Automatic Traffic Surveillance System for Vehicle Tracking and Classification

Jun-Wei Hsieh, *Member, IEEE*, Shih-Hao Yu, Yung-Sheng Chen, *Member, IEEE*, and Wen-Fong Hu

***Abstract*—This paper presents an automatic traffic surveillance system to estimate important traffic parameters from video sequences using only one camera. Different from traditional methods that can classify vehicles to only cars and noncars, the proposed method has a good ability to categorize vehicles into more specific classes by introducing a new "linearity" feature in vehicle representation. In addition, the proposed system can well tackle the problem of vehicle occlusions caused by shadows, which often lead to the failure of further vehicle counting and classification. This problem is solved by a novel line-based shadow algorithm that uses a set of lines to eliminate all unwanted shadows. The used lines are devised from the information of lane-dividing lines. Therefore, an automatic scheme to detect lane-dividing lines is also proposed. The found lane-dividing lines can also provide important information for feature normalization, which can make the vehicle size more invariant, and thus much enhance the accuracy of vehicle classification. Once all features are extracted, an optimal classifier is then designed to robustly categorize vehicles into different classes. When recognizing a vehicle, the designed classifier can collect different evidences from its trajectories and the database to make an optimal decision for vehicle classification. Since more evidences are used, more robustness of classification can be achieved. Experimental results show that the proposed method is more robust, accurate, and powerful than other traditional methods, which utilize only the vehicle size and a single frame for vehicle classification.**

***Index Terms*—Linearity feature, occlusions, shadow elimination, traffic surveillance, vehicle classification.**

I. Introduction

AN INTELLIGENT transportation system (ITS) is the application that incorporates electronic, computer, and communication technologies into vehicles and roadways for monitoring traffic conditions, reducing congestion, enhancing mobility, and so on. To achieve these goals, in past decades, there have been many approaches proposed for tackling related problems in ITS [1]–[19]. Among them, the vision-based approach has the advantages of easy maintenance and high flexibility in traffic monitoring and, thus, becomes one of the most popular techniques used in ITS for traffic controls. In the past, there have been many researchers devoting themselves to investigating different solutions [1]–[19] for extracting important real-time traffic information via image processing. For example, Beymer *et al.* [13] proposed a vehicle-tracking algorithm to estimate traffic parameters using corner features. In addition, Liao *et al.* [25] used entropy as an underlying measurement to calculate traffic flows and vehicle speeds. However, these approaches cannot further classify vehicles to more detailed types. Therefore, Baker *et al.* [4] proposed a three-dimensional (3-D) model-matching scheme to classify vehicles into various types like wagons, sedan, hatchback, etc. In addition, Lipton *et al.* [9] used maximum-likelihood estimation criteria with shape features to classify different targets into vehicles and humans. Furthermore, Gupte *et al.* [6] proposed a region-based approach to track and classify vehicles based on the establishment of correspondences between regions and vehicles.

For most traffic surveillance systems, three major stages are used to estimate desired traffic parameters, i.e., vehicle detection, tracking, and classification. For vehicle detection, most methods [6], [13], [25] assume that the camera is static and then desired vehicles can be detected by image differencing. Then, different tracking schemes like the Kalman filter [26] are designed to track each vehicle. After that, several vehicle features like shape, length, width, texture, etc., are extracted for vehicle classification. However, several environmental variations will heavily affect the accuracy and robustness of vehicle analysis. For example, vehicle occlusion will result in the failure of vehicle detection and further degrade the accuracy of vehicle classification and counting. The major factor to cause vehicle occlusions is shadows. To avoid this problem, most systems assume the analyzed sequence includes no shadows. Another environmental variation is perspective effects, which will let vehicle geometry features such as length, width, and height be not constant. Therefore, before recognition, a process of camera calibration for normalizing features should be applied in advance. For example, in [6] and [13], a manual method of camera calibration is presented to identify lane locations and corresponding lane widths. Then, different vehicles are tracked and recognized by a set of normalized features. Another condition will also increase the difficulty of vehicle classification. Usually, the camera is placed far from vehicles, and thus all types of vehicles will have similar features. Therefore, although different approaches [6], [9] claimed that they can classify different vehicles, only two categories are classified, i.e., cars and noncars. For classifying vehicles into more types, many advantages can be benefited from the use of 3-D vehicle features and models [4], [5], [20]. However, the inherent correspondence problem makes them unfeasible for real-time applications. On the other hand, when a vehicle is observed along a lane, it will have different appearances at different time

Manuscript received September 3, 2003; revised June 6, 2005. This work was supported in part by the National Science Council of Taiwan, Taiwan, R.O.C., under Grants NSC 92-2213-E-155-052 and 93-2213-E-155-026, and the Ministry of Economic Affairs under Contract 94-EC-17-A-02-S1-032. The Associate Editor for this paper was A. Eskandarian.

The authors are with the Department of Electrical Engineering, Yuan Ze University, Chung-Li 320, Taiwan, R.O.C. (e-mail: shieh@saturn.yzu.edu.tw; eeyschen@saturn.yzu.edu.tw; hubert@cyu.edu.tw).

Digital Object Identifier 10.1109/TITS.2006.874722

frames. With a tracking technique, all the appearances can be integrated together so that a trajectory is constructed. Since many cues can be found along this trajectory, a better judgment can be then made on recognizing this vehicle. However, in ITS, most approaches [1]–[16] used only one appearance to recognize vehicles. If the appearance is polluted by noise, recognition errors will be produced. In addition to the vision-based approach, there are many systems [2], [3], using other sensors like sonar or infrared to detect and track vehicles. It has been proved in [28] that the use of multiple sensors and multiple classifiers can provide better performances in object detection and classification. Therefore, we can build a surveillance system to monitor traffics using multiple sensors and classifiers. However, this scheme requires a complicated fusion technique for integrating information from multiple sensors. In addition, more sensors mean more costs required for building this system. Therefore, this paper focuses on developing a vision-based system to analyze vehicles using only one general camera.

In this paper, we propose a novel vehicle surveillance system to detect, track, and recognize vehicles from different video sequences. In this system, only one camera, still mounted on a pole and looking down on the observed highway, is used to monitor various vehicles. At the beginning, different vehicles are extracted from video sequences using the technique of image subtraction. Since this technique is sensitive to lighting changes, an adaptive background updating method is used for modeling the background. Then, a Kalman filter [26] is designed to track each detected vehicle and thus obtain its trajectory. After that, several vehicle features like vehicle size and linearity are extracted for vehicle classification. The proposed method makes several contributions on the improvement of an ITS including the accuracies and robustness of vehicle detection and recognition, respectively. First of all, this paper presents an automatic scheme to detect all possible lane-dividing lines by analyzing different vehicle trajectories. Thus, without involving any manual calibration process, different lane widths can be estimated. Second, the paper proposes a line-based shadow-elimination method to eliminate various unwanted shadows from video sequences. Since shadows have been removed, occluded vehicles caused by shadows can be well separated. Thus, further vehicle analysis like feature extraction and vehicle counting can be performed more accurately and robustly. Third, since different lane widths have been found, an automatic normalization scheme can be proposed for feature normalization. Fourth, we define a new feature called "vehicle linearity" to classify vehicles to more types. The new feature is very useful for discriminating "van truck" from "truck" even without using any 3-D information. Fifth, after feature extraction and normalization, an optimal classifier is then designed for more accurate vehicle classification. In practice, the vehicle features are easily affected by lighting changes or noise, which cannot be avoided even though various preprocessing techniques are used in advance. This paper presents two ideas for tackling this problem as follows. The first one is to introduce more training samples into each category of the database and the second one is to make a decision from more supports. Since a vehicle has many appearances when it moves along a road, they can be integrated together and provide more supports or evidences for the classifier to make a better decision. Based on these two ideas and the spirit of maximum-likelihood estimation, an optimization classifier can be then designed for more accurate vehicle categorization. Experimental results show that the proposed method offers great improvements in terms of accuracy, robustness, and stability in traffic surveillance.

The rest of this paper is organized as follows. In the next section, the procedures of the whole proposed system are described. Then, methods of image differencing and lane-dividing-line detection are discussed in Section III. Section IV describes details of our line-based shadow-elimination method. Then, details of feature extraction and vehicle classification are described in Section V. Section VI reports experimental results. Finally, a conclusion will be presented in Section VII.

