Data-driven Multi-level Segmentation of Image Editing Logs

Zipeng Liu

University of British Columbia

Leo Zhicheng Liu

Adobe Research

Tamara Munzner

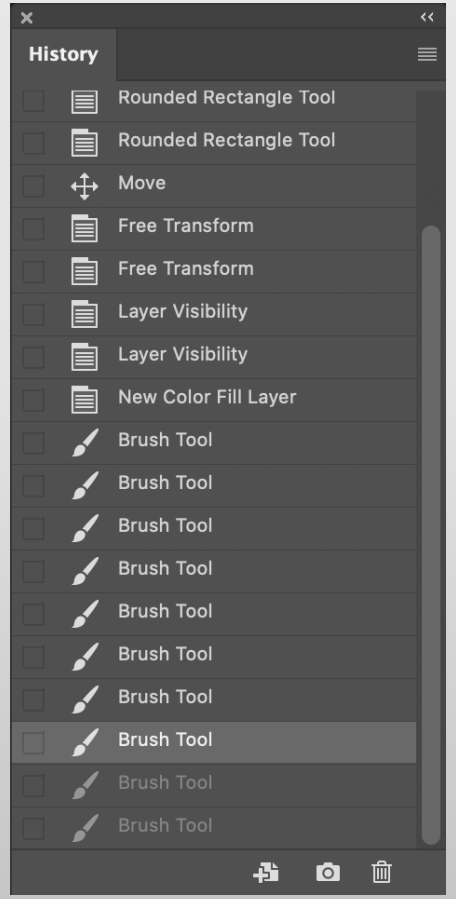
University of British Columbia

Paper published at CHI 2020

Presented at Adobe @ CHI online event on April 28 2020

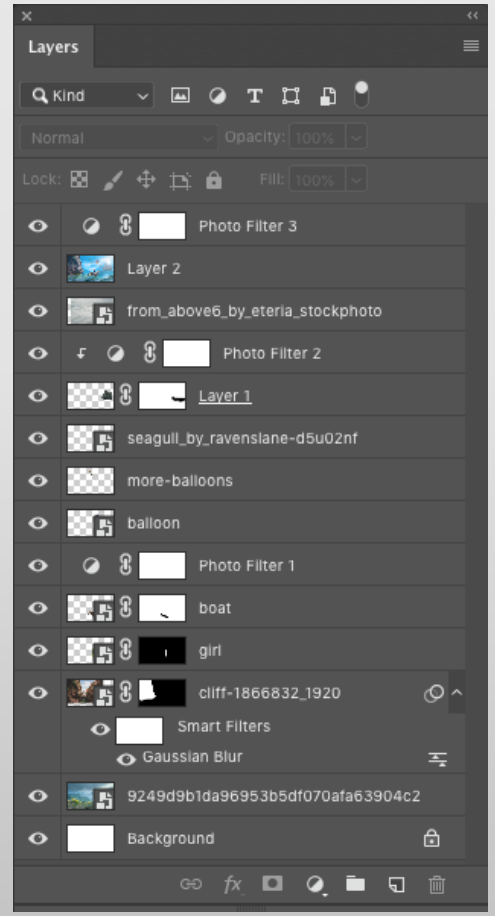
Complexity in professional creativity tool

