Observations of lane changing patterns on an uphill expressway

by

Anthony David Patire

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

Doctor of Philosophy

in

Engineering – Civil and Environmental Engineering

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Michael Cassidy, Chair
Professor Carlos Daganzo
Professor Avideh Zakhor

Fall 2010
Observations of lane changing patterns on an uphill expressway

Copyright © 2010

by

Anthony David Patire
Abstract

Observations of lane changing patterns on an uphill expressway

by

Anthony David Patire

Doctor of Philosophy in Engineering – Civil and Environmental Engineering

University of California, Berkeley

Professor Michael Cassidy, Chair

A mechanism is unveiled by which traffic congestion forms on a 3-lane, uphill expressway segment, and causes reductions in output flow. Vehicular lane-changing (LC) is key to the mechanism, particularly LC induced by speed disturbances (SDs) that periodically arise in the expressway’s median and center lanes. Early in the rush, when flow was relatively low in the shoulder lane, drivers readily migrated toward that lane to escape the oncoming SDs. The shoulder lane thus acted as a “release valve” for the high vehicular accumulations created by the SDs, such that forced vehicular decelerations were short-lived. The release valve failed only later in the rush, when flow increased in the shoulder lane in response to rising demand. LC induced by the SDs thereafter became disruptive: the decelerations they imposed spread laterally, and a persistent queue formed in all lanes. Long-run output flow dropped each day by 4% to 11% once the queue engulfed the base of the incline, and impeded vehicle ascent.

The more conspicuous details of this mechanism were observed in loop detector data measured over many days at the site, and are consistent with observations previously made at other sites. More subtle details became visible by examining thousands of vehicle trajectories extracted on a single day from a series of eleven roadside video cameras. (Video processing tools were developed and used for this purpose.) Many of the subtleties are compatible with an existing theory of multi-lane traffic. All of this suggests that the present findings can be generalized to other uphill expressway segments. Practical implications are discussed.
Contents

1 Introduction

1.1 Motivation ......................................................... 1
1.2 Scope ............................................................. 2
1.3 Technical Challenges ............................................ 2
1.4 Overview .......................................................... 3
  1.4.1 Summary of Congestion Stages ........................... 3
  1.4.2 Organization ................................................. 4

2 Literature Review

2.1 Traffic Operations ............................................... 6
  2.1.1 Uphill Bottlenecks ....................................... 6
  2.1.2 Speed Disturbances and Traffic Theory ................. 8
  2.1.3 Effects of Lane Changing ................................. 10
  2.1.4 Multi-lane Considerations ............................... 10

2.2 Video Processing ................................................ 11

3 Methods

3.1 The Site and its Data .......................................... 13
3.2 Video Tools ....................................................... 15
  3.2.1 Pre-processing ............................................ 15
3.2.2 Vehicle Re-identification ........................................... 18
3.2.3 Database Exploration .............................................. 23
3.2.4 Detailed Trajectories .............................................. 23

4 Empirical Evidence ...................................................... 25
  4.1 Chronic Macroscopic Features ...................................... 25
  4.2 Microscopic Spatio-temporal Inspection ............................. 28
    4.2.1 Stage One: Early rush ........................................ 30
    4.2.2 Stage Two: Persistent slowing ................................ 33
    4.2.3 Stage Three: Emergence of a single-pipe regime ............. 35
    4.2.4 Stage Four: Reduction in output flow ........................ 37

5 Discussion .............................................................. 39
  5.1 The Shoulder Lane as a Release Valve .............................. 40
  5.2 Comparisons with Theory .......................................... 41
  5.3 Implications for Traffic Management ............................... 42
  5.4 Future Work ........................................................ 43

Bibliography ............................................................ 45
# List of Figures

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Newell’s simplified car-following model</td>
<td>8</td>
</tr>
<tr>
<td>3.1</td>
<td>Tomei Expressway study site</td>
<td>14</td>
</tr>
<tr>
<td>3.2</td>
<td>Building a slit-scan image</td>
<td>16</td>
</tr>
<tr>
<td>3.3</td>
<td>The vid GUI</td>
<td>19</td>
</tr>
<tr>
<td>3.4</td>
<td>The tview GUI</td>
<td>21</td>
</tr>
<tr>
<td>3.5</td>
<td>Trajectories of the tview GUI</td>
<td>22</td>
</tr>
<tr>
<td>4.1</td>
<td>Time course of traffic states (upstream)</td>
<td>26</td>
</tr>
<tr>
<td>4.2</td>
<td>Time course of traffic states (downstream)</td>
<td>27</td>
</tr>
<tr>
<td>4.3</td>
<td>Velocity fields and (leftward) lane changes</td>
<td>29</td>
</tr>
<tr>
<td>4.4</td>
<td>Velocity fields and (rightward) lane changes</td>
<td>31</td>
</tr>
<tr>
<td>4.5</td>
<td>Cumulative oblique counts of vehicles (shoulder lane)</td>
<td>34</td>
</tr>
<tr>
<td>4.6</td>
<td>Cumulative oblique counts of vehicles (all lanes)</td>
<td>36</td>
</tr>
</tbody>
</table>
Acknowledgments

First, I would like to acknowledge my advisor, Michael Cassidy, for inspiring me to pursue a PhD and for supervising this work. I am indebted to him for his generosity and for making every effort when I needed help. It has been a pleasure to work with someone who takes such pride in his work.

Second, I would like to acknowledge Carlos Daganzo and Avideh Zakhor who have helped to shape this thesis through course work, discussion, and participating on my committee.

I would like to thank all of my colleagues, past and present, in the Transportation Program here at UC Berkeley. Of those who came before me, I am particularly indebted to Sue Ahn for countless discussions. Over the years, I have come to cherish the friendship of everyone in my cohort—you know who you are. I am also grateful to all the students who extended their friendship despite having to endure my teaching.

My experience in Berkeley could not be so rich were it not for my greater community of friends in the Bay Area and abroad. From this tangled web of MIT alums, dot-comers, gamers, and current and former expats, it seems that although addresses change, no one ever really leaves. These personal connections have helped me through several eventful chapters in my life, and enabled me to emerge with optimism for the future.

Finally, I’d like to thank my wife, Akane, and my family for their steadfast support.

This work was funded by the Volvo Research and Educational Foundations. I thank Professor Masao Kuwahara of Tokyo University, and Professor Hirokazu Akahane of the Chiba Institute of Technology for generously making available the raw videos and loop detector data.
Chapter 1

Introduction

1.1 Motivation

The general aim of this research was to understand driver behavior that causes congestion to form and flow to drop on uphill sections of expressways. Despite notable studies in Japan, uphill bottlenecks are not yet fully understood. Note that these types of bottlenecks occur without influence from merges, diverges, or lane reductions. As a result, lane change maneuvers at these uphill bottlenecks are purely optional in that they are independent from route choice. Bottlenecks caused by tunnels or horizontal curves share similar features, and insights gained upon study of one type of bottleneck might be applicable to others.

Once queues form upstream of bottlenecks, the discharge rates diminish substantially from the sustained rates that are measured prior to queue formations. Typically this “capacity drop” is on the order of 10 percent; however, queue discharge rates can vary markedly (and can include periods when discharge is particularly low). Even a drop as small as 10% can cause significant increases in commuter delays, and thus efforts to avert capacity drops are worthwhile endeavors.

In addition, the excess pollution and gasoline consumption that accompany traffic congestion also exact a toll on the environment and public health. While these effects are measurable and significant, there are also less tangible impacts on quality of life. Ultimately, better understanding of traffic phenomena will lead to better management and control strategies. These in turn will improve the efficiency and comfort of ground-based transportation systems.
1.2 Scope

The scientific goal and main contribution was to unveil a mechanism for traffic congestion on a hill, and to explain how cumulative impacts of individual vehicle actions result in the observed macroscopic effects. Toward this end, we cooperated with Japanese partners who installed a series of video cameras along the length of one uphill section marked by recurrent congestion. This site is free of ramp junctions. Yet in the absence of traffic collisions or other incidents, queues engulfing all three lanes form at the site and cause output flow to diminish.

The evidence for our conclusions derives from an analysis of loop detector and video data taken from this site.

1.3 Technical Challenges

Our chief technical challenge was the processing of raw video data. A total of eleven cameras were used to monitor the 1.3 km segment of expressway including the downhill approach and about half of the uphill climb. Roughly 30 minutes of video were selected from each camera for detailed analysis. Each vehicle needed to be re-identified in each of the cameras. Since flows were on the order of 5000 vph, one half-hour corresponds to about 2500 vehicles, and a total of 27500 vehicle observations. It would have been infeasible to extract detailed trajectories in a manual fashion. Video processing tools were therefore developed to automate much of the trajectory extraction process. Additional tools were developed to interact with the trajectories and videos with a user-friendly interface.

A substantial effort was invested in tool-building. Although a simple plot of trajectories in one lane does express what vehicles do, it does not express why. We found it necessary to build an interactive plot in which the trajectories of all lanes were clickable and matched up with the videos. This interactive plot enabled one to view the vehicle trajectories in the context of surrounding traffic, including brake lights, turning signals, and lane changes. We believe the tools developed here are valuable in their own right, and further hope that these will serve as building blocks for future tool-building efforts.

The physical scale of traffic phenomena studied here occurs at spatial distances of 500 m in length, and about 4 m in width (for one lane). Such an extreme aspect ratio is not well captured in a single video camera. Multiple video cameras are needed, and the most crucial function is to re-identify the same vehicle in multiple cameras. Rough trajectories spatially sampled at 120 m are surprisingly informative when augmented with interactive tools as described above.
1.4 Overview

A mechanism is unveiled by which congestion forms on a 3-lane, uphill expressway segment, and causes reductions in output flow. Vehicular lane-changing (LC) is observed to be key to the mechanism, particularly LC induced by speed disturbances (SDs) that periodically arise in the expressway’s median and center lanes. Furthermore, the traffic condition in the expressway’s shoulder lane is also found to be important.

Early in the rush, when flow was relatively low in the shoulder lane, drivers readily migrated toward that lane to escape the oncoming SDs. The shoulder lane thus acted as a “release valve” for the high vehicular accumulations created by the SDs, such that forced vehicular decelerations were short-lived. During this time, the shoulder lane accommodated high rates of lane-changing. The release valve failed only later in the rush, when flow increased in the shoulder lane in response to rising demand.

The resulting higher flows in the shoulder lane impeded drivers’ attempts to maneuver around the SDs in the median lane that occurred thereafter. These attempts disrupted traffic and spread the excess accumulation laterally across all lanes. Vehicles in this slowed queue tended to accelerate near the midpoint of the hill, most notably in the median lane. Further lane-changing, this time toward the median, reduced flow in the shoulder and forced upstream vehicles in the median and center lanes to slow down, thus fixing the head of the queue in place. Long-run output flow dropped each day by 4% to 11% once the queue engulfed the base of the incline, and impeded vehicle ascent.

The more conspicuous details of this mechanism were observed in loop detector data measured over many days at the site, and reinforce more qualitative descriptions previously reported of similar sites (Koshi et al., 1992). More subtle details became visible by examining thousands of vehicle trajectories extracted on a single day from a series of eleven roadside video cameras. Congestion formed in stages and many of the subtleties are compatible with an existing theory of multi-lane traffic (Daganzo, 2002).

1.4.1 Summary of Congestion Stages

Details of traffic flow are examined for a 3-lane expressway in Japan where a downhill segment is followed by a steep, 2.4% incline of extended length. The site is free of ramp junctions. Yet in the absence of traffic collisions or other incidents, queues engulfing all three lanes can form at the site and cause long-run output flow to diminish by 4% to 11%. Vehicle trajectories extracted from video data on a single day unveil causal mechanisms, including the key roles played by vehicular lane-changing (LC).
The process of congestion formation and output-flow reduction is described in the following four stages.

The first stage was marked by disproportionately low flows in the shoulder lane, and a predominantly leftward pattern of LC toward that lane.\(^3\) Traffic was by-and-large freely flowing, although ascending traffic occasionally became dense in the median lane, and to a lesser degree in the center lane. High density clusters coalesced into periodic speed disturbances (SDs) and propagated backward (against the flow of vehicles) as stop-and-go waves. Low shoulder-lane flow enabled many drivers to migrate leftward to avoid the SDs with minimal disruption to traffic. These migrations were beneficial: they promptly mitigated the high vehicle accumulations and attendant slowing in the center and median lanes, such that most vehicles ascended the incline at high speeds. Thus, the shoulder lane functioned as a “release valve.” This, coupled with the increased utilization of the shoulder lane brought by the LC pattern resulted in high average output flows from the hill.

During the second stage, LC started to become baneful. Rightward LC from shoulder to center lane became more prevalent as demand increased and more drivers were forced to travel in the shoulder lane. Median-lane vehicles therefore became boxed in: these drivers could not maneuver around incipient SDs, such that the decelerations in the median lane became longer-lived.

In the third stage, LC became even more disruptive by virtue of the high shoulder-lane flow. Migrations into that lane continued, and spread the queue laterally to all lanes. Downstream of this queue, median-lane vehicles accelerated more quickly than those in the center and shoulder lanes, thus inducing rightward LC maneuvers toward the faster-moving lane. Near the end of this stage, migrations into the (heavily utilized) median lane disrupted flows and hastened the queue’s longitudinal expansion.

By the fourth stage, the queue had extended to about 1 km in length, and output flows were diminished. The downstream LC pattern was predominantly rightward toward the median lane as before, and frequent SDs were now observed within the queue. Average headways thereafter increased, yielding a significant reduction in output flow that persisted for the lifetime of the queue.

### 1.4.2 Organization

The following chapter reviews the literature in this field. Previously reported observations of uphill bottlenecks are described, and a brief review of classical traffic theory is presented. This serves as the context in which to discuss refinements that have been proposed to account for the phenomena of SDs, or stop-and-go waves. Empirical effects of lane changing, and a multi-lane theory of traffic are also described.

\(^3\)Japanese drive on the left, such that the median lane is the right-most lane.
Finally, a brief survey of video processing techniques informs a discussion of design decisions made to overcome the technical challenges mentioned above.

In Chapter 3, the site used in the present study and its instrumentation are described. After explaining the video pre-processing steps, two GUI tools are introduced. The first GUI allows a user to supervise the re-identification process. The second enables visualization of the trajectories together with the video. Empirical evidence of the mechanism described above is furnished in Chapter 4. Interpretation of the evidence, its compatibility with the above-cited traffic theory, and implications for managing congestion are discussed in Chapter 5.
Chapter 2

Literature Review

The undertaking of this thesis was multi-disciplinary. In addition to a study of expressway traffic, there was a design component required to acquire the necessary data. The first part of this literature review provides a background of traffic operations that informs the scientific inquiry. The second part of this literature review presents a survey of video processing techniques and our observations as to their applicability to our specific requirements.

2.1 Traffic Operations

In this chapter, we begin with a review of previous work on uphill bottlenecks. We continue with a discussion of efforts to refine existing traffic theory to explain the generation and propagation of speed disturbances. Next, we consider the effects of LC maneuvers, and possible linkages between LC and SDs. Finally, we consider issues of multi-lane traffic, and theory formulated to address them.

2.1.1 Uphill Bottlenecks

Bottlenecks that occur at uphill sections of expressways are not yet fully understood, although there have been notable studies in Japan. In the Japanese literature, these bottlenecks are referred to as “sag bottlenecks.” This nomenclature emphasizes the fact that “sag sections” are typically found at the bases of hills or in valleys. We choose to call these “uphill bottlenecks,” as is the convention in the English literature. Note that these types of bottlenecks occur without influence from merges, diverges, or lane reductions. As a result, lane change maneuvers at these uphill bottlenecks are purely optional in that they are independent from route choice.
According to Koshi et al. (1992), the activation of an uphill bottleneck requires the speed drop of a so-called “lead” vehicle somewhere within a large, dense platoon that is climbing the uphill segment in the median lane. Under these conditions, even small speed changes made by that leader will propagate backward to following vehicles. Typically, these speed fluctuations tend to amplify, forming a backward-moving SD that results in low vehicle speeds or even complete stops upstream. Reportedly, a small queue may begin to form in this way. If there is another dense platoon of vehicles that catches up to the first, then these drivers will tend to abandon the median lane in droves. This lane-changing behavior will spread the queue across all lanes, and thus complete the transition to the congested state. Hatakenaka et al. (2006) acknowledges these and other findings for uphill bottlenecks, and notes that a detailed and definitive explanation of the mechanisms of congestion has not yet been offered.

After an uphill bottleneck is activated, the head of the queue remains near the bottom of the hill, although it is triggered downstream. Vehicle departures from the head of this queue are characterized by low accelerations. Drivers may not immediately realize that they are leaving the queue, and therefore fail to accelerate. To add to this, it is possible that heavy vehicles (such as trucks) or keijidousha (light cars with 0.6 L engines that are popular in Japan) are unable to accelerate to their desired velocities due to mechanical limitations, since they discharge from the queue at an uphill grade. Laval (2006) presents a theory to estimate the reduction in capacity of an uphill section caused by the presence of slow vehicles. This framework is further developed in Laval (2009). In theory, the slow vehicles may consist of any type of underpowered vehicle. In practice, it is difficult to determine precisely the distribution of desired speeds (or feasible crawl speeds). The link between heavy vehicles and the activation of an uphill bottleneck has not been proven conclusively. Furuichi et al. (2003) notes that severe congestion on hills occurs on holiday weekends when traffic demand consists primarily of passenger cars with few heavy vehicles.

A number of researchers have attempted to understand specifically what geometric features of the roadway cause some sags, but not others, to become bottlenecks. One group (Furuichi et al., 2003) found that vehicles climbing uphill tend to lose speed very gradually. They concluded that drivers are insensitive to gradual speed changes. As a result, most sag bottlenecks occur in situations where there is a relatively long (upwards of one kilometer) uphill grade, thus providing conditions conducive for subtle speed drops to occur. In fact, the length of the grade may be more important than its slope (Furuichi et al. 2003), but these initial results are not conclusive.
2.1.2 Speed Disturbances and Traffic Theory

The classical, first-order model of traffic theory is attributed to both Lighthill & Whitham (1955) and Richards (1956). This (LWR) theory assumes the existence of a functional relation between flow ($q$) and density ($\rho$). This concept of a fundamental diagram for traffic, first proposed by Greenshields (1935), specifies an equilibrium flow for every physically reasonable density. Empirical studies have since established that a reproducible bivariate relationship (that is approximately triangular for expressways) can be measured during sustained periods of nearly stationary traffic (Cassidy, 1998).

