

Effective moving cast shadow detection for monocular color traffic image sequences

George S. K. Fung
Nelson H. C. Yung
Grantham K. H. Pang
Andrew H. S. Lai
University of Hong Kong
Department of Electrical and Electronic
Engineering
Laboratory for Intelligent Transportation
Systems Research
Pokfulam Road, Hong Kong SAR
E-mail: skfung@eee.hku.edu

Abstract. For an accurate scene analysis using monocular color traffic image sequences, a robust segmentation of moving vehicles from the stationary background is generally required. However, the presence of moving cast shadow may lead to an inaccurate vehicle segmentation, and as a result, may lead to further erroneous scene analysis. We propose an effective method for the detection of moving cast shadow. By observing the characteristics of cast shadow in the luminance, chrominance, gradient density, and geometry domains, a combined probability map, called a shadow confidence score (SCS), is obtained. From the edge map of the input image, each edge pixel is examined to determine whether it belongs to the vehicle region based on its neighboring SCSs. The cast shadow is identified as those regions with high SCSs, which are outside the convex hull of the selected vehicle edge pixels. The proposed method is tested on 100 vehicle images taken under different lighting conditions (sunny and cloudy), viewing angles (roadside and overhead), vehicle sizes (small, medium, and large), and colors (similar to the road and not). The results indicate that an average error rate of around 14% is obtained while the lowest error rate is around 3% for large vehicles. © 2002 Society of Photo-Optical Instrumentation Engineers.
[DOI: 10.1117/1.1473638]

Subject terms: shadow detection; image sequence analysis; visual traffic surveillance; image segmentation; intelligent transportation systems.

Paper 010176 received May 21, 2001; revised manuscript received Aug. 23, 2001 and Oct. 2, 2001; accepted for publication Oct. 19, 2001.

1 Introduction

Visual traffic surveillance (VTS) is one of the major components within the research of intelligent transportation systems¹ (ITS). Its main purpose is to remotely acquire traffic image sequences from roadside surveillance cameras and interpret them into traffic parameters and vehicle behavior. To achieve these requirements, numerous image processing algorithms, including preprocessing and post-processing algorithms, have been developed.² Among them, segmentation algorithms that extract the objects of interest (such as moving vehicles) from the image background in an image sequence have recently been actively studied,^{3,4} and background subtraction is a common approach.^{5,6} However, most of these approaches suffer a major drawback. In outdoor daylight scenes, shadows cast by moving vehicles are often detected as part of the objects since shadows move in accordance with the movement of the objects. When the detected vehicles contain shadows, the calculated location, dimension, speed, and number often include large errors. For instance, in a traffic scene with long shadows, vehicle location may be incorrectly estimated on the shadow region rather than on the vehicle body by the center-of-gravity method. This also creates a multitude of problems associated with occlusion. Therefore, the accuracy and robustness of the algorithm may be seriously affected if the cast shadow is not detected and removed. Although numerous shadow detection methods have been proposed,^{7–18} they all suffer from a number of limitations that make them inef-

fective in practical outdoor environments. For example, some of them are limited to indoor environments only, and for those that can be used outdoors, environmental information is usually required. Thus, we are motivated to resolve this problem of separating cast shadows from the moving vehicles in a practical outdoor environment.

In this paper, we propose a method that can effectively detect the cast shadow of a moving vehicle from a monocular color traffic image sequence. We assume that the mask of the moving vehicle (the region both covering the vehicle and the cast shadow) and an estimated static background reference image are both available together with the input image sequence. This assumption is considered reasonable as many background estimation methods have been reported^{5,6} and can be used to determine the motion content of the image, and the moving foreground mask (MFM) can be directly computed from the reference background and the input image. Given the MFM, we obtain the shadow confidence score (SCS), which indicates the likelihood of shadow, based on the observations of the cast shadow characteristics. The edge pixels of the input image within the MFM are then computed and classified into object-edge pixels and non-object-edge pixels using the neighboring SCSs, where the object edge pixels are bounded by a convex hull. This convex hull denotes the vehicle region, while the remaining pixels of the MFM denote the cast shadow region. To evaluate the proposed method, we analyzed the selection of the parameters used to control the score functions, as well as the object edges classification. The pro-

posed method was tested on about 100 images taken under different lighting conditions (sunny and cloudy), viewing angles (roadside and overhead), vehicle sizes (small, medium, and large) and colors (similar to the road and not). The results indicate that an average error rate of around 14% is obtained while the lowest error rate is around 3% for large vehicles.

The paper is organized as follows. Section 2 gives a survey of various shadow detection approaches developed. The strengths and limitations of these approaches are considered. Section 3 presents the basic concept and methodology of the proposed method. Section 4 defines the computation of the SCS. Section 5 discusses the final steps in moving cast shadow detection. Section 6 outlines, evaluates, and analyzes a series of tests on typical outdoor traffic scenes. Finally, the paper is concluded in Sec. 7.

2 Related Works

Due to the importance of cast shadow detection, numerous shadow detection methods have been proposed in the last decade.^{7–18} Basically, the methods can be classified into three categories: (1) the single-frame approach, (2) the interframe- or reference-frame approach, and (3) other approaches.

2.1 Single-Frame Approach

Traditionally, cast shadow detection algorithms have been mostly developed based on the single-frame approach.^{7–10} As there are limited cast shadow characteristics that can be extracted from a single input frame, authors tend to make stricter assumptions as the basis for the algorithm development.

For instance, Scanlan et al.⁷ presented a shadow removal algorithm that employs a simple histogram modification function on the image intensity. They assumed that objects occupy only the uppermost intensity range of the image, and that the image is background-dominant. To preserve the image texture and edges, the image is partitioned into blocks. The mean intensity of the pixels in each block is computed and stored. They assumed that the median of the block intensity means, which is used as a scaling reference, is sufficient to isolate the object region from the background. All the blocks with mean intensity values below the median are considered as shadow elements and scaled to the median value. In theory, background shadows can be removed and the object regions are left untouched. Their airport aerial view experiment showed reasonably good results since the airplane occupies the uppermost intensity range of the image and is larger than the median of the respective mean image. It is, however, not indicative from their experiment that their algorithm will work equally well in other practical scenes involving objects that possess intensity content different from their assumptions.

Jiang and Ward⁸ presented a shadow identification and classification method for real images in a laboratory environment. In their method, the shadow intensity and shadow geometry are analyzed. The image is classified into object, self-shadow, and cast shadow regions based on a number of shadow hypotheses. However, their method is governed by assumptions such as shadow must cast on a flat surface, and there must be only one light source in the scene. Although satisfactory results were achieved in their simple indoor

images with a single-colored flat surface background and a single light source, these assumptions would likely restrict the method from being applicable in outdoor environments where complex lighting and unstructured background surfaces are both common.

Funka-Lea and Bajcsy⁹ presented an active shadow recognition method by combining color and geometric properties of the image. They suggested a number of cues that together point toward the identification of a shadow. One of the cues is that the intensity, hue, and saturation changes due to shadows tend to be predictable. The image is segmented by a color image segmentation method that recovers a single material surface as a single image region regardless of whether the surface is partially in shadow. The penumbra and umbra of shadows are then recovered based on an illumination model called the linear color cluster assumption for penumbras (LCCAP). To recover the geometric properties of the scene, such as the location of the light sources, an extendable probe is also used to actively obtain shadows in the scene. Both outdoor and indoor scenes were tested and the shadows were reasonably detected. However, due to the use of the linear color cluster assumption, their method is limited to relatively simple scenes. Moreover, the umbra and penumbra properties of shadow can hardly be maintained in complex outdoor scene.

Salvador et al.¹⁰ presented a method that is based on the use of invariant color models to identify and classify shadows in color images. The candidate shadow regions are first extracted by searching the edge map in the dark regions of the image. After color conversion to an invariant color model, the candidate shadow pixels are classified as self-shadow points or as cast shadow points based on the detected color edge of the image. The method was successfully applied to a number of indoor scenes with one or two simple objects and one light source. Similar to other single-frame approach methods, the application of their method is restricted by its assumptions that shadows are cast on a flat and nontextured surface, objects are uniformly colored, and a single light source illuminates the scene.

2.2 Interframe or Reference Frame Approach

Essentially, the interframe approach¹¹ and reference frame approaches^{12–15} are intrinsically very similar. Both approaches utilize multiple frames in an image sequence. They explicitly explore the change of the region properties under shadow as the criteria of the detection process. Their major difference is that the interframe approach uses the previous frames directly, while the reference frame approach uses the previous frames to generate a reference frame for comparison. These approaches are suitable for detecting cast shadows that are associated with moving objects captured by a static camera.

By using the interframe difference and explicitly detecting the penumbra and umbra properties of shadow, Stauder et al.¹¹ proposed a detection method for ideal indoor cast shadow. Their algorithm works well under the assumptions that there is a plane background and the light source is of nonnegligible size and intensity. According to their simulation results on three test sequences, their algorithm is able to detect single or multiple moving cast shadows in indoor video sequences with spotlights and cast shadows on the background. If the shadows are weak, their algorithm may

fail as the assumption for penumbra and umbra properties of shadow may not hold. Moreover, as interframe difference is used, shadows can be entirely detected only if their background is revealed entirely.

