IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Detection And Tracking Of Object In Real Time Video

Poonam B. Linghate[1],Prof. K.R.Ingole[2]

Department Of Computer Science and Engineering, Sipna COET

Abstract— Object tracking is the process of locating moving objects over time using the camera in video sequences. The objective of object tracking is to associate target objects in consecutive video frames. Object tracking requires location and shape or features of objects in the video frames. So, object detection and object classification is the preceding steps of object tracking in computer vision application. To detect or locate the moving object in frame, Object detection is first stage in tracking. It is challenging or difficult task in the image processing to track the objects into consecutive frames. Various challenges can arise due to complex object motion, irregular shape of object, occlusion of object to object and object to scene and real time processing requirements. This paper presents the various techniques of object tracking in video sequences.

Keywords: object Detection ,Object tracking,Real Time Video.

I. INTRODUCTION

Object detection and tracking are important and challenging tasks in many computer vision applications[1],[2][3]. Object detection involves locating objects. Webcam is used to capture image continuously and computer processes the image and shows the path of object in monitor.Object tracking is the process of locating moving objects over time using the camera in video sequences. The objective of object tracking is to associate target objects in consecutive video frames. Object tracking requires location and shape or features of objects in the video frames. So, object detection and object classification is the preceding steps of object tracking in computer vision application. Every tracking method requires an object detection mechanism either in every frame or when the object first appears in the real time video. Object tracking is the process of locating an object or multiple objects over time using a camera [4]. The high powered computers, the availability of high quality and inexpensive video cameras and the increasing need for automated video analysis has generated a great deal of interest in object tracking algorithms. There are three key steps in video analysis, detection interesting moving objects, tracking of such objects from each and every frame to frame, and analysis of object tracks to recognize their behavior. Therefore, the use of object tracking is pertinent in the tasks of, motion based recognition. Automatic detection, tracking, and counting of a variable number of objects are crucial tasks [5]. Webcam is used to capture image continuously and computer processes the image and shows the path of object in monitor.

The basic idea is to first detect interest regions (keypoints) that are covariant to a class of transformations. Then, for each detected regions, an invariant feature vector representation (i.e., descriptor) for image data around the detected keypoint is built. Two types of image features can be extracted form image content representation; namely global features and local features. Global features (e.g., color and texture) aim to describe an image as a whole and can be interpreted as a particular property of the image involving all pixels. While, local features aim to detect keypoints or interest regions in an image and describe them. In this context, if the local feature algorithm detects n keypoints in the image, there are n vectors describing each one's shape, color, orientation, texture and more. The use of global color and texture features are proven surprisingly successful for finding similar images, while the local structure oriented

features are considered adequate for object classification or finding other occurrences of the same object or scene.

II. BACKGROUND

Computer vision tasks with these "digital eyes", attaching "brains" to the imaging devices and thus, creating a very useful tool used for video surveillance, entertainment/augmented reality applications, autonomous vehicles and driver assistance systems, robotics and smart health care [1]. Visual surveillance systems, address real-time observation of objects in some environment leading to a description about the activities or interaction of the objects within the environment or among the objects. However, a human operator has either to watch a massive amount of video data in real-time with full attention to detect any anomalies or events, or the video data can only be used as evidence after the abnormal event has occurred, due to the lack of real-time automatic tracking analysis [2].object detection/tracking, human or object analysis and finally, activity or behavioural analysis. These blocks are implemented using computer vision techniques and algorithms alleviating the load on humans and enabling preventive acts or alarms when a specific event is detected [1]. Video object tracking is to associate target objects in consecutive video frames. The association can be especially difficult when the objects are moving fast relative to the frame rate. Another situation that increases the complexity of the problem is when the tracked object changes orientation over time. For these situations video tracking systems usually employ a motion model which describes how the image of the target might change for different possible motions of the object. A feedback fractional loop is introduced in original Kalman filter to increasing the object tracking accuracy [1].

III. PREVIOUS WORK DONE

In [1] Harpreet Kaur proposed a Fractional order gain Kalman filter (FOGKF) to avoid divergence of extended Kalman filter. It uses fractional derivative based method to improve performance by modifying the Kalman gain so that the sensitivity for large perturbations can be increased. The fractional order gain Kalman filter performs better even in the presence of for abrupt variations in the inputs. The gain value will never be too large due to the feedback loop.

In [2] Jean-Philippe Jodoin proposed an Urban tracker to track various a priori unknown road users. The method is based on tracking the resulting foreground blobs of pixels. Each blob is modeled by a collection of key points. Data association is performed from frame to frame, and a finite state machine (FSM) corrects the associations by handling blob merging, splitting and fragmenting.

In [3] Yongzheng Xu proposes an enhanced V-J scheme to improve the accuracy, computational time of original V-J scheme. It reduces false detection rates and it significantly saves computational time. The basic idea is to directly detect the orientation of the road and rotate the road according to the detected orientation only once. The proposed road orientation adjustment method then can be incorporated with the original V-J scheme to achieve better vehicle detection.

In [4] Jun Xiang proposes a multi-target tracking framework that combines HFRF-based data association and a modified RJMCMC algorithm for trajectory estimation. In the data association step, author adopted the HFRF framework, which has seen success in the joint object recognition and segmentation task.

In [5] Ahmed Elliethy proposes an Expectation Maximization (EM) framework for registering a vector road network to a WAMI aerial image frame using vehicle detections. The method is based on the intuitive synergy between the problems of registering of a (vector) roadmap to an image frame and the detection of on road vehicles in an image. The detection of on-road vehicles in an image allows us to register the image to a vector road map by aligning the detection locations with the roads.

In [6] Junlin Hu proposed a deep metric learning (DML) approach for robust visual tracking under the particle filter framework. The DML tracker can explicitly learn several hierarchical nonlinear transformations to map data points into another subspace via feed-forward neural network architecture so that these nonlinear transformations are explicitly solved by maximizing the interclass variations of negative pairs and minimizing the intra-class variations of positive pairs simultaneously. It overcomes both the nonlinearity and scalability problems of conventional metric learning methods and kernel based method.

In [7] Dong Wang present a robust tracking algorithm based on linear regression. Also introduce a linear regression method and least soft-threshold squares (LSS) method is introduced. The observation likelihood of each candidate is computed based on the LSS distance and is improved by introducing negative templates.

In [8] Bo Ma proposes the structural local sparse descriptor to represent the target region. A target represented by using the collection of local sparse descriptors, where each descriptor represents partial appearance of the target. For background clutter, the proposed tracker considers the background information to train our classifiers, which ensures the good separating capacity of distinguishing the target from the cluttered background.

In [9] Bohan Zhuang proposed a reversed multitask sparse tracking framework which projects the templates matrix (both positive and negative templates) into the candidates space. By selecting and weighting the discriminative sparse coefficients, the DSS map and pooling method lead to the best candidate. With this DSS map, candidates are evaluated in both directions: not only how similar it is to the target object but also how different it is from the background.

In [10] Nan Jiang proposes a nonparametric data-driven local metric adjustment method. It finds a spatially adaptive metric that exhibits different properties at different locations in the feature space, due to the differences of the data distribution in a local neighborhood. It minimizes the deviation of the empirical misclassification probability to obtain the optimal metric such that the asymptotic error as if using an infinite set of training samples can be approximated.

IV. EXISTING METHODOLOGIES

The performance depends on which technique use, amount of data required by technique, time and complexity parameters. Many method are proposed for object tracking and detection.

Object tracking and detection techniques	**Performance Evaluations**
Fractional order gain Kalman filter	In FOGKF a feedback loop is inserted to avoid the divergence. To calculate the modified Kalman gain three steps are used : The steady state Kalman gain of standard Kalman filter is calculated. The fractional derivative of previous Kalman gain is computed. The modified Kalman gain is the sum of the Kalman gain and the mean of the fractional derivative of previous Kalman gains. Kalman gain (K) is computed such that the cost function and given as $K_{new.}$ $K_{\mathrm{new}} = K + E\left\{ \sum_{\mathrm{j}=0}^{k} (-1)^{\mathrm{j}+1} \binom{\alpha}{j} K_{\mathrm{k}-\mathrm{j}} \right\}.$ The modified Kalman gain contains two terms. The first term represents gain of Kalman filter. In the second term, mean of fractional difference of previous values of Kalman gain is calculated. The $(-1)^{j+1}$ gives alternative positive and negative terms that make the value of the mean to be nominal. Fractional derivative of previous Kalman gain incorporates the variations of input signal with time [1].
DML Tracker	**Initialization:-** It is important to initialize parameters $W^{(k)}$ and $b^{(k)}$, $1 \leq k \leq K$ in our designed network for obtaining good performance. The weight $W^{(k)}$ at each layer is given by a uniform distribution as $W^{(k)} \sim U\left[-\frac{\sqrt{6}}{\sqrt{r^{(k)} + r^{(k-1)}}}, \frac{\sqrt{6}}{\sqrt{r^{(k)} + r^{(k-1)}}} \right]$ **Template Update:-** The template m_t is updated from several observation vectors (e.g., $y^o_t, \ldots, y^o_{t-2}, y^o_{t-1}, y^o_t$) following the mean update of the IVT as

