Proceedings of the IEEE
International Conference on Information and Automation
Ningbo, China, August 2016

A Survey on Multiple Object Tracking Algorithm*

Litong Fan[1], Zhongli Wang*[1,2], Member, IEEE and Baigen Cai[1,2], Member, IEEE

[1]*School of Electronic Information and Engineering, Beijing Jiaotong University, Beijing 100044, China*
[2]*Beijing Engineering Research Center of EMC and GNSS Technology for Rail Transportation, China*

{15120240 & zlwang & bgcai}@bjtu.edu

Chuanqi Tao, Zhiyi Zhang, Yinling Wang, Shanwen Li, Fengtian Huang, Shuangfu Fu, Feng Zhang

CSR Qindao Sifang Co. Ltd.Qindao, Shandong Province 266111, China

{wyl & softlsw}@cqsf.com

***Abstract* - Visual Multiple Object Tracking (VMOT) is an important computer vision task which has gained increasing attention due to its academic and commercial potential. There are many different approaches have been proposed to solve the problem. Compared with single object tracking which focuses on appearance model, motion model and other factors, multiple object tracking shares these common challenges, and has some more challenging tasks to be tackled, such as identifying among multiple objects, frequent occlusion due to the crowd, initializing and terminating of the object tracking, et al, which making it difficult to get a comprehensive understanding of this problem. In this paper, we try to make systematic review on VMOT. We review the recent advances in various aspects and introduce the classification of each aspect in detail. The main aspects include observation model and tracking algorithm. In each aspect, the existing methods are classified into different groups, and each group is discussed in details for their principles, advantages and drawbacks. This review work can provide researchers, especially new comers to the topic of VMOT, a general understanding of the state-of-the-arts, and help them to comprehend the aspects of a VMOT system and the inter-connection of these aspects.**

Index Terms* - *visual multiple object tracking, observation model, data association, Bayesian theory

I. Introduction

Visual multiple object tracking (VMOT) aims at locating multiple targets of interest, inferring their trajectories and maintaining their identities in a video sequence[1], which is the key techniques in many application, such as video surveillance, activity analysis, human-computer interaction, traffic control and robotics, motion-based recognition, video indexing and vehicle navigation[2]. Because of its wide application, VMOT is always a hot research topic in the field of computer vision.

VMOT is regarded as the natural extension of the well-studied single target tracking, both of them share something in common, for example, they both suffer from the translation, rotation and scale changes due to the relative motion of objects and camera, the camera distortion, information loss because of the projection from 3D to 2D. But VMOT is actually a more complicated and challenging problem. The task of VMOT is mainly partitioned to locating multiple objects, maintaining their identities and yielding their individual trajectories given an input video [3]. There are many issues bring the difficulties for multiple objects tracking, especially in crowded environments where many targets have similar appearance. Additionally, objects pose variation, shape deformation, and dynamic, cluttered backgrounds, in this situation, many researchers adapt multi-cues to describe the object [4][5]. In some cases, the targets may disappear in the field of view, frequently interaction among multiple objects and longtime occlude each other, motion is efficient to predict the objects to handle this miss detection [6][7]. Real time requirement in application is another challenge because of the number of objects.

Because of these challenges, VMOT attracts much attention and during the last two decade, a number of object tracking algorithms have been proposed aiming to solve one or more of these challenges [5].

There are some review papers on VMOT which present some classification methods. One of the conventional category methods to multiple objects tracking algorithm is recursive and non-recursive [4]. The recursive methods is that their current state is estimated only using information from previous frames, the non-recursive methods considers both past frames and future frames to estimate the state of objects. This category can cover all the methods of VMOT, but it is too broad to describe the principle of the algorithms. The other classification method is category free tracking (CFT) and association based tracking (ABT)[7]. The ABT approaches usually first localize objects in each frame and then link these object hypotheses into trajectories without any initialization labeling. CFT methods can be viewed as the extension of single object tracking to track multiple objects, i.e. multiple parallel single object trackers.

Generally, VMOT method is mainly composed of two components: observation model and tracking process. The observation model is a descriptor of objects, including some unique features which can used to distinguish the individuals, detect and track objects. Tracking is a process that links the observation to the trajectory (tracklets). Based on this, we

* This work was supported partly by Natural Science Foundation of China under Grant No.61573057, and partly by National Science and technology supporting project under Grant No.2015BAF08b01, and partly by the National Natural Science and High Speed Railway Joint fund project (U1334211) , and partly by the Basic scientific research Foundation of BJTU (2015JBJ002).

978-1-5090-4102-2/16/$31.00 ©2016 IEEE

introduce a classification for multiple tracking methods, they are tracking method based on Bayesian theory and the methods based on data association. The former treat tracking as a predict process, the commonality of these kinds of methods is that all trajectories are estimated with Bayesian theory. The later mainly associates the observation of targets to the trajectory by computing the affinity between them. It can be further divided into two groups according to the data range they used: local optimization-based data association and global optimization-based data association. Taxonomy of visual multiple object tracking in this paper can be summarized as in Figure1.