Kilger¹² presented a shadow handler in a video-based traffic monitoring system. The object and its shadow are detected by comparing the current frame with the background image and then described by a bounding box. From the geometry of the scene (heading of the observed road) and other global data (date and time), the searching criteria and direction for cast shadows are determined by the expected shadow direction derived from *a priori* knowledge about the scene. According to the expected position of the shadow, the edge histogram of the detected region is computed. The author argued that as shadow is relatively edgeless in nature, the subtraction of the edge histogram from the corresponding one in the background image indicates the presence of the vehicle. By detecting the first edge of the vehicle along the searching direction, the vehicle is then described by a refined bounding box. The algorithms were tested under normal traffic and daylight conditions for a number of image sequences. Given the expected shadow position and overhead camera setting, a high detection rate was reported. However, to guide the shadow search this algorithm requires environmental knowledge, which changes dynamically. As bounding box is used to describe the vehicle region and the vehicle appears in various poses depending on the moving direction of the vehicle, a significant portion of the shadow may still be present after the detection.

Gamba et al.¹³ realized two different shadow detection approaches: shadows associate with still objects in the scene, and shadows are more similar than the actual objects between corresponding targets. By exploiting the hue, luminosity, and saturation components, an algorithm that extracts a shadow model from a monocular color scene was presented. Their algorithm was tested on a number of image sequences of a parking lot and reported low misclassified pixels for most cases. Their analysis works well only when there are enough shadows in the reference image. Moreover, they have implicitly assumed that all the shadows are cast on the same kind of surface, which is in fact not true in most outdoor cases.

Mikic et al.¹⁴ presented an algorithm that statistically classifies pixels into the shadow, object, and background classes. In their approach, the color response of the camera is statistically predetermined as a diagonal matrix. Based on the given *a priori* probabilities of the pixel belonging to different classes, the pixels are classified by maximizing the *a posteriori* probability of the class membership. A spatial smoothing filter is used to filter the noisy shadow detection results computed from the previous stage. Their algorithm was successfully tested on a traffic scene with long shadows. However, computation of the diagonal matrix is highly dependent on the camera settings and may lead to performance degradation if there are changes in scene conditions.

Horprasert et al.¹⁵ presented an algorithm for detecting moving objects from a static background scene that contains shading and shadow. They developed a background subtraction algorithm that is able to cope with local illumination changes and global illumination changes. A compu-

tational color model is employed to decompose the distortion measurement of the difference between the input image and the background image into brightness distortion and chromaticity distortion. Based on their color model, their algorithm classifies pixels into (1) the original background class, (2) the shaded background or shadow class, (3) the highlighted background class, and (4) the moving foreground object class. The pixel classification procedure is defined by the decision rules using normalized brightness distortion and normalized chromaticity distortion. They successfully applied the method to image sequences of a person moving in indoor and outdoor scenes.

2.3 Other Approaches

Apart from the single-frame and interframe/reference-frame approaches, there are the model-based approach,¹⁶ the neural network approach,¹⁷ and the core-line approach.¹⁸

Instead of separating the shadow from the object, Koller et al.¹⁶ defined an illumination model and shadow formation model to handle the shadow in road traffic scene. A simple illumination model, which assumed parallel incoming light, was used and the visible contour of the 3-D vehicle model projected onto the street plane was computed. This approach is believed to be feasible only for limited real outdoor surveillance purposes, because of *a priori* parameter setting for the illumination direction, the unrealistic assumptions of the simplified illumination model, and the highly complex interpretation.

Tao et al.¹⁷ presented a shadow elimination method based on a fuzzy neural network approach for an outdoor mobile robot. A multiresolution approach is adopted to achieve the required accuracy, efficiency and speed of operation. The parameters of the membership functions of the fuzzy rules are optimized by a genetic algorithm. The inputs of the fuzzy neural network are the red, green, and blue component of the image pixels. The output is the result of shadow recognition. Without much details of the shadow discrimination process, it is suspected that the neural-network-based approach would require a large number of training samples captured under different lighting conditions to achieve reasonable results. The authors reported results by analyzing 100 images captured under different lighting conditions such as time of the day and year, but there was a lack of analytical and objective measures.

Sonoda and Ogata¹⁸ presented a technique of separating the moving object, standing in an erected posture, from its shadow. The authors assumed that the position of the light source with reference to the camera is given beforehand. Hence, the expected shadow direction is defined, from which two "core lines" that pass around the center of the extracted region of the object and the shadow are obtained automatically. Based on these lines and the matching with the predefined template, a "separation point" is decided to separate the object from its cast shadow. Their algorithm was successfully tested on an image sequence with three people walking across the scene. However, the applicability of this algorithm is strictly limited by the assumption that the targeted object is in an elongated shape and an erected posture.



Fig. 1 Double-deck bus example.

2.4 Summary

The single-frame approach utilizes only the information provided in a single image. To detect cast shadow with this limited information, some strict assumptions must be made. On the other hand, the interframe/reference-frame approach utilizes multiple previous frames to provide information over the temporal domain. The properties of the region under shadow and not under shadow can be extracted. Therefore, in most practical cases, the interframe/reference-frame approach achieves higher accuracy and robustness compared with the single-frame approach. Other approaches provide some robust methods but are limited by their specific applications. Overall, among all these approaches, the reference-frame approach is the most promising. However, most of the current reference-frame approaches consider only a specific aspect of the shadow, and do not fully utilize the spectrum of features that may be useful for the eventual classification of object and shadow. Therefore, it is our intention to include those unique shadow features and integrate them to give a combined confidence score to each pixel in the region of interest. To develop and test the method effectively, we must have an objective way to measure the error rate of the classification, and a large image data set to cover different outdoor environments.

3 Methodology

3.1 Observations

As defined in Ref. 8, there are two parts in a shadow: self-shadow and cast shadow. Self-shadow is the part of object that is not illuminated by direct light, while cast shadow is the region projected by the object in the direction of direct light. In this paper, our objective is to detect the cast shadow associated with a moving object. Although the formation of cast shadow depends on various environment factors, we observed that there are four generic features of cast shadow that can be considered.

To illustrate our observations, we use Fig. 1 as an ex-

ample. Figure 1(a) depicts a typical outdoor traffic scene with a double-deck bus under bright sunlight. A cast shadow falls on the road surface to the left of the bus. Lai and Yung¹⁹ presented a stationary background estimation algorithm for color image sequence. In the algorithm, the running mode and running average algorithms are employed as the estimation core. To select between the running mode and running average algorithm in each estimation iteration, a scoreboard is employed to keep track the pixel variations in the image sequence. The estimated background is depicted as shown in Fig. 1(b). Figure 1(c) shows the MFM, which is obtained by subtracting the background image from the input image, followed by mathematical morphological closing to join the disjoint regions that belong to the same object. This mask essentially defines the filter window for the subsequent processes. (Note that the MFM of the taxi at the upper left-hand corner is intentionally ignored to keep our illustration simple.) In Fig. 1(c), there are some holes inside the MFM which commonly occur in background subtraction because of the vehicle parts being similar to the corresponding background. The problem associated with these holes is resolved in the later stage by using convex hull to bound the object region.

Observation 1. *The luminance values of the cast shadow pixels are lower than those of the corresponding pixels in the background image.*

Since cast shadow is the region with lower illumination, the shadow region appears to be darker, as illustrated in Fig. 2. (All the figures in Fig. 2 are filtered by the MFM.) Figure 2(a) depicts the luminance of the double-deck bus within the mask. Figure 2(b) also depicts the luminance of the corresponding background within the mask is also depicted. Figure 2(c) shows the subtraction between the two. To represent negative value in Fig. 2(c), the gray level is shifted up by 128. Thus, for a pixel with gray level less than 128, the corresponding pixel in the input image is

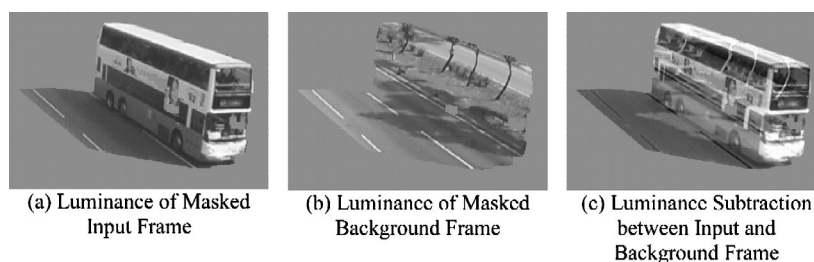


Fig. 2 Cast shadow luminance observation.

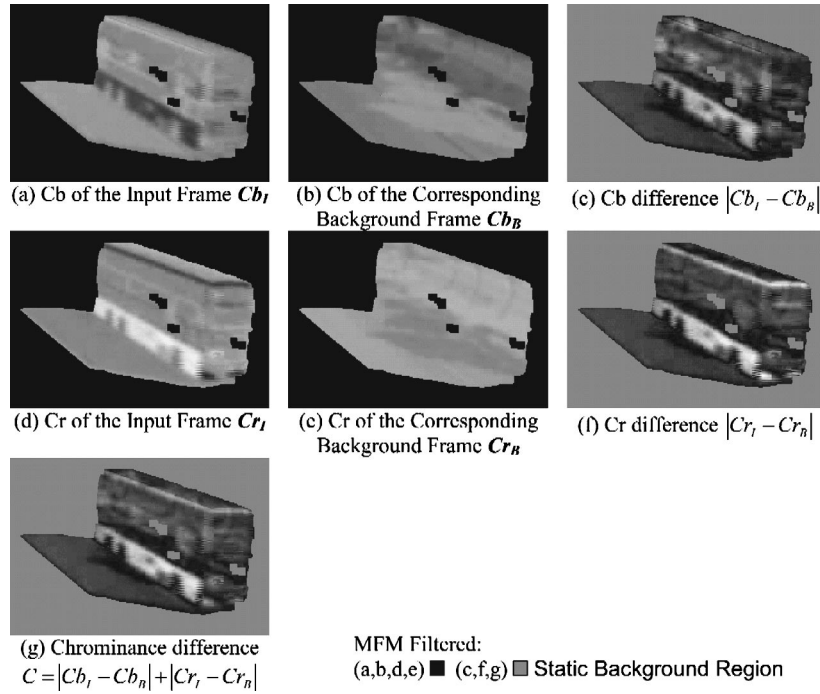


Fig. 3 Cast shadow chrominance observation.

lower than the background image. We can observe that the luminance of the input image is always lower than the background image in the cast shadow region.

