AtlantaNet: Supplementary Material

Giovanni Pintore[1 [0000−0001−8944−1045]], Marco Agus[2,1 [0000−0003−2752−3525]], and Enrico Gobbetti[1 [0000−0003−0831−2458]]

[1] Visual Computing, CRS4, Italy
giovanni.pintore@crs4.it enrico.gobbetti@crs4.it
[2] College of Science and Engineering, HBKU, Qatar
magus@hbku.edu.qa

1 Introduction

In this supplementary material, we include additional visual results integrating those presented in the submitted paper.

In the submitted version of the manuscript, we presented an extensive quantitative evaluation of our approach and a comparison with the most recent state-of-the-art methods [7,1,5,6]. Due to page limits, we showed visual results only for a selection of complex scenes, such as rooms bounded by vertical walls that do not form right angles or are curved – i.e., *Atlanta World* models, which are the main goal of our work.

Here, we provide more visual results from the Stanford 2D-3D-S [4] and MatterportLayout [8] datasets. The quantitative evaluation for these datasets is provided in the submitted paper (see Sec. 5.2).

Furthermore, as further support for the ablation study presented in Sec.5.3 of our submission, we present here additional *Atlanta World* cases using different setups for our method.

2 Results

MatterportLayout and Stanford. Fig. 1 and Fig. 2 show, respectively, visual results from the MatterportLayout [8] and Stanford 2D-3D-S [4] testing set, for which quantitative performances are presented in Tab. 2 of the main manuscript. The results demonstrate the validity of our approach, despite the fact that we do not perform any specific post-processing and cuboid regularization.

AtlantaLayout with ResNet101. In Fig. 3, we present some example from the AtlantaLayout testing set, which includes images from the original *Matterport3D* [3] and *Structured3D* [2] datasets annotated by us, adopting a *ResNet101* encoder (see Tab. 3 on the main manuscript). Using a deeper encoder expands the ability to capture more detailed structural features (see Sec. 5.3 on the main manuscript).

Ablation study. In Fig. 4 we present a visual comparison on the same Atlanta scene (Fig. 4(a)). We train the network (e.g., *ResNet50* backbone), respectively, without gradient loss and without fine tuning (Fig. 4(b)), with gradient loss without fine tuning (Fig. 4(c)) and with gradient loss and fine tuning (Fig. 4(d)). Visual results support proposed design choices.

