

Abstracting Collections of Objects and Scenes

Niloy J. Mitra

Smart Geometry Processing Group



- Melinos Averkiou
- James Hennessey
- Moos Heuting
- Bongjin Koo
- Aron Monszpart
- Nicolas Mellado (postdoc)

and *visitors/collaborators*



Microsoft Research

Understanding Scenes



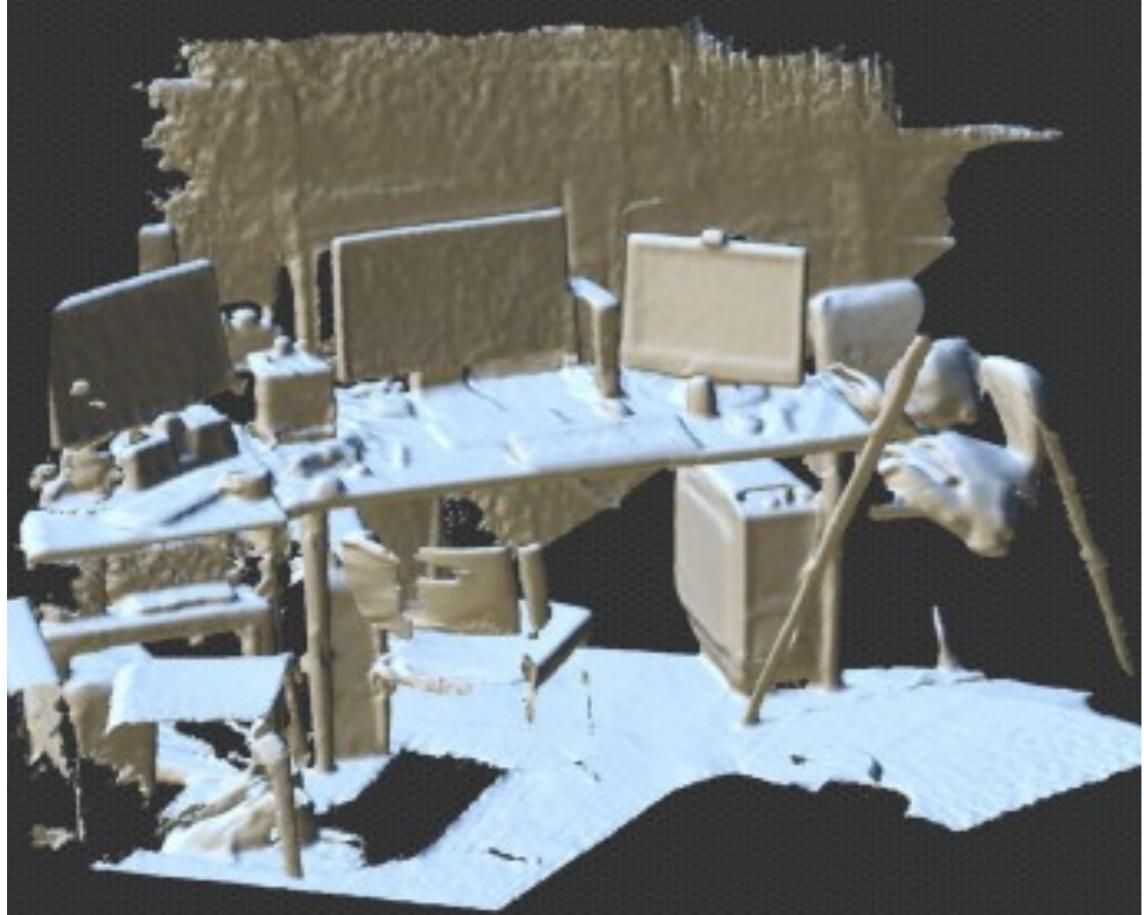
Understanding Scenes



At Different Scales

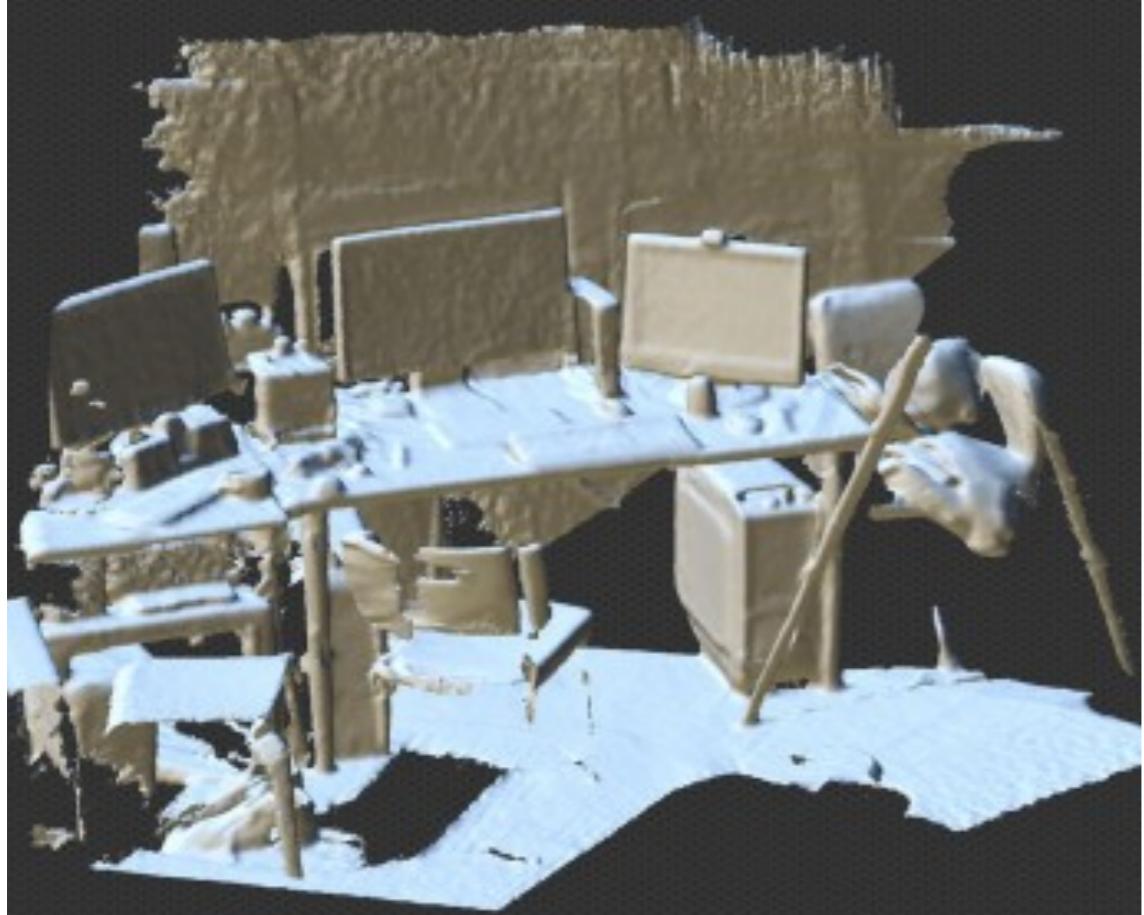


At Different Scales



few million points?

At Different Scales

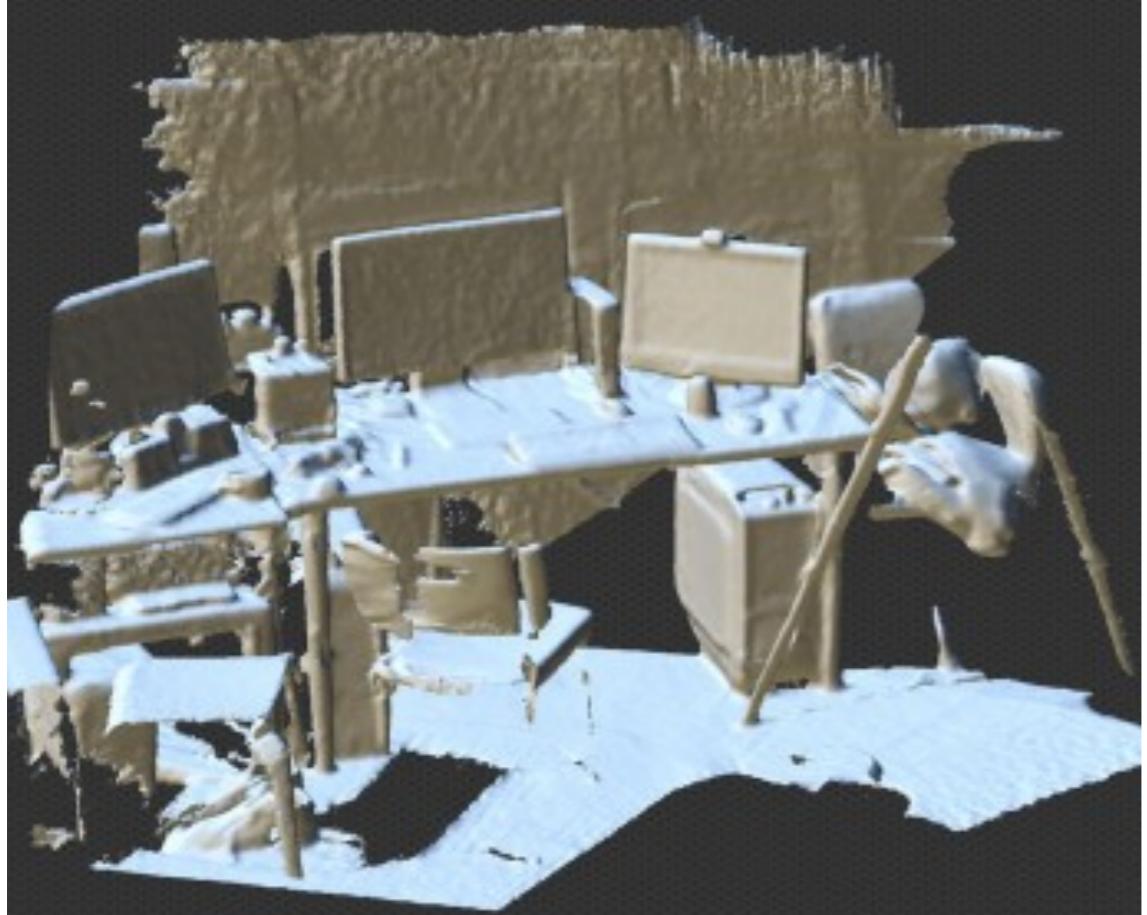


few million points?



100s million points?

At Different Scales



few million points?



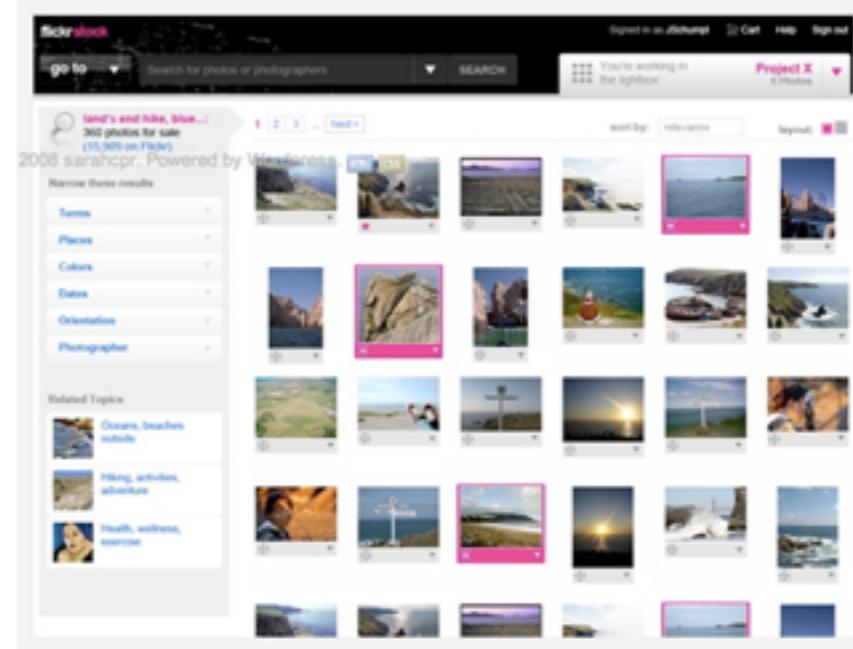
100s million points?

*100-1000s of **related** objects*

Data Sources

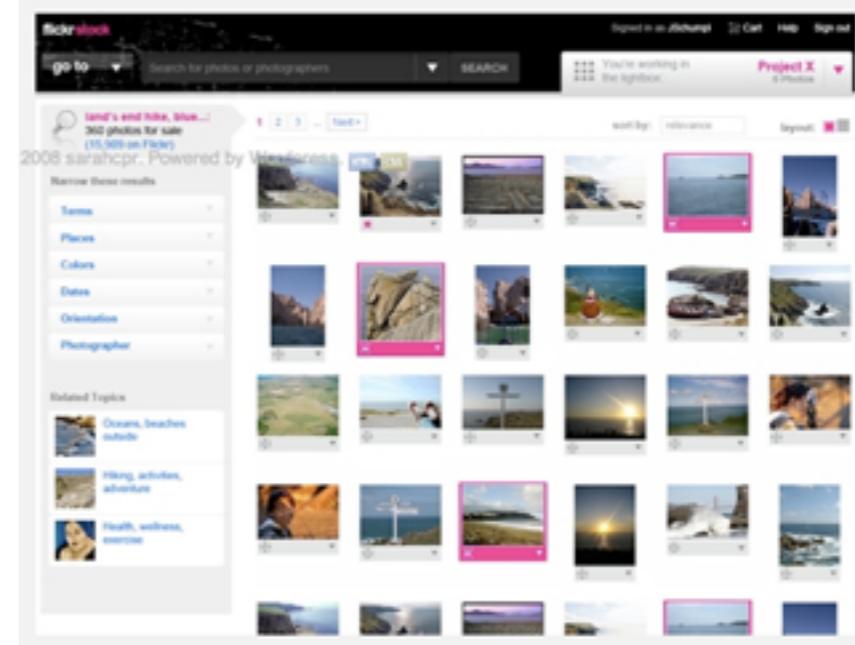
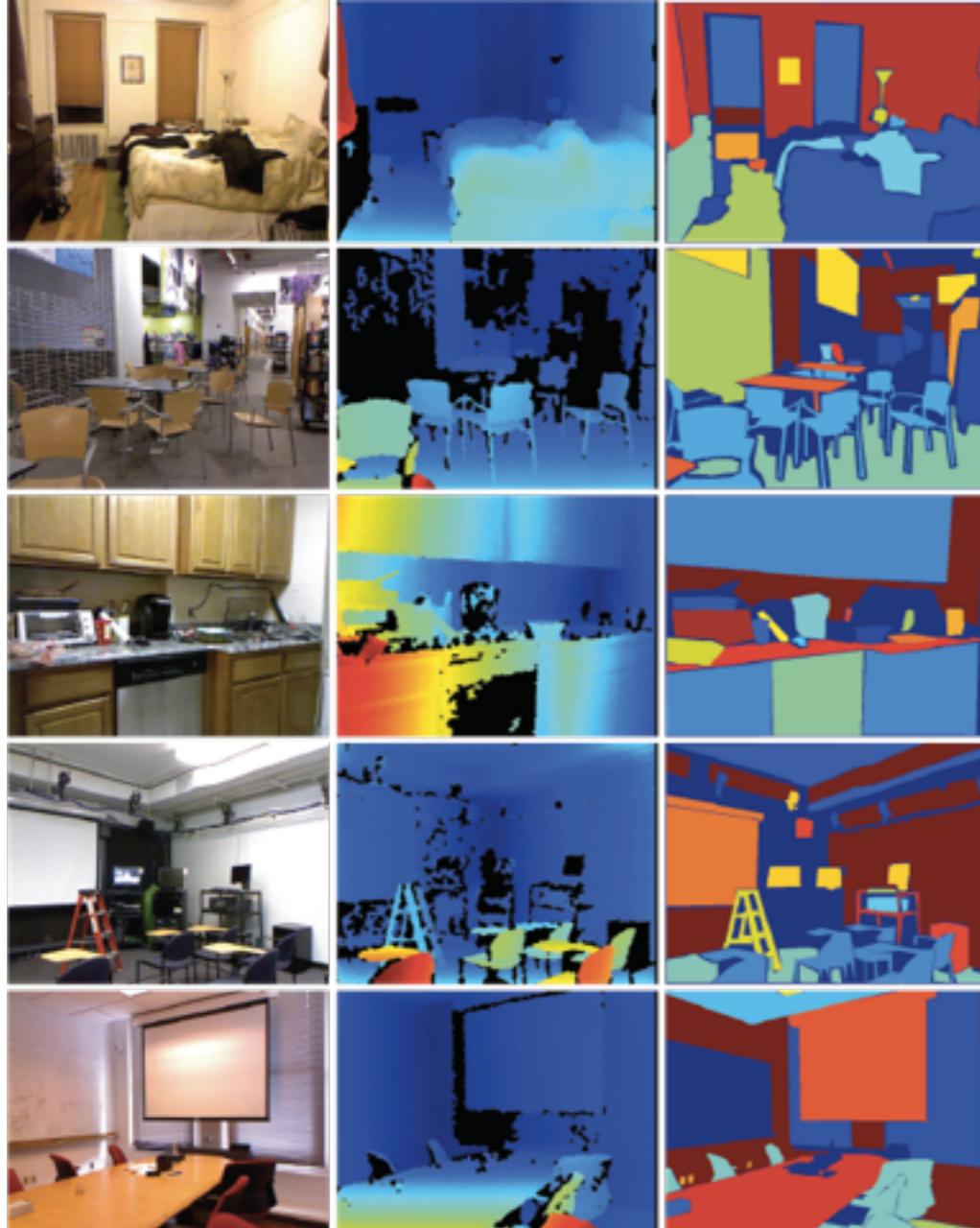


