International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 05, May 2017
www.ijarse.com
IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

OBJECT DETECTION AND TRACKING FOR VIDEO SURVEILLANCE

Mr. S. S. Athatey[1], Prof. Y. M. Kurwade [2] ,Dr. V. M. Thakare[3]

[1,2,3]SGBAU, Amravati (India)

ABSTRACT

Object detection and tracking is important and challenging tasks in many computervision applications such as surveillance, vehicle navigation, and autonomous robot navigation. This paper is focused on five existing object detection and tracking techniques. Also this paper presents the moving object detection and tracking using reference Background Subtraction. In proposed method, first frame of video is directly consider as Reference Background Frame and this frame is subtract from current frame to detect moving object and then set threshold T value. If the pixel difference is greater than the set threshold T, then it determines that the pixels from moving object, otherwise, as the background pixels. But this fixed threshold suitable only for an ideal condition is not suitable for complex environment with lighting changes. So, this paperused dynamic optimization threshold method to obtain a more complete moving objects.

Index Terms: Background subtraction,expectation maximization, object detection, tracking, urban tracker, viola-Jones.

I. INTRODUCTION

Object detection and tracking is 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. Image processing and computer vision has a wide range of applications in autonomous vehicle guidance, traffic flow, traffic data collection and road traffic monitoring. Kalman filter is one of the most widely used tracking algorithms. Prior knowledge of initial state, process noise covariance, measurement noise covariance as well as initial state error covariance give enhanced performance [1]. Urban tracker is used for tracking road traffic and it uses the simple background subtraction method for object detection [2].V-J algorithm is also achieved impressive performance but V-J method is sensitive to object orientations; therefore, it can only work when the orientations of vehicles are known [3]. Hough forest random field is based on tracking-by-detection approach for multitarget tracking [4].An Expectation Maximization (EM) framework for registering a vector road network to a WAMI aerial image frame using vehicle detections.The wide area motion imagery (WAMI) sensor, which records high-resolution, full-motion video over multiple square kilometers from an airborne platform. WAMI's level of detail is such that all individual vehicles are clearly visible [5].

In this paper, first take overview of five existing methods for object tracking and detection and then proposed a methodology to detect and track the object from video scene.The propose methodology is based on background subtraction method, with the help of background subtraction first detect the object from the video frame. In

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

second step label that object with the help of chain code method. This labeling is useful while tracking the moving object.

II. BACKGROUND

Automatic tracking is one of the critical issues in real-time traffic flow monitoring traffic analysis system. Therefore,vehicle behavior analysis is an important aspect in trafficmonitoring.Kalman filter is one of the most widely used tracking algorithms.Fractional order gain Kalman filter (FOGKF). The main novelty of FOGKF is the gain is modified by adding fractional derivative of previous Kalman gain to the Kalman gain. A feedback loop is inserted in FOGKF and it uses fractional derivative of previous kalman gain as feedback [1].

A fully automatic multiple object tracker, called as Urban Tracker, that is adapted to track various a priori unknown road users. The propose 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 [2].

V-J scheme to improve the original V-J scheme so that the enhanced one will be insensitive to on-road vehicles' in-plane rotation and achieve better accuracy and efficiency. The basic idea is to directly detect the orientation of the road and rotate the road according to the detected orientation only once [3].

A novel tracking-by-detection framework for multitarget tracking.A multi-target tracking scheme consisting of HFRF-based data association step, followed by an auxiliary visual tracking method. The data association problem is formulated as one of inference in the HFRF model, in which the CRF learning and inference are unified within the HF computational framework. The modified RJMCMC algorithm is used to handle the mutual-occlusion [4].

An Expectation Maximization (EM) framework for registering a vector road network to a WAMI aerial image frame using vehicle detections.The wide area motion imagery (WAMI) sensor, which records high-resolution, full-motion video over multiple square kilometers from an airborne platform. WAMI's level of detail is such that all individual vehicles are clearly visible [5].

