ISSN:0975-9646

Object Tracking in Video Using Mean Shift Algorithm: A Review

Aparna Shivhare and Dr. Vinita Choudhary
Department of Electronics and Communication Engineering,
Ujjain Engineering College, Ujjain, M. P. (India)

Abstract: One of the most popular areas of video processing is object tracking. The main purpose of object tracking is to estimate the position of the object in images in a continuous manner and reliably against dynamic scenes. This can be beneficially achieved by using the mean shift object tracking algorithm. In this approach, a rectangular target window is defined in an initial frame for a moving target in a video, and after that, the tracked object is separated from the back ground by processing the data within that window. It is considered as a new approach toward target localization and representation in visual tracking of objects. By using an isotropic kernel, target representations which are based on feature histogram are regularized with the help of spatial masking.

Keywords: video, object tracking, kernel, mean-shift, target localization and representation.

I. INTRODUCTION

Object Tracking is an important mechanism as it enables various applications such as: Retail space instrumentation – to analyze shopping behavior of customers, to enhance building and environment design; Security and surveillance – to recognize people, to provide better sense of security using visual information; Video abstraction – to obtain automatic annotation of videos, to generate object-based summaries; Medical therapy – to improve the quality of life for physical therapy patients and disabled people; Traffic management – to analyze flow, to detect accidents; Interactive games – to provide natural ways of interaction with intelligent systems such as weightless remote control; Video editing – to eliminate cumbersome human-operator interaction, to design futuristic video effects [1]. In other words, it provides ease of production and has great applicability to daily problems, e.g. plane detection, object-based video compression [9], perceptual user interface [4], adaptive traffic lights, human-computer interaction, driver assistance [10], surveillance [5], vehicle navigation, smart rooms [8], augmented reality [7] etc.

In video tracking, when the objects are moving fast relative to the frame rate then it is somewhat difficult to associate target location in consecutive video frames [2]. For this, various approaches for object tracking have been proposed. And which approach is to be implemented is completely depending upon the context in which the tracking is performed. Based on the changes to the properties of the object being tracked, object-tracking is a type of technique to track an object and to perform an immediate action on some other object which has no relation to the tracked object.

Fig. 1: Tracking a moving car

A typical visual tracker consists of two major components, which can be distinguished as:

A. Target Representation and Localization.
B. Filtering and Data Association.

Target Representation and Localization is a type of bottom-up process which has to deal with the variations in the appearance of the target. Filtering and Data Association is a type of top-down process which has to cope with the dynamics of the tracked object, evaluation of different hypothesis and learning of scene priors [1].

II. OBJECT TRACKING IN VIDEO SEQUENCE

Object tracking in video is defined as the process of locating the position of a moving object or multiple objects over the time using a camera. But it should not be confused with the Camera tracking. It has a number of uses, which includes: video communication and compression, human-computer interaction, security and surveillance, augmented reality, traffic control, medical imaging and video editing. Object tracking may be a time taking technique because of the quantity of data i.e., contained in video.

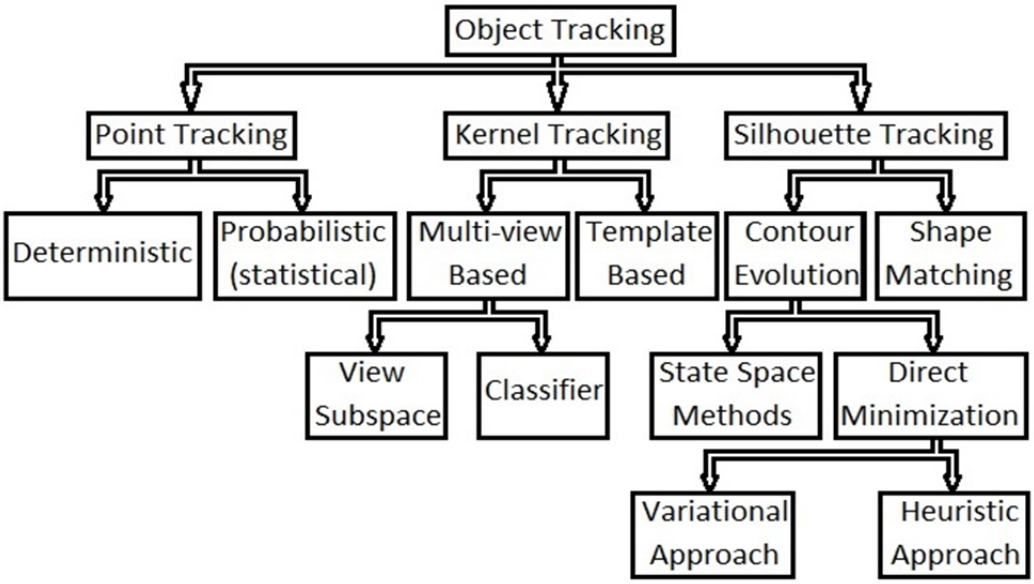

Fig. 2: Classification of Object Tracking

The main purpose of object tracking is to relate target objects in successive video frames. The relation can be specifically tough when the objects are in motion fast relative to the frame rate. Apart from this, another condition that steps up the difficulty of the issue is when the tracked object frequently changes orientation and location over time. For these conditions object tracking systems usually need a motion model which elaborates how the image of the target might change for various possible motions of the object.