Data Sources



[Flickr]

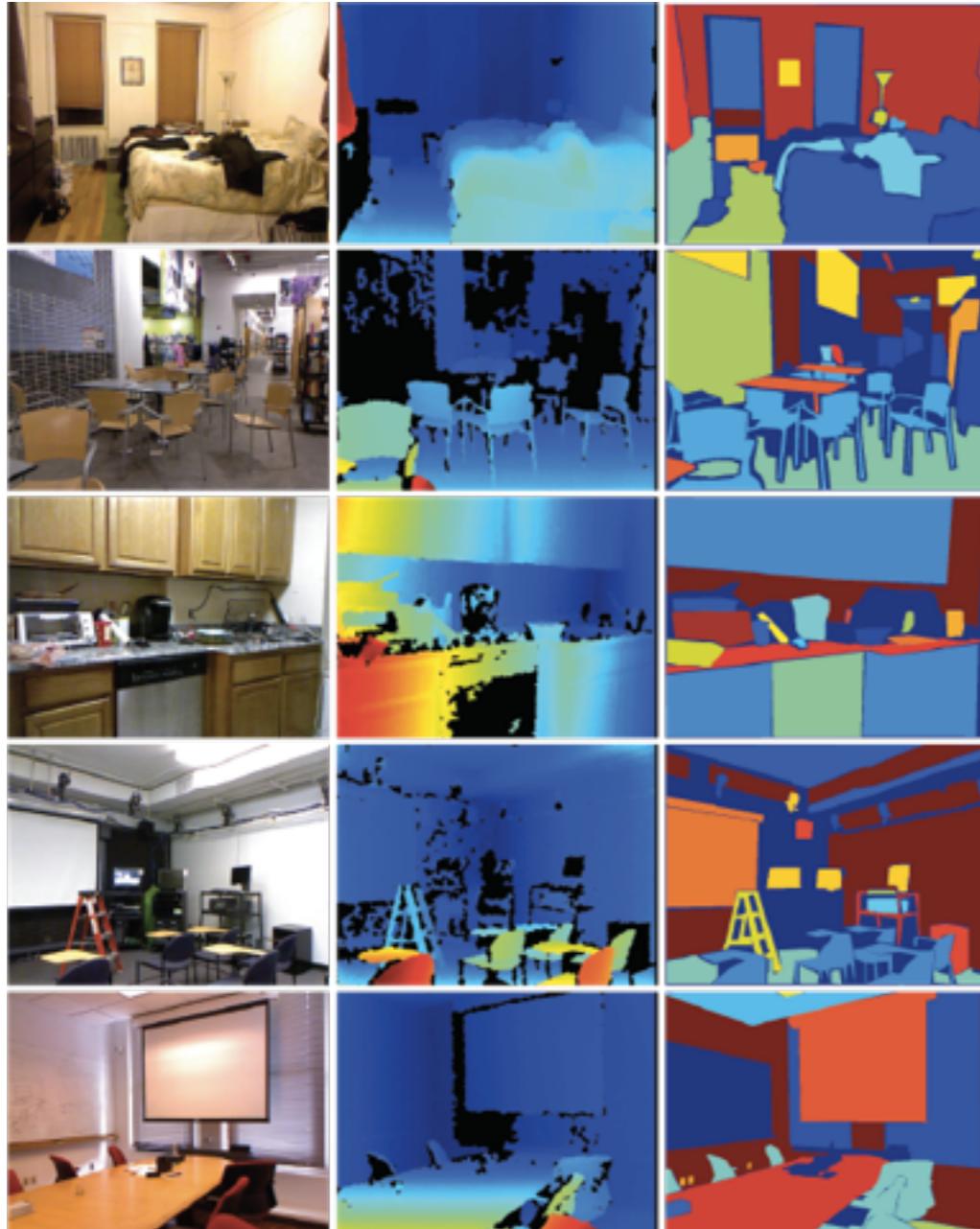
Data Sources



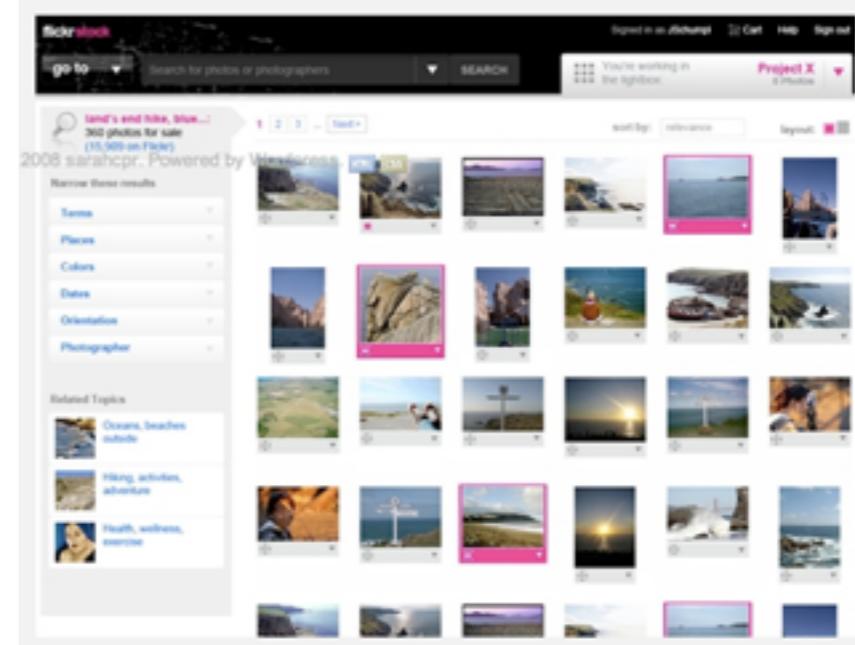
[Flickr]

[NYUx, UCL dataset]

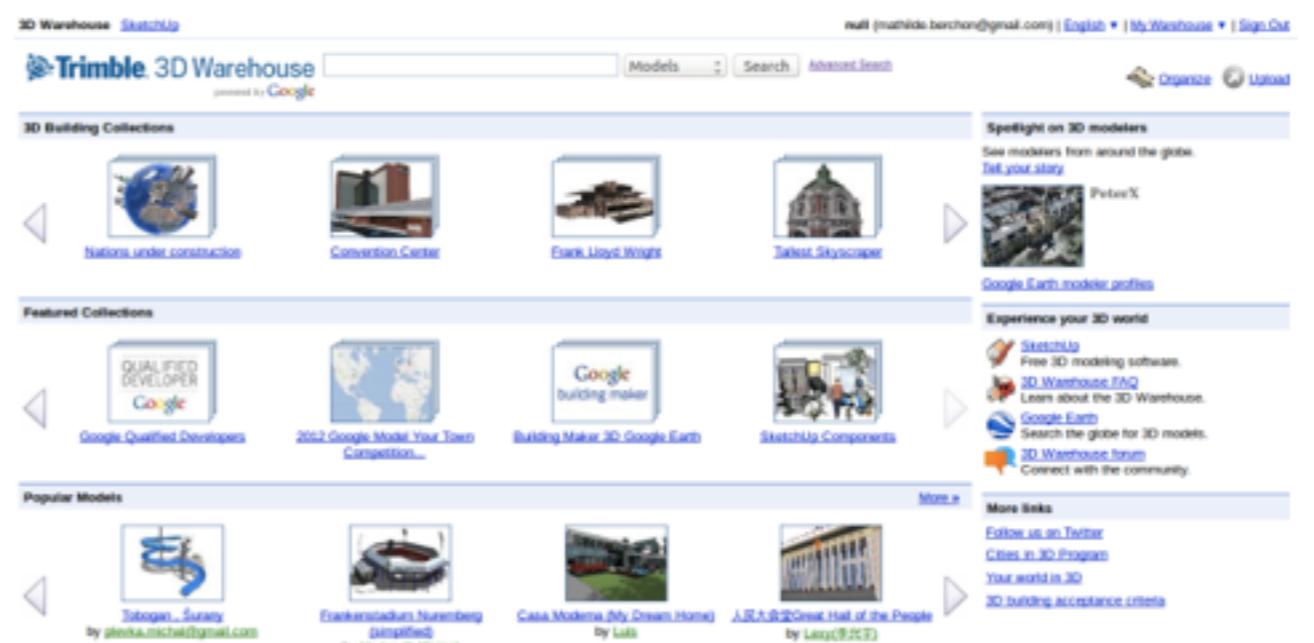
Data Sources



[NYUx, UCL dataset]

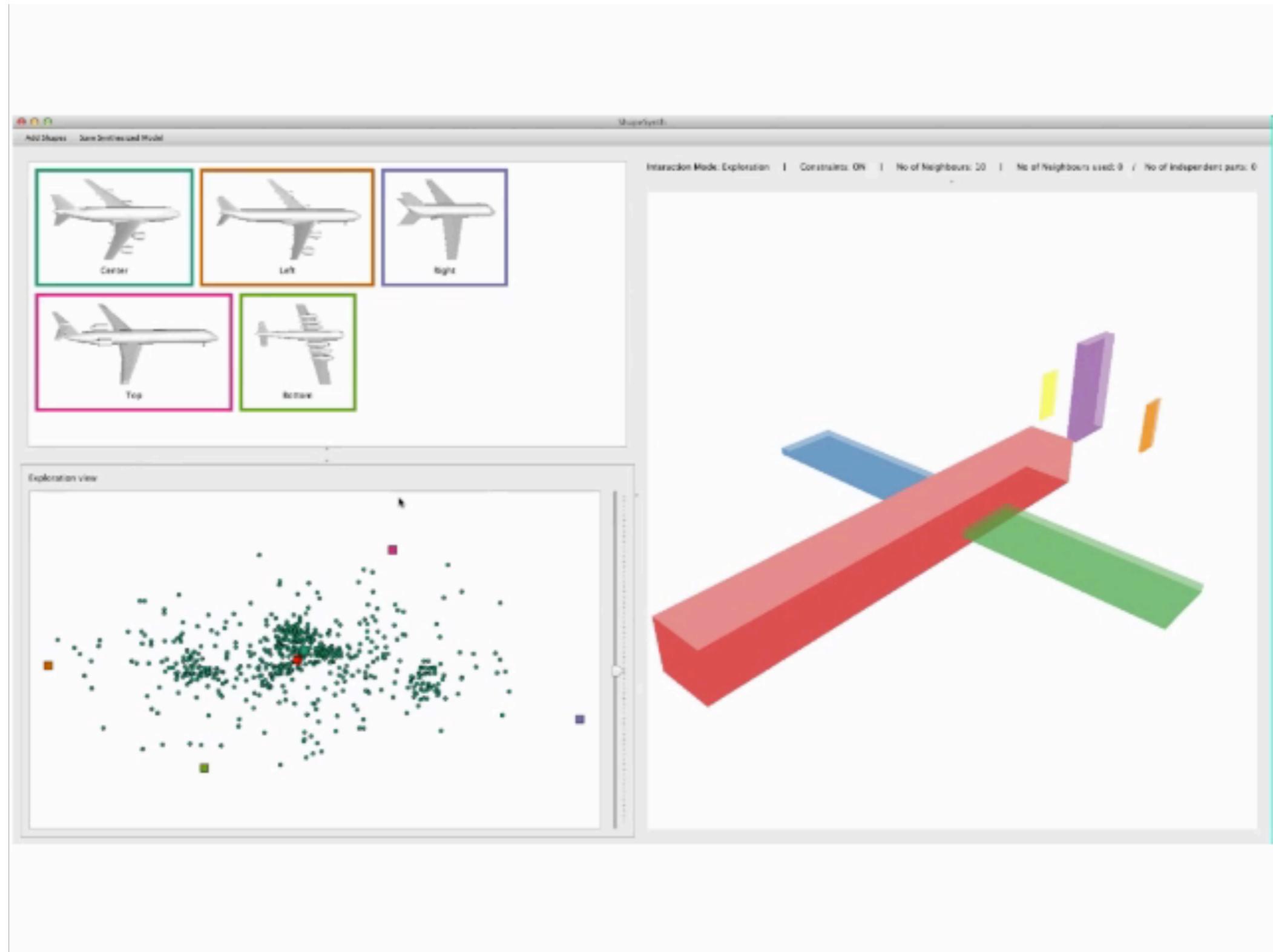


[Flickr]

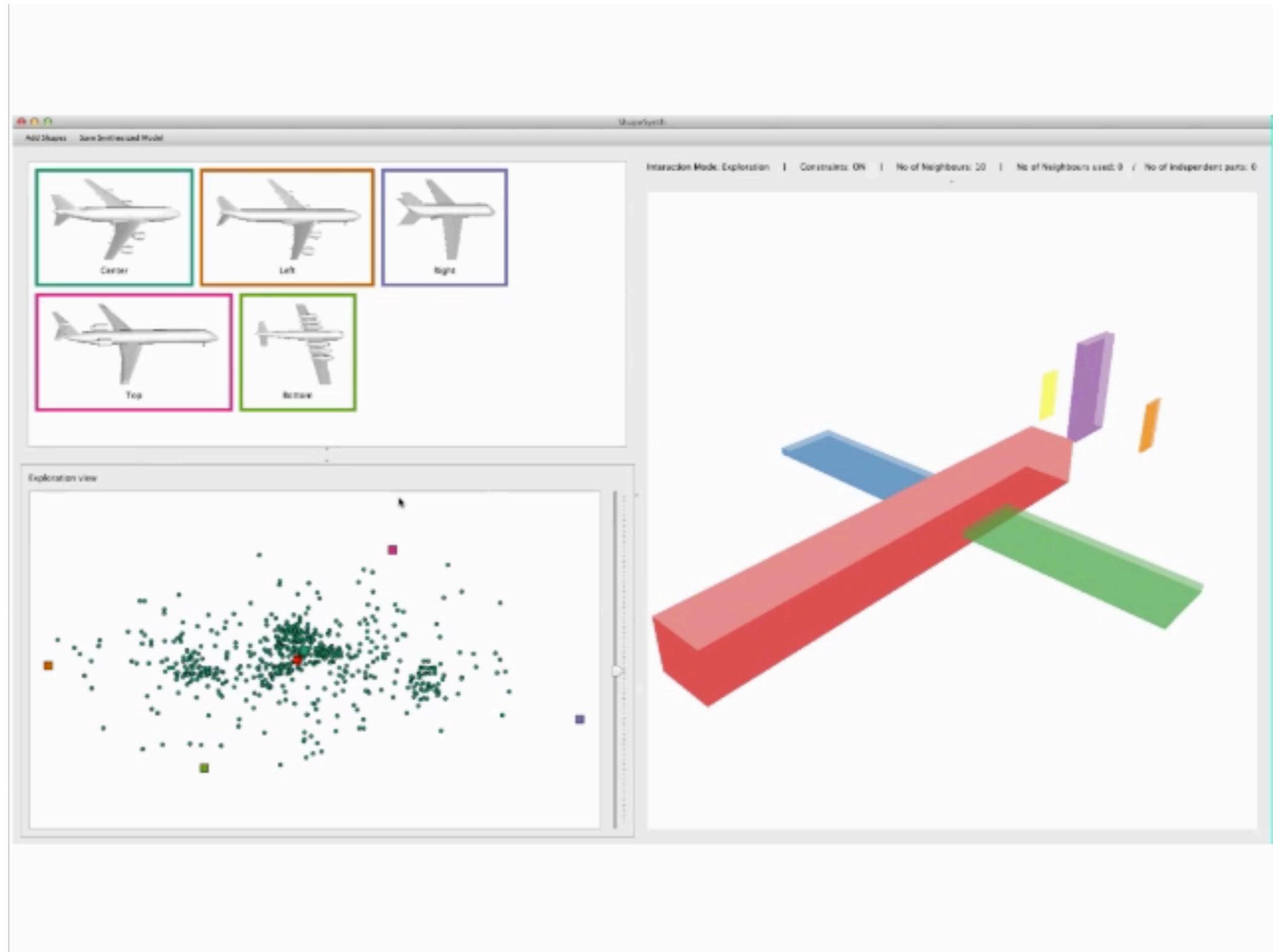


[Trimble warehouse]

