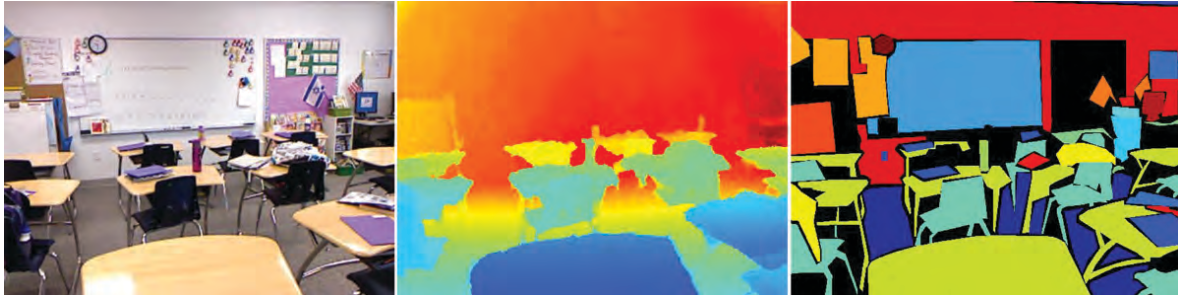# Datasets for A.I. Graphics

3D Scenes

NYU Depth Dataset V2 (2012) [Link]
1449 densely labeled pairs of aligned RGB and depth images from Kinect video sequences for a variety of indoor scenes.

SUNRGB-D 3D Object Detection Challenge [Link]
19 object categories for predicting a 3D bounding box in real world dimension
Training set: 10,355 RGB-D scene images, Testing set: 2860 RGB-D images

SceneNN (2016) [Link]
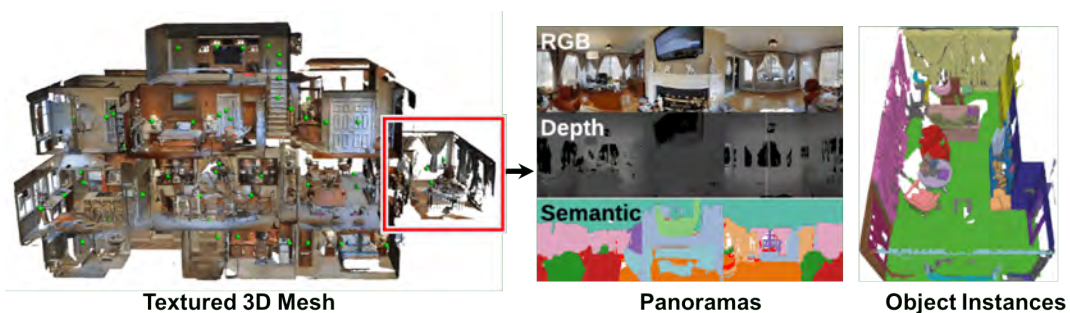100+ indoor scene meshes with per-vertex and per-pixel annotation.

ScanNet (2017) [Link]
An RGB-D video dataset containing 2.5 million views in more than 1500 scans, annotated with 3D camera poses, surface reconstructions, and instance-level semantic segmentations.

Matterport3D: Learning from RGB-D Data in Indoor Environments (2017) [Link]
10,800 panoramic views (in both RGB and depth) from 194,400 RGB-D images of 90 building-scale scenes of private rooms. Instance-level semantic segmentations are provided for region (living room, kitchen) and object (sofa, TV) categories.
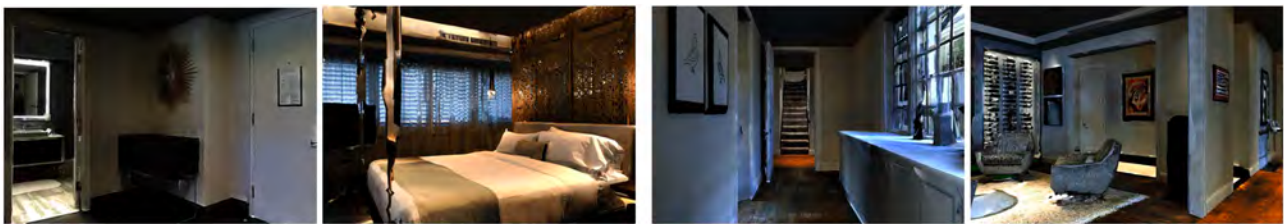

