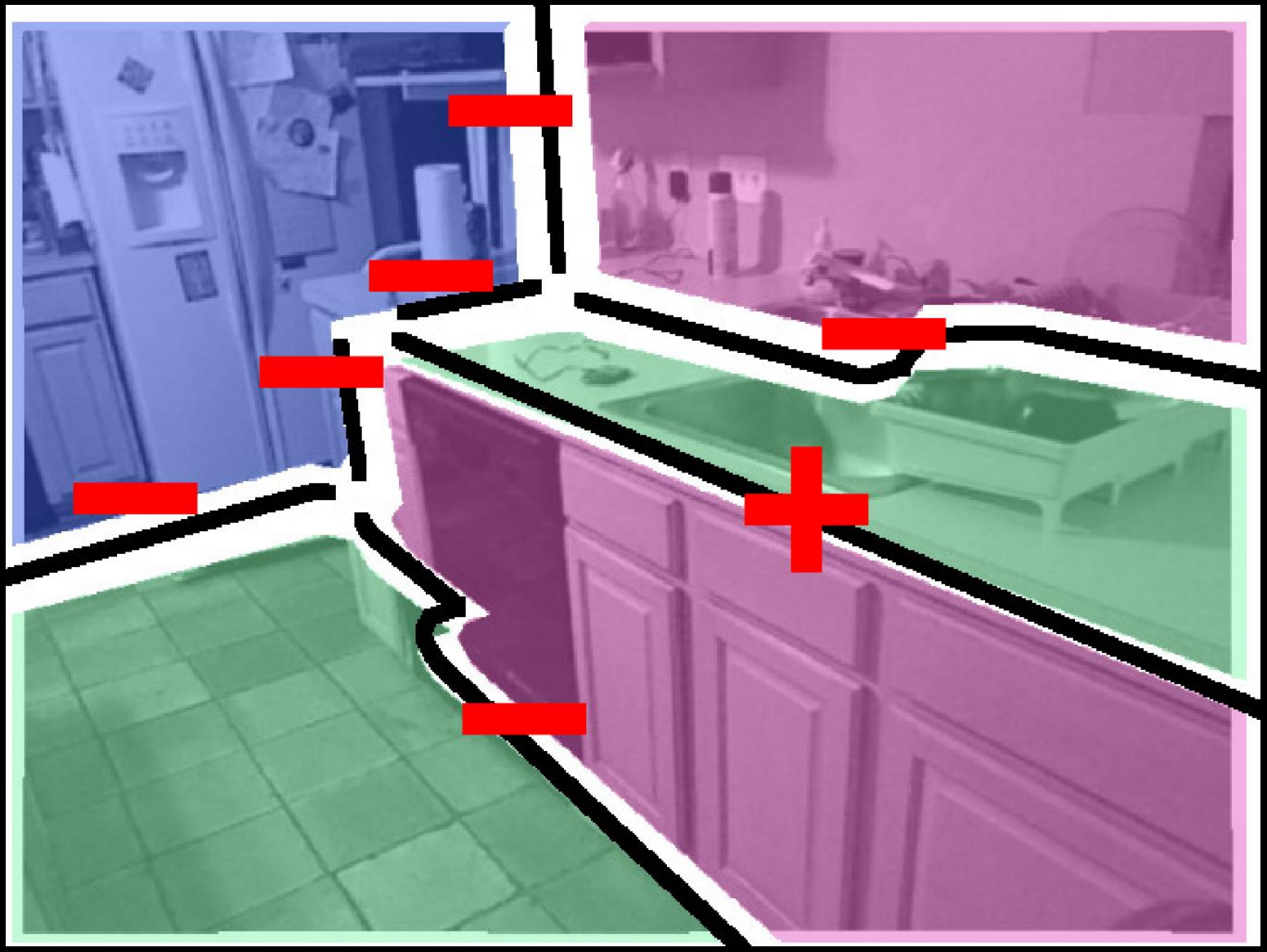
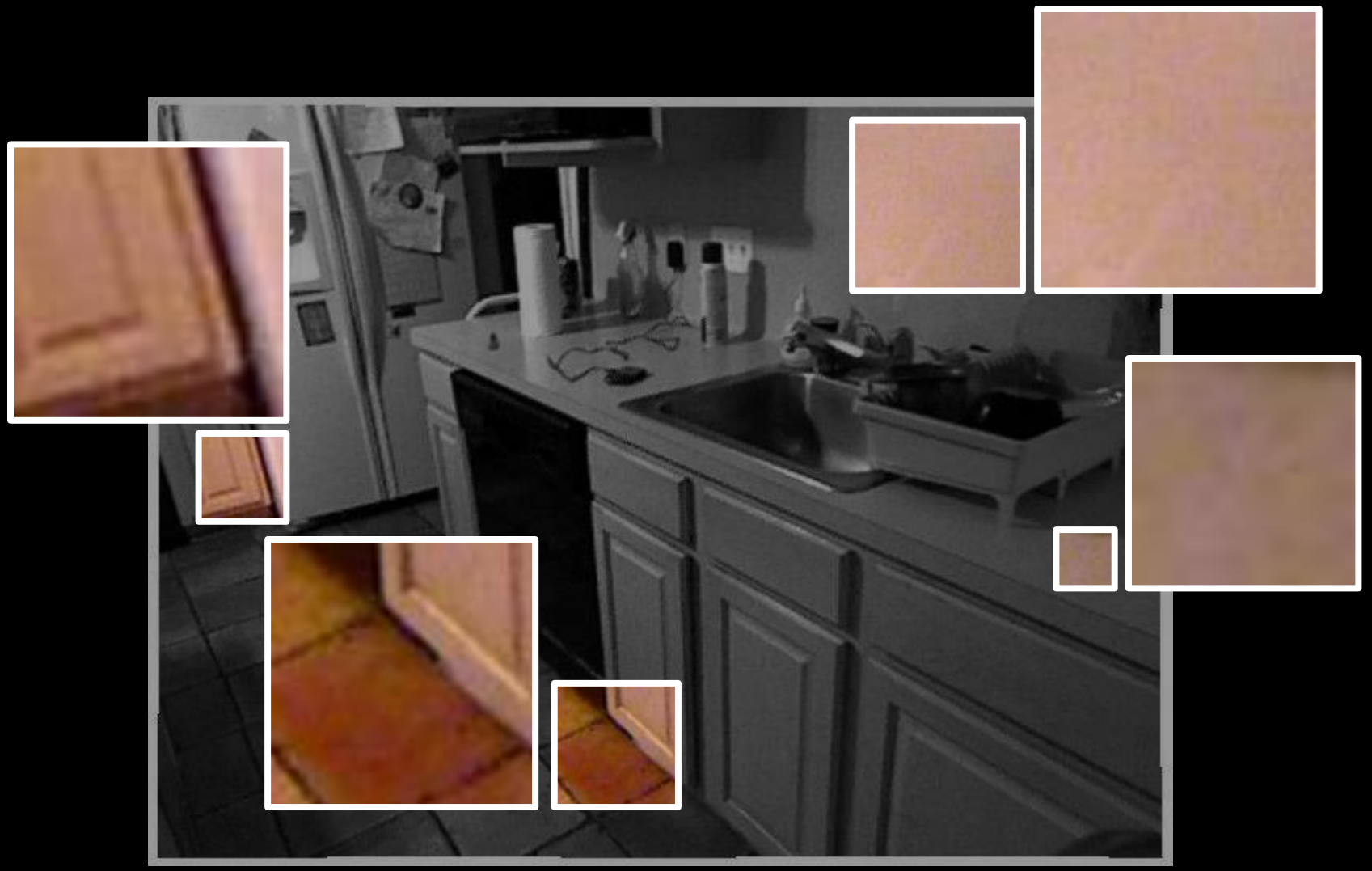
Unfolding an Indoor Origami World

David Fouhey, Abhinav Gupta, Martial Hebert

Carnegie Mellon
THE ROBOTICS INSTITUTE

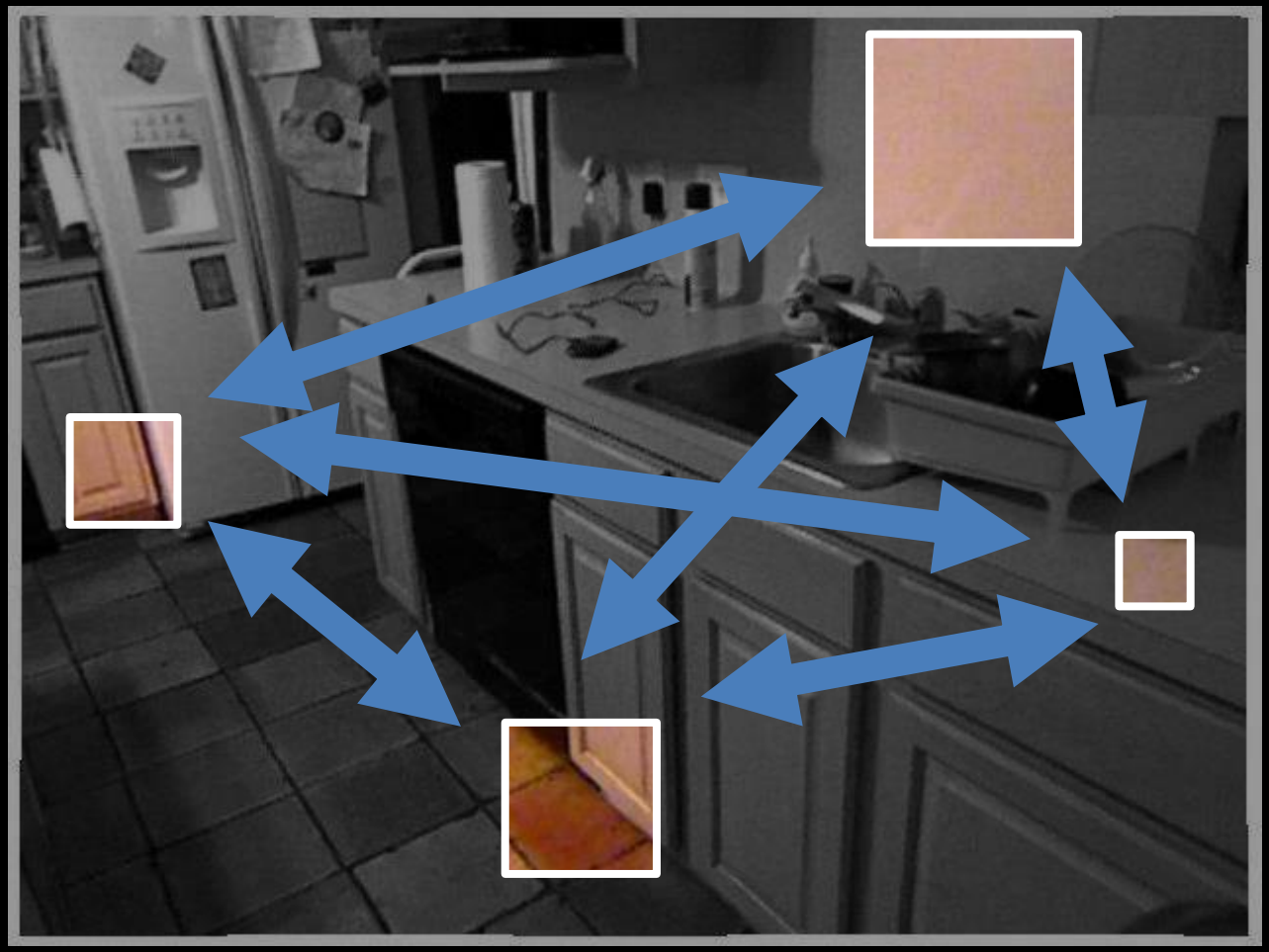


Local Evidence



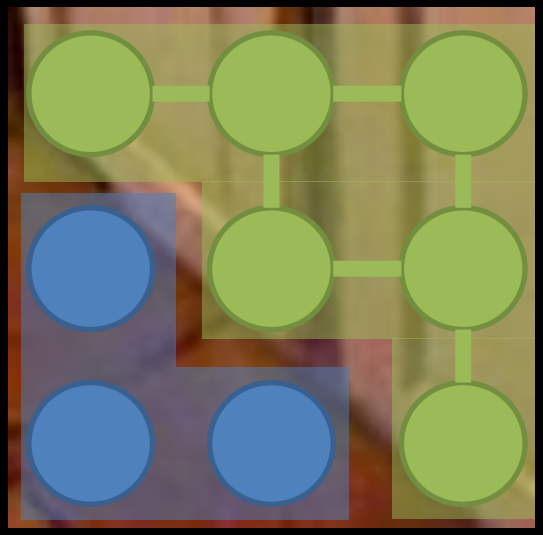
Hoiem et al. 2005, Saxena et al. 2005, Fouhey et al. 2013, etc.

Constraints



Constraints for Single Image 3D

Local Smoothness



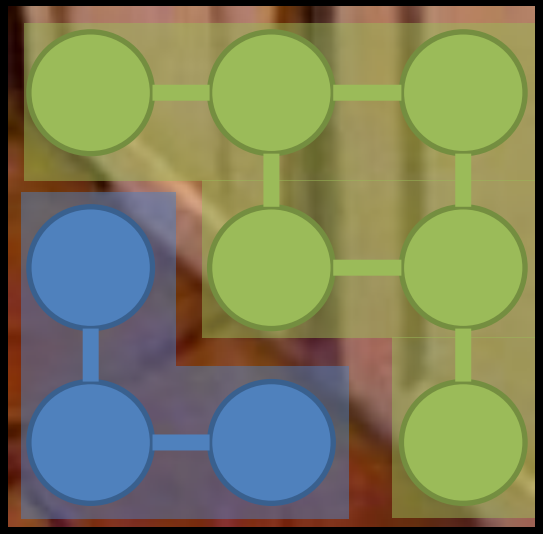
Low Level,
Generic

Hoiem et al. 2005, Saxena et al. 2005, 2008, Munoz et al., 2009, etc.

Constraints for Single Image 3D

Hoiem et al. 2005, Saxena et al. 2005, 2008, Munoz et al., 2009, etc.

Constraints for Single Image 3D



Low Level,
Generic

High Level,
Physical

Constraints for Single Image 3D

Local Smoothness

Low L… Level,
Generic Physical

Coughlan and Yuille 2000, etc.