Commands



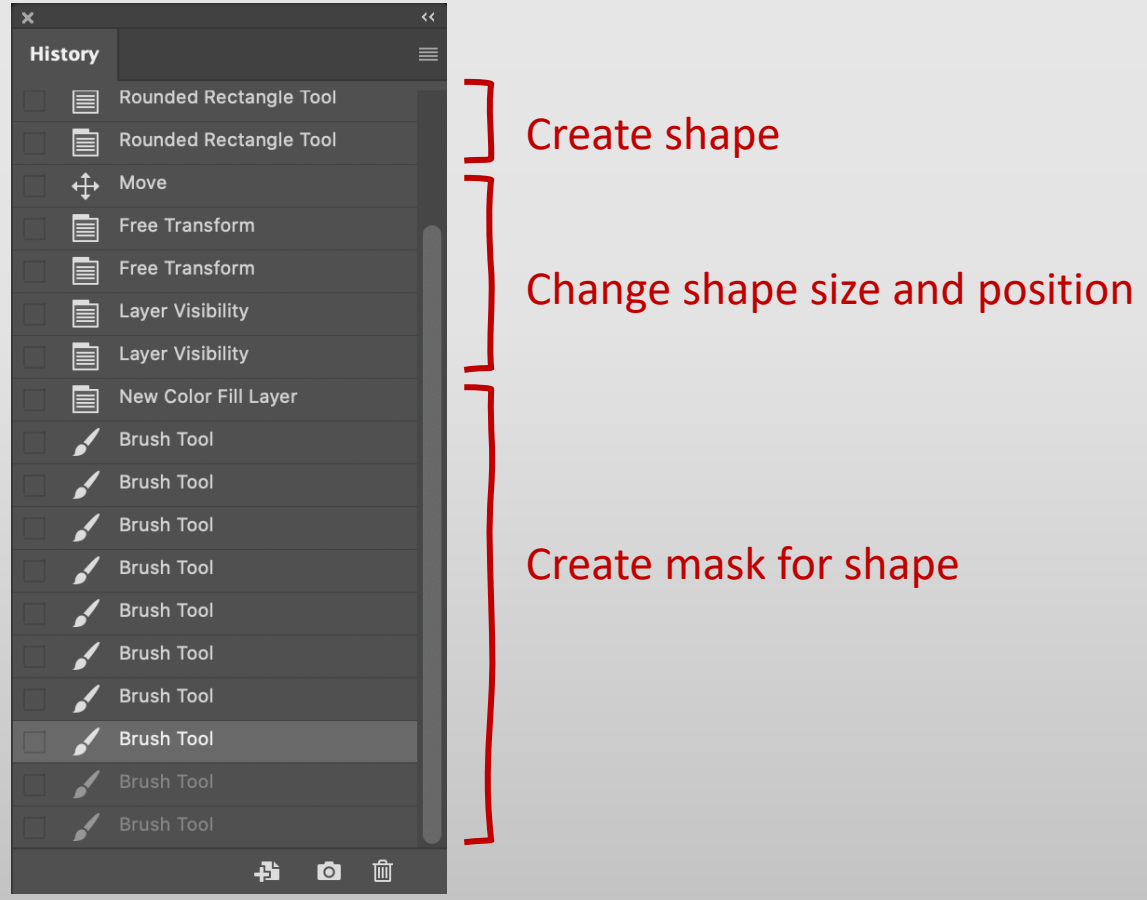
Pictures

Layers



User log segmentation can help

Smart undo: chunk of multiple coherent actions



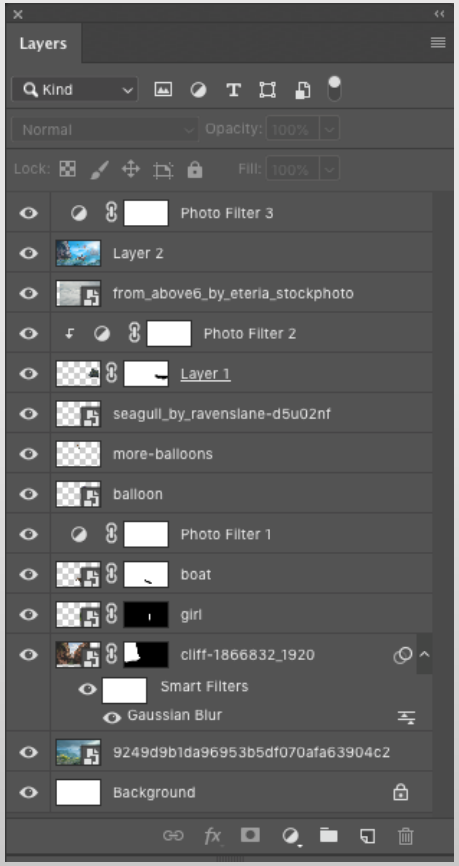
Definitions of segmentation

Session: poster creation

Events oooooooooooooooooooooooooooooooooooooooo

Definitions of segmentation

Session: poster creation

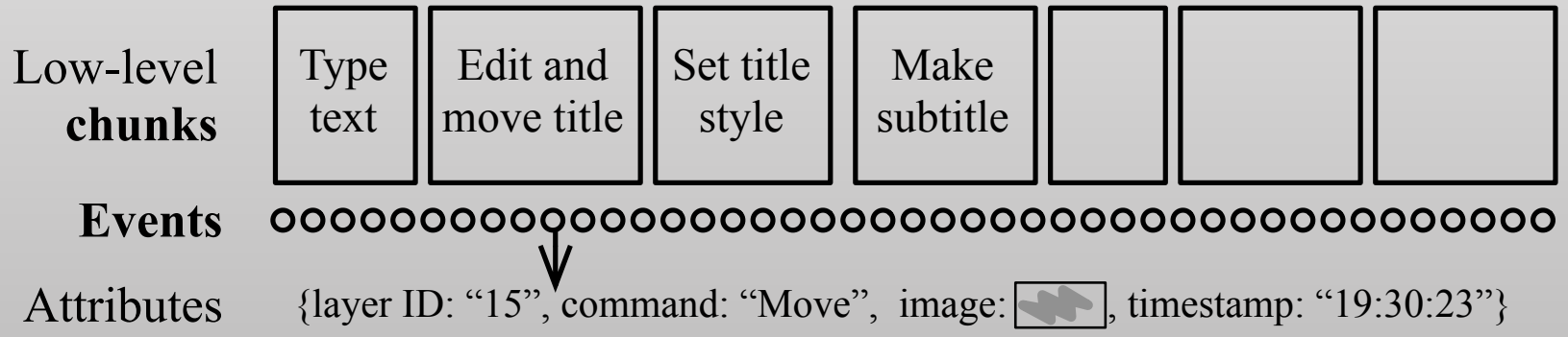


Events oooooooooooooooooooooooooooooooooooooo

Attributes {layer ID: "15", command: "Move", image:  , timestamp: "19:30:23"}

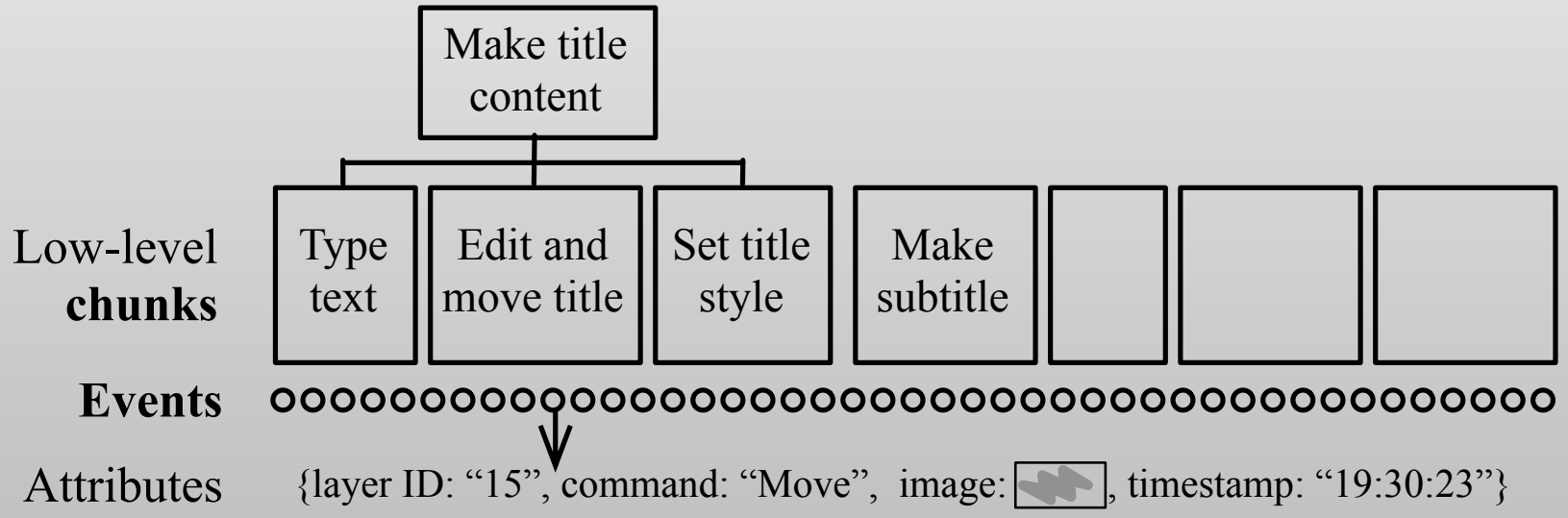
Definitions of segmentation

Session: poster creation

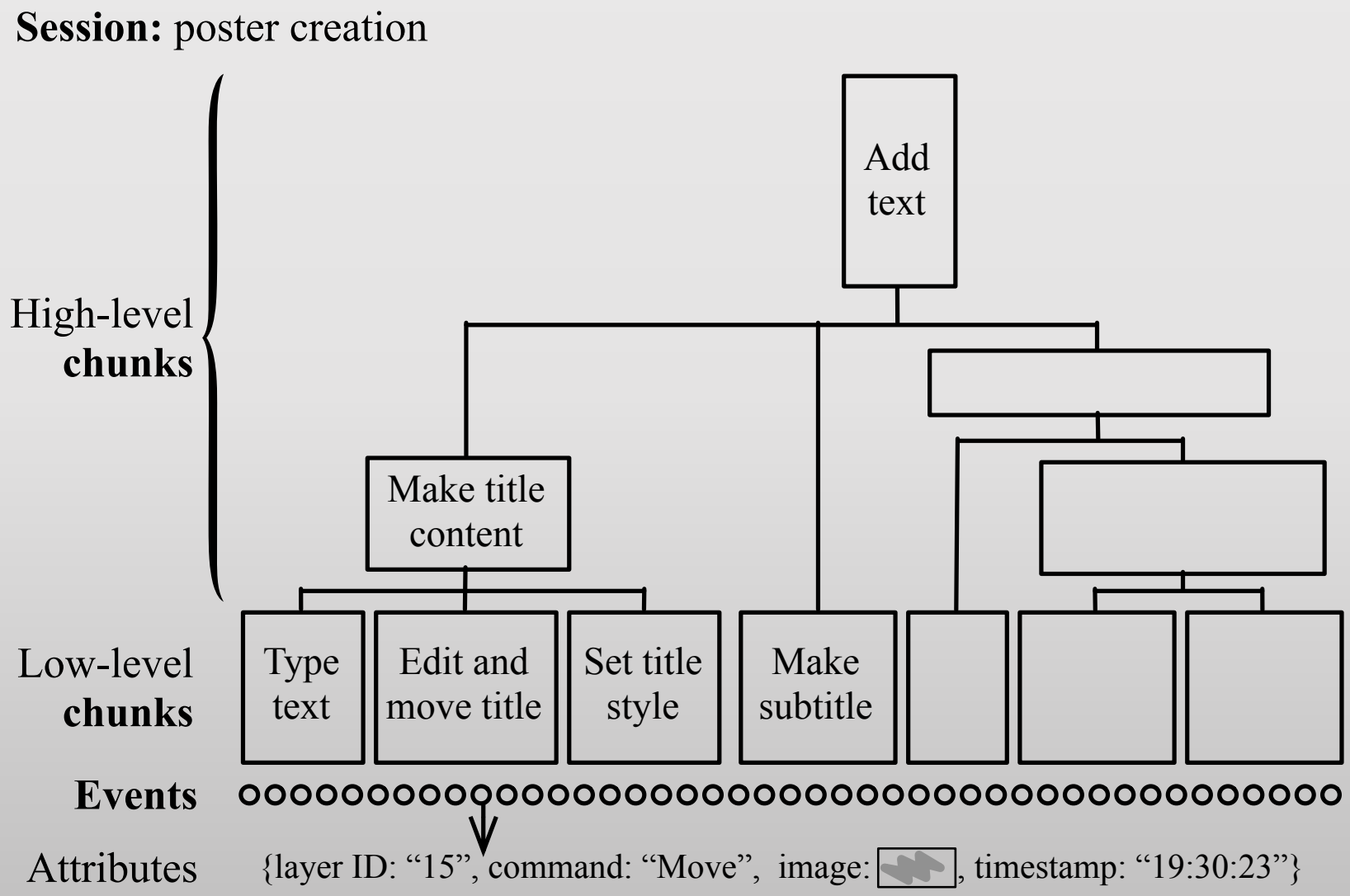


Definitions of segmentation

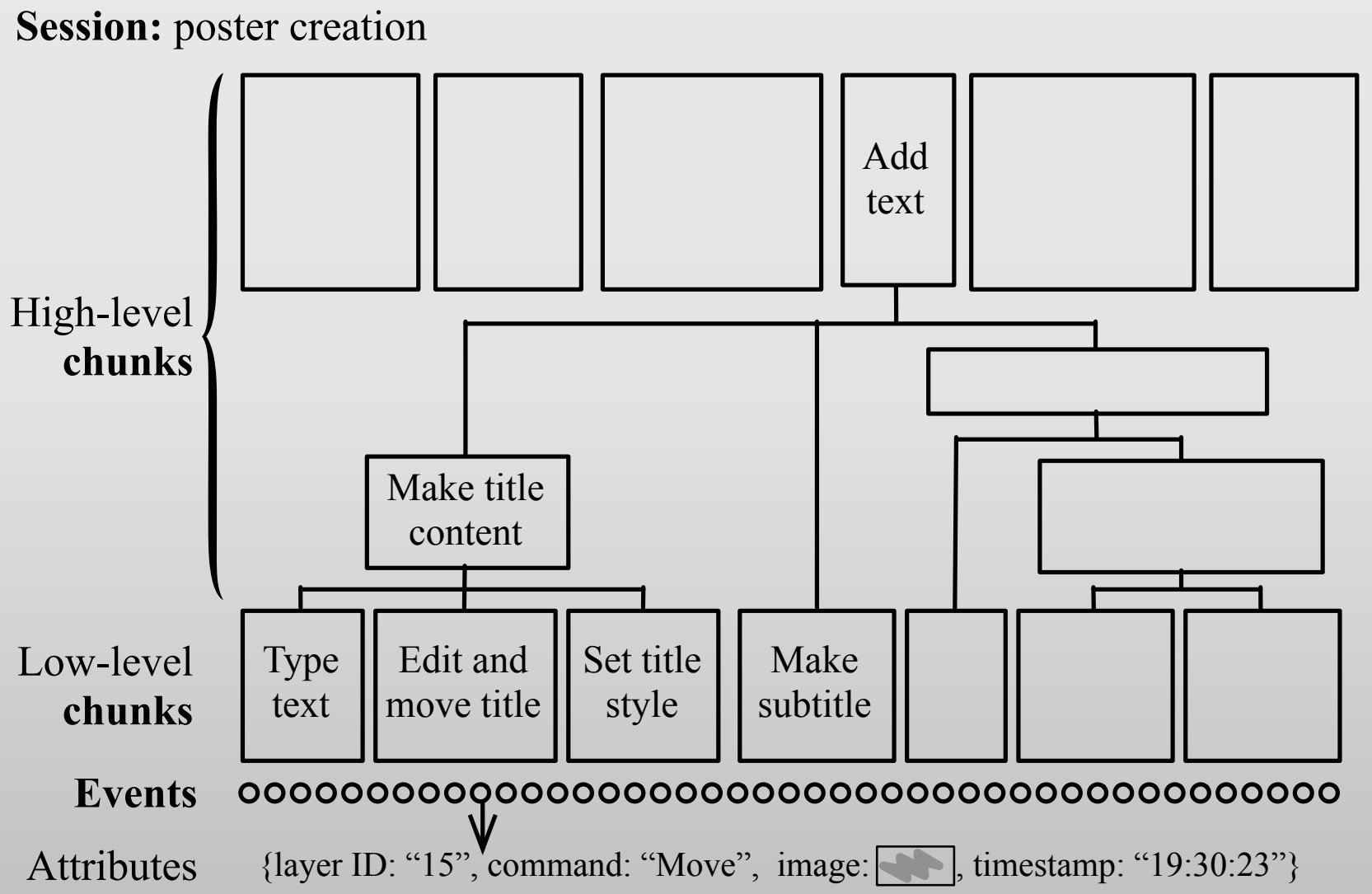
Session: poster creation



Definitions of segmentation



Definitions of segmentation



Related work

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

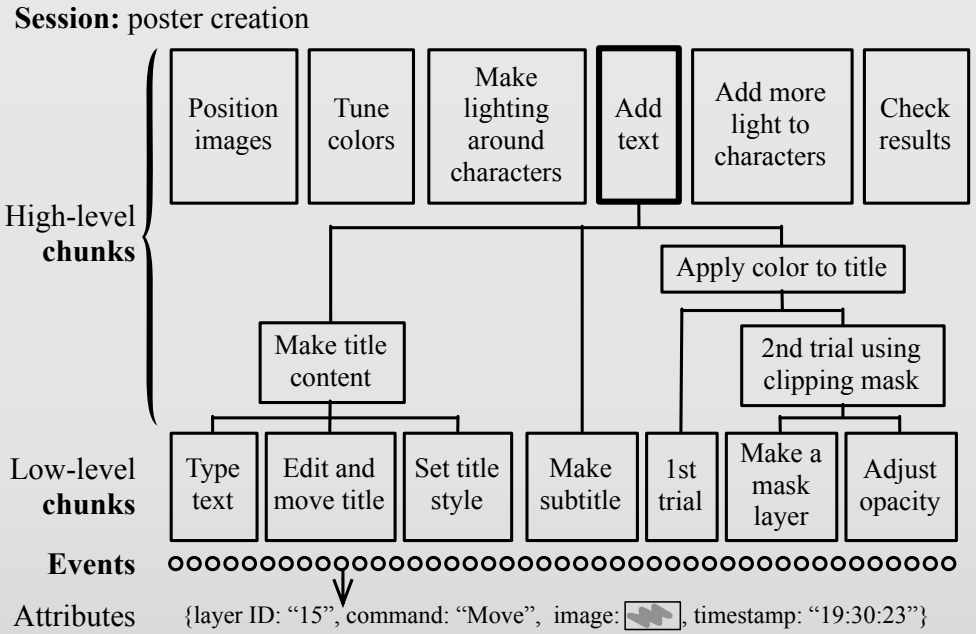
2. Failed to handle complex user behaviors
   - Polysemy
   - Errors and corrections

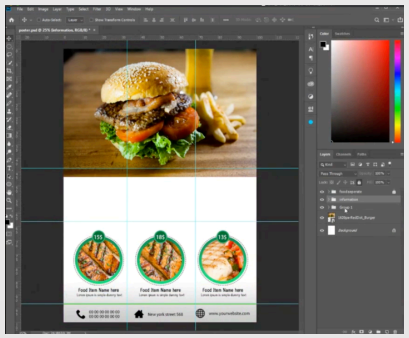
3. Ignored features specific to image editing
   - Layer

Contributions

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

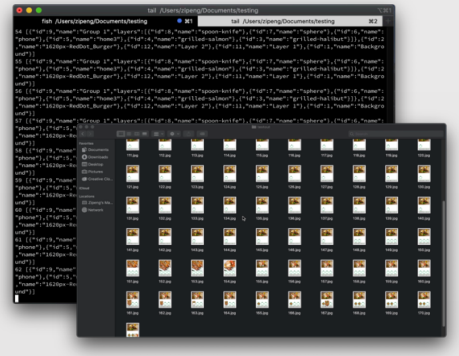
- Evidence for feature relevance
  - **Layer**
    - **First to use**
    - **Relevant**
  - Command and duration
    - Relevant
    - Aligned with previous work
  - Image content
    - No effect
    - Contrary to previous work





1. Instrument

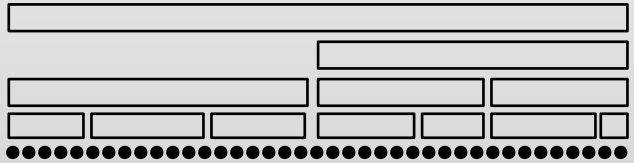
Command
Timestamp
Image
Layers …



2. Compute

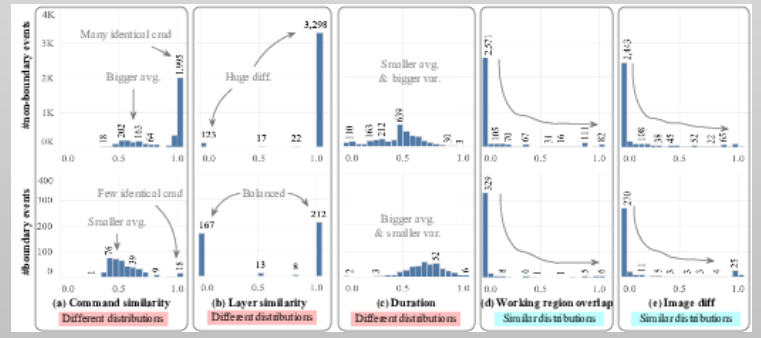
Features
(event similarity)

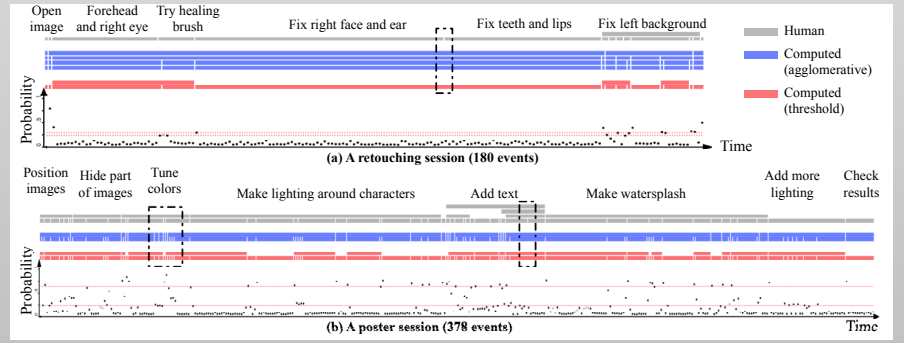
3. Segment

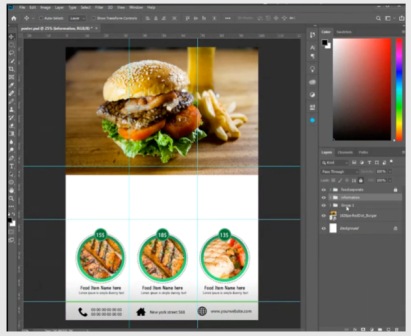


4. Analyze

5. Inspect

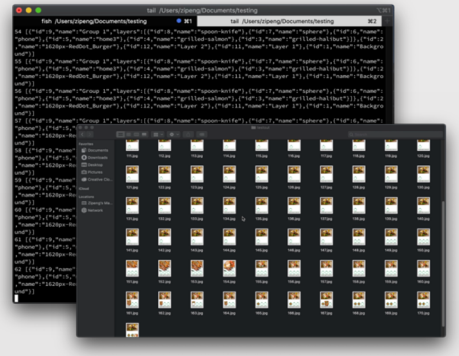






1. Instrument

Command
Timestamp
Image
Layers …



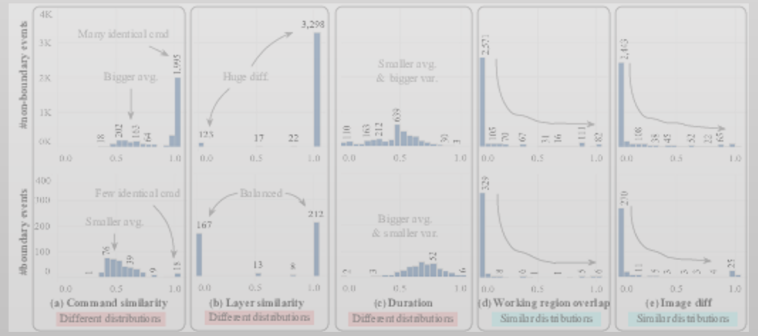
2. Compute

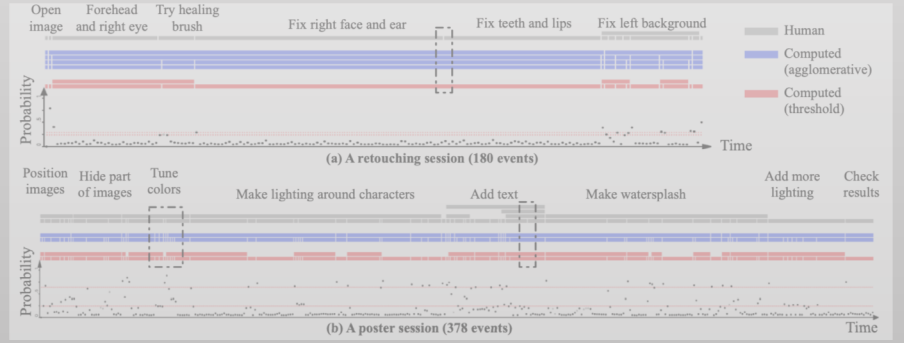
Features
(event similarity)

