Object Tracking using HOG and SVM

Siji Joseph[#1], Arun Pradeep[#2]

Electronics and Communication Engineering

Axis College of Engineering and Technology, Ambanoly, Thrissur, India

Abstract — *Object detection and tracking is a vital task among computer vision researchers. The main objective of object detection and tracking is to establish correspondence of objects and object parts between consecutive frames of video. It became difficult since the object faces some problems usually shape deformation, occlusion, scale change, drift etc. There are various algorithm used for tracking. In this project, tracking is performed by using HOG and SVM classifier. The Histogram of Oriented Gradients (HOG) is a feature/image descriptor used in image Processing and other visual areas for the purpose of object detection. The HOG is capable (Histogram of Oriented Gradients) to distinguish the target and the background with HOG visualization and the technique counts occurrences of gradient orientation in localized portions of an image. Support Vector Machine (SVM) is the classifier used for classification. SVM with a single kernel is used in this project. The radial basis function kernel (RBF kernel) SVM are used for training and tracking. The HOG and SVM combination makes our system more efficient. And also we use RGB histogram and SIFT to describe the image.*

Keywords — *Object tracking, HOG, SVM, RBF kernel, SIFT, RGB histogram .*

I. Introduction

Object detection and object tracking is one of the most important tasks in many applications. Applications like video indexing, sport competition, traffic monitoring, human-machine interaction, security, video surveillance, motion recognition etc. make use of object tracking. During object tracking the tracker may face several problems like pose change, scale variation, drift, partial occlusion etc. There are many algorithms used for this purpose. Some of them are very fruitful in specific fields. The successful tracking algorithms will rectify these problems efficiently. Superpixel tracking [1] presents discriminative appearance model based on superpixels in which the target and the background are distinguished with midlevel cues. It is effective in occlusion handling, drift, pose variation and scale change. But it has the disadvantage of low tracking speed. Template matching presents appearance model. Intensity, statistical feature distribution, and low-dimensional subspace representations are the basis of this method. This method is effective in object tracking but they are less effective in handling large appearance change and drift error .it is not suitable for tracking complex scenes. The visual tracking decomposition (VTD) approach uses the conventional particle filter framework extended with multiple motion and observation models. It is effective for handling appearance variation caused by pose change, lighting, scale variation and occlusion. The disadvantage of this method is the tracker cannot distinguish target and background patches of the entire frame. The fragment-based (Frag) tracking handles partial occlusion using histograms of local patches. In this method combination of matching local patches using a template is carried out. The template is not updated in this algorithm so it is incapable of handle appearance change due to large scale variation and shape deformation. The PROST method tracking-by-detection framework is extended with the multiple modules.it is used for reducing drift errors. The advantage of this method is, it can handle certain amount of drifts and shape deformation, but it is not able to handle targets undergoing non-rigid motion or large pose change.

II. Proposed Method

Object tracking in complex scenes effectively and efficiently using HOG and SVM is used in this project. The HOG and SVM combination makes tracking more efficient and simple. Here, three features are used namely, HOG, SIFT and RGB. These feature descriptors describes the object. Also Support Vector Machine (SVM) classifier is used for classification which makes the tracking faster.

A. Data Set

This project aims to associate object recognition and object tracking. The object tracking task deals with the task of identifying the object with respect to an offline database which contain description of different objects and track the upcoming image sequence. Here, images that having structural similarities are used. This project makes use of offline images which having such similar structural information from existing data sets and use it for online object tracking.

B. Object Detection

First step is object detection. In this detect the target object which is going to track. The normal detection is performed by warping of the images. Initially draw a box that fits to target object and initialize its state as the function of initial position, size and rotation angle. This process is manually done. The initial state can be represented by

$$x^1 = (p_x^1, p_y^1, s^1, \theta^1) \quad (1)$$

For increasing the training samples crop out a group of images for collecting positive samples and crop out a group of negative samples. Positive samples are used to distinguishing and tracking of target object and negative samples are used to background detection and occlusion detection. The positive samples should be greater than 0 that is $X^+ \geq 1$ and negative samples should be less than 0, that is $X^+ \leq 1$.

