Autoscanning for Coupled Scene Reconstruction and Proactive Object Analysis

Kai Xu, Hui Huang, Yifei Shi, Hao Li,

Pinxin Long, Jianong Caichen,

Wei Sun, Baoquan Chen

Shenzhen VisuCA Key Lab / SIAT

National University of Defense Technology

Shandong University

SIGGRAPH ASIA 2015 KOBE

sored by

Outline

- Background & motivation
- Method
- Results & evaluation
- Conclusion
- Future works

Background & motivation

- Fast development of
  - Commodity depth camera
  - ***Real-time reconstruction***

KinectFusion [Nießner et al. 2013]

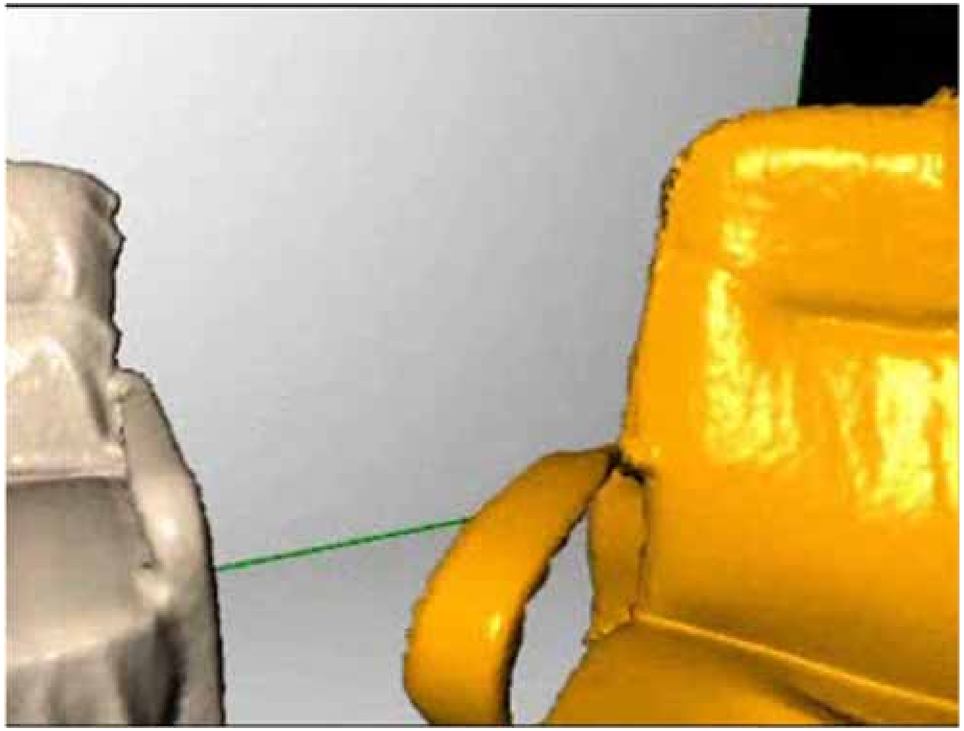
Microsoft Kinect

Background & motivation

- ***Real-time reconstruction*** allows for ...

Instant visual feedback

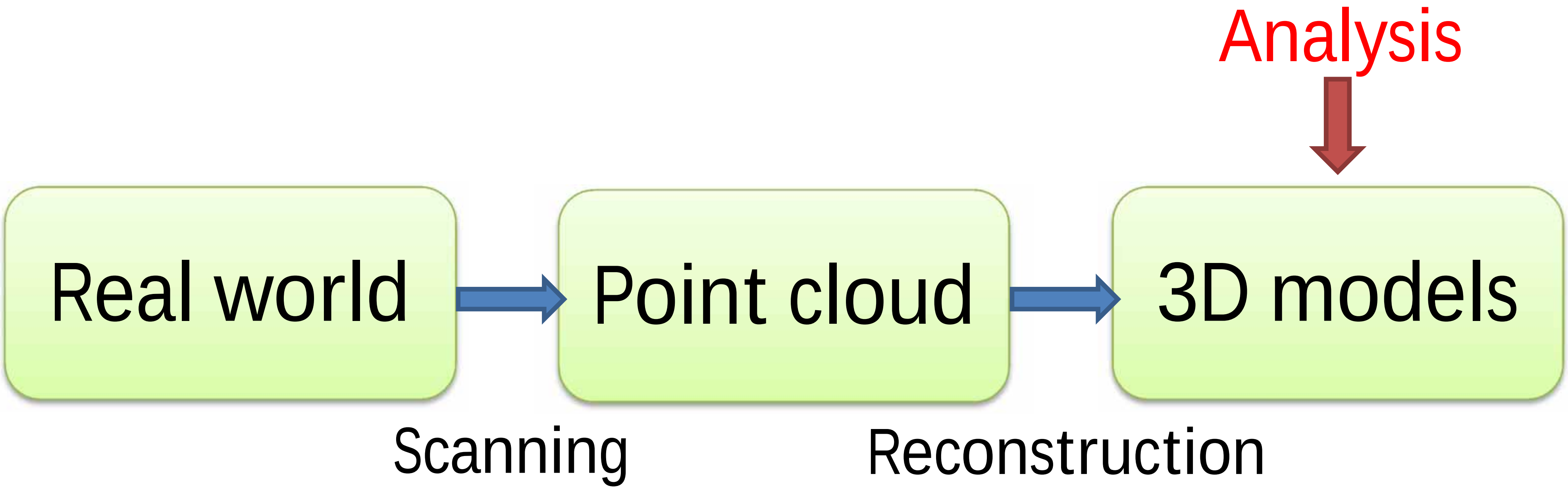
Online structural analysis [Zhang et al. 2014]



Online analysis!!

Background & motivation

- What difference does it make?



Background & motivation

- Why is it so exciting?

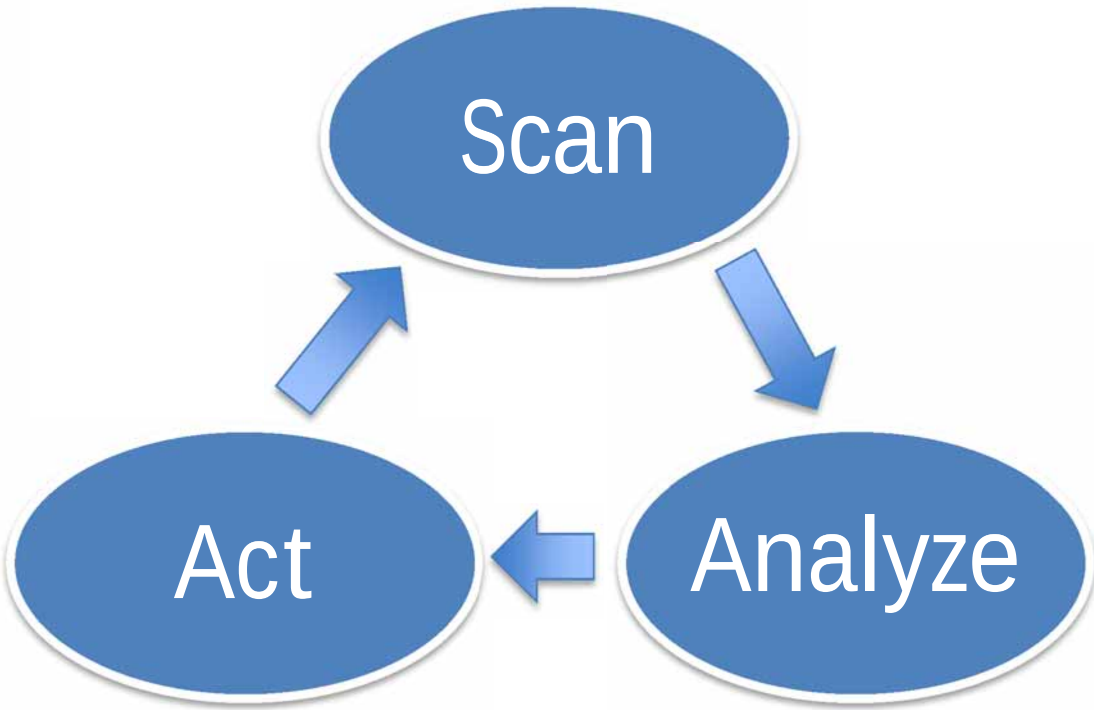
Online analysis guides autonomous scanning !!

- Online analysis tells the robot ***where to scan*** and ***when to stop***.
- Analysis will benefit from more and more complete reconstruction.

Background & motivation

- What's more exciting?

Online analysis guides autonomous scanning !!



Autoscanning replacing human scanning!

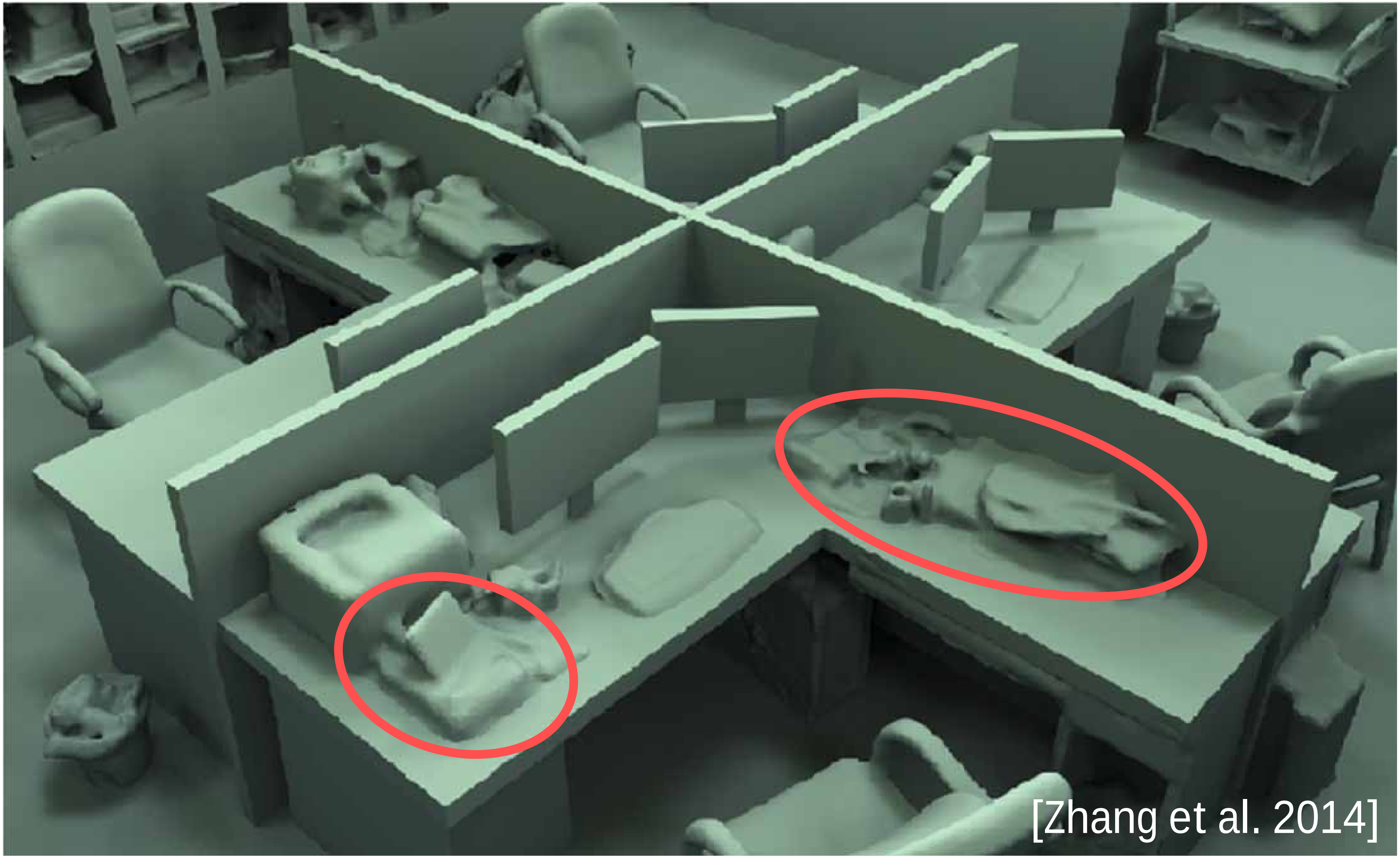
- What’s more exciting?

Online analysis guides autonomous scanning !!

Background & motivation

Background & motivation

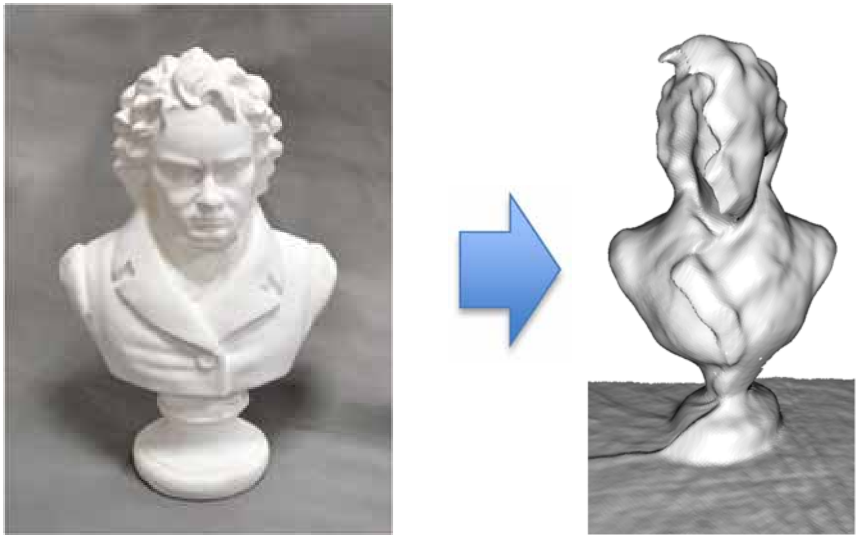
Background & motivation



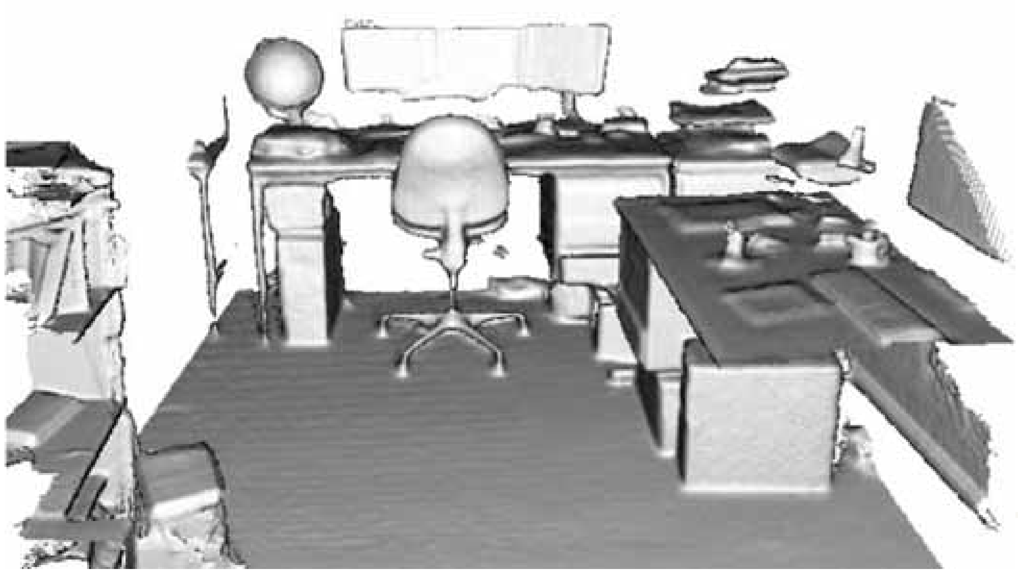
[Zhang et al. 2014]

Scene segmentation is not easy!

Low quality depth maps



Drifting issue with KinectFusion

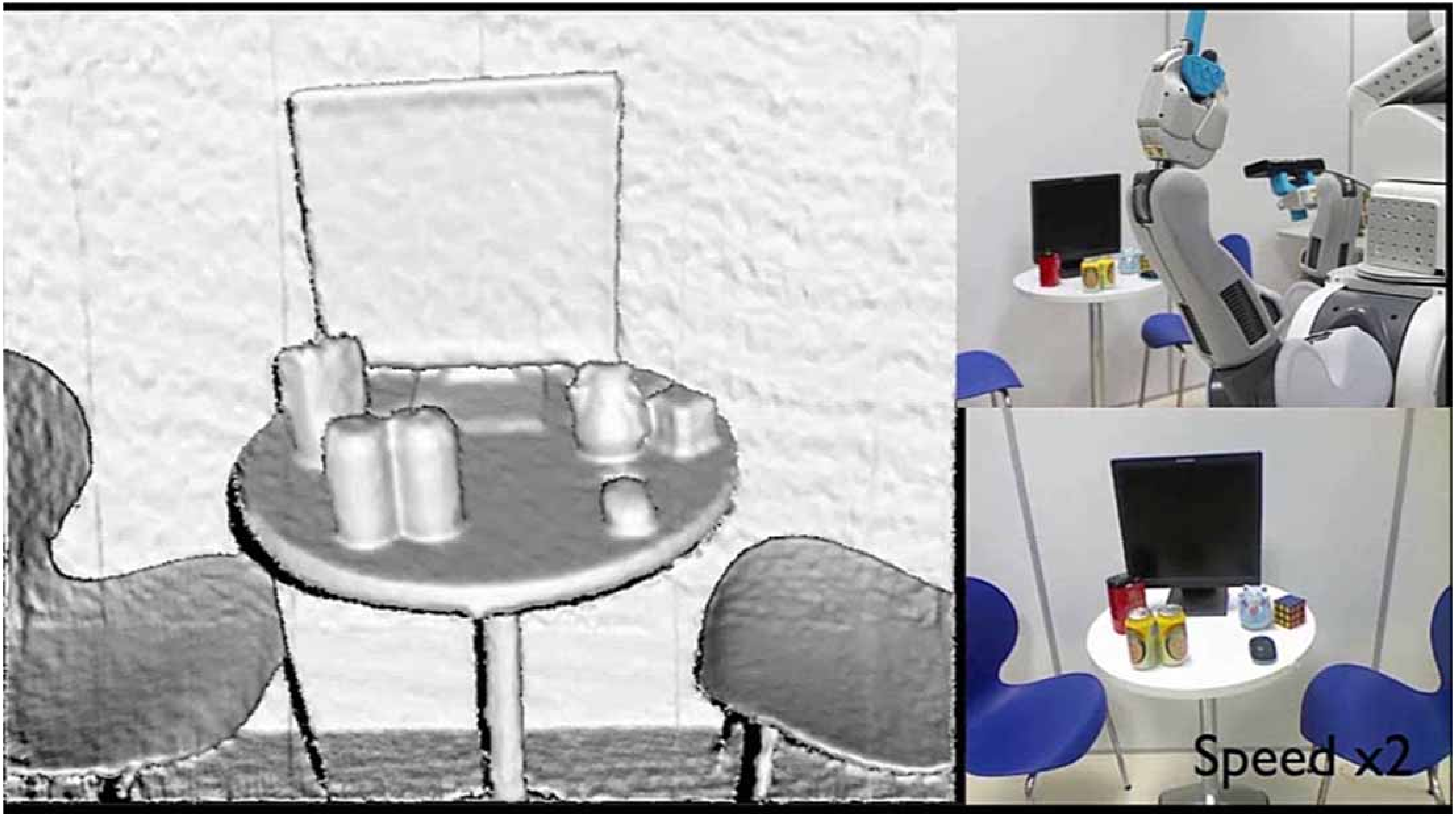


Objects in varying scales

Cluttered objects

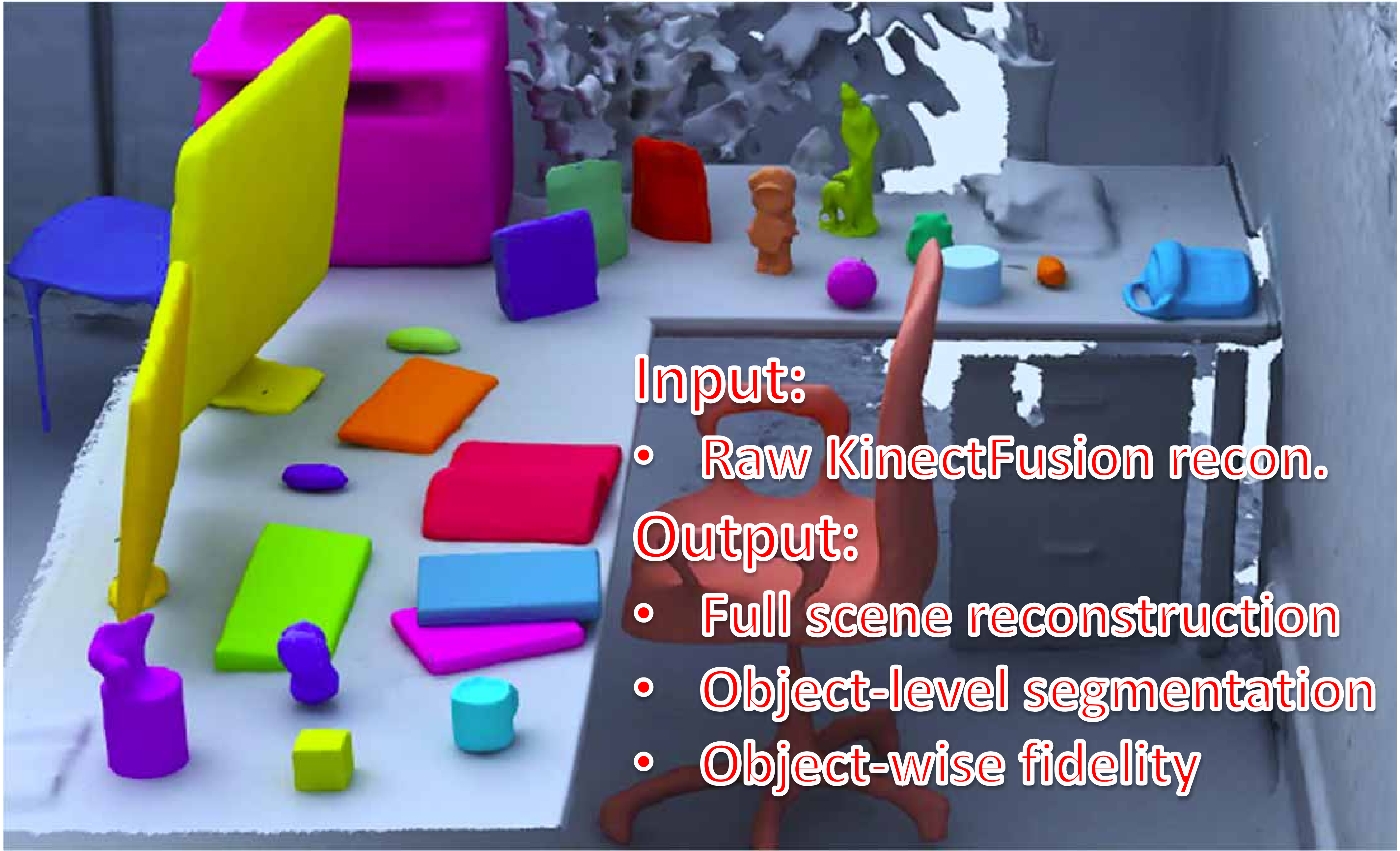

Our solution:
Proactive analysis
PR2

Push it!



Method

Problem statement



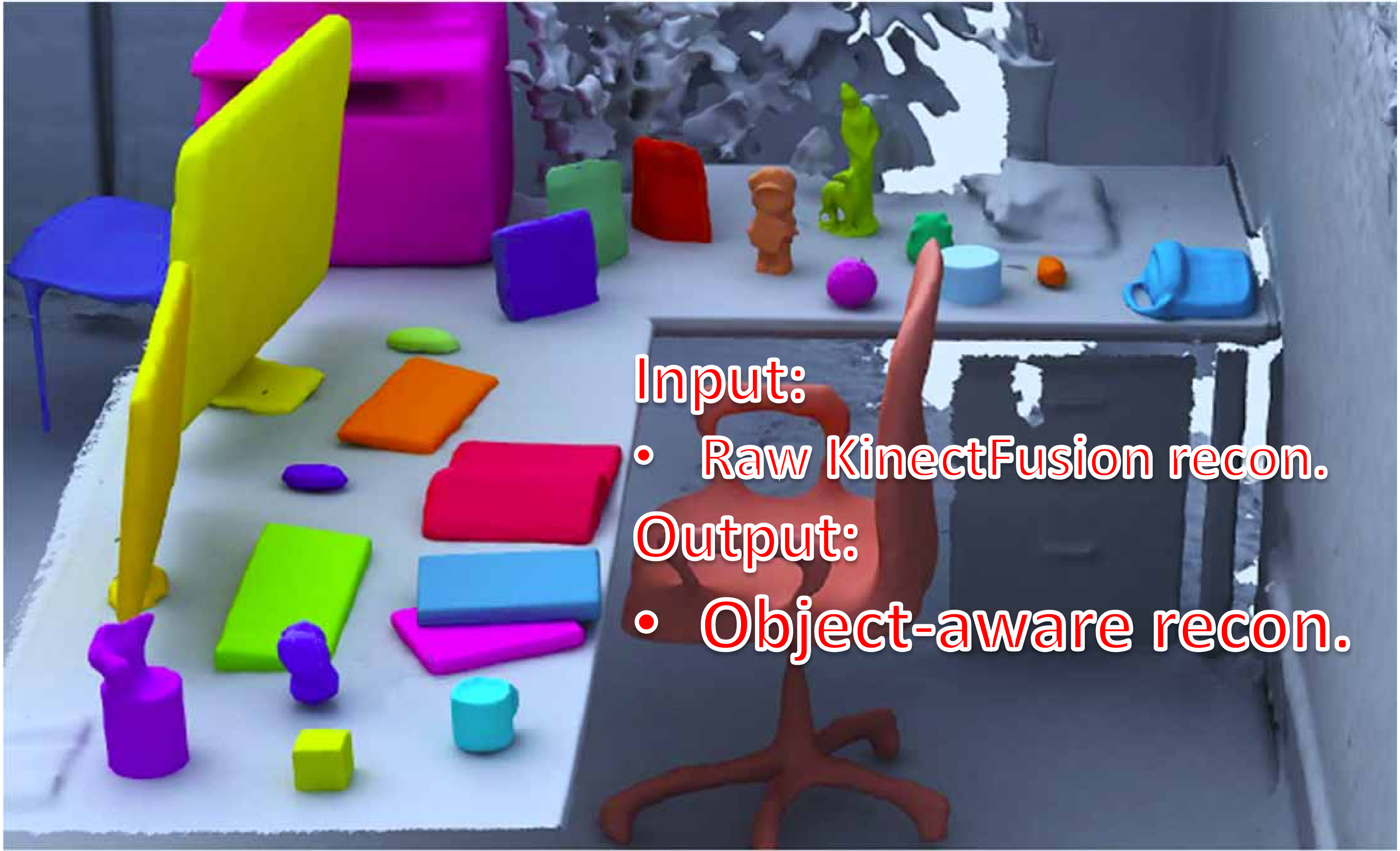
Input:

- Raw KinectFusion recon.

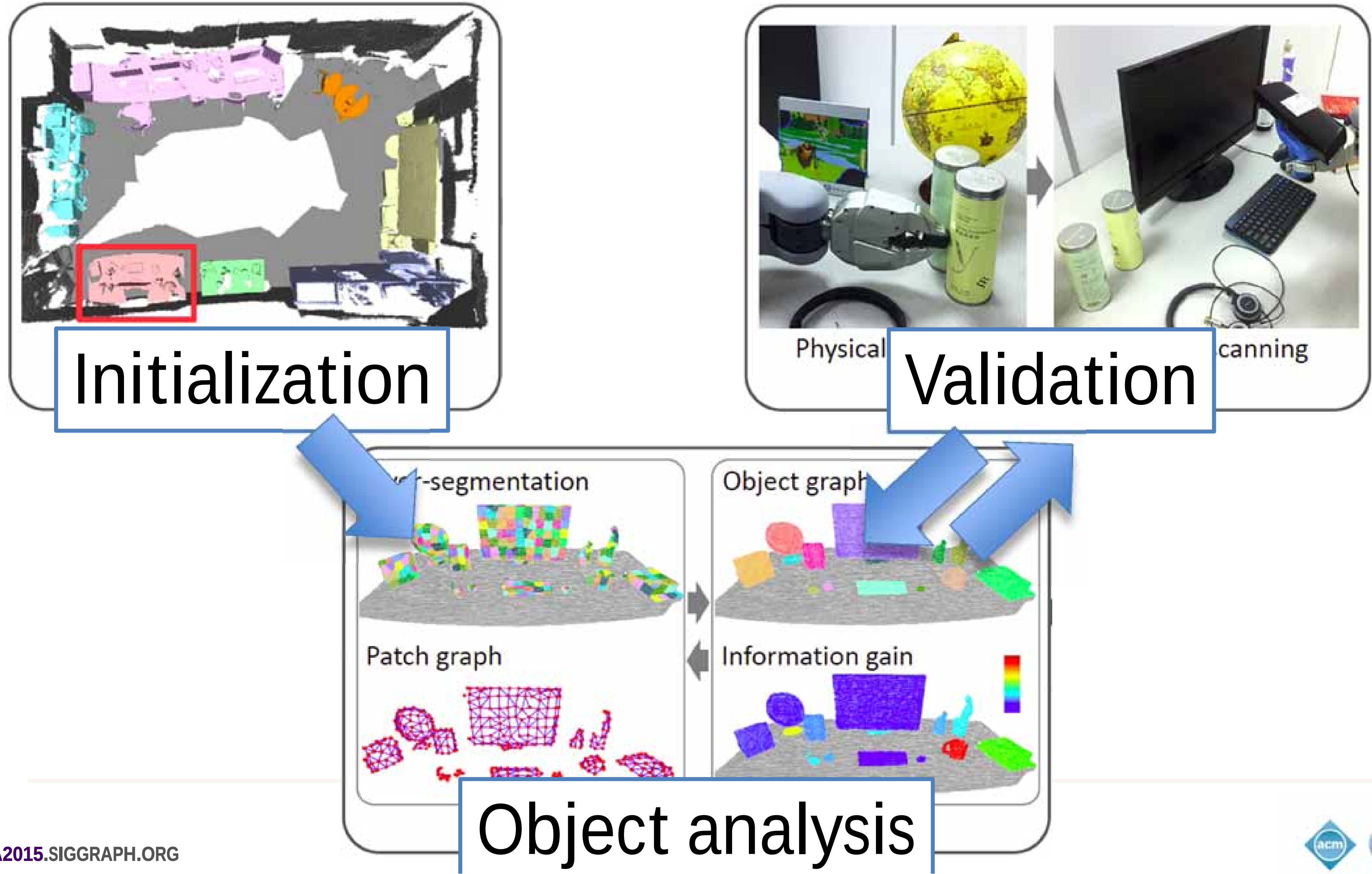
Output:

- Full scene reconstruction
- Object-level segmentation
- Object-wise fidelity

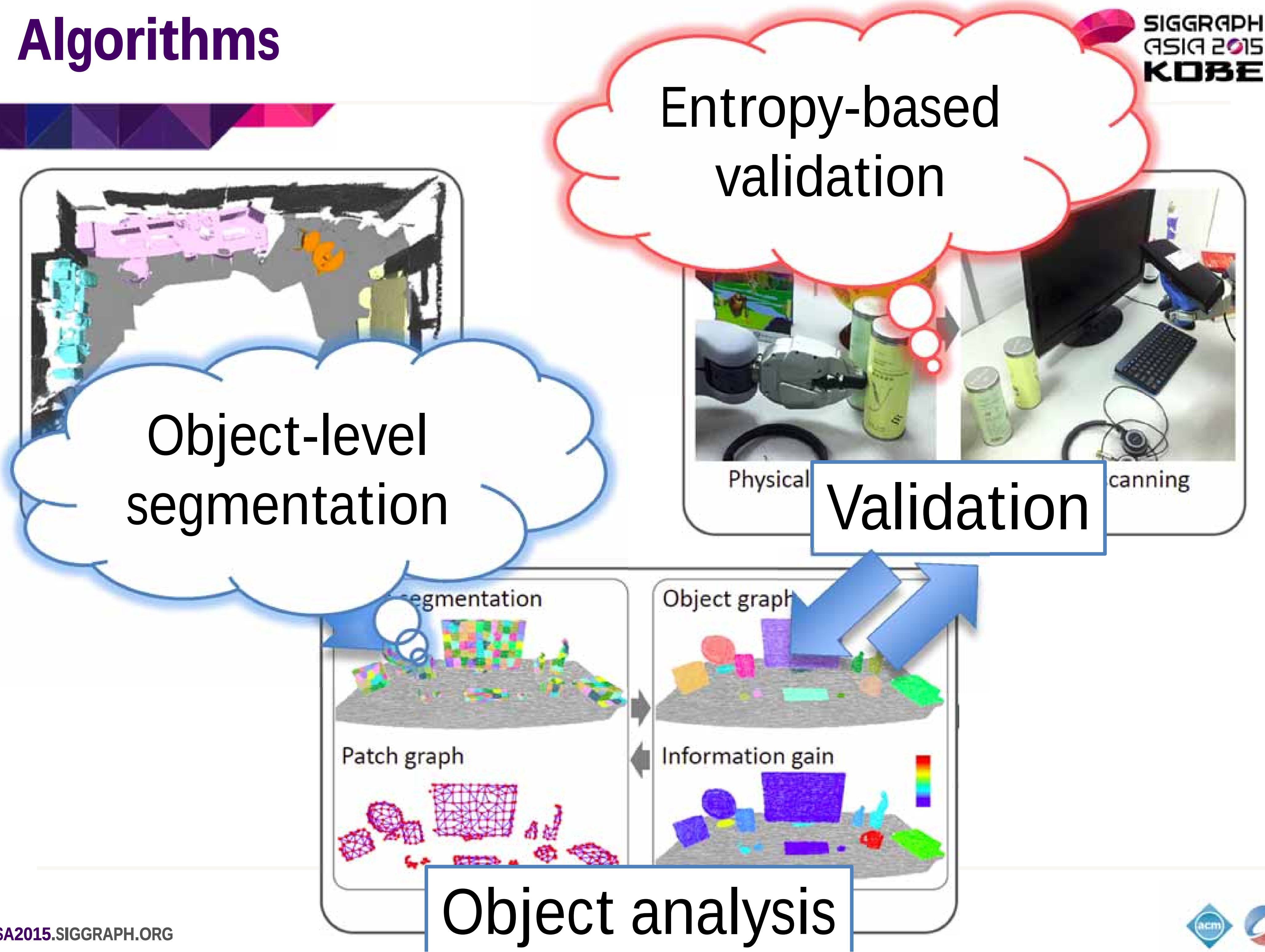
Problem statement



Method overview

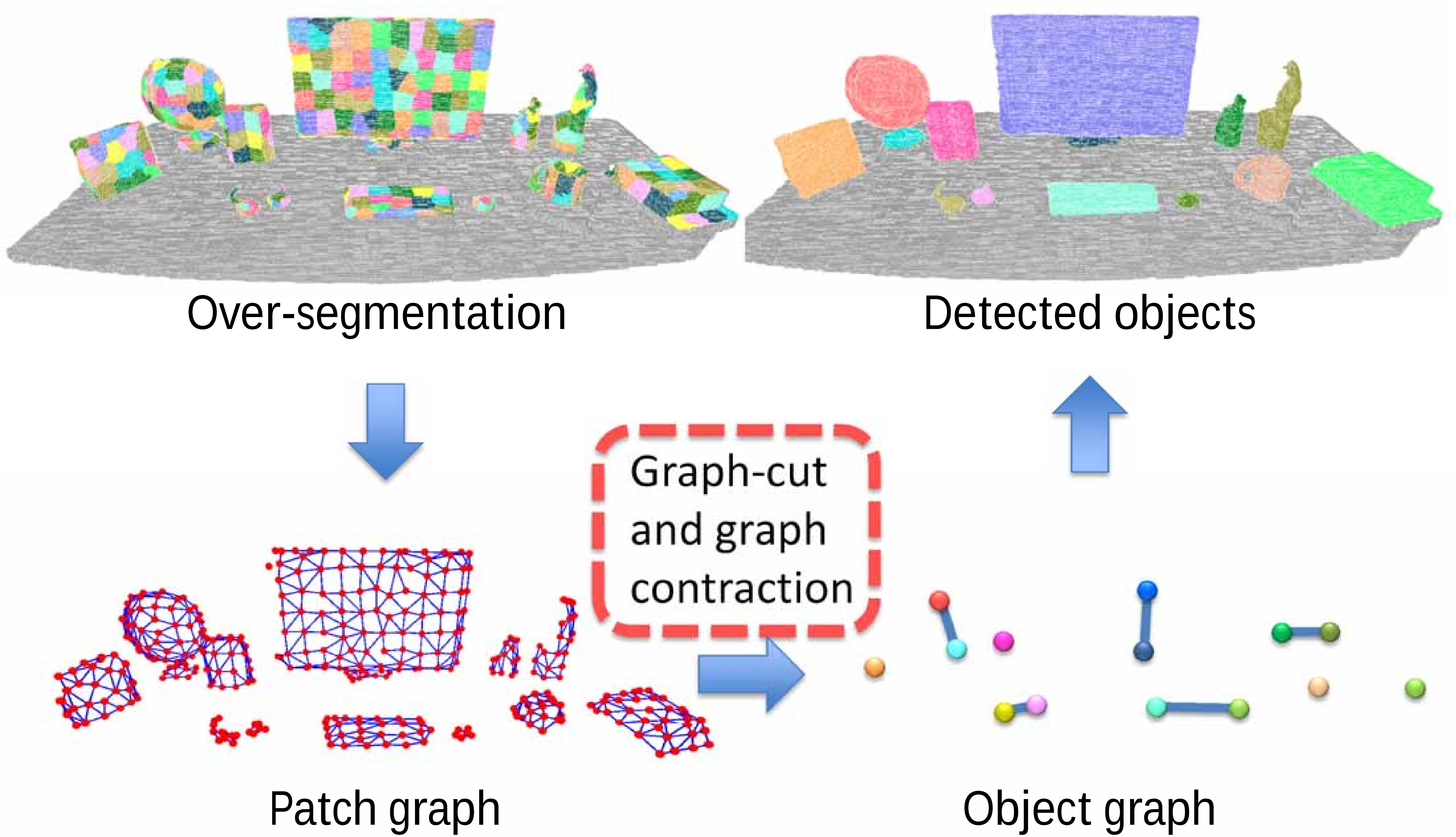


Algorithms


Entropy-based validation
Object-level segmentation
Physical
canning
Validation
segmentation
Object graph
Patch graph
Information gain
Object analysis

Algorithm I
Object-level segmentation

Object-level segmentation - Pipeline



Object-level segmentation

Generating object hypothesis

Multiple binary graph-cuts

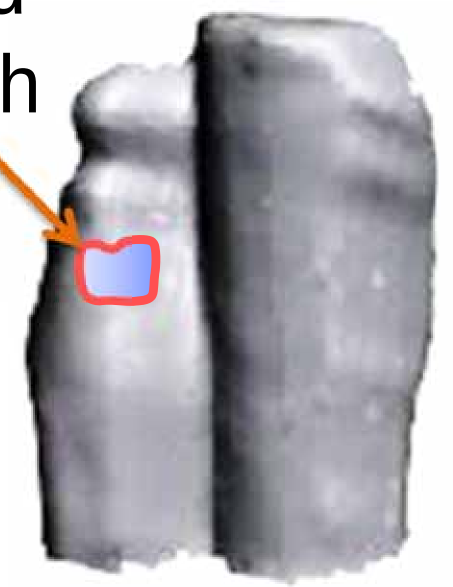
→

Selecting object hypothesis

One multi-class graph-cut

Generating object hypotheses

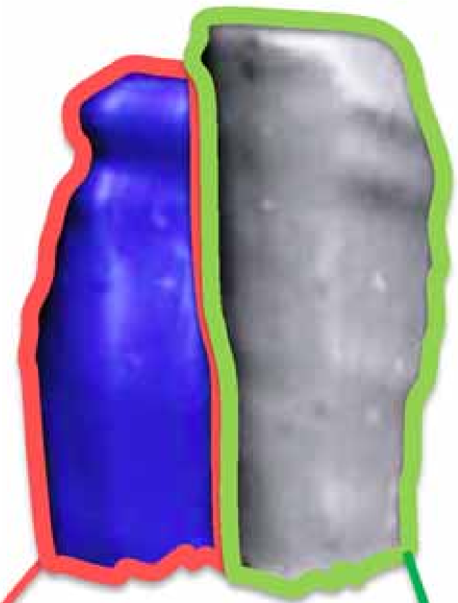
seed patch



Input scene

Binary graph-cut

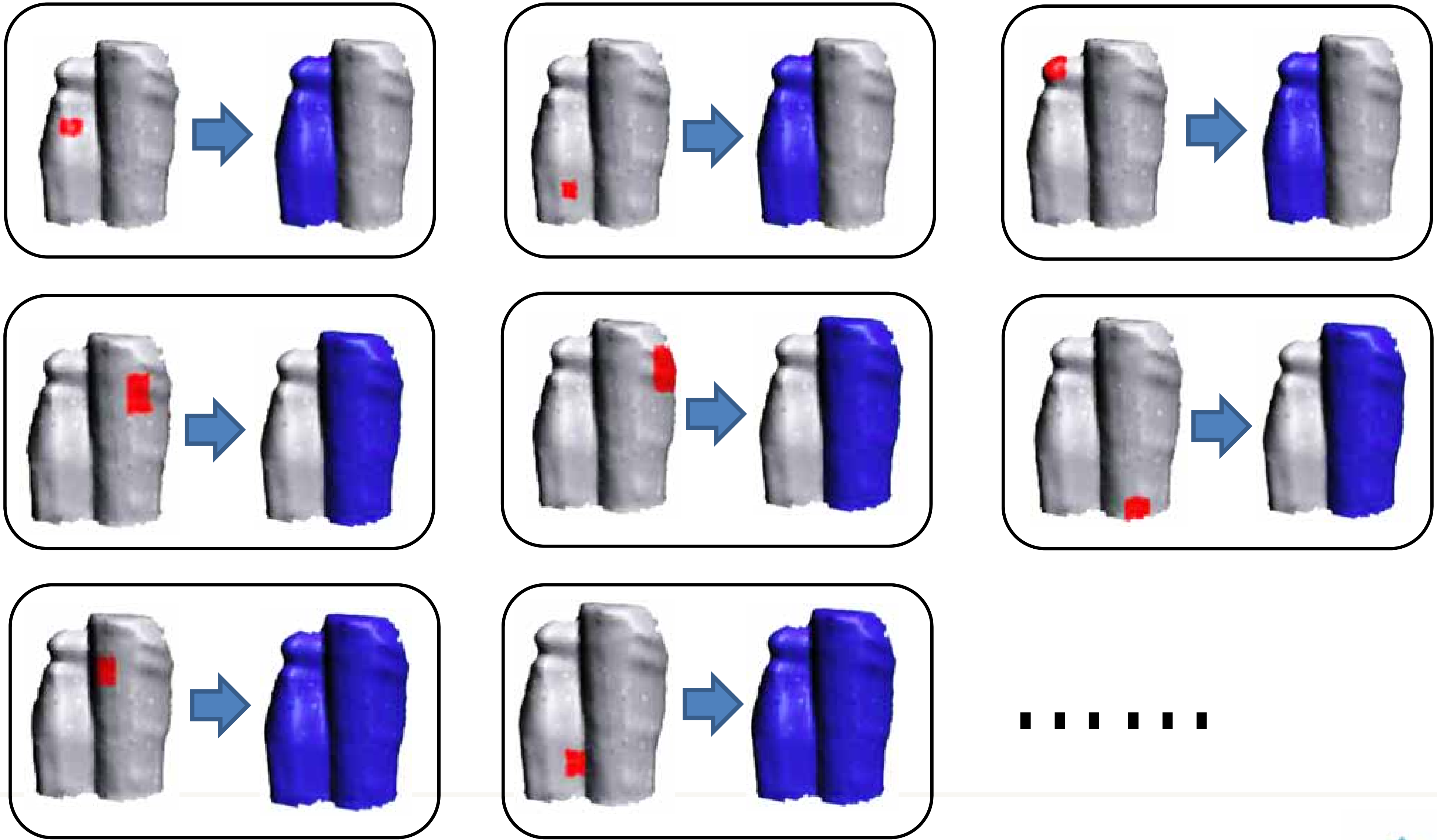
[Golovinskiy et al. 2009]



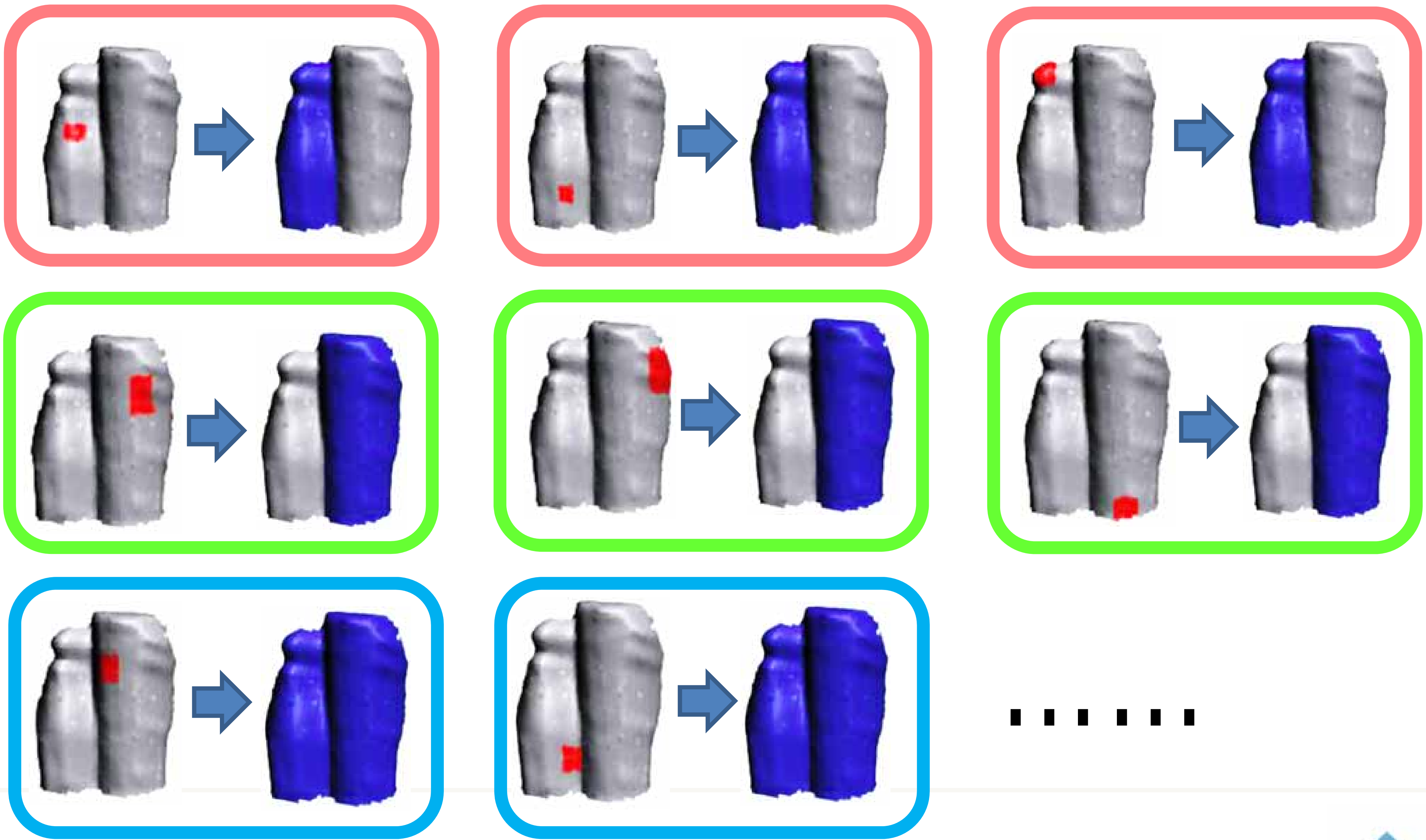
foreground background

a hypothetical object

Generating object hypotheses



Generating object hypotheses

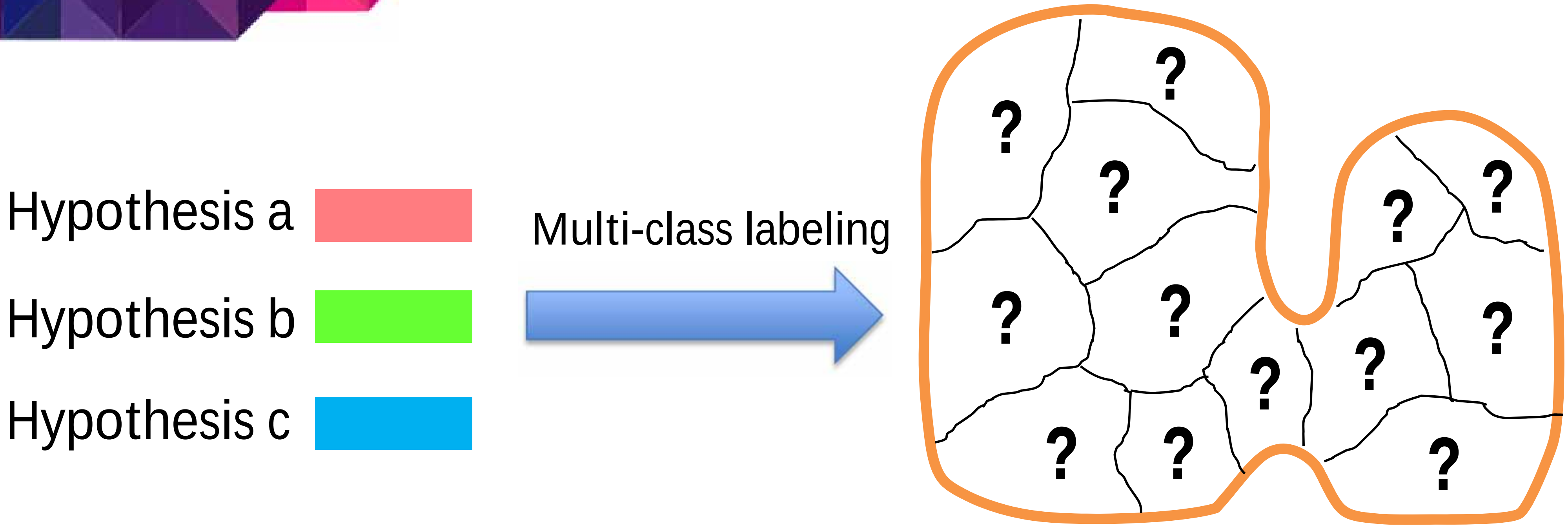


Generating object hypotheses

So we get three hypothetical objects:

The next step is to label the scene using the three labels above ...