3. Segment

4. Analyze

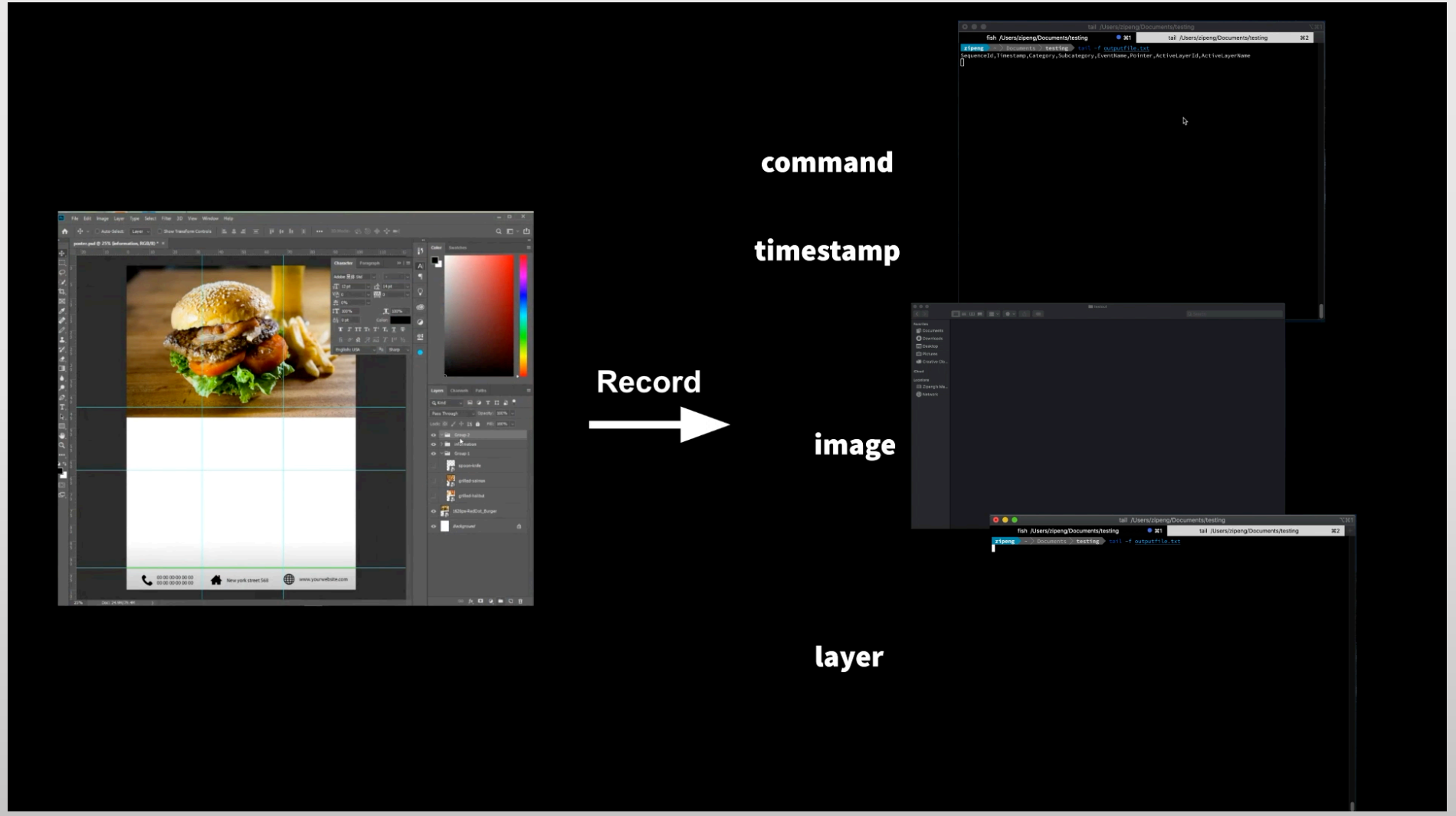
5. Inspect





1. Instrument | 2. Compute | 3. Segment | 4. Analyze | 5. Inspect

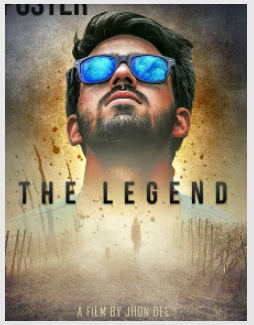
Data collection from PS experts in action



Scale

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

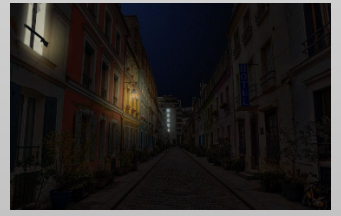
Poster creation:



Portrait retouching:

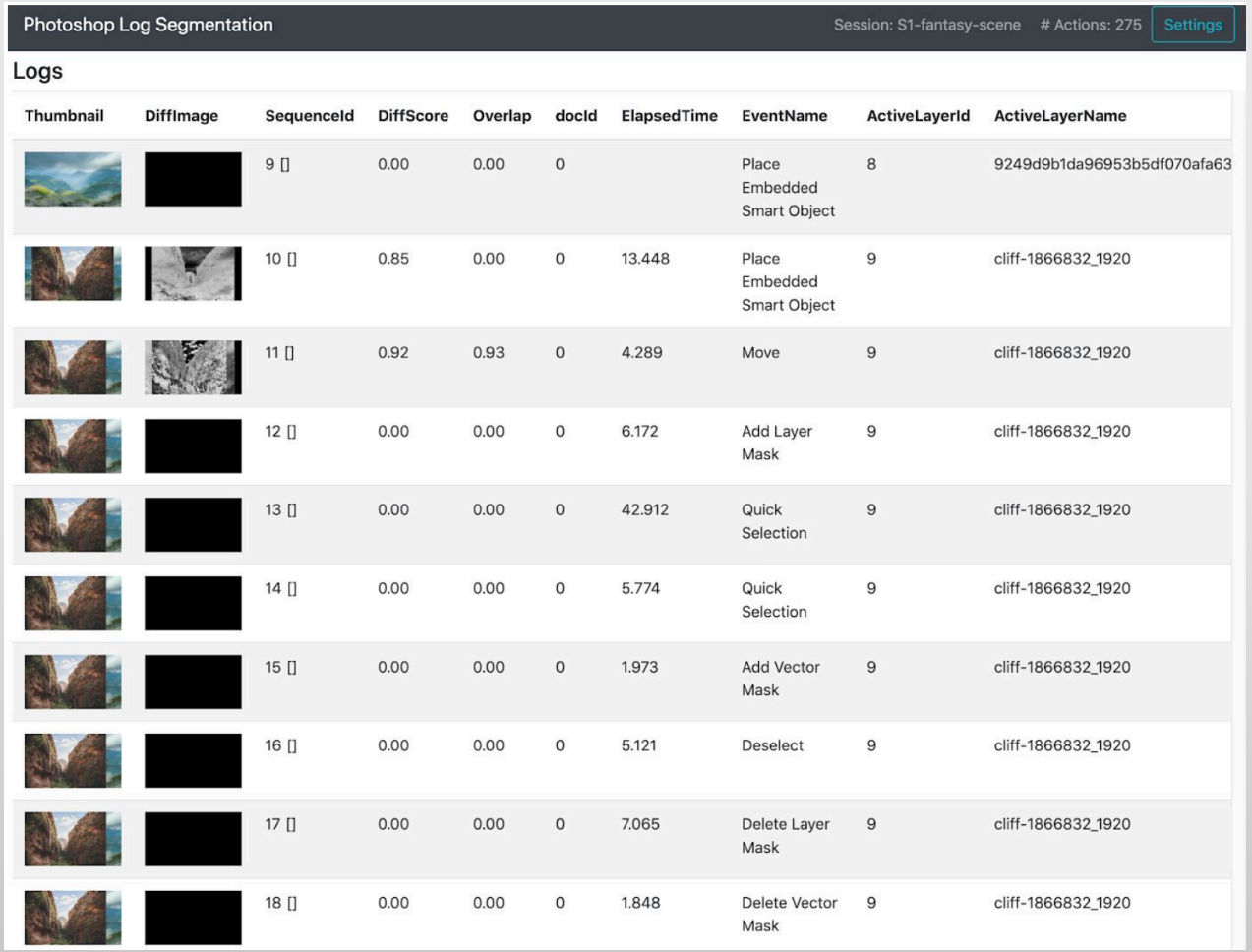


Special effect creation:

Labeling

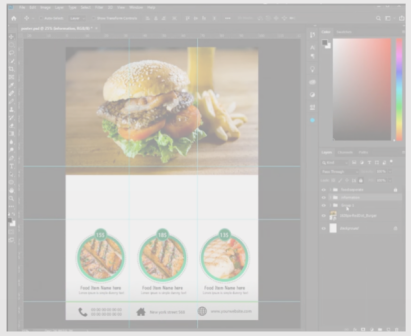
- Author manually segment
  - Event attributes
  - Think-aloud video / audio



Photoshop Log Segmentation — Session: S1-fantasy-scene — # Actions: 275 — Settings

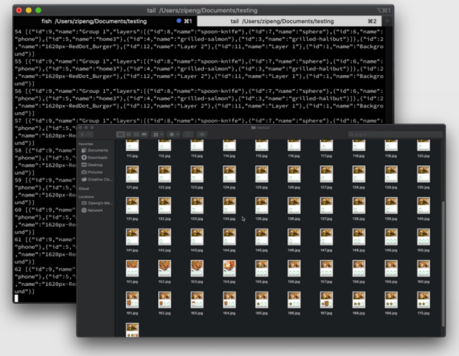
Logs

Thumbnail	DiffImage	SequenceId	DiffScore	Overlap	docId	ElapsedTime	EventName	ActiveLayerId	ActiveLayerName
		9 []	0.00	0.00	0		Place Embedded Smart Object	8	9249d9b1da96953b5df070afa63
		10 []	0.85	0.00	0	13.448	Place Embedded Smart Object	9	cliff-1866832_1920
		11 []	0.92	0.93	0	4.289	Move	9	cliff-1866832_1920
		12 []	0.00	0.00	0	6.172	Add Layer Mask	9	cliff-1866832_1920
		13 []	0.00	0.00	0	42.912	Quick Selection	9	cliff-1866832_1920
		14 []	0.00	0.00	0	5.774	Quick Selection	9	cliff-1866832_1920
		15 []	0.00	0.00	0	1.973	Add Vector Mask	9	cliff-1866832_1920
		16 []	0.00	0.00	0	5.121	Deselect	9	cliff-1866832_1920
		17 []	0.00	0.00	0	7.065	Delete Layer Mask	9	cliff-1866832_1920
		18 []	0.00	0.00	0	1.848	Delete Vector Mask	9	cliff-1866832_1920



1. Instrument

Command
Timestamp
Image
Layers …



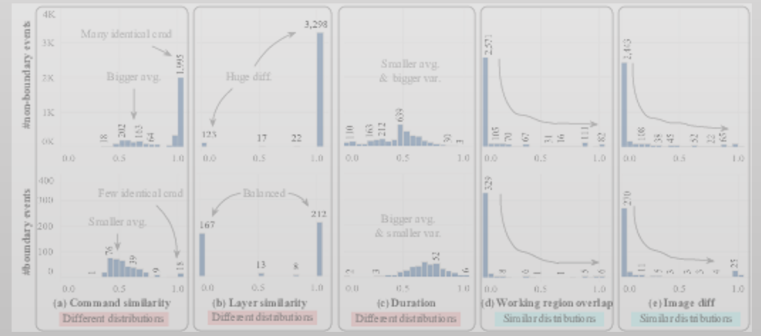
2. Compute

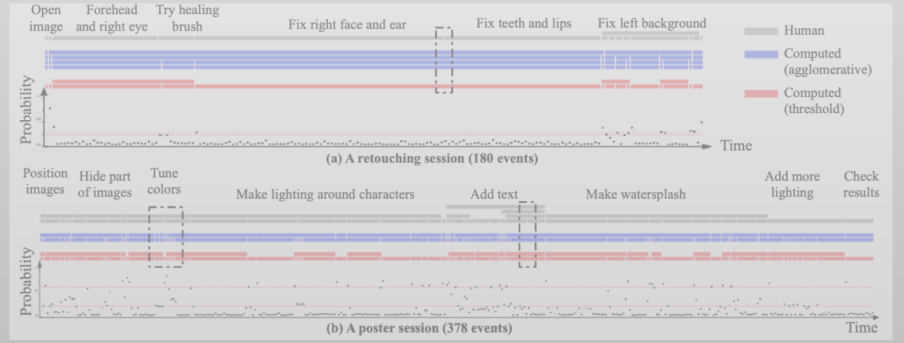
Features
(event similarity)

3. Segment

4. Analyze

5. Inspect





Feature = event similarity (event A, event B)

