Large Vehicle Recognition and Classification for Traffic Management and Flow Optimization in Narrow Roads

D. Triantafyllou, N. Kotoulas, S. Krinidis, D. Ioannidis, D. Tzovaras
Information Technologies Institute
Center for Research & Technology Hellas
Thessaloniki, Greece
{dtriant, kotoulas, krinidis, djoannid, dimitrios.tzovaras}@iti.gr

***Abstract*—A traffic management system for narrow roads that utilizes RGB cameras is proposed in this paper. The system is able to detect and recognize any large vehicles trying to cross narrow roads and control the traffic accordingly, for example prevent two large vehicles from entering the road from different directions simultaneously. The system is based on adaptive methodologies so that it can adjust to changing conditions of lighting. Moreover, a calibration technique combined with a fine tuning procedure overcome the limitations of a predefined camera setting. The goal of this paper is a system that optimizes the flow at the roads under monitoring with the use of traffic lights or VMS (variable-message signs). To our knowledge, there are no similar systems aiming at real time traffic management of narrow roads.**

***Index Terms*—traffic control, large vehicle detection and tracking, narrow roads**

I. Introduction

Increasing urban population has led to increasing traffic problems in the cities and a pressing need for incorporation of advanced technologies to traffic control. Older parts of cities often comprise narrow roads, which were initially sufficient for the transportation needs; nevertheless, due to increasing traffic, nowadays they constitute a severe cause of congestion. The problem augments when large vehicles, i.e. trucks and buses, attempt to cross these roads. Moreover, current traffic light management is usually based on timers lacking any intelligence.

In this paper a “smart” traffic management system for flow optimization in narrow roads is proposed. Its goal is the recognition of large vehicles, such as trucks and buses, so that traffic congestion in narrow roads is avoided. For example,in cases where the width of the road is suitable for only one large vehicle, traffic lights or VMS at the entrances of the road detect through the proposed system any buses or trucks and, if it is necessary, manage the circulation to avoid any unpleasant consequences of congestion.

This paper presents a full chain of all the necessary processes for the recognition of large vehicles, i.e. background extraction, blob detection, shadow extraction, tracking and classification of moving objects. The utilized methodologies are adaptive,such as VIBE [1], so that the system adjusts to changing lighting and background conditions. Furthermore classification is based on a calibration procedure that permits the extraction of moving objects dimensions on the road plane of any moving item. The system utilizes RGB cameras while its advantage is that the cameras’ setting can vary. To our knowledge, despite the prolific literature on vehicle recognition and classification, there are no similar systems aiming at real time traffic management of narrow roads.

The remainder of the paper is organized as follows. Section II presents related work while in section III there is an overview of the proposed system. Section IV describes the camera calibration procedure, section V the detection and tracking methodology and in section VI the classification of the tracked items is analysed. Finally, the paper concludes with experimental results in section VII and conclusions in section VIII .

II. Related work

Traffic monitoring has attracted the last few years the attention of many computer vision researchers leading to a variety of methodologies. Usually, the main concepts that preoccupy researchers are vehicle detection, tracking and sometimes classification.

Vehicle detection methodologies can be divided in three main categories according to [2] : 1) background subtraction, 2) frame differencing and motion based methods and 3) feature-based methods. In background subtraction methods there are several efforts to obtain an adaptive background. In [3], [4] Gaussian probability distribution of each pixel is utilized to achieve adaptability while in [5] a filtering method based on histograms is proposed. Frame differencing and motion based methodologies subtract subsequent frames to extract foreground or isolate blobs by analysing the orientations and speed of movements of sets of pixels [6], [7]. Furthermore, feature based methods use features such as Haar wavelets [8], edges [9] and Gabor filters [10] or PCA analysis [11] to achieve vehicle detection.

Tracking constitutes an essential part of traffic monitoring. Several approaches have been used by researchers utilizing criteria such as region and distances [12],contours [13], 3D-models [14], color [15] and features such as edge points and modified SIFT descriptors [16].

