Eurographics 2015

The 36th Annual Conference of the
European Association for Computer Graphics

Composition-Aware Scene Optimization for Product Images

Tianqiang Liu[1] Jim McCann[2,3] Wilmot Li[2] Thomas Funkhouser[1]

[1]Princeton University [2]Adobe Research [3]Disney Research

Motivation

Motivation

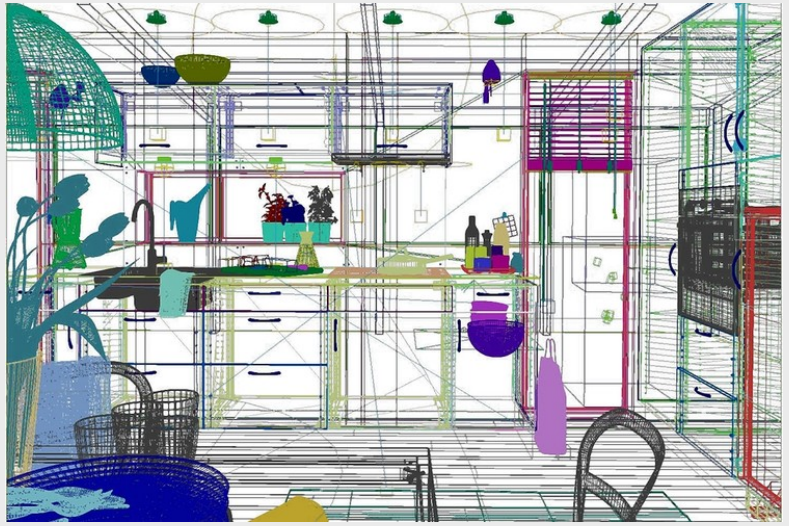
35% of scenes in IKEA catalogue are CGI.

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

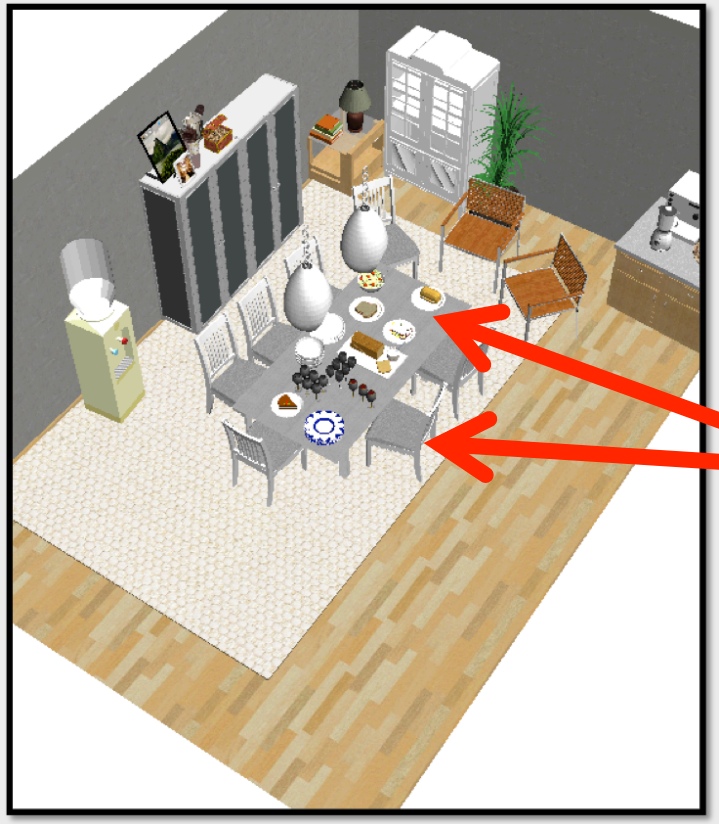
Artist's goal

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

Camera view

Artist's goal

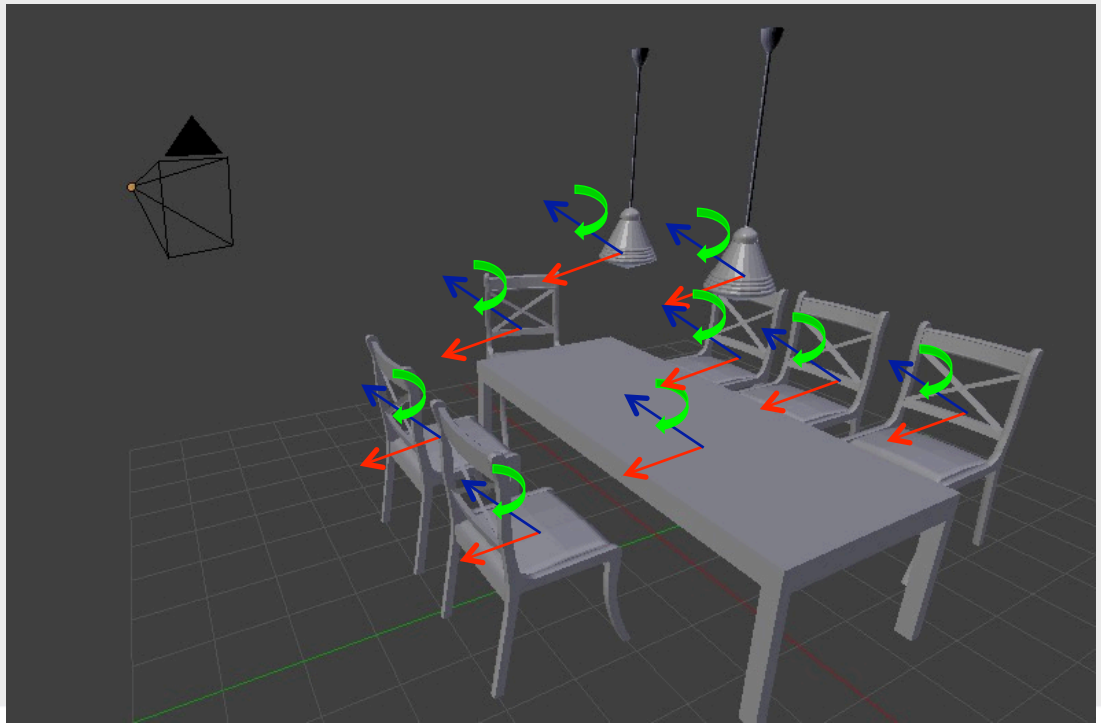
Output: a scene with optimized *object placement*, *materials* and *camera view* that produce an appealing 2D composition.

