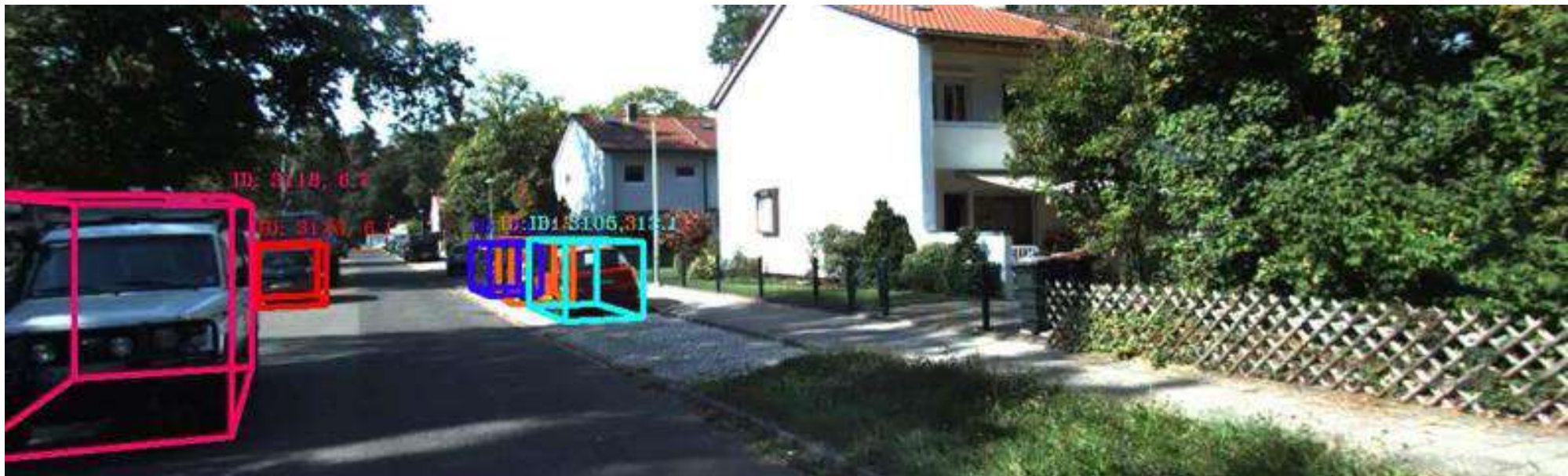


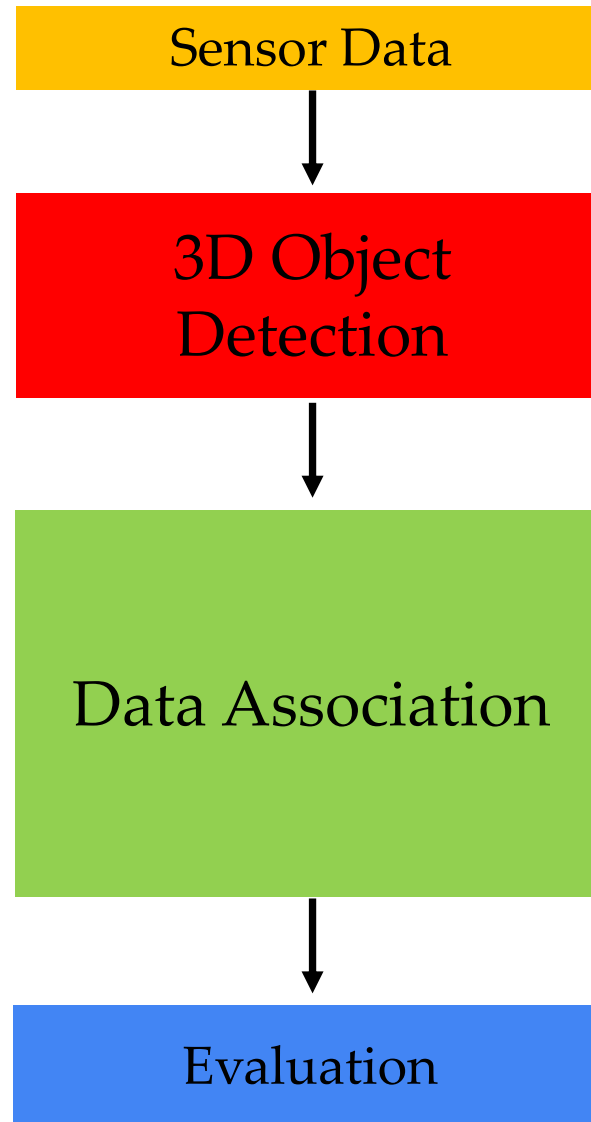
AB3DMOT: A Baseline for 3D Multi-Object Tracking and New Evaluation Metrics

Xinshuo Weng, Jianren Wang, David Held, Kris Kitani
Robotics Institute, Carnegie Mellon University