C. Image Representation and Feature Extraction

SVM is one of the best classifier used in pattern classification and image classification. Its main objective is to separate of a set of training images or data into two different classes, (x_1, y_1), (x_2, y_2), ..., (x_n, y_n) be the data where x_i in Rd, d-dimensional feature space, and y_i in {-1,+1}, the class label, with i=1….n. SVM model builds the optimal separating hyperplanes . Its operation is depended on a kernel function (K). All images or data , of which feature vector lies on one side of the hyper plane, are belong to class -1 and the others are belong to class +1,that is the image is belonging to target or background.

In SVM classifier, for tracking applications the data can be represented using kernels. Four kernels are used for mapping the input space to the kernel space. They are linear kernel, polynomial kernel, RBF kernel and sigmoid kernel. Here, RBF kernel is used for object tracking. For any object there are many features [2] or interesting points on the object, they can be extracted for providing description of the feature of the object. This description may helpful when attempting to track the target object in an image containing many other objects. Here for extracting features from the collected samples 3 feature descriptors are used. They are RGB, HOG and SIFT feature descriptors. The RGB histogram is a mixture of three histograms based on the RGB color space's R, G, and B channels, that is one histogram contain 3 color channels. It has no invariance related properties. HOG descriptors give the number of occurrences of gradient orientation in the image [3],[4]. Scale-invariant feature transform (or SIFT) is an algorithm, it is used for detecting and describing local features in an images.it is used mainly for handle the scale variation. Then the extracted features are given to kernel, in this method a single kernel SVM is used. Here, RBF kernel is used for classification.

D. Support Vector Machines

Support Vector Machines (SVMs) which are the most commonly used classifier [5], [6] in learning systems based on kernel methods. It is mainly used for classification [7] and regression. The kernel method is unsupervised learning method. The kernel SVM is a representation of data amounts in which a nonlinear projection of data into a high-dimensional space where it is easier to separate the two categories of data of image.

The radial basis function kernel, or RBF kernel is also known as Gaussian kernel, is a popular kernel function used in various kernelized learning algorithms for variety of applications [8]. It is widely used in support vector machine classifier for classification of data. The RBF kernel on two samples x and x', can be represented as feature vectors in some input space, and it is defined as

$$K(x, x') = e^{-\frac{\|x-x'\|^2}{2\sigma^2}} \quad (2)$$

$\| x - x' \|^2$ is the squared Euclidean distance between the two feature vectors and σ is a free parameter.

A parameter γ,

$$\gamma = \frac{1}{2\sigma^2} \quad (3)$$

The RBF kernel value decreases with distance and it ranges between zero and one. Kernel value equal to zero when it is in the limit and equal to one when x = x'.

Then the feature space of the kernel has an infinite number of dimensions i.e ; for σ=1 its expansion is

$$K(x, x') = e^{-\gamma \omega t} \| x - x' \|^2 = \sum_{j=0}^{\infty} \frac{(x^T x')^j}{j!} e^{(-\frac{1}{2}\|x'\|^2} e^{(-\frac{1}{2}\|x\|^2} \quad (4)$$

$$K(x, x') = \sum_{j=0}^{\infty} \sum_{n_i = j} \frac{(x^T x')^j}{j!} e^{(-\frac{1}{2}\|x\|^2} \frac{x_1^{n_1} \ldots \ldots x_k^{n_k}}{\sqrt{n_1! \ldots \ldots n_k!}} * e^{(-\frac{1}{2}\|x'\|^2} \frac{x_1'^{n_1} \ldots \ldots x_k'^{n_k}}{\sqrt{n_1! \ldots \ldots n_k!}} \quad (5)$$

E. Object Tracking

The object tracking done in mainly two phase. They are, training phase and tracking phase. In training stage, the learning of initial data samples is done. In the tracking stage, the newly arrived data is updated using the same learning as used in the training stage.

