JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Probability Based Approach for the Object Tracking

Anju Chaurasiya[1], Mr Wasif Khan[2]
Research Scholar[1],Assistant Professor[2]
Department of Computer Science and Engineering, SR Institute of Management and Technology, Dr. A.P.J. Kalam Technical University, Lucknow, Uttar Pradesh, India.
anjuchaurasiya.222@gmail.com, wasif.khan0105@gmail.com

Abstract

The techniques of object tracking concentrates on tracing the location of an object according to the probability distribution. The probability distribution estimates the location of objects on the based on their location distribution. In the proposed work, an improvement to the probability distribution algorithm is presented with the aim to track a multiple object. The methodology presented applies morphological segmentation and scan the entire object to classify the number of objects in each frame. The probability distribution method is implemented to forecast the location of the object in the subsequent frame. To collect the location of a large number of objects, the mean shift approach is implemented which will collect the information of many pixels. The simulation is done in MATLAB and it is analyzed the proposed algorithm yields better performance in comparison to the standard algorithmic approaches.

Keywords

Object, Detection, Tracking, Probability, Morphological.

Introduction

The number of online image-sharing websites and social networks is ever increasing. The machine learning and statistics dominate the realm of vision research in current times. The images and video are used for the detection, classification, and tracking of objects or episodes in order to "understand" the scenario at the ground level [1]. The field of computer vision is concerned with computer programming and designing algorithms to understand disclose the facts hidden in these images. Computer vision renders strength to a wide range of applications such as image search, robotic navigation, medical image analysis, image handling among others. In computer terms, an image is a scene that contains the objects of interest and the background characterized by all that other stuff in the photograph. The relationships and communications between these objects are main elements in visual comprehension. There are two main tasks in the computer vision: object detection and object recognition. Object detection defines the existence and/or radius of an object, and location in the photograph. Object recognition deals with identifying the object class in the database to which the object belongs. Object detection usually leads the way to object recognition. It can be thought of as dual-class object recognition, where one class is the representative of the object class while the other is of the non-object class. There are two more categories of object detection: soft detection and hard detection. Soft detection notices only the existence of an object, and hard detection detects both the presence and place of an object. Object detection fields are typically performed by determining each portion of a photograph to localize those parts whose photometric or geometric features correspond to those of the target object in the training database. This could be achieved by scanning an object template into a photograph at various locations, scales, and rotations, and an identity is declared if high-level match is found between the template and the photograph [2]. The resemblance between a template and an image area can be measured with their correlation (SSD). Therefore, the object detector aims to locate all object samples of one or more specified object classes, irrespective of scale, location, pose, position of view, partial obstruction, and lighting with respect to the camera. In the majority of computer vision systems, object detection is the first function performed because it permits more information to be obtained about the detected object and scene. Object detection has been employed in multifarious applications, the most commonly known of which are: (i) human–computer interaction (HCI), (ii) robotics (for example, service robots), (iii) consumer electronics (for example, smart-phones), (iv) security (for example, recognition, tracking), (v) retrieval (for example, search engines [3], photograph handling), and (vi) transportation (for example, automatic and supported driving).The requirements of each of these applications are different in terms of processing time (off-line, on-line, or real-time), robustness to obstructions, irreversibility to turns (for example, in-plane rotation), and detecting under pose variations. While most of the applications consider detecting a single object class (for example, face) and from a single scene (for example, frontal face), others need to detect numerous object classes (for example, vehicle, etc.), or a single class from many angles (for example, side and front view of vehicles). Generally speaking, the majority of systems can infer only one object class from a limited set of scenes and poses. Object detection aims at determining the location of objects in a particular photograph (object localization), and the class (object classification) to which each object belongs. Therefore, the pipeline of the classic object detection frameworks can be typically partitioned into three phases: informational region selection, feature extraction and classification as shown in figure 1.

