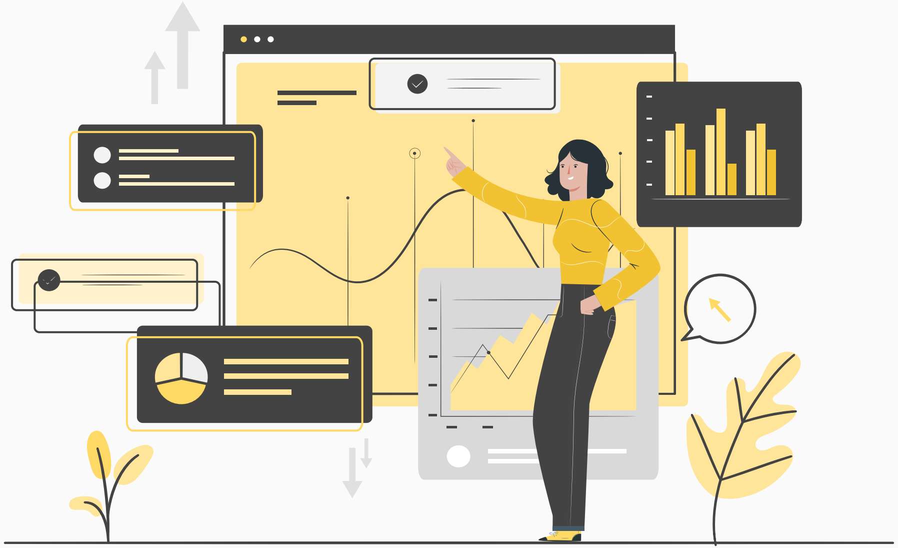
PAPER ANALYSIS



Presented by Yannis He

-

Paper: **Vision meets Robotics: The KITTI Dataset**

Authors: Andreas Geiger, Philip Lenz, Christoph Stiller and Raquel Urtasun

http://www.cvlibs.net/publications/Geiger2013IJRR.pdf

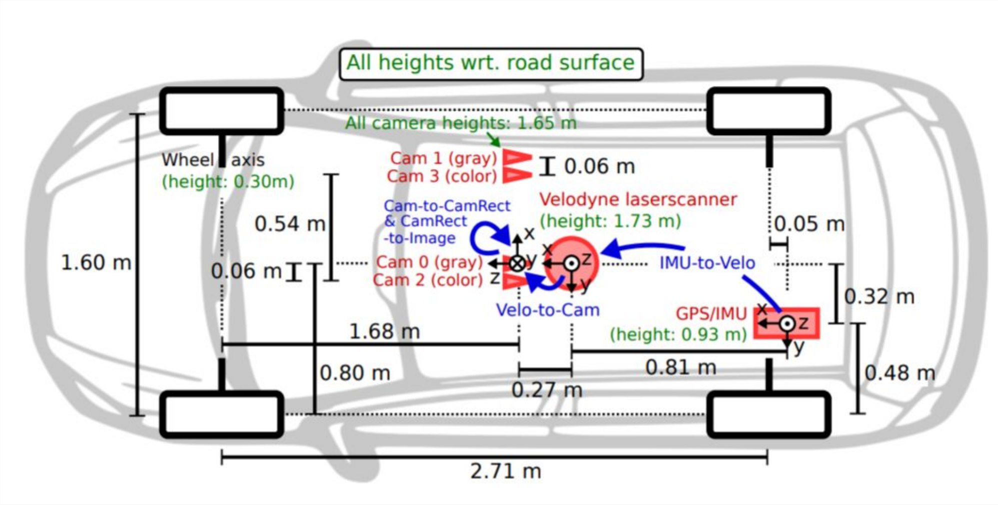
DATASET BRIEF SUMMARY

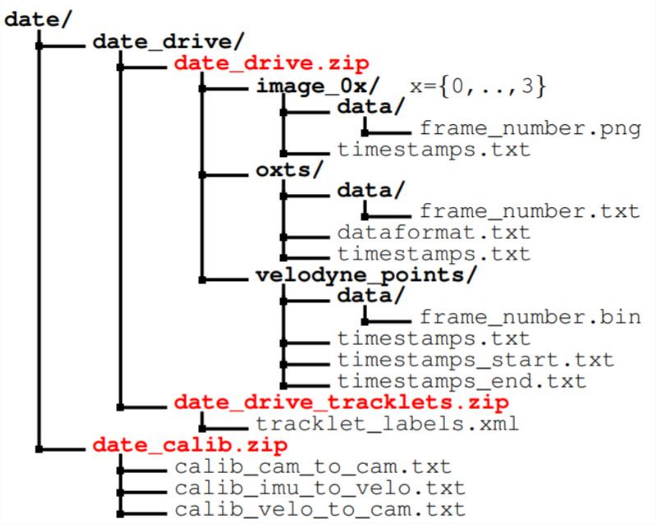
Info of KITTI Dataset:

- Collected around Karlsruhe, Germany
- 6 hours of traffic scenarios at 10-100 Hz with various sensor modalities
  - High-resolution color and grayscale stereo cameras
  - Velodyne 3D laser scanner
  - High-precision GPS/IMU inertial navigation system
    - IMU: Inertial Measurement Unit
- Diverse, real-world scenario:
  - Freeways over rural areas to inner city with static and dynamic objects
- Raw dataset:
  - Road, city, residential, campus, person
- Calibrated, synchronized, timestamped
- Each spin is considered as 1 frame
- Raw image sequences provided
- Image labels in form of 3D tracklets
- Contains online benchmark for stereo, optical flow, object detection and other tasks
- Total size: 180 GB
- Dataset can be downloaded here: http://www.cvlibs.net/datasets/kitti

SENSOR SETUP

- 2 × PointGray Flea2 grayscale cameras (FL2-14S3M-C), 1.4 Megapixels, 1/2" Sony ICX267 CCD, global shutter
- 2 × PointGray Flea2 color cameras (FL2-14S3C-C), 1.4 Megapixels, 1/2" Sony ICX267 CCD, global shutter
- 4 × Edmund Optics lenses, 4mm, opening angle ~ 90◦ , vertical opening angle of region of interest (ROI) ~ 35◦
- 1 × Velodyne HDL-64E rotating 3D laser scanner, 10 Hz, 64 beams, 0.09 degree angular resolution, 2 cm distance accuracy, collecting ~ 1.3 million points/second, field of view: 360◦ horizontal, 26.8◦ vertical, range: 120 m
- 1 × OXTS RT3003 inertial and GPS navigation system, 6 axis, 100 Hz, L1/L2 RTK, resolution: 0.02m / 0.1◦





DATA DESCRIPTION

- Sensor reading of a sequence are zipped as *[date]_[drive].zip*, where drive is placeholder for sequence number
- 3 time stamp files
  - *timestamps_start.txt*
    - Start position of spin
  - *timestamps_end.txt*
    - End position of spin
  - *timestamps.txt:*
    - Time where laser scanner is facing forward and triggering camers
    - Each line is composed of date and time in ours, minutes, and seconds
- Images:
  - 8-bit PNG
  - Rectified image
- OXTS (GPS/IMU):
  - 30 values for each frame
    - Altitude, global orientation, velocities, accelerations, angular rates, accuracies, satellite information
    - Acceleration and angular rate are reported in 2 coordinate systems
      - Body frame (x, y, z)
      - Global frame (f, l, u)
- Velodyne (laser scanner rotates continuously in counter-clockwise direction):
  - Stored as binary float point, including coordinate (x, y, z) and reflectance value ®
  - ~ 120,000 3D points and reflectance value per frame (or per file)

ANNOTATIONS & DEV KIT & SENSORS & NOTATION

Annotations:

- Annotations in form of 3D bounding box in Velodyne coordinates
- Class: Car, Van, Truck, Pedestrian, person, cyclist, Tram, Misc (trailers, segways)
  - Tracklets stored in *date_drive_tracklets.xml*
- Data values:
  - Class
  - 3D size (height, width, length)
  - Translation & rotation

Development Kit:

- MATLab demonstration code with C++ wrapper

Sensors Calibration:

- Camera: x = right, y = down, z = forward
- Velodyne: x = forward, y = left, z = up
- GPS/IMU: x = forward, y = left, z = up

Notations:

- Scaler: lower-case letter
- Vector: bold lower-case
- Matrix: boldface capitals
- 3D rigid-body transformation from coordinate system a to be: T_a^b

NOTATIONS

- $\mathbf{s}^{(i)} \in \mathbb{N}^2$ ....original image size (1392 × 512)
- $\mathbf{K}^{(i)} \in \mathbb{R}^{3\times3}$ ....calibration matrices (unrectified)
- $\mathbf{d}^{(i)} \in \mathbb{R}^5$ ....distortion coefficients (unrectified)
- $\mathbf{R}^{(i)} \in \mathbb{R}^{3\times3}$ ......rotation from camera 0 to camera i
- $\mathbf{t}^{(i)} \in \mathbb{R}^{1\times3}$ .....translation from camera 0 to camera i
- $\mathbf{s}^{(i)}_{rect} \in \mathbb{N}^2$ ....image size after rectification
- $\mathbf{R}^{(i)}_{rect} \in \mathbb{R}^{3\times3}$ ....rectifying rotation matrix
- $\mathbf{P}^{(i)}_{rect} \in \mathbb{R}^{3\times4}$ ......projection matrix after rectification

$i \in \{0, 1, 2, 3\}$

- 0: left grayscale
- 1: right grayscale
- 2: left color
- 3: right color