ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A ROBUST FRAMEWORK TO DETECT MOVING VEHICLES IN DIFFERENT ROAD CONDITIONS IN INDIA

[1]Dr.S.APPAVU ALIAS BALAMURUGAN, [2]BALAJI GANESH RAJAGOPAL, [3]Dr.KUMAR PARASURAM

[1]Professor and Research Director, E.G.S.Pillay Engineering College, Nagapattinam

[2]Research Scholar, Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli

[3]Assistant Professor, CITE, Manonmaniam Sundaranar University, Abishekapatti, Tirunelveli

E-mail: [1]datasciencbala@gmail.com, [2]balajiganeshr@gmail.com, [3]kumarcite@gmail.com

ABSTRACT

Traffic situation in India is a quite complex in nature when compared to the traffic models in other nations. It is very essential to model the traffic nature in Indian roadways, both rural and urban roads. Indian road conditions are predominantly occupies different classes of roads viz. single, double, multi-way, cross junctions etc. This research article addresses the different nature of Indian roads with an insight to model the traffic situations in different weather conditions also. The proposed system tries to solve the problem of counting and classifying the vehicles in Indian road conditions. The system uses color image based foreground moving object detection by preserving the color and model of the moving vehicles. The color image based background subtraction technique is supported by cascaded linear regression. The system also uses HoG for contour creation and extraction followed by morphological dilation to connect the missing pixels in the vehicle object. The framework uses adaptive Support Vector Machines to train and model the different classes of vehicles. It has been found that the proposed framework shows an accuracy of 92% in varying levels of traffic density, Illumination conditions.

Keywords: *Vehicle detection, Vehicle counting, Low quality video, Color image based background model, MoG, HoG, SVM Classifier*

1. INTRODUCTION

Road Traffic Surveillance in India is the key area of research in Intelligent Transportation Systems (ITS). As India is a democratic and developing country it paves way for vehicles transportation due to massive population. The growth in automobile sector is a positive sign for every developing nation but however there should be parallel growth in technology to monitor and track every vehicles. It involves traffic Analysis based on different situations like traffic load, vehicle movement, and considering all these factors in different climatic condition. The real scenario of road traffic approach is by considered with real road traffic environment and expose the output in the form of Traffic Flow, Vehicle Count and the Density of the Road in terms of High Congestion or Low Congestion. Introductory part of this chapter focuses on various environmental factors of Indian Traffic Conditions and highlights the behavior of the traffic surveillance in different climatic condition.

There are many existing solutions to analyze the Traffic Situations in the real road traffic, yet this work discusses the Indian Traffic Situation analysis by considering the various external factors which include the abnormal behavior of the road user. This work suggest some algorithms to analyze the Indian traffic conditions which in turn paves a way for generating input for more ITS Applications like vehicle identification, traffic surveillance using ATIS in various climatic condition. This Proposed System also predicts the Efficient Road Map Analysis which predicts the Dynamic road conditions and recommends the travel path with map between the specified Source and Destination Stations. But, the aforesaid applications and systems require intensive data sets in - terms of the entire geographical map with static cost parameter between the all the available points of travel.

This work comes out with the method for Traffic Surveillance on Indian Roads to identify vehicles based on its volume and count, which can be fed as input to various Intelligent Transportation problems. The resultant of this project will be a robust system for Traffic Surveillance at various climatic condition including both Day Time and Night Time. This system will generate useful traffic information like Vehicle structure, No. of each model of the vehicle and its direction of motion based on the pose of the vehicle.

There are numerous research methods and prototypes are available in literature which discusses the vehicle classification and Traffic Surveillance only on day and night time and the algorithms designed for day time surveillance never succeeded in different climatic condition. So, there arises a serious necessity for the development of an algorithm for traffic surveillance at Day &Night time in different weather condition. The existing work of Intelligent Transportation System showcases the emerging area to solve multiple transportation issues. Advanced Travelers Information System is a core area of ITS which gives amenities to the traveler by providing valuable information which can be used by the traveler for path selection, travel time forecasting and time management. As, all of the ATIS applications require large amount of information to conclude the decision for its users, the data procured by these ATIS applications are needed to be precise so that the concluded decision can be justified. On the other side, traffic surveillance systems aid commuters, provides valuable data for traffic cops, traffic infrastructure managers, enforces laws and encourages safe driving. Despite all these things the leverage of software systems for traffic management has emerged. But the focus is narrowed down to develop a core video processing algorithm, which serves the purpose of traffic monitoring. The terms and scope of this work sets the boundary for ITS and its architecture and defines the working principles of video surveillance.

In this technical world, Video surveillance system is a major system being adapted by many companies, organizations and governments. In traffic roads, government takes extra efforts in installing cameras for finding the various data like number of vehicles passing through the road, number of pedestrians walking along the road, etc. There are number of projects running throughout the globe for traffic monitoring system. Almost every projects running concentrate on daytime traffic monitoring. It mainly concentrates on the reflections coming from the road surfaces, which other projects didn’t give much importance.

The main aim of this work is to identify the number of vehicles being passed on the road by eliminating the reflections on the road surfaces. Vision based traffic surveillance systems extract useful and accurate traffic information for traffic flow control, such as vehicle count, vehicle speed, and vehicle classification. Most of the traditional methods concentrate on traffic monitoring in the daytime, and few works address the issue of nighttime traffic monitoring.

Video surveillance has significant application prospects such as security, law enforcement, and traffic monitoring. Visual traffic surveillance using computer vision techniques can be non-invasive, cost effective, and automated. Detecting and recognizing the objects in a video is an important part of many video surveillance systems which can help in tracking of the detected objects and gathering important information. In case of traffic video surveillance, vehicle detection and classification is important as it can help in traffic control and gathering of traffic statistics that can be used in intelligent transportation systems. Vehicle classification poses a difficult problem as vehicles have high intra-class variation and relatively low inter-class variation. In this work, it in investigate on five different object recognition techniques: PCA + DFVS, PA + DIVS, PCA + SVM, LDA, and constellation based modeling applied to the problem of vehicle classification. It also compares them with the state-of-the-art techniques in vehicle classification. In case of the PCA based approaches, we extend face detection using a PCA approach for the problem of vehicle classification to carry out multiclass classification.