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

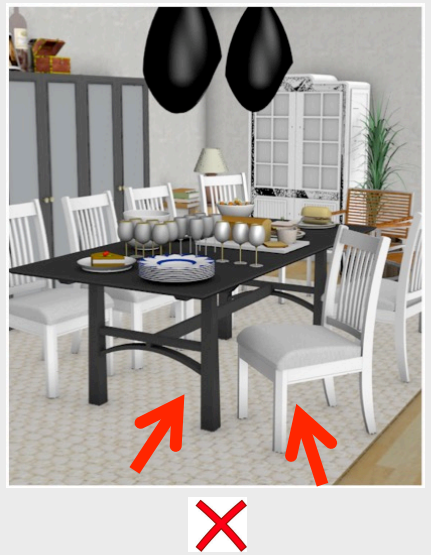


4*N + 6 parameters

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

Challenges

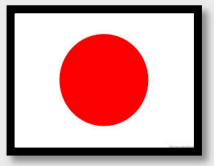
- Huge search space to explore
- Many principles/constraints to balance
- Requiring repeating work for customization

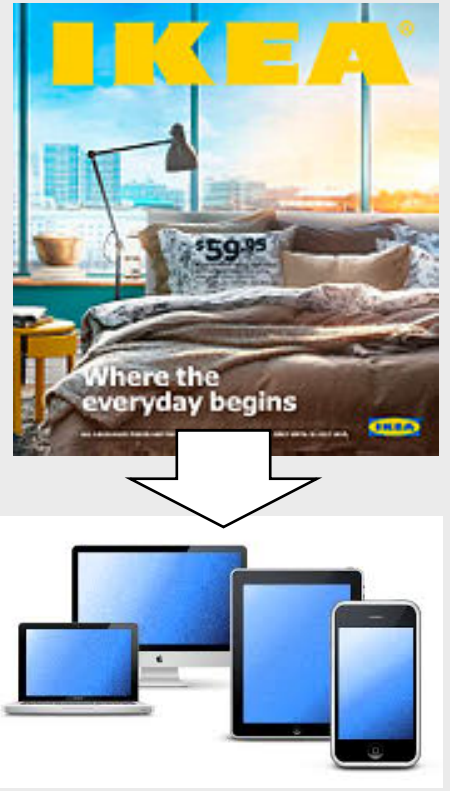


...

Challenges

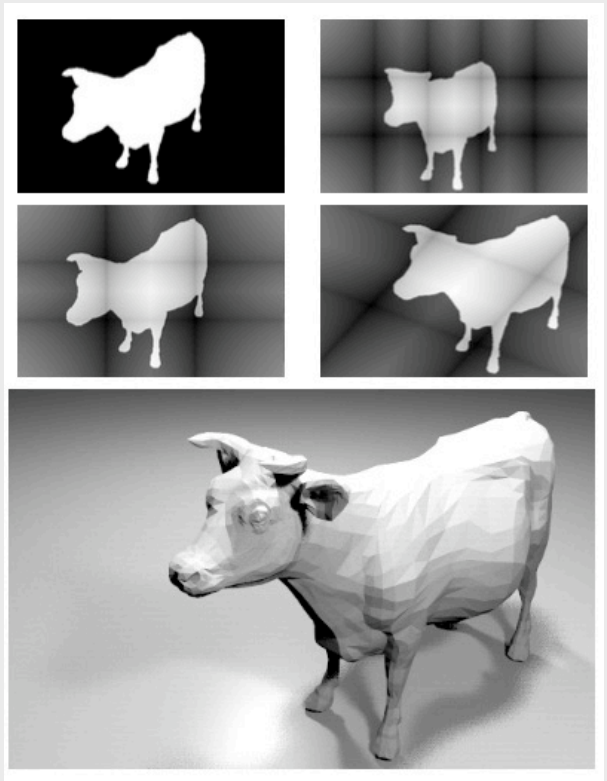
- Huge search space to explore
- Many principles/constraints to balance
- Requiring repeating work for customization



