Reconstruction and Annotation of 3D Cars

Vishal Batvia, Wen-Yen Yang, Jiacheng Xu and Ali Mahdavi Amiri

Simon Fraser University, CMPT 742

1 Introduction

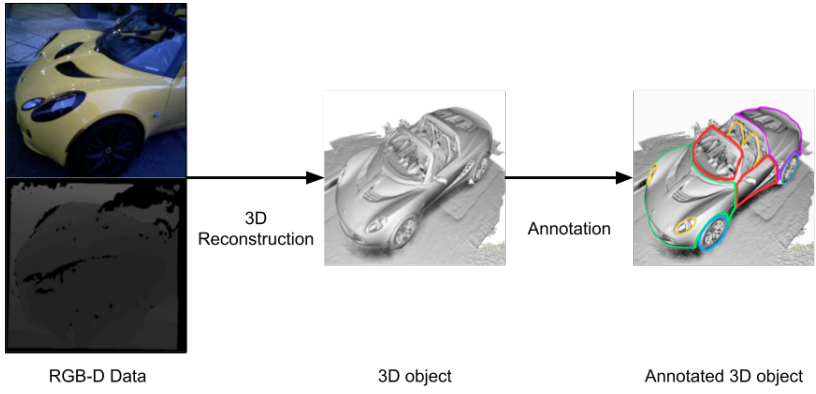


Fig. 1. The proposed workflow of the annotation interface for both reconstruction and annotation of 3D cars

Inferring detailed 3D features of objects is crucial for robotics, computer vision, simulation, and creating any form of 3D content. Thus it is crucial to obtain and create 3D data, which is a challenging problem, and thus, there is limited availability of 3D data. Deep learning models require enormous labelled data to gain high-quality performance, both expensive and time-consuming. The problem elevates further in a self-driving car and other problems related to cars in computer vision because of limited availability of 3D annotated cars dataset. This project proposes an interactive system for annotating 3D cars from both RGB-D and RGB datasets of cars. Our work's key idea is to exploit strong priors that humans have for the 3D world to annotate 3D cars interactively, as shown in fig 1.

2 Proposal

2.1 RGBD Reconstruction

3D reconstruction plays a crucial role in many applications for robotics, simulation self-driving cars and many more. The core of this project revolves around

the reconstruction of 3D cars. We propose to use both RGBD and RGB datasets of cars to reconstruct 3D cars and then annotate different parts of the cars using an interface. In our work, we plan to use a 3D Scan dataset containing 1132 full scans of cars with depth data [1]. For 2D images, we plan to use the cars dataset containing 16,185 images of cars [2]. For 3D reconstruction, we plan to work on Deep Networks like HoloGan and RGBD-GAN to generate 3D structures and opensource tools like Open3D. We plan to test and compare the reconstruction quality using both and produce a good quality reconstruction of the car, as shown in fig 2. For 2D to 3D Reconstruction, we plan to have a baseline from HoloGan and build upon it to have better reconstruction [3].

Open3D [8], an open-sourced library used to deal with 3D data, created by Qian-Yi Zhou and Jaesik Park and Vladlen Koltun will be used for model reconstruction. In this library, two different odometry methods[6] [4] are implemented to find the camera movement between two RGBD images can be used to align the point cloud created by the RGBD data. Figure 1 shows a car model created from 256 consecutive RGBD images with proper alignment.

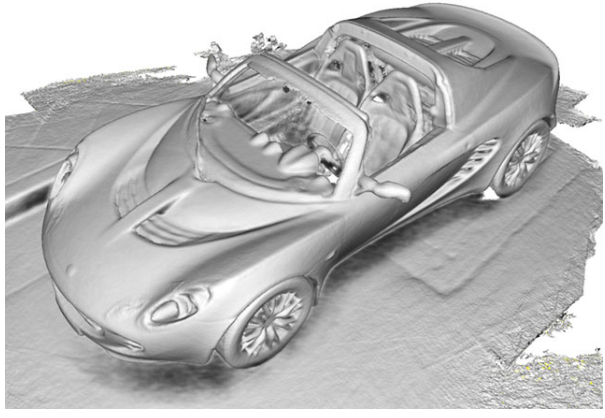
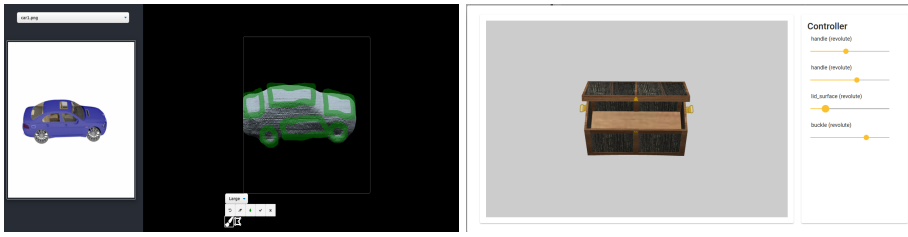


Fig. 2. 3D reconstruction of the car model.

