Learning to Track: Online Multi-Object Tracking by Decision Making

Yu Xiang[1,2], Alexandre Alahi[1], and Silvio Savarese[1]

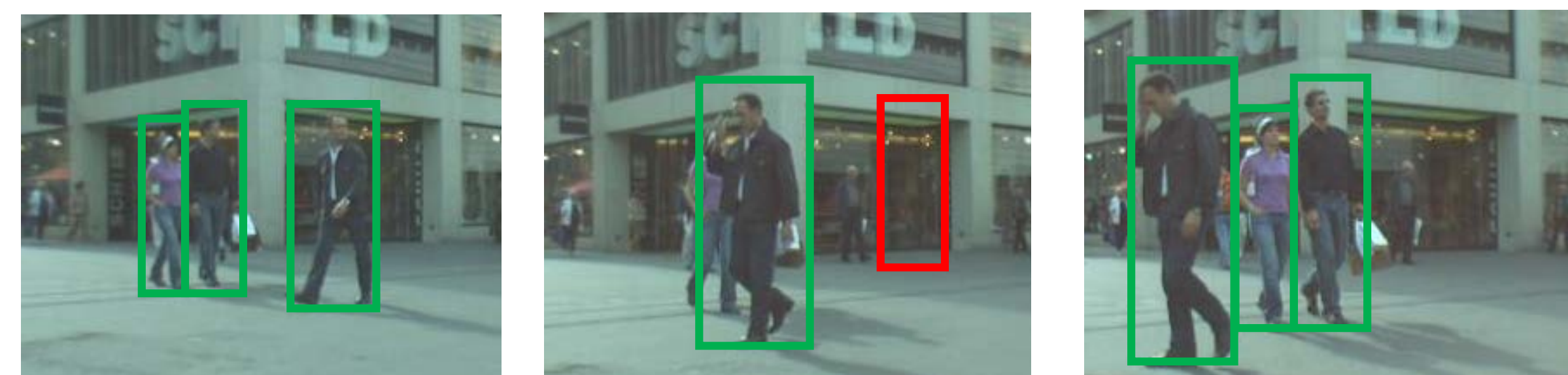[1]Stanford University, [2]University of Michigan