2. RELATED WORKS

Computer vision-based approaches for traffic flow surveillance combine information from cameras with other technologies, such as tags installed in vehicles, laser scanners that reconstruct the 3-D shape of the vehicles, or the Global Positioning System (GPS), to estimate the direction of the casted shadows. Compared with intrusive

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

technologies such as radar, inductive loop detectors (ILDs), or lasers, computer vision can be used to obtain richer information, such as analyzing the visual features of the vehicles (color, lights, plate number), apart from the geometry (vehicle volume). These advantages and the increasing computational power of processors have made vision-based systems an area of great interest for road operators, particularly in tolling application. As a mixture between these two groups, we have designed a robust and accurate vision-based system to obtain traffic data flow. On the one hand, it operates in real time and works in challenging scenarios (low-cost cameras, poor illumination, and, in the presence of shadows, unknown perspective. The novelty of our approach relies on a multilayer background subtraction procedure in which the segmentation thresholds can adapt robustly to illumination changes, maintaining a high sensitivity level to new incoming foreground objects and effectively removing moving casted shadows and headlight reflections on the road and in traffic jam situations, in contrast to existing approaches that do not cover all of these functionalities at the same time [6]. A tracking module provides the required spatial and temporal coherence for the classification of vehicles, which first generates a background model of the scene followed by bilateral filtering to remove the noise.

With advancement in image processing and computer vision techniques, much of the research has been done in the field of moving object's regions of change detection among multiple captured image sequences [13–16]. We categorise the vehicle detection and segmentation techniques based on the approach used in each technique into three methods which are Background Subtraction; Feature Extraction-based and Motion-based. A similar classification was done by [17] but it lacks comprehensive comparative analysis of the methods presented here.

Its advantage over previous approaches, such as [3]–[5] and [7], is that it can directly apply on existing 2-D tracks to infer the dimension lost due to camera projection. The handling of severe occlusions is out of the scope of this paper, but nevertheless, our approach can be used in practice by applying it to road viewpoints, such as those achievable from cameras installed in highway gantries, with a performance significantly beyond state-of-the-art vision approaches.

We include experimental results with varying weather conditions, on sunny days with moving directional shadows, headlight reflections on the road, rainy days, and traffic jams. We obtain vehicle counting and classification results comparable with those of ILD systems, which are currently the most widely used systems for these types of traffic measurements, while keeping the main advantages of vision-based systems, i.e., not requiring the cumbersome operation or installation of equipment at the roadside or the need for additional technology such as laser scanners, tags, or GPS[9].

A new Vision-based system for detection of Moving Vehicles from Low quality videos, for multiple vehicle types is developed in this paper. This system tracks every passing vehicle for several frames, and obtains the detection and counting of every passing vehicle. This system combines with the vehicle detection and outputs vehicle counting. A color image-based adaptive background subtraction is proposed to improve the reliability of object detection. The cross-lane vehicles are taken into consideration and the repeated counting for one vehicle is avoided, which significantly improves the accuracy of vehicle counting.

Owing to the complexity of the mixed traffic conditions, induction loops are not useful to collect such data. Many of these detectors are sensitive to road reflections, illumination changes. These detectors always ignore cross-lane vehicles and may repeatedly count one vehicle, which influences their accuracy of vehicle counting [10,12].

The geometric parameters of vehicles were exploited for classification, which makes the classification method universal and does not need training and learning. The features used in many existing algorithms could be robust when video quality is ideal. However, these features based on point, line, and image region could be highly unreliable with a low video quality[4,8]. Therefore, it could not be a good choice to use them for segmentation and tracking of individual vehicles and even 3-D reconstruction. Regression analysis can also be applied to count Vehicles in [16], which is similar to our proposed algorithm. As for vehicle detection, we compare the characteristics among different background subtraction methods, frame differential method, etc. and adopt a mean shift based background subtraction method to adapt the changeable traffic scenes[2].

Pang, Lam and Yung proposes a novel method for accurately counting the number of vehicles that are involved in multiplevehicle occlusions[1], based on the resolvability of each occluded vehicle, as seen in a monocular traffic image sequence[7,8]. A contour description model is utilized to describe the direction of the contour segments with respect to its vanishing points, from which individual contour description and vehicle count are determined.

Buch, Orwell and Velastin presents a detection and classification system for vehicles and pedestrians in urban traffic scenes[11]. It performs per frame vehicle detection and classification using 3D models on calibrated cameras. Motion silhouettes (from background estimation) are extracted and compared to a projected model silhouette to identify the ground plane position and class of vehicles and pedestrians[15]. One cue used to identify target objects in a video sequence is motion. To identify motion that denotes a target object, many background subtraction methods have been proposed; each has strengths, weaknesses and applicability to different scenarios[17]. Different background subtraction methods have been proposed in the literature are described and compared: frame differencing, temporal averaging, $\Sigma - \Delta$, mixture of Gaussians and Kernel Density Estimate[13]. These methods are applied to a selection of representative test footage and the results are presented and discussed so that an informed decision can be made when choosing a background subtraction method for highway surveillance.

Wren, Azarbayejani, Darrell and Pentland defines that Pfinder is a real-time system for tracking people. The system uses a multiclass statistical model of color and shape to obtain a 2D representation of head and hands in a wide range of viewing conditions[9].

Johansson, Wiklund, Forssen and Granlund presents a method that combines shadow detection and a 3D box model including shadow simulation, for estimation of size and position of vehicles[3]. The similarity measure is used in an optimization procedure to find the optimal box state. describe a tracking system that utilizes the estimated 3D boxes, including highlight detection, spatial window instead of a time based window for predicting the vehicle.

The shortcomings of the meanshift tracker, namely the selection of the bandwidth and the initialization of the tracker, are addressed with a fine estimation of the vehicle scale and kinematic model. Indeed, the projective Kalman filter[14] integrates the non-linear projection of the vehicle trajectory in its observation function resulting in an accurate localization of the vehicle in the image.

3. DESIGN METHODOLGY

In this paper, we present a robust vision-based system for vehicle tracking and classification devised for traffic flow surveillance. The system performs in real time, achieving good results, even in challenging situations, such as with moving casted shadows on sunny days, headlight reflections on the road, rainy days, and traffic jams, using only a single standard camera. We propose a robust adaptive multilayer segmentation strategy that detects foreground pixels corresponding to moving and stopped vehicles, even with noisy images due to compression. First, the approach adaptively thresholds a combination of luminance and chromaticity disparity maps between the learned background and the current frame. It then adds extra features derived from gradient differences to improve the segmentation of dark vehicles with casted shadows and removes headlight reflections on the road.

