

# STANDOFF VIDEO ANALYSIS FOR THE DETECTION OF SECURITY ANOMALIES IN VEHICLES

Satyam Srivastava and Edward J. Delp

Video and Image Processing Laboratory (VIPER)  
School of Electrical and Computer Engineering  
Purdue University  
West Lafayette, Indiana, USA

## ABSTRACT

Video surveillance systems are commonly used by security personnel to monitor and record activity in buildings, public gatherings, busy roads, and parking lots. These systems allow many cameras to be observed by a small number of trained human operators but suffer from potential operator fatigue and lack of attention due to the large amount of information provided by cameras which can distract the operator from focusing on important events.

In this paper, we propose the design of an autonomous video surveillance system which can operate from a standoff range that analyzes approaching vehicles in order to detect security anomalies. Such anomalies, based on dynamic analysis of the vehicle tracks, include unexpected slowing/stopping or sudden acceleration, particularly near check points or critical structures (e.g. government buildings). A human supervisor can be alerted whenever a significant event is detected and can then determine if the vehicle should be further inspected. Besides dynamic analysis, the system also estimates physical information about the vehicles such as make, body type and tire size. We describe low-complexity techniques to obtain the above information from two cameras.

## 1. INTRODUCTION

Due to the increase in volume of vehicles on the road, traffic monitoring has become both important and difficult. A traffic monitoring system together with an adequate response process would not only make the roads safer but can also potentially disrupt criminal and/or terrorist activities. Most technology solutions for vehicular monitoring involve significant human supervision or are passive forensic tools. Therefore, automated methods are needed to screen traffic in a non-obstructive manner and detect anomalies. Notable efforts at devising such systems include the Video Surveillance and Monitoring project (VSAM) [1], the European AVITRACK project [2], and the work by Coifman *et al.* [3].

In this paper, we describe a system that can autonomously observe vehicular traffic from a standoff range of approximately 50 feet (15m) and classify the vehicles in terms of “normal” behavior. The system is primarily based on video analysis but also provides interfaces for synergic operation with other types of sensors (including rumble strips for sensing mechanical abnormalities and synthetic

This work was funded by a grant from the US Department of Homeland Security and the US Naval Research Laboratory. Any opinions, findings, and conclusions or recommendations in this paper are those of the authors and do not necessarily reflect the views of the sponsors. Address all correspondence to E. J. Delp (ace@ecn.purdue.edu).

aperture radar). There are two premises on which this system is based. First, in most situations there is a logically acceptable range of behaviors that would be considered normal. If the response of an individual (vehicle) deviates from this range the behavior is considered anomalous. Secondly, there are many traits of anomalous behavior which can be ignored if they occur in isolation. However, the occurrence of multiple such behavior patterns is more likely to signal a potentially anomalous situation. This is described in more detail below.

## 2. OVERVIEW OF THE PROPOSED STANDOFF SURVEILLANCE SYSTEM

We begin this section with examples illustrating the two hypotheses of our approach. In normal circumstances vehicles are driven within the lane boundaries and lane changes are only occasional. However, a vehicle repeatedly ignoring the lane markings may be indicative of drunk/impaired driving or of road rage. Therefore, it is possible to identify anomalous driving behavior, in this case by observing the frequency of lane departures. Similarly, presence of relatively flat tires may not convey much information alone. But, a vehicle with visibly flat tires and moving below the speed limit could be overloaded and pose a threat. We select certain traits associated with a vehicle and define normal and abnormal “observations.” These can be later used to identify anomalies.

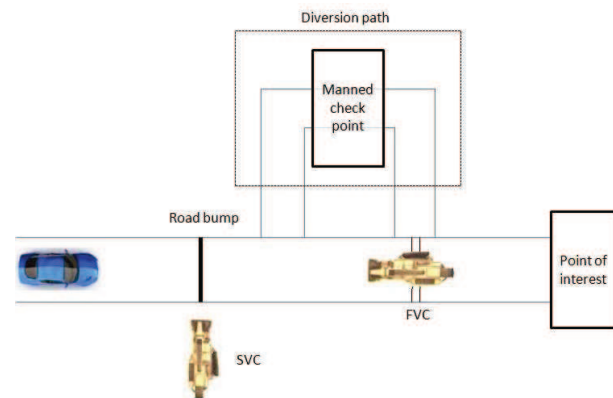


Fig. 1. An overview of the standoff surveillance system deployment. The distances are not drawn to scale.

An overview of our proposed system is illustrated in Figure 1.

The screening system can be deployed near the entrance to a point of interest (such as airport, sports venue, or a public building). The goal is to observe the vehicles without affecting the traffic flow. Our system uses two video cameras – the front view camera (FVC) and the side view camera (SVC). The actual position of these sensors is determined based on the desired field of view and camera resolution. It should be noted that the FVC is located at a fixed height above the road surface (although this is not evident in the top-view schematic) to allow sufficient clearance for vehicles.

The selection of two cameras and the particular configuration is described as follows. In our system, the purpose of multiple cameras is to generate complementary information (rather than redundant information). Thus, for a cost-effective design, the number of cameras should be small. It can be argued that a two-camera system provides considerably more information than a one camera system. However, the incremental gain with a three camera system is small. Furthermore, the front and side views provide more important information than alternative configurations of a two camera system. For example, a rear-view (instead of front-view) allows license plate detection but eliminates the scope for driver behavior monitoring.

For any vehicle, the system would produce one of the three “actions” – **Pass**, **Monitor**, and **Stop**. The vehicle designated for ‘Stop’ would be diverted for secondary screening while the ‘Pass’ vehicles will be allowed passage unobstructed. While an ideal system should produce minimum ‘Monitor’ labels, in practice, the passage of certain vehicles may need to be determined based on the application. In certain areas, all ‘Monitor’ vehicles may be screened; in others, they may all be allowed passage. The system would also enable human operator in making such a decision by providing them with the information associated with the vehicle. The exact mechanism for vehicle diversion and secondary screening is not discussed in this paper.

Using the two cameras we extract, using image analysis methods described in the next section, various information about the oncoming vehicles that can be placed into two categories.

**Physical Information:** The data obtained here is primarily used to characterize and/or identify a vehicle. This information includes vehicle body type, tire size, and make.

**Behavioral Information:** This set of information includes high-level analysis of the vehicle’s appearance and motion. It includes analysis of the velocity and trajectory of the vehicle. This information is directly used for assigning an action label for the vehicle.

### 3. IMAGE PROCESSING AND ANALYSIS

We can organize the various types of information extracted and/or inferred from our image analysis described below into an approximate hierarchy. This is shown in Figure 2. Note that certain information extracted from one camera is used to analyze the images from the other camera, thus highlighting the usefulness of the orthogonal camera configuration. A typical sequence of operations when a vehicle approaches would be as follows. The name of the camera which is primarily used for making a particular detection is provided inside parentheses. The image analysis methods used to perform these tasks are described in more detail below.

1. An oncoming vehicle is detected. (FVC)
2. The vehicle tracking and trajectory analysis systems are activated. An unexpected slowing/stopping may lead to the vehicle being marked for ‘Monitor’ while a sudden acceleration may result in a ‘Stop’ label. These decisions are based on the assumption that the driver may be searching for ways to evade the security systems or may be intending to crash through

them resulting in these, respective, patterns in the vehicle’s approach. (FVC)

3. When the vehicle enters the SVC’s field of view, its body type is determined. (SVC)
4. The vehicle’s tires are segmented and the tire size is estimated. (SVC)
5. The make of the vehicle is determined. (FVC)
6. The gap between the tires and the vehicle body (wheel well) is estimated and the front and rear values are compared. If the gap above the rear tire is significantly smaller, it can indicate a heavy load in the trunk which may lead to a ‘Monitor’ label. (SVC)

We now discuss the image analysis techniques to achieve the above tasks with a focus on low-complexity operations. Such low-complexity methods would better allow the system to be deployed for real-time operation. Note that some of the methods are described only briefly in this paper. Details can be obtained from the referenced works.

#### 3.1. Vehicle Detection

The task of vehicle detection can be viewed as one of figure-ground separation. Background subtraction is a popular approach to object detection and several enhancements have been proposed to the basic algorithm. A background subtraction operation consists of comparing the current video frame with a background model with respect to some properties like pixel intensity and color. The difference between the values is used to classify a pixel based on a decision threshold.

Let the current frame be represented by

$$C = \{c(i, j) : i = 0, 1, \dots, W - 1; j = 0, 1, \dots, H - 1\}. \quad (1)$$

Here,  $c(i, j)$  represents the value of the gray level at the pixel located at  $(i, j)$ . The frame is considered to be of width  $W$  and height  $H$ . Similarly, we can represent the background model as  $B = \{b(i, j)\}$ . Then, the output of background subtraction is a spatial mask  $F = \{f(i, j)\}$  defined as follows:

$$f(i, j) = \begin{cases} 0.0, & \text{if } |c(i, j) - b(i, j)| \leq \tau, \\ 1.0, & \text{otherwise.} \end{cases} \quad (2)$$

The symbol  $\tau$  represents a decision threshold.

Many background subtraction techniques allow modification of the background model and the decision threshold after each frame. This allows the system to correctly classify pixels even when the background is not truly static. These update tasks may be performed statistically [4, 5]. One can also formulate vehicle detection as a moving object detection problem and use techniques based on motion cues [6]. We use a background subtraction method which updates the background model and decision thresholds using the current classifications [7].

