Large Scale Augmented and Virtual Reality in Structured Environments

Joseph Menke, Oladapo Afolabi, Bill Zhou, Alex Yu, Allen Yang

OpenARK

Features:

- 3D Hand Tracking

- Building Scale SLAM

- Building Interior 3D Reconstruction

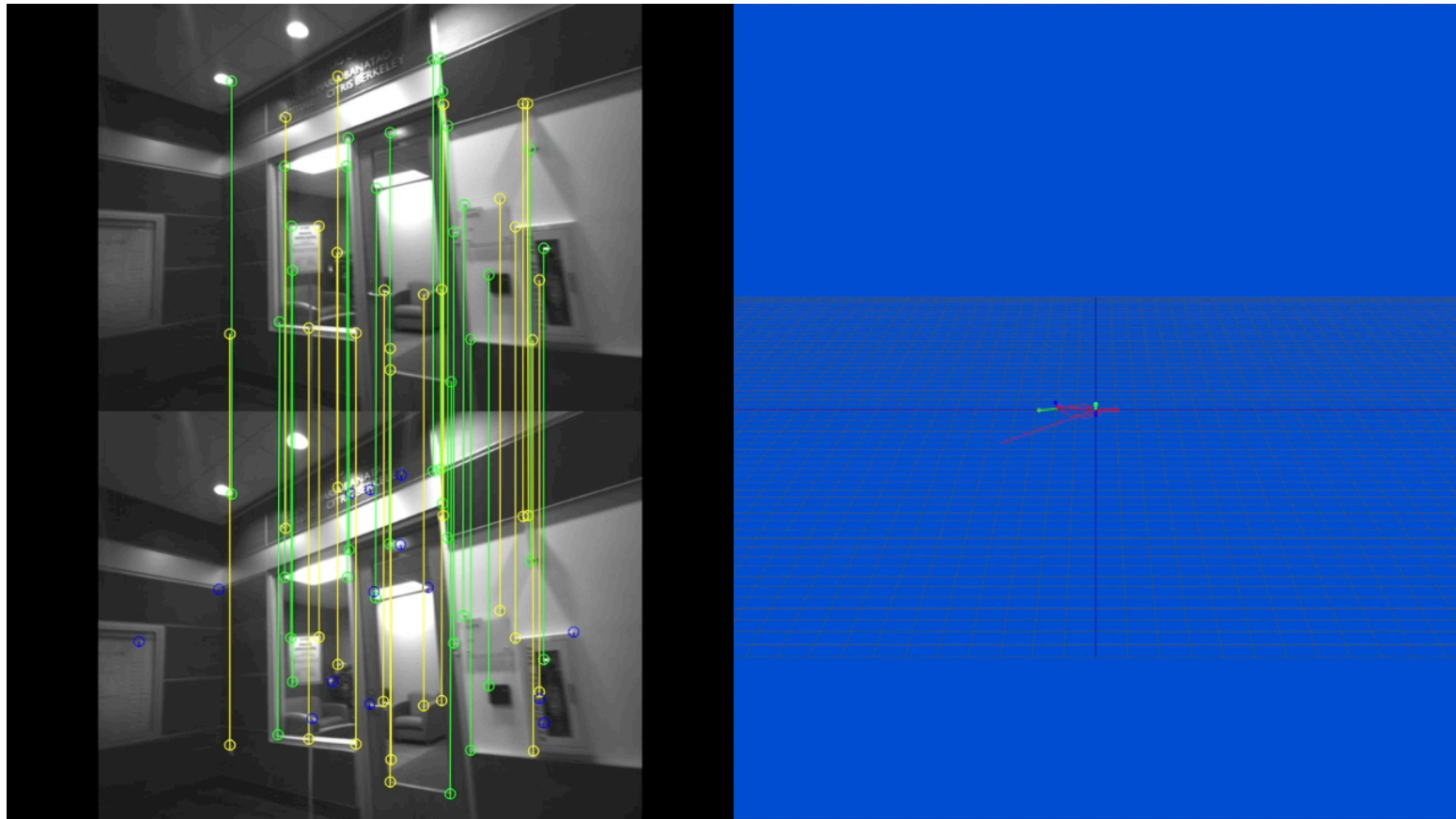- Articulated Avatar Tracking

