

Image-Based Traffic Monitoring With Shadow Suppression

Color analysis algorithms seem to allow more accurate estimation of traffic conditions when traffic control camera systems become confused by shadows.

By KAI-TAI SONG, Associate Member IEEE, AND JEN-CHAO TAI

ABSTRACT | For a vision-based traffic monitoring and enforcement system, shadows of moving objects often cause serious errors in image analysis due to misclassification of shadows and moving vehicles. An effective shadow suppression method is thus required to improve the accuracy of image analysis and this paper proposes a novel color-space ratio model for detecting shadow pixels in traffic imagery. The proposed approach does not require many image sequences for constructing the model. Instead the model can be easily built up using a shadow region in a single image frame. To increase the accuracy of shadow detection, we design two types of spatial analysis to verify actual shadow pixels. Comparative results show that the proposed method works better than several well-known methods. The proposed methods have been applied to an image-based traffic monitoring system for detecting shadow pixels in traffic imagery. The experimental results not only validate the feasibility of the proposed algorithm but also successfully estimate traffic parameters such as traffic flows, traffic densities, vehicle turn ratios and vehicle speeds, all with satisfactory accuracy.

KEYWORDS | Image processing; shadow suppression; traffic monitoring system; vehicle detection; visual tracking

I. INTRODUCTION

Vision-based traffic monitoring has become an active research area in recent years to support the development of intelligent transportation systems. Under the framework of advanced transportation management and information systems, services such as traveler information, route guidance, traffic control, congestion monitoring, incident detection, and system evaluation across complex transportation networks have been extensively studied to enhance traveling safety and efficiency [1]–[3]. For these advanced applications, various types of traffic information need to be collected on-line and distributed in real time. Automatic traffic monitoring and enforcement have also become increasingly important for road use and management.

Various sensor systems have been applied to estimate traffic parameters. Currently, magnetic loop detectors are the most widely used sensors, but they are difficult to install and maintain. It is widely recognized that image-based systems are flexible and versatile for advanced traffic monitoring and enforcement applications. Compared with loop detectors, image-based traffic monitoring systems (ITMS) provide more flexible solutions for estimating traffic parameters [4]–[8]. In ITMS, it is important to segment and track various moving vehicles from image sequences. Thanks to image tracking techniques, the pixel coordinates of each moving vehicle can be recorded in each time frame. Using calibrated camera parameters, such as focal length, tilt angle, installation height of the camera, and vanishing point of parallel lines in the scene, the pixel coordinates of moving vehicles can be transformed into their world coordinates [9]. Accordingly, useful traffic parameters, including vehicle speeds, vehicle travel directions, traffic flows, etc., can be obtained from image measurement. These quantitative traffic parameters are useful for traffic control and management [6].

Many approaches to vehicle segmentation have been studied for image-based traffic monitoring [6]–[8].

Manuscript received December 23, 2005; revised June 15, 2006. This work was supported in part by the National Science Council, Taiwan, R.O.C., under Grant NSC 94-2218-E-009-008 and by the Ministry of Education, Taiwan, R.O.C., under Grant EX-91-E-FA06-4-4.

K.-T. Song is with the Department of Electrical and Control Engineering, National Chiao Tung University, Hsinchu 300, Taiwan, R.O.C.
(e-mail: ktsong@mail.nctu.edu.tw).

J.-C. Tai is with the Department of Mechanical Engineering, Minghsin University of Science and Technology, Shinfeng, Hsinchu 304, Taiwan, R.O.C.
(e-mail: tjc.ec88g@nctu.edu.tw; tjc@must.edu.tw).

Digital Object Identifier: 10.1109/JPROC.2006.888403

Background removal is a powerful tool for extracting foreground objects from image frames. However, shadows of moving objects often cause serious errors in image analysis due to the misclassification of shadows and moving objects. In traffic imagery, shadows attached to their respective moving vehicles introduce distortions and cause problems in image segmentation, so cast-shadow suppression is an essential prerequisite for image-based traffic monitoring. In order to increase the accuracy of image analysis, it is desirable to develop a method to separate moving vehicles from their shadows.

Many algorithms have been proposed to detect and remove moving cast-shadows in traffic imagery [10]–[17]. Shadow suppression can be classified into two main categories: shape-based approaches and spectrum-based approaches. Shape-based methodologies employ *a priori* geometric information for the scenes, the objects, and the light-source location to resolve the shadow detection problem. Hsieh *et al.* employed lane features and proposed a line-based algorithm to separate all unwanted shadows from the moving vehicles [10]. Yoneyama *et al.* designed two-dimensional (2-D) joint-vehicle/shadow models to represent the objects and their attached shadows in order to separate the shadows from the objects [11]. Under specific conditions, such as when vehicle shapes and illumination directions are known, these models can accurately detect shadows. However, they are difficult to implement since the knowledge of lane features, object classes, and illumination conditions are not readily available in practical applications.

On the other hand, spectrum-based techniques utilize spectral information of lit regions and shadow regions to detect shadows [12], [13]. Compared with the shape-based approaches, inference based on color information is more explicit, because the spectral relationship between lit regions and object-shadow regions is affected only by illumination, not by object shapes or light source directions. Horprasert *et al.* proposed a computational RGB color model to detect shadows [14]. Brightness distortions and chromaticity distortions are defined and normalized to classify each pixel. However, the detection accuracy of the algorithm is sensitive to the background model and it is often difficult to obtain enough foreground-free frames to build a reliable background model in traffic scenes. Cucchiara *et al.* used the hue-saturation-value (HSV) color information to extract shadow pixels from previously extracted moving foreground pixels [15]. For this, a threshold operation is performed to find shadow pixels and the empirically determined thresholds dominate the detection accuracy. The shortcoming of this method is that it is not an analytical method. Salvador *et al.* employed invariant color features to detect shadows [16]. These features describe the color configuration of each pixel discounting shading, shadows and highlights. They are invariant to the change of illumination conditions, and therefore are powerful indexes for detecting shadows.

However, the computational load of this method is rather large. Bevilacqua proposed a gray-ratio-based algorithm to effectively detect shadows with an empirically assigned ratio threshold [17]. This gray-ratio-based analysis considerably shortens the computation time, but easily misclassifies objects as shadows, because pixels of different colors may have a similar gray value.

The objective of this study is to propose a new, analytically resolved shadow model for effectively detecting shadows in traffic imagery. Current color-constancy-based shadow suppression methods examine each pixel of an image to build shadow models. Analyzing the spectrum properties of each pixel covered by shadows in an image sequence is a time-consuming process. Further, it is difficult to obtain enough shadow properties of image pixels for probability analysis from image sequences. In traffic imagery, the region of interest (ROI) is the roadway in the background image, while light sources such as the sun and the sky are practically fixed. The road can be treated as a Lambertian surface since the sun direction does not change over a few frames and the light source from the sky is relatively uniform. Thus, it is not necessary to model each pixel respectively for shadow detection because the shadow property of each road pixel is similar under the same Lambertian condition [18]. Nevertheless, color values of pixels are different from each other in color space, so it is impossible to construct a unique model for each pixel merely by using color space. Investigating other shadow properties in order to resolve the unique model clearly deserves attention for cast-shadow detection.

In this paper, a shadow-region-based statistical non-parametric (SNP) approach is developed to construct a unique model for shadow detection of all pixels in an image frame. To effectively and analytically build a model for shadow detection, we do not establish the model by examining each pixel in image sequences; instead, we construct the model by using a shadow region in a single image frame. In our design, a color ratio between lit pixels and shadow pixels is utilized as an index to establish the unique model for different shadow pixels. We will later show that in all image sequences, the ratio can be considered constant under the Lambertian condition. The model generation procedure requires much less effort and hence much smaller computational load compared with currently available methods. To further improve the performance of shadow suppression, a postprocessing stage of spatial analysis is added to verify the actual shadow pixels. The shapes and boundary information of the detected shadow region are used to verify the actual shadow pixels.