The output of vehicle detection can be viewed as a silhouette of the vehicle. The silhouette also determines the vehicle’s exterior dimensions in image co-ordinates. By mapping these to ground co-ordinates, the true dimensions of the vehicle can be estimated. These transformations can be carried out with the use of calibrated cameras or by utilizing objects of known dimensions (fiducial markers). Some systems use the known size and separation of lane markers as reference to achieve co-ordinate transformation for speed estimation [8]. Perspective correction [9], if needed, would also be included in this process. We use a look-up table based approach which is described in Section 3.6.

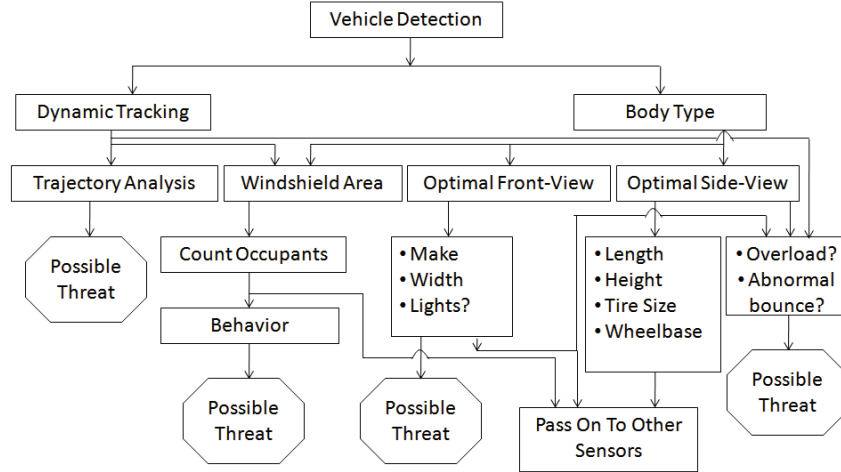


Fig. 2. An illustration of inter-dependencies in the information extracted from the cameras.

### 3.2. Vehicle Body-Type Determination

Determination of a vehicle’s type is a part of many traffic monitoring systems where it may be used to estimate the temporal composition of traffic. Some traffic monitoring systems are based on inductive loops and other non-video methods [10]. Video-based methods have also been proposed for the problem. In [11], a statistical model-based approach is used to detect and classify vehicles using a two stage process. Lai [12] describes a traffic analysis system using a video-based virtual loop technique which also classifies vehicles.

We use a template-based shape matching approach to classify a vehicle into one of four types – sedan, truck, SUV, and hatchback. Let  $S = \{s(i, j)\}$  be the silhouette of an oncoming vehicle as obtained from the foreground mask  $F$ . Also, let us represent the silhouettes of the template vehicles as  $T_1, T_2, T_3, T_4$ , where the subscripts correspond to the four classes above respectively and  $T_i = \{t(i, j)\}$ . These templates are obtained by thresholding selected images in which the vehicles differ considerably from a neutral background. The silhouettes are used to compute the sum of absolute difference (SAD) for each vehicle type:

$$\Delta_k = \sum |s(i, j) - t_k(i, j)|, \quad (3)$$

where the summation is over all pixels indexed by  $(i, j)$ . The class corresponding to the smallest SAD is selected.

### 3.3. Tire Size Estimation

Most traffic surveillance systems lack the capability to accurately segment the tires and estimate their size, but this is important to us. This information, along with the make (described next) and the exterior dimensions, can help in identifying the vehicle’s model. Secondly, we use the position of the tire centers to estimate the gap

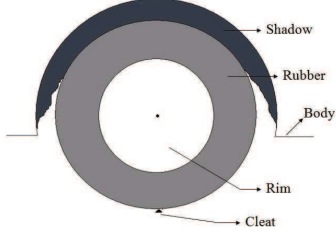
between the tires and the wheel well (Section 3.5). We use this to detect possible overloading. We now describe a method to segment the tires and determine their size.

There are two major challenges in accurate segmentation of a moving vehicle’s tires – the lack of information about both the position and size of the tire, and the different shapes of tire hubs. We note that despite the variation in the shapes of the hub, most tires have one common feature which is the prominent circular edge between the wheel and the rubber. Since most vehicles have either hub caps or alloy wheels, this observation holds true for them. The notable exceptions include vehicles with older wheels and missing hubcaps. Such cases can be handled separately if additional assumptions are made.

Here we describe the case of tires with visibly distinct hub. This situation is illustrated in Figure 3. Our goal is to accurately locate the circular edge because it provides us the knowledge of the tire’s center and the radius of the wheel. For the present discussion, consider (say) the front tire of the vehicle as seen during a particular video frame  $V_{SVC}(i, j)$  where  $(i, j)$  define the spatial position of a pixel in the frame. Let  $(x_t, y_t)$  be the true location of the front tire center in the frame and the true tire radius be  $r_t$  pixels. Further, let  $r_w$  be the true radius of the wheel as visible from a side view.