OpenARK is an open-source platform includes fundamental tools such as AR-based camera calibration and SLAM, and it also includes higher-level functions to aid human-computer interaction, such as 3D gesture recognition and multi-user collaboration.
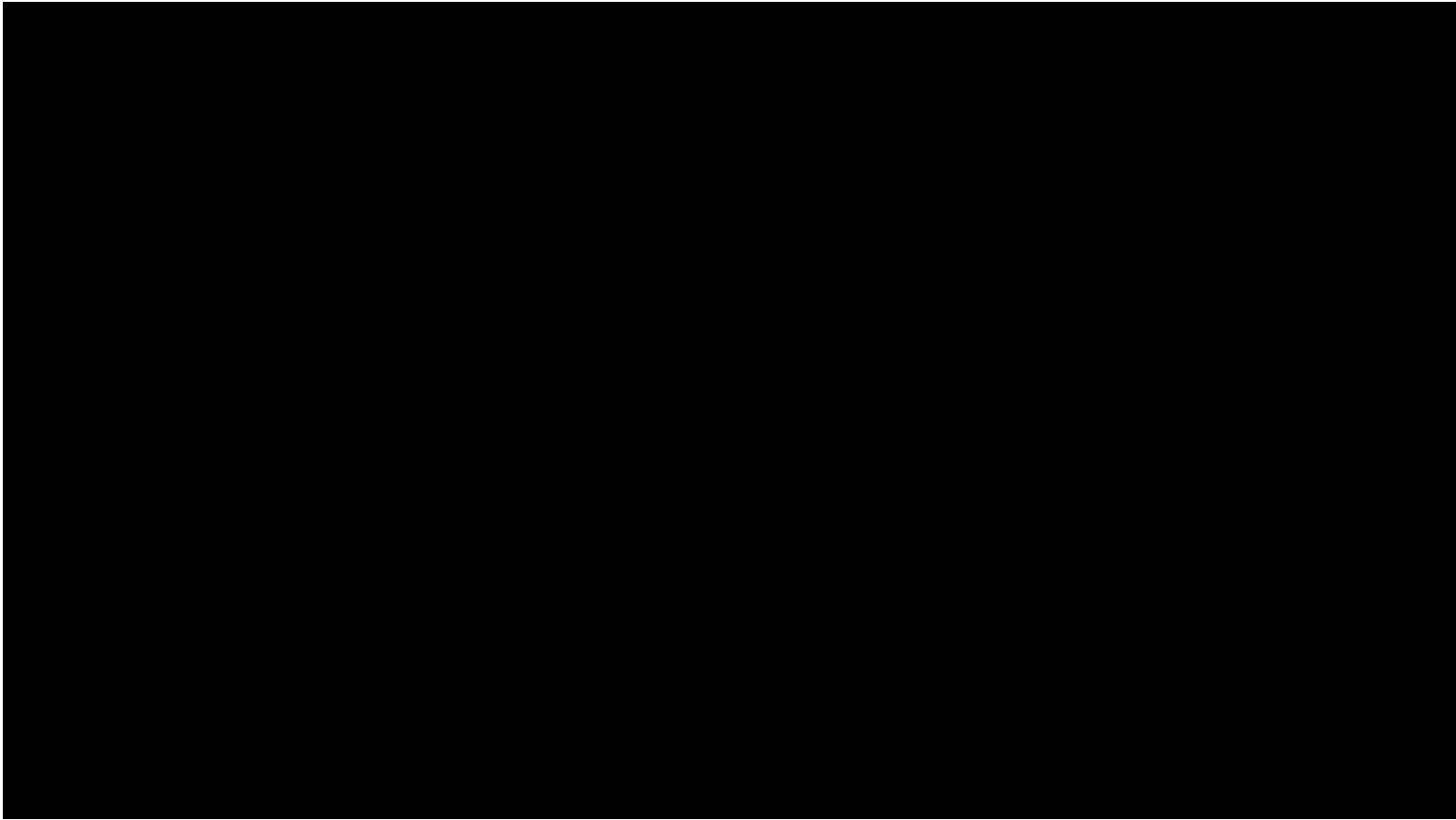
+ Hand Tracking

