INTERNATIONAL RESEARCH JOURNAL OF ENGINEERING & APPLIED SCIENCES

ISSN: 2322-0821(O)
ISSN: 2394-9910(P)
VOLUME 3 ISSUE 4
Oct 2015 - Dec 2015

ANALYSIS OF FRAMING OF MULTIPLE MOVING OBJECT USING MEAN SHIFT ALGORITHM

Traptimala Mishra[1], Md. Amzad Quazi[2], Md. Abdullah[3]

[1]*M. Tech Scholar, Department of Electronics & Communication Engineering, SIRT, Bhopal, M. P., India*
[2]*Asst. Professor, Department of Electronics & Communication Engineering, SIRT, Bhopal, M. P., India*
[3]*Asst. Professor Department of Electronics & Communication Engineering, SIRT, Bhopal, M. P., India*

Abstract -The goal of this paper is to review the different types of tracking methods, classify them into different categories and analysis their drawbacks & advantages. Object tracking, in general, is a challenging problem. Difficulties in tracking objects can arise due to abrupt object motion, changing appearance patterns of the object and the scene, non rigid object structures, object-to-object and object-to-scene occlusions, and camera motion. Tracking is usually performed in the context of higher-level applications that require the location and/or shape of the object in every frame. Typically, assumptions are made to constrain the tracking problem in the context of a particular application. In this survey, we categorize the tracking methods on the basis of the object and motion representations used, provide detailed descriptions of representative methods in each category, and examine their pros and cons. Moreover, I have discuss the important issues related to tracking including the use of appropriate image features, selection of motion models, and detection of objects and framing.

Index Terms— Mean shift, framing, Kalman filter, object detection, object tracking, occlusion.

I. INTRODUCTION

A. Concept of visual object tracking

Visual object tracking is an important task within the field of computer vision. It aims at locating a moving object or several ones in time using a camera. An algorithm analyses the video frames and outputs the location of moving targets within the video frame. So it can be defined as the process of segmenting an object of interest from a video scene and keeping track of its motion, orientation, occlusion etc. in order to extract useful information by means of some algorithms. Its main task is to find and follow a moving object or several targets in image sequences.

The proliferation of high-powered computers and the increasing need for automated video analysis have generated a great deal of interest in visual object tracking algorithms. The use of visual object tracking is pertinent in the tasks of automated surveillance, traffic monitoring, vehicle navigation, human-computer interaction etc. Automated video surveillance deals with real time observation of people or vehicles in busy or restricted environments leading to tracking and activity analysis of the subjects in the field of view. There are three key steps in video surveillance: detection of interesting moving objects, tracking of such objects from frame to frame, and analysis of object tracks to recognize their behavior.

Visual object tracking follows the segmentation step and is more or less equivalent to the "recognition" step in the image processing. Detection of moving objects in video streams is the first relevant step of information extraction in many computer vision applications. There are basically three approaches in visual object tracking. Feature based methods aim at extracting characteristics such as points, line segments from image sequences, tracking stage is then ensured by a matching procedure at every time instant. Differential methods are based on the optical flow computation, i.e. on the apparent motion in image sequences, under some regularization assumptions. The third class uses the correlation to measure inters image displacements. Selection of a particular approach largely depends on the domain of the problem.

The development and increased availability of video technology have in recent years inspired a large amount of work on object tracking in video sequences. Many researchers have tried various approaches for object tracking. Nature of the technique used largely depends on the application domain. Some of the research work done in the field of visual object tracking includes, for example: The block matching technique for object tracking in traffic scenes in a motionless airborne camera is used for video capturing. They have discussed the block matching technique for different resolutions and complexities.

Object tracking algorithm using a moving camera. The algorithm is based on domain knowledge and motion modeling. Displacement of each point is assigned a discreet probability distribution matrix. Based on the model, image registration step is carried out. The registered image is then compared with the background to track the moving object.

It uses object features gathered from two or more cameras situated at different locations. These features are then combined for location estimation in video surveillance systems. Another simple feature based object tracking method is explained. The method first segments the image into foreground and background to find objects of interest.

Then four types of features are gathered for each object of interest. Then for each consecutive frame the changes in features are calculated for various possible 2 directions of movement. The one that satisfies certain threshold conditions is selected as the position of the object in the next frame.

