A Hybrid Shadow Removal Algorithm for Vehicle Classification in Traffic Surveillance System

Long Hoang Pham, Hung Ngoc Phan, Duong Hai Le and Synh Viet-Uyen Ha

Abstract Shadow is one of the common parts in the natural scenes and has become an important topic in the field of computer vision. In many vision-based traffic surveillance systems, shadows interfere with fundamental tasks such as vehicle detection, classification, and tracking. Thus, it is necessary to suppress the effect of shadows. A difficult part of the shadow removal problem is how to accurately detect and remove shadow regions and recover the boundaries of the vehicles, while still achieving real-time processing performance. Many powerful methods have been proposed to solve this dilemma; however, instabilities at the boundaries of moving vehicles are still challenges. In this paper, an improved algorithm to remove shadow regions, and quickly recovering the boundaries of moving vehicles is presented in a detailed manner. The proposed method applies edge information with background subtraction to handle daytime traffic scenes. Our approach has demonstrated more accurate results than previous approaches regardless of lighting luminance levels or shadow orientations.

Keywords Traffic surveillance system · Shadow removal · Edge detection
Vehicle recovery · Daytime detection · Vietnam

1 Introduction

The past decade has seen increasingly rapid advances in the field of computer vision which in turn has led to a renewed interest in traffic surveillance systems. Vision-based traffic surveillance systems have the capability to provide fast and reliable information that is necessary for traffic management and congestion mitigation.

L. H. Pham (✉) · H. N. Phan · D. H. Le · S. Viet-Uyen Ha (✉)
School of Computer Science and Engineering, International University,
Vietnam National University HCMC, Block 6, Linh Trung,
Thu Duc District Ho Chi Minh City, Vietnam
e-mail: phlong@hcmiu.edu.vn

S. Viet-Uyen Ha
e-mail: hvusynh@hcmiu.edu.vn

V. Bhateja et al. (eds.), *Intelligent Engineering Informatics*, Advances in Intelligent Systems and Computing 695, https://doi.org/10.1007/978-981-10-7566-7_65

The main objective is to detect interesting objects (moving vehicles, people, and so on). Other targets include classifying objects based on their features and appearance (shape, color, texture, and area), counting and tracking vehicles (trajectory, motion), assessing the traffic situation (congestion, accident). While later processes are dependent on specific application requirements, the initial step of object detection must be robust and application-independent.

However, a major problem with this kind of application is the appearance of shadows in daytime scenes. In many traffic surveillance systems, shadows interfere with fundamental tasks such as moving vehicle detection, classification, and tracking. Firstly, cast shadows that appear next to the conveyances distort vehicles' shapes and confound vehicle classification process. Secondly, many vehicles are connected by shadows and thus are detected as one big vehicle, which affects the procedure of counting and tracking.

Many powerful methods have been proposed to solve the problem of shadow removal; however, instabilities at the boundaries of moving vehicles are still challenges. In this paper, we present a simple, effective, and robust algorithm which can remove shadow regions and recover vehicles' boundaries in real-time. Our approach includes three steps: (1) detecting moving vehicles; (2) subtracting shadow regions; (3) recovering vehicles' boundaries. The main contribution of this paper is that we successfully recover the moving vehicle boundaries by combining edge information from the input frame and the lightness component with better performance in the experimental results than related methods.

The rest of the paper is organized as follows. In Sect. 2, our method and relevant theories will be discussed in detail. Section 3 describes the experiments and results obtained. The conclusion is presented in Sect. 4.

2 The Proposed Method

2.1 Outline of the Algorithm

The outline of the proposed method is described as follows. Given a new frame captured by the camera (I), a pair of background (BG) and foreground (FG) images are obtained. Then, I and FG are used in the moving object extraction process. The output, MO, which is an RGB image containing moving objects over a black background, is then converted to grayscale (MO^{Gr}) and HSV color model from which the lightness component (MO^{V}) is extracted and threshold. The Canny edge maps of BG, FG, MO^{Gr}, MO^{V}, denoted by E_{BG}, E_{FG}, E_{Gr}, E_{V}, respectively, are generated. With E_{BG}, E_{FG}, and E_{Gr}, the edge pixels of shadows can be removed (E_S). Next, E_V is combined with E_S to refine the vehicle edge maps, RE_S. Finally, the post-processing process is performed to denoise and construct the final binary mask, MV.

