Data-driven Multi-level Segmentation of Image Editing Logs

Zipeng Liu
University of British Columbia

Leo Zhicheng Liu
Adobe Research

Tamara Munzner
University of British Columbia

Paper published at CHI 2020
Presented at Adobe @ CHI online event on April 28 2020
Complexity in professional creativity tool

**Commands**
- History
- Rounded Rectangle Tool
- Rounded Rectangle Tool
- Move
- Free Transform
- Free Transform
- Layer Visibility
- Layer Visibility
- New Color Fill Layer
- Brush Tool
- Brush Tool
- Brush Tool
- Brush Tool
- Brush Tool
- Brush Tool
- Brush Tool
- Brush Tool
- Brush Tool
- Brush Tool

**Pictures**
- Image 1
- Image 2
- Image 3

**Layers**
- Layer 1
- Layer 2
- Photo Filter 1
- Photo Filter 2
- Photo Filter 3
User log segmentation can help

Smart undo: chunk of multiple coherent actions

Create shape
Change shape size and position
Create mask for shape
Definitions of segmentation

**Session:** poster creation

**Events:** 

---
Definitions of segmentation

Session: poster creation

Events

Attributes

Definitions of segmentation

Session: poster creation

Low-level chunks

Events

Attributes

Definitions of segmentation

**Session:** poster creation

- **Low-level chunks**
  - Type text
  - Edit and move title
  - Set title style
  - Make subtitle

- **Attributes**
Definitions of segmentation

*Session*: poster creation

**High-level chunks**

- Make title content

**Low-level chunks**

- Type text
- Edit and move title
- Set title style
- Make subtitle

**Events**


**Attributes**

- Add text
Definitions of segmentation

Session: poster creation

High-level chunks

Add text

Low-level chunks

Type text
Edit and move title
Set title style
Make subtitle

Make title content

Events

Attributes

Related work

1. Limited to specific task
   • Portrait retouching [Chen et al. 2016]
   • Sketching [Zhao et al. 2015]
   • Poster creation
   • UI design ...

2. Failed to handle complex user behaviors
   • Polysemy
   • Errors and corrections

3. Ignored features specific to image editing
   • Layer
Contributions

• Multi-level segmentation model
  • Low level: for smart undo
  • High level: adjustable granularity

• Evidence for feature relevance
  • Layer
    • First to use
    • Relevant
  • Command and duration
    • Relevant
    • Aligned with previous work
  • Image content
    • No effect
    • Contrary to previous work
1. Instrument
Command
Timestamp
Image
Layers ...

2. Compute
Features
(event similarity)

3. Segment

4. Analyze

5. Inspect

(a) A retouching session (180 events)
(b) A poster session (378 events)

Feature

1. Instrument
Command
Timestamp
Image
Layers ...

2. Compute
Features
(event similarity)

3. Segment

4. Analyze

5. Inspect

(a) A retouching session (180 events)
(b) A poster session (378 events)

Feature

1. Instrument
Command
Timestamp
Image
Layers ...

2. Compute
Features
(event similarity)

3. Segment

4. Analyze

5. Inspect

(a) A retouching session (180 events)
(b) A poster session (378 events)

Feature

1. Instrument
Command
Timestamp
Image
Layers ...

2. Compute
Features
(event similarity)

3. Segment

4. Analyze

5. Inspect

(a) A retouching session (180 events)
(b) A poster session (378 events)

Feature

1. Instrument
Command
Timestamp
Image
Layers ...

2. Compute
Features
(event similarity)

3. Segment

4. Analyze

5. Inspect

(a) A retouching session (180 events)
(b) A poster session (378 events)

Feature

1. Instrument
Command
Timestamp
Image
Layers ...

2. Compute
Features
(event similarity)

3. Segment

4. Analyze

5. Inspect

(a) A retouching session (180 events)
(b) A poster session (378 events)

Feature

1. Instrument
Command
Timestamp
Image
Layers ...

2. Compute
Features
(event similarity)

3. Segment

4. Analyze

5. Inspect

(a) A retouching session (180 events)
(b) A poster session (378 events)

Feature
1. Instrument
   Command
   Timestamp
   Image
   Layers...

