# 3D Computer Vision and Deep Learning Applications
# Homework 03

**Ming-Yang, Ho (何明洋) R08945027**

## Code Execution

## Environments

open3d == 0.12.0

opencv-python == 4.5.1.48

numpy >= 1.19.5

### Run code

```
$python3 vo.py --input ./frames/ --camera_parameters ./camera_parameters.npy
```

will open 2 visualizers, one for image visualization and one for trajectory visualization

## Problem

*reference*

1. *https://github.com/FlagArihant2000/visual-odometry*
2. *https://github.com/v-shetty/Visual-Odometry-for-Monocular-Camera*
3. *https://stackoverflow.com/questions/63413018/opencv-triangulate-points-from-2-images-to-estimate-the-pose-on-a-third*

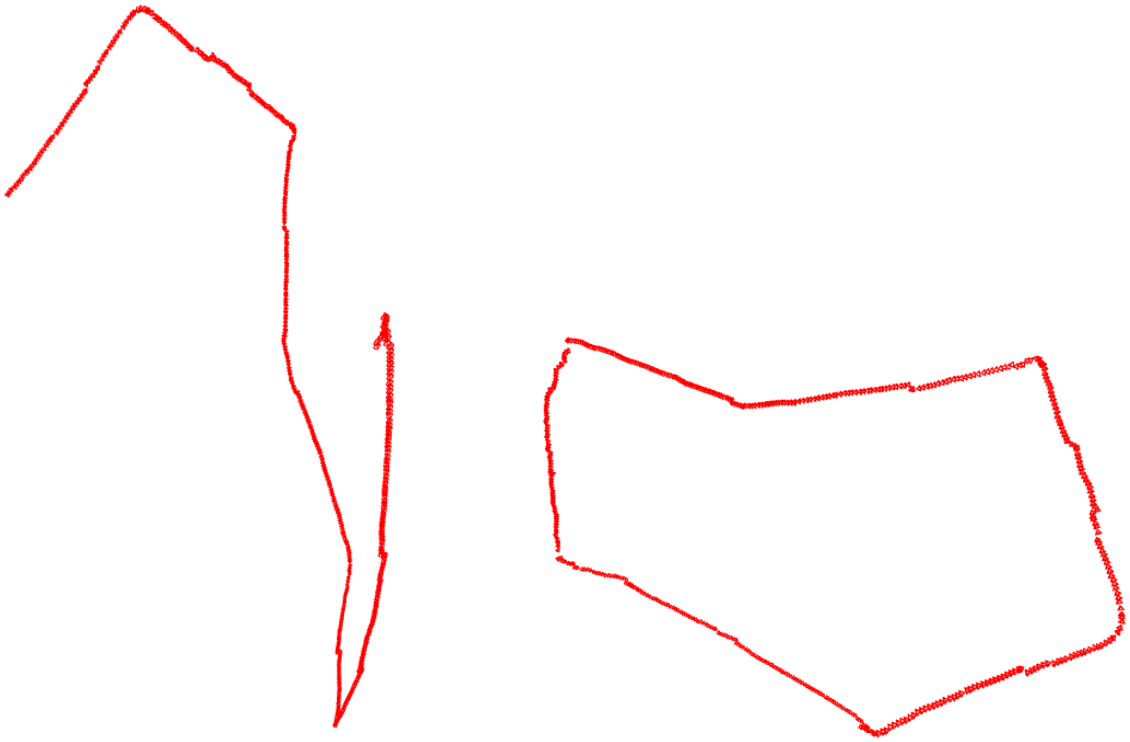
## Implementation details

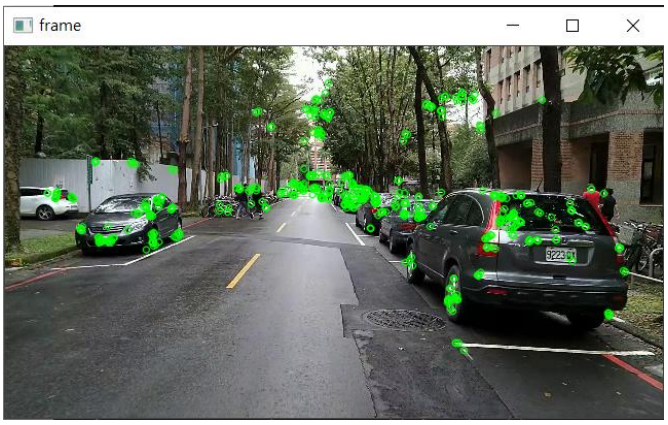
ORB is utilized as the feature descriptor.

For every two images pair, the rotation and translation are calculated by the full matched points.

To re-scale the translation, the previous matches points is kept and make an intersection with the current matched points. Use the intersection points to do triangulation and to calculate the ratio of distances. The median of them will be utilized to re-scale translation.

### Qualitative results







**Demo video**

Please see **screencast.mp4** in the folder

**Discussion**

As the screenshots shown, the fours turning events are well-captured and obviously about 90 degrees in the trajectory. However, the trajectory cannot be closed. One reason is that there is fluctuation on the trajectory

plane (imagined z-axis) and the ending point is much higher than the start point. Another reason is that during the trajectory from corner-2 to corner-3, there is an extreme skew path on this fragment. These inaccuracies might be owing to there are many dynamic objects (e.g. moving people and bicycles) that are detected as matches points and probably for rotation and translation estimation. Actually, it is arduous for a visual odometry scheme to close a cycle like this case that only use the previous and the current picture to calculate the rotation and translation matrix. For further improvement, the utilization of much previous points or the detection of closing (a similar screen is visited again) might be the solution.