2.2 Moving Object Detection

It is a common task in traffic surveillance systems to utilize background modeling to construct the background, *BG*, and to detect moving objects in traffic scenes. The effectiveness of the background model is evaluated through a binary foreground mask, *FG*, in which moving objects and their cast shadows are marked as white blobs. However, in outdoor scenes, a real background is not always available and can be affected by extrinsic factors including slow-moving or stationary objects, and camera vibration (e.g., strong wind, heavy vehicles). To account for these problems, we adopt the background subtraction algorithm proposed by Nguyen et al. [1]. The outcomes of this model acquire good precision and real-time performance, which are critical factors in real-world applications. Figure 1 shows the examples of *BG* and *FG*.

The next step is to refine *FG* to remove unwanted objects and noises. Regarding this issue, we adopt the observation zone technique presented in both [2, 3]. In this approach, the observation zones which are a region where vehicles traveling through have steady changing rates in their appearances are automatically defined on the camera angle. Particularly, this procedure is crucial to enhance the performance of the system through improving the quality of vehicle classification and reducing the computational efforts by focusing on a smaller subset of moving objects. Figure 1d illustrates the refined foreground mask, *RFG*, after applying observation zone technique on *FG*:

$$RFG(p) = \begin{cases} 255 & \text{if } FG(p) \in OZ, \\ 0 & \text{otherwise.} \end{cases} \quad \text{where } OZ \text{ is the observation zone.} \tag{1}$$

Then, *RFG* is combined with *I* to create the moving object mask, *MO*, by applying an AND operation: $MO(p) = RFG(p) \wedge I(p)$. Figure 2c illustrates examples of *MO*. Also, we convert *MO* to gray-level, MO^{Gr}, and HSV color model from which the lightness component, MO^{V}, is extracted.

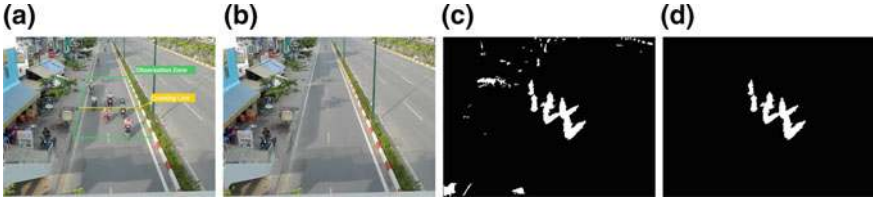

Fig. 1 **a** Observation zone and counting line. **b** Background model, *BG*. **c** Initial foreground mask, *FG*. **d** Refined foreground mask, *RFG*

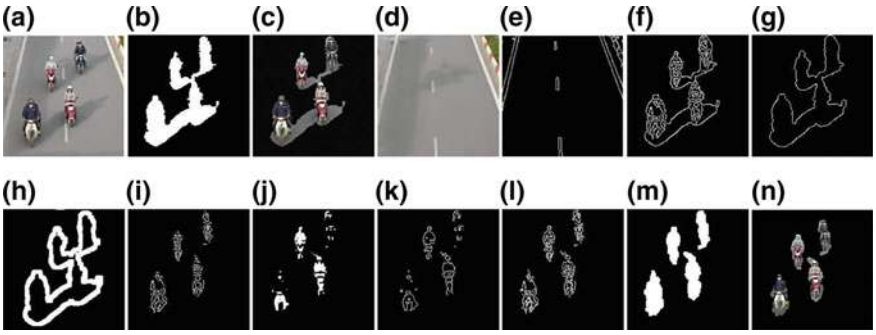

Fig. 2 Main steps in our algorithm with frame 451 of dataset PVD02. Note that the frame is cropped and scaled up for easier interpretation. **a** I. **b** RFG. **c** MO. **d** BG. **e** E_{BG}. **f** E_{Gr}. **g** E_{FG}. **h** DE_{FG}. **i** E_S. **j** MO^V. **k** E_V. **l** RE_S. **m** MV. **n** MV in RGB

2.3 Shadow Region Subtraction

Shadow region subtraction is the kernel of our contribution which comprises of two sections: (1) edge map generation; (2) shadow edge subtraction.

