15-887
Planning, Execution and Learning

Application:
Examples of Planning in Perception

Maxim Likhachev

Robotics Institute

Carnegie Mellon University

Two Examples

- Graph search for perception
- Planning for Active Perception

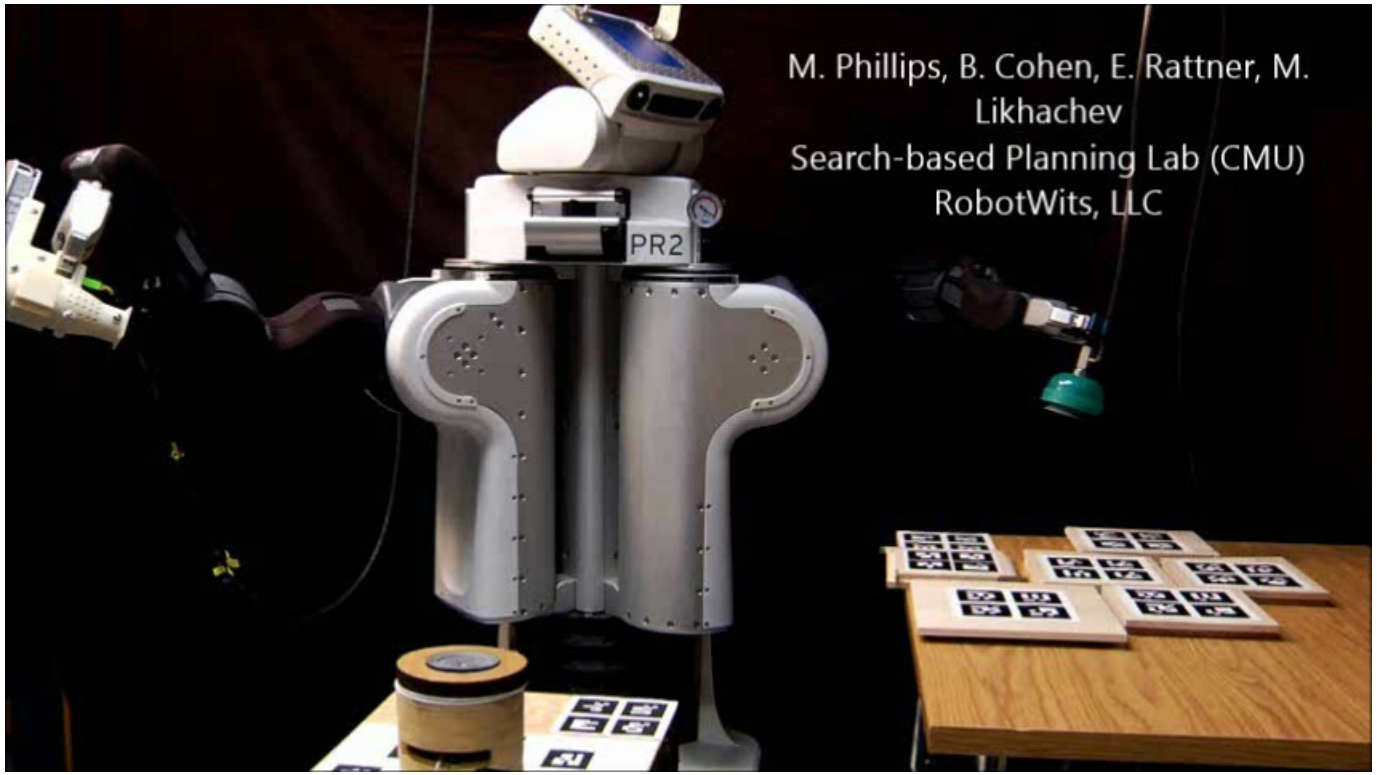
Two Examples

- Graph search for perception
- Planning for Active Perception

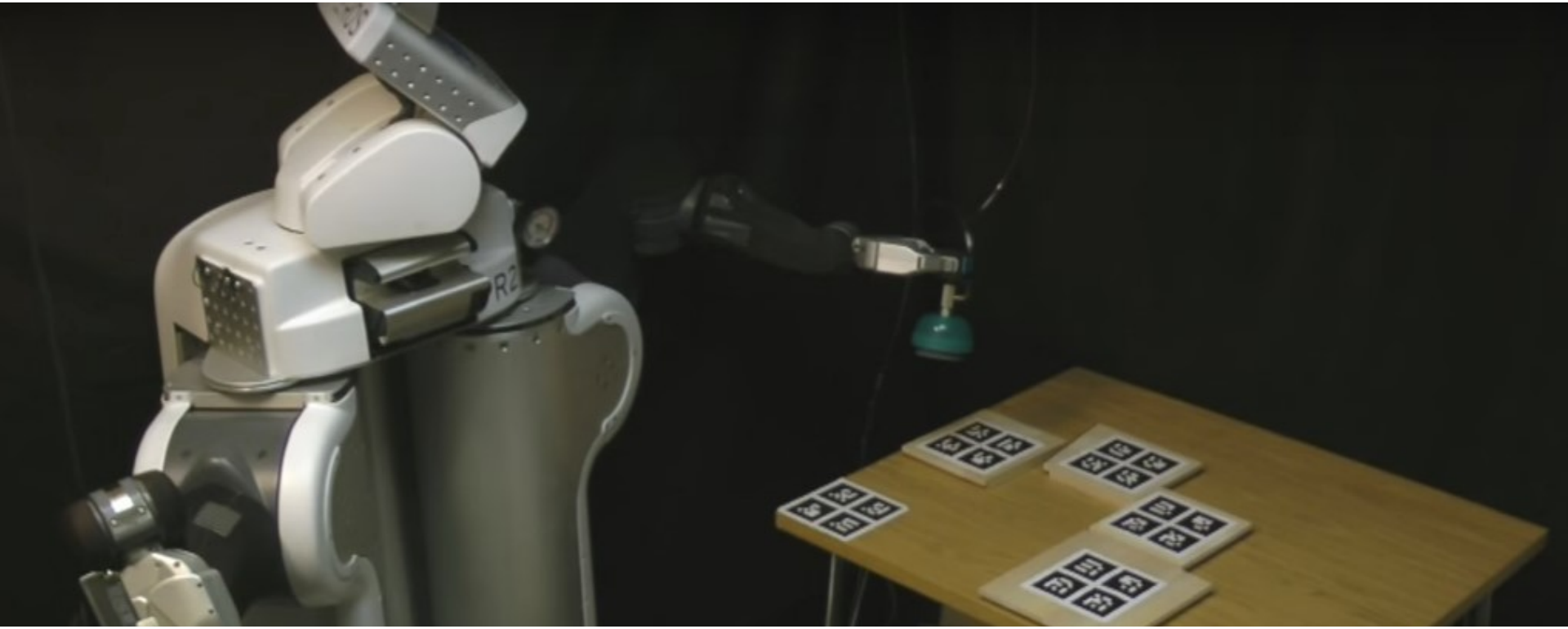
Perception via Search (PERCH)

Venkatraman Narayanan
PhD student at Search-based Planning Lab
Carnegie Mellon University

Birdhouse Project (Cohen et al.)



Perception is Brittle



Even for the (commonly occurring)

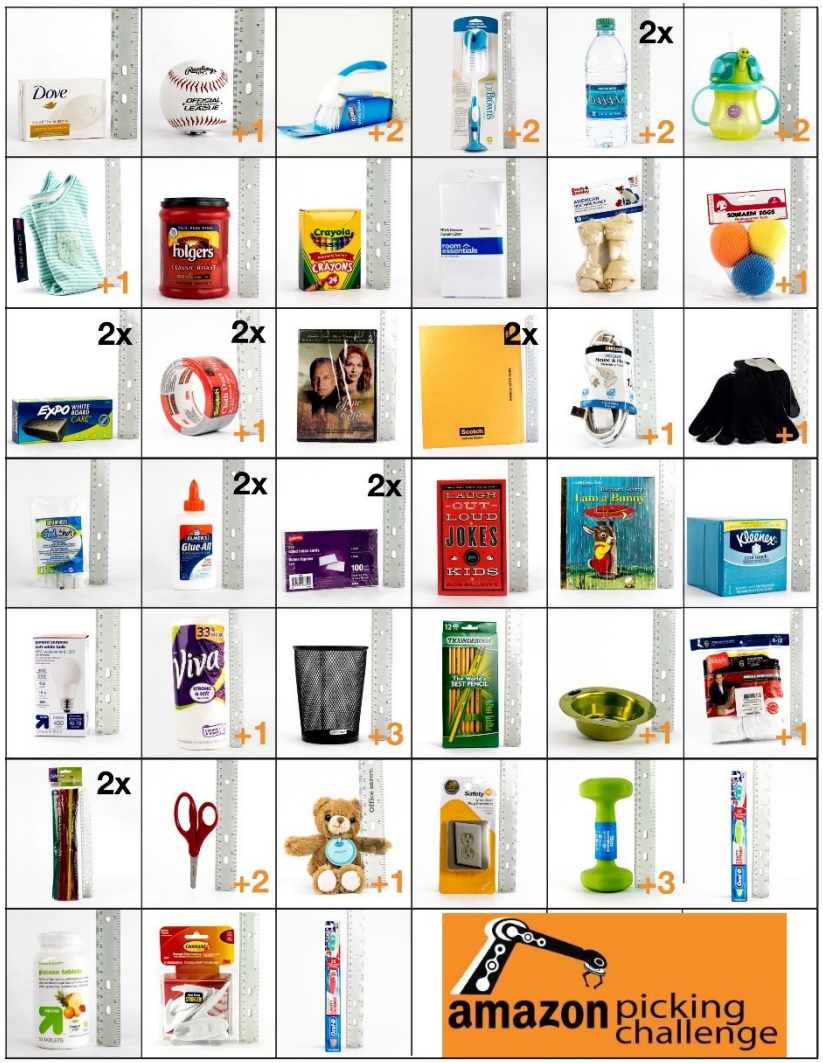
Object *Instance* Detection

(Objects with known 3D models)

Object Instance Detection

Joint work with Upenn (K. Daniilidis), GDLS, RR

Amazon Picking Challenge (APC)



APC: Perception is Hard

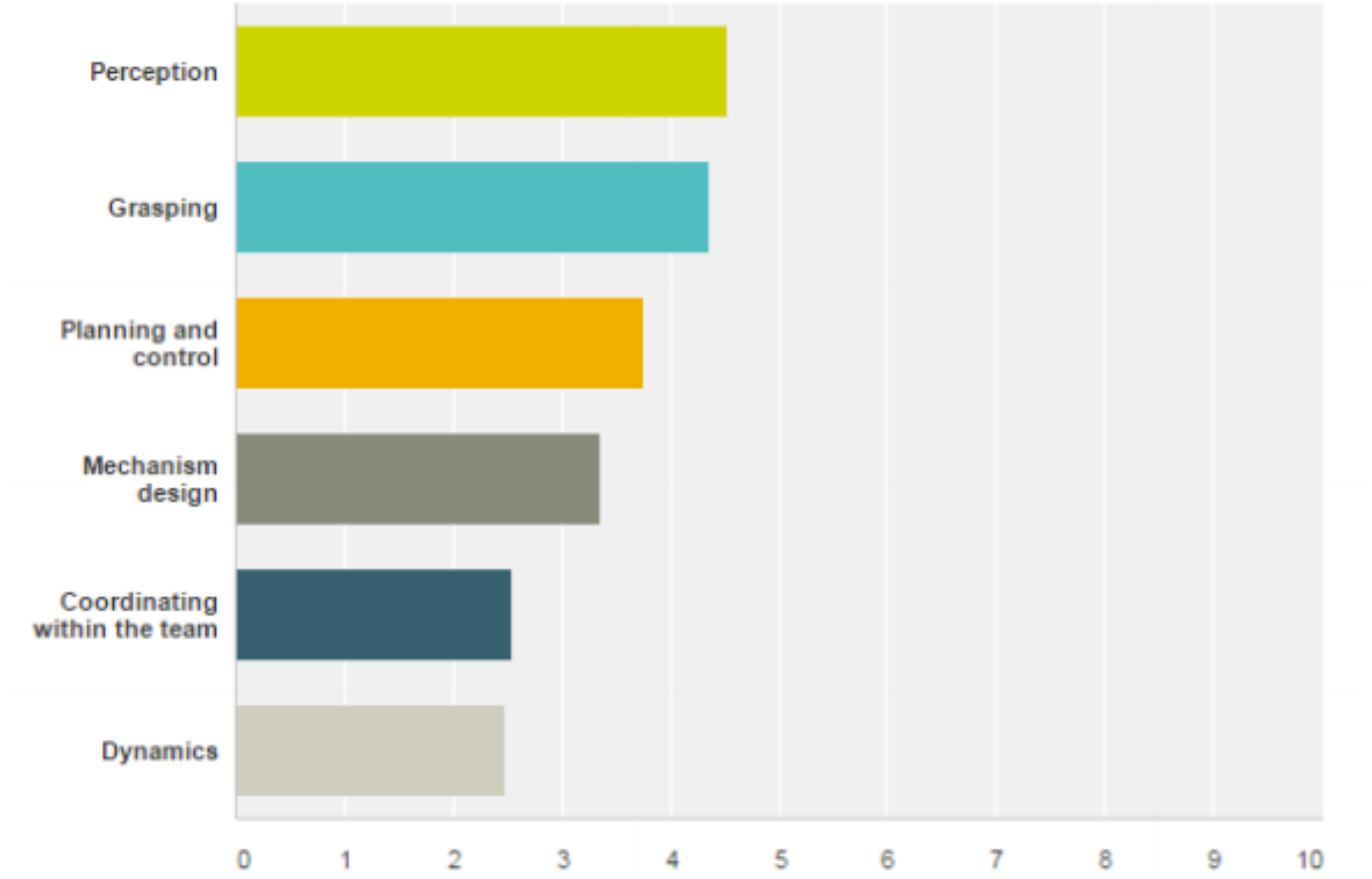


Fig. 6. "Please rank order the different aspects of the APC by their difficulty, starting with 'most difficult' at the top".