2.2 Annotation System

We want to create an annotation system that allows users to label the category, mobility, and texture. A coarse annotation will be achieved by a clustering method, and then a fine annotation will be finished by a human. As shown in fig 3 we plan to achieve an interface close to **(a)** where we can allow the user to load an RGB/RGB-D image and then annotate the 3D model produced for it using drawing tools.



(a) (b)

Fig. 3. (a) An implementation of 3D reconstruction and human correction [5]. **(b)** A 3D annotation and reconstruction dataset interface with annotation and features attached to the 3D objects [7]

3 Timeline

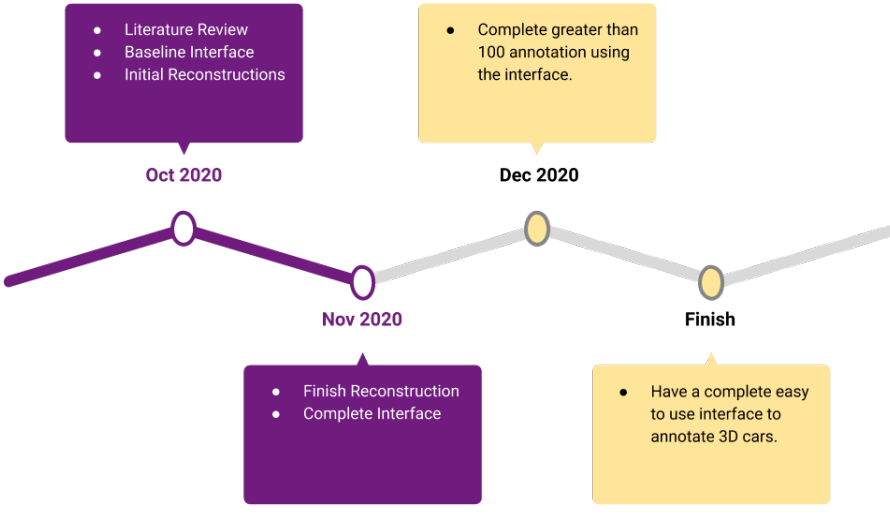


Fig. 4. Project Timeline

As shown in fig 4, we plan to go through the state of the art methods on 3D reconstruction as well as start the work of the annotation interface in the next two weeks. After that, in four weeks, we plan to finish the reconstruction and annotation interface with testing. In the last two weeks, we plan to annotate at least 100 3D cars and save the results.

References

1. Choi, S., Zhou, Q.Y., Miller, S., Koltun, V.: A large dataset of object scans. arXiv:1602.02481 (2016)
2. Krause, J., Stark, M., Deng, J., Fei-Fei, L.: 3d object representations for fine-grained categorization. In: 4th International IEEE Workshop on 3D Representation and Recognition (3dRR-13). Sydney, Australia (2013)
3. Nguyen-Phuoc, T., Li, C., Theis, L., Richardt, C., Yang, Y.L.: Hologan: Unsupervised learning of 3d representations from natural images (2019)
4. Park, J., Zhou, Q.Y., Koltun, V.: Colored point cloud registration revisited. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 143–152 (2017)
5. Shen, T., Gao, J., Kar, A., Fidler, S.: Interactive annotation of 3d object geometry using 2d scribbles (2020)
6. Steinbrücker, F., Sturm, J., Cremers, D.: Real-time visual odometry from dense rgb-d images. In: 2011 IEEE international conference on computer vision workshops (ICCV Workshops). pp. 719–722. IEEE (2011)
7. Xiang, F., Qin, Y., Mo, K., Xia, Y., Zhu, H., Liu, F., Liu, M., Jiang, H., Yuan, Y., Wang, H., Yi, L., Chang, A.X., Guibas, L.J., Su, H.: SAPIEN: A simulated part-based interactive environment. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (June 2020)
8. Zhou, Q.Y., Park, J., Koltun, V.: Open3D: A modern library for 3D data processing. arXiv:1801.09847 (2018)