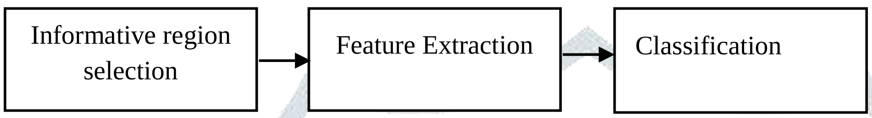

Figure 1: Traditional Object Detection Framework

All steps in the figure are explained below:

i. Informative region selection: Since individual objects may present at any location in the image and have differing aspect ratios or dimensions, scanning the entire image with a multi-scale sliding window is an organic selection. Although this detailed strategy can detect all feasible locations of objects, it also has apparent drawbacks [4]. The presence of multiple candidate windows increases its overhead in terms of computation and generates a lot of unnecessary windows. Nevertheless, the implementation of only a certain number of sliding window templates may generate inacceptable areas.

ii. Feature extraction: In order to recognize a variety of objects, it is required to retrieve visual properties that can provide a meaningful and robust representation. Features such as SIFT, HOG and HAR are representative. The reason behind this that these properties have potential to generate representations involving complex cells in the human brain [5]. Nevertheless, owing to the broad spectrum of scenes, lighting conditions and backgrounds, designing a powerful feature descriptor physically to fully determine all types of objects is not so easy.

iii. Classification: Furthermore, there is the need of a classification system to draw the line of distinction between the target object and all other classes, and to make the representation more categorical, meaningful and informative for visual recognition. Usually, Supported Vector Machines (SVMs), Ada Boost and Deformable Part-Based Models (DPMs) are selected on a priority basis. Among these classification frameworks, DPM is an adaptable model uniting object parts with deformation costs to manage simple deformations. In DPM, carefully designed low-level features and kinematic ally driven part decomposition are combined with the aid of a graphical model. And the discriminative learning of graphical models enables the construction of high- part-specific models with good precision for different object categories.

Literature Review

Ajmal Shahbaz, et.al (2020) suggested an algorithm of detecting a moving object for a moving camera [6]. A deep at rousspatial CNN (Convolutional Neural Network) was implemented in order to extract the attributes. A bilinear interpolated up-sampling network was adopted to extend the extracted feature which illustrated a moving object. A dilated convolution kernel was assisted in extracting the deep dilated features. The change detection dataset was applied to execute the top-ranked algorithms in the field of detecting a moving object. The suggested algorithm was capable of segmenting the foreground object at higher F-measure. The future work would emphasize on enhancing the algorithm on the basis of multi-scale feature and different up sampling methods.

Chinthaka Premachandra, et.al (2019) established a novel technique to detect the moving objects at road intersections with the help of 360-degree view camera THETA [7]. The installation of camera was done at a corner of the intersection. The images obtained from single THETA were useful for capturing the all directions of a road intersection. Moreover, a novel technique was put forward for detecting the moving objects from THETA images. The results of real time experimentation demonstrated the applicability of the established technique for detecting the moving objects such as pedestrians, bicycles and wheelchairs in successful manner at lower FPR (false positives rate).

Rahul, et.al (2018) introduced a system in which DL (deep learning) was integrated with stereovision to detect, tag and estimate the distance of objects in advance [8]. The objects were detected and recognized in the field of vision of the stereo camera using CNN (Convolutional Neural Network). Thereafter, the initial constructed 3D point cloud was employed to estimate the distance of those objects from the camera via triangulation technique. NVidia® Jetson TX1 was exploited with a Zed® stereo camera to deploy the introduced system. The results depicted that the introduced system yielded the accuracy of 84% with the average error 0.047 below.

Yizhou Wang, et.al (2021) developed a deep RODNet (radar object detection network) in which camera-radar fused algorithm was implemented to perform its cross-supervision so that the objects were detected from the RF (radio frequency) images in real-time [9]. Initially, this algorithm focused on converting the raw signals, obtained from milli meter-wave radars, into RF images in range-azimuth coordinates. Subsequently, a snippet of RF images was fed in this algorithm for input with the objective of predicting the likelihood of objects in the radar FoV (field of view). A dataset called CRUW was generated to evaluate the developed algorithm. The developed algorithm had offered precision of 86% and recall of 88% while detecting the objects.