[Correll et. al 2016]

We see

Depth Image

We know

- List of objects in the scene
- 3D models of those objects

We see

Depth Image

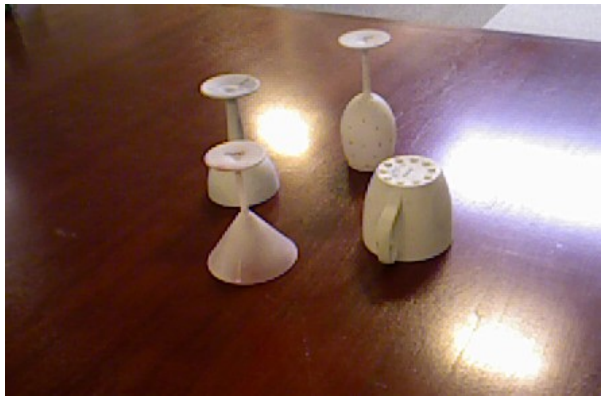
We know

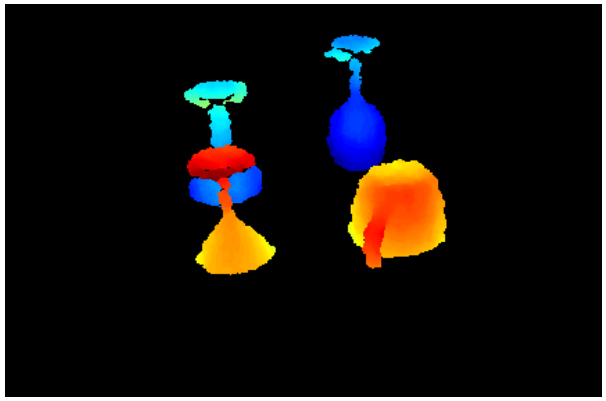
- List of objects in the scene
- 3D models of those objects

We do

Global reasoning

Deliberative Perception: Formulation





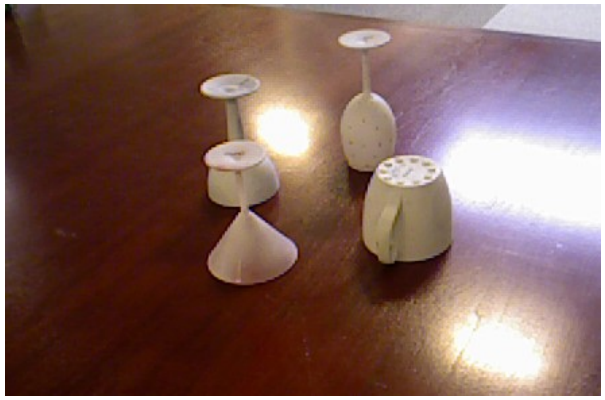
identify type and 3 DoF pose of all objects
in the scene (point cloud/depth image)

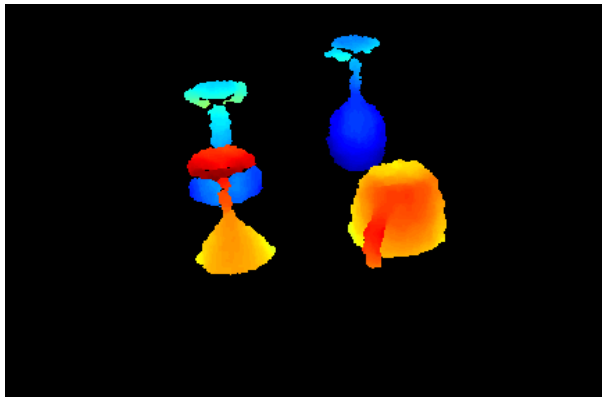
assuming robot knows

3D models of objects

6 DoF camera pose

Deliberative Perception: Formulation





Optimize over space of joint object configurations

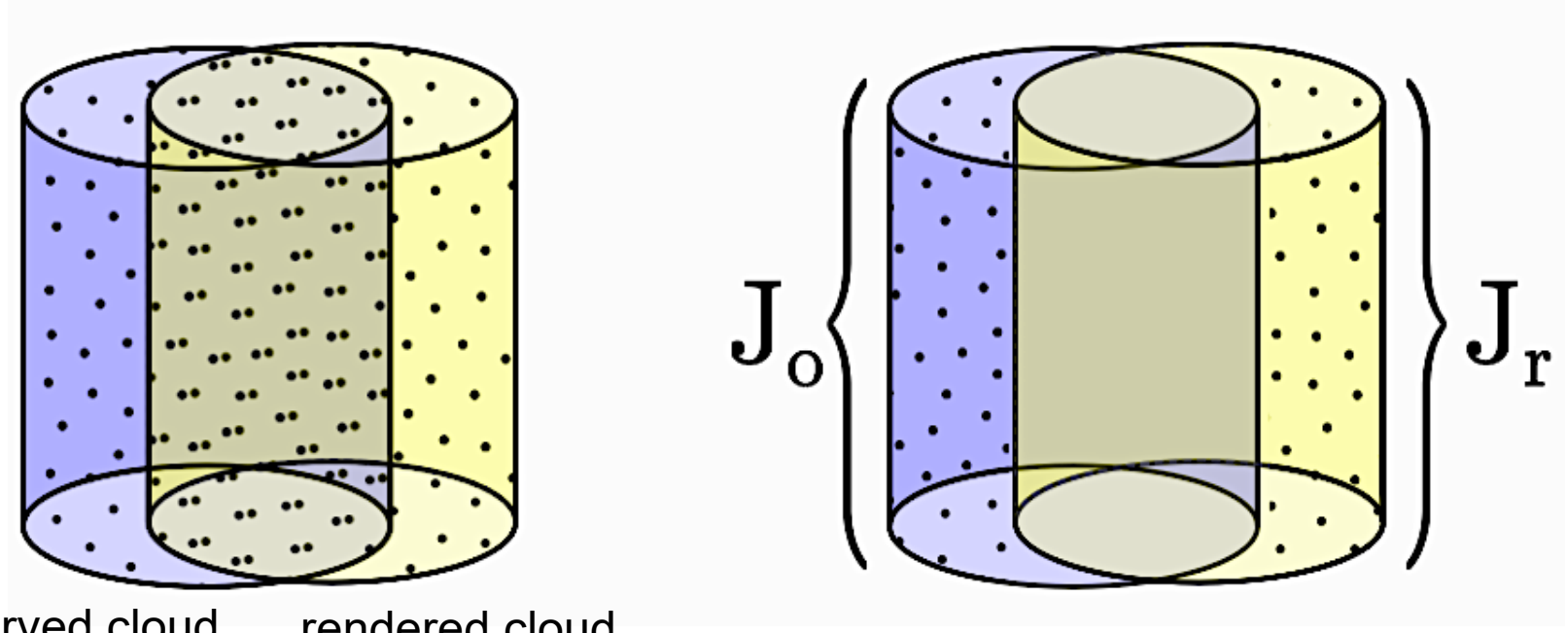
Ideas for the objective function?

Deliberative Perception: Formulation

$$\min_{O_{1:K}} J(O_{1:K})$$

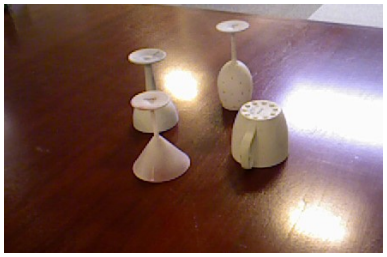
$$J(O_{1:K}) = J_{rendered}(O_{1:K}) + J_{observed}(O_{1:K})$$

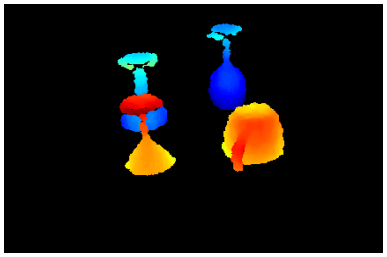
outliers in rendered cloud outliers in observed cloud



observed cloud rendered cloud

Deliberative Perception: Formulation





Optimize over space of joint object configurations

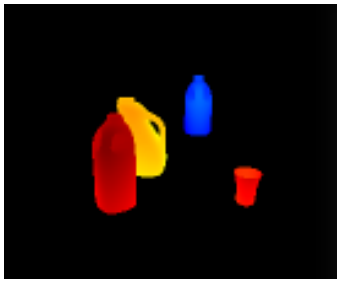
Ideas for better search-space?

combinatorial search space

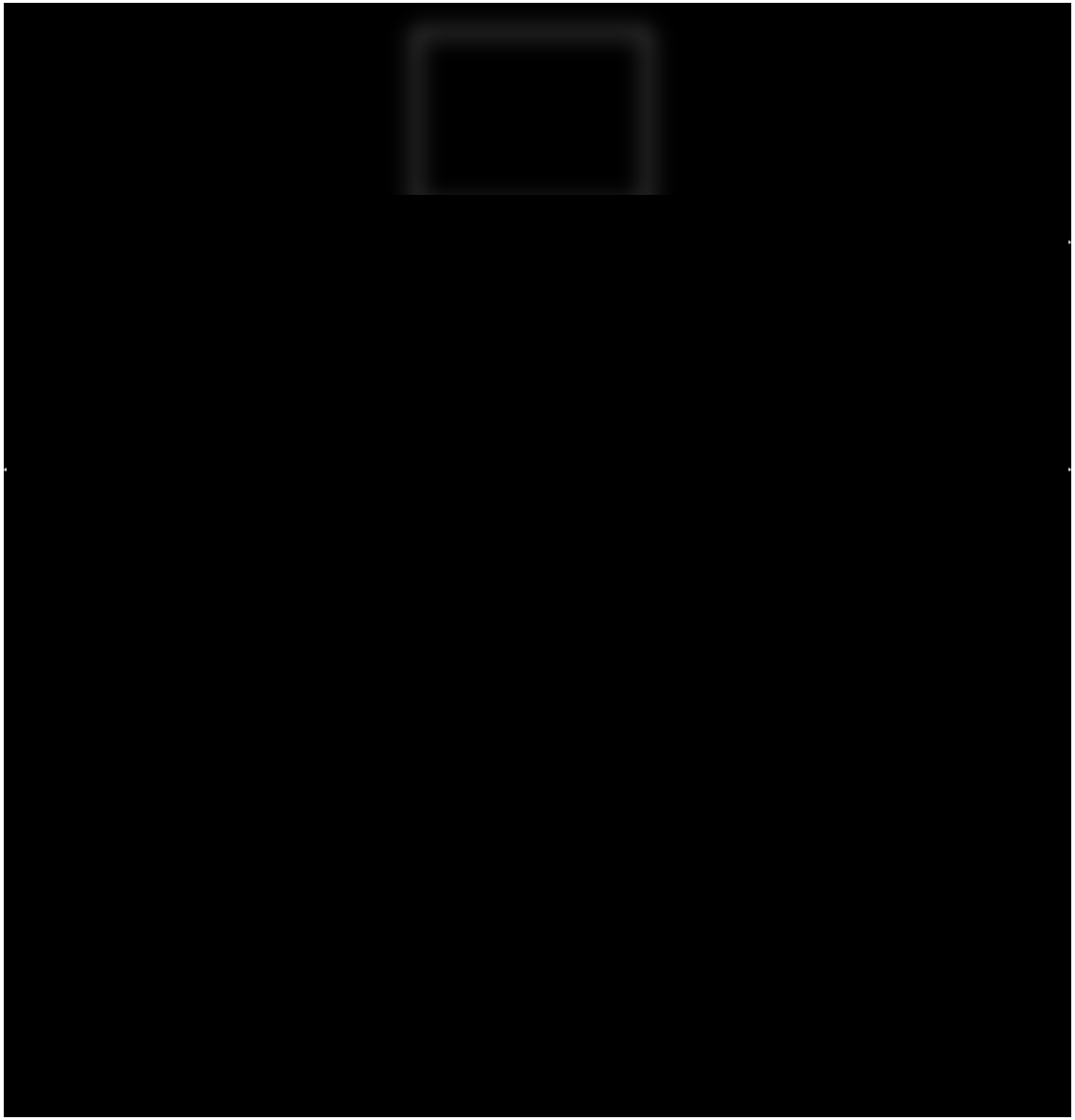
e.g.: 4 objects, 10 (x,y) locations, 10 orientations

