Sensor-independent Pedestrian Detection for Personal Mobility Vehicles in Walking Space Using Dataset Generated by Simulation (631)

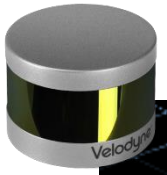
Takahiro Shimizu* , Kenji Koide**, Shuji Oishi**,
Masashi Yokozuka**, Atsuhiko Banno** and Motoki Shino*

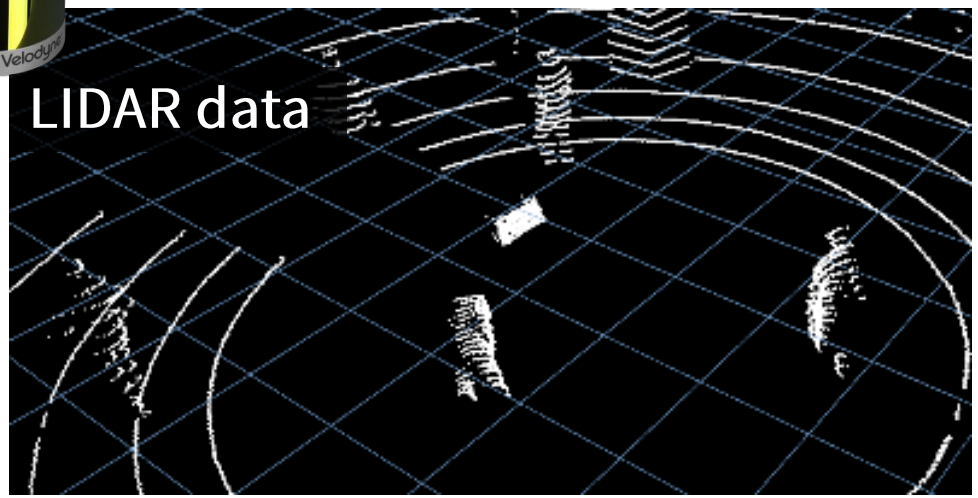
* The University of Tokyo
** National Institute of Advanced Industrial Science and Technology (AIST)

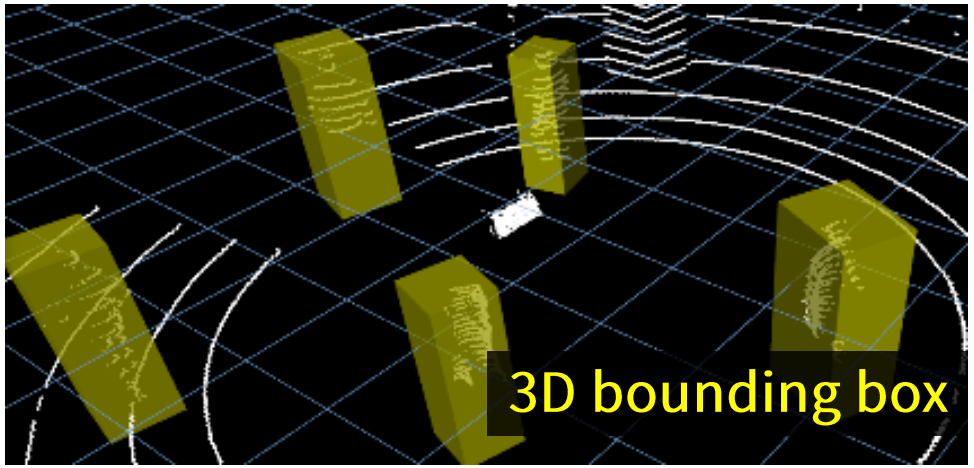
Introduction

◆Self-driving for wheal chairs

- Transportation for the elderly and the physically handicapped
- Accurate pedestrian detection in walking space is indispensable