Xiangmo Zhao, et.al (2020) projected a new technique for detecting and recognizing an object in which the complementary information captured from 2 kinds of sensors was fused [10]. At first, the object-region proposals were generated accurately using 3D (three-dimensional) LiDAR data. Afterward, this technique aimed to map these candidates onto the image space to select RoI (regions of interest) of the proposals and employ them as input in CNN (Convolutional Neural Network) to recognize the object further. The sizes of all the objects were recognized in accurate way after amalgamating the attributes of the last 3 layers of the CNN for extracting multi-scale attributes from the ROIs. The results of KITTI dataset depicted that the projected technique provided the recall of 95% and accuracy of 89.04% for recognizing cars and 78.18% for recognizing pedestrians.

Qian Yi, et.al (2018) presented a system on the basis of FPGA (Field－Programmable Gate Array) for detecting the moving objects [11]. First of all, the video images were captured usingD5M camera. After that, the image was cached into SDRAM memory. GMM based background subtraction technique was deployed for detecting the buffered image. The image display module utilized the processed image for exhibition. The hardware design was accomplished and the presented system was simulated on Matlab, Xlinx ISE and Modelsim. The experimental outcomes validated that the presented system was adaptable for detecting the moving targets in real time at higher accuracy.

Jianying Yuan, et.al (2021) designed a moving object detection technique on the basis of a vehicle-mounted binocular camera [12]. A method, of selecting feature matching points, was planned to acquire higher accuracy while estimating the camera ego-motion. Thereafter, the camera ego-motion and global mixed optical flow were considered to estimate the residual optical flow filed taken from the moving objects. In the end, every single moving object was extracted from the moving areas by integrating 3D (three-dimensional) and 2D (two-dimensional) information. The designed technique was capable of detecting any kind and any size of moving object theoretically based on the dense motion information related to the images. The results on KITTI database indicated that the designed technique offered 91% precision for detecting the moving objects.

Research Methodology

In this technique multi object tracking first gets the single frame, then at the same time they measure the point from which the image will pitch. After that, the overlaying image is constructed from that the image to find the object to be drawn. All process covered under image pre-processing should be known or object will be recognized. After that target identify the target in which assign different colour is assigned to each object along with different shape of rectangle to each object which are covered under extract shape. Then from database, the objects finalized or processing is performed using the colour feature or shape feature. If the target is found, the object driven is detected. It means that the object will move around so their detection will also move with them. After wards, then position is tracked with some end effect or in which they we will trace a line from the end with the same colour that is provided to them, after that their total time is calculated on the basis of fault rate detection and failure rate.

