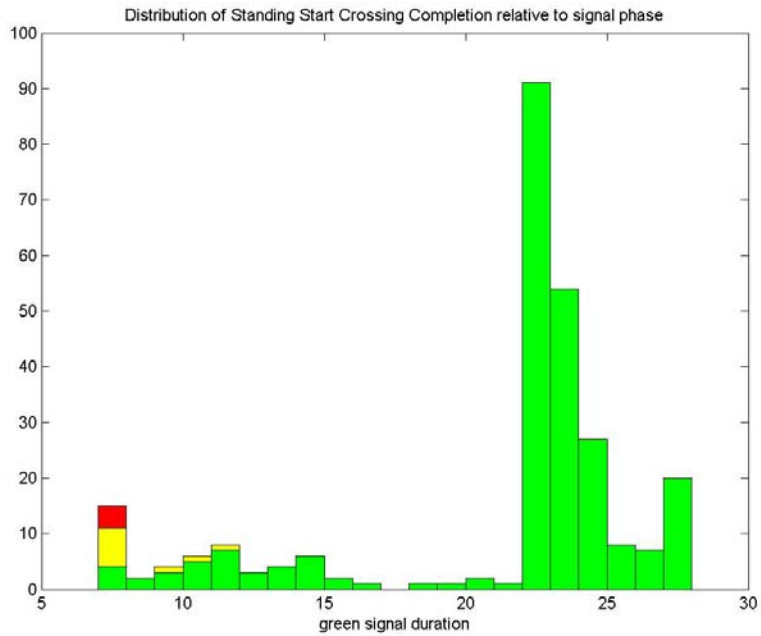


Figure 6-9 Distribution of duration of green phase (s) when bicyclists were crossing Telegraph at Russell St.

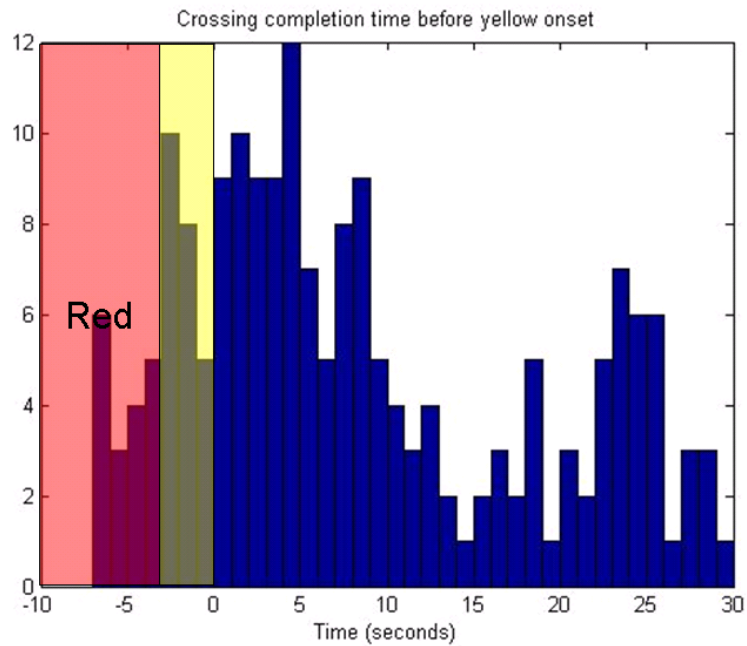


Figure 6-10 Completion times of bicyclist crossings from standing start, relative to yellow onset

6.4 Summary

In this study, a data set of bicyclist traffic video has been collected, and a vision aided accurate measurements of bicyclist crossing times has been performed. This video data reported here represent an unprecedentedly rich description of bicyclist intersection crossing behavior, particularly with respect to traffic signal phase. Based on the analyzed results from the vision based quantitative measurements, key findings from these observations were:

