ISSN 2229-5518

Multi-Vehicle Tracking Under Day and Night Illumination

Mohamed Taha [1], Hala H. Zayed[1], Taymoor Nazmy[2] and M. E. Khalifa[2]
[1] Computer Science Dept. Faculty of Computers & Informatics, Benha University.
[2] Faculty of Computer & Information Sciences, Ain Shams University.
{mohamed.taha, hala.zayed}@fci.bu.edu.eg, ntaymoor19600@gmail.com, esskhalifa@cis.asu.edu.eg

Abstract— Traffic surveillance plays a vital role in computer vision and Intelligent Transportation Systems (ITS). Image analysis provides several effective techniques to detect moving objects in images. Thus, it has been extensively used for traffic monitoring systems. Recently, the problem of detecting and tracking vehicles is an important emerging research area for intelligent transportation systems. Many algorithms have been developed to detect and track moving vehicles either in daytime or in nighttime. In fact, vehicle tracking in daytime and in nighttime cannot be approached with the same techniques, due to the extreme different illumination conditions. Building an integrated system to deal with daytime and nighttime is still a challenging problem especially when considering shadows at daytime, dim lighting at night, and real-time processing constraint.

In this paper, a vehicle tracking system is developed to deal with daytime and nighttime vehicles tracking. First, a daytime/nighttime detector is applied to the scene to determine the suitable technique. For daytime videos, shadows are removed from vehicles by applying a gamma decoding followed by a thresholding operation and employing an estimated background model of the video sequence. For nighttime videos, headlights and taillights are located and paired to initialize vehicles for tracking process. The experimental results have shown that the proposed method can effectively track vehicles in both daytime and nighttime..

Index Terms— Traffic Surveillance, Daytime/Nighttime Surveillance, Vehicles Tracking, Vehicles Detection, Daytime/Nighttime Tracking, Multi Objects Tracking.

1 Introduction

As the amount of vehicles on the roads is continuously increasing, the demand for accurate traffic monitoring and roadway surveillance systems is rising. Most of previous traffic surveillance systems are generally known to involve costly hardware equipment (loop detectors, ultrasonic detectors, microwave sensors, radar sensors and slit sensors) with complicated operation procedures [1]. Moreover, these hardware based systems are limited to the number of vehicles passing through the detection regions and They are difficult to apply for vehicle classification, vehicle speed detection, and vehicle motion analysis [2]. Problems associated with these systems have generated an interest in new vehicle detection technologies such as the use of computer vision and video image processing techniques. Compared to traditional sensors, video-based systems can capture a larger variety of desired information and less disturbing than other kinds of devices like loop detectors [3]. In general, computer vision systems are more flexible, more precise, more robust, easier to maintain, and less expensive to install [4].

The generic automatic surveillance from a single camera usually follows four steps: modeling of environments, detection of motion, classification of moving objects, and tracking. Sometimes, there is a fifth step to understand and analyze behaviors. In traffic surveillance, the camera is fixed above a road; therefore, the background seen in the camera images is nearly static. Vehicles tracking is an important step in any visual-based traffic surveillance systems and represents a challenging task for researchers. The difficulties in detecting and tacking vehicles reside into two main issues [5]: (1) the intra variations of vehicles including length, width, shape, color, etc. (2) the challenges of dynamic outdoor environment including illumination changes (day, night, twilight, cloudy, sunny), shadows, wind (shaking the camera), weather changes (rain, snow, fog), moving background (trees), etc. The feature extraction, recognition, and tracking are severely affected by shadowy and sunny locations, dim lighting at night, loss of color on a cloudy day, highlights on vehicles, and so on [6].

Despite being a recent active research, vehicle-tracking systems lack the flexibility to be used in real environments. Adaptation to various illumination conditions is an essential factor for practical use in these environments. The fundamental problem here is to identify vehicles in changing environment and illumination. Although there have been numerous techniques on general object recognition and tracking, not many of them could successfully be applied in both daytime and nighttime. Figure 1.a shows a merging problem of tracking two moving objects

- *Author name is currently pursuing masters degree program in electric power engineering in University, Country, PH-01123456789. E-mail: author_name@mail.com*
- *Co-Author name is currently pursuing masters degree program in electric power engineering in University, Country, PH-01123456789. E-mail: author_name@mail.com*
(*This information is optional; change it according to your need.)*

ISSN 2229-5518

during daytime. It occurs as a result of the fusion of the right vehicle shadow with the left vehicle. Cast-shadow can bring serious problems while extracting moving objects due to the misclassification of shadow points as foreground. In some cases when the shadows stretch, two or more independent objects can appear to be connected together. On the other hand, figure 1.b shows a nighttime traffic scene. Note that the image has very low contrast and a weak light sensitivity. Furthermore, there are strong reflections on the roads surface, which complicate the problem. The moving reflections of the headlights can introduce many foreground or background ambiguities. Under bad-illuminated condition in the nighttime road environment, the obvious features of vehicles, which are effective for detecting in daytime, become invalid in nighttime road environment.

This paper extends our previous work in [7] and [8] by integrating them into one comprehensive system. The paper proposes a vehicle tracking system for both daytime and nighttime. Under daytime conditions, the proposed system employs the estimated background model of the video sequence and applies a Gamma decoding followed by a thresholding for removing cast shadows from vehicles in outdoor environments. Additionally, to detect vehicles at nighttime road scenes, the proposed system extracts bright objects using connected components analysis and recognizes the paired lights of the vehicles using rule based component analysis approach. Also, a daytime/nighttime detector is first employed to determine whether the scene is during daytime or nighttime so that the suitable technique is applied.

(a) daytime scene (b) Nighttime scene

Fig.1. Two examples of traffic scenes

For clarity of presentation, the paper is organized as follows: Section 2 presents a review of literature in vehicle tracking during daytime and nighttime. Section 3 presents the proposed system in detail. Section 4 discusses the experimental results and the performance evaluation of the proposed method. Finally, the conclusion is presented in Section 5.

2 Related Work

Tracking of moving objects is an important step in visual-based surveillance systems and represents a challenging task for researchers. To track the physical appearance of moving objects (such as the vehicles) and identify them in a dynamic scene, it has to locate the position, estimate the motion of these blobs and follow these movements between two of consecutive frames in video scene [9]. In recent years, research in the area of vehicles tracking both during daytime and nighttime has grown rapidly. There is an urgent necessity for such systems in many applications such as driver assistance systems, surveillance systems and intelligent vehicles. The current research indicated that much has been achieved regarding vehicle detection during daytime [10, 11]. However, vehicle detection during nighttime is a more challenging issue. It has not been researched to the same extent as in daytime. Moreover, there are very less number of researchers who worked on developing systems adaptable for both daytime and nighttime scenarios at the same time.

