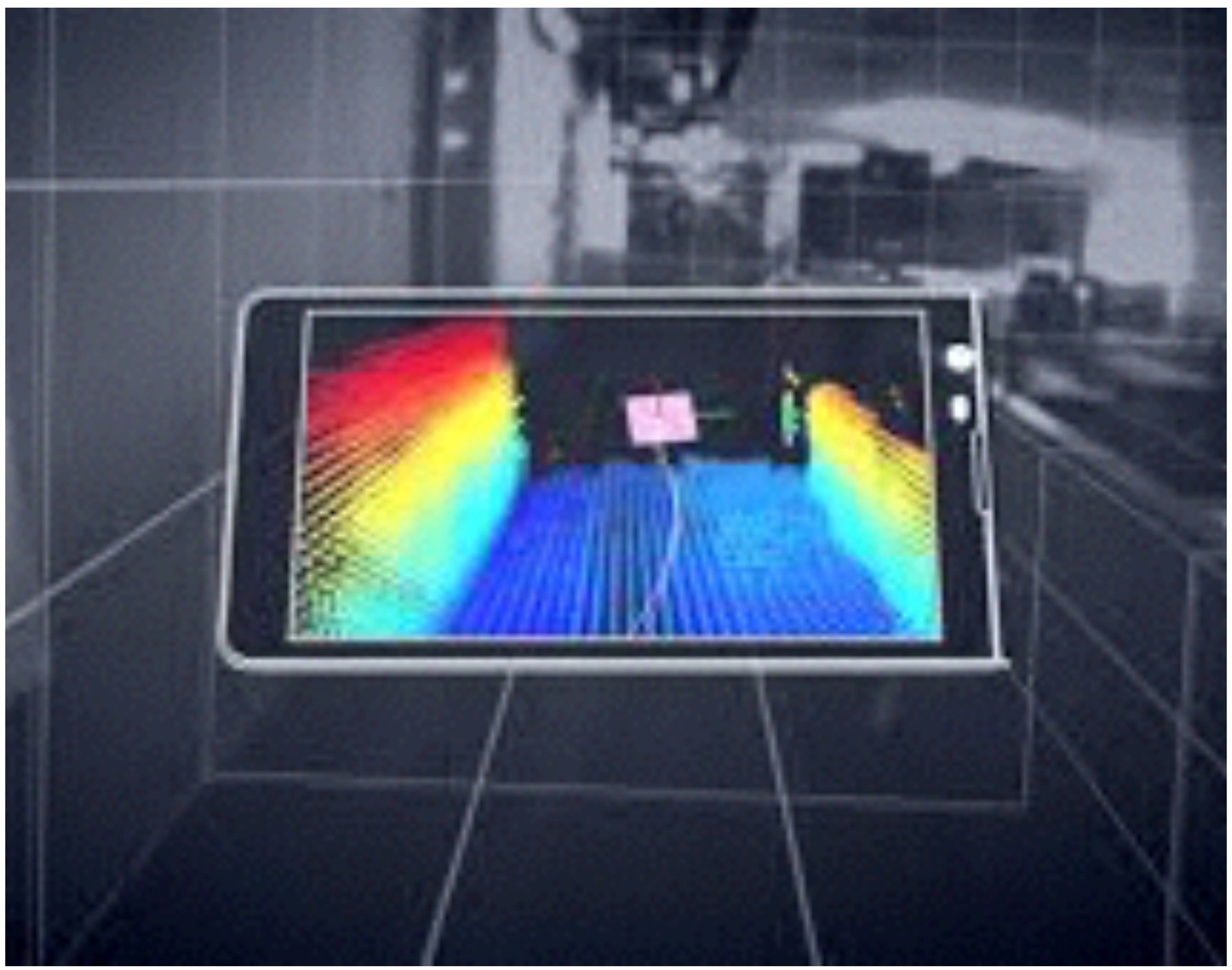
Part Structures in Large Collections of 3D Models

Vladimir Kim

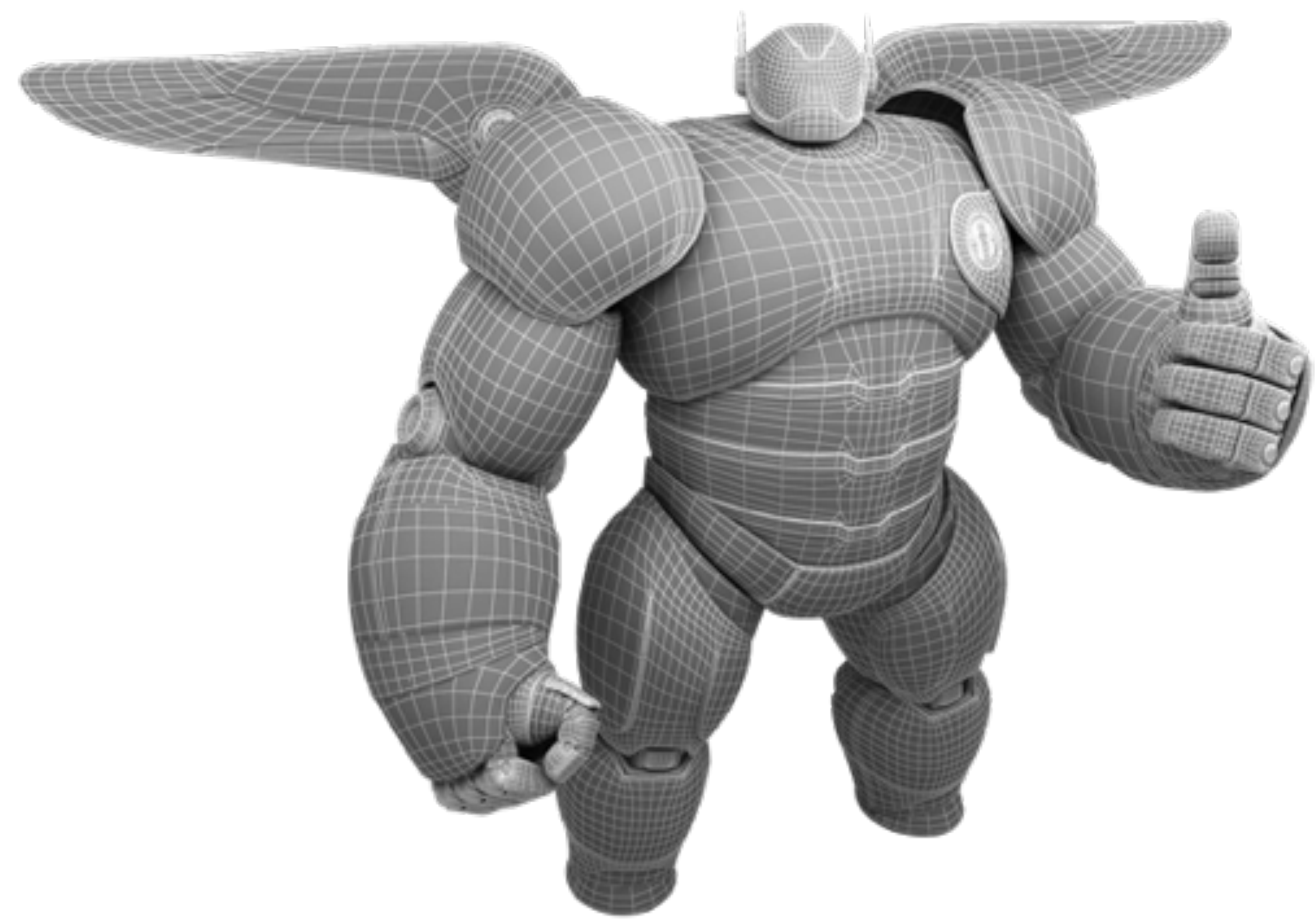
Adobe Research

Motivation

Explore, Analyze, and Create Geometric Data



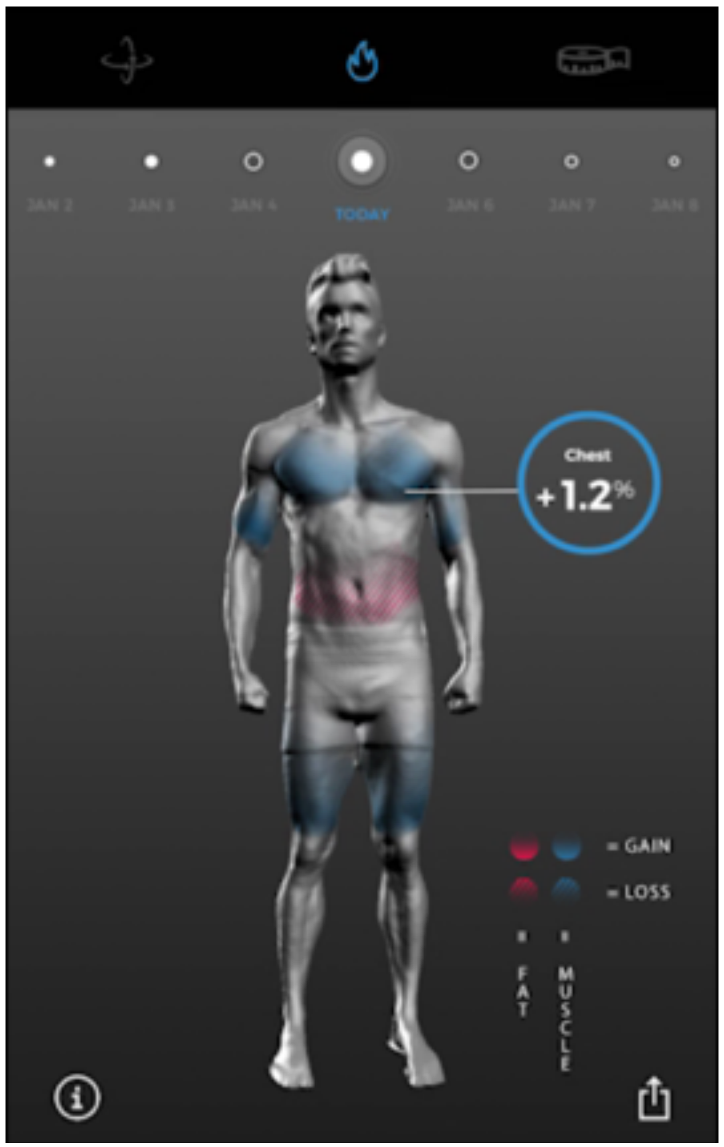
Real



Virtual

Motivation

Explore, Analyze, and Create Geometric Data

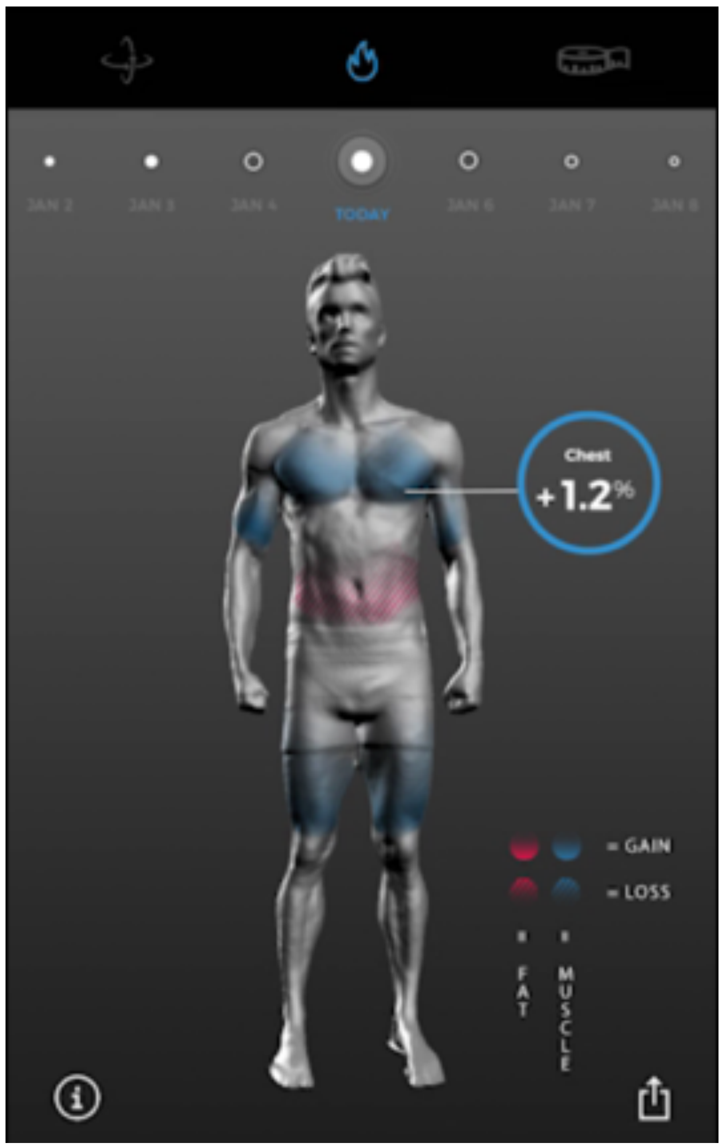


Personal data

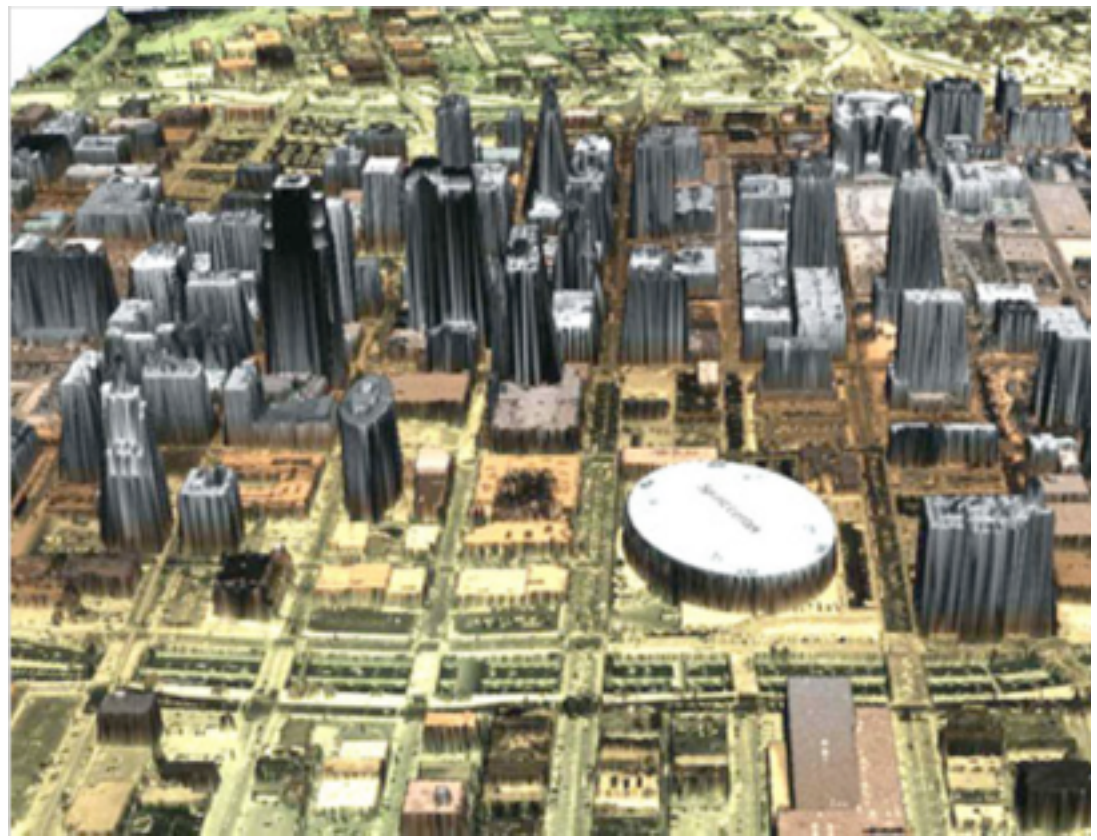
(image from naked.fit)

Motivation

Explore, Analyze, and Create Geometric Data



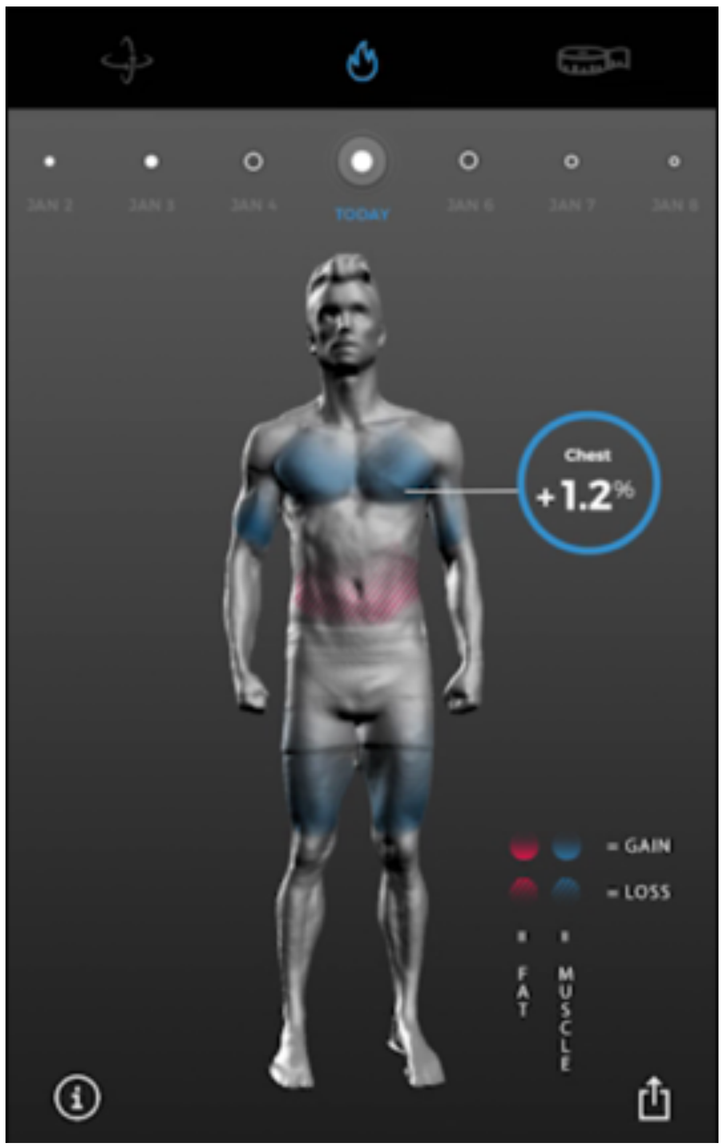
Personal data
(image from naked.fit)



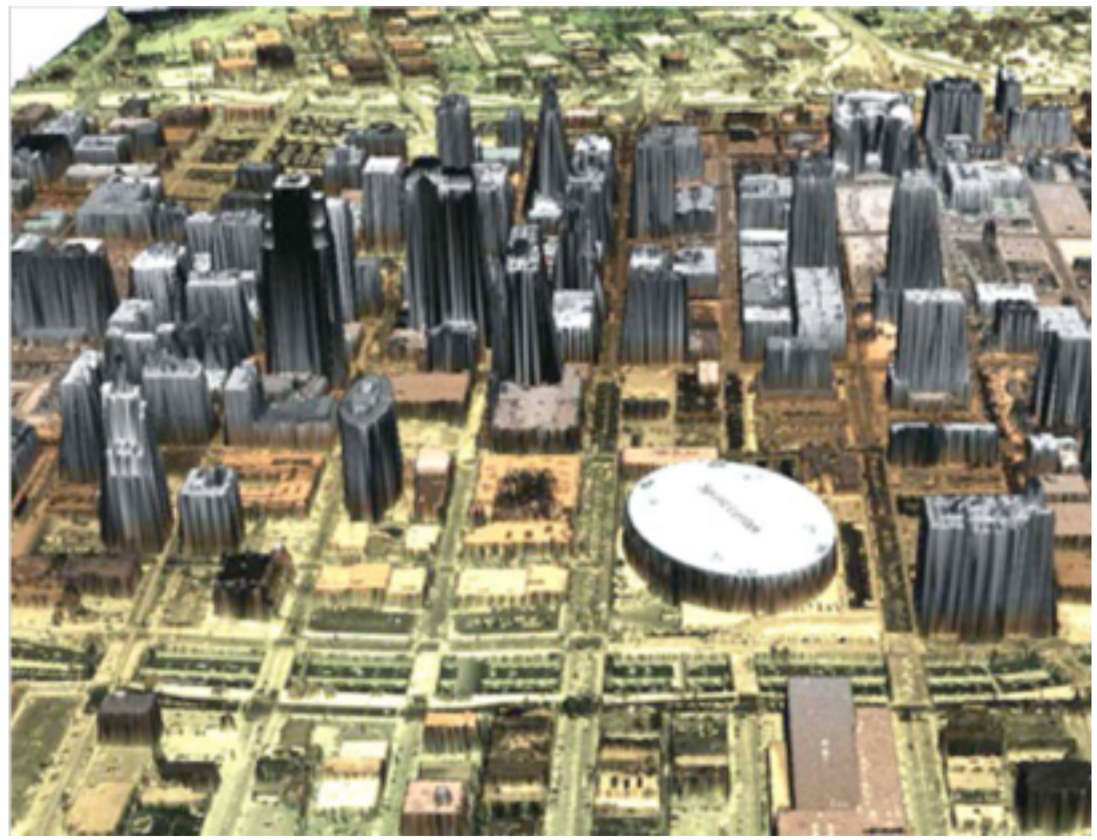
Scans of environments

Motivation

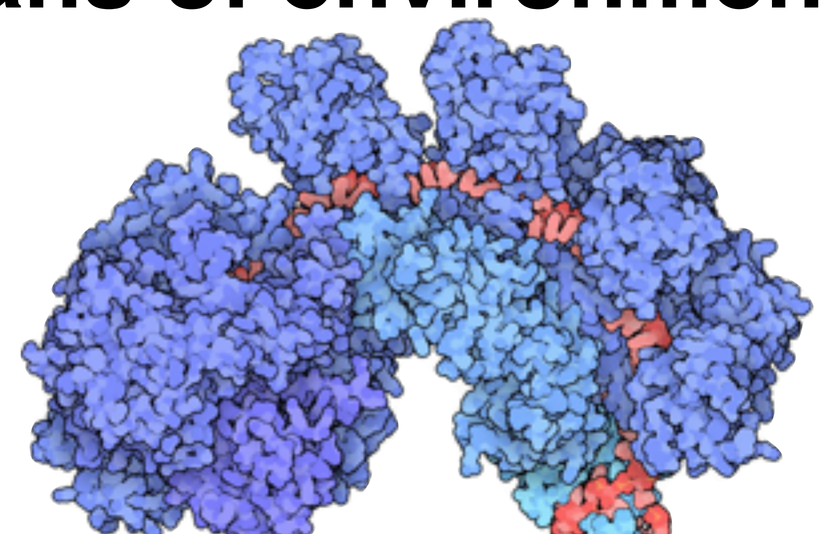
Explore, Analyze, and Create Geometric Data



Personal data
(image from naked.fit)