We model the wheel-rubber edge as concentric white and black circles on a neutral background as illustrated in Figure 4. The thickness of the circles  $t$  is kept constant whereas the radius can be changed during the estimation process. Let  $r$  be the outer radius of the white circle. Then, at the end of the processing,  $r$  would be the method’s estimate of  $r_w$ . The estimation is performed by searching for the edge model of a chosen radius at different positions in the frame.

Let the tire model be represented by  $T_{r,p,q}$  where  $r$  is the radius of the wheel and  $(p, q)$  is the hypothesized position of the wheel



**Fig. 3.** A graphical illustration of the system's view of a tire with bright hub.



**Fig. 4.** In the case of a tire with shiny hub, the wheel-rubber edge is modeled as concentric circles of chosen radii.

center. In order to superpose the edge model on the frame, we can define the model for the same width and height as the frame:

$$T_{r,p,q}(i,j) = \begin{cases} 0.0, & \text{if } r - t < \|(p,q) - (i,j)\| \leq r, \\ 1.0, & \text{if } r < \|(p,q) - (i,j)\| \leq r + t, \\ 0.5, & \text{otherwise.} \end{cases}$$

Then, the error associated with this set of values for  $(r, p, q)$  is computed by

$$\tilde{\delta}_{r,p,q} = \sum |V_{SVC}(i,j) - T_{r,p,q}(i,j)|,$$

where the summation is performed over all pixels in the frame for which the edge model is not neutral. That is, the difference is computed only if  $T_{r,p,q}(i,j) \neq 0.5$ . Finally, the best estimate of wheel parameters is given by,

$$(r, p, q)_{optimal} = \arg \min \{\tilde{\delta}_{r,p,q}\}.$$

It should be noted that the ranges of  $(r, p, q)$  to be searched over can be intelligently limited by using information about the vehicle's body type (as described in Section 3.2).

Once the position of the tire center is determined, the tire radius is determined by dropping straight lines away from the center. Multiple estimates of the radius are made by measuring the length of each such line between the center and the rubber-background edge. We drop lines only in the lower half of the tire and consider only the lines which are longer than  $r_w$ . Then, the best estimate of the tire radius is the shortest such length.

### 3.4. Make Recognition

We can classify the various make recognition techniques into two categories – the ones based on feature matching and the ones based on appearance matching. In [13] Petrovic and Cootes describe the

use of gradient features for vehicle identification. Methods based on scale invariant feature transform (SIFT) and oriented contours have also been used [14]. Appearance based methods use pixel intensities and provide more global features. Zafar discusses the use of appearance matching using principle component analysis (PCA) and two-dimensional linear discriminant analysis (2D-LDA) [15].

We determine the make of a vehicle without the use of any reference objects (such as front license plate) to obtain a region of interest (ROI). This is achieved by selecting a part of the vehicle image likely to contain the front grille based on the knowledge of its body type. For example, we observed that most trucks have a tall front grille compared to most sedans. We estimate edge features for the front grille area using the Sobel operator [16]. More sophisticated edge operators can be used, but since our goal is low complexity operation the simpler operator is used.

Let the horizontal and the vertical gradients at a pixel  $(i, j)$  be given by  $c_x(i, j)$  and  $c_y(i, j)$ . Then, the magnitude of the edge at  $(i, j)$  is given by

$$\sigma = \sqrt{c_x^2(i, j) + c_y^2(i, j)}, \quad (4)$$

and the orientation is given by

$$\theta = \arctan c_y(i, j)/c_x(i, j), \quad (5)$$

where  $-\pi/2 \leq \theta \leq \pi/2$ . This edge information is used to construct a  $N$ -point histogram of edge orientations which serves as the “make signature.” Thus, we obtain the histogram of the unknown vehicle and those of vehicles with a known make. Let these be represented by  $h_{test}$  and  $h_r$  where  $r \in \{1, 2, \dots, k\}$  and  $k$  makes are considered. The correlation coefficient is estimated as the measure of similarity between the unknown and the known histograms. Note that we normalize the histograms by the number of pixels in the region of interest to account for the unequal sizes of the ROI. Thus, we obtain the normalized histograms  $\tilde{h}$  for each histogram  $h$  being considered. The correlation coefficient is estimated as follows:

$$\rho_r = \frac{\tilde{h}_r^T \tilde{h}_{test}}{\sqrt{\tilde{h}_r^T \tilde{h}_r} \sqrt{\tilde{h}_{test}^T \tilde{h}_{test}}} \quad (6)$$

Then, the most likely make is given by,

$$\alpha = \arg \max_r (\rho_r), \quad (7)$$

where  $\alpha$  is the index number associated with each car maker.

### 3.5. Static Overload Detection

Addition of a large load into a vehicle's trunk causes two visible effects. The rear tires might appear more flat than the front tires, and the gap between the rear tire and the body (the wheel well) may appear smaller than that between the front tire and the body. While the flatness of the tires can be easily compensated by over-inflation, the lowering of the body cannot be avoided without modifying the suspension. Therefore, we focus only on measuring the gaps.