Most recent studies on vehicle detection during daytime adopt frame differencing, and background subtraction techniques to extract the features of moving vehicles from traffic scenes. However, their effectiveness depends on video image processing algorithms that are capable of reducing common problems such as shadows, occlusion, illumination, reflection, and camera shaking. Among of them, shadows have proven to be a large source of error in vehicles detection and tracking [12, 13]. In the traffic video scene, the existence of shadows might generate negative effect on pattern analysis.

Sun and Li [14] propose a moving cast shadow detection approach using combined color models. Firstly, the ratio of hue over intensity in HSI color model is used to detect the bright object pixels in foreground regions. Secondly, the theory of photometric color invariants in c1c2c3 color model is employed to distinguish the dark (similar to shadow) and colorful object pixels from shadow pixels. Finally, post processing is used to correct failed shadow and object detection in order to improve the accuracy of shadow detection,

In [15], Choi et al. present an adaptive shadow estimator to detect and eliminate the shadow of a moving object. Their method discriminates between the shadow and the moving object by cascading three estimators, which use the properties of chromaticity, brightness, and local intensity ratio. The method compensates for accumulated errors in the cascading process in the spatial adjustment step.

Moreover, Jia et al. [16] present an approach, which adequately considers color space information to detect moving cast shadows of vehicles in traffic videos. Firstly, RGB component ratios between frame and background as well as blue and red colors ratio (B/R ratio) are taken into account to detect shadows respectively. Then, they combine the two results for a refined shadow candidate. Finally, to improve the accuracy of shadow detection, post processing is adopted to correct the false detected pixels.

Although frame differencing and background subtraction techniques are effective for vehicle detection in daytime, they become inefficient in nighttime illumination conditions. This is because the background scenes are greatly affected by the varying lighting effect of moving

ISSN 2229-5518

vehicles [10]. Under nighttime, vehicle lights have been widely used as salient features for vehicle detection and tracking [11, 17]. Most of these methods use morphological operations to extract candidate headlight objects and then perform shape analysis, template matching, or pattern classification to find the paired headlights of moving vehicles. Nevertheless, there are many problems due to complex real-time conditions. Therefore, vehicle nighttime detection and tracking is still an open area with many potential for improvement.

Salvi [18] presents a traffic surveillance system for detecting and tracking moving vehicles in various nighttime environments. The algorithm is composed of four steps: headlight segmentation and detection, headlight pairing, vehicle tracking, and vehicle counting and detection. First, a fast segmentation process based on an adaptive threshold is applied to extract bright objects of interest. The extracted bright objects are then processed by a spatial clustering and tracking procedure that locates and analyzes the spatial and temporal features of vehicle light patterns, and identifies and classifies moving cars and motorbikes in traffic scenes. However, the classification function of the algorithm needs to be improved to enhance the classification capability on different vehicle types, such as buses, trucks, and light and heavy motorbikes

In [19], Wang et al. propose a region tracking-based vehicle detection algorithm during nighttime via image processing techniques. Their algorithm is based on detecting vehicle taillights and use it as the typical feature. The algorithm uses a global detection algorithm to detect and pair the taillights. When the vehicle is detected, a time series analysis model is introduced to predict vehicle positions and the possible region (PR) of the vehicle in the next frame. Then, the vehicle is only detected in the PR.

Zhang et al. [20] propose a nighttime traffic surveillance system, which consists of headlight detection, headlight tracking and pairing, camera calibration and vehicle speed estimation. First, a vehicle headlight is detected using a reflection intensity map and a reflection-suppressed map based on the analysis of the light attenuation model. Second, the headlight is tracked and paired by utilizing a bidirectional reasoning algorithm. Finally, the trajectories of the vehicle's headlight are employed to calibrate the surveillance camera and estimate the vehicle's speed. The disadvantage of this system is that when one headlight of the vehicle is occluded by other vehicles, it cannot be paired with other headlights.

While the prior art focuses mainly on detecting and tracking moving vehicles either during daytime or at night, there has been very small research on developing integrated systems that allow vehicles to be detected day and night. Although, there is a highly increasing demand for such applications, this type of systems is quite disregarded in the literature.

In [21], Cucchiara et al. describe two different sets of image analysis algorithms that have been used in a vehicular traffic tracking system for extracting vehicles from image sequences acquired in daytime and at night. In their system, a supervising level selects the set of algorithms to apply and performs vehicle tracking under control of a rule-based decision module. The image-processing modules extract visual data from the scene by spatio-temporal analysis during daytime and by morphological analysis of headlights at night.

Also, Chan et al. [22] propose an automatic system to detect preceding vehicles on the highway under various lighting and different weather conditions. Their system employs four cues (including underneath, vertical edge, symmetry and taillight) in order to adapt to different characteristics of vehicle appearance at daytime and nighttime. In addition, they adopt particle filter to integrate the four cues in order to accurately generate the vehicle distribution. This is done though four steps including initial sampling, propagation, observation and cue fusion, and evaluation.

Furthermore, Robert [23] presents a framework to detect vehicles, based on a hierarchy of features detection and fusion. The first layer of the hierarchy extract image features. The next layer fuses image features to detect vehicle features such as headlights or windshields. A last layer fuses the vehicle features to detect a vehicle with more confidence. His approach is road illumination agnostic and allows vehicles to be detected day and night. The vehicle features are tracked over frames. He uses a constant acceleration tracking model augmented with traffic domain rules to handle occlusion problems.

3 Proposed System

Previous studies indicate that illumination changes cause challenging problems for video surveillance algorithms. Hence, it is desired for such algorithms to maintain a consistent perception of a scene regardless of illumination variation. With this goal, the proposed system is developed to track vehicles during both daytime and nighttime. Figure 2 illustrates the system framework. A daytime/nighttime detector is first employed to distinguish daytime and nighttime scenes. If the scene is recognized as a daytime scene, then daytime procedures (daytime module) are applied. Otherwise, nighttime procedures (nighttime module) are invoked. This adaptive fusion can greatly improve the robustness and the flexibility of the system compared with the previous mostly non-adaptive schemes. The main motivation for differentiating the vehicle extraction procedure during daytime from the procedure during nighttime is that salient vehicle features are different in the two cases. Under daytime conditions, the proposed system employs the estimated background model of the video sequence and applies a Gamma decoding followed by a thresholding for removing cast shadows from vehicles in outdoor environments. However, under nighttime conditions, the proposed system extracts bright objects using connected components analysis and recognizes the paired lights of the vehicles using rule based component analysis approach..

