KI DATA TOOLING
The Data Kit for Automotive AI

Converting Real-World Measurements into Scenario-Based Simulations

Dominik Salles | FKFS

Scenarios & Synthetic Data

Introduction

Capturing and labeling real-world data is time consuming and expensive. Synthetic data can support the process of training, testing and understanding ML algorithms, but this kind of data needs to be validated. Therefore, we created digital twins and converted real-world measurements into scenario-based simulations. The real-data footage includes ego vehicle, labelled, infrastructure, weather, and traffic-light data.

Challenges & Methods: Real-Data Footage

Each data source needs to be preprocessed before it can be used for the scenario-based description.

Ego-Vehicle Data:

The ego vehicle's movement is described by measuring roll, pitch, heading, latitude and longitude. Although the GPS+RTK signal may be very precise, an offset in heading, latitude and longitude is subtracted for each sequence.

Labelled Data:

This data, or more precisely 3D bounding boxes, are used to define the surrounding dynamic objects. Human labelling errors are identified via SUMO before further processing.

Infrastructure Data:

Objects out of sight can still influence the sensors of the ego vehicle (e.g. mirroring effects). Infrastructure sensors and automatic object recognition are used for this purpose.

Weather Data:

This data may come from multiple sources, as one data source may not include all sensors needed for the weather recreation in the simulation.

Traffic-Light Data:

Each light signal must be recorded and paired with the corresponding light in OpenDRIVE.

The different data sources are depicted in figure 1.

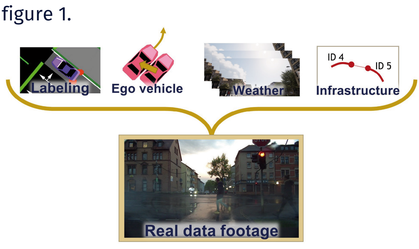

Figure 1: Preprocessing real data (© Robert Bosch GmbH | FKFS)

Challenges & Methods: Data Conversion

Apart from gathering and preprocessing the data, it also needs to be merged and smoothed.
As shown in figure 2, the data is captured at different timestamps and is synchronized with the real-data recording frequency of 10Hz. This is done by inter- and extrapolating the data points. The GPS signal is smoothed using additional techniques.

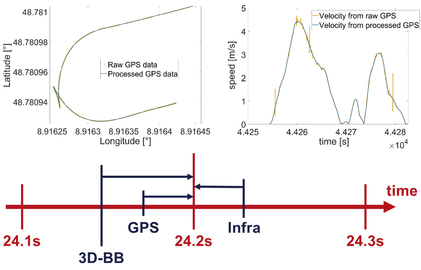

Figure 2: Smoothing and merging data (© FKFS)

The resulting output is written into OpenSCENARIO files. The sequences can then be replayed in a simulation framework, where the synthetic sensor data is generated [1].

References:

[1] D. Salles, L. Lang, M. Kehrer, and H.-C. Reuss, "A modular co-simulation framework with open source software and automotive standards," in *Int. Stuttgarter Symp.: Automobil-und Motorentechnik. Springer*, 2022, pp. 207–223.

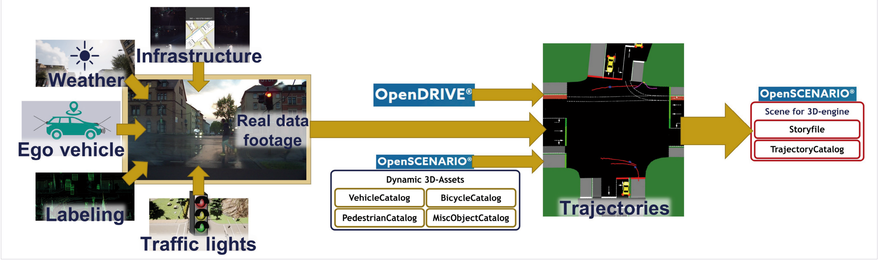

Figure 3: Scenario creation from real-world measurements (© TH Aschaffenburg | Robert Bosch GmbH | FKFS)

Partners

External partners

For more information contact:
dominik.salles@fkfs.de

KI Data Tooling is a project of the KI Familie. It was initiated and developed by the VDA Leitinitiative autonomous and connected driving and is funded by the Federal Ministry for Economic Affairs and Climate Action.

Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag