Title	Shadow detection for vehicles by locating the object-shadow boundary
Author(s)	So, AWK; Wong, KYK; Chung, RHY; Chin, FYL
Citation	Proceedings Of The Seventh Iasted International Conference On Signal And Image Processing, Sip 2005, 2005, p. 315-319
Issued Date	2005
URL	http://hdl.handle.net/10722/93382
Rights	Creative Commons: Attribution 3.0 Hong Kong License

SHADOW DETECTION FOR VEHICLES BY LOCATING THE OBJECT-SHADOW BOUNDARY

Angie W. K. So, Kwan-Yee K. Wong, Ronald H. Y. Chung, and Francis Y. L. Chin
Department of Computer Science
The University of Hong Kong, Pokfulam, Hong Kong
{wkso, kykwong, hychung, chin}@cs.hku.hk

ABSTRACT
We introduce in this paper a shadow detection method for vehicles in traffic video sequences. Our method approximates the boundary between vehicles and their associated shadows by one or more straight lines. These lines are located in the image by exploiting both local information (e.g. statistics in intensity differences) and global information (e.g. principal edge directions). The proposed method does not assume a particular lighting condition, and requires no human interaction nor parameter training. Experiments on practical real-world traffic video sequences demonstrate that our method is simple, robust and efficient under traffic scenes with different lighting conditions. Accurate positioning of target vehicles is thus achieved even in the presence of cast shadows.

KEY WORDS
Shadow detection, shadow identification, object detection, video segmentation

1 Introduction

In traffic surveillance applications [1, 2], object detection and segmentation are two fundamental steps in video analysis tasks such as vehicle tracking and classification. Background subtraction and its variants are a common approach for these steps [3, 4]. However, shadows which cause illumination distortions on the background, are often extracted with their associated moving objects in background subtraction. This may lead to problems such as detected object shape distortion and object merging, which introduce inaccuracies in object location and recognition. Over the past few decades, different approaches have been proposed in the literature for shadow detection. Comparisons of various approaches are reported in [5, 6].

Most of the existing shadow detection methods exploit information derived from pixel appearance changes in the presence of cast shadows. They can be classified into model-based and non-model-based methods [5]. By assuming a Lambertian planar surface locally, model-based methods rely on the notion that the luminance ratio (i.e. the ratio of the intensity value of a pixel when it is under shadow to that when it is under illumination) is a constant. These methods [7, 8, 9, 10] use a linear transformation to describe the reduction of pixel intensity in shadow regions. Current frame pixels with illumination reductions that follow the linear model are then identified as probable shadow pixels. In [7, 8], the coefficients of the linear transformation are statistically estimated over a spatial region in a video sequence. A supervised classification algorithm is then applied to perform shadow detection on a pixel by pixel basis. This approach is efficient in the shadow detection step. However, it requires time-consuming and tedious manual segmentation and labeling procedures in the training phase, which make it impractical for real applications. In addition, the explicit computation of the transformation coefficients essentially maps a specific image intensity to the shadowed state of a pixel, which makes this approach dependent on particular illumination conditions. Besides, even for the same scene, image intensity over a shadow region may change under various circumstances. For example, the shadow cast by a big vehicle like a bus can be slightly darker than that of a small vehicle in real-world traffic scenes. The scene illumination also changes temporally when a cloud blocks the sun. The approaches in [9, 10] tackle the above problems by adaptively updating the linear transformation coefficients over time. However, the correct initialization of these coefficient values still requires human attention. Further, shadow intensities that fall outside the modeled value cannot be detected, and darker object regions with intensity values similar to the modeled shadow intensity may often be misclassified. Non-model-based methods, on the other hand, detect shadows by locating pixels with small chromatic but significant illumination distortions [3, 11]. Efficient thresholding techniques are commonly applied on these distortion values for shadow detection, but it is difficult to determine appropriate thresholds which work well under different situations. As a result, the literature diverges in the equations for describing such distortions, and the color space in which these distortions are measured.

