

TOWARDS DETECTION OF MOVING CAST SHADOWS FOR VISUAL TRAFFIC SURVEILLANCE

G.S.K. FUNG, N.H.C. YUNG, G.K.H. PANG, A.H.S. LAI

The Laboratory for Intelligent Transportation Systems Research,
Department of Electrical and Electronic Engineering,
The University of Hong Kong, Pokfulam Road, Hong Kong SAR
E-mail: skfung,nyung,gpang,hslai@eee.hku.hk

Abstract

In this paper, an effective method for the detection of moving cast shadow for visual traffic surveillance is proposed. Based on the cast shadow observations in luminance, chrominance, gradient density and geometry domains, a combined probability map, called shadow confidence score, of the region belonging to the shadow is deduced. The object region is then separated from its shadow according to the score. The proposed method has been further tested on images taken under different lighting conditions (sunny and cloudy), viewing angles (roadside and overhead), and vehicle sizes (small, medium and large).

Keywords

Shadow Detection, Image Sequence Analysis, Visual Traffic Surveillance, Image Segmentation, Intelligent Transportation Systems

1 Introduction

Visual Traffic Surveillance (VTS) is one of the major components within the research of Intelligent Transportation Systems (ITS) [1]. Its main purpose is to remotely acquire traffic image sequences from roadside surveillance cameras and interpret them into traffic parameters and vehicle behavior by analyzing the image sequences. To achieve these requirements, numerous image processing algorithms, including pre-processing and post-processing algorithms, have been developed for outdoor environment [2]. Among them, segmentation algorithms that extract the concerned objects (such as moving vehicles) from the image background in an image sequence have been actively studied in recent years [3], of which background subtraction is a common approach in segmenting moving vehicles from the stationary background estimated over an image sequence [4]. 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 moving objects since shadows move in accordance with the movement of the objects. When the detected vehicles contain shadows, large errors may result with respect to the estimation of their location, dimension, speed and their total number. Although there have been numerous shadow detection methods proposed in the last decade [5-12], they all suffer from a number of limitations that make them ineffective in practical outdoor environment. Thus, we are motivated to resolve this problem of separating cast shadows from the vehicles in 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. First of all, we assume that the mask of the moving vehicle (the region covering the vehicle and the cast shadow) and an estimated static background reference image are both available together with the input image sequence. They can be used for determining the motion content of the image, and the Moving Foreground Mask (MFM) can be directly computed from the reference background and the input image. Secondly, from the MFM, we compute three score functions representing its luminance, chrominance and gradient density features. The results of the score functions are then combined to form a Shadow Confidence Score (SCS), which indicates the likelihood of shadow or otherwise. Thirdly, edge feature of the MFM is determined and classified into object edges and non-object edges using the SCS, where the object edges are then bounded by a convex hull. This convex hull denotes the vehicle, while the remaining pixels of the MFM denote the shadow. The proposed method has been tested on images taken under different lighting conditions (sunny and cloudy), viewing angles (roadside and overhead), and vehicle sizes (small, medium and large).

2 Methodology

2.1 Observations

In essence, shadows occur when objects partially or completely occlude direct light from a light source. As defined in [5], there are two parts in a shadow: the self-shadow and the cast shadow. The self-shadow is the part of object which is not illuminated by direct light. The 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 from the object. Although the formation of cast shadow depends on various environment factors, we have observed that there are four generic features of cast shadow that could be used to guide our proposed shadow detection method.

Our observations are:

Observation 1: *The luminance of the cast shadow is lower than the background.*

Observation 2: *The chrominance of the cast shadow is identical or only slightly shifted when compared with the background.*

Observation 3: *The difference in gradient density between the cast shadow and the background is relatively low. The difference in gradient density between the vehicle and background is relatively high (Gradient density is defined as the magnitude of the gradient averaged over a local area.)*

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

2.2 Proposed Method

Our proposed methodology, as shown in Figure 1, aims to extract the moving vehicles without the cast shadows from the stationary background in an image sequence. It is mainly divided into two parts: Shadow Confidence Score Computation and Moving Cast Shadow Detection.

