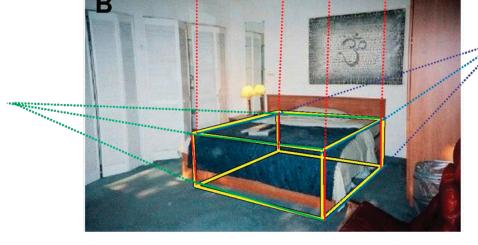
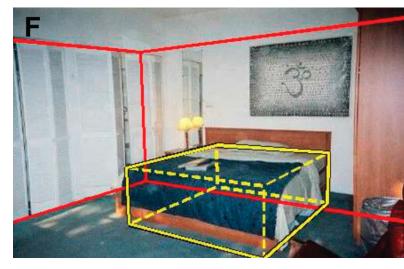


Motivations

1. Standalone boxy object detector

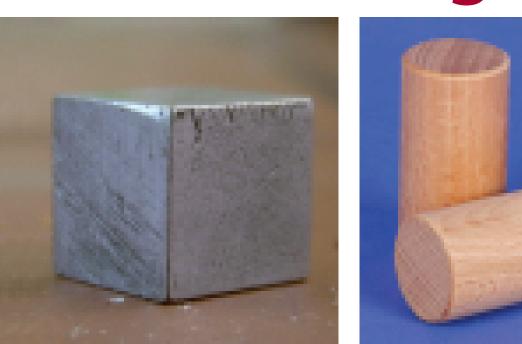




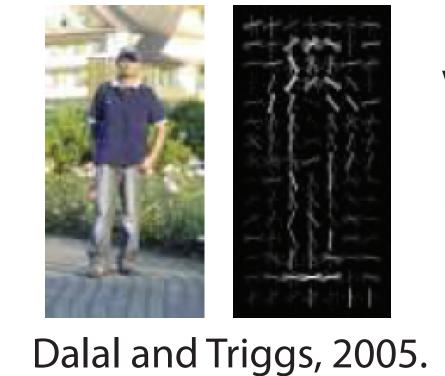
Thinking inside the box: Using appearance models and context based on room geometry. Hedau, Hoiem and Forsyth, 2010.

2. Cuboidness: not category



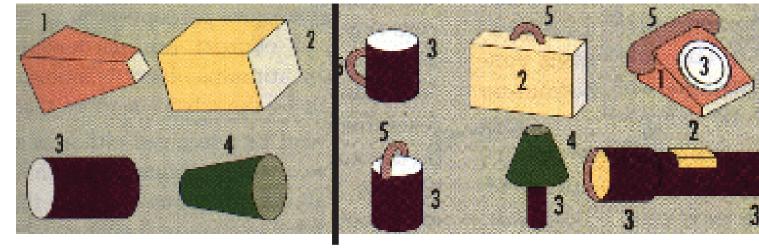


3. View-invariant 3D detector



View-based model? 3D view-invariant model? Mental rotation?

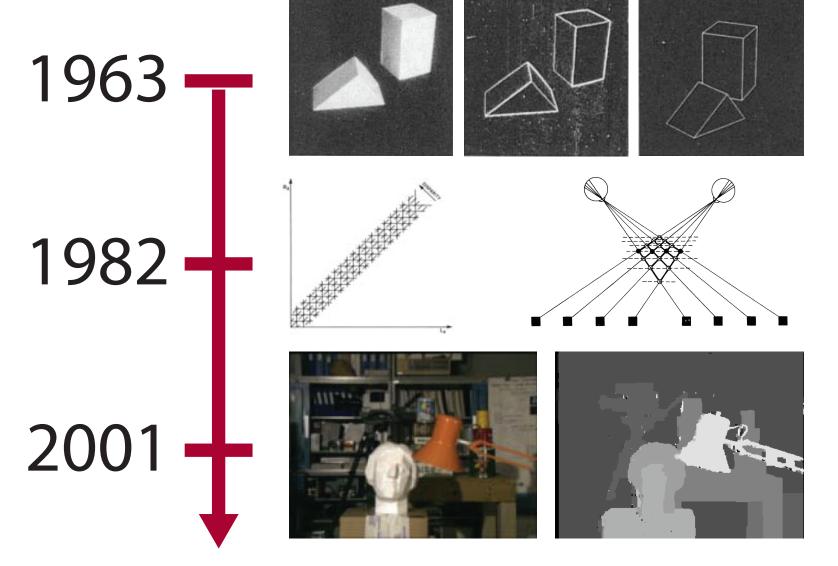
4. Geon and RBC: recognition by geometric primitive components



Recognition by components: a theory of human interpretation. Biederman, 1987.

