REAL-TIME VIDEO SURVEILLANCE SYSTEM FOR TRAFFIC MANAGEMENT WITH BACKGROUND SUBTRACTION USING CODEBOOK MODEL AND OCCLUSION HANDLING

Zakaria Moutakki[1,2], Imad Mohamed Ouloul[1], Karim Afdel[1], Abdellah Amghar[1]

[1] *Laboratory of Metrology and Information Processing, Department of Physics, Faculty of sciences,*
University Ibn Zohr, B.P 8106, Agadir, Morocco
[2] *zakaria.moutakki@edu.uiz.ac.ma*

The scope of this paper is a video surveillance system constituted of three principal modules, segmentation module, vehicle classification and vehicle counting. The segmentation is based on a background subtraction by using the Codebooks method. This step aims to define the regions of interest associated with vehicles. To classify vehicles in their type, our system uses the histograms of oriented gradient followed by support vector machine. Counting and tracking vehicles will be the last task to be performed. The presence of partial occlusion involves the decrease of the accuracy of vehicle segmentation and classification, which directly impacts the robustness of a video surveillance system. Therefore, a novel method to handle the partial occlusions based on vehicle classification process have developed. The results achieved have shown that the accuracy of vehicle counting and classification exceeds the accuracy measured in some existing systems.

Keywords: Intelligent transportation systems; Vehicle detection; Occlusion handling; Vehicle classification; Background subtraction

1. Introduction

In recent years, the world has experienced a significant growth in road traffic. This had imposed on road network managers to optimize the use of existing infrastructures and to provide users with reliable and comfortable conditions. Severe traffic disruptions, such as congestion and accidents, require real-time control of traffic conditions and the rapid implementation of suitable traffic management measures. To achieve this important target, road traffic video surveillance systems appear as an indispensable tool.

In general, traffic surveillance systems are based on sensors such as infrared sensors, radars, and cameras. However, the systems based on cameras are adequate due to their low installation and maintenance cost, as well as their functioning guarantees real-time traffic surveillance and control.

Several requirements must be satisfied to ensure the robustness and the performance of a video surveillance system based on a stationary camera. In real time, it must be able to detect and classify vehicles taking into account certain constraints that can influence its efficiencies, such as vehicle occlusion, traffic congestion, and illumination conditions. The general goal of our research is to develop surveillance system which ensures counting, tracking and classification of vehicles with a sufficient accuracy under the conditions previously cited.

The realized system is comprised of three modules: segmentation, classification and counting. In every module, a set of techniques is made to improve the system performance. Our method is able to detect and classify vehicles in real time. Thus, every vehicle is classified depending to its type, car, bike or truck. Then, vehicles is counted and tracked.

This paper is organized as follows. The section 3 presents the different methods used to detect vehicles, recognize their type and to handle partial occlusions. The test of the proposed system on several scenes and its comparison with other existing systems is presented in section 4. Finally, Section 5 contains a general conclusion.

2. Related Work

During the last decade, numerous methods of vehicle detection and classification were developed (Nicolas and Brulin, 2014; Kastrinaki *et al*., 2003; Atiq *et al*., 2010; Coifman *et al*., 1998). Generally, these methods start with motion detection phase. The objective in this phase is to detect moving vehicles

in the scene. This operation is frequently based on the background subtraction (Yu *et al.*, 2000; Mandellos *et al.*, 2011; Xia *et al.*, 2016) or Optical Flow (Ji *et al.*, 2006; Johansson *et al.*, 2009; Hossen and Tuli, 2016). In the systems based on background subtraction, the principal purpose of the algorithm is to model the background including stationary objects in the scene. Moving objects are then detected by utilizing pixel-wise subtraction between the current frame and background d model. The second widely used method is Optical Flow (Aslani and Mahdavi-Nasab, 2013). This technique measures the motion magnitude and direction for each image point with high accuracy which gives excellent results. However, Optical Flow needs a relatively important cost of time for the computation and, as a consequence; it does not meet the requirements within a system that operates in real time. The different extraction and segmentation methods produce a set of parameters for each vehicle: Speed, length, width and height of vehicle. The goal of the classification operation is to categorize the vehicle into a number of predefined types using these generated parameters (Fu *et al.*, 2016; Zhou and Cheung, 2016; Wen *et al.*, 2015).

