Micro Perceptual Human Computation for Visual Tasks

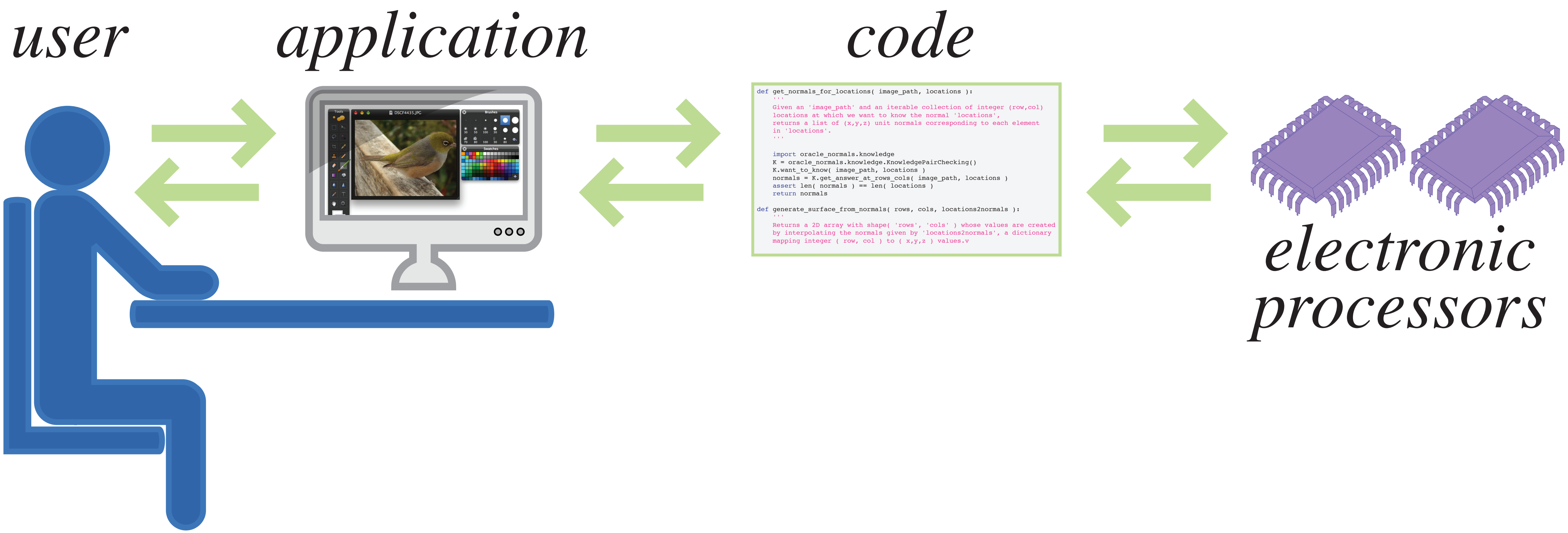
Yotam Gingold
*George Mason University**

Ariel Shamir
Herzliya IDC

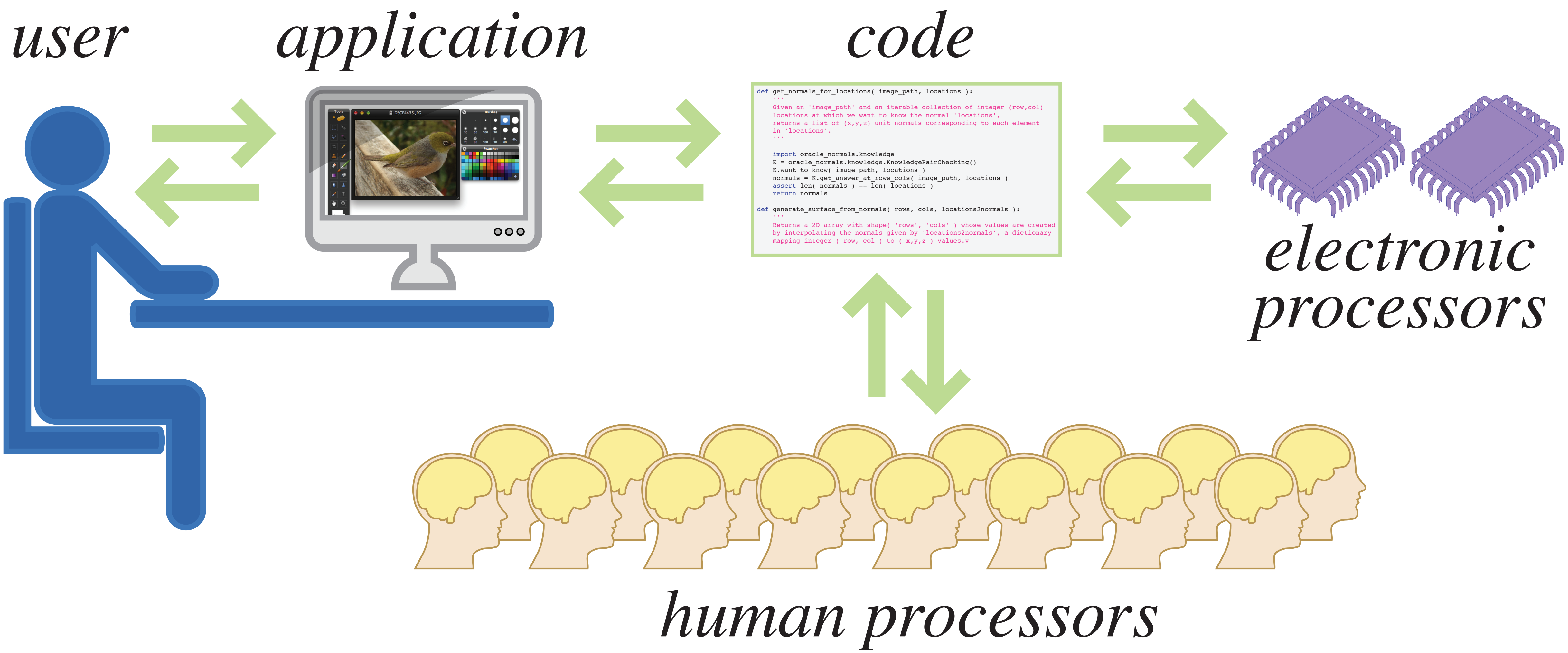
Daniel Cohen-Or
Tel-Aviv University

**Research performed while affiliated with Tel-Aviv University/Herzliya IDC/Rutgers/Columbia.*

Computation



Human Computation



The Human Advantage

Visual Perception

The Human Advantage

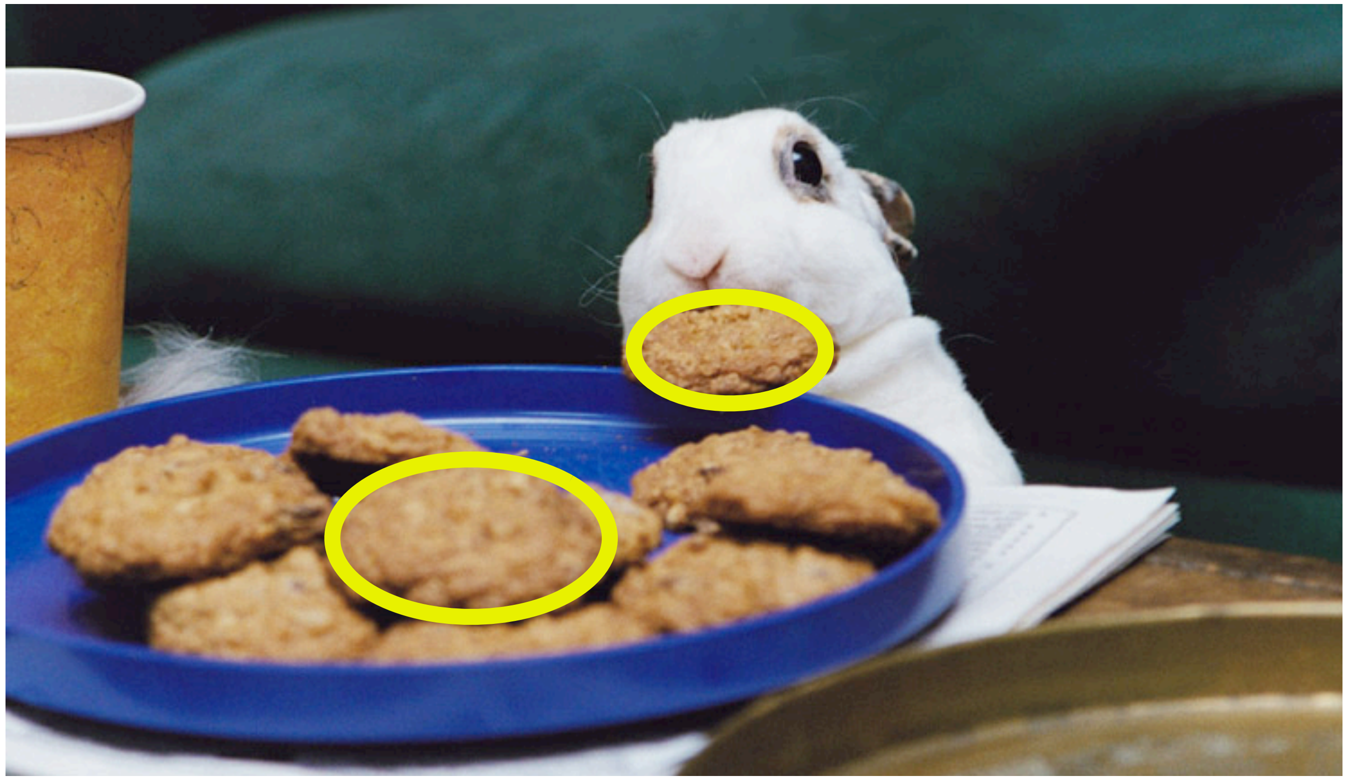
Visual Perception

- What is in this photo?

The Human Advantage

Visual Perception

- What is in this photo?
- Which object is farther away?



The Human Advantage

Visual Perception

- What is in this photo?
- Which object is farther away?
- Is this shape symmetric?

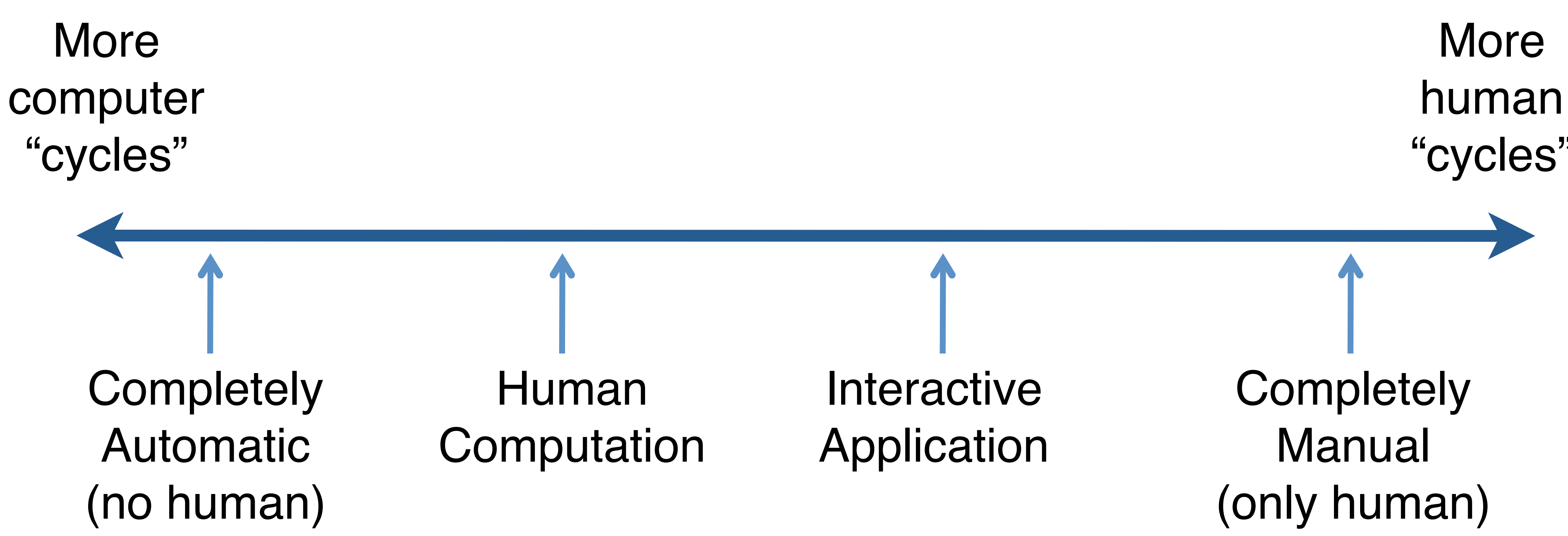
The Human Advantage

Visual Perception

- What is in this photo?
- Which object is farther away?
- Is this shape symmetric?
- What is the surface orientation (normal)?

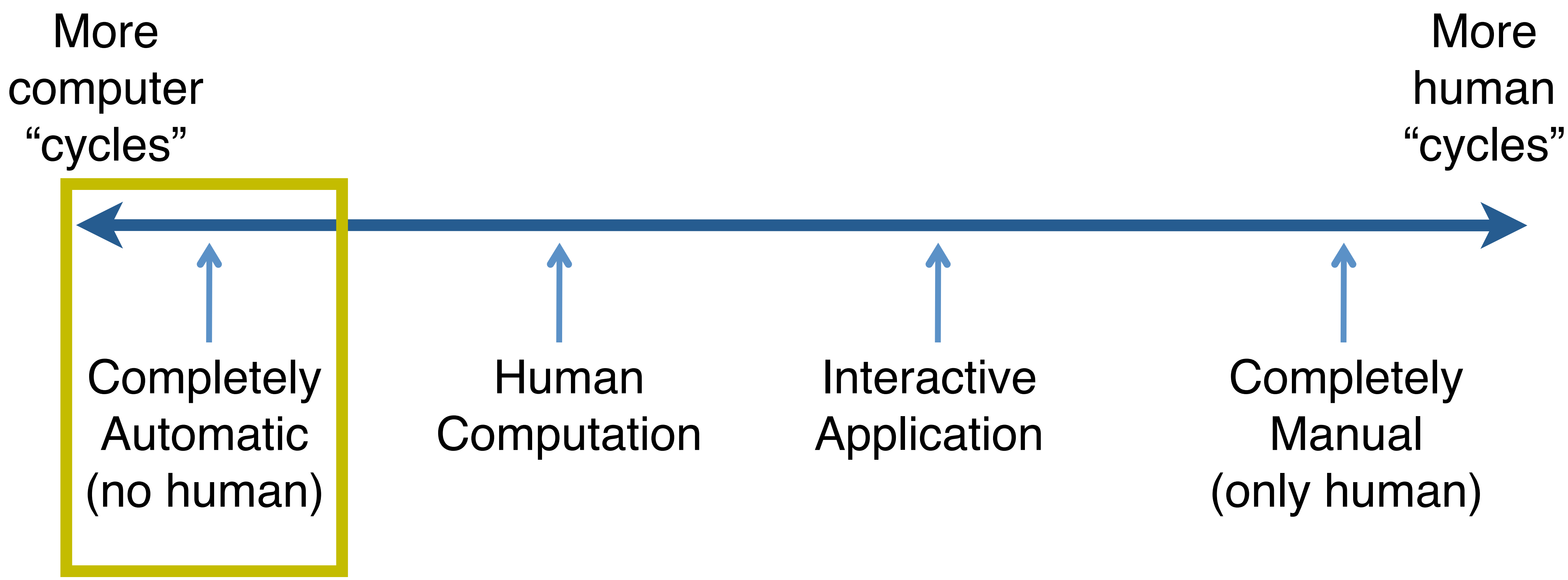
Range of Solutions

How much human and how much computer is involved?



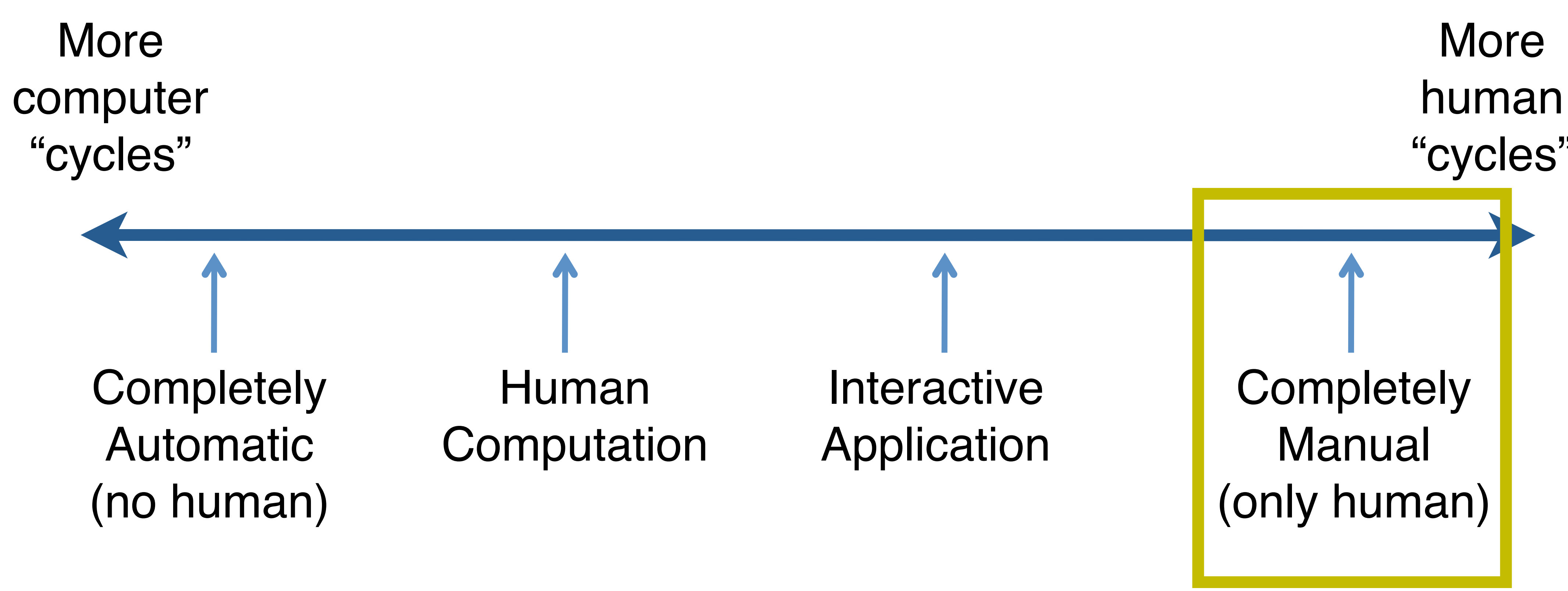
Range of Solutions

How much human and how much computer is involved?



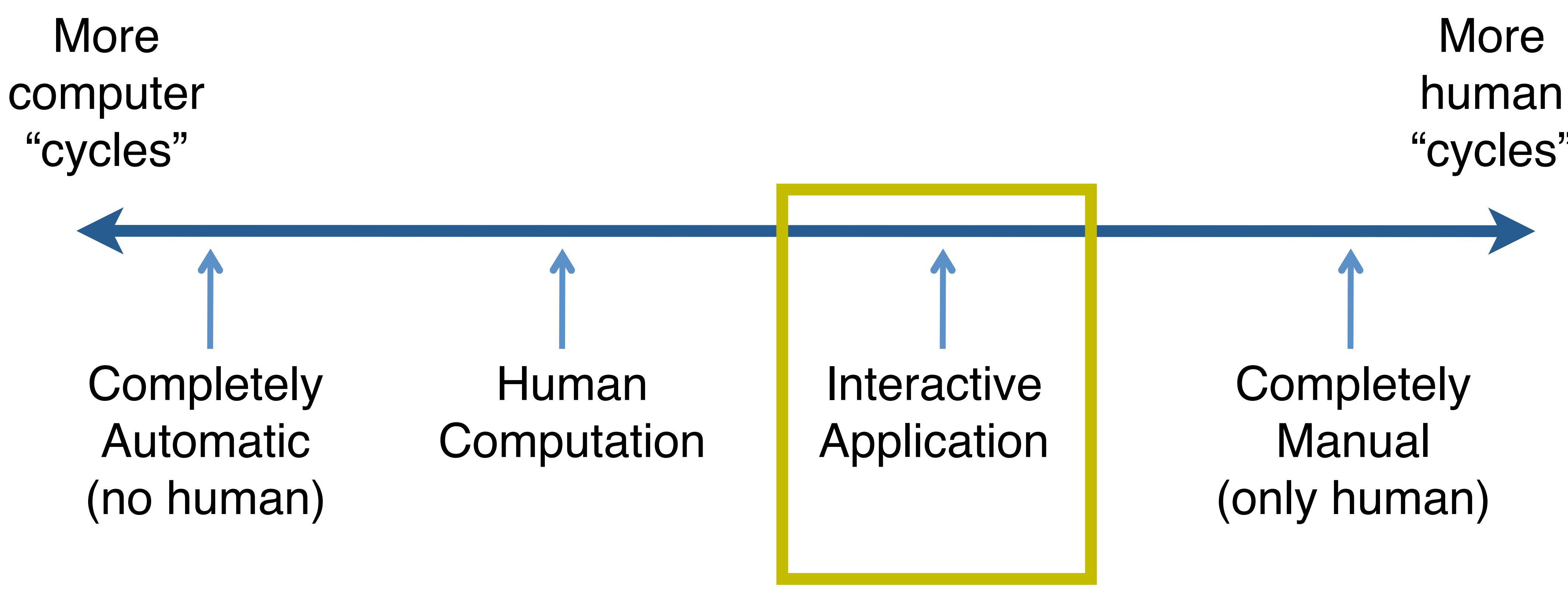
Range of Solutions

How much human and how much computer is involved?



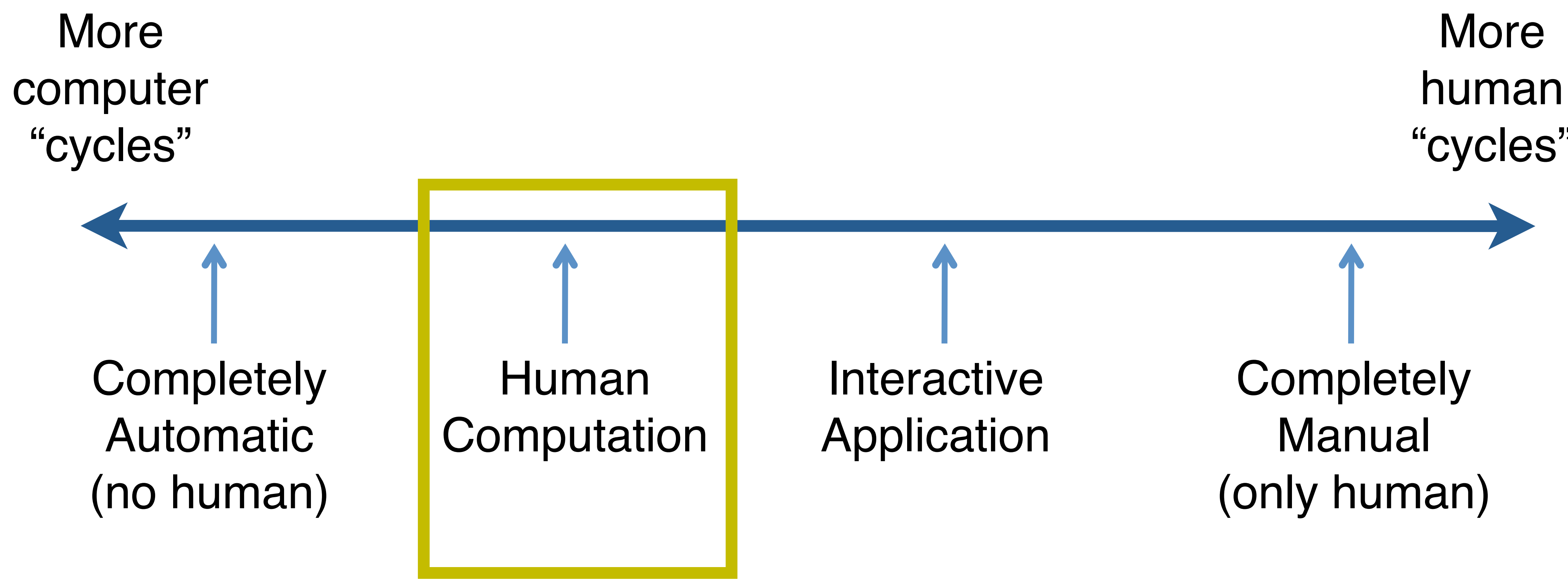
Range of Solutions

How much human and how much computer is involved?



