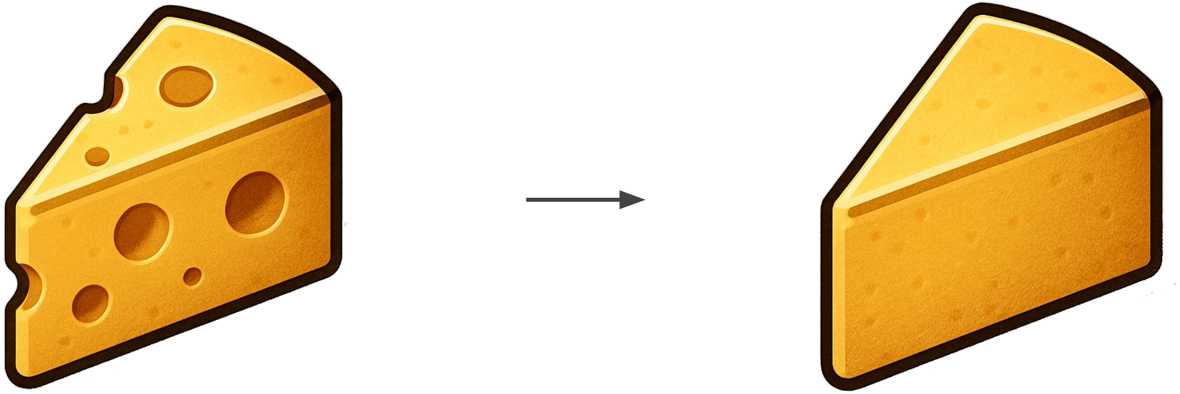
DOC-Depth:
A novel approach for dense depth ground truth generation

Simon de Moreau, Mathias Corsia, Hassan Bouchiba
Yasser Almehio, Andrei Bursuc, Hafid El-Idrissi, Fabien Moutarde

Motivation



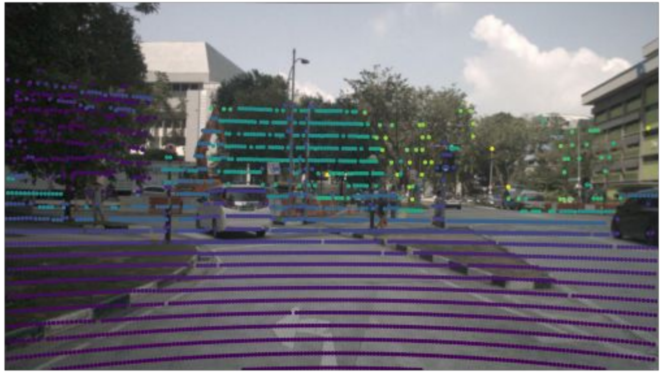
Fewer holes = More cheese!

Motivation

Depth information is essential to many robotics applications.

Available outdoor **depth ground truth is sparse**.

Waymo Open Dataset [1]



nuScenes dataset [2]

Can we complete the picture?

[1] Sun, P., et al, "Scalability in Perception for Autonomous Driving: Waymo Open Dataset," in CVPR, 2020.
[2] Holger Caesar, et al, "nuScenes: A multimodal dataset for autonomous driving," in CVPR, 2020.

Motivation

KITTI dataset proposed to **aggregate 11 frames** and **validate using stereo-vision** method.

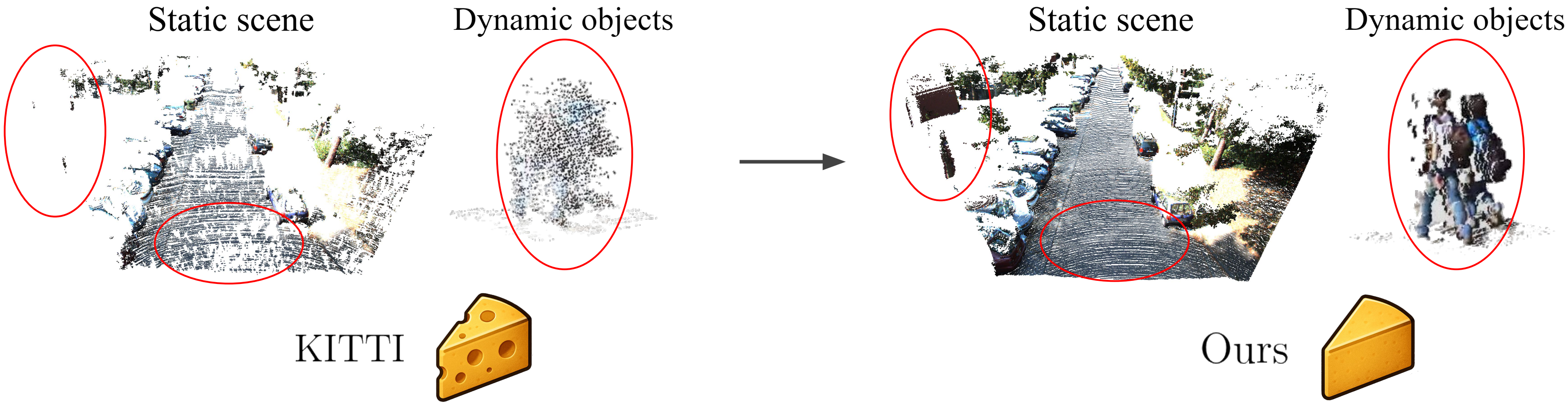
KITTI dataset [3]

But it **leads to inaccuracies** in the **ground-truth** that can **impair the performances** of AI models.

[3] Jonas Uhrig, et al, "Sparsity Invariant CNNs," in 3DV, 2017.

Motivation

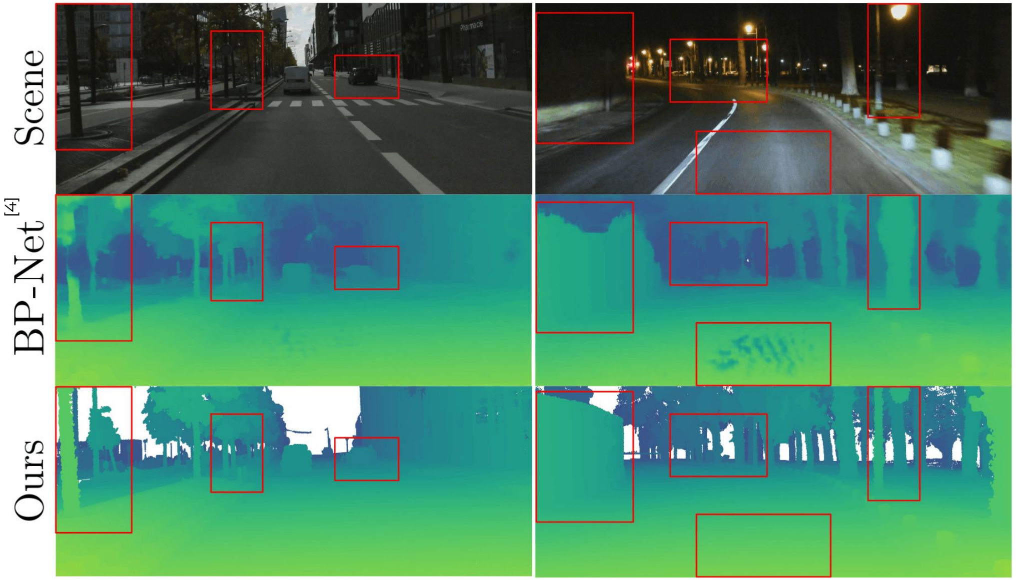
We can do better!

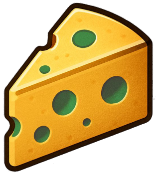


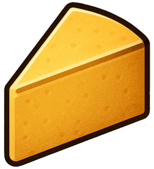
Our method produces **high quality** output **using the <u>same</u> LiDAR data.**

Motivation

Learning based method using **camera and LiDAR** for depth completion are very powerful.





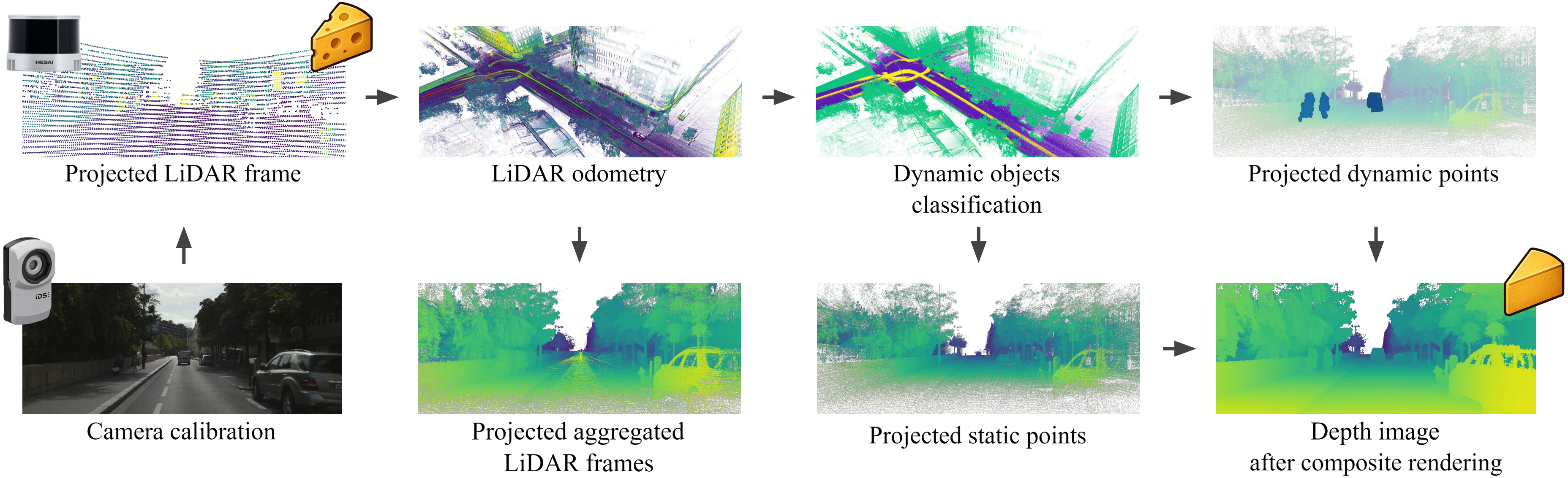


But they **struggle outside** the **training domain**.