Among the issues widely known in video surveillance of road traffic, partial occlusions appear as a major problem which negatively influences the accuracy of video surveillance systems. Thus, many methods have been developed to eliminate them. Some methods are based only on visible information of occluded objects (Song and Nevatia, 2007; Tang *et al.*, 2014; Shu *et al.*, 2012). Kanhere and Birchfield (2008) propose a real-time incremental segmentation and tracking system at low camera angles using stable features with partial occlusion handling; their method can detect the partially occluded using an estimation of the distribution of feature points such as vehicle corners. Huang and Liao (2004) presented a technique based on vehicle motion vectors. If one detected object has more than one motion vector, we are defiantly front of a vehicle occlusion. Thus, motion vectors belong to different vehicles. Tian *et al.* (2015) proposed a method based on designing and training specific grammars. It starts by dividing vehicle into semantic parts and sub-parts then to train the grammar productions for semantic parts and sub-parts by introducing novel pairwise SVM grammars and finally to detect the vehicle by applying the trained grammars. Using skeleton features, another study was proposed by Chen and Liu (2012) for occlusion handling. This method used the morphological skeleton operation for detection of the occlusion effect between the blobs by extracting the skeleton features which are used to segment the obstructed vehicles. Using independent component analysis (ICA) and Principal Component Analysis (PCA), Wang and Lien (2008) propose a statistical approach which detects the partially occluded vehicles using the local features located in significant sub-regions of the image.

3. Proposed System

Figure 1 shows the chain of different methods involved into the developed system. The proposed system includes three main modules. Each module contains a set of techniques in order to have a good quality of the motion detection and vehicle classification modules. The main modules of the proposed method are:

Segmentation module: The objective of that stage is to separate the regions of interest (ROIs) related to vehicles. The camera in our video surveillance system is stationary. It means that the background model contains all static objects in the scene. Taking this reality into account, we opted for the background subtraction as the first step of the segmentation phase to isolate moving objects in a mask. The subtraction of the static background is followed by a set of filtering operations in order to eliminate the segmentation artefacts. The detected contours in the filtered frames will be subject thereafter to a contour closing operation in order to model them by rectangles.

Vehicle classification: involves the use of the Histograms of Oriented Gradient (HOG) descriptors (Dalal and Triggs, 2005) and Support Vector Machine (SVM) (Cortes and Vapnik, 1995) to define the different types of vehicles (cars, trucks, and bikes). The classification phase is sensitive to the segmentation quality on the one hand, and on the other, on the presence of partial occlusion of vehicles. For this reason, we performed a technique to handle occluded vehicles based on classification process. Thus, two classes in the image database will be associated to occlusions (vertical and horizontal) In order to detect these occlusions and process them later.

Vehicle counting and tracking: are performed in a defined zone of the road. Any vehicle entering into this zone is taken into account by our system.
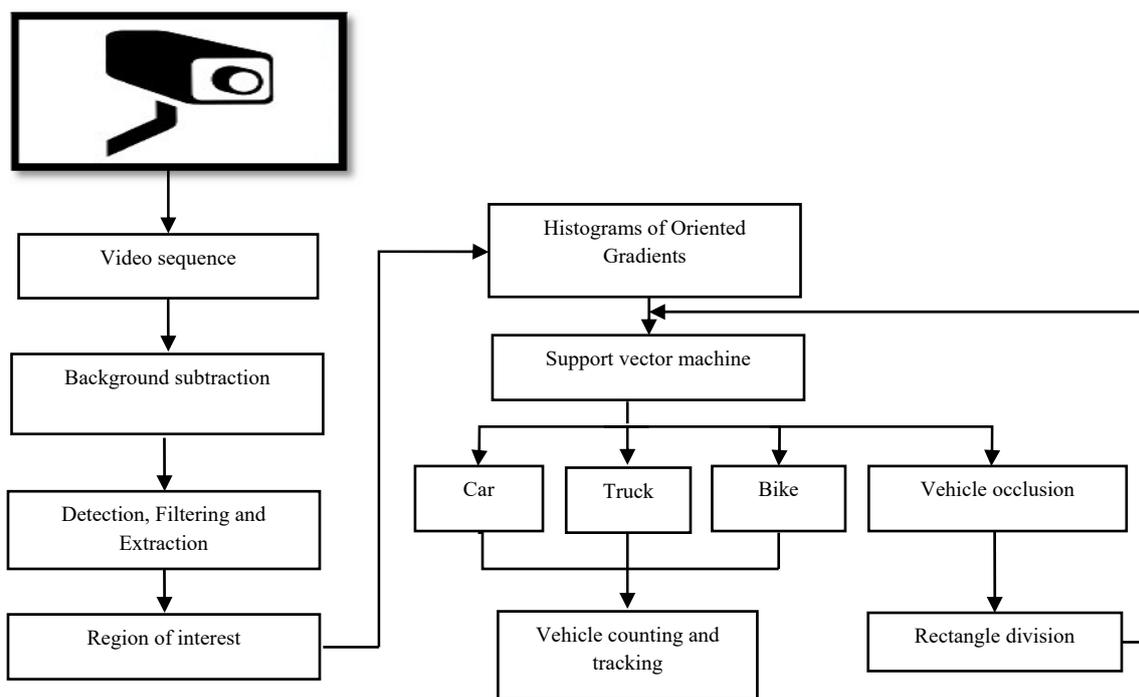
Figure 1. Flowchart of the proposed system

