

VFX Final Project

DeepVO: Towards end-to-end visual odometry with deep Recurrent Convolutional Neural Networks

S. Wang, R. Clark, H. Wen and N. Trigoni, *2017 IEEE International Conference on Robotics and Automation (ICRA)*,
Singapore, 2017

李佳蓮

李尚倫

陳健倫

Pytorch implementation
Rviz visualization

Outline

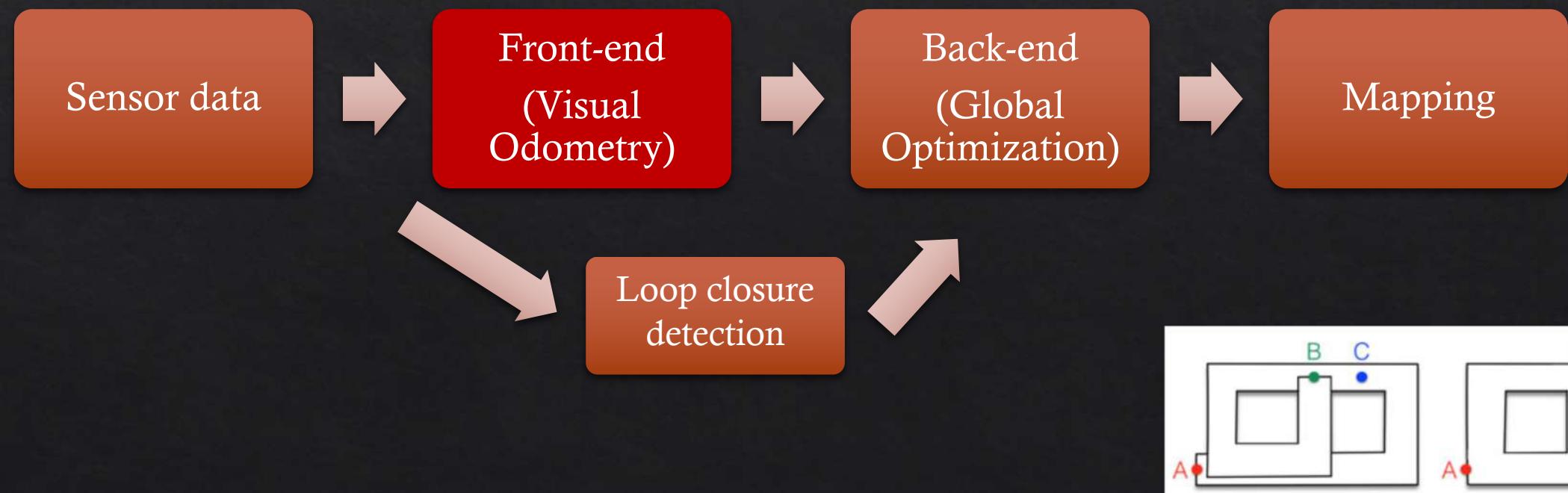
- ❖ What is visual odometry?
- ❖ Traditional vs. Deep Learning Method
- ❖ DeepVO model
- ❖ KITTI dataset
- ❖ Self made dataset through Blender
- ❖ Our prediction on KITTI
- ❖ Our prediction on our dataset

Outline

- ❖ What is visual odometry?
- ❖ Traditional vs. Deep Learning Method
- ❖ DeepVO model
- ❖ KITTI dataset
- ❖ Self made dataset through Blender
- ❖ Our prediction on KITTI
- ❖ Our prediction on our dataset

What is visual odometry?

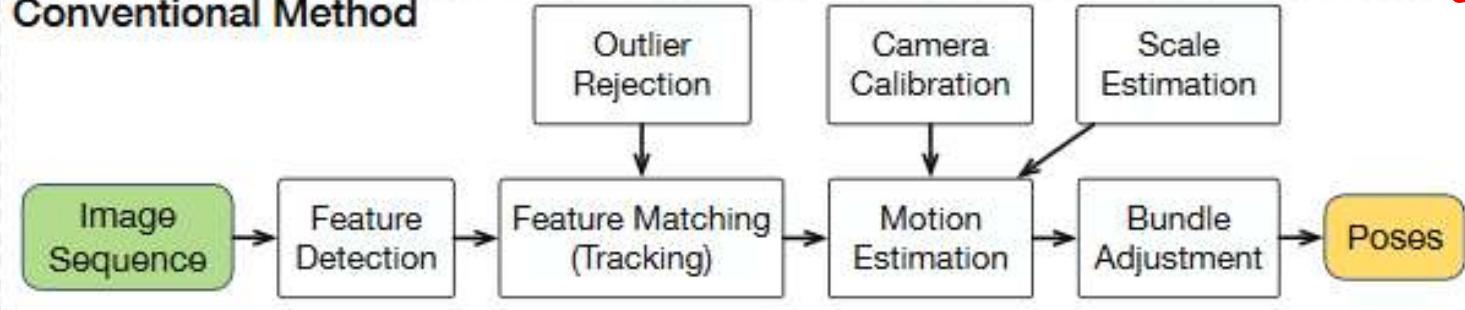
Visual SLAM (SFM) pipeline
(Simultaneous Localization and Mapping)



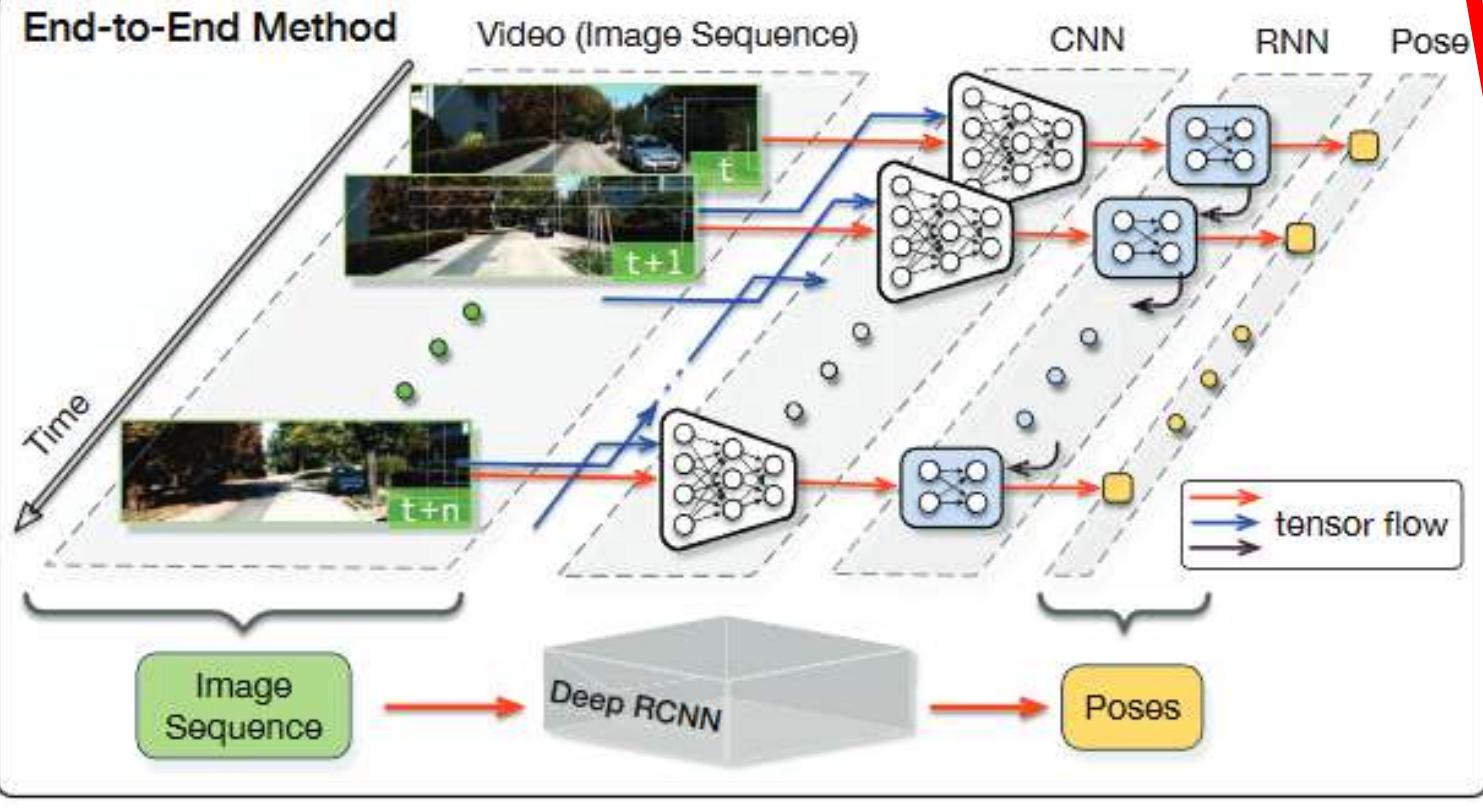
Outline

- ❖ What is visual odometry?
- ❖ Traditional vs. Deep Learning Method
- ❖ DeepVO model
- ❖ KITTI dataset
- ❖ Self made dataset through Blender
- ❖ Our prediction on KITTI
- ❖ Our prediction on our dataset

Conventional Method



End-to-End Method



Feature Detection:

Harris, SIFT, SURF, FAST, ORB...