100^4 scenes ~12 days @ 10 ms / render on single GPU

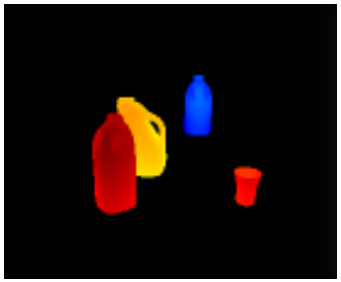
Deliberative Perception: Formulation



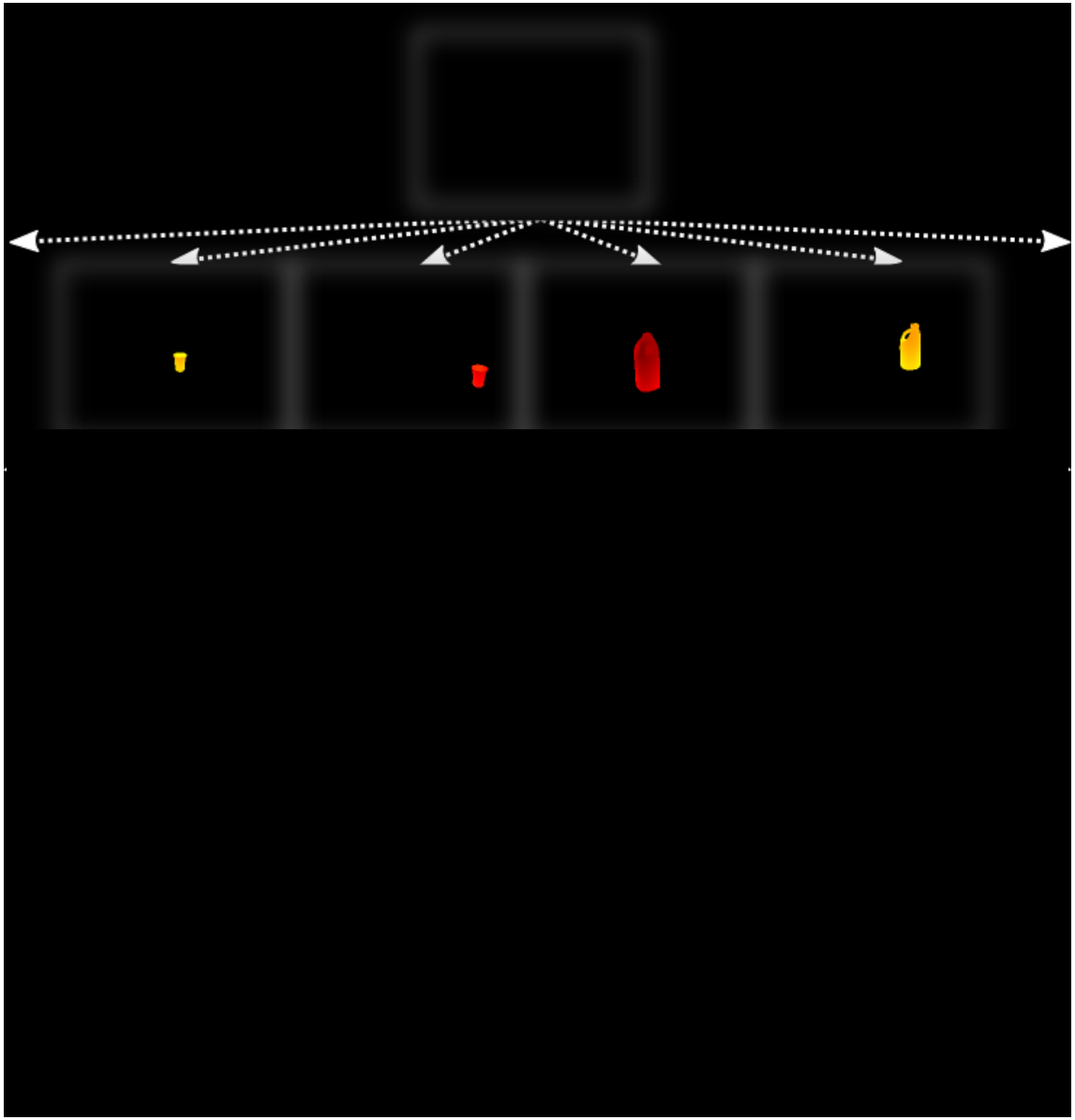
optimal solution



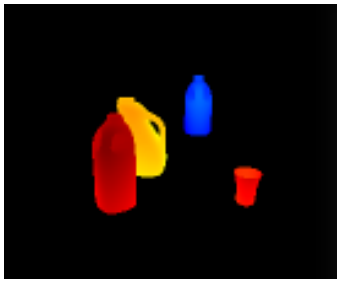
Deliberative Perception: Formulation



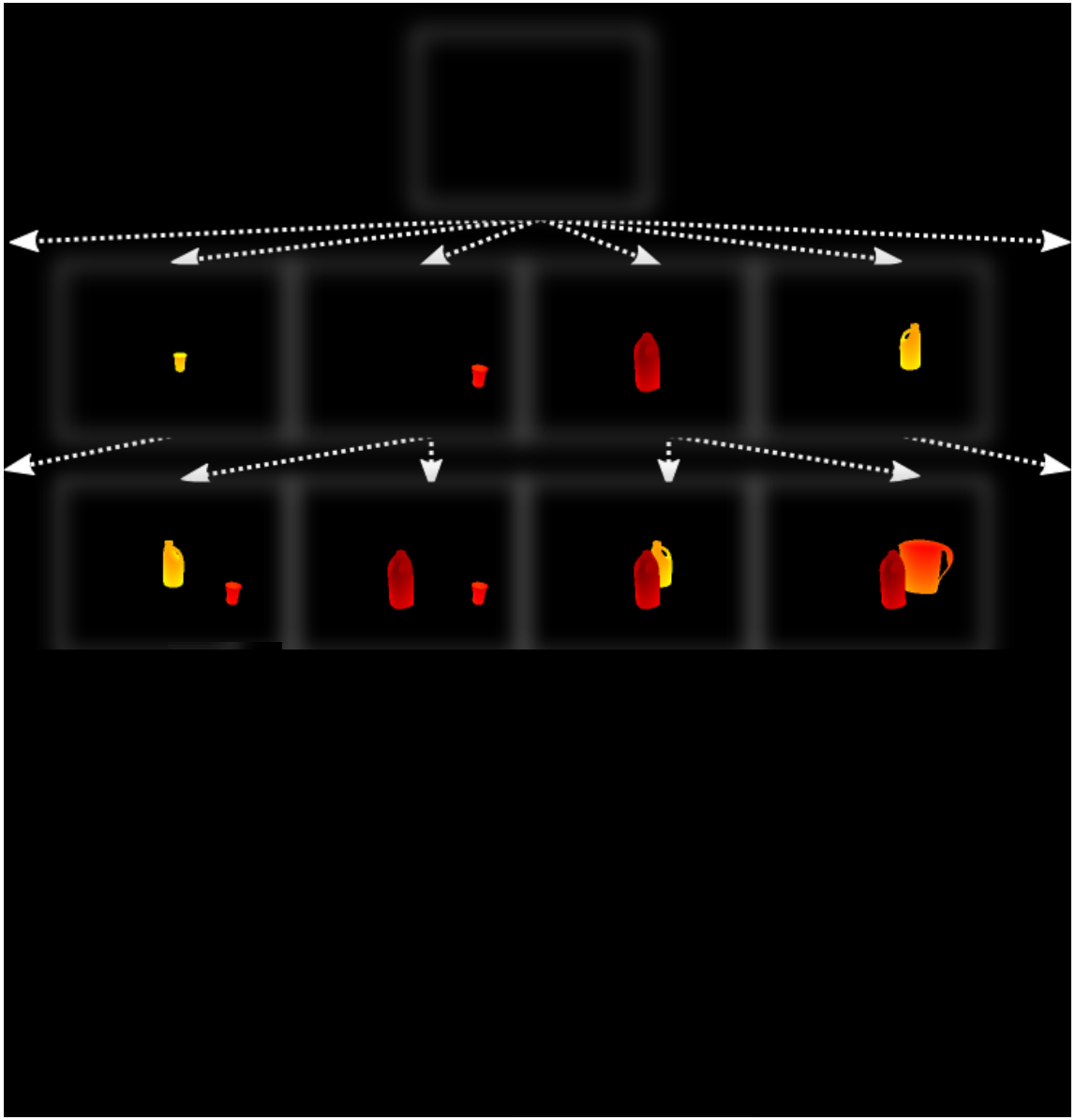
optimal solution



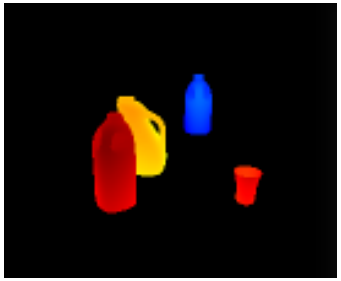
Deliberative Perception: Formulation



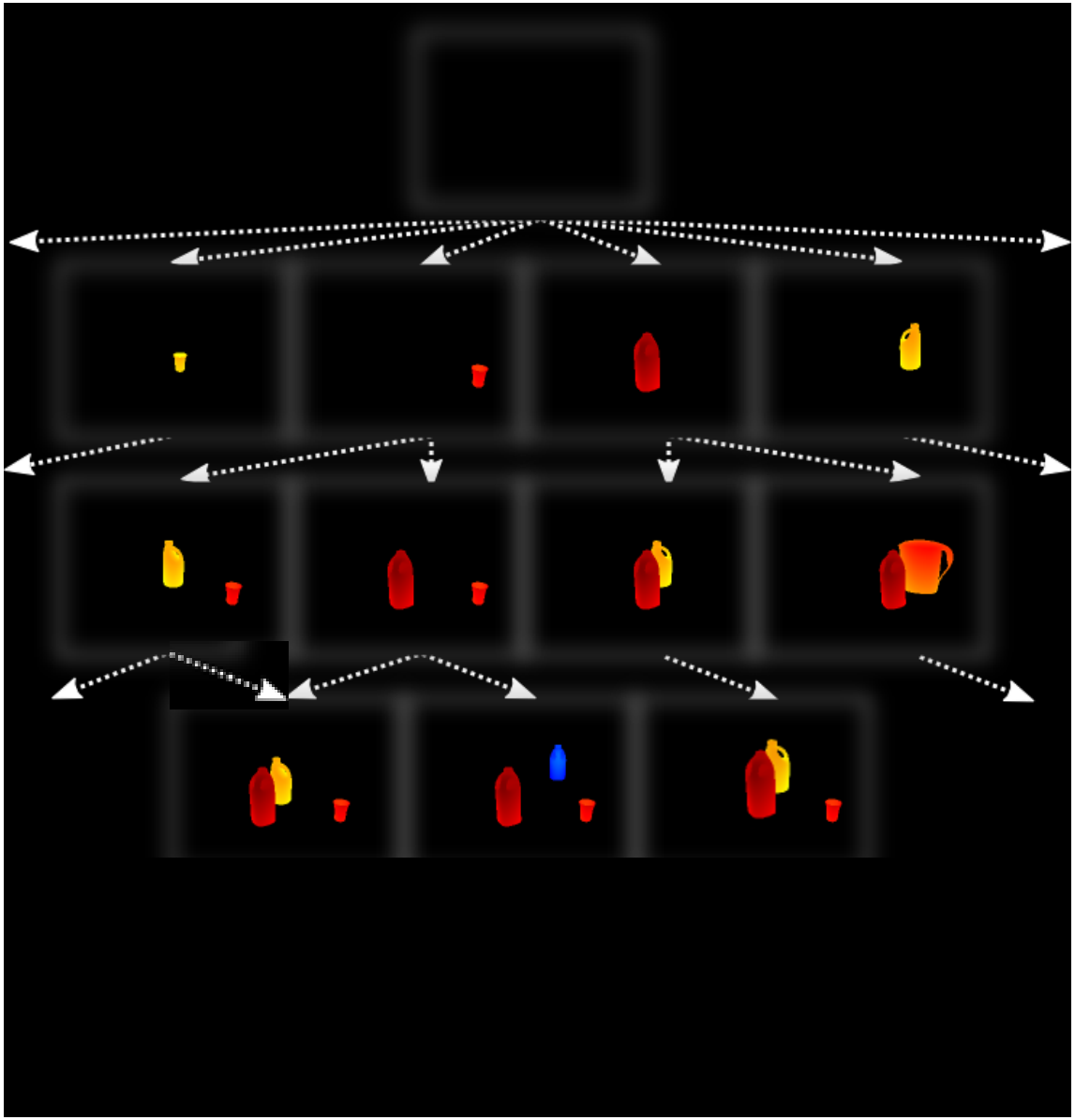
optimal solution