	$$\mathbf{m}_t = \frac{\theta t'}{\theta t' + \tau}\mathbf{m}_{t'} + \frac{\tau}{\theta t' + \tau}\sum_{i=0}^{\tau-1}\mathbf{y}_{t-i}^{o}$$ Where m_t is the template updated at the i^{th} frame. **Online Tracking process :-** When a new frame arrives, first randomly sample N particles around the current state of the tracked object according to a zero-mean Gaussian distribution under the particle filter framework. Then, the confidence of each particle is calculated at the top layer of the deep network. The particle with the highest confidence to object template is selected as the tracked object. Having localized the target object, sample positive and negative pairs to learn a set of nonlinear transformations using the proposed DML method [6].
Structural local sparse descriptor.	The number of templates is 8, training samples and candidate targets are all normalized to 32x32 pixels, and then 9 overlapped local patches with the size of 16x16 are extracted within the region with 8 pixels as step length. When constructing descriptors with local sparse codes, select 3 from 9 local sparse codes to perform average pooling, which results in 84 local descriptors. For learning the classifier,collect Np = 72 (9 positive samples per frame and 8 consecutive frames) positive samples and Nq = 150 negative samples. Choose 2/3 samples randomly from all training samples to get 84 weak classifiers, and select the best weak classifier [8].

Table 1: Performance of Existing Methods

VI. PROPOSED METHODOLOGIES

In order to detect and track moving objects based on motion, we consider amount of time taken by proposed method, number of moving objects detected, remove noise from frames and prediction of current pixel in next frame as an issues. Following issues consider for proposed method

1. **Execution Time:** It considers the processing time for Motion analysisTechniques..
2. **Detection and Tracking:** Detection of moving objects and track moving objects until it is present in the current frame based on motion are considered.
3. **Label on Moving Objects:** Once a moving objects is detected then label on that moving objects i.e. label on clusters of pixels.
4. **Prediction:** Predict the object locations in the next frame.

Tracking is the process of locating a moving object or multiple objects over time in a Real time video. Tracking an object is not the same as object detection. Object detection is the process of locating an object of interest in a single frame. Tracking associates detections of an object across multiple frames. Tracking multiple objects requires detection, prediction, and data association.

1. **Detection:** Selecting the right approach for detecting objects of interest depends on what you want to track and whether the camera is stationary.
2. **Prediction:** To track an object over time means that you must predict its location in the next frame. The simplest method of prediction is to assume that the object will be near its last known location. In other words, the previous detection serves as the next prediction. This method is especially effective for high frame rates. However, using this prediction method can fail when objects move at varying speeds, or when the frame rate is low relative to the speed of the object in motion.
3. **Data Association:** Data association is the process of associating detections corresponding to the same physical object across frames. The temporal history of a particular object consists of multiple detections, and is called a *track*. A track representation can include the entire history of the previous locations of the object. Alternatively, it can consist only of the object's last known location and its current velocity.

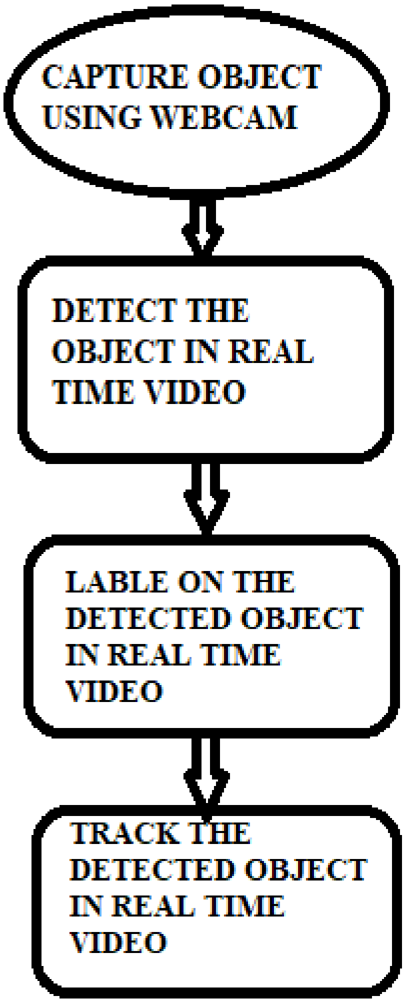

Figure 1: Block diagram of Proposed Methodology

In above figure, for detection and tracking of object in real time video, first step is object is capture by webcam. Webcam is used to capture image continuously and computer processes the image and shows the path of object in monitor. Cameras are imaging devices which deliver an abundance of information of which only a fraction may be needed for a specific control task. The camera output needs to be preprocessed to extract the relevant information for the tracking problem, e.g., the position of an object to be followed.

Secoand step is detect the captured object in a real time video.then label the detected object and match the keypoint of the capture object and detected object in a real time video.And track the object in real time video.