![Diagram](image)

Figure 2.1. Newell’s simplified car-following model. The thick lines labeled $x_{n-1}(t)$ and $x_n(t)$ are trajectories for the $(n-1)$th and $n$th vehicles at time $t$. The values of $\delta_n$ and $\tau_n$ characterize the behavior of the following vehicle given the trajectory of the leader. Finally, $s_n$ and $h_n$ denote the spacing and headway, respectively.

More recently, Newell (2002) proposed a simplified car-following model that provides a microscopic framework to describe the macroscopic behavior. This model begins by assuming the existence of a space displacement $\delta_n$, and a time displacement $\tau_n$ for vehicle $n$. See Figure 2.1] These two parameters are sufficient to characterize the car following behavior for each vehicle $n$ given the trajectory of vehicle $n - 1$. The trajectory $x_n(t)$ of the $n$th vehicle as a function of time, $t$, is

$$x_n(t + \tau_n) = x_{n-1}(t) - \delta_n,$$  \hspace{1cm} (2.1)

as shown in Figure 2.1.

In other words, the trajectory of the $n$th vehicle is exactly the trajectory of the
\((n - 1)\)th vehicle shifted down (upstream) by a distance, \(\delta_n\), and to the right by a time, \(\tau_n\). The values of each \(\delta_n\) and \(\tau_n\) are expected to differ from one vehicle to the next, as if sampled from a joint random probability distribution. However, the arithmetic means of these \(\delta_n\)’s and \(\tau_n\)’s are typically used in practice to specify what is called the “Newell” trajectory. Subsequent work has shown that this simple model characterizes well trajectories of vehicles discharging from long queues at signalized intersections (Ahn et al., 2004).

While the LWR (or Newell) model excels at predicting macroscopic traffic features, it neglects (by its nature) spontaneous speed disturbances that develop in stop-and-go traffic, and multi-lane considerations involving LC. Much effort has therefore been invested in developing car-following models that capture smaller-scale microscopic features, and yet remain consistent with established macroscopic findings.

Yeo & Skabardonis (2009) proposes an asymmetric theory of driving behavior, in which vehicles in congested traffic tolerate a range of \((q, \rho)\) pairs for a given velocity. This differs from the traditional notion of a fundamental diagram, in that the congested regime consists of two branches instead of only one. According to this theory, vehicles accelerate along one branch and decelerate along the other. In addition, hysteresis is possible, and oscillations in speed might be due to human error, i.e. driver anticipation and overreaction. In the context of an extended uphill section, asymmetries between acceleration and deceleration are likely to be exaggerated.

Laval & Leclercq (2009) formulates a model to show how stop-and-go oscillations may result from the reactions of heterogeneous drivers to minor speed variations on a hill. In particular, deviations from the Newell trajectory are categorized as being “timid” or “aggressive.” The study shows that stop-and-go oscillations result from these deviations, regardless of the specific details of the deviations.

Sugiyama et al. (2008) reports that SDs, or traffic oscillations, appear spontaneously in dense traffic. This finding was observed on a circular test track in which no hill was present. This result suggests SDs naturally result from human car-following behavior under suitable conditions. Building on this, Flynn et al. (2009) finds generic existence criteria for nonlinear traveling wave solutions, called “jamitons,” to hyperbolic continuum traffic equations. The study shows that under appropriate road conditions, these jamitons are attracting solutions. Other theoretical studies, such as Wilson (2008) and Ward (2010), have entailed stability analyses for classes of popular vehicle following models. Stop-and-go oscillations are explained as manifestations of linear instabilities that are present in the parameter space of the models themselves.
2.1.3 Effects of Lane Changing

In contrast with the previously cited literature, other studies report that LC maneuvers not only contribute to traffic oscillations, but may even be a primary cause (Mauch & Cassidy, 2002; Ahn & Cassidy, 2007).

Wang & Coifman (2008) reports that the relationship between speed and spacing is temporarily perturbed by LC maneuvers. Furthermore, the perturbation for exiting vehicles is much longer-lived than that for entering vehicles. As a result, an accommodation imbalance (as drivers adjust to vehicle insertions or desertions performed immediately in front of them) may propagate upstream as a source of instability within queues.

Studies also describe other deleterious effects of LC. For example, Cassidy & Rudjananoknad (2005) finds that LC is part of the mechanism that causes congestion to form and output flows to diminish (i.e., a “capacity drop”) at a merge bottleneck. By modeling a LC maneuver as a moving bottleneck with bounded acceleration, Laval & Daganzo (2006) theorizes that LC introduces voids in traffic flow, thus causing the capacity of fixed bottlenecks to drop. LC reportedly causes additional delay in queues, because delay caused by a vehicle insertion into one lane is greater in magnitude than the time savings brought by a vehicle desertion (Coifman et al., 2006).

Duret et al. (2008) supports this result and provides an excellent literature review.

2.1.4 Multi-lane Considerations

In the present study, we draw special attention to the ability of the shoulder lane to absorb LC when its flow and density are low. We interpret this behavior as a “release valve,” made possible by a bias in shoulder lane utilization that has been noted by others. For example, Mika et al. (1969) also studies a 3-lane expressway, and concludes that the shoulder lane typically has lower flow than do the center and median lanes. Tadaki et al. (2002) and Bertini et al. (n.d.) report findings that support this claim for a Japanese expressway and a German Autobahn, respectively.

Of further interest, some of the present findings can be explained using a theory of multi-lane traffic proposed in Daganzo (2002). The theory assumes the existence of at least two driver types: “slugs” who travel with a maximum speed $v_f$; and “rabbits” who travel with a maximum speed $V_f > v_f$. According to this theory, regime changes (between multi-pipe flow with distinct velocities in each lane, and single-pipe flow with similarly low speeds across all lanes) can have forward- or backward-moving interfaces. This feature can explain how a single-pipe (SP) queue can become fixed to the base of a hill, rather than to continue propagating upstream. Details are given in Chapter 5.
2.2 Video Processing

The purpose of this section is not to provide an exhaustive literature review of video processing techniques, but rather to frame the context in which a number of design decisions were made in the present work. In addition, we comment on why some approaches were less applicable than others for this work.

The specific problem of vehicle tracking has been studied by many researchers such as Kanhere et al. (2003) and Kim & Malik (2003). However our needs were more challenging, because in addition to having cameras at a low angle, the road segment under study was curved both horizontally and vertically. Since we lacked an appropriate three-dimensional model of this road segment, a mapping from camera coordinates to real-world coordinates could only be roughly approximated at best. It is also worth noting that no ground-truth trajectories were available (no probe vehicles). As this effort unfolded, tools to enable trajectory visualization alongside video navigation were necessary to verify correctness of extracted data, as we explain in Chapter 3.

A first goal was to perform motion segmentation, to subdivide an image from the video into parts that consist of moving objects of interest, and irrelevant detail. Toward this end a myriad of techniques exist, including phase-based, spectral-based, energy-based, Markovian, and differential methods such as the classical LK (Lucas-Kanade) algorithm (Lucas & Kanade, 1981). The LK algorithm does not perform well in instances of multi-modal motion. For that reason, the use of edge-avoidance to shift the LK integration window away from sharp edges has been explored in Jodoin & Mignotte (2006). Ultimately, we chose a differential method over a subset of the image for its speed and simplicity.

Once an object of interest has been identified, tracking algorithms can follow the trajectory of that object, within some bounds of performance. Most tracking techniques assume an a priori model of the object to be tracked, and motion segmentation is a step toward capturing the object and features to be tracked. Of course, we would have liked to have had completely automatic detection of moving objects as they entered the camera field of view. Toward this end, some have attempted to group segments with similar motions as in Ross (2004) where an improvement on the \( n : m \) matching technique is proposed. Others have grouped regions with similar spatial and temporal characteristics (Hsieh & Lee, 2006). In our case, vehicle trajectories were highly constrained, and the distinctions between motions of vehicles in adjacent lanes were only apparent at relatively long time intervals. In addition, traffic in our videos consisted of a significant fraction of vehicles with irregular shapes and sizes. These variations were exaggerated by the low angle of the camera point of view. Building a set of models for each vehicle type was beyond the scope of this effort.

We had the additional goal of re-identifying objects as they moved from camera to camera. With a high camera angle, this problem has been obviated by first stitching
together frames from multiple cameras to form a mosaic (Kim & Malik, 2003). When
this is not possible, the use of color spectrum histograms (Cheng & Piccardi, 2006) has
been proposed. In our case, the FOV (field of view) for each camera shared a region of
overlap with its immediate neighbors. This suggested the possibility of interpolating
the position of an object as it enters the FOV of a subsequent camera. In this way,
tracking could have served a secondary purpose to aid in re-identification. This turned
out to be feasible in light traffic. However, occlusion due to the low camera angle
became problematic as soon as conditions were dense enough for interesting traffic
phenomena to be observed.

With regards to tracking, particle filter methods have become increasingly popular
(Zhai et al., 2006), using color-based kernels (Lehuger & Lechat, 2006), or using both
dge and color information (Ross, 2006). Particle filtering is a method for imple-
enting a recursive Bayesian filter by using Monte Carlo simulation. The goal is to
construct an approximation of the posterior probability density function (PDF) of the
position of the object being tracked (Arulampalam et al., 2002). This approximation
is represented by a set of random samples (or particles) with associated weights. The
PDF is updated sequentially as new measurements become available.