Selecting object hypotheses



***Selecting** the most prominent hypotheses*
with a multi-class graph-cut

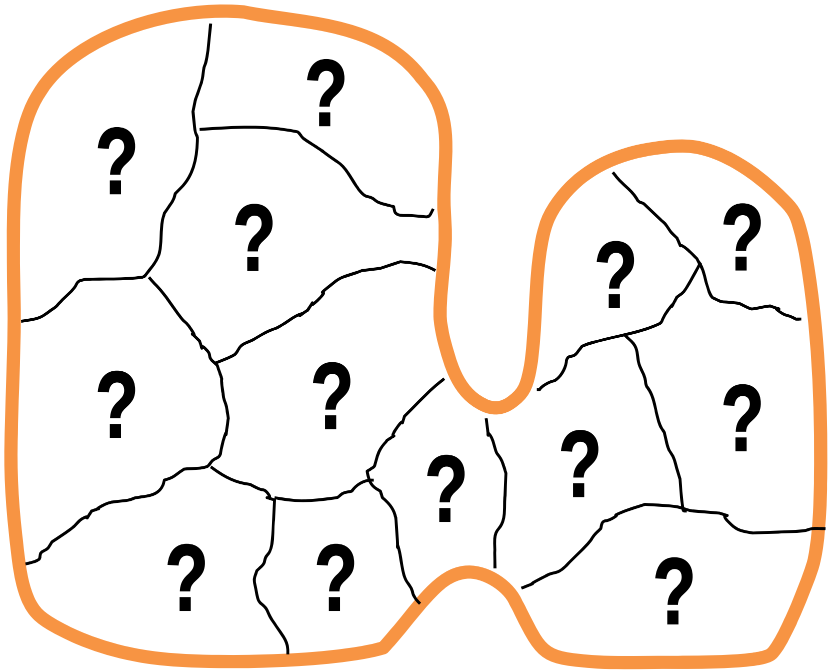
Selecting object hypotheses

Hypothesis a

Hypothesis b

Hypothesis c

Multi-class labeling



$$E(L) = \sum_{u \in \mathcal{V}_\mathrm{p}} E_d(l_u; P_u) + \sum_{(u,v) \in \mathcal{E}_\mathrm{p}} E_s(l_u, l_v)$$

multi-class graph-cut

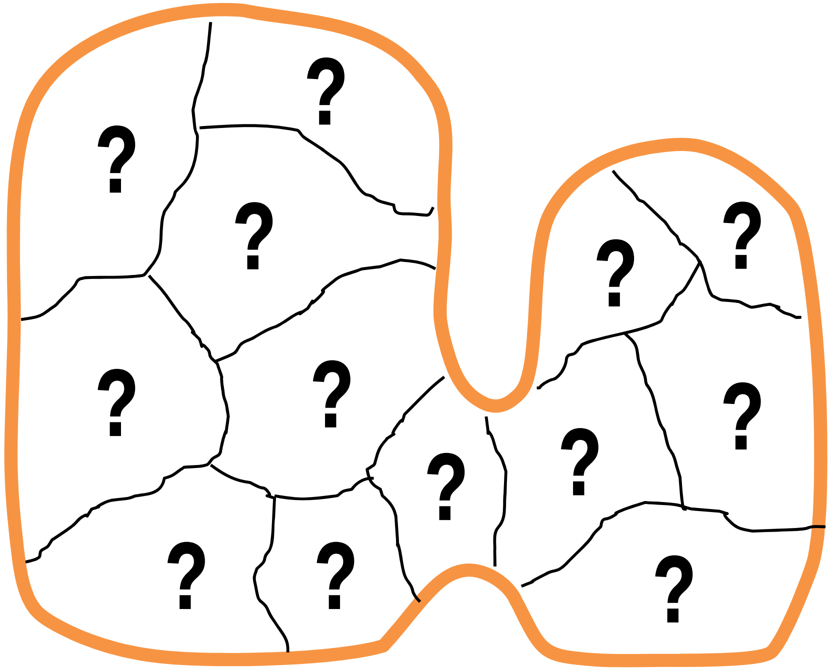
Selecting object hypotheses

Hypothesis a

Hypothesis b

Hypothesis c

Multi-class labeling



$$E(L)=\sum_{u\in\mathcal{V}_{\mathrm{p}}} E_d(l_u;P_u)+\sum_{(u,v)\in\mathcal{E}_{\mathrm{p}}} E_s(l_u,l_v)$$

$$E_d(l_u=\boldsymbol{a};P_u)=-\ln P(u\leftarrow \boldsymbol{a})$$

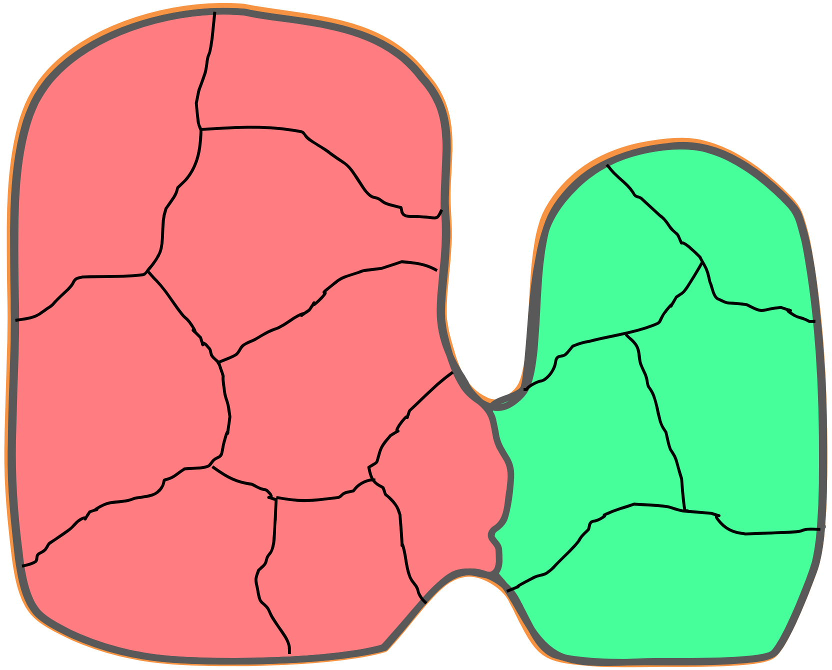
Selecting object hypotheses

Hypothesis a

Hypothesis b

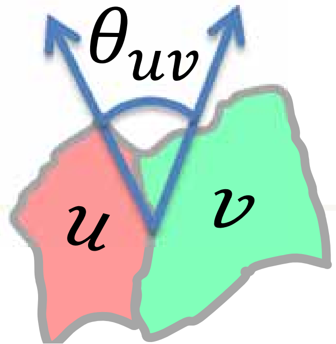
~~Hypothesis c~~

Multi-class labeling



$$E(L) = \sum_{u \in \mathcal{V}_p} E_d(l_u; P_u) + \sum_{(u,v) \in \mathcal{E}_p} \boxed{E_s(l_u, l_v)}$$

$$E_s(l_u \neq l_v) = \eta(1 - \cos(\theta_{uv}))$$

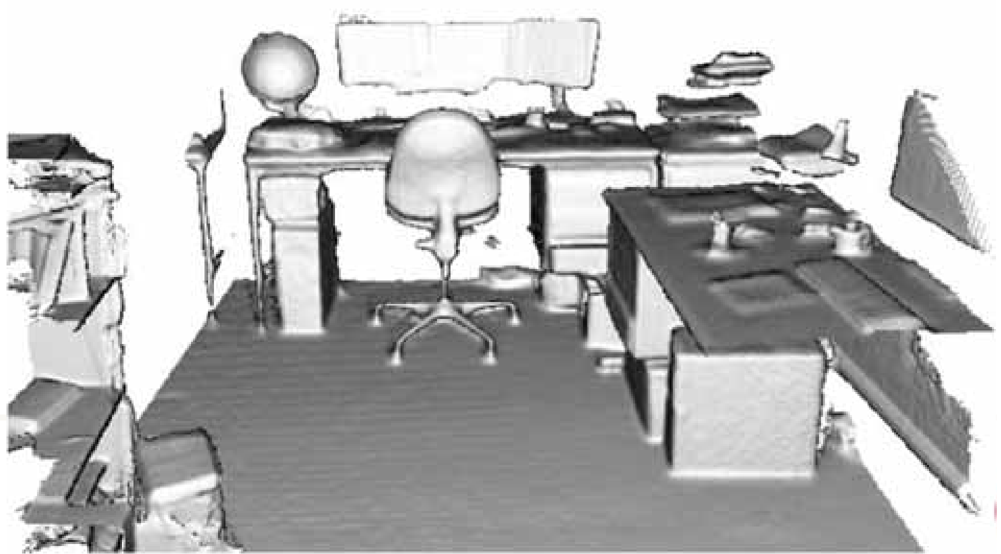


Object-level segmentation

Ours RANSAC

Learn segmentation from history ...

Can we learn from such case, and apply the learned knowledge in segmenting the rest part of the scene?



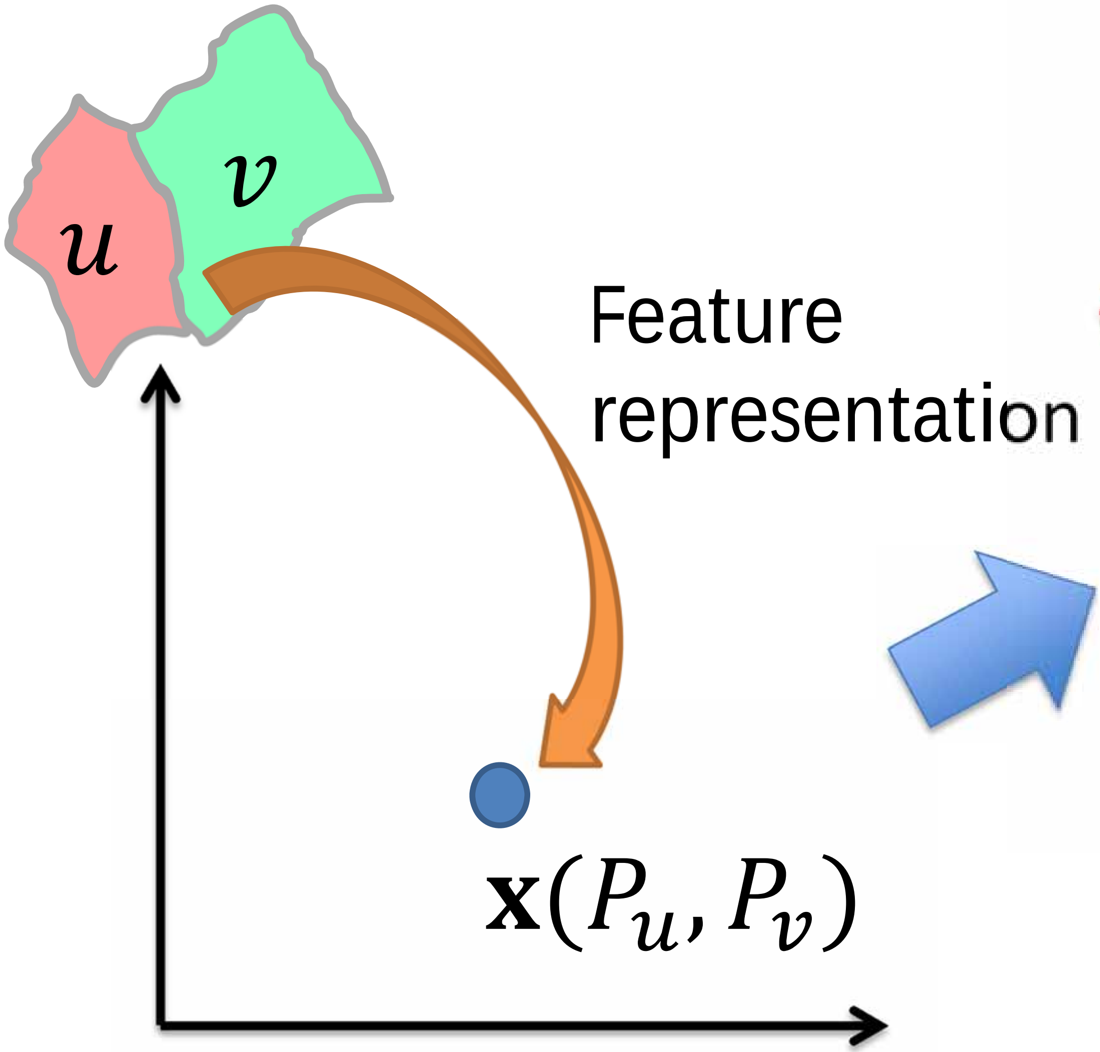
Online Learning of cut cost

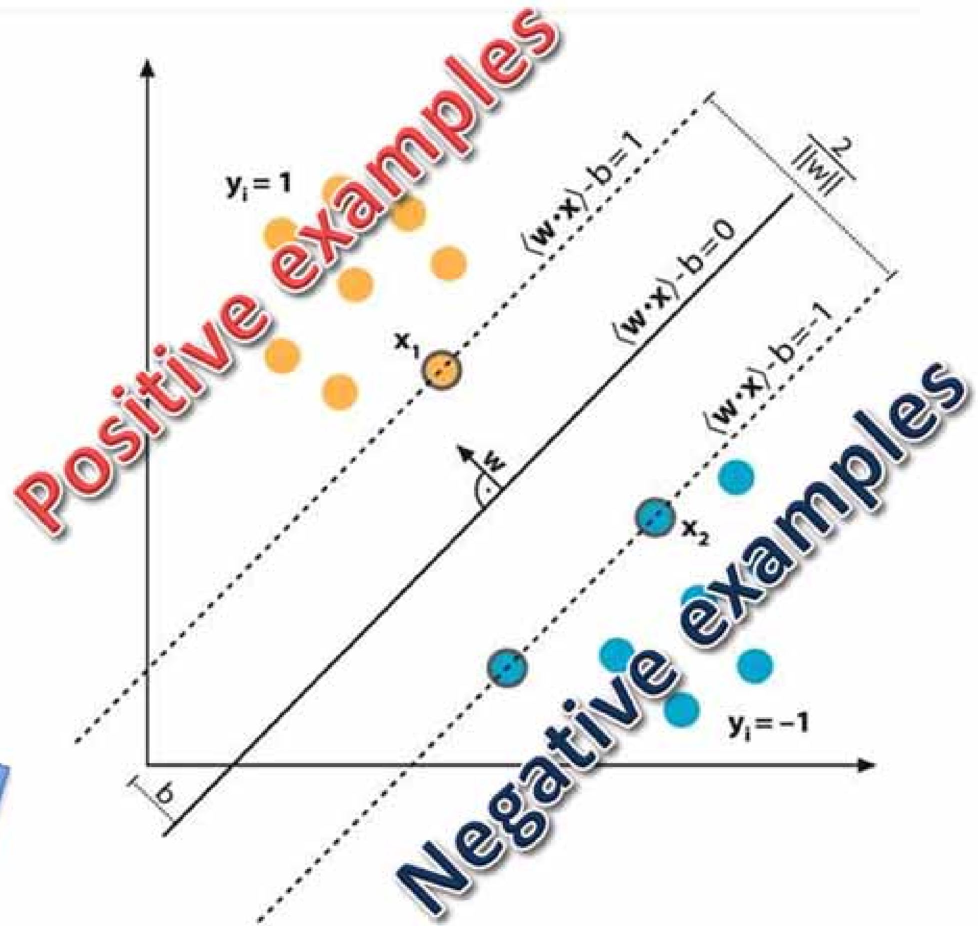
$$E(L) = \sum_{u \in \mathcal{V}_{\mathrm{p}}} E_d(l_u; P_u) + \sum_{(u,v) \in \mathcal{E}_{\mathrm{p}}} \boxed{E_s(l_u, l_v)}$$

$$E_s(l_u \neq l_v) = 1 - p(l_u \neq l_v | \mathbf{x}(P_u, P_v))$$

Predicted by learned SVM classifier

Online Learning of cut cost





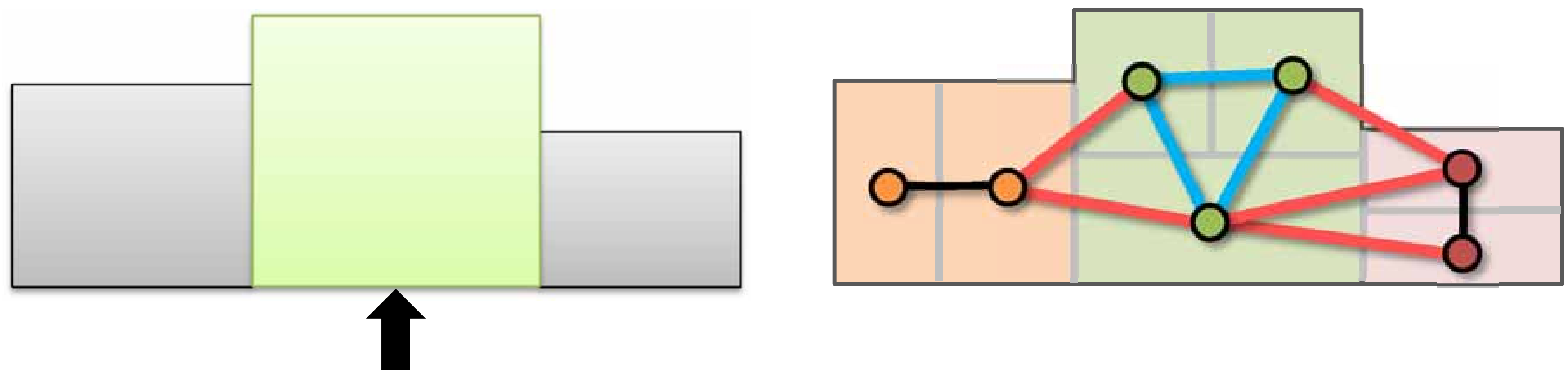
Linear classifier

Multiple Kernel Learning

Online Learning of cut cost

Collect training examples from pushes:

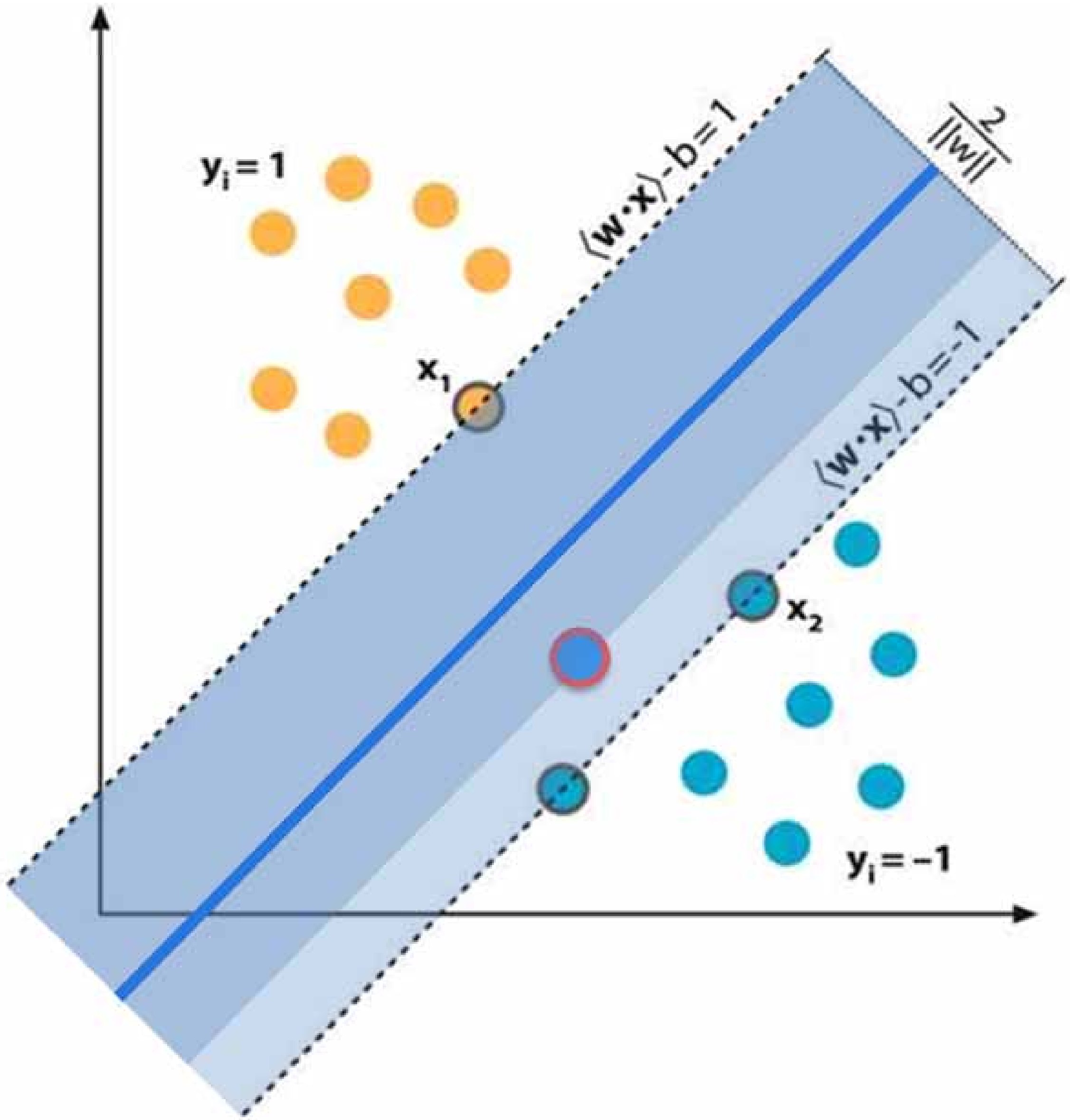
Positive examples: must cut



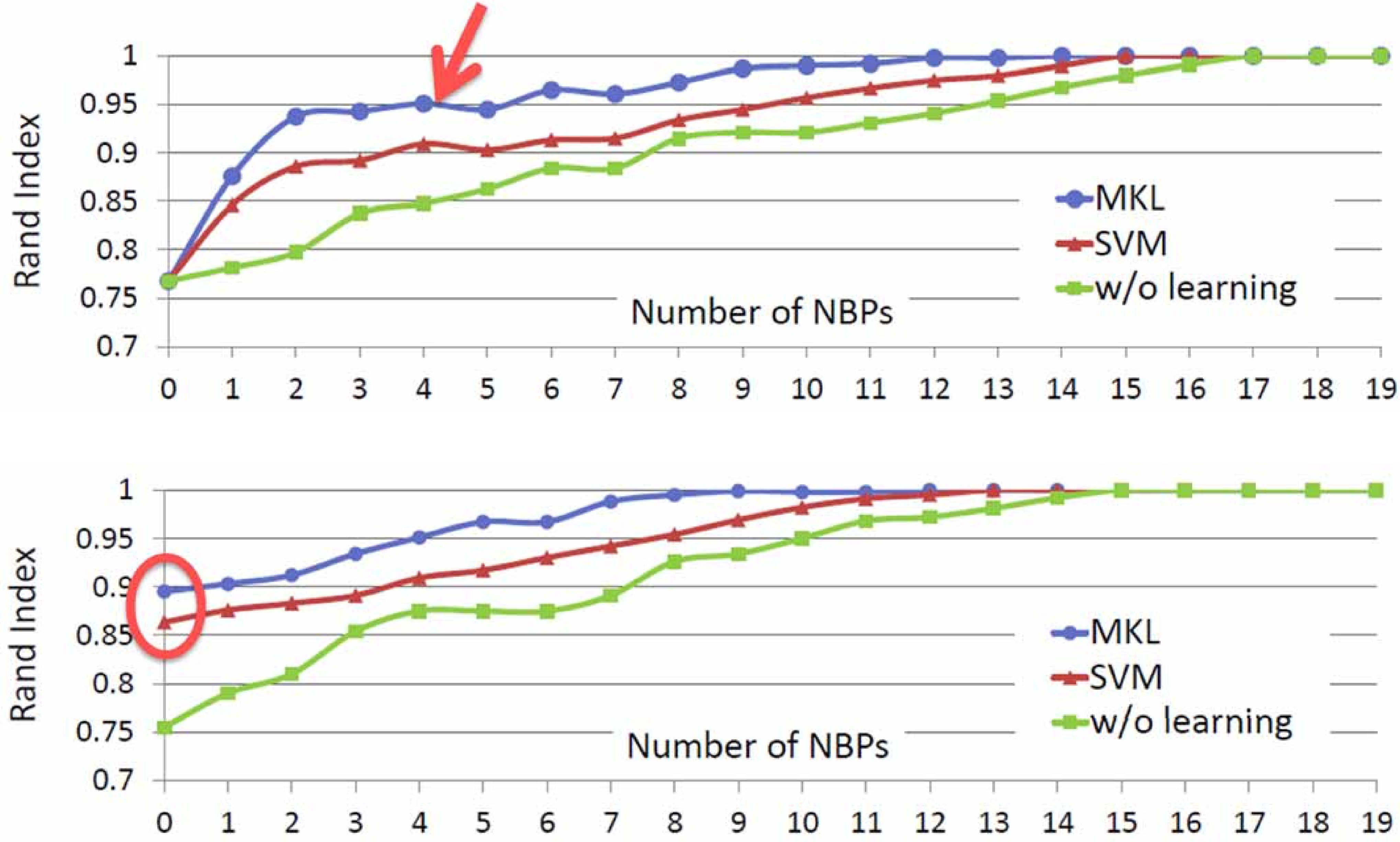
Negative examples: cannot cut

Online Learning of cut cost

- Incremental update of the classifier



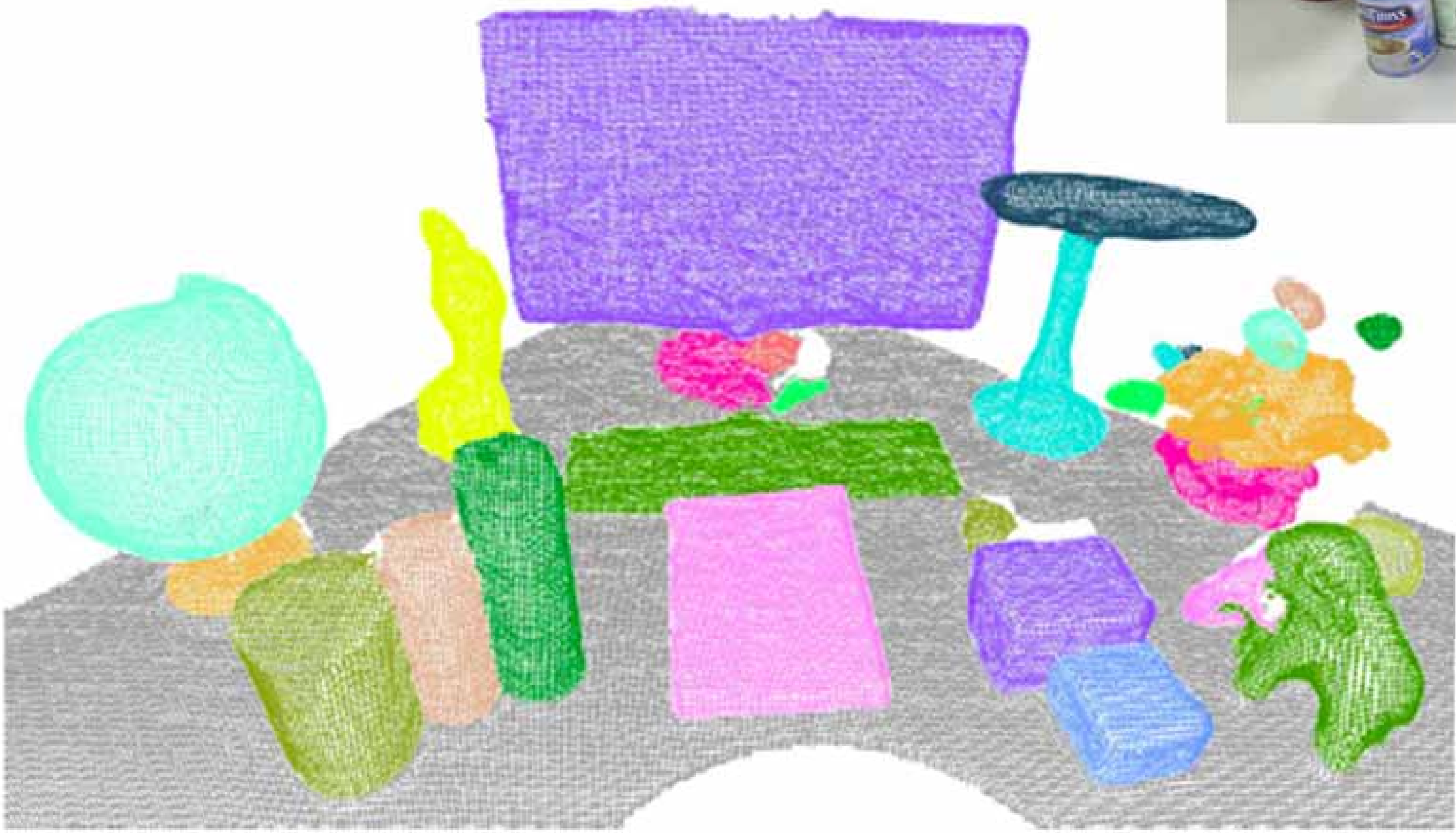
Online Learning of cut cost



Algorithm II

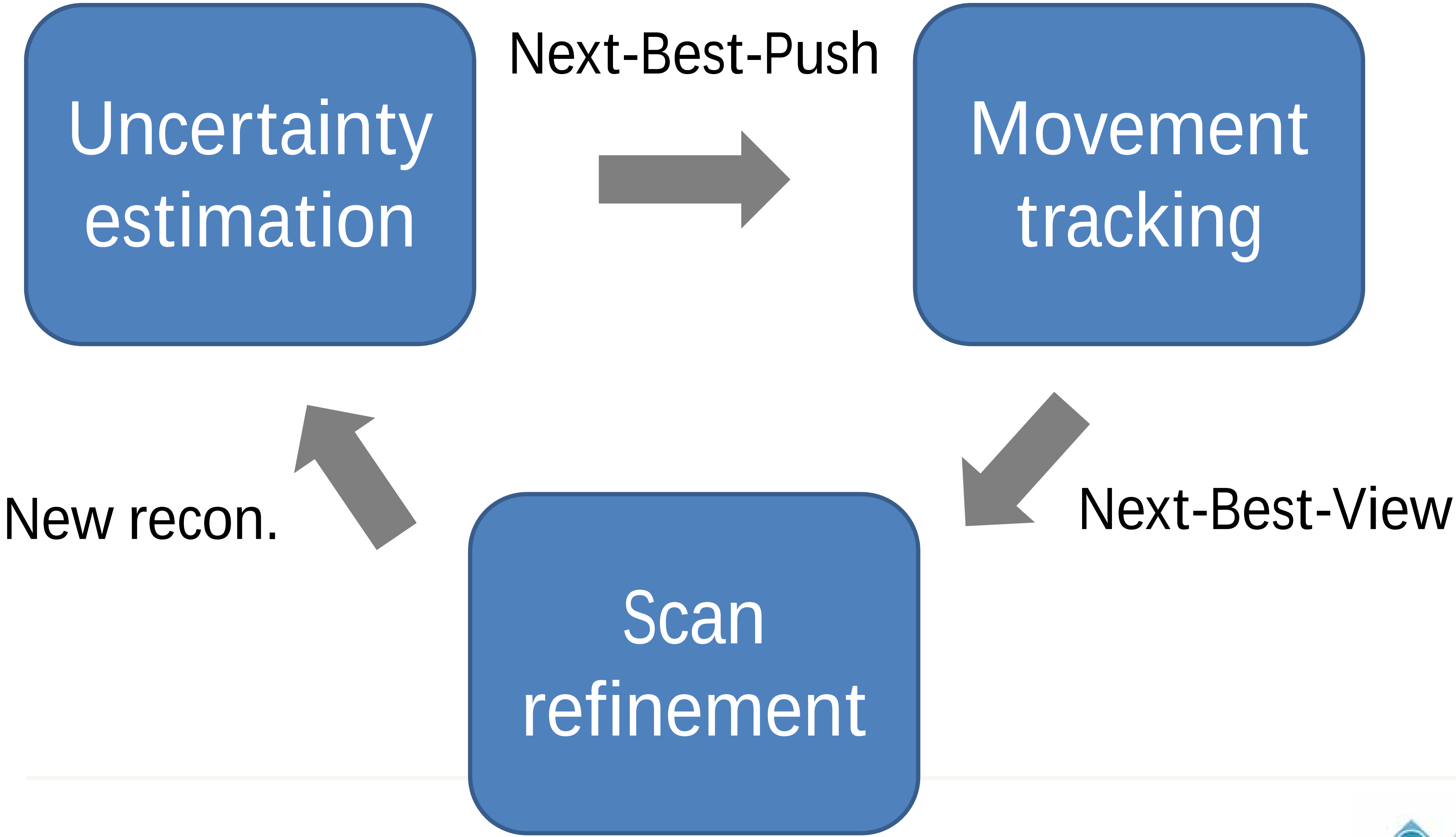
Entropy-based proactive validation

Where to push?



Object-level segmentation

Validation pipeline



Validation pipeline

Uncertainty estimation

Next-Best-Push

Movement tracking

Next-Best-View

Scan refinement

New recon.

Information gain maximization

- Maximize the **informa**

Shannon entropy

$$-\sum_i p(x_i)\log p(x_i)$$

$$I(X) = H(X) - H'(X)$$

info. gain — entropy before push — entropy after push

How to measure the entropy?

How to measure entropy?

- Our goal is *object-aware reconstruction*
- Two aspects:
  - Uncertainty of *S*egmentation
  - Uncertainty of *R*econstruction

$$H = H(\quad , \quad)$$

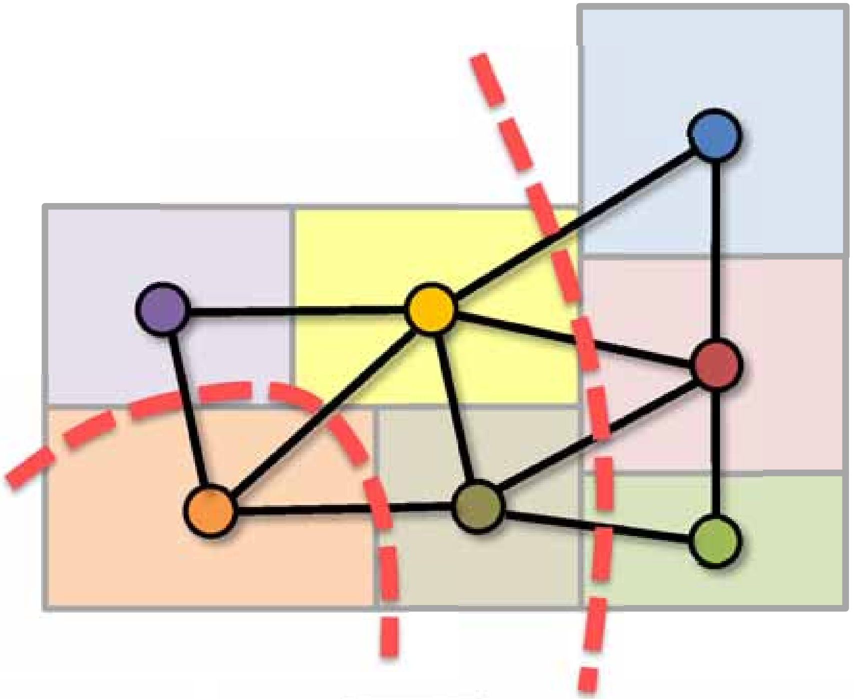
Joint entropy

$$H(S,R) = H(S) + \underline{H(R|S)}$$

Conditioned entropy

How to measure entropy?

$$H(S,R) = \boxed{H(S)} + H(R|S)$$



$$H(S) = -2\sum_{e\in\mathcal{E}^{\mathrm{p}}} \underline{p_c(e)} \log \underline{p_c(e)}$$

