Vehicle Detection and Tracking System for Real-Time Traffic Surveillance

Mrs. Shobha
Department of Information Science and Engineering
SDM Institute of Technology
Ujire, Karnataka, India

Abstract— Real-time vehicle detection, tracking and counting of vehicles is of great interest for researchers and is a need of the society in general for comfortable, smooth and safe movements of vehicles in cities. This proposed method for a video image processing algorithm which detects, tracks and finds the number of vehicles on a road. It converts RGB video frame to HSV color domain, which helps in differentiating the colors of the vehicles more absolutely. The noise is removed from each frame. The number of vehicles running in a video or in a particular lane is determined. This method proposed a novel idea to detect, track and count the vehicles on a using OpenCV and Python.

Keywords—Vehicledetection,Surveillance,OpenCV,Tracking

I. Introduction

Now-a-days, the rapid increase in the number of vehicles on the highway and urban roads have created many challenges regarding the proper management and control of the traffic. Detecting and tracking of vehicles using the traffic surveillance system give more promising way to manage and control the road traffic. Vehicle surveillance represents a challenging task of moving object segmentation in complex environment. The detection ratio of such algorithms depends upon the quality of the generated foreground mask. Therefore, there is a need for an efficient method for detection and tracking of vehicles which focuses on the trajectory of motion of the objects.

This system works with the integration of CCTV cameras for detecting the vehicles. The initial step will always be vehicle detection. Haar Cascades are used for detection of vehicle in the footage. Vehicle detection and counting algorithm is used to count the number of vehicles in the footage. Viola-Jones Algorithm is used in training these cascade classifiers. It is modified to find unique objects in the video, by tracking each vehicle in a selected region of interest. This is one of the fastest methods to correctly identify, track and count a vehicle object with accuracy.

The experimental results have established the superior performance of the proposed method. Automatic identification of vehicle data has been commonly used in the vehicle information system and intelligent traffic system. It has acquired more attention of researchers from the last decade with the advancement of digital imaging technology and computational capacity. Automatic vehicle detection systems are keys to road traffic control nowadays some applications of these systems are traffic response system, traffic signal controller, lane departure warning system, automatic vehicle accident detection and automatic traffic density estimation. An Automatic vehicle checking framework makes utilization of video information gained from stationary movement cameras, performing causal numerical operations over an arrangement of outlines got from the video to gauge the quantity of vehicles display in a scene. It is only the capacity of consequently remove and perceive the activity information. Stealing of vehicles is very common now days. It is very complicated to track a vehicle without adding any tracking device. We can easily track a vehicle if there's some tracking device integrated with it, but problem with tracking device is it can be easily removed or deactivated. Instead of using any device for tracking,tracking of vehicle number plate using CCTV footage can be used. For this purpose, android app can be used. The victim can register on the app with vehicles number with other information. As soon as he registers on the app, the information will be stored on the server. All the CCTV's connected to the server will look for the information in the live time using Image processing and will respond to the server as soon as it scans the mentioned information.

The server will then send a notice to the user with latest information given by CCTV. Using this technique, the problem of ambulance management in traffic can also solved. Most commonly the manual traffic surveillance is implemented where a human operator is needed to perform the task of monitoring the traffic changes by watching different streams of traffic coming from various cameras. The main disadvantage of using such system is that there is too much information for a human operator to process at a given time. Moreover, the operator needs to be vigilant to address a sudden disruption that makes it a tiresome task to perform. Semi-automatic system uses both human operator and low-level computer vision techniques for traffic surveillance. Here the extraction of useful information like classification and identification are done by the human operator. On the other hand, in fully automatic systems the entire task of object detection, recognition and extraction of useful information from given input are done by the machine using advanced computer vision techniques. Such systems are easy to install and can operate without any human intervention.

