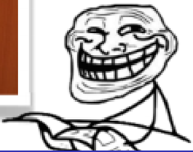
PanoContext: A Whole-room 3D Context Model for Panoramic Scene Understanding

Yinda Zhang Shuran Song Ping Tan† Jianxiong Xiao

Princeton University † Simon Fraser University

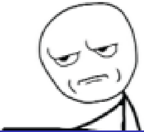
Importance of Context in Image Processing

Importance of Context in Image Processing



Importance of Context in Image Processing

Importance of Context in Image Processing



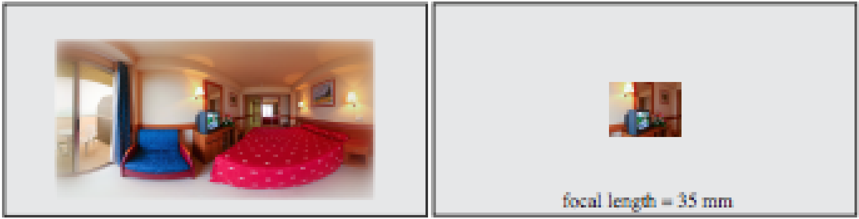
Importance of Context in Image Processing

Context Models

- Proposed in the past, but they do not work as well as expected
- WHY?

Field of View (FOV)

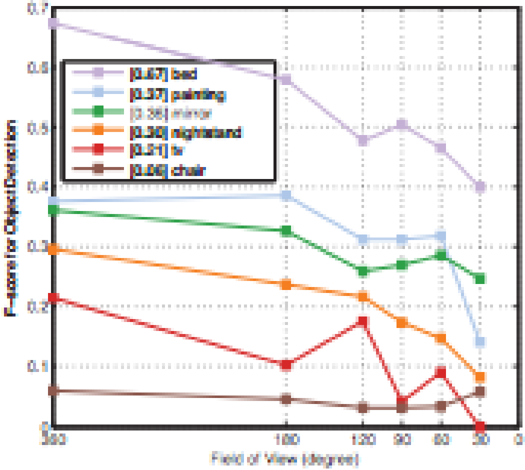
Small FOV $\Longrightarrow$ Hinders the contextual information



What your eyes see What a camera sees

Field of View (FOV)

Field of View (FOV)



F-score with different FOV

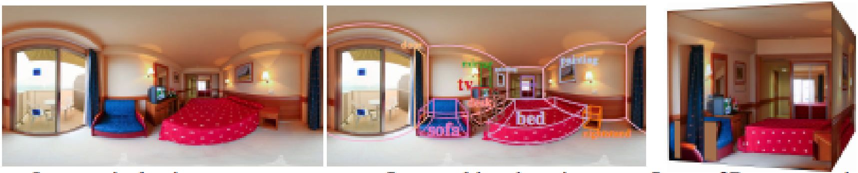
Panorama - 360° Horizontal and 180° Vertical

- Easily obtained by smartphones, special lenses, or image stitching
- All objects are usually visible despite occlusion
    - Enables the detection of the room layout and of the contextual information
- Panorama $\Longrightarrow$ Object Detection $\Longrightarrow$ Whole-room 3D Context Model

Input: a single-view panorama   Output: object detection   Output: 3D reconstruction

Panorama - 360° Horizontal and 180° Vertical

- Easily obtained by smartphones, special lenses, or image stitching
- All objects are usually visible despite occlusion
    - Enables the detection of the room layout and of the contextual information
- Panorama $\Longrightarrow$ Object Detection $\Longrightarrow$ Whole-room 3D Context Model



Input: a single-view panorama   Output: object detection   Output: 3D reconstruction

[1]Manhattan world assumption: assumes the scene consists of 3D cuboids aligned with the three principle directions

[2]Assume no floating objects

PanoContext: A Whole-Room 3D Context Model

Algorithm

1. Generate a set of hypotheses for room layout and objects
2. Rank these whole-room hypotheses holistically to determine the best hypothesis