- Even within a relatively homogeneous population of bicyclists, there was a wide range of speeds and start-up offset times;
- The bicyclist speeds and offset times were not correlated with each other, so their statistics could be analyzed independently;
- Start-up times were significantly longer than reported in previous studies, but we believe that our data are more accurate representations of reality because they are based on the individual bicyclist trajectories rather than being derived indirectly from ensemble statistics with large variability. A start-up time of at least 9 seconds was needed to represent the 90%ile of bicyclists crossing a high-speed, high-density arterial with limited visibility, while about 6 seconds was needed to represent a comparable percentile of bicyclists crossing a medium-density arterial with moderate-speed traffic and better visibility;
- About 90% of the primarily young adult commuter bicyclists in Palo Alto reached a steady cruising speed of at least 16.8 km/h during their crossing, while the comparable statistic for the more diverse bicyclists in Berkeley was 11.2 km/h.

This steady cruising speed is the parameter that has to be used to extrapolate the data to other intersections with different street widths;

- Grades on the roadways approaching the intersection significantly influence the speed of bicyclists making rolling approaches, particularly in the upper speed percentiles. An average grade around 3% led to a difference of 3 km/h in the median speed and 8.8 km/h in the 90%ile speed of the bicyclists approaching from opposite directions. Higher crowns on the surface of the street being crossed may also increase the offset time and the final crossing speed; and
- The contrasts between the Palo Alto and Berkeley data indicate the sensitivity of bicyclist crossing time statistics to differences in the bicycling population and the physical and operational characteristics of the intersection being crossed. Considerations such as the speed and density of the crossing traffic, the crown of the road surface and the ability of the bicyclists to see the cross traffic from their starting position can have a significant influence on the time needed to traverse an intersection.

This study provides a solid foundation for defining bicycle signal timing.

7. Conclusions and Future Work

This chapter provides a summary and conclusions of the dissertation, as well as possible future work.

7.1 Conclusions

This dissertation describes a vision and GPS equipped mobile traffic surveillance system that was developed to specifically obtain localized traffic conditions. It involves the development of a unique OODV system, the derivation of the mathematical description of the OODV projection model, the OODV based vehicle detection algorithm, the generation of the localized traffic map, the methodology of generating the traffic flow from localized density, the explanation of the relationship between the mobile traffic data and loop sensor data, and the experiment verification about this relationship.

The original goal of this research is to make use of mobile computer vision techniques to provide a direct measurement to the spatial and temporal varying traffic data, as a complement to existing traffic monitoring systems, such as a loop detector sensor system. We want to quantitatively prove and show that the mobile traffic data obtained from the mobile traffic surveillance system have some internal relationship with the loop data and the integration of the loop data and mobile traffic data will potentially provide a better understanding of the traffic conditions. The developed system has successfully met this goal and has been demonstrated for the initial trip testing between Riverside and Corona, in Southern California, United States along State Route-91.

The developed OODV system, in comparison against the rectilinear vision system, provides a much wider 360 degrees orthogonal view and in comparison against the other catadioptric vision systems, and eliminates the need of warping a polar-coordinate image to a perspective view. This OODV system has been theoretically and experimentally proven to approximate a perspective view with the view point locating at the center of the mirror.

Based on this unique OODV, a roadway traffic surveillance system was designed and implemented. It consists of three major components: 1) A GPS stamped roadway traffic data collection technique; 2) Post video processing that includes automatic vehicle detection/tracking with the ability to correct using interactive tools; and, 3) A traffic parameter (localized density) estimation process. Based on initial experimentation, the designed OODV system reports a 83.7% detection rate under the dynamic freeway environment. To achieve a higher detection and tracking accuracy, a human interactive detection module also has been implemented and added as an option to the OODV-based roadway traffic surveillance system.

The relative positions of detected vehicles in the surrounding environment have been represented in the vehicle coordinate system, to generate the surrounding traffic map, which contains localized traffic information.

A theoretical explanation about the relationship between loop data and mobile data is introduced. The traffic parameters are re-defined in the time-space diagram and thus a calculated method of traffic flow is proposed based on the localized traffic density and velocity. Experiments have verified this explanation by showing the traffic flow

calculated from the measurement of localized density stratify the flow distribution pattern measured directly from loop data.

In addition, several other conclusions from this dissertation can be made:

1). A vision based portable loop fault detection tool has been developed for the systematic loop fault detection at the control cabinet level. The experimental results show that the vehicle tracking technologies performs well as the baseline measurement in the portable tool for the loop fault diagnosis.

2). Vision techniques have been more and more involved in the safety applications in the ITS area. A quantitative bicycle travel time measurement system has been developed, to support the efforts in improving the traffic signal timing for the needs of bicyclists. The processed data have yielded cumulative distributions of the crossing speeds of bicyclists who did not have to stop at the intersection and the start-up times and final crossing speeds of the bicyclists who had to cross from a standing start. This study provides a foundation in recommendation of minimal green signal time in terms of safety purpose.

7.2 Publications Resulting From This Research

- [1]. Meng Cao, Anh Vu, Matthew Barth, “A Novel Omni-Directional Vision Sensing Technique for Traffic Surveillance”, *the 10th International IEEE Conference on Intelligent Transportation Systems*, Seattle, Sept. 2007
- [2]. XY Lu, Zu Kim, Meng Cao, Pravin Varaiya, “Portable Loop Fault Diagnosis Tool Development for Control Cabinet Level”, *2009 Annual Meeting of the Transportation Research Board*, January 10-15, in Washington, D.C., 2009

- [3]. Steven Shladover, Zu Kim, Meng Cao, M. Sharafsaleh, “Bicyclist Intersection Crossing Times: Quantitative Measurement Relative to Signal Phases”, *2009 Annual Meeting of the Transportation Research Board*, January 10-15, in Washington, D.C., 2009
- [4]. Anh Vu, Meng Cao, Matthew Barth, “Depth Map Reconstruction using Dual Collinear Optical Axis Catadioptric Cameras”, submitted to *IEEE Intelligent Vehicles Symposium, 2009*

7.3 Future Work

Thus far, only a few experimental spatial and temporal roadway traffic data have been collected, within limited time and region. Further long-term and extensive traffic videos should be collected and processed in future research. In this dissertation, the videos are processed in a combined automatic and interactive method, in order to ensure error free detection. This requires a huge operational work of a human operator. Thus, more robust and reliable computer vision techniques are worthy to be explored to provide complete automatic processing. In addition, with the collection of new roadway traffic data, the quantitative relationship and fusion model should be determined between the loop data and mobile data.

Bibliography

- [1] M. Barth, E. Johnston, and R. Tadi, "Using GPS Technology to Relate Macroscopic and Microscopic Traffic Parameters", *Transportation Research Record*, No. 1520, pp. 89- 96, Transportation Research Board, National Academy of Science, 1996.
- [2] Young-Kee Jung and Yo-Sung Ho, "Traffic Parameter Extraction using Video-based Vehicle Tracking", *IEEE/IEEJ/JSAI International Conference on Intelligent Transportation Systems*, 1999.
- [3] [http:// pems.eecs.berkeley.edu/Public/index.html](http://pems.eecs.berkeley.edu/Public/index.html)
- [4] Jun-Wei Hsieh, et al, "Automatic traffic surveillance system for vehicle tracking and classification", *IEEE Transaction on Intelligent Transportation Systems*, Vol. 7, pp.175-187, June 2006.
- [5] T. Gandhi and M. Trivedi, "Parametric egomotion estimation for vehicle surround analysis using omni-directional camera", *Machine Vision Applications*, Vol. 16, No. 2, pp 85-95.2005
- [6] A. Broggi, "Vision-based driving assistance", *Intelligent Systems and Their Applications, IEEE*, Volume 13, Issue 6, 1998.
- [7] A. Kuehnle, "Symmetry-based recognition for vehicle rears," *Pattern Recognition Letters*, vol. 12, pp. 249–258, 1991.
- [8] Guo, T. Fraichard, M. Xie, and C. Laugier, "Color modeling by spherical influence field in sensing driving environment," *IEEE Intelligent Vehicle Symposium*, pp. 249–254, 2000.
- [9] C. Tzomakas and W. Seelen, "Vehicle detection in traffic scenes using shadows," Tech. Rep. 98-06, Institute fur neuroinformatik, Ruhr-universitat, Bochum, Germany, 1998.
- [10] M. Bertozzi, A. Broggi, and S. Castelluccio, "A real-time oriented system for vehicle detection," *Journal of Systems Architecture*, pp. 317–325, 1997.
- [11] C. Goerick, N. Detlev and M. Werner, "Artificial neural networks in realtime car detection and tracking applications," *Pattern Recognition Letters*, vol. 17, pp.335–343, 1996.