II. Literature Survey

Authors[1], introduced Real-time Vehicle Detection and Tracking. The focus of this paper is to design and develop real-time vehicle detection and tracking algorithm. It would minimize the constraints like input image resolution, acquisition height, environmental noise etc. on the input file and still able to achieve satisfactory performance. Further, to detect the vehicles using videos of both low and high resolution captured by a stationary traffic surveillance camera. Afterwards, detection of vehicles can be tracked in the given video until detected vehicle is visible and count the total number of vehicles detected and tracked. The objective here is to improve the quality of the foreground mask generated by reducing unwanted noise and further consolidating the objects in generated foreground. The proposed method uses both low and high resolution video as an input argument. It performs the task of detection and tracking iteratively on every frame until the last frame is reached. The experiments have been performed in MATLAB version 7R2009b on an Intel i7 3.40GHz machine with 2 GB RAM. The proposed method was tested on real life situations of the highway traffic using the videos captured by the traffic surveillance camera. The videos which were used to test the algorithm can be broadly categorized into three types. (i) high definition video of single lane traffic, (ii) high definition video of two-lane traffic and (iii) low resolution video of traffic containing the curved subset of highway.

VPS Naidu et. al. [2] introduced Vehicle Tracking Using Image Processing. In this paper, authors address the issue of detecting vehicle from video frames. Accordingly, the vehicles can be easily tracked if there is some tracking device integrated with it, but problem with tracking device is it can be easily removed or deactivated. Instead of using any device for tracking, vehicles can be tracked using number plate from CCTV footage. For this purpose, android app has been used. The victim can register on the app with vehicles number with other information. As soon as he registers on the app, the information will be stored on the server. All the CCTV's connected to the server will look for the information in the live time using Image processing and will respond to the server as soon as it scans the mentioned information. The server will then send a notice to the user with latest information given by CCTV. The developed system first detects the vehicle and then captures the vehicle image. Vehicle number plate region is extracted using the image segmentation in an image. Optical character recognition technique is used for the character recognition. These systems can befully automated to operate 24/7 and monitor unauthorized parking and vehicle movements in environments such as Access control points, Distribution centres, Hospitals and car parking areas.

Authors [3], introduced Vehicle Detection, Tracking and Counting. This paper presents a low-cost camera-based algorithm in order to control traffic flow on a road. The algorithm is based on mainly three steps: vehicle detection, counting and tracking. Background subtraction is used to isolate vehicles from their background, Kalman filter is used to track the vehicles and Hungarian algorithm is exploited for association of labels to the tracked vehicles. This algorithm is implemented on both daytime and night time videos acquired from CCTV camera and IR camera. This paper uses Gaussian mixture model which has an advantage of detecting more minor details in foreground extraction because this method basically computes the PDF corresponding to every pixel in a frame which means that it is more flexible in terms of cluster covariance. Algorithm for vehicle detection is implemented using a MATLAB 2017a software. The results of the aforementioned algorithm are compared with the ground truth values of vehicles in frames. Algorithm yields accurate counting of vehicles in a video. Only the vehicles entering the region of interest (detection field) are detected, tracked and counted by this algorithm. The detection obtained through this framework when compared with the ground truth value gives minute error which shows the accuracy of this frame work.

Authors P. Spagnolo and others [4], introduced A novel approach in real-time vehicle detection and tracking using Raspberry Pi. This paper proposes a novel idea to detect, track and count the vehicles on a road and it has been implemented on Raspberry Pi 3 using OpenCV and C++. The proposed algorithm is built on Raspberry Pi with USB Camera to capture the traffic scene. Also, the pre-captured traffic videos can be run on the Raspberry Pi and can be analysed. The Raspberry Pi along with its camera kept at a remote place to capture the traffic videos and it can be controlled by the desktop or laptop or through the android devices and can analyse the traffic data. To access the remote Raspberry Pi from computers or any devices, a static IP address is assigned to Raspberry Pi and is connected to the private network, so that user can access information about the traffic from any remote place.

