International Journal of Electronics Engineering (ISSN: 0973-7383)
Volume 10 • Issue 1 pp. 207-212 Jan 2018-June 2018 www.csjournals.com

A Real Time Approach of Vehicle Detection Based On Feature Extraction

[1]Sujinder Singh, [2]Karamjit Kaur
[1]Research Scholar, [2]Assistant Professor
Department of ECE, Punjabi University, Patiala, INDIA

Abstract: Vehicle detection and tracking plays a successful and critical part in the zone of traffic surveillance framework where proficient activity administration and security is the fundamental concern. Programmed acknowledgment of vehicle information has been broadly utilized as a part of the vehicle data framework and keen movement framework. It has gained more consideration of analysts from the most recent decade with the headway of computerized imaging innovation and computational limit. Programmed vehicle recognition frameworks are keys to street movement control these days; a few utilizations of these frameworks are activity reaction framework, activity flag controller, path takeoff cautioning framework, programmed vehicle mischance discovery and programmed movement thickness estimation. This research work, proposed a technique to detect vehicles. A realtime dataset is created and stored in video format to utilize in future. This technique utilizes sift-surf technology to extract features of multiple vehicles. Then optimization of features is performed to select best features and to perform recognition.

Keywords: Vehicle detection, traffic surveillance, sift, surf, features optimization, recognition.

I. INTRODUCTION

Visual traffic surveillance has pulled in critical enthusiasm for PC vision, in light of its colossal application prospect. Creating vehicle directions from video information is additionally an imperative application and can be utilized as a part of examining activity stream parameters for ATMS. Foreground object detection is the backbone of most video surveillance applications. Foreground object detection is mainly concerned with detecting objects of interest in an image sequence.

Video based surveillance systems ordinarily screen the conduct of the objective in the assigned region utilizing single or numerous cameras. These frameworks require the improvement of picture handling and PC vision calculations to recognize, find and track a moving target. The utilization of various cameras empowers the reconnaissance of more prominent regions and the utilization of the excess data (gave by the recordings from various perspectives) can help beat a portion of the known confinements of single camera frameworks, for example, scene impediments. These confinements are principally because of two kinds of issues. There are sure zones that can't be handled by the cameras especially in jumbled scenes and can be tackled by utilizing various cameras and sensors. The below average of issues is because of the constrained determination of cameras. An intriguing answer for these issues could be utilizing straightforward yet viable particular sensors with various modalities to take care of the particular issues of the vision frameworks. Thusly, vision would in any case give abnormal state data, and low-level sensors would guarantee higher precision.

In this research work a technique is proposed to detect vehicles. This technique uses sift and surf approaches. Scale-invariant feature transform (or SIFT) is a calculation in PC vision to recognize and depict and portray nearby highlights in pictures. Fascinating focuses on the question can be removed to give a "component depiction" of the protest. This depiction, extricated from a preparation picture, would then be able to be utilized to distinguish the protest when endeavoring to find the question in test picture containing numerous different items. Speeded up Robust Features (or SURF) is a fast, performance scaled and rotation-invariant interest point detector and descriptor.

II. LITERATURE SURVEY

Chu et al. [1] proposed a novel vehicle detection plot in view of multi-errand profound convolutional neural systems (CNNs) and area of-intrigue (RoI) voting. In the outline of CNN engineering, creators enhance the regulated data with subcategory, district cover, jumping box relapse, and classification of each preparation RoI as a multi-errand learning system. This outline permits the CNN model to share visual information among various vehicle traits at the same time, and along these lines, location power can be adequately progressed. Furthermore, most existing strategies consider every rous autonomously, disregarding the pieces of information from its neighboring RoIs.

Poggio et al. [2] shown a general system for challenge revelation associated with auto ID in front and back view. The structure derives a lot of its vitality from a depiction that portrays a challenge class in wording a dictionary of neighborhood, orchestrated, multiscale control differentiates between adjacent regions that are figured using a Haar wavelet change. They used an outline based learning approach that evidently derives a model of a challenge class by means of setting up a SVM (Support Vector Machine) classifier using a significant game plan of positive and negative cases.

Rajagopalan et al. [3] modeled the appointment of auto pictures by learning higher demand bits of knowledge (HOS). Getting ready data trial of vehicles are clustered and the accurate parameters identifying with each gathering are assessed. Grouping relies upon a HOS-based decision measure which is obtained by deciding a course of action augmentation for the multivariate probability thickness function similar to the Gaussian limit and the Hermite polynomial. Online establishment learning is performed and the HOS-based closeness of a testing picture and every one of the gatherings of auto scattering and establishment course is figured. By then it is organized into auto or establishment.

