Google AI Open Images -Object Detection Track

Prof. Kishore M[1], Archana H[2], Bindushree S[3], Devika N[4], Hemanth Kumar[5]
[1, 2, 3] Asst.Professor, Dept of ECE
[4, 5] Dept of Electronics and Communication
[1, 2, 3, 4, 5] Atria Institute of Technology, Bangalore, Karnataka, India

Abstract- *Real time object detection and tracking are important and challenging tasks in many computer vision applications such as video surveillance, robot navigation and vehicle navigation. Object detection involves detecting the object in sequence of images. Every tracking mechanism requires object detection mechanism either in each frame. Object tracking is the process of locating an object or multiple objects using either static or dynamic camera. The availability of high-powered computers, high quality and inexpensive video cameras will increase need for automated video analysis. It has generated a great deal of interest in object detection and tracking algorithms. Even though high-powered computers are used for object detection and tracking algorithm, most of the object detection algorithms such as background subtraction, temporal difference, foreground extraction and simple differencing requires long time to detect object, requires more storage space and no robustness against illumination changes.*

Keywords- Open Images, artificial intelligence, Object detection, Computer vision, Open Image Dataset v4, Image processing, object recognition and segmentation.

I. INTRODUCTION

Object detection is an important, yet. It is a critical part in many applications such as image search, image auto-annotation and scene understanding; however, it is still an open problem due to the complexity of object classes and images.Object tracking has many practical applications (e.g. surveillance, HCI) and has long been studied in computer vision. Although there has been some success with building domain specific trackers (e.g. faces [6], humans [16]), tracking generic objects has remained very challenging Nowadays thousands of cameras in the world are collecting ahuge amount of data on a daily basis for surveillance. It becomes increasingly important to develop related methods that process these data automatically. Multi-camera object tracking plays an important role in video surveillance. The goal is not only to monitor pedestrians, but also to extract useful information for other tasks at semantically higher levels, such as personal behavior analysis. .Object Tracking is one of the most active research areas in computer vision, with numerous applications including augmented reality, surveillance, and object identification. Methods of extracting significant information from still images and videos captured in a constrained environment, are being studied for several decades to enhance images or to build automated systems.

II. LITERATURESURVEY

2.1 Author:Khushboo Khurana, Reetu Awasthi "Techniques for Object Recognition in Images and Multi-Object Detection "International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 4, April 2013

The modern world is enclosed with gigantic masses of digital visual information. Increase in the images has urged for the development of robust and efficient object recognition techniques. Most work reported in the literature focuses on competent techniques for object recognition and its applications. A single object can be easily detected in an image. Multiple objects in an image can be detected by using different object detectors simultaneously. The paper discusses various techniques for object recognition and a method for multiple object detection in an image. In particular useful would be methods that could automatically analyse the semantic contents of images or videos. The content of the image determines the significance in most of the potential uses. One important aspect of image content is the objects in the image. So, there is a need for object recognition techniques. Object recognition is an important task in image processing and computer vision. It is concerned with determining the identity of an object being observed in an image from a set of known tags. Humans can recognize any object in the real world easily without any efforts; on contrary machines by itself cannot recognize objects. Algorithmic descriptions of recognition task are implemented on machines; which is an intricate task. Thus, object recognition techniques need to be developed which are less complex and efficient.

In this paper, we have discussed various object detection techniques. The template matching technique requires large database of image templates for correct object recognition. Hence it must be used only when limited objects are to be detected. Global features and shape-based method can give better result and are efficient as compared to local

features. These techniques help in easy access of the images. They also find their application in fields such as biometric recognition, medical analysis, surveillance, etc. A method for multiple object detection is also presented

2.2 Author:Nidhi"Image Processing and Object Detection" International Journal of Applied Research 2015; 1(9): 396-399 With the advancement of modern technologies areas related to robotics and computer vision, real time image processing has become a major technology under consideration. So, I tried a novel approach for capturing images from the computer web cam in real time environment and process them as we are required. By using open source computer vision library (OpenCV for short), an image can be captured on the bases of its hue, saturation and colour value (HSV) range. The basic library functions for image handling and processing are used. Basic library functions are used for loading an image, creating windows to hold image at run time, saving images, and to differentiate images based on their colour values. I have also applied function to threshold the output image in order to decrease the distortion in it. While processing, the images are converted from their basic scheme Red, Green, and Blue (RGB) to a more suitable one that is HSV. The process of object detection analysis the input image and to determine the number, location, size, position of the objects. Object detection is the base for object tracking and object recognition, whose results directly affect the process and accuracy of object recognition. The common object detection method is: colour-based approach, detecting objects based on their colour values. The method is strong adaptability and robustness, however, the detection speed needs to be improved, because it requires test all possible windows by exhaustive search and has high computational complexity.So, detection of any object by the computer web cam is done on the basis on the various colour schemes values. [Figure 7] As the program was made just to detect the RED colour object, it does so very well and give in output a thresholder image stream on the monitor of the computer. The program can't detect any other colour object, as shown in above [figure 8]. The main problem I was facing while working with this project was as we developed program just for RED object detection, whenever we want to detect any object with different colour, we have to make changes, every time in the source code. The best solution for this problem that I came up with is why not to make our program enough flexible that we can make changes according to the colour of the object at runtime. So, I included the value bares for HUE, SATURATION, and value range on the top of the program at run time.