1) *Algorithm - Main Tracking Structure*

Input: data set, different features and kernel function.
Output: Tracking result of each frame

1. Input the initial frame.
2. Detect the target object using normal detection method like warping of images.
3. Obtain the initial position and initialize the tracking parameters.
4. Extract the positive and negative samples from the detected stage
5. Extract the features. HOG, SIFT and RGB histogram from the initial stage are extracted.
6. Train the target frame for SVM classifier.
7. For each new frame extract positive and negative samples.
8. Feed the samples to SVM classifier.
9. Use SVM output to generate the best classification result for each sample.
10. Obtain the tracking result by using SVM output. The positive samples are used for tracking and collect the remaining negative samples from the frame for track under occlusion.

III.RESULTS AND DISCUSSION

Object tracking using different data sets is done in this project. During the tracking process, many problems are faced. Those are given bellow:

- Illumination variation,
- Occlusion, scale change,
- Fast motion,
- Motion blur,
- Out-of-plane rotation,
- Out-of-view and
- Background clutters.

In this project the first phase is performed by Superpixel tracking. This method is effective in occlusion handling and other visual variations like drift, pose change etc. but its main disadvantage is low tracking speed. So for better performance, object tracking using HOG and SVM [9], [10] is used. Along with this for better performance SIFT and RGB histogram feature [10] descriptors are used. So this method can effectively solve the tracking problem as discussed above.

A. Object tracking

The target object can be efficiently tracked by using this method. Initially, the object of interest is determined and trained the initial frames using a single kernel SVM namely, RBF kernel SVM.

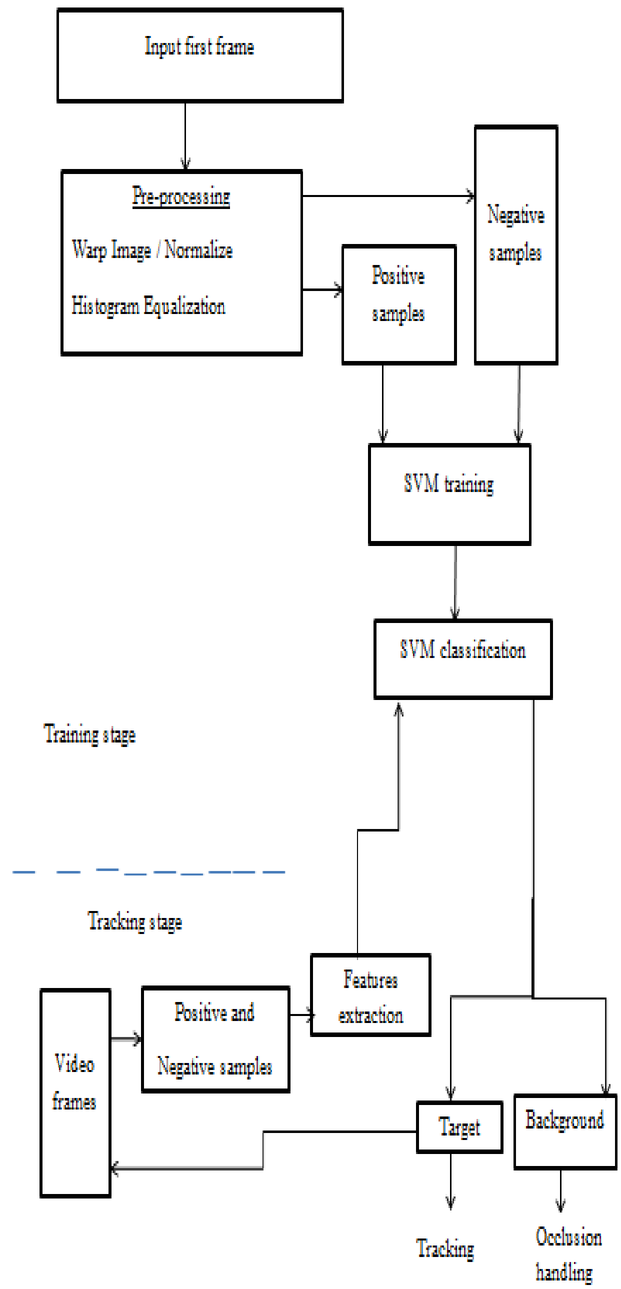

Fig.1: Block diagram of object tracking using HOG and SVM

Using the previous data obtained from training stage the new frames are updated. The movement of target in each frame is predicted in the tracking stage.