Similar to the problem of accurate tire extraction, the problem of visual determination of rear-heaviness has not been explored sufficiently. Recall Figure 3 where the side view of a vehicle's tire is illustrated. The region above the tire, marked as shadow is the gap we are interested in characterizing. Let us denote this value by  $\delta$ . We compare the  $\delta$  values for the front and rear tires. This approach is more appealing than comparing against a database of  $\delta$  values because it does not require precise knowledge of the make and model

(and the normal values associated with it), and the decision threshold can be specified directly in pixels, thus avoiding additional errors that may arise from the co-ordinate transformation.

Our approach to determining the thickness of the gaps is by using the information about the tire centers. Consider for example the rear tire of a vehicle which has been analyzed by the tire size estimation method described earlier. Let  $(x_t, y_t)$  be the location of the tire center and  $r_t$  be the best estimate of its radius. Our goal is to find a point  $(x^*, y^*)$  such that  $x^* = x_t$  and  $y_t - y^*$  is the smallest distance from the tire center to the vehicle body directly above it. Then, the gap thickness  $\delta$  is given by  $\delta = y_t - (y^* + r_t)$  (recall that  $y$  increases from top to bottom).

An identical set of operations would be performed on the front tire and the gap value recorded as  $\delta_{front}$ . Then, the decision about rear-heaviness is made if  $\delta_{front} - \delta_{rear} > \tau$  where  $\tau$  is an experimentally determined constant. Note that, the difference is not absolute. This is because, under normal conditions, some front engine cars may have lower front than the rear. Thus, the method may detect a false positive if absolute difference is taken.

### 3.6. Velocity And Trajectory Analysis

Video-based traffic monitoring systems have demonstrated an ability to estimate the speed of vehicles in the camera’s field of view. Coifman [3] describes a comprehensive traffic analysis system which estimates the average vehicle speeds. Mehrubeoglu [17] and Gramatikopoulos [8] also describe detailed methods for obtaining vehicle speeds in traffic videos. Although some non-video methods (such as laser speed guns) are more accurate than video-based methods, they are usually deployed for observing the speeds over a short portion of the road and require human supervision.

Our system analyzes a vehicle’s approach by observing the variations in the vehicle’s velocity. This requires tracking the vehicle using an object tracking method [18]. We use a particle filtering approach for vehicle tracking with color and edge orientation histograms as features [19, 20]. The tracker generates the vehicle’s temporal position in image co-ordinates. By mapping these to ground co-ordinates, the desired representation of the trajectory is obtained. We calibrate the camera measurements by driving a vehicle at a constant known speed. The image and ground positions are used to generate a 2D look-up table (LUT). This LUT is used to obtain the ground position for unknown trajectories.

Let  $d(n)$  represent the trajectory (in ground co-ordinates) of an oncoming vehicle, where  $n \in I$  and  $I = 1, 2, \dots, N$  is the index set of frames in which the vehicle is visible. We use the convention that  $d = 0$  when the vehicle has just entered the camera’s field of view and  $d = d_{max}$  when it has just exited. We approximate the instantaneous velocity as  $v(m) = d(m) - d(m-1)$  in “distance per frame” (which can be scaled to a convenient unit like feet per second by a constant multiplication) where  $m = 2, 3, \dots, N$ . Optionally, we can use a low-pass filter using the speeds from the adjacent frames. This operation would introduce a delay of one frame duration but makes the analysis robust to camera shake and/or tracker errors.

The lifetime of observation, given by the index set  $I$  is divided into two parts –  $R$  and  $S$ :

$$R = \{1, 2, \dots, N_R\} = \{i \in I : d(i) \leq d_{max}/2\}, \quad (8)$$

and  $S = I \setminus R$ . In other words, we treat the observations differently based on whether the vehicle has traversed half the field of view (and is close to the check point) or not. The set  $R$  serves as a training phase to determine the average speed of the vehicle when far from the check point while the set  $S$  represents the actual evaluation

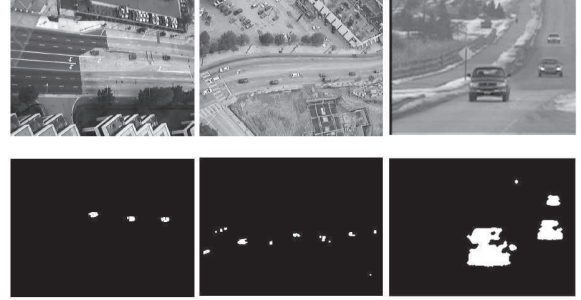


Fig. 5. Examples of vehicle detection.

phase in which the digression from the normal is studied. While the selection of  $d_{max}/2$  as the switch-over point is empirical, the idea of using a part of the trajectory for learning is useful because it allows the system to determine the “normal” response of vehicles on an individual basis. It does not assume that the vehicles are driven near the speed limits and reduces the dependence on the camera calibration step.