The mean-shift algorithm has applications in tracking as explained in Comaniciu et al.
(2000), and Comaniciu et al. (2003). In these references, a spatially smooth similarity
measure is defined, and the problem of tracking is reduced to that of moving in the
direction of a maximum gradient. This approach is less computationally expensive
than that of particle filtering.

We ultimately implemented the mean-shift algorithm with a color spectrum histogram
in hopes of combining it with the method of Cheng & Piccardi (2006). However,
with regards to the re-identification problem, it was more effective to exploit the
constrained nature of the trajectories as explained in Chapter 3.
Chapter 3

Methods

3.1 The Site and its Data

The site is a westbound, 3-lane stretch of the Tomei expressway (near Tokyo) instrumented with a series of eleven video cameras and two sets of loop detectors, as shown in Figure 3.1. The downhill approach from the Yokohama-Machida Interchange to the upstream detectors at kilo-post (KP) 21.5 is about 1.8 km in length. Camera surveillance extends from about KP 21.6 to KP 22.9, capturing the end of the shallow, downhill approach to the incline and the first 900 m of the hill itself. The incline has a slope of about 2.4%. The downstream detectors are located beyond the incline, at KP 24.0, on the next downhill section.

In addition to the vertical slope, this section of expressway curves gently to the right, as shown in the top portion of Figure 3.1. Traffic in the opposite direction is separated by a median with guard rails and a hedge. Recall that the Japanese drive on the left, so that the median is the right-most lane. As a result of the curve and the hedge, the line-of-sight for drivers in the median lane is limited somewhat, but not severely so.

Six months of loop detector data reveal 14 days in which congestion persisted on the incline in the absence of spillovers from downstream queues. There exist no nearby ramps available to general purpose traffic\(^1\). Therefore, the observed congestion is intrinsic to the site’s curvature.

Vehicle trajectories were extracted from videos taken during the onset of congestion on one day. A software tool was developed to facilitate semi-automatic identification of vehicles in one camera and reidentification in all other cameras. The result is

\(^1\)There are dedicated bus facilities serviced with shoulder-lane ramps near KP 24.0, but given the low flow of buses, it is extremely unlikely that this facility has any appreciable effect on congestion.
Figure 3.1. Tomei Expressway study site. There are three lanes of traffic moving from left to right in the figure. One (upstream) set of loop detectors is positioned near the first camera at KP 21.5 (kilo-post 21.5), and a second (downstream) set of loop detectors is positioned at KP 24.0 (beyond the uphill grade). A series of eleven cameras are placed about 120 m apart, on the left-side of the road facing forward from KP 21.57 to KP 22.73.
an internally consistent set of 2284 vehicle trajectories over the crucial 30 minutes leading up to and including the onset of persistent congestion.

### 3.2 Video Tools

A basic framework was built to facilitate the extraction and manipulation of vehicle trajectories using videos taken from a series of roadside cameras. The overall strategy was to automate all of the most tedious and repetitive tasks. In particular, GUI tools were built to enable a user to supervise the trajectory extraction process with a minimum of mouse clicks. The resulting framework may be applicable to other sites, including arterial intersections, with basic modifications.

One problem with video data is that it is very time-consuming for a user to actually watch. Therefore, our first goal was to enable the user to scan the contents of a video without actually watching the video. For this purpose, slit-scan images (Figure 3.2) were created to summarize several seconds of video data into one easily interpretable image. A second goal was to automate much of the trajectory extraction process. For this purpose, a GUI based program, *vid* was written in Matlab. A screen-shot of *vid* is shown in Figure 3.3. A third goal was to enable a user to interact with the trajectory data and the videos together, in an integrated environment. To do this, a GUI based program, *tview* was written. A screen-shot of *tview* is shown in Figure 3.4.

#### 3.2.1 Pre-processing

A video can be considered as a series of still images in time, as shown in Figure 3.2(a). In the figure, horizontal and vertical axes of the image plane are labeled $x$ and $y$, respectively. The roadside cameras are placed so that vehicles recede from view. As a result of the geometry, vehicle trajectories are highly constrained, and must move from the bottom of the frame to the top of the frame in each camera. A line crossing the image plane at a fixed $y$ corresponds roughly to a line on the road crossed by passing vehicles on the expressway.

One can fix $y$ to acquire one horizontal scan-line of the current frame, and to build a new image (referred to as a slit-scan image in more traditional photography) by concatenating the same scan-line from a number of successive frames, as shown in Figure 3.2(b). This slit-scan image captures each vehicle as it crosses a line on the road. For our application, video data was sampled at the NTSC frame rate of 29.97 Hz. A ten-minute video would result in a slit-scan image with 17982 pixels in one direction. Since this is unwieldy, slit-scan images were chopped up into epochs. A numerology was chosen with 111 frames per epoch, yielding 162 epochs per 10 minutes of video.
Figure 3.2. Building a slit-scan image. Video can be considered as a series of image frames in time. In (a), the $x$ and $y$ axes correspond to the horizontal and vertical directions in the frame of the image as seen by the camera. In (b), the value of $y$ is fixed, and a series of scan-lines in successive frames is concatenated to form a slit-scan image. In the slit-scan above, time flows from left to right.
In the first pre-processing step, a set of epoch images is built directly from the video data. For the present work, the term “epoch” or “epoch image” is used to mean one slit-scan image representing a duration of 1 epoch (about 3.7 seconds). Each camera has its own set of epochs. Depending on a vehicle’s arrival time and speed, its image may be divided into more than one epoch. Each epoch image summarizes exactly one epoch of video data where the vertical-axis in the frame has been traded for the time axis.

The screen-shot of $\text{tview}$ in Figure 3.4 shows two epochs in the top half of the window. Distorted vehicular shapes are numbered. Time flows from the top of the epoch to the bottom. One key benefit from this approach is that a user can readily understand the contents of a short video segment from a glance at the epoch image without taking the time to view the actual video. Notice that the road is slightly curved as shown in the image frames of Figure 3.2(a). In addition to moving in the positive $y$-direction, vehicles move in the negative $x$-direction as well. The trend of this motion appears in the epoch images by the slant of the vehicle. For example, the slower-moving vehicles that appear in the epochs of Figure 3.3 are stretched-out vertically. The somewhat faster-moving vehicles that appear in the epochs of Figure 3.4 appear to be compressed, and the angle of the slant is closer to the $x$-axis. This effect is exaggerated for tall vehicles such as trucks.

The vertical “location” of the bottom edge of a vehicle in an epoch image corresponds to the time when that vehicle passed the point on the roadway corresponding to the fixed location $y$ in the original frame of the camera. As shown in Figure 3.4, vehicle 312 arrives in the center lane, followed by vehicles 314, 317, 319, etc. The headway between two vehicles is measured as the number of pixels from rear-bumper to rear-bumper, where each pixel is about 0.0334 seconds.

In the second pre-processing step, a first-pass vehicle detection algorithm is applied to each epoch in order to build a new series of images referred to as “ghosts.” Each ghost is a binary image in which the background (asphalt) is black, and the foreground (vehicle) is white. To segment the images, background subtraction and a series of morphological operations are employed. During this step, occlusion caused by large trucks, or closely following vehicles can be problematic, so further logical steps are used to build up a set of hypotheses for vehicle locations.

Each hypothesis is a spatio-temporal index into the video, and contains a camera number, a scan-line, an epoch number, a vertical offset within the epoch, and a horizontal position within the epoch. The epoch number and vertical offset together encode a time-stamp, and identify the frame in the video when the hypothesized vehicle appeared at the scan-line. The scan-line is equal to the vertical position of the vehicle in the frame. The horizontal position in the epoch is equal to the horizontal position in the frame.

These two pre-processing steps must be repeated for the video from each camera. In
the present work, a total of eleven cameras are utilized. Upon completion of pre-
processing, the following are obtained: (1) eleven series of epochs, (2) eleven series
of ghosts, and (3) eleven series of vehicle location hypotheses (one for each camera).
At this point, data corresponding to vehicle counts has already been extracted, as if
each camera were a loop detector. Of course there are errors in the data, but these
errors consist mostly of missed vehicles caused by occlusion; as false positives were
rare with this method.

3.2.2 Vehicle Re-identification

The next goal is to trace a vehicle’s path as it moves from camera to camera. For now,
consider only one lane at a time. Assume that this task has already been completed
for vehicles 1 through \( n - 1 \) in all cameras. Assume also that vehicle \( n \) has been re-
identified for all cameras up to and including camera \( j - 1 \). We now concern ourselves
with vehicle \( n \) as it moves from camera \( j - 1 \) to camera \( j \). Recalling Newell’s car
following model, we exploit the fact that the relative positions of vehicles do not
change drastically from camera to camera, although vehicles do occasionally change
lanes, slow down, and speed up under the influence of nearby traffic conditions. Our
task is to re-identify vehicle \( n \) in camera \( j \).

In camera \( j - 1 \), vehicle \( n \) has already been observed following directly behind vehicle
\( n - 1 \). It seems reasonable to expect vehicle \( n \) to appear soon behind vehicle \( n - 1 \)
in camera \( j \) as well. Using this strategy, a recursive scheme is employed in which the
algorithm can choose a vehicle location hypothesis from the pre-processing step for
vehicle \( n \) in camera \( j \) given the location of previous vehicle \( n - 1 \) in both camera \( j \)
and camera \( j - 1 \).

