Length-Based Vehicle Classification Using Images from Uncalibrated Video Cameras

Ryan P. Avery, Yinhai Wang, and G. Scott Rutherford

Abstract—**Due to the marked difference in many characteristics between trucks and smaller vehicles, accurate and timely truck data are of significant importance. Unfortunately, few frequent and wide-area truck data are collected with the systems currently in place. Furthermore, the systems that are capable of truck data collection are typically expensive and limited in application. For this reason, wide-area truck data are typically collected every few years, although more timely truck data are desired. There is no doubt that continuous collection of truck data is beneficial to a variety of purposes. This paper presents an image processing algorithm for length-based vehicle classification using an image stream captured by an uncalibrated video camera. Although the current implementation separates vehicles based only upon length, the ultimate goal is to develop a system based upon the Highway Performance Monitoring System guidelines. The basis of the algorithm is to relatively compare vehicle lengths to each other to estimate truck volumes and eliminate the need for complicated system calibration. The algorithm was implemented in C#, a new programming language platform developed by the Microsoft Corporation. The system test revealed that the vehicle length classifications estimated by the algorithm do indeed satisfactorily resemble the actual observations. The proposed algorithm may enable the widely installed surveillance video cameras to count classified vehicles including trucks.**

I. Introduction

OVER the years it has become increasingly important for jurisdictions to keep track of vehicle travel patterns and volumes, especially as many agencies are shifting toward management of existing facilities rather than new construction. One of the most common means of achieving this aim is through the installation of inductive loop detectors. They have been around for several decades and have been established as the standard. If correctly installed and maintained, loop detectors are reliable and simple, which is why many organizations are not seeking something better. Nonetheless, loop detectors do have their shortcomings and advances in technology have led to many new creations that measure vehicle patterns and volumes just as accurately, but come with other very useful benefits.

The drawbacks to the traditional inductive loops are that they are intrusive and have to be cut into the pavement. This means that the roadways in which they are installed must undergo construction before vehicles can be counted. This also implies that there will be lane closures and added vehicle delay, which is costly for everyone. Another drawback is that wire loops have a fixed location; once put in place they cannot be adjusted to accommodate changing traffic conditions and more accurately measure traffic flow. Additionally, installation of loop detectors may decrease pavement life. As pavement wears and cracks loop detectors are subject to service interruptions and maintenance efforts, these actions cause further disruption and delay for vehicle traffic.

Video traffic detection via computer vision offers a solution to many of these problems and has received considerable attention in recent years. Lane closures are not necessary to implement this non-intrusive system, and little manpower is required for installation. However, when cameras are installed directly over traffic, a brief closure is generally undertaken to ensure safety; this closure is generally much shorter than that for loop installation. All that is needed is an elevated platform such as an overpass or a pole. A camera is then mounted on the platform and vehicle detection is performed via computer systems that may be local to the traffic detection area or located at a network-level traffic management center after the camera is calibrated. Typically, camera calibration requires measured road surface marks in the field of view and these marks may not be easily available. Previous research has demonstrated that calibration is also possible via knowledge of measures such as the camera elevation and tilt angle. The disadvantage of these methods is inflexibility – if the camera is to be adjusted, the calibration measurements must be retaken. Pumrin and Dailey [*1*] demonstrated that camera calibration without physical measurement is accurate enough for use in estimating mean vehicle speed. Programs such as the one discussed here require little calibration and offer the ability to adapt to countless different detection schemes.

Automated classification of trucks and heavy vehicles is important for several reasons. First, truck data are important for purposes such as transportation planning and

Manuscript received July 22, 2004. This work was supported in part by the U.S. Department of Transportation University Center (Federal Region 10) and by the Valle Scandinavian Scholarship and Exchange Program at the University of Washington.