The rest of the paper is organized as follows, in Section II, we describe some common observation model which used as an input for detector and tracker, and the different models used in VMOT. In Section III, both the method based on Bayesian theory and the method based on data association, including the main literatures related to these methods are summarized, their strength and weaknesses are also introduced. Section IV is the conclusion.

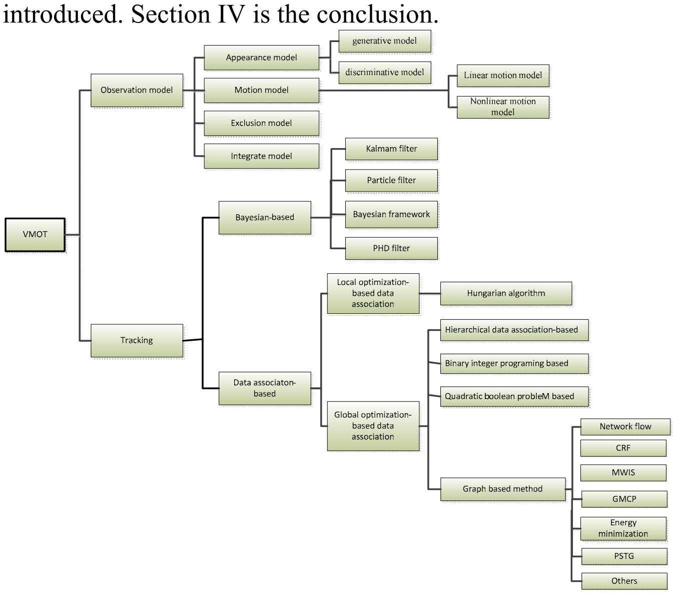

Fig.1 Taxonomy of visual multiple object tracking

II. OBSERVATION MODEL

Observation model is used for describing objects property, such as its appearance, velocity, and location. So the model is generally based on the feature, by capturing the special features to distinguish target and background, or different objects. There are some conventional features applied to describe object, such as colour, gradient, shape, texture, depth, super-pixel, motion, optic flow. And the appearance model may fuse multiple clues to achieve more robust tracking. The main features used in the existing literatures and the correspondent references are listed in TABLE I.

TABLE I THE FEATURES USED FOR OBJECT DISCRIPTION

feature type	representative	advantage	disadvantage	Reference
colour	Histogram of colour	efficient	Sensitive to lightning changes.	[4][8][10]
contour	Edge	less sensitive to illumination changes		[8][12]
gradient	HOG,SIFT	suit human detection		[4] [13]
shape		computational time low		
texture	LBP	accuracy	computation expensive	[8][9]
motion	optic flow	handle occlusion	computation expensive	[6][8][13][14]
Spatiotemporal information	super-voxel super-pixel	Handle partial occlusion and pose variations	can't handle complete occlusion	[4][15][16][17]
depth	3D	handle the drift and occlusion		[18]
multi-feature	colour with texture	robust	computation expensive	[7][12][17]

Using different representation and feature can construct different observation model. The observation model can be classified into four types: appearance model, motion model, exclusion model, integrated model.

A. Appearance model

In terms of the construction of appearance models, it can be categorized into generative model and discriminative model. The generative model describes the visual observations of a moving object, and then tracking is reduced to a search for an optimal state that yields an object appearance most similar to the appearance model. For the discriminative model, the tracking is viewed as a binary classification based on an optimal decision boundary which distinguishes the object from the background.

1) Generative appearance model. The useful methods for constructing the generative appearance models include color histograms [5], HOG, Gaussian mixture models, sparse representation, manifold learning and hidden Markov random field model.

Color histogram is robust to region scaling, rotation, and shape variations. Their limitation is that they usually ignore the spatial distribution of pixel values. Yang et al [7] use a sparse weight constraint in the joint sparse representation to dynamically select the templates from the template set and estimate the coefficients of the templates. Reference [19] proposes a feature learning algorithm for model-free multiple object tracking. Wang et al [20] propose a L1 minimization-based appearance model, which input is a set of candidate regions obtained from Robust Principle Component Analysis (RPCA) in each frame. Reference [21] proposes a hidden Markov random field model that represents the joint dependencies of cluster labels and tracklet linking associations.

2) Discriminative appearance model: Discriminative models select different discriminative features and construct a classifier in the feature space to distinguish between the features from the object and the features from the background. The typical methods are SVM [17][18], random forest, online boosting [4][22].

B. Motion model

Object motion model is also known as the dynamic model, which describes how an object moves. It is important for multiple object tracking since it can be used for predicting the potential position of objects in the future frames, reducing search space. In general, objects are assumed to move smoothly in image space. The popular motion models employed by multiple object tracking are divided into two classes: linear motion model and non-linear motion model.

1) Linear motion model. Tracklet velocity can be estimated using first order linear regression, based on the assumption that real motion is linear over short time-periods. In reference [13], an estimated velocity can be associated with every detection that is a member of a tracklet, and can be used to both refine the initial tracklets as well as to link widely separated tracklets. Reference [14] uses a constant velocity motion model for data association.