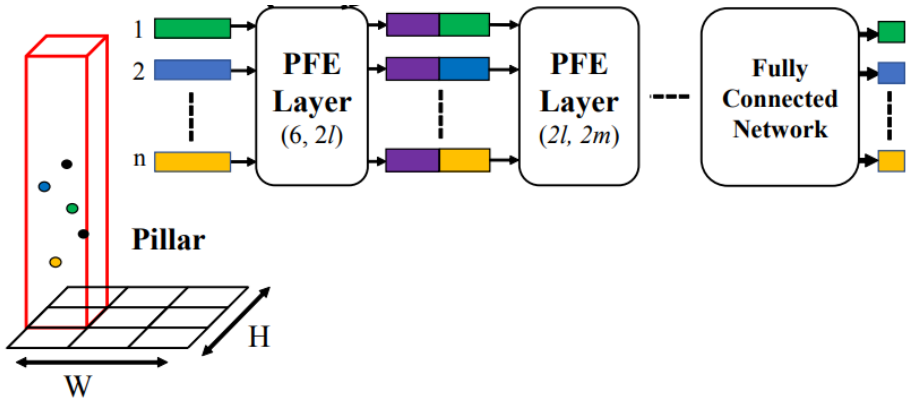


Strategy

◆ Laser intensity-free detection

- Recent methods rely on laser reflectance as additional information.
- These network may be affected by the difference of LIDAR models.

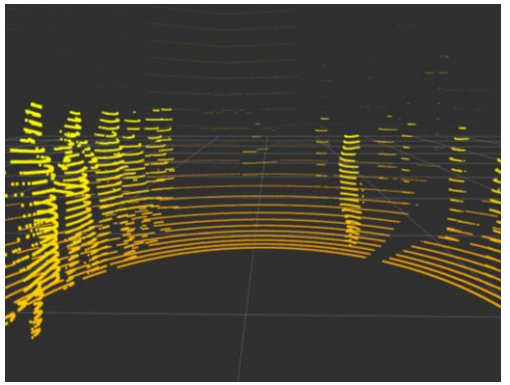
➡ Intensity-free Network



◆ Dataset in walking space

- The majority of datasets focuses on road scenes (e.g. KITTI[1]).
- They are not optimal for object detection in walking spaces.

➡ Walking space dataset for wheel chairs



[1] A. Geiger, P. Lenz, and R. Urtasun, “Are we ready for Autonomous Driving? The KITTI Vision Benchmark Suite.” In CVPR, pp. 3354–3361, 2012.

Proposed Network

◆ CosPointPillars

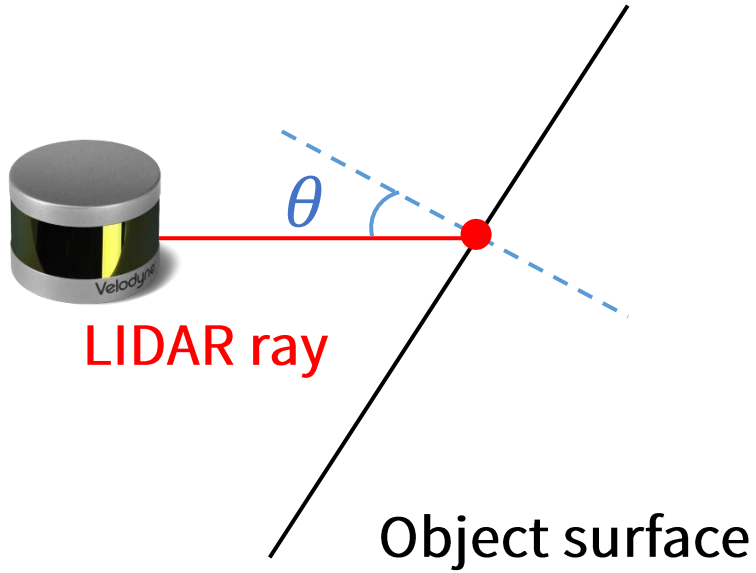
- As an alternative channel for reflection intensity, Cosine Estimation Network is added to PointPillars[4]

◆ Why Cosine?

- Lambertian model

$$Intensity \propto \frac{K_\lambda \cos\theta}{d^2}$$

θ: Incident angle
d: Distance
K_λ : Object reflectivity



Use cosθ instead of the laser reflectivity for a general detection network:

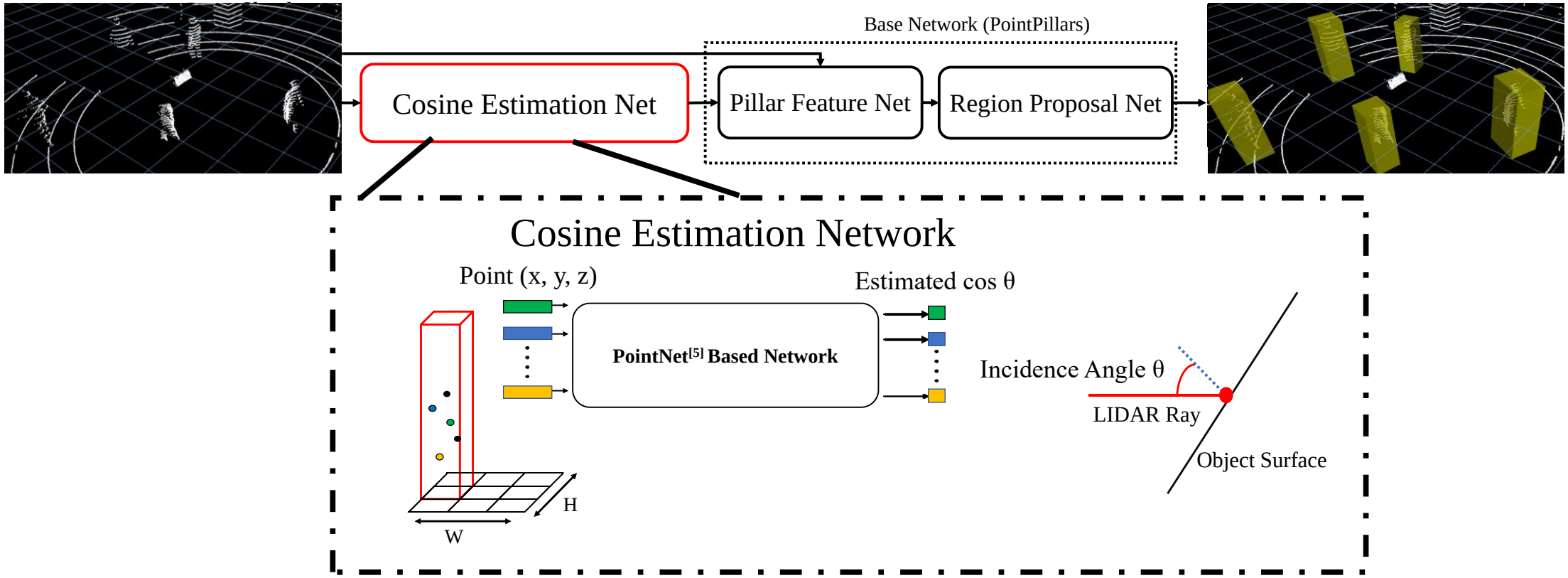
It reflects the local characteristics of the reflection intensity while it can be extracted from the positional relationship with neighboring points

[4]A. H. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang, and O. Beijbom,"Pointpillars: Fast Encoders for Object Detection from Point Clouds." In CVPR, 2019.

Proposed Network

CosPointPillars

- As an alternative channel for reflection intensity, Cosine Estimation Network is added to PointPillars[4]



[4]A. H. Lang, S. Vora, H. Caesar, L. Zhou, J. Yang, and O. Beijbom,"Pointpillars: Fast Encoders for Object Detection from Point Clouds." In CVPR, 2019.
[5]C. R. Qi, H. Su, K. Mo, and L. J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation, CVPR, 2017.

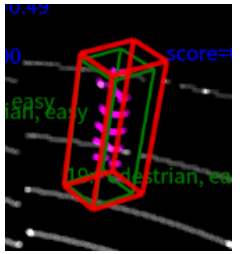
Evaluation

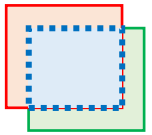
Datasets

- KITTI [Geiger, CVPR2012]
- nuScenes [Caesar, CVPR2020]

Metric

- IoU in 2D Birds' Eye View
- Avg. Precision





Detected BB
Ground Truth
IoU

Evaluation Results

- The accuracy of PointPillars largely deteriorates when the reflectance is not available
- CosPointPillars retains the accuracy by explicitly estimating the local geometrical features

