Beyond Point Clouds: Scene Understanding by Reasoning Geometry and Physics

Bo Zheng [1], Yibiao Zhao [2], Joey C. Yu [2], Katsushi Ikeuchi [1] and Song-Chun Zhu [2]

[1] The University of Tokyo, Japan, [2] University of California, Los Angeles

CVPR 2013 Portland, Oregon June 23-28

Overview of our approach

In this paper, we present an approach of scene understanding by reasoning physical stability of objects based on the input of point clouds. We utilize a simple observation:

- **By human design, objects in static scenes should be stable with respect to gravity.**

Our method consists of two major steps:

- **Geometric reasoning**: recovering solid 3D volumetric primitives from defective point cloud.
- **Physical reasoning**: grouping the unstable primitives to physically stable objects by optimizing the stability and the scene prior.

Our main contributions includes

- We define the physical stability function explicitly by studying minimum energy need to change the pose and position of an primitive (or object) from one equilibrium to another
- We introduce disconnectivity graph (DG) from physics (Spin-glass) to represent the energy landscapes.
- We solve the complex optimization problem of stability maximization by the sampling method Swendsen-Wang cut .

Geometric reasoning

Given a point cloud of scene, the goal of geometric reasoning is to recover a volumetric representation of object with physical properties, like volume, mass, support relation etc.

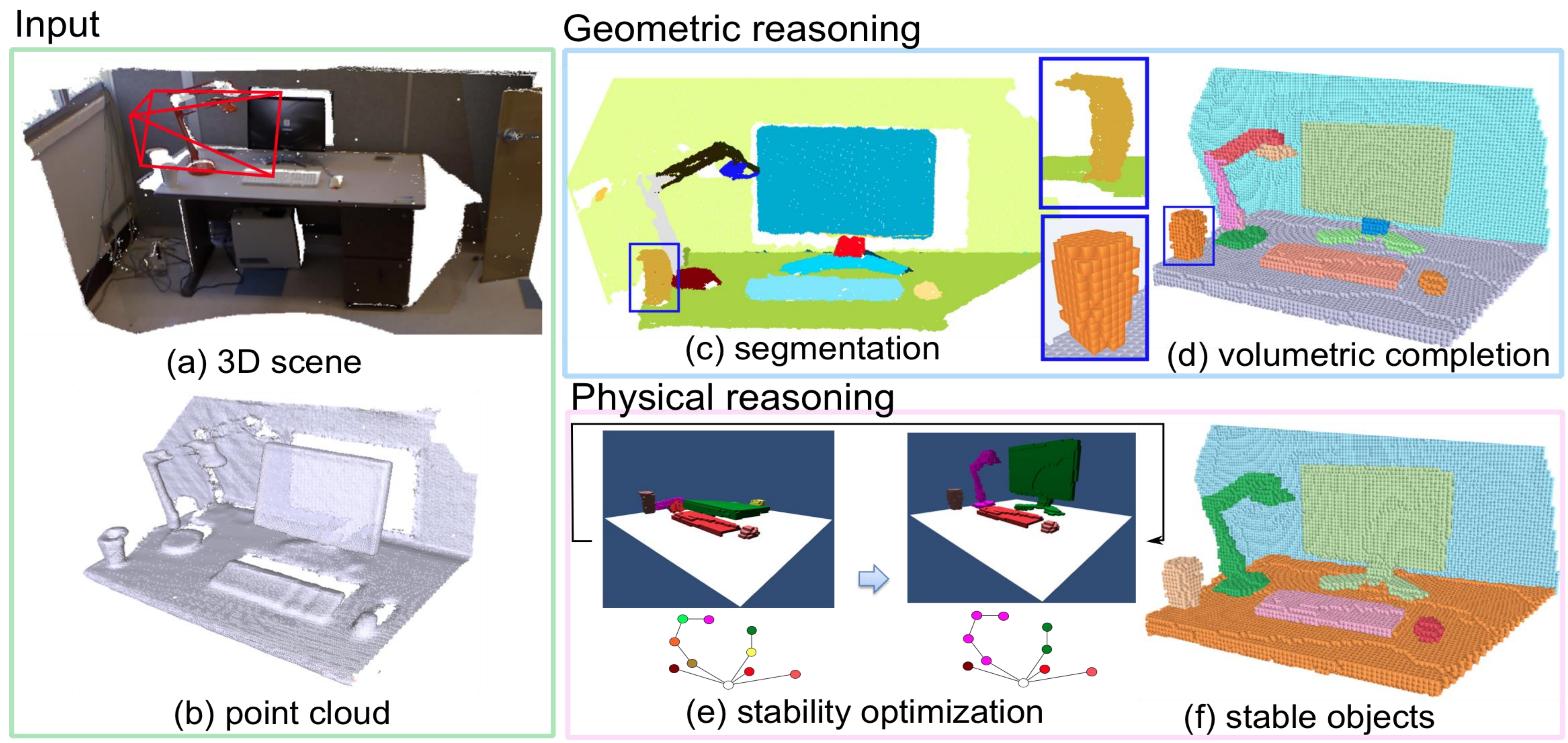
We first segment the point cloud with Implicit Algebraic Models (IAMs)

- Region growing segmentation by iterative IAMs fitting
- Further merging "convexly" connect regions.

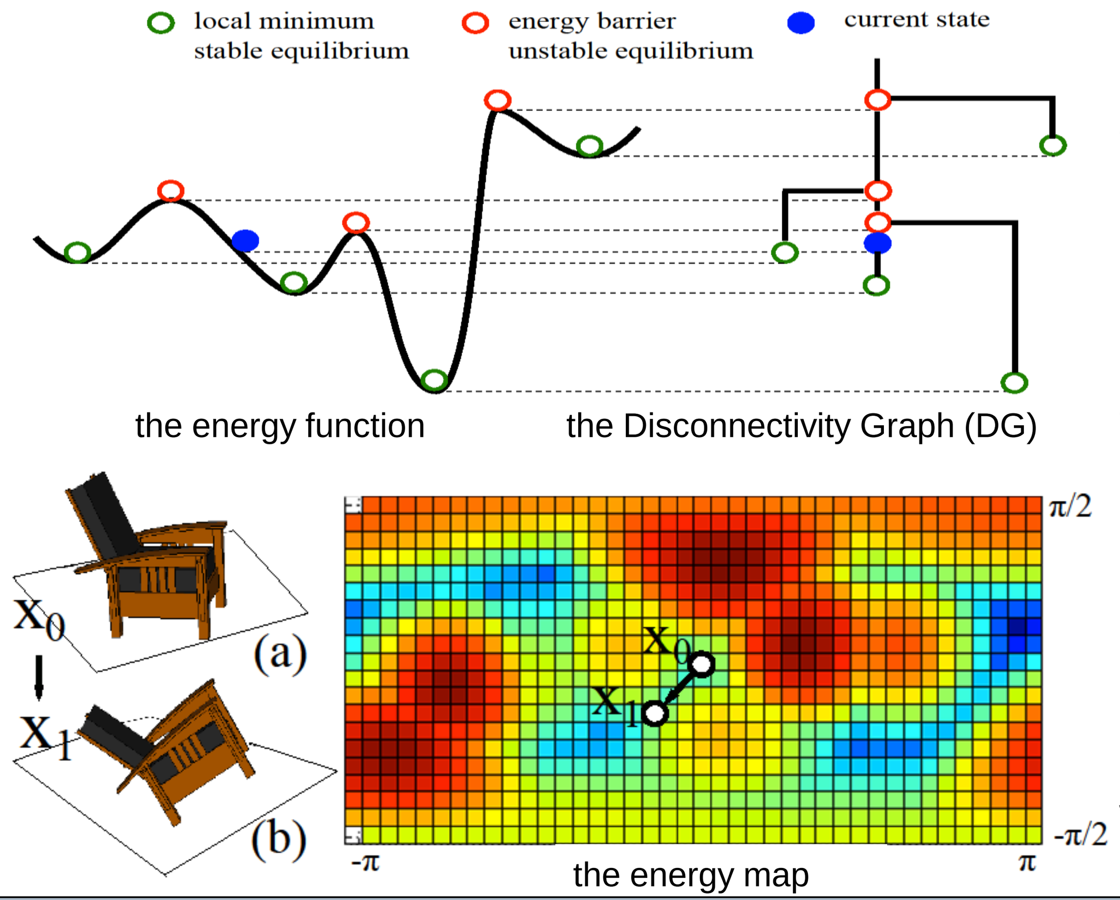
We then convert the defective point cloud segments to solid volumetric primitives.

- Estimation gravity direction and generating voxels.
- Estimating Invisible (occluded) space
- Filling missing voxels.

A illustration of our approach



Modeling object stability and Physical reasoning



- **Definition of stability**:

The stability $S(a, x_0, W)$ of an object a at state x_0 in the presence of a disturbance work W is the maximum energy that it can release when it moves out the energy barrier by the work W.

$$\mathrm{S}(a, \mathbf{x}_0, W) = \max_{\mathbf{x}'_0} \Delta\mathcal{E}(\tilde{\mathbf{x}} \to \mathbf{x}'_o)\delta([\min_{\tilde{\mathbf{x}}} \Delta\mathcal{E}(\mathbf{x}_o \to \tilde{\mathbf{x}})] \leq W)$$

The physical reasoning is then posed as a well-known graph partition problem, through which the unstable primitives can be grouped together to achieve the maximum global stability.

