# Survey on Video Object Tracking based on Object and Motion Representation

**B.Umamaheswari**

*Dept of CSE, Sri Krishna College of Engineering and Technology, Coimbatore, India*

*Abstract-* **Automated video analysis is important for many vision applications. Object tracking is the fundamental step in automated analysis. Object tracking is the tough problem. Difficulties in object due to object motion, the appearance changes in both the object and sight, nonrigid object structure and camera motion. Tracking is usually performed in the environment of top – level applications that require the location and shape of the object in every frame. In this paper, we present a brief analysis of tracking method on the basis of object and motion representation used.**

*Keywords-* **Moving object tracking, Background subtraction, Initialization, Updating Background, Motion based segmentation**

## I.INTRODUCTION

Object tracking is an important mission within the field of computer idea. There are three key steps in video analysis: detection of attractive moving object, tracking that object from frame to frame, analysis the tracked object to identify their behaviour. Therefore, the make use of object tracking is relevant in the tasks of:

- Motion-based identification - human identification based on way of walking, automatic object detection etc
- Automated examination - monitoring a scene to detect doubtful activities or doubtful events
- Video indexing - automatic explanation and recovery of the videos in multimedia database
- Traffic monitoring - real- time assembly of traffic information to direct traffic flow
- Vehicle navigation - video-based path planning and blockage evasion capabilities

A tracker assigns constant label to the tracked objects in different frames of a video. Moreover, depending on the tracking domain, a tracker can also offer object-centric information, such as direction, area, shape of an object. Tracking objects can be difficult due to:

- Noise in images
- Complex object motion
- Nonrigid or expressed nature of objects
- Partial and full object occlusion
- Difficult object shapes

One can make simpler tracking by impressive constraints on the motion and/or appearance of objects.

### *A. Representation of Object*

In a tracking scenario, the object for further analysis. The important to track in a specific area are boats on the sea, Vehicle on a road, people walking on road. Object can be represented by their shapes and appearance.

- By Points Representation -The object is represented by a point
- Ancient geometric shapes-Object shape is represented by a rectangle, ellipse.
- Object contour-It will just represent the boundary of an object
- Uttered shape models -Object are together of body parts that are held together with joints for example the human body leg, hand, upper body and feet connected by joints.

There are a number of ways to represent the appearance type of object. Some of the appearance representations are:

**Possibility densities of object appearance** it can either be parametric or non parametric and histogram [Comaniciu et al. 2003].
**Templates** are formed using simple geometric shapes [Fieguth and Terzopoulos 1997]. An advantage it carries both spatial and appearance information.
**Active appearance models** are generated at the same time modeling the object shape and appearance [Edwards et al. 1998].
**Multiview appearance models** encode different views of an object.

## II.BACKGROUND SUBTRACTION

 

Background subtraction becomes the initial step in tracking algorithm. In previous day until 90's background subtraction was known as a prevailing preprocessing step but limited within the indoor environment. In 1998, Stauffer and Grimson[1] offered the idea of representing each pixel by a mixture of Gaussians(MOG) and updating each pixel with new Gaussian through run-time. This allows background subtraction to be used in outdoor background.The technique by Stauffer and Grimson has today turn into the standard of background subtraction.After that number of advances have occurred which can be divided into Representation of background, categorization, background updating and background initialization.

### *A. Representation of background*

The MoG representation can be in RGB freedom,but we can also apply the other color freedom, notice[2] for a summary. Frequently a representation where the color and intensities separated and it is applied e.g., YUV [8],HSV[26],and normalized RGB[3],because this allows for detecting shadow-pixels incorrectly classified as object-pixels[13].Theoretically different representations have been developed. Elgammal et al. [24] use a kernel-based approach there they represent a background pixel by the individual pixels of the end *N* frames. Haritaoglu et al. [16] represent the minimum and maximum value mutually with the maximum allowed modify of the value in two consecutive frames. Eng et al. [22,23] divide a background model learned by a number of without overlapping blocks. The pixel inside each block is clustered according to its similarity. The representation of background in block by spatio-temporal representation.

The representation cannot be choose only based on the accuracy but also depend on the speed of the implementation and the application. It specifies that total accuracy of background subtraction is the combination of representation, categorization, updating and initialization.

### *B. Categorization*

After background subtraction there will be noise (i.e. false positive and negatives).By means of standard filtering techniques based on connected component analysis, Dimension, Median filter , Morphology and closeness can improve the result [26,24,18,3,11,12].Otherwise the neighboring pixels are possible to be both foreground or background can be used for categorization. To implement this idea the Markov Random fields is applied [10, 9].

The new methods have tried to identify the inaccurate pixels directly and they use classifiers to separate the pixels into a number of subcategory: shadows, unchanged background, and moving objectetc. [25, 26, 15]. Classifiers is based on color, gradients[3],flow information[26],and thresholding[22].

### *C. Update Background*

In the Outside background the value of a background pixel will modify frequently for that updating is required. If the background move slowly in the scene can be updated itself by including the current pixel value into the model [26, 24,3,7]. The different approach is used to measure the background overall average change in the scene compared to the usual background and this is used to update the model [20,11]. If no real-time needs are present, both past and future values used to update the background [21]. The model updates only the pixels that are classified as unchanged background is updated.

Fast changes in the scene are adapted by adding a new form to the model. For MoG model a new form is a Gaussian distribution. The further pixels that support this distribution because it will have more weight.

### *D. Initialization of background*

During an initialization time the background model need to be learned. In earlier approaches assumed that in a number of consecutive frames no moving object will present. But in real scenarios this assumption will not be valid and advanced methods focused on initialization of moving objects.

In the MoG representation during initialization moving object will be accepted to some extend while each foreground object will be represented by its own distribution which is possible to have a low weight. This wrong distribution is possible to produce false positive during categorization process. To find pixel that are true background pixels need a different approach and then only initialization can apply. This can be done by using temporal median filter if it is less than 50% of the values belong to foreground objects [22, 19, 16].

The new alternative first in the initialization phase divide the pixels into temporal Subintervals with similar values. Second find the background best subinterval with the minimum average motion [17] or with the maximum ratio between the number of samples and variance [6, 5].

### *E. Motion-based segmentation*

Foreground segmentation is based on motion that differs in the consecutive frames i.e., finding the person moving by its motion. The motion is measured using frame differencing.

Sidenbladh [14] calculate optical flow for many frame each containing a walking person. To detect walking person in the video the support vector machine (SVM) is used. When calculating optical flow there can be noise and instead, by using higher level entities the frame flow can be measured. Image differencing adapt easily when there is a changes in the scene, but pixel from the human body will not moved or related to their neighbours are not detected. The different type of frame differencing is used in Viola et al. [4].The tracked object is represented by rectangle shape. In the current frame a rectangle of pixels is compared with the corresponding rectangle pixel in the previous frame. The rectangle in the current frame can shift to up, down, left and right can do like this. Frame differencing is performed and the energy is low in the output the higher probability that person actually moved in the direction.

### *F. Advanced Tracking*

The foreground segmentation include to a large coverage been motivated by the increased focus on surveillance application. This is achieved to some extent within background subtraction where analysis of video sequence. Effort in this direction stated, no one is report an accurately independent system. In the most surveillance application several cameras are required to cover the scene of importance at an acceptable motion. The tracking different camera are analysed but it didn't give any satisfied result.

## III.CONCLUSION

The object tracking is one of the fundamental steps in video analysis and it is difficult in tracking because of the object movement. Here I have discussed the tracking by background subtraction and motion based segmentation it does not need any training dataset and it is not limited applicable for automated video analysis. As a further step the appearance based segmentation can be further study.

## IV.REFERENCES

*[1] C. Stauffer, W.E.L. Grimson, (June 1998)' Adaptive background mixture models for real-time tracking', in: Computer Vision and Pattern Recognition,Santa Barbara, CA.*

*[2] F. Kristensen, P. Nilsson, V. O wall ¨, (Jan 13–16, 2006), 'Background segmentation beyond RGB', in: Asian Conference on Computer Vision, LNCS 3852, Hyderabad, India.*

*[3] S.J. McKenna, S. Jabri, Z. Duric, H. Wechsler,(March 2000) 'Tracking Interacting People, in: International Conference on Automatic Face and Gesture Recognition', Grenoble, France.*

*[4] P. Viola, M.J. Jones, D. Snow, (2005) 'Detecting pedestrians using patterns of motion and appearance', International Journal of Computer Vision 63 (2) 153–161.*

*[5] H. Wang, D. Suter,( Jan 13–16, 2006)'A novel robust statistical method for background initialization and visual surveillance', in: Asian Conference on Computer Vision, Hyderabad, India, Lecturer Notes in Computer Science 3851.*

*[6]H. Wang, D. Suter, (Oct 15–16, 2005) 'Background initialization with a new robust statistical approach', in: International Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance,Beijing, China.*