The rest of the paper is organized as follows. In this paper, **Section II** gives us background details, **Section III** provides work which is done previously, **Section IV** gives idea about existing technology, in **Section V** analysis and discussion about techniques is carried out, proposed methodology is explained in **Section VI**, Possible outcomes and Result is described in **Section VII**, **Section VIII** concludes the paper.Finally, **Section IX**described future scope of the paper.

III.PREVIOUS WORK DONE

In research literature HarpreetKaur et.al.(2016)[1] 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.

Jean-Philippe Jodoin et.al.(2016)[2]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

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

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.

YongzhengXu et.al.(2016)[3]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.

Jun Xiang *et.al.*(2016)[4]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.

Ahmed Elliethy *et.al.*(2016)[5]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.

IV. EXISTING METHODOLOGIES

A. Fractional order gain Kalmanfilter :-This methodis used to detecting and tracking objects from the video scene. The method is based on the original Kalman filter algorithm. In FOGKF a feedback loop is inserted to avoid the divergence. The modified Kalman gain is estimated by minimizing the cost function of the FOGKF. 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 $K_{new.}$

$$K_{\text{new}} = K + E\left\{ \sum_{\text{j}=0}^{k} (-1)^{\text{j}+1} \binom{\alpha}{j} K_{\text{k}-\text{j}} \right\}.$$

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.

B. Urban Tracker :-Urban tracker is fully automatic multiple object tracker, that is adapted to track various a priori unknown road users. At first, foreground blobs are extracted using the ViBe background subtraction method. These foreground blobs are then modelled using their size and position, as well as FREAK keypoints. After that, “low-level” tracking is performed to match blobs across two consecutive frames. These blob matches correspond to what are called short tracklets (s-tracklets). To handle merging, splitting and fragmentation, s-tracklets are analyzed and then assigned to object tracks. After which track properties are updated. The track state is used for s-tracklet assignment and represents the track life cycle in the scene. The FSM is essential to determine when and how an s-tracklet can be added to another track. This step uses the s-tracklet information to resolve the ambiguity between the assignments of s-tracklets to tracks. At the end the final tracks are outputted.

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

They may differ from the ones computed at each frame because they are adjusted when merging and splitting are discovered.

```
procedure URBAN TRACKER
    for each frame do
    Extract foreground blobs
    Compute model for each blob
    Match blobs with those from previous frame
  (s-tracklets construction)
    Assign s-tracklets to object tracks
    for each s-tracklet do
      if s-tracklet is assigned to object track then
          Update track model
            Update track state
        else
            Create new track in hypothesis state
        end if
      end for
  end for
  Return final object tracks
end procedure
```

Algorithm 1: Overview of Urban Tracker

C. Enhanced V-J scheme :-The implementation of enhanced V-J scheme was done on UAV video datsets. UAV data collection is done by using a quadcopter airborne platform and a Gopro Hero Black Edition 3 arial camera. A 3-axis gimbal is mounted on the UAV to stabilize the videos and eliminate video jitters caused by UAV therefore greatly reducing the impact from external factors, such as wind. In addition, an On-Screen Display (OSD), an image transmission module and a video monitor are installed in the system for data transmission and airborne flying status monitoring and control.The accuracy of the proposed road orientation estimation method was evaluated by two typical indicators, mean error (ME) and the root mean square (RMS), which are defined as:

$$ME = \sum_{i}^{n} |\varphi_{\text{err}}(i)|, \quad RMS = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\varphi_{\text{err}}(i))^2}$$

where the error of road orientation estimation is defined as: ϕ_{err} (i) = ϕ_{est}(i)−ϕ_{GT}(i).ϕ_{est}(i) is the estimated road orientation of the i^{-th} frame, using the proposed road orientation estimation method; ϕ_{GT}(i) is the ground truth of road orientation of the i^{-th} frame, measured manually. n is the number of UAV images used for measuring the accuracy of the proposed road orientation estimation method. Lastely vehicle detection is done by evaluating four performances are chosen to evaluate the detection accuracy including: Detection speed in terms of frames per second (f/s), Correctness (Cor.), Completeness (Com.), and Quality (Qua.), as defined as:

$$Cor. = \frac{TP}{TP+FP}, \; Com. = \frac{TP}{TP+FN}, \; Qua = \frac{TP}{TP+FP+FN}$$

where T P is the number of “true” detected vehicles; FP is the number of “false” detected objects which are nonvehicle objects; and FN is the number of vehicles missed.

