SEGMENT-WISE ONLINE LEARNING BASED ON GREEDY ALGORITHM FOR REAL-TIME MULTI-TARGET TRACKING

Changhoon Lee, and Chang D. Yoo

Korea Advanced Institute of Science and Technology
Department of Electrical Engineering
373-1 Guseong-dong, Yuseong-gu, Daejeon, 305-701, Korea

ABSTRACT

This paper proposes a tracklet-based algorithm for online multiple-target tracking. The algorithm performs tracking in three steps: (1) tracklet initialization, (2) tracklet refinement, and (3) tracklet association. Given detection responses, tracklets are initialized by finding a near-optimum path in the min-cost flow network using a greedy-based algorithm. Based on an appearance-based model, the tracklets are refined so that the detection responses within the tracklet become more homogeneous. Finally, the tracklets are linked based on a novel affinity measure, then by optimizing a min-cost flow network with links, the final tracks are generated. For real-time multi-target tracking, every step is processed in a segment-wise manner. On popular public datasets and strictly in an online fashion, the proposed multi-target tracking algorithm performed comparable to that of many state-of-the-art algorithms.

***Index Terms*—** Multi-target, tracking, online, tracklet, greedy algorithm

1. INTRODUCTION

Multiple-target tracking has received a great deal of interest for applications in surveillance, traffic control, activity recognition and sports video analysis. Occlusion and variations in appearance and illumination render it a difficult problem. Various algorithms have been proposed, and some notable state-of-the-art algorithms are as follows. Wang *et al.* [1] use appearance-model based metric learning to determine the affinity between the tracklets in generating an occlusion robust tracking algorithm. Pauwels *et al.* [2] use a mixture model of dense motion and stereo cues with feedback to generate robust tracklets. These algorithms achieve their state-of-the-art performances in a manner which is not strictly online. This is a serious limitation that hinder their applicability to many applications.

This work was supported by ICT R&D program of MSIP/IITP. [B0101-15-0307, Basic Software Research in Human-level Lifelong Machine Learning (Machine Learning Center)]

This paper proposes an online tracklet-based multi-target tracking algorithm. Contributions of this paper are as follows. (1) A simple greedy-based algorithm is constructed to find a near-optimum path in the min-cost flow network, and this path is used for initializing the tracklets. The entire tracking process is performed in real-time and provides highly accurate tracks. (2) An affinity score between the tail of a tracklet and the head of the subsequent tracklet is estimated by performing a 1-norm on the outer product of the detection responses of the two tracklet regions. The score is approximated using probes at the two regions, and this allows accurate association between tracklets to be made with low complexity.

The rest of the paper is organized as follows: Section 2 discusses two maximum-a-posteriori (MAP) data association problems in the context of cost-flow network for tracklet initialization and tracklet association. Section 3 describes an online greedy learning algorithm. Section 4 discusses experimental results, and finally, Section 5 concludes the paper.

2. COST-FLOW NETWORK FOR TRACKLET INITIALIZATION AND ASSOCIATION

In a multi-target environment with frequent occlusion, the cost-flow network provides a systematic and efficient framework for inferring tracks that are relatively occlusion robust. Here, two separate cost-flow networks are constructed for tracklet initialization and association: the node represent detection response in one network and tracklet in the other. The details of the two networks are discussed below.

2.1. Cost-flow Network with Detection Responses

To initialize the tracklets, a network with object observation nodes analogous to that considered in [3] is constructed. Let $x = (p, s, t) \in \mathcal{X}$ be the space-time *location* of an object such that p, s and t are respectively the position, scale and frame index, and assume a human detector finds N potential locations $\{x_1, x_2, \ldots, x_N\}$ where $x_i \in \mathcal{X}$ for $i \in \mathbb{Z}_N$ [1]. Let y_i

