ISSN:0975-9646

Deepali Sale et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 7 (1) , 2016, 249-253

Multiclass Vehicle Detection and Tracking Using Local Visual Features

Prof. Deepali Sale, Namrata Bhole, Prof. S.D.Chavan

Dept. Electronics & Telecommunication,
Dr. D .Y. Patil Institute of Engineering and Technology Pune, Maharashtra, India.
SAVITRIBAI PHULE UNIVERSITY, PUNE, MAHARASHTRA, INDIA

***ABSTACT*- Vehicle detection and tracking applications play an important role for civilian and military applications such as in highway traffic surveillance control, management and urban traffic planning. Vehicle detection process on road are used for vehicle tracking, counts, average speed of each individual vehicle, traffic analysis and vehicle categorised objectives and may be implemented under different environmental changes. In this Review, We present a proposed overview of multiclass vehicles detection and tracking using local visual features. There are different types of local visual features among those we expected to use Scalable Invariant Features Transform feature for the detection and tracking purposes.**

***Keywords*-Vehicle detection, Tracking, SIFT (Scalable Invariant Features Transform) features**

1. INTRODUCTION

The moving object tracking in video pictures has attracted a great deal of interest in computer vision. Vehicle detection and tracking is done in two framework. I.e. Training phase and detection phase. In training phase we take multiple images as input images and extract multiple features to trained input sample. In detection phase main basic work is testing. These extracted features provide the evidences to compare unknown state of trained samples which indicate whether a pixel belongs to a vehicle or not. For tracking SIFT features descriptor is invariant to uniform scaling, orientation, and partially invariant to affine distortion and illumination changes. After that for each detected vehicle, feature matching and updating done. It extract SIFT features and establish vehicle information database matching them we track the vehicle. After applying mean shift algorithm, mean shift method, region based method and any other traditional method, these cannot deals with foreground motion of vehicle and also get complexity as background changes to overcome these drawback SIFT is used which can work well in free flowing traffic conditions and the result demonstrate flexibility and good generalization ability.

1.1. Problem Description

We are given a traffic video shot in India. The aim is to detect, track and recognize vehicles from the video. The problem becomes non trivial given the large volume of traffic it captures. Large number of objects leads to some objects fully visible and many objects partly visible due to occlusions from other objects. The problem of detection and tracking partially visible or hidden objects becomes really challenging. The camera is not fixed and there is lot of camera shake which further increases the complexity of the problem. The poor resolution of the video also compounds the detection problem. A vehicle tracking system combines the use of automatic vehicle location in individual vehicles with software that collects these fleet data for a comprehensive picture of vehicle locations. Modern vehicle tracking systems commonly use GPS or GLONASS and GSM technology for locating the vehicle, but other types of automatic vehicle location technology can also be used. When accuracy problems are occurring, the most common reason for the location-based error is because the device's antenna is likely struggling to acquire or maintain satellite signal. If the GPS tracker system or navigation unit is having problems keeping satellite signals one of the most common reasons for this is because of poor placement. To overcome such problems the field of image processing provide the tracking system based on visual transform features.

The objective is to derive relevant properties of a rigid object, moving within a three-dimensional unknown environment, from a sequence of stereo images and other sensor inputs. Each rigid object (vehicle) has specific, time-invariant properties (e.g. size, shape, color, weight etc.) as well as time-dependent properties, such as the relative position and orientation to another object, a particular configuration of the wheels in case of vehicles.

The results demonstrate flexibility and good generalization abilities of the proposed method on a challenging data set with aerial surveillance images taken at different heights and under different camera angles.

2. LITERATURE REVIEW

Here we will present the study on various techniques for vehicle detection and tracking system. The result of detection of vehicles is used as initialization process for tracking. From the decades there are various methods are used for the detection of vehicles. Different methods are

In [1, 2] Frame differencing method: This method detects moving vehicle regions by subtracting two consecutive image frames in the image sequence. Advantage: It works well in case of uniform illumination conditions.

Disadvantage: Frame differencing method does not work well if the time interval between the frames being subtracted is too large.