traffic operations. Safety is another reason for truck data collection, since a high percentage of all fatal accidents involve trucks. Pavement design is also influenced by truck volumes as both pavement life and design thickness depend upon the number of heavy vehicles traveling on the roadway. Geometric design features are also affected by roadways with high truck volumes; curb height and horizontal alignment are two examples. Environmentally it is also important to have an estimate of the magnitude of heavy vehicles on a facility since they produce more air pollution than smaller vehicles. Finally, if such information were collected over a large network of roadways, the information could be very useful in freight transportation analysis and planning. Unfortunately, wide-area truck data are traditionally collected only every 3-5 years and therefore do not reflect the variation of truck volumes with time nor are they suitable for ATMS applications [*2*]. Under federal requirements for the Highway Performance Monitoring System, states are required to perform classified vehicle counts on freeways and highways and provide this information to the Federal Highway Administration (FHWA) every year. These data are often collected via pneumatic tubes and is done only once per year at a select location of sites. This again prevents any detailed freight analysis from occurring since seasonal effects and variation between test locations are not reflected in the HPMS results. While more detailed methods are available (such as weight sensors), these tend to be installed solely at weigh stations that typically serve as ports-of-entry into a state. These few locations again cannot represent the variation in truck volumes and distribution that urban areas may experience. Installation of more weight sensors would be expensive and cause increased urban delay unless all such scales were weigh-in-motion. Other detection methods such as single-loop detectors, although common, cannot directly measure truck volumes. Dual-loop detectors and video image processors are capable of truck data collection but are not widely used. Commercially available video image processors typically require complicated calibrations in order to collect useful data.

Alternatively, surveillance cameras have commonly been installed for monitoring major roadways. If these cameras can be used for the collection of traffic volume data including truck data, this represents a major gain in data collection with limited capital expense. However, surveillance video cameras, which are operated by traffic operators and whose views are often re-pointed, are typically not calibrated and hence are not capable of collecting traffic data automatically. To better utilize these video sensors, the authors propose a new algorithm for vehicle classification. A major advantage of the presented algorithm is that the calibration is automatic and requires no manual calibration of points or objects. The following sections provide a brief overview of computer vision technology regarding vehicle detection, detail the methodology of the proposed truck classification, provide experimental results testing the proposed algorithm, and conclude with a description of the proposed system.

II. State of the Art

Computer vision is not an entirely new concept for vehicle detection; many agencies began investigating the possibilities of video detection 15 or more years ago. The first systems, however, were unable to function adequately under a variety of environmental conditions. Shadows affected detection, nighttime detection was troublesome and poor weather obscured vehicles. Thus, many agencies continued to use loop detector systems or considered other detection technologies, such as radar [*3*]. Over the years, however, many improvements have been made as advances in video image processing have been applied in the traffic detection arena. Commercial systems are currently available that are capable of eliminating false detection of shadows and function in weather to a significant degree of severity. These systems are field-proven and offer the advantage of preserving a continuous stream of information rather than the recording of discrete vehicle passages in other detection systems. Unfortunately, these commercial systems are fairly expensive to install and typically require concurrent installation of proprietary hardware and software, especially for intersection video traffic detection. Proprietary equipment prevents agencies from modifying or improving the algorithms used in traffic detection to better suit their needs. Although some vendors do allow for flexibility in hardware selection, the software remains immutable in its treatment of traffic detection and underlying assumptions [*4*]. This paper was written to offer a new and inexpensive means of vehicle detection and classification. With the goal of requiring only a standard video image stream, no proprietary equipment is necessary for the use of this software.

There have been several recent investigations into vehicle length measurement via computer vision. Lai, Fung, and Yung [*5*] demonstrated that accurate vehicle dimension estimation could indeed be performed from a single camera angle through the use of a set of coordinate mapping functions. Through the use of a shadow removal method (important to maintain the true vehicle dimensions) and convex hull to produce a vehicle mask, they were able to estimate vehicle lengths to within 10% in every instance. Their method, however, requires camera calibration in order to map image angles and pixels into real-world dimensions. Gupte, Masoud, Martin, and Papanikolopoulos [*6*], performed similar work by instead tracking regions and using the fact that all motion occurs in the ground plane to detect, track, and classify vehicles. Before vehicles may be counted and classified, their program must determine the relationship between the tracked regions and vehicles (e.g. a vehicle may have several regions or a region may have several vehicles). In a

20-minute trial of the program 90% of all vehicles were properly detected and tracked, and 70% of those vehicles were properly classified. The approach chosen for this program, however, is based on relative length comparison between vehicles. Relative comparisons to determine vehicle classification have been proven effective by Wang and Nihan [*7*] in development of more accurate speed estimation for single loop detectors. It is feasible in this application because the only desire is to classify vehicles by length; it is not necessary to know the actual length of each vehicle so long as it is properly classified. The principal assumption is that there exists a significant difference in mean vehicle length between regular vehicles and long vehicles (trucks). Wang and Nihan [*7*] illustrated that for a typical freeway location on I-5, the distribution of vehicle lengths clearly indicated a bi-normal distribution with two distinct peaks. Shorter vehicles had a mean length of 5.5 m (18.0 ft) and a standard deviation of 0.9 m (3.0 ft), while long vehicles had a mean length of 22.5 m (73.8 ft) and a standard deviation of 3.6 m (11.8 ft). This clearly illustrates that such an assumption regarding characteristic vehicle lengths is valid.

