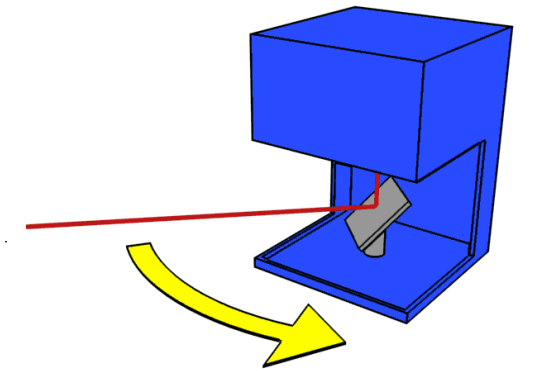
ELTE | FACULTY OF INFORMATICS

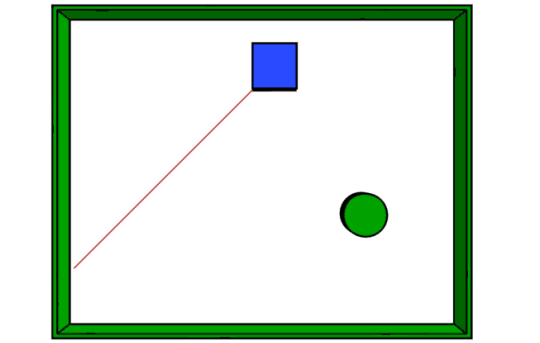
3D Point Cloud processing and analysis
Course_01
Introduction

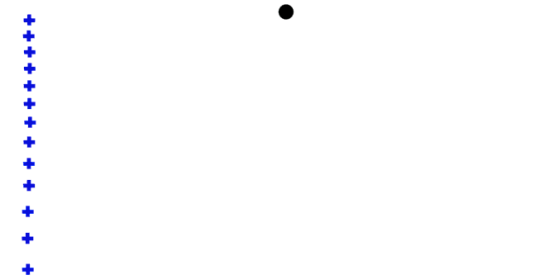
Massinissa Aouragh:
Faculty of Informatics, Department of Artificial Intelligence
Robert Bosch Kft
m2j7au@inf.elte.hu

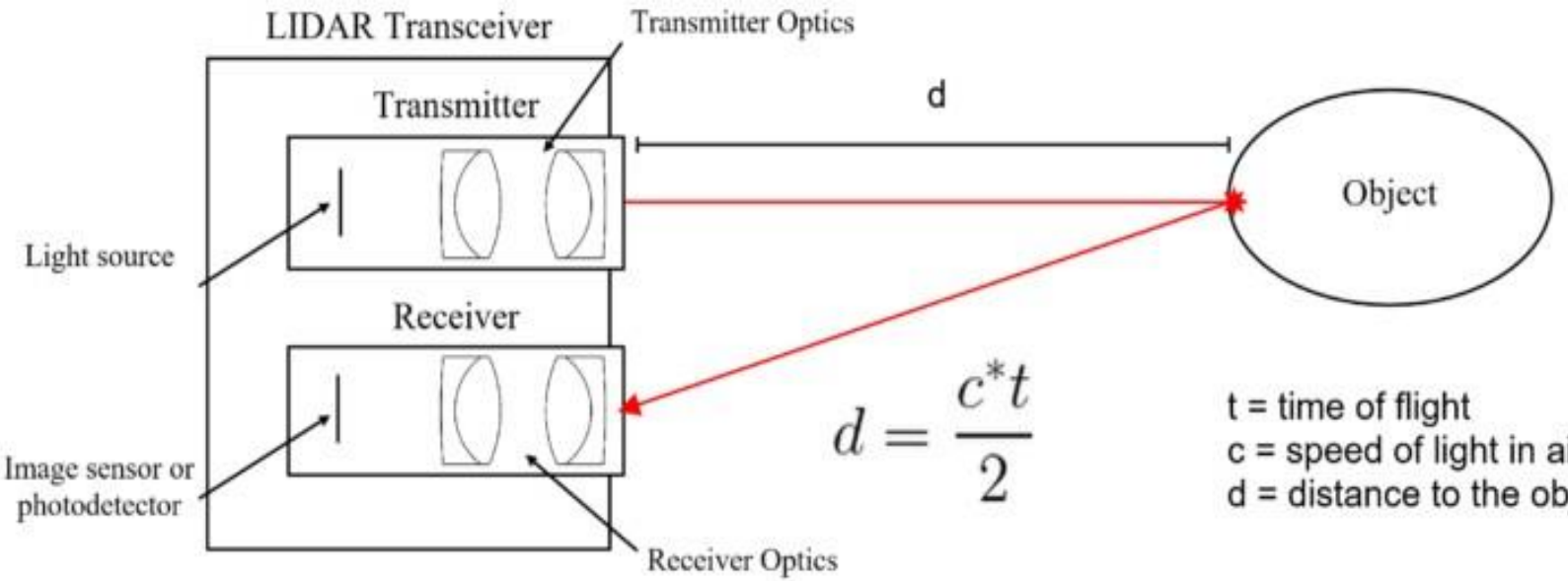
LiDAR (light detection and ranging)

- Active sensor
- Uses EM waves in the optical and infrared wavelengths
- Shorter wavelengths compared to radar
- Better angular resolution than radar but won’t see through fog or clouds








LIDAR Transceiver
Transmitter Optics
Transmitter
d
Object
Light source
Receiver
Image sensor or photodetector
Receiver Optics
$d = \frac{c^*t}{2}$
t = time of flight
c = speed of light in air
d = distance to the object


HDL-64E S2


Velodyne
VELARRAY
Velodyne
VELARRAY

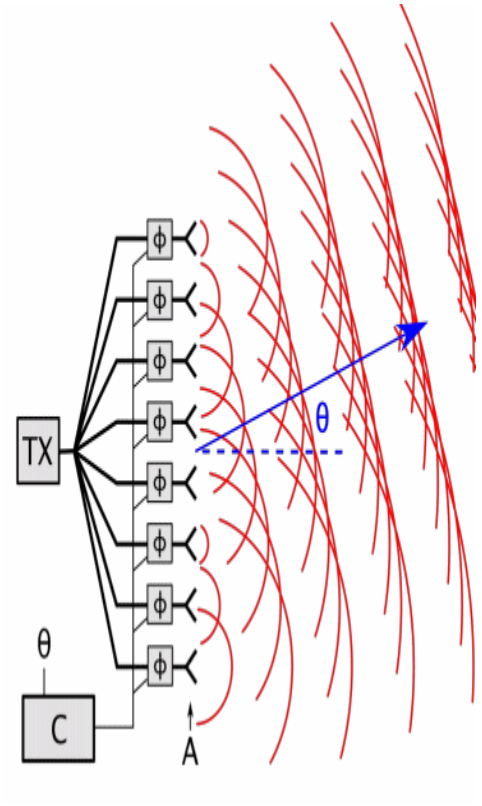
LiDAR types

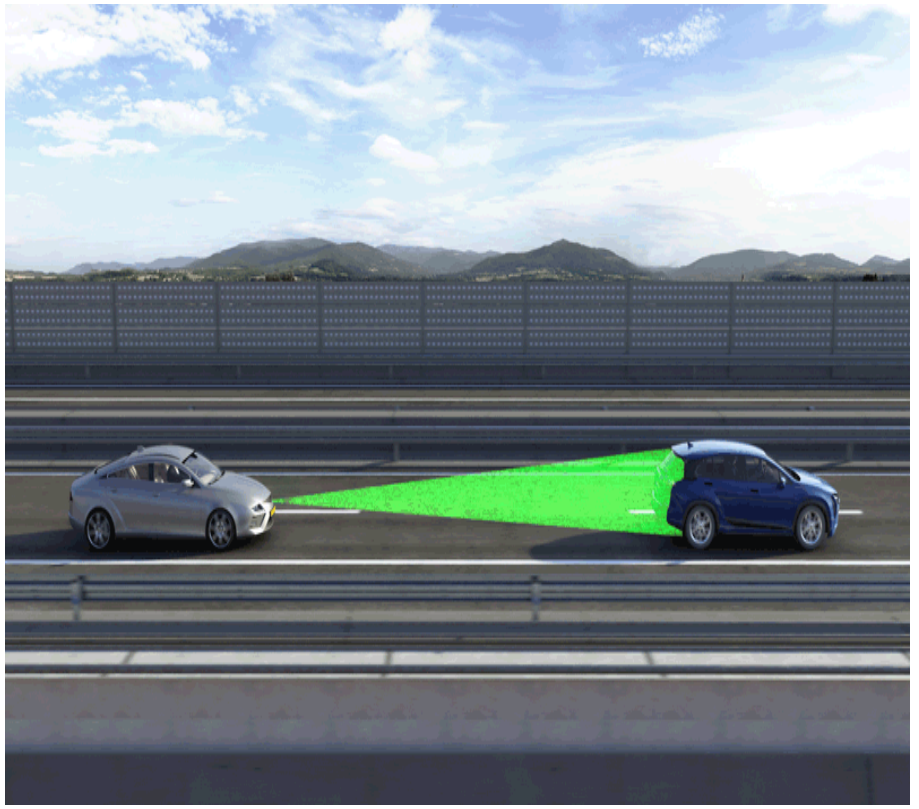
- **Mechanical scanning**: Set of lasers and detectors rotating mechanically at a frequency from 10 to 30 Hz.
- **Solid state LiDAR**: No mechanical rotating, different way to cover the entire scene:
  - MEMS (Micro-electromechanical system): moving mirrors
  - OPA (Optical phased array): Deliver light pauses to different directions. Direction is performed by adjusting the phase of the light

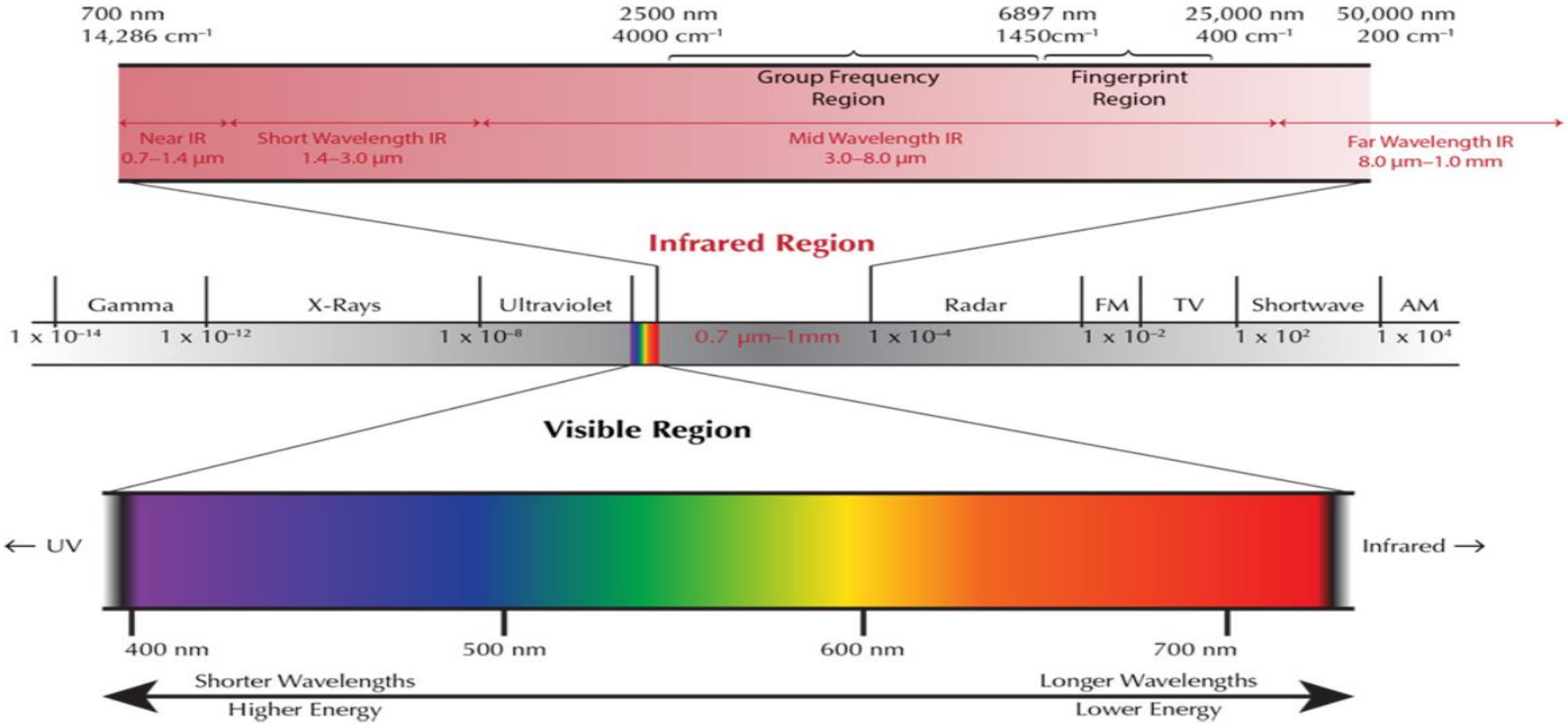
LiDAR types

- **Flash LiDAR:**
  - Operates like a flashing camera
  - Illuminates the scene by expending a laser beam
  - Generates the point cloud in one flash

LiDAR types







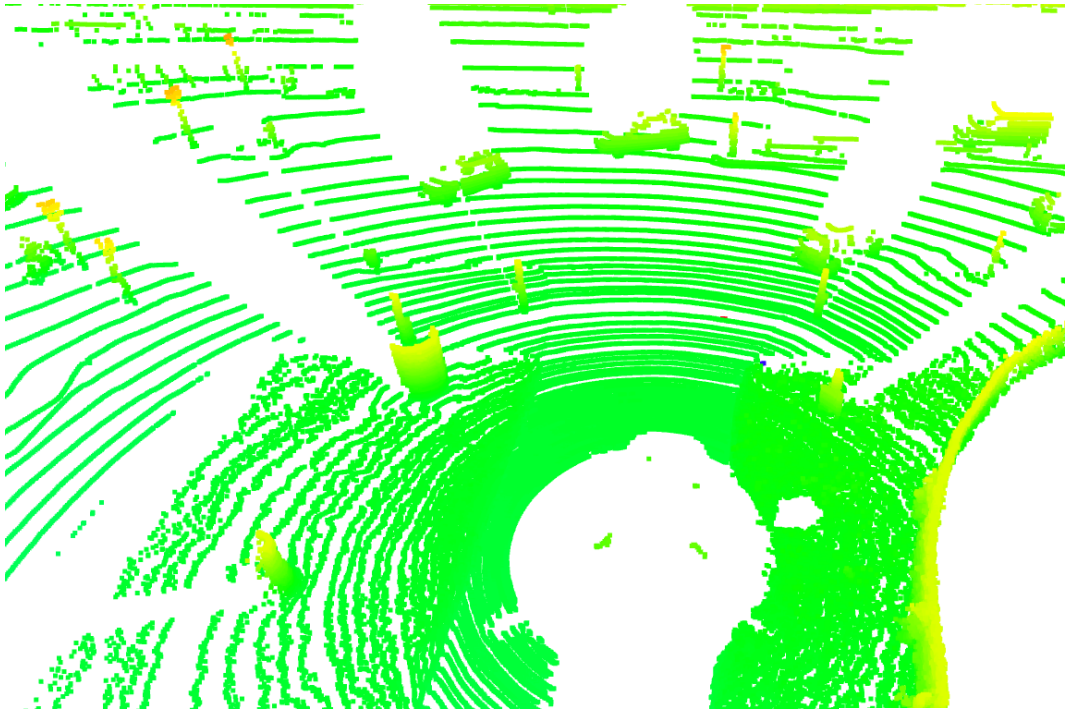
The electromagnetic spectrum, highlighting the visible and infrared regions. Illustration: Richard W. Hughes

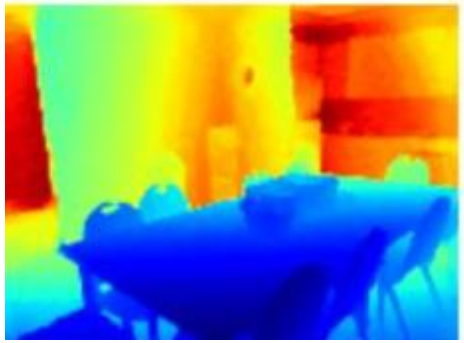
LiDAR terminology

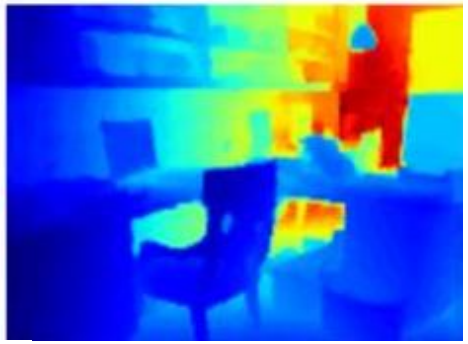
- **Pulse rate**: The rate at which the LiDAR is pulsing (KHz)
- **Scan frequency**: the speed of the scanner's oscillation
- **Scan angle or field of view**: distance to scan from one end to the other
- **Return number**: the number of returns coming from a single laser pulse.
- **Intensity Data:** The strength of the of the recorded returned signal. Show the reflectiveness of the object

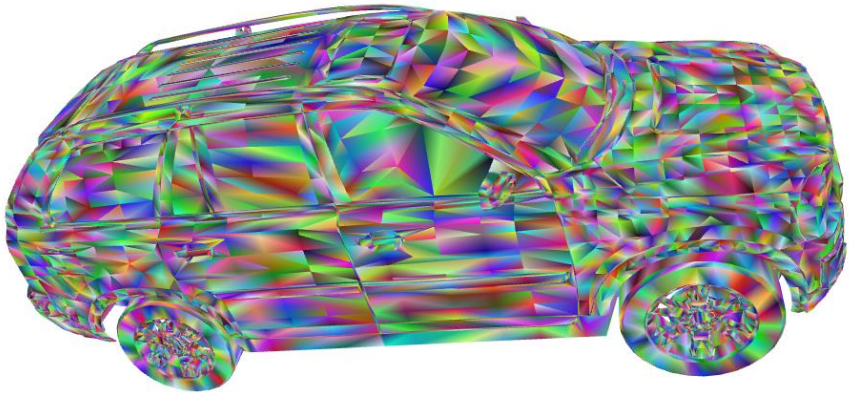
3D Point Clouds

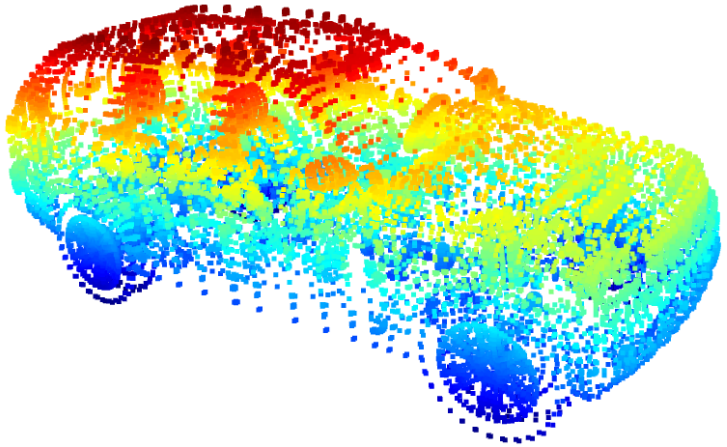
- Set of points in three-dimensional space
- Represent the surface of the scanned object
- Additional to XYZ point cloud can embed RGB and surface normal
- Unlike 2D images 3D point clouds are unorganized with no order
- Commonly acquired with LiDAR (light detection and ranging)





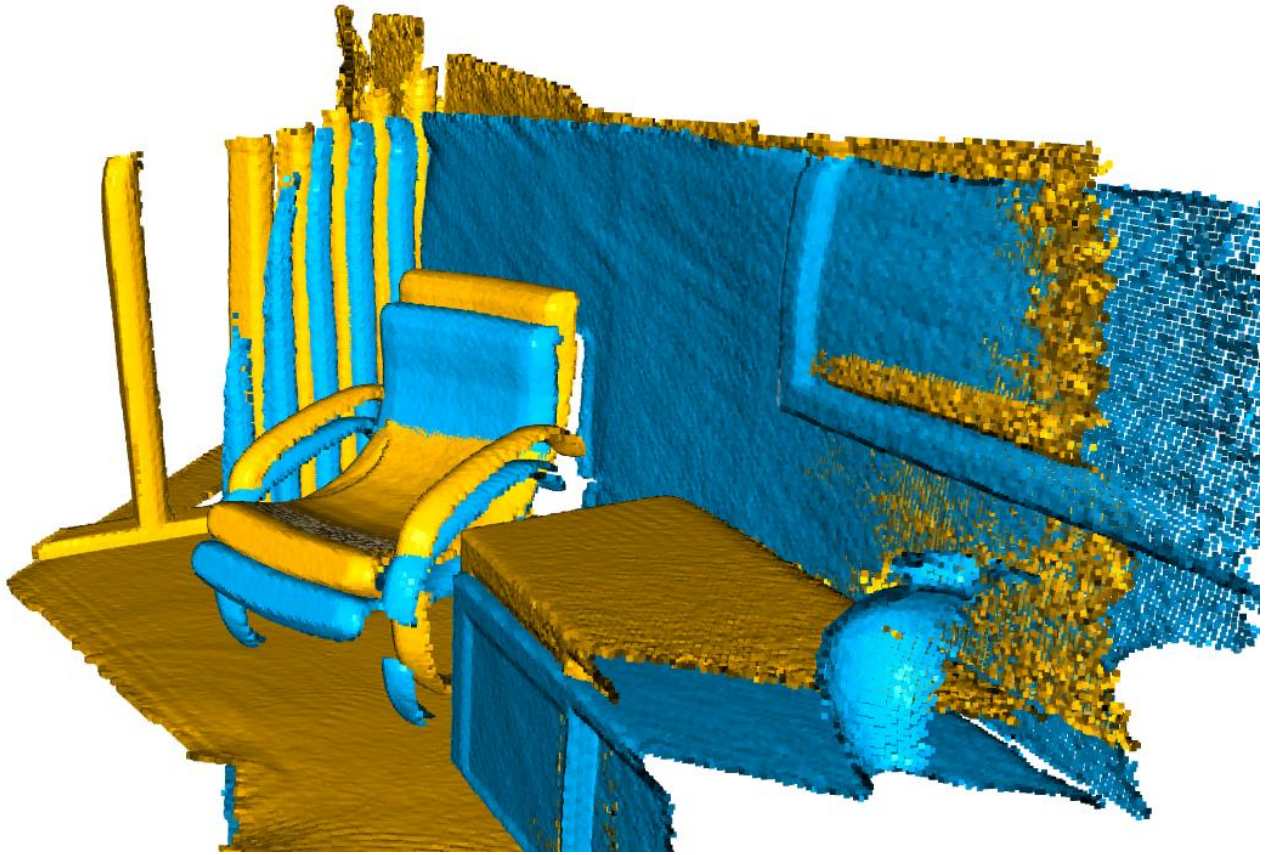


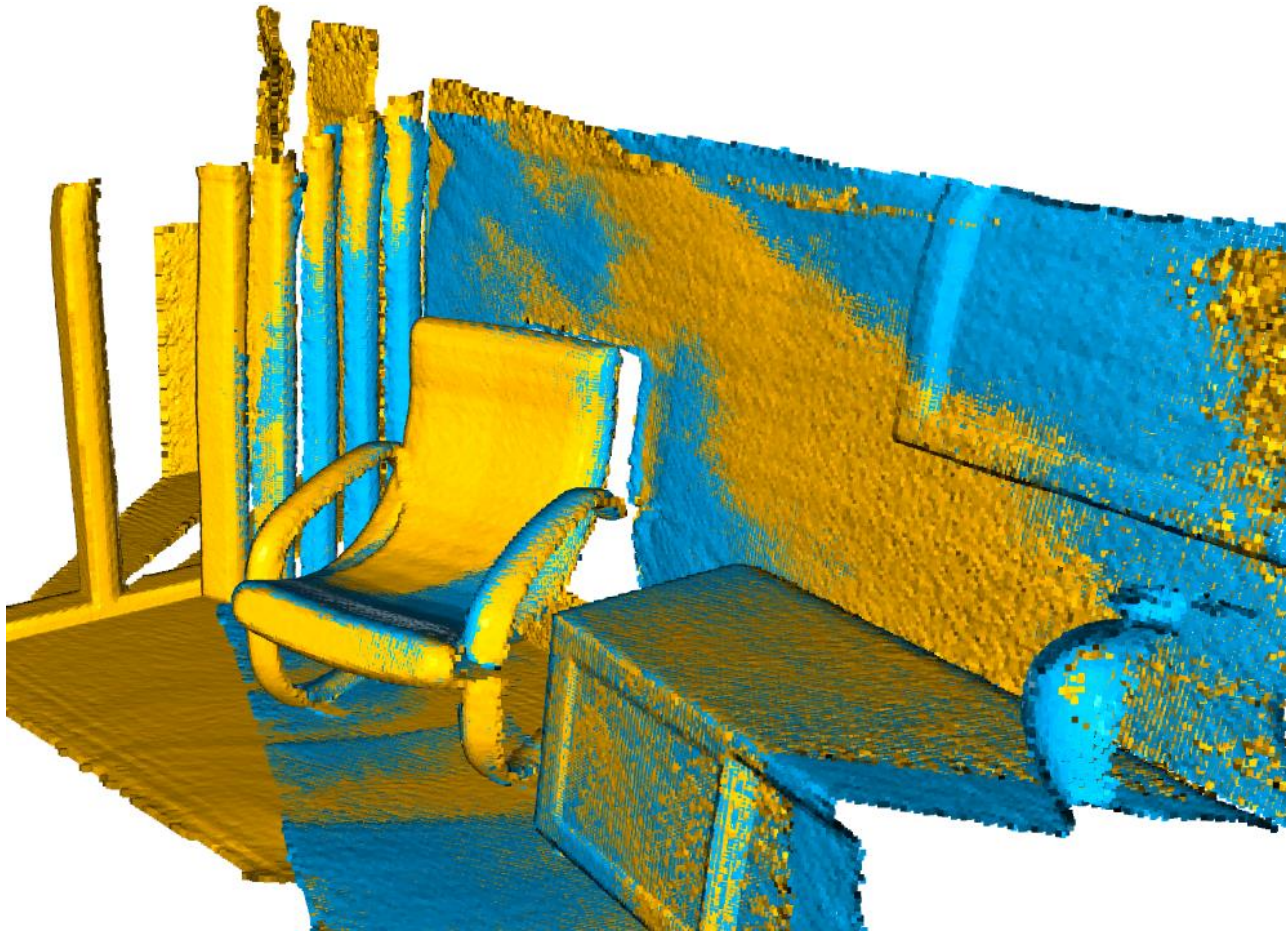




Point Cloud processing

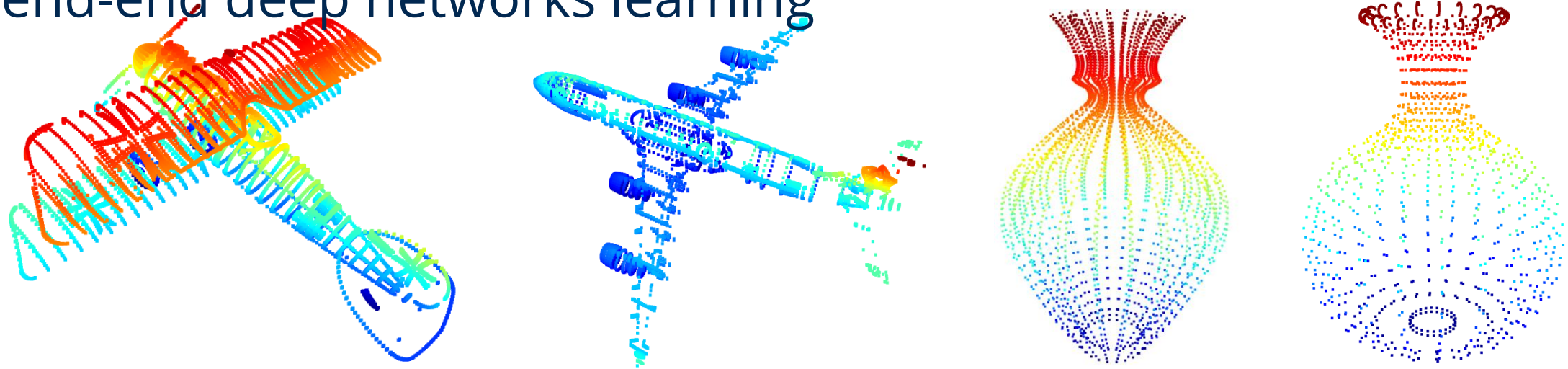
- **Registration:**
  3D alignment finding a spatial transformation that optimally aligns two-point clouds. Spatial transformation can be further divided:
  - Rigid body transformation: distance preserving between any two points (rotation and translation)
  - Non-rigid transformation: introducing different forms of deformations (scaling, perspective, affine)





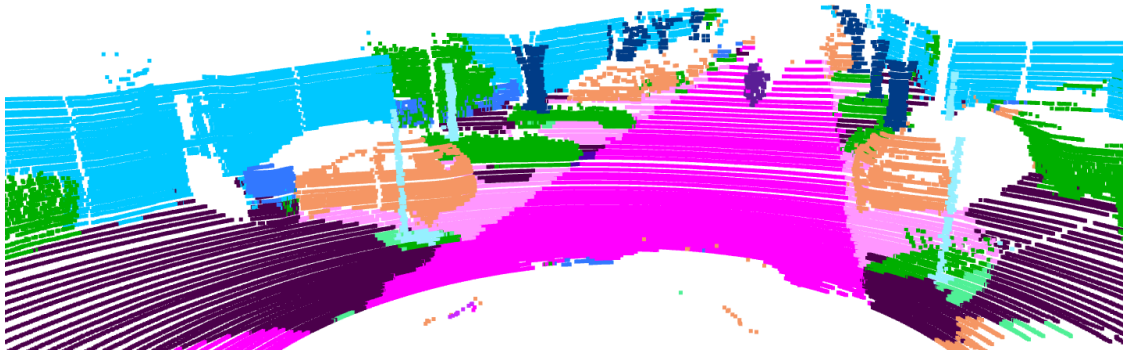
Classification

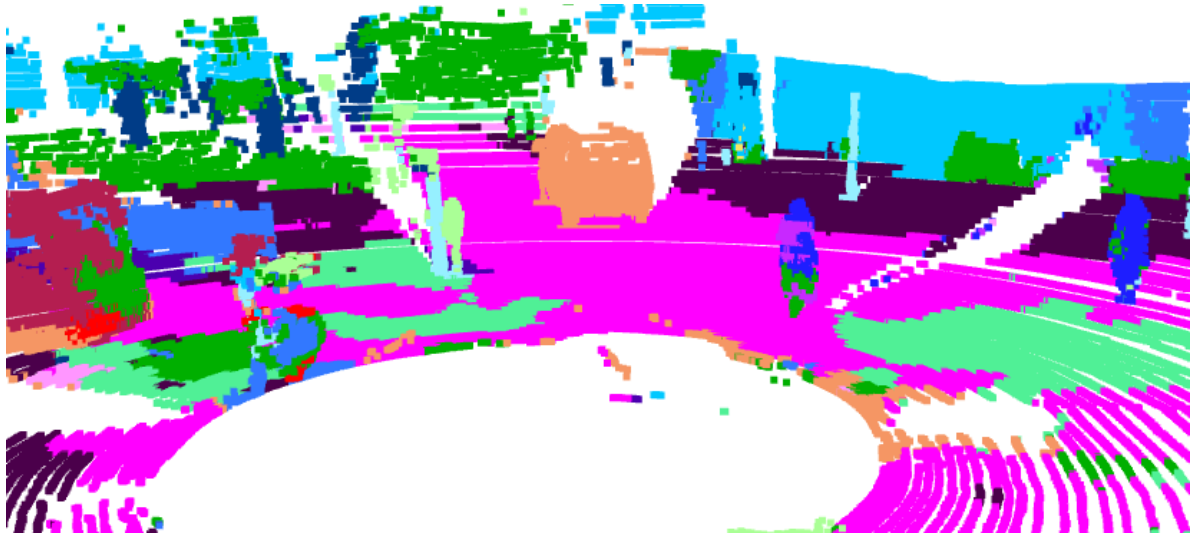
- The goal is to categorize an object represented by a set of 3D points into one of the predefined classes. The general pipeline is to extract global feature vectors.
- Features can be hand crafted or learned from data like in an end-end deep networks learning



Semantic segmentation

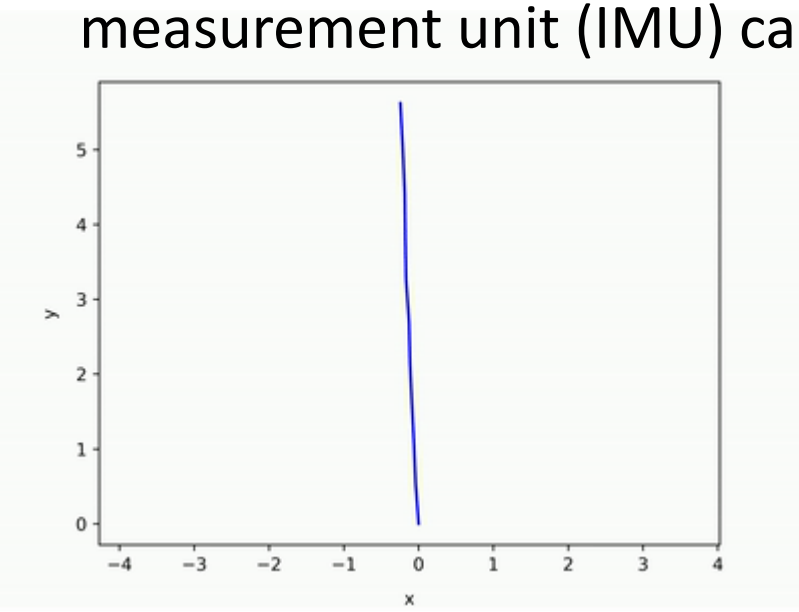
- Class label is assigned to each point in a point cloud
- More accurate point-level understanding of the objects
- Can be achieved with clustering-based methods or learning based methods

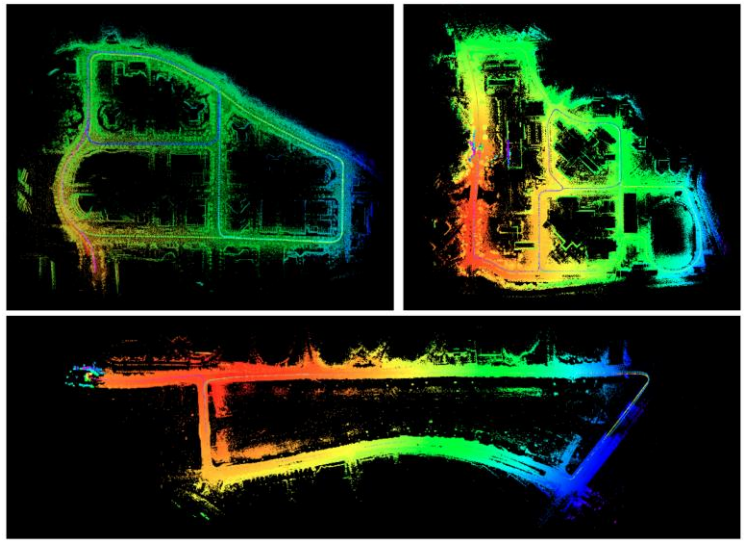




Odometry

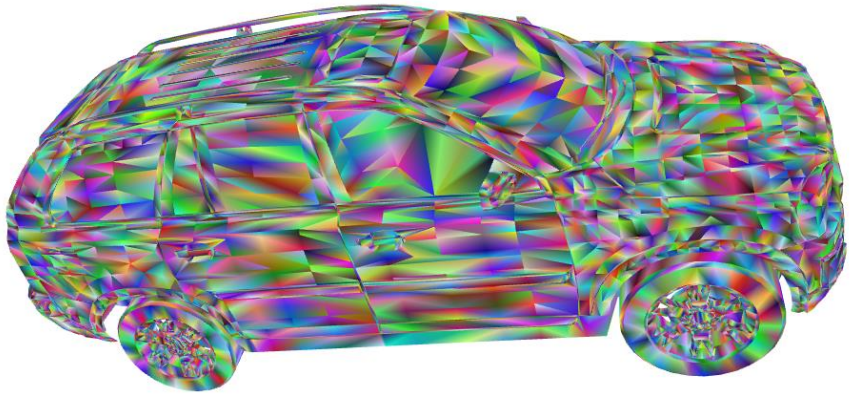
- The process of determining the position and orientation of an object as it traverses its environment
- To keep track of the position and orientation sensor like GPS, inertial measurement unit (IMU) can be used

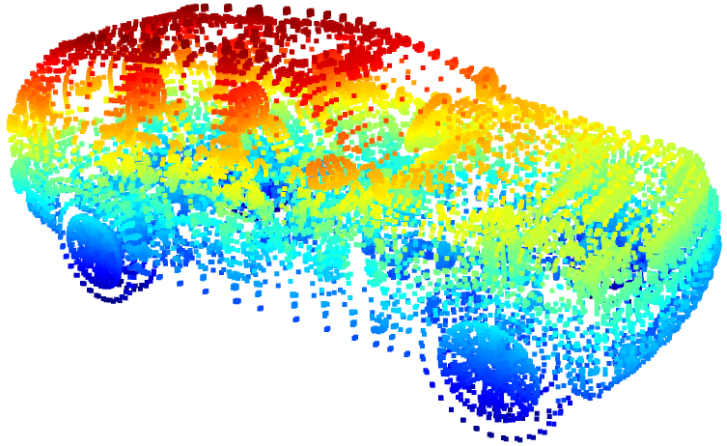




Datasets

- ModelNet40:
  - Compilation of around 12308 Computer-aided Design (CAD) models of common objects
  - Includes 40 object classes (Table, Chairs, Airplanes …)
  - Data is organized in vertices and faces for mesh construction



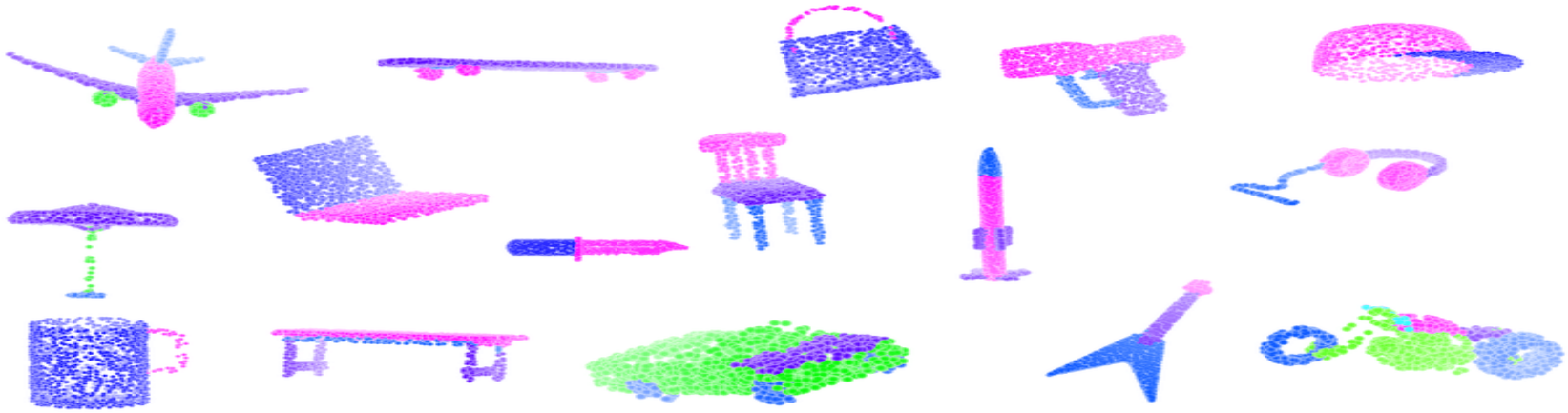


S3DIS (Stanford 3D Indoor Segmentation)

- A subset of the Stanford 2D-3D-Semantics dataset
- Contains point cloud scanned from 6 indoor areas with 271 rooms
- The point clouds are annotated with 13 categories (floor, wall, door ...)
- The data includes XYZ, RGB

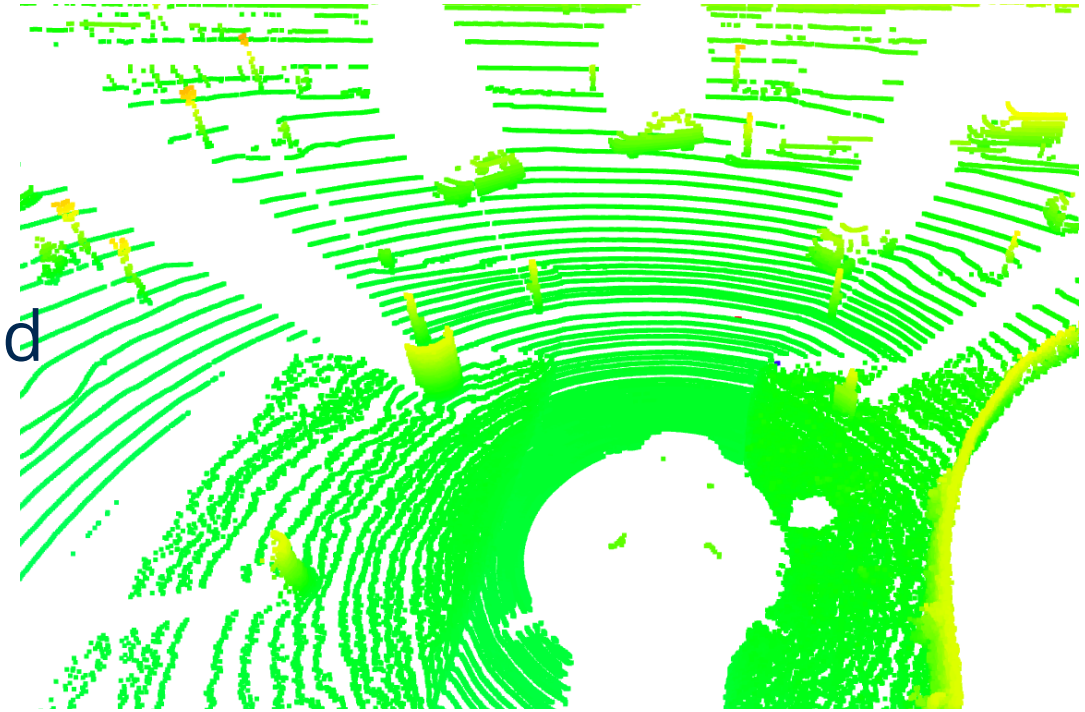
ShapeNet

- Similar to ModelNet40:
  - Compilation of around 57448 Computer-aided Design (CAD) models of common objects
  - Includes 55 object classes (Table, Chairs, Airplanes …)
  - Data is organized in 3D meshes in obj files



KITTI

- Developed for autonomous driving purposes
- Contains RGB images, depth maps and point clouds
- Data taken in the streets of Germany
- The point cloud is captured using a Velodyne LiDAR.



KITTI sensor setup

- One Inertial Navigation System
- One Lidar Velodyne HDL-64E
- Two Grayscale cameras
- Two Color cameras
- The laser scanner spins at 10 frames per second, capturing approximately 100k points per cycle.
- The vertical resolution of the laser scanner is 64.
- The camera images are cropped to a size of 1382 x 512 pixels.

KITTI sensor setup