The rest of this paper is organized as follows. Section II describes the design of background estimation and moving object segmentation. Section III focuses on shadow detection algorithms. Experimental studies of a traffic monitoring system with shadow suppression are presented

in Section IV, and the limitations of the illustrated results are also discussed. Section V summarizes the contributions of this work.

II. BACKGROUND ESTIMATION AND MOVING VEHICLE SEGMENTATION

A basic step in analyzing traffic scenes is to detect moving vehicles. Moving vehicles can be segmented by subtracting the background image from the current image frame. In general, a common assumption is that each pixel value of a background scene can be well described by a statistical model. If the statistical model of the background imagery can be obtained, moving objects can be segmented by checking the captured pixel value with its corresponding background model [19], [20].

A. Gaussian Mixture Background Modeling

In statistical approaches to background estimation, techniques adopting multimodel probability distribution have been widely used to represent the recorded intensities of a pixel. In recent years, Gaussian mixture model (GMM) approaches to obtaining reliable background images have gained increasing attention for ITMS [21]. The intensity and chromatic values of an image pixel are modeled as a mixture of K Gaussian distributions. A K -means approximation algorithm updates the mixture model parameters using new pixel values. The Gaussians are ordered in descending order by the value of w/σ , where w is an estimate of the weight (the portion of the data accounted for by the Gaussian) and σ is the standard deviation of the Gaussian. The most relevant Gaussians are then chosen to complete the background model.

GMM features effective background estimation under environmental variations through a mixture of Gaussians for each pixel in the image frame. In urban traffic, however, vehicles stop occasionally at intersections because of traffic lights or control signals. Such kinds of transient stops increase the weight of nonbackground Gaussians and degrade the accuracy of background generation.

B. Single Gaussian Background Modeling

For an image sequence captured by a static camera, pixel values may have complex distributions, and the intensity of a background pixel will dominate the largest Gaussian. The values of each color channel of a background pixel can be modeled as a single Gaussian model and therefore be determined by analyzing its corresponding color histogram [22], [23]. In histogram approaches, the mean of a Gaussian background model can be determined by searching for the maximum frequency in the color histogram, because each frequency of histogram is proportional to its appearance probability.

In practical traffic imagery, very often several intensities of values may have the same maximum frequency in

the histogram due to sensor noises of image acquisition devices. It is thus difficult to determine the mean value of a background Gaussian model from the histogram. In order to resolve this problem, a group-based histogram (GBH) has been proposed to estimate a single Gaussian background model [24]. In the GBH, an accumulative frequency is generated by summing the frequencies of each individual intensity and those of its neighboring intensities in the histogram. The GBH effectively smoothes the frequency curve of conventional histogram for more accurate estimation of the Gaussian mean. The value that has the maximum frequency in the GBH is treated as the mean value of the background model. After the mean value μ of the background model is found, the variance σ can be computed as follows:

$$\sigma^2 = \frac{1}{\sum_{v=\mu-3\sigma'}^{\mu+3\sigma'} n(v)} \times \sum_{x=\mu-3\sigma'}^{\mu+3\sigma'} (v - \mu)^2 n(v) \quad (1)$$

where $n(v)$ is the count of pixels with value v , and σ' is the maximum standard deviation of the Gaussian background models. (The value can be experimentally obtained by analyzing the background models from image sequences offline.)

C. Foreground Segmentation

Pixel values obtained from an image sequence consist of a principal background Gaussian and many minor foreground distributions. The types of each pixel value can be simply categorized as: static background objects and moving foreground objects. To cope with the variation in the background intensity, a moving object is segmented if its intensity is located beyond $\pm 3\sigma$ from the mean μ of the Gaussian background model. Therefore, foreground objects are detected by checking the difference between the input RGB value and the background RGB value of the current image. An erosion operation of morphology is employed to remove salt-and-pepper noise [25]. The final foreground segmentation result consists of moving objects and their cast shadows; further analysis is then performed to separate shadows from the moving objects.

III. CAST-SHADOW DETECTION IN TRAFFIC IMAGE

Fig. 1 illustrates the block diagram of the proposed shadow suppression procedure. The complete system consists of four modules: a background estimation module (Section II-B), a background removal module (Section II-C), a shadow detection module, and a shadow verification postprocessing module. The shadow detection module employs an RGB spectral-ratio model to identify shadows. In the shadow verification module, spatial

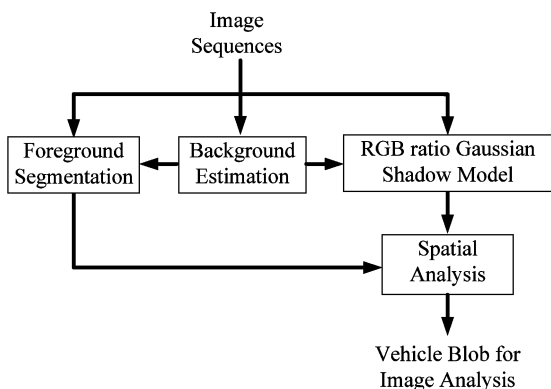


Fig. 1. Block diagram of the proposed shadow suppression method.

analysis schemes are applied to check whether a shadow pixel is true or not. Accordingly, moving objects and their cast-shadows can be separated.

A. Spectral Ratio Shadow Detection Algorithm

A traffic scene is illuminated by a faraway point source (the sun) and a diffuse source (the sky). Cast shadows in the scene are caused by sunlight occlusion. The distance between objects and their cast shadows is negligible in traffic scenes, compared with the distance between the light source and the objects. Thus, this type of the cast shadow is mostly an umbra or a strong shadow [12]. Shadow regions are darker than the background and their color spectrum also differs from that of the background. Since the RGB components of each pixel of the roadway differ from one another, it is impossible to use a unique model to detect the shadow of each pixel by merely using the RGB color space. Therefore, color space conversion or normalization is employed to find the model. To generate a shadow model, we hypothesize that the RGB ratio between lit regions and shadow regions is constant for all pixels of a traffic scene in image sequences. This hypothesis facilitates the construction of a unique model for shadow detection in an image frame. The model-built procedure can be effectively achieved by using a shadow region in the image, not from image sequences. Detailed reasoning of the constant-ratio hypothesis is presented in Appendix A.

According to the hypothesis, in each RGB channel, the color ratio between a lit pixel and a shadow pixel can be treated as a constant for every pixel in the traffic imagery.

As a result, a pixel is classified as a shadow if its RGB component satisfies

$$R_{\text{shadow}} = \alpha R_{\text{lit}} \quad \text{and} \quad G_{\text{shadow}} = \beta G_{\text{lit}} \\ \text{and} \quad B_{\text{shadow}} = \gamma B_{\text{lit}} \quad (2)$$

where $(R_{\text{lit}}, G_{\text{lit}}, B_{\text{lit}})$ is the RGB value of a lit pixel, $(R_{\text{shadow}}, G_{\text{shadow}}, B_{\text{shadow}})$ is the RGB value of a shadow pixel, and α, β, γ are constants whose values are smaller than one.

Based on (2), a shadow-region-based SNP method has been developed to construct the ratio model for shadow detection of all pixels in the image frame. Gaussian models are exploited to represent the constant RGB-color ratios between a lit pixel and a shadow pixel in this method. The unique ratio model can be found by analyzing shadow samples taken from the shadow region in an image frame. To cope with variations of the ratio, we use the Gaussian distribution inside $\pm 1.5\sigma$ (88.6%) as a threshold. Thus, a shadow pixel can be determined as shown in (3) at the bottom of the page, where $I_R(x, y)$, $I_G(x, y)$, and $I_B(x, y)$ are the input RGB values; $\mu_R(x, y)$, $\mu_G(x, y)$, and $\mu_B(x, y)$ denote the background RGB values at pixel position (x, y) ; r_R , r_G , and r_B are the RGB mean ratio values of a pixel when cast by a shadow in the image; σ_{rR} , σ_{rG} , and σ_{rB} represent the RGB standard deviation ratios of pixels when cast by a shadow in the background image.