[4] Tang, J., et al. "Bilateral Propagation Network for Depth Completion," in CVPR, 2024.

Method

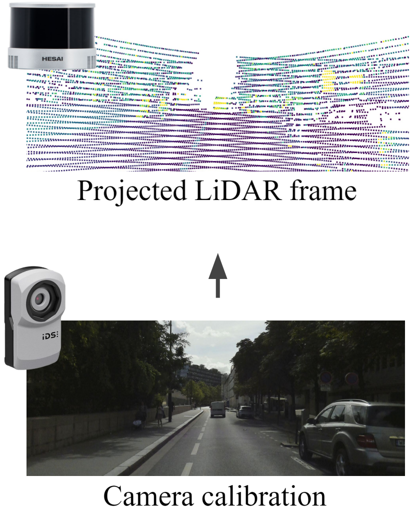
Overview

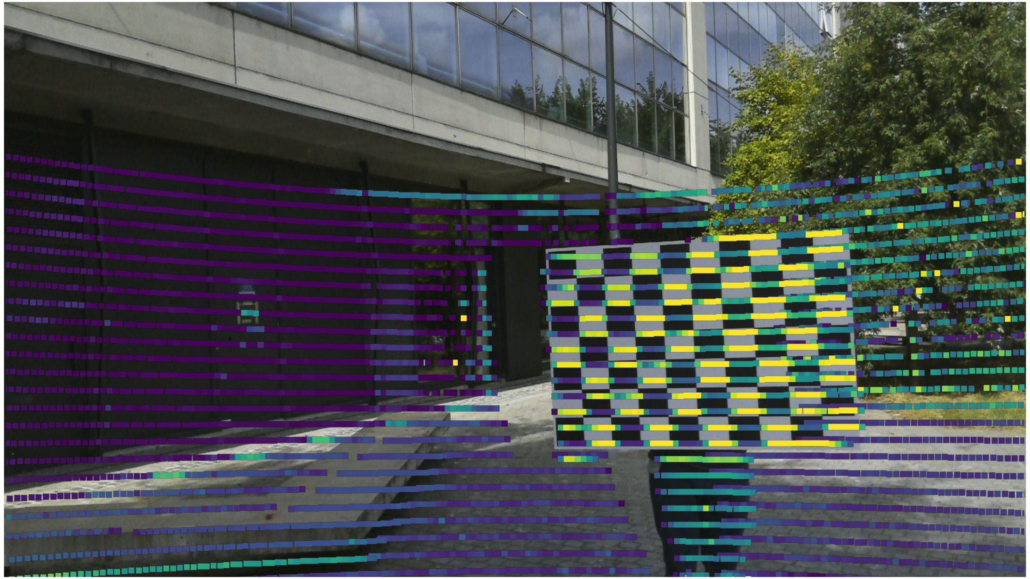


We propose a **learning-free** method **based only on LiDAR** measurements to generate **high quality dense depth ground truth**.

Method

Sensor Fusion



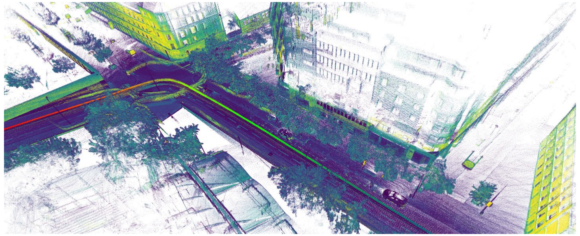


We **project the LiDAR points into the camera** point of view using the **intrinsic and extrinsic calibration**.

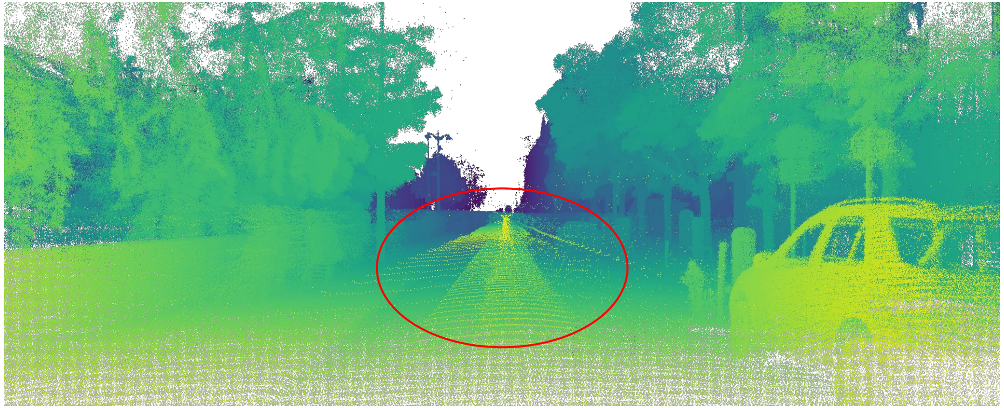
Method

3D densification

We **leverage the LiDAR odometry** to obtain a **dense reconstruction** of the scene.



LiDAR odometry

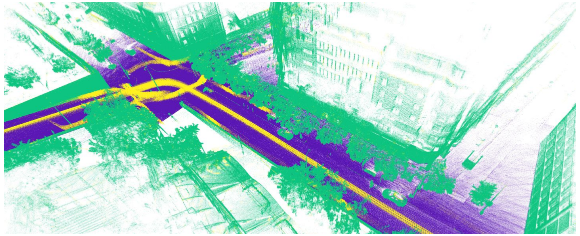


Projected aggregated
LiDAR frames

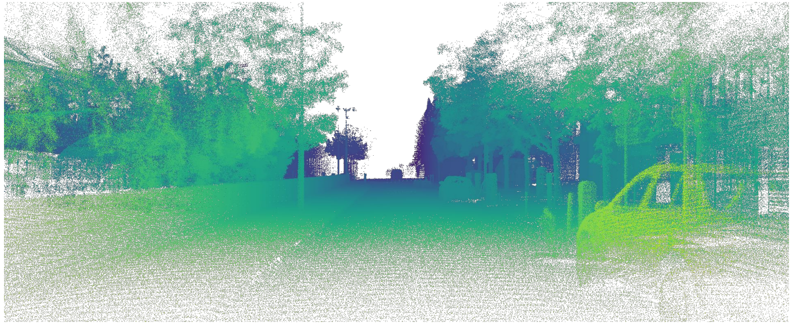
Dynamic objects leave **unwanted geometries producing occlusions** in the depth output.

Method

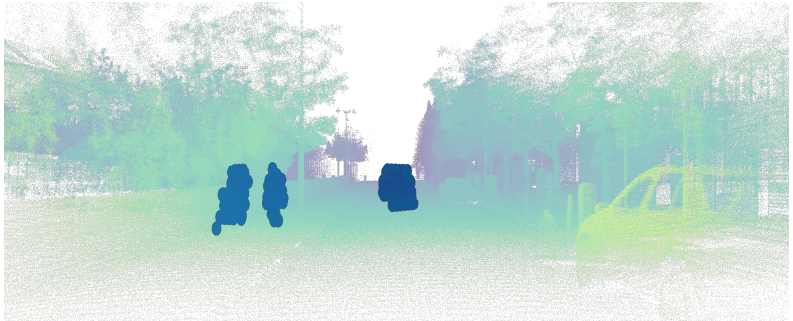
Dynamic object classification



Dynamic objects classification



Projected static points

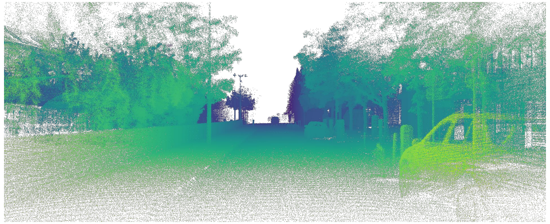


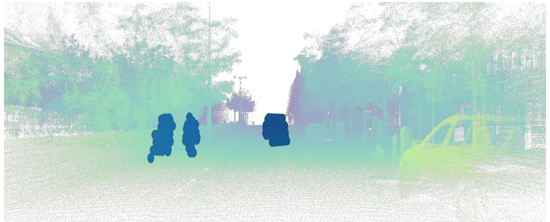
Projected dynamic points

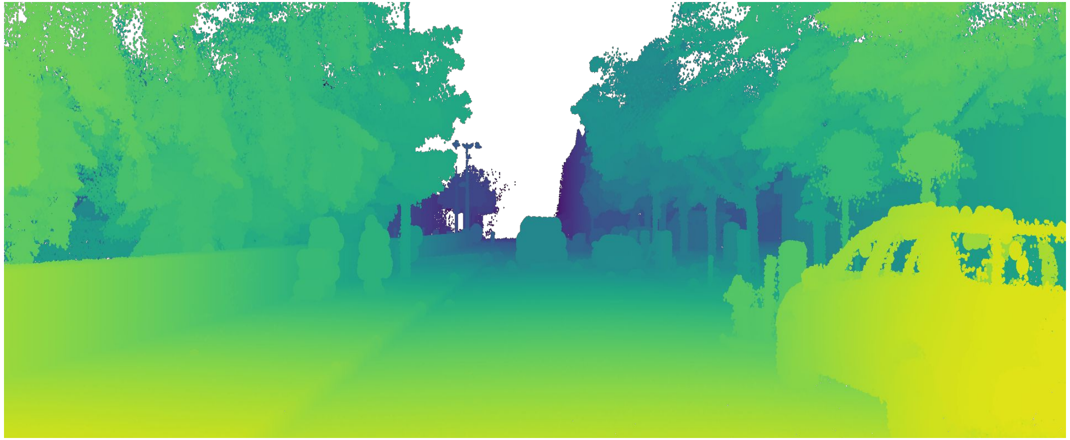
DOC classifies static and dynamic objects, allowing for dedicated rendering.

