11° ΔΙΕΘΝΕΣ ΣΥΝΕΔΡΙΟ για την ΕΡΕΥΝΑ ΣΤΙΣ ΜΕΤΑΦΟΡΕΣ
Καθαρές και Προσβάσιμες σε όλους Πολυτροπικές Μεταφορές
ICTR 2023
11th INTERNATIONAL CONGRESS on TRANSPORTATION RESEARCH
Clean and Accessible to All Multimodal Transport

Developing a Sensor Agnostic Infrastructure-Based Control Approach for Crossroad Assistance using Simulation

Lukas Merk, Yun-Pang Flötteröd, Andreas Leich, Sten Ruppe

Institute of Transportation Systems of the German Aerospace Center (DLR)

{lukas.merk, yun-pang.floetteroed, andreas.leich, sten.ruppe}@dlr.de

Abstract

Crossroad assistance has the potential to contribute to a fair distribution of road space and increase the safety of all road users. This paper presents an infrastructure-based control approach to support crossroad assistance by implementing remote-controlled turning maneuvers on a test track. The approach involves several steps, including the detection of traffic events using camera- based object recognition, inserting dynamic objects into the simulation via V2X transmission, adding lanes and traffic signal phases through V2X data forwarding to a high-precision simulation map, trajectory prediction of all dynamic objects in the simulation, sending a trajectory list to the test vehicle, and providing feedback through the vehicle's response to the trajectory list and re-detection of objects.

1 Introduction

Intersection assistance systems will emerge as a crucial component in preventing accidents and enhancing road safety. These systems aim to provide intelligent assistance at intersections. This paper explores the possibilities of an intelligent intersection assistance by creating a virtual representation of the real-time traffic conditions and continuously predicting the trajectories of all road users. By doing so, it can optimize driving patterns and, more importantly, proactively prevent accidents with timely interventions. The intersection assistance service acts as a comprehensive solution, aggregating all available information, applying traffic management measures, and actively influencing the traffic flow. Depending on the implementation and level of authority, this approach could potentially shift the responsibility of decision-making from the vehicle or driver to the local assistance system. For instance, in the case of automated vehicles, the intersection assistance system may dictate maneuvers, thus transferring responsibility from the vehicle's operator to the local assistance system.

2 Motivation

The study faced some difficulties in reliably tracking objects over several time steps due to issues such as occlusion and blind spots at the intersection. Simple approaches, such as applying a Kalman filter to extrapolate an object's motion vector, are not sufficient in more complex scenarios such as crossroads. Traditional methods may fail to identify an object with its previous ID after occlusion if the object has changed direction during the occlusion, which frequently occurs during turning maneuvers. Tracking supported by a microscopic traffic simulation tool, such as SUMO (Alvarez Lopez et al. 2018), can address this problem because the simulation knows the mentioned directional changes and driving lanes.

3 System Framework and Implementation

The use of Collective Perceptive Messages (CPM) for transmitting object lists makes the simulation independent of the deployed sensors. Thus, it is possible to replace or extend the object detection system with another solution. This infrastructure-based approach can support crossroad assistance by improving traffic flow and increasing the safety of all road users.

3.1 System constraints

In this research, several software components, developed at DLR's institute of transportation systems, have been extended to improve the overall functionality of the system.

3.1.1 BOB Traffic Monitoring

BOB (Building Blocks) is a DLR software framework that handles sensory inputs and provides higher level outputs. Specifically, in the context of this paper, BOB serves the function of camera-based object recognition and object tracking. To achieve seamless integration, BOB leverages the capabilities of the ROS (Robot Operating System).

In this setup, rather than a direct interface, V2X (Vehicle-to-Everything) messages are employed to facilitate the transportation of essential information from BOB to the Sumo interface TraCI via ROS. This approach ensures efficient and effective communication, enhancing the overall traffic monitoring process.V2X Framework

The Vehicle- to-Everything (V2X) framework (Bottazzi, Wesemeyer, Bargmann, & Ruppe, 2021) plays an important role in the infrastructure-based control approach. It is responsible for transmitting messages between the simulation and the test vehicle. To support the crossroad assistance system, the V2X framework has been extended with a project-specific MCM (Maneuver Coordination Message) V2X message. The extension has been implemented for both the air interface and the interface to ROS.

The MCM message extension includes two functions. The first function is the VehicleCapabilities container, which allows the test vehicle to communicate its limiting characteristics. These consist of the tolerated distance to the front and rear, the tolerated maximum and minimum speed, and the maximum tolerated acceleration in longitudinal and lateral di- rections. This data is necessary because different vehicles require very different control requirements. For example, a sports car may turn faster than a bus. In- formation about the external dimensions of the objects to be controlled will be obtained through object recognition or evaluation of CAM messages and is not part of the MCM extension.

The second function is the VehicleAdvice container, which has been extended with the TrajectoryAdvice fields. This allows the trajectory calculated in the simulation to be transmitted to the vehicle. It is a list containing Delta Position, Delta Time, Speed, and longitudinal acceleration.

By extending the V2X framework with these functions, the infrastructure-based control approach can communicate critical information to the test vehicle in real-time. The vehicle can then use this information to execute precise turning maneuvers and increase the safety of all road users.

3.1.2 Sumo traffic simulation

The traffic simulation is a critical component of the infrastructure-based control approach. It is responsible for simulating the traffic conditions at the inter- section and providing trajectory predictions for all dynamic objects in the simulation. To support the crossroad assistance system, the TraCI interface of the SUMO traffic simulation has been extended to receive CPM object data from Bob Traffic Monitoring and SPAT, MAP, CAM, and MCM messages for- warded from ROS. The interface has also been ex- tended to send MCM messages to the ROS system.

The trajectory prediction of dynamic objects is a challenging problem due to the complexity of traffic conditions at intersections. The infrastructure-based control approach addresses this problem by using a combination of object recognition, V2X transmission, and trajectory prediction in the SUMO traffic simulation. By predicting the trajectory of each dynamic object in the simulation, the system can provide ac- curate trajectory lists to the test vehicle, which can then execute precise turning maneuvers.

3.1.3 Selection of the demo intersection

The developed intersection assistance system is to be implemented in Brunswick, Germany, as part of the research. The Tostmannplatz is an ideal demonstration location due to its status as a research and innovation intersection of the DLR. Its topographic conditions are already well known from several other projects. MAPM (Map Message) and SPAT (Signal Phase Message) messages are available. Additionally, this intersection is equipped with a hemispherical camera and object detection hardware, which is a requirement for this project.

3.1.4 Generating test data

To evaluate the effectiveness of the infrastructure-based control approach, test data was generated at the Tostmannplatz intersection in Brunswick. The object detection system was put into operation, and CAM (Cooperative Awareness Message), MAPM, and SPATEM messages available on the air interface were transferred to ROS. Several turning maneuvers were manually performed at Tostmannplatz with the test vehicle to generate data that resembles the use case. This provided a recording of all on-site conditions that can be played back virtually as often as desired during further development of the simulation and control.

In addition, the camera used for object detection was calibrated using geo-referenced points from a satellite image. The calibration involved manually transferring the geo-coordinates to the camera image and calculating the camera angle and distortion resulting from the fisheye lens. The resulting systematic errors were based on inaccuracies in the compensation of lens curvature and the assumption that the road is a completely flat surface. Overall, the localization accuracy with a systematic offset of a maximum of a few centimeters is sufficient for the project.

The intersection assistance system uses the estimates of the tracking pipeline to initialize and up- date a traffic simulation. The tracking pipeline consists of a detector, a data association algorithm, and a Kalman filter. The detector identifies objects within the camera image, and the data association algorithm as- signs unique IDs to each object and tracks their movement over time. The Kalman filter is used to predict the future position of the objects based on their cur- rent position and velocity.

To test the effectiveness of the tracking pipeline, the test vehicle was driven through the intersection multiple times while performing different turning maneuvers. The resulting object lists and trajectory predictions were compared to ground truth data obtained through manual annotation. The results show that the tracking pipeline is effective in reliably tracking objects over multiple time steps, even in complex scenarios such as crossroads.

3.2 Calibration of Object Detection and Limits of Localization Accuracy in the Existing System

The camera calibration is based on easily identifiable geo-coordinates from a satellite image. These are manually transferred to the camera image, and the camera angle in height and the distortion resulting from the fisheye lens of the camera are automatically calculated. The resulting systematic errors are based on inaccuracies in the compensation of lens curvature and the assumption that the road is a completely flat surface. Overall, the localization accuracy with a systematic offset of a maximum of a few centimeters is sufficient for the project.