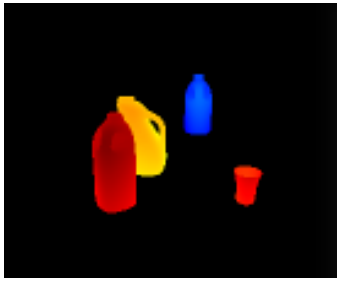
Deliberative Perception: Formulation



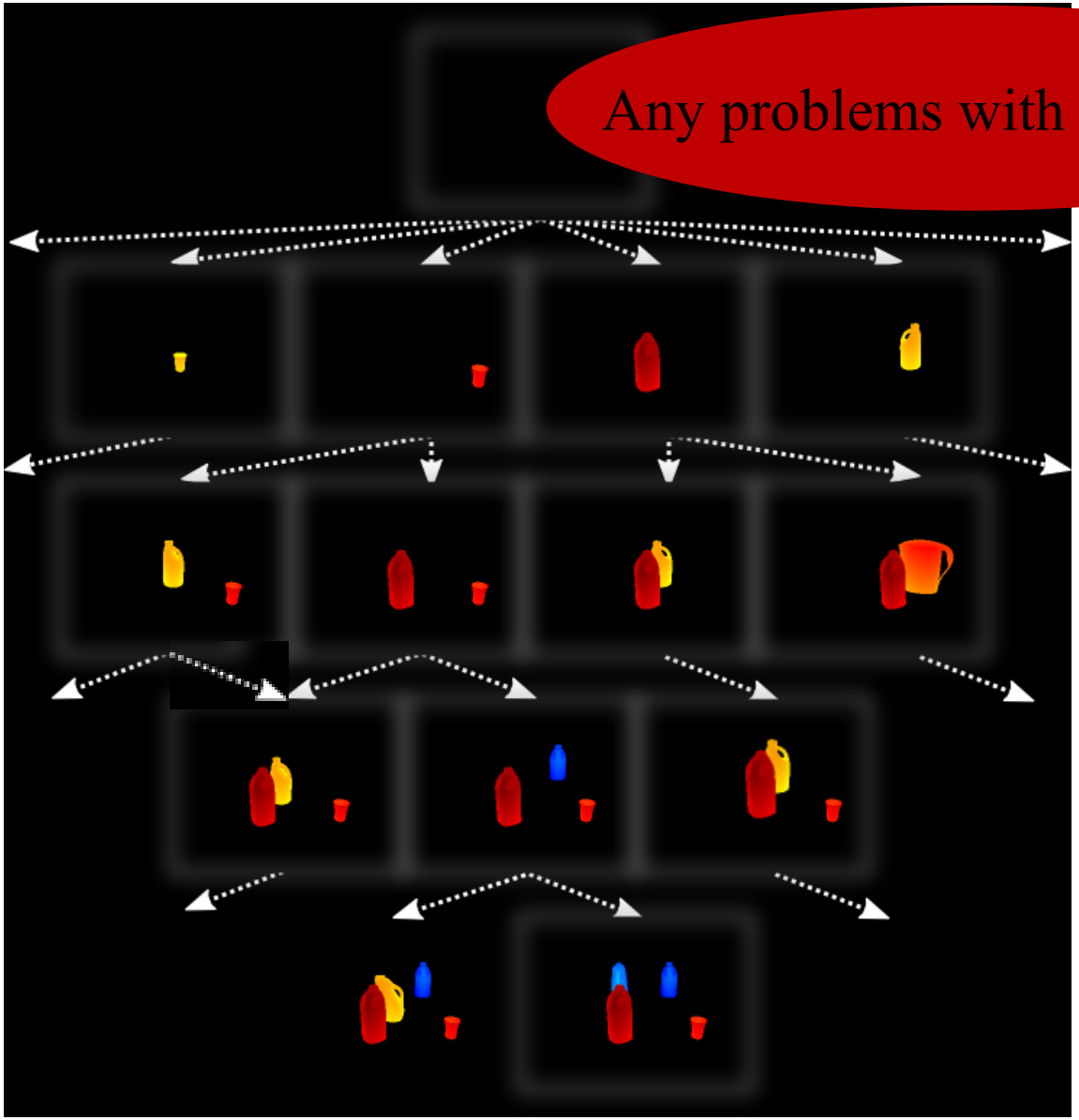
optimal solution



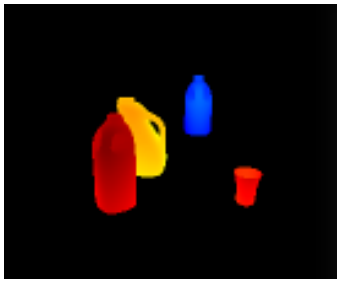
Deliberative Perception: Formulation



optimal solution



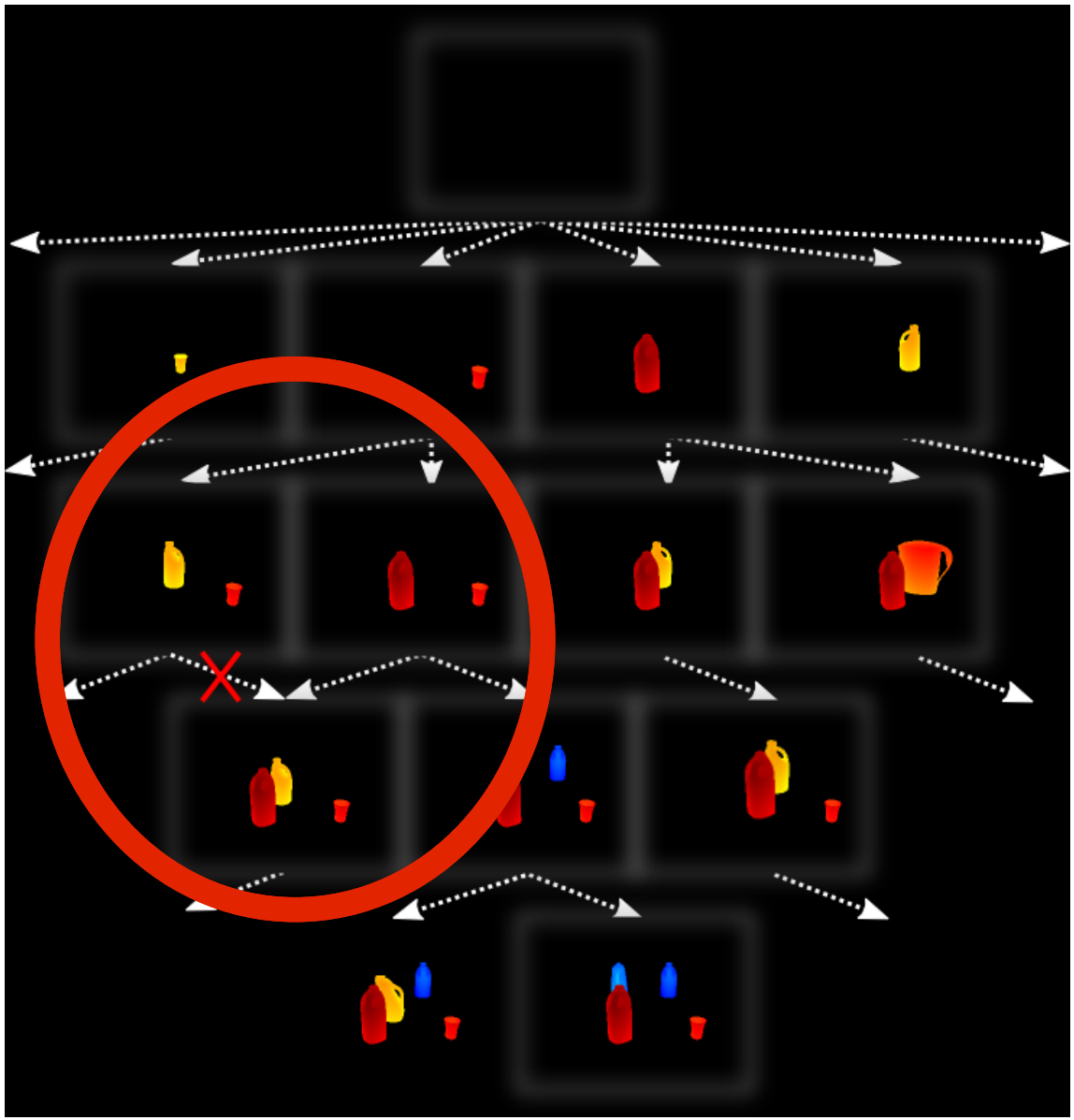
Deliberative Perception: Formulation



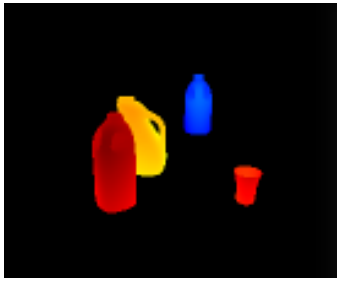
optimal solution

Monotone Scene Generation Tree

(only objects/poses not occluding already set objects are allowed = positive edgecosts)