The output is a vector of flags  $\Lambda = \{\lambda(n) : n = 1, 2, \dots, N\}$  where,

$$\lambda(n) = \begin{cases} 0, & \text{if the trajectory is normal,} \\ 1, & \text{if the trajectory is anomalous.} \end{cases} \quad (9)$$

By the assumption of normal approach during the training phase,  $\lambda(n) = 0 \forall n \in R$ . We also define the average speed,

$$\hat{v} = \frac{1}{N_R} \sum_{i=2}^{N_R} v(i). \quad (10)$$

Finally, for all  $i \in S$ ,

$$\lambda(i) = \begin{cases} 0, & \text{if } |v(i) - \hat{v}| < \epsilon \text{ and } v(i) > 0, \\ 1, & \text{otherwise.} \end{cases} \quad (11)$$

where  $\epsilon$  is the maximum allowable digression. This analysis can be done for an oncoming vehicle and the occurrence of a significant number of positive anomaly flags should alert the check point personnel.

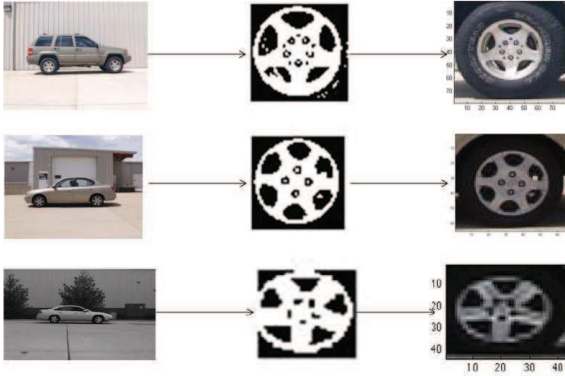
## 4. RESULTS

We tested the methods described above on videos of traffic recorded at different locations and with different weather conditions. Some methods were also tested using publicly available datasets [21]. Figure 5 shows examples of vehicle detection by background subtraction. All the foreground blobs correspond to vehicles but their sizes vary because of different camera distances and perspectives.

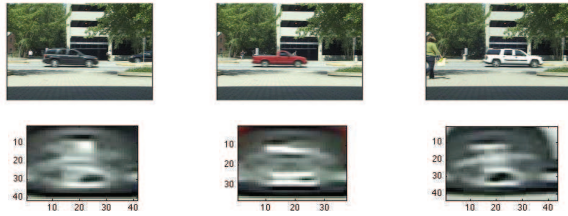
The methods for tire extraction were tested on both still images and video of vehicles. Figure 6 shows the extracted tires on a set of vehicle side view images. The estimated sizes were within  $\pm 1$  pixel tolerance.

The results of testing on videos are provided in Figure 7. Some vehicles in the test video were traveling as fast as 45 miles per hour. The size estimates were still within one pixel tolerance.

Similarly, the methods for body type identification were tested on both still images and videos. The results are presented in Figure 8 and 9, respectively. The vertical axis in the graphs represents the



**Fig. 6.** Examples of vehicle tires extracted from still side-view images. The middle column shows intermediate results in the form of wheel positions.



**Fig. 7.** Examples of vehicle tires extracted from video. In each case, the rear tire was segmented. Column 3 shows a failure case where the position of the tire center is incorrect.

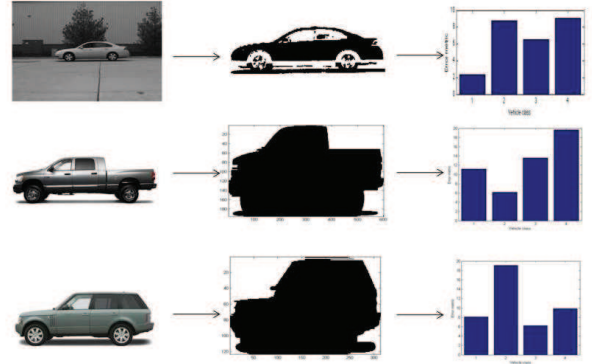
error metric (SAD) as defined in Equation 3. Thus the class with the smallest error is chosen.

Many experiments on vehicle make and model recognition use “clean” images of vehicles. Such images have high resolution and use vehicles which are stationary or near-stationary (for example, near the entrance of a parking lot or at a toll booth). Instead, we use video frames taken in outdoor conditions and on vehicles moving as fast as 45 miles per hour. In fact, the traffic video used in this experiment was recorded at the same time and location as the video used in the previous two experiments. Sample frames and the corresponding output are shown in Figure 10. The vertical axis represents the correlation coefficient as described in Equation 6. Thus the make with highest correlation coefficient is selected.