Authors[5], introduced Image Processing Based Vehicle Detection and Tracking System. The project aim proposes an accurate and effective moving vehicle detection method whichcan be used in complex traffic environment. It is broadly used in computer vision and video image. This paper detects and tracks vehicle for safety and traffic surveillance system. The conception of vehicle tracking in built upon the vehicle segmentation method. According to the authors Traffic detection method to track each vehicle, these methods show accurate and low error estimation result comparing with all the methods and the result depends on the quality of vehicle detection. It focuses in these areas, namely vehicle detection, tracking, and classification with appearance of shadow and partial occlusion.

III. Traffic: Major Problem in Cities

Insufficient traffic safety and traffic jams are major problems in cities around the world. Traffic jam is a situation when vehicles are stopped completely for some time period on roads. Also, vehicle have to wait for long period of time to move out of jam. Sometimes it becomes like congestion in traffic. This happens in transport network due to increasing vehicle and overuse of roads. Often it is due to slow speed, longer trip time and increased queues of vehicles. Traffic jam has tremendous impact on the life of people. It is one of most serious problems in big cities that people have to deal with it on daily basis they may get psychologically affected. It also negatively affects work, education and personal life of people and finally progress of the country.

Moving object detection involves locating objects in the frame of a video sequence. Every tracking method requires an object detection mechanism either in every frame or when the object first appears in the video. In moving object detection various background subtraction techniques available in the literature were simulated. Background subtraction involves the absolute difference between the current image and the reference updated background over a period of time. A good background subtraction should be able to overcome the problem of varying illumination condition, background clutter, shadows, camouflage, bootstrapping and at the same time motion segmentation of foreground object should be done at the real time. Disobeying the traffic signal rules may lead to congestion and increase in number of accidents. Drivers who fail to obey traffic signals put not only themselves but other motorists in danger of accidents and serious injuries. Violation in traffic law are very common in highly populated country like India.

IV. Concepts Used

A. *Segmentation*

The Segmentation is the process of partitioning a digital image into multiple regions or set of pixels. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyse. The result of segmentation is a set of segments that collectively cover the entire scene or a set of contours extracted from the scene. Each of the pixel in a region is similar with respect to some characteristics or computed property such as colour, intensity or texture. Adjacent regions are significantly different with respect to the same characteristic(s).

B. *Segmentation of Objects in Image Sequences*

Images are segmented into objects to achieve efficient compression by coding the contour and texture separately. As the purpose is to achieve high compression performance, the objects segmented may not be semantically meaningful to human observers. The more recent applications, such as content-based image/video retrieval and image/video composition, require that the segmented objects be semantically meaningful. Finding moving objects in image sequences is one of the most important tasks in computer vision and image processing. Background subtraction approach means to compute the stationary background image and to identify the moving objects as that pixel in the image that differ significantly from the background. Background subtraction can provide an effective means of locating a moving object.

C. Gaussian Mixture Model Algorithm

In order to give a better understanding of the algorithm used for background subtraction the following steps were adopted to achieve the desired results:

1. Firstly, we compare each input pixels to the mean 'μ' of the associated components. If the value of a pixel is close enough to a chosen component's mean, then that component is considered as the matched component. In order to be a matched component, the difference between the pixel and mean must be less than compared to the component's standard deviation scaled by factor D in the algorithm.
2. Secondly, update the Gaussian weight, mean and standard deviation (variance) to reflect the new obtained pixel value. In relation to non-matched components the weights 'w' decreases whereas the mean and standard deviation stay the same. It is dependent upon the learning component 'p' in relation to how fast they change.
3. Thirdly, here we identify which components are parts of the background model. To do this a threshold value is applied to the component weights 'w'.
4. Fourthly, in the final step we determine the foreground pixels. Here the pixels that are identified as foreground don't match with any components determined to be the background.