We observe that the underlying problem in the existing methods is to determine a shadow point from the values of intensity reduction caused by shadowing in a limited local neighborhood. This problem is inherently as difficult as deciding if an isolated patch belongs to a shadow region by judging only from its color or intensity. On the other hand, humans perform shadow classification with recognition of the object casting the shadow and the whole shadow region. This indicates that object and global information

provide crucial cues for shadow detection. Therefore, we are motivated to propose a global shadow detection method by considering the spatial relationship between the object and its shadow. In essence, we perform shadow detection by locating the object-shadow boundary, which can be well approximated by one or more straight lines when the object is a vehicle or a polyhedra in general. In other words, we exploit not only information on pixel-wise intensity reductions, but also their spatial relationships caused by shadowing as demonstrated by empirical experiments [8, 9].

In summary, our approach has the following merits:

1. It is non-model-based, and does not require any human interaction or parameter training as in other approaches [8, 10].
2. It does not make any assumptions about the position of the light source [9], and thus allows variation of lighting conditions, scenes, and objects.
3. It is not purely pixel-based and considers global information, which is less error-prone.
4. It is simple and efficient as it requires no additional pre- and post-processing steps.

Section 2 summarizes the observations on which our approach is based, and our method is described in Section 3. Experimental results are given in Section 4, followed by conclusion and discussion on future work in Section 5.

2 Observations

Typical images of vehicles in a traffic video sequence along with their cast shadows are shown in fig. 1. The following observations are made:

Shadow properties

1. Shadow regions are darker. In RGB space, intensities in all three channels are reduced [7, 8, 12].
2. Shadow regions represent the same background surface under a reduced illumination, and share similar textures to the background.

Target object properties

1. The object-shadow boundary can be well approximated by one or more straight lines. These lines usually follow one of the segments of the vehicle boundary. This is illustrated in fig. 1.
2. The directions of the edges in an edge map are dominated by the directions of vehicle boundary segments, which provide good estimates for the directions of the object-shadow boundary segments.

The next section describes in detail our shadow detection method, which exploits these properties in locating the object-shadow boundary using a sliding window. In our discussion, we assume there is only one shadow cast by a vehicle in a traffic scene under a single light-source.

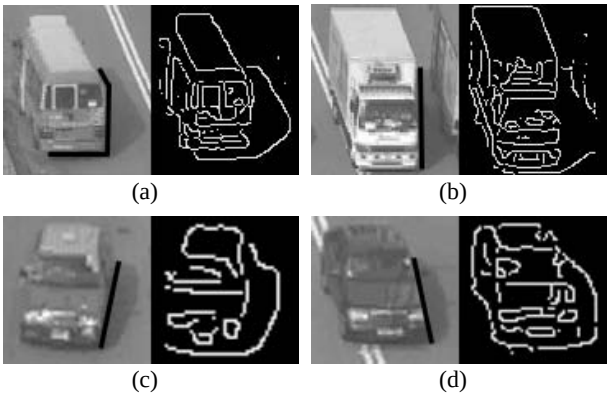


Figure 1. Example images of vehicles with their cast shadows, and their corresponding edge maps. The object-shadow boundary can in general be approximated by one or more straight lines, which are indicated by the thick black lines in the images. The directions of these lines can often be determined from the edge map.

3 Shadow detection method

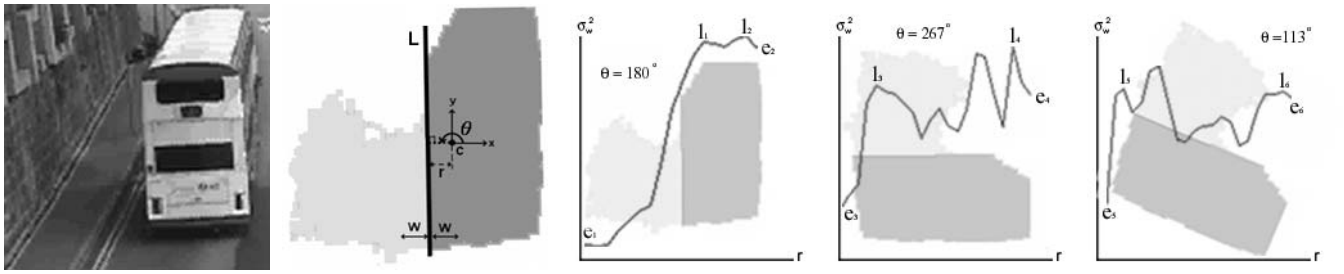
3.1 Moving object extraction

We use a Gaussian mixture background model [4] in RGB space for modeling the background scene. Background subtraction is applied on current frames to identify foreground pixels. After noise removal with morphological filtering, each region of connected foreground pixels of considerable size gives a binary object mask M.