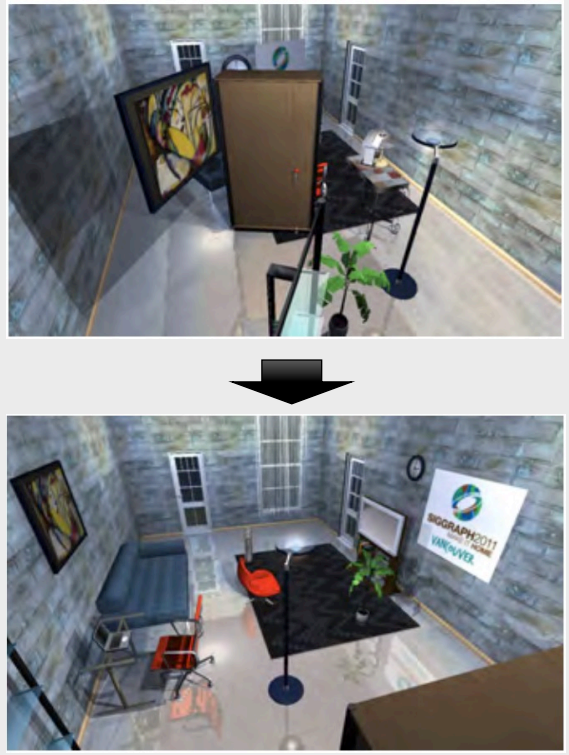


Previous Work

Image optimization
[Liu et al. 2010]



Camera optimization
[Gooch et al. 2001]



Scene optimization
[Yu et al. 2011]

Key Idea

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

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

Key Idea

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

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

θ_i : orientation of object i

Key Idea

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

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

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

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

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

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

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

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 Never been considered before

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

Overview

- Composition rules and constraints
- Optimization
- Applications

Terms for composition rules and constraints

1. Object placement within 2D frame E_{op}
2. Object saliency within the 2D frame E_{os}
3. Image composition E_{ic}
4. Camera placement E_{cp}
5. Object constraints within the 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 the 2D frame

- Visibility
- Object size

Term 2: Object saliency within the 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 the 3D scene

- Collision relationships
- Support relationships
- Semantic constraints

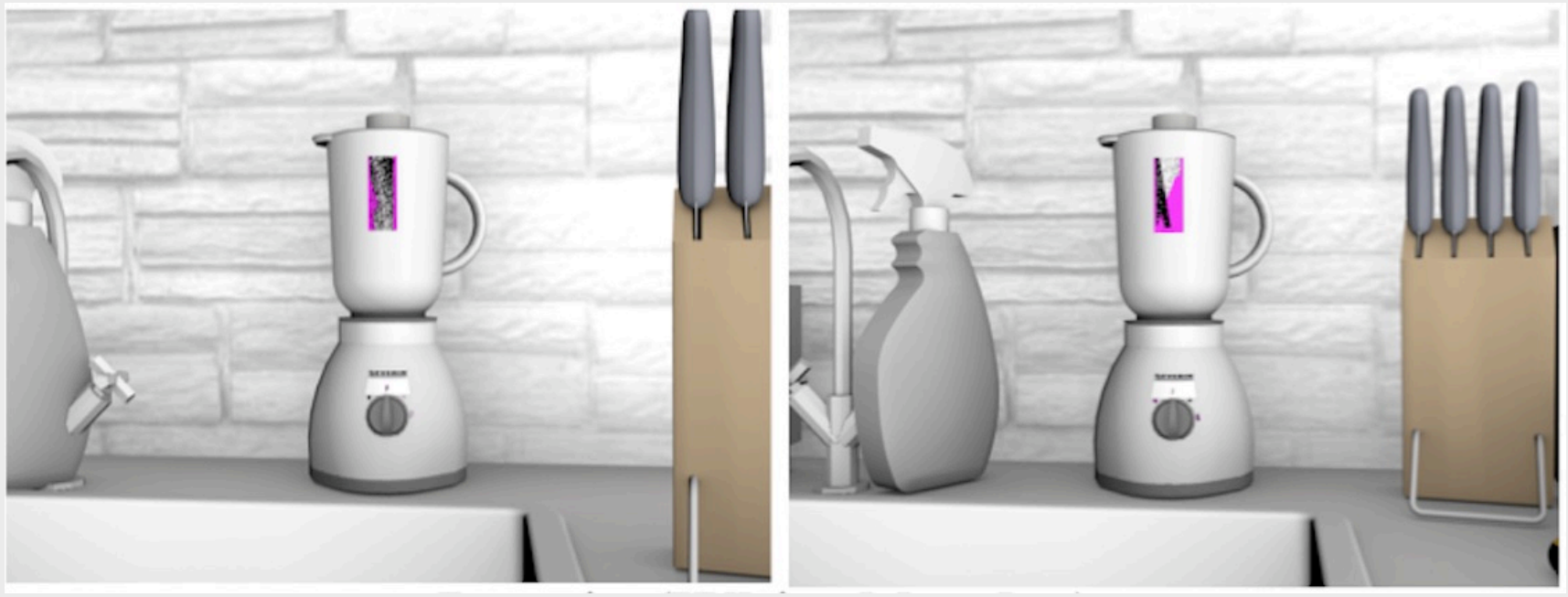
Term 5: Object constraints within the 3D scene

