

Real Time Unattended Object Detection and Tracking Using MATLAB

Sagar Sangale¹, Sandip Rahane²

P.G. Student, Department of Electronics Engineering, Amrutvahini College of Engineering, Sangamner,
Maharashtra, India¹

Assistant Professor, Department of Electronics Engineering, Amrutvahini College of Engineering, Sangamner,
Maharashtra, India²

ABSTRACT: A key technology to fight against terrorism and crime for public safety moving object detection and tracking is become very popular and one of the challenging research topic in various security areas of computer vision and video surveillance applications. To detect and track moving objects which is based on mean shift and particle filter is implemented using MATLAB 2012b. In this method, we proposed static webcam for capturing live scene as well as stored database of segmented videos. Issues related with object detection and tracking is jointly employed using mean shift and particle filters. Tracker will estimate the dynamic shape and random appearance of objects. Tracking requires location and shape of object in every segmented frame. Rectangular bounding box is used to improve the estimate of its shape and appearance. Target object may experience partial occlusions, intersection with other objects with similar colour distributions, abrupt motion speed changes, and cluttered background. Test result shows improvement in terms of robustness, tracking drifts, accuracy of tracked bounding box to occlusions.

KEYWORDS: Real time object tracking, Unattended object, anisotropic mean shift, object intersection, live learning of objects, partial occlusion, particle filters.

I. INTRODUCTION

To design smart computer vision based video surveillance systems which can contribute to the safety of people in the home and in public places such as, airports, railway stations, shopping malls and other public related places has now made it possible with the advent of smart multi mega pixel consumer cameras with higher processing capabilities. The critical threat of public safety due to terrorist attacks especially, explosive attacks with unattended packages are repeatedly concentrated on such public places. A key function in such a computer vision based video surveillance system is the understanding of human behaviour in relation with objects left unattended in public places. In this context, real time smart visual surveillance object detection and tracking systems for human behaviour understanding have drawn much attention of researchers and investigated worldwide as an active research topic [1]. Thus, a primary goal of video surveillance is to obtain a live description of what is happening in a monitored area and take (or trigger) appropriate action against pertained object and human misbehaviour. Moving object detection and tracking surveillance system provides crucial information about the object behaviour, interaction and the relationship between objects of interest to high level processing since automated video surveillance systems has modules from low-level such as object detection, object tracking, classification, event analysis, efficiency and robustness in each module are particularly important.

The first step in a video surveillance system is usually requires the extraction of foreground regions i.e. region of interest containing moving objects. Different strategies have been adopted, depending upon whether static or dynamic multi mega pixel cameras are being used. One method to extract foreground object region in each frame by using adaptive background subtraction is then followed by foreground object tracking of the scenes with stable background usually that applies to videos captured by static cameras [2]. Videos captured by moving cameras is another way to use

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 6, June 2016

a tight rectangular shape bounding box around each target object of interest and track these boxes through frames when the background is dynamic[3]. Tracking multiple objects in poor imaging conditions remains a challenging task, especially in case of partial occlusions, objects intersection; objects pose changes, deformations, abrupt motion speed changes, complex background or dynamic complex background. Tracking of isolated objects is relatively simpler than others. Visual features of individual object may become ambiguous and uncertain; e.g. when a scene contains several objects of intersect even fully by several nearby objects but distant object from the camera can be occluded partially due to drastically change in pose and shape of objects. During occlusion the motion speed and direction of objects may change due to that despite many research works tracking of multiple moving objects through complex scenarios remains a challenging task [1].

The computational complexity and efforts involved in simply following someone through an extended video sequence is enormous, and a truly reliable and robust tracker has yet to be developed. The main aim of our proposed MATLAB based system for automated analysis of unattended objects in relation with human behaviour in the surrounding area is explored. In real time unattended object detection and tracking system inter-relation with human suspicious behaviour requires a sufficiently high accuracy and the computational complexity. For such a system, we need to analyse not only the motion of people and objects, but also the posture of the person such as carrying bags, leaving the bags unattended, as the postures of the persons can provide important aspects for understanding of their motives and intentions. Hence, accurate detection and recognition of various targeted objects and human postures in relation with objects contribute to the scene understanding. An independent tracker is used to detect and track each individual object is one category of existing methods and by using multiple independent trackers multiple object tracking can be performed based on Gaussian mixture models. There some different methods have been employed which includes detection and tracking of moving objects such as Eigen objects appearance modelling, Kalman filtering approach, optical flow method, anisotropic mean shift and particle filters method etc.