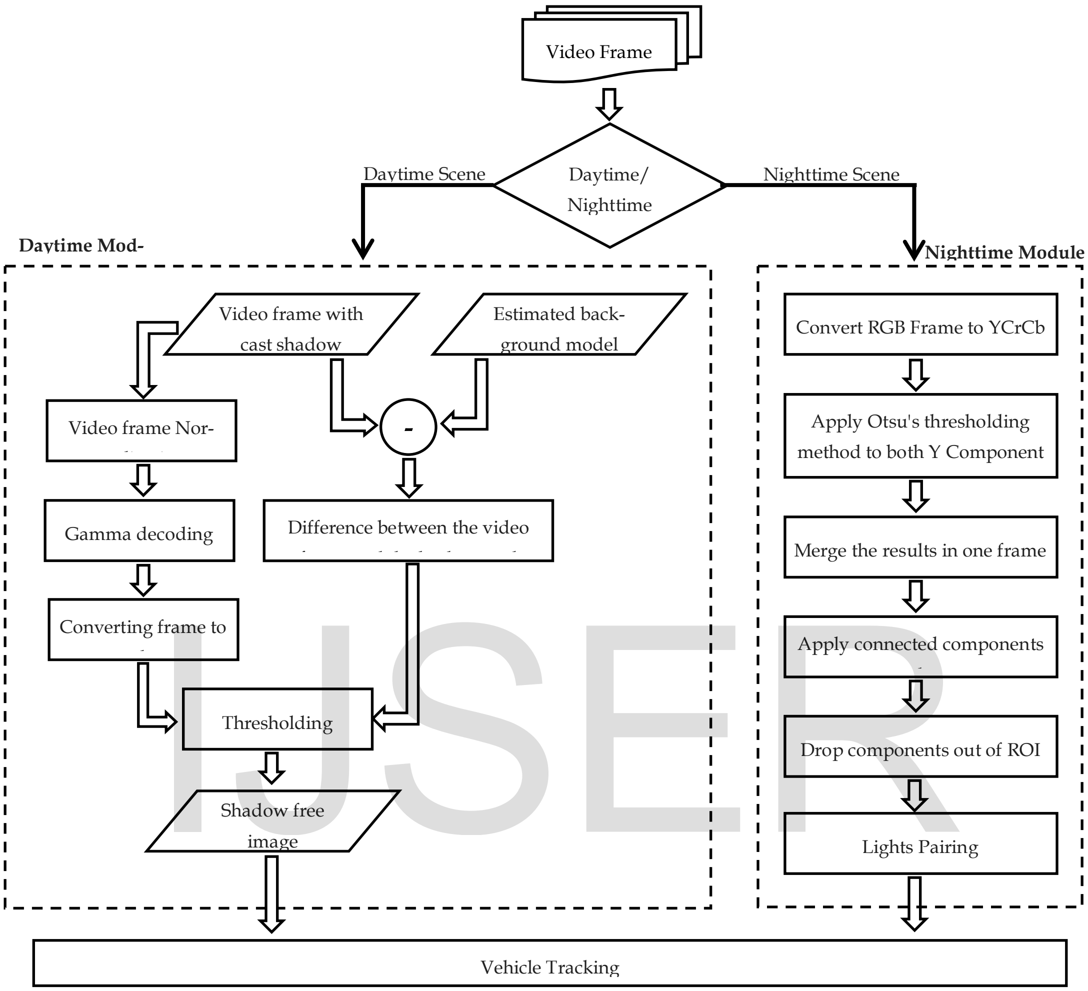

Fig. 2. The block diagram of the proposed system

3.1 Daytime/Nighttime Detector

When detecting vehicles using computer vision techniques, the external environment around the moving vehicle varies constantly and significantly. Nighttime images of a scene from a surveillance camera have lower contrast and higher noise than their corresponding daytime images of the same scene due to low illumination. Hence, it is difficult to use only one particular feature or method to recognize a vehicle image under these different environments. Therefore, it is necessary to recognize a current environment and to provide one of multiple different algorithms that is suitable for the current environment.

Conventionally, it is required to determine whether the current image is a daytime image or a nighttime image according to the lightness of the image to choose a vehicle detection method suitable for the current environment. In the proposed system, a daytime/nighttime detector presented in [24] is built for this purpose. The basic idea of the detector is that compared to daytime images, there is a distinct lack of color information in nighttime images, specifically in the blue and green regions of the spectrum [24]. This is due to the ambient illumination in nighttime images is largely from artificial sources, while the ambient illumination in daytime images is largely from natural light. In other words, nighttime images tend to be largely consist of pixels with hues extensively in the red and yellow regions of the spectrum, while ordinarily exposed daytime images tend to have a significant number of pixels with hues extensively in the green and blue regions of the spectrum.

Moreover, daytime images tend to have a relatively great number of pixels, which are sufficiently bright or light, while conversely nighttime images tend to have a relatively great number of pixels, which are relatively dark or not sufficiently bright or light. Accordingly, the detector takes advantage of a combination of these parameters (i.e., pixel hue and brightness or lightness) to distinguish daytime and nighttime images.
The detector first converts the image frame from RGB color space to HSV (Hue-Saturation-Value) color space. Then, a histogram is calculated for the H and V components, by statistically binning the values of the respective components for the pixels of the image. Using H-histogram, the detector determines the first parameter (n_H) which represents the amount of pixels in the image that are sufficiently red or yellow. That is, it counts all the pixels having an H-component below a predetermined low hue threshold (H_{Low_TH}) as well as all the pixels having an H-component above a predetermined high hue threshold (H_{High_TH}). Similarly using V-histogram, the detector determines the second parameter (n_V) which represents the amount of pixels in the image that are sufficiently light or bright. That is, it counts all the pixels having a V-component above a predetermined value threshold (V_{TH}). As stated in [24], when the hue or H-component is mapped on a scale from zero to l, the optimum threshold values obtained from experiments are approximately 0.2 and 0.8 for (H_{Low_TH}) and (H_{High_TH}) respectively. In other words, the thresholds essentially specify the lower 20% and upper 20% of the hue scale, i.e., the regions of the spectrum that are sufficiently red or yellow. Consequently, if (n_H) is less than a determined hue count threshold (Th_H) or if (n_V) is substantially greater than a determined value count threshold (Th_V), then the image is classified as a daytime image; otherwise if (n_H) is above (Th_H) and (n_V) is below (Th_V), then the image is suitably classified as a nighttime image.