Method

Composite rendering


Projected static points

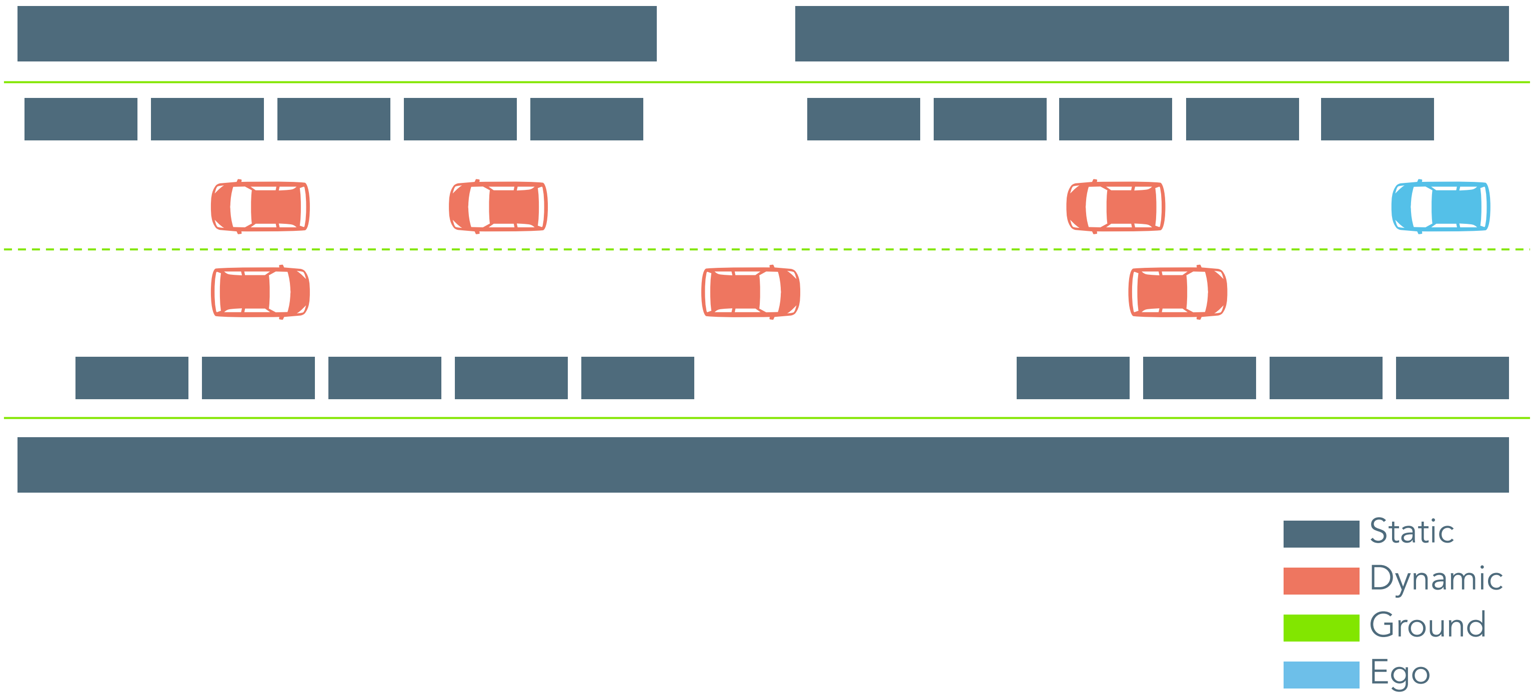

Projected dynamic points


Depth image
after composite rendering

Static structures are sharply reconstructed and **dynamic points are precisely positioned** by using only the current frame. **Point size is adjusted to close gaps**.

