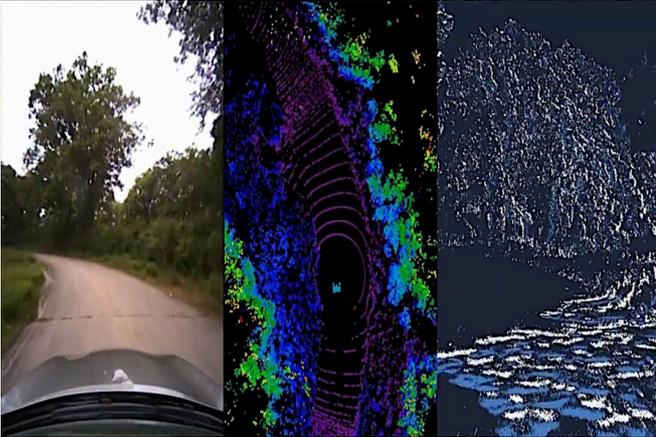


Researchers release open-source photorealistic simulator for autonomous driving

21 June 2022, by Rachel Gordon



VISTA 2.0 is an open-source simulation engine that can make realistic environments for training and testing self-driving cars. Credit: MIT CSAIL

Hyper-realistic virtual worlds have been heralded as the best driving schools for autonomous vehicles (AVs), since they've proven fruitful test beds for safely trying out dangerous driving scenarios. Tesla, Waymo, and other self-driving companies all rely heavily on data to enable expensive and proprietary photorealistic simulators, since testing and gathering nuanced I-almost-crashed data usually isn't the most easy or desirable to recreate.

To that end, scientists from MIT's Computer Science and Artificial Intelligence Laboratory (CSAIL) created "VISTA 2.0," a data-driven simulation engine where vehicles can learn to drive in the real world and recover from near-crash scenarios. What's more, all of the code is being open-sourced to the public.

"Today, only companies have software like the type of simulation environments and capabilities of

VISTA 2.0, and this software is proprietary. With this release, the [research community](#) will have access to a powerful new tool for accelerating the research and development of adaptive robust control for autonomous driving," says MIT Professor and CSAIL Director Daniela Rus, senior author on a paper about the research.

VISTA 2.0 builds off of the team's previous model, VISTA, and it's fundamentally different from existing AV simulators since it's data-driven—meaning it was built and photorealistically rendered from real-world data—thereby enabling direct transfer to reality. While the initial iteration supported only single car lane-following with one [camera sensor](#), achieving high-fidelity data-driven simulation required rethinking the foundations of how different sensors and behavioral interactions can be synthesized.

Enter VISTA 2.0: a data-driven system that can simulate complex sensor types and massively interactive scenarios and intersections at scale. With much less data than previous models, the team was able to train autonomous vehicles that could be substantially more robust than those trained on large amounts of real-world data.

"This is a massive jump in capabilities of data-driven simulation for autonomous vehicles, as well as the increase of scale and ability to handle greater driving complexity," says Alexander Amini, CSAIL Ph.D. student and co-lead author on two new papers, together with fellow Ph.D. student Tsun-Hsuan Wang. "VISTA 2.0 demonstrates the ability to simulate sensor data far beyond 2D RGB cameras, but also extremely high dimensional 3D lidars with millions of points, irregularly timed event-based cameras, and even interactive and dynamic scenarios with other vehicles as well."

The team was able to scale the complexity of the

interactive driving tasks for things like overtaking, following, and negotiating, including multiagent scenarios in highly photorealistic environments.

Training AI models for autonomous vehicles involves hard-to-secure fodder of different varieties of edge cases and strange, dangerous scenarios, because most of our data (thankfully) is just run-of-the-mill, day-to-day driving. Logically, we can't just crash into other cars just to teach a [neural network](#) how to not crash into other cars.

Recently, there's been a shift away from more classic, human-designed simulation environments to those built up from real-world data. The latter have immense photorealism, but the former can easily model virtual cameras and lidars. With this [paradigm shift](#), a key question has emerged: Can the richness and complexity of all of the sensors that autonomous vehicles need, such as lidar and event-based cameras that are more sparse, accurately be synthesized?

Lidar sensor data is much harder to interpret in a data-driven world—you're effectively trying to generate brand-new 3D point clouds with millions of points, only from sparse views of the world. To synthesize 3D lidar point clouds, the team used the data that the car collected, projected it into a 3D space coming from the lidar data, and then let a new virtual vehicle drive around locally from where that original vehicle was. Finally, they projected all of that [sensory information](#) back into the frame of view of this new virtual [vehicle](#), with the help of neural networks.

Together with the simulation of event-based cameras, which operate at speeds greater than thousands of events per second, the simulator was capable of not only simulating this multimodal information, but also doing so all in real time—making it possible to train neural nets offline, but also test online on the car in augmented reality setups for safe evaluations. "The question of if multisensor simulation at this scale of complexity and photorealism was possible in the realm of data-driven simulation was very much an open question," says Amini.

With that, the driving school becomes a party. In

the simulation, you can move around, have different types of controllers, simulate different types of events, create interactive scenarios, and just drop in brand new vehicles that weren't even in the original data. They tested for lane following, lane turning, car following, and more dicey scenarios like static and dynamic overtaking (seeing obstacles and moving around so you don't collide). With the multi-agency, both real and simulated agents interact, and new agents can be dropped into the scene and controlled any which way.

Taking their full-scale car out into the "wild"—a.k.a. Devens, Massachusetts—the team saw immediate transferability of results, with both failures and successes. They were also able to demonstrate the bodacious, magic word of self-driving car models: "robust." They showed that AVs, trained entirely in VISTA 2.0, were so robust in the real world that they could handle that elusive tail of challenging failures.

Now, one guardrail humans rely on that can't yet be simulated is human emotion. It's the friendly wave, nod, or blinker switch of acknowledgement, which are the type of nuances the team wants to implement in future work.

"The central algorithm of this research is how we can take a dataset and build a completely synthetic world for learning and autonomy," says Amini. "It's a platform that I believe one day could extend in many different axes across robotics. Not just [autonomous driving](#), but many areas that rely on vision and complex behaviors. We're excited to release VISTA 2.0 to help enable the community to collect their own datasets and convert them into virtual worlds where they can directly simulate their own virtual autonomous vehicles, drive around these virtual terrains, train [autonomous vehicles](#) in these worlds, and then can directly transfer them to full-sized, real self-driving cars."

More information: [VISTA 2.0](#)

[VISTA 2.0: An Open, Data-driven Simulator for Multimodal Sensing and Policy Learning for Autonomous Vehicles](#), arXiv:2111.12083v1 [cs.RO]. arxiv.org/abs/2111.12083

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology

APA citation: Researchers release open-source photorealistic simulator for autonomous driving (2022, June 21) retrieved 6 July 2022 from <https://techxplore.com/news/2022-06-open-source-photorealistic-simulator-autonomous.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is provided for information purposes only.