In our proposed robust visual tracking method that jointly employs an-isotropic mean shift and particle filters, our main objective includes comparison of existing work with mean shift and particle filters such as: 1) The joint particle filters and mean shift tracking scheme, introduction to the full combination of functionalities and adjustable/flexible bounding box parameters; 2) Partitioning a rectangular bounding box, introducing and deriving multi-mode anisotropic mean shift; 3) For the reference object distribution; introducing live learning method of target and reference objects and 4) The method of computing bounding box parameters through decomposition of Eigen vectors for geometry of partitioned areas and weighted average of parameters of bounding box. Following are some of the main merits of the proposed robust scheme shows: a) Robustness in terms of long-term partial occlusions and intersections of objects; b) Efficient particle filter tracking by using a small number of particles; c) Robust mean shift through tunable parameters, anisotropic kernel and partitioning bounding boxes, tight tracked box due to accurate box parameter estimation.

II. LITERATURE SURVEY OF RELATED WORK

Real time object detection and tracking system first video is converted into number of consecutive frames then it will processes frames for locating moving objects from the scene. Real time object tracking is a challenging problem in the field of computer vision such as motion-based recognition, automated surveillance, traffic monitoring, and object based video compression etc. Mean shift has drawn much interest in object tracking by maximizing the Bhattacharyya coefficient between reference and the target objects. Early work was proposed by Comaniciu [3], Collins extended his work for mean shift by introducing kernel bandwidth normalization. It performs an extensive search of object within a range of bounding box scales and it is computationally intensive. Yilmaz proposed to use an asymmetric kernel for the mean shift for level setting. Using anisotropic mean shift method the centre, size, shape and orientation of the bounding box are simultaneously estimated during the tracking was proposed by Sumin in [6].

According to the object deformation Parameswaran proposed mean shift with a tuneable bandwidth. Xu proposed how to make the tracker more robust by computing the Bhattacharyya coefficient in case of joint spatial and color histograms of non-overlapping multi-scale regions. Image and video segmentation was proposed by Wang using an isotropic kernel mean shift, and Maggio used overlap sub-regions from an ellipse bounding box. All above proposed methods uses the anisotropic kernel for the mean shift. A major barrier of using particle filters for visual object tracking is its heavy computation, in which large number of particles are required to characterize a visual object as the

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 6, June 2016

result of a large size state vectors are used. Early work of particle filters have been proposed for visual tracking by Wang [5], by separating the state vectors into shape and appearance based (Eigen) sub-vectors thereby it reduces the number of particles required in the particle filter in which appearance is treated by linear models. Further improvement is made in visual object tracking by combining mean shift and particle filters is also called as the hybrid method which is integrated as single scheme. To track the human hands motion [6] by embedding the mean shift in particle filters where particles with higher weights from the mean shift are combined in the observation model, and it reduces the degeneracy and requires fewer particles than the conventional particle filter [4]. By applying the mean shift on particles with large weights is also as called elite particles to weight particles by using the observation model which is proposed by Zhong.

Wang [5] extended an adaptive size of bounding box for evaluation of particle weights by using zeroth moment of pixel's likelihood weight, and a unique proposal distribution. It is important to mention that none of the above joint mean shift and particle filter method contain a full combination of all rectangular bounding box parameters (i.e. width, height, central location and orientation) of as the automatic tuning parameters in tracking, even though some of these parameters are used in either mean shift, or particle filter methods. Motivated by this, we propose a novel joint particle filter and mean shift scheme for target video object tracking. In our method, which is an extension of our earlier work in [4], each rectangular bounding box is partitioned into sub-regions such as sectors and ellipse to allow multiple modes in the mean shift. Further, live learning of the reference object distribution is performed in different way of to formulate the estimation five parameters of the rectangular box such as shape parameters is given by using the result from the kernel bandwidth matrix estimates. In the proposed method, the particle filter estimates the pdf of the state vector containing the shape parameters of rectangular bounding box, whereas mean shift improves the estimation of to track the box by combining the object appearance within the bounding box.

III. REAL TIME MOVING VISUAL ABONDEND OBJECT DETECTION AND TRACKING USING COMBINED MULTI-MODE ANISOTROPIC MEAN SHIFT AND PARTICLE FILTERS

In this section we are describing the proposed tracking scheme using anisotropic mean shift embedded in particle filters. The basic idea of moving object tracking is to use the object appearance and its shape. The dynamic shape is estimated by the particle filter, and dynamic appearance is allocated by the mean shift. Further live learning approach is employed for updating the reference object distribution. The main objective of the tracker is to allocate the best region of interest i.e. object region (or bounding box) in each image frame is done by using particle filter that tracks the affine bounding box and weights to moving object box locations that are used as particle weights.

