

# Problem Analysis of Multiple Object Tracking System: A Critical Review

Md Zahidul Islam<sup>1</sup>, Md Shariful Islam<sup>2</sup>, Md Sohel Rana<sup>3</sup>

Associate Professor, Dept. of Information & Communication Engineering, Islamic University, Kushtia, Bangladesh<sup>1,2</sup>

Lecturer, Dept. of Computer Science & Engineering, Uttara University, Dhaka, Bangladesh<sup>3</sup>

**Abstract:** Simultaneous tracking of multiple objects is a state-of-art problem in the field of computer vision. The problems of tracking single object and tracking multiple objects are not same. For tracking multiple objects, lots of problem can arise due to abrupt object motion, multiple object interaction, drifting of object etc. Some recent work has proposed solution to minimize those problems. The main goal of this article is analyzing the recent work in multiple objects tracking to handle those challenges. Here, Different proposed penalty functions are discussed to handle the multiple objects tracking problem. We also comprehensively discussed contribution and limitation of recent advancement in the field of multiple object tracking.

**Keywords:** Tracking problems, Tracklet, Hijacking problem, Drifting problem.

## I. INTRODUCTION

Multiple object tracking has attained growing attention to the current researcher. With the development of camera quality and efficiency tracking get more attention to the researcher. Recently, rigorous advancement has attained, although still there is lots of problems remain unsolved. There are many difficulties to track an object. Those are listed below:

- Random motion of object
- Partial or full occlusion of object
- Change of illumination or background appearance changes
- Object articulation

There two basic steps for object tracking. Those are:

- 1) Object segmentation
- 2) Object tracking

In object segmentation stage, the object is segmented from its background. Many methods are used for object segmentation. In tracking system, the mostly used object segmentation method is graph-cut method [1]. In object tracking steps, segmented objects are tracked in each consecutive video frame.

For tracking multiple objects, the system has to face each problem individually for each object in the scene. Moreover some others difficulties has to handle, those are enlisted in figure 2. For multiple object tracking the tracker has to be robust than it can avoid the occlusion or effect from other objects in the scene. For huge number of object the systems can be too slow. Specially, when considered a tracking system with optimization for each tracker.

For tracking multiple objects, the tracker has to handle both issues for single object tracking and multiple object tracking. Many other reviews show the both tracking single object and multiple object tracking problem combinedly. Tae-Kyun Kim et el. classified both single object and multiple object tracking problem [2]. But, our review described and analyzed the unique problem of

multiple objects tracking system. The paper is focused on to briefly describe those problems and analyze the researchers work to overcome those issues.

The paper is organized as: section 2 describes the different object representation techniques for tracking, section 3 describes the classification of multiple object tracking problems, and section 4 describes the recent contribution to handle the situation of MOT problems.

## II. OBJECT REPRESENTATION

For development a robust tracking system, object representation is very important. Many representation techniques are used. Figure 1 show recent object tracking systems that is attracted to the researchers.

Below describes various object representation methods.

- Shape based representation: Object shape is an important cue to represent an object for tracking application. For Histogram oriented gradient uses gradient cue to represent the object. Object contour or silhouette is also used for tracking the object.

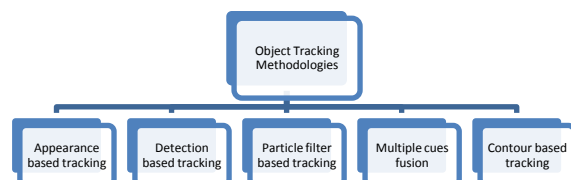


Fig1. Object tracking methodologies

- Appearance based representation: Object shape tracking is not always efficient for tracking object. Especially for the articulating object like animals or human shape can change in random manner. Color cues give better performance for such type of object tracking. In image processing, Object is represented in RGB color space. However, HSV color space is more uniform, but also noise sensitive. Color histogram is used to match the object with reference object.

- Motion based tracking: For articulated object, object motion is very effective cue for tracking. Optical flow is better for non rigid classification. Residual flow can be used for calculating objects rigidity, like human has higher residual flow.
- Detection based tracking: Detection based tracking means object detection at each frame of video sequence. In case of detection based object tracking objects are detected in each frame. Then, detected object in each frame is link together to make the object trajectory. Here, if the detector failed to detect the object, then tracker will also fail to track the object. Basically, In case of multiple object tracking, detection based tracking is used to verify the object that has been tracked correctly.
- Multiple cues fusion: For robust tracking object tracking multiple cues like shape, color, motion are combined together where each cues has individual weights. Specially, particle filter based object multiple cues is fused to track the object more robustly.

III.MOT PROBLEMS

Using camera, tracking multiple objects faces many problems. The problems are categories in figure 2.

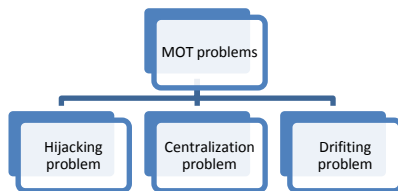


Fig 2: Multiple object tracking problems

Before going to details the phenomena of track lets and trajectory needs to discuss. Track let is segmented object traveling path that represent consecutive states of an object. By joining all track lets, objects trajectory is found. Figure 3 shows the track let and trajectory representation of an object.

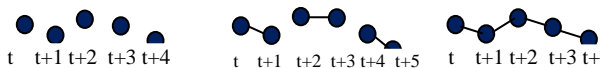


Fig 3: States of object in consecutive frame from t to t4 (left), track lets of an object (middle), trajectory of an object.

A. Hijacking problem

If two similar colour objects come too close, the tracker of one object can jump to associates object. This problem is called hijacking problem. Figure 4(a) shows the hijacking problem. Here, two similar colour objects come closer and one of the tracker jumps to other object. For this reason, the tracker misses the correct object to track.

B. Centralization Problem

Centralization problem occurred when two objects centralized into same position and the tracker misses to track the object each object individually when those object split from each other. Figure 4(b) shows the centralization problem. Here, two objects overlapped each other, when those object detached; the tracker can detect the wrong object.

C. Drifting problem

Drifting problem occurred when an object abruptly changes its direction to reverse. In that case, it's become very difficult to track the object; because the motion model doesn't work. Figure 4(c) shows the object drifting problem. Here, the dotted rectangle represents predicted object location in next frame that is predicted by object motion model.

(a) Hijacking problem



(b) Centralization problem



(c) Drifting Problem



Fig4. Multiple object tracking problems

IV.RECENT WORKS AND LIMITATIONS

Konrad Schindler et el. proposed energy based multiple objects tracking system [3]. There they proposed energy based tracking system. The energy defines the object trajectories. Energy equation is constructed summation of the five models. Those models are: observation model, dynamic motion model, mutual exclusion model, persistence model and regularization. The system tries to minimize the energy of to track the object correctly. Exclusion model is used to handle hijacking problem.

The energy function of exclusion model is:

$$E_{exc}(X) = \sum_{t=1}^F \sum_{i \neq j} \frac{s_g^2}{\|x_i^t - x_j^t\|^2} \tag{1}$$

Here,  $s_g$  is object scaling factor.  $x_i^t$  is the states of i-th object and  $x_j^t$  is the states of j-th object. When two object comes to close energy functions value comes close to infinity. In that case, the system tracks an identical object.

Although the system is good for restore the tracker and automatic initialization of tracker. But using this system it's difficult to track two associate objects moving together. Another problem of this tracker is drifting problem. If an object drifts to fast then, the tracker misses the object to track. The major drawbacks of track let based tracker is that the system cannot track the object continuously. Hence, if the tracker start to track two same objects, sometimes it can't be possible to identify which object is refers to which trajectory. Because, the track let can break and lost the trajectory. But, most significant issues that improved in this paper are that, if once the object is initialized, the tracker can track the object smoothly.

Chil-Woo Lee et el. give a good solution to remove centralization and hijacking problem [4]. They have explained two different areas when two objects come closer. One is interacting area and other is contacting area. They proposed Penalty function in particle filter based

multiple objects tracking system. They reduced the weight of a particle when the neighbouring object comes closer to objects. Their penalty function  $h(x_t^i)$  for a particle  $x_t^i$  is defined as:

$$h(x_t^i) = 1 - (D - g(x_t^i))D^{-1} \quad (2)$$

Here,  $D$  is the maximum distance of interaction area and function  $g$  is the distance of the neighbouring target. The proposed system is good for two interacting object when objects. But, this system is not good to handle drifting problem.

Mario Sznai et al. proposed similar appearance multi target tracking based on objects motion model and tracklets [5]. They generated short tracklets on the basis of non-conflict object detection. Two objects will be non-conflicting if the ratio of their distance is less than 0.3. They proposed an equation which calculates the similarity between two tracklets. The similarity between two tracklets  $\alpha_i$  and  $\alpha_j$  is defined as  $P_{ij}$ :

$$P_{ij} = \begin{cases} -\infty & \text{if } \alpha_i \text{ and } \alpha_j \text{ conflicts} \\ \frac{\text{rank}(H_{\alpha_i}) + \text{rank}(H_{\alpha_j})}{\min \beta_i^j \text{rank}(H_{\alpha_{ij}})} - 1 & \text{otherwise} \end{cases} \quad (3)$$

Here,  $H_{\alpha_i}$  is Hankel matrix that represents the sequence of measurements of tracklets  $\alpha_i$ .  $\alpha_{ij}$  is joint tracklets where  $\beta_i^j$  is padding between  $\alpha_i$  and  $\alpha_j$ . If two object follows same trajectory then  $P_{ij}=1$ . If two objects refers different tracklet, then  $\text{rank}(H_{\alpha_{ij}}) \leq \text{rank}(H_{\alpha_i}) + \text{rank}(H_{\alpha_j})$ . The proposed system can track the full trajectory of an object and doesn't interchange the tracker, if two objects cross each others. But, if any object changes its motion randomly, then it's become difficult to track the object smoothly.

Sinisa Todorovic et al. proposed an multiple object tracking system where they formulate data association for by finding maximum-weight independent set of the graph [6, 9]. They detected objects into two consecutive frames. Here, in graph every node is made by matching between two consecutive frames. Match score defines the weight of the nodes. Nodes edges are those nodes whose violate constrain. Then, maximum weight independent set is applied and the graph nodes are learned for multiple object tracking. Here, multiple object tracking problems are handled by learning object appearance and motion model. They showed how to handle long term occlusion by using the algorithm iterative maximum weight independent set of the graph.

Afshin Dehghan et al. proposed a multiple object tracking system where targets identity based network flow method is used to track multiple objects [7]. Here structured learning and inference is used to detect an object. They have used spatial constrains optimization to track highly overlapped object. They defined observation edges as equation (4).

$$p_{ij}^{q+1} = [p_{ij}^q + \theta^q [y_i^t \cap y_j^t - 0.5]^+ e^{((y_i^t \cap y_j^t) - 0.5)/2}]^+ \quad (4)$$

Here,  $y_i^t \cap y_j^t$  is the overlap between neighbouring bounding boxes in same frame.  $p_{ij}^q$  penalize the observation which is associated with cost  $c_{ij}$ . The spatial constrain  $c_{ij}$  is update at each iteration.

Siyu Tang proposed subgraph multicut detection based tracking system [8]. Subgraph multicut has the property that can define the association and temporal problem combinedly. Here, shortest path algorithm is used to remove the overlapping of tracklets. For getting optimal hypothesis, set of detection is removed from a set of detection.

## V. RESULT EVALUATION

Accuracy of multiple object trackers (MOTA) is measured by equation:

$$MOTA = 1 - \frac{\sum_t f n_t + f p_t + m m_t}{\sum_t g t_t} \quad (5)$$

Where  $f n_t$ ,  $f p_t$ ,  $m m_t$  and  $g t_t$  are false negatives, false positives, mismatches and ground truth at frame  $t$ , respectively.

Table 1 shows MOTA tracking percentage for different tracking system. Here, the video sequence is considered for TUD human crossing sequence. This video sequence is a good choice to find out effectiveness of a tracking sequence, spatially when considered hijacking or centralization problem is one of the major issues for tracking. It shows the recent two works that is target identity network flow [7] and subgraph decomposition methods shows [8] better results to track crossing people each others.

			MOTA
Anton	Andriyenko	Energy minimization method [3]	60.5%
Caglayan	Dicle	Iterative Hankel Total Least Squares method [5]	88.0%
Brendels	maximum weight independent set method [6]		85.9%
Afshin Dehghan's	target identity network flow method [7]		92.9%
Siyu Tang	subgraph decomposition method [8]		80.9%

Table1. TUD crossing video sequence

## VI. CONCLUSION

This paper presented a critical review of multiple object tracking problems. The review illustrated the key issues of multiple object tracking and recent development of MOT problems. Although some good achievement has gained to the development of multiple objects tracking system, there are still some issues remained unsolved. Tracking during crossing object or during object comes too close has been improved its result. Specially, Afshin Dehghan proposed the best method which has MOTA accuracy about 92.9% for crossing people. But, still it's too difficult to track an object if the object moves so fast and changes its direction randomly. Recent works has shows the algorithm can handle very smoothly if two objects cross each other. It

still be difficult to track an object if the situation is like that the two similar object drifts after centralization. These issues are remaining unsolved and still there is lots of task has to do to acquire more precision in different situation of tracking. To find out the best motion model for each object in scene is the main issue to handle the entire MOT problem. To improve the motion model, we recommend to use tracklet based motion model combinedly with intra-frame motion change cues, that will improve the motion model and can be effective to handle MOT problems.

### REFERENCES

- [1] Chari, Visesh, et al. "On Pairwise Costs for Network Flow Multi-Object Tracking." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015.
- [2] Luo, Wenhan, Xiaowei Zhao, and Tae-Kyun Kim. "Multiple object tracking: A review." arXiv preprint arXiv:1409.7618 (2014).
- [3] Andriyenko, Anton, and Konrad Schindler. "Multi-target tracking by continuous energy minimization." Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011.
- [4] Oh, Chi-Min, et al. "Centralization problem of contacting interaction in multiple object tracking." Frontiers of Computer Vision (FCV), 2015 21st Korea-Japan Joint Workshop on. IEEE, 2015.
- [5] Dicle, Caglayan, Octavia Camps, and Mario Sznai. "The way they move: tracking multiple targets with similar appearance." Computer Vision (ICCV), 2013 IEEE International Conference on. IEEE, 2013.
- [6] Brendel, William, Mohamed Amer, and Sinisa Todorovic. "Multiobject tracking as maximum weight independent set." Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on. IEEE, 2011.
- [7] Dehghan, Afshin, et al. "Target identity-aware network flow for online multiple target tracking." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015.
- [8] Tang, Siyu, et al. "Subgraph Decomposition for Multi-Target Tracking." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015.
- [9] Brendel, William, and Sinisa Todorovic. "Video object segmentation by tracking regions." Computer Vision, 2009 IEEE 12th International Conference on. IEEE, 2009.

### BIOGRAPHIES



**Md Zahidul Islam** has received his B.Sc. and M.Sc. degree from the Department of Applied Physics and Electronic Engineering, University of Rajshahi (RU), Bangladesh in 2000 and 2002 respectively. In 2003, he has joined as a Lecturer in the Department of Information & Communication

Engineering, Islamic University (IU), Kushtia, Bangladesh. He has done his Ph.D research on Visual Object Tracking System from the Department of Computer Engineering at Intelligent Image Media and Interface Lab, Chonnam National University (CNU), South Korea. In August 2011, Dr. Islam has been successfully awarded his PhD from the same department. Besides, he has done his research internship in 3D Vision Lab in Samsung Advanced Institute of Technology (SAIT), Suwon, South Korea. Dr. Islam has also other research interests like computer vision, 3D object, human and motion tracking and tracking articulated body and genetic algorithm etc. Currently he is an Associate Professor and head of the Department of Information &

Communication Engineering, Islamic University (IU), Kushtia, Bangladesh.



**Md. Shariful Islam** received the B.Sc. and M.Sc. degrees from the Department of Applied Physics, Electronics and Communication Engineering (Erstwhile Electronics & Applied Physics) from Islamic University, Kushtia-7003, Bangladesh, in 2001 and 2003

respectively.. He joined as a lecturer in the Dept. of Information and Communication Engineering, IU, Kushtia-7003, Bangladesh, in September 2004 and currently, he is working as an Associate Professor in the same department. Now, he is doing his PhD work on Integral Imaging Three-dimensional display (3D) system. His research interests are in the area of Integral Imaging 3D display technology, Image processing.



**Md Sohel Rana** has completed his M.Sc. in 2015 from the Department of Information & Communication Engineering, Islamic University, Kushtia, Bangladesh. He completed his B.Sc. from the same university at

2013. He is currently working in Visual Object Tracking. He also has research interest in Machine Learning, Pattern Recognition and Medical Imaging. Currently, he is a lecturer of the department of Computer Science and Engineering, Uttara University, Dhaka, Bangladesh.