Simultaneous Localization and Mapping

S. Leutenegger et al "Keyframe-based visual–inertial odometry using
nonlinear optimization" 2015

+ 3D Reconstruction
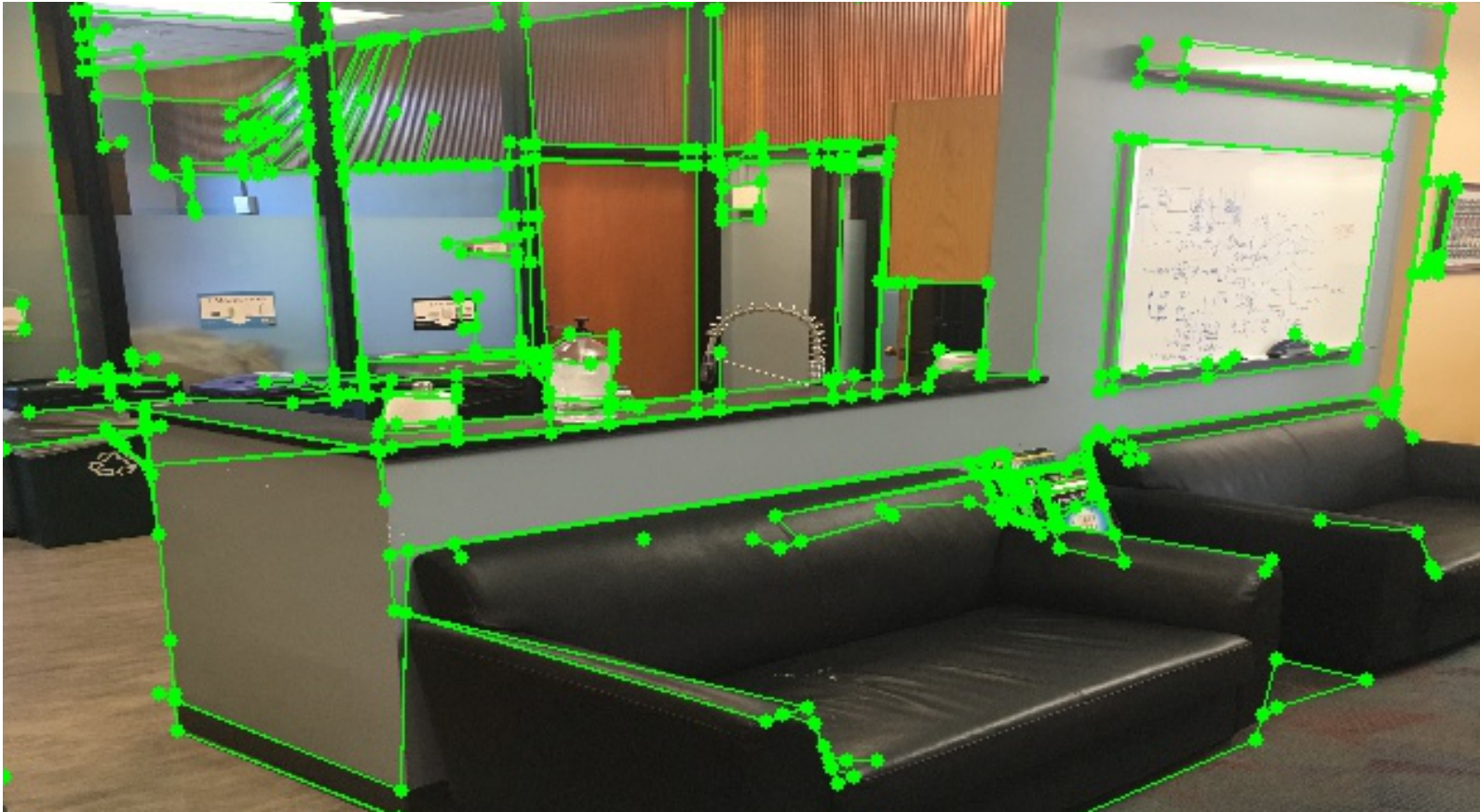
+ Avatar Tracking
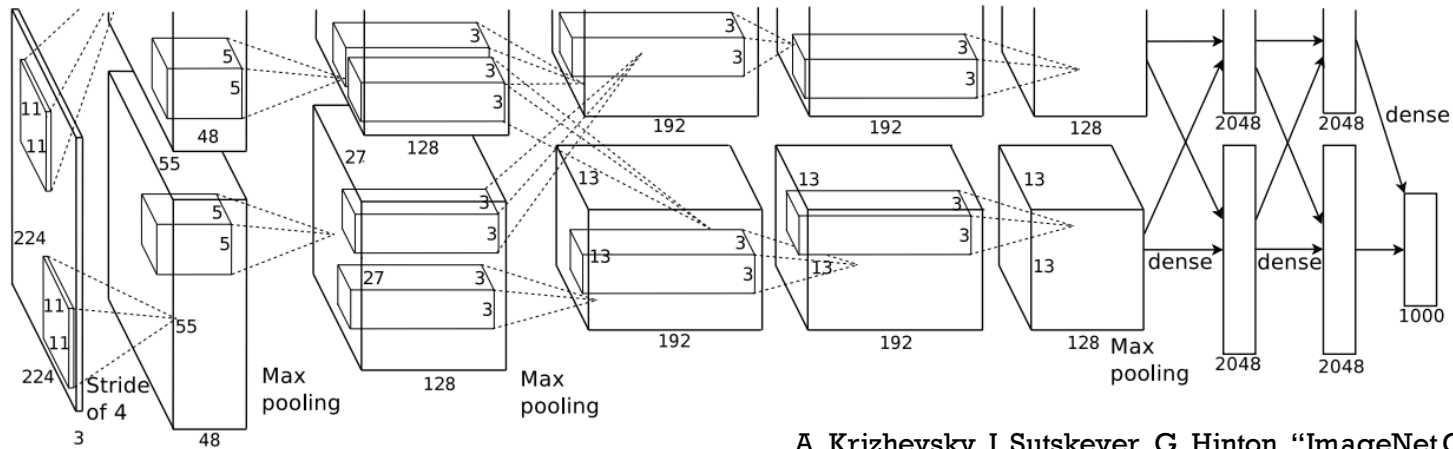
+ OpenARK

+

Where do we go from here?