3.2 Daytime Module

The proposed method incorporates estimated background model information and gamma decoding to detect cast shadow. This integration has greatly increased the scope of applicability and brought significant enhancements in the shadow-free images and the time of processing. The proposed method shows a significant elimination of the shadows in the frames of image sequences. First, the difference between the video frame and the background image is computed in a pixel-by-pixel manner and the absolute value is taken. Large values in the resultant difference image indicate the existence of foreground objects. However, small values are usually noise due to environmental factors such as illumination and background clutter. Hence, these values can be further ignored. The resultant image is then converted to black and white image to facilitate its manipulation. This difference image will be used later in the thresholding process. In the next step, the video frame is normalized by dividing each pixel value by its corresponding value of the background image. This is a pre-processing step to apply Gamma decoding.
To focus the shadow areas, Gamma decoding is applied on the normalized frame. This nonlinear operation is used to improve the fidelity of the brightness value magnitudes. Using the following equation, Gamma values (γ) larger than one make the image darker, while values smaller than one make dark regions lighter.

$$g(x, y) = A \times f(x, y)^{\gamma}$$

Where f(x,y) denotes the intensity of the pixel (x, y) in the normalized frame, g(x,y) indicates the corresponding result after Gamma decoding. In the proposed algorithm, a value larger than one is used to make shadow more apparent. By this process, the shadow pixels are getting darker and hence it can be easily separated. It compresses the low value pixels and stretches high value pixels. The input and output values of the Gamma equation are non-negative real values. The value of the constant A in the same equation controls pixel value range. In our proposed algorithm, the typical value one is used for A. In this case, inputs and outputs are typically in the range from 0 to 1. If a value greater than 1 is used for A, the range will be expanded to be from zero to A. It should be mentioned that in order to stress the shadow areas, using Gamma decoding is more effective than adjusting image brightness. Applying Gamma decoding does not change the level of detail in the image. It adjusts the RGB value of each pixel in an image but not by the same amount. However, using brightness in the same case can make details wash out or fade to white or black. Since brightness just adds or subtracts the same value to/from each pixel, the image may lose information at the extremes.
After Gamma decoding, the resultant video frame is converted from RGB to grayscale image. The resultant grayscale image has a range from zero to 255. To make the thresholding process more robust, the threshold value should be automatically selected with each frame. The manual threshold setting method and offline learning based method cannot adapt to the variation of the environment in real-time. So in the proposed algorithm, a dynamic threshold is calculated using Otsu's method [25, 26]. It is designed to select the optimum threshold for separation into two classes based upon maximizing the variance between them. It involves iterating through all the possible threshold values and calculating a measure of spread (intra-class variance) for the pixel levels each side of the threshold, i.e. the pixels that fall either in foreground or in background. The aim of this step in the algorithm is to find the threshold value where the sum of foreground and background spreads is at its minimum. It does not depend on modelling the probability density functions; however, it assumes a bimodal (i.e., two classes) distribution of gray-level values.
Therefore, thresholding process is performed over the

obtained grayscale image. The new image frame (shadow free image) will have all pixels with values greater than the dynamic threshold. This process ensures the removing of cast shadows because the value of shadow pixels will be less than the threshold value. However, it keeps the pixels of both foreground image and background image. In object tracking, foreground objects are more important. Hence, the difference image obtained in step one is used as a filter during the thresholding process. That is, the thresholding is performed only on the pixels whose corresponding pixels on the difference image are white. By this way, thresholding is restricted to the foreground pixels only instead of the whole image pixels.

3.3 Nighttime Module

Under the nighttime traffic environment, the camera images have very low contrast and a weak light sensitivity. In these conditions, the vehicles can only be identified by locating their headlights and rear lights. These are the only visual features of the vehicles at night and under darkly illuminated conditions. In order to detect the vehicle lights, it is common for image processing techniques to use some form of thresholding. However, the RGB color space is not ideal for the task of color thresholding. It is difficult to set and manipulate color parameters due to high correlation between the Red, Green, and Blue channels. Hence, the proposed method starts with converting the color space of the video frame from RGB to YCbCr. Y is the luminance component while Cb and Cr are the Blue-difference and Red-difference Chroma components. The main advantage of converting the frame to YCbCr color space is that this color space is characterized by its ability to separate the light intensity component from the chrominance. The Y component gives all information about the brightness, while the Cb (Blue) and Cr (Red) components are independent from the luminosity influence. However, in the RGB color space, each component (Red, Green and Blue) has a different brightness.

In nighttime traffic, vehicle lights appear as the brightest pixels in the video frames. Headlight objects are bright and therefore appear white in color while the core part of rear light object is red. In order to detect these pixels, the proposed method uses Otsu's thresholding technique [25, 26] to both Y component and Cr component. Vehicle headlights can be detected by applying the Otsu's technique to the luminance component Y while the rear lights (red light sources) can be detected by applying the Otsu's technique to the Red-difference Chroma channel Cr. The thresholding is performed using the following equation:

$$g(x,y) = \begin{cases} 1 & f(x,y) > T \\ 0 & otherwise \end{cases}$$

Where f(x,y) denotes the intensity of the pixel (x, y) in the video frame, g(x,y) indicates the corresponding segmentation result after thresholding and T is the threshold value. To make the thresholding process more robust, the threshold value T should be automatically selected with each frame. The manual threshold setting method and offline learning based method cannot adapt to the variation of the environment in real-time. So in the proposed algorithm, a dynamic threshold is calculated using Otsu's method. In order to extract all the bright pixels (headlights and rear lights) found in the frame, the thresholding results of both the luminance(Y) and the red component (Cr) are then combined to form a single frame.

It should be noted that some interferential objects, such as street lamps and traffic lights, are also detected in the frame especially in urban roads. Utilizing the thresholding method extracts all the bright pixels. Hence, further filtering is required as there are many potential light sources that are not vehicle lamps. To filter out these objects, the proposed method applies two consecutive steps: connected component analysis [27, 28] and Region Of Interest (ROI) filtering. First, a connected component extraction process is performed to locate the connected components of the bright objects. Extracting these components clarifies the significant features of location, dimension, shape, and pixel distribution associated with each connected component. Second, the ROI filtering is applied to each video frame to exclude any connected component with a location out of the detection region. The detection region should cover the lanes that are being monitored. It is usually set at the lower part of the image. It can be predetermined either manually during the setup of the surveillance camera or automatically by using lane detection algorithms [29, 30]. Then, the detection is only performed in the ROI. Hence, after masking outside the ROI, the scene becomes simpler, since out-of-ROI distracting objects, such as street lamps, are removed. The ROI not only can reduce complexity in searching for vehicle candidates but also can decrease the false positive detection rate. At the same time, ROI definition speeds up the processing time as only a part of original image is processed.