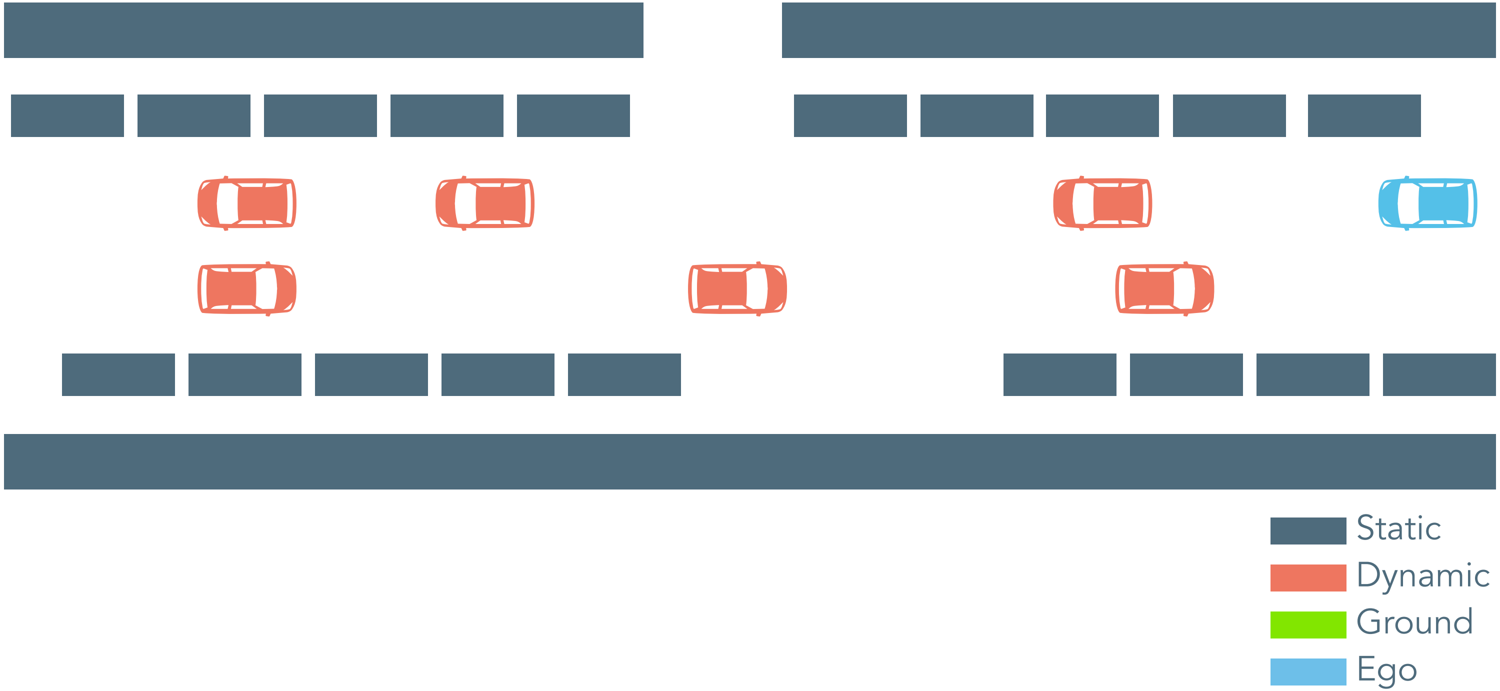
DOC - Dynamic Object Classification

Ground segmentation



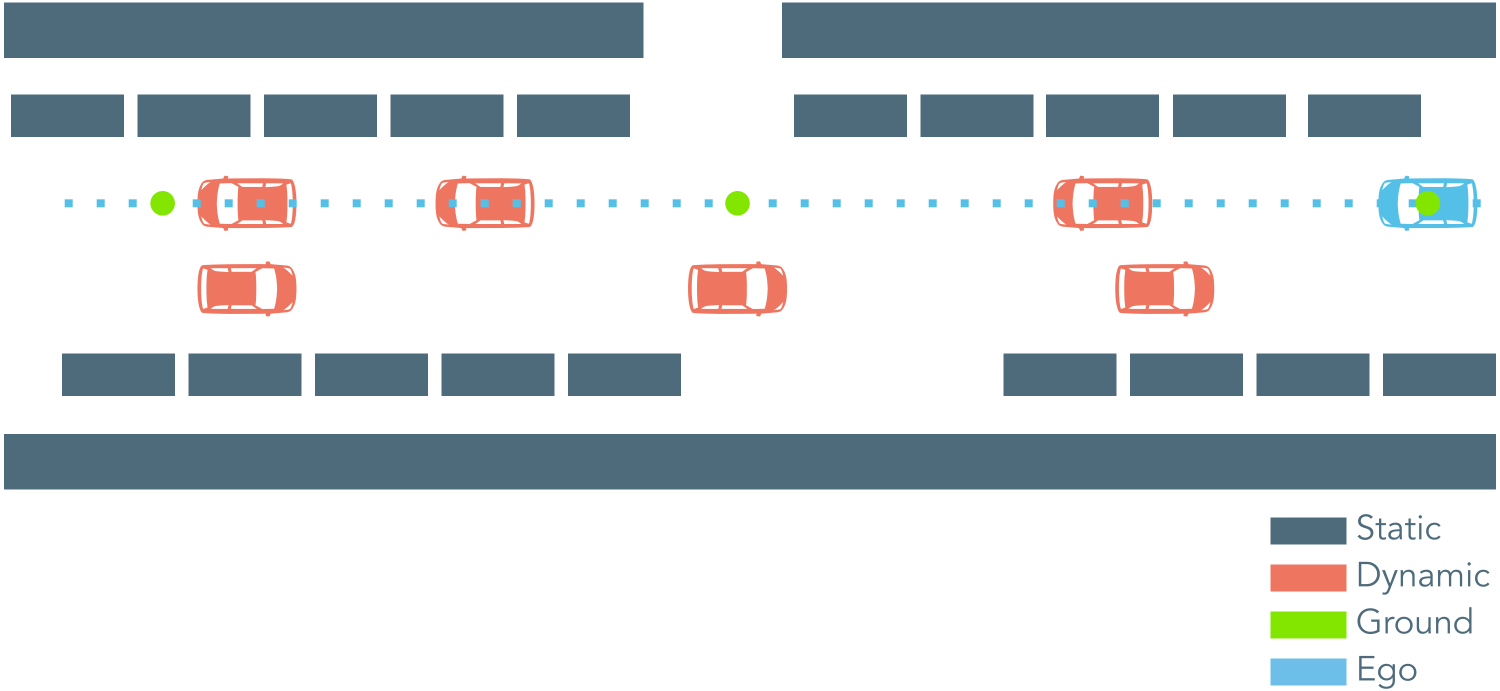
DOC - Dynamic Object Classification

Voting method



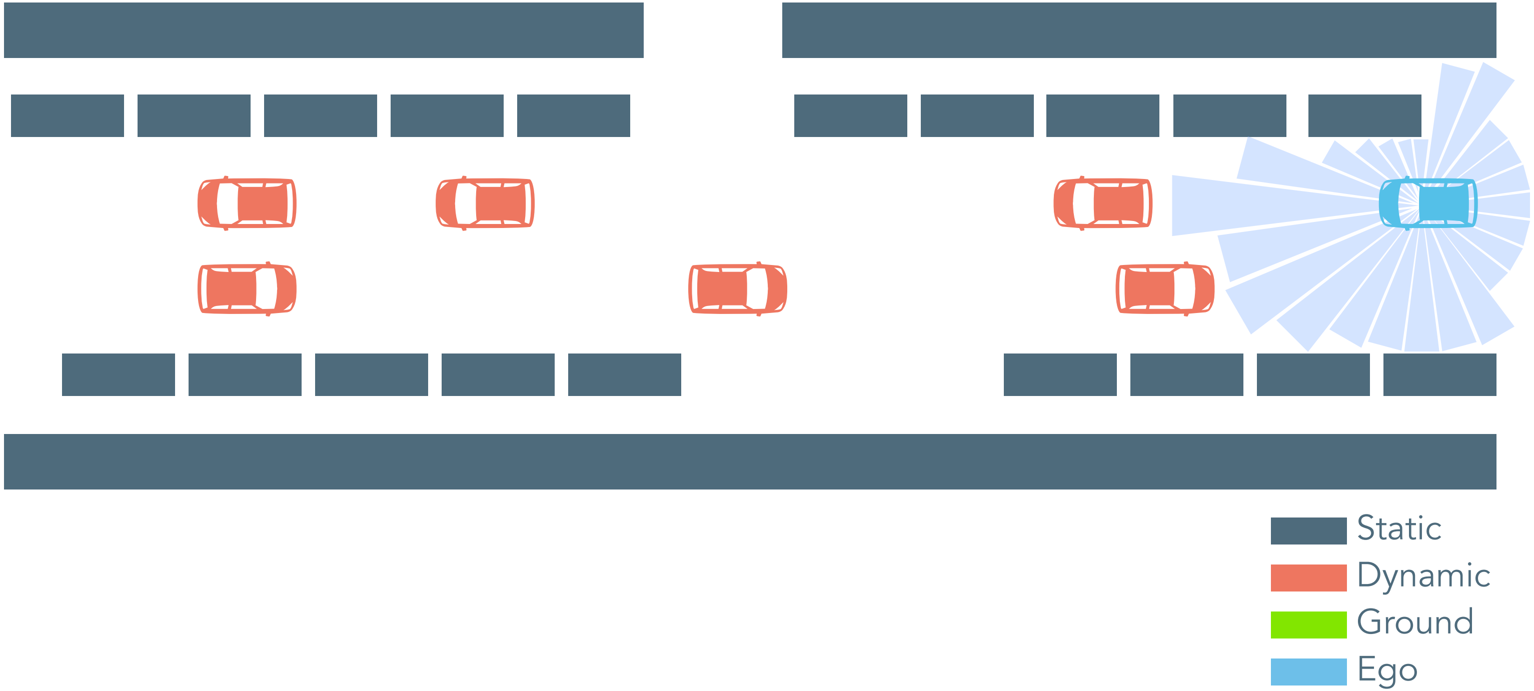
DOC - Dynamic Object Classification

Voting method



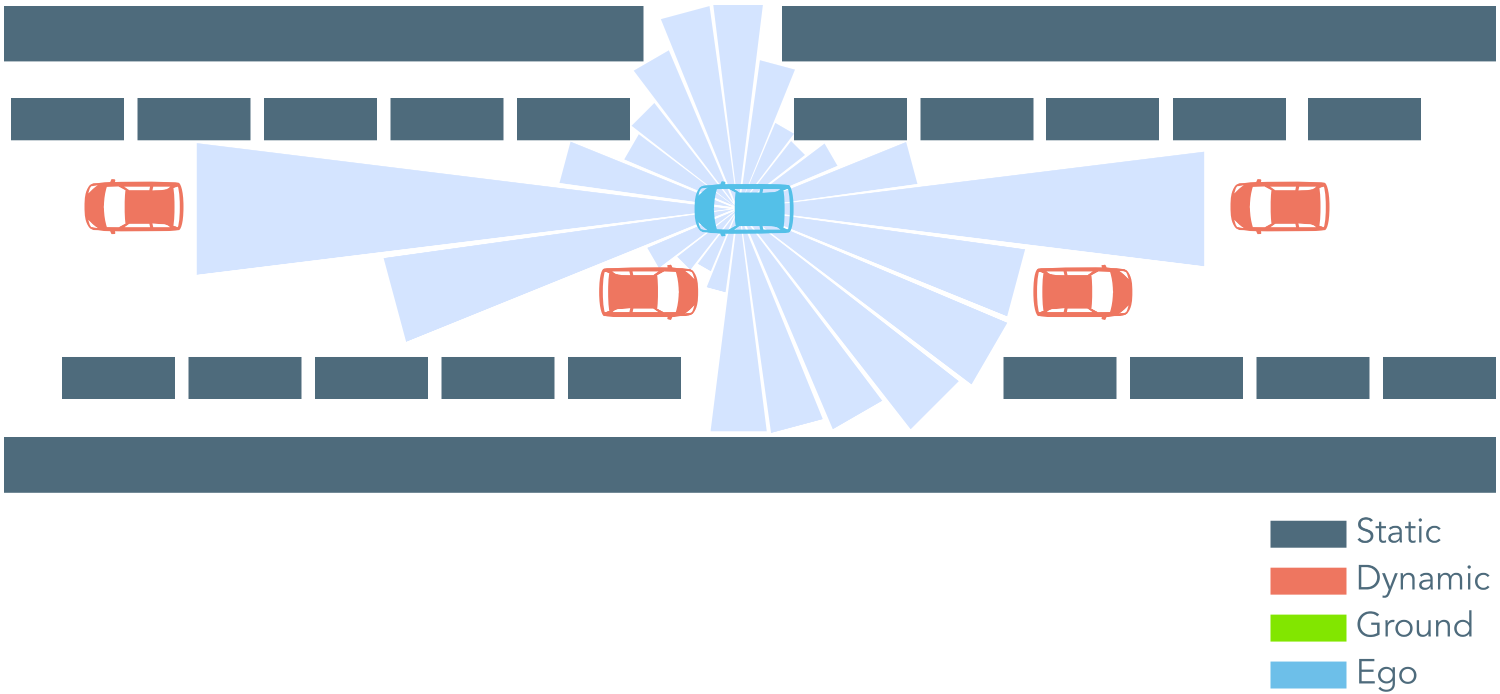
DOC - Dynamic Object Classification

Voting method



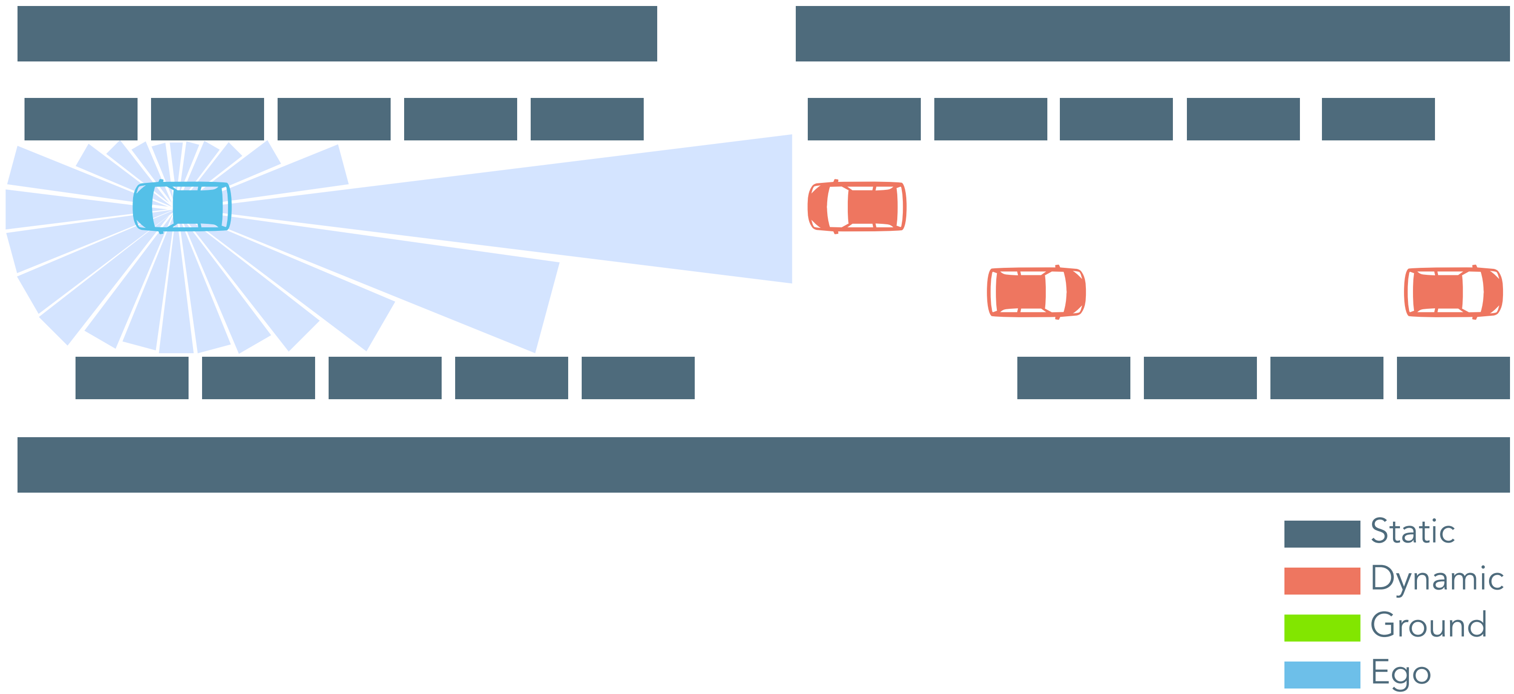
DOC - Dynamic Object Classification

Voting method



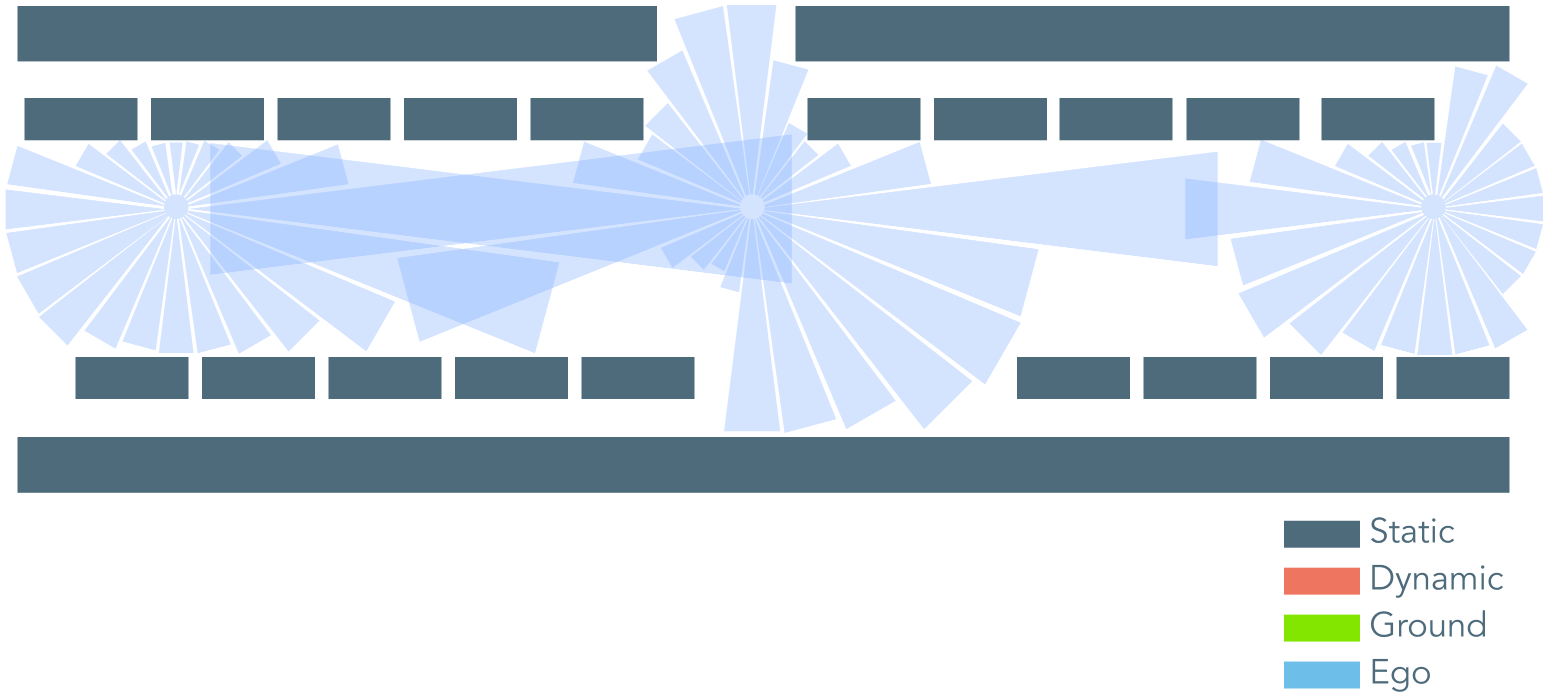
DOC - Dynamic Object Classification

Voting method



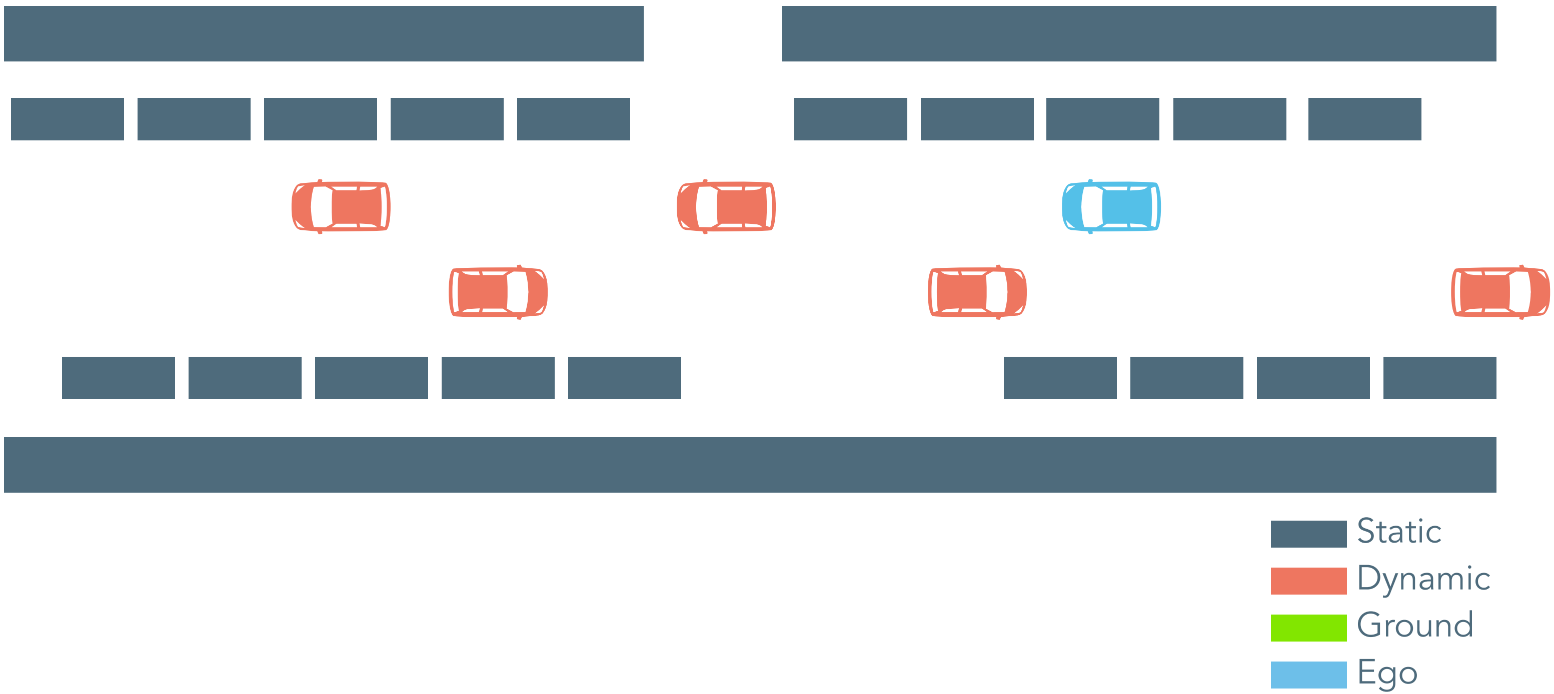
DOC - Dynamic Object Classification

Voting method



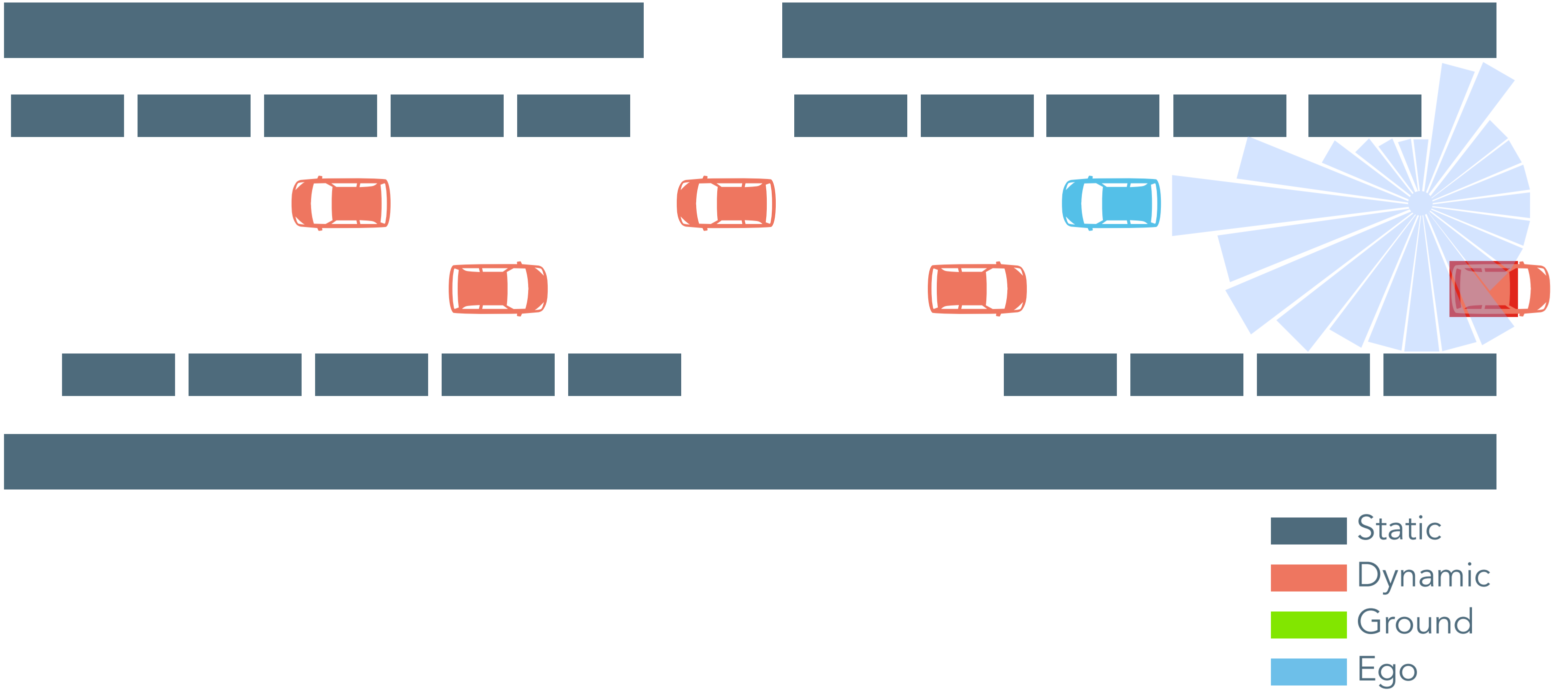
DOC - Dynamic Object Classification

Voting method