One Type Of Structure

+ Detecting Structure

■ **Big Data:**

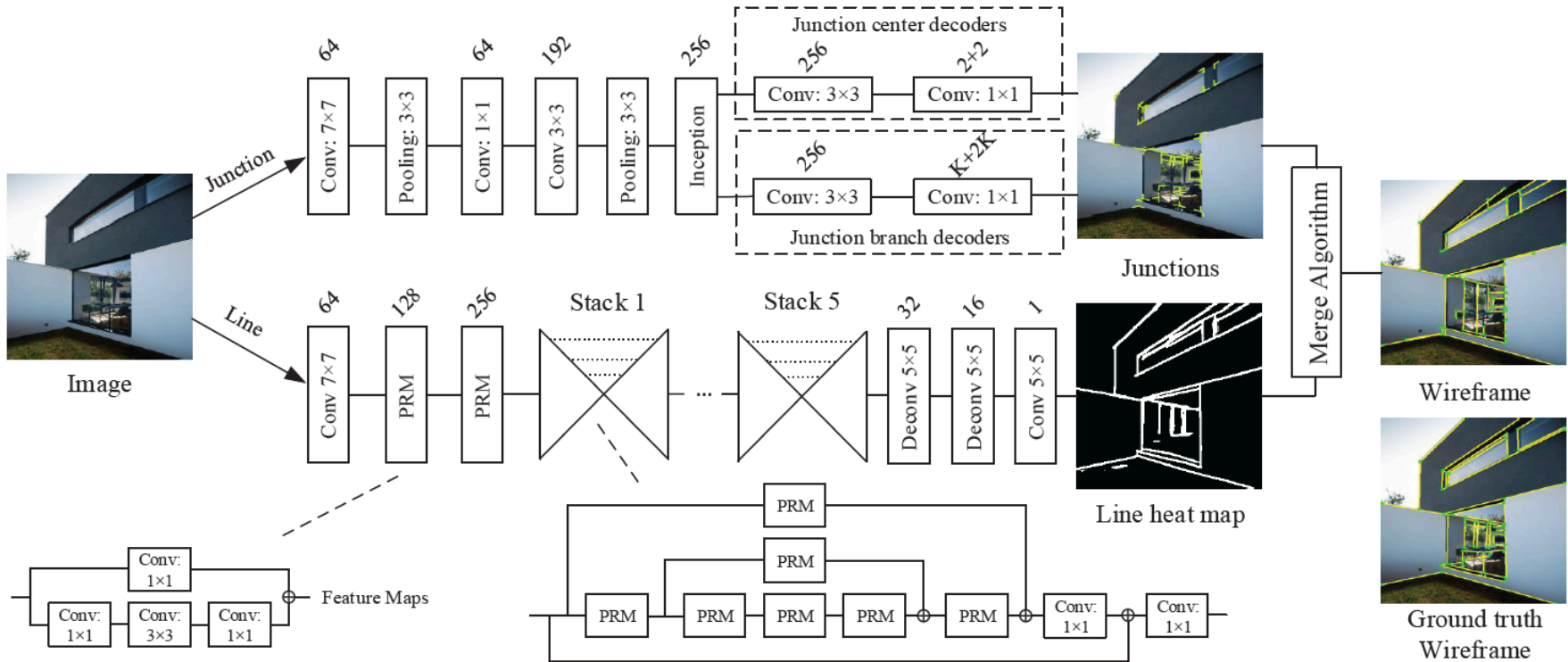

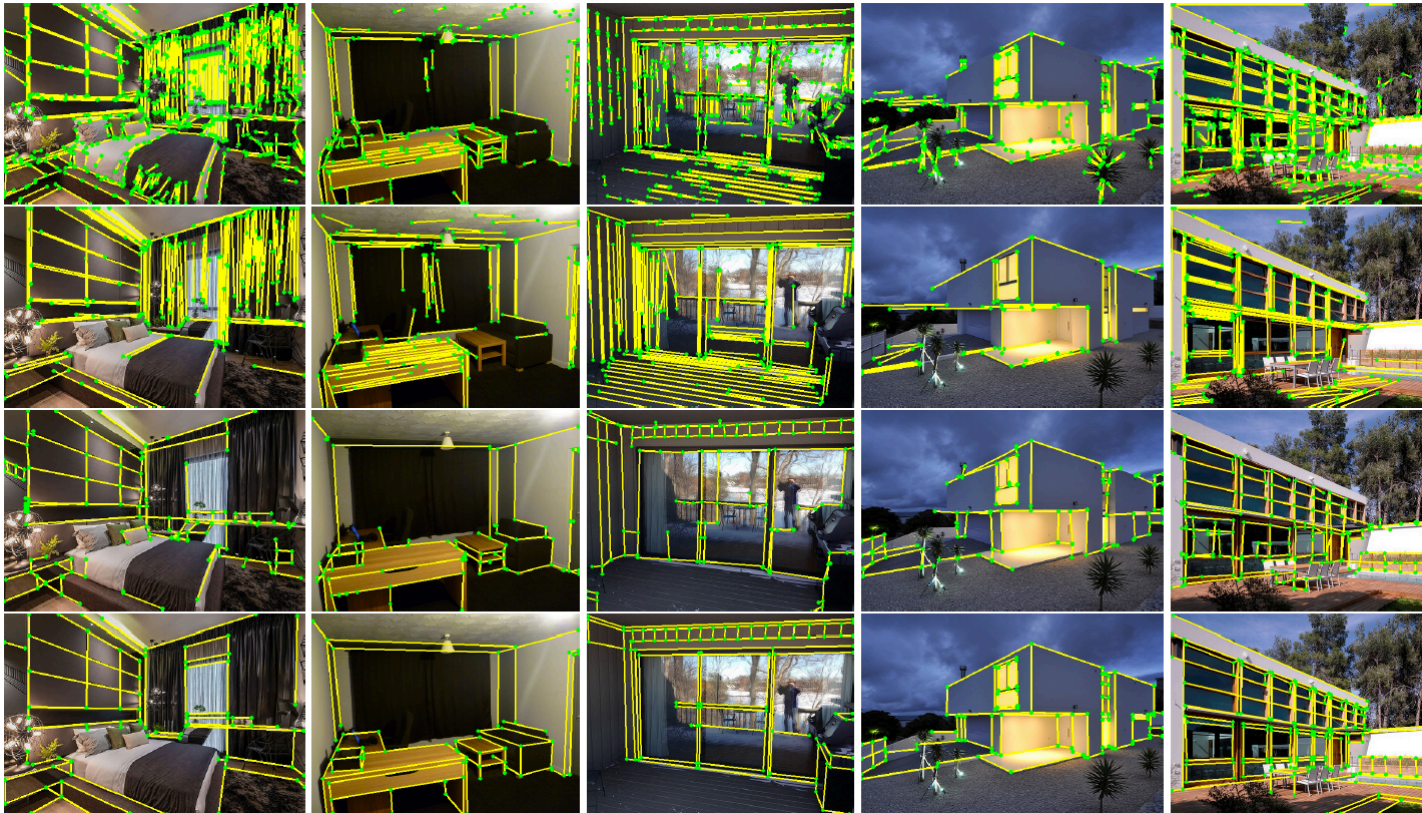A. Krizhevsky, I. Sutskever, G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks", NIPS 2012

