

Supplementary for DepthGAN: GAN-based Depth Generation from Semantic Layouts

1. 3D Scene Generation Process

In this section, we explain the construction process of the point clouds from our generated depth maps. As shown in Fig. 1. For the perspective view, given the generated depth map, we simply use a *pinhole camera model* to reconstruct the point clouds with a given camera intrinsic. And we can fuse the corresponding semantic layout or the generated appearance to the point clouds. Moreover, we show the generated depth quality by projecting the generated depth maps back to the panorama with reverse gnomonic projection [1]. The panorama images naturally contain fixed camera parameters, thus we can construct the point clouds from the depth maps. Note that the panorama is just used to show the accuracy of the generated depth map. Since the panorama nature of Structured3D [2], we do not need any extra information on the panorama in our training and evaluation processes.

References

- [1] K. Tateno, N. Navab, and F. Tombari. Distortion-aware convolutional filters for dense prediction in panoramic images. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 707–722, 2018. 1
- [2] J. Zheng, J. Zhang, J. Li, R. Tang, S. Gao, and Z. Zhou. Structured3d: A large photo-realistic dataset for structured 3d modeling. In *Computer Vision - ECCV 2020 - 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part IX*, volume 12354 of *Lecture Notes in Computer Science*, pages 519–535. Springer, 2020. 1

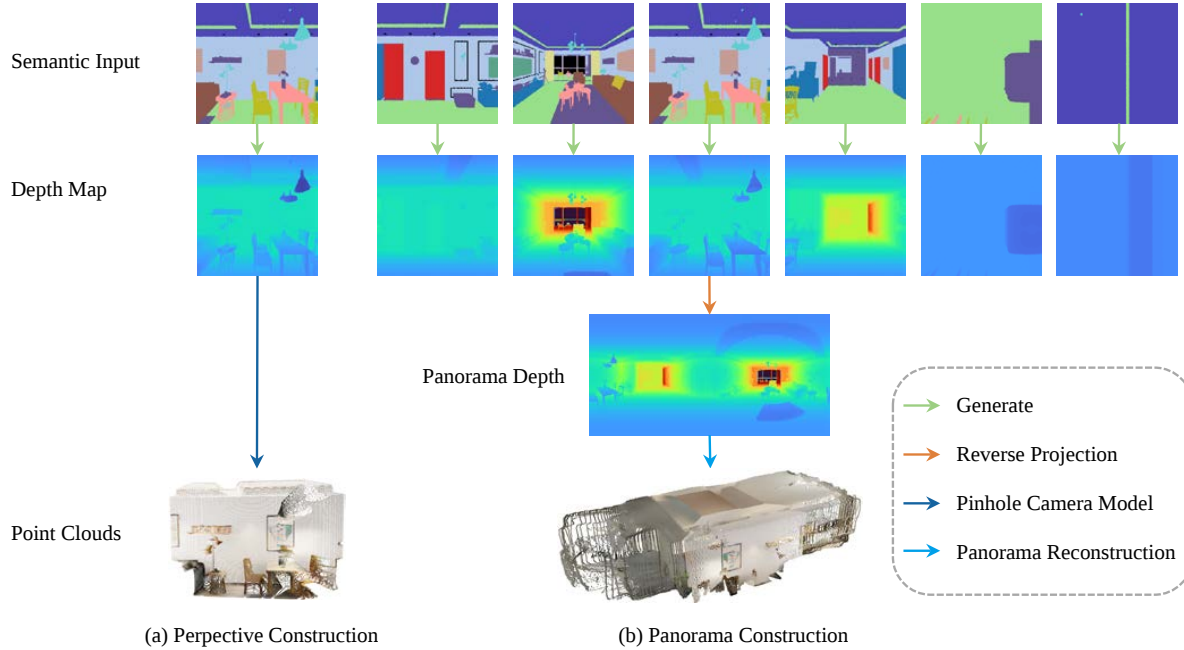


Figure 1. The construction process. (a) Perspective view construction through a *pinhole camera model*. (b) Panorama construction from the generated depth maps by the reverse geometric projection. The appearance of the point cloud is from the dataset. We can also generate appearance, which is described in the subsection of depth-appearance generation.