



Multi-sensor Point Cloud Fusion

Scene perception is undoubtedly dependent on the correctness and completeness of the acquired data. As such, 3D data form of the scene present a comprehensive outlook from the environment. Capturing complete 3D information can help to extract features that are highly useful for vision tasks, like 3D object detection in autonomous driving and robotics. To this end, there are various technologies to record the 3D data, *i.e.* LiDAR, time-of-flight (ToF) cameras, and stereo devices. Nevertheless, to obtain a complete 3D form of the scene, these technologies are not sufficient, yielding sparse, incomplete and unevenly distributed point clouds.

Fusion of multiple sensors is a workaround to tackle this problem. Accordingly, this thesis aims to capture data from three sensors and fuse their information to form more reliable data. Namely, there are three Azure Kinect cameras mounted in a full horizontal baseline with known displacements from each other. After recording the data, fusion of point clouds from different pairs is investigated and the final data is projected to a 2D image plane to create the depth map. This map can be compared against the depth that is already recorded from the sensors. For point cloud fusion, which requires the registration as the primary step, the standard ICP (Iterative Closest Point) and some deep learning-based methods, like 3DRegNet [1] or Deep Global Registration [2] can be used. Cross-source fusion (multiple types of devices) is beyond the scope of this thesis.

[1] “3DRegNet: A Deep Neural Network for 3D Point Registration”, G. Pais, *et al.*, CVPR 2020.

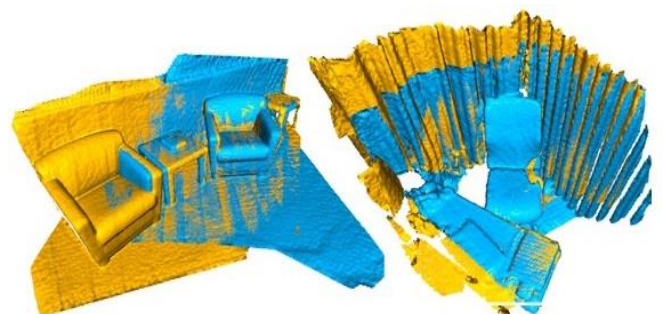
[2] “Deep Global Registration”, C. Choy, *et al.*, CVPR 2020.

Requirements:

- Python programming
- Knowledge of DNN

Contact

Dr. Faranak Shamsafar
Sand 1, Room A304
Tel. (07071) 29-70574
faranak.shamsafar@uni-tuebingen.de



Source: <https://arxiv.org/abs/2004.11540>