- Collision relationships
- Support relationships
- Semantic constraints

Term 5: Object constraints within the 3D scene

- Collision relationships
- Support relationships
- Semantic constraints

Term 6: Regularization



Overview

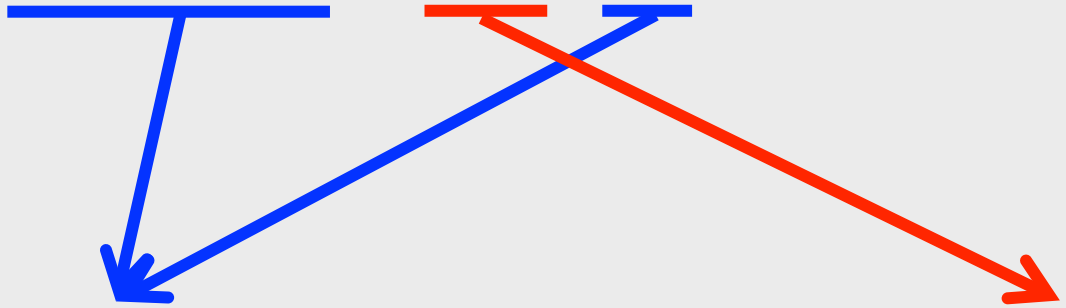
- Composition rules and constraints

→ • Optimization

- Applications

Energy function

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



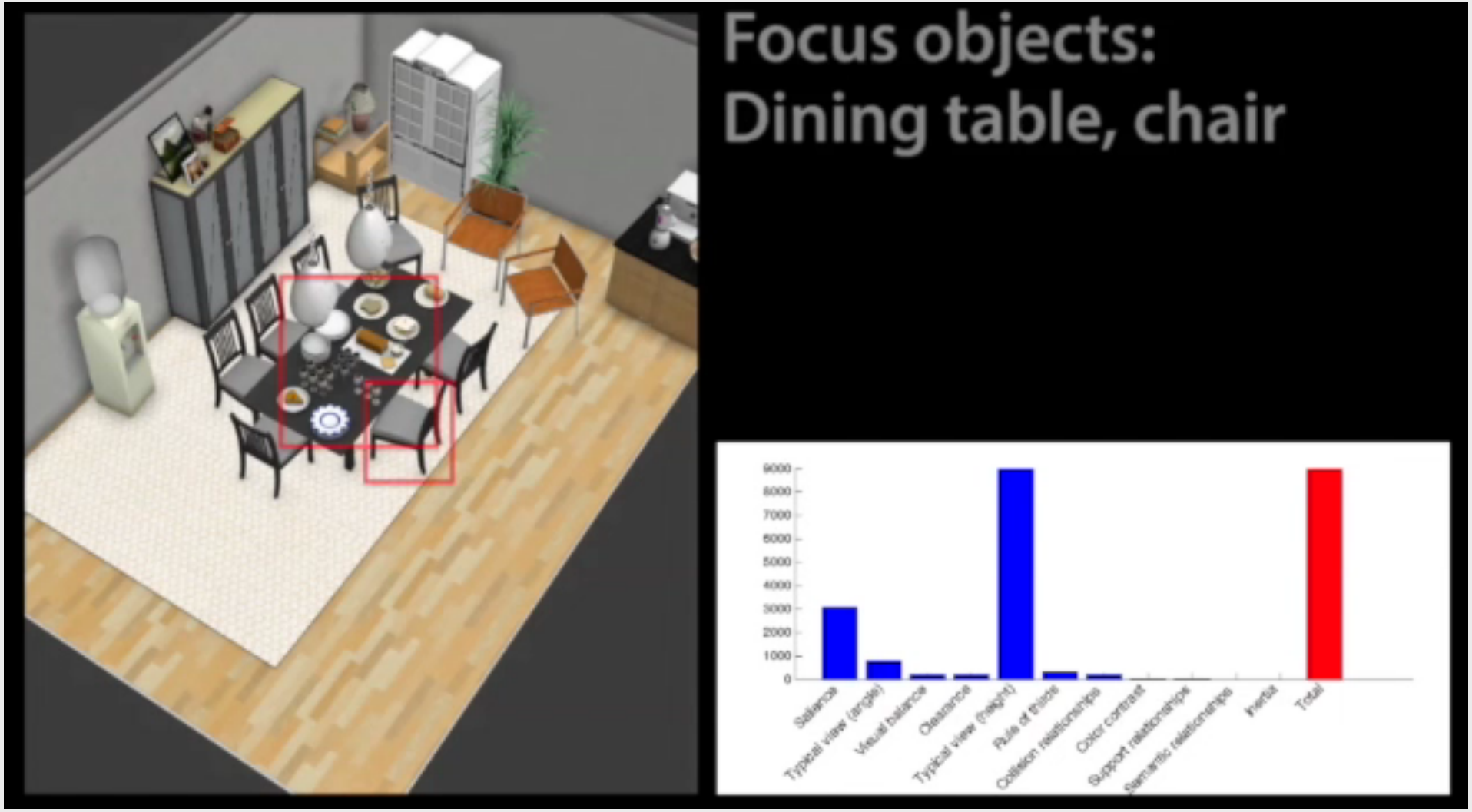
Continuous variables

Discrete variables

Optimization

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

Example 1


Focus objects:
Dining table, chair
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
Saliance
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