Background modelling by Gaussian mixtures is a pixel based process. Let x be a random process representing the value of a given pixel in time. A convenient framework to model the probability density function of x is the parametric Gaussian mixture model where the density is composed of a sum of Gaussians. Let p(x) denote the probability density function of a Gaussian mixture comprising K component densities:

$$p(x)=\Sigma w_k N(x;\mu_k,\sigma_k) \; k=1 \text{ to } K \quad (4.1)$$

where w_k are the weights, and $N(x;\mu_k,\sigma_k)$ is the normal density of mean μ_k and covariance matrix $\Sigma_k = \sigma_k I$,(I denotes the identity matrix).

First, the parameters are initialized with $w_k = w_0$, $\mu_k = \mu_0$ and $\sigma_k = \sigma_0$. If there is amatch, i.e.

$$||x - \mu j||/\sigma j < \tau \text{ for some } j \in [1..K] \quad (4.2)$$

where $\tau(> 0)$ is some threshold value, then the parameters of the mixture are updated as follows:

$$w_k(t) = (1-\alpha)w_k(t-1) + \alpha M_k(t), \quad (4.3)$$

$$\mu_k(t) = (1-\beta)\mu_k(t-1) + \beta x, \quad (4.4)$$

$$\sigma 2\, k(t) = (1-\beta)\sigma 2\, k(t-1) + \beta\, ||(x-\mu k(t))||2, \quad (4.5)$$

where Mk(t) is equal to 1 for the matching component j and 0 otherwise. If there is no match, the component with the lowest weight w_0 is re-initialized with $w_k = w_0$, $\mu_k = x$ and $\sigma_k = \sigma_0$. The learning rate α is constant and β is defined as:

$$\beta = \alpha N(x;\mu_k,\sigma_k) \quad (4.6)$$

Finally, the weights w_k are normalized at each iteration to add up to 1.

D. Temporal Differencing

Figure In temporal differencing, moving regions are detected by taking pixel-by-pixel differenceof consecutive frames (two or three) in a video sequence. Temporal differencing is the most common method for moving object detection in scenarios where the camera is moving.Here the moving object is detected by taking the difference of consecutive image frames t-1 and t.

E. Foreground Detection

In this step, it identifies the pixels in the frame. Foreground detection compares the videoframe with the background model, and identify candidate foreground pixels from the frame.Commonly- used approach for foreground detection is to check whether the pixel issignificantly different from the corresponding background estimate.

F. Background Subtraction

After camera calibration, the background subtraction is done to identify and extract any object from the video frame. Before applying the Background subtraction video is converted form RGB frame to grey scale frame. Figure 3 and Figure 4 shows the outcome of the converted frame from RGB to grey scale.

After conversion into the grey scale, the Mask analyser has been applied on video frame to differentiate foreground mask from the background mask. Initially the background mask is totally black. However, when something was detected moving on the background mask, it is subtracted from it.

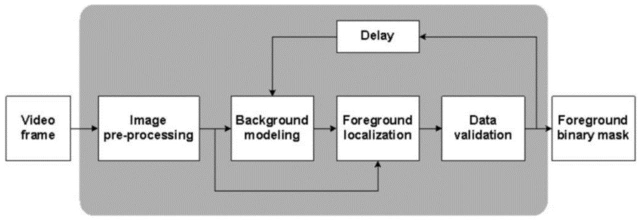


Fig 1 Background Subtraction Model.

G. Background Subtraction

After the object detection is achieved, the problem of establishing a correspondence between object masks in consecutive frames should arise. Obtaining the correct track information is crucial for subsequent actions, such as object identification and activity recognition. For this situation, Kalman filtering technique is used. The Kalman filter is a recursive two-stage filter. At each iteration, it performs a predict step and an update step.

The predict step predicts the current location of the moving object based on previous observations. For instance, if an object is moving with constant acceleration, we can predict its current location, based on its previous location, using the equations of motion.