- [12] T. Kalinke, C. Tzomakas und W. v. Seelen, "A texture-based object detection and an adaptive model-based classification," *IEEE International Conference on Intelligent Vehicles*, pp. 143–148, 1998.
- [13] R. Cucchiara and M. Piccardi, "Vehicle detection under day and night illumination," *International ICSC Symposium on Intelligent Industrial Automation*, 1999.
- [14] Giachetti, M. Campani, and V. Torre, "The use of optical flow for road navigation," *IEEE transactions on robotics and automation*, vol. 14, no.1, pp. 34–48, 1998.
- [15] J. Weng, N. Ahuja, and T. Huang, "Matching two perspective views," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 14, pp. 806– 825, 1992.
- [16] D. Koller, N. Heinze, and H. Nagel, "Algorithmic characterization of vehicle trajectories from image sequence by motion verbs," *IEEE International Conference on Computer Vision and Pattern Recognition*, pp. 90– 95, 1991.
- [17] B. Heisele and W. Ritter, "Obstacle detection based on color blob flow," *IEEE Intelligent vehicle symposium*, pp. 282–286, 1995.
- [18] Benschraier, M. Bertozzi, A. Broggi, P. Miche, S. Mousset and G. Moulminet, "A cooperative approach to vision-based vehicle detection," *IEEE Intelligent Transportation Systems*, pp. 209–214, 2001.
- [19] A.N.Rajagopalan and R.Chellappa, "Vehicle Detection and tracking in video", *International Conference on Image Processing,, Volume 1*, pp 351 – 354, 2000
- [20] N. Matthews, P. An, D. Charnley, and C. Harris, "Vehicle detection and recognition in greyscale imagery," *Control Engineering Practice*, vol. 4, pp. 473–479, 1996.
- [21] He Zhiwei, Liu Jinlin, Li Peihong, "New Method of Background Update for Video-based Vehicle Detection", *IEEE Intelligent Transportation Systems Conference*, 2004
- [22] Lei Xie, Guangxi Zhu, Yuqi Wang, Haixiang Xu, Zhenming Zhang, "Robust Vehicles Extraction in a Video-based Intelligent Transportation Systems", *International Conference on Communications, Circuits and Systems*, 2005
- [23] Alan J.Lipton, Hironobu Fujiyoshi, Raju S Patil, "Moving Target Classification and Tracking from Real-time Video", *Fourth IEEE Workshop on Applications of Computer Vision*, 1998

