# SUN3D

# A large-scale RGB-D video database with

### Camera pose

### Object labels

### Capturing the full 3D extent of many places

### place-centric rather than view-based, containing full 3D models of spaces

# Two task

### Hand-labeling videos

### Structure from motion

# 3D Obejct and Scene Dataset

### RGB-D object dataset

###### 11-ICRA-A large-scale hierarchical Multiview RGB-D object dataset

### The NYU Depth dataset

###### 12-ECCV-Indoor segmentation and support inference from rgbd images

### Others

###### 11-ICCVW-A category-level 3-d object dataset: Putting the kinect to work

###### 11-NIPS-Semantic labeling of 3d point clouds for indoor scenes

###### 12--Kinect@Home

# Full 3D models with semantics

### RGB-D images

### Camera poses

### Object segmentations

### Point clouds registered into a global coordinate frame

# Works combine recognition and reconstruction

### 11-CVPR-Semantic structure from motion

### 13-CVPR-SLAM++: Simultaneous localization and mapping at the level of objects

### 09-ICCV-Multiple view semantic segmentation for street view images

### 13-CVPR-Joint detection, tracking and mapping by semantic bundle adjustment

### 13-CVPR-Joint 3d scene reconstruction and class segmentation

# Real-time SLAM

#### 10--RGB-D mapping: Using depth cameras for dense 3d modeling of indoor environments

#### 12-ICRA-An evaluation of the RGB-D SLAM system

#### 11-ISMAR-Kinectfusion: Real-time dense surface mapping and tracking

# Data-driven bruuce-force SfM

## ||Tracking (Registering neighboring frames)||

### Matching key-points using SIFT and remove poor matches using the ratio test

### 3-point(3D) RANSAC

## ||Loop closure||

### Bag of Words model to compute a feature vector for each frame in the tf-idf manner

### compute the dot product between all pairs obtain the score matrix for possible pairs

### With the score matrix, we use Gaussian smoothing, non-maximum suppression, and then dilation to pick the list of possible pairs

### frame-to-frame registration

##### more than 25 SIFT key-point correspondences

## ||Optimization (Joint 2D+3D bundle adjustment)||

## ||Depth map improvement||

### TSDF

# Multi-view object annotation

## ||Interaction at each frame||

### Propagation algorithm will try to propagate labels from keyframes to this new frame

##### Correct all mistakes if there are any

##### Continue to other frames

## ||Polygon reconstruction||

### …

## ||Annotation propagation||

### Retrieve the closest two keyframes based on frame number

### Reproject the 3D object polygons into the current frame using the estimated poses

### Multiple polygons are merged together by the union of the polygons from multiple frames

## ||Conflict list||

### when a user corrects a major mistake in a frame

##### the algorithm checks to see which keyframes the wrong label is propagated from

##### places them into the conflict list with the current frame.

##### All nearby frames will exclude frames from the conflict list during propagation

## ||Instance naming||

### ask the user to name an object first for its object category

# Generalized bundle adjustment

### 

# Plane semantics

# Object context