Kanade et al. [4] embraced a view-based system. They developed one individual marker for every one of the coarsely quantized viewpoints. By then, they used the histograms of some precisely picked wavelet features and their relative zones to show the auto and non-auto scattering expecting the histograms are quantifiably self-ruling. Vehicle area in raised pictures is tolerably constrained by the point of view and the assurance.

Zhao et al. [5] shown a system to recognize explorer cars in hoisted pictures along the road heading where automobiles appear as meager dissents. They started from mental tests to find basic features for human revelation of cars. In perspective of these recognitions, they picked the farthest point of the auto body, the cutoff of the front windshield, and the shadow as the features. They used a Bayesian framework to facilitate all features in conclusion use it to perceive automobiles. In case the vehicle ought to be perceived from a video stream, by then development signs can be utilized. In a static camera configuration, moving articles can be perceived by establishment subtraction, while for a moving camera a photo stabilizer is required first. In conditions of solidly moving things, a moving blob may not come close to one single inquiry. In this way, a more ordered examination like the technique in static picture auto area should be used.

Chunrui et al. [6] developed another division system for gathering of moving vehicles. They used fundamental association with get the desired match. The results showed up in the paper are for the level point of view of the vehicles and no quantitative results were given.

Gupte et al. [7] proposed a system for vehicle area and portrayal. The took after vehicles are requested into two groupings: automobiles and non-cars. The portrayal relies upon vehicle estimations and is completed at a greatly coarse granularity – it can simply isolate automobiles from non-cars. The fundamental idea of this paper is to figure the length and stature of a vehicle, as showed by which a vehicle is designated an auto or non-auto.

Avely et al. [8] used a practically identical approach where the vehicles are assembled in view of length using an uncalibrated camera. In any case, this system furthermore describes the vehicles into two coarse social occasions – short vehicles and long vehicles. Remembering the ultimate objective to achieve a superior level game plan of vehicles, they require a more present day system that can recognize the never-ending traits for each vehicle class considered.

Zhang et al. [9] In their work they used a PCA-based vehicle portrayal structure. They executed two game plan estimations – Eigen vehicle and PCA-SVM to portray vehicle objects into trucks, voyager cars, vans, and pick-ups. These two systems abuse the perceiving vitality of Principal Component Analysis (PCA) at different granularities with different learning instruments. Regardless of the way that the systems themselves are captivating, the results disregard to achieve high accuracy. The execution of such counts in like manner depends upon the precision of vehicle institutionalization. As showed up in their paper, such systems can't describe the vehicles enthusiastically.

Hsieh et al. [10] introduced another gathering computation in perspective of features as fundamental as "appraise" and another segment "linearity" to arrange vehicles. They have conveyed imperative results, yet the subject of recuperating the "linearity" incorporate into frontal view remains unanswered.

III. PROPOSED TECHNIQUE

This research work, proposed a technique to detect vehicles. A realtime dataset is created and stored in video format to utilize in future. This technique utilizes sift-surf technology to extract features of multiple vehicles. Then optimization of features is performed to select best features and to perform recognition.

The methodology of the proposed technique consists of following steps:

1. Input the user defined Image
2. Select the Number Plate Portion
3. View Image
4. Perform Preprocessing
5. Split the Image according to number of Lines.
6. Split the Character of Number Plate by using Bounding Box Technique
7. Extract the Features from Image
8. Train and Test the Feature of character extracted
9. Show the output Image

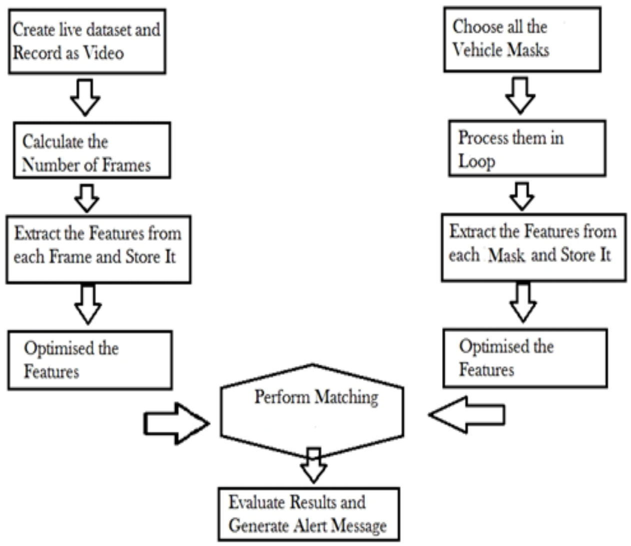

Figure 1: Proposed Technique

The methodology of the proposed technique consists of two stages. Both stages are processed in different manner and results obtained from both phases are then matched. First part consists of following steps:

- A live dataset is created and recorded as a video.
- In this video, numbers of frames are calculated.
- From each frame, features are extracted and stored for further use.
- These features are then optimized.

