A Real-Time Vision System for Nighttime Vehicle Detection and Traffic Surveillance

Yen-Lin Chen, *Member, IEEE*, Bing-Fei Wu, *Senior Member, IEEE*, Hao-Yu Huang, and Chung-Jui Fan

***Abstract*—This paper presents an effective traffic surveillance system for detecting and tracking moving vehicles in nighttime traffic scenes. The proposed method identifies vehicles by detecting and locating vehicle headlights and taillights using image segmentation and pattern analysis techniques. First, a fast bright-object segmentation process based on automatic multilevel histogram thresholding is applied to effectively extract bright objects of interest. This automatic multilevel thresholding approach provides a robust and adaptable detection system that operates well under various nighttime illumination conditions. The extracted bright objects are then processed by a spatial clustering and tracking procedure that locates and analyzes the spatial and temporal features of vehicle light patterns, and identifies and classifies moving cars and motorbikes in traffic scenes. The proposed real-time vision system has also been implemented and evaluated on a TI DM642 DSP-based embedded platform. The system is set up on elevated platforms to perform traffic surveillance on real highways and urban roads. Experimental results demonstrate that the proposed traffic surveillance approach is feasible and effective for vehicle detection and identification in various nighttime environments.**

***Index Terms*—Traffic information system, traffic surveillance, nighttime surveillance, vehicle detection, vehicle tracking.**

I. Introduction

DETECTING and recognizing vehicles are an important emerging research area for intelligent transportation systems. Previous studies on this topic have discussed traffic surveillance [1]–[21], driver assistance systems and autonomous vehicle guidance [22]–[33], and road traffic information systems [34]–[36]. In traffic surveillance applications, information about moving vehicles may be obtained from loop detectors, slit sensors, or cameras. Among the aforementioned sensors, camera-based systems can provide much more traffic analysis information, including traffic flow, vehicle classification, and vehicle speed.

Manuscript received October 4, 2009; revised February 12, 2010 and April 27, 2010; accepted June 2, 2010. Date of publication July 26, 2010; date of current version April 13, 2011. This work was supported by the National Science Council of Taiwan under Grants NSC-98-2220-E-027-010, NSC-99-2221-E-009-159, and NSC-99-2622-E-027-019-CC3.

Y.-L. Chen is with the Department of Computer Science and Information Engineering, National Taipei University of Technology, Taipei 10608, Taiwan (e-mail: ylchen@csie.ntut.edu.tw).

B.-F. Wu and H.-Y. Huang are with the Institute of Electrical and Control Engineering, National Chiao Tung University, Hsinchu 30050, Taiwan (e-mail: bwu@cssp.cn.nctu.edu.tw; grant@cssp.cn.nctu.edu.tw).

C.-J. Fan is with Acer Unipac Optronics (AUO) Corporation, Hsinchu 30078, Taiwan (e-mail: cjfan@cssp.cn.nctu.edu.tw).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TIE.2010.2055771

Due to the falling costs and growing power of computers, vision-based technologies have become popular solutions for traffic surveillance and intelligent transportation systems. Many researchers have developed valuable techniques for detecting and recognizing vehicles and obstacles in traffic images [2]–[21], [34]–[36]. By monitoring illumination changes in prespecified detection regions, techniques based on virtual slit or virtual loop detectors [2], [34]–[36] can rapidly detect vehicles moving through their predetermined monitoring regions. However, these methods are limited to the number of vehicles passing through the specific detection regions and are difficult to apply for vehicle classification, vehicle speed detection, and vehicle motion analysis.

To more efficiently obtain traffic information from moving vehicles, techniques based on frame differencing [3]–[9] have been applied to differentiate moving vehicles from motionless background scenes based on change detection or other statistical models. Other studies [3][4] use spatial–temporal difference features to segment moving vehicles, while the methods in [5]–[9] utilize techniques based on background subtraction to extract moving vehicles. These methods can be efficiently applied to daytime traffic scenes with stationary and unchanged lighting conditions. However, spatial–temporal difference features are no longer reliable when vehicles stop or move slowly in congested traffic areas, and vehicles may be falsely detected as background regions and missed. Moreover, in poorly illuminated or nighttime conditions, background scenes are substantially affected by the lighting effect of moving vehicles, making the reliable hypotheses of background models which are effective for vehicle detection during daytime invalid. Thus, most of the aforementioned frame-differencing techniques may be unreliable for handling nighttime and congested traffic environments.

To efficiently deal with slowly moving or stationary vehicle and nighttime traffic scenes, researchers have developed model- and feature-based techniques [10]–[15] to detect and track vehicles. Gardner and Lawton [10] and Sullivan *et al.* [11] applied 3-D vehicle shape models to detect and track vehicles in traffic scenes. Although 3-D models can provide high accuracy in detecting vehicles in less crowded traffic scenes, it requires a large set of vehicle models and has high computational costs. Moreover, the effectiveness of this approach relies greatly on detailed vehicle shapes and thus is not reliable in congested and complicated traffic scenes with cluttered vehicles with indistinct shapes. The methods in [12]–[14] employ an active contour model to extract the contours of vehicles in traffic scenes and then use these extracted contours to detect and track vehicles. Contour-based methods can provide greater computational

efficiency in detecting and tracking vehicles. However, their detection effectiveness depends greatly on the appropriate initialization of the contour models applied to the vehicles. Thus, these methods are sensitive to noisy objects, occlusions, and variations in illumination. Tsai *et al.* [15] proposed an approach for detecting vehicles in still images based on color and edge features. This approach can detect vehicles without motion information, allowing static or slowly moving vehicles to be efficiently detected from image sequences. Most of the aforementioned methods rely on some hypothetical vehicle appearance cues, which are only valid and efficient in daytime with sufficient ambient illuminations.

However, at night, and under darkly illuminated conditions in general, headlights and taillights are the only salient features of moving vehicles. In addition, there are many other sources of illumination that coexist with vehicle lights in nighttime traffic scenes, including street lamps, traffic lights, and ground-level road reflector plates. These nonvehicle sources of illumination make it very difficult to obtain cues for detecting vehicles in nighttime traffic scenes. To detect salient objects in nighttime traffic surveillance, Beymer *et al.* [16] presented a feature-based technique that extracts and tracks the corner features of moving vehicles instead of their entire regions. Their approach works in both daytime and nighttime traffic environments and is more robust to partial or complete occlusions. However, this technique suffers from high computational costs, as it must simultaneously process numerous features of moving vehicles and is unable to classify the types of vehicles detected. Huang *et al.* [17] proposed a method based on block-based contrast analysis and interframe change information. This contrast-based method can effectively detect outdoor objects in a given surveillance area using a stationary camera. However, contrast and interframe change information are sensitive to the lighting effects of moving vehicle headlights, resulting in erroneous vehicle detection.

Recently, vehicle lights have been used as salient features for nighttime vehicle detection applications for traffic monitoring systems [18]–[21] and driver assistance systems [29]–[33]. For traffic surveillance applications, these methods use morphological operations to extract candidate headlight objects and then perform shape analysis [18], template matching [19], or pattern classification [20] to find the paired headlights of moving vehicles. Nevertheless, since the aforementioned methods attempt to locate paired headlights in every single frame, the ambiguities of pairing headlights may cause many false detections and occlusions when vehicles move alongside each other. Robert [21] proposed a nighttime vehicle detection system that first detects paired headlights, uses a supervised machine learning system to verify and recognize actual vehicles, and finally applies a tracking approach to resolve occlusions. However, the detection performance of this approach is limited by its reliance on suitable prior training samples of eigenvehicles and the undergone large computational costs of performing multiple pattern classifiers. Moreover, most of the aforementioned methods are unable to deal with vehicles with single headlights, such as motorbikes.

This paper presents effective nighttime vehicle detection, tracking, and identification approaches for traffic surveillance by locating and analyzing the spatial and temporal features of vehicle lights. By improving our previous work in [37], this paper presents a system comprising the following processing stages for efficiently detecting and classifying moving vehicles at night. First, a fast bright-object segmentation process based on automatic multilevel histogram thresholding is performed to extract pixels of bright objects from the captured image sequences of nighttime traffic scenes. The advantage of this automatic multilevel thresholding approach is its robustness and adaptability when dealing with various illumination conditions at night. Then, to locate the connected components of these bright objects, a connected-component analysis procedure is applied to the bright pixels obtained by the previous stage. A spatial clustering process then groups these bright components to obtain groups of vehicle lights for potential moving cars and motorbikes. Next, a feature-based vehicle tracking and identification process is applied to analyze the spatial and temporal information of these potential vehicle light groups from consecutive frames. This step also refines the detection results and corrects for errors and occlusions caused by noise and errors during the segmentation and spatial clustering processes. Thus, actual vehicles and their types can be efficiently detected and verified from these tracked potential vehicles to obtain accurate traffic flow information. We also implemented the proposed real-time vision system on a TI DM642 DSP-based embedded platform and set up the system on elevated platforms near highways or urban roads. Experimental results show that the proposed traffic surveillance system is feasible and effective for vehicle detection and identification in various nighttime environments.

(a) (b)

Fig. 1. Typical examples of nighttime traffic scenes. (a) Sample nighttime traffic scene 1: An urban road. (b) Sample nighttime traffic scene 2: A highway.

The rest of this paper is organized as follows. Section II describes the lighting object segmentation approach for adaptively extracting possible vehicle lights. Section III presents a spatial clustering process on the obtained lighting objects. Then, Section IV proposes a feature-based vehicle tracking and identification process for tracking and identifying actual vehicles based on spatial and temporal information. Next, Section V illustrates the comparative experimental results. Finally, Section VI states the conclusions of this study.

II. Lighting Object Extraction

The first step in detecting and extracting moving vehicles from nighttime traffic scenes is to segment the salient objects of moving vehicles from traffic image sequences. Fig. 1 shows samples of typical nighttime traffic scenes from an urban road and highway under different environmental illumination conditions. These sample figures depict that, in typical nighttime traffic scenes, there are moving cars and motorbikes on the road, and under poorly or brightly environmental illuminated

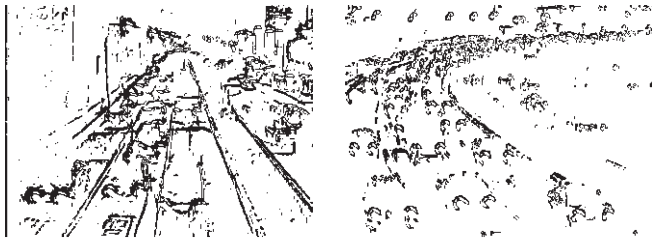


Fig. 2. Edge features extracted from sample nighttime traffic scenes in Fig. 1. (a) Edge features extracted from sample scene 1 in Fig. 1(a). (b) Edge features extracted from sample scene 2 in Fig. 1(b).