A feedback-based method for object tracking in presence of occlusions in this method several performance evaluation measures for tracking are placed in a feedback loop to track non-rigid contours in a video sequence.

B. Applications of visual object tracking

Visual object tracking has many applications. Some important applications are:

(1) Automated video surveillance: In these applications computer vision system is designed to monitor the movements in an area (shopping malls, car parks, etc.), identify the moving objects and report any doubtful situation. The system needs to discriminate between natural entities and humans, which require a good visual object tracking system.

(2) Robot vision: In robot navigation, the steering system needs to identify different obstacles in the path to avoid collision. If the obstacles themselves are other moving objects then it calls for a real-time visual object tracking system.

(3) Traffic monitoring: In some countries highway traffic is continuously monitored using cameras. Any vehicle that breaks the traffic rules or is involved in other illegal act can be tracked down easily if the surveillance system is supported by an object tracking system.

(4) Animation: Visual object tracking algorithm can also be extended for animation.

(5) Government or military establishments.

To sum up, visual object tracking is applied to a wide range of fields nowadays, such as multimedia, video data compression, industry production, and military affairs and so on.

Accordingly, it is of great real significance and application value to investigate in visual object tracking.

The detection and tracking of motion object in real time image sequences is the important task in image processing, computer vision, mode identification etc. It flexibly combines the technologies of image processing, auto control and information science, forms a new technology of real time detection of motion object, extraction location information of the object and tracking of it. Furthermore, rapid progress in technologies of signal processing, sensor and new material provides reliable software and hardware for the capturing and processing of image in real time.

C. Difficulties

In general, trackers can be subdivided into two categories. First, there are generic trackers which use only a minimum amount of a priori information, e.g., the mean-shift approach by Comaneci and the color-based particle filter developed by Perez. Secondly, there are trackers that use a very specific model of the object, like e.g. the spline representation of the contour by Izard.

The objects found in video trackers are often being tracked in "difficult" environments characterized by the variable visibility (e.g. shadows, occlusions) and the presence of spurious (e.g. similarly-colored) objects and backgrounds. As a result, visual object tracking still suffers from a lack of robustness due to temporary occlusions, objects crossing, changing lighting conditions, secularities and out-of plane rotations. The main difficulty in video tracking is to associate target locations in 3 consecutive video frames, especially when the objects are moving fast relative to the frame rate. Here, video tracking systems usually employ a motion model which describes how the image of the target might change for different possible motions of the object to track.

Many algorithms have been developed and implemented to solve the difficulties that a raised from the video tracking process, such as SIFT (Scale Invariant Feature Transform), KPSIFT (key point-preserving-SIFT), PDSIFT (partial-descriptor-SIFT),RANSAC(Random Sample Consensus), mean shift, optical flow, GDOH (gradient distance and orientation histogram), KALAMN etc. The role of the tracking algorithm is to analyze the video frames in order to estimate the motion parameters. These parameters characterize the location of the target.

D. Algorithms For Object Tracking

1) Scale Invariant Feature Transform (SIFT)

Scale-invariant feature transform (or SIFT) is an algorithm in computer vision to detect and describe local features in images. The algorithm was published by David Lowe in 1999.

Applications include object recognition, robotic mapping and navigation, image stitching, 3D modeling, gesture recognition, video tracking, individual identification of wildlife and match moving.

SIFT key points of objects are first extracted from a set of reference images and stored in a database. An object is recognized in a new image by individually comparing each feature from the new image to this database and finding candidate matching features based on Euclidean distance of their feature vectors. From the full set of matches, subsets of key points that agree on the object and its location, scale, and orientation in the new image are identified to filter out good matches. The determination of consistent clusters is performed rapidly by using an efficient hash table implementation of the generalized Hough transform. Each cluster of 3 or more features that agree on an object and its pose is then subject to further detailed model verification and subsequently outliers are discarded. Finally the probability that a particular set of features indicates the presence of an object is computed, given the accuracy of fit and number of probable false matches. Object matches that pass all these tests can be identified as correct with high confidence.

