International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-5, March 2020

Tracking and Detection of Vehicles using Locality Sensitive Histogram (LSH) Feature Extraction

Bhavya R., Geetha K. S.

Abstract: ***Detection and tracking has become a vital chore in most of the computer vision applications. It analyzes the behavior of the object and detects when it appears in other frames. In this paper, a locality sensitive histogram (LSH) algorithm along with SVM is used to detect and track the objects. Locality Sensitive Histogram is used for feature extraction and detection. It is computed at each pixel location, by adding a floating-point value to bin, which is its unique nature. The extracted features are subjected to Linear SVM classifier and then the object is tracked by eliminating false positives. This proposed method precisely tracks and detects the object well with different challenges. Experimental results demonstrate the performance of the proposed algorithm with an accuracy of 89% considering several challenging factors. Evaluation of various other algorithms using different performance parameters is also tabulated in the diagram and shows that the proposed method is topmost performer in tracking the objects. This method can be utilized to track different objects of different scale and track efficiently.***

Keywords: ***Detection, Locality Sensitive Histogram, SVM, Tracking.***

I. INTRODUCTION

Robust visual tracking is still a challenging task as it is expected to handle the sudden changes in the appearance of the target object. Histogram being the strongest statistical tool is being widely used for various Visual analysis and more into object tracking applications. The challenging factors of Object tracking are addressed by extracting various features and proposing different models to represent target objects. In Generative tracking algorithms, it uses a particular feature space like spatiogram/histogram methods to represent the object. These methods lose spatial information and is less effective in handling occlusions, appearance changes [2], [4]. In case of Discriminative algorithms, it considers tracking by differentiating foreground from background as a binary classification problem and classifiers are usually updated for appearance changes, thereby estimating the location of the target object [5]. In this paper, Histogram is used to extract pixel distributions as a function of color variations. Instead of using the local histogram algorithm which used pixels from local neighborhood region, this paper uses Locality Sensitive Histogram (LSH) [1], considering the contributions from every pixels of the image. However, LSH is computed not by considering the frequency of occurrences of each intensity value to a bin, but in addition by adding a floating-point value to the bin. This floating-point value is reduced further as the distance of the pixel reaches closer to the point where LSH is computed. This method thus removes the background effect which would thus improve the accuracy and efficiency of tracking. Additionally, to differentiate the foreground and background, the discriminative classifier using Support Vector Machine (SVM) is used, by classifying the training data to give better prediction accuracy. After training, sliding window approach is used to track the objects which performs effectively under different appearances. This method gives good responses even for tracking objects in color images with sudden changes in appearances.

This paper is organized as follows: Section II describes about the proposed model for object tracking. Section III gives details about the experimental results and evaluation. Section IV concludes with future direction of work.

II. PROPOSED METHOD

A. Dataset

The dataset taken as input represents images containing vehicles and non-vehicles of different categories with different appearances and illumination variations. The dataset contains more than 10000 images of different size and aspect ratio.

B. Preprocessing

The most important step is to maintain uniform aspect ratio and size. This is achieved by down-scaling the images into 64x64 size before subjecting to feature extraction. These input images were categorized as vehicles and non-vehicles and taken to the working directory for further progress.

C. Feature extraction using LSH

A unique feature, called Locality Sensitive Histogram (LSH) is used for feature extraction. LSH effectively handles appearance variations as it utilizes every pixel in the image, operates similar to conventional image histograms. For each intensity value a floating-point number is added to the bin. This exponentially reduces as it moves away from the pixel location where it is computed [1].

Revised Manuscript Received on March 30, 2020.
* Correspondence Author
Bhavya Rudraiah*, Research Scholar, Department of ECE, R V College of Engineering, Bangalore, India. E-mail: bhavyanadgouda@gmail.com
Dr. Geetha K. S., Professor, HOD, Department of ECE, R V College of Engineering, Bangalore, India. E-mail: geethaks@rvce.edu.in

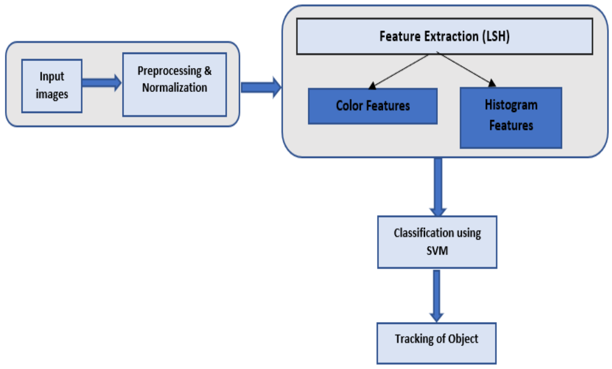

Fig 1: Proposed block diagram.

LSH at a particular pixel p is computed by [1]:

$$H_p^E(b) = \sum_{q=1}^{W} \alpha^{|p-q|} \cdot Q(I_q, b), \; b = 1, \ldots, B \qquad (1)$$

where α ε (0,1) is a weight controlling parameter, which decreases as the pixel moves away from the target. For a 1D image, the LSH is computed as:

$$H_p^E(b) = H_p^{E,left}(b) + H_p^{E,right}(b) - Q(I_p, b) \qquad (2)$$

Where

$$H_p^{E,right}(b) = Q(Ip,b) + \alpha . H_{p+1}^{E,right}(b) \qquad (3)$$

$$H_p^{E,left}(b) = Q(Ip,b) + \alpha . H_{p-1}^{E,left}(b) \qquad (4)$$

Based on the above equations it can be concluded that, pixels on right side of **p** contribute to right side LSH and pixels on the left side of **p** contribute to left side LSH. Hence combining both LSHs, results in total LSH which includes contributions from all the pixels [1]. For color images, there is an increase in bins count exponentially when represented in full color space. In order to reduce this, spatial binning of 16x16 is used to reduce the feature vector size. The color-space is converted from RGB- YUV and stacked as the feature list resulting with 64x64 size feature vector.

D. Classification using SVM

Classification is a task of assigning label to the images. The obtained feature vectors are input to Linear SVM classifier [7]. Linear SVM Classifier is a machine learning algorithm that solves multiclass classification problem. Using large datasets, it designs Linear SVM that implements proprietary version of a cutting plane algorithm [13]. Let us consider dataset as **{x_i, y_i},** where $\mathbf{x_i}$ is the input vector space, $\mathbf{x_i}$ **ε X** and $\mathbf{y_i}$ **ε {-1,1},** the label associated for classification. In order to find hyperplane **(W, x) – b,** where **W, b** is determined using the learning algorithm. The decision rule for this classifier is:

$$f(x) = sgn((W,x) - b) \qquad (5)$$

In case of linear classification problem, there would be infinitely many solutions. Hence to choose a best hyper plane, additional constraints like generalization error has to be minimized and distance between the support vectors has to be maximized [6]. To obtain a best hyper plane, Lagrange multipliers technique is used for **W** which is a linear combination of input vectors given as:

$$W = \sum_{i=1}^{N} \alpha_i x_i y_i \qquad (6)$$

where $\boldsymbol{\alpha_i}$ – Lagrange multiplier of $\mathbf{i^{th}}$ vector.
Hence the decision rule is:

$$f(x) = sgn\left(\left(\sum_{i=1}^{N_s} \alpha_i y_i x_i\right) - b\right) = sgn\left(\sum_{i=1}^{N_s} \alpha_i y_i (x_i, x) - b\right) \qquad (7)$$

With this implication, before training the classifier, the feature vectors were subjected to pre-processing which includes normalization and scaling using standard scaling, and splitting of dataset into training and test data. This scaling was done to avoid overfitting and test the model on the test data. Later, the Linear SVM classifier is trained by giving the feature vectors of the input images. This trained classifier gives better performance with classification accuracy of 84%-90%. After the classification the trained classifier is used to track the objects which is briefly described in the next section.

E. Object Tracking

In this paper, feature-based tracking is used to track the objects. The feature vectors which identify the object are used by the classifier to track the object. Tracking is done using sliding window search method which detects the object. Sliding windows plays a major role to localize exactly where the object is present in an image. Sliding window method uses a rectangular box that slides across an image. The classifier is used to determine if the window has the object of interest or not. This method results with a window list of varying sizes which includes the object or not with certain features. Multiple detections obtained due to overlapping can be resolved into single detection based on selecting the heat map of the image and finally detecting the required object. Once the object has been detected, false positives were removed using an averaging approach in the video pipeline which gives the total heatmap and applying the threshold to draw final box on the detected object. SVMs performance mainly depends on parameters like feature selection and tuning. These parameters contribute towards the accuracy and generalization of the classifier. Training and classification are other parameters which is important in the implementation of SVMs [7].

III. RESULTS AND DISCUSSIONS

The performance evaluation and efficiency of the proposed model is discussed and evaluated in this section. The proposed algorithm is implemented using Python and developed on Intel i7 CPU. The proposed technique correctly detects and tracks the object continuously by considering LSH for feature.

Performance Analysis: The performance of the proposed method is evaluated on several videos.

International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-5, March 2020

It contains some challenging factors like pose and scale variation and difference in motion. The videos were collected from different repositories.