In [3, 4] Background subtraction method: this method is one of the widely used methods to detect moving vehicle regions. In this method, the background of images is subtracted to highlight the object. It subtracts the generated background image from the input image frame to detect the

moving vehicle regions. This difference image is then threshold to extract the vehicle regions.
Advantage: It is very faster method Works for stationary background.
Disadvantage: Background frames are not adaptive to the environment changes which may create non-existent vehicle regions.
In [5, 6] Feature based method: This method made use of sub-features to detect moving vehicle regions. These features are grouped by analysing their motion between consecutive frames. Thus a group of features segments a moving vehicle from the background.
Advantages: The problem of occlusion between the vehicle regions can be handled well. Feature based methods have less computational complexity compared to background subtraction method.
Disadvantage: If the features are not grouped accurately, then it may be fails in detecting vehicles correctly.
In [7, 8] Motion based method: This method assumes that vehicles tends to move in a consistent direction over time and that foreground motion has different saliency.
Advantage: It is less sensitive to noise and very effecting on small moving objects.
Disadvantage: calculation of motion information consumes time, and it cannot be used to detect static obstacles after vehicle detection, vehicle tracking process is carried out. The approaches of vehicle tracking are as follow: Active contour based method [9, 10]: This method represents vehicle by bounding contour of the object and dynamically update it during the tracking. The advantage of active contour tracking over region-based tracking is the reduced computational complexity.
The disadvantage of the method is their inability to accurately track the occluded vehicles and tracking need to be initialized on each vehicle separately to handle occlusion better.
Feature based method [11, 12]: This method extracts suitable features from the vehicle regions and these features are processed to track the vehicles correctly. The advantage is it has low complexity and also can handle occlusions well. The disadvantage is the recognition rate of vehicles using two-dimensional image features is low, and the problem that which set of sub features belong to one object is complex.
Model based method [13]: This method tracks vehicle by matching a projected model to the image data. The advantages of model based vehicle tracking is it is robust to interference between nearby images and also be applied to vehicle classification. But the disadvantage of this method has high computational cost and they need detailed geometric object model to achieve high tracking accuracy.
Region based method or sliding window based: This method subtracts image frame containing vehicles from the background frame which is then further processed to obtain vehicle regions (blobs). Then these vehicle regions are tracked.
Pixel wise classification method [14]: The novelty lies in the fact that, in spite of performing pixel wise classification, relations among neighbouring pixels in a region are preserved in the feature extraction process. We consider features including vehicle color and local features. For vehicle color extraction, we utilize a color transform to separate vehicle color and non-vehicle color effectively. Advantage of It can work well in free flowing traffic conditions, and the results demonstrate flexibility and good generalization abilities of the proposed method on a challenging data set with aerial surveillance images taken at different heights and under different camera angle.

3. DESIGN OVERVIEW

The moving object tracking in video pictures has attracted a great deal of interest in computer vision. For object recognition, navigation systems and surveillance systems, object tracking is an indispensable first step. Object tracking has significance in real time environment because it enables several important applications such as Security and surveillance to recognize people, to provide better sense of security using visual information, In Medical therapy to improve the quality of life for physical therapy patients and disabled people, In Retail space instrumentation to analyse shopping behavior of customers to enhance building and environment design, Video abstraction to obtain automatic annotation of videos, to generate object based summaries, Traffic management to analyse flow, to detect accidents, Video editing to eliminate cumbersome human operator interaction, to design futuristic video effects. This project presents a real time multiclass vehicle detection and tracking system. The system uses machine learning and feature analysis to detect and track the vehicles on the road. Following fig shows the proposed block diagram for vehicle detection and tracking mechanism.