Larger similarity → more likely same chunk

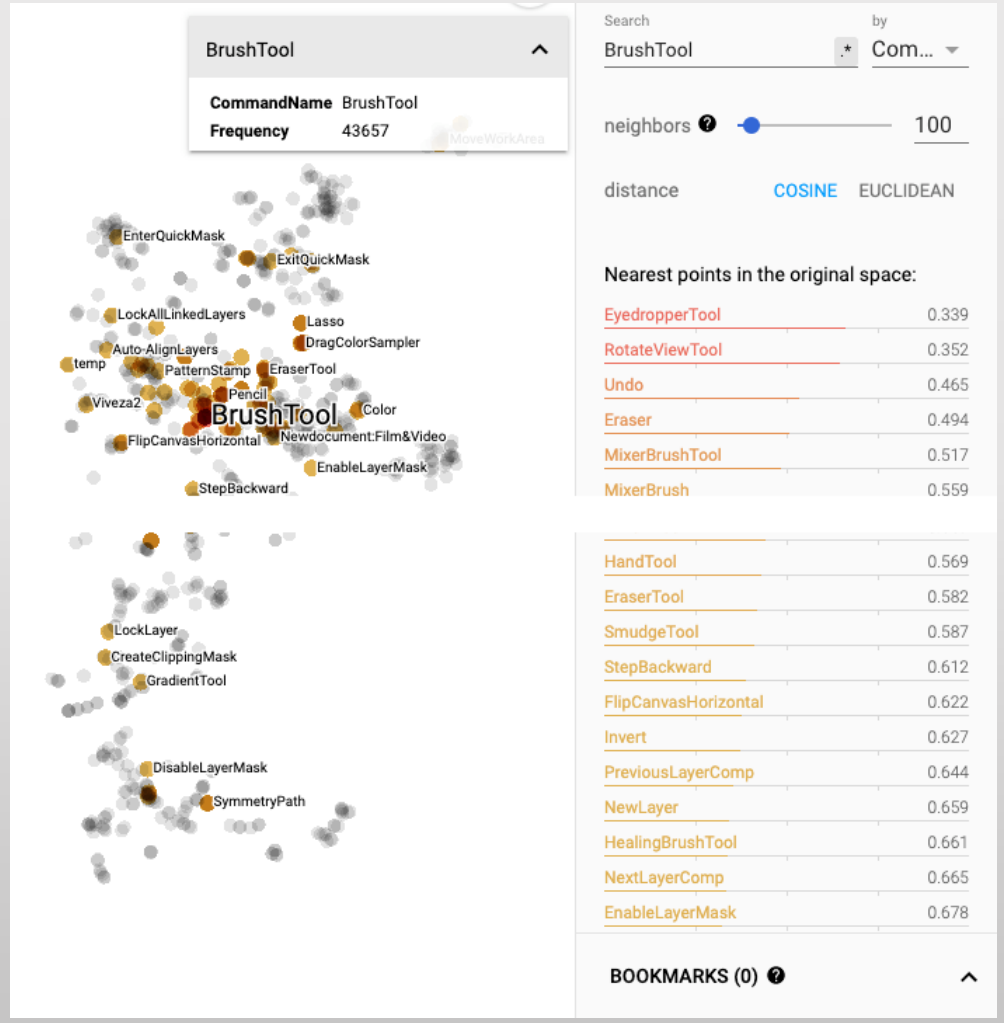
1. Command similarity: NLP

Large database of command logs (100 million)

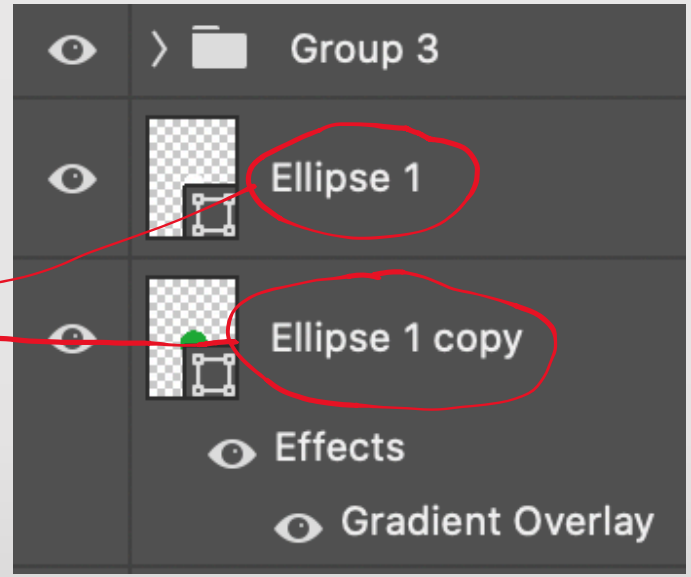
word2vec

Semantic vector space (command space)

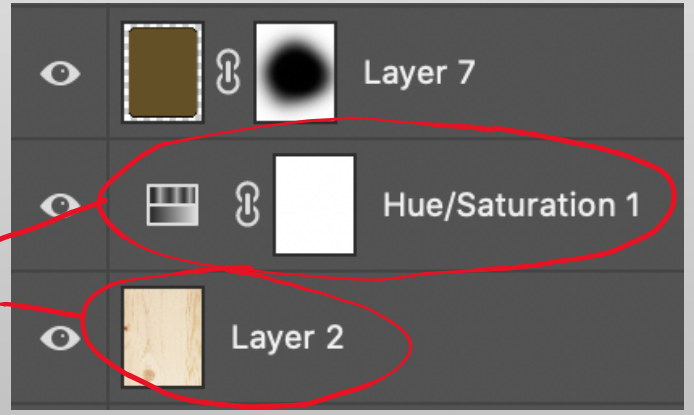
Closer in vector space → larger similarity



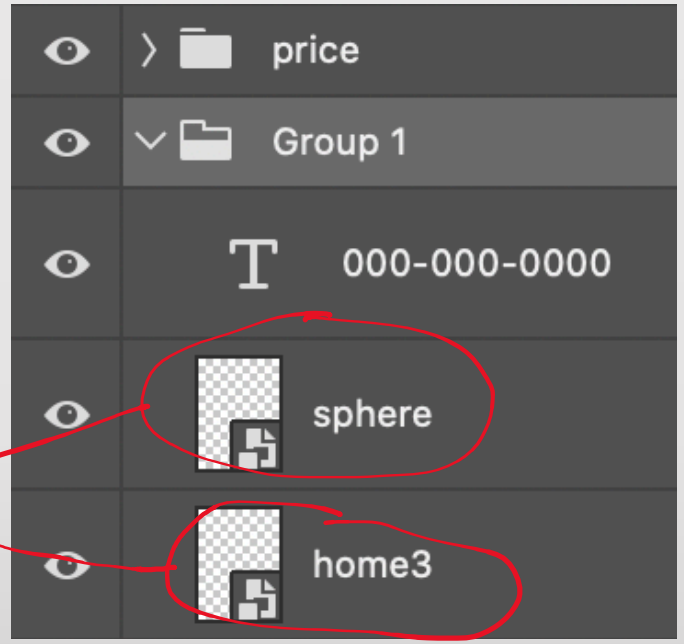
2. Layer similarity: rule-based



Duplicate layers



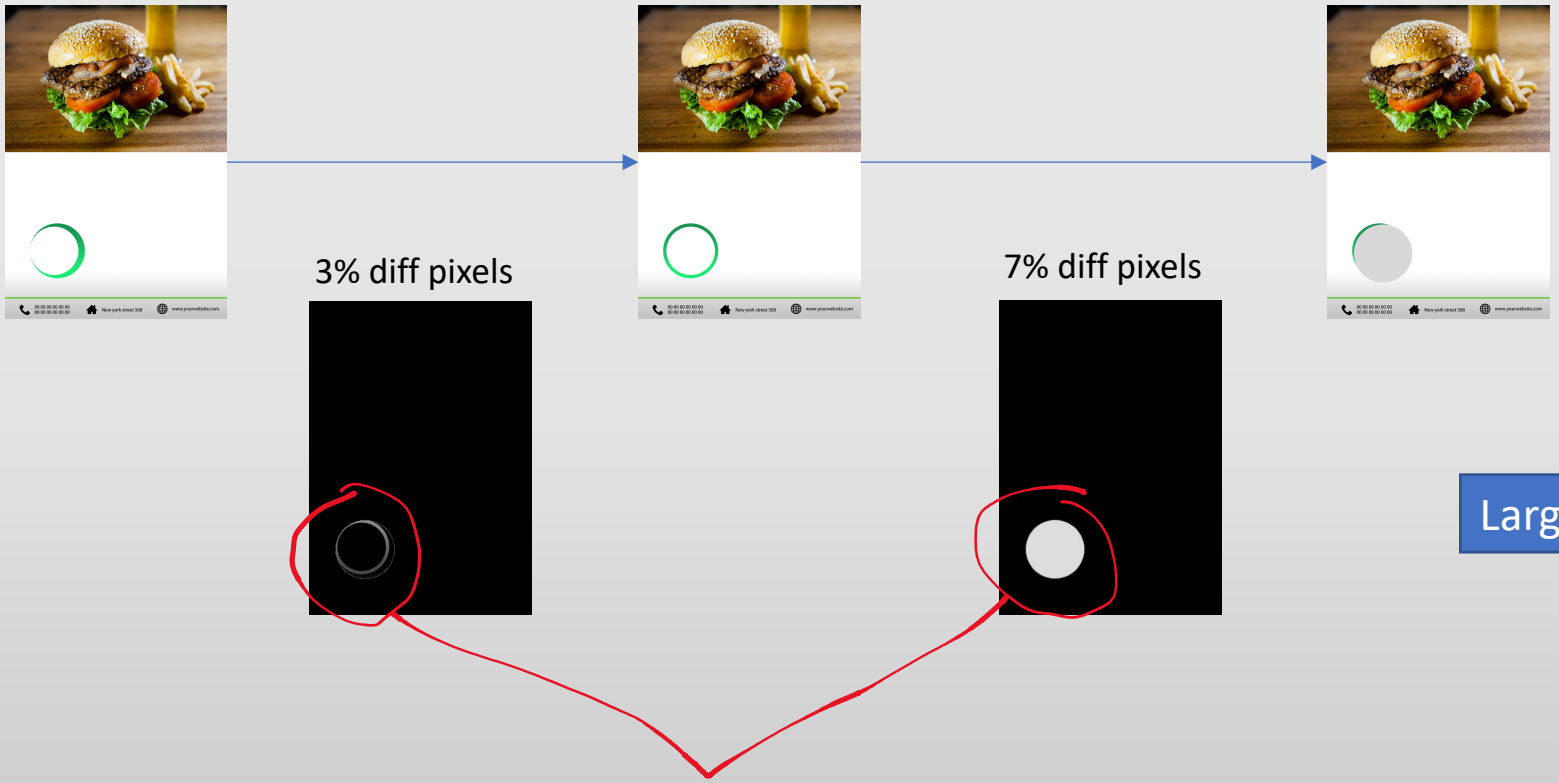
Adjustment layer



Layer in the same group

Stronger layer relationships → larger similarity

3. Image-based similarity

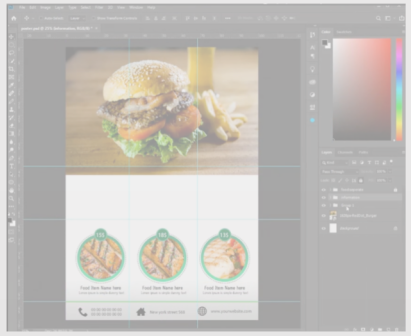


Larger image diff → smaller similarity

Larger overlap → larger similarity

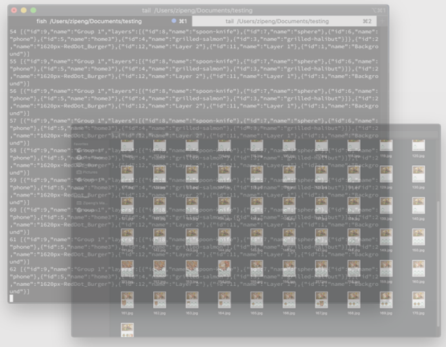
4. Duration

Larger duration → smaller similarity



1. Instrument

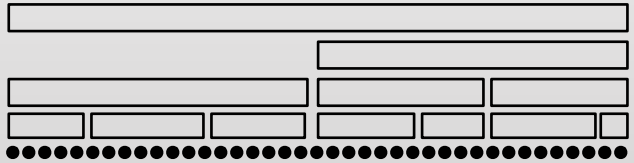
Command
Timestamp
Image
Layers …



2. Compute

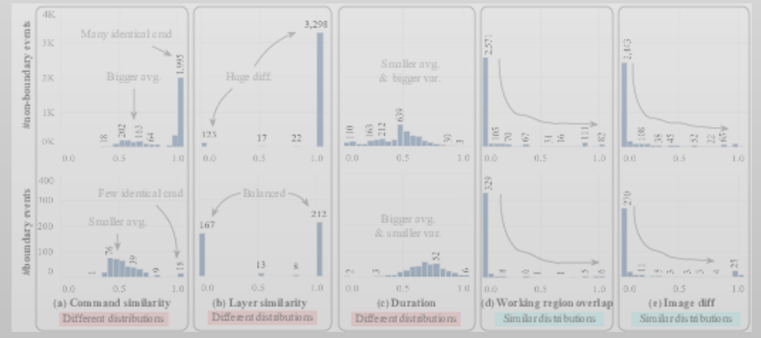
Features
(event similarity)