3.3 Localization Estimations by the Tracking Pipeline

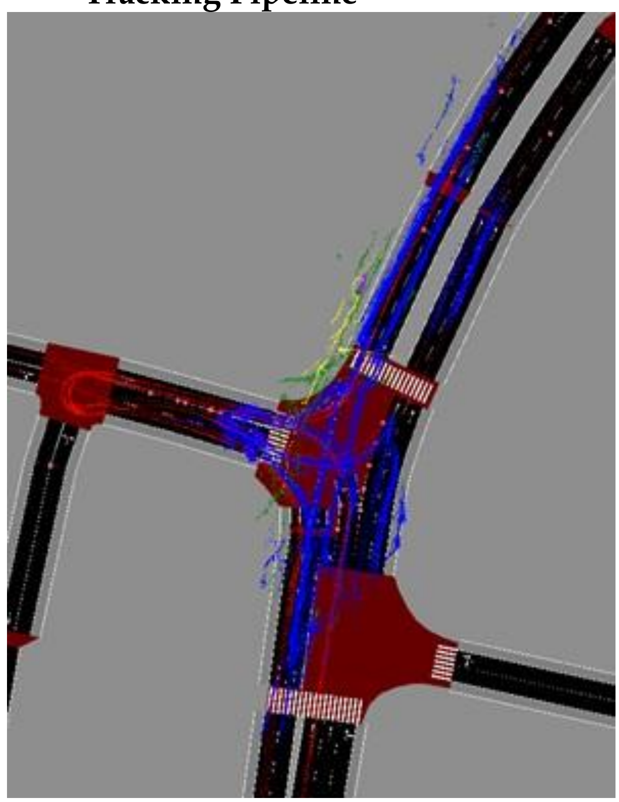

Figure 1 Comparison of CAM based trajectories and trajectories from CPM (CAM red, CPM vehicles in blue)

The intersection assistance system uses the estimates of the tracking pipeline to initialize and update a traffic simulation. The following sources of errors exist in tracking:

- False associations: The trajectory of an object may be falsely assigned detections of another object, especially when traffic is dense and the viewing angle is flat.

- Fragmentations: Due to partial occlusions of the traffic participants, a trajectory can be fragmented into several pieces.

- ID switches: It frequently happens that a traffic participant is tracked with an ID and later receives a different ID, e.g., because the trajectory could not be continued due to missing measurements (occlusion) and a new ID is later assigned.

The influences of errors are summarized in a quality metric, the so-called MOTA. The MOTA of the tracking algorithm used is around MOTA ~= 77.1 [?]. Thus, it is necessary to develop strategies for implementing intersection assistance that can handle errors that are still unavoidable according to the state of the art.

3.4 Trajectory Prediction through Simulation

The prediction of trajectories through simulation is a promising approach for enhancing the performance of autonomous vehicles (AVs). AVs rely on real-time sensor data to detect and respond to the movements of surrounding traffic participants, such as adjusting their speed and acceleration. However, the available information is limited by the capabilities of the vehicle's sensor system. The application of V2X communication is one way to expand the local information scope, enabling better route planning at decision points such as intersections, junctions, and turning areas. Prediction of future movements of local traffic participants using a vehicle external simulation is a feasible way to fore- cast trajectories when corresponding real-time input data is available. The predicted trajectories and sub- sequent trajectory proposals can then be used by AVs for proactive trajectory planning and speed control, such as early lane changes and energy-efficient speed management.

One of the objectives of this research is to develop an interface for extracting the classes, trajectories, speeds, and destination directions of traffic participants from a given data source and inserting them into the simulation at the respective time stamps. As simulations can run faster than real-time events, simulated trajectories can be used as predicted trajectories for AVs. To achieve this, video data collected from the camera at the Tostmannplatz research node in Brunswick was utilized as the data source, and the DLR's microscopic traffic simulation suite SUMO was used as the simulator. The overall process of the simulation is illustrated in Figure 2. This paper first describes the network preparation work and available video data in Rosbag format, followed by an explanation of the conducted work and considered points. Finally, a summary of the current research outcome is presented.

3.5 Object Alignment

The implementation of the proposed method was carried out using Python and SUMO 1.11. The focus was on (1) setting up the processing interface between the simulation and video data source and (2) publishing simulated (predicted) vehicle trajectories in a ROS topic. The processes were implemented for online and offline applications, as shown in Figure 2. For each subscribed message, a simulation is started for 30 timestamps (3 seconds) in response to the given AP requirement. Objects in each message are parsed and inserted into the simulation network at the first timestamp. Then, matching objects are simulated, and the corresponding data is published at each timestamp until 30 timestamps are reached. The entire process runs continuously until it is interrupted for online application or all messages in a specific rosbag file are processed for offline application. The processing time for each cycle is currently under 0.1 seconds for the example rosbag file mentioned in Section 1.2, meeting the given requirement. The following sections describe how objects can be matched and inserted into the simulation.