The next step of the proposed method is to pair the identified vehicle lights in order to start tracking. The proposed method adopts the rule based component analysis approach [19, 32] where the identified vehicle lights can be paired with each other if certain rules are satisfied. The pairs of vehicle lights must have some common properties. Hence, two connected components are said to belong to the same vehicle if the following rules are satisfied [19, 32]:

- The components must be horizontally close to each other and the vertical and horizontal positions should be considered.
- The components are of similar size.
- The width to height ratio of the bounding box enclosing the two components must be greater.

ISSN 2229-5518

- Area of the pixels must be similar.
- The symmetry condition must be satisfied.

Finally, the proposed method uses Kalman Filter (KF) to perform the tracking process. Vehicles are tracked using the four parameters of a bounding box surrounding the lamp pair (x-position, y-position, width, and height). Kalman filter is composed of two steps [32]: prediction and correction. In the prediction step, the location of an object being tracked is predicted in the next frame while in the correction step, the object in the next frame within designated region is identified. A set of KFs is used to keep track of a variable and unknown number of moving targets [33]. Each time a new observation is received, it is associated to the correct track among the set of the existing tracks or if it represents a new target, a new track has to be created. Several advantages are gained for using Kalman filter [32, 33]. First, prediction using the basic Kalman filter is extremely fast and requires little memory. This makes it a convenient form for online real time processing. Second, it is easy to formulate and implement given a basic understanding. Third, an error estimate is associated with each prediction. Fourth, these predictions can be computed recursively, bounding the time and memory needed for computation.

4 Experimental Results

In order to analyze the performance of the proposed system, several experiments were conducted to evaluate vehicles tracking during daytime and nighttime. The experiments were implemented on a 2.27GHz Intel Core i5 PC with 4GB memory, running under Windows 8 Enterprise. The algorithms were coded using MATLAB 8.1.0.604 (R2013a).

4.1 Daytime Module Evaluation

To evaluate quantitatively the performance of the shadow removal algorithm employed in the daytime module, the shadow detection rate η (Eta) and shadow discrimination rate ζ (Zeta) are calculated. These two metrics are used widely as standard metrics for testing the performance of shadow detection algorithms [7, 12-16]. They are defined as follows:

$$\text{Shadow detection rate } (\eta) = \frac{TP_s}{(TP_s + FN_s)}$$

$$\text{Shadow discrimination rate } (\xi) = \frac{TP_f}{(TP_f + FN_f)}$$

where *TP* and *FN* stand for true positive and false negative pixels with respect to either shadows (*s*) or foreground objects (*f*). This means TP_s is the number of pixels which are determined correctly as shadow pixels; TP_f is the number of pixels which are determined correctly as foreground object pixels. FN_s is the number of errors in which a shadow pixel is defined as an object pixel, and FN_f is the number of false detection which identified an object pixel as a shadow pixel. The shadow detection rate is concerned with labelling the maximum number of cast shadow pixels as shadows. The shadow discrimination rate is concerned with maintaining the pixels that belong to the moving object as foreground. In general, η is reduced with increasing ζ, and ζ is reduced with increasing η; thus, η and ζ are a reciprocal relationship.

The proposed system has been tested on two benchmark video sequences "highway I" and "highway III". This dataset is provided by Computer Vision and Robotics Research Laboratory of UCSD (http://cvrr.ucsd.edu/aton/shadow). The ground truth masks are obtained from http://arma.sourceforge.net/shadows/. Moreover, test videos presented in [16] are used to evaluate the performance of the proposed algorithm. It includes three video sequences: video 1, video 2, and video 3. Figure 3 shows the results of applying the proposed algorithm on these different video sequences. The first row shows the original frames in videos, the second is shadow detection results, and the third is ground truths. From the figure, it is obvious that the algorithm is feasible in detecting the shadow regions. Table 1 lists comparative results of the proposed algorithm with some state-of-the-art methods [14-16]. It achieves an average shadow detection rate (η) equal to 93% and an average shadow discrimination rate equal to 88%. In addition, the average of the two rates is often used as a single performance measure [12-16]. The combined score is about 91%, which seems so promising results comparing to the existing methods. The results prove that the algorithm habits excellent detection performance and is superior to other existing algorithms.

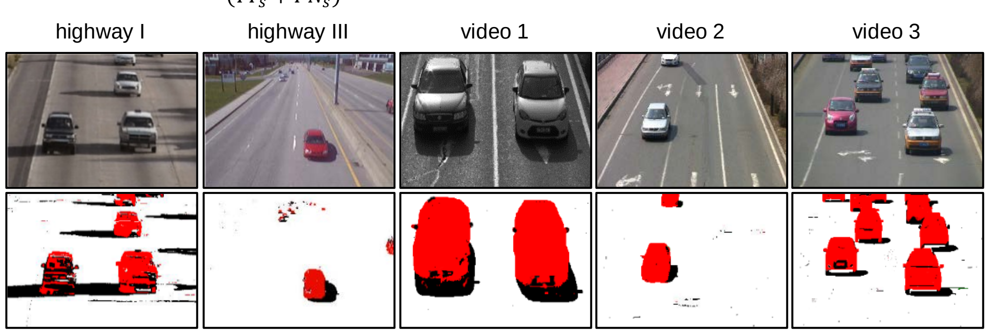

Fig. 3. The results of running the proposed system on different video sequences during daytime: original frame, shadow detection result, ground truth image from top to bottom, respectively

TABLE 1. COMPARISON OF THE PROPOSED SYSTEMS WITH EXISTING METHODS WORKING IN DAYTIME CONDITIONS

Algorithm	highway I			highway III			Video 1			Video 2			Video 3		
	η (%)	ζ (%)	Mean (%)	η (%)	ζ (%)	mean (%)	η (%)	ζ (%)	mean (%)	η (%)	ζ (%)	mean (%)	η (%)	ζ (%)	mean (%)
Sun and Li [14] (2010)	89	47	68	72	63	68	84	65	75	85	49	67	83	48	66
Choi et al. [15] (2010)	86	89	88	84	91	88	72	72	72	88	92	90	90	77	84
Jia et al. [16] (2013)	87	85	86	85	81	83	89	93	91	90	98	94	94	82	88
Proposed	92	88	**90**	93	87	**90**	94	89	**92**	97	91	**94**	91	86	**89**

Furthermore, to study the effect of using the proposed shadow removal algorithm on the tracking process, Figure 4 compares the tracking results without and with applying our shadow removal algorithm, respectively. Note the colored rectangles indicate the detected objects. In each frame shown in the figure, two vehicles appear as one object in the tracking process due to shadow stretching. In addition, in frames number 6 and 245, some moving shadows (identified by orange and blue rectangles respectively) have been wrongly detected as foreground object. On the other hand, we could see that the proposed algorithm has successfully detected the real objects and neutralized the negative influences by shadows.