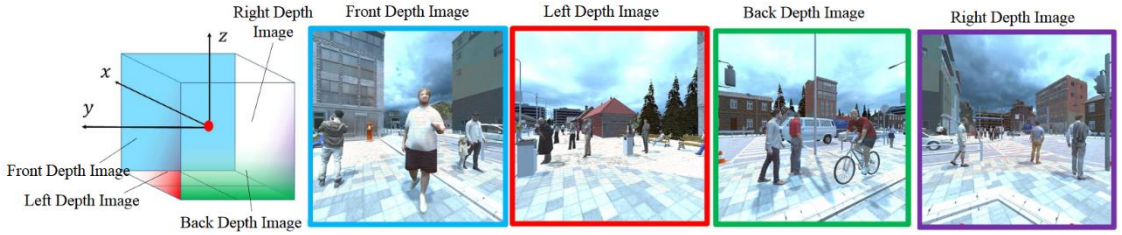
➡ Sensor-independent detection performance

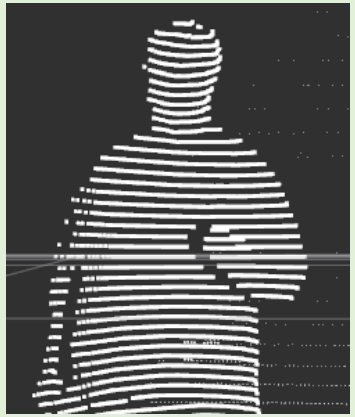
Network	Reflectance	KITTI (IoU 0.5)			NuScenes	
		Easy	Moderate	Hard	IoU 0.5	IoU 0.15
PointPillars	w/	84.36	79.97	76.82	-	-
	w/o	80.22	75.92	74.48	30.23	55.29
CosPointPillars	w/o	82.35	77.29	75.94	32.83	56.94

Pedestrian Detection in a Walking Space

◆ Realistic 3D LiDAR Simulation

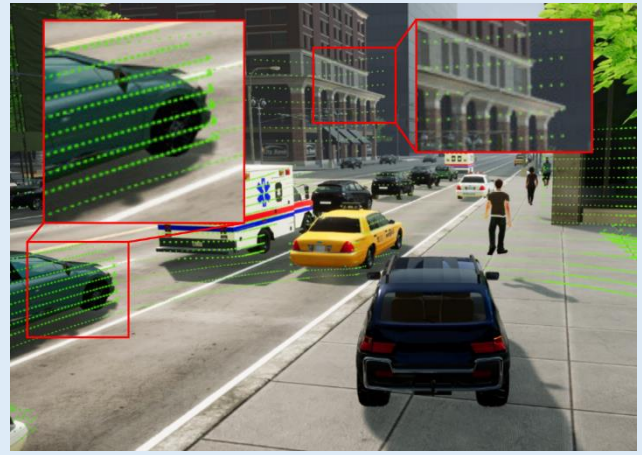
1. Generate Omni-directional depth images
2. Perform ray-casting on the depth images
3. Generate annotation (Labels and BBoxes)





Realistic Point Cloud

Our SimDataset (Ray-casting on depth images)