[1]For postive integer n, let $\mathbb{Z}_n$ be the set $\{1, 2, \ldots, n\}$

denote the feature vector at location x_i. A tracklet defined as a fragmented track or trajectory is a frame-ordered list of space-time locations such that the k^{th} tracklet is defined as $V_k = \{x_{k_i} | x_{k_i} \in \mathcal{X}, k_i \in S_k, i \in \mathbb{Z}_{|S_k|}\}$, where $S_k \subset \mathbb{Z}_N$. Here, S_k is a set of indices for estimated detection responses(space-time location) on the k^{th} tracklet, and $|S_k|$ is the cardinality of S_k. The set of all tracklets found in a video is denoted as V. In the tracklet initialization in Section 3.2, the objective is to estimate a set of tracklets V that maximizes its posterior probability. This can be formulated as minimizing the following objective function:

$$\mathbf{f}^* = \underset{\mathbf{f}}{\operatorname{argmin}} \left(\sum_i c_i^s f_i^s + \sum_{i,j \in \mathbf{E}} c_{ij} f_{ij} + \sum_i c_i f_i + \sum_i c_i^t f_i^t \right), \tag{1}$$

$$\begin{aligned} \text{where} \quad & c_i^s = -\log P_s(x_i),\ c_{ij} = -\log P(x_j|x_i), \\ & c_i = -\log l(y_i),\ c_i^t = -\log P_t(x_i), \\ \text{s.t.} \quad & f_{ij}, f_i, f_i^s, f_i^t \in \{0,1\}, \\ \text{and} \quad & f_i^s + \sum_j f_{ji} = f_i = f_i^t + \sum_j f_{ij}. \end{aligned}$$

where P_s, P_t and P are the start, terminate, and transition probabilities of an object. Here l, f_i and f_{ij} are respectively the feature likelihood, a binary indicator variable that takes the value 1 when x_i is included in a track and 0 otherwise, and binary indicator that takes the value 1 when x_i and x_j are included in consecutive frames of a track and 0 otherwise. Equation (1) is formulated using Integer Linear Programming (ILP) analogous to that used in [4]: the min-cost flows from the start node to the termination node should be obtained.

2.2. Cost-flow Network with Tracklet Units

To associate the tracklets, a network with tacklet nodes is constructed. Let $U = \{u_i | i \in \mathbb{Z}_L\}$ be a collection of L refined initialized tracklets discussed in Section 3.3. A single trajectory hypothesis is defined as a set of ordered sequence of tracklets such that the k^{th} trajectory is denoted as $T_k = \{u_{k_i} | u_{k_i} \in U, k_i \in R_k, i \in \mathbb{Z}_{|R_k|}\}$ where $R_k \subset \mathbb{Z}_L$. Here, R_k is a set of indices for estimated tracklet on the k^{th} trajectory, and $|R_k|$ is the cardinality of R_k. A set of all trajectories is defined by $\mathcal{T} = \{T_1, T_2, \ldots, T_{\mathcal{N}}\}$.

Tracklet association in Section 3.5 is performed by maximizing the a posterior probability (MAP) of $\mathcal{T}$ given U with the assumption that the likelihoods of u_{k_i} are conditionally independent such that

$$\begin{aligned} \mathcal{T}^* &= \underset{\mathcal{T}}{\operatorname{argmax}}\, P(\mathcal{T}|U) = \underset{\mathcal{T}}{\operatorname{argmax}}\, P(U|\mathcal{T})P(\mathcal{T}) \\ &= \underset{\mathcal{T}}{\operatorname{argmax}} \prod_i P(u_i|\mathcal{T})P(\mathcal{T}). \end{aligned} \tag{2}$$

Let us assume that tracklets are independent and one tracklet must belong to one trajectory. Then, we can further decompose the above equation as follows:

$$\mathcal{T}^* = \underset{\mathcal{T}}{\operatorname{argmax}} \prod_i P(u_i|\mathcal{T}) \prod_{T_k \in \mathcal{T}} P(T_k) \tag{3}$$

$$= \underset{\mathcal{T}}{\operatorname{argmax}} \prod_{T_k \in \mathcal{T}} P(T_k), \tag{4}$$

$$\text{s.t.} \quad T_k \cap T_l = \Phi, \forall\, k \neq l.$$

We assume that all the tracklets U are reliable and have no false alarms. Therefore Equation (3) can be simplified with $P(u_i|\mathcal{T}) = 1$. A priori probability of each single trajectory is modeled as a Markov chain defined by the product of the start probability $P_s(u_{k_1})$, the transition probability $P(u_{k_{n+1}}|u_{k_n})$, and the termination probability $P_t(u_{k_{N_k}})$ such that

$$\begin{aligned} P(T_k) &= P(u_{k_1}, u_{k_2}, \ldots, u_{k_{N_k}}) \\ &= P_s(u_{k_1}) \left(\prod_{n=1}^{N_k - 1} P(u_{k_{n+1}}|u_{k_n}) \right) P_t(u_{k_{N_k}}). \end{aligned} \tag{5}$$

