Scenario-based Validation of Autonomous Vehicles using Augmented Reality

Jean-Baptiste Horel[1], Alessandro Renzaglia[2], Radu Mateescu[1] and Christian Laugier[1]

***Abstract*— Validation of autonomous vehicles (AVs) is a critical task in their development and for their approval on public roads. Scenario-based testing is the state-of-the-art validation method and is recommended by international automotive regulators. While simulated execution of critical scenarios is essential, it cannot yet fully replace real-world testing, which however remains tedious, time-consuming, and resource-intensive. In this work, we propose an enhanced methodology using Augmented Reality to bridge the gap, providing an intermediate testing method that enables comprehensive real-world testing with reduced cost and improved realism. To demonstrate this methodology, we conducted tests in a controlled environment using six critical scenarios selected from road crash studies.**

I. INTRODUCTION

Validation of perception and navigation modules in autonomous vehicles remains a challenge for the development and widespread deployment of autonomous driving technologies. Simulated environments are a fundamental tool because they are cheap, easy to use, and can reproduce a large number of highly varied scenarios in different environmental conditions. However, their representation of the real world and the real behavior of the system under test is far from perfect and cannot be the sole source of validation. Real-world experiments remain necessary, but are significantly more expensive and difficult to perform. Controlled environments allow for a limited range of reproducible situations, while testing on real roads can be dangerous and requires countless hours of driving to cover a significant amount of test space, as the most relevant scenarios for validation are also the least likely to occur.

This work presents a scenario-based validation methodology where a limited set of critical scenarios are selected and reproduced in controlled environments thanks to a real-time Augmented Reality (AR) framework [1]. This approach allows us to easily and safely put the vehicle under test in potentially dangerous and complex situations, such as collision or near-collision scenarios with other vehicles and pedestrians, without any risk. By combining the advantages of simulated and real systems, the results are as reliable as fully real experiments, but cheaper, risk-free, and easier to obtain. Our work focuses on the testing and validation of a Lidar-based perception module based on probabilistic occupancy grids, which allows the estimation of collision risk and triggers emergency maneuvers. Our AR framework

[1]Univ. Grenoble Alpes, Inria, 38000 Grenoble, France.
[2]Inria, INSA Lyon, CITI, UR3720, 69621 Villeurbanne, France.
This work was supported by the PRISSMA project, co-financed by the French Grand Défi on Trustworthy AI for Industry.

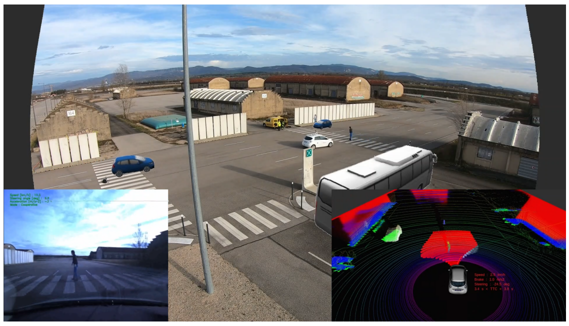

Fig. 1. View of Scenario 6 execution with the real Zoé AV (white car), a real pedestrian target, and virtual actors (blue car, white bus, and pedestrian). The bottom left inset image shows the AV front camera while the bottom right inset image is a visualization of the AV perception and planning.

provides a flexible way to introduce any virtual element in real time in the data of the LiDAR sensors of the vehicle. In our study, we selected 5 critical scenarios from traffic collision studies [2] and the resulting data are analyzed using a recently proposed metric specifically designed for autonomous navigation based on probabilistic occupancy grids [3]. The rest of this extended abstract presents more in detail this proposed methodology and the attached video submission shows the real experiments conducted in a road testing facility and the preliminary results that we obtained.

