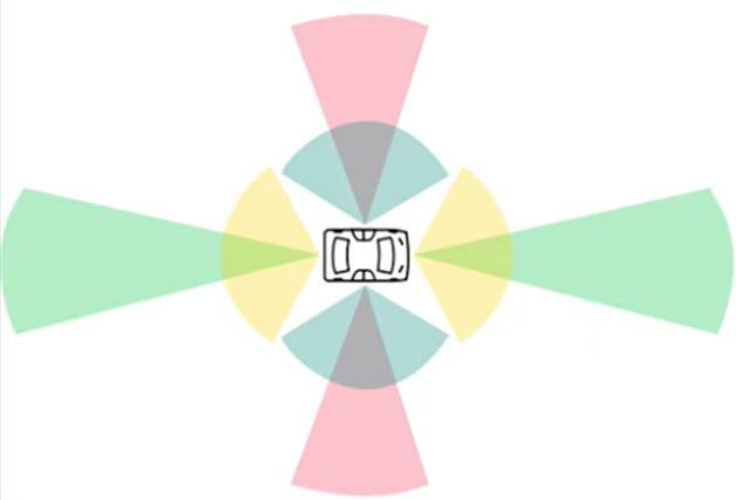
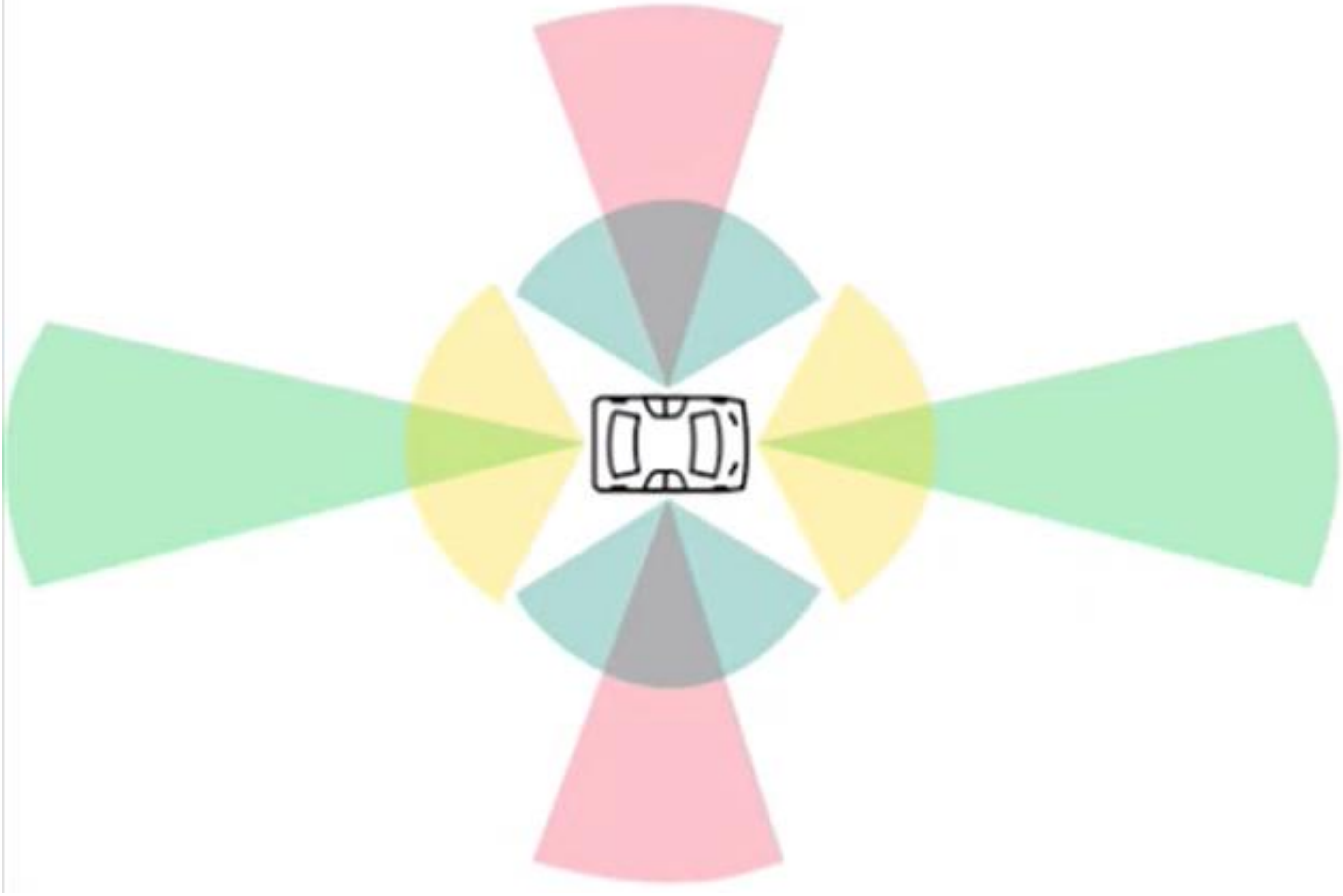
L1.2 AV Sensors, HD Maps

Zonghua Gu 2023

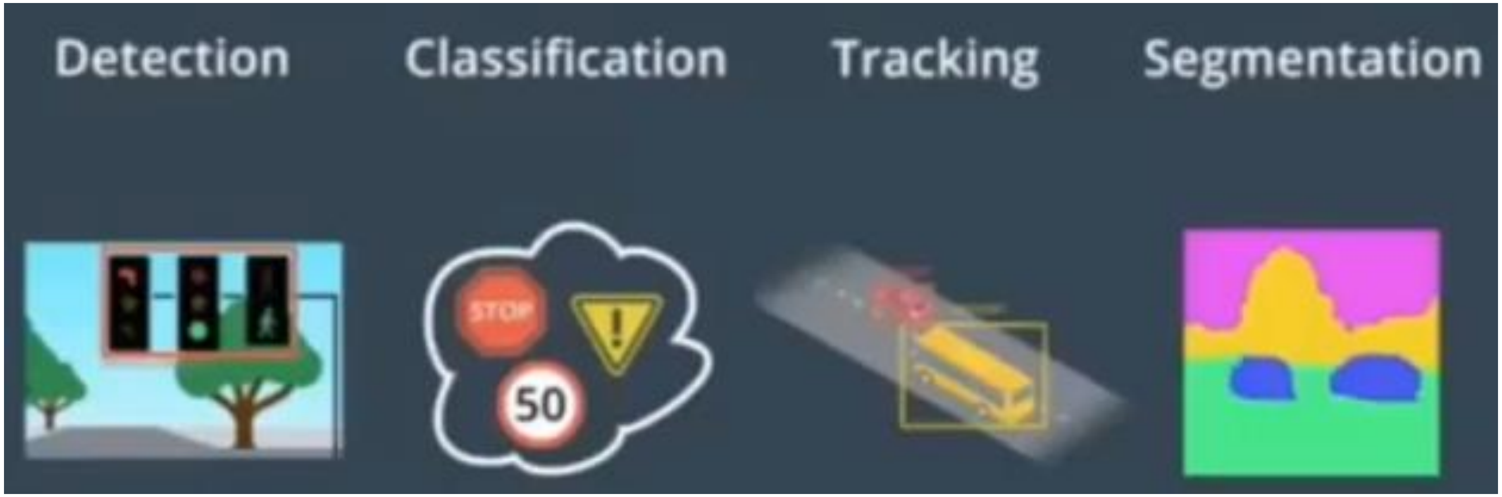


AV SENSORS

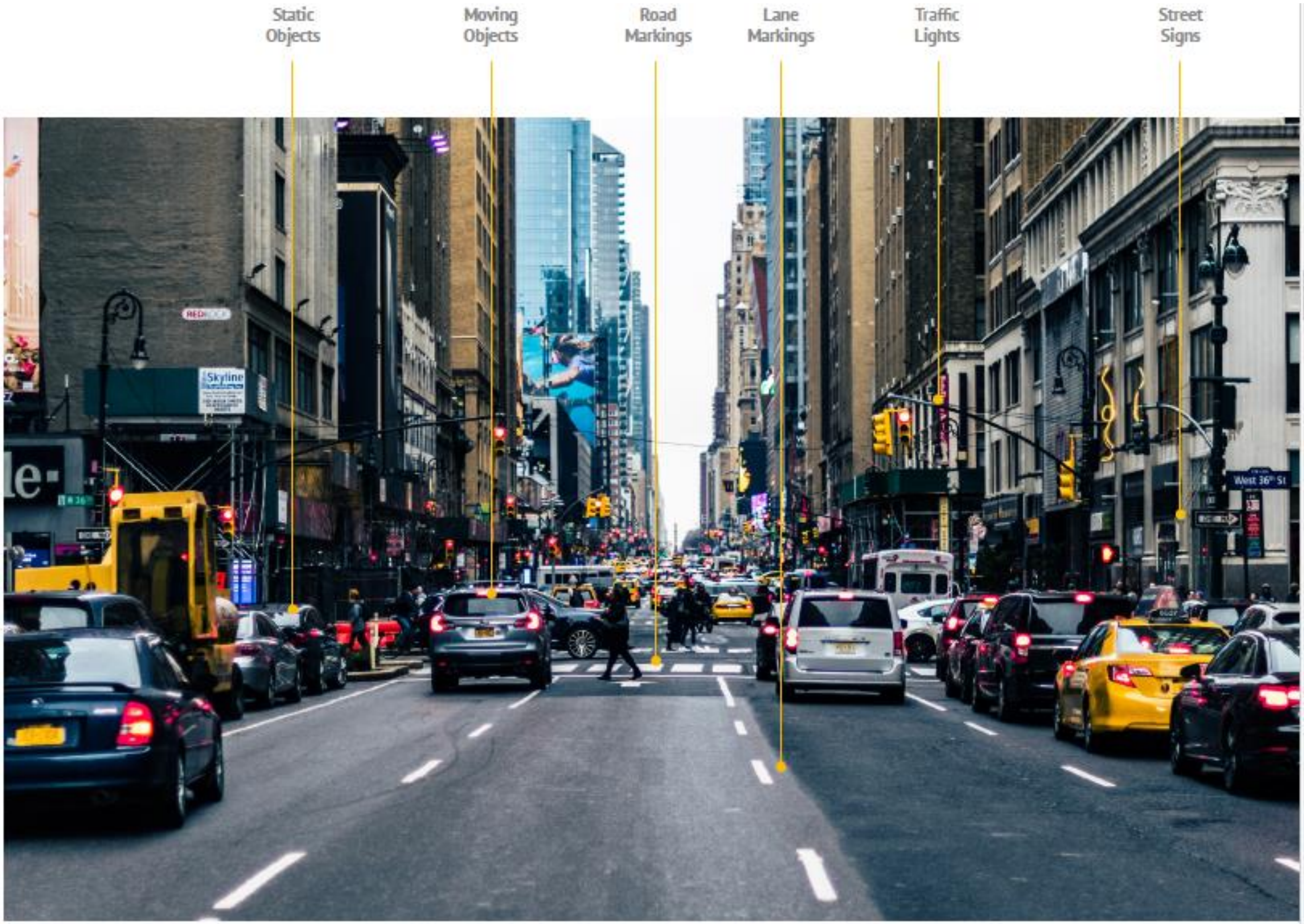


Perception Tasks

- 4 main perception tasks
  - Detection
    - Detect the existence of an object in the environment
  - Classification
    - Identify what the object is, e.g., traffic sign, traffic light, pedestrian
  - Tracking
    - Track a moving object across time
  - Segmentation
    - Semantic segmentation: classify each pixel to its semantic category, e.g., road, car, sky…
    - Instance segmentation: classify each pixel to an object instance, e.g., car1, car2…
- Mobileye's Autonomous Car What the System Sees
  - https://www.youtube.com/watch?v=jKfwHsHUdVc

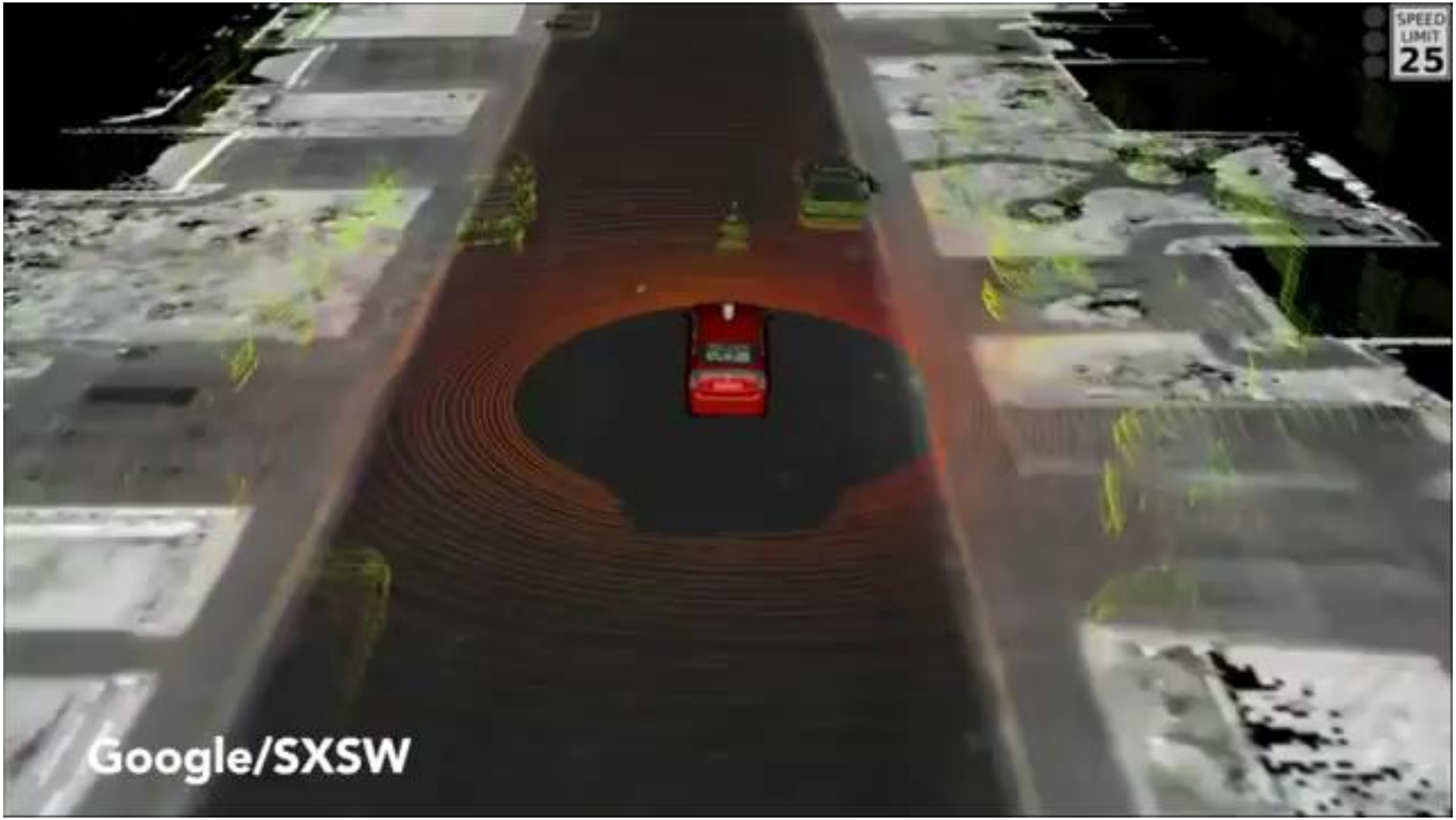


The variety of static and moving objects that an AV needs to detect and recognize

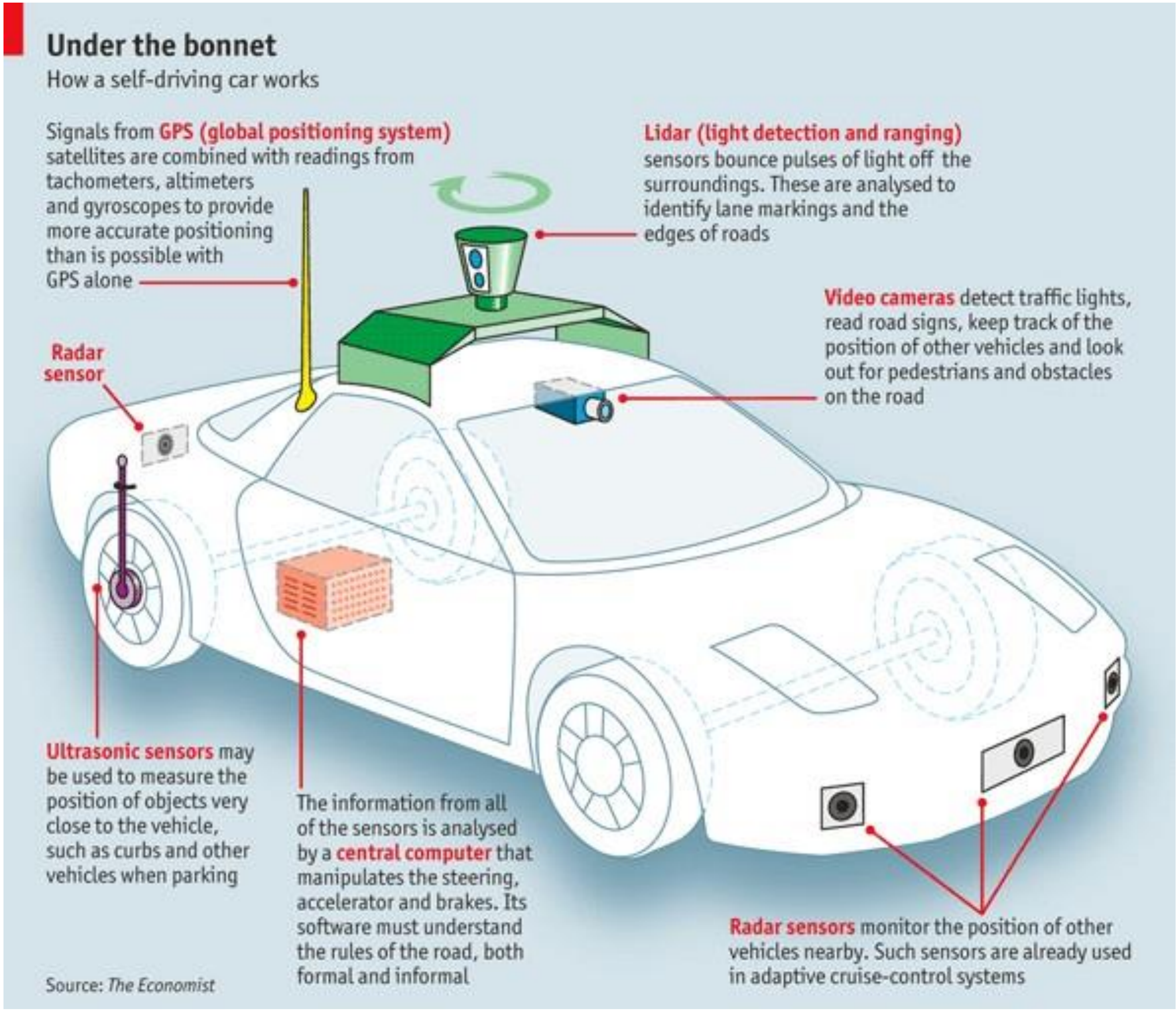


Perception is Challenging

- The long tail distribution is challenging: anything can happen on the road!
- Video from 2015, recorded by Google's AV.



The Typical AV Sensor Configuration



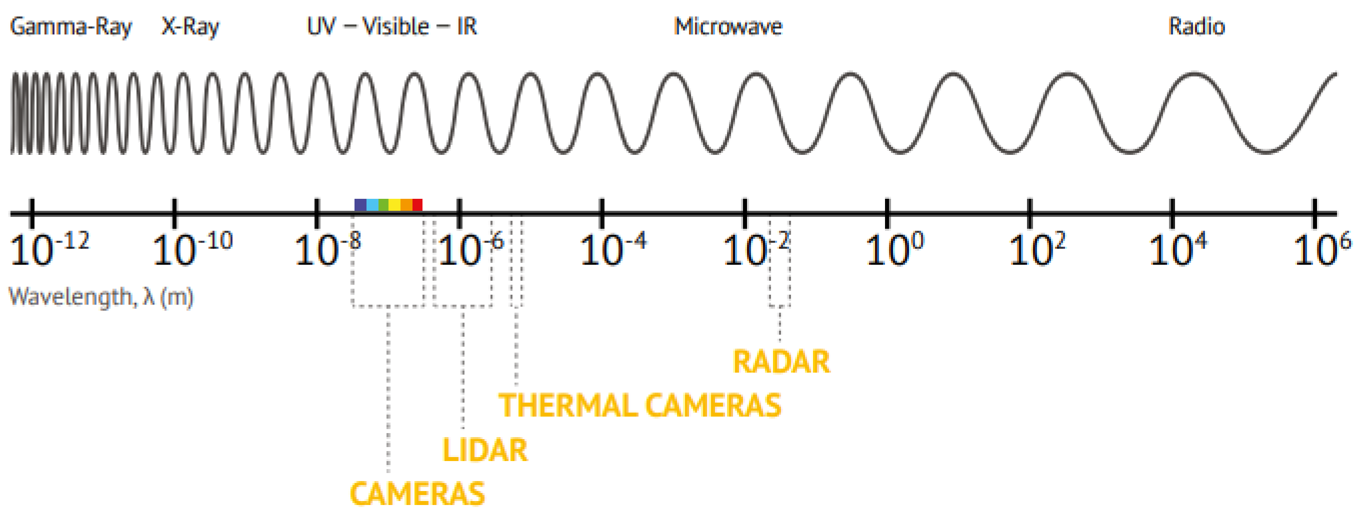
Configuration of Sensors of Some Research AVs

	Uber	Waymo	GM Cruise	Navya Autonomy Cab	Drive.ai	Nissan	Tesla Autopilot V9
Cameras	8	8	16	6	10	12	8
Lidars	1	6	5	10	4	6	0
Radars	4	4	8	4	2	9	1

- Tesla is one of the few AD companies that do not use Lidar.
- Elon Musk, 2017:
  - "Once you solve cameras for vision, autonomy is solved; if you don't solve vision, it's not solved … You can absolutely be superhuman with just cameras."
  - "In my view, Lidar is a crutch that will drive companies to a local maximum that they will find very hard to get out of. Perhaps I am wrong, and I will look like a fool. But I am quite certain that I am not."

Passive vs. Active Sensors

- Passive sensors detect existing energy, like light or radiation, reflecting from objects in the environment.
  - Cameras
- Active sensors (also called range sensors) send their own signal and sense its reflection
  - Lidar, Radar, ultrasound

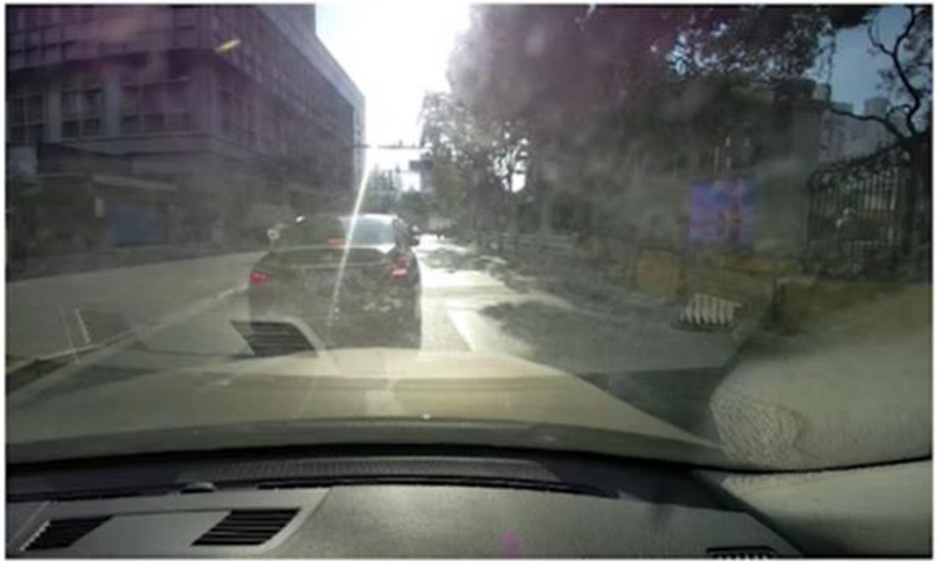


The electromagnetic spectrum and its usage for perception sensors .[16]

https://www.wevolver.com/article/2020.autonomous.vehicle.technology.report

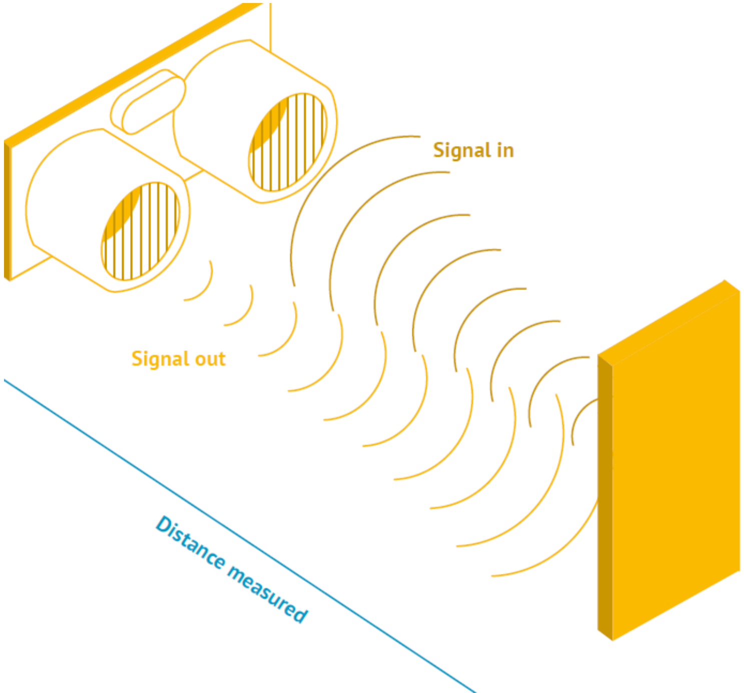
Camera

- Pro: cheap, versatile, stereo vision w. two cameras
- Con: easily affected by illumination conditions, needs additional light at night
- Key parameters
  - Resolution
    - e.g.,1080p HD cameras provide 1920x1080-pixel resolution, or 2.1 megapixels.
  - Field of View (FOV)
    - The extent of the observable world that is seen at any given moment
    - Given same resolution, wider FOV results in large image distortion.
  - Dynamic range
    - Maximum difference between the darkest and lightest pixel intensities in an image, measured in dB. An AV needs HDR (High-Dynamic-Range) cameras with at least 100dB.



Range Sensors

- They rely on Time of Flight (ToF) to measure distance (range), a key element for localization and environment modeling
  - Lidar uses electromagnetic waves.
  - Radar uses radio waves
  - Ultrasonic uses sound waves
- The traveled distance of a wave is given by $d = \frac{v*t}{2}$
  - d: distance
  - v: speed of wave propagation
  - t: ToF (roundtrip)

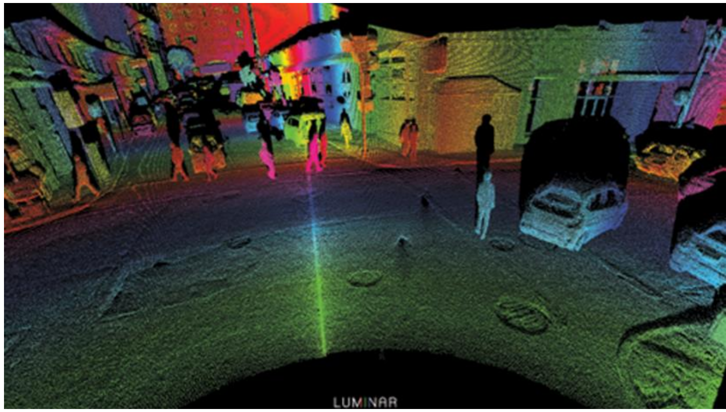


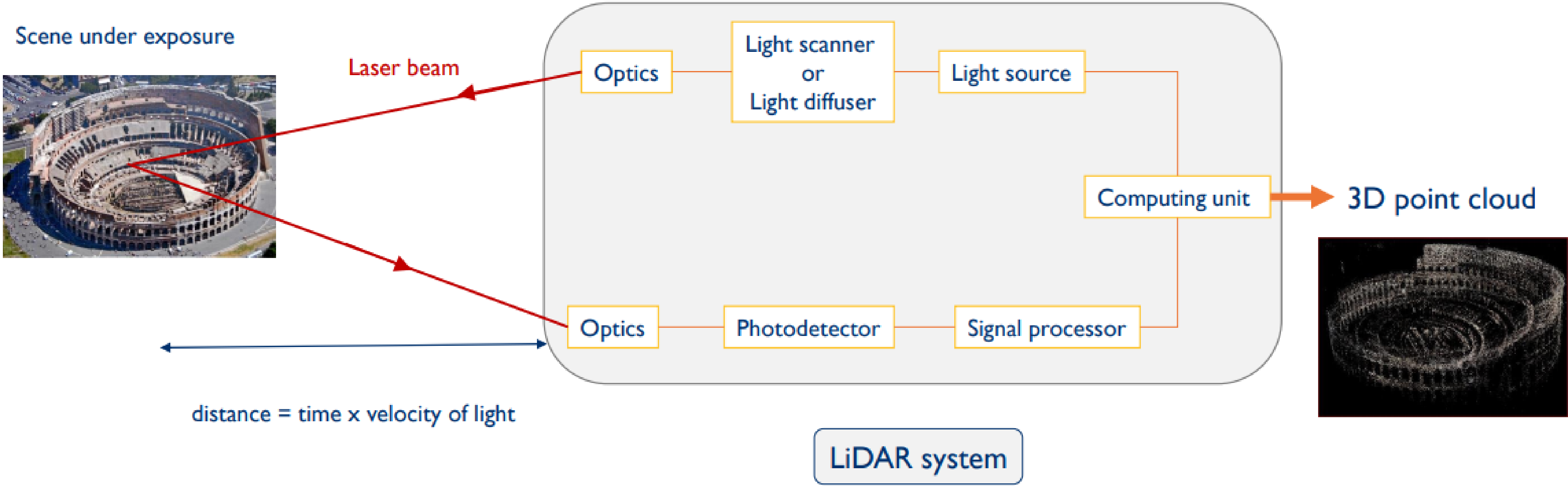
Range Sensors

- Propagation speed v
  - Sound: 0.3 m/ms
  - Electromagnetic wave (incl. light): 0.3 m/ns
    - 1 M times faster than sound
- To travel 3 meters:
  - 10 ms for ultrasonic sensor
  - 10 ns for Lidar
  - Measuring time of flight with electromagnetic signals is not an easy task. Hence Lidars are expensive and delicate
- The quality of range sensors mainly depends on:
  - Inaccuracies in the time of fight measurement (laser range sensors)
  - Opening angle of transmitted beam (especially ultrasonic range sensors)
  - Interaction with the target (surface, specular reflections)
  - Variation of propagation speed (sound)
  - Speed of vehicle and target

Lidar

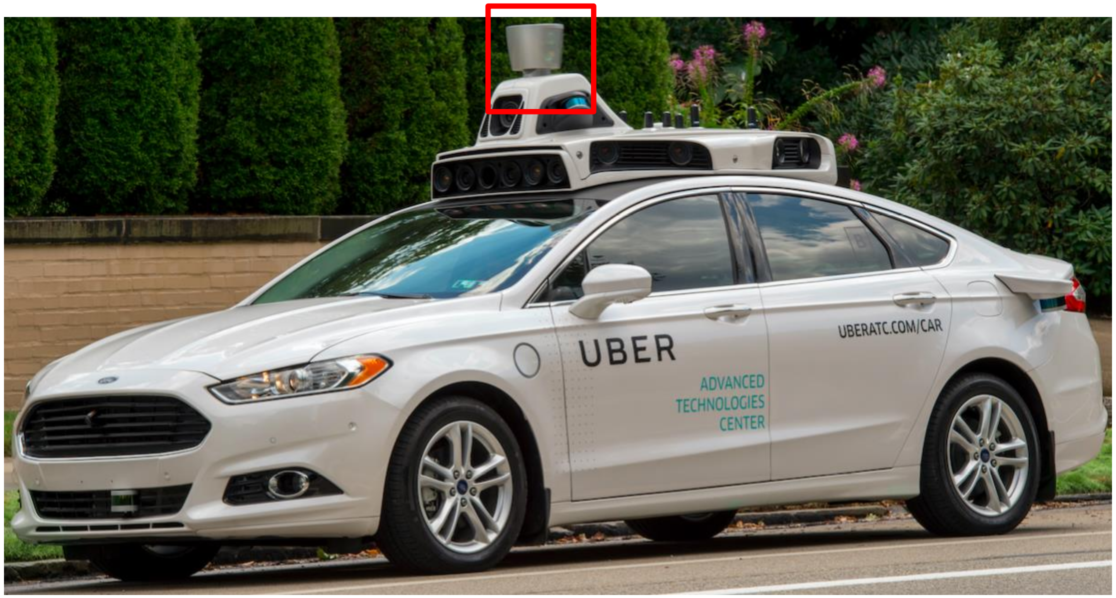
- Lidar (Light Detection and Ranging Device) sends millions of light pulses per second in a well-designed pattern to generate “Point Clouds” that describe the 3D geometry of the surrounding environment
- Pro: independent of lighting conditions, precise distance measurements for 3D perception
- Con: expensive, medium resolution
- Key parameters:
  - Laser beam count
  - Rotation Speed
  - FOV
  - Range distance (from tens to hundreds of meters)



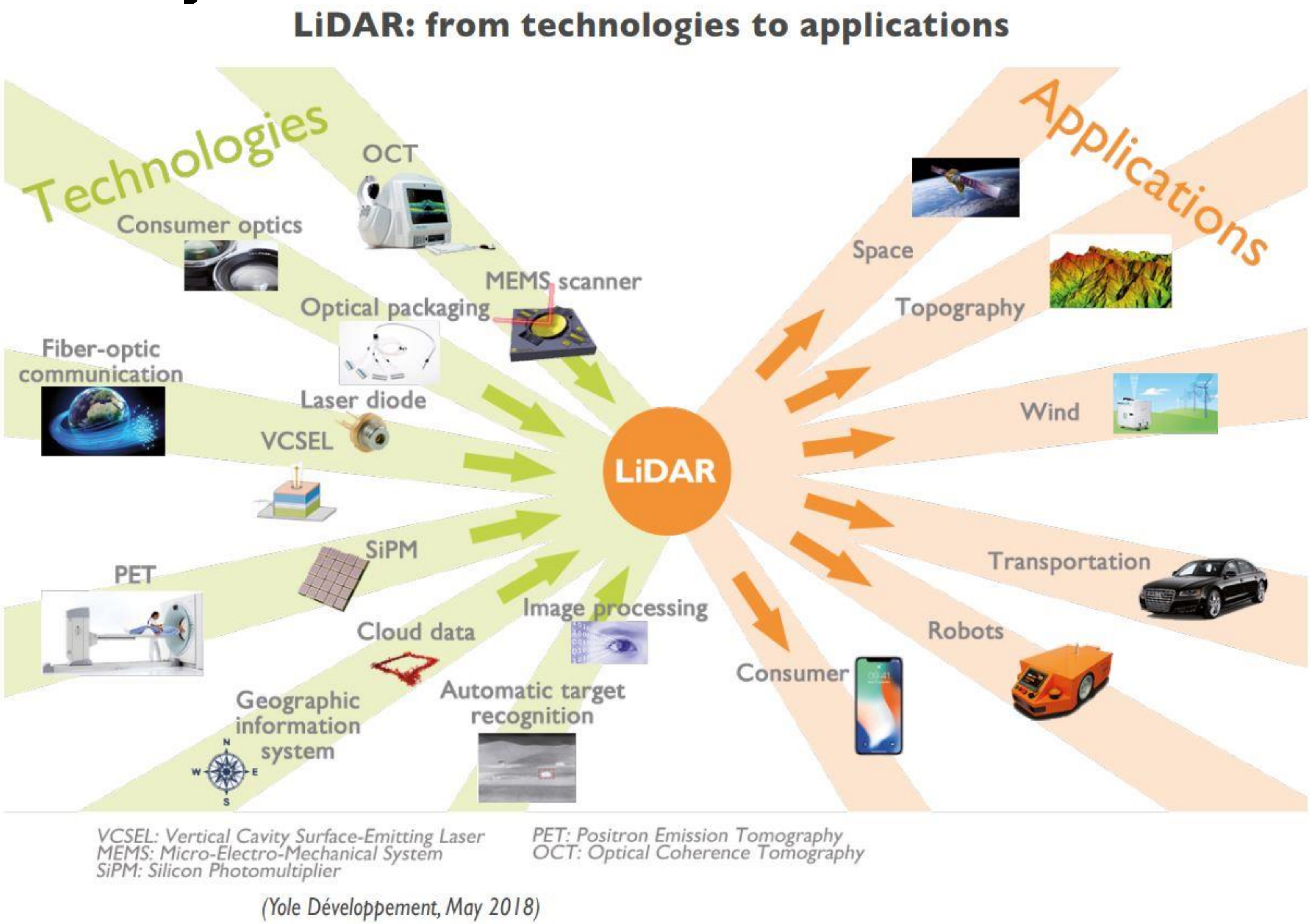


Velodyne Lidar

- The high-end Velodyne HDL-64E with 64 laser emitters
  - Rotation rate up to 15 Hz
  - FOV is 360° horizontally and 26.8° vertically
  - Angular resolution is 0.09° and 0.4° respectively
  - Delivers ~1.3M data points per second
  - Expensive (~USD$40-80K) (cheaper versions available)



Many Variants of Lidars in the Market



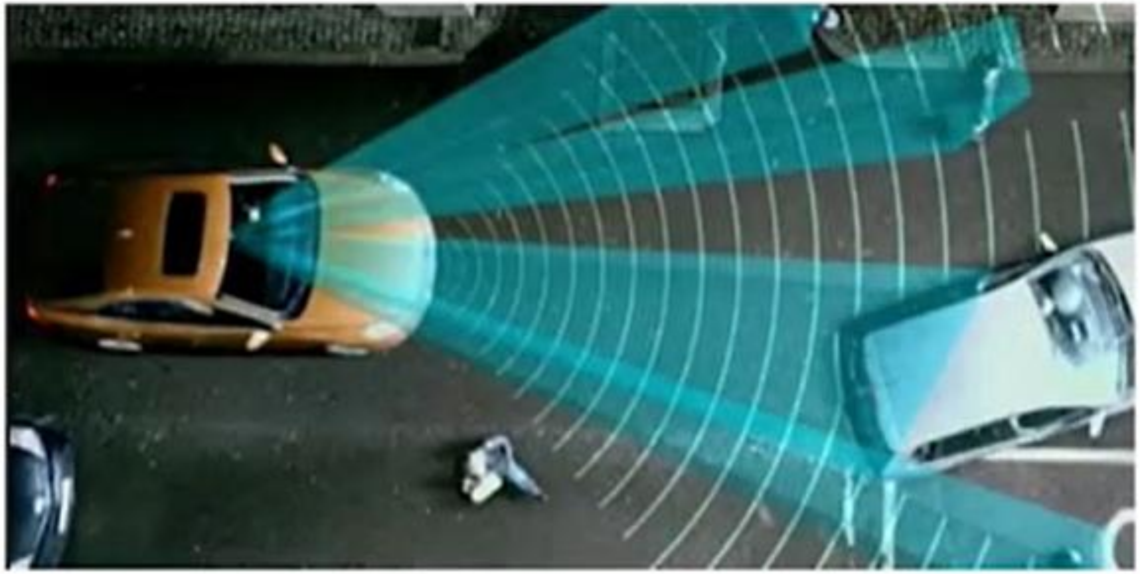
http://www.f4news.com/2018/05/04/yole-on-lidar-market/

Apple iPAD Pro has Built-in Lidar (2020)

- It allows users to scan a depth-accurate depiction of the environment.
- Main application: Augmented Reality (AR)
  – Needs depth information to place virtual objects in the environment.

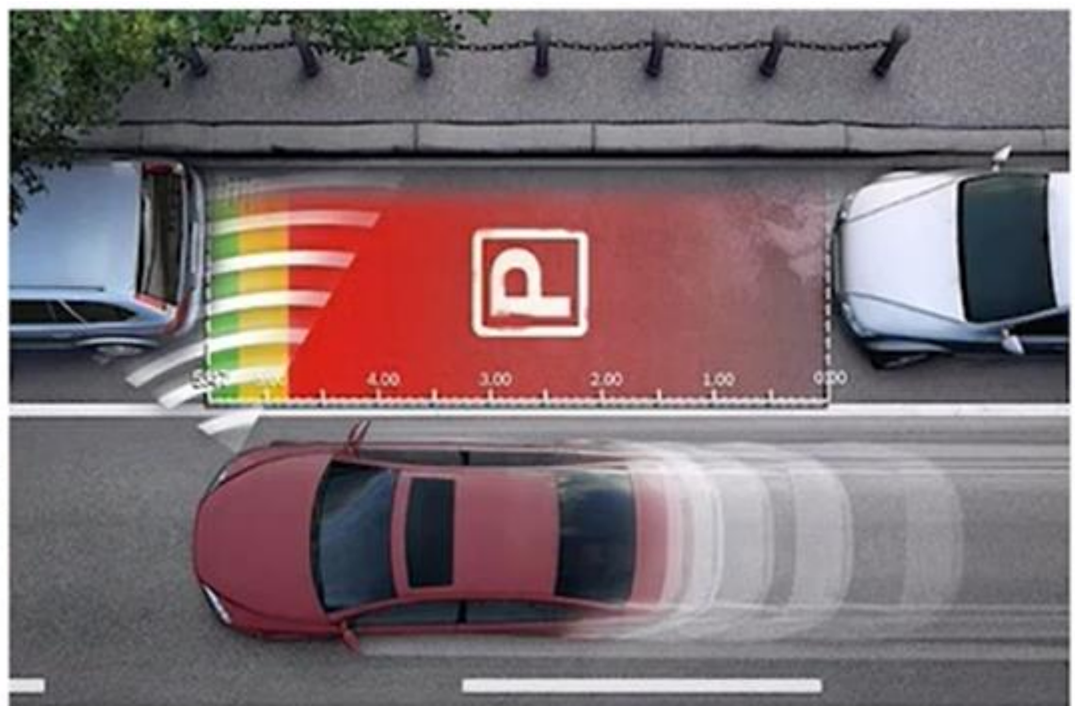
mmWave Radar

- Pro: provides both position and relative speed information; can operate in varied conditions (low-lighting, rain, fog…)
- Con: low resolution
- Key parameters
  - Sensing distance, FOV, Position and velocity accuracy
- Two types
  - Short-medium range with wide FOV
  - Long range with narrow FOV

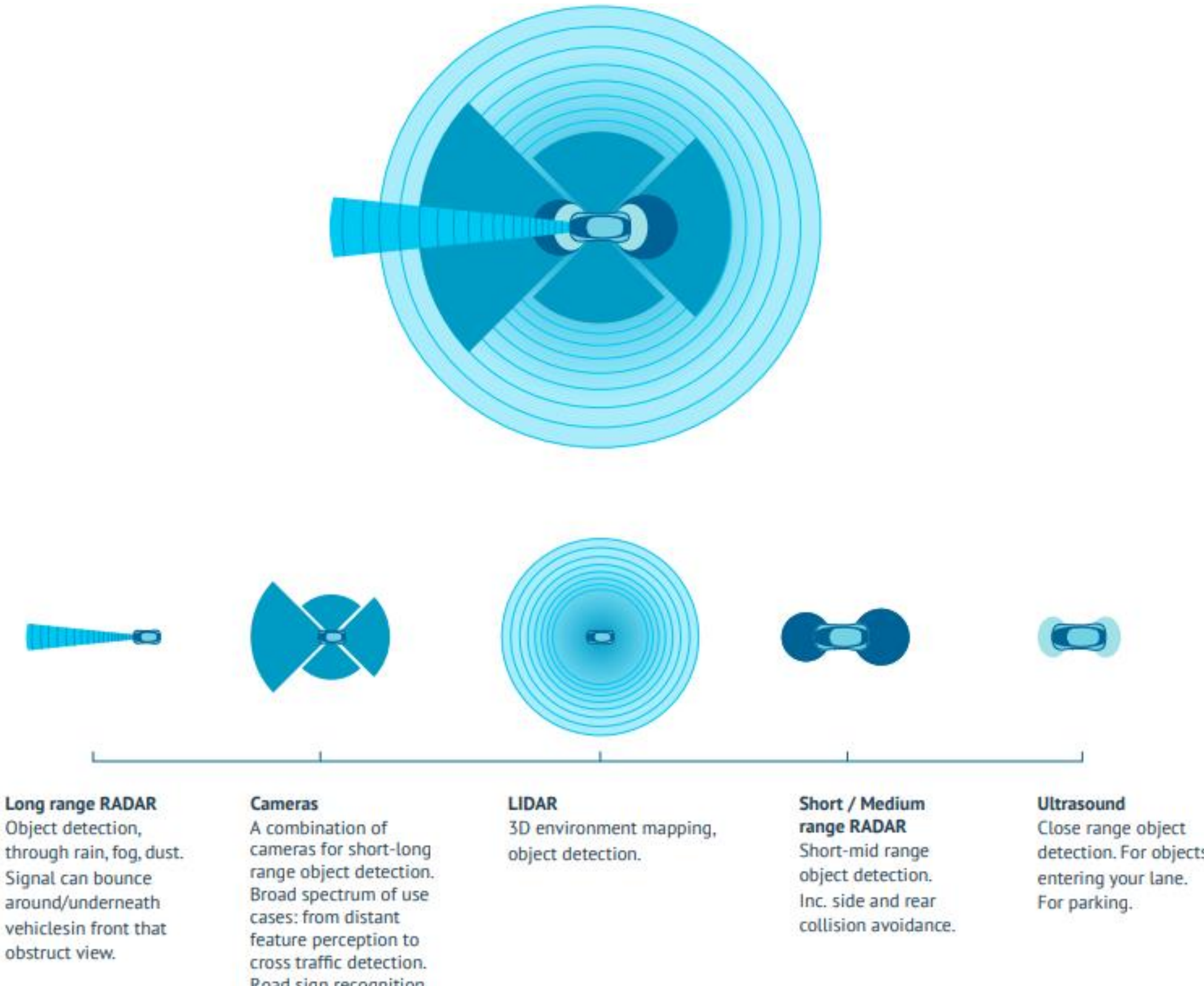


Ultrasound

- Pro: not affected by lighting conditions, rain or fog
- Con: short sensing range (mainly used for parking assistance)
- Key parameters
  - Sensing range
  - FOV



Comparison of Sensing Ranges



https://www.wevolver.com/article/2020.autonomous.vehicle.technology.report

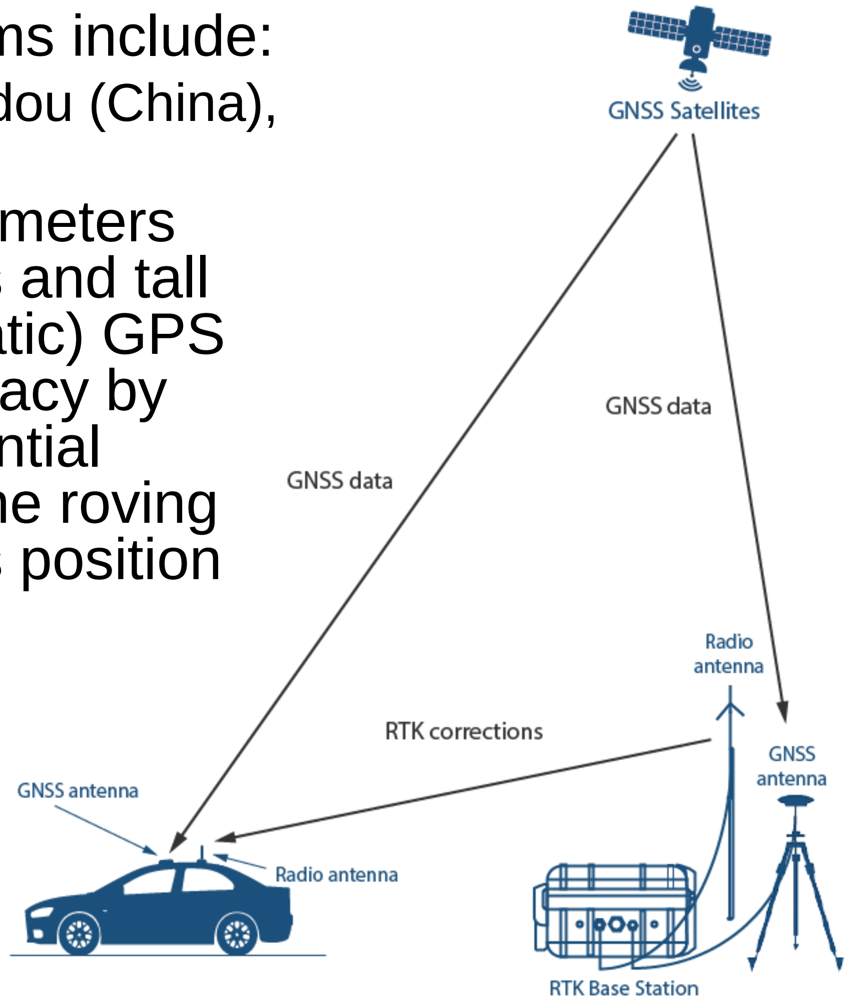
Other Comparisons

- Each sensor has its strengths and weaknesses
- Sensor fusion crucial for robust perception

	Camera	**Lidar**	**Radar**
Sensing Range	Mixed	Mixed	Good
Functioning in bad weather	Poor	Mixed	Good
Functioning in poor lighting	Mixed	Good	Good
Object Detection	Mixed	Good	Mixed
Object Classification	Good	Mixed	Poor
Lane Tracking	Good	Poor	Poor

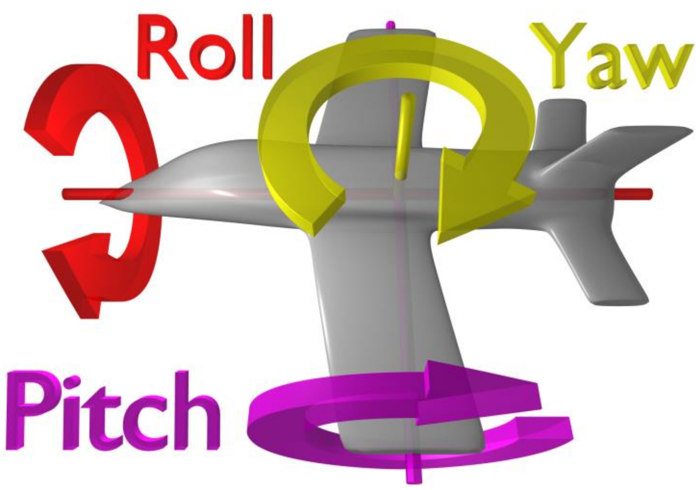
GNSS/GPS

- GNSS (Global Navigation Satellite System) provides localization service for outdoor applications. Current GNSS systems include:
  - GPS (USA), Galileo (Europe), Beidou (China), GLONASS (Russia)
- Conventional GPS provides a few meters accuracy, affected by cloud covers and tall buildings; RTK (Real-Time Kinematic) GPS can provide centimeter-level accuracy by calculating and transmitting differential correction data via radio to allow the roving GPS system (vehicle) to correct its position



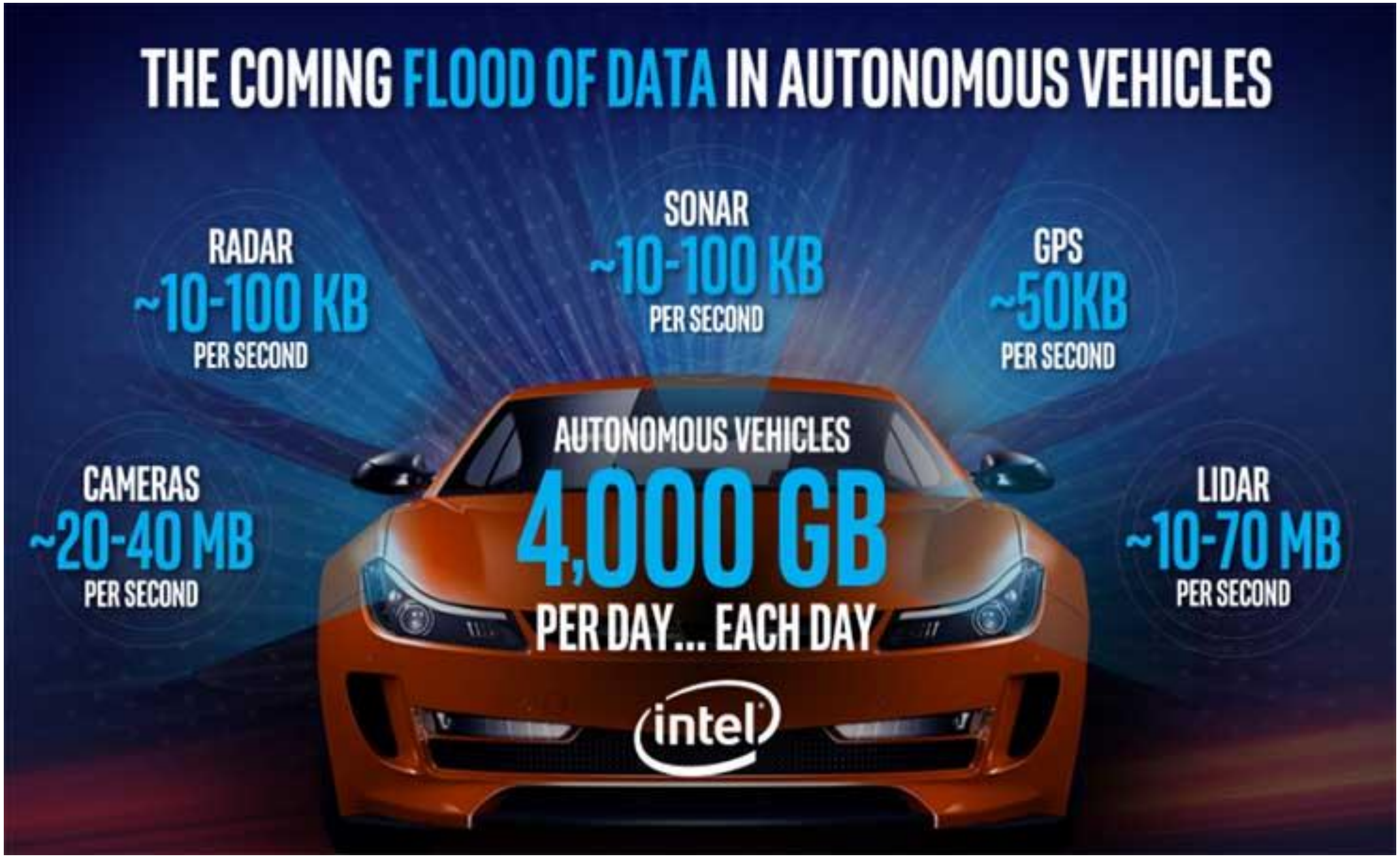
IMU

- IMU (Inertial Measurement System) measures acceleration (linear and angular) and orientation (yaw, pitch, roll)
  - For ground vehicles on 2D plane, only yaw is relevant.
  - For aerial vehicles in 3D space, all three are relevant
- Often combined with GNSS/GPS to form INS (Integrated Navigation System), using sensor fusion to achieve higher estimation accuracy



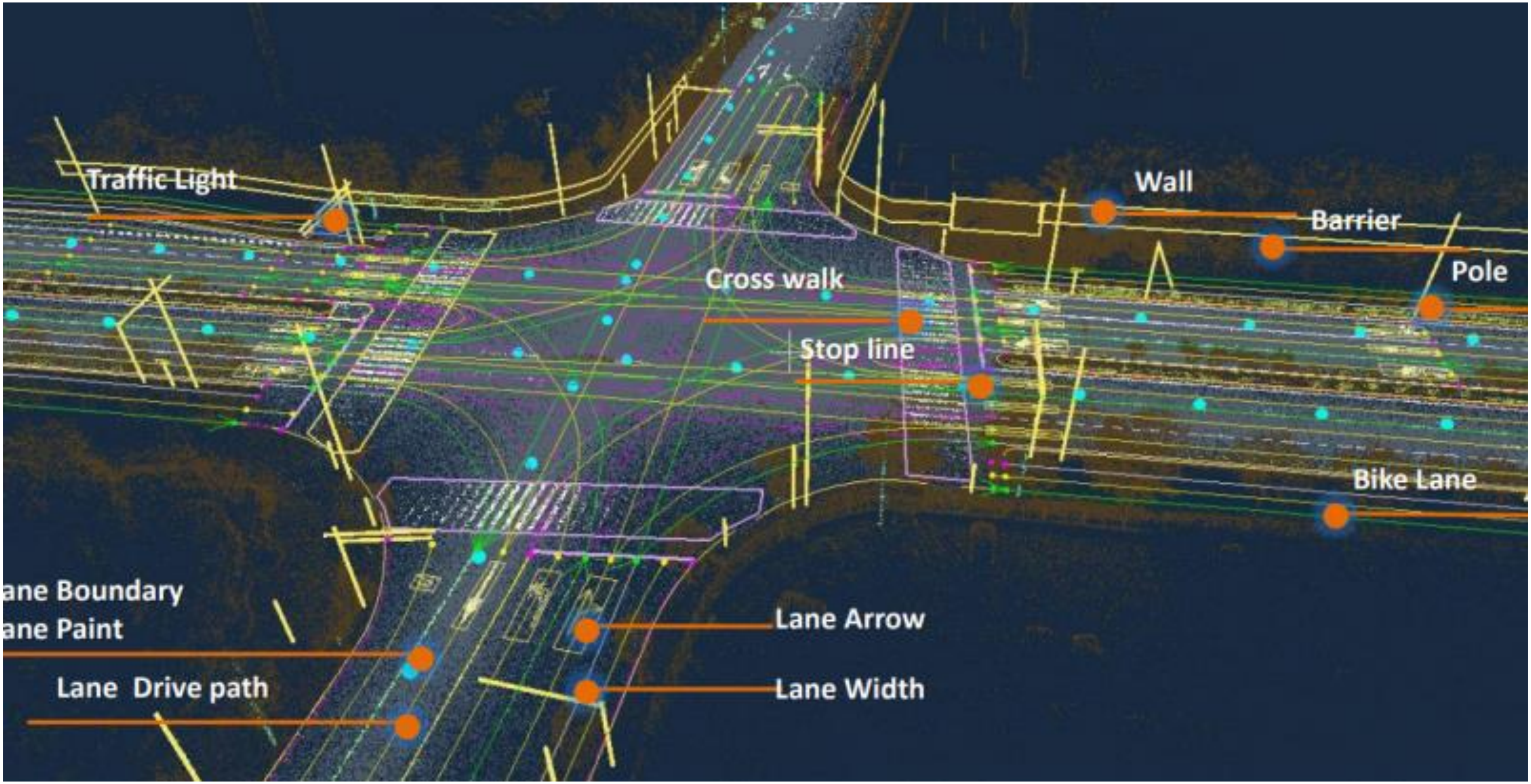
AV Sensors Generates Big Data

- Sensors, esp. cameras and lidars, generate the most amount of data
- Sensor data must be processed in real-time by perception algorithms



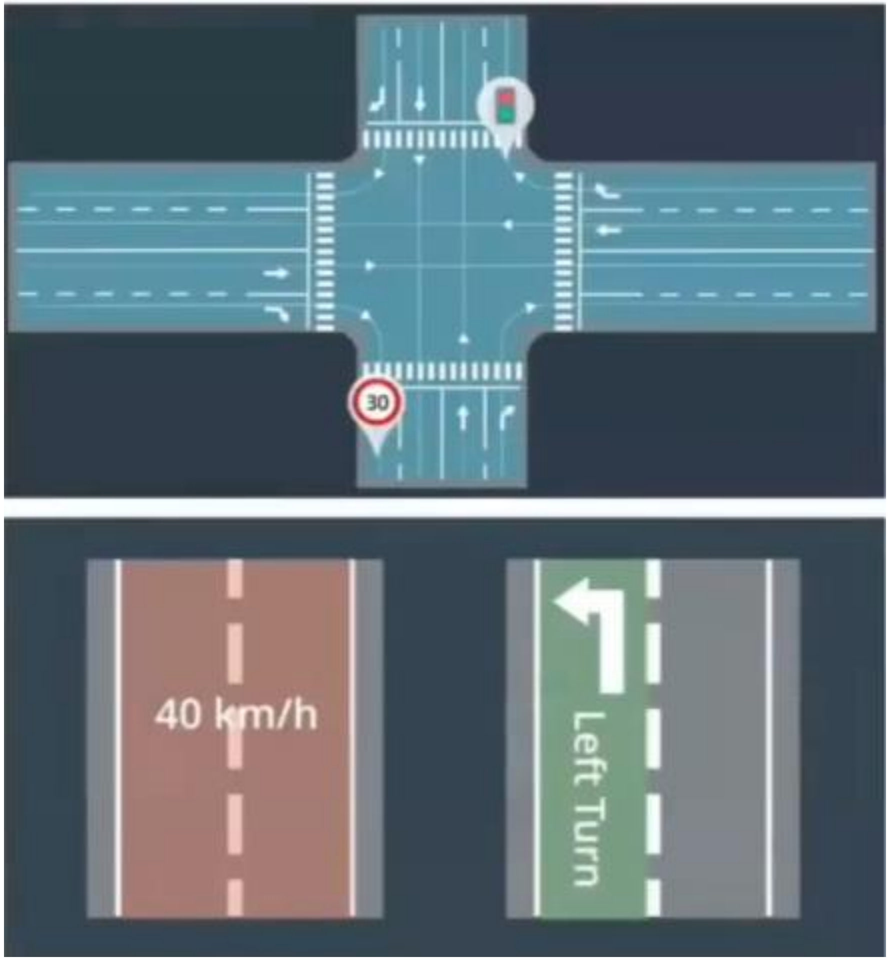
https://datacenterfrontier.com/autonomous-cars-could-drive-a-deluge-of-data-center-demand/

High-Definition (HD) Maps



HD Maps

- Different from navigation maps (e.g., Google Maps) designed for human eyes, HD maps are designed for processing by computers
  - Highly-accurate (centimeter-level) 3D representation of the road network, e.g., cross section layout; locations of traffic lights/signs; semantic information on certain road segments (speed limits…)
- Benefits
  - Help reduce Region-of-Interest (ROI) for detection of traffic elements
  - Help with AV localization based on known object positions
  - help with recognition of lane center line



Are HD Maps Necessary?

- Humans don’t need HD maps to drive
- HD maps vs. on-board sensing
  - Use of HD maps limits the area of operation; mapless systems allow universal operation anywhere.
  - Use of HD maps reduces the computational burden on on-board sensing: the more that has already been mapped out, the easier it is for the on-board system to focus on the moving parts; mapless systems must figure out everything on-the-fly with no prior knowledge of the environment
- It is generally agreed that L4-L5 levels of automation cannot work without HD maps, at least for now; L2-L3 levels may work without them (e.g., Tesla does not use HD maps)