Constraints for Single Image 3D

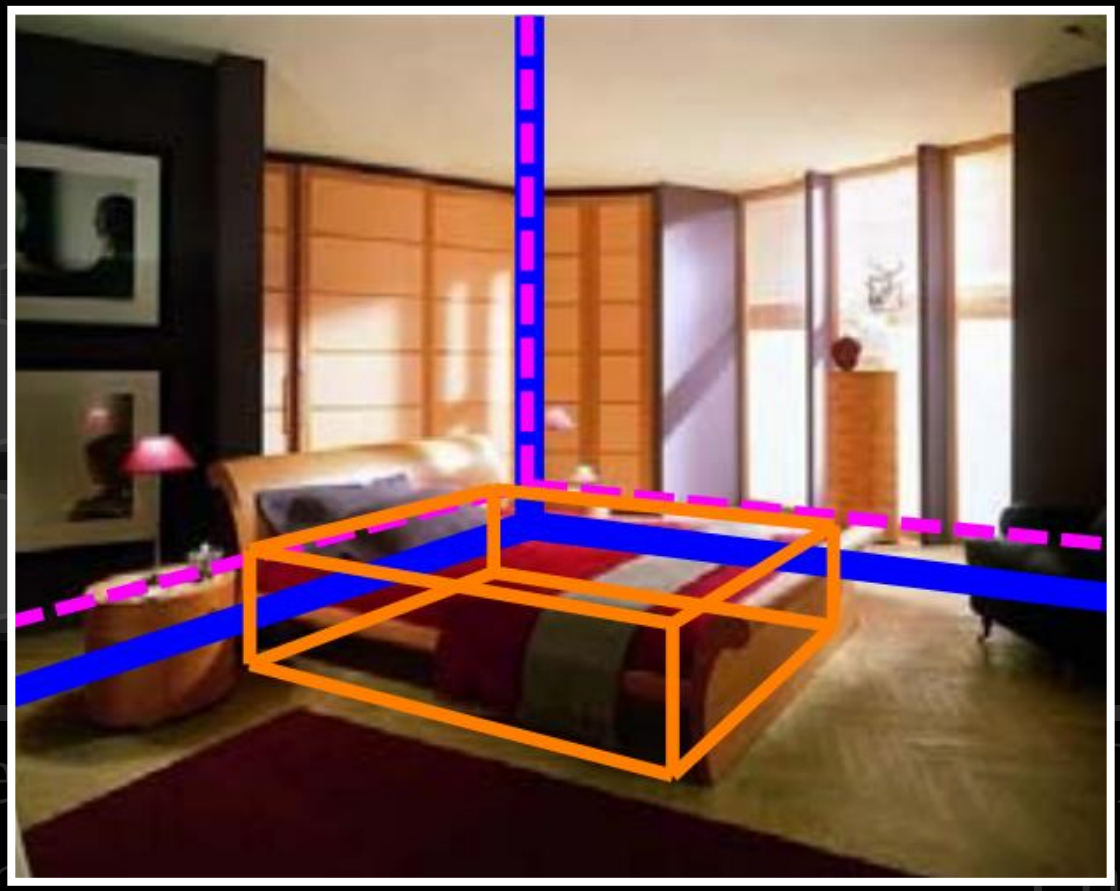
Local Smoothness

Low Lev Level,
Generic Physical

Hedau et al. 2009, Del Pero et al., 2011, Wang et al., 2012, Schwing et al. 2012, etc.

Constraints for Single Image 3D

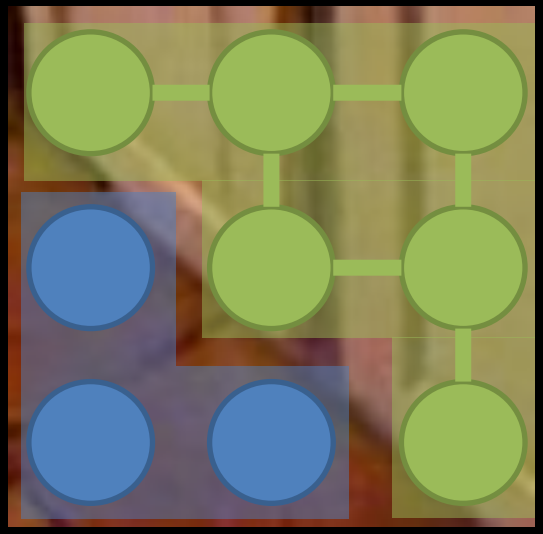
Local Smoothness

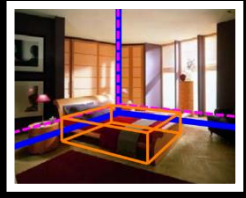


Low Le… Level, Physical
Generic

Lee et al. 2010, Xiao et al. 2012, Zhao et al. 2013, Schwing et al., 2013, etc.

Constraints for Single Image 3D

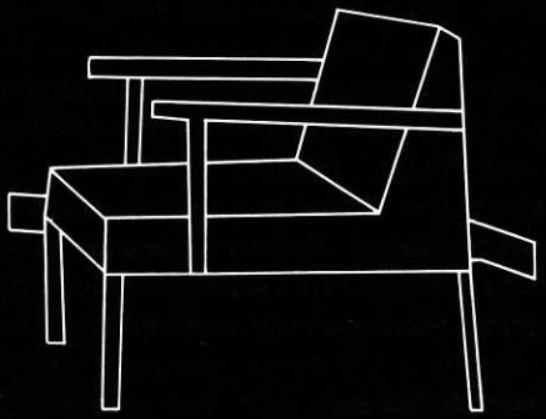


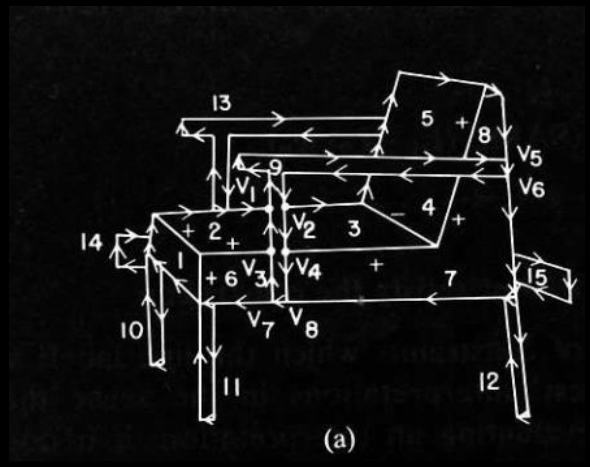


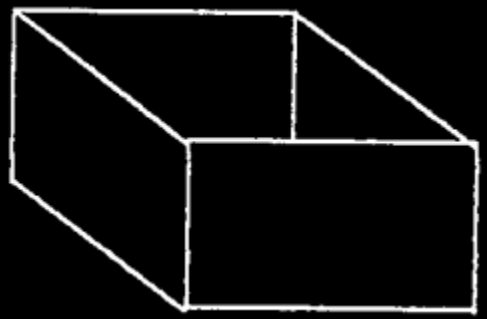
Low Level,
Generic

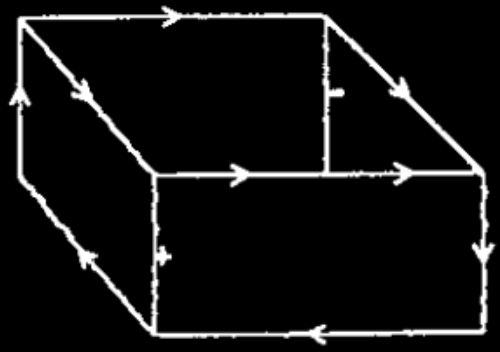
High Level,
Physical

Mid-level in the Past



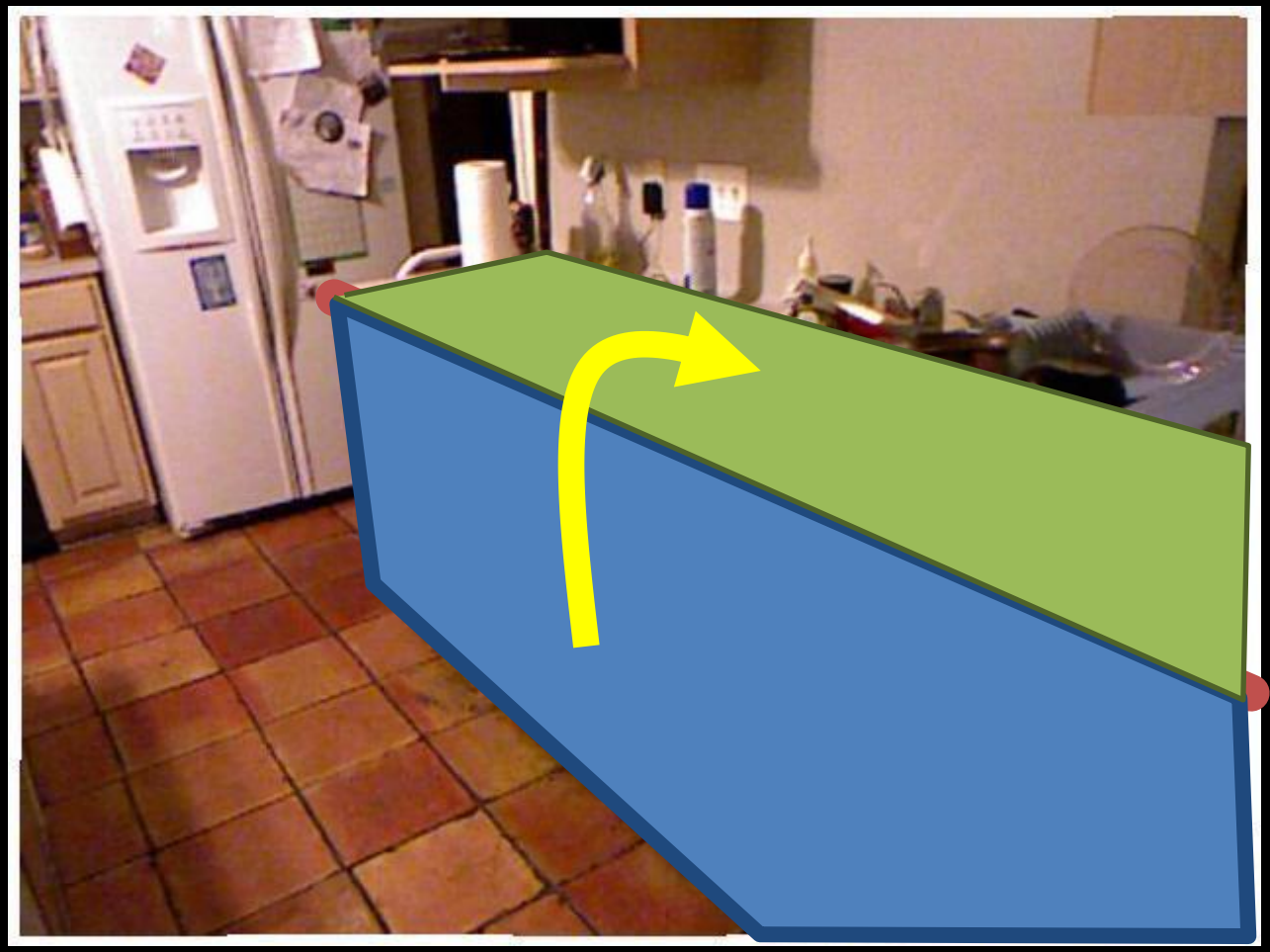






Huffman 71, Clowes 71, Kanade 80, 81 Sugihara 86, Malik 87, etc.

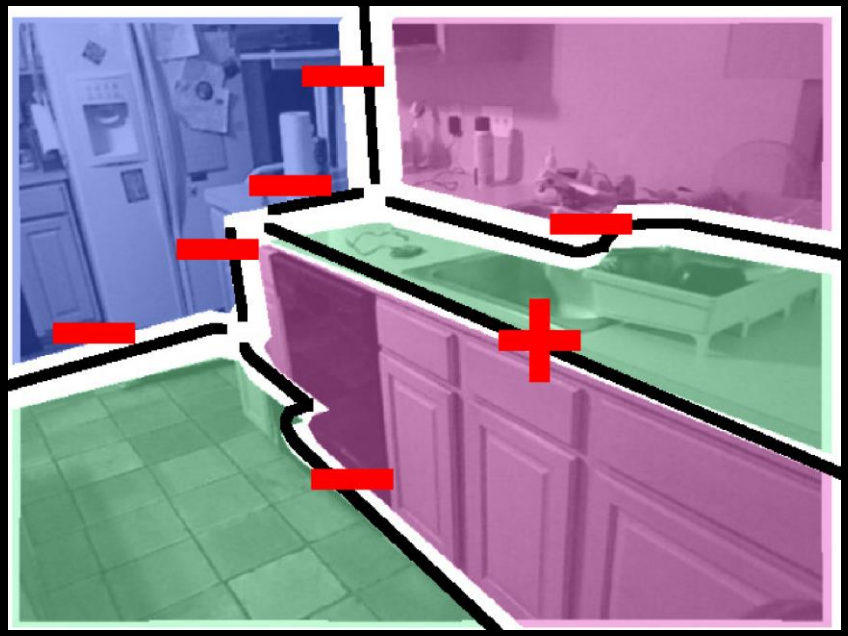
Our Mid-Level Constraints



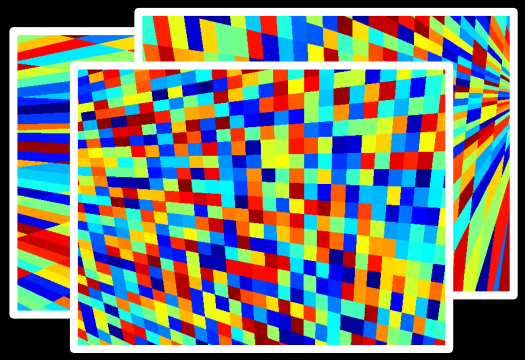
This Work

Input:
Single Image

Output:
Discrete Scene Parse



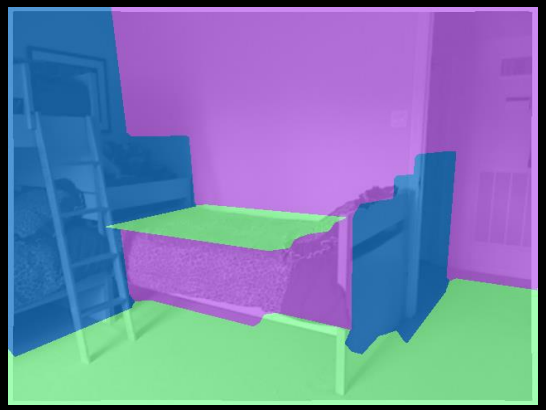
Overview



Parameterization

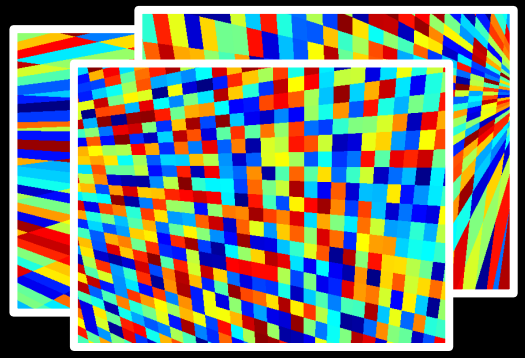
$$\arg\max_{\mathbf{x}\in\{0,1\}^n} \mathbf{c}^T\mathbf{x} + \mathbf{x}^T\mathbf{H}\mathbf{x}$$
$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{1}$$

Formulation



Experimental Results

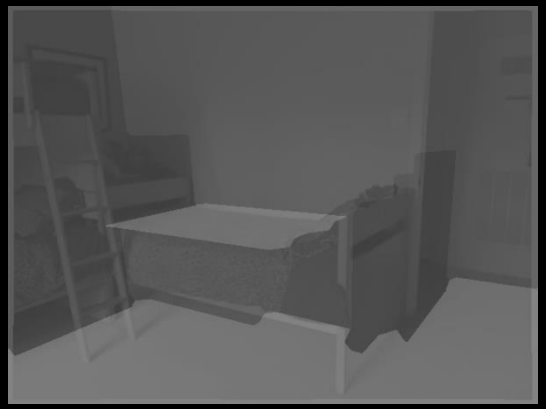
Overview



Parameterization

$$\arg\max_{\mathbf{x}\in\{0,1\}^n} \mathbf{c}^T\mathbf{x} + \mathbf{x}^T\mathbf{H}\mathbf{x}$$
$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{1}$$