B. Occlusion Handling

As discussed earlier, one of main problem in object tracking is occlusion. Object tracking using HOG and SVM can effectively handle heavy occlusion. It can handle severe occlusion as shown in fig .3. It can track up to 310 frames of tom1 data set.

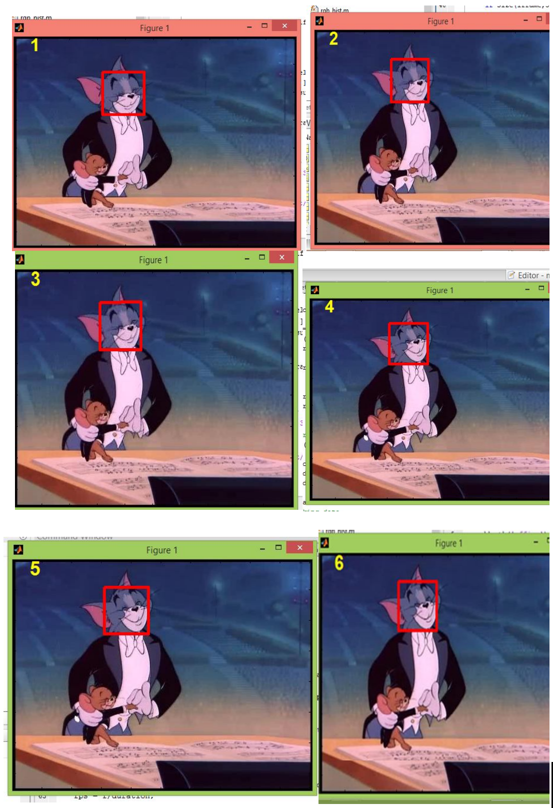

Fig.2: First six output frames of object tracking using HOG and SVM

C. Tracking speed

Now, conduct the tracking using SPT and our proposed method. The object tracking using HOG and SVM is very faster than tracking using superpixels. Tom1 data set was given for both tracking method. The total time taken for the object tracking is shown in the table I.

D. Quantitative Evaluation

Track the target object of the same image set using both methods. Here evaluation is done by graphical method for which a fixed number of frames are taken and then the time taken for tracking of each frame.

From the studies, it is clear that the object tracking using HOG and SVM is offers high tracking speed. In fig. 4, a graph superpixel tracking and object tracking using HOG and SVM From the studies, it is clear that the object tracking using HOG and SVM is offers high tracking speed. In fig. 4, a graph superpixel tracking and object tracking using HOG and SVM is plotted.

Fig.3: Effective occlusion handling in object tracking using HOG and SVM

TABLE I

TRACKING SPEED OF TRACKING USING SPT AND PROS POSED METHOD

Method used for object tracking	**Total time elapsed for tracking in seconds.**
Object tracking using HOG and SVM	714.0668 seconds
Superpixel object tracking	5672.456782 seconds.

E. Error Rate and Target Detection Success Rate

In this project, two parameters are used to measure the success of project. The first parameter is the Error Rate (ER) which is defined as number of false detection in the frame is dividing by the total number of frames.

$$\text{Error Rate (ER)} = \frac{Number\ of\ false\ detection}{Total\ number\ of\ frames} * 100\ \% \qquad (6)$$

Here, number of false detection means those frames that are tracked by during the execution of the program. The total number of frames is the summation of total frames used in the object tracking.