Classification is an important stage in many traffic monitoring applications. Several features such as Haar wavelets [8] and shape [17] are used to distinct the moving objects into categories while support vector machines [8], neural networks [18] and higher order statistics [19] perform the classification task.

III. TRAFFIC CONTROL SYSTEM OVERVIEW

The most important part of the proposed traffic control system is the accurate detection and classification of vehicles. The main idea to fulfil this task is the detection of vehicles in the images as they are provided by the traffic camera and their classification based on their dimensions when they are transformed to "bird-eye" view.

The aforementioned procedure includes several sub-tasks that have to be accomplished correctly in order to achieve robustness. First of all, a calibration method is deployed to allow the transformation of the view under monitoring to "bird-eye-view". Once this procedure is concluded an iterative stage begins (Figure 1). Each acquired image undergoes dynamic background subtraction, blob extraction, shadow removal, vehicle tracking, transformation to "bird-eye" view and classification of the detected items. Since the traffic light system should be able to operate under any weather conditions, the implemented algorithms are adaptive and robust to different light conditions. Moreover, since the camera might be installed with different configurations the proposed methodology is tolerant to various settings.

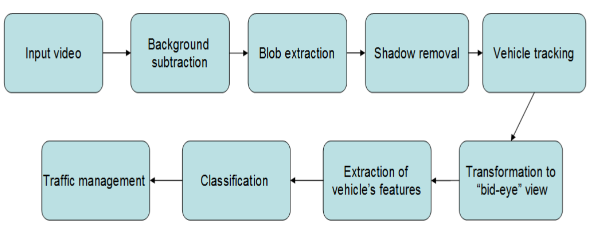


Fig. 1. Traffic control system overview.

IV. CAMERA CALIBRATION PROCEDURE

The transformation of an image and its context to "bird-eye-view" requires an initial calibration procedure. The goal is to acquire correct calibration data through an easy procedure and without traffic interruptions. Therefore, based on epipolar geometry theory, a rectangular plane located on the road with a priori known dimensions is utilised for the calibration. Given the coordinates of its four corners on the camera image and its real world dimensions calibration data are extracted. Nevertheless, the acquired data refer only to the road plane and do not include 3D information.

In Figure 2 the calibration procedure is depicted. Four points that form a rectangle of known dimensions are chosen on the road plane (Figure 2.a). Based on this information the camera image is transformed to "bird-eye-view" (Figure 2.b). In cases where physical measurement of the rectangles dimensions is not possible, existing references could be used as an alternative to achieve similar results, e.g. there is legislation that determines a traffic lane's width and other distances referring to road marks, vehicle size etc.

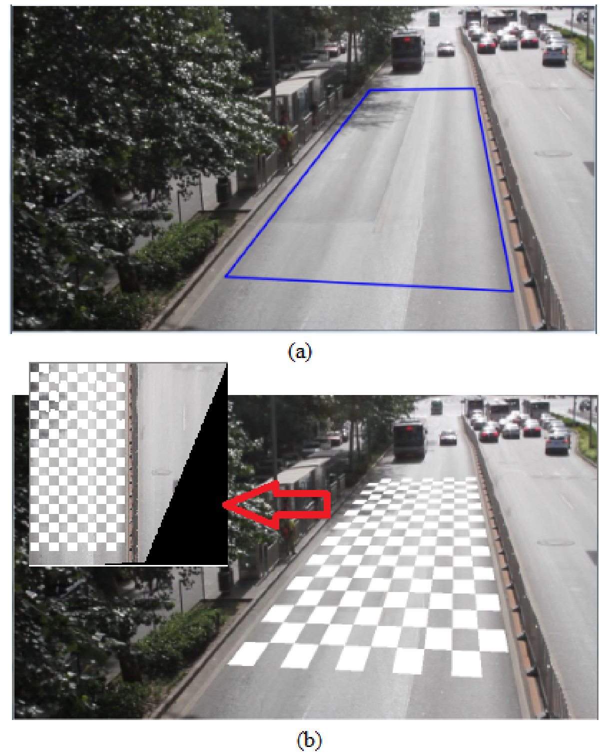


Fig. 2. Calibration procedure: (a) four points forming a rectangular are selected on the road plane, (b) transformation of the selected plane to "bird-eye-view". Notice the correspondence between the chessboard as it is depicted on the initial and "bird-eye-view" image.

