International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 08 Issue: 07 | July - 2024 SJIF Rating: 8.448 ISSN: 2582-3930

A Review of Real Time Image Processing for Object Detection

Prof. Shivaji Goroba Shinde1 ; Mr. Shubham Suresh Patil

Shivajishinde261@gmail.com ; shubhpatil2211@gmail.com

TPCT's College of Engineering , Osmanabad (Dharashiv), India.

Abstract

In past days, capture images with very high quality and good size is so easy because of rapid improvement in quality of capturing device with less costly but superior technology. Videos are a collect of sequential images with a constant time interval. So video can provide also more information about our object when scenarios about to changing with respect to time. Therefore, manually handling videosit can be quite impossible. That time all that need an automatic devise to process these videos. In this thesis one such attempt has been made to track objects in videos. Many algorithms and technology have been developed to automate monitoring the object in a video file. Object detection and tracking is a one of the challenging task in computer vision. Mainly there are three basic steps in video analysis: Detection of objects of Interest from moving objects, Tracking of that interested objects in consecutive frames, and Analysis of object tracks to understand their behavior

Some common choice to choose suitable feature to categories, visual objects are intensity, shape, color and feature points. In this thesis, we studied about mean shift tracking based on the color pdf, optical flow tracking based on the intensity and motion; SIFT tracking based on scale invariant local feature points.

Keywords: real-time, object detection, tracking, surveillance

INTRODUCTION

Object detection and tracking is an important challenging task within the area in Computer Vision that try to detect, recognize and track objects over and over a sequence of images called video. It helps to understand, describe object behavior instead of monitoring computer by human operators. It aims to locating moving objects in a video file or surveillance camera. Object tracking is the process of locating an object or multiple objects using a single camera, multiple cameras or given video file.

Invention of high quality of the imaging sensor, quality of the image and resolution of the image are improved, and the exponential increment in computation power is required to be created of new good algorithm and its application using object tracking. In Object Detection and Tracking we have to detect the target object and track that object in consecutive frames of a video file.

Challenges of Object Detection and Tracking Object tracking fundamentally entails estimating the location of a particular region in successive frames in a video sequence. Properly detecting objects can be a particularly challenging task, especially since objects can have rather complicated structures and may change in shape, size, location and orientation over subsequent video frames. Various algorithms and schemes have been introduced in the few

The most popular application in this area is vision-based surveillance, to help understand the movement patterns of people with suspicious actions. Traffic scene analysis is also a well-known application, to get the tracking information for keeping the vehicles in lane and preventing the accidents. Thus, object detection and tracking under dynamic conditions is still a challenge for real-time performance which requires the computational complexity to be minimum. Various methods for object detection have been proposed; such as feature-based, template-based object detection and background subtraction [1]. But selection of the best technique for a specific application is relative and dependent upon the hardware resources and scope of the application. Feature-based detection searches for corresponding features in successive frames, including Harris corner, edges, SIFT, contours or colour pixels. Background subtraction is a popular method which uses static background and calculating the difference between the hypotheses background and the current image. This approach is fast and good for fixed background but it cannot deal with the dynamic environment, with different illumination and motions of small objects