Equation (4) can be reformulated as a min cost-flow problem to minimize the cost of flow from start s to termination t in a network flow graph. This can formulated as minimizing the following objective function:

$$\boldsymbol{\mathcal{F}}^* = \underset{\boldsymbol{\mathcal{F}}}{\operatorname{argmin}} \left(\sum_i \mathcal{C}_i^s \mathcal{F}_i^s + \sum_{i,j} \mathcal{C}_{ij} \mathcal{F}_{ij} + \sum_i \mathcal{C}_i^t \mathcal{F}_i^t \right), \tag{6}$$

$$\begin{aligned} \text{where} \quad & \mathcal{C}_i^s = -\log P_s(u_i), \mathcal{C}_{ij} = -\log P(u_j|u_i), \\ & \mathcal{C}_i^t = -\log P_t(u_i), \\ \text{s.t.} \quad & \mathcal{F}_i^s, \mathcal{F}_{ij}, \mathcal{F}_i^t \in \{0,1\}, \\ \text{and} \quad & \mathcal{F}_i^s + \sum_j \mathcal{F}_{ji} = \mathcal{F}_i^t + \sum_j \mathcal{F}_{ij}, \end{aligned}$$

Here $\mathcal{F}_i^s$, $\mathcal{F}_i^t$ and $\mathcal{F}_{ij}$ are binary indicator variables.

3. ONLINE LEARNING WITH GREEDY ALGORITHM

The proposed algorithm performs tracking in three steps: (1) tracklet initialization, (2) tracklet refinement, and (3) tracklet association. Given detection responses obtained as described in [5], a min-cost flow problem is formulated and solved using a greedy-based algorithm to initialize the tracklets. The tracklets are refined in a segment-wise manner using the appearance-based model discussed below such that it matches the true trajectory. Finally, the transition costs in the min-cost flow network of tracklets are evaluated.

3.1. Appearance-based Model

An appearance-based model is constructed for extracting discriminative features of the targets. To obtain a strong appearance-based cues, we start from a rich set of basic color, shape and texture features to describe a person's appearance. For the color feature, RGB, YCbCr and HSV color histograms are extracted into a 144-element vector. To capture shape information, we adopt the Histogram of Gradients (HOG) feature to form a 3968-element vector. In addition, two types of texture features are extracted using the Schmid and Gabor filters to form a 336-element vector. Thus, a person present in an image is represented as a feature vector of 4448-dimension. These feature vectors are used in tracklet refinement and computation of tracklet affinity score, and the distance between two features is calculated as the Euclidean distance of two vectors.

3.2. Tracklet Initialization with Greedy-based Algorithm

In the proposed algorithm, tracklets found up to the $(t-1)^{\text{th}}$ frame is used to estimate tracklets up to the t^{th} frame. The main idea of the proposed greedy algorithm for generating the initial tracklets is to elongate the existing tracklets so that it matches the true trajectory. This method is analogous to the algorithm in [4] in some ways, but the main difference is that tracklet initialization is performed in a segment-wise and iterative manner whereas in [4], the whole sequence is processed in a batch manner.

3.3. Tracklet Refinement

Tracklets initialized using only spatio-temporal information are refined such that tracklets are split into shorter tracklets with higher purity by considering both occlusions and interactions between objects. These refined shorter tracklets of higher purity are denoted as u_i. A probe determined as the detection response with the highest detection score of the M head frames of the tracklet, is selected to represent the tracklet and the distance between the feature extracted from the probe and other parts of the tracklet are evaluated and when the distance of $\mathcal{K}$ consecutive frames exceeds the threshold τ, the tracklet is split into two. We iterate this process multiple times until the tracklets is "pure". In our implementation, $M = 10$, $\mathcal{K} = 5$, and $\tau = 90$.

