# VECtor: A Versatile Event-Centric Benchmark for Multi-Sensor SLAM

Supplementary Material

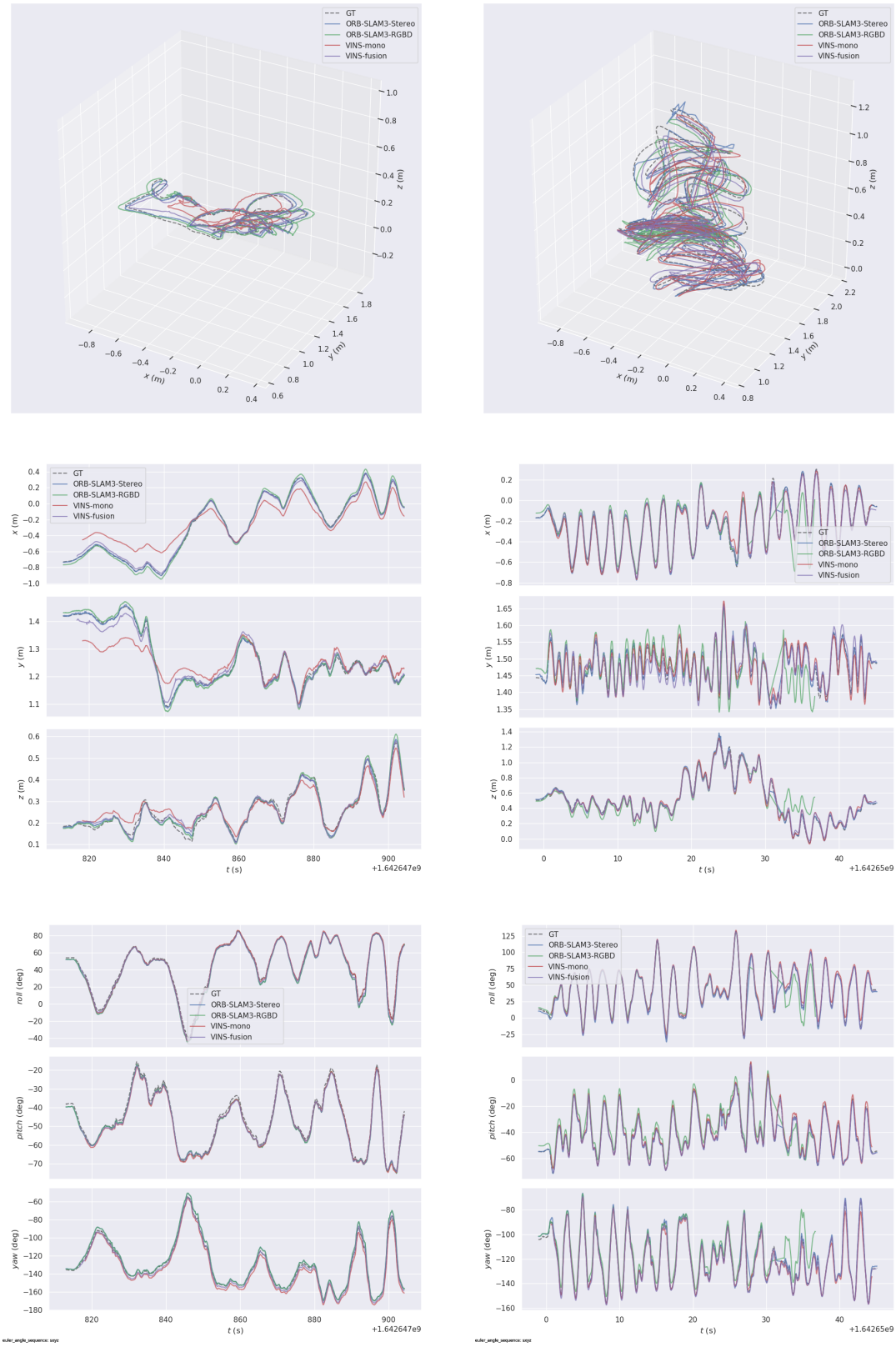
## Small-scale Data Sequences

Table 1: State-Of-The-Art algorithms' Performance

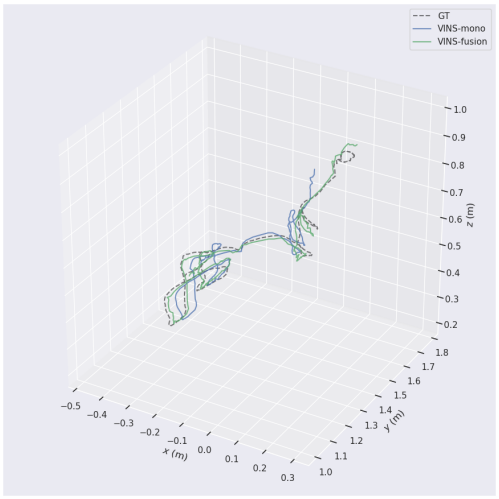
| Data Sequence | ORB-SLAM3 Stereo | ORB-SLAM3 RGBD | VINS mono | VINS fusion |
|---|---|---|---|---|
| *desk-normal* | ✔ | ✔ | ✔ | ✔ |
| *desk-fast* | ✔ | ✖ | ✔ | ✔ |
| *hdr-normal* | ✖ | ✖ | ✔ | ✔ |
| *hdr-fast* | ✖ | ✖ | ✔ | ✔ |
| *mountain-normal* | ✖ | ✔ | ✔ | ✔ |
| *mountain-fast* | ✖ | ✔ | ✔ | ✖ |
| *robot-normal* | ✔ | ✔ | ✔ | ✔ |
| *robot-fast* | ✔ | ✔ | ✔ | ✔ |
| *sofa-normal* | ✔ | ✔ | ✔ | ✔ |
| *sofa-fast* | ✔ | ✔ | ✔ | ✔ |

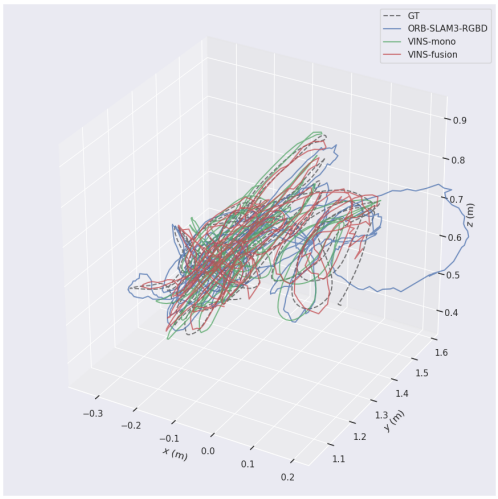
- **ORB-SLAM3-Stereo** takes the readings from the regular stereo camera.
- **ORB-SLAM3-RGBD** takes the readings from the left regular camera, and the depth readings reprojected onto the left regular frame.
- **VINS-mono** takes the readings from the left regular camera and the IMU.
- **VINS-fusion** takes the readings from the regular stereo camera and the IMU.
- ✔ indicates a good result performed by this algorithm.
- ✔ indicates an average result performed by this algorithm.
- ✖ indicates an incomplete result performed by this algorithm.
- ✖ indicates no result can be generated by this algorithm.
- All trajectories are first transformed to the same reference frame as the ground truth poses by extrinsics, then further aligned with all poses by Umeyama's SE(3) method.