II. VALIDATION METHODOLOGY

The proposed validation methodology can be divided into four main components: scenario definition, testing means, dataset generation and metric evaluation.

a) Scenario selection and generation: AVs are expected to drive on public roads and must deal with dangerous human actions and the crashes they can cause. Road crashes have been studied and recorded for several decades. Recent studies [2] classified crashes and identified the most occurring and critical scenarios into comprehensive libraries for AV validation. While not being the most challenging scenarios, they are the ones we know AVs will be exposed to, therefore AV's safety must be evaluated on them. However, manually specifying concrete tests from scenario definitions to obtain a high level of coverage remains a tedious and time-consuming activity. For this reason, we developed an automated scenario generation method based on formal conformance testing tools [4], [5]. Starting from an abstract definition of a critical scenario, as identified from cited studies, the method generates all concrete test cases covering the scenario, which can be executed with various testing means (simulation, real world, or augmented reality).

b) Simulation, Augmented reality and real world: Generated scenarios can be executed in both real and virtual environments. Simulation allows us to easily and massively evaluate the AV on every selected scenario while Real-world testing remains a necessity to confirm the realism of virtual tests. The in-between solution we propose leverages AR [1]. This method is a middle-ground of cost and realism, between simulation and real world. Virtual vehicles or pedestrians can be added to the real scene at the sensor level. Obstacles (e.g., pedestrians, cyclists, vehicles) are safe to collide while testing the AV in real conditions. It is not a replacement for simulation and real tests but the advantage of AR is to gather more data on the real AV, by testing more scenarios than typical real tests enable.

c) Dataset generation: During experiments with the AV, data is recorded for offline evaluation and Ground Truth (GT) generation. AR allows using simulation data (e.g., virtual actor poses, shapes, velocities) to generate GT for the perception and prediction. We can also safely measure if a collision occurred during a scenario or the time to collision with other actors. Performing the test in a controlled environment, a virtual model of the test environment is finally necessary to create a GT of the surrounding environment (buildings, road typology, sidewalks, etc.).

d) Evaluation metrics: The AV experiments focus on occupancy grid (OG) based perception, but the methodology can also apply to planning and control evaluation with dedicated metrics [6]. For perception, we designed a metric [3] dedicated to the similarity evaluation of two probabilistic OGs (i.e., an inference of the environment and a desired GT). This metric is well suited for AV validation: instead of doing a cell-wise comparison of OGs, it simulates a navigation algorithm on each grid to compare the corresponding navigation behavior.

III. EXPERIMENTS & RESULTS

a) Experiment: We applied the methodology defined in the previous section to an autonomous Renault Zoé. The autonomous navigation of this vehicle under test is achieved using an RTK GPS and an IMU for localization, 3D LiDARs for perception and planning, and automated steering, throttle and brakes commands.

We conducted the experiments in a controlled environment, the urban area of the dedicated facility Transpolis (https://transpolis.fr/). Figure 1 shows the test environment.

We tested 6 critical scenarios taken from the results of a French study on crash reports [2], listed in Table I and presented in the attached video. Scenarios 1 to 5 were executed with AR virtual actors while scenario 6 mixed AR with a real pedestrian target (in front of the AV in Figure 1). Scenario 6 is comparable to scenarios 4 and 5, allowing us to specifically compare their outcomes to study AR and real-world tests.

We recorded a total of 120 scenarios, 1H of driving, and 850GB of data. The AV's software is ROS-based [7], all data were recorded in ROS format with the `rosbag` tool.

	Scenario name	#Tests	Collision rate
1	Front collision	20	30%
2	Rear collision	20	50%
3	Side collision	20	0%
4	Pedestrian crossing	20	13%
5	Occluded pedestrian crossing	20	24%
6	Real pedestrian target	20	50%

TABLE I

NUMBER OF RECORDINGS PER SCENARIO AND COLLISION RATES

We used the Cartographer SLAM algorithm [8] on a dedicated bag to generate a GT occupancy grid of the whole test area. The grid contains only the occupancy of the environment static elements, only the actors in the scenarios being expected to be dynamic. AR actors' GT is easily generated (see II-.0.c) and the pedestrian target, whose shape is known, has an accurate position and speed tracking system.