2.3 Author:Fare's Jalled& Cybernetics Ilia Voronkov, Moscow Institute of Physics & Technology, Department of Radio Engineering "Object Detection Using Image Processing" arXiv.1611.077 [cs.CV] 23 Nov 2016.

An Unmanned Ariel vehicle (UAV) has greater importance in the army for border security. The main objective of this article is to develop an OpenCV-Python code using Haar Cascade algorithm for object and face detection. Currently, UAVs are used for detecting and attacking the infiltrated ground targets. The main drawback for this type of UAVs is that sometimes the object is not properly detected, which thereby causes the object to hit the UAV. This project aims to avoid such unwanted collisions and damages of UAV. UAV is also used for surveillance that uses Voila-jones algorithm to detect and track humans. This algorithm uses cascade object detector function and vision. train function to train the algorithm. The main advantage of this code is the reduced processing time. The Python code was tested with the help of available database of video and image, the output was verified.Image processing-based UAV is not completely operational as it is there is a manual intervention of a camera and joy stick. It will reduce the man work time and complexity of the work. In some cases, UAVs use very costly laser sensors and multiple sensor integrated systems to detect the objects and people. This project will be useful in replacing the laser sensor and servile the location using cheaper systems. results showed that the main aim was to detect the objects and the output objects were detected from the real scene. The face detection program can be implemented to detect and follow people in case of surveillance and other domains. We introduced a tool to explain some of the success of object detection systems.

2.4 Author: Liming Wang1, Jianbo Shi2, Gang Song2, and I-fan Shen1,Fudan University & University of Pennsylvania, 3330 Walnut Street, Philadelphia."Object Detection Combining Recognition and Segmentation".

Here they develop an object detection method combining top-down recognition with bottom-up image segmentation. There are two main steps in this method: a hypothesis generation step and a verification step. In the top-down hypothesis generation step, we design an improved Shape Context feature, which is more robust to object deformation and background clutter. The improved Shape Context is used to generate a set of hypotheses of object locations and figure ground masks, which have high recall and low precision rate. In the verification step, we first compute a set of feasible segmentations that are consistent with top-downobjecthypotheses,thenweproposeaFalsePositivePruning(FPP)procedure to prune out false positives. We exploit the fact that false positive regions typically do not align with any feasible image segmentation. Experiments show that this

simple framework is capable of achieving both high recall and high precision with only a few positive training examples and that this method can be generalized to many object classes.
Current approaches ([1][2] [3][4][5] [6][7] [8] [9][10]) to object detection can be categorized by top-down, bottom-up or combination of the two. Top-down approaches ([11][2][12]) often include a training stage to obtain class-specific model features or to define object configurations. Hypotheses are found by matching models to the image features. Bottom-up approaches start from low-level or mid-level image features, i.e. edges or segments ([8][5][9] [10]). These methods build up hypotheses from such features, extend them by construction rules and then evaluate by certain cost functions. Thethirdcategoryofapproachescombiningtop-downandbottom-upmethodshave become prevalent because they take advantage of both aspects. Although top-down approach scans quickly drive attention to promising hypotheses, they are prone to produce many false positives when features are locally extracted and matched. Features within the same hypothesis may not be consistent with respect to low-level image segmentation. On the other hand, bottom-up approaches try to keep consistency in low level image segmentation, but usually need much more efforts in searching and grouping. Here they used method contains three major parts: code book building, top-down recognition using matching and voting, and hypothesis verification, as depicted in Fig.1. The object models are learned by building a codebook of local features. We extract improved SC as local image features and record the geometrical information together with object figure-ground masks. The improved SC is designed to be robust to shape variances and background clutters. For, our experiments show that only a few training examples suffice to encode local shape information of objects. In this paper, we developed an object detection method of combining top-down model-based recognition with bottom-up image segmentation. Our method not only detects object positions but also gives the figure-ground segmentation mask. We designed an improved Shape Context feature for recognition and proposed a novel FPP procedure to verify hypotheses.