The advantage of this new approach is that it is less demanding computationally than many other approaches. Furthermore, it does not require calibration since all comparisons are relative, which is useful when developing a mobile traffic detection system fit for various different applications.

III. Algorithm and Implementation

This section describes the logic and reasoning behind the development of the computer vision vehicle detector and classification by length. It is separated into two distinct processes: background extraction and vehicle detection. A description of the various additional program features follows the discussion of the specific algorithms implemented in the design.

A. Background Extraction

Typically, a computer vision based detection system requires a background image that represents the base state of the area under observation. In the case of traffic detection, it is rarely possible to obtain an image of the observation area without vehicles that also maintains similar lighting and environmental qualities. Thus, it is necessary to extract the background image from the video stream itself. This is accomplished in an iterative fashion using the pixels that make up an image. Each pixel contains a red, green, and blue component that, when combined, represent the color and brightness of that particular cell of the image. Each pixel in the current image is compared to the same pixel location in a previous image. If the absolute differences between the red, green, and blue color amounts (hereafter referred to as RGB values) in that pixel location for the two images are within the desired threshold level (a threshold of 10 is used in this paper), the pixels are similar. RGB values are based on a unit-less scale from 0 to 255; 0 represents no presence of that hue while 255 represents full presence of that hue. Since the images were taken at different times, it is likely that the only instance the pixels at a location would be similar is when a vehicle is not occupying that pixel in either frame. This implies that the given pixel represents a part of the background image, and is then locked as a part of the background image and is no longer checked in the next iteration. This process is repeated until a sufficiently large percentage (99.95%) of pixels have been confirmed as being part of the background. At this point the background image is saved for use in vehicle detection. Currently the background detection is run upon program execution and whenever requested by the user.

B. Vehicle Detection

There are several key considerations when implementing any detection algorithm, and they vary depending on the specific task. First, when counting vehicles, it is important to count each vehicle only once. To this end, the program discussed in this paper utilizes a line referred to hereafter as a registration line. This ensures that vehicles will only be counted as they pass over the line. When measuring the length of a vehicle, the line along which vehicles are measured, known hereafter as the longitudinal line, must be placed along the line of travel for that lane. Also, it is advisable to place both the registration and longitudinal lines near the foreground of the image to optimize the detection algorithms since vehicles in the foreground are better defined and occupy more pixels than those closer to the horizon. Figure 1 demonstrates an ideal detector setup and illustrates a few of the terms discussed above. The description of the vehicle detection algorithm is separated into two sections: vehicle count and length classification.

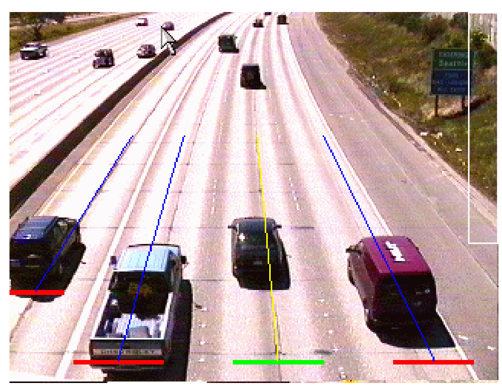

Fig. 1. Example of detector setup.

1) Vehicle Count

The vehicle count algorithm is performed whenever the program is set to detect vehicles and at least one registration line is turned on. During each frame, the RGB values of each pixel on a registration line (the line upon which vehicles are counted) are compared to the RGB values of each pixel in the same location in the background image. If the absolute difference of all RGB values is lower than the threshold (20 is used in this paper), the pixels are considered similar and no vehicle is detected. Alternatively, if the values exceed the threshold, the pixels are considered different. The threshold of 20 was chosen through calibration of the program as a balance between being small enough to detect vehicles that are colored similar to the background image while not being so sensitive as to detect many pixels that are not vehicles. If 40% of the pixels on a registration line are found to be different, the program considers a vehicle to be present. This parameter was again chosen through calibration and is based upon requiring enough pixels to be different to ensure the presence of a vehicle but also allow for detection of smaller vehicles that would obviously not cover the entire registration line. In order to prevent over-counting vehicles, a vehicle is counted only if no vehicle was present in the previous frame. This is reasonable in even very congested situations because the lower speeds in those situations means that a vehicle does not travel as far per frame. This still enables the space between vehicles to be detected. In stopped traffic the detector will not over-count either, which is an advantage over loop detectors that do tend to overestimate the vehicle count in such situations. Each registration line changes from green to red whenever a vehicle is present on the line to provide a visual cue to the observer. The output provided in this algorithm is the total count of each detector, the total number of unique vehicles detected in the current frame, and the sum count of all vehicles detected.

