Journal of Innovations in Engineering & Applied Science (JIEAS)
Vol.1, Issue.1, Nevember 2018, pp. 16-22
ISSN: xxxx-xxxx

Robust Approach for Multiple Object Tracking

Dr. P. P. Halkarnikar*, Mr.Ishwarappa R. K., Mr.Nareshkumar R. M.****
Department of Computer Engineering
Dr.D.Y.Patil Institute of Engineering, Management & Research, Akurdi, Pune-44
pp.halkarnikar@gmail.com*, ishwar.kalbandi@gmail.com, nareshkumarmustary@gmail.com****

Article Info

Article history:

Received Sept. 01st , 2018
Revised Oct. 12th, 2018
Accepted Oct. 30th, 2018

Keyword:

Association rules,
Multiple object tracking,
Occlusion detection

ABSTRACT

Tracking applications does video analysis in which object is detected and tracked over time. The applications like security system, traffic monitoring, and content analysis require not only object detection but robust tracking of detected object. Many tracking algorithms give misleading result when multiple objects are present in video frames. The performance degradation is due to overlapping or occlusion of objects with each other. In this paper we present the robust algorithm to handle these cases of multiple object tracking. For robust tracking different association rules are applied, while tracking the objects. Situations of object merge and split are handled by this algorithm in order to accurately track the objects. Association rules are developed for tracking the multiple objects from one frame to next frame. When association rules are applied total frame processing time increases when objects are merged or split, otherwise time taken is same as object detection under normal conditions. Linked list type memory structures are used to save memory required for saving attributes of objects for association. The linked list will also expand and collapse as number of objects changes in given consecutive video frames. Object attributes are stored as one node of the linked list along with its object ID. Split and merge conditions of objects are indicated by flag attribute. This linked list is updated for every video frame for tracking of the objects. The result of proposed system shows good result while tracking the multiple objects in situation like occlusion or overlapping with in a frame. The results on simulated video are listed in this paper.

1. INTRODUCTION

In all object tracking applications robust detection and tracking of object is essential. The applications like security system, radar, traffic analysis and video content analysis need separate individual object detection and its tracking. These tracking algorithms for tracking object should be able to deal with various dynamic situations in the video scene. These situations in the scene disturb stages of object detection and its tracking. Some of the situations are like dynamic background changing, ambient light conditions, changing environmental conditions or multiple objects creating clutter of the objects. These disturb the detection of object. Multiple objects present in frame appeared to be merged due to self-shadow or occlusion. [1] The objects appear split due to occlusion or pixels detected from object region. This will lead to the wrong object tracking. The objects appear cluttered either with background or with other objects present in frame. These situations make system to wrongly count the numbers of objects present in frame. In this paper we have proposed the algorithm to track the objects in video reliably in environment where multiple objects are present. To achieve this we propose to store object attributes like position, area and flags along with object ID number. Flags are added in attributes to handle situations like merging and splitting of objects present in the video frame. To store information of objects in terms of attributes, we propose linked list type data structure for all the objects detected in video frame. This information is stored as node of list when object enters the frame and updated per frame of the video. As number of objects increase the linked list will expand in memory to accommodate the objects present in video frame. As object get disappear from the frame, object ID and other information stored associated with that object will get deleted from the list which will free the memory occupied by the node of the object. This will increase the efficiency of memory utilized in the system. The comparison of linked list of consecutive frames is done according to the association rules proposed in this paper. Euclidian distance between the object position in current frame and previous frame is considered for continuation of the object ID. Distance within given threshold will continue the object ID of previous frame. The other rules are applied for comparison if distance is reduced or increased. The processing time per frame remain almost constant for normal situation and increases only if merge or split situation is detected in the frame. The robust tracking of multiple objects is gives good results in the applications such as security, traffic monitoring, mob monitoring and analysis of human behavior in group.

2. RELETED WORK

Tracking multiple objects with single video camera leads to additional issues which are different than single object tracking with video camera. The segmentation and tracking of the object get disturb due to clutter and occlusion with other objects present in frame. Reliability of the system depends upon how it handles these issues. Following prominent issues are found in literature. [2] [3]

Issues during multiple object tracking

1) Object Clutter.

2) Shadow clutter

3) Clutter with background or camouflage.

4) Overlapping of object or occlusion.

1) Objects Clutter: When two or more object come very close to each other, it is difficult to recognize them as separate objects. Then tracker looses the objects locations and re-initialization of parameters takes place.

