

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor & UGC Approved Journal)

Website: www.ijirset.com

Vol. 6, Issue 9, September 2017

Implementation of Optical Flow, Sliding Window and SVM for Vehicle Detection and Tracking

Mohammad Baji¹, Dr. I. SantiPrabha²

M. Tech scholar, Department of E.C.E,U.C.E.K,Jawaharlal Nehru Technological University Kakinada, A.P, India¹

Professor, Department of E.C.E,U.C.E.K,Jawaharlal Nehru Technological University Kakinada, A.P, India²

ABSTRACT: Vehicle detection and tracking are two most challenging tasks of traffic surveillance in transport system. Traditional vehicle detection and tracking methods are computationally expensive and become inefficient in case where light intensity is low and occlusion of vehicles is high. This paper includes detection of moving vehicles, counting of vehicles and tracking of detected vehicles. Here detection is carried out by Optical Flow Method in which a moving vehicle is extracted from the relative motion by segmenting the region representing the same optical flows. The optical flow image uses Horn-Schunk technique in finding the optimum parameters for providing the smoothness. The detected vehicles are tracked using Linear Support Vector Machine Classifier (Linear SVM) with a set of features like Histograms of Oriented Gradients (HOG), Histograms of color features and Geometric features of vehicles, with this details the classifier is trained. This trained classifier is used to search for vehicles in a video, based on the sliding – window technique then its heat map is generated to know the position of a vehicle. Combine this heat maps gives the exact location of the vehicle in images. Experimental results demonstrate that the proposed method is effective in detection & counting and tracking of vehicles.

KEYWORDS: Detection and Tracking, Horn-Schunk method, Optical Flow, Morphological operation, HOG, SVM, Sliding-Window technique.

I. INTRODUCTION

Road Safety with the amount of vehicles is becoming more serious problem. The collision is a main reason for traffic accidents, especially on the highway. Therefore, developing an automotive safety driver assistance system is very important to reduce the accidents. Vehicle detection and tracking techniques are in demand due to numerous applications in Self driving car, traffic surveillance, toll collection, parking lots, and transportation system. For example, counting vehicles over a period of time on a busy intersection helps the concerned authority to efficiently control the duration of traffic signal on a road to reduce the level of traffic congestion during busy hours. Automated counting and tracking of vehicles in self driving car. At present, moving vehicles detection method for video are mainly temporal difference between two consecutive frames [1], image subtraction with background [2], Gaussian Mixture model (GMM) [7] and optical flow estimation [3].

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor & UGC Approved Journal)