Formulation



Experimental Results

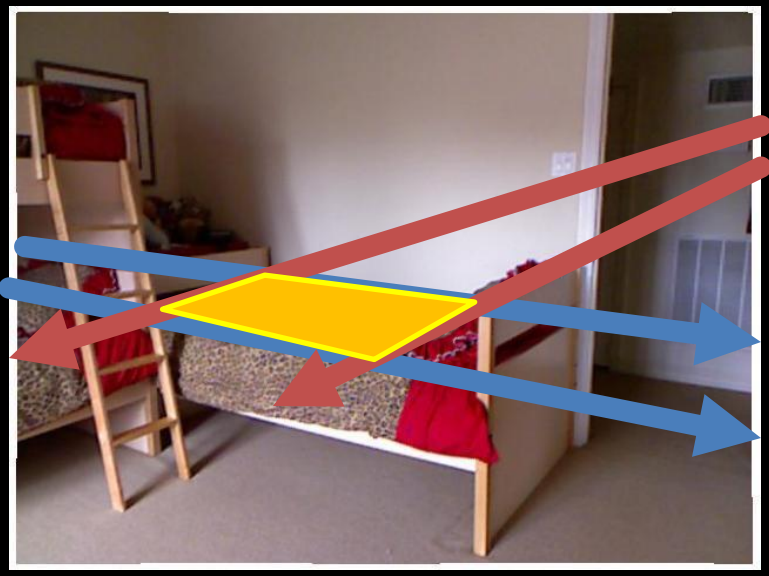
Parameterization

Parameterization

vp_2

vp_3

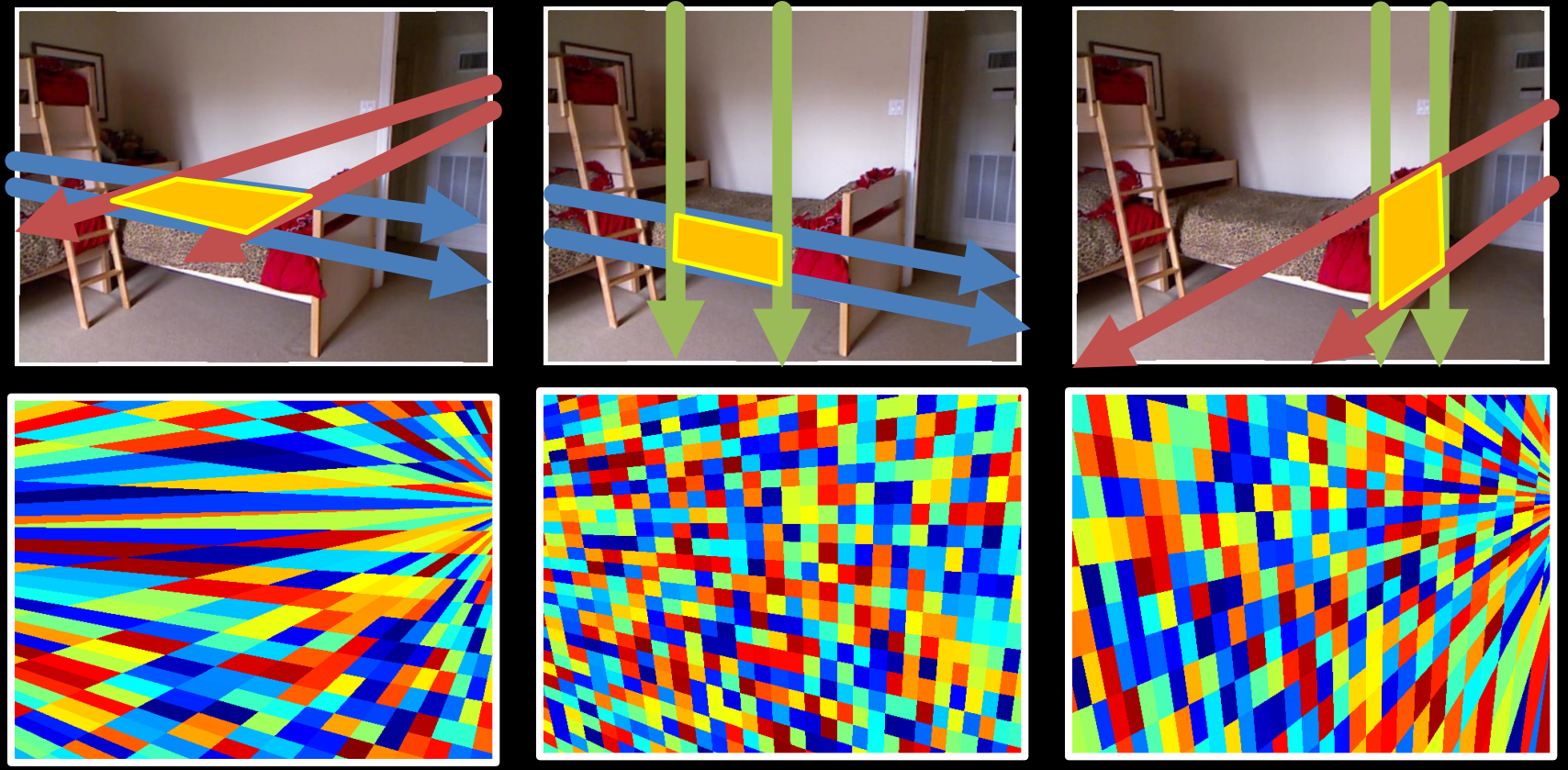
vp_1



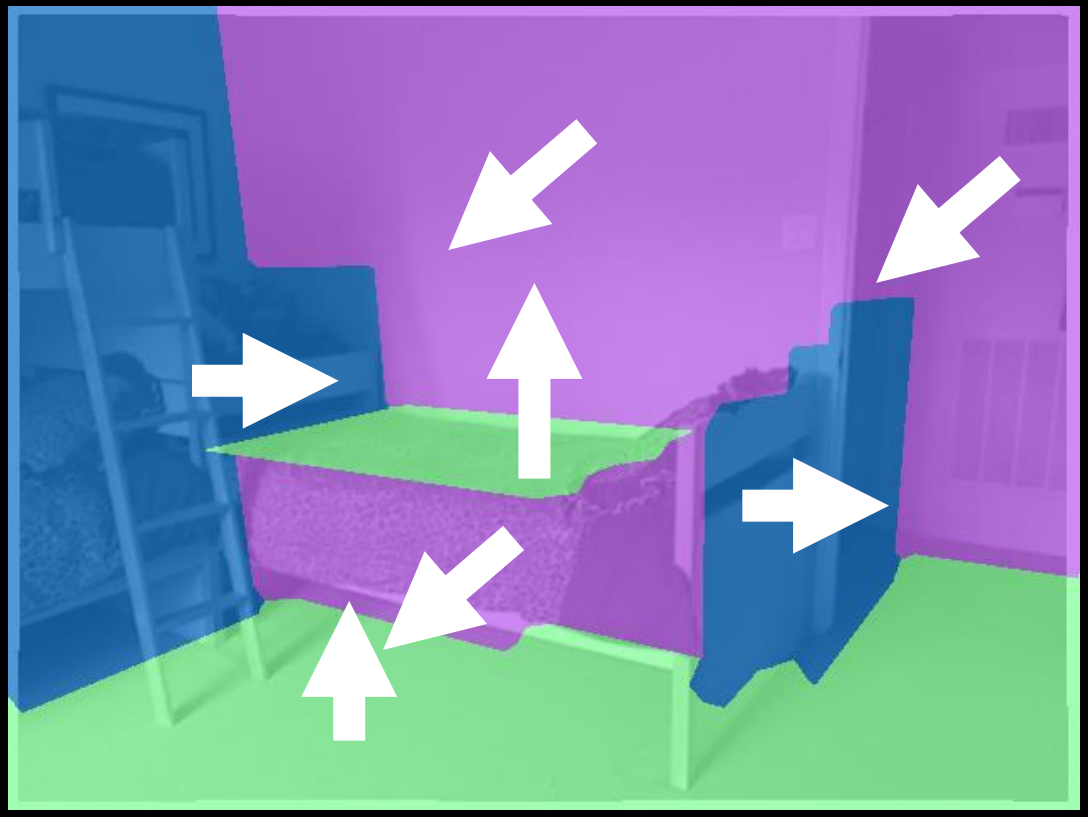
VP Estimator from Hedau et al., 2009

Parameterization

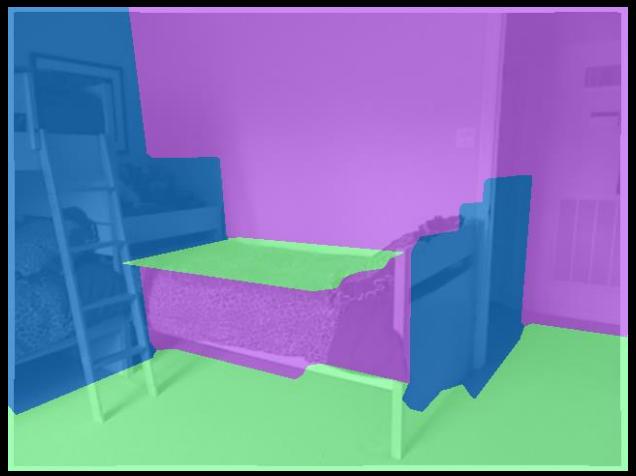
Two VPs give grid cell

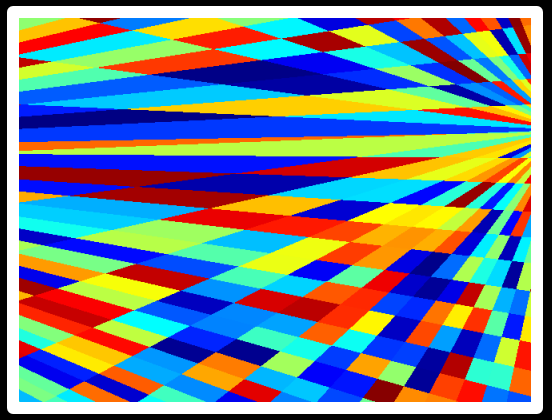


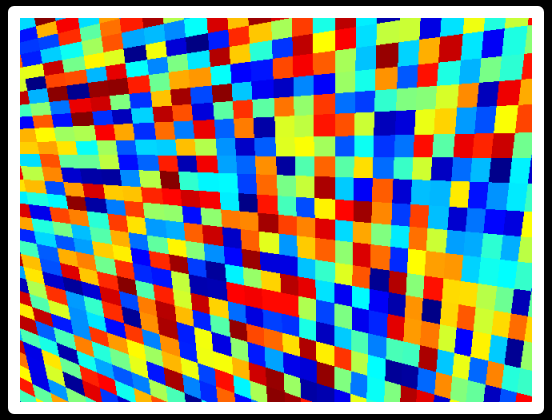
Encoding Surface Normals

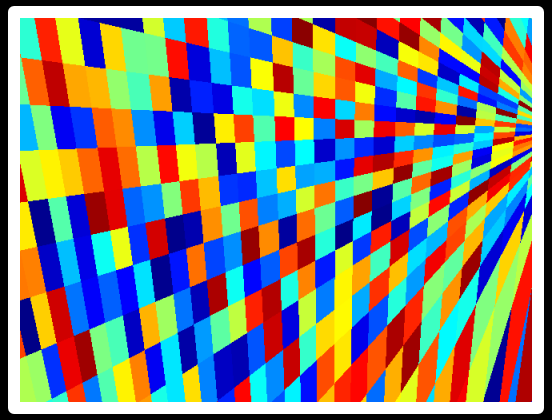


Encoding Surface Normals

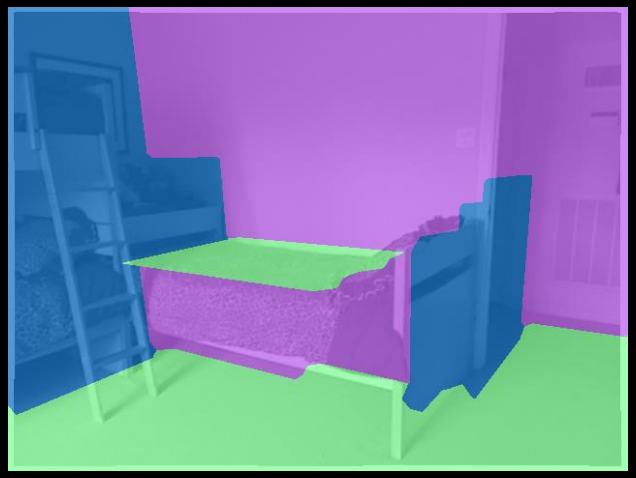


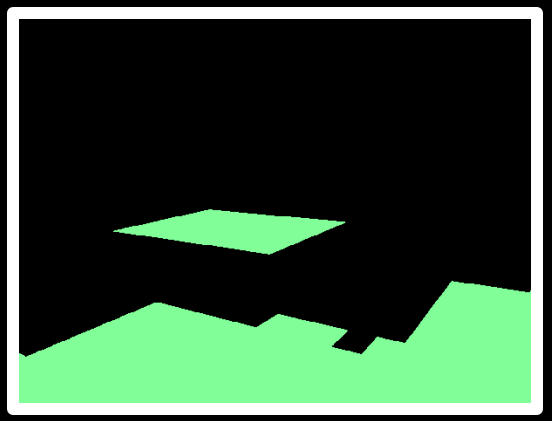


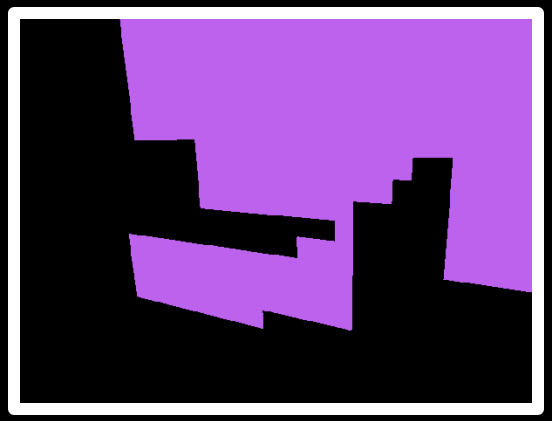


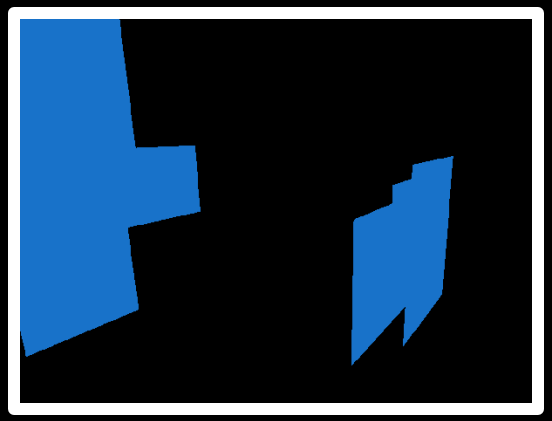


Encoding Surface Normals

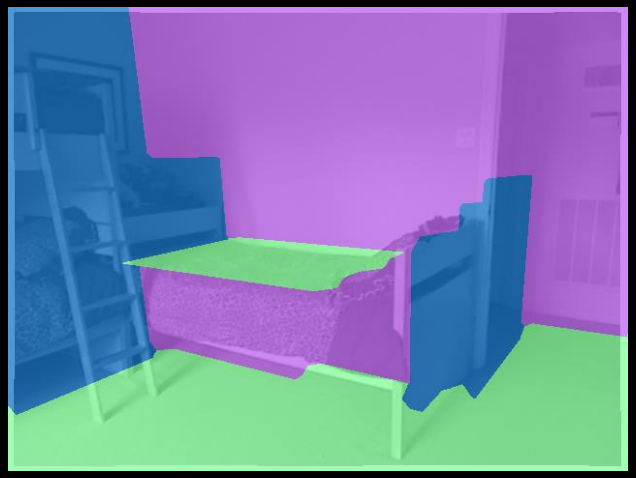


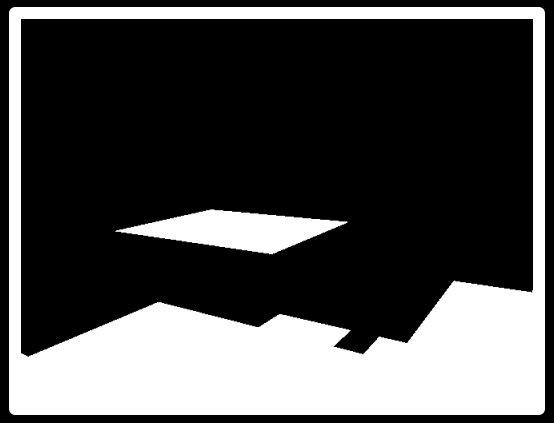




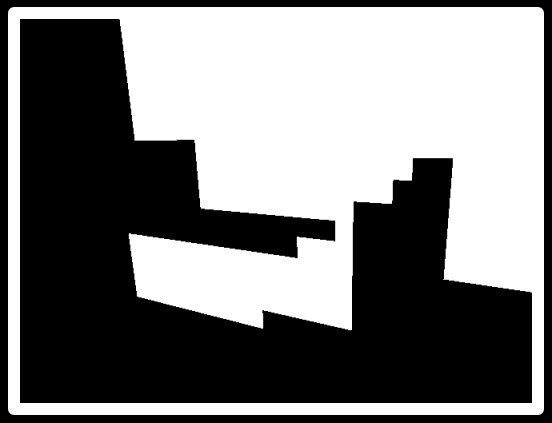


Encoding Surface Normals

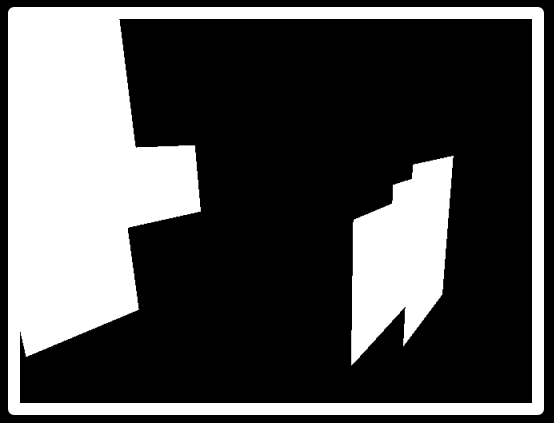




$x_1, \ldots, x_{400}$

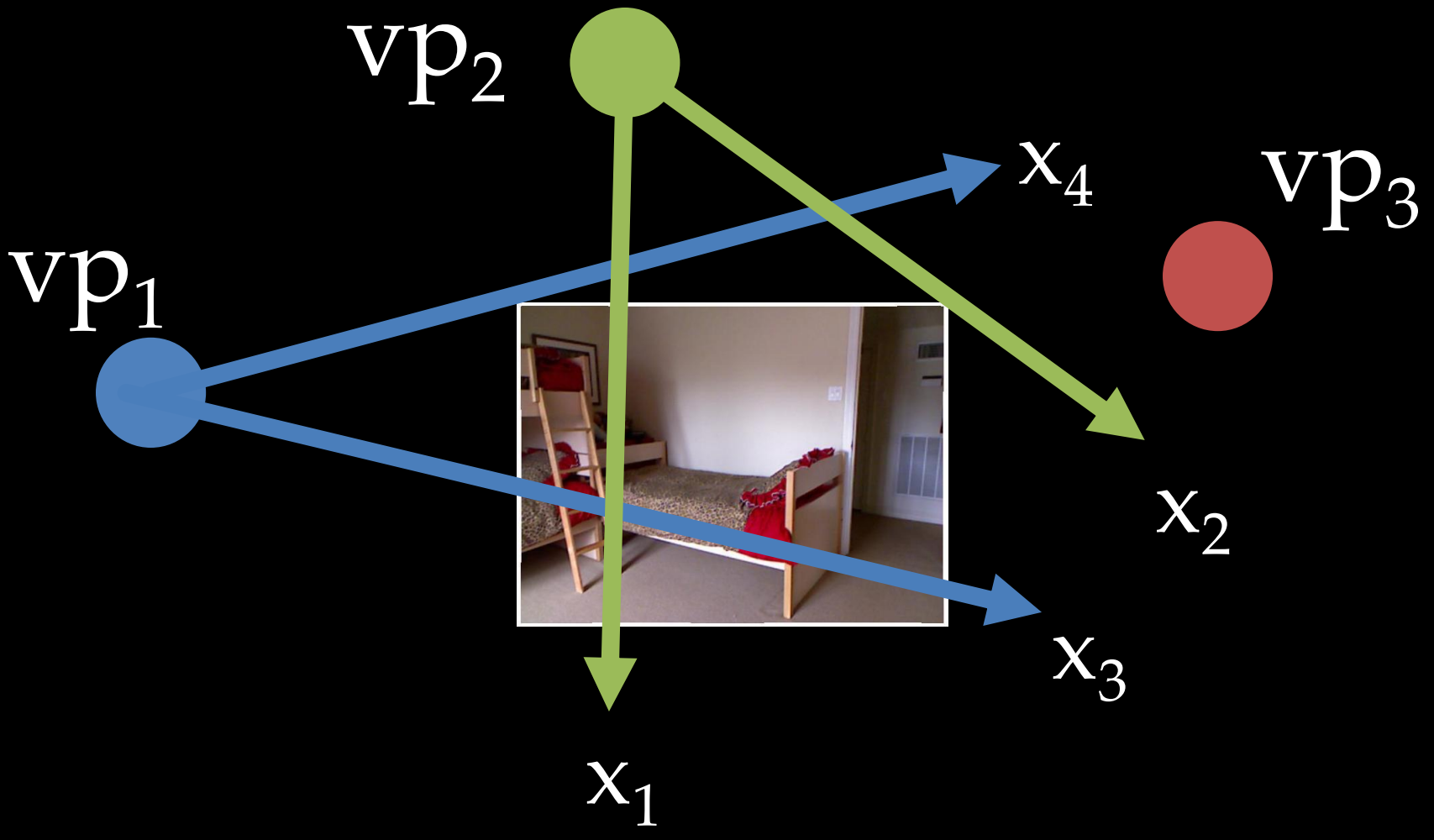