2) Non-linear motion model. There are many cases which the linear motion model cannot deal with. Moreover, non-linear motion model can produce more accurate motion affinity between tracklets. Reference [23] proposes a method using motion dynamics as a clue to distinguish targets with similar appearance. Reference [24] describes an online approach to learn non-linear motion patterns and robust appearance models for multi-target tracking in a tracklet association framework. Reference [25] considers motion context from multiple objects which describes the relative movement between objects and construct a Relative Motion Network (RMN) to factor out the effects of unexpected camera motion for robust tracking.

C. Exclusion model

The exclusion model usually takes spatiotemporal information and other constrains into account. Exclusion is an obvious constraint when seeking a solution to the VMOT problem due to the collisions [26]. Given multiple detection responses and multiple trajectory hypotheses, reference [27] takes two constraints into consideration, they are detection-level exclusion and trajectory-level exclusion. Reference [22] constructs multiple graphs to model the spatial-temporal relationship, appearance, exclusion, and propagate labels in the graphs. They also using the principle of reference [26] to construct the exclusion graph and give the detection responses with the same time stamp (occurring at the same time) different labels. In reference [28], they propose spatio-temporal constraint embodies the tracklet association among different targets, e.g. paratactic or serial relation between two trajectories. There are some conventional constrains not introduced above, such as trajectory smooth, integrity of trajectory and et al.

D. Integrated model

In some situations, appearance models alone are not adequate to discriminate the objects, particularly for separating instances of the same types (e.g., pedestrians), since their shapes and textures look similarly. On the other hand, it is difficult to detect target when they were longtime occluded. So the integrated model can handle this problem efficiently. Reference [29] measure the similarity between two hypotheses by fusing several intrinsic properties, such as the color, size, spatio-temporal position and motion. In reference [10], the appearance model of an object using a color model, a sparse appearance model, a motion model and spatial information has been proposed. Reference [30] proposes an online learned Hough forest framework which effectively combines motion and appearance information for discrimination between two tracklets.

III. TRACKING METHODS

With the advances in object detection, *tracking-by-detection* method has recently become a popular paradigm for object tracking [9][65]. Most of the recent approaches that aim to solve the VMOT problem follow two main steps: object detection and data association. In the detection stage, a pre-trained object detector is first applied to find some potential object locations in each frame of input video. Once the object candidates are found, it is represented by observation model. And tracking is mostly on linking different observation model to different targets. There are two main classes of tracking methods: Bayesian theory based and data association based, depending on the way they link the observation to target.

A. Bayesian theory-based

There are four tracking methods based on Bayesian filter theory, Kalman filter, particle filter, Bayesian framework and the RFSs Bayesian multiple target filtering-PHD.

1) Kalman filter. Kalman filtering is an efficient way to address multi-target tracking when the number of objects remains small. However, when the number of objects increases, identity switches become more frequent and are difficult to correct due to the recursive nature of the method [31]. Reference [32] use Kalman filter to establish object motion model, using the current object's information to predict object's position, so that it can reduce the search scope and search time of moving object to achieve fast tracking. In [33], measurement noise covariance will be adapted by result of a local template matching when occlusion happens.

2) Particle filter. One of the main advantages of the particle filter-based solutions is the ability to handle nonlinear systems and multi-modal distributions. However, particle filters are not suitable for high-dimensional state spaces as their computational complexity tends to exponentially grow with the number of state parameters [34]. In [35], integral images is used for efficiently computing the color features and edge orientation histograms, which allows a large amount of particles and a better description of the targets. Okuma et al [36] propose a mixture particle filter which is ideally suited to multi-target tracking as it assigns a mixture component to each player, and construct the proposal distribution using a mixture model that incorporates information from the dynamic models of each player and the detection hypotheses generated by Adaboost. Gonzalez-Duarte et al [37] reports a rao-blackwellized particle filter (RBPF) model for multiple object tracking. It is possible for RBPF method to compute some steps of particle filtering with different state estimators instead of using sampling methods for all steps.

3) Bayesian framework. Bayesian framework is a kind of probability graph model. Tracking can be formulated with a Bayesian framework by maximizing the posterior distribution of the state of the target at time t given past and current measurements. Reference [25] incorporates the RMN model within the Bayesian filtering framework and a data association method for the online multi-object tracking. Reference [26] presents a novel Bayesian framework for tracking pedestrians.

*4) PHD filter.*In recent years, the random finite set (RFS) theory for multi-target tracking has attracted considerable attention, which considers multiple object states and their observations as random finite sets (RFSs). The probability hypothesis density (PHD), the cardinality PHD (CPHD), and multi-Bernoulli are sub-optimal approximation, but more tractable alternative to the RFSs Bayesian multiple target filtering.

