Analyzing, Optimizing and Synthesizing Scenes by Reasoning About Relationships Between Objects

Tianqiang Liu

Adviser: Thomas A. Funkhouser

July 23rd, 2015

PRINCETON UNIVERSITY

3D virtual scenes

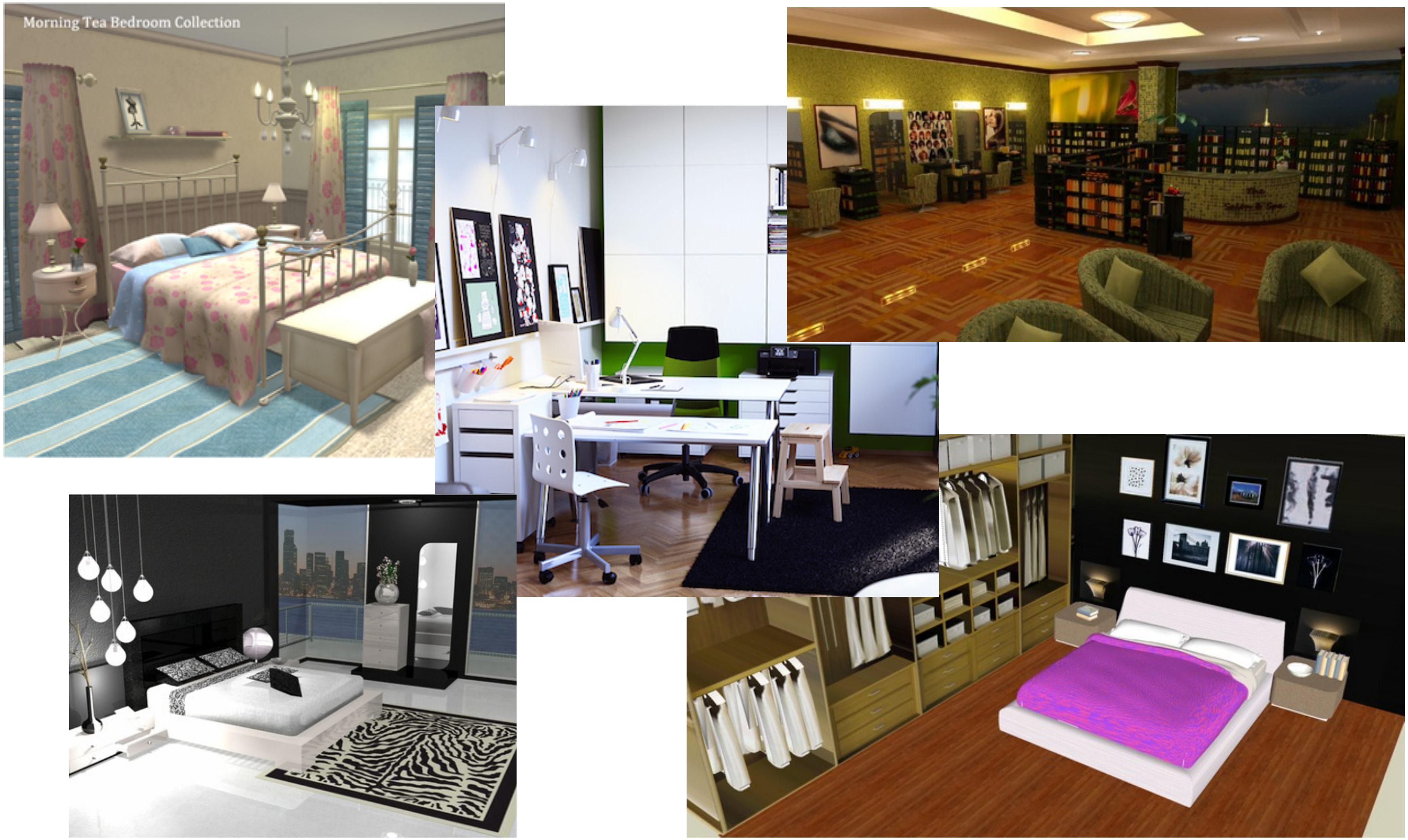


Image courtesy: Winter Thorn, IKEA, Studio Bottini, Surya M.

Manually scene modeling is tedious

- Traversing large 3D databases
- Choosing materials for each object
- Positioning objects in the scene

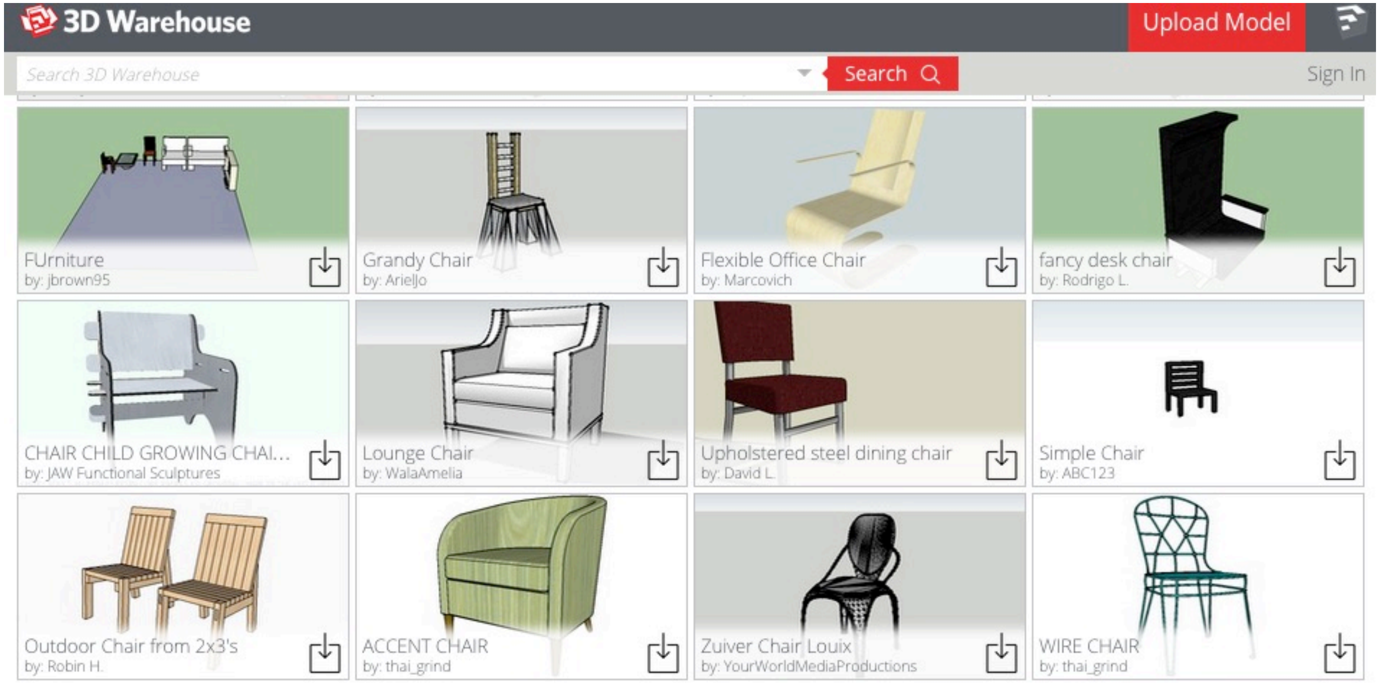


Image courtesy: Trimble 3D Warehouse

Manually scene modeling is tedious

- Traversing large 3D databases
- Choosing materials for each object
- Positioning objects in the scene

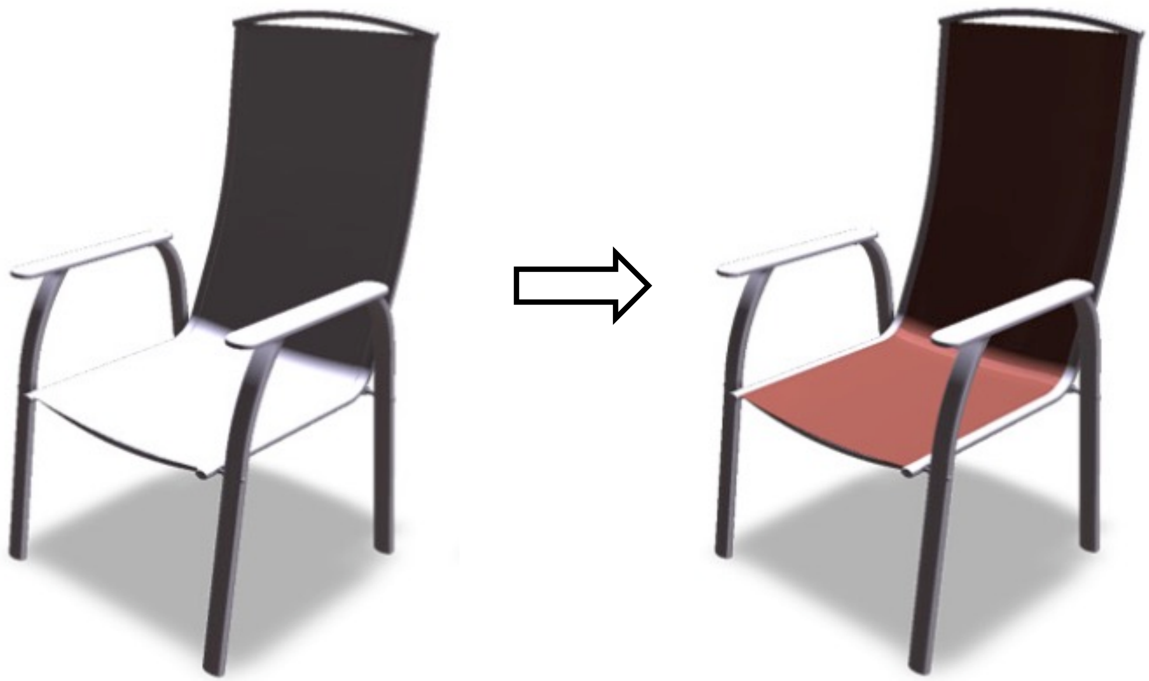


Image courtesy: Jain et al.

Manually scene modeling is tedious

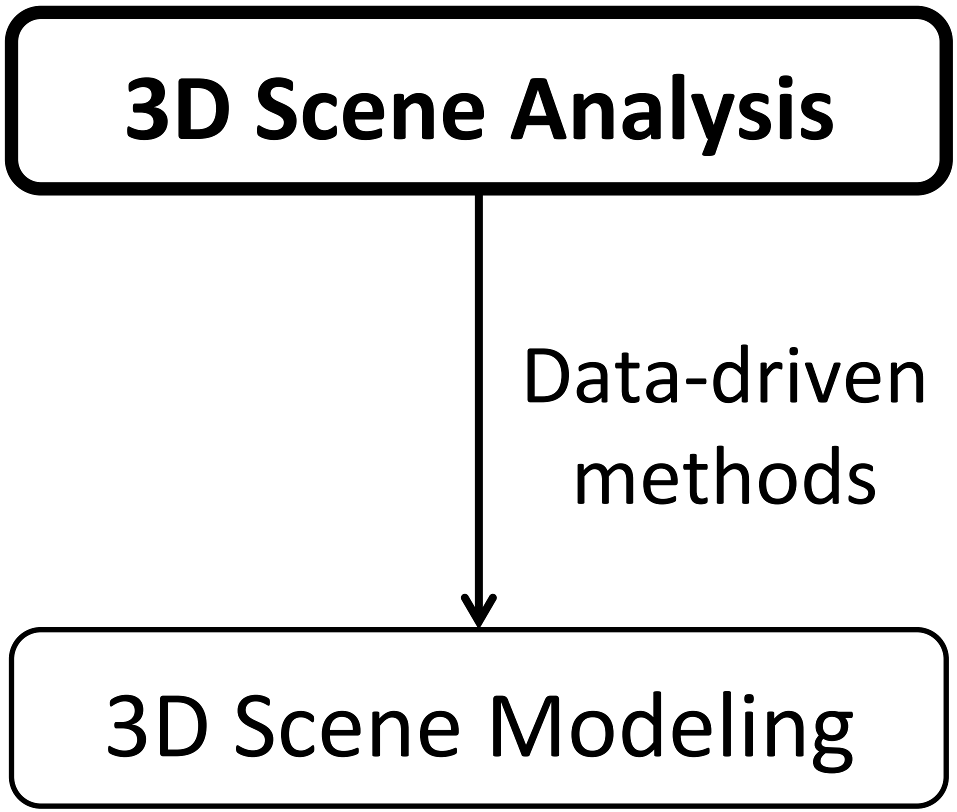
- Traversing large 3D databases
- Choosing materials for each object
- Positioning objects in the scene

Image courtesy: Yu et al.

Introduction

3D Scene Modeling

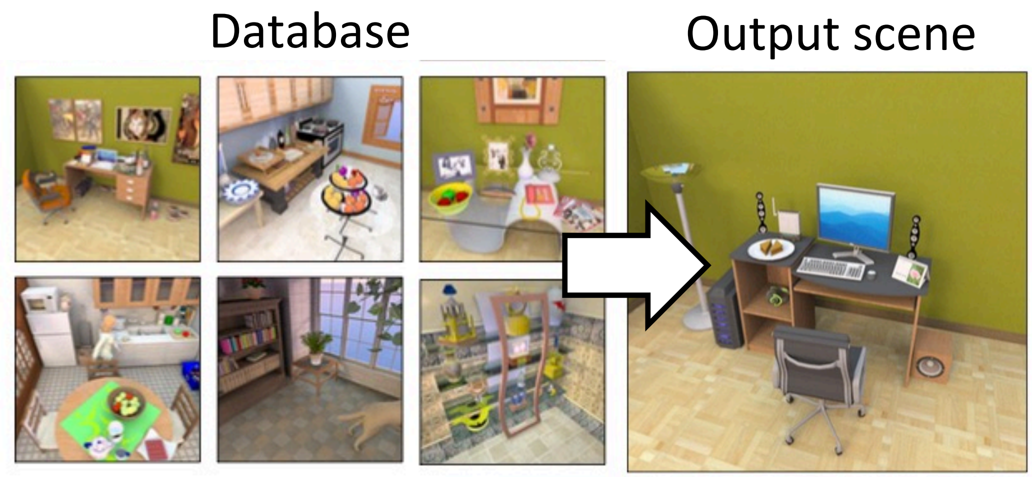
Introduction



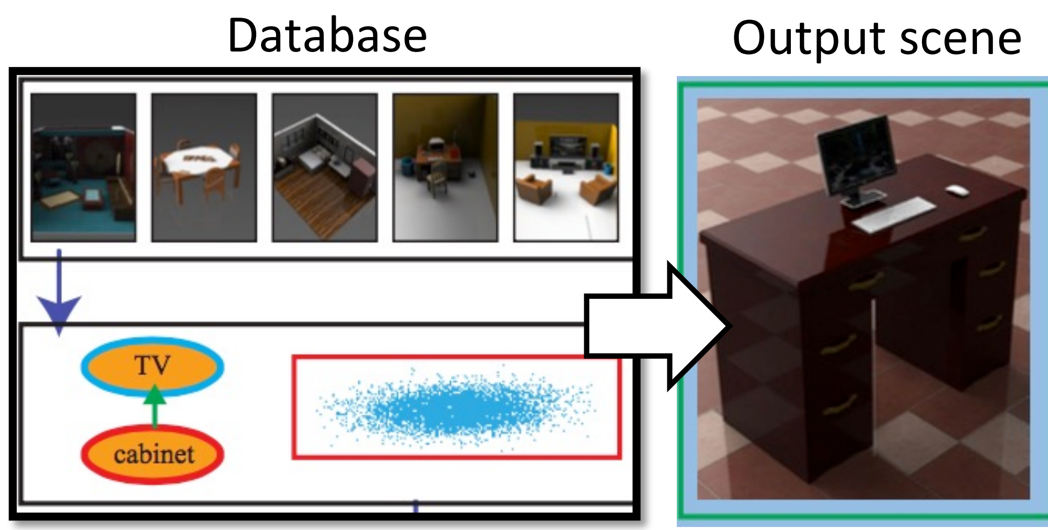
Related work: Data-driven scene modeling

Previous work requires

- Perfect segmentation
- Perfect annotation

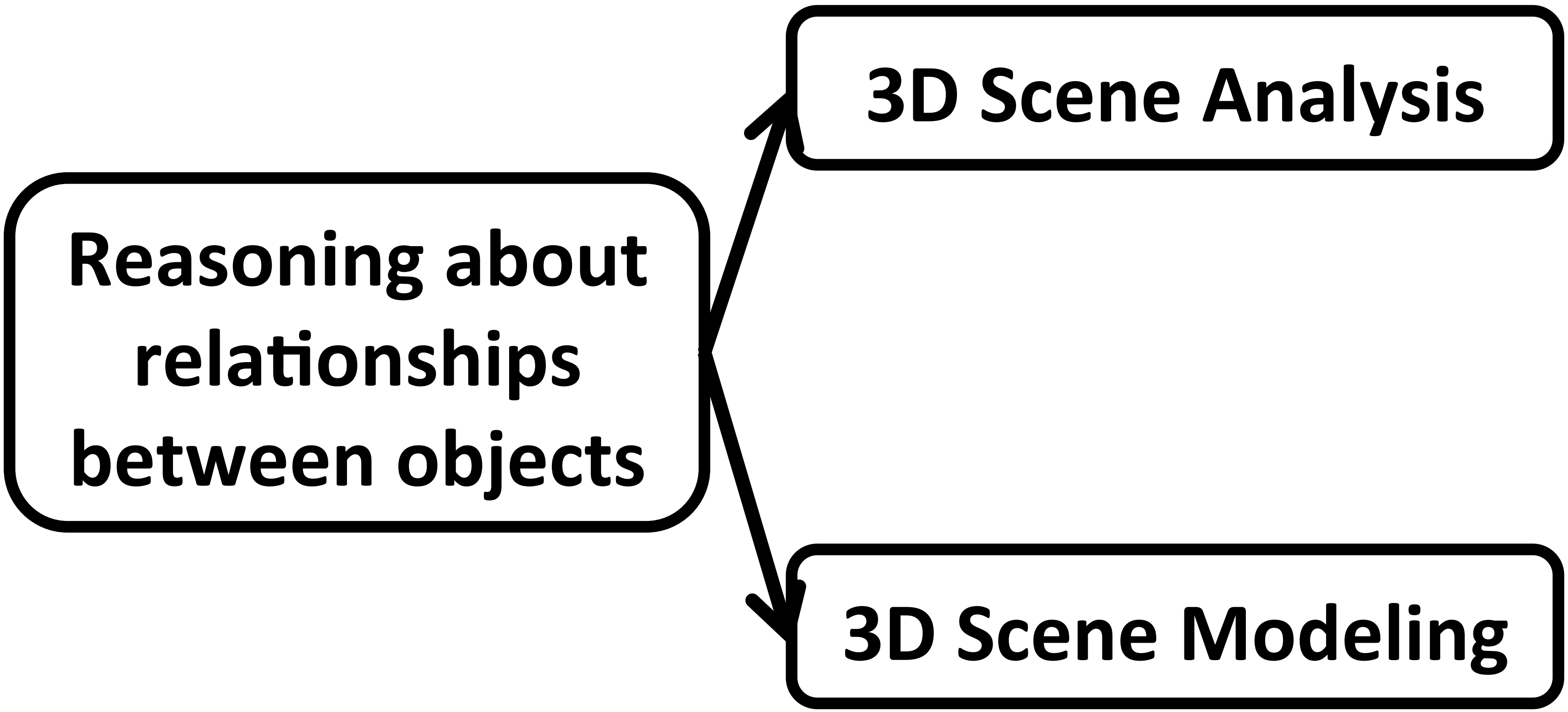


[Fisher et al. 2012]



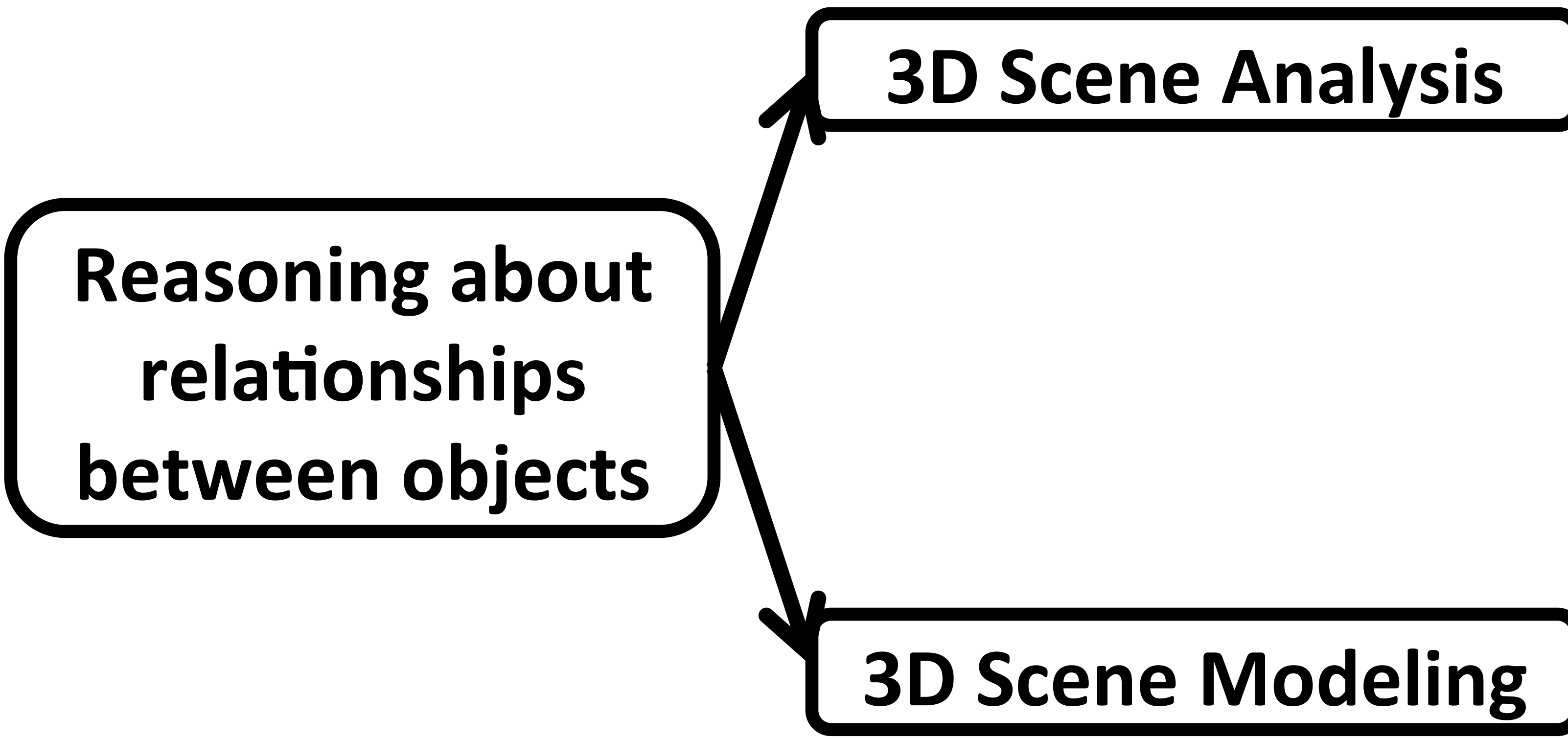
[Xu et al. 2013]

Key idea


3D Scene Analysis
Reasoning about relationships between objects
3D Scene Modeling

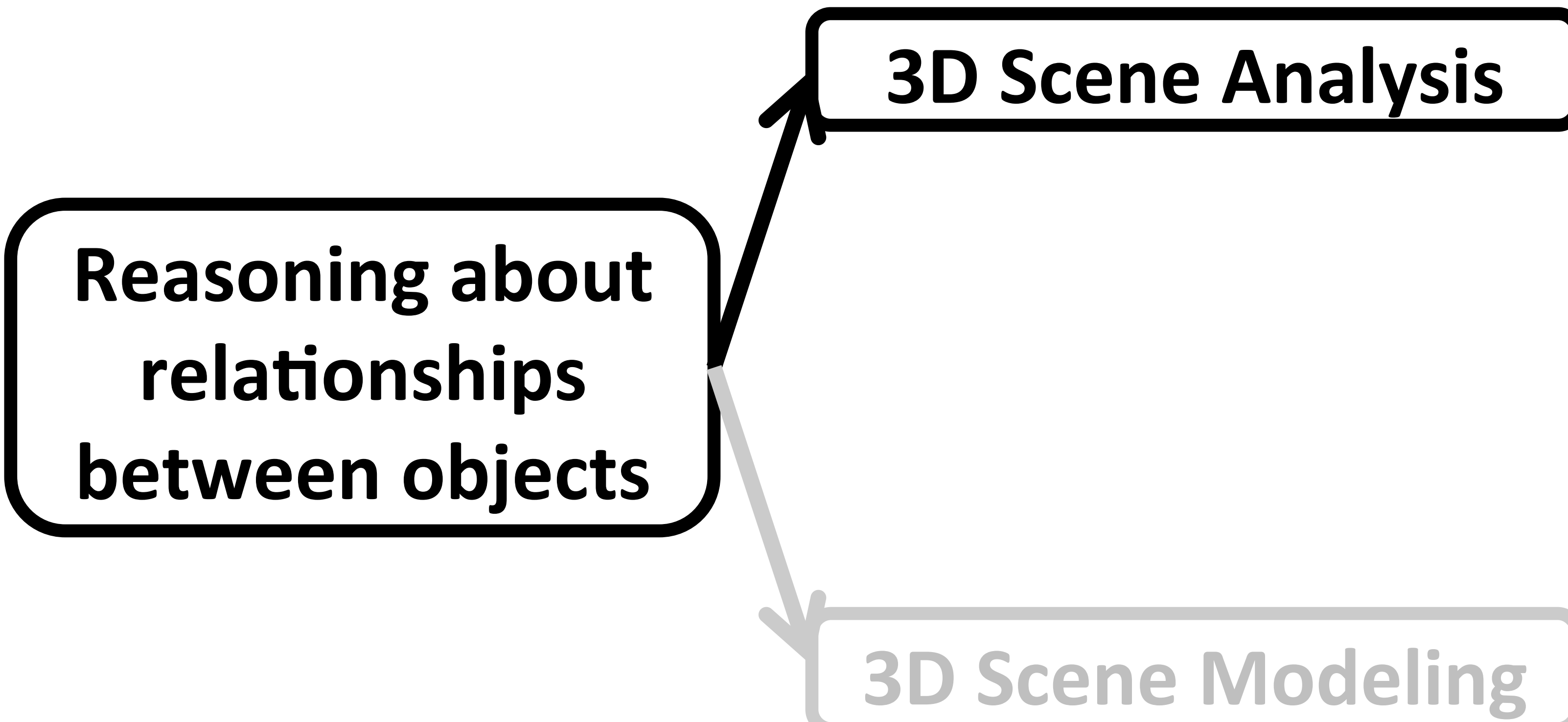
Outline

- Analyzing 3D scenes by modeling hierarchical structure
- Composition-aware scene optimization for product images
- Style compatibility for 3D furniture models



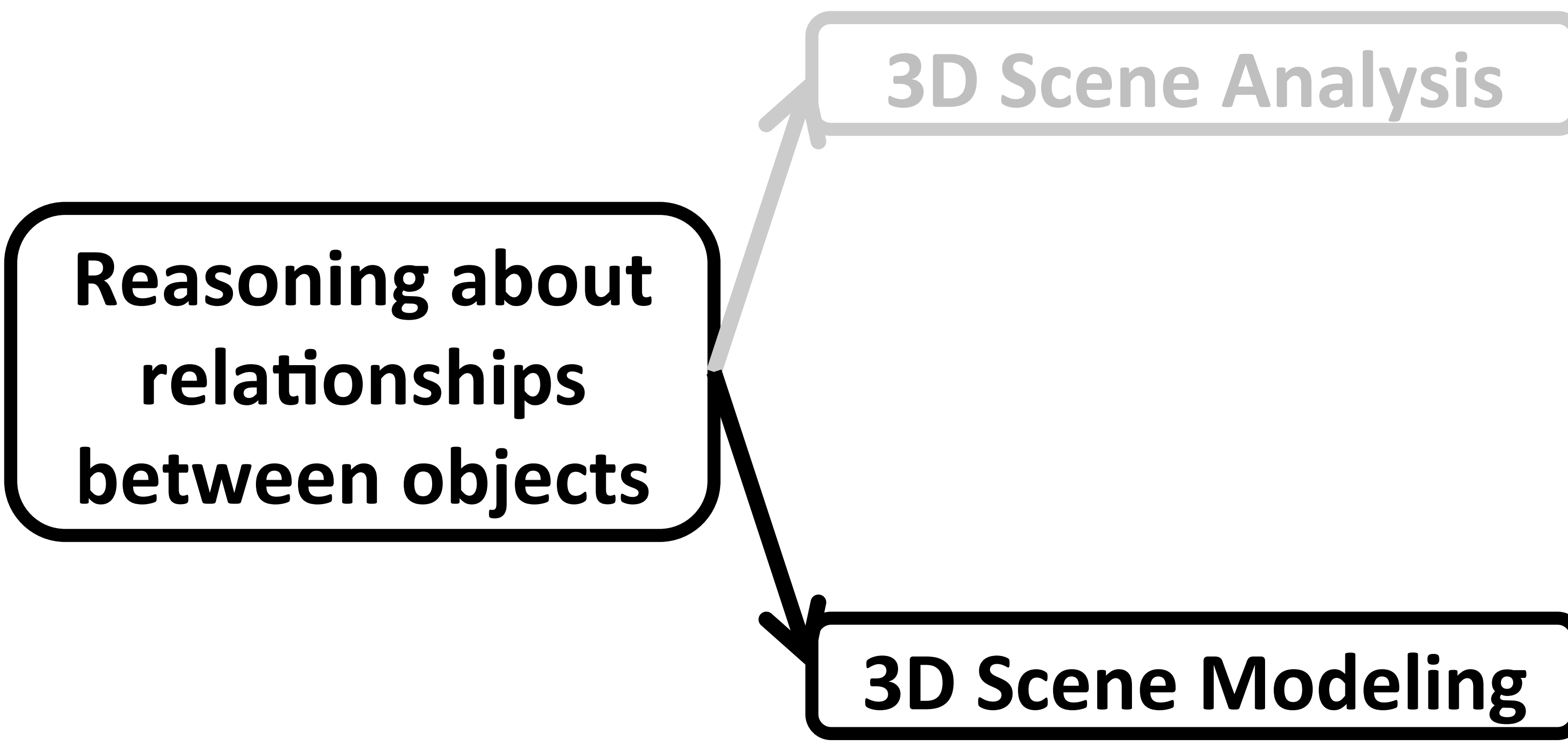
Outline

- Analyzing 3D scenes by modeling hierarchical structure
- Composition-aware scene optimization for product images
- Style compatibility for 3D furniture models



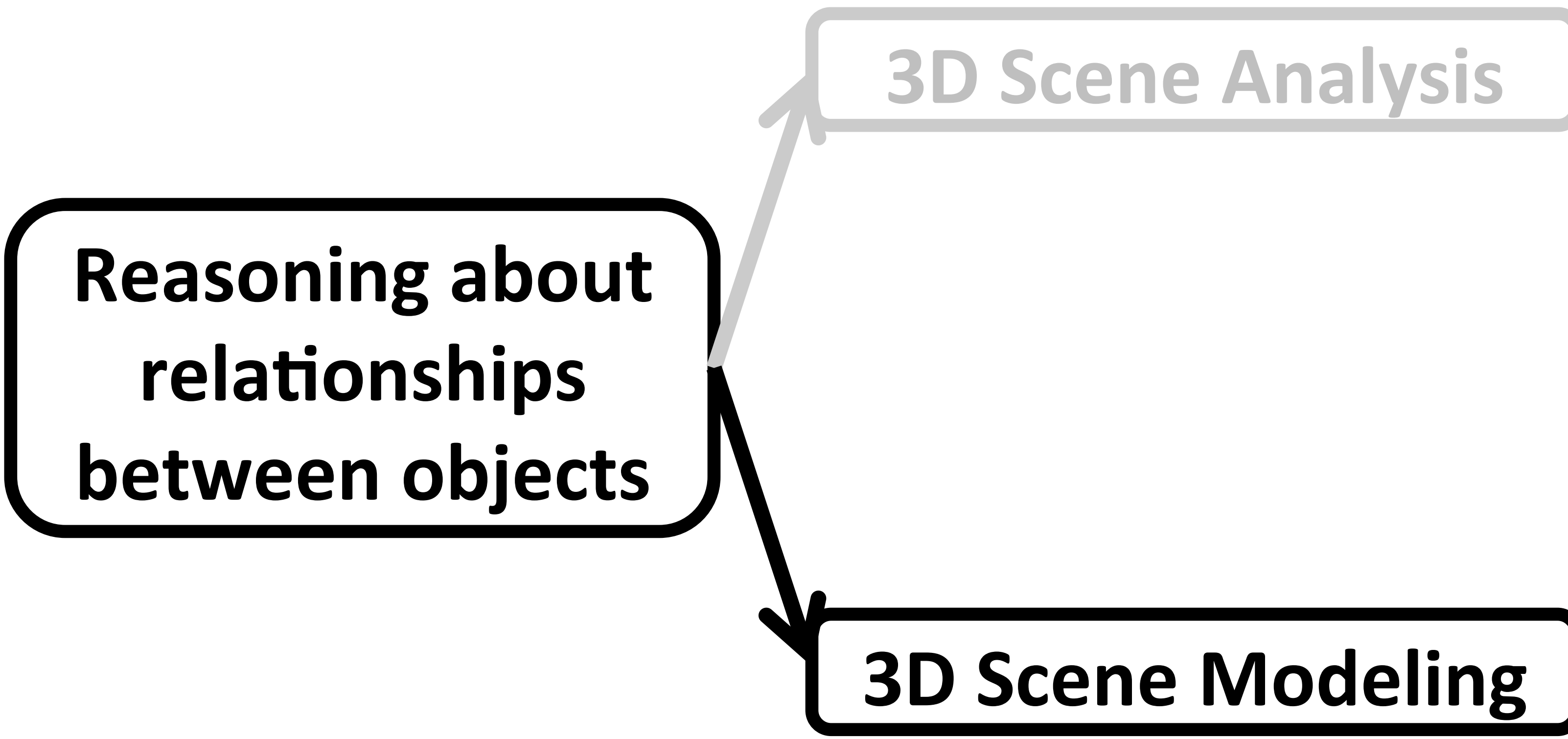
Outline

- Analyzing 3D scenes by modeling hierarchical structure
- Composition-aware scene optimization for product images
- Style compatibility for 3D furniture models



Outline

- Analyzing 3D scenes by modeling hierarchical structure
- Composition-aware scene optimization for product images
- Style compatibility for 3D furniture models

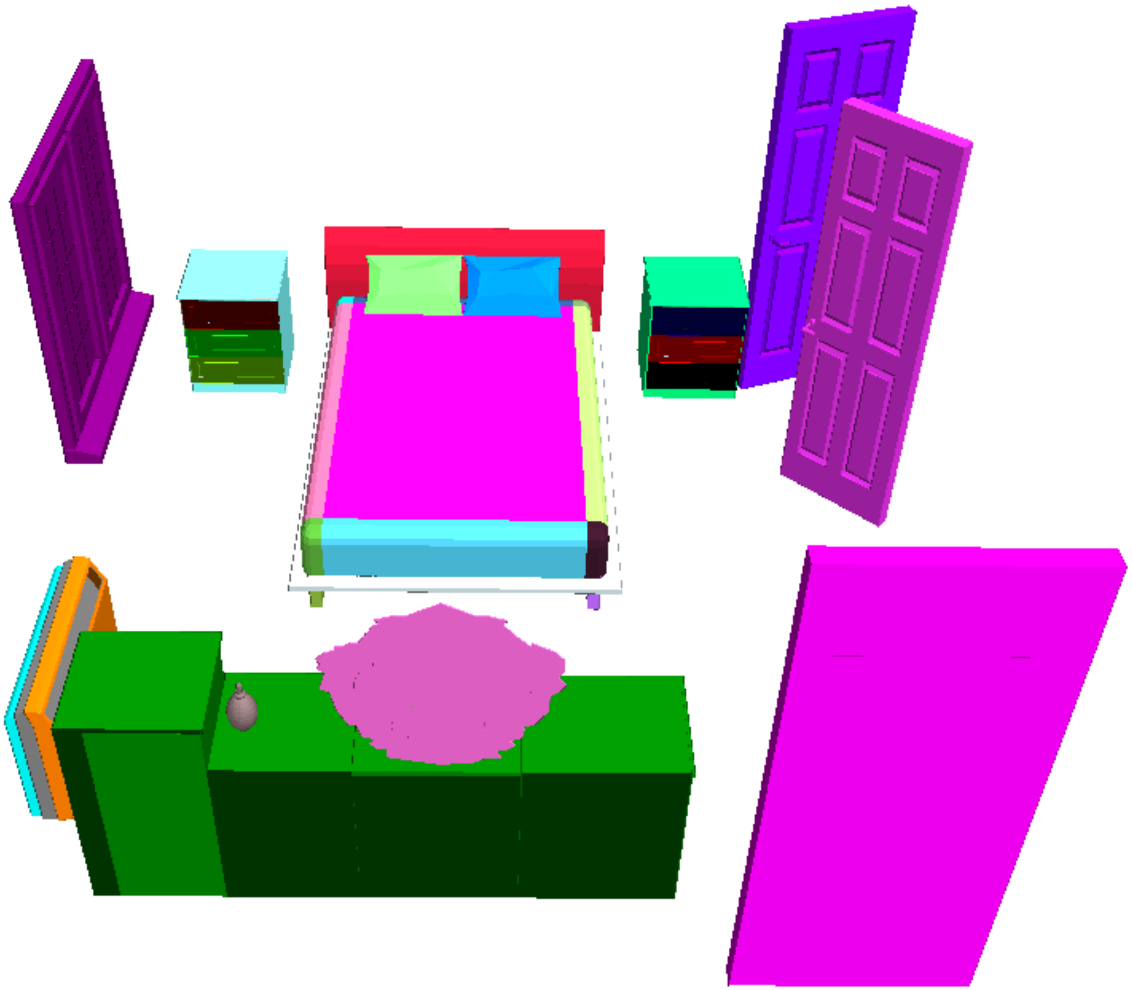


Outline

- Analyzing 3D scenes by modeling hierarchical structure
- Composition-aware scene optimization for product images
- Style compatibility for 3D furniture models