II. Overview of the Proposed System

In this paper, we propose a novel traffic surveillance system for detecting, tracking, and recognizing vehicles from different video sequences. Fig. 1 shows the flowchart of this proposed system. This system includes an initialization stage to obtain the information of lane-dividing lines and lane width. Fig. 1(a) shows details of this initialization stage. In this stage, different vehicles are first extracted through image differencing, and then a vehicle histogram is obtained by accumulating the number of vehicles passing a position. Thus, the information of lane-dividing lines can be found by analyzing this histogram and will be later used for shadow removing and feature normalization. Fig. 1(b) shows details of the whole system to obtain all desired traffic parameters step by step. The initial inputs of the proposed system include lane width, lane-dividing lines, and the analyzed video sequence. Then, according to the initial parameters, each moving vehicle can be extracted and analyzed for traffic-parameter estimation. Unfortunately, due to shadows, several vehicles will be occluded together and cannot be well separated. Therefore, before tracking and recognizing, each extracted vehicle will be fed into a shadow-elimination process for reducing the occlusion effect to a minimum. However, we should emphasize that an occlusion may be caused by shadows or perspective effects. Since general occlusion is still an open problem in computer vision, our proposed elimination method can only deal with the case of occlusions caused by shadows. Then, a tracking technique is applied for obtaining the trajectory of each passing vehicle. Furthermore, through line fitting and a connected component analysis [30], two useful features, i.e., linearity and size, are extracted from each vehicle. According to the chosen features and trajectories, an optimal classifier can be designed to categorize vehicles into more detailed classes. Once each vehicle is recognized, different traffic parameters can be estimated. For example, given a road and its length, we can easily estimate its volume, density, mean speed, and lane flow. In addition, since vehicles can be well classified to different categories, the flow segregated by vehicle type can also be well estimated. In what follows, Section III will first describe details of our method to find lane-dividing lines. Other procedures like shadow elimination, tracking, feature extraction, and vehicle classification will be described in Sections IV–VI, respectively.

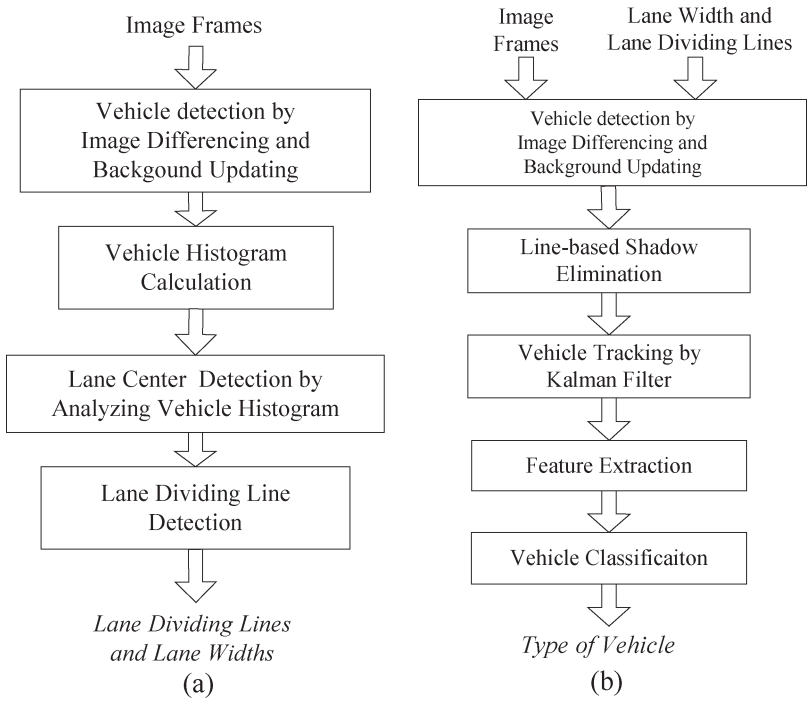

Fig. 1. Flowchart of the proposed surveillance system. (a) Initialization stage to detect dividing lane lines and lane widths. (b) Procedures of the proposed traffic surveillance system.

III. Detection of Lane-Dividing Lines

As described before, for classification, each vehicle should be first detected and tracked from video frames. For simplifying the problems of vehicle segmentation, this paper assumes all the analyzed frames are captured by a still camera. When the camera is static, different moving objects can be detected through background subtraction. Assume that I_k and B_k are intensities of the kth frame and background, respectively. The difference image $D_k(x, y)$ used to detect moving objects can be defined as follows:

$$D_k(x,y) = \begin{cases} 0, & \text{if } |I_k(x,y) - B_k(x,y)| \le T_{\mathrm{d}} \\ 1, & \text{otherwise} \end{cases} \quad (1)$$

where T_{d} is a predefined threshold and chosen as the average of the difference image $D_k(x, y)$. After subtraction, a series of simple morphological operations is applied for noise removing. Fig. 2 shows the result of vehicle detection after several morphological operations. Fig. 2(c) is the result of image subtraction based on Fig. 2(a) and (b). Fig. 2(d) is the result of noise removing using some morphological operations. Then, based on (1), each vehicle can be detected from video sequences and used for further analyzing and recognizing. However, if shadows exist, each vehicle will be occluded by another vehicle. In addition, its geometrical features will also change when it moves along a road. For recognizing each vehicle more accurately, this paper takes advantages of lane-dividing lines to tackle the two problems, i.e., shadow elimination and feature normalization. Therefore, in what follows, we will present an automatic method to detect lane-dividing lines from a vehicle histogram.

In general, if a vehicle moves regularly, its center will be very close to one of the lane centers. When more vehicles are collected, their trajectories will be gradually close to the central lines of lanes. Therefore, we can construct a vehicle histogram for estimating all desired lane centers. The vehicle histogram is obtained by recording the frequencies of vehicles moving at different positions across different frames. Fig. 3 shows the accumulation result of vehicles moving at different positions when two thousand training frames were used. When accumulating, if two vehicles are occluded together, they still are considered as one for this accumulation. After accumulation, we can obtain the desired vehicle histogram $H_{\mathrm{vehicle}}(x, y)$, where the entry (x, y) records the number of vehicles whose centers are at (x, y).

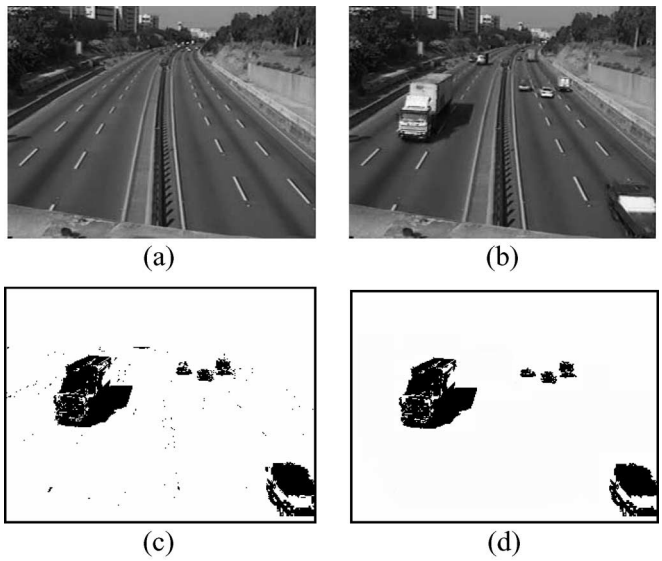

Fig. 2. Results of vehicle extraction. (a) Background. (b) Frame of vehicles. (c) Differencing and binarization between (a) and (b). (d) Result after morphological operations.

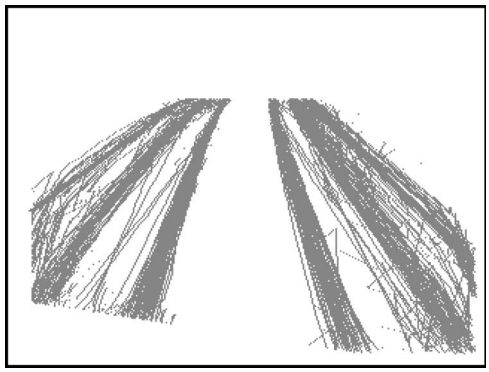

Fig. 3. Histogram of different vehicles moving at different lanes.

Assume there are N_L lanes in the analyzed highway. Clearly, each row of H_{vehicle} should have N_L peaks corresponding to different lane centers. Thus, different lane centers can be found from the peaks of H_{vehicle}. Assume v_k is a vehicle in the training set with the center (x_{v_k}, y_{v_k}). In addition, H_{vehicle} has the dimension $N_{\text{col}} \times N_{\text{row}}$. Then, the algorithm to detect lane-dividing lines can be illustrated in detail as follows.

Lane-Dividing-Line Detection Algorithm

Input: all training vehicles v_k.

Step 1) Initialize all entries of H_{vehicle} to be zero.

Step 2) For all vehicles v_k, calculate $H_{\text{vehicle}}(x_{v_k}, y_{v_k}) += 1$.

Step 3) Smooth the histogram H_{vehicle} using the equation

$$\overline{H}_{\text{vehicle}}(i,j) = 1/5 \sum_{k=-2}^{2} H_{\text{vehicle}}(i+k,j) \quad \text{for all } i \text{ and } j.$$

Step 4) Get the average value of $\overline{H}_{\text{vehicle}}$ at the jth row, i.e.,

$$T_H^j = \frac{1}{N_{\text{col}}} \sum_i \overline{H}_{\text{vehicle}}(i,j).$$

Step 5) For each pixel (i,j) along the jth row, if $\overline{H}_{\text{vehicle}}(i,j)$ is a local maximum and larger than T_H^j, set $H_{\text{vehicle}}(i,j)$ to 1; otherwise, set $H_{\text{vehicle}}(i,j)$ to zero.