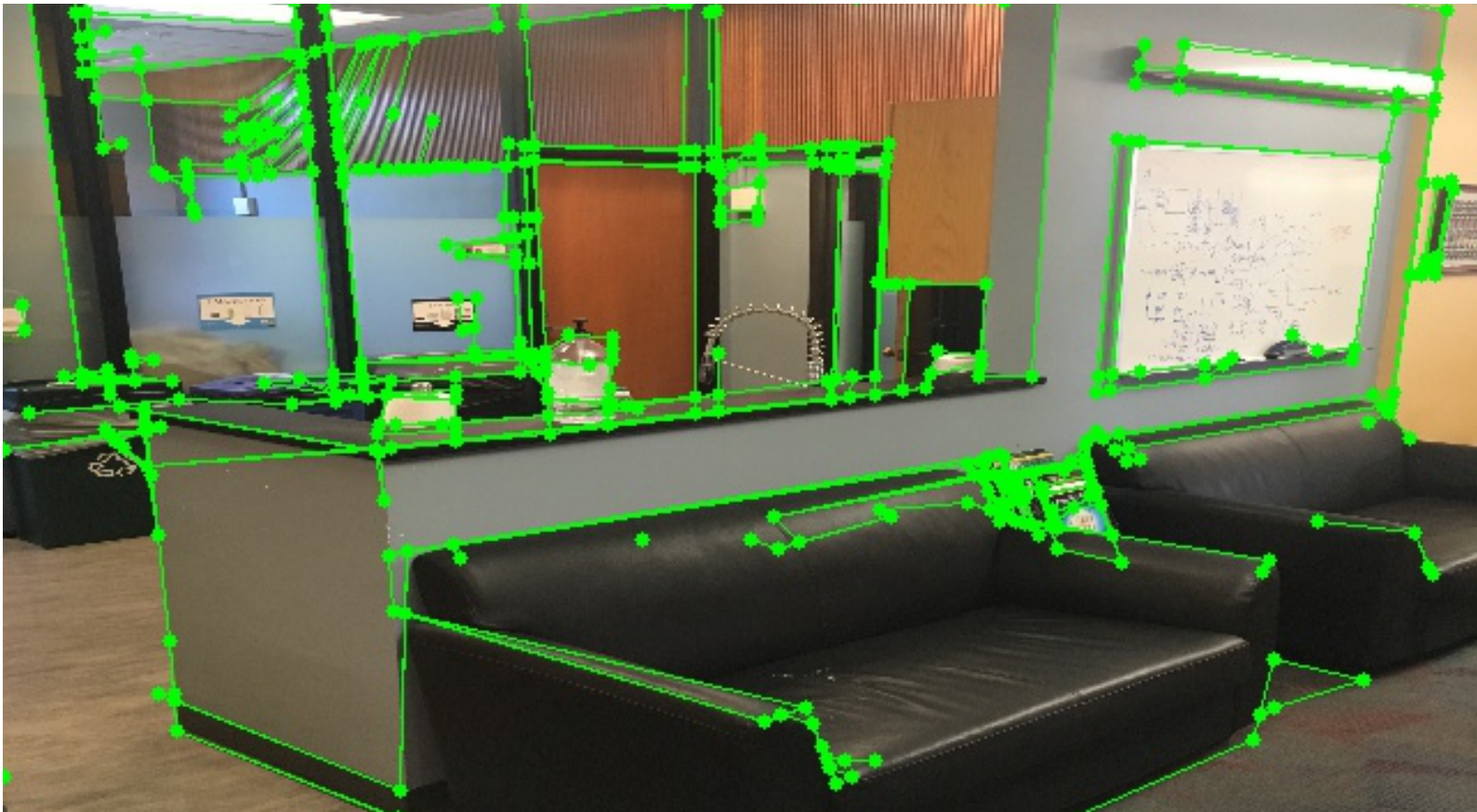■ **Cheap Stereo and Depth Cameras:**

Learning to Detect Wireframes

Image

Junction

Line

Conv: 7×7	Pooling: 3×3	Conv: 1×1	Conv 3×3	Pooling: 3×3	Inception
64		64	192		256

Junction center decoders
Conv: 3×3 (256) — Conv: 1×1 (2+2)

