Composition-Aware Scene Optimization for Product Images

Tianqiang Liu\textsuperscript{1} Jim McCann\textsuperscript{2,3} Wilmot Li\textsuperscript{2} Thomas Funkhouser\textsuperscript{1}
\textsuperscript{1}Princeton University \textsuperscript{2}Adobe Research \textsuperscript{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
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’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
- \( \theta_i \): 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
- \(\theta_i\): orientation of object \(i\)
- \(m_i\): material of object \(i\)
Key Idea

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

- \( x_i, y_i \) : position of object i on its supporting surface
- \( \theta_i \) : 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
- \( \theta_i \): 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
- \(\theta_i\) : 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

Eurographics 2015
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

(Could have been off-loaded to the computer)
Application 2: Retargeting for different aspect ratios
Application 2: Retargeting for different aspect ratios
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
Application 4: Text-incorporated composition

Extra terms for overlaid text
• Contrast term

Input
Application 4: Text-incorporated composition

Extra terms for overlaid text
- Contrast term
- Variance term
Application 4: Text-incorporated composition

Input

Camera only
Application 4: Text-incorporated composition

Retargeting for different text layouts. The artist provides a rough position for the text box and specifies the champagne bottle and the goblet as focus objects. Then our optimization adjusts object positions, view-point and text positions to increase contrast, reduce clutter and remove occlusion of the focus objects.
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

<table>
<thead>
<tr>
<th>ID</th>
<th>Ours</th>
<th>Camera Only</th>
<th>No preference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expert 1</td>
<td>22</td>
<td>12</td>
<td>2</td>
</tr>
<tr>
<td>Expert 2</td>
<td>17</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td>Expert 3</td>
<td>22</td>
<td>11</td>
<td>3</td>
</tr>
<tr>
<td>Expert 4</td>
<td>21</td>
<td>12</td>
<td>3</td>
</tr>
</tbody>
</table>
Amazon Mechanical Turk Study
If null hypothesis is there is no preference, 
• Our method is preferred in 26/36 cases.
If null hypothesis is there is no preference,
• Our method is preferred in 26/36 cases.
• No statistical significance in 8 cases.
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
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.