- [24]Gangyi Jiang, Shengnan Wang, Mei Yu, Tae-Young Chio, Yong-Deak KIM, “New method of Vision Based vehicle detection and tracking in complicated background”, *IEEE Region 10 Conference*, 2004.
- [25]Mao-Chi Huang and Shwu-Huey Yen, “A real-time and color-based Computer Vision for Traffic Monitoring System”. *IEEE International Conference on Multimedia and Expo (ICME)*, 2004.
- [26]Xin Li, XiaoCao Yao, Yi L.Murphey, Robert Karlsen and Grant Gerhart, “A real-time Vehicle Detection and Tracking System in Outdoor Traffic Scenes”. *Proceedings of the 17th International Conference on Pattern Recognition (ICPR)*, 2004.
- [27]Koller, D, Weber, J, Huang, T, Malik, J, Ogasawara, G, Rao, B, Russell, “Towards robust automatic traffic scene analysis in real-time”. *ICPR*, Israel, vol 1, pp 126-131, 1994.
- [28]Kai She, George Bebis, Haisong Gu and Ronald Miller, “Vehicle Tracking Using on-line Fusion of Color and Shape Features”, *IEEE Intelligent Transportation System Conference*, 2004.
- [29]Jianguang Lou, Tieniu Tan, Weiming Hu, Hao Yang and Steven J. Maybank, “3D Model-Based Vehicle Tracking”, *IEEE Transactions on image processing*, vol.14, 2005.
- [30]D.Koller, K.Daniilidis and H.H.Nagel, “Model-based object tracking in monocular image sequences of road traffic scenes”, *Int. J.Comput. Vis.*, vol.10, 1993.
- [31]D.Koller, J.Weber and J.Malik, “Efficient and robust vehicle localization”, *Proc. 3rd Eur.Conf. Computer Vision*, 1994.
- [32]H. Ishiguro, “Development of Low-Cost Compact Omnidirectional Vision Sensors and their applications”, *Proc. Int. Conf. Information Systems, Analysis and Synthesis*, pp.433-39, 1998.
- [33]S.Baker and S.Nayar. “A theory of catadioptric image formation”, *Proc. Int. Conference on Computer Vision*, pp. 35-42,1998.
- [34]R.Hicks, “Reflective surfaces as computational sensors”, *Proceedings of the 2nd Workwhop on Perception of Mobile Agents, CVPR99*, pp. 82-86,1999.
- [35]R.Hicks and R.Bajcsy, “Catadioptric sensors that approximate wide-angle perspective projections”, *IEEE Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 545 – 551, 2000.

- [36] T. Gandhi and M. Trivedi, "Vehicle Surround Capture: Survey of Techniques and a Novel Omni-Video-Based Approach for Dynamic Panoramic Surround Maps" *IEEE Transaction. on Intelligent Transportation Systems*, vol. 7, Sept. 2006.
- [37] W. Kruger, "Robust real-time ground plane motion compensation from a moving vehicle", *Machine Vision and Application*, 1999.
- [38] M. Bertozzi and A. Broggi, "GOLD: A Parallel Real-Time Stereo Vision System for Generic Obstacle and Lane Detection", *IEEE Transactions on Image Processing*, vol. 7, Jan, 1998.
- [39] W. Kruger, "Robust real-time ground plane motion compensation from a moving vehicle", *Machine Vision and Application*, 1999.
- [40] Kim, Z., "Realtime Object Tracking based on Dynamic Feature Grouping with Background Subtraction", *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, 2008.
- [41] Kim, Zu, "Geometry of Vanishing Points and its Application to External Calibration and Realtime Pose Estimation", *Institute of Transportation Studies, Research Reports*, Paper UCB-ITS-RR-2006-5, July 1, 2006.
- [42] Lu, X. Y., Varaiya, P., Horowitz, R. and Palen, J., "Faulty loop data analysis/correction and loop fault detection", *15th ITS World Congress*, November 16-20, 2008.
- [43] Kell, J., Fullerton, I., and Mills, M., *Traffic Detector Handbook*, Second Edition, Federal Highway Administration, Washington, DC, 1990.
- [44] Ingram, J., *The Inductive Loop Vehicle Detector: Installation Acceptance Criteria and Maintenance Techniques*, California Department of Transportation, Sacramento, CA, 1976.
- [45] Jacobson, L, Nihan, N., and Bender, J., "Detecting Erroneous Loop Detector Data in a Freeway Traffic Management System", *Transportation Research Record 1287*, TRB, Washington, DC, pp 151-166, 1990.
- [46] Cleghorn, D., Hall, F., and Garbuio, D., "Improved Data Screening Techniques for Freeway Traffic Management Systems", *Transportation Research Record 1320*, TRB, Washington, DC, pp 17-31, 1991
- [47] Nihan, N., "Aid to Determining Freeway Metering Rates and Detecting Loop Errors", *Journal of Transportation Engineering*, vol 123, no 6, ASCE, pp 454-458, 1997.

- [48]Chen, L., and May, A., “Traffic Detector Errors and Diagnostics”, *Transportation Research Record 1132, TRB*, Washington, DC, pp 82-93, 1987.
- [49]Karl Petty, “Freeway Service Patrol (FSP) 1.1: The Analysis Software for the FSP Project”, California PATH Research Report, UCB-ITS-PRR-95-20.
- [50]Skabardonis, A., Petty, K., Noeimi, H., Rydzewski, D. and Varaiya, P. “I-880 Field Experiment: Data-Base Development and Incident Delay Estimation Procedures”, *Transportation Research Record, TRB*, pp 204-212, 1996.
- [51]Coifman, B., “Using Dual Loop Speed Traps to Identify Detector Errors”, *78th TRB Annual Meeting*, Washington, DC., Jan. 1999.
- [52]Wachtel, A., J. Forester and D. Pelz, “Signal Clearance Timing for Bicyclists”, *ITE Journal*, pp. 38 – 45,1995.
- [53]Rubins, D.I. and S. Handy, “Times of Bicycle Crossings: Case Study of Davis, California”, *Transportation Research Record 1939*, Washington DC, pp. 22-27, 2005.
- [54]Bertozzi, M., Broggi,A., Fascioli, A., “Vision-based intelligent vehicles:State of the art and perspectives”, *Robot. Auton.Syst.* vol.32, no.1, pp.1-16, Jul.2000.
- [55]Kastrinaki, V, Zervakis, M, “A survey of video processing techniques for traffic applications”, *Image and Vision Computing*, 2003.
- [56]Gandhi,T, Trivedi,M, “Vehicle Surround Capture: Survey of Techniques and a Novel Omni-Video-Based Approach for Dynamic Panoramic Surround Maps”, *IEEE Transactions on Intelligent Transportation System*, Vol.7, No.3, 2006.
- [57]Mao-Chi Huang and Shwu-Huey Yen, “A real-time and color-based Computer Vision for Traffic Monitoring System”. *IEEE International Conference on Multimedia and Expo (ICME)*, 2004.
- [58]Xin Li, XiaoCao Yao, Yi L.Murphey, Robert Karlsen and Grant Gerhart, “A real-time Vehicle Detection and Tracking System in Outdoor Traffic Scenes”. *Proceedings of the 17th International Conference on Pattern Recognition (ICPR)*, 2004.
- [59]Koller, D, Weber, J, Huang, T, Malik, J, Ogasawara, G, Rao, B, Russell, “Towards robust automatic traffic scene analysis in real-time”. *ICPR*, Israel, Vol 1, pp 126-131, 1994.
- [60]Kai She, George Bebis, Haisong Gu and Ronald Miller, “Vehicle Tracking Using on-line Fusion of Color and Shape Features”, *IEEE Intelligent Transportation System Conference*, 2004.

- [61] Jianguang Lou, Tieniu Tan, Weiming Hu, Hao Yang and Steven J. Maybank, "3D Model-Based Vehicle Tracking", *IEEE Transactions on image processing*, VOL.14, 2005.
- [62] Z. Kim, "Robust Lane Detection and Tracking in Challenging Scenarios," *IEEE Trans. on Intelligent Transportation Systems*, vol. 9, no. 1, pp. 16-26, 2008.
- [63] XY Lu, Zu Kim, Meng Cao, Pravin Varaiya, "Portable Loop Fault Diagnosis Tool Development for Control Cabinet Level", *2009 Annual Meeting of the Transportation Research Board*, January 10-15, in Washington, D.C., 2009.
- [64] Steven Shladover, Zu Kim, Meng Cao, M. Sharafsaleh, "Bicyclist Intersection Crossing Times: Quantitative Measurement Relative to Signal Phase", *Proc. Transportation Research Board*, 2009
- [65] http://www.vision.caltech.edu/bouguetj/calib_doc/
- [66] C.F. Daganzo, "Fundamentals of Transportation and Traffic Operations", Pergamon, ISBN: 0080427855

Appendix A

Derivation of Eqn.3-1

For the pin-hole model, as illustrated in Figure 3-3 , the ray starts from the view point of the camera; goes through one point in the CCD array; hits the point (x, y) on the mirror and then reflects; the reflected ray intersects the reference plane. We define u as the CCD distance from the hitting point to the center of CCD, and $d(u)$ as the corresponding real world distance between the center of the reference plane and the hitting point of the reflected ray and reference plane. $\alpha, \beta, \gamma, \theta$ are also defined as in the Figure 3-3.

The goal is to find a mathematical description for the mirror surface so that u and $d(u)$ satisfy: $d(u) = a \times u$, where a is a constant.

The following equations can be derived from the geometrical relationship in Figure 3-3:

$$\tan \beta = \frac{dy}{dx} \quad (\text{Eqn. A-1})$$

$$\tan \alpha = \frac{y-t}{x} \quad (\text{Eqn. A-2})$$

$$\tan \theta = \frac{y}{d(u) - x} \quad (\text{Eqn. A-3})$$

$$\therefore 2 \times \beta = \alpha - \theta$$

$$\therefore \tan(2 \times \beta) = \tan(\alpha - \theta) \quad (\text{Eqn. A-4})$$

$$\Leftrightarrow \frac{2 \times \tan(2 \times \beta)}{1 - \tan^2(\beta)} \quad (\text{Eqn. A-5})$$

Substitute (Eqn. A-1), (Eqn. A-2), (Eqn. A-3) into (Eqn. A-5), we have:

$$\frac{2y'}{1-y'^2} = \frac{\frac{y-t}{x} - \frac{y}{d(u)-x}}{1 + \frac{y-t}{x} \times \frac{y}{d(u)-x}} \quad (\text{Eqn. A-6})$$

$$\text{Because } \frac{u}{f} = \frac{x}{y-t} \Rightarrow u = \frac{fx}{y-t} \Rightarrow d(u) = au = \frac{afx}{y-t} \quad (\text{Eqn. A-7})$$

Substitute (Eqn. A-7) into (Eqn. A-6) and simplify, we have

$$\frac{2y'}{1-y'^2} = \frac{-2xy^2 + (afx + 3xt)y - (aft + t^2)x}{y^3 - 2ty^2 + (t^2 - x^2)y + (af + t)x^2} \quad (\text{Eqn.A-8})$$

Eqn. A-8 mathematically describes the surface shape of the mirror, however for such a differential equation; there is not a close form solution. The only way to create such a shape is by using the numerical method.

$$\text{We define } g = \frac{-2xy^2 + (afx + 3xt)y - (aft + t^2)x}{y^3 - 2ty^2 + (t^2 - x^2)y + (af + t)x^2} \text{ and } y' = \frac{dy}{dx} \approx \frac{\Delta y}{\Delta x}, \text{ so Eqn.D-8}$$

became: $\frac{2y'}{1-y'^2} = g$ with initial condition $x = 0, y(0) = cont$, so:

$$g(\Delta y)^2 + 2\Delta x \Delta y - g(\Delta x)^2 = 0$$

$$\Rightarrow \Delta y = \frac{-2\Delta x + \sqrt{4(\Delta x)^2 + 4g^2(\Delta x)^2}}{2g} = \frac{(\sqrt{1+g^2} - 1)\Delta x}{g} \quad (\text{Eqn.A-9})$$

Appendix B

Least Square Error Line Fitting

Suppose the signal model is $s[n] = an + b$, where a and b are the parameters to be estimated. Here I tried to minimize the squared difference between the given data $x[n]$

and the assumed signal data. The Least Square error criterion means minimizing the following:

$$J(\theta) = \sum_{n=1}^N (x[n] - s[n])^2$$

Here the parameters to be estimated are vectors as $\theta = (a, b)^T$, so we have

$$s = H\theta$$

where H is the know $N \times 2$ matrix. So we have:

$$\begin{aligned} J(\theta) &= \sum_{n=1}^N (x[n] - s[n])^2 = (x - H\theta)^T (x - H\theta) \\ &= x^T x - x^T H\theta - \theta^T H^T x + \theta^T H^T H\theta = x^T x - 2x^T H\theta + \theta^T H^T H\theta \end{aligned}$$

The gradient is:

$$\frac{\partial J(\theta)}{\partial \theta} = -2H^T x + 2H^T H\theta = 0$$

So we have:

$$\hat{\theta} = (H^T H)^{-1} H^T x$$