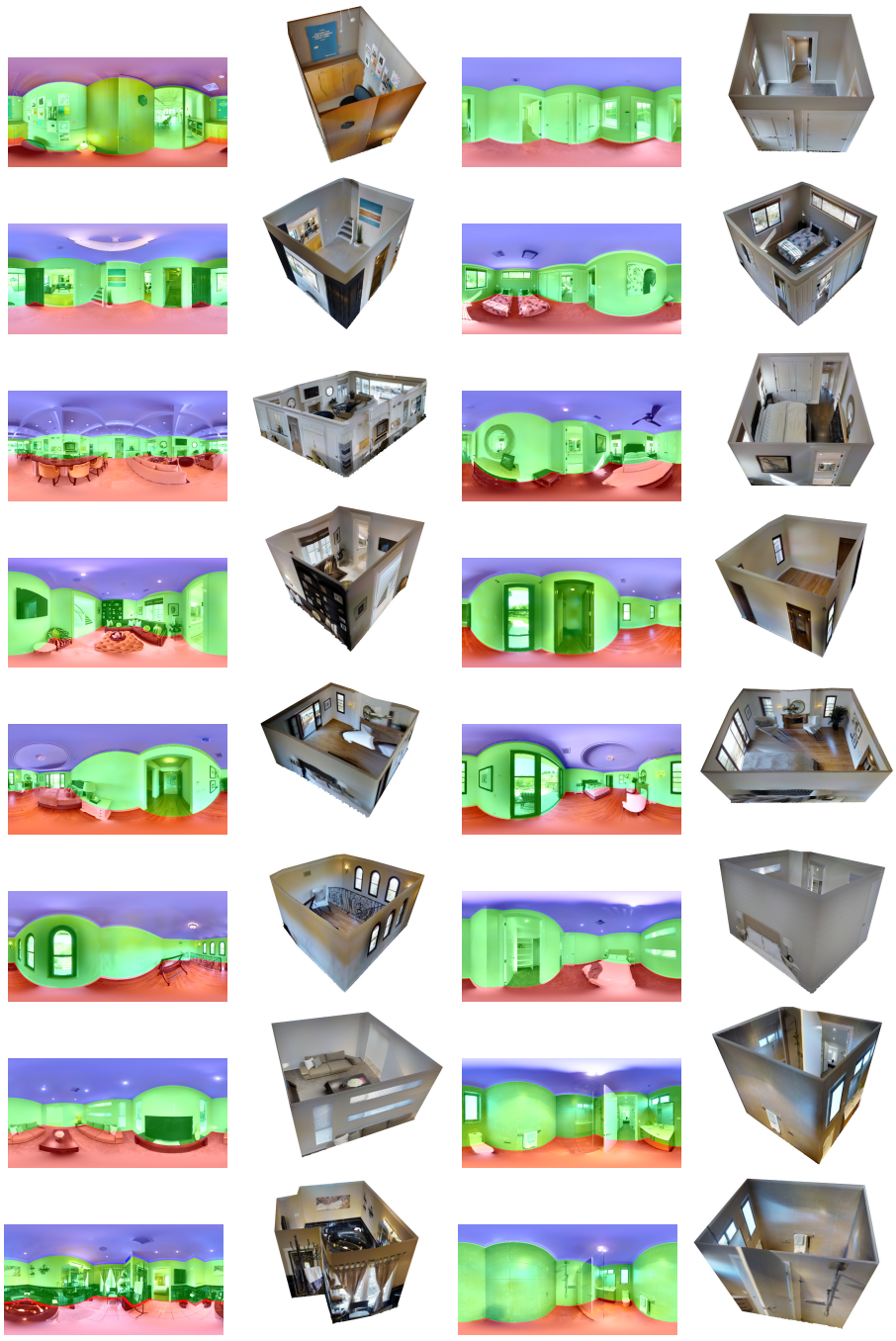


Fig. 1. Qualitative results on the MatterportLayout [8] testing set. We show the original panoramic image annotated with our reconstruction followed by the recovered 3D layout.

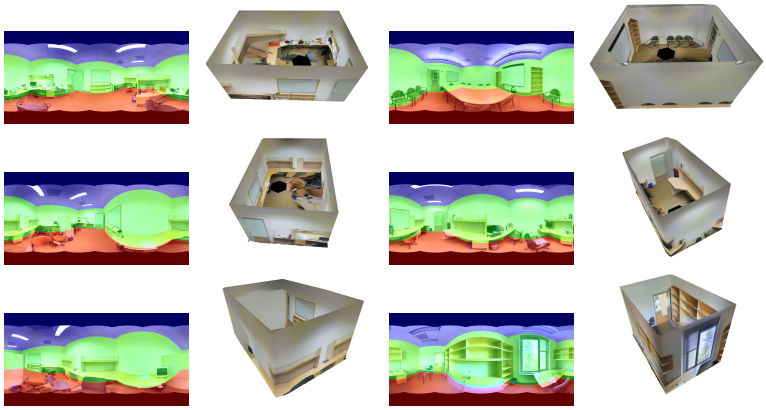


Fig. 2. Qualitative results from the Stanford [4] testing set. We show the original panoramic image labeled with our reconstruction and the textured 3D layout.

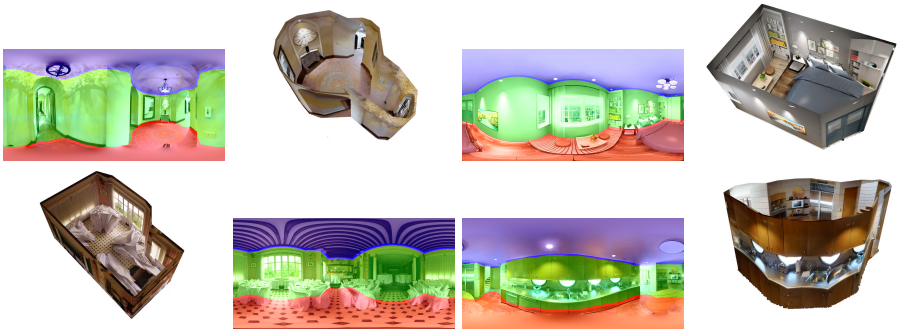
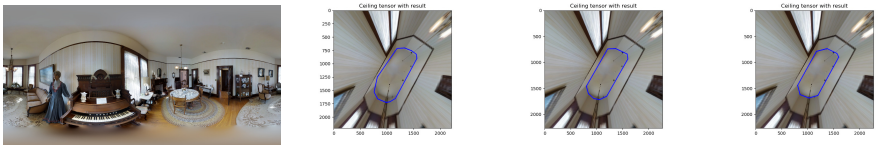


Fig. 3. Qualitative results from the AtlantaLayout testing set. We show the original panoramic image labeled with our reconstruction and the textured 3D layout. Results are obtained using the *ResNet101* encoder.



(a) (b) (c) (d)

Fig. 4. Qualitative comparison using gradient loss and fine tuning. We show visual comparison on the ceiling transform adopting different design choices (inferred shape in blue).

References

1. Fernandez-Labrador, C., Fácil, J.M., Perez-Yus, A., Demonceaux, C., Civera, J., Guerrero, J.J.: Corners for layout: End-to-end layout recovery from 360 images. arXiv:1903.08094 (2019)
2. Kujiale.com: Structured3D Data. `https://structured3d-dataset.org/` (2019), [Accessed: 2019-09-25]
3. Matterport: Matterport3D. `https://github.com/niessner/Matterport` (2017), [Accessed: 2019-09-25]
4. Stanford University: BuildingParser Dataset. `http://buildingparser.stanford.edu/dataset.html` (2017), [Accessed: 2019-09-25]
5. Sun, C., Hsiao, C.W., Sun, M., Chen, H.T.: HorizonNet: Learning room layout with 1D representation and pano stretch data augmentation. In: Proc. CVPR (June 2019)
6. Yang, S.T., Wang, F.E., Peng, C.H., Wonka, P., Sun, M., Chu, H.K.: DuLa-Net: A dual-projection network for estimating room layouts from a single RGB panorama. In: Proc. CVPR (2019)
7. Zou, C., Colburn, A., Shan, Q., Hoiem, D.: LayoutNet: Reconstructing the 3D room layout from a single RGB image. In: Proc. CVPR. pp. 2051–2059 (2018)
8. Zou, C., Su, J.W., Peng, C.H., Colburn, A., Shan, Q., Wonka, P., Chu, H.K., Hoiem, D.: 3d manhattan room layout reconstruction from a single 360 image (2019)