2) Shadow clutter: Illumination angle of the object with source creates the shadow. In object detection shadow pixels are detected as moving object pixels, so treated as object in frame. Shadow has got two effects during tracking. First, shadow changes the shape of object. Trackers which uses template of object for detection get failed in this situation. Second, when shadow of one object fall on another object, both objects appear as one object in frame. This situation is called as shadow clutter.

3) Clutter with background: Sometimes object color and shape matches with background in such way that it is difficult isolate the object from the background.

4) Overlapping of object or occlusion: When number of object in camera frame gets increased, they overlap each other. Due to this one of the object get occluded by other. Partial or full occlusions of objects take place depending upon the percentage of overlapping. When partial occlusion is present objects involved get merged, there by changing individual shape. In full occlusion track of occluded object become difficult.

The issues above are addressed by many researchers in their research papers. Attempt to overcome these issues introduces extra computation and complexity. [4] Parameters required to detect and handle the issues are large as compared to normal single object detection. To avoid that either the issues are not covered or partial solutions are suggested in multiple object detection development. Fast and robust algorithm is open problem in this area.[5]

Kalman filter is a popular technique for object tracking. This can be extended for multiple objects tracking by employing Kalman filter array for each object. In [2] authors used Kalman filter for tracking. To avoid clutter due to shadow they have adopted RGB colour model for object detection. To handle occlusion condition template matching is done. Appearance feature vector is taken to exactly locate the objects. When partial occlusions take place Kalman filter loose the control. Hence Mahalanobis distance between features vectors are considered for the objects. As it is difficult to maintain this distance without reference, so authors have proposed this method for short duration. In [4] Kalman filter along with colour background model is used for object detection and tracking. Tracking is done through Kalman filter attached to individual blobs. For multiple objects tracking, array of Kalman filter parameters are used for individual object. Each pixels group is identified by separate ID. For detection of overlap between the objects pixels information is not used as it is not reliable source. Association of object from frame to frame is done by object ID allotted during detection phase. Appearance and exit of object is done by comparing number of object detected in consecutive frames. If number of objects is less then difference IDs are discarded from the list. If number of objects is greater than previous frame then new IDs allocated. In [5] Kalman filter and elliptical templates are used for human tracking. As multiple human objects are involved occlusion disturbs the tracking of human. To detect the occlusion a ratio based on visible objects and pixels within elliptical shapes is considered. Two thresholds are assumed for occlusion detection. If the ratio is within these thresholds then partial occlusion else complete occlusion is considered.

In paper [7] hybrid strategy is used for tracking objects. Object information as well as region information is combined for exact location of the objects. Interaction of both the information is used for finding objects in case of merge and split of pixel region due to occlusion. Association of region descriptor is carried out for identification of objects. Region segmentation is done to find out region of interest in next frame. Within this region of interest object identification is done to track exact object. This is multilevel or cascaded matching helps to increase reliability of tracker. To know similarity of objects Mahalanobis distance is calculated per frame, which consider variance of feature vector. Association between both the region vectors and object vectors help to increase the robustness of object tracing in case of clutter. In [8] sports videos are considered for multiple object tracing. Football video contains player which are tracked using Bayesian inference filter. The multiple objects are tracked as if individual object by particle filter and maintaining the track information of each object. When occlusion occurs instead of single track, merged objects are treated as single object and tracked independently.

3. PROPOSED ALGORITHM

This section deals with procedure of detecting multiple objects from the input video. The proposed algorithm deals with situation in which objects get merged, overlapped or split. The steps involved for tracking in this algorithm are divided as object segmentation, validation and feature association. The tracking of objects is done by association of object and their properties.

A. Segementaion of objects:

The background subtraction module separates the background steady pixel and dynamic foreground pixels. There is no information about the object detected in the region. Therefore it is necessary to devise the algorithm which can distinguish number of objects present separately. Major hurdle in separating the multiple objects are shadow, occlusion and overlapping of moving objects. Gaussian mixture model with shadow removal [9] shows good result but unable to handle the situation of occlusion and overlapping. Occlusion and overlapping generates the object merging. Sometime object pixel blob get separated by stationary object or moving objects which creates split situation. To handle situations of split and merge condition following algorithm is proposed.