yuxiang@umich.edu, {alahi, ssilvio}@stanford.edu
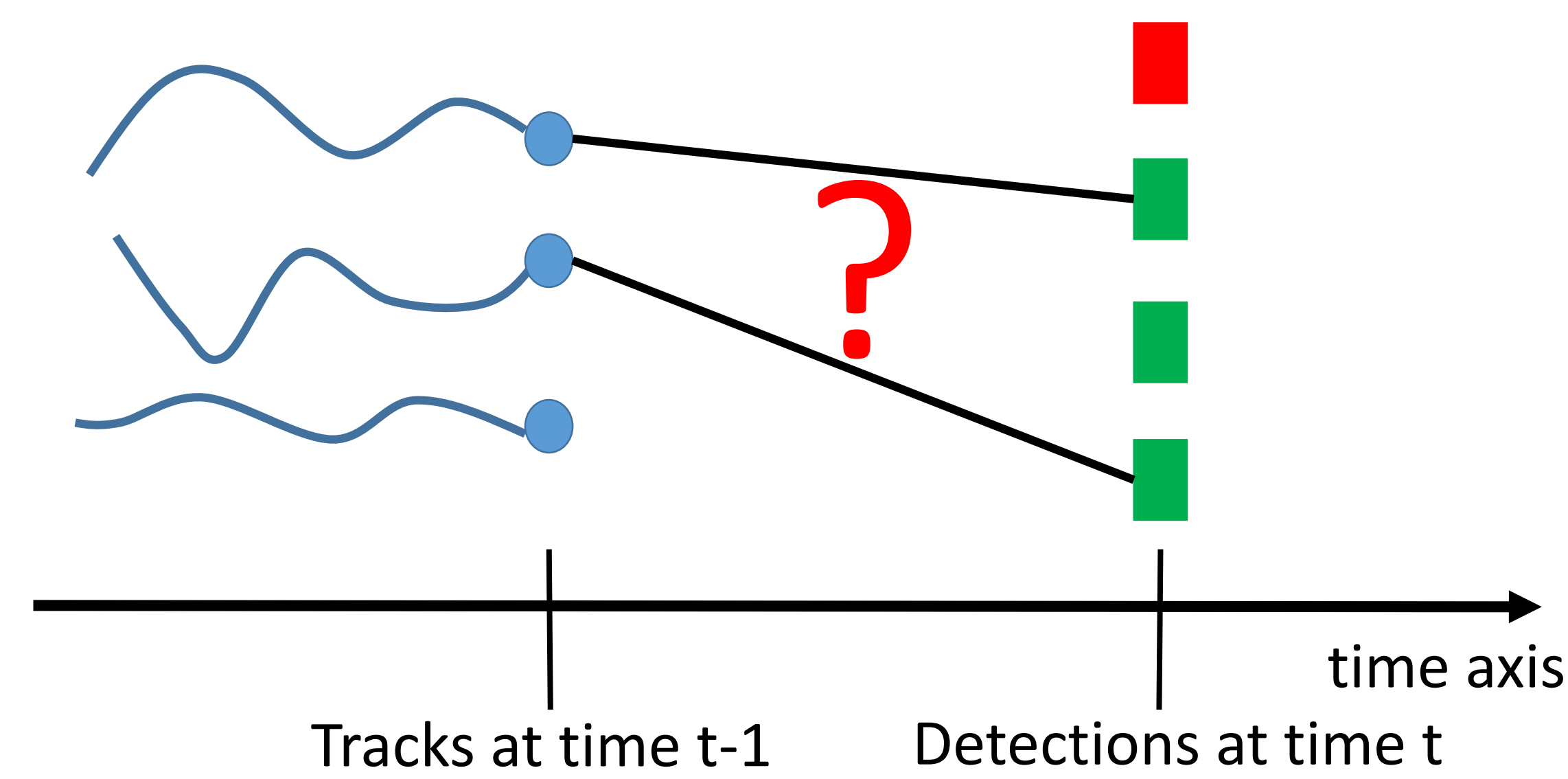