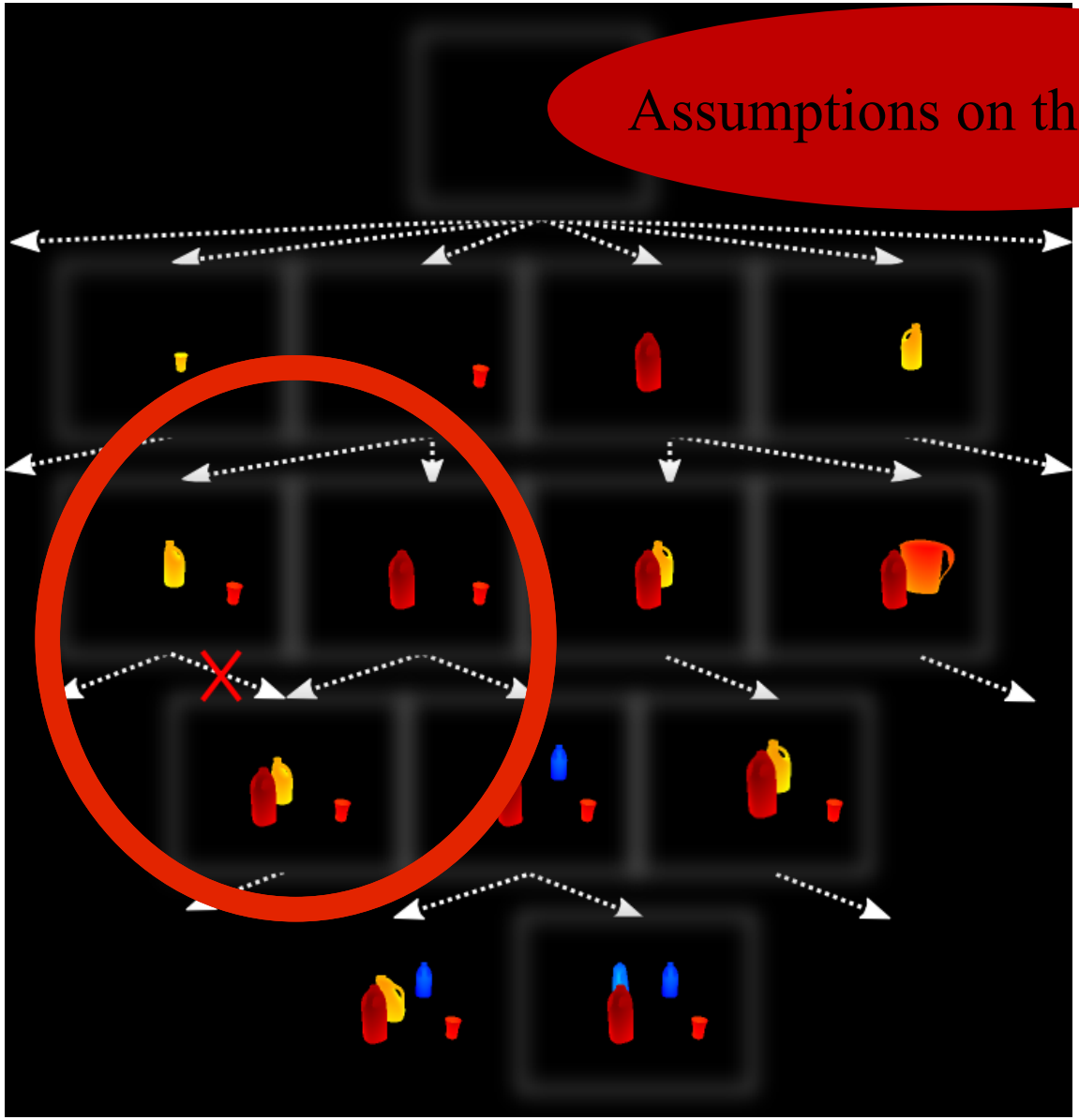
Deliberative Perception: Formulation



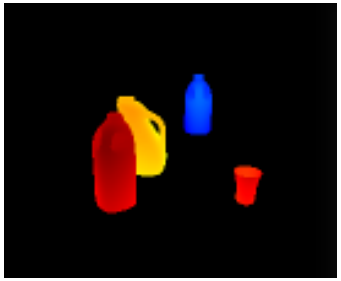
optimal solution

Monotone Scene Generation Tree

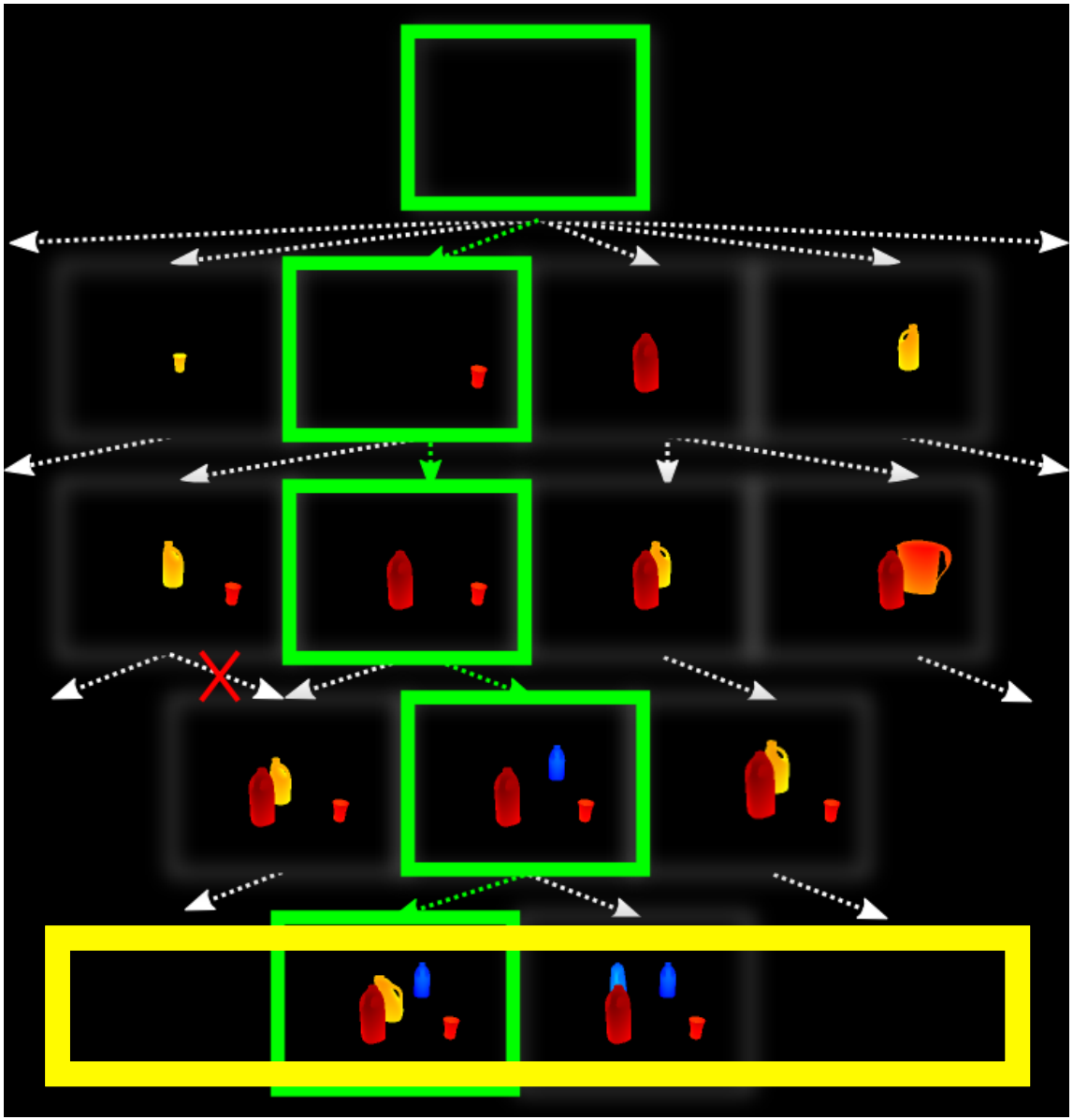
(only objects/poses not occluding already set objects are allowed = positive edgecosts)



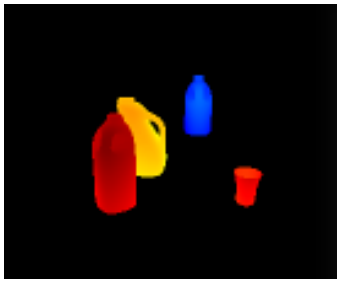
Deliberative Perception: Formulation



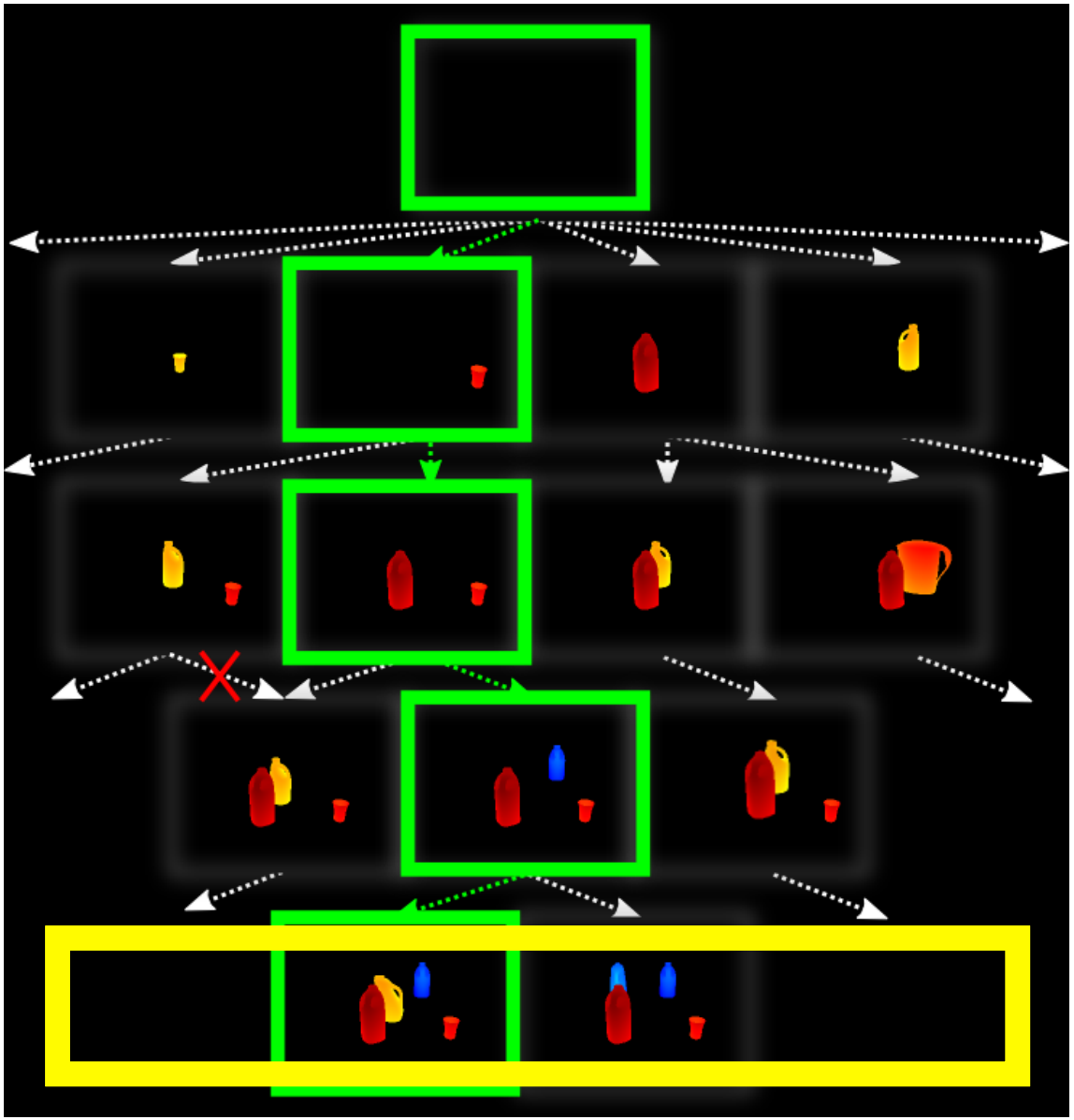
optimal solution



Deliberative Perception: Formulation



optimal solution



Want edge costs to:

Be non-negative

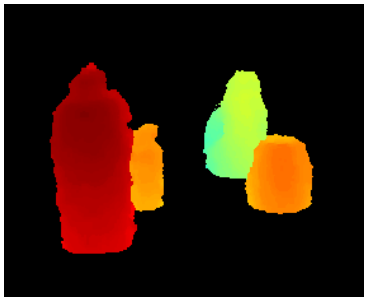
Add up to optimization objective

Deliberative Perception: Formulation

$$J(O_{1:K}) = J_{rendered}(O_{1:K}) + J_{observed}(O_{1:K})$$

outliers in rendered cloud outliers in observed cloud

input rgb (unused)



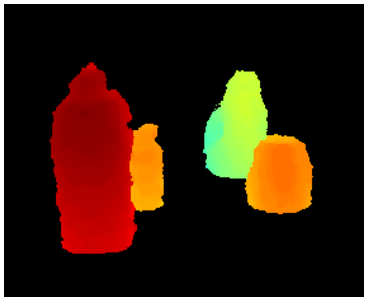
input depth

Deliberative Perception: Formulation

$$J(O_{1:K}) = J_{rendered}(O_{1:K}) + J_{observed}(O_{1:K})$$

outliers in rendered cloud outliers in observed cloud

input rgb (unused)



input depth

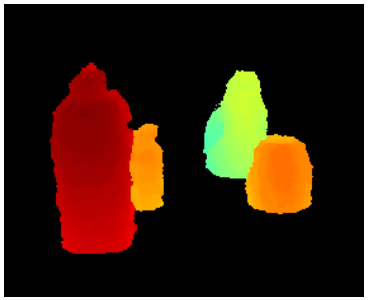
$J_{rendered}(O_1)$
+
$J_{observed}(O_1)$