$x_{401}, \ldots, x_{800}$



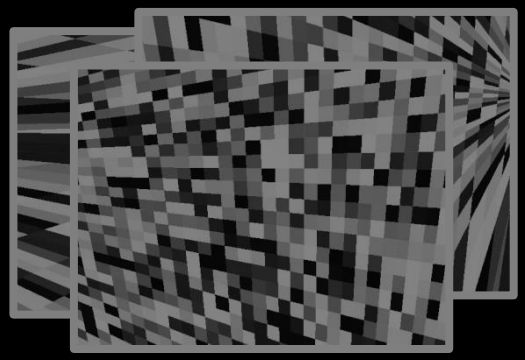
$x_{801}, \ldots, x_{1200}$

Related Parameterizations



Hedau et al., 2009; Wang et al. 2010, Schwing et al., 2012, 2013

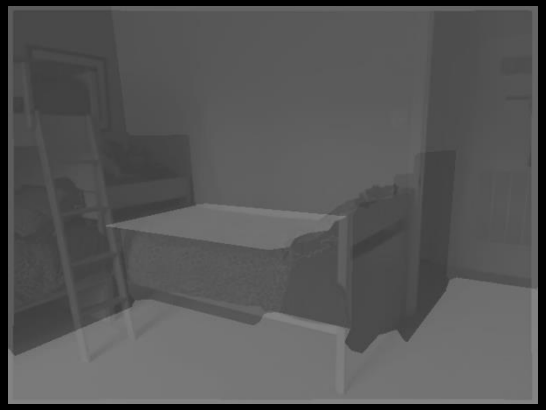
Overview



Parameterization

$$\arg\max_{\mathbf{x}\in\{0,1\}^n} \mathbf{c}^T\mathbf{x} + \mathbf{x}^T\mathbf{H}\mathbf{x}$$
$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{1}$$

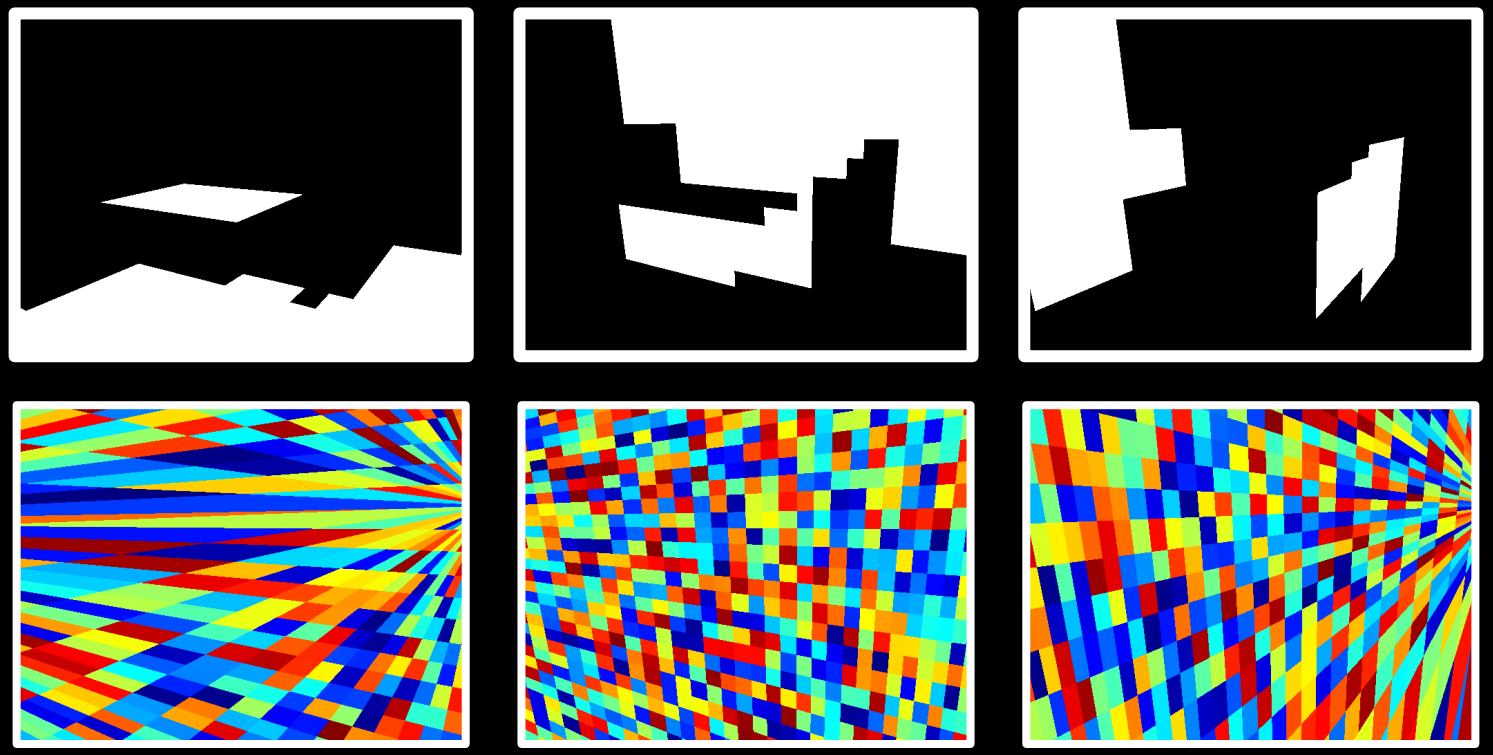
Formulation



Experimental Results

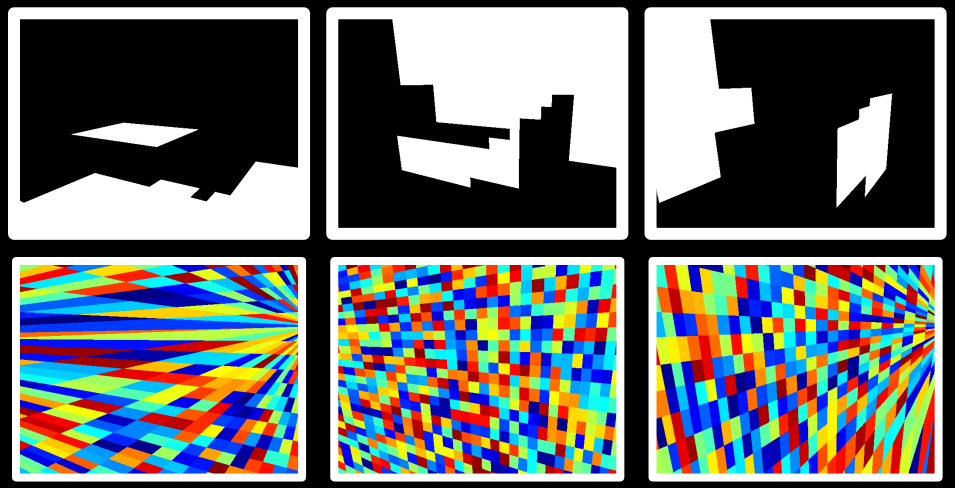
Parameterization

$$\mathbf{x} \in \{0,1\}^n$$



Formulation

$$\arg\max_{\mathbf{x}\in\{0,1\}^n} \mathbf{c}^T\mathbf{x} + \mathbf{x}^T\mathbf{H}\mathbf{x} \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{1}$$

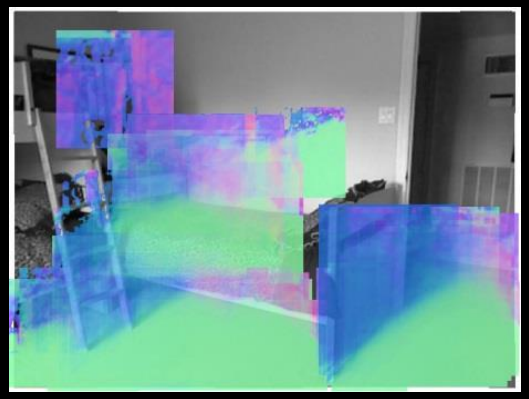


Unaries

$$\arg\max_{\mathbf{x}\in\{0,1\}^n} \mathbf{c}^T\mathbf{x} + \mathbf{x}^T\mathbf{H}\mathbf{x} \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{1}$$