3.1. Image segmentation

3.1.1.   *Background subtraction*

The first step of the segmentation phase consists in extracting moving objects from the image sequence. There have been multiple proposed methods of background subtraction in the literature (Horprasert *et al.*, 1999; Stauffer and Grimson, 1999; Kim *et al.*, 2005). The principle behind these methods is to define a model corresponding to the static background in the scene by learning and subtract this model from the image sequence to isolate moving objects.

A test was conducted to analyze the results of the background subtraction using these three methods. The first method (Fig. 2(b)) models the background by a unimodal distribution of a Gaussian per pixel (Horprasert *et al.*, 1999). The second method (Fig. 2(c)) uses a Gaussian mixture for each pixel (Stauffer and Grimson, 1999). The third method is based on Codebook technique (Kim *et al.*, 2005). In short, the Codebook method consists conserving a compressed representation of the background model based on the quantification and the grouping of the observed values during an initialization period. The filtering of the observed values belonging to moving objects can be performed by a periodicity measurement. The test shows that the result given by the Codebook method (Fig. 2(d)) is most efficient because moving objects corresponding to vehicles appear clearly in foreground mask with minimal noise.
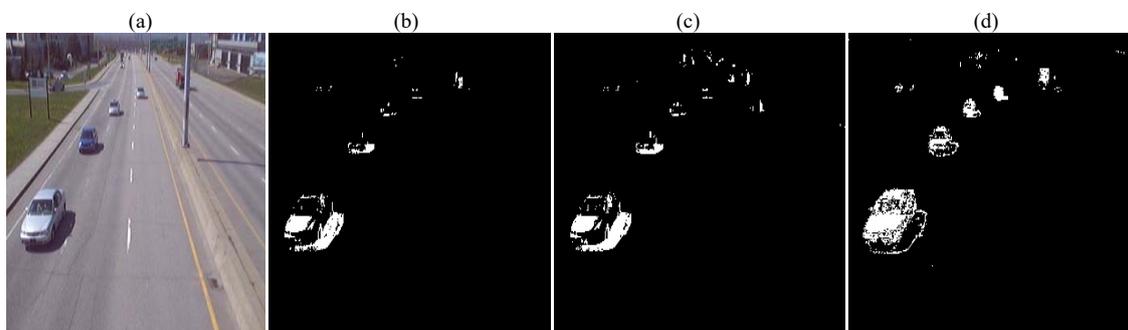

Figure 2. Motion detection using background subtraction
    (a) Video sequence
    (b) Background subtraction based on one Gaussian per pixel
    (c) Background subtraction using MOG method
    (d) Background subtraction using Codebook method

3.1.2.   Filtering and extraction

To eliminate the noise due to the suppression of static background, we have used a set of filtering and extraction techniques:

- Step 1: Morphological operations, such as dilatation and erosion (Fig. 3(b)).
- Step 2: Contour detection using Canny (1986) (Fig. 3(c)).
- Step 3: Calculation of contour coordinates in frames.
- Step 4: Elaboration of a rectangular mask associated to vehicles (Fig. 3(d)).
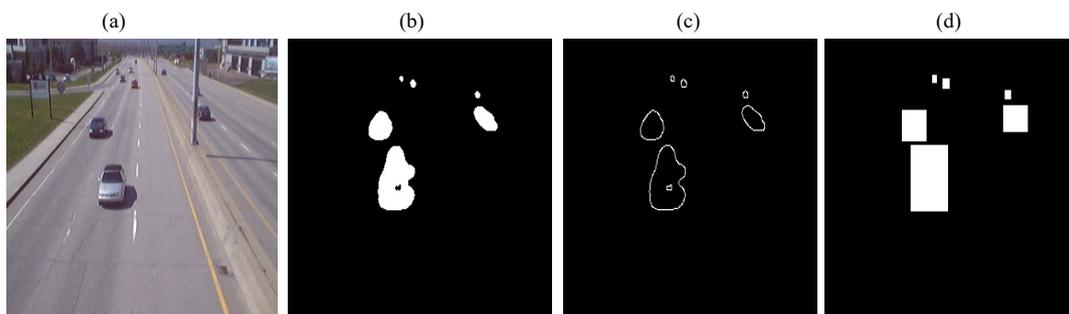