A. Video processing

Video processing is a specific type of signal processing that frequently uses video filters having the input and output signals as video streams or video files. Video processing methods are employed in DVDs, video scalars, video codec, television sets, video players, VCRs, and other devices [11]. This technique also needs a stream processing construction, in which video frames coming from an ongoing stream are taken into account for processing one or more at an instant. This type of method is found to be somewhat difficult in systems that have current video or where the video data is so huge that cramming the overall set into the counter is incompetent [12].

B. Algorithms

In order to operate object tracking, an algorithm examines consecutive video frames and produces the motion of targets between the frames as output. There are a wide variety of algorithms, each possessing advantages and disadvantages. There are two main components of a visual object tracking system: target representation and localization, as well as filtering and data association.

Target representation and localization gives numerous tools for recognizing the moving object. Locating and tracking the target object completely is based on the algorithm. For example, the use of blob tracking is helpful for analyzing human motion as a person's contour alters effectively. Generally, the estimated complexity for these algorithms is less. Some common target representation and localization algorithms are as follows:

1) *Mean-Shift tracking:* It is also known as Kernel-Based tracking. It is an iterative positioning method built on the augmentation of a parallel measure (Bhattacharyya coefficient) [6].
2) *Contour tracking:* It is also known as Condensation Algorithm and is used to estimate the object boundary. Contour tracking process iteratively develops an initial contour occurred from the foregoing frame to its new location in the present frame. This algorithm of contour tracking directly outputs the contour by reducing the contour energy using gradient descent.

III. MEAN-SHIFT BASED TRACKING ALGORITHM

In 1975, Fakunaga and Hostetter proposed an algorithm called as Mean-shift (MS) object-Tracking Approach. It is basically iterative expectation maximization – clustering algorithm executed within local search regions [3]. In other words, it is a type of non-parametric clustering algorithm that does not need prior information of the number of clusters and also does not constrain the shape of the clusters. That is, the Mean-shift algorithm is a non-parametric density gradient estimator.

The following steps are iterated in order to track the object by using the Mean-Shift algorithm:

A. Select a search window size and the initial position of the search window.
B. Estimate the mean position in the search window.
C. Center the search window at the mean position estimated in Step B.
D. Repeat Steps B and C until the mean position moves less than a preset threshold. That is, until convergence is achieved.