Although simple, this method was fast and worked quite well. Recall that cameras
are only separated by 120 m, and vehicles typically cover this distance in about 4 s
to 12 s, depending on prevailing speeds. Lane change maneuvers on this time-scale
were not numerous enough to justify a more involved strategy.

If all the hypotheses were correct, the re-identification process would consist of collect-
ing the hypotheses that correspond to the same vehicle. The collection of hypotheses
would then become a collection of spatio-temporal observations (i.e., a trajectory).
Unfortunately, not all hypotheses were correct and therefore vehicle matching needed
to be supervised to guarantee 100% accuracy. For this purpose, a GUI named \texttt{vid}
was written to display epochs in cameras \( j - 1 \) and \( j \), a screen-shot of vehicle \( n \) in
camera \( j - 1 \), and a screen-shot of the hypothesized vehicle \( n \) in camera \( j \).

The \texttt{vid} GUI is displayed in Figure 3.3. In this example, the top and bottom halves of
the GUI correspond to cameras 3 and 4, respectively. Two epochs and a screenshot of
Figure 3.3. The vid GUI. The top and bottom halves of the GUI correspond to cameras $j - 1$ (camera 3), and camera $j$ (camera 4), respectively. Vehicles $n - 1$ and $n$ correspond to vehicle ID 17 and ID 22, respectively. Two epochs and a screenshot of vehicle 22 are shown at the top, and two epochs and a screenshot of the hypothesized vehicle 22 (indicated with a question mark) are shown at the bottom.
vehicle 22 are shown at the top, and two epochs and a screenshot of the hypothesized vehicle 22 (indicated with a question mark) are shown at the bottom.

As shown in the top status panel of vid, epoch 403 is the current epoch for camera 3. The status panel also indicates that the top screen-shot displays a frame from video camera 3 taken at 07:04:45:15. In the screen-shot, vehicle number 22 out of 57 is marked with a cyan ellipse and a crosshair. The crosshair is drawn again in the epoch. The crosshair indicates the spatio-temporal observation of vehicle number 22.

As shown in the bottom status panel of vid, epoch 405 is the current epoch for camera 4. The status panel also indicates that the bottom screen-shot displays a frame from video camera 4 taken at 07:04:49:26. In the screen-shot, a vehicle is marked with a magenta ellipse and a cyan crosshair. Once again, this crosshair is drawn in the current epoch as well. The crosshair indicates the spatio-temporal hypothesized vehicle number 22. In order for observations to be precise, it is important for observations of vehicles in successive cameras to be consistent, with the crosshair in the lower left-hand corner of the vehicle, and the ellipse to be a similar size.

In addition to the other markers, vid draws a cyan line from vehicle $n - 1$ to vehicle $n$ in both sets of epochs. The user has several options depending on the result. If the result is correct, he simply accepts the observation as valid and continues to the next vehicle. If the result is not correct, he can fix it. Typically, the hypothesis is good enough to enable the GUI to index the correct epoch (in this case, it has successfully found vehicle 22 in epoch 405 of camera 4). Within the epoch, one can easily find the correct vehicle, as the downstream vehicles in that space-time vicinity have already been numbered and identified in the epoch image displayed for the user. One click in the epoch indexes into the video to yield a corrected screen-shot.

Sometimes a fine-tune adjustment is necessary to get the crosshair in the right place. A version of the Lucas-Kanade algorithm (indexing in frame number instead of $y$ to move the image in the vertical direction) was employed to make these adjustments. To assure 100% self-consistency, this process was repeated for each of the 11 cameras in the series. Occasionally, large trucks would completely occlude smaller vehicles for a subset of camera locations. In this case, the user would specify a best guess for the location of the smaller vehicle.

The process outlined above was performed on batches of about 20 epochs, re-identifying vehicles one lane at a time. Rough vehicle trajectories, including all lane change maneuvers are captured upon completion of this re-identification task.
Figure 3.4. The tview GUI. The GUI displays two epochs and a screen-shot. The status panel indicates that vehicle 321 out of 2284 arrived at camera 7 at 06:43:57:25. Controls enable the plotting and navigation of trajectories in a separate window. Shown here, vehicle 321 is about to change lanes, its trajectory is displayed in Figure 3.5.
Figure 3.5. Trajectories of the tview GUI. The trajectory window shows three time-space diagrams. The median, center, and shoulder lanes are shown in the top, center, and bottom stages, respectively. The trajectory of vehicle 321 is shown in dark black, and its lane-changing maneuver (from the center lane to the median lane) in the current camera is shown in yellow.
3.2.3 Database Exploration

Another GUI, named tview, was written to enable exploration of the data-set. In one window, trajectories are plotted on a 3-level time-space diagram, with one level for each lane (see Figure 3.5). Lines traversing from one level to another indicate a lane-change maneuver. In another window, epochs and screenshots are displayed as in vid. Clicking on a trajectory chooses an observation of that vehicle at that camera. In the trajectory window, that trajectory is drawn in dark black, and the segment of that trajectory in that camera is highlighted in yellow. The main GUI window (see Figure 3.4) displays the epoch containing that vehicle alongside a screenshot of that vehicle in that camera.

Clicking in the epoch has one of two functions that depends on the GUI mode. In view mode, clicking in the epoch chooses the closest vehicle in the time-space vicinity of the click. This action highlights the corresponding trajectory and displays a screenshot of that vehicle as described above. In search mode, clicking in the epoch will index into the video and choose a screenshot from the video. The vertical, and horizontal positions of the click in the epoch choose the frame number within the video, and horizontal offset within the indexed frame, respectively. In this way, the user can simultaneously navigate the video and trajectory.

This visualization tool was crucial, as it enabled the investigation of local traffic features anywhere within the space-time boundaries of the surveillance zone such as, lane change maneuvers, brake lights, turn signals, behavior of drivers in adjacent lanes, etc. This capability was deemed to be more useful in the scientific inquiry than the acquisition of more detailed trajectories.

3.2.4 Detailed Trajectories

To improve the resolution of the trajectories, the mean-shift algorithm was first implemented using a color histogram. This algorithm was augmented in two ways. First, a linear motion model was added to reduce jitter in the trajectory. Second, an adaptive sizing function was added to compensate for the fact that vehicles progressively appear smaller as they recede in the FOV of the camera.

The augmented algorithm was then applied to several hundred vehicles traversing each of the 11 cameras. Good results were obtained when vehicles were about 40 m to 80 m away from the camera. Beyond this range, tracking quality dropped off quickly due to occlusions or tracking drift. As cameras were placed 120 m apart, this resulted in alternating sequences of good trajectory data (for about 40 m) and

---

2The first version of this software was developed in cooperation with Rodny Rodriguez as part of a class project for Electrical Engineering 225B, Digital Image Processing.
less good trajectory data (for about 80 m). Trajectory data were then filtered and
smoothed, partially taking into account the quality of the measurements.

Spline interpolation of the original 11 data points in the rough trajectories yielded
comparable results. Therefore, rough trajectories were used in the remainder of this
work. This was reasonable because aggregate traffic patterns such as SDs occur with
spatial periods of 1 km or more, and temporal periods of about 2 min to 5 min.
Chapter 4

Empirical Evidence

The data presented below reveal the mechanism of congestion formation and output flow reduction. Clues to the mechanism are found in loop detector data across days (Section 4.1). Vehicular trajectories from one day reveal details in four stages (Section 4.2).

4.1 Chronic Macroscopic Features

Loop data reveal repeatable patterns that are now examined lane by lane. Figures 4.1 and 4.2 present average speeds versus flows measured by the upstream and downstream detectors at KP 21.5 and KP 24.0, respectively. Red circles, blue squares, and green diamonds are 5-minute aggregated loop detector data for shoulder, center, and median lanes, respectively. Several hours of data are shown from each of the 14 days when persistent congestion was observed in all three travel lanes at the upstream detectors, while absent at the downstream detectors.

For each lane, observed values for speeds and flows fall along separate but distinct arcs. On the day for which video data are available (December 23, 2005), time-of-day tags from 6:30 to 7:10 hrs are plotted at 10-minute intervals. Each tag shows the measured traffic state at the corresponding 5-minute interval. On each of the 14 days, the evolution of traffic states leading to congestion followed along the same three arcs; i.e., the pattern was reproducible across days. In particular, speed reductions in all lanes always coincided with increased flow in the shoulder lane.