Figure 3. Results of the filtering operation applied on video sequence
(a) Video sequence (b) Morphological operations (c) Canny edge detector (d) contour closing

3.2. Vehicle classification

Vehicle classification is an indispensable tool in road traffic video surveillance systems, it is used to define the type of each vehicle (car, motorcycle, truck) crossing the road.

3.2.1.   Learning and training

During the segmentation phase, we realized a mask containing the regions of interest associated with vehicles modelled by rectangles. To isolate the image of a vehicle, we have conducted the operator "and" between the mask and the source image. The vehicle image obtained will then be transmitted to the SVM classifier to select its type. Fig. 4 reveals the steps executed to distinguish the type of vehicles.
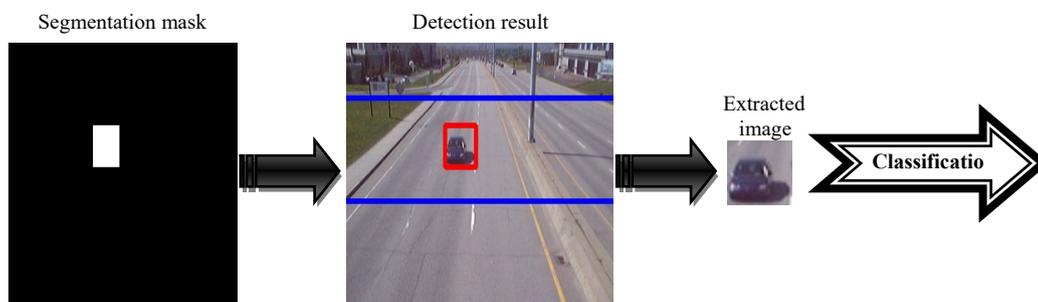

Figure 4. Vehicle classification steps

The classification process begins with a learning stage. For that purpose, we defined an image database size 32*32 (Fig. 5). The images constituting the learning base should represent all existing types of vehicles (car, bike, and truck). To classify also the occluded vehicles, two classes are added to the learning database.

Figure 5. Database used in the learning phase

Thus, each detected object will be classified to one of the five classes defined in the database (Table 1). Each class will contain images representing a defined type. This operation aims in one hand to regroup the partial occlusions in two classes in order to detect them and process them subsequently, on the other hand, to classify the other vehicles to its appropriate type.

Thus, the determination of classes depends on two factors:

- Existing vehicle types: car, bike, and truck.
- Vertical and horizontal partial occlusions.

Table 1. Determined classes and their appropriate images

Classes	Class 1	Class 2	Class 3	Class 4	Class 5
Image					
Classification result	Car	Bike	Truck	Vertical occlusion	Horizontal occlusion

To characterize the different types of vehicles and partial occlusions, we used HOG descriptor combined with SVM. HOG can describe the local appearance and the form of object based on an intensity distribution of gradients or based on the contour direction. HOG starts with the division of image to a set of cells. The combination of histograms computed in each cell forms the HOG descriptor. The next step is the combination of the feature vectors using SVM which is a simple method of classification; it is characterized by its power of generalization.

3.2.2. *Occlusion handling*

The basic idea of our method is to use the classification phase as a tool to verify if a partial occlusion is present in the scene. Figure 6 shows the flowchart used by our system. If the generated type of vehicle is car, truck or bike, the system proceeds directly to vehicle counting and tracking module. Otherwise, if there's a presence of a vertical partial occlusion in the scene, the height of the rectangle associated with this occlusion will be divided on two to get two rectangles one per vehicle. However, if the system detects horizontal occlusion, the width of the rectangle associated with the occlusion will be divided on two. The newly detected objects will be classified again to determine its type.