All SIFT algorithm have the major stages of computation used to generate the set of image features:

- Scale-space extreme detection: The first stage of computation searches over all scales image locations. It is implemented efficiently by using a difference-of-Gaussian function to identify potential interest points that are invariant to scale and orientation.
- Key point localization: At each candidate location, a detailed model is fit to determine location and scale. Key points are selected based on measures of their stability.

- Orientation assignment: One or more orientations are assigned to each key point location based on local image gradient directions. All future operations are performed on image data that has been transformed relative to the assigned orientation, scale, and location for each feature, thereby providing invariance to these transformations.
- Key point descriptor: The local image gradients are measured at the selected scale in the region around each key point. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination.

SIFT features have many advantages such as follows:

SIFT features are all natural features of images. They are favorably invariant to image translation, scaling, rotation, illumination, viewpoint, noise etc.

Good specialty, rich in information, suitable for fast and exact matching in a mass of feature database.

Fertility. Lots of SIFT features will be explored even if there are only a few objects.

Relatively fast speed. The speed of SIFT even can satisfy real time process after the SIFT algorithm is optimized.

Better expansibility. SIFT is very convenient to combine with other eigenvector, and generate much useful information.

2) Kalman Filter

Kalman filter technique is used to estimate the state of a linear system where state is assumed to be distributed by a Gaussian. In 1960, R.E. Kalman published his famous paper describing a recursive solution to the discrete-data linear filtering problem. Object tracking is performed by predicting the object's position from the previous information and verifying the existence of the object at the predicted position. Secondly, the observed likelihood function and motion model must be learnt by some sample of image sequences before tracking is performed . The Kalman filter is a set of mathematical equations that provides an efficient computational (recursive) means to estimate the state of a process in several aspects: it supports estimations of past, present, and even future states, and it can do the same even when the precise nature of the modeled system is unknown. The Kalman filter estimates a process by using a form of feedback control. The filter estimates the process state at some time and then obtains feedback in the form of noisy measurements. The equations for Kalman filters fall in two groups: time update equations and measurement update equations. The time update equations are responsible for projecting forward (in time) the current state and error covariance estimates to obtain the a priori estimate for the next time step. The measurement update equations are responsible for the feedback. That is used for incorporating a new measurement into the a priori estimate to obtain an improved a posteriori estimate. The time update equations can also be thought of as predictor equations, while the measurement update equations can be thought of as corrector equations.

3) CAM Shift

It is Continuously Adaptive Mean Shift Tracking. It is based on an adaptation of Mean Shift that, given a probability density image, finds the mean (mode) of the distribution by iterating in the direction of maximum increase in probability density. Unlike Mean Shift that uses Static Distributions, it uses continuously adaptive probability distributions (that is, distributions that may be recomputed for each frame). It is one of the simplest methods and supplies reliable and robust results, if the colors in the background differ significantly from those in the target object.

- Set the region of interest (ROI) of the probability distribution image to the entire image.
- Select an initial location of the Mean Shift search window. The selected location is the target distribution to be tracked.
- Calculate a color probability distribution of the region centered at the Mean Shift search window.
- Iterate Mean Shift algorithm to find the centroid of the probability image. Store the zeroth moment (distribution area) and centroid location.
- For the following frame, center the search window at the mean location found in Step 4 and set the window size to a function of the zeroth moment. Go to step 3.

4) ABC shift

It is acronym for Adaptive Background CAM shifts tracking. CAMSHIFT was designed for close range tracking from a stationary camera and failed when color of tracked object resembled its background. ABC shift achieves robustness against camera motion and other scene changes by continuously relearning its background model at every frame. It tracks efficiently even when the object shares common color with the background. Therefore, it enables robustness in difficult scenes where the tracked object moves past backgrounds with which it shares significant colors.

- Identify an object region in the first image and train the object model, P(C|O).
- Center the search window on the estimated object centroid and resize it to have an area r times greater than the estimated object size.
- Learn the color distribution, P(C), by building a histogram of the colors of all pixels within the search window.
- Use Bayes' law, equation (1), to assign object probabilities, P(O|C), to every pixel in the search window, creating a 2D distribution of object location.
- Estimate the new object position as the centroid of this distribution and estimate the new object size (in pixels) as the sum of all pixel probabilities within the search window.
- Repeat steps 2-6 until the object position estimate converges.
- Return to step 2 for the next image frame.