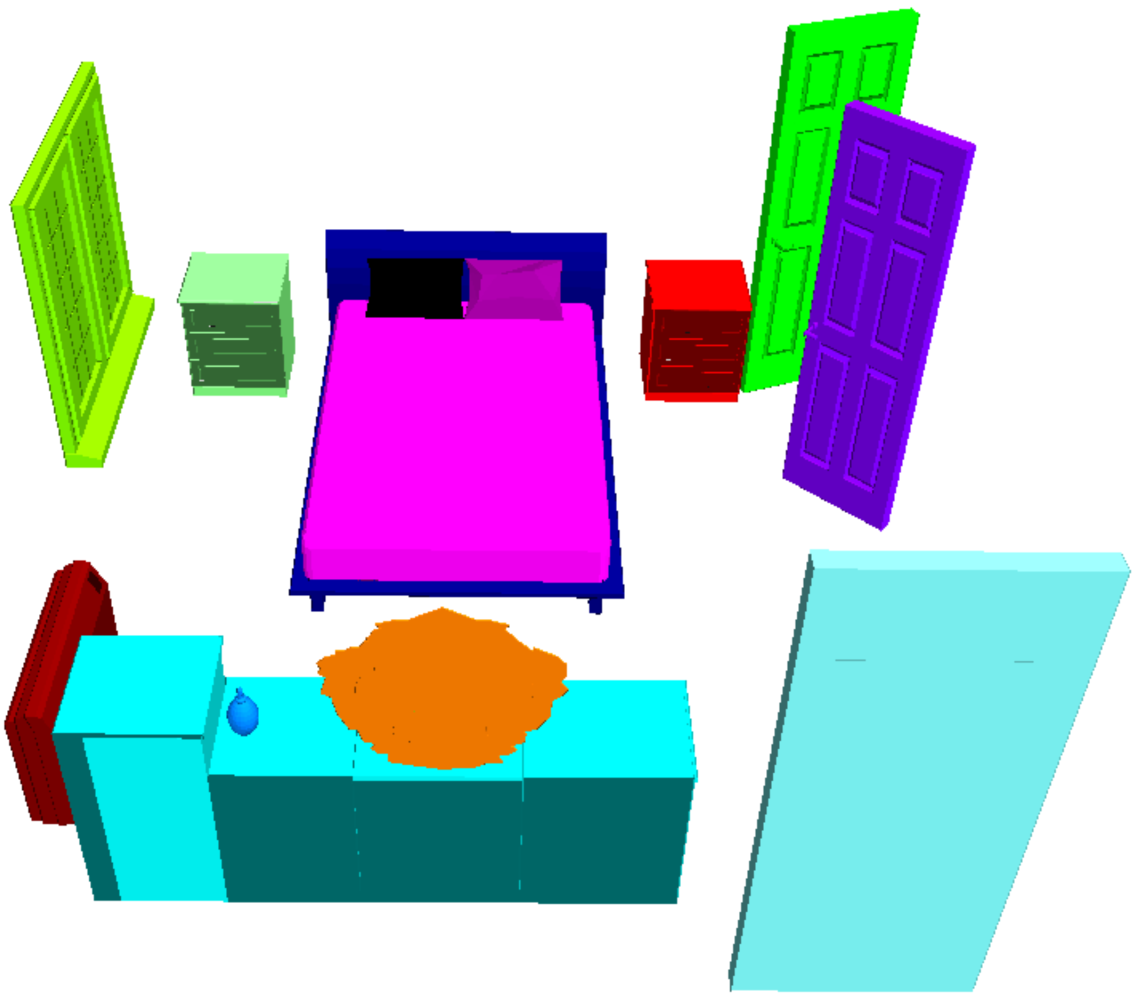
Goal

Input: A scene from Trimble 3D Warehouse



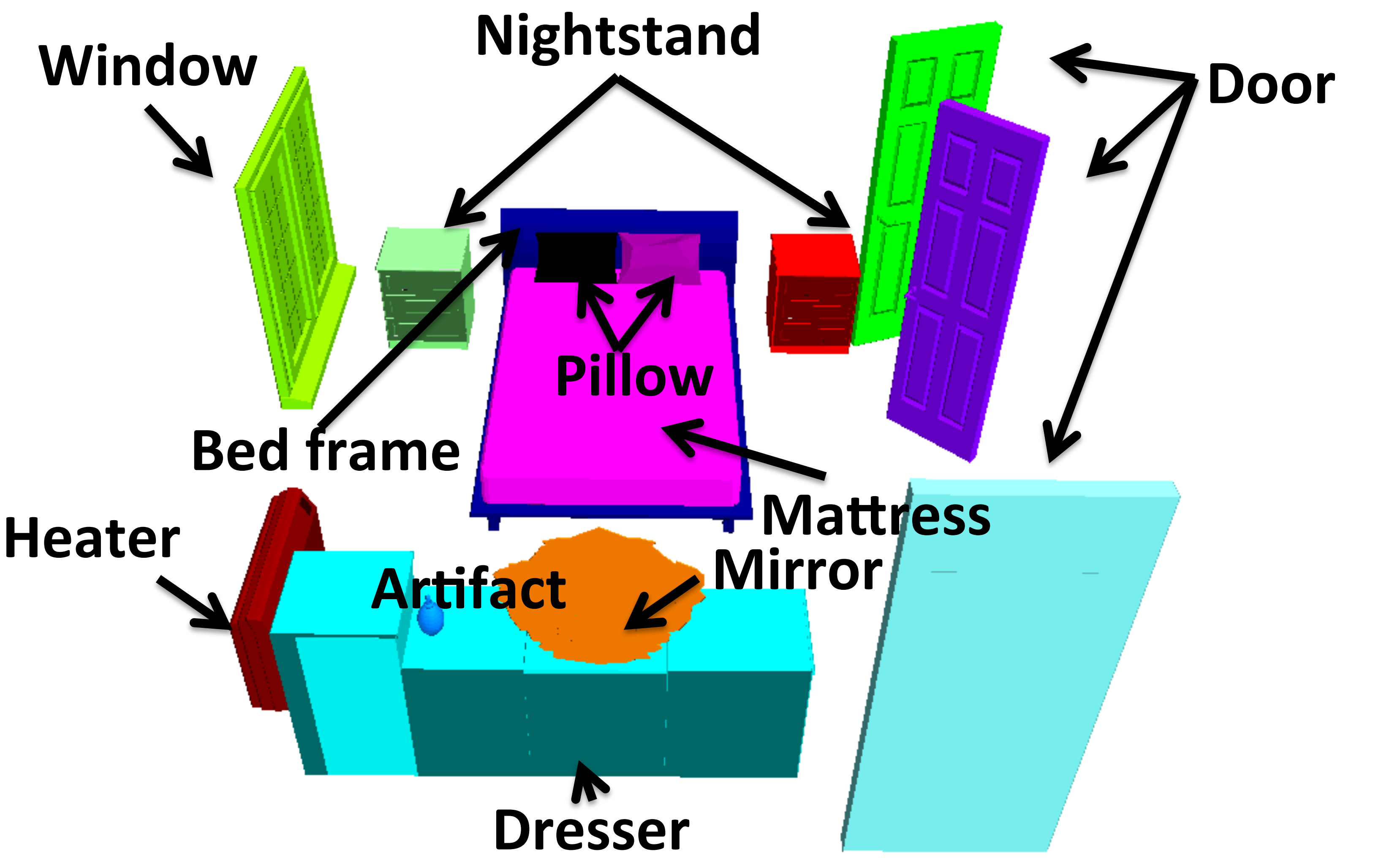
Goal

Output 1: Semantic segmentations



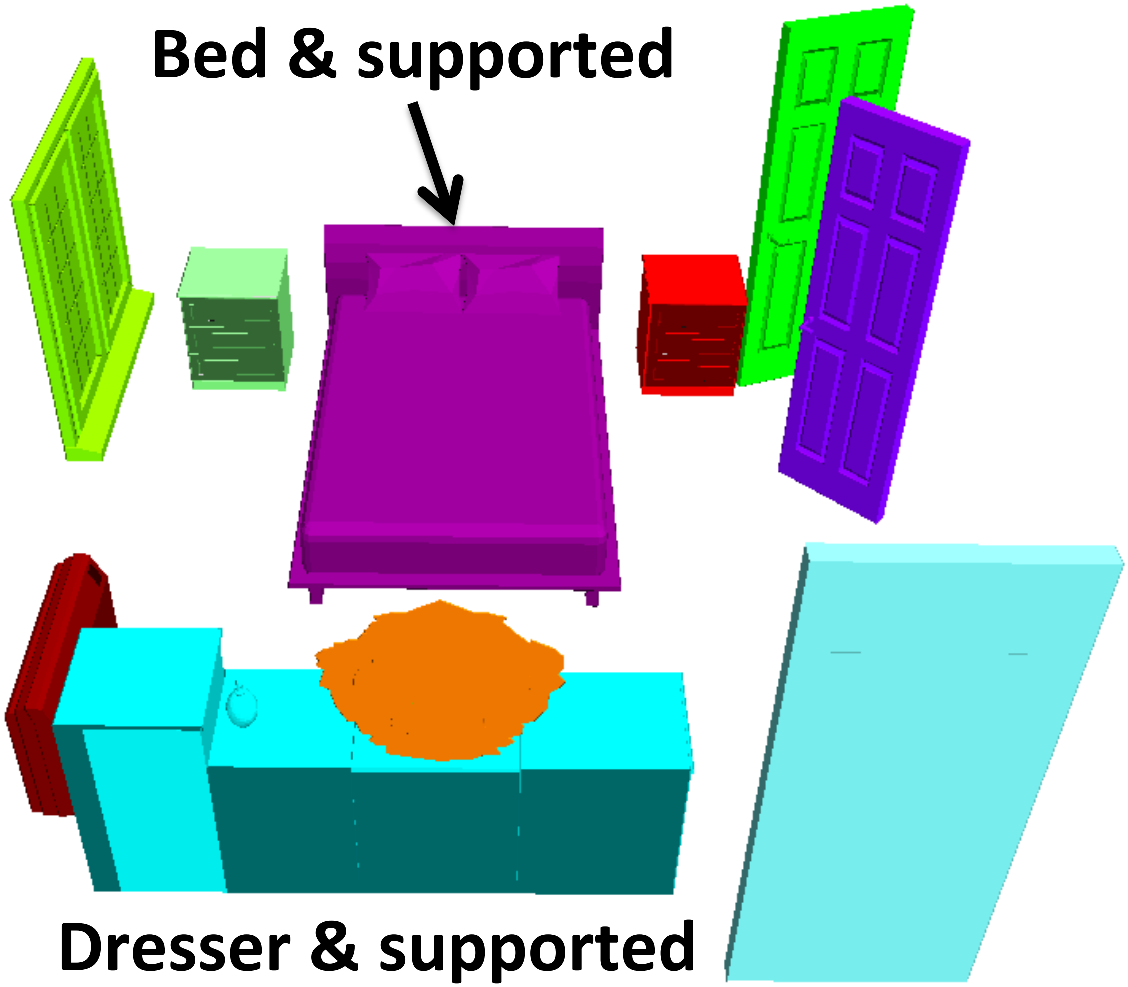
Goal

Output 2: Category labels.



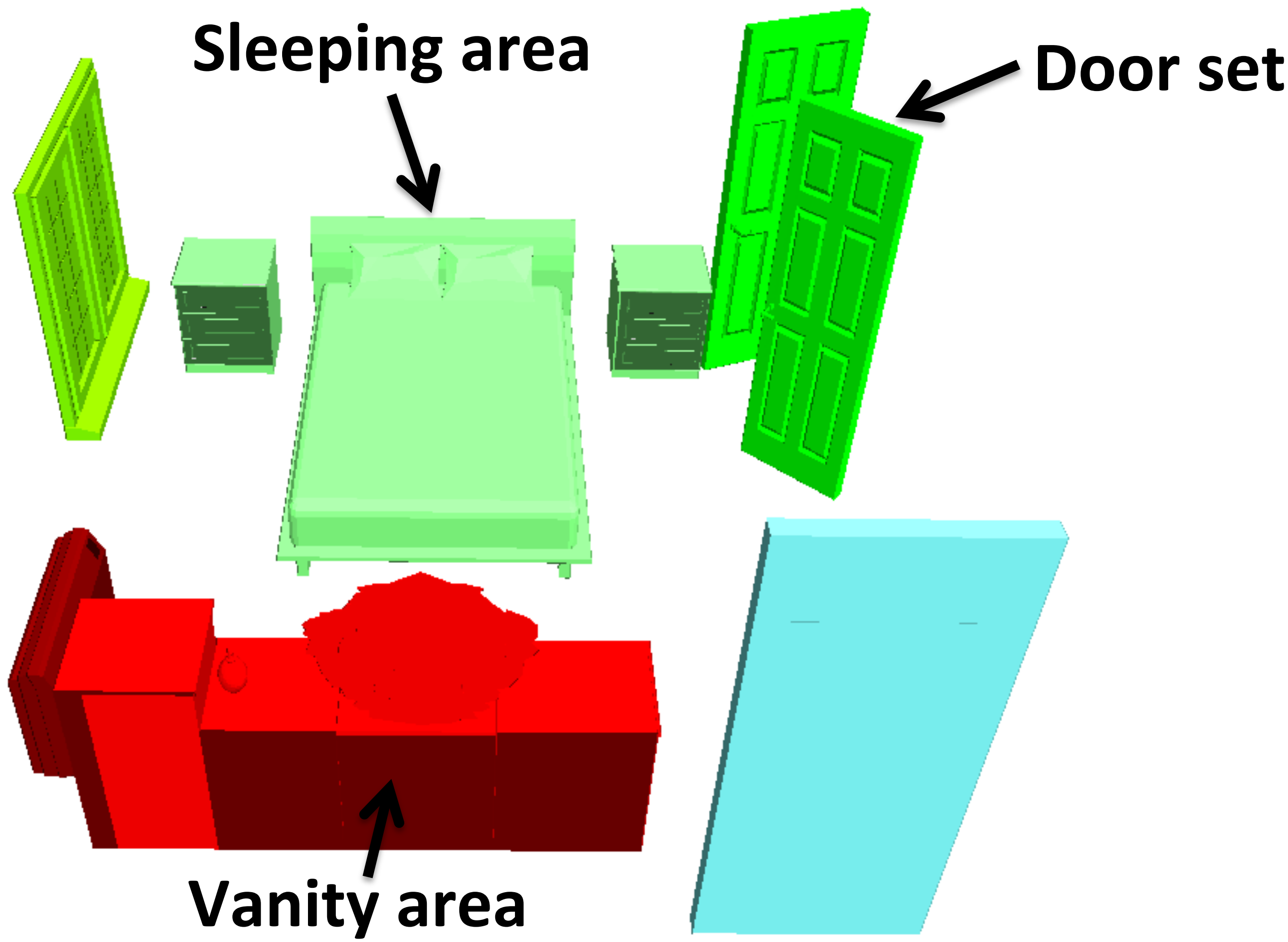
Goal

Output 2: Category labels at different levels.



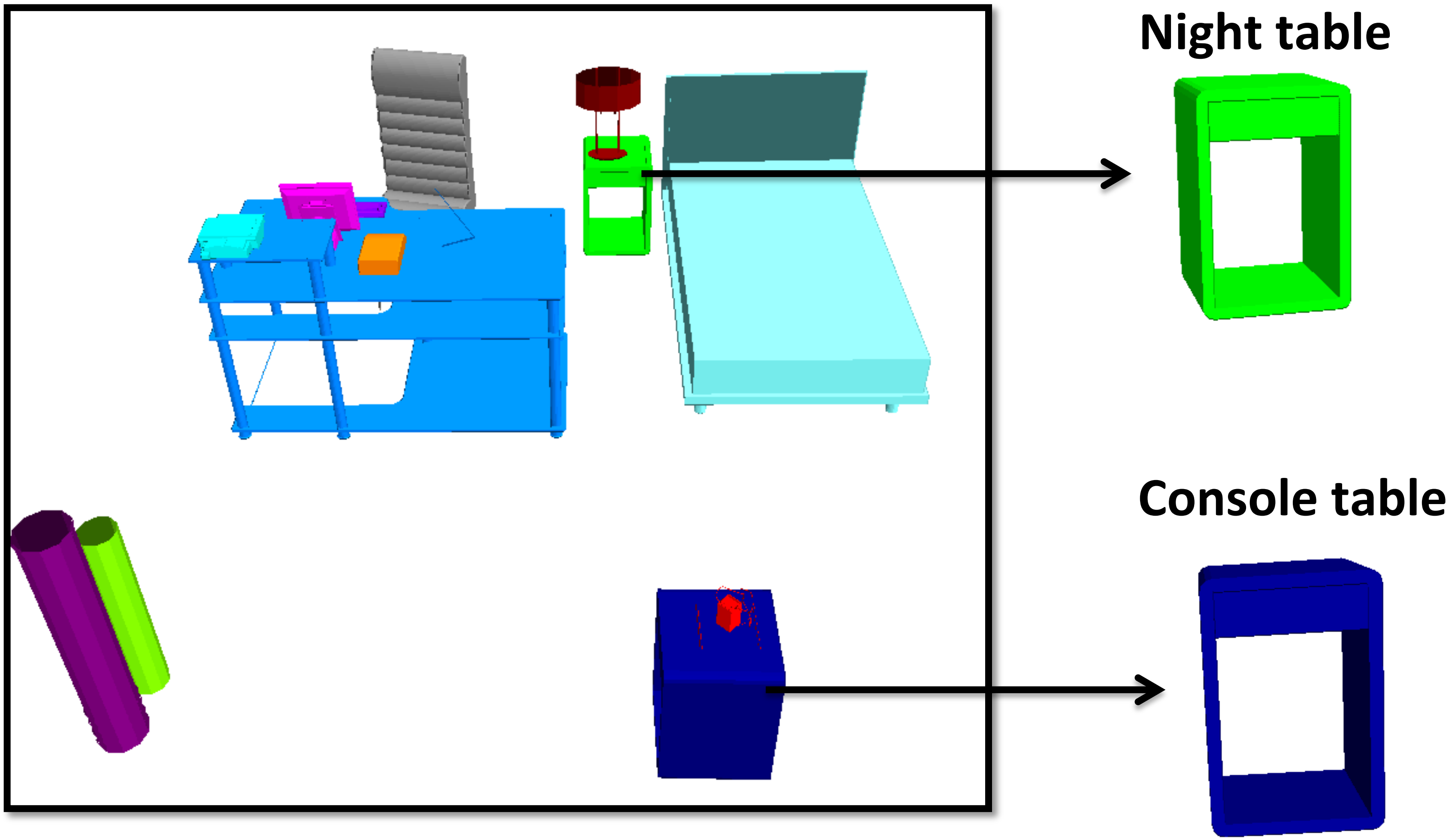
Goal

Output 2: Category labels at different levels.



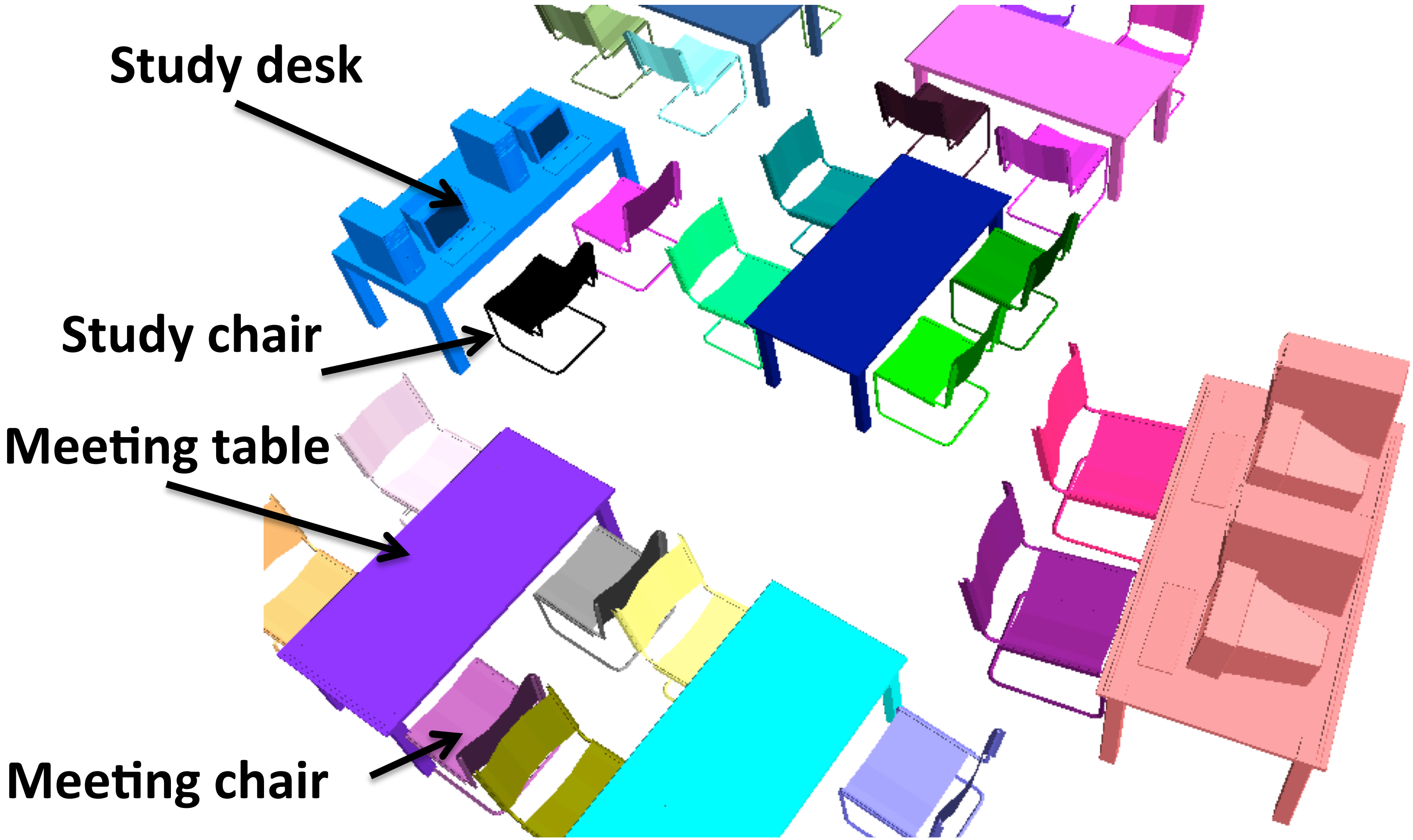
Challenges

Shape is not distinctive.



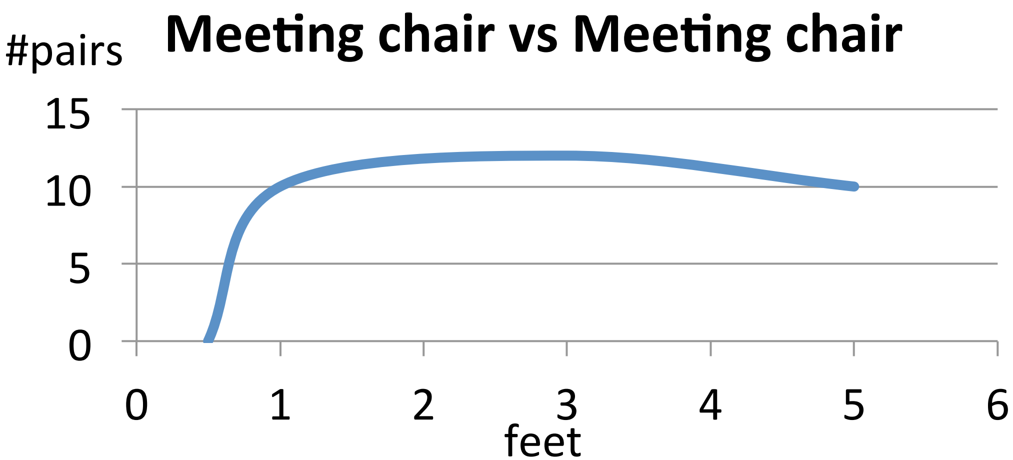
Challenges

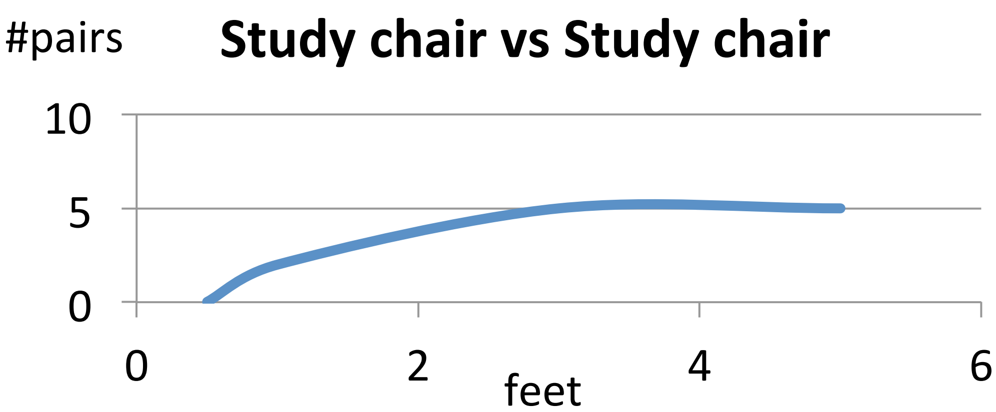
Contextual information



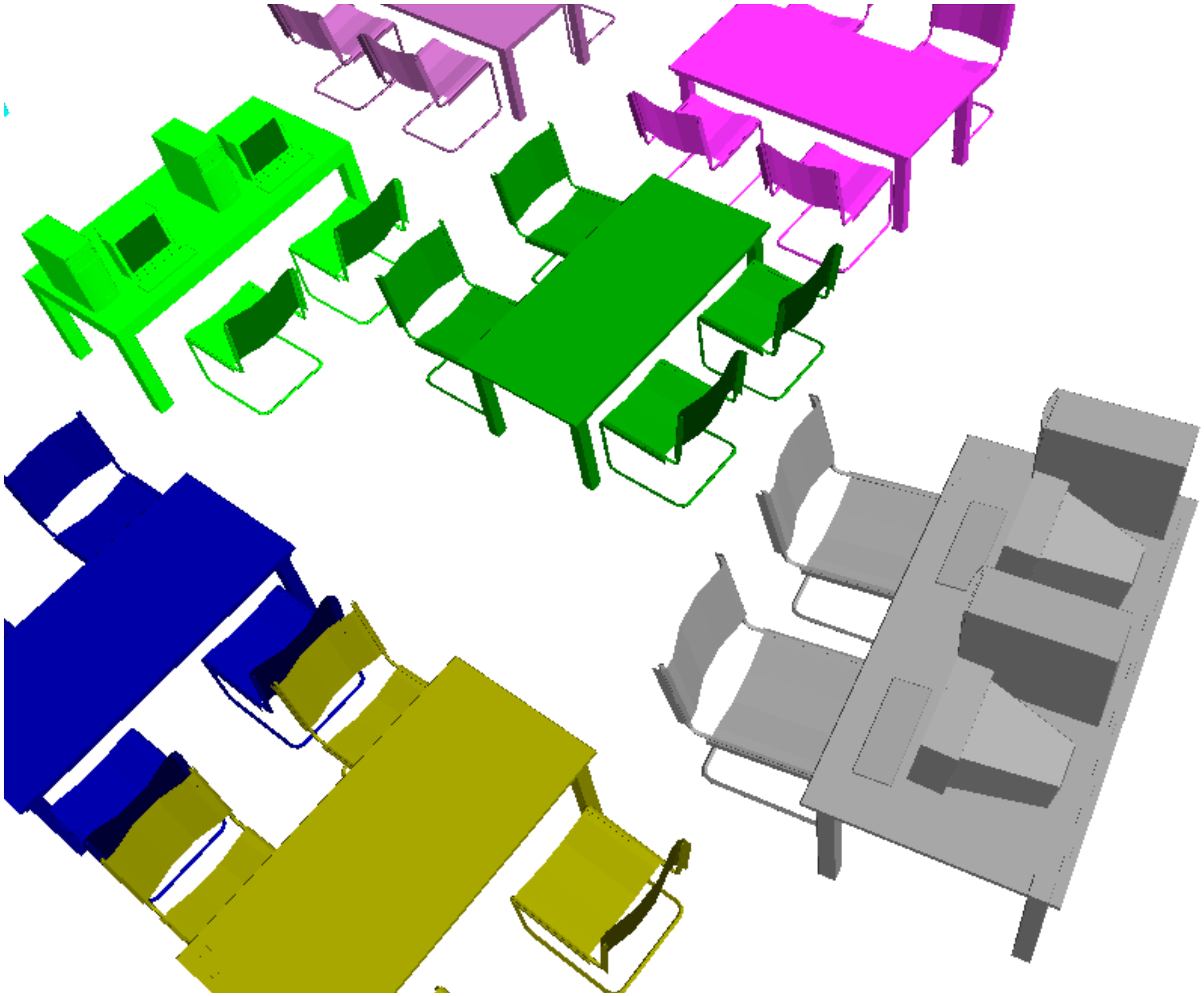
Challenges

All-pair contextual information is not distinctive.

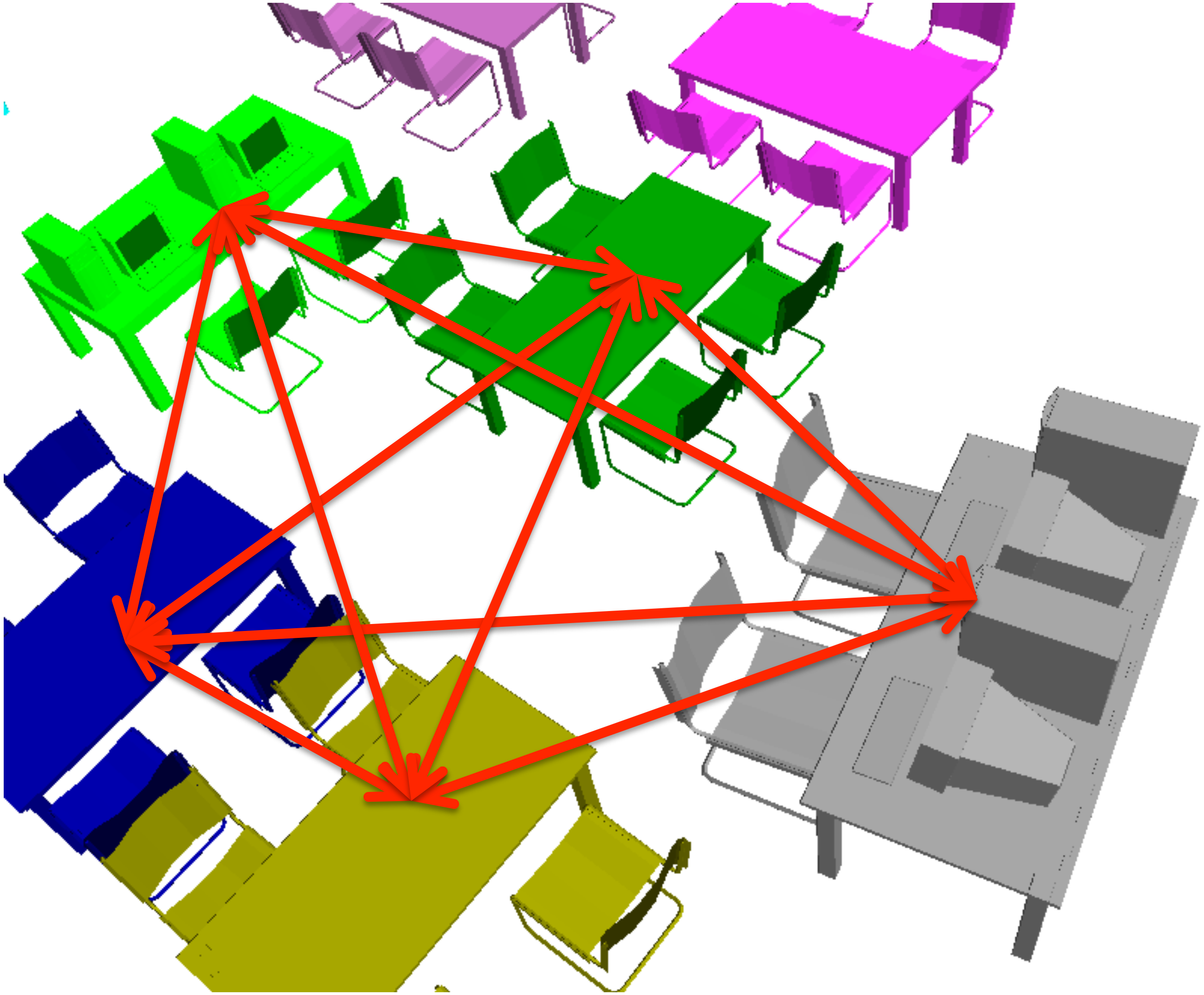




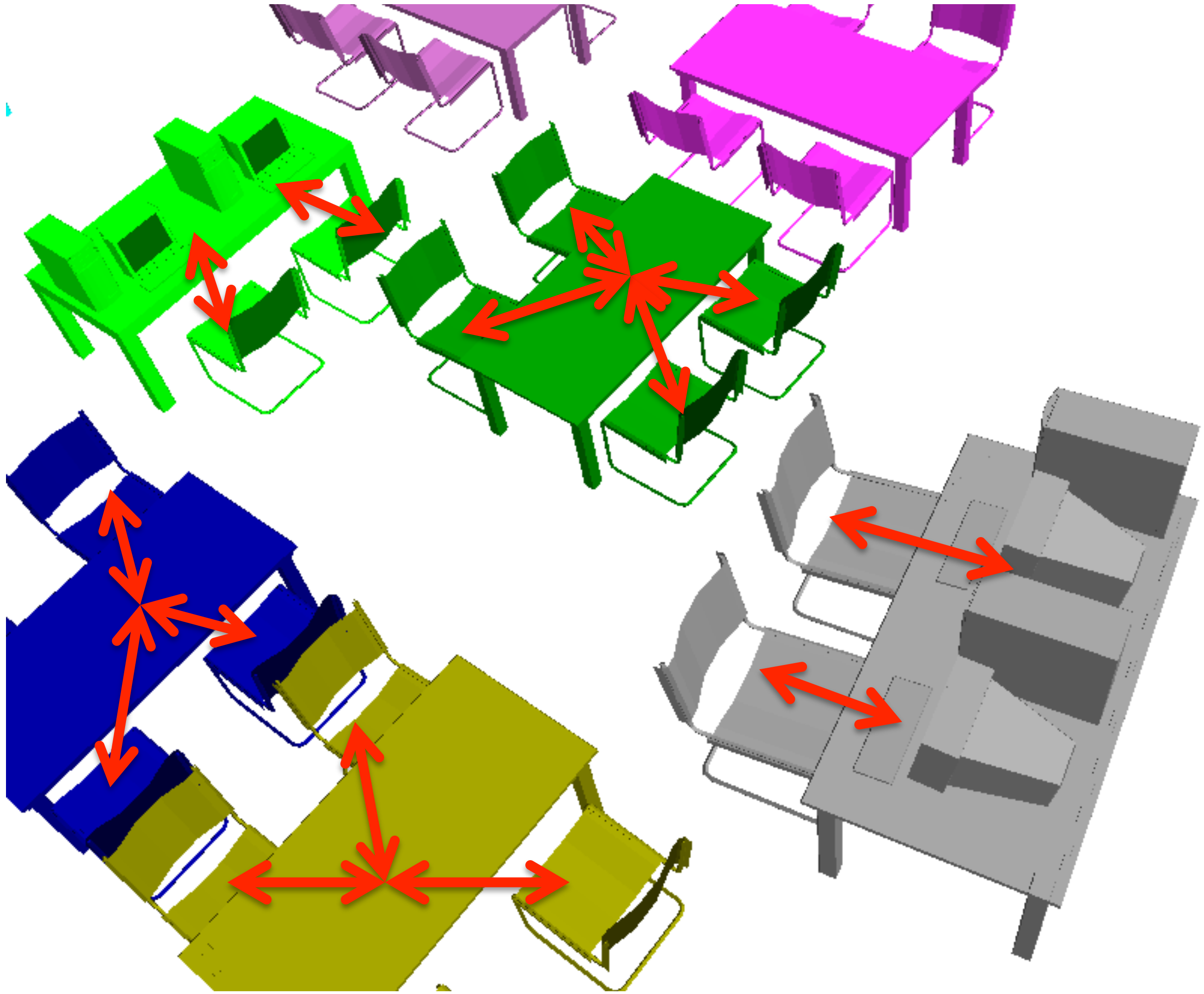
Challenges



Challenges

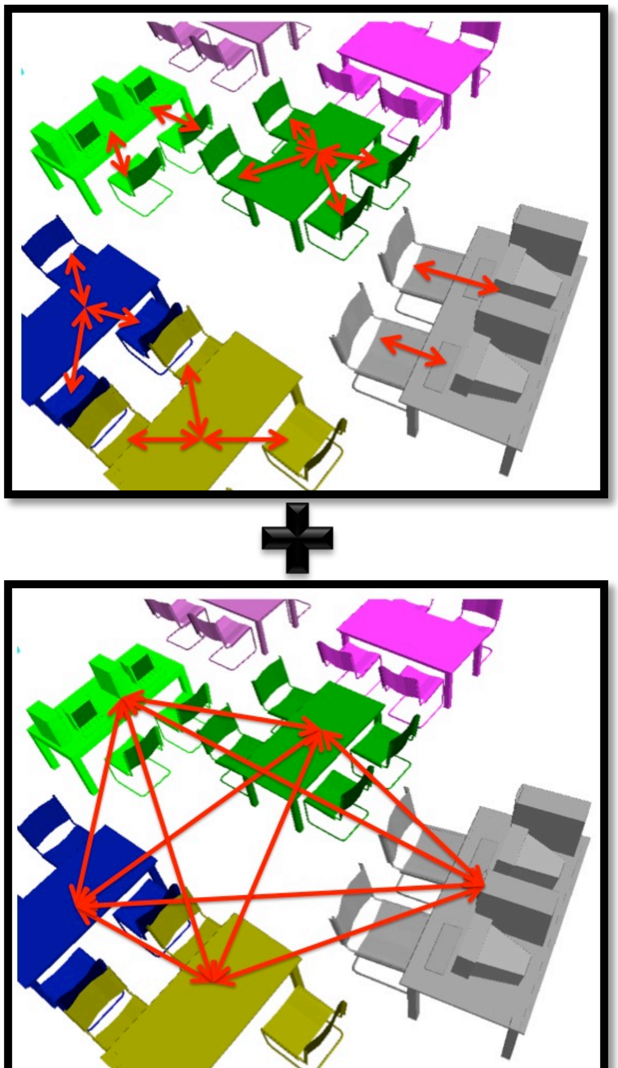


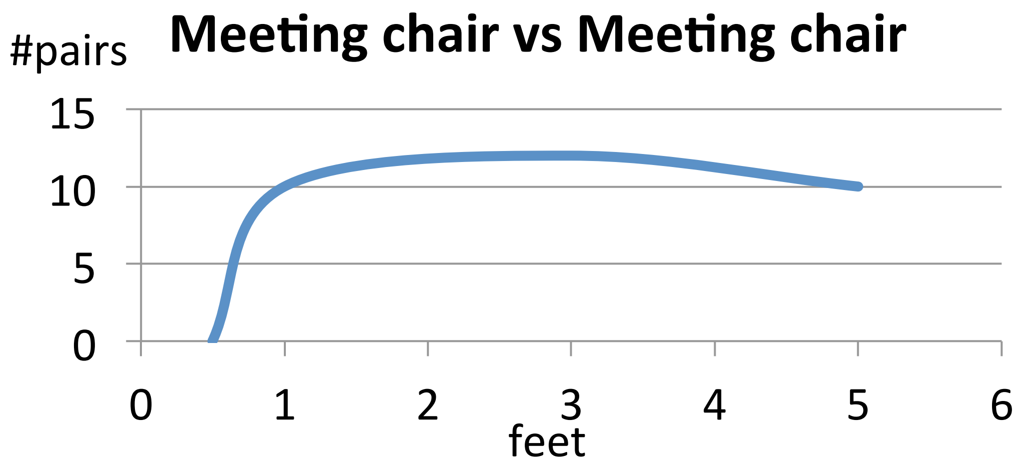
Challenges

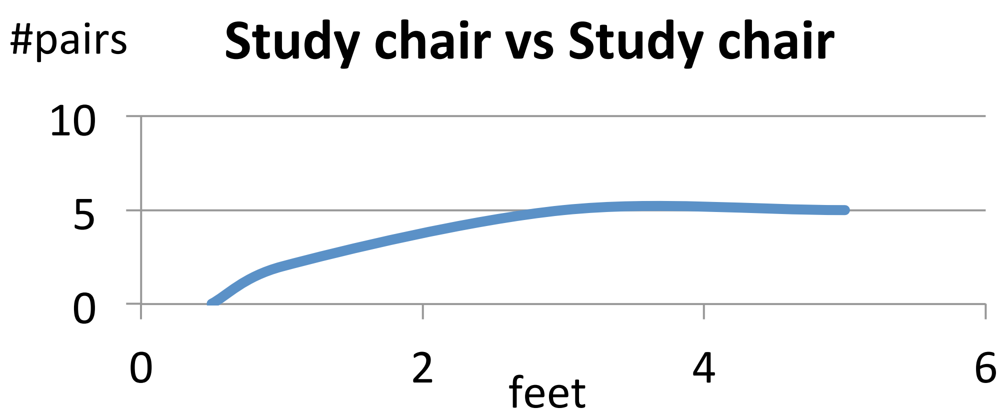


Challenges

All-pair contextual information is not distinctive.

