International Journal of Latest Trends in Engineering and Technology
Special Issue SACAIM 2017, pp. 150-154
e-ISSN:2278-621X

COMPARATIVE STUDY ON OBJECT DETECTION AND TRACKING

LeedaJovita Rodrigues[1], Cynthia D'Souza N[2] & Nausheeda B.S[3]

Abstract- Object detection is a computer technology related to computer vision and image processing that deals with detecting instances of semantic objects of a certain class (such as humans, buildings, or cars) in digital images and videos. Well-researched domains of object detection include face detection and pedestrian detection. Object detection has applications in many areas of computer vision, including image retrieval and video surveillance. In a tracking scenario an object can be defined as anything that is of interest for further analysis. Objects can be represented by their shapes. Object shape representations commonly employed for tracking are: points and primitive geometric shapes. Object detection and tracking are important and challenging tasks in many computer vision applications such as surveillance, vehicle navigation, and autonomous robot navigation. Video surveillance in a dynamic environment, especially for humans and vehicles, is one of the current challenging research topics in computer vision. It is a key technology to fight against terrorism, crime, public safety and for efficient management of traffic.

1. INTRODUCTION

The video cameras are the most commonly used sensors in a large number of applications which are starting from surveillance to smart rooms for video conferencing. Continuously moving object detection means to detect moving objects from the background image to the continuous video image. Currently used methods in moving object detection are mainly frame subtraction method, then background subtraction method. Frame subtraction method is considering the difference between two consecutive frames to determine the presence of moving objects. WhereasBackground subtraction orforeground detection, is a technique wherein an image's foreground is extracted for further processing (object recognition etc.). Generally, an image's regions of interest are objects (humans, cars, text etc.) in its foreground. After the stage of image preprocessing (which may include image denoising, post processing like morphology etc.) object localization is required which may make use of this technique.

One of the most popular methods is Frame difference using online K-means approximation. This is a technique which is used for computation cost and memory requirements are very low. However, Due to the loss of color information the foreground segmentation accuracy unavoidable decreases. To solve the problem of resource requirement and accuracy, this paper proposes a background subtraction method by Kaman filter method.

Following are the basic steps for tacking an object:

1)Object Detection: Object Detection is to identify objects of interest in the video sequence and to cluster pixels of these objects. Object detection can be done by various techniques such as frame differencing, Optical flow and Background subtraction.

2) Object Classification:Object can be classified as vehicles, birds, floating clouds, swaying tree and other moving objects. The approaches to classify the objects are Shape-based classification, Motion-based classification, Color based classification and texture based classification.

3) Object Tracking: Tracking can be defined as the problem of approximating the path of an object in the image plane as it moves around a scene. The approaches to track the objects are point tracking, kernel tracking and silhouette. Following are some of the challenges that should be taken care in object tracking as described in [3]:

1. Loss of evidence caused by estimate of the 3D realm on a 2D image,
2. Noise in an image,
3. Difficult object motion,
4. Imperfect and entire object occlusions,
5. Complex objects structures.

2. OBJECT DETECTION

[1] Department of Information Technology& Bioinformatics, AIMIT, St.Aloysius College,Beeri, Mangalore
[2] Department of Information Technology & Bioinformatics, AIMIT, St. Aloysius College,Beeri, Mangalore
[3] Department of Information Technology& Bioinformatics, AIMIT,St.Aloysius College, Beeri, Mangalore

Every tracking method requires an object detection mechanism either in every frame or when the object first appears in the video. A common approach for object detection is to use information in a single frame. However, some object detection methods make use of the temporal information computed from a sequence of framesto reduce the number of false detections. For object detection, there are several common object detection methods described in.

1. Point Detectors-Point detectors are used to find interesting points in imageswhich have an expressive texture in their respective localities. A desirablequality of an interest point is its invariance to changes in illumination and camera viewpoint. In literature, commonly used interest point detectors includeMoravec's detector, Harris detector, KLT detector, SIFT detector.
2. Background Subtraction-Object detection can be achieved by building are presentation of the scene called the background model and then findingdeviations from the model for each incoming frame. Any significant changein an image region from the background model signifies a moving object.The pixels constituting the regions undergoing change are marked for further. This process is referred to as the background subtraction.
3. Segmentation-The aim of image segmentation algorithms is to partitionthe image into perceptually similar regions. Every segmentation algorithmaddresses two problems, the criteria for a good partition and the methodfor achieving efficient partitioning.