Figure 6 and 7 show the results of the technique used to process vertical and horizontal occlusions.
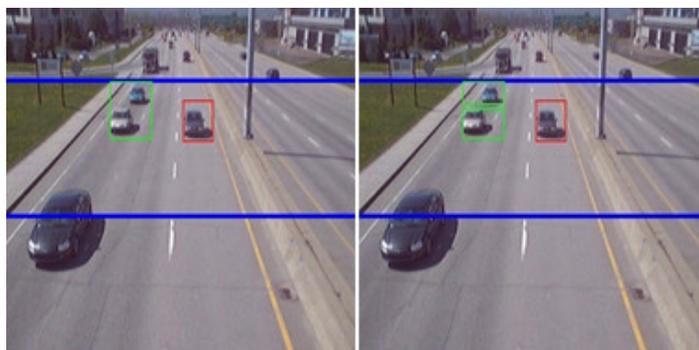

Figure 6. Results of the proposed method for vehicle occlusion processing in the case of a vertical occlusion
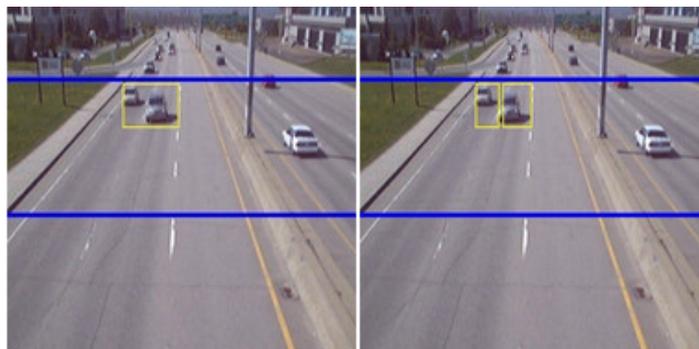

Figure 7. Results of the proposed method for vehicle occlusion processing in the case of a horizontal occlusion

3.3. Vehicle counting and tracking

After the classification phase, each vehicle is counted and tracked. The process is based on a set of vectors in 2-dimensions. Three of these vectors add at each frame the positions (x,y) of vehicles in the scene depending on its type.

We use another vector which responsible of adding the positions of new vehicles present in the scene and removing the positions of vehicles which left the surveillance zone at each frame. Using this vector we can define the number reached by each type of vehicles, in one hand. We can know at each frame the position of vehicles in the scene, in the other hand.

4.   Results and Discussion

To evaluate the quality of this study, we reserved this section to represent the experimental results acquired during the execution of the proposed system. Then, we introduce a comparative study between our method and other existing methods.

4.1. System test

An implementation of the system has been tested on different road scenes. Figure 8 illustrates an example of the test conducted in a road scene. A surveillance zone is designated by the blue horizontal lines. Each vehicle enters this zone will be detected and classified, and then the vehicle will be counted and tracked. The colour of the detection rectangle around each vehicle varies depending on the type of the detected vehicle. Counting results are displayed in a white bar above the image sequence indicating the current number reached by the traffic flow for each vehicle type.
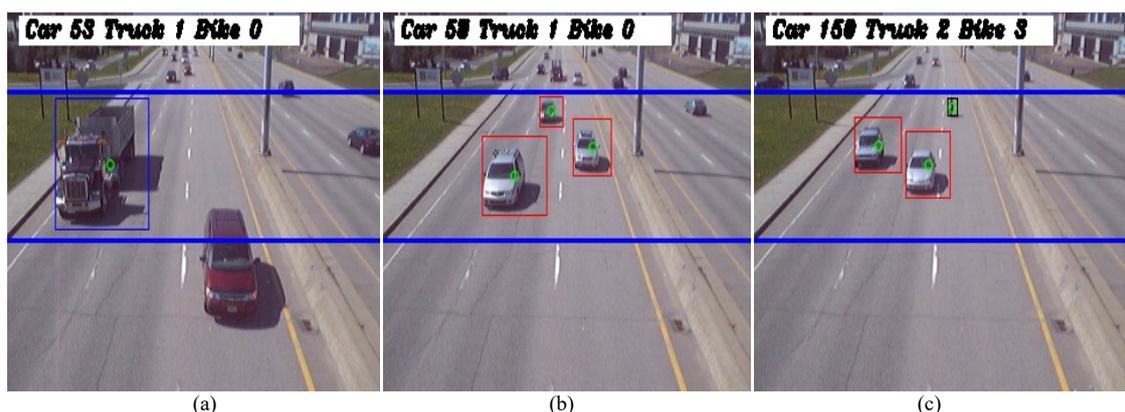

Figure 8. Results of the proposed system
802th frame (b) 825th frame (c) 1776th frame

4.2. Accuracy of traffic counting

To evaluate the robustness and efficiency of our system, we conducted an analytic study of the counting accuracy of vehicles. The measures are performed twice in our system, before and after the treatment of occlusions. We have noticed that the occlusion processing technique had a significant impact on vehicle counting accuracy. Indeed, the accuracy has increased from 85.05% before occlusion processing to 99.45% after processing. Table 2 shows the results of vehicle counting before and after occlusion processing.