B. Registration of objects:

After extracting foreground pixels which are detected by frame differencing of consecutive frames. Each object is detected as homogeneous region of connected pixels. The properties of each object is collected and compared with properties of the objects stored in previous frames to track the objects. The properties or descriptor used for registering the objects are centroid, area, size, shape, vertices of the edges etc. either all or some of the descriptor are used to trace the object under different situations. The application of the object detecting system decides the number of descriptor used for verification of the object. For faster detection of objects it is essential to use optimum number of descriptor for object registration. The objects are labeled and data association is carried out from frame to frame to identify exact object. In proposed system each object detected in scene is marked by object ID. Every object ID is associated with feature vector consisting of object descriptor. To save the descriptor of objects in memory linked list data structure is used. It helps in flexible data storage and faster comparison of the feature. As number of objects appearing and disappearing from the frame it is necessary to vary the storage space. Therefore linked list is preferable over fixed size data storage. Object tracking is two stage processes. In first step object descriptor are projected from current frame. The projected features are compared and object are identified in second stage and updated for next step. For each frame n, objects are defined by object partition Π_o^n , and object properties are defined by Π_ϕ^n .

The information of current frame 'n' of objects present Oi(n) is given by descriptor Φi(n). This operation can be defined as object descriptor projection. The object descriptor is defined as

$$\Phi_i = (\Phi_i^1(n), \Phi_i^2(n), \Phi_i^3(n), \ldots \ldots, \Phi_i^k(n))$$

………. (1)

Where k are number of features describing the object i at given frame n. The number of feature selected will decide the complexity of object detection.

Let object ID is $\Phi_i^1(\mathrm{n})$, the centroid of the object,$(\Phi_i^2(n), \Phi_i^3(n))$, the area of the object $\Phi_i^4(n)$, and the mean displacement of the object $\Phi_i^5(n)$ ….(2)

The number of feature selected and type of features selected depends upon the application to be developed. More features will increase the complexity of execution and requires more memory space for storage. The accuracy of object tracking will enhance with more feature registered for comparison. An updated object descriptor in next frame is defined as

$$\Phi_i(n+1) = (\Phi_i^1(n+1), \Phi_i^2(n+1), \Phi_i^3(n+1), \ldots\ldots \Phi_i^K(n+1)) \qquad \ldots\ldots (3)$$

The updated values of features become initial values of next frame during processing, which reduces the memory space requirement for the feature vector.

C. Feature association :

The correspondence of the video objects in successive frames is achieved through association of object features. Tracking the object with feature association leads to the flexible technique that overcomes the problem of occlusion and merging of the objects due to proximity. Given the object feature of new frame and object feature of previous frame the proposed method performs two tasks

i) It defines the association between the object feature in current frame (n+1) and previous frame (n).

ii) It provides an effective initialization of the features comparison in successive frames.

iii) It does object segmentation and validation between consecutive frames and assigns proper object ID to each object in frame.

The object segmentation and validation between consecutive frames is achieved by object ID defined. Feature association operates at low level and validate feature with specific descriptor. The two steps generate final correspondence between the objects present in consecutive frames.

D. Object segmentation validation:

Object segmentation validation initializes the object attribute collection. Every separate object is given the object ID which helps in increases the accuracy of object detection in case of merge and split of the objects. For association of the objects features of current frame are compared with feature of previous frame. As generally frame to frame variations are not large unless drastic condition takes place. The feature vector of current frame becomes values for comparison for next frame. It is stored as linked list structure defined by object ID. Only two linked list are sufficient for tracking of objects from frame to frame to detect the changes. If numbers of linked list are increased then array of the entire previous linked list can be compared with correct frame number for more accurate detection of objects. In this context a new object identify as new node in linked list. This node is not associated with any tracked object as new ID is allotted for it. Here linked list structure gives many advantages over other storage techniques as space requirement are defined dynamically and comparison can be done on any attribute as per rule of association.

E. Object Tracking:

The projection feature of all the objects in next frame are not taken for association but only $(X_i , Y_i)^n$, A_i^n and d_i^n are considered for comparison. Then tentative correspondence of the objects in frame n and next n+1 is established with collected objects of next frame that is Π_o^{n+1}. For convenience in our proposed work we have taken centroid, area and mean distance of objects in consecutive frames for association

$$\Phi_i(n) = [(X_i, Y_i)^n, A_i^n, d_i^n] \quad \ldots\ldots(4)$$

As stated above number of features will change according to the application of system. The feature use for tracking may get modified due to viewing angle, illumination changes or motion change of the objects. Updated feature descriptor is defined as

$$\Phi_i\,(\mathrm{n+1}) = [X_i', Y_i', A_{i,}', d_i'] \quad \ldots\ldots(5)$$

The correspondence is established as

$$[(X_i', Y_i')A_i', d_i', n+1)]: [(X_{i\,,}Y_{i\,,})A_{i,}\ \ d_i, n] \text{ …….(6)}$$

This procedure works well if objects are correctly segmented and remain same in each frame. But in reality the objects may appear merge due to occlusion. Some objects split due to occlusion by stationary objects. Sometimes due to noise in image object gets split into different blob which is temporary phenomena and may get original object in short time. An occlusion takes place when two or more objects overlap each other or by getting close to each other. All the possible situations can be detected by conditional validation of the object features.

F. Association rules for object detection :

For applying the rules for detection of different situations creating clutter, blobs are considered as smallest unit. Blob is set of connected pixels in given frame. This can be treated as separate objects.
Blob are formed by following situations in video

1) Single separate object.
2) Multiple object merge together due to proximally or occlusion.
3) Part of object split due to occlusion by stationary object.

Following rules are formed to detect object in above situations. Every object is identified by unique object ID.

Rule I

For identification of object minimum distance between objects in consecutive frames is considered as first condition. Values of centroid stored are used for calculation of distance in all the objects present in next frame.

$$d_i = \sqrt{(X_i - X_i')^2 + (Y_i - Y_i')^2)} \text{ …………(7)}$$

Out of the distances calculated with each object the smallest is considered for object similarity.
If $d_i < Th_1$ Then $O_i(n+1) = O_i(n)$
else
new object initiated $O_{k+1}(n+1)$. ……..(8)
if distance is within threshold then object is matched and tracked with same object ID. No other feature is tested as object ID remain same. Only updating of initialization values take place. Threshold depends upon the application and motion factor of the object.

Rule II

If (d(O_i^n, O_j^n) < Th_1) AND ($O_i^n, O_j^n \epsilon O^n$)
Then condition of object merged or merged may have occurred. We have to compare (A_i ,n) : (A_i,n+1).
If (A_i(n) < A_i(n+1)) AND (A_j(n) < A_j(n+1))
then split flag is set for object ID i, j in frame n+1.
If (A_i(n) > A_i(n+1)) AND (A_j(n) > A_j(n+1))
then merge flag is set for object ID i, j in frame n+1.
In both merge and spilt condition new object ID generated in frame n+1. This rule verifies the condition of partial or full occlusion. To test whether object disappeared from frame separate timer is attached. When time expires object is deleted from list and new object created is taken as true object from frame n+1. In this case rule I will be applied from next frame. If merge or split flag is set comparisons are carried out according to rule III.

Rule III

If merge flag is set then $O_i^n \notin O^{n+1}$, but it is not immediately removed from list of objects. Timer object is attached to this object to test if it is disappeared or occlude by some other objects. if object reappears flags get reset. New object due to rule II will follow rule I and rule II as if independent object in list. When timer attached to object expires, object will be removed from list. If flags are set all rules are applied to that object. In case of clear flag condition only rule I is sufficient to track the object.

Applying above association rule objects are tracked using matching of objects from current frame with existing objects in previous frame.

4. STORAGE STRUCTURE

To improve the tracking of the objects from frame to frame, linked list type of object storage is used. Advantage of linked list is that it is linear data storage hence depending upon the number of objects present the length of linked list can be varied which save the memory space. It also keeps the algorithm expandable to match with number of objects present in frame. List accommodates dynamic existence of object due to merge and split. Objects inserted in list are consecutively connected if they are spatially together. Every node of the linked list has header and feature vector stored. Node have separate field for every attribute, flag and the pointer point to next node in the list. If new object is detected it is added to the end of the list. If object present in the list is spilt or merge additional node are generated with new object ID. Merge flag is set and it is inserted next to the node having distance closed to the existing object. Note that new object introduced has object ID next to last object present.

5. EXPERIMENTAL RESULTS

In this section we have concentrated on the condition of overlapping of objects causing clutter of objects. To handle situation of occlusion it is necessary to keep track of individual object. The pixel blob formation during detection stage indicate single blob as objects are merged. By keeping track of object properties and applying association rule it is possible to keep track even in occlusion as discussed in previous sections. To test this algorithm we used simulated video. Profiles of this video is as shown in table I. Experiments are carried out offline on Intel i3 processor with 2 GB ram.

TABLE I

Name of Video	Frame Size	Frame Rate	Duration
T_m2o2.avi	320x240	15fps	0.08 sec

The video contains to object moving with constant speed. Split and merge conditions are simulated for testing. Figure 1 show the video frames of simulated video. The split and merge conditions are simulated and algorithm response tested for number of object detected.

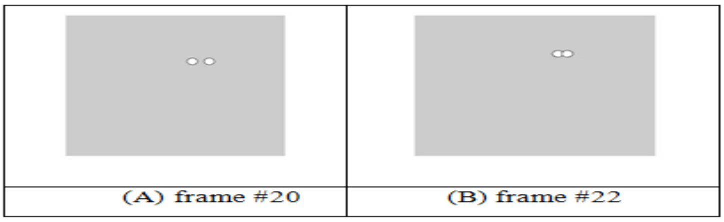

Figure 1 (A) Sample frame 20 showing normal moving objects. (B) Sample frame 22 showing partial overlapping.

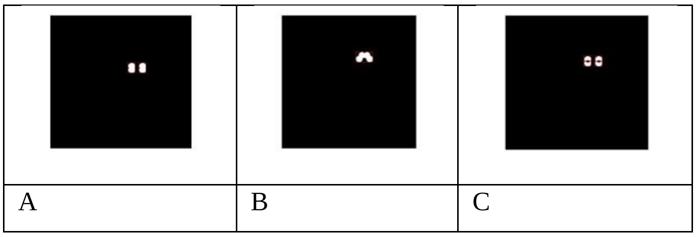

Figure 2(A) object blog detected using background subtraction algorithm. (B) Merged Flag set as object blog is one. (C) Split condition due to discontinuity in blog set split flag.

Results are verified with keeping the track of no of objects present in the frame. The association rule proposed above is applied with in area of interest instead of entire frame. Figure 2. show the video frames of simulated video. The split and merge conditions are simulated and algorithm response tested for number of object detected. The table II show contains of linked list with object ID for some sample conditions. Table show results of applying proposed association rule for detection of split and merge on simulated video of two objects moving in same direction with constant speed. Though new object created by split are counted they are not finalized as split flag is set. Similarly during merging of objects count get reduced but not finalized as merge flag is set. The setting of flags can be used for detection of anomalies in surveillance.

6. CONCLUSION

In this paper we propose the robust algorithm for tracking multiple objects in video. The conditions like occlusion and cast shadow disturb the detection process when multiple objects are present. Merge and split of object blob formed by detected pixels take place. Therefore it becomes difficult to track the object in tracking stage. Here we experimented efficient association rule technique to keep track the objects. The conditional application of association rules removes the overburden on tracking process. The results show good track during split and merge of objects.

REFERENCES

[1] D. Parks and S. Fels, "Evaluation of Background Subtraction Algorithms with Post-Processing," Proc. IEEE Int'l Conf. Advanced Video and Signal Based Surveillance, pp. 192-199, 2008.

[2] Hieu T. Nguyen, Arnold W.M. Smeulders, “ Fast Occluded Object Tracking by a Robust Appearance Filter”, IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 26, No. 8, pp 1099 – 1104, August 2004.

[3] Auguste Genovesio, Jean-Christophe Olivo-Marin, “Split and Merge Data Association Filter for Dense Multi-target Tracking”, Proceedings of the 17th International Conference on Pattern Recognition 2004 (ICPR’04), Vol. 4, pp. 677-680, August 2004.

[4] Jong Sun Kim, Dong Hae Yeom, and Young Hoon Joo, “Fast and Robust Algorithm of Tracking Multiple Moving Objects for Intelligent Video Surveillance Systems”, IEEE Transactions on Consumer Electronics, Vol. 57, No. 3, pp. 1165 – 1170, August 2011.

[5] Tao Zhao, Ram Nevatia, “Tracking Multiple Humans in Complex Situations”, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26, No. 9, pp. 1208-1222, September 2004.

[6] P. P. Halkarnikar and Dr. S.N. Talbar, “Multiple Object Detection in Outdoor Videos: Issues and Methodology”, IRACST - International Journal of Computer Science and Information Technology & Security (IJCSITS), Vol. 3, No.1, pp. 96-100, February 2013.

[7] V. Venkataraman, G. Fan, J. P. Havlicek, Xin Fan,Yan Zhai, M. B. Yeary, “Adaptive Kalman Filtering for Histogram-Based Appearance Learning in Infrared Imagery”, IEEE Transactions on Image Processing, Vol. 21, No. 11, pp. 4622-4635, November 2012.

[8] Emilio Maggio, Elisa Piccardo, Carlo Regazzoni, Andrea Cavallaro, “Particle PHD Filtering for Multi-Target Visual Tracking”, IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP 2007), Vol. 1, pp. 1101-1104, 2007.

[9] Hinz, S., "Detection and counting of cars in aerial images," IEEE International Conference on Image Processing, 2003. (ICIP 2003), vol.3, pp. 997-1000, 14-17 Sept. 2003.