Scans of environments

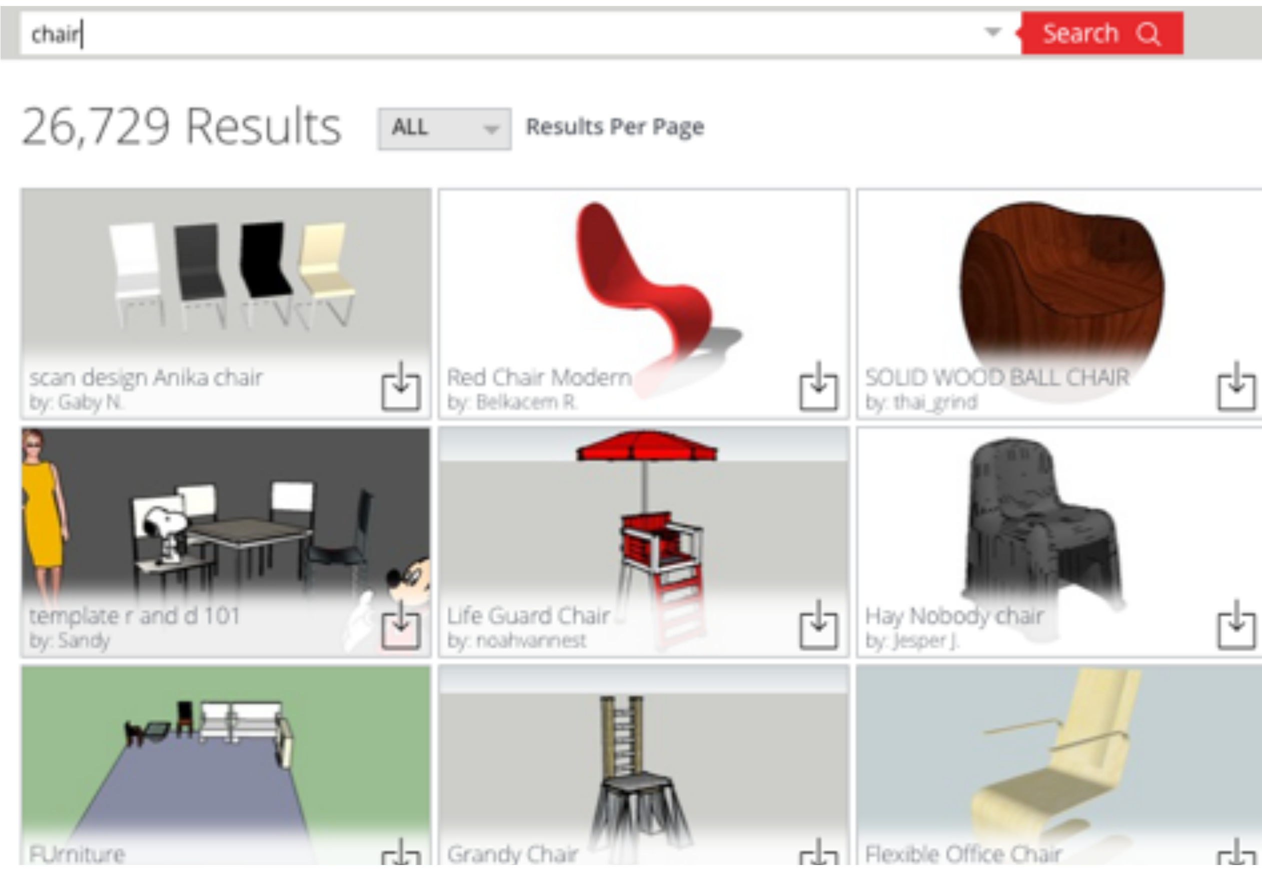


Scientific data

Medical Imaging

Motivation

Explore, Analyze, and Create Geometric Data

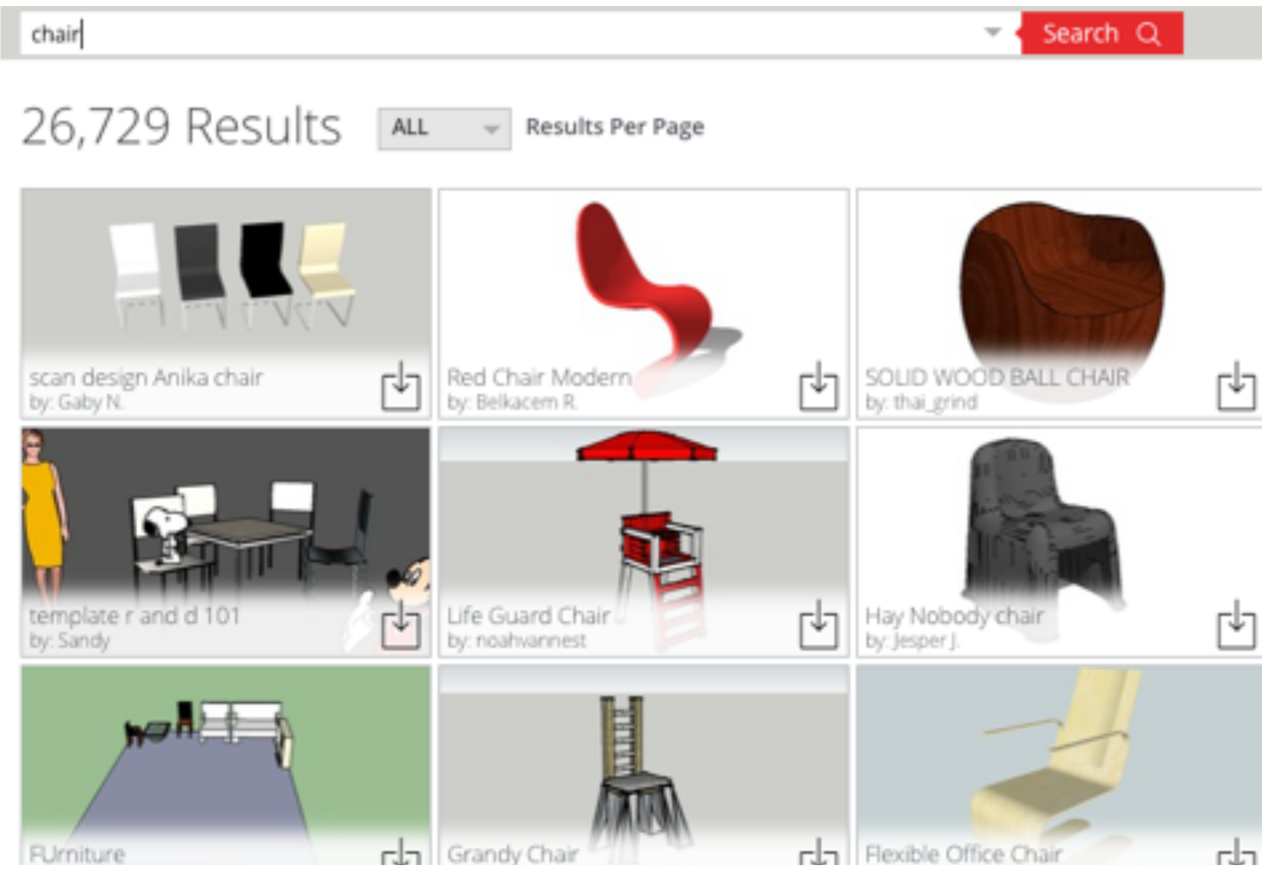


CG data

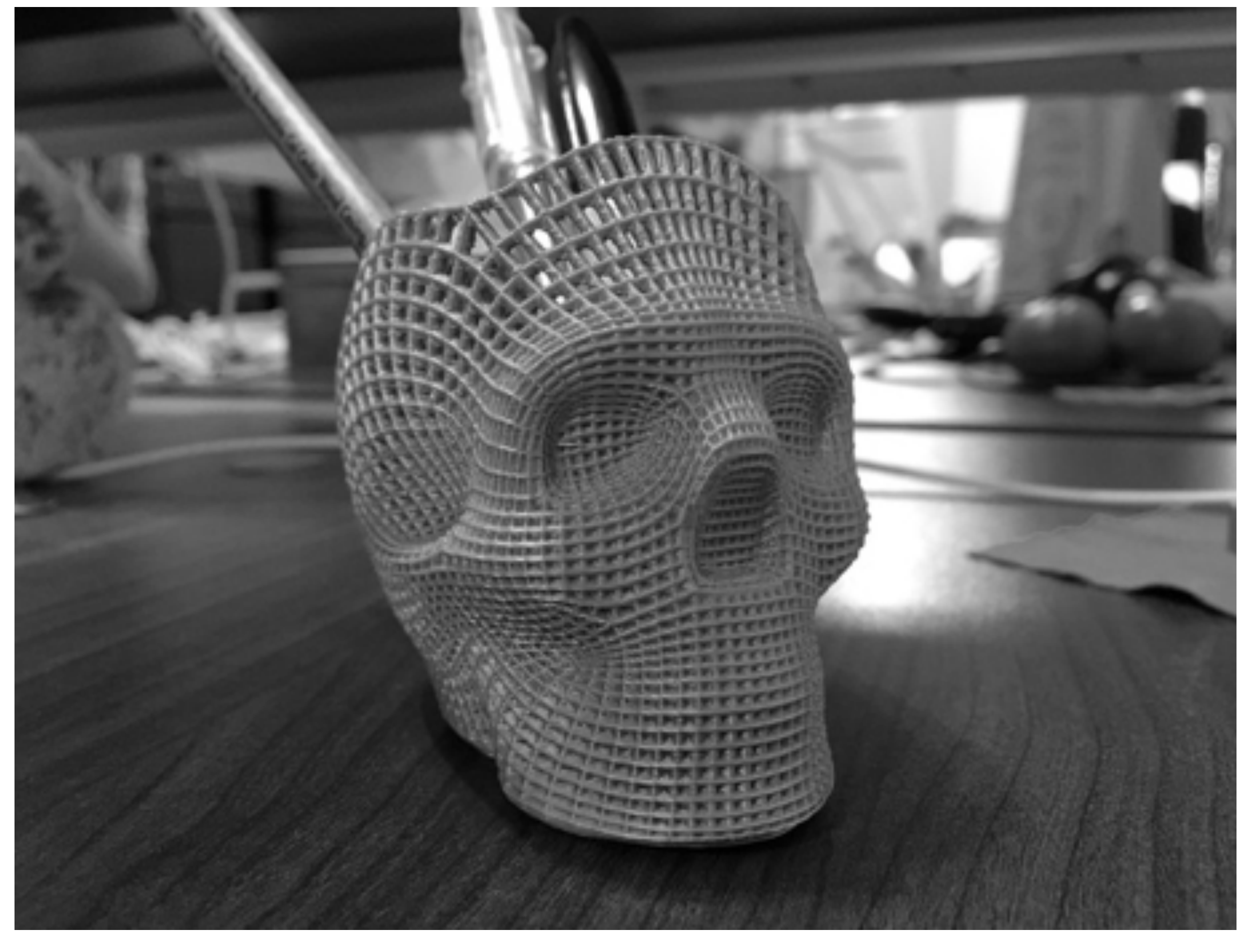
(image from 3dwarehouse.sketchup.com)

Motivation

Explore, Analyze, and Create Geometric Data



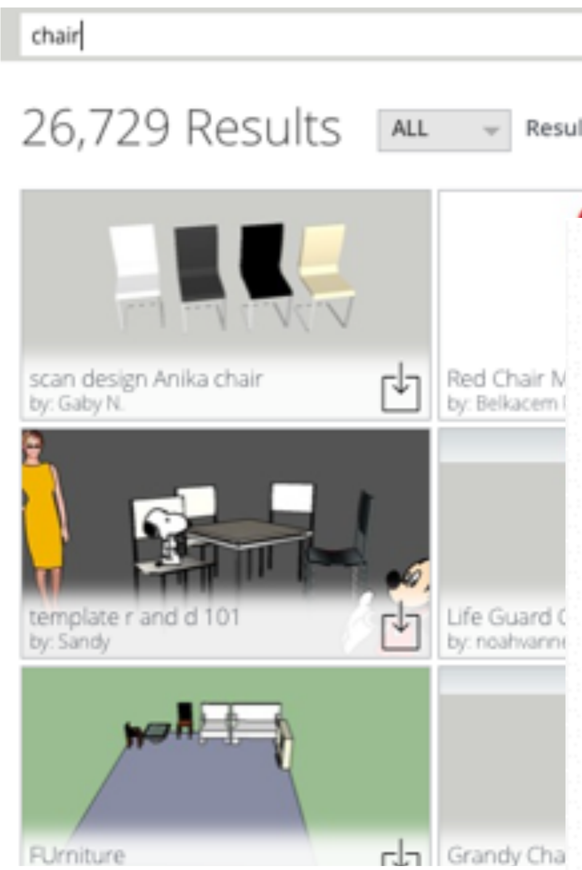
CG data
(image from 3dwarehouse.sketchup.com)



3D printing data
(image from thingiverse.com/)

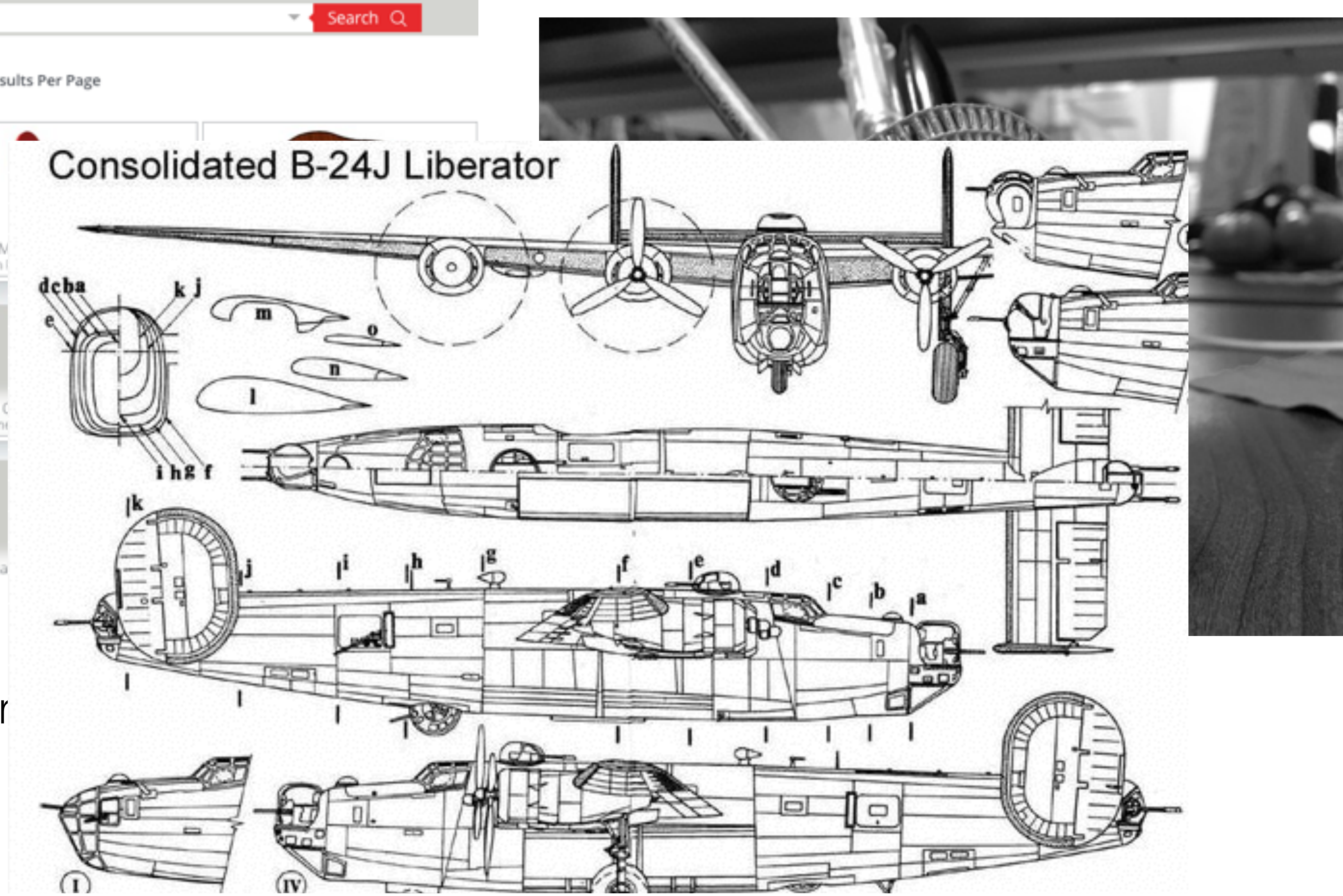
Motivation

Explore, Analyze, and Create Geometric Data



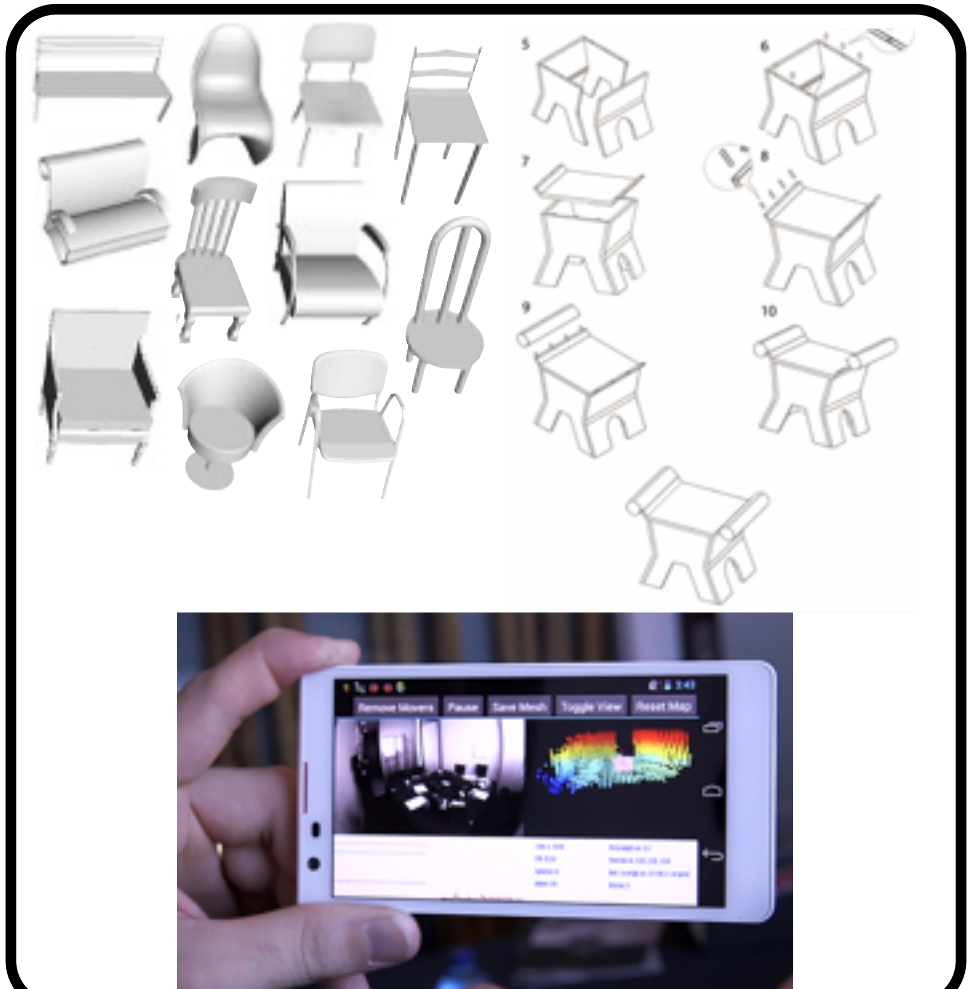
CG data

(image from 3dwar



CAD models

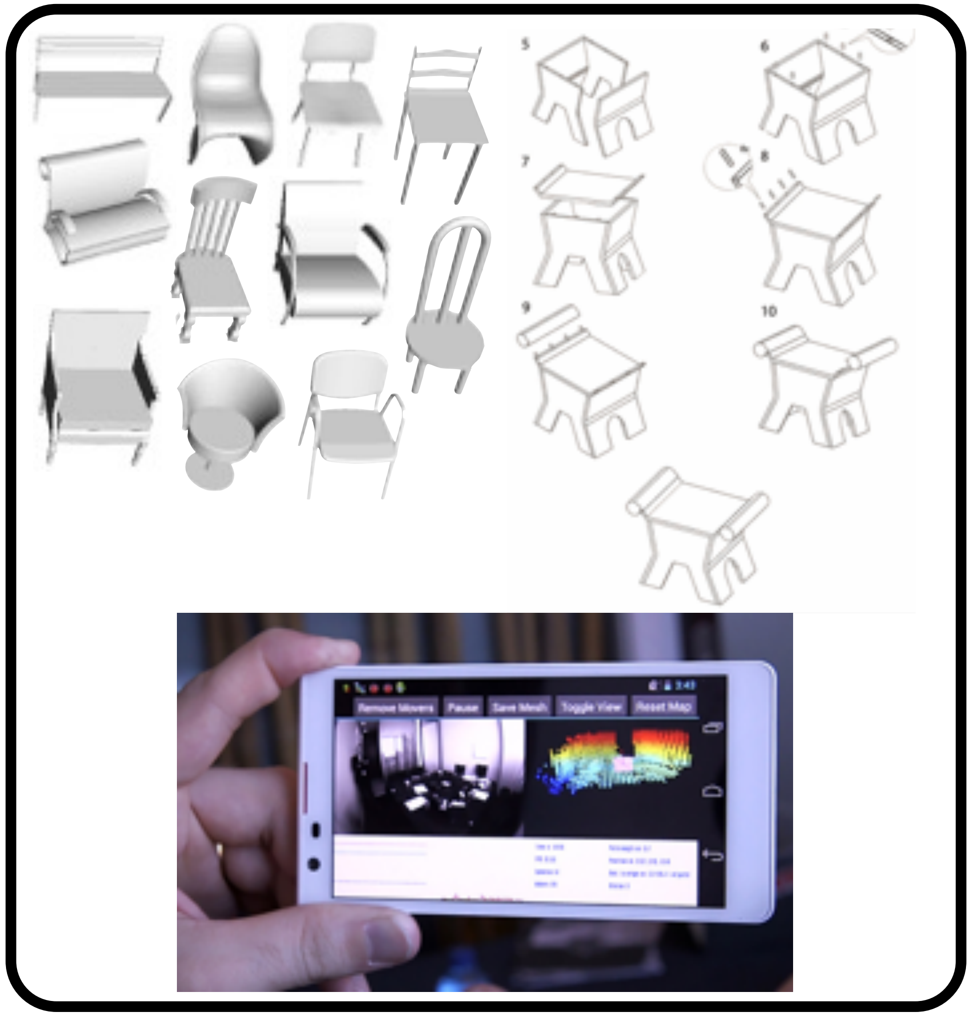
Goal



3D Data

→ **Structural Model** → **Applications**

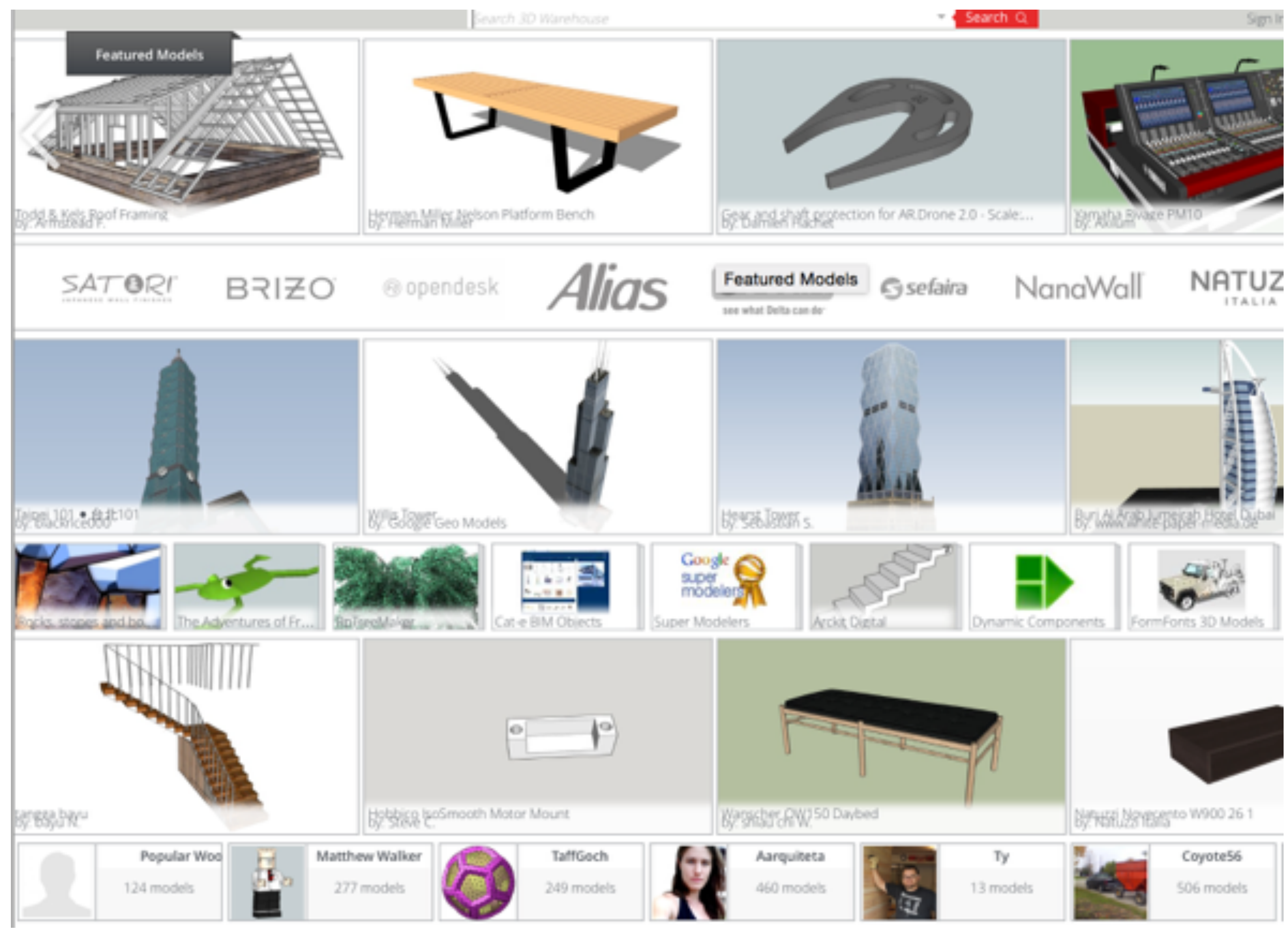
Goal



3D Data

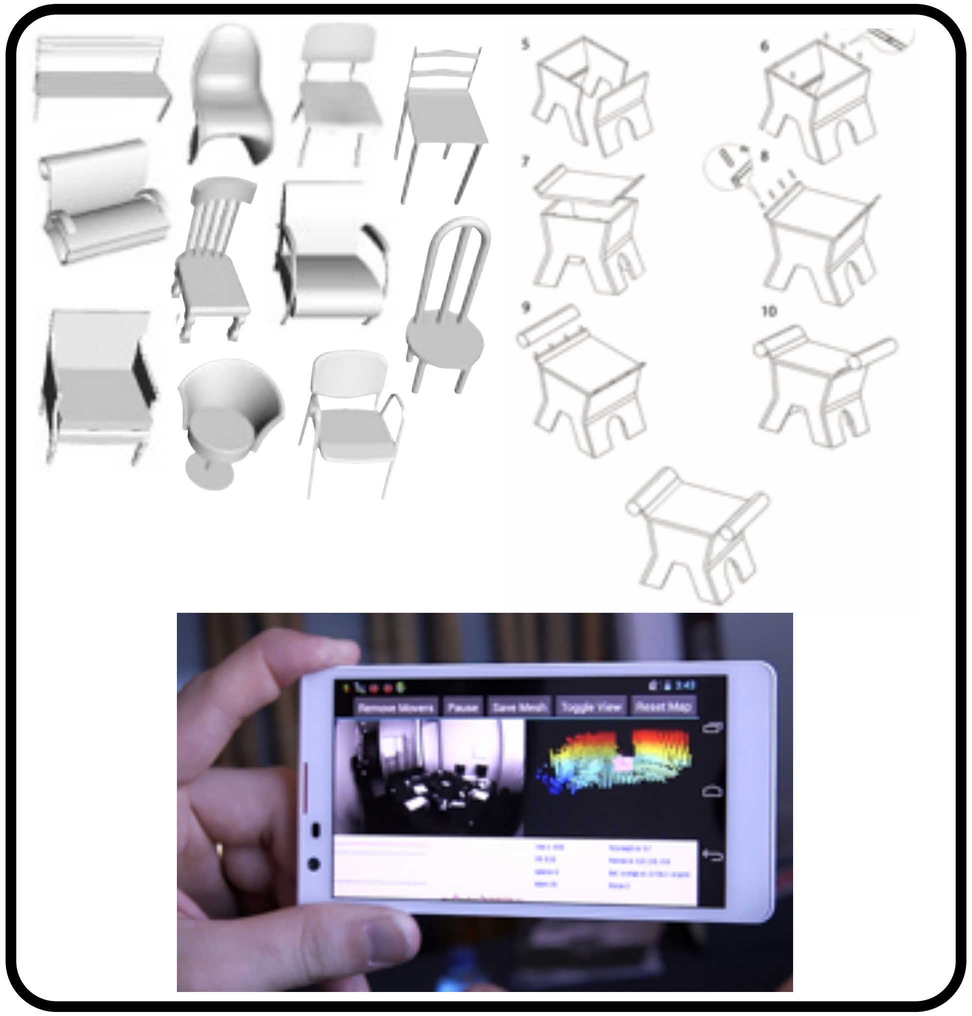
Structural Model

Explore



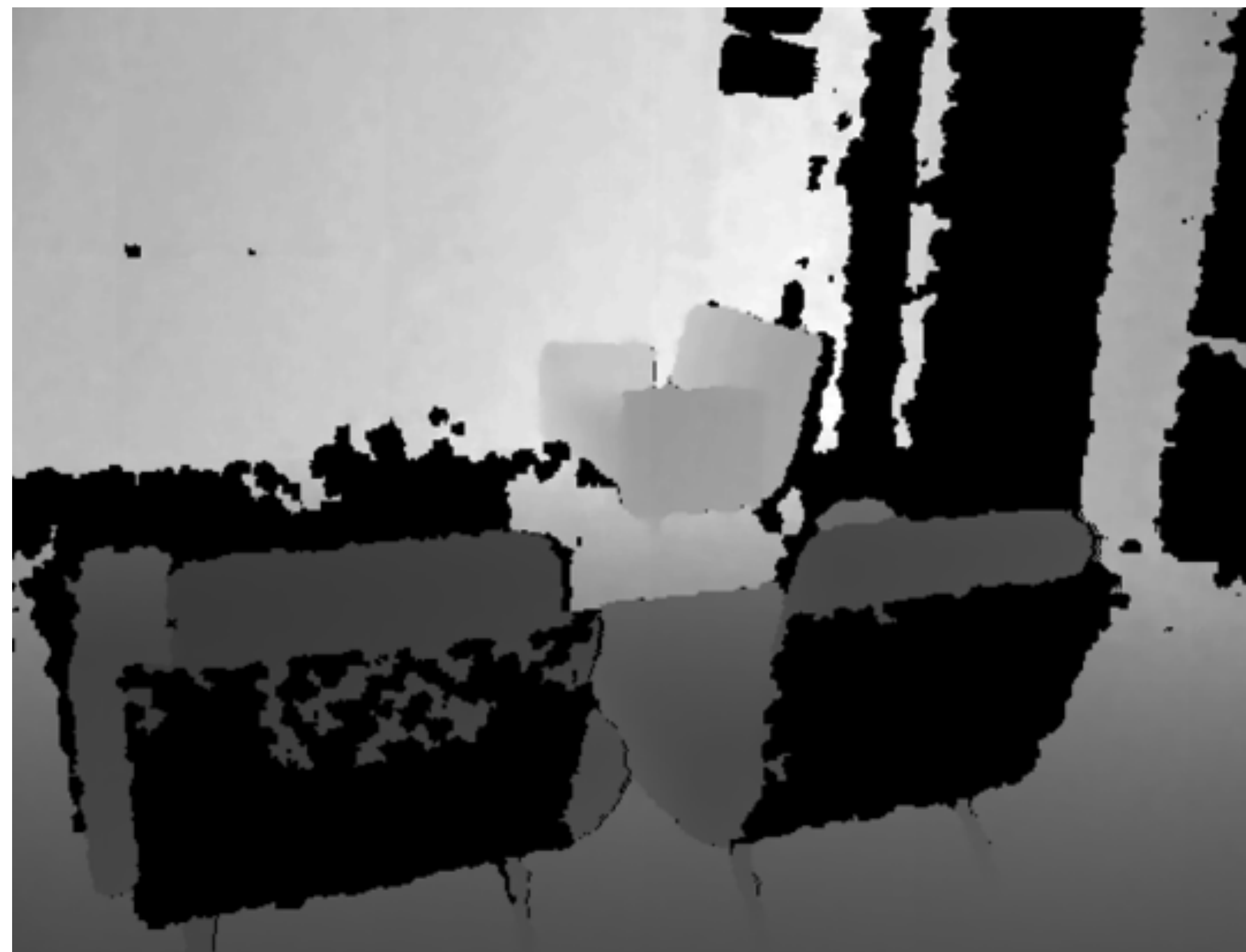
Organize and explore large collections of shapes

Goal



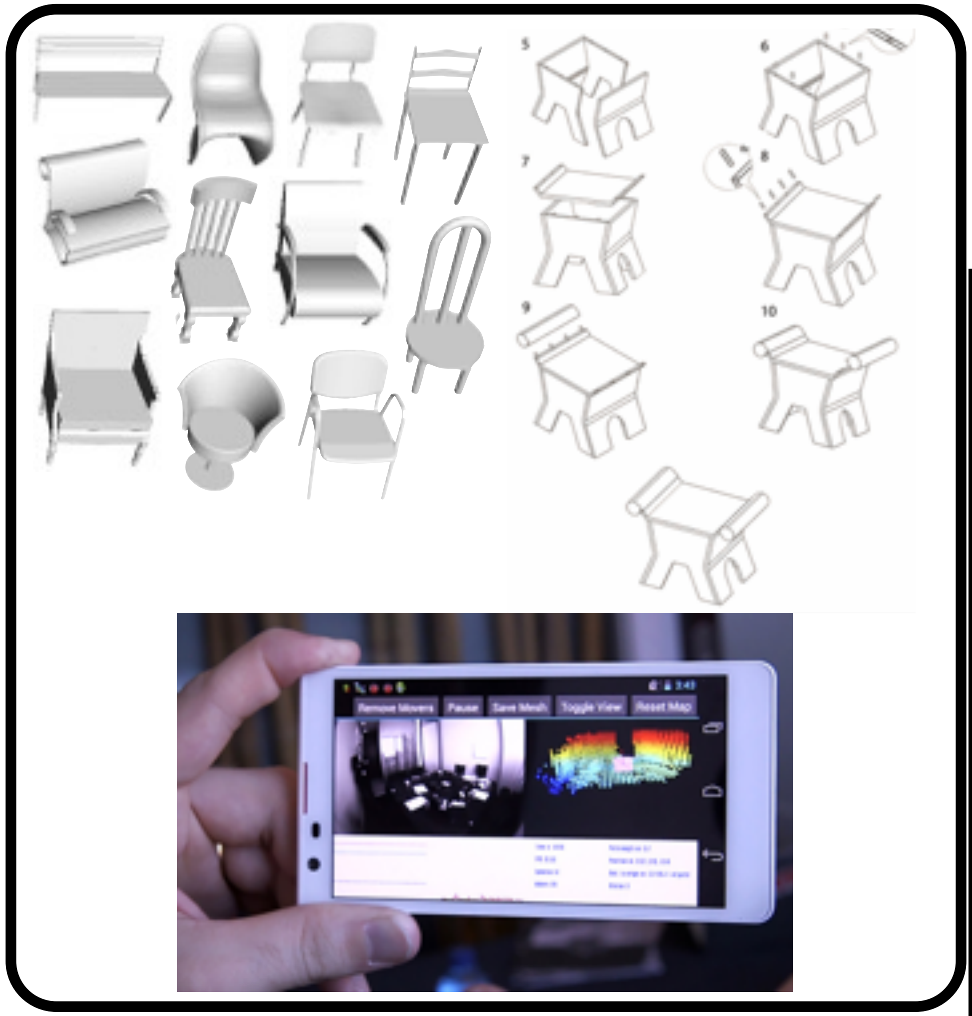
3D Data

→ **Structural Model** → **Analyze**



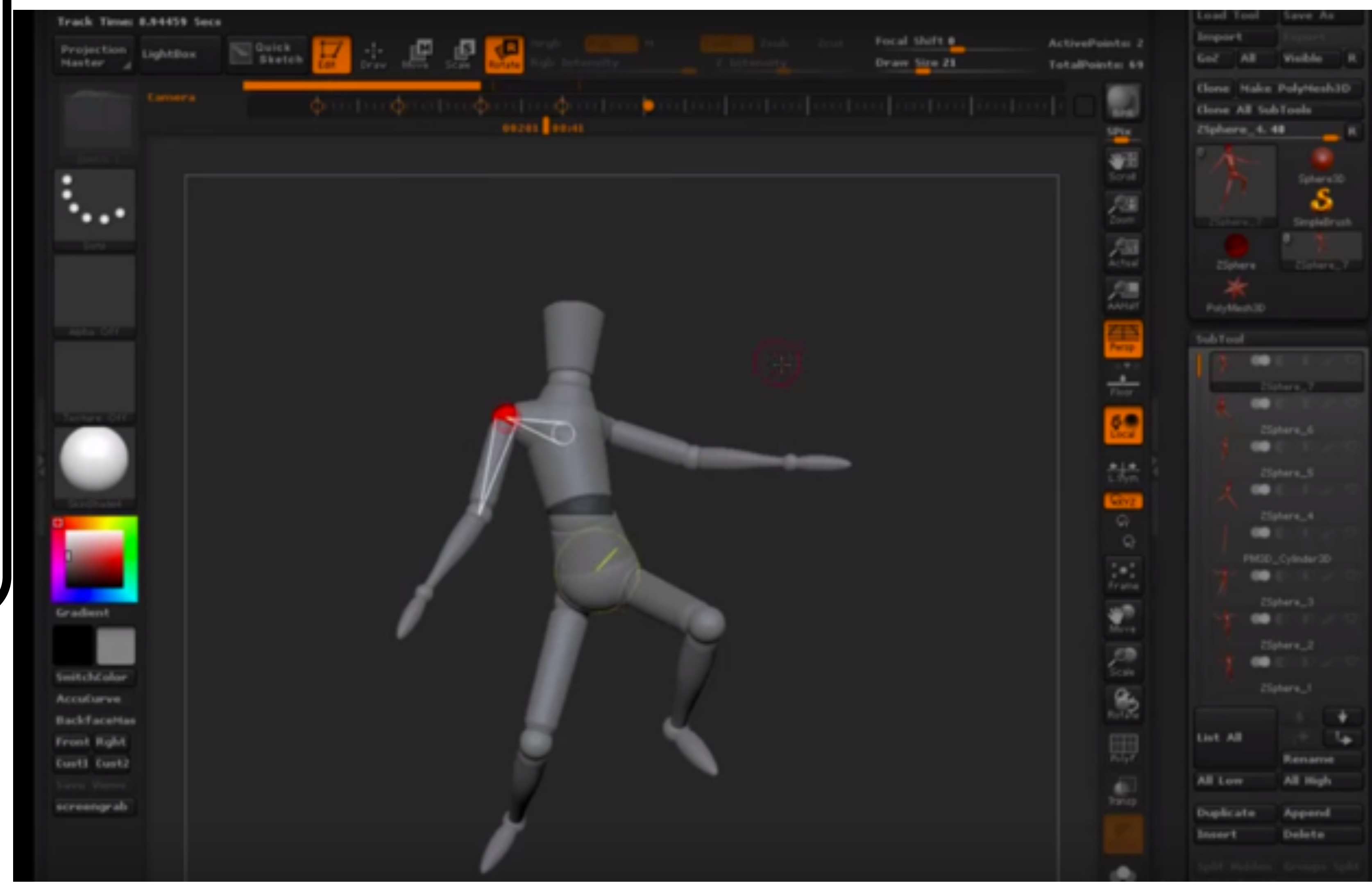
Understand and label novel geometric data

Goal



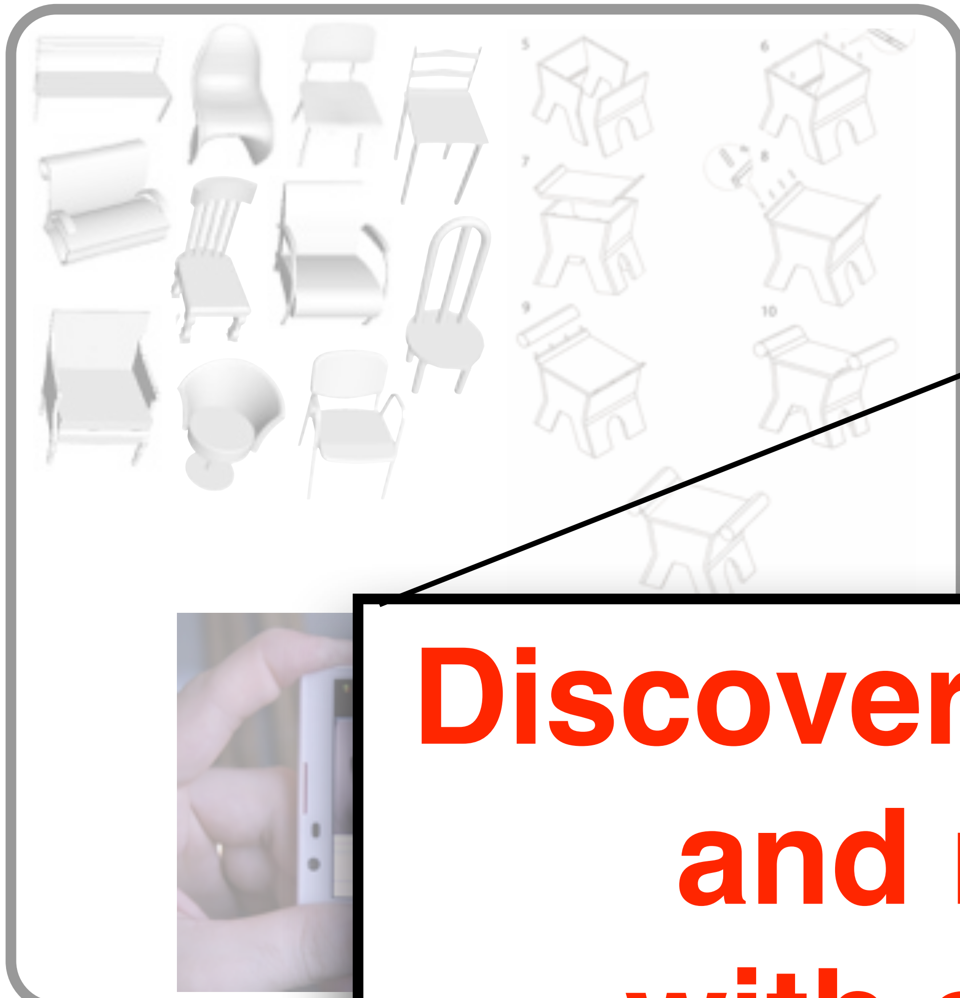
3D Data

→ **Structural Model** → **Create**



Make 3D modeling more accessible to non-experts

Goal



Structural Model

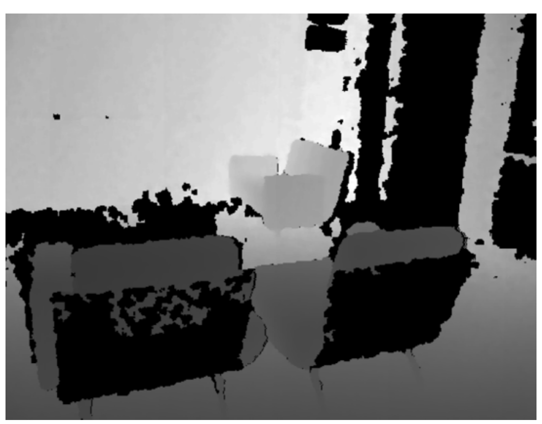
Create

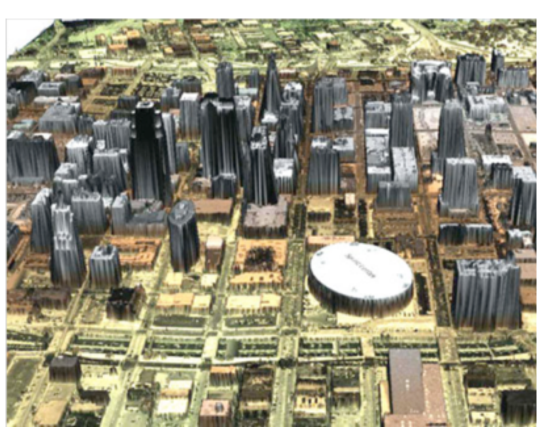
Discover repeating patterns and represent them with a concise model

Make 3D modeling more accessible to non-experts

Challenges

Discover common structure





Diversity

TurboSquid

3D Warehouse

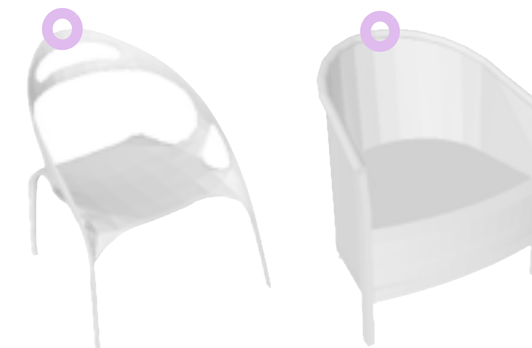
shapeways

Thingiverse

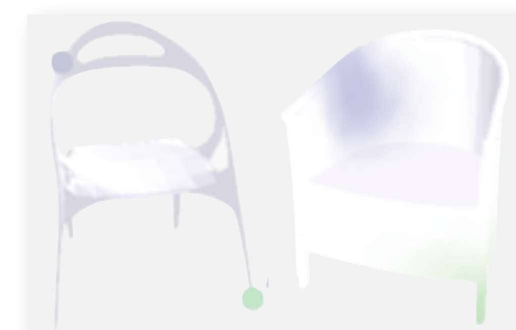
ShapeNet

Project Tango

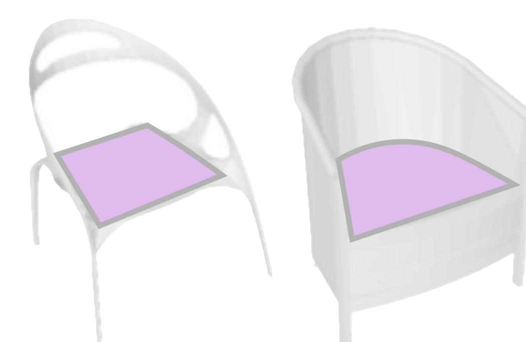
Scale



Point-to-point



Region-to-region

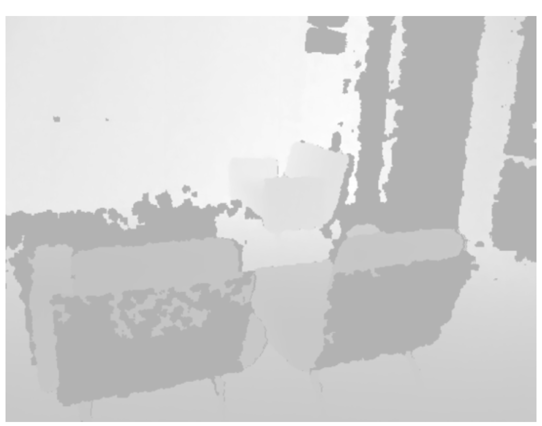


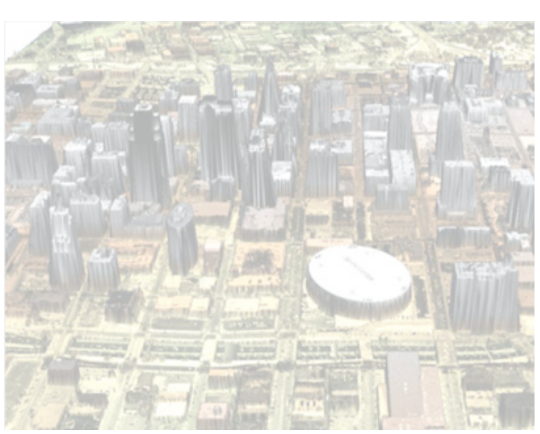
Part-to-part

Ambiguity

Challenges

Discover common structure





Diversity

TurboSquid

3D Warehouse

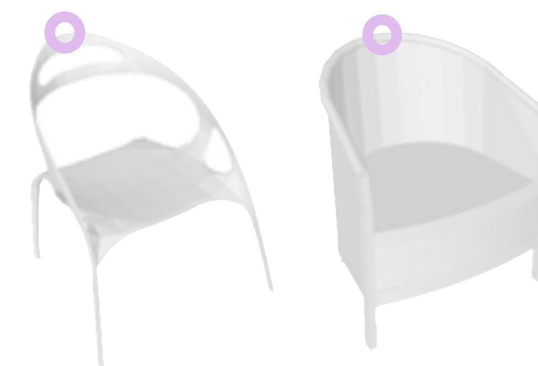
shapeways

Thingiverse

ShapeNet

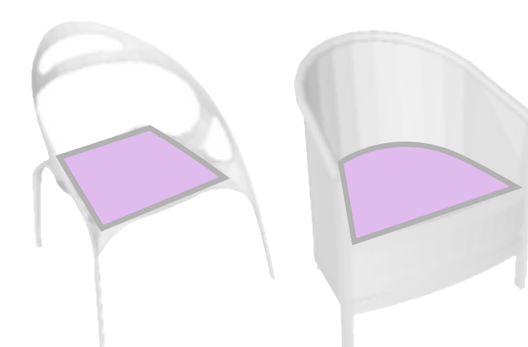
Project Tango

Scale



Point-to-point

Region-to-region

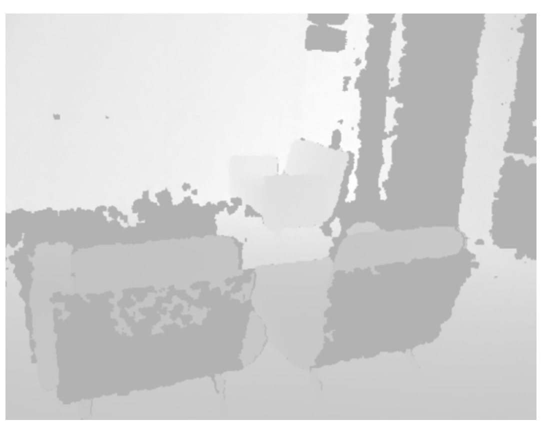


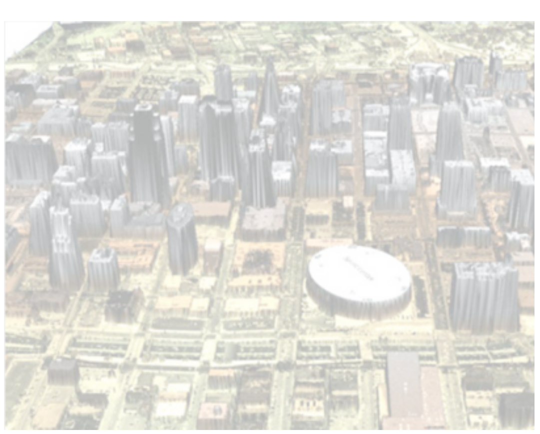
Part-to-part

Ambiguity

Challenges

Discover common structure





Diversity

TurboSquid

3D Warehouse

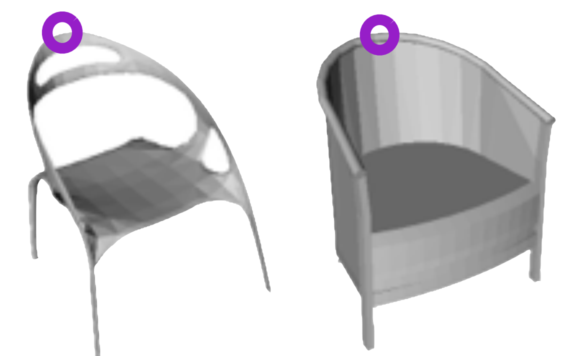
shapeways

Thingiverse

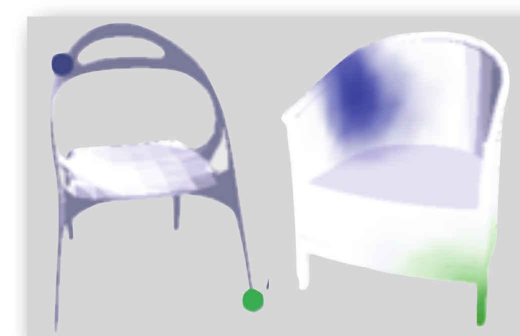
ShapeNet

Project Tango

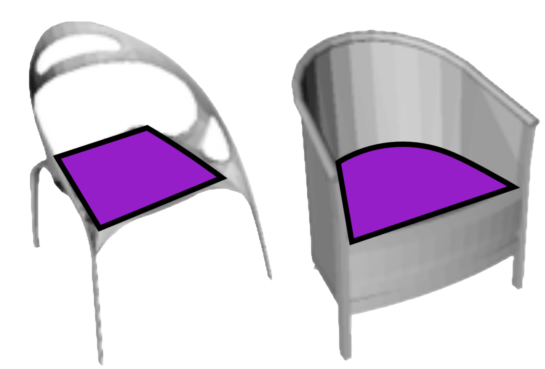
Scale



Point-to-point



Region-to-region

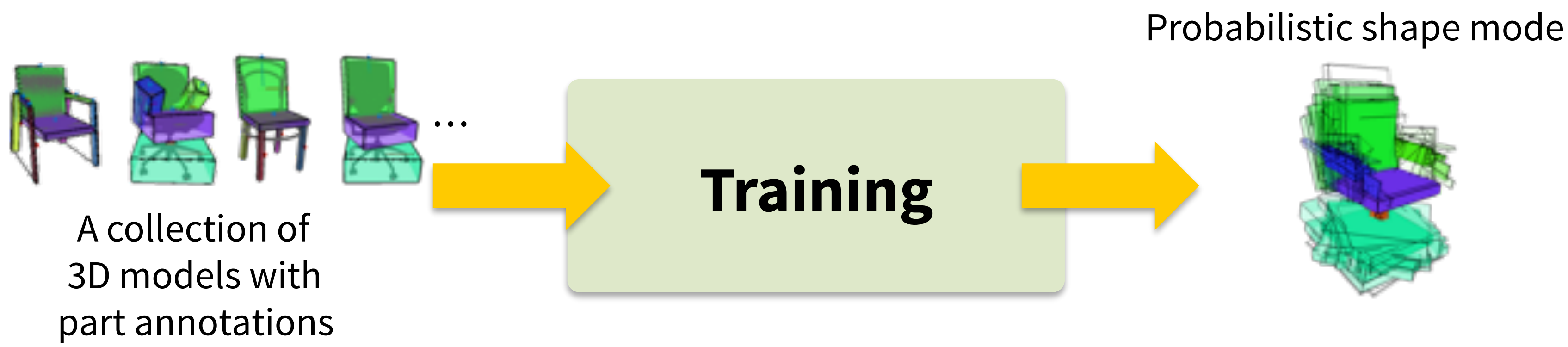


Part-to-part

Ambiguity

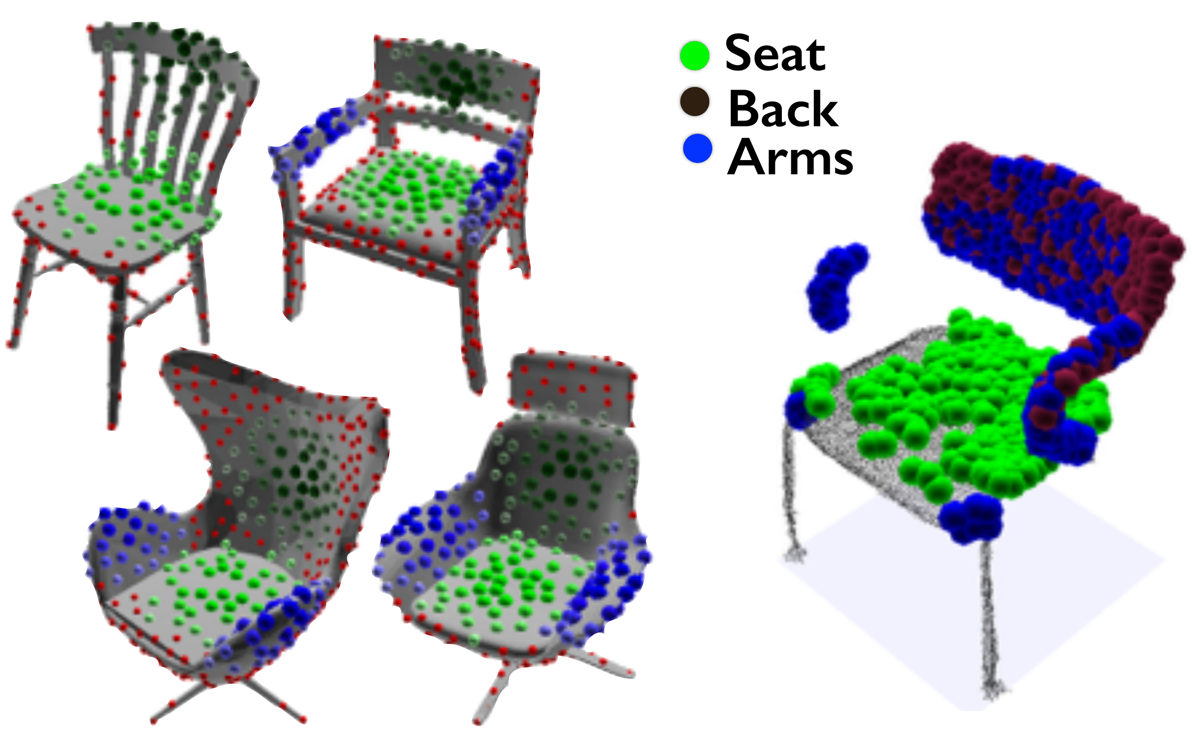
Part-based Template

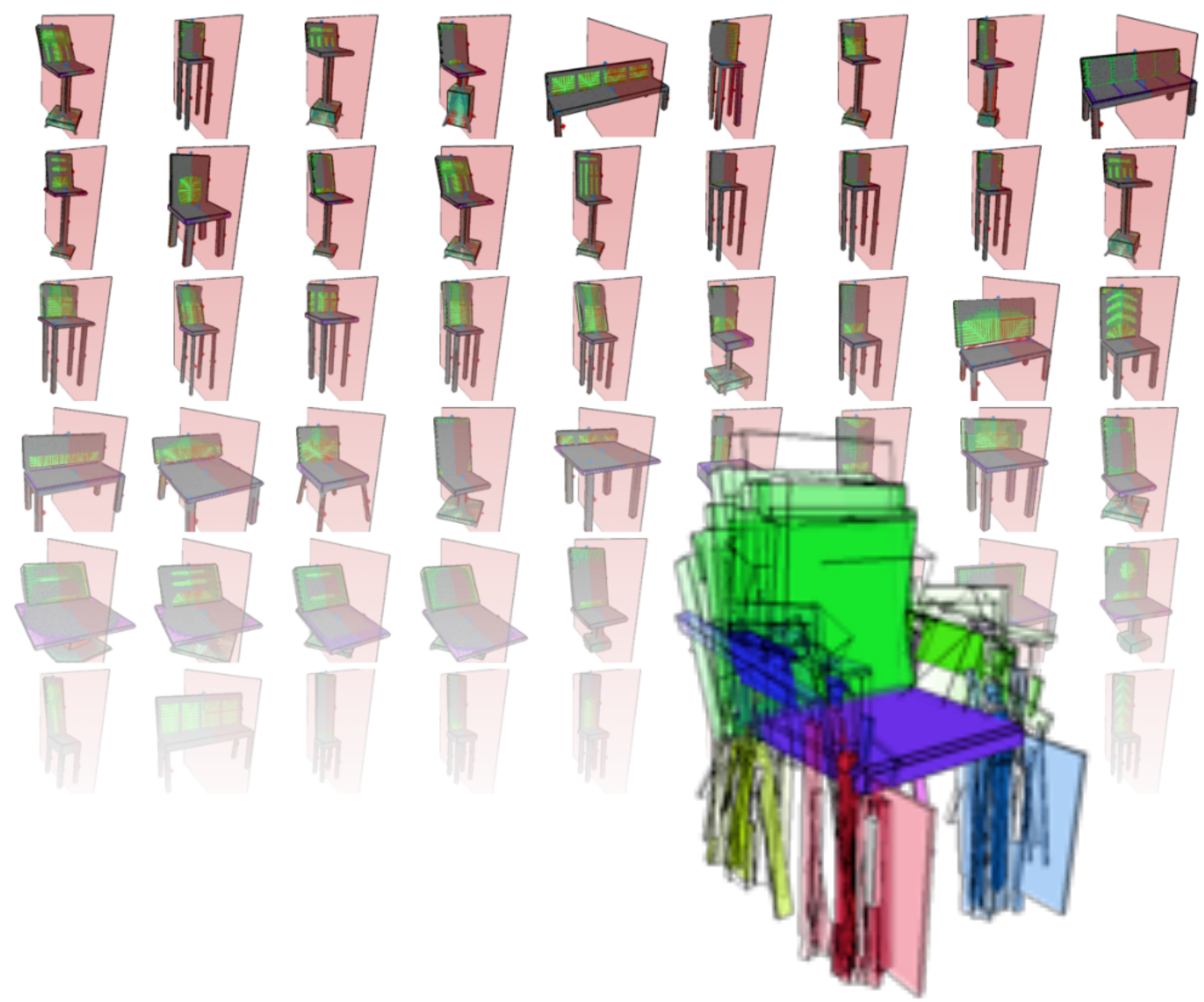
Template is learned from segmented objects



Part-based Template

Template is learned from segmented objects



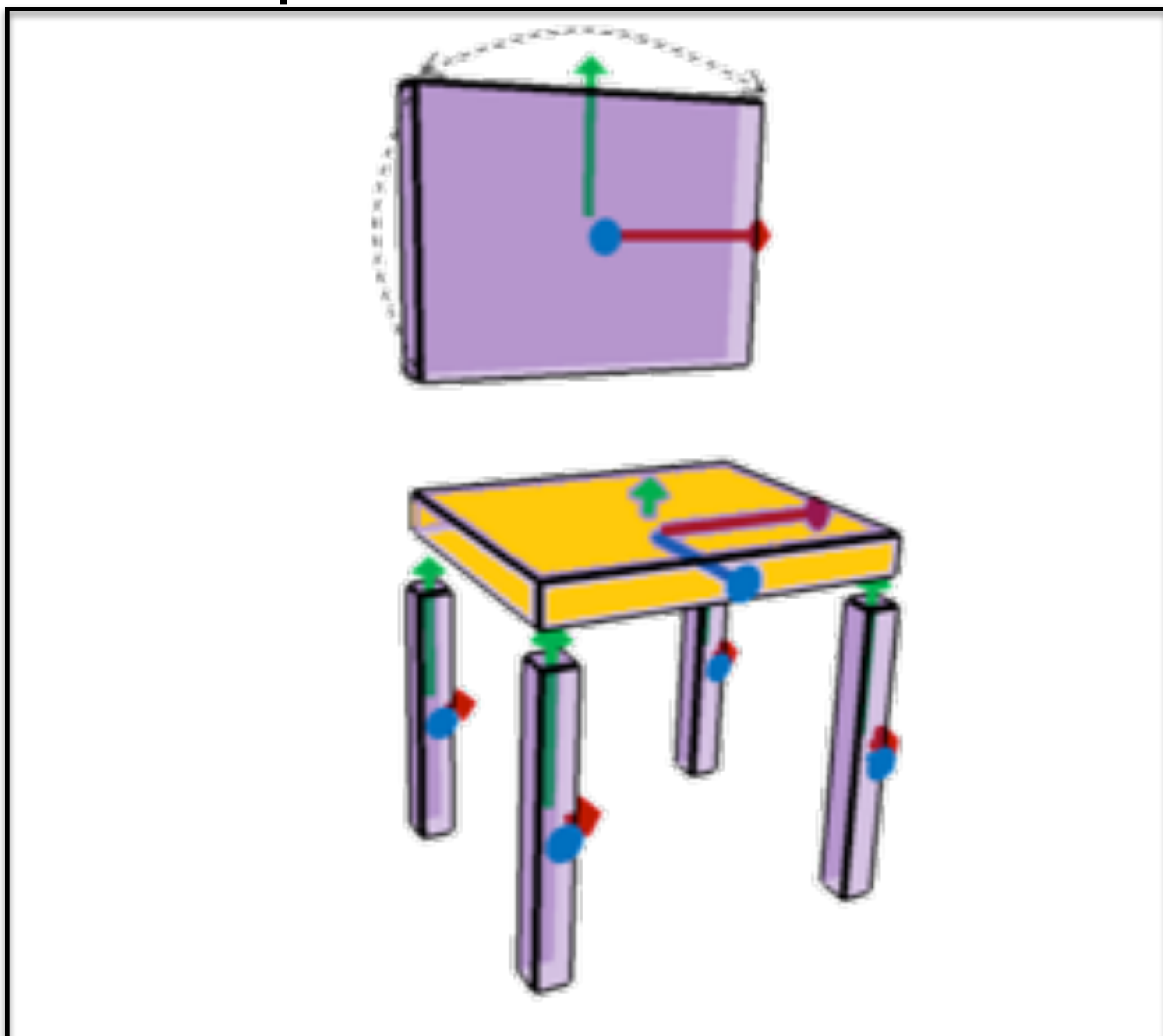


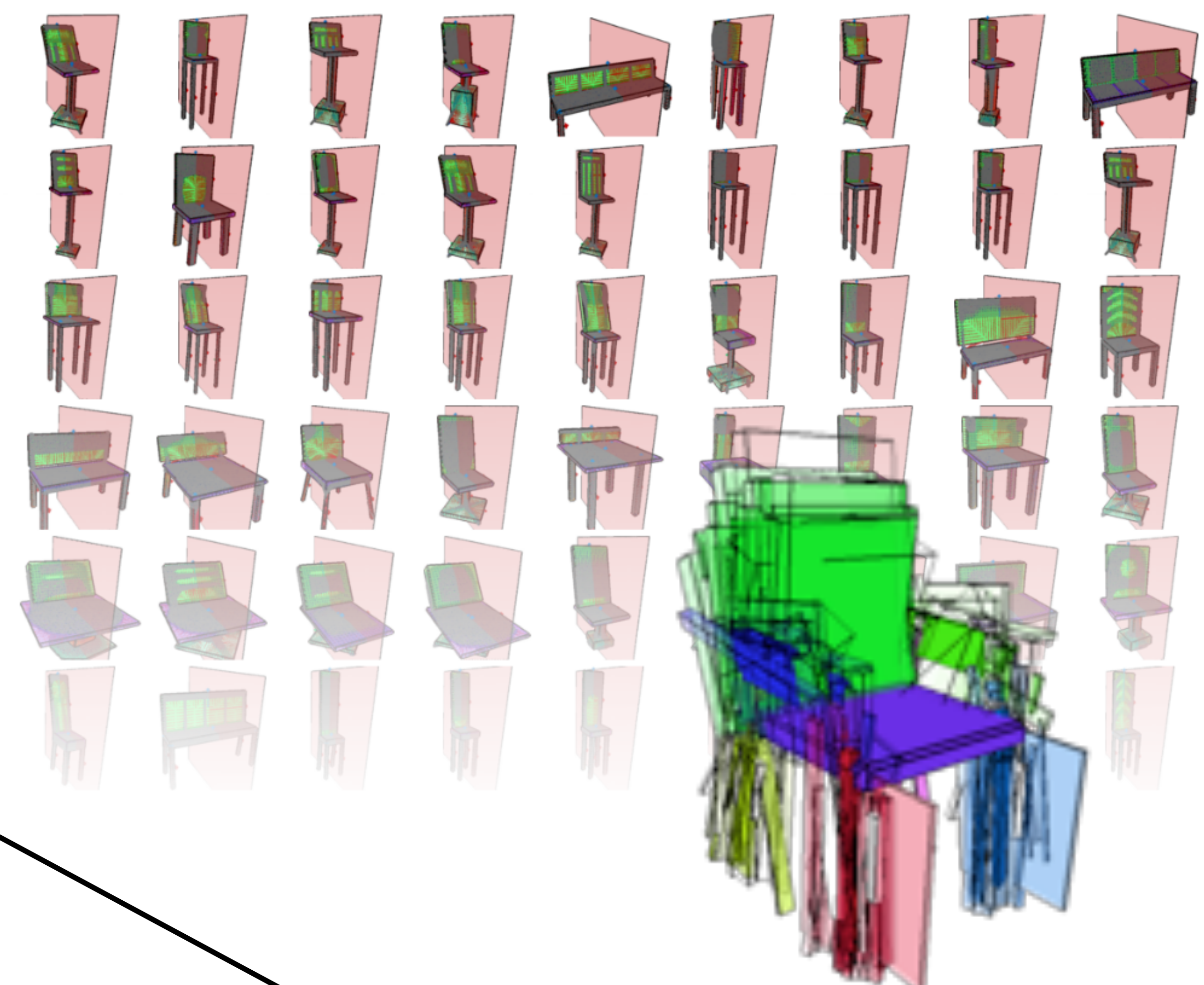
Per-point classifiers

Pairwise part relations

Part-based Template

Template is learned from segmented objects





Pairwise part relations

Each part is a local coordinate system
Represent all other parts in this coordinate system
Probabilistic distribution (e.g., Gaussian)

Structure Inference

Fitting template for structure inference

➡ **Input**

- Initialization
- Part labels and orientations
- Point segmentation
- Pose optimization
- Additional part candidates

Structure Inference

Fitting template for structure inference

- ➡ **Input**
- Initialization
- Part labels and orientations
- Point segmentation
- Pose optimization
- Additional part candidates

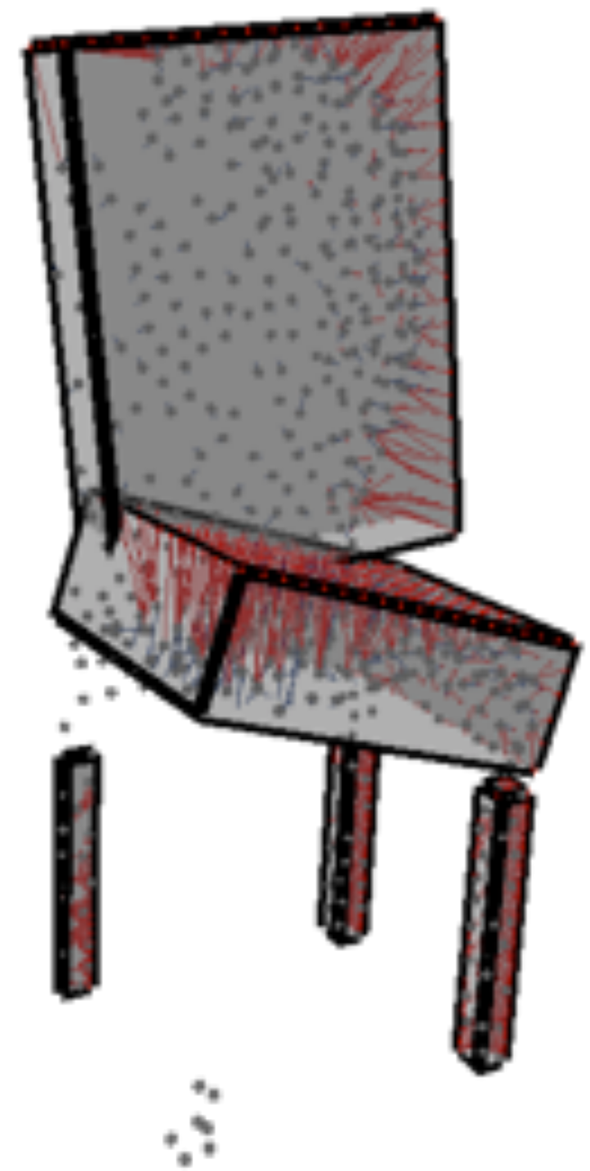
$$E = E_{\mathrm{pnt}} + E_{\mathrm{dist}} + E_{\mathrm{part}}$$

Classifiers	Template distance	Part Relations

Structure Inference

Fitting template for structure inference

- Input
- ➡ **Initialization**
- Part labels and orientations
- Point segmentation
- Pose optimization
- Additional part candidates



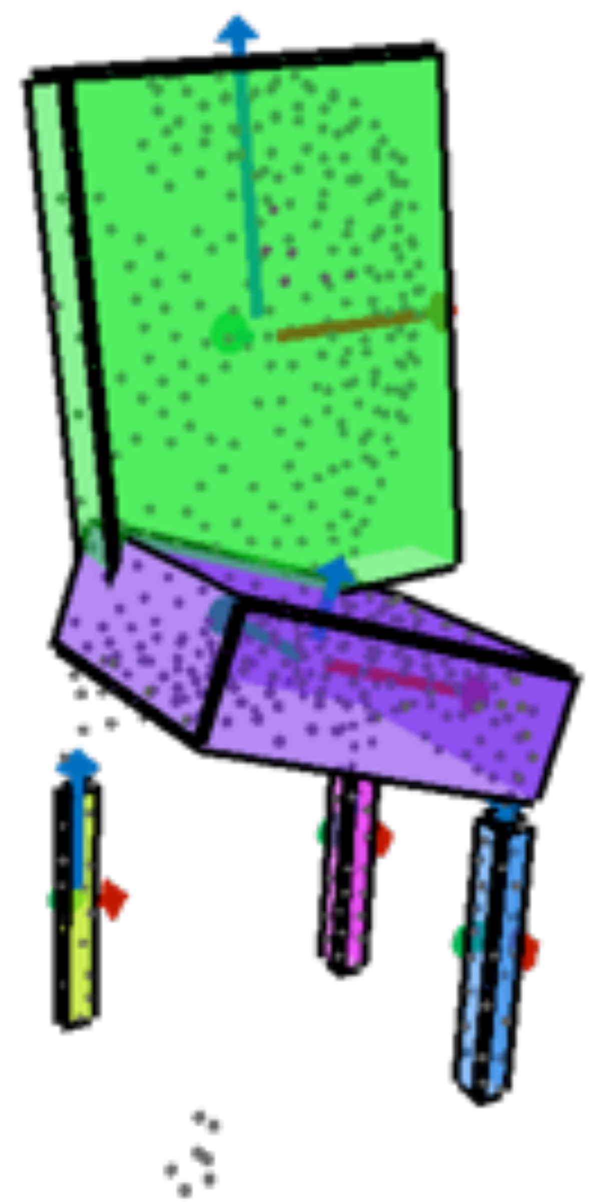
Opt: Clustering

$$E = \boxed{E_{\text{pnt}}} + E_{\text{dist}} + E_{\text{part}}$$

Structure Inference

Fitting template for structure inference

- Input
- Initialization
- ➡ **Part labels and orientations**
- Point segmentation
- Pose optimization
- Additional part candidates



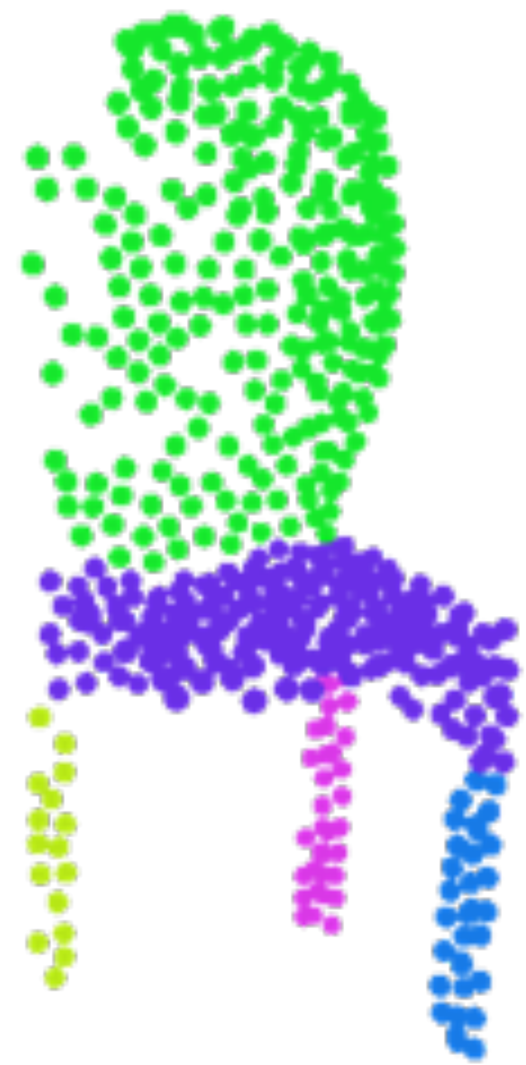
Opt: CRF (at segment level)

$$E = \boxed{E_{\text{pnt}} + E_{\text{dist}} + E_{\text{part}}}$$

Structure Inference

Fitting template for structure inference

- Input
- Initialization
- Part labels and orientations
- ➡ **Point segmentation**
- Pose optimization
- Additional part candidates



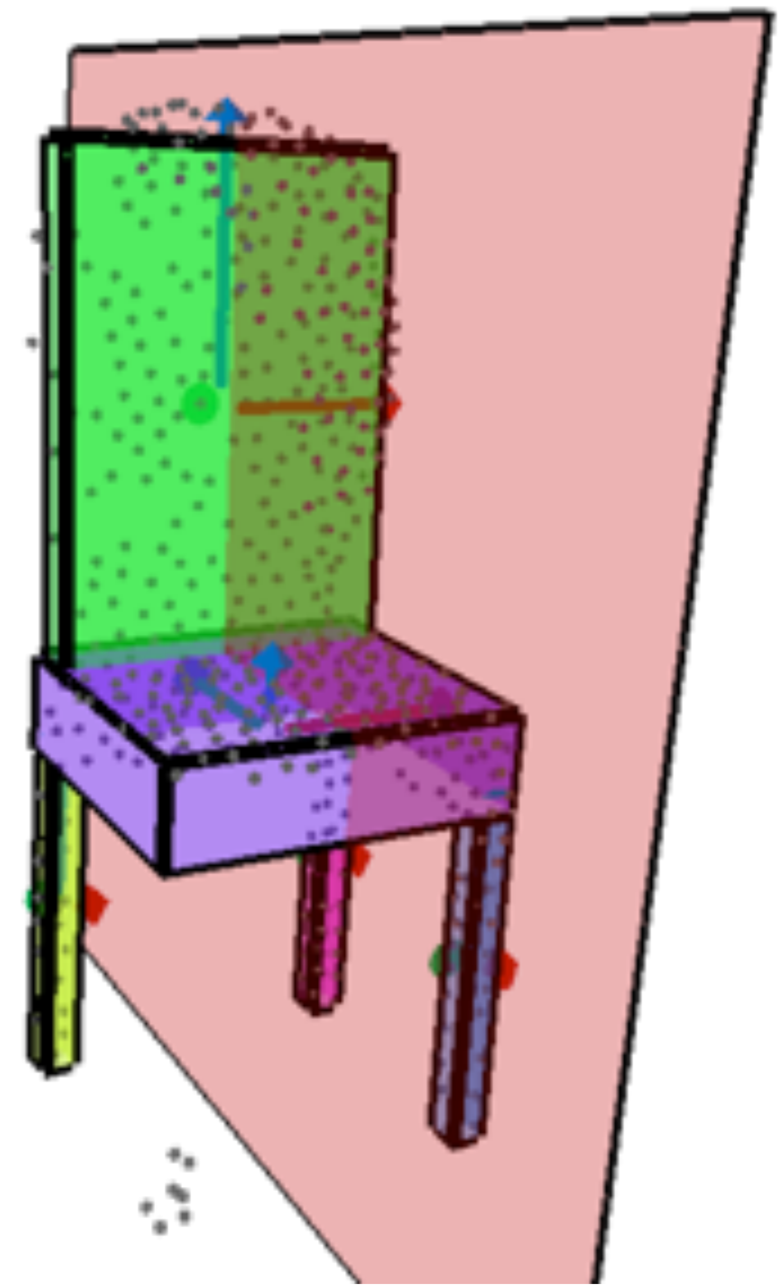
Opt: CRF (at point level)

$$E = \boxed{E_{\text{pnt}} + E_{\text{dist}}} + E_{\text{part}}$$

Structure Inference

Fitting template for structure inference

- Input
- Initialization
- Part labels and orientations
- Point segmentation
- ➡ **Pose optimization**
- Additional part candidates



Opt: Gradient Descent

$$E = E_{\mathrm{pnt}} + \boxed{E_{\mathrm{dist}} + E_{\mathrm{part}}}$$

Structure Inference

Fitting template for structure inference

- Input
- Initialization
- Part labels and orientations
- Point segmentation
- Pose optimization

➡ **Additional part candidates**

Opt: greedy sampling

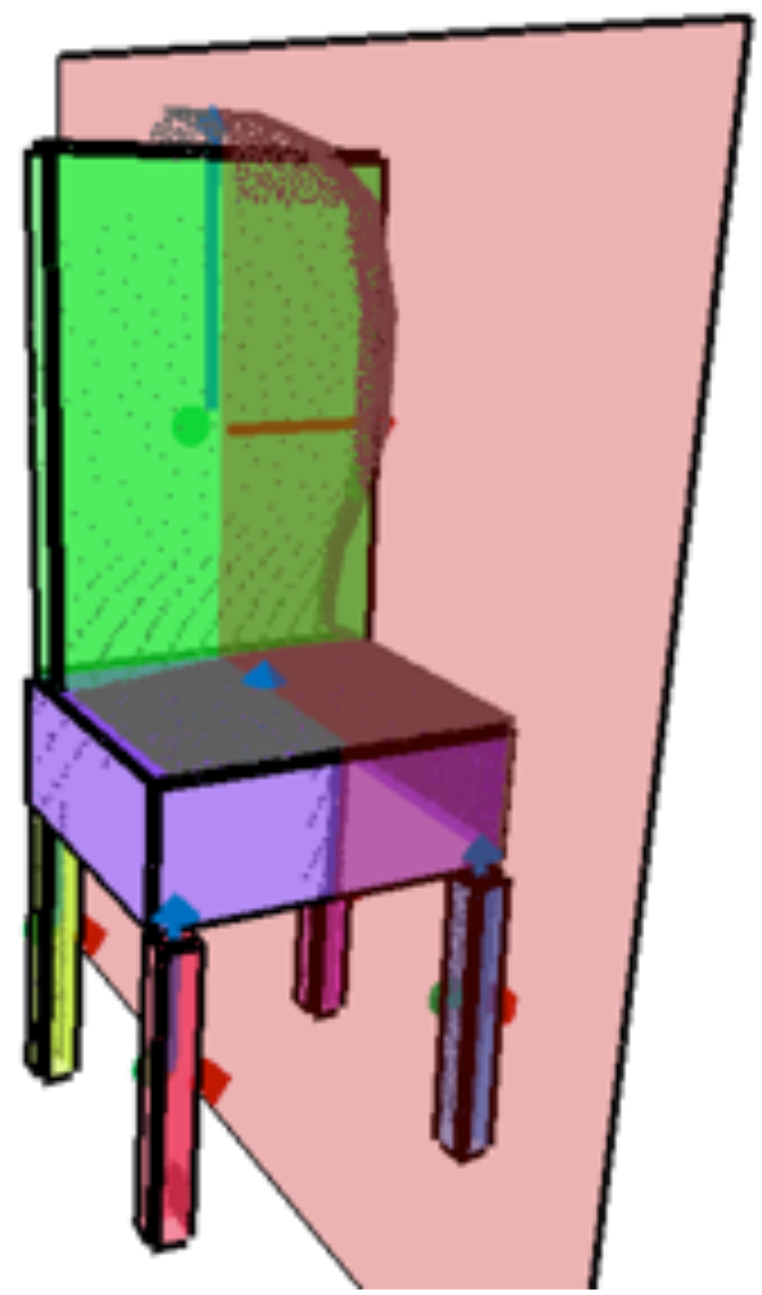
$$E = E_{\text{pnt}} + E_{\text{dist}} + \boxed{E_{\text{part}}}$$

Structure Inference

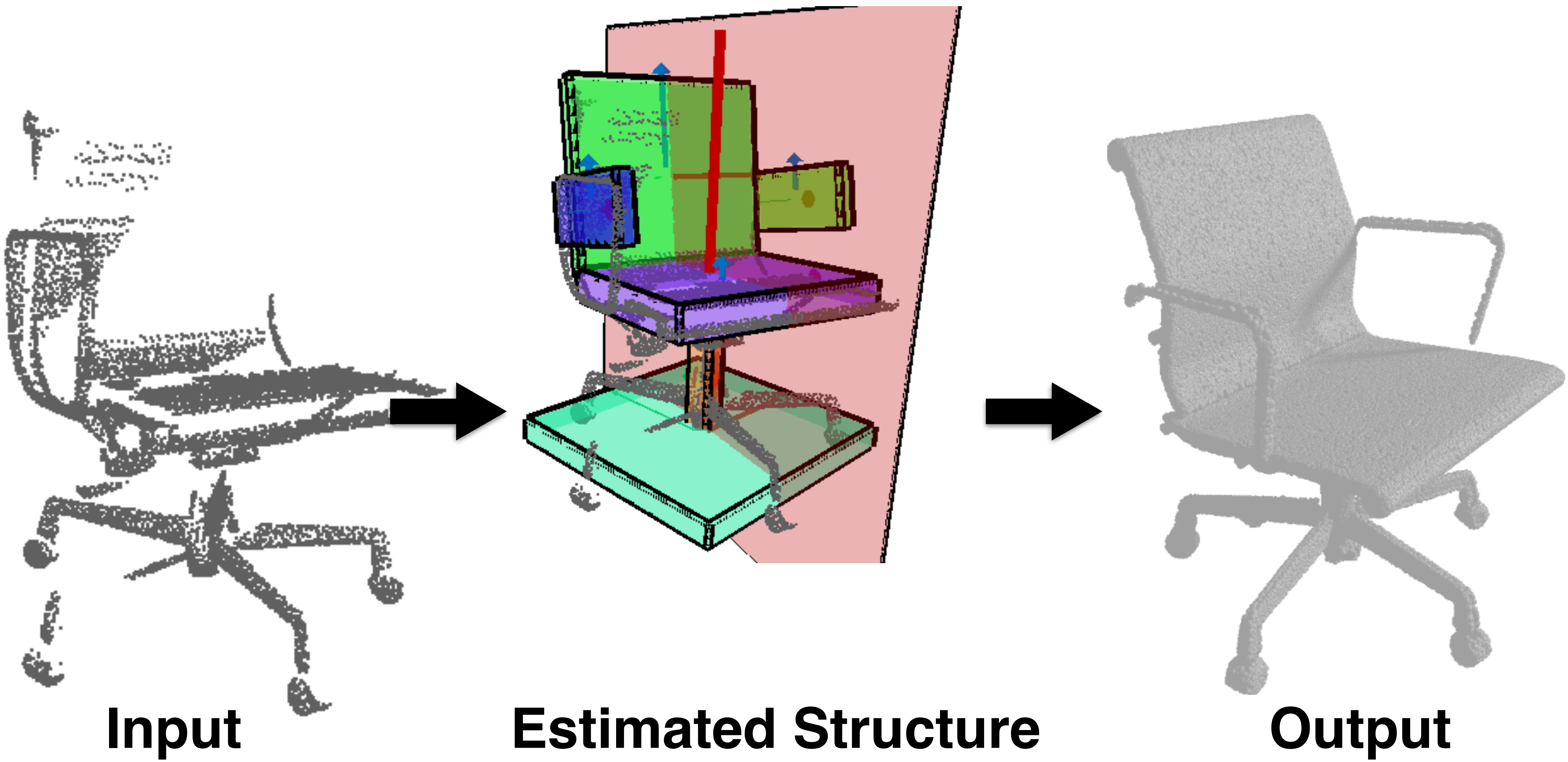
Fitting template for structure inference

- Input
- Initialization
- Part labels and orientations
- Point segmentation
- Pose optimization
- Additional part candidates