Feature Description:

Harris, SIFT, BRIEF...

Feature Matching:

Brute-Force, FLANN...

Motion Estimation:

Epipolar constraint, Triangulation, Perspective-n-Point(PnP), Iterative Closest Point(ICP)...

Non-linear Optimization:

Bundle Adjustment(BA), ...

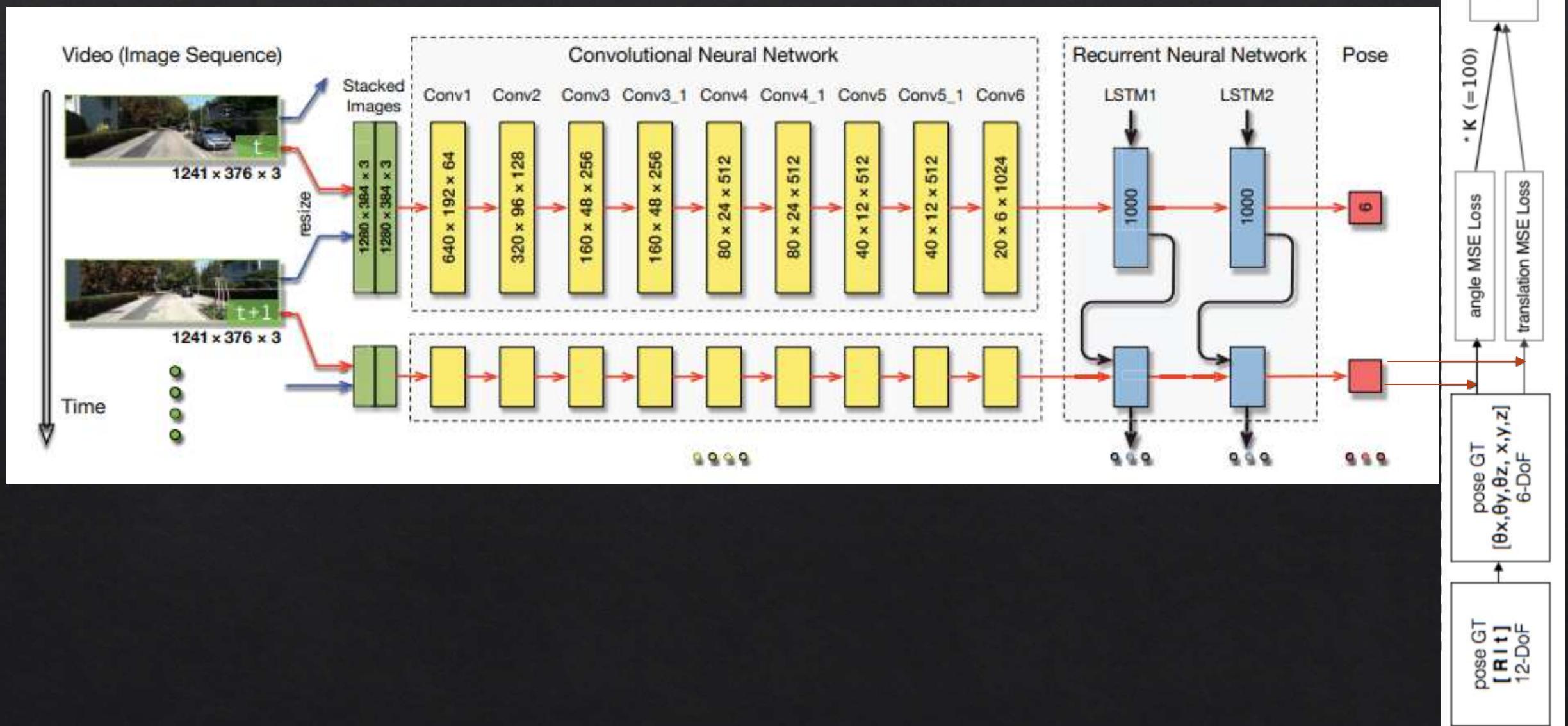
Pose:

Quaternion, Euler angels, Rotation matrix ...

Outline

- ❖ What is visual odometry?
- ❖ Traditional vs. Deep Learning Method
- ❖ DeepVO model
- ❖ KITTI dataset
- ❖ Self made dataset through Blender
- ❖ Our prediction on KITTI
- ❖ Our prediction on our dataset

DeepVo Model



Outline

- ❖ What is visual odometry?
- ❖ Traditional vs. Deep Learning Method
- ❖ DeepVO model
- ❖ KITTI dataset
- ❖ Self made dataset through Blender
- ❖ Our prediction on KITTI
- ❖ Our prediction on our dataset

KITTI Dataset: Odometry benchmark

A. Geiger, P. Lenz, C. Stiller, R. Urtasun, "Vision meets robotics: The KITTI dataset", *Int. J. Robot. Res.*, vol. 32, no. 11, pp. 1231-1237, 2013.

	Method	Setting	Code	Translation	Rotation	Runtime	Environment	Compare
1	RLO		code	0.00 %	0.0000 [deg/m]	0.05 s	GPU @ 1.0 Ghz (C/C++)	<input type="checkbox"/>
M. Dimitrievski., M. Dimitrievski., D. Hamme., D. Hamme., P. Veelaert., W. Philips. and W. Philips.: Robust Matching of Occupancy Maps for Odometry in Autonomous Vehicles . Proceedings of the 11th Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications - Volume 3: VISAPP, (VISIGRAPP 2016) 2016.								
2	V-LOAM			0.55 %	0.0013 [deg/m]	0.1 s	2 cores @ 2.5 Ghz (C/C++)	<input type="checkbox"/>
J. Zhang and S. Singh: Visual-lidar Odometry and Mapping: Low drift, Robust, and Fast . IEEE International Conference on Robotics and Automation(ICRA) 2015.								
3	LOAM			0.57 %	0.0013 [deg/m]	0.1 s	2 cores @ 2.5 Ghz (C/C++)	<input type="checkbox"/>
J. Zhang and S. Singh: LOAM: Lidar Odometry and Mapping in Real-time . Robotics: Science and Systems Conference (RSS) 2014.								
4	IMLS-SLAM++			0.61 %	0.0014 [deg/m]	1.3 s	1 core @ >3.5 Ghz (C/C++)	<input type="checkbox"/>
5	SOFT2			0.65 %	0.0014 [deg/m]	0.1 s	2 cores @ 2.5 Ghz (C/C++)	<input type="checkbox"/>
I. Cvijić, J. Česić, I. Marković and I. Petrović: SOFT-SLAM: Computationally Efficient Stereo Visual SLAM for Autonomous UAVs . Journal of Field Robotics 2017.								
6	IMLS-SLAM			0.69 %	0.0018 [deg/m]	1.25 s	1 core @ >3.5 Ghz (C/C++)	<input type="checkbox"/>
J. Deschaud: IMLS-SLAM: Scan-to-Model Matching Based on 3D Data . 2018 IEEE International Conference on Robotics and Automation (ICRA) 2018.								
7	MC2SLAM			0.69 %	0.0016 [deg/m]	0.1 s	4 cores @ 2.5 Ghz (C/C++)	<input type="checkbox"/>
F. Neuhaus, T. Koss, R. Kohnen and D. Paulus: MC2SLAM: Real-Time Inertial Lidar Odometry using Two-Scan Motion Compensation . German Conference on Pattern Recognition 2018.								
36	VINS-Fusion		code	1.09 %	0.0033 [deg/m]	0.1s	1 core @ 3.0 Ghz (C/C++)	<input type="checkbox"/>
T. Qin, J. Pan, S. Cao and S. Shen: A General Optimization-based Framework for Local Odometry Estimation with Multiple Sensors . 2019.								
42	ORB-SLAM2		code	1.15 %	0.0027 [deg/m]	0.06 s	2 cores @ >3.5 Ghz (C/C++)	<input type="checkbox"/>
R. Mur-Artal and J. Tardós: ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras . IEEE Transactions on Robotics 2017.								
54	RTAB-Map		code	1.26 %	0.0026 [deg/m]	0.1 s	1 core @ 2.5 Ghz (C/C++)	<input type="checkbox"/>
79	VISO2-S		code	2.44 %	0.0114 [deg/m]	0.05 s	1 core @ 2.5 Ghz (C/C++)	<input type="checkbox"/>
A. Geiger, J. Ziegler and C. Stiller: StereoScan: Dense 3d Reconstruction in Real-time . IV 2011.								
102	VISO2-M		code	11.94 %	0.0234 [deg/m]	0.1 s	1 core @ 2.5 Ghz (C/C++)	<input type="checkbox"/>
A. Geiger, J. Ziegler and C. Stiller: StereoScan: Dense 3d Reconstruction In Real-time . IV 2011.								
106	DeepVO			24.55 %	0.0489 [deg/m]	1 s	1 core @ 2.5 Ghz (Python)	<input type="checkbox"/>