Table 2. Results of vehicle counting before and after occlusion processing

	Vehicle class	Count	Error	Accuracy
Vehicle counting rate before occlusion processing	Car	165	19	88.48%
	Bike	5	0	100%
	Truck	3	1	66.67%
	Average			**85.05%**
Vehicle counting rate after occlusion processing	Car	184	3	98.37%
	Bike	5	0	100%
	Truck	4	0	100%
	Average			**99.45%**

4.3. GRAM Road-Traffic Monitoring dataset (GRAM-RTM)

GRAM Road-Traffic Monitoring (GRAM-RTM) (Guerrero-Gomez-Olmedo *et al*., 2013) is a dataset designed for real time tracking vehicle systems. It consists of 3 video sequences, recorded under different conditions. The first video has been recorded in a sunny day; it contains 7520 frames with a resolution of 800x480. The second video (9390 frames), has been recorded in a similar location but during a cloudy day at 1200x720. The third video sequence (23435 frames), has been recorded in a busy intersection with a resolution of 600x360.

Another test of the accuracy of system was performed based on GRAM-RTM dataset. The surveillance area will be defined using a mask (Fig. 9). The surveillance zone where the system will operate is shown in Figure 10.

Figure 9. Mask used to define the surveillance zone

Figure 10. Surveillance zone

In two different video sequences of GRAM-RTM dataset, the method used was able to classify the detected objects to four types: Car, bike, vertical and horizontal occlusion (Table 3). The first two classes will be processed immediately within the counting module. The two remaining classes which present the occlusions will be treated and classified to handle these occlusions.

Table 3. Determined classes and their appropriate images

Classes	Class 1	Class 2	Class 4	Class 5
Image				
Classification result	Car	Bike	Vertical occlusion	Horizontal occlusion

Figure 11 presents the results obtained during the execution of the realized system. The detection rectangles emitted take two colours relatively to the type of the detected vehicle. The colour of the detection rectangle around each vehicle changes depending on the type of the detected vehicle. Counting results are displayed in a white bar above the image sequence indicating the current number reached by the traffic flow for each vehicle type.
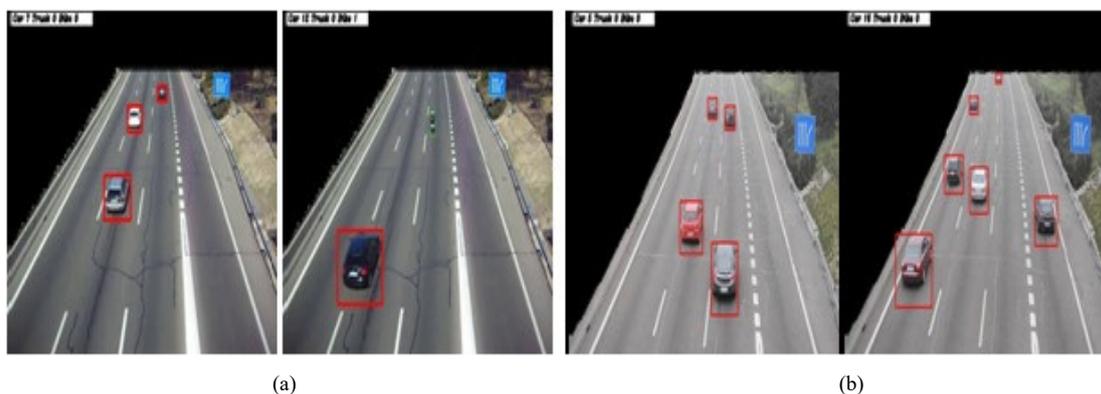
<div align="center">(a)</div> <div align="center">(b)</div>

Figure 11. Results of system test using GRAM-RTM dataset (a) Scene 1, (b) Scene 2

4.4. Comparative study

A Comparative study on different road traffic video surveillance systems is developed in Table 4. This study was conducted to compare the counting efficiency of our system to that measured in other existing systems.