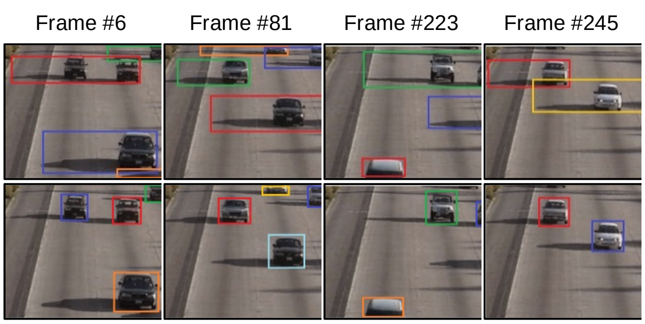

Fig. 4. The tracking results on "highway I" video sequence first row: without shadow removal, second row: with the proposed algorithm

4.2 Nighttime Module Evaluation

Several experiments were conducted to evaluate the nighttime module. Establishing standard test beds is a fundamental requirement to compare algorithms. Unfortunately, there is no standard dataset to compare the results and efficiency of nighttime vehicles detection and tracking algorithms. This was a major difficulty in order to compare the performance of the proposed method with its counterparts. Most of algorithms found in the literature record their videos by their own. To evaluate the performance of the proposed method, two sets of experiments are conducted. The first set of experiments are performed over a self-collected and prepared dataset. It consists of 14 video clips containing both urban and highway scenes (downloaded from http://www.videoblocks.com/). All the video sequences are with a frame rate equal to 30 frames per second and the size of the grabbed image sequence is 480 × 270 pixels with 24-bit true color. The video clips are selected with different traffic density (high – medium- low). The ground truth for each video clip of the dataset was prepared manually. A detailed description of the dataset is found in [8]. The second set of experiments are performed over the testing video data used in [19]. It consists of four video sequences captured in an urban roadway environment. Two videos of them (video a and video b) contain only one moving vehicle in the scene while the others (video c and video d) contain from two to four moving vehicles in the scene. The frame rate of each video is 30 frames per second and the size of the grabbed image sequence is 720 × 480 pixels with 32-bit true color.

Figure 5 shows the results of applying the proposed tracking method to both highway road and urban road traffic scenes of our dataset. The first row shows the original frame of a nighttime surveillance video. The second row displays the results after applying the first five steps of the proposed method: applying Otsu's thresholding to both Y Component & Cr Component, merging the results in one frame, applying connected components analysis, and excluding components out of ROI. The third row shows the results of pairing detected vehicle lights where the green rectangles indicate the vehicle lights that have been paired. The fourth row shows the tracking results. The experimental results demonstrate that the proposed method can robustly detect and track vehicles in different nighttime traffic environments.

Table 2 shows quantitative results of the proposed method for vehicle tracking in different nighttime traffic environments. The average tracking rates of the proposed method are 96.27% and 95.76% for both urban and highway scenes respectively in our dataset. Almost all the vehicle lights can be detected and the false tracking of vehicles occurs when the vehicles move side by side or when there exist some moving reflection objects on the road. This in turn may cause some false pairing. However, the effect of interferential objects such as street lamps are attenuated by the step of excluding all detected components outside the ROI. .

Above LA Highway Traffic | Highway LA Overpass | LA Highway Bend | Night Time Traffic on Snowy Downtown Street in Homer | Slow Moving Los Angeles Traffic | Traffic on Busy Times Square Street

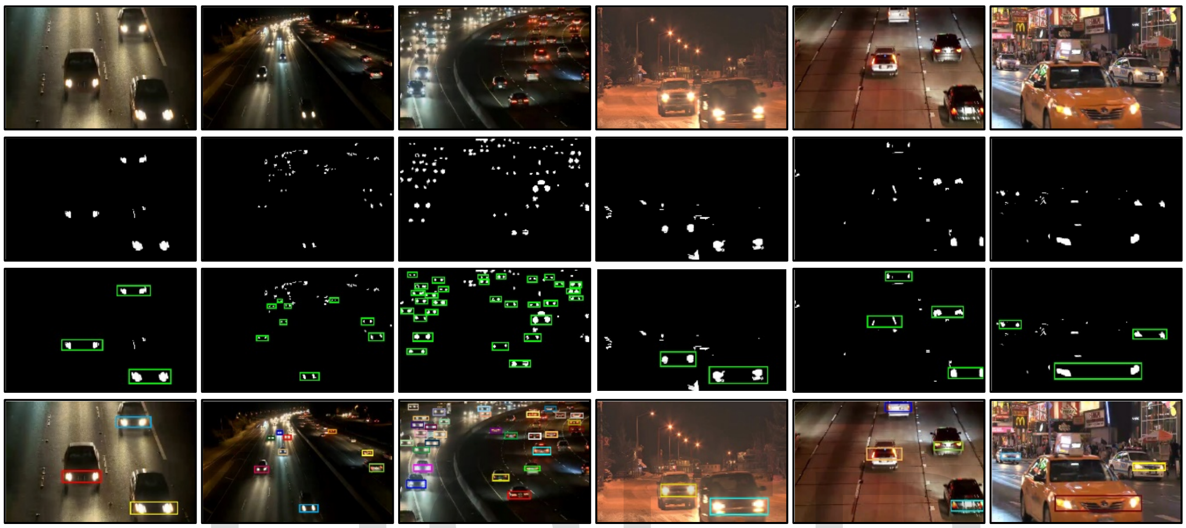

Fig. 5. Moving vehicles tracking in highway and urban road traffic scenes of the dataset