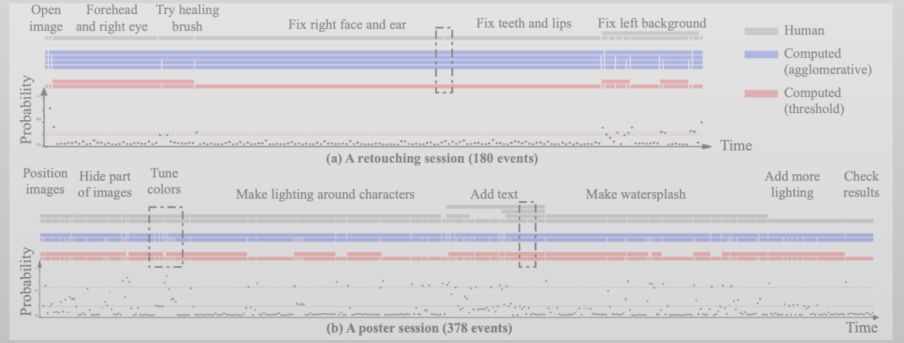
3. Segment



4. Analyze

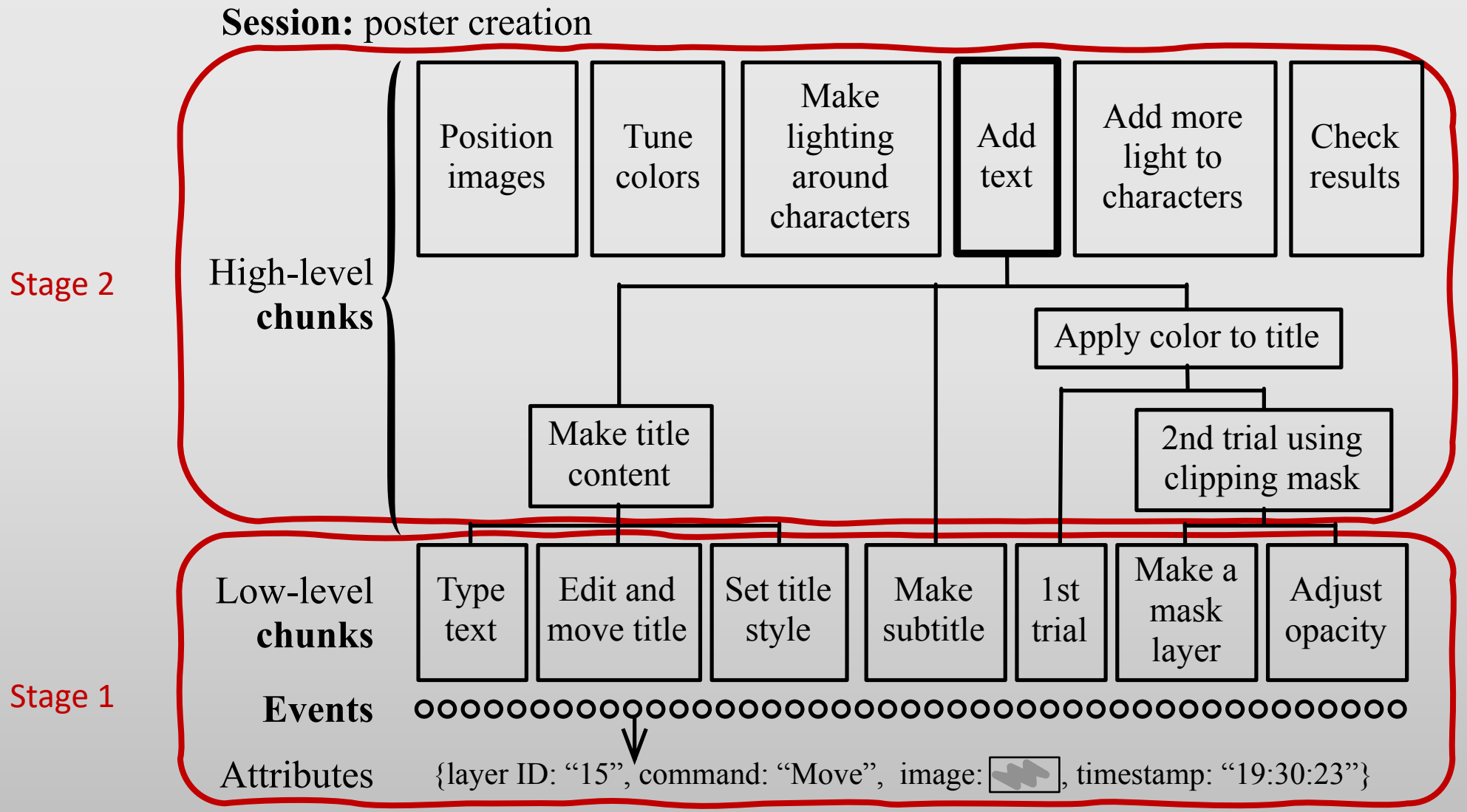
5. Inspect





1. Instrument | 2. Compute | 3. Segment | 4. Analyze | 5. Inspect

Algorithm: two stage approach

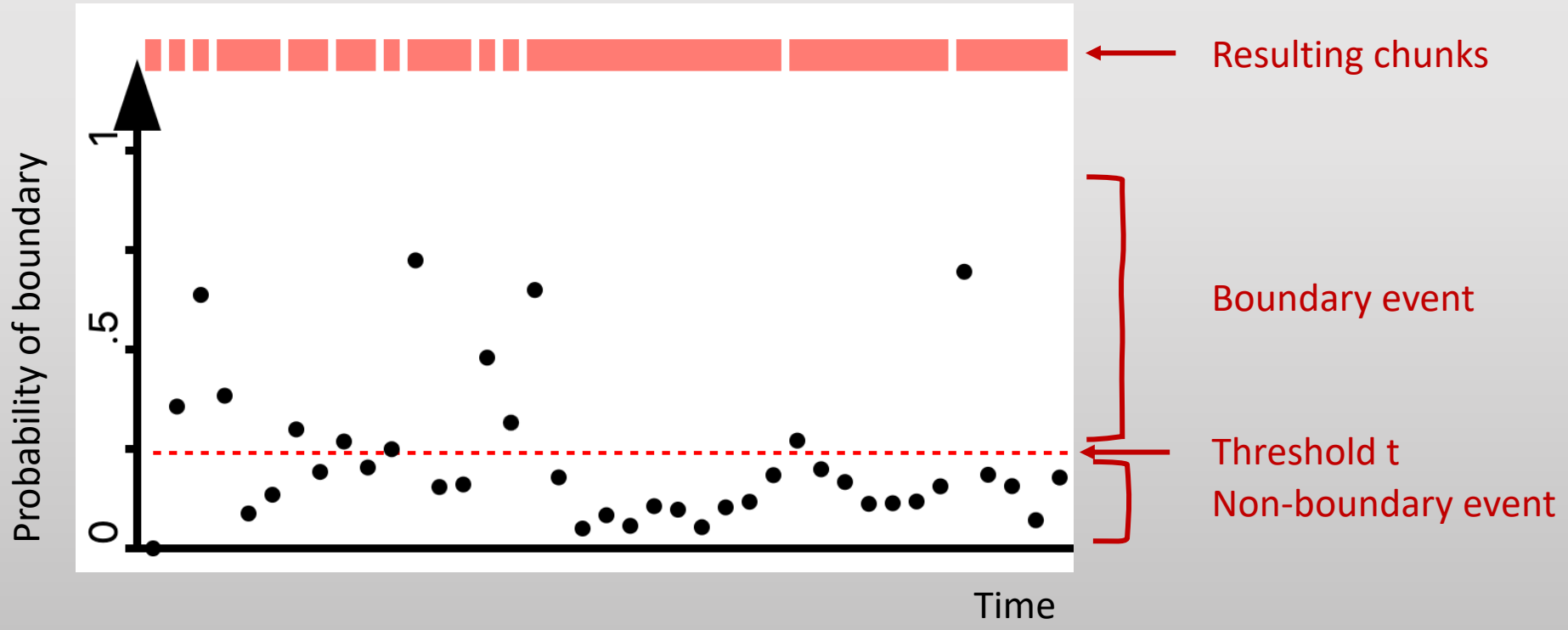


Low level: binary classification

- 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

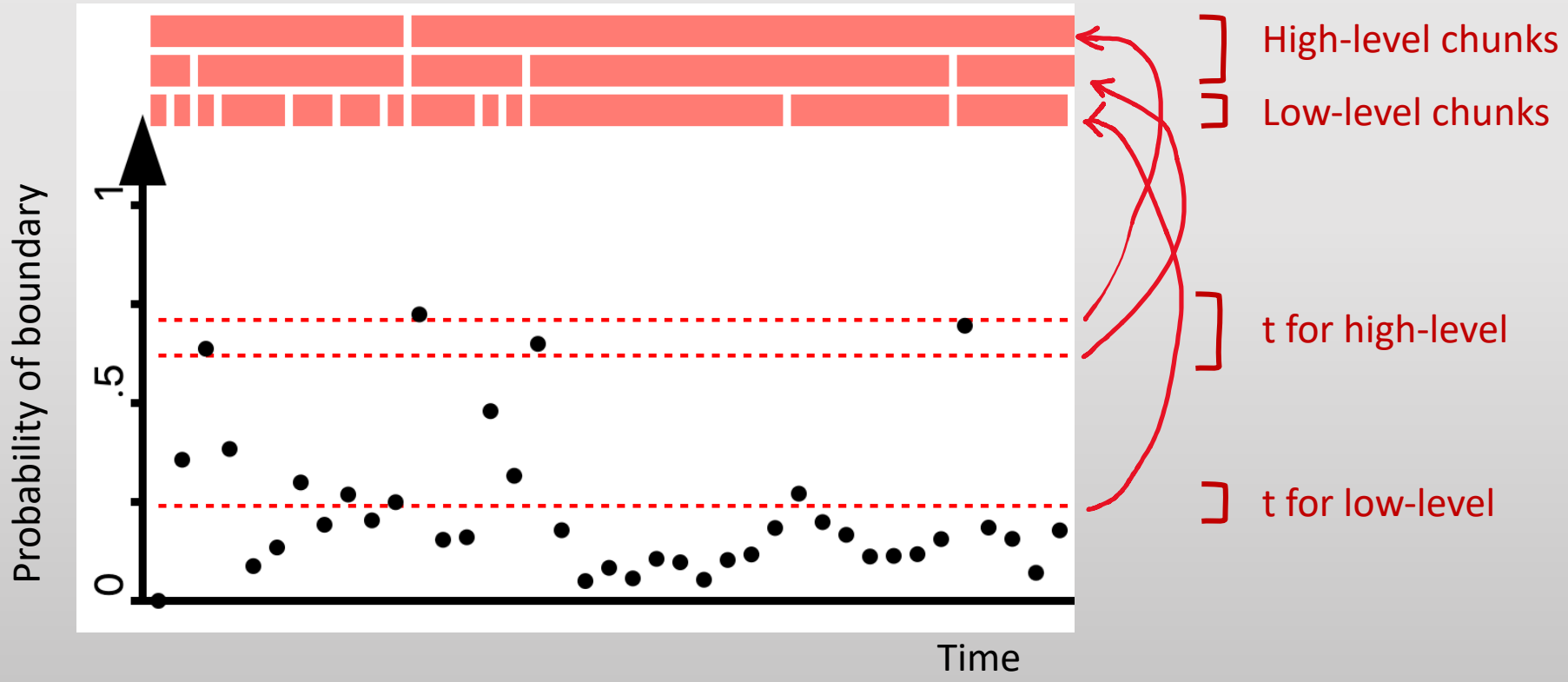
Low level: binary classification

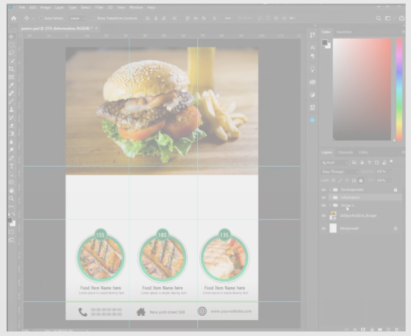
- Cost for smart undo:
  - Missed boundary (false negative) > over segmentation (false positive)
  - Favor recall over precision (use F2 metric)



High level: multi-tier thresholds

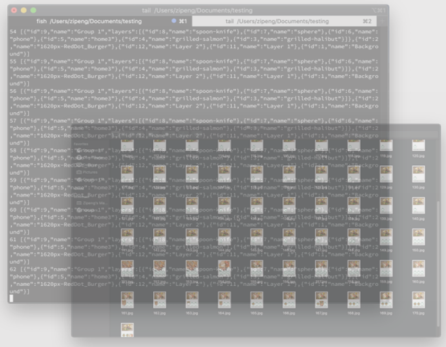
- Threshold t = granularity of segmentation





1. Instrument

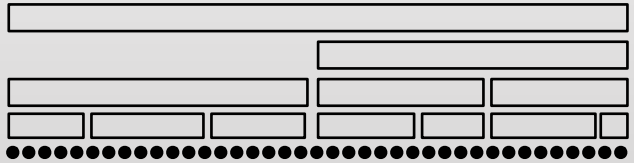
Command
Timestamp
Image
Layers …



2. Compute

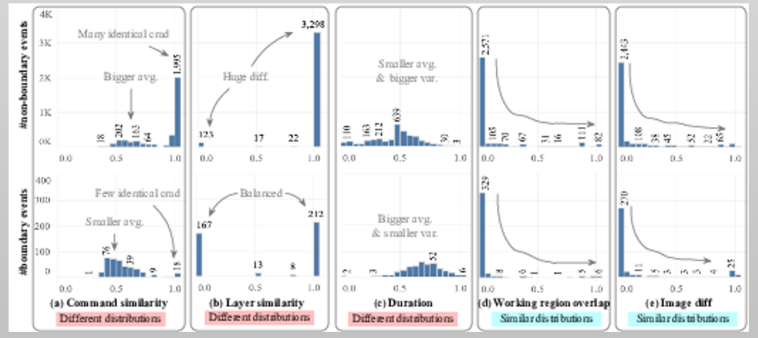
Features
(event similarity)