In this study, the process of aligning object data with the simulation in the context of trajectory planning is described. The object data is first parsed and stored in a timestamp-based map, which is then used in conjunction with the TraCI application to start the simulation. The position of each object is matched with the nearest lane in the simulation at defined time intervals, with a search radius of 1 meter. Only cars are currently considered due to the accuracy limitations of the tracked object positions. If the corresponding lane is valid for the object class, such as a passenger, the alignment is deemed successful. The distance from the matching point to the start of the matching lane is also calculated and used to insert the object at the precise location in the simulation network. Each object in the simulation requires a destination. If there is no information available on the object's movement direction or destination for objects with matching lanes and positions, an estimate is made and assigned to the respective object. Possible target edges for lanes approaching or near the Tostmannplatz intersection are predefined in a map based on allowed turning marks in practice. If multiple target edges exist, a target edge is randomly selected. When adding an object to the simulation, the object ID, speed, start position, start lane ID, start and target edges are specified. The corresponding route is then calculated by invoking the TraCI function rerouteTraveltime, as SUMO requires route information. Currently, added objects are only loaded into the simulation at the next timestamp. The corresponding time adjustment for inserting objects is also considered. If an object is already in the simulation, its speed and target are updated at each respective timestamp based on extracted object data such as position and velocity. If there is no extracted data available for a particular inserted object the updated object list is returned to BOB, and the process repeats for the next time step. By incorporating feedback from the simulation, the prediction can be refined and the object persistence improved. This pipeline is currently under development and will be tested in future work.

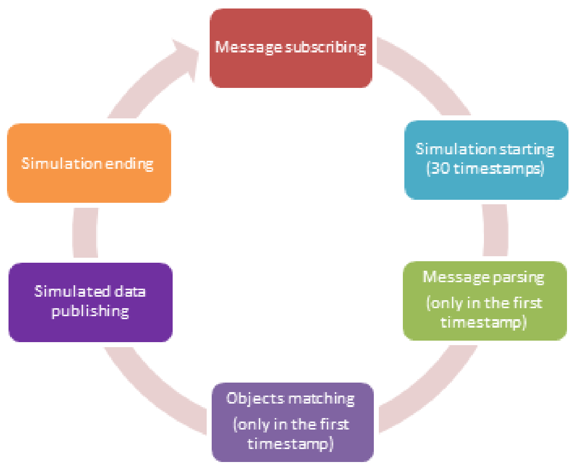

Figure 2 Process of the simulation

3.6 Latency Considerations

In order for the intersection assistance system to calculate a valid trajectory for the vehicle, several latency-prone steps are involved in the processing chain. The camera operates with a temporal resolution of 15 frames per second. Changes in the traffic scenario that occur in a time frame below 1/15 seconds cannot be detected by the system in principle. Additionally, there are other delays such as object recognition, transmission to the SUMO interface, and calculations within the simulation itself. Even after the generation of the trajectory through the transmission as MCM, latencies of several milliseconds are expected. Finally, the delays in the test vehicle's system itself cannot be underestimated.

3.7 Tracking Objects Reliably

In Section 3.2 and 3.3, it was shown that tracking objects reliably over several time steps is challenging. Problems can arise, for example, due to obstruction and blind spots at the intersection. Simple approaches, such as applying a Kalman filter to extrapolate the motion vector of an object, are insufficient in more complex scenarios, such as an intersection. For example, in the conventional method, an object cannot be tracked with its previous ID after an obstruction if it has changed direction during the obstruction, which frequently happens during a turning maneuver. A tracking method supported by SUMO can provide a solution here, as the aforementioned direction change and driving lanes in the simulation are known.

3.8 Structure of the Pipeline for Object Recognition

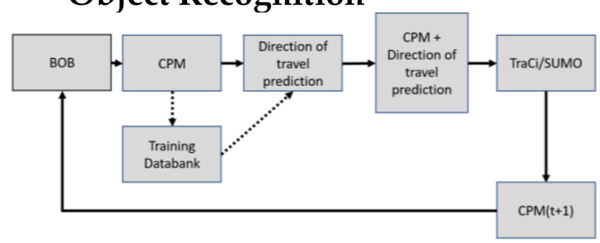

Figure 3 Planned structure of the pipeline for object recognition with improved object persistence through feedback of the prediction.