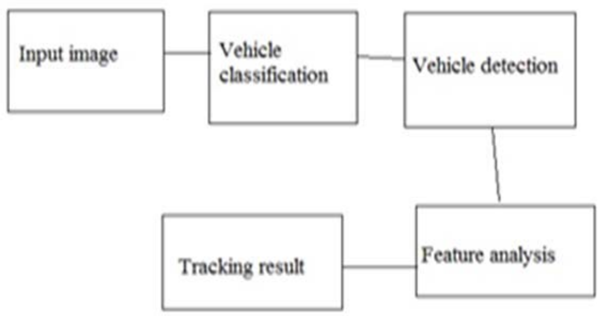

Fig 1: block diagram of proposed work for the detection and tracking of vehicle

Fig shows the overall framework of proposed work. Input image of a video sequence taken from roadway vehicles, system shows the types and locations of the vehicles in the images, then a feature information description of the detected vehicles is obtained, and finally this description is used to match the detected vehicles in the next frame. The framework contains three main processes: vehicle classification, vehicle detection, and vehicle tracking. In the vehicle classification process, using offline learning to create multiclass classifier, once the created multiclass classifier recognizes a potential vehicle in an image, the system generates a train sample for a corresponding vehicle detector. The vehicle detectors were then trained by online learning based on these generated train samples. In the vehicle detection process, using the trained vehicle detectors to classify and locate vehicles from video sequence, while at the same time the vehicle detectors will continue to be trained to improve detection ability. In the vehicle tracking process, the tracker analyses the feature information of the detected vehicles in the previous image frames and matches the feature information of the detected vehicles in the current image. If the matching result is accurate, the tracker outputs the label information for the detected vehicle. The detection result provides the system used for tracking. For tracking sift method is used. This project provides the single and complete framework of detection and tracking.

4. PROPOSED SCHEDULING METHOD

In this project, we design a new vehicle detection framework that preserves the advantages and avoids their drawbacks.

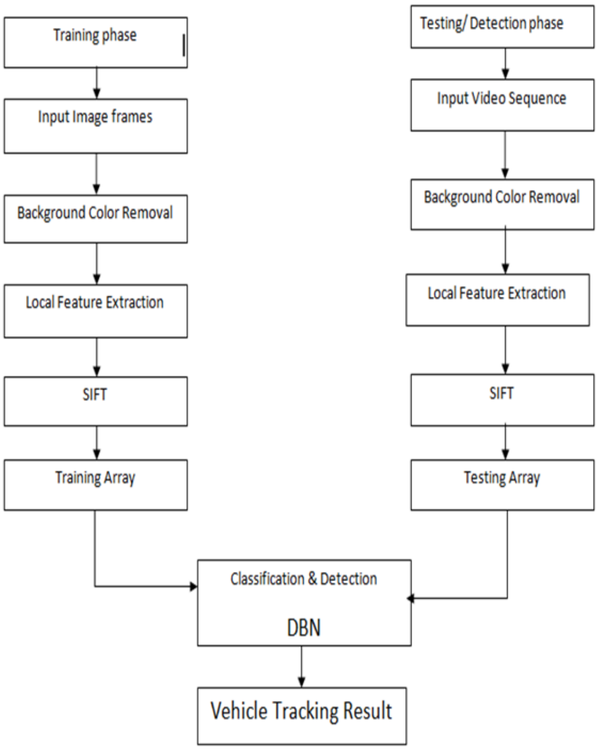

Fig2. Proposed system framework

In this project, we design a new vehicle detection framework that preserves the advantages of the existing works and avoids their drawbacks. The modules of the proposed system frame- work are illustrated in above Fig.
The framework can be divided into -

- Training phase and
- Detection phase.

In the detection phase, we first perform background color removal. Afterwards, the same feature extraction procedure is performed as in the training phase. The extracted features serve as the evidence to infer the unknown state of the trained DBN, which indicates whether a pixel belongs to a vehicle or not. The distinguishing feature of the proposed framework is that the detection task is based on pixel wise classification. However, the features are extracted in a neighbourhood region of each pixel. Therefore, the extracted features comprise not only pixel-level information but also relationship among neighbouring pixels in a region. Such design is more effective and efficient than region-based or multi scale sliding window detection methods.

4.1. Background Color Removal:

Since non vehicle regions cover most parts of the entire scene in aerial images, we construct the color histogram of each frame and remove the color that appear most frequently in the scene. Take Fig. 3. For example, the color are quantized into 48 histogram bins. Among all histogram bins, the 12th, 21st, and 6th bins are the highest and are thus regarded as background color and removed. These removed pixels do not need to be considered in subsequent detection processes. Performing background color removal cannot only reduce false alarms but also speed up the detection process.

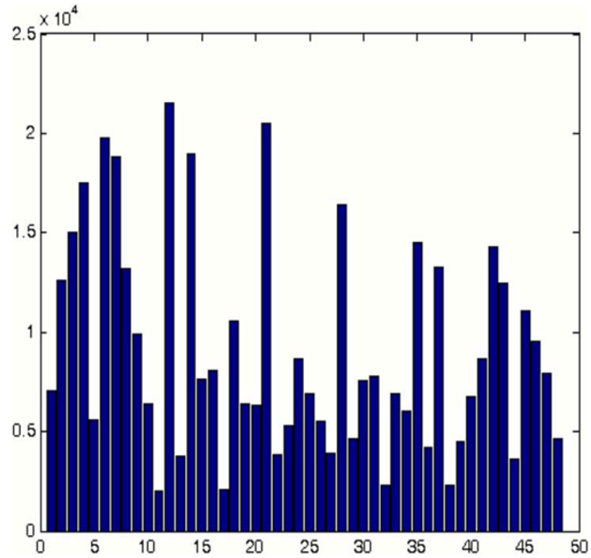

Fig 3: color histogram frame

4.2. Feature Extraction:

Feature extraction is performed in both the training phase and the detection phase. We consider local features and color features in this project.