For our project,
Error Rate (ER) = 2.9 %

Second parameter Object Detection Success Rate (ODSR) .It is defined as the number of frames that tracks the target faces.

$$\text{Object Detection Success Rate} = \frac{\textit{Number of frames that detect target region}}{\textit{Total number of frames}} *100\% \quad (7)$$

For our project,
Object Detection Success Rate (ODSR) = 97%

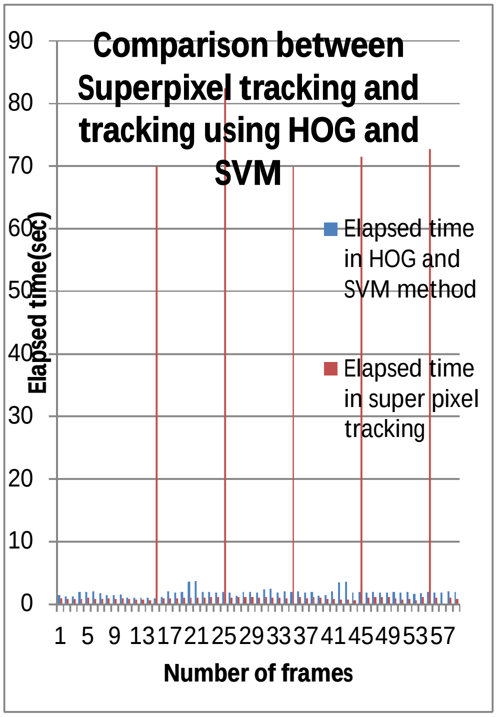

Fig.4: Time duration comparison between Superpixel tracking and tracking using HOG and SVM.

IV. CONCLUSION AND FUTURE WORK

In this project, an efficient and effective object tracking is done. It may track object in complex scenes by creating an appearance model based on histogram of oriented gradients along with single kernel algorithm. The HOG-SVM combination makes efficient and simple algorithm. It detects target objects even in visual difficulties. Here RGB histogram, SIFT feature descriptors also used for describing the image. The training is performed in the first phase and update scheme is also included to account for appearance variation of target object during tracking. Experimental results show that this tracking method performs more effectively and efficiently comparing to the other methods in handling occlusion, rapid motion, harder background, scale changes and it is faster comparing to other methods. Tracking accuracy and speed of this method is better than the other methods.

In this project, the images for tracking are taken from database (from internet).in future work, Real time images can be used for object tracking. Here single RBF kernel is used as classifier. Tracking using different kernels for classification using Support Vector Machine can also be used further. Also, for classification we can use other classifiers like ANN, ELM etc.

REFERENCES

[1] Fan Yang, Huchuan Lu, and Ming-Hsuan Yang, Robust Superpixel Tracking, IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 23, NO. 4, page 1639- 1651, 2014.

[2] Navneet Dalal and Bill Triggs,Leslie Lamport, Histograms of Oriented Gradients for Human Detection, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR, vol.1, pp. 886-893, 2005.

[3] Miss. A. Sanofer Nisha and Mrs. K. Thulasimani, Part-Based Pedestrian Detection And Tracking Using HOG-SVM Classification, International Journal of Computer Science and Mobile Applications, Vol.2 Issue. 1, pg. 142-155, 2014.

[4] Robert T. Collins and Yanxi Liu, On-Line Selection of Discriminative Tracking Features,in Proc. IEEE Int. Conf. Comput.Vis, pp. 346352 2nd edition,2003.

[5] Kwon and K. M. Lee, Visual Tracking Decompositionm, in Proc. IEEE Conf. Comput. Vis. Pattern Recognition,pp. 12691276 , 2010.

[6] Avidan S, Support Vector Tracking,In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 184191, 2001.

[7] Reecha P. Yadav, Vinu chackravarthy and Sunita P. Ugale, Implementation of Robust HOG-SVM based Pedestrian Classification,International Journal of Computer Applications (0975 8887) ,Volume 114 No. 19,pp 10-16, 2015.

[8] Er. Navjot Kaur and Er. Yadwinder Kaur , Object classification Techniques using Machine Learning Mode, International Journal of Computer Trends and Technology (IJCTT) Volume 18 Number 4, 2014.

[9] Deepthi V.K.P and Mr. Mohammed Anvar P.K Visual Tracking Using HOG and SVM, International Journal of Advanced Engineering Research and Technology (IJAERT),Volume 4 Issue 5,2016.

[10] Durgesh K. Srivastava and Lekha Bhambhu, Data Classification Using Support Vector Machin, Journal of Theoretical and Applied Information Technology,pp 1-7.