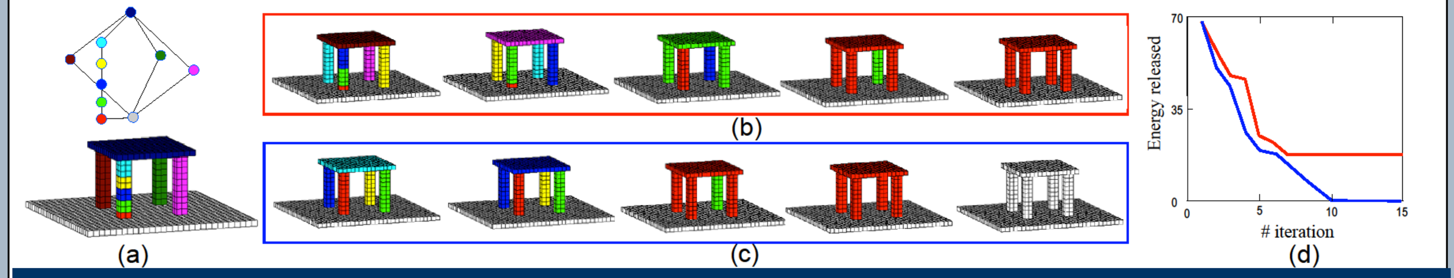
- **Inference of Maximum stability**:

$$E(L|G;W) = \sum_{O_i \in L} (\mathrm{S}(O_i, \mathbf{x}(O_i), W) + \mathcal{F}(O_i))$$

where L is the labels of graph partition, $x(O_i)$ is the current state of grouped object O_i, and $F(O_i)$ represents a penalty function for the object geometric prior e.g. the size and shape complexity.

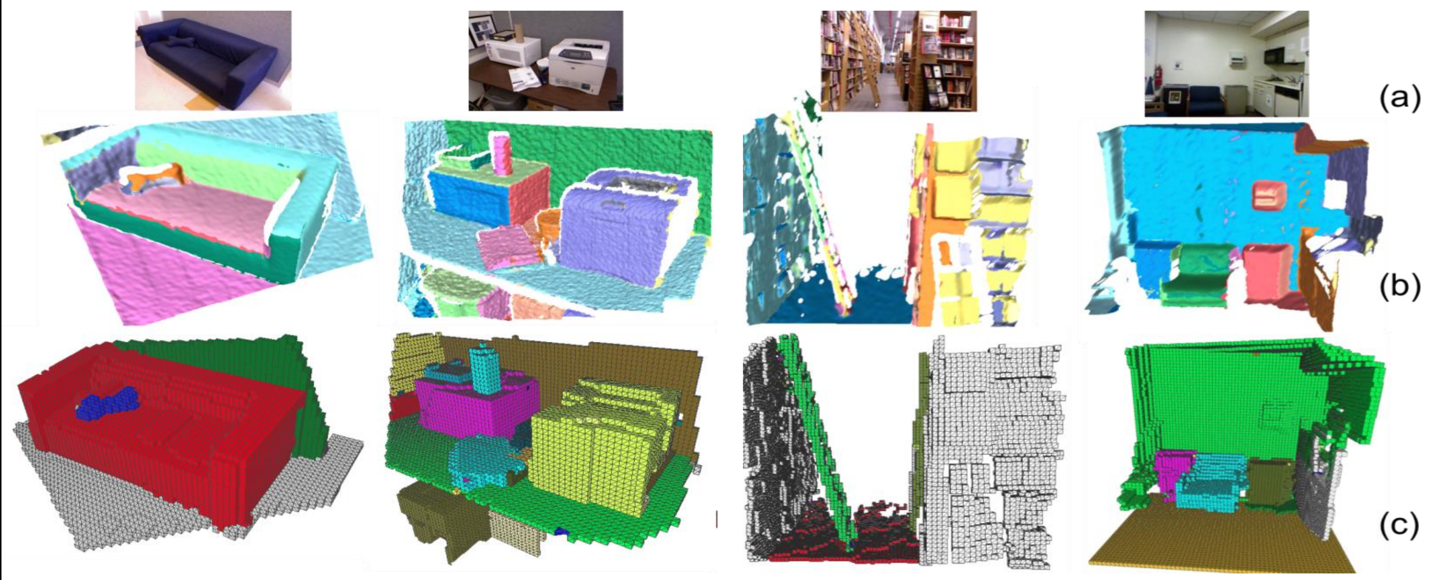
Problem Inference

The Swendsen-Wang cut is applied to solve the complex optimization problem by the sampling the partitioning of unstable primitives..

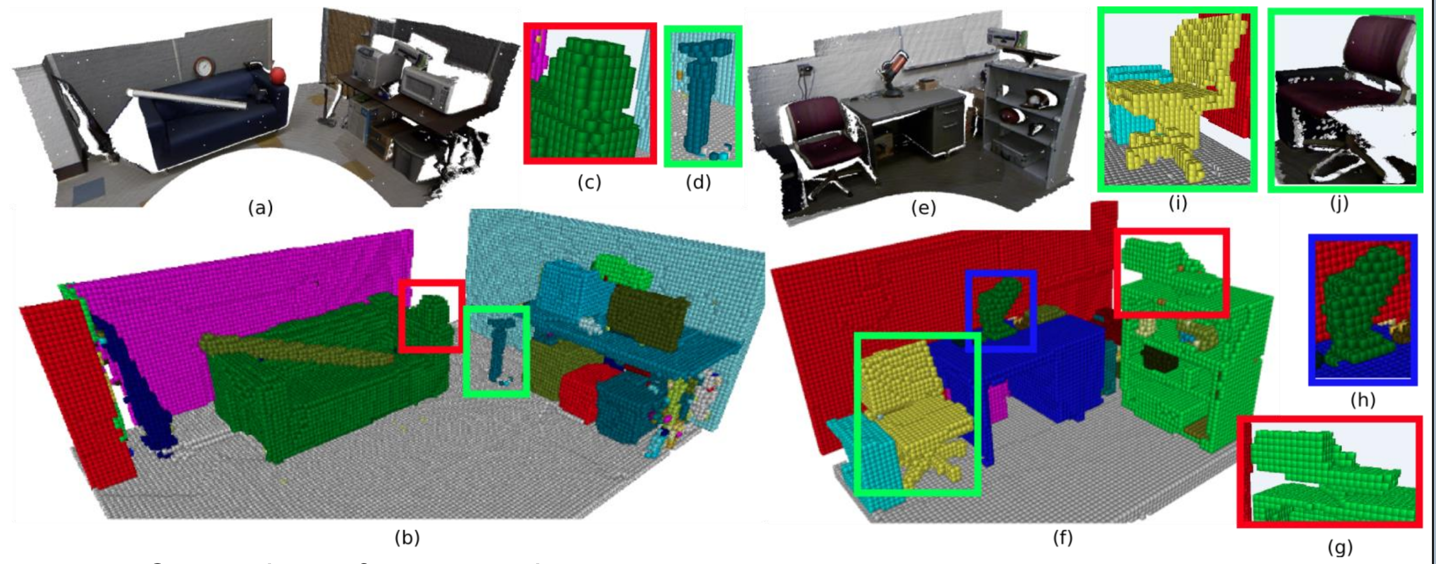


Experiment

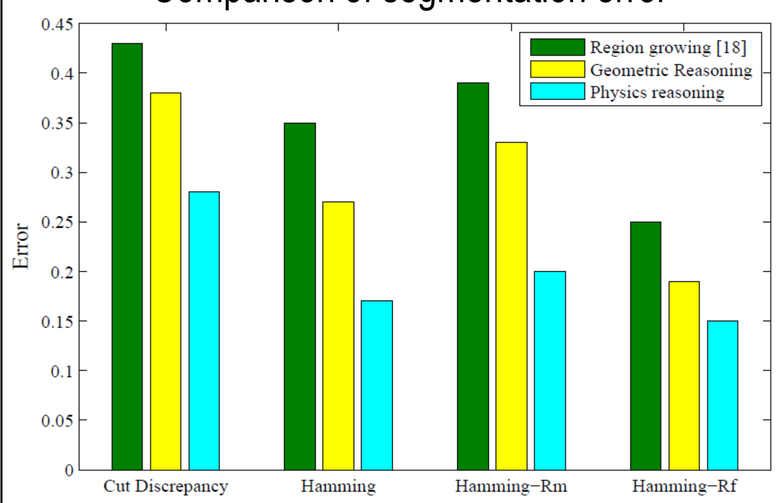
Qualitative results: (a) rgb images (b) results after geometric reasoning (c) results after physical reasoning



Other qualitative results on large scale point clouds captured by KinectFusion



Comparison of segmentation error



Comparison of missing voxel recovery

	Octree [19]	Invisible space	Vol. com.
Precision	98.5%	47.7%	**94.1%**
Recall	7.8%	95.1%	**87.4%**

Comparison of physical relation inference

relations	Discriminative	Greedy	SWC
fixed joint	20.5%	66%	**81.8%**
support	42.2%	60.3%	**78.1%**

Project Page

http://www.stat.ucla.edu/~ybzhao/research/physics/