Autonomous vehicle companies like Waymo use simulation environments to train, test, and validate their systems before those systems are deployed to real-world cars. There are countless ways to design simulators, including simulating mid-level object representations, but basic simulators omit cues critical for scene understanding, like pedestrian gestures and blinking lights. As for more complex simulators like Waymo’s CarCraft, they’re computationally demanding, because they attempt to model materials highly accurately to ensure sensors like lidars and radars behave realistically.

In SurfelGAN, Waymo proposes a simpler, data-driven approach for simulating sensor data. Drawing on feeds from real-world lidar sensors and cameras, the AI creates and preserves rich information about the 3D geometry, semantics, and appearance of all objects within the scene. Given the reconstruction, SurfelGAN renders the simulated scene from various distances and viewing angles.

Waymo SurfelGAN

Above: The first column shows surfel images (more on those below) under novel view, while the second column is the synthesized result from SurfelGAN. The third column is the original view.

Image Credit: Waymo

“We’ve developed a new approach that allows us to generate realistic camera images for simulation directly using sensor data collected by a self-driving vehicle,” a Waymo spokesperson told VentureBeat via email. “In simulation, when a trajectory of a self-driving car and other agents (e.g. other cars, cyclists, and pedestrians) changes, the system generates realistic visual sensor data that helps us model the scene in the updated environment … Parts of the system are in production.”

SurfelGAN

SurfelGAN makes use of what’s called a texture-enhanced surfel map representation, a compact, easy-to-construct scene representation that preserves sensor information while retaining reasonable computational efficiency. Surfels — an abbreviated term for “surface element” — represent objects with discs holding lighting information. Waymo’s approach takes voxels (units of graphic information defining points in 3D space) captured by lidar scans and converts them into surfel discs with colors estimated from camera data, after which the surfels are post-processed to address variations in lighting and pose.

To handle dynamic objects like vehicles, SurfelGAN also employs annotations from the Waymo Open Dataset, Waymo’s open source corpus of self-driving vehicle sensor logs. Data from lidar scans of objects of interest are accumulated so that in simulation, Waymo can generate reconstructions of cars and pedestrians that can be placed in any location, albeit with imperfect geometry and texturing.

One module within SurfelGAN — a generative adversarial network (GAN) — is responsible for converting surfel image renderings into realistic-looking images. Its generator models produce synthetic examples from random noise sampled using a distribution, which along with real examples from a training data set are fed to discriminators, which attempt to distinguish between the two. Both the generators and discriminators improve in their respective abilities until the discriminators are unable to tell the real examples from the synthesized examples with better than the 50% accuracy expected of chance.

Waymo SurfelGAN

The SurfelGAN module trains in an unsupervised fashion, meaning it infers patterns within the corpora without reference to known, labeled, or annotated outcomes. Interestingly, the discriminators’ work informs that of the generator — every time the discriminators correctly identify a synthesized work, they tell the generators how to tweak their output so that they might be more realistic in the future.