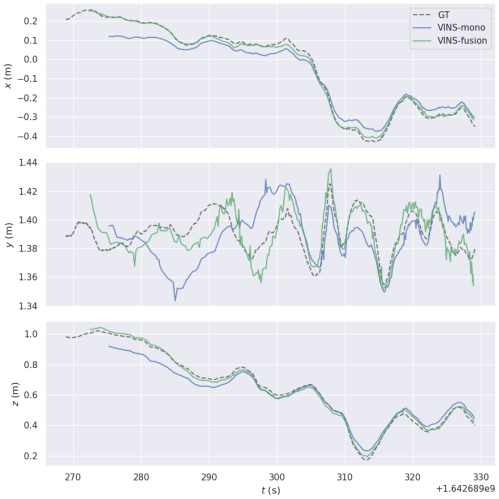
## desk-normal (left column) and desk-fast (right column)

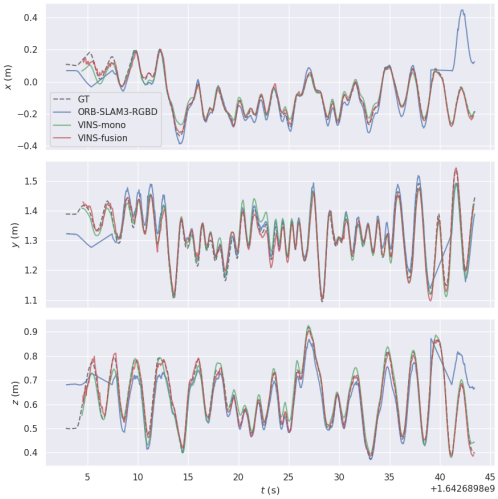


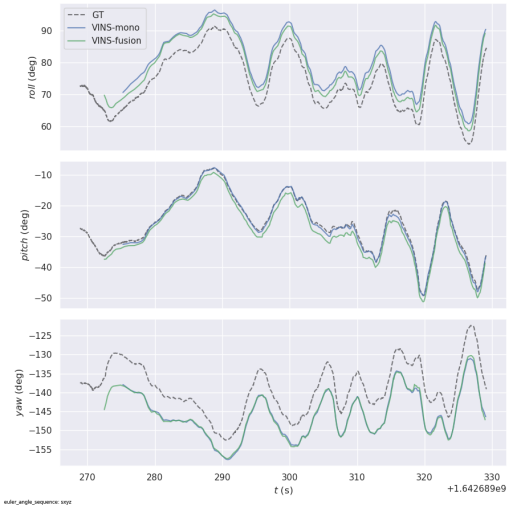
## hdr-normal (left column) and hdr-fast (right column)

