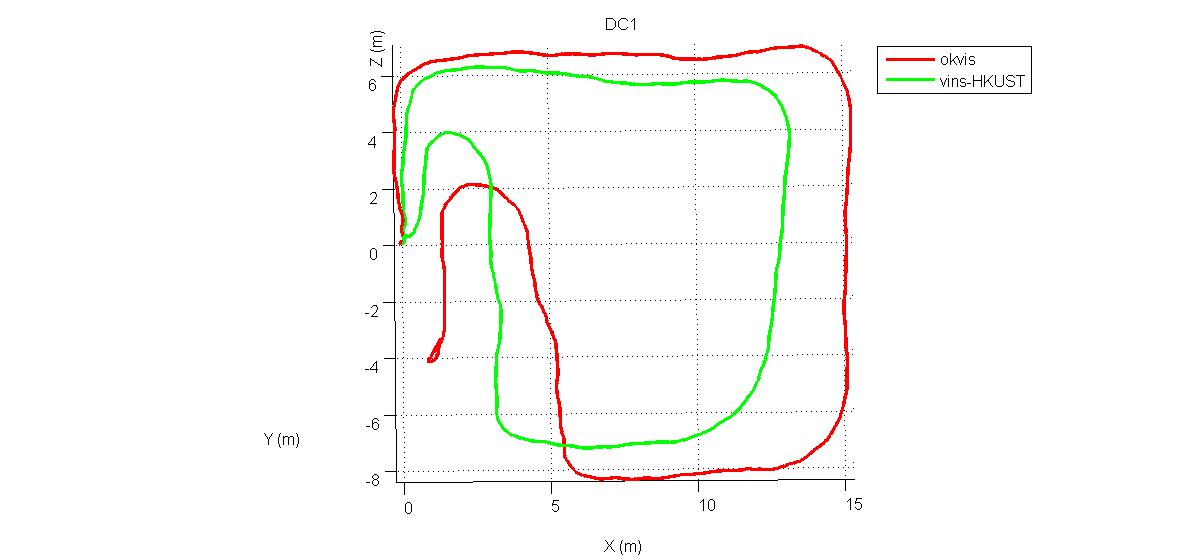
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| DataSets | Length | People in Sight | Okvis EPEN | Vins HKUST |
| DC\_downward | ~ 60 m | Less | ~4.195 m | ~0.553 m |
| DC2 | ~ 70 m | Normal | Failed | ~ 3.41 m |
| DC3 | ~ 110 m | Normal | Failed | ~ 10.34 m |
| DC4 | ~ 70 m | Many | Failed | Failed |
| AT1 | ~ 70 m | Many | Failed | Failed |
| AT2 | ~ 70 m | Many | Failed | Failed |

People in sight indicates how many and how often that people were observed by the camera in the video. For the datasets DC4, AT1, AT2, many people appear in the camera’s view, the VIO methods all failed (very large pose estimation error).

Trajectories comparison is shown in the figures below using datasets DC\_downward (DC1), DC2, DC3, respectively. The red is the result from OKVIS[1], while the green is from vins-HKUST[2], which performs much better than OKVIS.

[1] Stefan Leutenegger, Simon Lynen, Michael Bosse, Roland Siegwart and Paul Timothy Furgale. Keyframe-based visual–inertial odometry using nonlinear optimization. The International Journal of Robotics Research, 2015.

[2] Qin, Tong, and Shaojie Shen. "Robust Initialization of Monocular Visual-Inertial Estimation on Aerial Robots." *Proc. of the IEEE/RSJ Int. Conf. on Intell. Robots and Syst., Vancouver, Canada*. 2017.



Trajectory Comparison for Dataset downward