3.4. Transition Cost Computation Based on Tracklet Affinity Score

To determine the transition cost in the network discussed above, affinity scores between reliable tracklets pairs u_i and u_j are calculated. In order to determine u_i and u_j, the following conditions are checked with u_i belonging to either the current local segment L_n or the previous segment L_{n-1} while u_j belonging to the current segment L_n. Let t_i^e denote the end time of u_i while t_j^s be the start time of u_j then $t_j^s - t_i^e = \Delta t$ is the time gap between the tracklets. Also, let $p_i^{t_i^e}$ and $p_j^{t_j^s}$ denote for the positions of detection responses of u_i at the end time and u_j at the start time, respectively. Furthermore, let the instant velocity of u_i at the end time be v_i^e, and then reliable tracklets are determined satisfying the following conditions:

1. $\Delta t > 0$; there are no temporal overlaps between two tracklets.
2. $\Delta t < S$; the time gap between two tracklets does not exceed the maximum gap S.
3. $||p_j^{t_j^s} - (p_i^{t_i^e} + v_i^e \Delta t)|| < \gamma \Delta t$; the positional error estimate must be within the maximum range $\gamma \Delta t$.

For all possible transitions, transition costs are evaluated. First, similar to the interval of M frames at the head of a tracklet considered in the refinement procedure, a probe with the highest detection score within the interval of M frames at the tail end of the tracklet is considered. Let the feature vector at the tail probe of u_i be $\mathbf{X}_{TP_i}$ and that at the head probe of u_j be $\mathbf{X}_{HP_j}$. Let λ_j and λ_i denote the lengths of u_j and u_i, respectively, then we can calculate $m = \lceil \rho \lambda_j \rceil$, $n = \lceil \rho \lambda_i \rceil$, where $\rho < 1$. The m and n are the lengths of the matching intervals at the head of u_j and at the tail of u_i, respectively.

Next, we calculate the feature distance between a probe in one tracklet and the detection responses within the matching interval in the other tracklet. Accordingly, we formulate

$$d_{ij}^t = ||\mathbf{X}_{TP_i} - \mathbf{X}_j^t||,\; d_{ji}^{t'} = ||\mathbf{X}_{HP_j} - \mathbf{X}_i^{t'}||. \tag{7}$$

where d_{ij}^t and $d_{ji}^{t'}$ indicate each feature distance, and $\mathbf{X}_j^t$ and $\mathbf{X}_i^{t'}$ are the feature vectors of detection responses in the matching intervals at the head of u_j and at the tail of u_i, respectively. The average distance is defined as $\bar{d}_{ij} = \left(\sum_t d_{ij}^t + \sum_{t'} d_{ji}^{t'}\right) / (m+n)$ which can be evaluated at low complexity cost.

Finally, affinity score between i^{th} and j^{th} tracklets is defined as $S_{ij} = (\bar{d}_{ij})^{-1}$. We observe that the smaller $\bar{d}_{ij}$ is, the higher the score is. The transition cost has the relationship $\mathcal{C}_{ij} = -\log S_{ij}$.

3.5. Tracklet Association with Greedy-based Algorithm

In this step, the tracks found in the previous segment L_{n-1} is used to estimate the tracks in the current segment L_n. The greedy algorithm for generating final tracks tries to elongate existing tracks and find updated tracks. Similar to tracklet initialization, the minimum cost from the start node to node u_i, $\mathcal{S}_i$, is recursively calculated as $\mathcal{S}_i = \min(\pi, \mathcal{C}_i^s)$, where $\pi = \min_{j \in N(u_i)} \mathcal{C}_{ij} + \mathcal{S}_j$ and $N(u_i)$ is the set of nodes that can alter node u_i. Let $\mathbf{u}^n = \{u_1^n, \ldots, u_J^n\}$, where u_j^n represents the j^{th} node in the n^{th} segment, $\mathcal{S}^n = \{\mathcal{S}_1^n, \ldots, \mathcal{S}_J^n\}$ where $\mathcal{S}_j^n$ indicates the min-cost from the start node to u_j^n. Let $\mathcal{P}_j^n$ be the shortest path from start node to tracklet node u_j^n, then this include all tracklet nodes on the shortest path. The nodes of path $\mathcal{P}_j^n$ in the n_0^{th} segment are denoted as $\mathcal{P}_j^n(n_0)$. Residual graph $\mathcal{G}_r$ is the graph in which some tracks are removed from the original graph $\mathcal{G}$. Whenever $\mathcal{P}_j^n$ is assigned to be a track, $\mathcal{P}_j^n$ is removed from the graph, and the costs in the residual graph $\mathcal{G}_r$ is updated.

