ZOOX INC.
May 2020

Key: Important terms | Stats / Need to verify | Need more info

CLAMS

CLAMS stands for calibration, localization, and mapping, simultaneously. The concept stems from Zoox co-founder and CTO Jesse Levinson's Ph.D. work at Stanford University, which involved adding a calibration stage into the traditional robotics SLAM stack. Using calibrated sensor data, Zoox's in-house CLAMS software tool allows our vehicles to accurately, precisely, and consistently understand where they are and how they are moving through the world. That information feeds into every other system in Zoox's autonomy stack.

Once the vehicle is calibrated we use the sensor data it accumulates to determine where and how the vehicle has moved through the world, i.e., localize the vehicle. That information is then processed offline by a mapping system which outputs geographic maps of the world. When the vehicle is ready for deployment, we load those maps onto the vehicle, verify that the vehicle is calibrated, and localize the vehicle to the maps so that it can navigate the mapped area autonomously.

Calibration

Calibration involves aggregating sensor data to figure out where the sensors are on the vehicle. Similar to the way your brain calibrates the information received by your two eyes to deliver a crisp, singular vision of the world, our system uses perception data to calculate where each sensor is mounted on the vehicle and determine how their fields of vision overlap. Calibration is a safety-critical component of the vehicle's autonomy stack—to safely navigate the world, the vehicle's sensor data must align into a coherent view.

Our calibration system is infrastructure-free—an approach that is unique to Zoox—meaning we can calibrate our sensors just about anywhere without boards or targets. As our vehicle drives through the world, its sensors take in tons of data about the geometry of the streets, buildings, and other landmarks it passes. The calibration system uses that data to understand how all of these geometric features relate to each other and to the vehicle's sensors. Because our sensor calibration process does not rely on infrastructure, it is easily scalable.

Reliable and super-precise sensor calibration is what allows us to build solid mapping and localization systems.

Mapping

Instead of relying on dedicated mapping vehicles, we simply drive one of our normal test vehicles manually around the area we want to map. Using calibrated sensor data from that drive, our system outputs dense geometric maps that display a high-definition 3D representation of the world. In addition to altitude and depth information, the 3D maps also capture the retroreflective intensity of colorful road markings like lane dividers and crosswalks. The Zoox Road Network (ZRN) is then built out on top of the 3D map. The ZRN is a graph structure with semantic information about the geometry, topology, and policy of the mapped area, such as street names, how and where roads are connected to each other, how many lanes they have, speed limits, etc. These two maps are layered together and sent to the Autonomous Stack for vehicles to download.

Because streets change and maps eventually go out of date, we've built our mapping software to be resilient to structural changes in the world. As our vehicles drive around and notice discrepancies between existing maps and sensor data, they contact the TeleOps team and flag areas of the map as out of date. Our cartography team then uses the vehicles' log files to update the map before sending out the update to the rest of the fleet. This self-healing capability is a part of normal operations and is only possible because of Zoox's full-stack approach.

Localization

The Six Degrees of Freedom LIDAR Localizer determines the vehicle's location and pose, which contribute to two localization coordinate frames, smooth and global, that are needed for driving.

In the same way humans use our eyes and inner ears to sense motion and orientation, the vehicle is equipped with a variety of sensors that contribute to an understanding of how the vehicle has moved through the world, including wheel speeds and odometry, steering angles, IMU sensors (which include gyroscopes and accelerometers), and GPS. In the future, we may incorporate radar odometry and a visual localizer into the sensor suite. These inertial measurements are added up to create the smooth frame, a continuous odometric frame that keeps pace with the vehicle.

However, because sensors accrue inaccuracies over time due to measurement noise, we also need to determine the true position of the vehicle using the LIDAR Localizer. By comparing information from the vehicle's sensors to the 3D map, the localizer pinpoints where the vehicle is in relation to the 3D architecture of its environment and how it is oriented (i.e., pose estimation) with respect to six degrees: the x-axis, y-axis, z-axis, roll, pitch, and yaw.

Reconciling the vehicle's true position with its sensed position requires constant minuscule corrections, and because these corrections disrupt the smoothness of the vehicle's trajectory, we process them in the global frame. The global frame uses perception data to continuously jump/snap/correct the vehicle's position by millimeters, allowing the system to know where the vehicle is with extreme accuracy.