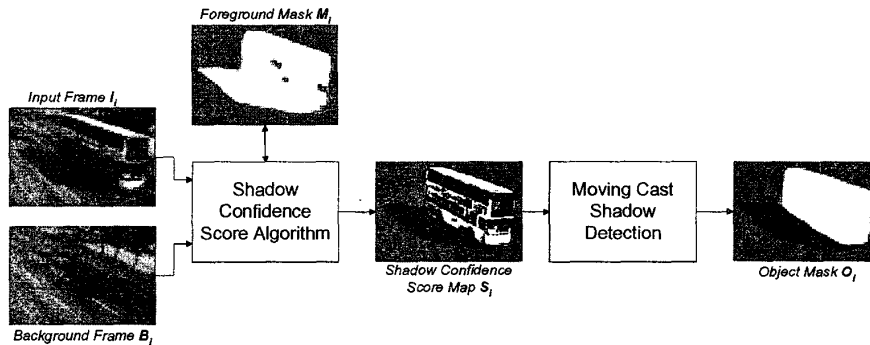


Fig. 1: Proposed Method

Based on the observations of cast shadow, we transform the Input Frame I_i , the Background Frame B_i and the Moving Foreground Mask (MFM) M_i into Shadow Confidence Score Map S_i in the Shadow Confidence Score Computation, in which, the input frame is subtracted from the background frame in the luminance, chrominance and gradient density domains. By mapping through various shadow score functions, these subtraction results are transformed into the shadow confidence scores, which provide indication of the likelihood of the pixels belonging to the cast shadow region. Based on the Shadow Confidence Score Map S_i , the Object Mask O_i is determined in the Moving Cast Shadow Detection. In this detection algorithm, we obtain the edges that belong to the vehicle through a threshold filtering by the shadow confidence score. The convex-hull of these vehicle edges is then calculated. This convex-hull of these vehicle edges defines the Object Mask O_i , whereas the remaining pixels in the MFM become the shadow region.

3 Shadow Confidence Score Computation

Let the current image and the background image be defined respectively as follows:

$$I_i(x, y) = \begin{cases} I_{I,i}(x, y) \\ cb_{I,i}(x, y), cr_{I,i}(x, y) \\ g_{I,i}(x, y) \end{cases}, \quad (1)$$

$$B_i(x, y) = \begin{cases} I_{B,i}(x, y) \\ cb_{B,i}(x, y), cr_{B,i}(x, y) \\ g_{B,i}(x, y) \end{cases}, \quad (2)$$

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, $I_{I,i}(x, y)$ 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) , $g_{I,i}(x, y)$ is the gradient density at pixel (x, y) in the input frame.

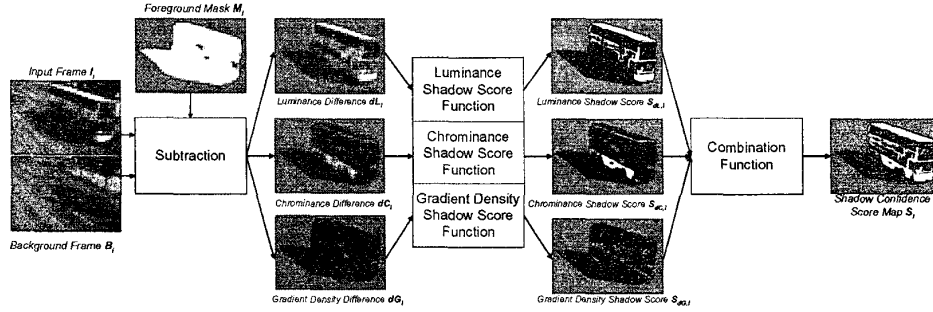


Fig. 2: Shadow Confidence Score Computation

Let the Moving Foreground Mask (MFM) $M_i(x,y)$ be defined as

$$M_i(x,y) = \begin{cases} 1, & |I_i(x,y) - B_i(x,y)| > T_{BG} \\ 0, & \text{otherwise} \end{cases}, \quad (3)$$

where T_{BG} is the threshold used in background subtraction. To indicate whether the region should be classified as cast shadow, a shadow confidence score, S , is defined. If the region is likely to be a cast shadow, a high score will be given to that region. On the other hand, if the region is likely to be object or background, a lower score will be given. The score is a probability value ranged from 0 to 1 inclusive.