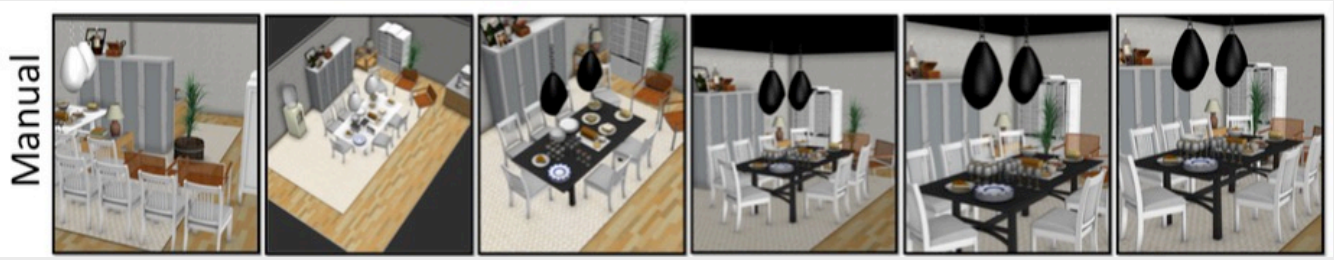
User study

Application 1: Refining rough compositions

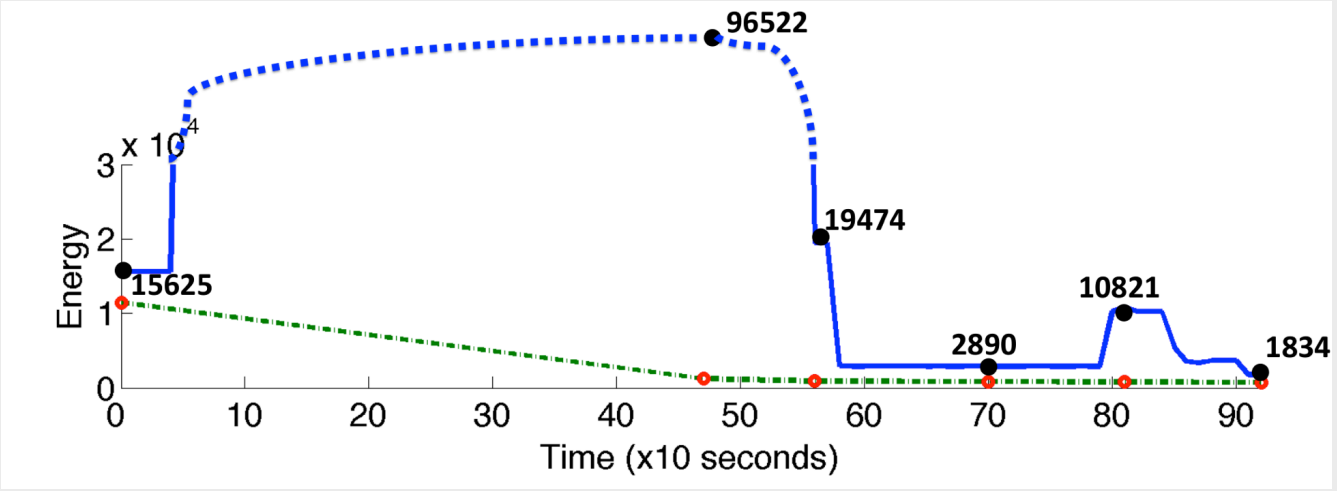
User study

Reference

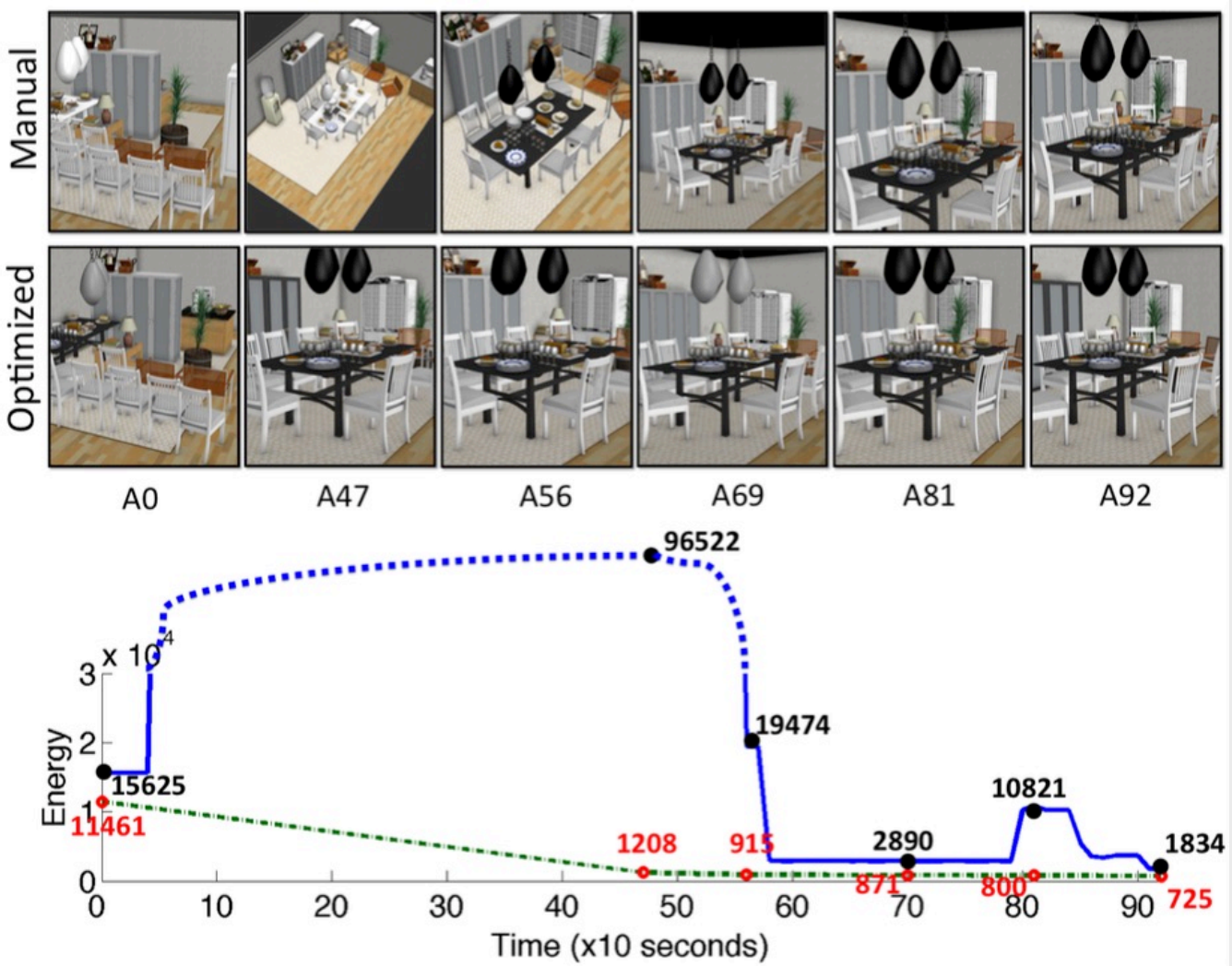
Application 1: Refining rough compositions