Deliberative Perception: Formulation

$$J(O_{1:K}) = J_{rendered}(O_{1:K}) + J_{observed}(O_{1:K})$$

outliers in rendered cloud outliers in observed cloud

input rgb (unused)



input depth

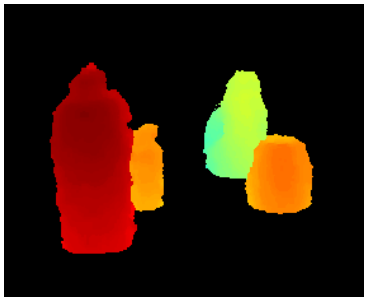
$$J_{rendered}(O_1) + J_{observed}(O_1) + J_{rendered}(O_2) + J_{observed}(O_2)$$

Deliberative Perception: Formulation

$$J(O_{1:K}) = J_{rendered}(O_{1:K}) + J_{observed}(O_{1:K})$$

outliers in rendered cloud outliers in observed cloud

input rgb (unused)



input depth

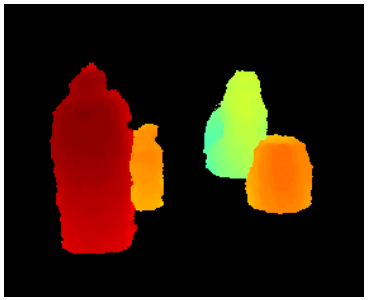
$$J_{rendered}(O_1) + J_{observed}(O_1) + J_{rendered}(O_2) + J_{observed}(O_2) + J_{rendered}(O_3) + J_{observed}(O_3)$$

Deliberative Perception: Formulation

$$J(O_{1:K}) = J_{rendered}(O_{1:K}) + J_{observed}(O_{1:K})$$

outliers in rendered cloud outliers in observed cloud

input rgb (unused)



input depth

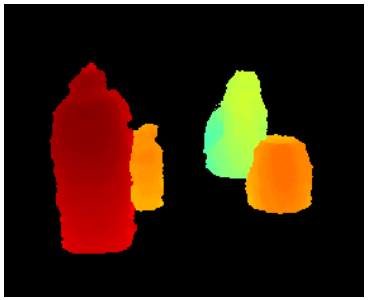
$$J_{rendered}(O_1) + J_{observed}(O_1) + J_{rendered}(O_2) + J_{observed}(O_2) + J_{rendered}(O_3) + J_{observed}(O_3) + J_{rendered}(O_3) + J_{observed}(O_3)$$

Deliberative Perception: Formulation

$$J(O_{1:K}) = J_{rendered}(O_{1:K}) + J_{observed}(O_{1:K})$$

outliers in rendered cloud outliers in observed cloud

input rgb (unused)