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

D. Multitarget tracking using HFRF :-In these method The data association problem is formulated as one of inference in the HFRF model, in which the CRF learning and inference are unified within the HF computational framework. The modified RJMCMC algorithm to handle the mutual-occlusion reasoning more efficiently. For SW-cut inference initialize the labels of all nodes and edges using the Bayesian decision on the corresponding posteriors p_i and p_{ij}. By assuming that $\sum_{vi \in Head\ i1} Yi \leq 1$ and $\sum_{vi \in Tail\ i2} Yi \leq 1$ where $Head_{i1} = \{(T_{i1} \rightarrow T_j) \in V\}$ and $Tail_{i2} = \{(T_j \rightarrow T_{i2}) \in V\}$. The first constraint limits that any tracklet T_{i1} may be linked to at most one distinct tracklet, and the second constraint limits that at most one tracklet may be linked to any tracklet T_{i2}. Average error (AE) is defined in a similar way as MOTP $\mathrm{AE} = \frac{\sum_t d_t}{\sum_t c_t}$ where c_t is a binary variable indicating whether the tracking result is matched to the truth target at time t and$\sum_t ct$ denotes the number of matches. d_t is, for each match, the overlap ratio between tracking result and its corresponding true target.

E. Expectation Maximization (EM) framework :- 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. Author formulate the problem as the minimization of a joint probabilistic objective function that combines (a) the classification of vehicle detections as true on-road vehicles vs. other detections and (b) a penalty for misalignment between the putative on-road vehicle detections and the vector roadmap under a parametric transformation. An explicit algorithm for registration is then developed in an Expectation Maximization (EM) framework that alternates between estimation of posterior probabilities that individual detections of vehicles correspond to on-road vehicles and the minimization of the weighted sum of minimum squared Euclidean distances from detection locations to the corresponding nearest points on the network of roads, where the weights are the estimated posterior probabilities that the detections correspond to on-road vehicles.

V. ANALYSIS AND DISCUSSION

This section describes brief analysis of the above five method. The table shows the comparison between five existing methods and also shows advantages and disadvantages of five methods.

Object Detection and Tracking Technique	**Advantages**	**Disadvantages**
Fractional order gain Kalman filter	RMSE is improved upto 17% by using the. It performs better even in the presence of for abrupt variations in the inputs.	TheKalman gain of the present state depends upon the previous Kalman gain. If previous Kalman gain is not accurate then calculated new Kalmangain have errors and produces false results.
Urban Tracker	For urban tracking it requires no prior knowledge. The algorithm gives balanced performance.	It requires more computation time. The algorithm doesn't have any prior knowledge due to this you must set some parameters manually.
Enhanced V-J scheme	The enhanced V-J, which incorporates the road orientation adjustment method, is insensitive to road orientation	The Quality (88.60%) of our method is slightly lower than KCF (Quality, 89.59%).