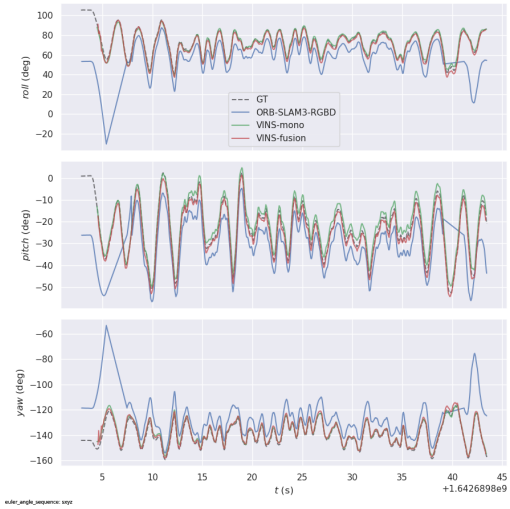


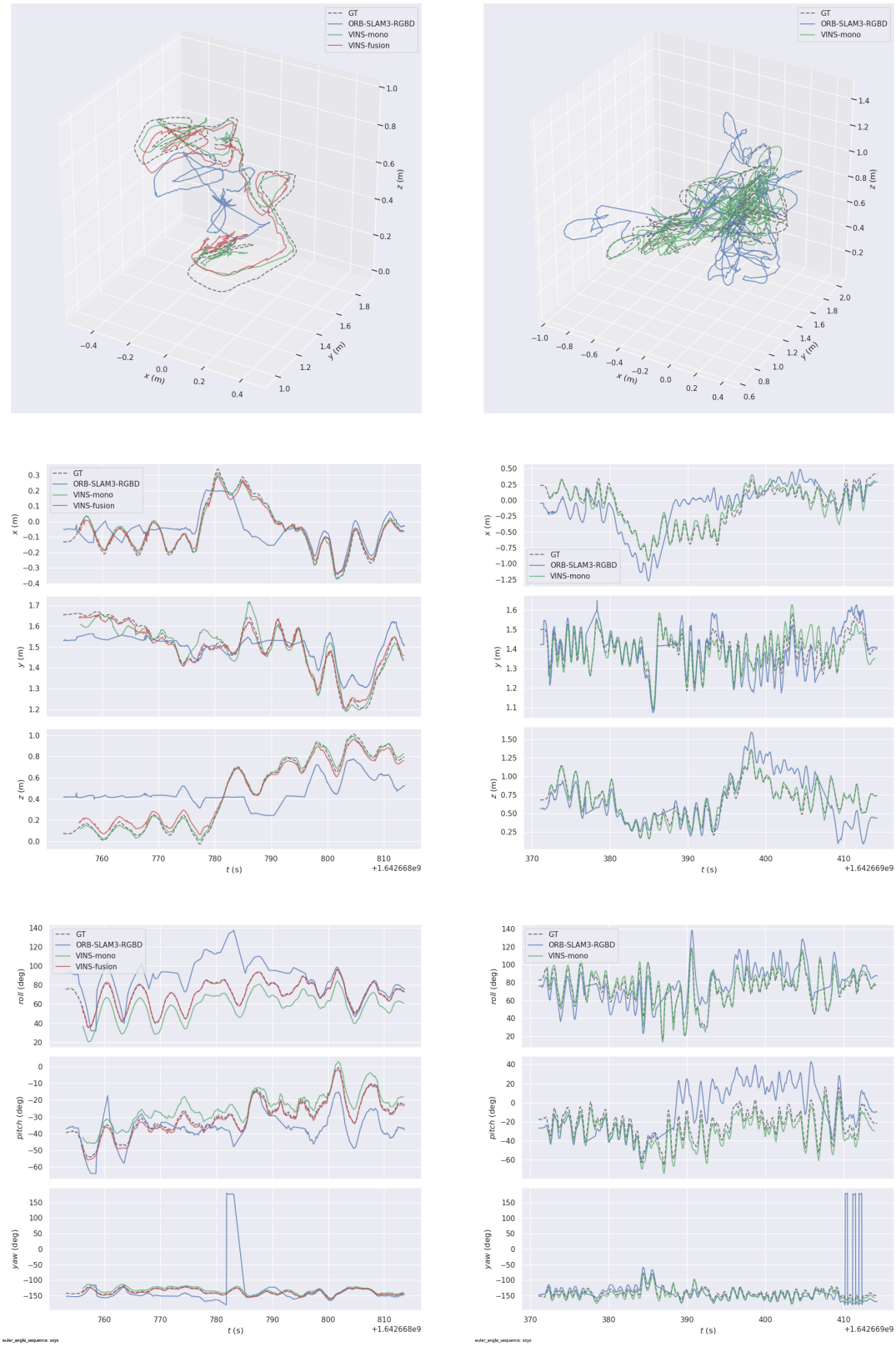



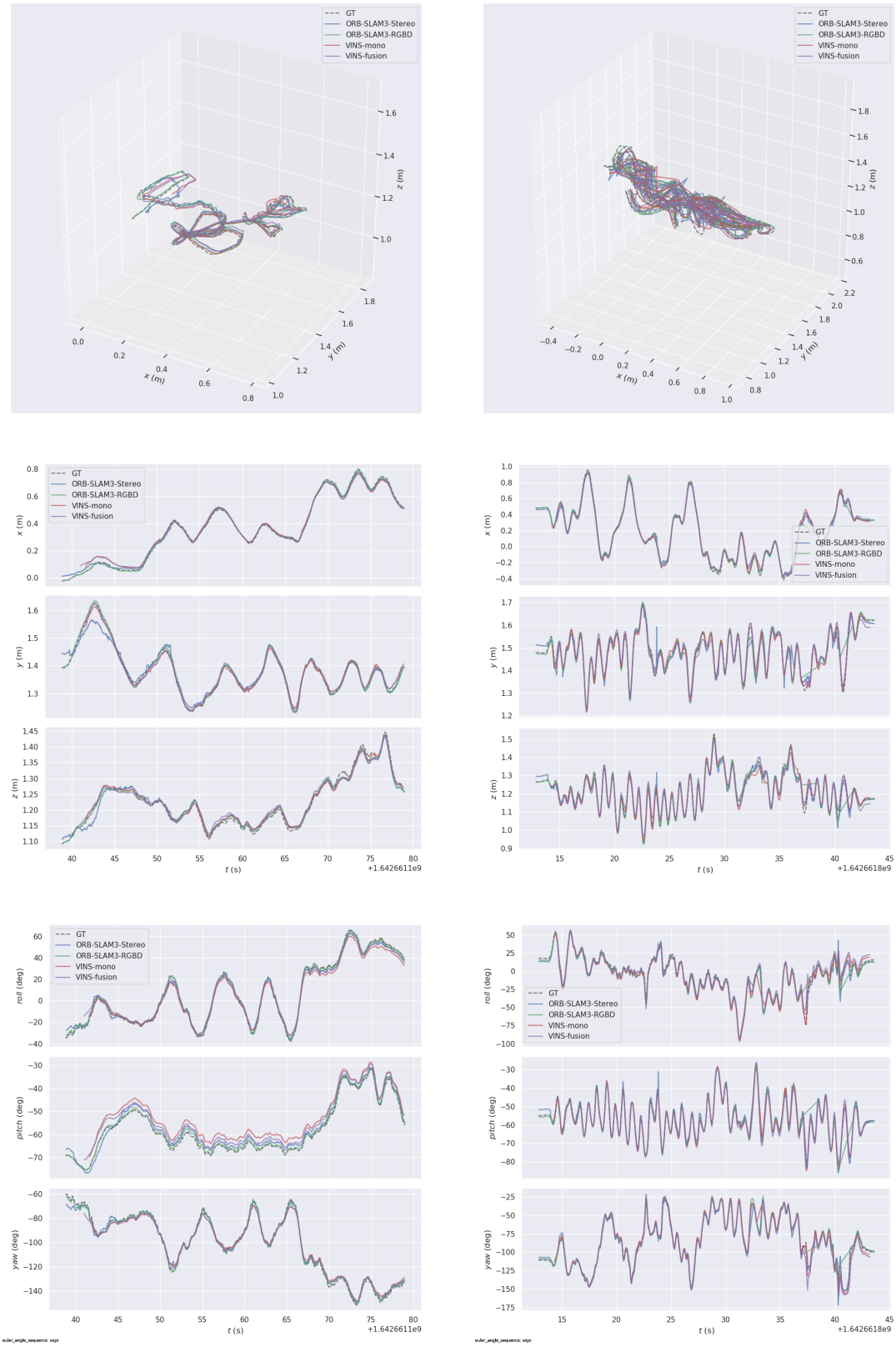