V. DETECTION AND TRACKING PROCEDURE

Once calibration is completed, an iterative procedure comprising detection of moving objects and tracking commences.

A. Detection of moving objects

The first processing step on an image by the proposed algorithm is background subtraction so that blobs of interest can be detected. ViBe [4] algorithm is considered suitable for the application since its updating mechanism makes it adjustable to changing conditions of lighting. Moreover, it can incorporate rather quickly any long lasting changes of the background.

Often, blobs that occur from background subtraction constitute noise or smaller parts of a moving item that need to be merged. Therefore, small blobs are deleted while bigger ones are merged with each other if they fulfil certain proximity criteria. To determine the distance between two blobs two

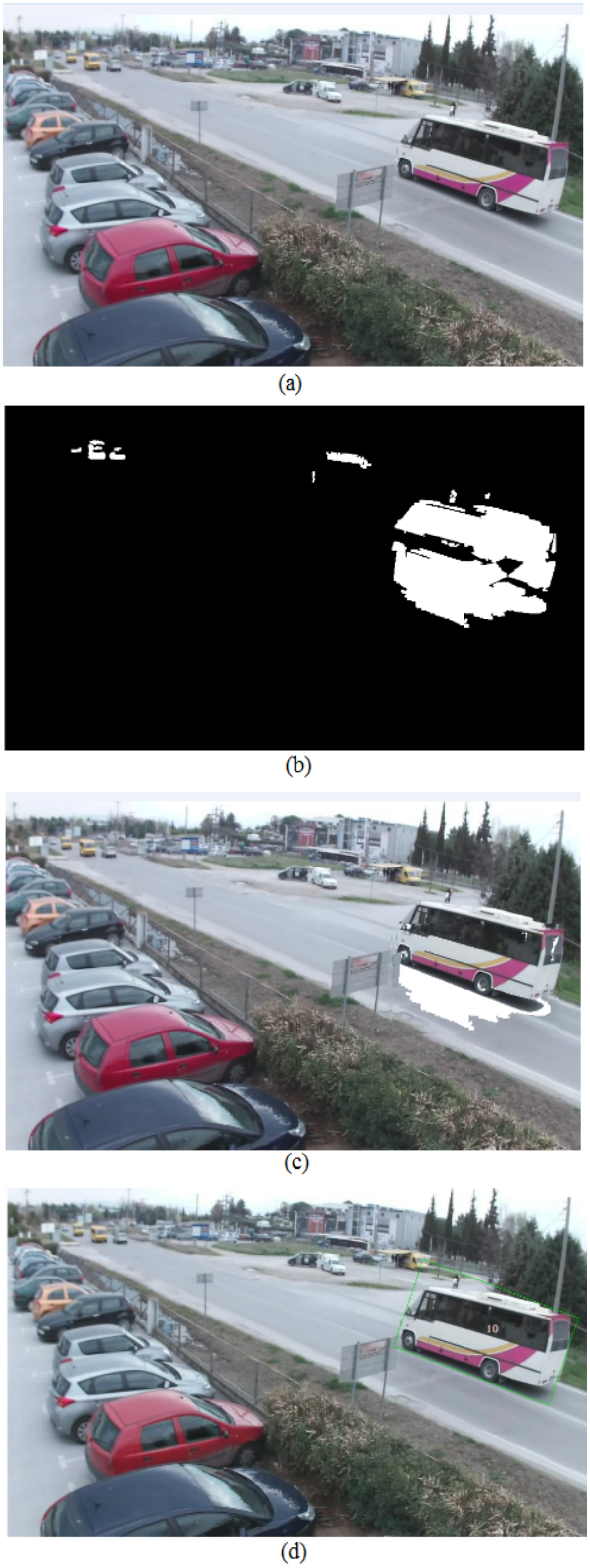


Fig. 3. Vehicle detection and shadow removal procedure:(a) initial image of a bus, (b) the detected blobs, (c) the detected shadow is depicted with white color, (d) the final result of the detection and shadow extraction procedure is marked with a green rectangle.