Filling in the Void



Filling in the Void



‘Seeing’ the Unseen



[Siggraph Asia 2014]

'Seeing' the Unseen



[Siggraph Asia 2014]

'Seeing' the Unseen



[Siggraph Asia 2014]

'Seeing' the Unseen



[Siggraph Asia 2014]

'Lifting' a Single Image



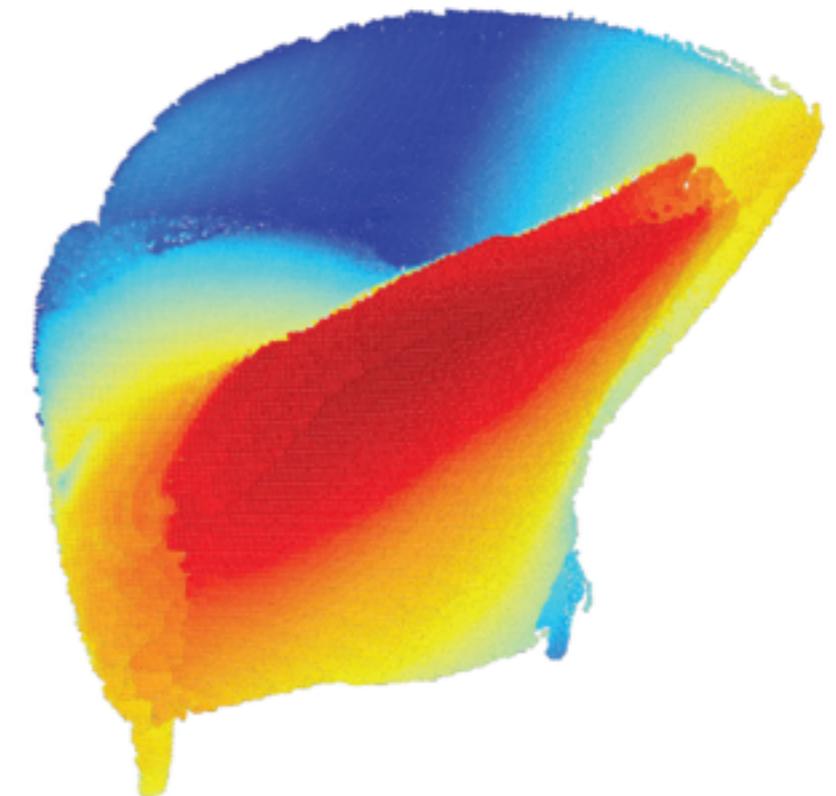
[Siggraph 2014]

'Lifting' a Single Image



[Siggraph 2014]

'Lifting' a Single Image

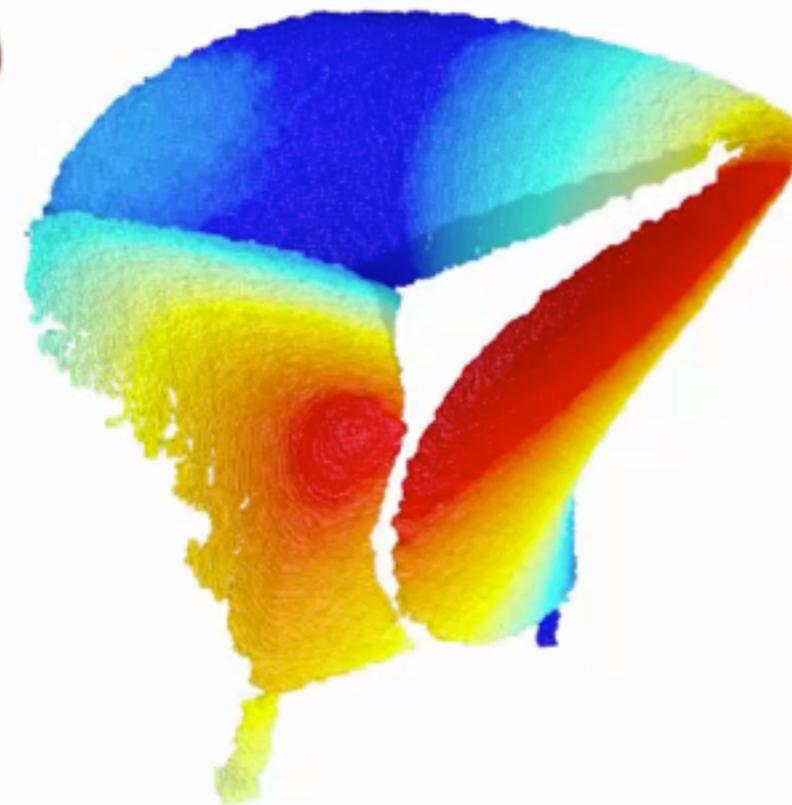


[Siggraph 2014]