Table 4. Results of the comparative study of the proposed system and some existing systems

Comparative methods	(Huang *et al*., 2012)	(Lei *et al*., 2008)	(Pornpanomchai *et al*., 2008)	(Mohana *et al.*, 2009)	Proposed system
Type studied of vehicles	Cars and bikes	Cars	Cars	Cars	Cars, Bikes and trucks
Segmentation method	Background subtraction	Background subtraction	Background subtraction	Optical flow	Background subtraction
Classification method	Feature extraction	X	X	X	HOG + SVM
Vehicle counting accuracy	96.9%	Scene 1: 90.41% Scene 2: 70.31% Scene 3: 98.39% Scene 4: 80.00%	Afternoon: 91.98% Evening : 96.35%	94.04%	99.45%

5. Conclusion

Recent video surveillance systems of road traffic are subject to certain limitations due to vehicle occlusion, traffic congestion, and illumination conditions. In this paper, we presented a video surveillance system for vehicle detection and classification, which is able to process vertical and horizontal occlusions. The proposed method starts with a motion detection phase based on background subtraction using Codebooks, followed by a set of filters in order to eliminate any kind of noise and generate a mask containing the regions of interest in the form of rectangles. The second module in our system is designed to specify the type of the detected vehicle (car, bike, and truck) using HOG descriptor and SVM classifier. Using also HOG and SVM, a technique for the detection of partial occlusions was developed. This technique can detect occluded vehicles and process them subsequently by acting on the shape of the detected object. The experimental section shows that our method can achieve better counting accuracy in comparison with other existing systems. As perspectives, we look forward to resolve the problem of the shadow of vehicles which may also influence the vehicle counting efficiency, and consequently the robustness of the entire system.

References

1. Aslani, S. and Mahdavi-Nasab, H. (2013) Optical flow based moving object detection and tracking for traffic surveillance. *International Journal of Electrical, Electronics, Communication, Energy Science and Engineering,* 7, 789-793.
2. Atiq, H. M., Farooq, U., Ibrahim, R., Khalid, O. and Amar, M. (2010) Vehicle detection and shape recognition using optical sensors: a review. In *Machine Learning and Computing (ICMLC)*, 2010 Second International Conference on, IEEE, pp. 223-227.
3. Canny, J. (1986) A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, pp. 679-698.

4. Chen, C. and Liu, S. (2012) Detection and segmentation of occluded vehicles based on skeleton features. In *Instrumentation, Measurement, Computer, Communication and Control (IMCCC)*, 2012 Second International Conference on (pp. 1055-1059). IEEE.

5. Coifman, B., Beymer, D., McLauchlan, P. and Malik, J. (1998) A real-time computer vision system for vehicle tracking and traffic surveillance. *Transportation Research Part C: Emerging Technologies*, 6, 271-288.

6. Cortes, C. and Vapnik, V. (1995) Support-vector networks. *Machine Learning*, 20, 273-297.

7. Dalal, N. and Triggs, B. (2005) Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition*, 2005. CVPR 2005. IEEE Computer Society Conference on (pp. 886-893). IEEE volume 1.

8. Fu, H., Ma, H., Liu, Y. and Lu, D. (2016) A vehicle classification system based on hierarchical multi-svms in crowded traffic scenes. *Neurocomputing*, 211, 182-190.

9. Guerrero-Gomez-Olmedo, R., Lopez-Sastre, R.J., Maldonado-Bascon, S. and Fernandez-Caballero, A. (2013) Vehicle tracking by simultaneous detection and viewpoint estimation. In *IWINAC 2013*, *Part II, LNCS 7931* (pp. 306-316).

10. Horprasert, T., Harwood, D. and Davis, L. S. (1999) A statistical approach for real-time robust background subtraction and shadow detection. In *IEEE ICCV* (pp. 1-19). Volume 99.

11. Hossen, M. K. and Tuli, S. H. (2016) A surveillance system based on motion detection and motion estimation using optical flow. In *Informatics, Electronics and Vision (ICIEV)*, 2016 5th International Conference on (pp. 646-651). IEEE.

12. Huang, C.-L. and Liao, W.-C. (2004) A vision-based vehicle identification system. In Pattern Recognition, 2004. ICPR 2004. *Proceedings of the 17th International Conference* on (pp. 364-367). IEEE volume 4.