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

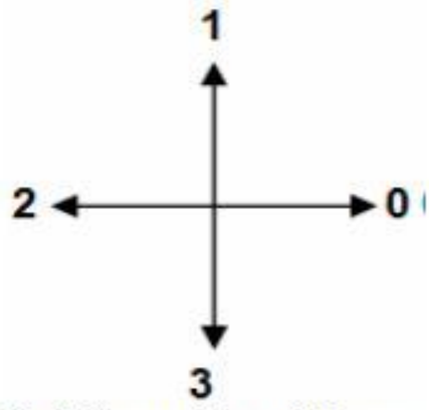
	changes. The enhanced V-J yields the best Completeness (92.49%) of all the seven methods. The proposed framework achieved an average Quality of 88.60% with an average tracking speed of 5.30 f/s. Compared with other methods.	An enhanced V-J vehicle detection method uses only low-altitude UAV images. The enhanced V-J method still has difficulties to address the illumination and turning problems
Multitargettracking using HFRF	The proposed method is useful to track multi-target at same time. It is best over other compared methods and also it gives better computation time than other algorithm. The proposed framework achieved an average Quality of 88.60% with an average tracking speed of 5.30 f/s. Compared with other methods.	For integrate tracklet association, explicit occlusion reasoning, and missed detec- tions recovering there no function is designed.
Expectation Maximization framework	Resolution optimized to detect moving vehicles and dismounts. Unlimited pan and digital zoom with in entire field of view. Ability to cue high-resolution full mot-ion video cameras.	The proposed method require high internet bandwidth to track the vehicles. The proposed method requires more resources and it is very costly in real time implementation.

Table 1: Comparisons between different Object Detection and tracking techniques.

VI. PROPOSED METHODOLOGY

A very simple method for representing the form and position of an object in an image consists in coding the edge information of the same object. The most widespread method for memorizing a list of points (with position but no colour information) is commonly known as “chain code". The basic idea is to go along the object contour and to code progressively the direction to be followed. Once identified the coordinates of a point in the object contour (generally the highest and the left-most), the successive point is identified only on the basis of the direction to be followed which links the barycentre of the pixels. The directions allowed are usually limited, so as to improve coding efficiency. Some of the most frequently observed conventions are shown in Figures 1 and 2.

International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 05, May 2017
www.ijarse.com
IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346



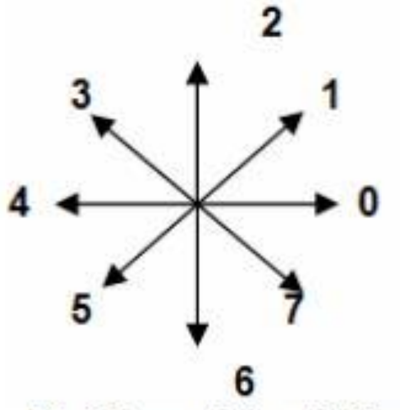


Fig 1. Connectivity neighbours   Fig 2 Connectivity neighbours

a) in case of 4-connectivity it is possible to proceed only horizontally or vertically;

b) in case of 8-connectivity it is possible to proceed diagonally, horizontally and vertically.

In the background subtraction method, consider that the whole scene from two parts, the background, the foreground. The background is a static scene and which can be seen, foreground is the moving objects which are interested in the video surveillance, such as vehicles, pedestrians etc. however due to the scene of the monitor changes over time, the foreground stagnation in the picture for a long time could be treated as part of the background, so updating of the reference image periodically is essential for moving object detection. Updating of reference image can be achieved through the frame difference method.The first step is to take input video from static cameras. For processing the video files, convert video into frames and from frames to images. Next step is take first frame as a Background frame and next is current frame and then apply subtraction operation. Background frame is subtracted from current frame. Then Threshold operation is performed and foreground object is detected. After object detected last step is track object in video.

VII. ALGORITHM ON OBJECT DETECTION AND TRACKING IN VIDEO AND LABELLING ON MOVING OBJECT

Step 1: First convert an input signal or video into video sequences.

Step 2: Convert input video into frames and the apply the background subtraction method to extract the foreground object.

Step3: Applythe proper threshold operation to finding the object in the frame.

Step 4: Once an moving object detected by background subtraction method apply an chain code method to labelled on moving object.

Step 5: After detecting and labelling the object next task is to track that object. Tracking can be done by using any object tracking technique (i.e. Kalman filter, Urban tracker, V-J algorithm).

Step 6: Labelling method on moving object will help in many applications, it will labelled on moving object in video with their estimate speed and direction on which object is moving.

