EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN

Mathematisch-Naturwissenschaftliche Fakultät
Fachbereich Informatik
Kognitive Systeme (Prof. A. Zell)

Bachelorarbeit

Trinocular Stereo Vision

Stereo vision is a method to infer the third dimension of a scene using two 2D images. In particular, by computing the disparity of a matching point in the left and right images, the distance of the point from the camera viewpoint can be estimated. Stereo vision can be applied in a wide range of applications, including autonomous driving, robot navigation and bin picking.

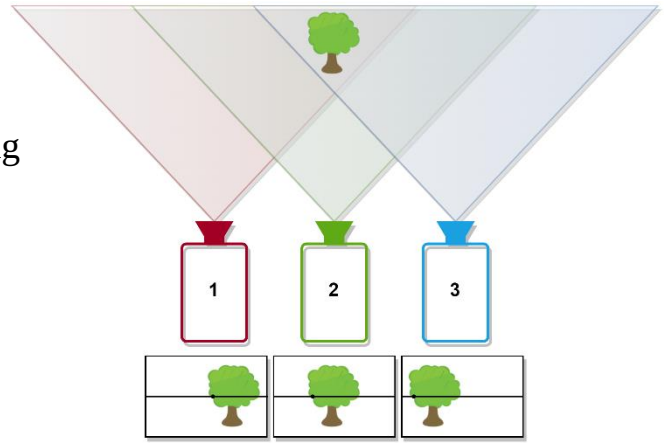
This technique, however, suffers from the challenge of occlusion, where only one camera is able to observe a particular part of the scene. Besides, with a fixed baseline, the range of the estimated depth is limited and the error of depth estimation increases quadratically as the object is more distant.

In this work, stereo vision is investigated by adding a viewpoint from a third camera. By having three cameras, there are shorter and longer baselines, which help in handling the problems of occlusion and lower depth accuracy of distant objects.

As a first step, this research includes capturing images with three synchronized cameras, e.g. by three Azure Kinect cameras, and applying a trinocular rectification on them, so that the disparity amount is limited to the epipolar lines. The cost volume between different image pairs can be computed by a classical matching function (e.g. by Census transform) and merged into a global cost volume. To infer the depth, a classical regularization pipeline, like Semi-Global Matching is adapted. This shall be compared against the attitude in which a deep network is applied on the global cost volume for regressing the disparity.

Requirements:

- Python programming
- Knowledge of deep learning
- Familiar with PyTorch



Contact

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