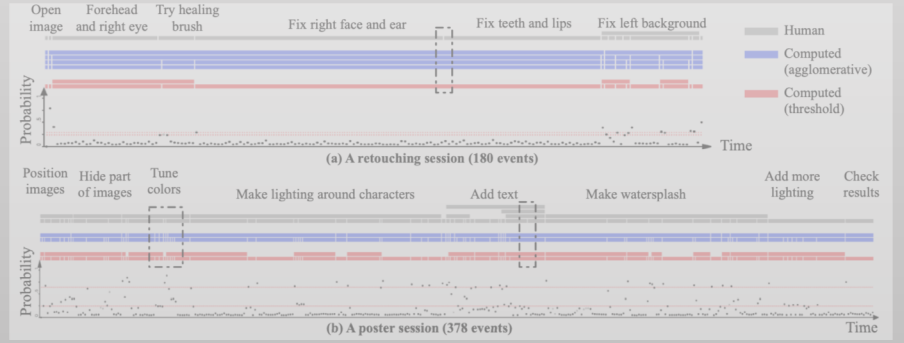
3. Segment



4. Analyze

5. Inspect





1. Instrument | 2. Compute | 3. Segment | 4. Analyze | 5. Inspect

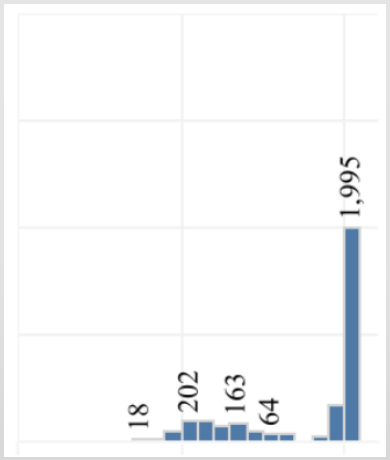
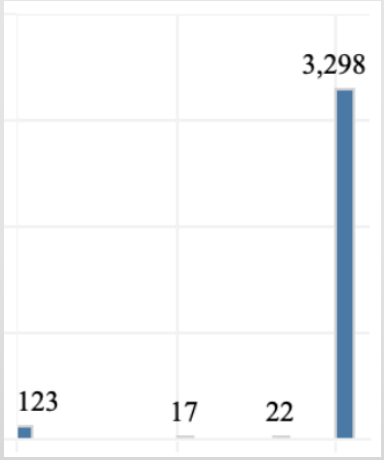
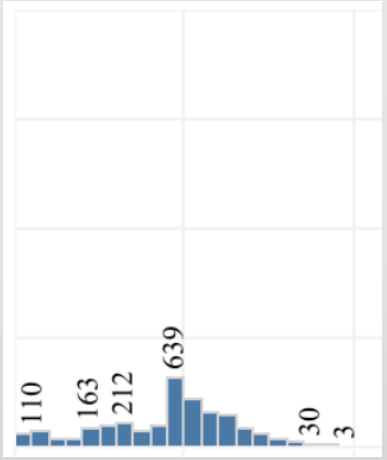
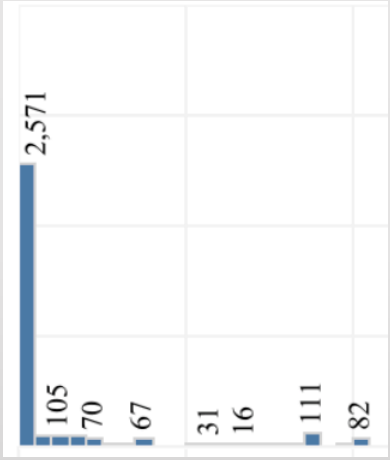
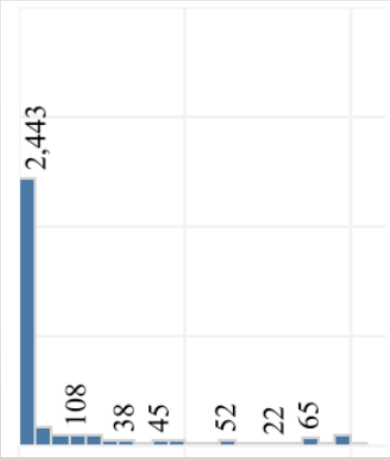
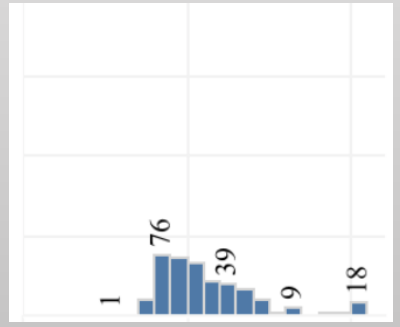
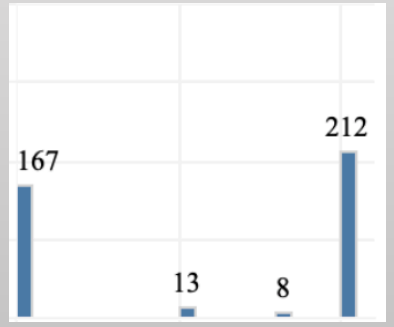
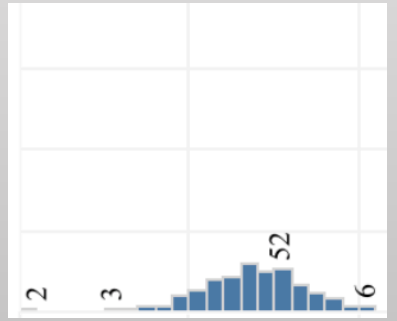
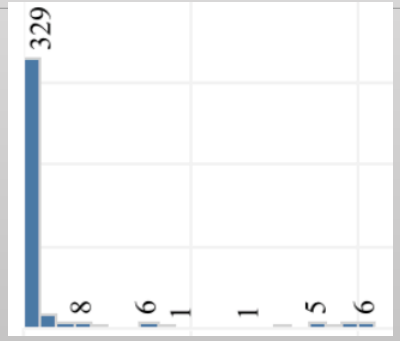
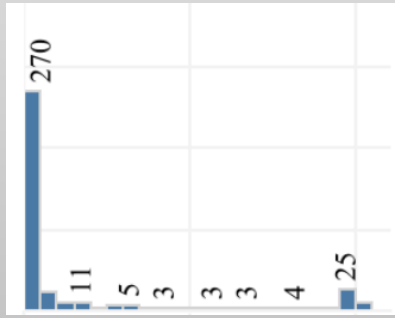
Quantitative analysis

	Command similarity	Layer similarity	Duration	Working region overlap	Image diff
Relevance					

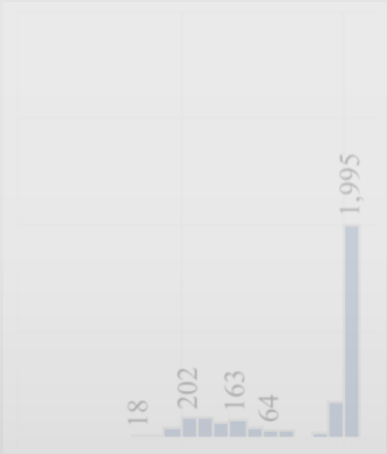
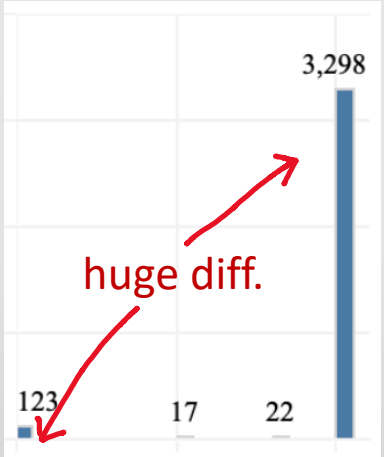
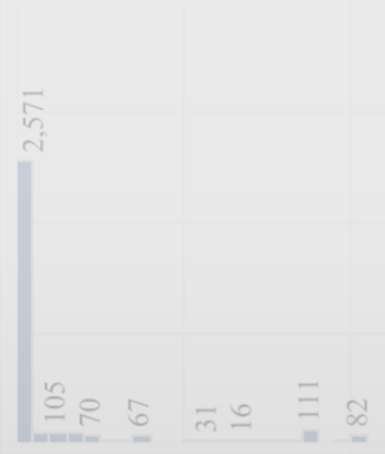
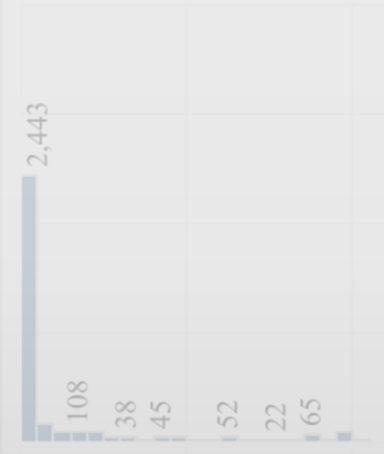
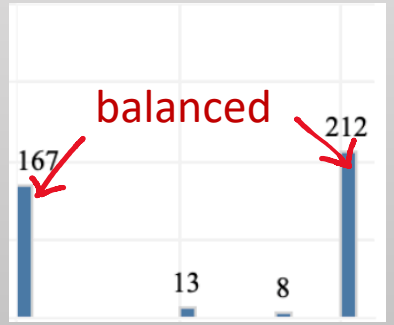
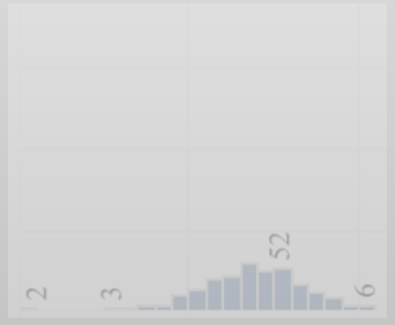
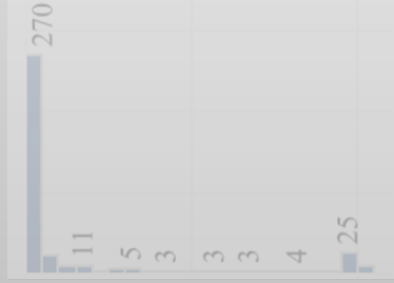
Quantitative analysis

	Command similarity	Layer similarity	Duration	Working region overlap	Image diff
Relevance					
Linear Coefficient	-2.57	-1.74	+1.55	-0.18	-0.07

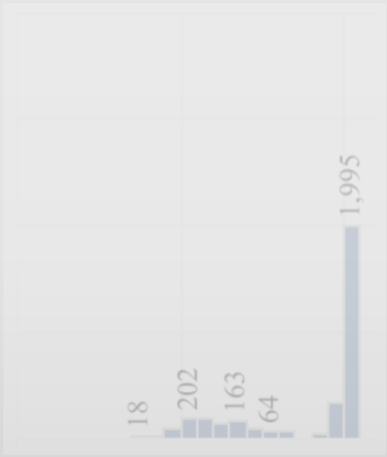
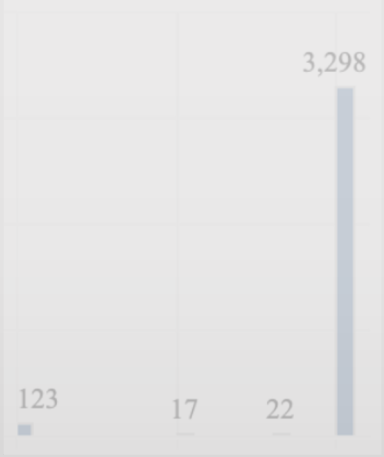
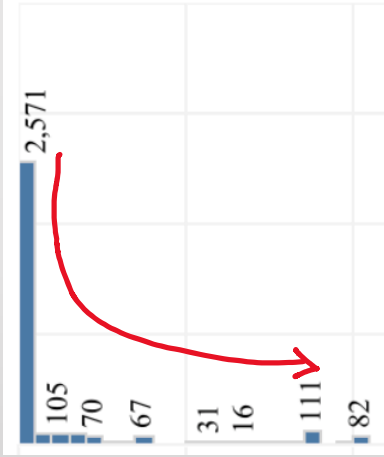
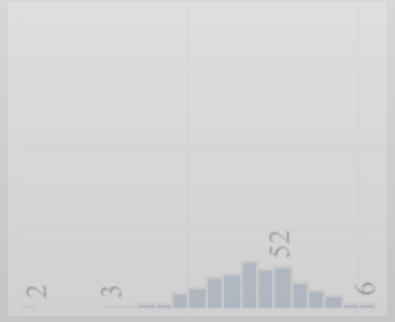
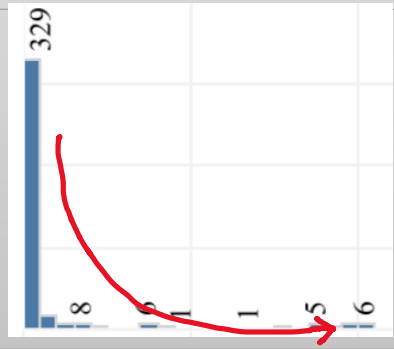
Qualitative analysis

	Command similarity	Layer similarity	Duration	Working region overlap	Image diff
Relevance					
Linear coefficient	-2.57	-1.74	+1.55	-0.18	-0.07
Distribution of non-boundary events					
Distribution of boundary events					
Top vs. down					