*[7] C. Stauffer, W.E.L. Grimson, (June 1998) 'Adaptive background mixture models for real-time tracking', in: Computer Vision and Pattern Recognition,Santa Barbara, CA.*

*[8] C.R. Wren, A. Azarbayejani, T. Darrell, A.P. Pentland, Pfinder, (1997)' real-time tracking of the human body', Transactions on Pattern Analysis and Machine Intelligence 19 (7) 780–785.*

*[9] Y. Sheikh, M. Shah, (2005) 'Bayesian modelling of dynamic scenes for object detection', IEEE Transactions on Pattern Analysis and Machine Intelligence 27 (11) 1778–1792.*

*[10] K. Schindler, H. Wang, (Jan 13–16, 2006) 'Smooth foreground-background segmentation for video processing', in: Asian Conference on Computer Vision, LNCS 3852, Hyderabad, India.*

*[11] T. Yang, S.Z. Li, Q. Pan, J. Li,(June 20–25, 2005) 'Real-time multiple objects tracking with occlusion handling in dynamic scenes', in: Computer Vision and Pattern Recognition, San Diego, CA.*

*[12] T. Zhao, R. Nevatia,(2004)'Tracking multiple humans in complex situations', IEEE Transactions on Pattern Analysis and Machine Intelligence 26 (9) 1208–1221.*

*[13] A. Prati, I. Mikic´, M.M. Trivedi, R. Cucchiara, (2003) ' Detecting moving shadows: algorithms and evaluation', IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (7) 918–923.*

*[14] H. Sidenbladh, (Aug 2004) 'Detecting human motion with support vector machines', in: International Conference on Pattern Recognition,Cambridge, UK.*

*[15] T. Horprasert, D. Harwood, L.S. Davis,(Sep 1999) 'A statistical approach for real-time robust background subtraction and shadow detection', in:IEEE ICCV'99 Frame-Rate Workshop, Corfu, Greece.*

*[16] I. Haritaoglu, D. Harwood, L.S. Davis,(2000) 'W4: real-time surveillance of people and their activities', IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (8) 809–830.*

*[17] D. Gutchess, M. Trajkovic, E.C. Solal, D. Lyons, A. Jain,( July 9–12, 2001) 'A background model initialization algorithm for video surveillance', in: International Conference on Computer Vision, Vancouver, Canada.*

*[18] P. Guha, A. Mukerjee, K.S. Venkatesh,(Oct 15–16,2005) 'Efficient occlusion handling for multiple agent tracking by reasoning with surveillance event primitives', in: Workshop on Visual Surveillance and Performance Evaluation of Tracking and Surveillance, Beijing, China.*

*[19] B. Gloyer, H.K. Aghajan, K.Y.S. Siu, T. Kailath,(1995) 'Video-based freeway monitoring system using recursive vehicle tracking', in: IS&T-SPIE Symposium on Electronic Imaging: Image and Video Processing.*

*[20] P. Fihl, R. Corlin, S. Park, T.B. Moeslund, M.M. Trivedi,(Nov 6–8, 2006) 'Tracking of individuals in very long video sequences', in: International Symposium on Visual Computing (ISVC), Stateline, Lake Tahoe, Nevada, USA.*

*[21] P. Figueroa, N. Leite, (2006) 'R.M.L. Barros, Background recovering in outdoor image sequences: an example of soccer players segmentation', Image and Vision Computing 24 (4) 363–374.*

*[22] H.L. Eng, K.A. Toh, A.H. Kam, J. Wang, W.Y. Yau, (Oct 13–16, 2003) 'An automatic drowning detection surveillance system for challenging outdoor pool environments', in: International Conference on Computer Vision,Nice, France.*

*[23] H.L. Eng, J. Wang, A.H.K.S. Wah, W.Y. Yau, (2006) 'Robust human detection within a highly dynamic aquatic*

 

*environment in real time',IEEE Transactions on Image Processing 15 (6) 1583–1600.*

*[24] A. Elgammal, D. Harwood, L. Davis, (June 2000) 'Non-parametric model for background subtraction', in: European Conference on Computer Vision, Dublin, Ireland.*

*[25] M. Chen, G. Ma, S. Kee, (Jan 2005) 'Pixels classification for moving object extraction', in: IEEE Workshop on Motion and Video Computing (MOTION'05), Breckenridge, Colorado.*

*[26] R. Cucchiara, C. Grana, M. Piccardi, A. Prati,(2003) 'Detecting moving objects, ghosts, and shadows in video streams', IEEE Transactions on Pattern Analysis and Machine Intelligence 25 (10) 1337–1342*

 