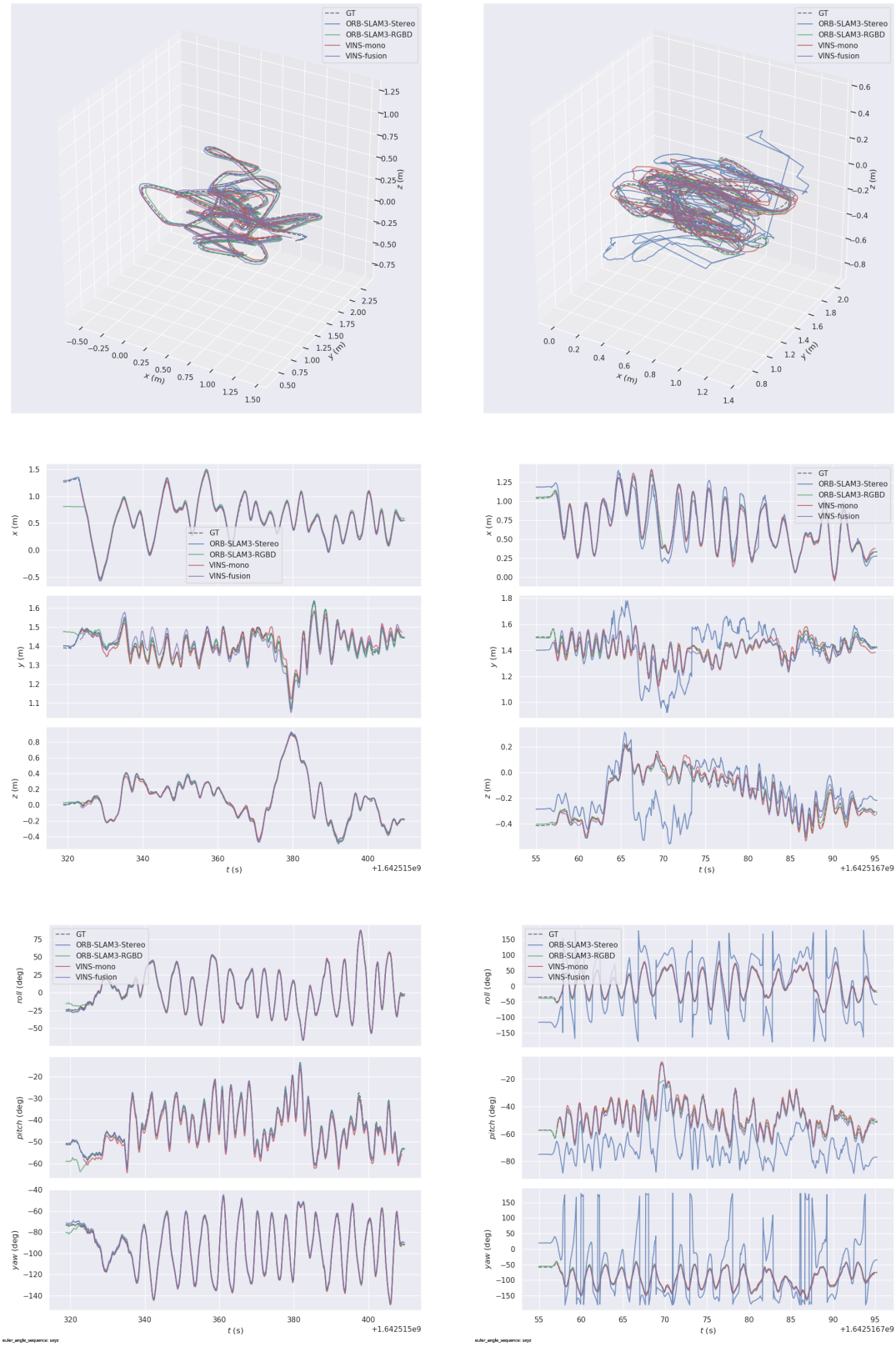
**mountain-normal (left column) and mountain-fast (right column)**



**robot-normal (left column) and robot-fast (right column)**



**sofa-normal (left column) and sofa-fast (right column)**



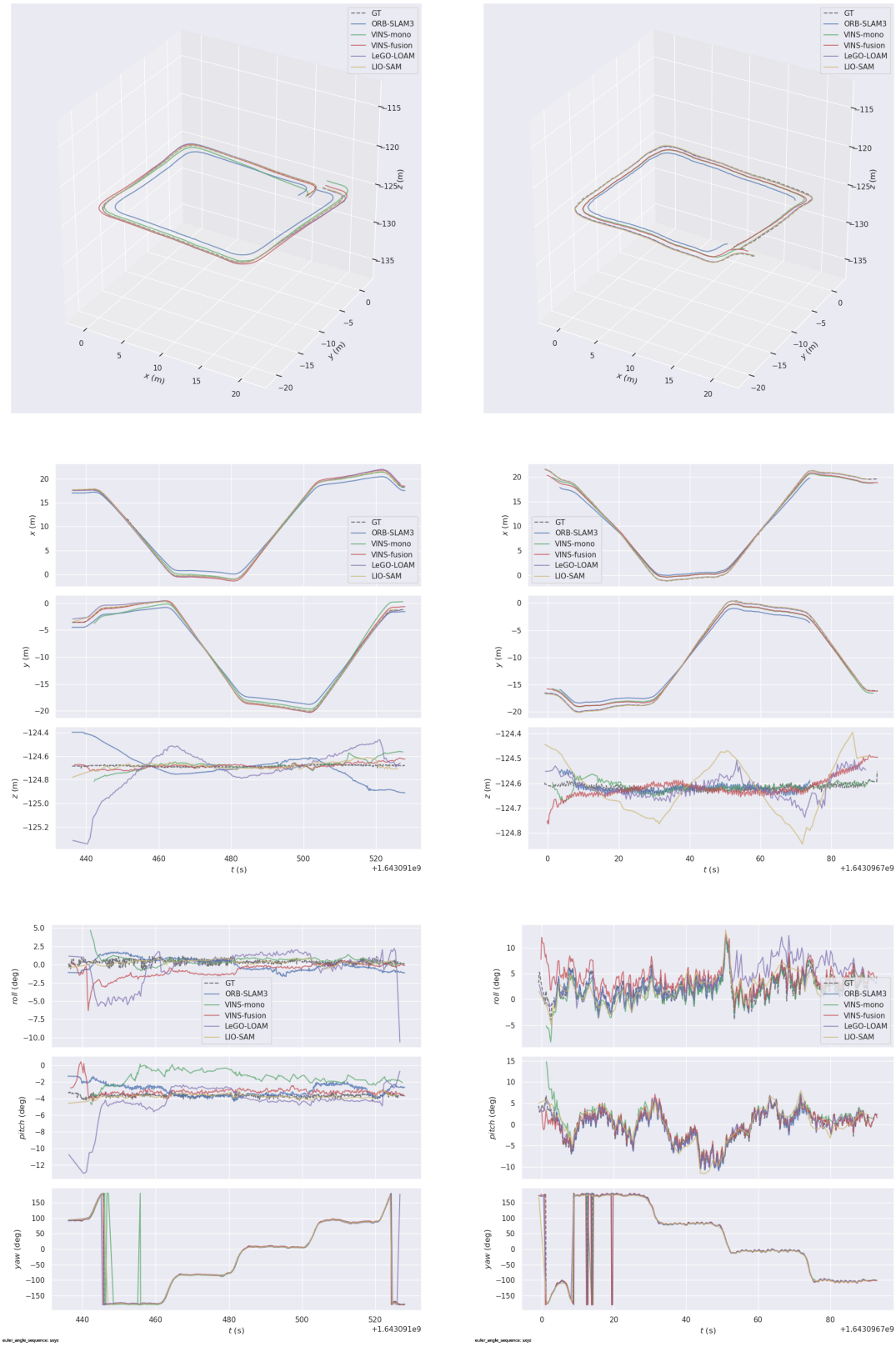
**Large-scale Data Sequences**

Table 2: State-Of-The-Art algorithms' Performance

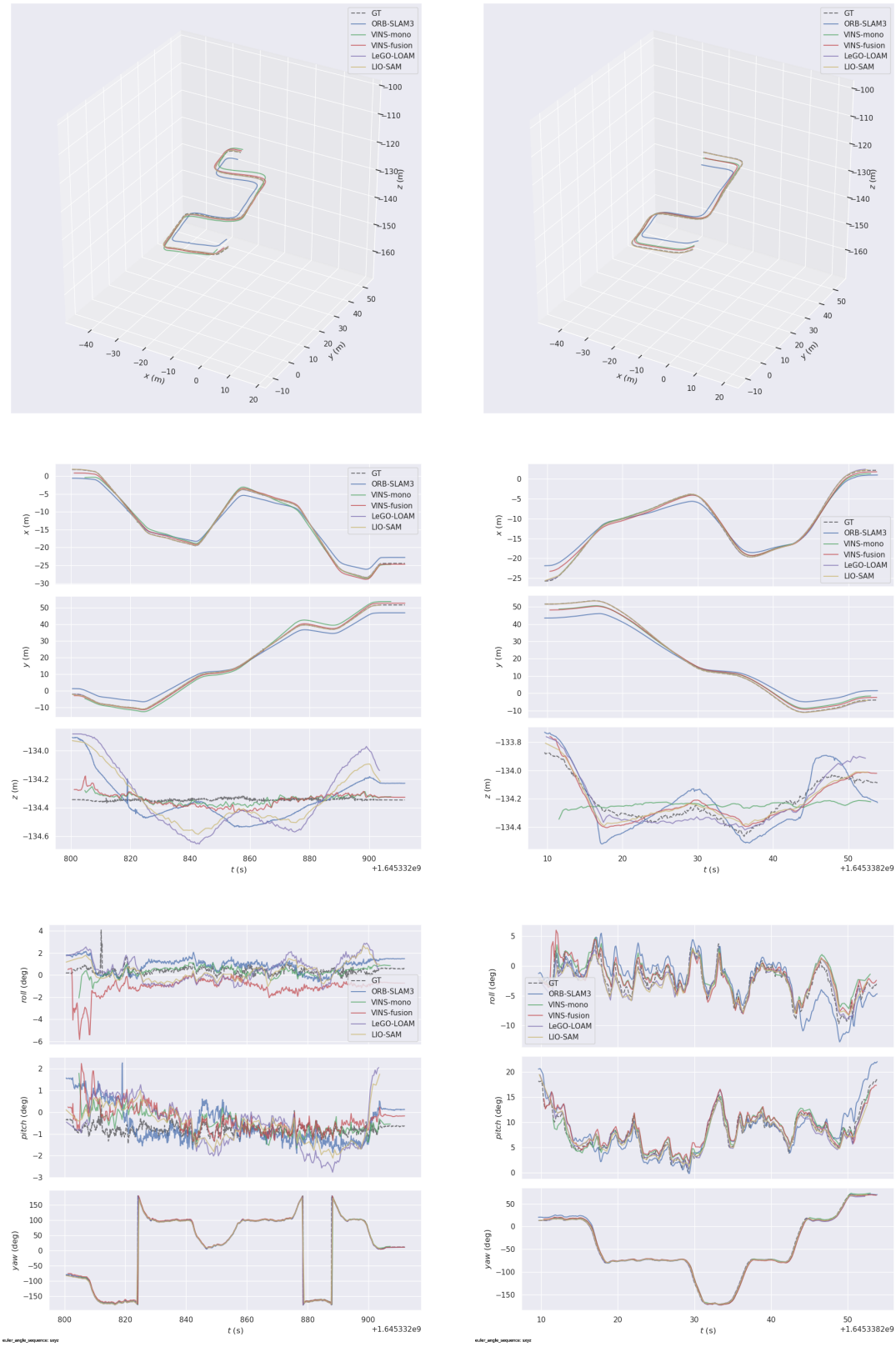
| Data Sequence | ORB-SLAM3 Stereo | VINS mono | VINS fusion | LeGO-LOAM | LIO-SAM |
|---|---|---|---|---|---|
| *corridors-dolly* | ✔ | ✔ | ✔ | ✔ | ✔ |
| *corridors-walk* | ✖ | ✔ | ✔ | ✔ | ✔ |
| *school-dolly* | ✔ | ✔ | ✔ | ✔ | ✔ |
| *school-scooter* | ✔ | ✔ | ✔ | ✔ | ✔ |
| *units-dolly* | ✖ | ✔ | ✔ | ✔ | ✔ |
| *units-scooter* | ✔ | ✔ | ✔ | ✔ | ✔ |

- **ORB-SLAM3-Stereo** takes the readings from the regular stereo camera.
- **VINS-mono** takes the readings from the left regular camera and the IMU.
- **VINS-fusion** takes the readings from the regular stereo camera and the IMU.
- **LeGO-LOAM** takes the readings from the LiDAR.
- **LIO-SAM** takes the readings from the LiDAR and the IMU.
- ✔ indicates a good result performed by this algorithm.
- ✔ indicates an average result performed by this algorithm.
- ✖ indicates an incomplete result performed by this algorithm.
- ✖ indicates no result can be generated by this algorithm.
- All trajectories are first transformed to the same reference frame as the ground truth poses by extrinsics, then further aligned with all poses by Umeyama's SE(3) method.

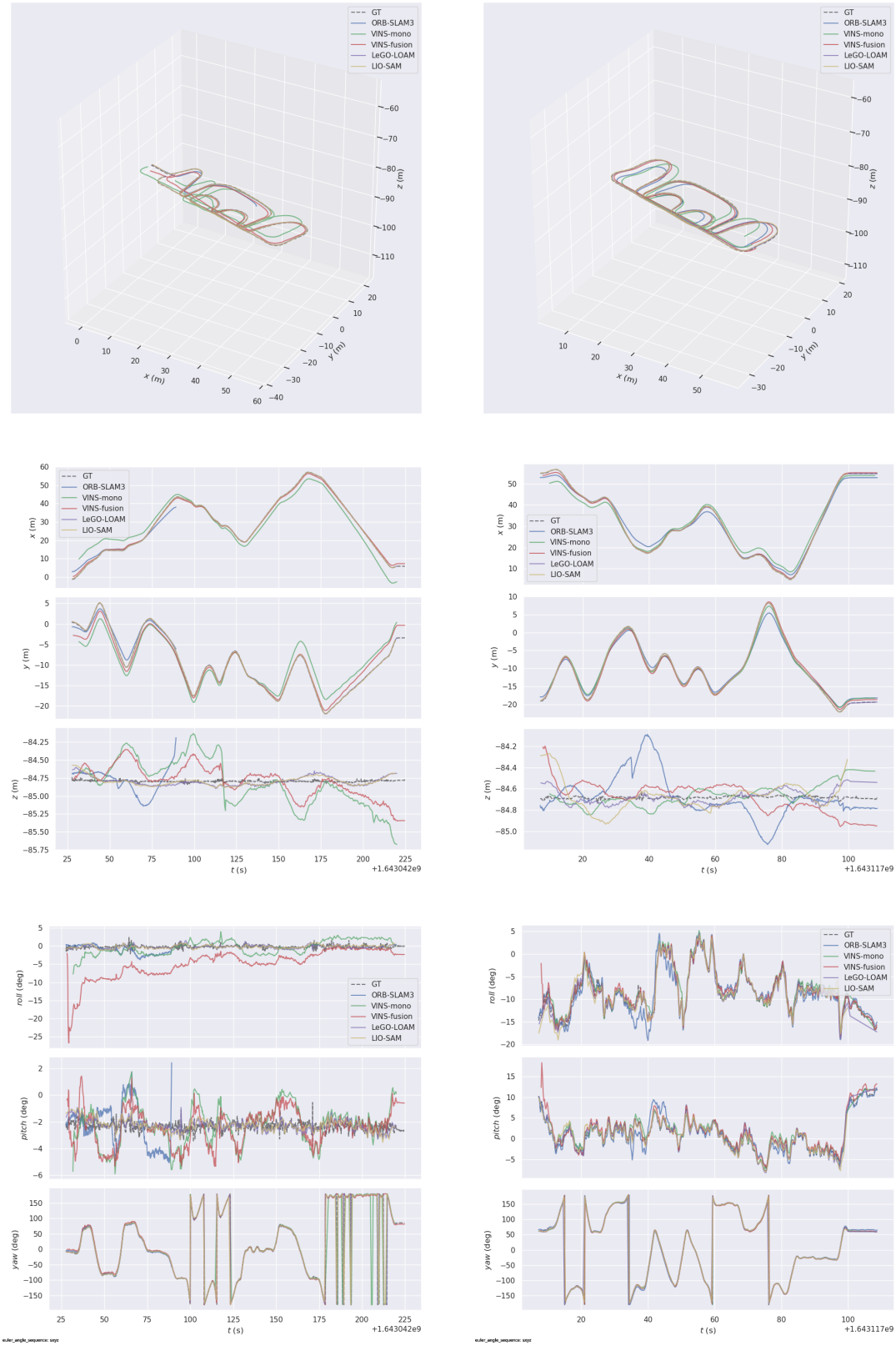
**corridors-dolly (left column) and corridors-walk (right column)**



**school-dolly (left column) and school-scooter (right column)**



**units-dolly (left column) and units-scooter (right column)**



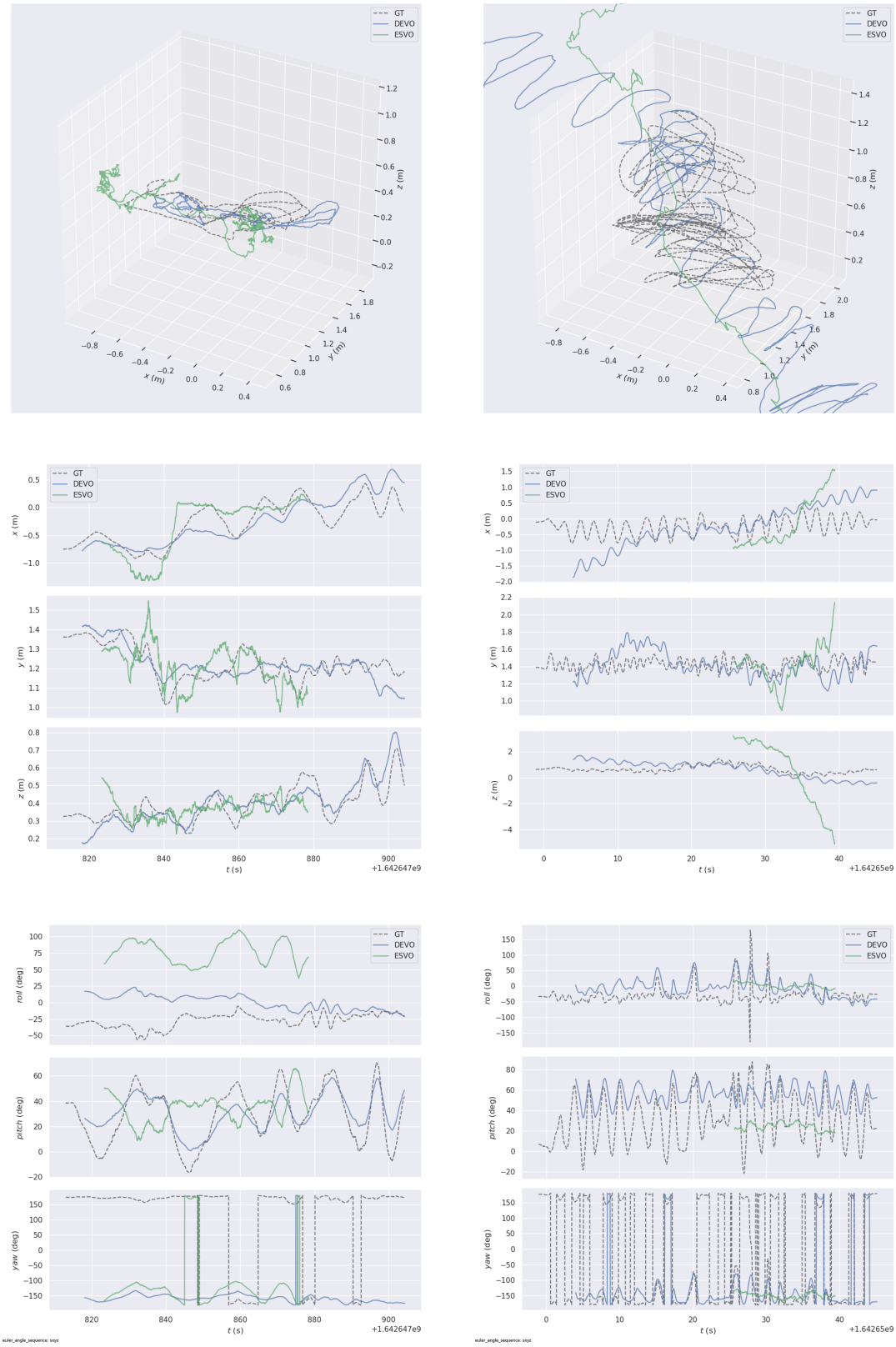
**Small-scale Data Sequences**

Table 3: Event-based algorithms' Performance

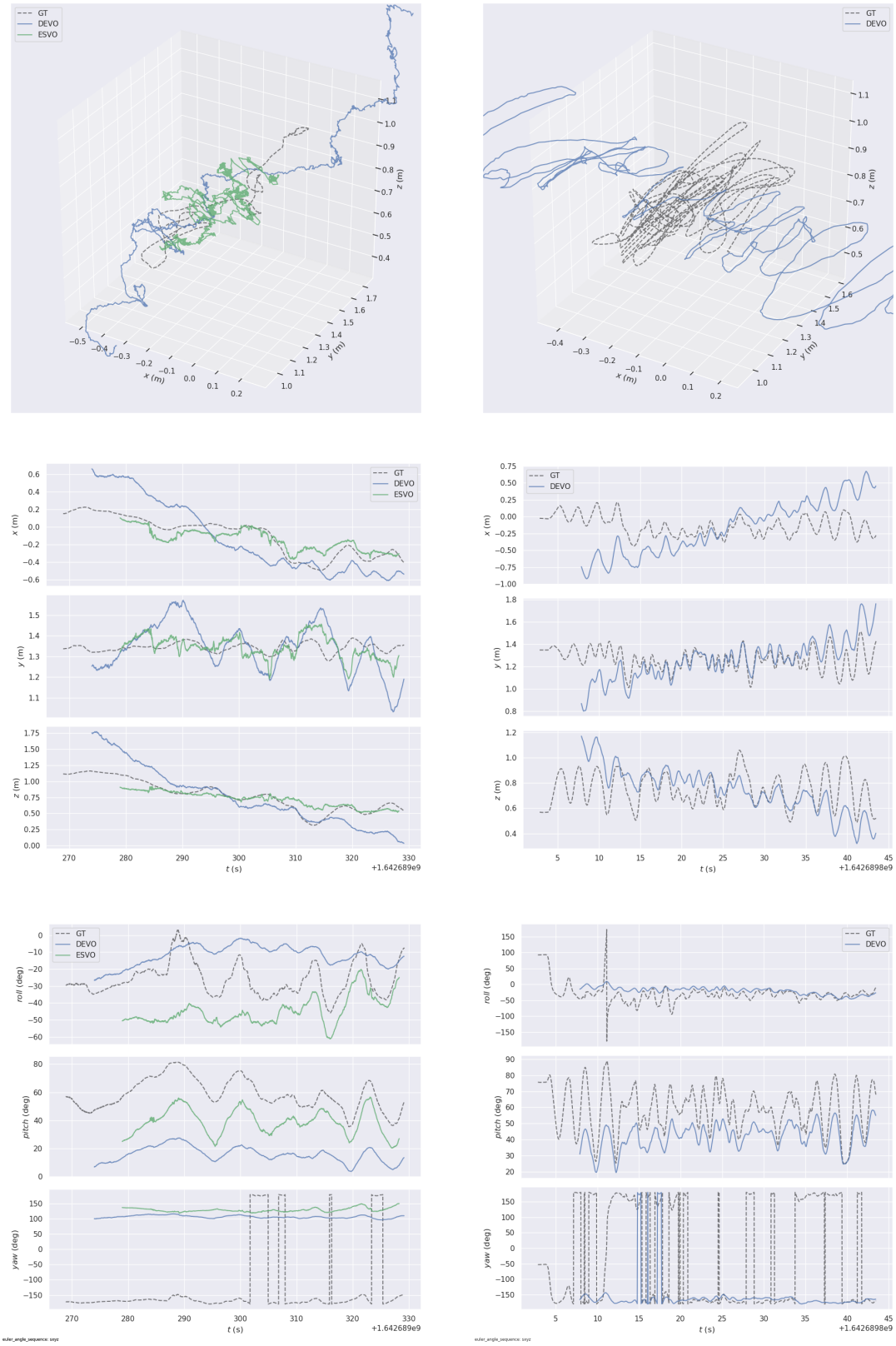
| Data Sequence | DEVO | ESVO |
|---|---|---|
| *desk-normal* | ✖ | ✖ |
| *desk-fast* | ✖ | ✖ |
| *hdr-normal* | ✖ | ✖ |
| *hdr-fast* | ✖ | ✖ |
| *mountain-normal* | ✖ | ✖ |
| *mountain-fast* | ✖ | ✖ |
| *robot-normal* | ✖ | ✖ |
| *robot-fast* | ✖ | ✖ |
| *sofa-normal* | ✖ | ✖ |
| *sofa-fast* | ✖ | ✖ |

- **DEVO** takes the readings from the left event camera, and the depth readings reprojected onto the left event frame.
- **ESVO** takes the readings from the event stereo camera.
- ✖ indicates a relatively poor, yet complete result performed by this algorithm.
- ✖ indicates an incomplete result performed by this algorithm.
- ✖ indicates no result can be generated by this algorithm.
- All trajectories are first transformed to the same reference frame as the ground truth poses by extrinsics, then further aligned with all poses by Umeyama's SE(3) method.

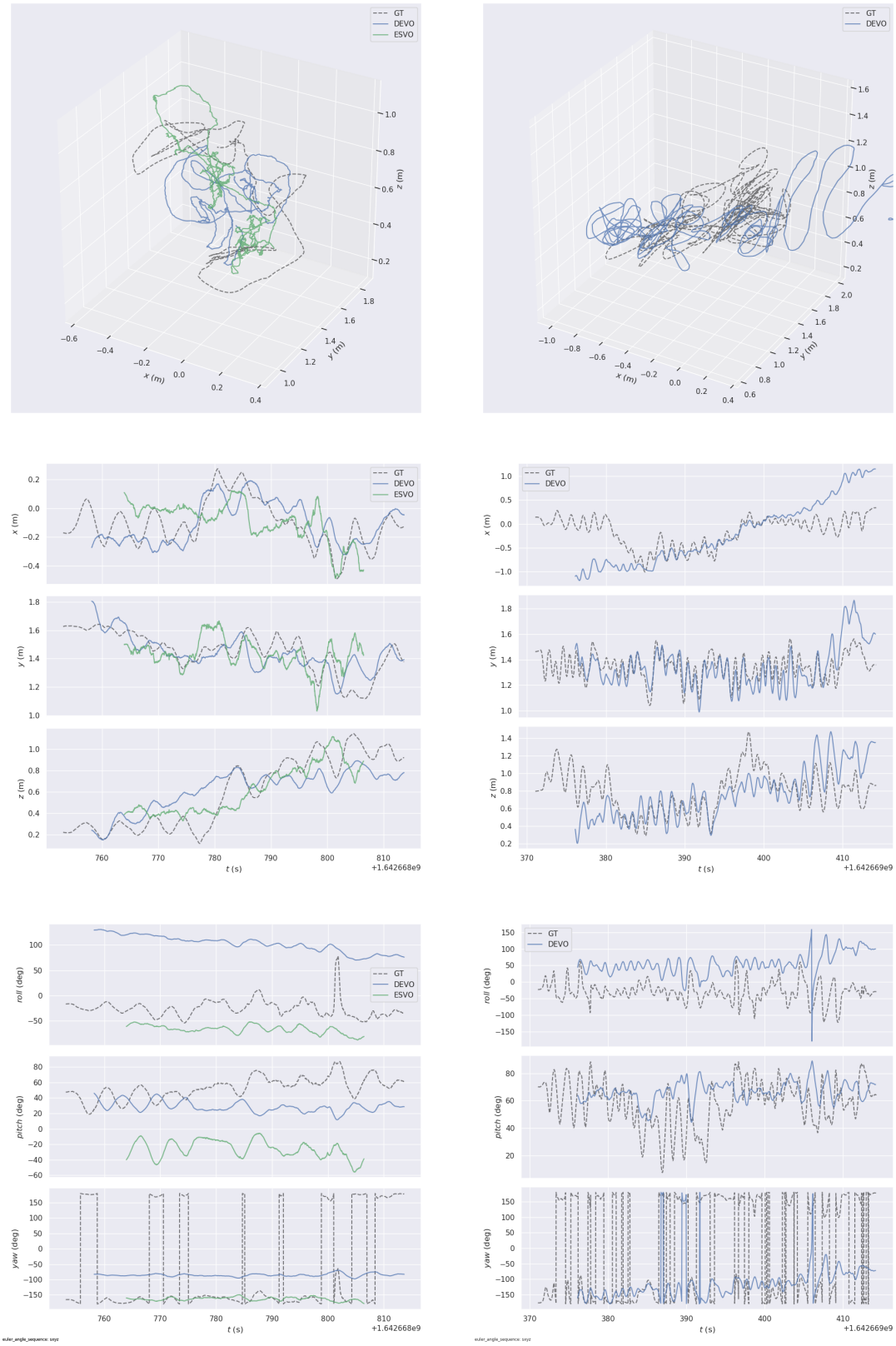
**desk-normal (left column) and desk-fast (right column)**



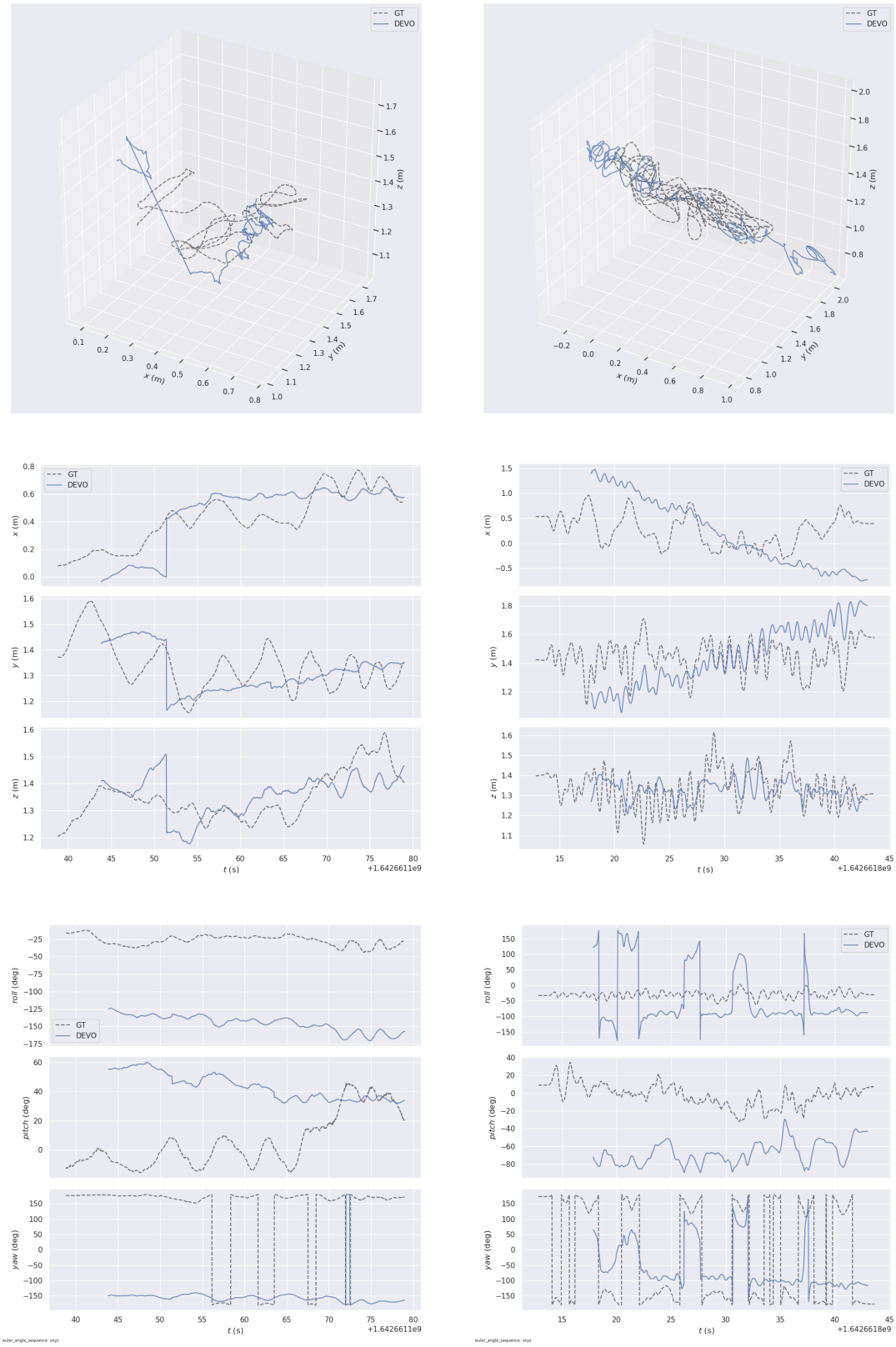
## hdr-normal (left column) and hdr-fast (right column)



**mountain-normal (left column) and mountain-fast (right column)**



**robot-normal (left column) and robot-fast (right column)**



**sofa-normal (left column) and sofa-fast (right column)**