PlaneNet [3]

Piece-wise Planar Reconstruction from a Single RGB Image

Jia Zheng

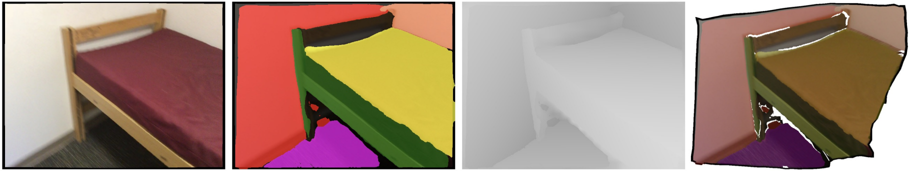
August 19, 2018

Outline

1. Introduction
2. Methods
3. Experiments

Task

Infer a set of plane parameters and corresponding plane segmentation masks from a single RGB image.



From left to right: an input image, a piece-wise planar segmentation, a reconstructed depthmap, and a texture-mapped 3D model.

Challenge

We do not know

1. the number of planes to be inferred,
2. the order of planes to be regressed in the output feature vector.

Pipeline

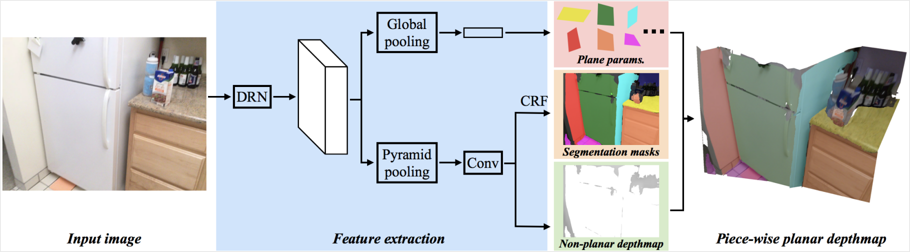


Figure: PlaneNet predicts plane parameters, their probabilistic segmentation masks, and a non-planar depthmap from a single RGB image.

Plane parameter branch

Predict a fixed number (K) of planar surfaces $\mathcal{S} = \{S_1, \cdots S_K\}$. Each surface S_i is specified by the three plane parameters P_i (i.e., encoding a normal and an offset).

An order-agnostic loss function based on the Chamfer distance metric for regressed plane parameters:

$$\mathcal{L}^P = \sum_{i=1}^{K^*} \min_{j \in [1,K]} \| P_i^* - P_j \|_2^2 .$$

Plane segmentation branch

Standard cross entropy loss to supervise the segmentation

$$\mathcal{L}^M = \sum_{i=1}^{K+1} \sum_{p \in I} (\mathbf{1}(M^{*(p)} = i) \log(1 - M_i^{(p)}))$$

Non-plane depth branch

Define the loss as sum of squared depth differences weighted by probabilities

$$\mathcal{L}^D = \sum_{i=1}^{K+1} \sum_{p \in I} (M_i^{(p)} (D_i^{(p)} - D^{*(p)})^2)$$

Data

Generate 51,000 ground truth piece-wise planar depthmaps from ScanNet [1]. We fit planes to a consolidated mesh and project them back to individual frames.

Reconstruction results

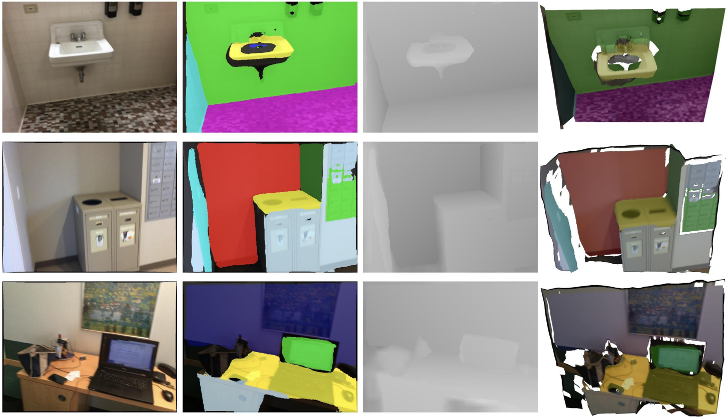


Figure: Piece-wise planar depthmap reconstruction results. From left to right: input image, planar segmentation, depthmap reconstruction, and 3D rendering of our depthmap.

Plane segmentation accuracy

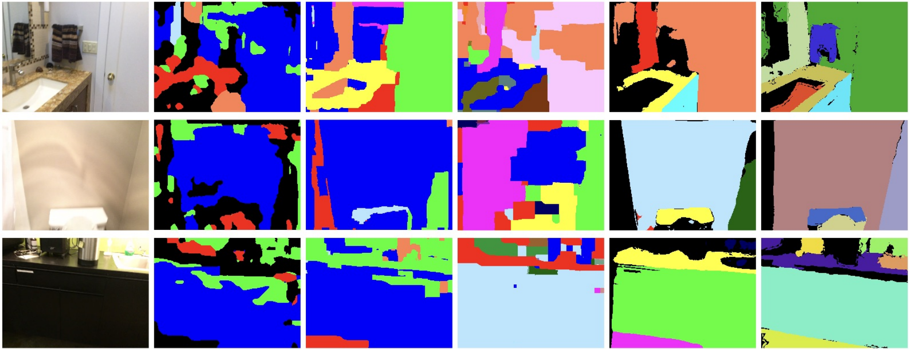


Figure: Qualitative comparisons. From left to right: input image, [4], [2], [5], [3], the ground-truth.

Plane ordering consistency

Figure: Room layout estimation. We manually select the entries of planes that correspond to the ceiling, the floor, and the left/middle/right walls.

Failure modes

1. Generate two nearly co-planar vertical surfaces in the low-light region.
2. Can not distinguish a white object on the floor from a white wall.
3. Miss a column structure on a wall due to the presence of object clutter.

Texture editing applications

From top to bottom, an input image, a plane segmentation result, and an edited image.

Summary

Table: Comparisons with the state-of-the-art planar reconstruction algorithms.

Method	Liu et al. [3]	Yang and Zhou [6]
	Top-down approach Fixed number of planes	
Data	Ground-truth 3D planes	RGB-D
Semantic Info.	No	Yes

Reference

Angela Dai et al. “ScanNet: Richly-annotated 3D Reconstructions of Indoor Scenes”. In: *CVPR*. 2017.

Yasutaka Furukawa et al. “Manhattan-world Stereo”. In: *CVPR*. 2009.

Chen Liu et al. “PlaneNet: Piece-wise Planar Reconstruction from a Single RGB Image”. In: *CVPR*. 2018.

Nathan Silberman et al. “Indoor Segmentation and Support Inference from RGBD Images”. In: *ECCV*. 2012.

Sudipta N Sinha, Drew Steedly Microsoft, and Richard Szeliski. “Piecewise Planar Stereo for Image-based Rendering”. In: *ICCV*. 2009.

Fengting Yang and Zihan Zhou. “Recovering 3D Planes from a Single Image via Convolutional Neural Networks”. In: *ECCV*. 2018.

Thanks

Thanks for Attention!