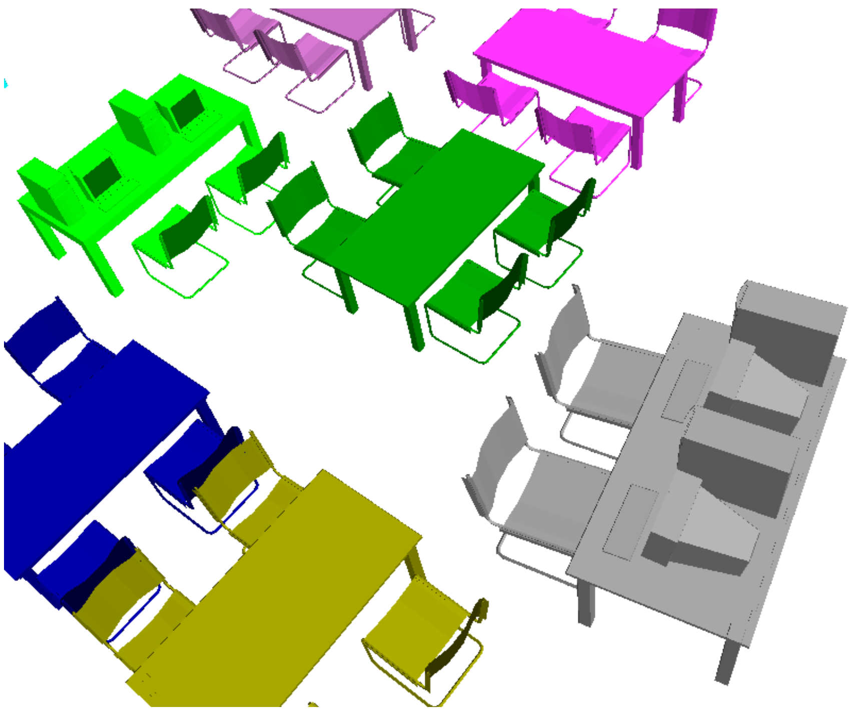




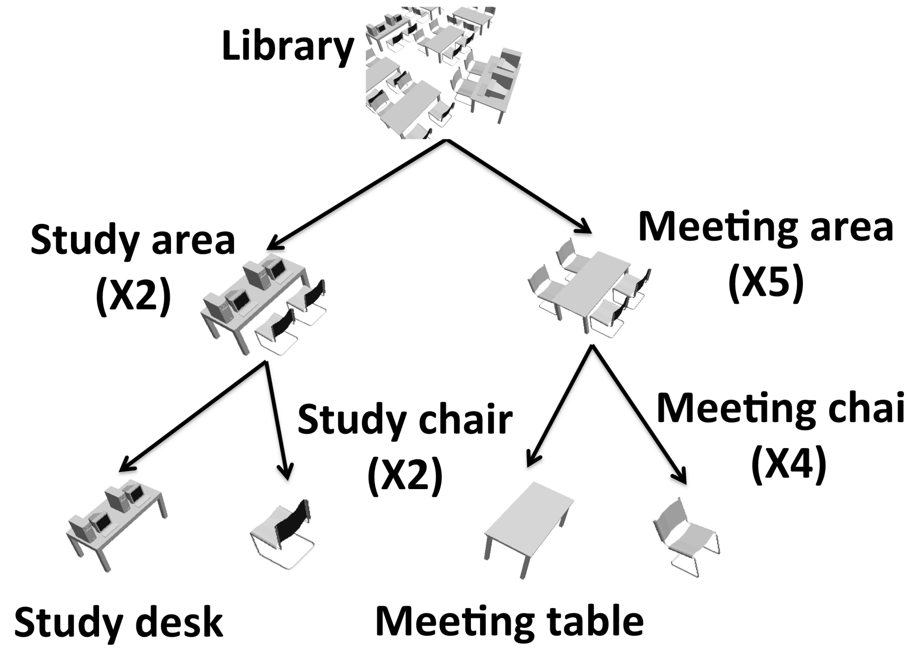


Key Idea

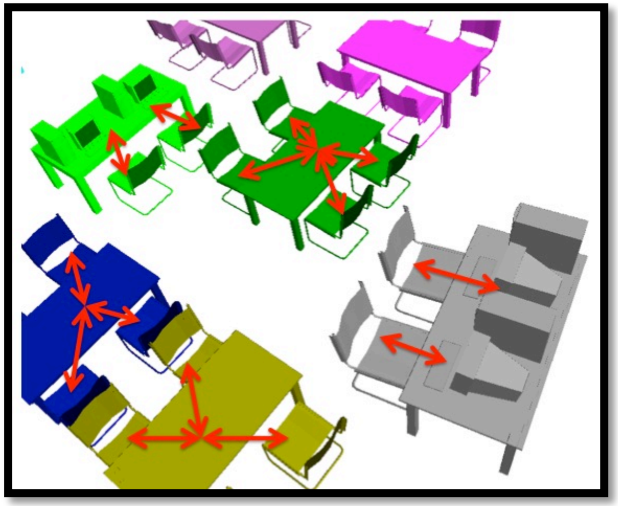
Semantic groups

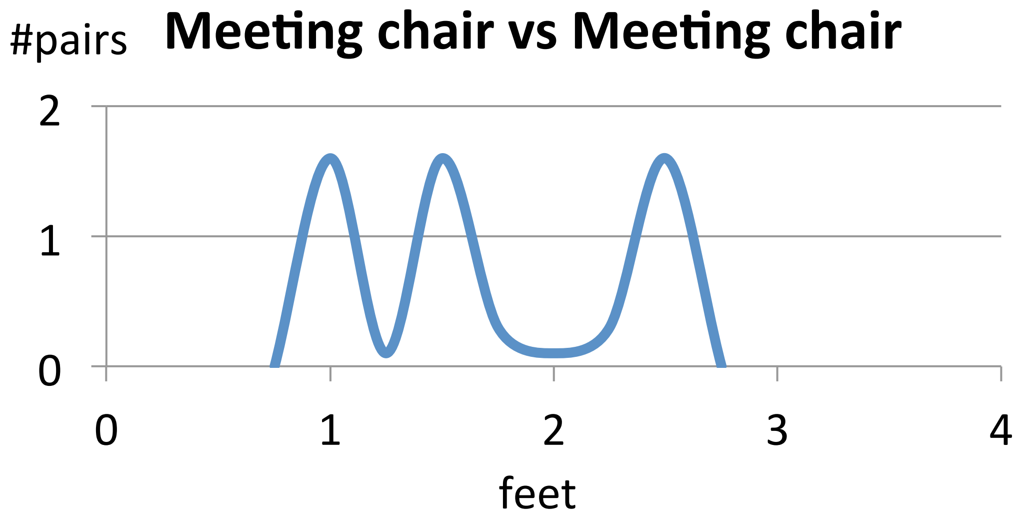


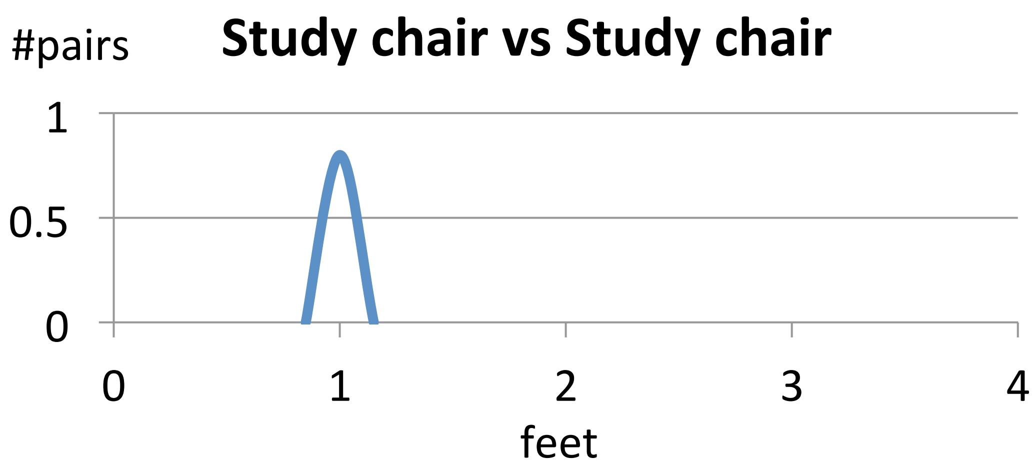
Semantic hierarchy



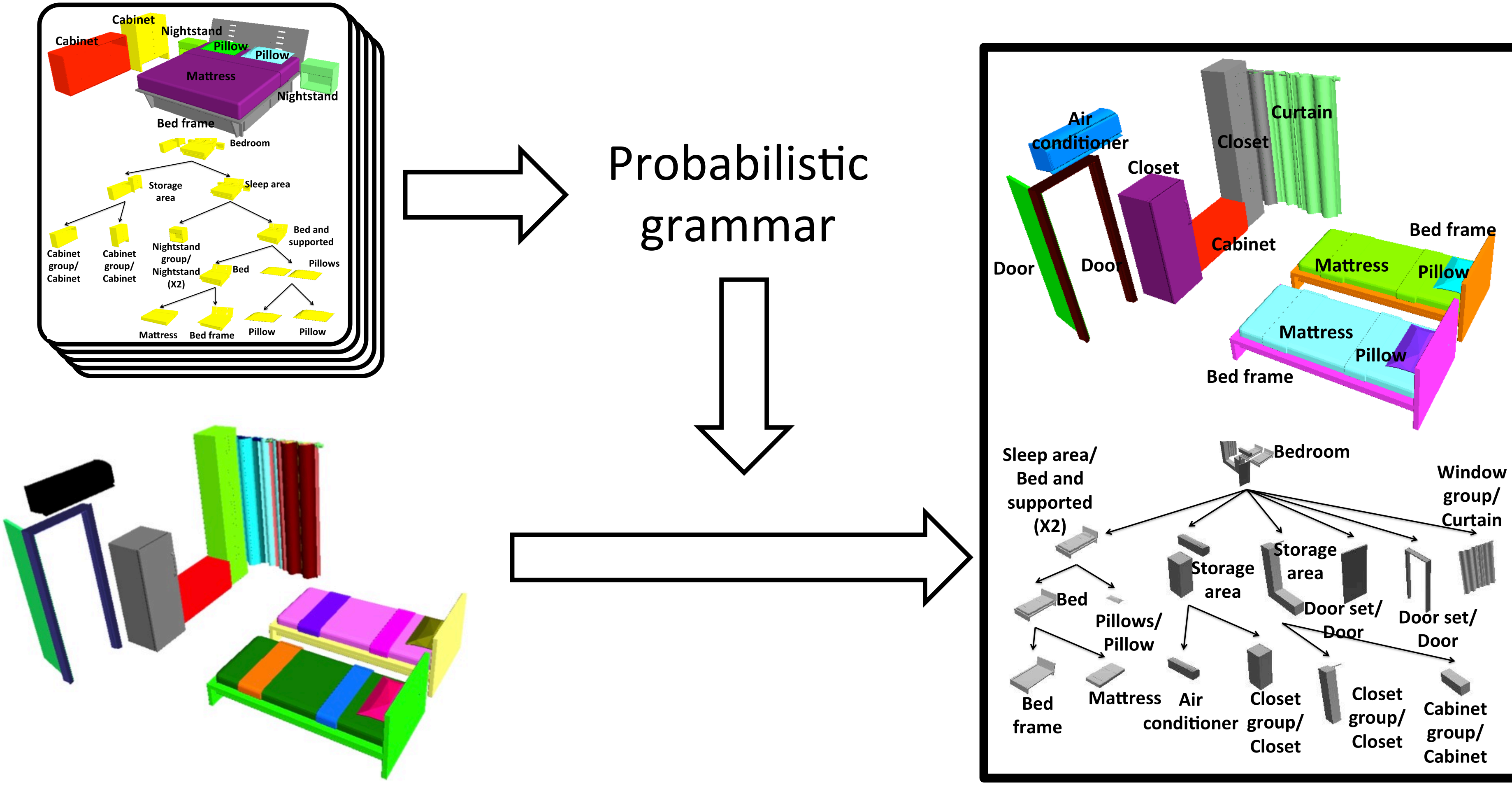
Key idea




#pairs
Meeting chair vs Meeting chair
2
1
0
0
1
2
3
4
feet


#pairs
Study chair vs Study chair
1
0.5
0
0
1
2
3
4
feet

Pipeline

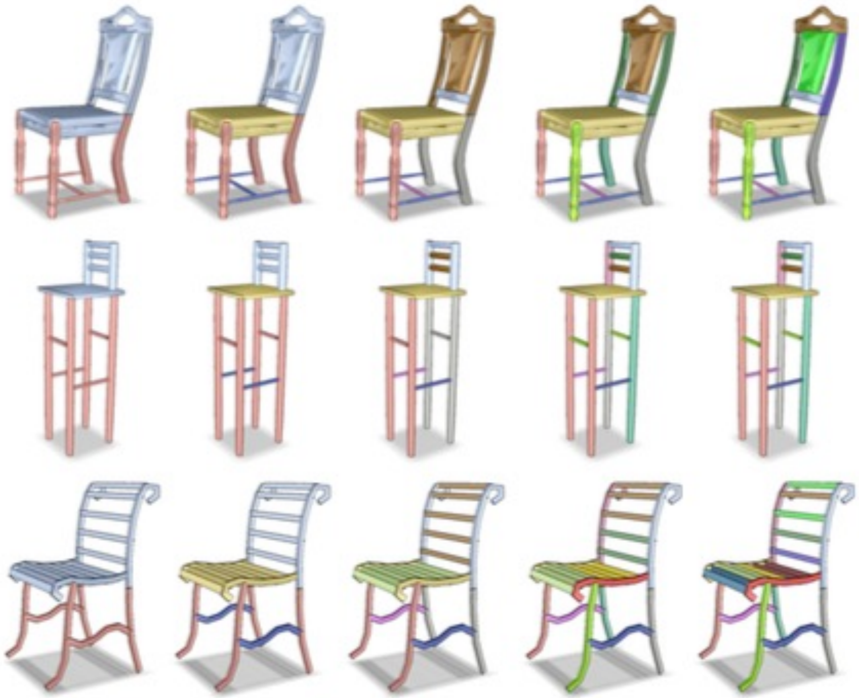


Cabinet
Cabinet
Nightstand
Pillow
Pillow
Mattress
Nightstand
Bed frame
Bedroom
Storage area
Sleep area
Cabinet group/ Cabinet
Cabinet group/ Cabinet
Nightstand group/ Nightstand (X2)
Bed and supported
Bed
Pillows
Mattress
Bed frame
Pillow
Pillow
Probabilistic grammar
Air conditioner
Closet
Curtain
Closet
Cabinet
Door
Door
Bed frame
Mattress
Pillow
Mattress
Pillow
Bed frame
Sleep area/ Bed and supported (X2)
Bedroom
Window group/ Curtain
Storage area
Storage area
Bed
Pillows/ Pillow
Door set/ Door
Door set/ Door
Bed frame
Mattress
Air conditioner
Closet group/ Closet
Closet group/ Closet
Cabinet group/ Cabinet

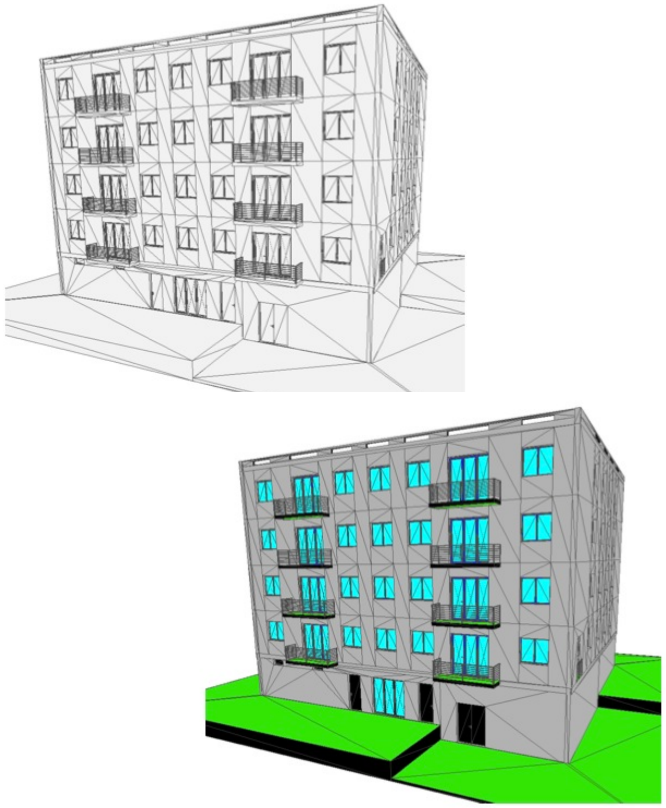
Related work

Van Kaick et al. 2013

Related work



Van Kaick et al. 2013



Boulch et al. 2013

Overview

Grammar Structure

Learning a Probabilistic Grammar

Scene Parsing

Results

Probabilistic grammar

Labels

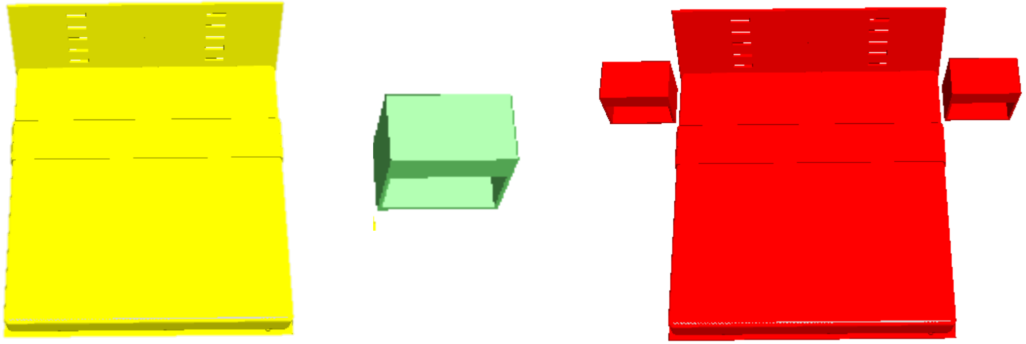
Rules

Probabilities

Labels

Examples:

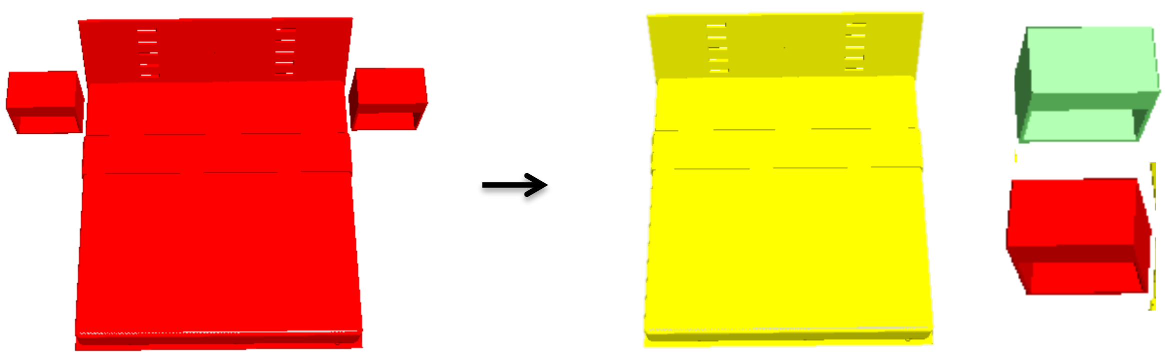
bed, night table, sleeping area



Rules

Example:

sleeping area $\rightarrow$ bed, night table



Probabilities

Derivation probabilities

Cardinality probabilities

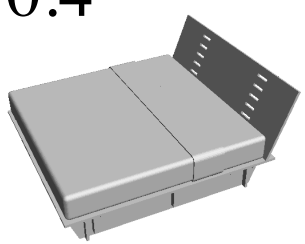
Geometry probabilities

Spatial probabilities

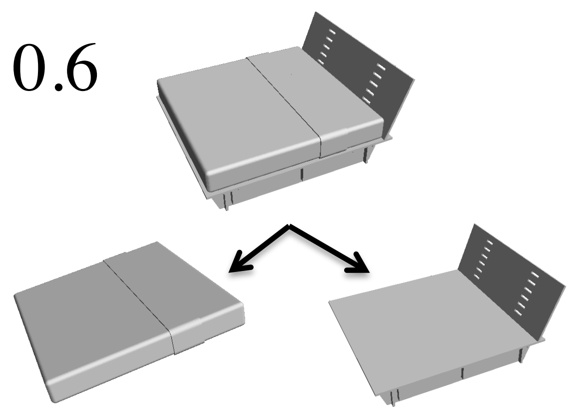
Derivation probability P_{nt}

$$\text{bed} \xrightarrow{0.6} \text{bed frame, mattress}$$

$P = 0.4$

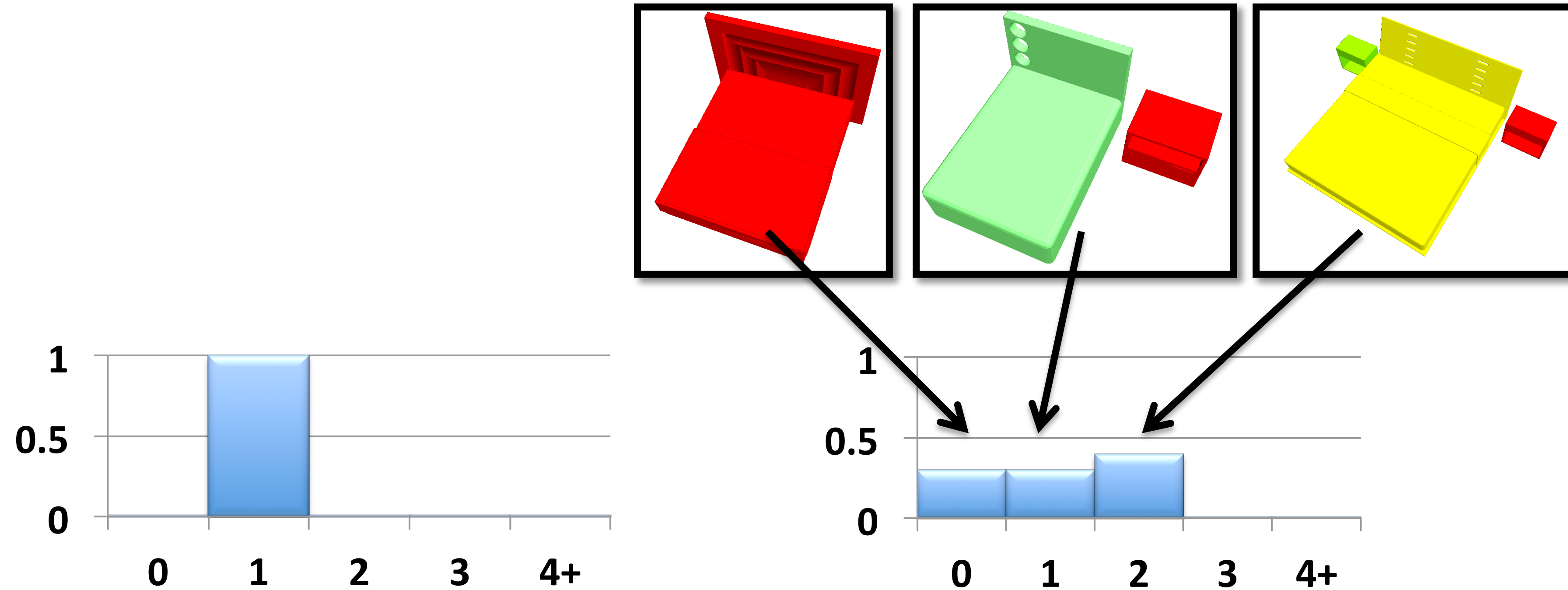


$P = 0.6$



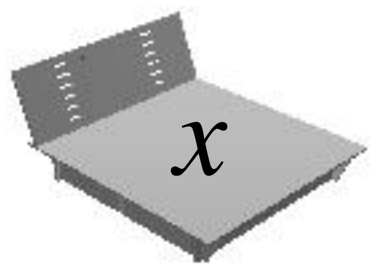
Cardinality probability P_{card}

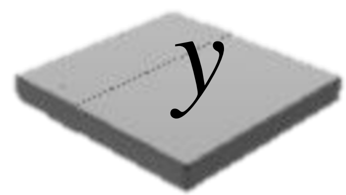
sleeping area → bed, night table



$P_{card}(bed \mid sleepingarea)$ $P_{card}(nighttable \mid sleepingarea)$

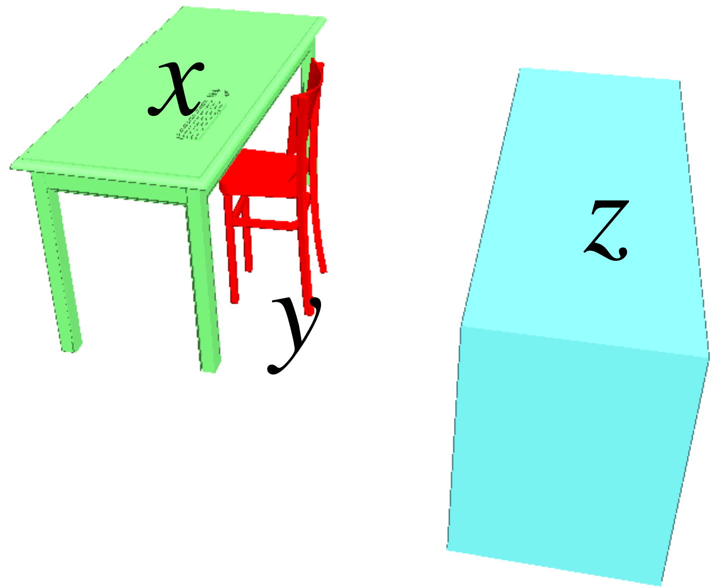
Geometry probability P_g





$$P_g(x \mid bedframe) > P_g(y \mid bedframe)$$

Spatial probability P_s



$$P_s(x, y \mid desk, chair, studyarea) > P_s(z, y \mid desk, chair, studyarea)$$

Overview

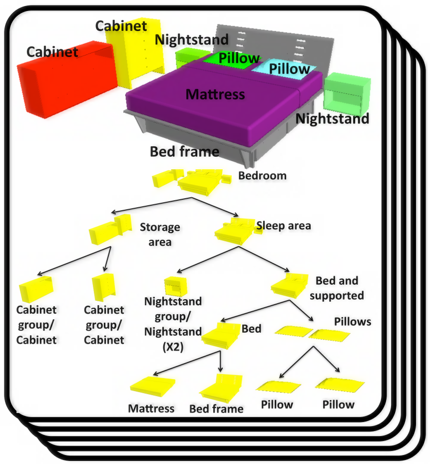
Grammar Structure

→ Learning a Probabilistic Grammar

Scene Parsing

Results

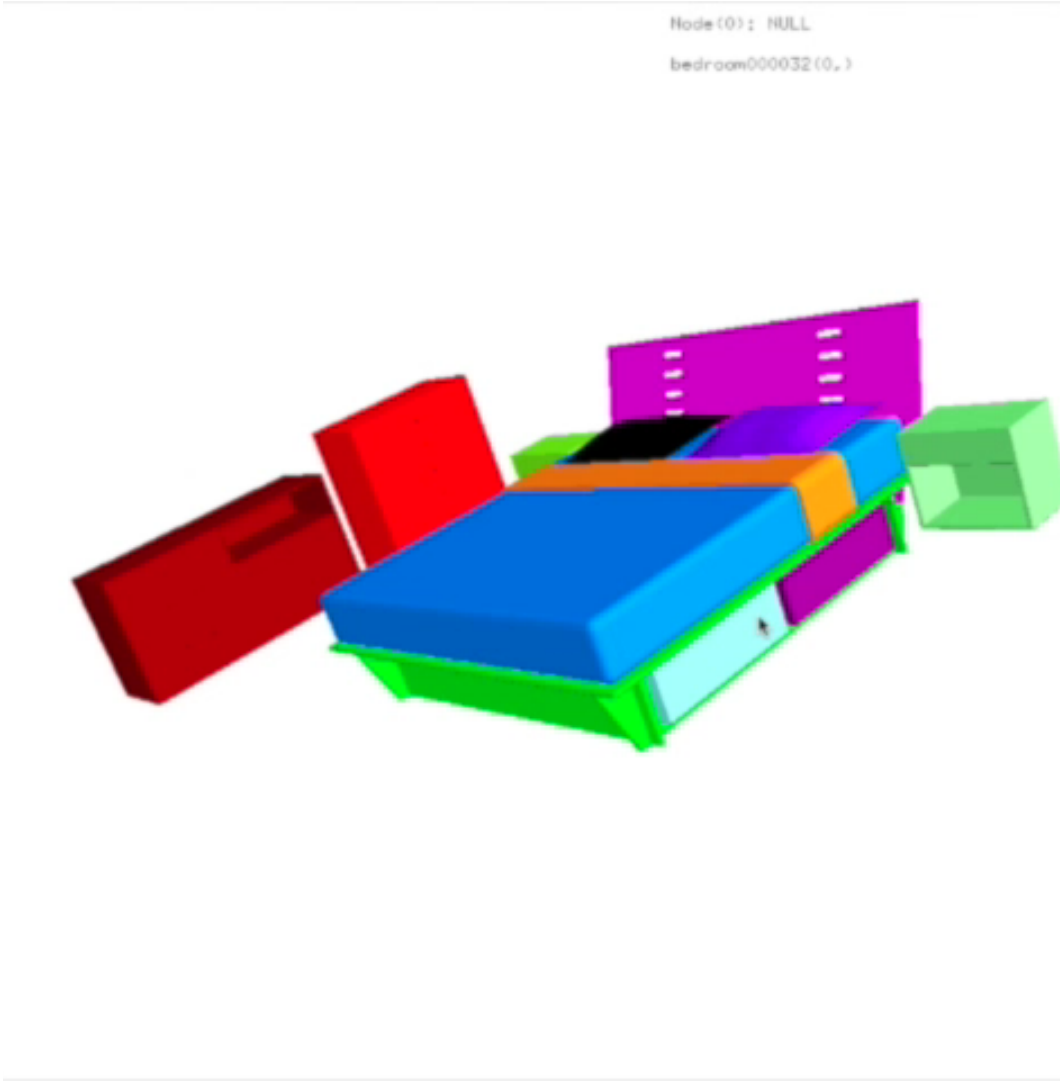
Pipeline



Probabilistic grammar

Learning a probabilistic grammar

Identify objects



Speed X 5

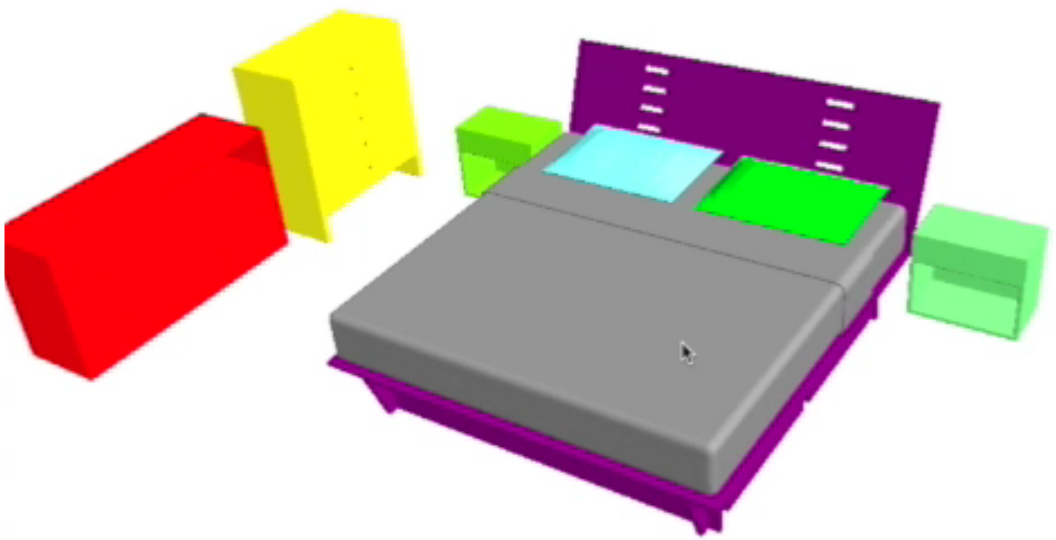
Learning a probabilistic grammar

Label objects

Speed X 5

Learning a probabilistic grammar

Group objects



Speed X 5

Learning a probabilistic grammar

Grammar generation

→ Labels all unique labels

Rules

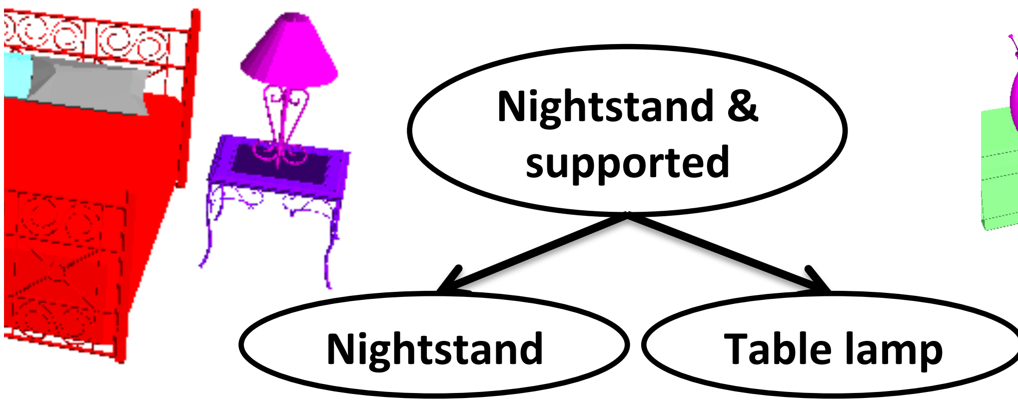
Probabilities

Learning a probabilistic grammar

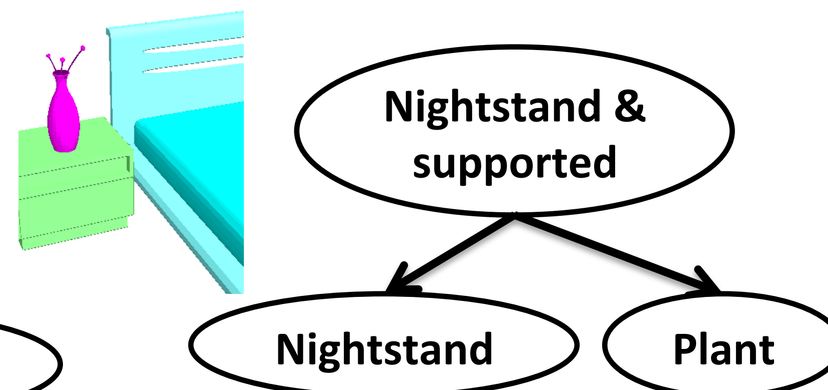
Grammar generation

- Labels
- Rules — concatenating all children for each label
- Probabilities

Training example 1:



Training example 2:



Learning a probabilistic grammar

Grammar generation

Labels

Rules

→ Probabilities

P_{nt}, P_{card} : learning from occurrence statistics

P_g : estimating Gaussian kernels

P_s : kernel density estimation

Overview

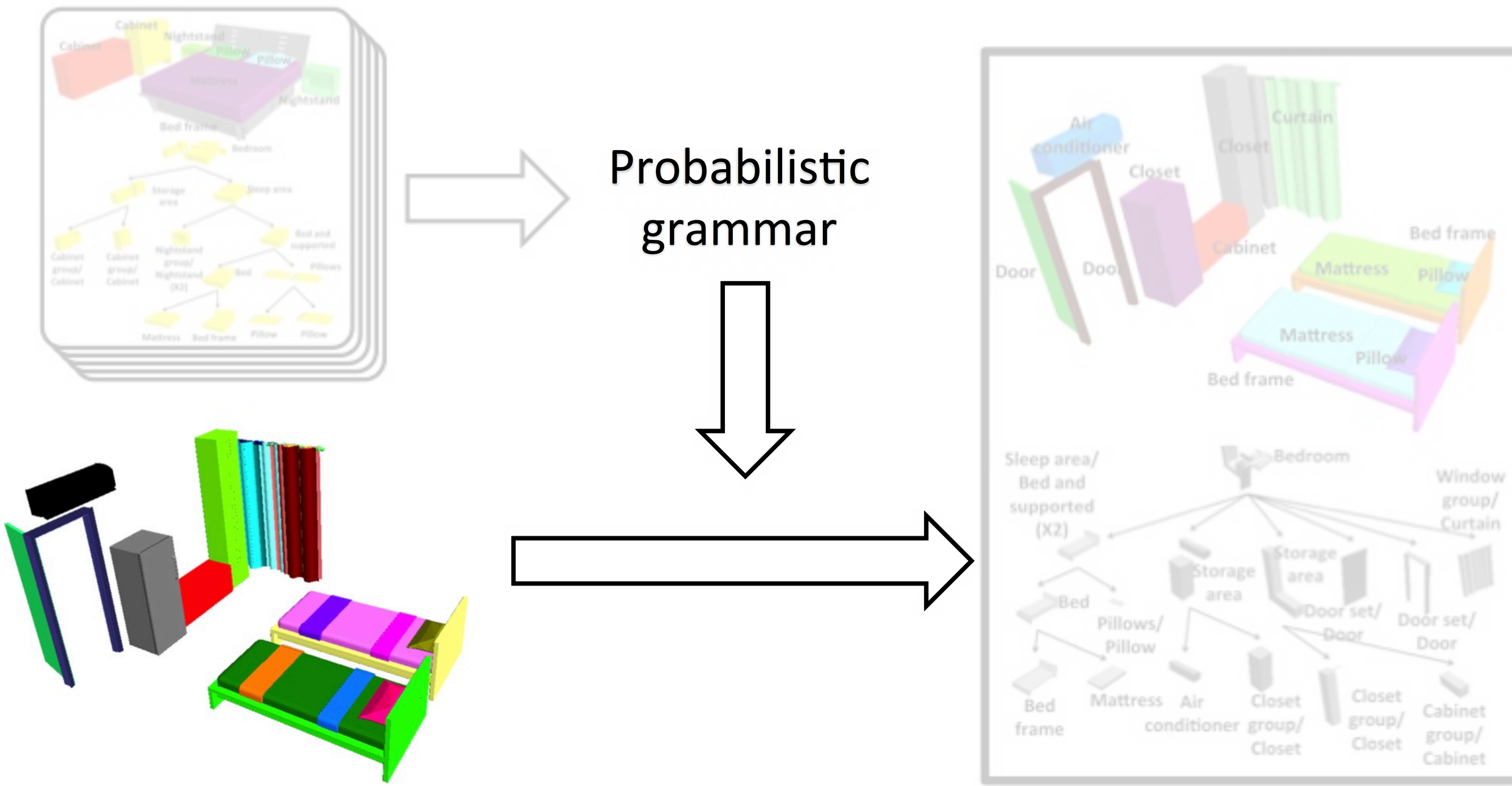
Grammar Structure

Learning a Probabilistic Grammar

→ Scene Parsing

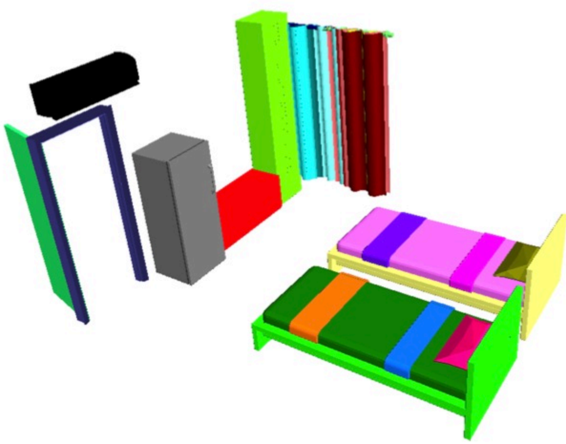
Results

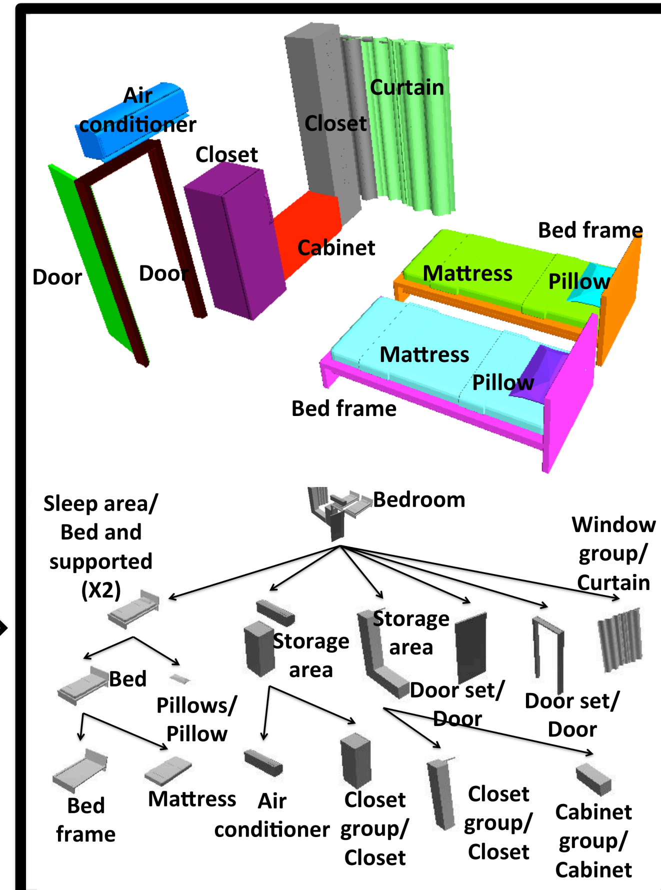
Pipeline


Probabilistic grammar
Air conditioner
Closet
Curtain
Door
Cabinet
Bed frame
Mattress
Pillow
Sleep area/ Bed and supported (X2)
Bedroom
Window group/ Curtain
Storage area
Bed
Pillows/ Pillow
Door set/ Door
Air conditioner
Closet group/ Closet
Cabinet group/ Cabinet

Pipeline

Probabilistic grammar





Scene parsing

Objective function

$$H^* = \text{argmax}_H \, P(H \mid S, G)$$

- H is the unknown hierarchy
- S is the input scene
- G is the probabilistic grammar

Scene parsing

After applying Bayes' rule

$$H^* = \text{argmax}_H\, P(H \mid G)P(S \mid H, G)$$

Scene parsing

After applying Bayes' rule

$$H^* = \operatorname{argmax}_H P(H \mid G) P(S \mid H, G)$$

Prior of hierarchy $$P(H \mid G) = \prod_{x \in H} P_{prod}(x)^{T(x)}$$

Scene parsing

After applying Bayes' rule

$$H^* = \arg\max_H P(H \mid G) P(S \mid H, G)$$

Prior of hierarchy $$P(H \mid G) = \prod_{x \in H} P_{prod}(x)^{T(x)}$$

$P_{prod}(x)$: probability of a single derivation

formulated using P_{nt}, P_{card}

Scene parsing

After applying Bayes' rule

$$H^* = \operatorname{argmax}_H P(H \mid G) P(S \mid H, G)$$

Prior of hierarchy $\quad P(H \mid G) = \prod_{x \in H} P_{prod}(x)^{T(x)}$

$T(x)$ compensates for decreasing probability as H has more internal nodes.

Scene parsing

After applying Bayes' rule

$$H^* = \operatorname{argmax}_H P(H \mid G) P(S \mid H, G)$$

Likelihood of scene

$$P(S \mid H, G) = \prod_{x \in H} P_g(x)^{T(x)} P_s^*(x)^{T(x)}$$

Scene parsing

After applying Bayes' rule

$$H^* = \operatorname{argmax}_H P(H \mid G) P(S \mid H, G)$$

Likelihood of scene

$$P(S \mid H, G) = \prod_{x \in H} P_g(x)^{T(x)} P_s^*(x)^{T(x)}$$

$P_g(x)$: geometry probability

Scene parsing

After applying Bayes' rule

$$H^* = \mathrm{argmax}_H \, P(H \mid G) P(S \mid H, G)$$

Likelihood of scene

$$P(S \mid H, G) = \prod_{x \in H} P_g(x)^{T(x)} P_s^*(x)^{T(x)}$$

$P_s^*(x)$: sum of all pairwise spatial probabilities $P_s(x)$

Scene parsing

We work in the negative logarithm space

$$E(H) = \log P(H \mid G) P(S \mid H, G)$$

$$= -\sum_{x \in H} T(x) \log\left(P_{prod}(x) P_g(x) P_s^*(x)\right)$$

Scene parsing

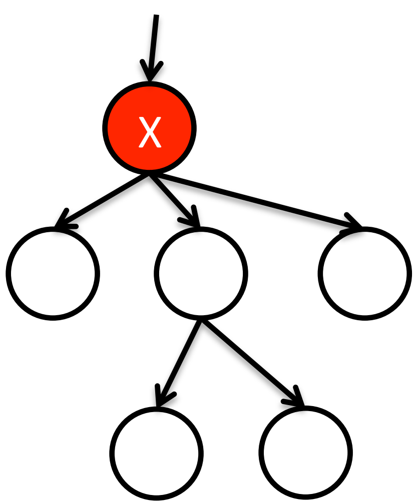
Rewrite the objective function recursively

$$E(H) = \overline{E}(R)$$

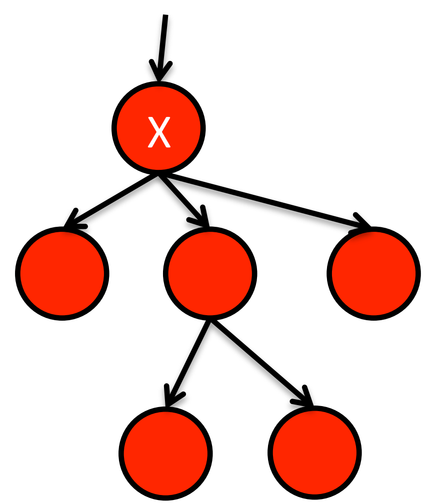
$$\overline{E}(x) = E(x) + \sum_{y \in x.children} \overline{E}(y)$$

where R is the root of H, $\overline{E}$ is the energy of a sub-tree.

$E(x)$



$\overline{E}(x)$



Scene parsing

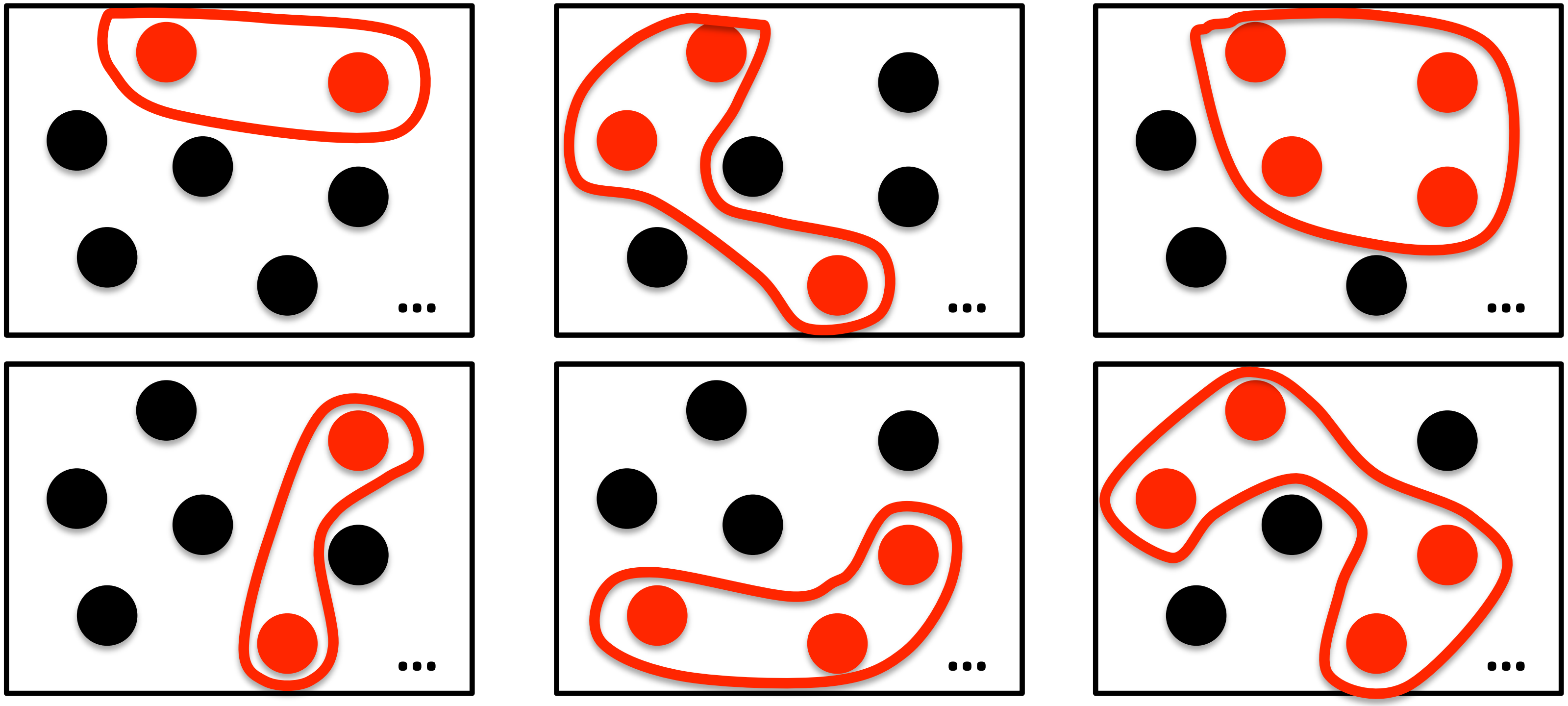
The search space is prohibitively large …

Problem 1: #possible groups is exponential.

Problem 2: #label assignments is exponential.

Scene parsing

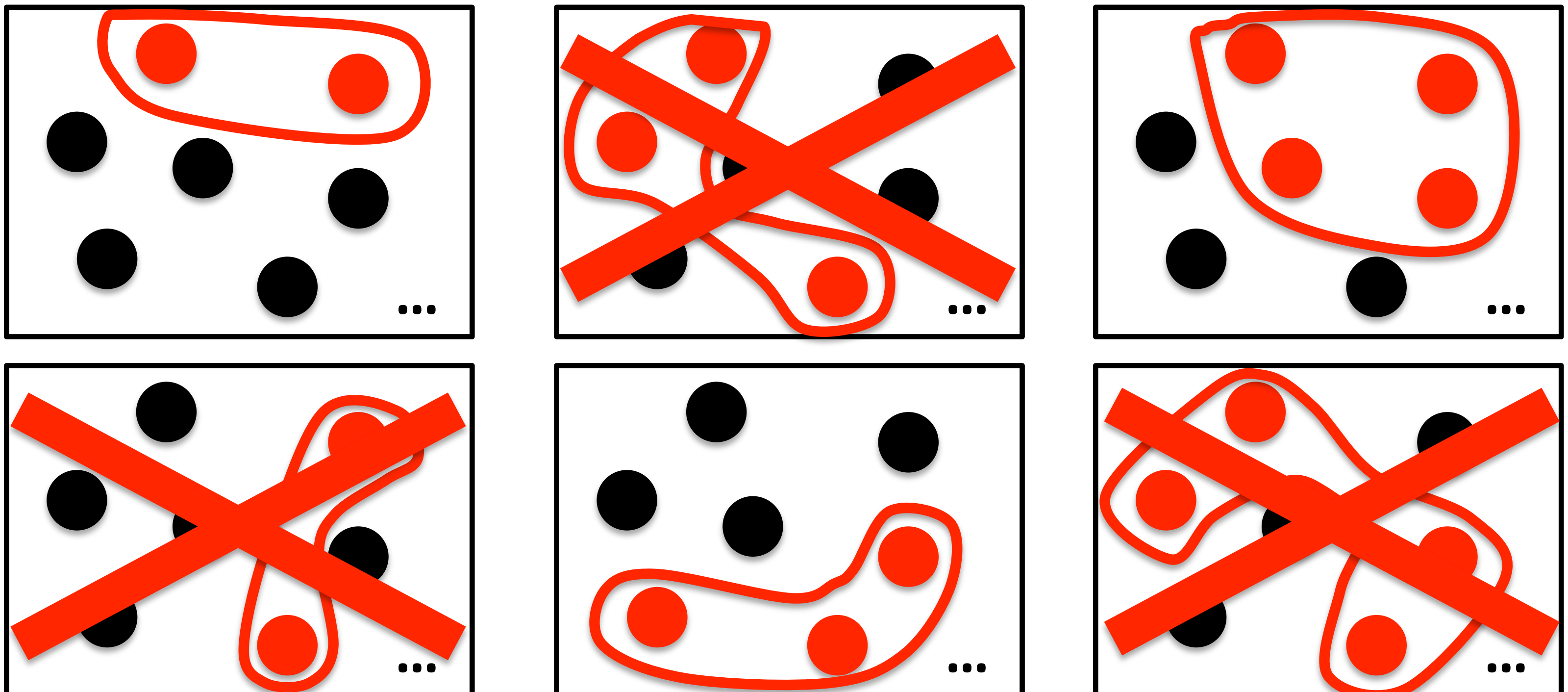
Problem 1: #possible groups is exponential.



Scene parsing

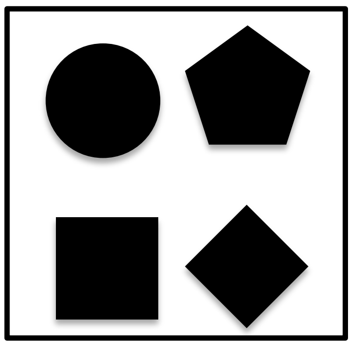
Problem 1: #possible groups is exponential.

Solution: proposing candidate groups.

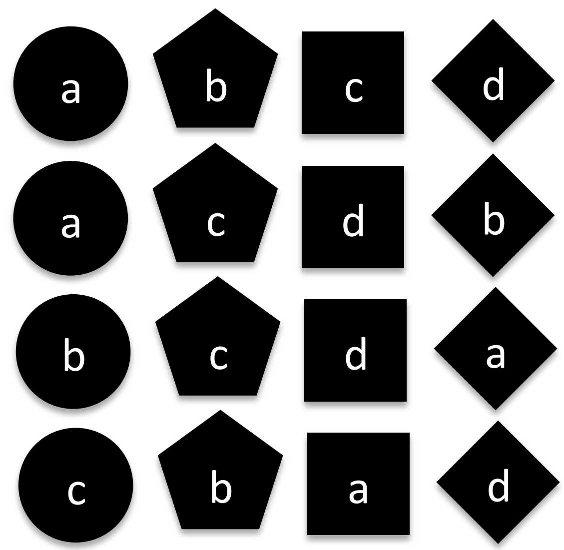


Scene parsing

Problem 2: #label assignments is exponential.



Rule: $r \rightarrow a, b, c, d$

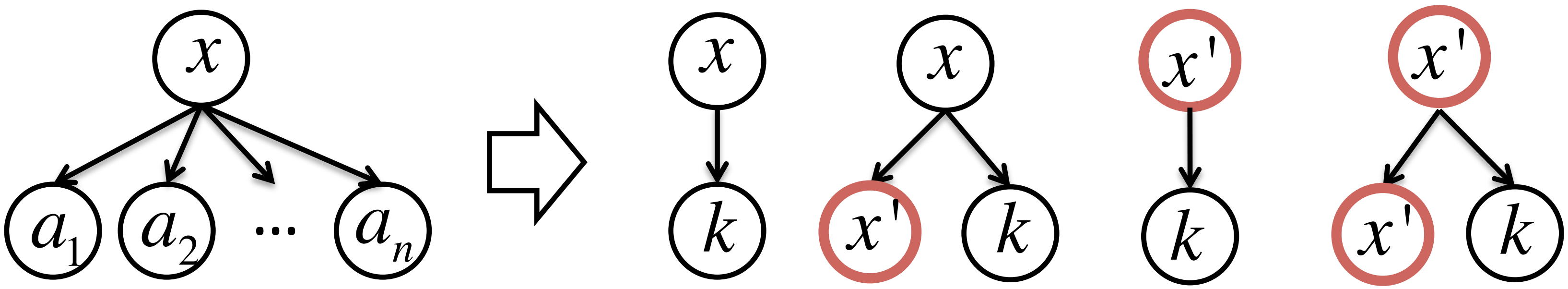


…

Scene parsing

Problem 2: #label assignments is exponential.

Solution: bounding #RHS by grammar binarization



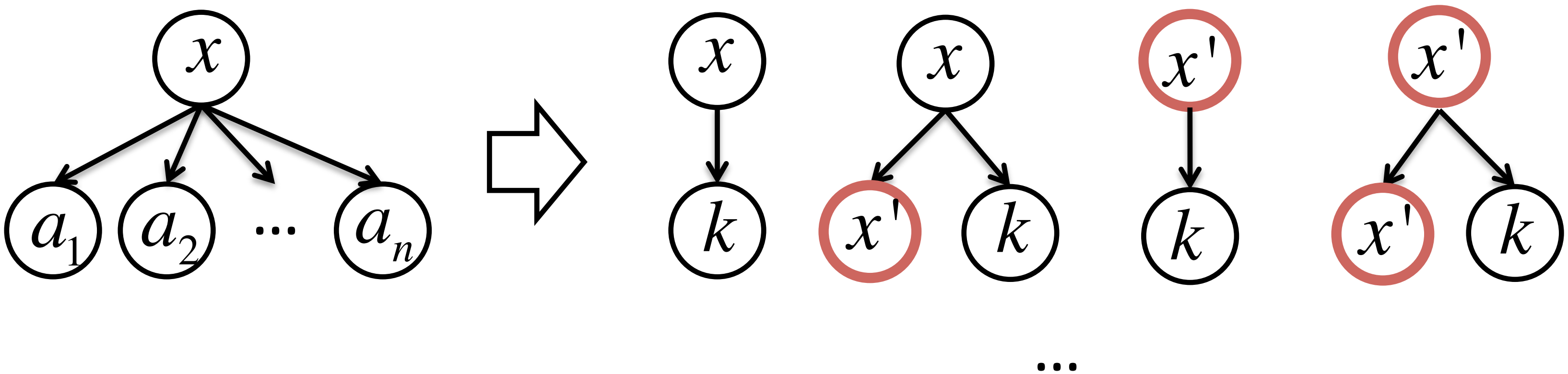
…

where x' is partial label of x, $k \in \{a_1, a_2, \ldots, a_n\}$

Scene parsing

Problem 2: #label assignments is exponential.

Solution: bounding #RHS by grammar binarization



where x' is partial label of x, $k \in \{a_1, a_2, \ldots, a_n\}$

Now #rules and #assignments are both polynomial.

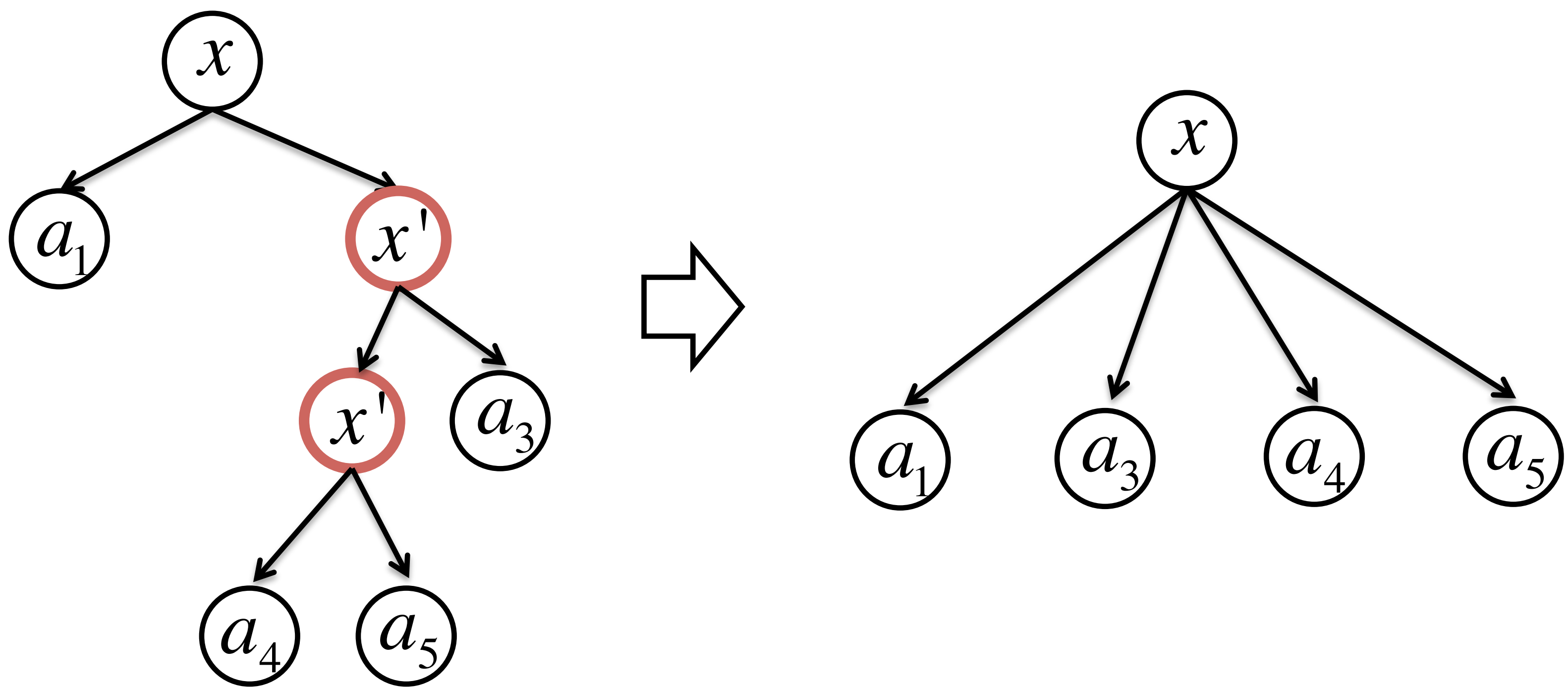
The problem can be solved by dynamic programming.

Scene parsing

Problem 2: #label assignments is exponential.