Unaries

$$\arg\max_{\mathbf{x}\in\{0,1\}^n} \mathbf{c}^T\mathbf{x} + \mathbf{x}^T\mathbf{H}\mathbf{x} \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} \leq 1$$



Any
3D Evidence

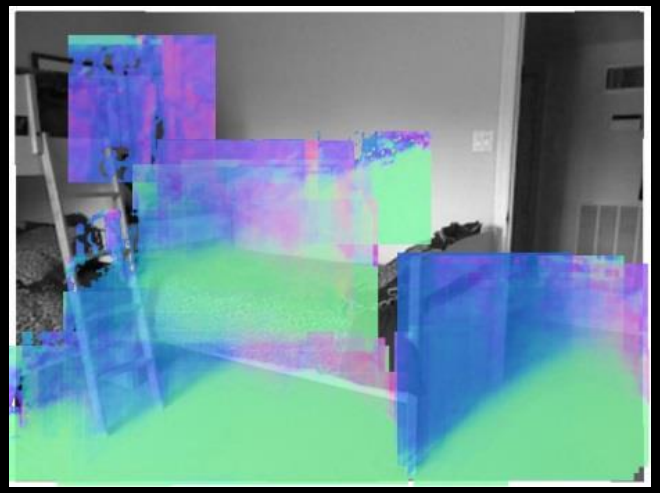
High c

Low c

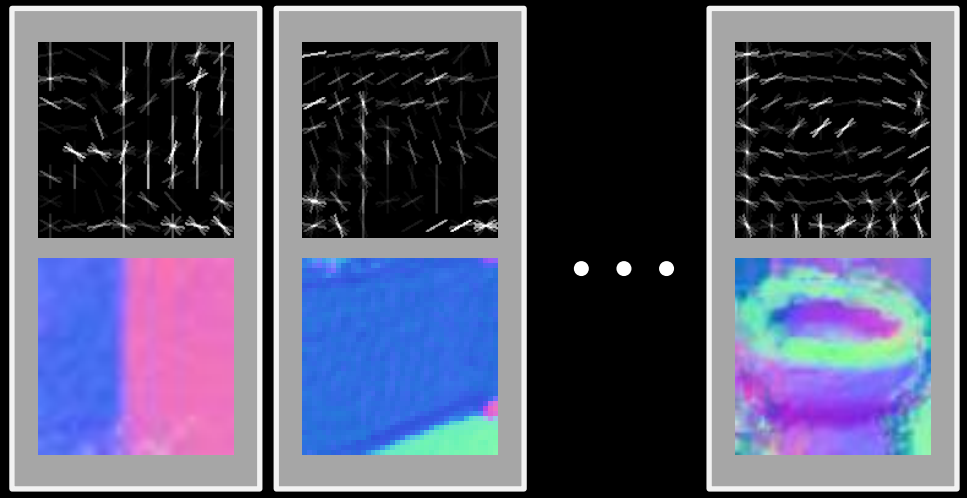
Unaries

Local: Data-Driven 3D Primitives

Input



3D Primitive Bank



Fouhey, Gupta, Hebert, 2013

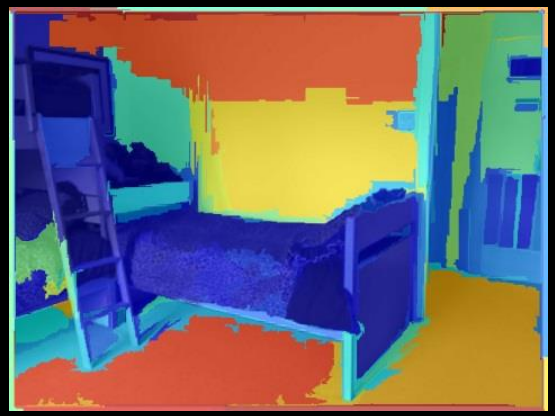
Unaries

Global: Cuboid Fit + Clutter Mask

Input

Predicted Walls

Clutter Mask

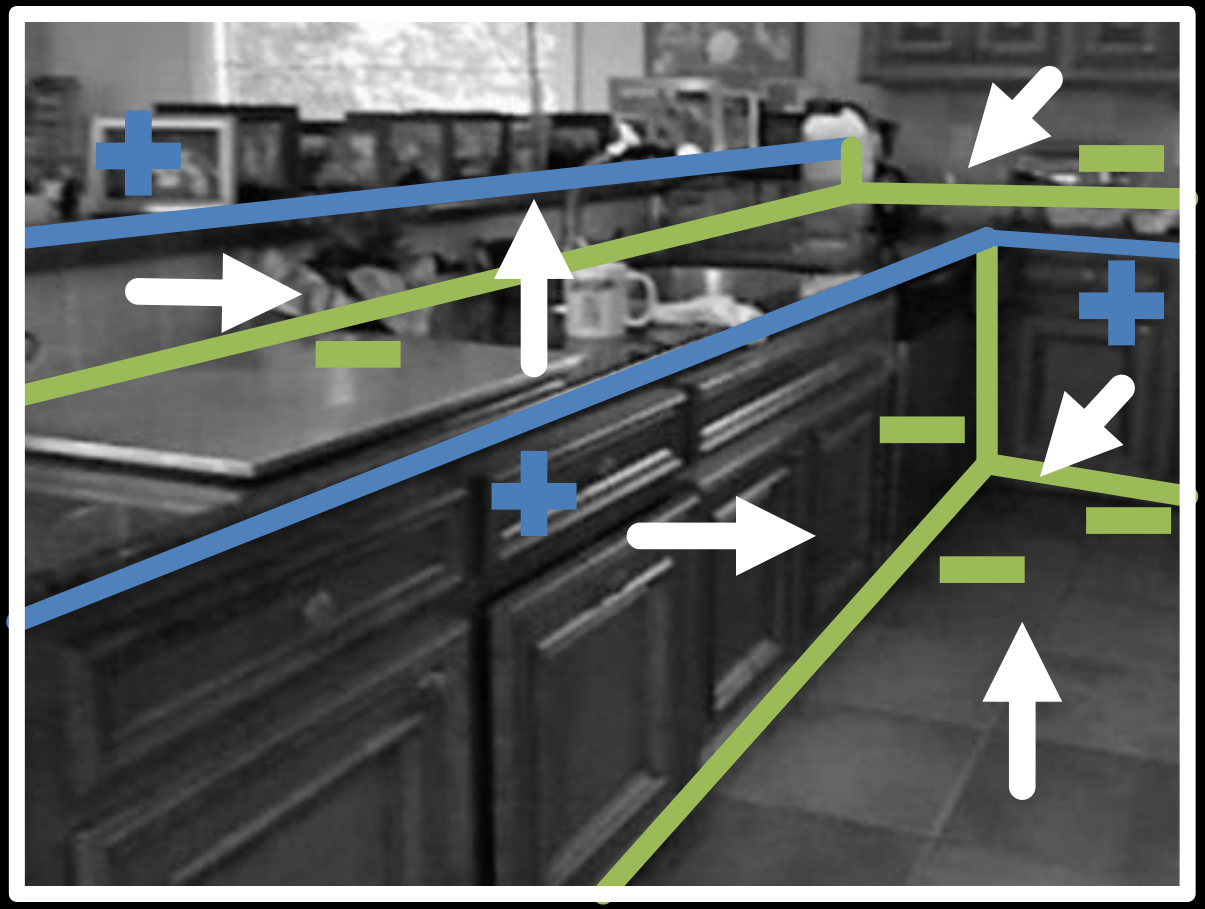


Hedau, Hoiem, Forsyth, 2009

Binaries

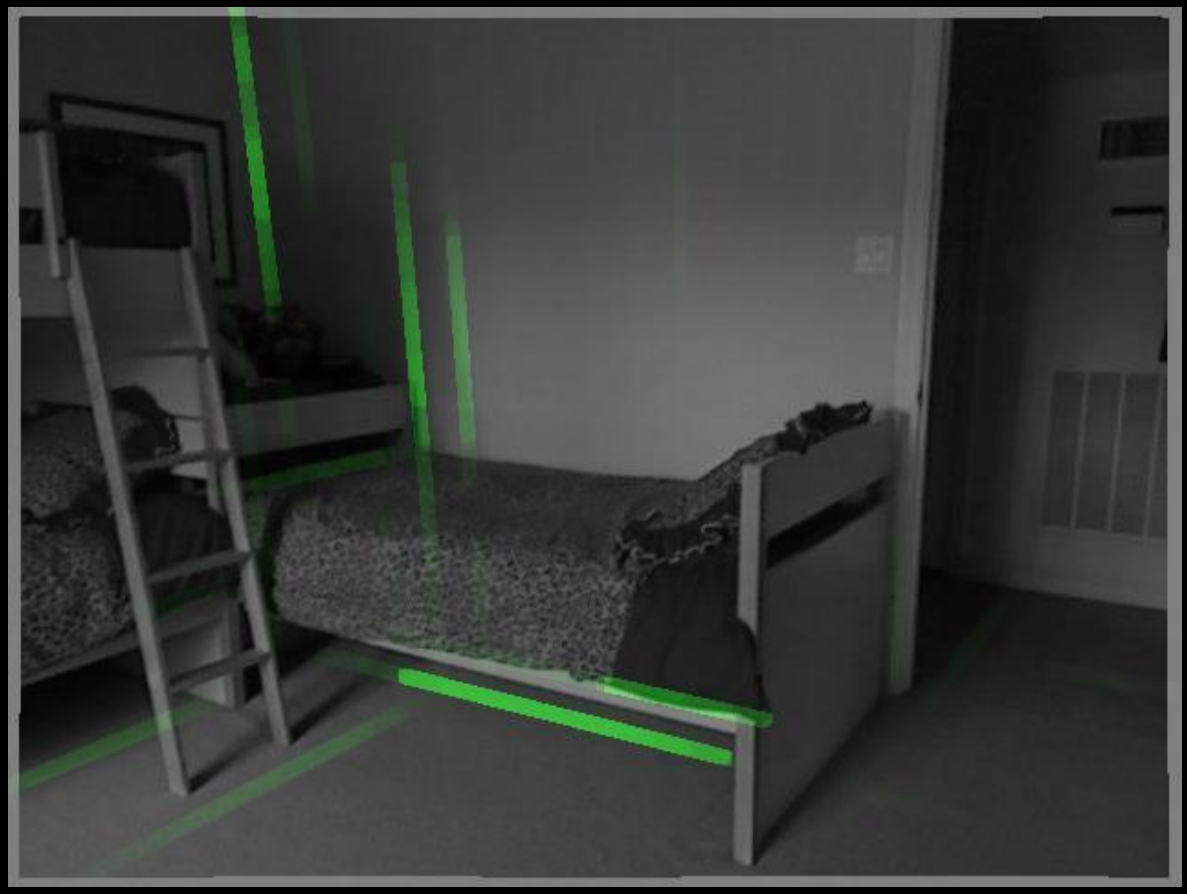
$$\arg\max_{\mathbf{x}\in\{0,1\}^n} \mathbf{c}^T\mathbf{x} + \mathbf{x}^T\mathbf{H}\mathbf{x} \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{1}$$

Convex/Concave Constraints



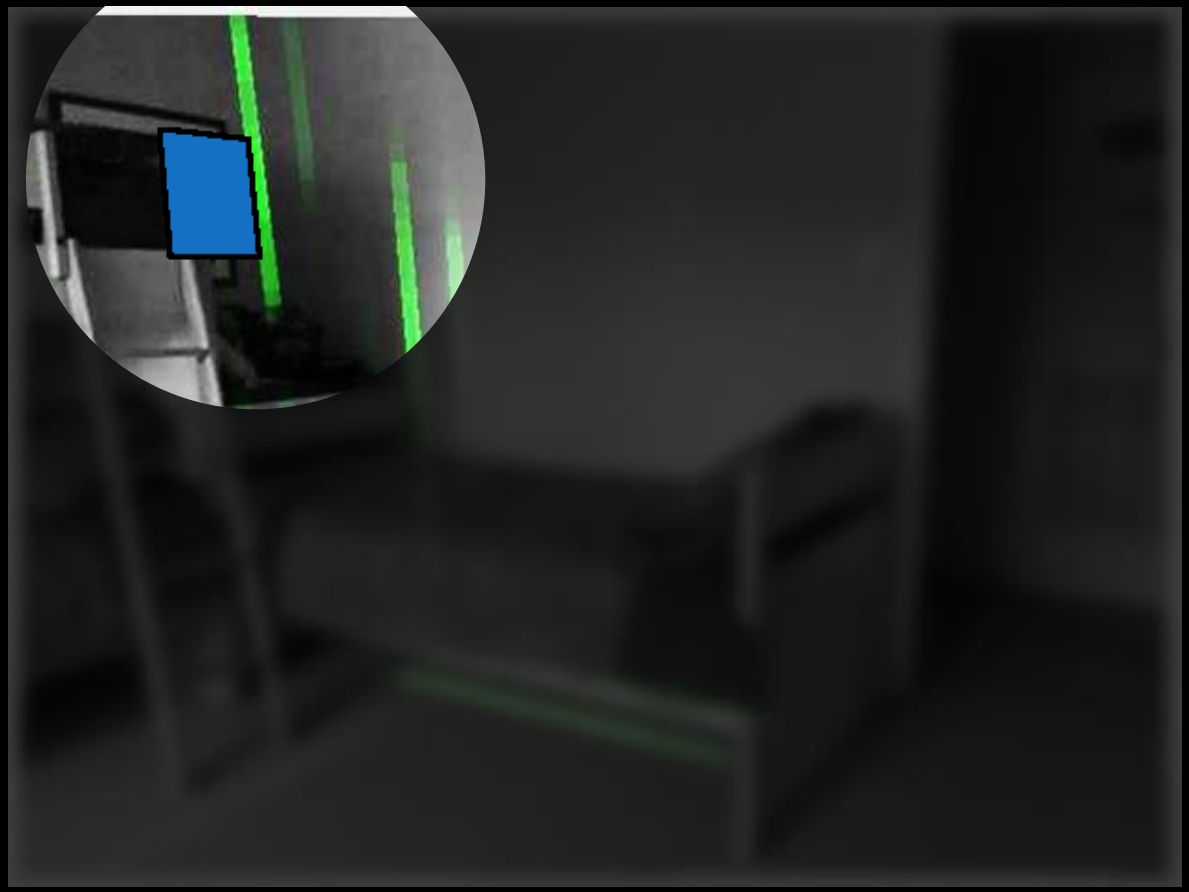
Convex (+) Concave (-)

Convex/Concave Constraints



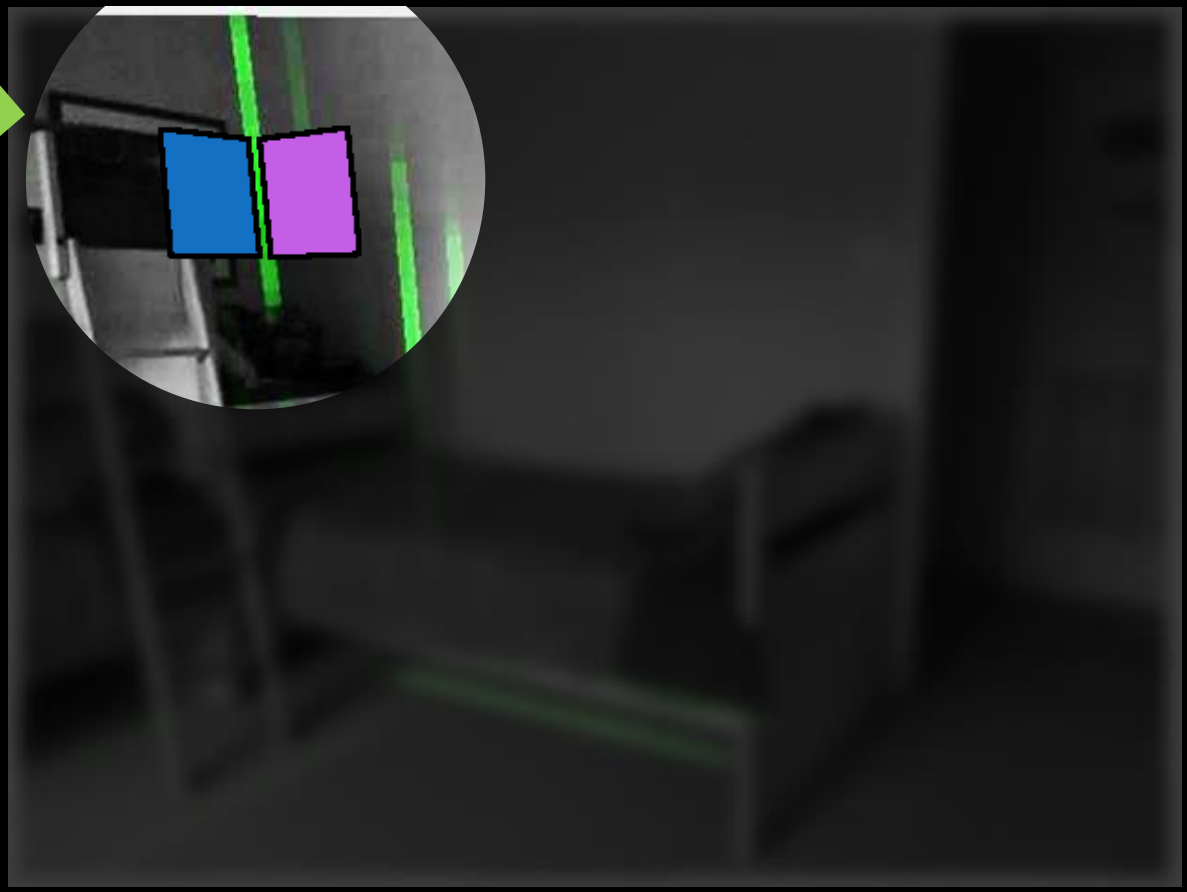
Detected Concave (-)

Convex/Concave Constraints



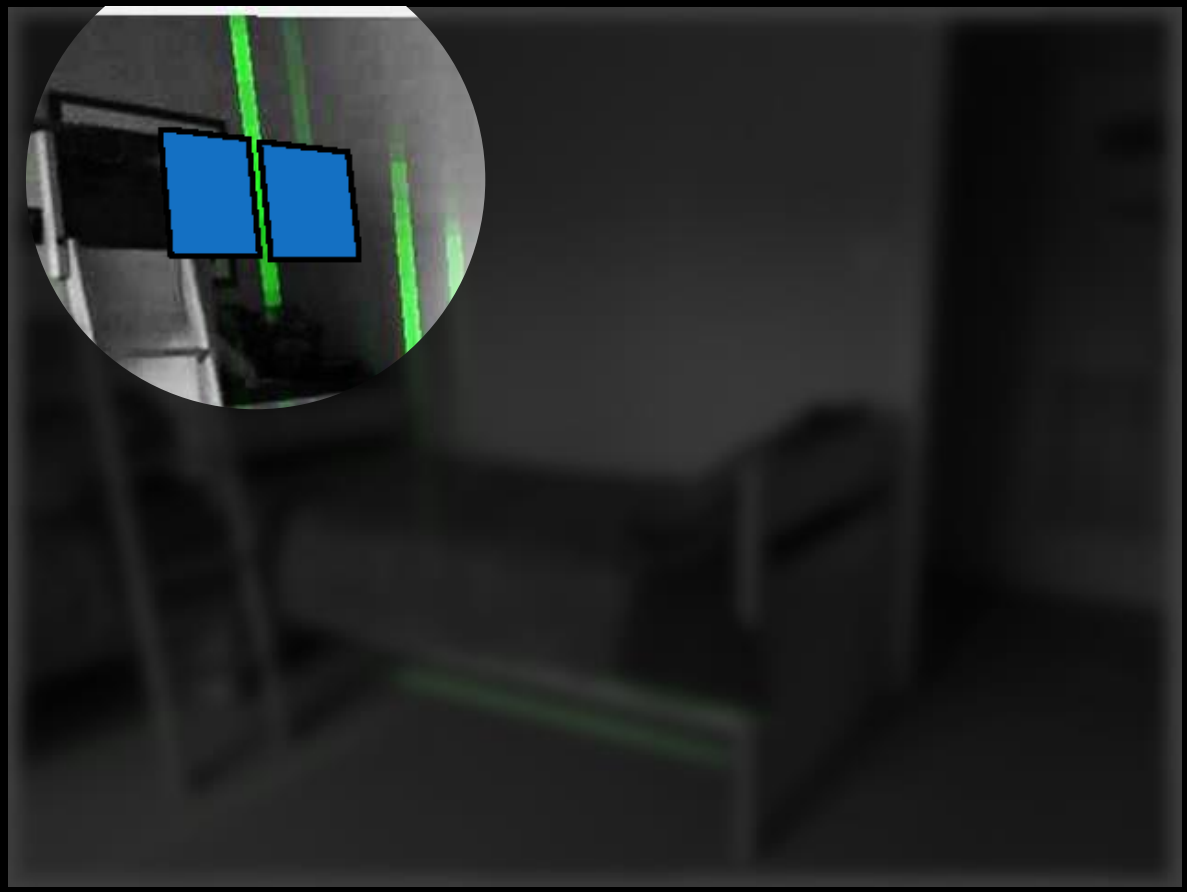
Detected Concave (-)

Convex/Concave Constraints



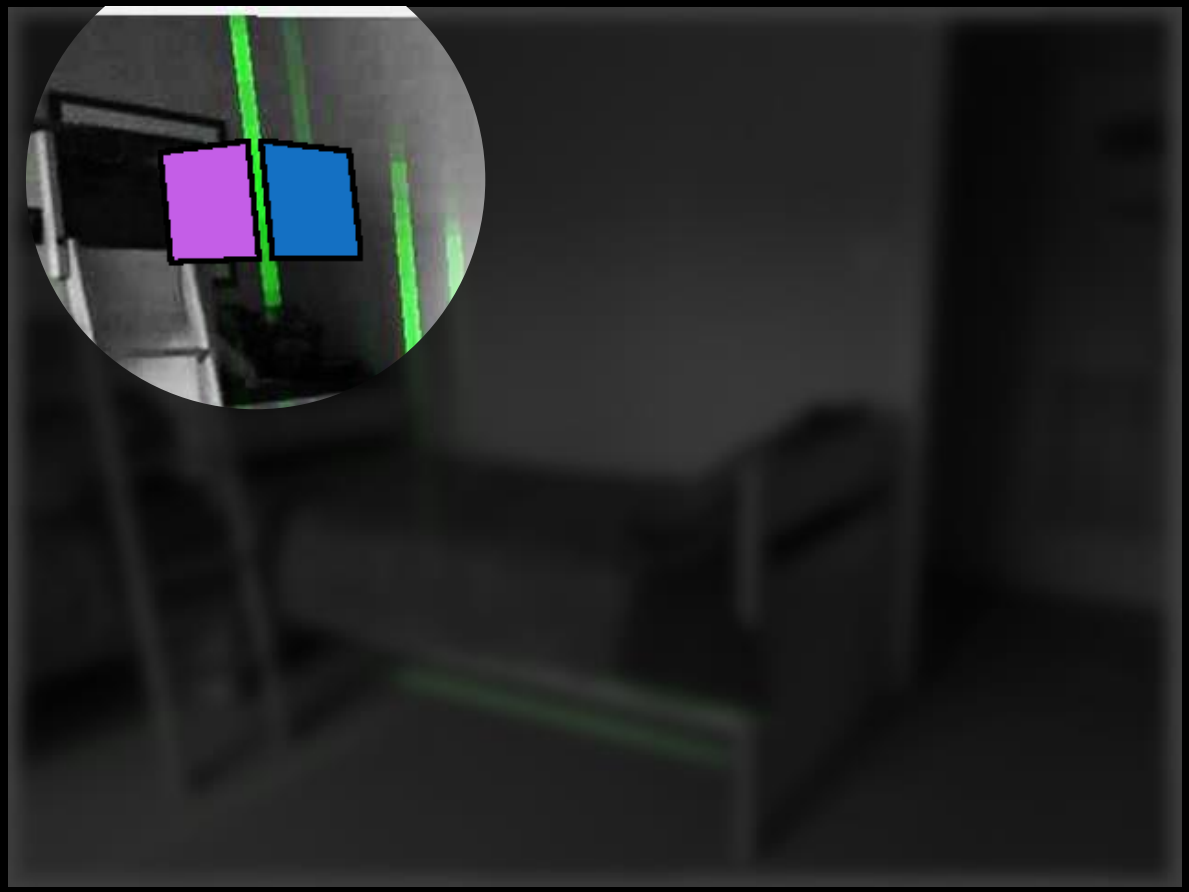
Detected Concave (-)

Convex/Concave Constraints



Detected Concave (-)

Convex/Concave Constraints

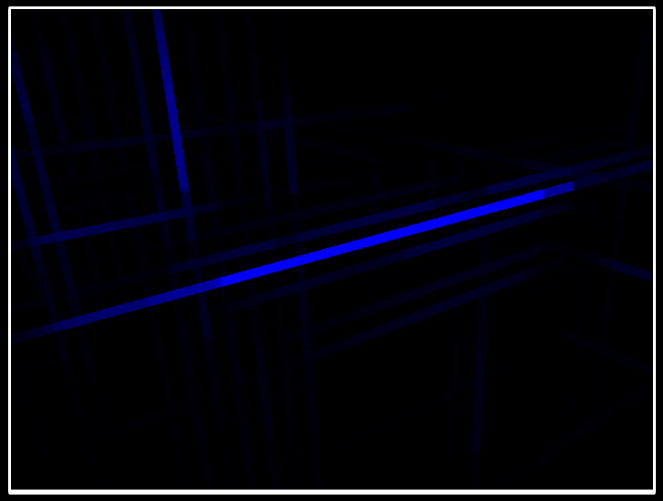


Detected Concave (-)

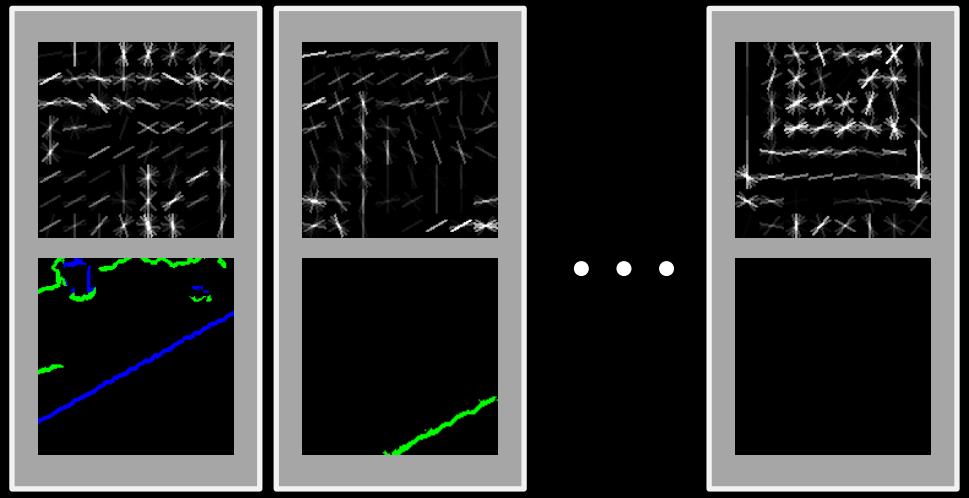
Detecting Convex/Concave

Use 3DP to Transfer Discontinuities

Input

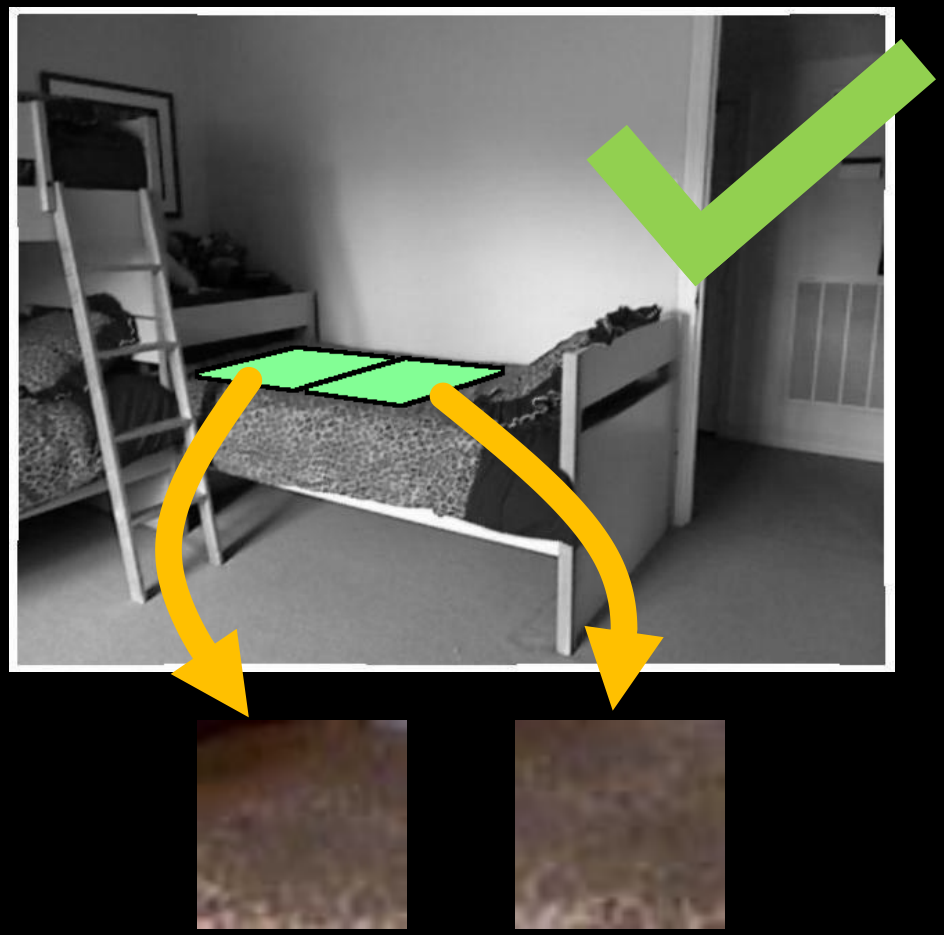


3D Primitive Bank



Ground-Truth Discontinuities similar to Gupta, Arbelaez, Malik, 2013
3DP from Fouhey, Gupta, Hebert, 2013