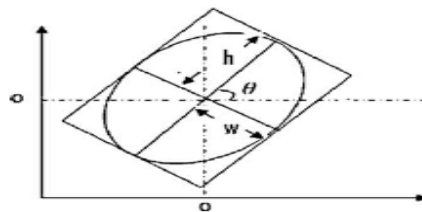


Fig.1. Definition of bounding box w.r.t to Width, Height and Orientation [4].

To compute the parameters from a partitioned box:

According to Cartesian coordinate system to calculate the bounding box orientation as shown in Fig.1. Let 'θ' be the angle between the long axis of kernel bandwidth 'Σ' and the horizontal axis. The height 'h' and width 'w' of the rectangular bounding box can be defined as in terms of its radii along the long and short axes of ellipse of the rectangular bounding box. Given the kernel bandwidth 'Σ', kernel bandwidth follows that the box parameters 'h', 'w', and 'θ' are related by

$$\Sigma = R^T(\theta) \begin{bmatrix} \left(\frac{h}{2}\right)^2 & 0 \\ 0 & \left(\frac{w}{2}\right)^2 \end{bmatrix} R(\theta) \quad \& \quad R = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \quad (1)$$

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 6, June 2016

Where ‘R’ is the rotation matrix, to estimate these parameters of bounding box, Eigen vector decomposition is applied to ‘Σ’ such as $\Sigma = VAV^{-1}$.

$$A = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad V = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix}$$

The orientation ‘θ’, height ‘h’ and width ‘w’ of the bounding box are obtained by,

$$\theta = \tan^{-1}\left(\frac{v_{21}}{v_{11}}\right), \quad h = 2\sqrt{\lambda_1}, \quad w = 2\sqrt{\lambda_2} \quad (2)$$

Where v_{11} & v_{21} be the two components from the largest eigenvector, also λ_1 & λ_2 are the two Eigen values.

The main novelties of the joint proposed scheme are as follows: 1) It uses set five parameter such as (width, height, central location and orientation) of rectangular box as a fully tunable variables; 2) By partitioning a rectangular box into sub-regions, deriving equations for multi-mode anisotropic mean shift; 3) An efficient approach for live learning of reference object distributions is employed; and 4) By relating the rectangular bounding box parameters with the mean shift can be estimated by applying Eigen decomposition, exploiting geometry of partitioned sub-regions, and using weighted average of parameters. The real time moving object detection and tracking is designed to reduce the tracking drifts in offline as well as real time live videos scenes to tackle the problems of single object tracking and to provide further tracking robustness in terms of a) Long-term partial occlusions (poor imaging conditions) and b) Intersections of objects. An efficient moving visual tracking system can be employed by using particle filters with a small number of particles, and live learning of reference object distribution.

IV. REAL TIME OBJECT TRACKING BY EMBEDDING MULTI-MODE MEAN SHIFT WITH PARTICLE FILTER

This section describes the moving object detection and particle filter tracking through embedding the multi-mode anisotropic mean shift in it. Let the state vector, or particles in the particle filter are defined as the shape of bounding box.

$$S_t = [y_t^1 \ y_t^2 \ w_t \ h_t \ \theta_t]^T \quad (3)$$

These particles are initially generated according to the Gaussian distribution in under the Brownian motion model [6]. The key connection between the joint particle filter and mean shift tracking is lies in these two methods, where the object appearance within the bounding box is taken into account in for the estimation of posterior box shape by the state vector. Therefore, the state vector of particle filter is simplified to, $X_t = S_t$ which has a small size. Roughly speaking, initial particles first pass through a multi-mode anisotropic mean shift, resulting in the improved particles then by moving the box toward locations where the object of interest appears most similar to the reference object. The posterior estimates of rectangular bounding box for these improved particles are then used by the particle filter.

V. ALGORITHM OF THE PROPOSED JOINT TRACKING SCHEME

System Initialization: Choose an object region (for frame $t=0$), compute q_0 for the reference object, and compute $\alpha_i, m_i, i = 1 \dots \dots M$ for the given partition. **For** frame $t=0, 1, \dots$, do:

- i) Generate particles $S_{t+1}^{(j)} \sim P(S_{t+1} | S_t^{(j)})$, set weights $w_t^{(j)} = \frac{1}{p}$, **For** Particles $j=1 \dots \dots Np$ do:
- ii) Mean shift iterations for $y_{t+1}^{(j)}$ using Eq. (10) from the reference [4].
- iii) Iteratively compute: $\sum_{i,t+1}^{(j)}, i = 1 \dots \dots M$ Using Eq. (11) from the reference [4], until converge;

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 6, June 2016

- iv) Compute $w_{t+1}^{(j)}, h_{t+1}^{(j)}, \theta_{t+1}^{(j)}$ by Eq. (14) and (15) from the reference [4].
- v) Update $S_{t+1}^{(j)}$ in Eq. (17) of the reference [4].
- vi) Update the particle weight by w_{t+1}^j using Eq. (17) and (18) of the reference [4].