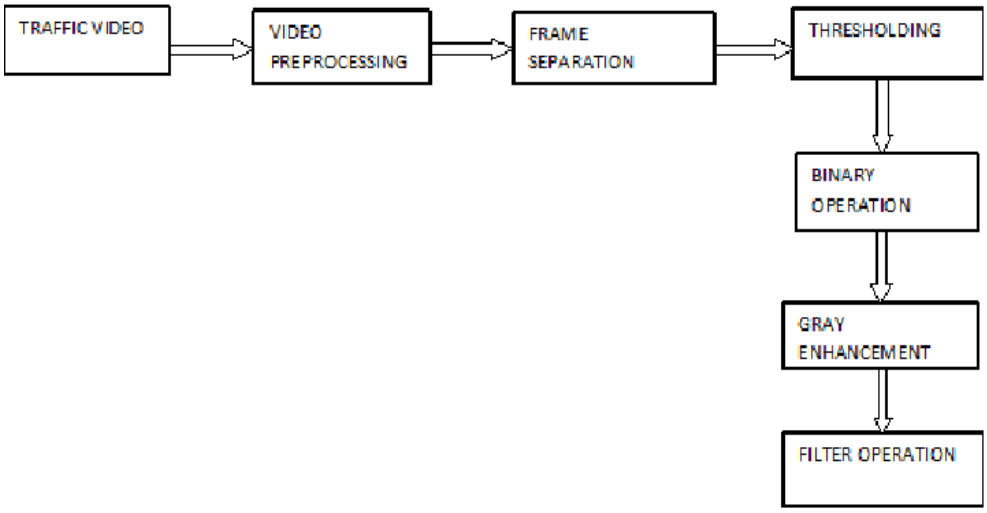


Figure 1 : The Basic flow diagram of Object tracking

This paper implements a method to track and recognize the object in a surveillance area. We analyze usual pixel-approach. Camera system (webcam) acts as a sensor to track the object in an surveillance area. Edge detection is an image segmentation process is implemented to have clear knowledge on real edges of real time video. Background separation...

As a name suggested it is a nothing but separation of foreground object from the background in a sequence of video frame and the fundamental logic is that the difference between current frame and reference frame and mainly reference frame is also known as a background image

Any object tracking algorithm will contain errors which will eventually cause a drift from the object of interest. The better algorithms should be able to minimize this drift such that the tracker is accurate over the time frame of the application. In object tracking the important challenge that has to consider while the operating a video tracker are when the background is appear which is similar to interested object or another object which are present in the scene.

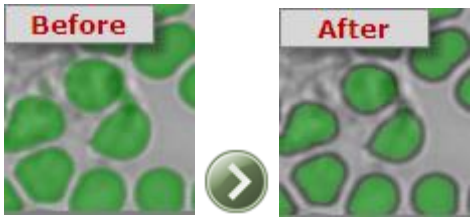
Thresholding is one of the segmentation techniques that generates a binary image (a binary image is one whose pixels have only two values – 0 and 1 and thus requires only.

Image segmentation is the technique of subdividing an image into constituent sub-regions or distinct objects. The level of detail to which subdivision is carried out depends on the problem being solved. That is, segmentation should stop when the objects or the regions of interest in an application have been detected.

Segmentation of non-trivial images is one of the most difficult tasks in image processing. Segmentation accuracy determines the eventual success or failure of computerized analysis procedures. Segmentation procedures are usually done using two approaches – detecting discontinuity in images and linking edges to form the region (known as edge-based segmenting), and detecting similarity among pixels based on intensity levels (known as threshold-based segmenting).

Binary Image Operations are a collection of digital filters, image arithmetic and image processing techniques. Apply simple modifications such as open and fill holes to count the number of cells. Develop complex Work files that utilize multiple binary layers to differentiate nuclei inside cells.

Shrinks both the interior and exterior boundaries of objects in the displayed binary image by one pixel for each pass applied. Erode can be used to shrink small objects until they disappear so that they will not be measured or counted later. Erode can have multiple passes applied, although the shape of objects after many passes may change.



Grey level thresholding is a simple lookup table, which partitions the gray levels in an image into one or two categories - those below a user-selected threshold and those above. Thresholding is one of many methods for creating a binary mask for an image

Method

Systems for object tracking might be arranged into emulating four classifications as per the tool utilized throughout tracking.

In above diagram has multiple different method to use in a image processing and video detection method so the main detection is that object detection all other are the subpart of the image detection .These strategies give a productive approach to decipher and investigate movement in a frame sequence of video. A frame district might be characterized as a set of pixels having homogeneous attributes. It could be determined by image segmentation, which might be focused around different object characteristics like color, edges and so forth. Basically, a district would be the image range secured by the projection of the object of investment onto the frame plane. On the other hand, a locale could be the bouncing box of the anticipated object under examination.