Textured 3D Mesh   Panoramas   Object Instances

SUNCG: A Large 3D Model Repository for Indoor Scenes (2017) [Link]
The dataset contains over 45K different scenes with manually created realistic room and furniture layouts. All of the scenes are semantically annotated at the object level.
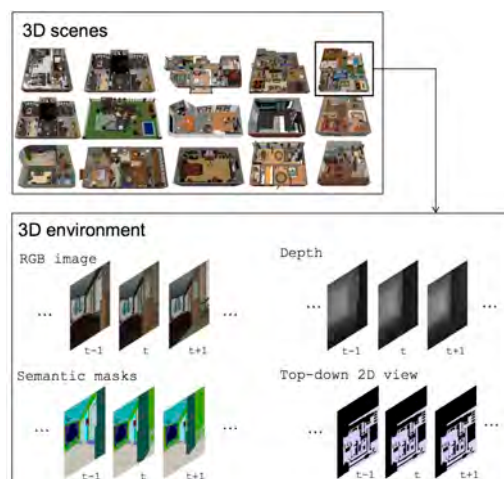

MINOS: Multimodal Indoor Simulator (2017) [Link]
MINOS is a simulator designed to support the development of multisensory models for goal-directed navigation in complex indoor environments. MINOS leverages large datasets of complex 3D environments and supports flexible configuration of multimodal sensor suites. MINOS supports SUNCG and Matterport3D scenes.
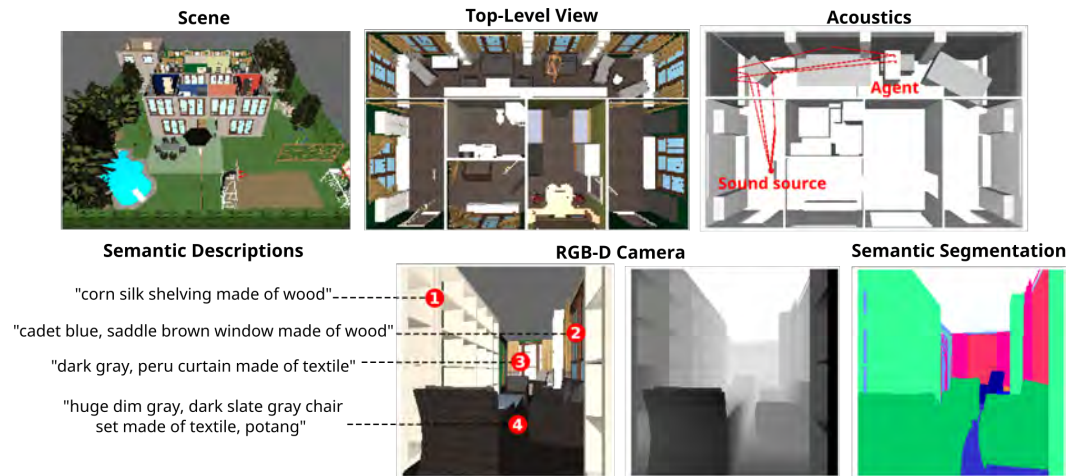

Facebook House3D: A Rich and Realistic 3D Environment (2017) [Link]
House3D is a virtual 3D environment which consists of 45K indoor scenes equipped with a diverse set of scene types, layouts and objects sourced from the SUNCG dataset. All 3D objects are fully annotated with category labels. Agents in the environment have access to observations of multiple modalities, including RGB images, depth, segmentation masks and top-down 2D map views.

HoME: a Household Multimodal Environment (2017) [Link]

HoME integrates over 45,000 diverse 3D house layouts based on the SUNCG dataset, a scale which may facilitate learning, generalization, and transfer. HoME is an open-source, OpenAI Gym-compatible platform extensible to tasks in reinforcement learning, language grounding, sound-based navigation, robotics, multi-agent learning.
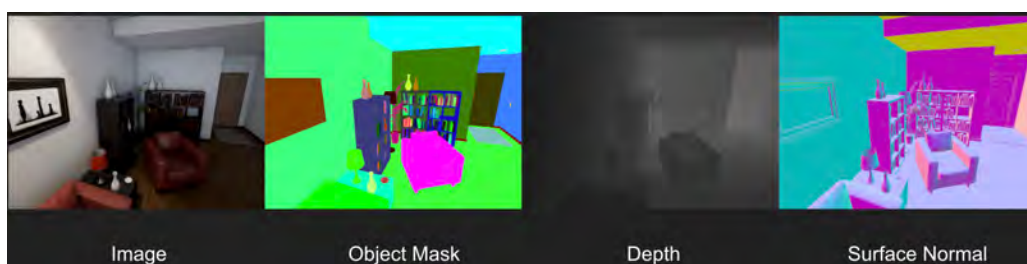

AI2-THOR: Photorealistic Interactive Environments for AI Agents [Link]