Our algorithm mainly consists of four steps. The first step is the background estimation and foreground segmentation. Background estimation is a necessary pre-processing step for most vision-based systems. Background could be estimated by a simple averaging. The averaging method has little computational cost; however, it could not be robust to different operating conditions. To improve robustness, the background pixel could also be modeled as a single Gaussian, a mixture of Gaussians. The vehicle motion is small in one frame interval, so the movement frame is relatively small in consecutive frames, which can be used as the feature to track vehicles. If the system wants to track the ith vehicle object in the frame k, the pixel-based distances between the FPVH of this vehicle object and the FPVH of all the vehicle objects in the next frame are calculated. The distance can be expressed as follows

Step1: Read Video Input

Step2: Preprocess Video (Frame selection, Noise Removal, Histogram Equalization using bilateral filtering)

Step3: Frame Blurring to select ROI & Image Enhancement

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Step4: Adaptive threshold is used to prepare Foreground masks

Step5: Background Subtraction Using Multilayer Background Subtractions Model

Step6: Image In-painting using MoG to Separate Background & Foreground Moving Objects

Step7: Cascaded Regression to Improve Detection Accuracy of Foreground Objects

Step8: Contour Creation of Blob Detection to the Moving objects Frames

Step9: Count the Number of Blobs Result in Number of Vehicles From Image Frames

Step10: Finally the counts of the vehicles in each frame to be saved in Log file (debug.log)

3.1 Foreground mask and edge:

All detection techniques are based on modeling the background of the image, i.e. set the background and detect which changes occur. Defining the background can be very difficult when it contains shapes, shadows, and moving objects. In defining the background it is assumed that the stationary objects could vary in color and intensity over time. Scenarios where these techniques apply tend to be very diverse. There can be highly variable sequences, such as images with very different lighting, interiors, exteriors, quality, and noise. In addition to processing in real time, systems need to be able to adapt to these changes.

3.2 Environment parameters for Foreground Mask and Edge

Here are the descriptions on the parameters

1. image : Input 8-bit 5-channel image.
2. mask : Input/output 8-bit single-channel mask. The mask is initialized by the function when mode is set to MORPH_CROSS. Its elements may have one of following values:
3. cross : ROI containing a segmented object. The pixels outside of the ROI are marked as obvious background. The parameter is only used when mode== MORPH_CROSS.
4. bgdModel : Temporary array for the background model. Do not modify it while you are processing the same image.
5. fgdModel : Temporary arrays for the foreground model. Do not modify it while you are processing the same image.
6. iterCount : Number of iterations the algorithm should make before returning the result. Note that the result can be refined with further calls with mode==INIT_WITH_MASK.
7. MORPH_CROSS The function initializes the state and the mask using the provided rectangle. After that it runs iterCount iterations of the algorithm.

3.3 Multi-layer Background Subtraction Algorithm

Background subtraction is a major preprocessing step in many vision based applications. In this paper, we present a new multilayer segmentation architecture to fuse different image cues, which combine bottom-up and top-down strategies to solve global/local illumination changes, and to obtain a conditional background model learning. As shown in the experiments carried out and it improves existing segmentation approaches, resulting in a step forward on the current state of the art in vehicle counting and classification using surveillance cameras.

In this paper we present a process that takes advantages of background changeability enabling automating and improving the quality of image segmentation. It is based on an approach that adapts a widely used method for detecting moving objects from a video called background subtraction (foreground detection), to the image segmentation framework. Background subtraction combines local and global thresholding techniques to take advantage of the computational efficiency of the former and of the accuracy of the latter. It is found that the algorithm in figure 1 provides good results in segmentation and allows to automate the image segmentation process when foreground color of images is not constant, as well as to speed it up significantly.

In this algorithm, pixel vectors are formed each for luminance and chrominance. High level noises are removed in the pixel vectors. Intensity levels of the pixels are compared for the shadow, white, foreground, vehicle masks. The shadow and background masks are combined for improving the efficiency of the algorithm. Sobel edge detection algorithm is applied to retain the edges of the highlighted vehicle objects in the image. Watershed segmentation is used to identify and fill the foreground, vehicle and white blobs in the foreground image. Finally mask is reinforced in the direction the shadow. Background mask is updated the mask is returned to perform subtraction.

Journal of Theoretical and Applied Information Technology
15th January 2019. Vol.97. No 1
© 2005 – ongoing JATIT & LLS
ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$D_l \leftarrow Per$ xy pixel : $\left|I_{xy}^{y} - B_{xy}^{y}\right| - k_l.(\Sigma_l)xy$

$D_c \leftarrow Per$ xy pixel : $\sqrt{\left(I_{xy}^{c_x} - B_{xy}^{c_x}\right)^2 + \left(I_{xy}^{c_y} - B_{xy}^{c_y}\right)^2} - k_c.(\Sigma_c)$

Apply noise removal filtering to D_l and D_c
$t_l \;\&\; t_c \leftarrow t_{(l,c)} = \min(D_{(l,c)} + o_{(l,c)})$
if $(D_l)_{xy} > t_l$ U $(D_c)_{xy} > t_c$ (per xy pixel) **then**
if $B_{xy}^{y} > I_{xy}^{y} \cap \left|B_{xy}^{y} - I_{xy}^{y}\right| < o_s$ **then**
$Mask_{xy}$ = shadow
else if $B_{xy}^{y} > I_{xy}^{y} \cap \left|B_{xy}^{y} - I_{xy}^{y}\right| \geq o_s$ **then**
$Mask_{xy}$ = black
else if $I_{xy}^{y} - B_{xy}^{y} > o_h \cap I_{xy}^{y} \leq o_w$ **then**
$Mask_{xy}$ = highlight
else if $I_{xy}^{y} > o_w$ **then**
$Mask_{xy}$ = white
else
$Mask_{xy}$ = foreground
end if
if $Mask_{xy}$ = (shadow U black) ∩ $(D_c)_{xy} > t_c$ ∩ $(D_l)_{xy} - t_l < (D_c)_{xy} - t_c$ **then**
$Mask_{xy}$ = foreground
end if
else
$Mask_{xy}$ = background
end if
E_I, E_B ← Apply $Sobel_{xy}$ to I and B with t_g
Add (E_I – E_B) to Mask as foreground
Crop highlight regions in Mask
Watershed (t_f): fill foreground & highlight & white blob in between as foreground
Reinforce Mask from shadow directions
Update B, Σ_l, Σ_c
return Mask