Algorithm 1 Greedy algorithm for tracklet association.

Definition:
Initialize the graph $\mathcal{G} = \Phi$, and the tracks $\mathcal{T} = \Phi$
while segment n in the given video sequence **do**
    V = do tracklet initialization
    $\mathbf{u}^n$ = do tracklet refinement
    Add $\mathbf{u}^n$ to $\mathcal{G}$
    Calculate $\mathcal{S}^n$ of $\mathcal{G}$ using DP
    Let $\mathcal{G}_r = \mathcal{G}$
    while $\mathcal{S}^n_{j^*} < 0$ where $j^* = \text{argmin}_j\ \mathcal{S}^n_j$ **do**
        if $\exists\, l$ s.t. $\mathcal{P}^n_{j^*}(n-1) \subset T_l$ **then**
            $T_l \leftarrow T_l \cup u^n_{j^*}$
            $\mathcal{G}_r =$ Build the residual graph $(\mathcal{G}_r, T_l)$
        else
            $\mathcal{T} \leftarrow \mathcal{T} \cup \mathcal{P}^n_{j^*}$
            $\mathcal{G}_r =$ Build the residual graph $(\mathcal{G}_r, \mathcal{P}^n_{j^*})$
        end if
        Update $\mathcal{S}^n$ of $\mathcal{G}_r$
    end while
end while
Output: Final tracks $\mathcal{T} = \{T_1, T_2, .., T_{\mathcal{N}}\}$

4. EXPERIMENTAL RESULTS

The performance of the proposed algorithm is evaluated against four popular multi-person tracking datasets: TUD-Stadtmitte, TUD-Crossing, ETH-BAHNHOF and ETH-SUNNYDAY. The performance is evaluated using evaluation tools [6], ground truth [7], and the popular evaluation metrics defined in [8]. The same pre-trained human detector used in [5] is used for fair comparison.

4.1. Comparison with Other Methods

The performance of the proposed algorithm is evaluated and compared to various state-of-the-art algorithms as shown in Table 1. The baseline result (BL) was obtained using the DP algorithm [3] for tracklet initialization. The BL and proposed algorithm (PA) results were obtained in real-time and in online manner while others were not. In terms of recall, PA performed better than [3] by 6.9%, 11.9%, 10.7% and 14.6% on the four datasets. The PA used the greedy algorithm to estimate the tracks in real-time based on [3], and it outperformed the algorithm based on globally optimal min-cost flow on all the measures.

4.2. Computational Speed

The computational speed depends on the number of targets in a video clip. The proposed method is implemented on a 2.6 GHz PC with 8 GB RAM using MATLAB under Mac OS X. The average speed is around 40 frames per second. This excludes detection and feature extraction time. The speed can be improved with code optimization and exploiting parallelization.

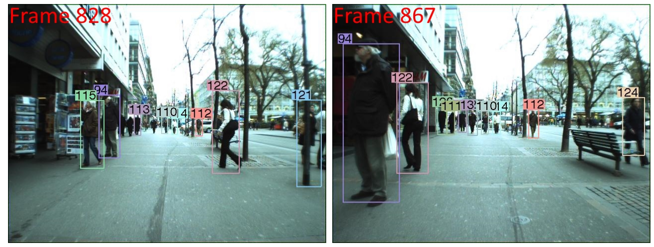

Fig. 1. Qualitative results on BAHNHOF sequences. We show robust tracking result while preserving their identities even significant occlusions due to index number 122 passing through from right to left in the view.