Website: www.ijirset.com

Vol. 6, Issue 9, September 2017

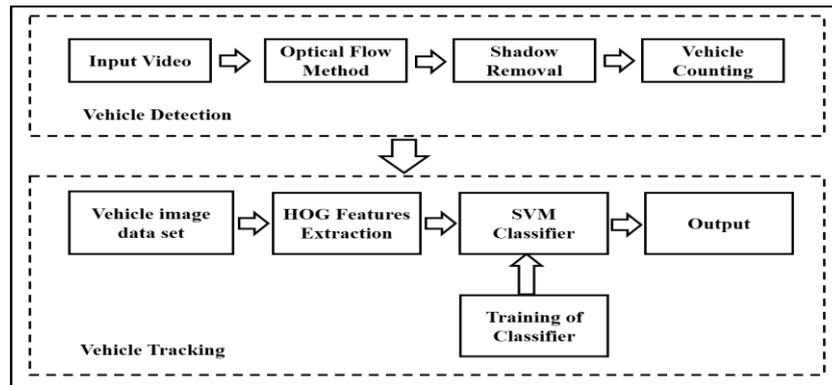


Fig. 1. Block Diagram of vehicle detection and tracking system

Due to the higher detection accuracy of optical flow, it is more suitable for multi-objective moving analysis in complex scenes. Through optical flow estimation, motion parameters of moving objects can be obtained and at the same time, phenomena of occlusion and overlapping of objects may be avoided by applying the morphological operation, which is erosion followed by dilation. Mean filter is used to predict the centroid of each track in the current frame and update to next frame by deleting the previous frame details. Histogram of Oriented Gradients (HOG) [4] is a feature descriptor used in computer vision and image processing for the purpose of object detection. HOG features are calculated by taking orientation histogram of edge intensity in a local region. The extracted HOG features from all locations of a dense grid on a image region and combined features are classified by using linear SVM [5]. The simplest classifier available is the k-nearest neighbor (k-NN) [6] which doesn't have any separate training stage. SVM is a supervised learning method. SVM algorithm uses a set of training examples which creates a model that is used for classify objects in accordance with the data set created.

II. LITERATURE SURVEY

The concept for using the SVM classifier is taken from paper Saran K. B. and Sreelekha G in which they proposed the SVM classifier for human detection. The idea of HOG feature descriptor is taken from the paper N. Dalal and B. Triggs which HOG features from all locations of dense grid on a image region and the combined features are classified by using linear SVM. For the detection of vehicle in an video sequence is taken from paper [3] which uses the optical flow method. Here optical flow uses the Horn Schunck method for the smoothness of the images velocity flow magnitude.

III. RELATED WORK

Proposed method is based on detection of vehicles using optical flow, vehicle counting and vehicle tracking which are explained below.

1. Optical Flow Method:

One of the important blocks presented in the above scheme is called optical flow analysis block. The main purpose of this block is to determine the existence of possible contacts in the incoming video signal and process them in such manner that the median filter is able to detect them with minimal error. The Horn-Schunck algorithms qualify the smoothness of velocity field using the magnitude of gradient of optical flow velocity defined as:

$$f_x u + f_y v + f_t = 0$$

Where u and v are velocity vectors corresponding to Optical flow and $f(x,y,t)$ is brightness of (x,y) at time t .

$$\text{Optimal value } E(w) = \iint \left((f_x u + f_y v + f_t)^2 + \alpha (|\nabla u|^2 + |\nabla v|^2) \right) dx dy$$

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor & UGC Approved Journal)

Website: www.ijirset.com

Vol. 6, Issue 9, September 2017

Where $f_{x,y}$ is partial derivative of f with respect to x,y and $|\nabla u| = \sqrt{u_x^2 + u_y^2}$.

2. Vehicle Counting:

Vehicle is represented using white pixels and the background is represented using black pixels. So the maximum intensity will be at the vehicle position. Thus the centroids can be used to mark the position of vehicles, which is represented as arithmetic mean of x and y directions. Suppose A denotes the entire region, then the centroids (\bar{x} and \bar{y}) are given by:

$$\bar{x} = \frac{1}{|A|} \sum_{x,y \in A} x \quad ; \quad \bar{y} = \frac{1}{|A|} \sum_{x,y \in A} y$$

It can be represented as

$$\bar{x} = \frac{1}{|A|} \sum_{x,y \in A} x^1 y^0 = \frac{M_{10}}{M_{00}} \quad ; \quad \bar{y} = \frac{1}{|A|} \sum_{x,y \in A} x^0 y^1 = \frac{M_{01}}{M_{00}}$$

Where M_{00} denotes the zero order moment, which gives the area ($|A|$) of the region A . So each vehicle's binary mask will have one centroid. The number of vehicles in a frame will be equal to the number of centroids within that frame. For that the position or perspective of the camera is required. For example, in the camera view, if the motion of vehicles is in the vertical direction, then consider two columns of pixels in the frame as real axes (i.e. parallel to x -axis).

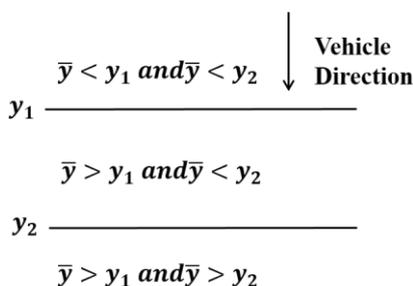


Fig. 2. Representation of real axes for vehicle counting

Where \bar{y} denotes the centroid of vehicle regions and y_1 and y_2 represents two rows of pixels in the frame representing the real axes.