3.2 Object-shadow boundary location

In the followings, we attempt to detect the shadow region by locating the object-shadow boundary on M. This boundary is approximated by a line $L: r = x\cos\theta + y\sin\theta$, where r is the distance of the line L along its normal direction from the centroid c of M to this line, and θ is the angle that the normal of the line L makes with the positive x-axis. This is illustrated in fig. 2(b). Referring to the figure, pixels on one side of L shaded in deep gray are labeled as *object*, and those on the other side shaded in light gray are labeled as *shadow*. The assignment of these labels is discussed at the end of this section. We assume temporarily that a shadow region is present in M, which is verified later (to be described in Section 3.3).

Assuming object property 2 in Section 2 holds, we deduce θ by first performing Canny edge detection for every pixel $p \in M$. Hough transform is then applied on the resultant edge map, and the cells of the accumulator array in radius-angle space are incremented, by taking into account the local intensity gradients dy and dx of individual edge pixels. Cell counts with the same angle at different radius are summed. The three most dominated angle directions $\theta = \arctan\frac{dy}{dx}$ ranging from 0° to 360° (wrapped around



(a) (b) (c) (d) (e)

Figure 2. Locating the object-shadow boundary: (a) shadow detection region in the original frame; (b) an object mask obtained by background subtraction. The object-shadow boundary is approximated by a line L with parameters r and θ, $2w$ is the width of the sliding window used in locating L; (c), (d) and (e) the plots for the variance in image-background intensity differences within the sliding window σ_w^2 against r for three different values of θ.

at $180°$) are determined by selecting angles with the highest sum of cell counts.

Next, we search for r along each of the three directions θ with a small sliding window of width $2w$, where w is a fixed perpendicular distance from L as shown in fig. 2(b). The variance of image-background intensity difference σ_w^2 for pixels $p \in M$ within the sliding window is computed. Fig. 2(c), 2(d), and 2(e) plot the values of σ_w^2 against r for different values of θ.

We expect to observe characteristic patterns in the plots when θ is chosen for L at the desired direction in separating the object and the shadow region. Consider:

Case 1: The object region is rich in texture
From shadow property 2 in Section 2, a shadow region shares similar texture with the background, while a textured object usually has a much larger texture difference from the background. σ_w^2 in the shadow region tends to be small, and that in the object region tends to be high and fluctuates to reflect the object textures. This typical curve is shown in fig. 2(c), in which the object-shadow boundary is located at a particular value of r with a large gradient of σ_w^2.

Case 2: The object region is plain in texture
In both the shadow and the object region, σ_w^2 is expected to be comparably small. Yet, the image-background intensity differences in the shadow region should be different from those in the object region in general. Thus, as the sliding window is moved to cover the object-shadow boundary, σ_w^2 is expected to attain its maximum value.

Note that if the texture difference and intensity difference are similar in both the object region and shadow region, our method would fail, but it would be equally difficult for humans to identify the object-shadow boundary.

When θ is chosen at a value such that L does not align with the actual object-shadow boundary in the image, we expect the plot of σ_w^2 against r to have large random variations. Fig. 2(d) and 2(e) illustrate this situation.