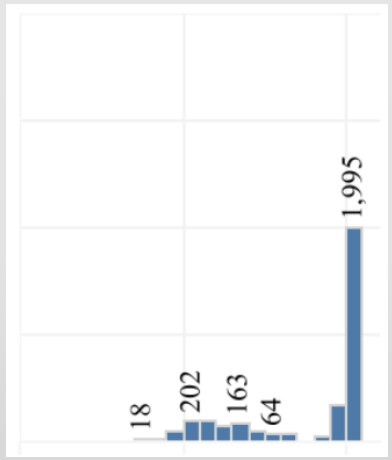
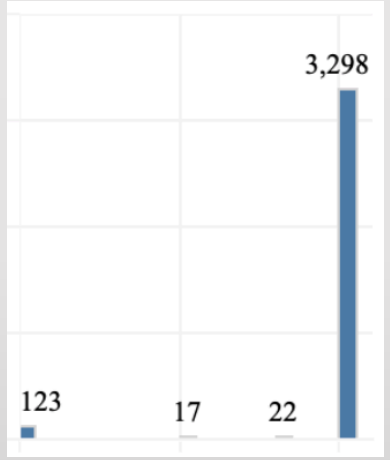
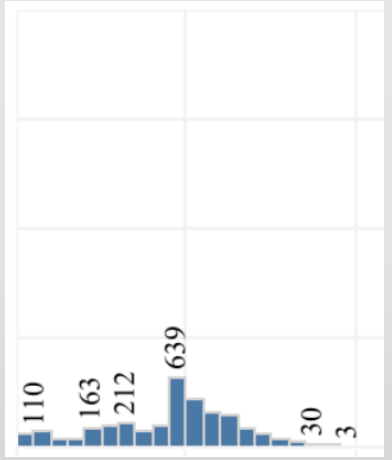
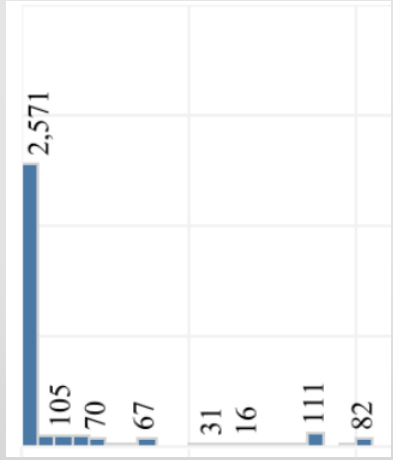
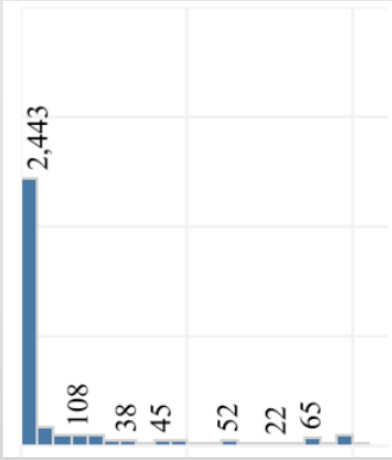
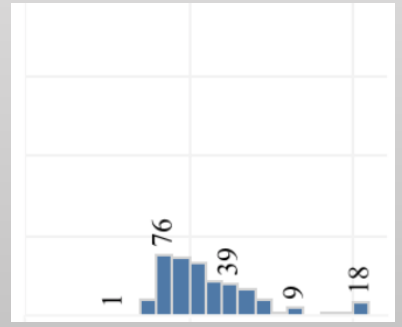
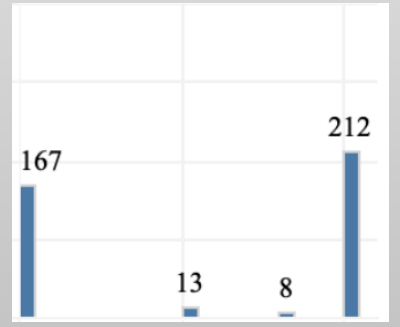
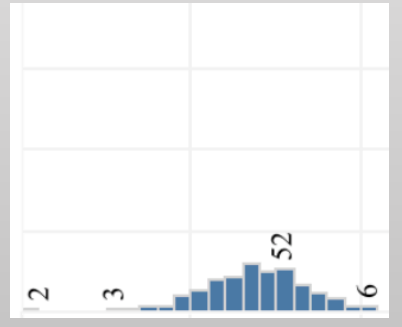
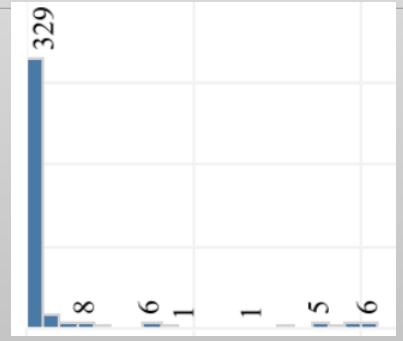
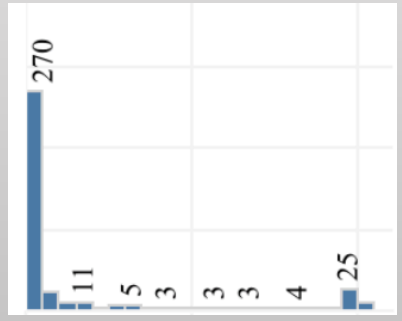
Qualitative analysis

	Command similarity	Layer similarity	Duration	Working region overlap	Image diff
Relevance					
Linear coefficient	-2.57	-1.74	+1.55	-0.18	-0.07
Distribution of non-boundary events					
Distribution of boundary events					
Top vs. down					

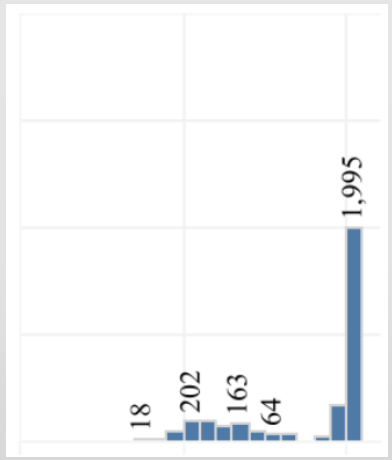
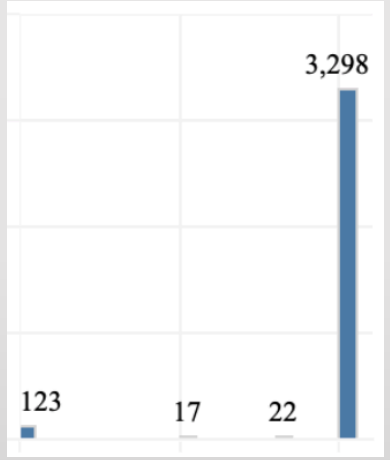
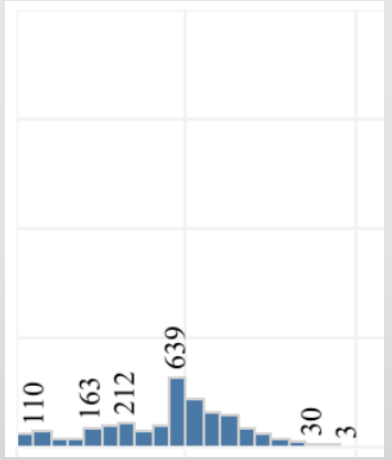
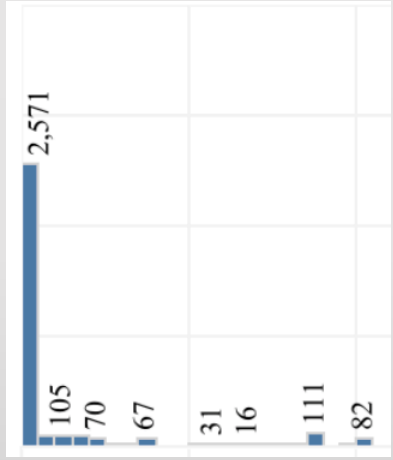
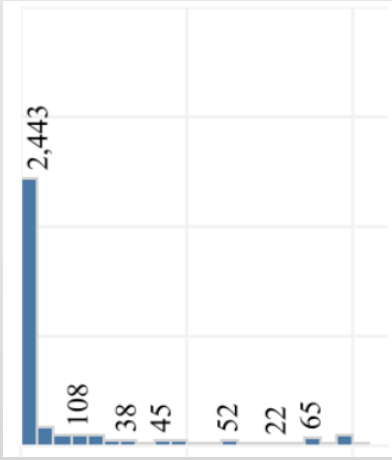
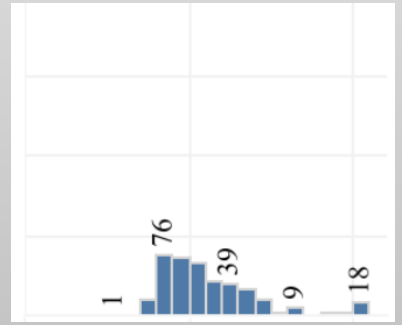
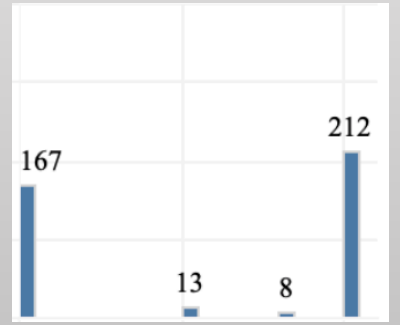
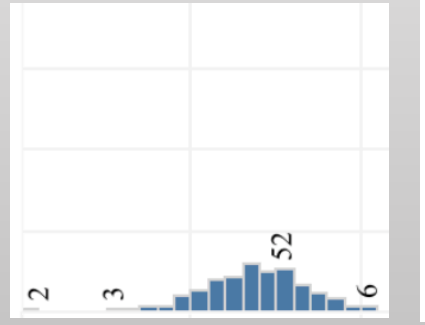
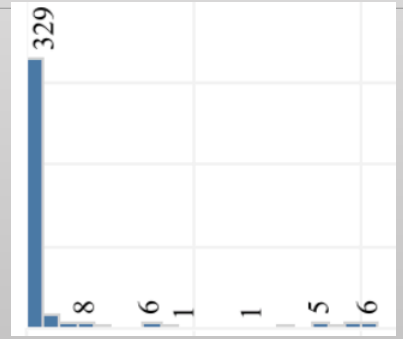
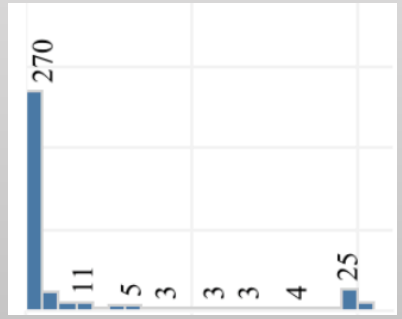
Qualitative analysis

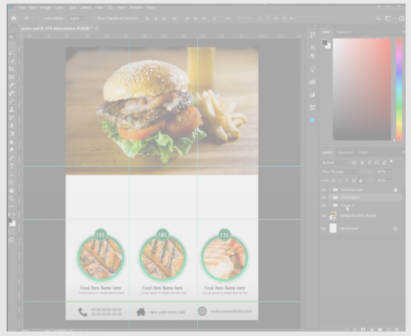
	Command similarity	Layer similarity	Duration	Working region overlap	Image diff
Relevance					
Linear coefficient	-2.57	-1.74	+1.55	-0.18	-0.07
Distribution of non-boundary events					
Distribution of boundary events					
Top vs. down					

Qualitative analysis

	Command similarity	Layer similarity	Duration	Working region overlap	Image diff
Relevance					
Linear coefficient	-2.57	-1.74	+1.55	-0.18	-0.07
Distribution of non-boundary events					
Distribution of boundary events					
Top vs. down	Different distributions			Similar distributions	

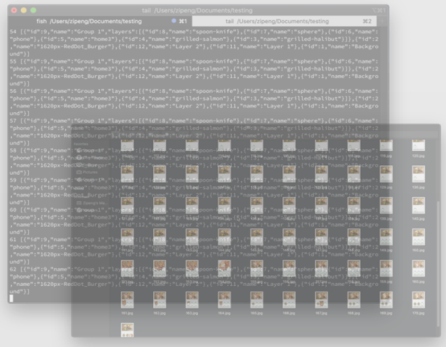
Feature relevance analysis

	Command similarity	Layer similarity	Duration	Working region overlap	Image diff
Relevance category	Most important	Important		No effect	
Linear coefficient	-2.57	-1.74	+1.55	-0.18	-0.07
Distribution of non-boundary events					
Distribution of boundary events					
Top vs. down	Different distributions			Similar distributions	



1. Instrument

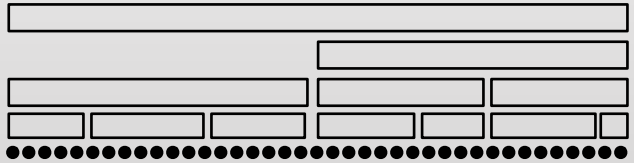
Command
Timestamp
Image
Layers …



2. Compute

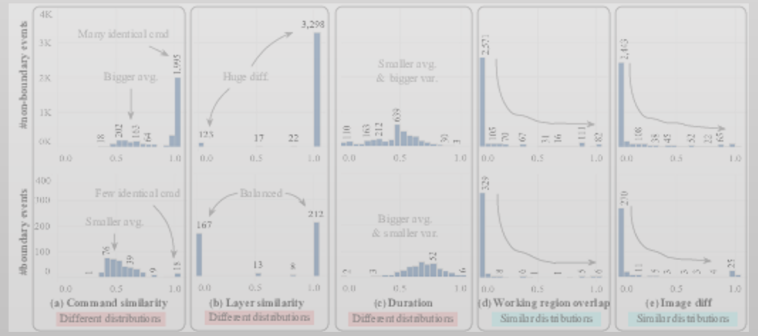
Features
(event similarity)