The earliest portion of a rush (at around 6:30 hrs) on the upstream end of the site was characterized by multi-pipe traffic, in that speeds and flows were markedly different in each lane. In Figure 4.1, average speeds ranged from about 100 kph in the median lane to about 80 kph in the shoulder lane. Flows were especially high in the median lane.
Figure 4.1. Time course of traffic states (upstream). Speeds versus flows measured at KP 21.5 are plotted. Red circles, blue squares, and green diamonds are 5-minute aggregated loop detector data for shoulder, center, and median lanes, respectively. Shown here are several hours of data from each of the 14 days when persistent congestion was observed in all three travel lanes at the upstream detectors, while absent at the downstream detectors. On the day for which video data are available (December 23, 2005), time-of-day tags from 6:30 to 7:10 hrs are plotted at 10-minute intervals. Each tag shows the measured traffic state at the corresponding 5-minute interval. These tags illustrate the time course of traffic on that day.
Figure 4.2. Time course of traffic states (downstream). Speeds versus flows measured at KP 24.0 are plotted. Red circles, blue squares, and green diamonds are 5-minute aggregated loop detector data for shoulder, center, and median lanes, respectively. Shown here are several hours of data from each of the 14 days when persistent congestion was observed in all three travel lanes at the upstream detectors, while absent at the downstream detectors. On the day for which video data are available (December 23, 2005), time-of-day tags from 6:30 to 7:10 hrs are plotted at 10-minute intervals. Each tag shows the measured traffic state at the corresponding 5-minute interval. These tags illustrate the time course of traffic on that day. The speed limit is 100 kph.
lane, rising from around 2200 vph at 6:30 hrs to about 2500 vph at 6:40 hrs. The median lane thus carried almost 50% of the expressway’s total flow. Flows in the shoulder lane, on the other hand, were low at this time (around 1000 vph), and constituted only about 20% of the total flow.

Note that high flow in the median lane was observed to persist in the presence of low shoulder lane flow (under 1200 vph). Under these conditions, total output flows ranging from 5000 vph to 5500 vph were sustained. Persistent output flow reductions (from 4% to 11%) consistently coincided with a rapid increase in shoulder-lane flow. The shoulder lane’s increased utilization was repeated on all days shown in Figures 4.1 and 4.2 and turned out to be a key part of the congestion mechanism.

As the rush wore on, traffic at the upstream detectors gradually transformed into a SP, congested state with speeds and flows that were much more balanced across lanes. Figure 4.1 reveals that by 7:10 hrs, speeds in all lanes were about 40 kph (with the low speed indicating the presence of congestion); and flows ranged only from about 1800 vph in the median lane to about 1400 vph in the shoulder lane. The latter lane eventually carried about 30% of the expressway’s total flow.

The pattern of traffic measured at the downstream detectors (near KP 24) also provides clues to the mechanism. Figure 4.2 reveals that traffic was initially in a multi-pipe, uncongested state, as it was upstream. However, traffic at the downstream location never exhibited a SP, congested regime: once having passed through the SP queue that had formed later in the rush, drivers evidently re-distributed themselves across the three expressway lanes. Once vehicles reached the downstream detector, their speeds and flows were highest in the median lane and lowest in the shoulder lane. These matters will now be explored in greater detail using the traffic data collected from videos.

4.2 Microscopic Spatio-temporal Inspection

Figures 4.3 and 4.4 show velocity fields calculated for each of the three lanes as per generalized definitions (Edie, 1963) using a moving time-space window of 6.4 s in duration and 40 m in extent. Regions of high and low speed are colored green and red, respectively. Regions devoid of vehicle trajectories appear as white streaks.

In Figure 4.3, triangular arrowheads indicate leftward lane-changing maneuvers that are spatially quantized to the nearest camera. The color of the arrowhead distinguishes insertions into the lane (in black) from desertions (in white). Figure 4.4

1 The ratio (40 m : 6.4 s) corresponds to a backward wave speed of 22.5 kph. Assuming constant travel speeds of 90 kph, the moving window will always overlap at least one trajectory if headways are under 8 s. If the window is smaller, then more white streaks (zero density in the window) appear in the shoulder lane. If the window is bigger, then details are smoothed somewhat.
Figure 4.3. Velocity fields and leftward lane changes (December 23, 2005). Regions of high and low speed are colored green and red, respectively. Using vehicle trajectories, velocities are calculated as per generalized definitions ([Edie 1963]) using a moving time-space window of 6.4 s in duration and 40 m in extent. Regions devoid of vehicle trajectories appear as white streaks. Locations of reddish colored speed disturbances (SDs) are numbered. The first emergence of a single-pipe, congested regime is labeled SP. Triangular arrowheads indicate vehicular insertions into the lane (in black) or desertions (in white). A template of congestion stages is defined in (d).
shows rightward LC maneuvers; the velocity field and labeling are consistent with Figure 4.3. LC maneuvers that occur closely in space and time may overlap in the figures.

SDs are manifest as transient disruptions in the velocity field. Trajectories reveal that a small drop in speed by a vehicle (causing a delay on the order of 0.3 s over a distance of 1 km) is amplified by successive vehicles and results in a SD as described by Koshi et al. (1992), and Laval & Leclercq (2009). On average, vehicles respond to incipient SDs by braking later than predicted by Newell’s simplified car following model (Newell, 2002). As a result, successive vehicles decelerate more quickly and over shorter distances to avoid collision. A series of these perturbations coalesce into a high-density, backward-moving SD.

For the purposes of this study, SDs are identified as they propagate upstream between KP 22.2 and KP 21.73 (the base of the hill). A SD is designated as such when for any 240 meter stretch between those two KPs, the vehicular travel times in one lane increase by at least 10 s and then recover by at least 10 s. Six instances of this were identified; these SDs are numbered chronologically, as they arrive at KP 22, and appear as reddish swaths in Figures 4.3 and 4.4.

A time-space diagram annotating four stages of congestion is provided as a template in Figure 4.3(d) and repeated in Figure 4.4(d). These diagrams also demark three spatio-temporal regions in which key traffic phenomena are observed. In region A, LC maneuvers out of the median lane become conspicuously sparse. Particularly disruptive LC maneuvers occur in regions B and C.

Cumulative oblique counts of vehicles, plotted in the usual way (Cassidy & Windover, 1995), are furnished in Figures 4.5 and 4.6. In Figure 4.5, shoulder lane count is measured at KP 21.7. In Figure 4.6, vehicle counts across all lanes are provided for: KP 22, KP 22.45, KP 22.7, and KP 22.8. Plots for the downstream positions (KP 22.7 and KP 22.8) are almost exactly coincident, indicating an absence of vehicular delays. Vertical displacements between any two of the other curves indicate excess vehicle accumulation between the corresponding measurement locations.

Details will be examined in the following four sections, one stage at a time. Key findings, containing lane-specific details, are numbered to clarify the presentation.

### 4.2.1 Stage One: Early rush

Inspection of Figures 4.3–4.6 will reveal the following features of this early stage (with time-space boundary as shown in the template of Figures 4.3 and 4.4). Traffic in the shoulder lane was freely flowing while SDs periodically propagated through the center and median lanes. Surges in leftward LC out of those lanes and toward the shoulder
Figure 4.4. Velocity fields and rightward lane changes (December 23, 2005). Triangular arrowheads indicate lane desertions (in white) or lane insertions (in black). The velocity field and labeling are consistent with Figure 4.3. For ease of reference, the template of congestion stages is illustrated again in (d).
lane coincided with the early SDs (and were performed without measurable impedi-
ment), as some drivers bypassed the transient slow-downs. Those drivers who were
slowed promptly recovered their speeds once having passed through the SDs. In their
wakes, output flows in the center and median lanes increased and excess accumu-
lations abated. The directionality of LC reversed as some drivers then maneuvered
rightward toward those now faster-moving lanes.

Observation 1  Traffic in the shoulder lane was by-and-large freely flowing, with low
average flow and forward-moving voids.

The greenish shades that dominate Figure 4.3(a) during the first stage unveil the high
speeds that initially persisted in the shoulder lane. Flow in that lane was low at the
time (about 1000 vph), and the white, forward-moving streaks reveal voids within
which no vehicles were present.

Observation 2  SDs coalesced gradually and then propagated backward, primarily af-
fecting the median and center lanes.

The three swaths of red in Figures 4.3(b) and 4.3(c) unveil periodic SDs that tem-
porarily slowed traffic in the center lane, and even more so in the median lane.

Observation 3  The highest concentrations of leftward LC (toward the shoulder lane)
coincided with the presence of the early SDs.

The three early SDs were each accompanied by brief surges in leftward LC maneuvers.
Each surge of this kind persisted for about one minute at rates of 15 to 30 LC
maneuvers per lane-km per minute. Note how the white arrowheads in Figure 4.3(c)
cluster along the reddish swaths of SDs 1, 2, and 3. These same LC maneuvers
appear as black arrowheads in Figure 4.3(b) indicating insertions into the center
lane. Leftward insertions into the shoulder lane, indicated by black arrowheads in
Figure 4.3(a), were absorbed with minimal disruption to shoulder-lane traffic; i.e.,
these insertions were not accompanied by low shoulder-lane speeds.

Observation 4  Output flow increased and accumulation diminished in the immediate
wake of each early SD.

The slopes of the downstream-most curves in Figure 4.6 reveal that output flows
increased (by about 800 vph) in the immediate wake of SDs 1, 2 and 3.

\[2\] The flow increase after SD2 was smaller, as it was cut off by SD3. SD3 occurred about two
minutes after SD2.
output flows cleared away excess vehicle accumulation as evident in Figure 4.6 by the wax and wane of vertical displacements between curves.

Observation 5 After traveling downstream of a SD, vehicles migrated rightward toward the median (reversing the directionality of the LC pattern).

The black triangles of Figures 4.4(b) and 4.4(c) indicate rightward LC maneuvers into the center and median lanes, respectively. These insertions primarily occurred into the greenish regions of those figures, in which the speeds in the wake of the early SDs were high. Note also Figure 4.4(a) from 6:42 to 6:44 hrs, in which white arrowheads proliferate in the tranquil lull between SD1 and SD2. These rightward LC maneuvers coincided with a short-lived surge in shoulder lane input flow, as revealed by a short-term rise in the slope of the oblique count curve in Figure 4.5.