Smoothness



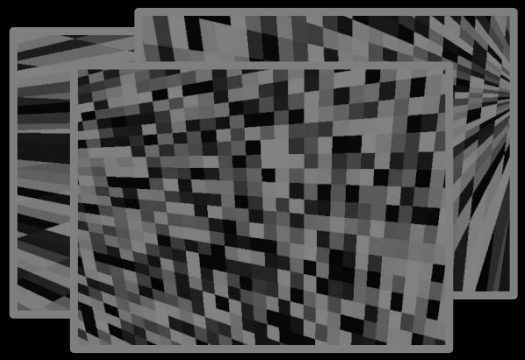
Constraints

$$\arg\max_{\mathbf{x}\in\{0,1\}^n} \mathbf{c}^T\mathbf{x} + \mathbf{x}^T\mathbf{H}\mathbf{x} \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{1}$$

Solving the Model

$$\underset{\mathbf{x}\in\{0,1\}^n}{\arg\max}\, \mathbf{c}^T\mathbf{x} + \mathbf{x}^T\mathbf{H}\mathbf{x} \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{1}$$

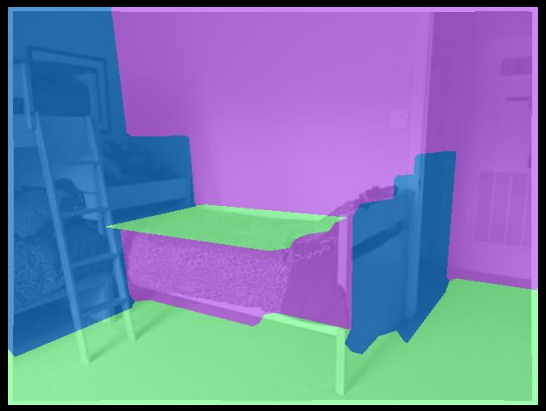
Overview



Parameterization

$$\arg\max_{\mathbf{x}\in\{0,1\}^n} \mathbf{c}^T\mathbf{x} + \mathbf{x}^T\mathbf{H}\mathbf{x}$$
$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{1}$$