AI2-THOR is a photo-realistic interactable framework for AI agents. There are a total 120 scenes in version 1.0 of the THOR environment covering four different room categories: kitchens, living rooms, bedrooms, and bathrooms. Each room has a number of actionable objects.
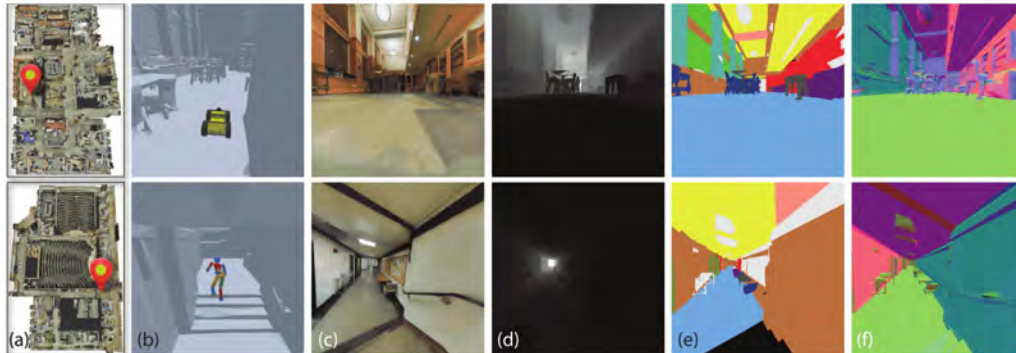

UnrealCV: Virtual Worlds for Computer Vision (2017) [Link][Paper]

An open source project to help computer vision researchers build virtual worlds using Unreal Engine 4.

Gibson Environment: Real-World Perception for Embodied Agents (2018 CVPR) [Link]
This platform provides RGB from 1000 point clouds, as well as multimodal sensor data: surface normal, depth, and for a fraction of the spaces, semantics object annotations. The environment is also RL ready with physics integrated. Using such datasets can further narrow down the discrepency between virtual environment and real world.
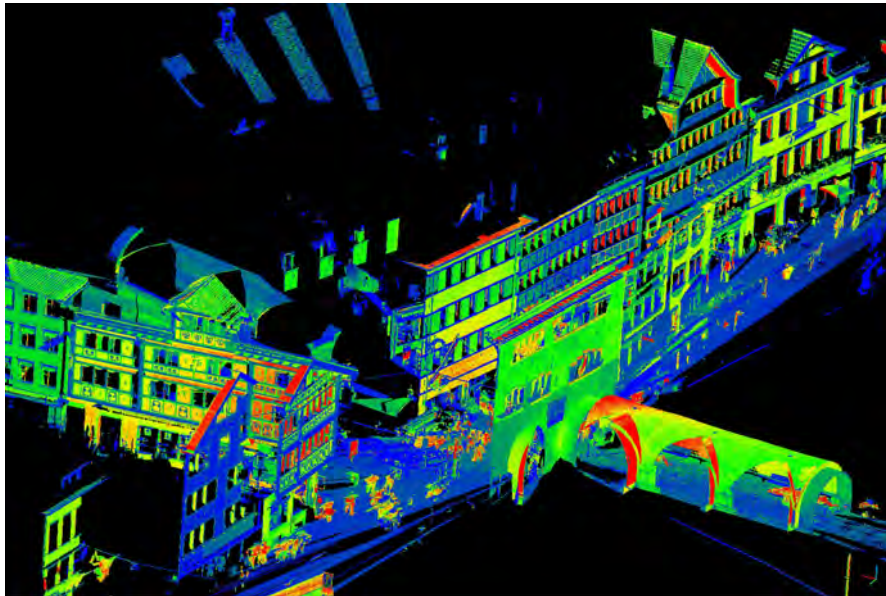

InteriorNet: Mega-scale Multi-sensor Photo-realistic Indoor Scenes Dataset [Link]
System Overview: an end-to-end pipeline to render an RGB-D-inertial benchmark for large scale interior scene understanding and mapping. Our dataset contains 20M images created by pipeline: (A) We collect around 1 million CAD models provided by world-leading furniture manufacturers. These models have been used in the real-world production. (B) Based on those models, around 1,100 professional designers create around 22 million interior layouts. Most of such layouts have been used in real-world decorations. (C) For each layout, we generate a number of configurations to represent different random lightings and simulation of scene change over time in daily life. (D) We provide an interactive simulator (ViSim) to help for creating ground truth IMU, events, as well as monocular or stereo camera trajectories including hand-drawn, random walking and neural network based realistic trajectory. (E) All supported image sequences and ground truth.

Semantic3D[Link]

Large-Scale Point Cloud Classification Benchmark, which provides a large labelled 3D point cloud data set of natural scenes with over 4 billion points in total, and also covers a range of diverse urban scenes.

Structured3D: A Large Photo-realistic Dataset for Structured 3D Modeling [Link]