Observation 2. *The chrominance values of the cast shadow pixels are identical or only slightly different from those of the corresponding pixels in the background image.*

To illustrate this, the chrominance feature is depicted in Fig. 3. In this paper, the color model YCbCr is used to separate the luminance and chrominance components of the images. Other invariant color models, such as $c_1c_2c_3$, can potentially be employed¹⁰ to separate the chrominance component from the luminance component. Figures 3(a) and 3(b) show the Cb representation of the input and background images. The Cb absolute difference between input and background images is depicted in Fig. 3(c). Similarly, the Cr representation of the input and background images and the Cr absolute difference are depicted in Figs. 3(d) and 3(f). The sum of the Cb and Cr differences is depicted in Fig. 3(g). For a typical white light source or sunlight, a decrease in illumination for the shadow region will cause only a slight change in chrominance from the correspond-

ing background. As observed in Fig. 3, there is only a slight change at the cast shadow region but a large change at the lower part of the bus, which is orange in color.

Observation 3. *The difference in gradient density values of the cast shadow pixels and the corresponding background pixels is relatively low. The difference in gradient density values between the vehicle pixels and the corresponding background pixels is relatively high.*

Let $e(x,y)$ be the magnitude of the gradient response at pixel location (x,y) as given by

$$e(x,y) = |\text{grad } l(x,y)| = \left[\left(\frac{\partial l}{\partial x} \right)^2 + \left(\frac{\partial l}{\partial y} \right)^2 \right]^{1/2}, \quad (1)$$

where $l(x,y)$ is the luminance value at pixel location (x,y) , and $g(x,y)$ is the average of $e(x,y)$ over a spatial window area as given by

$$g(x,y) = \frac{1}{(2N+1)^2} \sum_{k=x-N}^{x+N} \sum_{l=y-N}^{y+N} e(k,l), \quad (2)$$

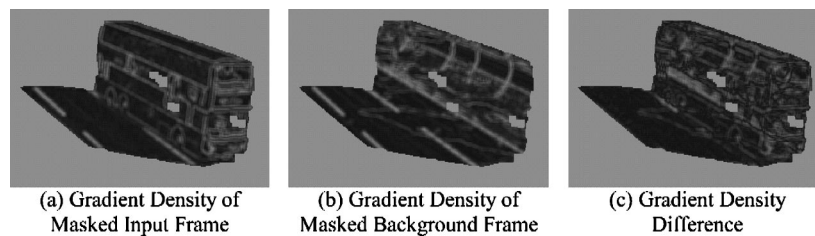
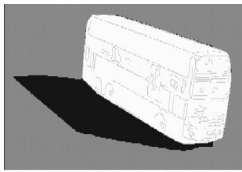


Fig. 4 Cast shadow gradient density observation.



(■ Static Background Region, □ Object Region, ■ Cast Shadow Region)

Fig. 5 Vehicle convexity observation.

where $(2N+1)$ is the size of the window.

Figure 4 illustrates the gradient density change at the cast shadow area. As shown in Eqs. (1) and (2), gradient density is defined as the magnitude of the gradient averaged over its neighboring pixels. Figures 4(a) and 4(b) depict the gradient densities of the input image and the background image. The absolute difference between them is shown in Fig. 4(c). As observed in Fig. 4(c), there is no significant gradient density difference in the cast shadow region. However, in the vehicle region, the gradient density difference between the input and background images varies significantly. Since the shadow is cast on the background region, the original gradient property of the background is not affected too much by the shadow. However, the gradient distribution of the vehicle is mostly different from the background. Therefore, most pixels in the vehicle region are likely to have large gradient density difference between the input and background images.

Observation 4. *The vehicle is approximately bounded by its convex object mask. The cast shadow is always an extension of this object mask.*

Figure 5 shows the three regions including the background, object, and cast shadow of the input image. For the bus shown in Fig. 5, the convex-hull can accurately define the object region. Hence, the remaining foreground region can be classified as the cast shadow region. Theoretically, each vehicle can be approximately bounded by the convex-hull of the vehicle edge pixels. The shadow is then defined as the remaining region within the foreground mask.

3.2 Proposed Method

Our proposed method, as shown in Fig. 6, aims to extract the moving vehicles without the cast shadows from the sta-

tionary background in an image sequence. It is mainly divided into two parts: the SCS computation and moving cast shadow detection.

Based on the observations of cast shadow as discussed in the last section, we transform the input image, the background image, and the MFM into a SCS map in the SCS computation, in which the input image is subtracted from the background image in the luminance, chrominance, and gradient density domains. By mapping through various shadow score functions, these shadow scores in different domains for each pixel are combined and transformed into an overall SCS, which indicates the likelihood of the pixel belonging to the cast shadow region. Based on the SCS map, the object mask is determined in the moving cast shadow detection. In this detection algorithm, we obtain the edge pixels that belong to the vehicle through a threshold filtering by their shadow confidence scores. The convex-hull of these vehicle edges is then determined and is used to define the object mask, whereas the remaining pixels in the MFM are classified as the shadow region.

4 SCS Computation

Let the input image and the background image be described, respectively, as

$$I_i(x,y) = \begin{Bmatrix} l_{I,i}(x,y) \\ cb_{I,i}(x,y), cr_{I,i}(x,y) \\ g_{I,i}(x,y) \end{Bmatrix}, \quad (3)$$

$$B_i(x,y) = \begin{Bmatrix} l_{B,i}(x,y) \\ cb_{B,i}(x,y), cr_{B,i}(x,y) \\ g_{B,i}(x,y) \end{Bmatrix}, \quad (4)$$

where $x=0, \dots, W-1$, $y=0, \dots, H-1$, i is the frame number, W is the width of the image, H is the height of the image, $l_{I,i}(x,y)$ is the luminance at pixel (x,y) , $cb_{I,i}(x,y)$ and $cr_{I,i}(x,y)$ are the chrominance values at pixel (x,y) , and $g_{I,i}(x,y)$ is the gradient density at pixel (x,y) in the input image. In addition, let the MFM $M_i(x,y)$ be defined as

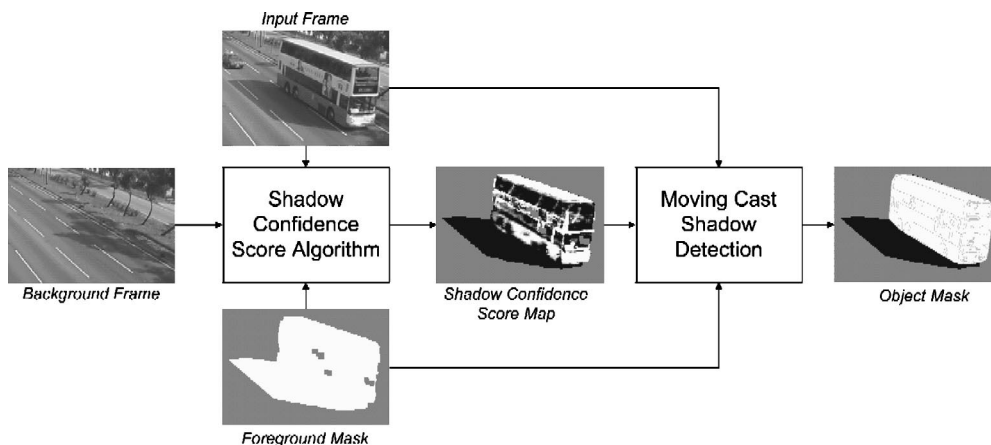


Fig. 6 Proposed method.

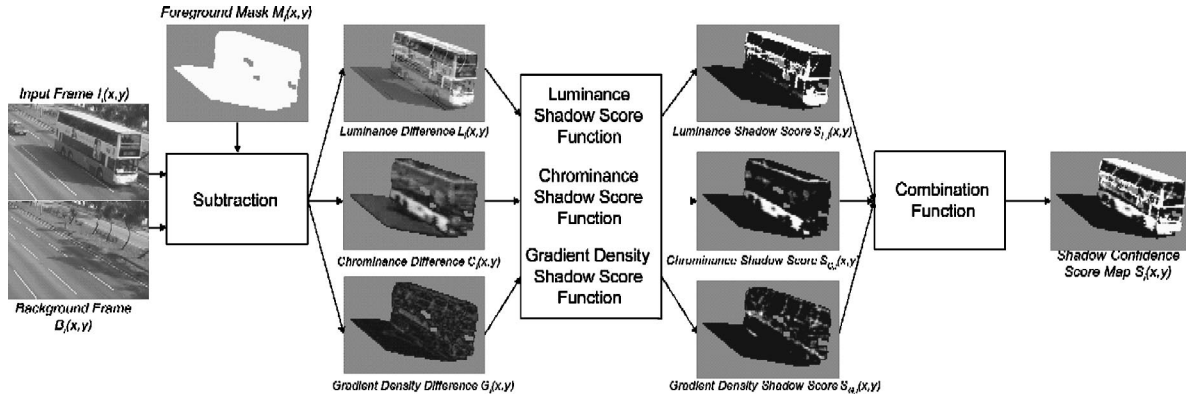


Fig. 7 SCS computation.

$$M_i(x,y) = \begin{cases} 1 & |I_i(x,y) - B_i(x,y)| > T_{BG} \\ 0, & \text{otherwise,} \end{cases} \quad (5)$$

where T_{BG} is the threshold used in background subtraction.

