Efficient 3D Room Shape Recovery from a Single Panorama

# Abstract

### Our algorithm can automatically infer a 3D shape from a collection of partially oriented **superpixel facets** and **line** **segments**

### The core part of the algorithm is a constraint graph, which includes

##### Lines and superpixels as vertices

##### Their geometric relations as edges

### Based on the constraint graph by solving all the geometric constraints as constrained linear least-squares

### The selected constraints used for reconstruction are identified using an occlusion detection method with a Markov random field

# Introduction

## ||Indoor World model||

### Consists of a single floor, a single ceiling, and vertical walls

###### 06-CVPR-A dynamic bayesian network model for autonomous 3d reconstruction from a single indoor image

###### 09-CVPR-Geometric reasoning for single image structure recovery

### Estimating a cuboid shape that fits the room layout

###### 09-ICCV-Recovering the spatial layout of cluttered rooms

###### 10-ECCV-Discriminative learning with latent variables for cluttered indoor scene understanding

###### 10-ECCV-Thinking inside the box: Using appearance models and context based on room geometry

###### 10-NIPS-Estimating spatial layout of rooms using volumetric reasoning about objects and surfaces

###### 13-CVPR-Understanding indoor scenes using 3d geometric phrases

###### 12-CVPR-Efficient structured prediction for 3d indoor scene understanding

###### 13-ICCV-Box in the box: Joint 3d layout and object reasoning from single images

###### 13-CVPR-Manhattan junction catalogue for spatial reasoning of indoor scenes

###### 14-ECCV-Panocontext: A whole-room 3d context model for panoramic scene understanding

### Most of these approaches work in a discretized manner

##### Results are selected from a set of candidates based on certain scoring functions

##### The generation rules of the candidates limit the scope of these algorithms

## ||Manhattan world assumption||

### Constraint graph, which includes all the lines and superpixels as vertices is constructed

### A 3D reconstruction is performed in an iterative manner, which can solve constraints as constrained linear least squares (CLLS)

### We propose an occlusion detection method using a Markov random field (MRF) to select plausible constraints for reconstruction

# Related work

## ||single-view reconstruction (SVR)||

### Geometric approaches

##### Recovered indoor layout by computing orientation maps (OMs) from lines

09-CVPR-Geometric reasoning for single image structure recovery

##### Reconstructed symmetric objects by detecting symmetric lines

11-CVPR-Symmetric piecewise planar object reconstruction from a single image

##### Recognized cuboids from single photographs based on both the appearance of corners and edges as well as their geometric relations

12-NIPS-Localizing 3d cuboids in single-view images

##### Lifted 2D lines to 3D by identifying true line intersections in space

13-ICCV-Lifting 3d manhattan lines from a single image

### Semantic approaches

##### Hoiem et al estimated geometric context label on each image pixel

05-TOG-Automatic photo pop-up

07-IJCV-Recovering surface layout from an image

##### Delage et al inferred room layout via floor–wall boundary estimation

06-CVPR-A dynamic bayesian network model for autonomous 3d reconstruction from a single indoor image

##### Gupta et al proposed blocks world to predict 3D arrangement by explaining their volumetric and physical relations

10-ECCV-Blocks world revisited: Image understanding using qualitative geometry and mechanics

### Semantic approaches (Manhattan direction)

##### Hedau et al utilized structured learning to improve prediction accuracy

09-ICCV-Recovering the spatial layout of cluttered rooms

##### The inferences of both indoor objects and the box-like room layout have been continuously improved thereafter to enhanced object presentation, novel features or more efficient inference techniques

10-ECCV-Discriminative learning with latent variables for cluttered indoor scene understanding

10-ECCV-Thinking inside the box: Using appearance models and context based on room geometry

10-NIPS-Estimating spatial layout of rooms using volumetric reasoning about objects and surfaces

13-CVPR-Understanding indoor scenes using 3d geometric phrases

12-CVPR-Efficient structured prediction for 3d indoor scene understanding

13-ICCV-Box in the box: Joint 3d layout and object reasoning from single images

13-CVPR-Manhattan junction catalogue for spatial reasoning of indoor scenes

##### Zhang recently stressed the limitation of a narrow view angle imposed by standard camera, illustrated the advantage of panoramic images over normal photographs

14-ECCV-Panocontext: A whole-room 3d context model for panoramic scene understanding

##### Cabral took multiple indoor panoramas as input and utilized structural cues from single image for floorplan reconstruction

14-CVPR-Piecewise planar and compact floorplan reconstruction from images

### Our work focuses on 3D room shape recovery

##### Related to

13-ICCV-Lifting 3d manhattan lines from a single image

A constraint graph is proposed, its vertices are fully orientated lines, and its edges are intersections/incidences between lines

##### Additional consideration

Including lines with unknown orientation, and planar superpixels with varying degrees of freedom (DOF)

##### Extended constraints

include connections between two superpixels or between a line and a superpixel

# Contributions

### An expressive constraint graph

### An iterative algorithm is proposed to solve the constraint graph as CLLS

### Identify occluding lines using an MRF

# Preprocessing

## ||Glance||

### Input is a panorama that covers a 360˝ horizontal field of view represented in an equirectangular projection

### One-to-one correspondence occurs between a panorama pixel and a 3D view direction

##### Use the term angle distance to measure the distance between two pixels by computing the angle of their directions

##### Use angle length to measure the length of a pixel sequence (like a line segment or a superpixel boundary) by accumulating the vector angles between adjacent pixels

## ||Similar approach||

#### 14-ECCV-Panocontext: A whole-room 3d context model for panoramic scene understanding

##### Used to detect lines, estimate Manhattan vanishing points, and identify the spatial directions of lines from the panorama

##### The vanishing point direction which is the most vertical in space is denoted as the vertical direction of the scene

##### Panorama version of graph cut is to over segment the panorama into superpixels

04-IJCV-Efficient graphbased image segmentation

## ||Priors for orientations of the superpixels||

### Wall prior

##### Assume regions near the horizon to be parts of the walls

##### super pixels whose angle distances to horizon are assigned to be vertical

### Vanishing point prior

##### assume that the superpixels with Manhattan vanishing points should face the corresponding Manhattan direction

### Geometric context prior(floor)

##### Geometric context prior

10-ECCV-Thinking inside the box: Using appearance models and context based on room geometry

##### Follows the same strategy with

14-ECCV-Panocontext: A whole-room 3d context model for panoramic scene understanding

##### We calculate the average label scores in each superpixel, and use the label with the maximum score to orient the superpixel

##### Only superpixels that have high scores in two labels are considered: ground and wall

# Constraint Graph

## ||Graph||

### Vertices

##### Lines

Two types

##### Superpixels

Three types with different DOF

### Edges

##### The constraints of connectivity

Adjacent superpixels

Intersections of lines

Adjacent lines lines and superpixels

##### The constraints of coplanarity

Adjacent facets that may lie on the same plane

## ||Vertex||

### Notion

##### the vector that encodes the unknown parameters of vectex

##### reprensent the set of known values of

##### is the DOF of

##### represents the distance from the plane to the viewpoint

##### is the unit normal

##### is the unit direction vector that the surface plane must be parallel with

### 

##### for plane equation

### Three types

##### : ,

##### : ,

##### ,

### Line

##### Supporting plane contains line and orthogonal to plane that contains viewpoint and line

## ||Constraint solving||

### Connection constraints

##### Notion

The set of connection points

Plane coefficients ,

[ignore constants]

##### Formulation

### Coplanarity constraints

##### Formulation

### Minimizing

##### Repeat

## ||Occlusion||

### Occlusions are likely to be covered by detected line segments, follow rule of coherence

### Notion

##### To each line with DOF=1

##### Label , front or behind

### Three kinds of evidence are used in occlusion detection

##### Orientation violation between lines and superpixels

##### T-junctions formed by lines

##### The coherence between collinear lines

### MRF to infer the of each by minimizing the objective

### Orientation Violation Cost

##### is the spatial orientation of line

##### be two closest vanishing points of , and satisfy

Lies on the {left, right} of

its spatial direction is orthogonal to

##### Neighborhoods

Sweep toward to an angle , denoted as

##### is then defined as the weighted number of violations

is a threshold

is the pixel weight used to rectify panorama distortion

is 1 if given orientation of the superpixel with conflicts with the direction of (A conflict occurs if the superpixel has DOF=1 and its plane normal coincides with the direction of )

is defined as

### T-junction

##### Commonly used

13-ICCV-Lifting 3d manhattan lines from a single image

##### T-junction cost is defined

, is the number of T-junction

### Coherence cost

##### For each collinear lines and

##### 

### Solution ( Convex max product)

###### 11-ICML-Convex max-product algorithms for continuous mrfs with applications to protein folding

## ||Graph construction||

### Notion

##### Connection points

##### Angle length of the boundary

### A superpixel pair that shares a same boundary is collected into and if its boundary is not covered by any occluding line

### A line and a superpixel that are adjacent in the panorama are collected into if the superpixel lies on the **front** side of the line

### Recognizing constraints between lines

##### Pair of lines with different orientations as intersecting if the angle distance between the line segments is

##### Collinearity is considered for lines with the same orientation that are separated by an angle

##### Connect the nearest points of these two lines and check whether the connection is intercepted by any occluding line

##### If any of is an occluding line, then we place them into only if

the other line lies on the front side of the occluding line

their connection is not intercepted

##### If neither is an occluding line, then is collected only if their connection is not intercepted

# 

## 