(1) Edge Map Generation. With respect to this issue, we make use of Canny edge detection. During the process of Canny edge detector, two threshold values (TH_1 and TH_2) are used to filter out the edge pixels with weak gradient values, which are caused by noise and color variation, and preserve the edge pixels with high gradient values. If the pixel value is smaller than the lower threshold (TH_1), it will be suppressed. If the edge pixel value is higher than the upper boundary (TH_2), it is marked as strong edge pixels. If the edge pixel value is between TH_1 and TH_2, it is marked as weak edge pixels. The two threshold values are empirically determined values, which will need to be defined when applying to different images. The problem becomes determining the optimal values for the thresholds when processing multiple frames captured under varying lighting conditions. We solve this problem by taking the median of I and then construct the upper and lower thresholds based on a percentage of this median:

$$\begin{cases} TH_1 = \max\{0, (1.0 - s) * M(I)\} \\ TH_2 = \min\{255, (1.0 + s) * M(I)\} \end{cases} \tag{2}$$

where $M(I)$ denotes the median of I and s is an optimal value used to vary the percentages. Typically, a lower value of s indicates tighter threshold, whereas a larger value of s gives wider threshold. In practice, $s = 0.52$ tends to give good results on the datasets we are working with. Table 1 summarizes the lower and upper thresholds calculated in our experiments with different lighting conditions.

The results, $E_{BG}, E_{FG}, E_{Gr}, E_V$, are the binary masks of values 0 and 255, as shown in Fig. 2e, f, g, k.

Table 1 Summary of TH_1 and TH_2 values calculated using Eq. 2

Video sequences	Conditions	TH_1	TH_2
Highway3	Strong shadows	66	216
VVK01	Morning, faint shadows	80	255
PVD01	Afternoon, strong shadows	69	228
PVD02	Afternoon, cloudy, varied shadows	65	215

(2) Shadow Edge Subtraction. From Fig. 2f, three observations can be made: (1) The cast shadows present sharp edges because the illumination source is far from the objects; (2) the vehicle has significant edges; however, the corresponding shadow is edgeless; (3) the edge of the cast shadow fastens on the boundary region of the moving foreground mask.

In order to remove the edges of shadows, we first compute the boundary of foreground mask, E_{FG}, which represents the outline of both the vehicle and its shadow. Then, the dilated boundary of the foreground mask which is more than one pixel thick is acquired:

$$DE_{FG} = E_{FG} \oplus E \tag{3}$$

where E is a 5×5 dilated structure element and $\oplus$ denotes morphological dilation.

Also, the edge map of the background model, E_{BG}, is computed to remove the edges created by background textures, such as road markings and pedestrian crossing pavements. Finally, the edge pixels of moving vehicles, E_S, are the interior edges of E_{Gr} (as shown in Fig. 2i), that is:

$$E_S(p) = \begin{cases} 255 & E_{Gr}(p) = 255 \text{ and} \\ & DE_{FG}(p) = 0 \text{ and} \\ & E_{BG}(p) = 0, \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

2.4 Vehicle Boundary Recovery

After the operator of Sect. 2.3, almost all the edge pixels in cast shadow areas are eliminated. However, the edges near the boundary of the vehicle have also been removed in the process, as shown in Fig. 2i. The recovery process consists of refining the vehicle boundary and recovering the vehicle mask.

To obtain additional edge information for the vehicle boundary, we look at the color aspect of shadow. We can use the value channel to acquire additional informa-

tion to refine the vehicle boundary. The value mask, V, is established by applying a binary threshold with a value of 130 on the moving object mask, MO, which is shown in Fig. 2j. This threshold value is chosen by some initial experiments, and then they are fixed for all experiments. Then, the edge map E_V is obtained as in Fig. 2k. Then, we combine E_S with E_V to create the refined edge map RE_S, that is, $RE_S(p) = E_S(p) \vee E_V(p)$ as shown in Fig. 2l.

Then, the recovery of vehicle shape is performed. First, a morphological erosion is applied using the square structure element E with the size of 5×5. Second, the contour of each vehicle is extracted by removing small connected regions and repainting the vehicle shapes. Then, a morphological dilation operation with the same square structure element, E, is used to compensate the effect of the morphological erosion. The final extraction of moving vehicles is denoted by MV, which is shown in Fig. 2m, n. Illustrations of MV show that our algorithm can remove all shadow areas while still preserving the vehicle features.

3 Experiments and Discussion

To evaluate the proposed algorithm, we performed some experiments using the Highway3 video from dataset ATON [4], which is often used in recent publications regarding traffic surveillance system, and our datasets captured in Ho Chi Minh City, Vietnam. The video sequences were captured at daytime under a different lighting condition with the resolution of 640×480 and at the frame rate of 30 fps. We also compare our approach with related methods: Chromacity [5], Geometry [6], Edge [7], Physical [8], SR Texture [9], and LR Texture [10] on a system having a configuration of Intel Core i7 2630QM and 8GB of RAM.