Range of Solutions

How much human and how much computer is involved?



Key Question

What is the minimum amount of information a human could provide in order to solve the original problem?

Key Question

What is the minimum amount of information a human could provide in order to solve the original problem?

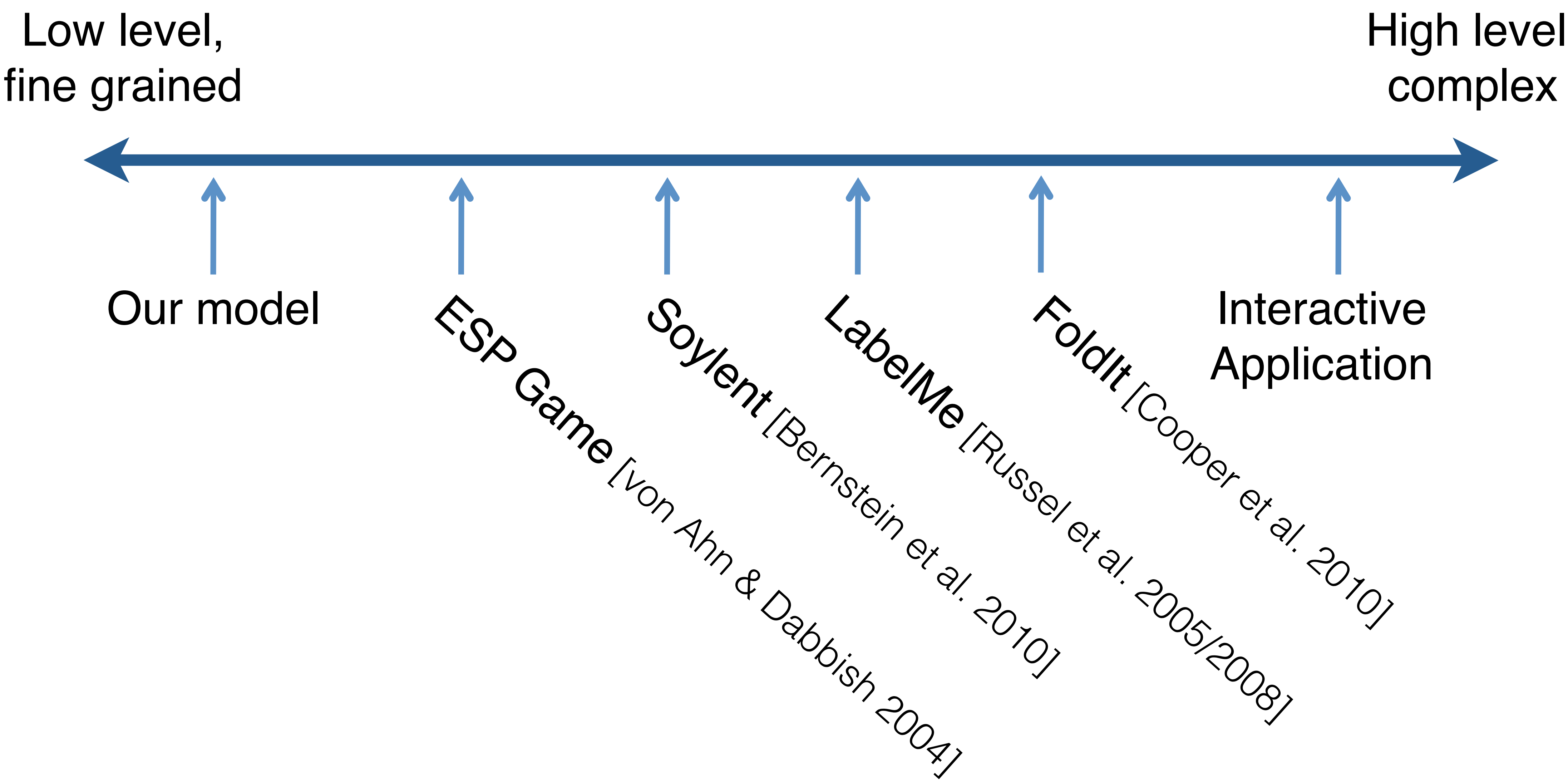
- Rephrase the algorithm in terms of the smallest piece of information that without it the problem could not be solved.

Key Question

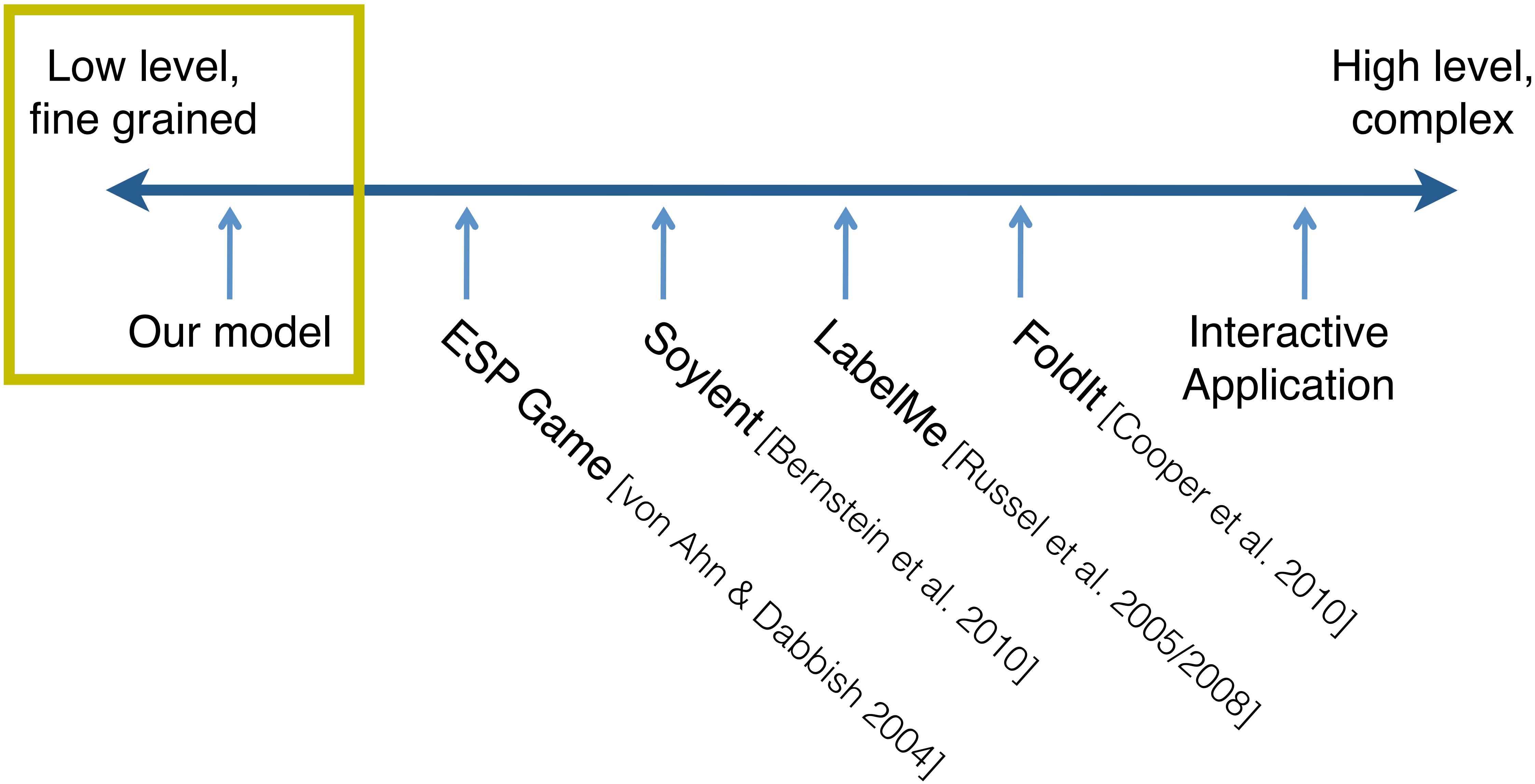
What is the minimum amount of information a human could provide in order to solve the original problem?

- Rephrase the algorithm in terms of the smallest piece of information that without it the problem could not be solved.
- Use only as much human computation as necessary, and no more than is sufficient.

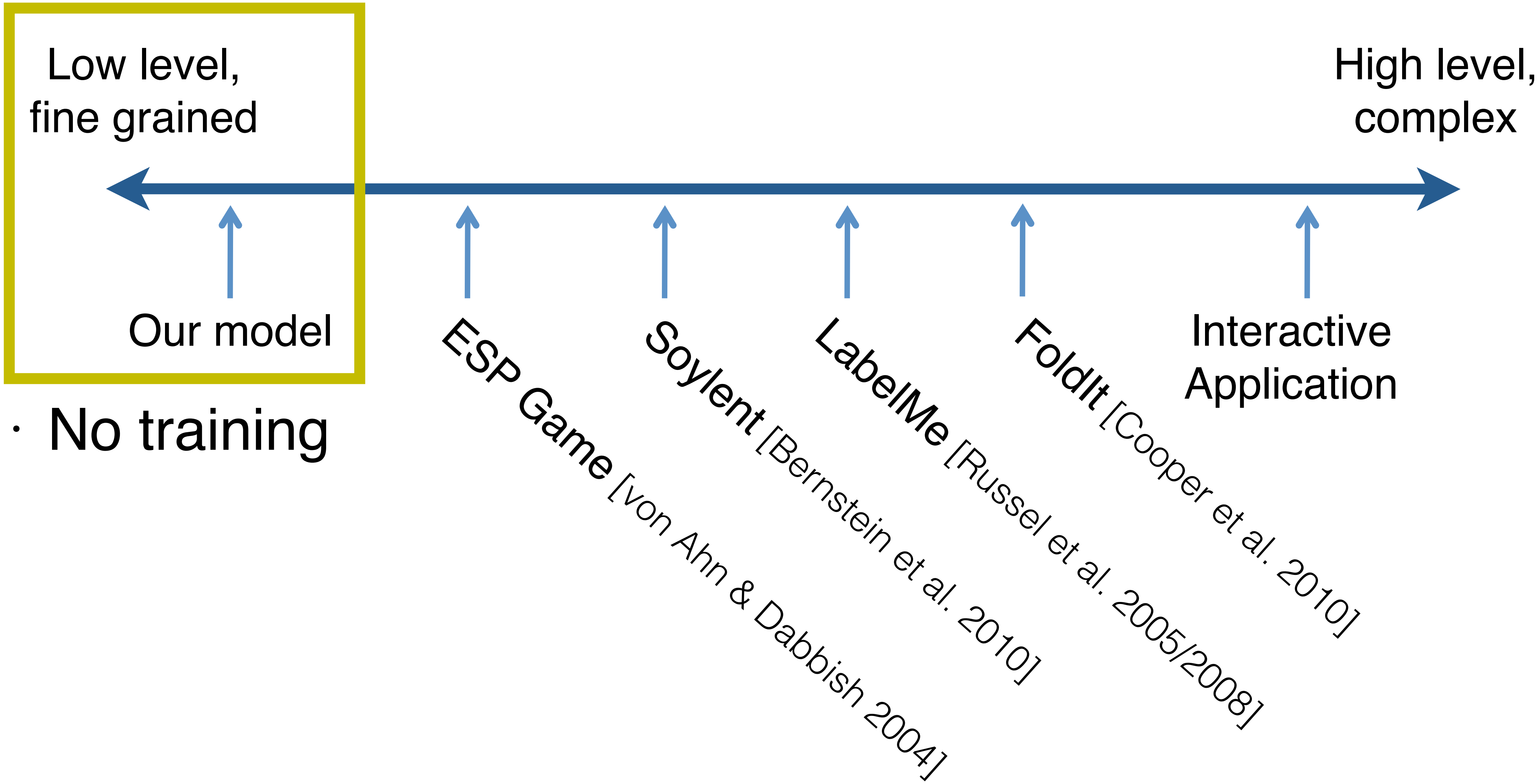
Type of Human Cycles



Type of Human Cycles



Type of Human Cycles



- No training

Type of Human Cycles

Low level, fine grained

High level, complex

Our model

ESP Game [Von Ahn & Dabbish 2004]

Soylent [Bernstein et al. 2010]

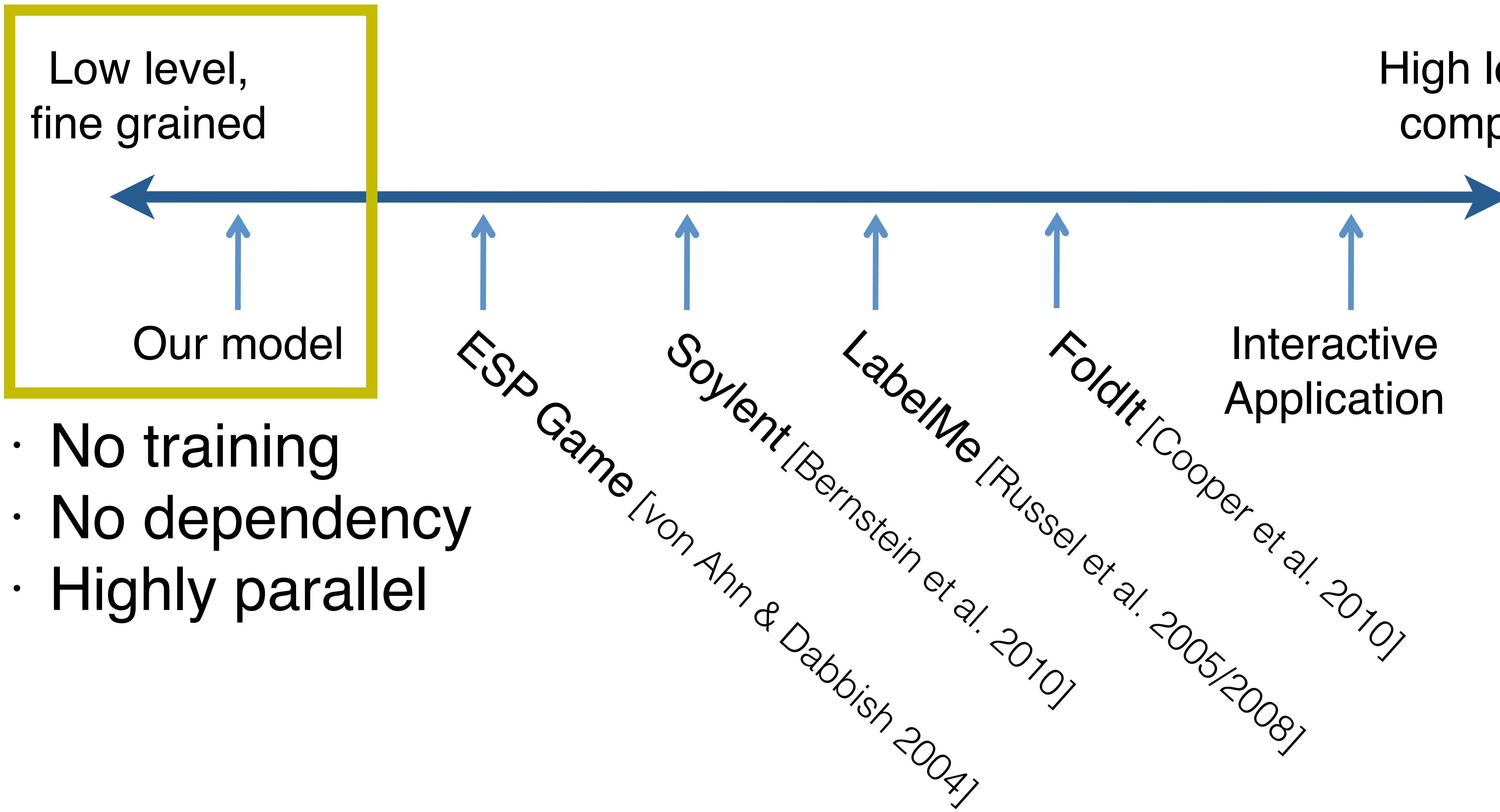
LabelMe [Russel et al. 2005/2008]

FoldIt [Cooper et al. 2010]

Interactive Application

- No training
- No dependency

Type of Human Cycles



- No training
- No dependency
- Highly parallel

Guidelines for Designing Micro-Tasks

Guidelines for Designing Micro-Tasks

Task must be simple (instantaneous)

Guidelines for Designing Micro-Tasks

Task must be simple (instantaneous)

Task must be specific (well-defined)

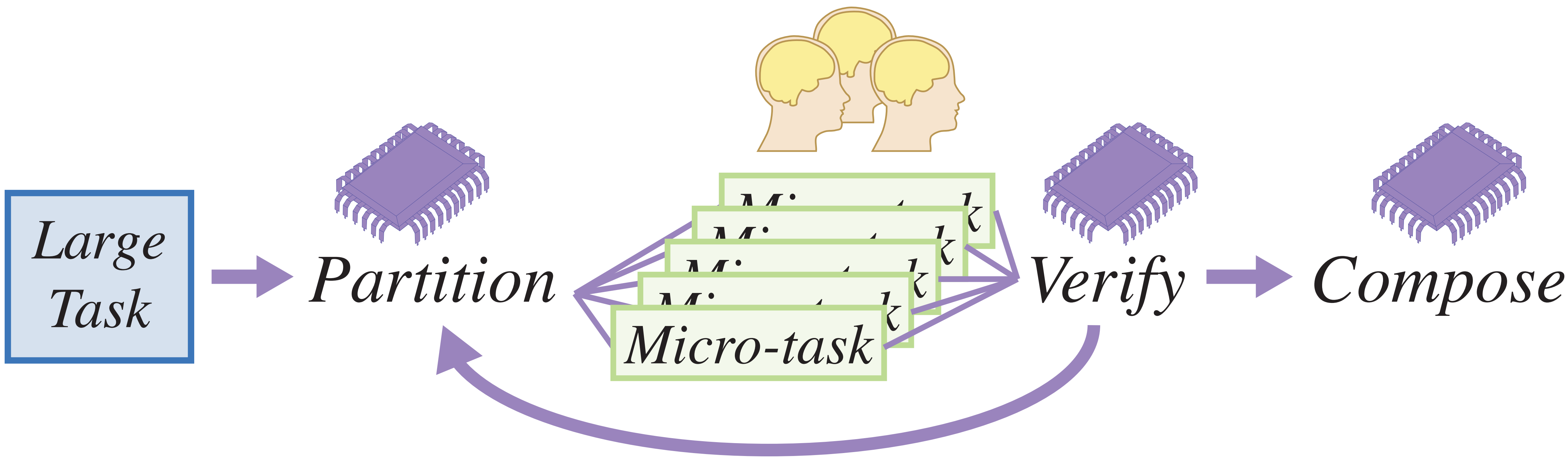
Guidelines for Designing Micro-Tasks

Task must be simple (instantaneous)

Task must be specific (well-defined)

Task must be reliable (humans can do it)

Algorithm Design Pattern



Issues

Issues

Motivation

Issues

Motivation	money (via Amazon Mechanical Turk)

Issues

Motivation	money (via Amazon Mechanical Turk)
Efficiency	

Issues

Motivation	money (via Amazon Mechanical Turk)
Efficiency	Massive parallelism Extremely simple visual queries

Issues

Motivation	money (via Amazon Mechanical Turk)
Efficiency	Massive parallelism Extremely simple visual queries
Quality Control	

Issues

Motivation	money (via Amazon Mechanical Turk)
Efficiency	Massive parallelism Extremely simple visual queries
Quality Control	Batches:

1	2	3	4	5	6

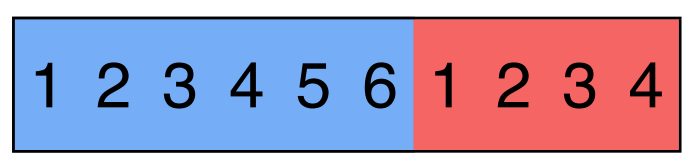
queries

Issues

Motivation money (via Amazon Mechanical Turk)

Efficiency Massive parallelism
Extremely simple visual queries

Quality Control Batches:



queries

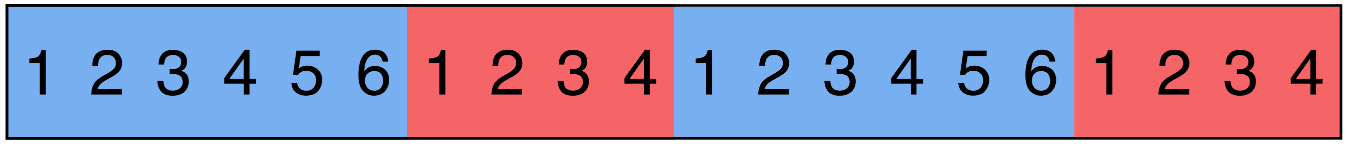
queries
with known answers

Issues

Motivation — money (via Amazon Mechanical Turk)

Efficiency — Massive parallelism
Extremely simple visual queries

Quality Control — Batches:



Issues

Motivation	money (via Amazon Mechanical Turk)
Efficiency	Massive parallelism Extremely simple visual queries
Quality Control	Batches:

4	3	2	6	4	1	2	6	2	4	1	3	3	5	1	3	4	1	5	2

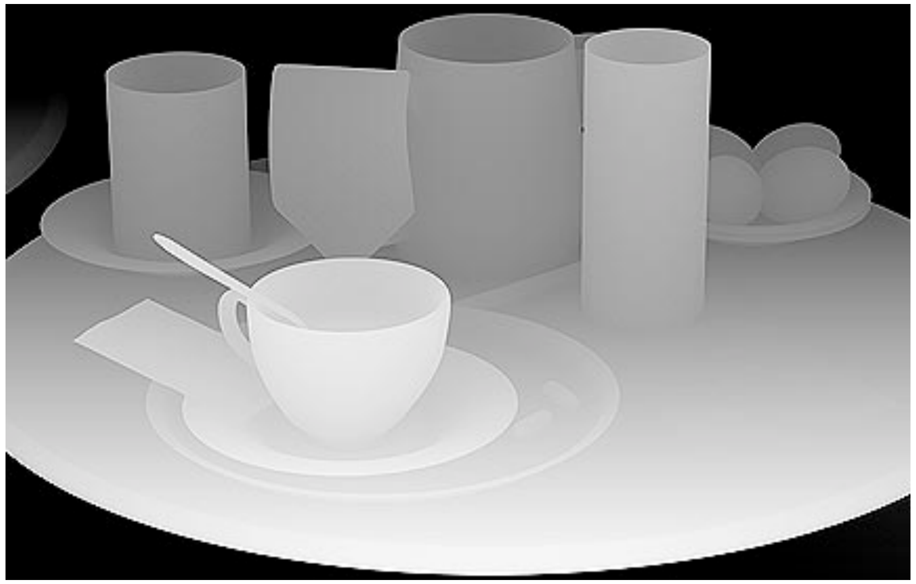
Three Example Algorithms

Given an image, create

Three Example Algorithms

Given an image, create

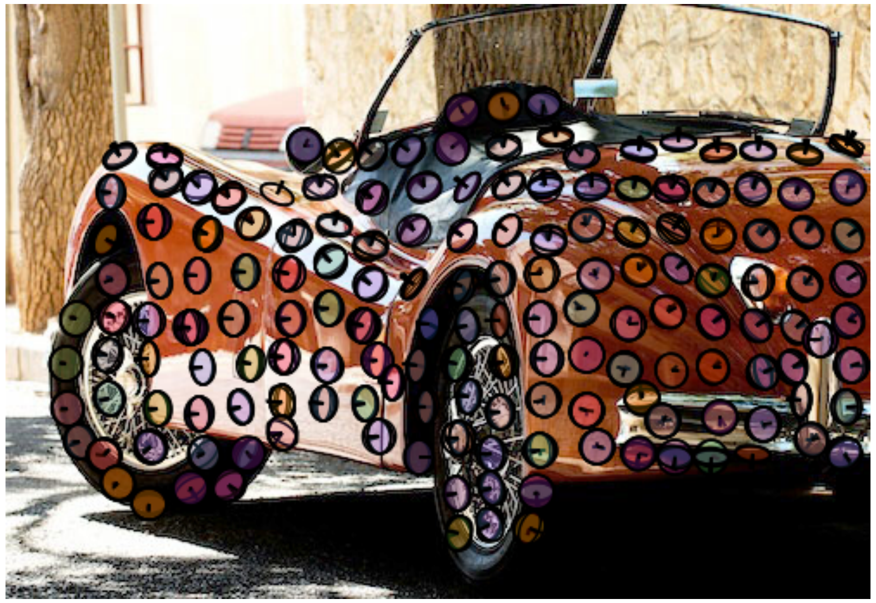
- depth layers



Three Example Algorithms

Given an image, create

- depth layers
- a normal map

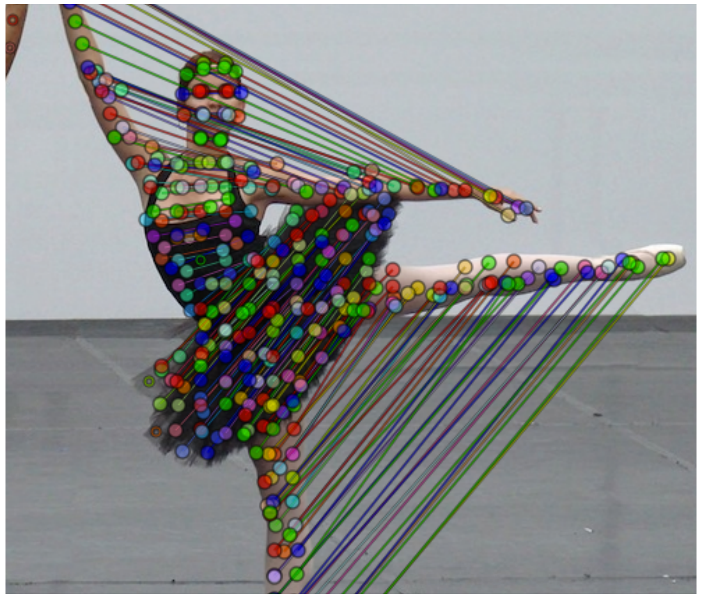


[Pedro Ribeiro Simões]

Three Example Algorithms

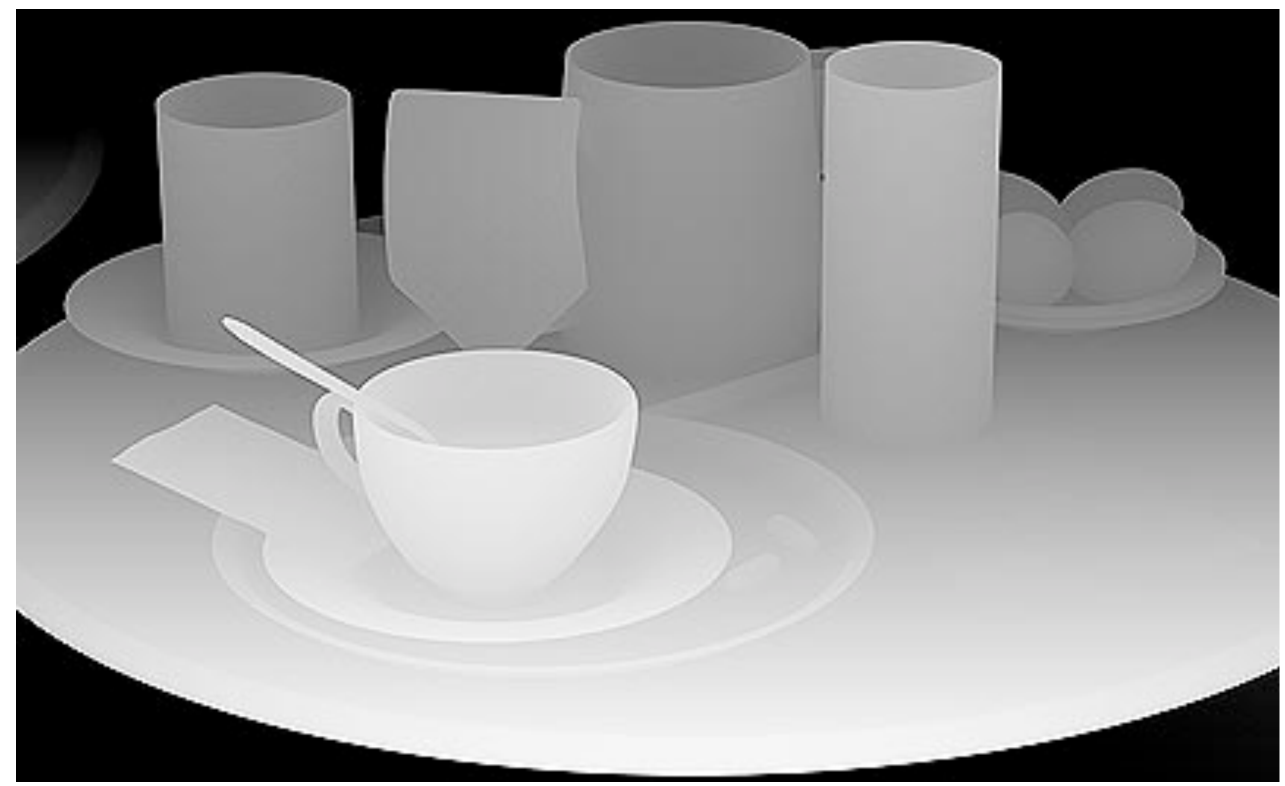
Given an image, create

- depth layers
- a normal map
- a bilateral symmetry map



[flickr user dalbera]

Algorithm 1: Depth Layers



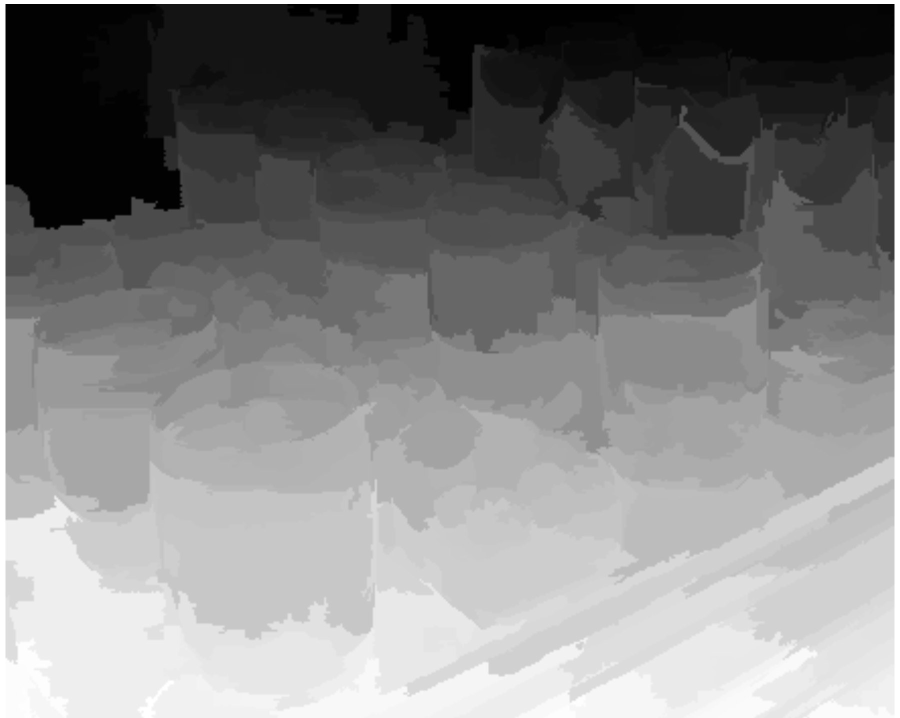
Automatic methods

e.g. [Hoiem et al. 2005; Assa and Wolf 2007; Saxena et al. 2009]

Automatic methods

e.g. [Hoiem et al. 2005; Assa and Wolf 2007; Saxena et al. 2009]

Depth increases in the up direction



Make3D [Saxena et al. 2009]

Automatic methods

e.g. [Hoiem et al. 2005; Assa and Wolf 2007; Saxena et al. 2009]

Depth increases in the up direction

Color similarity implies depth similarity

Make3D [Saxena et al. 2009]

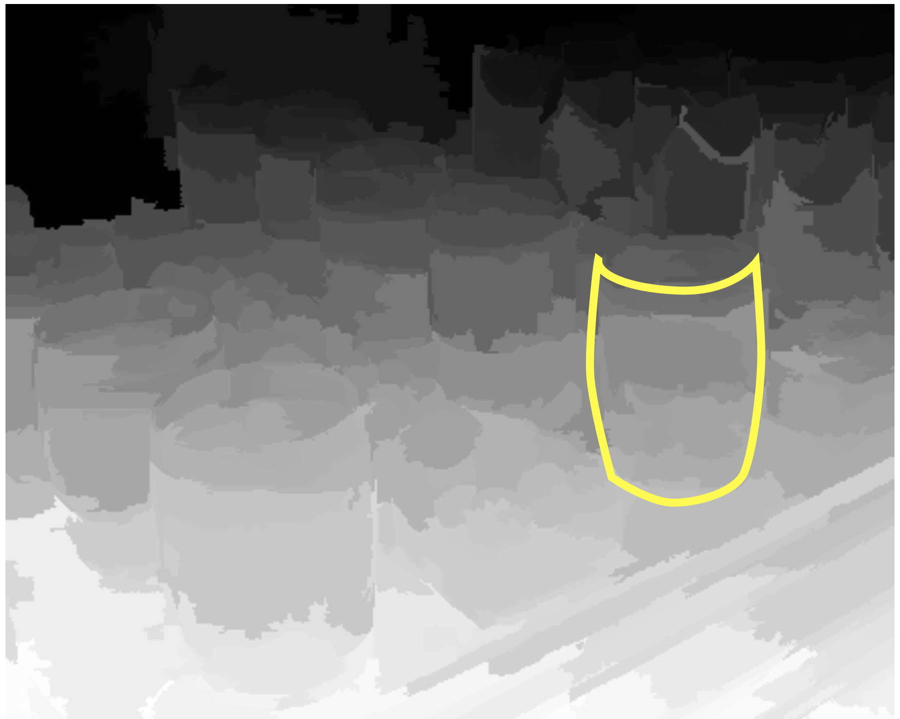
Automatic methods

e.g. [Hoiem et al. 2005; Assa and Wolf 2007; Saxena et al. 2009]

Depth increases in the up direction

Color similarity implies depth similarity

Not always correct



Make3D [Saxena et al. 2009]

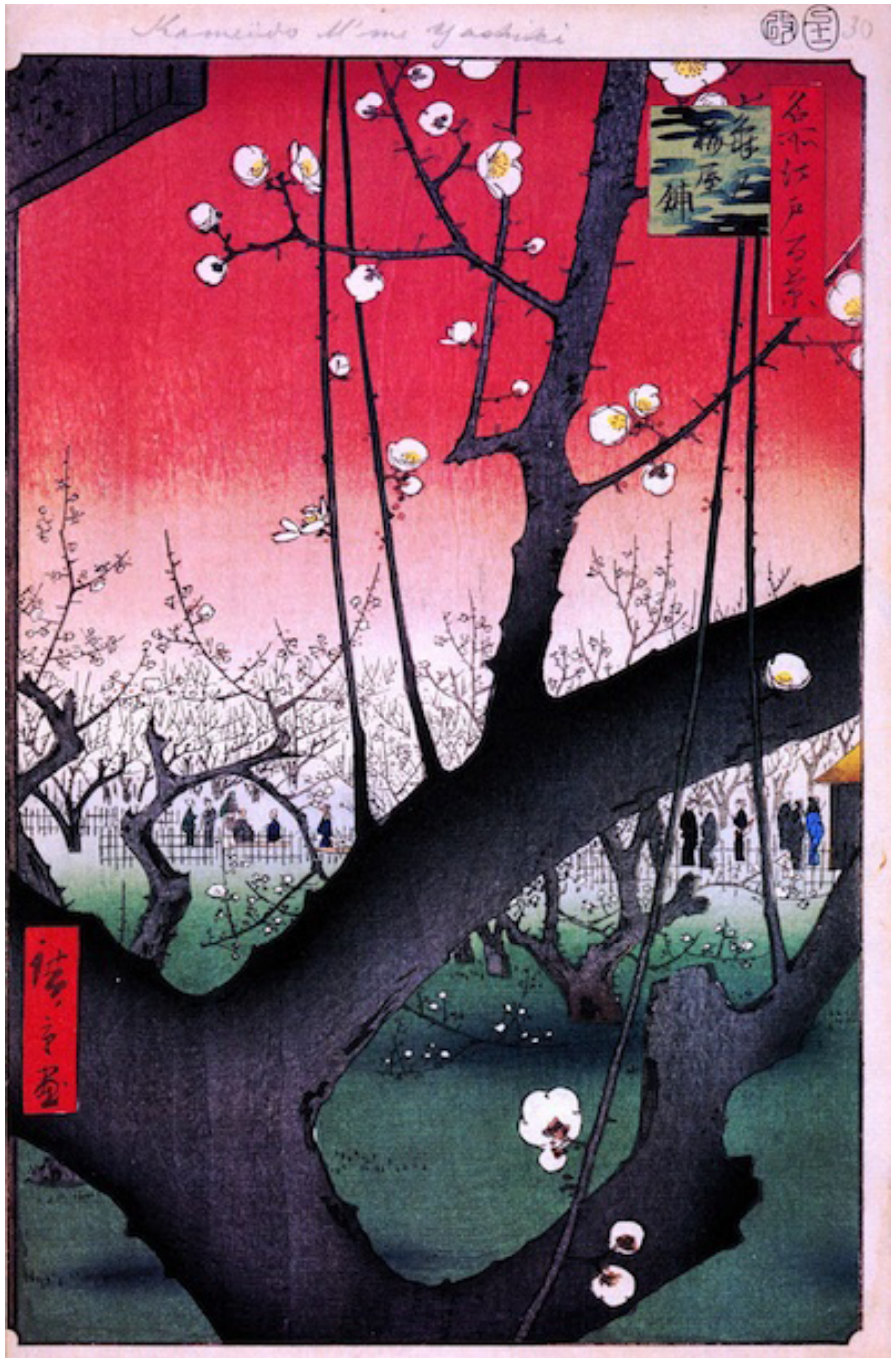
Automatic methods

Depth increases in the up direction

Color similarity implies depth similarity

Not always correct

Some images are very challenging (art)



[Hiroshige]

Micro-Task

Micro-Task

Ask “What is the depth of a pixel?”

Micro-Task

Ask "What is the depth of a pixel?"

- Too fine, can be ambiguous

Micro-Task

Ask "What is the depth of a pixel?"

- Too fine, can be ambiguous

Ask "What is the depth of an object?"

Micro-Task

Ask "What is the depth of a pixel?"

- Too fine, can be ambiguous

Ask "What is the depth of an object?"

- Segmentation is too difficult

Micro-Task

Ask "What is the depth of a pixel?"

- Too fine, can be ambiguous

Ask "What is the depth of an object?"

- Segmentation is too difficult

Ask "What is the depth of a patch in the image?"

Micro-Task

Ask “What is the depth of a pixel?”

- Too fine, can be ambiguous

Ask “What is the depth of an object?”

- Segmentation is too difficult

Ask “What is the depth of a patch in the image?”

- Getting better… but humans are not good at assessing absolute depth

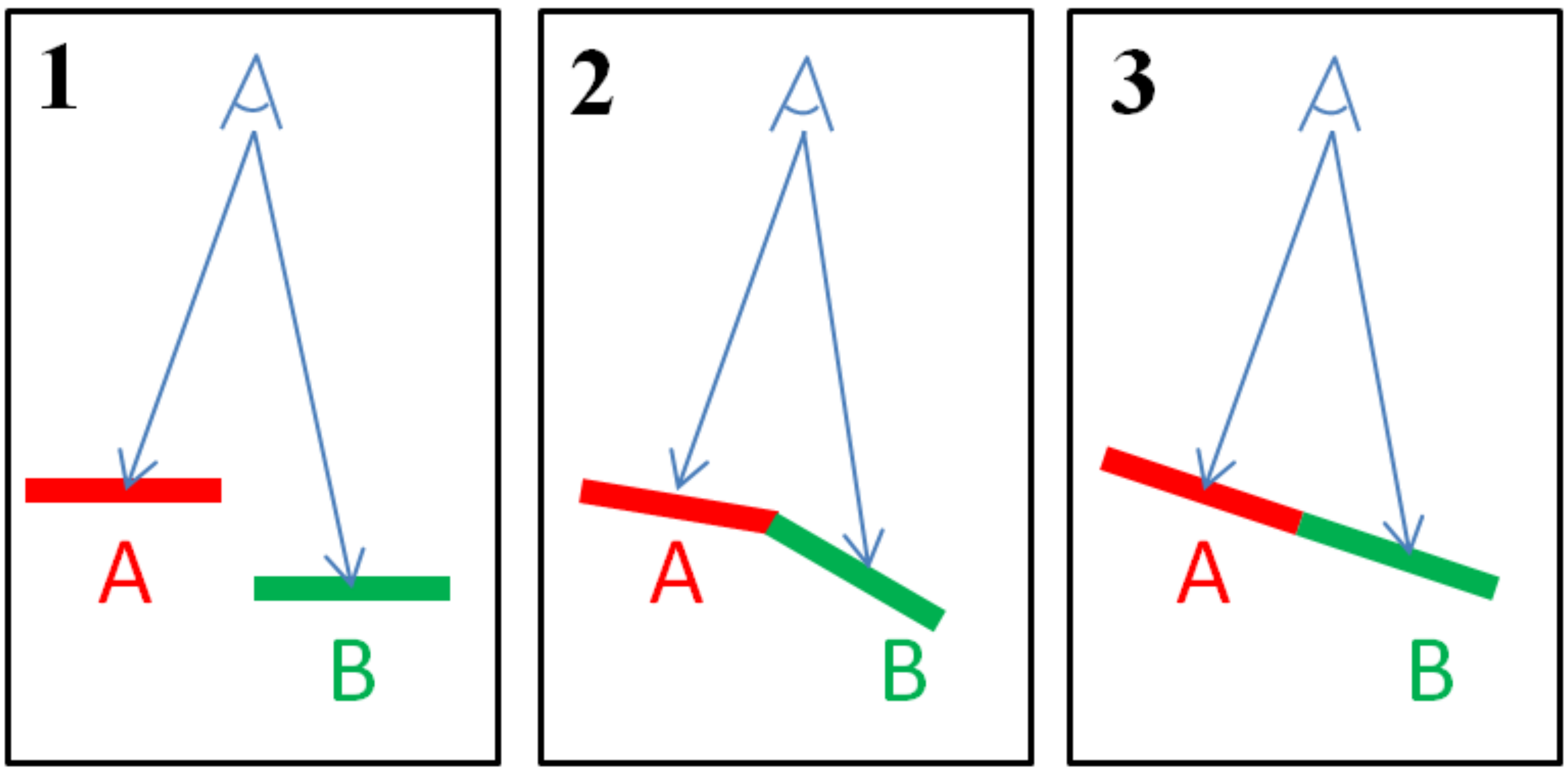
Relative Ordering

Ask “Which is closer?” on neighboring patches

Relative Ordering

Ask “Which is closer?” on neighboring patches

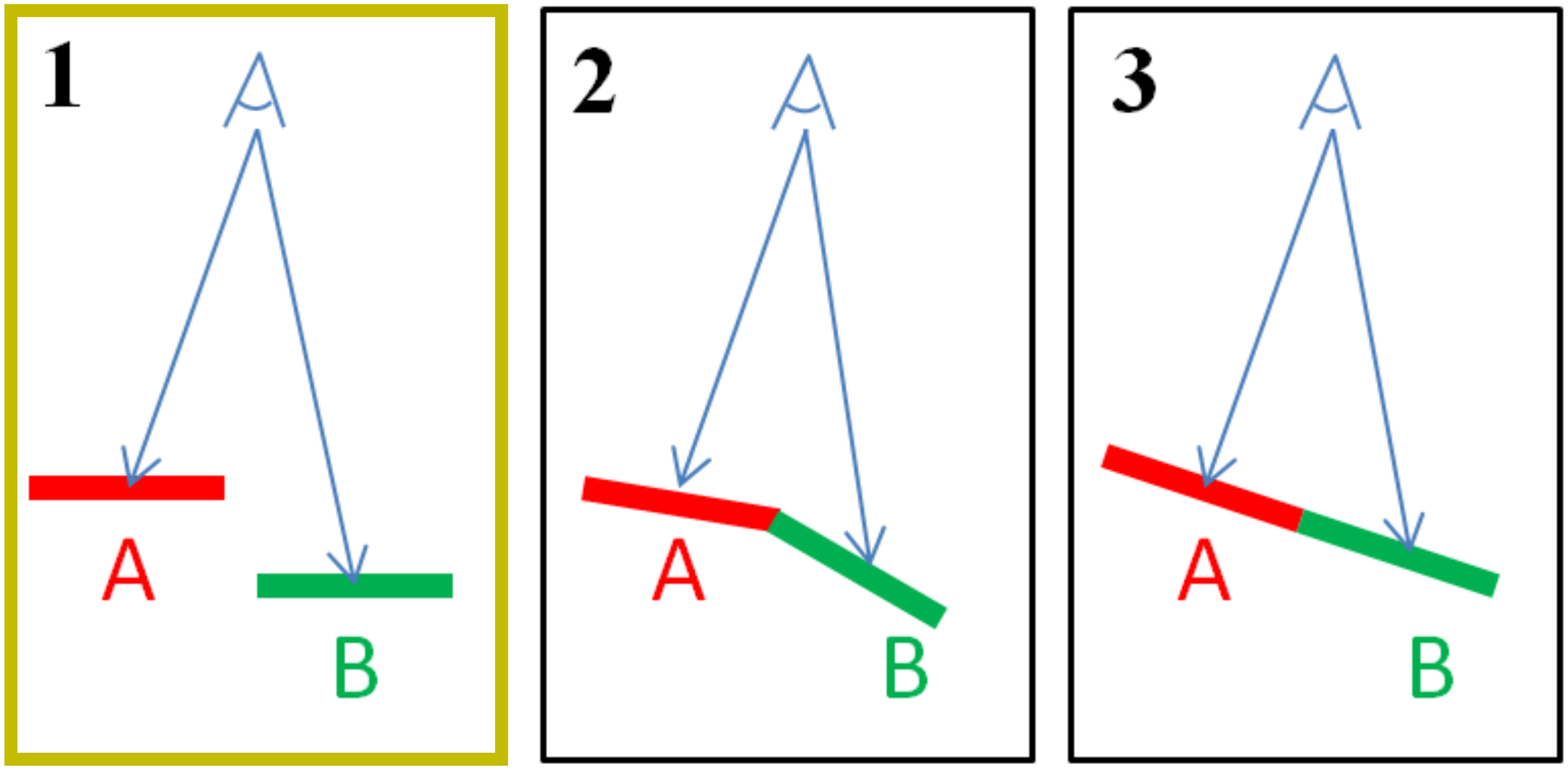
- Reliable, but not well-defined. A is closer than B:



Relative Ordering

Ask “Which is closer?” on neighboring patches

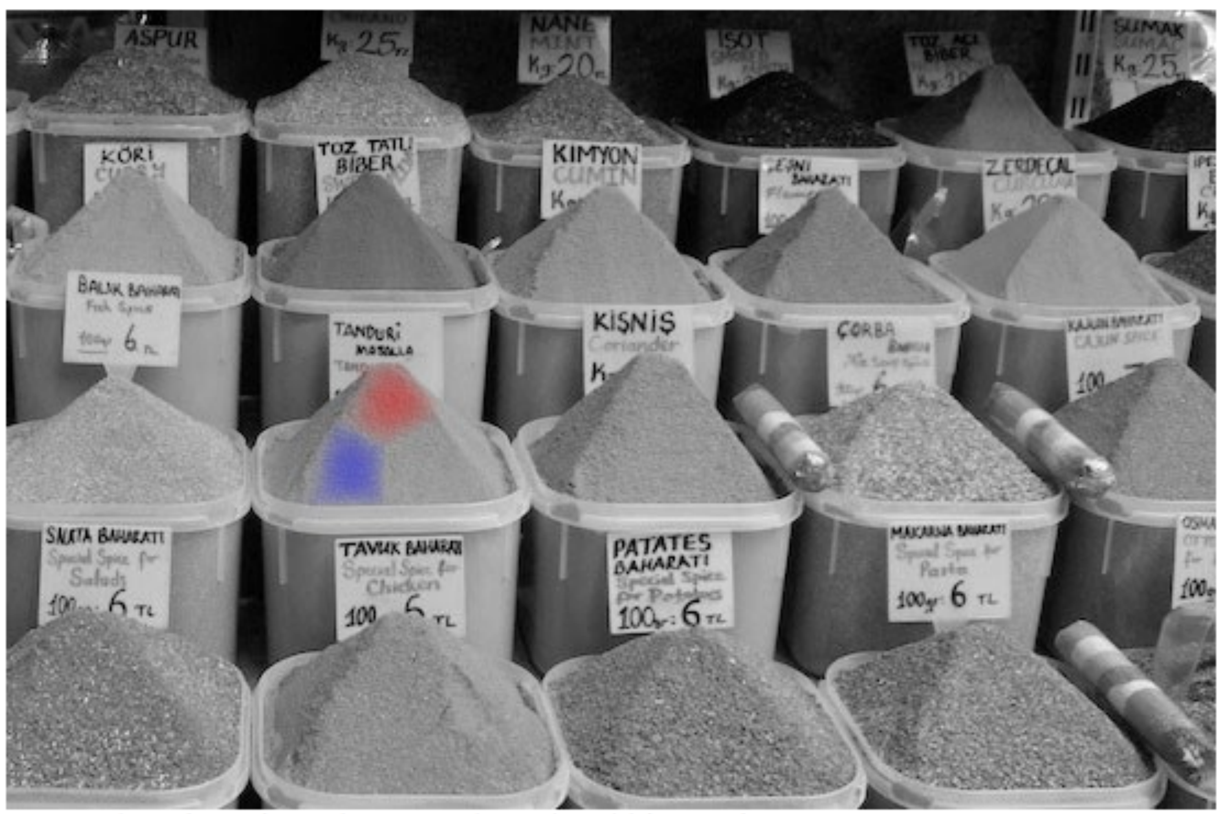
- Reliable, but not well-defined. A is closer than B:



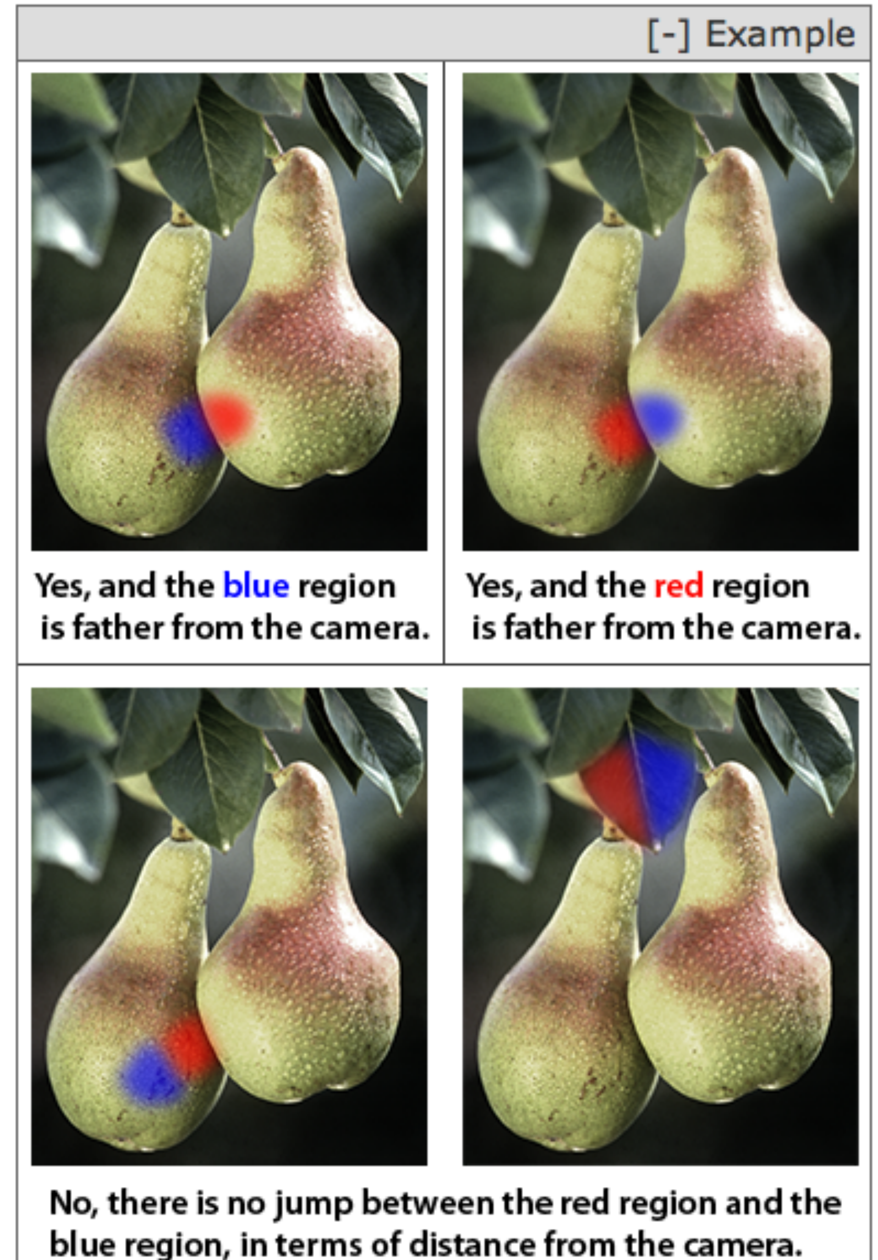
Our Micro-Task

Is there a jump between the red region and the blue region, in terms of distance from the camera?

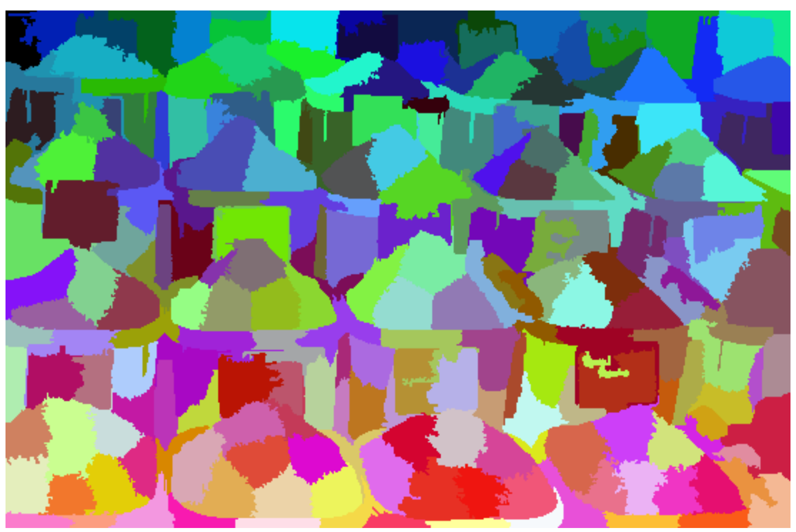
Place the mouse over an image to hide the highlighted regions.



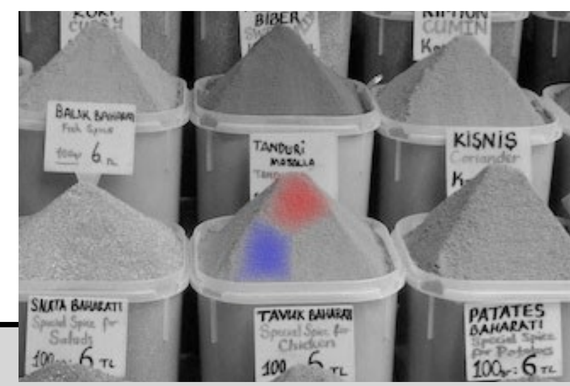
○ **No**, there is no jump between the red and blue regions.
○ Yes, and the **blue** region is farther from the camera.
○ Yes, and the **red** region is farther from the camera.



Algorithm



Patch Segmentation



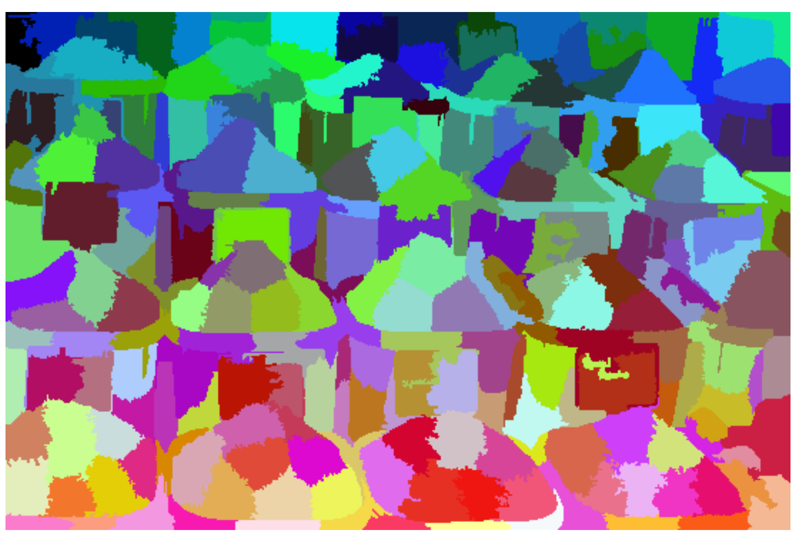
Micro-tasks (human)

Laplace equation

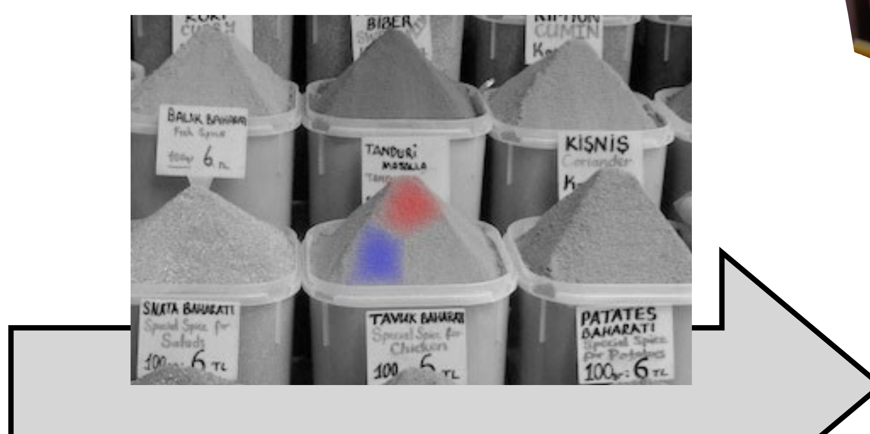
$$\Delta f = 0$$

with constraints

Algorithm



Patch Segmentation



Micro-tasks (human)

Laplace equation

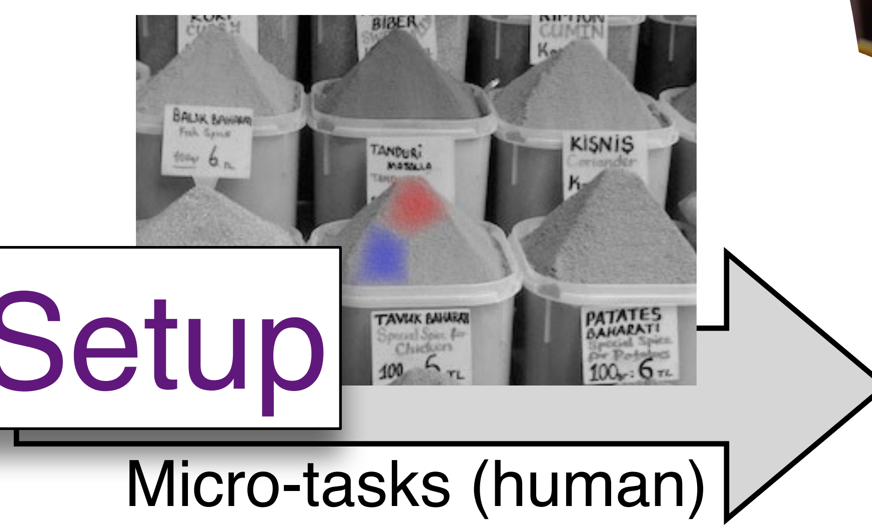
$$\Delta f = 0$$

with constraints

Algorithm

Quality Control Setup

Patch
Segmentation

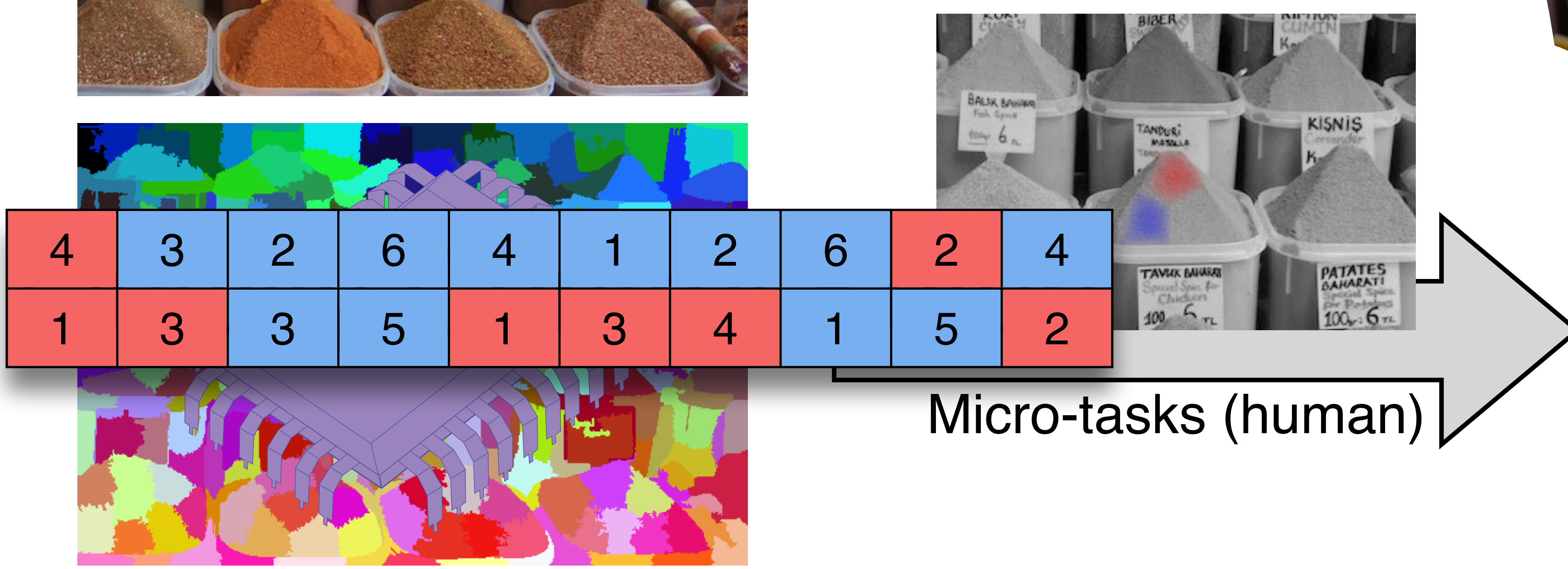


Laplace equation

$$\Delta f = 0$$

with constraints

Algorithm



Patch Segmentation

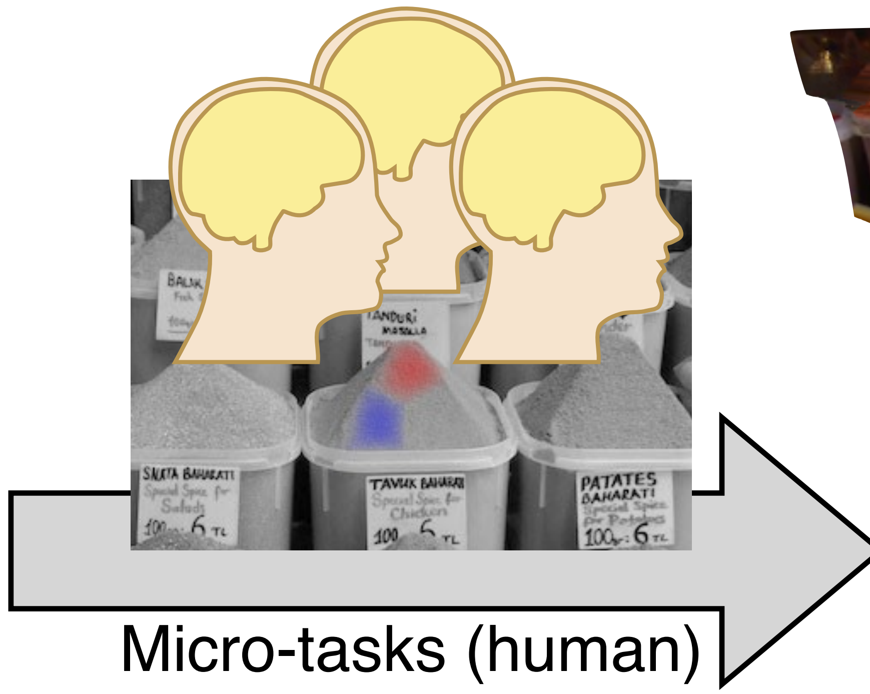
Laplace equation

$$\Delta f = 0$$

with constraints

Algorithm

Patch
Segmentation

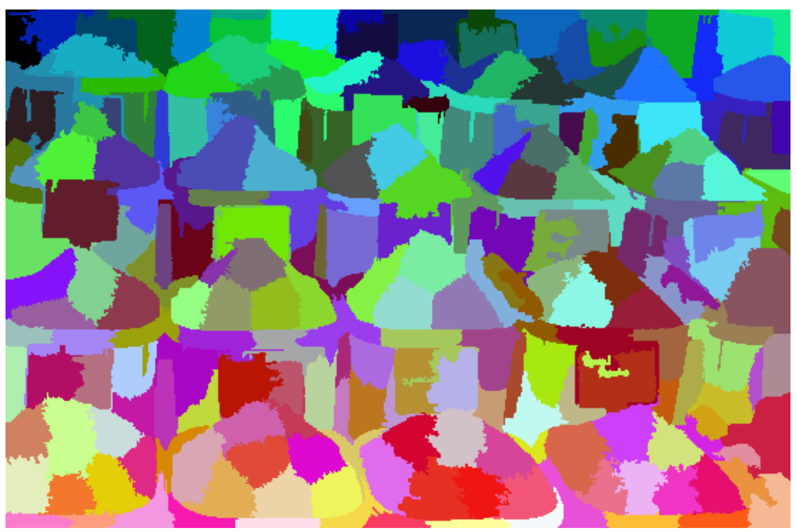


Laplace equation

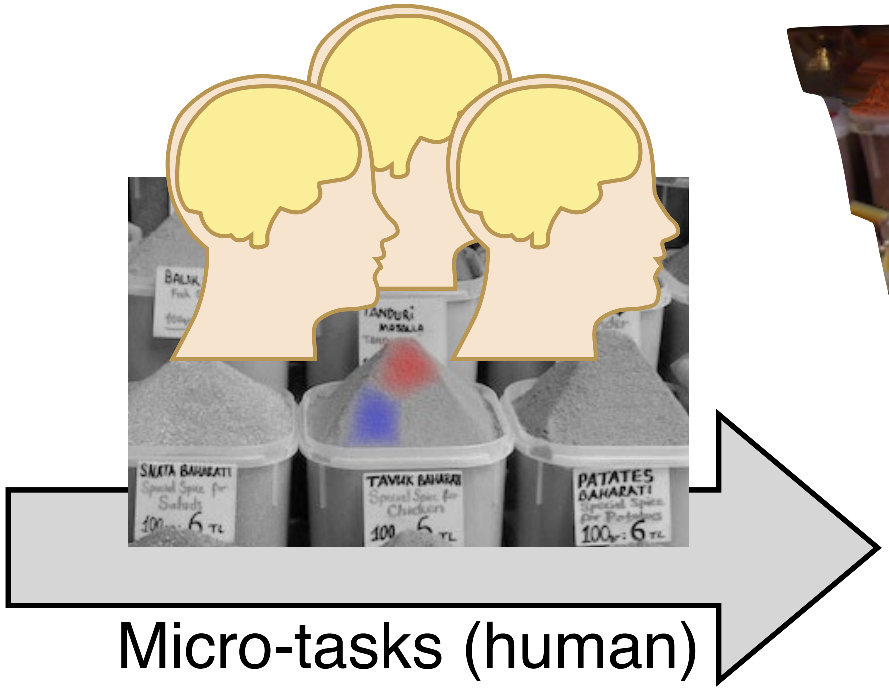
$$\Delta f = 0$$

with constraints

Algorithm

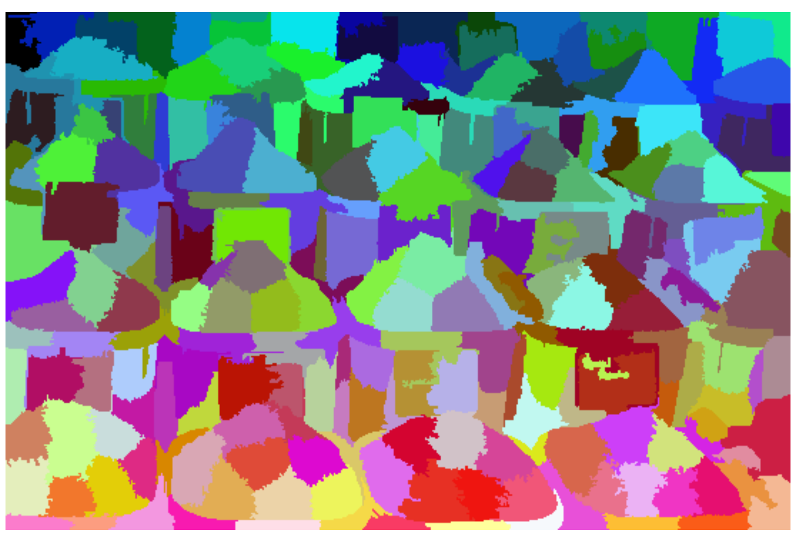


Patch
Segmentation

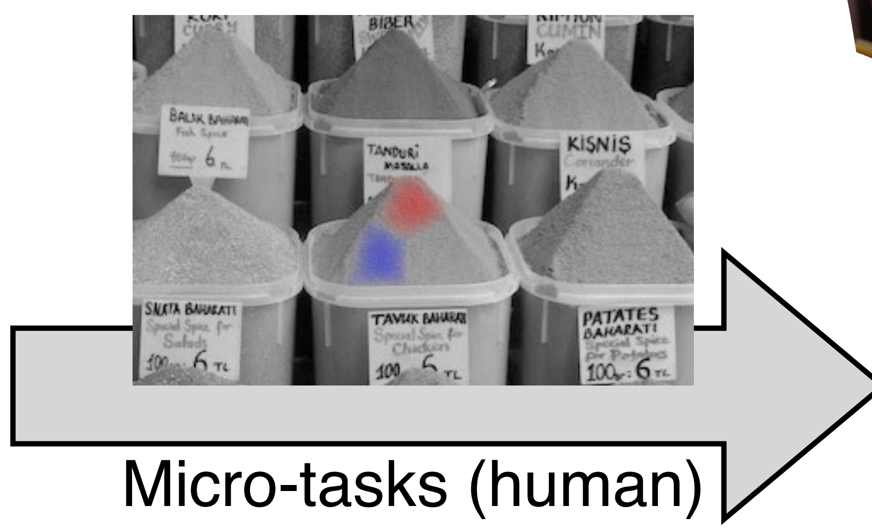


Laplace equation
$$\Delta f = 0$$
with constraints

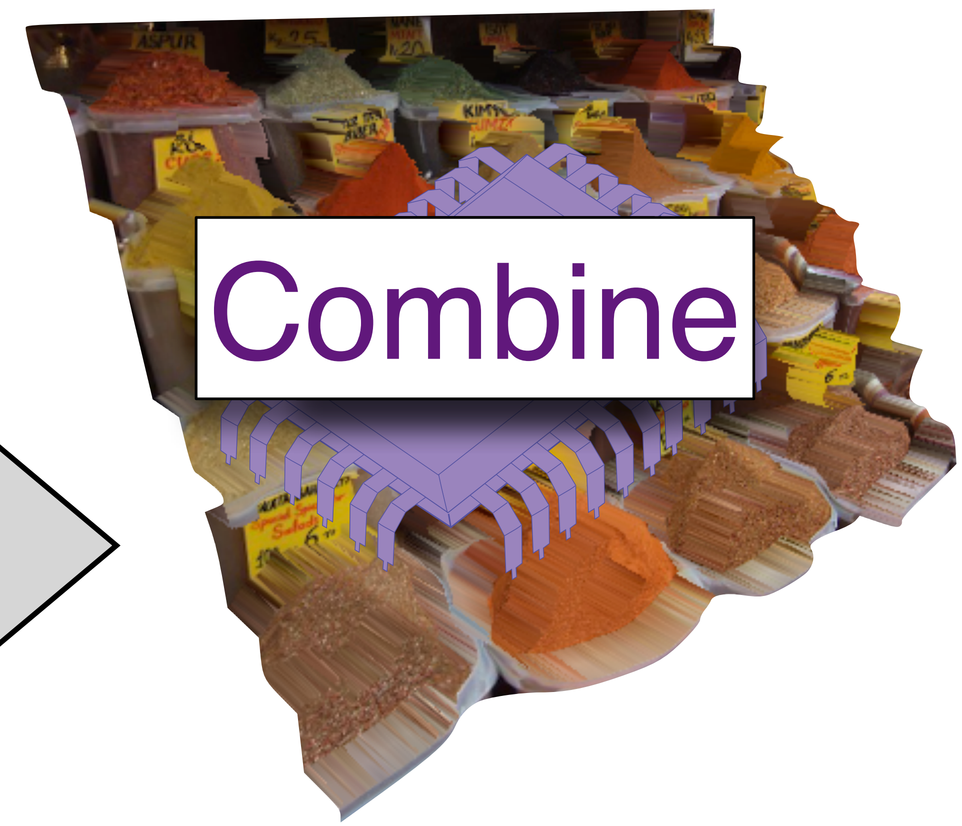
Algorithm



Patch
Segmentation



Micro-tasks (human)

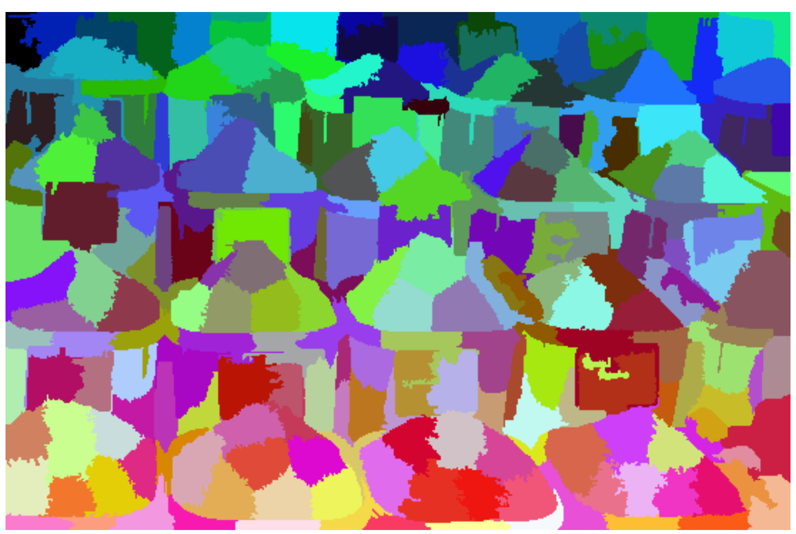


Laplace equation

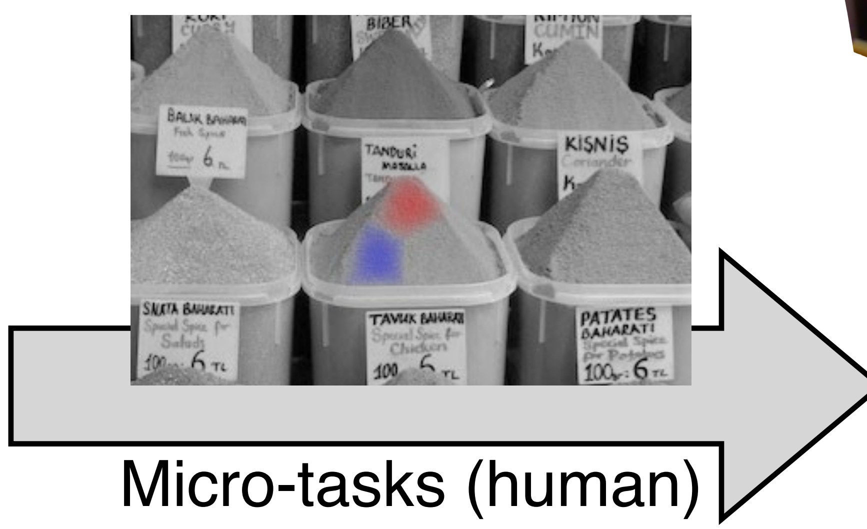
$$\Delta f = 0$$

with constraints

Algorithm



Patch Segmentation



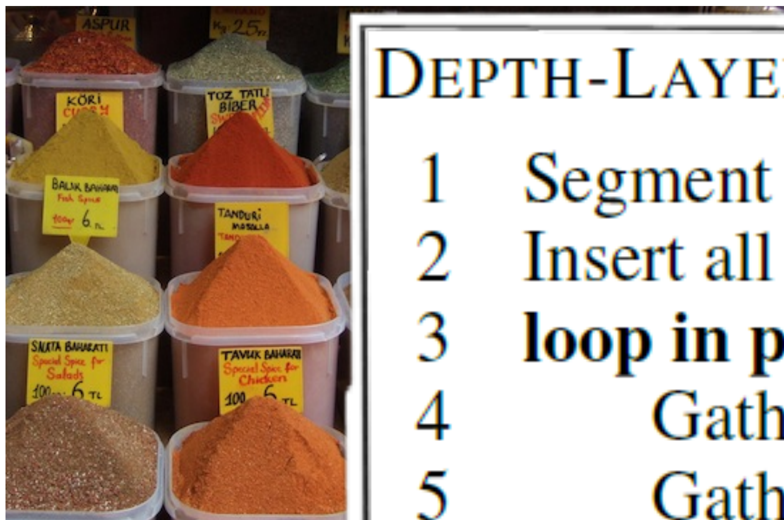
Micro-tasks (human)

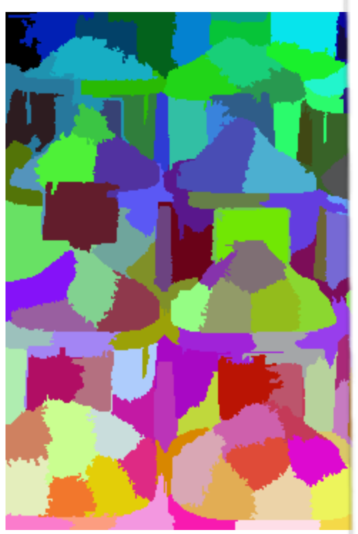
Laplace equation

$$\Delta f = 0$$

with constraints

Algorithm





DEPTH-LAYERS(image I, sentinel queries S)

```
Segment I into regions (using mean-shift and SLIC)
Insert all pairs of neighboring regions into Q
loop in parallel until each pair has been visited N times
    Gather K random pairs from Q
    Gather M random pairs from S
    for each pair: Build the visual query & Duplicate it
    Mix the 2K + 2M queries
    results = send all queries to an HP
    if average(consistent(results)) ≥ 0.75
        for each pair
            Add consistent results to the list of votes
            Increment #visited
for each pair of neighboring regions
    final_result = majority(list of votes)
Solve the Laplace equation to construct a depth map
```