4.2.2 Stage Two: Persistent slowing

In stage two, higher shoulder lane input flow was followed by fewer leftward LC maneuvers out of the median lane. Vehicles passing through SD4 no longer recovered speeds quickly. Output flow did not increase so as to clear away excess vehicle accumulation.

Observation 6 A rise in shoulder lane flow occurred near 6:48 hrs and this higher flow persisted, as shown in Figure 4.5.

Observation 7 A rise in rightward LC occurred after 6:48 hrs, in the wake of SD3. These rightward maneuvers were more numerous than those accompanying earlier SDs (Observation 5).

Note the long, slender cluster of white arrowheads in Figure 4.4(b) demarking rightward center to median LC maneuvers, which perforate the boundary between stages one and two. Soon afterward, black arrowheads in the same figure demark rightward shoulder to center LC maneuvers. Seven of these latter maneuvers entered the center lane along the leading edge of SD4 (between KP 21.9 and 22.4). These migrating vehicles evidently blocked those in the median lane from making leftward LC maneuvers (see Observation 8).

Observation 8 Unlike previous SDs, there was a scarcity of leftward LC out of the median lane along SD4. In particular, note the dearth of arrowheads in the portion of Figure 4.3(c) corresponding to Region A in the template.
Figure 4.5. Cumulative oblique counts of vehicles (shoulder lane only), plotted in the usual way, (Cassidy & Windover 1995). Shoulder lane count is measured at KP 21.7. With the exception of the short-lived surge between 6:42 and 6:44 hrs, long-run shoulder-lane flow was low (1100 vph as evident from the curve’s slope) until about 6:48 hrs. Afterward, long-run shoulder-lane flow was high (1400 vph) with a notable surge coinciding with the emergence of a single-pipe regime, marked SP.
Region A demarks a time-space region in which relatively few median to center LC maneuvers were observed (few white arrowheads in Figure 4.3(c)). Therefore, few step increases in vehicular spacings appeared in the median lane in the wake of SD4. Vehicles did not accelerate quickly in that lane as evidenced in Figure 4.3(c) by the reddish shading at and upstream of KP 22.5. Again in Region A, there were numerous LC maneuvers out of the center and into the shoulder lane, as shown by the white arrowheads in Figure 4.3(b). However, these were insufficient to clear the excess accumulation brought by SD4, as evidenced by what now become persistent vertical deviations between curves in Figure 4.6.

### 4.2.3 Stage Three: Emergence of a single-pipe regime

LC maneuvers become more disruptive in stage three. In the shoulder lane, high flow persisted and large gaps between platoons disappeared. Continued leftward LC spread the queue to the shoulder lane, resulting in the emergence of a SP, congested regime. Rightward LC in region C of Figure 4.4(c) exacerbated SD6 in the median lane.

**Observation 9** Leftward center- to shoulder-lane vehicle insertions coincided with reduced speeds in the shoulder lane and high shoulder-lane flow.

LC maneuvers into the shoulder lane occurred in reduced numbers (note from Fig 4.3(a) the diminished density of black arrowheads in Region B of the template) as compared to those that clustered around the early SDs. However, insertions of this type that did occur coincided with high shoulder-lane flow (about 1400 to 1600 vph, as shown in Figure 4.5) and the disappearance of gaps between platoons (no further white streaks in Figure 4.3(a)). As a result, LC maneuvers into the shoulder lane became disruptive: they resulted in a lateral spreading of the queue, as evidenced by the reddish shading corresponding to Region B in Figure 4.3(a).

**Observation 10** Sustained slowing persisted in all lanes near the base of the incline and did not propagate backward as did the earlier SDs.

Persistent reddish zones in Figure 4.3, most notably in Region B, indicate a queue’s presence in all lanes. The zones grew over the course of several minutes, as LC maneuvers (noted in Observation 9) continued in the vicinity.

**Observation 11** Rightward center- to median-lane vehicle insertions occurred at the head of SD5 and the leading edge of SD6.
Figure 4.6. Cumulative oblique counts of vehicles (all lanes), plotted in the usual way, (Cassidy & Windover, 1995). Vehicle counts across all lanes are provided for: KP 22, KP 22.45, KP 22.7, and KP 22.8. Plots for the downstream positions (KP 22.7 and KP 22.8) are almost exactly coincident, indicating an absence of vehicular delays. Before 6:50 hrs, the curves show intermittent excess vehicle accumulation due to vehicular delays (i.e., vertical displacements between curves) resulting from SDs. The excess accumulation in the zone between KP 22 and KP 22.45 was not sustained until after 6:50 hrs.
In region C a recurrence of rightward LC was observed, as evidenced by the cluster of arrowheads in Figures 4.4(b) and 4.4(c). SD5 expanded longitudinally from Region C and marked the boundary between stages three and four.

4.2.4 Stage Four: Reduction in output flow

The queue became well established in stage four and output flow diminished. Low speeds prevailed upstream of KP 22.6 and extended upstream of the surveillance region. Output flows from this queue were reduced from the earlier observed rates. The LC pattern at the downstream end of the SP queue was predominantly rightward (shoulder to center, and center to median). Multiple SDs occurred within the queue, which surely caused driver workload to increase.

Observation 12 SD5 and SD6 were qualitatively similar to SD4, repeating the same pattern of observations.

Again, relatively few leftward LC maneuvers were observed along SD5 and SD6 (few white arrowheads in Figure 4.3(c)), much as in Observation 8. Low speeds persisted in all lanes in the wake of these latter SDs, with significant speed reductions in the shoulder lane as well; see Figure 4.3(a). Excess vehicle accumulation increased, as evidenced by progressively larger vertical displacements between curves in Figure 4.6.

Observation 13 The queue further extended in length and was characterized by disruptive SDs.

The queue extended beyond the base of the incline (KP 22), stretching more than 1 km in length by 6:58 hrs (shown by reddish shading in Figures 4.3(a), 4.3(b), and 4.3(c)). Vehicles arrived to the base of the incline at low speeds, and experienced multiple SDs within the queue, as evidenced by fluctuating bands of red in Figures 4.4(b) and 4.4(c). Rightward vehicle insertions were prevalent in the acceleration region downstream of KP 22.4 (marked by the arrowheads in the upper right corner of Figure 4.4(b)). The SDs were apparently exacerbated by the continued LC, and all this appears to have increased driver workload, as evidenced by slightly increased headways maintained by drivers climbing the hill; see observation 14. This increase was most pronounced in the median lane, where larger headways resulted in a long-run reduction in that lane’s output flow of 10%.

3Backward propagating disturbances in addition to SD5 and SD6 formed during this stage. However, these did not meet the speed recovery condition in the definition of Section 4.2. These disturbances occurred at twice the frequency of the SDs during the early rush.
Observation 14  *Output flows diminished.*

Over the shorter run, total output flow across all lanes dropped by 750 vph, or 14%, which persisted until about 7:00 hrs; note the reduced slopes in the downstream-most curves of Figure [4.6] from about 6:55 to 7:00 hrs. This transient flow reduction mostly occurred in the center and median lanes, and contributed to longitudinal queue growth. Over the longer run, vehicles discharged from the queue at about 4800 vph—a 6% reduction from the longer run output flow measured during the first three stages of congestion. As evident in Figure [4.6], the reduced longer run average output flow persisted until the rush subsided and the queue dissipated.
Chapter 5

Discussion

A clear, back-and-forth LC pattern was faithfully reproduced for each of SDs 1, 2, and 3. Leftward then rightward LC maneuvers occurred in sequence and in significant numbers as described in Observations 3 and 5. All of these LC maneuvers were optional in that they were independent of route choice; i.e., there are no junctions for about 10 km downstream. The SDs were the largest and most obvious traffic features in the videos. There was a clear tendency to vacate the median lane when it slowed due to an early SD and then to return to that lane when its speeds recovered, i.e., the overwhelming majority of LC maneuvers materialized in lockstep with these SDs, and appeared to be directly caused by them.

Observations 5 and 7 support the interpretation that a significant fraction of vehicles arriving at KP 21.7 in the shoulder lane, and in the wake of SD1 and SD3, were what Daganzo (2002) refers to as rabbits, and not slugs. These rabbits would not ordinarily choose to use the shoulder lane. As evidence, notice that SD1 and SD3 propagate upstream of the surveillance region in the median lane. It seems reasonable to speculate that the increases observed in the shoulder-lane’s input flow were the result of increasing numbers of rabbits temporarily utilizing that lane. Once downstream of the SD that prompted their tactical move, the rabbits returned to their preferred lanes. The shortsightedness of this tendency resulted in Observations 7 and 8, in which conflicting rightward LC maneuvers produced persistent slowing and excess accumulation.

Similar to claims in Koshi et al. (1992), we find that a SD can arise on a hill when flow becomes high in the median lane, and that this can trigger a persistent SP queue. Building on Koshi’s earlier work, we further find that the congestion mechanism does not entail one, but rather a series of SDs. Within one day, reproducible patterns are observed in the videos. As described above, these patterns evolve as conditions change. Macroscopic details presented in Section 4.1 are consistent with the detailed
patterns revealed by the trajectories, and with those reported by others at other sites, suggesting that the pattern is reproducible across both days and sites.