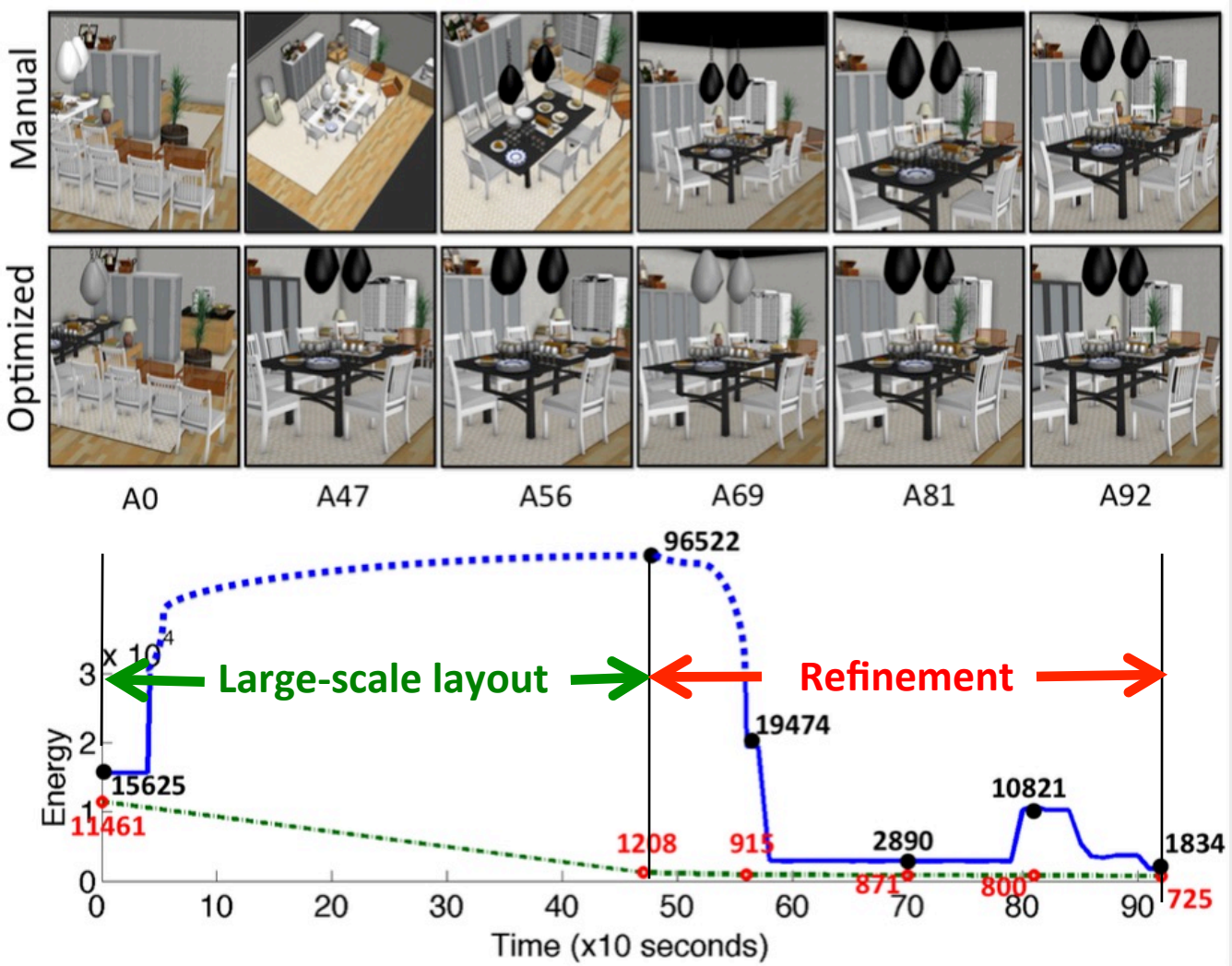
Application 1: Refining rough compositions



