A review of different techniques for real time object recognition

Piyush Sinha
RegNo 11708835
Lovely Professional University, Phagwara, Punjab

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 videos. Every tracking mechanism requires object detection mechanism either in each frame or when an object appears newly on the video sequence. 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. 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. Recently computer vision research has to address the Multiple object detection and tracking in dynamic environment.

Keywords— Object Tracking Algorithms, Object Detection, frame differencing.

Introduction

The activity of looking for a specific object among others is really simple for a human brain. We do it all the time, we are used to this process. However, computers have obvious challenges with this seemingly easy task. That's why computer engineers around the world are trying their best to train Artificial Intelligence how to find the needed objects in pictures. And this is no small task for developers. To train the AI tool to detect certain objects, you have to show these objects first. In other words, you should 'feed' AI with the labeled data – images containing the needed objects, item coordinates, location, and class labels. The most frequently asked question here is "How many images are needed?" The answer is the more, the better. Also, you should choose images with different locations of the object, so that items change their coordinates and sizes during machine learning.

It will help AI understand that even though this object can be located in different places on the image and be both big and small, these changes don't affect its class. So, as you can see, it is a time-consuming process that requires lots of resources and efforts. But let's look on the bright side. Artificial Intelligence is already making quite a progress here. With GPUs – Graphics Processing Units – deep learning has become much faster and easier. GPU is an electronic circuit that allows to manipulate the memory and accelerate graphics processing.

Literature Review

In computer vision field, object detection and tracking plays a vital role. Object detection means locating/identifying objects in frame of video sequence and whereas tracking is the process of locating moving object or multiple objects over a period of time using camera. Technically, tracking is estimating trajectory or path of an object in the image. The availability of high power computers, high quality and low cost camera will lead the great deal of interest in object tracking algorithms. Three main key steps for video analysis are: Detection of Interesting moving Objects, Tracking of such objects from frame to frame, Analysis of Object tracks to recognize their behaviour. The main application areas of object detection and tracking are: Motion based recognition, automated surveillance, video indexing, human-computer interaction, traffic monitoring, vehicle navigation and etc.

But object detection and tracking is multifaceted course of action when projecting 3D world on 2D image because of the loss of information, noise present in an image, complex object motion, articulated nature of object, complex object shapes and occlusion. Almost all tracking algorithms assume that the object motion is smooth with no abrupt changes. But practically it is impossible. Even though so many difficulties are exist in object detection and tracking, a number of object detection and tracking algorithms are proposed and implemented. Each and every algorithms are varied with respect to the object representation used, image features and the motion, appearance and shape of the object be modelled for detection and tracking. The shape representation should be combined with appearance representation. There are a various ways to represent the appearance feature of objects.

Proposed Methodology

In computer vision field, object detection and tracking plays a vital role. Object detection means locating/identifying objects in frame of video sequence and whereas tracking is the process of locating moving object or multiple objects

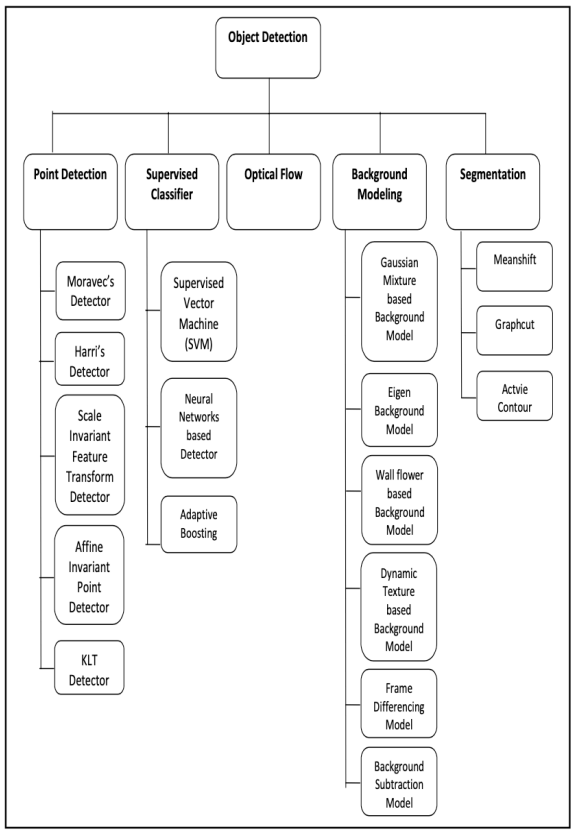
over a period of time using camera. The main application areas of object detection and tracking are: Motion based recognition, automated surveillance, video indexing, human-computer interaction, traffic monitoring, vehicle navigation and etc. But object detection and tracking is multifaceted course of action when projecting 3D world on 2D image because of the loss of information, noise present in an image, complex object motion, articulated nature of object, complex object shapes and occlusion. There are a various ways to represent the appearance feature of objects. Some of the common appearance representation in the case of object tracking is described in which are as follows:

A. Object Representation

An entity of an interest is an object. Object can be represented by their shapes and features. There are various object representations available for tracking. The selection of object representation for tracking is based on application. Some of the object representation techniques are: Points representation which is suitable for tracking an object which occupies small region in an image, Primitive geometric shape representation is suitable for tracking both rigid and non rigid objects, Object silhouette representation is suitable for tracking complex non rigid shapes, articulated shape model is suitable for tracking articulated object with torso, legs, hands and etc, skeletal model is suitable for tracking both articulated and rigid objects.

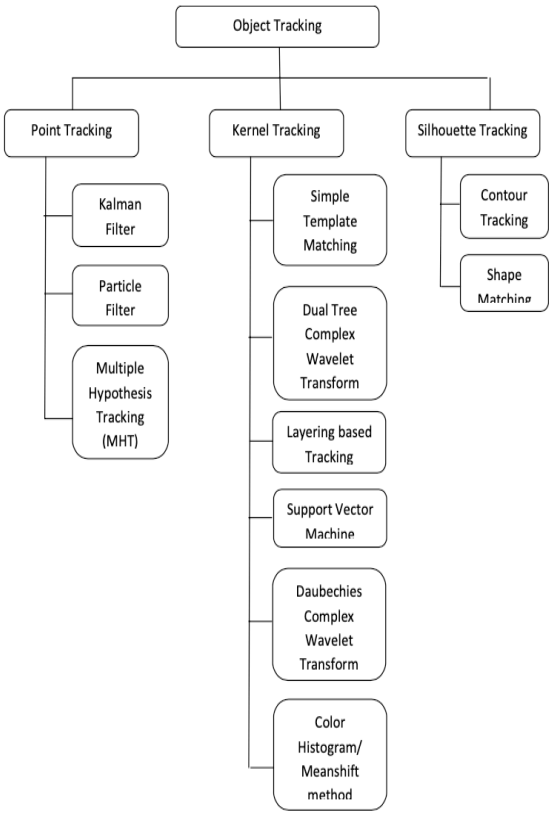
B. Object Detection

The object detection is the process of locating objects in the frame of video sequence. Every tracking algorithm requires an object detection mechanism either in every frame or when an object occurs newly in a frame. Most of the object detection mechanism used information from single frame for detecting an object. But some of the object detection mechanism used temporal information which is computed from sequence of frames. It will reduce the false detection rate.



C. Object Tracking

An object tracker aims to generate trajectory of an object over time by locating its position in every frames of video [1]. Generally, object tracking has two main tasks. One is to detect the object and another one is to establish the correspondence between the object in every frames of video. These tasks can be performed either separately or jointly. The first case, the object in region has been detected by detector mechanism and later connection between objects in sequence of frame has been done by tracker. In the later case, the object region and correspondence is jointly estimated. It can be done by iteratively updating object location and region information obtained from previous frames.



a. Point Tracking

Moving object can be represented by point in an image structure. An identification of point in a moving object is done by threshold value. Point tracking is difficult when occlusion will be appeared. Some of the points tracking algorithms are: Kalman Filter, Particle Filter and Multiple Hypothesis tracking.