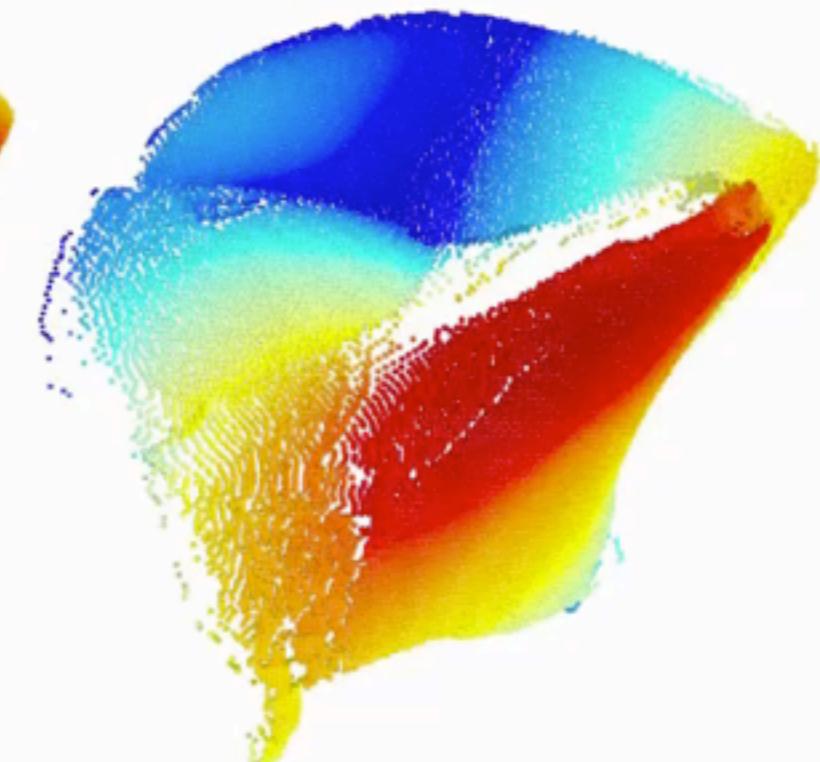
Hallucinated Depth



Input Image



Kinect Scan

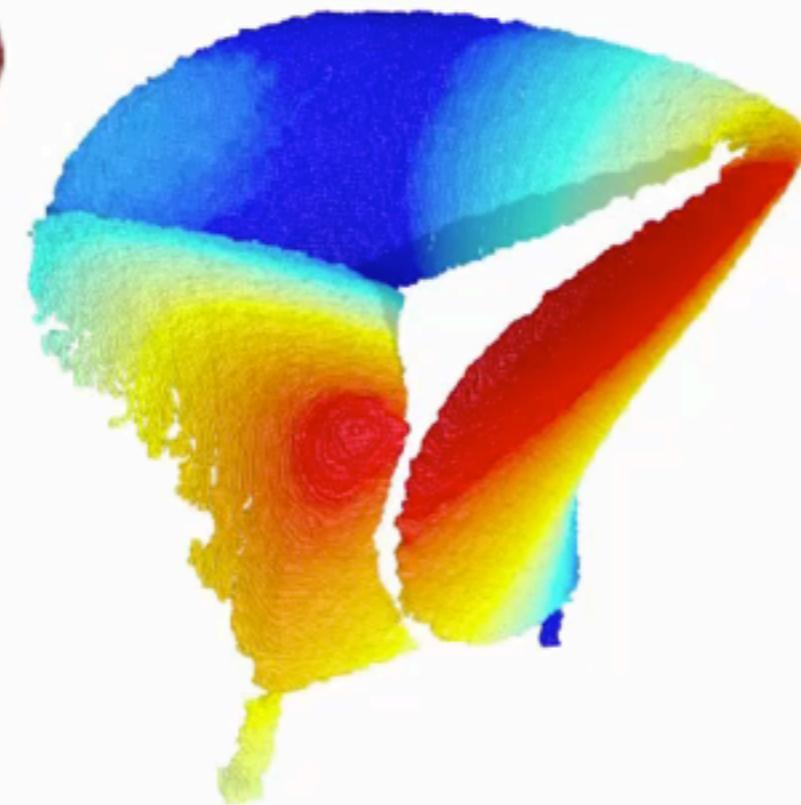


Depth Recovery

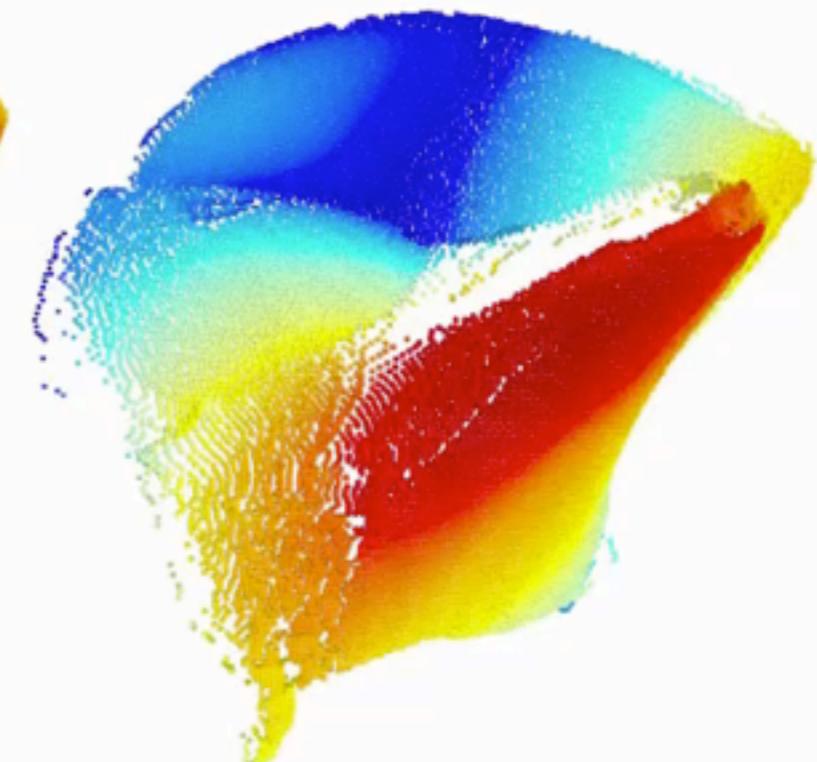
Hallucinated Depth



Input Image

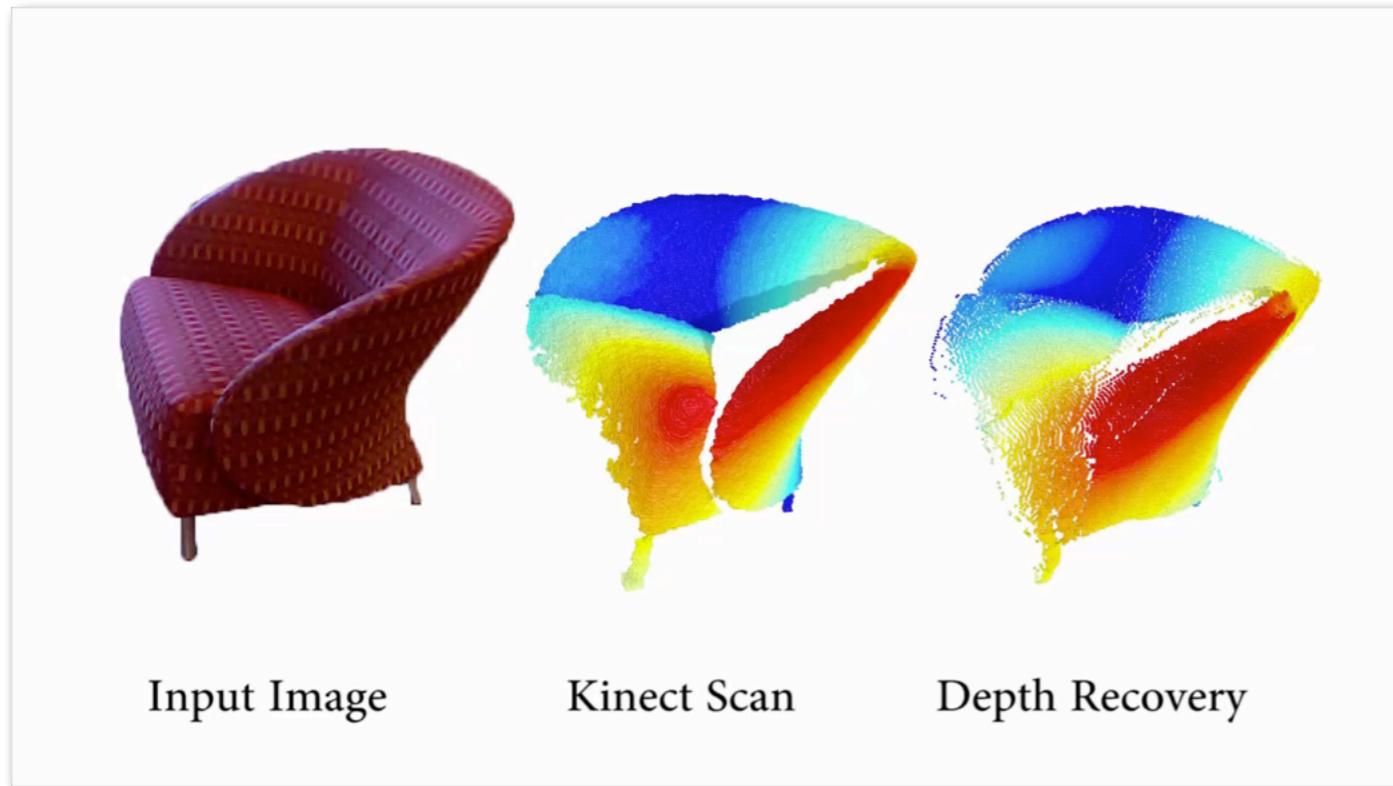


Kinect Scan



Depth Recovery

Hallucinated Depth



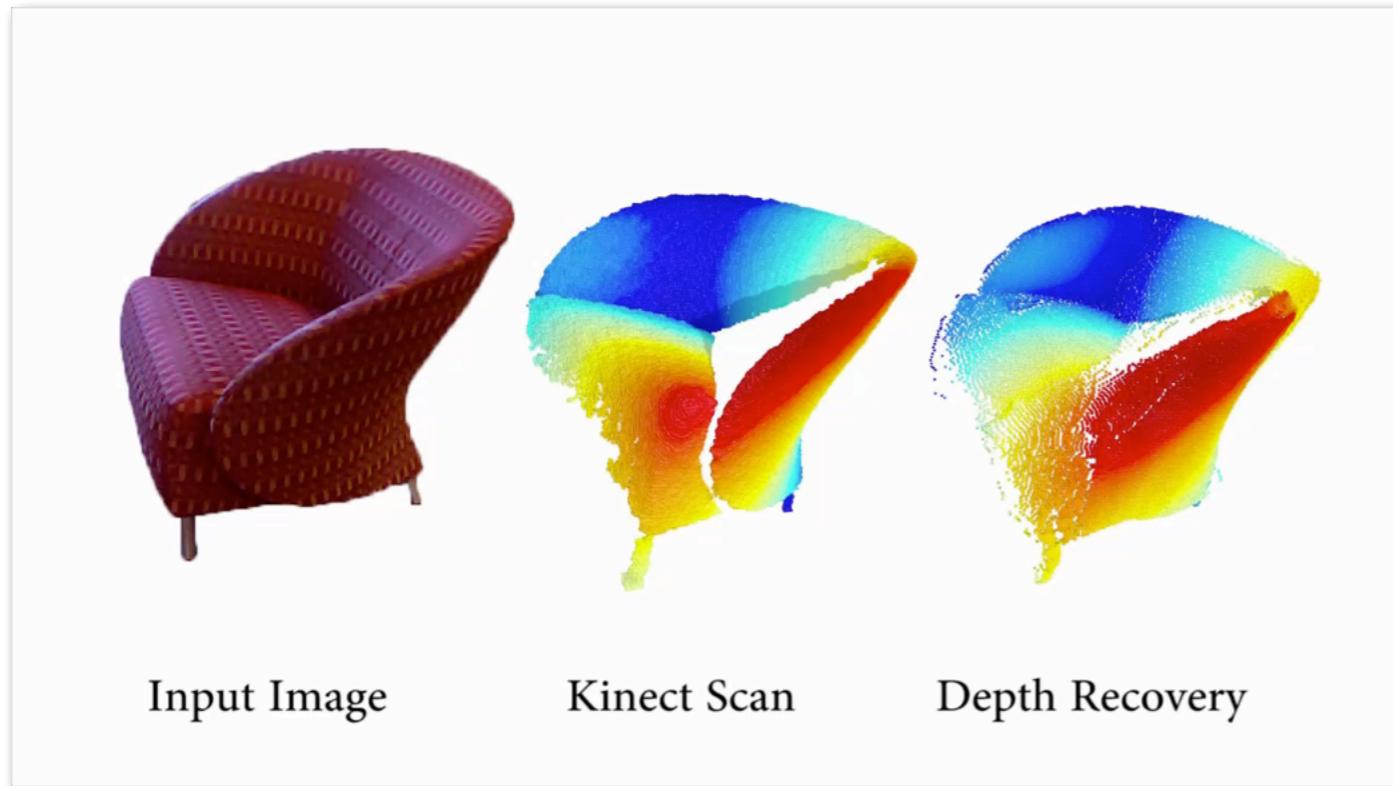
Novel View Synthesis



Relighting



Hallucinated Depth



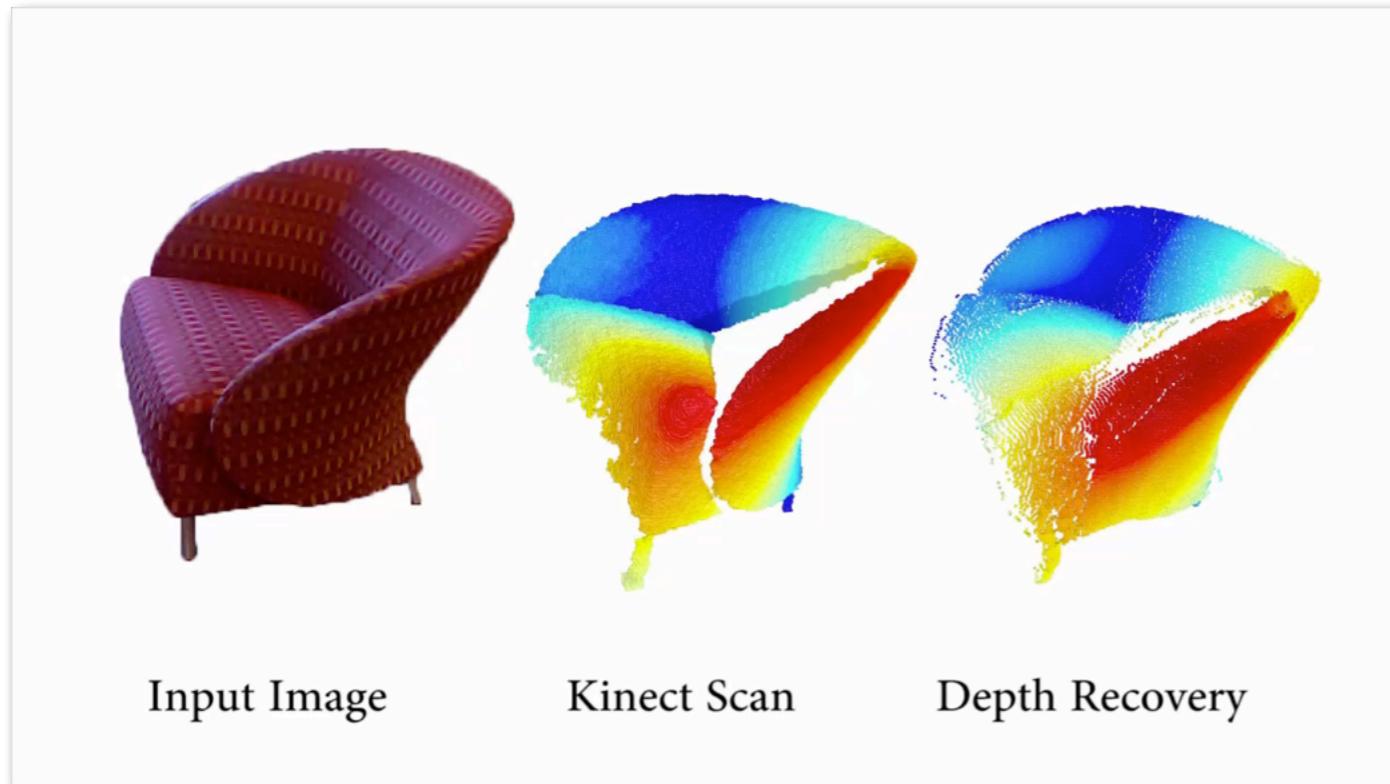
Novel View Synthesis



Relighting



Hallucinated Depth



Novel View Synthesis



Relighting



How to Encode?

parts + relations

Data → Knowledge?



• • •



Data → Knowledge?



• • •



geometric differences can hide functional correlation

Data → Knowledge?



• • •



geometric differences can hide functional correlation

raw measurement → (elements+relations) + variations

Data → Knowledge?



• • •



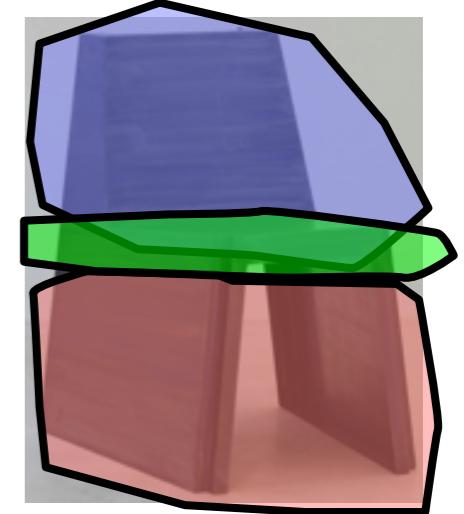
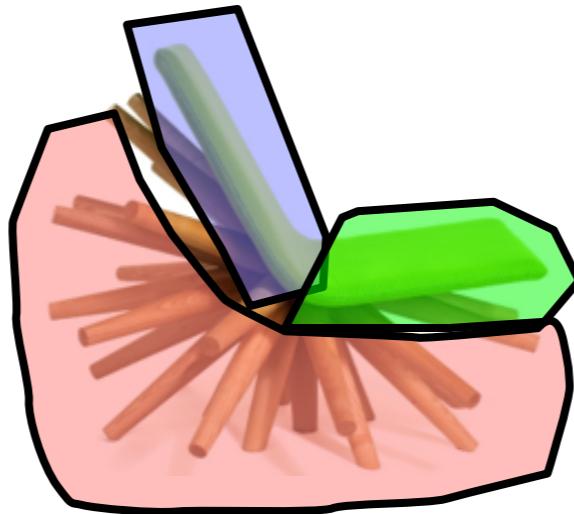
geometric differences can hide functional correlation

raw measurement → (elements+relations) + variations

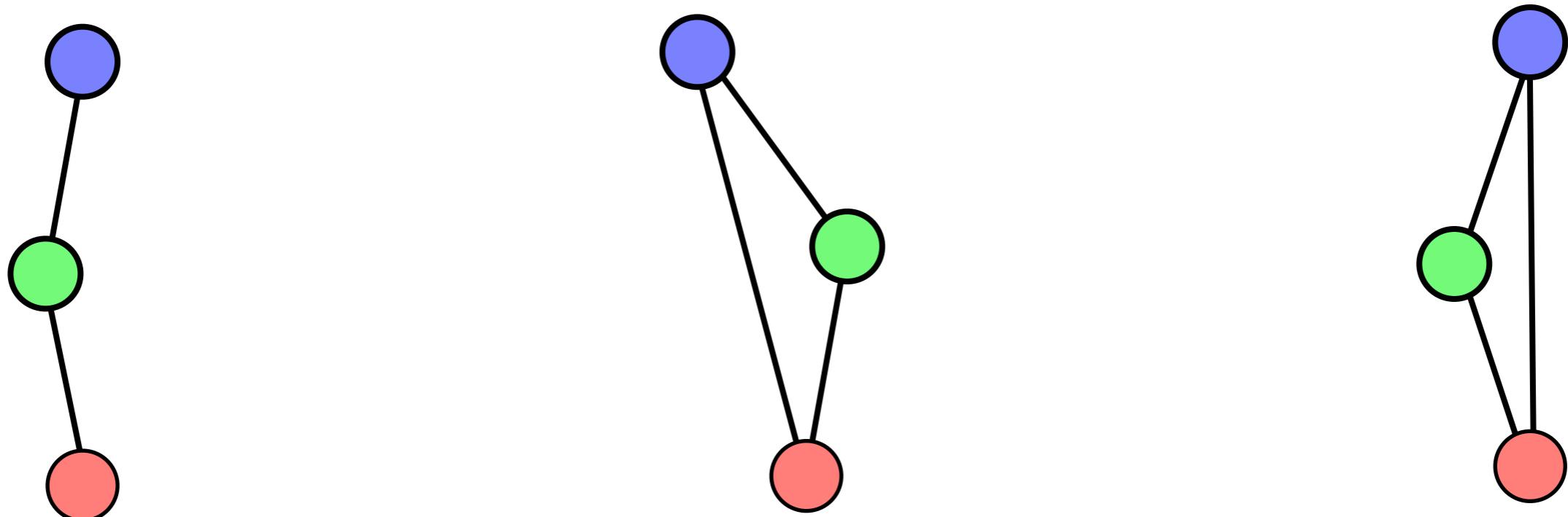
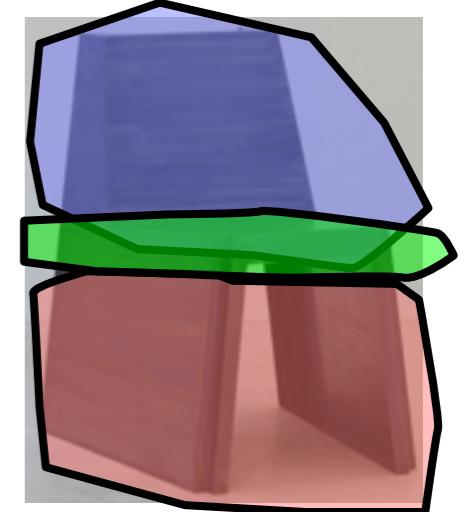
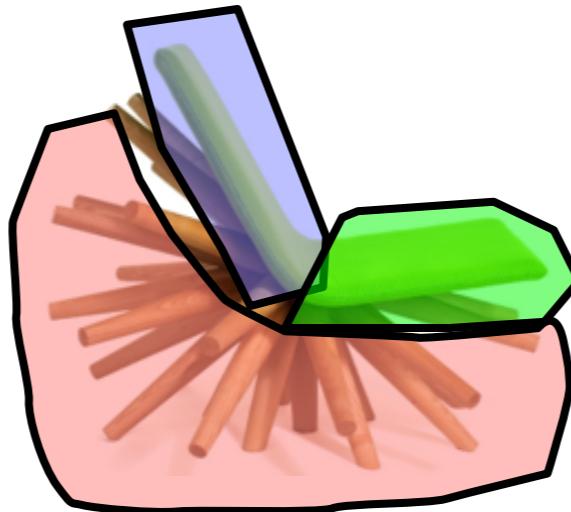
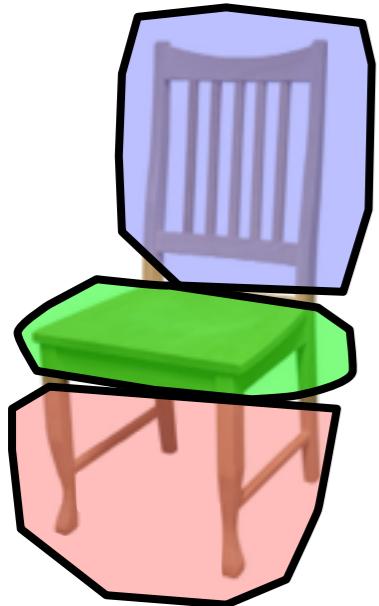
Instead, Focus on Relations