3. Vehicle Tracking:

Vehicles tracking is an important task in the traffic video surveillance, since it can be used to get the type of vehicles or any other information regarding a particular vehicle that is passed through the region of interest. The tracking is done using Linear Support Vector Machine (SVM) with histograms of oriented gradients (HOG) [4] and geometric features has been used.

3.1 Vehicle tracking using SVM

SVM is supervised learning model with associated learning algorithms that analyze data used for classification and regression analysis. Linear SVM is a binary classifier where positive and negative samples are used for training. This SVM is fed with feature extracted from HOG.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor & UGC Approved Journal)

Website: www.ijirset.com

Vol. 6, Issue 9, September 2017

Features Extraction:

A. Histograms of Oriented Gradients

The concept of HOG is presented by Dalal et al., [4], which is an image descriptor mainly used for the Object detection. HOG is calculated by dividing the image into small blocks with 50% overlap. These blocks are again divided into small cells. G_x and G_y are the respective horizontal and vertical gradients of the image, then the gradient magnitude (M) and orientation (θ) can be calculated as:

$$M = \sqrt{G_x^2 + G_y^2}$$

$$\theta = \arctan\left(\frac{G_y}{G_x}\right)$$

A voting is conducted for each gradient orientation. The vote is the gradient magnitude. According to the number of votes obtained, the gradient orientation is quantized into nine bins ranging from 0 to 180. Thus it will form a histogram. The histograms of all the blocks are concatenated to form the feature vector. Since HOG deals with the gradient of the image, it tells how the edges of an object are oriented. Intuitively it gives the information regarding the shape of the vehicle. The HOG of all the vehicle images in the dataset are found and stored for training the neural network. When a new vehicle is detected in a frame, the vehicle is cropped and pre-processed to find the HOG. The size of each image is converted into 32 x 32 and is divided into 8 x 8 blocks containing 2 x 2 cells. The gradient of these cells are used to find the HOG. The HOG of the detected vehicle is fed to the trained SVM and the output determines the class of that vehicle.

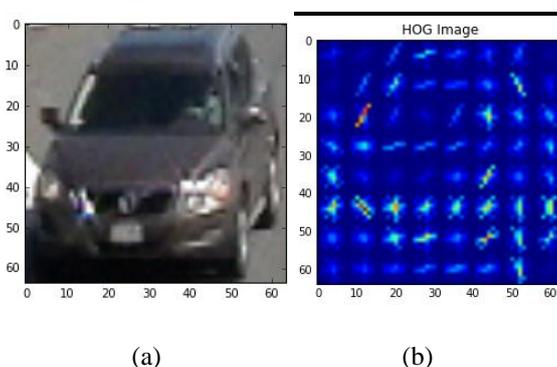


Fig. 3. (a) Test image for features extraction (b) HOG image of Fig. (a)

B. Geometric features:

The dataset contains images of vehicles. When the sizes of vehicles are compared they differ by a lot. Length and breadth of any vehicle will not be same. With these geometric features such as height, length, breadth and area, we can classify the image. The ratio of length to height is given by

$$RLH = \frac{l}{h}$$

The ratio of perimeter to area is given by

$$RPA = \frac{[2(l + h)]^2}{l * h}$$

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor & UGC Approved Journal)

Website: www.ijirset.com

Vol. 6, Issue 9, September 2017

C. Sliding Window Search:

We have to process the image of the vehicle that is in motion in a video. Now we need to extract the feature of the full resolution images into some sub regions and look for the car is present or not. To find the sub-regions of the images a sliding window approach is used. To improve the performance minimal search of sub-region should be maintained and to avoid for looking of car in region of non-interest.

We need to compute the feature vector of the each sub region and feed to the linear SVM, which predict if the car is present or not in the images. By recording all the positions of positive detections in each frame of the video we generate a heat map by fixing a threshold to identify vehicle positions and create the blobs. Using this individual blobs of heat map, combine them into single unit to construct the region of the vehicle.