3.1 Subjective Evaluation

Experiments were conducted on four traffic datasets. We have also compared our results with other methods. For all frames in these video sequences, our algorithm can achieve satisfactory results of removing shadows from moving vehicles.

On the first row of Fig. 3, we performed experiments on dataset Highway3. The result of the LR Texture method still contains some noises on the right of the vehicle which are created by the textures on the road. Moreover, the Edge method could not recover the left side of the vehicle after subtracting the shadow areas. In our method, because the road textures are removed during the shadow edge subtraction process, the vehicle boundary is refined using additional information. Hence, we could obtain a cleaner vehicle extraction. The second row of Fig. 3 illustrated the experiments on dataset VVK01, which was captured in the morning with faint shadows under low lighting conditions. In this scenario, all methods, except for our and LR Texture, fail to deliver satisfying results and misclassify parts of the bus as shadows. Compar-

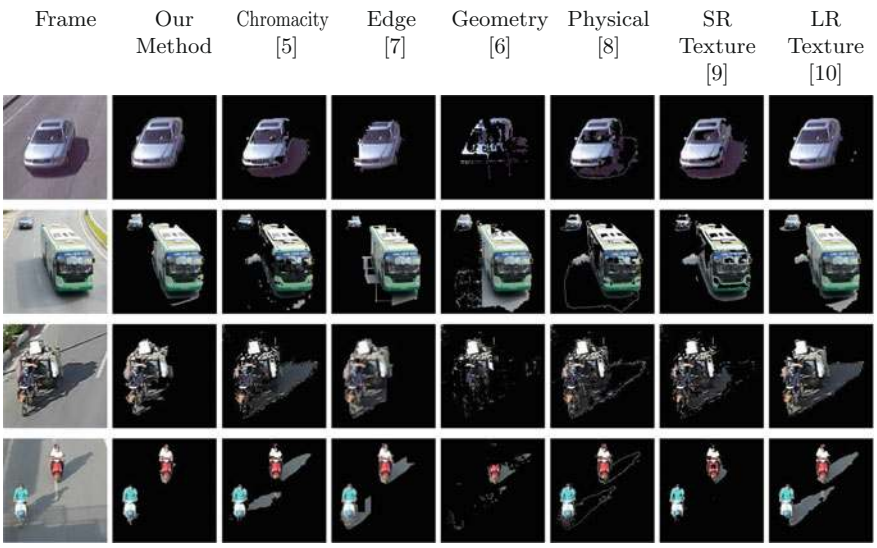

Fig. 3 Comparison results of shadow detection and removal. First row: the car from frame 1171 of dataset Highway3. Second row: the bus from frame 3750 of dataset VVK01. Third row: the unconventional vehicle from frame 1059 of dataset PVD01. Fourth row: the motorbikes from frame 948 of dataset PVD02

ing with LR Texture results, our method presents a better result. The third row of Fig. 3 is an experiment carried out on dataset PVD01, which was captured in the afternoon with strong shadows casting alongside the vehicles. We demonstrate the effectiveness of our shadow removal method under a special case with an unconventional vehicle. Although the Edge method can remove the shadows on the right side, it fails to detect the shadows under the vehicle, but our method can remove all the shadows. These results show that our algorithm can dynamically adapt to the real-world traffic scenes and therefore outperforms other methods.

3.2 Objective Evaluation

The error rates ε, proposed by [11], of the binary mask of moving vehicles, MV, are used to show the effectiveness of our algorithm. It is defined as follows:

$$\varepsilon = \frac{N_e}{N_I} \tag{5}$$

where N_I is the frame size and N_e is the number of pixels in MV that are different from the reference alpha plane.