Step 6) Apply a connected component analysis [30] to H_{vehicle} for finding all its isolated segments. Eliminate each segment if its length is short, i.e., less than 0.5 T_L, where T_L is the average length of all segments.

Step 7) Merge any two adjacent segments if they are very close to each other. Set all the remained segments as the lane centers.

Step 8) Let $C_{L_k}^j$ be the center of the kth lane at the jth row. Then, each lane-dividing line (except the most left and right ones) can be determined as follows:

$$\text{DL}_k^j = \frac{1}{2}\left(C_{L_{k-1}}^j + C_{L_k}^j\right)$$

where DL_k^j is the point at the jth row of the kth dividing line. In addition, the lane width $w_{L_k}^j$ of the kth lane at the jth row is obtained as

$$w_{L_k}^j = \left|C_{L_k}^j - C_{L_{k-1}}^j\right|.$$

(a) (b)

Fig. 4. Result of detecting lane centers and lane-dividing lines. (a) Result of detecting lane centers from a vehicle histogram. (b) Result of detecting lane-dividing lines from (a).

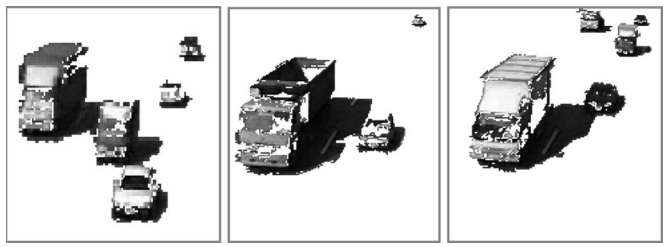

Fig. 5. Different cases of occlusions caused by shadows.

Step 9) For the most left and right dividing lines, i.e., DL_0 and DL_{N_L}, their positions at the jth row can be extended from DL_1^j and $\text{DL}_{N_L-1}^j$, respectively, as follows:

$$\left(x_{\text{DL}_0^j}, j\right) = \left(x_{\text{DL}_1^j} - w_{L_0}^j, j\right) \text{ and}$$
$$\left(x_{\text{DL}_{N_L}^j}, j\right) = \left(x_{\text{DL}_{N_L-1}^j} + w_{L_0}^j, j\right).$$

In Fig. 4, according to the trajectories of moving vehicles, six major lane central lines were obtained and shown in Fig. 4(a). Then, by tracing the middle points of any two adjacent central lines, four lane-dividing lines were first obtained. The other four lane-dividing lines at the boundaries can be obtained using an extending technique.

IV. Shadow Detection and Occlusion Elimination

As described before, due to perspective effects or shadows, different vehicles may be occluded together. Fig. 5 shows different cases of occlusions when shadows exist. In this section, a novel shadow-elimination method will be proposed for reducing the effects of shadow occlusions to a minimum. For the occlusions caused by perspective effects, since it is still an open problem in computer vision [30], this paper will not pay any attention to solving it. In the past, there have been many shadow-elimination methods [21]–[24] proposed in the literatures. However, these approaches model shadows only based on color features and some ones require more than one camera for shadow elimination. Then, these approaches impose photometric constraints locally on individual points and then identify shadow pixels based on local *a priori* thresholding. However, color features cannot provide enough information to discriminate black objects from shadows. Since only color features are used, it is not surprising that only very limited results

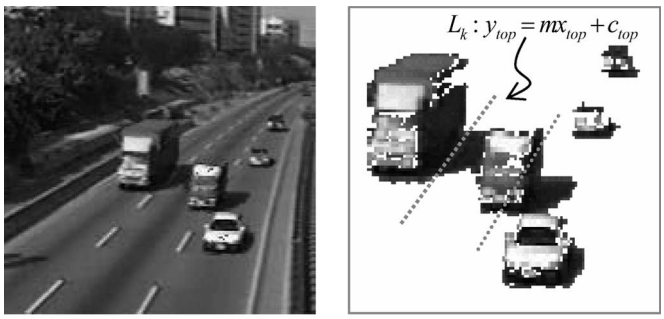

Fig. 6. Three vehicles are occluded together due to shadows. (a) Original image. (b) Result after image differencing. Two lane-dividing lines pass through this occluded region.

were achieved by the above approaches. Therefore, different from traditional methods [21]–[24], this paper considers that shadow geometry is also an important cue for shadow modeling and elimination. Then, by taking advantage of the found lane-dividing lines, this paper will propose a simple but effective line-based method for shadow elimination.

Actually, the proposed method uses two kinds of lines to eliminate unwanted shadows. The two lines are, respectively, parallel and vertical to the lane-dividing lines. The different type of lines uses different scanning methods to eliminate unwanted shadows. Fig. 6 is an example to illustrate details of our shadow-elimination method. If only color feature is considered, the black vehicle regions will also be classified as shadows. However, there are two lane-dividing lines passing this occluded region, as shown in Fig. 6(b). Then, we can use the lines, parallel to the dividing lines, to horizontally eliminate different unwanted shadows from left to right. Similarly, we can benefit from the lines vertical to the dividing lines for another scanning method to eliminate shadows. In what follows, we discuss the scheme to eliminate shadows using the first type of lines.

In Fig. 6(b), assume that the straight line L_k passes the occluded region $R_{\rm O}$ and is determined by the kth dividing line DL_k. Let U_k be the set of points that appear both in DL_k and $R_{\rm O}$. Then, L_k can be approximated by U_k using a line-fitting method. Assume U_k has N data points (x_i, y_i) and will be fitted to a straight-line model:

$$y = m_k x + b_k. \tag{2}$$

Then, the parameters m_k and b_k can be obtained by minimizing the error function [29]:

$$E(b_k, m_k) = \sum_{i=1}^{N} (y_i - b_k - m_k x_i)^2. \tag{3}$$

By setting the first derivatives of $E(b_k, m_k)$ with respect to m and b to zero, respectively, m and b can be obtained by

$$m_k = \frac{1}{K}\left(N\overline{X}\,\overline{Y} - \sum_{i=1}^{N} x_i y_i\right) \text{ and}$$
$$b_k = \frac{1}{K}\left(\overline{X}\sum_{i=1}^{N} x_i y_i - \overline{Y}\sum_{i=1}^{N} x_i^2\right) \tag{4}$$

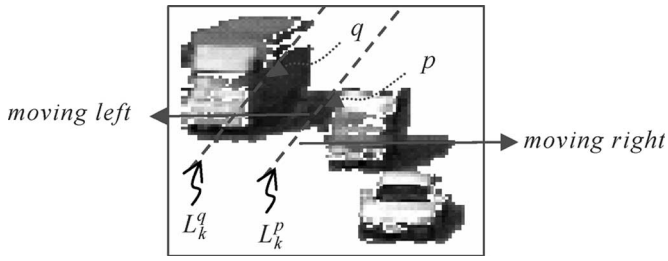

Fig. 7. Different pixels determine different lines with the same slope of one dividing line. Different shadows can be well discriminated from these lines.

where $\overline{X} = (1/N)\sum_{i=1}^{N} x_i$, $\overline{Y} = (1/N)\sum_{i=1}^{N} y_i$, and $K = N\overline{X}\,\overline{X} - \sum_{i=1}^{N} x_i^2$. Based on m_k, a set of lines parallel to L_k can be generated for shadow elimination by changing the value of b. In (2), given any point p in $R_{\rm O}$, a new value of b can be generated. Assume $b_k^{\min}$ and $b_k^{\max}$ are, respectively, the minimum and maximum values of b for all points in $R_{\rm Occlusion}$, i.e.,

$$b_k^{\min} = \min_{p \in R_{\rm O}} y_p - m_k x_p \quad \text{and} \quad b_k^{\max} = \max_{p \in R_{\rm O}} y_p - m_k x_p. \tag{5}$$

Thus, given a new b in the range $[b_k^{\min}, b_k^{\max}]$, a new line L_k^b with the same m_k but different b can be generated. Let $S_{L_k^b}$ be the set of pixels appearing both in L_k^b and $R_{\rm O}$. For a pixel p in $S_{L_k^b}$, if only its intensity is used, the probability of p being a shadow pixel is measured by

$$P(\text{shadow}|p) = \exp\left(-\frac{(I(p) - m_{\rm shadow})^2}{\sigma_{\rm shadow}^2}\right)$$

where $m_{\rm shadow}$ and $\sigma_{\rm shadow}$ are the intensity mean and variance of shadows, respectively, and $I(p)$ is the intensity of p. $m_{\rm shadow}$ and $\sigma_{\rm shadow}$ can be experimentally obtained from thousands of shadow images through a training process. Then, if $P(\text{shadow}|p) > 0.8$, p is said to be a shadow pixel. However, this rule cannot discriminate a nonshadow pixel from shadows if it is black. The problem can be easily tackled if p is classified according to a set of evidences (or pixels). The evidence of p is collected along a line L_k^p, which is determined by p and the slope m_k of L_k. If all evidences in L_k^p show that p is a shadow, p will be a shadow. Otherwise, it will be a nonshadow one. In Fig. 7, given a pixel p, we can use p and the slope m_k of L_k to determine the scanning line L_k^p. Since all points in L_k^p are classified as shadows, p will be a shadow pixel. Conversely, for a nonshadow pixel q (although it is black), only parts of evidences in L_k^q (determined by q) show that q is a shadow. The rejections from other nonshadow pixels imply that q is a nonshadow pixel and should be classified as a vehicle pixel. Let $S_{L_k^p}$ be the set of pixels appearing both in L_k^p and $R_{\rm O}$. In our proposed method, p is classifies as a shadow if all pixels r in $S_{L_k^p}$ should satisfy that $P(\text{shadow}|r) > 0.8$. In addition, the line L_k^p is called a shadow line if all pixels in $S_{L_k^p}$ are shadows.

If p is given along the direction parallel to the x-axis, the line L_k^p will be horizontally moved. When moving, if L_k^p is a shadow line, all elements in $S_{L_k^p}$ will be eliminated. Thus, different shadow pixels can be detected and eliminated line

by line. In what follows, details of our horizontal line-based shadow-elimination method are described.

Horizontal Line-Based Shadow-Elimination Algorithm

Input: a vehicle V.

Step 1) Determine all the lane-dividing lines that pass V and collect them as the set DL.

Step 2) If DL is empty, V is not occluded by shadows and go to step 4) for stopping.

Step 3) For each dividing line DL_k in DL:

a) Determine the set U_k of points that appear both in DL_k and V.

b) Determine m_k and b_k of DL_k from U_k by (4).

c) Determine $b_k^{\min}$ and $b_k^{\max}$ by the following equations:

$$b_k^{\min} = \min_{p \in R_{\mathrm{O}}} y_p - m_k x_p \quad \text{and} \quad b_k^{\max} = \max_{p \in R_{\mathrm{O}}} y_p - m_k x_p.$$

d) Let $b = b_k$.
While $(b \geq b_k^{\min})$, do the following steps:
i) Determine the line L_k^b by $y = m_k x + b$.
ii) Determine the set $S_{L_k^b}$, i.e., the pixels appearing in both L_k^b and V.
iii) If all pixels r in $S_{L_k^b}$ satisfy that $P(\text{shadow}|r) > 0.8$, eliminate all pixels in $S_{L_k^b}$ from V and set $b = b - 1$. Otherwise, set b to $b_k^{\min} - 1$ for stopping.

e) Let $b = b_k + 1$.
While $(b \leq b_k^{\max})$, do the following steps:
i) Determine the line L_k^b by $y = m_k x + b$.
ii) Determine the set $S_{L_k^b}$, i.e., the pixels appearing both in L_k^b and V.
iii) If all pixels r in $S_{L_k^b}$ satisfy that $P(\text{shadow}|r) > 0.8$, eliminate all pixels in $S_{L_k^b}$ from V and set $b = b + 1$. Otherwise, set b to $b_k^{\max} + 1$ for stopping.

Step 4) Stop and quit.

Fig. 8 shows the results obtained by the proposed horizontal shadow-elimination method. Fig. 8(b) and (d) are the elimination results of Fig. 8(a) and (c), respectively. Since only one scanning direction is used, it is not surprising that the car A still contains some unwanted shadows [see Fig. 8(b) and (d)]. These problems can be easily solved if a vertical scanning method is adopted. Actually, the method uses a set of lines vertical to the lane-dividing lines to gradually eliminate shadows. If the camera is placed with a view parallel to the lanes, the set of used lines will be vertical to the y-axis. For computation efficiency, this paper chooses the lines vertical to the y-axis rather than the lane direction for shadow elimination.

When eliminating shadows, the scanning line is moved vertically. If it is a shadow line, all its elements will be eliminated. Thus, the shadows of the car A in Fig. 8(b) can be successfully eliminated as the result in Fig. 9(a). However, in Fig. 8(d), the remained shadow connects not only to the vehicle A but also to vehicle B. Thus, when scanning, each scanning line will include not only shadows but also other vehicle pixels. Therefore, none of the scanning lines is a shadow and no

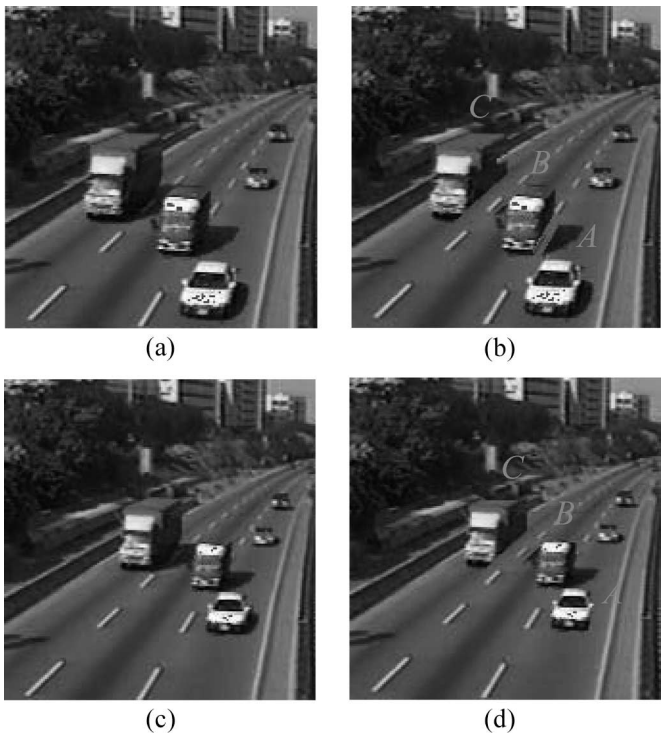

Fig. 8. Result of horizontal shadow elimination. (a) and (c) Original images. (b) Result obtained from (a). (d) Result obtained from (c).

shadow pixel in this occluded region can be eliminated. In practice, two cases will cause this double connection. The first one is due to the occlusion caused not only by shadows but also perspective effects. The second one is due to the vehicle occupying two lanes. For the first one, since the occlusion caused by perspective effects is still an open problem in computer vision [30], this paper will not tackle this kind of shadow problem. For the second one, it is seldom and irregular for a driver to drive a vehicle occupying two lanes. Since the second case seldom happens, it will not significantly affect the accuracy of further vehicle counting and recognition. However, this paper will tackle this problem by finding different separation lines through a y-projection technique. The technique projects all boundary points on the y-axis for obtaining a boundary histogram. Then, different separation lines can be found from the peaks of the histogram for vehicle separation. In Fig. 10, the separation line is the one that has a large peak in the boundary histogram. By setting all shadow elements in this line to the background, different parts (A and B) in the occluded region can be well separated.

In our vertical shadow-elimination algorithm, the y-projection technique is first applied to obtaining a boundary histogram from which different separation lines can be found. These lines will be further examined whether they can successfully separate occluded vehicles to different parts by setting their shadow elements to the background. For a separation line, if this examination is passed, it will be an important key to separate occluded vehicles to different parts. If unfortunately none of the separation lines is found, the occlusion will be caused by perspective effects and will not be

(a) (b)

Fig. 9. Elimination results of Fig. 8(b) and (d) using the proposed vertical shadow-elimination algorithm.

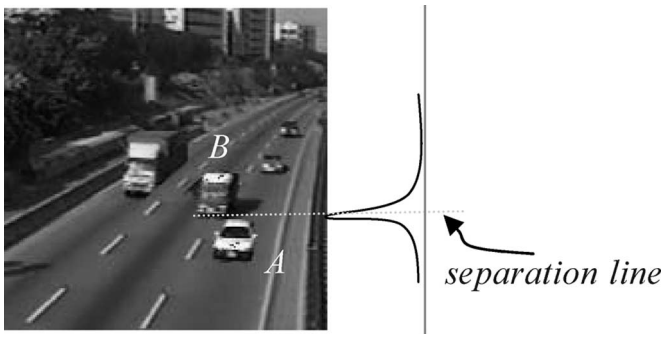

Fig. 10. Separation line will separate different vehicles into different parts.

tackled in this paper. In what follows, details of our proposed vertical shadow-elimination algorithm are described.

Vertical Line-Based Shadow-Elimination Algorithm

Input: a vehicle V with the set B_V of boundary points

Step 1) For all boundary points in V, obtain their y projection, i.e., $H(i) = \sum_{p(x,y)\in B_V \text{ and } y=i} 1$.

Step 2) Find all positions whose values in $H(i)$ are local maxima. For each local maximum, if its value is larger than a threshold $T_{H(i)}$, collect its position to the set Y_H, where $T_{H(i)}$ is the average value of $H(i)$.

Step 3) For each element y_k in Y_H, define a horizontal line L_{y_k} with the equation: $y = y_k$. Collect the line L_{y_k} to the set S_{sep} if L_{y_k} is a shadow line.

Step 4) Separate V into several parts $\{v_0, \ldots, v_{n_{\text{sep}}}\}$ using the lines in S_{sep} which has n_{sep} elements.

Step 5) For each part v_k in V, do:

a) Get the range of the y coordinates of pixels in v_k, i.e., $[y_{v_k}^{\min}, y_{v_k}^{\max}]$.

b) Generate each scanning line L_{y_i} with the equation $y = y_i$ for all integers $y_i \in [y_{v_k}^{\min}, y_{v_k}^{\max}]$.

c) Eliminate all pixels of L_{y_i} from v_k with a top-down scanning method for $y_i \in [y_{v_k}^{\min}, y_{v_k}^{\max}]$ if L_{y_i} is a shadow line. Otherwise, go to step d).

d) Eliminate all pixels of L_{y_i} from v_k with a bottom-up scanning method for each $y_i \in [y_{v_k}^{\min}, y_{v_k}^{\max}]$ if L_{y_i} is a shadow line. Otherwise, go to step e).

e) Apply the horizontal shadow-elimination algorithm to v_k.

Based on this algorithm, the remained shadows in Fig. 8(b) and (d) can be well eliminated as the results shown in Fig. 9(a) and (b), respectively.

V. Feature Extraction and Vehicle Classification

Once different vehicles have been extracted, we should classify them into different categories. In this paper, two features, including the size and the "linearity" of a vehicle, are used to classify vehicles into different categories, i.e., cars, minivans, trucks, and van trucks. Since a vehicle has different appearances when it moves, an optimal classifier can be then designed by integrating these appearances and features together. In what follows, details of feature extraction and the integration scheme of vehicle classification are discussed.

A. Feature Extraction

This paper uses the features "size" and "linearity" to classify vehicles. For the first feature, due to the perspective effect, the size of a vehicle is not constant and will gradually change when it moves. To keep its invariance, we can use the information of lane width to normalize it in advance. In the past, the lane width was usually manually obtained by the camera calibration process. However, we can take advantage of lane-dividing lines (found in Section III) to obtain all lane widths automatically. In addition to the size normalization, a new feature called "vehicle linearity" is also proposed for classifying vehicles into more detailed types. In what follows, we will first discuss the problem of size variation and then describe details of the "linearity" feature for vehicle classifications.

In Section III, a histogram-based approach has been proposed for finding a set of lane-dividing lines $\{\text{DL}_i\}_{i=0,1,\ldots,N_{\text{DL}}-1}$. This paper measures the width of the ith lane by calculating the distance between the ith and $(i+1)$th dividing lines. In other words, the width of the ith lane at point p is determined by

$$W_{\text{Lane}_i}(p) = \left|X_{\text{DL}_i}(y_p) - X_{\text{DL}_{i+1}}(y_p)\right|$$

where y_p is the y coordinate of p and $X_{\text{DL}_i}(j)$ is the x coordinate of the jth point on the ith lane-dividing line. Assume that a vehicle v with the size s_v moves along the ith lane. Its size will gradually change according to its different positions on this lane. If c_v is the central point of v, the size change of v will be proportioned to the square of $W_{\text{Lane}_i}(c_v)$. According to this fact, s_v can be then normalized as follows:

$$\overline{s}_v = s_v / W_{\text{Lane}_i}^2(c_v). \tag{6}$$

Then, the $\overline{s}_v$ forms a good feature for vehicle classification.

In addition to the vehicle size, the "linearity" feature is also very important for discriminating trucks from van trucks (or buses). In Fig. 11, the truck and bus have similar sizes and speeds but different up-slanted edges. If the "linearity" of an up-slanted edge can be defined and estimated, the work of classifying vehicles into different classes like buses or trucks will become easy. Given a vehicle H_i, its up-slanted edge U_{H_i} can be obtained by tracing all the boundary pixels of H_i. When

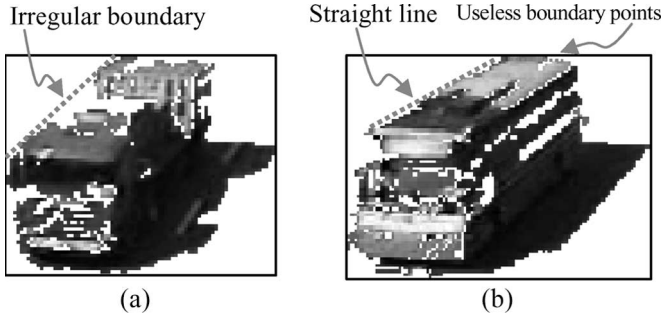

Fig. 11. Different up-slanted boundaries of a truck and a bus. (a) Irregular upper slanted boundary of a truck. (b) Straight upper slanted boundary of a bus.

Fig. 12. Results of up-slanted boundary detection. (a) Result of a truck. (b) Result of a bus.

scanning H_i, only the boundary pixels far from the vehicle bottom are recorded as elements in U_{H_i}. Fig. 11(b) shows the scanning result of a bus. However, the set U_{H_i} still will include many useless points for estimating the linearity feature of H_i. Assume that B_{H_i} is the minimum bounding box of H_i. Only the pixels far from B_{H_i} are useful for estimating the linearity feature. Let $d_{B_H}(p)$ be the minimum distance of a point p to all boundaries of B_H and $d_{B_{H_i}}^{\max} = \max_{q \in U_{H_i}} \{d_{B_{H_i}}(q)\}$. Then, a new set $\overline{U}_{H_i}$ of up-slanted boundary pixels can be obtained as follows:

$$\overline{U}_{H_i} = \{p | p \in U_{H_i}, d_{B_H}(p) > 0.25 d_{B_h}^{\max} \text{ and } P(\text{shadow}|p) < 0.8\}$$

where the last condition enforces shadow pixels being filtered out. Fig. 12 shows the results of up-slanted boundary detection of a truck and a bus.

Once $\overline{U}_{H_i}$ is obtained, all the points in $\overline{U}_{H_i}$ can be used to estimate the linearity of H_i. Assume $\overline{U}_{H_i}$ has N data points (x_i, y_i) and will be fitted to a straight-line model:

$$y = mx + b. \tag{7}$$

The parameters m and b can be obtained from (4) with a similar formulation. Then, the linearity of H_i is defined as

$$\text{Linearity}(H) = \exp\left(-\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i - mx_i - b)^2}\right). \tag{8}$$

Based on (6) and (8), the size and "linearity" features of a vehicle can be well extracted for further vehicle classification.

B. Classification by Vehicle Library

In the previous section, two features, i.e., size and linearity, have been described to represent vehicles. With these features, this section will design an optimal classifier for categorizing vehicles into different classes. Different from other vehicle-classification schemes [6], this paper makes two contributions to improving this classification work. First, it is known that a vehicle has many different appearances when it moves along a lane. In the past, most classification methods recognized a vehicle based on only one frame. If this frame is corrupted by noise, these approaches will fail to recognize the desired vehicles. If the designed classifier can integrate more cues from different frames (or appearances), more robustness and accuracies of vehicle classification can be then achieved. The second contribution is to let the designed recognizer make an optimal decision from different views. We build different templates in each vehicle category for overcoming the view changes of a vehicle. Each template captures different representative features of its corresponding vehicle type under different lighting conditions. Then, better decisions can be made if the designed classifier can consider different contributions of each template in each category.

Assume that there are K classes in the database and the kth class VC_k has n_k templates. Let V_j^k denote the jth vehicle in VC_k and $f_r(V_j^k)$ be the rth feature of V_j^k. Then, the variance of the rth feature in VC_k can be obtained as

$$\sigma_{r,k} = \sqrt{\frac{1}{n_k}\sum_{j=1}^{n_k}\left(f_r\left(V_j^k\right) - m_r^k\right)^2}$$

where $m_r^k = (1/n_k)\sum_{i=1}^{n_k} f_r(V_i^k)$. Then, given a vehicle H_i and a template V_j^k in VC_k, the similarity between H_i and V_j^k can be measured by

$$S_k\left(H_i, V_j^k\right) = \exp\left(-\sum_{r=1}^{2}\frac{\left(f_s(H_i) - f_r\left(V_j^k\right)\right)^2}{\sigma_{r,k}^2}\right) \tag{9}$$

where $r = 1$ for the size feature and $r = 2$ for the linearity feature. Then, the similarity of H_i to a vehicle class VC_k is defined by

$$S(H_i|\text{VC}_k) = \frac{1}{n_k}\sum_{V_j^k \in \text{VC}_k} S_k\left(H_i, V_j^k\right).$$

Furthermore, the probability of H_i belonging to VC_k is calculated by

$$P(\text{VC}_k|H_i) = \frac{S(H_i|\text{VC}_k)}{S_{\text{sum}}(H_i)} \tag{10}$$

where $S_{\text{sum}}(H_i) = \sum_k S(H_i|\text{VC}_k)$. Based on (10), H_i will be classified to class l if

$$l = \arg\max_k P(\text{VC}_k|H_i). \tag{11}$$

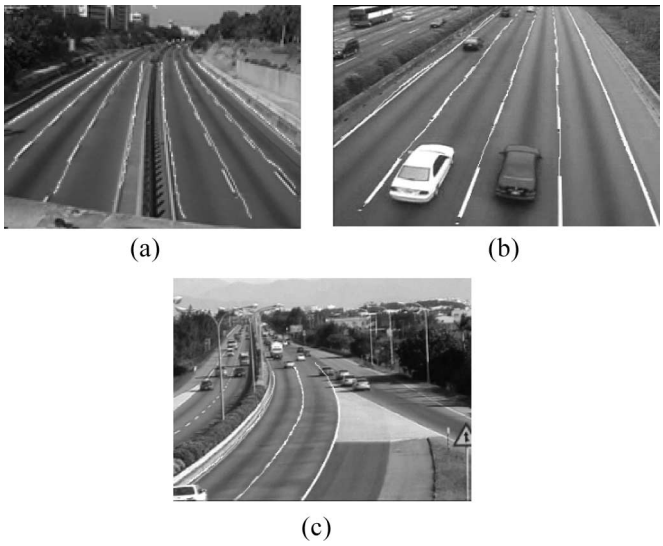

Fig. 13. Different detection results of dividing lines.

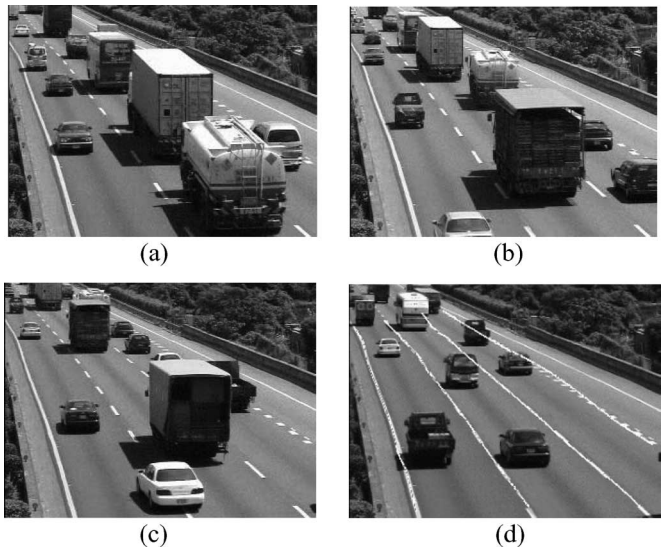

Fig. 14. Detection result of dividing lines when various lane changes happen during rush hour. A moving oil tanker has a lane change between (a) and (b). (c) A white car having lane change. (d) Detection result of lane-dividing lines.

Equation (11) classifies vehicles based on only one appearance of the verified vehicle. However, due to noise, only one vehicle appearance cannot provide enough support for robust vehicle classification. Assume that H_i^t denotes the result of H_i detected at the tth time frame. If all samples H_i^t of H_i can be used and integrated, the accuracy of vehicle classification can be much improved. Let Q_{H_i} be the set of all appearances H_i^t. For each H_i, the system classifies it into class l when

$$\forall k \neq l, \sum_{H_i \in Q_{H_i}} P(\mathrm{VC}_l|H_i) \geq \sum_{H_i \in Q_{H_i}} P(\mathrm{VC}_k|H_i). \quad (12)$$

Based on (12), each input vehicle can be correctly classified to its corresponding class.

VI. Experimental Results

In order to analyze the performance of the proposed method, four image sequences were used. All the sequences were acquired from Taiwan's highways with a fixed camera. Figs. 13 and 14 show different shots of these four sequences. In these sequences, they included 2000, 5500, 3500, and 5400 frames, respectively. For the first set of experiments, the performance of our proposed detection algorithm of lane-dividing lines was examined. In this set of experiments, all vehicles, even occluded, were extracted by image differencing and used as seeds to obtain their corresponding vehicle histograms from which desired lane-dividing lines can be found. Fig. 13 shows the detection results of dividing lines extracted from the first three vehicle sequences. Clearly, all the desired dividing lines were correctly extracted. When traffic is heavy, different lane changes will happen and will disturb the work of vehicle voting. Another experiment was performed to demonstrate the superiority of our proposed method to detect lane-dividing lines from the above condition. Fig. 14 shows the result detected from the fourth sequence, which was captured in rush hour and included various lane changes. In Fig. 14(a) and (b), there was a lane change caused by the oil tanker. In Fig. 14(c), another case of lane change was caused by the white car. Fig. 14(d) is the detection result of lane-dividing lines using our histogram-based method. In real conditions, although the traffic jam will include many lane changes, it also brings more regular vehicles moving on the same lane for the vehicle voting. It is noticed that regular vehicle movements always happen more frequently than irregular ones. The existence of lane change can be considered as a kind of noise. Since our method is statistics based, the unexpected effect of lane change can be easily removed using a smoothing technique (see step 3) in this algorithm). The accuracy of our algorithm can be easily verified by comparing the differences between the true lane-dividing lines and the estimated ones.

In the second set of experiments, the performance of our proposed shadow-elimination algorithm was examined. Fig. 8 shows the first case of shadows and its corresponding result of shadow elimination is shown in Fig. 9. Fig. 15 shows another two cases of vehicle shadows. Fig. 15(a) and (c) are two original images, and Fig. 15(b) and (d) are the results of shadow elimination obtain from Fig. 15(a) and (c), respectively. Clearly, the van truck in Fig. 15(c) was successfully separated from another truck. Once the vehicle shadows are successfully removed, the following vehicle counting and feature extraction can be achieved more accurately. For most vehicle-recognition systems, the size feature is the most important feature but easily affected by shadows. In our proposed ITS, since vehicle shadows are eliminated in advance, the vehicle size can be more accurately estimated than other methods. Thus, in this paper, significant improvements of accuracy in vehicle recognition can be achieved.

Fig. 16 shows the detection results of vehicles occluded by shadows. These vehicles were difficultly separated if the information of lane-dividing lines was not used. In Fig. 16(c), only two lanes were observed and the vehicles above the red line were ignored due to their small sizes. Table I shows the comparisons of vehicle counting between different methods

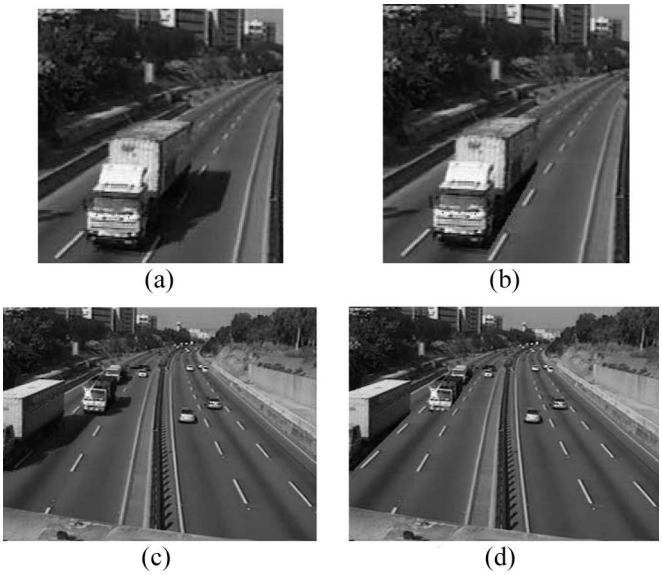

Fig. 15. Results of shadow elimination. (a) and (c) Original images. (b) and (d) Results of shadow elimination obtained from (a) and (c), respectively.

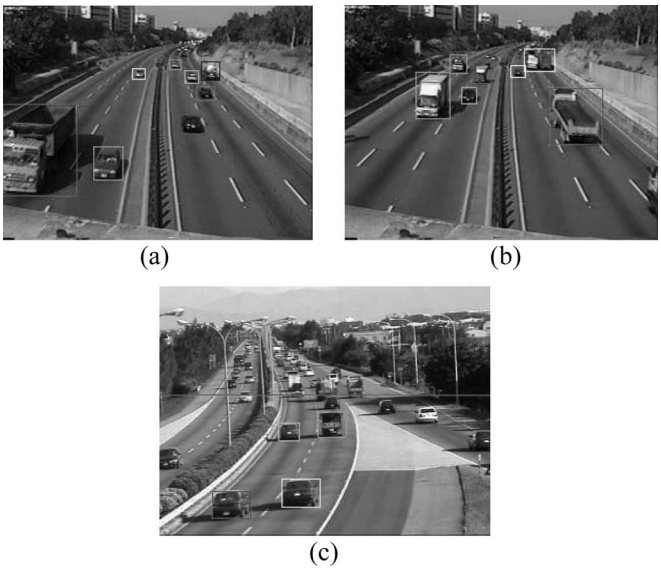

Fig. 16. Different detection results of vehicles when shadows exist. For all these cases, each vehicle can be well detected after shadow elimination.