3. OBJECT TRACKING

The aim of an object tracker is to generate the route of an object over time bylocating its position in every frame of the video. But tracking has two definitionsone is in literally it is locating a moving object or multiple object over a period oftime using a camera. Another one in technically tracking is the problem of estimatingthe route or path of an object in the image plane as it moves around a scene.The tasks of detecting the object and establishing a correspondence between theobject instances across frames can either be performed separately or jointly. In thefirst case, possible object region in every frame is obtained by means of an objectdetection algorithm, and then the tracker corresponds objects across frames. In thelatter case, the object region and correspondence is jointly estimated by iterativelyupdating object location and region information obtained from previous frames.There are different methods of Tracking. Point is tracking- Tracking can be formulated as the correspondence of detectingobjects represented by points across frames. Point tracking can be divided intotwo broad categories, i.e. Deterministic approach and Statistical approach.Objects detected in consecutive frames are represented by points, and theassociation of the points is based on the previous object state which can includeobject position and motion.

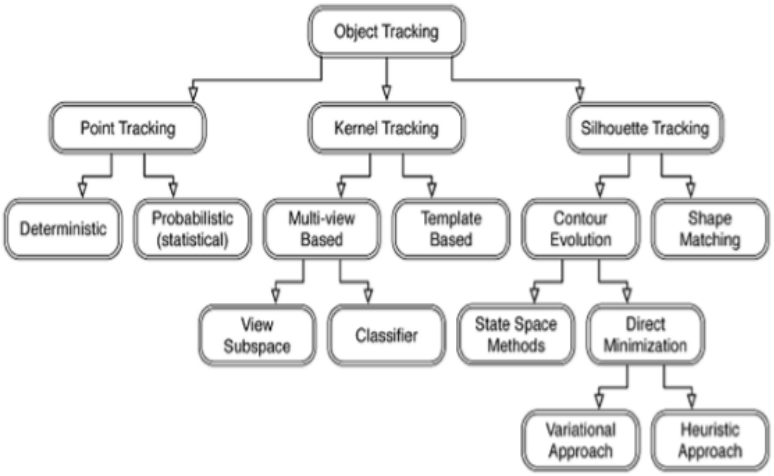

Figure.1 Object tracking

Based on the appearance representation used i.e. Template and Density-basedAppearance Model and Multi-view appearance model.Silhouette Tracking- It Provides an accurate shape description of the targetobjects. The goal of silhouette tracker is to find the object region in each frameby means of an object model generated using the previous frames. Silhouettetrackers can be divided into two categories i.e. Shape matching and Contourtracking.Object tracking consists in estimating of the trajectory of moving objects in thesequence of images. The most important is the automation of object tracking is achallenging task. Dynamics of multiple parameters, changes representing featuresand motion of the objects and temporary partial or full occlusion of the trackedobjects have to be considered.

4. FRAME DIFFERENCE

As the name suggest Background subtraction is the process of separating out foreground objects from the background in a sequence of video frames. There are many methods exist for background subtraction, with each having different strengths and weakness in terms of computational requirements and performance. Varying complexity can be used to implement Background subtraction.

1) Low complexity-using frame difference method.
2) Medium complexity using approximate median method.
3) High complexity using mixture of Gaussian method.

Frame difference method is the simplest method which can be used for to detect object in the surveillance video captured by immobile camera. Frame difference method is the simplest method for the reason that it has great detection speed, which can be implemented easily on the hardware and has been used widely. While detecting moving object by frame difference method, in the difference image, the image part which is unchanged will be eliminated while the changed part remains. This change is caused by movement or noise, so it calls for a binary process upon the difference image to distinguish the moving objects and noise. Furthermore, connected component labeling is also needed to acquire the smallest rectangle containing the moving objects.

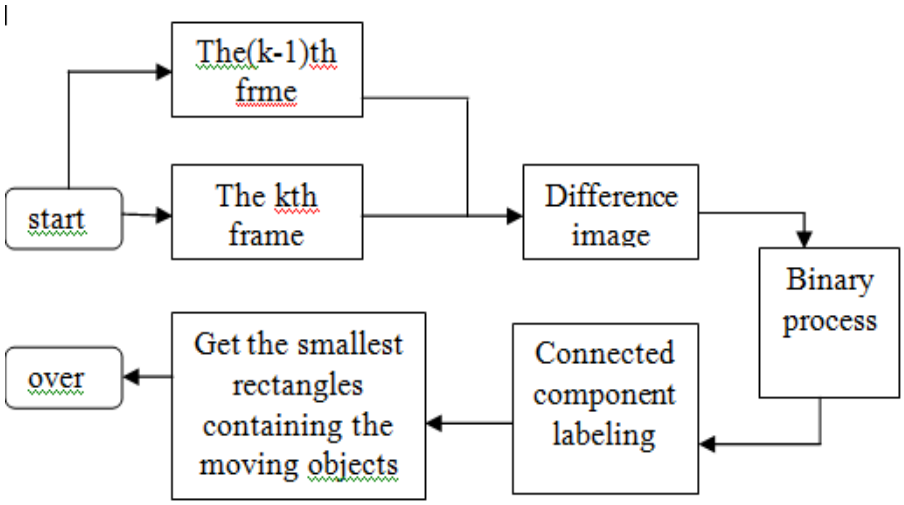

Figure.2 The flow chart of frame difference method

```
fr_bw=rgb2grey(fr); //convert to grayscale
fr_diff=abs(double(fr_bw)-double(bg_bw));
        //if fr_diff>thresh pixel in foreground
                for j=1: width
                        for k=1:height
                                if((fr_diff(k, j)>thresh))
                                        fg(k,j)=fr_bw(k,j);
                                else
                                        fg(k,j)=0;
                                end
                        end
                end
```