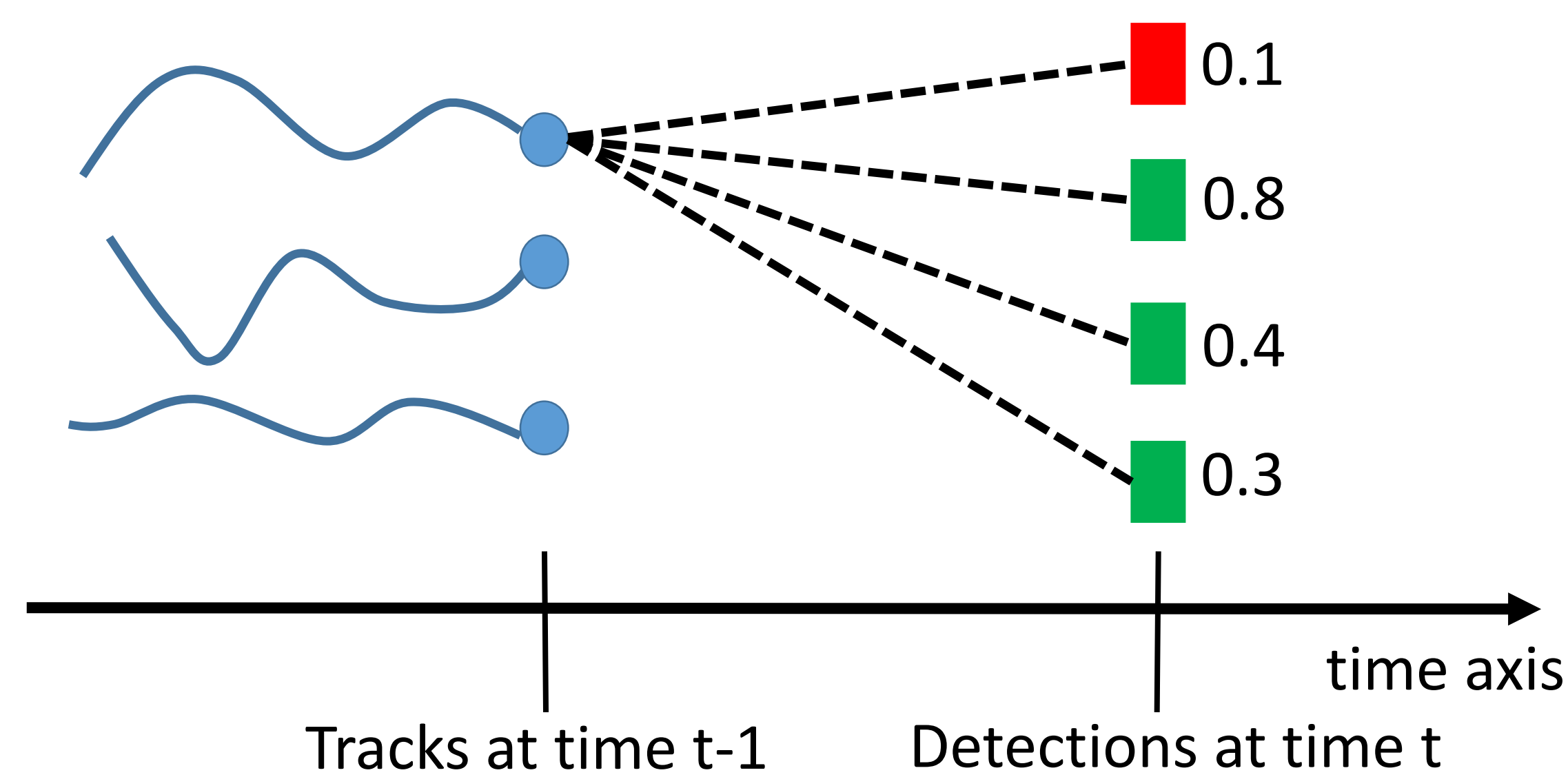
■ Goal: Online Multi-Object Tracking