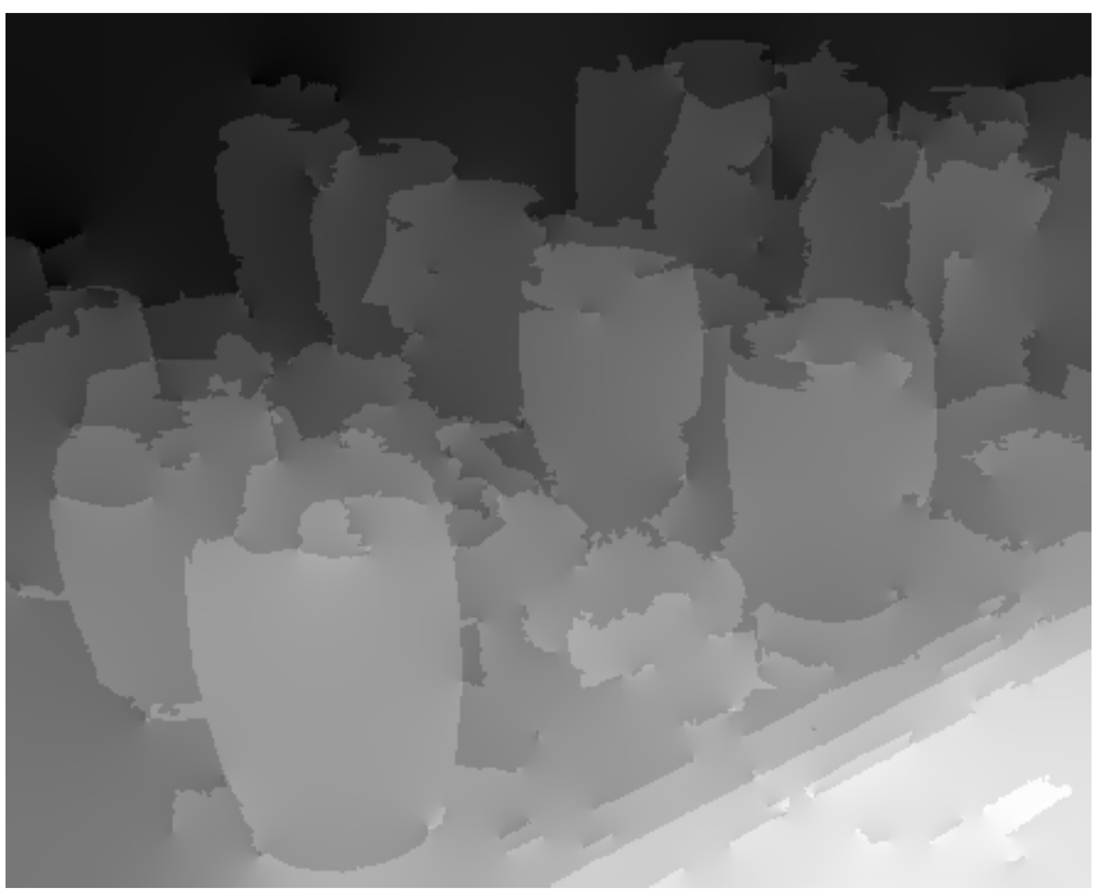
Pa

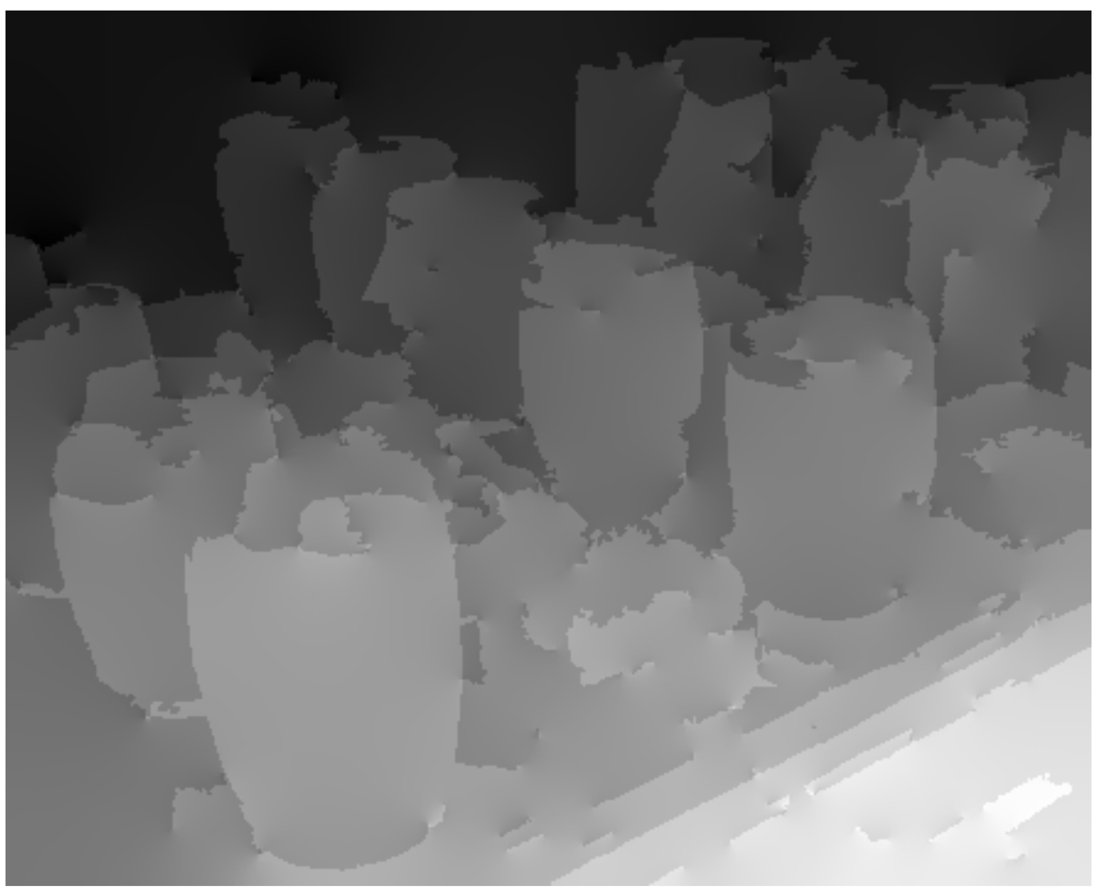
Segm

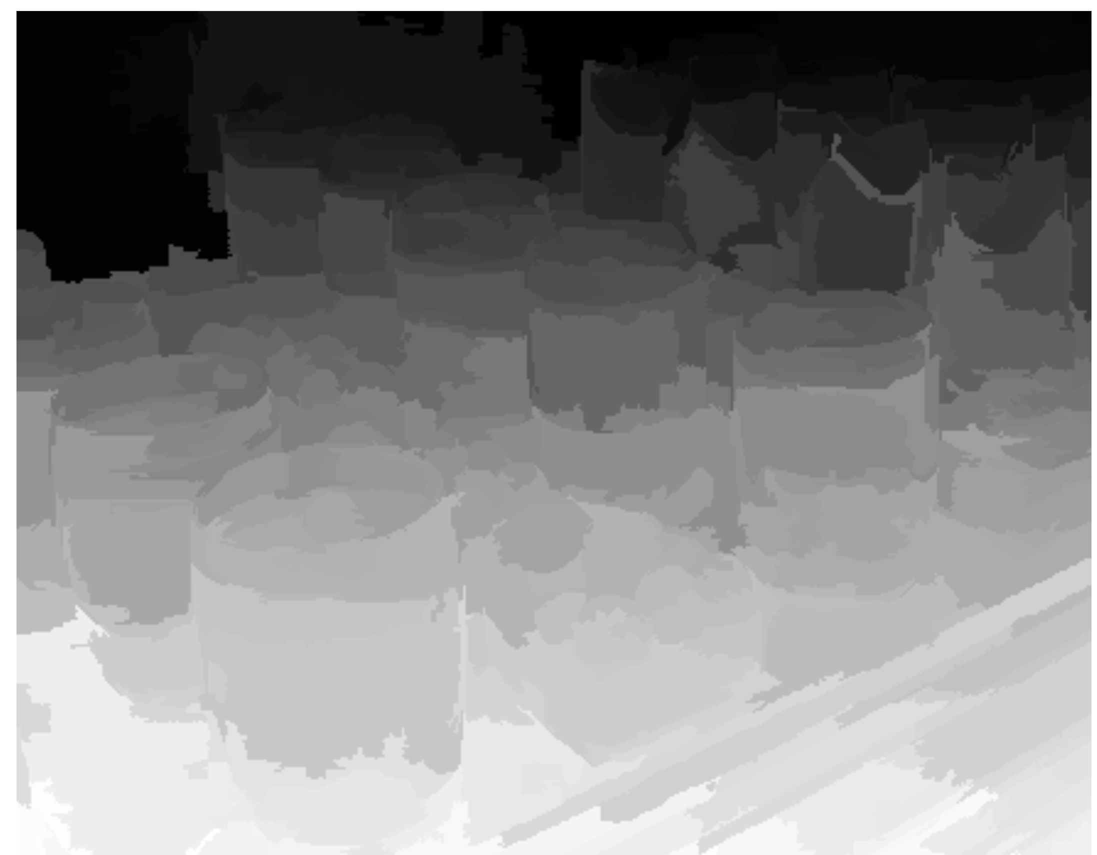
equation

$= 0$

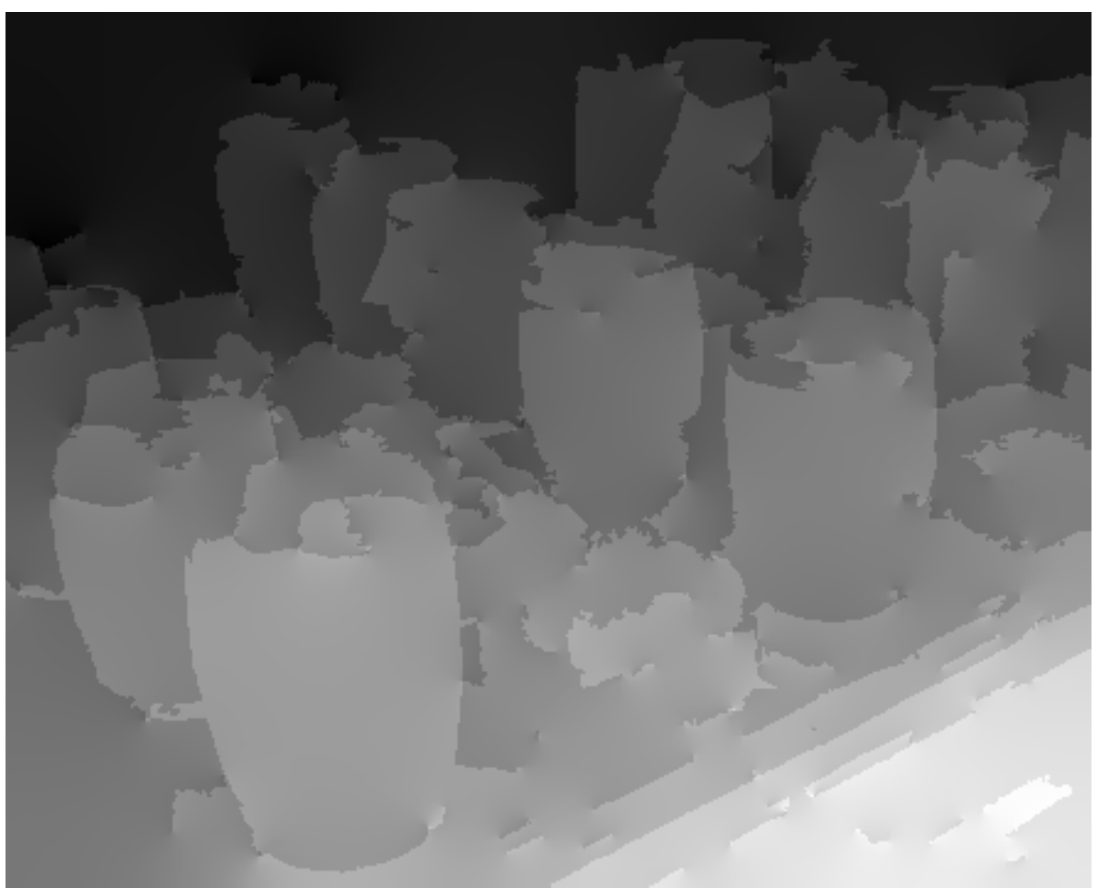
constraints



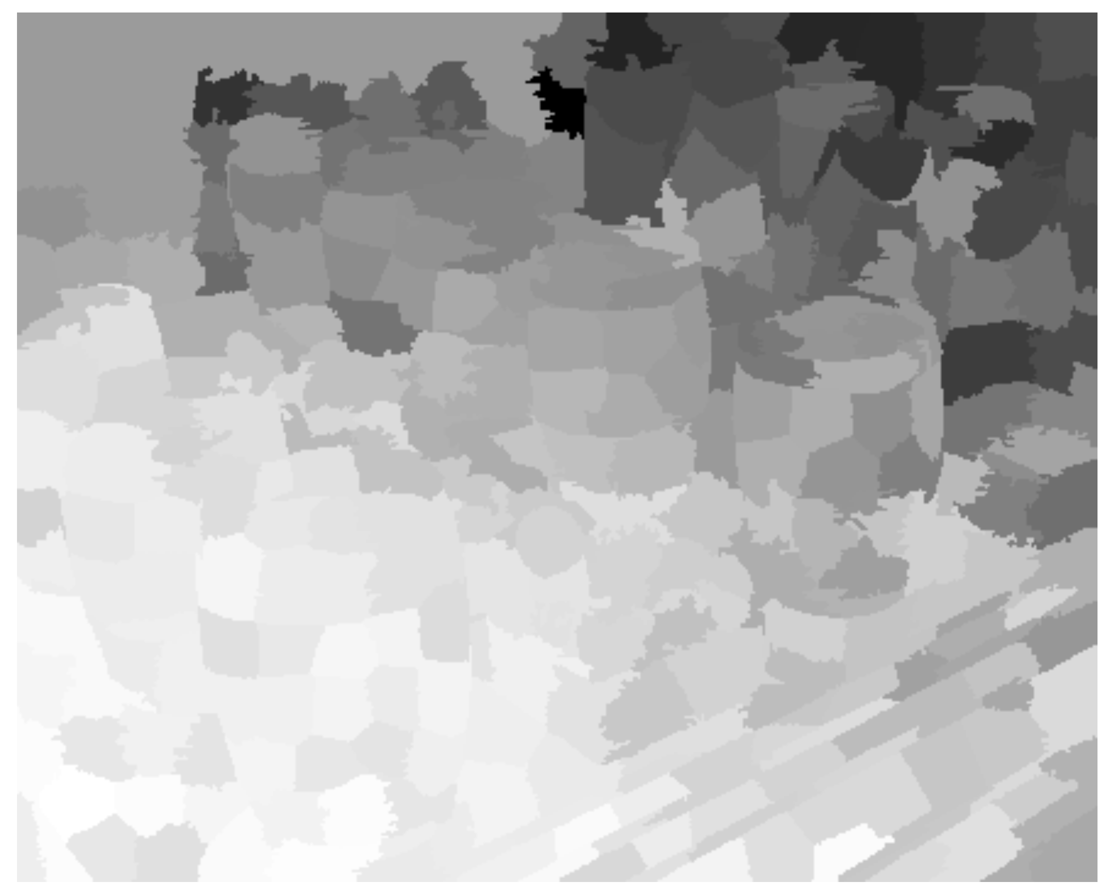




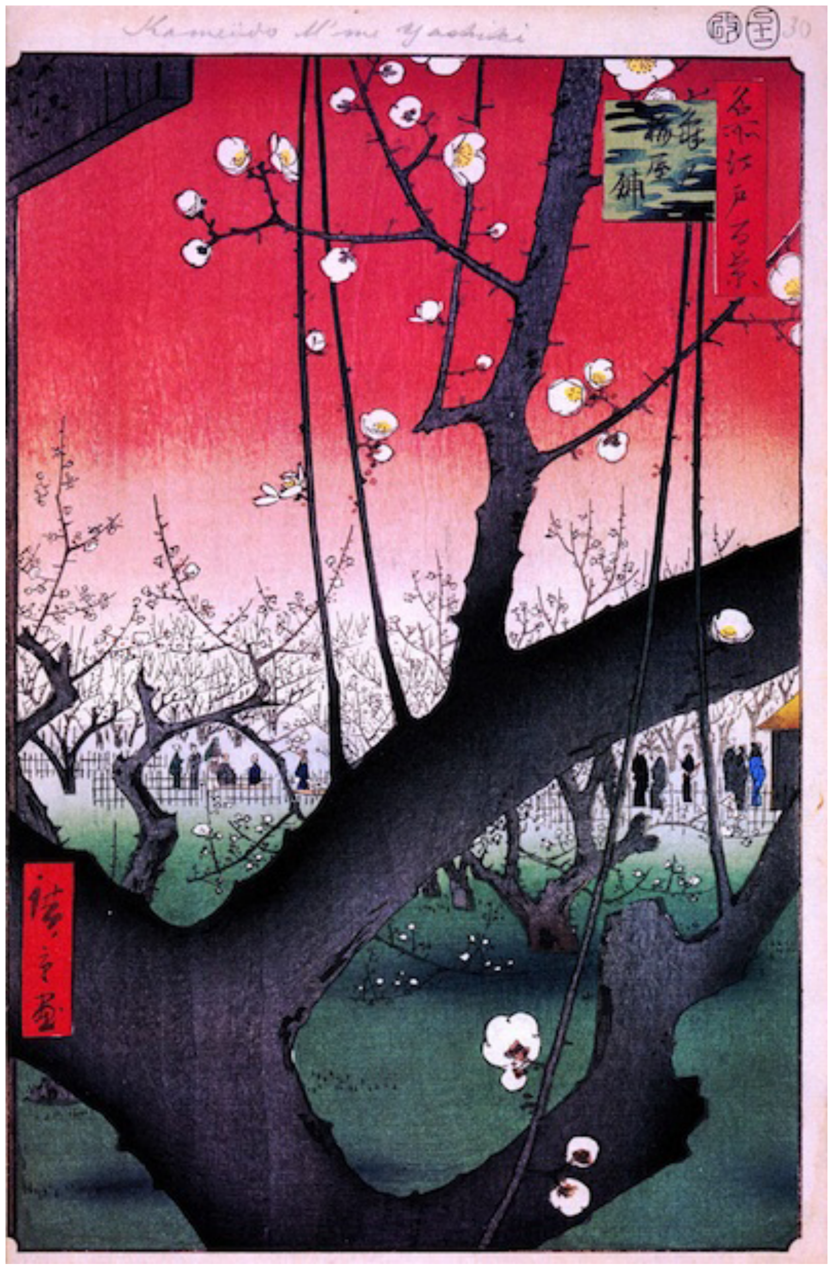
Automatic (Make3D)

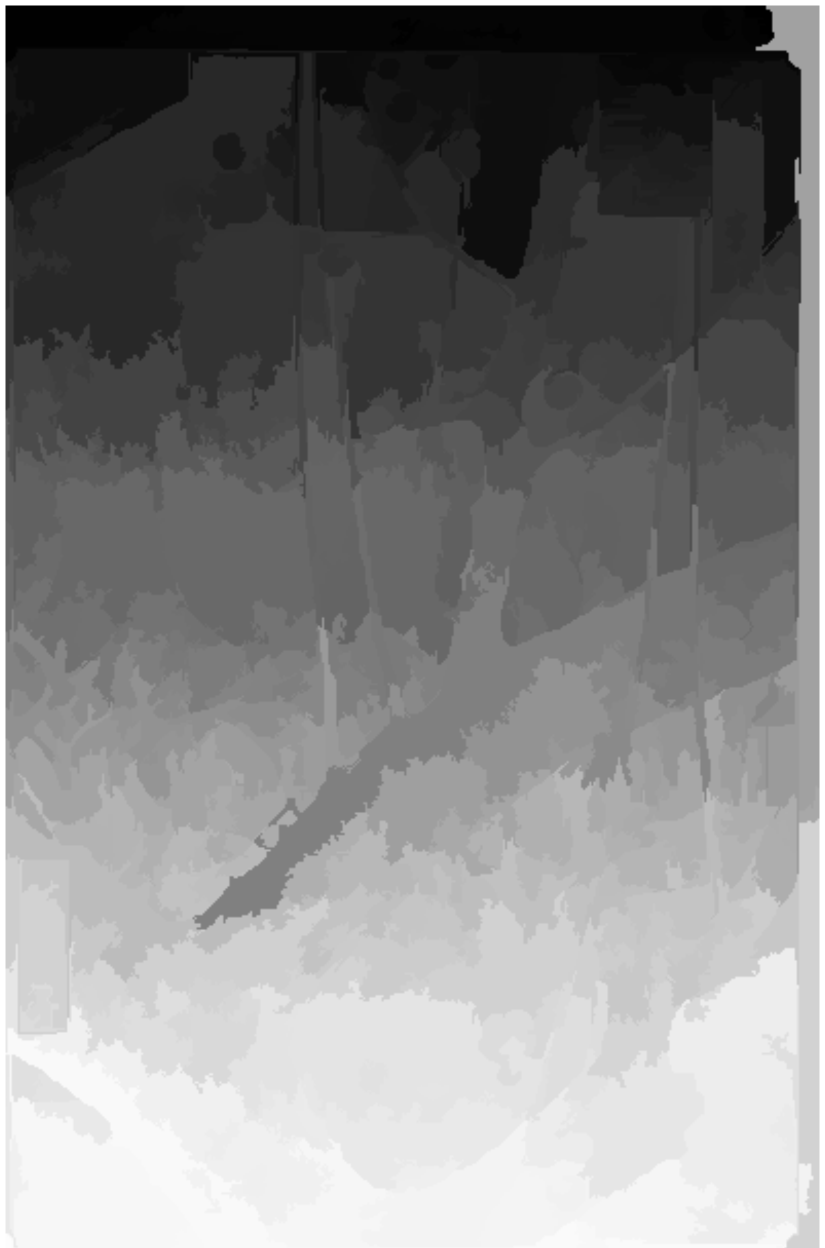


discrete depth

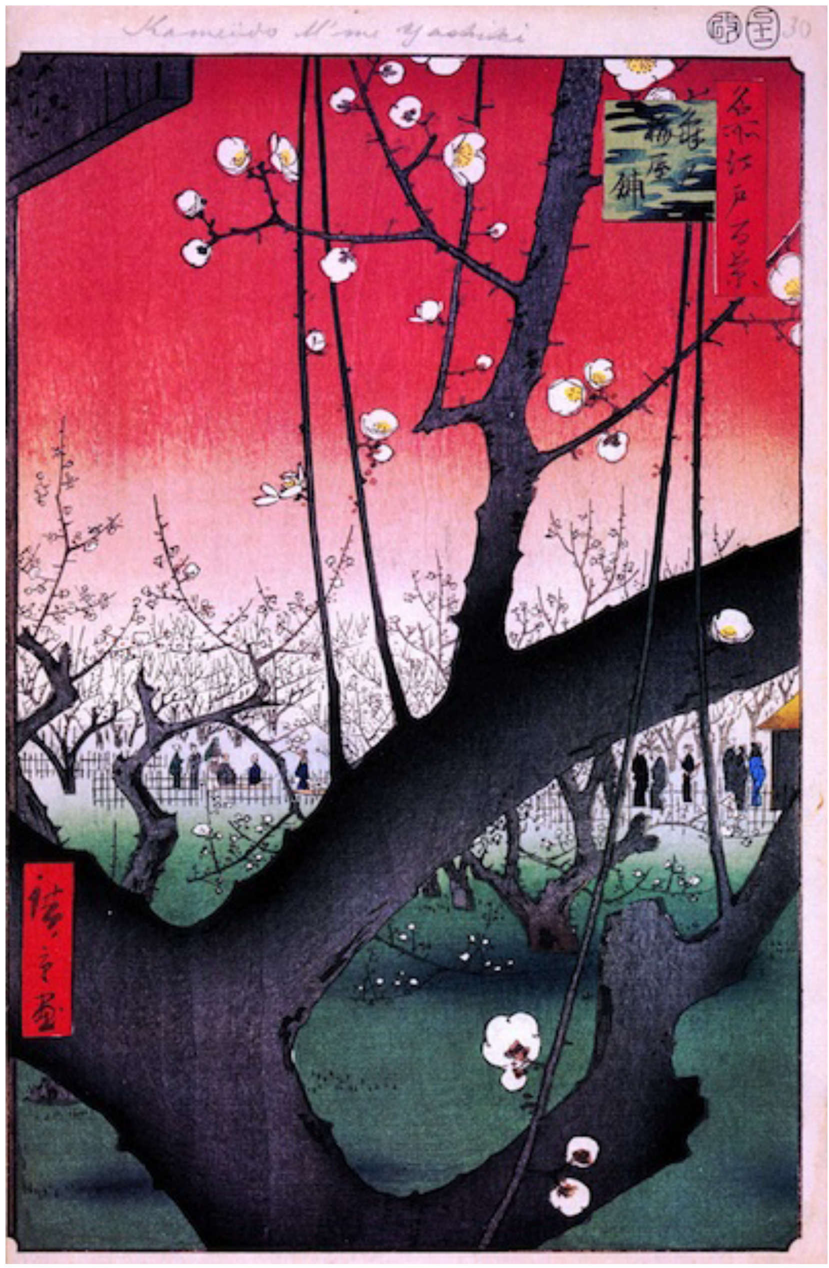


absolute depth





Automatic (Make3D)

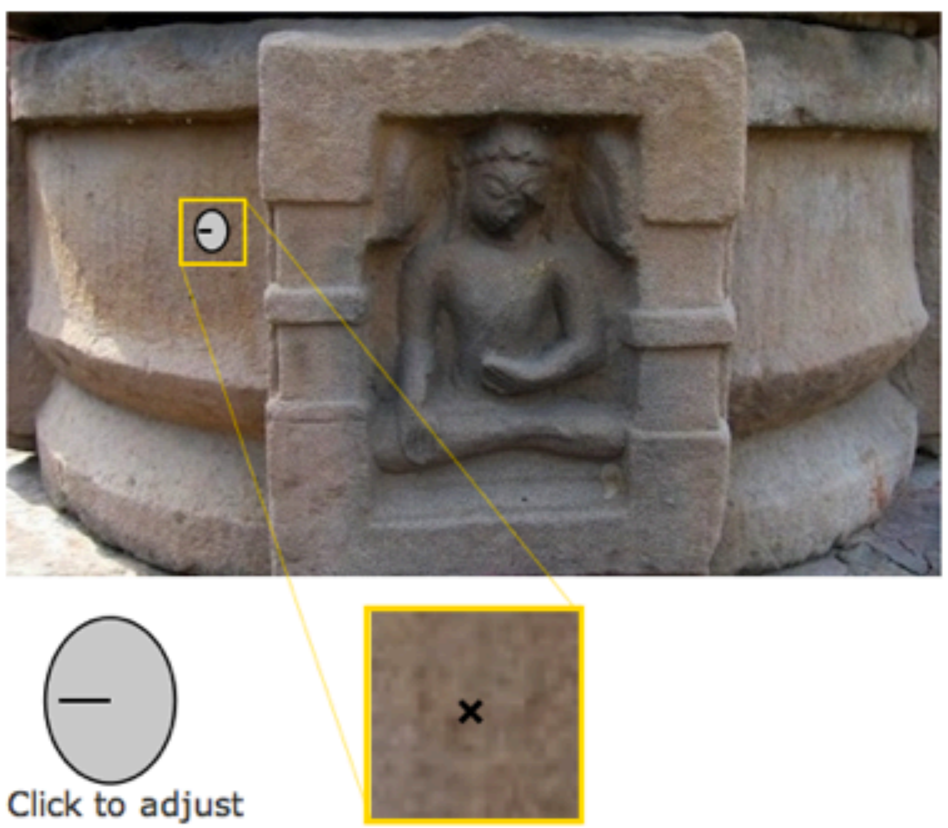


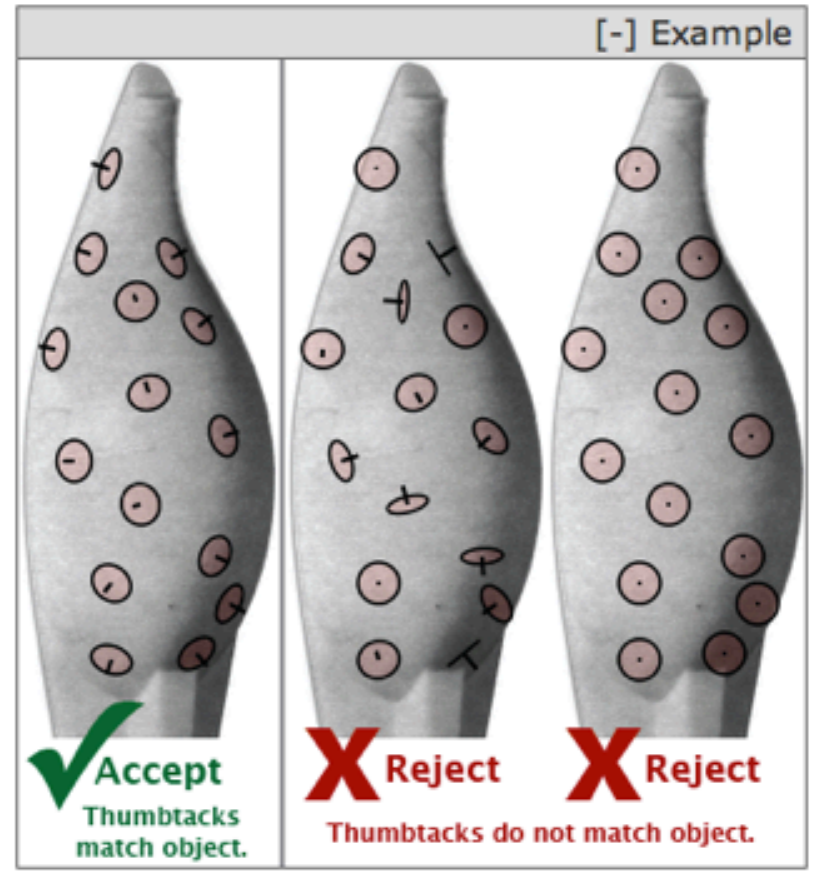
Algorithm 2: Normal Map

Orient the thumbtacks flush against the surface.

The thumbtack's pin should point away from the surface behind it. See the Example for good and bad examples.

Thumbtacks may appear at the same location multiple times. We check for consistency and may reject inconsistent HITs.

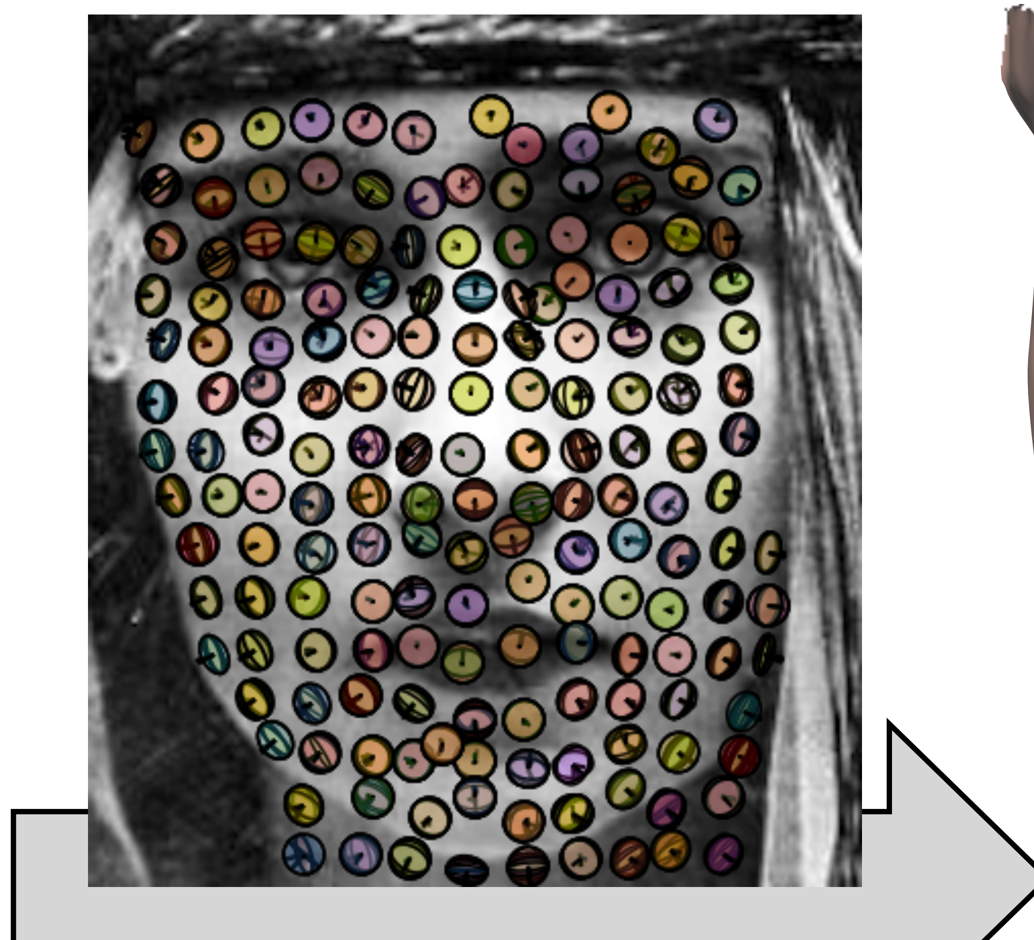




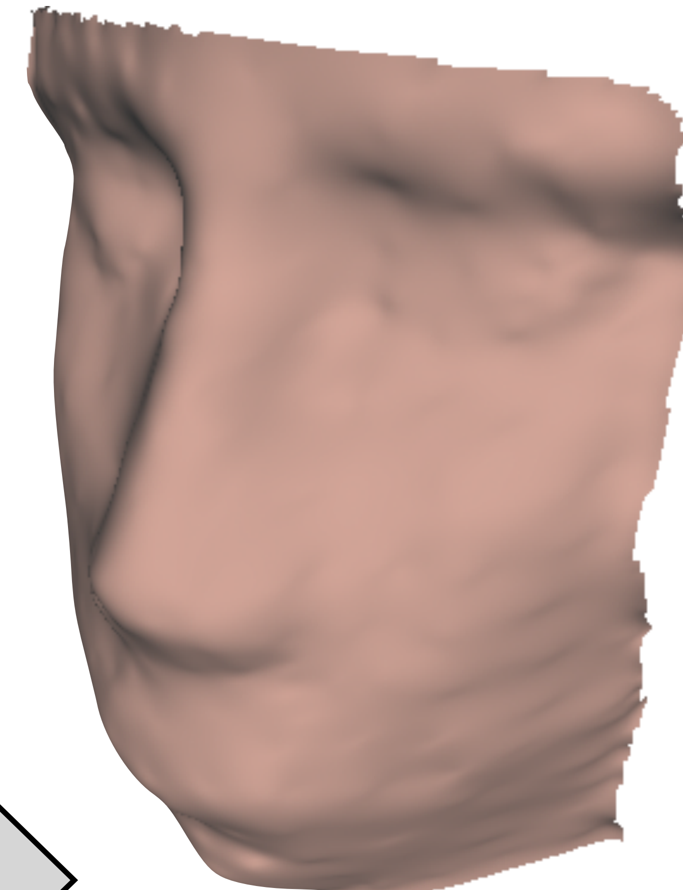
Hide thumbtack

Algorithm

Select normal locations



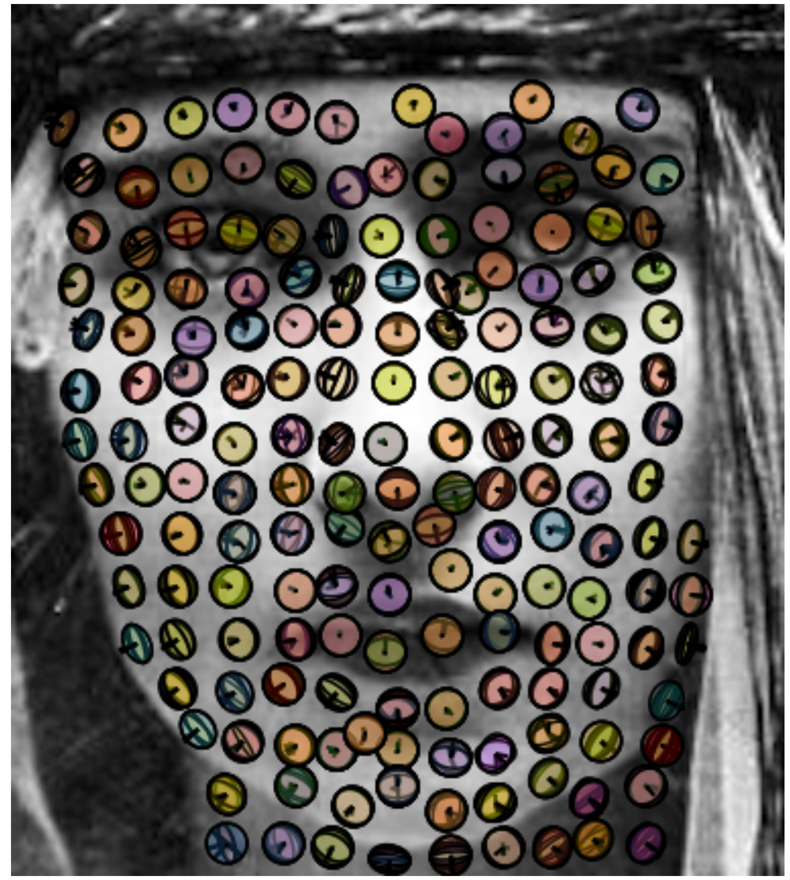
Micro-tasks (human)