Local Feature Analysis: Corners and edges are usually located in pixels with more information. Edge detection is an essential pre-processing step in many computer vision algorithms. Within this project we implement one of these methods, the Canny Edge Detector. The Canny edge detector is a popular method for detecting edges that begins by smoothing an image by convolving it with a Gaussian of a given sigma value. Based on the smoothed image, derivatives in both the x and y direction are computed; these in turn are used to compute the gradient magnitude of the image. Once the gradient magnitude of the image has been computed, a process called 'non maximum suppression' is performed; in which pixels are suppressed if they do not constitute a local maximum. The final step in the canny edge detector is the hysteresis operator, in which pixels are marked as either edges, non-edges and in-between, this is done based on threshold values.

The next step is to consider each of the pixels that are in-between, if they are connected to edge pixels these are marked as edge pixels as well. The result of this edge detector is a binary image in which the white pixels closely approximate the true edges of the original image.

4.3. DBN (Dynamic Bayesian Network):

Next we use pixel wise classification for detection by using DBN. The design of the DBN model is illustrated in Fig. 4. Node VT indicates if a pixel belongs to a vehicle at time slice. The state of VT is dependent on the state of Vt-1. Moreover, at each time slice t, state has influences on the observation nodes St, Ct, ET, at and Zt. The observations are assumed to be independent of one another. Discrete observation symbols are used in our system. We use means to cluster each observation into three clusters, i.e., we use three discrete symbols for each observation node. In the training stage, we obtain the conditional probability tables of the DBN model via expectation-maximization algorithm by providing the ground-truth labelling of each pixel and its corresponding observed features from several training videos. In the detection phase, the Bayesian rule is used to obtain the probability that a pixel belongs to a vehicle.

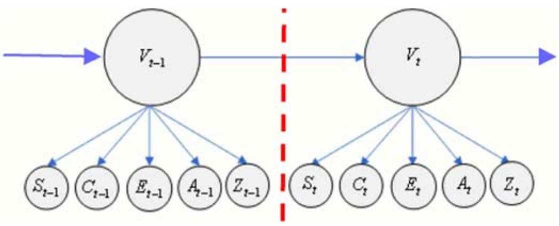

Fig4. DBN model for pixel wise classification

4.4. SIFT feature analysis:

SIFT (Scale Invariant Feature Transform) is a well-established local feature descriptors, which was proposed in 1999. For tracking the vehicles SIFT feature analysis is used. Due to SIFT feature descriptor is invariant to uniform scaling, orientation, and partially invariant to affine distortion and illumination changes, it has been widely applied to object tracking and image matching. For multiclass vehicle tracking, we need a kind of feature which can describe different vehicles accurately, the SIFT feature is very suitable in the circumstance. The SIFT algorithm includes four steps: Scale-space extreme detection, Feature point localization, Orientation assignment and Generation of feature point descriptors.

4.5. Feature matching and updating:

For each vehicle detected from multiclass detection framework, extract SIFT feature and establish vehicle information database (VID). The VID consists of four parts: vehicle class, vehicle number, vehicle location (rectangle coordinates) and SIFT feature point descriptor (feature priority, feature point coordinate, orientation and scale), each vehicle detected from multiclass detection framework is tracked in a new video frame sequences by separately comparing its feature point with the same class vehicle from the VID. The Euclidean distance is introduced as a similarity measurement of feature characters. By matching them we track the object.

5. CONCLUSION

By using this proposed approach, we get high processing speed and more hit rate. Also get flexibility and good generalization abilities of the proposed method on a challenging data set with different images taken at different heights and under different camera angles.

REFERENCES

[1] D. Dailey, F.W. Cathy, S. Pumrin, An algorithm to estimate mean traffic speed using uncalibrated cameras, IEEE Transactions on Intelligent Transportation Systems 1 (2) (2000) 98–107.