2.5Authors: B. Babenko, M.-H. Yang, and S. Belongie, in Proc. IEEE Conf. Computer. Vision Pattern Recogn., Jun. 2009," , "Visual tracking with online multiple instance learning," Incremental Multiple principle component analysis"

The main objective of visual tracking is to deal with the streams of an image that changes according with the time sequence. For this purpose, a novel incremental MPCA is used for online learning streams over dynamic tensor concept.
Tensor is defined as geometric representation of objects and determines the linear relationship between the matrices in an image. After applying this process Bayes interference framework is applied. Different learning methods have been implemented for visual tracking. Here different algorithms are presented for modeling appearance variations called as Eigen tracking algorithm but it is not felicitous in real time utilization because of the constancy in assumptions of a subspace vectors. To avoid all these limitations an online learning mechanism is provided which serves to satisfy the devaluation of imperfections in an image.

These includeLoss of spatial information

- Flatting of matrices into vectors due to intensive computation
- Mainly Principal Component Analysis is used to recognize the areas of face patterns in an image. As name suggests Principle Component Analysis is used for feature extraction and in this process traditional 2DPCA and 2DLDA but not satisfied to superlative extent.2DPCA is abbreviated as two-dimensional principle component analyses and is developed for feature extraction of an image. In this process a covariance matrix is constructed and takes less time to process all the functions. Another method called concurrent subspace analysis is used in visual tracking and its function is used to process single tensor in different directions. But all these technique does not serve to maximum extent as a color object is viewed as 3 order tensor and this was implemented in Incremental Multiple Component Analysis (IMPCA). Figure below suggests the flow chart of IMPCA. Following are the steps of implementation of IMPCA

- First an image is given as a input to IMPCA tracker
- By the process of sampling the image is divided based on the possible candidate states
- Evaluation process takes place for the selected state using tensor subspace
- Best state is selected using Bayes interference and this process goes on until the tracker updates using IMPCA and goes to next frame.

Bayesian interference mainly depends on hidden state Markov model. Here, in this process hidden states are not visible but output is dependent on state in which the model passes. Bayes interference is determined as the states that occur before and after condition is satisfied to the odd variables

2.6Authors: Weihua Chen Lijun Cao Xiaotang Chen Kaiqi Huang"A novel solution for multi-camera object tracking" Conference Paper · January 2015: IEEE International Conference on Image Processing, ICIP 2014

Frag Track

To track an object in a video progression a novel algorithm is used called as Frag track. Here, an object is represented by multiple fragments or patches of an image. Analysed with traditional algorithms in Frag track each and every position in an object is determined and different scales are obtained by taking the histogram of each position of an object and comparing it with histogram of original object. Here, an application is developed based on the integral datagram structure and the region based on the histogram used to extract 2015 multiple regions in rectangular manner.
Many difficulties have been raised and can be overcome by using Frag track.

These difficulties are

1. Occurring of partial occlusions
2. Spatial distribution of pixel intensities
3. Computational cost

2.7Authors: F. Tang, S. Brennan, Q. Zhao, and H. Tao, in Proc."Co-tracking using semi- supervised support vector machines,"IEEE Int. Conf. Computer. Vision, 2007, pp. 1–8.

Semi Supervised Support Vector Machines

To track foreground and background object of an image a semi supervised support vector machine is used to track the object. Initially, a small number of labeled samples are used and treats a new sample as unlabeled data sequence co training frame work using semi supervised approach is used by combining different features of classifiers by the process of classifier weighing method based on multiple features. By doing this process the background and object can be easily separated. Tracking is an important task of finding the state of an object which can be position, scale, velocity etc. Many drawbacks have been occurred from the previous algorithms and this task has become a challenging thought.

Limitations

- Leads to difficult in object re-acquisition
- Object appearance changes has become very complicated
- Difficult in background clutter
- Hard to predict the state of an object
- Incomplete occlusions

2.8Author: Yi Wei1, Nenghui Song1, Lipeng Ke2, Ming-Ching Chang1, Siwei Lyu11University at Albany, State University of New York 2University of Chinese Academy of Sciences "Street object detection/tracking for ai city traffic analysis".

Smart transportation based on big data traffic analysis is an important component of smart city. With millions of ubiquitous street cameras and intelligent analytic algorithms, public transit systems of the next generation can be safer and smarter.In the AI City Track 1 contest on visual detection, we built a competitive street object detector for vehicle and person localization and classification. In the AI City Track 2 contest on transportation applications, we developed a traffic analysis framework based on vehicle tracking that can assist the surveillance and visualization of the traffic flow. Both developed methods demonstrated practical, and competitive performance when evaluated with state-of-art methods on real-world traffic videos provided in the challenge contest.In this paper, we describe a practical system for street object detection, tracking and traffic analysis. the visualization of our tracking algorithm on the UA-DETRAC and AIC1080 datasets. The tracking results are continuous on the UA-DETRAC dataset, while they are occasionally discontinuous on the AIC1080 dataset due to frequent occlusions and low discriminative appearances After vehicle tracking is performed, each vehicle is assigned an unique ID for its moving trajectories which can be used for vehicle counting. In addition to the visualization of tracking on the input video, we further generate a topdown aerial view of the traffic flow. Specifically, we perform a manual site calibration by calculating a camera projection matrix, which establish a mapping between pixels and the physical world.