■ Tracking by Detection

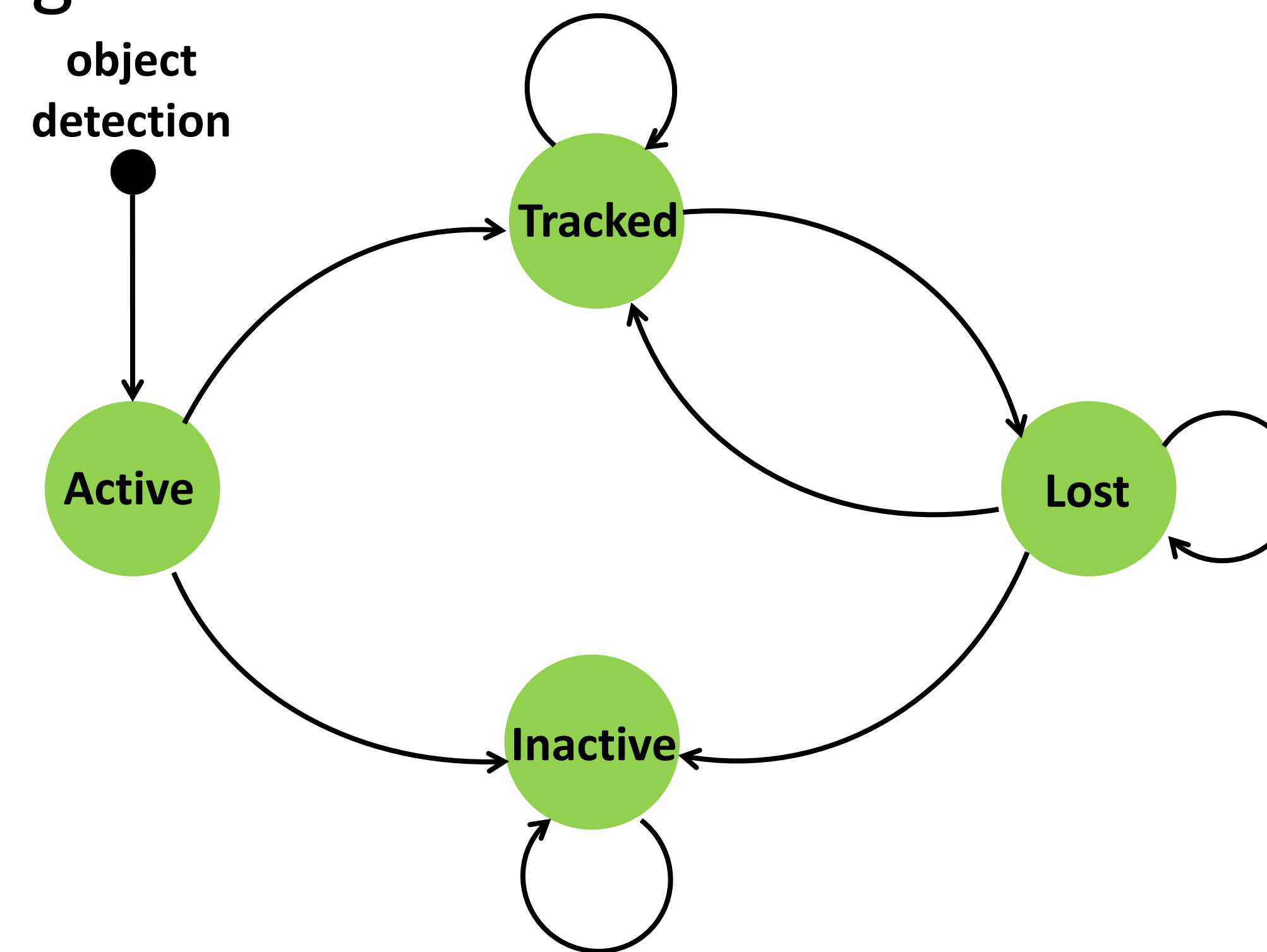## ■ Challenge: data association under occlusion and noisy detections

Tracks at time t-1   Detections at time t

■ Learning to Track: learn a similarity function for data association

0.1
0.8
0.4
0.3

Tracks at time t-1   Detections at time t

■ Our Solution: Tracking by Decision Making

The target is tracked   The target is occluded   The target is tracked again
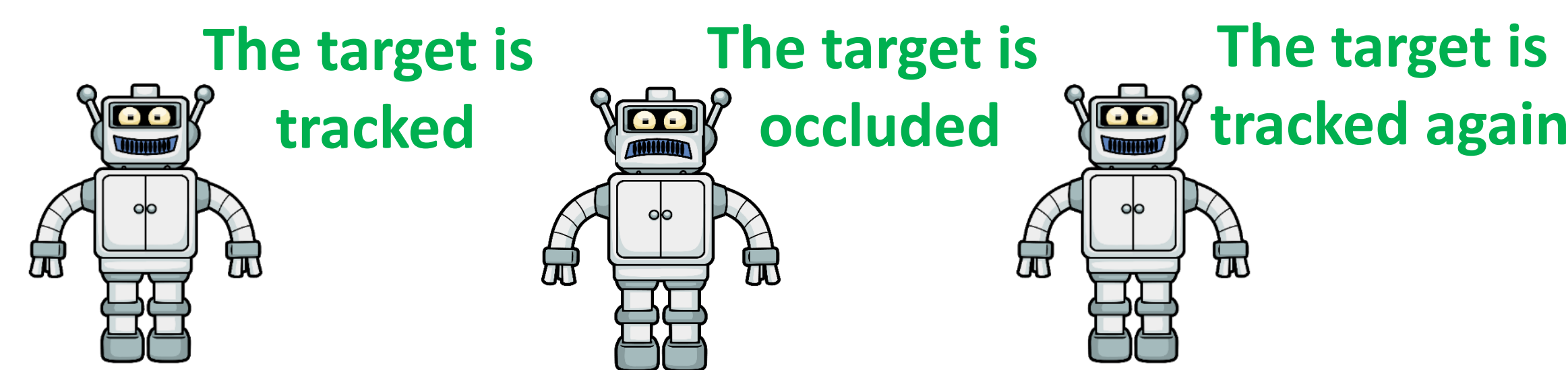
■ Markov Decision Process for a Single Target

