RESEARCH ARTICLE OPEN ACCESS

Real Time Implementation of Object Tracking and Classification: A Review

Amit S. Kakad, Santosh S. Mhaske
Dept of Electronics & Telecommunication, MGI-COET, Shegaon, India
ameetkakad@gmail.com
Dept of Electronics & Telecommunication, MGI-COET,Shegaon, India

Abstract—
The goal of this paper is to review the progress of real time object detection and tracking for security systems and video surveillance applications. Here in one approach tracking is achieved through background subtraction and in other approach edges are detected and than tracking is achieved using Kalman filter. Also an overview of classical steps in video classification would be discussed for video surveillance.
Keywords—Object Tracking; kalman filter; video classification.

I. Introduction

Object tracking refers to finding the location of an object when it moves from frame to frame. In case of tracking, object could be anything like vehicles on road, plane in sky or person walking on road. It is important to develop a real time working algorithm. No of tracking algorithms have been developed [1][2]. Particle filter[4] is also a powerful method to identify the target. Particle filter is not suitable for the real time implementation as it is computationally expensive.

In tracking two problems are considered[3]: prediction and correction. Predict problem deals with predicting the location of an object being tracked in the next frame, that is identify a region in which probability of finding the object is high whereas correction problem deals with identifying the object in the next frame within designated region. A well known solution for prediction is Kalman filter[6], a recursive estimator of state of a dynamic system. To predict search region more effectively, mean shift is combined with kalman filter[5].

II. Tracking Methodologies

2.1.Background Estimation

The proposed method uses both region based and feature based tracking algorithms for tracking cars. For prediction step we use region based and predict the position of object in frame at time t+1. We consider the center of the object as a point. Then we use Kalman filter in frame at time t for predict the position of this point in frame at time t+1. then we use the distance and color feature for correction in frame t+1, we match the color of vehicle and check its distance. And we find if it predicts correctly or not. If it has errors, this step corrects the errors.

First we should detect vehicles, we use an adaptive background model.In this section first we describe an adaptive background generation and moving objects detection.Then the tracking algorithm is described.Notice that in this article camera position assumed stationary. At first, background frame B0(x, y) is initialized by the first frame of image sequence.

$$B_0(x,y) = I_0(x,y)$$

For determined changes in frame mask $M_n(x,y)$ is defined by thresholding the difference between three consecutive frames.

$$M_{n+1}(x,y) = \begin{cases} 0 & \text{f } |I_{n+1}(x,y) - I_n(x,y)| < 1 \text{ or } |I_{n+1}(x,y) - I_{n-1}(x,y)| < 1, \\ 1 & \text{Otherwise} \end{cases}$$

2.2.Object Detection

The most used approach to detect objects in a video sequence is background subtraction. There are situations for which a poor implementation of the scheme causes an erroneous/coarse segmentation, so that consequent processes of tracking and recognition often fail. An automatic approach preserves from such inconvenience. We assume B (x, y) , I (x, y) the gray-valued frame and the corresponding adaptive background model. A binary mask D (x, y) n is performed to segment out objects of interest (foreground detection)

$$D_n(x,y) = \begin{cases} 0 & \text{if } S_n(x,y) < Th \\ 1 & \text{Otherwise} \end{cases}$$

Where $S_n(x, y)$ and then compute whit following relations:

$$S_n(x, y) = | I_n(x, y) - B_n(x, y) |$$

$$MED = \text{median}[S_n(x, y)] \qquad \forall (x, y) \in I_n ,$$

$$MAD = \text{median}[S_n(x, y) - MED] \qquad \forall (x, y) \in I_n$$

2.3. **Kalman Filter**

Kalman filter is a probabilistic prediction method for object tracking. Main problems that can be solved by using kalman filter in tracking is

- The object can be tracked if it moves beyond the searched region

- Variation factors such as lighting and occlusion which effect the appearance of target.

The Kalman filter estimates a process by using a form of feedback control which means the filter estimates the process state at some time and then obtains feedback in the form of (noisy) measurements. The equations associated with Kalman filter fall into two groups: time update equations and
measurement update equations.
The time update equations point towards the current state estimate and error covariance prediction is used to obtain the main estimates for the next time step. Then measurement update step is responsible for the feedback by calculating thekalman gain. The time update equations is associated withprediction, while the measurement update equations is associated with correction equations.

*2.3.1.***Time Update**

In the case of kalman filters each iteration begins with
predicting the process's state using a linear dynamics model.