To indicate whether a pixel should be classified as cast shadow, an SCS $S_i(x,y)$ is defined. If the pixel is likely to be part of a cast shadow, a high $S_i(x,y)$ will be given to that pixel. On the other hand, if the pixel is likely to be part of an object or background, a low $S_i(x,y)$ will be given. The score is a probability value ranging from 0 to 1 inclusive.

As shown in Fig. 7, the luminance, chrominance, and gradient density values for each pixel are calculated from the input and background images. To calculate the overall score $S_i(x,y)$, the three mapping functions are defined: luminance score $[S_{L,i}(x,y)]$ versus luminance difference, chrominance score $[S_{C,i}(x,y)]$ versus chrominance difference, and gradient density score $[S_{G,i}(x,y)]$ versus gradient density difference. Then, the combined SCS $[S_i(x,y)]$ is computed by combining these three individual scores.

4.1 Luminance Score $S_{L,i}(x,y)$

The luminance score $S_{L,i}(x,y)$ can be defined as follows: let $L_i(x,y)$ be the luminance difference between the i 'th input image and the i 'th background image at location (x,y) as given by

$$L_i(x,y) = I_{L,i}(x,y) - I_{B,i}(x,y), \quad \forall(x,y) \text{ where } M_i(x,y) = 1, \quad (6)$$

$$S_{L,i}(x,y) = \begin{cases} 1, & L_i(x,y) \leq 0 \\ [T_L - L_i(x,y)]/T_L, & 0 < L_i(x,y) < T_L \\ 0, & L_i(x,y) \geq T_L. \end{cases} \quad (7)$$

The mapping function of luminance score $[S_{L,i}(x,y)]$ against luminance difference $[L_i(x,y)]$ is depicted in Fig. 8, where T_L is a predefined parameter to accommodate the acquisition noise in luminance domain. As discussed in observation 1, the luminance values of the pixels in the input image is lower than that of the corresponding pixels in the background image at the shadow. Therefore, a pixel with a

negative luminance difference value means that the cast shadow criterion is satisfied and it is likely to be a cast shadow. On the other hand, if the luminance value of a pixel in the input image is higher than that of the corresponding pixel in the background image (positive luminance difference value), it does not satisfy the shadow criterion and $S_{L,i}(x,y)$ of the pixel tends to zero. For $L_i(x,y)$ between 0 and T_L , a linear mapping from 0 to 1 is chosen to provide a smooth transition.

4.2 Chrominance Score $S_{C,i}(x,y)$

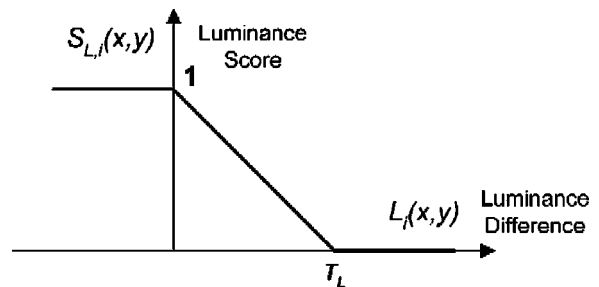
The chrominance score $S_{C,i}(x,y)$ can be defined as follows: let $C_i(x,y)$ be the chrominance difference between the i 'th input image and the i 'th background image at location (x,y) as given by

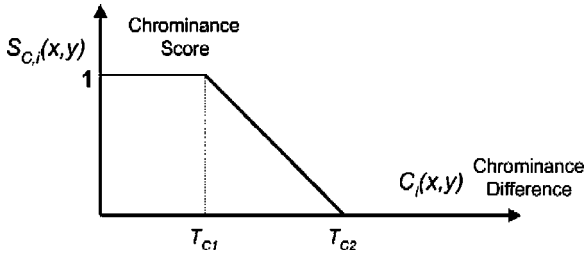
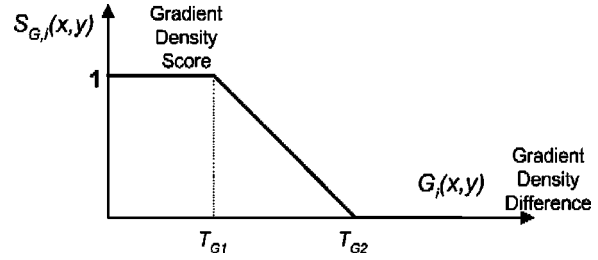
$$C_i(x,y) = |cb_{L,i}(x,y) - cb_{B,i}(x,y)| + |cr_{L,i}(x,y) - cr_{B,i}(x,y)|, \quad \forall(x,y) \text{ where } M_i(x,y) = 1, \quad (8)$$

$$S_{C,i}(x,y)$$

$$= \begin{cases} 1, & C_i(x,y) \leq T_{C1} \\ [T_{C2} - C_i(x,y)]/(T_{C2} - T_{C1}), & T_{C1} < C_i(x,y) < T_{C2} \\ 0, & C_i(x,y) \geq T_{C2}. \end{cases} \quad (9)$$

The mapping function of chrominance score $[S_{C,i}(x,y)]$ against chrominance difference $[C_i(x,y)]$ is depicted in Fig. 9, where T_{C1} and T_{C2} are predefined parameters to

Fig. 8 Function $S_{L,i}(x,y)$ versus $L_i(x,y)$.


 Fig. 9 Function $S_{C,i}(x,y)$ versus $C_i(x,y)$.

 Fig. 10 Function $S_{G,i}(x,y)$ versus $G_i(x,y)$.

accommodate the tolerance to acquisition noise in the chrominance domain. As discussed in observation 2, the chrominance value of a pixel in the input image is approximately the same as that of the corresponding pixel in the background image at the cast shadow region. Thus, we have observed that the change will occur only in the luminance dimension and there should be a very small change or no change in chrominance level. Therefore, for $C_i(x,y)$ less than T_{C1} , $S_{C,i}(x,y)$ is set to 1 (high score) since it satisfies the shadow criteria (small change) in the chrominance domain. For $C_i(x,y)$ larger than T_{C1} , $S_{C,i}(x,y)$ is set to 0 because of the large change in the chrominance domain. Smooth transition from 0 to 1 is implemented for the rest of the range of chrominance difference. For simplicity, $T_{C2} = 2 \times T_{C1}$ is chosen in this paper.

4.3 Gradient Density Score $S_{G,i}(x,y)$

The gradient density score $S_{G,i}(x,y)$ can be defined as follows: let $G_i(x,y)$ be the gradient density difference between the i 'th input image and the i 'th background image at location (x, y) as given by

$$G_i(x,y) = g_{L,i}(x,y) - g_{B,i}(x,y),$$

$$\forall (x,y) \text{ where } M_i(x,y) = 1, \quad (10)$$

$$S_{G,i}(x,y) = \begin{cases} 1, & G_i(x,y) \leq T_{G1} \\ [T_{G2} - G_i(x,y)] / (T_{G2} - T_{G1}), & T_{G1} < G_i(x,y) < T_{G2} \\ 0, & G_i(x,y) \geq T_{G2}. \end{cases} \quad (11)$$

Gradient densities $g_{L,i}(x,y)$ and $g_{B,i}(x,y)$ are the averages of the gradient magnitude over a spatial window area in the input and background images. The mapping function of gradient density score $[S_{G,i}(x,y)]$ against gradient density difference $[G_i(x,y)]$ is depicted in Fig. 10, where T_{G1} and T_{G2} are predefined parameters. As defined in observation 3, after subtraction of gradient density values of the corresponding pixels in the input and background images, the gradient density values are mostly canceled out in the cast shadow region. However, in the object region, there is significant difference between the input and background images in gradient density. Therefore, a pixel with small gradient density difference value is more likely to be part of the shadow and $S_{G,i}(x,y)$ is set to 1. For a pixel with high gradient density difference, it is likely to be part of an ob-

ject and $S_{G,i}(x,y)$ is set to 0. Smooth transition from 0 to 1 is implemented for the rest of the range. For simplicity, $T_{G2} = 2 \times T_{G1}$ is chosen in this paper.

4.4 Combined SCS $S_i(x,y)$

After the three scores, $S_{L,i}(x,y)$, $S_{C,i}(x,y)$, and $S_{G,i}(x,y)$, are calculated in the three difference domains, the total $S_i(x,y)$ is computed by combining them:

$$S_i(x,y) = \mathfrak{J}[S_{L,i}(x,y), S_{C,i}(x,y), S_{G,i}(x,y)]. \quad (12)$$

Since each domain is a necessary requirement for the pixel to be classified as cast shadow, hence function \mathfrak{J} is chosen to be a logical AND function. Mathematically, direct multiplication of $S_{L,i}(x,y)$, $S_{C,i}(x,y)$, and $S_{G,i}(x,y)$ can be used, such that, $S_i(x,y) = S_{L,i}(x,y) \times S_{C,i}(x,y) \times S_{G,i}(x,y)$.

5 Moving Cast Shadow Detection

In the detection part, the cast shadow is separated from the object based on the shadow confidence score $S_i(x,y)$ and the object edge pixels of the foreground masked input image (see Fig. 11). All the pixels with significant gradient values are detected using the edge detector within the MFM. These pixels are denoted as $E_i(x,y)$. For pixels with high gradient values, a thresholding T_s is applied to discard those with high shadow confidence score levels. This test also eliminates noise and edge pixels that do not belong to the vehicle. If a high gradient pixel has high shadow confidence score value, it will be discarded; otherwise, it will be retained:

$$E_i(x,y) = \begin{cases} 0 & \text{(discarded) for } S_i(x,y) \geq T_s \\ 1 & \text{(retained) for } S_i(x,y) < T_s. \end{cases} \quad (13)$$

As discussed in observation 4, the vehicle can be segmented out in the foreground mask by bounding convex hull on the vehicle edge pixels. The remaining pixels within the foreground mask are then classified as cast shadow pixels.

6 Results and Discussions

In this section, 100 test samples are used to systematically determine the parameters T_L , T_{C1} , T_{G1} , and T_s . The results of a test case are used to illustrate how the proposed method works step by step. Finally, the effect behavior of our method due to environment factors and vehicle factors

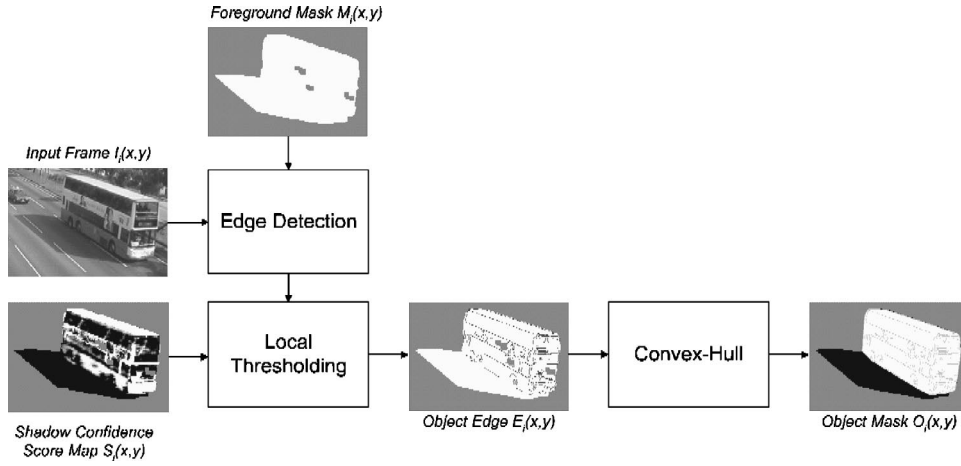


Fig. 11 Moving cast shadow detection.

are detailed. The strengths and limitations of the proposed method are objectively analyzed based on these results.

6.1 Evaluation Conditions

To test the effectiveness and robustness of the proposed method, some typical outdoor traffic image sequences on different roads in Hong Kong were captured and tested. These image sequences were all captured by a Panasonic DV camera, and transferred to a PC by a Canopus DV-Rex through IEEE-1394. The image sequences were captured under different lighting conditions, including sunny, cloudy, and different time of the day, with the camera position either overhead or by the roadside. The camera was set to a good coverage of the road lengthwise. Among these image sequences, different vehicle samples were selected under two different vehicle factors. These factors include vehicle types, such as small and large vehicles, and vehicle colors. For evaluation purposes, reference object masks of the vehicles without their cast shadows are required for the calculation of classification error. A reference object mask for each vehicle sample is defined manually by combining the visual observation on the images and the knowledge about the vehicle:

$$R_i(x,y) = \begin{cases} 1 & \text{if } (x,y) \text{ is in vehicle region defined} \\ & \text{in the reference object mask} \\ 0 & \text{otherwise,} \end{cases} \quad (14)$$

$$O_i(x,y) = \begin{cases} 1 & \text{if } (x,y) \text{ is in vehicle region in the} \\ & \text{object mask obtained by proposed method} \\ 0 & \text{otherwise,} \end{cases} \quad (15)$$

$$\text{FNM}_i(x,y) = \begin{cases} 1 & \text{if } R_i(x,y) = 1 \text{ and } O_i(x,y) = 0 \\ 0 & \text{otherwise,} \end{cases} \quad (16)$$

$$\text{FPM}_i(x,y) = \begin{cases} 1 & \text{if } R_i(x,y) = 0 \text{ and } O_i(x,y) = 1 \\ 0 & \text{otherwise,} \end{cases} \quad (17)$$

$$\text{FN}_i = \frac{\sum_{k=0}^{xsize-1} \sum_{l=0}^{ysize-1} \text{FNM}_i(k,l)}{\left[\sum_{m=0}^{xsize-1} \sum_{n=0}^{ysize-1} R_i(m,n) \right]}, \quad (18)$$

$$\text{FP}_i = \frac{\sum_{k=0}^{xsize-1} \sum_{l=0}^{ysize-1} \text{FPM}_i(k,l)}{\left[\sum_{m=0}^{xsize-1} \sum_{n=0}^{ysize-1} R_i(m,n) \right]}, \quad (19)$$

where

$R_i(x,y)$ = reference object mask of the i 'th frame at pixel (x, y)

$O_i(x,y)$ = object mask of the i 'th frame at pixel (x,y)

$\text{FNM}_i(x,y)$ = false negative map of the i 'th frame at pixel (x,y)

$\text{FPM}_i(x,y)$ = false positive map of the i 'th frame at pixel (x, y)

FN_i = false negative error rate of the i 'th frame

FP_i = false positive error rate of the i 'th frame

By comparing the object mask computed by our method and the reference object mask, two error rates, false positive (FP_i) and false negative (FN_i) as defined in Eqs. (18) and (19), can be calculated. Note that the FP_i error rate is defined as the ratio of the number of nonvehicle pixels that are incorrectly classified as vehicle pixels to the total number of vehicle pixels. The FN_i error rate is defined as the ratio of the number of vehicle pixels that are incorrectly

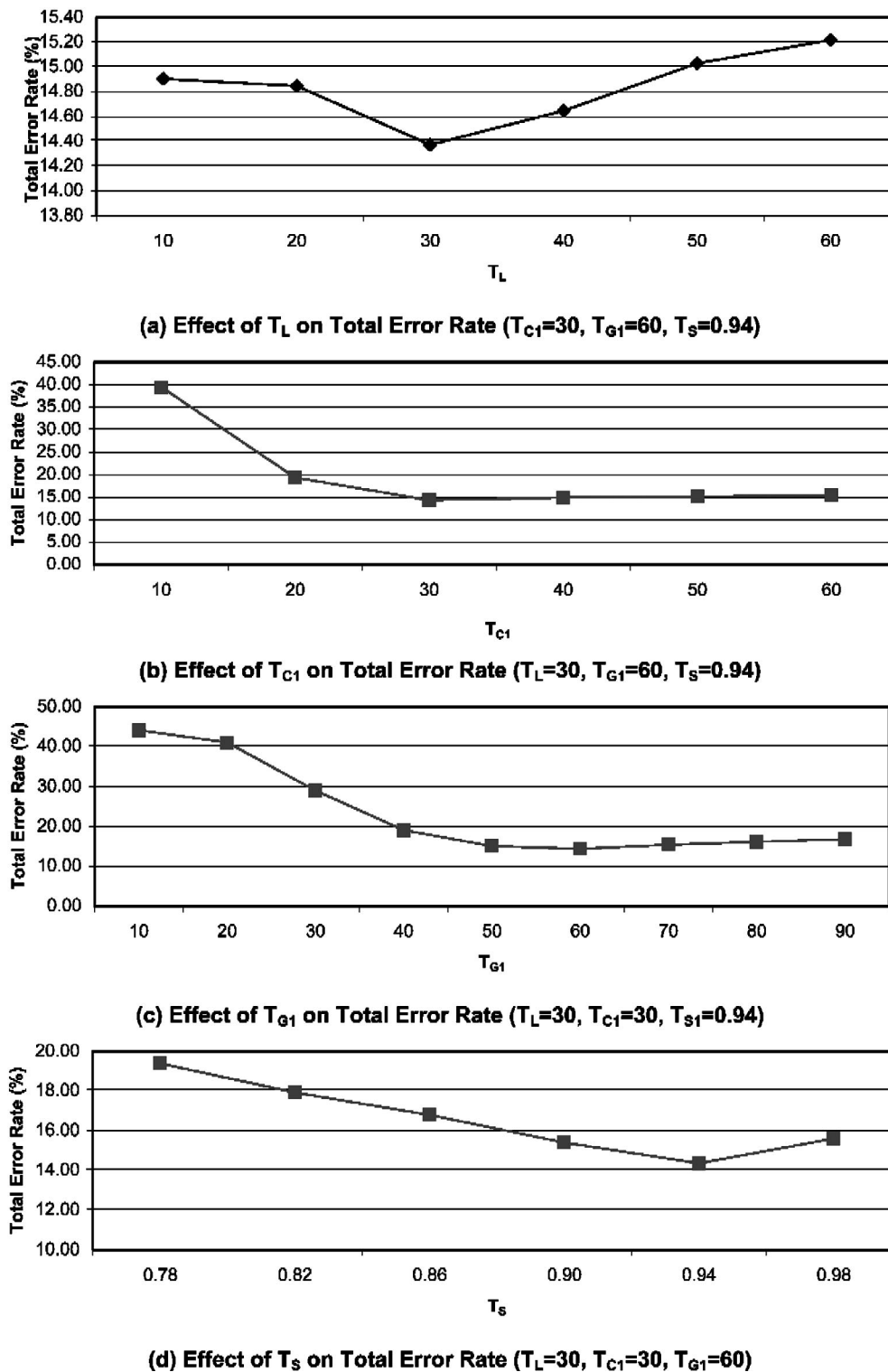


Fig. 12 Impact of the performance by varying various parameters setting.

classified as nonvehicle pixels to the total number of vehicle pixels. To objectively evaluate the error performance of the method, the total error rate is taken, which is the mean of the FP_i and FN_i error rates of the 100 test cases.