2. Compute
   Features
   (event similarity)

3. Segment

4. Analyze

5. Inspect
Data collection from PS experts in action
Scale

• 16 sessions from 13 PS experts
  • ~ 30 min / session
• 5.7k events
  • ~ 300 / session

Poster creation:

Portrait retouching:

Special effect creation:

Labeling

- Author manually segment
- Event attributes
- Think-aloud video / audio

### Photoshop Log Segmentation

<table>
<thead>
<tr>
<th>Thumbnail</th>
<th>Event</th>
<th>SequenceId</th>
<th>DiffScore</th>
<th>Overlap</th>
<th>docId</th>
<th>ElapsedTime</th>
<th>EventLayerId</th>
<th>ActiveLayerName</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Place Embedded Smart Object</td>
<td>9</td>
<td>0.00</td>
<td>0.00</td>
<td>0</td>
<td>13.448</td>
<td>8</td>
<td>9249d9fe8069693b3d0f0706f06f3</td>
</tr>
<tr>
<td></td>
<td>Move</td>
<td>11</td>
<td>0.92</td>
<td>0.93</td>
<td>0</td>
<td>4.289</td>
<td>9</td>
<td>9249d9fe8069693b3d0f0706f06f3</td>
</tr>
<tr>
<td></td>
<td>Add Layer Mask</td>
<td>12</td>
<td>0.00</td>
<td>0.00</td>
<td>0</td>
<td>6.172</td>
<td>9</td>
<td>9249d9fe8069693b3d0f0706f06f3</td>
</tr>
<tr>
<td></td>
<td>Quick Selection</td>
<td>13</td>
<td>0.00</td>
<td>0.00</td>
<td>0</td>
<td>42.912</td>
<td>9</td>
<td>9249d9fe8069693b3d0f0706f06f3</td>
</tr>
<tr>
<td></td>
<td>Quick Selection</td>
<td>14</td>
<td>0.00</td>
<td>0.00</td>
<td>0</td>
<td>5.774</td>
<td>9</td>
<td>9249d9fe8069693b3d0f0706f06f3</td>
</tr>
<tr>
<td></td>
<td>Add Vector Mask</td>
<td>15</td>
<td>0.00</td>
<td>0.00</td>
<td>0</td>
<td>1.973</td>
<td>9</td>
<td>9249d9fe8069693b3d0f0706f06f3</td>
</tr>
<tr>
<td></td>
<td>Deselect</td>
<td>16</td>
<td>0.00</td>
<td>0.00</td>
<td>0</td>
<td>5.121</td>
<td>9</td>
<td>9249d9fe8069693b3d0f0706f06f3</td>
</tr>
<tr>
<td></td>
<td>Delete Layer Mask</td>
<td>17</td>
<td>0.00</td>
<td>0.00</td>
<td>0</td>
<td>7.665</td>
<td>9</td>
<td>9249d9fe8069693b3d0f0706f06f3</td>
</tr>
<tr>
<td></td>
<td>Delete Vector Mask</td>
<td>18</td>
<td>0.00</td>
<td>0.00</td>
<td>0</td>
<td>1.848</td>
<td>9</td>
<td>9249d9fe8069693b3d0f0706f06f3</td>
</tr>
</tbody>
</table>

1. Instrument
2. Compute
3. Segment
4. Analyze
5. Inspect
1. Instrument
   Command
   Timestamp
   Image
   Layers ...

2. Compute
   Features
   (event similarity)

3. Segment

4. Analyze

5. Inspect

Feature = event similarity (event A, event B)

Larger similarity $\Rightarrow$ more likely same chunk
1. Command similarity: NLP

Large database of command logs (100 million)

word2vec

Semantic vector space (command space)

Closer in vector space → larger similarity
2. Layer similarity: rule-based

Duplicate layers

Stronger layer relationships $\rightarrow$ larger similarity

Layer in the same group

Adjustment layer
3. Image-based similarity

- Larger image diff $\rightarrow$ smaller similarity
- Larger overlap $\rightarrow$ larger similarity
4. Duration

Larger duration $\rightarrow$ smaller similarity
1. Instrument
- Command
- Timestamp
- Image
- Layers...