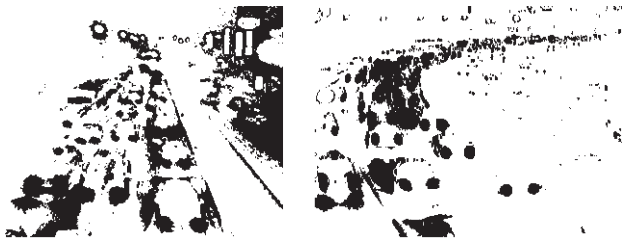


Fig. 3. Foreground object images extracted from sample nighttime traffic scenes in Fig. 1 by background subtraction techniques. (a) Foreground image extracted from sample scene 1 in Fig. 1(a) by background subtraction. (b) Foreground image extracted from sample scene 2 in Fig. 1(b) by background subtraction.

conditions, vehicle lights are the only valid salient features. In addition to the vehicle lights, some lamps, traffic lights, and signs are also visible sources of illumination in the image sequences of nighttime traffic scenes.

As Section I mentions, most recent studies on vehicle detection adopt edge maps, frame differencing, and background subtraction techniques to extract the features of moving vehicles from traffic scenes. However, as Figs. 2 and 3 show, these techniques, which are effective for vehicle detection in daytime, become invalid in nighttime illumination conditions. This is because the background scenes are greatly affected by the varying lighting effect of moving vehicles. Fig. 2 shows that the edge features of moving vehicles extracted from nighttime traffic scenes are mostly adjoined and blurred. There are also numerous nonvehicle lighting objects, such as reflective lane markings, road surfaces, and reflected beams caused by nearby moving vehicle lights. Similarly, Fig. 3 shows the foreground object images extracted by background-subtraction-based methods. This figure shows that moving vehicles in the extracted foreground object images are mostly smeared, with large spurious nonvehicle and background regions. This is because the illumination intensities of these background regions are sharply varied due to fluctuating lighting effects, preventing such techniques from obtaining an appropriate background model.

Therefore, performing vehicle detection and recognition for nighttime traffic surveillance requires an effective approach for correctly and rapidly locating and extracting the salient features of vehicle lights under poorly illuminated conditions. This would enable the efficient extraction and segmentation of the object regions of moving vehicles. Therefore, this section presents a fast bright-object segmentation process based on automatic multilevel histogram thresholding. The proposed method extracts the bright-object pixels of moving vehicles from image sequences of nighttime traffic scenes.

Fig. 4. Results of performing the bright-object segmentation process on the traffic-scene images in Fig. 1. (a) Lighting objects extracted from Fig. 1(a). (b) Lighting objects extracted from Fig. 1(b).

The first step in the bright-object extraction process is to extract bright objects from the road image to facilitate subsequent rule-based classification and tracking processes. To reduce the computational complexity of extracting bright objects, we first extracted the grayscale image, i.e., the Y-channel, of the grabbed image by performing an RGB-to-Y transformation. To extract bright objects from a given transformed gray-intensity image, the pixels of bright objects must be separated from other object pixels of different illuminations. For this purpose, we have presented a fast effective multilevel thresholding technique [38]. In this paper, this effective multilevel thresholding technique is applied to automatically determine the appropriate levels of segmentation for extracting bright-object regions from traffic-scene image sequences. More detailed descriptions of this multilevel thresholding technique can be found in [38].

By applying this multilevel thresholding technique, the lighting object regions of moving vehicles can be efficiently and adaptively segmented under various environmental illumination conditions in different nighttime traffic scenes as in Fig. 1(a) and (b). As a result, lighting objects can be appropriately extracted from other objects contained in nighttime traffic scenes. Accordingly, as Fig. 4(a) and (b) shows, performing this lighting object segmentation process successfully separates the lighting objects of interest in Fig. 1 into thresholded object planes under different environmental illumination conditions in nighttime traffic scenes.

III. Spatial Classification Process of Lighting Objects

To extract/obtain potential vehicle light components from the detection zone in the bright-object plane, a connected-component extraction process [39] can be performed to label and locate the connected components of the bright objects. Extracting the connected components reveals the meaningful features of location, dimension, and pixel distribution associated with each connected component. The location and dimension of a connected component can be represented by the bounding box surrounding it.

Since various nonvehicle light components, such as traffic lamps, road signs, road reflector plates, reflected beams, and some other illuminant objects, coexist with actual vehicle

Fig. 5. Illustration of detection area of traffic scenes.

Fig. 6. Illustration of extracted bright components in the detection area.

lights, we applied a spatial classification process to preliminarily detect potential vehicle lights and filter out nonvehicle components. These detected potential vehicle lights are then processed by the following vehicle light tracking and identification process to identify the actual moving vehicles.

To preliminarily screen out nonvehicle illuminating objects, such as street lamps and traffic lights located at the top side of traffic scenes, and to effectively and rapidly locate the sufficiently reliable and clear features of moving vehicles, and efficiently save the redundant computational costs for the embedded system implementation, we apply a detection area for each traffic scene. This detection area is the midline of the traffic-scene image and bounded by the most left and right lanes, as Fig. 5 shows. These lane boundaries were determined by performing a lane detection process derived from our previous study [28] in the system initialization. The connected-component extraction and spatial classification processes are only performed on the bright objects located in the detection area, as Fig. 6 shows.

To facilitate the description of the proposed spatial classification processes, we first define the bright connected components and their groups as follows.

1) C_i denotes the ith lighting component to be processed.
2) CS_k denotes the kth set of bright components $CS_k = \{C_i, i = 0, 1, \ldots, p\}$, while the amount of its contained lighting components is denoted as $|CS_k|$.
3) The locations of a certain component C_i employed in the spatial classification process are their top, bottom, left, and right coordinates, denoted as t_{C_i}, b_{C_i}, l_{C_i}, and r_{C_i}, respectively.
4) The width and height of a bright component C_i are denoted as $W(C_i)$ and $H(C_i)$, respectively.
5) The horizontal distance D_h and the vertical distance D_ν between a pair of ith and jth lighting components are defined as

$$D_h(C_i, C_j) = \max(l_{C_i}, l_{C_j}) - \min(r_{C_i}, r_{C_j}) \quad (1)$$

$$D_\nu(C_i, C_j) = \max(t_{C_i}, t_{C_j}) - \min(b_{C_i}, b_{C_j}). \quad (2)$$

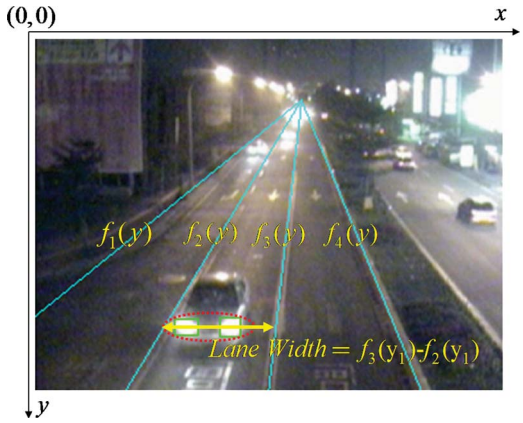


Fig. 7. Illustration of the approximated lane widths in the image coordinate.

If two bright components are overlapping in the horizontal or vertical direction, then the value of the $D_h(C_i, C_j)$ or $D_\nu(C_i, C_j)$ will be negative.

6) Hence, the measures of overlapping between the horizontal and vertical projections of the ith and jth bright components can be respectively computed as

$$P_h(C_i, C_j) = \frac{-D_h(C_i, C_j)}{\min[W(C_i), W(C_j)]} \quad (3)$$

$$P_\nu(C_i, C_j) = \frac{-D_\nu(C_i, C_j)}{\min[H(C_i), H(C_j)]}. \quad (4)$$

Fig. 7 shows the image coordinate system used for vehicle detection. In this image coordinate system, the vehicles located at a relatively distant place on the road will appear in a higher location and become progressively smaller until converging into a vanishing point. Therefore, the driving lanes stretched from the vanishing point can be modeled by a set of line equations by

$$f_l(y) = \frac{y - c_l}{m_l}; \; l = 1, 2, \ldots, L \quad (5)$$

where y denotes the vertical coordinate; m_l and c_l are the slope and intercept of the lth driving lane, respectively; and L represents the number of driving lanes. Here, the driving lanes are obtained by using the lane detection method of our previous study [28] in the system initialization process.

The approximate lane width associated with a bright component C_i at a distance on the image coordinate, denoted by $LW(C_i)$, can be obtained by

$$LW(C_i) = |f_{l+1}(C_Y(C_i)) - f_l(C_Y(C_i))| \quad (6)$$

where $C_Y(C_i)$ represents the vertical position of the component C_i on the image coordinate and is defined by $C_Y(C_i) = (t_{C_i} + b_{C_i})/2$.

Based on the aforementioned definitions of bright components, a preliminary classification procedure can be applied to the obtained bright components to identify potential vehicle light components and filter out most nonvehicle-illuminant light components, such as large ground reflectors and beams. For this purpose, a bright component C_i is identified as a potential vehicle light component if it satisfies the following conditions.

1) Since most vehicle lights have a nearly circular shape, the enclosing bounding box of a potential vehicle light component should form a square shape, i.e., the size-ratio feature of C_i must satisfy the following condition:

$$\tau_{RL} \leq W(C_i)/H(C_i) \leq \tau_{RH} \tag{7}$$

where the thresholds τ_{RL} and τ_{RH} for the size-ratio condition are set as 0.8 and 1.2, respectively, to determine the circular-shaped appearance of a potential vehicle light.

2) A vehicle light object should also have a reasonable area compared to the area of the lane. Thus, the area feature of C_i must satisfy the following condition:

$$\tau_{AL} < A(C_i) < \tau_{AH} \tag{8}$$

where the thresholds τ_{AL} and τ_{AH} for the area condition are determined as $\tau_{AL} = (LW(C_i)/8)^2$ and $\tau_{AH} = (LW(C_i)/4)^2$, respectively, to adaptively reflect the reasonable area characteristics of a potential vehicle light.

Accordingly, if two neighboring bright components C_i and C_j satisfy the following conditions, they are categorized as a homogeneous potential vehicle light set and are merged and clustered as a potential vehicle light set CS.

1) They are horizontally close to each other, i.e.,

$$D_h(C_i, C_j) < \min\left[W(C_i), W(C_j)\right]. \tag{9}$$

2) They are also vertically close to each other, i.e.,

$$D_v(C_i, C_j) < 2.0 \times \left(\min\left[H(C_i), H(C_j)\right]\right). \tag{10}$$

3) Two vertically overlapping bright objects with high horizontal projection profiles should be grouped the same group CS

$$P_h(C_i, C_j) > T_{hp} \tag{11}$$

where the threshold T_{hp} is chosen as 0.6 to reflect the vertical alignment characteristics of compound vehicle lights.

Figs. 8 and 9 show the results of the spatial clustering process. This process yields several sets of potential vehicle components CSs in the detection area, and these are labeled as $\boldsymbol{P}$ in the following tracking processes. For example, consider the bottom-right car with a set of compound headlights (marked by a yellow circle). Fig. 8 shows that its meaningful light components are preliminarily refined and grouped into sets of potential vehicle components, in which the light components of the bottom-right car are grouped into two potential vehicle component sets. This stage also filters out some nonvehicle bright components, such as reflected beams on the ground. Fig. 9 shows another sample of the spatial clustering process of bright components, in which the reflections of the headlights of the bottom-right car are excluded from the resulting potential vehicle component sets.

Note that the current stage does not yet merge the vehicle light sets on the two sides of the vehicle body into paired groups. This is because vehicles, which have paired light sets, and motorbikes, which have single-light sets, both exist in most nighttime road scenes. Therefore, without motion information in the subsequent frames, it is difficult to determine if the approaching light sets represent paired lights belonging to the same vehicle. Thus, the vehicle light tracking and identification process described in the following section is applied to these potential vehicle light sets to identify actual moving vehicles and motorbikes.

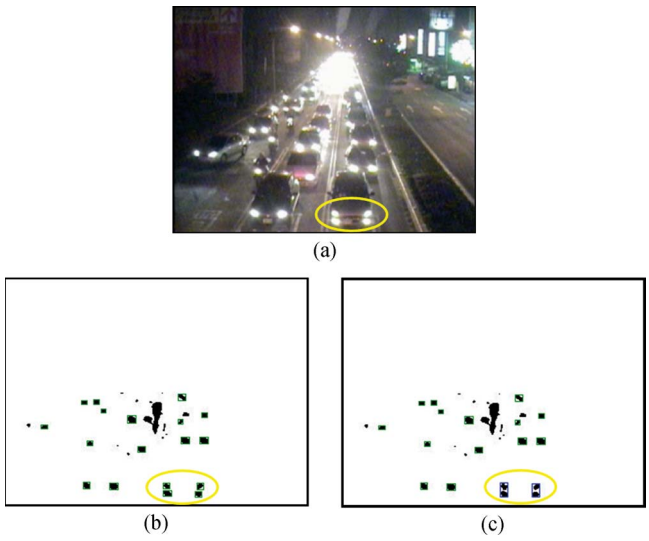


Fig. 8. Results of the spatial clustering of bright components. (a) Original traffic-scene image. (b) Extracted bright components after performing the preliminary classification procedure. (c) Sets of potential vehicle components obtained by (9)–(11).

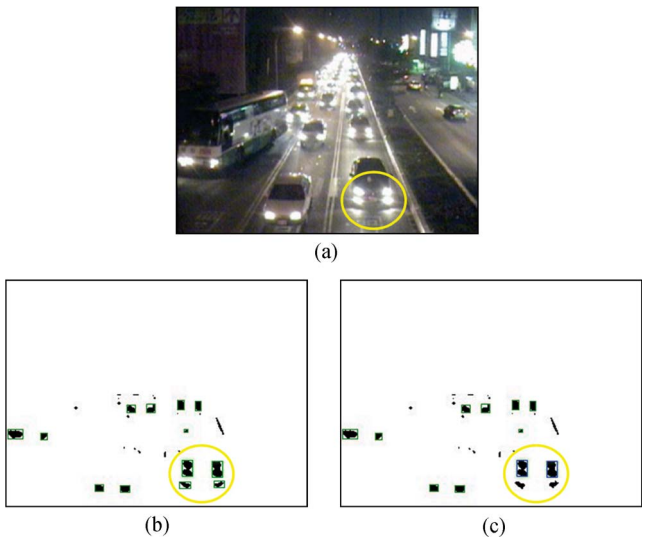


Fig. 9. Results of the spatial clustering of bright components when the reflected beams are filtered. (a) Original traffic-scene image. (b) Extracted bright components after performing the preliminary classification procedure. (c) Sets of potential vehicle components obtained by (9)–(11).

IV. Tracking and Identification of Potential Vehicles and Motorbikes

The aforementioned processes identify potential vehicle components entering the detection area, which are represented by lighting object sets in each image frame. However, since complete information for determining the types of potential

vehicles may not be immediately obtained from single image frames, a component-based tracking and grouping procedure is applied to analyze the motion information of these potential vehicle components based on consecutive image frames. This procedure accurately identifies and classifies moving vehicles and motorbikes. This tracking information is used to refine the detection results of potential vehicle components and correct the errors caused by occlusions, noise, and errors in the bright-object segmentation and spatial classification processes. The tracked potential vehicle components that move rigidly together are grouped as whole moving vehicles by evaluating their common motion information. After the whole moving vehicles are obtained, it is possible to identify the types of tracked vehicles.

The proposed vehicle tracking and identification process includes three phases. First, the phase of potential vehicle component tracking process is associated with the motion relation of vehicle components in succeeding frames by analyzing their spatial and temporal features. Then, the phase of motion-based grouping process is applied to the tracked vehicle components to construct whole moving vehicles. These moving vehicles are then tracked in the vehicle tracking phase. Finally, the vehicle recognition phase identifies and classifies the types of tracked vehicles.

A. Tracking Process of Potential Vehicle Components

When a potential vehicle component is initially detected in the detection area, a tracker will be created to associate this potential vehicle component with those in subsequent frames based on spatial–temporal features. The features used in the tracking process are described and defined as follows.

1) P_i^t denotes the ith potential vehicle component appearing in the detection zone in frame t. The location of P_i^t employed in the tracking process is represented by its central position, which can be expressed by

$$P_i^t = \left(\frac{l\left(P_i^t\right) + r\left(P_i^t\right)}{2}, \frac{t\left(P_i^t\right) + b\left(P_i^t\right)}{2} \right). \tag{12}$$

2) The tracker TP_i^t represents the trajectory of P_i, which has been tracked in sequential frames 1 to t, and is defined as

$$TP_i^t = \left\langle P_i^1, P_i^2, \ldots, P_i^t \right\rangle. \tag{13}$$

3) The overlapping score of the two potential vehicle components P_i^t and $P_j^{t'}$, detected at two different times t and t', can be computed using their area of intersection

$$S_o\left(P_i^t, P_j^{t'}\right) = \frac{A\left(P_i^t \cap P_j^{t'}\right)}{\mathrm{Max}\left(A\left(P_i^t\right), A\left(P_j^{t'}\right)\right)}. \tag{14}$$

In each recursion of the tracking process for a newly incoming frame t, the potential vehicle components appearing in the incoming frame, denoted by $\mathbf{P}^t = \{P_i^t | \, i = 1, \ldots, k'\}$, will be analyzed and associated with the set of potential vehicle components tracked in the previous frame $t-1$, denoted by $\mathbf{TP}^{t-1} = \{TP_j^{t-1} | \, j = 1, \ldots, k\}$. The set of tracked potential vehicles $\mathbf{TP}^t$ will then be updated according to the following process.

During the tracking process, a potential vehicle component might be in one of the three possible tracking states. The component tracking process applies different relevant operations according to the given states of each tracked potential vehicle component in each frame. The tracking states and associated operations for the tracked potential vehicle components are as follows.

1) **Update**: When a potential vehicle component $P_i^t \in \mathbf{P}^t$ in the current frame matches a tracked potential vehicle component $TP_j^{t-1} \in \mathbf{TP}^{t-1}$, then the tracker updates the set of the tracked potential components $\mathbf{TP}^t$ by associating P_i^t with the tracker TP_j^t if the following *tracker matching condition* is satisfied. This matching condition is

$$S_o\left(P_i^t, TP_j^{t-1}\right) > \tau_{mp} \tag{15}$$

where τ_{mp} is a predefined threshold that represents the reasonable spatial–temporal coherence for P_i^t to be associated with TP_j^{t-1}. For performing under free-flown traffic scenes with sufficiently high frame-grabbing rate, i.e., at least 15 frames per second (FPS), the movement of a potential component between two subsequent frames will probably be less than its size. Thus, a value of $\tau_{mp} = 0.25$ is experimentally determined to obtain sufficiently intact tracks.

2) **Appear**: If a newly appearing potential vehicle component $P_i^t \in \mathbf{P}^t$ does not match any $TP_j^{t-1} \in \mathbf{TP}^{t-1}$ at the previous time, then a new tracker is created for this potential vehicle component and appended to the updated set $\mathbf{TP}^t$.

3) **Disappear**: An existing tracker of potential vehicle component $TP_j^{t-1} \in \mathbf{TP}^{t-1}$ cannot be matched by any newly coming potential vehicle components $P_i^t \in \mathbf{P}^t$. A tracked potential vehicle component may sometimes be temporarily sheltered or occluded in some frames and will soon reappear in subsequent frames. Thus, to prevent this vehicle component from being regarded as a newly appearing potential vehicle, its tracker is retained for a span of 0.5 seconds, i.e., $0.5 \cdot FPS$ frames, where FPS denotes the grabbing frame rate (FPS) of the charge-coupled device (CCD) camera, to appropriately cope with vehicles leaving straightforward or making turns. If a tracker of potential vehicle component TP_j^{t-1} cannot be matched with any potential vehicles $P_i^t \in \mathbf{P}^t$ for more than five succeeding frames, then this potential vehicle component is judged to have disappeared, and its tracker is removed from the tracker set $\mathbf{TP}^t$ in the following frames.

Fig. 10 shows that, after performing the component tracking process, the potential vehicle components entering the detection area, including cars and motorbikes with different amounts of vehicle lights, are tracked accordingly. These potential component tracks are then analyzed and associated by the following motion-based grouping process.

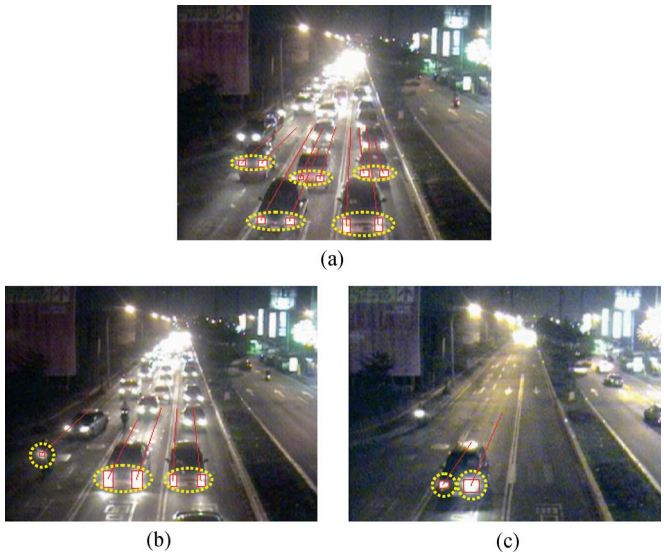


Fig. 10. Results of the component tracking process of potential vehicle components. (a) Tracked potential vehicle components of moving cars with symmetric headlight pairs. (b) Tracked potential vehicle components of moving cars along with motorbikes. (c) Tracked potential vehicle components of moving cars with asymmetric headlights.

Fig. 11. Illustration of motion-based grouping process.

B. Motion-Based Grouping of Vehicle Components

With the tracks of potential vehicle components, the subsequent motion-based grouping process groups potential vehicle components belonging to the same vehicles. For this purpose, potential vehicle components with rigidly similar motions in successive frames are grouped into a single vehicle. Fig. 11 shows this concept.

The pairing tracks of nearby potential vehicle components TP_i^t and TP_j^t are determined to belong to the same vehicle if they continue to move coherently and reveal homogeneous features for a period of time. The coherent motion of vehicle components can be determined by the following coherent motion conditions.

1) They are consistently moving together on the same driving lane for a period of time. First, their spatial motion coherence can be determined by the following *spatial coherence criterion*, including

$$D_h\left(TP_i^{t-\tau}, TP_j^{t-\tau}\right) < \frac{LW\left(TP_i^{t-\tau}\right) + LW\left(TP_j^{t-\tau}\right)}{2}$$
$$D_v\left(TP_i^{t-\tau}, TP_j^{t-\tau}\right) < \frac{\left(\min\left[H\left(TP_i^{t-\tau}\right), H\left(TP_j^{t-\tau}\right)\right]\right)}{2}. \quad (16)$$

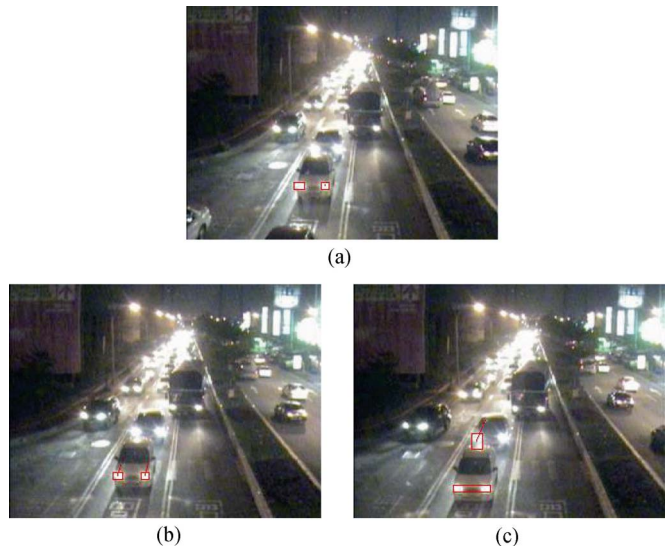


Fig. 12. Example of the motion-based grouping process. (a) Extracted potential vehicle components of separate headlights. (b) Tracked potential vehicle components of separate headlights. (c) Component group track of a potential vehicle obtained by the motion-based grouping process.

Then, the following *lane coherence criterion* is used to determine whether they are moving on the same given lane:

$$f_l\left(C_Y\left(TP_i^{t-\tau}\right)\right) < C_X\left(TP_i^{t-\tau}\right) < f_{l+1}\left(C_Y\left(TP_i^{t-\tau}\right)\right)$$
$$f_l\left(C_Y\left(TP_j^{t-\tau}\right)\right) < C_X\left(TP_j^{t-\tau}\right) < f_{l+1}\left(C_Y\left(TP_j^{t-\tau}\right)\right) \quad (17)$$

where l represents the lth driving lane shown in Fig. 7; $C_X(TP_i^t)$ denotes the horizontal position of the component TP_i^t on the image coordinate [as C_Y defined for (6)] and is defined by $C_X(TP_i^t) = (l_{TP_i^t} + r_{TP_i^t})/2$. Here, $\tau = 0, \ldots, n-1$, n is also determined to be the frames of a duration of 0.5 s (i.e., $0.5 \cdot FPS$ frames), to properly reflect the sufficient sustained time of their coherent motion information in most traffic flow conditions, including free-flowing and congestion cases.

2) They have similar heights for a span of time, i.e.,

$$H\left(TP_S^{t-\tau}\right)/H\left(TP_L^{t-\tau}\right) > T_h \quad (18)$$

where $TP_S^{t-\tau}$ is the one with the smaller height among the two potential vehicle components $TP_i^{t-\tau}$ and $TP_j^{t-\tau}$ at time $t-\tau$, while $TP_L^{t-\tau}$ is the larger one. To reasonably reveal the alignment features of paired vehicle lights, T_h is chosen to be 0.6.

If the tracks TP_i^t and TP_j^t meet the aforementioned coherent motion conditions, they are merged into the same "component group track" of a potential vehicle, denoted by TG_k^t. After performing the motion-based grouping process, a set of K component group tracks, denoted by $\mathbf{TG}^t = \{TG_k^t \,|\, k = 1, \ldots, K\}$, which consist of two or more vehicle components, can be obtained for the subsequent tracking process. Fig. 12 shows the motion-based grouping process on the vehicle component tracks. In this figure, two headlights of a white car are

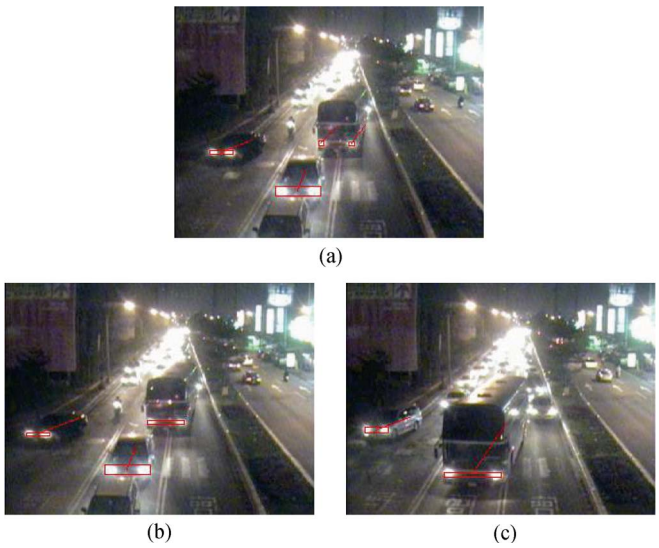


Fig. 13. Example of the tracking process of vehicle component groups. (a) Tracked potential vehicle components of the bus' headlights. (b) Headlights of the bus are merged into a component group by the motion-based grouping process. (c) Tracked potential component group of the bus.

first detected as two potential vehicle components after upon entering the detection area [as in Fig. 12(a)]. Two separate trackers for these two potential vehicle components are then created [as in Fig. 12(b)], and they are accordingly grouped after they continue to move coherently for a period of time [as in Fig. 12(c)]. Notably, in Fig. 12(c), one larger headlight of the following car on the same lane is just detected as a potential vehicle component and tracked. The headlight pair of this car will subsequently be detected, tracked, and grouped as the subsequent car [as shown in Fig. 13(a)].

C. Tracking Process of Vehicle Component Groups

When a potential vehicle represented by a component group is being tracked across the detection area, the segmentation process and the motion-based grouping process can cause some occlusion problems, such as follows: 1) Two vehicles that are simultaneously moving parallel on the same lane are too close to each other (particularly large vehicles, such as buses, vans, or lorries, parallel moving with nearby motorbikes), and they may be occluded for a while because this may not be completely avoided in the *spatial coherence criterion* based on the lane information during the motion-based grouping process, and 2) some large vehicles may have multiple light pairs and therefore may not be immediately merged into single groups during the motion-based grouping process. Therefore, using the potential vehicle tracks of component groups $TG_k^t \in \mathbf{TG}^t$ obtained by the motion-based grouping process, the component group tracking process can update the position, motion, and dimensions of each potential vehicle. This process progressively refines the detection results of potential vehicles using spatial–temporal information in sequential frames. This section describes the tracking process for component groups of potential vehicles, which handles the aforementioned occlusion problems.

First, the possible location of each tracked component group of a potential vehicle in the current frame t will be preliminarily estimated by an adaptive search window based on motion information from the previous frame. To rapidly determine the search window of a tracked vehicle component group, its motion vector is first computed as

$$\begin{aligned}\Delta x_k^{t-1} &= C_X\left(TG_k^{t-1}\right) - C_X\left(TG_k^{t-2}\right)\\ \Delta y_k^{t-1} &= C_Y\left(TG_k^{t-1}\right) - C_Y\left(TG_k^{t-2}\right)\end{aligned} \tag{19}$$

where $C_X(TG_k^t)$ and $C_Y(TG_k^t)$ represent the horizontal and vertical positions of the tracked component group TG_k^t on the image coordinate, respectively, and are defined by $C_X(TG_k^t) = (l_{TG_k^t} + r_{TG_k^t})/2$ and $C_Y(TG_k^t) = (t_{TG_k^t} + b_{TG_k^t})/2$, respectively. A displacement factor (w_1, w_2), which reflects the possible position of the potential vehicle in the current frame, can then be respectively computed as

$$w_1 = 1 + \frac{\Delta x_k^{t-1}}{\left\|\Delta x_k^{t-1}, \Delta y_k^{t-1}\right\|} \qquad w_2 = 1 + \frac{\Delta y_k^{t-1}}{\left\|\Delta x_k^{t-1}, \Delta y_k^{t-1}\right\|} \tag{20}$$

where $\|\Delta x_k^{t-1}, \Delta y_k^{t-1}\|$ indicates the Euclidian distance between TG_k^{t-1} and TG_k^{t-2}. The center of the search window of a tracked potential vehicle in the current frame can then be determined as $(w_1 \times C_X(TG_k^{t-1}), w_2 \times C_Y(TG_k^{t-1}))$, and its width and height can be defined as $1.5 \times W(TG_k^{t-1})$ and $3 \times H(TG_k^{t-1})$, respectively.

Accordingly, the possible positions of tracked potential components TP_i^t, which are matched with a tracked potential component group TG_k^t in the current frame, can be more rapidly and correctly obtained in the search window. A tracked component group TG_k^t appearing in the search window may be in one of four possible states associated with its own component tracks $TP_i^t, \ldots, TP_{i+n}^t$. This potential vehicle tracking process conducts different operations according to the current state of TG_k^t.

1) **Update**: All of the grouped component tracks $TP_i^{t-1}, \ldots, TP_{i+n}^{t-1}$ owned by a tracked component group TG_k^{t-1} in the previous frame still exactly and respectively match a set of vehicle component tracks $TP_{i'}^t, \ldots, TP_{i'+n}^t$ in the current frame within the search window. In other words, they all satisfy the following *group matching condition*:

$$S_o\left(TP_{i'}^t, TG_k^{t-1}\right) > \tau_{mg}. \tag{21}$$

The vehicle tracker then updates the component group TG_k^t of a potential vehicle to include the renewed group of $TP_{i'}^t, \ldots, TP_{i'+n}^t$. Here, the threshold τ_{mg} reflects a reasonable spatial–temporal coherence confirmation for $TP_{i'}^t, \ldots, TP_{i'+n}^t$ to be continuously associated with the same group as TG_k^{t-1}. Like the *tracker matching condition* in (15), for efficient performance under free-flowing traffic scenes with at least a 15 FPS frame-grabbing rate, τ_{mg} should be reasonably firmer than the value of *tracker matching* criterion parameter τ_{mp} in (15). This will ensure that the updated TG_k^t is sufficiently coherent with the associated group of $TP_{i'}^t, \ldots, TP_{i'+n}^t$. Thus, the value of $\tau_{mg} = 0.5$ is experimentally chosen to obtain adequately intact track groups.

