Semantic Segmentation of Urban Video Sequences

Bachelor Thesis / Master Project

Description:
Accurate 3D tracking in urban environments is becoming more and more important, in particular for applications such as outdoor augmented reality, autonomous driving and mobile robotics. However, since most tracking approaches suffer from errors that accumulate over time (drifting), we want to exploit semantic image information in order to compensate for such inaccuracies. In this project, we want to develop a CNN-based segmentation model that identifies different classes related to buildings in urban environments (e.g., facade, edges surrounding the facade, background). Specifically, the model should be applied to video sequences for training and testing, taking temporal consistency into account.

Objective:
- Review literature about LSTM (Long Short Term Memory) networks
- Choose a deep learning framework (e.g., Caffe\(^1\) or TensorFlow\(^2\)) and implement an LSTM-layer
- Create a CNN-based segmentation model that takes temporal consistency via LSTM into account

Qualifications:
- Experience in Python
- Interest in deep learning
- Interest in video segmentation

Contact ICG:
Martin Hirzer
hirzer@icg.tugraz.at

\(^1\)caffe.berkeleyvision.org
\(^2\)www.tensorflow.org