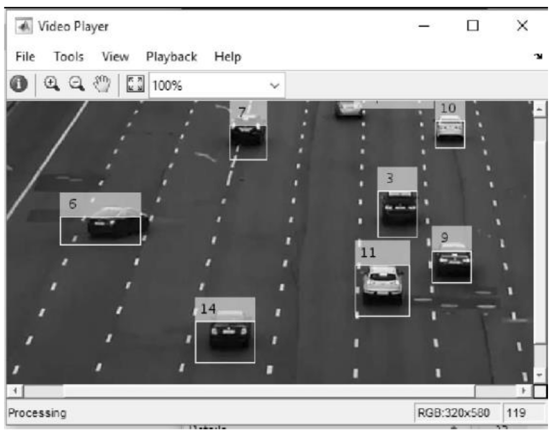


Fig 2 Detection and Tracking of objects.

H. Blob Analysis

For image processing, a blob is defined as a region of connected pixels. Blob analysis is the identification and study of these regions in an image. The algorithms discern pixels by their value and place them in one of two categories: the foreground (typically pixels with a nonzero value) or the background (pixels with a zero value).

Blob analysis is used in finding blobs whose spatial characteristics satisfy certain criteria. In many applications where computation is time consuming, blob analysis is used toeliminate blobs that are of no interest based on their spatial characteristics, and keep only the relevant blobs for further analysis. It can also be used to find statistical information such as the size of the blobs or the number, location, and the presence of blob regions.

Another typical problem of any motion detection system is its reliability in the presence of sudden changes in light conditions. This is typical of indoor environments (owing to light switches) but also of outdoor contexts, owing, for example, to a sudden cloud. In such cases, the system should be able to re-establish normal conditions as soon as possible. This kind of situation can be easily detected by monitoring continuously the variations between consecutive images of moving points.

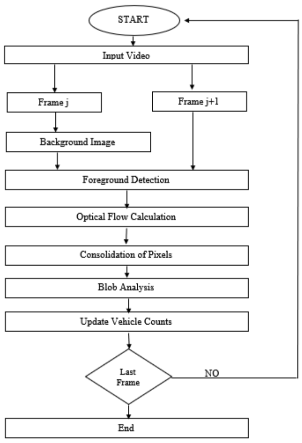


Fig 3 Flow Chart of Proposed Method.

V. Vehicle Detection and Counting Algorithm

One of the promising tasks is estimating traffic density on highways, road intersections and streets of a city in order to increase efficiency of transport system management. The road traffic analysis systems can be used both for solving problems in real time (for an adaptive traffic control) and for obtaining of traffic flow statistics. Processing and analysis of statistical information allows us to propose effective measures to increase passenger and cargo traffic and traffic safety, for example, installation of traffic lights and regulation of their work, modification of a road intersection, building additional lanes (for common use and for public transport) and others.

The developed vehicle detection and counting algorithm requires the definition of special regions of interest (the sensors) in the image. The sensor is installed on each road lane (example is presented in the section "Experimental results"). Each sensor is divided into two zones (usually the entry zone and exit zone). This allows us to determine the movement direction of a passing vehicle, as well as perform a rough estimation of it speed if the distance between the zones has been specified.

The initialization stage of the algorithm consists in the background modeling in each zone. At this stage the algorithm determines frames without motion. For this purpose the number of moving points on each frame is estimated:

- Calculating the difference between the current frame and previous frame.
- Obtaining the number of moving points by thresholding.
- If the number of moving points is greater than the threshold, depending on the resolution of the image, then it is considered that motion is detected in the zone. Otherwise, there is no motion.

The zones are synchronized to determine the moment when a vehicle left the entry zone and then left the exit zone. This step is necessary to exclude the selection of a reference frame at the time when a vehicle stopped in the zone. If there is no motion in the zone within the specified short time interval, then a reference frame is selected for background estimation.

The specific thresholds are defined to determine points belonging to the shadowor excessive lighting. Filtering such points is necessary to avoid false detections. If a sufficiently large area of the zone is occupied by points related to an object, then the zone goes into the "vehicle detected" state. The final decision about vehicle detection is being taken at the sensor level.