5.3D reconstruction is not just low-level

Given a depth map, we still cannot manipulate an object. Because there is no high level representation of the 3D world. Obtaining a good depth map = = Inventing a digital camera.



Cuboids

Pyramid

Machine perception of three-dimensional solids. Roberts, 1963.

Stereo correspondence.

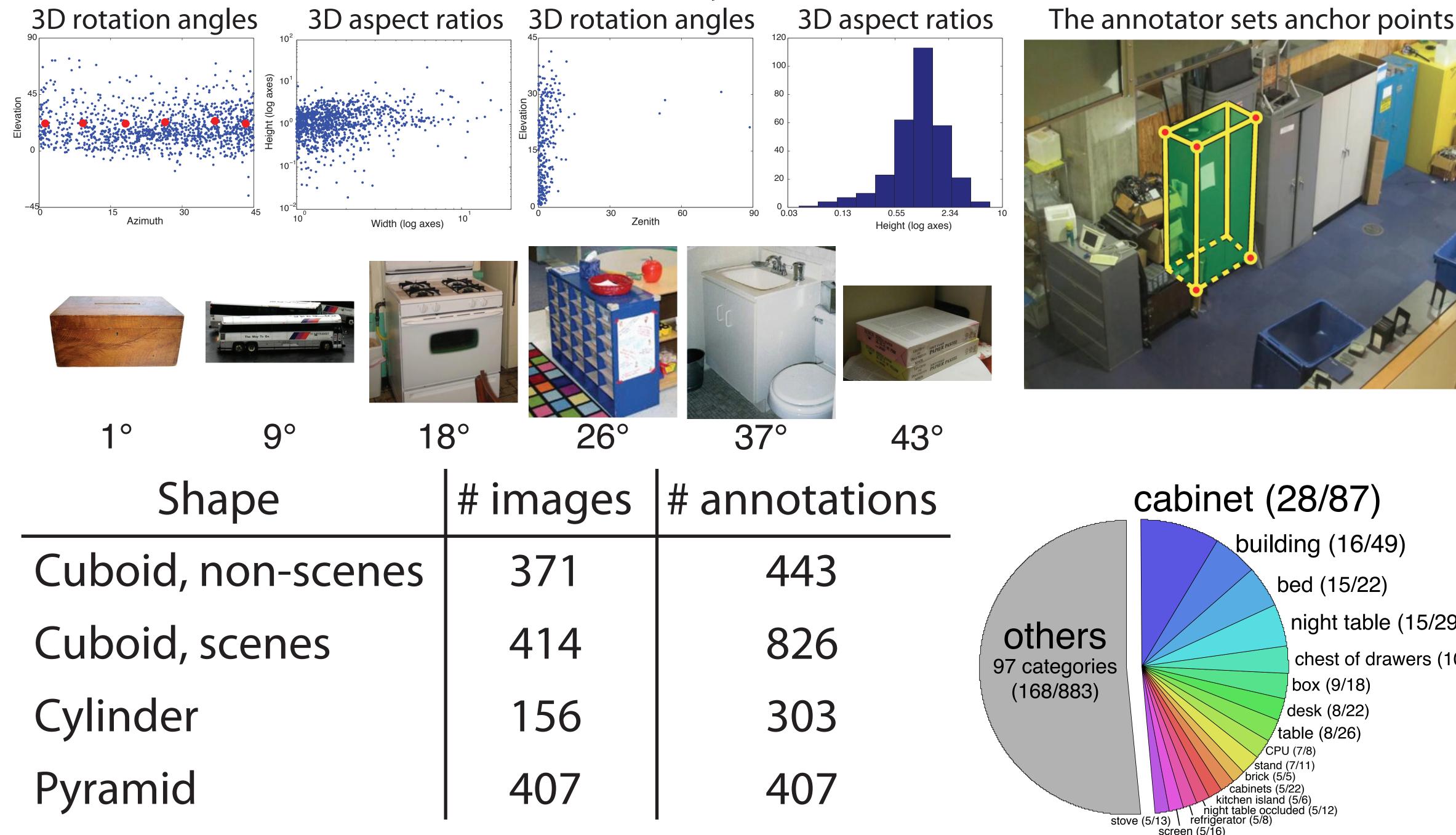
Dev 1974; Marr and Poggio 1976; Marr, 1982.