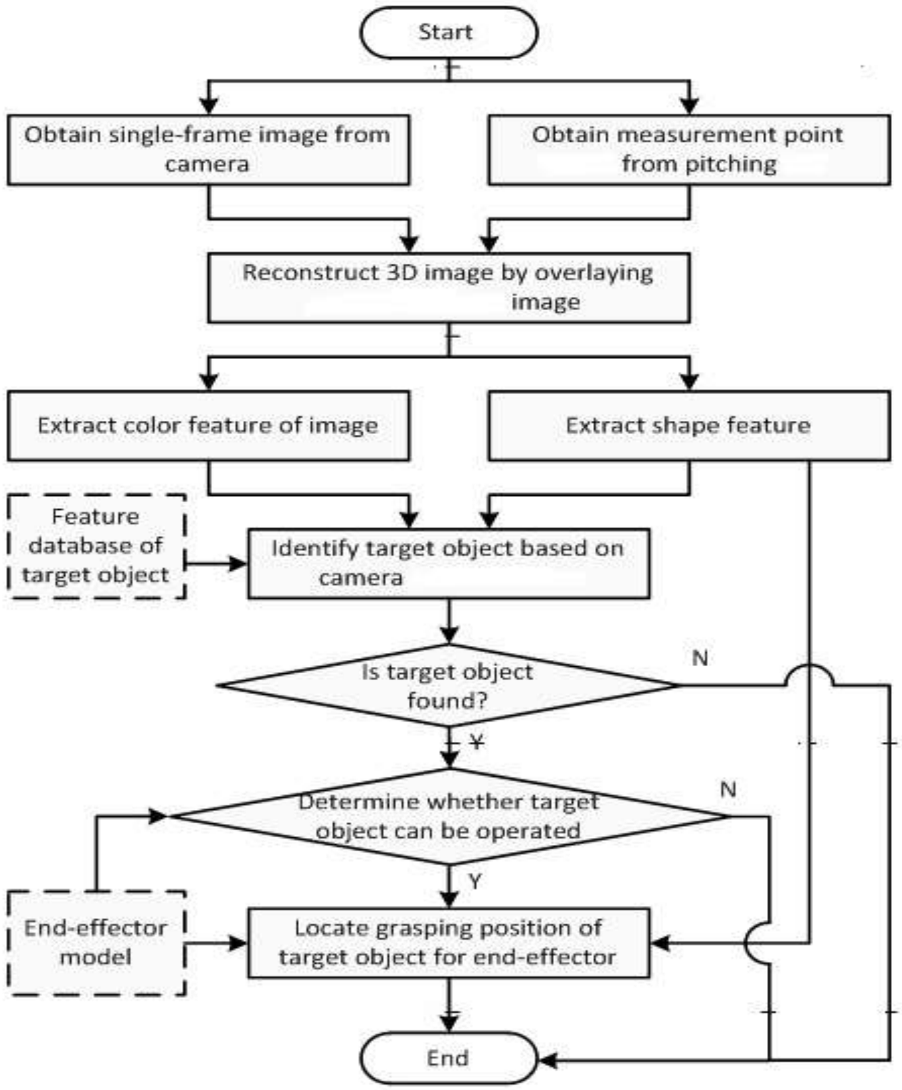

Figure 2: Proposed Methodology

Result and Discussion

MATLAB is considered as a Matrix Laboratory. It is an interactive program to conduct numerical computation and visualize the data. The programming capabilities of this language generate a tool which is adaptable for all areas of science and engineering. A complete set of reference-standard algorithms, functions, and apps for image processing, analysis, visualization, and algorithm development is obtained from the Image Processing Toolbox. It assists in enhancing an image, de-blur an image, extract the feature, mitigate the noise, segment an image, geometric transformations, and register an image.

In this, a technique has been introduced for tracking multiple objects. The existing technique utilized probability technique, which only worked on tracking single object from the video. The enhanced probability based algorithm will able to track multiple objects and performance of proposed algorithm is higher as compared to the existing algorithm concerning time, fault detection rate and accuracy rate. The performance of existing and proposed algorithm has been represented graphically

Parameter	Open CV Tracker	Probability Tracker	Enhanced Probability Tracker
TME	15	21	28
CT	0.12	0.34	0.53
TC	2.22	2.7	3.1
CDR	0.307	0.308	0.3091
FAT	0.90	0.30	0.45
TF	34.34	17.19	7.27

Table 1: Comparison of algorithms

The table 1 denoted the performance analysis of three algorithms and these algorithms are Open CV tracker algorithm, Probability based Tracker and enhanced probability tracker is implemented. The performance of these mentioned algorithms is evaluated with regard to TME, CT, TC, CDR, FAT and TF parameters. This indicated that higher performance of enhanced probability algorithm than other algorithms.

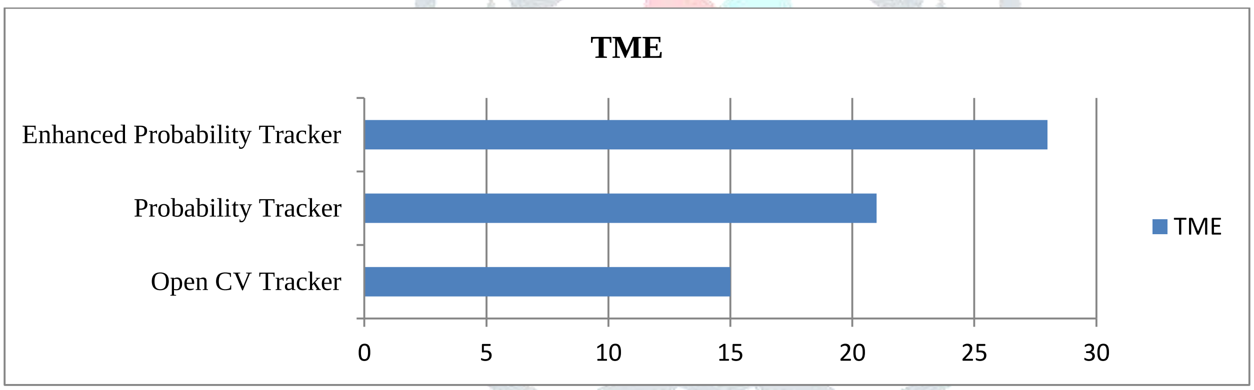

Figure 3: TME Comparison

The figure 3 represented that the open CV tracker, Probability tracker algorithm and enhanced tracker algorithm are compared. This depicted that TME value of object tracking algorithm is higher in comparison to the other algorithms