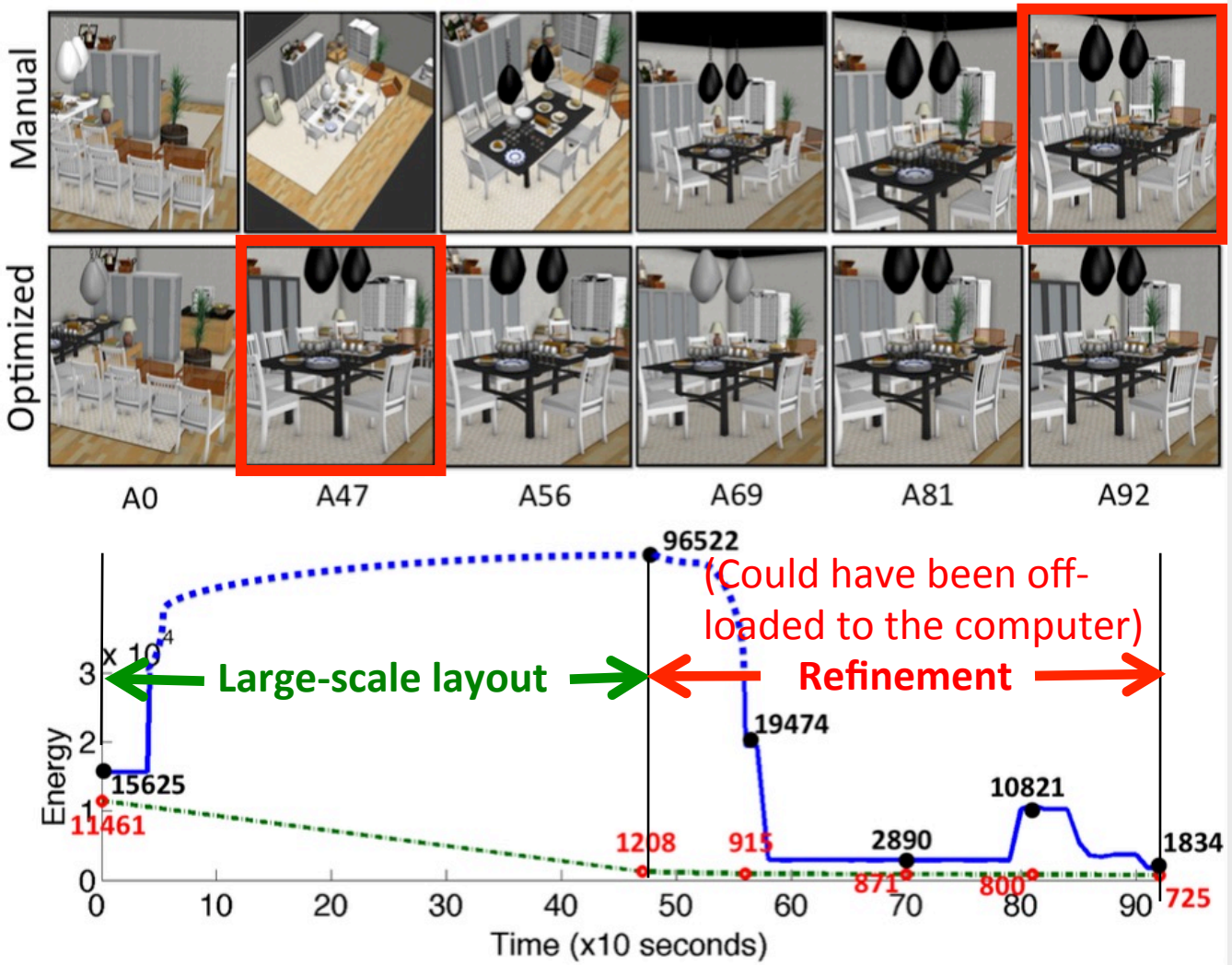
Application 1: Refining rough compositions



Application 1: Refining rough compositions



Application 1: Refining rough compositions



Application 2: Retargeting for different aspect ratios

?

Application 2: Retargeting for different aspect ratios

Input (4:3)

Camera-only

Application 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*

Application 4: Text-incorporated composition

Application 4: Text-incorporated composition

Input

Application 4: Text-incorporated composition

Input

Extra terms for overlaid text

- Contrast term

Application 4: Text-incorporated composition

Input

Extra terms for overlaid text

- Contrast term
- Variance term

Application 4: Text-incorporated composition

Input

Camera only

Application 4: Text-incorporated composition

Input

Camera only

Our result

Application 5: Generating detail images from an overview

Application 5: Generating detail images from an overview

Application 5: Generating detail images from an overview

Application 5: Generating detail images from an overview

(a) *Overview*

Application 5: Generating detail images from an overview

(a) *Overview*

(b) *Speaker*

Application 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

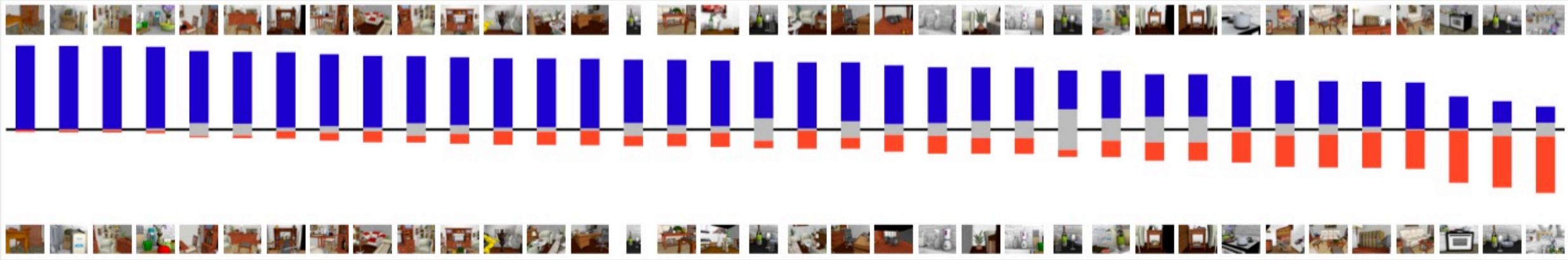
Study

Living room

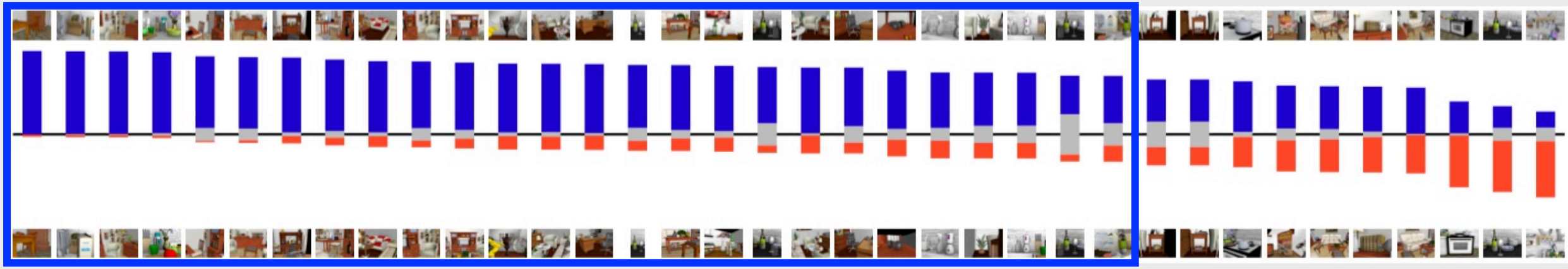
Expert Study Results

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



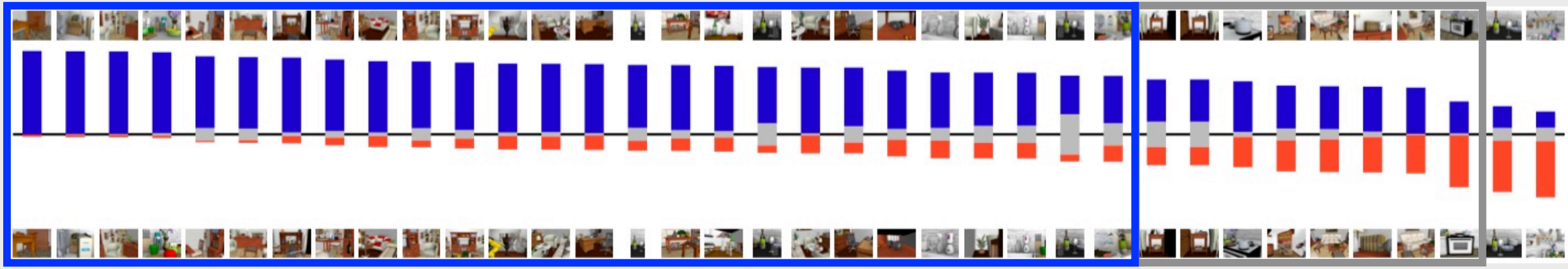
Amazon Mechanical Turk Study



If null hypothesis is there is no preference,

- Our method is preferred in 26/36 cases.

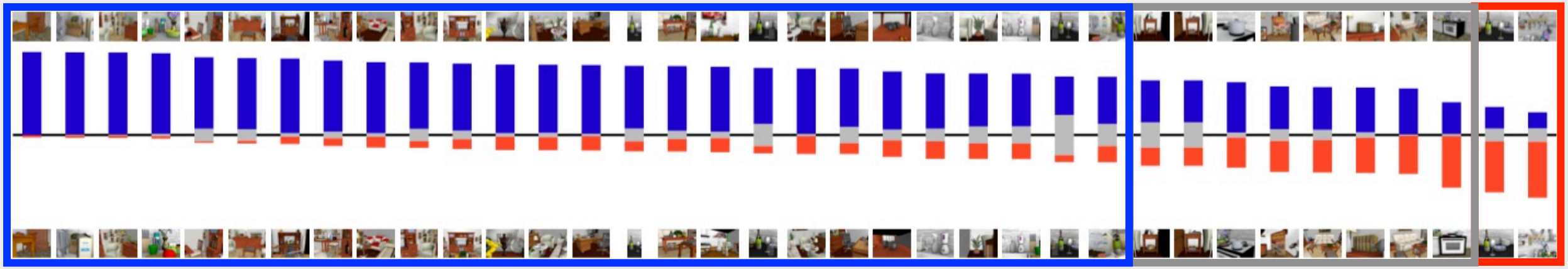
Amazon Mechanical Turk Study



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

- Moving objects and changing materials significantly improves the quality of compositions.
- Our optimization framework benefits a variety of applications.

Limitation and future work

- Interactive scene optimization
- Global illumination
- Additional composition rules

Limitation and future work

- Interactive scene optimization
- Global illumination
- Additional composition rules

Limitation and future work

- Interactive scene optimization
- Global illumination
- Additional composition rules

Limitation and future work

- Interactive scene optimization
- Global illumination
- Additional composition rules

Symmetry

Vanishing points

Acknowledgement

Interviews

- Martin Enthed and Helen Crowther

Catalogue images

- IKEA

Funding

- NSF, Intel, Adobe

Thank you!

Take-away message

- Moving objects and changing materials significantly improves the quality of compositions.
- Our optimization framework benefits a variety of applications.

(a) *Overview* (b) *Speaker* (c) *Shelf*