The PHD filter is an approximation of the multi-target Bayes filter, but it does not have a close form solution. The Gaussian mixture PHD filter is an analytic solution to the PHD filter for linear Gaussian multi-target models [38][39]. By using Gaussian component labels in GM-PHD filter, the identities of individual target can be obtained. Feng et al [40]propose a robust particle PHD filter which is based on forming variational approximation to the joint distribution of states and noise parameters at each frame separately; the state is estimated with a particle PHD filter and the measurement noise variances used in the update step are estimated with a fixed point iteration approach. Yoon et al propose to fuse global and local observations to make hybrid observations and incorporate these hybrid observations into PHD filtering [41].

B. Data association- based

Data association-based method formulates the tracking problem as selecting and clustering of the detected object over time. We divide these methods into two classes: local optimization-based data association and global optimization-based data association. Usually the association can be solved in an optimization framework using carefully designed cost functions.

1) Local optimization-based data association. In this kind of methods, data association only consider a few frames for solving the association problem, for example, two frames or a few more frames. The best example of such approaches is bi-partite matching and its extensions. Among them, Hungarian algorithm is a typical one, which is widely used in compute the affinity between tracklets and detection observation. Wu and Nevatia[42] define an affinity measurement based on position, size and color, and use the Hungarian algorithm to associate object hypotheses and detection responses at neighboring frames. Stauffer [43] first obtains tracklets by performing a conservative frame-to-frame correspondence, and then associates these tracklets by the Hungarian algorithm with an extended transition matrix that considers initialization and termination of each tracklet. Riahi and Bilodeau [10] calculating the affinity between the set of tracked targets and detection responses, a matching step is achieved by Hungarian algorithm. Zhong et al [44] considers the target tracking as a bipartite graph matching problem where the nodes of the bipartite graph correspond to the targets in two neighboring frames, and the edges correspond to the degree of the similarity measure between the targets in different frames. The correspondence problem is solved by the dynamic Hungarian algorithm.

2) Global optimization-based data association. Local optimization based approach can be replaced by global trajectory optimization over batches of frames. Dynamic Programming or Linear Programming can be applied to solve such global optimization frameworks. The global optimization methods mainly include graph-based, hierarchical data association-based, quadratic boolean problem-based and binary integer programing-based method.

a. Graph based approach

Graph based framework operate on directed graphs whose nodes represent places where objects have been detected [45]-[58]. Many graph-based approaches are proposed, such as network flow(NF), maximum weight independent set(MWIS),conditional Random Field(CRF), generalized maximum clique problem(GMCP), paratactic–serial tracklet graph(PSTG) and et al.

---Network flows. It is a kind of global optimal procedure and use a directed acycle graph where each edge has capacity. In the VMOT problem, nodes in the graph for network flow are usually low-level observations (detection responses or the short tracklets). Two nodes are connected by an arc means that they may represent the same target and the cost of arcs or edges in the graph indicates the agreement between two detection observations. To meet the flow balance requirement, a source node and a sink node corresponding to the start and the end of a trajectory are added to the original graph (see Fig.2). One trajectory corresponds to one flow path in the graph. The flow transited from the source node to the sink node equals to the number of trajectories or the objects in the video, and the cost to transit the flow from the source node to the sink node is the likelihood of all the association hypotheses.

Many papers have formulated the VMOT as the network flow problem, but they using different algorithm to find the global optimized solution. Most of these methods are summarized in Table II.

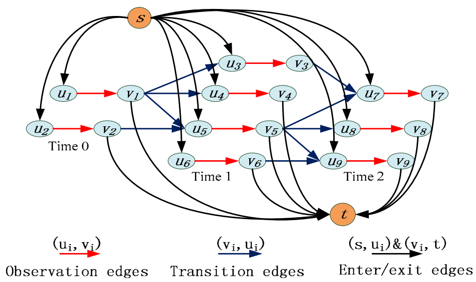

Fig.2 The model of network flow[45]

TABLE II. The comparison among the methods which solve the NF

the papers using NF	the methods to solve the NF	distinguishing aspect
Jiang H et al,2007[52]	linear programing	models object tracking as a multi-path searching problem
Zhang et al,2008[45]	push relable	first proposed the model
Berclaz J et al,2011[49]	k-shortest paths	first proposed
Pirsiavash et al,2011[46]	greedy method	first proposed
Wu Z et al,2012[50]	min-cost flow algorithm	first proposed
Butt A et al,2013[47]	Lagrangian relaxation	first proposed
Chari V et al,2014[51]	min-cost flow algorithm	Explicit Occlusion Model(EOM)
Xi Z et al,2015[2]	improved greedy algorithm	A* algorithm with dynamic weights
Dehghan A et al,2015[55]	Lagrangian relaxation	Target Identity
He Z et al,2015[48]	greedy method	integrated observation model handle occlusion,, a linear hypothesis method is proposed to fill up the gaps in the trajectories
McLaughlin Net al,2015[5]	linear programming	incorporating a motion model

-Conditional Random Field (CRF). This approach defines a graph $G = (V,E)$, where V is the set of vertexes and E is the set of edges between vertexes. Each vertex in the graph is defined as a pair of tracklets, and a label is predicted to indicate whether this pair of tracklets can be linked or not. These labels compose the label map which corresponds to the optimal association of the tracklets for the VMOT problem.

