

http://vision.princeton.edu

PRINCETON VISION GROUP PanoContext: A Whole-room 3D Context Model for Panoramic Scene Understanding NUS

Jianxiong Xiao (Princeton) Ping Tan (NUS) Yinda Zhang (Princeton) Shuran Song (Princeton)



Motivation

Narrow feld of view hinders context information:

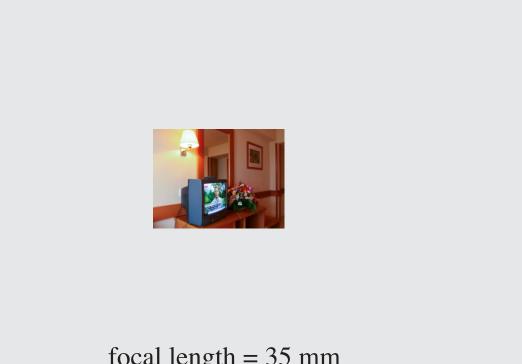
- 1. little interplay among objects: small # of objects per image
- 2. occurrance is unpredictable: missing bed in bedroom

Data driven context model

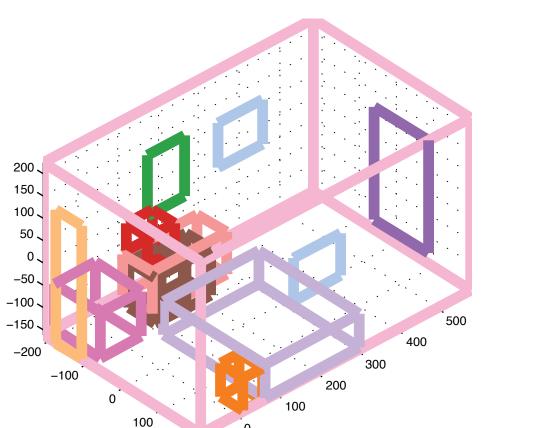
- 1. should be built in 3D space
- 2. should be learnt from big data







focal length = 35 mmWhat a camera see



Whole-room model

Our Task



Input: a single-view panorama

Output: object detection



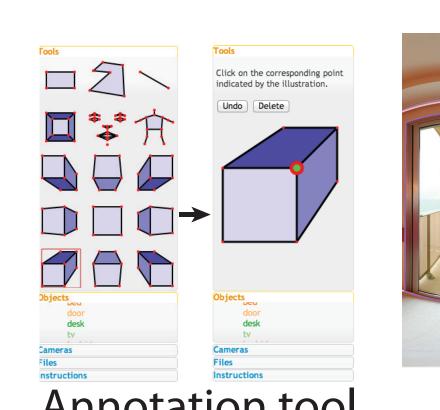
Output: 3D reconstruction

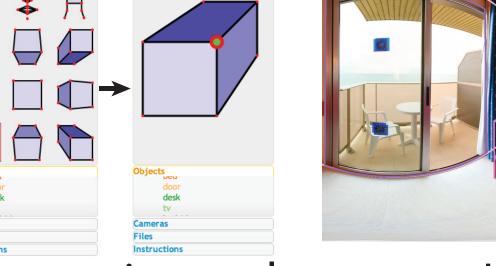
SUN360 Dataset & Ground Truth





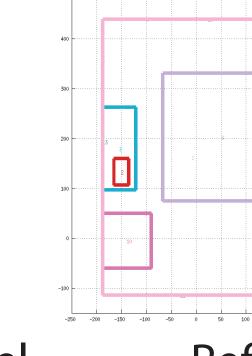
http://vision.cs.princeton.edu/projects/2012/SUN360/data/