Figure 1 Algorithm for Background Subtraction

3.4 Contour Creation

Contours can be explained simply as a curve joining all the continuous points (along the boundary), having same color or intensity. The contours are a useful tool for shape analysis and object detection and recognition. A contour line or is-online of a function of two variables is a curve along which the function has a constant value. It is a cross-section of the three-dimensional graph of the function f(x, y) parallel to the x, y plane. Contour creation, is an artistic technique used in the field of art in which the artist sketches the contour of a subject by drawing lines that result in a drawing that is essentially an outline; the french word contour meaning, “outline.” The purpose of contour drawing is to emphasize the mass and volume of the subject rather than the detail. The focus is on the outlined shape of the subject and not the minor details. The flow of the proposed framework is detailed in figure 8 with its high level architecture in figure 9

```
contours, hierarchy = cv2.findContours(fg_mask)
cv2.RETR_EXTERNAL,
cv2.CHAIN_APPROX_SIMPLE)
   log.debug("Found %d vehicle contours.",
len(contours))
   matches = [ ]
   for (i, contour) in enumerate(contours):
      (x, y, w, h) = cv2.boundingRect(contour)
      contour_valid = (w >=
MIN_CONTOUR_WIDTH) and (h >=
MIN_CONTOUR_HEIGHT)
      log.debug("Contour #%d: pos=(x=%d,
y=%d) size=(w=%d, h=%d) valid=%s", i, x, y, w,
h, contour_valid)
      if not contour_valid:
        continue
     centroid = get_centroid(x, y, w, h)
      matches.append(((x, y, w, h), centroid))
```

Figure 2. Algorithm for Vehicle contour creation

Once the background subtracted image is obtained, vehicle contours needs to be found. Contours centroid can also be calculated and contours can be matched to keep track of the vehicle. The algorithm shows in figure 2 is implemented with OpenCV library in python which yields good results for the contour creation. The width and height of the contours can be made static to determine the size of the contours.

3.5 Cascaded Regression Model

Compared to detection, tracking has both additional challenges and opportunities. Arguably the most important aspect in this domain is updating a tracker's models as tracking progresses, also known as incremental tracking.

The recent appearance of regression-based methods that directly infer the whole shape has revolutionized the facial land marking problem, and they have quickly become the state-of- the-art approach. The most notable exemplar is the

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Supervised Descent Method. Its main characteristics are the use of the cascaded regression approach, the use of the full appearance as the inference input, and therefore mentioned aim to directly predict the full shape.

In this article we argue that the key aspects of the algorithm are the use of cascaded regression and the avoidance of the constrained optimization problem that characterized most of the previous approaches. Instead, we show that, surprisingly, it is possible to achieve correct, the enhanced output is obtained based on the lower and higher size of the frames, these processes are repeated until processing all the frames from the video.

3.6 Vehicle Detection and Vehicle Counting

The Steps involved in Vehicle Detection and Vehicle Counting are summarized as follows:
Step 1: First create the divider line for the detection purpose
Step 2: Then, detect the mask of the each processed image
Step 3: Now remove the background from the frame mask
Step 4: Then foreground mask filter from the subtracted frame
Step 5: Then Centroid assigned for each detected mask for each frames in the video

4 RESULTS AND DISCUSSIONS

We generated our own dataset from cameras monitoring road traffic in Chennai, India. To ensure that data are temporally uncorrelated, we have sampled frames at 0.5 fps from multiple video streams. We have gathered 200 video sequences using Oncam Grand eye 360° with fisheye lens as an omnidirectional camera, and Samsung SNP-3500 as a PTZ camera for the experiments. The cameras were mounted at the front side of a building where there is comparatively more traffic circulation rather than other buildings. The developed algorithm is tested on all captured video sequences in offline using a standard personal computer. The frame rate of the video sequence is 25 fps and resolution is 640x480. The results obtained in different weather conditions and for different road types is given in Annexure-I.

4.1 Extract Frames

The human visual system appears to be capable of temporally integrating information in a video sequence in such a way that the perceived spatial resolution of a sequence appears much higher than the spatial resolution of an individual frame. While the mechanisms in the human visual system that do this are unknown, the effect is not too surprising given that temporal adjacent frames in a video sequence contain slightly different, but unique, information.

Bayesian frame extraction algorithm was used to estimate a single high resolution frame given a short low-resolution video sequence. Provided that the object motion has sub pixel resolution, the estimate computed by the frame extraction algorithm has the potential to be substantially improved over single frame interpolations. The number of videos captured in different scenario can be found in Table 1.

Table.1 No. of videos captured in different environments and their contribution to the Training Data set

Video Scenario	Input Videos (200)	% in the Training Data set
Single Way at Normal Scenario	50	25%
Double way at Normal Scenario	60	30%
At Snow Days	20	10%
At Winter Season	20	10%
Night Time Video	50	25%

4.2 Noise Reduction in Video

Video acquisition, transmission, and recording systems are not perfect. Consequently, video sequences may be subject to different types of noise and distortion. For digital video, most of the unwanted artifacts result from side effects of the lossy compression techniques is used to store or transmit the video using fewer bits.

The main types of artifacts that may be present in digital video sequences are as follows: Block Edge Effect (or Simply Blockiness) are Caused by coarse quantization of DCT coefficients in block-based compression schemes (e.g., M-JPEG and MPEG-1), which result in noticeable intensity discontinuities at the boundaries of adjacent blocks in the decoded frame. False Edges (or False Contouring) are caused by coarse quantization of amplitude levels.

Its visible effect in the frame is the appearance of edges at places where a smooth intensity transition should occur. Ringing is caused by the use of ideal filters which produces a rippling along high contrast edges. Blurring and Color Bleeding are caused by imperfections in compression algorithms. This Results in loss of

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

spatial details, running or smearing of color in areas with complex textures or edges. Mosquito Noise is caused by imperfections in compression algorithms. Finally it produces fluctuation of luminance/chrominance levels around high-contrast edges or moving video objects.

