Nikolas Müller

SELF-SUPERVISED DETECTION AND POSE ESTIMATION OF LOGISTICAL OBJECTS IN 3D SENSOR DATA

Challenges

- High Variety
- Indefinite Amount
- Occlusions
- Symmetries
- Complex features
- Difficult data acquisition



Challenges 3D Sensor Data

- Better Receptive Field for Convolutions
- Better Generalization
- Simpler Sensor Model
- \rightarrow Self-Supervision from Simulation

3D Sensor Data Scene Construction

- Stochastic Process
- Reference already placed objects
- Maintain object relations



3D Sensor Data Scene Construction Scene Capturing

• Depending on Use-Case

• Model effects like noise and missing measurements



• KPFCNN [1] with replicated Decoder

Scene Construction
Scene Capturing
Architecture



[1] Thomas, Hugues, et al. "Kpconv: Flexible and deformable convolution for point clouds." Proceedings of the IEEE International Conference on Computer Vision. 2019.

Architecture Class Partitioning

Split Segmentation into classified point sets



Class Partitioning Location Vote Clustering

- Using DBSCAN for performance and noise compensation
- Parameters can be learned with Bayesian Optimization
- Performed for each class separately



Class Partitioning Location Vote Clustering Orientation Point Vote Matching

- Using DBSCAN for performance and noise compensation
- Parameters can be estimated with Bayesian Optimization
- Performed for each class separately



Vote Clustering Orientation Point Vote Matching <u>Results</u>

evaluated on 4000 randomly generated scans



[2] He, Yisheng, et al. "PVN3D: A deep point-wise 3D keypoints voting network for 6DoF pose estimation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020.

Orientation Point Vote Matching Results Summary

- Self-supervised training with Stochastic Scene Construction
- Novel Fully Convolutional Voting Network on Point Clouds
- Novel Postprocessing with (H)DBSCAN(*) and Vote Matching
- Compensation for Noise, Symmetries and Occlusions
- Sensor Independency
- Outperforms pure 3D variant of current state-of-the-art