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


* 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.

- **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.

---

[Intensity-free Network](#)

Walking space dataset for wheel chairs

---

Proposed Network

**CosPointPillars**
- As an alternative channel for reflection intensity, Cosine Estimation Network is added to PointPillars\(^4\)

**Why Cosine?**
- Lambertian model

\[
\text{Intensity} \propto \frac{K_\lambda \cos \theta}{d^2}
\]

\(\theta: \text{Incident angle}\)
\(d: \text{Distance}\)
\(K_\lambda: \text{Object reflectivity}\)

Use \(\cos \theta\) 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

Proposed Network

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

## Evaluation

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

### Metric
- IoU in 2D Birds’ Eye View
- Avg. Precision

### 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

<table>
<thead>
<tr>
<th>Network</th>
<th>Reflectance</th>
<th>KITTI (IoU 0.5)</th>
<th>NuScenes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Easy</td>
<td>Moderate</td>
</tr>
<tr>
<td>PointPillars w/</td>
<td></td>
<td>84.36</td>
<td>79.97</td>
</tr>
<tr>
<td>w/o</td>
<td></td>
<td>80.22</td>
<td>75.92</td>
</tr>
<tr>
<td>CosPointPillars w/o</td>
<td></td>
<td>82.35</td>
<td>77.29</td>
</tr>
</tbody>
</table>
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)

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