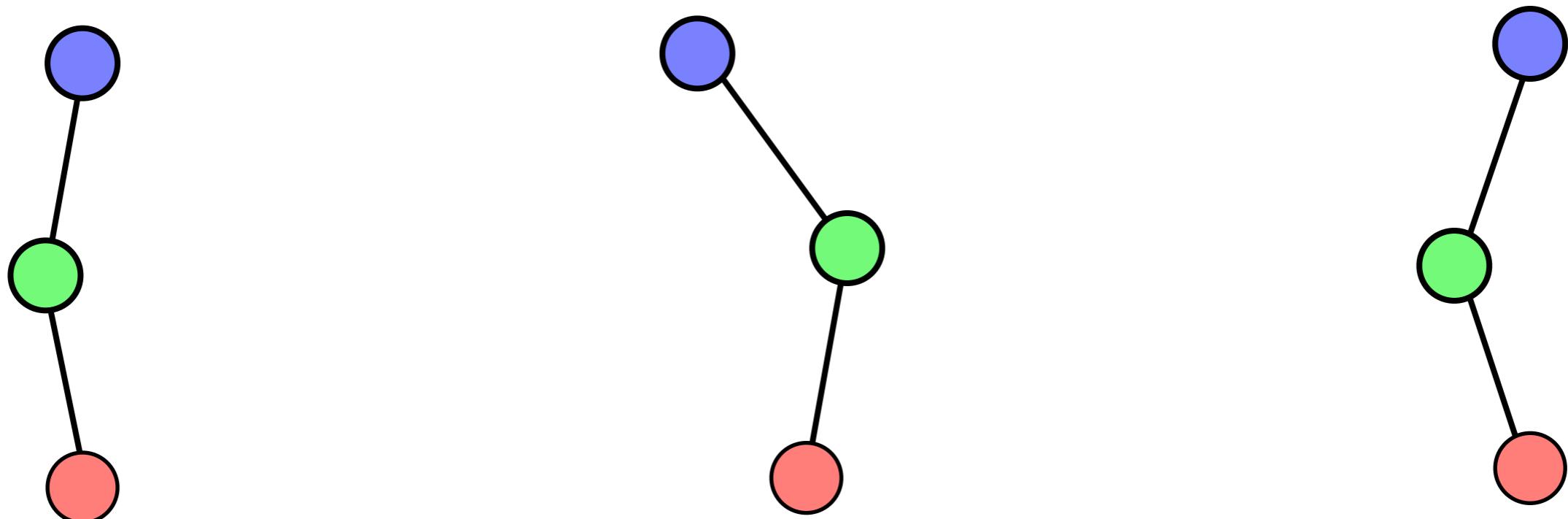
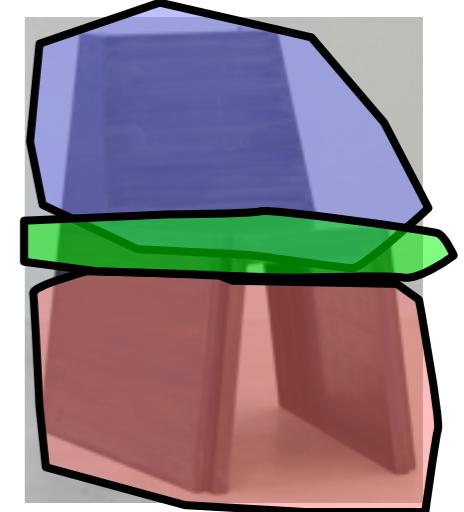
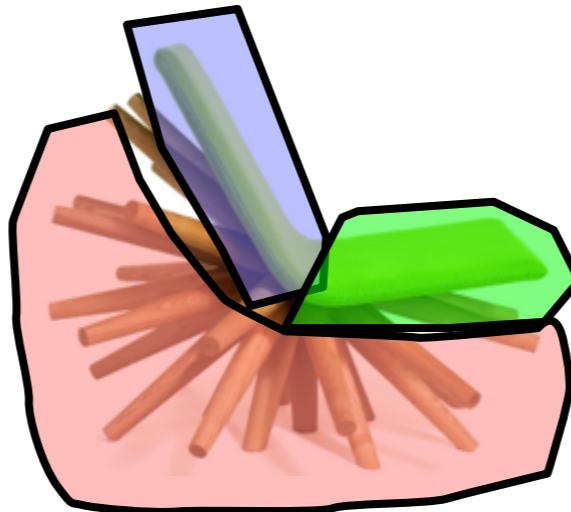
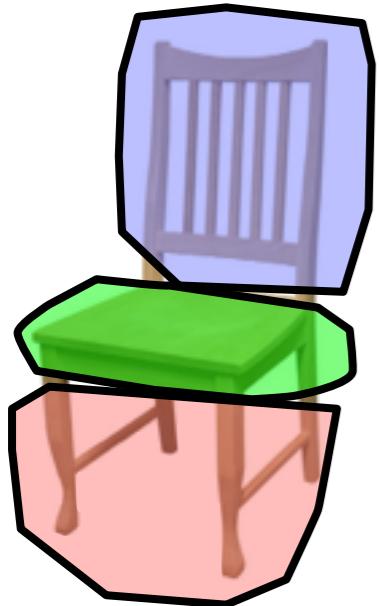
Instead, Focus on Relations



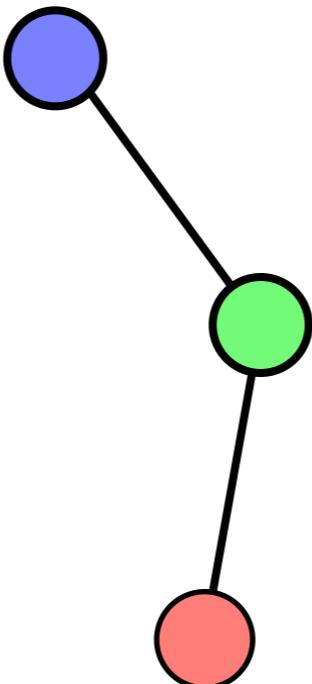
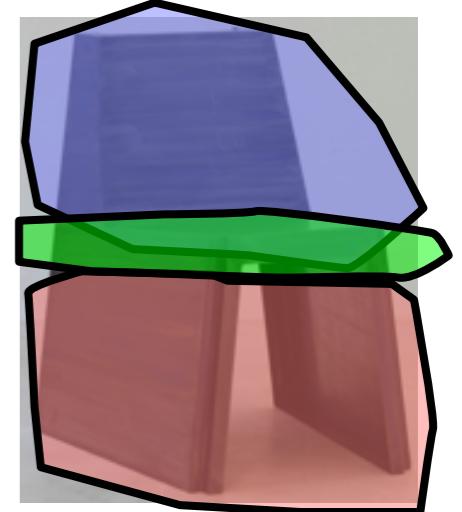
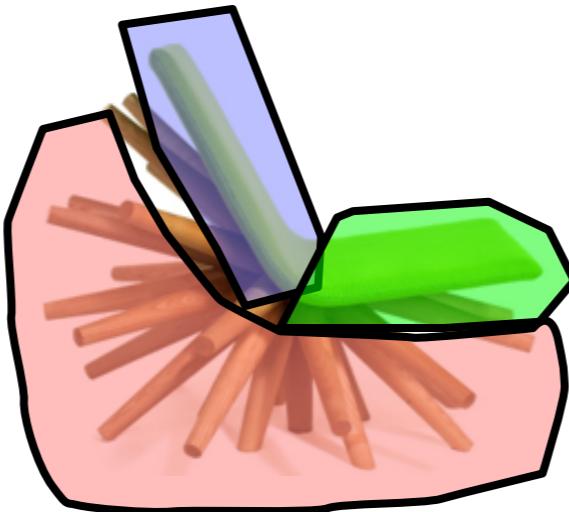
Instead, Focus on Relations



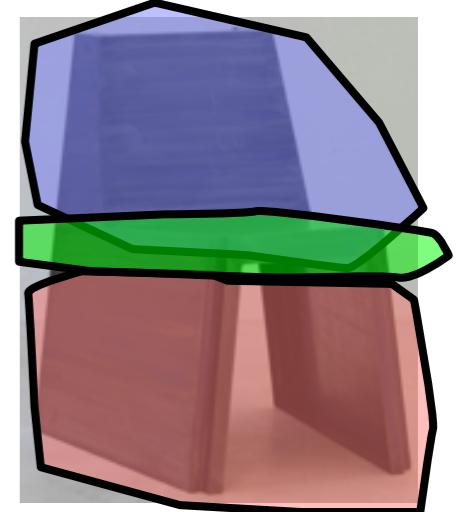
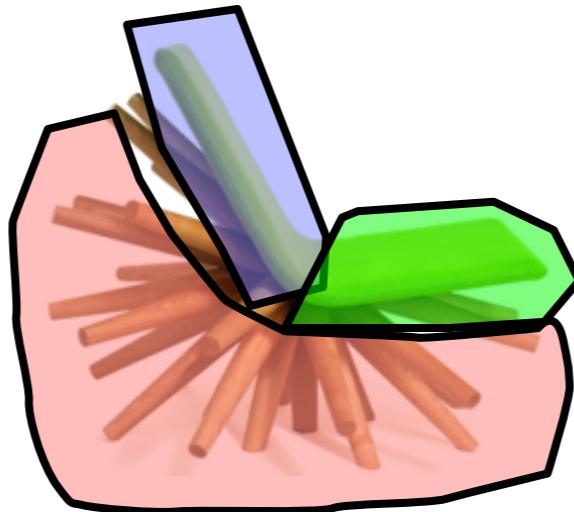
Instead, Focus on Relations



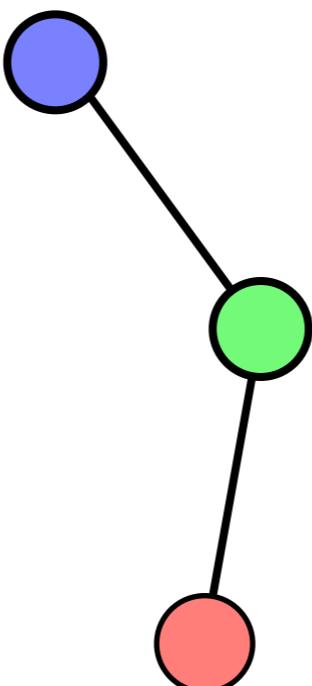
Instead, Focus on Relations



Instead, Focus on Relations



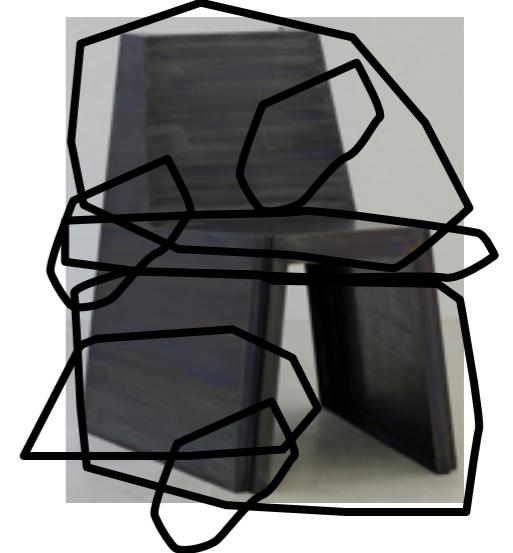
relation graph + variations ~ captures function?



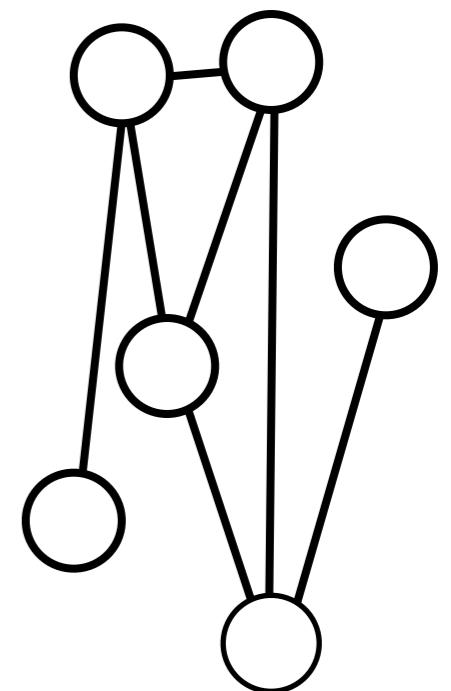
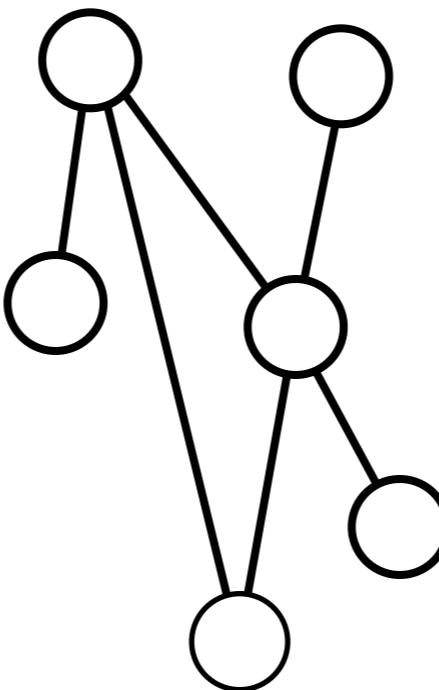
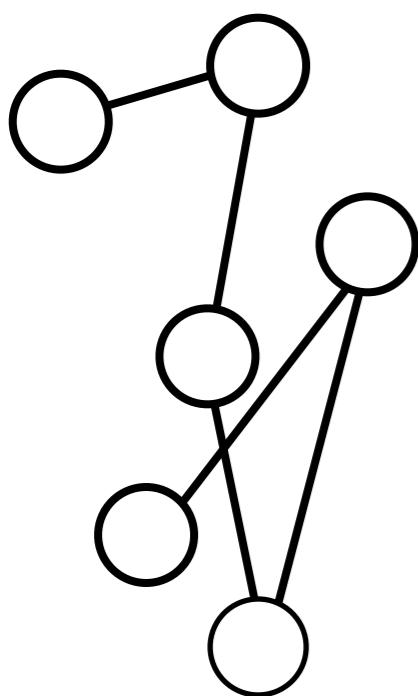
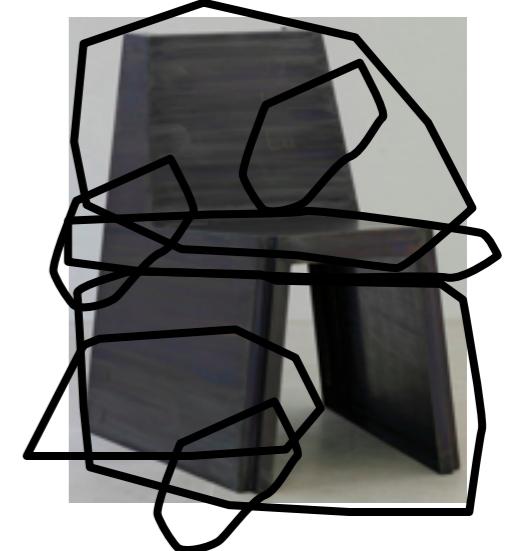
why is this Difficult?



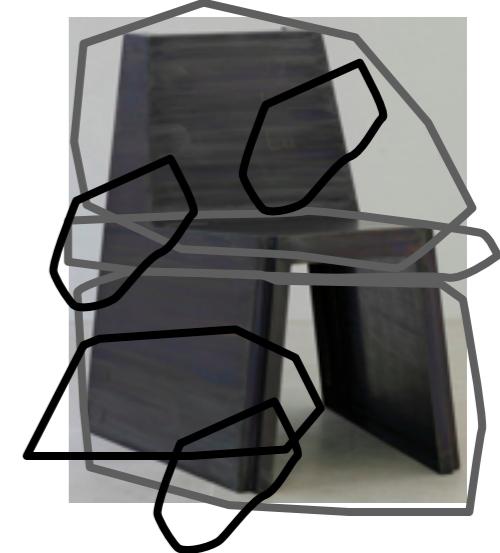
why is this Difficult?



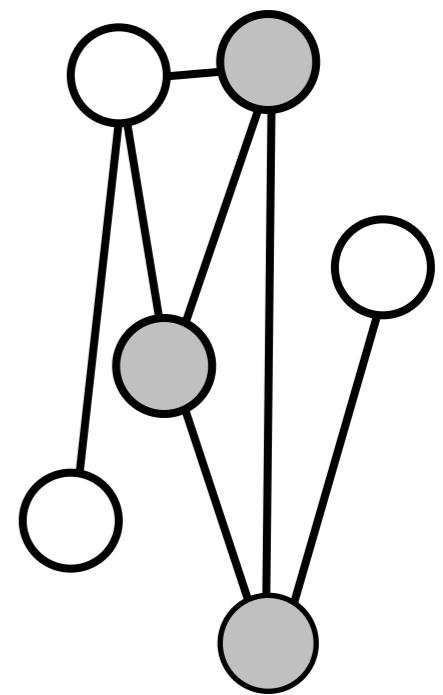
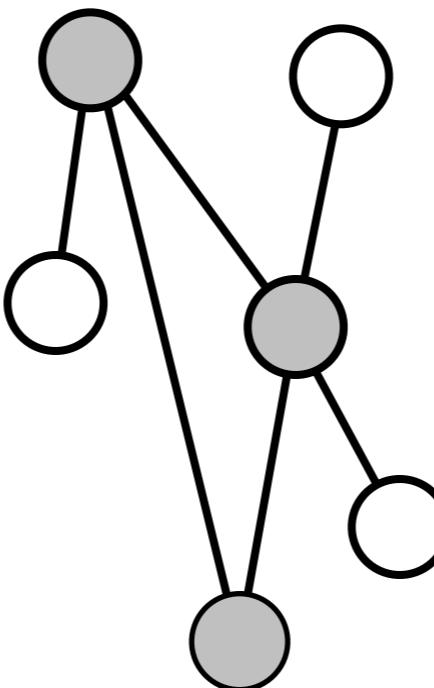
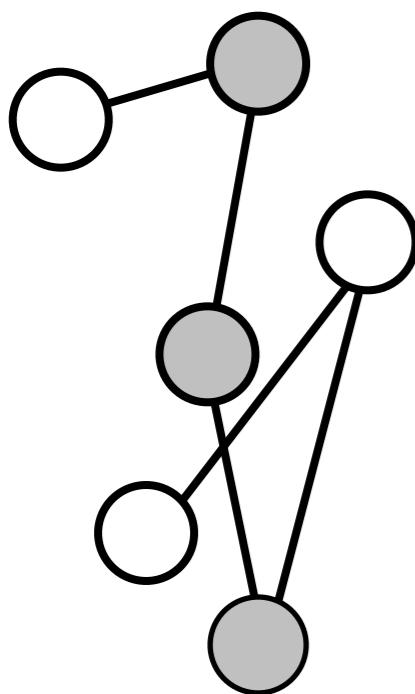
why is this Difficult?