Solution: bounding #RHS by grammar binarization

Convert the result to a parse tree of the original grammar



Overview

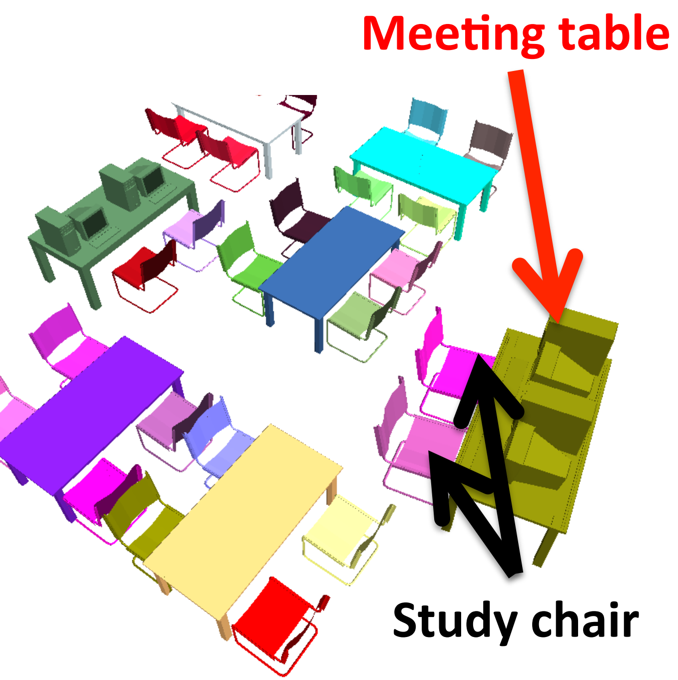
Grammar Structure

Learning a Probabilistic Grammar

Scene Parsing

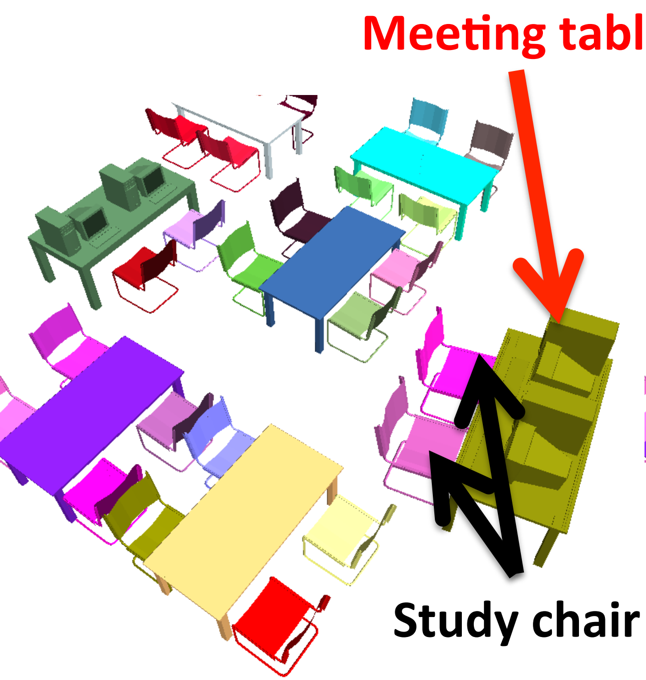
→ Results

Benefit of hierarchy

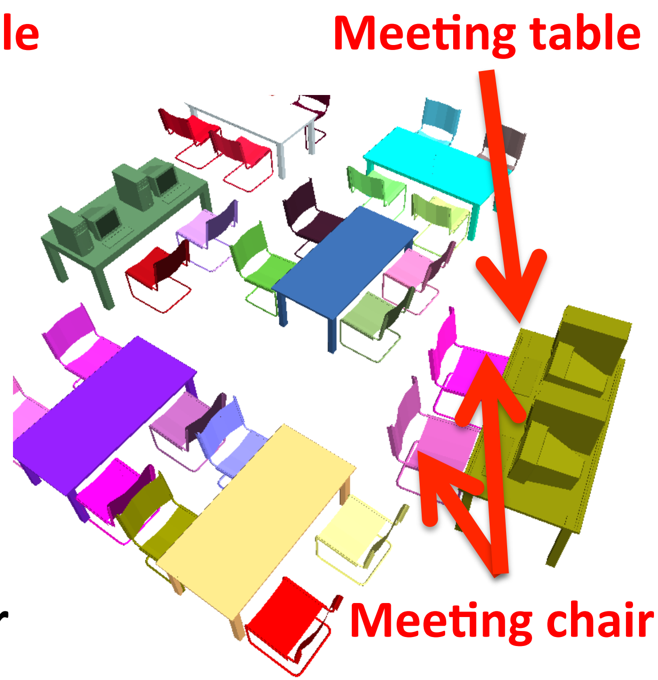


Shape only

Benefit of hierarchy

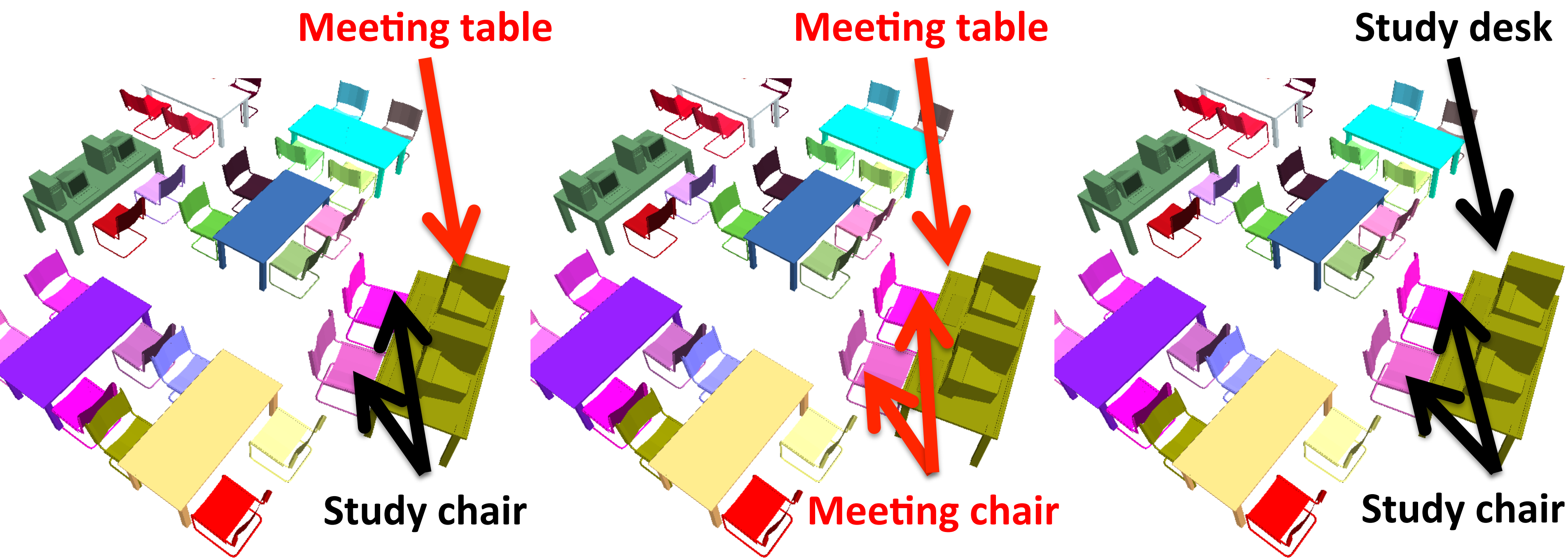


Shape only



Flat grammar

Benefit of hierarchy

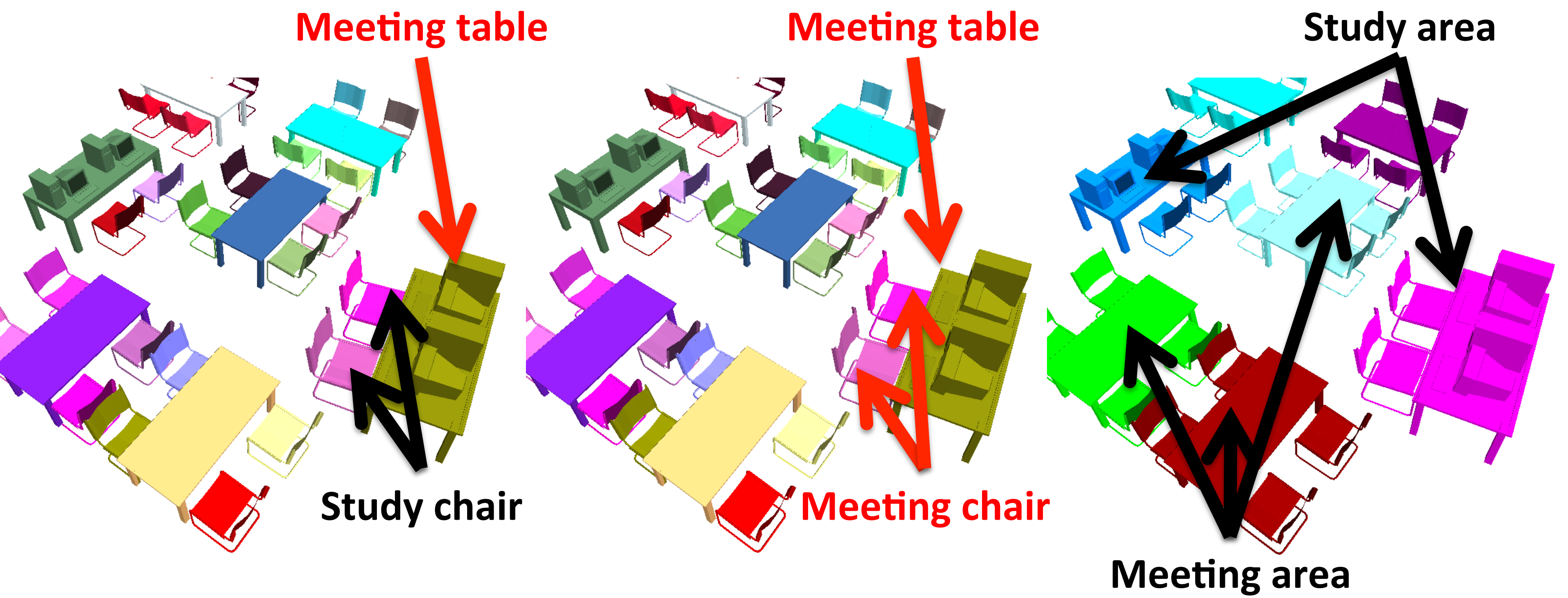


Shape only

Flat grammar

Ours

Benefit of hierarchy

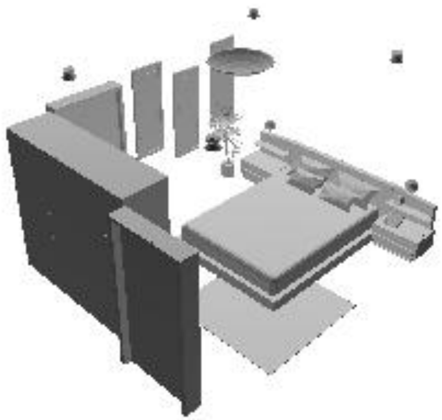


Shape only

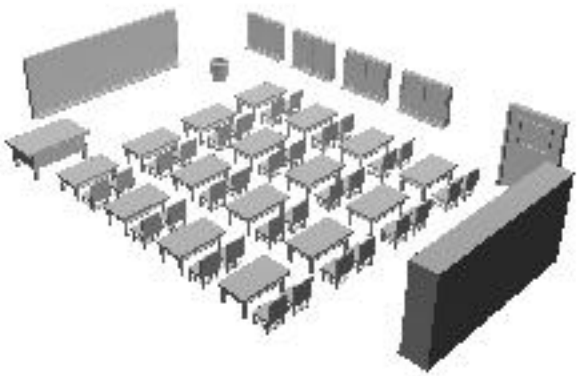
Flat grammar

Ours

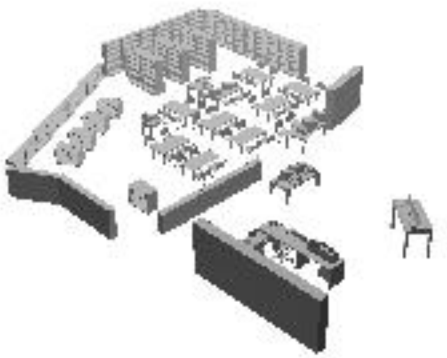
Datasets



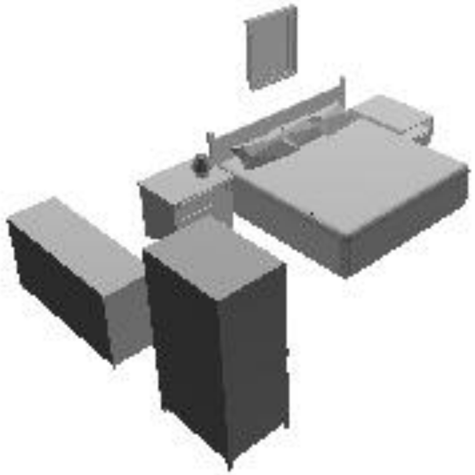
77 bedrooms



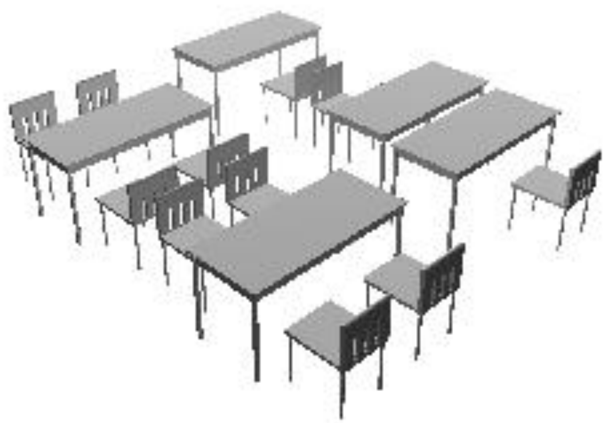
30 classrooms



8 libraries

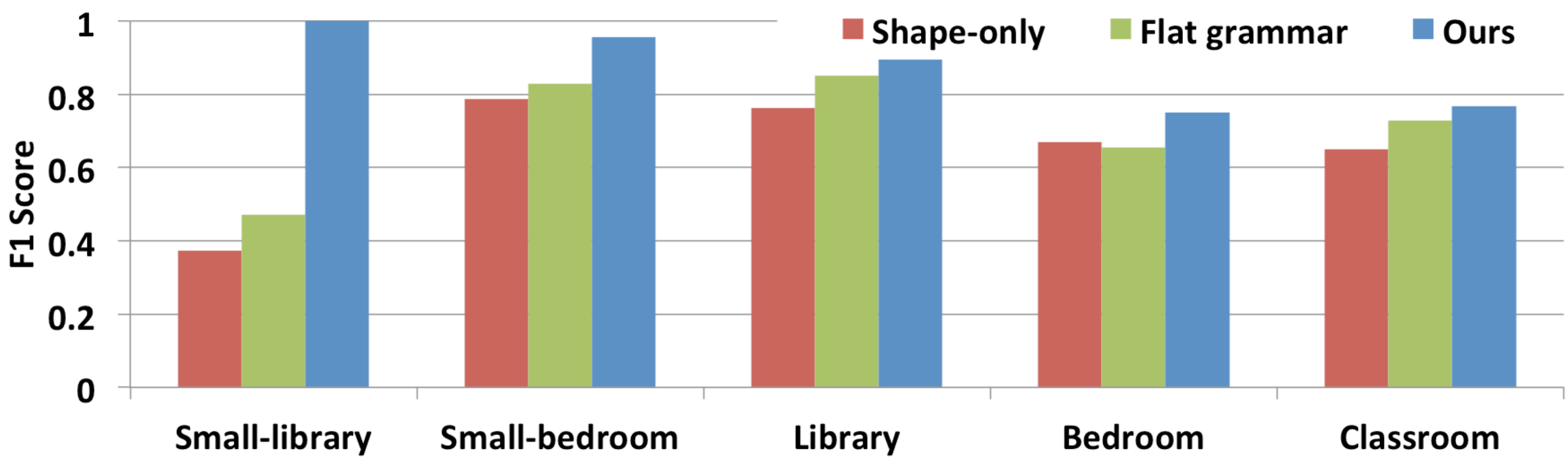


17 small bedrooms



8 small libraries

Benefit of hierarchy



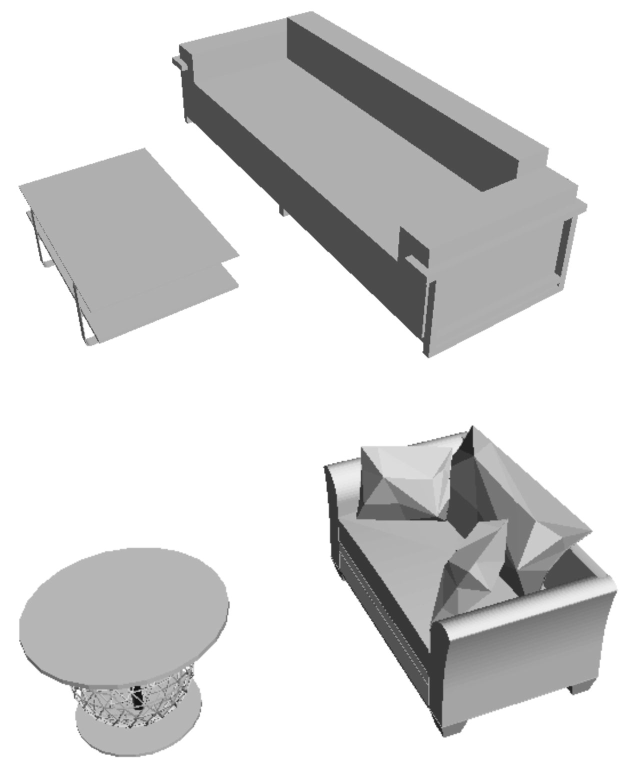
Object classification

Summary

- Modeling hierarchy improves scene understanding.

Limitations and future work

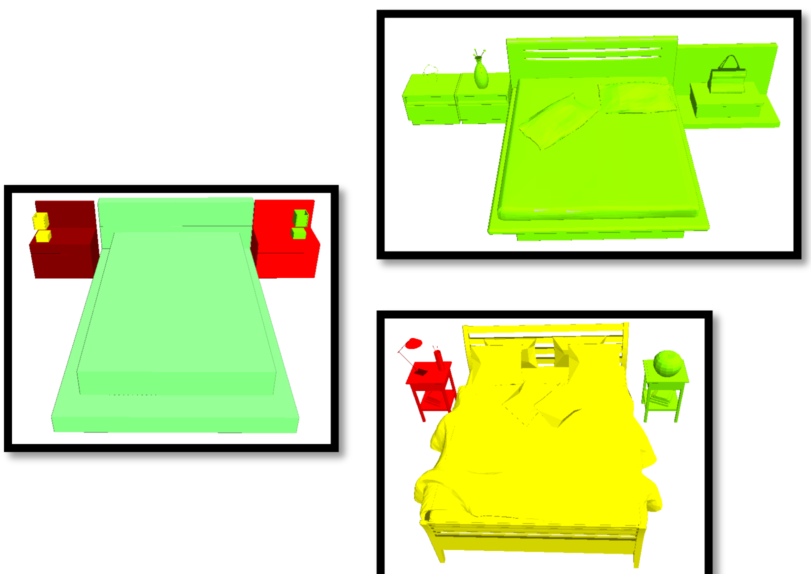
- Modeling correlation in probabilistic grammar
- Grammar learning from noisy data
- Applications in other fields



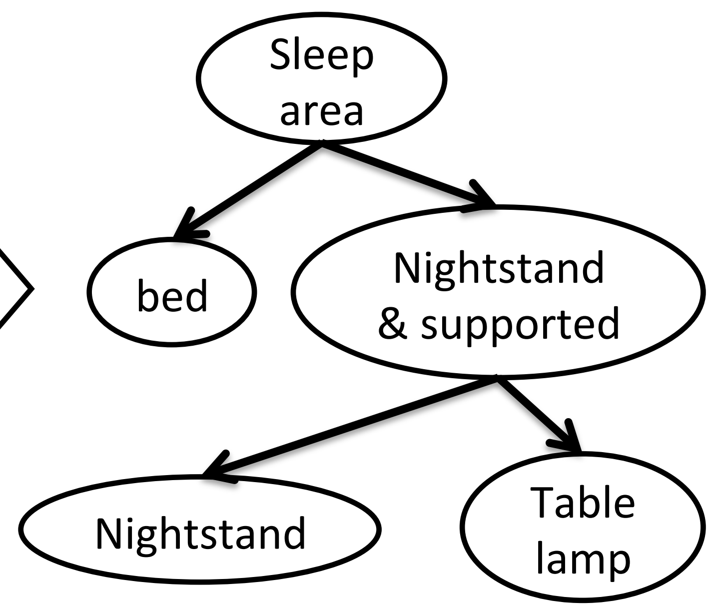
Limitations and future work

- Modeling correlation in probabilistic grammar
- Grammar learning from noisy data
- Applications in other fields

Input scene graphs

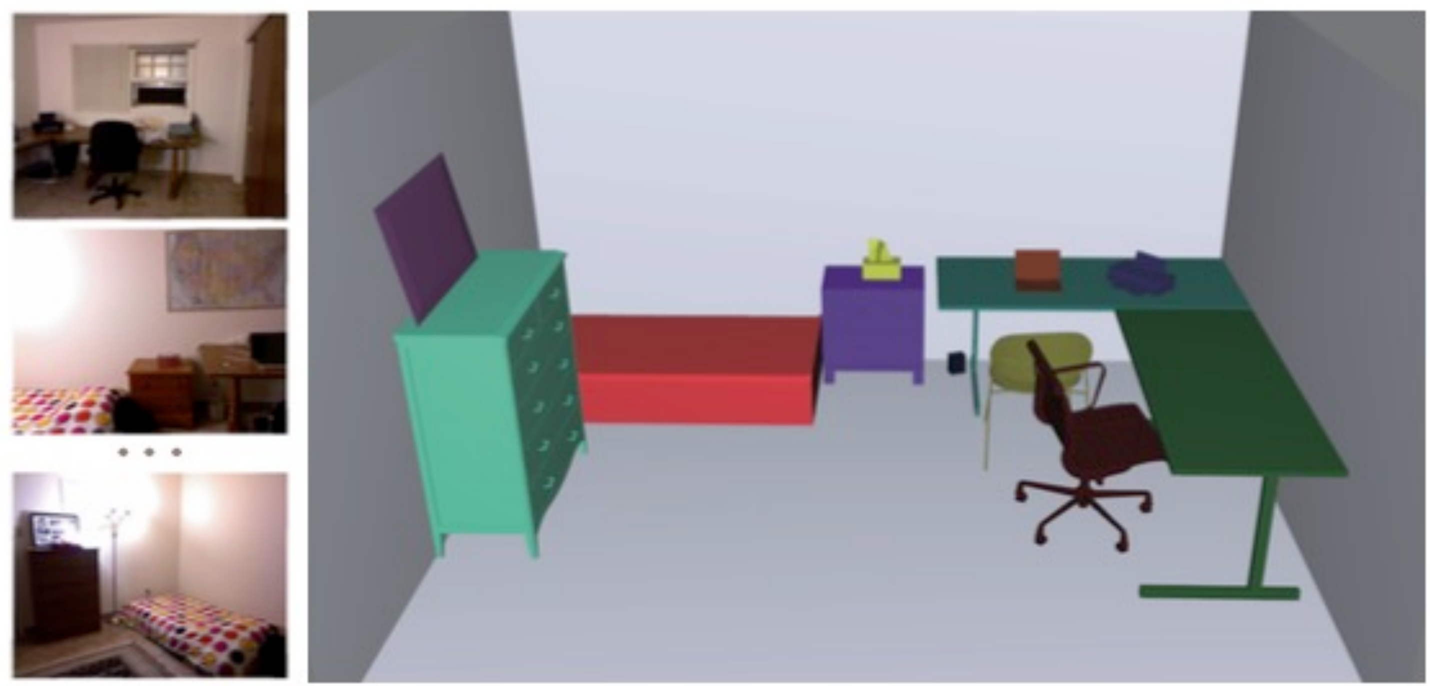


Grammar



Limitations and future work

- Modeling correlation in probabilistic grammar
- Grammar learning from noisy data
- Applications in other fields



Modeling from RGB-D data [Chen et al. 2014]

Outline

- Analyzing 3D scenes by modeling hierarchical structure
- Composition-aware scene optimization for product images
- Style compatibility for 3D furniture models

Motivation

Motivation

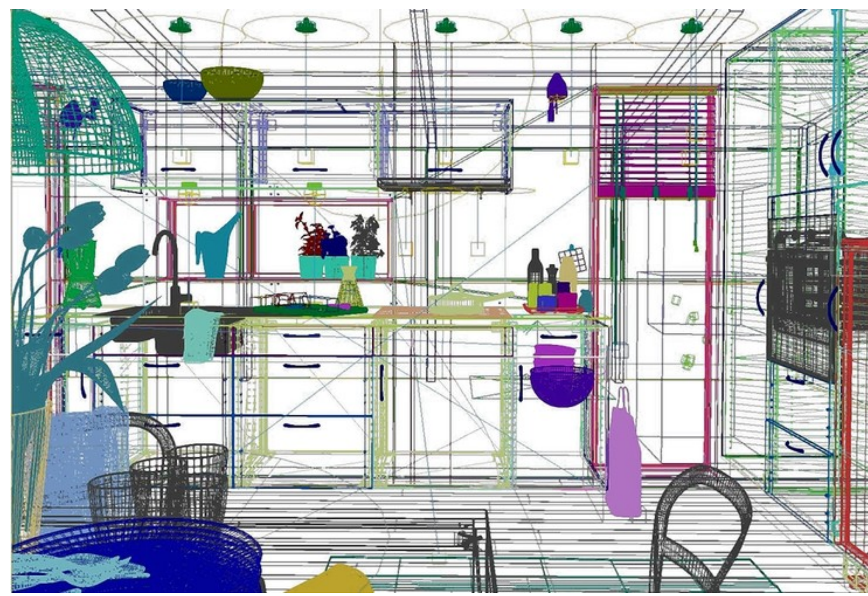
35% of scenes in IKEA catalogue are CGI.

Advantages

- Much less expensive
- Much easier for customization

Advantages

- Much less expensive
- Much easier for customization

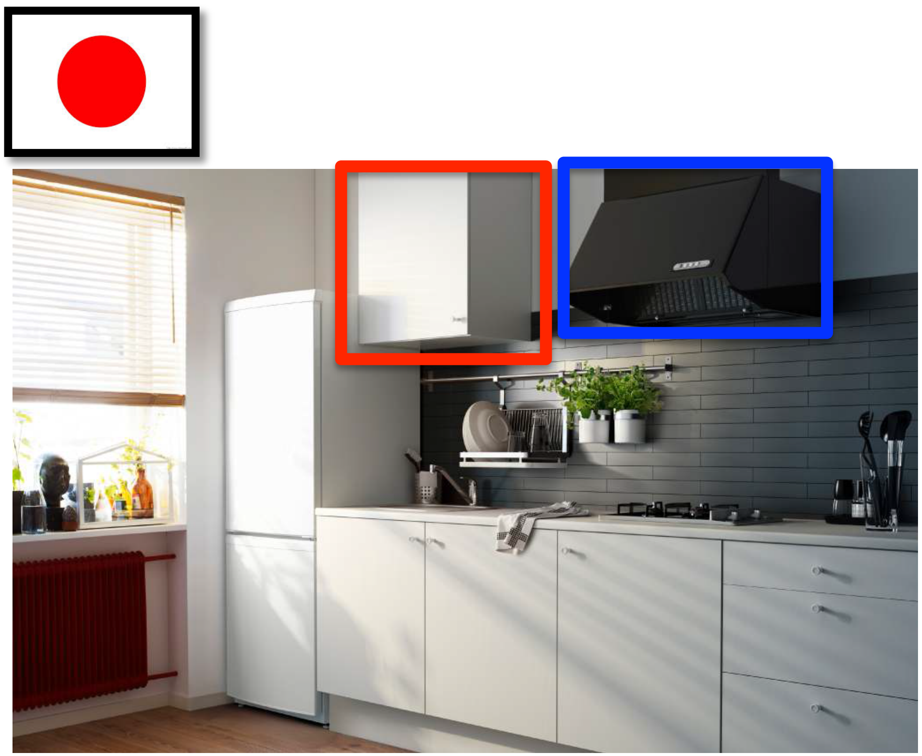


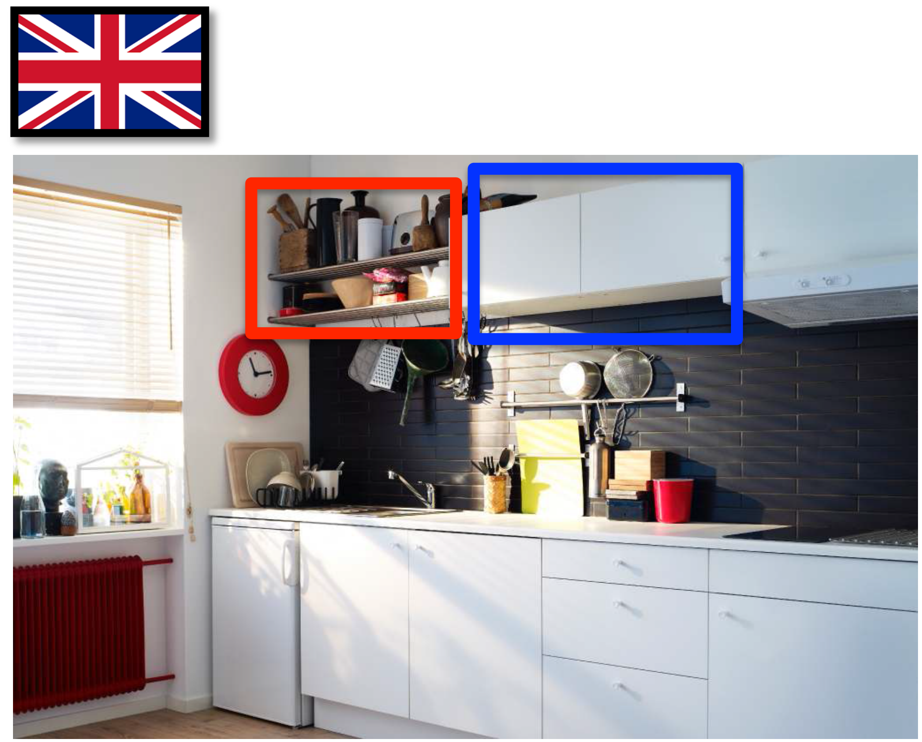
Advantages

- Much less expensive
- Much easier for customization

Advantages

- Much less expensive
- Much easier for customization





Artist’s goal

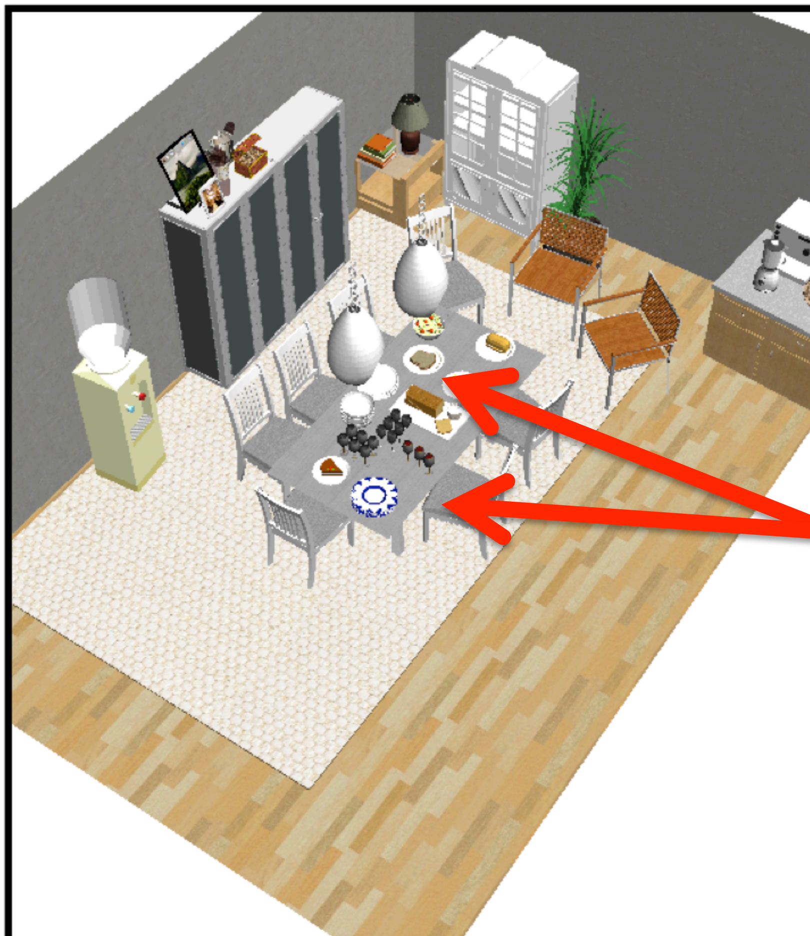
Artist's goal

Input: a rough scene, objects to highlight, and an initial camera view

Rough layout

Artist's goal

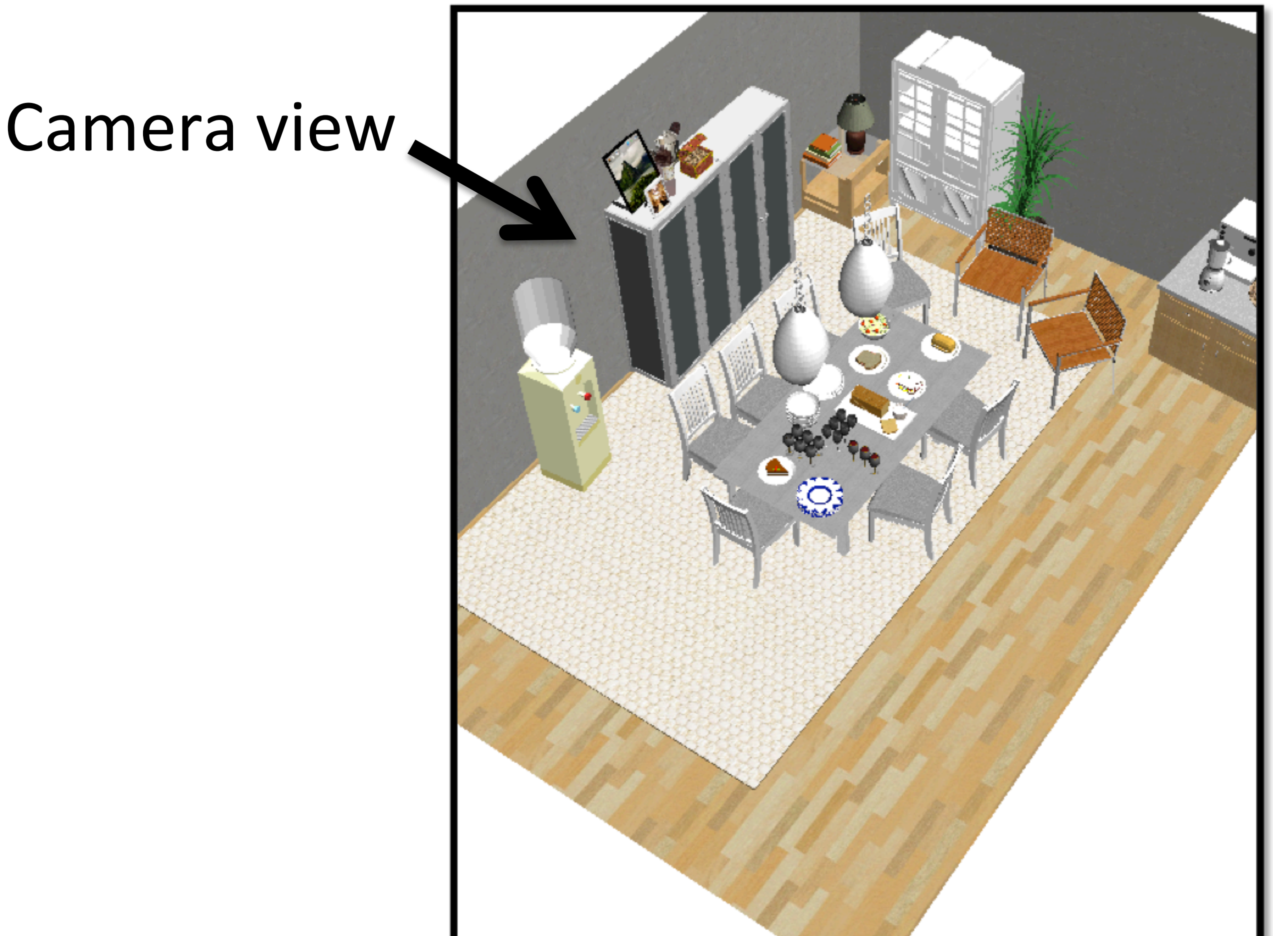
Input: a rough scene, objects to highlight, and an initial camera view



Highlight this chair and this table

Artist’s goal

Input: a rough scene, objects to highlight, and an initial camera view



Artist goal

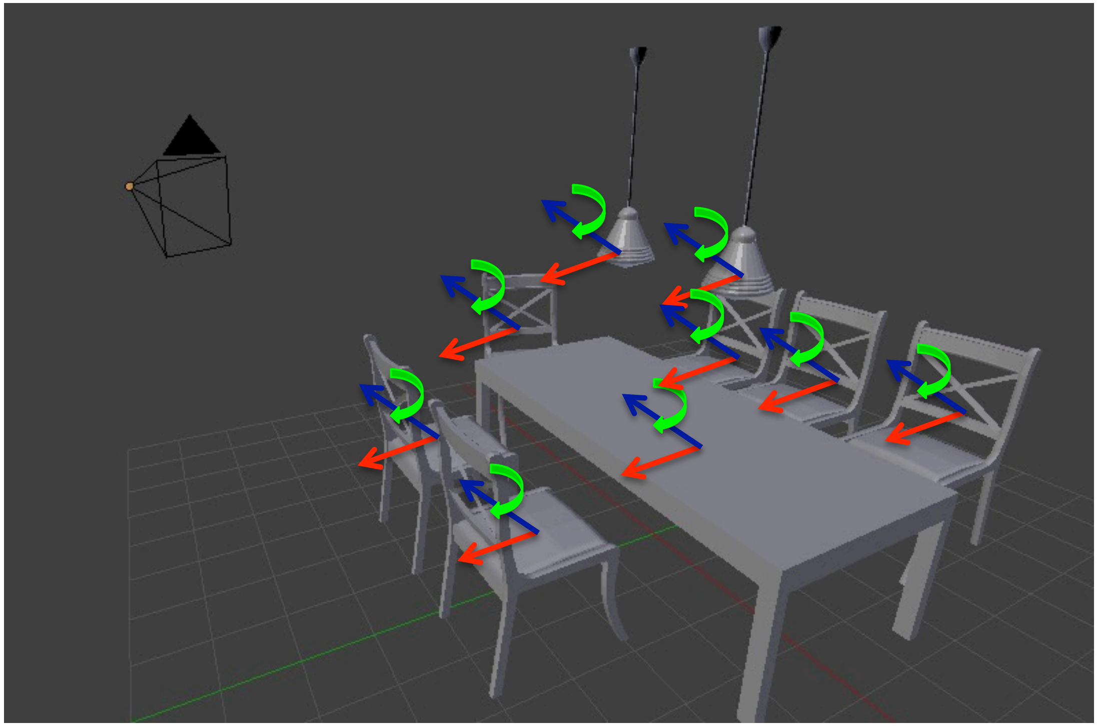
Output: a scene with optimized ***object placement***, ***materials*** and ***camera view*** that produce an appealing 2D composition.

Challenges

- Huge search space to explore
- Many principles/constraints to balance
- Requiring repeating work for customization

Challenges

- Huge search space to explore
- Many principles/constraints to balance
- Requiring repeating work for customization



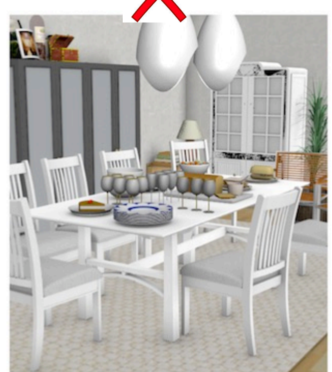
4*N + 6 parameters

- 3 DOF per object
- 1 material per object
- 6 DOF for camera

Challenges

- Huge search space to explore
- Many principles/constraints to balance
- Requiring repeating work for customization

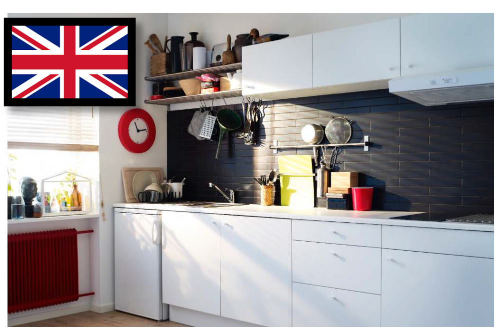
...

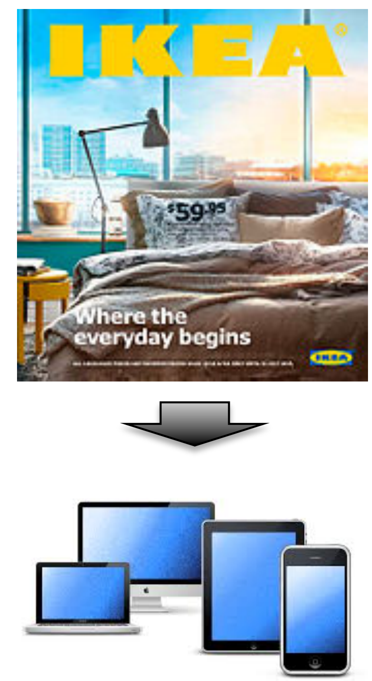


...