Table 2 Average error rates of shadow detection and removal algorithms

Methods	Highway3	VVK01	PVD01	PVD02
Chromacity [5]	0.0072	0.0025	0.0017	0.0055
Edge [7]	0.0145	0.0044	0.0008	0.0049
Geometry [6]	0.0022	0.0069	0.0005	0.0046
Physical [8]	0.0049	0.0015	0.0009	0.0021
SR Texture [9]	0.0095	0.0106	0.0011	0.0027
LR Texture [10]	0.0094	0.0227	0.0019	0.0068
Our method	0.0021	0.0024	0.0006	0.0004

Table 3 Average processing time

Step	VVK01 (ms)	PVD01 (ms)	PVD02 (ms)
Background subtraction	12.40	15.62	15.93
Foreground object extraction	6.53	6.20	5.49
Canny edge detection	3.42	3.45	3.29
Shadow edge subtraction	0.22	0.28	0.28
Vehicle edge refinement and post-processing	1.81	1.85	1.78
Total	24.38	27.40	26.77

Table 2 reports the error rates for all methods used in these experiments. We can see that our method does not always have the lowest error rates in all video sequences. However, they are relatively equivalent to the best results. Also, the average error rates of each method are: 0.138% (our method), 0.443% (Chromacity [5]), 0.355% (Geometry [6]), 0.615% (Edge [7]), 0.235% (Physical [8]), 0.598% (SR Texture [9]), and 1.020% (LR Texture [10]). It shows that our algorithm can robustly detect shadows with minimum error rates in different scenarios.

The average processing time in each step of our algorithm can be found in Table 3. The average time needed to process one frame is ranging from 24.38 to 28.61 ms. Our algorithm has been optimized to run in the parallel fashion so that it only introduces a small delay to the overall processing time. Therefore, we are confident that after incorporating with our algorithm, the surveillance traffic system can achieve real-time processing capability (fps $\geq$ 25).

4 Conclusion

In this paper, we have proposed a novel shadow removal algorithm in daytime traffic scenes. The algorithm is based on edge information from both the input frame and the lightness component of HSV color model. The advantages of our method

are: (1) The algorithm is robust to a variety of shadow orientations, shapes, and appearances under different lighting conditions; (2) the algorithm precisely removes shadows from both smooth and textured backgrounds. Experiments show that our algorithm performs better than previous method as it can run in real-time speed when processing single or multiple traffic sequences.

Acknowledgements The study was supported by Science and Technology Incubator Youth Program, managed by the Center for Science and Technology Development, Ho Chi Minh Communist Youth Union, the contract number is "20/2017/ HĐ-KHCN-VU".

References

1. Nguyen, T.P., Tran, D.N.-N., Huynh, T.K., Ha, S.V.-U.H.: Disorder detection approach to background modeling in traffic surveillance system. J. Sci. Technol. Vietnamese Acad. Sci. Technol. **52**(4A), 140149 (2014)
2. Pham, L.H., Duong, T.T., Tran, H.M., Ha, S.V.U.: Vision-based approach for urban vehicle detection and classification. In: 2013 Third World Congress on Information and Communication Technologies (WICT 2013), pp. 305–310 (2013)
3. Ha, S.V.-U., Pham, L.H., Tran, H.M., Ho-Thanh, P.: Improved vehicles detection and classification algorithm for traffic surveillance system. J. Inf. Assurance Secur. **9**(5), 268277 (2014)
4. ATON datasets by UCSD. http://cvrr.ucsd.edu/aton/shadow/
5. Cucchiara, R., Grana, C., Piccardi, M., Prati, A.: Detecting moving objects, ghosts, and shadows in video streams. IEEE Trans. Pattern Anal. Mach. Intell. **25**(10), 1337–1342 (2003)
6. Hsieh, J.-W., Wen-Fong, H., Chang, C.-J., Chen, Y.-S.: Shadow elimination for effective moving object detection by Gaussian shadow modeling. Image Vis. Comput. **21**(6), 505–516 (2003)
7. Xiao, M., Han, C.-Z., Zhang, L.: Moving shadow detection and removal for traffic sequences. Int. J. Autom. Comput. **4**(1), 3846 (2007)
8. Huang, J.B., Chen, C.S.: Moving cast shadow detection using Physics-based features. In: 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, CVPR Workshops 2009, pp. 2310–2317 (2009)
9. Leone, A., Distante, C.: Shadow detection for moving objects based on texture analysis. Pattern Recognit. **40**4 (2007)
10. Sanin, A., Sanderson, C., Lovell, B.C.: Improved shadow removal for robust person tracking in surveillance scenarios. In: 2010 20th International Conference on Pattern Recognition, pp. 141–144 (2010)
11. Chien, S., Ma, S., Chen, L.: Efficient moving object segmentation algorithm using background registration technique. IEEE Trans. Circuits Syst. Video Technol. **12**(7), 577586 (2002)