Second part consists of following steps:

- All vehicle masks are selected.
- These vehicle masks are then processed in loop.
- From each mask, features are extracted and stored for further use.
- These features are then optimized.

The optimized features from part one and part two of the algorithm are matched using minimum Euclidean distance. Then results are evaluated and an alert message is generated.

SIFT-SURF: Scale Invariant Feature Transform (SIFT) are invariant to rotation, scaling and partly illumination. Also invariant to 3D projective transform. Speed-up robust features (SURF) are a scale and in-plane rotation invariant feature. It contains interest point detector and descriptor. The detector locates the interest points in the image, and the descriptor describes the features of the interest points and constructs the feature vectors of the interest points. Like using SIFT feature in face recognition, SURF features should be extracted from images through SURF detectors and descriptors.

SURF Detector: The SURF detector concentrates on blob-like structures in the picture. These structures can be found at corners of articles, yet additionally at areas where the impression of light on specular surfaces is maximal (i.e. light dots). In 1998, Lindeberg saw that the Monge-Ampère administrator and Gaussian subsidiary channels could be utilized to find highlights. In particular, he identified blobs by convolving the source picture with the determinant of the Hessian (DoH) framework, which contains diverse 2-D Gaussian second request subordinates. This metric is then partitioned by the Gaussian's fluctuation, σ^2, to standardize its reaction:

$$\mathrm{DoH}\,(x,y,\sigma) = \frac{G_{xx}(x,y,\sigma).G_{yy}(x,y,\sigma) - G_{xy}(x,y,\sigma)^2}{\sigma^2}$$

where, $$G_{ij}(x,y,\sigma) = \frac{\partial N(0,\sigma)^2}{\partial i.\partial j} * \mathrm{image}(x,y)$$

The local maxima of this filter response occur in regions where both G_{xx} & G_{yy} are strongly positive, and where G_{xy} is strongly negative. In this manner, these extrema happen in districts in the picture with extensive force inclination varieties in various ways, and also at saddle focuses. Outwardly, this implies blob-like structures allude to corners and spots.

The other motivation behind why numerous component location plans depend on Gaussian channels is to dispose of boisterous information by obscuring the picture. As a side-effect, Gaussian obscuring features picture points of interest at or close to a solitary remarkable scale. As Marr wrote in the Theory of Edge Detection, "no single channel can be ideal all the while at all scales, so it takes after that one should look for a method for managing independently with the progressions happening at various scales." He goes ahead to propose to remove highlights utilizing the consolidated data of numerous reactions, created utilizing a similar group of channels however at various scales. This subsequent heap of convolutions is commonly alluded to as the scale space.

SURF's creators and numerous others have discovered that the quantity of results in scale-space diminishes exponentially with expanding scale measure. One potential clarification is that a lot of Gaussian obscuring can normal out almost all valuable data in pictures. Thusly, looking through the scale space directly can be exceptionally inefficient computationally. As an option, SURF presents the idea of scale octaves: every octave straightly tests a little locale of the scale space, however this area estimate is relative to the scale extents utilized. Each adjoining octave copies the district size and inspecting augmentation utilized as a part of the past one, to decrease the measure of inquiry essential at bigger scales.

The main negative side-effect of utilizing octaves is that outcomes found at bigger scales can conceivably have more mistake, because of the bigger augmentations utilized as a part of their relating octaves. To adjust, the SURF locator inserts the directions of any neighborhood maxima found into the sub-pixel and sub-scale run.

Finally, SURF's creators propose to make the qualification between splendid blobs found on a dim foundation, versus dull blobs found on a brilliant foundation. This property can be spoken to by the indication of the Laplacian, as demonstrated as follows:

$$\mathrm{sgn}\{G_{xx}(x, y, \sigma) + G_{yy}(x, y, \sigma)\}$$

The SURF detector algorithm can thus be summarized by the following steps:

1. Form the scale-space response by convolving the source image using DoH filters with different σ
2. Search for local maxima across neighbouring pixels and adjacent scales within different octaves
3. Interpolate the location of each local maxima found
4. For each point of interest, return x, y, σ, the DoH magnitude, and the Laplacian's sign

SURF Descriptor: To portray each element, SURF condenses the pixel data inside a nearby neighborhood. The initial step is deciding an introduction for each component, by convolving pixels in its neighborhood with the even and the vertical Haar wavelet channels. Appeared in Figure 4.2, these channels can be thought of as piece based strategies to process directional subordinates of the picture's force. By utilizing force changes to portray introduction, this descriptor can depict includes in a similar way paying little mind to the particular introduction of items or of the camera. This rotational invariance property permits SURF highlights to precisely recognize questions inside pictures taken from alternate points of view.