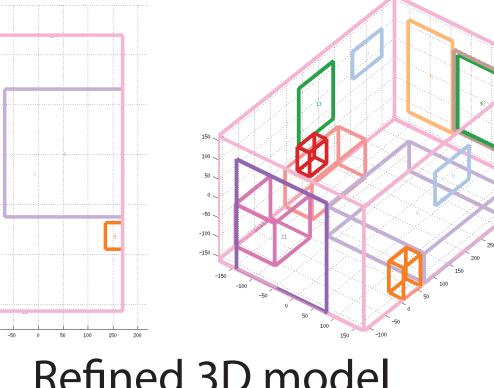












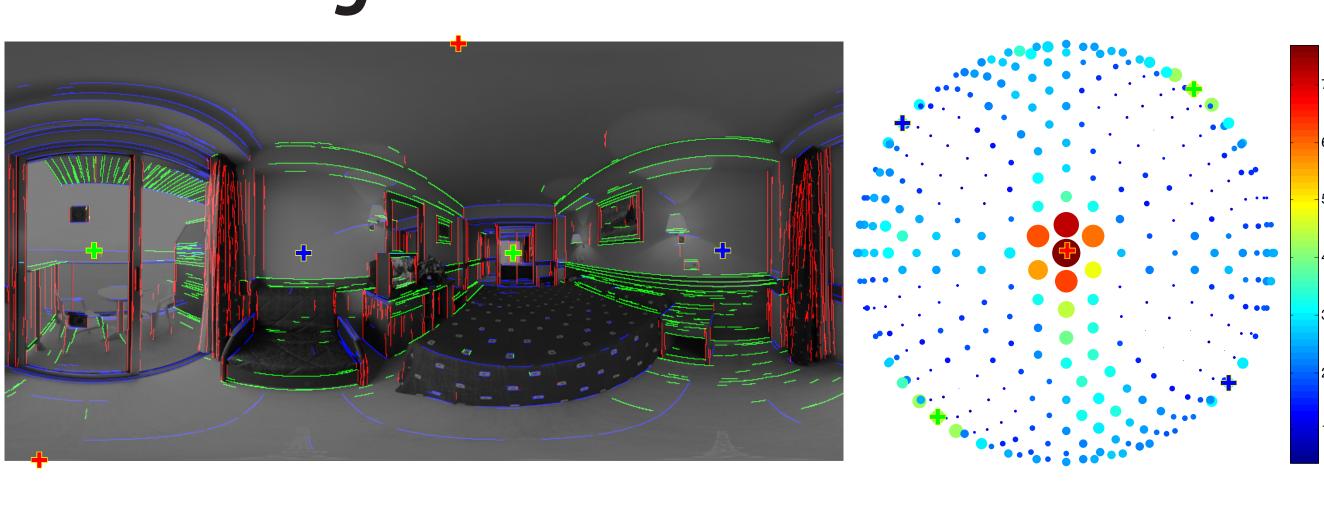
User annotation

Minimize reprojection error under constraints:

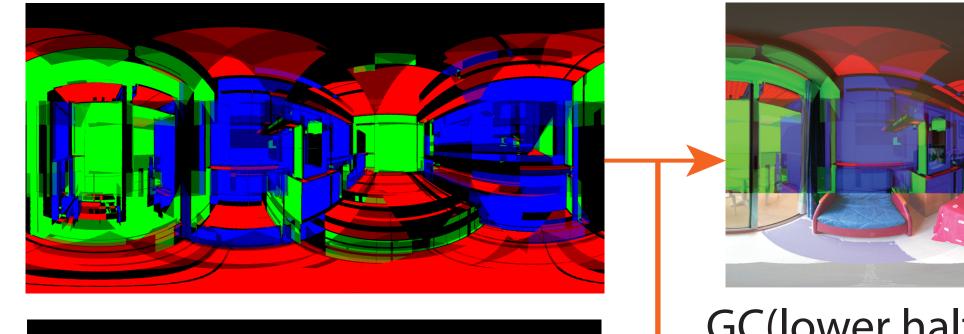
- 1. Annotation should be perfect cuboids. 2. Objects are axis-aligned if possible. 3. Objects cannot be outside of the room.
 - 4. No small gaps to walls.