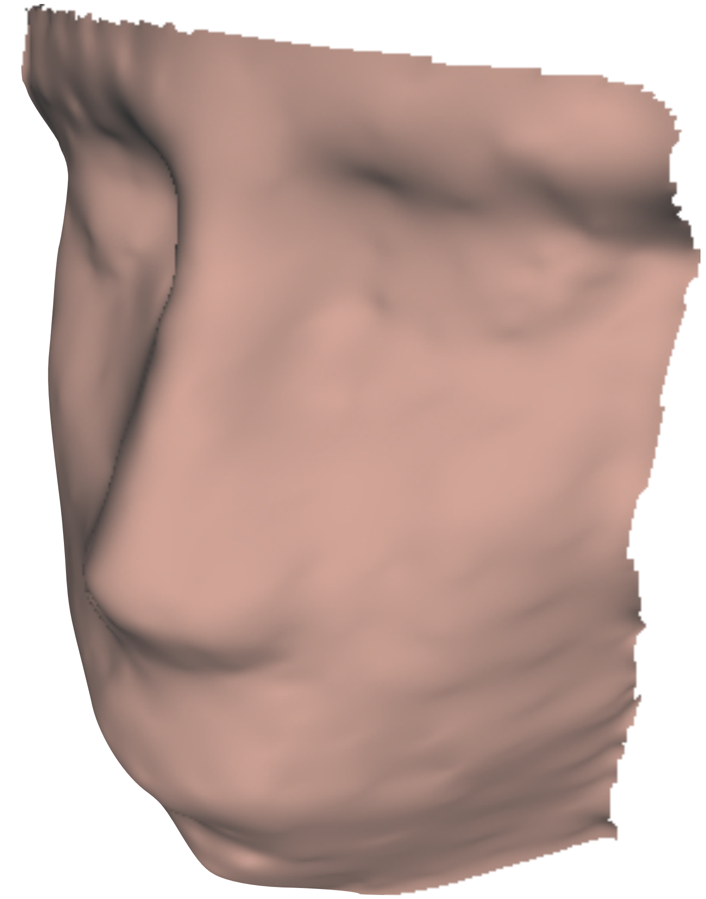


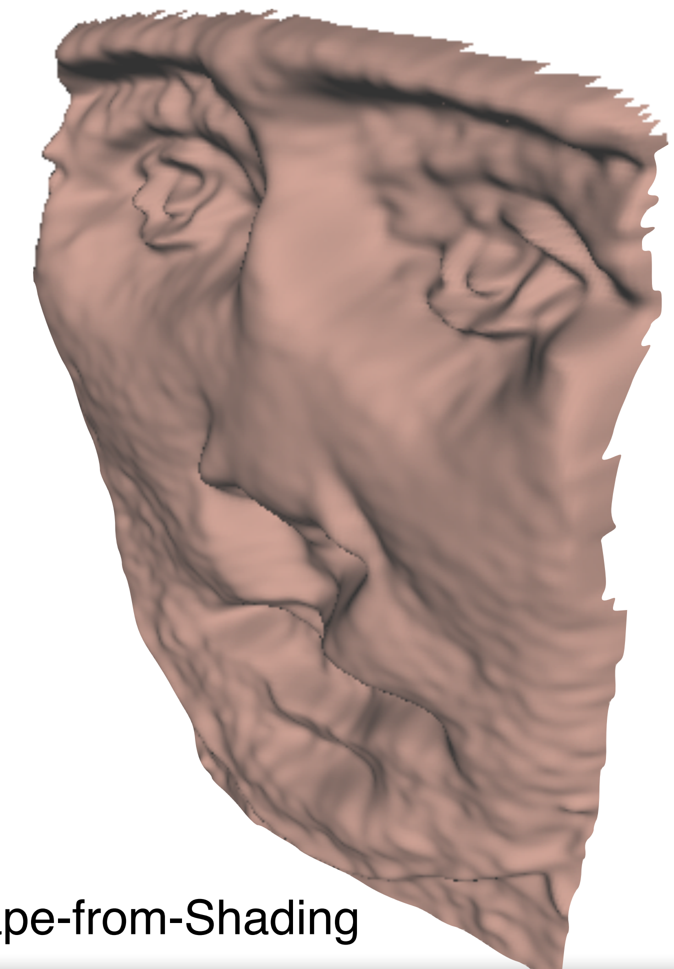
Bi-Laplace equation

$$\Delta^2 f = 0$$

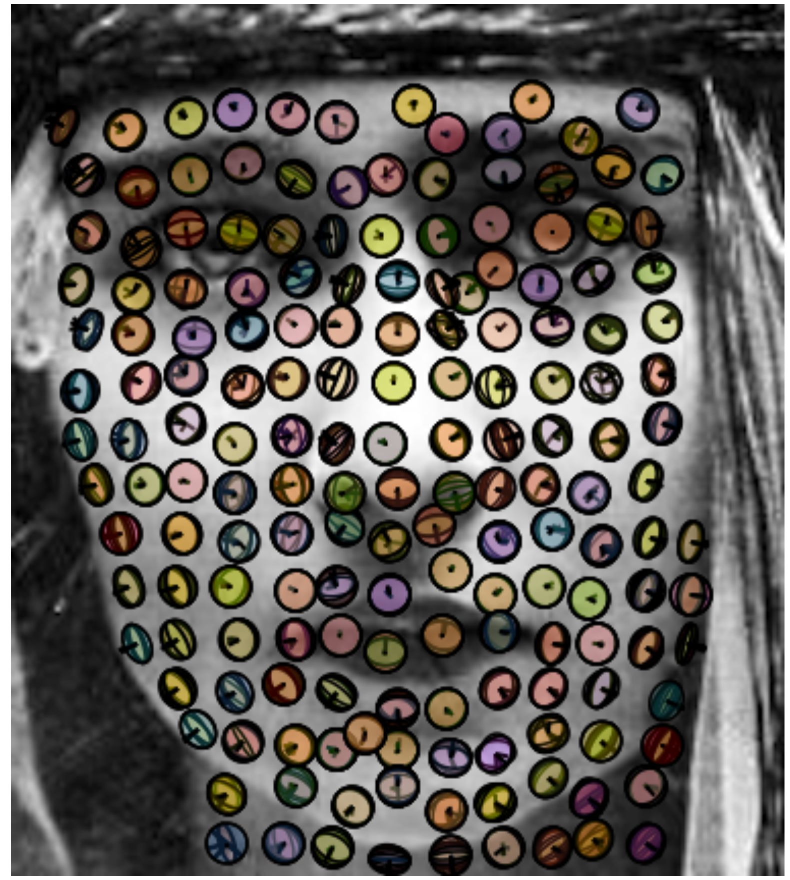
with constraints

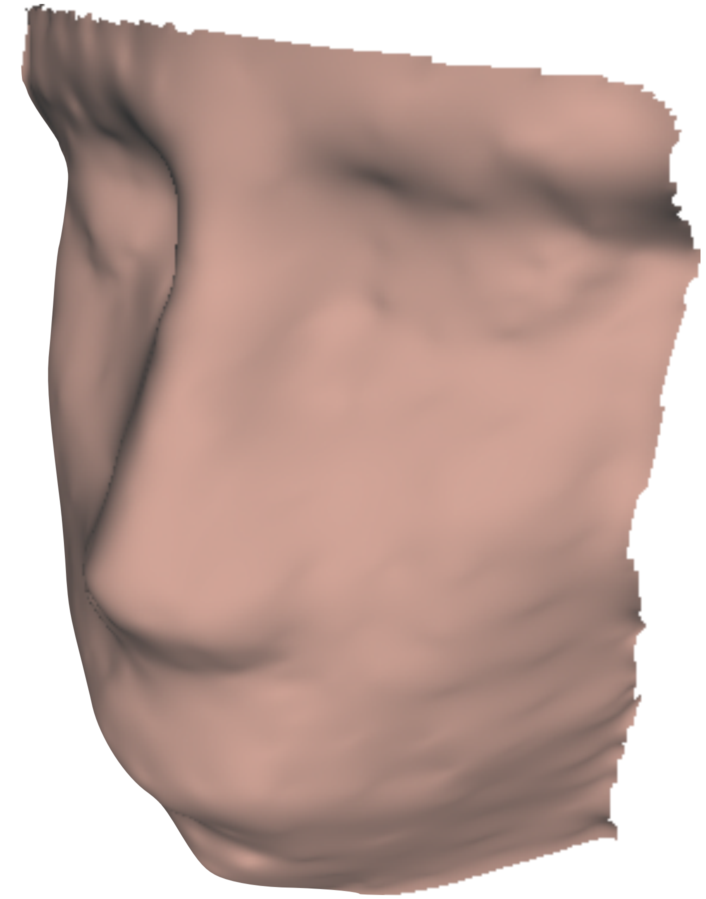


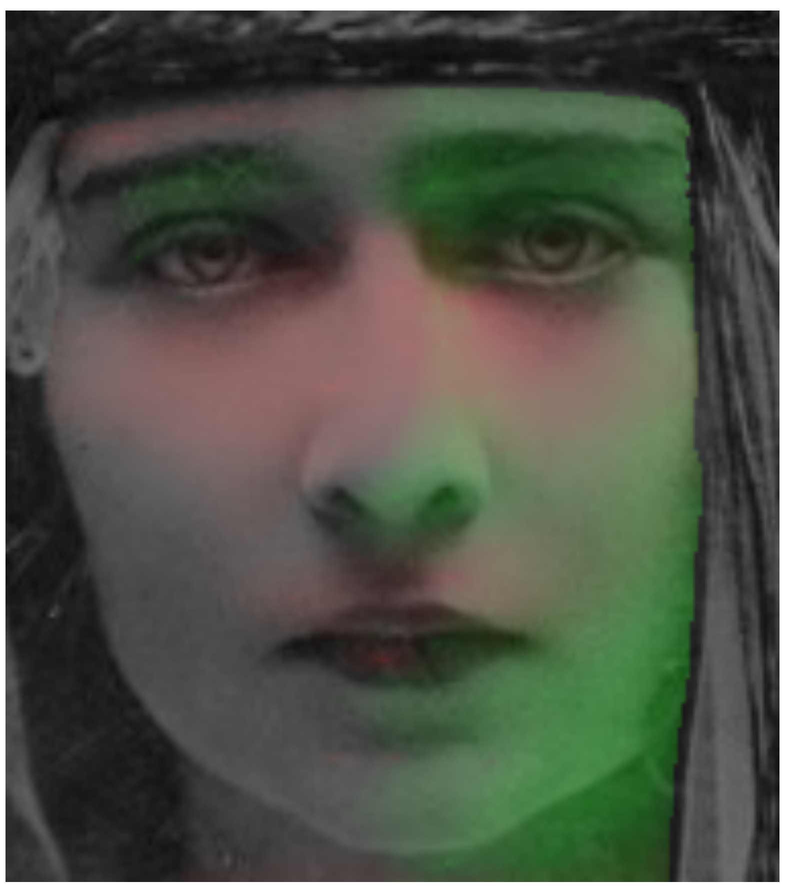




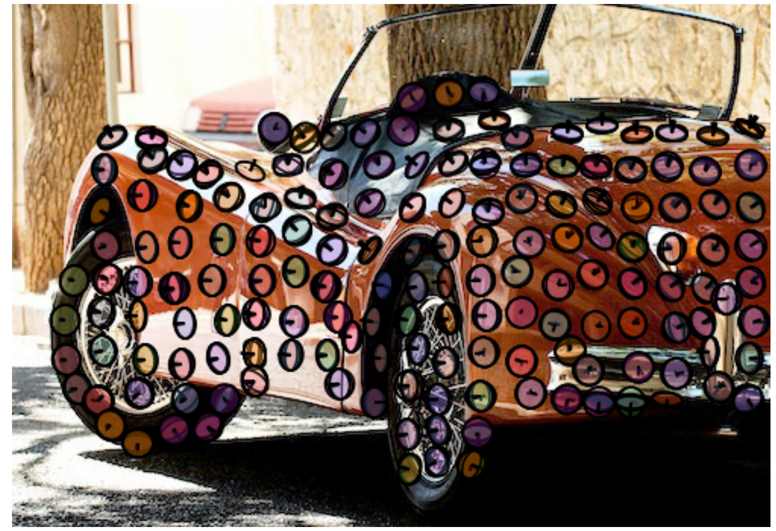
Shape-from-Shading



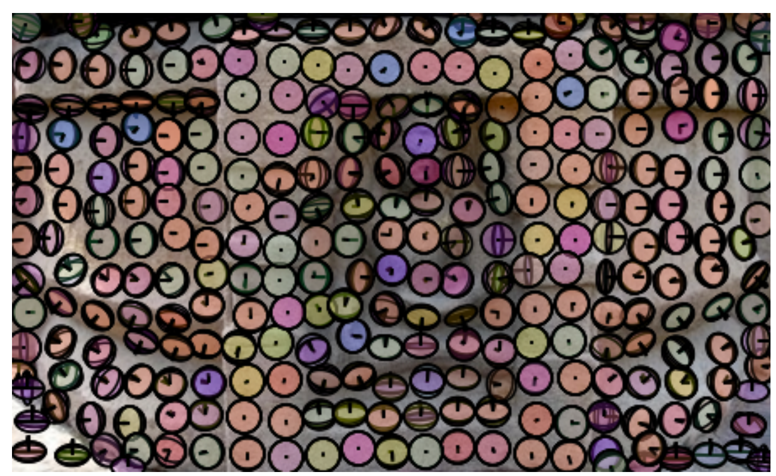


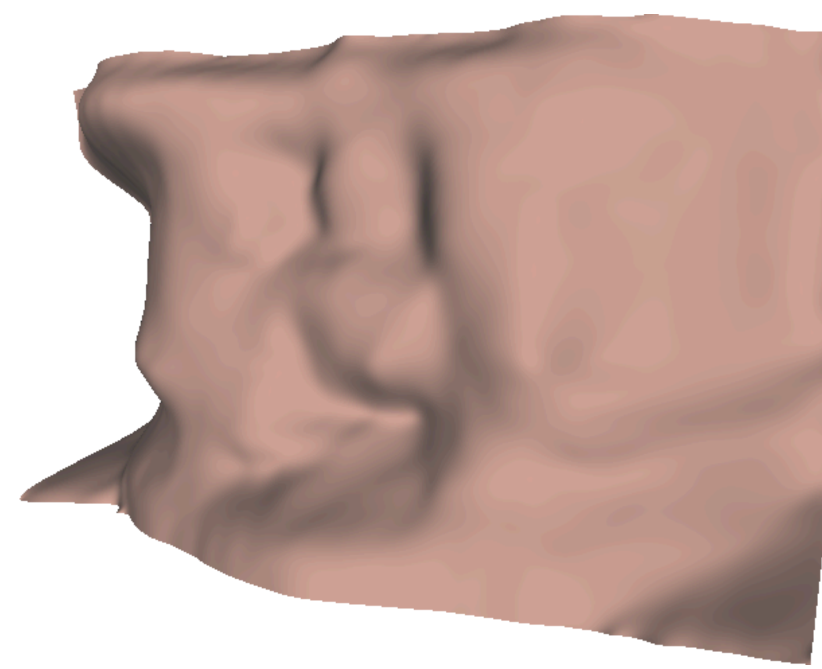


[Pedro Ribeiro Simões]

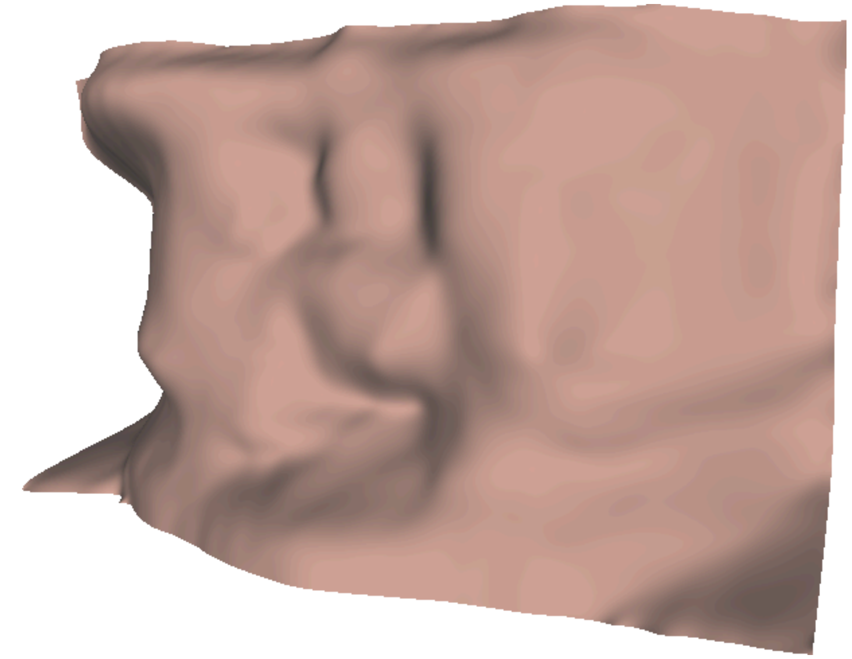


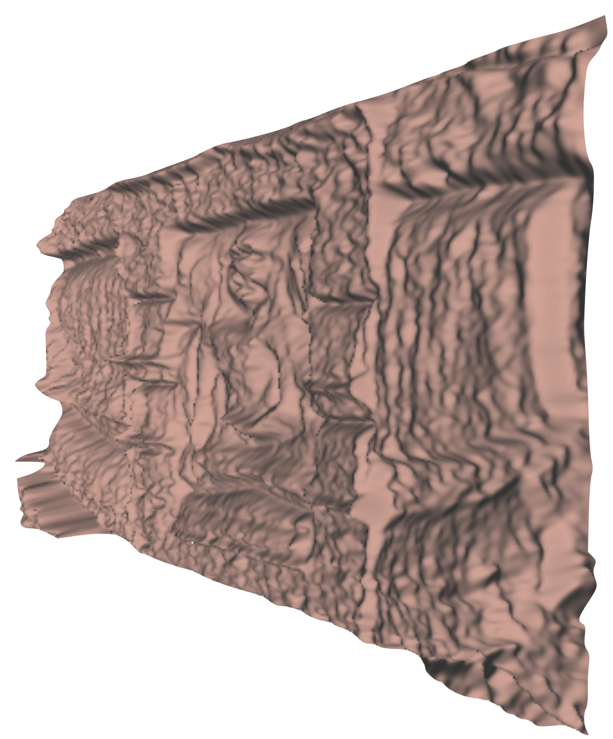
[Warren Apel]

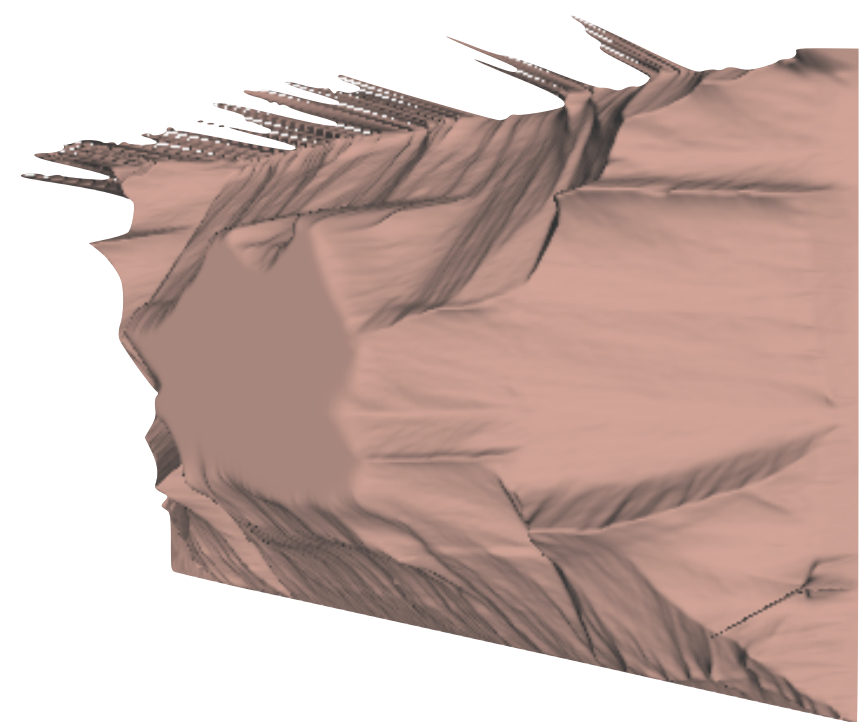




[Warren Apel]



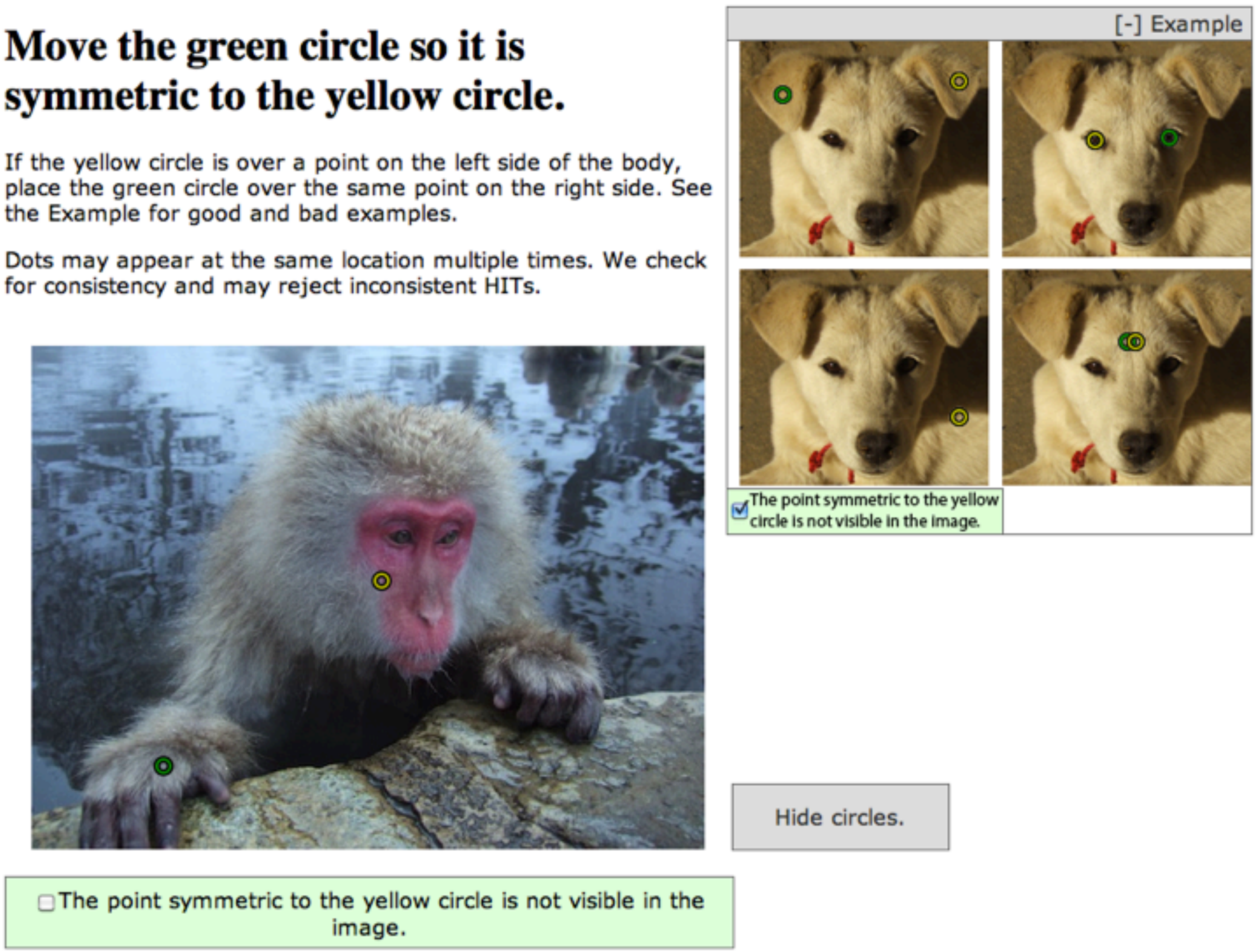


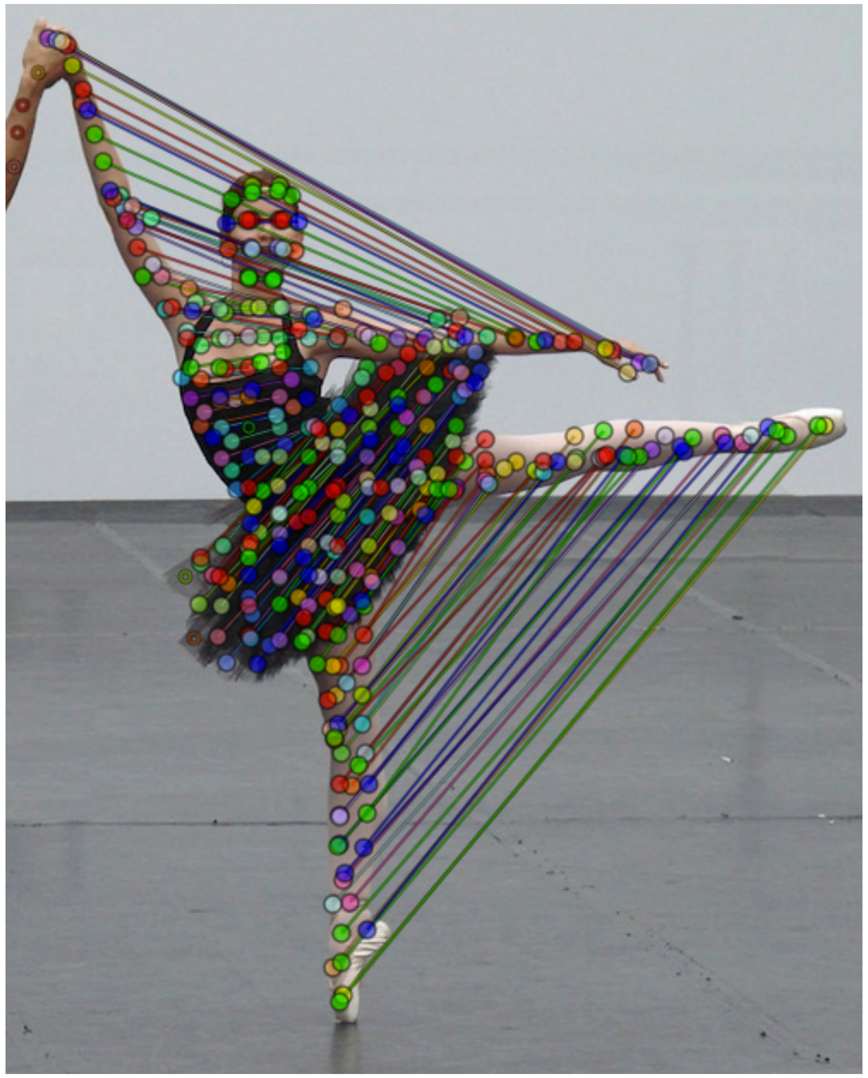


Make3D

Shape-from-Shading

Algorithm 3: Bilateral Symmetry Map




[flickr user dalbera]