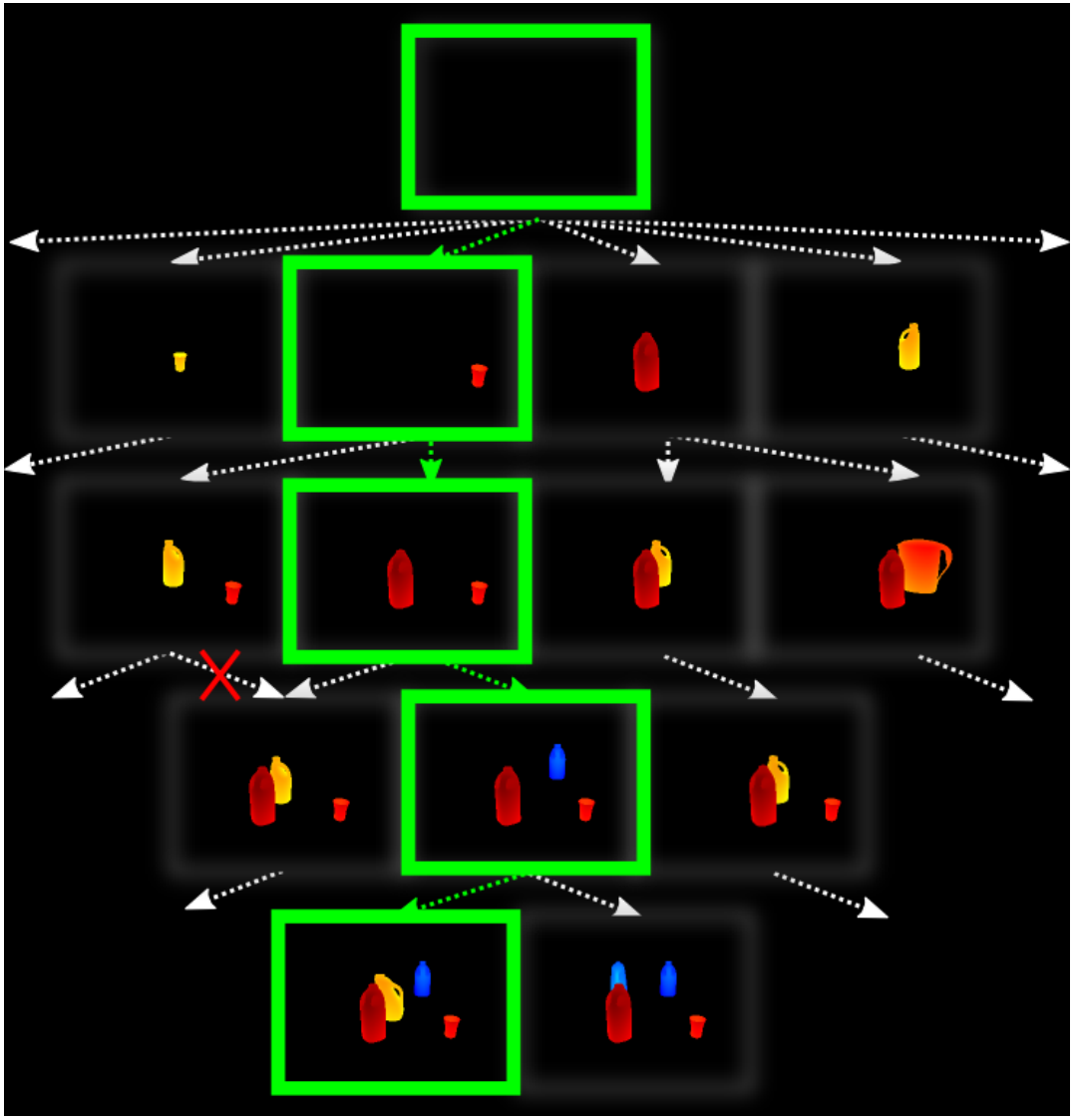
input depth

$$J(O_{1:K}) = \Sigma_i J_{rendered}(O_i) + J_{observed}(O_i)$$

s.t. O_{i+1} does not occlude O_j for all $j < i+1$

PERCH: Perception via Search

Monotone Scene Generation Tree



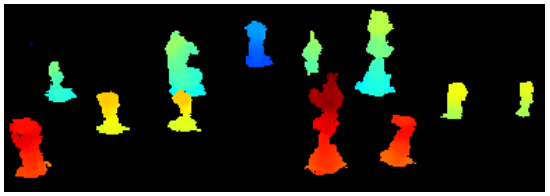
Edge cost from level i-1 to i

$J_{rendered}(O_i) + J_{observed}(O_i)$

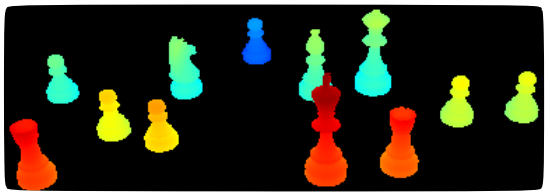
PERCH, Narayanan & Likhachev, ICRA '16

PERCH: Qualitative Results

Input Depth Image

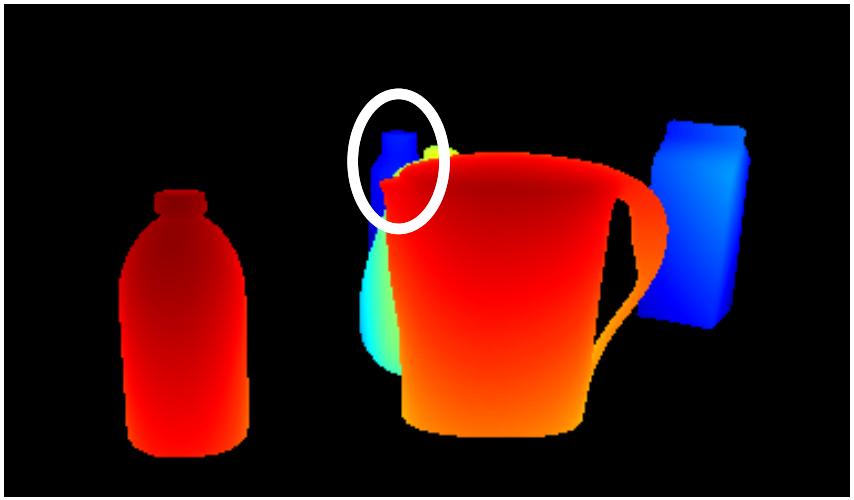


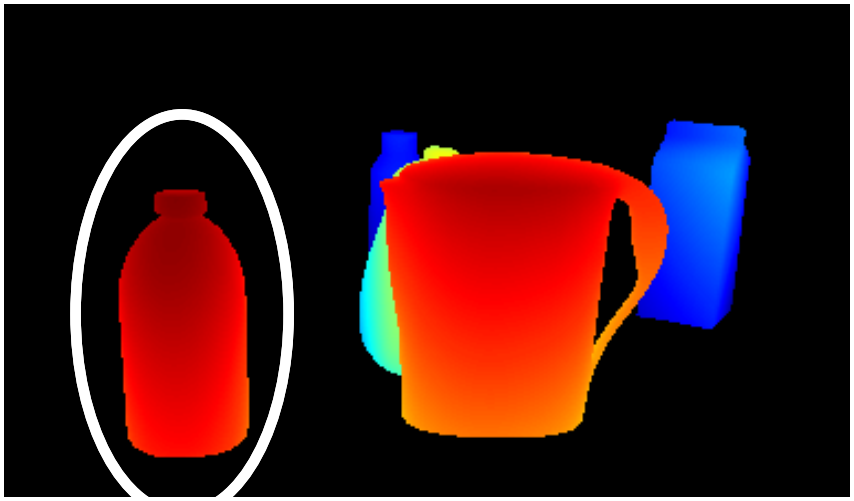
Output Depth Image



PERCH, Narayanan & Likhachev, ICRA ’16

Leveraging Discriminative Guidance

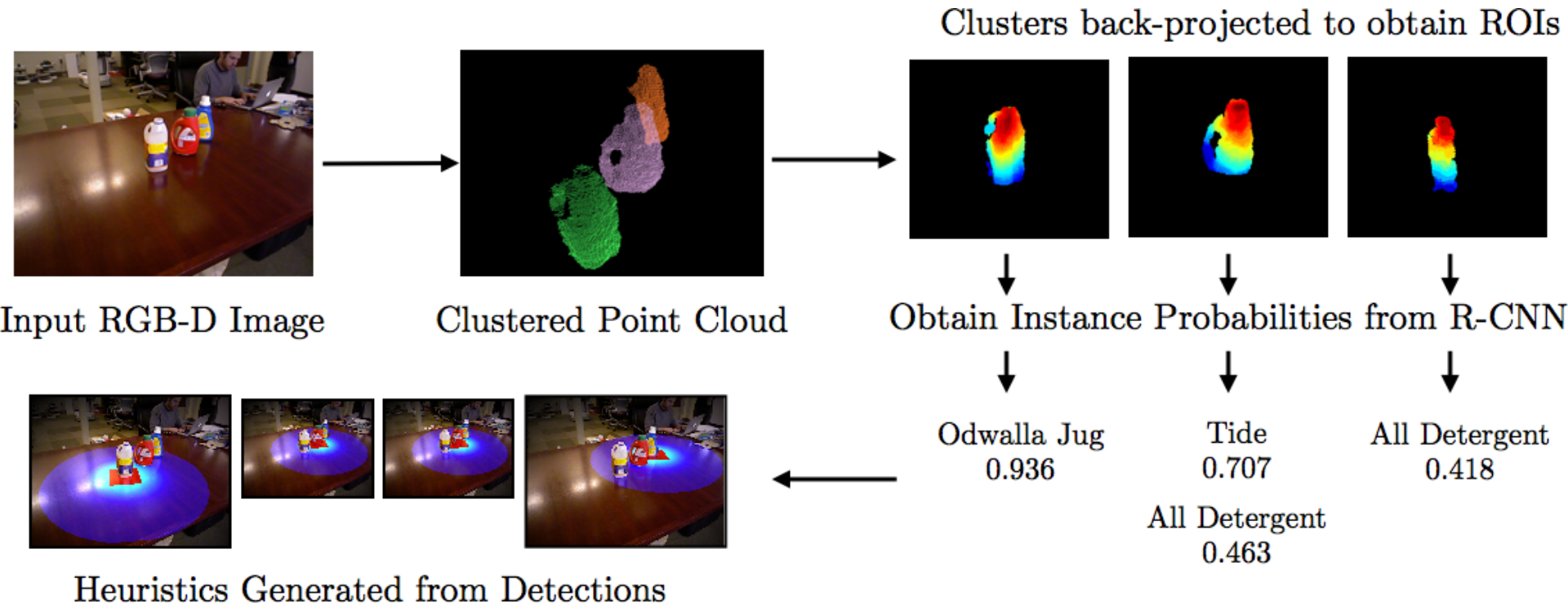




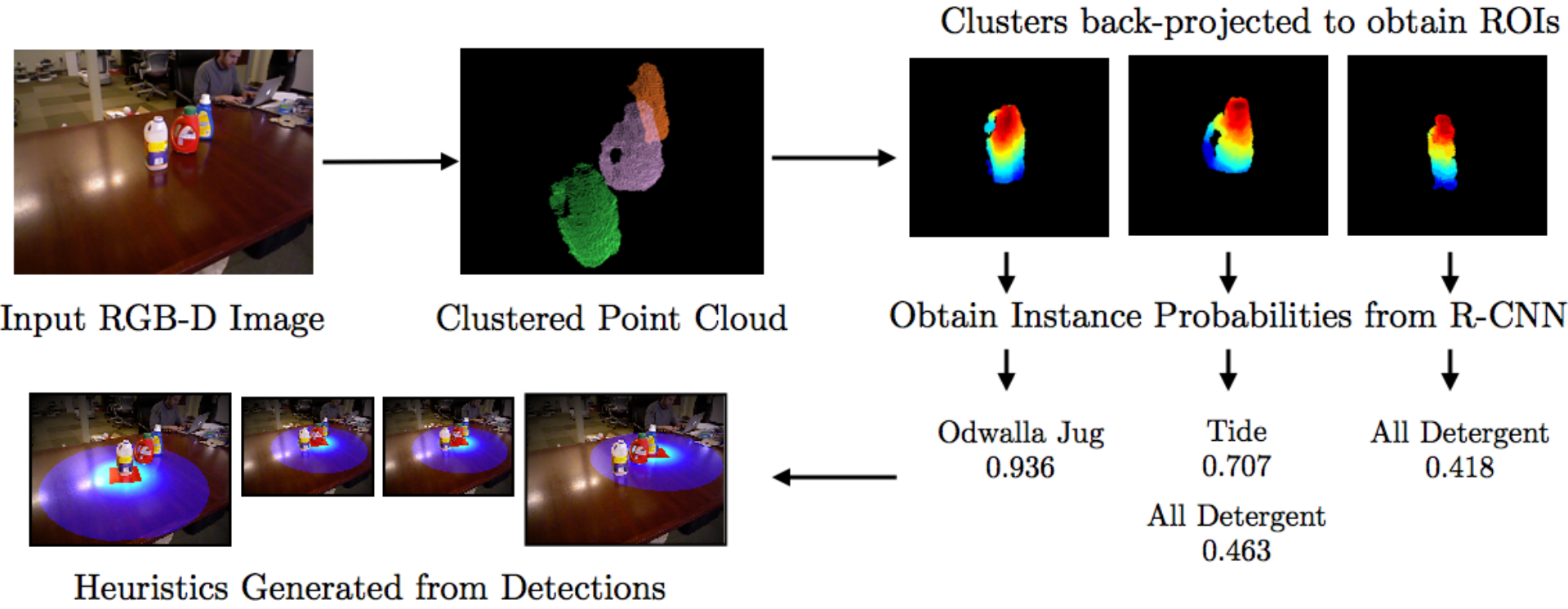
Use discriminative learners as heuristics to guide global deliberative search

D2P, Narayanan & Likhachev, RSS ’16

Discriminatively-guided Deliberative Perception (D2P)



Discriminatively-guided Deliberative Perception (D2P)



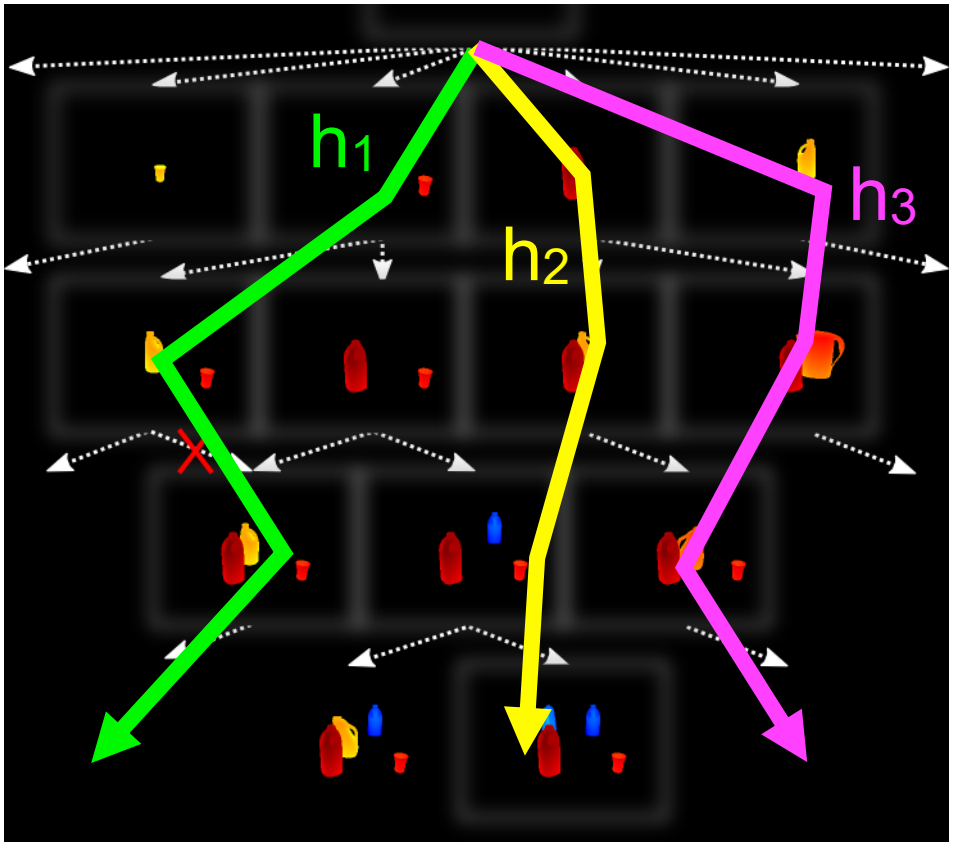
For each hypothesis ‘i’

$heuristic_i(s)$ = distance between predicted and hallucinated location

D2P, Narayanan & Likhachev, RSS ’16

Discriminatively-guided Deliberative Perception (D2P)

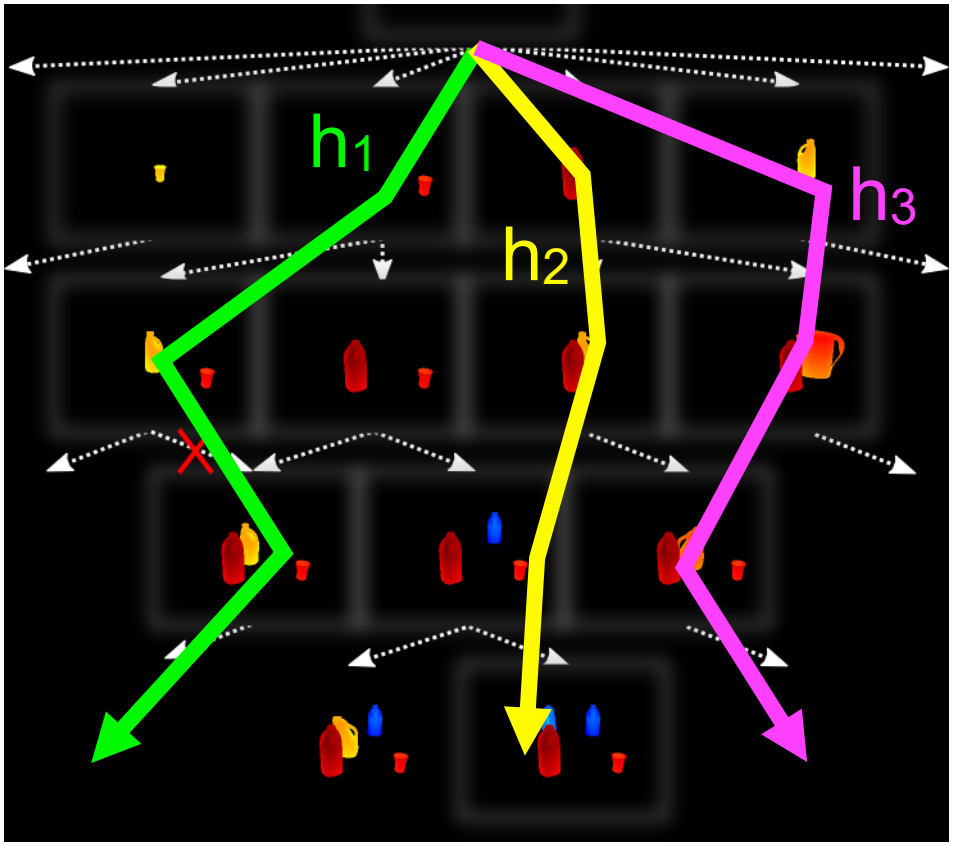
Heuristics are multi-modal hypotheses
Do not want to combine into “one” heuristic



Discriminatively-guided Deliberative Perception (D2P)

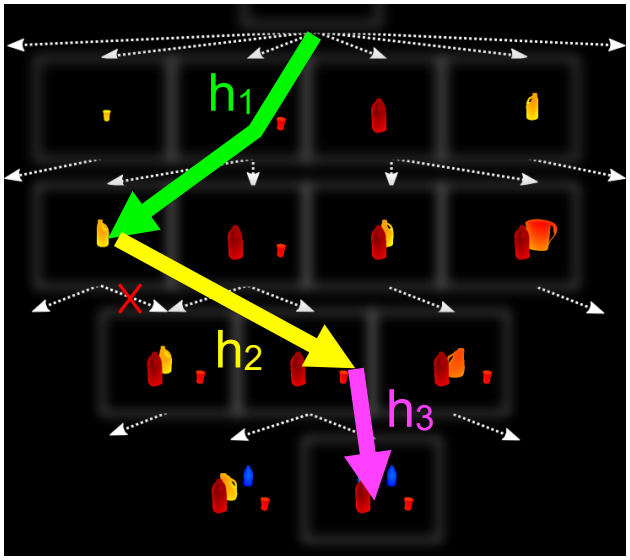
Heuristics are multi-modal hypotheses

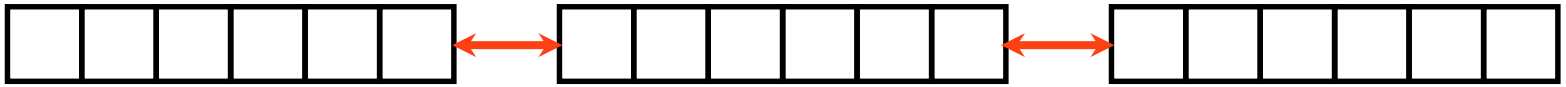
Do not want to combine into “one” heuristic



What else can we use?