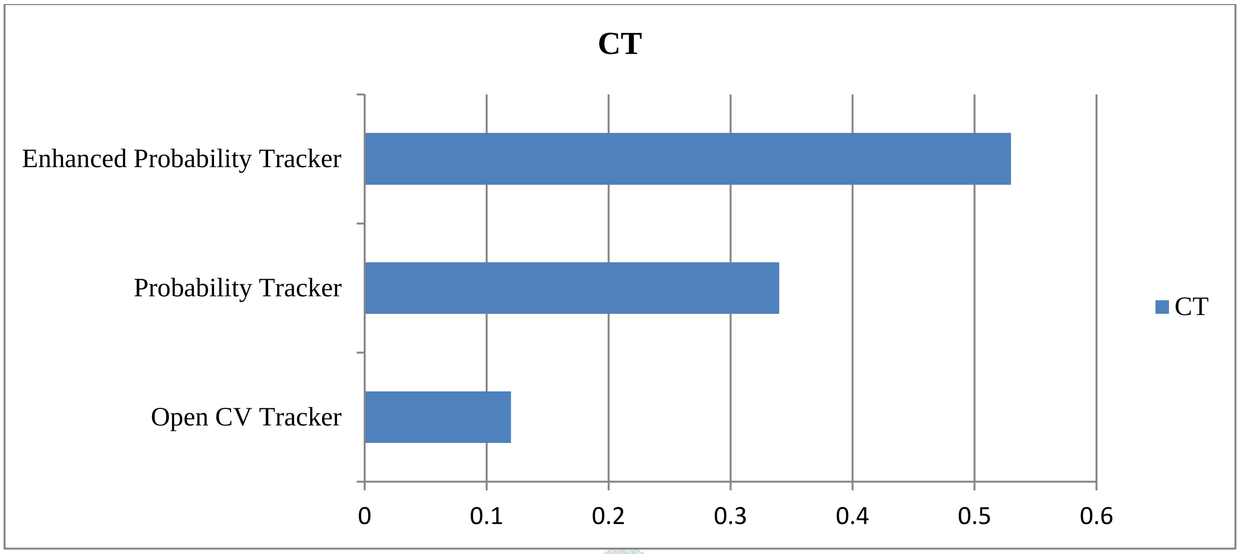

Figure 3: CT comparison

The figure 3 denoted that the open CV tracker, Probability tracker algorithm and enhanced tracker algorithm are compared. This depicted that CT value of object tracking algorithm is higher in comparison to the other algorithms