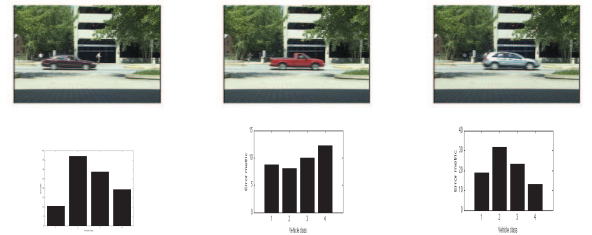
In order to study the static impact of loading on the vehicle, we imaged three vehicles with and without a payload of 500 pounds of sand bags in the trunk. The vehicles are shown in Figure 11 along with the difference image. The difference image is obtained by thresholding the absolute difference between the images under loaded and unloaded conditions. The white patches above the tires represent the lowering caused by the payload.

As discussed earlier, we expected to see larger gaps above the front tire (compared to the rear) after loading, and equal or smaller gaps without loading. Table 1 shows these measurements for the test vehicles. It can be seen that our hypothesis holds true in all three cases. Thus, comparing the gaps can serve as a useful yardstick to detect overloading.

Finally, we present the results of trajectory analysis using a ve-



**Fig. 8.** Examples of vehicle type identification from still side-view images. The middle column shows the silhouettes obtained from vehicle detection. The graphs represent the SAD values ( $y$ -axis) for each vehicle type ( $x$ -axis).



**Fig. 9.** Examples of vehicle type identification from video. The graphs represent the SAD values for each vehicle type.

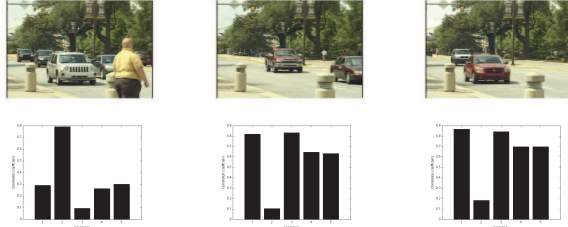
hicle driven to simulate different anomalous maneuvers. Figure 12 shows the trajectories for the different cases. The blue circles represent positions of the vehicle judged “normal” by the system whereas the red crosses indicates positions which correspond to anomalies. It can be seen that our proposed method of analysis and visualization can successfully detect digressions from the normal approach and allow quick characterization by a human operator.

## 5. CONCLUSIONS AND FUTURE WORK

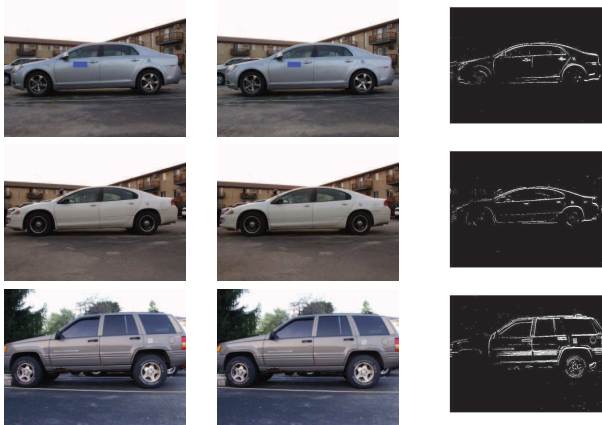
In this paper, we have described an automated video surveillance system which can screen vehicular traffic from a standoff range. We describe an orthogonal two-camera deployment scheme which obtains videos corresponding to the front and side views of a vehicle. We also described low-complexity methods which can be used to extract desired information from the camera data. Using these methods, we successfully determined the body type, make, and tire sizes

**Table 1.** The gap above tires measured in pixels for the test vehicles.

Vehicle	Front	Rear	Front	Rear
Vehicle 1	56	57	55	49
Vehicle 2	53	53	54	48
Vehicle 3	70	72	71	67



**Fig. 10.** Examples of vehicle make recognition using edge orientation histograms and automatic region selection. The bottom row represents the correlation coefficients ( $y$ -axis) corresponding to each car make ( $x$ -axis).

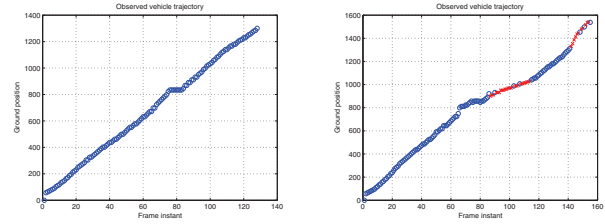


**Fig. 11.** Illustration of the visual impact of loading on test vehicles. The difference image (column 3) captures the displacement caused by loading. Columns 1 and 2 show the unloaded and loaded vehicles.

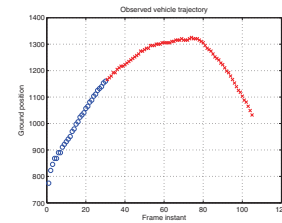
of vehicles in real traffic videos. We also used the velocity analysis method to detect anomalies in the approach of a vehicle. Our work can be extended by adding analysis techniques for extracting more information from videos and testing the robustness of the methods to changes in operating conditions (weather, traffic density, and lighting).

## 6. REFERENCES

- [1] R. T. Collins *et al.*, “A system for visual surveillance and monitoring,” Final Report CMU-RI-TR-00-12, Carnegie Mellon University, 2000.
- [2] J. Aguilera *et al.*, “Visual surveillance for airport monitoring applications,” *Proceedings of the Computer Vision Winter Workshop*, Telc, Czech Republic, February 2006.
- [3] B. Coifman, D. Beymer, P. McLauchlan, and J. Malik, “A real-time computer vision system for vehicle tracking and traffic surveillance,” *Transportation Research: Part C*, vol. 6, no. 4, pp. 271–288, 1998.
- [4] C. Stauffer and W. E. L. Grimson, “Learning patterns of activity using real-time tracking,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747–757, August 2000.
- [5] N. Ohta, “A statistical approach to background suppression for surveillance systems,” *Proceedings of the International Con-*



(a) Test vehicle approaches at uni- (b) The vehicle slows down, and form 30 miles per hour. then accelerates sharply.



(c) The vehicle makes a U-turn, and drives away.

**Fig. 12.** Trajectories of the test vehicle represented using the decision flags. Abnormal movements are indicated with red crosses.

*ference on Computer Vision*, Vancouver, Canada, July 2001, pp. 481–486.

- [6] A. Mittal and N. Paragios, “Motion-based background subtraction using adaptive kernel density estimation,” *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Washington, DC, June–July 2004, pp. 302–309.
- [7] K. K. Ng and E. J. Delp, “Object tracking initialization using automatic moving object detection,” *Proceedings of the IS&T/SPIE Electronic Imaging: Visual Information Processing and Communication*, vol. 7543, San Jose, California, January 2010.
- [8] L. Grammatikopoulos, G. Karras, and E. Petsa, “Automatic estimation of vehicle speeds from uncalibrated video sequences,” *Proceedings of the International Symposium of Modern Technologies, Education and Professional Practice in Geodesy and Related Fields*, Sofia, Greece, November 2005, pp. 332–338.
- [9] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. New York: Cambridge University Press, 2003.
- [10] S. Y. Cheung, S. Coleri, B. Dundar, S. Ganesh, C. Tan, and P. Varaiya, “Traffic measurement and vehicle classification with single magnetic sensor,” *Transportation Research Record*, vol. 1917, no. 19, pp. 173–181, 2005.
- [11] X. Song and R. Nevatia, “A model-based vehicle segmentation method for tracking,” *Proceedings of the International Conference on Computer Vision*, vol. 2, Beijing, China, October 2005, pp. 1124–1131.
- [12] A. H. S. Lai and N. H. C. Yung, “Vehicle type identification through automated virtual loop assignment and block-based direction-biased motion estimation,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 1, no. 2, pp. 86–97, June 2000.

- [13] V. S. Petrovic and T. F. Cootes, "Vehicle type recognition with match refinement," *Proceedings of the International Conference on Pattern Recognition*, vol. 3, Cambridge, UK, August 2004, pp. 95–98.
- [14] L. Dlagnekov, "Video-based car surveillance: License plate, make and model recognition," Masters Thesis, University of California, San Diego, 2005.
- [15] I. Zafar, E. A. Erdisinghe, S. Acar, and H. E. Bez, "Two dimensional statistical linear discriminant analysis for real-time robust vehicle type recognition," *Proceedings of the IS&T/SPIE Electronic Imaging: Real Time Image Processing*, vol. 6496, San Jose, California, January-February 2007.
- [16] I. Sobel and G. Feldman, "A  $3 \times 3$  isotropic gradient operator for image processing," Unpublished, presented at Stanford Artificial Project, 1968.
- [17] M. Mehrubeoglu and L. McLauchlan, "Determination of vehicle speed in traffic video," *Proceedings of the IS&T/SPIE Electronic Imaging: Real Time Image and Video Processing*, vol. 7244, San Jose, California, January 2009.
- [18] A. Yilmaz, O. Javed, and M. Shah, "Object tracking: A survey," *ACM Computing Surveys*, vol. 38, no. 4, December 2006.
- [19] M. S. Arulampalam, S. Maskell, and N. Gordon, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174–188, February 2002.
- [20] K. K. Ng and E. J. Delp, "New models for real-time tracking using particle filtering," *Proceedings of SPIEIS&T Electronic Imaging: Visual Communications and Image Processing*, vol. 7257, San Jose, California, January 2009.
- [21] The Next Generation Simulation (NGSIM) Community, "Data sets," Web archive: <http://ngsim-community.org/>, Accessed: May 2010.