

SceneCAD: Predicting Object Alignments and Layouts in RGB-D Scans

Supplementary Material

1 Dataset

In this supplemental document, we provide additional details for our layout dataset that is used for training. Figure 1 shows an illustration of the real-world layout annotations comprising of more than 1000 individual ScanNet [1] scenes. In addition to layouts from real-world scans, we also extract layouts from the synthetic SUNCG dataset [2]; see Figure 2. From these dataset annotations, we create ground truth targets, as visualized in Figure 3, for our hierarchical layout estimation training.

In Figure 4, we illustrate visually the different kinds of object-to-layout relationship classes. Note that an object can have relationships with multiple layout elements.

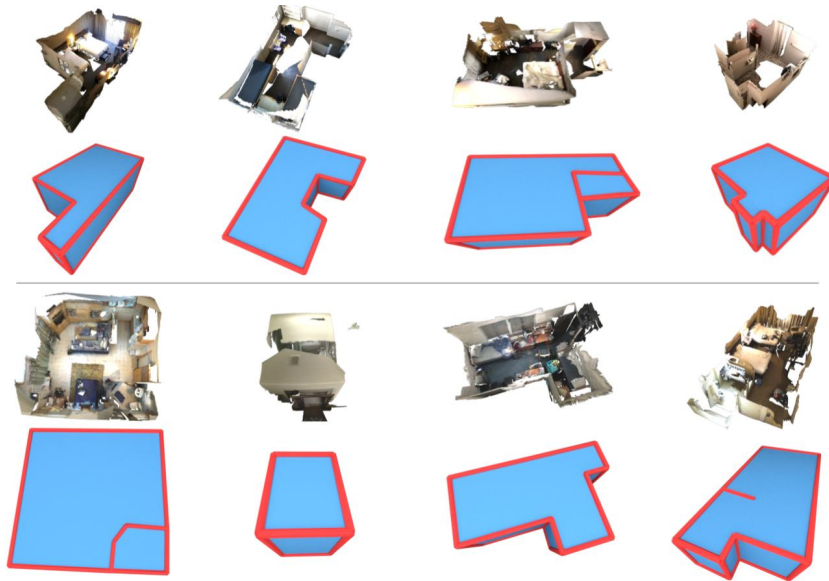


Fig. 1. Samples of manually annotated layouts on ScanNet [1]. Annotations include wireframes and room-level CAD shells of the scenes representing walls, floors, and ceilings.

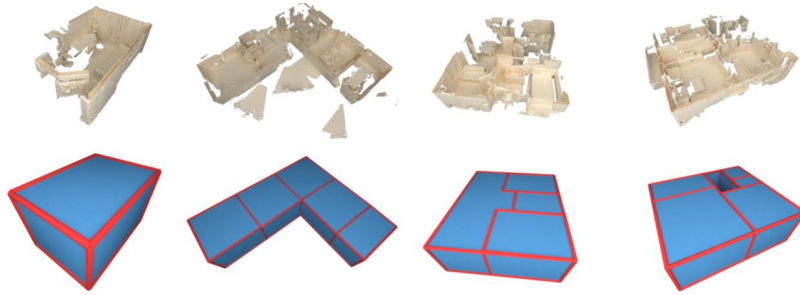


Fig. 2. Samples of automatically parsed layouts from SUNCG [2]. Layouts include wireframes and room-level CAD shells.

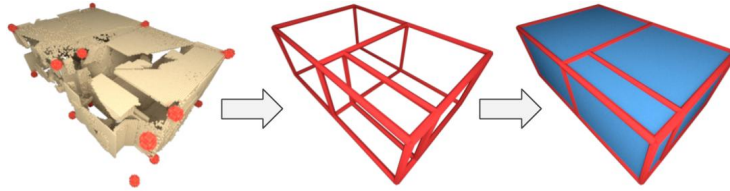


Fig. 3. Sample targets of the layout estimation. The pipeline starts with a corner point estimation (left). Then, valid edges are estimated from the detected corners producing a wireframe (middle). Finally, valid layout quads are predicted from the edge candidates.

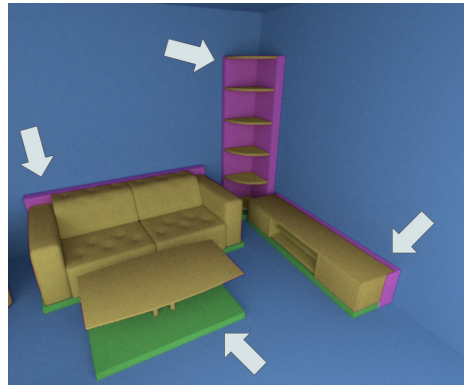


Fig. 4. Groundtruth sample from the dataset. Green tile correspond to *vertical support* relationship and pink tile correspond to *horizontal touch* relationship.

References

1. Dai, A., Chang, A.X., Savva, M., Halber, M., Funkhouser, T., Nießner, M.: Scannet: Richly-annotated 3d reconstructions of indoor scenes. In: Proc. Computer Vision and Pattern Recognition (CVPR), IEEE (2017)
2. Song, S., Yu, F., Zeng, A., Chang, A.X., Savva, M., Funkhouser, T.: Semantic scene completion from a single depth image (2017)