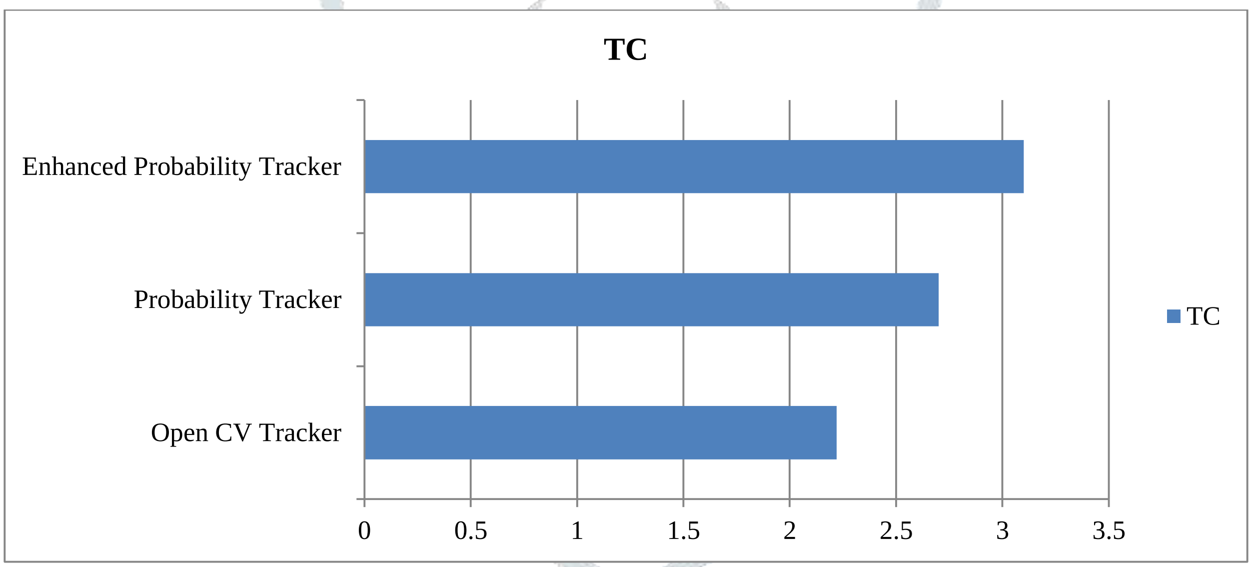

Figure 4: TC comparison

The figure 4 defined that the open CV tracker, Probability tracker algorithm and enhanced tracker algorithm are compared. This depicted that TC value is higher that implies the supremacy of object tracking algorithm over other algorithms

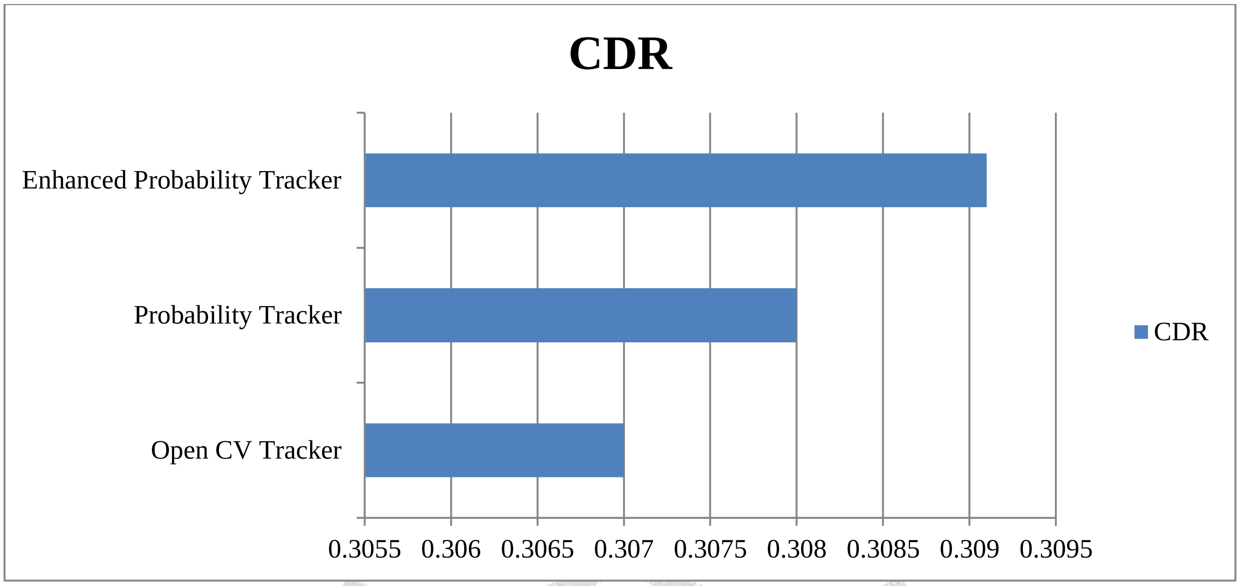

Figure 5: CDR comparison

The figure 5 described that the open CV tracker, Probability tracker algorithm and enhanced tracker algorithm are compared. This depicted that CDR value is higher that implies the supremacy of object tracking algorithm over other algorithms

