International Journal of Recent Advances in Multidisciplinary Topics
Volume 3, Issue 4, April 2022
https://www.ijramt.com | ISSN (Online): 2582-7839

Algorithms for Object Tracking Under Occlusion Using Multiple Cameras – A Survey

B. C. Bharath[1*], Latha Anuj[2]
[1]PG Student, Department of Information Science and Engineering, Dayananda Sagar College of Engineering, Bangalore, India
[2]Assistant Professor, Department of Information Science and Engineering, Dayananda Sagar College of Engineering, Bangalore, India

***Abstract*: Computer Vision based object tracking are gaining more demand in the field of video surveillance. However, tracking of multiple moving objects under occlusions is inefficient with single camera-based setups. Multiple item tracking under long-term or even complete occlusion is a time-consuming process with unreliable results. Multiple camera-based multiple objects tracking with object re-identification in multiple cameras can be incredibly effective when dealing with occlusion situations. This paper reviews multiple approaches for detection and tracking using multiple-camera based setups.**

***Keywords*: Object detection, object tracking, occlusion, re-identification, multiple cameras.**

1. Introduction

The world has evolved to modern era, where anything and everything is digital. Gone are those days when one used to put physical effort in order to do a specific work. As the technology is evolved, the needs are also increasing, and with that need for securing things is also increasing. In applications such as monitoring in public places, traffic surveillance, military purposes, banks, railway stations, navy ports, malls, airports, highways, campuses, etc., surveillance plays a main role. The demand for technologies aiding detection of objects, and efficient tracking is increasing, they require systems to be precise in any conditions and the system must be robust to any changes in the given conditions.

Here let's consider surveillance of vehicles in city traffic and also highways conditions. Let's consider some uses of detection and tracking of vehicles: It eases the job of police department to identify any issues (accidents, traffic violations, etc) and attend them immediately, and even in the case of tracking down an accused driving a vehicle under surveillance; and in either case a delivery company (food, logistic, groceries, etc.) able to track their vehicles around the city for efficient and accurate deliveries. To achieve all the requirements, we must be able to cross all the challenges and build an effective model.

A. Challenges in Object Tracking

To help the persons to conduct object tracking in an effective manner one must develop and train a model which is efficient, accurate and robust in nature, by overcoming all the challenges in building a model. This section discusses some major challenges faced in object tracking.

To detect and track down objects, these surveillance models take inputs from video cameras placed at different positions in a city or a highway. The model developed will take that video input from specific camera and applies algorithm to the stream for detection. The first challenge is the quality of input video, if it is of poor-quality model must perform additional work to process it so that the input is converted to clear form so that it can be processed to detect objects.

Once the objects are detected, the objects must be assigned IDs with bounding boxes for identification purposes, as there is a chance of multiple similar objects being present in the coverage area of input. The system must classify the objects with specific categories (class) and give them IDs accordingly. Therefore, clear classification of objects is another challenge as the object of a certain class comes in different form and shapes.

The next challenge is deformation, since the objects in the ROI of camera tend to move frequently, hence, there is a change in shape of subject being tracked. It may result in identifying the object as a new one with new ID or also lead to misclassification of objects.

Apart from quality of the video input, the natural parameters like illumination conditions are also considered as a challenge; as lighting can have a large influence in classifying objects accurately and in some severe conditions, objects might not be visible to detect. The model must be able to handle changes in illumination conditions.

Sometimes the objects tend to be blunt with the environment or background conditions, the model must be able to apply appropriate filters to detect the object efficiently.

Some objects tend to move, and their speed varies for each one of them, the model must be able to identify objects moving at different speeds.

Occlusion is considered as the major challenge in object tracking. We can say occlusion can be caused due to three reasons, they are; first, it can be caused due to natural conditions such as severe illumination conditions; second, another reason is object moving out of the ROI of the camera input either for short or long period of time; third, even if the object is in the ROI of the camera input, the other objects in the scene may block the visual of the object to be tracked.

To overcome the challenge of occlusion as it is a major challenge in detection and tracking of objects, multiple-cameras

*Corresponding author: bharathbc37@gmail.com

must be involved in the process which are placed in either overlapping or non-overlapping conditions. Where the system must be trained to identify the objects which are in motion and track their path among different cameras and must also able to reidentify the objects with the same IDs assigned which is already tracked among different cameras.