2. Compute
- Features
  (event similarity)

3. Segment

4. Analyze

5. Inspect
Algorithm: two stage approach

Session: poster creation

Stage 2
High-level chunks
- Position images
- Tune colors
- Make lighting around characters
- Add text
- Add more light to characters
- Check results
- Make title content
- Apply color to title
- 2nd trial using clipping mask

Stage 1
Low-level chunks
- Type text
- Edit and move title
- Set title style
- Make subtitle
- 1st trial
- Make a mask layer
- Adjust opacity

Events
- 1. Instrument
- 2. Compute
- 3. Segment
- 4. Analyze
- 5. Inspect

Attributes
Low level: binary classification

• Problem:
  • boundary (start of a chunk) ↔ non-boundary

• Data:
  • 5.7k events
    • Features: similarities between current and previous events
    • Manual segmentation as ground truth
  • Partition: train – validate – test

• SVM with linear kernel
Low level: binary classification

- Cost for smart undo:
  - Missed boundary (false negative) > over segmentation (false positive)
  - Favor recall over precision (use F2 metric)
High level: multi-tier thresholds

- Threshold $t = \text{granularity of segmentation}$
1. Instrument

Command
Timestamp
Image
Layers...

2. Compute

Features
(event similarity)

3. Segment

4. Analyze

5. Inspect

1. Instrument

2. Compute

3. Segment

4. Analyze

5. Inspect
Quantitative analysis

<table>
<thead>
<tr>
<th>Command similarity</th>
<th>Layer similarity</th>
<th>Duration</th>
<th>Working region overlap</th>
<th>Image diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

1. Instrument  
2. Compute  
3. Segment  
4. Analyze  
5. Inspect
## Quantitative analysis

<table>
<thead>
<tr>
<th>Relevance</th>
<th>Command similarity</th>
<th>Layer similarity</th>
<th>Duration</th>
<th>Working region overlap</th>
<th>Image diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear Coefficient</td>
<td>-2.57</td>
<td>-1.74</td>
<td>+1.55</td>
<td>-0.18</td>
<td>-0.07</td>
</tr>
</tbody>
</table>
## Qualitative analysis

<table>
<thead>
<tr>
<th></th>
<th>Command similarity</th>
<th>Layer similarity</th>
<th>Duration</th>
<th>Working region overlap</th>
<th>Image diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevance</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear coefficient</td>
<td>-2.57</td>
<td>-1.74</td>
<td>+1.55</td>
<td>-0.18</td>
<td>-0.07</td>
</tr>
</tbody>
</table>

**Distribution of non-boundary events**

**Distribution of boundary events**

**Top vs. down**

1. Instrument  
2. Compute  
3. Segment  
4. Analyze  
5. Inspect
## Qualitative analysis

<table>
<thead>
<tr>
<th>Command similarity</th>
<th>Layer similarity</th>
<th>Duration</th>
<th>Working region overlap</th>
<th>Image diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relevance</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linear coefficient</td>
<td>-2.57</td>
<td>-1.74</td>
<td>+1.55</td>
<td>-0.18</td>
</tr>
</tbody>
</table>

### Distribution of non-boundary events

- **h**uge diff.

### Distribution of boundary events

- **balanced**

### Top vs. down

1. Instrument
2. Compute
3. Segment
4. Analyze
5. Inspect
## Qualitative analysis

<table>
<thead>
<tr>
<th>Relevance</th>
<th>Command similarity</th>
<th>Layer similarity</th>
<th>Duration</th>
<th>Working region overlap</th>
<th>Image diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear coefficient</td>
<td>-2.57</td>
<td>-1.74</td>
<td>+1.55</td>
<td>-0.18</td>
<td>-0.07</td>
</tr>
</tbody>
</table>

**Distribution of non-boundary events**

**Distribution of boundary events**

**Top vs. down**

1. Instrument
2. Compute
3. Segment
4. Analyze
5. Inspect
## Qualitative analysis

<table>
<thead>
<tr>
<th></th>
<th>Command similarity</th>
<th>Layer similarity</th>
<th>Duration</th>
<th>Working region overlap</th>
<th>Image diff</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Relevance</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Linear coefficient</strong></td>
<td>-2.57</td>
<td>-1.74</td>
<td>+1.55</td>
<td>-0.18</td>
<td>-0.07</td>
</tr>
</tbody>
</table>