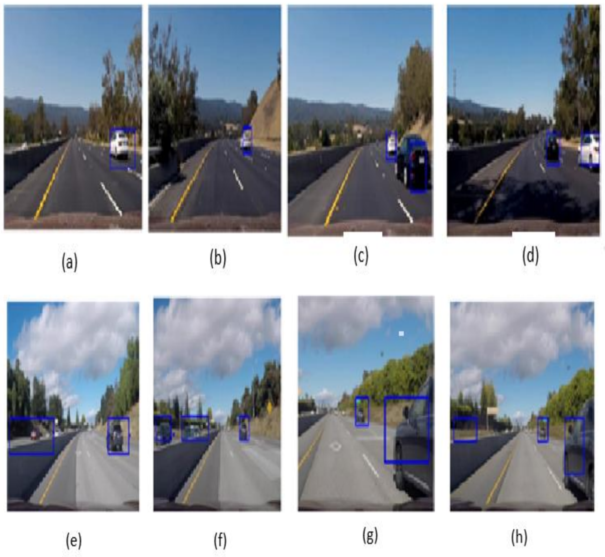

Fig.2: (a)-(h): Screenshots of tracking moving objects using the proposed technique in different sequences is illustrated. The first row Fig (a)- Fig (d) represents object tracking in different frames of video sequence 1. The second row Fig (e)- Fig (h) represents tracking results for video sequence 2. These video sequences have several challenging factors like pose variation, scale variation, motion differences. Blue box represents the tracked object in all the frames collected from different video sequence.

All of these videos have more than 500 frames. The classifiers which uses different features for detection and tracking is taken into consideration for evaluation. The screenshots shown in Fig.2 (a) – (h) represents the results obtained after the implementation of the proposed method which takes Locality Sensitive Histogram for feature extraction. The video sequences considered have some challenging issues like illumination, pose and scale variation. The first row Fig.2(a)–(d) represents video sequence 1 and second row Fig.2(e)–(h) represents video sequence 2, where these videos have around 500 frames each. The blue color box represents the detected object taken at different frames which include single/multiple objects in the frame. These objects which appear with different pose and scale variation are also tracked with the proposed method effectively in this paper.

Fig.3 & 4 illustrates the performance plot of few classifiers such as KNN (k-nearest neighborhood), MLP (Multilayer Perceptron), J48(Decision tree), which uses color, shape and texture-based features for object detection and tracking. Fig.3 illustrates the performance of several different classifiers with different feature extractions. Fig.4 illustrates the performance plot of the proposed method, using LSH as feature for detection and tracking. From the plot in Figure 4, it can be inferred that there is a remarkable performance in classification accuracy when the k-Nearest Neighborhood uses the proposed method, LSH as feature for classification. It results in a classification accuracy of **84.25%.** Similarly, Decision tree and Support Vector Machine results with a classification accuracy of **88%** and **89.75%** respectively. These classifiers are highly precise in identifying the positive samples when it considers Locality Sensitive Histogram for feature extraction than compared to color, shape or texture features. The accuracy is improved since most of the false positives were averaged out with threshold defined and thereby identifying vehicles clearly even under certain challenges. The other classifiers use different combination features like color, shape, texture for classification. K-Nearest Neighborhood results in the classification accuracy of 45%-63%. J48 of 65% and MLP with 45% accuracy. These results infer that the classifiers are able to correctly predict the samples within range of 45% - 65 %.

It can be concluded from fig.3 & 4, by comparing all the parameters, that the classifier which uses LSH gives better performance with increased accuracy, and precision. The proposed object tracker performs well in classification and can also be extended by considering illumination invariant features. This is due to the unique property of LSH as, it not only considers the pixel values but rather convey the color information also which is very effective to detect and track the objects. With these results, it can be stated that the proposed method in this paper performs well than the other methods included in the research.

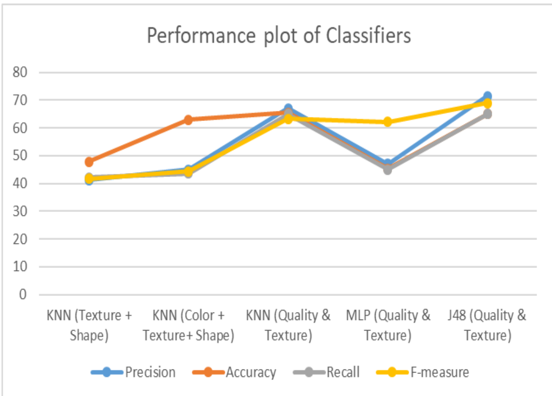

Fig.3: Performance Plot of Classifiers using Color, Shape and Texture as Features.