Using Multiple Heuristics: MHA*



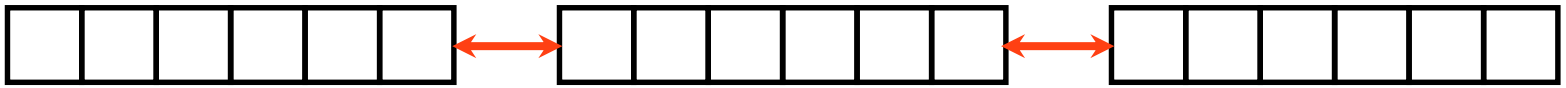


g(s)+w·h1(s) g(s)+w·h2(s) g(s)+w·h3(s)

shared g-values

Using Multiple Heuristics: MHA*

$OPEN_1 = OPEN_2 = OPEN_3 = \{s_{start}\}$
while s_{goal} not expanded
  for i in 1 to 3
    expand from $OPEN_i$



$g(s)+w \cdot h_1(s)$ $g(s)+w \cdot h_2(s)$ $g(s)+w \cdot h_3(s)$

shared g-values

Deliberative Perception: Properties

Guarantees on solution quality

$$\min_{O_{1:K}} J(O_{1:K})$$

$$J(O_{1:K}) = J_{observed}(O_{1:K}) + J_{rendered}(O_{1:K})$$

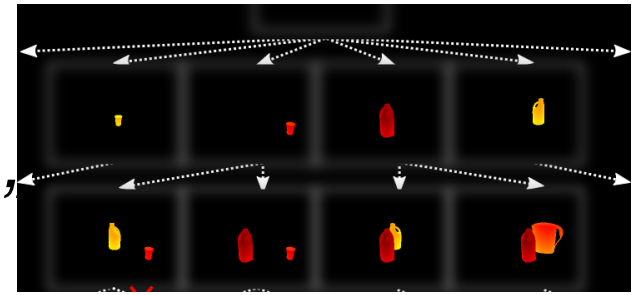
Deliberative Perception: Properties

Guarantees on solution quality

$$\min_{O_{1:K}} J(O_{1:K})$$

$$J(O_{1:K}) = J_{observed}(O_{1:K}) + J_{rendered}(O_{1:K})$$

Introduce notion of "completeness" for multi-object instance recognition



Implementation Details

Cost Computation

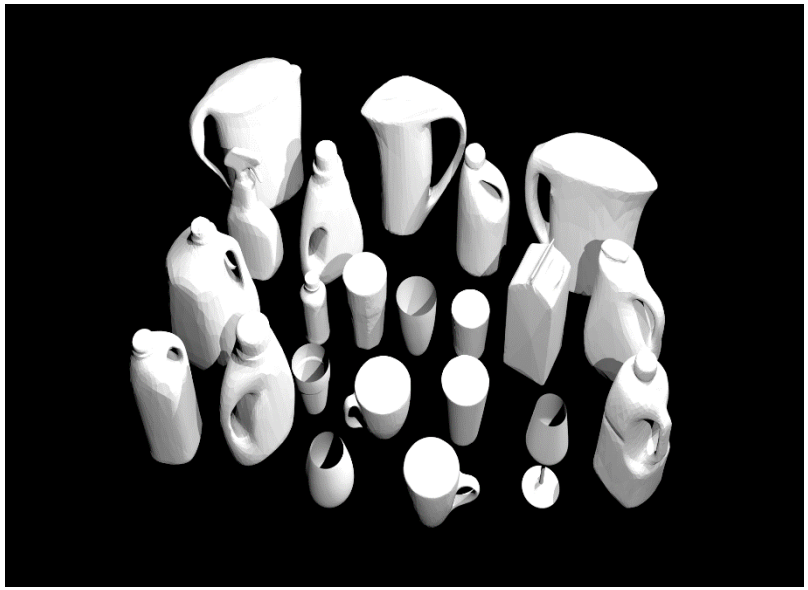
- ICP to eliminate discretization artifacts
- Cache depth images

Search

- Lazy edge evaluation
- Parallelize successor generation

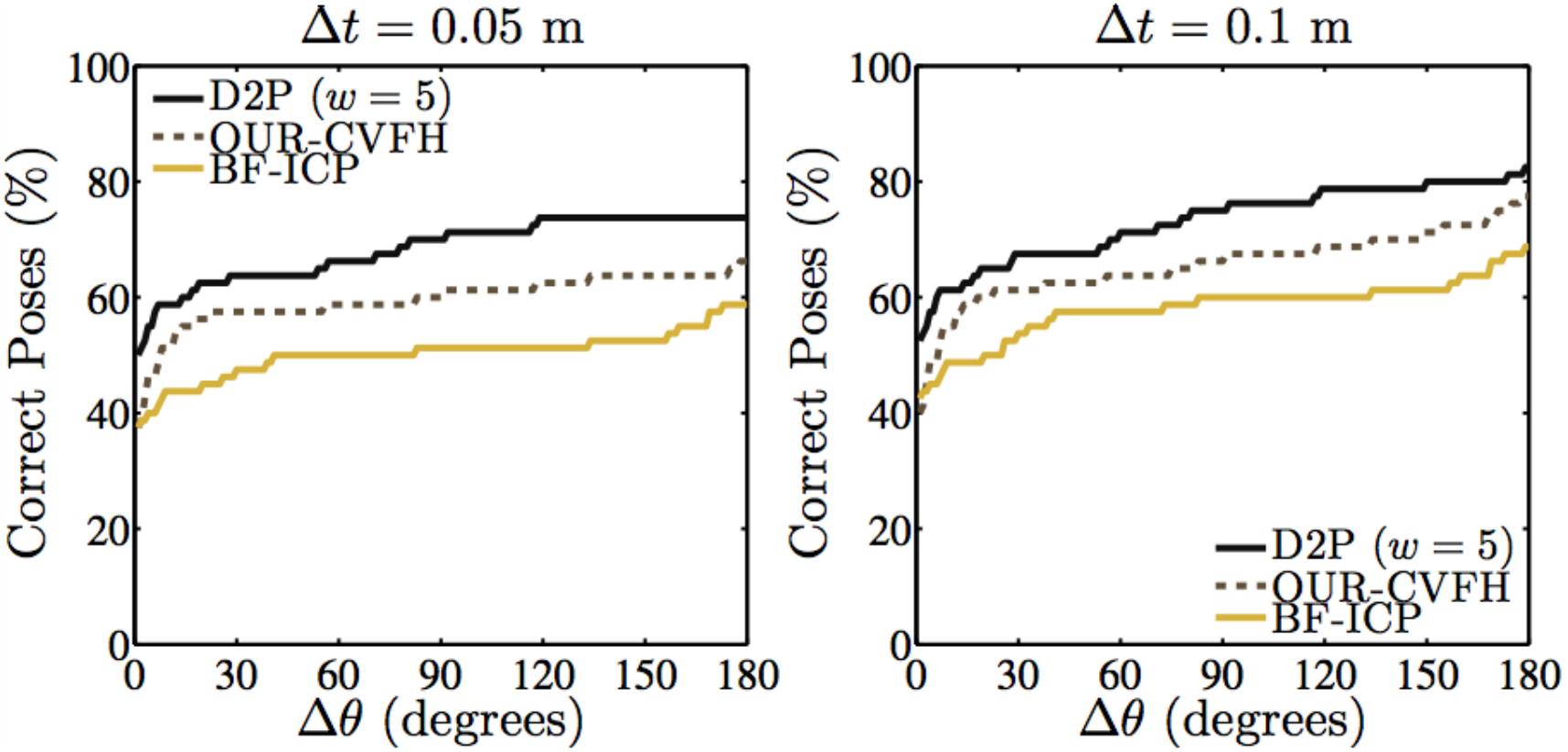
Dataset

- Household objects occlusion dataset
- 36 objects models
- 82 object instances in 23 scenes
- Ground truth pose for all objects available



[Point Cloud Library, Aldoma et al., IEEE RAM ’12]

Experiments: Baseline Comparions



Baselines: OUR-CVFH, Brute-force ICP (without rendering)

Summary

- Deliberative Perception: search for best “explanation” of the observed scene
- Discriminative guidance from *any* and *multiple* statistical learners
- Multi-Heuristic A* (MHA*) for graph search with distinct hypotheses
- Theoretical guarantees on solution quality
- Notion of completeness for multi-object pose estimation

Bibliography

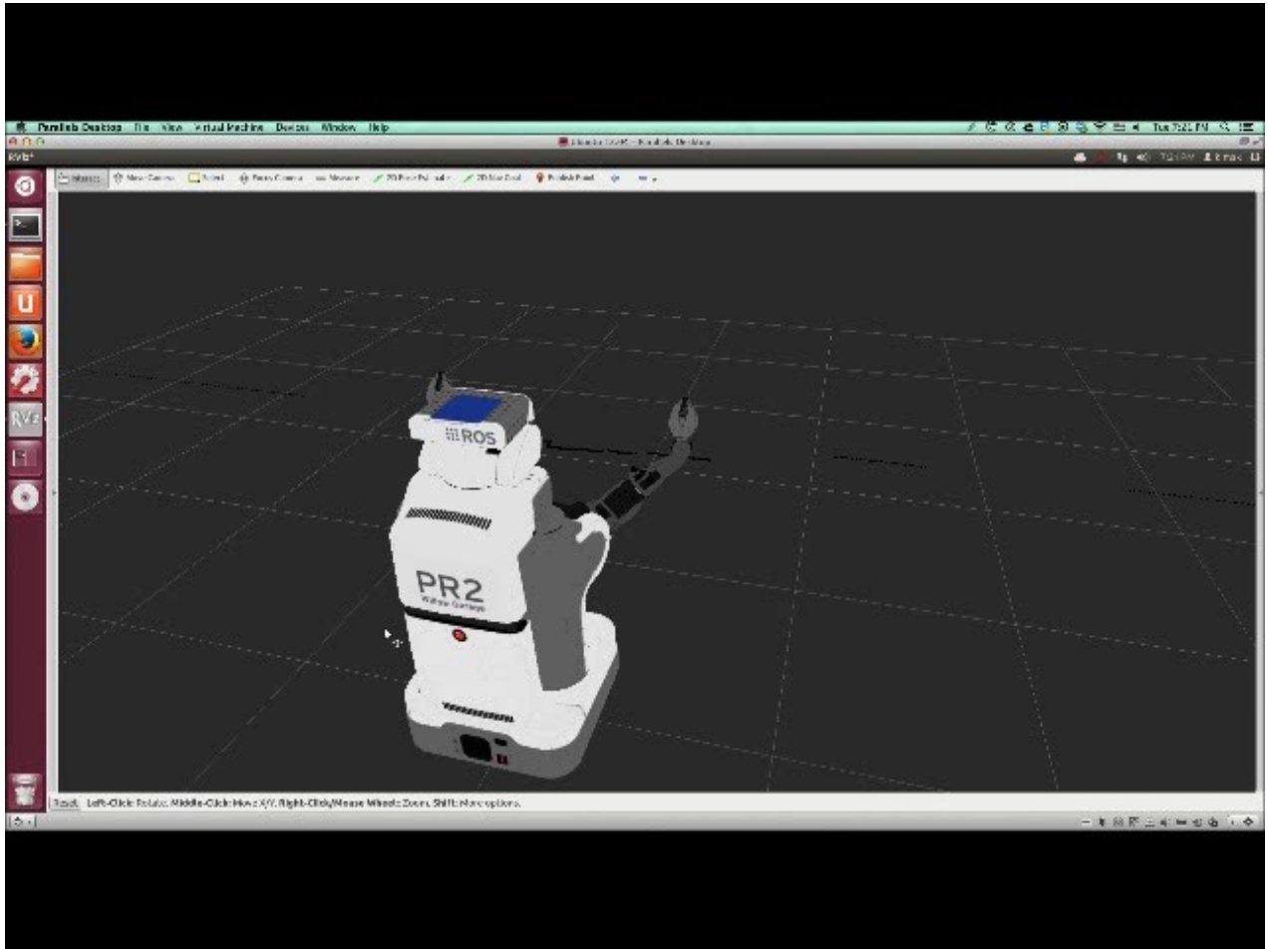
- PERCH: Perception via Search, Narayanan and Likhachev, ICRA ’16
- Discriminatively-guided Deliberative Perception (D2P), Narayanan and Likhachev, RSS ’16
- Multi-Heuristic A*, Aine, Swaminathan, Narayanan, Hwang, and Likhachev, RSS ’14, IJRR ’16
- Improved Multi-Heuristic A*, Narayanan, Aine, and Likhachev, SoCS ’15
- Graph Search with Edge Existence Priors, Narayanan and Likhachev, AAAI ’17 (submitted)
- Deliberative Object Pose Estimation in Clutter, Narayanan and Likhachev, ICRA ’17 (submitted)

Two Examples

- Graph search for perception
- Planning for Active Perception

Active Perception

- What should be a sequence of viewpoints to get full certainty about the object of interest?



Movie generated by S. Kim

Active Perception

- What should be a sequence of viewpoints to get full certainty about the object of interest?

Movie generated by S. Kim

Summary

- Techniques for planning are a form of optimization and are often applicable beyond pure planning