International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-10 Issue-10, August 2021

Object Detection and Tracking using Multiple Features Extraction

Bhavya. R, Geetha K S

***Abstract:** In most of the video analysis applications, object detection and tracking play vital role. Most of detection and tracking algorithms fail to predict multiple objects with varying orientation. In this paper, the goal is to identify and track multiple objects using different feature extraction methods like Locality Sensitive Histogram, Histogram of Oriented Gradients and Edges. These features are subjected to train classifier that can detect the object of different orientations. Experimental results and performance evaluation depicts the proposed method which uses LSH performs well with an increased accuracy of 98%. This method can precisely track the object and can be utilized to track under different scale and pose variations.*

***Keywords:** Decision Tree, Edges, Histogram of oriented Gradients, Locality Sensitive Histogram, Support Vector machine.*

I. INTRODUCTION

Multiple Object Tracking is still an exigent approach as it is challenging due to long term occlusion, multiple targets, slow moving objects, etc. Feature extraction being an important step in object detection and tracking applications, can reduce the complexity and improvise robustness in various environments. In addition to handling specific problems associated with the appearance of the target object, different tracking models represent the target objects in various representation schemes. Few tracking algorithms operates on subspace model which is time consuming while other group of methods use image statistics with adaptive schemes to represent target object [1].

In this paper, three features like: Locality Sensitive Histogram, Histogram of Oriented Gradients, Edge features are extracted and subjected to classification. Locality sensitive histogram computes histogram by considering all the pixels in an image. It takes into account contributions from all the pixels by adding floating point value to each bin. This enables to efficiently track the target object with multiple overlapping regions [3]. Histogram of oriented gradients is a superior feature descriptor which computes the gradient magnitude and angle, performs orientation binning and finally normalizes the blocks. This reduces the computational complexity and is a well-recognized descriptor for tracking [2].

Manuscript received on August 16, 2021.
Revised Manuscript received on August 23, 2021.
Manuscript published on August 30, 2021.
* 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

Edges represent the variations in intensity at the borders of the object. The paper is organized as follows: Section II explains the proposed architecture, and different feature extraction methods. Section III explains tracking of objects using the features. Section IV explains the experimental results.

II. PROPOSED METHOD

In this paper, the proposed algorithm efficiently tracks the moving objects using the extracted features. The extracted features of the moving object is considered as a tracking model which is represented by rectangular box. Tracking is accomplished by using the trained classifier, which can distinguish the moving object from the background in target frame.

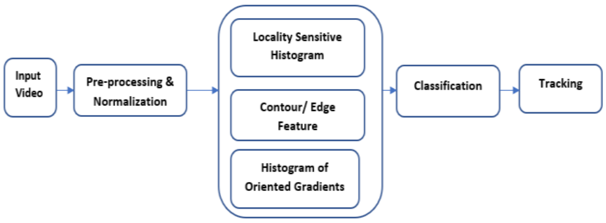

Fig 1: Proposed Model Architecture.

The features extracted for classification are Locality Sensitive Histogram, Histogram of Oriented Gradients, Edges along with R, G, B, for classification. Fig.1 represents the architecture of the proposed method. The proposed system consists of Preprocessing and Normalization, Feature Extraction, Object detection and tracking. Dataset collected contains vehicles of different size and different categories. All of the images vary itself with different illumination and appearances. All these images are subjected to maintain uniform aspect ratio.

Preprocessing and Normalization: Color palettes are used to describe the colors of the images. The primary colors used are Red, Green, and Blue. To process the images with these colors is difficult, as they are highly correlated. Hence the color images are further converted to grayscale that aids in analyzing the image and implement recognition algorithms. Feature extraction plays a vital role in the proposed system. The different features extracted are locality Sensitive Histogram, Histogram of Oriented Gradients, Contour/Edge extraction.

These features are explained in following sections in detail. The extracted features are subjected to classification and further the objects are tracked.

Retrieval Number: 100.1/ijitee.J943708101021
DOI: 10.35940/ijitee.J9437.08101021
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication

A. Locality Sensitive Histogram (LSH)

LSH, considers the contributions from all the pixels in an image rather than limiting itself to the neighborhood [3]. It adds a floating-point value to the bin, for each frequency of occurrences of intensity value. LSH is calculated at pixel p by:

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

where α ε (0,1) is controls the decreasing weight as pixel moves from center.