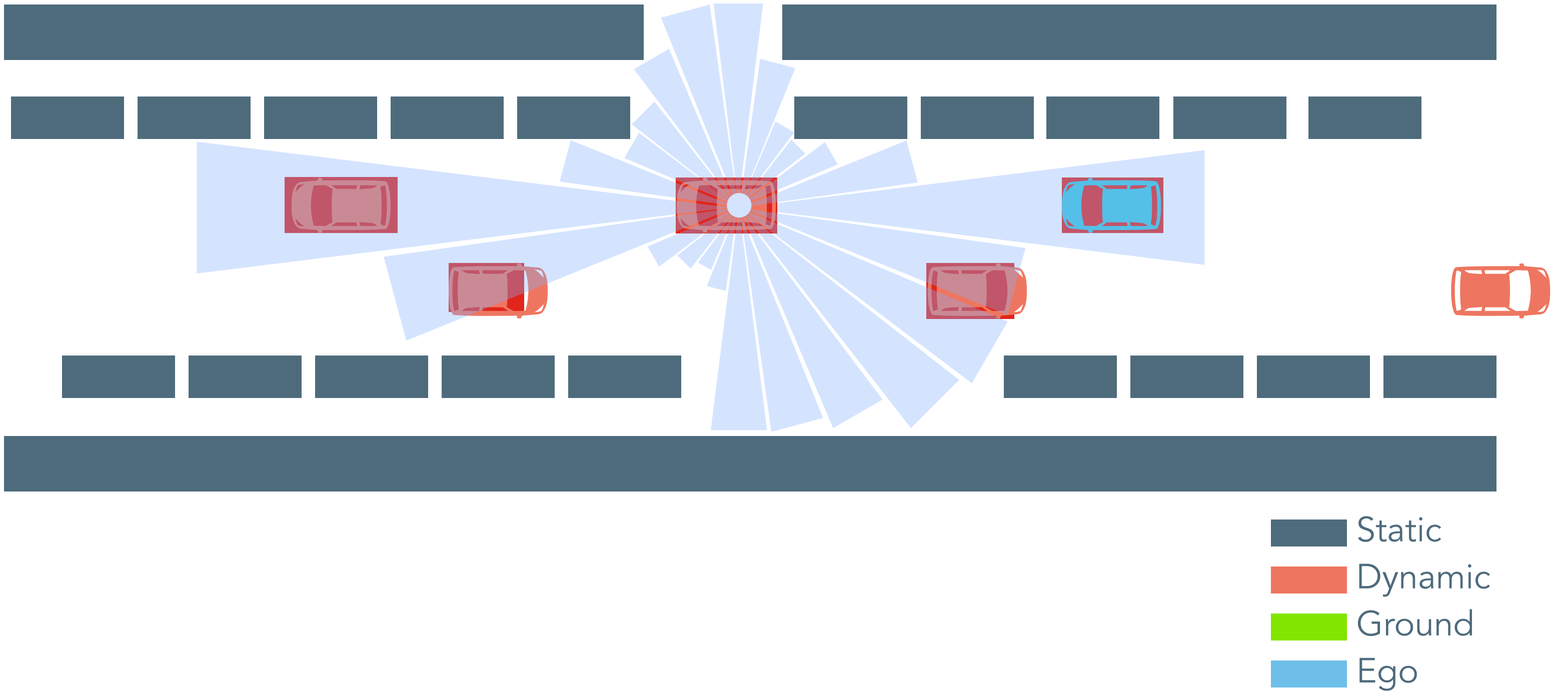
DOC - Dynamic Object Classification

Voting method



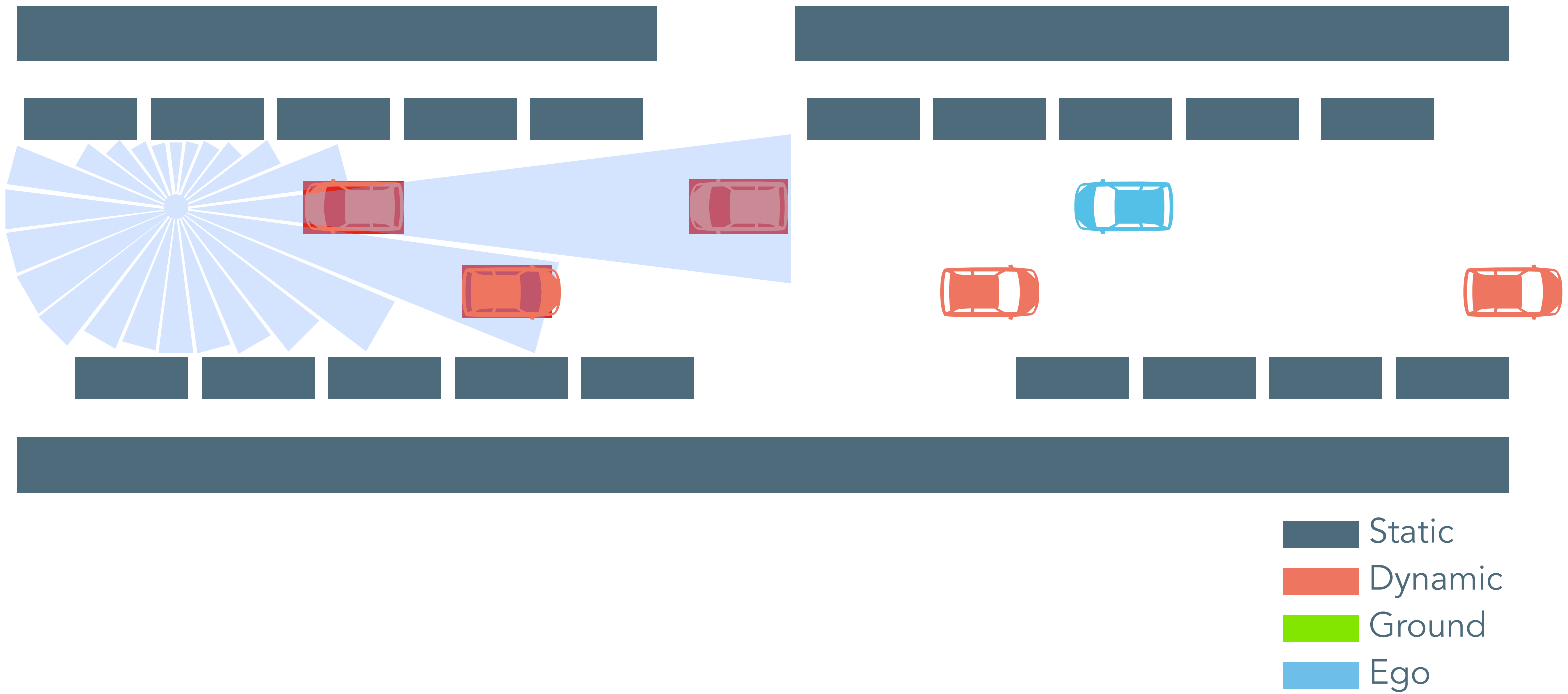
DOC - Dynamic Object Classification

Voting method



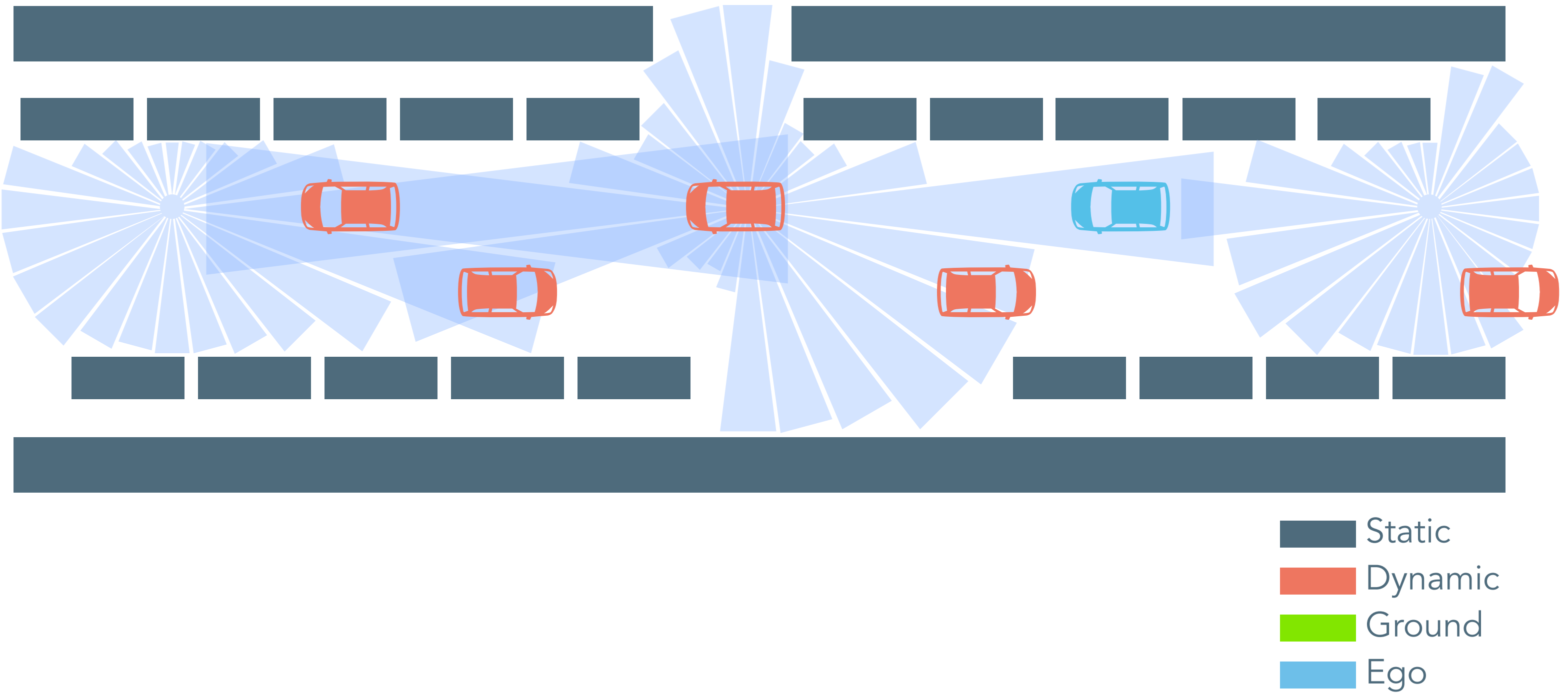
DOC - Dynamic Object Classification

Voting method



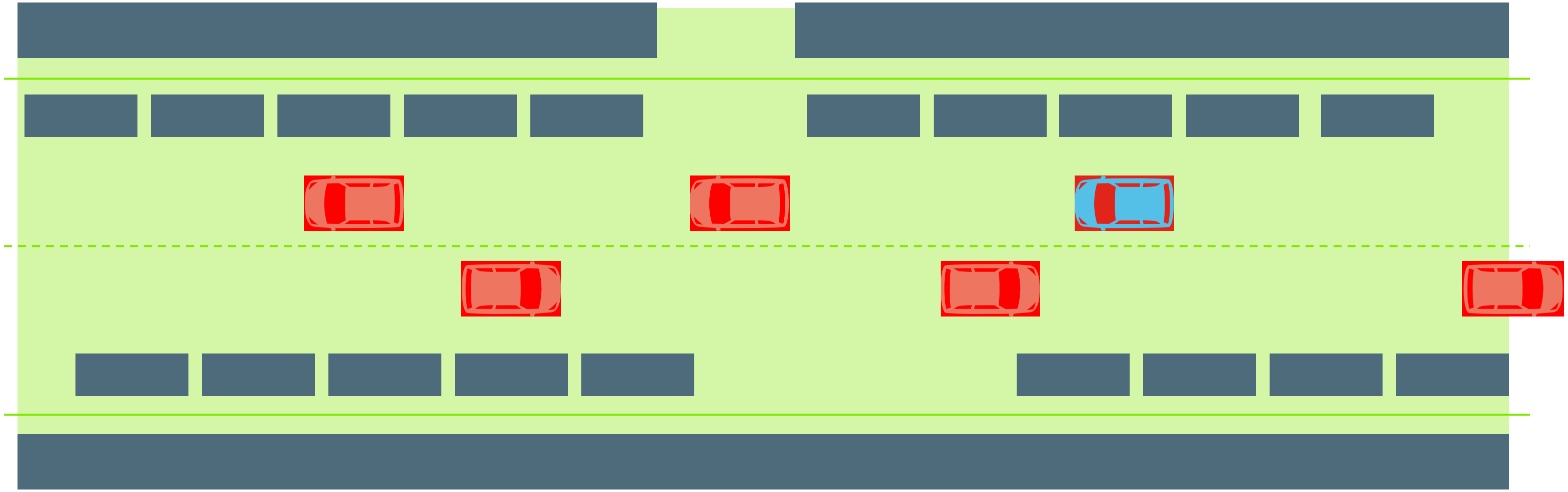
DOC - Dynamic Object Classification

Voting method



DOC - Dynamic Object Classification

Voting method



Static
Dynamic
Ground
Ego

Results on SemanticKITTI [5]

Dynamic objects classification

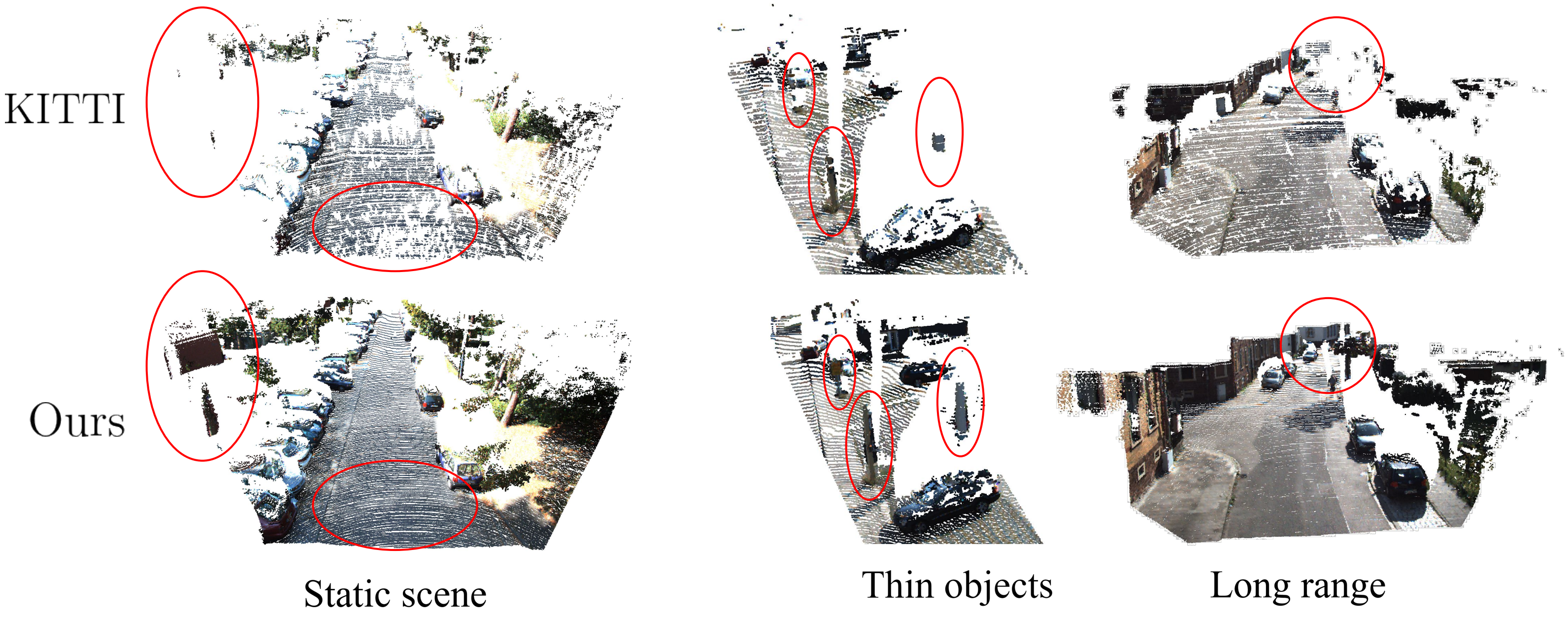
Sequence	Method	SA (%) ↑	DA (%) ↑	F1-score ↑
00	ERASOR [6]	66.70	98.54	0.7955
	MapCleaner [7]	98.89	98.18	0.9853
	Dynablox [8]	96.76	90.68	0.9362
	BeautyMap [9]	96.76	98.38	0.9756
	DOC (Ours)	**99.73**	**98.99**	**0.9935**