VI. Experiment

Detecting and tracking of vehicles using the traffic surveillance system give more promising way to manage and control the road traffic. In the proposed method Real time vehicle detection and tracking method is used. This algorithm minimizes the constraints like input image resolution, acquisition height, environmental noise etc. on the input file and still able to achieve satisfactory performance. The main purpose is to detect and track vehicles entering an intersection and track them robustly in real-time. The objective here is to improve the quality of the foreground mask generated by reducing unwanted noise and further consolidating the objects in generated foreground. Background subtraction algorithm in the field of image processing and computer vision where an image's foreground is extracted for further processing. This algorithm is used to detect objects by comparing two different frames with matrix distance differences.

The steps of the proposed model as in figure 7.1 are as follows:

Step 1: The video stream from the stationary video surveillance camera is used as an input argument to the system. The input video stream is decomposed into frames where the frames can vary from 1 to n and then system is initialized after that it continuously accepts the video input till the last frame is not reached.

Step 2: All the incoming frames are converted into grey scale format sequentially and then passes through the median filter to remove impulse noise.

Step 3: Frame differencing process is used to segregate the moving pixels from the stationary pixels in the sequence of frames. Based on the difference of the relative positions of the pixels in the subsequent image frames the region of interest is identified. Only those set of pixels which have recorded some displacement and have a pixel density above the pre-defined threshold level are selected as a candidate objects, discarding the remaining pixels. Here the current image C(x, y) is compared with the reference background image B(x, y) and the resultant of this process would produce I (x, y) as the generated foreground mask where x and y represent the coordinates of the pixels. This background image is updated at regular interval so that it did not become obsolete with changing environment.The threshold is calculated by computing the mean of all the pixels. A possible coherent manner between two consecutive frames is assumed. This variance is used to segregate the objects which can be probable vehicles from the input frame.

Step 4: The direction of motion of group of pixels is determined using Optical flow. Optical flow is used to determine the direction of motion of the group of pixels. It is represented as complex variables like: $z = a + ib$. The optical flow calculations are done to capture the degree of displacement of the pixel density between the two subsequent frames. It uses flow vectors of the moving objects over time. Since the optical flow is highly sensitive to noise that is the reason, the median filter is used in the Step 2. The optical flow calculation is used to further segregate the objects which cannot be vehicles; we segregate it computing the motion flow vector corresponding to each pixel of moving object. The threshold is determined by comparing the complex conjugate of the optical flow with the corresponding running mean of the pixels. This calculation further refines the region of interest as itdiscards the various moving objects like animals, moving trees etc. as noise. After applying the velocity threshold, we get the structuring element, which contains the probable vehicles.

Step 5: The structuring element calculated in Step 4 contains random distribution of density that could influence the efficiency by increasing false positives and false negatives in the image. In order to reduce the discrepancies and sequence of morphological operation is used to consolidate the pixel density around the segmented region using structuring element to analyse the image in the neighbourhood of that element. It is placed in all possible places in the image and is compared with the neighbourhood pixels. These operations test whether the structuring element fits within neighbourhood or hits (intersects) the neighbourhood.

Morphological erosion operation is used in this method to eliminate the noise due to the sharp edges found in road environment like road lane markings, sign posts etc. The erosion operation determines only those pixels where the structuring element fits in the given image. Therefore, modified value at locations where structuring element fits the image. It modifies the image matrix f (x, y) that enhanced the quality in terms of noise suppression, preserving the contour and enhance the detection efficiency of the system. The modified foreground mask generated after applying this operation is comparatively better in terms of pixel quality and free from the noise due to surrounding environment.