There are three different methods as follows :

1. **Contour-based methods** :

An option method for concocting an object tracking algorithm is by representation of object utilizing contour shape information 15 Chapter 2 Object detection and tracking and tracking it time to time, hence recovering both its position and shape. Such a demonstrating technique is more entangled than displaying whole locales. Then again, contour based tracking are typically more hearty than region based object tracking algorithm, on the grounds that it could be adjusted to adapt to halfway impediments. Additionally the outline information is unfeeling to light varieties.

2. Feature point-based methods :

Feature point-based object tracking could be characterized as, the endeavor to recuperate the motion parameters of a characteristic point in a feature succession. All the more formally, let f = f0, f1, , fn means the N frames of a video file sequence and pi (xi, yi) , i = 0, 1,...,N indicate the positions of the same characteristic point in those frame. The current task is to focus a motion vector di (dx,i, dy,i) that best decides the position of the feature points in the following frame, mi+1 (xi+1, yi+1), that is: mi+1 = mi + di. The interested object to be tracked is generally characterized by the bouncing box or the curved structure of the tracked feature point.

3. **Template-based methods** :

Template-matching procedures are generally utilized by numerous scientists to perform object tracking. Template based tracking is a nearly identified with region based tracking on the grounds that a template is basically a model of the picture area to be tracked. These routines include two steps for tracking; introduction step took after by matching step. In the first step template might be instated by different on-line and off-line strategies. Throughout matching, it includes the procedure of seeking the interested object to focus the image district that looks like the layout, taking into account a likeness or separation measure. In present commitment article following is attained utilizing characteristic point-based technique.

Challenges of Object Recognition:

- Since we take the output generated by last (fully connected) layer of the CNN model is a single class label. So, a simple CNN approach will not work if more than one class labels are present in the image.
- If we want to localize the presence of an object in the bounding box, we need to try a different approach that not only outputs the class label but also outputs the bounding box locations.

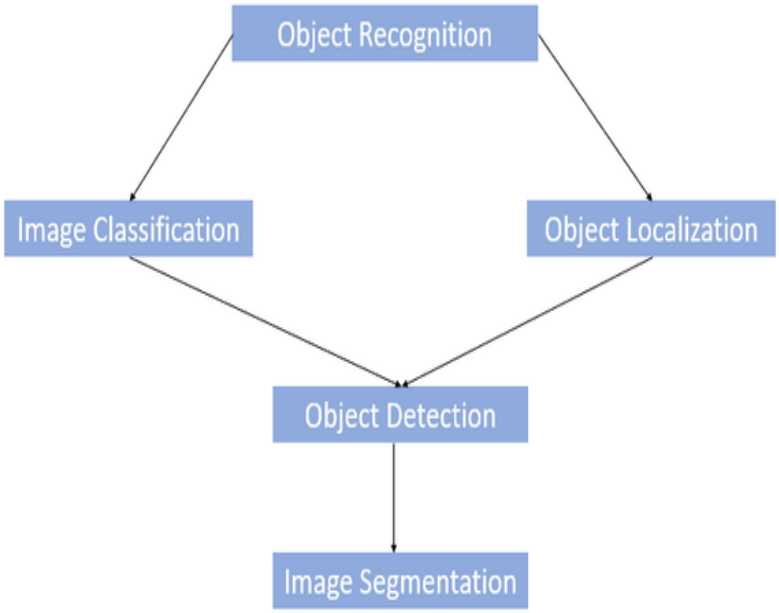


Figure : challenges in objectRecognition

Image Classification :

In Image classification, it takes an image as an input and outputs the classification label of that image with some metric (probability, loss, accuracy, etc). For Example: An image of a cat can be classified as a class label "cat" or an image of Dog can be classified as a class label "dog" with some probability.

Object Localization:

This algorithm locates the presence of an object in the image and represents it with a bounding box. It takes an image as input and outputs the location of the bounding box in the form of (position, height, and width).