6.2 Selection of T_L , T_{C1} , T_{G1} , and T_S

In this section, we explicitly evaluate the error performance of the proposed method due to the different parameter set-

tings. In our proposed method, there are four parameters: T_L , the parameter for luminance score $[S_{L,i}(x,y)]$ function; T_{C1} , the parameter for chrominance score $[S_{C,i}(x,y)]$ function; T_{G1} , the parameter for gradient score $[S_{G,i}(x,y)]$ function; and T_S , the threshold to retain the object edge pixels. As there is a lack of a theoretical approach to determine the optimal parameter setting for this kind of problem,

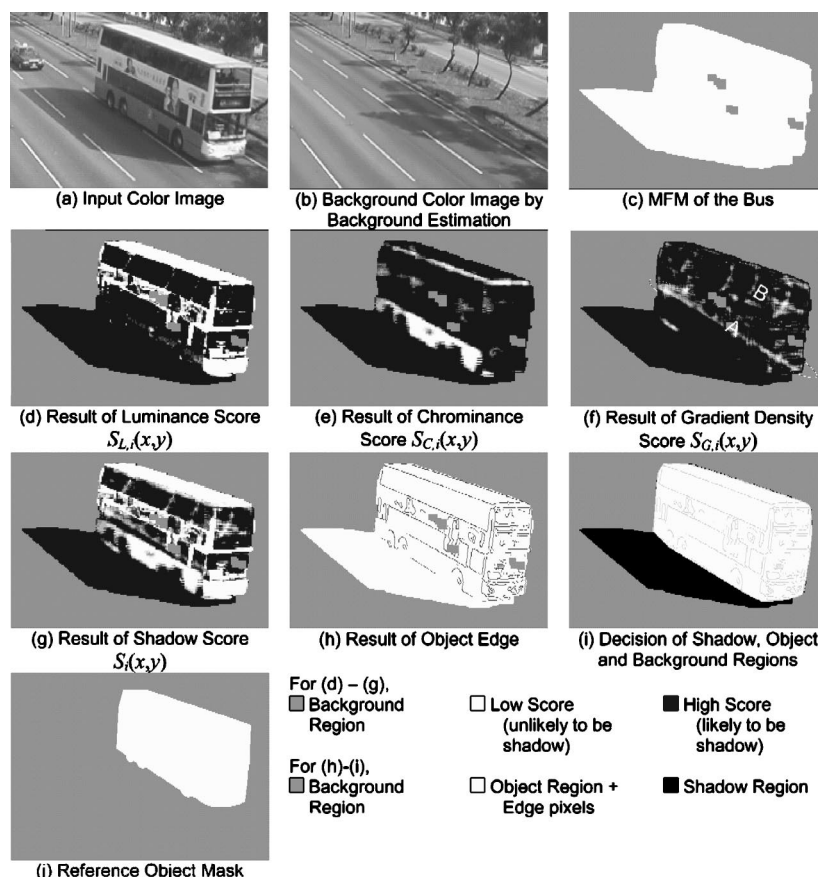


Fig. 13 Results for a bus.

we chose to embark on an experimental approach. We have performed a brute force search over the Cartesian combinations of parameter settings. The near optimal setting is determined as the setting that achieved the lowest error rate averaged over the test images.

To illustrate the trends of the error performance due to different parameter settings, a subset of test results is shown in Fig. 12. The results are illustrated in four graphs. For each graph, one of the parameters is varied while the others are fixed. Figure 12(a) shows the effect on the total error rate of varying T_L . A total error rate ranging from 14.37 to 15.22% was obtained for $T_L=10$ to 60, which indicates that the selection of T_L setting will unlikely affect the error performance much. However, given the minimum at $T_L=30$, this seems to offer a good choice. Figure 12(b)

shows the effect on the total error rate of varying T_{C1} . The total error rate hits the high of 39.50% at $T_{C1}=10$, then settles down to approximately 15% for T_{C1} beyond 30. The high error rate reported in the low T_{C1} setting is mainly caused by the misclassification of shadow region as object region. As mentioned in observation 2, there is a chrominance shift at the cast shadow region. A larger T_{C1} will enable our method to accept higher tolerance in the chrominance shift. Figure 12(c) shows the effect on the total error rate of varying T_{G1} is depicted. Similar to varying T_{C1} , varying T_{G1} also starts with high error rate of 44.12% at $T_{G1}=10$ and ends at approximately 15% at $T_{G1}=50$ to 90. There is a minimum at $T_{G1}=60$, but the difference is very small. It is also caused by incorrectly classifying shadow

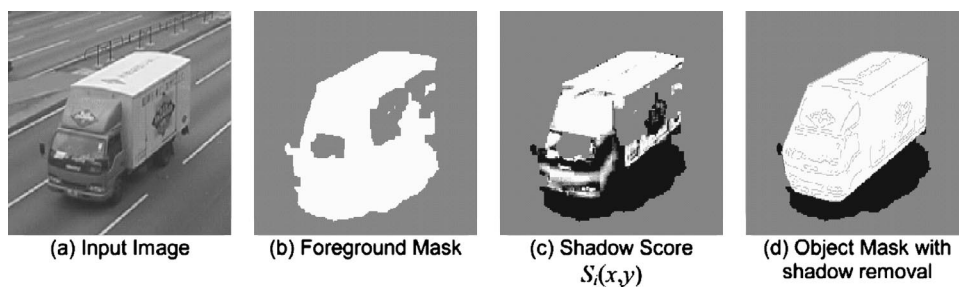


Fig. 14 Results for a truck image taken on a cloudy day.

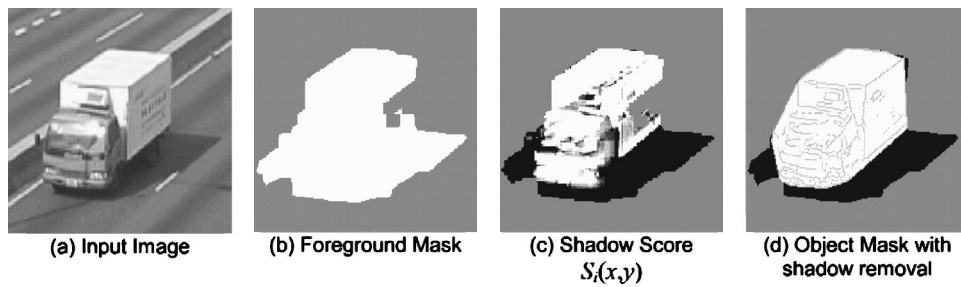


Fig. 15 Results for a truck image taken on a sunny day.

region as object region. The high error rate reported for a low T_{G1} setting is because the pixels on the shadow border are misclassified as vehicle edge pixels under sunny condition. Figure 12(d) shows the effect on the total error rate of varying T_S . The small error rate range of 14.37 to 19.40% indicates a mild effect of T_S on the error performance. A minimum is detected at $T_S=0.94$, giving approximately 14% error.

By setting $T_L=30$, $T_{C1}=30$, $T_{G1}=60$, and $T_S=0.94$, the mean of total error rate averaged over 100 samples is 13.95% and the standard deviation is 7.65%. The best error rate is 3.23% and the worst is 33.46%.

6.3 Test Case

Figure 13 shows one of the test cases. The settings $T_L=30$, $T_{C1}=30$, $T_{G1}=60$, and $T_S=0.94$ are used. The output images at different stages of the proposed method are described step by step.

An image of a white and orange double-deck bus, which was captured on a sunny day, is shown in Fig. 13(a). In Fig. 13(b), the background color image was generated by the background estimation method of Ref. 19. After subtracting the background frame from the input frame, the MFM after the morphological closing transform is shown in Fig. 13(c), where the white region is the foreground and the gray region is the background region. In Figs. 13(d), 13(e), and 13(f), the foreground masked results for $S_{L,i}(x,y)$, $S_{C,i}(x,y)$, and $S_{G,i}(x,y)$ are shown, respectively. In Fig. 13(d), some parts of the bus, such as the windows and the orange part, are recognized as shadow since they have a luminance level similar to the background image. Therefore, luminance can provide only a limited indication of the SCS. In Fig. 13(e), based on the chrominance value, the orange part of the bus is clearly classified as nonshadow because its chrominance value is significantly different from the background. In Fig. 13(f), the pixels with large gradient density difference are clearly marked as non-shadow. They include the blocked road lane mark (region A) and the tree branches (region B). By combining the results for $S_{L,i}(x,y)$, $S_{C,i}(x,y)$, and $S_{G,i}(x,y)$, the total SCS $S_i(x,y)$ is shown in Fig. 13(g), in which only the cast shadow region has a high $S_i(x,y)$ value (shown in black) while the vehicle region is surrounded by low $S_i(x,y)$ values. However, the windscreen and the windows were incorrectly interpreted as shadow since they exhibit shadow-like characteristics. In Fig. 13(h), the edge pixels that have shadow scores lower than threshold T_S are retained as object edge pixel. In Fig. 13(i), the background, object, and

shadow are shown in gray, white, and black, respectively, after performing the convex-hull on the object edge pixels. It is important to use the convex-hull to bound the vehicle to recover the inner misclassified region. Comparing this with the reference object mask, as depicted in Fig. 13(j), the false positive error rate is 2.18% and the false negative error rate is 1.05%.