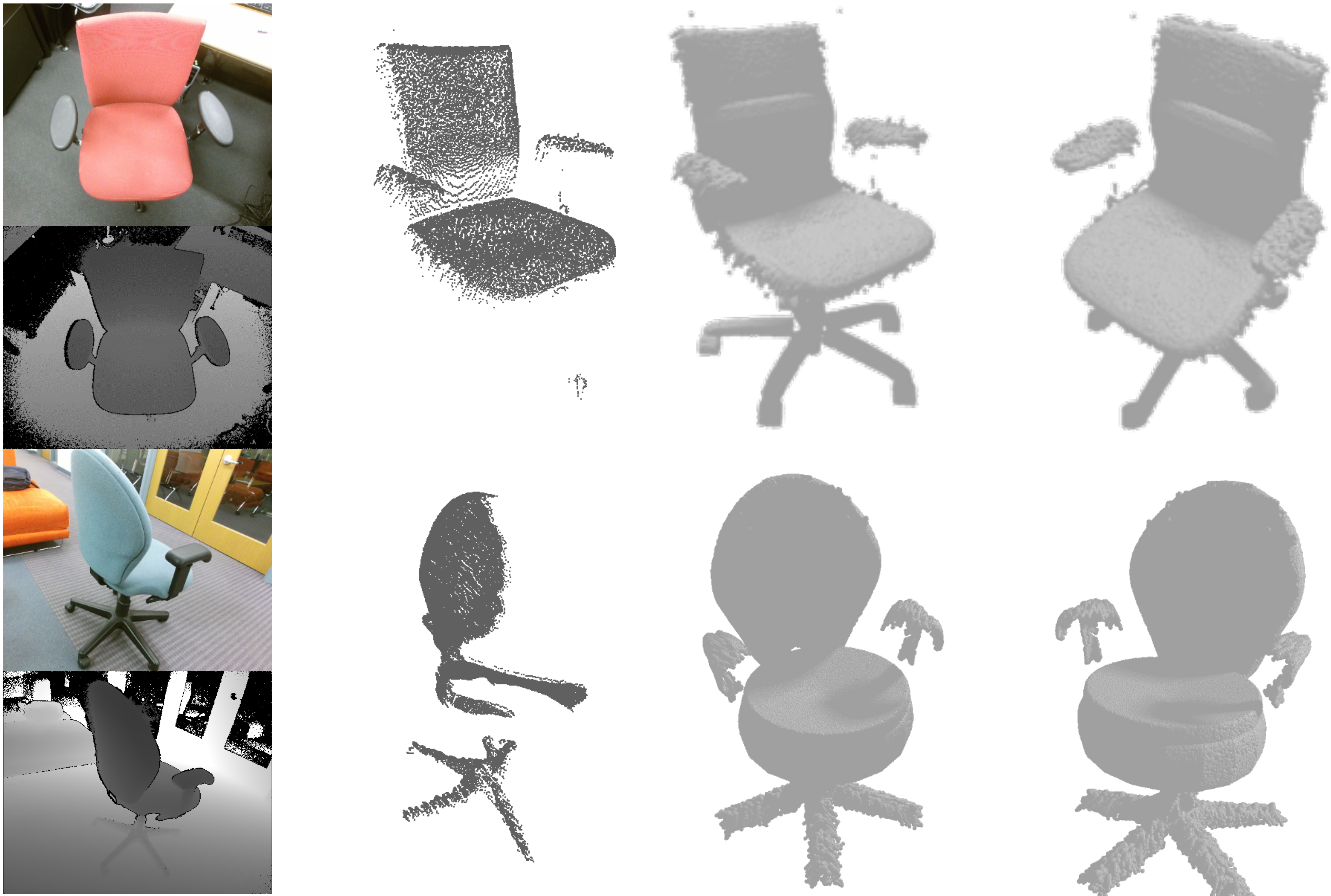
Final Result:



Shape Completion



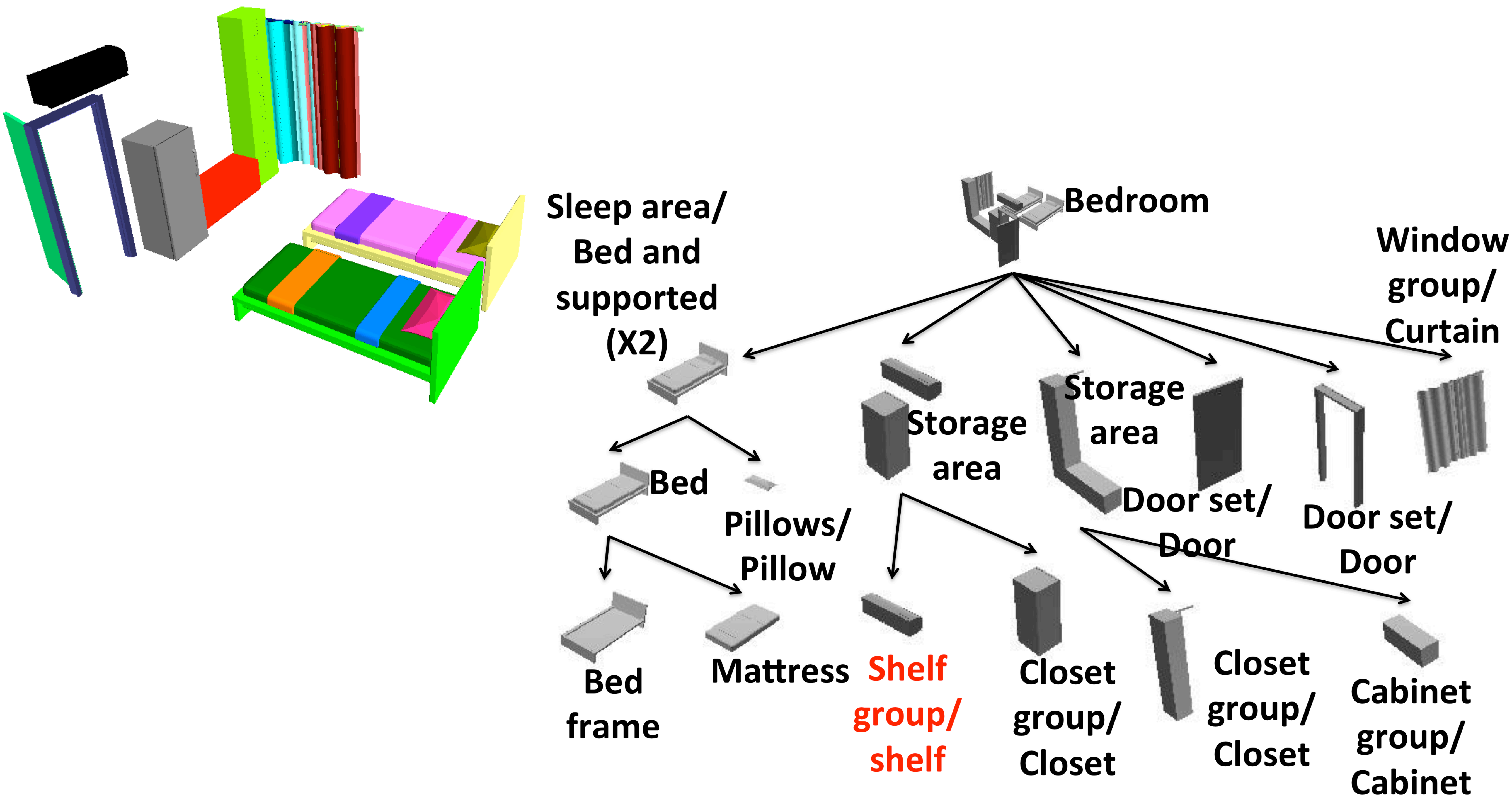
Shape Completion



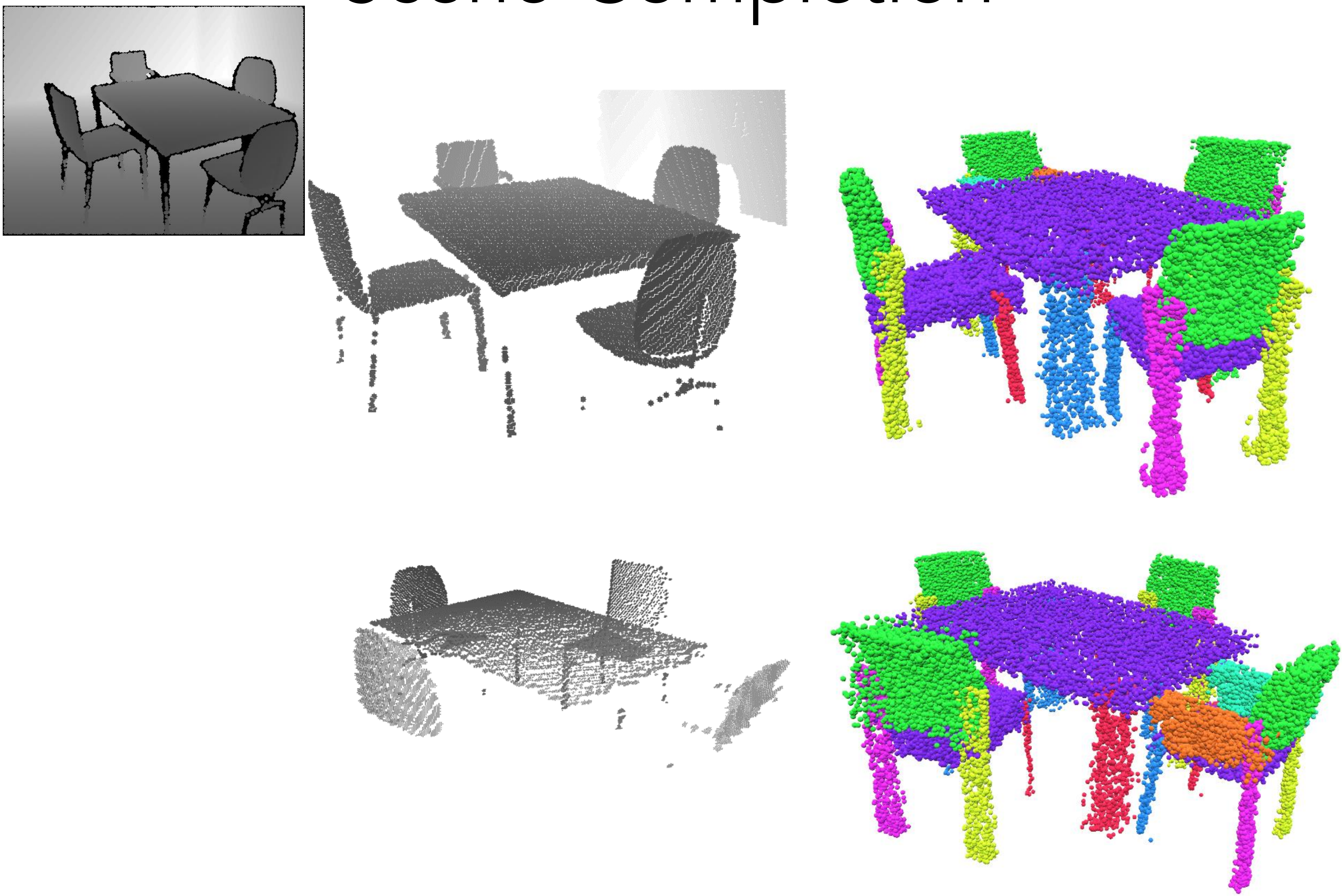
Shape Completion

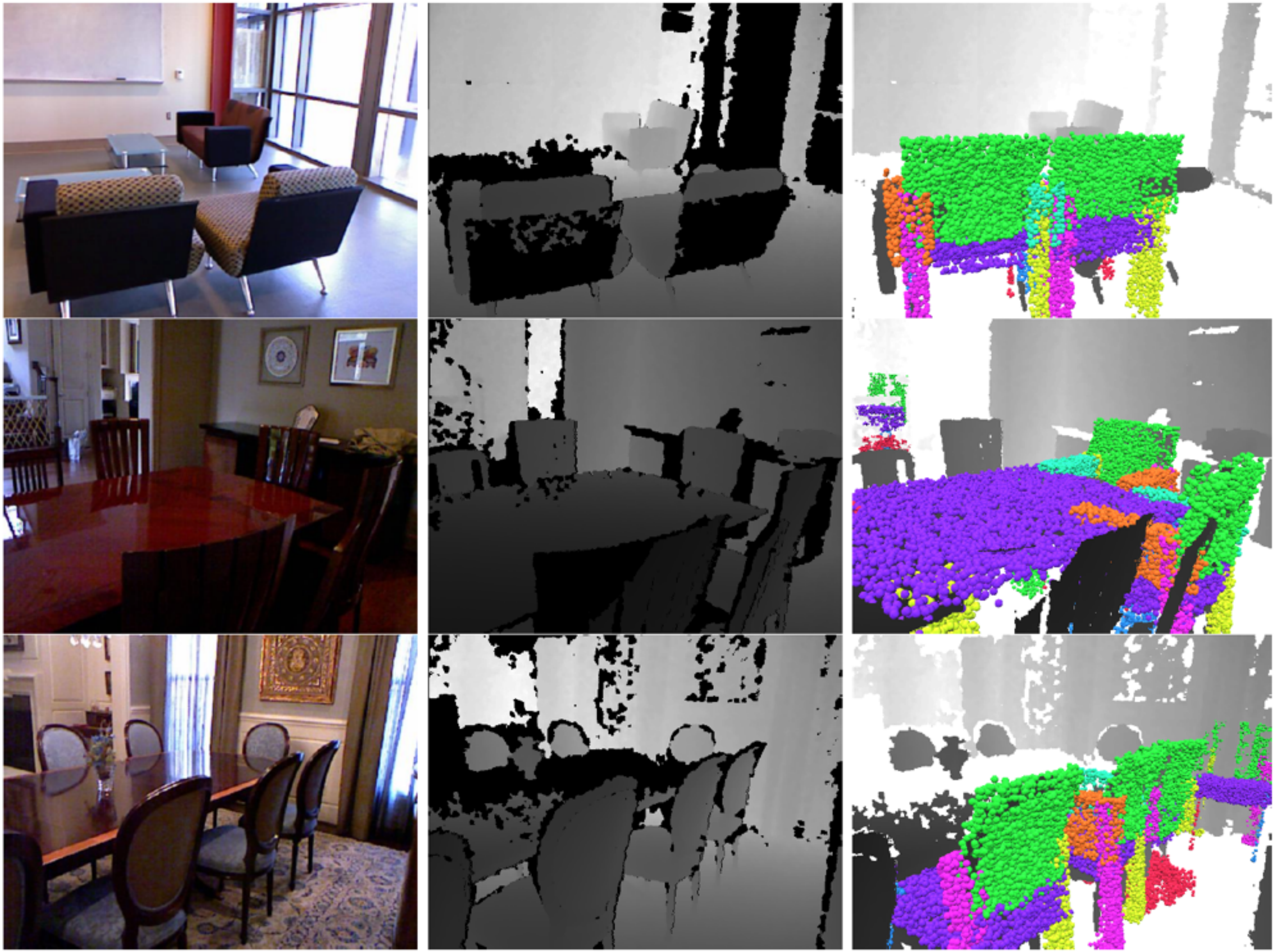
Hierarchical Templates

Use probabilistic grammars for hierarchies

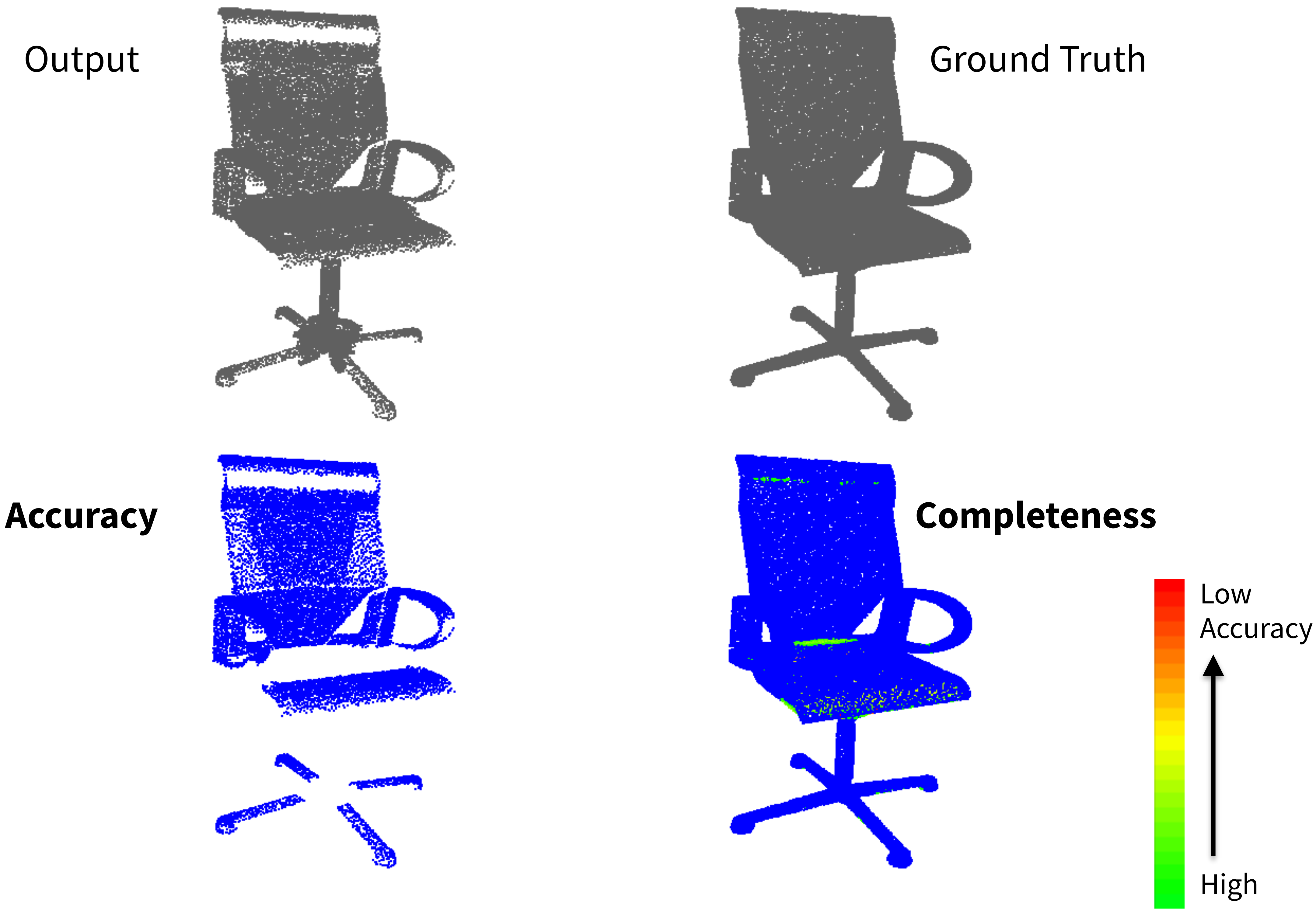


Scene Completion





Evaluation



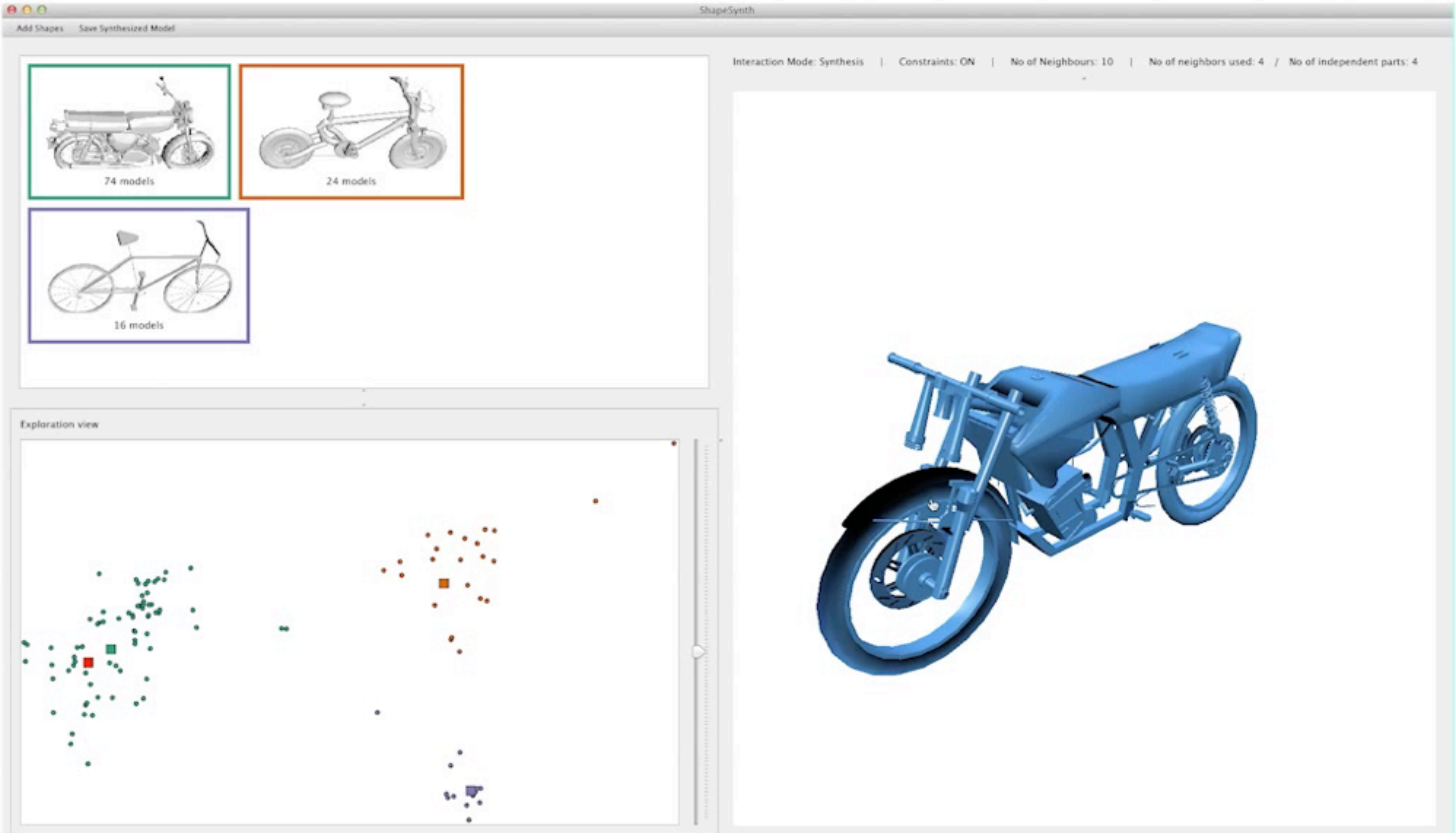
Output
Ground Truth
Accuracy
Completeness
Low
Accuracy
High

Comparisons

	Accuracy	Completeness
Shape Completion		
Ours	85%	**78%**
[Podolak'06] (symmetry)	**99%**	69%
[Shen'12] (closest parts)	62%	60%
Scene Completion		
Ours	**62%**	**82%**
[Gupta'15] (closest object)	57%	52%

Exploration & Synthesis

Using part-based templates for synthesis


ShapeSynth
Add Shapes
Save Synthesized Model
Interaction Mode: Synthesis | Constraints: ON | No of Neighbours: 10 | No of neighbors used: 4 / No of independent parts: 4
74 models
24 models
16 models
Exploration view

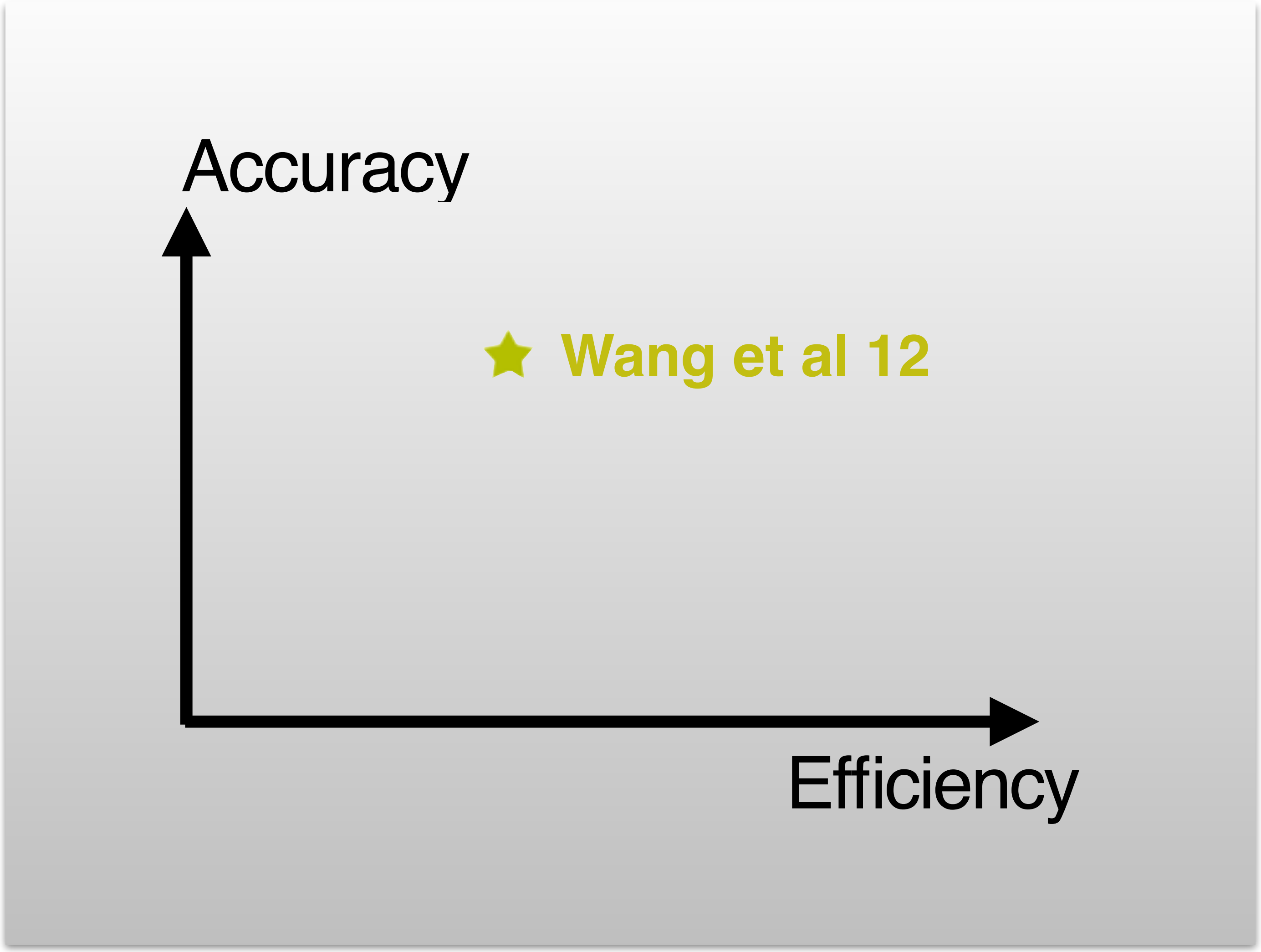
Training Data

Where do we get segmented shapes?

Training Data

Where do we get segmented shapes?