object detection

Tracked

Active

Lost

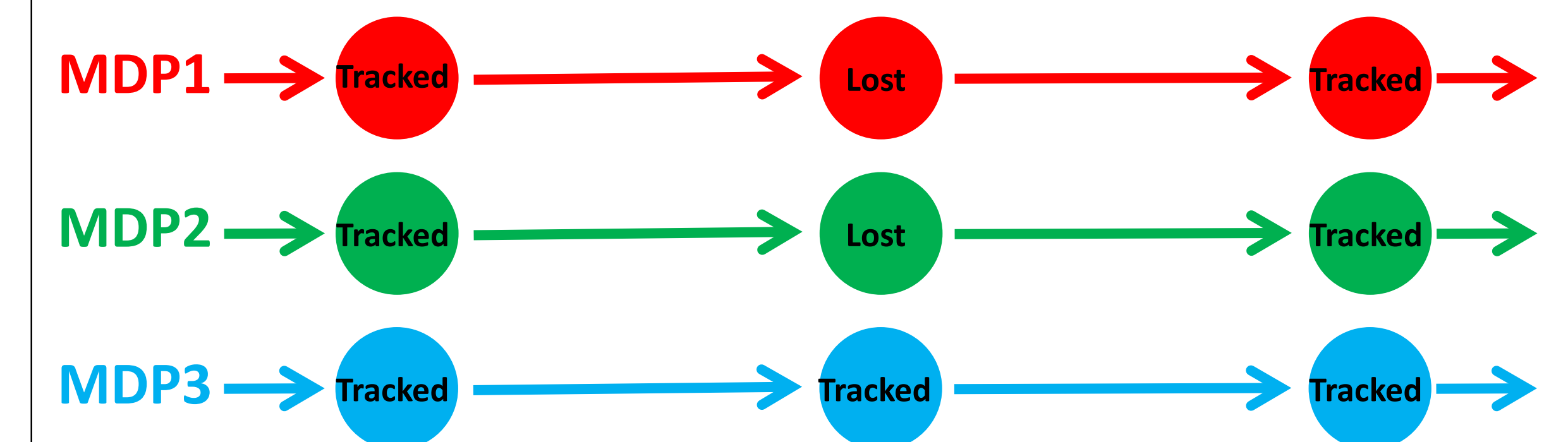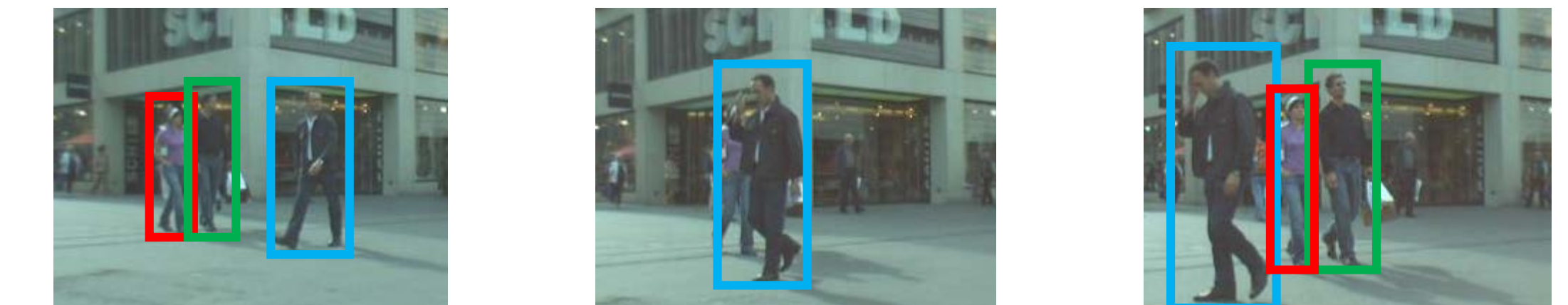Inactive