PanoContext: A Whole-Room 3D Context Model

1. Generate a set of whole-room hypotheses

1. Generate a set of whole-room hypotheses

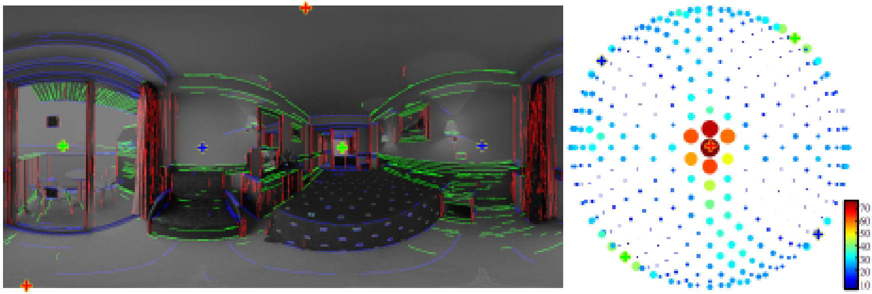
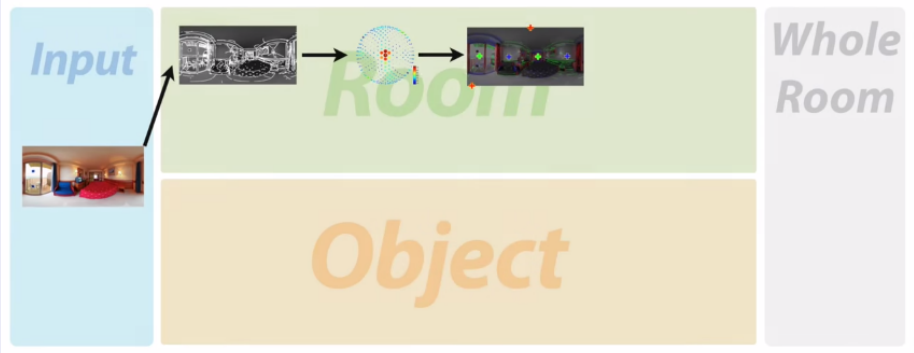


Figure : Hough transform for vanishing point detection

PanoContext: A Whole-Room 3D Context Model

1. Generate a set of whole-room hypotheses



PanoContext: A Whole-Room 3D Context Model

1. Generate a set of whole-room hypotheses

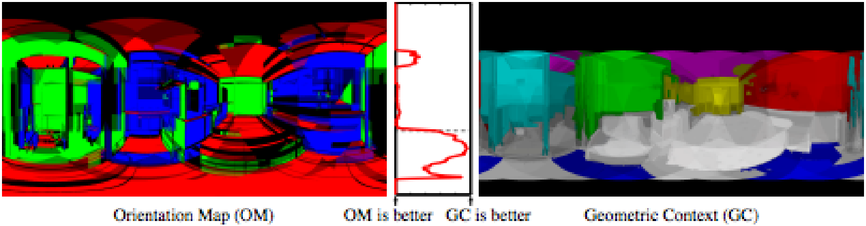
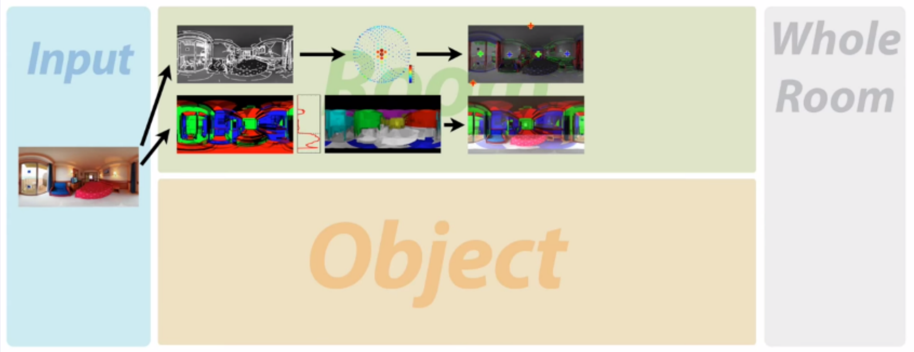


Figure : Comparison of OM and GC. OM works better on the top half of the image, while GC provides better normal estimation at the bottom

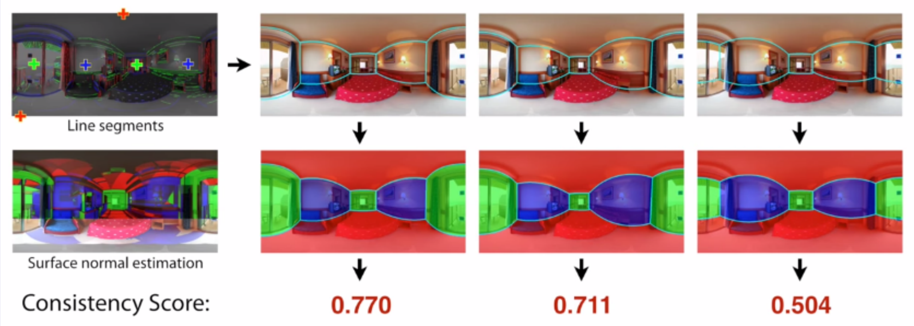
PanoContext: A Whole-Room 3D Context Model