Cut probability

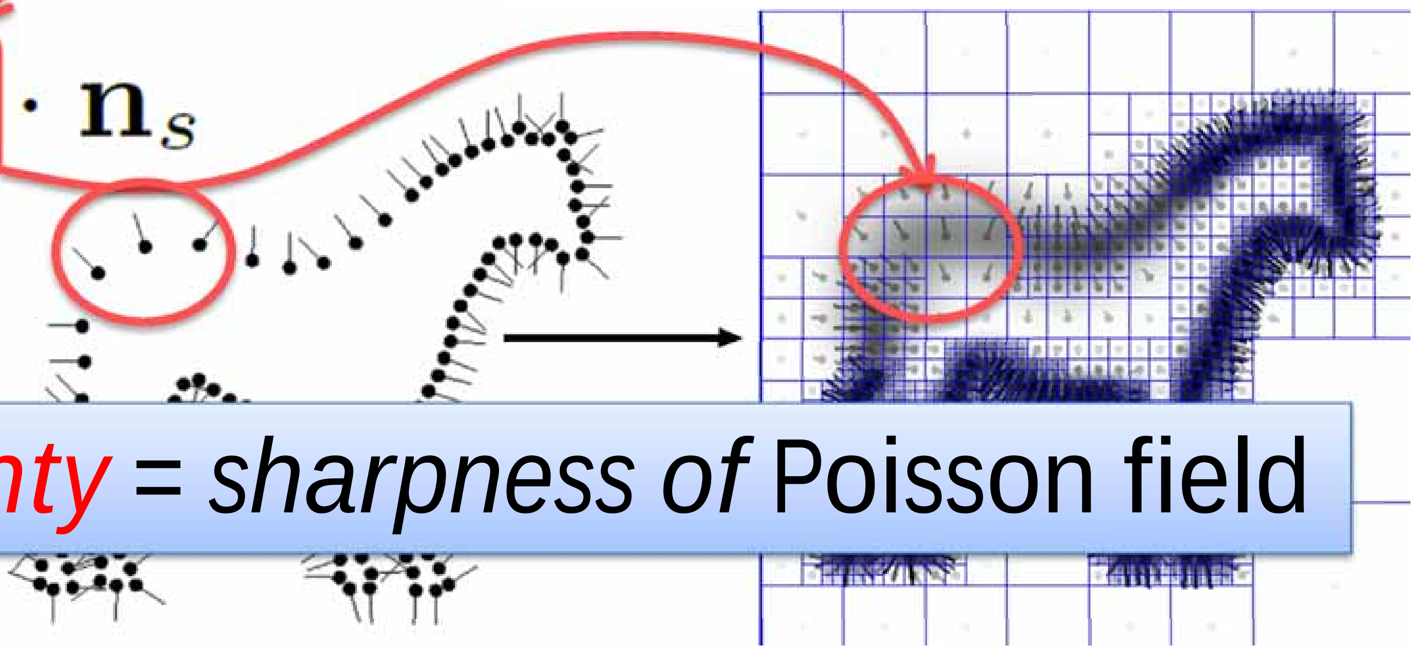
How to measure entropy?

$$H(S,R) = H(S) + \boxed{H(R|S)}$$

$$H(R) = -\sum_{s\in\Omega} g(\boxed{c(s)}) \log g(c(s))$$

$$c(s) = \boxed{\Gamma(s)} \cdot \mathbf{n}_s$$

[Kazhdan et al. 2006]



recon. *certainty* = *sharpness of* Poisson field

How to measure entropy?

$$H(S,R) = H(S) + \boxed{H(R|S)}$$

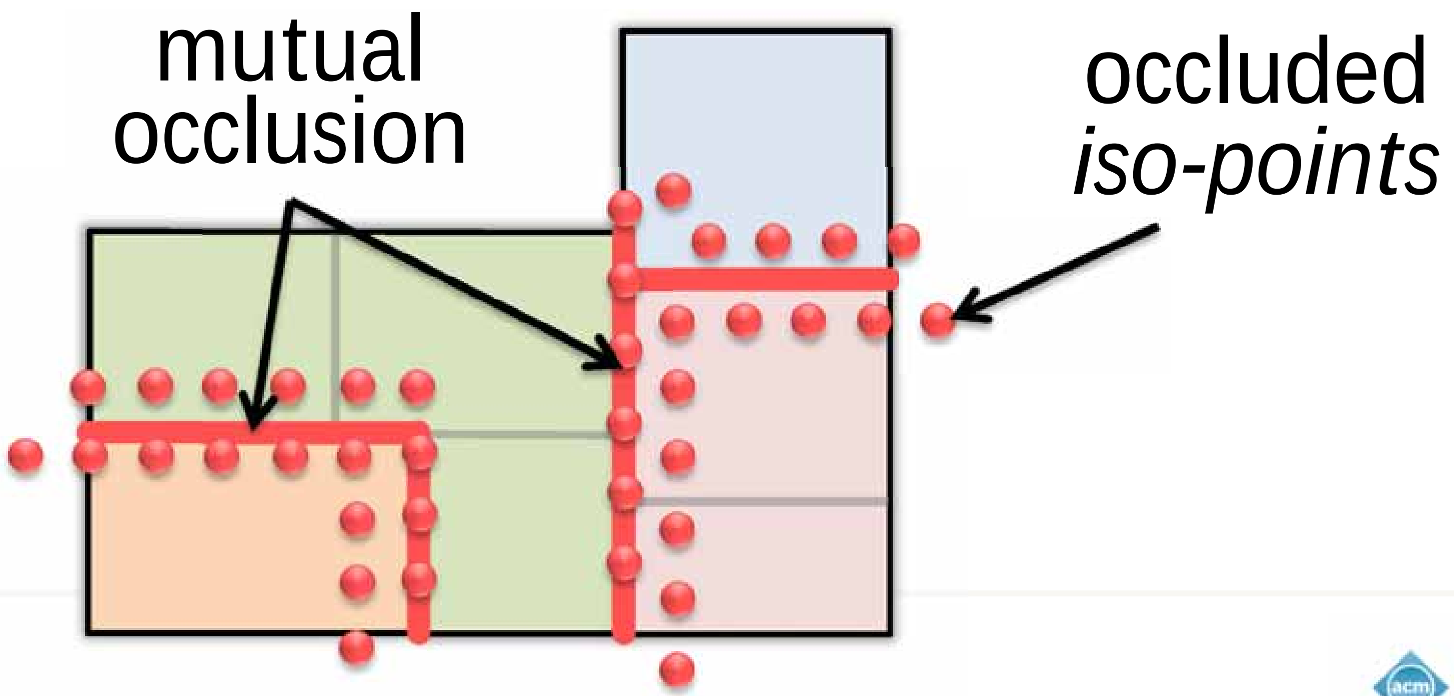
$$\underbrace{H(R|S)}_{\text{Conditioned entropy}} = -\sum_{e \in \mathcal{E}^{0}(S)} p_{\mathrm{e}}(e) \sum_{s \in \bar{\Omega}(e)} g(c(s)) \log g(c(s))$$

Given an ***object*** *segmentation*, how much uncertainty is there in the ***object-wise*** *reconstruction*?

How to measure entropy?

$$H(S,R) = H(S) + \boxed{H(R|S)}$$

$$H(R|S) = -\sum_{e\in\mathcal{E}^{o}(S)} p_{\mathrm{e}}(e) \sum_{s\in\bar{\Omega}(e)} g(c(s)) \log g(c(s))$$



How to select the Next Best Push?

Search over all push points:

$$u^* = \arg\max_u I(S, R | \langle \mathbf{p}_u, \mathbf{d}_u \rangle)$$

Posterior entropy

$$I(S, R | \langle \mathbf{p}_u, \mathbf{d}_u \rangle) = \underbrace{H(S, R)}_{\text{before push is performed}} - \underbrace{H'(S, R | \langle \mathbf{p}_u, \mathbf{d}_u \rangle)}_{\text{after push is performed}}$$

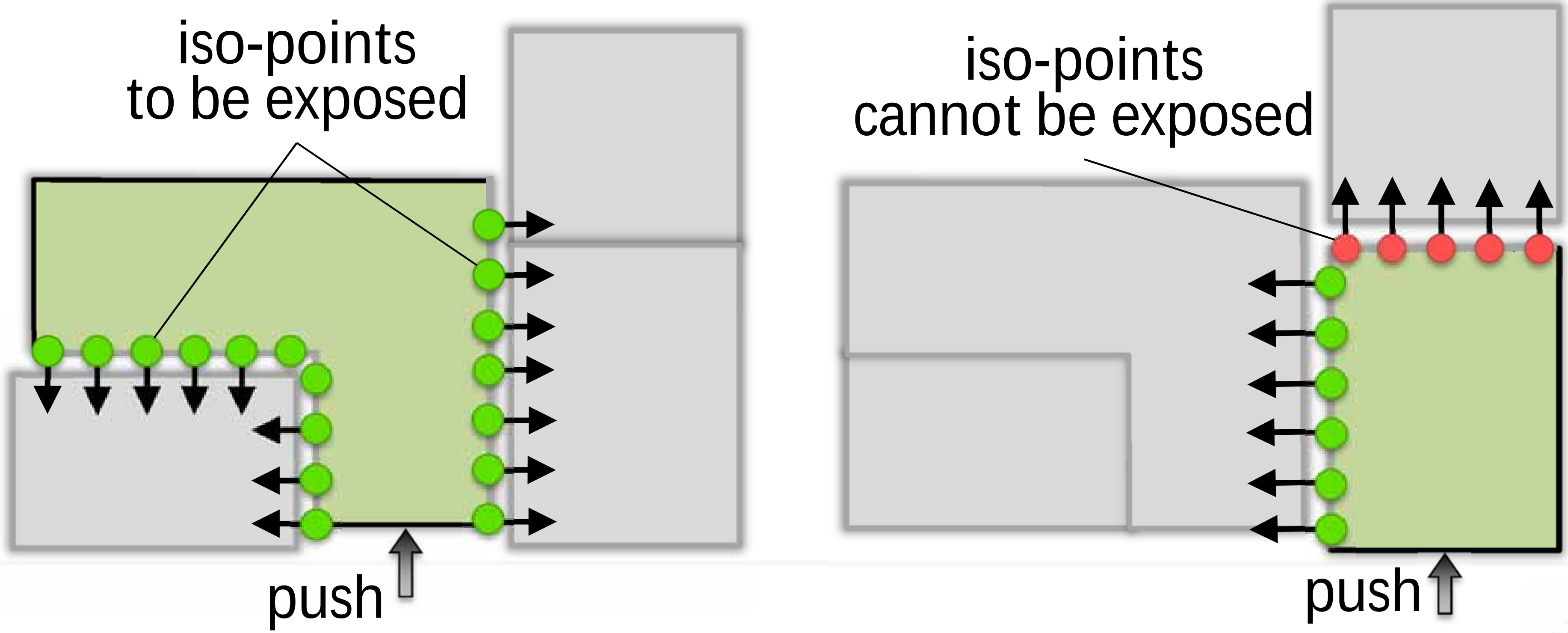
before push is performed

after push is performed

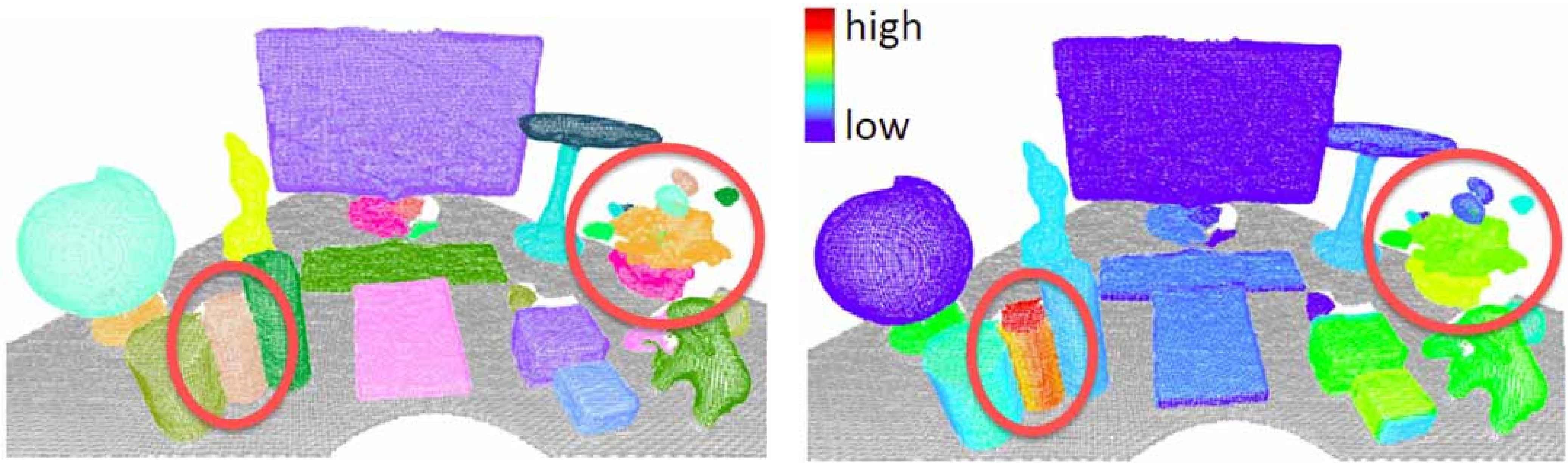
How to compute the posterior entropy

Given a push, how much uncertainty can be reduced by it?

How many occluded iso-points can be exposed by it?