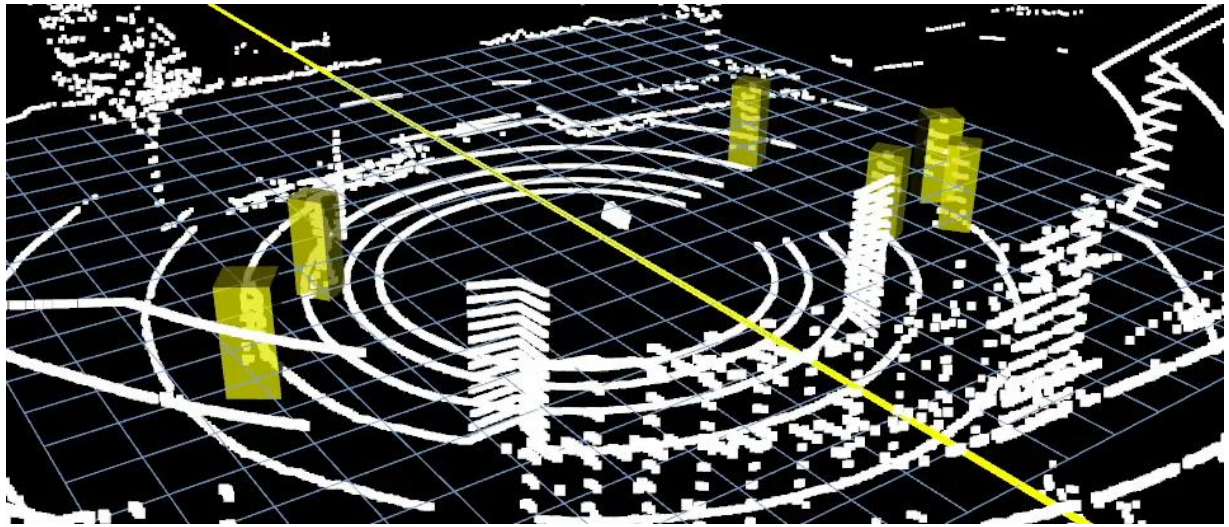
AirSim (Rough collision models)

- We generated a pedestrian detection dataset with over 22k frames and 120k labels

Pedestrian Detection in a Real Environment

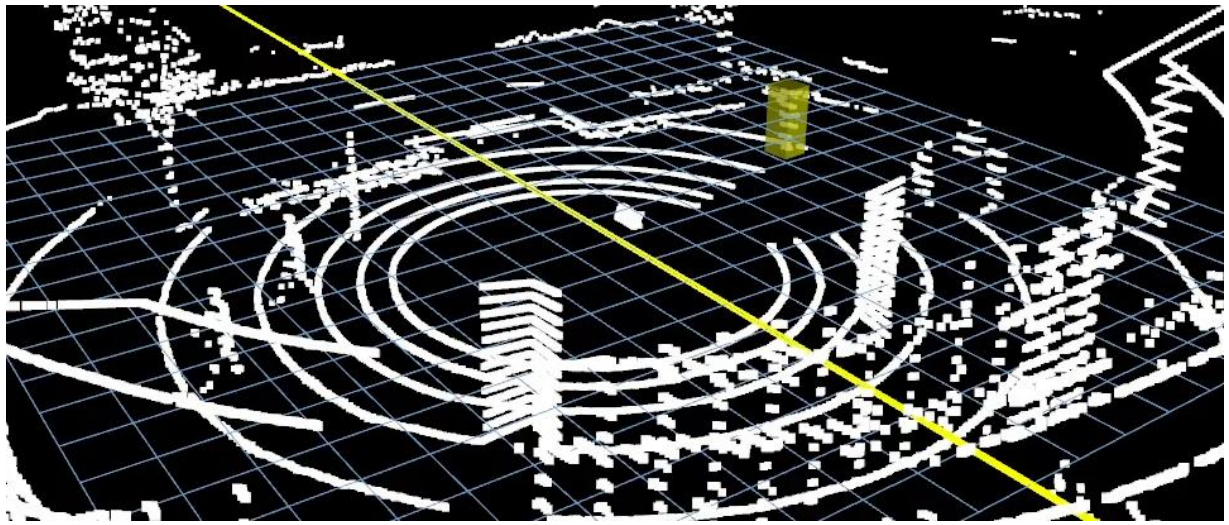
Network trained on KITTI

- Failed to detect nearby pedestrians
- KITTI is taken on a roadway scene and doesn't contain nearby pedestrian data



Network trained on SimDataset

- Nearby pedestrians are robustly detected
- Simulation-based approach enables us to generate a tailor-made dataset for a specific use scenario



Conclusion

- We proposed **CosPointPillars**, a reflectance-intensity-free 3D pedestrian detection network
- CosPointPillars explicitly estimates the cosine local geometric features to compensate for the removed reflectance intensity information
- A large-scale simulation-based pedestrian dataset was created to apply CosPointPillars to a real use scenario
- We succeeded in improving the pedestrian detection accuracy in a real walking space environment