1. Generate a set of whole-room hypotheses



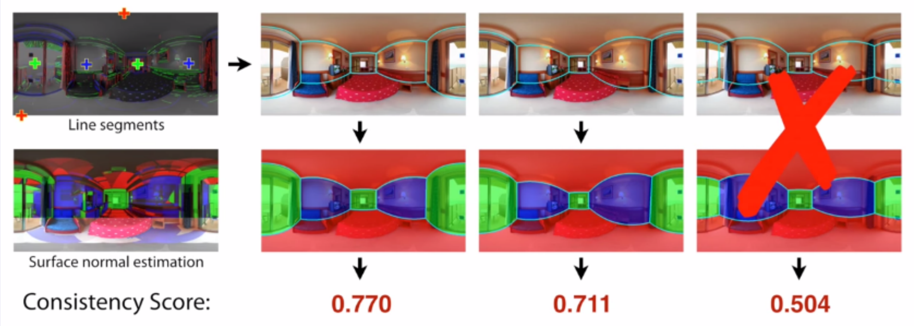
PanoContext: A Whole-Room 3D Context Model

1. Generate a set of whole-room hypotheses



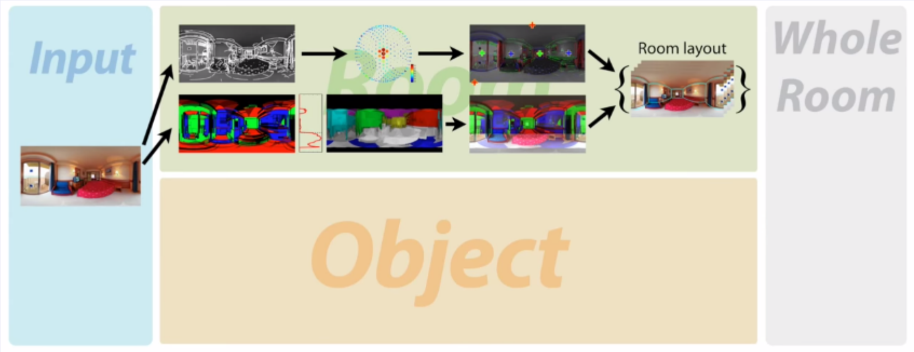
PanoContext: A Whole-Room 3D Context Model

1. Generate a set of whole-room hypotheses



PanoContext: A Whole-Room 3D Context Model

1. Generate a set of whole-room hypotheses



PanoContext: A Whole-Room 3D Context Model

1. Generate a set of whole-room hypotheses

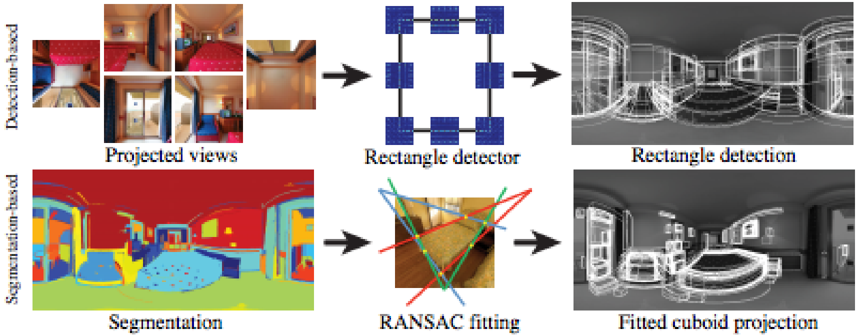
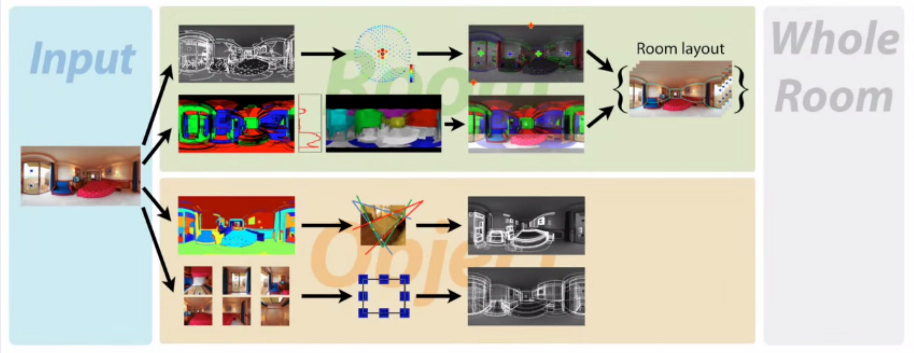