2) **Shelter/Absorb**: The grouped component tracks $TP_i^{t-1}, \ldots, TP_{i+n}^{t-1}$ owned by TG_k^{t-1} in the previous frame now have fewer component tracks $TP_{i'}^{t}, \ldots, TP_{i'+m}^{t}$ (where $m < n$) in the current frame within the search window. The *group matching condition* (21) of the component group TG_k^{t-1} with $TP_{i'}^{t}, \ldots, TP_{i'+m}^{t}$ will be respectively checked, and the component tracks that satisfy the matching condition will remain associated with the renewed TG_k^t. The tracks of unexpectedly disappeared or absorbed components missing from TG_k^t are retained in the TG_k^t until they are regarded as disappeared components and removed by the potential vehicle component tracking process.
3) **Extend/Split**: The grouped component tracks $TP_i^{t-1}, \ldots, TP_{i+n}^{t-1}$ owned by TG_k^{t-1} in the previous frame are now extended or split into more component tracks $TP_{i'}^{t}, \ldots, TP_{i'+m}^{t}$ (where $m > n$) in the current frame within the search window. The *group matching condition* (21) of TG_k^{t-1} with $TP_{i'}^{t}, \ldots, TP_{i'+m}^{t}$ will be respectively checked, and the component tracks which coincide with TG_k^{t-1} will remain associated with the renewed TG_k^t. The tracks of newly appearing or split components are not matched with TG_k^{t-1}, and the *motion-based grouping process* (16)–(18) will be applied to these nonmatched component tracks to determine if they have coherent motion property with TG_k^{t-1}. The component tracks that have coherent motion will be assigned to the updated TG_k^t, and the others will be detached as orphan component tracks.
4) **Exit**: Once a tracked potential component group TG_k^t has moved across the boundary of the detection area, the potential vehicle component tracking process determines that all of its component tracks have disappeared.

According to the examples in Fig. 12, Fig. 13 shows the examples of the potential vehicles analyzed by the component group tracking process. In this example, two headlights of a bus are first detected and tracked as two separate potential vehicle components after entering the detection area [as in Fig. 13(a)]. They are then merged into a component group by the motion-based grouping process [as in Fig. 13(b)], and its component group is accordingly tracked as a potential vehicle [as in Fig. 13(c)]. After the potential vehicles are tracked for a certain time, the following verification and classification process is performed on these tracked potential vehicles to identify the actual vehicles and their associated types.

D. Vehicle Identification and Classification From the Tracking Process

During the tracking process of the potential vehicles, a rule-based vehicle verification and classification process is applied to each of the potential components and component groups of potential vehicles tracked for more than ten frames. This process determines whether it comprises a car or a motorcycle, and filters out other on-road nonvehicle-illuminant objects.

1) *Car identification*. First, to identify moving cars in a frame, we can reasonably assume that a group of lighting components has a higher possibility of being a car. Therefore, a tracked component group TG_k^t which has been consistently tracked by the component group tracking process for a span of more than ten frames after being created by the motion-based grouping process can be nominated as a moving car candidate. Accordingly, if TG_k^t contains a set of actual vehicle lights that reveal an actual car, then TG_k^t must satisfy the following discriminating rules of statistical features.
   a) Since a moving car can be approximately modeled as a rectangular patch, the enclosing bounding box of a potential car should form a horizontal rectangular shape, i.e., the size-ratio feature of the enclosing bounding box of TG_k^t must satisfy the following condition:
   $$\tau_{r1} \le W\left(TG_k^t\right)/H\left(TG_k^t\right) \le \tau_{r2} \tag{22}$$
   where the threshold values τ_{r1} and τ_{r2} of the size-ratio condition are selected as 2.0 and 8.0, respectively. These values are based on our analysis of the typical rectangular-shaped front and rear appearances of paired headlights and taillights on most cars.
   b) The number of lighting components of TG_k^t should also be symmetrical and well aligned, and the number of these components should be in reasonable proportion to the size of the size-ratio feature of its enclosing bounding box. Thus, the following alignment condition should be satisfied:
   $$\tau_{a1}\left(\frac{W\left(TG_k^t\right)}{H\left(TG_k^t\right)}\right) \le \left|TG_k^t\right| \le \tau_{a2}\left(\frac{W\left(TG_k^t\right)}{H\left(TG_k^t\right)}\right) \tag{23}$$
   where τ_{a1} serves as a criterion to reveal the thinner and slenderer rectangularly aligned characteristics for detecting taillight sets in the rear view of moving cars, while τ_{a2} reflects the thicker rectangularly aligned characteristics for detecting headlight sets in the front view of moving cars. According to our analysis of typical visual alignment characteristics of most moving cars appearing in nighttime traffic scenes, the thresholds τ_{a1} and τ_{a2} are determined to be 0.4 and 2.0, respectively, to suitably identify moving cars in both front and rear views.
   c) Moreover, since a moving car occupies a considerable area on a driving lane, the width of TG_k^t of a potential car should occupy a reasonable ratio with respect to the lane width, i.e., the following width condition should be met:
   $$\tau_{w1} \cdot LW\left(TG_k^t\right) \le W\left(TG_k^t\right) \le \tau_{w2} \cdot LW\left(TG_k^t\right) \tag{24}$$
   where $LW(TG_k^t)$ is the approximate lane width associated with TG_k^t at its position on the image coordinate and is computed by (6), and the thresholds τ_{w1} and τ_{w2} are determined as 0.5 and 0.9, respectively, to reveal the width property of the most types of different sized cars, such as sedans, buses, and trucks, appearing in traffic scenes with typical lane widths.
2) *Motorbike identification*. To identify motorbikes, we can adopt the fact that a motorbike usually appears as a

single and nearly square-shaped or vertical rectangular-shaped lighting component in nighttime traffic scenes. Thus, a single tracked component TP_i^t which has not been associated to any component groups and been consistently and alone tracked by the vehicle component tracking process for a significant span of more than 1 s, i.e., $1.0 \cdot FPS$ frames, can be identified as a moving motorbike candidate. Therefore, if a single tracked component is actually a motorbike, then the size-ratio feature of its enclosing bounding box should reflect a square or vertical rectangular shape and should satisfy the following discriminating rule:

$$\tau_{m1} \leq W\left(TP_i^t\right)/H\left(TP_i^t\right) \leq \tau_{m2} \tag{25}$$

where the threshold values τ_{m1} and τ_{m2} on the size-ratio condition are selected as 0.6 and 1.2, respectively, to suitably identify the shape appearance characteristic of the motorbikes, which are obviously different from those of the cars.

The aforementioned discriminating rules can be obtained by analyzing many experimental videos of real nighttime traffic environments, in which vehicle lights appear in different shapes and sizes, and move in different directions at different distances. The threshold values utilized for these discriminating rules were determined to yield good performance in most general cases of nighttime traffic scenes.

3) *Motorbike identification*. A tracked component group or single potential component of a potential vehicle will be identified and classified as an actual car or a motorbike based on the aforementioned vehicle classification rules. When a classified vehicle leaves the detection area, the count of its associated vehicle type is then incremented and recorded to update the traffic flow information. Thus, each detected vehicle is guaranteed to be counted once, and the redundant counting of vehicles can be efficiently avoided.

In summary, Fig. 14 shows the overall process flow of the proposed nighttime traffic surveillance system.

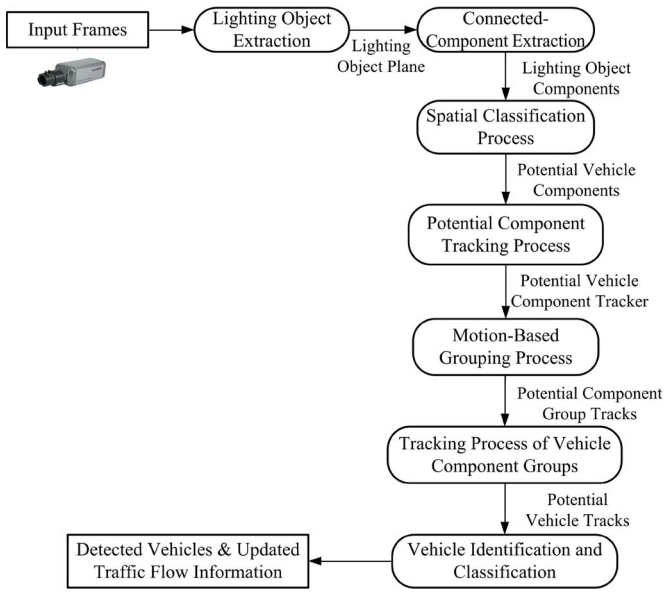


Fig. 14. Block diagram of the proposed nighttime traffic surveillance system.

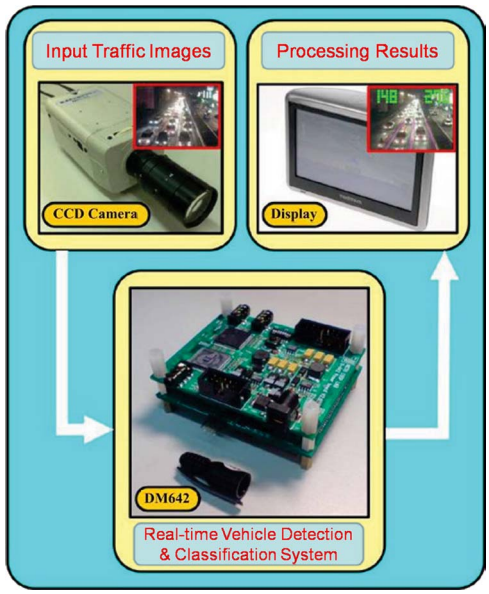


Fig. 15. Illustration of the proposed real-time system.

V. Experimental Results

This section describes the implementation of the proposed vehicle detection, tracking, and classification system on a DSP-based real-time system. We conducted various representative real-time experiments to evaluate the vehicle detection and classification performance obtained by the proposed system. The proposed real-time vision system was implemented on a TI DM642 DSP-based embedded platform, operated at 600 MHz with 32-MB DRAM, and set up on elevated platforms near highways and urban roads.