IV. EXPERIMENTAL RESULTS

For implementation of the proposed technique, MATLAB tool is used. The performance parameters for the evaluation of technique can be measured in terms of precision and recall.

Precision: Precision refers to the closeness of two or more measurements to each other. The precision of a measurement is a measure of the reproducibility of a set of measurements. It is the fraction of relevant instances among the retrieved instances. The formula to calculate precision is as follows:

$$\mathbf{Precision} = \frac{\mathbf{TP}}{\mathbf{TP + FP}}$$

where, TP is true positive, FP is False Positive.

Recall: Recall is the fraction of the relevant documents that are successfully retrieved. The formula to calculate recall is as follows:

$$\mathbf{Recall} = \frac{\mathbf{TP}}{\mathbf{TP + FN}}$$

where, TP is true positive, FN is False Negative.

Table 1: Showing comparison of the proposed technique with the existing technique based on precision and accuracy

Parameter	Existing Technique	Proposed Technique
Precision	0.97	0.98
Recall	0.78	0.95

The table above shows the comparison of the proposed technique with the existing technique on the basis of precision and recall. The precision of the existing technique is 0.97 whereas the precision of the proposed technique is 0.98. Recall of the existing technique is 0.78 whereas the recall of the proposed technique is 0.95. The precision and recall of the proposed technique is better than the existing technique.

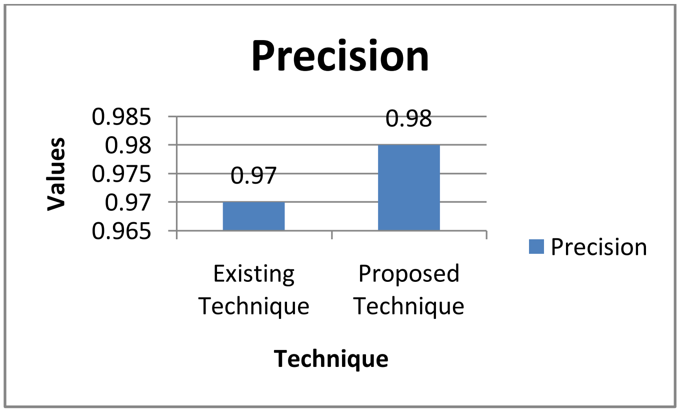

Figure 2: Showing comparison of the proposed technique with the existing technique based on precision

The figure above shows the comparison of the proposed technique with the existing technique on the basis of precision. The precision of the proposed technique is 0.98 whereas precision of the existing technique is 0.97.

International Journal of Electronics Engineering (ISSN: 0973-7383)
Volume 10 • Issue 1 pp. 207-212 Jan 2018-June 2018 www.csjournals.com

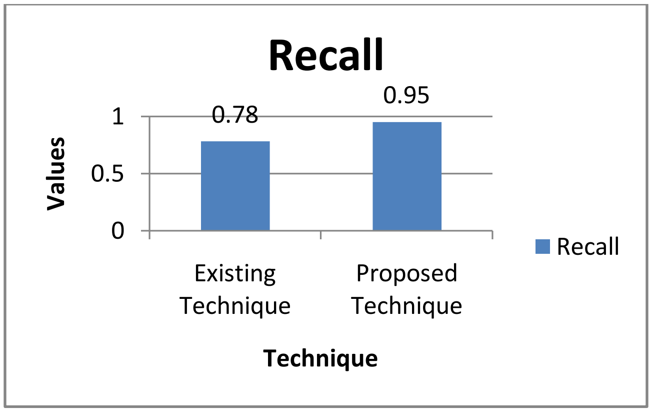

Figure 3: Showing comparison of the proposed technique with the existing technique based on recall

The figure above shows the comparison of the proposed technique with the existing technique on the basis of recall. The recall of the proposed technique is 0.95 whereas recall of the existing technique is 0.78.

Table 2: Showing accuracy comparison of proposed technique with existing techniques

Method	EASY	MODERATE	HARD
baseline	87.33	86.67	76.78
overlap prediction	88.22	87.23	77.42
RV	88.37	87.38	77.47
MLL	90.97	89.17	77.94
MTL	90.52	89	78.78
MTL and subNMS	90.66	89.15	79.27
ALL	90.88	90.66	84.33
Proposed Method	96	96	96

Above table shows the comparison of proposed technique with existing techniques on the basis of accuracy. Proposed technique gives highest accuracy in all cases than the existing technique.