As shown in Figure 2, the luminance, chrominance, and gradient density levels for each pixel in the input and background images are calculated. Then, the subtraction of the two images is calculated in the luminance, chrominance, and gradient density dimensions. To calculate the score S , the three mapping functions are defined, which are Luminance Score (S_L) vs Luminance Difference, Chrominance Score (S_C) vs Chrominance Difference and Gradient Density Score (S_G) vs Gradient Density Difference. Then, the overall score (S) is computed by combining these three individual scores.

Let $L_i(x,y)$ be the luminance difference between the i th input frame and the i th background at location (x,y) as given by

$$L_i(x,y) = I_{i,i}(x,y) - I_{B,i}(x,y), \quad \forall(x,y) \text{ where } M_i(x,y) = 1, \quad (4)$$

$$S_{L,i} = \begin{cases} 1, & L_i(x,y) \leq 0 \\ (T_L - L_i(x,y)) / T_L, & 0 < L_i(x,y) < T_L \\ 0, & L_i(x,y) \geq T_L \end{cases}. \quad (5)$$

T_L is a predefined parameter to accommodate the acquisition noise. As discussed in **Observation 1**, the luminance level should be lower in the input image when compared with the background image at the shadow. Therefore, for negative luminance difference value, the cast shadow criterion is satisfied and the region is most likely to be a cast shadow. On the

other hand, if the luminance level is higher in the input image comparing with the background image (positive luminance difference value), it does not satisfy the shadow criteria and $S_{L,i}$ tends to zero.

Let $C_i(x,y)$ be the chrominance difference between the i th input frame and the i th background at location (x,y) as given by

$$C_i(x,y) = \frac{|cb_{I,i}(x,y) - cb_{B,i}(x,y)| + |cr_{I,i}(x,y) - cr_{B,i}(x,y)|}{\forall(x,y) \text{ where } M_i(x,y) = 1}, \quad (6)$$

$$S_{C,i} = \begin{cases} 1, & C_i(x,y) \leq T_{C1} \\ \frac{T_{C2} - C_i(x,y)}{T_{C2} - T_{C1}}, & T_{C1} < C_i(x,y) < T_{C2} \\ 0, & C_i(x,y) \geq T_{C2} \end{cases}. \quad (7)$$

T_{C1} and T_{C2} are predefined parameters to accommodate the tolerance to chrominance change. As discussed in **Observation 2**, the chrominance of the input and background images should be the same or slightly shift in the cast shadow where there is less light shining on it. Thus, we have observed that the change will only occur in luminance dimension and there should be very small or no change in chrominance level. Therefore, for $C_i(x,y)$ less than T_{C1} , $S_{C,i}$ is set to 1 (high score) since it satisfies the shadow criteria (small change) in the chrominance dimension. For $C_i(x,y)$ larger than T_{C1} , $S_{C,i}$ is set to 0 because of the large change in chrominance dimension. $T_{C2} = 2 \times T_{C1}$ in this paper.

Let $G_i(x,y)$ be the chrominance difference between the i th input frame and the i th background at location (x,y) as given by

$$G_i(x,y) = g_{I,i}(x,y) - g_{B,i}(x,y), \quad \forall(x,y) \text{ where } M_i(x,y) = 1, \quad (8)$$

$$S_G = \begin{cases} 1, & G_i(x,y) \leq T_{G1} \\ \frac{T_{G2} - G_i(x,y)}{T_{G2} - T_{G1}}, & T_{G1} < G_i(x,y) < T_{G2} \\ 0, & G_i(x,y) \geq T_{G2} \end{cases}. \quad (9)$$

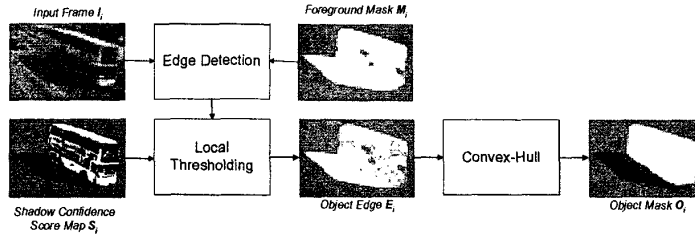


Fig.3: Moving Cast Shadow Detection

Gradient density (g_i) is the average of the magnitude of the gradient over a spatial window area. T_{G1} and T_{G2} ($= 2 \times T_{G1}$) are predefined parameters. As defined in **Observation 3**, after subtraction of gradient density in the input and background images, the gradient density level is mostly cancelled out in the cast shadow. However, in the object region, there is significant difference between the input and background images in gradient density. Therefore, for small gradient density difference value, the region is more likely to be shadow and $S_{G,i}$ is set to 1. For high gradient density difference, the region is likely to be an object and $S_{G,i}$ is set to 0.