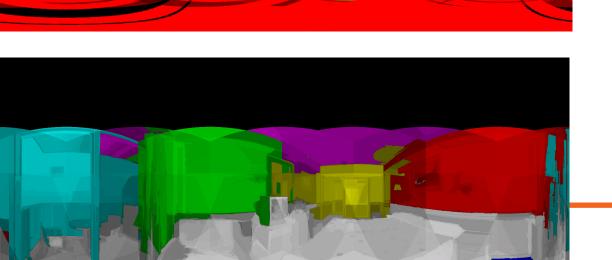
Method

1. Vanishing Point Estimation:



2. Room Layout Estimation:







GC(lower half)+OM(upper half)



Hough Transform:

Vote line segments to uniformly

and select the three orthogonal

directions with maximal sum of

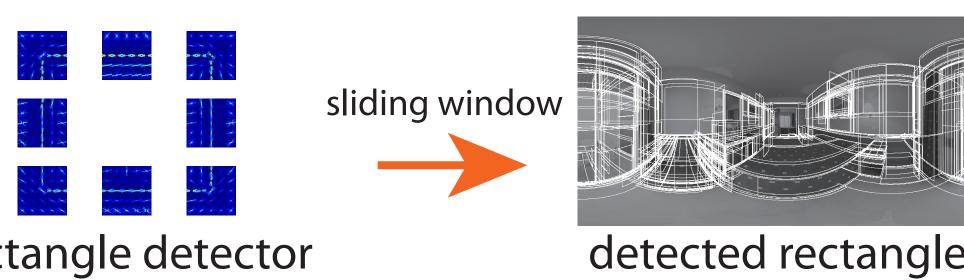
votes as the vanishing direction.

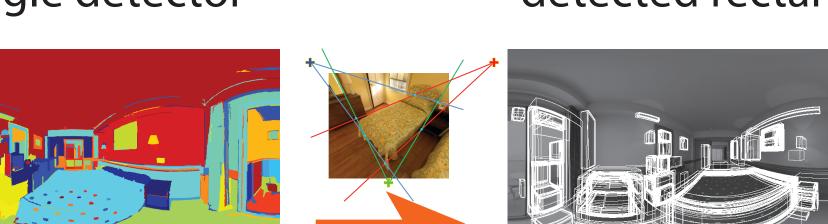
sampled directions on sphere,

5-line sampling:

1. Randomly sample 5 line segments to form a room layout hypothesis. . Use the surface normal consistency with GC and OM to rank these hypotheses and choose the top 50. This guarantees recall.

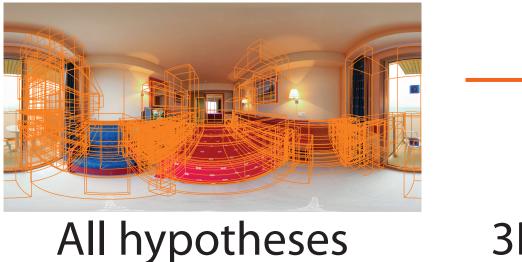
3. Object Hypothesis Generation:









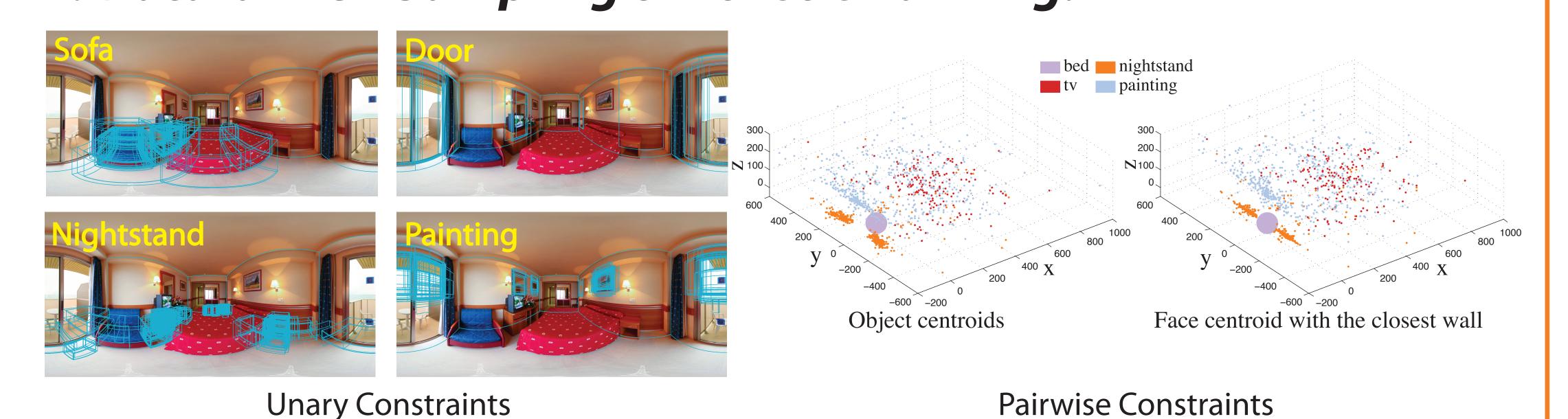


Cuboid hypotheses:



- 1. Run rectangle detector by sliding window. 2. Fit hexagon on segments by sampling 6 lines, two from each vanishing points.
- 3. Pop up 2D rectangles and hexagons to 3D space given a room layout.

4. Data-driven Sampling & Holistic Ranking:



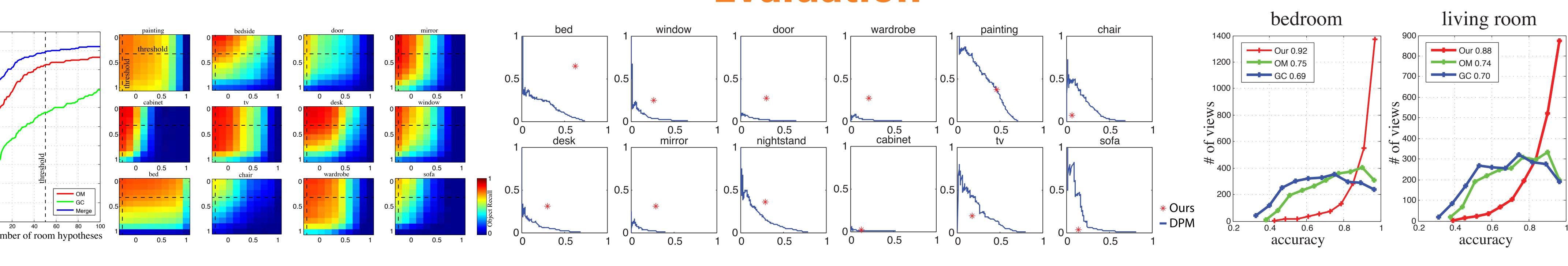
Whole-room sampling to generate complete scene hypothesis:

- 1. Randomly select a room layout according to their score evaluated by GC and OM.
- 2. Decide the number of instances for each type of objects.
- 3. Decide the sampling order for different object types according to statistic prior.
- 4. Start from the first object. Search for cuboids of the selected object type, and randomly choose a cuboid according to the unary constraints (rectangle scores, unary size distribution, random forest score).
- 5. Go to the next object, randomly select cuboids further considering the pairwise constraint with already selected objects. Repeat until all the types are sampled.



Holistic Ranking:

- 1. Convert all bottom up and top down information of all object categories to a feature vector.
- 2. In training dataset, use matching cost with ground truth to sparate sampled scenes to pos. and neg.
- 3. Train a binary SVM classifier. 4. During testing, take the top several scenes with highest SVM score.
- 5. Randomly disturb these top rank scenes, take the scene with the highest SVM score.



Room Layout & Object Recall X: Image Score, Y: Semantic Score

Object Detection Compare with DPM X: Recall, Y: Precision

Pixelwise Accuracy of **Room Layout Estimation**

References

- 1. PanoContext: Accepted by ECCV 2014. Code and dataset will be available soon.
- 2. Project webpage: http://panocontext.cs.princeton.edu
- 3. J. Xiao, J. Hays, B. C. Russell, G. Patterson, K. A. Ehinger, A. Torralba and A. Oliva, Basic level scene understanding: categories, attributes and structures, Frontiers in Psychology, Volume 4, Number 00506, 2013

More Results

Input Image & Reprojected 2D Result



3D Result