Challenges

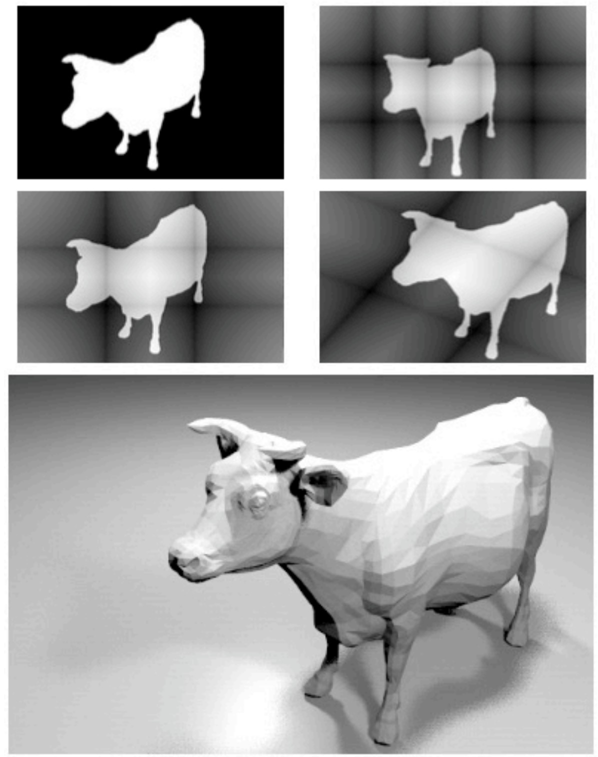
- Huge search space to explore
- Many principles/constraints to balance
- Requiring repeating work for customization



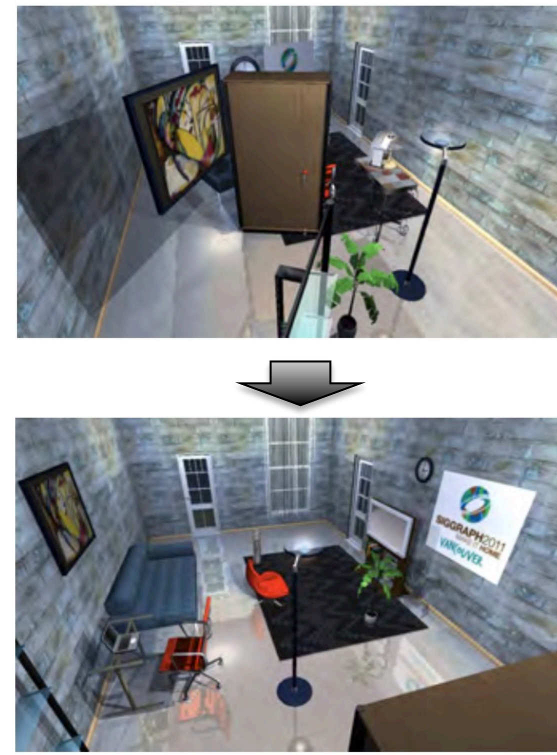


Related work

Image optimization
[Liu et al. 2010]



Camera optimization
[Gooch et al. 2001]



Scene optimization
[Yu et al. 2011]

Key idea

$$E(\{\boxed{x_i, y_i}, \theta_i\}, \{m_i\}, C) = E_{op} + E_{os} + E_{ic} + E_{cp} + E_{3d} + E_r$$

x_i, y_i : position of object i on its supporting surface

Key idea

$$E(\{x_i, y_i, \boxed{\theta_i}\}, \{m_i\}, C) = E_{op} + E_{os} + E_{ic} + E_{cp} + E_{3d} + E_r$$

x_i, y_i : position of object i on its supporting surface

θ_i : orientation of object i

Key idea

$$E(\{x_i, y_i, \theta_i\}, \boxed{\{m_i\}}, C) = E_{op} + E_{os} + E_{ic} + E_{cp} + E_{3d} + E_r$$

x_i, y_i : position of object i on its supporting surface

θ_i : orientation of object i

m_i : material of object i

Key idea

$$E(\{x_i, y_i, \theta_i\}, \{m_i\}, \boxed{C}) = E_{op} + E_{os} + E_{ic} + E_{cp} + E_{3d} + E_r$$

x_i, y_i : position of object i on its supporting surface

θ_i : orientation of object i

m_i : material of object i

C : camera parameters

Key idea

$$E(\{x_i, y_i, \theta_i\}, \{m_i\}, C) = \boxed{E_{op} + E_{os} + E_{ic} + E_{cp} + E_{3d} + E_r}$$

x_i, y_i : position of object i on its supporting surface

θ_i : orientation of object i

m_i : material of object i

C : camera parameters

$E_{op}, E_{os}, E_{ic}, E_{cp}, E_{3d}, E_r$: terms for composition rules

Key idea

$$E(\{x_i, y_i, \theta_i\}, \{m_i\}, C) = E_{op} + E_{os} + E_{ic} + E_{cp} + E_{3d} + E_r$$

Never been considered before

x_i, y_i : position of object i on its supporting surface

θ_i : orientation of object i

m_i : material of object i

C : camera parameters

$E_{op}, E_{os}, E_{ic}, E_{cp}, E_{3d}, E_r$: terms for composition rules

Overview

→ Composition rules and constraints

Optimization

Applications

Composition rules

1. Object placement within 2D frame E_{op}
2. Object saliency within 2D frame E_{os}
3. Image composition E_{ic}
4. Camera placement E_{cp}
5. Object constraints within 3D scene E_{3d}
6. Regularization E_r

Term 1: Object placement within 2D frame

- Rule of thirds
- Centeredness
- Clearance

Term 1: Object placement within 2D frame

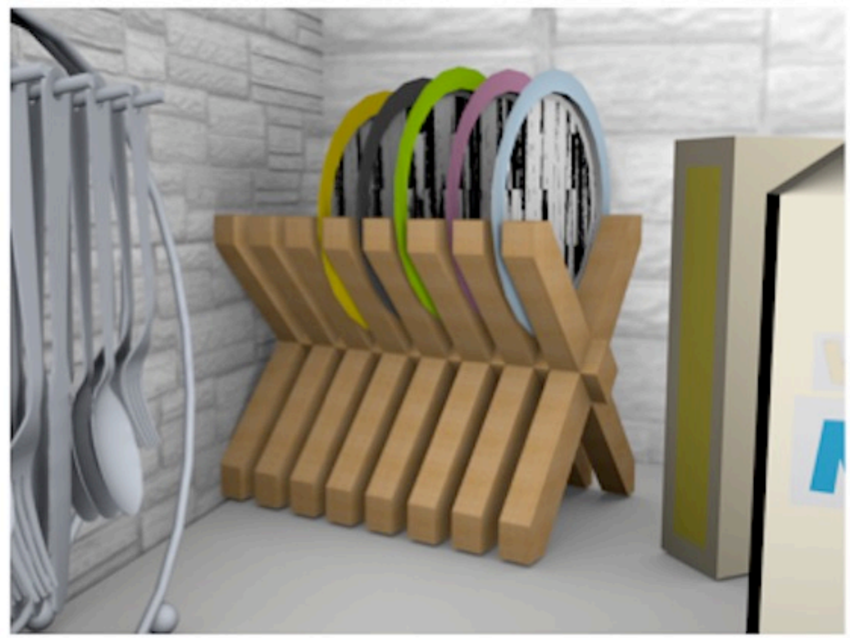
- Rule of thirds
- Centeredness
- Clearance

Term 1: Object placement within 2D frame

- Rule of thirds
- Centeredness
- Clearance

Term 2: Object saliency within 2D frame

- Visibility
- Object size



Term 2: Object saliency within 2D frame

- Visibility
- Object size

Term 3: Image composition

- Visual balance
- Color contrast

Term 3: Image composition

- Visual balance
- Color contrast

Term 4: Camera placement

- Canonical views
- Typical views

Term 4: Camera placement

- Canonical views
- Typical views

Term 5: Object constraints within 3D scene

- Collision relationships
- Support relationships
- Semantic constraints

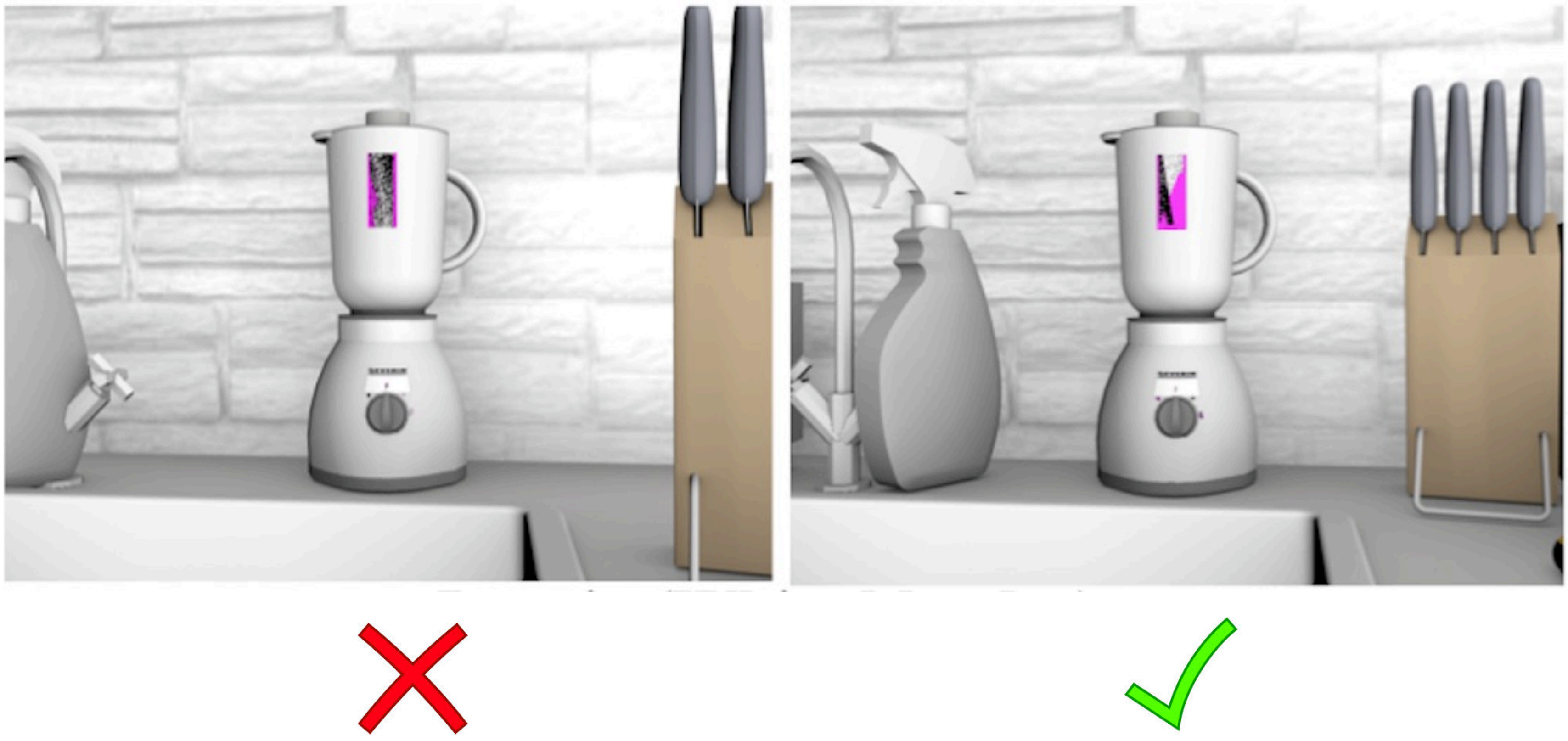
Term 5: Object constraints within 3D scene

- Collision relationships
- Support relationships
- Semantic constraints

Term 5: Object constraints within 3D scene

- Collision relationships
- Support relationships
- Semantic constraints

Term 6: Regularization



Overview

Composition rules and constraints

→ Optimization

Applications

Energy function

$$E(\{x_i, y_i, \theta_i\}, \{m_i\}, C) = E_{op} + E_{os} + E_{ic} + E_{cp} + E_{3d} + E_r$$

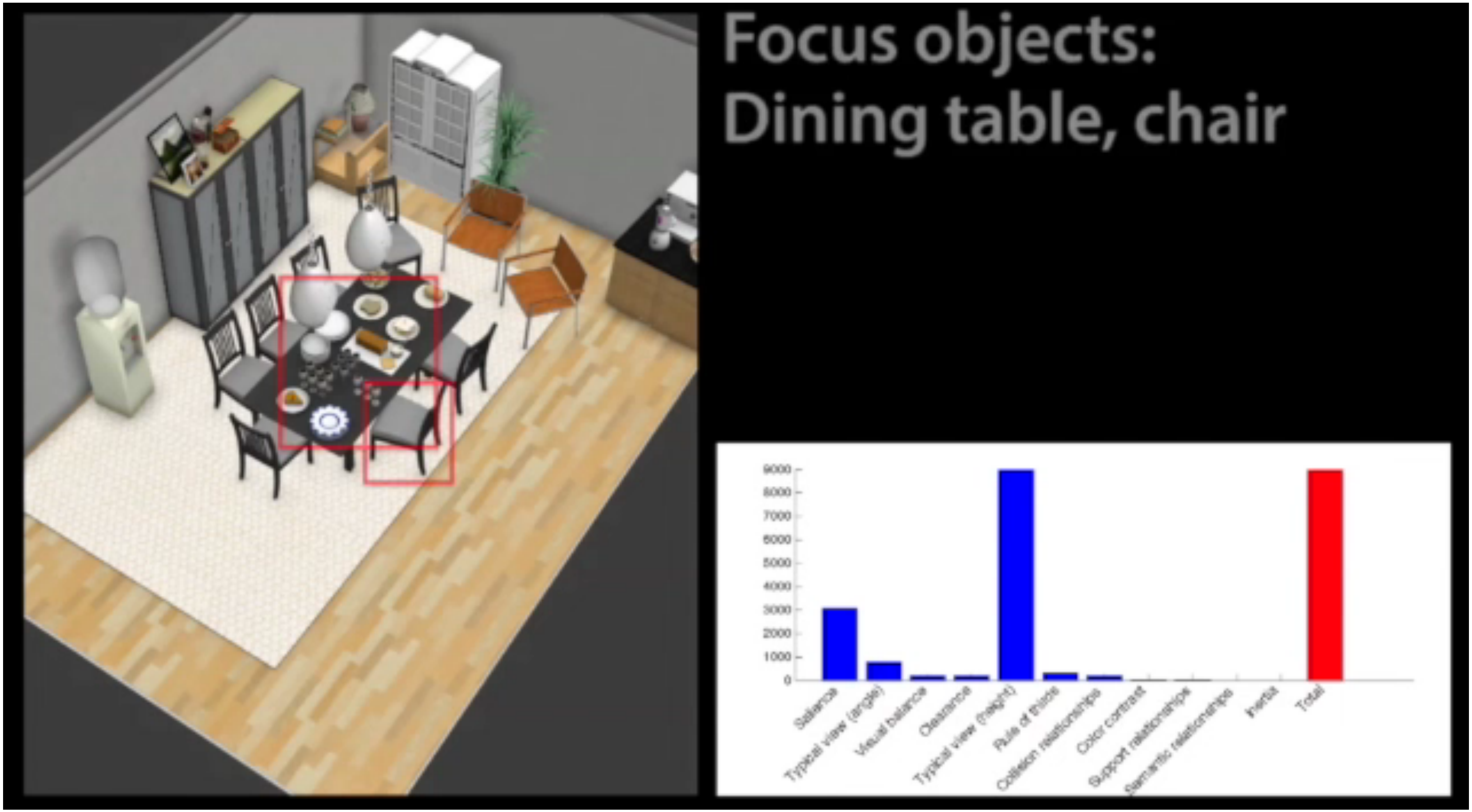
Continuous variables

Discrete variables

Optimization

- Continuous optimization – camera view and object placement
- Discrete optimization – materials

Example


Focus objects:
Dining table, chair
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
Saliency
Typical view (angle)
Visual balance
Clearance
Typical view (height)
Rule of thirds
Collision relationships
Color contrast
Support relationships
Semantic relationships
Inertia
Total

Overview

Composition rules and constraints

Optimization

→ Applications

Applications

1. Refining rough compositions
2. Retargeting for different aspect ratios
3. Retargeting for different cultural preferences
4. Text-incorporated composition
5. Generating detail images from an overview

Application 1: Refining rough compositions

Rough composition

Optimized composition

Application 1: Refining rough compositions

User study

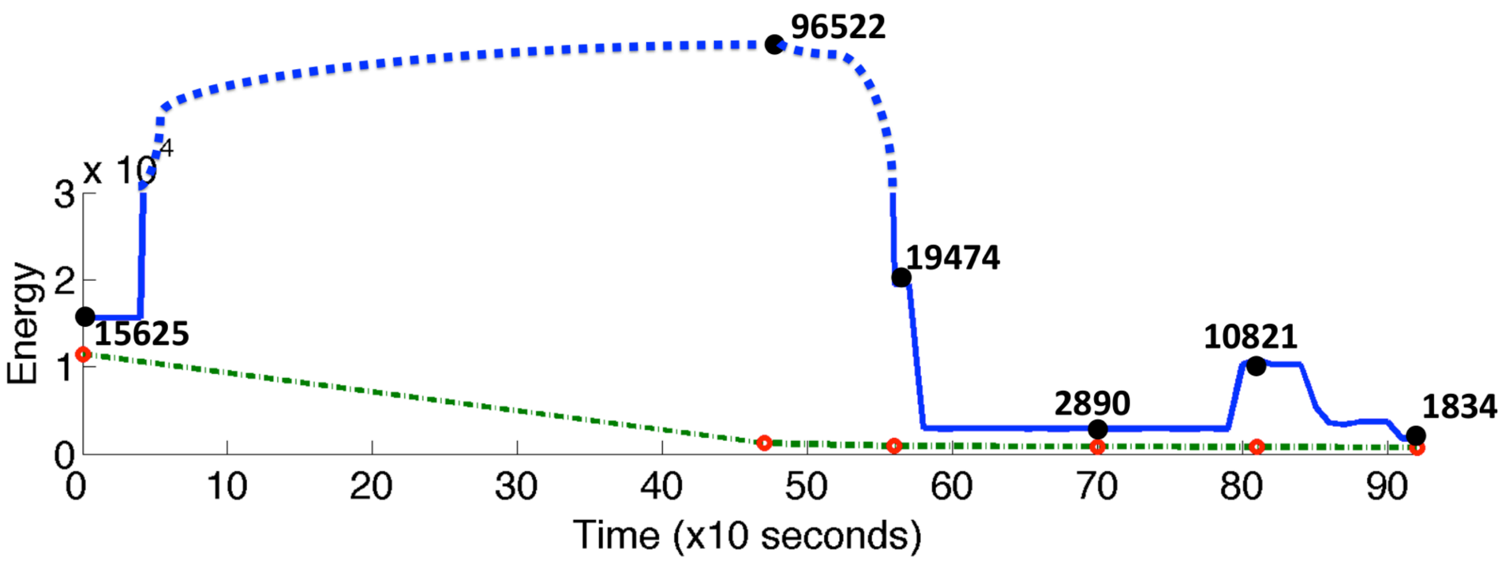
Application 1: Refining rough compositions

User study

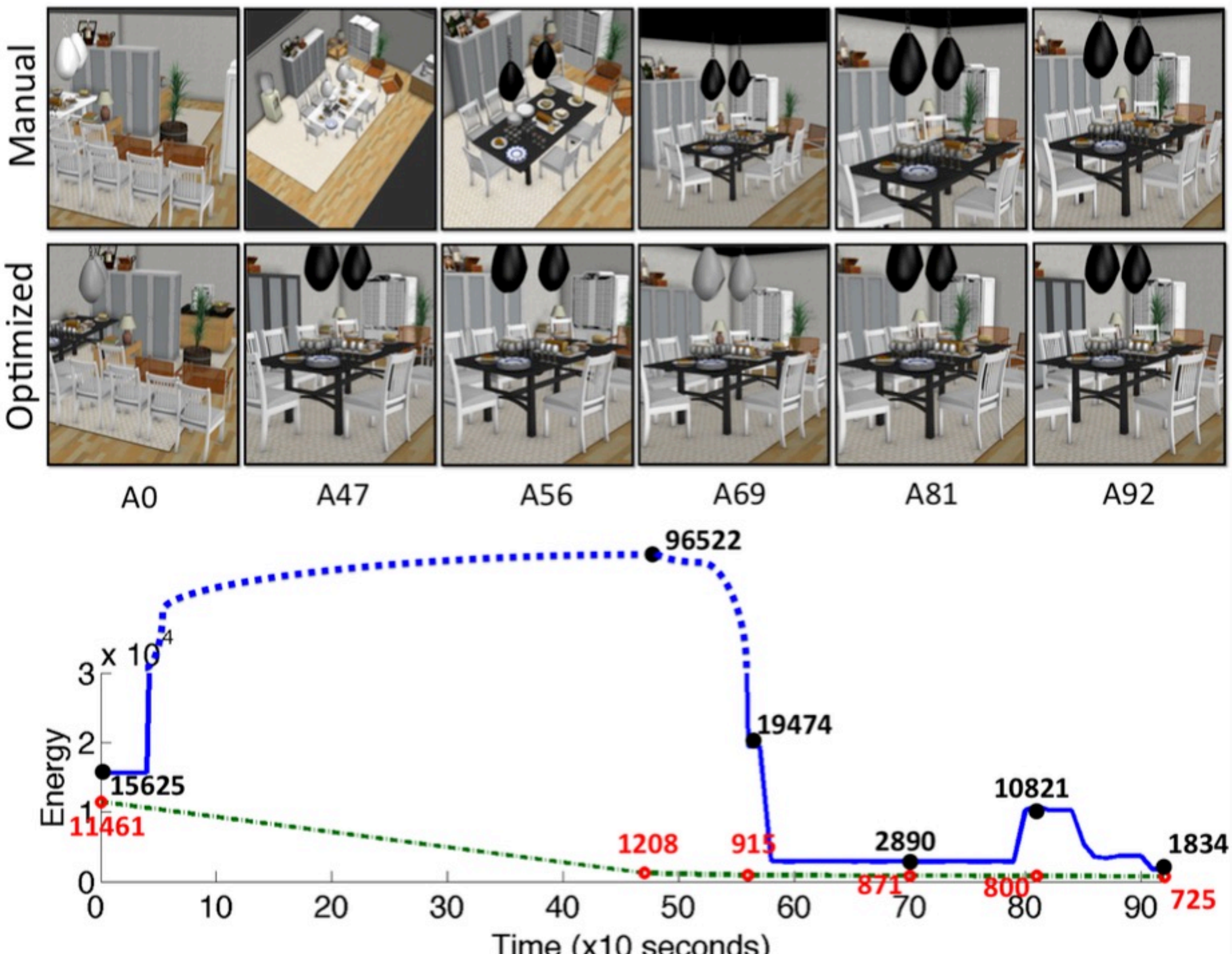
Reference

Application 1: Refining rough compositions

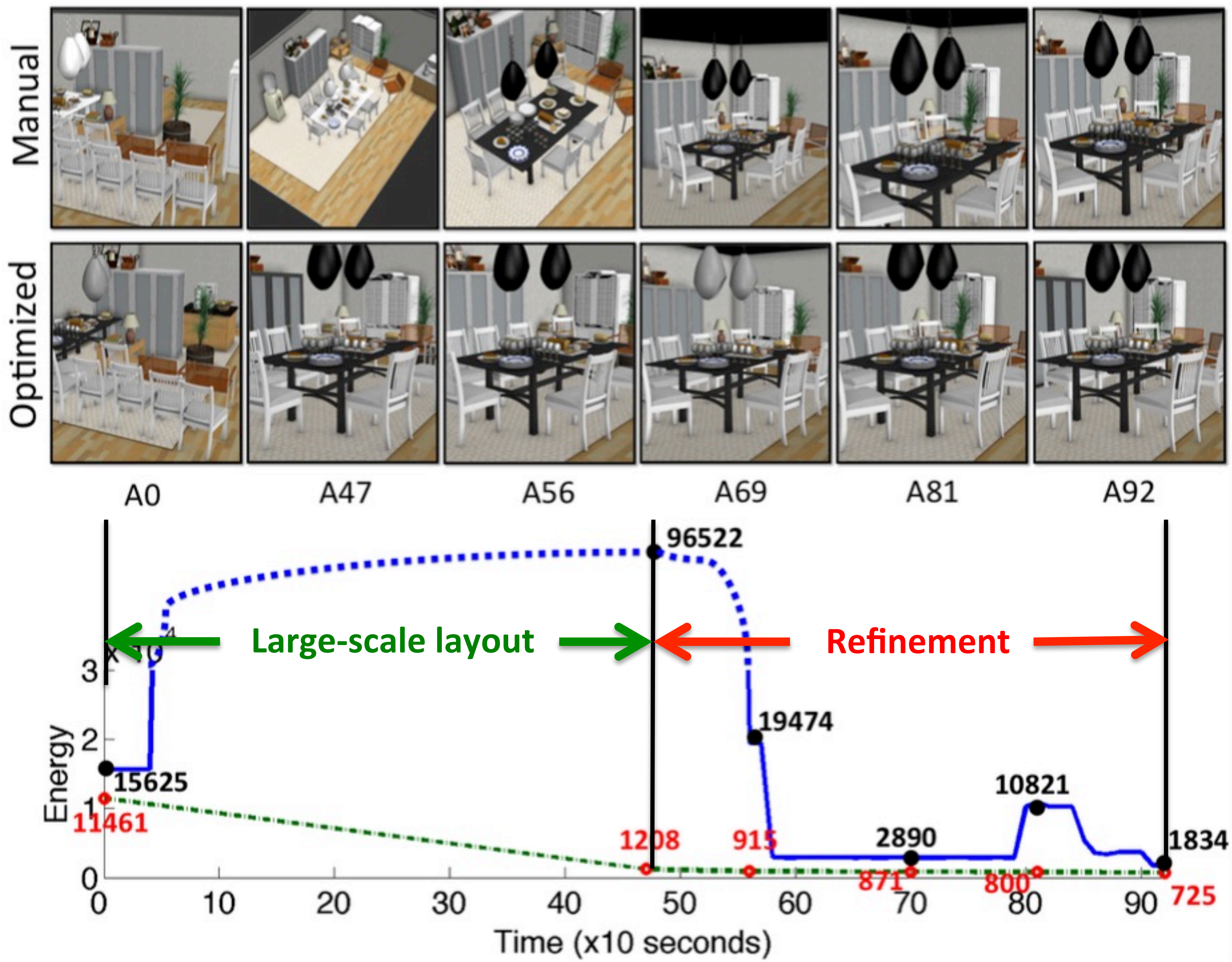
Application 1: Refining rough compositions



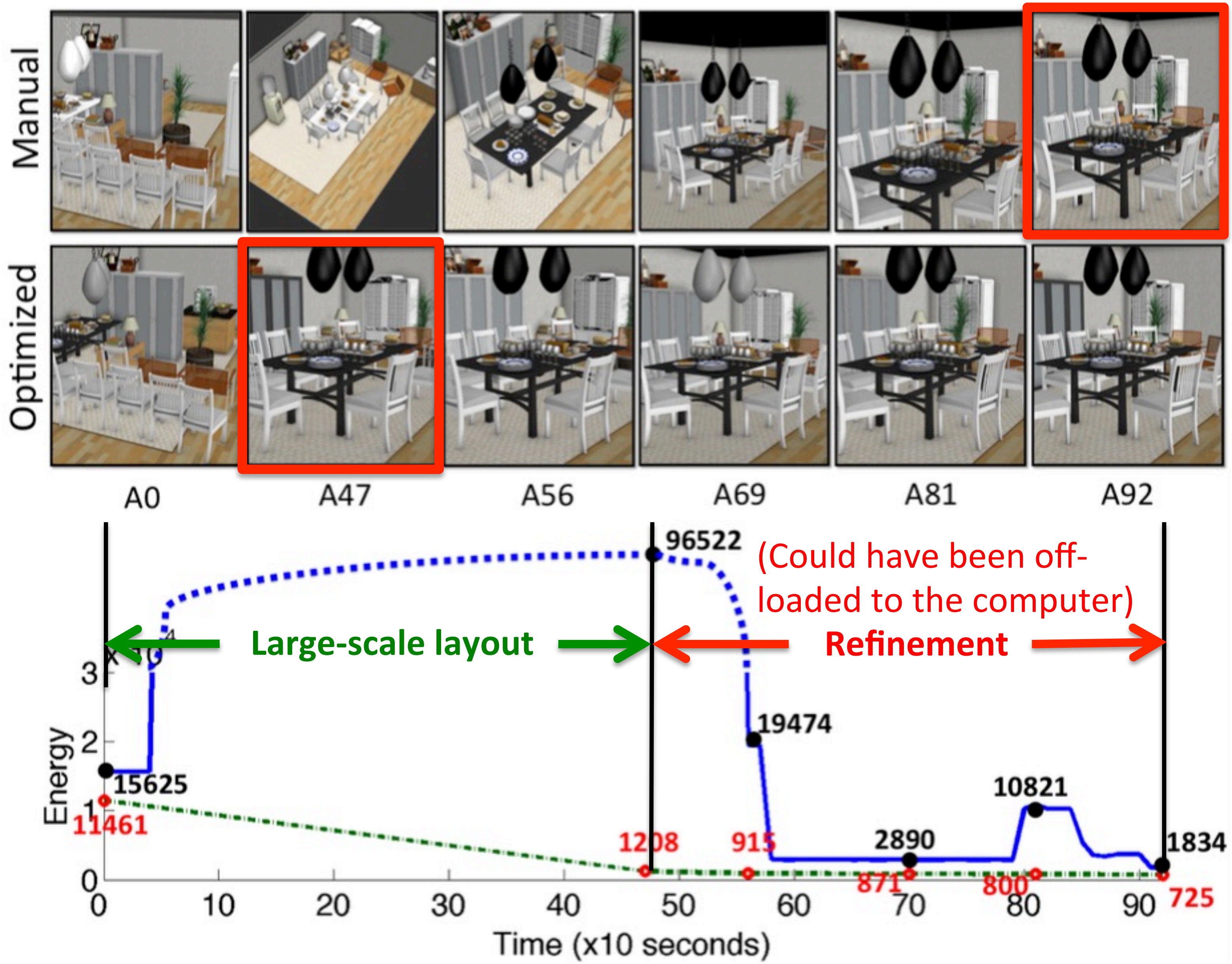
Application 1: Refining rough compositions



Application 1: Refining rough compositions



Application 1: Refining rough compositions



App 2: Retargeting for different aspect ratios

?

App 2: Retargeting for different aspect ratios

Input (4:3)

Camera-only

App 2: Retargeting for different aspect ratios

Input (4:3) Camera-only Ours (1:2)

App 3: Retargeting for different cultural preferences

App 3: Retargeting for different cultural preferences

(a) *Original* **(b)** *Objects replaced*

App 3: Retargeting for different cultural preferences

(a) *Original* **(b)** *Objects replaced* **(c)** *Optimized*

App 4: Text-incorporated composition

App 4: Text-incorporated composition

Input

App 4: Text-incorporated composition

Input

Extra terms for overlaid text

- Contrast term

App 4: Text-incorporated composition

Input

Extra terms for overlaid text

- Contrast term
- Variance term

App 4: Text-incorporated composition

Input

Camera only

App 4: Text-incorporated composition

Input

Camera only

Our result

App 5: Generating detail images from an overview

App 5: Generating detail images from an overview

App 5: Generating detail images from an overview

App 5: Generating detail images from an overview

(a) *Overview*

App 5: Generating detail images from an overview

(a) *Overview*

(b) *Speaker*

App 5: Generating detail images from an overview

(a) *Overview*

(b) *Speaker*

(c) *Shelf*

A perceptual study

Comparing the results of our method and optimizing camera only.

Kitchen

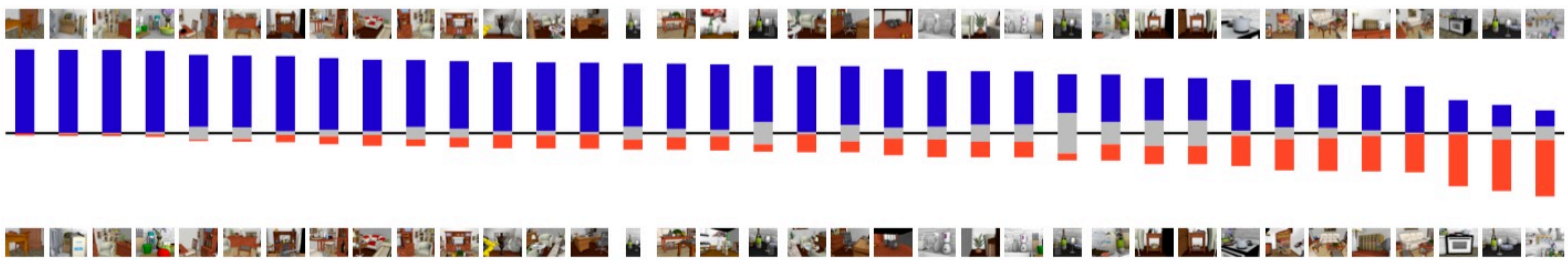
Study

Living room

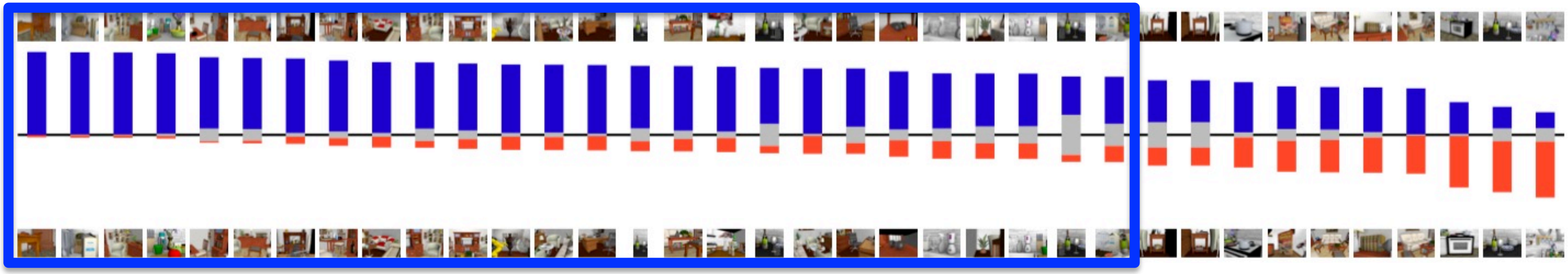
Expert study results

ID	Ours	Camera Only	No preference
Expert 1	22	12	2
Expert 2	17	14	3
Expert 3	22	11	3
Expert 4	21	12	3

Amazon Mechanical Turk study



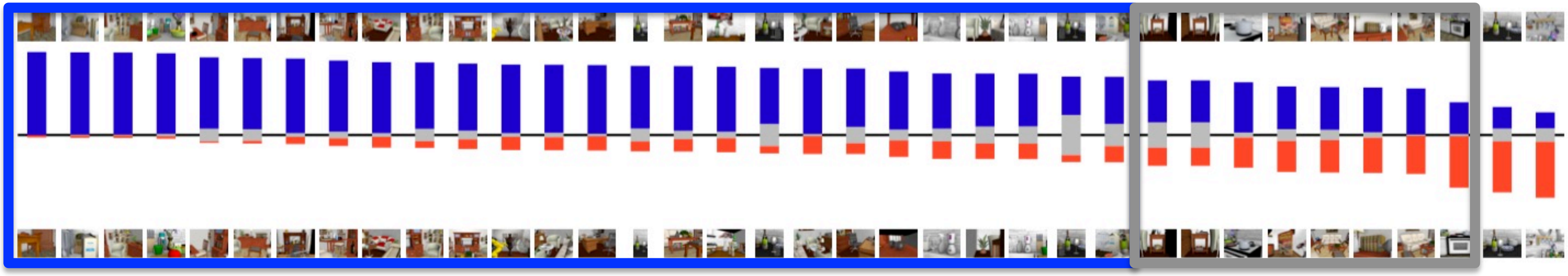
Amazon Mechanical Turk study



If null hypothesis is there is no preference,

- Our method is preferred in 26/36 cases.

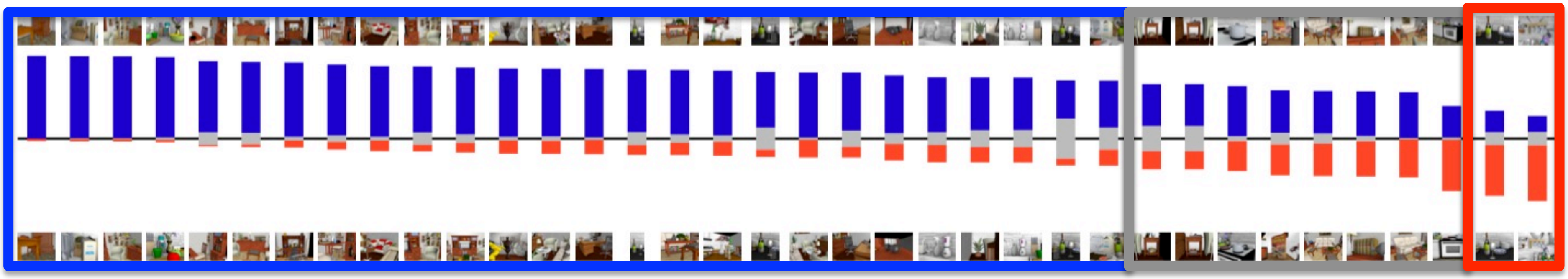
Amazon Mechanical Turk study



If null hypothesis is there is no preference,

- Our method is preferred in 26/36 cases.
- No statistical significance in 8 cases.

Amazon Mechanical Turk study



If null hypothesis is there is no preference,

- Our method is preferred in 26/36 cases.
- No statistical significance in 8 cases.
- Camera only is preferred in 2 cases.

Summary

- Reasoning about relationships between objects in the image space and the scene space helps create good compositions.
- Moving objects and changing materials significantly improves the quality of compositions.
- Our optimization framework benefits a variety of applications.