Yang and Nevatia [53] propose a framework for considering both the affinities between the tracklets and the dependencies among local associations by a Conditional Random Field (CRF) model. The CRF model considers both global descriptors for distinguishing different targets as well as pairwise descriptors for differentiating difficult pairs [7]. Milan et al [27] address mutual exclusion using a mixed discrete-continuous CRF that explicitly models two types of constraints, the exclusion between conflicting observations with super modular pairwise terms, and the exclusion between trajectories by generalizing global label costs to suppress the co-occurrence of incompatible labels (trajectories). In [17], they propose a new CRF model that exploits a lot more of the image evidence, including high-level detector responses and low-level super pixel information, fully automated segmentation and tracking of an unknown number of targets.

-Maximum weight independent set (MWIS).The MWIS is the heaviest subset of non-adjacent nodes of an attributed graph. Concerning the VMOT problem, the nodes represent candidate matches from every two consecutive frames, referred as tracklets, node weights encode the similarity of the corresponding matches; and edges connect nodes. Given this graph, the data association problem is modeled as the MWIS problem [11].

-Generalized Maximum Clique Problem (GMCP). GMCP-based data association method is a fully global framework. This method of data association assumes no simplification in the problem formulation and considers a model which is closer to the tracking scenario in real world. In the graph, the input to data association problem for finding tracklets is a graph $G = (V,E\ w)$, where V, E and w denote the set of nodes, set of edges and weights of edges, respectively. Reference [54] divides the input video into a number of segments and finds the tracklet of pedestrians within each segment using the proposed global method for tracklet generation utilizing GMCP. Reference [55] formulate tracking as a GMCP which based on reference [54], their graph incorporates the pairwise relationship between all the observations in a batch of frames and the cost function allows to incorporate higher order costs between all the candidates in a track.

-Energy minimization. This method focus on designing an energy function that corresponds to a more complete representation of the problem, rather than one that is amenable to global optimization. It defines a global energy function which depends on all targets at all frames within a temporal window, and thus represents the existence, motion and interaction of all interest objects in the scene, the crucial property of an energy function is to approximate the true situation sufficiently well. In the graph, the energy function is made up of some terms, such as observation term based on image data and some constrain in tracking scene [13], then using some algorithm to find the energy minima as the best trajectory. In [13][56] , to find strong local minima of the proposed nonconvex energy, they construct a suitable optimization scheme that alternates between continuous conjugate gradient descent and discrete trans-dimensional jump moves.

-Paratactic serial tracklet graph (PSTG) [28]. PSTG theory is proposed for inter-tracklet analysis in multi-target tracking to avoid tracking failure caused by long-term occlusion within group or crossing trajectories. PSTG is defined to describe the correlation among all tracklets in spatio-temporal domain to model the mutual influence among trajectories. The paratactic tracklets may have similar motion patterns and the serial tracklets may represent the same target. Furthermore, a PSTG-based multi-label optimization algorithm which embodys proposal labeling, group labeling and track ID labeling is presented to make the trajectory estimation more accurate.

-Other graph based method. Wen et al construct an undirected hypergraph, in which the nodes represent the tracklets, and the hyperedges are constituted by multiple tracklets across the temporal domain [57]. Segal et al interpret data association and tracking as a single Switching Linear Dynamical System (SLDS) based on latent data association [58]. Reference [22] constructs the graph by the locally linear embedding (LLE) of either the spatio-temporal or the appearance features associated to the detections. Once the graph has been defined, the multi-object tracking is formulated as the problem of finding a label assignment that is consistent with the constraints captured by each of the graphs. Chen et al [16] formulate tracking as labeling mid-level features by object identifiers, and specify a new approach, called constrained sequential labeling(CSL), which uses a cost function to sequentially assign labels while respecting the implications of hard constraints computed via constraint propagation

b. Quadratic boolean problem-based

Leibe et al [59]simultaneously optimize detection and tracking, coupled into a Quadratic Boolean Problem and solved with EM algorithm. The main benefit of quadratic programming is the ability to represent constraints between pairs of trajectories, such as coupling them to encourage similar motions.

c. Hierarchical data association-based

The hierarchical data association-based method is first proposed by Huang et al [63], in which they formulate tracking as a MAP problem and solve it using Hungarian algorithm to associate tracklets together. Similar to [60],other methods like in reference[1],also employ hierarchical association approach to multiple target tracking by progressively linking detection responses into longer track fragments, the difference between them is the model they used and the way they compute affinity. Huynh et al [61] proposes an efficient algorithm to learn both discriminative and generative appearance models online for different targets in a hierarchical framework. Within a time sliding window, detection responses are gathered to form short track fragments based on spatial-temporal information. These tracklets provide affinity and discriminative information, which can be used to collect samples for learning models. To learn appearance model, reference [62]uses AdaBoost to maximize discriminative capability. Kalman filter is also used to estimate the motion as an important factor to associate tracklets together.

d. Binary integer programing-based

Data association problem can be formulated as a Binary integer programing which pursues the minimum cost data associations among target measurements via one-to-one, one-to-*N*, and *N*-to-one associations. Reference [63] proposed a procedure for obtaining a Lagrangian dual relaxation solution for the binary integer programming problem. The procedure guarantees that its solution is integer-valued, and the duality gap of the solution is smaller than that of the simple LP relaxation solution.