The proposed approach provides a solution to smart transportation, street surveillance, traffic safety, and can ultimately lead to a smarter city. In the future, we will continue to improve the capability and robustness of vehicle detection, classification, and tracking against real-world scenarios and applications.

III. PROPOSED BLOCK DIAGRAM

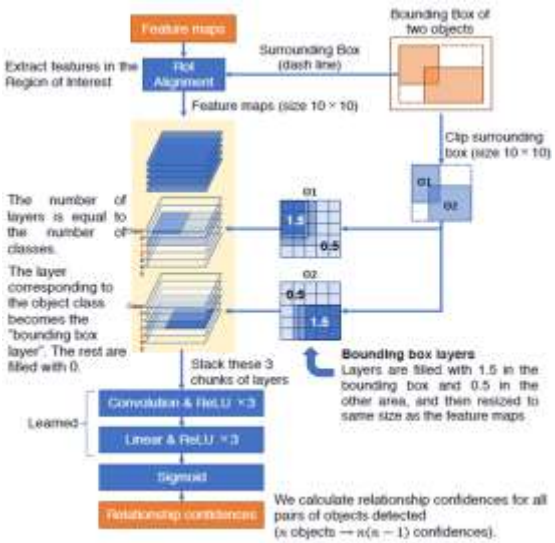

Fig. 1 Block Diagram of Object Identification

Firstly, extract image from real time system through camera, then feature mapping takes places by comparing with dataset, if it maps with the stored dataset, then the layer corresponding to object classes become bounding box layer. The bounding boxes are filled with 1.5 in the bounding box and 0.5 in the other area and resized to same size. Then it starts labelling for each and every object in an image. object detection for each object is done then starts pairing for similar object if founds then pair the object and makes a visual relationship of object in an image.

IV. CONCLUSION

In this paper various object detection and tracking are studied and a number of shortcoming and limitations were highlighted in each and every technique. we compared different techniques for object detection and visual relationship for pairing based on size, shape, colour of objects in an image. With the help of methods used in this paper and adding addition large database and improved techniques for fast accurate object detection and visual relationship of objects in an image. It can be summarized as Google AI Open Images - Object Detection Track is a simplest method provides computationally efficiency compared to all other methods. In future advance study may carried out to find efficient algorithm to reduce the tracking drift

REFERENCES

[1] B. Babenko, M.-H. Yang, and S. Belongie, "Visual tracking with online multiple instance learning," in Proc. IEEE Conf. Comput. Vision Pattern Recogn., Jun. 2009, pp. 983–990.

[2] H. Grabner, C. Leistner, and H. Bischof, "Semi-supervised on-line boosting for robust tracking," in Proc. Eur. Conf. Comput. Vision, 2008, pp. 234–247

[3] "A novel solution for multi-camera object tracking" Conference Paper · January 2015: Authors: Weihua Chen Lijun Cao Xiaotang Chen Kaiqi Huang IEEE International Conference on Image Processing, ICIP 2014

[4] Khushboo Khurana, Reetu Awasthi "Techniques for Object Recognition in Images and Multi-Object Detection "International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 4, April 2013

[5] Nidhi "Image Processing and Object Detection" International Journal of Applied Research 2015; 1(9): 396-399

[6] Fare's Jalled& Cybernetics Ilia Voronkov, Moscow Institute of Physics & Technology, Department of Radio Engineering "Object Detection Using Image Processing" arXiv.1611.077 [cs.CV] 23 Nov 2016.

[7] Liming Wang1, Jianbo Shi2, Gang Song2, and I-fan Shen1, Fudan University &University of Pennsylvania, 3330 Walnut Street, Philadelphia."Object Detection Combining Recognitionand Segmentation".

[8] Yi Wei1, Nenghui Song1, Lipeng Ke2, Ming-Ching Chang1, Siwei Lyu1 1University at Albany, State University of New York 2University of Chinese Academy of Sciences "Street object detection/tracking for ai city traffic analysis".

[9] Kumar S. Ray and Soma Chakraborty "An Efficient Approach for Object Detection and Tracking of Objects in a Video with Variable Background ", arXiv:170620652 v1 [cs.CV] 11 may 2017.

[10] Andrew D Bagdanor "Colour attributes for object detection", may 2018