Limitations and future work

- Interactive scene optimization
- Global illumination
- Additional composition rules

Limitations and future work

- Interactive scene optimization
- Global illumination
- Additional composition rules

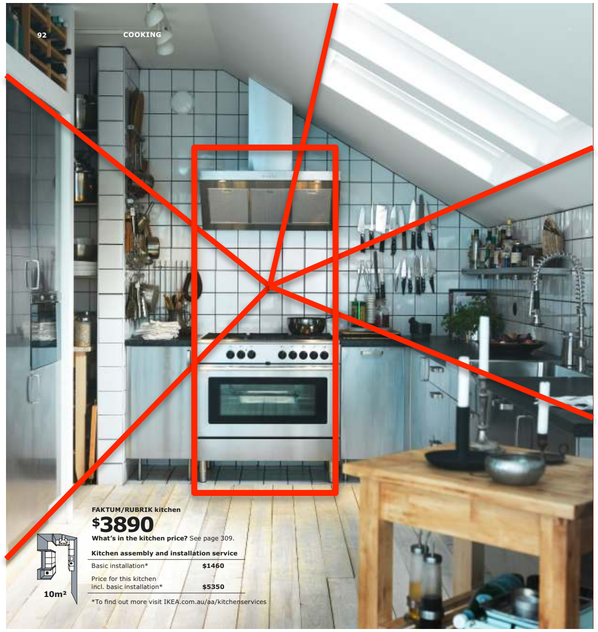
Limitations and future work

- Interactive scene optimization
- Global illumination
- Additional composition rules

Limitations and future work

- Interactive scene optimization
- Global illumination
- Additional composition rules

Symmetry

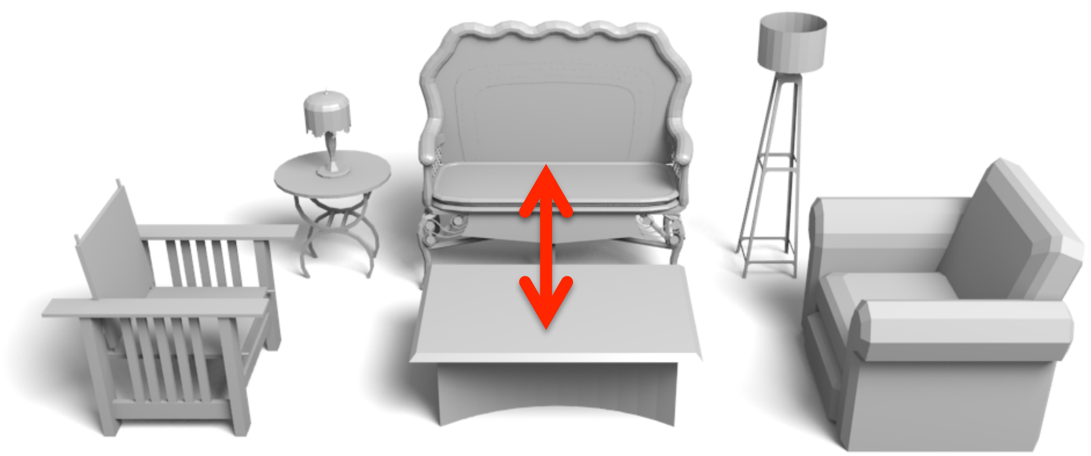


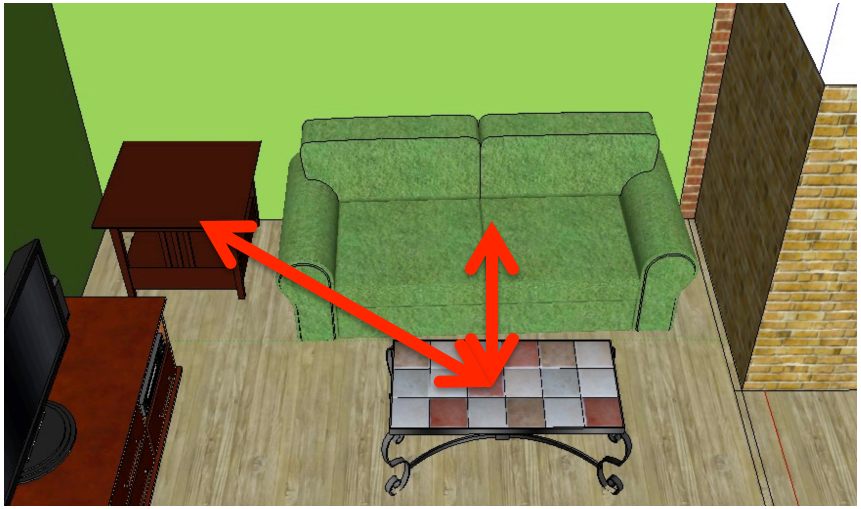
Vanishing points

Outline

- Analyzing 3D scenes by modeling hierarchical structure
- Composition-aware scene optimization for product images
- Style compatibility for 3D furniture models

Motivation





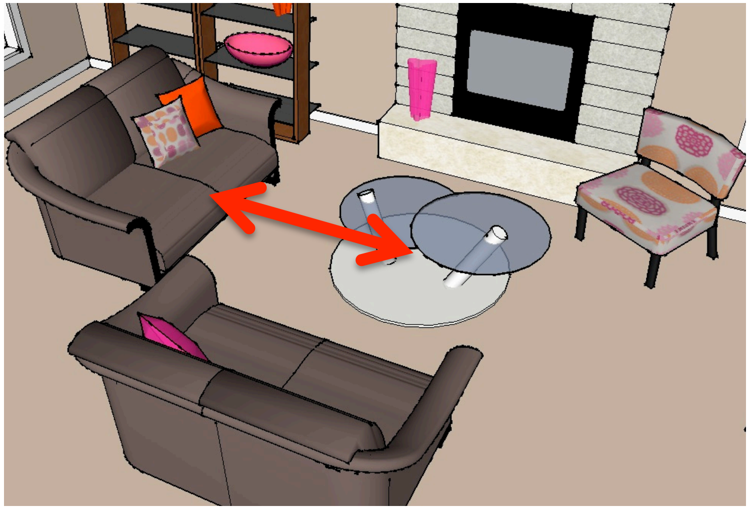
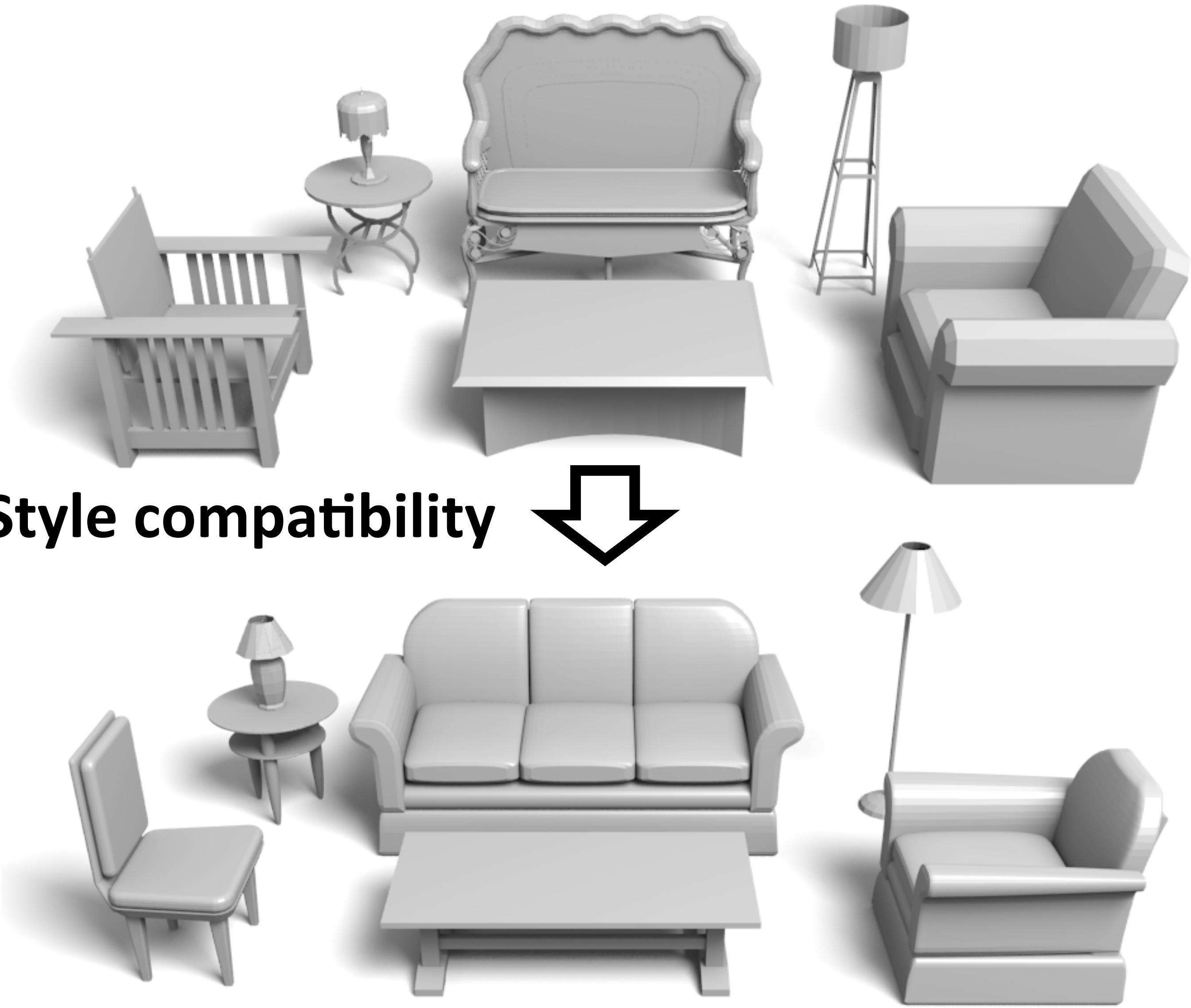


Image courtesy: smartnick100, Designer_Tina, Xu et al.

Motivation



Goal

Modeling pairwise style compatibility

How likely would a person put the two furniture pieces together in the same room if he was furnishing an apartment?

Goal

Modeling pairwise style compatibility

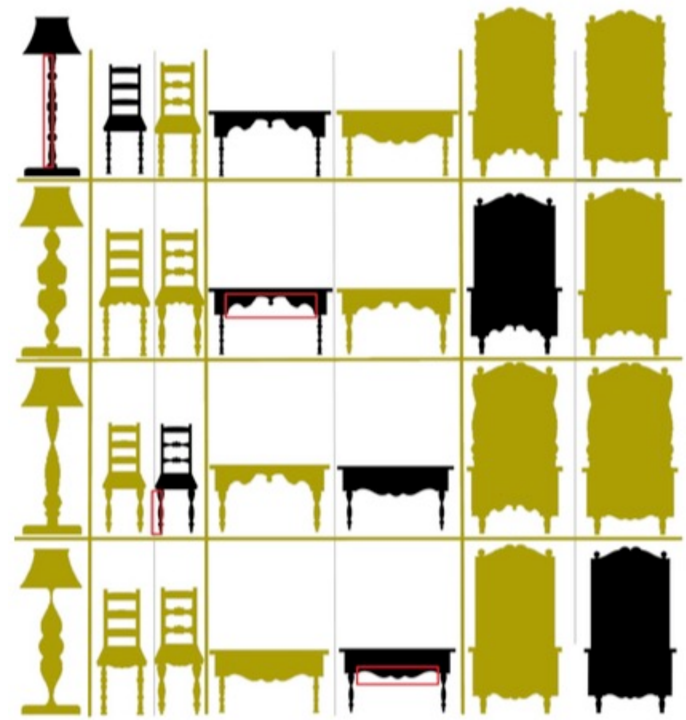
 

Extract feature vectors

$$F(x_i, x_j) = Scalar$$

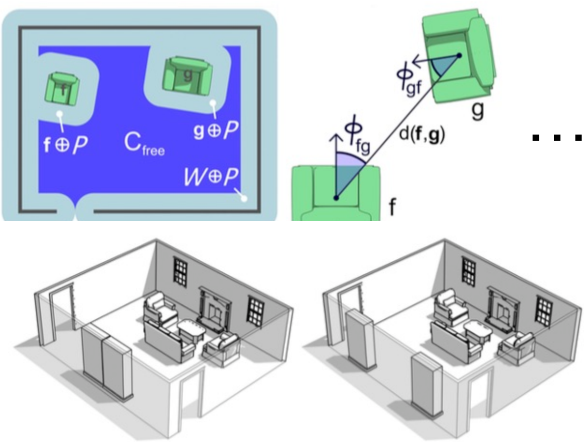
Previous work – shape style

[Xu et al. 2010]



[Li et al. 2013]

Previous work – virtual world synthesis



[Merrell et al. 2011] [Fisher et al. 2012] [Xu et al. 2013]

Challenges

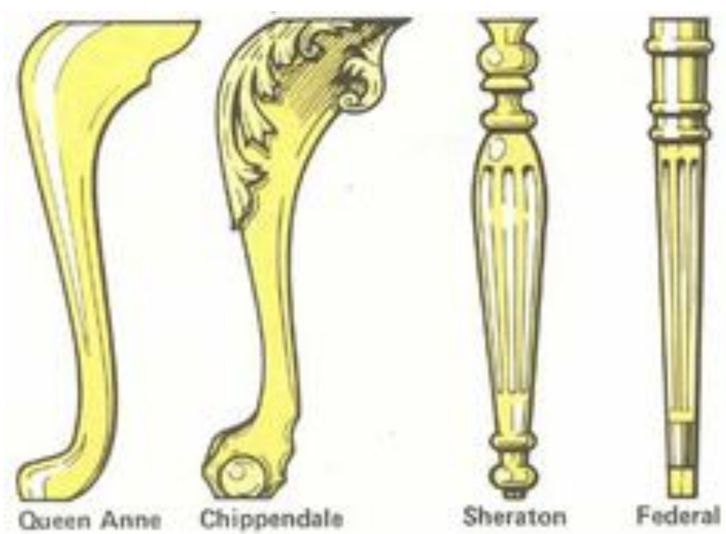
- Hard to design a hand-tuned function
- Coupled with functionality
- Requiring comparisons across object classes

Challenges

- Hard to design a hand-tuned function
- Coupled with functionality
- Requiring comparisons across object classes

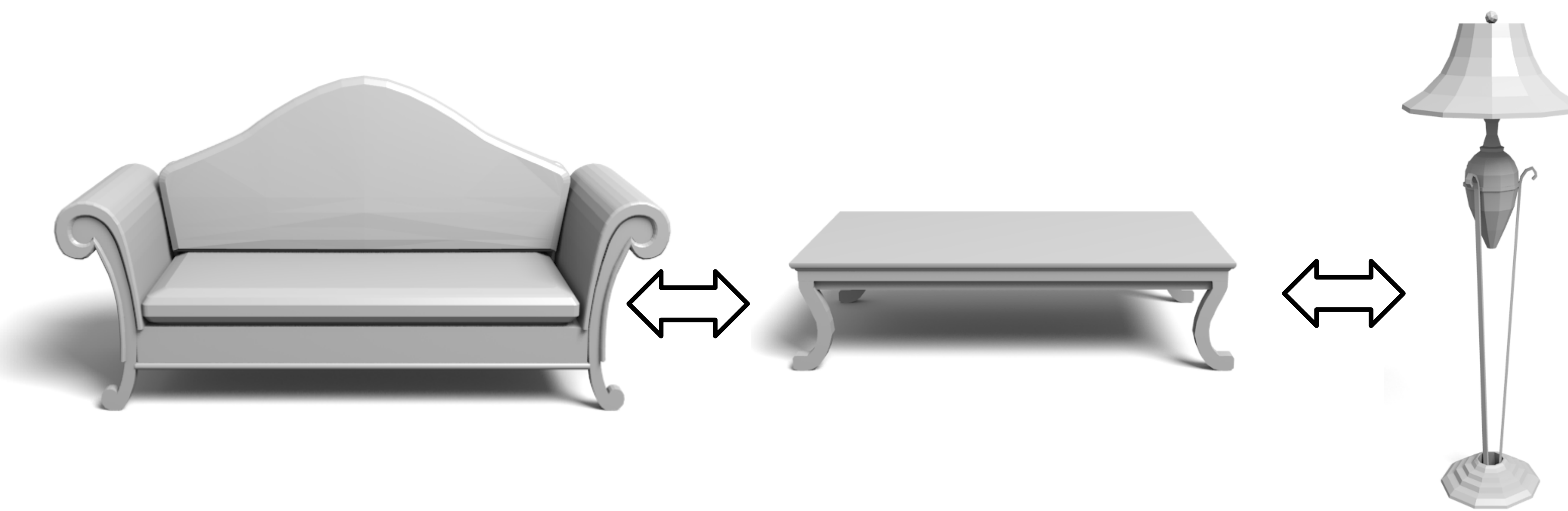
Challenges

- Hard to design a hand-tuned function
- Coupled with functionality
- Requiring comparisons across object classes



Challenges

- Hard to design a hand-tuned function
- Coupled with functionality
- Requiring comparisons across object classes

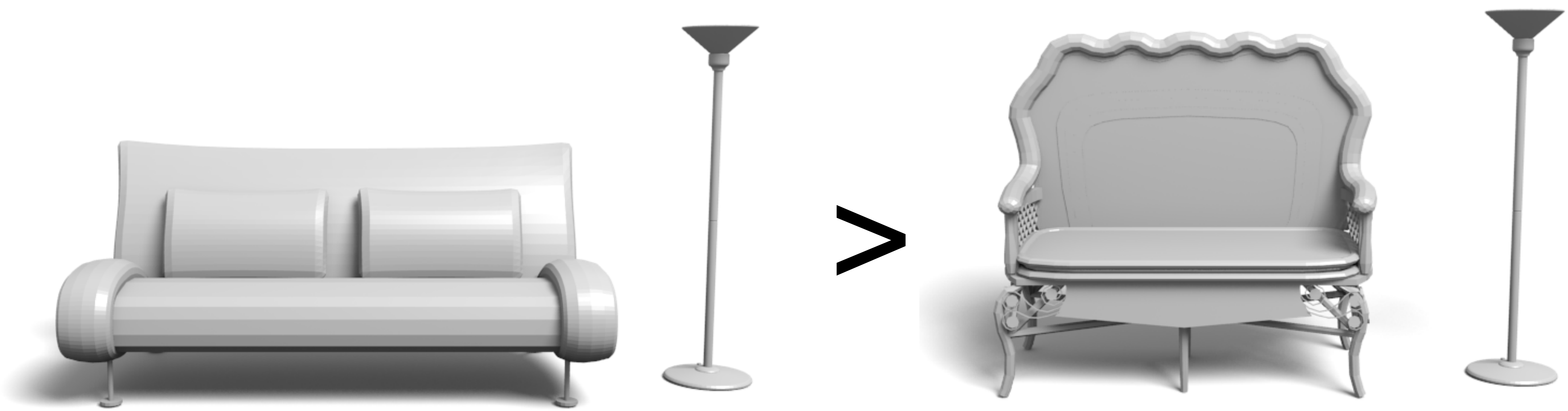


Key Ideas

- Crowdsourcing compatibility preferences
- Part-aware geometric features
- Learning object-class specific mappings

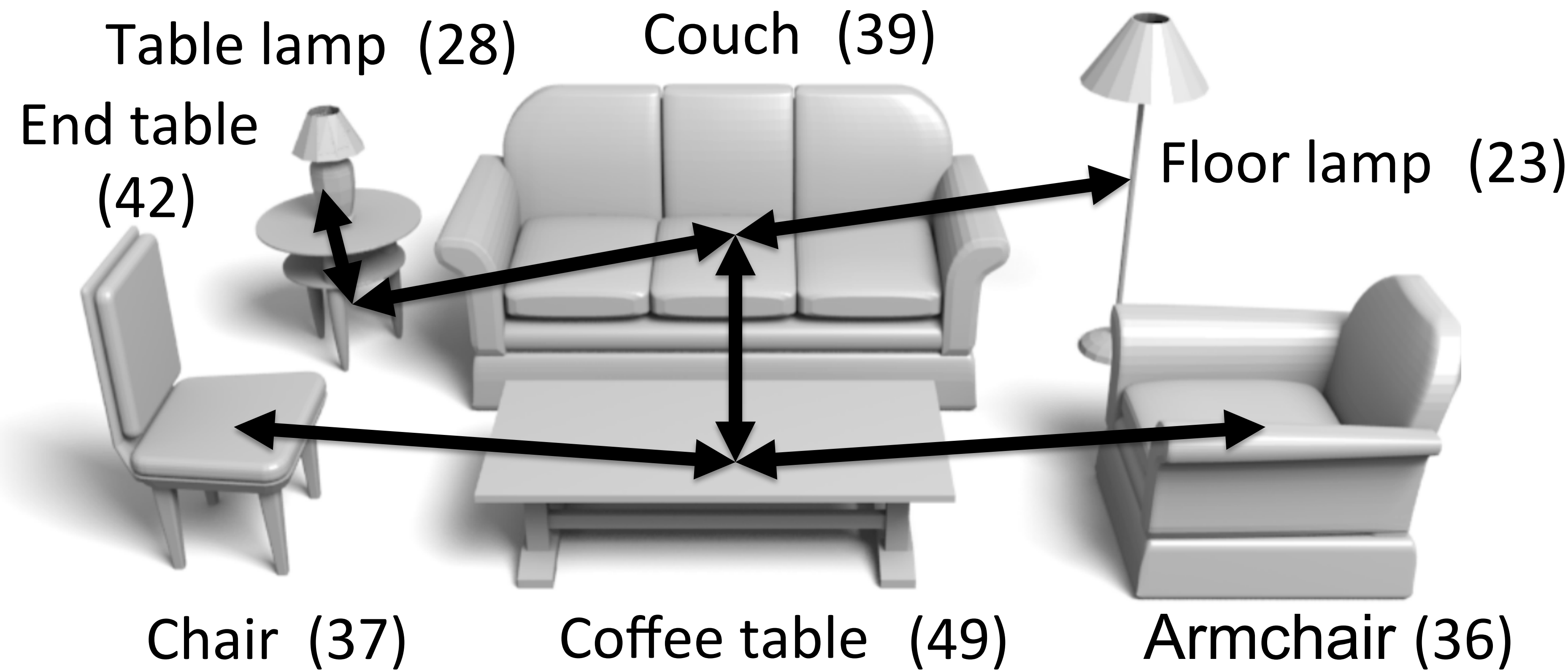
Key Ideas

- Crowdsourcing compatibility preferences
- Part-based geometric features
- Learning object-class specific mappings



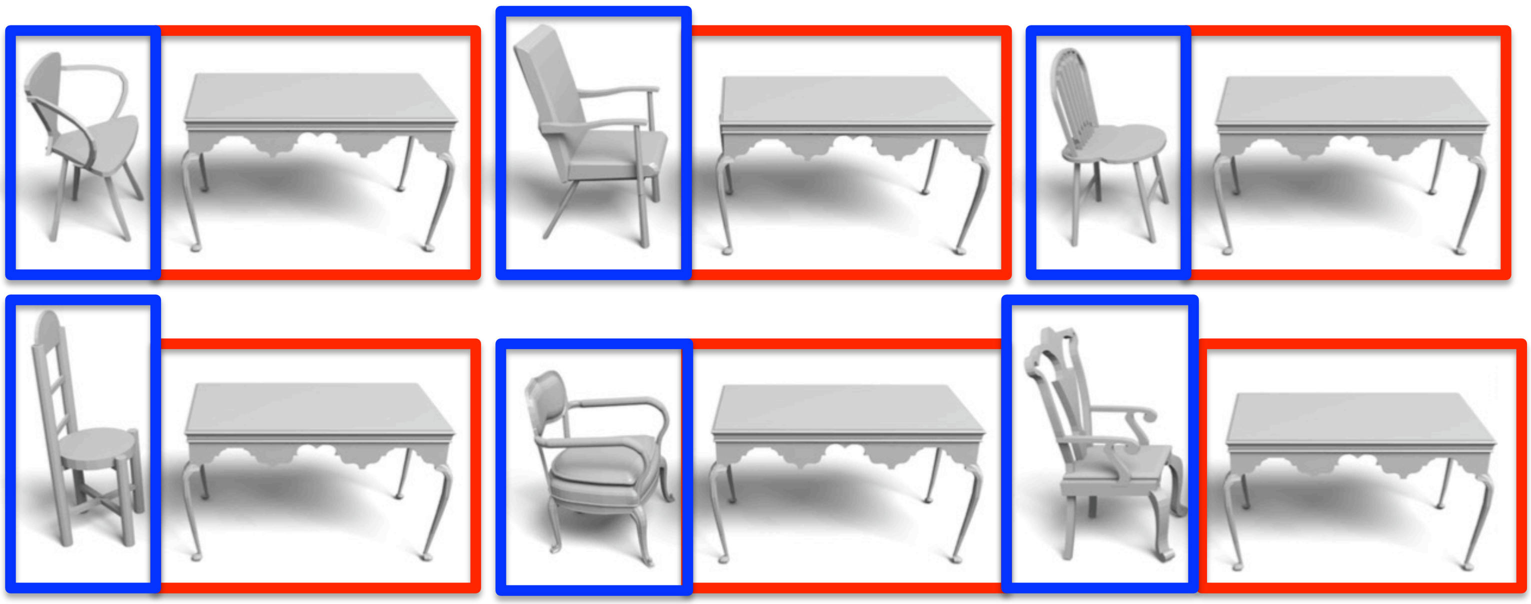
Crowdsourcing compatibility preferences

Living room



Crowdsourcing compatibility preferences

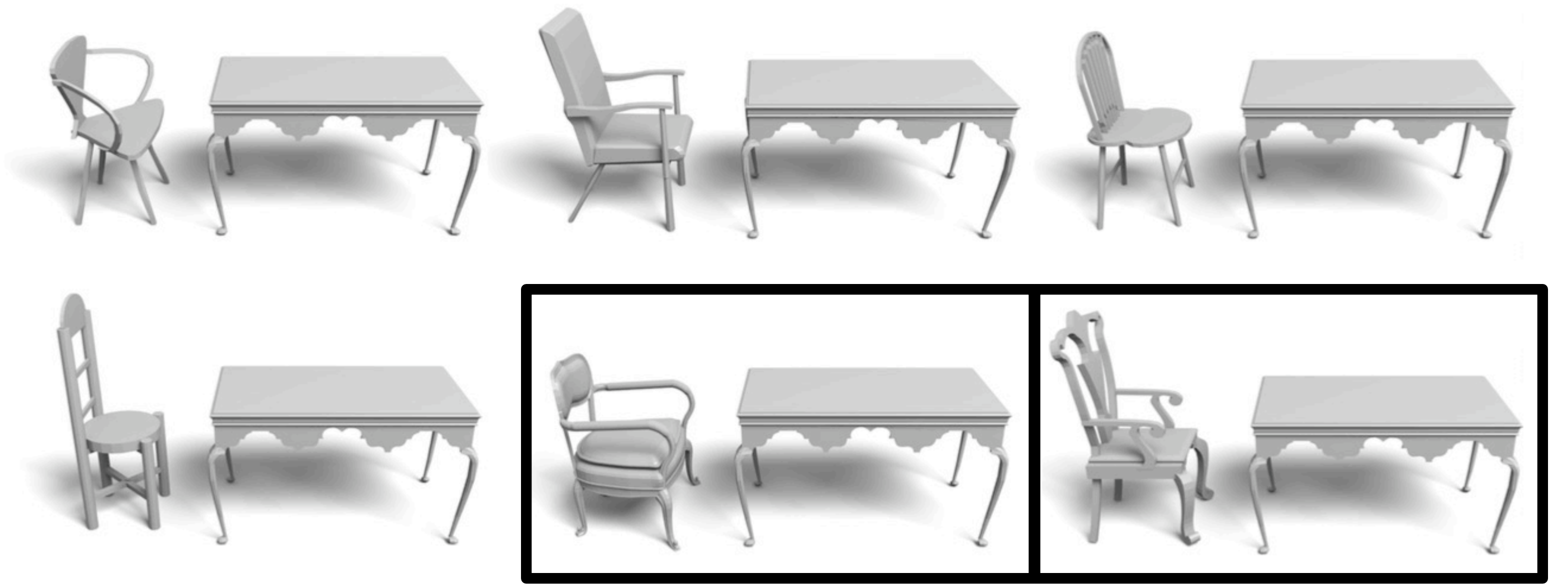
Design of user study [Wilber et al. 2014]



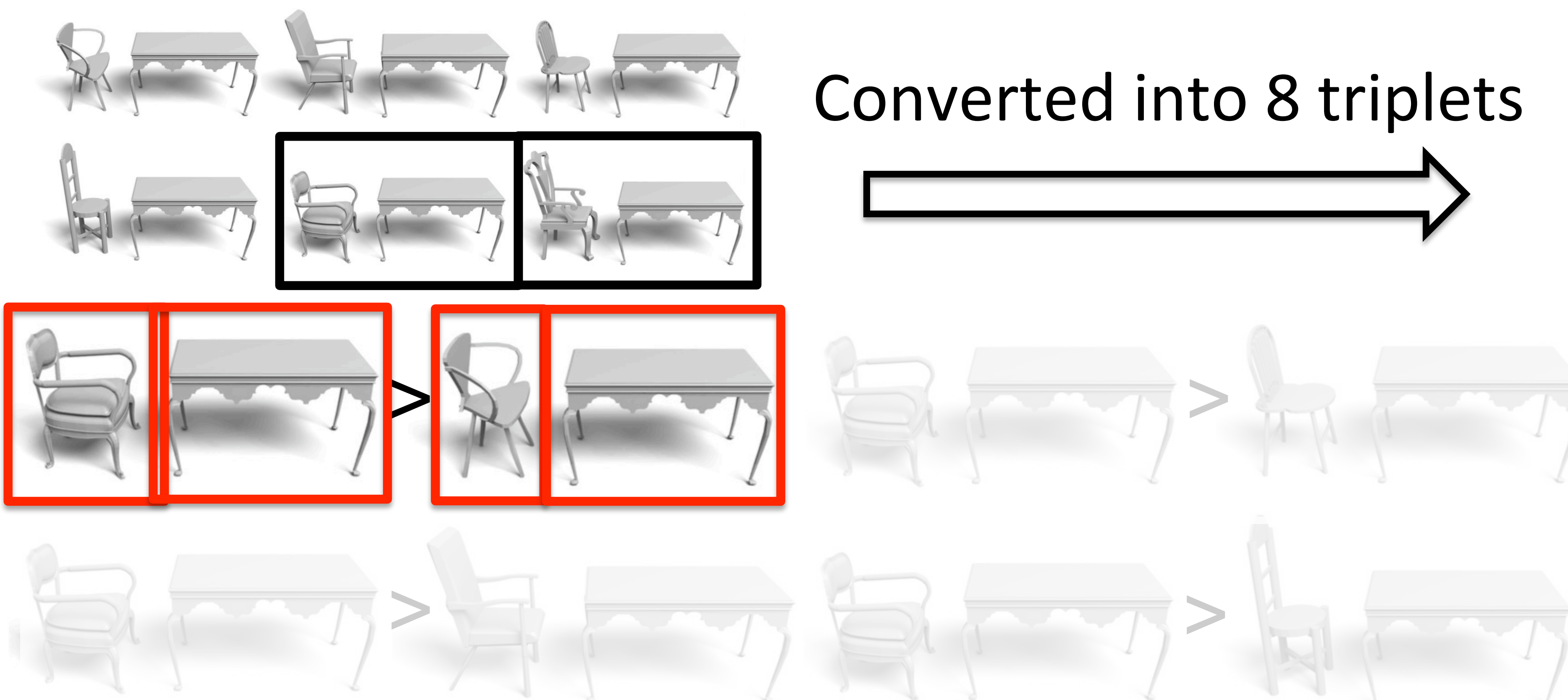
Please select the two most compatible pairs.

Crowdsourcing compatibility preferences

Rater's selection

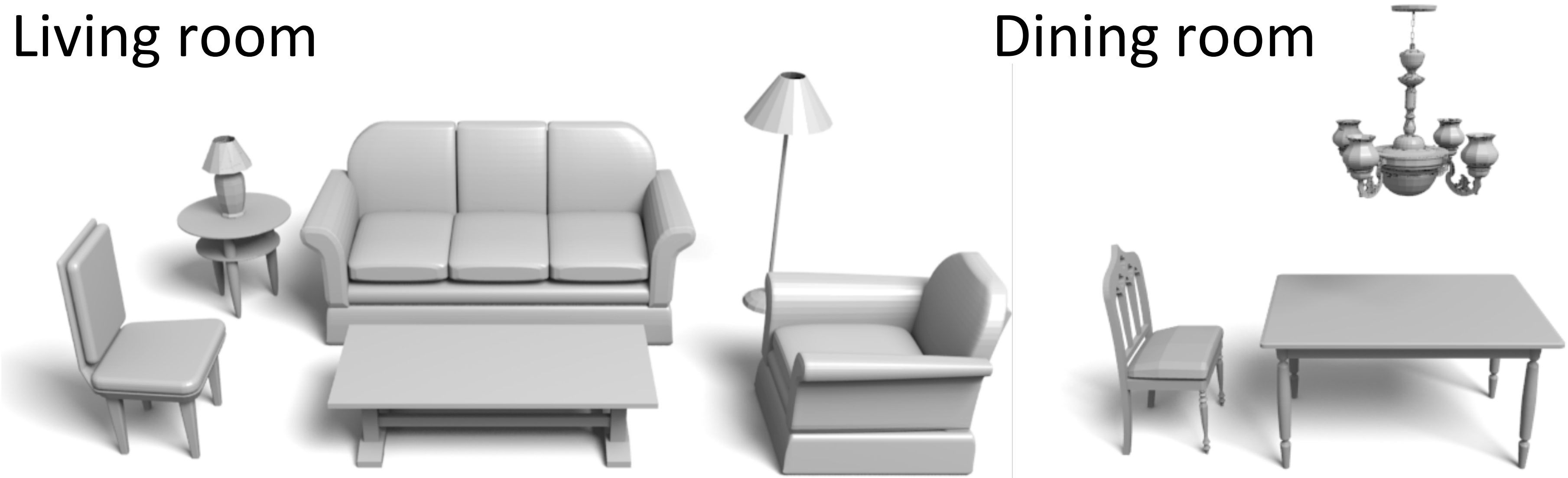


Crowdsourcing compatibility preferences



and 4 more triplets ...

Crowdsourcing compatibility preferences



Collected *63,800* triplets for living room
and *20,200* for dining room

Key Ideas

- Crowdsourcing compatibility preferences
- Part-aware geometric features
- Learning object-class specific mappings

Part-aware geometric features

Contemporary

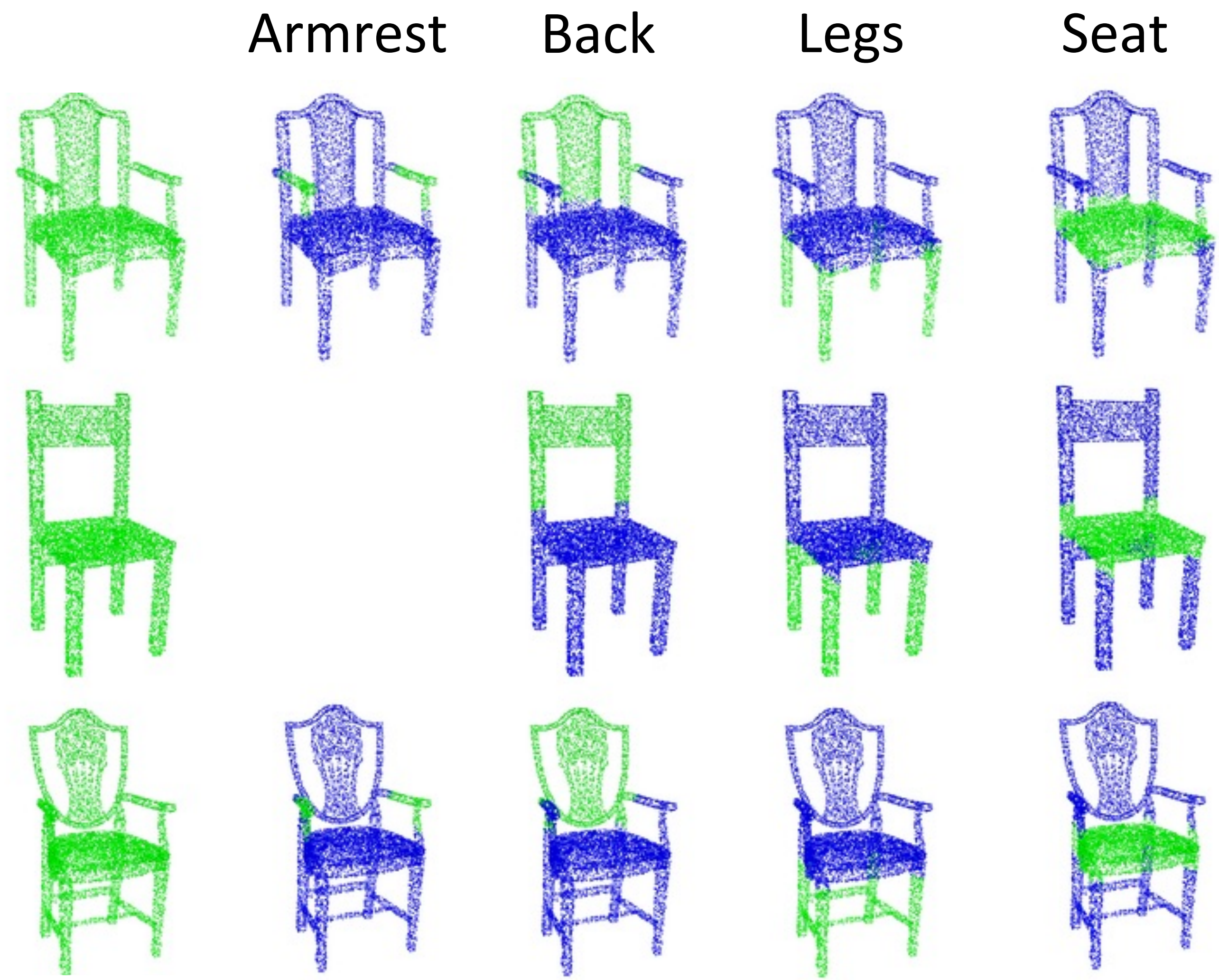
Antique

Part-aware geometric features

- Consistent segmentation
- Computing geometry features for each part
- Concatenating features of all parts