Early in the rush, when flow was low in the shoulder lane, the LC induced by the SDs was beneficial: it promoted high output flow by enabling the extra accumulation brought by the SDs to promptly dissipate, and by redistributing vehicles into the underutilized shoulder lane. When flow in that lane increased later in the rush, LC became baneful: it resulted in diminished output flows by spreading the SDs laterally, and by disrupting flows in the SP queue that formed as a result. To our knowledge, the current findings are the most detailed of their kind.

Further discussion on this mechanism, and in particular the role of the shoulder lane, is offered in Section 5.1. The present observations are compared against theory in Section 5.2. Implications for traffic management are examined in Section 5.3, and finally, directions for future work are discussed in Section 5.4.

5.1 The Shoulder Lane as a Release Valve

When shoulder-lane flow was low (around 1000 vph), the predominant leftward pattern of LC gave rise to the periodic surges in output flow evident in Figure 4.6, and thus to a higher long-run average rate. Vehicles redistributed toward the less utilized lanes, instantly leaving behind larger spacings (especially in the median lane). Vehicles remaining in the median lane accelerated quickly, soon reaching free-flow speeds. Careful study of oblique count curves (like the one in Figure 4.5) revealed that output flows initially increased in the shoulder lane, like a release-valve, and that this was soon followed by secondary increases in the center and median lanes. As a result, total output flow increased in the wake of each SD, and cleared the excess accumulation as evident in the early portions of the curves in Figure 4.6. As will be discussed further in Section 5.2, this release valve action in the shoulder lane enabled many of the more aggressive drivers (i.e., “rabbits”) to bypass oncoming SDs.

The release valve failed only when flow in the shoulder lane reached about 1400 vph and large gaps between platoons were no longer available to accommodate the leftward LC. As drivers continued thereafter to migrate into the shoulder lane, speeds in that lane diminished. This resulted in the regime change from multi-pipe to a SP queue, thus closing off the rabbit “escape route.”

The SP regime emerged during the final collapse of the release valve. During the transition, some of the highest output flows were observed; i.e., the rate across all lanes reached 5300 vph for a brief 5-minute period. Yet this rate soon diminished markedly for the reasons described in conjunction with stage four (Section 4.2.4).

In light of this, we reinterpret (at least for the present site) what has widely been
called a "capacity drop" as a failure of the shoulder lane to accommodate beneficial LC; i.e., as a failure of the release valve. This is not a trivial distinction. The implication is that before this bottleneck activates, drivers temporarily tolerate non-equilibrium headways (and smaller spacings at high speeds) over spatial zones of up to several hundred meters. Flows measured under these conditions do not represent a "capacity" or level of efficiency that drivers are willing to tolerate indefinitely. As a queue extends beyond 500 meters or so (the extent of a typical SD), drivers appear to become less tolerant of non-equilibrium headways.

5.2 Comparisons with Theory

The multi-lane traffic theory of [Daganzo (2002)] can explain elements of the observed mechanism. According to this theory, drivers self-segregate in lanes based on their desired speeds. Rabbits will confine themselves to the passing lane(s) whenever they enjoy a speed advantage in doing so. Slugs, who are content to travel at lower speeds, use the shoulder lane only. When vehicles discharge from a SP queue and into a multi-pipe regime, the direction of the kinematic wave signaling a speed recovery (whereby rabbits migrate out of the shoulder lane and accelerate) is governed by the discharge rate of rabbits.

According to the theory, rabbits will react to SDs in the passing lane(s) by immediately redistributing themselves across all lanes, and this will produce a congested SP queue in all lanes. If the discharge rate of rabbits is low, the recovery wave (i.e., the head of the SP queue) can move forward or remain stationary in space. Rabbits will migrate back to the passing lane(s) upon traveling through the recovery wave. The SP queue left behind can expand longitudinally in the upstream direction.

All of the above is qualitatively consistent with what was observed in the latter stages of the rush, as was described in Chapter 4, and under the provision that the median and center lanes constitute passing lanes. It bears mention that the observed transition to the SP regime required time; the transition was marked by the time-varying LC patterns noted in Observations 3 and 5. This time requirement may partially explain why the theory did not match the present observations made in the early stages of the rush, as discussed below.

The theory posits that when the discharge rate of rabbits is high, the recovery wave from a SD will propagate backward (upstream). In fact, backward-moving recoveries were observed in the early portion of the rush as is clear in Figures 4.3(b) and 4.3(c). However, the SDs and their backward-moving recoveries were confined to the passing lanes; i.e., the LC among rabbits in response to the early SDs did not induce a SP queue whereby low speeds spread to the shoulder lane.

This apparent discrepancy in the theory is not surprising, and can be explained in...
part by the release valve through which many rabbits were able to migrate leftward to bypass the oncoming SDs. Moreover, the longitudinal expanses of the early SDs were small, on the order of the characteristic length of an interface between traffic states, about 500 m (Daganzo, 2002). It seems that these SDs came and went before rabbits could redistribute themselves in sufficient numbers to create a stable equilibrium (i.e., a SP queue).

The postulate that a SP regime will invariably result from rabbit migration seems to be based on findings in Mika et al. (1969); see observation B3 of Daganzo (2002). Tellingly perhaps, these earlier findings relied upon detector data that were averaged over one-minute intervals. This sampling interval may have been too coarse to unveil the details observed in connection with the early SDs. Key details of traffic dynamics at the onset of congestion may not occur at scales that are well approximated by SP models, or by stationary traffic states.

Another explanation might be that there was an initial undersupply of rabbits, and that slugs did not confine themselves to the shoulder lane. To make this explanation fit the observations, three flavors of slugs would be needed, one for each lane. Slugs stubbornly choose a lane, and stay in that lane regardless of opportunities for greater speeds in other lanes. This feature would capture the real life tendency of drivers who sometimes hesitate to change lanes and therefore become temporarily stuck in a slow moving lane. This type of behavior might explain how, in the first stage of congestion, high speeds are sustained in the shoulder lane, and low speeds in the median and center lanes propagate upstream in the form of a SD. Despite a substantial number of observed LC maneuvers, no SP regime forms in this stage because there are simply not enough rabbits to equalize speeds across lanes. Such a model would differ from the original theory (Daganzo, 2002) in that stationary states in a semi-congested regime could exist at speeds below the desired speed of slugs, $v_f$.

5.3 Implications for Traffic Management

Any traffic control strategy that could postpone the formation of a SP queue and prolong the period of high output flows would benefit travelers on the present site and might be suitable for broader application as well. The site is equipped with a paved shoulder along the left edge of the expressway that is reserved for use by emergency vehicles. Under appropriate circumstances, this road space could also be used by regular traffic to clear the queue on the hill, and allow traffic to recover higher speeds and flows. Videos reveal that some drivers already adopt this unsanctioned use of the shoulder. One key consideration would be to determine where and under what circumstances vehicles in the shoulder would be allowed to return to the shoulder lane.
Recall that rightward LC maneuvers from the shoulder lane were observed to block leftward LC from the median lane. It might be advantageous to discourage premature rightward LC on the uphill segment into the more highly utilized lanes. Perhaps lane markings could be adopted to specify the direction of permitted (or recommended) LC maneuvers.

The onset of SDs and the resulting LC behavior may be related to line-of-sight limitations. The uphill segment curves rightward horizontally, and limits the line-of-sight more severely for drivers in the median lane, than it does for those in the shoulder lane. Unfortunately, it is the median-lane drivers who require greater vigilance to maintain the short headways that they adopt (on the order of 1.5 seconds). Furthermore, these drivers seem unable to respond gracefully to small perturbations, and thus they trigger the observed SDs. If downstream traffic conditions were communicated in real-time, as drivers approach and ascend the hill, perhaps the severity of output flow disruptions that result from SDs could be diminished.

Finally, the most effective method of management might be through pricing. The site used in this study is a tolled facility, and most vehicles are ETC equipped. It would be a relatively simple matter to charge an express fee for low trip times. Such a policy has the potential to shift the distribution of desired speeds, and alter the trend in lane utilization as traffic gets heavier. Starting from the uncongested state, as flow increases, speeds fall off. As speeds fall off in the center lane, more drivers switch to the median lane, eventually resulting in the chain of events leading to congestion. Perhaps a charge for high speeds could reduce demand for the median lane, and encourage behavior in which all lanes are effectively utilized.

5.4 Future Work

Ongoing and future work includes pursuing opportunities for experimentation and mitigation of known bottlenecks. Further empirical work is necessary for continued progress in this field. In particular, it is clear from the present work that LC and SDs occur together in the aggregate. As more data become available, attempts to incorporate LC into continuum traffic models will be more fruitful as a line of inquiry.

The expressway segment studied in the present work displayed strikingly different flows and speeds for each of the three different lanes. Lane utilization was always particularly unbalanced whenever flows were high. Since a SP continuum model will never capture dynamics that it does not model, this author argues that future continuum models must be multi-lane in nature.

One output from this work is a human-verified, internally consistent set of re-identified vehicles. This data set is a significant resource, not only for additional traffic studies, but for future efforts in video processing.
Finally, for lack of a better data collection method, future means for studying traffic phenomena will continue to involve multiple video cameras at a low angle. Efforts are ongoing to apply the tools to other types of sites.
Bibliography