**Distribution of non-boundary events**

**Distribution of boundary events**

**Top vs. down**

- **Different distributions**
- **Similar distributions**

1. Instrument  
2. Compute  
3. Segment  
4. Analyze  
5. Inspect
## Feature relevance analysis

<table>
<thead>
<tr>
<th>Relevance category</th>
<th>Command similarity</th>
<th>Layer similarity</th>
<th>Duration</th>
<th>Working region overlap</th>
<th>Image diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Most important</td>
<td>-2.57</td>
<td>-1.74</td>
<td>+1.55</td>
<td>-0.18</td>
<td>-0.07</td>
</tr>
<tr>
<td>Important</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No effect</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Linear coefficient

- Distribution of non-boundary events
- Distribution of boundary events
- Top vs. down

### Notes
- Different distributions
- Similar distributions

### Commands
- Instrument
- Compute
- Segment
- Analyze
- Inspect
1. Instrument
   Command
   Timestamp
   Image
   Layers ...

2. Compute
   Features
   (event similarity)

3. Segment

4. Analyze

5. Inspect
Example: poster creation

- Position images
- Tune colors
- Make lighting around characters
- Add text
- Make watersplashes
- More lights
- Check results

**Positioning:**
- Human labeled chunks
- Predicted chunks

**Probes:**
1. Instrument
2. Compute
3. Segment
4. Analyze
5. Inspect
Take-away

• Multi-level segmentation model for image editing logs
  • Low level: **smart undo**
  • High level: more use cases
    • Tutorial generation
    • Visual summary
    • Design alternatives

• Evidence for feature relevance
  • **Layer**
    • First to use
    • Relevant
  • Command and duration
    • Relevant
    • Aligned with previous work
  • Image content
    • No effect
    • Contrary to previous work
Data-driven Multi-level Segmentation of Image Editing Logs.

Zipeng Liu, Leo Zhicheng Liu, Tamara Munzner.
Presented at Adobe @ CHI online event on April 28 2020
http://www.cs.ubc.ca/labs/imager/tr/2020/logseg/

Take-away:

- Multi-level segmentation model for image editing logs
  - Low level: smart undo
  - High level: more use cases

- Evidence for feature relevance
  - Layer: first to use; relevant
  - Command and duration: relevant
  - Image content: not relevant
2. Layer similarity: rule-based

Stronger layer relationships $\rightarrow$ larger similarity

<table>
<thead>
<tr>
<th>Relationship</th>
<th>Description</th>
<th>Similarity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Same layer</td>
<td>$A = B$</td>
<td>1.0</td>
</tr>
<tr>
<td>Duplicate layer</td>
<td>$A$ is a copy of $B$</td>
<td>0.8</td>
</tr>
<tr>
<td>Adjustment layer</td>
<td>$A$ is an adjustment layer of $B$</td>
<td>0.5</td>
</tr>
<tr>
<td>Grouped layer</td>
<td>$A$ and $B$ are located in the same layer group</td>
<td>$\leq 0.5$</td>
</tr>
<tr>
<td>Other diff. layer</td>
<td>none of the above</td>
<td>0.0</td>
</tr>
</tbody>
</table>
Example 1: portrait retouching

Open image
Forehead and right eye
Try healing brush
Fix right face and ear
Fix teeth and lips
Fix left background

Highest Probability of Errors:

- Forehead and right eye
- Fix right face and ear
- Fix left background

Try healing brush:

- Forehead and right eye
- Fix right face and ear
- Fix left background

Fix right face and ear:

- Forehead and right eye
- Fix right face and ear
- Fix left background

Fix teeth and lips:

- Forehead and right eye
- Fix right face and ear
- Fix left background

Fix left background:

- Forehead and right eye
- Fix right face and ear
- Fix left background
Reflections on real-world user behavior

- Mistakes
- Interleaving subtasks
- Trial-and-error experiments
- Fuzzy boundaries