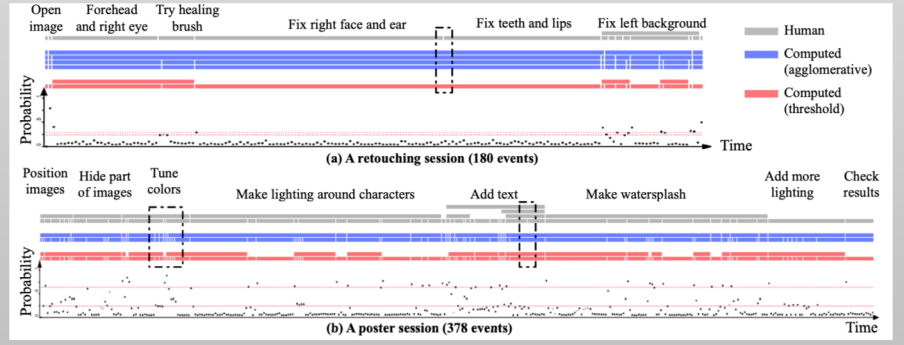
3. Segment



4. Analyze

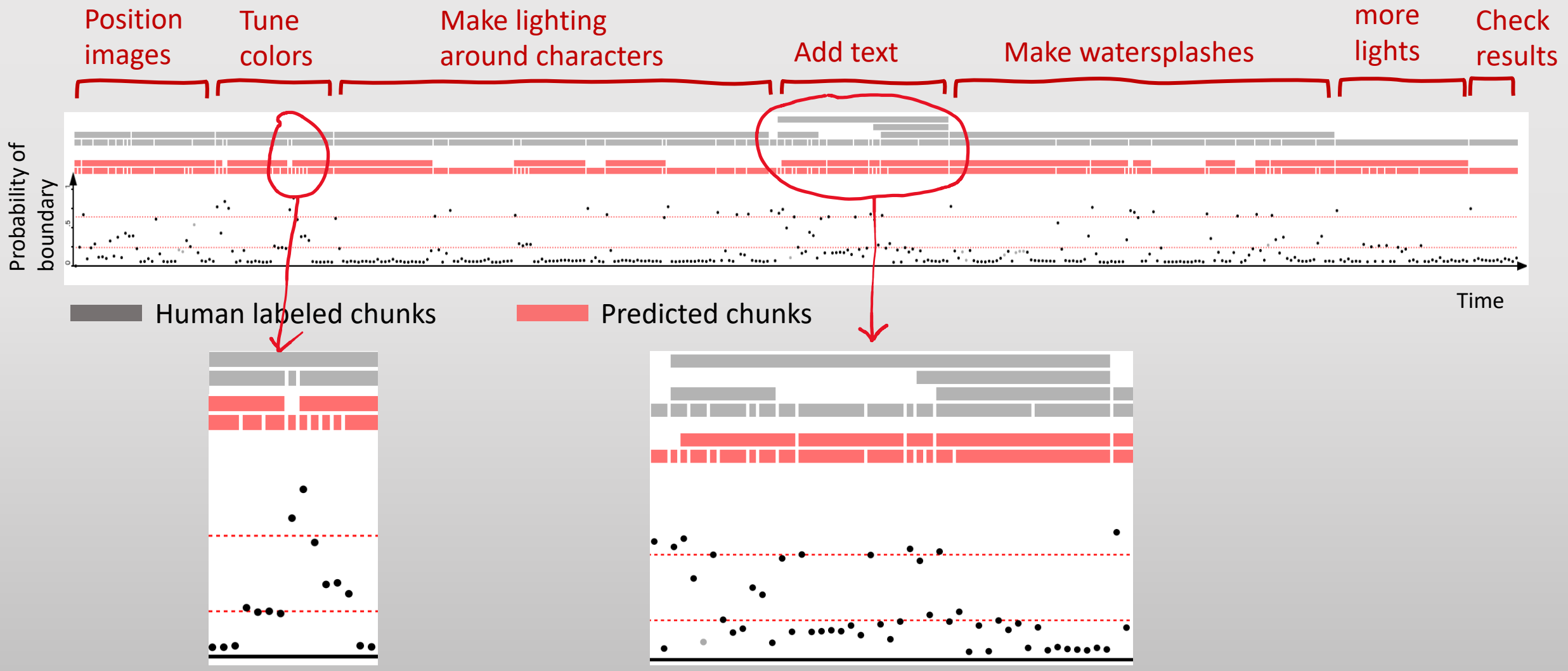
5. Inspect





1. Instrument | 2. Compute | 3. Segment | 4. Analyze | 5. Inspect

Example: poster creation



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

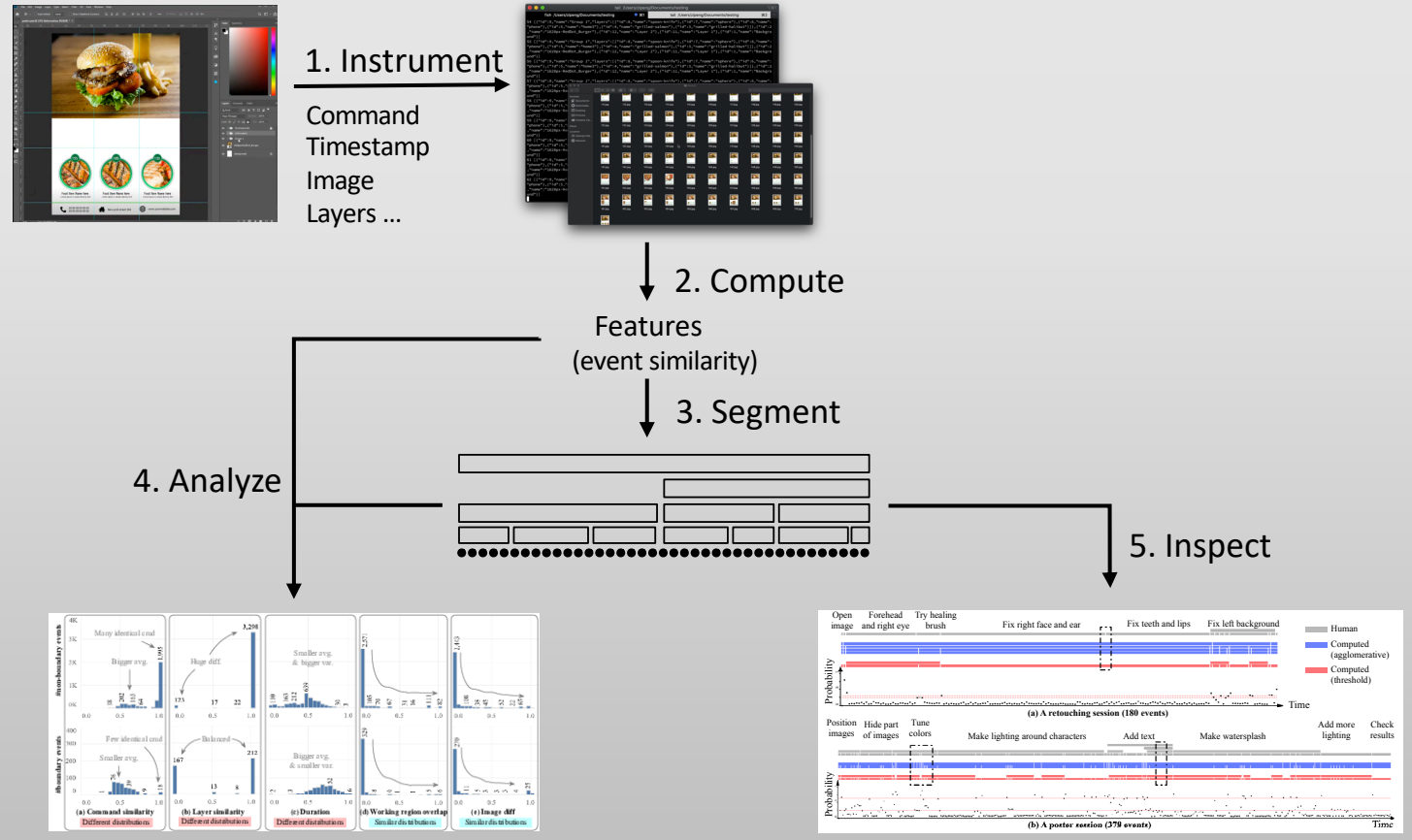
Data-driven Multi-level Segmentation of Image Editing Logs.

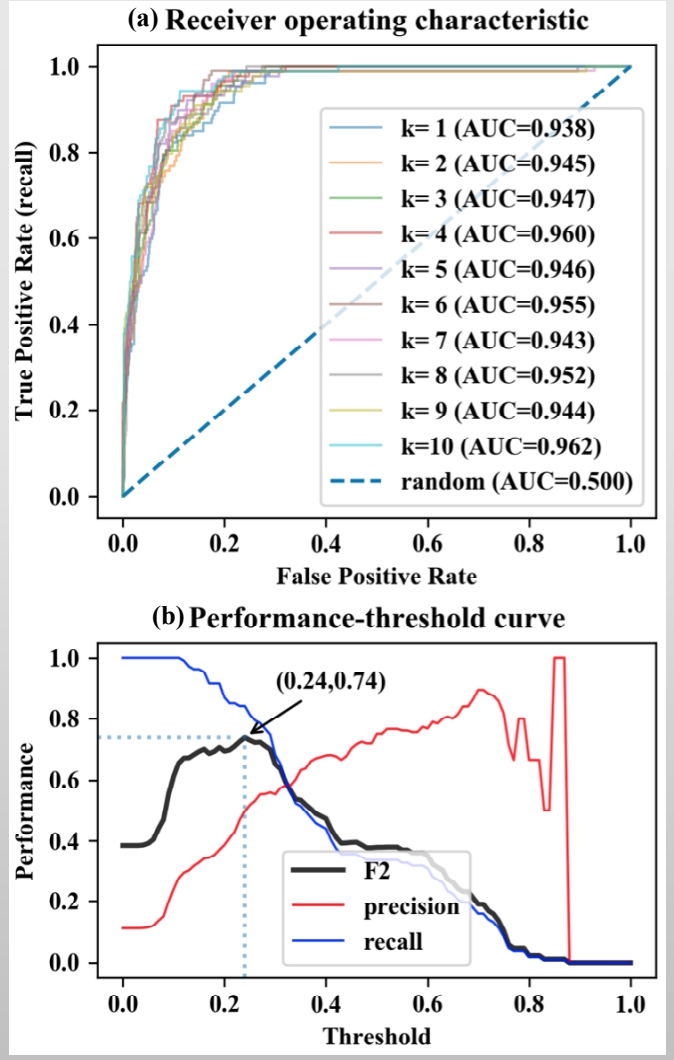
Zipeng Liu, Leo Zhicheng Liu, Tamara Munzner.
Proc. CHI Conf. Human Factors in Computing Systems (CHI), 2020
Presented at Adobe @ CHI online event on April 28 2020
http://www.cs.ubc.ca/labs/imager/tr/2020/logseg/

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

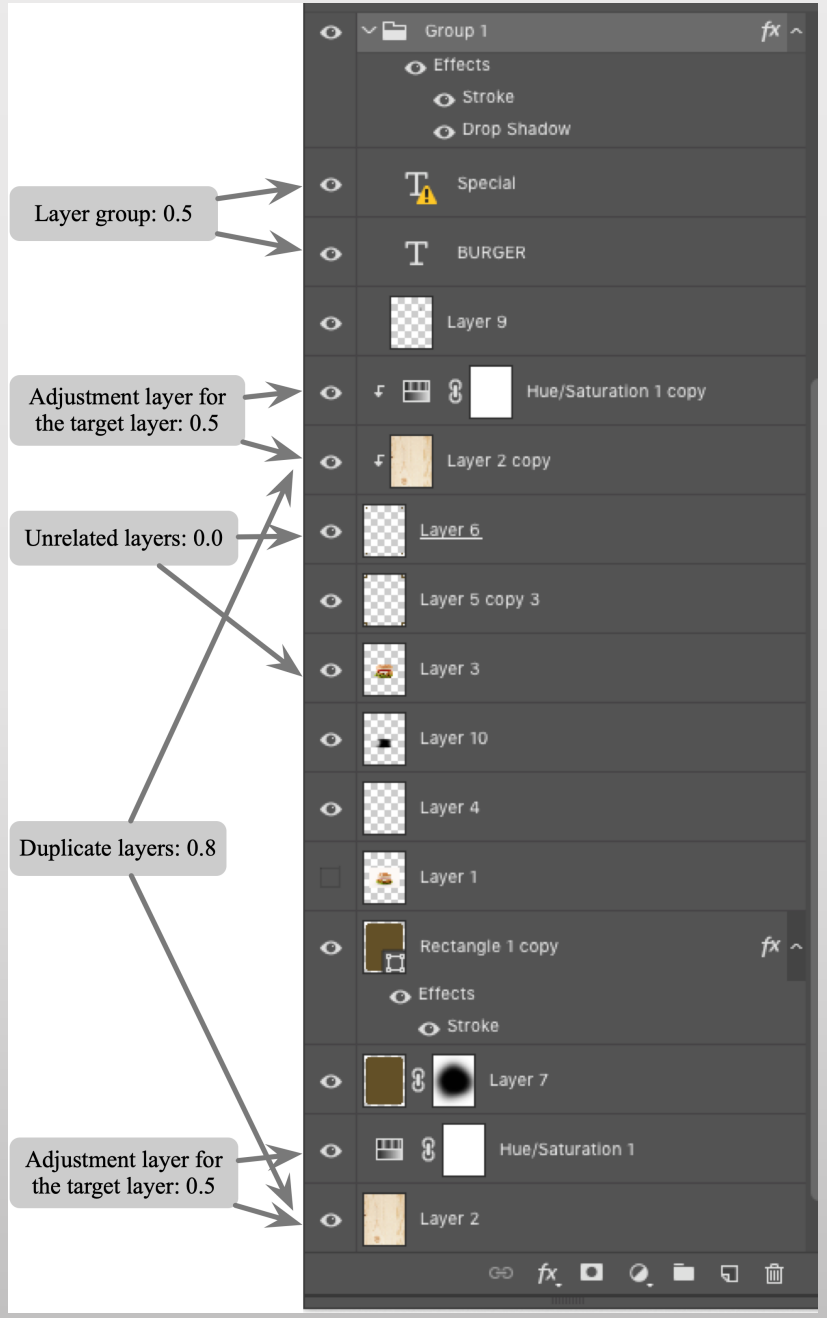



(a) Receiver operating characteristic
True Positive Rate (recall)
False Positive Rate
k= 1 (AUC=0.938)
k= 2 (AUC=0.945)
k= 3 (AUC=0.947)
k= 4 (AUC=0.960)
k= 5 (AUC=0.946)
k= 6 (AUC=0.955)
k= 7 (AUC=0.943)
k= 8 (AUC=0.952)
k= 9 (AUC=0.944)
k=10 (AUC=0.962)
random (AUC=0.500)
(b) Performance-threshold curve
Performance
Threshold
(0.24,0.74)
F2
precision
recall

2. Layer similarity: rule-based

Stronger layer relationships → larger similarity

Relationship	Description	Similarity
Same layer	A = B	1.0
Duplicate layer	A is a copy of B	0.8
Adjustment layer	A is an adjustment layer of B	0.5
Grouped layer	A and B are located in the same layer group	<=0.5
Other diff. layer	none of the above	0.0



Example 1: portrait retouching

Reflections on real-world user behavior

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