Learning to Reconstruct 3D Manhattan Wireframes from A Single Image

Yichao Zhou* Haozhi Qi* Yuexiang Zhai* Qi Sun‡ Zhili Chen ‡† Li-Yi Wei‡ Yi Ma*

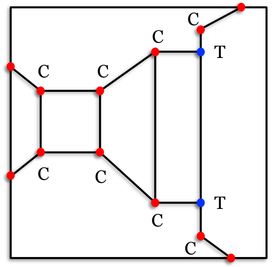
*University of California, Berkeley ‡Adobe Research †Bytedance Research

ICCV 2019 Seoul, Korea

Wireframe Representation

Our representation of wireframes is based on the notation from graph theory:

- Let $W = (V, E)$ be a wireframe;
- V is the set of junction indices;
- $E \subseteq V \times V$ is the set of lines;
- For each $\forall i \in V$:
  - $\boldsymbol{p}_i$ represents its coordinate;
  - $\boldsymbol{z}_i$ represents its depth;
  - $\boldsymbol{t}_i \in \{C, T\}$ represents its type.

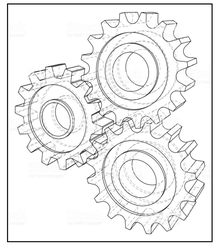


Why Wireframe?

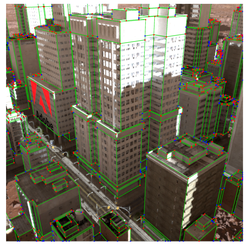
- Editable CAD representation;
- Clean geometry from priors of man-made environments;
- Compact, easy for content sharing and transmission;
- High-level structuralized features, as opposed to local corner features such as SIFT or line segments;
- Now possible with recent advances in deep learning.

Applications

(a) Augmented Reality

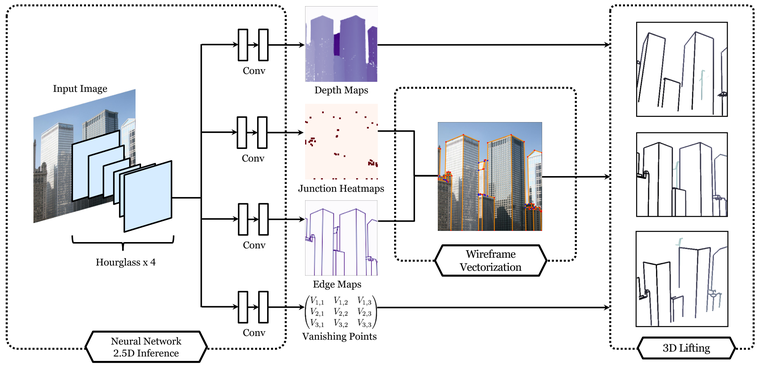


(b) CAD Reconstruction



(c) 3D Editing

Methods



Datasets

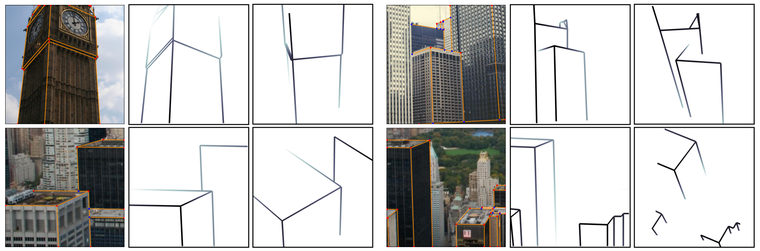
Synthetic SceneCity Urban 3D (SU3) Dataset

23,000 images rendered with procedural generated cities.

Real-world 3D Wireframe Dataset

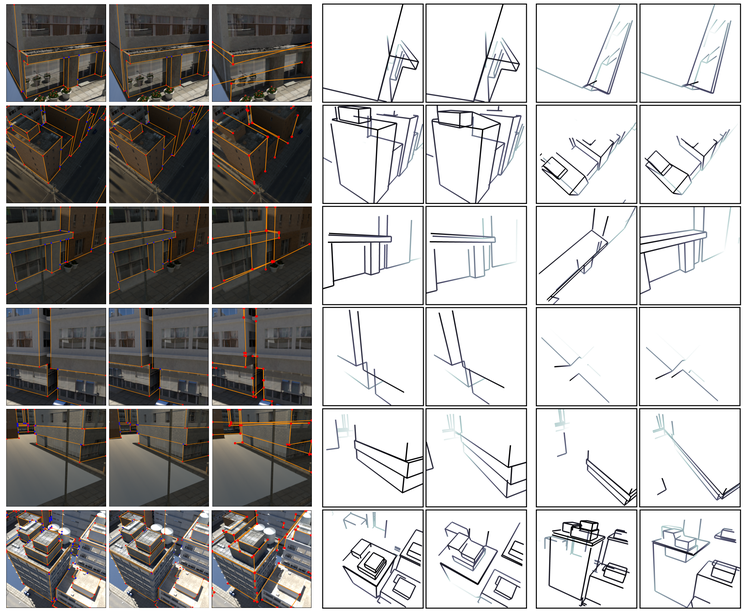
200 images from MegaDepth [1] with human-labelled wireframes.

Results on the Real-world Wireframe Dataset



Ground Truth Inferred Novel views Ground Truth Inferred Novel views

Results on Our Synthetic SceneCity Urban 3D (SU3) Dataset



GT 2D Our 2D 2D [2] GT 3D Our 3D GT 3D Our 3D

Acknowledgement

This work is partially supported by Sony US Research Center, Adobe Research, FHL Vive Center for Enhanced Reality, Berkeley BAIR, and Bytedance Research Lab. We also thank Kenji Tashiro of Sony for helpful discussions.

Reference

1. Zhengqi Li and Noah Snavely. MegaDepth: Learning single-view depth prediction from internet photos. In CVPR, 2018.
2. Kun Huang, Yifan Wang, Zihan Zhou, Tianjiao Ding, Shenghua Gao, and Yi Ma. Learning to parse wireframes in images of man-made environments. In CVPR, 2018.

Berkeley UNIVERSITY OF CALIFORNIA