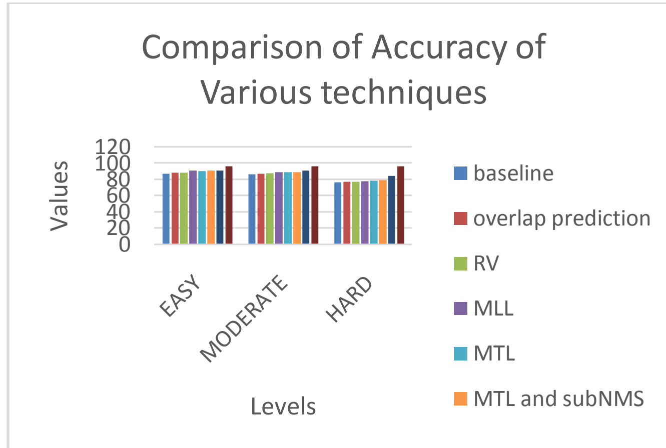

Figure 4: Showing comparison of the proposed technique with existing technique on the basis of accuracy

Above figure shows comparison of accuracy in proposed and existing techniques. It is clear from the graph that proposed technique gives highest accuracy as compared to existing techniques and hence performs better.

V. CONCLUSION

Vehicle recognition is an innovation which its point is to find and demonstrate the vehicle measure in advanced images. Detecting and following vehicles from transportation observation recordings is fundamental for applications going from activity line identification, volume figuring to episode and vehicle distinguishing proof. In this research work, a technique to detect vehicles is proposed. A real time dataset is created and stored in video format to utilize in future. This technique utilizes sift-surf technology to extract features of multiple vehicles. Then optimization of features is performed to select best features and to perform recognition. The technique is implemented using MATLAB tool. The performance parameters for the evaluation of technique can be measured in terms of precision and recall. The precision of the proposed technique is 0.98 whereas precision of the existing technique is 0.97. The recall of the proposed technique is 0.95 whereas recall of the existing technique is 0.78. Experimental results demonstrate that the proposed technique outperforms the existing technique in terms of precision and recall. In future, to further enhance the results, we may use artificial intelligence techniques, also vehicle detection is implemented as police department security system and we may use live cctv cameras instead of video footage.

REFERENCES

[1] Wenqing Chu , Yao Liu, Chen Shen, Deng Cai, and Xian-Sheng Hua, “Multi-Task Vehicle Detection With Region-of-Interest Voting”, IEEE Transactions on Image Processing, 27(1): 432-441,January 2018.

[2] C. Papageorgiou, and T. Poggio, “A Trainable System for Object Detection,” in International Journal of Computer Vision, 38(1):15-33, 2000.

[3] A. Rajagopalan, P. Burlina and R. Chellappa, “Higher Order Statistical Learning for Vehicle Detection in Images,” in Proceedings of IEEE International Conference on Computer Vision, 2:1204-1209, 1999.

[4] H. Schneiderman and T. Kanade, "A Statistical Method for 3D Object Detection Applied to Faces and Cars," in Proceedings of International Conference on Computer Vision and Pattern Recognition, 1:746-751, 2000.

[5] T. Zhao and R. Nevatia, "Car detection in low resolution aerial images," in Proceedings of International Conference on Image Processing, 710-717, 2001.

[6] Z. Chunrui and M.Y. Siyal, "A new segmentation technique for classification of moving vehicles," in Proceedings of Vehicular Technology Conference, 1:323-326, 2000.

[7] S. Gupte, O. Masoud, R. F. K. Martin, and N. P. Papanikolopoulos, “Detection and Classification of Vehicles,” in IEEE Transactions on Intelligent Transportation Systems, 3(1):37-47, 2002.

[8] R. P. Avely, Y. Wang, and G. S. Rutherford, "Length-Based Vehicle Classification Using Images from Uncalibrated Video Cameras," in Proceedings of lntelllgent Transportation Systems Conference, 2004.

[9] C. Zhang, X. Chen, W. Chen, "A PCA-based Vehicle Classification Framework," in Proceedings of International Conference on Data Engineering Workshops, 17-17, 2006.

[10] J. Hsieh, S. Yu, Y. Chen, and W. Hu, "Automatic Traffic Surveillance System for Vehicle Tracking and Classification," in Transactions on Intelligent Transportation Systems, 7(2):175-187, 2006.