TABLE I
ACCURACY COMPARISONS OF VEHICLE COUNTING WITH OR WITHOUT SHADOW ELIMINATION. THE COUNTING ERRORS ARE MAINLY CAUSED BY VEHICLE OCCLUSION

Counts / Sequence	Actual Number of Vehicles	Estimation with Shadow Elimination		Estimation without Shadow Elimination	
		Number	Accuracy	Number	Accuracy
Video 1	4945	3687	74.56%	2695	65.79%
Video 2	12521	10382	82.91%	8992	71.81%
Video 3	2977	2728	91.66%	2489	83.60%
Total	20443	16797	82.16%	14446	69.34%

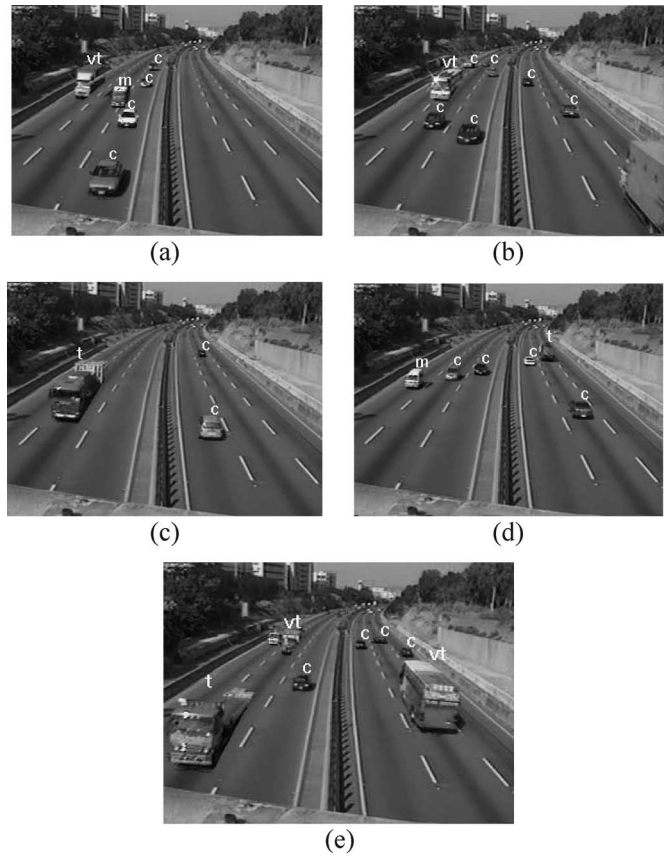

Fig. 17. Results of vehicle classification. The symbols "c," "m," "vt," and "t" denote the types of cars, minivans, van trucks, and trucks, respectively.

if our shadow-elimination method was applied or not. In this table, when counting vehicles, no recognition process was used. The inaccuracy in vehicle counting was mainly caused by the occlusions due to perspective effects. Clearly, the counting method with shadow elimination performed much better than the one without shadow elimination.

After vehicle extraction, a novel classification scheme is then used for vehicle classification. In this paper, according to the size and "linearity" features, four categories, i.e., car, minivan, van truck (including bus), and truck, were used for vehicle classification. Fig. 17 shows the classification results of vehicles appearing in the analyzed video sequences. Each vehicle after classification was labeled with a symbol, where the symbols "c," "m," "t," and "vt" mean a "car," "minivan," "truck," and "van truck," respectively. If only the size feature of the vehicle is used, it is difficult to discriminate a bus from a truck. Fig. 17(a) is the recognition result of Fig. 6. If the shadow-elimination algorithm is not applied, the three occluded vehicles in Fig. 6 will not be correctly classified.

Table II shows the comparisons of vehicle classification between different methods with or without shadow elimination. In this table, the linearity feature was included for discriminating trucks and van trucks to different classes. On the other hand, in order to reflect the real effect of shadow elimination, this table did not include the case of vehicles if they were occluded by perspective effects. Clearly, if a shadow-elimination process is adopted, the robustness and accuracy of recognition can be much enhanced. Table III shows the comparison results when the linearity feature was used or not, where the symbols m, vt, and t were used for denoting the "minivan," "van truck," and "truck," respectively. Since the category "car" was little

TABLE II
ACCURACY COMPARISONS OF VEHICLE RECOGNITION BETWEEN METHODS WITH OR WITHOUT SHADOW ELIMINATION

Analysis Sequence	Actual Vehicle Number				Recognition without Shadow Elimination					Recognition with Shadow Elimination				
	Car	MV	VT	T	Car	MV	VT	T	Accuracy	Car	MV	VT	T	Accuracy
Video 1	2157	1111	190	164	1867	822	155	128	82.1%	2064	922	170	145	91.1%
Video 2	7835	1212	1107	1113	6276	848	874	868	78.7%	7129	945	952	946	88.5%
Video 3	1225	564	670	278	1004	301	529	219	75.0%	1127	451	582	244	87.8%

TABLE III
ACCURACY COMPARISONS OF VEHICLE-RECOGNITION METHODS WHEN THE LINEARITY FEATURE IS USED OR NOT

Analysis Videos	Actual Vehicle Number			Recognition with Vehicle Size				Recognition with Vehicle Length				Recognition with Vehicle Size and Linearity Feature			
	m	*vt*	*t*	*m*	*vt*	*t*	Accur.	*m*	*vt*	*t*	Accur.	*m*	*vt*	*t*	Accur.
V1	1111	190	164	732	115	101	64.7%	725	109	107	64.2%	922	170	145	84.4%
V2	1212	1107	1113	751	652	650	59.8%	745	639	641	58.9%	945	952	946	82.8%
V3	564	670	278	341	395	179	60.5%	334	384	172	58.8%	451	582	244	84.45%

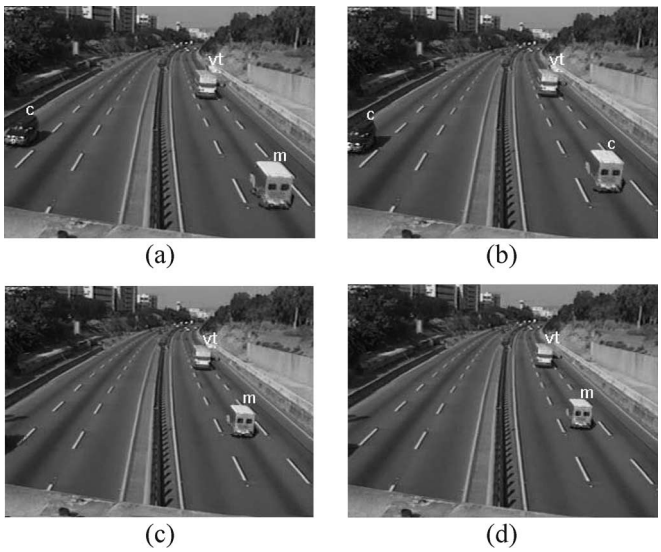

Fig. 18. Results of truck classification across different frames. (a), (c), and (d) Correct vehicle classification; (b) misclassification. If an integration classification scheme is used, the error in (b) can be avoided.

related to the linearity feature, this table did not include it for this comparison. In this table, for comparisons, two features "vehicle size" and "vehicle length" were also used. Since each vehicle has similar width, the feature of "vehicle length" had similar performances to the feature of vehicle size. Clearly, when the linearity feature was used, improvements in recognition accuracy were achieved, especially in the car types "truck" and "van truck."

On the other hand, when classifying vehicles, if more than one frame is used, the accuracy of recognition can also be further improved. For example, the minivan in Fig. 18(a), (c), and (d) were correctly classified but the one in Fig. 18(b) was misclassified. However, the misclassification in Fig. 18(b) can be avoided if we can integrate the results of Fig. 18(a), (b), and (d) together. The appearances of a vehicle moving at the road can be integrated together using the technique of the Kalman filter [26]. Then, a better decision in vehicle classification can be made by the above integration. Table III lists the classification results of vehicles when only one frame was used. Table IV lists the classification result when our integration technique was adopted or not (see Section V-B). Obviously, our proposed integration scheme performed much more accurately than the scheme using only one frame to classify vehicles. The superiority of the proposed method had been successfully verified through the preceding experimental results.

VII. CONCLUSION

In this paper, we have proposed a novel vehicle-classification scheme for estimating important traffic parameters from video sequences. In this approach, for robustness consideration, a background update method is first used to keep background static. Then, desired vehicles can be detected through image differencing and then tracked by a Kalman filter. Furthermore, through shadow elimination and feature extraction, several features including vehicle size and linearity features can be extracted. Then, based on these features, an optimal classifier is designed to accurately categorize input vehicles into different categories like cars, buses, and trucks. The contributions of this paper can be summarized as follows.

1) A new algorithm for detecting lane-dividing lines was proposed. With the information, vehicle occlusions caused by shadows can be easily solved.
2) A new normalization method was proposed to normalize vehicle size without using any 3-D information. Hence, the vehicle size can be kept constant for more accurately classifying vehicles.
3) A new shadow-elimination algorithm was proposed to remove unwanted shadows from video sequences. Due to the elimination, the occlusion caused by shadows can be reduced to a minimum.
4) A new feature called "linearity" was defined to classify vehicles. Experimental results have proved that the new feature works well to categorize vehicles into more detailed types.
5) An optimal classifier was designed to overcome different environmental variations. Since the proposed classifier can integrate different supports from vehicles, improvements of accuracy in vehicle classification have been achieved.

TABLE IV
ACCURACY COMPARISONS OF VEHICLE-RECOGNITION METHODS WITH ONE OR MULTIPLE FRAMES

Analysis Sequence	Actual Vehicle Number				Recognition with One Frame					Recognition with Multiple Frames				
	car	*m*	*vt*	*t*	*car*	*m*	*vt*	*t*	Accur.	*car*	*m*	*vt*	*t*	Accur.
Video 1	310	142	32	22	290	113	25	16	87.7%	301	129	29	20	94.8%
Video 2	1994	162	143	146	1814	124	121	120	89.1%	1860	139	131	135	92.6%
Video 3	167	78	89	41	151	63	75	34	86.1%	159	70	81	38	92.8%

Experimental results have shown that our method is superior in terms of accuracy, robustness, and stability in vehicle classification.

REFERENCES

[1] Y. K. Jung, K. W. Lee, and Y. S. Ho, "Content-based event retrieval using semantic scene interpretation for automated traffic surveillance," *IEEE Trans. Intell. Transp. Syst.*, vol. 2, no. 3, pp. 151–163, Sep. 2001.

[2] B. Rao, H. F. Durrant-Whyte, and J. Sheen, "A fully decentralized multi-sensor system for tracking and surveillance," *Int. J. Rob. Res.*, vol. 12, no. 1, pp. 20–44, Feb. 1993.

[3] M. Bertozzi, A. Broggi, and A. Fascioli, "Vision-based intelligent vehicles: State of the art and perspectives," *Robot. Auton. Syst.*, vol. 32, no. 1, pp. 1–16, 2000.

[4] G. D. Sullivan, K. D. Baker, A. D. Worrall, C. I. Attwood, and P. M. Remagnino, "Model-based vehicle detection and classification using orthographic approximations," *Image Vis. Comput.*, vol. 15, no. 8, pp. 649–654, Aug. 1997.

[5] T. Yoshida *et al.*, "Vehicle classification system with local-feature based algorithm using CG model images," *IEICE Trans. Inf. Syst.*, vol. E85D, no. 11, pp. 1745–1752, Nov. 2002.

[6] S. Gupte, O. Masoud, R. F. K. Martin, and N. P. Papanikolopoulos, "Detection and classification of vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 3, no. 1, pp. 37–47, Mar. 2002.

[7] O. Masoud, N. P. Papanikolopoulos, and E. Kwon, "The use of computer vision in monitoring weaving sections," *IEEE Trans. Intell. Transp. Syst.*, vol. 2, no. 1, pp. 18–25, Mar. 2001.

[8] C. E. Smith, C. A. Richards, S. A. Brandt, and N. P. Papanikolopoulos, "Visual tracking for intelligent vehicle-highway systems," *IEEE Trans. Veh. Technol.*, vol. 45, no. 4, pp. 744–759, Nov. 1996.

[9] A. J. Lipton, H. Fujiyoshi, and R. S. Patil, "Moving target classification and tracking from real-time video," in *Proc. IEEE Workshop Appl. Comput. Vis.*, 1998, pp. 8–14.

[10] D. Gao and J. Zhou, "Adaptive background estimation for real-time traffic monitoring," in *Proc. IEEE Conf. Intell. Transp. Syst.*, Aug. 2001, pp. 330–333.

[11] M. Fathy and M. Y. Siyal, "A window-based image processing technique for quantitative and qualitative analysis of road traffic parameters," *IEEE Trans. Veh. Technol.*, vol. 47, no. 4, pp. 1342–1349, Nov. 1998.

[12] S. Kamijo, Y. Matsushita, K. Ikeuchi, and M. Sakauchi, "Traffic monitoring and accident detection at intersections," *IEEE Trans. Intell. Transp. Syst.*, vol. 1, no. 2, pp. 108–118, Jun. 2000.

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

[14] R. Cucchiara, M. Piccardi, and P. Mello, "Image analysis and rule-based reasoning for a traffic monitoring system," *IEEE Trans. Intell. Transp. Syst.*, vol. 1, no. 2, pp. 119–130, Jun. 2002.

[15] Y. Iwasaki, "A measurement method of pedestrian traffic flow by use of image processing and its application to a pedestrian traffic signal control," in *Proc. IEEE Conf. Intell. Transp. Syst.*, Oct. 1999, pp. 310–313.

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

[17] B. Coifman, D. Beymer, P. McLauchlan, and J. Malik, "A real-time computer vision system for vehicle tracking and traffic surveillance," *Transp. Res. Part C*, vol. 6, no. 4, pp. 271–288, 1998.

[18] G. S. K. Fung, N. H. C. Yung, and G. K. H. Pang, "Vehicle shape approximation from motion for visual traffic surveillance," in *Proc. IEEE Conf. Intell. Transp. Syst.*, Aug. 2001, pp. 608–613.

[19] J. E. Boyd, J. Meloche, and Y. Vardi, "Statistical tracking in video traffic surveillance," in *Proc. IEEE Conf. Comput. Vis.*, Sep. 1999, pp. 163–168.

[20] M. Haag and H. Nagel, "Combination of edge element and optical flow estimates for 3D-model-based vehicle tracking in traffic image sequences," *Int. J. Comput. Vis.*, vol. 35, no. 3, pp. 295–319, Dec. 1999.

[21] X. Tao, M. Guo, and B. Zhang, "A neural network approach to elimination of road shadow for outdoor mobile robot," in *Proc. IEEE Int. Conf. Intell. Process. Syst.*, Beijing, China, 1997, pp. 1302–1306.

[22] K. Onoguchi, "Shadow elimination method for moving object detection," in *Proc. 14th Int. Conf. Pattern Recog.*, 1998, vol. 1, pp. 583–587.

[23] J. Stauder, R. Mech, and J. Ostermann, "Detection of moving cast shadows for object segmentation," *IEEE Trans. Multimedia*, vol. 1, no. 1, pp. 65–76, Mar. 1999.

[24] Y. Ivanov, A. Bobick, and J. Liu, "Fast lighting independent background subtraction," *Int. J. Comput. Vis.*, vol. 37, no. 2, pp. 199–207, Jun. 2000.

[25] W. L. Hsu, H. Y. Liao, B. S. Jeng, and K. C. Fan, "Real-time traffic parameter extraction using entropy," *Proc. Inst. Elect. Eng.—Vis. Image Signal Process.*, vol. 151, no. 3, pp. 194–202, Jun. 2004.

[26] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Trans. ASME—J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, Mar. 1960.

[27] B. N. Nelson, "Automatic vehicle detection in infrared imagery using a fuzzy inference based classification system," *IEEE Trans. Fuzzy Syst.*, vol. 9, no. 1, pp. 53–61, Feb. 2001.

[28] L. Bruzzone, D. F. Prieto, and S. B. Serpico, "A neural-statistical approach to multitemporal and multisource remote-sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 37, no. 3, pp. 1350–1359, May 1999.

[29] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C—The Art of Scientific Computing*. Cambridge, U.K.: Cambridge Univ. Press, 1992.

[30] M. Sonka, V. Hlavac, and R. Boyle, *Image Processing, Analysis and Machine Vision*. London, U.K.: Chapman & Hall, 1993.

Jun-Wei Hsieh (M'06) received the Bachelor's degree in computer science from TongHai University, Taichung, Taiwan, R.O.C., in 1990 and the Ph.D. degree in computer engineering from the National Central University, Taoyuan, Taiwan, in 1995. He was a recipient of the Phai–Tao–Phai award when he graduated.

From 1996 to 2000, he was a Researcher Fellow at the Industrial Technology Researcher Institute, Hsinchu, Taiwan, and led a team to develop video-related technologies. He is presently an Assistant Professor at the Department of Electrical Engineering, Yuan Ze University, Chung-Li, Taiwan. His research interests include content-based multimedia databases, video indexing and retrieval, computer vision, and pattern recognition.

Dr. Hsieh received the Best Paper Award for the 2005 IPPR Conference on Computer Vision, Graphics, and Image Processing.

Shih-Hao Yu was born in Taiwan, R.O.C., on September 19, 1977. He received the B.S. degree from Fu Jen Catholic University, Hsinchu, Taiwan, in 2000 and the M.S. degree from Yuan Ze University, Chung-Li, Taiwan, in 2002, all in electrical engineering.

His research interests include image processing, computer vision, and pattern recognition.

Yung-Sheng Chen (M'93) was born in Taiwan, R.O.C., on June 30, 1961. He received the B.S. degree from Chung Yuan Christian University, Chung-Li, Taiwan, in 1983 and the M.S. and Ph.D. degrees from National Tsing Hua University, Hsinchu, Taiwan, in 1985, and 1989, respectively, all in electrical engineering.

In 1991, he joined the Electrical Engineering Department, Yuan Ze Institute of Technology, Chung-Li, where he is now a Professor. His research interests include human visual perception, computer vision, circuit system, and teaching web design.

Dr. Chen received the Best Paper Award from the Chinese Institute of Engineers in 1989 and an Outstanding Teaching Award from Yuan Ze University in 2005. He has been listed in the *Who's Who of the World* since 1998. He is a member of the IPPR of Taiwan, R.O.C.

Wen-Fong Hu was born in Taiwan, R.O.C., on April 2, 1965. He received the B.S. and M.S. degrees in electronic engineering and automatic control engineering from Feng Chia University, Taichung, Taiwan, in 1987 and 1990, respectively. He is currently working toward the Ph.D. degree in the Department of Electrical Engineering, Yuan Ze University, Chung-Li, Taiwan.

His research interests include neural network and image processing.