Figure : Two ways to generate object hypotheses

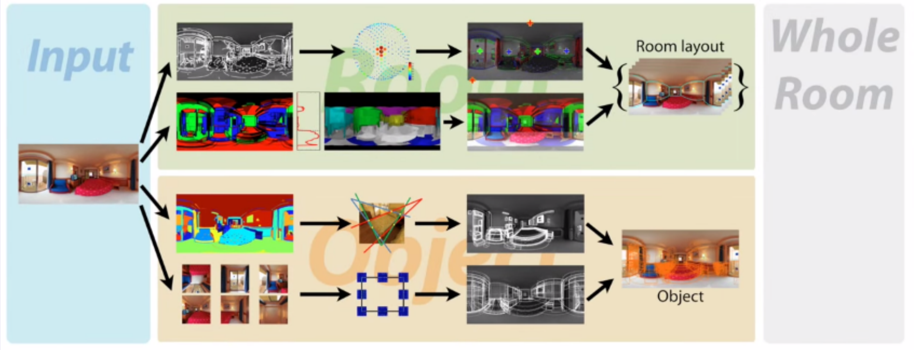
PanoContext: A Whole-Room 3D Context Model

1. Generate a set of whole-room hypotheses



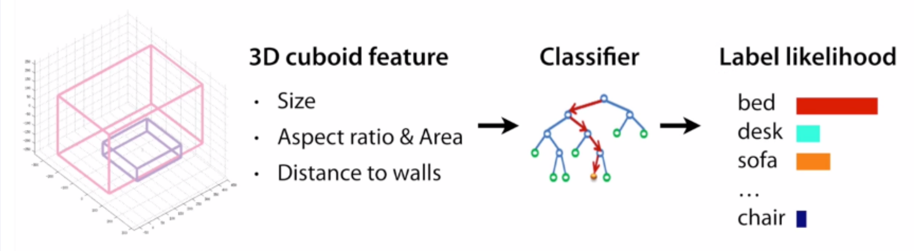
PanoContext: A Whole-Room 3D Context Model

1. Generate a set of whole-room hypotheses



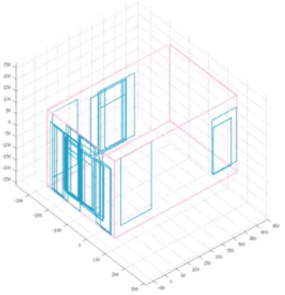
PanoContext: A Whole-Room 3D Context Model

1. Generate a set of whole-room hypotheses



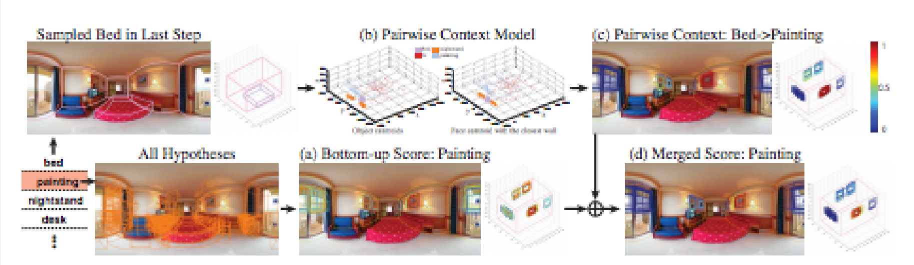
[1]Semantic Label

1. Generate a set of whole-room hypotheses



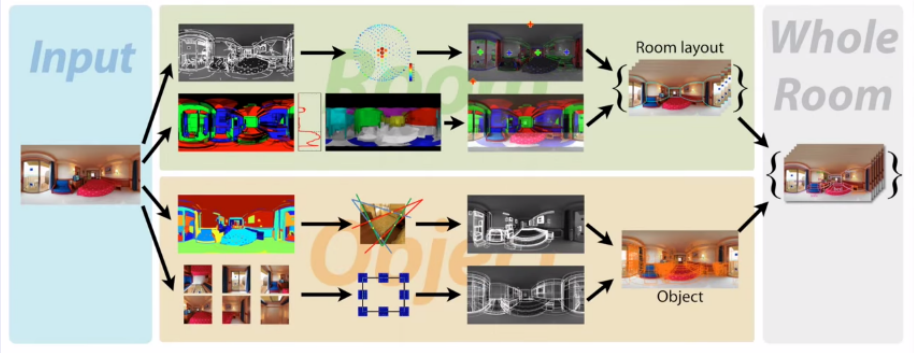
PanoContext: A Whole-Room 3D Context Model

1. Generate a set of whole-room hypotheses



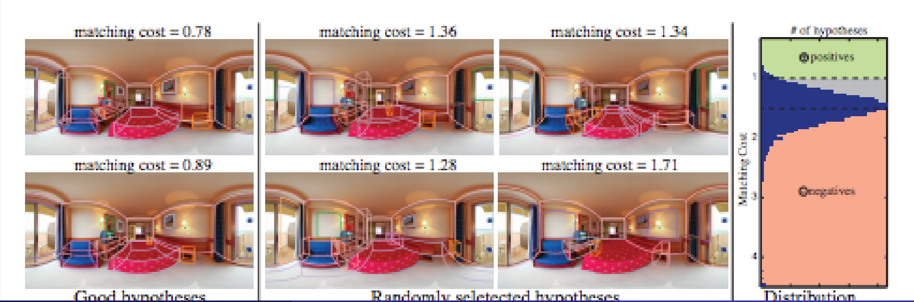
[1] Pairwise Constraint

1. Generate a set of whole-room hypotheses



2. Determine the best whole-room hypotheses

- Train a linear SVM model to rank the whole-room hypotheses and choose the best hypothesis
- Want the matching cost (difference between whole-room hypothesis and its ground truth) to be as low as possible



Summary: How does context help?

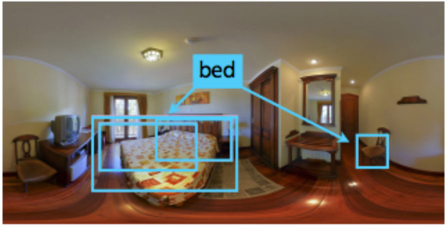
- Helps to determine the size of objects

DPM: Wrong relative size

PanoContext

Summary: How does context help?

- Helps to determine the size of objects
- Helps to determine the correct number of objects

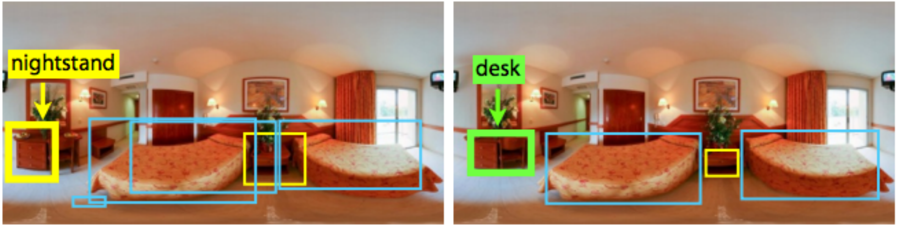


DPM: Wrong number of objects

Our detection

Summary: How does context help?

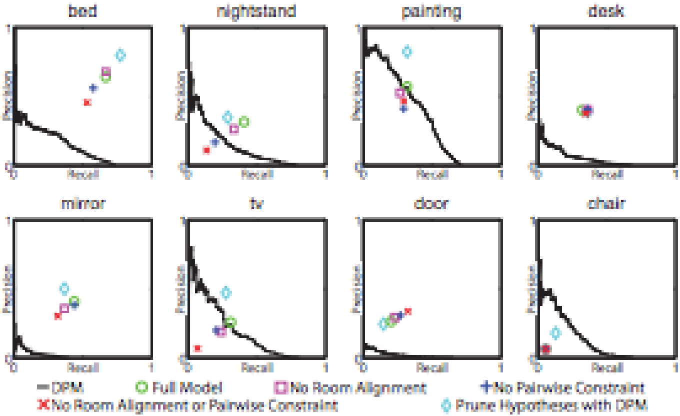
- Helps to determine the size of objects
- Helps to determine the correct number of objects
- Helps the determine the relative position of objects



DPM: Wrong relative position

Our detection

Summary: How does context help?



Precision-recall comparison with DPM

[1] DPM: Deformable Part Model

[2] Felzenszwalb et. al: Discriminative training with partially labeled data

Contributions

- Context model is fully in 3D
- First annotated panorama dataset

Questions?