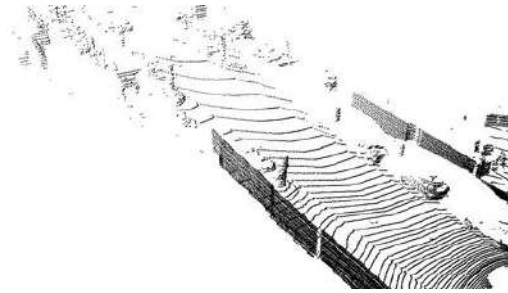
European Conference on Computer Vision (ECCV) Workshops, 2020



Standard 3D MOT Pipeline



Standard 3D MOT Pipeline



LiDAR point clouds

Sensor Data



3D Object Detection



Data Association

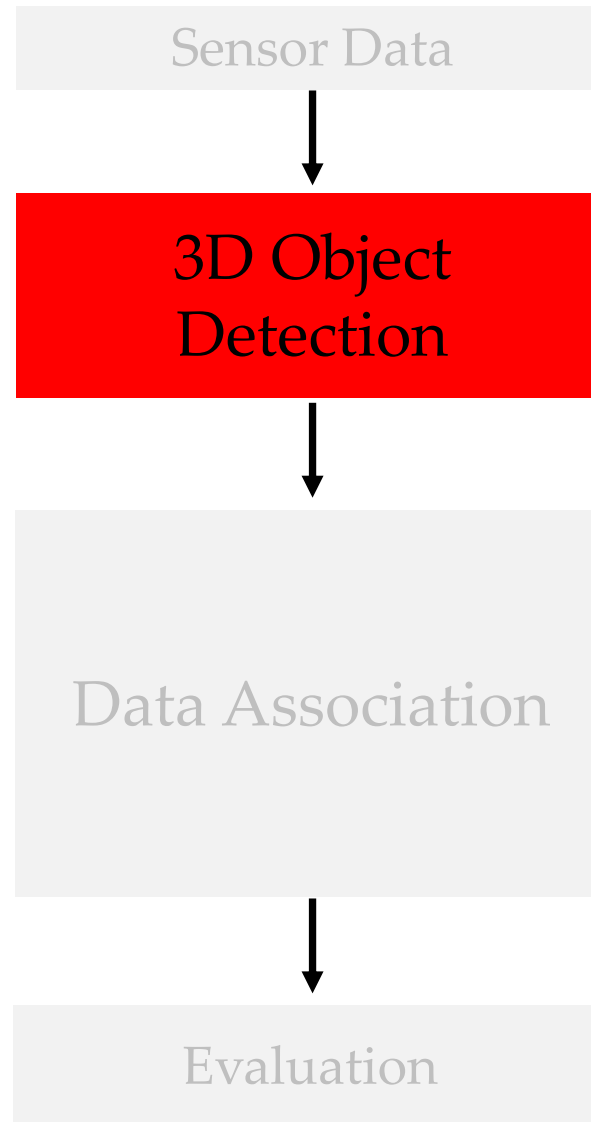


Evaluation



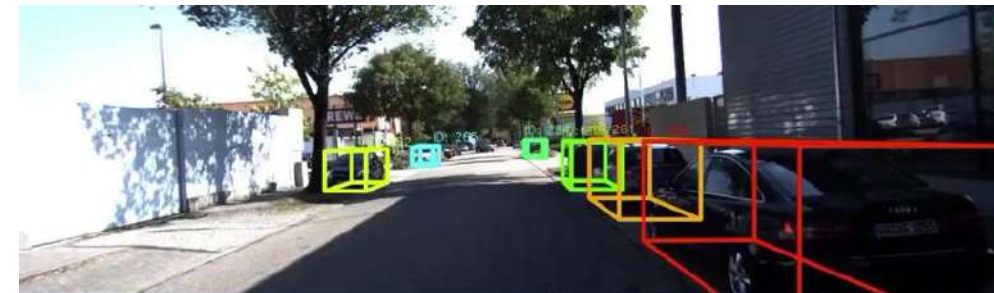
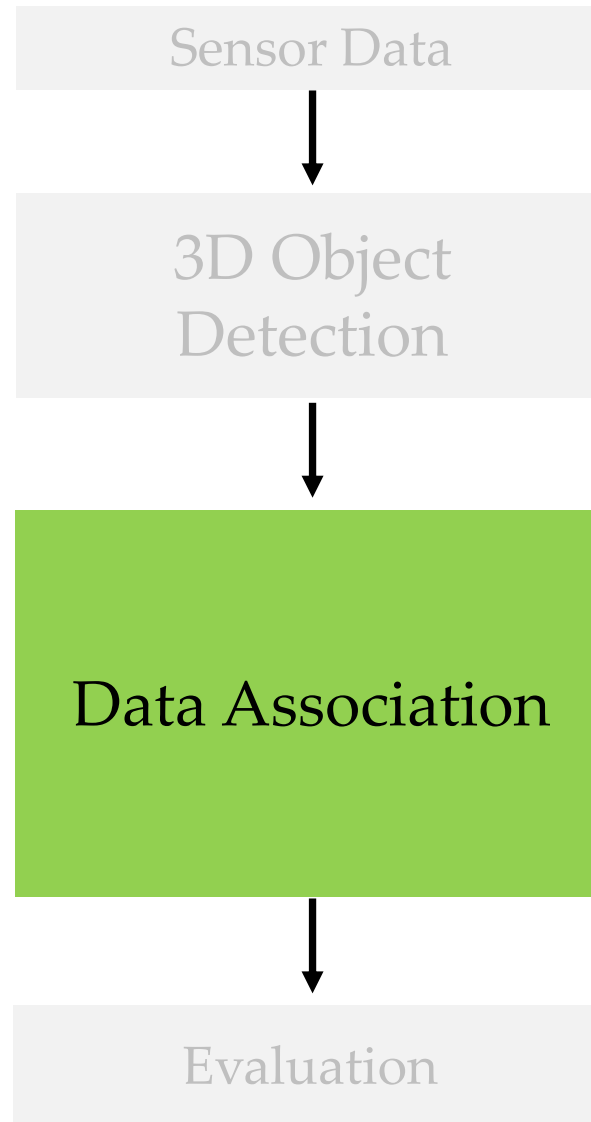
RGB frames