Algorithm of the frame difference

The algorithm is relatively simple.
1. Translate the received frame "fr" to grayscale (here we assume a color RGB sensor)
2. Subtract the current frame from the related model "bg_bw" (in this case it's just the previous frame)
3. For each pixel, if the difference between the present frame and related "fr_diff (j,k)" is greater than a threshold "thresh", the pixel is considered part of the foreground.

Fig.3 Original frame

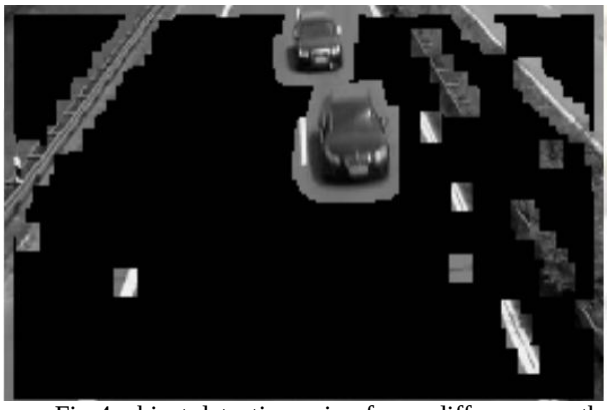

Fig.4. object detection using frame difference method

5. KALMAN FILTER

Kalman filters technique is also used to implement foreground object detection. Especially Kalman filter is used in tracking explicit curves and Kalman filter framework for the recovery of moving objects. But, the framework does not model dynamic, textured backgrounds. It is also work on dynamic videos. The algorithm of kalman filter is given below. It will read the vehicle video and convert video in to frames i.e :

I(x,y) For k=1:totalframes

Ik=FkIk-1+BkUk
Pk=FkPk-1F T k+Qk
Yk=zk-HkIk
Sk=HkPkH T k+Rk
Kk=PkH T kSk -1
Ik=Ik-1+KkYk

let Q=1e-5. (We could certainly let Q=0 but assuming a small but non-zero value gives us more flexibility in "tuning" the filter as we will demonstrate below.) Let's assume that from experience we know that the true value of the random constant has a standard normal probability distribution, so we will "seed" our filter with the guess that the constant is 0. Ik and zk are the actual state and measurement vectors. Ik and zk are the approximate state and measurement vectors. H is the Jacobian matrix of partial derivatives of h with respect to x.

$$H_{[i,j]} = \frac{\partial h_{[i]}}{\partial x_{[j]}}(\tilde{x}_{k,}, 0)(1)$$

The image Ik, represents the estimation of image at frame k given observation up to and including at frame k Pk is the error covariance matrix that is measure of the estimate accuracy of the state estimate. In the update phase, the current prediction is combined with current observation information to update the state estimate. Typically, the two phases alternate, with the prediction advancing the state until the next scheduled observation, and the update incorporating the observation. However, this is not necessary; if an observation is unavailable for some reason, the update may be skipped and multiple prediction steps performed. Likewise, if multiple independent observations are available at the same time, multiple update steps may be performed typically with different observation matrices Hk The formula for the updated estimate and covariance above is only valid for the optimal Kalman gain[1].

Fig.5 Original frame

Fig.6. object detection using kalman filter method

6. CONCLUSION

Different methods for object detection are frame difference, optical flow and background subtractions. Object tracking and detection can be performed using various methods like kalmanfilter and Frame Difference.Kalman filter has better accuracy compare to frame difference method. Furthermore, the segmentation has been improving and the object detection more smooth.

7. REFERENCES

[1] C.Srinivas Rao, P.Darwin. "Frame Difference And Kalman Filter Techniques For Detection Of Moving Vehicles In Video Surveillance ",IJERA,Vol. 2, Issue 6, November- December 2012,India

[2] Yerra chalapathi1, Dr. I Santi prabha2," Moving Vehicle Detection Using Kalman Method ",Volume 1, Issue 2, July 2012,India

[3] Himani S. Parekh1, Darshak G. Thakore 2, Udesang K. Jaliya 3," A Survey on Object Detection and Tracking Methods ", IJIRCCE ,Vol. 2, Issue 2, February 2014,India.