Fig. 2(a)–(d) illustrates an example of the RGB Gaussian shadow model. In Fig. 2(a), 100 samples of shadow pixels in an image frame are selected to build the Gaussian RGB ratio model, as depicted by small white dots in the shadow region of the figure. To validate the Gaussian model constructed from the shadow region, we tested shadow data at three points [indicated by large white dots and labeled as S1, S2, and S3, respectively, in Fig. 2(a)]. The shadow data is recorded only if vehicle's shadow is cast on the points. Furthermore, all the recorded data are also combined (referred to as S_total) and used to verify the hypothesis of constant color ratio. The Gaussian RGB ratio models of recorded data and shadow-region data (labeled by region) are depicted in Fig. 2(b)–(d), respectively. The mean and the standard deviation of the RGB ratio of each sample are presented in Table 1. In this design, a pixel is regarded as a shadow pixel if its RGB ratio satisfies (3). The results computed from the four groups of samples (S1, S2, S3, and S_total) demonstrate that the accuracy is higher than 86%, as shown in Table 1. This

$$S(x, y) = \begin{cases} 1, & \text{if } \left| \frac{I_R(x, y)}{\mu_R(x, y)} - r_R \right| < 1.5\sigma_{rR} \text{ and } \left| \frac{I_G(x, y)}{\mu_G(x, y)} - r_G \right| < 1.5\sigma_{rG} \text{ and } \left| \frac{I_B(x, y)}{\mu_B(x, y)} - r_B \right| < 1.5\sigma_{rB} \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

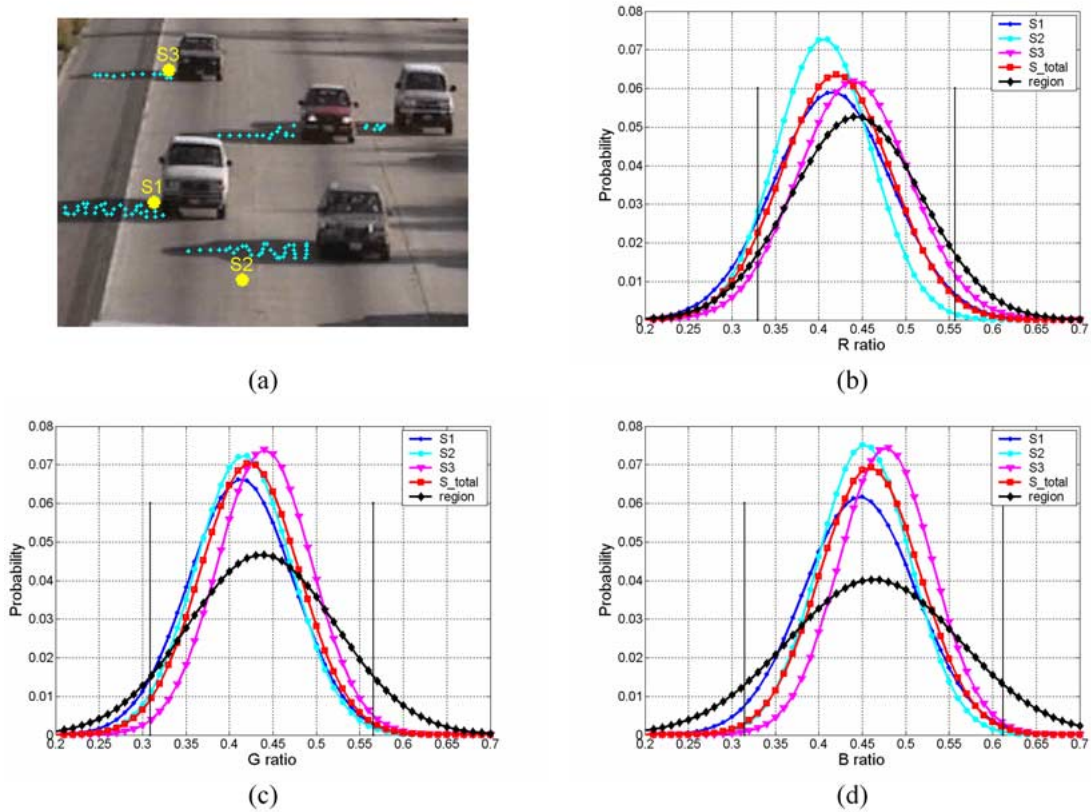


Fig. 2. Gaussian models of RGB ratio of recorded samples and shadow-region data.

Table 1 Gaussian Models of RGB Ratio of Recorded Data and the Shadow-Region Data

Item	Mean	Standard Deviation	Accuracy
R ratio of region	0.443	0.076	86.64%
R ratio of S1	0.416	0.067	88.16%
R ratio of S2	0.406	0.055	91.58%
R ratio of S3	0.440	0.064	92.16%
R ratio of S_total	0.420	0.063	91.15%
G ratio of region	0.437	0.085	86.64%
G ratio of S1	0.413	0.060	96.39%
G ratio of S2	0.416	0.055	96.70%
G ratio of S3	0.440	0.055	97.65%
G ratio of S_total	0.423	0.057	97.25%
B ratio of region	0.463	0.099	86.64%
B ratio of S1	0.447	0.065	97.47%
B ratio of S2	0.452	0.053	99.41%
B ratio of S3	0.477	0.054	99.30%
B ratio of S_total	0.459	0.058	99.02%

reveals the effectiveness of the hypothesis so the shadow-region-based RGB ratio model can be used to determine the shadow in image sequences. Fig. 3 illustrates an example of shadow detection. By using (3) and the Gaussian RGB ratio model of shadow-region data, one can detect the shadow as shown in Fig. 3(a). Fig. 3(b) shows the detected moving object. The result of SNP shadow suppression is shown in Fig. 3(c). It can be seen in Fig. 3(c) that some shadow pixels are not recognized as expected. This is mainly caused by the uncertainties in image sensing. We handle this insufficiency and improve the performance of shadow suppression by adding a post-processing spatial analysis step.

B. Spatial Analysis for Shadow Verification

Two types of shadow detection errors may commonly occur, namely shadow detection failure and object detection failure. The first type of error occurs if a shadow pixel has the ratios outside the detection range of the shadow model, causing the shadow to be unrecognizable. This can be seen in Fig. 3(c), where there are some shadow pixels that are to be unrecognized. The second type of detection failure occurs as the RGB color ratios of an object pixel lies inside the detection range of the shadow model.

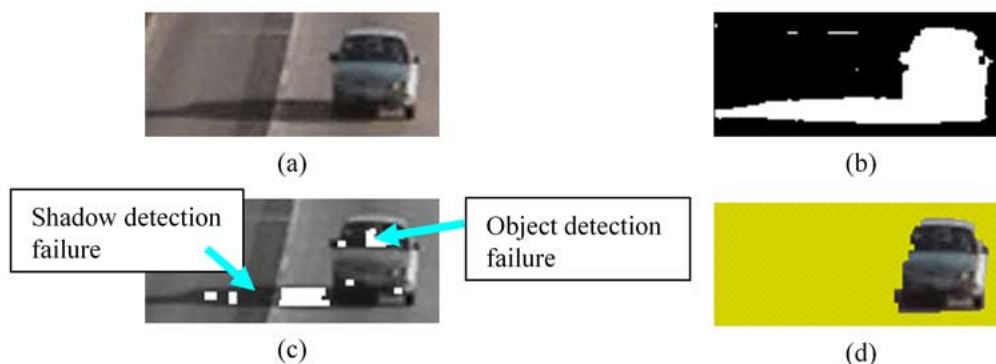


Fig. 3. Explanation of shadow suppression steps. (a) Original image. (b) Moving object segmentation result of background removal. (c) Shadow segmentation results of spectral ratio shadow detection. Detected shadows are indicated by white area. (d) Segmentation results of shadow suppression after spatial analysis.