b. Kernel Tracking

Kernel tracking tracks the moving objects which are represented by growing object region from one frame to another. The geometric object representation is common in real time. But the restriction here is that, the part of the moving object defined may be left outside of the region or some background object may be covered by the region. The various kernel tracking methods are :Simple template matching, Dual Tree Complex Wavelet Transform, Layering based tracking, Support Vector Machine and Color Histogram or Mean Shift Method.

c. Silhouette Tracking

Simple geometric based object representation is not suitable for complex objects representation such as hands, fingers, shoulders and etc. Silhouette based method gives an accurate shape description for the object which find the objects in each and every frame by means of an object model generated in the previous

frame. Silhouette based method is suitable for the variety of object shape; occlusion and object split & merge. The various silhouette tracking methods are described: Contour tracking and Shape matching.

Results and Discussion

COMPARISION OF REAL TIME OBJECT DETECTION AND TRACKING ALGORITHMS

a. Object Detection Algorithms:

An accuracy of detecting object in background subtraction based object detection algorithms are low to moderate and it requires moderate computational time. An optical flow based object detection algorithms require high computational power and the detection of object is moderate. This algorithm produces the complete movement information about the object. But it requires complex calculation. Frame differencing based object detection algorithm detect the object with high accuracy. The computational requirement for this algorithm is low to moderate. It is very easiest method. It will support only for static background and not for moving background.

b. Object Tracking Algorithms:

Some of the point tracking algorithms viz kalman filtering algorithm can track single objects and it cannot handle occlusion. But the response colour image/video to gray scale. Hence time of detecting and tracking an object is good. Some other point tracking algorithms viz Multiple Hypothesis Tracking (MHT) and particle filter based.

on code book background can track multiple objects and handle occlusion. But it needs high computation and memory requirements.

Most of the kernel tracking algorithm can track single object and partial occlusion handling. Histogram based kernel tracking algorithm runs very fast which is suitable for the models which are having dominant color. Once the spatial information of target is lost, this algorithm will not give good performance. Daub CxWT kernel tracking algorithm reduces the false tracking of object and this algorithm could not detect and track the object in which shapes undergo change from one frame to another.

Silhouette tracking can handle variety of object shapes and they have the capability to deal with object split and merge.

CONCLUSION

In this paper, I did the widespread literature survey on both object detection and tracking algorithms and its various requirements. Based on the survey, we observed the following issues on real time object detection and tracking: 1. Most of the existing algorithms work on gray scale image/video. It losses some

information when it is converted from the detection and tracking is complex. 2. All tracking algorithm assumes that the object motion is smooth with no abrupt changes. 3. Object detection and tracking algorithms requires more computational and memory requirements when increase the amount of data contained in the video. 4. Some of the detection and tracking algorithms can detect and track multiple objects and handle occlusion. But it needs more computational and memory requirement. A different object tracking algorithm can handle the varying illumination, background clutter, camouflage, bootstrapping and handling occlusion separately. It is hard to implement all these in one algorithm.

Based on the observation made, we concluded that real time object detection and tracking algorithms require parallel programming platform viz either CUDA or WebGL or WebCL or Grid or Cloud computing platform for satisfying its computational requirements.

References

- Alper Yilmaz, Omar Javed and Mubarak Shah, —Object Tracking: A Survey",ACM Computing Surveys (CSUR),38(4):13, 2006.

- Chen Yan-yan, Chen Ning, Zhou Yu-yang,Wu Ke-han, and Zhang Wei-wei,‖ Pedestrian Detection and Tracking for Counting Applications in Metro Station‖, Discrete Dynamics in Nature and Society,Volume 2014, Article ID 712041, 11 pages,http://dx.doi.org/10.1155/2014/712041,2014

- J. Jiménez, A.M. López, J. Cruz, et.al., —A Web platform for the interactive visualization and analysis of the 3D fractal dimension of MRI data‖, Journal of Bio Medical Informatics,Elsevier,2014.

- WeizhiNie , AnanLiu , YutingSu , HuanboLuan and ZhaoxuanYang,,‖ Single/cross-cameramultiple-persontrackingbygraphmatching‖, Journal of Neurocomputing, Elseveir, 2014.