HAL
open science

KITTI-CARLA: a KITTI-like dataset generated by CARLA Simulator

Jean-Emmanuel Deschaud

▶ To cite this version:

Jean-Emmanuel Deschaud. KITTI-CARLA: a KITTI-like dataset generated by CARLA Simulator. 2021. hal-03445117

HAL Id: hal-03445117
https://hal.science/hal-03445117v1

Preprint submitted on 23 Nov 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

KITTI-CARLA: a KITTI-like dataset generated by CARLA Simulator

Jean-Emmanuel Deschaud[1]

arXiv:2109.00892v1 [cs.CV] 17 Aug 2021

***Abstract*— KITTI-CARLA is a dataset built from the CARLA v0.9.10 simulator [1] using a vehicle with sensors identical to the KITTI dataset [2]. The vehicle thus has a Velodyne HDL64 LiDAR positioned in the middle of the roof and two color cameras similar to Point Grey Flea 2. The positions of the LiDAR and cameras are the same as the setup used in KITTI.**

The objective of this dataset is to test approaches of semantic segmentation LiDAR and/or images, odometry LiDAR and/or image in synthetic data and to compare with the results obtained on real data like KITTI. This dataset thus makes it possible to improve transfer learning methods from a synthetic dataset to a real dataset.

We created 7 sequences with 5000 frames in each sequence in the 7 maps of CARLA providing different environments (city, suburban area, mountain, rural area, highway...). The dataset is available at `http://npm3d.fr`

I. INTRODUCTION

Our approach is to generate a synthetic dataset in a realistic simulated environment and having the ground truth for semantic segmentation, instance segmentation, odometry poses.

II. CARLA

CARLA is an open source simulator for autonomous vehicles which allows to simulate multiple sensors, especially cameras and LiDARs (Figure 1). We use it to simulate a LiDAR and two color cameras boarded on the roof of a car traveling randomly in accordance with the traffic laws. The simulation is launched with a fixed time step of 0.001s. This means that the complete physics will be updated at every step: we generate all the data at each step, then move on to the next step until the end of the simulation. This frequency of 1000Hz makes it possible to simulate the rolling-shutter effect of LiDAR sensors.

III. SETUP

The simulated LiDAR is vertical and mounted on top of the roof of the vehicule. It has the same characteristics as a Velodyne HDL-64E (64 channels, from $-24.8°$ to $+2°$ vertical field of view, 10Hz revolution frequency, 80m maximum range so around 130k points per frame). The LiDAR is attached to the vehicle body with a global translation of ($X = 0.0m$, $Y = 0.0m$, $Z = 1.80m$) and is pointing backwards (Figure 2).

Let's define a frame as a complete revolution of the data coming from the LiDAR. During the revolution, the LiDAR attached to the vehicle moves. Because LiDAR frequency is 10Hz and CARLA simulation is 1000Hz, at each simulator step the LiDAR returns 1/100 of a frame. During this interval, the physics is not updated so all the points in a measurement reflect the same "static picture" of the scene taken at a same time.

Fig. 1. Examples of maps from the CARLA Simulator used to create our synthetic dataset. Respectively from top to bottom: city (Town01), highway (Town05) and farm (Town07).

We provide with the dataset all the poses of the LiDAR at 1000Hz allowing to know the ground truth of the poses to generate a point cloud. We provide a python script to compute the position of each point in world coordinate (thanks to the LiDAR transformation saved at each step).

The two simulated cameras have a resolution of 1392x1024 pixels with 72 degrees FOV. They work at 10Hz. Cameras are attached to the vehicle body with a

[1] MINES ParisTech, PSL University, Centre for Robotics, 75006 Paris, France {jean-emmanuel.deschaud@mines-paristech.fr}

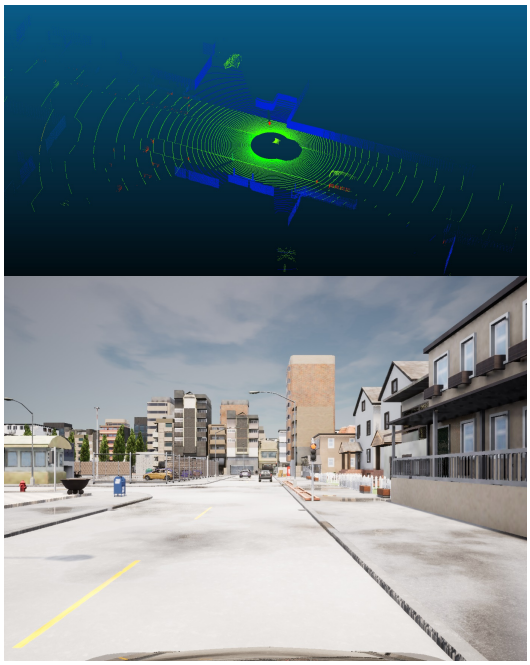

Fig. 2. Top, one frame of the LiDAR with ground truth semantic segmentation in color. Bottom, an image from the camera 0.

global translation of ($X = 0.30m$, $Y = 0.0m$, $Z = 1.70m$) for camera 0 and ($X = 0.30m$, $Y = 0.50m$, $Z = 1.70m$) for camera 1 (Figure 2).

IV. DATASET

We recorded 7 sequences using the 7 maps from Town01 to Town07 (Figure 1). We generated 5000 frames per sequence and for each of them we saved the full LiDAR transformations in world coordinates and corresponding timestamps in a ASCII file. We also saved the timestamps of the camera images and the calibration between the LiDAR and the 2 cameras in ASCII files.

REFERENCES

[1] A. Dosovitskiy, G. Ros, F. Codevilla, A. Lopez, and V. Koltun, "CARLA: An open urban driving simulator," in *Proceedings of the 1st Annual Conference on Robot Learning*, ser. Proceedings of Machine Learning Research, S. Levine, V. Vanhoucke, and K. Goldberg, Eds., vol. 78. PMLR, 13–15 Nov 2017, pp. 1–16. [Online]. Available: http://proceedings.mlr.press/v78/dosovitskiy17a.html

[2] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? the kitti vision benchmark suite," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*, 2012, pp. 3354–3361.