IV. CONCLUSION

In this paper, we make a comprehensive review on proposed visual multiple object tracking. The review is based on the key issues of VMOT-observation model and tracking methods. As for the key aspects of VMOT approaches, we have presented the recent developments in details with classification and case study. We endeavor to organize them systematically and denote them to the reference papers where they appear. The readers can refer to the correspondent references if they find the method interested. Also, it includes our personal ideas through the whole paper. So we would appreciate very much for receiving comments from the readers that have different opinions.

REFERENCES

[1] C. Huang, Y. Li, and R. Nevatia,"Multiple target tracking by learning-based hierarchical association of detection responses," *Pattern Analysis & Machine Intelligence IEEE Transactions on,*vol. 35, no. 4, pp. 898-910, 2013.

[2] Z. Xi, D. Xu, W. Song, et al,"A *, algorithm with dynamic weights for multiple object tracking in video sequence," *Optik - International Journal for Light and Electron Optics*,vol. 126, no.20, pp. 2500-2507, 2015.

[3] W. Luo, X. Zhao,K. TaeKyun,"Multiple Object Tracking: A Review," *Eprint Arxiv*, 2014.

[4] C. H. Kuo, C.Huang, and R.Nevatia, "Multi-target tracking by on-line learned discriminative appearance models," *IEEE Conference on Computer Vision & Pattern Recognition*, vol. 238, pp. 685-692, 2010.

[5] G. S.Walia,and R. Kapoor, "Recent advances on multicue object tracking: a survey," *Artificial Intelligence Review,*vol. 46, no. 1, pp. 1-39,2016.

[6] N. Mclaughlin, J. M. D.Rincon, & P.Miller, "Enhancing Linear Programming with Motion Modeling for Multi-target Tracking," *Applications of Computer Vision (WACV), 2015 IEEE Winter Conference onIEEE*,pp.71-77, 2015.

[7] B.Yang, R.Nevatia,"Multi-target tracking by online learning a CRF model of appearance and motion patterns," *International Journal of Computer Vision,*vol. 107, no.2,pp. 203-217,2014.

[8] W. Hu, W. Li, X. Zhang, et al,"Single and Multiple Object Tracking Using a Multi-Feature Joint Sparse Representation," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 37, no. 4, pp. 816-833,2015.

[9] C. Jia, Z. Wang, X. Wu, and B.Cai, "A Tracking-Learning-Detection (TLD) method with local binary pattern improved," *IEEE International Conference on Robotics and Biomimetics,IEEE*,2015.

[10] D. Riahi, G.A.Bilodeau, "Multiple object tracking based on sparse generative appearance modeling," *IEEE International Conference on Image Processing*IEEE, 2015.

[11] W. Brendel, M. Amer, S.Todorovic,"Multiobject tracking as maximum weight independent set," *Computer Vision and Pattern Recognition IEEE Conference on. IEEE*,pp.1273-1280, 2011.

[12] T. Suwannatat, K. Chinnasarn, N. Indra-Payoong,"Multi-features particle PHD filtering for multiple humans tracking," *International Computer Science and Engineering Conference. IEEE*, 2015.

[13] A. Andriyenko, K.Schindler,"Multi-target tracking by continuous energy minimization," *IEEE Conference on Computer Vision & Pattern Recognition*, 2011, pp.1265-1272.

[14] R. T. Collins, "Multitarget data association with higher-order motion models,"*Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE,* 2012, pp. 1744-1751.

[15] H.S. Ben, J. Berclaz, F. Fleuret, et al,"Multi-Commodity Network Flow for Tracking Multiple People,"*IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 36, no. 8, pp. 1624-27, 2014.

[16] S. Chen, A. Fern, S.Todorovic,"Multi-object Tracking via Constrained Sequential Labeling,"*Computer Vision and Pattern Recognition. IEEE,* 2014, pp. 1130-1137.

[17] A. Milan, L. Lealtaixe, K. Schindler, et al,"Joint tracking and segmentation of multiple targets,"*Computer Vision and Pattern Recognition. IEEE*, 2015.

