

Diminished Reality

Mentor: Shun-Cheng Wu
shuncheng.wu@tum.de

Goal Given an input image, this project focuses on removing a selected object from the image while in-paint the removed region so as to keep the image realistic. A possible extension will be to also complete depth information on the removed region.

Description In contrast to augmented reality, this project aims to remove objects from the target scene. This increases the flexibility in many applications. For example, we want to change the layout of a classroom, with this technique, we can remove all the tables and replace them with totally different ones. Moreover, if depth information is available, this project aims to also complete the removed depth information so that the depth map is complete and can be used for detecting collision and applying animations.

This project consists of three steps:

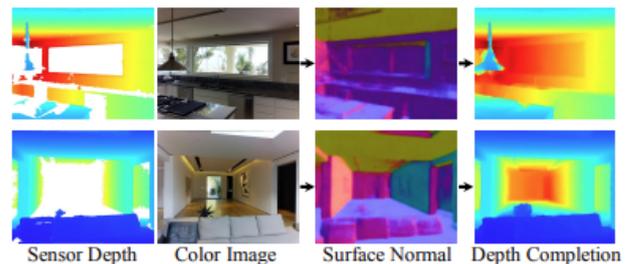
1. Detect an object and its region on the input image.
2. Remove and in-paint the removed region.
3. If a depth map is available, complete or predicts depth information.



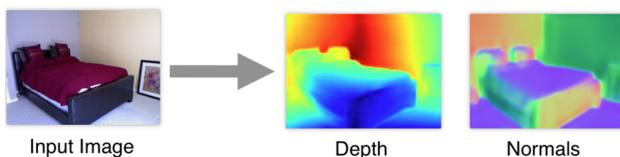
(a) MaskRCNN



(b) Image inpainting



(c) Depth completion



(d) Depth prediction

Here is one possible solution for this project:

- Using a object detection network to identifying object (e.g. Mask-RCNN[2], YOLO[3])
- Using a image in-painting network to recover the missing region (e.g. FreeForm [4]).
- Using a depth completion network [5] or depth prediction network [1] to get a complete depth information.

Requirements

- Basic knowledge in Pytorch (or Tensorflow)
- Basic knowledge in Python
- Basic knowledge in C++

References

- [1] David Eigen and Rob Fergus. Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture. In *Proceedings of the IEEE international conference on computer vision*, pages 2650–2658, 2015.
- [2] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask r-cnn. In *Proceedings of the IEEE international conference on computer vision*, pages 2961–2969, 2017.
- [3] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.
- [4] Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas S Huang. Free-form image inpainting with gated convolution. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 4471–4480, 2019.
- [5] Yinda Zhang and Thomas Funkhouser. Deep depth completion of a single rgb-d image. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 175–185, 2018.