b) Qualitative results: The experiment gave us a better understanding of the AV perception and navigation performances. As shown in Table I, the collision rate varies from one scenario to another, it reflects both the safety of the AV and the criticality of the scenario for the AV. For example, the collision rate of scenario 5 is higher than that of scenario 4 because scenario 5 is, by design, more critical due to the occlusion of a pedestrian by a bus. The experiment demonstrated the AV's efficiency in performing emergency stops to avoid forward obstacles. However, it revealed a deficiency in predictive capabilities for dynamic avoidance, as seen in scenario 2.

c) Quantitative results: The primary goal of the experiment is to evaluate the AV's perception using the metric described in II-.0.d on the recorded perception OGs and the generated GT. This final step of the methodology is currently in progress. We have successfully generated the GT OGs for scenarios 1 to 5 and are now integrating the pedestrian target data to generate the GT for scenario 6. Leveraging the modularity of ROS and AR, we plan to execute these scenarios in simulation. This will enable us to quantitatively compare simulation, AR and real-world tests, specifically focusing on scenarios 4 and 5 (with AR virtual pedestrians) and scenario 6 (with a real pedestrian target).

IV. CONCLUSION

This work proposed a novel methodology for the safe and efficient validation of autonomous vehicles (AV) in controlled environments. Its main steps are: *i)* an automated concrete scenario generation method based on a formal conformance testing tool, *ii)* an Augmented Reality (AR) framework to merge virtual and real objects in real time to easily and safely recreate critical scenarios, *iii)* the definition of a specific metric designed to evaluate probabilistic occupancy grids for autonomous navigation. Tests conducted with a real AV demonstrated the feasibility of the proposed methodology and allowed the identification of both correct and improvable behaviors of the AV.

The obtained preliminary results will be followed by a more extensive and quantitative analysis of the recorded AV data. This will allow a more in-depth evaluation of both the perception and navigation algorithms and our AR framework.

REFERENCES

[1] T. Genevois, J.-B. Horel, A. Renzaglia, and C. Laugier, "Augmented Reality on LiDAR data: Going beyond Vehicle-in-the-Loop for Automotive Software Validation," in *IV 2022 - 33rd IEEE Intelligent Vehicles Symposium IV*. Aachen, Germany: IEEE, Jun. 2022, pp. 1–6. [Online]. Available: https://inria.hal.science/hal-03703227

[2] V. Ledoux, R. Krishnakumar, and V. Hervé, "Livrable l2.8 situations d'interactions accidentogènes : enjeux," 2019. [Online]. Available: https://surca.univ-gustave-eiffel.fr/livrables-et-publications/wp2-etat-de-lart-donnees-accidentologiques/

[3] J.-B. Horel, R. Baruffa, L. Rummelhard, A. Renzaglia, and C. Laugier, "A navigation-based evaluation metric for probabilistic occupancy grids: Pathfinding cost mean squared error," in *IEEE International Conference on Intelligent Transportation Systems (ITSC)*, 2023, pp. 3148–3153.

[4] J.-B. Horel, C. Laugier, L. Marsso, R. Mateescu, L. Muller, A. Paigwar, A. Renzaglia, and W. Serwe, "Using formal conformance testing to generate scenarios for autonomous vehicles," in *Design, Automation & Test in Europe Conference & Exhibition (DATE)*. IEEE, 2022, pp. 532–537.

[5] J.-B. Horel *et al.*, "Verifying collision risk estimation using autonomous driving scenarios derived from a formal model," *Journal of Intelligent & Robotic Systems*, vol. 107, no. 4, p. 59, 2023.

[6] H. Caesar *et al.*, "Nuplan: A closed-loop ml-based planning benchmark for autonomous vehicles," in *CVPR ADP3 workshop*, 2021.

[7] M. Quigley, K. Conley, B. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, A. Y. Ng *et al.*, "Ros: an open-source robot operating system," in *ICRA workshop on open source software*, vol. 3, no. 3.2. Kobe, Japan, 2009, p. 5.

[8] W. Hess, D. Kohler, H. Rapp, and D. Andor, "Real-time loop closure in 2d lidar slam," in *2016 IEEE International Conference on Robotics and Automation (ICRA)*, 2016, pp. 1271–1278.