Plot of information gain

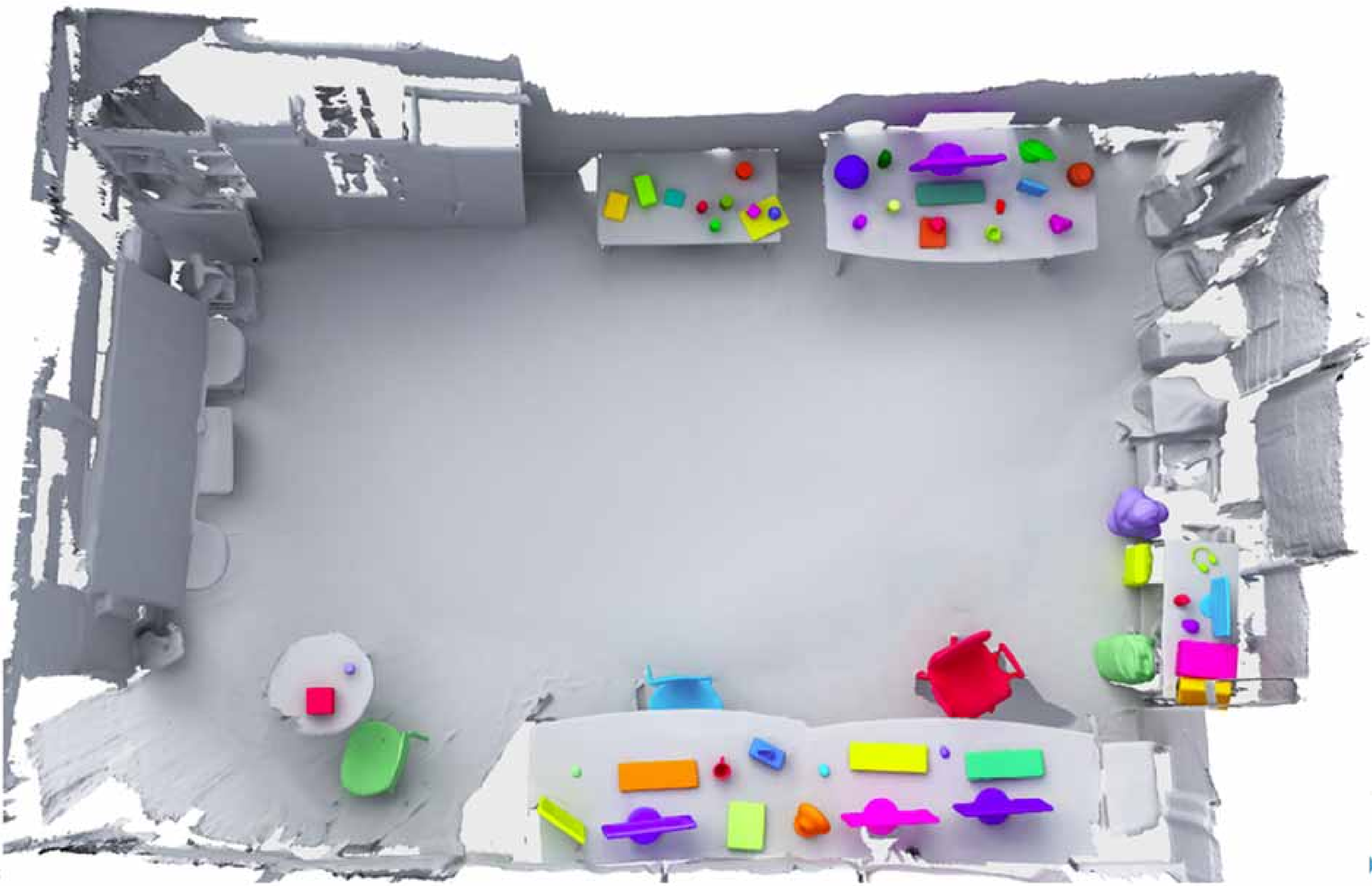


Object-level segmentation　　　　Information gain

Results and evaluation

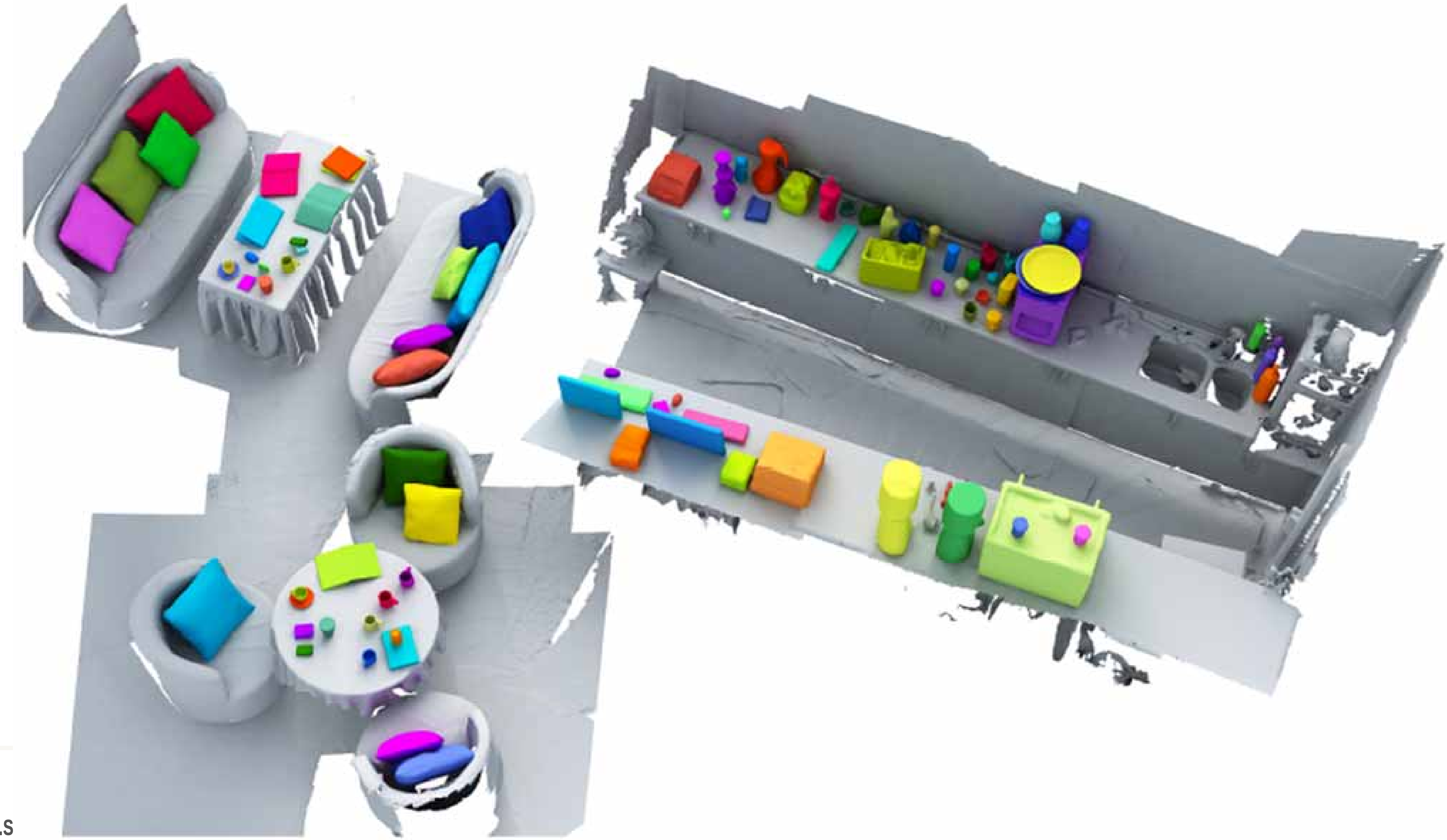
Results

- Test on real-life scenes



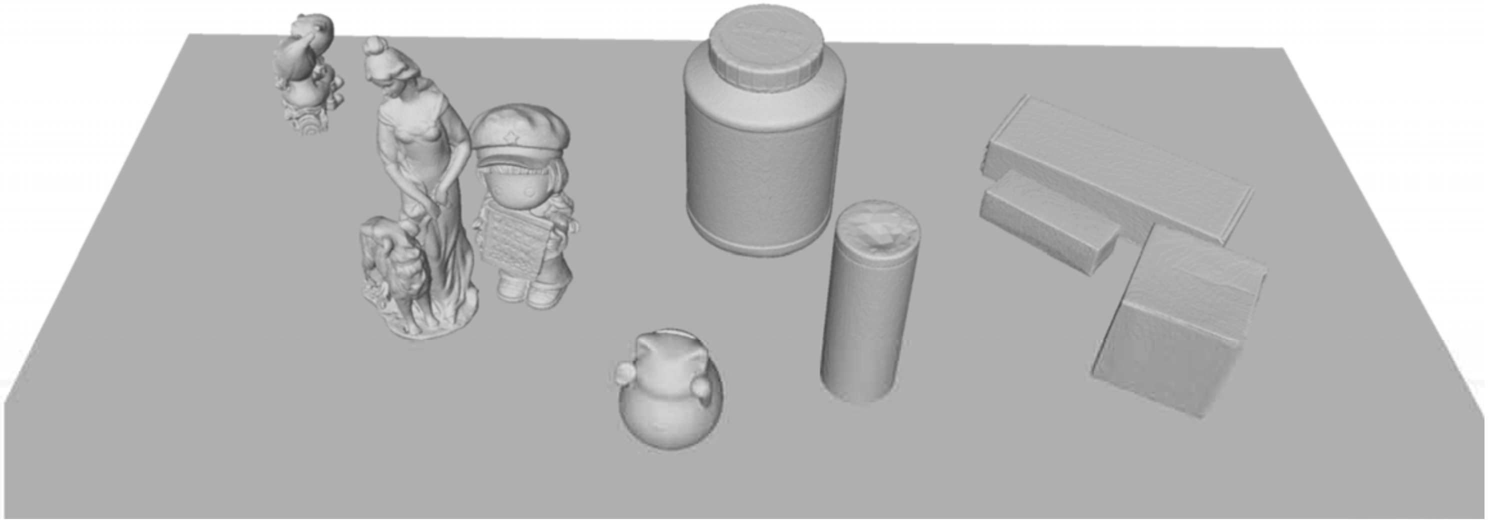
Results

- Test on real-life scenes



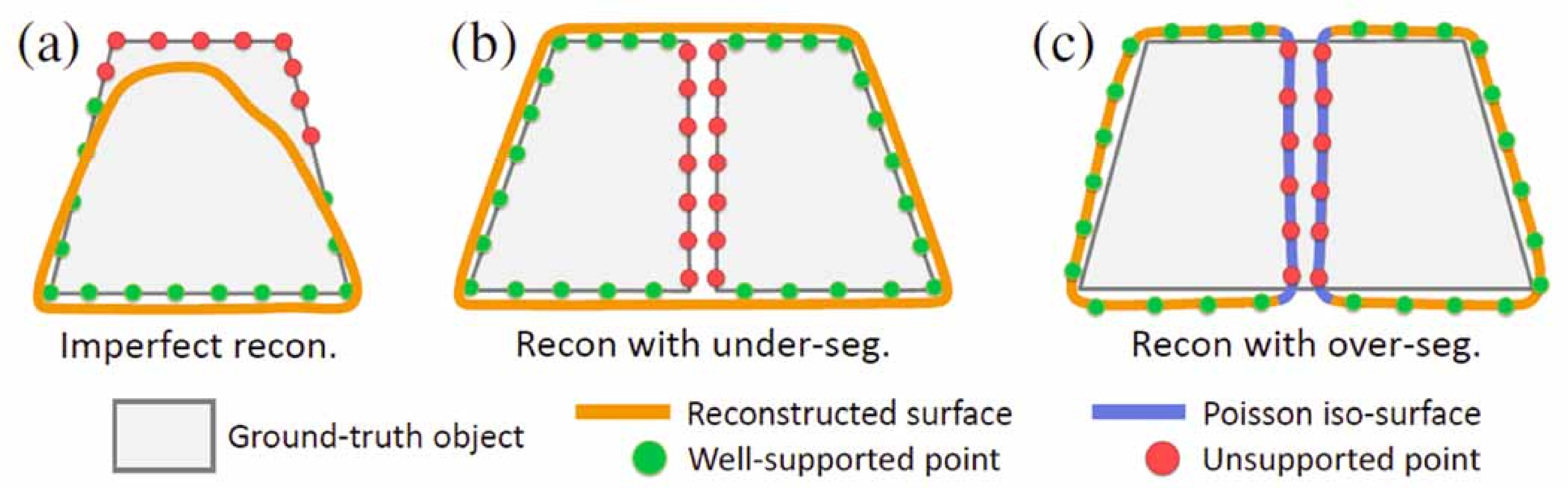
Quantitative evaluation

- Ground-truth data



A quantitative measure

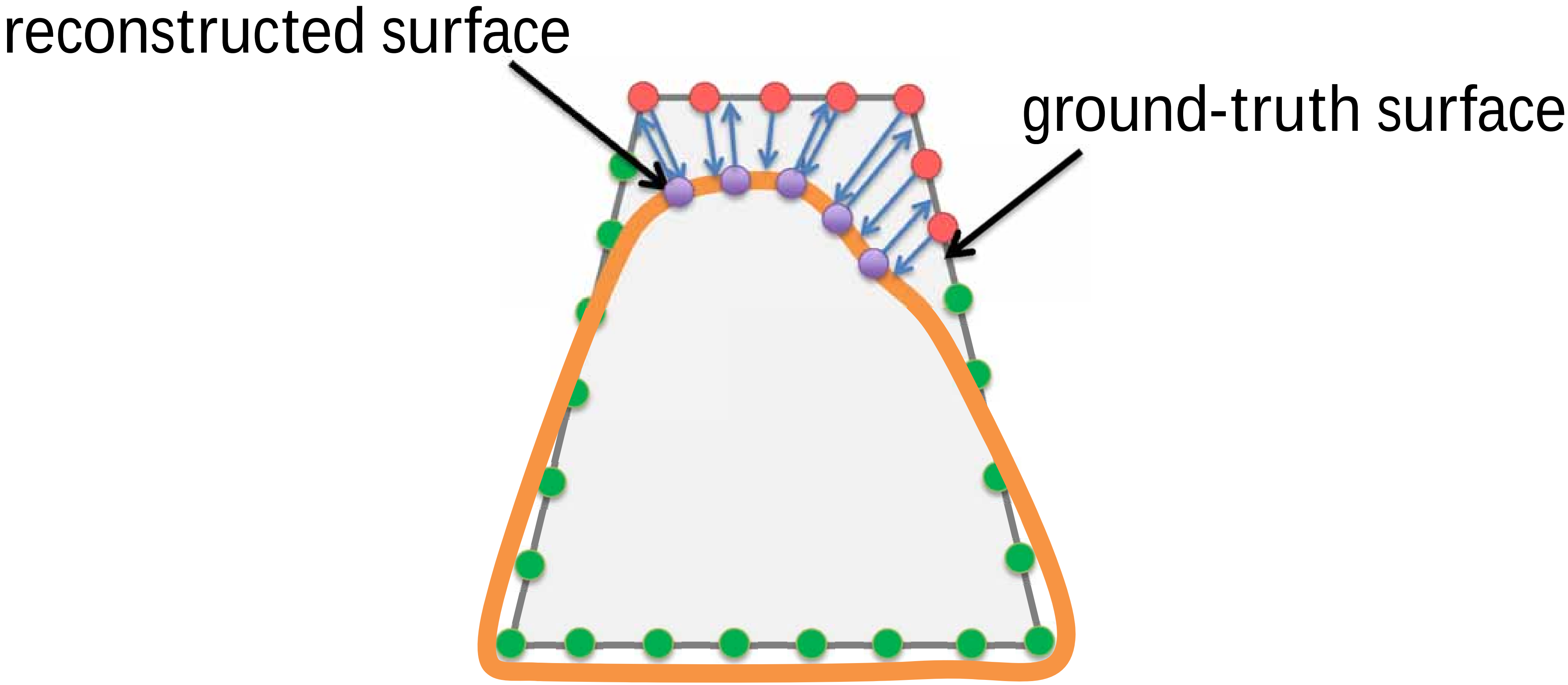
- *bilateral support* between ground-truth and reconstructed surfaces



$$\Pi(S,T) = \frac{\sum_{(p_i,\mathbf{n}_i)\in T} \pi(p_i,\mathbf{n}_i,S)}{|T|} + \frac{\sum_{(p_j,\mathbf{n}_j)\in S} \pi(p_j,\mathbf{n}_j,T)}{|S|}$$