Junction branch decoders
Conv: 3×3 (256) — Conv: 1×1 (K+2K)

Junctions

Conv 7×7 (64) — PRM (128) — PRM (256) — Stack 1 ... Stack 5 — Deconv 5×5 (32) — Deconv 5×5 (16) — Conv 5×5 (1)

Line heat map

Merge Algorithm

Wireframe

Ground truth Wireframe

Feature Maps

Conv: 1×1
Conv: 1×1 — Conv: 3×3 — Conv: 1×1

PRM
PRM
PRM — PRM — PRM — PRM — PRM — Conv: 1×1 — Conv: 1×1

K. Huang, Y. Wang, Z. Zhou, T. Ding, S. Gao, Y. Ma, "Learning to Parse Wireframes in Images of Man-Made Environments", CVPR 2018

Learning to Detect Wireframes

LSD

MCMLSD

Yi Ma

Ground Truth

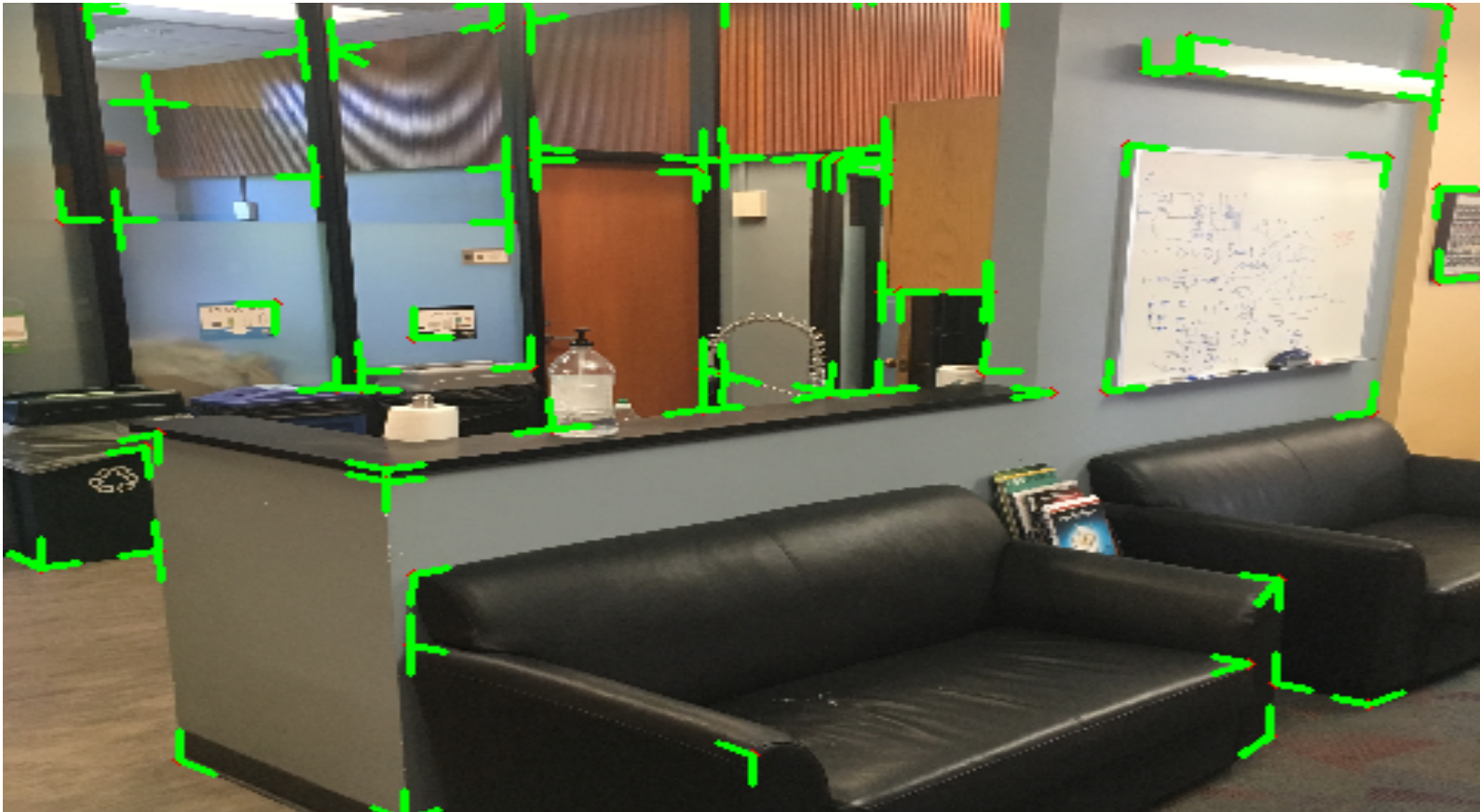K. Huang, Y. Wang, Z. Zhou, T. Ding, S. Gao, Y. Ma, "Learning to Parse Wireframes in Images of Man-Made Environments", CVPR 2018
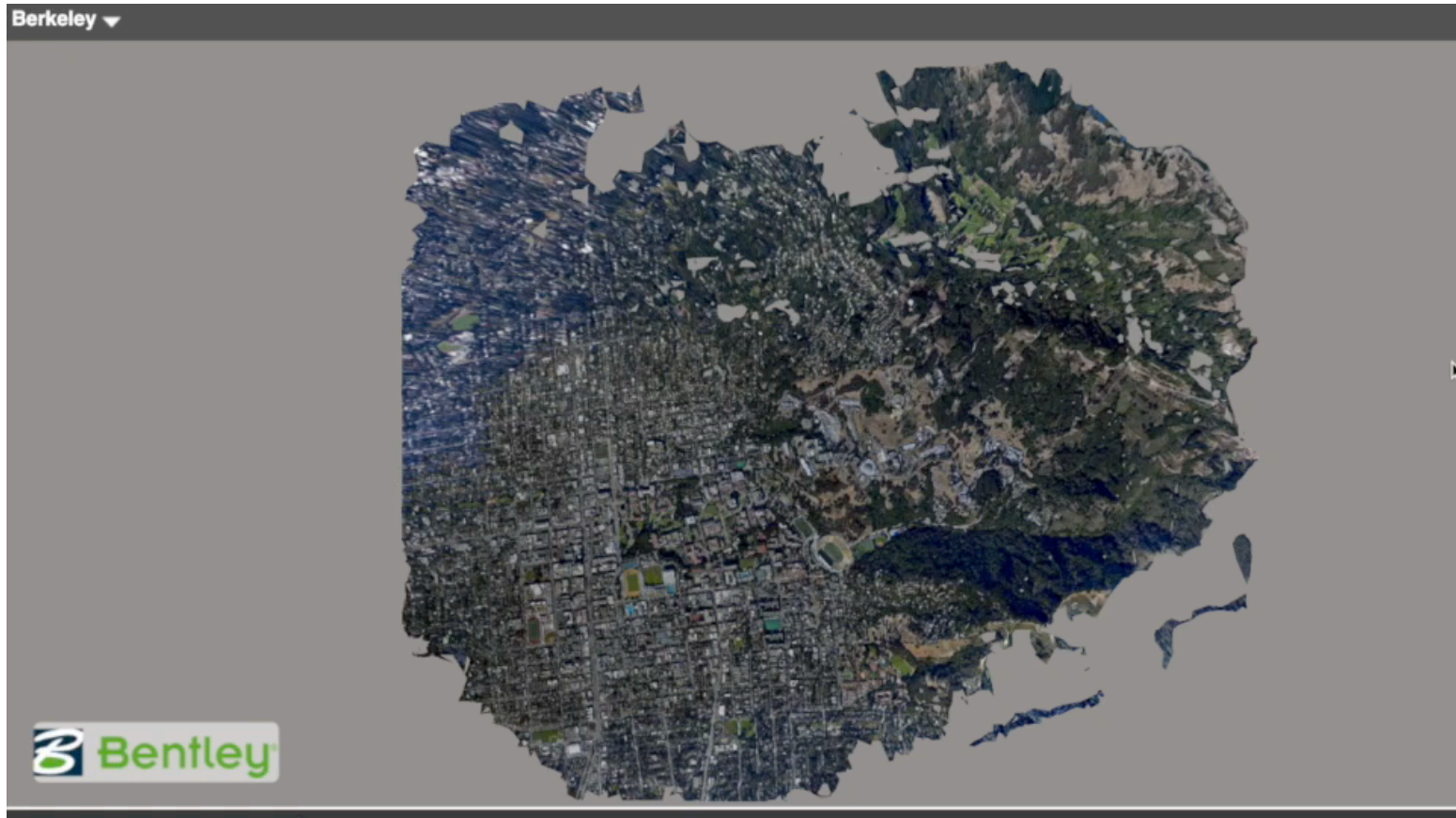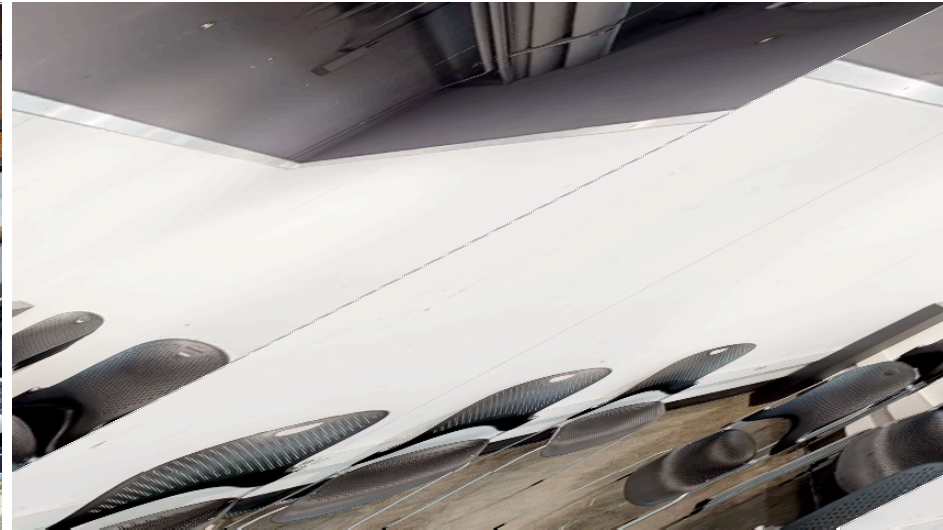
+ Heuristics