After the three scores, $S_{L,i}$, $S_{C,i}$ and $S_{G,i}$, are calculated for the three difference dimensions, the total S_i is computed by combining all the three scores. Since each dimension is a necessary requirement for the region to be classified as cast shadow, hence direct multiplication of $S_{L,i}$, $S_{C,i}$ and $S_{G,i}$, $S_i(x,y) = S_{L,i}(x,y) \times S_{C,i}(x,y) \times S_{G,i}(x,y)$, is used in this paper.

4 Moving Cast Shadow Detection

In the detection part, as depicted in Figure 3, the cast shadow is separated from the object based on the shadow confidence score S_i and object edges of the foreground masked input frame. Firstly, all the pixels with significant gradient level are detected using the edge detector within the MFM. These pixels are denoted as E_i . Secondly, for each pixel with high gradient level, a thresholding test is applied to filter out the pixel with low shadow confidence score level. The threshold is denoted as T_S . This test removes noise and edge pixels which do not belong to the vehicle. If the high gradient pixel has high shadow confidence score level, the pixel will be discarded; otherwise, it will be retained.

$$E_i(x,y) = \begin{cases} 0 & \text{(discarded) for } S_i \geq T_S \\ 1 & \text{(retained) for } S_i < T_S \end{cases} \quad (10)$$

Thirdly, as discussed in **Observation 4**, the vehicle can be segmented out from the foreground mask by bounding convex hull on the vehicle edge

pixels. The remaining foreground mask is then classified as cast shadow region.

5 Results and Discussions

An image of a white and orange double-deck bus, which captured on a sunny day, is shown in Fig 4(a). In Fig 4(b), the background color image was generated by the background estimation. After subtracting the background frame from the input frame, the MFM after morphological closing transform is shown in Fig 4(c), where the white region is the foreground and the gray region is the background region. In Fig 4(d), (e) and (f), the results of S_L , S_C and S_G are shown respectively. In Fig 4(d), some parts of the bus, such as the windows and the orange part, are recognized as shadow since they have similar luminance level as the background image. Therefore, luminance can only provide limited indication on the shadow confidence score. In Fig 4(e), based on the chrominance value, the orange part of the bus is clearly classified as non-shadow because its chrominance value is significantly different from the background. In Fig 4(f), the regions with large gradient density difference are clearly marked as non-shadow. They include the blocked road lane mark region (Region A) and the tree branches region (Region B). By combining the results of S_L , S_C and S_G , the total shadow confidence score S is shown in Fig 4(g), in which, only the cast shadow region has high S value (which is mostly white) while the vehicle region is mostly covered with low S value. However, the windscreen and the windows have been falsely interpreted as shadow since they exhibit shadow-like features. In Fig 4(h), the edge pixels which have shadow score lower than threshold T_S are retained as object edge pixel. In Fig 4(i), the background, object and shadow are shown in gray, white and black colors respectively after performing the convex hull on the object edge pixels. It is important to use convex hull to bound the vehicle to recover most of the inner misclassified region.