rectangles that characterize a blob are defined: 1) a rectangle enclosing the blob with edges parallel to the image frame, 2) a rectangle enclosing the blob with the minimum possible area. The combination of these two rectangles captures sufficiently the blobs' shape, hence, they are used to calculate the distance separating two blobs. In particular, the distance between two blobs is defined as:

$$d = min(max(D_p), max(D_m)), \tag{1}$$

where D_p and D_m refer to the distances of the parallel to the image and minimum area rectangles respectively. Thus, to merge two blobs their distance d should be smaller than a predefined threshold.

It has to be mentioned that this solution is preferred over calculating the distance between blobs' centroids due to the variety of the vehicles' sizes.

B. Shadow removal

Vehicles often create shadows that are included in the detected blobs, enlarge their size and could affect their classification leading to false negatives. Thus, a shadow removal technique is utilized [20]. In particular, the RGB image of the monitored area is transformed to the HSV color space that is closer to human perception of color. Comparing to the background information and utilizing predefined thresholds for each HSV channel shadow areas of the blobs are extracted. In order to avoid removing pixels that belong to the detected vehicle and not its shadow, spatial averaging facilitates the separation of the blob in two parts: shadow and vehicle.

Figure 3 depicts the steps of blob detection and shadow removal procedure. In Figure 3.b the blobs extracted from background subtraction are illustrated while Figure 3.c depicts with white color the detected shadow of the bus. In Figure 3.d the shadow has been removed and the detected bus is marked by a green rectangle.

C. Tracking of moving objects

Tracking is based on distance criteria. During blob matching priority is given to bigger blobs since in this implementation the correct detection and tracking of large vehicles is highly important for traffic management. In cases of cars passing by or crossing each other partial or total occlusion might occur. To deal with these events tracking information combined with Kalman filters for position prediction is utilized so that tracking of vehicles is not interrupted.

VI. Classification

Detected vehicles are classified in two major categories: 1) small/average sized vehicles, that include cars (sedan, SUV), motorcycles, 2) large vehicles, such as big trucks, vans and buses. The criteria of the classification is the width and hight of the minimum area rectangle of the tracked item. Since the original image cannot provide correct and steady dimensions of the tracked vehicles, they are transformed to "bird-eye-view". At this point one would think that the classification

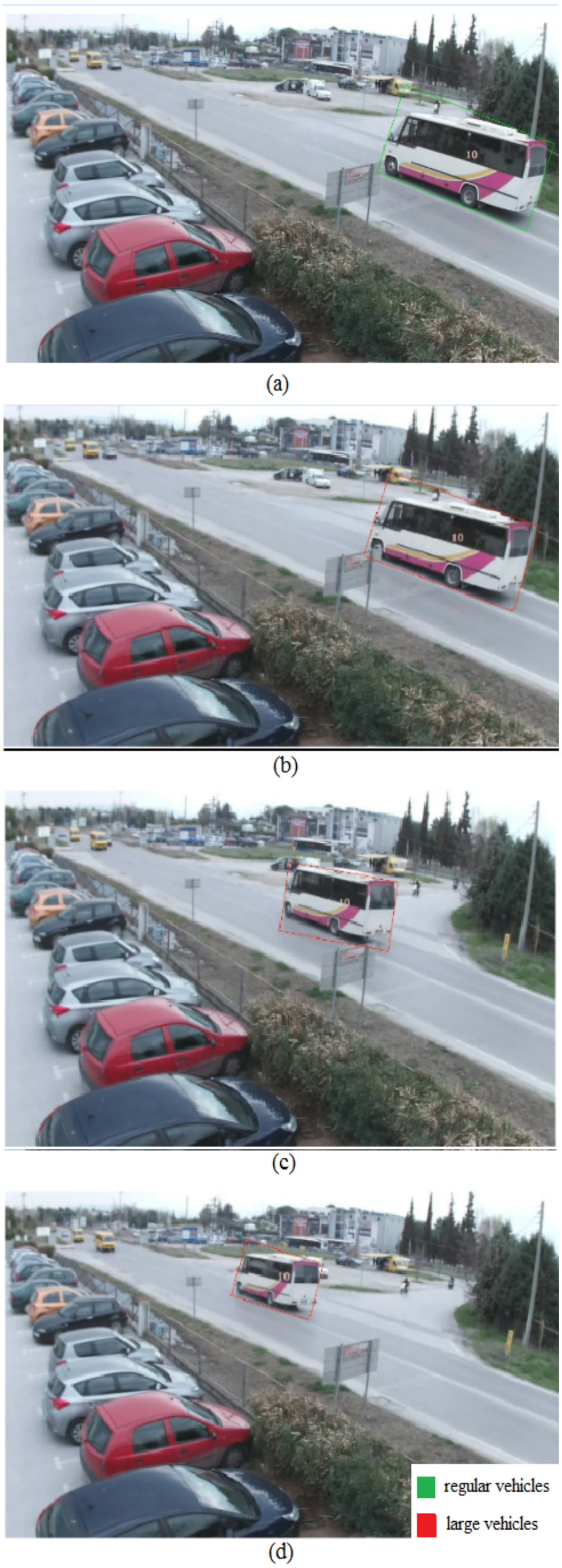


Fig. 4. Vehicle tracking and classification: (a) initially, the bus is characterized as simple car since it is affected by previous frames where it was not fully visible, (b-d) the bus now is classified as big vehicle.

could be easily made utilizing average values of the two vehicle categories as they are determined by their manufacturers. Nevertheless, the changing view of the vehicle that might occur as it proceeds in combination with some deformation from "bird-eye-view" transformation (as it is mentioned in section IV "bird-eye-view" transformation refers only to the road plane)dictate the need for a fine-tuning procedure that will determine the thresholds that define each category's width and height. To achieve this, once the calibration procedure is completed, few samples of the first category are utilized to refine the width and height values that will be used to separate the two categories.

Finally, an object's class is determined through a voting procedure. As frames proceed, the mean value of the tracked vehicle's "bird-eye-view" hight and width defines the per frame vote for the object's class. Once the object disappears from the image the final decision for the vehicle's classification is made.

VII. Experimental Results

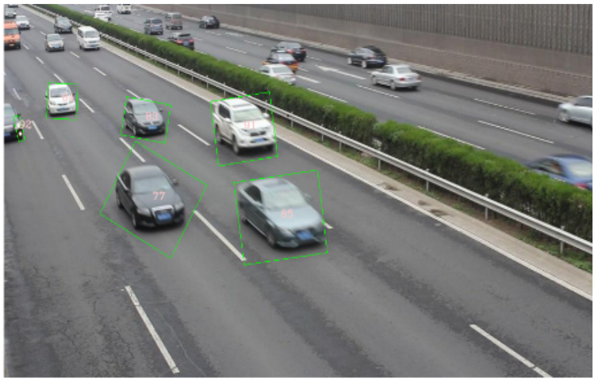
The proposed methodology was tested in more than 22000 images of UA-DETRAC dataset[1] [21] detecting and tracking approximately 3000 vehicles.

Since the goal of the methodology proposed in this paper is to manage traffic lights according to the existence of large vehicles, attention must be paid to missed large vehicles or false positives that have occurred. Particularly, there was only one case of failing to detect a large vehicle. In the particular case, the bus made a stop in the middle of its route, VIBE absorbed it to the background and when it started moving again there was not sufficient space left to allow its correct recognition. Furthermore, false positives occurred from cases where two or more cars were moving very close to each other for the majority of the frames that captured them, heading to the same direction with similar speed. This situation created a huge blob for several frames leading to miss-classification. Nevertheless, these are cases that are not in the scope of the proposed application since it refers to narrow roads.