[18] Y. Chen, Y. Shen, X. Liu, et al,"3D object tracking via image sets and depth-based occlusion detection," *Signal Processing*,vol. 112, pp. 146-153, 2015.
[19] L. Wang, N. T. Pham, T. Ng, et al,"Learning deep features for multiple object tracking by using a multi-task learning strategy,"*IEEE International Conference on Image Processing. IEEE*, 2015, pp. 838-842.
[20] X. Wang, Q.Wang,"Coupled data association and l1 minimization for multiple object tracking under occlusion,"*SPIE/COS Photonics Asia. International Society for Optics and Photonics*, vol. 9273, 2014.
[21] B. Wu, S. Lyu, B. G. Hu, et al,"Simultaneous Clustering and Tracklet Linking for Multi-face Tracking in Videos," *IEEE International Conference on Computer Vision. IEEE Computer Society*, 2013, pp. 2856-2863.
[22] K. C. A. Kumar, C. D. Vleeschouwer, "Discriminative Label Propagation for Multi-object Tracking with Sporadic Appearance Features,"*Computer Vision, 2013 IEEE International Conference on.* IEEE, 2013, pp. 2000-2007.
[23] C. Dicle, O. I. Camps,M. Sznaier,"The Way They Move: Tracking Multiple Targets with Similar Appearance," *ICCV*, 2013, pp. 2304-2311.
[24] B. Yang, R.Nevatia,"Multi-target tracking by online learning of non-linear motion patterns and robust appearance models,"*IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society*, 2012, pp. 1918-1925.
[25] J. H. Yoon,M. H. Yang, J. Lim, et al, "Bayesian multi-object tracking using motion context from multiple objects,"*Applications of Computer Vision (WACV), 2015 IEEE Winter Conference on. IEEE*, 2015, pp. 33-40.
[26] L. Kratz , K.Nishino "Tracking Pedestrians Using Local Spatio-Temporal Motion Patterns in Extremely Crowded Scenes, " P*attern Analysis & Machine Intelligence IEEE Transactions on*, vol. 34, no.5, pp. 987-1002, 2011.
[27] A. Milan, K. Schindler, S.Roth,"Detection- and Trajectory-Level Exclusion in Multiple Object Tracking," *IEEE Conference on Computer Vision & Pattern Recognition, IEEE*, 2013,pp. 3682-3689.
[28] J.Chen,H. Sheng, C. Li, et al,"PSTG-based multi-label optimization for multi-target tracking,"*Computer Vision and Image Understanding*, vol.144, pp. 217-227,2016.
[29] Z. He, Y. Cui, H. Wang, et al,"One global optimization method in network flow model for multiple object tracking,"*Knowledge-Based Systems*, 2015.
[30] J. Xiang, N. Sang, J.Hou ,"An online learned hough forest model for multi-target tracking,"*Image Processing (ICIP), 2014 IEEE International Conference on. IEEE*, 2014, pp. 2398-2402.
[31] J. Berclaz, F. Fleuret, E. Türetken, et al,"Multiple object tracking using k-shortest paths optimization,"*Pattern Analysis and Machine Intelligence, IEEE Transactions on,*vol. 33, no. 9, pp. 1806-1819,2011.
[32] X. Li, K. Wang, W. Wang, et al,"A multiple object tracking method using Kalman filter"*IEEE International Conference on Information and Automation. IEEE*, 2010, pp. 1862-1866.
[33] M. Azari, A. Seyfi, A. H. Rezaie,"Real Time Multiple Object Tracking and Occlusion Reasoning Using Adaptive Kalman Filters,"*Machine Vision and Image Processing (MVIP), 2011 7th Iranian. IEEE*, 2011, pp.1-5.
[34] A. Vatavu, R. Danescu, S.Nedevschi,"Stereovision-Based Multiple Object Tracking in Traffic Scenarios Using Free-Form Obstacle Delimiters and Particle Filters,"*Intelligent Transportation Systems, IEEE Transactions on*, vol. 16, no. 1, pp. 498-511,2015.
[35] C. Yang, R. Duraiswami, L.Davis,"Fast Multiple Object Tracking via a Hierarchical Particle Filter,"*IEEE International Conference on Computer Vision. IEEE*, 2005, pp.212-219.
[36] K. Okuma, A. Taleghani, N. D. Freitas, et al,"A boosted particle filter: Multitarget detection and tracking,"*Computer Vision-ECCV 2004. Springer Berlin Heidelberg*, 2004,pp. 28-39.
[37] S. Gonzalez-Duarte, M. I. Chacon-Murguia,"Rao-blackwellized particle filter for multiple object tracking in video analysis,"*International Conference on Electrical Engineering, Computing Science and Automatic Control. IEEE*, 2014, pp. 328-31.
[38] H. Zhang, J. Yang,H. Ge, et al,"An improved GM-PHD tracker with track management for multiple target tracking,"*International Conference on Control, Automation and Information Sciences. IEEE,* 2015.
[39] M. Khazaei, M. Jamzad,"Multiple human tracking using PHD filter in distributed camera network,"*Computer and Knowledge Engineering , 2014 4th International eConference on. IEEE,* 2014, pp. 569-574.