A) State Prediction: For each time step T, a Kalman filter first
makes a prediction X_t and it is given by

$$X_t = AX_{t-1} + Bu_t$$

xk -1 is process state at time t-1, A is process transition matrix
and B is control vector. Ut which converts control vector into state space. In our model of moving objects on webcam images, state is a 4-dimensional vector [x; y; dx; dy] where x and y shows the coordinates of the object's centre and dx and dy shows its velocity.

B) **Error Covariance Prediction:** The Kalman filter concludes the time update steps by estimating the error covariance forward by one time step

$$P_t = AP_{t-1}A^T + Q$$

Pt-1 is a matrix representing error covariance in the state prediction at time t-1and Q is the process noise covariance. Lower the value of prediction error covariance Pt we can trust more on prediction of the state xt . If the process is precisely modeled, then the prediction error covariance will be low.

2.3.2 Measurement Update

By using the time update step kalman filter will predict the state xt -1 and find the error covariance at time k. then after during the measurement update steps kalman filter uses measurements to correct its prediction

A). Kalman Gain

Kalman filter computes a Kalman gain K_t, which is used for
correcting the state estimate xt

$$K_k = P_k H^T (H P_k H^T + R_k)^\wedge - 1$$

Where H is a matrix used for converting into measurement space from state space and R is measurement noise covariance.

B). State Update: Using Kalman gain $K_t$and measurements Z_t from time step t, where we can update the state estimate:

$$X_t = X_{t-1} + X_t(Z_t - HX_{t-1})$$

C). Error Covariance Update: The final step of the Kalman filter's iteration is to update the error covariance Pt.

$$Pt = (1-KtH)\ Pt-1$$

If the measurements are accurate then the updated error covariance will be decreased.

III. OBJECT CLASSIFICATION IN VIDEO SEQUENCES

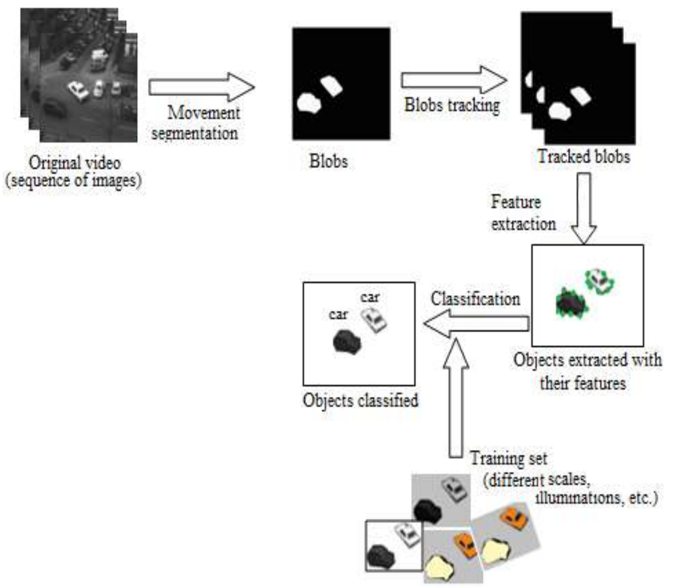

Figure 1. The mechanism of a classification system in a video sequence

3.1.Movement Segmentation

It identifies moving objects in a video sequence. It consists of detecting regions of interest rather than considering the whole image of sequence. This uses generally two types of information: temporal and spatial[7].

- Temporal method calculates absolute difference between two or three frames of sequence and applies a movement threshold to determine which regions have undergone a significant movement.
- Optical Flow is a visual displacement field which considers variations in motion as a displacement of points in image and assumes that color of pixel is independent of time.

3.2.Feature Extraction

Bileschi, and Wolf [8] divide features into two broad categories: histograms of oriented edges and patch based features.

A) Histograms of Oriented Edges

A Histogram of oriented edges is described as a weighted histogram wherein each histogram bin collects votes from gradients near particular locations and particular orientations.

The most popular example of this approach is the SIFT (Scale Invariant Feature Transform). Other examples are the geometric blur, histogram of oriented gradient (HOG) and Hu moments.

- Advantages: This representation is discriminative and tolerates a large number of transformations that the image can undergo (scale, translations, etc.).
- Disadvantages: Has sometimes a heavy computation.

B) Patch based Features

A patch based feature vector is an image description which depends on comparing the image with a set of stored image crops, also known as templates or fragments. Different implementations select different balances between invariance and the representation of geometric structure.

- Advantages: The use of multiple patches allows a better distinction between images.
- Disadvantages: Need to have a set of prototypes.

Many other features are present in the literature, such as symmetry, velocity, size and so on.

