

Spatio-Temporal Depth Estimation in Real-Time

Project Management and Software Development
for Medical Applications

General Info

Project Title: Depth Estimation with Spatio-Temporal Self-Supervision in Real-Time

Contact Person: Benjamin Busam

Contact Email: b.busam@framos.com

Project Abstract

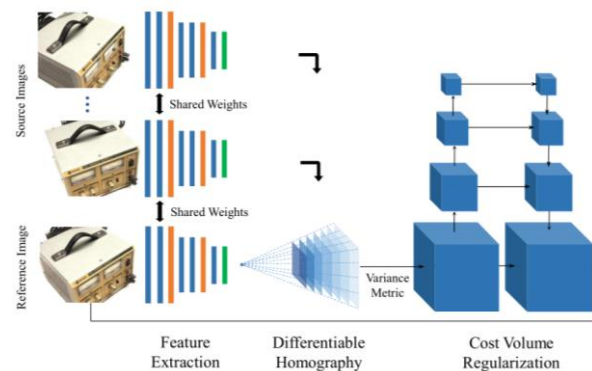
Using a modern deep learning pipeline based on the seminal works StereoNet and MVSNet, the goal of the project is to establish an end-to-end architecture capable of inferring depth images within a SLAM stream in real-time.

Background and Motivation

Depth estimation is an essential capability for reliable 3D vision systems reaching from applications in the operating room and medical tracking systems to autonomously driving vehicles.

This project leverages the capabilities of recent SLAM approaches (e.g. R. Mur-Artal, 2017) to improve the depth estimation of temporally connected video sequences.

Through a SLAM pipeline, the relative pose between frames can be determined. With these transformations, the method of differentiable homography warping (see figure) is used to build a 3D cost volume upon the current camera frame. A guided hierarchical refinement stage (S. Khamis, 2018) helps predicting the depth map while maintaining thin structures in real-time.



[Y. Yao. 2018.](#)

Student's Tasks Description

After getting familiar with the pipelines of MVSNet and (Active)StereoNet, the first target for the student is to fuse the pipelines in a common architecture and to implement supervised training on synthetic data from the CARLA simulator (<http://carla.org>).

Accuracy and speed are tested, and the architecture is refined. The pipeline is evaluated on publically available real-world datasets and extended to learn the scale in a monocular SLAM scenario. The results are compared against depth-from-mono approaches as well as recent multi-modal odometry pipelines.

Finally, a system prototype is set up to visualize the results and to allow for real-time illustration of the principle.

After this project, the student is familiar with TensorFlow and has an overview over recent self-supervised depth estimation methods. Moreover, the constant coordination with FRAMOS gives some insight into the working procedures in the industry.

Please send the completed proposal to b.busam@framos.com. Please note that this proposal will be evaluated by the BMC coordinators and will be assigned to a student only in case of acceptance.



Technical Prerequisites

- Knowledge of TensorFlow and Python
- Basic skills in linear algebra

Optional requirements that can also be learned during the project are:

- Ability to work with versioning tools such as GIT

References

S. Khamis et al. "StereoNet: Guided Hierarchical Refinement for Real-Time Edge-Aware Depth Prediction". ECCV 2018.

Y. Zhang et al. „ActiveStereoNet: End-to-End Self-Supervised Learning for Active Stereo Systems“. ECCV 2018.

Y. Yao et al. "MVSNet: Depth Inference for Unstructured Multi-view Stereo". ECCV 2018.

R. Mur-Artal et al. „ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras“. ToR 2017.