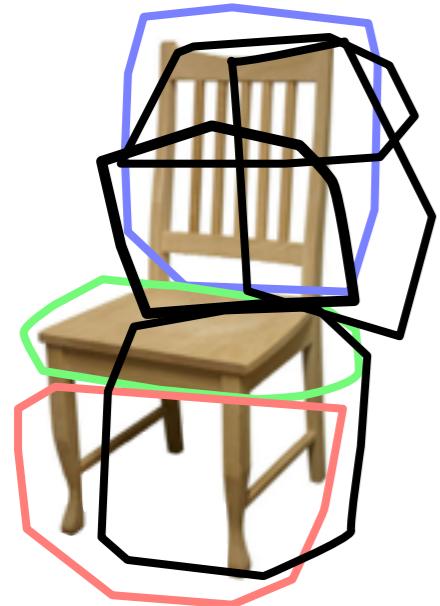
why is this Difficult?



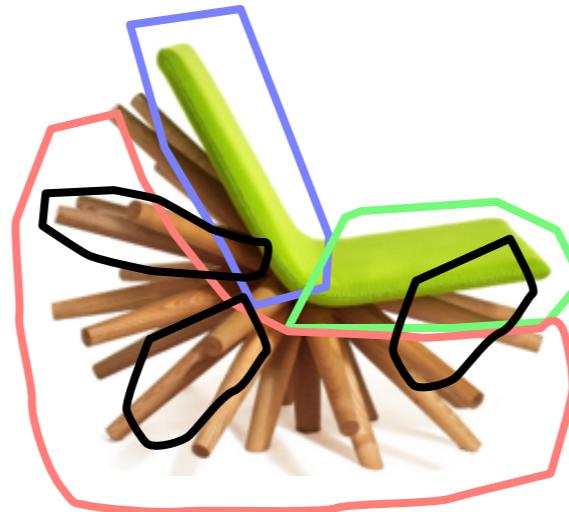
segmentation



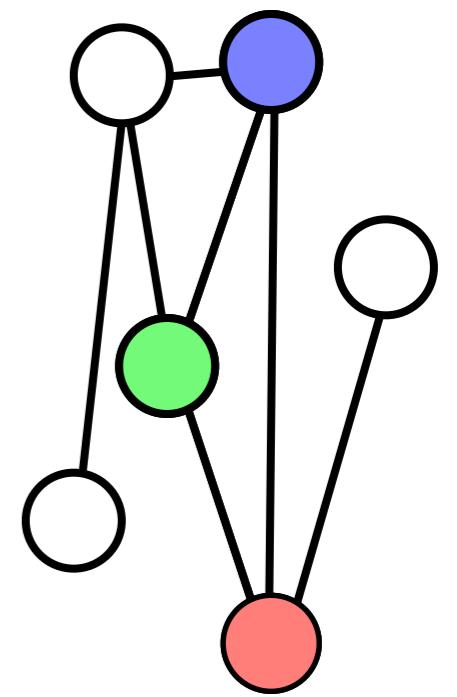
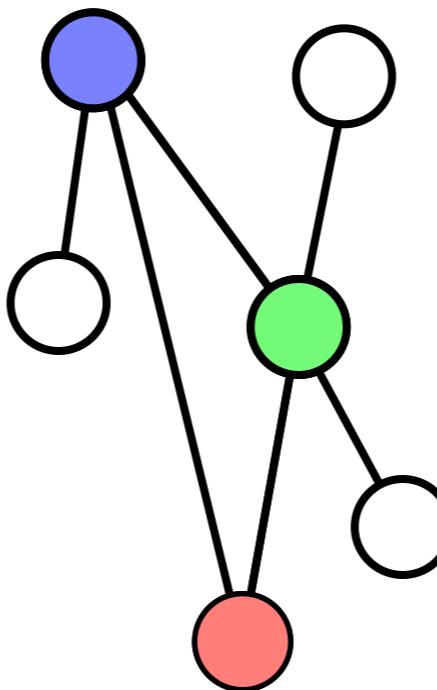
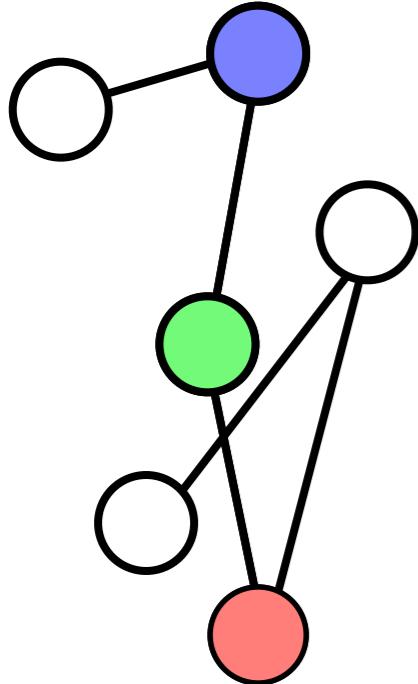
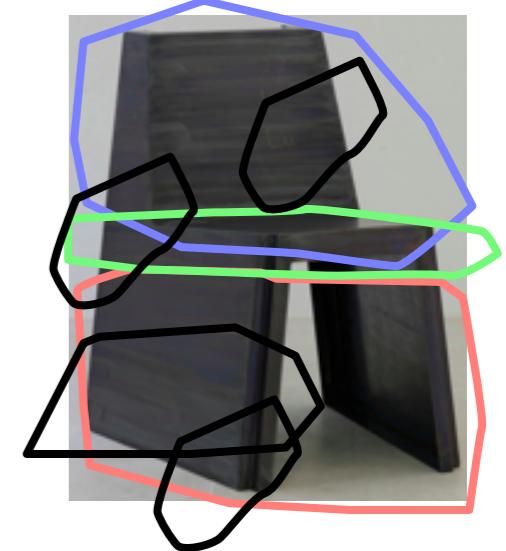
why is this Difficult?



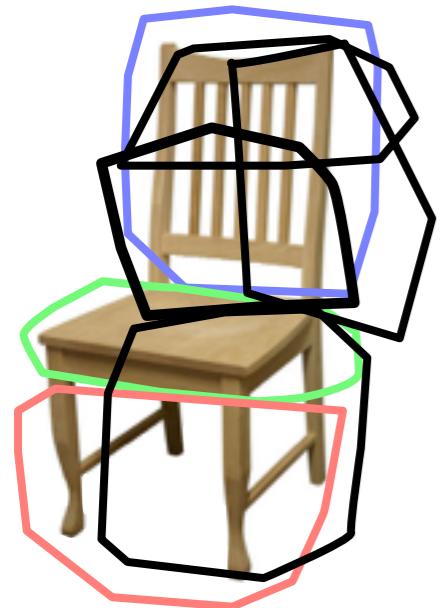
segmentation



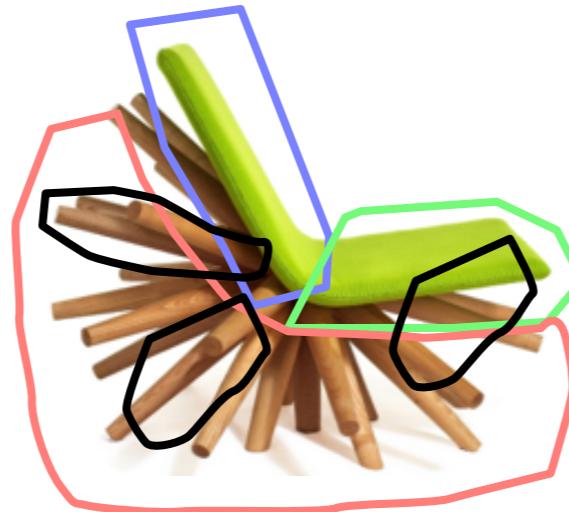
correspondence



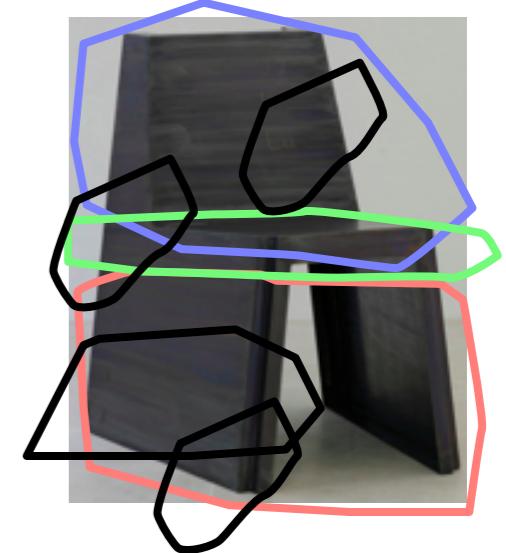
why is this Difficult?



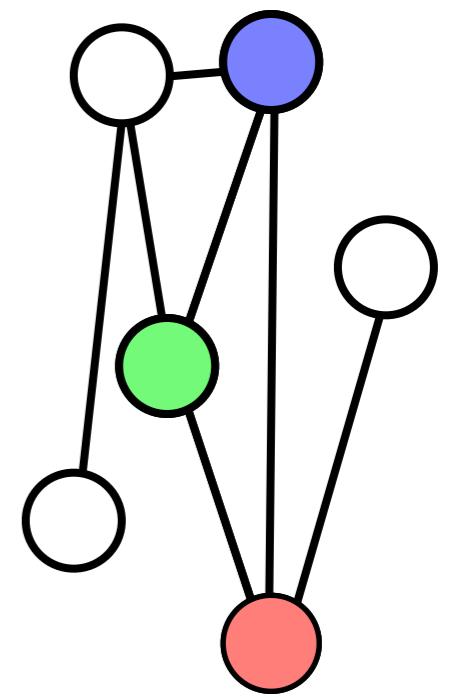
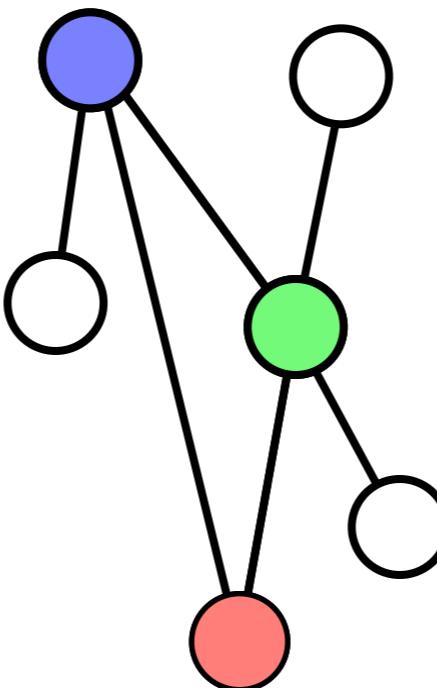
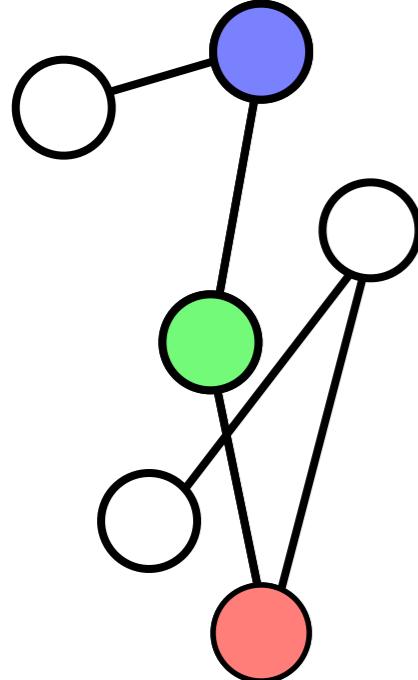
segmentation



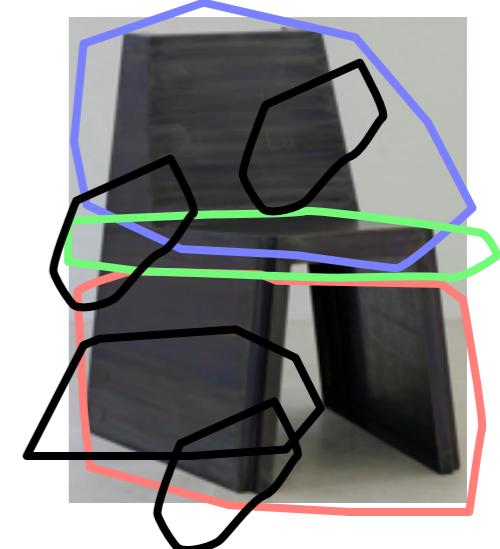
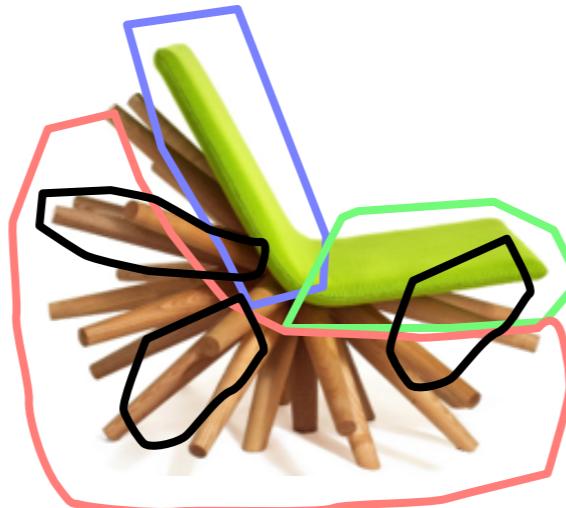
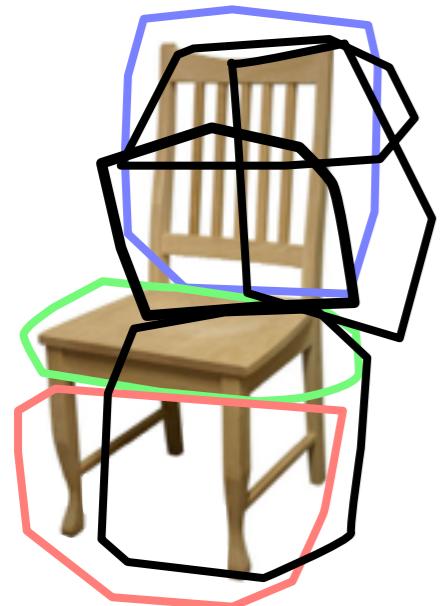
correspondence



variations



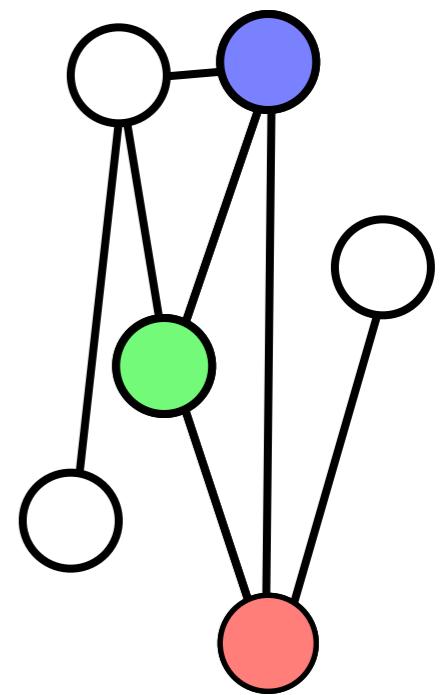
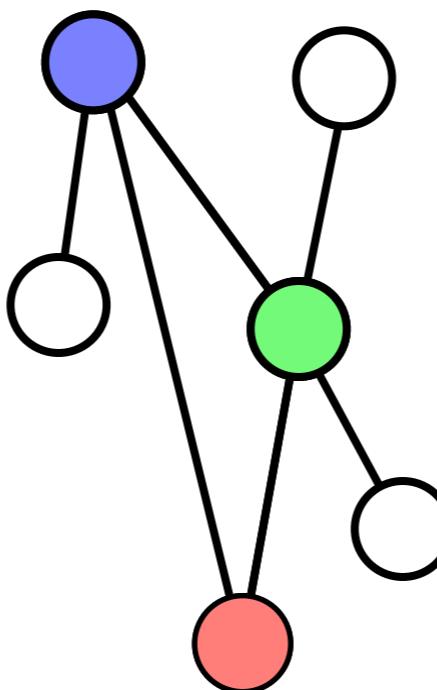
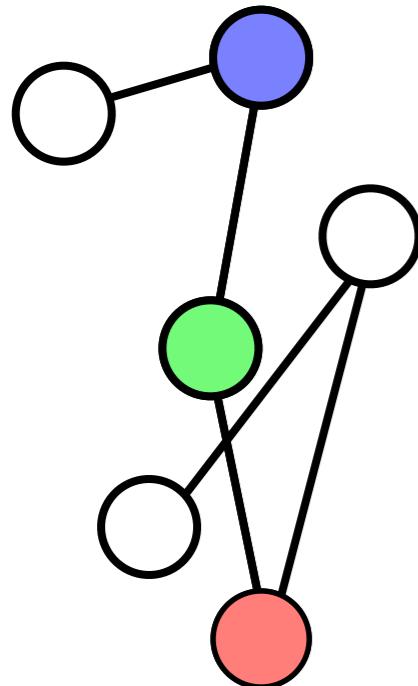
why is this Difficult?



segmentation

correspondence

variations



Relations \Rightarrow Redundancy



Relations \Rightarrow Redundancy



symmetry

Relations \Rightarrow Redundancy



symmetry

Relations \Rightarrow Redundancy



symmetry



Relations \Rightarrow Redundancy



symmetry



Relations \Rightarrow Redundancy



symmetry

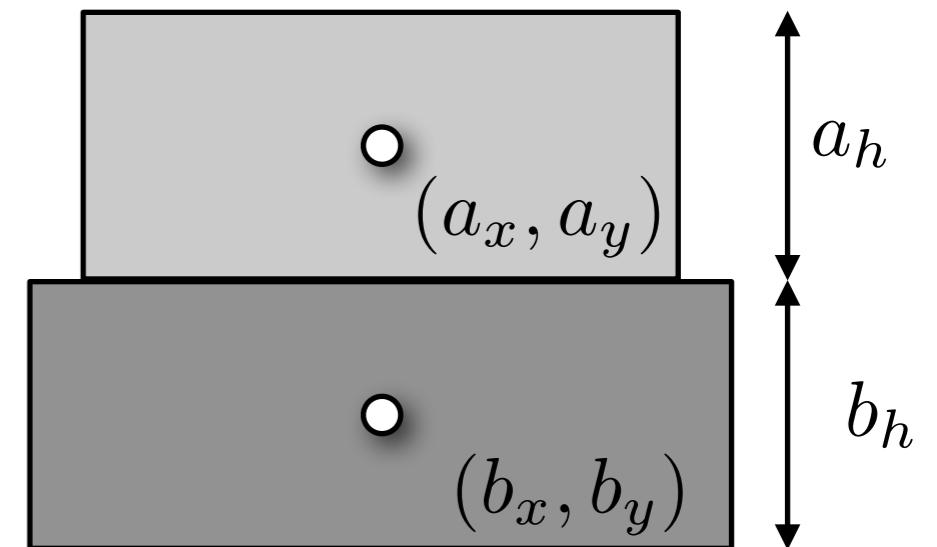


contact

Relations \Rightarrow Redundancy



symmetry



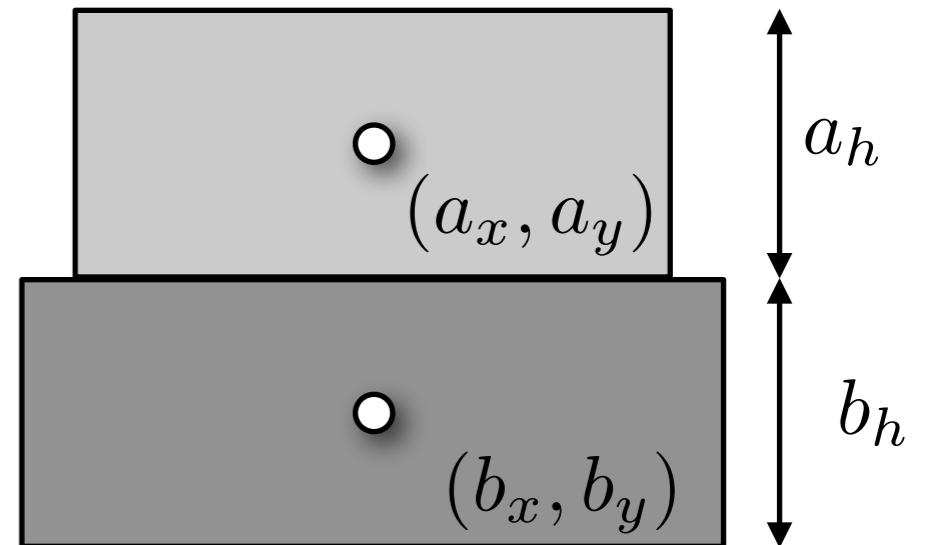
contact

Relations \Rightarrow Redundancy



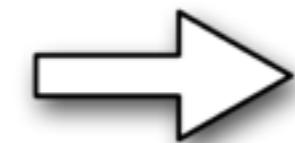
symmetry

$$|a_y - b_y| = \frac{a_h + b_h}{2}$$

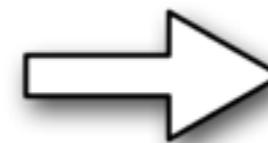
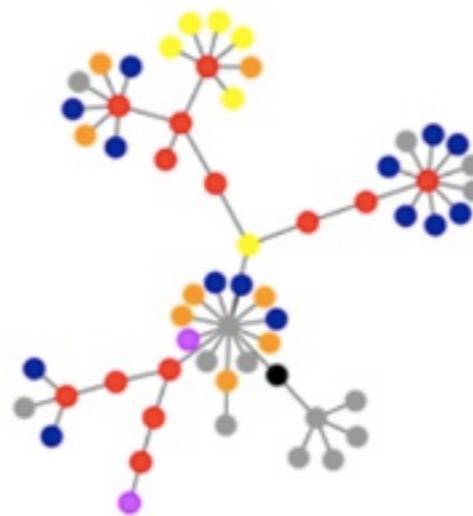


contact

In Search of Simplicity



*encoding
structure*

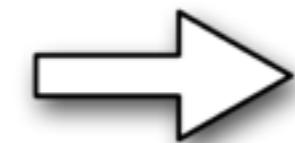
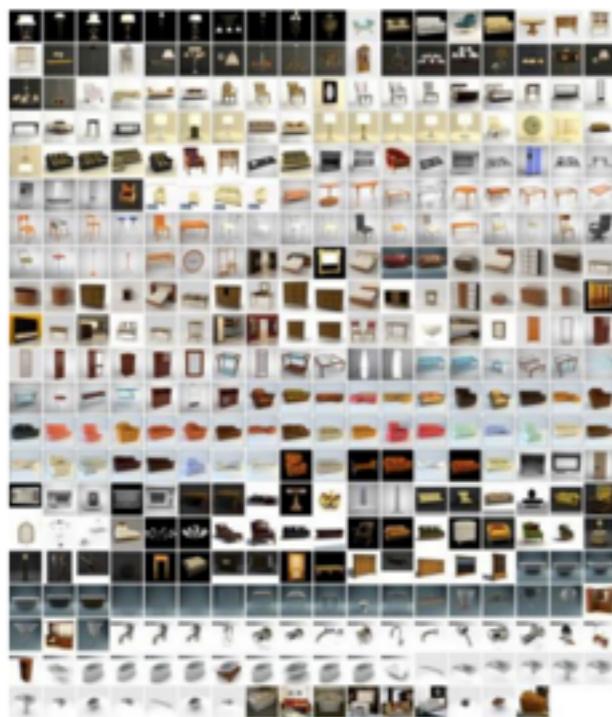


analysis

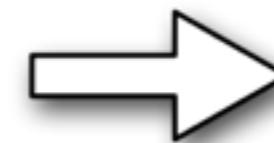
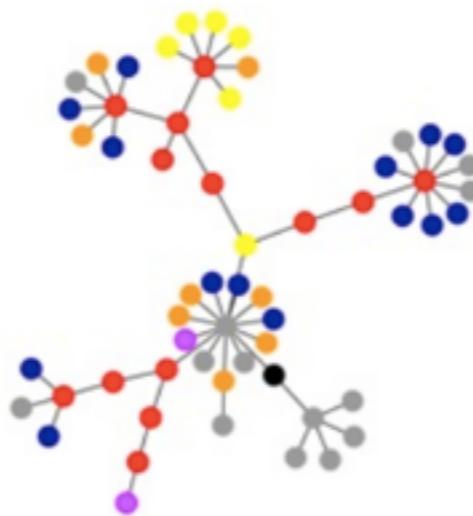
manipulation



In Search of Simplicity



*encoding
structure*



analysis

manipulation



capturing, encoding, and using non-local (algebraic) relations

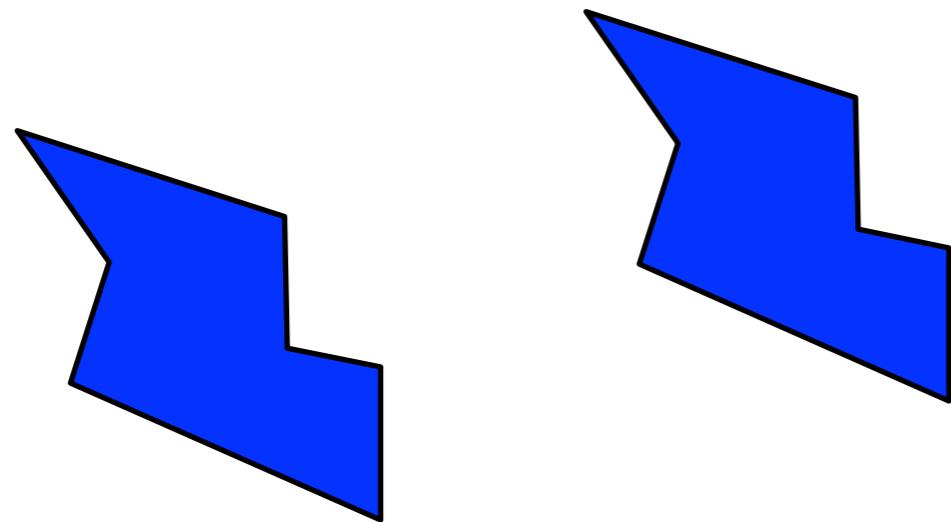
Overview



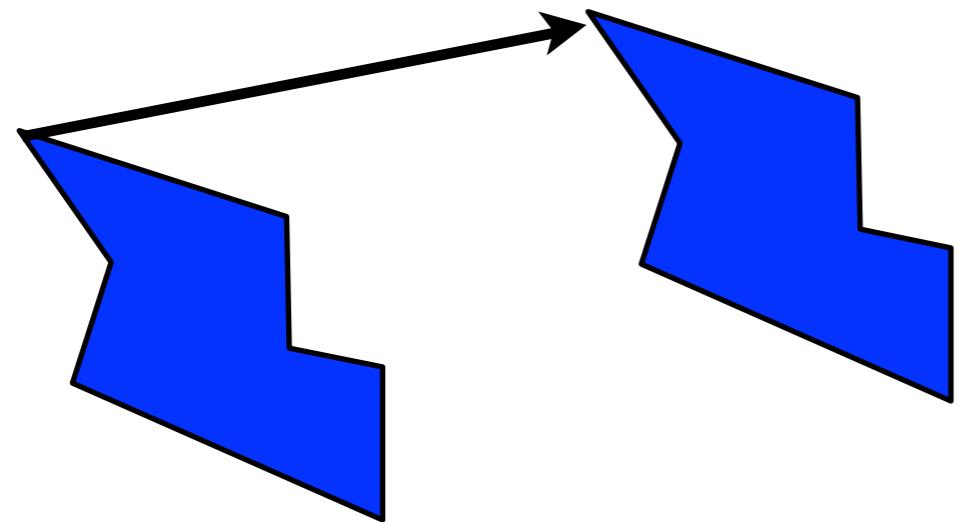
- Motivation
- Factored representations
 - in individual models
 - in model collections
 - in scene collections
- Applications
 - data exploration
 - ‘lifting’ to 3D