International Journal of Advance Research in Science and Engineering
Vol. No.6, Issue No. 05, May 2017
www.ijarse.com
IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

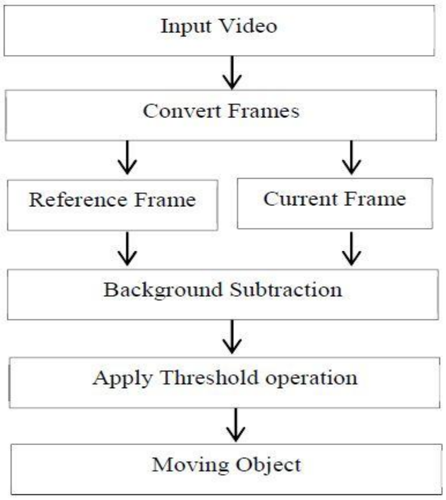


Fig 3.Flow Chart of Background Subtraction Method

VII. POSSIBLE OUTCOME AND RESULTS

To track moving objects labeling becomes as essential process. Object labeling becomes important because each object should be represented by a unique label with a condition that the object must preserve its label without any change. From the moment object enters the scene at frame F0 that is the starting frame till it leaves the scene at the end frame. The propose method will have successfully labeled on moving object that are detected by different motion analysis techniques and achieve satisfactory results.

VIII. CONCLUSION

This paper focused on the study of different moving object detection and tracking techniques i.e. Fractional order gain Kalman filter, Urban Tracker, V-J algorithm, Expectation Maximization (EM) framework. The performance of the different algorithms over the set of Video Sequences and Images can be extrapolated to the results obtained over real ones. However, the outputs of these methods can be detection of moving object in video sequences and labeled on the moving object, which provide the best performance according to the simplicity of the algorithm and the accuracy of the results. This paper proposed an algorithm which improves performance and accuracy of different object detection techniques by labeling on moving object that links in video. To track moving objects labeling becomes as essential process. Object labeling becomes important because each object should be represented by a unique label with a condition that the object must preserve its label without any change. From the moment object enters the scene at frame F_0 that is the starting frame till it leaves the scene at the end frame. Object detection in video sequences has several important applications such as object tracking, visual surveillance systems and dynamic three-dimensional scene analysis.

IX.FUTURE SCOPE

In this paper the proposed methodology only considers the ideal environmental conditions. Future study tries to overcome this issue.

IJARSE
ISSN (O) 2319 - 8354
ISSN (P) 2319 - 8346

REFERENCES

[1] HarpreetKaur and J. S. Sahambi, "Vehicle Tracking in Video Using Fractional Feedback Kalman Filter",IEEE Transaction on Computational Imaging, Vol. 2, No. 4, PP. 550-561, December 2016.

[2] Jean-Philippe Jodoin, Guillaume-AlexandreBilodeau and Nicolas Saunier,"Tracking All Road Users at Multimodal Urban Traffic Intersections", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 17, No. 11, PP. 3241-3251, November 2016.

[3] YongzhengXu, Guizhen Yu, Member, Xinkai Wu, Yunpeng Wang and Yalong Ma, "An Enhanced Viola–Jones Vehicle Detection Method From Unmanned Aerial Vehicles Imagery",IEEE Transaction on Intelligent Transportation Systems, Vol. , No. , PP. 1-12, October 2016.

[4] Jun Xiang, Nong Sang, JianhuaHou, Rui Huang and ChangxinGao, "Multi-target Tracking Using Hough Forest Random Field", IEEE Transaction on Circuits and Systems For Video Technology, Vol. 26, No. 11, PP. 2028-2042, November 2016.

[5] Ahmed Elliethy and Gaurav Sharma, "Automatic Registration of Wide Area Motion Imagery to Vector Road Maps by Exploiting Vehicle Detections", IEEE Transaction on Image Processing, Vol. 25, No. 11, PP. 5304-5315, November 2016.