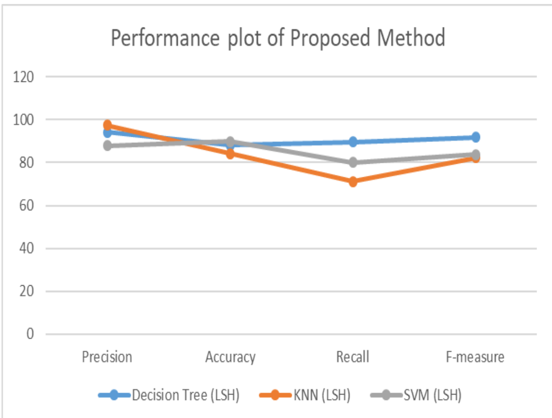

Fig.4: Performance Plot of the Classifiers considering LSH as Features.

Table I: Performance measure of Different classifiers.

Classifiers (Features)	Precision	Accuracy	Recall	F-measure
KNN (Texture + Shape**)**	41.13	47.88	42.19	41.65
KNN **(**Color + Texture+ Shape)	44.99	63.01	43.62	44.29
KNN **(**Quality & Texture**)**	67	65.5	65	63.2
MLP **(**Quality & Texture)	47	45.3	45	62.1
J48 **(**Quality & Texture)	71.4	65	65	69

Table II: Performance measure of proposed method with different classifiers.

Classifiers (Features)	Precision	Accuracy	Recall	F-measure
Decision Tree (LSH)	**94.05**	**88**	**89.63**	**91.78**
KNN (LSH)	**97.31**	**84.25**	**71.07**	**82.15**
SVM (LSH)	**87.8**	**89.75**	**80.1**	**83.7**

The Performance measure of different classifiers with different features are tabulated in Table I. Using the proposed method, the performance measure of classifiers for LSH is tabulated in Table II.

IV. CONCLUSION

In this paper, the proposed object tracking algorithm which uses Locality Sensitive Histogram feature for tracking outperforms well with an accuracy of 90% by considering some challenging factors like pose variation, scale variation. This method gives better results by using SVM and other different classifiers which yields an accuracy and precision of 90% and 94% when compared to other existing methods. The classification results are accurate due to reduced false prediction. Therefore, detection and tracking still a hotspot in most of the surveillance applications. Hence, future direction would be to outspread the research of detection and tracking of the object by considering challenging factors like crowded scene, illumination variation, abrupt motion and occlusion.

REFERENCES

1. S. He, Rynson W.H. Lau, Q.Yang, J Wang, M Yang, *"Robust Object tracking via Locality Sensitive Histograms"*. In IEEE Transactions On Circuits And Systems For Video Technology, 2016.
2. J. Kwon and K. Lee, *"Visual tracking decomposition*". In Proc. IEEE CVPR, pages 1269–1276, 2010.
3. Z. Hong, X. Mei, D. Prokhorov, D. Tao. "*Tracking via robust multi-task multi-view joint sparse representation"*. In Proc. ICCV, pages 649–656, Dec 2013.
4. X.Jia, H.Lu, M.-H.Yang. "*Visual tracking via adaptive structural local sparse appearance model"*. In Proc. IEEE CVPR, pages 1822– 1829, June 2012.
5. T.Mahalingam, M. Subramaniom, "*A Robust Single and Multiple moving object detection , tracking and Classification"*, Applied Computer and Informatics, 2018.
6. Benjamin Castaneda, "*Support Vector Machines in a real time tracking architecture"*, Rochester Institute of Technology RIT Scholar Works,2004.
7. Seemanthini K, Dr.Manjunath.S.S, " *Human Detection and Tracking using HOG for Action Recognition"*, International Conference on Computational Intelligence and Data Science (ICCIDS 2018), Elsevier.

AUTHORS PROFILE

Bhavya R., is a Research Scholar at Dept of ECE, RV College of Engineering, Bangalore, INDIA. She has completed her Master of Technology in VLSI Design and Embedded Systems in 2010 from Visvesvaraya Technological University. She has served as an Assistant Professor for 5 years in Engineering Colleges. Her Research Interests includes Image Processing, Signal Processing and Video Processing.

K. S. Geetha, is Professor and Head of Department of Electronics and Communication Engineering, R V College of Engineering, Bangalore, India. She has received her B. E and MTech in Electronics Engineering from National Institute of Engineering, Mysore, India. Her research interests include Digital Signal Processing, Image and Video processing, Large Area Flexible Microelectronics. Publications which include various international journals and international conference proceedings.