Fig. 4. (a) Pipeline: search vehicle (b) Heat map of Fig. (a)

IV. EXPERIMENTAL RESULTS

The proposed optical flow model provides the detecting, counting and shadow removal by using the motion vector estimation vector. Below Fig. 5(a) show the vehicle in video and (b) represent the binary mask of (a) from which the centroid is calculated. One centroid represent one vehicle in the proposed method.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor & UGC Approved Journal)

Website: www.ijirset.com

Vol. 6, Issue 9, September 2017



Fig. 5. Vehicle detection result from video (a) Actual frame (b) binary mask after morphological operation

Table I Comparison of detection accuracy of KNN and FBS with SVM and HOG

Tracking	KNN and FBS	SVM and HOG
Total no. of vehicles	58	58
Correct Detection	45	50
Detection Accuracy (%)	77.58	86.20

The images used for creating data set are obtained from the GTI vehicle image database, the KITTI vision benchmarks suite and other websites. The performance of a classifier is determined by the correctness in the classification. Which is quantified by the recognition accuracy. Table: I shows the classification accuracy of SVM with HOG of feature vector compared with KNN classifier and FBS.

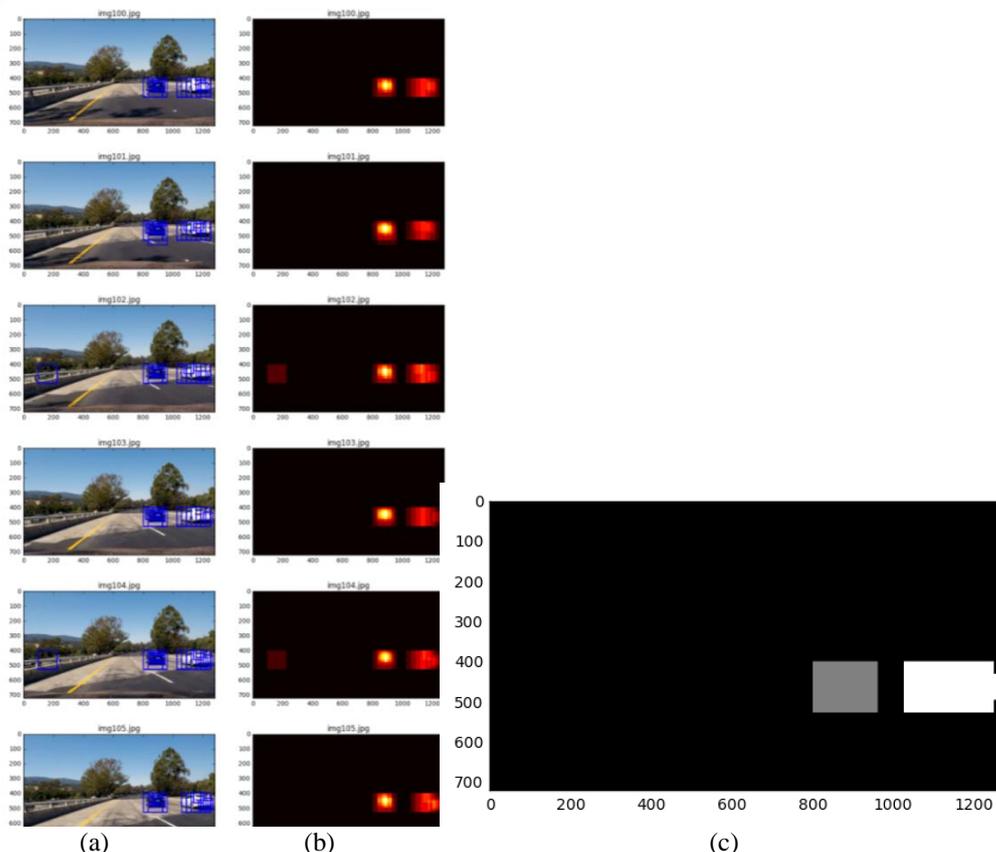


Fig. 6.(a) Tracking of vehicles by using sliding window (b) Heat map of Fig. (a), (c) region of vehicle from combined heat maps