Dense stereo matching. Boykov, Veksler and Zabih, 2001.

SUNprimitive Dataset

Cylinders





407

407

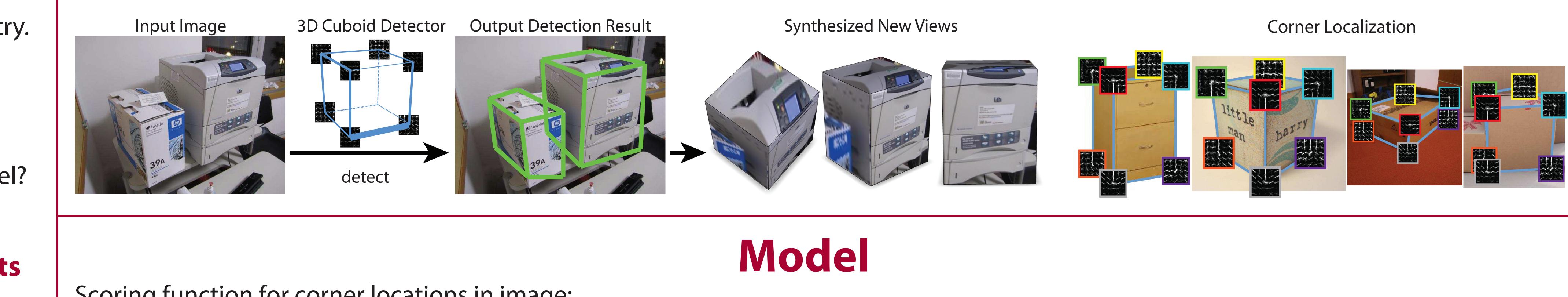
Localizing 3D Cuboids in Single-view Images

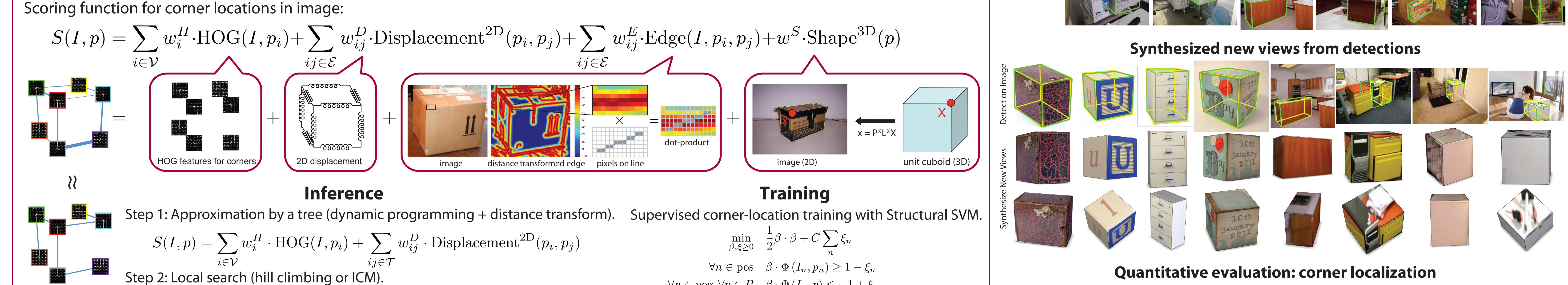
Jianxiong Xiao (MIT)