TABLE 2: TRACKING RATE OF THE PROPOSED METHOD FOR OUR DATASET

No	Video sequence name	No of vehicles	No of correctly tracked vehicles	Tracking rate (%)
1	Above LA Highway Traffic	23	20	86.96%
2	Highway LA Overpass	53	52	98.11%
3	LA Highway Bend Traffic	96	94	97.92%
4	LA Highway Bend	64	63	98.44%
5	Slow Moving 101 North Traffic	41	39	95.12%
6	Slow Night Commute In Cali	74	74	100.00%
7	Cars On LA Highway	112	109	97.32%
8	Night Time Traffic on Snowy Downtown Street in Homer	4	4	100.00%
9	Slow Moving Los Angeles Traffic	35	32	91.43%
10	Nighttime Traffic in Aspen	7	7	100.00%
11	Roadway Traffic at Night in Snowy Small Town	3	3	100.00%
12	Seattle Airport Control Tower and Traffic at Night	5	5	100.00%
13	Taxi Cabs and Traffic in Times Square	9	8	88.89%
14	Traffic on Busy Times Square Street	10	9	90.00%

The following part evaluates the performance of the proposed method and compares it to the region tracking-based vehicle detection algorithm presented by Wang et al. [19]. Figure 6 shows the comparative results of nighttime vehicle tracking for running the two methods on the test sequences used in [19]. The first column of the figure shows the original frame. The second column shows the results of applying the region tracking-based vehicle detection algorithm. The third column shows the results of applying the proposed method. As it can be noted from the figure, the proposed method successfully detects and tracks all vehicles appeared in the scene. However, the region tracking-based algorithm does not perform well in detecting all vehicles under some complicated nighttime traffic scenes, and some vehicles are missed. This is because the proposed method applies the adaptive thresholding step to both Y Component and Cr Component of the video frame and merge the two results in one frame. However, the region tracking-based vehicle detection algorithm applies the thresholding on the gray scale image of the video frame. Gray scale image based segmentation succeeds in segmenting white pixels but fails in segmenting red pixels especially in low illumination conditions. Pixels with high red color component and low green and blue color components are bright red in color but their corresponding gray scale values can be low. Hence, these pixels face a difficulty to be detected.

Original frame | Region Tracking-Based Vehicle Detection Algorithm [9] | Proposed Method

Fig. 6. Comparative results of vehicle detection and tracking in nighttime traffic scenes on test sequences used in [19]

Table 3 shows the tracking rates achieved when running both the Region Tracking-Based Vehicle Detection Algorithm [19] and the proposed method on the testing sequences used in [19]. As the table indicates, both algorithms succeed in tracking all the vehicles appeared in the video a and video b because both videos contain only one vehicle to be tracked. However, when the number of vehicles increases, the performance of the Region Tracking-Based Vehicle Detection Algorithm is degraded while the proposed method still successfully detects and tracks almost all vehicles.

TABLE 3: TRACKING RATE OF THE PROPOSED METHOD FOR TEST SEQUENCES USED IN [19].

No	*Video sequence name*	*Tracking rate (%)*	
		Region Tracking-Based Vehicle Detection Algorithm [19]	***Proposed method***
1	Video a	100%	100%
2	Video b	100%	100%
3	Video c	98.67%	99.1%
4	Video d	93.45%	97.23%

5 CONCLUSION

This paper presents a comprehensive system to tracking vehicles under various environmental conditions. The proposed system can be used in both daytime and nighttime conditions. A specific detector has been employed to assess the lighting conditions of the images captured by the camera. It detects whether the scene is daytime or nighttime conditions. According to the output of the detector, the suitable tracking module is triggered. Under daytime conditions, the proposed system employs the estimated background model of the video sequence and applies a Gamma decoding followed by a thresholding for removing cast shadows from vehicles in outdoor environments. Additionally, under nighttime conditions, the proposed system extracts bright objects using connected components analysis and recognizes the paired lights of the vehicles using rule based component analysis approach. Experimental results demonstrate that the proposed method is feasible and effective for vehicle detection and identification in various daytime and nighttime environments.

ACKNOWLEDGMENTS

We would like to express our sincere appreciation to Dr. Jianqiang Wang at Tsinghua University and Dr. Xiaoyan Sun at Suzhou INVO Automotive Electronics Co., for their help in providing their dataset to be used in testing the performance of the proposed method.

REFERENCES

[1] D. Neelima, Gowtham Mamidisetti, "A Computer Vision Model for Vehicle Detection in Traffic Surveillance," International Journal of Engineering Science & Advanced Technology (IJESAT), Volume 2, Issue 5, pp. 1203 – 1209, Sep-Oct 2012.

[2] Yen-Lin Chen, Bing-Fei Wu, Hao-Yu Huang, and Chung-Jui Fan, "A Real-Time Vision System for Nighttime Vehicle Detection and Traffic Surveillance," *In Proceedings of IEEE Transactions on Industrial Electronics*, Volume 58, Issue 5, pp.2030-2044, May 2011.

[3] Hsu-Yung Cheng, Po-Yi Liu, Yen-Ju Lai, "Vehicle Tracking in Daytime and Nighttime Traffic Surveillance Videos," In Proceedings of the 2nd International Conference on Education Technology and Computer (ICETC), Shanghai, China, Volume 5, pp. 122-125, June 2010.

[4] Tomás Rodríguez, Narciso García, "An Adaptive, Real-Time, Traffic Monitoring System," In the International Journal of Machine Vision and Applications, Springer-Verlag, Volume 21, Issue 4, pp. 555–576, June 2010.

[5] Kostia Robert, "Video-based Traffic Monitoring at Day and Night Time," In Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems (ITSC '09), St Louis, USA, October 2009.

[6] Amirali Jazayeri, Hongyuan Cai, Jiang Yu Zheng, and Mihran Tuceryan, "Vehicle Detection and Tracking in Car Video Based on Motion Model," In IEEE Transactions on Intelligent Transportation Systems, Vol. 12, No. 2, pp. 583-595, June 2011.

[7] Mohamed Taha, Hala H. Zayed, M. E. Khalifa, and Taymoor Nazmy, "Moving Shadow Removal for Multi-Objects Tracking in Outdoor Environments," In The International Journal of Computer Applications (IJCA), Foundation of Computer Science, New York, USA, Volume 97, Number 10, pp. 43-51, July 2014.

[8] Mohamed Taha, Hala H. Zayed, M. E. Khalifa, and Taymoor Nazmy, "An Efficient Method for Multi Moving Objects Tracking at Nighttime," In The International Journal of Computer Science Issues (IJCSI), Volume 11, Issue 6, November 2014.

[9] Raad Ahmed Hadi, Ghazali Sulong, and Loay Edwar George, "Vehicle Detection and Tracking Techniques: a Concise Review," In the International Journal of Signal & Image Processing (SIPIJ), Volume 5, Number 1, pp. 1-12, February 2014.

