A Review for Enhancement Methods for Object Detection and Tracking Methods in Visual Surveillance System of Cameras Network

Wael Mahdi Brich
Computer Science Department, College of Computer Science & Information Technology, University of Babylon, Najaf, Iraq.
E-mail: waal_555@yahoo.com

Israa Hadi Ali
Computer Science Department, College of Computer Science & Information Technology, University of Babylon, Babylon ,Iraq.
E-mail: Israa_hadi@itnet.uobabylon.edu.iq

Received May 06, 2021; Accepted August 08, 2021
ISSN: 1735-188X
DOI: 10.14704/WEB/V18SI05/WEB18221

Abstract

The intersection of computer vision and sensor network Contributed to design and spread of camera networks. Camera network-based visual monitoring is an active and important field. A wireless camera sensor network can be classified as a centralized or as a distributed network. In this research explains overview for surveillance systems, cameras network topology and the recent methods are used for association different observations in the same view and acrossblind area from different cameras views for the same object. For video tracking enhancementa new idea will be suggested for using new features with M-RCNN for improve the object identification and prediction where the object may be reappeared in network.

Keywords

Surveillance System, Camera Networks, Vedio Tracking, Network Topology, M-RCNN.

Introduction

Surveillance systems based on multiple cameras are very important in our life because it used for monitoring and analysis for many objects in different places at the same time. In addition to the above, these systems can be detected the strange and anomalies events (H. Bouma, J. Baan& P.T. Eendebak, 2015). Depend on camera distribution there are three types of cameras network, the first kind is called centralized network when each camera sent it scene information for central server for processing. The second type is named decentralized network when numbers of cameras are grouped in cluster and corporative between them to sending it scene information for a single camera of the cluster called vision center or leader. The third type is distributed network when each camera won an embedded unit processing for process the scene information and exchanges it with other camera neighbors. The distributed network can be overlapping (interference cameras FOV) or nonoverlapping (blind area between cameras FOV) (Mingli Song& Dacheng Tao,2013). Covert Large-scale area by cameras network for monitoring stay high challenge in surveillance systems when environment light and object pose change across FOV, add on that the difficult to association observations data between two trajectories for the same object across nonoverlapping cameras FOV. The area within FOV called intra view and the area when the object moving between two views named enter view. To improve the video tracking on a large-scale area required to select a better method that can solve the handoff problems between cameras (Dimitrios Makris, Tim Ellis & James Black, 2004).

There are many approaches used statistical methods and data transition model for estimating the location of an object with in-camera FOV and across a blind area between two cameras neighbors. The first step for design a monitoring system is topology inference and cameras celebration for all cameras in network depend on estimation methods that used path weights on a graph. The topology graph represents each entrance and exit node when the point where the object leaves the FOV is called the exit node and the point wherethe object inters the FOV named entrance node. Similar trajectories for the same object across different FOV must be detected and joint on a long track (Corey Snder & Minh N. Do, 2019).

This paper follows the structure: the first section explains the concept and importance of surveillance cameras systems and cameras network. Survey for techniques background to apply camera network (overlapping and nonoverlapping) for object detection and tracking. And literature review of recent methods used different tools for re-identification and association trajectories data for same object across different cameras FOV. At last section it presents a table contains advantages and disadvantages for recent methods and the new idea will be suggested for enhancement video tracking depend on new features using M-RCNN.

Surveillance Systems

Computer Vision (CV) technologies have advanced dramatically over the last decade in several methods. Many algorithms have been proposed in (CV) and have significantly improved the performance for many applications which depend on the use of a visual method or manual movement to communicate meaning.

Distributed cameras monitoring is a suitable canonical example of this application class. Visual information plays a necessary function in monitoring applications. Fourcomponents represent the cornerstone for the optimal design of any surveillance camera system. These components explain the subsection of this part. Fig (1) illustrates the relation of dependency and limitations between four components, as the network topologyinference based on camera celebration when camera covert local area in optimal view and depend on communication (adjust the distance between cameras to get on good communicate with as possible for sparse distance). Balancing is required between these components when a short distance between cameras means more power for communication but less sparse (Henry Detmold & Anton van den, 2007) (Dhanya Devarajan & Zhaolin Cheng, 2008).

Sparse Cameras Network

sparse cameras are used for analysis and to reduce the size of video images, determining what video elements of the video stream will be need embedding and at any degree resolution before transmitting. Since the high paced modern field of life needed communication in almost every aspect, features of key for an image captures system is a step by step becoming internet communication and portability. Currently, the main employment of wireless cameras is indoor and wide-area surveillance systems (Omar Javed & Khurram Shafique, 2008) (Mingli Song & Dacheng Tao, 2013).

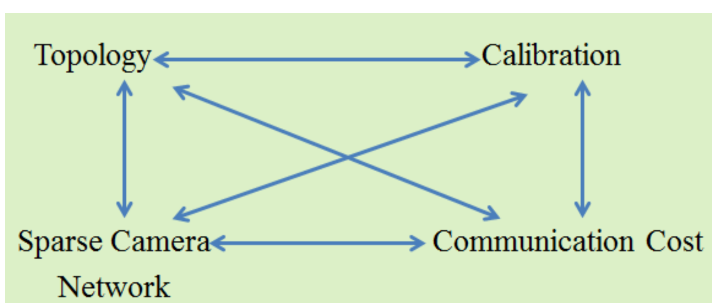

Fig. 1 Illustrates the dependency and limitations between four components

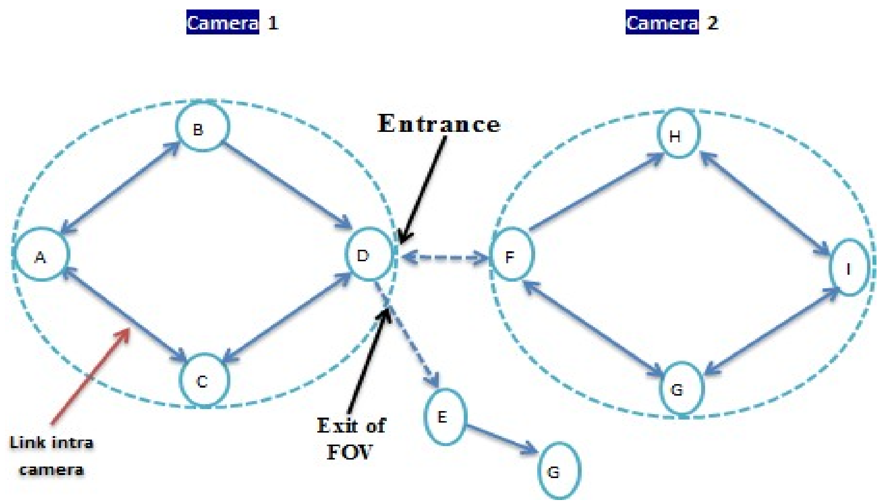

Fig. 2 Illustrates topology representation for exit and entrance nodes of camera and the edges linking nodes

Topology of Network

Topology component meaning how can distribute the numbers of cameras on one area and compute the relations between them. Topology identifies position and FOV for each camera in-network when camera views may be taking different situations. Spatial adjacency means connecting pathway directly between two camera views there is no otherview field between two camera neighbors. When a target leaves a FOV, it may birth in some of the other adjacent camera views with certain likelihoods as illustrates in fig (2) (Bir Bhanu &Amit Roy, 2011). When topology inference must calculate the celebration information and communication distance between cameras on the network, on a large areascale the distance as possible between cameras which used to covert this area. Create topology graph when each source and sink represent a node and the distance between nodes represent as a weighted link depends on many features (position, direction, speed, time transition, traffic, etc..) using estimation methods (Dimitri Marinakis & Gregory Dudek, 2006).

Camera Celebration

Camera celebration is an important and necessary step that must be used when designing a monitoring system. This step employ calculates the accurate coordinates of the image after transfer the information of the 3D-real word to 2D camera projection, the transfer operation based on two terms of parameters called extrinsic and intrinsic parameters. The intrinsic parameters are the same as focal distance combined with local extrinsic parameters same as height and tilt angle to process camera celebration. Camera calibrationmethods can be classified in explicit versus implicit, methods using 3D rather than planer point arrays and linear techniques as illustrated in fig (3) (Xiaotang Chen & Kaiqi Huang, 2013) (Dhanya Devarajan & Zhaolin Cheng, 2008).

Communication Cost

When design a video monitoring system a communication bandwidth and space storage are important considerations for design requirements. The complexity of sceneinformation, frame rate, and compression ratio represents the factor that must be doing balancing to get on good design. In wireless cameras network when reducing the communication between disjoint cameras on the network, will increase the performanceof the video tracking (Omar Javed & Zeeshan, 2014).

Data Association Across Cameras (Fov)

Distributed cameras on one large area depend on dependency and limitations of celebration, topology, and sparse network for covering different locations; this distribution can be overlapping when cameras FOV are interference and nonoverlapping when there is a distance between cameras FOV. The first type can be tracking the target based on the appearance and shape of the object using detection methods to select the unique features of the target for tracking across FOV. Object tracking in one FOV called intra tracking when the target appears in camera view, second type must track an object across a blind area between nonoverlapping cameras FOV; in this case for joint same trajectories for the same object across the blind area, the system must use estimation methods for prediction where the object may reappear on a network.

Kalman Filter and Particle Filter are the two most basic stochastic tracking methods, as well as two representatives an observations joints method – Multi-Hypothesis Tracking (MHT) and Joint Probabilistic Data Association Filters (JPDAF) methods (Paulo Freitas & Paulo Menezes, 2012). In multi-cameras environments when tracking is decentralized there are some prominent tracking algorithms as graph matching, practical filter and Kalman filter. Graph matching mean link up the trajectories for each group of tracks and create super-tracks, then associating the super tracks between center fusions over the network (Mingli Song& Dacheng Tao, 2013). Particle filter using to compute local state estimates which are later can be merged in one track at the fusion center.

Kalman filter can be defined when a camera is chosen as a leader of cluster which gathered each the information measurements of an object to analyze and approximate its position at any time depend on a Kalman filter and delivers that estimate to a central base station for processing (Yuan-Kai Wang & Ching-Tang Fan, 2012). Collaborative among cameras network when adjoin cameras are exchanged information matrix and covariance matrix of target improve the prediction process and reduce the error for object re-identification (Xiaotang Chen & Kaiqi Huang, 2013).

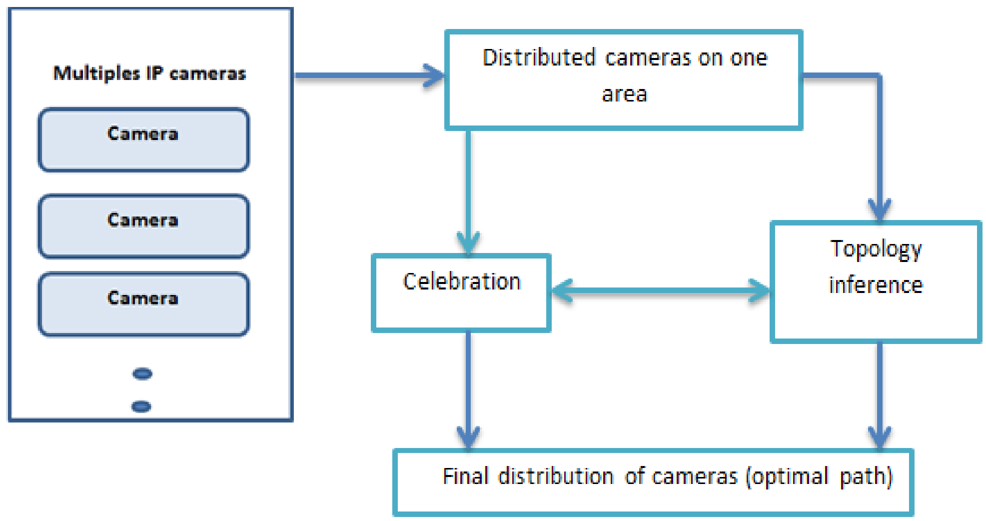

Fig. 3 Illustrates the block diagram for camera celebration and topology inference

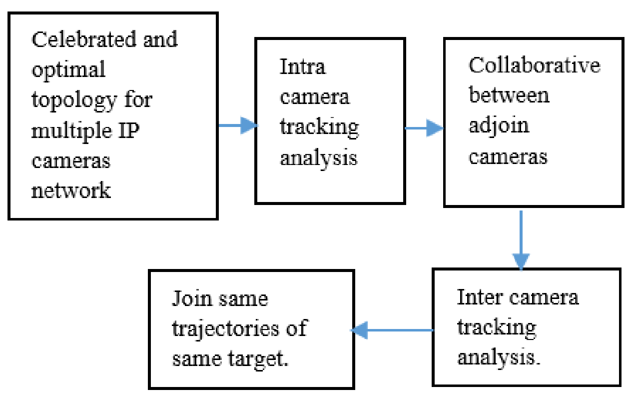

Fig. 4 illustrates the steps for monitoring camera system over large area scale

Tracking of Object Across Cameras Network

After completing the camera celebration and topology inference for the IP cameras network and represented the environment by optimal distribution for cameras based on thelocal camera position and the relation between cameras with other adjoint cameras, the tracking process is beginning when a target is detected in one camera (Andrzej Czyzwski & Karol LISOWSKI, 2014) (Li Hou & Wanggen Wan, 2017). The object moving withina camera and calculate the trajectory information of the target, many methods used to generate trajectory and discriminate this information from other information for other objects moving at one time as illustrated in fig (4) (Li He & Guoling Liu, 2015). The second stage presents tracking of target depend on the information gathered from moving object across camera FOV as (position, direction, transfer time, appearance) and based on the estimation methods for prediction at any place the object may be reaper over the network (Amit K. Roy & Bi Song, 2012) (Xiaojing Chen & Bir Bhanu, 2017). For optimal global view for moving object must join the same trajectories for the same object across nonoverlapping cameras FOV, closed blind area between nonoverlapping camera FOV will improve the tracking process, many features depend on target behavior andplace geometrical used for improving prediction process for target (Weihua Chen & Xiaotang Chen, 2016) (Jorge Nino & Vedran Jelaca, 2011).

Survey for Techniques Background to Object Tracking Across IP Cameras Network

In this section, different techniques of recent methods of cameras topology inference and object tracking across multiple cameras over large area scale using IP cameras network for the period between of (2017-2021).

(Yeong-Jun Cho, Kuk-Jin Yoon, 2017) the method proposes a novel method for effective person re-identification. Based on distance in camera network topology. In first step is calibrating each camera and complete guesstimate relative scale ratio between neighbor cameras based on the human height's information. the calibration which results from numbers of cameras that lead to compute the velocity of each person and deduce the distance between neighbor cameras to produce optimal distance. The method can be used adaptively to each person according to its velocity and cope with different transition time of people between non-overlapping cameras.

(Behzad Bozorgtabar, Roland Goecke, 2017) Visual tracking of multiple persons at the same time is an essential operation for behavior study. The approach reveals that multi-tracking of target in network with non-overlapping FOVs can be formulated in a context, where the association among all given objects which may be within and across cameras is performed instantaneously. The approach supports to overcome on the weakness of multi-camera-based tracking. The method begins with the initial object track and incorporates equally inter and intra-camera data link into a track of global view.Inside the context lies target-particular model learning for data association in inter camera.Furthermore, local clues such as location of point on trajectories that give a better estimate the affinity between any couples of tracks within a specific camera view.

(Quoc Cuong Le, Donatello Conte, 2018) the paper presents a novel technique of cooperative cameras tracking in a coordinated overlapping cameras network. It proposes an effective target association method a cross camera depend on the tracking results of each target on each FOV. The method naturally holds obstacle and shared target occlusions. The outcomes of tracking for each FOV are gathered and associated completely into targets. The individual cameras based on feedback from the centralassociation to improve tracking unseen targets, even in the case of full occlusion.

(Yihua Tan, Yuan Tai and Shengzhou Xiong, 2018) In this paper, named NCA-Net is designed based on a feature extraction network to enhancement the performance of several objects tracking in number of cameras by avoiding the situation of insufficient strength caused by hand-crafted features. The network grouped two metrics general learning and features learning using (CNN) model and the cost function like neighborhood components analysis (NCA). The cost function is modified from the likelihood loss of NCA aiming at object tracking.

(Yingkun Xu, Xiaolong Zhou, 2019) this paper is summaries deep learning depend on multi-object tracking techniques, which are good ranked in public standards. This approach has three phases. First, the usage of robust deep learning for multi-object tracking is organized, and the process of deep feature transferring, neural net embedding, and end-to-end network training are analyzed depend on existing techniques, by which theprocedures to construct new tracking frame are motivated. Second, can examine the roles of deep networks in tracking framework and discover the problems of training in these networks. Third, differences between these multi-object tracking techniques are submitted and reorganized according to general datasets and some evaluations.

(Jing Li In, Jing Xu, 2020) in this research, the researcher present extend the single-camera AOT to a multi-camera setting, where these cameras tracking a moving target in a collaborative fashion. To accomplish efficient collaboration among cameras, the method shows and propose a novel Pose Assisted Multi-Camera Collaboration model, which helps a camera to collaborate with the other cameras by sharing camera fronts for effective object tracking. In this model, each camera is supplied with two switcher and controllers. The pose-based controller moves the camera in accordance with the poses of the other cameras. At each stage, the switcher decides which action to take from the two controllers according to the visibility of the target.

(Olly Styles, Tanaya Guha, 2020) the paper is present a new process for trajectory prediction of object(human) in a multi-camera situation, which is named multicamera trajectory forecasting (MCTF). To enable more research on MCTF, its show a big dataset, WNMF, which was labeled based on a semi-automated data mark technique which is developed. Moreover, the approach offered numerous baseline outcomes for forecasting the following camera in which a target in this case a (person) may be re-birth inside a cameras network.

(Parshwa Shah, Vandit Gajjar, 2021) The idea of approach (PeR-ViS) recovers the person in scene monitoring depend on the semantic features as a Height; Cloth color of Torso, pattern, and kind; Cloth color of leg, pattern, and Gender. The main advantage from this filtering series is that height, Torso Color, kind, and Pattern are simple differentiable compared to other descriptors, where the dense crowd is present, leg patch will not easily extract. Thus, the select of this filtering series is very essential in this work.

Discussion the Methods and Suggested Idea

Most of researchers previously worked on video tracking in computer vision, and they are also applying a Various method used different tools for spatial-temporal estimations for target location and using different algorithms for target re-identification on network.Many recent methods discussed in section (survey for techniques background) When eachmethod has several advantages and disadvantages discussed in a table (1). It used for association different observations across blind area from different cameras scenes for the same object in tracking pathway. The concept of suggestion research is using a new feature (geometrical features and the features which depend on target behavior) for prediction where the object may be reappeared in the large camera network. Depend onthe summation of weights for features that collected from a scene to estimate the place where the object may be reappeared. Depending on new features to increase the efficiencyof video tracking when monitoring a target across cameras network. Geometrical features mean the relation spatial with traffic and the important paths in place which covered by camera FOV, when the target is human can be used a new feature as direction of object onpath, direction movement of head for target, traffic cues in streets and hand movement of target and comparative the distance (after number of frames) from current position of object to points on scene boundary where the object may be leave the scene. Each feature has value of weight when sum all weights represent the final value for prediction forwhere the target may be reappearing on network. M-RCNN will used for detection and tracking (Kaiming He & Georgia Gkioxari, 2018).

Table 1 Survey for recent methods of monitoring camera system with advantages and disadvantages

Reference of Method	Advantages	Disadvantages	Year
(Yeong-Jun& Kuk-Jin)	1- Multiple features of object appearances are combined in one by the averaged summation 2- depend on speed object in two cameras to estimate (blind area distance) to inference topology	The proposed structure does not objective to find the pure scale of the space world coordinate system and actual walking rates of people, but just to match the measurements across camera.	2017
(Roland Goecke & Behzad Bozorgtabar)	1-it is used for control on the fragility of multi-camera based on operation of tracking. 2- can be resolves data linking in intra, inter-cameras At same time	The operate time of the proposed system largely depends on the number of targets (graph nodes) and on the disjoint camera views	2017
(Yihua Tan & Yuan Tai)	Combining the deep features with these subtly constructing tracking systems can efficiently reduce the performance loss produced by artificial appearance feature and display a tactic that solves the training complexity in problem of convergence.	More complexity	2018
(Quoc Cuong & Donatello Conte)	presented a new robust technique online multi-observation multi-object tracking approach that easily handles with the difficult occlusion in tracking.	The association approach in the original MDP technique cannot improve the tracking state by recover of target when it has been unseen for a long period.	2018
(Bonan Cuan)	Wonderful advances of learning algorithms (deep learning system) are produced in the fields of image identification, detection of object and object re-identification, which is also beneficial to the enhancement of multi-object tracking.	Usually, it is not easy to obtain multi-object tracking results by just one network because there are some interconnected sub-modules in MOT tracking.	2019
(Jing Li & Fangwei Zhong)	AOT method has been extensively deployed in a variety of real-world applications, for instance controlling a mobile robot to track a moving target for getting filmic shots or rotating a 3-axis stabilized camera to auto-follow a pedestrian or track a face.	It is costly and hard to learn a general collaborative protocol under the high-dimensional visual observation, particularly with the expanding in number of cameras.	2020
(Olly Styles & Tanaya Guha)	It is need predicting the future coordinates of an moving object even before it arrives the boundaries camera view. The forecasted location may then work as a prior for the object detection algorithm, to reducing the search in space area for complete detection process.	Stopped object in blind area for long period causes error in result.	2020
(Vandit Gajjar & Parshwa Shah)	It solves a problem of person recovery in video monitoring with a semantic feature by build a deep learning algorithm-based cascade filtering approach (PeR-ViS), the approach applies (Mask R-CNN).	Increase the features to improve the description of person it means increase the computation complexity.	2021

Conclusion

In the current proposal, many recent algorithms (published in 2017-2020) presented and discussed. Most papers presented in the last years focused on using estimation methods for prediction the target at any place of camera network may be re-appear and re-identification the features of same object across cameras on network. These papers discussed the tracking across large area scale and the challenges which appear as reason for changes in illumination, pose, clutter, occlusion with object moving. Different methods used different types of tools for tracking and joint same trajectories for same object over network. And explain the suggested idea for improve the tracking when using new features selected from trajectory of object with geometrical features to robust the forecasting of moving object across blind area between nonoverlapping cameras network.

References

Den Hollander, R.J., Bouma, H., Baan, J., Eendebak, P.T., & Van Rest, J.H. (2015). Automatic inference of geometric camera parameters and inter-camera topology in uncalibrated disjoint surveillance cameras. *In Optics and Photonics for Counterterrorism, Crime Fighting, and Defence XI; and Optical Materials and Biomaterials in Security and Defence Systems Technology XII*, 9652, 96520D. http://dx.doi.org/10.1117/12.2194435.

Mingli, S., & Dacheng, T. (2013). *Sparse Camera Network for Visual Surveillance -- A Comprehensive Survey,* IEEE.

Makris, D., Ellis, T., & Black, J. (2004). Bridging the gaps between cameras. *In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition.*

Snyder, C., & Do, M. (2019). Streets: A novel camera network dataset for traffic flow. *Advances in Neural Information Processing Systems*, *32*, 10242-10253. https://doi.org/10.13012/B2IDB-3671567_V1

Detmold, H., Van Den Hengel, A., Dick, A., Cichowski, A., Hill, R., Kocadag, E., & Munro, D.S. (2007). Topology estimation for thousand-camera surveillance networks. *In First ACM/IEEE International Conference on Distributed Smart Cameras*, 195-202. https://doi.org/10.1109/ICDSC.2007.4357524

Devarajan, D., Cheng, Z., & Radke, R.J. (2008). Calibrating distributed camera networks. *Proceedings of the IEEE*, *96*(10), 1625-1639. https://doi.org/10.1109/JPROC.2008.928759

Javed, O., Shafique, K., Rasheed, Z., & Shah, M. (2008). Modeling inter-camera space–time and appearance relationships for tracking across non-overlapping views. *Computer Vision and Image Understanding*, *109*(2), 146-162. https://doi.org/10.1016/j.cviu.2007.01.003.

Bhanu, B., Ravishankar, C.V., Roy-Chowdhury, A.K., Aghajan, H., & Terzopoulos, D. (Eds.). (2011). *Distributed video sensor networks*. Springer Science & Business Media, 349-371. https://doi.org/10.1007/978-0-85729-127-1_24.

Marinakis, D., & Dudek, G. (2006). A practical algorithm for network topology inference. *In Proceedings IEEE International Conference on Robotics and Automation,* 3108-3115.

Chen, X., Huang, K., & Tan, T. (2014). Object tracking across non-overlapping views by learning inter-camera transfer models. *Pattern Recognition*, *47*(3), 1126-1137. https://doi.org/10.1016/j.patcog.2013.06.011.

Omar, J., Zeeshan, R., Khurram, S., & Mubarak, S. (2014). Tracking Across Multiple Cameras with Disjoint Views. *Proceedings Ninth IEEE International Conference on Computer Vision*. https://doi.org/10.1109/ICCV.2003.1238451

Freitas, P., Menezes, P., & Dias, J. (2012). Online topological mapping of a sparse camera network. *In Doctoral Conference on Computing, Electrical and Industrial Systems*. Springer, Berlin, Heidelberg, 229-240.

Wang, Y.K., Fan, C.T., & Huang, C.R. (2012). A large scale video surveillance system with heterogeneous information fusion and visualization for wide area monitoring. In *2012 Eighth International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 178-181. https://doi.org/10.1109/IIH-MSP.2012.49.

Czyzewski, A., & Lisowski, K. (2014). Employing flowgraphs for forward route reconstruction in video surveillance system. *Journal of Intelligent Information Systems*, *43*(3), 521-535. https://doi.org/10.1007/s10844-013-0253-8.

Hou, L., Wan, W., Hwang, J. N., Muhammad, R., Yang, M., & Han, K. (2017). Human tracking over camera networks: a review. *EURASIP Journal on Advances in Signal Processing*. https://doi.org/10.1186/s13634-017-0482-z.

He, L., Liu, G., Tian, G., Zhang, J., & Ji, Z. (2019). Efficient multi-view multi-target tracking using a distributed camera network. *IEEE Sensors Journal*, *20*(4), 2056-2063. https://doi.org/10.1109/JSEN.2019.2949385.

Roy-Chowdhury, A.K., & Song, B. (2012). Camera networks: The acquisition and analysis of videos over wide areas. *Synthesis Lectures on Computer Vision*, *3*(1), 1-133. https://doi.org/10.2200/S00400ED1V01Y201201COV004

Chen, X., & Bhanu, B. (2016). Integrating social grouping for multitarget tracking across cameras in a CRF model. *IEEE Transactions on Circuits and Systems for Video Technology*, *27*(11), 2382-2394.

Chen, W., Cao, L., Chen, X., & Huang, K. (2016). An equalized global graph model-based approach for multicamera object tracking. *IEEE Transactions on Circuits and Systems for Video Technology*, *27*(11), 2367-2381. https://doi.org/10.1109/TCSVT.2016.2589619

Castañeda, J.N., Jelaca, V., Frías, A., Pizurica, A., Philips, W., Cabrera, R.R., & Tuytelaars, T. (2011). Non-overlapping multi-camera detection and tracking of vehicles in tunnel surveillance. *In International Conference on Digital Image Computing: Techniques and Applications*, 591-596. https://doi.org/10.1109/DICTA.2011.105.

Yeong-Jun, C., & Kuk-Jin, Y. (2017). Distance-based Camera Network Topology Inference for Person Re-identification. *Computer Vision Laboratory, GIST.*

Bozorgtabar, B., & Goecke, R. (2017). MSMCT: Multi-state multi-camera tracker. *IEEE Transactions on Circuits and Systems for Video Technology*, *28*(12), 3361-3376.

Le, Q.C., Conte, D., & Hidane, M. (2018). Online multiple view tracking: Targets association across cameras. *In 6th Workshop on Activity Monitoring by Multiple Distributed Sensing (AMMDS 2018)*.

Tan, Y., Tai, Y., & Xiong, S. (2018). NCA-net for tracking multiple objects across multiple cameras. *Sensors*, *18*(10), 3400.

Bonan, C. (2020). *Deep similarity metric learning for multiple object Tracking.* doctoral thesis.

Li, J., Xu, J., Zhong, F., Kong, X., Qiao, Y., & Wang, Y. (2020). Pose-assisted multi-camera collaboration for active object tracking. *In Proceedings of the AAAI Conference on Artificial Intelligence*, 34(1), 759-766. https://arxiv.org/pdf/2001.05161.

Styles, O., Guha, T., Sanchez, V., & Kot, A. (2020). Multi-camera trajectory forecasting: Pedestrian trajectory prediction in a network of cameras. *In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 1016-1017.

Shah, P., Garg, A., & Gajjar, V. (2021). PeR-ViS: Person Retrieval in Video Surveillance using Semantic Description. *In WACV (Workshops)*, 41-50.

He, K., Gkioxari, G., Dollar, P., & Girshick, R. (2018). Mask R-CNN. Facebook AI research (FAIR). *arXiv preprint arXiv:1703.06870*. https://github.com/facebookresearch/Detectron.

Shahmoradi, L., Ehtesham, H., Mehraeen, E., Rostampour, N., Tahmasbian, S., & Ghasempour, M. (2019). Instructional design, development and evaluation of congenital hypothyroidism registry system. *Webology, 16*(2), 275-288.