6.4 Effect of Environment Factors

6.4.1 Lighting conditions

It is important that the method works well under different lighting conditions. We tested our method under two different lighting conditions: cloudy and sunny. On a cloudy day, as shown in Fig. 14(a), the cast shadow is poorly defined since it is mainly caused by ambient light. However, the computed foreground mask depicted in Fig. 14(b) clearly includes a large region as shadow. The SCS is shown in Fig. 14(c). In Fig. 14(d), by the proposed method, the object mask is successfully segmented by removing the shadow region. However, the right-hand mirror and the lower part of the tires are misclassified as shadow regions. In Fig. 15, a similar truck image was taken on a sunny day. The shadow is clearly defined and exhibits a significant change in luminance. Similarly, as shown in Fig. 15(d), the lower part of the tires is misclassified as shadow region. Moreover, we noted that there is a shadow region correctly detected next to the right front of the truck. This shadow is cast on the concrete structure of the road but not on the road surface. Potentially, our proposed method is capable of detecting multiple cast shadows.

In our test cases, there are vehicle samples taken under different degree of cloudy condition and others taken under different degree of sunny condition. As shown in Table 1, the error rates are both approximately equal to 14% under

Table 1 Effect on the error performance under different lighting conditions.

	Cloudy	Sunny
Vehicle samples tested	20	79
Total error rate		
Minimum	5.58%	3.23%
Maximum	33.25%	33.46%
Average	14.01%	13.93%
Standard deviation	7.34%	7.77%

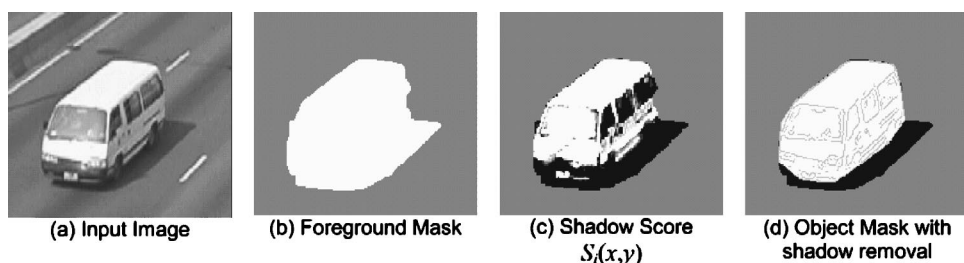


Fig. 16 Results for a van captured by a roadside camera.

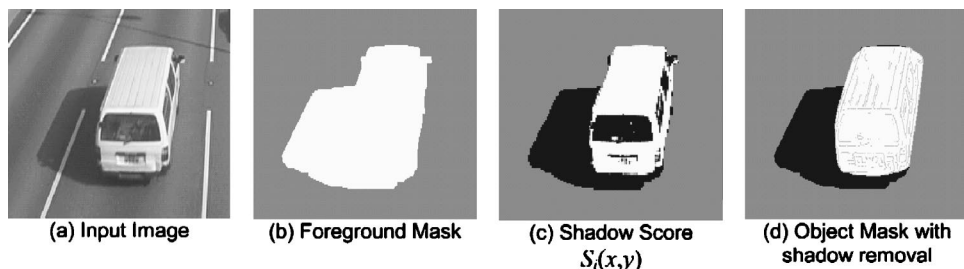


Fig. 17 Results for a van captured by an overhead camera.

different lighting conditions. From the results, it seems that our method copes with the cloudy and sunny lighting conditions reasonably well.

6.4.2 Effect of viewing angles

Traffic surveillance cameras are mostly installed by the roadside or overhead. In Figs. 16 and 17, the results for a van captured by a roadside camera and an overhead camera are shown. In both cases, most of the van is successfully segmented with its shadow removed. In Fig. 16(d), a small part of the dark gray bumper is misclassified as shadow region since it exhibits shadow characteristics. In Fig. 17(a), there is a white road lane mark partially covered by the cast shadow of the van. As shown in Fig. 17(d), this lane mark does not affect our method and is correctly classified as shadow region. As shown in Table 2, under different viewing angle, the total error rates are approximately equal to 14%. From the results, we find that the error performance of our method is not sensitive to the viewing angle.

Table 2 Effect on the performance under different viewing angles.

	Roadside	Overhead
Vehicle samples tested	79	20
Total error rate		
Minimum	3.23%	5.55%
Maximum	33.46%	28.31%
Average	14.06%	13.49%
Standard deviation	7.90%	6.70%

6.5 Effect of Vehicle Factors

6.5.1 Vehicle types

There are many different vehicles on the road. Very broadly, a vehicle can be roughly classified as a motorbike, a small vehicle (sedan, hatch, station-wagon, taxi), and a large vehicle (van, truck, minibuss, bus). In Figs. 18–20, the results of applying our method to a motorbike, a sedan, and a bus are depicted, respectively.

In Fig. 18(d), the computed object mask is under segmented since the motorbike does not have a cuboid-like vehicle body. In Fig. 19(d), since our proposed method uses the convex-hull to define the region border, small vehicles, such as a sedan, are not convex objects. Besides, the vehicle side mirror further increases the false positive error. As shown in Fig. 20(d), our method performs better on large vehicles since they are mostly cuboid and can be described by a convex-hull. In Fig. 20(d), there are holes in the shadow region since the input image is very similar to the background image at these holes. Since a vehicle without a cast shadow is the objective of our method, there is no impact of having holes in the shadow region.

Table 3 Effect on the performance for different vehicle class.

	Motorbike	Small Vehicle	Large Vehicle
Vehicle samples tested	7	45	47
Total error rate			
Minimum	15.22%	5.58%	3.23%
Maximum	30.42%	33.46%	33.25%
Average	22.80%	15.07%	11.55%
Standard deviation	5.84%	7.43%	6.94%

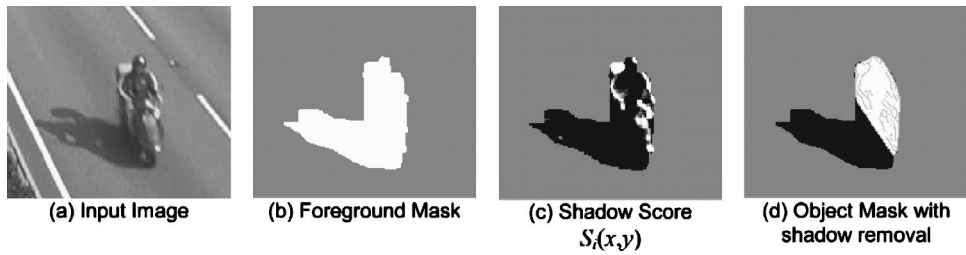


Fig. 18 Results for a motorbike.

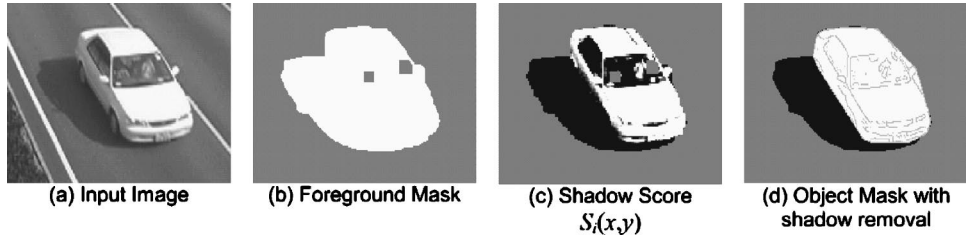


Fig. 19 Results for a sedan (small vehicle).

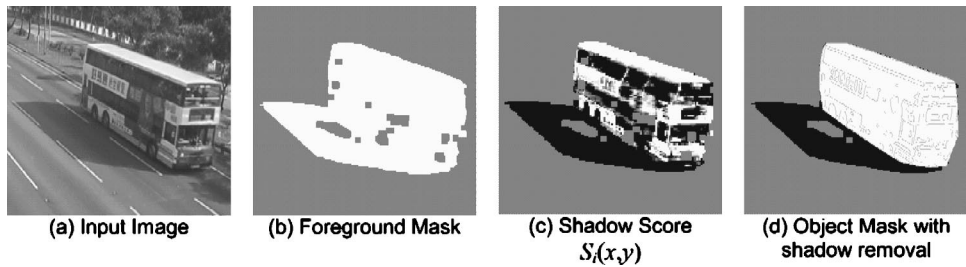


Fig. 20 Results for a bus (large vehicle).

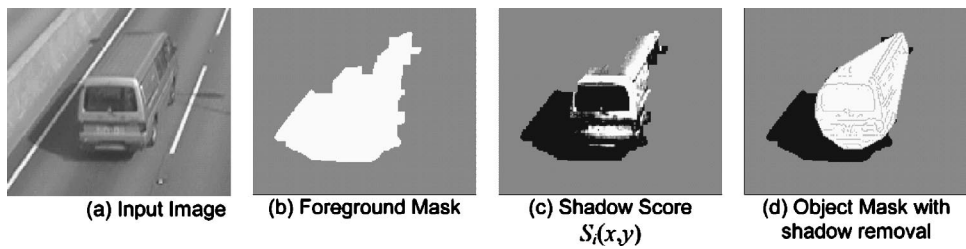


Fig. 21 Results for a gray van (similar to road surface color).

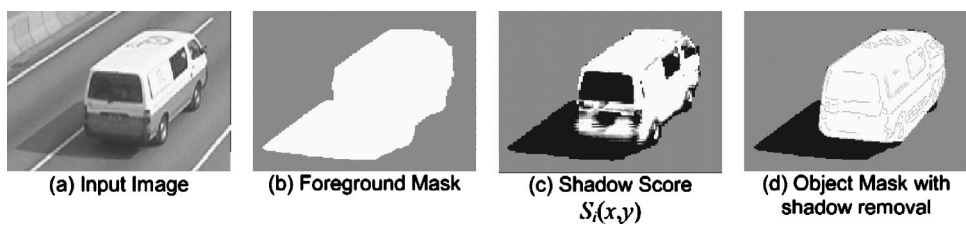


Fig. 22 Results for a white/blue van (not similar to road surface color).

Table 4 Effect on the performance for different vehicle color.

	Similar	Nonsimilar
Vehicle samples tested	8	91
Total error rate		
Minimum	9.68%	3.23%
Maximum	33.46%	33.25%
Average	18.00%	13.59%
Standard deviation	8.03%	7.55%

Our argument is further supported in Table 3, in which our method achieves the lowest error rate for a large vehicle and the highest error rate for a motorbike.

6.5.2 Vehicle colors

We roughly divide vehicle colors into two classes: similar to road surface color and not similar to road surface color. In Figs. 21 and 22, the results for a gray van and a colored van are shown. A large part of the gray van is not correctly classified since this part is not successfully extracted by background subtraction. Our observation is further confirmed in Table 4. Those vehicles that have colors similar to the road surface color will likely be undersegmented.

7 Conclusions

We presented a cast shadow detection method for estimating a vehicle outline. It can effectively separate cast shadow from the vehicle under different environment and vehicle factors. In our method, an SCS is computed in three different domains: luminance, chrominance, and gradient density. Based on the shadow confidence score and the object edge pixels, the cast shadow is separated from the vehicle using convex-hull within the foreground mask.

We tested our proposed method on 100 different vehicle samples captured in typical outdoor traffic scenes. From our results and analysis of various vehicle samples, the proposed method can successfully separate the cast shadow and a moving vehicle. With settings of $T_L=30$, $T_{C1}=30$, $T_{G1}=60$, and $T_S=0.94$, the mean of error rate is 13.95% and the standard deviation is 7.65%. The lowest error rate is around 3% for large vehicles. By observing the effect of the error performance of our method under varying environment and vehicle factors, including different lighting conditions, camera view angles, vehicle types, and vehicle colors, we found that our proposed method is reasonably robust for various outdoor daylight environments and vehicles.

For future research, we will further investigate the use of a more efficient search algorithm to select an appropriate parameter setting for a typical traffic scene. In addition, to further improve the robustness of our proposed method, we intend to work on a continuous sequence of images of the same vehicle. Since the vehicle is a rigid body, the extracted object masks should be consistent over the image sequence. By analyzing the change of the object mask over an image sequence, we can further remove undesired shadow regions and recover the missing vehicle regions.

References

1. N. H. C. Yung and A. H. S. Lai, "A system architecture for visual traffic surveillance," in *Proc. 5th World Congr. on Intelligent Transport Systems*, ITS Congress Association (1998).
2. N. Hoose, *Computer Image Processing in Traffic Engineering*, Research Studies Press, London (1991).
3. A. Makarov, "Comparison of background extraction based intrusion detection algorithm," in *Proc. Int. Conf. on Image Processing*, Vol. 1, pp. 521–524 (1996).
4. A. Neri, S. Colonnese, and G. Russo, "Video sequence segmentation for object-based coders using higher order statistics," in *Proc. Int. Symp. on Circuits and Systems*, Vol. 2, pp. 1245–1248 (1997).
5. N. H. C. Yung and A. H. S. Lai, "Detection of vehicle occlusion using a generalized deformable model," in *Proc. IEEE Int. Symp. on Circuits and Systems*, *Proc. IEEE* 4, pp. 154–157 (1998).
6. J. Malik, S. Russel, J. Weber, T. Huang, and D. Koller, "A machine vision based surveillance system for California roads," PATH Project MOU-83 Final Report, University of California (1994).
7. J. M. Scanlan, D. M. Chabries, and R. W. Christiansen, "A shadow detection and removal algorithm for 2-D images," in *Proc. Int. Conf. on Acoustics, Speech, and Signal Processing*, pp. 2057–2060 (1990).
8. C. Jiang, and M. O. Ward, "Shadow identification," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pp. 606–612 (1992).
9. G. Funka-Lea and R. Bajcsy, "Combining color and geometry for the active, visual recognition of shadows," in *Proc. Int. Conf. on Computer Vision*, pp. 203–209 (1995).
10. E. Salvador, A. Cavallaro, and T. Ebrahimi, "Shadow identification and classification using invariant color models," in *Proc. Int. Conf. Acoustics, Speech, and Signal Processing*, Vol. 3, pp. 1545–1548 (2001).
11. J. Stauder, R. Mech, and J. Ostermann, "Detection of moving cast shadows for object segmentation," *IEEE Trans. Multimedia* 1(1), 65–76 (1999).
12. M. Kilger, "A shadow handler in a video-based real-time traffic monitoring system," in *Proc. IEEE Workshop on Applications of Computer Vision*, pp. 11–18 (1992).
13. P. Gamba, M. Lilla, and A. Mecocci, "A fast algorithm for target shadow removal in monocular colour sequence," in *Proc. Int. Conf. on Image Processing*, pp. 436–439 (1997).
14. I. Mikic, P. C. Cosman, G. T. Kogut, and M. M. Trivedi, "Moving shadow and object detection in traffic scenes," in *Proc. Int. Conf. Pattern Recognition*, pp. 321–324 (2000).
15. T. Horprasert, D. Harwood, and L. S. Davis, "A statistical approach for real-time robust background subtraction and shadow detection," in *Proc. Int. Conf. Computer Vision, FRAME-RATE Workshop* (1999).
16. D. Koller, K. Daniilidis, and H. H. Nagel, "Model-based object tracking in monocular image sequences of road traffic scenes," *Int. J. Comput. Vis.* 10(3), 257–281 (1993).
17. X. Tao, M. Guo, and B. Zhang, "A neural network approach to the elimination of road shadow for outdoor mobile robot," in *Proc. Int. Conf. on Intelligent Processing Systems*, Vol. 2, pp. 1302–1306 (1997).
18. Y. Sonoda and T. Ogata, "Separation of moving objects and their shadows, and application to tracking of loci in the monitoring images," in *Proc. Int. Conf. on Signal Processing*, Vol. 2, pp. 1261–1264 (1998).
19. A. H. S. Lai and N. H. C. Yung, "A fast and accurate scoreboard algorithm for estimating stationary backgrounds in an image sequence," in *Proc. IEEE Int. Symp. on Circuits and Systems*, Vol. 4, pp. 241–244 (1998).



George S. K. Fung received his BSc degree from the Monash University, Australia, and his MPhil degree from the Hong Kong Polytechnic University. He is currently a PhD candidate in the Department of Electrical and Electronic Engineering, the University of Hong Kong. His research interests include pattern recognition, digital image processing, and intelligent transportation systems.



Nelson H. C. Yung received his BSc and PhD degrees from the University of Newcastle-Upon-Tyne while he was a Croucher Foundation Scholar in 1982 and 1985, respectively. He was a lecturer at the same university from 1985 until 1990. From 1990 to 1993, he was a senior research scientist with the Department of Defence, Australia. He joined the University of Hong Kong (HKU) in late 1993 as an associate professor. He is the founding director

of the Laboratory for Intelligent Transportation Systems Research at HKU. He has coauthored a computer vision book and has published over 100 journal and conference papers in the areas of digital image processing, parallel algorithms, visual traffic surveillance, and autonomous vehicle navigation. He is a reviewer for the *IEEE Transactions of SMC*, *Signal Processing*, *IEE Pt. G*, *SPIE Optical Engineering*, *International Journal of Vehicle Mechanics and Mobility*, *HKIE proceedings*, *Microprocessors and Microsystems*, and *Robotics and Autonomous Systems Journal*. He is a member of the advisory panel of the *ITS Strategy Review*, Transport Department, HK-SAR, and a council member of ITS-HK and chairs the Computer Division of the International Institute for Critical Infrastructures. He is a chartered electrical engineer, a member of the HKIE and IEE, and a senior member of the IEEE. He was regional secretary of the IEEE Asia-Pacific Region from 1995 to 1997. His biography has been published in *Who's Who in the World* since 1998.

expert systems for control system design, intelligent control, and intelligent transportation systems.



Andrew H. S. Lai received his BEng degree from the University of Hong Kong in 1994, his MSc degree from the University of Surrey in 1995, and his PhD degree from the University of Hong Kong in 2000. He is currently an honorary assistant professor with the University of Hong Kong and a senior research engineer with the Laboratory for Intelligent Transportation Systems Research. His research interest includes visual traffic surveillance, digital

image processing, and wireless data communication.



Grantham K. H. Pang received his PhD degree from the University of Cambridge in 1986. He was with the Department of Electrical and Computer Engineering, University of Waterloo, Canada, from 1986 to 1996 and joined the Department of Electrical and Electronic Engineering at the University of Hong Kong in 1996. Since 1988 he has published more than 120 technical papers and has authored or coauthored five books. His research interests include

machine vision for surface defect detection, optical communications,