Initially, the detection area for each traffic scene was first determined using a lane detection process derived from our previous study [28]. The detection area was located along the midline of the traffic-scene image, bounded by the most left and right lane boundaries (as Fig. 5 shows), and divided into driving lanes (as Fig. 7 shows). To make the proposed system operate well, the CCD camera should be set up on an elevated platform with a sufficient height to capture an appropriate region for covering all the driving lanes to be monitored, and the view angles of the CCD camera should be adjusted to be oriented to the monitored region for suitably obtaining reliable and clear features of vehicle lights. Fig. 15 shows the processing architecture of the proposed real-time system. The frame rate of this vision system is 30 true-color FPS, and each frame in the grabbed image sequences measures 320 by 240 pixels. The computation required to process one input frame depends on the traffic-scene complexity. Most of the computation time is spent on the connected-component analysis and the spatial clustering process of lighting objects. For an input video sequence with 320×240 pixels per frame, the proposed real-time system takes an average of 26.3 ms to process each frame on the 600-MHz TI-DM642 DSP-based embedded platform. This

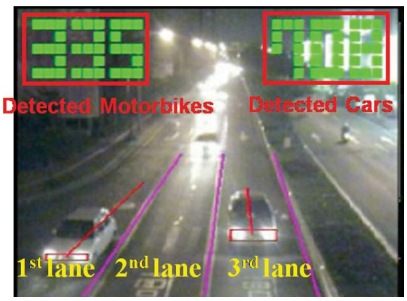


Fig. 16. Illustration of the vehicle detection and classification results.

minimal computation cost ensures that the proposed system can effectively satisfy the demand of real-time processing at more than 30 FPS.

The proposed system was tested on several videos of real nighttime highway and urban traffic scenes in various traffic conditions in Taiwan. Fig. 16 shows that the proposed system counts the numbers of detected cars and motorbikes appearing in each driving lane of the detection area and displays the number of detected cars on the top right of the screen and the amount of detected motorbikes on the top left.

For the quantitative evaluation of vehicle detection performance, this study adopts the Jaccard coefficient [40], which is commonly used for evaluating performance in information retrieval. This measure is defined as

$$J = \frac{T_p}{T_p + F_p + F_n} \tag{26}$$

where T_p (true positives) represents the number of correctly detected vehicles, F_p (false positives) represents the number of falsely detected vehicles, and F_n (false negatives) is the number of missed vehicles. We determined the Jaccard coefficient J for the vehicle detection results of each frame of the traffic video sequences by manually counting the number of correctly detected vehicles, falsely detected vehicles, and missed detections of vehicles in each frame. The average value of the Jaccard coefficients J was then obtained from all frames of the video sequences by

$$\bar{J} = \sum_{N} J/N \tag{27}$$

where N is the total number of video frames. Here, the ground truth of detected vehicles was obtained by manual counting.

Figs. 17–20 and Tables I–IV exhibit the most representative experimental samples of traffic scenes with different traffic flows and different environmental illumination conditions on performance evaluation. First, Fig. 17 shows a nighttime urban traffic scene with a dark environmental illuminated condition and low traffic flow. As Fig. 17 shows, although nonvehicle illuminating objects and reflected beams on the ground coexist with the vehicle in this scene, the proposed system correctly detected and tracked nearly all moving cars and motorbikes on a free-flowing urban road by locating, grouping, and classifying their vehicle lights. However, a few detection errors occurred when some cars with broken (single) headlights were misclassified as motorbikes. Table I depicts the quantitative results of

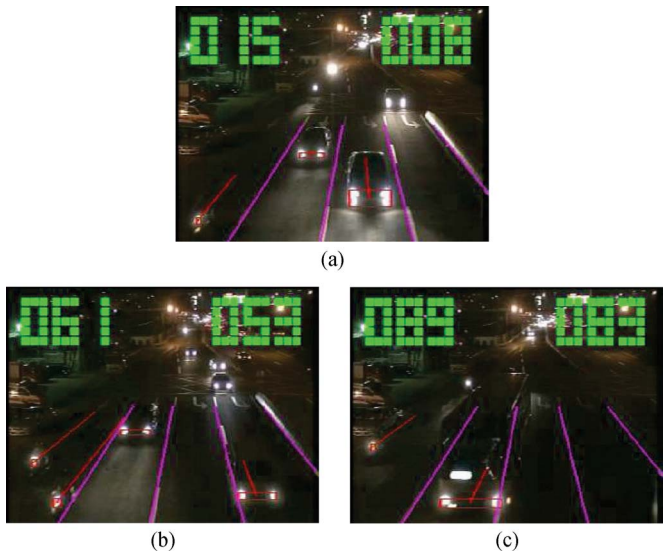


Fig. 17. Results of different car/motorbike detection and classification results for a nighttime urban traffic scene with a dark environmental illuminated condition.

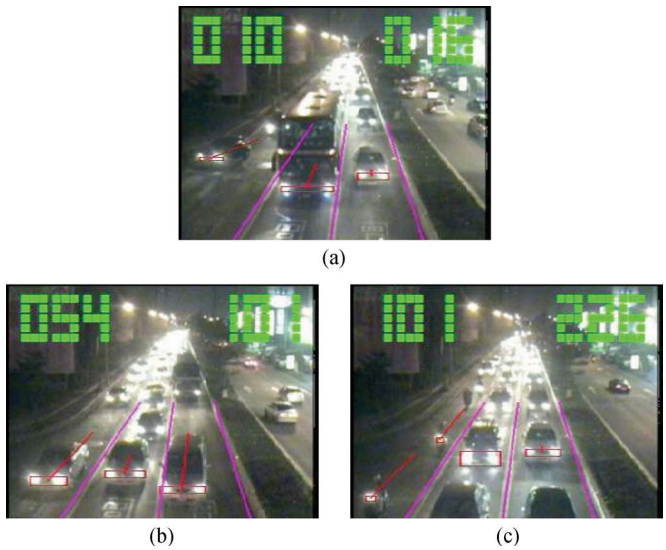


Fig. 18. Snapshots of different car/motorbike detection and classification results of a busy urban road under bright illumination condition.

the proposed approach for vehicle detection and tracking on this urban road.

Fig. 18 shows another sample of a more complicated traffic scene from a nighttime urban road at rush hour under a bright environmental illumination condition. Due to traffic signal changes, the vehicles, including large and small cars, and motorbikes, stop and move intermittently. As Fig. 18 shows, most of these cars and motorbikes are correctly detected, tracked, and classified, although many nonvehicle illuminating objects, such as street lamps, reflected beams, and road reflectors on the ground, appear very close to the lights of the detected vehicles. Moreover, as shown in Fig. 18(b) and (c), most vehicles driving very close to nearby lanes are also successfully discriminated and detected. Table II shows the experimental data of the proposed approach on vehicle detection and tracking for the traffic scene in Fig. 18. Notably, in this traffic scene, many cars

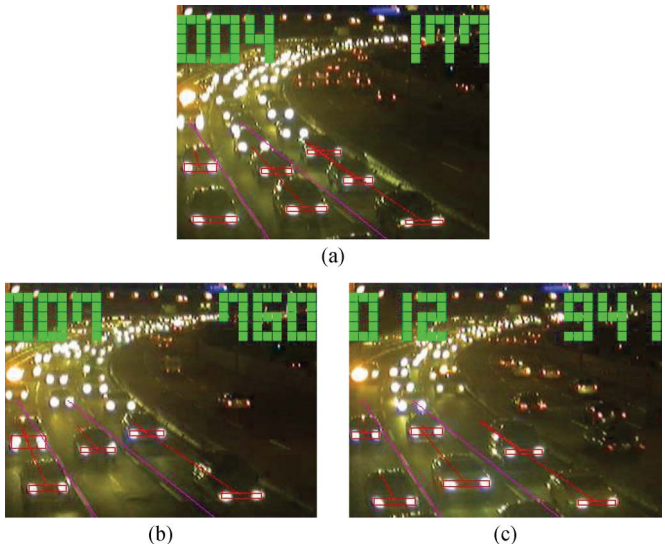


Fig. 19. Results of vehicle detection and classification for a congested highway traffic scene under a light environmental illumination condition.

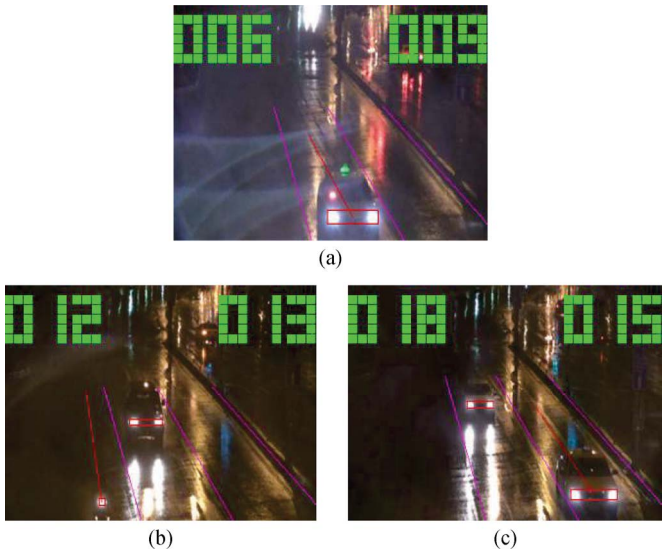


Fig. 20. Snapshots of different car/motorbike detection and classification results for a nighttime urban traffic scene in a rainy day.

and motorbikes are mixed and closely comoved (particularly in the first lane), and many cars and motorbikes intending to turn right also drive in the first lane. Although this lane has more complicated traffic flow patterns and vehicle features than the other two lanes, and causes a few more misdetections than the other two lanes, the proposed system still appropriately detected and classified most of the moving cars and motorbikes.