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor & UGC Approved Journal)

Website: www.ijirset.com

Vol. 6, Issue 9, September 2017

The data set of a vehicle and non-vehicle examples are used to train your classifier. These example images come from a combination of the GTI vehicle image database, the KITTI vision benchmark suite. The data set contains of 8792 car images and 8968 not cars images. HOG features are obtained by using 9 orientations, 8 pixels per cell and 2 cell per block. The Features vector is mapped into single array element and fed to trainer. SVM classifier took near 6minutes during training the data to the above set. After the sliding search in fig (a) it is heat map as shown in fig (b). Combine to form the position of vehicles.

Computational Time:

Vehicle detection and tracking is implemented using Python 2.7.3 with OpenCv 3.0.0 and is supported with 3 GB RAM Intel i3 processor. The code took 6 minutes to train the SVM classifier with above data set. Computational time dependence on the data set, processor speed and ram memory. Once the SVM is trained, the network can be ready to use for any number videos until the processor is shut down.

Table II Comparison of computation time of KNN with SVM for a video sequence

Classification Method	Computation Time (sec)
KNN	16.13
SVM	7.25

V.CONCLUSION

In this paper, vehicle detection and tracking has been implemented which focuses on a real time application of self -driving cars and highway traffic. The experimental results validate that the Optical flow gives better performance for detecting the vehicles in a real time videos. The value of optimal parameters for each displacement of Horn-Schunk optical flow have been identified and are necessary in determining image flow sequence providing image with optimal smoothness. For tracking of vehicle, Linear SVM with histogram of oriented gradients (HOG) features, color histogram, and geometry as features are used. The proposed method gives better accuracy compared to k-NN classifier with the Frame Background Subtraction (FBS). The classifier is restricted to moving objects and non-moving objects. This can be extended by developing multi-class SVM which classify the vehicles in different categories (i.e., motorbike, cars, buses, trucks etc.). Some of the challenges in practical cases can also be considered such as night time traffic, different weather conditions and different time instants for future studies.

REFERENCES

- [1] K. Park, D. Lee and Y. Park, "Video-based detection of street-parking violation", in Proc. Int. Conf. Image Process., Computer Vision., Pattern Recognition, Las Vegas, NV, vol. 1, pp. 152–156, 2007.
- [2] P. G. Michalopoulos, "Vehicle detection video through image processing: The autoscope system," IEEE Trans. Veh. Tech., vol.40, no.1, pp. 21–29, Feb. 1991.
- [3] H. Yalcin, M. Herbert, R. Collins, and M. J. Black, "A flow-based approach to vehicle detection and background mosaicking in airborne video", in Proc. CVPR, San Diego, CA, 2005, vol. 2, pp. 1202.
- [4] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, vol. 1, pp. 886–893, June 2005.
- [5] B. Morris and M. Trivedi, "Improved vehicle classification in long traffic video by cooperating tracker and classifier modules," IEEE Int. Conf. on Video and Signal Based Surveillance, vol. 9, p. 9, November 2006.
- [6] NiluthpolChowdhuryMithun, Nafi Ur Rashid, and S. M. MahbuburRhman, "Detection and classification of vehicles from video using multiple time spatial images", IEEE Transactions on Intelligence Transactions Systems, vol. 13, no. 3, September 2012.
- [7] Saran K. B. and Sreelekha G, "Traffic Video Surveillance: Vehicle Detection and Classification," International Conference on control, communication and computing India, pp. 516-521, Nov. 2015.
- [8] Shiva Kamkar and Reza Safabakhsh, "Vehicle detection, counting and classification in various conditions," IET Intel. Transp. Syst., vol. 10, Iss. 6, pp. 406-413, 2016.