[40] P. Feng,W. Wang, S. M. Naqvi, et al,"Variational Bayesian PHD Filter with Deep Learning Network Updating for Multiple Human Tracking,"*Sensor Signal Processing for Defence.IEEE,* 2015.
[41] J. H. Yoon, K. J. Yoon, D. Y. Kim,"Multi-object tracking using hybrid observation in PHD filter,"*Image Processing , 2013 20th IEEE International Conference on. IEEE,* 2013, pp. 3890-3894.
[42] B.Wu,R.Nevatia,"Detection and Tracking of Multiple, Partially Occluded Humans by Bayesian Combination of Edgelet based Part Detectors," *International Journal of ComputerVision*, 2007.
[43] C.Stauffer, "Estimating tracking sources and sinks,". *IEEE Workshop on Event Mining in Video*, 2003
[44] J. Zhong, J. Tan, Y. Li, et al,"Multi-Targets Tracking Based On Bipartite Graph Matching," *Cybernetics & Information Technologies*, vol. 14,no. 5, pp.78-87,2014.
[45] L. Zhang, Y.Li, R. Nevatia,"Global data association for multi-object tracking using network flows,"*Computer Vision and Pattern Recognition, 2008,*IEEE Conference on. IEEE, 2008, pp. 1-8.
[46] H. Pirsiavash,D. Ramanan, C. Fowlkes,"Globally-Optimal Greedy Algorithms for Tracking a Variable Number of Objects," *Proceedings of the 2011 IEEE Conference on Computer Vision and Pattern Recognition. IEEE Computer Society*, 2011, pp. 1201-1208.
[47] A. Butt, R.Collins,"Multi-target tracking by lagrangian relaxation to min-cost network flow," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition,*2013, pp.1846-1853.
[48] A. Dehghan, Y. Tian,P. H. S. Torr, et al,"Target Identity-aware Network Flow for online multiple target tracking,"*Computer Vision and Pattern Recognition. IEEE*, 2015, pp. 1146-1154.
[49] J. Berclaz, F. Fleuret, E. Türetken, et al,"Multiple object tracking using k-shortest paths optimization,"*Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 2011, 33(9): 1806-1819.
[50] Z. Wu, A. Thangali, S. Sclaroff, et al,"Coupling detection and data association for multiple object tracking," *Computer Vision and Pattern Recognition, 2012 IEEE Conference on. IEEE,* 2012, pp. 1948-1955.
[51] V. Chari, S. Lacostejulien, I. Laptev, et al,"On Pairwise Costs for Network Flow Multi-Object Tracking," *Mathematics*, 2014.
[52] H. Jiang, S. Fels, J. Little,"A Linear Programming Approach for Multiple Object Tracking,"*IEEE Conference on Computer Vision & Pattern Recognition.* 2007, pp.1-8.
[53] B. Yang, C. Huang, R.Nevatia,"Learning affinities and dependencies for multi-target tracking using a CRF model,"*IEEE Conference on Computer Vision & Pattern Recognition. IEEE*, 2011:1233-1240.
[54] A. R. Zamir, A. Dehghan, M.Shah,"GMCP-Tracker: Global Multi-object Tracking Using Generalized Minimum Clique Graphs,"*European Conference on Computer Vision,* 2012, pp.343-356.
[55] A. Dehghan, S. M. Assari, M. Shah,"GMMCP tracker: Globally optimal Generalized Maximum Multi Clique problem for multiple object tracking,"*IEEE International Conference on Computer Vision and Pattern Recognition*, 2015.
[56] A. Milan, S. Roth, K.Schindler,"Continuous energy minimization for multitarget tracking," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 36,no.1, pp. 58-72.2014.
[57] L. Wen, W. Li, J. Yan, et al,"Multiple target tracking based on undirected hierarchical relation hypergraph," *Computer Vision and Pattern Recognition, 2014 IEEE Conference on. IEEE*, 2014, pp. 1282-1289.
[58] A. V. Segal, I.Reid,"Latent Data Association: Bayesian Model Selection for Multi-target Tracking," *Proceedings of the 2013 IEEE International Conference on Computer Vision*. 2013, 2904-2911.
[59] B. Leibe, K. Schindler, L. Van Gool, "Coupled Detection and Trajectory Estimation for Multi-Object Tracking," *IEEE International Conference on Computer Vision*. 2007:1-8.
[60] C. Huang, B. Wu, R. Nevatia,"Robust Object Tracking by Hierarchical Association of Detection Responses," *European Conference on Computer Vision,*vol.5303, pp. 788-801, 2008.

[61] L.Huynh, D. Nguyen, T. B. Dinh, et al,"Online multiple object tracking by hierarchical association of detection responses," *The Nineth International Conference on Advances in Mobile Computing and Multimedia,*pp. 175-181, December, *2011,*
[62] C. Park, T. J. Woehl, J. E. Evans, et al, "Minimum cost multi-way data association for optimizing multitarget tracking of interacting objects," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* vol. 37, no.3, pp. 611-624, 2015.
[63] Z. Kalal, K. Mikolajczyk, J. Matas, "Tracking-Learning-Detection," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 2012, 34(7):1409-1422.