Blotches are Large, uncorrelated, bright or dark spots, typically present in film material present in the video sequences due to mishandling or aging and finally Intensity Flicker which are Unnatural temporal fluctuation of frame intensities that do not originate from the original scene.

Finally, the preprocessed frames are obtained after applying the transformations. After preprocessing, the quality of the frame is enhanced by implement the video enhancement process. Typically, video enhancement is defined as the process of enhancing the features of the video for improving the quality and visibility of the image. The contrast stretching and histogram equalization process are applied in the enhancement stage.

Based on the lower and upper threshold limit, the contrast of the video is stretched. Moreover, histogram is the statistical probability distribution of each grey level in the image, which also increases the contrast of the image. Then, it changes the brightness of the image by preserving the brightness.

A picture-perfect video contains equal number of pixels with all its grey levels. In this algorithm, the filtered output Y is given as the input, then the dimension and type of the frames are verified. If both are correct, the enhanced output is obtained based on the lower and higher size of the frames, these processes are repeated until processing all the frames from the video.

4.3 Extract HOG, SVM Classifier

Classes of objects such as a vehicle that we are training vary so much in color. In contrast, Structural cues like shape give a more robust representation. Gradients/derivative of specific directions captures some notion of shape. To allow for some variability in shape, we'll use the Histogram of Oriented Gradients (HOG) instead. The idea of HOG is instead of using each individual gradient direction of each individual pixel of an image, we group the pixels into small cells of n x n pixels. For each cell, we compute all the gradient directions and grouped into a number of orientation bins.

We sum up the gradient magnitude in each sample. So stronger gradients contribute more weight to their bins, and an effect of small random orientations due to noise is reduced. This histogram gives us a picture of the dominant orientation of that cell. Doing this for all cells gives us a representation of the structure of the image. The HOG features keep the representation of an object distinct but also allows for some variations in shape.

We can specify the number of orientations, pixels_per_cell, and cells_per_block in computer the hog features of a single channel of an image. The number of orientations is the number of orientation bins that the gradients of the pixels of each cell will be split up in the histogram. The pixels_per_cells is the number of pixels of each row and column per cell over each gradient the histogram is computed. The cells_per_block specifies the local area over which the histogram counts in a given cell will be normalized. We performed HOG to all channels of the image in HSL format.

The Hue value is its perceived color number representation based on combinations of red, green and blue, the Saturation value is the measure of how colorful or or how dull it is, and Lightness how closer to white the color is. The main intuition is that shapes of the change in all these measurements provide a good representation of the shape of a vehicle.

4.4 Day Time Video

Out of the 200 video we have 140 video data sets are captured during day time. These video data sets comprises of video at different types of roads like single way, double way, intersections as given in Table I. Our Algorithm deduces the types of video based on the HOG and selects the filtering method to remove noise in the video. Figure 3 shows the results of the vehicle counting algorithm in two way roads.

Figure 3 (a) Input video frame (b) Multi-cue color background subtracted image (c) Vehicles Identified and counted in Two-way roads in Highway roads and City roads in India

4.5 Night time Video

Out of the 200 video we have 20 video data sets captured during night time. Reflection Suppression, Shadow removal & Head light detection were applied in order to detect the moving vehicles. The self-diagnostic abilities of our algorithm based on HOG and SVM classifies the Input video as Night Time video. The vehicle headlights were detected using Hidden Markov model from first-fourth order neighborhood pixels.

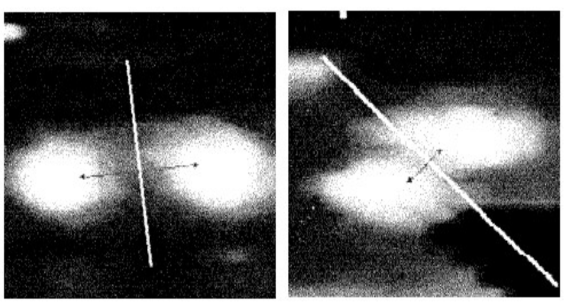

Figure 4. Head light pairing using distribution of Head lights in the neighboring pixels using distribution of Headlights on Intensity, Reflection Intensity and Reflection Suppressed Map

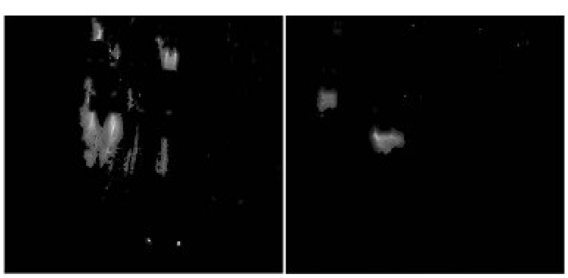

Figure 5. Reflection Suppressed Map

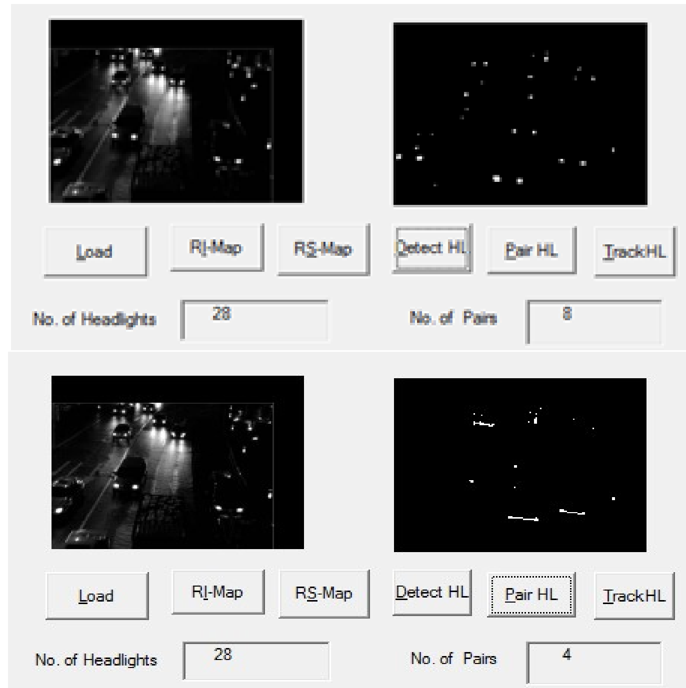

Figure 6. No. of Headlights Detected using the Distribution of Intensity, Intensity Map and Reflection Suppressed Map