2) *Length Classification*

For the purpose of this paper, a long vehicle or truck is considered to be any vehicle with a length exceeding 40 ft (12.2 m). This value was chosen because work performed by Wang and Nihan [*7*] indicated that this is a reasonable value to break the bimodal vehicle length distribution into two subpopulations, short and long vehicles. Also, this is consistent with the loop detection system implemented by the Washington State Department of Transportation (WSDOT) that uses this value as the boundary between short vehicles (Bin1&Bin2) and long vehicles (Bin3&Bin4). The length classification algorithm runs whenever the program is set to detect vehicles and at least one longitudinal line (the line along the direction of vehicle travel upon which vehicle lengths are measured) is turned on. In addition, length classification may only be performed on a lane when counting is also enabled on the same lane. When a vehicle exits the registration line corresponding to the longitudinal line (that is, the first frame a registration line is unoccupied after being occupied for at least one frame), the length classification algorithm is run to measure the length of the vehicle in pixels. This makes the lengths of all the vehicles in a lane be measured at almost same starting point so that the measured lengths are comparable. Note that this is a relative length, and does not represent the actual length of the vehicle. The length algorithm merely steps along the longitudinal line counting the number of different pixels. A minimum of five consecutive pixels must be different from their background values to begin counting. Five consecutive pixels must also be similar to their background values to stop counting. This five pixel threshold was chosen to be long enough to prevent noise from being picked up as part of the vehicle length while also ensuring that short vehicles such as motorcycles can be picked up.

Once a vehicle's length in pixels is obtained, it is stored in an array for later comparison with other vehicles. Fifteen vehicles must be measured before the array is filled. Fifteen was chosen in order to provide reasonably frequent updates of long vehicle counts. In future work with longer samples a higher number may be chosen. Once filled, the array of lengths is sorted in ascending order. Vehicle lengths shorter than a third of the length of the longest vehicle are rejected as short vehicles that were improperly measured or the result of detection error. This is permissible because although a long vehicle may physically be four times as long as a short vehicle, the camera angle and perspective distort this relationship to the point where long trucks only appear to be 1.5-2 times longer. Furthermore, since the algorithm is used only to count long vehicles that have already been detected, the incidental rejection of an abnormally short but valid vehicle measurement such as a motorcycle will have no impact on the resulting calculation of long vehicles. Rejection of these short vehicles focuses the variation in the sample on longer vehicles.

From the remaining vehicle lengths, the mean, standard deviation, and range are calculated. These measures are used as characteristic measures rather than trying to fit the data to a normal distribution. Wang and Nihan [*7*] demonstrated that vehicle lengths follow a bimodal distribution. However, since the vehicles are not yet separated into their respective populations, this algorithm does not take advantage of that fact. Situations in which the range is greater than 75% of the mean vehicle length are considered to contain trucks. In this case, the vehicle lengths greater than one standard deviation above the mean are counted as trucks. When all lengths in the array are checked and trucks are recorded, the array is cleared for future use by the next 15 vehicles. Note that this process of relative comparison means that truck counts for a lane are not provided until 15 vehicle observations have been made. Users are cautioned from significantly changing the position of longitudinal line when running the program; pixel length is greatly influenced by both the camera angle and the longitudinal line's position near the foreground or horizon of the picture, and modifying the position of the line affects the relative comparison to determine long vs. short vehicles. The output of this algorithm is the total count of all long vehicles (trucks) detected.

C. Program Features

There are several features built into the program that aid in the detection of vehicles. First, the program interface affords the user the capability to use up to eight vehicle detectors. The registration and longitudinal lines of each detector may be adjusted to fit a variety of different locations and situations. For a given detector, the longitudinal line may be disabled so as to only maintain a count of traffic in that zone. An automatic gain control (AGC) area is also used to control for rapid lighting changes. The AGC area used for global light changes is bounded by a white box and is also capable of being adjusted by the user. It is important to note that this AGC area must be placed in areas of the picture where the background always remains the same (e.g. no vehicles pass over the area) in order for the filter to work accurately. This is because the AGC works by measuring the RGB differences between the bounded area in the presented image and the background image. The average difference is calculated from this area, which can only represent environmental effects since the pixels in this area do not change due to physical effects. This average lighting affect is used to adjust the detection of vehicles to filter out lighting effects and thereby more accurately detect the presence of a vehicle.

Another option available to the user is the ability to choose to view a pixel map rather than the regular bitmap image, as illustrated in Figure 2. During background extraction, the pixel map illustrates confirmed pixels as blue and unconfirmed pixels appear white. If instead used during vehicle detection, the pixel map illustrates what the computer “sees” by coloring pixels similar to the background image blue and coloring those different from the background image white. Note that during vehicle detection the pixel map illustrates only a portion of the foreground in order to limit the drain on computer resources. Also, although the detection lines do not appear on the pixel map, vehicle detection is still performed. Users may reset all counts to zero by selecting the reset button.

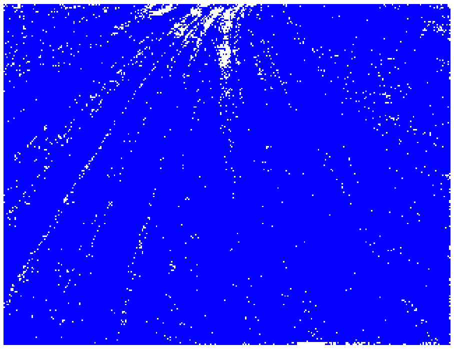

Fig. 2. Pixel map display.

IV. Experimental Results

Evaluation of the system was performed using a video stream digitized into bitmap images files. The video was taken between 11:30 AM and 12:30 PM on I-5 at 145th Street in Seattle, Washington, on June 11th, 1999. The stream consisted of 4500 frames at a rate of 15 frames per second, resulting in 300 seconds (5 minutes) of traffic observation. Given the camera location at this site, vehicle occlusion was rare. Also, although the weather was sunny, the time of day during which the video stream was taken resulted in shadows that did not tend to stray into other lanes; this would otherwise have produced spurious vehicle counts. Thus, some of the typical issues associated with video image processing (effect of shadows, weather, vehicle occlusion, etc) that were not apparent in the media have been left unresolved in order to instead develop the capability of vehicle length classification. Selection of this site therefore enabled these issues to be addressed at a future point when the system is applied to other locations.

Table 1 summarizes the results from the test video stream. There was an overall detection error of only 2.5 percent, and trucks were properly identified approximately 92 percent of the time. Table 2 tabulates all the detection errors encountered in the test. From these errors it is noted that improvements should be made to avoid vehicle occlusion (especially on the outside lanes) and improve the light filter. Additionally, improvements should be made to better detect merging and/or small vehicles. It should also be noted that loop detection systems perform no better in lane merge situations and often may not detect motorcycles

TABLE I

SUMMARY OF RESULTS

Vehicles Detected				
Lane	**Actual**	**Program**	**Error**	**Accuracy**
HOV	62	69	7	88.71%
GP Lane 3	154	154	0	100.00%
GP Lane 2	134	134	2	98.51%
GP Lane 1	94	94	2	97.87%
Total	**444**	**451**	**11**	**97.52%**
Trucks Detected				
Lane	**Actual**	**Program**	**Error**	**Accuracy**
HOV	1	1	0	100.00%
GP Lane 3	8	8	0	100.00%
GP Lane 2	20	17	3	85.00%
GP Lane 1	8	8	0	100.00%
Total	**37**	**34**	**3**	**91.89%**

as well. There were also three errors in truck classification. Seven large trucks were included in one of the 15-vehicle groups for lane two. With such a large number of trucks, the mean vehicle length was driven high enough that in processing the vehicles only the four longest trucks were recognized properly. This problem can be solved by allowing each detector to keep a longer memory of the distribution of vehicle lengths rather than starting fresh every fifteen vehicles. This will also allow instantaneous classification of a presented vehicle after enough vehicles have been detected in a lane to provide a reasonable approximation of the vehicle length distribution.