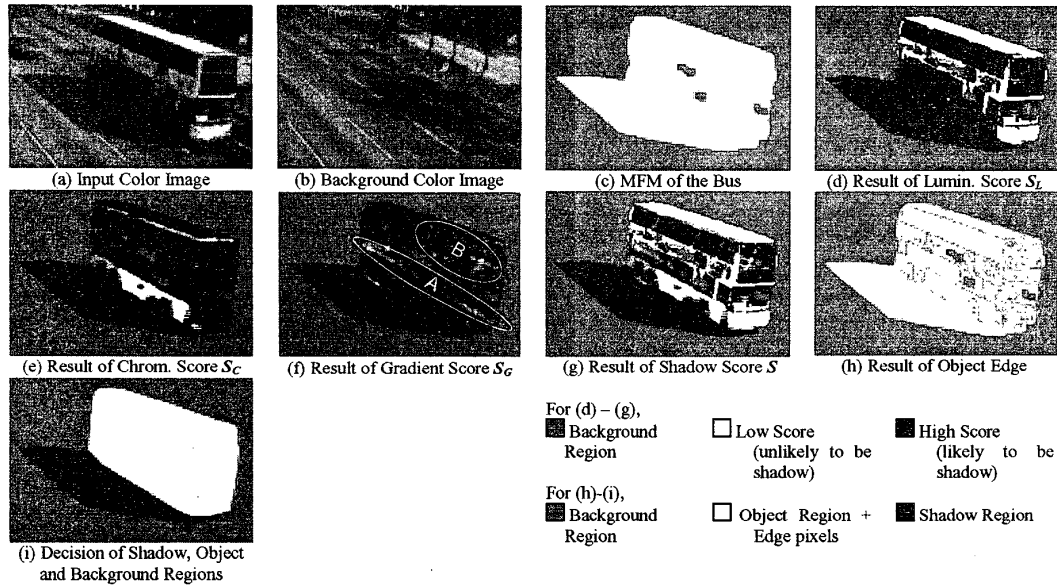


Fig. 4: Results on a Bus

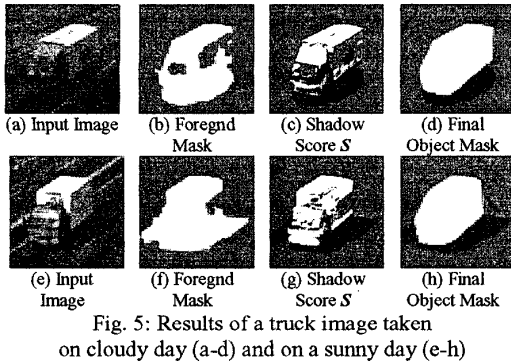


Fig. 5: Results of a truck image taken on cloudy day (a-d) and on a sunny day (e-h)

It is important that the method works well under different lighting conditions. We have tested our method under two different lighting conditions: cloudy and sunny. On a cloudy day, as shown in Fig 5(a), the cast shadow is poorly defined since it is mainly caused by ambient light. However, the computed foreground mask depicted in Fig 5(b) clearly includes a large region as shadow. The shadow confidence score S is shown in Fig 5(c). In Figure 5(d), by the proposed method, the object mask is successfully segmented by removing the shadow region. However, the right side mirror and the lower part of the tires are misclassified as shadow region. In Fig 5(e), 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 Figure 5(h), the lower part of the tires is misclassified as shadow region. Moreover, we noted that there is a shadow region correctly detected in 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 to detect multiple cast shadows.

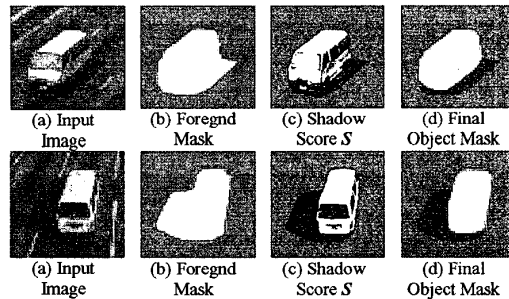


Fig. 6: Results of a van captured by a roadside camera (a-d) and an overhead camera (e-h)

Traffic surveillance cameras are mostly installed by the roadside or overhead. In Fig 6, results of a van captured by a roadside camera and an overhead camera are shown. In both cases, most of the van is successfully segmented with shadow removed. In Figure 6(d), a small part of the dark gray bumper is misclassified as shadow region since it exhibits shadow characteristics. In Figure 6(e), there is a white road lane mark partially covered by the cast

shadow of the van. As shown in Figure 6(h), this lane mark does not affect our method and is correctly classified as shadow region.