Bryan C. Russell (Intel Labs)

Task

Given a single-view input image, our goal is to detect the 2D corner locations of the cuboids depicted in the image. With the output part locations we can subsequently recover information about the camera and 3D shape via camera resectioning.





night table (15/29) chest of drawers (10/10)



Antonio Torralba (MIT)

Ground truth

 $\forall n \in \operatorname{neg}, \forall p \in P \quad \beta \cdot \Phi(I_n, p) \leq -1 + \xi_n$

Results

SUNprimitive Dataset + Code: http://SUNprimitive.csail.mit.edu

Evaluation

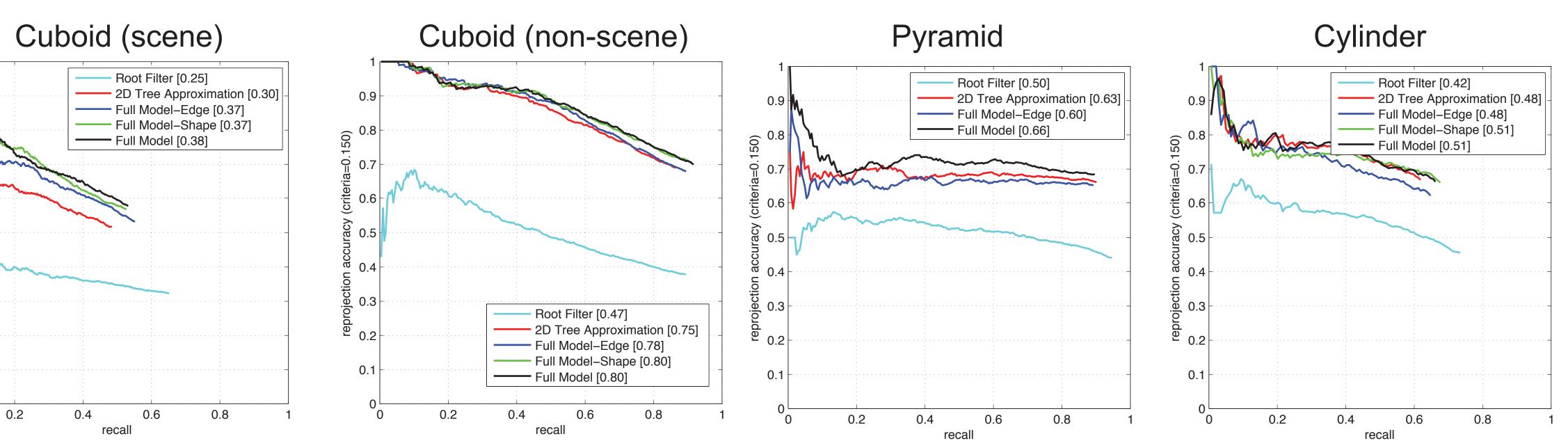
Comparison with 2D parts model



Quantitative evaluation: corner localization

• Corner location is correct if it lies within threshold.

• Threshold is set to a fraction of the ground truth bounding box.



Acknowledgements

J.X. is supported by Google U.S./Canada Ph.D. Fellowship in Computer Vision.

B.C.R. was funded by the Intel Science and Technology Center for Pervasive Computing (ISTC-PC). ⁻his work is funded by ONR MURI N000141010933 and NSF Career Award No. 0747120 to A.T.

References

[1] Articulated pose estimation using flexible mixtures of parts. Yang and Ramanan. In CVPR, 2011 [2] Recognizing scene viewpoint using panoramic place representation. Xiao, Ehinger, Oliva and Torralba. In CVPR, 2012.

[3] Reconstructing the world's museums. Xiao and Furukawa. In ECCV, 2012.

[4] Basic level scene understanding: From labels to structure and beyond. Xiao, Russell, Hays,

Ehinger, Oliva, and Torralba. In SIGGRAPH Asia, 2012.