LSH for a 1D image is obtained from:

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

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

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

The above equations compute LSH on right and left side of pixel **p** which contributes to right side LSH and left side LSH respectively. Summing up the left and right LSH, results in total LSH which considers contributions from all the pixels, thereby dropping the pixel contribution exponentially that is far away from the center. This reduces the computational complexity of LSH.

B. Histogram of Oriented Gradients (HOG)

Histogram of oriented gradients is the most powerful feature vector/descriptor and is an improvement of SIFT descriptor that **calculated** gradient histogram using spatial normalization [5]. The basic idea is to characterize the object appearance and shape by local intensity gradient or edge directions, without defining the parameters. This is implemented by partitioning the image window into small cells, computing 1D histogram of gradient or edge directions for each cell. In order to handle illuminance variation, the histogram is combined and further normalized from all the cells in the block.

The gradients of the grayscale image I is obtained as follows:

$$I_x(r, c) = I(r, c+1) - I(r, c-1) \qquad (5)$$

$$I_y(r, c) = I(r-1, c) - I(r+1, c) \qquad (6)$$

The gradient with polar co-ordinates, angle is between 0 to 180 degrees.

$$\mu = \sqrt{I_x^2 + I_y^2} \qquad (7)$$

$$\theta = \frac{180}{\pi} (\tan_2^{-1}(I_y, I_x) \bmod \pi \qquad (8)$$

The combined histogram is characterized and normalized using Euclidean form as:

$$b \leftarrow \frac{b}{\sqrt{\|b\|^2 + \varepsilon}} \qquad (9)$$

The final HOG feature obtained by combining the features of the normalized blocks is:

$$h \leftarrow \frac{h}{\sqrt{\|h\|^2 + \varepsilon}}, h_n \leftarrow min(h_n, \tau) \qquad (10)$$

C. Edge Detection

Edge is a key element that detects the boundaries of the object and identify them. Edges in an image typically conveys the discontinuities in intensity, identifying object properties like shape, size, structure, reflection, intensity [4]. Using the thresholding approach, the first order derivative operator called Robert operator is used. This gradient operator considers the maximum and minimum intensity values around a pixel and then determine if the pixel is an edge or not. Gradient magnitude is calculated by:

$$|G| = \sqrt{Gx^2 + Gy^2} \approx |G_x| + |G_y| \qquad (11)$$

Angle orientation is computed by:

$$Robert's, \theta = arctan\left(\frac{G_y}{G_x}\right) \qquad (12)$$

These edges are used as feature vectors for object detection and tracking in the proposed model.

D. Decision Tree

Decision tree is a supervised learning algorithm which uses predictive modelling approach in data mining, statistics and machine learning applications. It is a tree like structure representing the flowchart which has root node, branches, and leaf nodes. It uses divide and conquer technique as a strategy. Root node is the topmost node in the decision tree. From the root node to a leaf node, path is traced which carries the prediction of that class for the tuple.

An input tuple X, considers the attribute values of the tuple and is tested against the decision tree. It holds the associated class labels with splitting criterion. Starting from the root node, splitting criteria is used to determine if the tuple belongs to a class. These decision trees can be converted into classification rules. The internal node of a decision tree contains an attribute test value, branch refers to the outcome of the test and class label is represented by terminal node. There are several decision trees methods which is faster in constructing the trees and provides much better accuracy. Considering this classifier, the datasets are trained with the feature vectors derived from the input images. The decision tree classifies with an accuracy of 96%. With this trained classifier, the input images are subjected to tracking as explained in the next section.

III. OBJECT TRACKING

In this paper, the objects are tracked using Sliding window method.