4.6 Weather condition – snow and rain time Video

For testing our algorithm for snow video, we have taken the standard testing video dataset available from the library of university of New york. The extracted frames low frequency parts are discarded and only High frequency components are considered. From the High frequency components of the snow image, HOG was applied to describe each snow/rain streaks in the image. Then, Optimized SVM classifier is used to classify the pixel the snow and rain streak pixel. The moving objects are detected using bi-lateral filtering. To eliminate non-vehicles, histogram of every detected rectangle's background model is calculated. A vehicle or non-vehicle decision is made based on the intensity distribution of the background model for the corresponding rectangle. Moreover, height to width ratio of detected rectangles gives information about the classification of detected objects. The snowfall video frame and the resultant frame after moving vehicle detection is shown in figure 7(a) and 7(b).

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

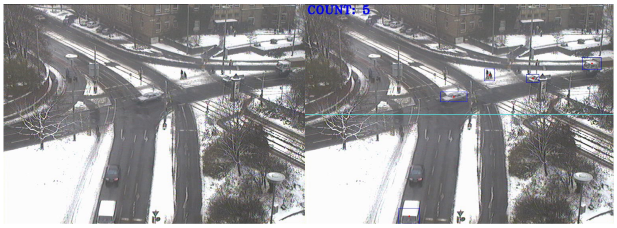

Figure 7(a). The results of the algorithm which removes the snow streaks from the input image and identified the moving vehicles.

Figure 7(b). The results of the algorithm which removes the rain streaks from the input image and identified the moving vehicles.

5. PERFORMANCE MEASURES

For testing the performance of our approach, we performed two kinds of tests: 1) a background subtraction performance to check the sensitivity of our approach to scene variations with respect to recent alternatives and 2) a counting and classification test through a set of videos with different and challenging situations for vision-based systems.

5.1 Performance of Background Subtraction Algorithm

We can measure the sensitivity of background subtraction methods in detecting low contrast targets against the learned background. The lower the threshold values that are needed for detecting foregrounds, the more sensitive the background subtraction approach will be, and hence the better the results. Therefore, for this test, it is not necessary to observe scenes with moving vehicles. Observing background scenes is enough as foreground pixel candidates are artificially generated.

5.2. Counting of Vehicles in different weather conditions

An image of each of the four videos used for this test: 1) a nightfall with diffuse ambient light in which vehicles have headlights turned on (called Nightfall); 2) a sunny day where vehicles project shadows (called Sunny); 3) a rainy day in which the sun appears and causes vehicles to project shadows (called Transition); and 4) a rainy day that includes a traffic jam in which, at times, vehicles are completely stationary in some lanes (called Jam). Each video has duration of 5 min, and the traffic flow in all of them is significantly dense, with several passing maneuvers and vehicles in parallel.

In this test, we have compared our segmentation approach with other vision-based recent alternatives, maintaining the same tracking procedure for the following categories of traffic videos (a) Normal-single way (b) Normal-double way (c) Night Time, and (d) Snow/Rainfall for vehicle counting.

The main difficulty in the Nightfall sequence is correctly segmenting the headlight reflections on the road with respect to the vehicle. The classification errors in light vehicles are mainly derived from this issue as the measured length in some cases includes parts of reflections. In these cases, the reflection intensity is very high, and therefore, some pixels are not classified as highlight. Therefore, the reflection is not completely cropped. In the case of MCB and MFS approaches, as they do not apply any post-processing for removing this kind of region, they have more false positives and classification errors.

In the Sunny sequence, the main difficulty comes from the detection of dark vehicles that project shadows. In some cases, there may be some dark vehicles that do not have sufficient gradient features to identify that they are vehicles. MCB has particular problems in distinguishing these cases as it does not take into account cues other than color.

In the Transition sequence, the main challenges come from the visible slipstreams behind vehicles due to the wet road and the sudden illumination changes, but the proposed system behaves well in most cases. On the contrary, both MCB and MFS have problems with sudden illumination changes as they do not handle them explicitly, provoking more false positives.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

In this section, we evaluate most promising detection algorithms, leaving out their slightly improved variants, from each of the three classifications presented in the previous section. The chosen algorithms for comparison are (1) Gaussian Mixture Model foreground detection, (2) Histogram of Gradients feature detection and (3) Detection based on adaptive motion histogram. The video feed from a stationary camera installed on a highway is first pre-processed to extract frames and thena algorithms. All algorithms were implemented in OpenCV running on a desktop machine with Intel Core i7 processor and 8 GB RAM. The input data of video feed from a camera is collected over the period of 3 weeks under varying illumination and traffic density conditions. Multiple cases were tested for different scenarios and the output from each module is evaluated in terms of well-known performance metrics of Precision and Recall. Both the metrics are defined as:

Precision is the ratio of vehicle detected correctly to the total number of detections in the scene i.e.

$$Precision=tp/(tp+fp)$$

Recall is the ratio of the number of vehicle correctly detected to the actual number of vehicle in the scene.

$$Recall= tp/(tp+fn)$$

where, tp is true-positive, fp is false-positive and fn is false-negative.

The different scenarios under which the algorithms are tested are:
(1) Traffic Density: The detection accuracy depends on the density of vehicles on the road inside the coverage area. It can below or high. Since at different times of the day the traffic density vary the detection algorithms performance also varies.
(2) Illumination Conditions: Lighting affects vision based object detection algorithms. A vehicle detection algorithm which is robust under both bright and dim light condition is preferred in general.
(3) Occlusion: Vehicles or environmental objects may occlude other vehicles. An algorithm's suitability in real life depends on its performance under occluding conditions.

The first sets of experiments were performed under different traffic densities. The low case refers to less than or equal to 4 vehicles in the frame, and high case refers to more than 8 vehicles in the frame during peak rush hours. Table 2 gives the performance of each algorithm in terms of recall and precision. From Table 2, it is evident that HoG performance is better than the other two even when the traffic density is high. GMM precision and recall suffers under high density as it counts vehicles moving close to each other as one.

Table 2.
Impact Of Traffic Density

Algorithm	Low (<4 vehicles)		High (> 8 vehicles)	
	Precision (%)	Recall (%)	Precision (%)	Recall (%)
Mixture of Gaussians	85.71	80.0	82.75	77.41
Histogram Oriented Gradients	92.0	90.19	86.79	88.49
SVM Classification	85.71	88.88	81.35	84.21

Table 3 shows the performance results under different illumination conditions i.e. during day time and night time. It is shown from Table 2 that both precision and recall is less for all the algorithms during afternoon as compared to evening because of strong shadows at day time thereby increasing false positive detections.

Table 3.
Impact Of Illumination Conditions

Algorithm	Afternoon		Evening	
	Precision (%)	Recall (%)	Precision (%)	Recall (%)
Mixture of Gaussians	87.17	77.27	89.47	82.92
Histogram Oriented Gradients	84.50	86.95	88.23	90.90
SVM Classification	89.55	90.85	93.75	92.30

Next, we consider the performance under both occluded and non-occluded case. Table 4 shows the results obtained. It is evident that the performance metrics of each algorithm is higher in the non-occluded case when compared to that of occluded

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

case. In the presence of occlusion, HoG algorithm is more robust than rest as it presents better precision and recall.

Table 4.
Impact Of Occlusion

Algorithm	Occlusion		No Occlusion	
	Precision (%)	Recall (%)	Precision (%)	Recall (%)
Mixture of Gaussians	81.48	68.75	84.61	75.86
Histogram Oriented Gradients	86.27	88.0	91.67	89.79
SVM Classification	84.44	86.36	90.47	88.37

6. CONCLUSION AND FUTURE DIRECTIONS

The Proposed Solution for the Analysis of the Traffic Situation is efficient and produces approximated results on all the video sequence. This Method deduces, whether the road traffic is congested or free flow. This algorithm also detects the Traffic Flow and Occupancy Rate given traffic scene, which makes further analysis of the road traffic. The method describes the step by step process in the analysis and so, there are several midterm results like Foreground Object Extraction in the proposed solution which can be useful in all video surveillance systems. The outputs of this system can be a mandatory input for all kind of Advanced Traveler Information Systems like Travel Time Forecasting, Efficient Road Map Analysis.

The proposed solution can be extended through edge detection, bit masking and more license plate recognition techniques to recognize the vehicles that are violating the traffic. It can also be extended to classify the vehicles based on their structure and total mass. This input can be used for various vehicle classification algorithms. We have presented three different vehicle detection methods used for real-time traffic parameters extraction in video-based surveillance systems, one of the integral components of Intelligent Transportation System (ITS). Vehicle detection is a critical yet challenging step, and its performance varies under different practical scenarios and environment conditions. We assessed the performance of three major vehicle detection algorithms under varying illumination, traffic density and occlusion.

We observed that Histogram of Gradients (HoG) based detection is more robust than Mixture of Gaussians (MoG). And SVM based detection under high traffic density, and occlusion, making it a preferred candidate in these applications. Our research work gives a better insight of different vehicle detection methods and provides a benchmark for performance improvement in vehicle detection under different applications.

7. ACKNOWLEDGEMENT

This work is supported in part by the Science and Engineering Research Board under Department of Science and Technology (DST-SERB), Government of India. The authors wish to acknowledge the financial support from the Department of Science and Technology (DST-SERB), New Delhi, India for the Major Research Project entitled 'Development of Intelligent Traffic Analysis Platform based on the Analysis of Trajectory Big Data' vide reference F.No : ECR/2017/000733.

REFERENCES

[1] "Multilevel Framework to Detect and Handle Vehicle Occlusion," Intelligent Transportation Systems, IEEE Transactions on, vol. 9, pp. 161-174, 2008.

[2] N. K. Kanhere and S. T. Burchfield, "Real-Time Incremental Segmentation and Tracking of Vehicles at Low Camera Angles Using Stable Features," Intelligent Transportation Systems, IEEE Transactions on, vol. 9, pp. 148-160, 2008.

[3] N. K. Kanhere, "Vision-based detection, tracking and classification of vehicles using stable features with automatic camera calibration," ed, 2008, p. 105.

[4] J. Goo, et al., "Tracking and Segmentation of Highway Vehicles in Cluttered and Crowded Scenes,"in Applications of Computer Vision, 2008. WACV 2008. IEEE Workshop on, 2008, pp. 1-6.

[5] N. A. Mandellos, et al., "A background subtraction algorithm for detecting and tracking vehicles,"Expert Systems with Applications, vol. 38, pp. 1619-1631, 2011.

[6] W. Tao and Z. Zhigang, "Real time moving vehicle detection and reconstruction for improving classification," in Applications of Computer Vision(WACV), 2012 IEEE Workshop on, 2012, pp.497-502.

[7] L. Jin-Cyuan, et al., "Image-based vehicle tracking and classification on the highway," in Green Circuits and Systems (ICGCS), 2010 International Conference on, 2010, pp. 666-670.

[8] Jin Mei-Shan, Meng Qing-hui,Huang Fei,Li. Research on the moving vehicle detection technology[C]. The IEEE 2nd International Conference on Consumer Electronics, Communication and Networks, 2012:2046-2048.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[9] Qianwei Zhou, Baoqing Li, Zhijun Kuang, Dongfeng Xie, Guanjun Tong, Liping Hu, Xiaobing Yuan. A quarter-car vehicle model based feature for wheeled and tracked vehicles classification. Journal of Sound and Vibration, 2013, Vol.332(26)

[10] Henry X. Liu, Jie Sun. Length-based vehicle classification using event-based loop detector data.Transportation Research Part C, 2014, Vol.38.

[11] KaewTraKulPong, P., Bowden, R.: 'An improved adaptive background mixture model for real-time tracking with shadow detection'. European Workshop on Advanced VideoBased Surveillance Systems, 2012

[12] Gurwicz, Y., Yehezkel, R., Lachover, B.: 'Multiclass object classification for real-time video surveillance system', Pattern Recognit. Lett., 2011, 32, pp. 805–815.

[13] Shi Zhong-ke Yuan Ji-Wei. A method of vehicle tracking based on gm(1,1). Control and Decision, 2(3), 2011.