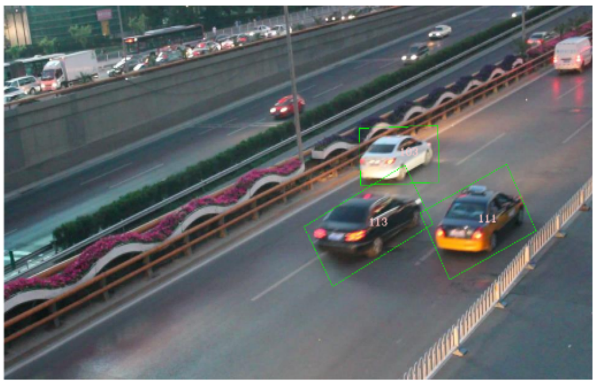
In Figure 5 detection and classification examples with different camera settings and light conditions are illustrated. Furthermore, the performance of the proposed application regarding the final classification of large vehicles is given in table I. Precision is 0.93 , recall is 0.68 while F1-score is 0.79. The low recall value in comparison to the precision can be explained by the difference of the large and regular vehicles that usually exist in a road, i.e. there are few large vehicles to provide true positives in contrast to the large number of regular cars that can cause false negatives.

Furthermore, the proposed method was also tested in 15000 images of a dataset collected in a narrower road with two lanes, one for each direction (Figure 4). Once the case of two

[1]UA-DETRAC dataset includes a lot of wide roads and big intersections in urban environments in contrast to the application scenario of our methodology. Therefore, overcrowded cases near traffic lights with several cars deployed next to each other are excluded from our test set since they are completely out of scope.

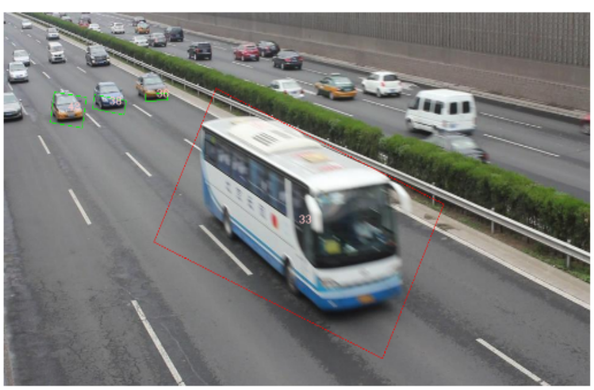


(a)



(b)

(c)



(d)

Fig. 5. Classification examples. Large vehicles are marked with red color while smaller cars with green.

vehicles moving in parallel lanes was eliminated, the results were highly improved. In particular precision reached the value of 1, since all large vehicles were recognized, recall 0.95 is while F1-score is 0.97 (table I). Nevertheless, evaluation in a more extended dataset is required and planned as future work. In addition, a frame rate of 8.7 fps is achieved so that the proposed methodology performs in real time. Future plans include a more parallel and gpu oriented version of the application aiming at higher performance.

TABLE I
PERFORMANCE

	Precision	Recall	F1 - score
online dataset	**0.94**	**0.68**	**0.79**
narrow road dataset	**1**	**0.95**	**0.97**

VIII. CONCLUSIONS

A system for large vehicle recognition in order to achieve traffic management and flow optimization in narrow roads is proposed in this paper. The full chain of the recognition procedure from the background subtraction, to the detection, tracking and classification is analysed. The method manages to adjust in different light conditions while it incorporates rapidly any background changes. Moreover, the proposed combination of calibration with a fine tuning procedure permits different camera settings. The results acquired from the experiments are promising since false positives or misses mainly refer to situations, such as bus-stop or multiple cars moving in parallel in close distance for many frames, that do not occur in narrow roads.