Object Detection:

Object Detection algorithms act as a combination of image classification and object localization. It takes an image as input and produces one or more bounding boxes with the class label attached to each bounding box. These algorithms are capable enough to deal with multi-class classification and localization as well as to deal with the objects with multiple occurrences.

Image Segmentation:

Image segmentation is a further extension of object detection in which we mark the presence of an object through pixel-wise masks generated for each object in the image. This technique is more granular than bounding box generation because this can helps us in determining the shape of each object present in the image because instead of drawing bounding boxes , segmentation helps to figure out pixels that are making that object. This granularity helps us in various fields such as medical image processing, satellite imaging, etc. There are many image segmentation approaches proposed recently. One of the most popular is Mask R-CNN proposed by *K He et al*. in 2017.

- **Instance Segmentation:** Multiple instances of same class are separate segments i.e. objects of same class are treated as different. Therefore, all the objects are coloured with different colour even if they belong to same class.
- **Semantic Segmentation:** All objects of same class form a single classification ,therefore , all objects of same class are coloured by same colour

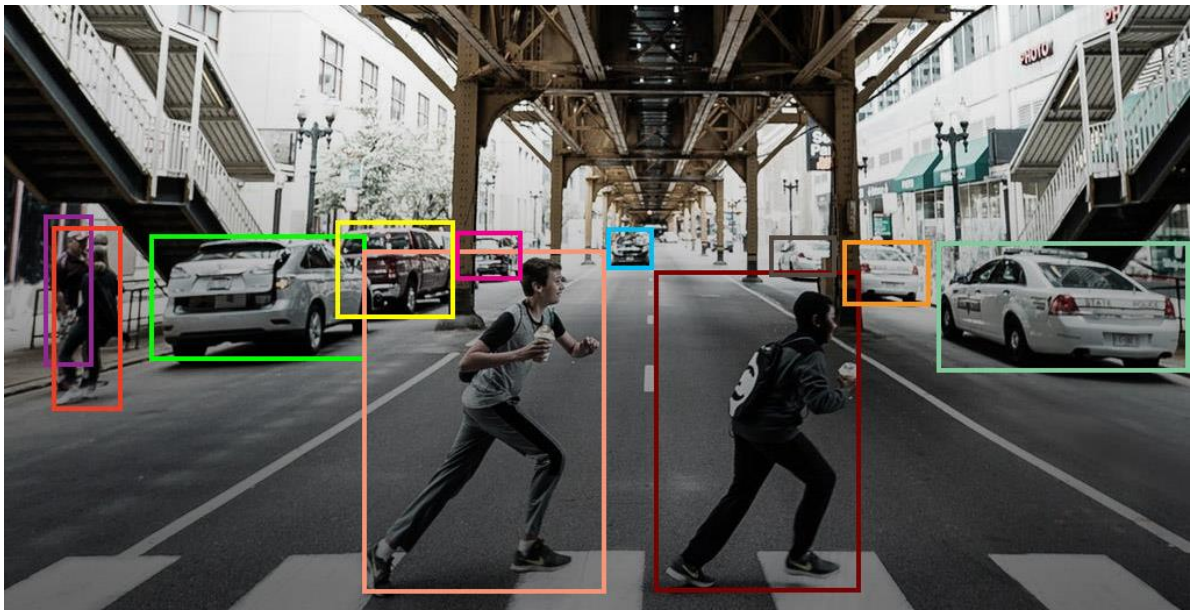


Figure 1 : Real Time Object Detection

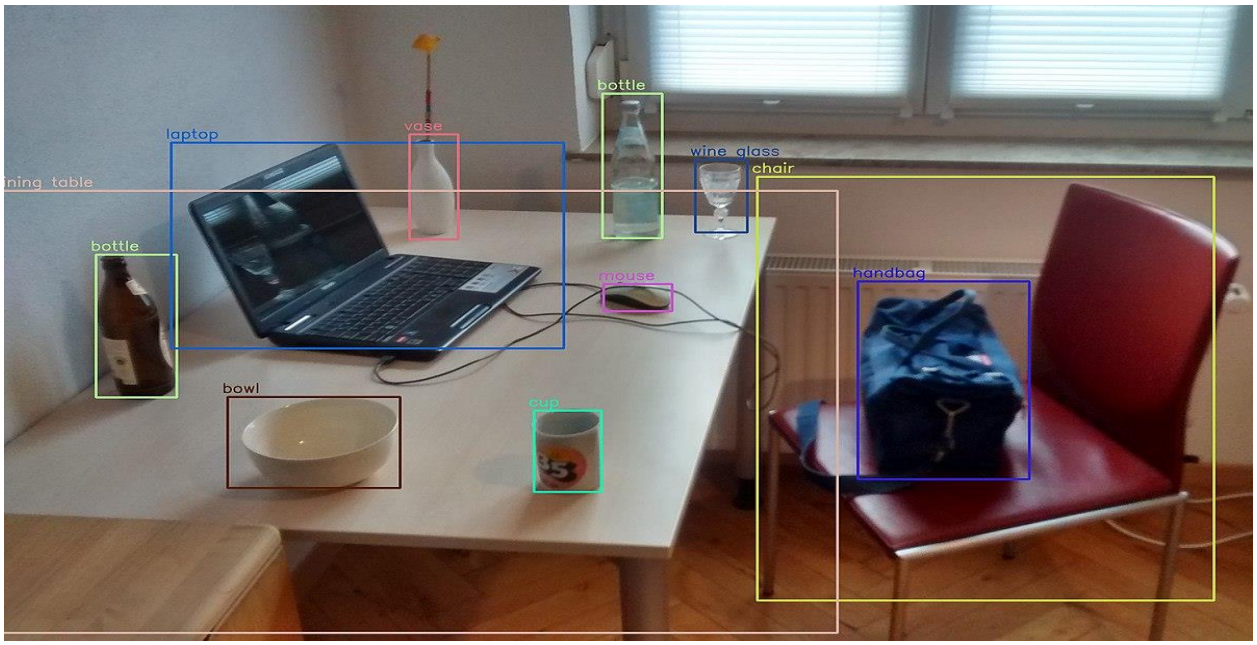


Figure 2: Real Time Image Processing and Image Detection

CONCLUSION