[14] KaewTraKulPong, P., Bowden, R.: 'An improved adaptive background mixture model for real-time tracking with shadow detection'. European Workshop on Advanced Video Based Surveillance Systems, 2012

[15] Jung, C.: 'Efficient background subtraction and shadow removal for monochromatic video sequences', IEEE Trans. Multimed., 2009, 11,(3), pp. 571–577.

[16] Shafie, A., Ali, M., Hafiz, F., Ali, R.: 'Smart video surveillance system for vehicle detection and traffic flow control', J. Eng. Sci. Technol.,2011, 6, (4), pp. 469–480.

[17] L. Unzueta et al., "Adaptive multicue background subtraction for robust vehicle counting and classification," IEEE Trans. Intell. Transp. Syst.,vol. 13, no. 2, pp. 527–540, Jun. 2012.

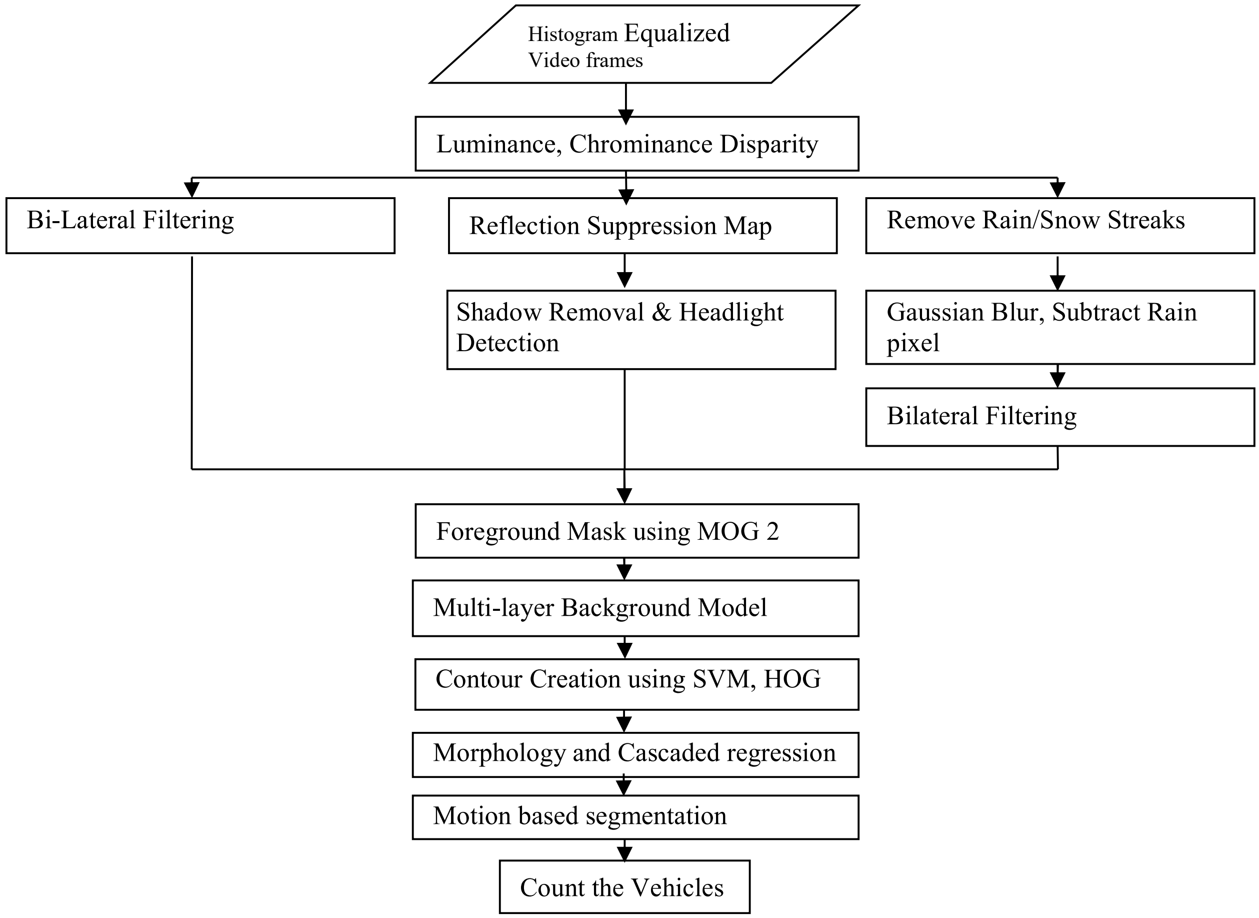

Figure 8 Process flow

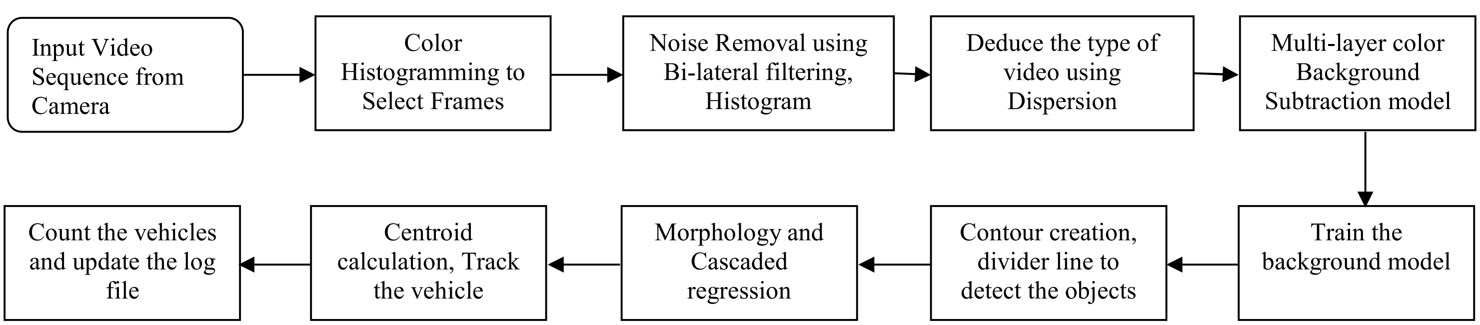

Figure 9 High Level Architecture of the proposed system