■ Inverse Reinforcement Learning

$$R_{\text{Active}}(s,a) = y(a)\big(\mathbf{w}_{\text{Active}}^T \phi_{\text{Active}}(s) + b_{\text{Active}}\big)$$

$$R_{\text{Tracked}}(s,a) = \begin{cases} y(a), & \text{if } e_{\text{medFB}} < e_0 \text{ and } o_{\text{mean}} > o_0 \\ -y(a), & \text{otherwise}, \end{cases}$$

(a) target template   (b) stable prediction   (c) unstable prediction

Frame 50   Frame 50   Frame 51   Frame 51   Frame 50   Frame 57   Frame 57

$$R_{\text{Lost}}(s,a) = y(a)\Big(\max_{k=1}^{M} \big(\mathbf{w}^T \phi(t,d_k) + b \big) \Big)$$

■ Online Multi-Object Tracking with MDPs

MDP1 → Tracked → Lost → Tracked →

MDP2 → Tracked → Lost → Tracked →

MDP3 → Tracked → Tracked → Tracked →

■ Experiments

Tracker	Mode	MOTA	MOTP	MT	ML	IDS	Frag
[1]	Batch	14.5	70.8	6.0%	40.8%	4,537	3,090
[2]	Online	15.1	70.5	3.2%	55.8%	**637**	1,716
[3]	Batch	15.9	70.9	6.4%	47.9%	1,939	1,963
[4]	Batch	18.2	71.2	2.8%	54.8%	1,148	2,132
[5]	Online	18.6	69.6	5.3%	53.3%	684	1,282
[6]	Batch	19.3	70.7	8.5%	46.5%	813	1,023
[7]	Batch	22.5	**71.7**	5.8%	63.9%	697	**737**
[8]	Batch	23.1	70.9	4.7%	52.0%	1,018	1,061
Ours	Online	**30.3**	71.3	**13.0%**	**38.4%**	680	1,500

TUD-Crossing #31   PETS09-S2L2 #68   PETS09-S2L2 #111   ETH-Jelmoli #82   ETH-Linthescher #51   ETH-Crossing #97

AVG-TownCentre #52   ADL-Rundle-1 #232   ADL-Rundle-3 #183   Venice-1 #235   KITTI-16 #90, KITTI-19 #281

[1] H. Pirsiavash, D. Ramanan, and C. C. Fowlkes. Globally-optimal greedy algorithms for tracking a variable number of objects. In CVPR, pages 1201–1208, 2011.
[2] S.-H. Bae and K.-J. Yoon. Robust online multi-object tracking based on tracklet confidence and online discriminative appearance learning. In CVPR, pages 1218–1225, 2014.
[3] A. Geiger, M. Lauer, C. Wojek, C. Stiller, and R. Urtasun. 3d traffic scene understanding from movable platforms. TPAMI, 36(5):1012–1025, 2014.
[4] C. Dicle, O. I. Camps, and M. Sznaier. The way they move: Tracking multiple targets with similar appearance. In ICCV, pages 2304–2311, 2013.
[5] J. H. Yoon, M.-H. Yang, J. Lim, and K.-J. Yoon. Bayesian multiobject tracking using motion context from multiple objects. In WACV, pages 33–40, 2015.
[6] A. Milan, S. Roth, and K. Schindler. Continuous energy minimization for multitarget tracking. TPAMI, 36(1):58–72, 2014.
[7] A. Milan, L. Leal-Taix´e, K. Schindler, and I. Reid. Joint tracking and segmentation of multiple targets. In CVPR, pages 5397–5406, 2015.
[8] L. Leal-Taixe, M. Fenzi, A. Kuznetsova, B. Rosenhahn, and S. Savarese. Learning an image-based motion context for multiple people tracking. In CVPR, pages 3542–3549, 2014.