KI WISSEN
Automotive AI Powered by Knowledge

2.8 Tracking and Trajectory Prediction Using Scene Semantics

Nils Kornfeld | DLR

Tracking Pipeline

To be able to predict the future motion of road users it is important to gain a reliable model of the previous motion of all dynamic objects present in a traffic scene. To create this dynamic environment model the tracking pipeline pictured in Figure 1 is used:

- Objects of interest are detected by a neural network.
- Detections are fed as measurements to a multiple object tracking system.
- An unscented Kalman Filter estimates the state of each road user.
- The association is based on the euclidian distance between actual and predicted measurements.

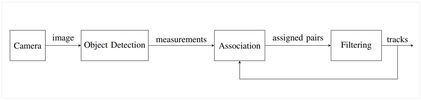

Figure 1: Multiple object tracking pipeline (© Deutsches Zentrum für Luft- und Raumfahrt, e. V.)

- Prediction is achieved with a physically plausible motion model for each kind of road user.
- Object dimensions are estimated as well.

The future development of the generated trajectories can be predicted by the existing Kalman Filter according to the underlying motion model or by a neural network used for prediction of a scene evolution[1].

Figure 2: Semantic segmentation image from synthetically created data in CARLA[2] (© Deutsches Zentrum für Luft- und Raumfahrt, e. V.)

Incorporating Scene Semantics

Scene semantics like knowledge of the 3-dimensional geometry of the environment, angle of view as well as the path and direction of traffic lanes can be used to aid both tracking and trajectory prediction:

- Prior domain knowledge can be used to improve the performance of object detectors[3]
- The measurement model uses information about peripheral infrastructure occlusion and viewing angle.
- If available, information about the direction and path of traffic lanes is used to find a plausible initial heading angle for tracked road users.
- The semantic information can be obtained from maps, annotated real world data sets or synthetically generated data with semantic annotation (see Figure 2).

The incorporation of viewing angle and geometric occlusion information in the measurement model enables the tracker to predict plausible measurements of vehicles which are only partially visible (see lower right corner of Figure 3) and to further predict trajectories of objects that are occluded by static parts in an image (see upper left corner of Figure 3).

Figure 3: Exemplary Tracking and trajectory prediction results of one partially and one fully occluded object (© Deutsches Zentrum für Luft- und Raumfahrt, e. V.)

Conclusion

Incorporating knowledge about a traffic scene into a multiple object tracking system, which uses a combination of data driven and traditional filtering approaches can help to overcome some of the system‘s inherent shortcomings.

References:

[1] Nils Kornfeld, Zachary Feng,: A Latent Variable Model State Estimation System for Image Sequences, 2019
[2] Alexey Dosovitskiy, German Ros, Felipe Codevilla et al.: CARLA: An Open Urban Driving Simulator, 2017
[3] Gurucharan Srinivas: Leveraging Knowledge for Traffic Sign Detection, 2024

Partners

External partners

For more information contact:
nils.kornfeld@dlr.de
andreas.leich@dlr.de
karsten.liesner@dlr.de

KI Wissen 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

Funded by the European Union NextGenerationEU

on the basis of a decision by the German Bundestag