5) PAM shift

It is short for Path Assigned Mean shift tracking. It is a fast tracking method. In PAMS assignment, all points along the path towards the mode point are assigned to that final mode value. Points already assigned modes are eliminated from the mean shift process and are not traversed in the future. Since large swathes of feature space vectors are now assigned in one iteration step, the complete mean shift process converges much faster.

- Select a point site (p, q) at random in the image.
- Extract the colour values vector of the pixel at that point IU,V(p,q).
- Find the j neighbourhood vector, I(U,V)j(t), within the colour bandwidth, hc.
- Compute the Center of Mass, CoMc in the colour domain.
- Translate by the mean shift vector, mc(U, V).
- Repeat 3 and 5 till convergence to stationary mode vector,I(Um,Vm). Assign the final mode vector,I(Um,Vm), to the entire mean shift path, U t=i t=0 I(U,V)j (t), where i is the number of iterations to convergence.

6) The Mean Shift Algorithm

The Mean Shift algorithm is a robust, non-parametric technique that climbs the gradient of a probability distribution to find the mode (peak) of the distribution (Fukunaga, 1990). Mean Shift was first applied to the problem of mode seeking by Cheng (1995). Particle filtering based on color distributions and Mean Shift is described by Isard and Blake (1998) and extended by Nummiaro et al. (2002). Kernel based object tracking (including adaptive scale and background-weighted histogram extensions) is described by Comaniciu et al. (2003). CamShift is primarily intended to perform efficient head and face tracking in a perceptual user interface (Bradski, 1998). It is based on an adaptation of Mean Shift that, given a probability density image, finds the mean (mode) of the distribution by iterating in the direction of maximum increase in probability density (Intel Corporation, 2001). The primary difference between Cam Shift and the Mean Shift algorithm is that Cam Shift uses continuously adaptive probability distributions (that is, distributions that may be recomputed for each frame) while Mean Shift is based on static distributions, which are not updated unless the target experiences significant changes in shape, size or color. Since CamShift does not maintain static distributions, spatial moments are used to iterate towards to mode of the distribution. This is in contrast to the conventional implementation of the Mean Shift algorithm where target and candidate distributions are used to iterate towards the maximum increase in density using the ratio of the current (candidate) distribution over the target.

II. Literature Review

Mahesh Kumar Chouhan [1] studied movable object tracking is an interesting subject in the field of video tracking and its applications. When object is moving then its tracking is a challenging task in much vision area. In movable object tracking two tasks are perform first detection and second following the object path. This paper introduces an adjusted background histogram with mean shift for movable object tracking. In the proposed method, the adjusted background-histogram (ABH) plays an important role to reduce the interference of background in target localization in mean shift method. In this propose method a formula is generated by transforming only the target model. The proposed method has good accuracy to track a moving object in successive frames under some difficulties such as appearance changes due to image noise, etc. The advantages of this method are faster convergence and correct location of object as compare to mean shift. We will employ this approach for increasing the capability of moving object tracking. Here we introduced a new proposed method i.e. mean shift with adjusted background histogram (MS with ABH) and that is use the background information. By this background information we can increase the object tracking performance of mean shift method. This proposed method reduces the relevance of background information. By this proposed method the target model histogram is transforms and its probability is reducing. This proposed method is reducing the number of iteration step of mean shift method and increase the object tracking accuracy. Future work consists in considering tracking of multi objects.

Marcus A. Brubaker [2] introduced the basic elements of modern approaches to pose tracking. Using the probabilistic formulation introduced in this chapter one should be able to build a state-of-the-art framework for tracking relatively simple motions of single isolated subjects in a compliant (possibly instrumented) environment. The more general problem of tracking arbitrary motion in monocular image sequences of unconstrained environments remains a challenging and active area of research. While many advances have been made, and the progress is promising, no system to date can robustly deal with all the complexities of recovering the human pose and motion in an entirely general setting. While the need to track human motion from images is motivated by a variety of applications, currently there have been relatively few systems that utilize the image based recovery of the articulated body pose for higher-level tasks or consumer applications. This to a large extent can be attributed to the complexity of obtaining an articulated pose in the first place. Nevertheless, a few very promising applications in biomechanics (Corazza, Muendermann, Chaudhari, Demattio, Cobelli, and Andriacchi, 2006) and human computer interfaces (Demirdjian, Ko, and Darrell, 2005; Ren, Shakhnarovich, Hodgins, Pfister, and Viola, 2005; Sukel, Catrambone, Essa, and Brostow, 2003) have been developed. The articulated pose has also proved useful as a front end for action recognition applications (Ning, Xu, Gong, and Huang, 2008). We believe that as the technologies for image-based recovery of articulated pose grows over the next years, so will the applications that utilize that technology.

Tushar S. Waykole [3] studied about Background Subtraction methods are widely used for detecting and tracking moving objects in videos. It is useful in many applications such as traffic monitoring, video surveillances. The accurate tracking and detection of moving object is the challenging aspect of such approach. This work proposes a general purpose method which combines the advantage of spatial-temporal differencing with the basic background subtraction method. The results are promising on the proposed method as compared with the basic model of background subtraction and mean shift method In this paper we have read about a new method for detecting and tracking of moving objects in video combining the basic background subtraction method is with the spatiotemporal analysis and the results are promising. We have evaluated the performance of the algorithm using sample videos. The performance is compared with the well known mean shift method and the basic background subtraction method. As can be seen, the proposed method is faster and efficient. Future scope will be to resolve the problems of locating multiple moving objects separately in the real time scenarios.

Miss. Sadaf Khan [5] The complex task in a running video is to efficiently track the target with respect to its vary in position. The Mean Shift algorithm can be applied robustly and quickly to track the object. Local binary pattern is the technique through which a neighbourhoods of a pixel is determined by its binary derivatives. So the LBP can be employed in the field of pattern and movement recognition. In the proposed work mean shift algorithm will effectively track the target object applied along with LBP and joint colour histogram in real time videos. In proposed work we are working on mean shift algorithm. Our emphasis will be on improvising the conventional mean shift algorithm with the application of LBP technique histogram for colour feature extraction. Further modification will be applied to LBP and colour histogram to make the object tracking more robust in comparison to the Simple RGB method along with the application for texture feature extraction and joint colour histogram and further with correlation.

Mr. Suraj R.Jaronde [6] not always possible to provide labeled data for training because it requires substantial human effort, expensive tests, and disagreement among experts. Labeling is not possible at instance level. To overcome these problem multiple instance learning (MIL) method is introduced which actively trained the data in online manner and combine with discriminative classifier which separate the object from its background and provide positive and negative bags. The fisher information criteria is used to train dataset in online manner which perfectly describe the label of positive content in positive label bag and negative content in negative label bags. The use of actively trained classifier helps to improve the efficiency of tracking object in motion. Thus we proposed a robust tracker based on online discriminative appearance model. We develop an online active feature selection approach via minimizing Fisher information criterion and show that this method could also be used to selectively sample the independent dataset used for Active Learning and query it for retraining. Integration of lane detection, trajectory learning and pedestrian detection are some other key.

III. CONCLUSION

According to literature review of papers and methods for object segmentation, concluded some advantage of mean shift algorithm over other algorithms.

- Mean shift is an application-independent tool suitable for real data analysis.
- Does not assume any predefined shape on data clusters.
- It is capable of handling arbitrary feature spaces.
- The procedure relies on choice of a single parameter: bandwidth.

REFERENCES

[1] Mahesh Kumar Chouhan, “Movable Object tracking by using Mean Shift method with Adjusted Background Histogram”, Volume 2, Issue 7, July 2012, ISSN: 2277 128X.

[2] Marcus A. Brubaker, “Video-Based People Tracking”, 2012.

[3] Tushar S. Waykole, “Detecting and Tracking of Moving Objects from Video”, International Journal of Computer Applications (0975 – 8887) Volume 81 – No 18, November 2013.

[4] “Moving Object detection and tracking in video streaming”, 2013.

[5] Xiaojing Zhang “Object Tracking Approach based on Mean Shift Algorithm”, JOURNAL OF MULTIMEDIA, VOL. 8, NO. 3, JUNE 2013.

[6] Miss. Sadaf Khan , “Object Tracking in Real Time Videos using Mean Shift Algorithm”, IJSRD - International Journal for Scientific Research & Development| Vol. 3, Issue 04, 2015 | ISSN (online): 2321-0613