Fig. 19 shows another experimental scene of a congested nighttime highway at rush hour under a light environmental illumination condition. These images were obtained by a closed-circuit television (CCTV) camera. Since motorbikes are not allowed to drive on highways in Taiwan, only cars appeared in this highway traffic scene. This figure shows that even though multiple vehicles are stopped or moving slowly close to each other in this congested traffic scene [particularly in Fig. 19(a) and (c)], the proposed method still successfully detects and tracks almost all vehicles. Table III shows the quantitative

TABLE I
EXPERIMENTAL DATA OF THE PROPOSED APPROACH ON TEST SEQUENCE 1 FOR THE URBAN ROAD SCENE IN FIG. 17

Lane	Detected Vehicles	Actual Vehicles
Lane 1	131	137
Lane 2	111	113
Lane 3	67	69
Lane 4	36	36
Total No. Cars	163	165
Total No. Motorbikes	184	190
Detection Score *J* of Cars	98.79%	
Detection Score *J* of Motorbikes	96.84%	
Time span of the video	20 minutes	

TABLE II
EXPERIMENTAL DATA OF THE PROPOSED APPROACH FOR TEST SEQUENCE 2 OF THE URBAN ROAD IN FIG. 18

Lane	Detected Vehicles	Actual Vehicles
Lane 1	921	969
Lane 2	292	300
Lane 3	228	233
Total No. Cars	887	909
Total No. Motorbikes	584	593
Detection Score *J* of Cars	97.58%	
Detection Score *J* of Motorbikes	98.48%	
Time span of the video	50 minutes	

TABLE III
EXPERIMENTAL DATA OF THE PROPOSED APPROACH FOR TEST SEQUENCE 3 OF THE HIGHWAY SCENE IN FIG. 19

Lane	Detected Vehicles	Actual Vehicles
Lane 1	1392	1428
Lane 2	1527	1535
Lane 3	1495	1536
Total No. Cars	4397	4499
Detection Rate *J* of Cars	97.73%	
Time span of the video	50 minutes	

TABLE IV
EXPERIMENTAL DATA OF THE PROPOSED APPROACH FOR TEST SEQUENCE 4 OF THE RAINY URBAN ROAD IN FIG. 20

Lane	Detected Vehicles	Actual Vehicles
Lane 1	4	4
Lane 2	7	7
Lane 3	6	6
Total No. Cars	17	17
Total No. Motorbikes	22	20
Detection Score *J* of Cars	100%	
Detection Score *J* of Motorbikes	90.91%	
Time span of the video	15 minutes	

results of the proposed approach for vehicle detection on a nighttime highway. Due to the unsatisfactory view angle of the CCTV camera, the first lane is partially occluded. Thus, the vehicle light sets of some few detected cars may be occluded and misclassified as single-light motorbikes. However, this does not significantly influence the determination of typical traffic

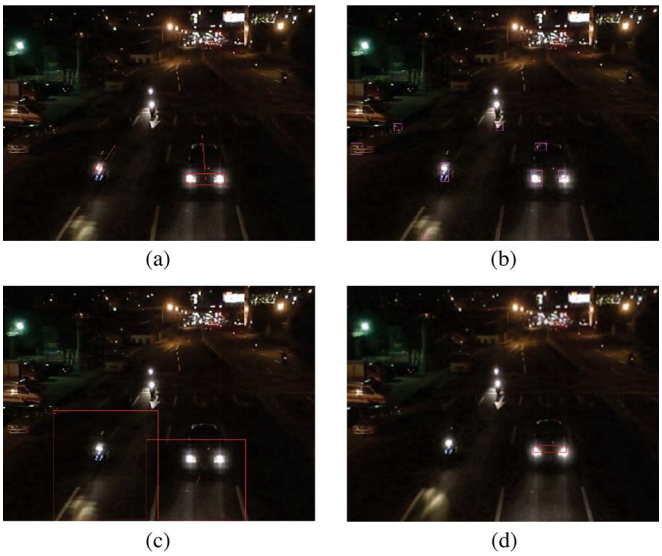


Fig. 21. Comparative results of vehicle detection on test sequence 1 for an urban road scene. (a) Proposed method. (b) Contrast-based method. (c) Background-subtraction-based method. (d) Pairing-light-based method of Wang *et al.*

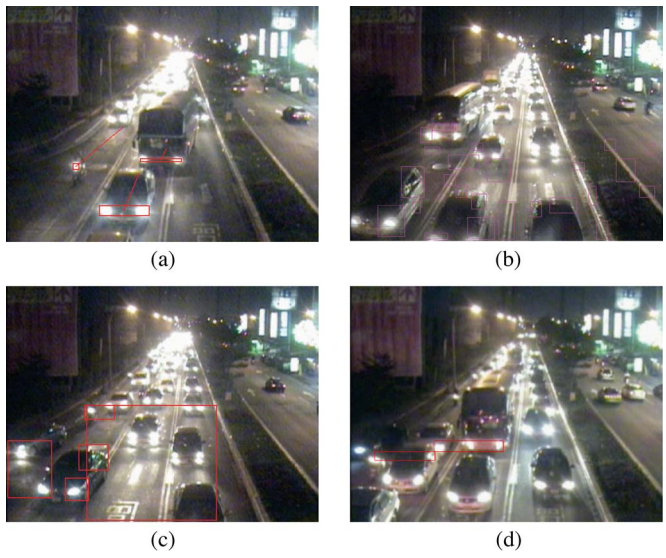


Fig. 22. Comparative results of vehicle detection on test sequence 2 for an urban road scene. (a) Proposed method. (b) Contrast-based method. (c) Background-subtraction-based method. (d) Pairing-light-based method of Wang *et al.*

flow parameters, including congestion, throughput, and queue length.

Fig. 20 shows the alternative experimental samples of vehicle detection on a nighttime urban road in a rainy day. In the rainy condition, the most significant difficulty is the numerous scattering glares caused by reflected lights on the ground. As Fig. 20 shows, although there are many spurious lighting objects appeared close to actual vehicle lights in the traffic scene, the proposed system is able to avoid most of the reflected glares and appropriately detect, classify, and track most cars and motorbikes. Table IV depicts the results of vehicle detection in the rainy traffic scene. Although few motorbikes with smaller single headlights being misdetected due to large reflected glares, the vehicles in each lane are almost properly detected because glaring effects can be efficiently suppressed by the proposed adaptive lighting object segmentation, spatial classification, and motion-based tracking methods.

The aforementioned experimental traffic video sequences in various environmental illumination conditions and different road environments were also employed for a comparative performance evaluation of vehicle detection. The following part evaluates the performance of the proposed system and compares it to the contrast-based method of Huang *et al.* [17], the background-subtraction-based method of Wu *et al.* [8], and the pairing-light-based method of Wang *et al.* [30]. Figs. 21–23 show the representative comparative results of nighttime vehicle detection for the aforementioned three test sequences produced by the proposed approach, the contrast-based method, the background-subtraction-based method, and the pairing-light-based method.

Fig. 21(a), Fig. 22(a), and Fig. 23(a) show that, although the vehicles of interest move and stop at various speeds and are interfered with by many nonvehicle illuminating objects, the proposed system successfully detects cars and motorbikes with different moving speeds despite the various difficulties associated with traffic-scene complexity. By comparison, Fig. 21(b),

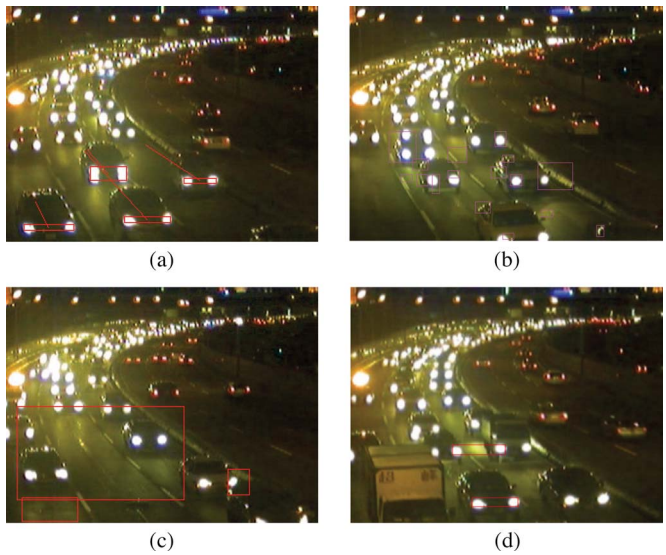


Fig. 23. Comparative results of vehicle detection on test sequence 3 for a congested highway traffic scene. (a) Proposed method. (b) Contrast-based method. (c) Background-subtraction-based method. (d) Pairing-light-based method of Wang *et al.*

Fig. 22(b), and Fig. 23(b) show that the contrast-based method does not perform well in detecting vehicles under some complicated nighttime traffic scenes, and several detected vehicles are broken and blurred by spurious detections of nonvehicle illuminating objects. This is because the block-based detection scheme of the contrast-based method is often unable to accurately locate the object regions of nighttime vehicles. The detection results in Fig. 21(c), Fig. 22(c), and Fig. 23(c) show that the vehicles of interest are seriously smeared with large spurious nonvehicle and background regions. This is because sharp changes in illumination intensities caused by fluctuating vehicle lights frustrate background-based models. Fig. 21(d), Fig. 22(d), and Fig. 23(d) show that, while the pairing-light-based method can efficiently detect vehicles under free-flowing

TABLE V
COMPARATIVE DATA ON THE VEHICLE DETECTION PERFORMANCE OF THE PROPOSED SYSTEM, THE CONTRAST-BASED METHOD, THE BACKGROUND-SUBTRACTION-BASED METHOD, AND THE PAIRING-LIGHT-BASED METHOD OF WANG *et al.*

Detection Rate *J*(%)	The proposed	Contrast-based Method [17]	Background Subtraction Method [8]	Pairing-light method [30]
Test sequence 1 (Urban road 1, 20 min.)	96.3%	76.9%	64.5%	43.93%
Test sequence 2 (Urban road 2, 50 min)	97.9%	43.2%	less than 30%	less than 30%
Test sequence 3 (Highway CCTV, 50 min)	96.9%	56.7%	less than 30%	59.88%

conditions, it does not perform well on detecting vehicles under traffic congestion conditions, and many occlusions and misdetections occur when they are closely moved with large amount. Moreover, most motorbikes are missed because they have only single headlights.

Table V depicts the quantitative evaluation results for the proposed approach, contrast-based method of Huang *et al.* [17], background subtraction method of Wu *et al.* [8], and pairing-light-based method of Wang *et al.* [30]. To fairly conduct a comparative evaluation of computational costs, all these methods were implemented on a 2.4-GHz Pentium-IV personal computer platform with a frame resolution of 320 $\times$ 240 true-color pixels per frame. The average computing times of the proposed approach, the method of Huang *et al.* [17], the method of Wu *et al.* [8], and the method of Wang *et al.* [30] are 12, 41, 18, and 10 ms, respectively. Therefore, according to Table V and the computational timings, although the proposed system requires a little more computing time than the pairing-light-based method of Wang *et al.*, the proposed system can provide better vehicle detection performance for nighttime traffic surveillance than other existing methods. Accordingly, the experimental and comparative results of numerous different traffic scenes can demonstrate that the proposed system can provide fast, effective, and robust vehicle detection and identification performance on different nighttime traffic environments, including various illumination conditions, traffic flows, and road environments for nighttime traffic surveillance.

VI. CONCLUSION

This paper has proposed an effective nighttime vehicle detection and tracking system for identifying and classifying moving vehicles for traffic surveillance. The proposed approach uses an efficient and fast bright-object segmentation process based on automatic multilevel histogram thresholding to extract bright objects from nighttime traffic image sequences. This technique is robust and adaptable when dealing with varying lighting conditions at night. A spatial analysis and clustering procedure is applied to group lighting objects into groups of vehicle lights for potential moving cars and motorbikes. Next, a new effective feature-based vehicle tracking and identification process analyzes the spatial and temporal information of these potential vehicle light groups from consecutive frames, refines the detection results, and corrects for errors and occlusions. Actual vehicles and their types can then be efficiently detected and classified from these tracked potential vehicles to obtain traffic flow information from traffic monitoring images. The proposed nighttime vehicle detection and classification approaches were also implemented on a TI DM642 DSP-based real-time vision system and tested with real highway and urban nighttime traffic scenes. Experimental results and comparison with existing methods have shown that the proposed system is effective and offers advantages for vehicle detection and classification for traffic surveillance in various nighttime environments. For further studies, the vehicle type classification function can be further improved and extended by integrating some sophisticated machine learning techniques such as support vector machine classifiers on multiple features, including vehicle lights and vehicle bodies, to further enhance the classification capability on more detailed vehicle types, such as sedans, buses, trucks, lorries, and light and heavy motorbikes.

REFERENCES

[1] V. Kastrinaki, M. Zervakis, and K. Kalaitzakis, "A survey of video processing techniques for traffic applications," *Image Vis. Comput.*, vol. 21, no. 4, pp. 359–381, Apr. 2003.

[2] W.-L. Hsu, H.-Y. , M. 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.

[3] 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. 2000.

[4] M.-C. Huang and S.-H. Yen, "A real-time and color-based computer vision for traffic monitoring system," in *Proc. IEEE Int. Conf. Multimed. Expo*, May 2004, pp. 2119–2122.

[5] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 1999, pp. 246–252.

[6] J. Kong, Y. Zheng, Y. Lu, and B. Zhang, "A novel background extraction and updating algorithm for vehicle detection and tracking," in *Proc. IEEE Int. Conf. Fuzz. Syst. Knowl. Discovery*, 2007, pp. 464–468.

[7] 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.

[8] B.-F. Wu, S.-P. Lin, and Y.-H. Chen, "A real-time multiple-vehicle detection and tracking system with prior occlusion detection and resolution," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol.*, Dec. 2005, pp. 311–316.

[9] J. Zhou, D. Gao, and D. Zhang, "Moving vehicle detection for automatic traffic monitoring," *IEEE Trans. Veh. Technol.*, vol. 56, no. 1, pp. 51–59, Jan. 2007.

[10] W. F. Gardner and D. T. Lawton, "Interactive model-based vehicle tracking," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 18, no. 11, pp. 1115–1121, Nov. 1996.

[11] 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.

[12] D. Koller, J. Weber, T. Huang, J. Malik, G. Ogasawara, B. Rao, and S. Russell, "Towards robust automatic traffic scene analysis in real-time," in *Proc. Int. Conf. Pattern Recog.*, 1994, vol. 1, pp. 126–131.

[13] N. Peterfreund, "Robust tracking of position and velocity with Kalman snakes," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 21, no. 6, pp. 564–569, Jun. 1999.

[14] S.-T. Tseng and K.-T. Song, "Real-time image tracking for traffic monitoring," in *Proc. IEEE 5th Int. Conf. Intell. Transp. Syst.*, 2002, pp. 10–14.

[15] L.-W. Tsai, J.-W. Hsieh, and K.-C. Fan, "Vehicle detection using normalized color and edge map," *IEEE Trans. Image Process.*, vol. 16, no. 3, pp. 850–864, Mar. 2007.

[16] D. Beymer, P. McLauchlan, B. Coifman, and J. Malik, "A real-time computer vision system for measuring traffic parameters," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, Jun. 1997, pp. 495–510.

[17] K. Huang, L. Wang, T. Tan, and S. Maybank, "A real-time object detecting and tracking system for outdoor night surveillance," *Pattern Recognit.*, vol. 41, no. 1, pp. 432–444, Jan. 2008.
[18] M. Dufaut, R. Taktak, and R. Husson, "Vehicle detection at night using image processing and pattern recognition," in *Proc. IEEE Int. Conf. Image Process.*, 1994, vol. 2, pp. 296–300.
[19] R. Cucchiara and M. Piccardi, "Vehicle detection under day and night illumination," in *Proc. 3rd Int. ICSC Symp. Intell. Ind. Autom. Soft Comput.*, 1999, pp. 1–4.
[20] W. Wang, C. Shen, J. Zhang, and S. Paisitkriangkrai, "A two-layer nighttime vehicle detector," in *Proc. Int. Conf. DICTA*, 2009, pp. 619–626.
[21] K. Robert, "Night-time traffic surveillance a robust framework for multi-vehicle detection, classification and tracking," in *Proc. IEEE Conf. AVSS*, 2009, pp. 1–6.
[22] S. Tsugawa, "Vision-based vehicles in Japan: Machine vision systems and driving control systems," *IEEE Trans. Ind. Electron.*, vol. 41, no. 4, pp. 398–405, Aug. 1994.
[23] Z. Sun, G. Bebis, and R. Miller, "On-road vehicle detection: A review," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 5, pp. 694–711, May 2006.
[24] Broggi, M. Bertozzi, A. Fascioli, and G. Conte, *Automatic Vehicle Guidance: The Experience of the ARGO Autonomous Vehicle*. Singapore: World Scientific, 1999.
[25] H. Mori, M. Charkari, and T. Matsushita, "On-line vehicle and pedestrian detections based on sign pattern," *IEEE Trans. Ind. Electron.*, vol. 41, no. 4, pp. 384–391, Aug. 1994.
[26] A. Broggi, M. Cellario, P. Lombardi, and M. Porta, "An evolutionary approach to visual sensing for vehicle navigation," *IEEE Trans. Ind. Electron.*, vol. 50, no. 1, pp. 18–29, Feb. 2003.
[27] T. Bücher, C. Curio, J. Edelbrunner, C. Igel, D. Kastrup, I. Leefken, G. Lorenz, A. Steinhage, and W. Seelen, "Image processing and behavior planning for intelligent vehicles," *IEEE Trans. Ind. Electron.*, vol. 50, no. 1, pp. 62–75, Feb. 2003.
[28] B.-F. Wu, C.-T. Lin, and Y.-L. Chen, "Dynamic calibration and occlusion handling algorithms for lane tracking," *IEEE Trans. Ind. Electron.*, vol. 56, no. 5, pp. 1757–1773, Apr. 2009.
[29] M. Y. Chern and P. C. Hou, "The lane recognition and vehicle detection at night for a camera assisted car on highway," in *Proc. IEEE Int. Conf. Robot. Autom.*, 2003, vol. 2, pp. 2110–2115.
[30] C.-C. Wang, S.-S. Huang, and L.-C. Fu, "Driver assistance system for lane detection and vehicle recognition with night vision," in *Proc. IEEE/RSJ Int. Conf. Intell. Robot. Syst.*, 2005, pp. 3530–3535.
[31] Y.-L. Chen, Y.-H. Chen, C.-J. Chen, and B.-F. Wu, "Nighttime vehicle detection for driver assistance and autonomous vehicles," in *Proc. 18th IAPR ICPR*, 2006, vol. 1, pp. 687–690.
[32] Y.-M. Chan, S.-S. Huang, L.-C. Fu, and P.-Y. Hsiao, "Vehicle detection under various lighting conditions by incorporating particle filter," in *Proc. IEEE Intell. Transp. Syst. Conf.*, 2007, pp. 534–539.
[33] A. M. López, J. Hilgenstock, A. Busse, R. Baldrich, F. Lumbreras, and J. Serrat, *Nighttime Vehicle Detection for Intelligent Headlight Control*, vol. 5259. Berlin, Germany: Springer-Verlag, 2008, ser. Lecture Notes in Computer Science, pp. 113–124.
[34] P. G. Michalopoulos, "Vehicle detection video through image processing: The autoscope system," *IEEE Trans. Veh. Technol.*, vol. 40, no. 1, pp. 21–29, Feb. 1991.
[35] 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.
[36] 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 Trans. Intell. Transp. Syst.*, vol. 1, no. 2, pp. 86–97, Jun. 2000.
[37] Y.-L. Chen, B.-F. Wu, and C.-J. Fan, "Real-time vision-based multiple vehicle detection and tracking for nighttime traffic surveillance," in *Proc. IEEE Int. Conf. SMC*, San Antonio, Texas, 2009, pp. 3452–3458.
[38] B.-F. Wu, Y.-L. Chen, and C.-C. Chiu, "A discriminant analysis based recursive automatic thresholding approach for image segmentation," *IEICE Trans. Inf. Syst.*, vol. E88-D, no. 7, pp. 1716–1723, Jul. 2005.
[39] K. Suzuki, I. Horiba, and N. Sugie, "Linear-time connected-component labeling based on sequential local operations," *Comput. Vis. Image Understand.*, vol. 89, no. 1, pp. 1–23, Jan. 2003.
[40] P. Sneath and R. Sokal, *Numerical Taxonomy. The Principle and Practice of Numerical Classification*. New York: Freeman, 1973.

Yen-Lin Chen (S'02–M'07) was born in Kaohsiung, Taiwan, in 1978. He received the B.S. and Ph.D. degrees in electrical and control engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2000 and 2006, respectively.

From February 2007 to July 2009, he was an Assistant Professor with the Department of Computer Science and Information Engineering, Asia University, Taichung, Taiwan. Since August 2009, he has been an Assistant Professor with the Department of Computer Science and Information Engineering, National Taipei University of Technology, Taipei, Taiwan. His research interests include image and video processing, pattern recognition, embedded systems, document image analysis, and intelligent transportation systems.

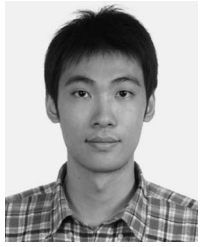
Dr. Chen is a member of the International Association for Pattern Recognition and the Institute of Electronics, Information and Communication Engineers. In 2003, he received the Dragon Golden Paper Award sponsored by the Acer Foundation and the Silver Award of Technology Innovation Competition sponsored by AdvanTech.

Bing-Fei Wu (S'89–M'92–SM'02) was born in Taipei, Taiwan, in 1959. He received the B.S. and M.S. degrees in control engineering from National Chiao Tung University (NCTU), Hsinchu, Taiwan, in 1981 and 1983, respectively, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, in 1992.

Since 1992, he has been with the institute of electrical and control engineering, NCTU, where he is currently a Professor. He has been involved in the research of intelligent transportation systems for many years and leads a research team to develop the first Taiwan smart car, called TAIWAN *i*TS-1, with autonomous driving and active safety system. His current research interests include data compression, vision-based vehicle driving safety, intelligent vehicle control, multimedia signal analysis, embedded systems, and chip design.

Hao-Yu Huang was born in Kaohsiung, Taiwan, in 1983. He received the B.S. degree in electrical and control engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2004, where he is currently working toward the Ph.D. degree in electrical and control engineering.

His research interests include audio and video coding, and embedded system design.



Chung-Jui Fan was born in Taipei, Taiwan, in 1985. He received the B.S. degree in mechanical and electromechanical engineering from National Sun Yat-Sen University, Kaohsiung, Taiwan, in 2006, and in electrical and control engineering from National Chiao Tung University, Hsinchu, Taiwan, in 2008.

He is now with the Acer Unipac Optronics (AUO) Corp., Taiwan, as an Engineer. His research interests include intelligent transportation systems and vehicle detection.