01	ERASOR	98.12	90.94	0.9439
	MapCleaner	**99.74**	94.98	0.9730
	Dynablox	96.33	68.01	0.7373
	BeautyMap	99.17	92.99	0.9598
	DOC (Ours)	99.66	**96.89**	**0.9825**
05	ERASOR	69.40	99.06	0.8162
	MapCleaner	99.14	97.92	0.9852
	Dynablox	97.80	88.68	0.9302
	BeautyMap	96.34	98.29	0.9731
	DOC (Ours)	**99.69**	**99.07**	**0.9937**
Average	ERASOR	78.07	96.18	0.8618
	MapCleaner	99.25	97.02	0.9812
	Dynablox	96.96	82.46	0.8912
	BeautyMap	97.42	96.55	0.9698
	DOC (Ours)	**99.69**	**98.31**	**0.9899**

DOC outperforms previous methods of dynamic object classification on SemanticKITTI.

[5] J. Behley, et al, "SemanticKITTI: A Dataset for Semantic Scene Understanding of LiDAR Sequences," in ICCV, 2019.
[6] H. Lim, S. Hwang, H. Myung. "ERASOR: Egocentric Ratio of Pseudo Occupancy-Based Dynamic Object Removal for Static 3D Point Cloud Map Building," in IEEE RA-L, 2021.
[7] H. Fu, H. Xue, G. Xie. "MapCleaner: Efficiently Removing Moving Objects from Point Cloud Maps in Autonomous Driving Scenarios," in Remote Sensing, 2022.
[8] Schmid, et al. "Dynablox: Real-time Detection of Diverse Dynamic Objects in Complex Environments," in IEEE RA-L, 2023.
[9] Jia, M., et al. "BeautyMap: Binary-Encoded Adaptable Ground Matrix for Dynamic Points Removal in Global Maps," in IEEE RA-L, 2024.

Results on KITTI

Whole scenes



Our ground truth is **highly denser** (71% of pixels vs 16.1% for KITTI) and **all structures** of the scene are **preserved**. It ensures **better AI model performance** while being **more data-efficient**.

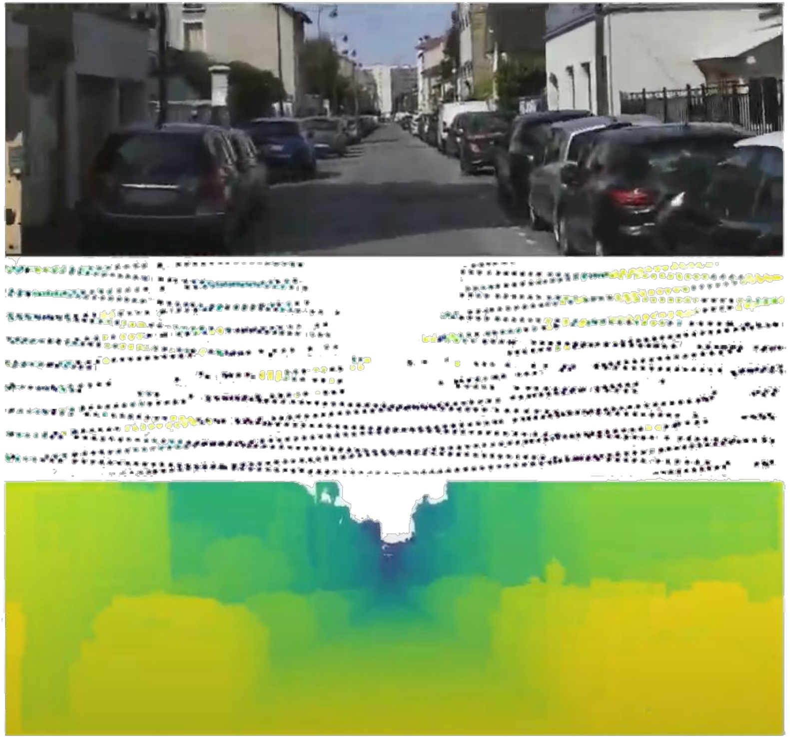
Results on KITTI

Dynamic objects

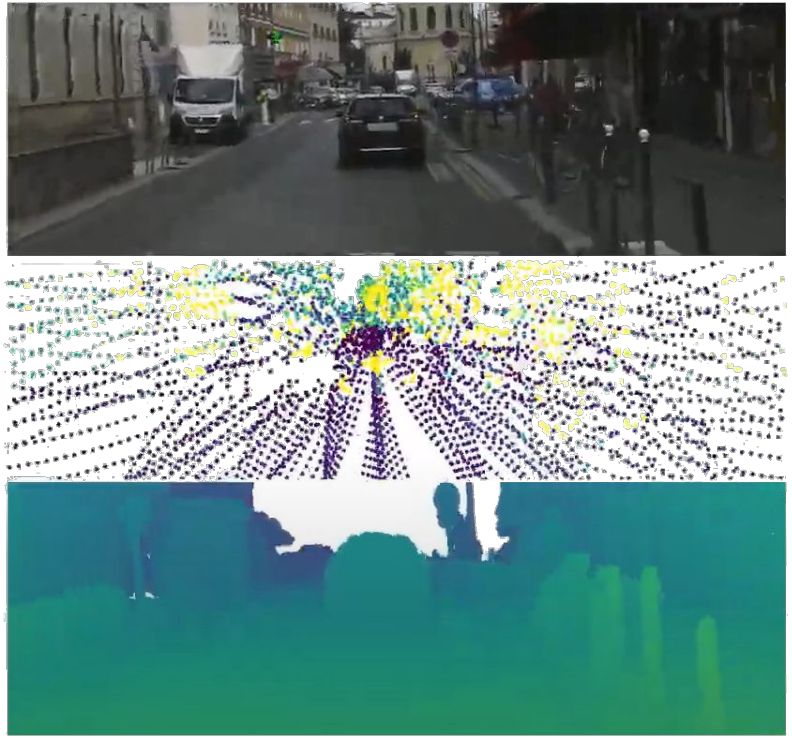
DOC-Depth ensures **accurate dynamic object reconstruction.**

Results on novel datasets

The method is **LiDAR agnostic.**



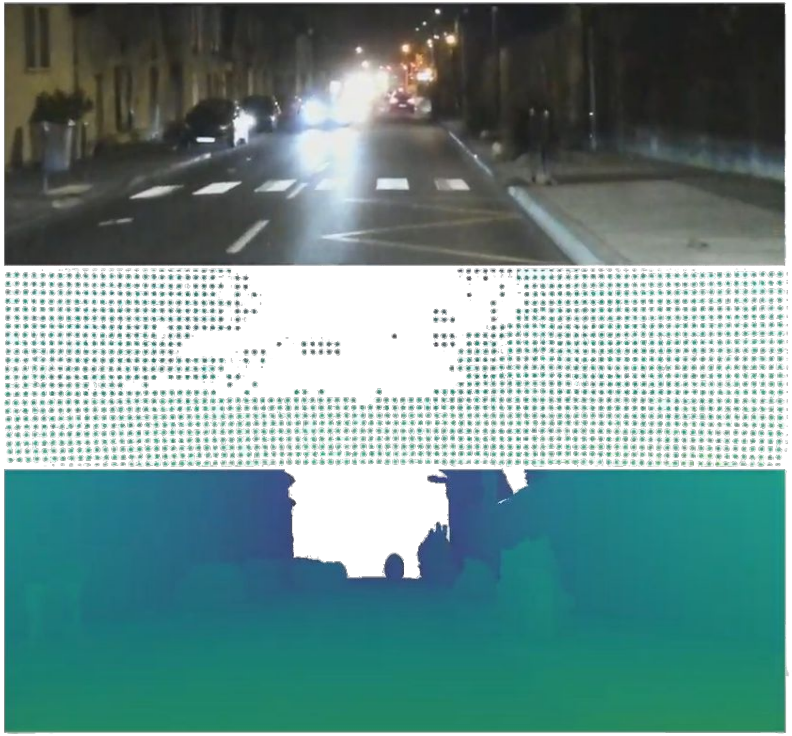
(a) 2 Hesai Pandar-XT32 - High Range



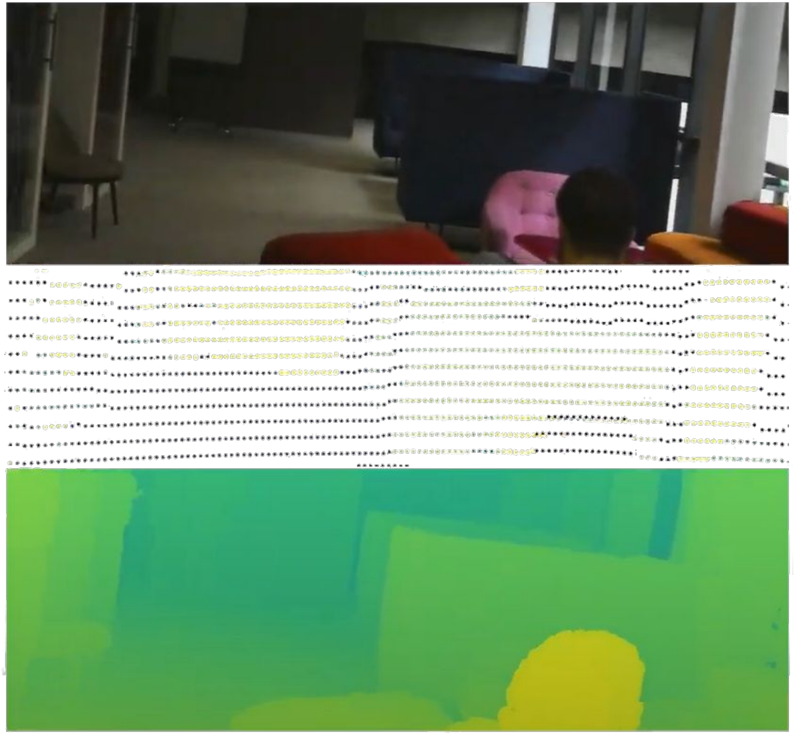
(b) Livox Avia - Specific Pattern

Results on novel datasets

The method **works in diverse environments.**

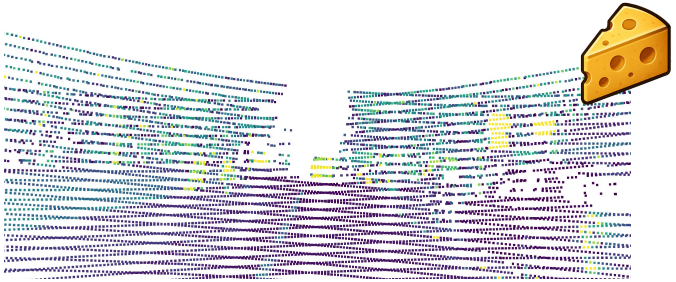


(c) Ouster OS1-128 - Nighttime

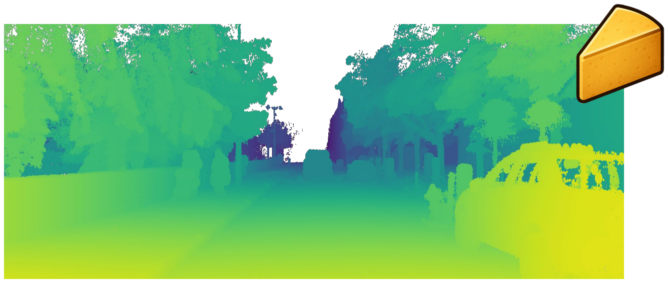


(d) Hesai Pandar-XT32 - Indoor

Contributions



LiDAR frame



DOC-Depth output

- **DOC-Depth** produces **reliable and high quality dense depth** ground truth.
- **DOC** is a **state-of-the art** method of **dynamic object classification** for LiDAR data.
- Our method is **easy-to-deploy, scalable and LiDAR agnostic**.
- **Software components** and **dense depth KITTI ground truth** are available for the research community.

Meet the team

Simon de Moreau

Valeo | MINES PARIS | PSL

Mathias Corsia

EXAWIZ

Hassan Bouchiba

EXAWIZ

Yasser Almehio

Valeo

Andrei Bursuc

valeo.ai

Hafid El-Idrissi

Valeo

Fabien Moutarde

MINES PARIS | PSL

Check out the project:

Project page

Code

Simon de Moreau: simon.de_moreau@minesparis.psl.eu