END {j}

- i) Compute the posterior estimate of the state vector s_{t+1} using Eq. (20) of the reference [4].
- ii) Resampling...
- iii) If 't' is the end boundary of the interval sized L (i.e. mod (t, L) = 0) then live learning of q_i^j using Eq. (21), (23) of the reference [4] is then satisfied.

END {t}.

A. Matlab Implementation Of Moving Object Detection And Tracking Using Static As Well As Dynamic Camera For Offline Stored Video And Live Webcam Scene.

To design smart computer vision based video surveillance systems which will contribute to the safety of people for this to monitor the areas like airports, railway stations, shopping malls and other public related places where security is the major, so here we have proposed the real time unattended object detection and tracking system which will monitor this areas for that there must be one of the administrator who will keep the record of this auspicious activities for tracking the live scenes as well as database of stored recordings. Fig.2. shows the GUI of admin login page of the system in which user name and password is case sensitive, for log in to the system admin or the user has to enter correct credentials so that system will enter in to its main page for further processing of the scenes.

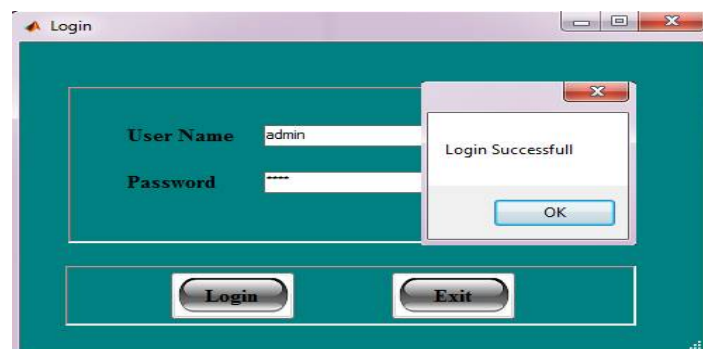


Fig.2. System Login Page

In our proposed MATLAB based object detection and tracking system according to algorithm mentioned in section V, Fig. 3. Shows the obtained result in which offline stored video captured at railway station using stationary camera is fetched as input to the system, where system respond and shows unattended object for long time such as travelling bag of the passenger is being detected and tracked.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 6, June 2016

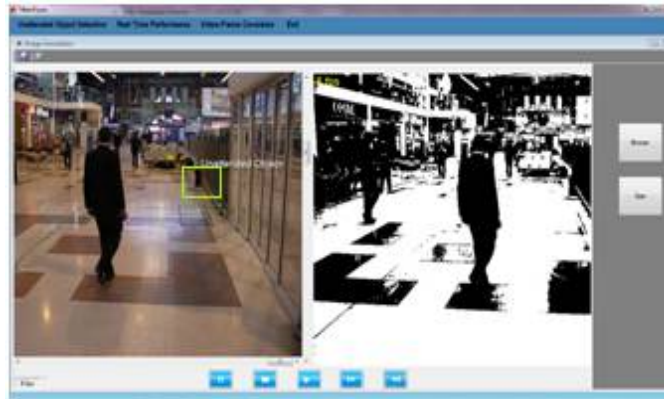


Fig.3. Unattended Object Detection and Tracking at Railway Station (Offline Stored Video Input).

Moving object detection and tracking is one of the major first steps in video surveillance systems which have attracted a great interest from computer vision and image processing researchers due to its application in areas, like traffic monitoring and image recognition. In case of real time moving object detection from the scene involves identification of an object in consecutive frames but in case of object tracking which uses to monitor the movements with respect to the region of interest of the object. Single object and multiple object movements in a frame with respect to the computed vectors are segmented with the help of specified threshold limit. The extracted movements are tracked using mean shift and particle filter algorithm. Tracking from scene can be possible due to segmented continuous video frame sequences of the real time captured offline stored database video as an input to the system is shown in Fig.4.



Fig. 4. Unattended Object (Travelling Bag) Detection and Tracking at Railway Station from Segmented Continuous Video Frame Sequences. (Offline Stored Video Input to the System).

Thresholding is one of the most powerful and important tools in image processing of computer vision for image segmentation. The importance of segmented images obtained from thresholding has the advantages as it requires smaller storage space, performs fast processing and eases in manipulation compared with gray level image. Fig. 4 shows the travelling bag from the scene as an unattended object is detected; as bag is not attended for long time then system administrator store its template (Fig. 5) for tracked object using system interface tabs and fires security alarm for tracked auspicious objects.