Statistics

example	micro-tasks used	total $ cost	successful micro-task duration		algorithm delay until % complete	
			avg	median	50%	100%
normal map	1620–4340	$5.04–10.76	8.8 s	8.1 s	1.1–5.0 hrs	2.8–15.1 hrs
depth layers	2669–7620	$6.41–17.15	6.2 s	5.5 s	0.95–1.6 hrs	3.7–8.0 hrs
symmetry map	1020–1740	$3.24–3.92	9.0 s	8.5 s	0.4–1.6 hrs	0.7–4.9 hrs

Statistics

example	micro-tasks used	total $ cost	successful micro-task duration		algorithm delay until % complete	
			avg	median	50%	100%
normal map	1620–4340	$5.04–10.76	8.8 s	8.1 s	1.1–5.0 hrs	2.8–15.1 hrs
depth layers	2669–7620	$6.41–17.15	6.2 s	5.5 s	0.95–1.6 hrs	3.7–8.0 hrs
symmetry map	1020–1740	$3.24–3.92	9.0 s	8.5 s	0.4–1.6 hrs	0.7–4.9 hrs

Statistics

example	micro-tasks used	total $ cost	successful micro-task duration		algorithm delay until % complete	
			avg	median	50%	100%
normal map	1620–4340	$5.04–10.76	8.8 s	8.1 s	1.1–5.0 hrs	2.8–15.1 hrs
depth layers	2669–7620	$6.41–17.15	6.2 s	5.5 s	0.95–1.6 hrs	3.7–8.0 hrs
symmetry map	1020–1740	$3.24–3.92	9.0 s	8.5 s	0.4–1.6 hrs	0.7–4.9 hrs

Statistics

example	micro-tasks used	total $ cost	successful micro-task duration		algorithm delay until % complete	
			avg	median	50%	100%
normal map	1620–4340	$5.04–10.76	8.8 s	8.1 s	1.1–5.0 hrs	2.8–15.1 hrs
depth layers	2669–7620	$6.41–17.15	6.2 s	5.5 s	0.95–1.6 hrs	3.7–8.0 hrs
symmetry map	1020–1740	$3.24–3.92	9.0 s	8.5 s	0.4–1.6 hrs	0.7–4.9 hrs

Summary

HC algorithms can work where automatic algorithms still cannot.

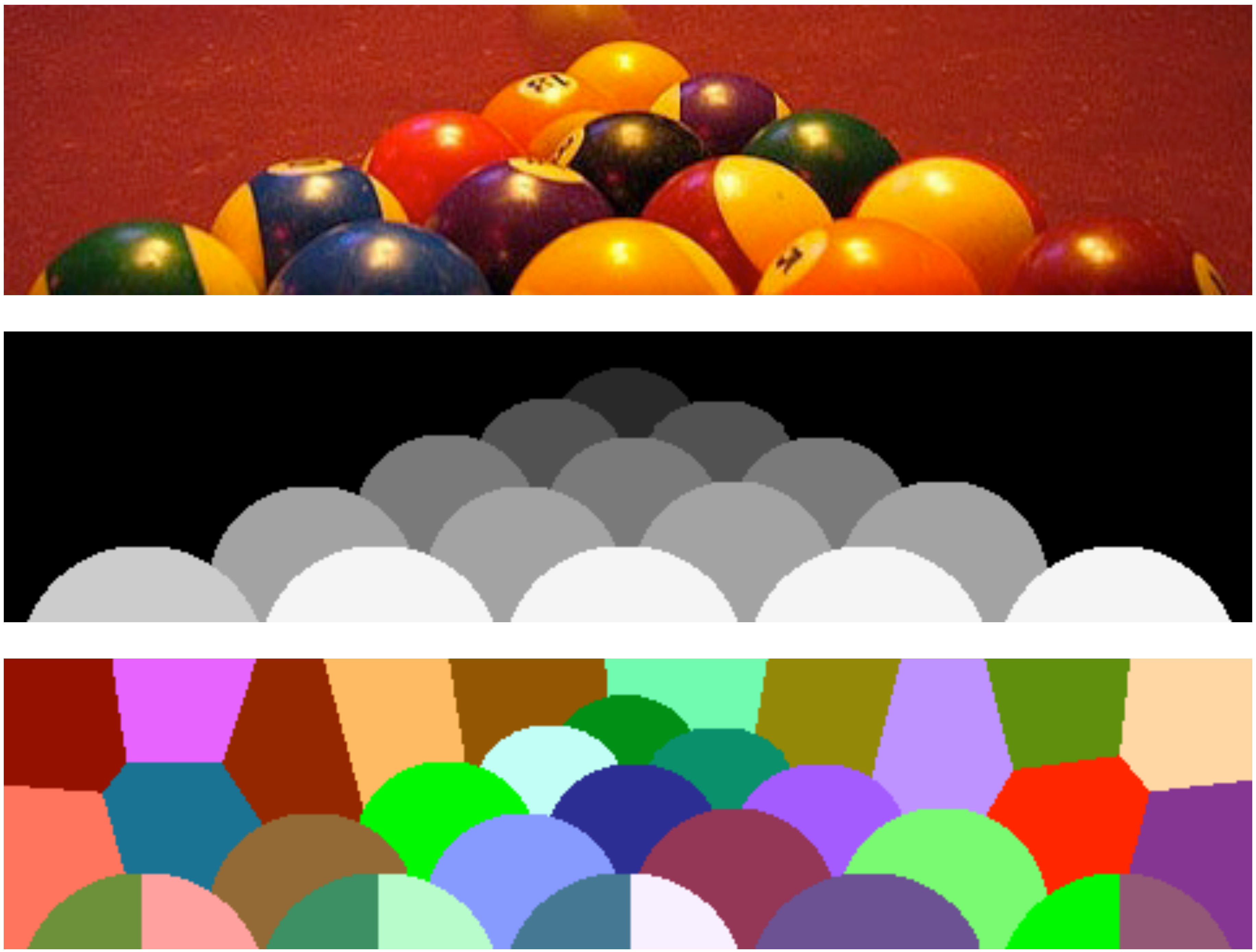
Identify the essential difficulty, and rephrase the algorithm in terms of micro human perception.

Problem	Micro-task	Combining Algorithm
depth layers	identify depth jumps	laplace equation
normal map	orient thumbtacks	bi-laplace equation
symmetry map	position point pair	none

If this were a Photoshop plug-in, would you use it?

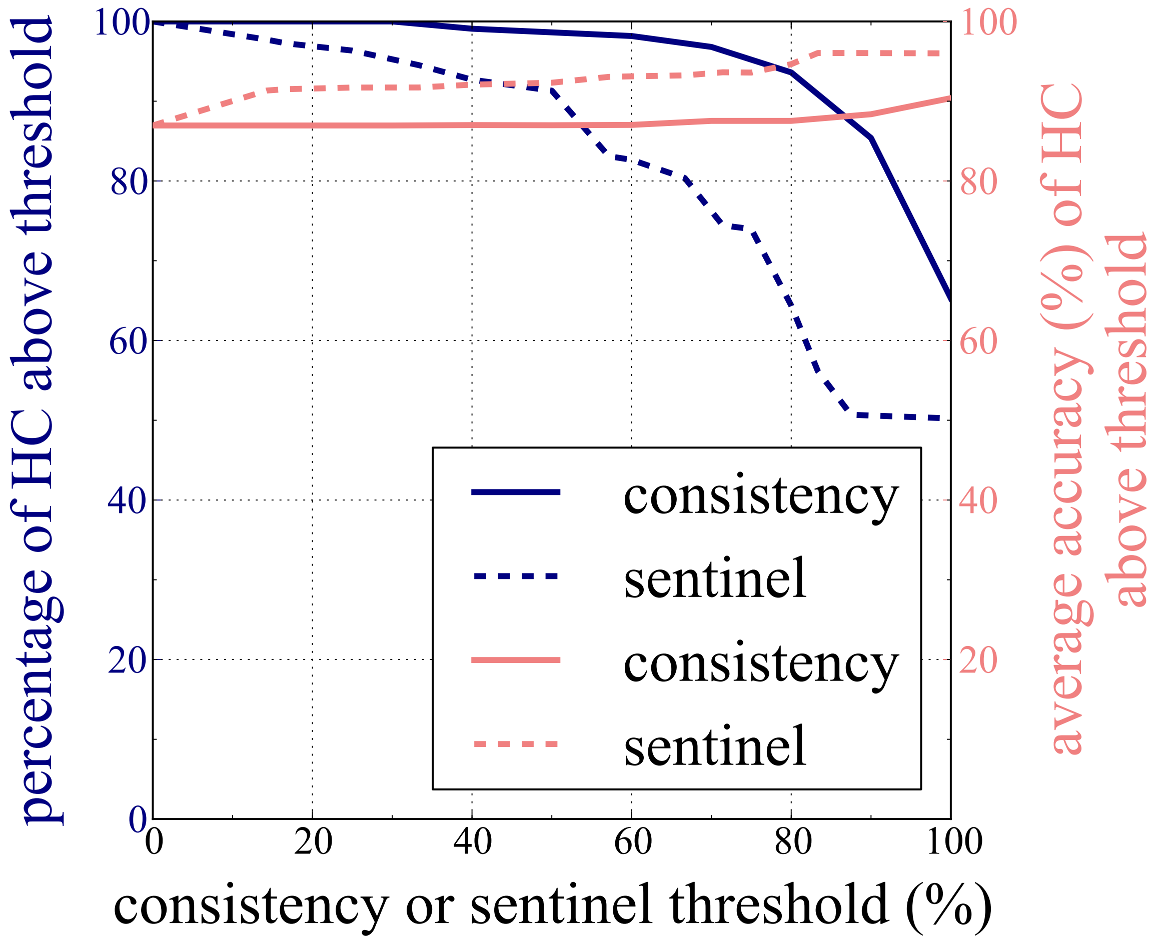
End

Accuracy

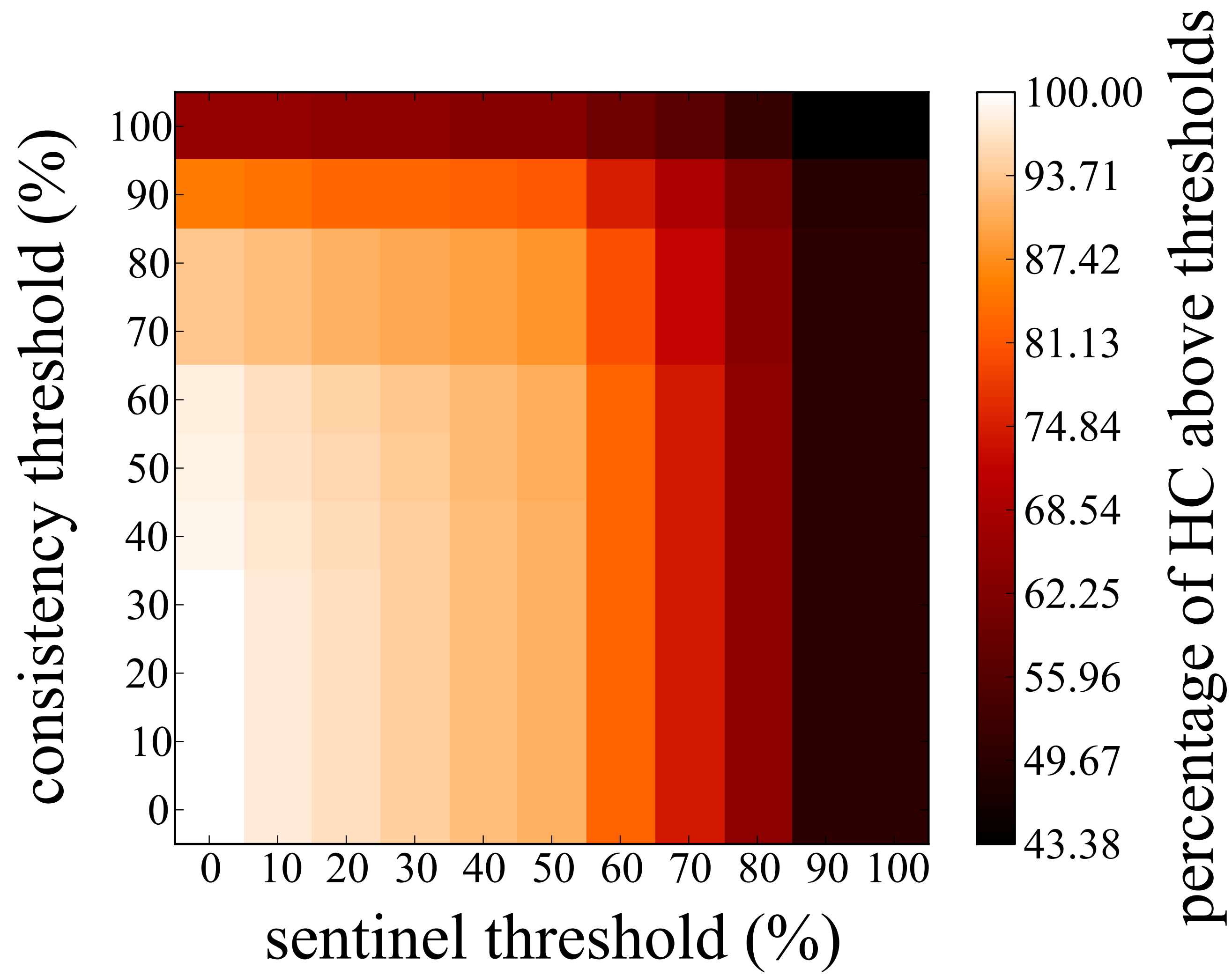


[flickr user fictures]

Accuracy

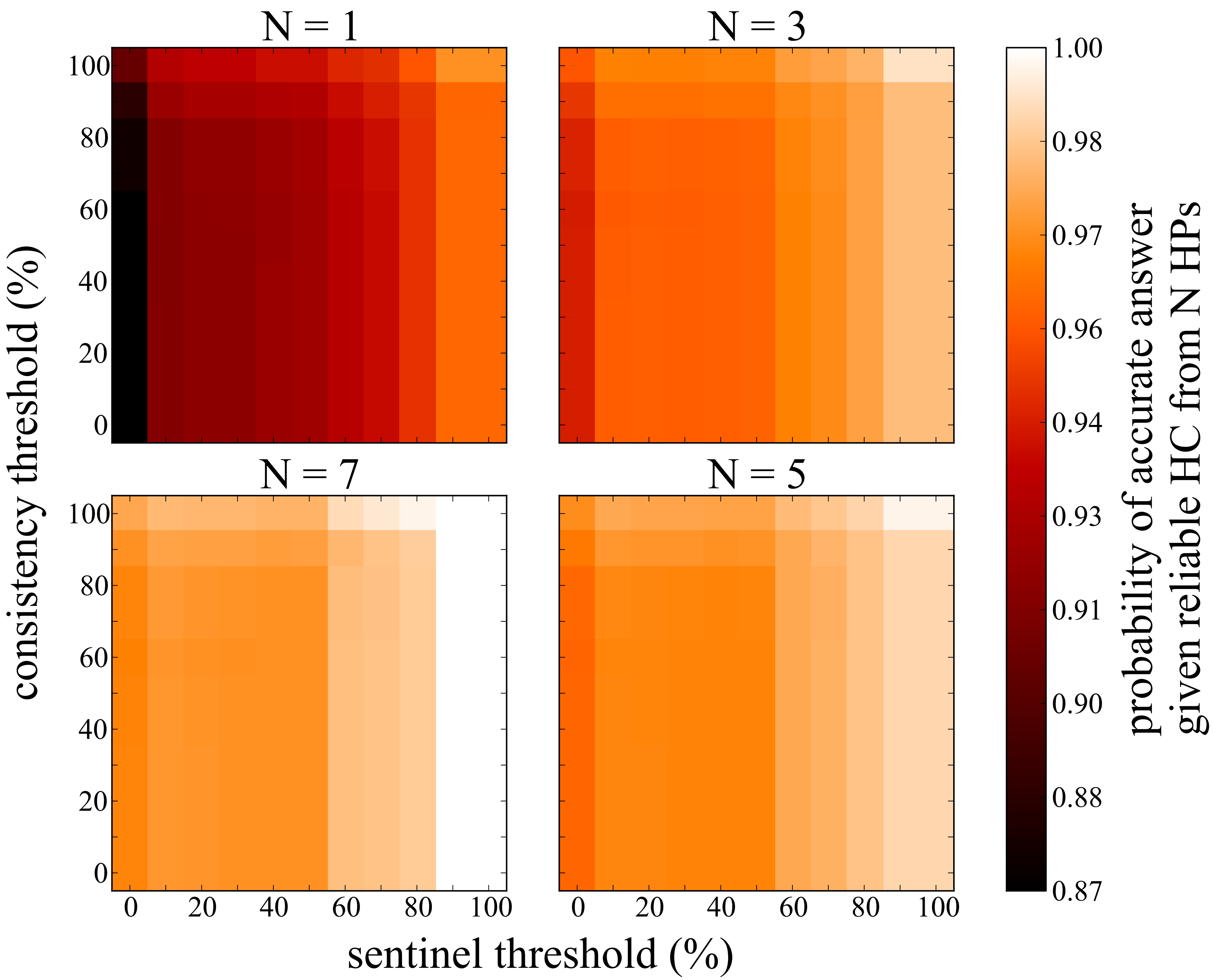


percentage of HC above threshold
100
80
60
40
20
0
0
20
40
60
80
100
consistency or sentinel threshold (%)
average accuracy (%) of HC above threshold
consistency
sentinel
consistency
sentinel



consistency threshold (%)
100
90
80
70
60
50
40
30
20
10
0
0 10 20 30 40 50 60 70 80 90 100
sentinel threshold (%)
100.00
93.71
87.42
81.13
74.84
68.54
62.25
55.96
49.67
43.38
percentage of HC above thresholds

Accuracy


N = 1
N = 3
N = 7
N = 5
100
80
60
40
20
0
0
20
40
60
80
100
consistency threshold (%)
sentinel threshold (%)
1.00
0.98
0.97
0.96
0.94
0.93
0.91
0.90
0.88
0.87
probability of accurate answer
given reliable HC from N HPs

Cost and Reliability

example	micro-tasks used	ratio of used per executed	$ per micro-task	total $ cost
normal map	1620–4340	0.60	.002–.003	$5.04–10.76
depth layers	2669–7620	0.76	.002	$6.41–17.15
symmetry map	1020–1740	0.93	.002	$3.24–3.92

Table 1: *Micro-tasks*

example	total HPs	% completely unreliable	average reliability for reliable HPs	micro-tasks per HP	
				avg	median
normal map	61	42%	89%	123	33
depth layers	48	35%	87%	193	63
symmetry map	19	24%	99%	97	20

Table 2: *Human Processors*

Timing

	successful micro-task duration		algorithm delay until % complete	
example	avg	median	50%	100%
normal map	8.8 s	8.1 s	1.1–5.0 hrs	2.8–15.1 hrs
depth layers	6.2 s	5.5 s	0.95–1.6 hrs	3.7–8.0 hrs
symmetry map	9.0 s	8.5 s	0.4–1.6 hrs	0.7–4.9 hrs

Related Work (1/6)

Many kinds of collective intelligence

- open-source software, Wikipedia, PageRank, supervised learning, elections?

Modern assembly line (Ford Motor Company 1908–1915)

Interchangeable parts:

- Adam Smith on division of labor (1776)
- Terracotta army (3rd century BC)
- Venetian Arsenal (ship building)

Related Work (2/6)

Online:

- [von Ahn 2008]
- [Little et al. 2010a,b] and [Bernstein 2010]
- [Bigham et al. 2010] and [Bernstein 2011]
- [Sorokin et al. 2010]
- many more recent/contemporary applications

Recast existing experiments

- [Koenderink et al. 1992], [Cole et al. 2009]
- [Chen et al. 2009]

Related Work (3/6)

Training data:

- ESP Game [von Ahn and Dabbish 2004], …
- LabelMe [Russel et al. 2008; Yuen et al. 2009]
- Hands by Hand [Spiro et al. 2010]

Using HC data gathered offline:

- [Talton et al. 2009]
- [Kalogerakis et al. 2010] using [Chen et al. 2009]

Related Work (4/6)

Depth Layer Algorithm

- automatic: [Hoiem et al. 2005; Assa and Wolf 2007; Saxena et al. 2009]
- manual: [Oh et al. 2001; Ventura et al. 2009; Sykora et al. 2010]

Normal Map Algorithm

- manual: [Wu et al. 2008]

Symmetry Map Algorithm

- automatic: [Chen et al. 2007]

Related Work (5/6)

History

- “When Computers Were Human” [Grier 2005]
- Genetic Algorithms
    - [Sims 1991]
    - Interactive Genetic Algorithm [Takagi 2001]
    - Human-Based Genetic Algorithms [Kosorukoff 2001]
    - Electric Sheep
- Open Mind Initiative
- collaborative filtering: [Goldberg et al. 1992; Adomavicius and Tuzhilin 2005]

“Human Computation” [von Ahn 2005]

Related Work (6/6)

Recent survey: [Quinn and Bederson 2011]

Market properties:

- [Ipeirotis 2010; Chilton et al. 2010; Faridani et al. 2011; Mason and Suri 2011; Mason and Watts 2010]

Surface perception:

- [Koenderink et al. 1992; Belheumer et al. 1997; Koenderink et al. 2001]

Shape-from-Shading:

- [Durou et al. 2008]

Theoretical Limits

125–180 seconds (median) / 20 questions = 6.25–9 seconds per perception for our tasks

7 billion humans (does not include other animals capable of similar tasks)

(number of humans) / (seconds per perception) ~= 1 billion perceptions per second