REFERENCES

[1] M. Van Droogenbroeck, O. Barnich, "A Disruptive Method for Background Subtraction,"In T. Bouwmans, F. Porikli, B. Hoferlin, and A. Vacavant, editors, Background Modeling and Foreground Detection for Video Surveillance, chapter 7. Chapman and Hall/CRC, June 2014.

[2] M. Van Droogenbroeck, O. Barnich, "A Disruptive Method for Background Subtraction,"Signal and Image Processing : An International Journal (SIPIJ) Vol.5, No.1, 2014.

[3] G. Xiang, T.E. Boult, F. Coetze and V. Ramesh, "Error Analysis of Background Adaption ," Computer Vision and Pattern Recognition, pp. 503–510, 2000.

[4] I. Haritaoglu , D. Harwood and L. S. Davis, "W4: Real-time surveillance of people and their activities,",IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, pp.809–830.

[5] N. Mandelos, I. Keramitsoglou and C. Kiranoudis, "A background subtraction algorithm for detecting and tracking vehicles," in Expert Systems with Applications, vol. 38, 2011, pp. 1619–1631.

[6] L. Vasu, "An effective step to real-time implementation of accident detection system using image processing," in Master of Science, Oklahoma State University, USA, 2010.

[7] W. Zhang and Q.M. Jonathan "Moving vehicles detection based on adaptive motion histogram," in Digit. Signal Process., vol. 20, pp. 793–805, 2010.

[8] C. P. Papageorgiou and T. Poggio "A Trainable System for Object Detection in Images and Video Sequences," Massachusetts Institute of Technology Center for Biological and Computational Learning, vol. 1673., 2000.

[9] Z. Tao and R. Nevatia "Car detection in low resolution aerial image," in Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on, 2001, pp. 710–717 vol.1.

[10] T. M. Deng and B. Li "A Detection Method of Traffic Parameters Based on EPI," Procedia Engineering, vol. 29, pp. 3054–3059, 2012.

[11] W. Chi-Chen Raxle and J. J. Lien "Automatic Vehicle Detection Using Local Feature," in Intelligent Transportation Systems, IEEE Transactions on, vol. 9, pp. 83–96, 2008.
[12] W. Chi-Chen Raxle, S. Huang and C. Tseng "Automatic Vehicle Detection Using Local Feature," in Green Circuits and Systems (ICGCS), 2010 International Conference on, 2010, pp. 666–670.
[13] R. Rad and M. Jamzad "Real time classification and tracking of multiple vehicles in highways," in Pattern Recognition Letters, vol. 26, pp. 1597–1607, 2005.
[14] F. Bardet T. Chateau and D. Ramadasan "Unifying real-time multi-vehicle tracking and categorization," in Intelligent Vehicles Symposium, 2009 IEEE, 2009, pp. 197–202.
[15] H. Mao-Chi and Y. Shwu-Huey "A real-time and color-based computer vision for traffic monitoring system," in Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on, 2005, pp. 1185–1192, Vol. 2.
[16] M. Xiaoxu and W. E. L. Grimson "Edge-based rich representation for vehicle classification," in Intelligent Vehicles Symposium, 2009 IEEE, 2009, pp. 197–202.
[17] J. Lipton, H. Fujiyoshi, and R.S. Patil "Moving target classification and tracking from real-time video," in Workshop Applications of Computer Vision, pages 129-136, 1998.
[18] D. M. Xa, J.-M. Lee and Y. D. Kim "Neural-edge-based vehicle detection and traffic parameter extraction," in Image and Vision Computing, vol. 22, pp. 899–907, 2004.
[19] A.N. Rajagopalan, P. Burlina and R. Chellappa "Higher order statistical learning for vehicle detection in images," in Computer Vision, 1999, 1999, pp.1204–1209 vol.2.
[20] E. A. Mosabbeb, H. Abbasian and M. Fathy "Clustering Based Non-parametric Model for Shadow Detection in Video Sequences ," in IPCV, 2009.
[21] UA-DETRAC challenge , http://detrac-db.rit.albany.edu/download