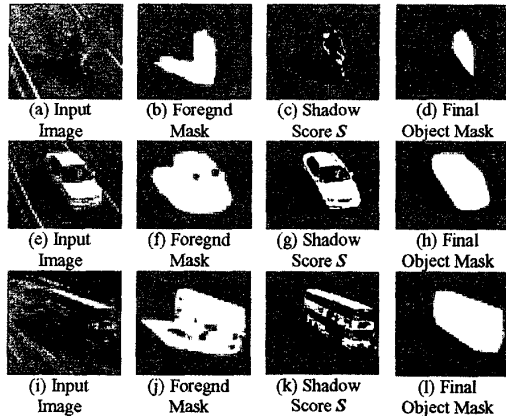


Fig. 7: Results of a motorbike (a-d), a sedan (small vehicle) (e-h) and a bus (large vehicle) (i-l)

On the road, there are many different vehicle models. Very broadly, vehicle can be roughly classified as motorbike, small vehicle (sedan, hatch, station-wagon, taxi) and large vehicle (van, truck, minibus, bus). In Fig 7, the results of apply our method to a motorbike, a sedan and a bus are depicted respectively. In Figure 7(d), the computed object mask is under segmented since the motorbike does not have a cuboid-like vehicle body. In Fig 7(h), since our proposed method use convex hull to define the region border, small vehicles, such as sedan, are not convex object. In addition, the side mirror of the vehicle further increases the false positive error of our method. As shown in Figure 7(l), our method performs better on large vehicles since they are mostly in the shape of cuboid and can be reasonably described by convex hull. In Figure 7(l), there are holes in the shadow region since the input image is very similar to the background image at these holes. Since the vehicle without cast shadow is the objective of our method, there is no impact of having holes in the shadow region.

6 Conclusions

In conclusion, we have presented a cast shadow detection method for estimating the vehicle outline which can effectively discriminate against shadow under different environment and vehicle factors. In our method, a shadow confidence score is computed in three different dimensions: luminance, chrominance and gradient density. Then, based on the shadow confidence score and the object edge pixels,

the cast shadow is separated from vehicle using convex-hull in the foreground mask.

We have tested our proposed method on different vehicle samples under typical outdoor scenes. From our results and analysis on various vehicle samples, the proposed method can successfully separate the cast shadow and a moving vehicle. By observing the effect on the error performance of our method under varying environment and vehicle factors, including lighting conditions, camera view angles, and vehicle types, we found that our proposed method is reasonably robust in various outdoor daylight environments and vehicles.

References

- [1] N. H. C. Yung, A. H. S. Lai, "A System Architecture for Visual Traffic Surveillance," *Proceedings of 5th World Congress on Intelligent Transport Systems*, 1998.
- [2] N. Hoose, *Computer Image Processing in Traffic Engineering*, Research Studies Press Ltd., London, 1991.
- [3] Makarov, "Comparison of Background Extraction Based Intrusion Detection Algorithm," *Proceedings of International Conference on Image Processing, I*, 521-524 (1996).
- [4] A. H. S. Lai, and N. H. C. Yung, "A Fast and Accurate Scoreboard Algorithm for Estimating Stationary Backgrounds in an Image Sequence," *Proceedings of IEEE International Symposium on Circuits and Systems*, 4, 241-244 (1998).
- [5] C. Jiang, and M. O. Ward, "Shadow Identification," *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 606-612 (1992).
- [6] G. Funka-Lea and R. Bajcsy, "Combining Color and Geometry for the Active, Visual Recognition of Shadows," *Proceedings of International Conference on Computer Vision*, 203-209 (1995).
- [7] J. Stauder, R. Mech, and J. Ostermann, "Detection of Moving Cast Shadows for Object Segmentation," *IEEE Transactions on Multimedia*, 1 (1), 65-76 (1999).
- [8] M. Kilger, "A Shadow Handler in a Video-Based Real-Time Traffic Monitoring System," *Proceedings of IEEE Workshop on Applications of Computer Vision*, 11-18 (1992).
- [9] G. G. Sexton, and X. Zhang, "Suppression of Shadow for Improved Object Discrimination," *Proceedings of IEEE Colloq. Image Processing for Transport Applications*, 9/1-9/6 (1993).
- [10] P. Gamba, M. Lilla, and A. Mecocci, "A Fast Algorithm for Target Shadow Removal in Mono-ocular Colour Sequence," *Proceedings of International Conference on Image Processing*, 436-439 (1997).
- [11] I. Mikic, P. C. Cosman, G. T. Kogut, and M. M. Trivedi, "Moving Shadow and Object Detection in Traffic Scenes," *Proceedings of International Conference on Pattern Recognition*, 321-324 (2000).
- [12] D. Koller, K. Daniilidis, and H. H. Nagel, "Model-Based Object Tracking in Monocular Image Sequences of Road Traffic Scenes," *International Journal of Computer Vision*, 10(3), 257-281 (1993).