Dataset	Method	Rec.	Prec.	FAR	GT	MT	PT	ML	Frg	Ids
Stadtmitte	[9]	74.7	84.2	0.87	10	50.0	50.0	0.0	8	10
	[10]	81.0	99.5	0.03	10	60.0	30.0	10.0	0	1
	[7]	87.0	96.7	0.18	10	70.0	30.0	0.0	1	0
	[1]	98.0	99.3	0.04	10	100	0.0	0.0	3	0
	[3]	80.2	88.4	0.65	10	80.0	20.0	0.0	13	11
	BL	82.3	94.8	0.28	10	70.0	30.0	0.0	10	10
	PA	87.1	98.1	0.11	10	80.0	20.0	0.0	10	10
Crossing	[11]	78.8	56.6	-	-	42.3	30.8	26.9	8	5
	[3]	70.9	98.4	0.06	13	38.5	53.8	7.7	29	29
	BL	79.9	94.6	0.25	13	53.8	38.5	7.7	25	25
	PA	82.8	80.7	1.09	13	61.5	30.8	7.7	21	25
BAHNHOF	[11]	82.4	80.6	-	-	70.3	25.5	4.2	81	68
	[3]	59.9	95.1	0.26	94	33.0	51.0	16.0	166	131
	BL	69.7	88.8	0.74	94	45.7	42.6	11.7	149	76
	PA	70.6	86.6	0.93	94	46.8	41.5	11.7	152	80
SUNNYDAY	[11]	90.4	75.6	-	-	80.0	13.3	6.7	4	3
	[3]	64.4	96.8	0.12	30	33.3	40.0	26.7	9	13
	BL	76.6	91.9	0.36	30	46.7	30.0	23.3	4	13
	PA	79.0	89.9	0.47	30	50.0	26.7	23.3	9	18

Table 1. Results on four publicly available sequences.

5. CONCLUSION

This paper proposes a tracklet-based algorithm for online multiple-target tracking. The algorithm performs tracking in three steps: (1) tracklet initialization, (2) tracklet refinement, and (3) tracklet association. Given detection responses, tracklets are initialized by finding a near-optimum path in the min-cost flow network using a greedy-based algorithm. Based on appearance-based model, the tracklets are refined so that the detection responses within the tracklet become more homogeneous. Finally, the tracklets are linked based on a novel affinity measure, then by optimizing a min-cost flow network with links, the final tracks are generated. For real-time multi-target tracking, every step is processed in a segment-wise manner. On popular public datasets and strictly in an online fashion, the proposed multi-target tracking algorithm performed comparable to that of many state-of-the-art algorithms.

6. REFERENCES

[1] Bing Wang, Gang Wang, Kap Luk Chan, and Li Wang, "Tracklet association with online target-specific metric learning," in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 2014, pp. 1234–1241.

[2] Karl Pauwels, Lorenzo Rubio, and Eduardo Ros, "Real-time model-based articulated object pose detection and tracking with variable rigidity constraints," in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 2014, pp. 3994–4001.

[3] Hamed Pirsiavash, Deva Ramanan, and Charless C Fowlkes, "Globally-optimal greedy algorithms for tracking a variable number of objects," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 1201–1208.

[4] Tae-Ho Kim, Changhoon Lee, and Chang D Yoo, "Greedy algorithm for real-time multi-object tracking," in *Image Processing (ICIP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 398–402.

[5] RB Girshick, PF Felzenszwalb, and D McAllester, "Discriminatively trained deformable part models, release 5," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2012.

[6] Anton Milan, Konrad Schindler, and Stefan Roth, "Detection-and trajectory-level exclusion in multiple object tracking," in *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*. IEEE, 2013, pp. 3682–3689.

[7] Bo Yang and Ram Nevatia, "An online learned crf model for multi-target tracking," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 2034–2041.

[8] Yuan Li, Chang Huang, and Ram Nevatia, "Learning to associate: Hybridboosted multi-target tracker for crowded scene," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*. IEEE, 2009, pp. 2953–2960.

[9] Anton Andriyenko, Konrad Schindler, and Stefan Roth, "Discrete-continuous optimization for multi-target tracking," in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*. IEEE, 2012, pp. 1926–1933.

[10] Cheng-Hao Kuo and Ram Nevatia, "How does person identity recognition help multi-person tracking?," in *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*. IEEE, 2011, pp. 1217–1224.

[11] Laura Leal-Taixé, Michele Fenzi, Alina Kuznetsova, Bodo Rosenhahn, and Silvio Savarese, "Learning an image-based motion context for multiple people tracking," in *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. IEEE, 2014, pp. 3542–3549.