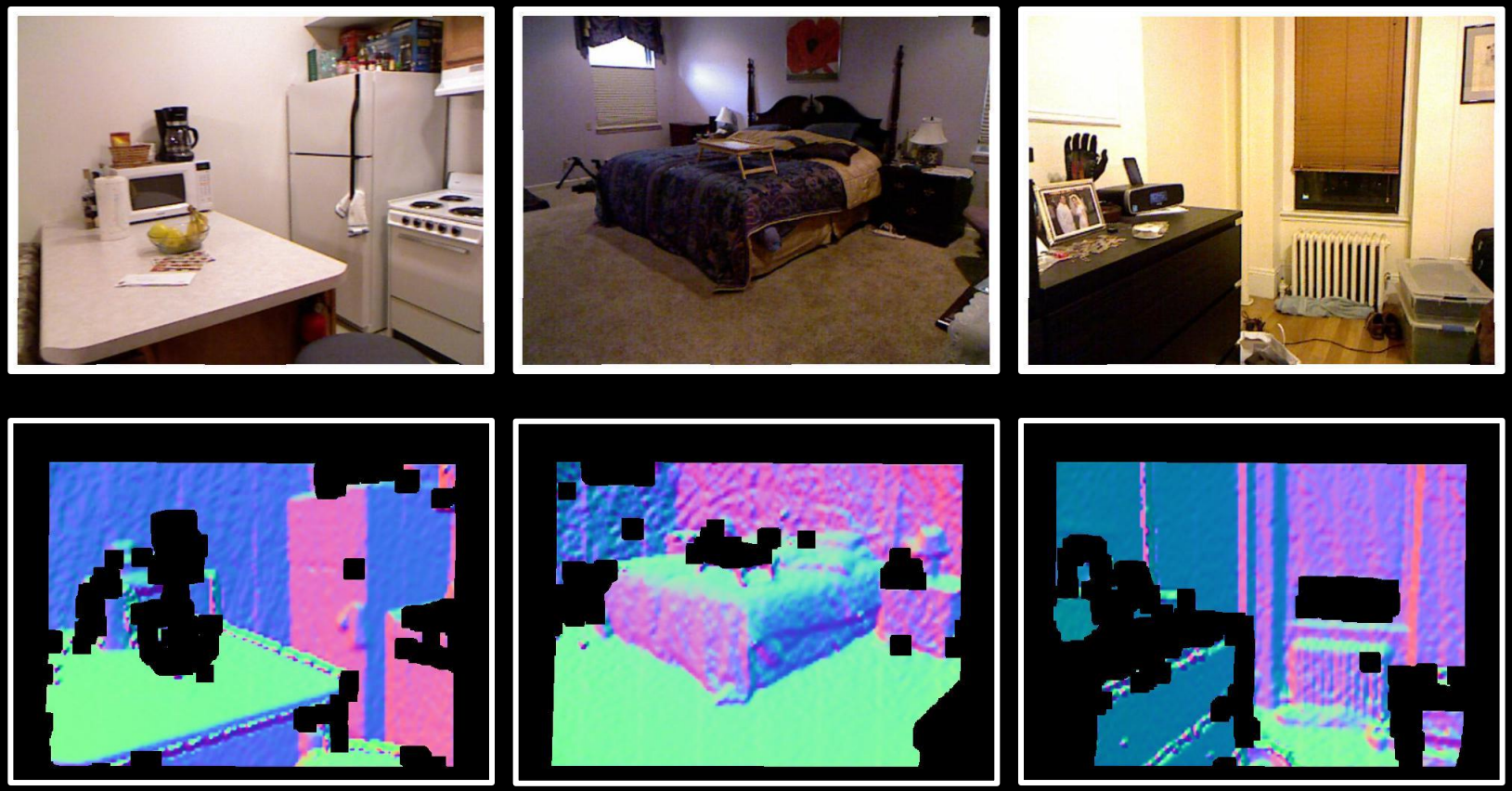
Formulation



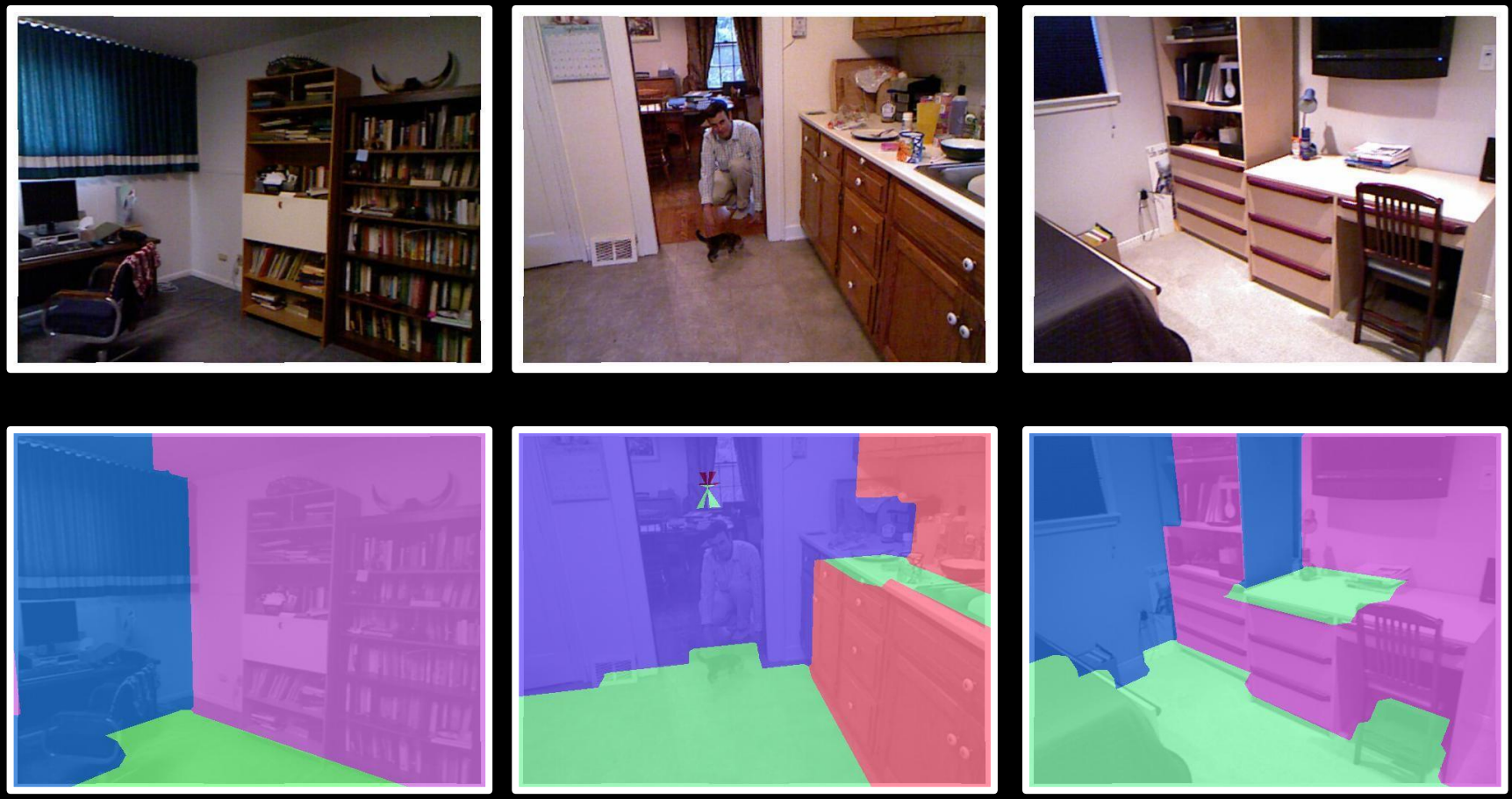
Experimental Results

Dataset

NYU Depth v2: 795 Train, 654 Test

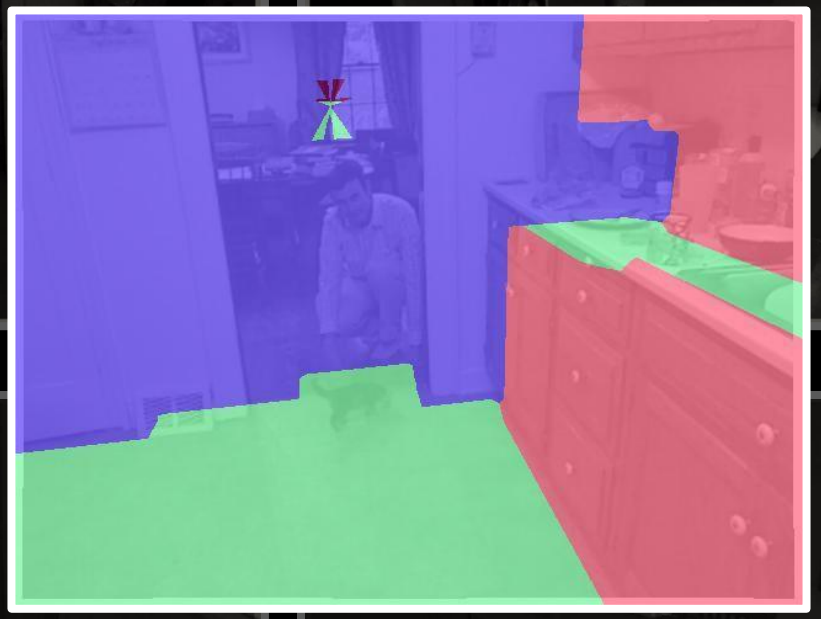


Qualitative Results

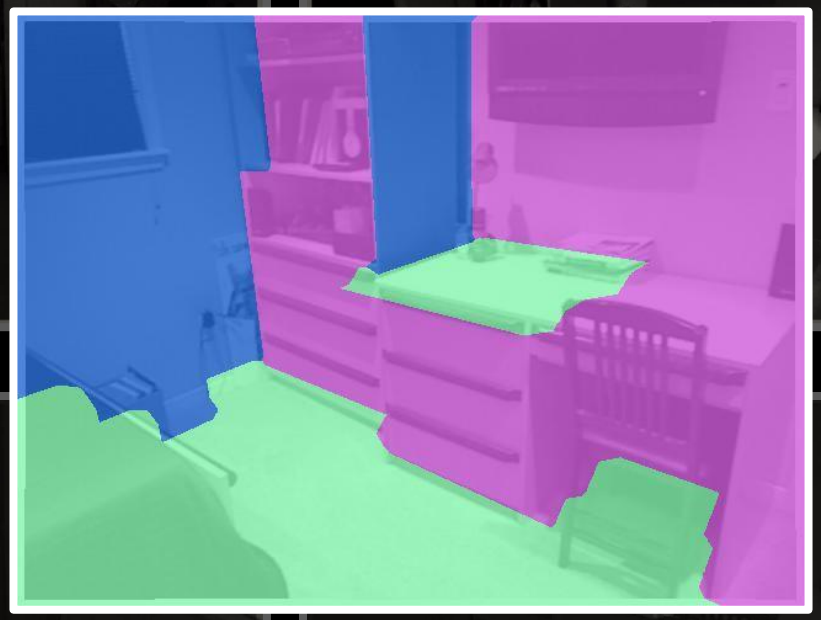


Qualitative Results

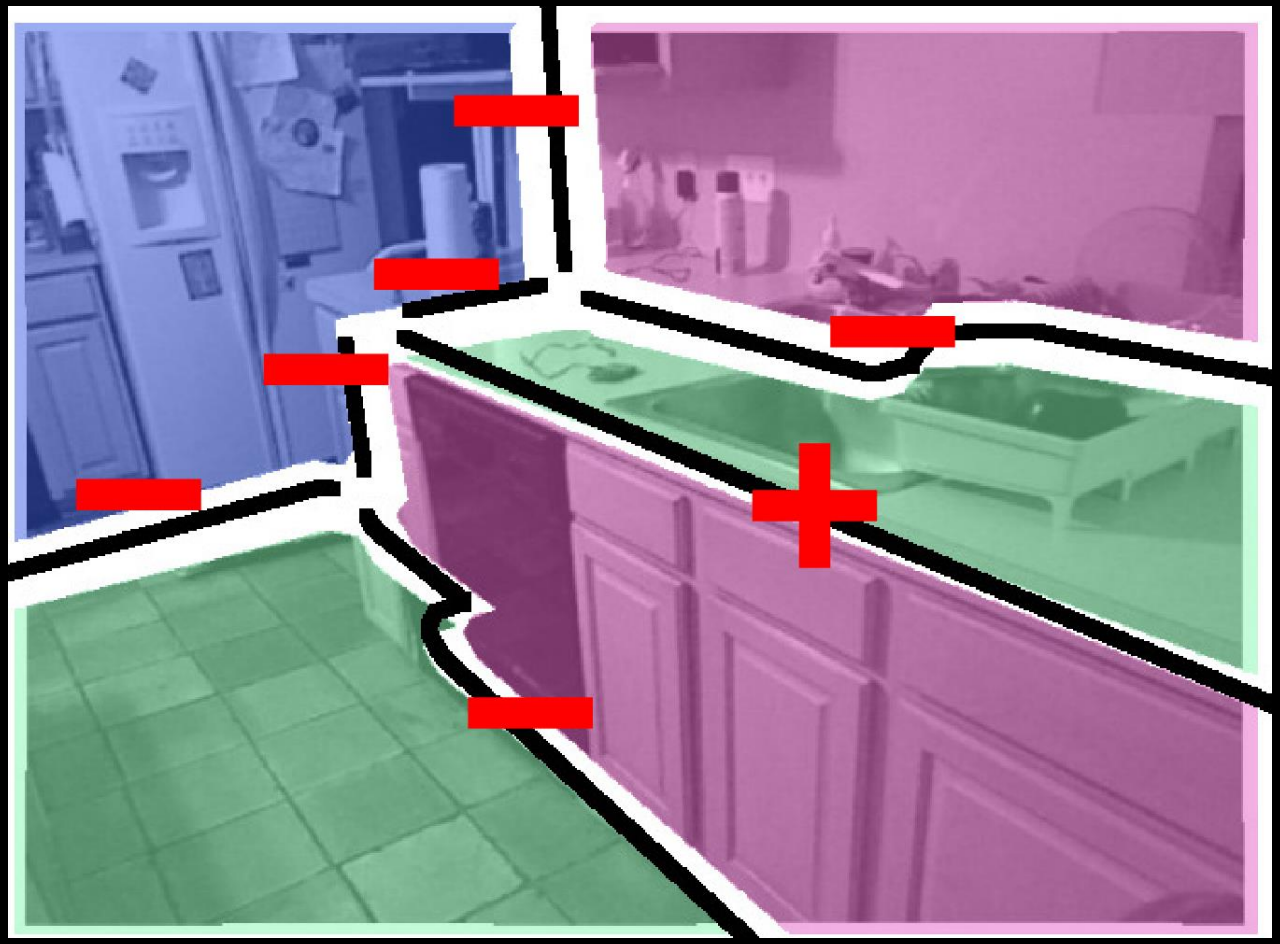
Qualitative Results



Qualitative Results



Surface Connection Graphs

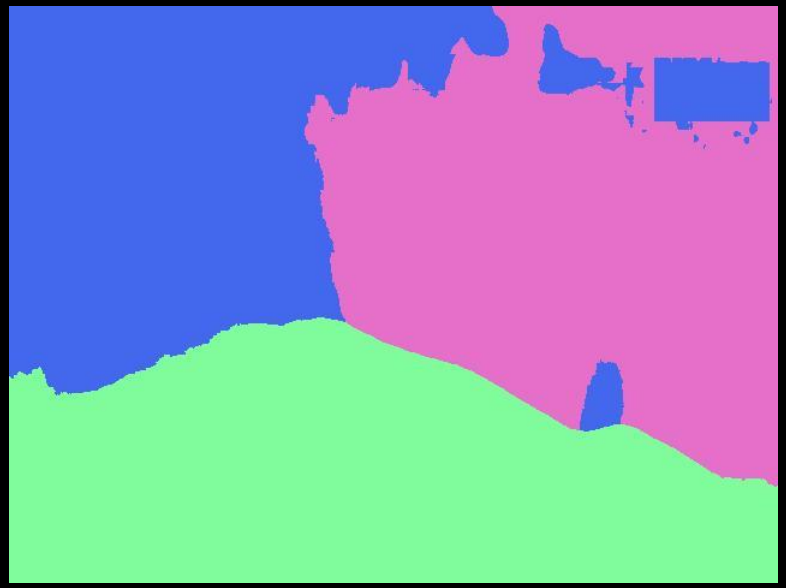


+ Convex − Concave

Baseline

Primary Baseline: 3D Primitives

Input



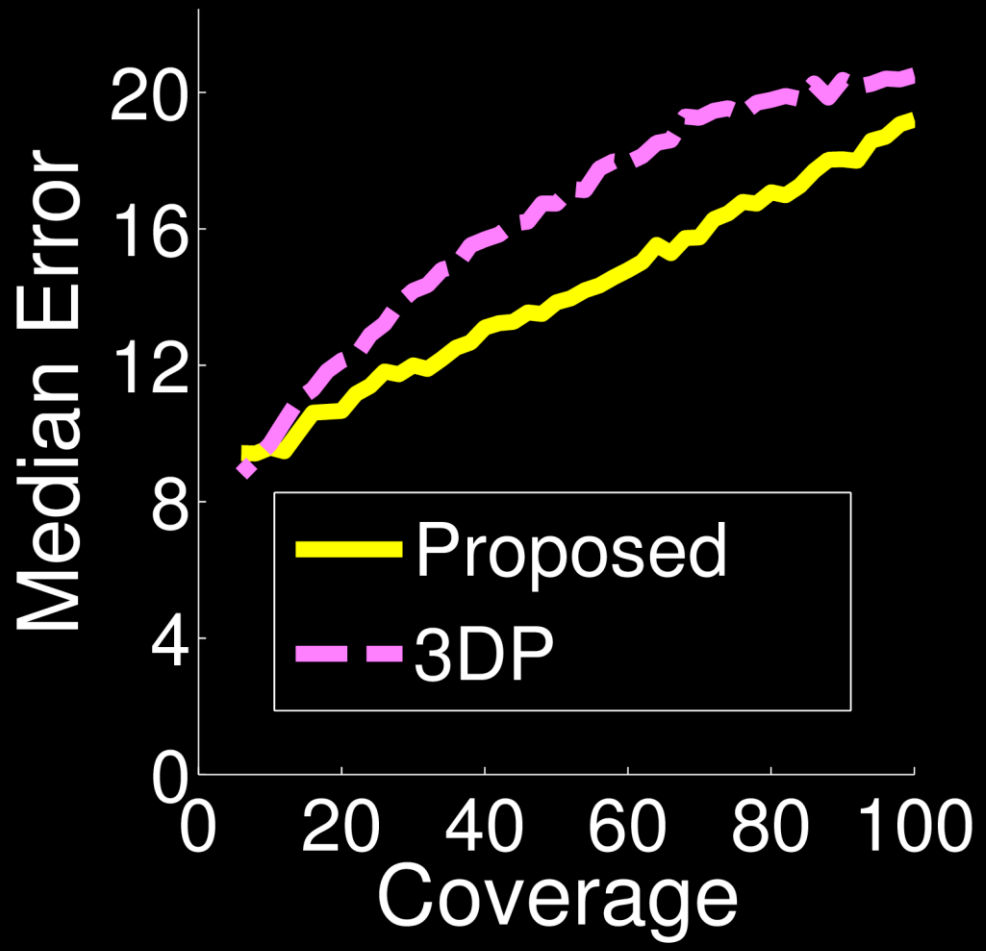
Output

Fouhey, Gupta, Hebert, 2013

Quantitative Results

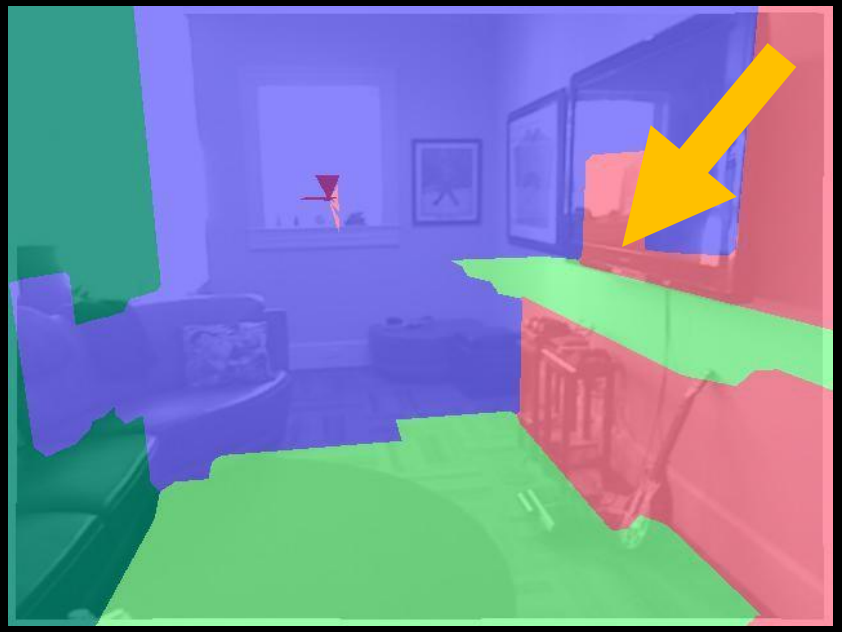
	Summary Stats (^{0}) (Lower Better)			% Good Pixels (Higher Better)		
	Mean	Median	RMSE	11.25^0	22.5^0	30^0
Proposed	**35.1**	**19.2**	**48.7**	**37.6**	**53.3**	**58.9**
3DP	36.0	20.5	49.4	35.9	52.0	57.8
Hedau et al.	40.0	23.5	54.1	34.2	49.3	54.4
Lee et al.	43.3	36.3	54.6	18.6	38.6	49.9

Quantitative Results



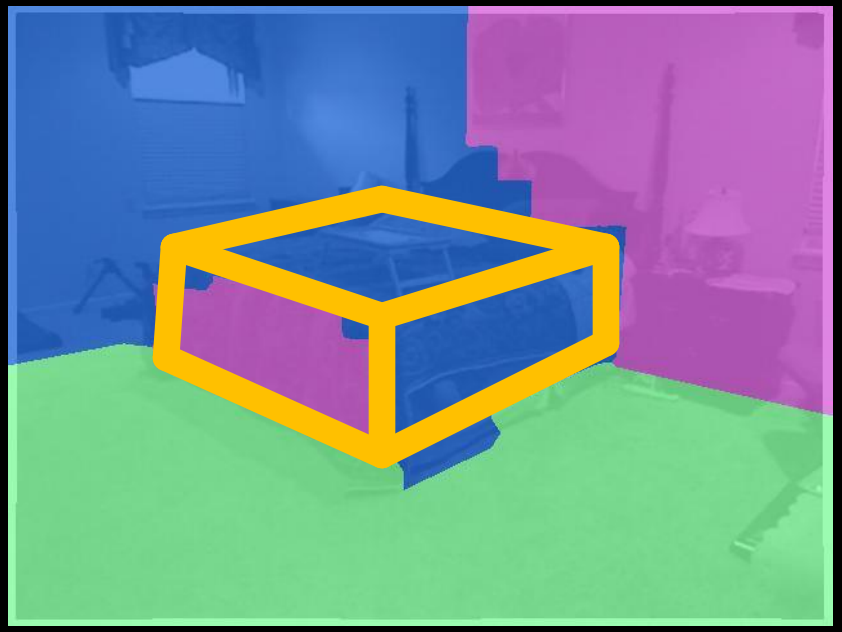
Failure Modes

Mistaken but Confident Evidence

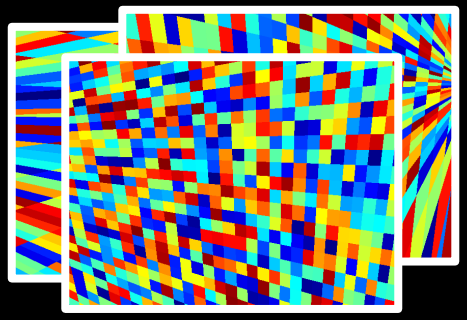


Failure Modes

Missing High-Level Modeling



Conclusions

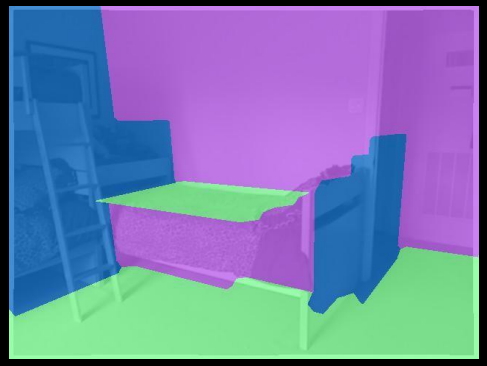


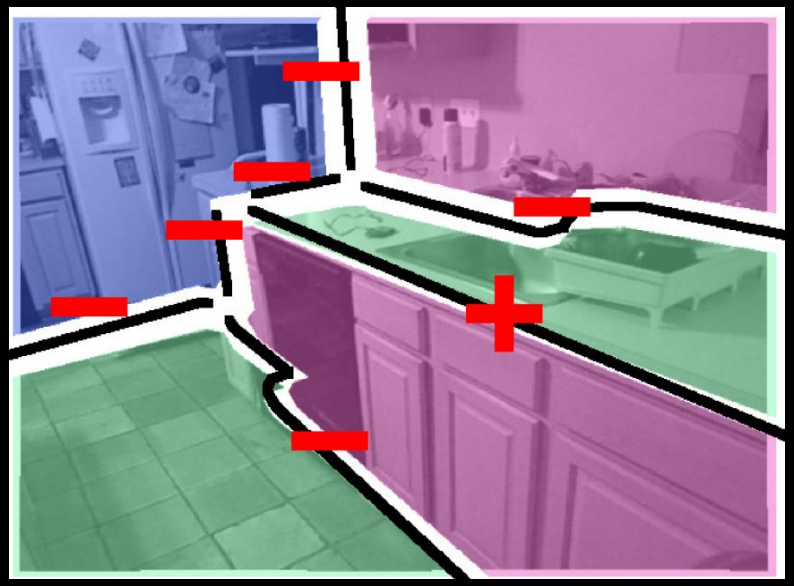
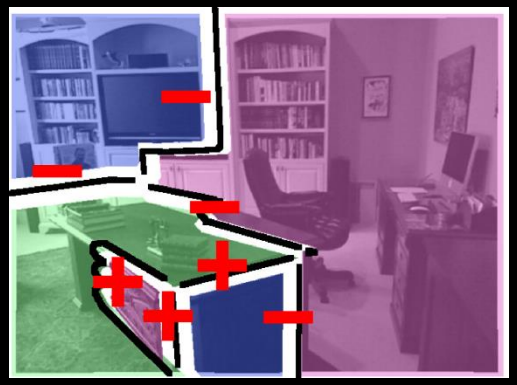
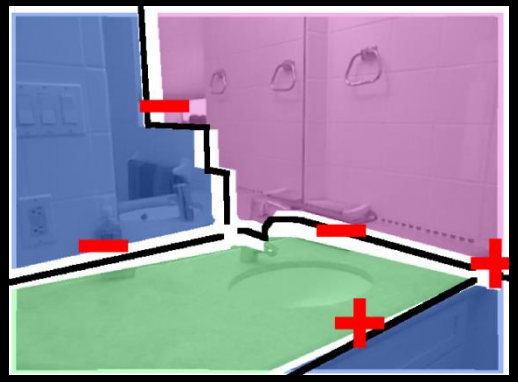
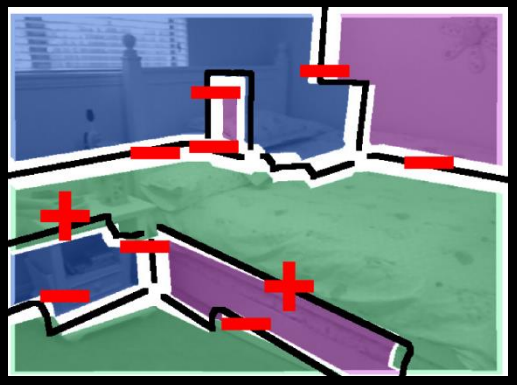
Parameterization

Single Image

$$\arg\max_{\mathbf{x}\in\{0,1\}^n} \mathbf{c}^T\mathbf{x} + \mathbf{x}^T\mathbf{H}\mathbf{x}$$
$$\text{s.t.} \quad \mathbf{A}\mathbf{x} \leq \mathbf{1}$$

Formulation



Discrete Parse

Thank You