This occurs especially when the ratios are higher than the mean of Gaussian ratio model and still inside the detection range; there are almost no shadow pixels in this region as shown in the sample plots of Fig. 2(b)–(d). For instance, in Fig. 3(c), partial pixels of the vehicle are misclassified as shadow pixels. To improve the accuracy of shadow detection, we propose adding a postprocessing spatial analysis for shadow confirmation. The spatial analysis is performed to confirm the true shadows as well as the true objects according to their geometric properties.

1) *Size Discrimination of Moving Object Candidates*: In the process of shadow detection, actual shadows sometimes break into small isolated shadow blobs. Generally, the sizes of these small shadow blobs are smaller than the detected moving vehicles in the image sequence. These small blobs will not be considered as moving object candidates. Thus, one can discriminate the small blobs of shadow from the big blobs of moving objects by using the size information. In this design, all blobs of moving object candidates are grouped into different regions using a connected components labeling algorithm [26]. The regions that have small sizes are recognized as shadow regions.

2) *Border Discrimination of Moving Cast Shadow Candidates*: The moving blobs segmented by background removal consist of shadow pixels and object pixels. In practice, the true shadow pixels will cluster in the fringes of the blobs. If a part of the detected vehicle is misclassified as a shadow, most of the boundary of this region will be located inside the candidate foreground, as shown in Fig. 3(b) and (c). If the shadow candidate is a true shadow, more than a half of the boundary should be adjacent to the boundary of foreground candidates. Thus, one can use the boundary information of a shadow-candidate region to confirm whether the shadow is a true shadow or not [17]. In this design, the boundaries of

foreground candidate are segmented by Sobel edge detection algorithm [25]. Next, each distinct candidate shadow region is determined by using a connected component's labeling algorithm. Sobel edge detection is also used to find the edge of each distinct shadow region. The number N_f of boundary shadow pixels that are adjacent to the boundary of a foreground-candidate region and the number N_s of all boundary shadow pixels are computed. The shadow is considered a true shadow if the ratio N_f/N_s is greater than 50%. Fig. 3(d) depicts the confirmation result of spatial analysis of Fig. 3(c). It can be seen that the accuracy of shadow detection is greatly improved in comparison with the original result. More detailed experimental results are presented in Section III-C.

C. Comparison Results

For traffic monitoring and enforcement applications, shadow suppression effectively reduces misclassification and erroneous counting of moving vehicles. The goal of shadow suppression is to minimize both the *false negative* (FN_S , the shadow pixels misclassified as background/foreground) and the *false positive* (FP_S , the background/foreground pixels misclassified as shadow pixels) detections. In order to systematically evaluate the performance of the proposed method, we adopt two metrics, namely the shadow detection rate η and the object detection rate ξ [13] for quantitative comparison

$$\eta = \frac{TP_S}{TP_S + FN_S} \quad (4)$$

$$\xi = \frac{TP_F}{TP_F + FN_F} \quad (5)$$

where TP_S and TP_F are respectively the number of shadow pixels and foreground pixels correctly identified; FN_S and

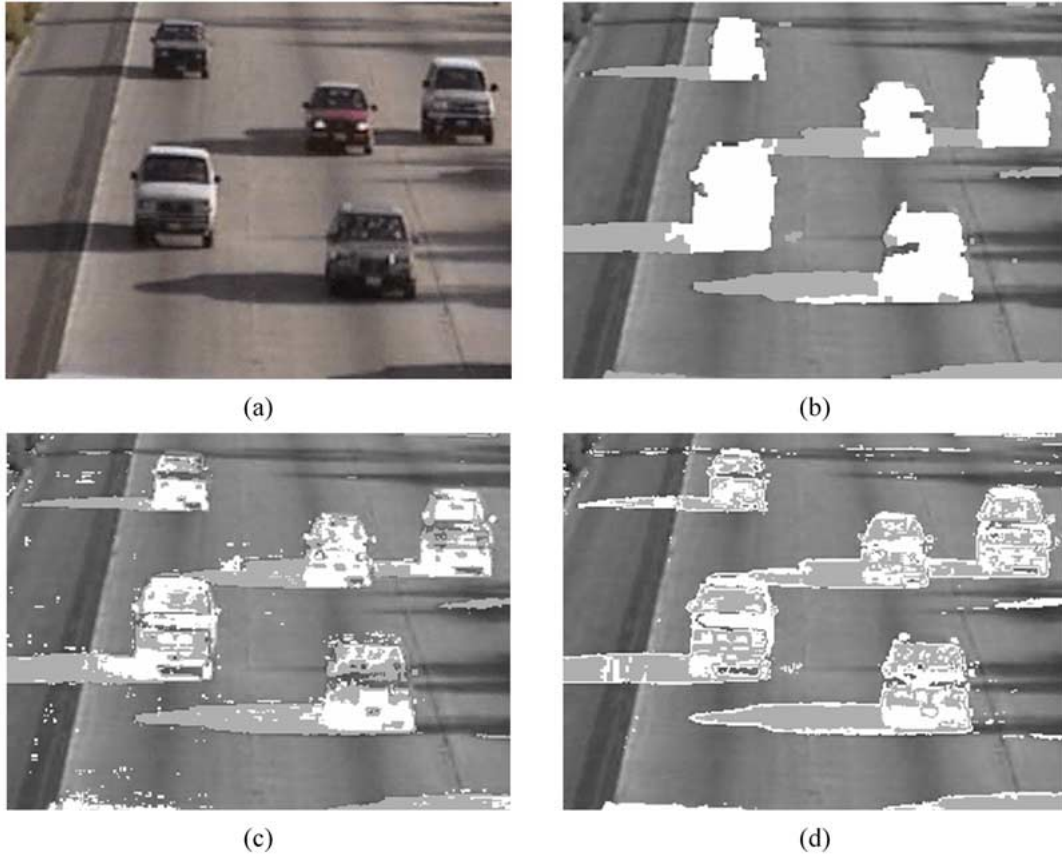


Fig. 4. Comparison results of shadow suppression method (white pixels represent the moving vehicle; the gray pixels represent the attached shadow). (a) Original image. (b) Proposed method. (c) SNP method. (d) DNM method.