Learned Wireframes

+ Think Bigger

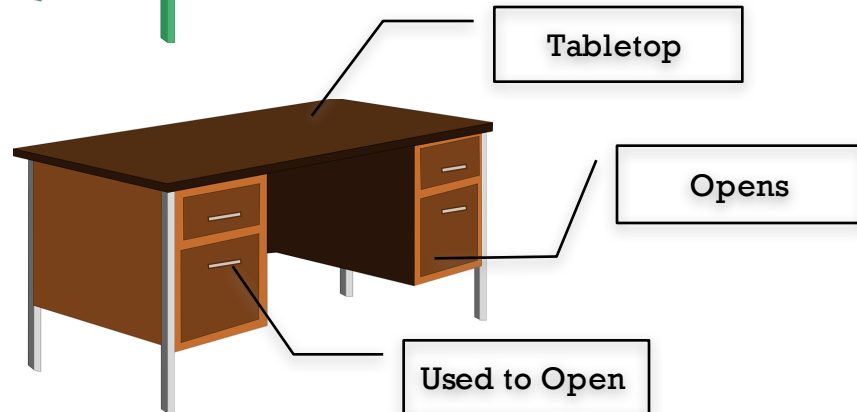# Berkeley Dataset Generation
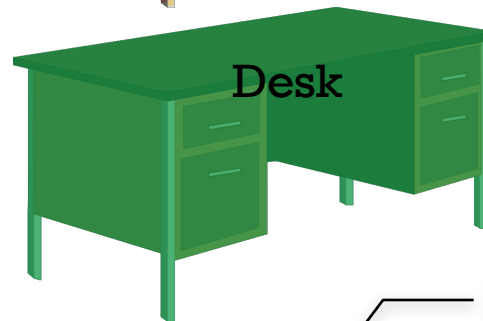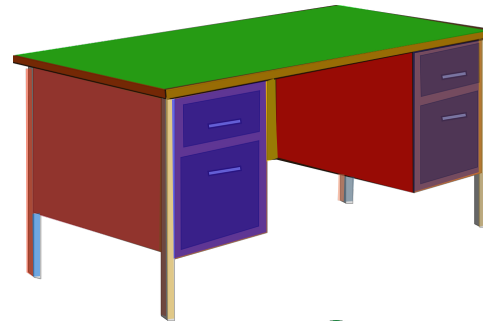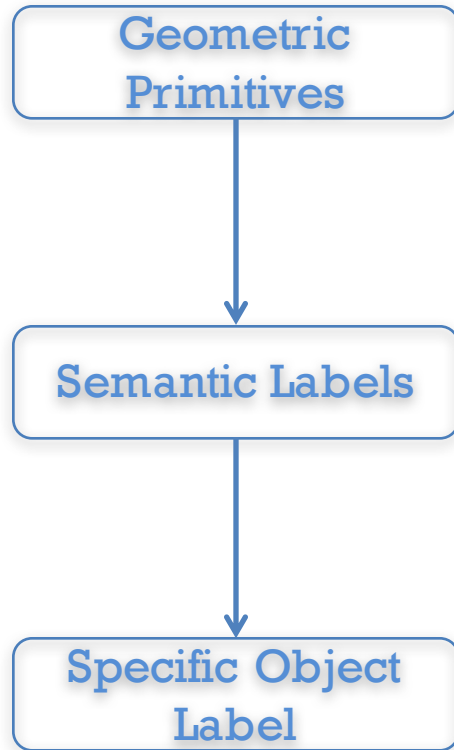
Mapping Interiors

- PX-80 Commercial 3D Reconstruction System provides ground truth information for the dataset.

- Intel Realsense d435i commercial depth camera allows for evaluation against a practical system

- PointGrey wide baseline stereo camera setup enables testing new deep learning algorithms with stereo constraints.

+ # How to Label?

+ Semiautomatic Labeling

Geometric Primitives

Semantic Labels

Specific Object Label

Desk

Tabletop

Opens

Used to Open

+ Summary

- Through Open-Source projects such as OpenARK we are working to advance the state of AR/VR by bringing more features available to researchers, hobbyists, and companies.

- We are working to tackle fundamental challenges in computer vision to allow for better and more useful 3D reconstruction.

- If you would like to collaborate contact:
  - Joseph Menke: joemenke@berkeley.edu
  - Allen Yang: yang@eecs.berkeley.edu

+

Thank You!