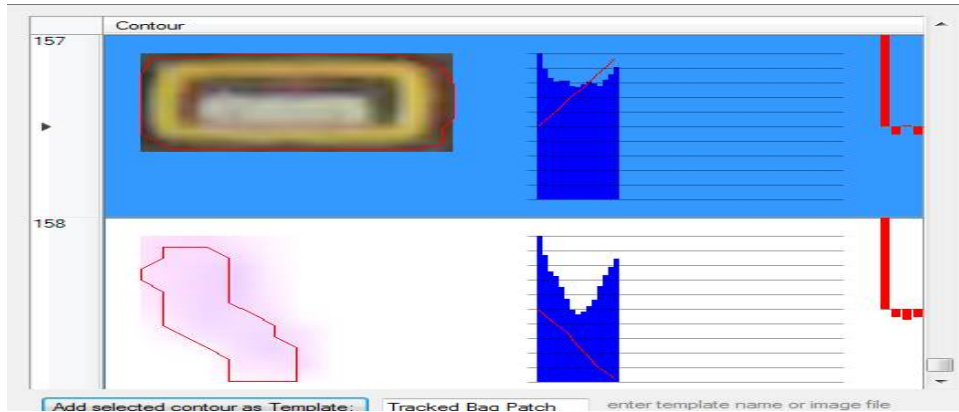


Fig.5. Live Template Selection and Adding Along the Contour of an Object
(Bag Patch as Template for Live Tracking).

Our MATLAB based system is operating two modes one is Real time (Live Scene) and offline stored database video. In Real time mode system that processes 800*600, 640*480 and 320*240 resolution video sequences and provide the location of a predefined objects from the dataset template as shown in Fig.6.



Fig. 6. Live Object Tracking Using Webcam (Real Time).

Fig.5. shows template for the bag patch is added so whenever bag will come across either stationary or moving camera our proposed Mean Shift and Particle Filter algorithm will be get executed in which dynamic shape of object along with rectangular bounding box is calculated by the particle filter, and the dynamic appearance is allocated by the mean shift that is embedded in the particle filter is shown in fig. 6.

VI. CONCLUSION

In our MATLAB based Real Time Unattended Moving Object Detection and Tracking system we have tracked single object through complex multiple scenarios from videos captured by a single dynamic as well as stationary cameras with stationary or complex moving backgrounds. Test result shows that the proposed system is very robust which is resulting in considerable improvement by reducing somewhat tracking drifts and capable of tracking the object of intersection, long-term partial occlusions in scene, object pose or shape changes, fast motion changes, and

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 6, June 2016

complex background. The proposed tracking scheme nearly reaches its limit when video scenes contain too many similar objects in terms of their appearance distributions in video scene, frequent intersections of objects at background and occlusions. Other limitation such as the computational speed and tracking drifts of the system requires future improvement.

REFERENCES

- [1] Moeslund T. B., Hilton A., and Kruger V., "A survey of advances in vision-based human motion capture and analysis," *Computer Vision and Image Understanding*, vol. 104, pp. 90-126, 2006.
- [2] L. Li, Huang W., Gu I. Y. H., and Tian Q., "Statistical modelling of complex background for foreground object detection," *IEEE Transactions on Image Process.*, vol. 13, no. 11, pp. 1459-1472, November 2004
- [3] Comaniciu D., Ramesh V., and Meer P., "Kernel-based object tracking," *IEEE Transactions on Pattern Anal. Mach. Intell.*, vol. 25, no. 5, pp. 564-577, May 2003.
- [4] Khan Z. H., Yu-Hua Gu I., and Andrew G., "Robust visual object tracking using multi-mode anisotropic mean shift and particle filters", *IEEE Transaction on Circuits and System for video tech.*, vol. 21, no.1, January 2011.
- [5] Wang Z., Yang X., Xu Y., and Yu S., "Camshift guided particle filter for visual tracking," *IEEE Workshop on Signal Process. Syst.*, pp. 301-306, October 2007.
- [6] Sumin Q. and Xianwu H., "Hand tracking and gesture recognition by anisotropic kernel mean shift," *IEEE Int. Conf. Neural Netw. Signal Process*, vol. 25. , pp. 581-585., June 2008.
- [7] Maggio E. and Cavallaro A., "Hybrid particle filter and mean shift tracker with adaptive transition model," *IEEE Int. Conf. Acoust. Speech Signal Process*, vol. 2, pp. 221-224, March. 2005.