Standard 3D MOT Pipeline



Detection results

Standard 3D MOT Pipeline



3D MOT results

Standard 3D MOT Pipeline

Sensor Data



3D Object
Detection



Data Association



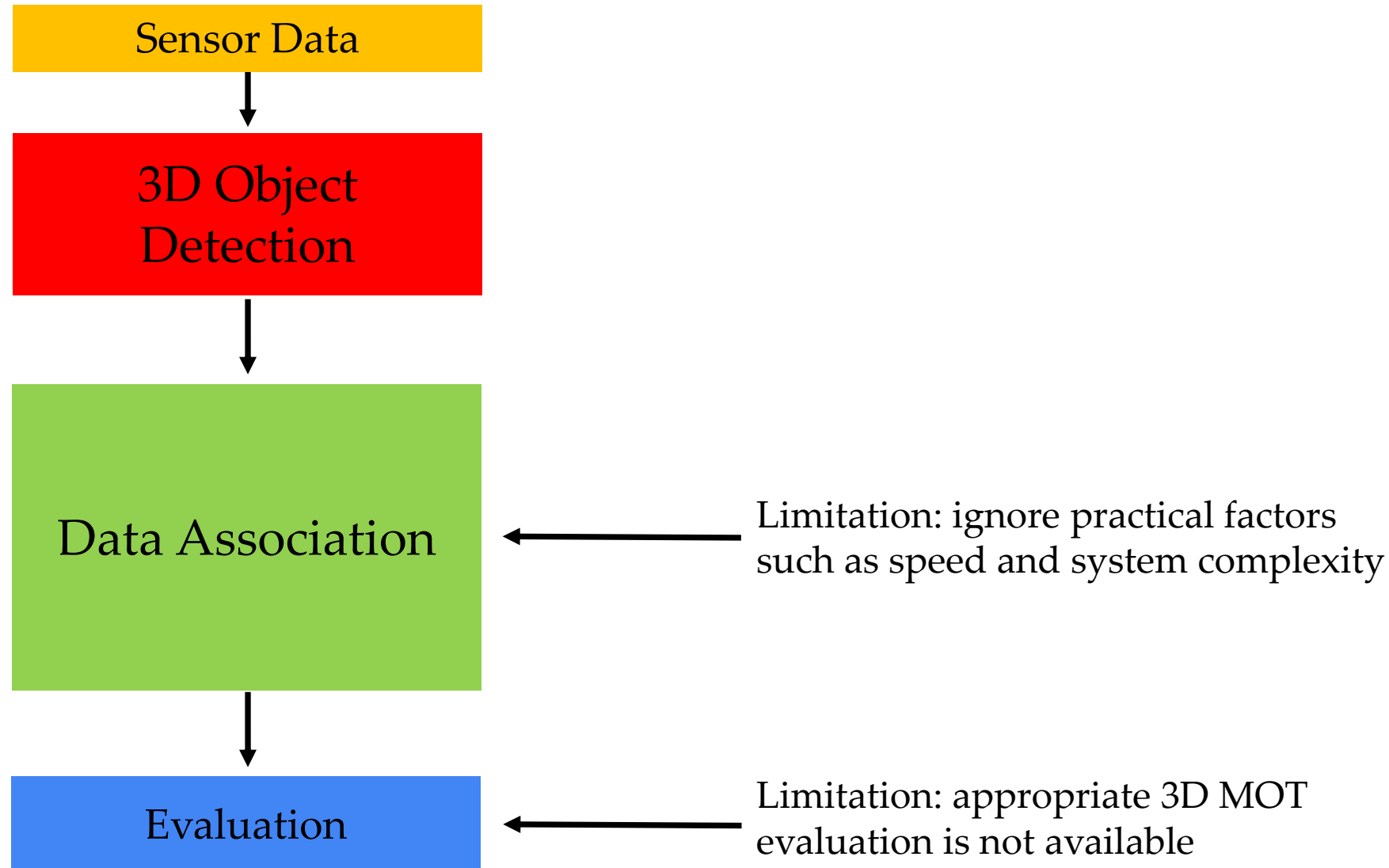
Evaluation

Evaluation:

1. MOTA: MOT accuracy
2. MOTP: MOT precision
3. IDS: # of identity switches
4. FRAG: # of trajectory fragments
5.

← Also important!

Standard 3D MOT Pipeline

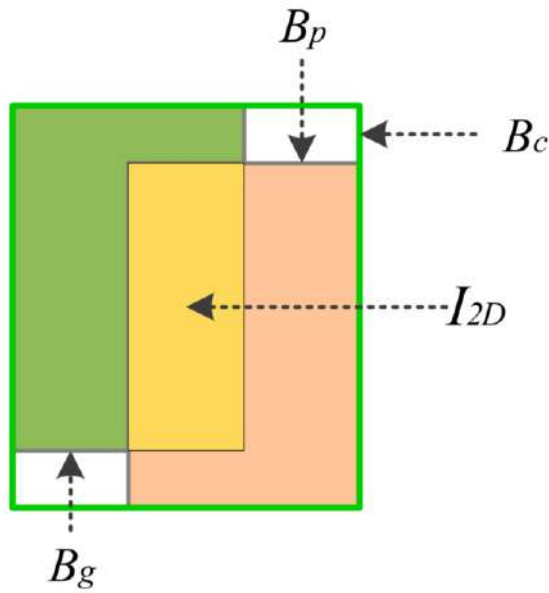


Our Contributions

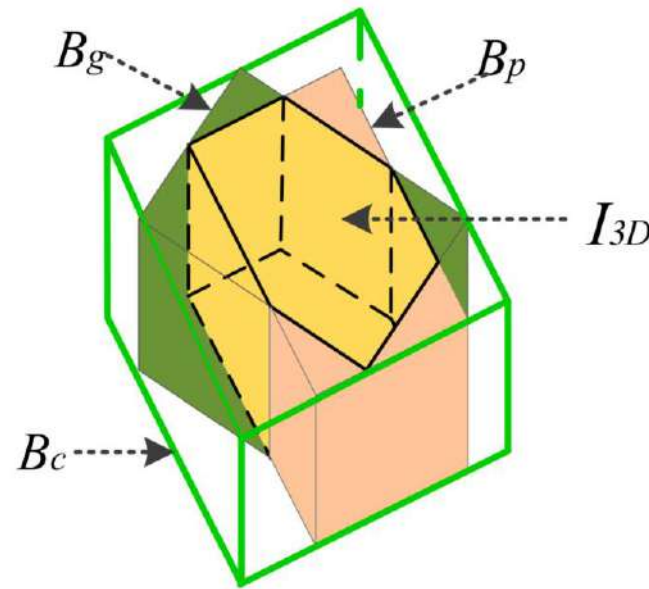
1. A 3D MOT evaluation tool along with three integral metrics
2. A strong and simple 3D MOT system with the fastest speed (207.4 FPS)

What are the Issues of 3D MOT Evaluation?

- Matching criteria: IoU (intersection of union)
- For the pioneering 3D MOT dataset KITTI, evaluation is performed in the 2D space
 - IoU is computed on the 2D image plane (not 3D)
- The common practice for evaluating 3D MOT methods is:
 - Project 3D trajectories onto the image plane
 - Run the 2D evaluation code provided by KITTI



IoU in 2D space



IoU in 3D space

B_p : the predicted box

B_g : the ground truth box

B_c : the smallest enclosing box

I_{2D} , I_{3D} : the intersection

Our Solution: Upgrade the Matching Criteria to 3D

- Replace the matching criteria (2D IoU) in the KITTI evaluation code with 3D IoU
 - <https://github.com/xinshuoweng/AB3DMOT> (800+ stars)
- Work with nuTonomy collaborators and use our 3D MOT evaluation metrics in the nuScenes evaluation with the matching criteria of center distance
 - <https://www.nuscenes.org/>



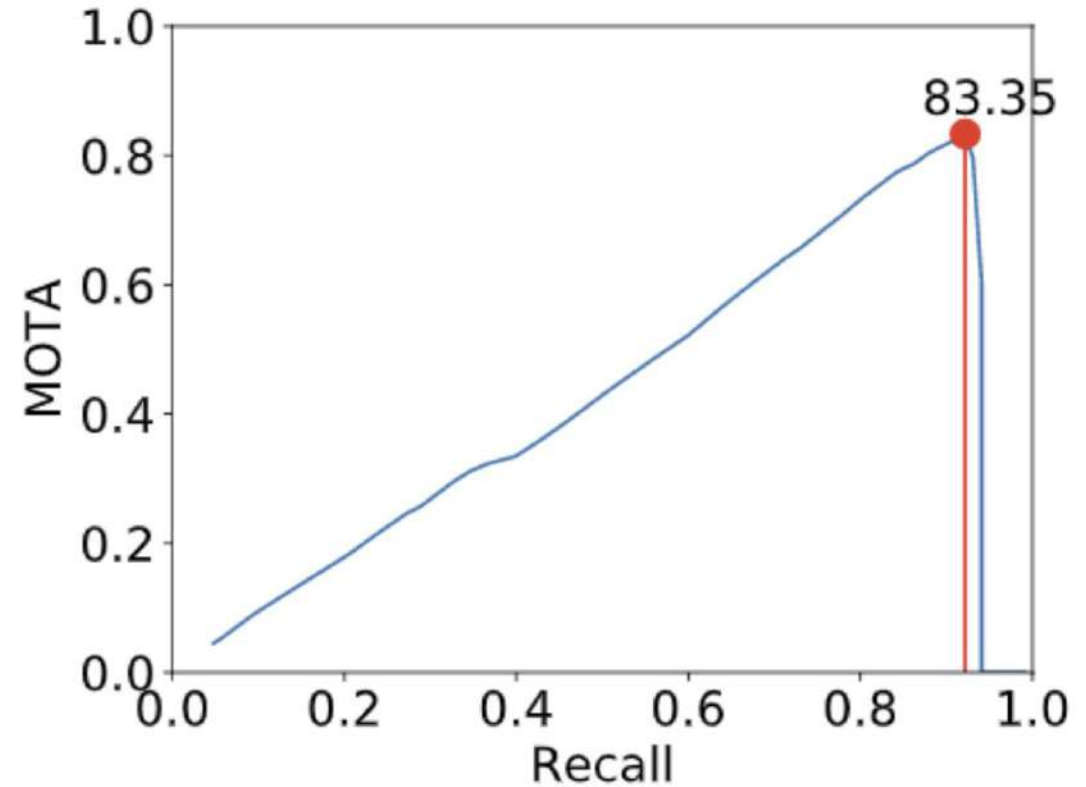
Our released new evaluation code



nuScenes 3D MOT evaluation with our metrics

What are the Issues of Evaluation?

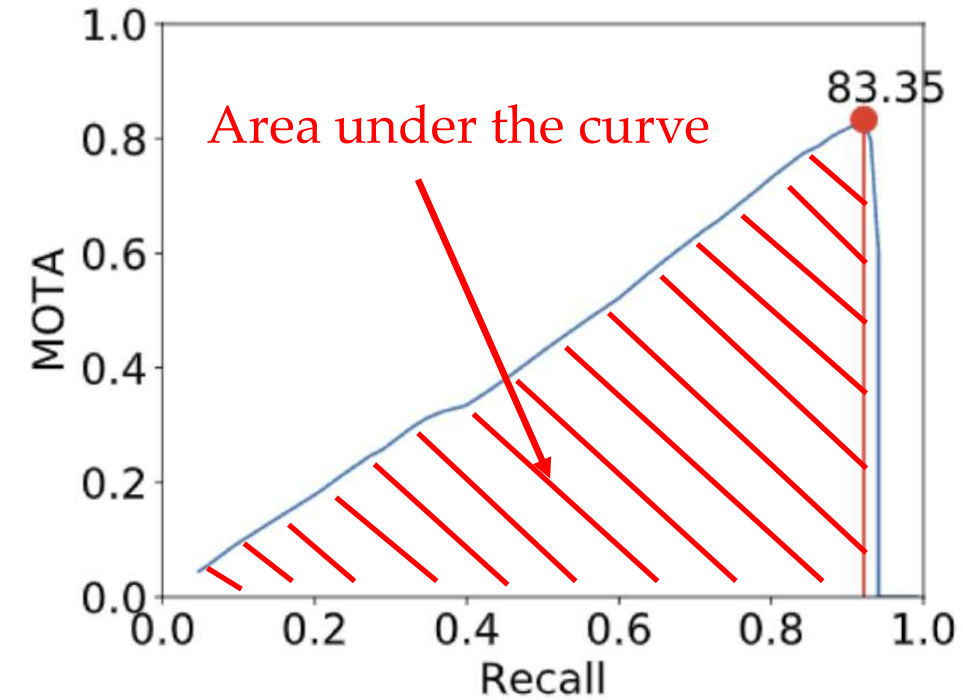
- Are we done with the evaluation? Can we further improve the current metrics?
 - E.g., MOTA (multi-object tracking accuracy)
 - $MOTA = 1 - \frac{FP + FN + IDS}{num_{gt}}$
- Performance is measured at a single recall point
 - Cannot understand the full spectrum of accuracy of a MOT system



MOTA over Recall curve

Our Solution: Integral Metrics

- MOTA is measured at a single point on the curve
- What can we do to improve the evaluation metrics?
- Compute the integral metrics through the area under the curve, e.g., average MOTA (AMOTA)
 - Analogous to the average precision (AP) in object detection
 - Can measure the full spectrum of MOT accuracy



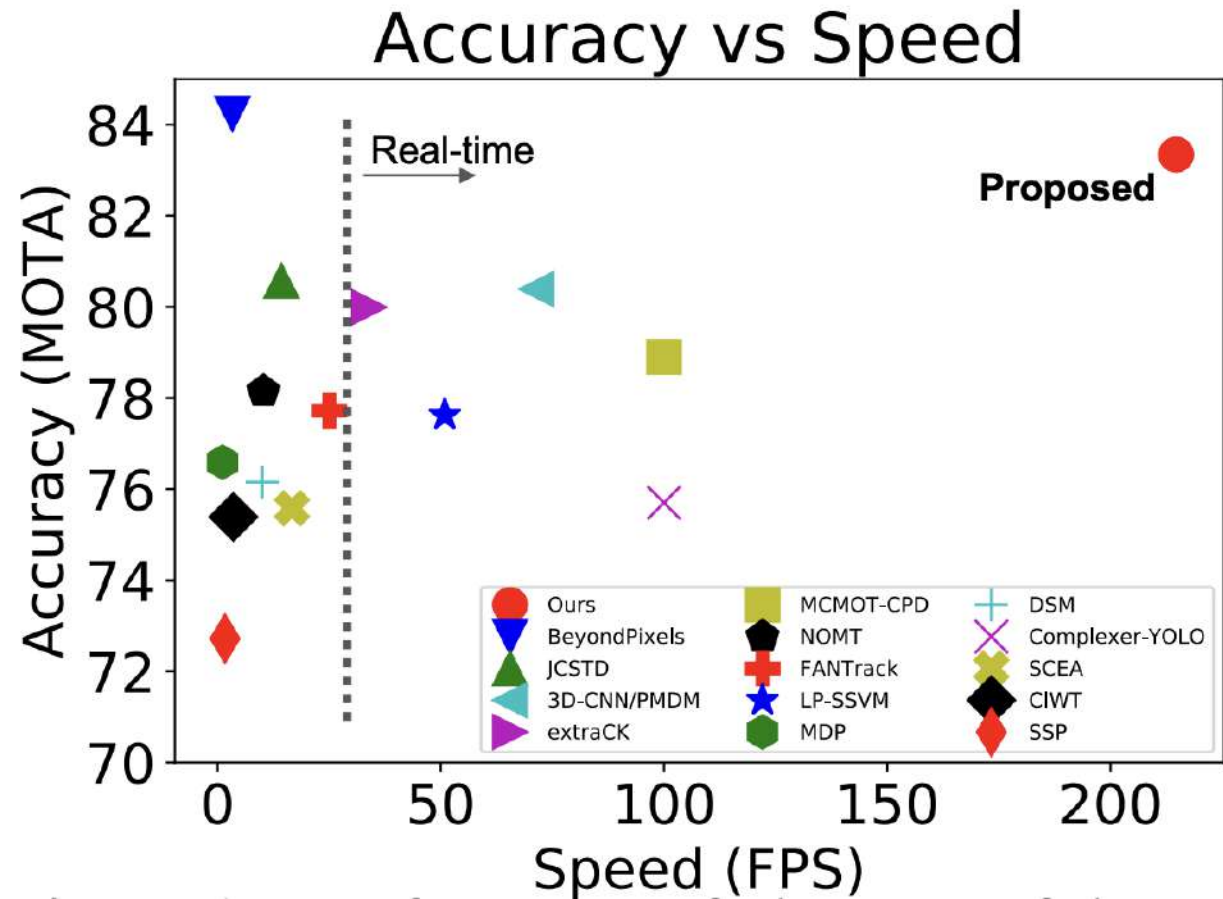
MOTA over Recall curve

Our Contributions

1. A 3D MOT evaluation tool along with three integral metrics
2. A strong and simple 3D MOT system with the fastest speed (207.4 FPS)

AB3DMOT: A Baseline for 3D Multi-Object Tracking

- Motivation
 - Reduce system complexity of 3D MOT methods
 - Increase the computational efficiency (i.e., run time speed)
- Simple design: 3D Kalman filter + Hungarian algorithm
 - 3D Kalman filter
 - Extension of standard 2D Kalman filter
 - Add object's 3D property into the state space
- High speed:
 - 207.4 FPS on the KITTI dataset for Cars
 - 470.1 FPS on the KITTI dataset for Pedestrians
 - 1241.6 FPS on the KITTI dataset for Cyclists
- Strong 3D MOT performance competitive to more complicated systems



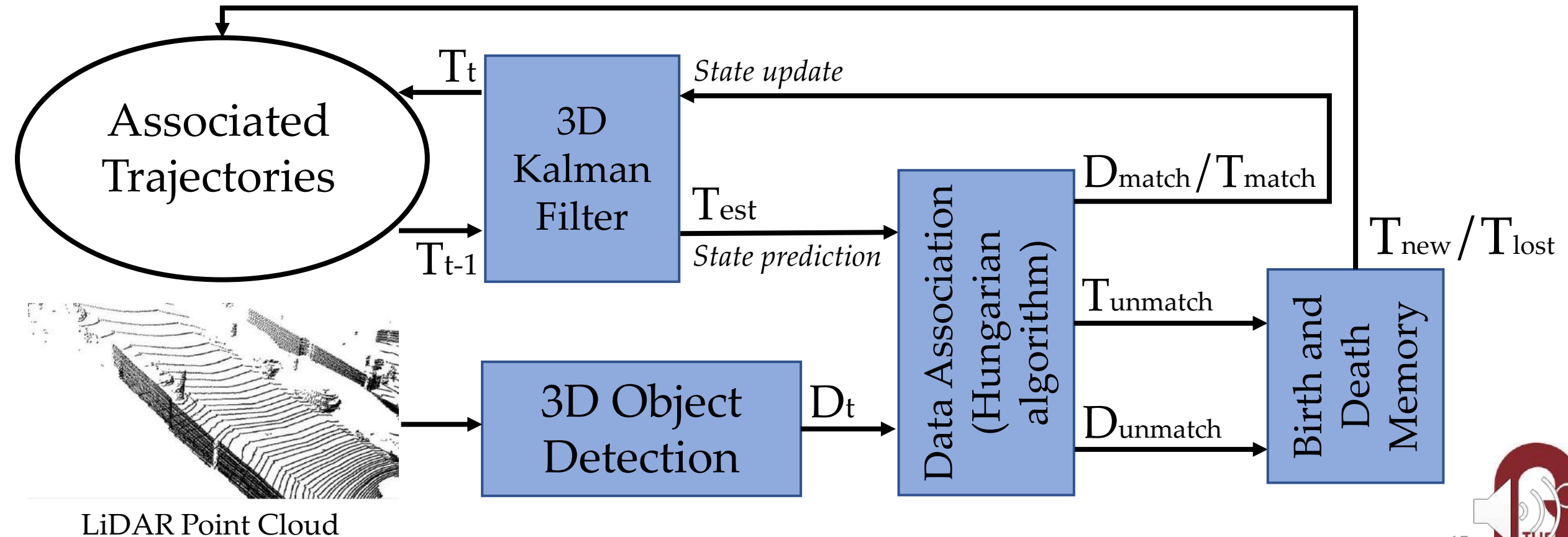
KITTI MOT leaderboard by end of 2019

AB3DMOT: A Baseline for 3D Multi-Object Tracking

- System pipeline (5 modules)
 - 3D object detection
 - Hungarian algorithm
 - Birth and death memory

3D Kalman filter: state prediction

3D Kalman filter: state update



Quantitative Results

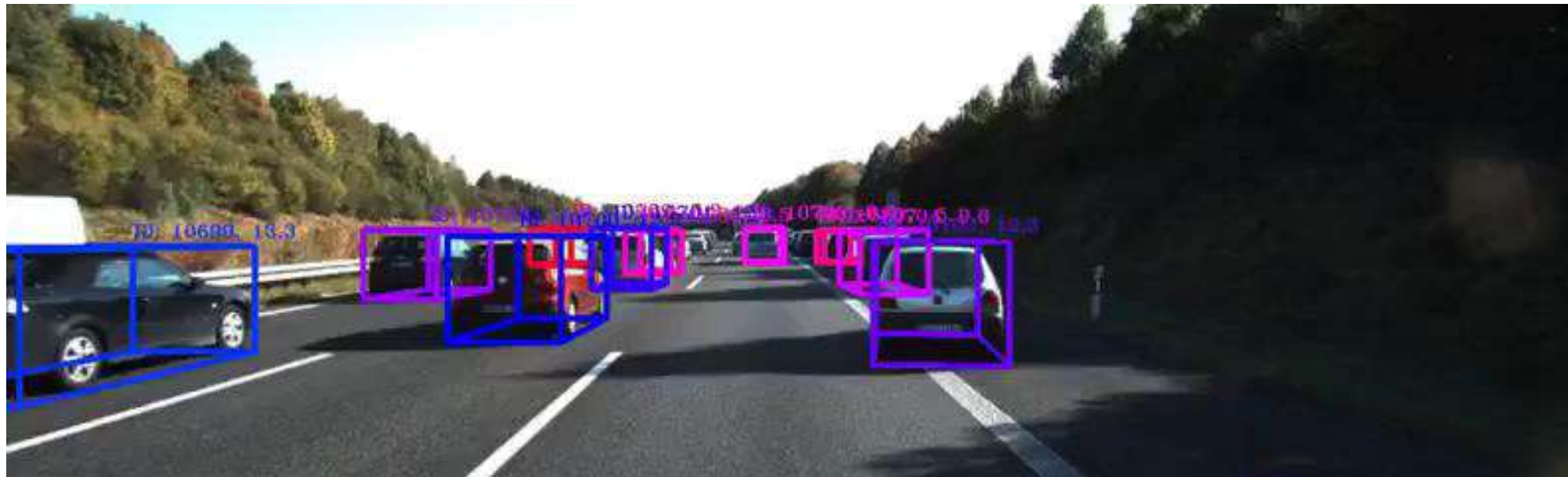
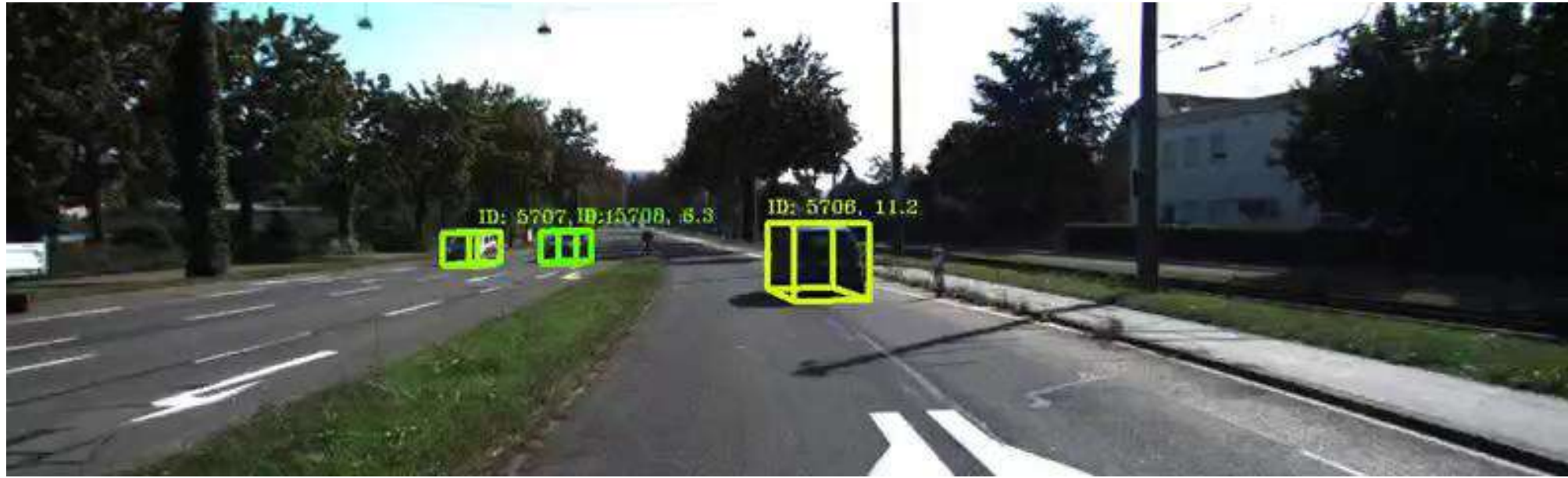
3D MOT Evaluation on KITTI for Cars

- Our 3D MOT system runs at the fastest speed without the need of a GPU
- Our simple system outperforms two more complicated 3D MOT systems

Method	Input Data	Matching criteria	sAMOTA \uparrow	AMOTA \uparrow	AMOTP \uparrow	MOTA \uparrow	MOTP \uparrow	IDS \downarrow	FRAG \downarrow	FPS \uparrow
mmMOT [32] (ICCV'19)	2D + 3D	IoU _{thres} = 0.25	70.61	33.08	72.45	74.07	78.16	10	55	4.8 (GPU)
FANTrack [17] (IV'20)	2D + 3D	IoU _{thres} = 0.25	82.97	40.03	75.01	74.30	75.24	35	202	25.0 (GPU)
Ours	3D	IoU _{thres} = 0.25	93.28	45.43	77.41	86.24	78.43	0	15	207.4 (CPU)

Qualitative Results

Qualitative Results for Cars



AB3DMOT: A Baseline for 3D Multi-Object Tracking and New Evaluation Metrics

Xinshuo Weng, Jianren Wang, David Held, Kris Kitani
Robotics Institute, Carnegie Mellon University

European Conference on Computer Vision (ECCV) Workshops, 2020