[2] T. N. Schoepflin, D. J. Dailey. Dynamic camera calibration of roadside traffic management cameras for vehicle speed estimation, IEEE Transactions on Intelligent Transportation Systems 4 (2) (2003) 90–98.

[3] C.Wern, A. Azarbayejani, T. Darrel, A. Petland, Pfinder, real-time tracking of human body, IEEE Transactions on Pattern Analysis and Machine Intelligence 19 (4) (1997) 780–785. [7] X. Gao, T. Boult, F. Coetzee, V. Ramesh, Error analysis of background adaption, IEEE Conference on Computer Vision and Pattern Recognition, 2000, pp. 503–510.

[5] D. Koller, K. Danilidis, H. H. Nagel, Model-based object tracking in monocular image sequences of road traffic scenes, International Journal of Computer Vision 10 (3) (1993) 257-281.

[6] S. M. Smith, J. M. Brady, ASSET-2: real-time motion segmentation and shape tracking, IEEE Transactions on Pattern Analysis and Machine Intelligence 17 (8) (1995) 814-820.

[7] A. Bainhridge-Smith, R. G. Lane, Deremining optical flow using a differential method, Image and Vision Computing 15 (1) (1997) 11-22.

[8] L. Wilson, Detecting salient motion by accumulating directionally-consistent flow, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (8) (2000) 774–780.

[9] 14. D. Koller, J. Weber, T. Huang, J. Malik, G. Ogasawara, B. Rao, S. Russell, Toward robust automatic traffic scene analysis in real-time, IEEE Conference on Pattern Recognition, 1994, pp. 126–131.

[10] N. Paragios, R. Deriche, Geodesic active contours and level sets for the detection and tracking of moving objects, IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (3) (2000) 266–280.

[11] L. Huang, M. Barth, Real-time multi-vehicle tracking based on feature detection and color probability model, IEEE Conference on Intelligent Vehicles Symposium, 2010, pp. 981–986.

[12] B. Coifman, D. Beymerb, P. McLauchlan, J. Malik, A real- time computer vision system for vehicle tracking and traffic surveillance, Transportation Research Part C: Emerging Technologies 6 (4) (1998) 271–288.

[13] M. Haag and H. Nagel, Combination of edge element and optical flow estimate for 3D model-based vehicle tracking in traffic image sequences, International Journal of Computer Vision 35(3) (1999) 295–319.

[14] D. J. Dailey, F. W. Cathy, S. Pumrin, An algorithm to estimate mean traffic speed using uncalibrated cameras, IEEE Transactions on Intelligent Transportation Systems 1 (2) (2000) 98– 107.

[15] Hsu-Yung Cheng, Member, IEEE, Chih-Chia Weng, and Yi-Ying Chen,"Vehicle Detection In Aerial Surveillance Using Dynamic Bayesian Networks, IEEE, Vol.21, No 4, April 2012.

[16] Zhiming Qian, Hongxing Shi, Jiakuan Yang and Lianxian Duan,"Video-Based Multiclass Vehicle Detection and Tracking", IJCSI Intrnational Journal of Computer Science, Vol.10, Issue 1, No 3, January 2013

[17] Gagan Bansal and Sandeep Mullur," Vehicle detection And Tracking", Johns Hopkins University.

AUTHORS BIOGRAPHY:

Prof. Deepali R Sale is pursuing her Ph.D. in image processing (image fusion). She has completed her M.E. (Electronics) from College of Engineering, Pune. She is presently working as an Assistant Professor in Pad. Dr. D. Y. Patil Institute of Engineering and Technology, Pimpri, Pune. She has more than 10 years of Teaching Experience. She has about 20 papers published in National /International Journals and conferences. She is a member of IETE, India, and IEEE, US.

Ms. Namrata Bhole is pursuing her M.E. Electronics and Telecommunication in Communication Network from Pad. Dr. D. Y. Patil Institute of Engineering and Technology, Pimpri, Pune. She received B.E. Degree in Electronics and Communication Engineering from North Maharashtra University, Maharashtra, India.

Mr. S. D. Chavan, working as Associate Professor M. E. (Electronics), PhD (Appeared) Electronics. Experience consists of more than 14 years of Academics and 5 years of Industrial experience. Professional Body membership includes LMISTE, MIETE.