The multi-task learning network for panoptic segmentation and 3D shape reconstruction (EN)

Nienburger Straße 1, 30167 Hannover
Fakultät für Bauingenieurwesen und Geodäsie

Institut für Photogrammetrie und GeoInformation

Prof. Dr.-Ing. habil. Christian Heipke

Tuan Nguyen

Tel. +49 511 762 2488
Fax +49 511 762 2483
E-Mail:
tuan.nguyen@ipi.uni-hannover.de

The perception network aims at the development of methods for improving collaborative positioning of vehicles in areas without adequate GNSS coverage on the basis of stereo image sequences. Focusing on an improvement of the perception of the environment, the poses and shapes of vehicles in the surrounding of the ego-position are of particular interest. Thus, the CNN-based method integrates the detection and semantic classification of objects via panoptic segmentation (Kirillov, Alexander, et al., 2019) with the reconstruction of the 3D geometry of these objects. Assuming that the scene is dynamic, with respect to ego-motion as well as moving objects, temporal information is considered to be beneficial and is taken into account in the network. The end-to-end architecture is depicted in Figure 1, where the red rectangle represents the anticipated output generated by the network.

In this context, the end-to-end deep network can not only estimate the panoptic segmentation mask (Nguyen et al., 2024) but also reconstruct the 3D shape of an object using the Active Shape Model (ASM) approach (Coenen et al., 2019). A multi-task learning network is an efficient framework for this purpose, utilizing a shared backbone to process the input and generate outputs through multiple branches. Initially, the panoptic segmentation method should be developed as a baseline, followed by the addition of an extra branch to predict the 3D shape of a car.

While several public datasets and methods for panoptic segmentation are available, there is a notable lack of 3D shape models compatible with ASM in these datasets. To address this gap, our goal is to create 3D shape model annotations on public datasets (e.g., KITTI) for training the network. This annotation process could be fully automated using LiDAR data.

This thesis will be supervised by Tuan Nguyen.

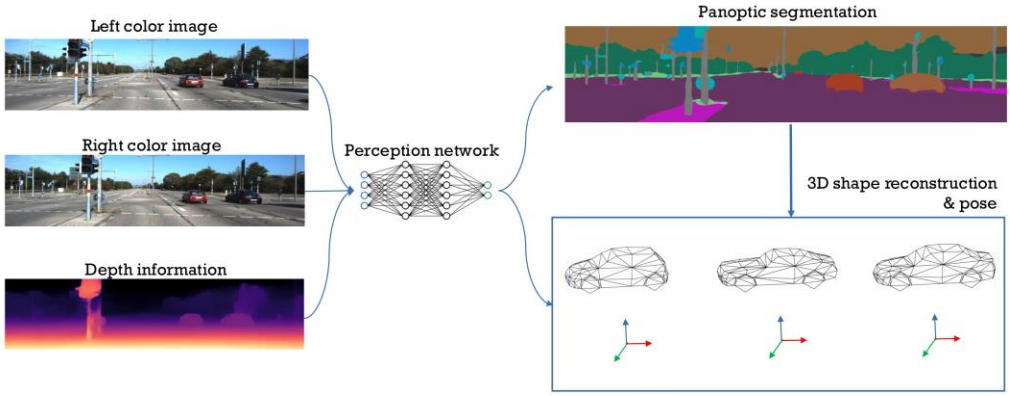

Figure 1. The perception network with an input as sequence of color and depth image.

Besucheradresse:
Nienburger Straße 1
30167 Hannover
www.ipi.uni-hannover.de

References

1. Kirillov, A., He, K., Girshick, R., Rother, C. and Dollár, P., 2019. Panoptic segmentation. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition (pp. 9404-9413).
2. Coenen, M., and Franz, R., 2019 . Probabilistic vehicle reconstruction using a multi-task CNN. Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops.
3. Nguyen, T. and Mehltretter, M. and Rottensteiner, F., 2024. Depth-Aware Panoptic Segmentation. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences (pp. 153-161).