Quantitative evaluation

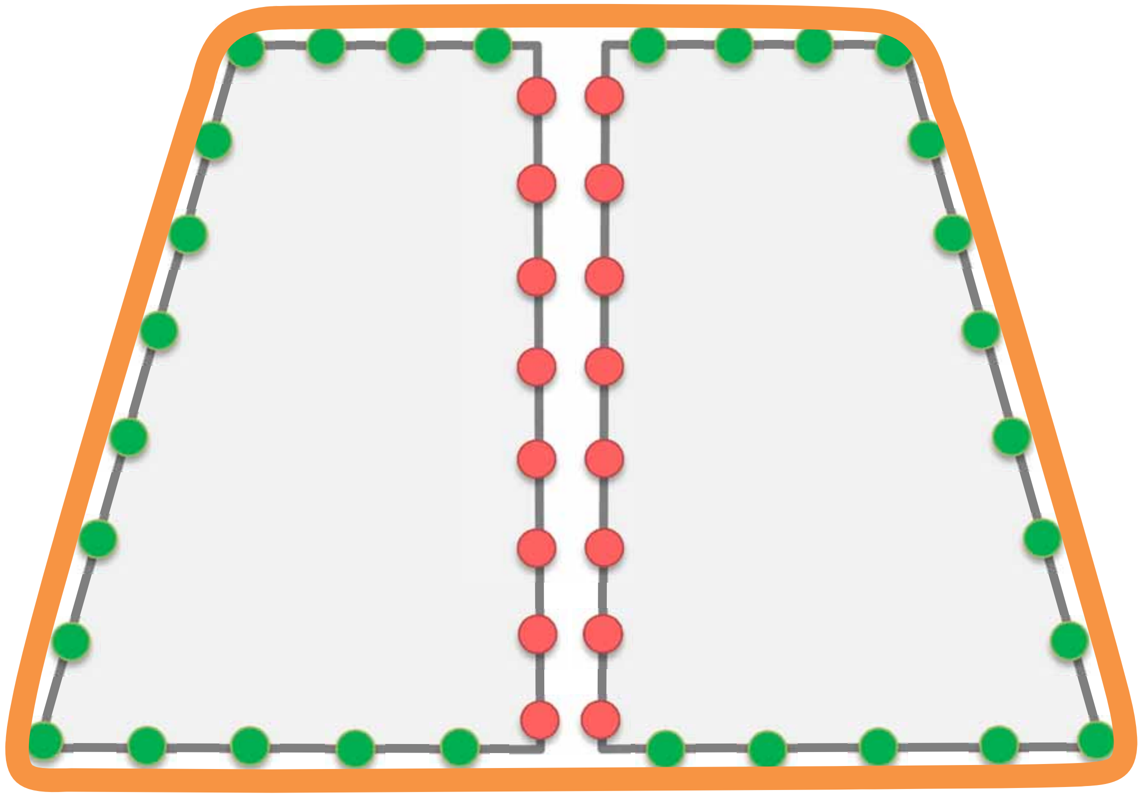
- **Bilateral support**



Imperfect reconstruction

Quantitative evaluation

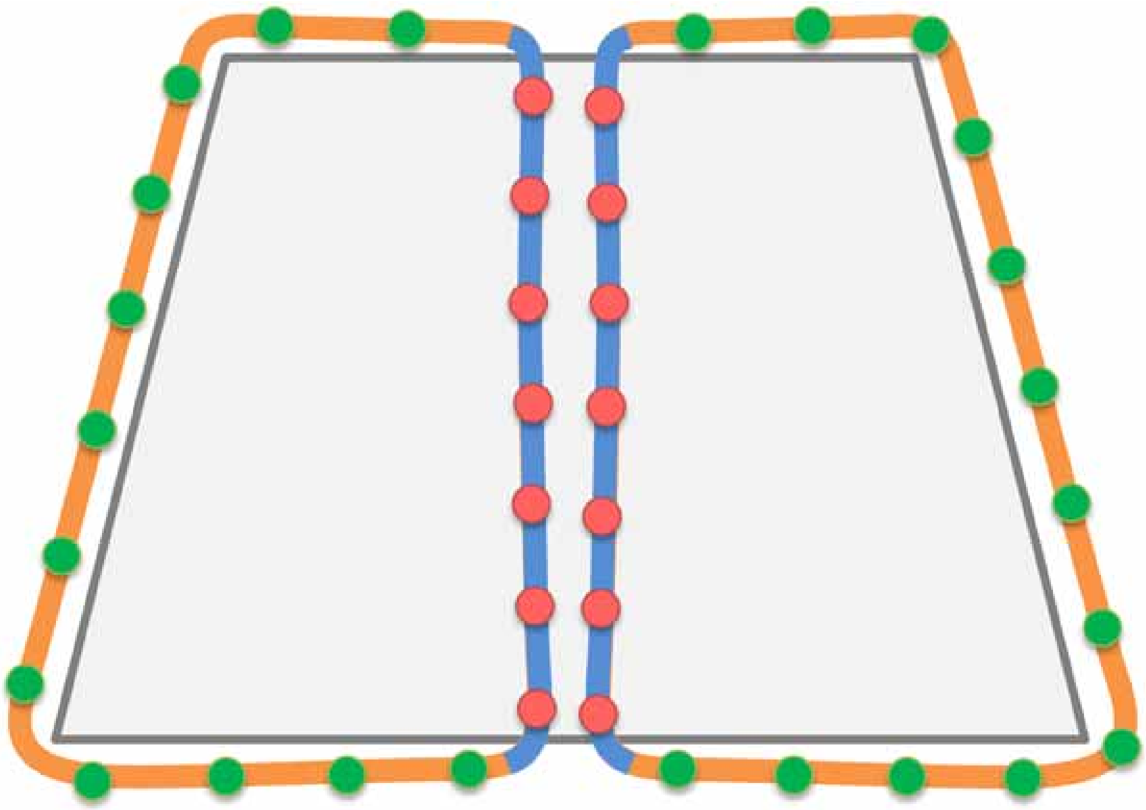
- **Bilateral support**



Reconstructed with *under*-segmentation

Quantitative evaluation

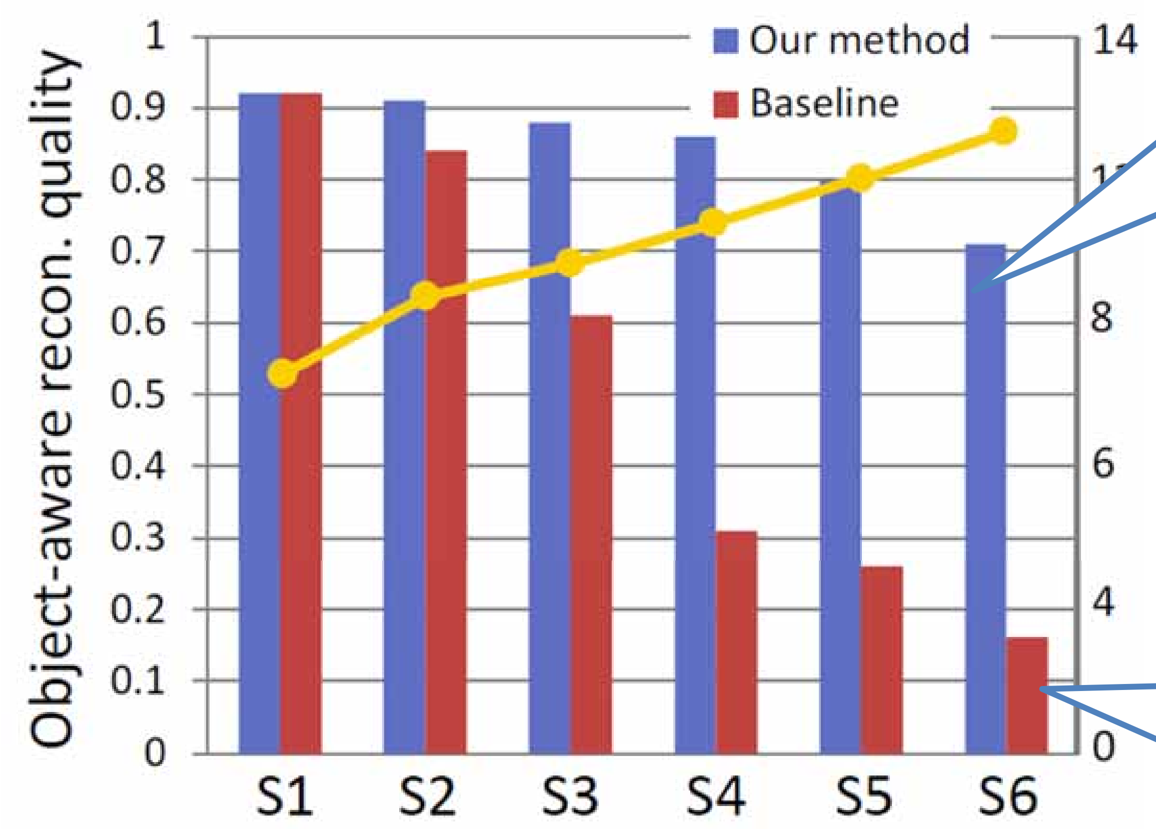
- **Bilateral support**

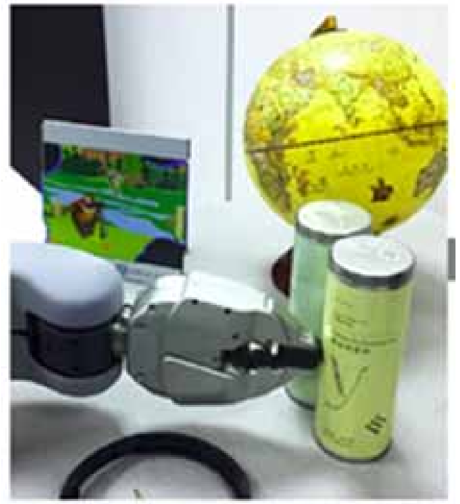


Reconstructed with *over*-segmentation

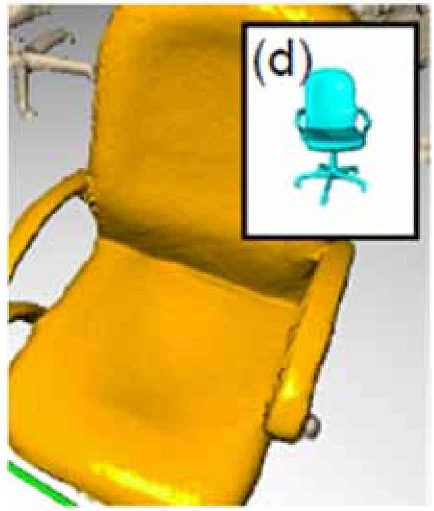
Quantitative evaluation

- Compare to [Zhang et al. 2014]





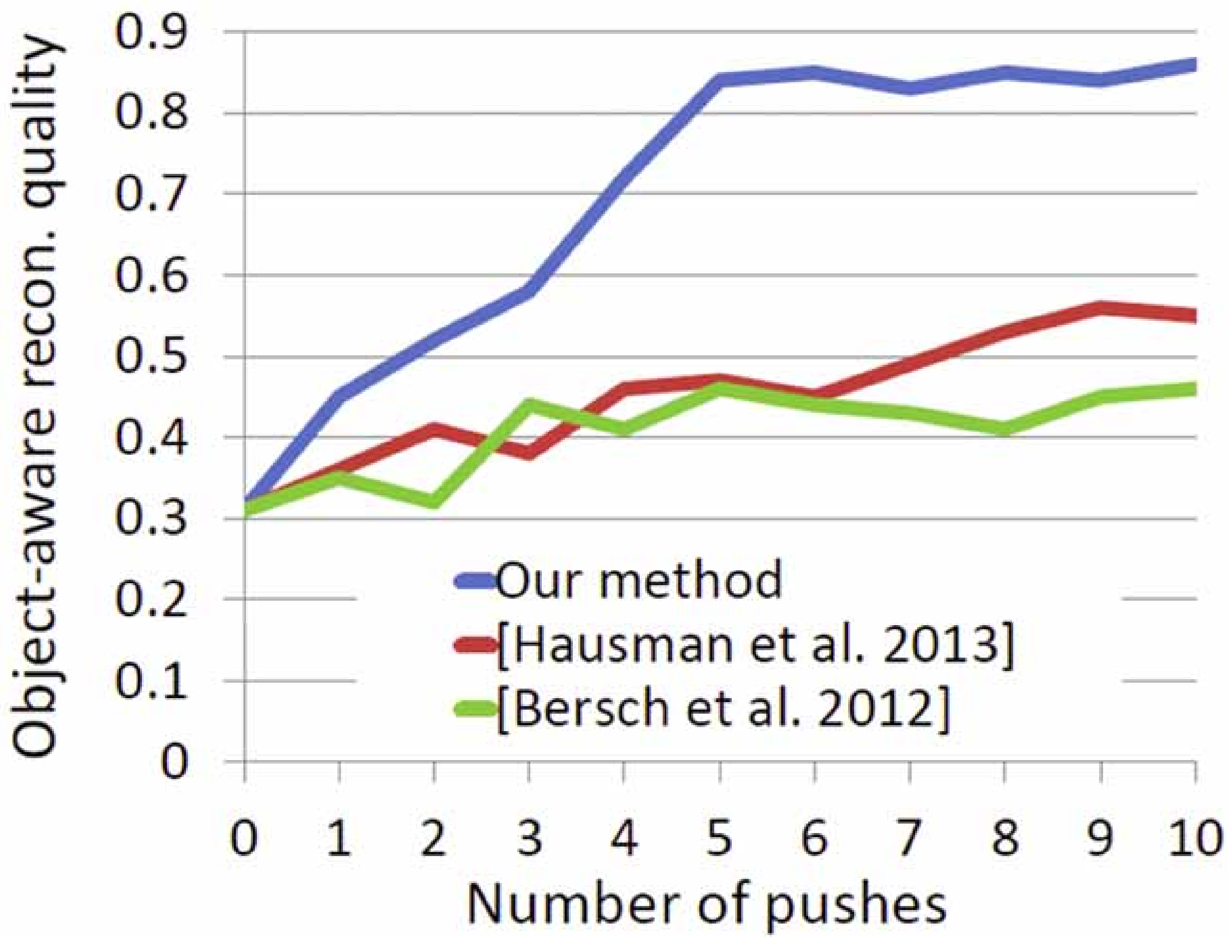
Our method

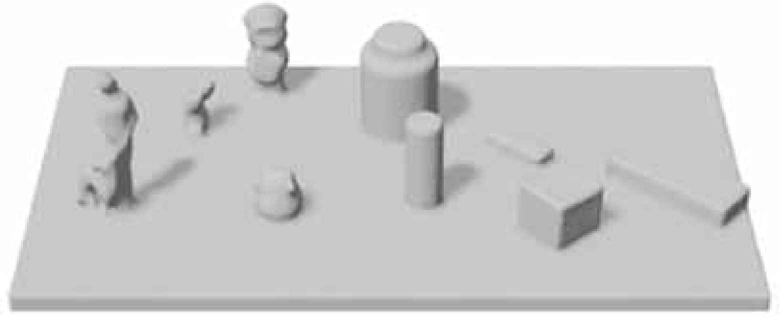


[Zhang et al. 2014]

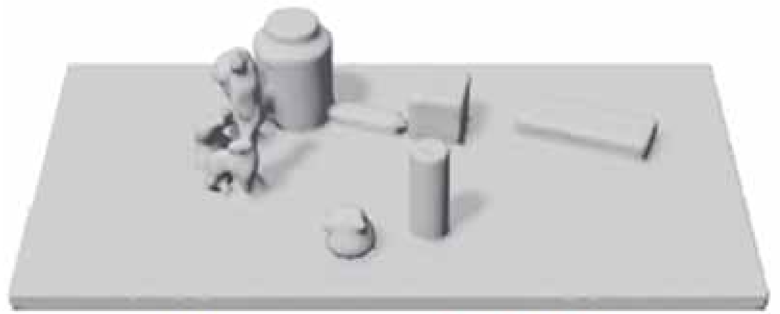
Quantitative evaluation

- Comparing to alternative robot active analysis





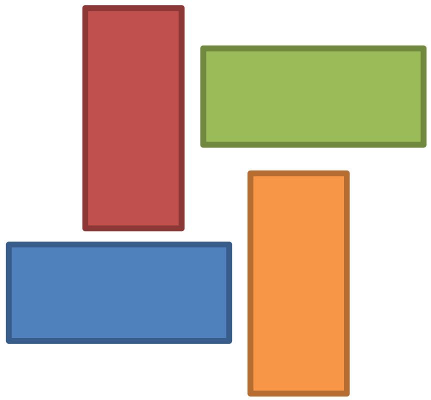
Our method



[Hausman et al. 2013]

Limitations

- Extraction rate: 50%~80%
- Push scheduling
- Vertical support
- Object stack
- Non-rigid objects

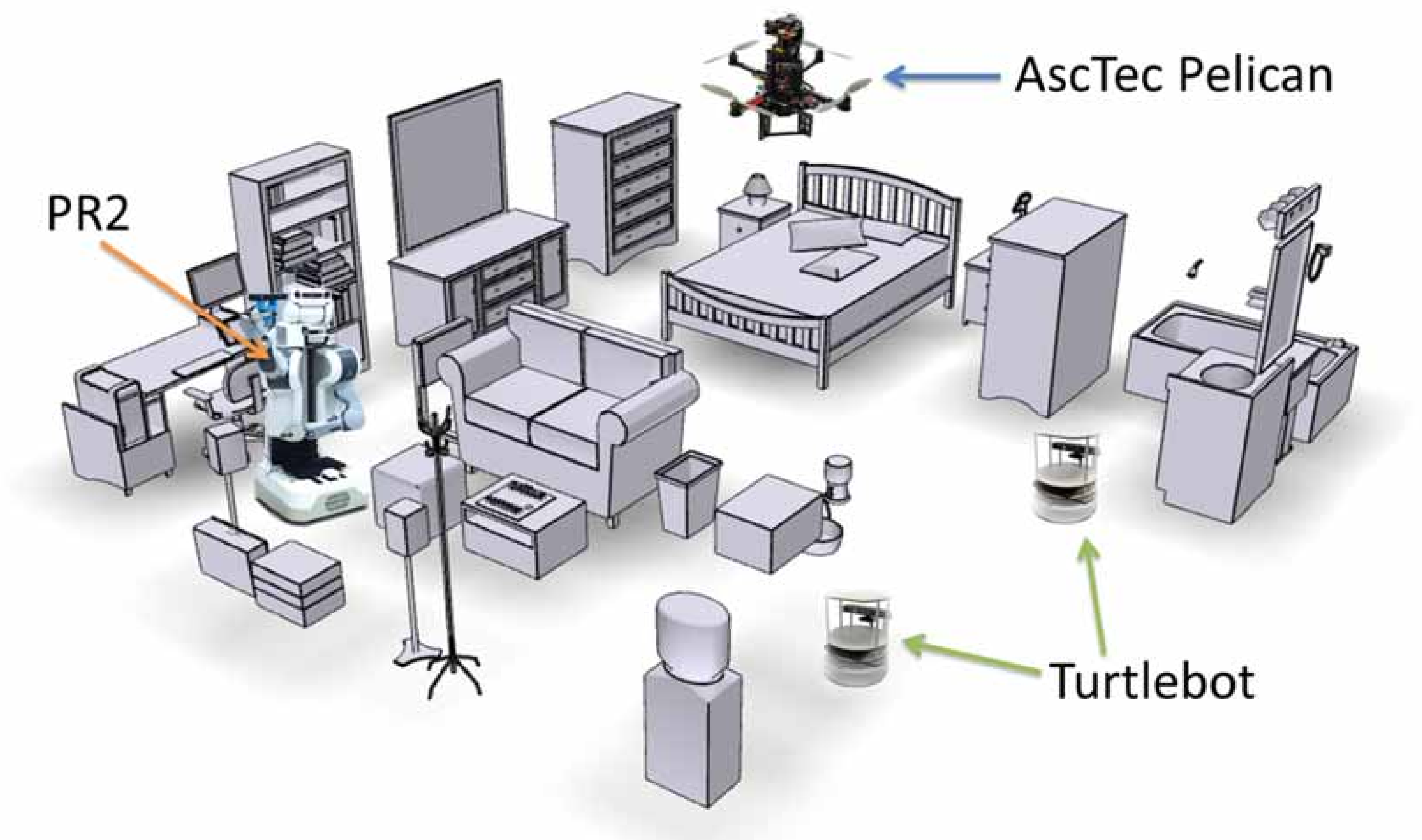


Conclusion

- A **new paradigm** of 3D acquisition
  - Autonomous scene scanning
    - Scene level: Large scale
    - Object level: Detailed
  - Proactive validation
    - Physically-validated segmentation

Future works

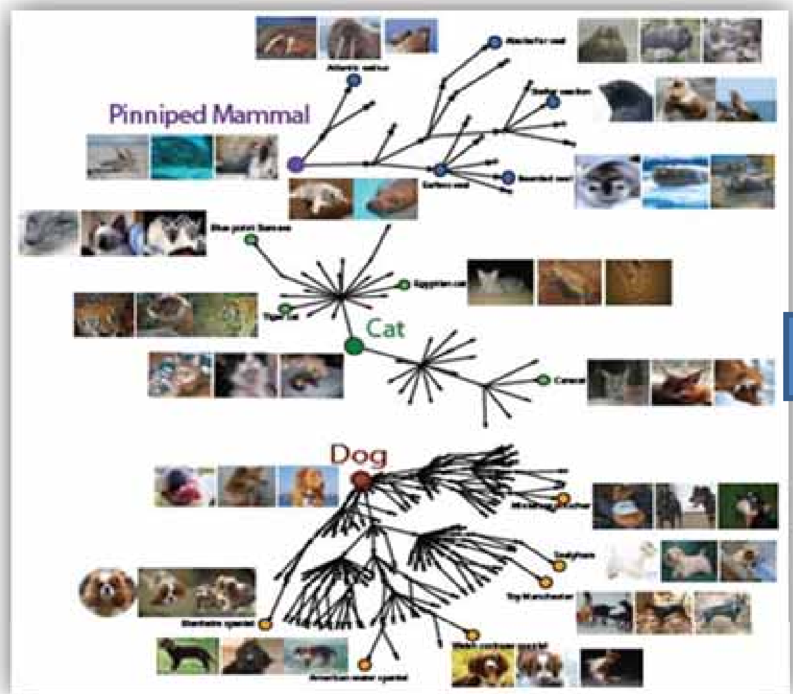
- Multiple robots?



Future works

- Recognizing while scanning

What is this?



Data-driven !

Acknowledgement

- Anonymous reviewers
- Hao (Richard) Zhang and Daniel Cohen-Or
- Research grants
  - NSFC
  - National 973 Program
  - Guangdong Science and Technology Program
  - Shenzhen VisuCA Key Lab
  - The Fundamental Research Funds of Shandong University
  - SIAT Innovation Program for Excellent Young Researchers

Thank you!

Code of object analysis:

www.kevinkaixu.net

Physically feasible push

- Heuristic rules