After computing the values of σ_w^2 against r for different values of θ as shown in fig. 2(c), 2(d), and 2(e), the first local maxima point l_i from each end point e_i of the curves are located. We then obtain two segments $e_i l_i$ at each angle θ, which represent the search range of parameter r of L. Assuming there is only one cast shadow associated with the vehicle object, the object-shadow boundary may occur only in one of these two segments at each angle. To further limit the search range of r, we select one of these two segments at each angle θ by computing a score which is proportional to (1) the magnitude of the difference in σ_w^2 between e_i and l_i, and (2) the number of pixels that fall in the aforementioned sliding window with intensity reduction in all RGB channels (shadow property 1, Section 2). The segment with a higher score at each angle θ is selected. The gradient of σ_w^2 along these selected segments at different values of θ are computed, and the pair of value of $r = r_s$ and $\theta = \theta_s$ with the maximum gradient is chosen. This pair of θ_s and r_s gives a candidate solution of L as the object-shadow boundary.

By considering the segment $e_i l_i$ that contains the candidate solution, if the value of r at e_i of this segment is smaller than r_s, all pixels $p \in M$ on the left side of L are labeled as *shadow*. Otherwise, those on the right side are. Unlabeled pixels $p \in M$ are labeled as *object*.

3.3 Verifying the presence of shadow

Recall that we have assumed in the search for L that a shadow region is present in M. To verify the validity of our assumption, two global criteria are evaluated:

1. All pixels on the *shadow* side of L should have reduced RGB values. (Section 2, shadow property 1)

2. The value of the variance of image-background intensity differences of pixels on the *shadow* side of L is small. This value is smaller than a threshold T. (Section 2, shadow property 2)

A candidate solution which does not satisfy these two criteria is rejected. If no solution for L is produced under

these criteria, our method decides no shadow exists in the foreground mask M under consideration.

3.4 Locating multiple lines

In typical traffic scenes, the cast shadow of a vehicle often appears in an "L-shape", for example the shadow in fig. 4(b) of the bus shown in fig. 4(a). In this case, the object-shadow boundary can be better approximated by two lines L_1 and L_2 instead of a single line. After locating the first line L_1 with the method as described in Section 3.2, the second line L_2 can be located at a different angle similarly. The line L_2 which corresponds to the greatest gradient of σ_w^2, and at the same time satisfying the following criteria is taken as a solution:

1. All criteria as described in Section 3.3
2. L_1 and L_2 intersect, and their intersection point is within the object mask M

If no solution for L_2 is produced under these criteria, our method may determine a single line L_1 is satisfactory in approximating the object-shadow boundary, or there is no shadow exists in the foreground mask M under consideration.

4 Experimental results

Fig. 3 and 4 show the shadow detection results of our proposed method for two outdoor traffic video sequences *Traffic1* and *Traffic2*. These sequences represent different illumination conditions, with visibly darker cast shadows in the *Traffic1* sequence. Without requiring parameter tuning, our method is able to detect shadow regions shaded in gray as shown in fig. 3(d),4(d), and 4(h).

The size of parameter w is set to be 5% of the dimension of M in our experiment. This window size works well according to our experimental results, in the sense that it can capture the large changes in variance σ_w^2 when the window is moving across the object-shadow boundary, without over-smoothing or under-smoothing effects. Also, the smallest possible w is desirable because the computation cost of our method is proportional to the size of w. The typical value of w is 4 pixels for our experimental video sequences with frame size 352×288 pixels. In our experiment, the value of the threshold T that is applied on the variance of image-background intensity differences of pixels on the *shadow* side is 100. This value is determined experimentally using 3 different video sequences recorded from different traffic scenes.

In the sequence *Traffic1* with 18 vehicles in total, shadows are correctly detected for 77.8% of vehicles, 22.2% are incorrectly detected, and 0% is not detected. The statistics for the *Traffic2* sequence with 81 vehicles in total are 81.5%, 8.6%, and 9.9% respectively. On a 1 GHz computer with 256 MB memory, this method can process 10–12 frames per second with the processing time proportional to the number of foreground objects present.