[10] Sayanan Sivaraman and Mohan M. Trivedi, "A Review of Recent Developments in Vision-Based Vehicle Detection," IEEE Intelligent Vehicles Symposium, June 2013.

[11] Kovacic, Kristian, Ivanjko, Edouard and Gold, Hrvoje. "Computer Vision Systems in Road Vehicles: A Review," *In the Proceedings of the Croatian Computer Vision Workshop*, October 2013.

[12] Yang Wang, "Real-Time Moving Vehicle Detection with Cast Shadow Removal in Video Based on Conditional Random Field," In IEEE Transactions on Circuits and Systems for Video Technology, Volume 19, Issue 3, pp.437-441, March 2009.

[13] Ye-Peng Guan, "Texture and Space-Time Based Moving Objects Segmentation and Shadow Removing," In Proceedings of the Third

ISSN 2229-5518

International Conference on Advances in Swarm Intelligence (IC-SI'12), Shenzhen, China, Volume Part II, pp. 244-253, June 2012.

[14] Bangyu Sun, Shutao Li, "Moving Cast Shadow Detection of Vehicle Using Combined Color Models," In the Proceedings of the Chinese Conference on Pattern Recognition (CCPR), China, pp. 1–5, October 2010.

[15] JinMin Choi, Yung Jun Yoo, Jin Young Choi, "Adaptive Shadow Estimator for Removing Shadow of Moving Object," In the Journal of Computer Vision and Image Understanding, Volume 114, Issue 9, pp.1017-1029, September 2010.

[16] Yuhong Jia, Xiaoxi Yu, Jiangyan Dai, Wanjun Hu, Jun Kong, and Miao Qi, "A Novel Moving Cast Shadow Detection of Vehicles in Traffic Scene," Intelligent Science and Intelligent Data Engineering Lecture Notes in Computer Science, Springer, Volume 7751, pp. 115-124, 2013.

[17] Padmavathi. S, Keerthana Gunasekaran, "Night Time Vehicle Detection for Real Time Traffic Monitoring Systems: A Review" In the International Journal of Computer Technology & Applications, Volume 5, Number 2, PP. 451-456, April 2014.

[18] G. Salvi, “An Automated Nighttime Vehicle Counting and Detection System for Traffic Surveillance,” In Proceedings of the 2014 International Conference on Computational Science and Computational Intelligence (CSCI '14), Volume 01, Volume 01, March 2014.

[19] Jianqiang Wang, Xiaoyan Sun, Junbin Guo, “A Region Tracking Based Vehicle Detection Algorithm in Nighttime Traffic Scenes”, In the International Journal of Sensors, Volume 13, Issue 12, pp. 16474-16493, December 2013.

[20] Wei Zhang, Q. M. J. Wu, Guanghui Wang, and Xinge Yu, "Tracking and Pairing Vehicle Headlight in Night Scenes," In Proceedings of IEEE Transactions on Intelligent Transportation Systems, Volume 13, Issue 1, March 2012.

[21] Rita Cucchiara, Massimo Piccardi, and Paola Mello, "Image Analysis and Rule-Based Reasoning for a Traffic Monitoring System," In IEEE Transactions on Intelligent Transportation Systems, Volume 1, Number 2, June 2000.

[22] Yi-Ming Chan, Shih-Shinh Huang, Li-Chen Fu, "Vehicle Detection under Various Lighting Conditions by Incorporating Particle Filter," In Proceedings of the IEEE Intelligent Transportation Systems Conference Seattle, WA, USA, September 2007.

[23] Kostia Robert, "Video-Based Traffic Monitoring at Day and Night Vehicle Features Detection Tracking," In Proceedings of the 12th International IEEE Conference on Intelligent Transportation Systems (ITSC'09), St. Louis, USA, pp. 1-6, October 2009.

[24] Bhaskar Saha, Daniel Davies, and Ajay Raghava, "Day Night Classification of Images Using Thresholding on HSV Histogram" US Patent 2013/0335556 A1, December 19, 2013.

[25] S. Kumar, M. Pant, A. Ray, "Differential Evolution Embedded Otsu's Method for Optimized Image Thresholding," In Proceedings of the 2011 World Congress on Information and Communication Technologies (WICT), Mumbai, India, pp. 325 - 329, December 2011.

[26] Mehmet Sezgin,and Bulent Sankur, "Survey Over Image Thresholding Techniques and Quantitative Performance Evaluation," In the Journal of Electronic Imaging, Volume 13, Issue 1, pp. 146-168, January 2004.

[27] Lifeng Hea, Yuyan Chaob, Kenji Suzukic, and Kesheng Wud, "Fast Connected-Component Labeling," In the International Journal of Pattern Recognition, Volume 42, Issue 9, pp. 1977–1987, September 2009.

[28] Kenji Suzukia, Isao Horibaa, and Noboru Sugieb, "Linear-Time Connected-Component Labeling Based on Sequential Local Operations," In the International Journal of Computer Vision and Image Understanding, Volume 89, Issue 1, pp. 1–23, January 2003.

[29] Aharon Bar Hillel, Ronen Lerner, Dan Levi, Guy Raz, "Recent Progress in Road and Lane Detection: a Survey," In Machine Vision and Applications Journal, Volume 25, Issue 3, pp. 727-745, April 2014.

[30] Sibel Yenikaya, Gökhan Yenikaya, Ekrem Düven, “Keeping the Vehicle on the Road: A Survey on On-Road Lane Detection Systems,” In the International Journal of ACM Computing Surveys (CSUR), Volume 46, Issue 1, Article No. 2, October 2013.

[31] Shifu Zhou, Jianxiong Li, Zhenqian Shen and Liu Ying “A Night time Application for a Real-Time Vehicle Detection Algorithm Based on Computer Vision”, Research Journal of Applied Sciences, Engineering and Technology, Volume 5, Number 10, pp. 3037-3043, March 2013.

[32] Vahid Fathabadi, Mehdi Shahbazian, Karim Salahshour, Lotfollah Jargani, "Comparison of Adaptive Kalman Filter Methods in State Estimation of a Nonlinear System Using Asynchronous Measurements," In Proceedings of the World Congress on Engineering and Computer Science(WCECS 2009), Vol II, San Francisco, USA, October 2009.

[33] Amir Salarpour, Arezoo Salarpour, Mahmoud Fathi, and MirHossein Dezfoulian, "Vehicle Tracking Using Kalman Filter and Features," In the International Journal of Signal & Image Processing (SIPIJ), Volume 2, Number 2, June 2011.