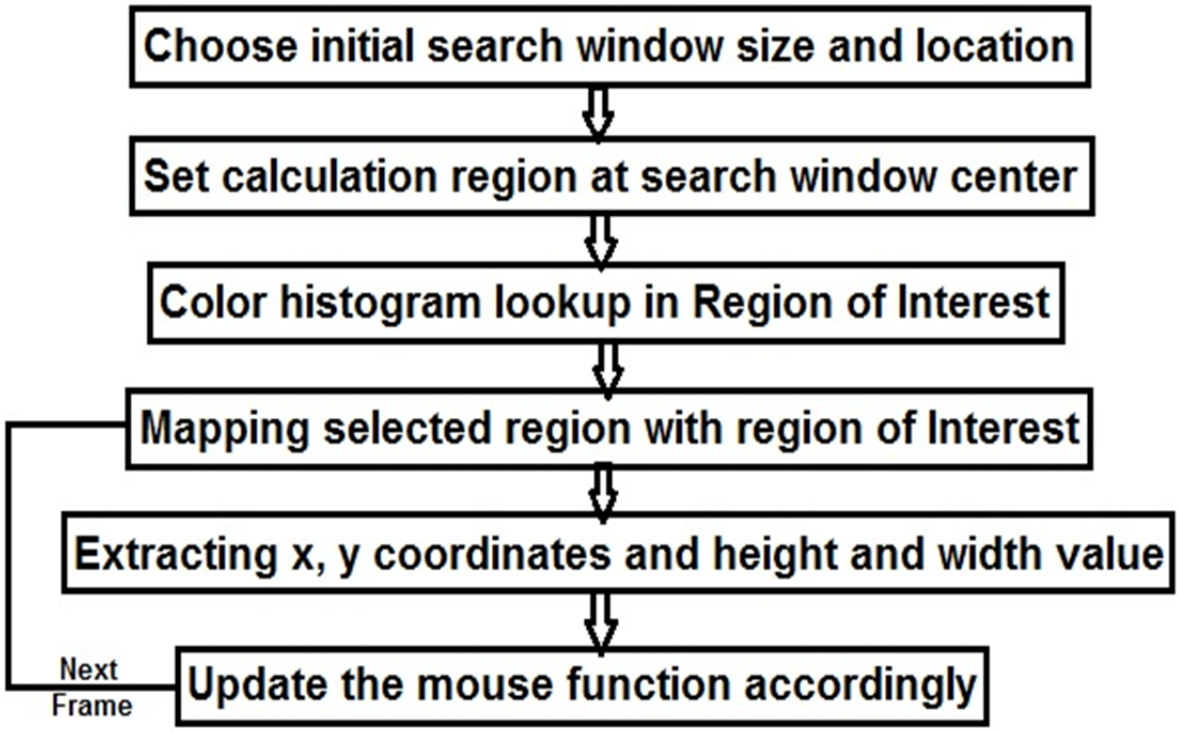

Fig. 3: Procedure of Object Tracking

After the Mean-Shift Algorithm is executed on various videos it is concluded that when the target moves so fast that the target area in the two neighboring frame will not overlap, tracking object often converges to a wrong object. Because of this issue, traditional Mean-Shift Algorithm gets failed to track fast moving object. But there are some solutions like combining Kalman filter or particle filter with Mean-shift Algorithm [4]. The center of convergence will become more accurate as at first it predicts the direction and speed of the object and then adjusting the search window center of the mean shift convergence. However, such methods are unsuited to be using in real time tracking systems as these require high CPU costing to estimate the moving object location. The kernel-based tracking algorithm, when incorporated with prior task-specific information, can gain adequate results. This enhanced model could profitably detect and track a human subject in arbitrary motion and in a situation where there is a certain alter in radiance. It was constituted that a well built alteration in radiance induced the procedure to experience quite major deformation in the probability distribution image. Hence, the non-adaptive behavior of the mean shift algorithm may conduct to a wrong tracking conclusion.

IV. CONCLUSION

Object tracking in an untidy surrounding remains a demanding investigation subject. In this paper we reviewed the mean shift algorithm with some definite improvements. At first, the frames of the video file are created. And after that, the window size is estimated to track a target precisely when the target's shape and location are altering. Numerous consequences can be obtained by implementing mean-shift algorithm over the target object. According to the motion model, one should begin the tracker in several different positions in the neighborhood of basin of attraction if the movement of the target from frame to frame is well known to be greater than the operational basin of attraction.

If total closure is present, one should take on a more revealing motion filter. In the same way, one should check that the selected target description is adequately different for the application domain or not. Hence our review of tracking an object will surely make us to analyze new areas of investigation and moreover, also helps to improve its applications in the already existing areas.

REFERENCES

[1] Comaniciu, V. Ramesh and Peter Meer, "*Kernel-Based Object Tracking*," IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 25, NO. 5, MAY 2003.

[2] W. Hu, T. Tan, L. Wang, S. Maybank, "*A survey on visual surveillance of object motion and behaviors*", IEEE Trans. Syst. Man Cyber.-C 34 (3) (2004) 334–352.

[3] R. T. Collins, "*Mean-shift blob tracking through scale space*", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2003.

[4] R. Bradski. "*Computer vision face tracking for use in a perceptual user interface*", Intel Technology Journal, 2nd Quarter, 1998.

[5] R. Collins, A. Lipton, H. Fujiyoshi, and T. Kanade, "*Algorithms for cooperative multisensor surveillance*, "*Proceedings of the IEEE*, vol. 89, no. 10, pp. 1456–1477, 2001.

[6] Comaniciu, D.; Ramesh, V.; Meer, P., "*Real-time tracking of non-rigid objects using mean shift*," Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on, vol.2, no., pp.142, 149 vol.2, 2000.

[7] V. Ferrari, T. Tuytelaars, and L. V. Gool, "Real-*time affine region tracking and coplanar grouping*," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, Kauai, Hawaii, volume II, 2001, pp.226–233

[8] S. Intille, J. Davis, and A. Bobick, "*Real-time closed-world tracking*," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, San Juan, Puerto Rico, 1997, pp. 697–703.

[9] D. Bue, D. Comaniciu, V. Ramesh, and C. Regazzoni, "*Smart cameras with real-time video object generation*," in *Proc. IEEE Intl. Conf. on Image Processing*, Rochester, NY, volume III, 2002, pp.429–432.
[10] U. Handmann, T. Kalinke, C. Tzomakas, M. Werner, and W. von Seelen, "*Computer vision for driver assistance systems*," in *Proceedings SPIE*, volume 3364, 1998, pp. 136–147.
[11] Wang, Yao, Jörn Ostermann, and Ya-Qin Zhang. *Video Processing and Communications*. Signal Processing Series. Upper Saddle River, N.J.: Prentice Hall, 2002. ISBN 0-13-017547-1.
[12] http://in.mathworks.com/solutions/image-video-processing/video-processing.html