



# High Accuracy Depth Data for Indoor Scenarios

The necessity for scene perception in 3D in various computer vision tasks has brought up different technologies. LiDARs, time-of-flight (ToF) cameras and stereo devices are the main three categories as 3D sensors. While LiDARs are the preferred devices for outdoors environments, ToF cameras are the common sensors for indoor settings. Stereo devices can operate in almost any setting but with difficulties in poorly-textured areas, repetitive patterns and occlusion. Nevertheless, even in the best suited setting for these techniques, 3D data is corrupted with noise, erroneous data or missing points. Therefore, fusion of these technologies is a more sophisticated approach to capture data with high accuracy and confidence.

In this thesis, fusion of passive stereo and active ToF 3D data is investigated. Namely, there are two Microsoft Azure Kinect cameras displaced with a fixed baseline. The Azure Kinect is an RGB-D sensor that perceives depth by casting illumination in the near-IR (NIR) spectrum onto the scene and measuring the time as a reference for depth. However, lighting variations, color, material and distance of objects can cause problems in correct estimation. To improve the depth map, the estimated disparity from the stereo setup can be exploited. The goal of this thesis is to combine two depth maps from Azure Kinect cameras and the disparity map from their triangulation for high accuracy data capture. The multimodal combination is investigated by both traditional methods and deep learning-based approaches. For pre-training the deep model, SYNTH3 [1] dataset can be used.

[1] "Deep Learning for Confidence Information in Stereo and ToF Data Fusion", G. Agresti, et al., ICCVWS 2017.

Requirements:

- Python programming
- Knowledge of deep learning

## Contact

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