FN_F are respectively the number of shadow pixels and foreground pixels falsely identified. A comparison with two well-known methods has been carried out to validate the performance of the proposed algorithm. An SNP approach [27] and a deterministic nonmodel (DNM)-based approach [15] were selected for the performance comparison. The SNP approach treats object colors as a reflectance model from the Lambertian hypothesis. It uses the normalized distortion of brightness α'_i and distortion of the chrominance CD'_i , computed from the difference between the background color of a pixel and its value in the current image, to classify a pixel in four categories, as shown in (6) at the bottom of the page. The

DNM method works in the HSV color space. Shadow detection is determined according to the following equation:

$$S(x, y) = \begin{cases} 1, & \text{if } \alpha \leq \frac{I_k^V(x, y)}{B_k^V(x, y)} \leq \beta \text{ and } I_k^S(x, y) - B_k^S(x, y) \leq \tau_S \\ & \text{and } D_H(x, y) \leq \tau_H \\ 0, & \text{otherwise} \end{cases} \quad (7)$$

where $D_H(x, y) = \min(|I_k^H - B_k^H|, 360 - |I_k^H - B_k^H|)$, $I_k(x, y)$ and $B_k(x, y)$ are the pixel values at (x, y) coordinate in the

$$C(i) = \begin{cases} \text{Foreground :} & CD'_i > \tau_{CD} \quad \text{or} \quad \alpha'_i < \tau_{\alpha lo} \quad \text{else} \\ \text{Background :} & \alpha'_i < \tau_{\alpha 1} \quad \text{and} \quad \alpha'_i > \tau_{\alpha 2} \quad \text{else} \\ \text{Shadowed background :} & \alpha'_i < 0 \quad \text{else} \\ \text{Highlighted background :} & \text{otherwise} \end{cases} \quad (6)$$

Table 2 Comparison of Computation Time for Shadow Spectral Analysis

Algorithm	Proposed	SNP	DNM
Time (msec)	3.4	93	15.6
Specification	Image size: 352 × 240 pixels, CPU type: Intel Pentium 4, 2.4GHz, RAM: 448MB.		

input image (frame k) and in the background model (computed at frame k), respectively.

For an objective comparison, the model of background image is updated in advance for all algorithms in the test. Fig. 4 shows the test results of these methods (Proposed, DNM [15], and SNP [27]) using the benchmark sequences Highway-I [13]. First, we check the computation time of each algorithm, as shown in Table 2. The proposed algorithm requires the least computation time. It takes only 3.6%–21% of the time needed by the other two methods because it merely utilizes division operation to obtain the shadow information. The SNP algorithm takes the longest time due to its complex normalization procedure, which consists of square, square root, addition, and division operations.

Second, we examine the shadow detection rate and the object detection rate of each algorithm. The ground truth for each frame is required for calculating the quantitative metrics (4) and (5). We manually and accurately classify the pixels into foreground, background, and shadow regions in the image sequences. Figs. 5 and 6 show the comparison results of shadow detection rate and object

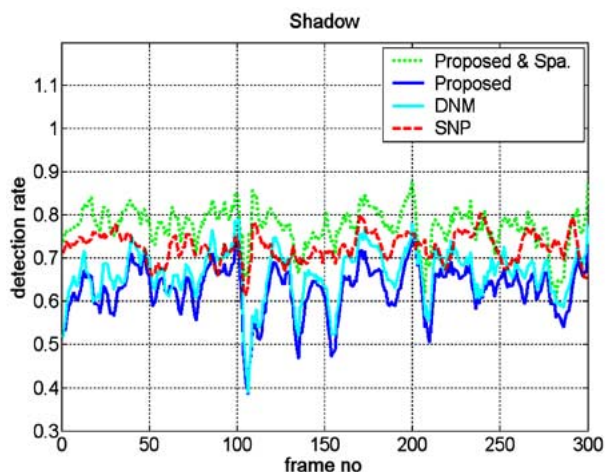


Fig. 5. Comparison results of shadow detection rate between the proposed method, DNM, and SNP.

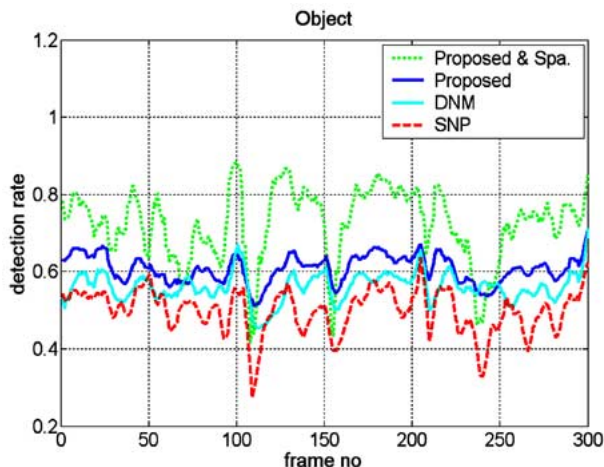


Fig. 6. Comparison results of object detection rate between the proposed method, DNM, and SNP.

detection rate of the proposed method, SNP and DNM, respectively. The accuracy corresponding to the plots in Figs. 5 and 6 are listed in Table 3. The results evaluated by Prati *et al.* [13] (obtained by analyzing dozens of frames for each video sequence representative of different situations) are also listed in Table 3 for a fair comparison. The experimental results show that the proposed method provides more reliable object detection results than the two state-of-the-art methods. The results obtained from each algorithm with or without spatial analysis are also listed in the table to check the effectiveness of spatial analysis. The merit of the proposed spatial analysis is that it can be combined with existing shadow suppression methods for further improvement of the performance. It can be seen that the spatial analysis significantly improves the performance of shadow detection. For instance, the shadow detection rate of the DNM algorithm is increased from 57% to 80%. The proposed method outperforms the other two methods in both shadow suppression and moving object detection. Its advantage is that it uses a ratio model that is constructed from only a single image frame. A video clip of the experimental results can be found at: http://isci.cn.nctu.edu.tw/video/PIEEE_WMV/PIEEE_shadow1.wmv.

IV. TRAFFIC MONITORING EXPERIMENTAL RESULTS

In this section, two experimental results of traffic monitoring with shadow suppression are presented. The video images under various illumination conditions were recorded in advance from an expressway, as well as at an urban intersection. Shadows attached to their respective moving vehicles degrade the performance of traffic monitoring if the ITMS has no shadow suppression

Table 3 Mean Accuracy of Shadow Detection Results and Object Detection Results

Method	Video-Sequence Shadow Detection Rate ^{*1}	Video-Sequence Object Detection Rate ^{*1}	Ten-Frame Shadow Detection Rate ^{*2}	Ten-Frame Object Detection Rate ^{*2}
Proposed & Spa.	77.5%	72.2%	76.86%	80.52%
Proposed	62.9%	60.7%	71.97%	70.14%
SNP	72.6%	49.9%	74.48% (81.59% ^{*3})	59.39% (63.76 ^{*4})
DNM	66.4%	56.07%	67.94% (69.72% ^{*3})	68.57% (76.93% ^{*4})
<p>*1: Mean accuracy of detection results obtained by analyzing 300 frames for each video sequence.</p> <p>*2: Mean accuracy of detection results obtained by analyzing tens of frames for each video sequence representative of different situations.</p> <p>*3: Results from [13].</p> <p>*4: Results from [13], the false positives belonging to the background were not considered in the computation of the object detection rate.</p>				

function. These experiments both validate the feasibility of the proposed algorithm and also show that the ITMS achieves high accuracy for estimating of traffic parameters such as traffic flow, traffic density, vehicle speeds, and vehicle turn ratio.

A. Image Tracking for Traffic Parameter Estimation

A method has been developed to simultaneously detect multiple vehicles and initiate tracking of vehicles on a multilane road [28]. To track vehicles, we adopted *B*-spline active contours to represent vehicle contours in the image plane [29]. Exploiting a linear transformation, the vehicle contour can be transformed to a shape-space vector with six elements. This simplifies the postprocessing of contour tracking in the image plane and the vehicle contour will be restricted to vary steadily by the shape-space vector. In the process of vehicle detection, the location of a vehicle can be detected by using a detection window or region-of-interest technique. Its size can be also determined by using the region which the vehicle image occupies. If a vehicle leaves the region, an initial contour is generated accordingly for tracking operations. Once initialized, the vehicle contour will be tracked and updated iteratively. Dynamic models are designed to predict the shape-space vector [30]. An image measurement procedure is responsible for obtaining the best-fitted curve of a vehicle outline in an image according to the vehicle contour generated from the predicted shape-space vector. A Kalman filter is designed to combine the information from the predicted states and the best-fitting measurement states. Assuming the vehicles lie in a flat plane and the camera has been calibrated, the pixel coordinates of the vehicle can be transformed into its world coordinates [31]. As vehicles in an image sequence are successfully tracked, traffic parameters such as traffic

flow, vehicle speed, and traffic density can be obtained through simple computation.