In this paper, review on different object detection, tracking, recognition techniques, feature descriptors and segmentation method which is based on the video frame and various tracking technologies. This approach used towards increase the object detection with new ideas. Furthermore, tracking the object from the video frames with theoretical explanation is provided in bibliography content. The bibliography content is the most significant contribution of research since it will lead to a new area of research. We have identified and discussed the limitation/future scope of various methods. Also, we have noted some methods which give accuracy but have high computational complexity. Specifically, the statistical methods, background subtraction, temporal differencing with the optical flow was discussed. However, these technique needs to concentrate towards handling sudden illumination changes, darker shadows and object occlusions (Susar and Dongare, 2015).

FUTURE SCOPE

Design and simulation of complex video sequence and test them using same tracking algorithm. In the potential scenario, occlusion is used for an object with the same color for the moving objects or else using bigger occlusion with longer occlusion time. Increasing the number of the object help to identify the efficiency and functionality of the tracking algorithm. ¬

Weight parameters are needed to be added for individual intensity levels of each pixel. In an image, if an intensity value is assigned as foreground based on the current frame then it has less probability that foreground also has similar pixel coordinate so that BG weightage for the pixel is set to the minimum than the initial value. Through adding weightage lower than the initial value provides the advantage of removing the old pixel value with least probability rather than the evolved scene.

Need to focus towards enhancing the variance data of each channel based on the Mahalanobise distance calculation. By this, can able to adopt a change in the rapid scene through Euclidean distance algorithm.

REFERENCES

1. Aggarwal, A., Biswas, S., Singh, S., Sural, S. & Mujumdar, A.K., 2006. Object Tracking Using Background Subtraction and Motion Estimation in MPEG Videos, in 7th Asian Conference on Computer Vision. Springer Berlin Heidelberg, pp. 121–130. doi:10.1007/11612704_13
2. Amandeep, Goyal, M., 2015. Review: Moving Object Detection Techniques. Int. J. Comput. Sci. Mob. Comput. 4, 345 – 349.
3. Athanesious, J. & Suresh, P., 2012. Systematic Survey on Object Tracking Methods in Video. Int. J. Adv. Res. Comput. Eng. Technol. 1, 242–247
4. Avidan, S., 2004. Support vector is tracking. IEEE Trans. Pattern Anal. Mach. Intell. 26, 1064–1072. doi:10.1109/TPAMI.2004.53
5. Badrinarayanan, V., Perez, P., Le Clerc, F. & Oisel, L., 2007. Probabilistic Color and Adaptive Multi-Feature Tracking with Dynamically Switched Priority Between Cues, in: 2007 IEEE 11th International Conference on Computer Vision. IEEE, pp. 1–8. doi:10.1109/ICCV.2007.4408955
6. Balasubramanian, A., Kamate, S., & Yilmazer, N., 2014. Utilization of robust video processing techniques to aid efficient object detection and tracking. Procedia Comput. Sci. 36, 579–586. doi:10.1016/j.procs.2014.09.057
7. Blackman, S.S., 2004. Multiple hypotheses tracking for multiple target tracking. IEEE Aerosp. Electron. Syst. Mag. 19, 5–18. doi:10.1109/MAES.2004.1263228

8. Chakravarthy, S., Aved, A., Shirvani, S., Annappa, M., & Blasch, E., 2015. Adapting Stream Processing Framework for Video Analysis. Procedia Comput. Sci. 51, 2648–2657. doi:10.1016/j.procs.2015.05.372
9. Chate, M., Amudha, S., & Gohokar, V., 2012. Object Detection and tracking in Video Sequences. ACEEE Int. J. Signal Image Process.
10. Collins, R.T., Yanxi Liu, & Leordeanu, M., 2005. Online selection of discriminative tracking features. IEEE Trans. Pattern Anal. Mach. Intell. 27, 1631–1643. doi:10.1109/TPAMI.2005.205
11. Comaniciu, D., Ramesh, V. & Meer, P., 2003. Kernel-based object tracking. IEEE Trans. Pattern Anal. Mach. Intell. 25, 564–577. doi:10.1109/TPAMI.2003.1195991
12. Comaniciu, D., Ramesh, V. & Meer, P., 2000. Real-time tracking of nonrigid objects using mean shift, in: IEEE International Proceedings of the Computer Vision and Pattern Recognition. IEEE Computer Society, DC, pp. 142–14
13. Cucchiara, R., Grana, C., Piccardi, M. & Prati, A., 2003. Detecting moving objects, ghosts, and shadows in video streams. IEEE Trans. Pattern Anal. Mach. Intell. 25, 1337–1342. doi:10.1109/TPAMI.2003.1233909
14. Feris, R.S., 2000. Detection and Tracking of Facial Features in Video Sequences, in: Lecture Notes in Artificial Intelligence. Springer-Verlag Berlin Heidelberg, pp. 197–206.
15. Gang, M., Zhiguo, J., Zhengyi, L., Haopeng, Z. & Danpei, Z., 2010. Fullviewpoint 3D space object recognition based on kernel locality preserving projections. Chinese J. Aeronaut. 23, 563–572. doi:10.1016/S1000-9361(09)60255-7
16. Guan, N., Tao, D., Luo, Z. & Shawe-Taylor, J., 2012. MahNMF: Manhattan Non-negative Matrix Factorization.
17. Guo, L., Liao, Y., Luo, D. & Liao, H., 2012. Generic Object Detection Using Improved Gentle boost Classifier. Phys. Procedia 25, 1528–1535. doi:10.1016/j.phpro.2012.03.272
18. Haritaoglu, I., Harwood, D. & Davis, L.S., 2000. W/sup 4/: real-time surveillance of people and their activities. IEEE Trans. Pattern Anal. Mach. Intel. 22, 809–830. doi:10.1109/34.868683