Retrieval Number: 100.1/ijitee.J943708101021
DOI: 10.35940/ijitee.J9437.08101021
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication

It is one of the feature-based tracking approaches which uses a rectangular window of fixed width and height. This detects the presence of object in the region of interest using the trained classifier. This classifier is trained with the feature vectors representing the object. This helps the window region to localize the object in the image. Multiple search windows of varying size is listed which can detect the object of interest. Multiple windows are obtained by defining scaling factors for each window, thereby resulting in overlapping regions. These overlapping regions are combined and false detections can be removed by using heatmap and thresholding approach. Scaling of different window sizes in repeated multiple times to generate multiple windows in the search method. This reduces the calculation time and thereby increasing the throughput rate. Multiple detections are resolved into one single detection with suitable threshold and detect the object to be tracked with a bounding box around it. Once the object is detected, pipeline creator function is used which loads the input video, with multiple parameters. This function processes all the frames in video, detects and tracks the object with the above-mentioned techniques and thereby creating a video with detected object. Pipeline function, with heatmap and multiple detections, the output video stream holds the detected object with bounding boxes around it. The experimental evaluation and results obtained after the implementation is tabulated and explained in the following section.

IV. EXPERIMENTAL RESULTS AND EVALUATION

In this section, the performance, efficiency and effectiveness are evaluated by considering few performance measuring parameters. The proposed detection and tracking algorithms are implemented using Machine Learning techniques and Python 3.7 on Intel i7CPU. The training dataset contains several images collected from different repositories. These collections contain images of vehicles belonging to different categories considering abrupt motion, illumination variation, different pose, scale with appearance and background variation. The number of frames in each video is approximately 1000frames. Dataset consists of 640x480 pixels. The results are evaluated both quantitatively and qualitatively by visualizing the results and considering few measuring performance metrics. Qualitative results are visualized in the video where the object is tracked efficiently subjected to challenges. The input dataset is preprocessed for better accuracy. All of these images with variable aspect ratio are normalized and downscaled to a size of 64x64 to maintain uniformity. Processed images are further subjected to feature extraction. In the proposed system, Locality sensitive histogram, HOG and Edge detection are the features considered. With these extracted feature vectors, the classifier is trained which performs well with an accuracy in range of 94% to 97%. This trained classifier localizes and detects the objects, thereby tracking the object in the given input video.

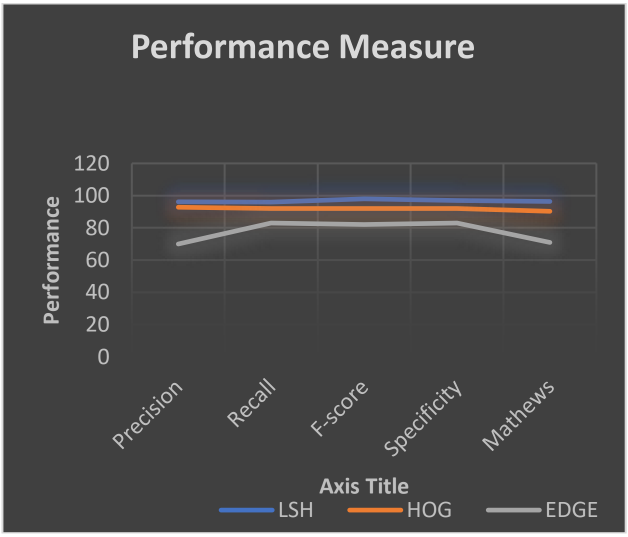

Fig 2: Performance plot of three feature extraction methods.

Table 1

Results of LSH									
Classifiers	**Precision**	**Recall**	**F-score**	**Specificity**	**PWC**	**Kappa score**	**FPR**	**FNR**	**Accuracy**
Decision Tree	0.97	0.969	0.984	0.969	1.265	0.2826	0.03	0	**98.73**
SVM	0.961	1	0.98	1	1.886	0.107	0	0.0357	**98.11**
Results of HOG									
Classifiers	**Precision**	**Recall**	**F-score**	**Specificity**	**PWC**	**Kappa score**	**FPR**	**FNR**	**Accuracy**
Decision Tree	0.86	0.78	0.85	0.787	11.392	0.282	0.212	0.043	88.6
SVM	0.928	0.92	0.92	0.92	3.773	0.641	0.071	0.025	96.22
Results of Edge									
Classifiers	**Precision**	**Recall**	**F-score**	**Specificity**	**PWC**	**Kappa score**	**FPR**	**FNR**	**Accuracy**
Decision Tree	0.8	0.92	0.85	0.92	6.25	0.74	0.07	0.05	93.75
SVM	0.7	1	0.82	1	11.62	0.61	0	0.16	88.37

Retrieval Number: 100.1/ijitee.J943708101021
DOI: 10.35940/ijitee.J9437.08101021
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication

Fig. 3: Tracking and Detection of Multiple vehicles using Locality Sensitive Histogram, Histogram of Oriented Gradients, and Edges.