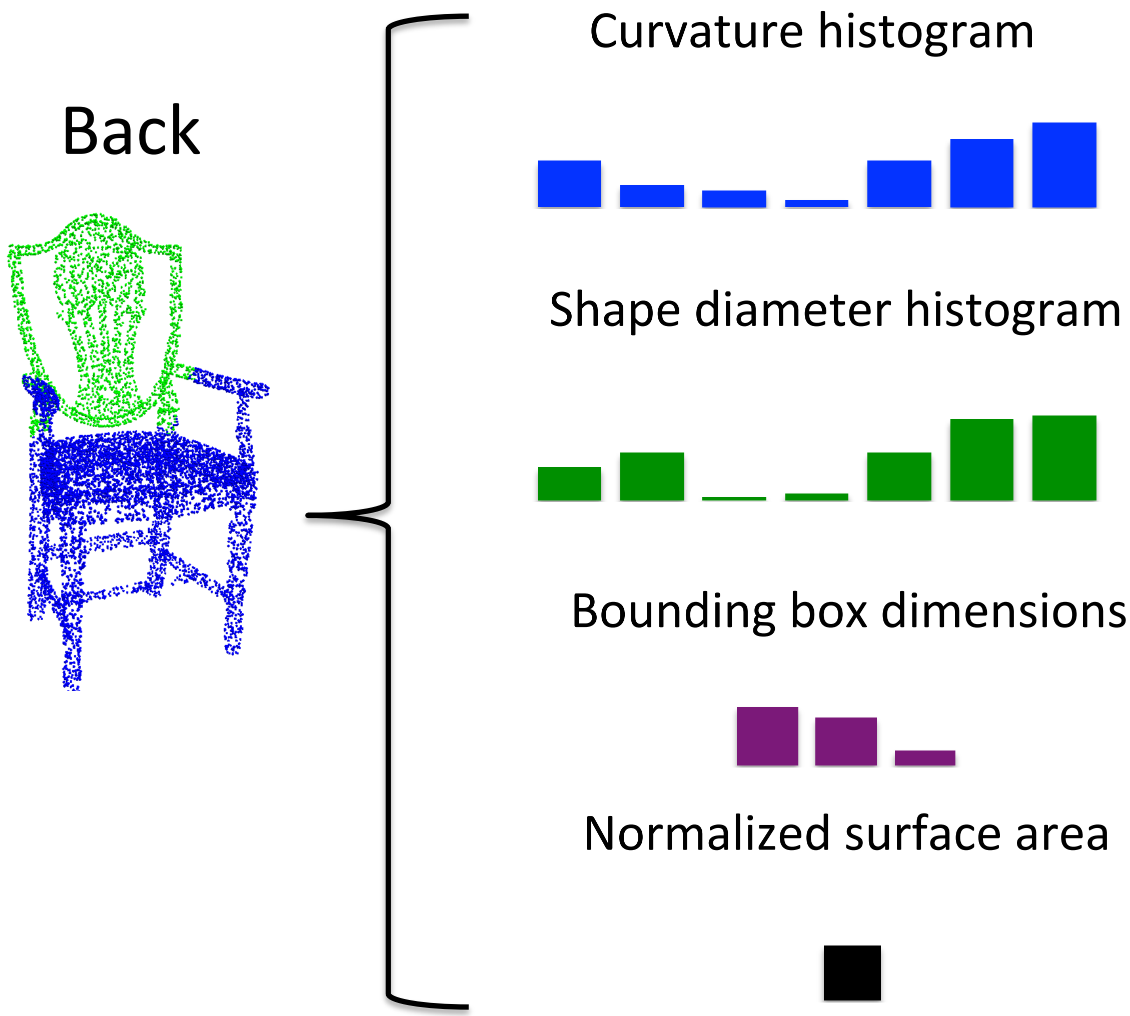
Part-aware geometric features

Step 1: Consistent segmentation [Kim et al. 2013]



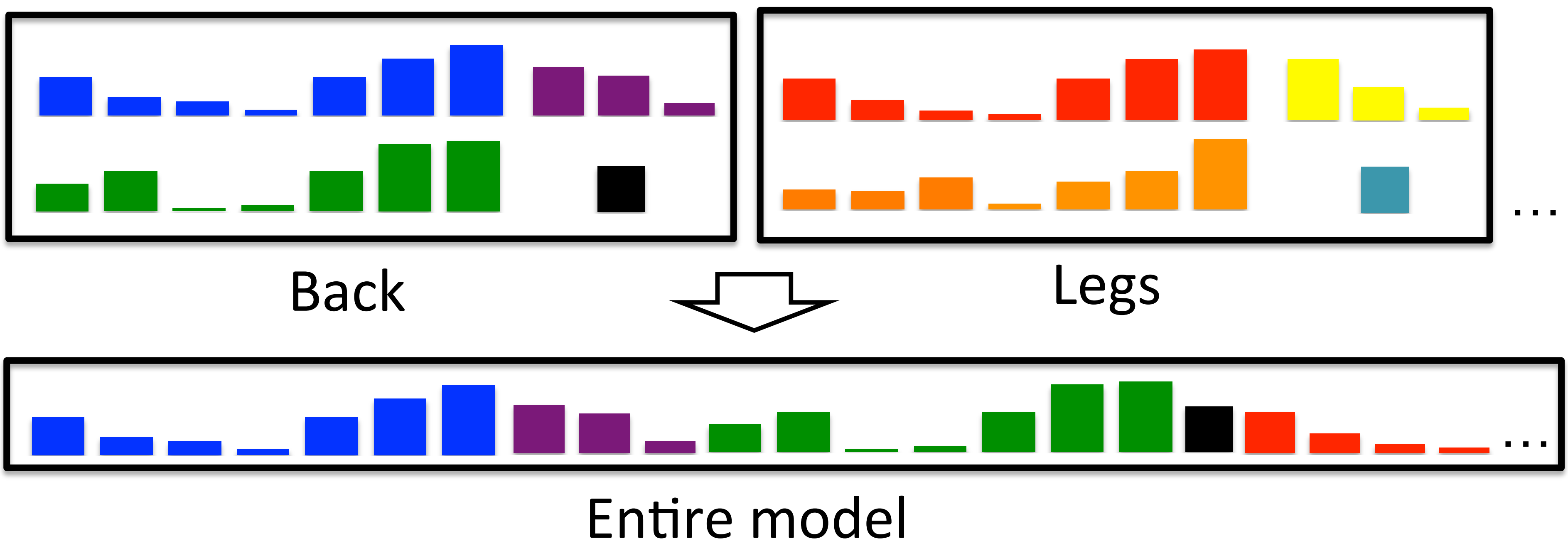
Part-aware geometric features

Step 2: Computing geometric features for each part



Part-aware geometric features

Step 3: Concatenating features of all parts



Key Ideas

- Crowdsourcing compatibility preferences
- Part-aware geometric features
- Learning object-class specific mappings

Learning object-class specific mappings

Previous approach [Kulis 2012]:

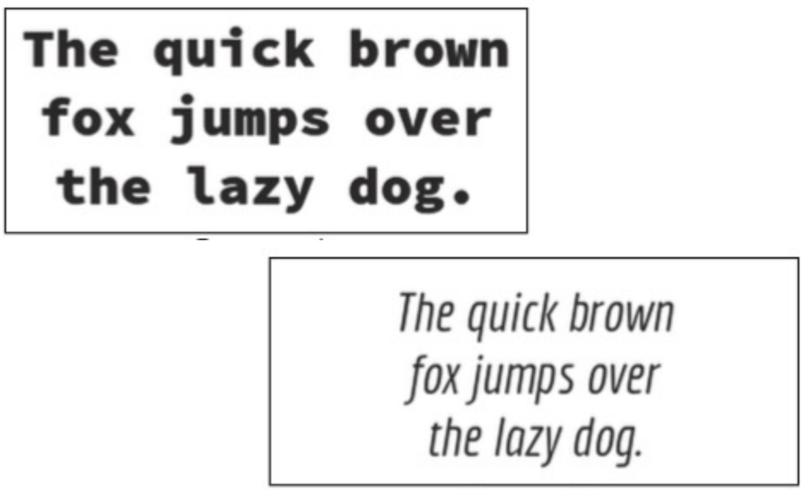
$$d_{symm}(x_i, x_j) = \left\| W(x_i - x_j) \right\|_2$$

d_{symm} is the compatibility distance

x_i, x_j are feature vectors of two shapes

Learning object-class specific mappings

Previous approach [Kulis 2012]:



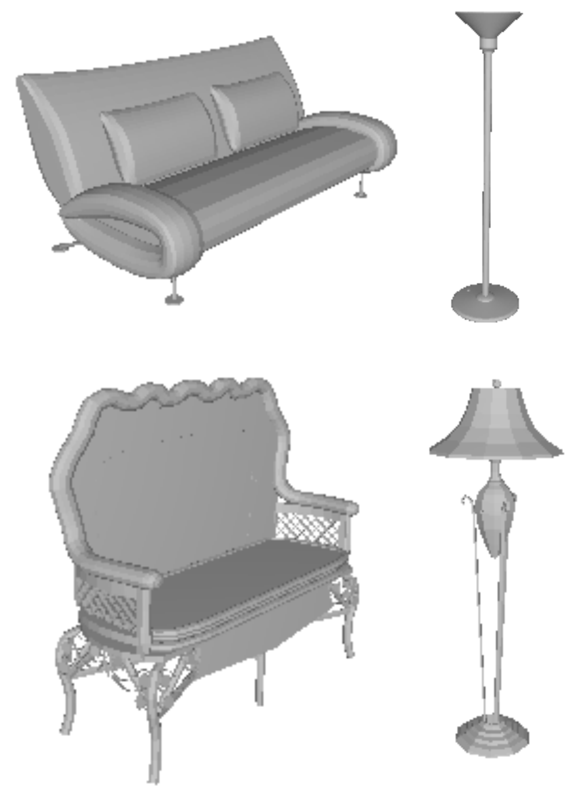
Fonts
[O'Donovan et al. 2014]

Illustration styles
[Garces et al. 2014]

Learning object-class specific mappings

Assumptions of the previous approach

- Feature vectors have the same dimensionality
- Corresponding dimensions are comparable



4 parts 3 parts

Learning object-class specific mappings

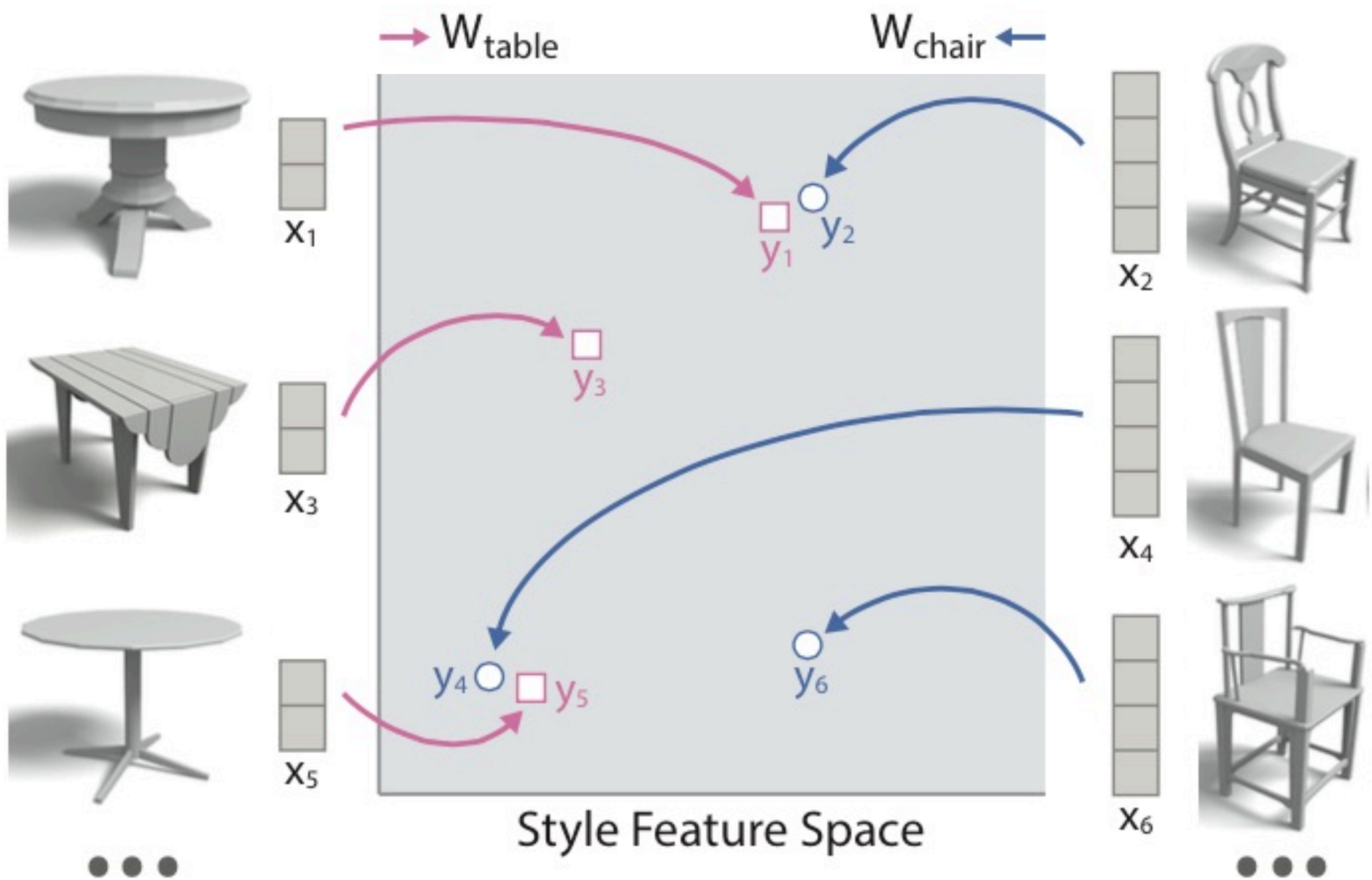
Our approach:

$$d_{asymm}(x_i, x_j) = \left\| W_{c(i)} x_i - W_{c(j)} x_j \right\|_2$$

$c(i)$ is the object class of x_i

$c(j)$ is the object class of x_j

Learning object-class specific mappings



Learning object-class specific mappings

Learning procedure [O'Donovan et al. 2014]

- Using a logistic function to model rater's preferences
- Learning by maximizing the likelihood of the training triplets with regularization

Results of triplet prediction

Test set: triplets that human agree upon

- 264 triplets from dining room
- 229 triplets from living room

Method	Dining room	Living room
Chance	50%	50%
No part-aware, Symmetric	63%	55%
Part-aware, Symmetric	63%	65%
No part-aware, Asymmetric	68%	65%
Part-aware, Asymmetric (Ours)	73%	72%
People	93%	99%

Applications

- Style-aware shape retrieval
- Style-aware furniture suggestion
- Style-aware scene building

Applications

- Style-aware shape retrieval
- Style-aware furniture suggestion
- Style-aware scene building

Style-aware shape retrieval

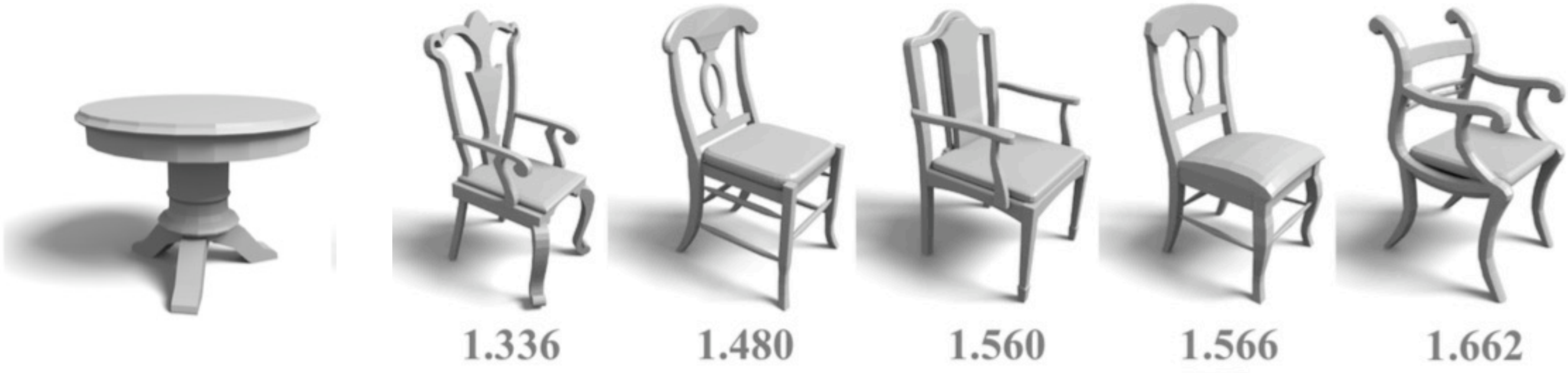
Query model

Dining chair

?

Style-aware shape retrieval

Query model Dining chair



Style-aware shape retrieval

Query model

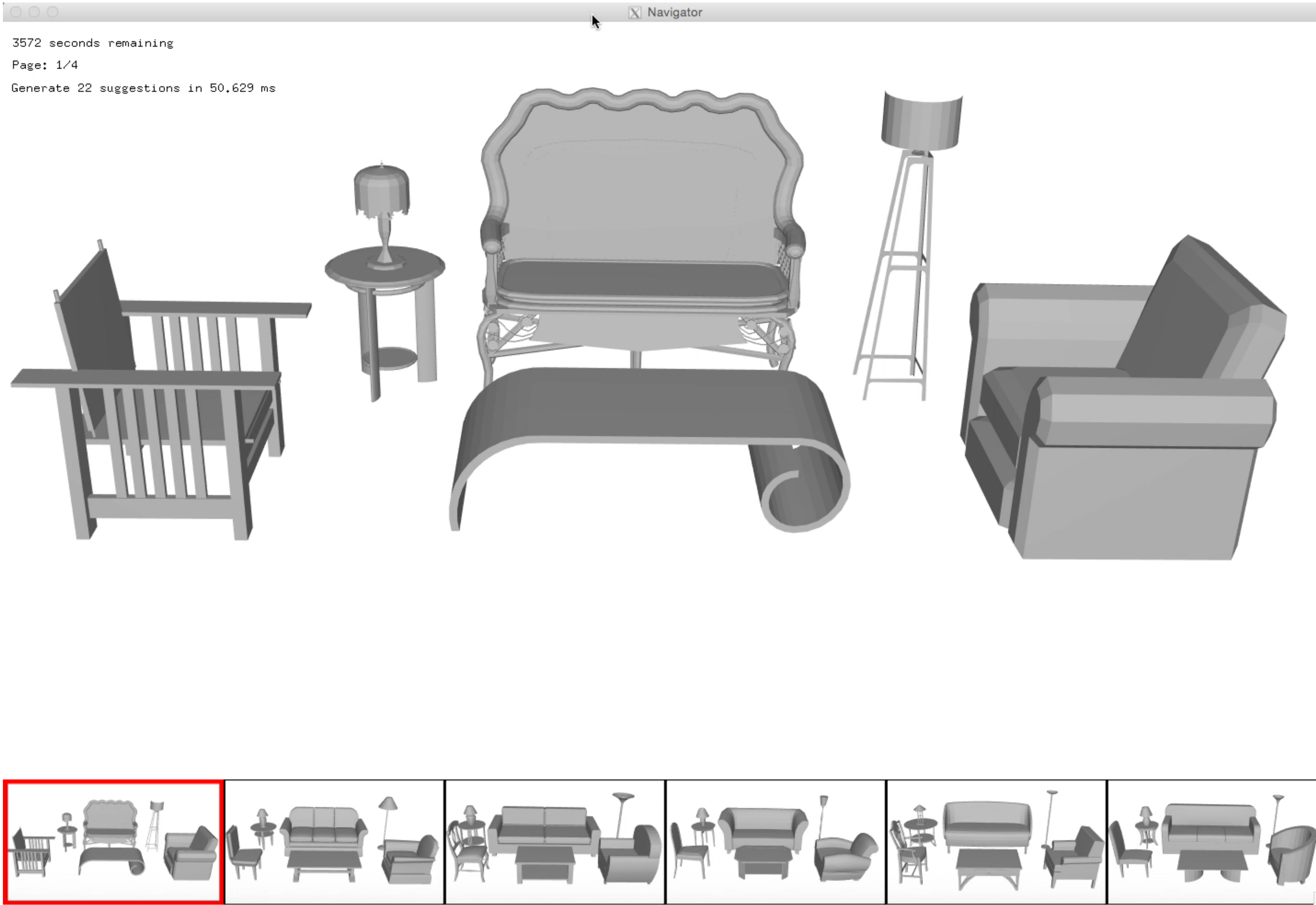
Dining chair

1.336 1.480 1.560 1.566 1.662

(Most incompatible chairs)

2.790 2.847 3.149 3.246 3.525

Style-aware scene building

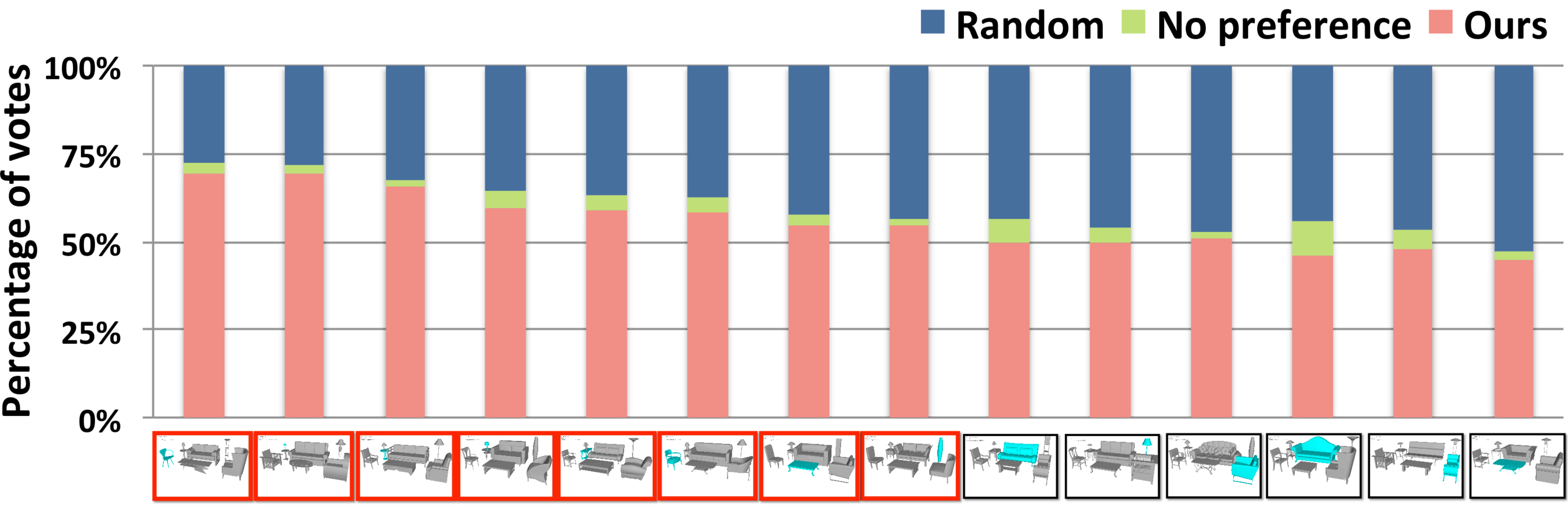


Style-aware scene building

User study

- 12 participants, each works on 14 tasks.
- Half of the tasks are assisted by our metric, and the other half are not.
- Results from the two settings are compared on Amazon Mechanical Turk.

Style-aware scene building



Summary

- It is possible to learn a compatibility metric for furniture of different classes.
- The learned compatibility metric is effective in style-aware scene modeling.

Limitations and future work

- Modeling fine-grained style variations
- Investigating style compatibility in other domains

Limitations and future work

- Modeling fine-grained style variations
- Investigating style compatibility in other domains

Duncan Phyfe style with eagle motif
(Courtesy: Carswell Rush Berlin)

Sheraton style with lyre motif

Limitations and future work

- Modeling fine-grained style variations
- Investigating style compatibility in other domains

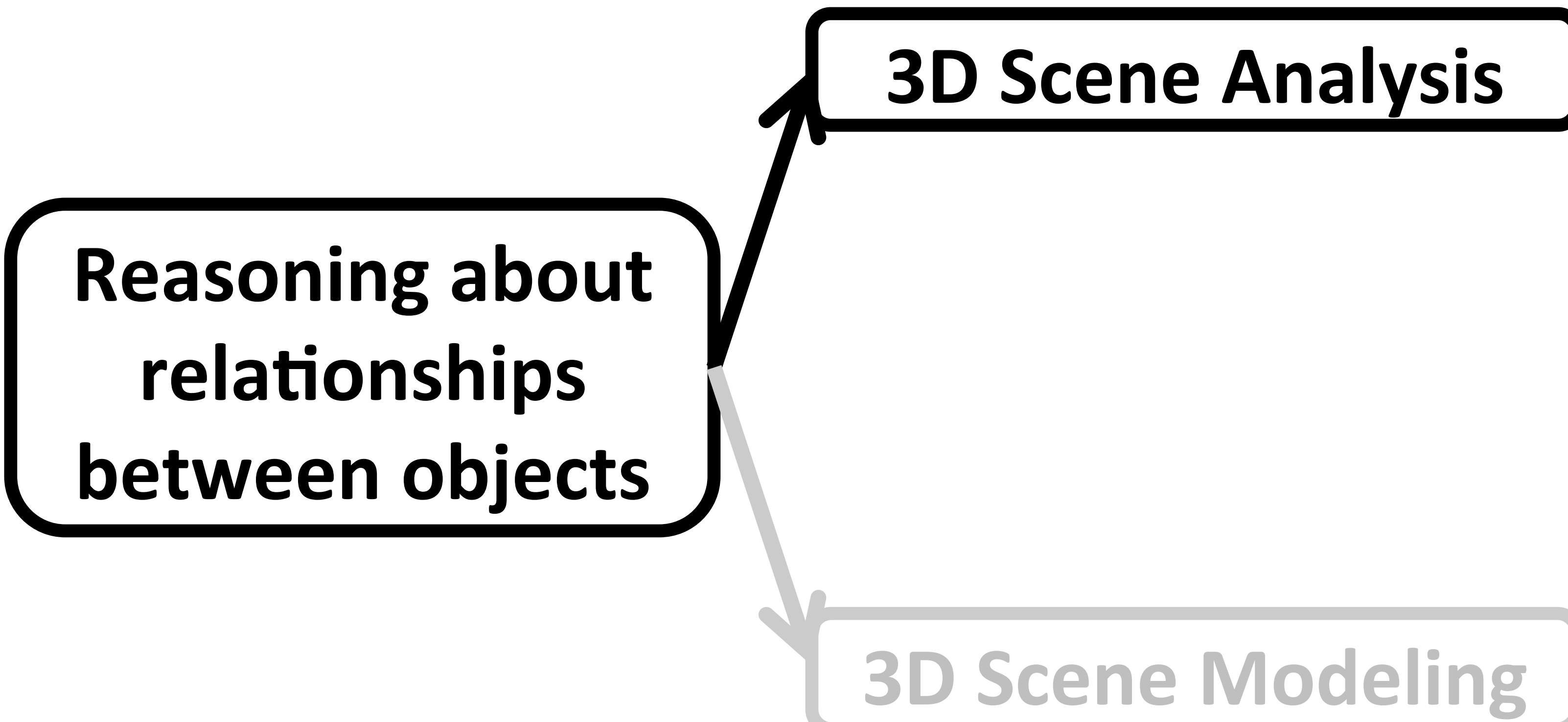
Outline

- Analyzing 3D scenes by modeling hierarchical structure
- Composition-aware scene optimization for product images
- Style compatibility for 3D furniture models

Summary of my thesis

Relationships between objects are

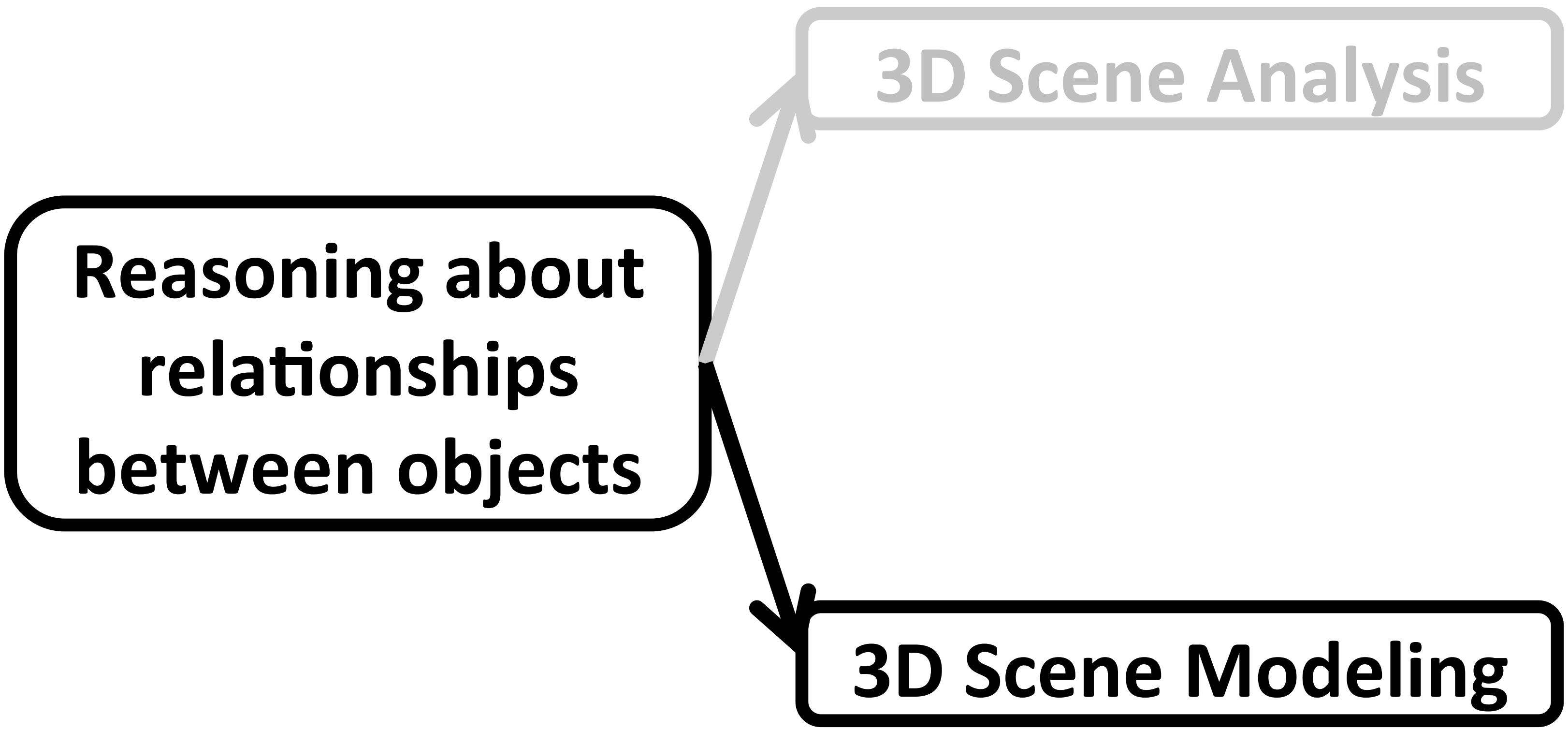
- a strong cue for scene understanding
- a strong factor for scene plausibility and aesthetics



Summary of my thesis

Relationships between objects are

- a strong cue for scene understanding
- a strong factor for scene plausibility and aesthetics



Future work

- Other sources for data-driven scene modeling
- Other factors related to scene plausibility

Future work

- Other sources for data-driven scene modeling
- Other factors related to scene plausibility

Image courtesy: IKEA

Future work

- Other sources for data-driven scene modeling
- Other factors related to scene plausibility

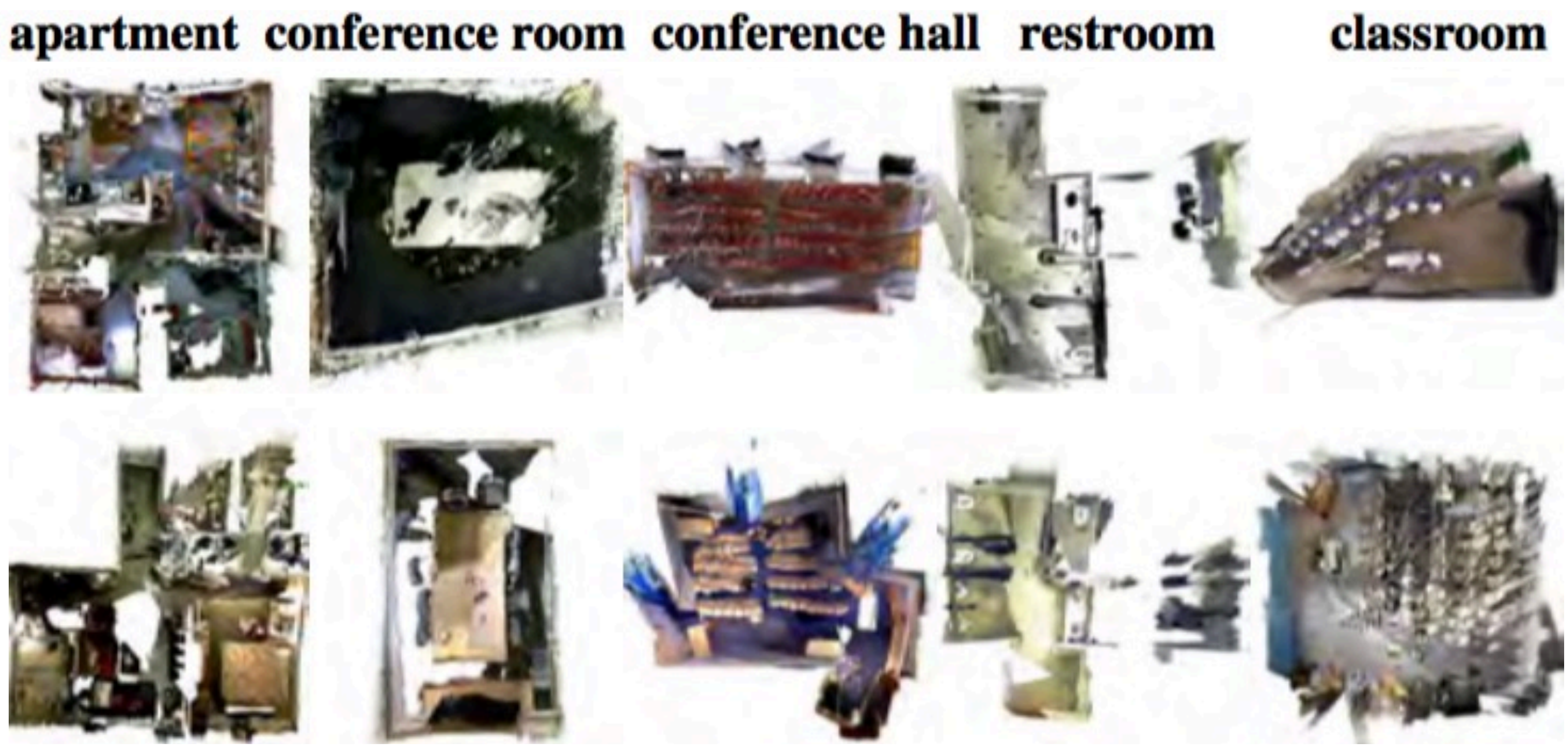


Image courtesy: Xiao et al.

Future work

- Other sources for data-driven scene modeling
- Other factors related to scene plausibility

Materials are strongly related to style compatibility

Acknowledgements

Adviser: Thomas Funkhouser

Mentors:
Wilmot Li, Jim McCann, Aaron Hertzmann




Acknowledgements

Collaborators

Niloy Mitra

Vladimir Kim

Sid Chaudhuri

Qi-Xing Huang

Acknowledgements

Princeton Graphics Group

Acknowledgements

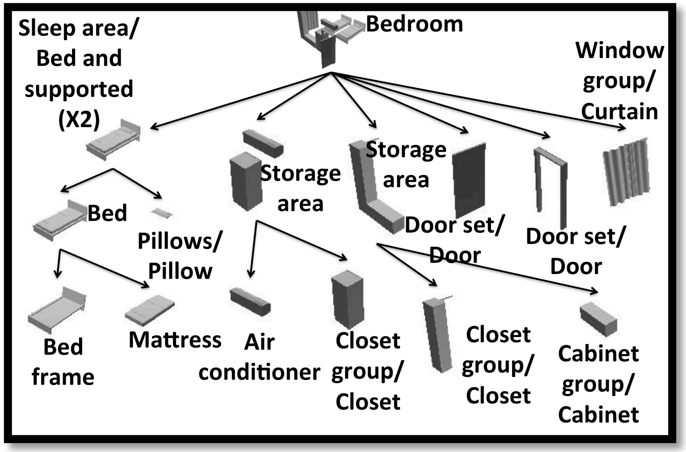
Family

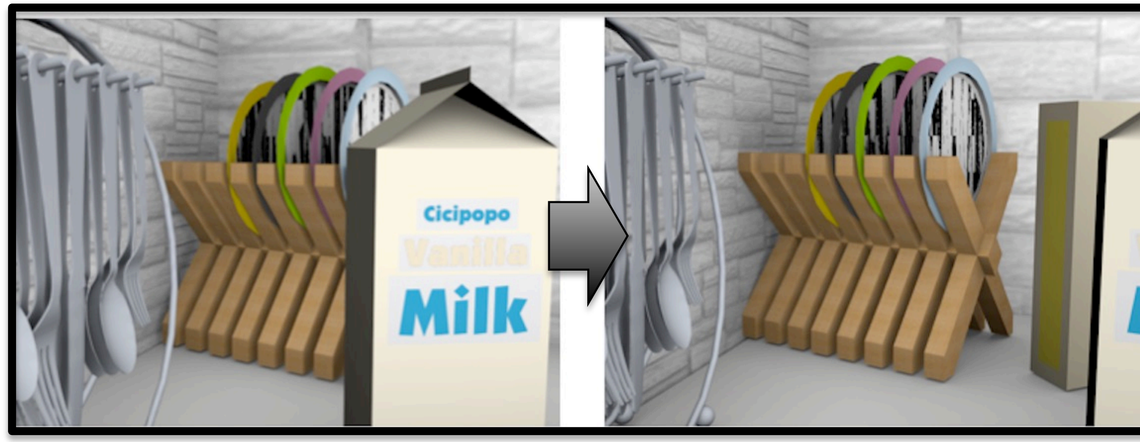
Acknowledgements

Funding agencies

- Adobe
- NSF, ERC Starting Grant, Intel, Google

Thank you!


Bedroom
Sleep area/
Bed and
supported
(X2)
Window
group/
Curtain
Storage
area
Storage
area
Door set/
Door
Door set/
Door
Bed
Pillows/
Pillow
Bed
frame
Mattress
Air
conditioner
Closet
group/
Closet
Closet
group/
Closet
Cabinet
group/
Cabinet


Cicipopo
Vanilla
Milk


VET TES EN
NOV TAM TVM
DEI
SVB NVMINE
VIGET
®