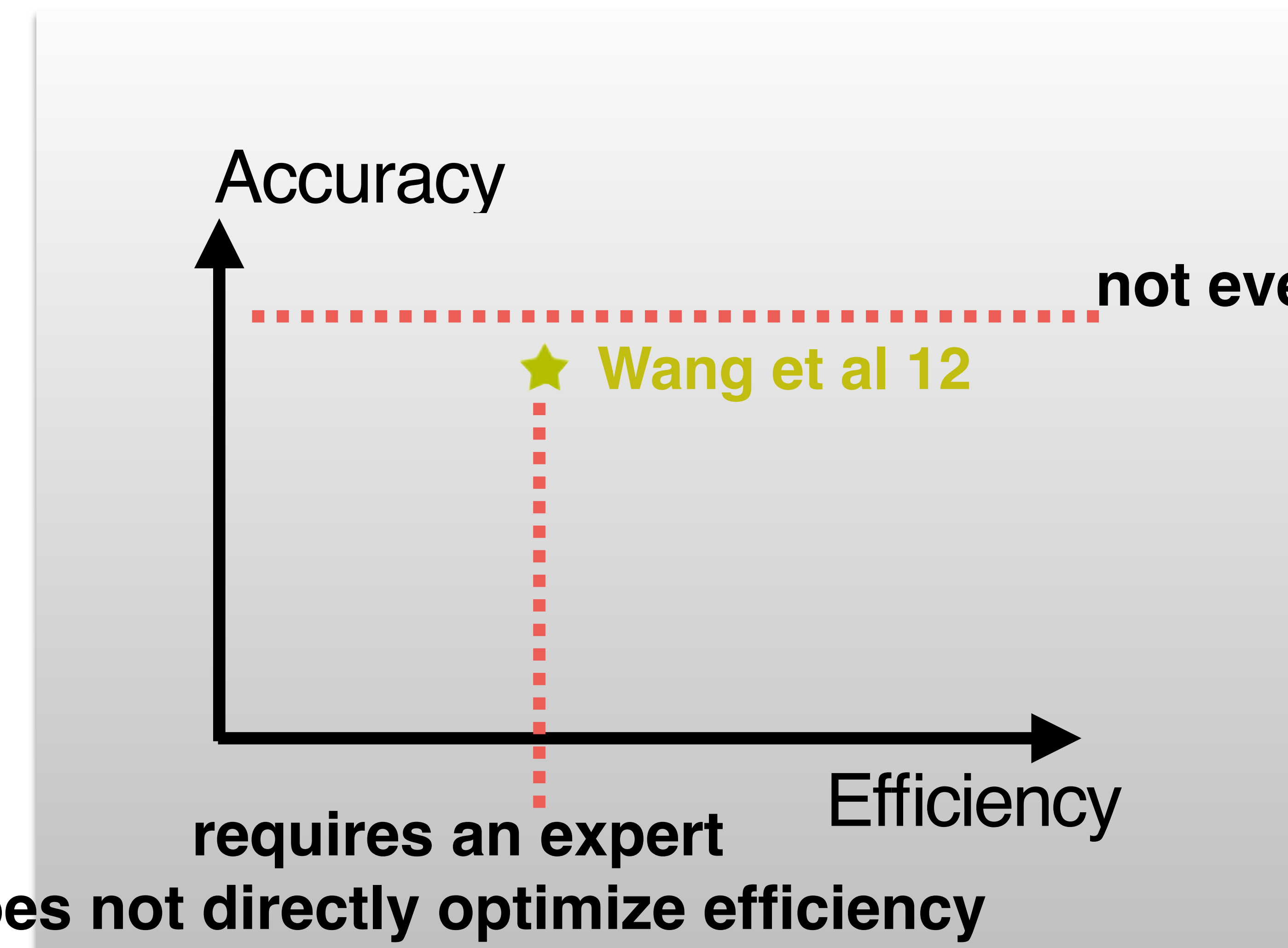
- Active learning for labeling [Wang'12]
    - Dataset: 1090 meshes, 2-5 parts per mesh, watertight manifold



Training Data

Where do we get segmented shapes?

- Active learning for labeling [Wang'12]
  - Dataset: 1090 meshes, 2-5 parts per mesh, watertight manifold



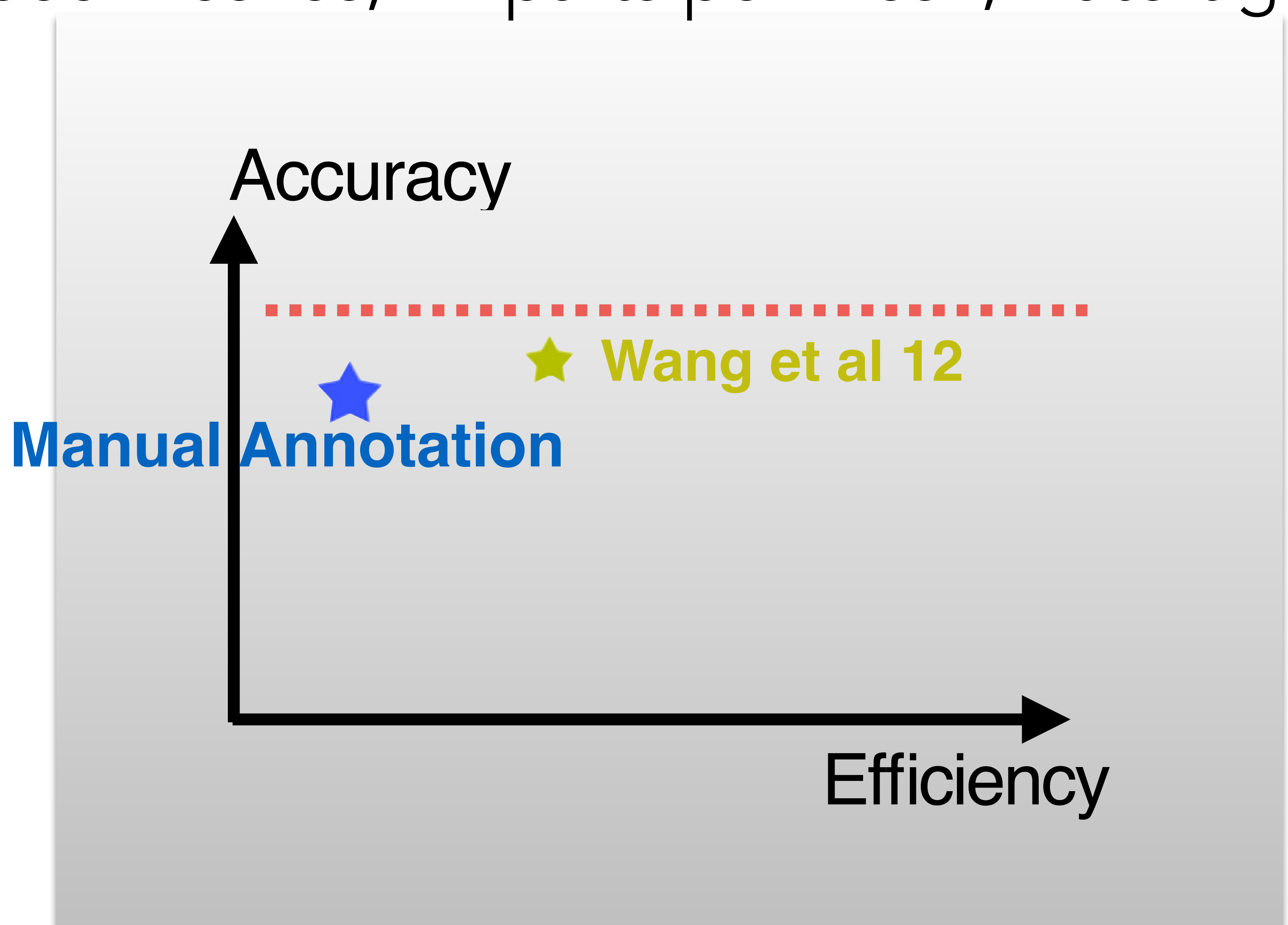
not every model is verified by a human

does not directly optimize efficiency
(active learning objective: train the best classifier)

Training Data

Where do we get segmented shapes?

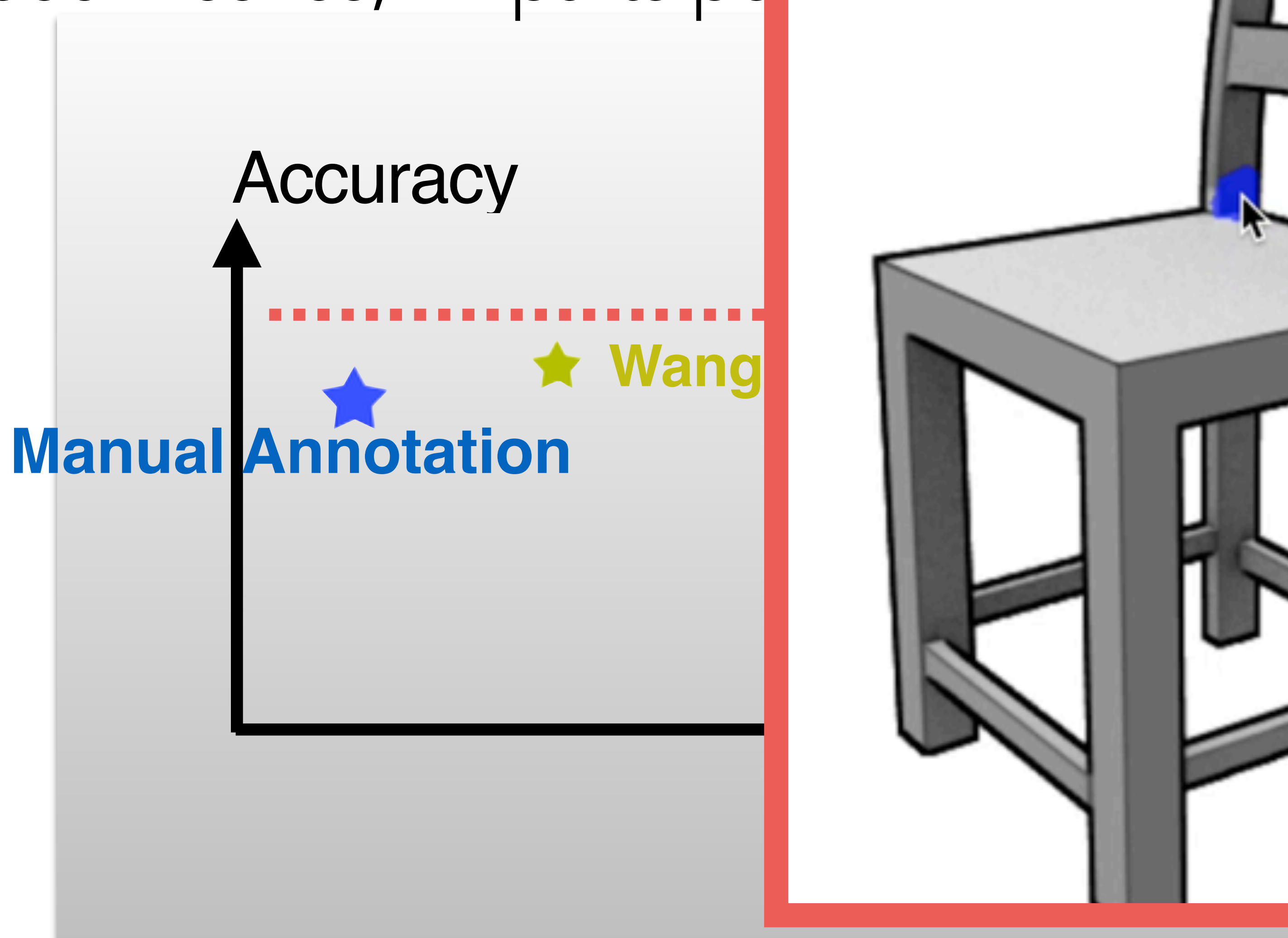
- Active learning for labeling [Wang'12]
  - Dataset: 1090 meshes, 2-5 parts per mesh, watertight manifold
- Manual annotation [Chen'09]
  - Dataset: 380 meshes, 19 parts per mesh, watertight manifold

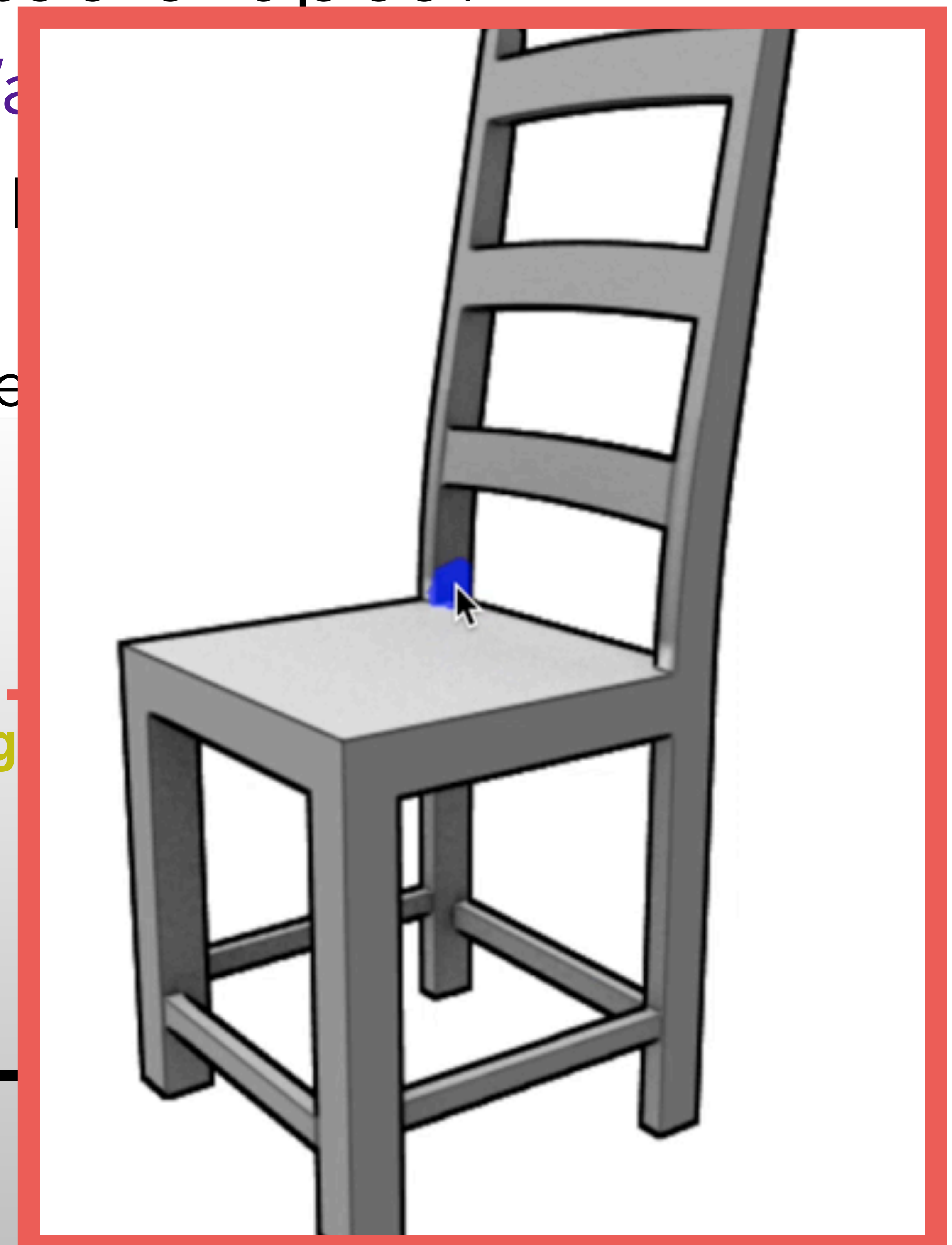


Training Data

Where do we get segmented shapes?

- Active learning for labeling [Wa
  - Dataset: 1090 meshes, 2-5 parts
- Manual annotation [Chen'09]
  - Dataset: 380 meshes, 19 parts pe



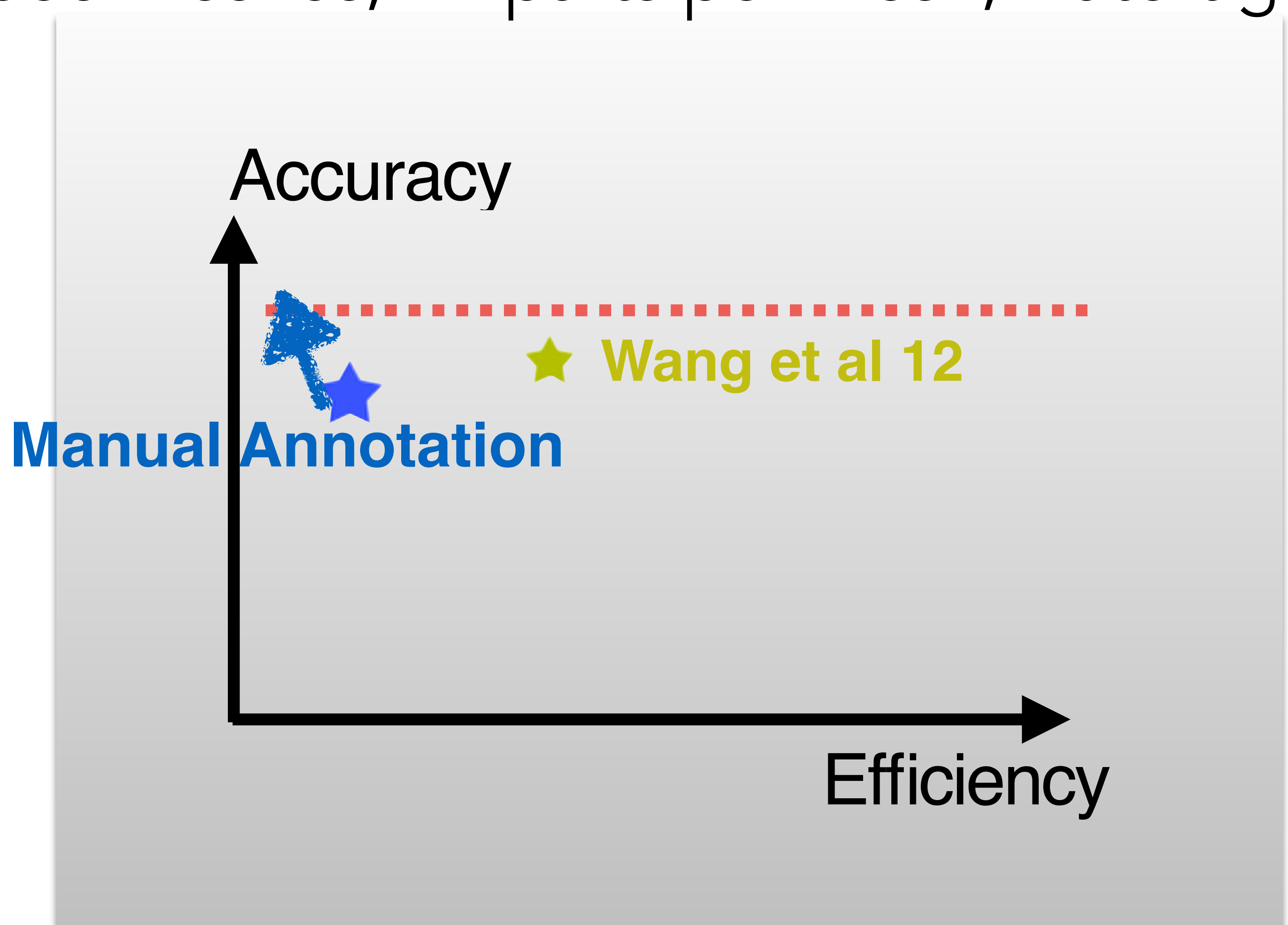


Training Data

Where do we get segmented shapes?

- Active learning for labeling [Wang'12]
  - Dataset: 1090 meshes, 2-5 parts per mesh, watertight manifold
- Manual annotation [Chen'09]
  - Dataset: 380 meshes, 19 parts per mesh, watertight manifold

Verify all results!



Training Data

Where do we get segmented shapes?

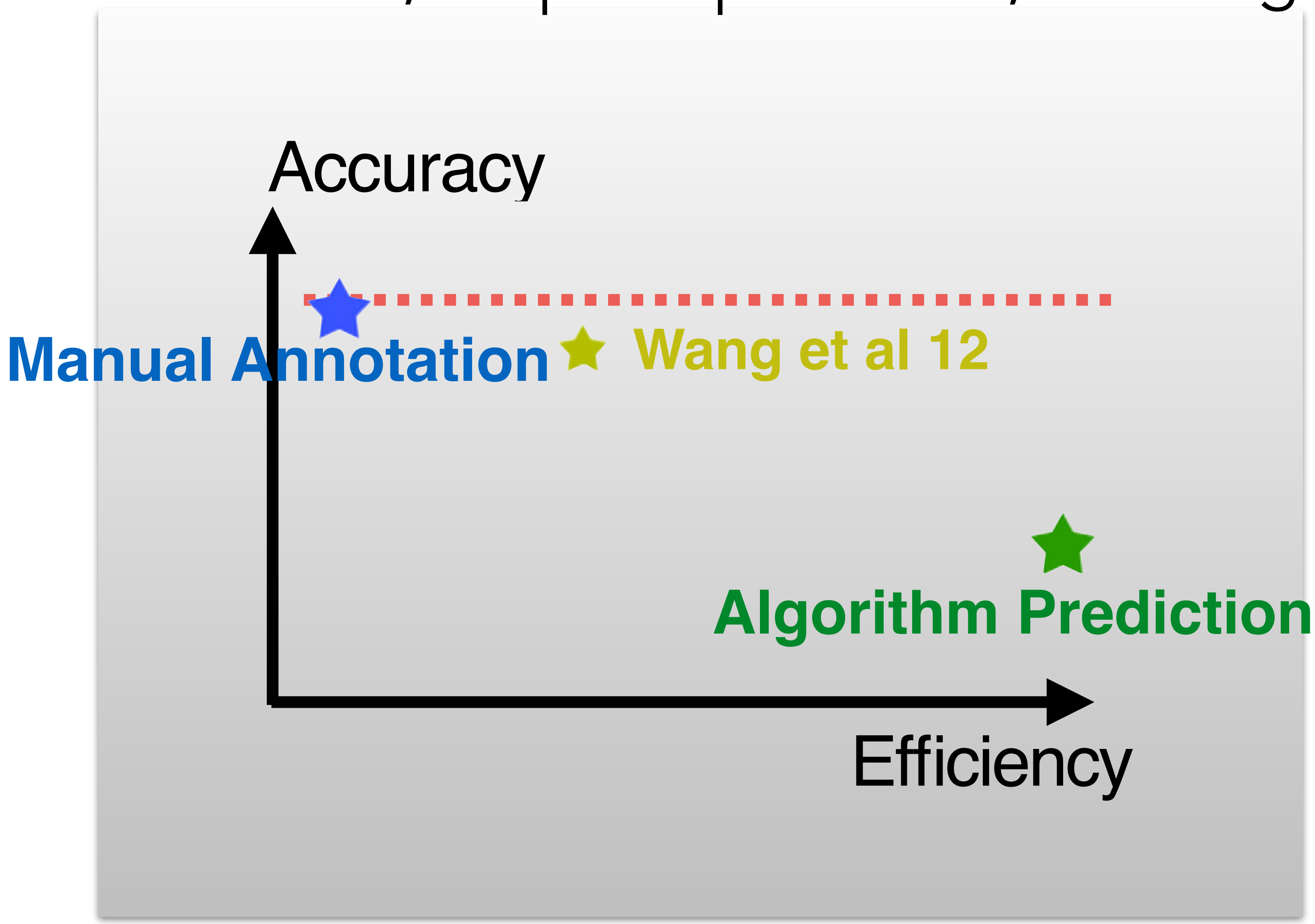
- Active learning for labeling [Wang'12]
  - Dataset: 1090 mesh
- Manual annotation
  - Dataset: 380 meshe

Verify all results!

Training Data

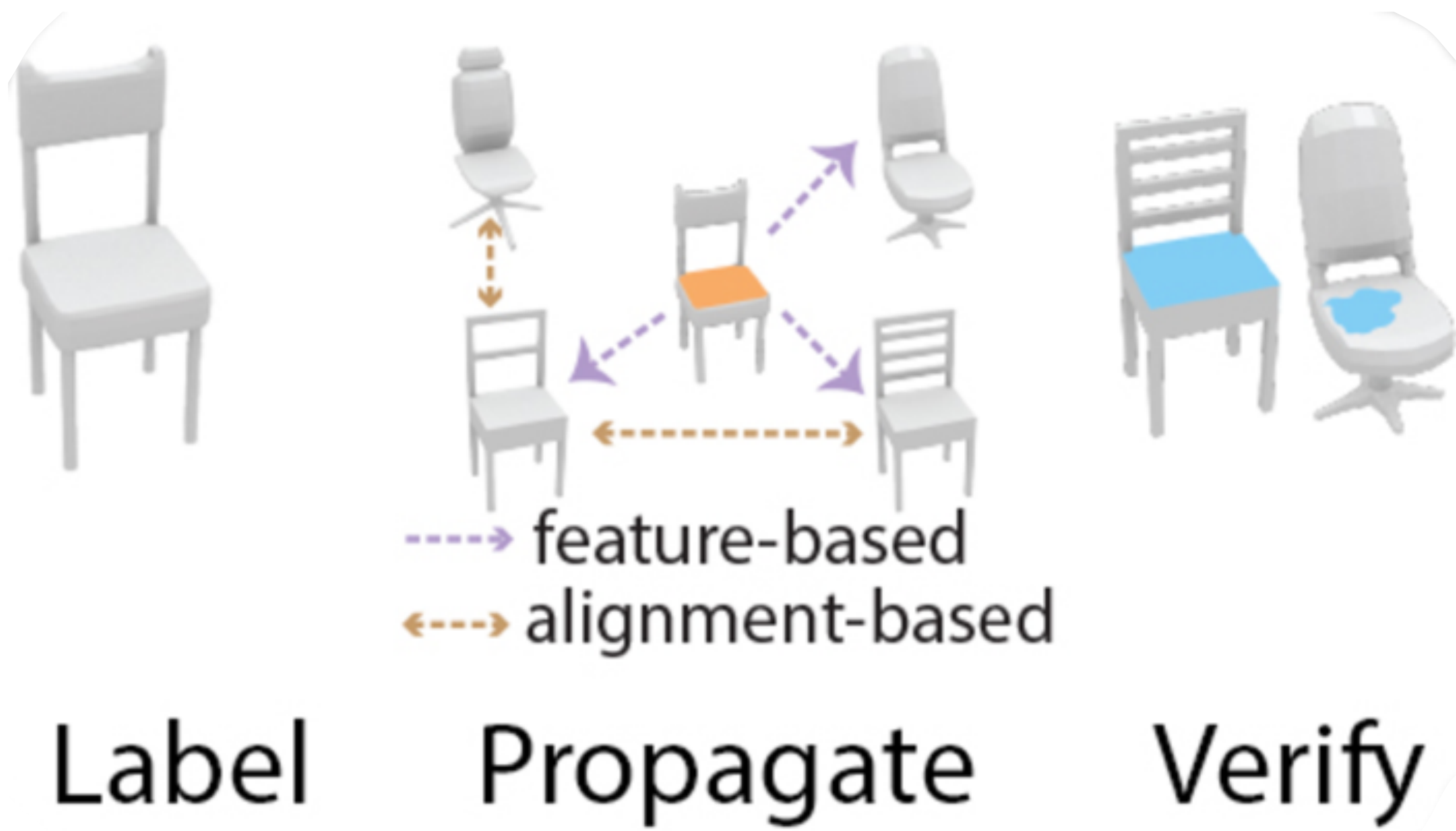
Where do we get segmented shapes?