2. Literature Survey

This section provides a brief of methods proposed by different authors in order to carry out multi-camera object tracking.

In an approach to detect vehicles in crossroads, Liwei Liu et al. [1], he presents a robust multi-view vehicle detection and tracking method in Cross roads. It's a real-time, online processing system that can successfully handle view alterations and occlusions. The dual-layer occlusion handling technique proposed here can handle both partial and full occlusions, and the two-stage view selection is efficient in fusing many detectors. Experiments in various weather circumstances (snowy, bright, and cloudy) show that our system is successful and efficient.

In another approach using multi-camera data fusion with machine learning, Hao Wu et al. [2], demonstrates a multi-camera vehicle detection system using an MVRPN/CNN pipeline. In real-world traffic scenarios, the suggested method detects partially and severely obstructed vehicles. To overcome the issues of vehicle tracking, a vehicle detection system that can use temporal video frames is being developed. To improve bounding-box predictions in future studies, a multi-view bounding-box regression will be incorporated into the process.

O. Messoussi et al. [3], uses the popular tracking-by-detection paradigm to present a MOT approach for tracking vehicles in urban environments in online use cases that integrates spatial information with re-identification attributes. On the UA-DETRAC benchmark, data association function design performs very well for long-term occlusions and substantial shifts due to fast motion.

Chih-Wei Wu et al. [4], proposes vehicle Re-ID system. Here they propose an adaptive feature learning technique based on the space-time before to harvest virtually limitless training samples from the target movies to solve the lack of labeled training data and visual domain mismatch between datasets. On the vehicle Re-ID dataset, the AFL approach has the same level of success. When they tested the concept using human Re-ID datasets and applied it to the vehicle Re-ID system.

D. Sudha et al. [5], Proposed a Hybrid Algorithm for Multiple Vehicle Detection and Tracking in a Day-Time Environment, which shows that the method proposed clearly estimates the information about the object region with the help of the machine aware system learning, particle grouping and bus topology approach, and also the working nature of the particle filter enables efficient processing of a non-linear system.

Jonah Ong et al. [6], develops a traceable 3D occlusion model, and obtained Bayesian algorithm for multi-view multi-object filtering which only requires monocular detector training, without depending on the configurations of multiple-cameras. By this algorithm multi-camera system can operate uninterruptedly in the case of extension of network or even while reconfiguring the system.

Reyes Rios Cabrera et al. [7], presents a framework for surveillance in tunnel with non-overlapping cameras. It has same function in all the cameras for detection, tracking and matching as well, which can help in reducing the computation time and yielding promising results in tracking which helps to maintain a good accuracy.

Latha Anuj et al. [8], proposes a model with DTM-CNN approach where pre-training is done over a large dataset. Use of shared layer structure in the model helped in identification of small dimensional objects even under the conditions of occlusion. In case of change in sequence of tracking, the system updates online where DTM-CNN forms a new network by combining the shared layers, which indicates the system is robust to change in conditions.

Han Wu et al. [9], proposes a design for predicting the lost objects while tracking under occlusion using motion prediction strategy. It employs a pre-trained Yolov4-tiny for detection, and the SORT-YM algorithm produces efficient results and improves the time for tracking.

Pavana Pradeep et al. [10], developed a simple detection technique YLLO for object detections for even in the conditions when the edge controller in the model receives the video stream where a fall in frames transmitted is seen. The system manages to maintain a good accuracy above 85% for object identification even tin the condition mentioned.

Swati A. Sagar et al. [11], proposed a modern system for communication between cameras in network about the object tracking information, where each cameras in the network are equipped with individual memory, processor, and communication medium, and are also capable of individually tracking the objects from input video stream.

Suresh Kumar Jha et al. [12], proposes a working model for detection and tracking process from surveillance video inputs, as these videos are considered with low resolution and are challenging in the process. The multi-view approach uses in the model helped to overcome the challenge faced by the author and also is the reason for efficiency of the system.

Mukesh Prasad et al. [14], developed a system to detect vehicles from multi point of views, and used an active learning algorithm for labelling and two methods of filtering are used for filtering the background making the system robust to changes in nature and viewpoints.

Pengfei Ren et al. [15], based on spatial-temporal filtering the author proposes a vehicle tracking system. The spatial-filter used in this MTMC model is efficient in detection with lesser mismatches.

Andreas Specker et al. [16], proposes a model with MTMC tracking which can handle occlusions and also employs a cross-camera clustering hierarchically and also has the feature of vehicle Re-IDing.

Xiao Tan et al. [18], the author proposes a multi-camera tracking model with re-identification feature. It can deal with tracking loss with orientation prediction. Classification of vehicles and also based on their body types is employed to improve tracking in large traffic data.

3. Discussions

Improvement in dual-layer occlusion handling provides progressive association for tracking occluded targets, which overcomes most of the fragments in the frames of the input video stream [1].

The proposed model works good for detection of objects from top angle. In addition to this work, to track objects from other angles the model must be trained to detect efficiently [2].

The work needs further improvement in terms of speed with regarding to frames processed, as well as adding more information to improve tracking performance [3].

The missing and required work here is to collect diverse labelled data and develop unsupervised learning techniques for the vehicle Re-ID task. The system must also aim to systematically and quantitatively analyse each stage of the vehicle Re-ID system on a large-scale labelled dataset in the future [4].

Even in the presence of particle filter, the system is quite slow for real time problems such as handling the symbolic constants and thus needs a lot of calculation to fix this requirement. In addition to this, the Kalman filter implemented in this system fails unexpectedly while tracking the objects which are under non-linear motion [5].

When demonstrated the proposed algorithm to detect it worked well in the fewer occlusion environment when compared with high occlusion environments, the tracing time would increase as the crowd or occlusion increased [6]

Even after reducing computation time by using the same feature functions for all the cameras in the setup, an issue of errors from camera to camera was not addressed which can mislead in wrong identification and mis tracking [7].

With the good tracking speed SORT-YM algorithm does a good job in tracking objects, but the system must improve the Kalman filter to enhance the robustness in tracking even under complex situations, and the other increasing the efficiency of the object detection algorithm and the underlying network [9].

With the capability of maintain a good accuracy in the problematic condition, with the author specifies that the system can be improved with the use of techniques such as inter-camera feature correspondence and homograph transformation of overlapping camera views [10].

Cameras with individual memory, processor can be a vital improvement in the network. But when seen from the aspect of operation the system must be capable of upgrading the camera modules efficiently as each contain individual processors and during failure in the processor, communication disrupts and replacement may cost more than in general networks with simple camera modules [11].

The system proposed works efficiently on the low-resolution videos, the requirement needed to be addressed is the increase in the system's performance speed and increase in processing of frames per second [12].

The system proposed by the author using spatial- temporal filtering, can perform efficient detection without mismatching in multiple-cameras. The system must be addressed in different deployment scenarios as it is worked on a single track of data [15].

The MTMC tracking system proposed employs Re-IDing feature to reduce mismatching and mis-tracking among multiple cameras. And this system is better in when compared to the system proposed in [15], and has better performance parameters [16].

4. Conclusion

In this paper, we have discussed on different approaches for object detection and tracking with multiple-cameras. The ability of each proposed systems and the required improvements are also mentioned in this paper.

As previously considered that occlusion as a major challenge in tracking objects; multi-camera-based setups enable to overcome the performance shortcomings of single camera-based setups. Where multi-camera setups are able to detect the objects from multiple view-points, and across multiple-cameras with proper re-identification of objects and re-assigning of proper IDs to the previously tracked objects. The multi-camera setup also addresses the issue of tracking objects under both long term and short-term occlusion conditions, by employing appropriate filters, estimation schemes, path predictors, etc lead to precise and promising tracking.

3D approach for detection and classification of objects can be seen in the future scope. As once the object detected is abstracted in 3D form, it can be easier for the system to track down the object even in different view-points among multiple-camera setup. Communication between cameras can improve tracking process and it can be considered as an additional feature for applications such as Smart Cities.

References

[1] L. Liu, J. Xing and H. Ai, "Multi-view vehicle detection and tracking in crossroads," The First Asian Conference on Pattern Recognition, 2011, pp. 608-612.

[2] Hao Wu, Xinxiang Zhang, Brett Story, Dinesh Rajan, "Accurate Vehicle Detection Using Multi-Camera Data Fusion and Machine Learning", ICASSP 2019.

[3] Oumayma Messoussi, Felipe Gohring de Magalh˜aes, Francois Lamarre, Francis Perreault, Ibrahima Sogoba, Guillaume-Alexandre Bilodeau, and Gabriela Nicolescu, "Vehicle Detection and Tracking from Surveillance Cameras in Urban Scenes", Polytechnique Montreal, Montreal, Quebec H3T 1J4, Canada.

[4] Chih-Wei Wu, Chih-Ting Liu, Cheng-En Chiang, Wei-Chih Tu, Shao-Yi Chien, "Vehicle Re-Identification with the Space-Time Prior", MOST 107-2633-E-002-001, National Taiwan University, Intel Corporation, and Delta Electronics.

[5] D. Sudha, J. Priyadarshini, "A New Hybrid Proposed Algorithm for Multiple Vehicle Detection and Tracking in a Day-Time Environment", International Journal of Recent Technology and Engineering, Volume 8, Issue 2, July 2019.

[6] Jonah Ong, Ba-Tuong Vo, Ba-Ngu Vo, Du Yong Kim and Sven Nordholm, "A Bayesian Filter for Multi-view 3D Multi-object Tracking with Occlusion Handling", IEEE.

[7] Reyes Rios Cabrera, Tinne Tuytelaars and Luc Van Gool, K. U. Leuven, "Efficient Multi-Camera Detection, Tracking, and Identification using a Shared Set of Haar-Features", ESAT – PSI, Belgium. IBBT, Belgium, ETH-Zurich BIWI Switzerland.

[8] Latha Anuj and M. T. Gopalakrishna, "Discriminatively Trained Multi-source CNN Model for Multi-camera Based Vehicle Tracking Under Occlusion Conditions", Smart Intelligent Computing and Applications, Springer Nature Singapore Pte Ltd. 2020.

[9] Han Wu, Chenjie Du, Zhongping Ji, Mingyu Gao and Zhiwei He, “SORT-YM: An Algorithm of Multi-Object Tracking with YOLOv4-Tiny and Motion Prediction”, Electronics 2021, 10, 2319.

[10] Pavana Pradeep, Amitangshu Pal, Krishna Kant, “Resource Efficient Edge Computing Infrastructure for Video Surveillance”, Computer and Information Sciences, Temple University, Philadelphia, USA.

[11] Swati A. Sagar and Anilkumar Holambe, “A Noval System Architecture for Multi Object Tracking Using Multiple Overlapping and Non-Overlapping Cameras”, Volume 13, Number 3 (2017) pp. 275-283.

[12] Sureshkumar Jha, Rohan Sawant, and Parth Shinde, Rakhi Kalantri, Shagufta Rajguru, “Study on Design and Implementation of Distributed Multiple Camera Surveillance and Tracking System”, ITM Web of Conferences 40, 03036 (2021) ICACC-2021.

[13] Ms. Latha Anuj M T Gopala Krishna, “Multiple Camera Based Multiple Object Tracking Under Occlusion: A Survey”, International Conference on Innovative Mechanisms for Industry Applications (ICIMIA 2017).

[14] Mukesh Prasad, Chih-Ling Liu, Dong-Lin Li, Chandan Jha, Chin-Teng Lin, “Multi-view Vehicle Detection based on Part Model with Active Learning”, IJCNN 2018.

[15] Pengfei Ren, Kang Lu, Yu Yang, Yun Yang, Guangze Sun, Wei Wang, Gang Wang, Junliang Cao, Zhifeng Zhao, Wei Liu, “Multi-Camera Vehicle Tracking System Based on Spatial-Temporal Filtering”, Nanjing Fiberhome Tiandi CO., LTD Nanjing, Jiangsu, China.

[16] Andreas Specker, Daniel Stadler, Lucas Florin, Jurgen Beyerer, “An Occlusion-aware Multi-target Multi-camera Tracking System”, Fraunhofer Center for Machine Learning.

[17] Latha Anuj, M T Gopalakrishna, C Naveena, and Sharath Kumar Y H, “V-DaT: A Robust method for Vehicle Detection and Tracking”, Vol.12 No.2 (2021),2492-2505, Turkish Journal of Computer and Mathematics Education.

[18] Xiao Tan, Zhigang Wang, Minyue Jiang, Xipeng Yang, Jian Wang, Yuan Gao, Xiangbo Su, Xiaoqing Ye, Yuchen Yuan, Dongliang He, Shilei Wen and Errui Ding, “Multi-camera vehicle tracking and re-identification based on visual and spatial-temporal features”, Baidu Research, Baidu Inc.