Future testing of this prototype system should include testing on several different locations and camera positions to demonstrate adaptability. Considerable time will need to be spent on testing the sensitivity of the calibration parameters discussed in the algorithm development. Another future step is to remove vehicle shadows in order to more accurately model the true length of each vehicle. This is important because taller vehicles cast longer shadows which may skew the length estimation to an unsatisfactory level. Additionally, the changing position of the sun may result in shadow occlusion which produces false vehicle detection in neighboring lanes.

Ultimately a much more detailed classification of vehicles is desired, such as one that uses the Federal Highway Performance Monitoring System (HPMS) as a basis for vehicle classification. To accomplish this goal, robust pattern matching algorithms will have to be developed. In order to maximize the intelligence of the program, it is desired to enable the program to "learn" patterns for vehicles to better match a presented vehicle to a predefined category. This type of intelligent adaptation can also be applied to the length algorithm to classify vehicles by length into more categories as well as update the vehicle counts immediately after the presentation of the vehicle. This method also gains the obvious advantage of using knowledge from all previous vehicles to make a decision rather than by segmenting the population into bins independent of each other. Another improvement involves updating the background image when the AGC detects large enough lighting changes over a specified period of time. This update should be performed while vehicles are being detected so as to avoid any gaps in data collection.

TABLE II
TABULATION OF ERRORS

Lane	Frame	Error	Description
HOV	1360	1	Occlusion from neighboring lane
HOV	1722	1	Occlusion from neighboring lane
HOV	2058	1	Occlusion from neighboring lane
HOV	2389	1	Unidentified lighting effects
HOV	2686	1	Occlusion from neighboring lane
HOV	3040	1	Occlusion from neighboring lane
HOV	3046	1	Occlusion from neighboring lane
GP Lane 2	1258	-1	Missed lane change
GP Lane 2	2387	1	Unidentified lighting effects
GP Lane 1	3172	1	Unidentified lighting effects
GP Lane 1	3884	-1	Missed lane change

V. CONCLUSIONS

Wide-area truck data are seldom collected in most transportation agencies today. Collection of frequent long vehicle data offers many advantages to transportation agencies regarding travel patterns, safety, seasonal variations, roadway design, and transportation planning. It is clear that the method presented in this paper for truck detection via relative length comparison is indeed a viable alternative for data collection, as illustrated by the truck classification accuracy of nearly 92%. Furthermore, it is advantageous to utilize existing surveillance equipment in a new manner to offer a new rich source of data at minimal cost. The benefits of flexibility in algorithm design and ability to use uncalibrated video cameras further expand the usefulness of this prototype system. Continued investigation into applications and development of the algorithms presented here show promise of development of accurate new vehicle detection and classification systems that have the potential to be used in many jurisdictions not only in the United States, but around the world.

ACKNOWLEDGMENT

The authors appreciate Mr. Hu Dong who participated in developing the frame of this prototype system.

REFERENCES

[1] S. Pumrin and D. J. Dailey, "Dynamic Camera Calibration in Support of Intelligent Transportation Systems," in *Transportation Research Record*, Issue 1804, 2002, pp. 77-84.

[2] Y. Wang and N. L. Nihan, "Dynamic Estimation of Freeway Large Truck Volume Based on Single Loop Measurements," in *ITS Journal*, May 2002, pp. 1-4.

[3] A. N. Weber, "Verification of Radar Vehicle Detection Equipment Study SD98-15 Final Report," South Dakota Department of Transportation, 1999.

[4] Trafficon, *Product Demonstration*, More Hall Room 234, University of Washington Campus, Seattle, WA, June 4, 2003.

[5] A. H. S. Lai, G. S. K. Fung, and N. H. C. Yung, "Vehicle Type Classification from Visual-Based Dimension Estimation" in *IEEE Intelligent Transportation Systems Conference Proceedings*, Oakland, CA, August 2001, pp. 201-206.

[6] S. Gupte, O. Masoud, R. F. K. Martin, and N. P. Papanikolopoulos, "Detection and Classification of Vehicles," in *IEEE Transactions on Intelligent Transportation Systems*, Vol. 3, No. 1, March 2002, pp. 37-47.

[7] Y. Wang and N. L. Nihan, "Can Single-Loop Detectors Do the Work of Dual-Loop Detectors?" in *ASCE Journal of Transportation Engineering*, 129(2), 2003, pp. 169-176.