The results of tracking the object in video is shown in Fig 3. The first column in Fig 3 represents the frames of the input video where the object of interest to be tracked is present. The second column represents the heatmap of the image which combines the detections of the object overlapping, which eliminates false positives. The third column indicates the object tracked using different features like LSH, HOG and Edge. This is identified by a blue colored bounding box identifying multiple category vehicles in the video frames. The performance measure of the classifiers are tabulated in the Table 1, 2, 3.

Retrieval Number: 100.1/ijitee.J943708101021
DOI: 10.35940/ijitee.J9437.08101021
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication

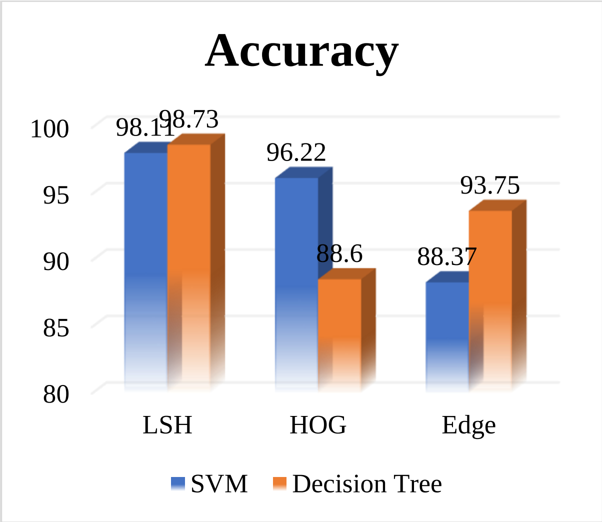

Fig 4: Accuracy of Different classifiers using different feature extraction methods.

Table 1 gives us the quantitative measure of the classifier using LSH as feature extraction. This classifier performs extensively well in tracking and detecting the object when compared to other feature extraction methods. LSH results with an accuracy of **98%** using both decision tree and Support vector machine. Feature extraction using Histogram of oriented gradients also detects and tracks the object with **88%** to **96%** accuracy. Edge detection and tracking using this method performs with a minimal accuracy of **88%** to **93%**. With this tabulation, it can infer that the classification using LSH performs well in tracking and detecting the object with increased accuracy. The graphical representation of the performance of these trackers is shown in Fig 2 and Fig 4. Locality Sensitive histogram performs well by effectively detecting the objects in the frame when compared to other feature extraction methods.

V. CONCLUSION

The proposed method detects and tracks the object using multiple features. Different feature extraction techniques are used to analyze the tracking and detection of vehicles. Using LSH as feature extraction method, tracking and detection results with a higher accuracy of **98.73%**. With this accuracy, the proposed method can detect multiple vehicles of different categories.

The performance parameters like Precision, Recall, F-score, Specificity, False positive rate is used to validate the results obtained using the proposed method. These validations clearly depict that the proposed method outperforms well when compared to other existing methods. In-order to increase the robustness of the tracker, input videos with different challenging factors like illumination variation can be considered for future work.

REFERENCES

1. T.Mahalingam, M. Subramaniom, "*A Robust Single and Multiple moving object detection , tracking and Classification*", Applied Computer and Informatics, 2018.
2. 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.
3. S. He, Rynson W.H. Lau, Q.Yang, J Wang, M Yang, "*Robust Object tracking via Locality Sensitive Histograms*", IEEE Transactions On Circuits and Systems For Video Technology, 2016.
4. N.M. Raikar, Dr. Megha P Arakeri,, "*Development of Vehicle Speed Estimation Technique using Image Processing*", International Journal of Engineering Research & Technology, ISSN : 2278-0181, Vol.9, Issue 09, September -2020.
5. Wei Zhou, Shengyu Gao, Ling Zhang and Xin Lou, *"Histogram of Oriented Gradients Feature Extraction from Raw Bayer Pattern Images"*, IEEE Transactions on Circuits and Systems, 2020.

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.

Retrieval Number: 100.1/ijitee.J943708101021
DOI: 10.35940/ijitee.J9437.08101021
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication