Monocular 3D Object Detection for indoor objects

Computer Vision (CS6350)

TPA-6

Problem statement:

3D object detection is a basic problem in many applications such as autonomous navigation and house-keeping tasks in robotics. In the real world, robotics involves decision-making and interaction with objects which involve the identification and localization of the object. Hence, in this assignment, the task is to extract 3D object bounding boxes and meshes from a single image of an indoor scene.

Input:

Figure 1:Single RGB Indoor image (input)

Output:

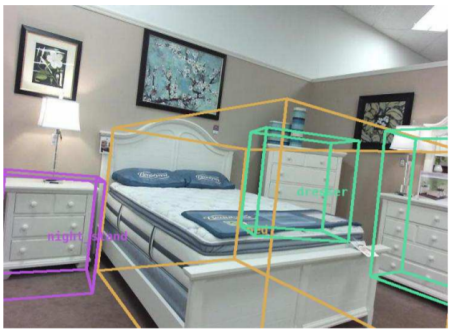

Figure 2: 3D Object bounding boxes

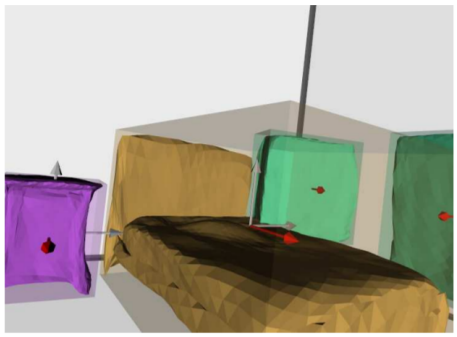

Figure 3:Reconstruct object meshes

Another Example:

(a) *(b)* *(c)*

Figure 4: Input (a) ,3D objects bounding boxes (b), Reconstruct object meshes (c)

Dataset:

1. SUN RGB-D: A RGB-D Scene Understanding Benchmark Suite - https://rgbd.cs.princeton.edu/
2. Pix3D: Dataset and Methods for Single-Image 3D Shape Modeling - http://pix3d.csail.mit.edu/

References:

[1]. Nie, Yinyu, et al. "Total3dunderstanding: Joint layout, object pose and mesh reconstruction for indoor scenes from a single image." *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2020.

[2]. Shuran Song, Samuel P Lichtenberg, and Jianxiong Xiao. Sun rgb-d: A rgb-d scene understanding benchmark suite. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 567–576, 2015

[3]. Xingyuan Sun, Jiajun Wu, Xiuming Zhang, Zhoutong Zhang, Chengkai Zhang, Tianfan Xue, Joshua B Tenenbaum, and William T Freeman. Pix3d: Dataset and methods for single-image 3d shape modeling. In

Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 2974–2983, 2018

[4]. Siyuan Huang, Siyuan Qi, Yixin Zhu, Yinxue Xiao, Yuanlu Xu, and Song-Chun Zhu. Holistic 3d scene parsing and reconstruction from a single rgb image. In Proceedings of the European Conference on Computer Vision (ECCV), pages 187–203, 2018

[5]. Saumitro Dasgupta, Kuan Fang, Kevin Chen, and Silvio Savarese. Delay: Robust spatial layout estimation for cluttered indoor scenes. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 616–624, 2016

[6]. Arun Mallya and Svetlana Lazebnik. Learning informative edge maps for indoor scene layout prediction. In Proceedings of the IEEE international conference on computer vision, pages 936–944, 2015

[7]. Lin Li, Salman Khan, and Nick Barnes. Silhouette-assisted 3d object instance reconstruction from a cluttered scene. In Proceedings of the IEEE International Conference on Computer Vision Workshops, pages 0–0, 2019