- Active learning for labeling [Wang'12]
  - Dataset: 1090 meshes, 2-5 parts per mesh, watertight manifold
- Manual annotation [Chen'09]
  - Dataset: 380 meshes, 19 parts per mesh, watertight manifold



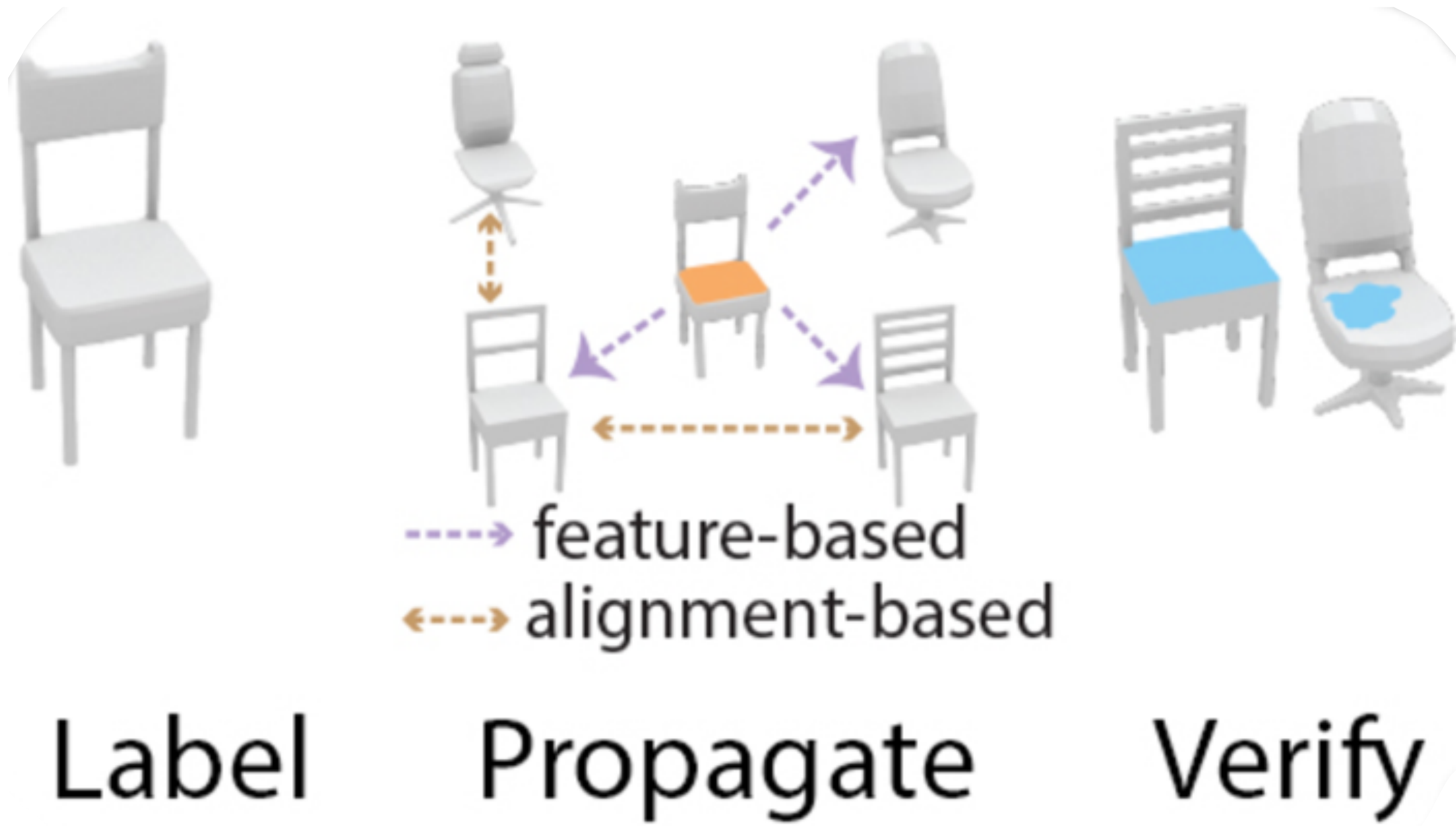
Our Approach

Verify algorithmic predictions



Our Approach

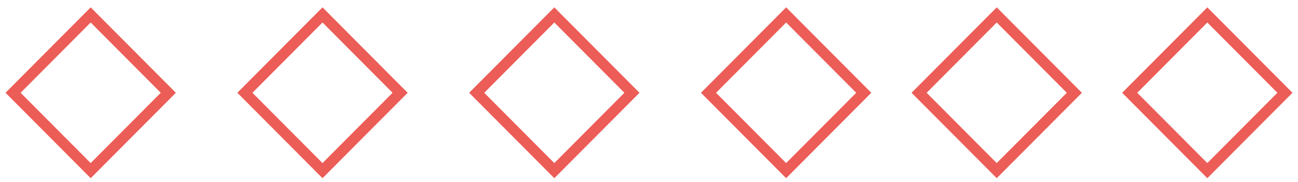
Verify algorithmic predictions



Which models to label & which labels to verify?

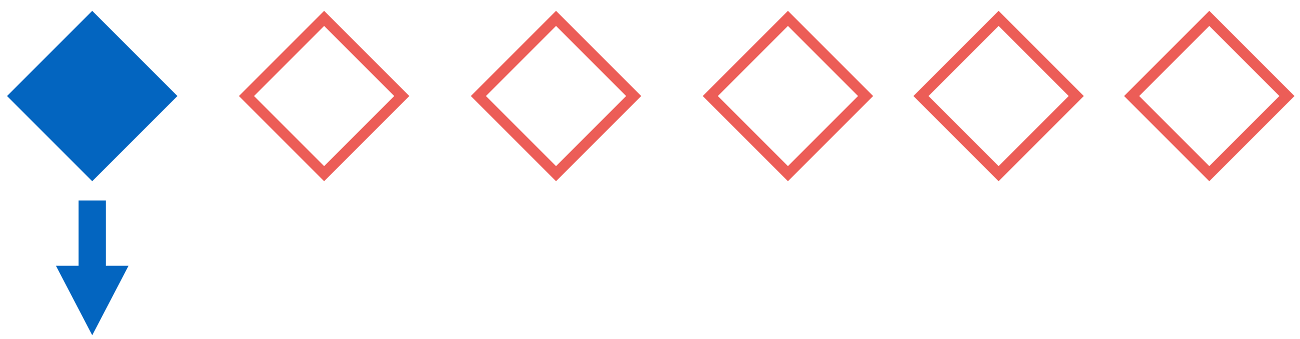
Utility Function

several shapes



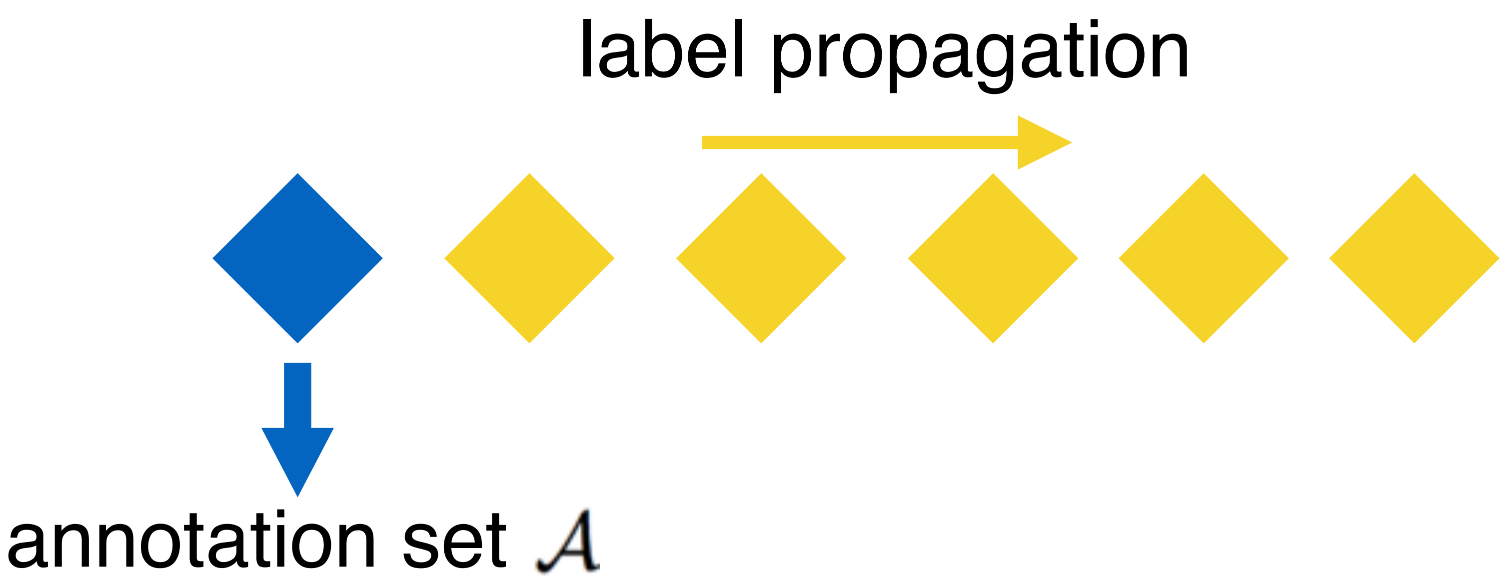
Utility Function

several shapes

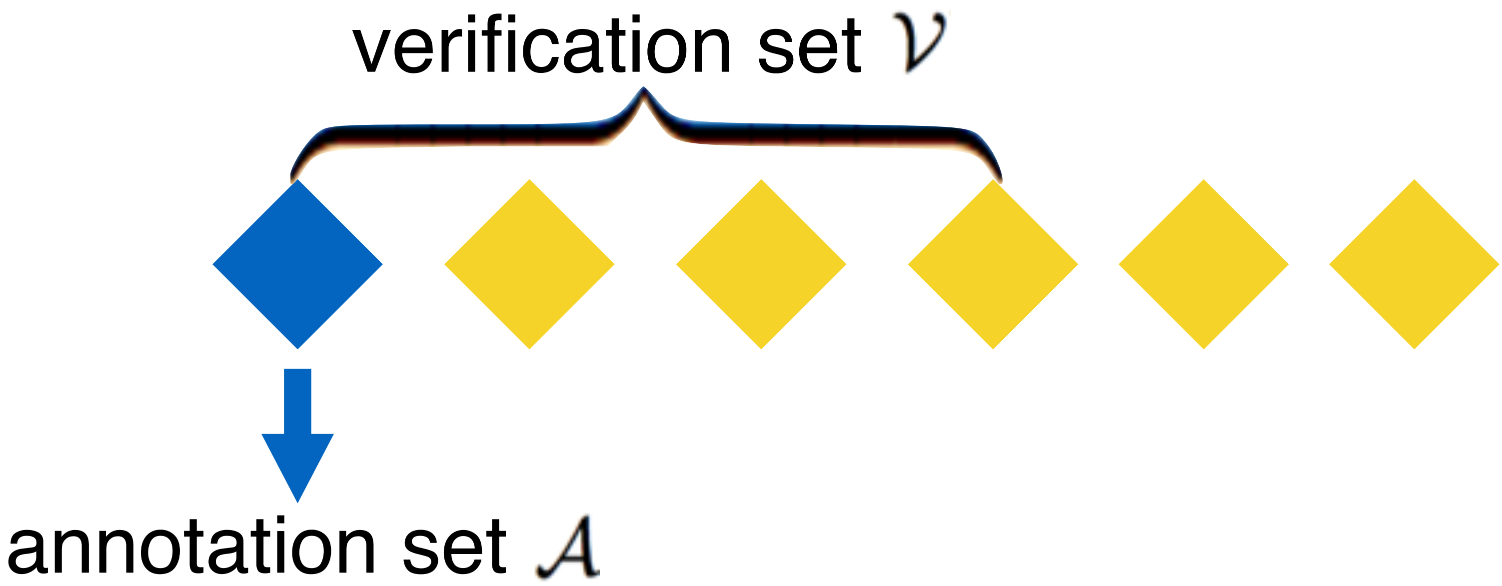


annotation set $\mathcal{A}$

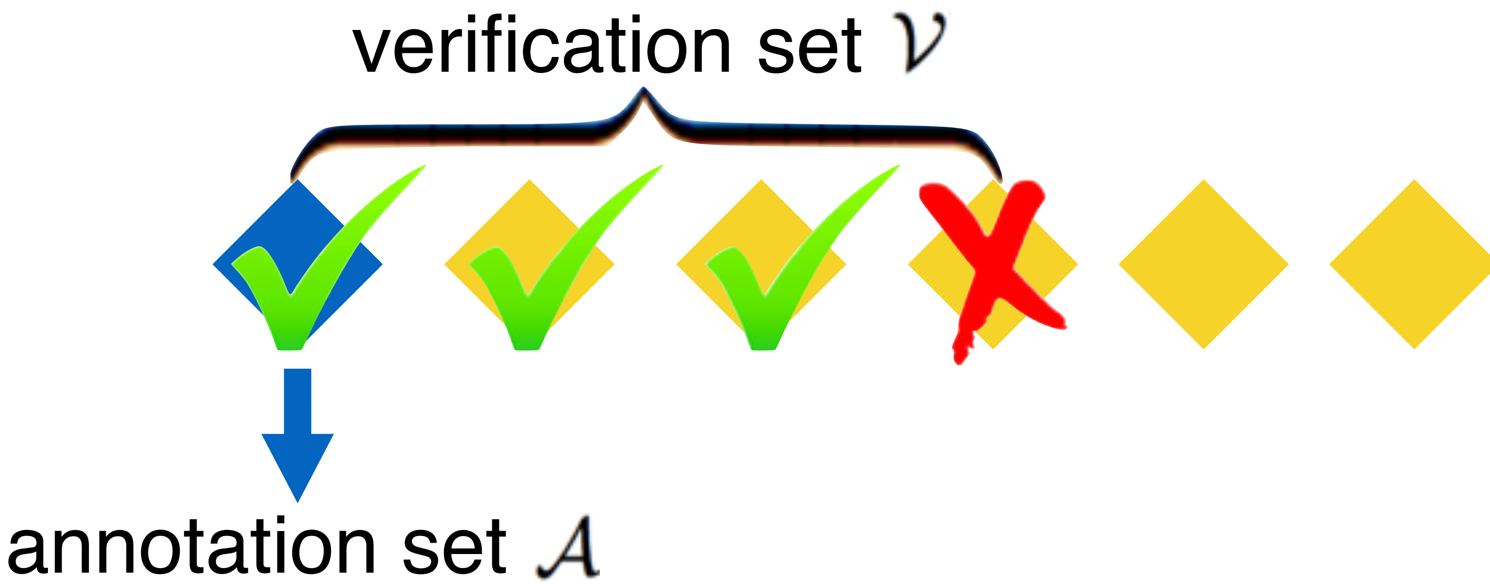
Utility Function



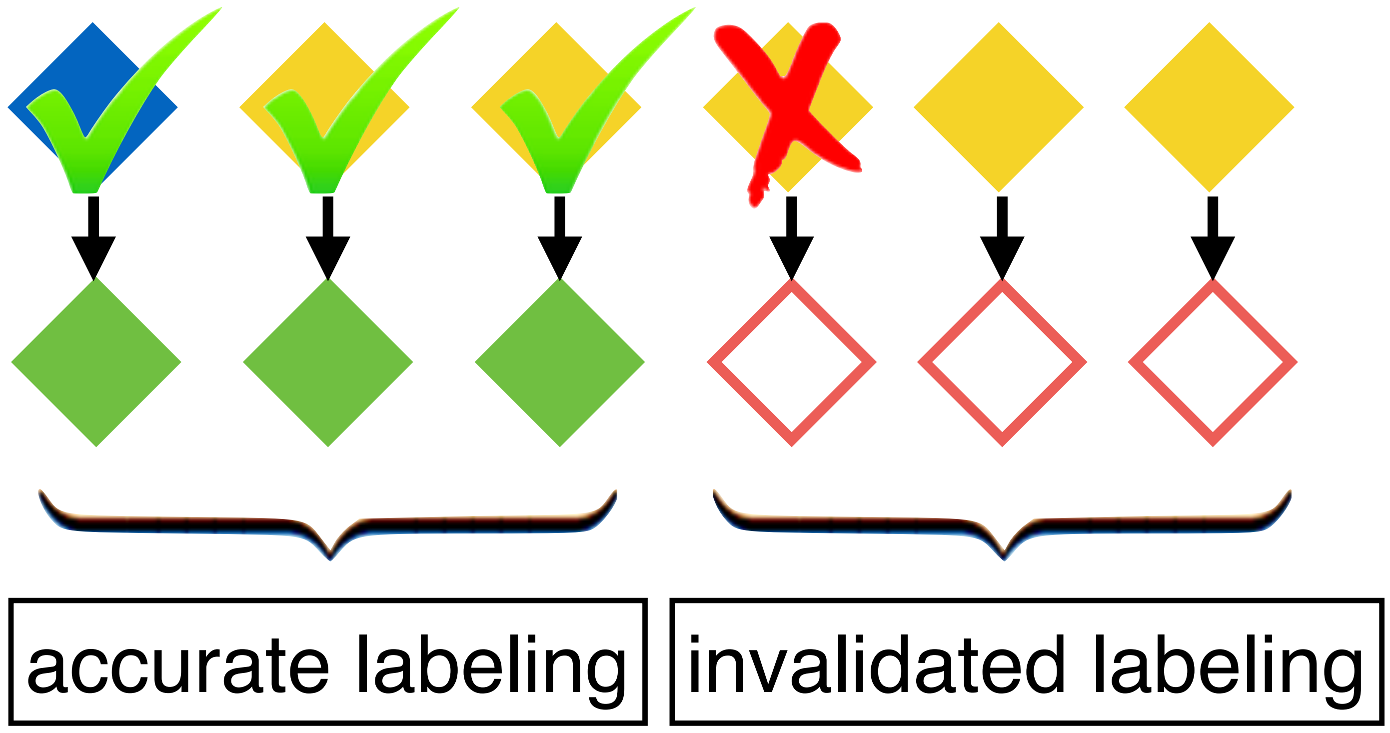
Utility Function



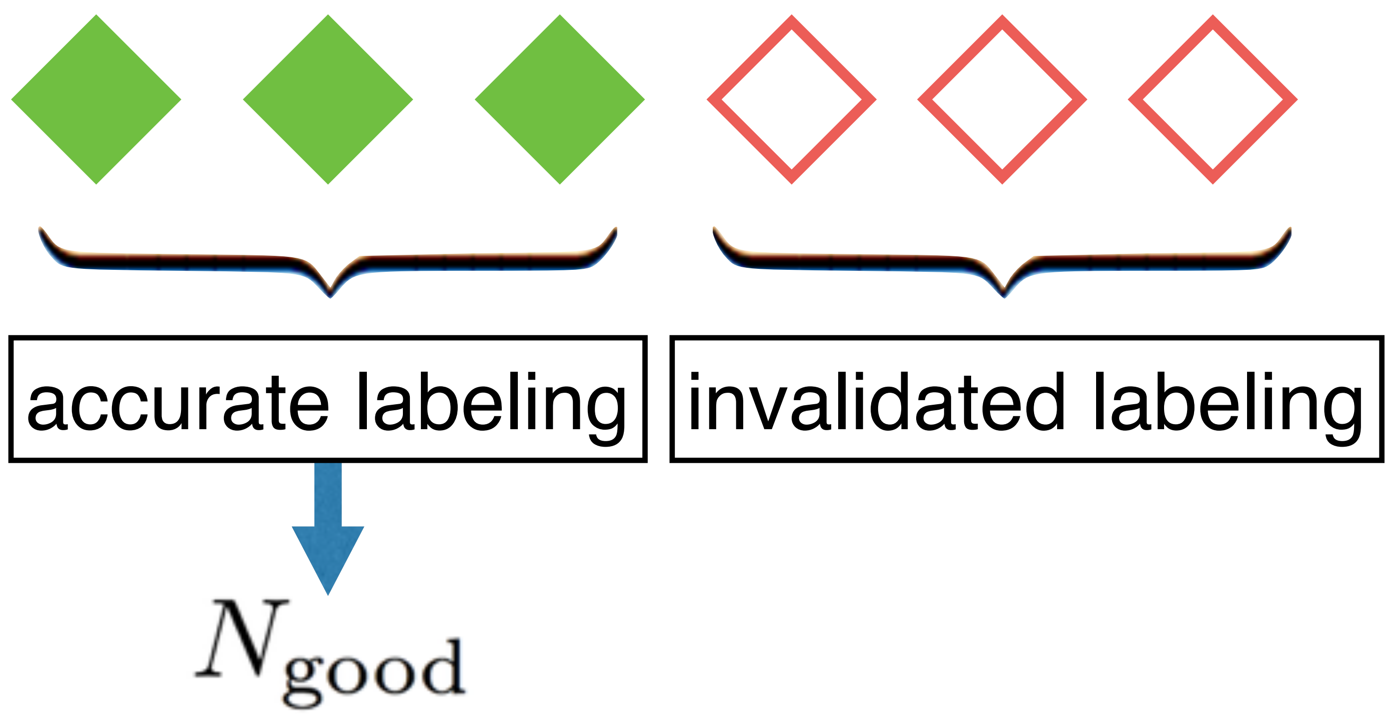
Utility Function



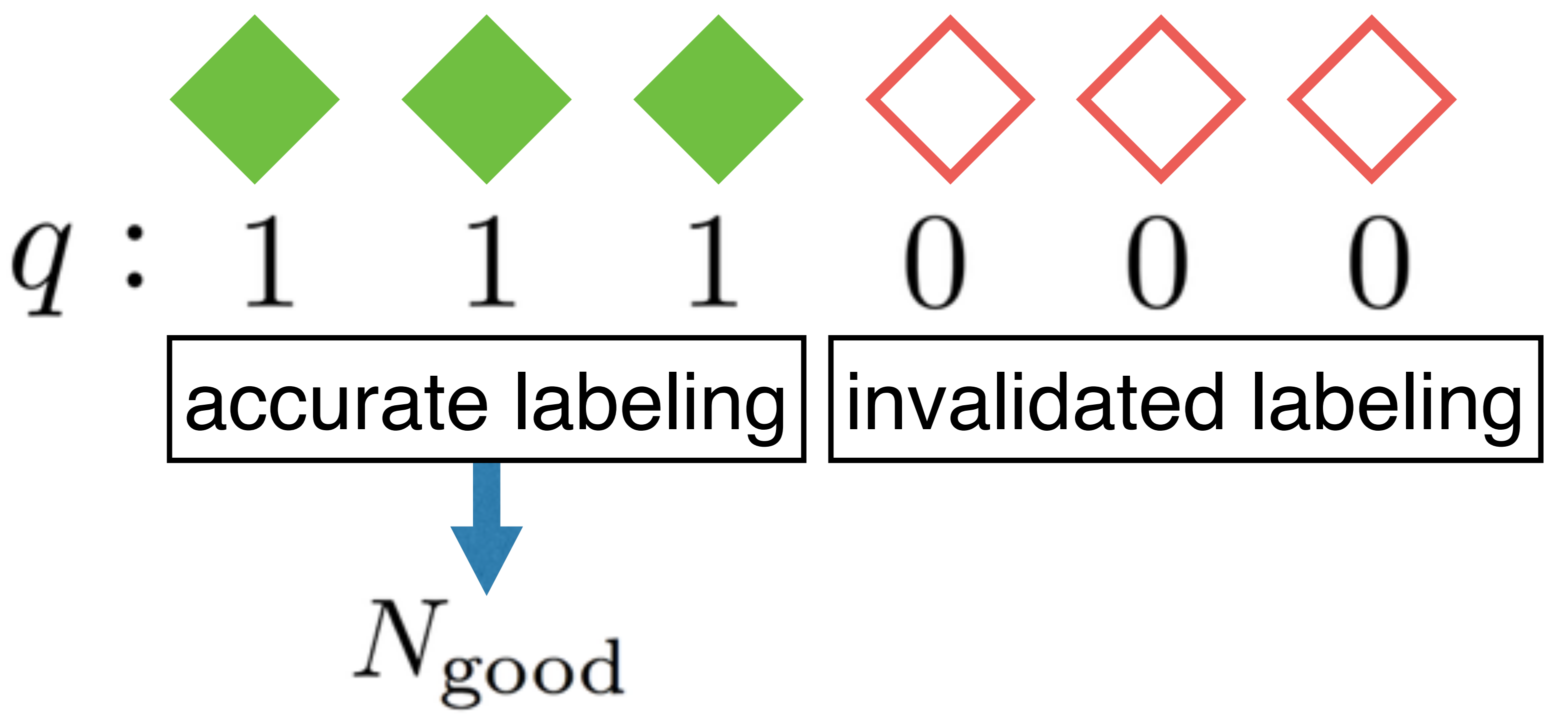
Utility Function



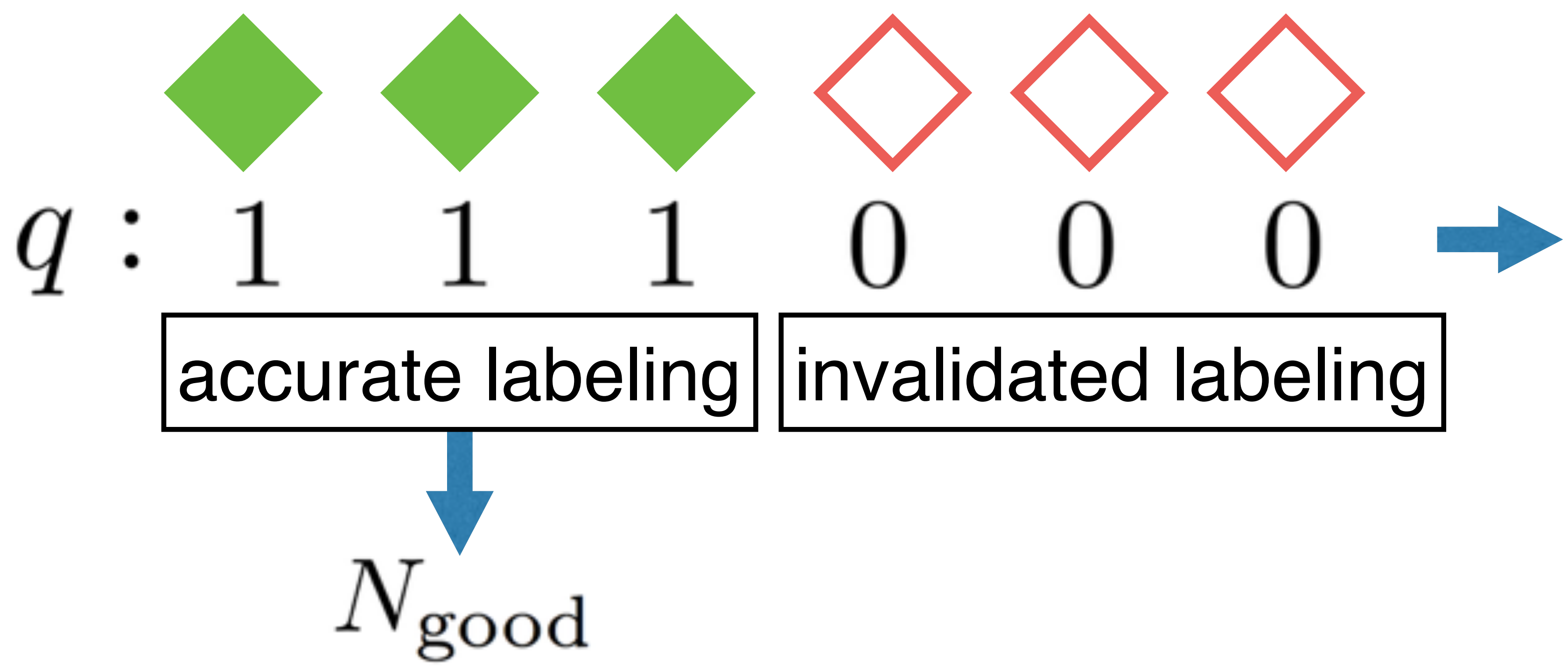
Utility Function



Utility Function

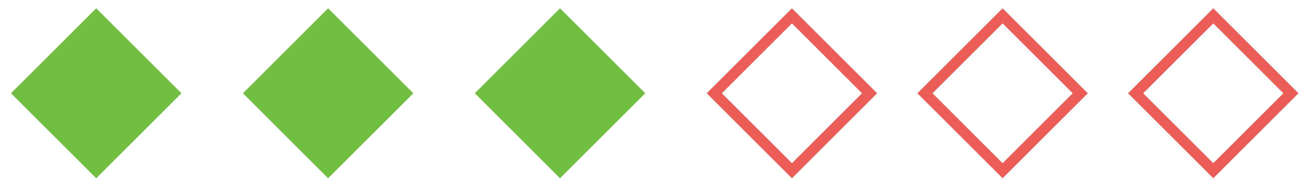


Utility Function



A function of
annotation set $\mathcal{A}$
and verification set $\mathcal{V}$

Utility Function



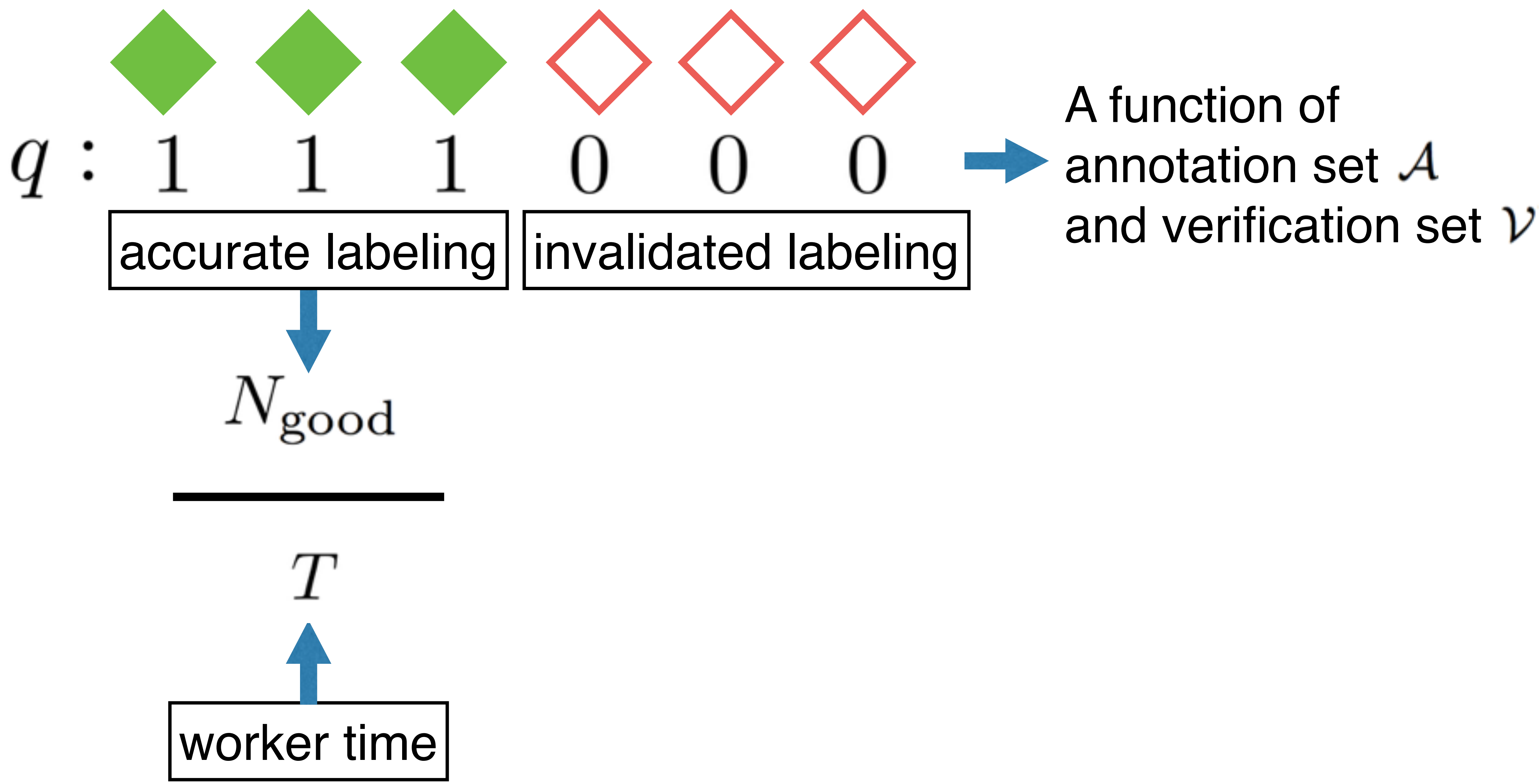
$q: 1 \quad 1 \quad 1 \quad 0 \quad 0 \quad 0$

accurate labeling | invalidated labeling

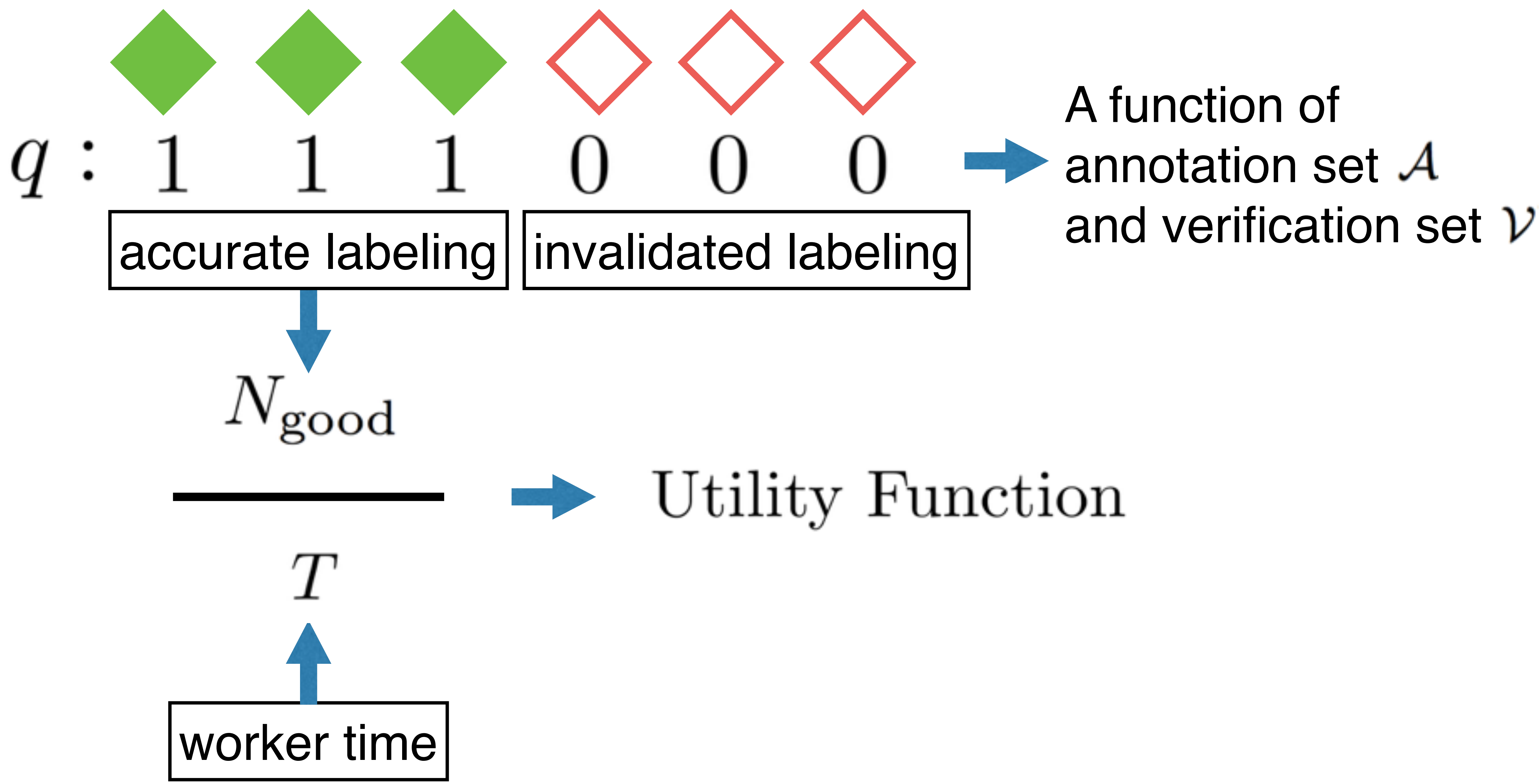
A function of annotation set $\mathcal{A}$ and verification set $\mathcal{V}$

$$N_{\text{good}} = \sum q(\mathcal{A}, \mathcal{V})$$

Utility Function



Utility Function

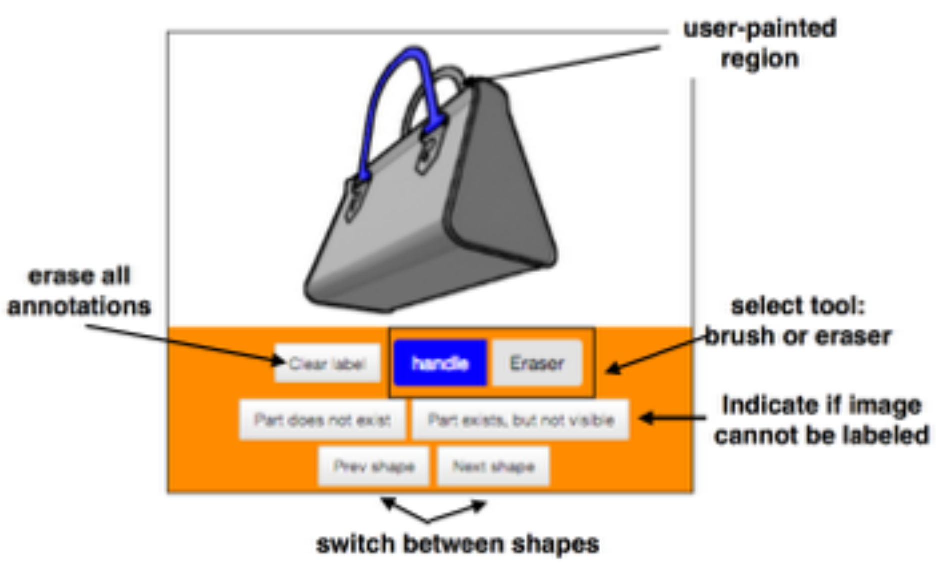


Utility Function

$$T = t_{\text{annotation}}(\mathcal{A}) + t_{\text{verification}}(\mathcal{V}, \{q\})$$

Utility Function

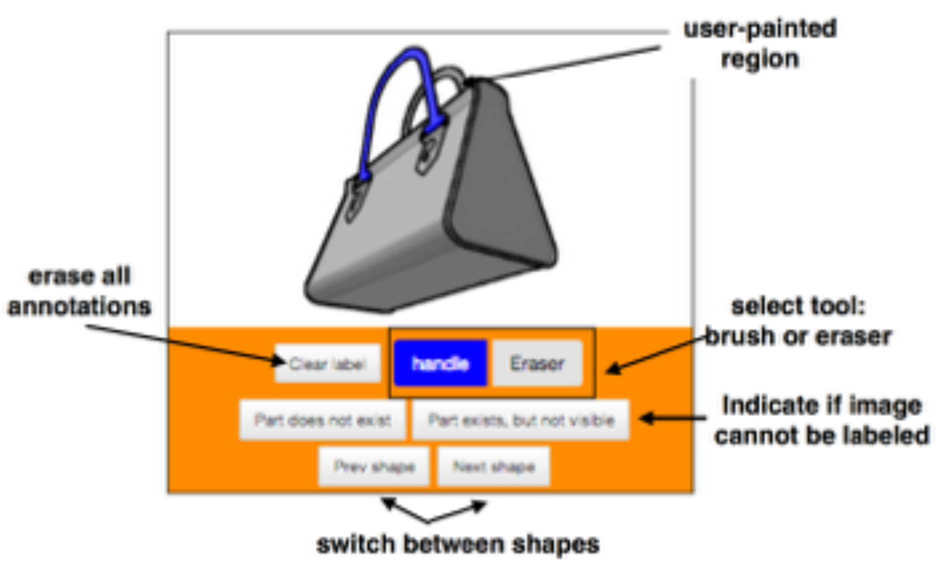
$$T = t_{\text{annotation}}(\mathcal{A}) + t_{\text{verification}}(\mathcal{V}, \{q\})$$



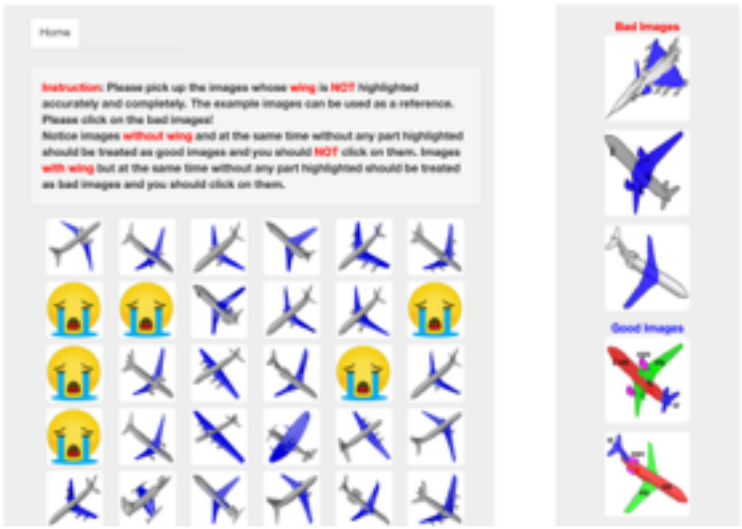
$$\tau_{\text{annotation}} \approx 30s/\text{label}$$

Utility Function

$$T = t_{\text{annotation}}(\mathcal{A}) + t_{\text{verification}}(\mathcal{V}, \{q\})$$



$$\tau_{\text{annotation}} \approx 30s/\text{label}$$



$$\tau_{\text{identification}} \approx 0.3s/\text{label}$$

$$\tau_{\text{click}} \approx 1.1s/\text{label}$$

Utility Function

$$\frac{N_{\text{good}}}{T}$$

Utility Function

$$\frac{N_{\text{good}}}{T}$$

$\{q\}$ is not known before selecting $\mathcal{A}$ and $\mathcal{V}$

Utility Function

$$\frac{N_{\text{good}}}{T}$$

$\{q\}$ is not known before selecting $\mathcal{A}$ and $\mathcal{V}$

Estimate probability and use expectation

Utility Function

$$\frac{N_{\text{good}}}{T}$$

$\{q\}$ is not known before selecting $\mathcal{A}$ and $\mathcal{V}$

Estimate probability and use expectation

$$\boxed{\frac{\mathbb{E}[N_{\text{good}}]}{\mathbb{E}[T]}}$$

Utility Function

$$\frac{N_{\text{good}}}{T}$$

$\{q\}$ is not known before selecting $\mathcal{A}$ and $\mathcal{V}$

Estimate probability and use expectation

$$\boxed{\frac{\mathbb{E}[N_{\text{good}}]}{\mathbb{E}[T]}}$$

Algorithm behavior

Utility Function

$$\operatorname*{argmin}_{\mathcal{A},\mathcal{V}} \frac{\mathbb{E}[N_{\mathbf{good}}]}{\mathbb{E}[T]}$$

Utility Function

$$\underset{\mathcal{A},\mathcal{V}}{\operatorname{argmin}} \frac{\mathbb{E}[N_{\text{good}}]}{\mathbb{E}[T]}$$

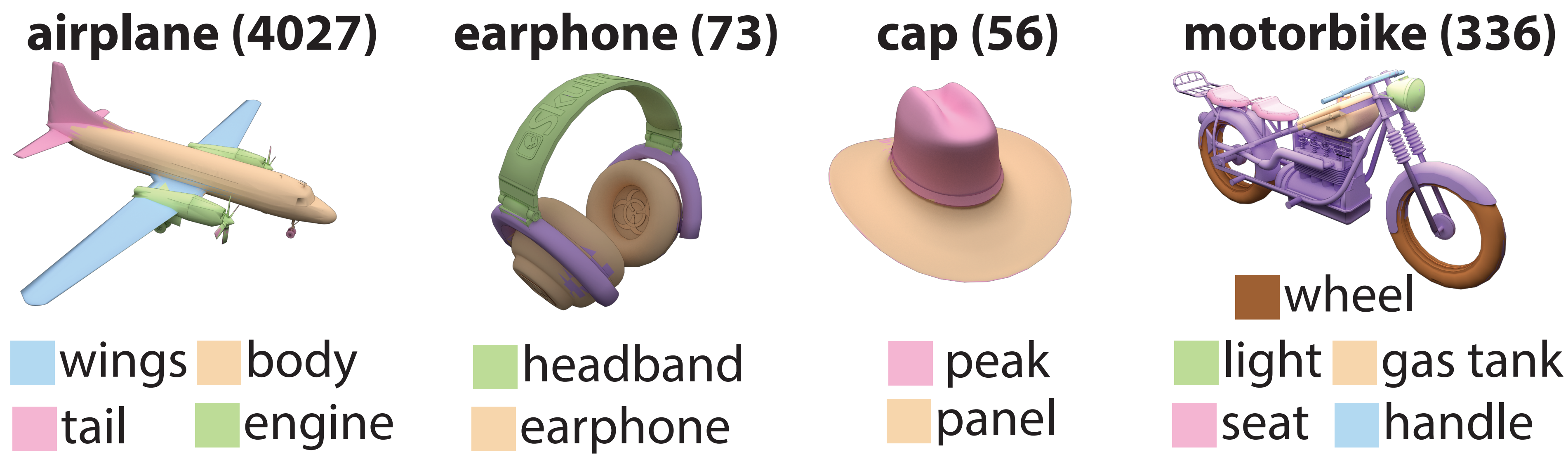
Select Annotation Set: ***Beam Search***

Select Verification Set: ***Greedy Search***

Results

Dataset

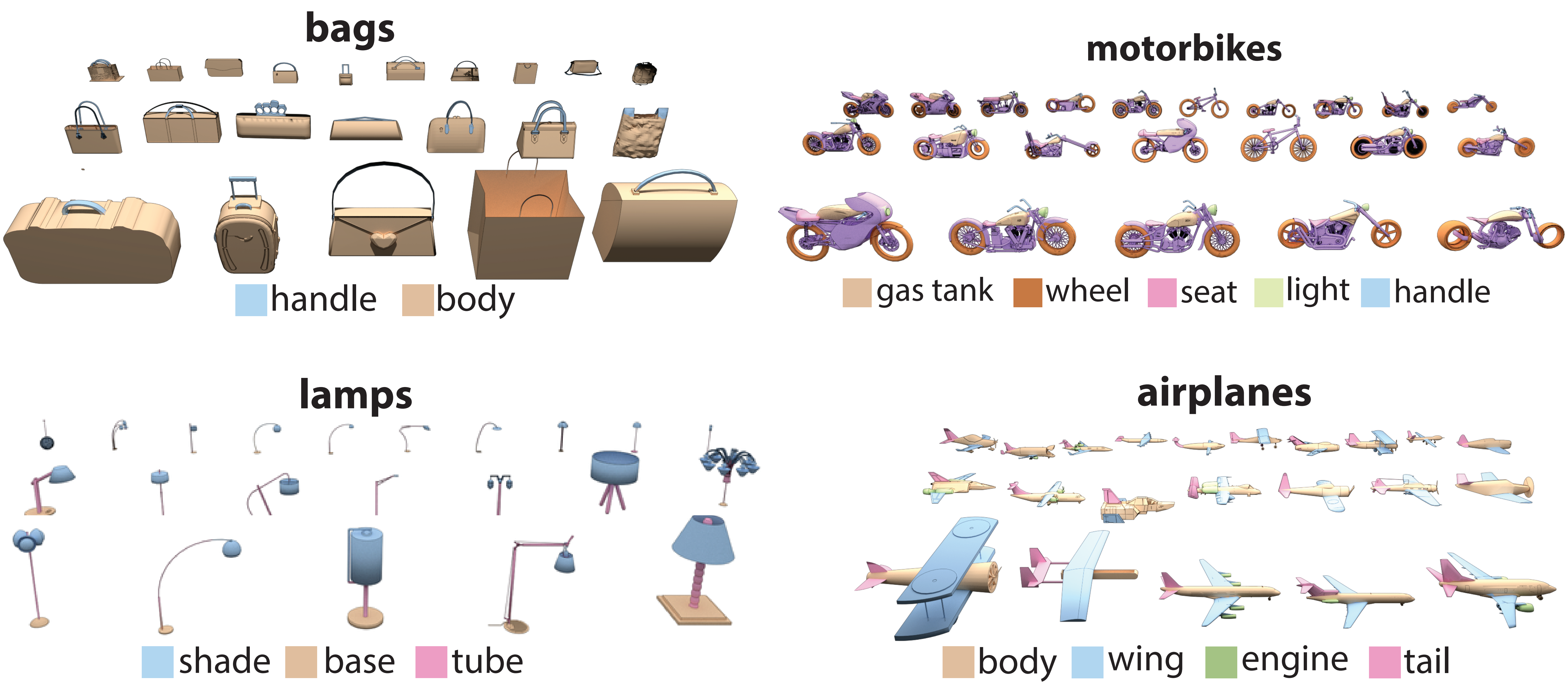
- 30,000 shapes
- 90,000 parts



Results

Dataset

- 30,000 shapes
- 90,000 parts



Time Saving

Comparison on manifold shapes

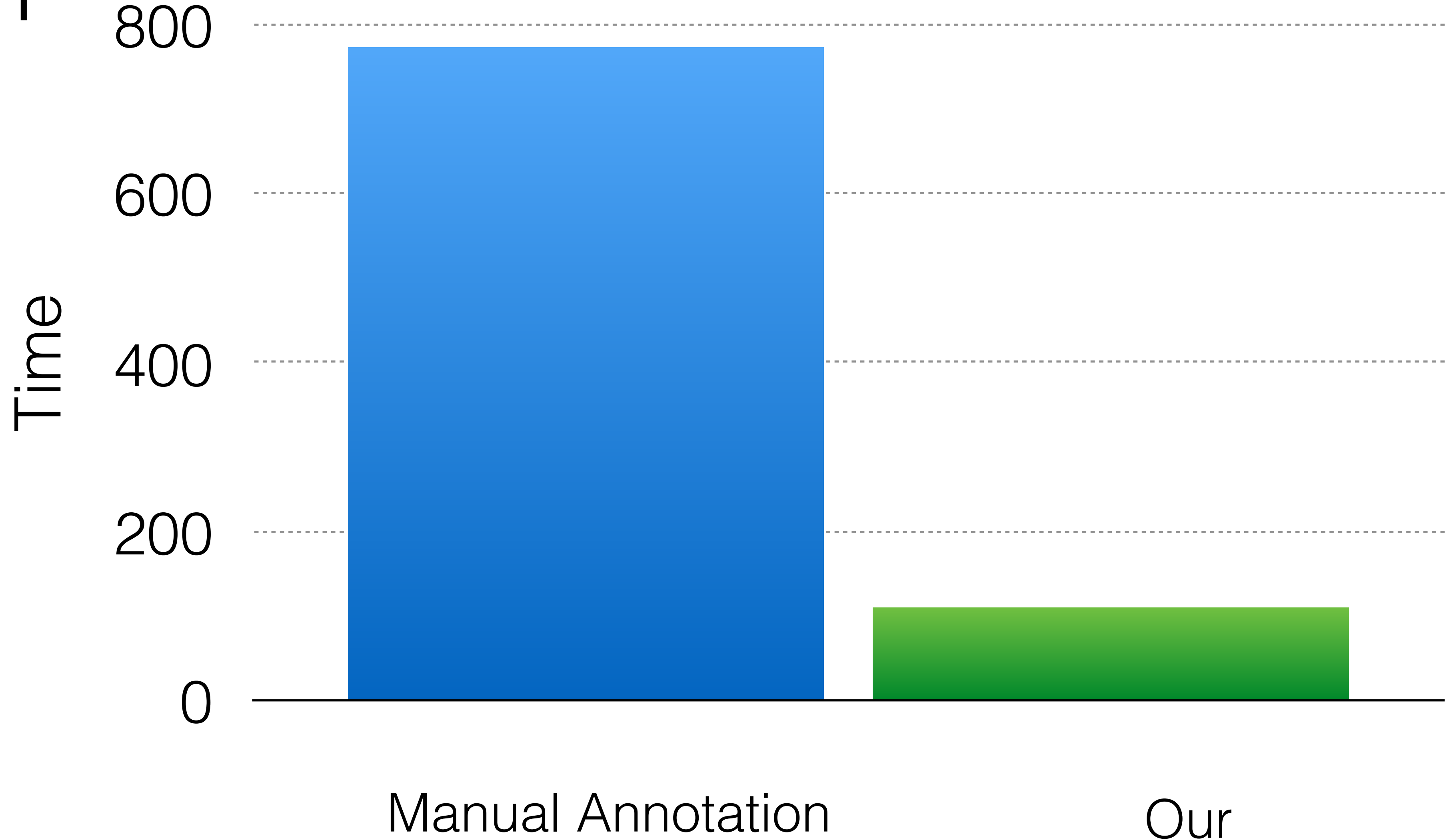
- Chen'09 (manual) : about x12 more expensive
- Wang'12 (active) : about x2 more expensive

Time Saving

Comparison on manifold shapes

- Chen'09 (manual) : about x12 more expensive
- Wang'12 (active) : about x2 more expensive

Comparison on data "in the wild"



Summary

Large collections of 3D models are available (and more are coming!)

- 3D modeling repositories
- 3D printing datasets
- Kinect scans
- Laser scans of cities,
- Online shopping catalogues
- Medical imaging data
- Protein databases
- …

Summary

Large collections of 3D models are available (and more are coming!)

Part-based models are powerful, but need data

Summary

Large collections of 3D models are available (and more are coming!)

Part-based models are powerful, but need data

Getting verified results might save you time

- if you verify automatic labels

Future Work

Beyond geometry

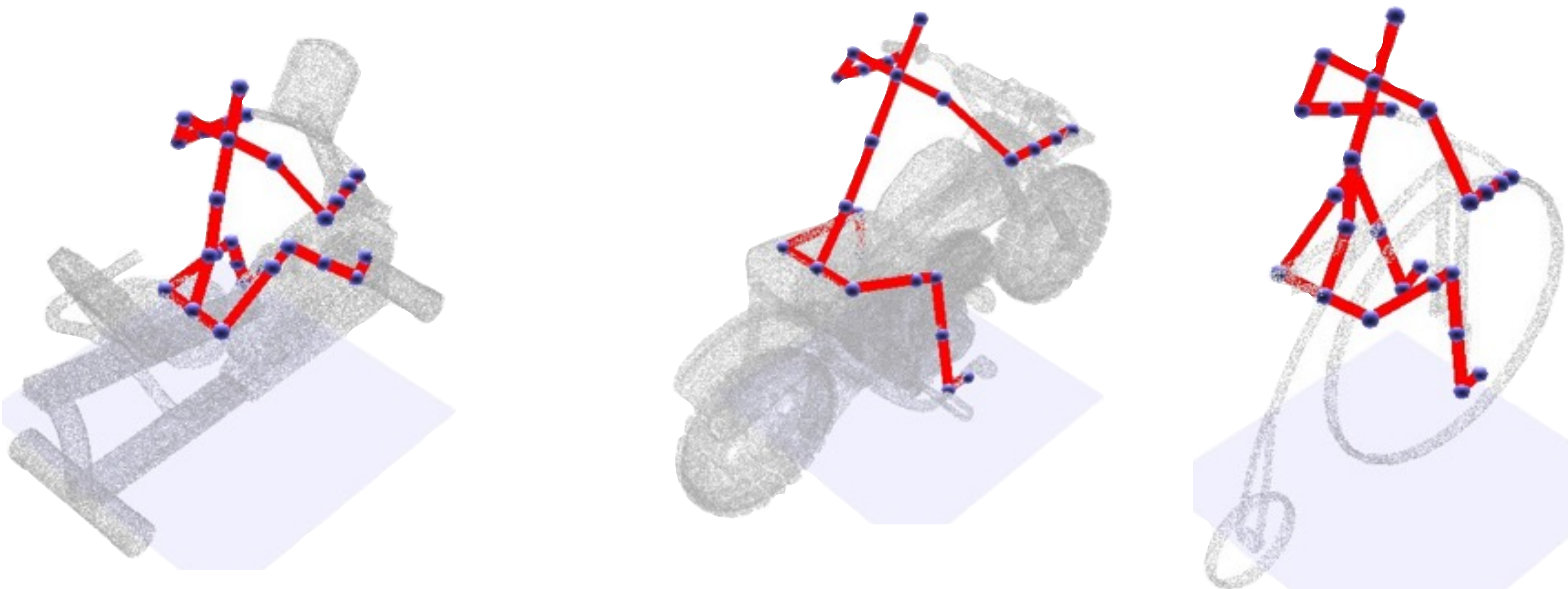
- Understand function
- Consider physical materials that make up objects
- Model mechanics and articulations

Future Work

Beyond geometry

➡ **Understand <u>function</u>**

- Consider physical <u>materials</u> that make up objects
- Model <u>mechanics</u> and articulations



Future Work

Better structural models

- Geometry processing layers in deep neural networks