13. Huang, D.-Y., Chen, C.-H., Hu, W.-C., Yi, S.-C., Lin, Y.-F. *et al*. (2012) Feature-based vehicle flow analysis and measurement for a real-time traffic surveillance system. *Journal of Information Hiding and Multimedia Signal Processing*, 3, 279-294.

14. Ji, X., Wei, Z. and Feng, Y. (2006) Effective vehicle detection technique for traffic surveillance systems. *Journal of Visual Communication and Image Representation*, 7, 647-658.

15. Johansson, B., Wiklund, J., Forssen, P.-E. and Granlund, G. (2009) Combining shadow detection and simulation for estimation of vehicle size and position. *Pattern Recognition Letters*, 30, 751-759.

16. Kanhere, N. K. and Birchfield, S. T. (2008) Real-time incremental segmentation and tracking of vehicles at low camera angles using stable features. *IEEE Transactions on Intelligent Transportation Systems*, 9, 148-160.

17. Kastrinaki, V., Zervakis, M. and Kalaitzakis, K. (2003) A survey of video processing techniques for traffic applications. *Image and Vision Computing*, 21, 359-381.

18. Kim, K., Chalidabhongse, T. H., Harwood, D. and Davis, L. (2005) Real-time foreground-background segmentation using codebook model. *Real-Time Imaging*, 11, 172-185.

19. Lei, M., Leoch, D., Gouton, P. and Madani, K. (2008) A video-based real-time vehicle counting system using adaptive background method. In *Signal Image Technology and Internet Based Systems, 2008. SITIS'08. IEEE International Conference* on (pp. 523-528). IEEE.

20. Mandellos, N. A., Keramitsoglou, I. and Kiranoudis, C. T. (2011) A background subtraction algorithm for detecting and tracking vehicles. *Expert Systems with Applications*, 38, 1619-1631.

21. Mohana, H., Ashwathakumar, M. and Shivakumar, G. (2009) Vehicle detection and counting by using real time traffic ux through differential technique and performance evaluation. In *Advanced Computer Control, 2009. ICACC'09. International Conference* on (pp. 791-795). IEEE.

22. Nicolas, H. and Brulin, M. (2014) Video traffic analysis using scene and vehicle models. *Signal Processing: Image Communication*, 29, 807-830.

23. Pornpanomchai, C., Liamsanguan, T., and Vannakosit, V. (2008) Vehicle detection and counting from a video frame. In *Wavelet Analysis and Pattern Recognition, 2008. ICWAPR'08. International Conference* on (pp. 356-361). IEEE volume 1.

24. Shu, G., Dehghan, A., Oreifej, O., Hand, E. and Shah, M. (2012) Part-based multiple-person tracking with partial occlusion handling. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference* on (pp. 1815-1821). IEEE.

25. Song, X. and Nevatia, R. (2007) Detection and tracking of moving vehicles in crowded scenes. In *Motion and Video Computing, 2007. WMVC'07. IEEE Workshop* on (pp. 4{4). IEEE.

26. Stauffer, C. and Grimson, W. E. L. (1999) Adaptive background mixture models for real-time tracking. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference* on. (pp. 246-252). IEEE volume 2.

27. Tang, S., Andriluka, M. and Schiele, B. (2014) Detection and tracking of occluded people. *International Journal of Computer Vision*, 110, 58-69.

28. Tian, B., Tang, M. and Wang, F.-Y. (2015) Vehicle detection grammars with partial occlusion handling for traffic surveillance. *Transportation Research Part C: Emerging Technologies*, 56, 80-93.

29. Wang, C.-C. R. and Lien, J.-J. J. (2008) Automatic vehicle detection using local features a statistical approach. *IEEE Transactions on Intelligent Transportation Systems*, 9, 83-96.

30. Wen, X., Shao, L., Xue, Y. and Fang, W. (2015) A rapid learning algorithm for vehicle classification. *Information Sciences*, 295, 395-406.

31. Xia, Y., Shi, X., Song, G., Geng, Q. and Liu, Y. (2016) Towards improving quality of video-based vehicle counting method for traffic flow estimation. *Signal Processing*, 120, 672-681.

32. Yu, M., Jiang, G. and Yu, B. (2000) An integrative method for video based traffic parameter extraction in its. In *Circuits and Systems, 2000. IEEE APCCAS 2000. The 2000 IEEE Asia Pacific Conference* on (pp. 136-139). IEEE.

33. Zhou, Y. and Cheung, N. M. (2016) Vehicle Classification using Transferable Deep Neural Network Features. arXiv, 2016, vol. 1601.