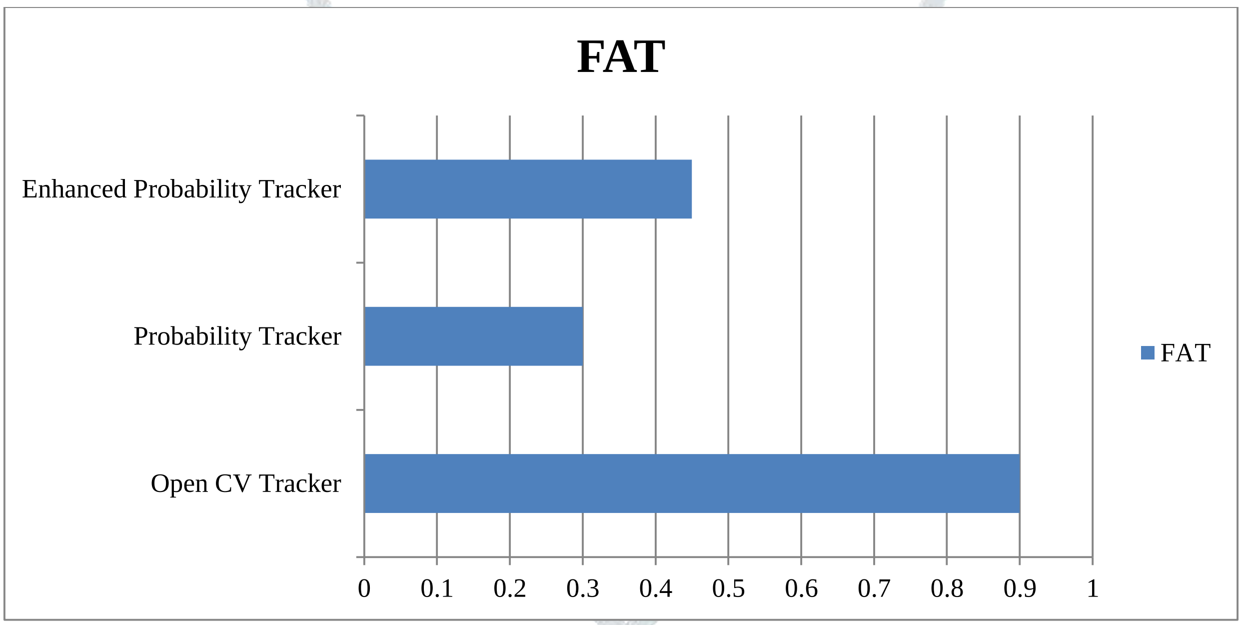

Figure 6: FAT comparison

The figure 6 illustrated that the open CV tracker, Probability tracker algorithm and enhanced tracker algorithm are compared. This depicted that FAT value is lower that implies the supremacy of object tracking algorithm over other algorithms