Practical experiments with traffic parameter extraction have been conducted to evaluate the shadow suppression and the tracking performance by using a recorded video of traffic scenes. The frame rate adopted in the experiment is 15 frames/s and image capture size is 352×240 pixels. Fig. 7 illustrates an example of the image tracking of cars and trucks on an intercity expressway. In the video scene, the shadow strength is medium and the shadow size is large. The ITMS will fail to detect and track the moving vehicles if there is no shadow suppression procedure in use. Fig. 7(b)–(l) shows that the proposed method successfully separates the shadows from the moving objects. In Fig. 7(b), two cars are detected in the detection window and another car is tracked. One car leaves the detection window and an active contour is initiated for tracking in Fig. 7(d). Three cars are tracked in Fig. 7(f). One car is tracked and a truck is detected in the detection window, as shown in Fig. 7(h). The truck leaves the detection window and an initial contour is generated for tracking in Fig. 7(j). In Fig. 7(l), the truck is tracked and a car is detected in the detection window. The experimental results demonstrate that the traffic monitoring system can successfully track multiple moving vehicles with shadow suppression. Partial experimental results can be found in the video clip at: http://isci.cn.nctu.edu.tw/video/PIEEE_WMV/PIEEE_track.wmv.

In this example, useful traffic parameters have been estimated. In the time span of 23 s, a total of 16 vehicles were detected (the ground truth is 16 vehicles). Table 4 shows the experimental results of the estimation of vehicle speed. In this table, the ground truth, which was manually measured from image sequences, is also presented for

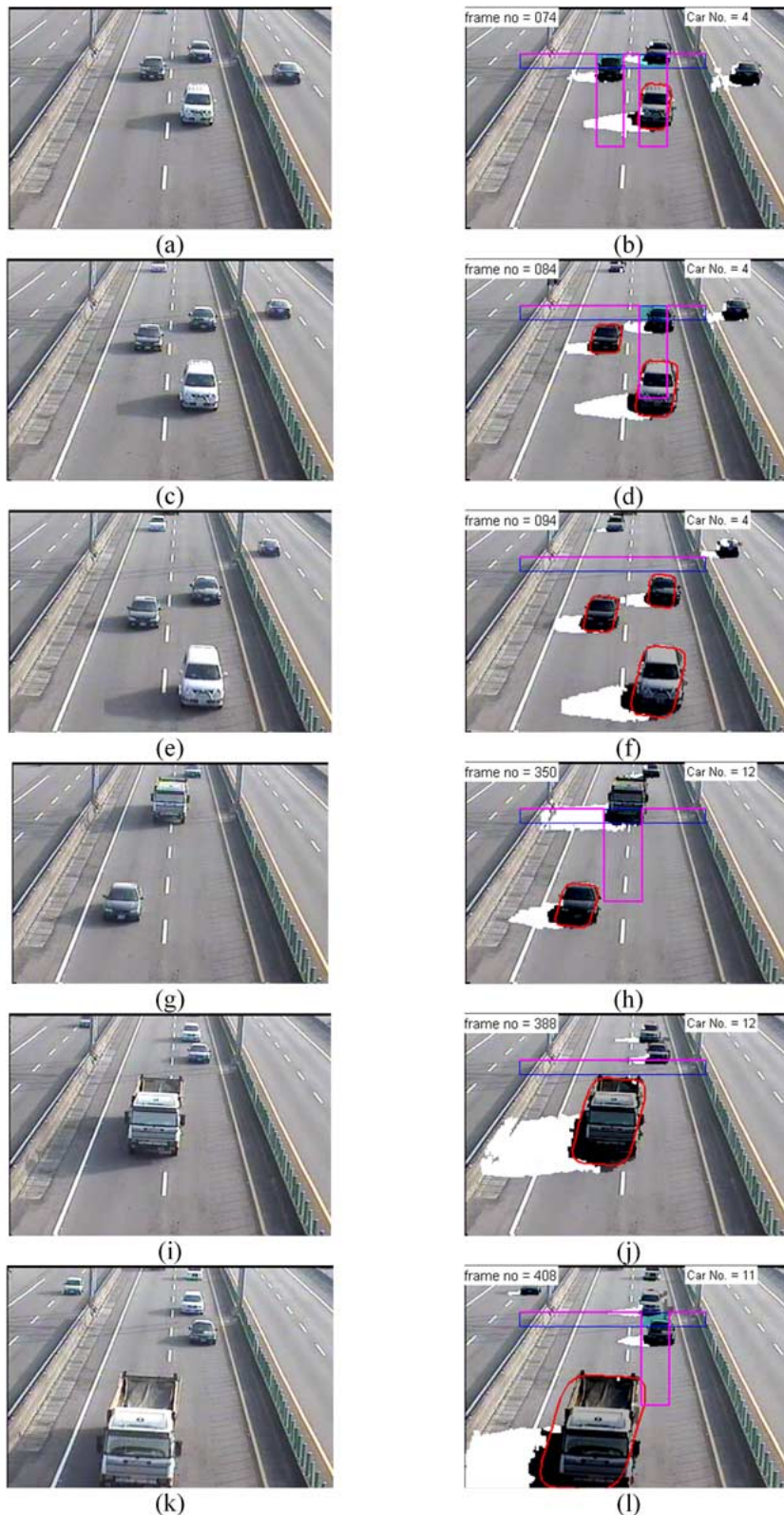


Fig. 7. Experimental results of vehicle tracking with shadow suppression on an expressway. (Detected shadows are indicated by white area and detected vehicles are indicated by darker areas.)

Table 4 Experimental Results of Vehicle Speed Estimation

Vehicle Index	Estimated Speed (km/hr)	Ground Truth (km/hr)	Error
1	84.1	80.6	4.3%
2	86.0	84.9	1.4%
3	86.4	85.7	0.8%
4	84.9	86.8	-2.2%
5	86.8	86.4	0.4%
6	101.3	103.5	-2.2%
7	82.9	87.6	-5.3%
8	100.6	95.7	5.1%
9	73.1	73.1	0.0%
10	75.0	76.8	-2.3%
11	77.0	77.3	-0.4%
12	79.3	77.6	2.2%
13	84.1	82.9	1.4%
14	54.9	54.0	1.7%
15	66.0	66.6	-0.9%
16	64.8	64.5	0.5%
Average	80.4	80.2	0.3%

comparison. Because the initial contour generated by the proposed method is well-suited for the active-contour-based image tracking with shadow suppression; the error of average speed estimation is within 5%. The traffic parameters are estimated as follows: the average speed is 80.4 km/h (the ground truth is 80.2 km/h), the traffic flow is 2504.3 car/h.

B. Turn Ratio Estimation

In this experiment, the left turn ratio of oncoming vehicles is estimated at a T-shaped intersection, as shown in Fig. 8. Two detection windows are set up to detect vehicles turning left and those moving straight ahead, respectively. Note that vehicles passing through the right region involve three types of motions—oncoming vehicles turning left, opposite direction ongoing vehicles moving straight, and vehicles of the cross-lane moving straight. Only the oncoming vehicles turning left need to be counted for the left-turn ratio estimation. It is clear that the motion directions of different type of vehicles are considerably different. The motion direction can be used to monitor the left-turn oncoming vehicles. The motion field can be computed by identifying corresponding pairs of feature points in two successive image frames taken at time t and $t + \Delta t$. In this design, the Harris corner detector is adopted to extract the corner points from traffic imagery because of its superior

repeatability, robustness to viewpoint changes and invariance to illumination variations [32].

Fig. 8(a)–(c) depicts the test results of shadow suppression and the motion vector estimation. The image sequences were captured shortly before sunset. In this scene, the shadow strength is high, and the shadow size is medium. The actual shadows have been successfully separated from the moving vehicles by the proposed method. In the figure, the small arrows indicate the detected motion direction of the vehicles. An average motion vector is obtained by considering the motion vectors of all feature points in the specific region. The direction of the large arrow shows the average motion vector of vehicles in the ROI. Fig. 8(a) shows two oncoming vehicles passing the ROI. One vehicle moves left and passes through the ROI in Fig. 8(b). Fig. 8(c) shows an ongoing vehicle passing through the ROI. The detected motion accurately reflects the motion type of vehicles in this region and the shadows have been suppressed for traffic monitoring. According to the information on the estimated direction, the system can identify the origin of the vehicles and determine the turn ratio accordingly. In this test, the actual left turn ratio is 7% and the calculated value is 7.2% with an error of 3.7%. A video clip of partial experimental results can be found at: http://iscn.nctu.edu.tw/video/PIEEE_WMV/PIEEE_ratio.wmv.

C. Discussion

In this presentation, our illustrations of results exclusively refer to cases where vehicles are clearly separated from each other. When vehicles are very close to each other and their shadow size is large, very often their shadows overlap other vehicles in the image frame. Under these circumstances, the geometric properties of true shadows will not be useful for shadow verification, as described in Section III-B. For instance, if the cast shadow of a vehicle is not recognizable and the shadow is adjacent to other vehicles, the shadows cannot be segmented properly by using size discrimination. Furthermore, the ratio will be smaller than 50% if a true shadow is surrounded by vehicles. It is easy to show that border discrimination will also misclassify the true shadow as an object. To solve the occlusion problem, edge and color information on vehicles and shadows can be exploited to verify shadow and moving vehicles. How to achieve effective vehicle tracking once vehicles or/and their shadows begin to overlap each other remains an open question.

V. CONCLUSIONS AND FUTURE DIRECTIONS

In this paper, an SNP method has been developed to construct a color ratio model for shadow detection of traffic imagery. This shadow detection method outperforms two well-known approaches in both shadow suppression rate and computation time. The proposed



Fig. 8. Experimental results of turn ratio estimation with shadow suppression at an intersection. Three types of driving direction of vehicles are distinguished: (a) moving right, (b) moving left, (c) moving straight ahead.

method has been successfully applied to an ITMS. Traffic parameters such as vehicle speed and traffic flow can be obtained with good accuracy even under shadow conditions. The absolute error rate is less than 5.3% for vehicle speed estimation and within 3.7% for the turn ratio estimation.

For future studies, more emphasis needs to be directed to increasing the robustness of the shadow detection and vehicle detection under occlusion and various illumination conditions. On one hand, the Gaussian ratio model built under a specific illumination condition might fail under a considerably different illumination. In traffic monitoring applications, it will be beneficial to build a database of

ratio models for different illumination conditions. On the other hand, shadow pixels that lie close to the moving vehicles or overlap other vehicles might be misclassified as the moving-vehicle pixels. The pixels within one shadow region may have similar color information in traffic imagery. The color distribution can be used to find uniform subregions and hence the uniform subregions can be used to verify the actual shadow region [33]. Heavy occlusion of vehicles influences the accuracy of image measurement, so methods should be developed to distinguish individual vehicles. Color information for individual tracked vehicles can be a useful tool for solving this problem [34]. ■

APPENDIX

RGB COLOR RATIO MODEL OF SHADOW PIXELS

In a daytime outdoor environment, there are effectively two light sources, namely a point light source and a diffuse extended light source. In the following derivation, the road is assumed to be Lambertian with a constant reflectance in a traffic scene. The radiance L_{lit} of the light reflected at a given point on a surface in the scene is formulated as [35]

$$L_{lit}(\lambda, i, e, g) = L_s(\lambda, i, e, g) + L_b(\lambda, i, e, g) + L_a(\lambda) \quad (A.1)$$

where $L_s(\lambda, i, e, g)$, $L_b(\lambda, i, e, g)$, and $L_a(\lambda)$ are the surface, body, and ambient reflection terms, respectively; λ is the wavelength; i is the angle of incidence between the illumination direction and the surface normal at a considered point; e is the reflection angle between the surface normal and the viewing direction; g is the phase angle illumination direction and the viewing direction. When the occlusion of sunlight creates shadows, the radiance L_{shadow} of the reflected light becomes

$$L_{shadow}(\lambda, i, e, g) = L'_a(\lambda) \quad (A.2)$$

where $L'_a(\lambda)$ is the ambient reflection term in the presence of the occluding object. To simplify the analysis, we assume that the ambient light coming from the sky is not influenced by the presence of the occluding objects, that is, $L'_a(\lambda) = L_a(\lambda)$.

The model is derived based on an RGB color space. The color components of the reflected intensity reaching the RGB sensors at a point (x, y) in the 2-D image plane can be expressed as

$$C_i = \int_{\Lambda} E(\lambda, x, y) S_{C_i}(\lambda) d\lambda \quad (A.3)$$

where $C_i \in \{R, G, B\}$ are the sensors responses, $E(\lambda, x, y)$ is the image irradiance at point (x, y) point, and $S_{C_i}(\lambda)$ is the spectral sensitivities of the RGB sensors of a color camera, respectively. Moreover, Λ is determined by $S_{C_i}(\lambda)$, which is nonzero over a bounded interval of wavelengths λ . We assume that the scene radiance and the image irradiance are the same because of Lambertian scenes under uniform illumination [34]. For a point in direct light, the sensor measurements are

$$C_i(x, y)_{lit} = \int_{\Lambda} [L_s(\lambda, i, e, g) + L_b(\lambda, i, e, g) + L_a(\lambda)] S_{C_i}(\lambda) d\lambda. \quad (A.4)$$

When a point is in the shadow, the sensor measurements are

$$C_i(x, y)_{shadow} = \int_{\Lambda} L_a(\lambda) S_{C_i}(\lambda) d\lambda. \quad (A.5)$$

Since the reflection terms $L_s(\lambda, i, e, g)$, $L_b(\lambda, i, e, g)$, and $L_a(\lambda)$ of a road surface are similar in a traffic scene for each object point of the road surface, the RGB measurement ratio between the lit and the shadow condition is approximately constant

$$\frac{C_i(x, y)_{shadow}}{C_i(x, y)_{lit}} = \text{const}. \quad (A.6)$$

that is, $R_{shadow} = \alpha R_{lit}$, $G_{shadow} = \beta G_{lit}$, and $B_{shadow} = \gamma B_{lit}$, where α, β, γ is less than 1.

Acknowledgment

The authors would like to thank the anonymous reviewers for their constructive comments and suggestions.

REFERENCES

- [1] B. McQueen and J. McQueen, *Intelligent Transportation Systems Architectures*. Norwood, MA: Artech House, 1999, pp. 19–49.
- [2] K. Hayashi and M. Sugimoto, "Signal control system (MODERATO) in Japan," in *Proc. IEEE Int. Conf. Intell. Transport. Syst.*, Tokyo, Japan, 1999, pp. 988–992.
- [3] G. K. H. Pang, K. Takabashi, T. Yokota, and H. Takenaga, "Adaptive route selection for dynamic route guidance system based on fuzzy-neural approaches," *IEEE Trans. Veh. Technol.*, vol. 48, no. 6, pp. 2028–2041, Nov. 1999.
- [4] V. Kastrinaki, M. Zervakis, and K. Kalaitzakis, "A survey of video processing techniques for traffic applications," *Image and Vision Computing*, vol. 21, no. 4, pp. 359–381, Dec. 2003.
- [5] 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.
- [6] N. J. Ferrier, S. M. Rowe, and A. Blake, "Real-time traffic monitoring," in *Proc. 2nd IEEE Workshop on Applications of Computer Vision*, Sarasota, FL, 1994, pp. 81–88.
- [7] D. Koller, K. Daniilidis, and H. H. Nagel, "Model-based object tracking in monocular image sequences of road traffic scenes," *Int. J. Comput. Vision*, vol. 10, no. 3, pp. 257–281, Jun. 1993.
- [8] H. Veeraraghavan, O. Masoud, and N. Papanikolopoulos, "Computer vision algorithms for intersection monitoring," *IEEE Trans. Intell. Transp. Syst.*, vol. 4, no. 2, pp. 78–89, Jun. 2003.
- [9] K. T. Song and J. C. Tai, "Dynamic calibration of pan-tilt-zoom cameras," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, 2006.
- [10] J. W. Hsieh, S. H. Yu, Y. S. Chen, and W. F. Hu, "A shadow elimination method for vehicle analysis," in *Proc. IEEE Int. Conf. Pattern Recognition*, Cambridge, U.K., 2004, pp. 372–375.
- [11] A. Yoneyama, C. H. Yeh, and C. C. J. Kuo, "Moving cast shadow elimination for robust vehicle extraction based on 2-D joint vehicle/shadow models," in *Proc. IEEE Int. Conf.*

- Advanced Video and Signal Based Surveillance*, Miami, FL, 2003, pp. 21–22.
- [12] S. Nadimi and B. Bhanu, “Physical models for moving shadow and object detection in video,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 8, pp. 1079–1087, Aug. 2004.
- [13] A. Prati, I. Mikic, M. M. Trivedi, and R. Cucchiara, “Detecting moving shadows: Algorithms and evaluation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 7, pp. 918–923, Jul. 2003.
- [14] T. Horprasert, D. Harwood, and L. S. Davis, “A robust background subtraction and shadow detection,” in *Proc. 4th Asian Conf. Computer Vision*, Taipei, Taiwan, R.O.C., 2000, pp. 983–988.
- [15] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, “Detecting moving objects, ghosts and shadows in video streams,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 25, no. 10, pp. 1337–1342, Oct. 2003.
- [16] E. Salvador, A. Cavallaro, and T. Ebrahimi, “Cast shadow segmentation using invariant colour features,” *Comput. Vision Image Understanding*, vol. 95, no. 2, pp. 238–259, Aug. 2004.
- [17] A. Bevilacqua, “Effective shadow detection in traffic monitoring applications,” *J. Winter School of Computer Graphics*, vol. 11, no. 1, pp. 57–64, Feb. 2003.
- [18] Y. Sato and K. Ikeuchi, “Reflectance analysis under solar illumination,” in *Proc. IEEE Workshop Physics-Based Modeling and Computer Vision*, Cambridge, MA, 1995, pp. 180–187.
- [19] N. L. Seed and A. D. Houghton, “Background updating for real-time image processing at TV rates,” *SPIE Image Processing, Analysis, Measurement and Quality*, vol. 901, pp. 73–81, 1988.
- [20] C. Eveland, K. Konolige, and R. Bolles, “Background modeling for segmentation of video-rate stereo sequences,” in *Proc. IEEE Conf. Computer Vision Pattern Recognition*, Santa Barbara, CA, 1998, pp. 266–271.
- [21] C. Stauffer and W. E. L. Grimson, “Learning patterns of activity using real-time tracking,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 747–757, Aug. 2000.
- [22] D. Gutches, M. Trajkovic, E. Cohen-Solal, D. Lyons, and A. K. Jain, “A background model initialization algorithm for video surveillance,” in *Proc. IEEE Int. Conf. Computer Vision*, Vancouver, BC, Canada, 2001, pp. 733–740.
- [23] P. Kumar, S. Ranganath, H. Weimin, and K. Sengupta, “Framework for real-time behavior interpretation from traffic video,” *IEEE Trans. Transp. Syst.*, vol. 6, no. 1, pp. 43–53, Mar. 2005.
- [24] J. C. Tai and K. T. Song, “Background segmentation and its application to traffic monitoring using modified histogram,” in *Proc. IEEE Int. Conf. Networking Sensing Control*, Taipei, Taiwan, R.O.C., 2004, pp. 13–18.
- [25] R. C. Jain, R. Kasturi, and B. G. Schunck, *Machine Vision*. New York: McGraw-Hill, 1995.
- [26] L. G. Shapiro and G. C. Stockman, *Computer Vision*. Englewood Cliffs, NJ: Prentice-Hall, 2001.
- [27] I. Haritaoglu, D. Harwood, and L. S. Davis, “W⁴: Real-time surveillance of people and their activities,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 22, no. 8, pp. 809–830, Aug. 2000.
- [28] J. C. Tai, S. T. Tseng, C. P. Lin, and K. T. Song, “Real-time image tracking for automatic traffic monitoring and enforcement applications,” *Image Vis. Comput.*, vol. 22, no. 6, pp. 485–501, Jun. 2004.
- [29] A. Blake and M. Isard, *Active Contours*. London: Springer Press, 1998.
- [30] J. C. Tai and K. T. Song, “Automatic contour initialization for image tracking of multi-lane vehicles and motorcycles,” in *Proc. IEEE 6th Int. Conf. Intell. Transport. Syst.*, Shanghai, China, 2003, pp. 808–813.
- [31] A. H. S. Lai and N. H. C. Yung, “Lane detection by orientation and length discrimination,” *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 30, no. 4, pp. 539–548, Aug. 2000.
- [32] C. J. Harris and M. Stephens, “A combined corner and edge detector,” in *Proc. 4th Alvey Vision Conf.*, Manchester, U.K., 1988, pp. 147–151.
- [33] D. Comaniciu and P. Meer, “Mean shift: A robust approach toward feature space analysis,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 5, pp. 603–619, May 2002.
- [34] P. Favaro and S. Soatto, “A variational approach to scene reconstruction and image segmentation from motion blur cues,” in *Proc. IEEE Int. Conf. Comp. Vis. and Patt. Recog.*, Washington, DC, 2004, pp. 631–637.
- [35] E. Salvador, “Shadow Segmentation and Tracking in Real-World Conditions,” Ph.D. dissertation, Signal Processing Institute, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, 2004.

ABOUT THE AUTHORS

Kai-Tai Song (Associate Member, IEEE) was born in Taipei, Taiwan, R.O.C., in 1957. He received the B.S. degree in power mechanical engineering from National Tsing Hua University, Hsinchu, Taiwan, R.O.C., in 1979 and the Ph.D. degree in mechanical engineering from the Katholieke Universiteit Leuven, Belgium, in 1989.

He was with the Chung Shan Institute of Science and Technology from 1981 to 1984. Since 1989, he has been on the faculty and is currently a Professor in the Department of Electrical and Control Engineering, National Chiao Tung University, Taiwan. His areas of research interest include mobile robots, image processing, visual tracking, sensing and perception, embedded systems, intelligent system control integration, and mechatronics.

Dr. Song served as the chairman of the Society of IEEE Robotics and Automation, Taipei Chapter, from 1998 to 1999.



Jen-Chao Tai received the B.S. degree in power mechanical engineering from National Tsing Hua University, Hsinchu, Taiwan, R.O.C., in 1985 and the M.S. and Ph.D. degrees in electrical and control engineering from National Chiao Tung University, Hsinchu, Taiwan, R.O.C., in 1992 and 2006, respectively.

He has been with the Department of Mechanical Engineering, Minghsin University of Science and Technology, Hsinchu, Taiwan, R.O.C., since 1992. He is now an Associate Professor in the Department. His research interests include image processing, visual tracking, vision-based traffic parameter estimation, dynamic camera calibration, real-time imaging system, and mechatronics.