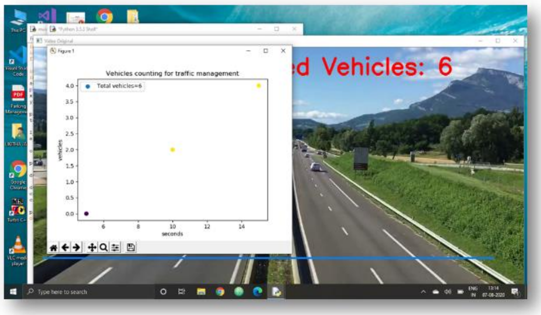
Step 6: Blob Analysis is used to detect any type of 2-dimentional shape of an image. It is used to eliminate the objects which cannot be vehicles like flying birds, pedestrian, etc. For identification of vehicles a threshold is set based on the ratio of the area of blob to area of the bounding box surrounding the blob, (the ratio $\geqslant 0.4$).

Step 7: After performing blob analysis the output is generated. Two variables are maintained to detect vehicles and count the number of vehicles detected and tracked. Whenever a new blob is detected as vehicles it first check to see if that blob is already registered in the buffer, if not then increments the count by one assuming it as a new vehicle else it is treated as a part of registered blob and it is ignored. Finally, the output is presented with bounding box surrounding the detected vehicles, and it is tracked in the video until the vehicle is visible.

VII. Results and Conclusion

In this work background subtraction algorithm is successfully applied in a continuous image. This approach is used as the main tracking algorithm, with morphological and median filtering to remove noise. The success of the foreground and background segmentation and found the object coordinates. Trained the model for their ability to learn complex functions, generalize effectively, tolerate noise and support parallelism.

Fig 4 Figure showing detection of vehicles of frame



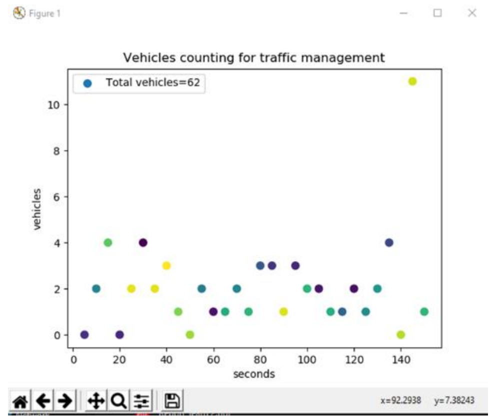


Fig 5 Graph showing density of vehicles when vehicle count is 62

The vehicle count has been effectively increased after the camera calibration. Whereas the results was not so encouraging before calibration. On the other hand, sometimes two or more vehicles moving in the same direction with same velocity is detected as one vehicle due to partial and complete occlusion. Comparison and integration of these pictures from different angle can address problem of occlusion. However, further research work is required to address this. The statistical noise and some inaccuracies related to recursive method are also affected the vehicle detection at some stages. This problem can be reduced by using another Bayesian recursive algorithm like particle filter.

References

[1] K. Meenatchi and P. Subhashini, “Multiple Object Tracking and Segmentation in Video Sequences” International Journal of Advance Research in Computer Science and Management Studies, Volume 2, Issue 5, May 2014

[2] VPS Naidu and J.R. Rao, “Object Tracking using Image Registration and Kalman Filter”.

[3] Mr. D. W. Chinchkhede and Mr. N. J. Uke, “Image segmentation in video sequences using Modified Background Subtraction,” International Journal of Computer Science & Information Technology (IJCSIT) Vol 4, No 1, Feb 2012.

[4] P. Spagnolo, T.D’ Orazio and M. Leo, A. Distante, “Moving object segmentation by background subtraction and temporal analysis,”Istituto di Studi sui Sistemi Intelligenti per l’Automazione—CNR, Via Amendola 122/D-I, 70126 Bari, Italy.

[5] P. Wayne Power and Johann A. Schoonees, “Understanding Background Mixture Models for Foreground Segmentation,” Proceedings Image and Vision Computing New Zealand 2002 University of Auckland, Auckland, New Zealand 26–28th November 2002.