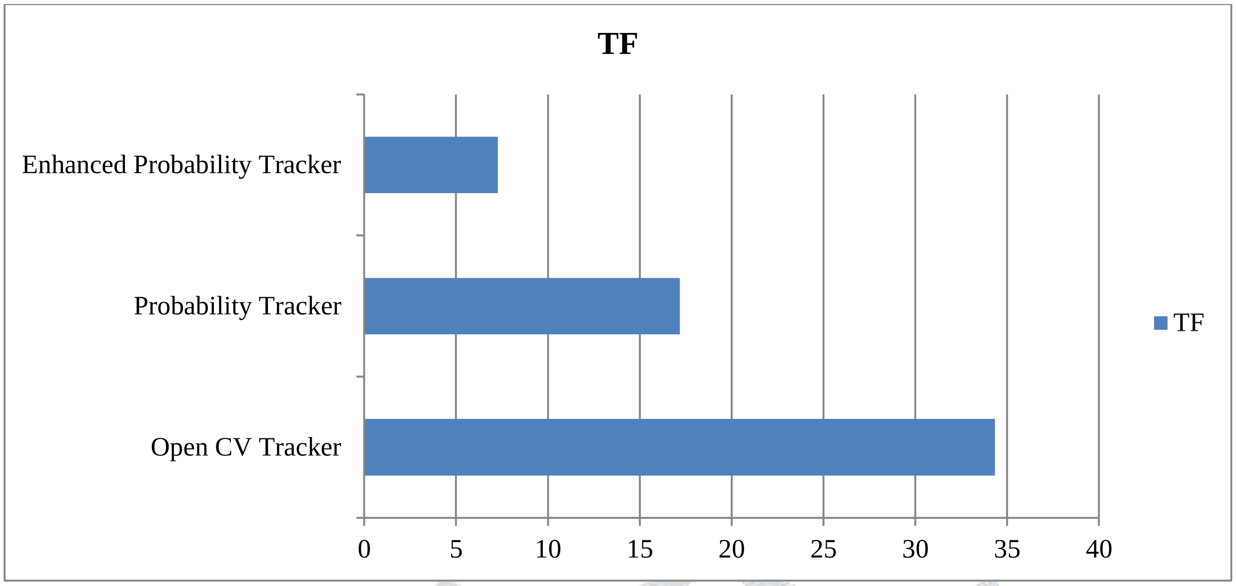

Figure 7: TF comparison

The figure 7 represented that the open CV tracker, Probability tracker algorithm and enhanced tracker algorithm are compared. This depicted that TF value is lower which means that it is effective object tracking algorithm in contrast to other algorithms.

Conclusion

The object tracking is a process in which spatial and temporal changes are monitored in sequence of images, such as the presence of object, its position, size, shape, etc. In the proposed improvement, method of filtering is adopted to eliminate the noise from the frames of videos. Multiple objects are tracked using method of mean shift in which pixels which have similar location are put together. For the detection of multiple objects, probability distribution approach is implemented to detect the future location of the object having similar properties. The simulation results indicated that the introduced algorithm yielded 90% accuracy to track multiple objects from the video. The work future will focus on enhancing the introduced algorithm for tracking the object having no similar sizes.

References

[1] S. Mane and S. Mangale, "Moving Object Detection and Tracking Using Convolutional Neural Networks," 2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS), 2018, pp. 1809-1813

[2] H. -C. Shin, K. -I. Lee and C. -E. Lee, "Data Augmentation Method of Object Detection for Deep Learning in Maritime Image," 2020 IEEE International Conference on Big Data and Smart Computing (BigComp), 2020, pp. 463-466

[3] H. Anwar, K. Javed, S. Rubab and M. J. Khan, "Towards Efficient Aerial Object Detection: Techniques, Datasets and Challenges," 2021 International Conference on Communication Technologies (ComTech), 2021, pp. 61-66

[4] R. Chang et al., "Underwater object detection with efficient shadow-removal for side scan sonar images," OCEANS 2016 - Shanghai, 2016, pp. 1-5,

[5] M. Noman, V. Stankovic and A. Tawfik, "Object Detection Techniques: Overview and Performance Comparison," 2019 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT), 2019, pp. 1-5

[6] Ajmal Shahbaz, Kang-Hyun Jo, "Moving Object Detection based on Deep Atrous Spatial Features for Moving Camera", 2020, IEEE 29th International Symposium on Industrial Electronics (ISIE)

[7] ChinthakaPremachandra, S. Ueda, Y. Suzuki, “Road Intersection Moving Object Detection by 360-degree View Camera”, 2019, IEEE 16th International Conference on Networking, Sensing and Control (ICNSC)

[8] Rahul, Binoy B. Nair, “Camera-Based Object Detection, Identification and Distance Estimation”, 2018, 2nd International Conference on Micro-Electronics and Telecommunication Engineering (ICMETE)

[9] Yizhou Wang, Zhongyu Jiang, Yudong Li, Jenq-Neng Hwang, Guanbin Xing, Hui Liu, “RODNet: A Real-Time Radar Object Detection Network Cross-Supervised by Camera-Radar Fused Object 3D Localization”, 2021, IEEE Journal of Selected Topics in Signal Processing

[10] Xiangmo Zhao, Pengpeng Sun, Zhigang Xu, Haigen Min, Hongkai Yu, “Fusion of 3D LIDAR and Camera Data for Object Detection in Autonomous Vehicle Applications”, 2020, IEEE Sensors Journal

[11] Qian Yi, “Design of Moving Object Detection System Based on FPGA”, 2018, 10th International Conference on Communications, Circuits and Systems (ICCCAS)

[12] Jianying Yuan, Gexiang Zhang, Fengping Li, Jiajia Liu, Lin Xu, Sidong Wu, Tao Jiang, Dequan Guo, YuruiXie, “Independent Moving Object Detection Based on a Vehicle Mounted Binocular Camera”, 2021, IEEE Sensors Journal.

.