In the proposed system, BOB initiates the process by creating an object list from the camera image at time (t) and subsequently converts it into a CPM.

Secondly, a trained neural network attaches its prediction for the direction of travel to each object (the development of the neural network for travel prediction is out of scope for this paper). It sends this prediction together with the CPM to the SUMO simulation. SUMO adds the objects to its simulation and assigns the driving directions from the prediction to the objects, SUMO simulates the progress of the traffic scenario for the time (t + 1) and writes a CPM for time (t + 1). This CPM is converted back into the format of an object list that BOB requires. Finally, BOB compares the new camera image at time (t +1) with the prediction from SUMO and assigns the objects with the same object ID as in the previous loop. This matching should be greatly improved since we now have a precise prediction to where object have moved in between two timesteps. CPM to Transmit Object lists.

By using CPM to transmit the object lists, the simulation can be kept independent of the sensors. This configuration allows, for example, BOB to be replaced or extended with a commercial object recognition solution or other new approaches for generating object lists.

By transmitting an optimized trajectory through MCM to one or more vehicles, the system creates a secondary feedback loop. At time t+1, the object recognition can be compared with the simulation's prediction, and additionally, the advised trajectory of the test vehicle can be compared to its actual trajectory. It is anticipated that the time required for a test vehicle to respond to a suggested trajectory is significantly longer than one timestep in the object detection loop. Further research will determine whether this system is inherently self-stabilizing or necessitates substantial tuning of the test vehicle's input.

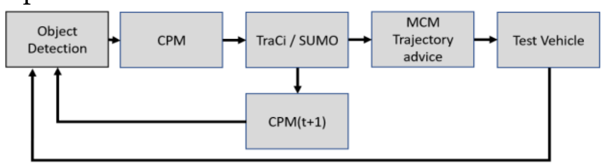

Figure 4 shows a system consisting of a fast simulated and a slow real-world feedback loop for trajectory advice

3.9 Further development

The DLR Institute of Transportation Systems has been at the forefront of research and development in the field of automated driving for several years. Leveraging our extensive expertise and knowledge, we have successfully integrated external sensor input to enhance the capabilities of automated vehicles (Lapoehn, Heß, Böker, Böhme, & Schindler, 2021). Building on this strong foundation, our ongoing efforts aim to pave the way to meet the requirements of German level 4 automated driving regulations, specifically focusing on the advancement of remote-operated driving.

As automated driving technologies rapidly progress, the concept of remote-operated driving emerges as a promising solution to address certain challenges associated with fully autonomous vehicles. Remote operation allows for human intervention and oversight, providing an additional layer of safety and control during complex driving scenarios or in situations where the vehicle encounters unprecedented obstacles. Furthermore, it enables seamless transition from autonomous mode to remote control, thus ensuring a smooth and efficient driving experience.

The research presented in this paper serves as a critical stepping stone towards the development and implementation of remote-operated driving. By utilizing trajectories generated from advanced simulation techniques, we aim to bridge the gap between virtual environments and real-world driving scenarios. These simulated trajectories are carefully crafted to mirror real-world conditions and intricacies, allowing us to test and validate automated vehicles extensively in controlled test scenarios.

With the successful integration of simulation-generated trajectories into the operation of autonomous vehicles, we expect to achieve significant advancements in the accuracy, safety, and reliability of automated driving systems. This research opens up new avenues for optimizing vehicle control algorithms and decision-making processes, ultimately leading to a higher level of autonomy and enhanced performance in a wide range of driving situations.

In conclusion, the findings presented in this paper provide a solid foundation for the continued development of remote-operated and infrastructure-assisted driving. We are confident that our research will inspire further studies, collaborations, and innovations in the field.

4. References-Bibliography

Bottazzi, M., Wesemeyer, D., Bargmann, M., & Ruppe, S. (2021). DLR V2X Software Framework: A CCU-agnostic solution for transceiving V2X messages. Hamburg: ITS World Congress 2021.

Lapoehn, S., Heß, D., Böker, C., Böhme, H., & Schindler, J. (2021). Infrastructure-aided Automated Driving in Highly Dynamic Urban Environments. Hamburg: ITS World Congress 2021.

Pablo Alvarez Lopez, M. B.-W.-P. (2018). Microscopic Traffic Simulation using SUMO. IEEE Intelligent Transportation Systems Conference (ITSC).