train on KITTI dataset video: 00, 01, 02, 05, 08, 09

valid on KITTI dataset video: 04, 06, 07, 10

test on KITTI dataset video: 04, 05, 07, 09, 10



0001



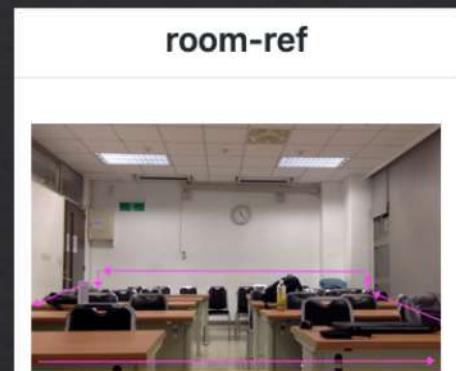
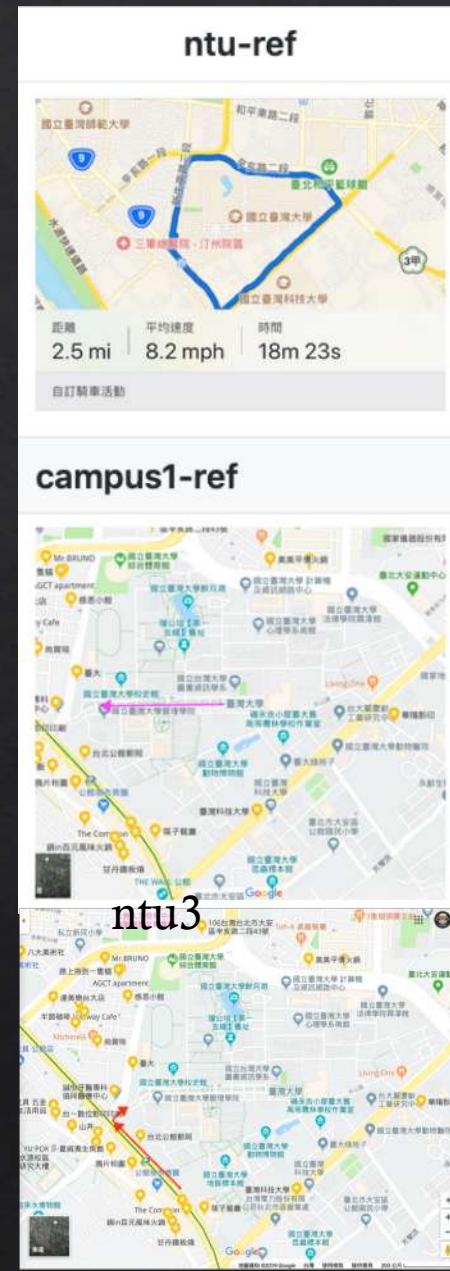
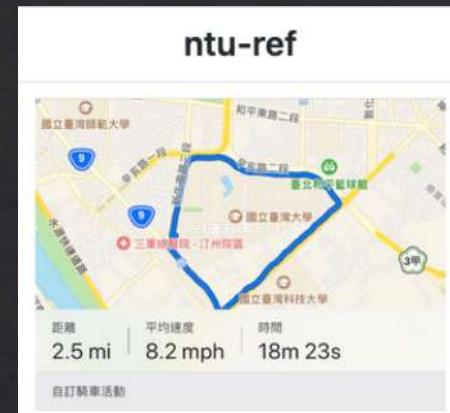
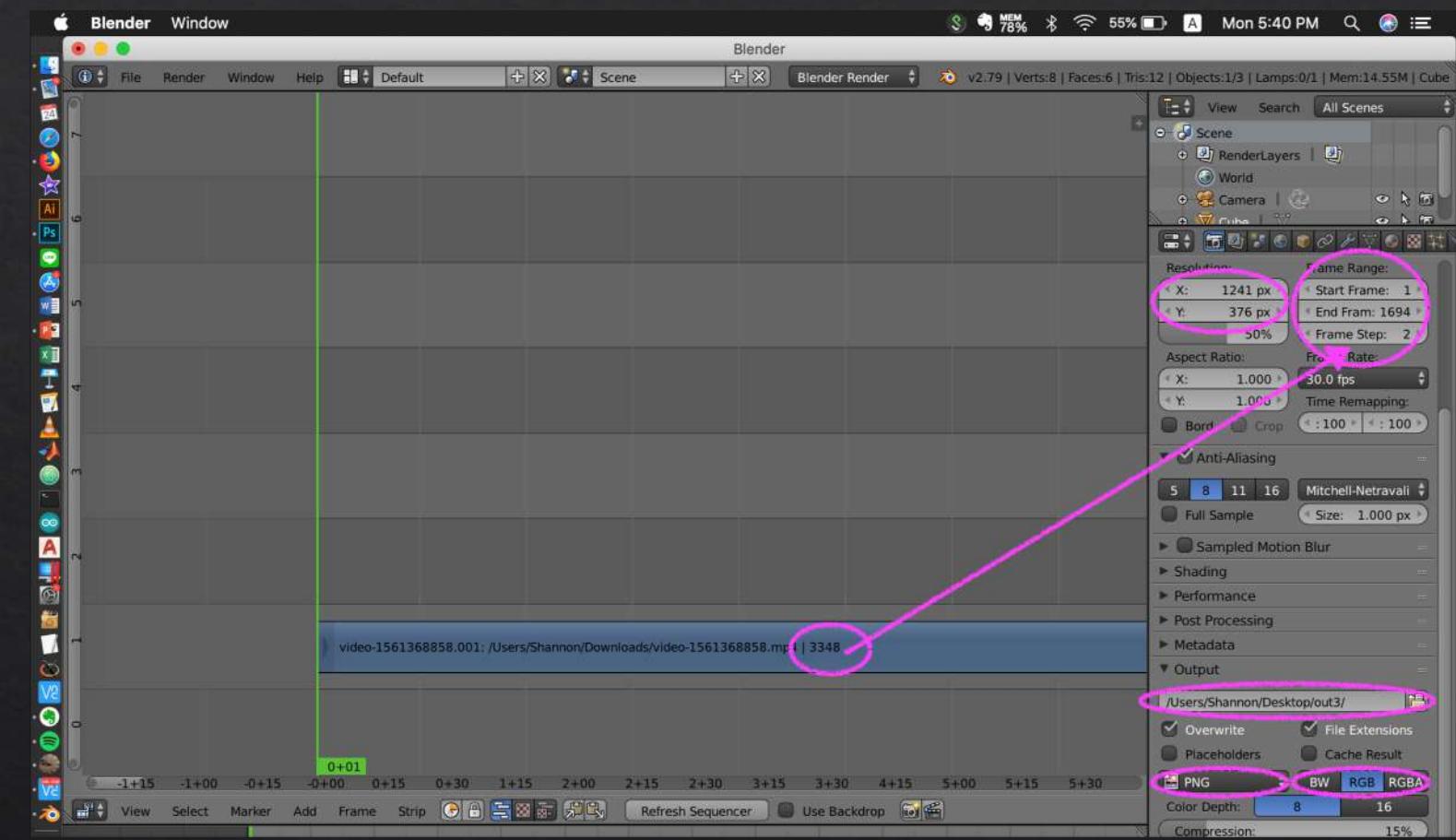
0005



0009

Outline

- ❖ What is visual odometry?
- ❖ Traditional vs. Deep Learning Method
- ❖ DeepVO model
- ❖ KITTI dataset
- ❖ Self made dataset through Blender
- ❖ Our prediction on KITTI
- ❖ Our prediction on our dataset

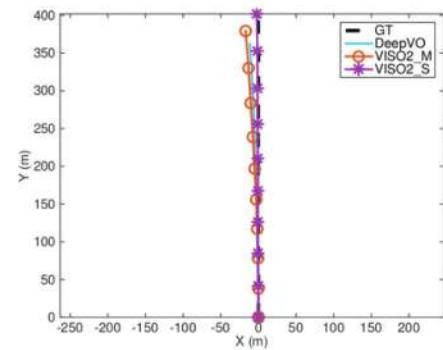


Self made dataset through Blender

Outline

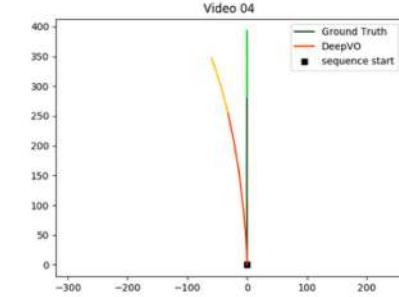
- ❖ What is visual odometry?
- ❖ Traditional vs. Deep Learning Method
- ❖ DeepVO model
- ❖ KITTI dataset
- ❖ Self made dataset through Blender
- ❖ Our prediction on KITTI
- ❖ Our prediction on our dataset

paper result

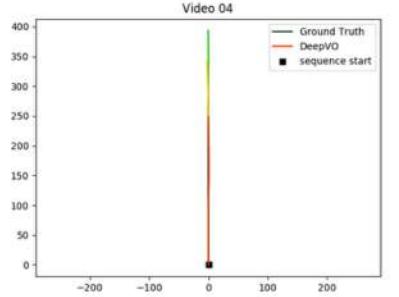


(a) Sequence 04.

pre-trained model from alexart13



our model



04

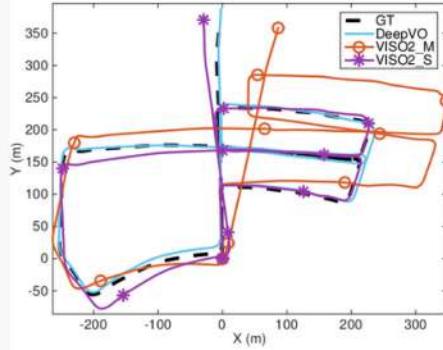
05

07

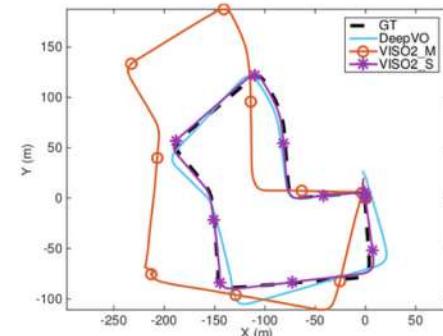
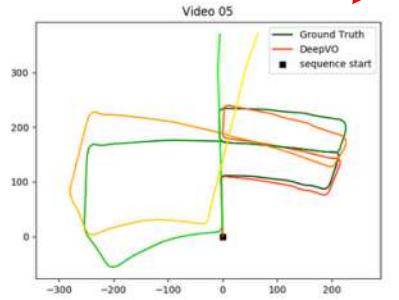
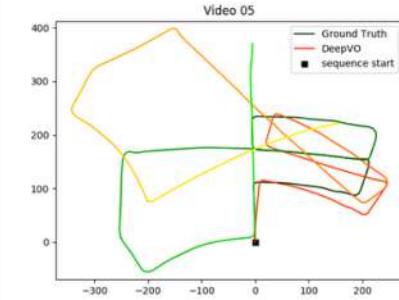
10 09

Green for GT
Red for DeepVO

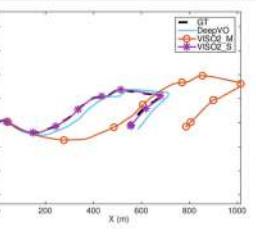
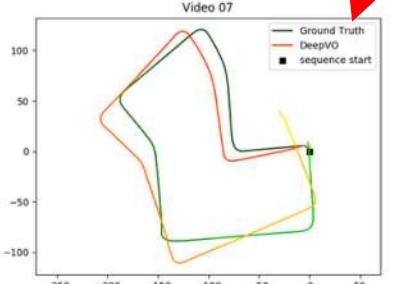
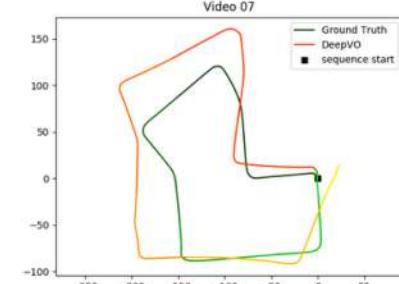
Black Dash for GT
Blue for DeepVO
Red VISO2_M
Purple VISO2_S



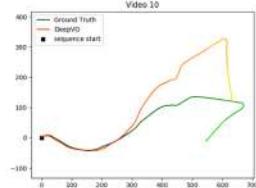
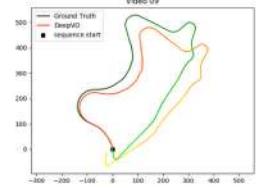
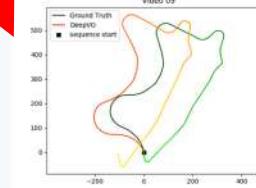
(b) Sequence 05.



(c) Sequence 07.

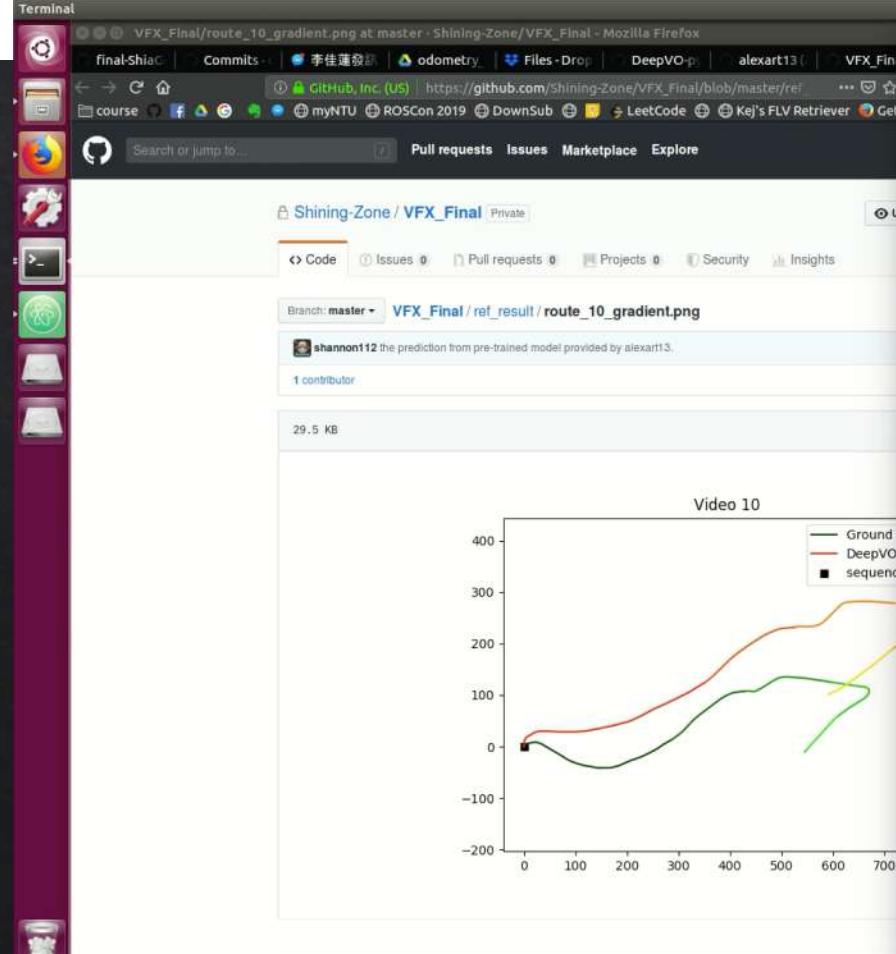
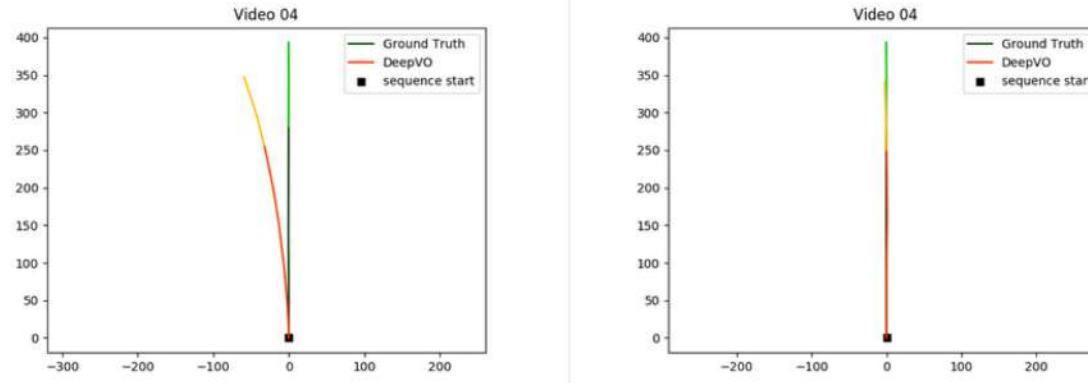
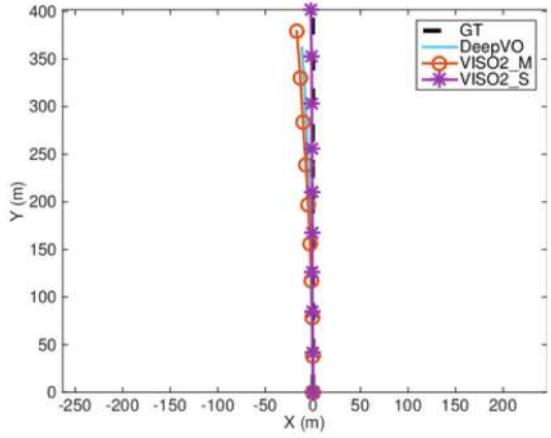


(d) Sequence 10.



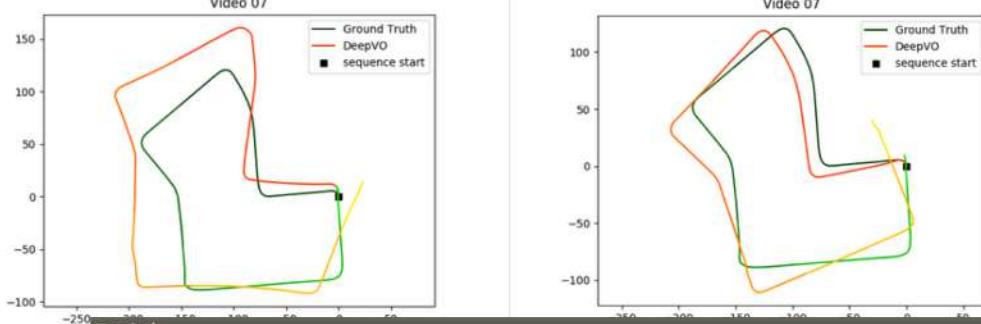
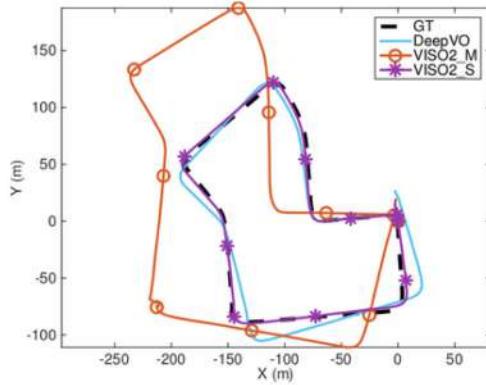
Visualize
KITTI 04
on Rviz

(a) Sequence 04.



Green for GT
Yellow is Our
Red is pretrain

Visualize KITTI 07 on Rviz



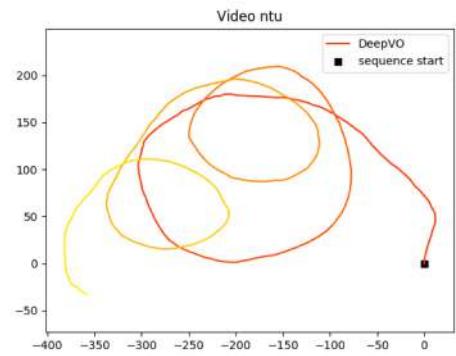
Green for GT
Yellow is Our
Red is pretrain

```
sult/out_05.txt
[ INFO] [1561469835.648185284]: get the odomFile at: /home/shannon/Documents/VFX_Final/GT_pose_rpyxyz/GT_05.txt
[ INFO] [1561469835.648350854]: Odometry points is loaded successfully...
[ INFO] [1561469835.648385139]: Odometry points number: 2761
[ INFO] [1561469835.649302967]: get the imageFile at: /home/shannon/Documents/DeepVO-pyroch/KITTI/Images/05
[ INFO] [1561469835.651633912]: get the odomFile at: /home/shannon/Documents/VFX_Final/re
result/out_05.txt
[ INFO] [1561469835.658867614]: Odometry points is loaded successfully...
[ INFO] [1561469835.658991958]: Odometry points number: 2761
[ INFO] [1561469835.659765808]: get the imageFile at: /home/shannon/Documents/DeepVO-pyro
ch/KITTI/Images/05
[ INFO] [1561469835.664648730]: Odometry points is loaded successfully...
[ INFO] [1561469835.664691558]: Odometry points number: 2761
[ INFO] [1561469835.665699794]: Image is loaded successfully...
[ INFO] [1561469835.665724244]: Image number: 2761
[ INFO] [1561469835.665803845]: Image is loaded successfully...
[ INFO] [1561469835.665849598]: Image number: 2761
[ INFO] [1561469835.667339511]: get the imageFile at: /home/shannon/Documents/DeepVO-pyro
ch/KITTI/Images/05
[ INFO] [1561469835.672331054]: Image is loaded successfully...
[ INFO] [1561469835.672605286]: Image number: 2761
odom_kitti_publisher_pn-3 process has died [pid 27791, exit code -11, cmd /home/shanno
m_kitti_publisher_pn _log=/home/shannon/.ros/log/57c9cdf4-974e-11e9-90fe-1cb72caaff51/odom_kitti_
publisher_pn-3.log]
odom_kitti_publisher_pn-2 process has died [pid 27790, exit code -11, cmd /home/shannon/Do
cuments/ros-kinetic/devel/lib/odometry_visualizer/odom_kitti_publisher _name=odom_kitti_
publisher_pn _log=/home/shannon/.ros/log/57c9cdf4-974e-11e9-90fe-1cb72caaff51/odom_kitti_
publisher_pn-2.log]
log file: /home/shannon/.ros/log/57c9cdf4-974e-11e9-90fe-1cb72caaff51/odom_kitti_
publisher_pn-2.log
odom_kitti_publisher_GT-4 process has died [pid 27797, exit code -11, cmd /home/shannon/Do
cuments/ros-kinetic/devel/lib/odometry_visualizer/odom_kitti_publisher_GT _name=odom_kitti_
publisher_GT _log=/home/shannon/.ros/log/57c9cdf4-974e-11e9-90fe-1cb72caaff51/odom_kitti_
publisher_GT-4.log]
log file: /home/shannon/.ros/log/57c9cdf4-974e-11e9-90fe-1cb72caaff51/odom_kitti_
publisher_GT-4.log
^C[rviz-6] killing on exit
[car2cam_broadcaster_GT-5] killing on exit
[rosout-1] killing on exit
[master] killing on exit
shutting down processing monitor...
... shutting down processing monitor complete
done
shannon@shannon-desktop:~/Documents/ros-kinetic$ rosrun odometry_visualizer odometry_k
itti_visualizer.launch
... logging to /home/shannon/.ros/log/b03af228-974f-11e9-90fe-1cb72caaff51/roslaunch-shan
non-desktop-28227.log
Checking log directory for disk usage. This may take awhile.
Press Ctrl-C to interrupt
WARNING: disk usage in log directory [/home/shannon/.ros/log] is over 1GB.
It's recommended that you use the 'rosclean' command.
```

Outline

- ❖ What is visual odometry?
- ❖ Traditional vs. Deep Learning Method
- ❖ DeepVO model
- ❖ KITTI dataset
- ❖ Self made dataset through Blender
- ❖ Our prediction on KITTI
- ❖ Our prediction on our dataset

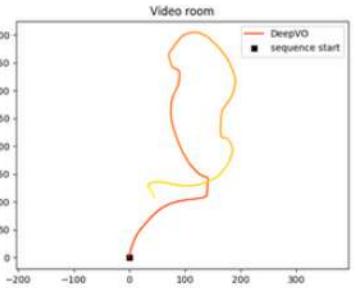
ntu



ntu-ref



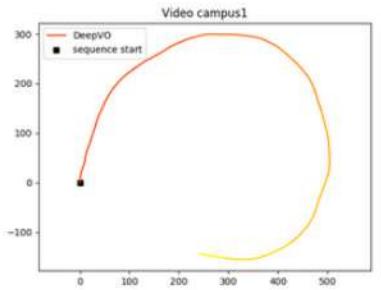
room



room-ref



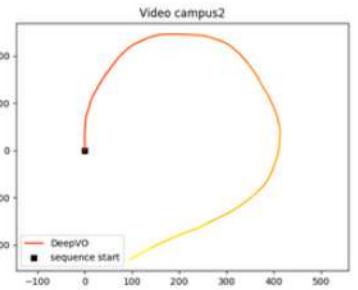
campus1



campus1-ref



campus2

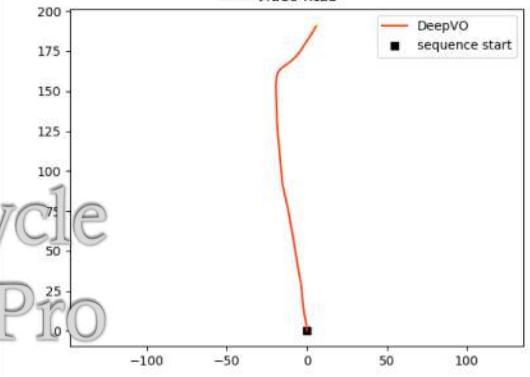


campus2-ref



->
By
motorcycle
with goPro

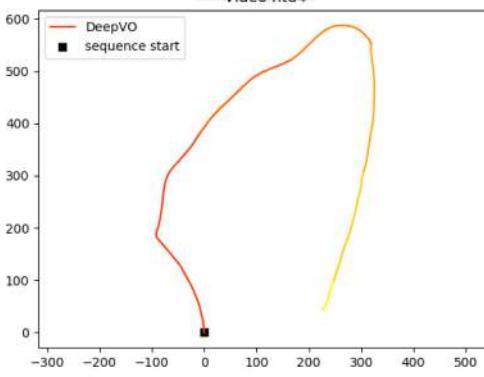
ntu3



ntu3_ref



ntu4



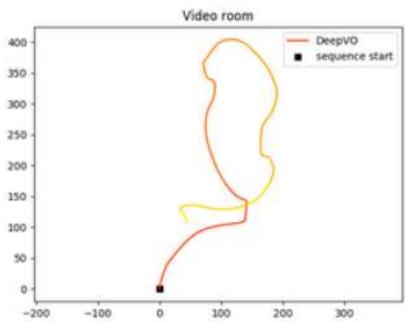
ntu4_ref



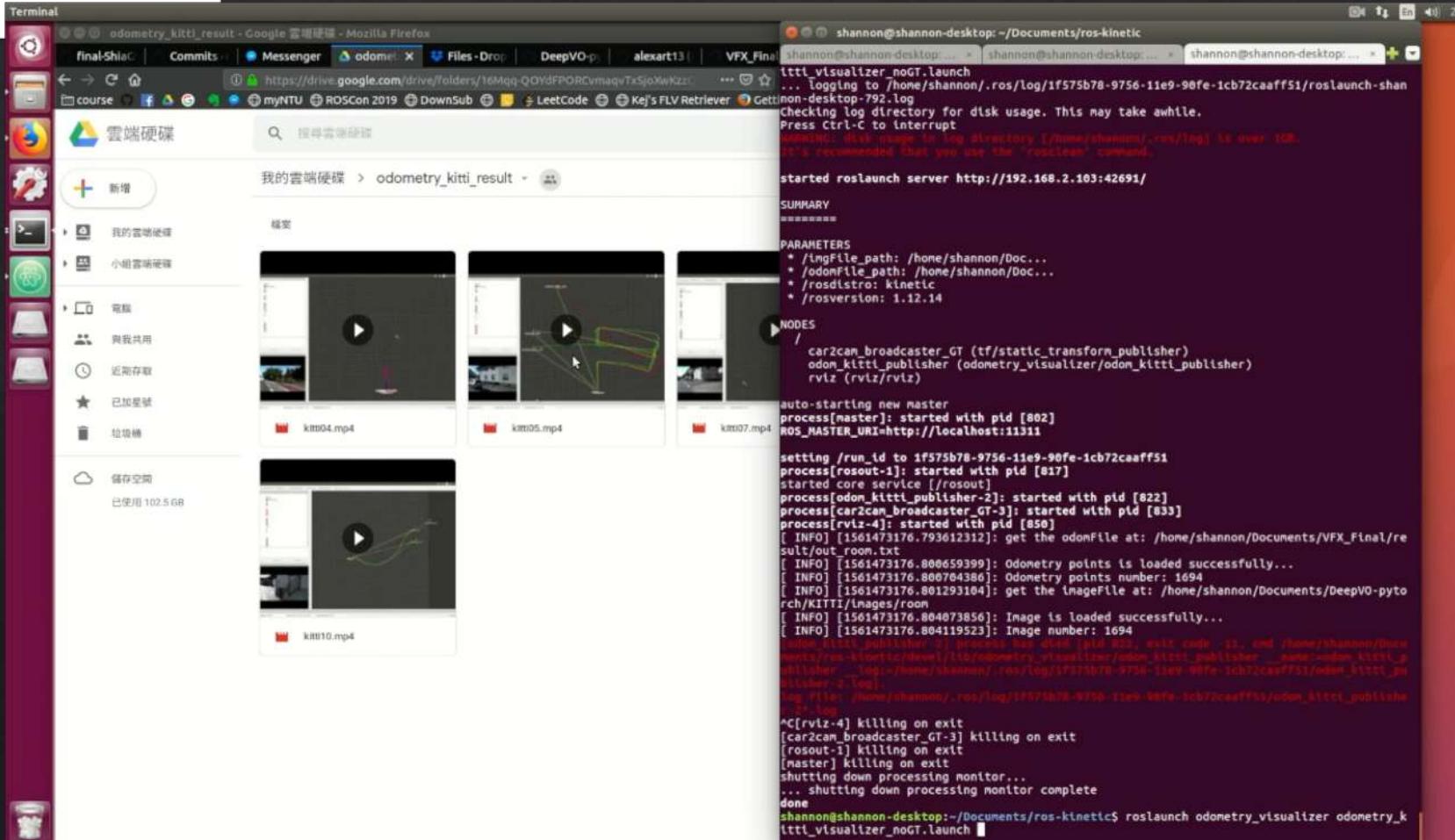
By riding
bicycle and
walking with
iphone

Visualize Our dataset room (walking) on Rviz

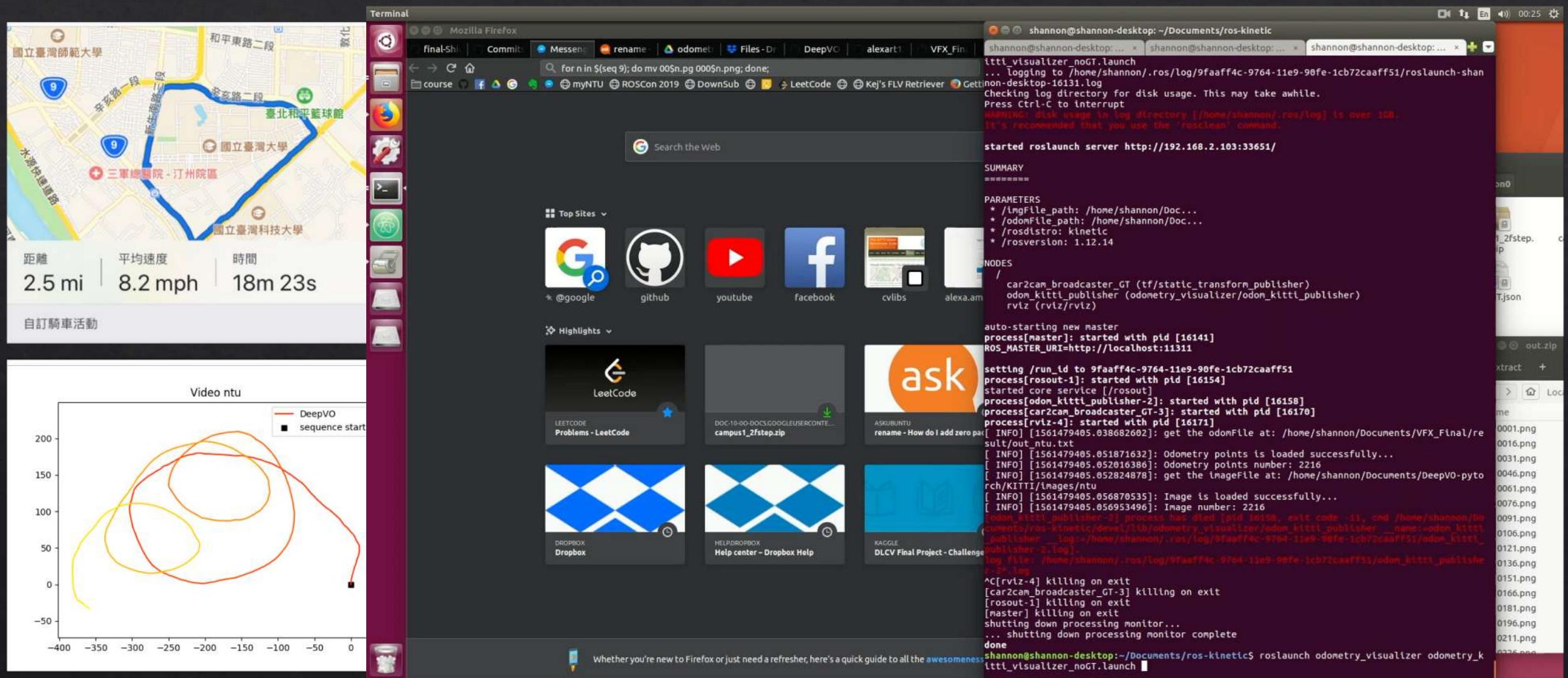
room



room-ref



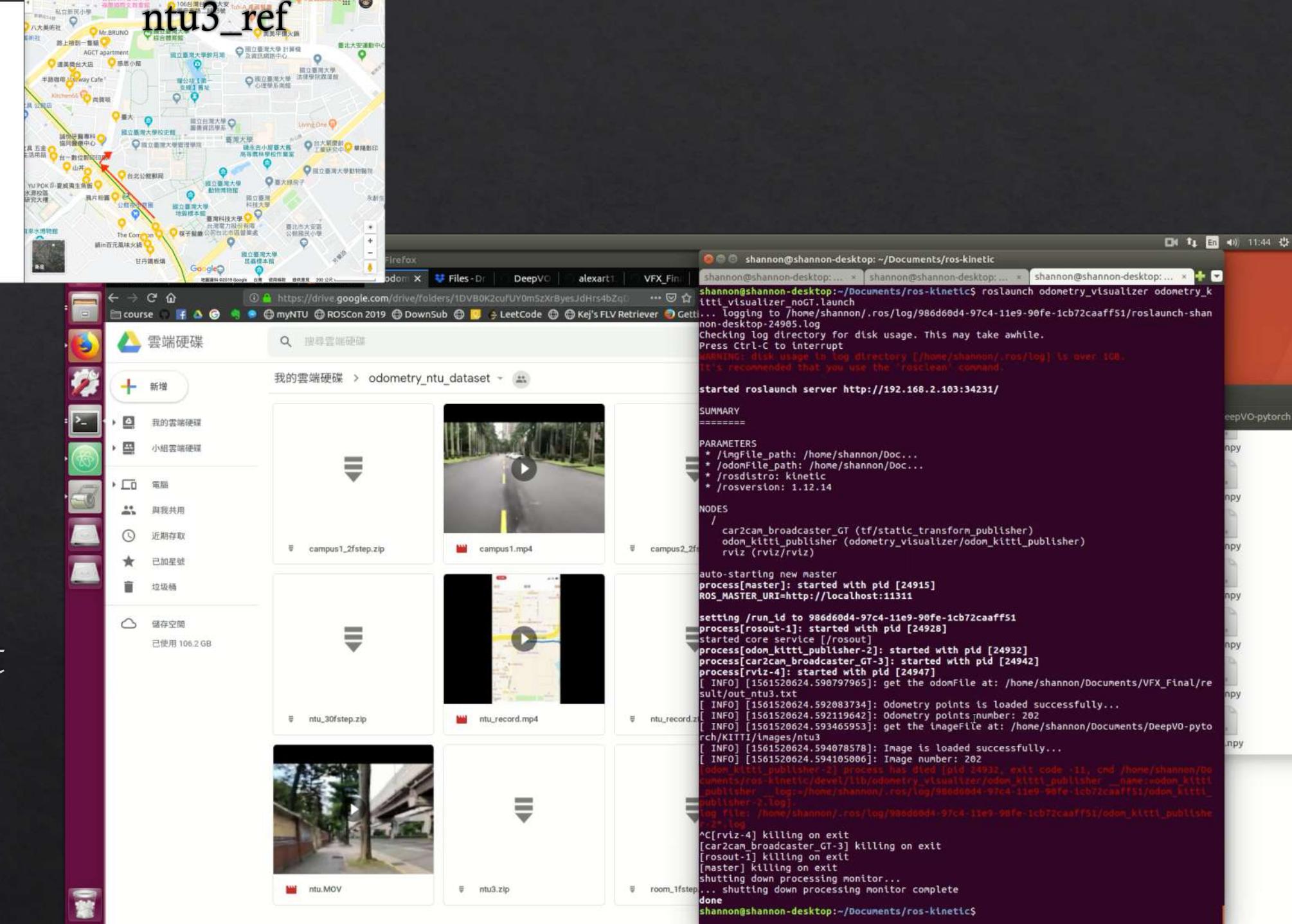
Demo on ntu(cycling) in our dataset



Visualize Our dataset ntu3 (biking) on Rviz

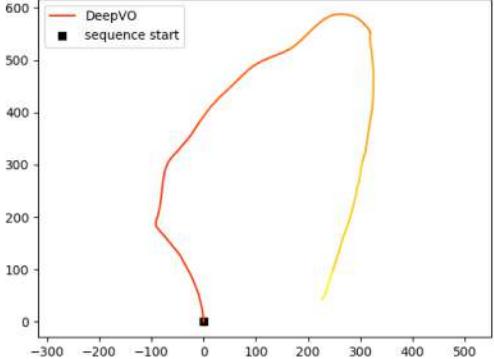
ntu3

biking

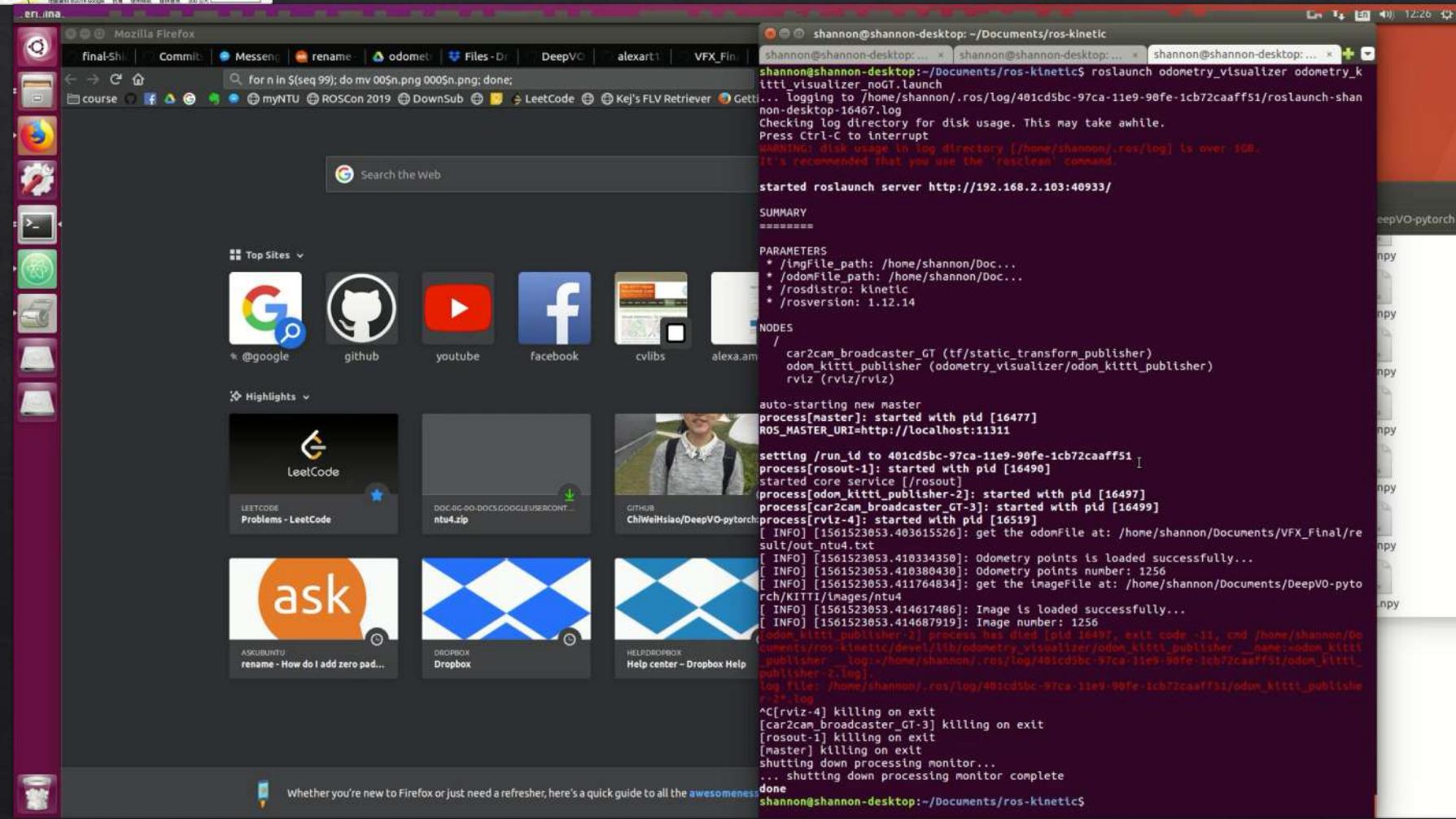


Visualize Our dataset ntu4 (biking) on Rviz

ntu4



ntu4_ref



Thanks

Raw
Image

第一張 第二張 第三張 第四張 第五張 第六張 第七張

Raw
Label

1 2 3 4 5 6 7

Image

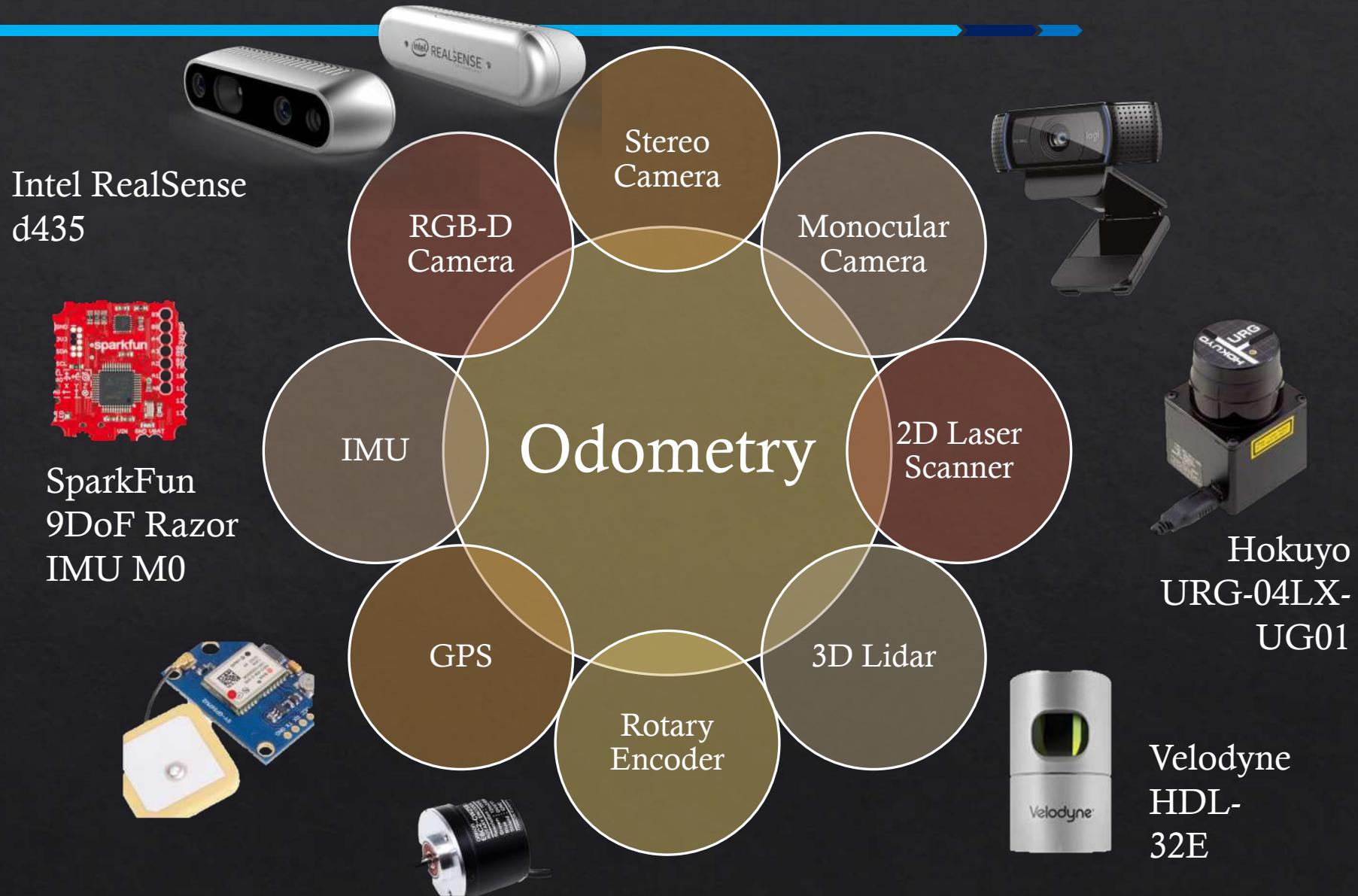
第一張 第二張 第三張 第四張 第五張 第六張 第七張
第一張 第二張 第三張 第四張 第五張 第六張 第七張

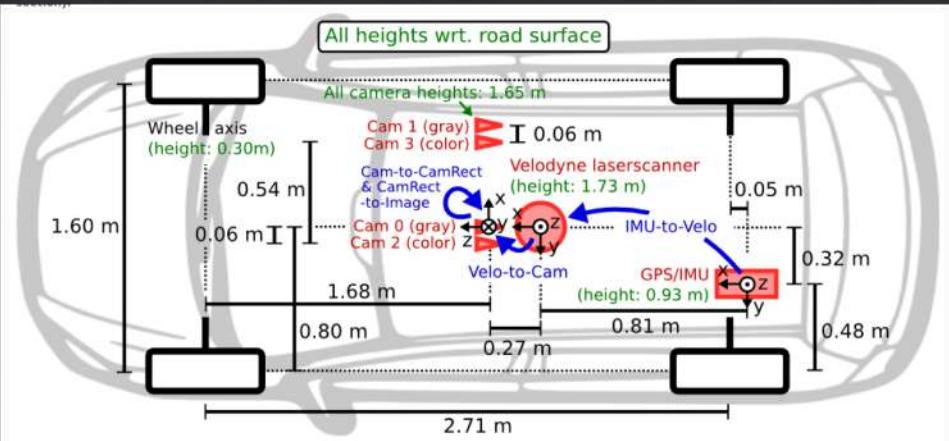
Label

1 2 3 4 5 6 7

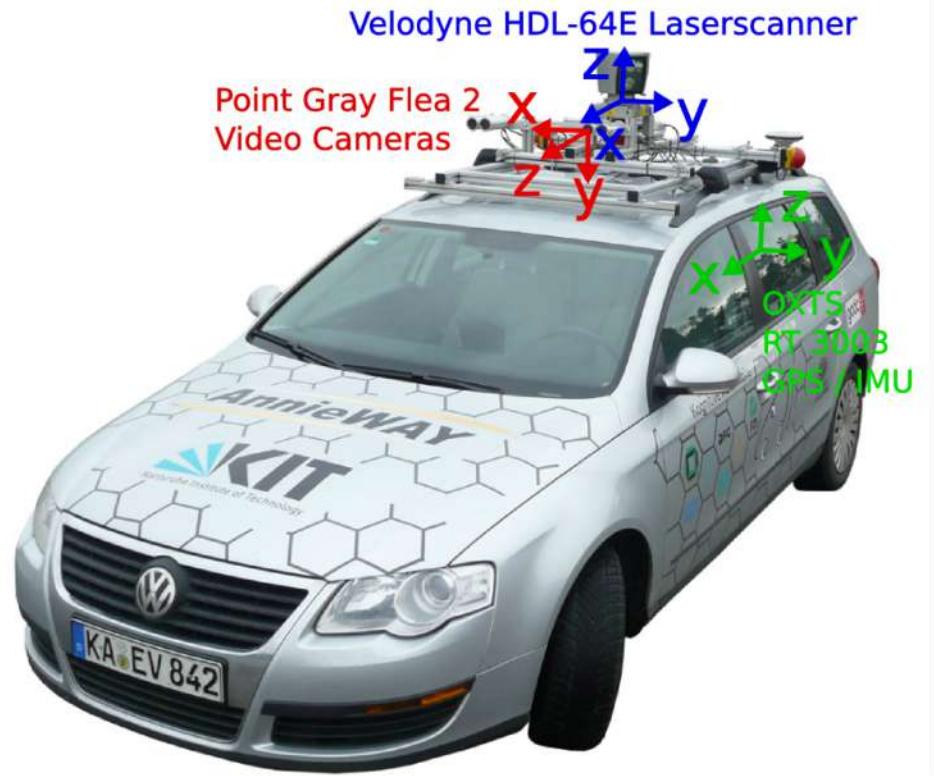
相對座標

Odometry related Sensors



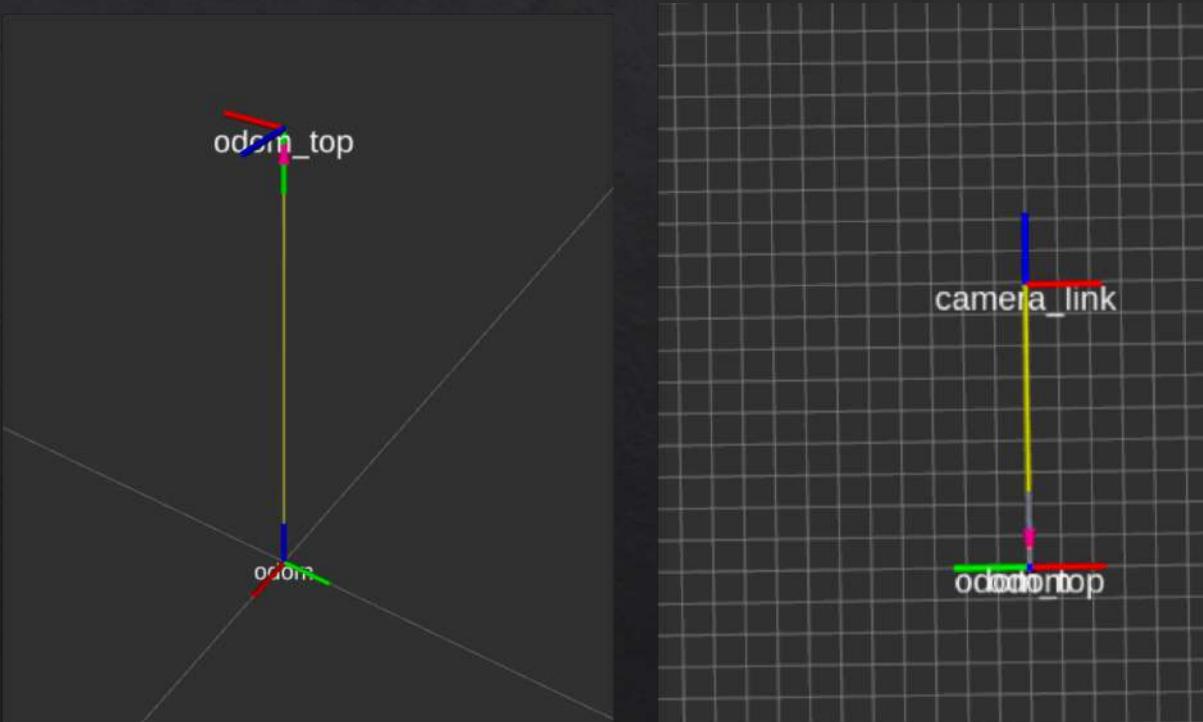


This figure shows our fully equipped vehicle:



Monocular Visual Odometry

- We only using Cam3 image
- Camera pose Z axis is forward



Original pose and camera pose on Our visualizer on Rviz