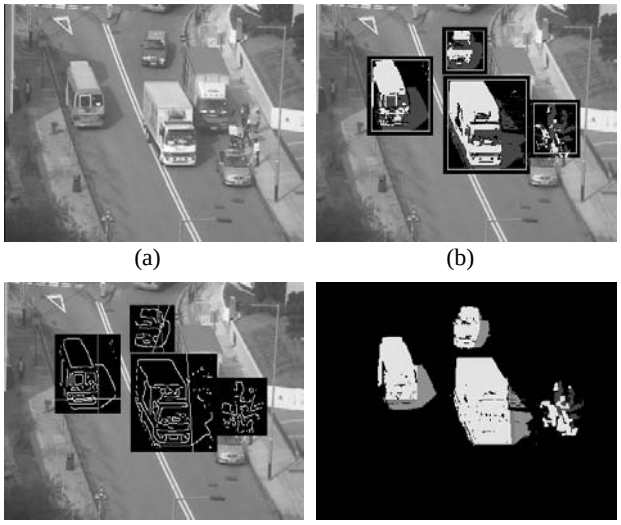


Figure 3. Shadow detection results in *Traffic1* sequence: (a) original video frame 532; (b) foreground regions detected by background subtraction, with probable shadow pixels having intensity reductions in all RGB channels colored in dark gray; (c) the edge maps. The lines L_i located by our method to approximate the object-shadow boundary are shown; (d) the refined object masks, with the *shadow* region in gray and the *object* region in white; Note that the lorry and the car parking along the road side on the right are modeled as background, which are not detected.

5 Conclusion and future work

In this paper, we have proposed a robust, simple, and efficient shadow detection method for vehicles in outdoor traffic scenes. Our method considers a global spatial relationship between an object and its shadow, requires no training nor parameter tuning, and makes no assumption about the light source. This method is therefore less sensitive to different scene illumination conditions as demonstrated by our experimental results.

However, one of the limitations of our method is that shadow regions of dimensions smaller than the sliding window size $2w$ cannot be detected. Besides, the difficulty in determining the accurate location of the object-shadow boundary increases with strong variations in background texture, together with a very large intensity reduction caused by shadowing. In this case, the value of σ_w^2 may not be small in the shadow region as expected, making the characteristic pattern of σ_w^2 against r as discussed in Section 3.2 can be unidentifiable. Moreover, we have assumed in our method that a foreground region containing only a vehicle and its shadow can be obtained by background subtraction, but this assumption may not hold in heavy traffic conditions where occlusions between vehicles are frequent.

Despite these limitations, our proposed method could be applied efficiently to practical traffic surveillance sys-

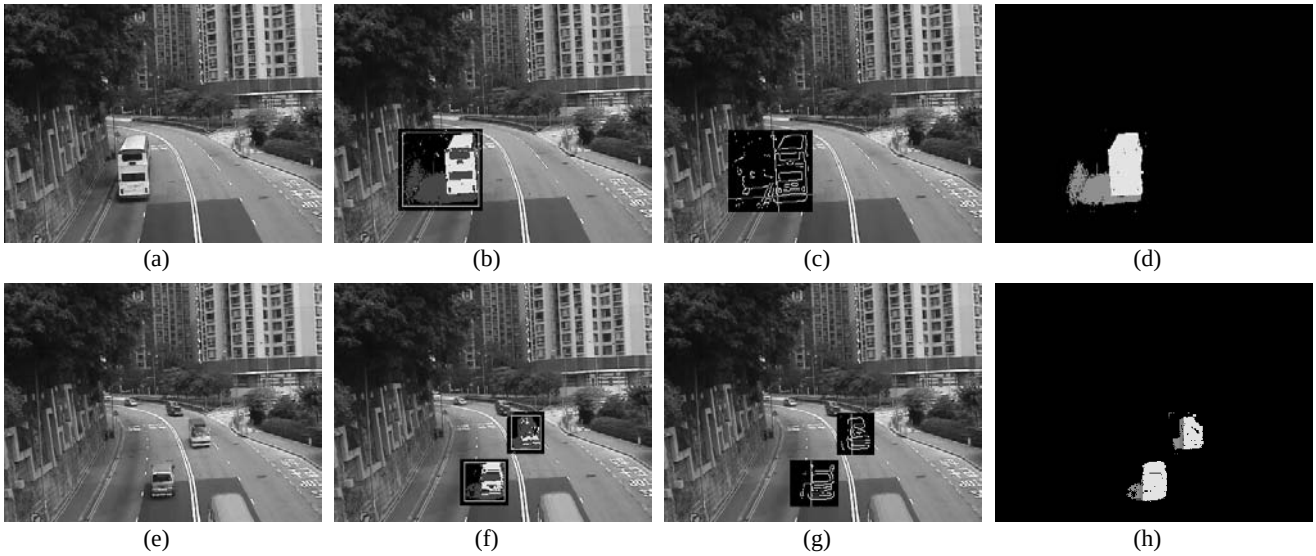


Figure 4. Shadow detection results in *Traffic2* sequence: (a)–(d) images showing a bus in frame 181 with an "L-shaped" shadow. Two lines are computed to approximate the object-shadow boundary; (e)–(h) images showing two vans in frame 703. A single line is sufficient for the approximation; This video sequence shows a scene under a different illumination condition from that in the *Traffic1* sequence as shown in fig. 3. Satisfactory results are obtained without any parameter tuning.

tems monitoring scenes with moderate traffic density and under a single light source.

Our work can be further improved in the future by including a local refinement stage for shadow detection after locating the object-shadow boundary. Intensity statistics and edge information on shadow points as labeled by L can be collected. A similarity measurement for spatial connectivity of shadow points can then be defined, and this allows a shadow region to be detected up to pixel accuracy. Furthermore, temporal information can be exploited during the location of L, such that r_s and θ_s can be modeled as a Kalman filter's states of a tracked vehicle, and be searched in the neighborhood of their predicted values. Finally, more work is anticipated for detecting shadows in foreground masks with multiple vehicles, and for extending our current method to detect non-vehicle objects.

Acknowledgment

This research is co-sponsored by Multivision Intelligence Surveillance Limited and the Innovation and Technology Commission of the Government of the Hong Kong Special Administrative Region, under the Grant UIM/118.

References

[1] A. Rajagopalan and R. Chellappa, "Vehicle detection and tracking in video," in *ICIP*, 2000, pp. I:351–354.

[2] H. Yang, J. Lou, H. Sun, W. Hu, and T. Tan, "Efficient and robust vehicle localization," in *ICIP*, 2001, pp. II: 355–358.

[3] T. Horprasert, D. Harwood, and L.S. Davis, "A statistical approach for real-time robust background subtraction and shadow detection," in *ICCV Frame-Rate Workshop*, 1999.

[4] C. Stauffer and W.E.L. Grimson, "Adaptive background mixture models for real-time tracking," in *CVPR*, 1999, pp. II: 246–252.

[5] A. Prati, I. Mikic, M.M. Trivedi, and R. Cucchiara, "Detecting moving shadows: algorithms and evaluation," *PAMI*, vol. 25, no. 7, pp. 918–923, July 2003.

[6] S. Nadimi and B. Bhanu, "Physical models for moving shadow and object detection in video," *PAMI*, vol. 26, no. 8, pp. 1079–1087, August 2004.

[7] I. Mikic, P.C. Cosman, G.T. Kogut, and M.M. Trivedi, "Moving shadow and object detection in traffic scenes," in *ICPR*, 2000, pp. Vol I: 321–324.

[8] K. Siala, M. Chakchouk, O. Besbes, and F. Chaieb, "Moving shadow detection with support vector domain description in the color ratios space," in *ICPR*, 2004, pp. IV: 384–387.

[9] J.M. Pinel and H. Nicolas, "Shadows analysis and synthesis in natural video sequences," in *ICIP*, 2002, pp. III: 285–288.

[10] Y. Wang, T. Tong, and K.F. Lo, "A probabilistic method for foreground and shadow segmentation," in *ICIP*, 2003, pp. III: 937–940.

[11] R. Cucchiara, C. Grana, M. Piccardi, and A. Prati, "Detecting moving objects, ghosts, and shadows in video streams," *PAMI*, vol. 25, no. 10, pp. 1337–1342, October 2003.

[12] S. Nadimi and B. Bhanu, "Moving shadow detection using a physics-based approach," in *ICPR*, 2002, pp. II: 701–704.