3.3. Object Classification

A) Support Vector Machine SVM

This algorithm was introduced by [9]. It solves the problem of binary classification (+1, -1) and aims to define a hyperplane of the formula:

$$y(x) = \text{sign}\ (w\times x + b)\ (1)$$

where w and b are the parameters of the decision surface, x is the vector of descriptors and y(x) the binary decision of the classifier. This hyperplane allows better separating the set of classes in the basis of support vectors by maximizing the margin between these vectors and the hyperplane. Decision bounds separating the data are defined by the kernel function:

$$k(x,y) = \exp(-c||x-y||2)\ (2)$$

Garcia-Pedrajas and Ortiz-Boyer proposed in 2006 an extension of the SVM classifier to resolve a multi classification problem. The approach was taken back in 2010 by Gurwicz and al. [10]. However, it is worth noting that the SVM classifier is the most widespread classification algorithm in the area of pattern classification and more precisely in object classification in images and videos. More information on SVM can be found on [11].

B) Bayesian Network

Proposed by Pearl [12], this classifier aims to predict, based on a training set, to which class belongs any new observation of an object. Bayesian network is a directed cyclic graph with a set of conditional probabilities P(Xi|Pai) where Pa is the parent set of the node X in the graph. The likelihood probability is given by:

$$P(X) = P(X1,\ldots, Xn)= \Pi\ P(Xi\ |\ Pai) \text{ with i from1 to n.}$$

And the object is assigned to the class with best likelihood probability.

C) K Nearest Neighbors

This algorithm is among the easiest algorithms and generally allows a comparative study of classification algorithms. The goal of this algorithms is to estimate the value of the unknown probability density function at a given point *x*. According to the *k*-nearest neighbor estimation technique, the following steps are performed:

- Choose a value for *k*.

- Find the distance between x and all training points x_i , $i = 1, 2, \ldots, N$. Any distance measure can be used (e.g., Euclidean, Mahalanobis, etc.). But Euclidean distance is usually used.
- Find the k-nearest points to x.
- Compute the volume $V(x)$ in which the k-nearest neighbors lie.

If the Euclidean distance is employed and the distance between the k-furthest neighbor and x is ρ, the volume $V(x)$ is equal to:

$V(x) = 2\rho$ in the 1-dimensional space
$V(x) = \pi\rho 2$ in the 2-dimensional space

IV. CONCLUSION

The goal of this paper was to give an overview of moving objects detection, tracking and classification. It discussed what is object tracking and how does it work. In one approach it detected the edges and then tracked it using kalman filter and on other hand object was detected using Background subtraction and then tracked using Kalman filter , also in complex situations it can be done using a particle filter. Finally more work based on different algorithm described in this paper can be combined for better and accurate tracking and classification.

References

[1] A.Yilmaz, O.Javed, and M.Shah, “Object tracking: A survey,” ACM Computing Surveys, vol. 38, no. 4, 2006.
[2] Sherin Cherian, C.Senthil Singh, “Real Time Implementation of Object Tracking through webcam”
[3] Amir Salarpour, Arezoo Salarpour, Mahmoud Fathi, and MirHossien Dezfoulian, “Vehicle Tracking using Kalman Filter and Features” Signal & Image Processing: (SIPIJ) Vol.2, No.2,June 2011.
[4] Yu Huang, Joan Llach, Chao Zhang, “A method of small object detecion and tracking based on Particle filters” Thomson Corporate Research Princeton, NJ08540, US.
[5] D. Comaniciu , v Ramesh, (2000) “ Mean shift and optimal prediction for efficient object tracking”, IEEE Int. Conf. Image Processing, vol. 3, pp. 70-73, Vancouver, Canada.
[6] M.Farrokhi, B.Tokaman, “ A Kalman Filter Based Methodfor real Time Visual Tracking of a Moving Object Using Pan and Tilt Platform,” International Journal of Scientific & Engineering Research Volume 3, Issue 8, August 2012
[7] Insaf Setitra, “Object Classificatiion in videos-An Overview” Journal of Automation and Control Engineering, Vol, 1, No. 2, June 2013
[8] S. Bileschi and L. Wolf, “Image representations beyond histograms of gradients: The role of Gestalt descriptors,” in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2007, pp. 1-8.
[9] V. N. Vapnik, *The Nature of statistical learning theory*, New York: Springer-Verlag, 1995.
[10] Y. Gurwicz, R. Yehezkel, and B. Lachover, “Multiclass Object Classification for Real-Time Video Surveillance Systems,” *Pattern Recognition Letters*, vol. 32, pp. 805-815. 2011.
[11] C. J. C. Burges, “A tutorial on support vector machines for pattern recognition,” *Data Mining and Knowledge Discovery*, 1998.
[12] J. Pearl, “Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference,” *Morgan-Kaufman*, 1988.