individual models

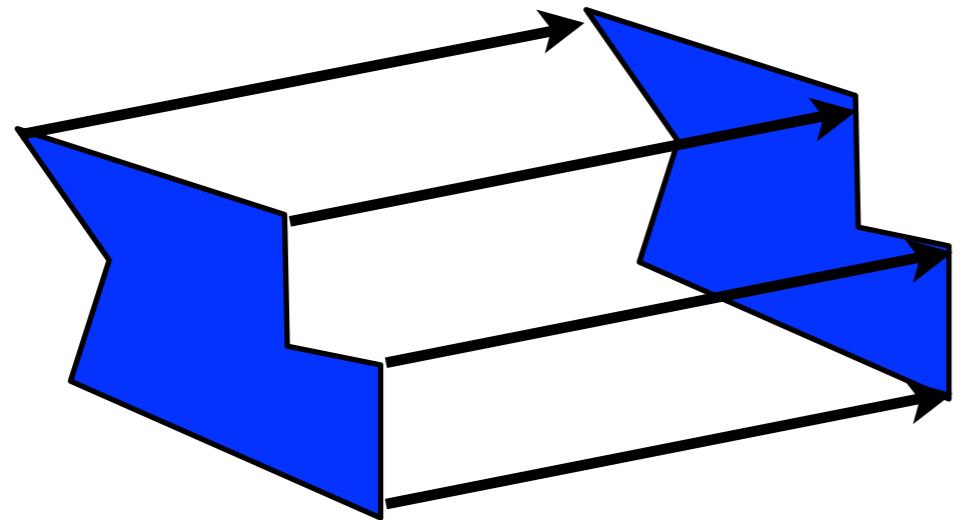
Exact Matching



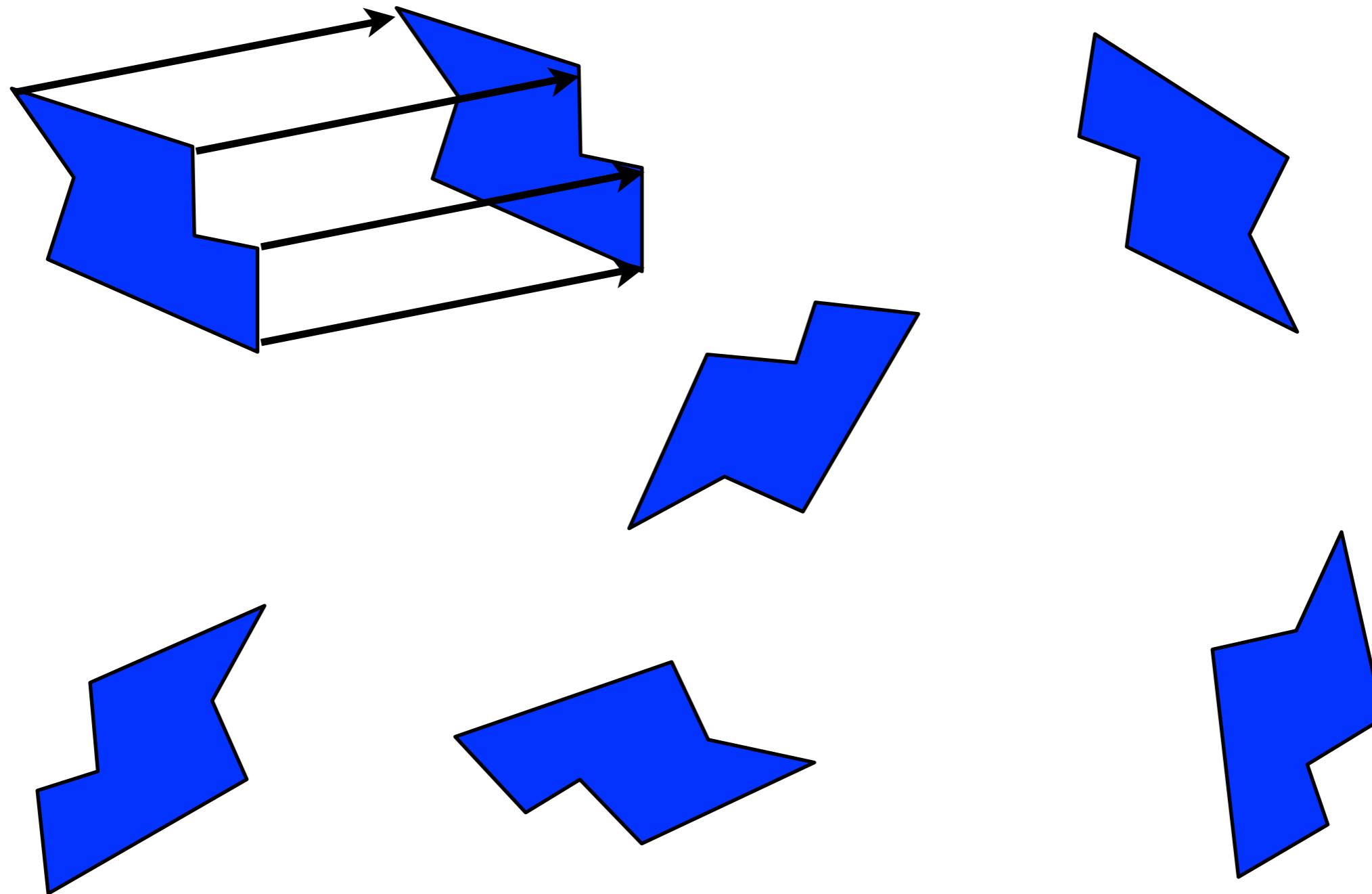
Exact Matching



Exact Matching



Exact Matching

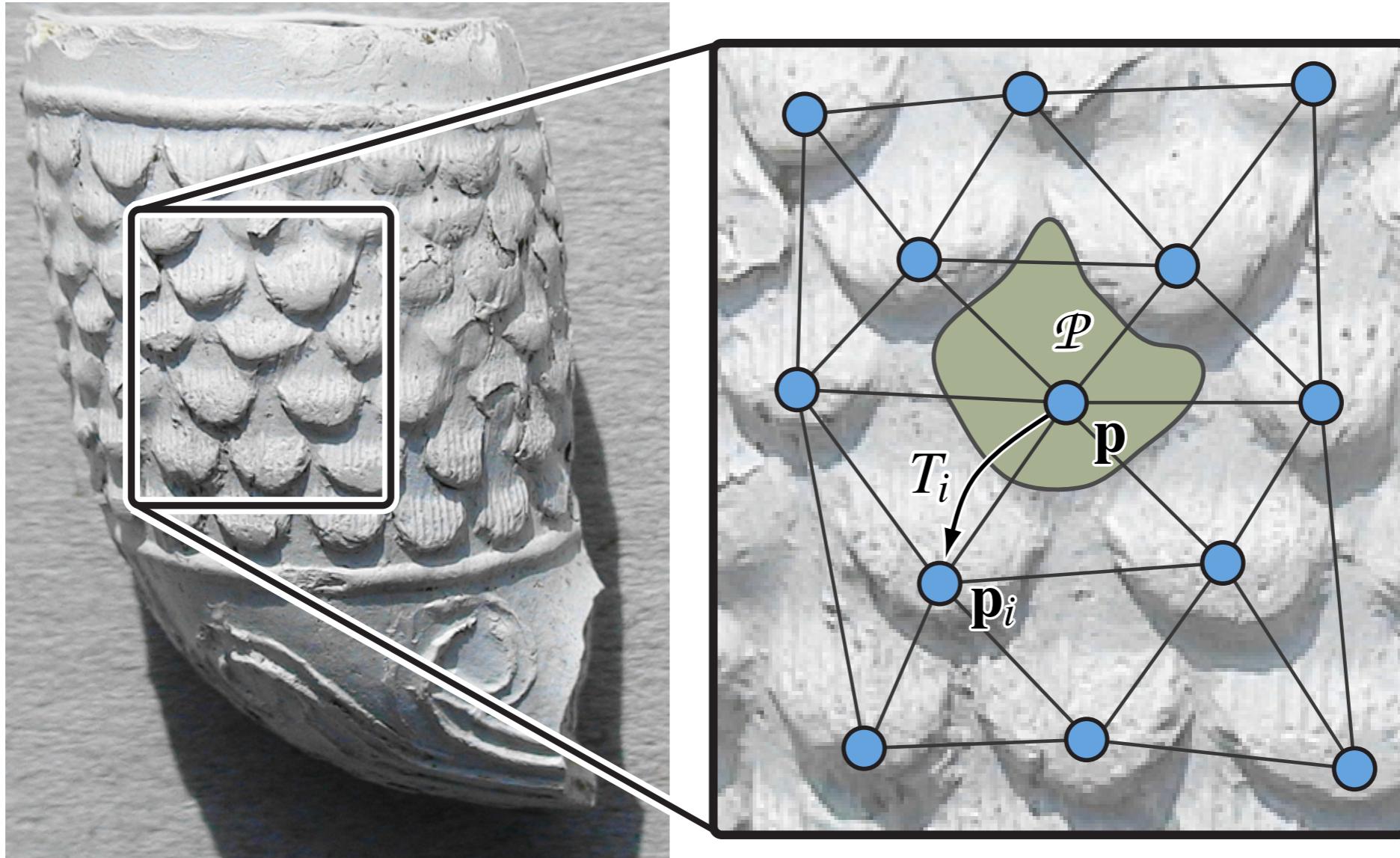


How about Near-Regularity?

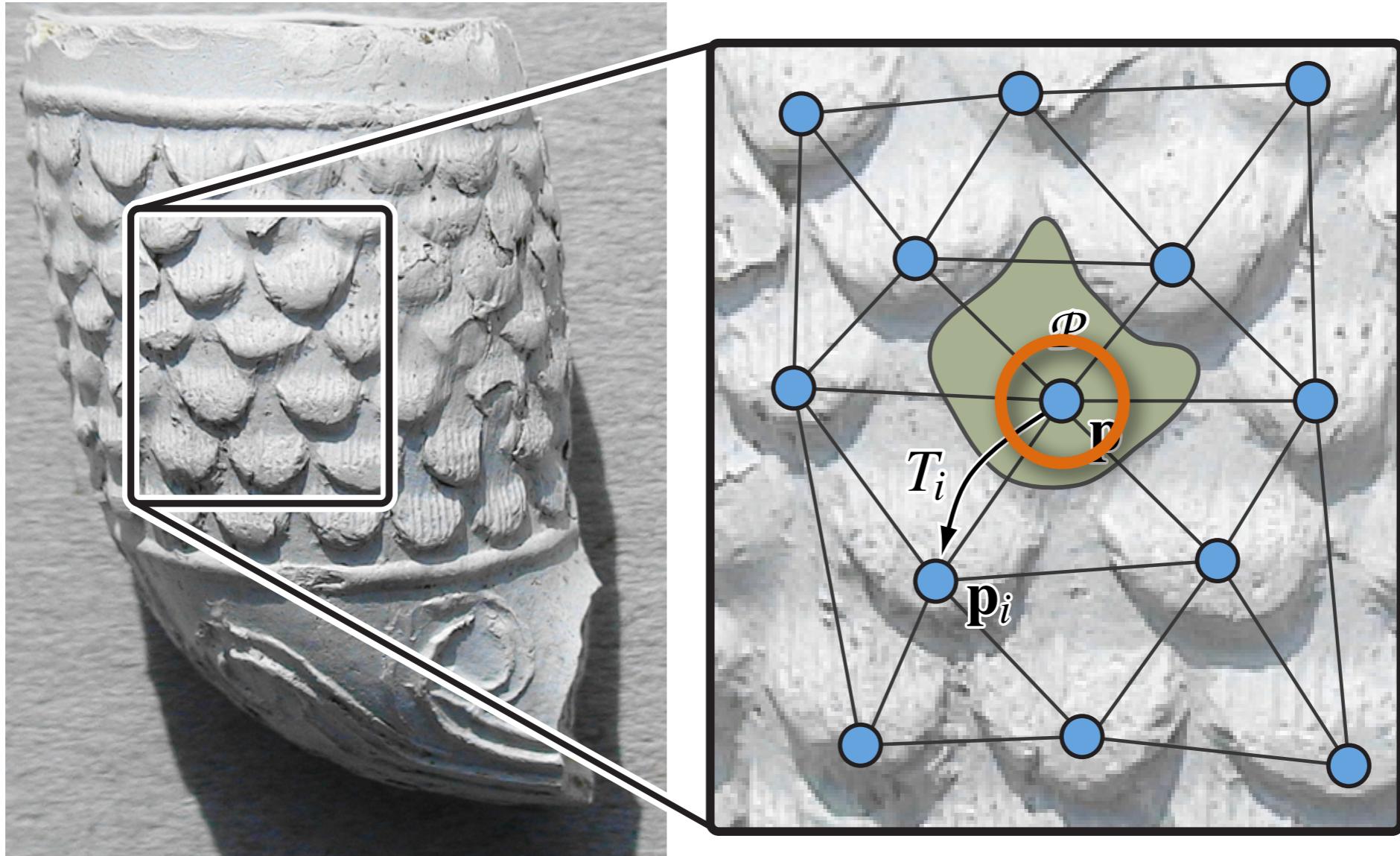


[Transactions on Graphics 2014]

Quantifying Regularity

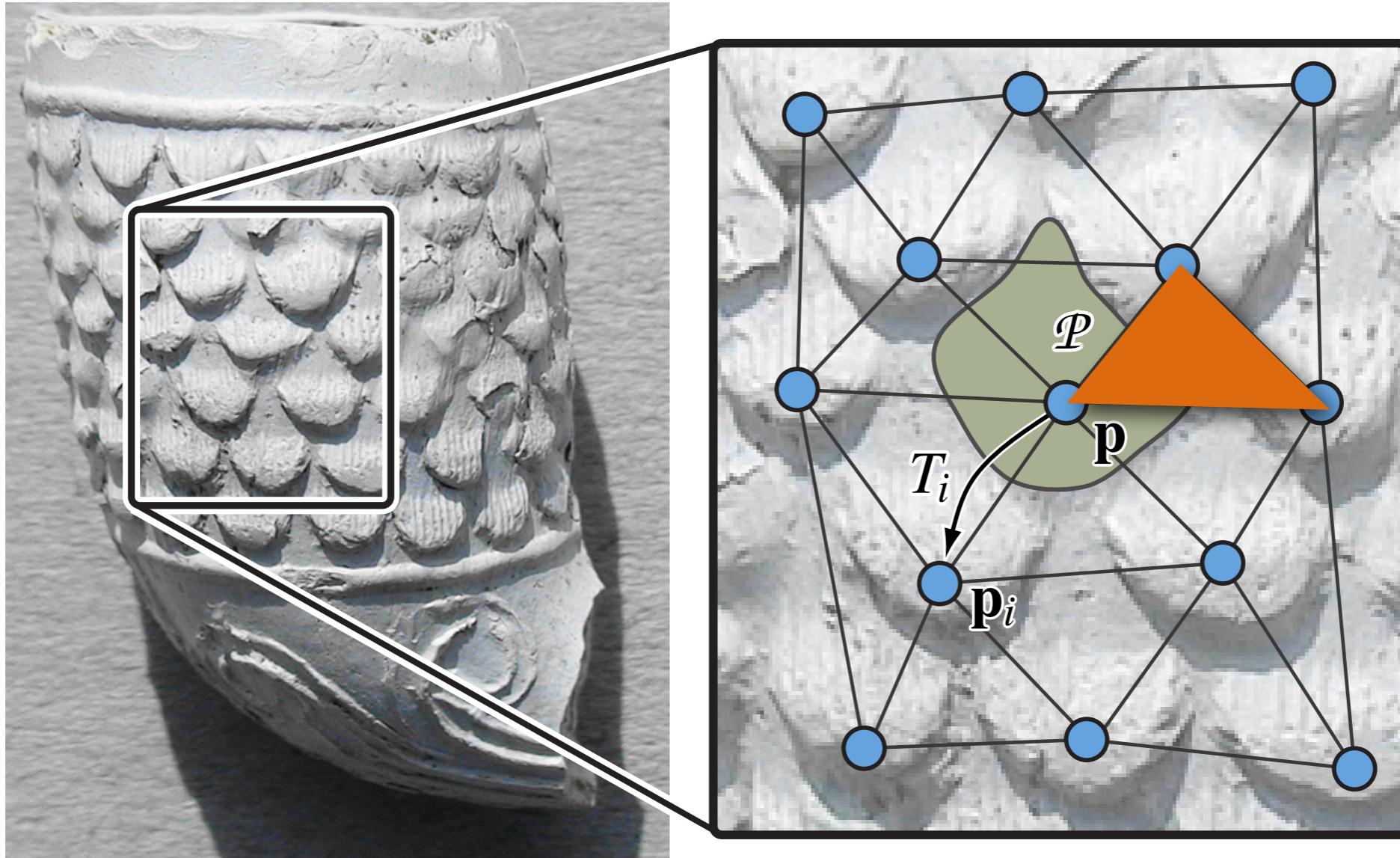


Quantifying Regularity



vertex regularity

Quantifying Regularity



face regularity

Extraction as Selection



$$\min_{z_P, x_v \in \{0,1\}}$$

Regularity Energy

s.t.

patch collection
results in a manifold mesh

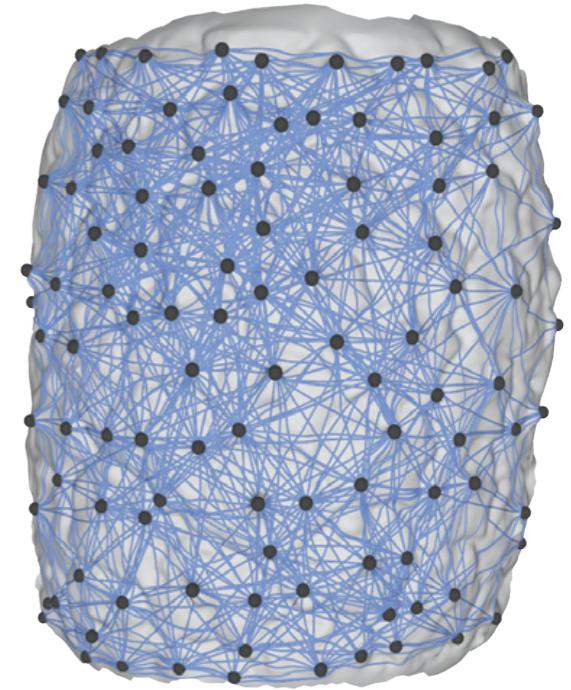
Candidate Curves/Patches



Candidate Curves/Patches



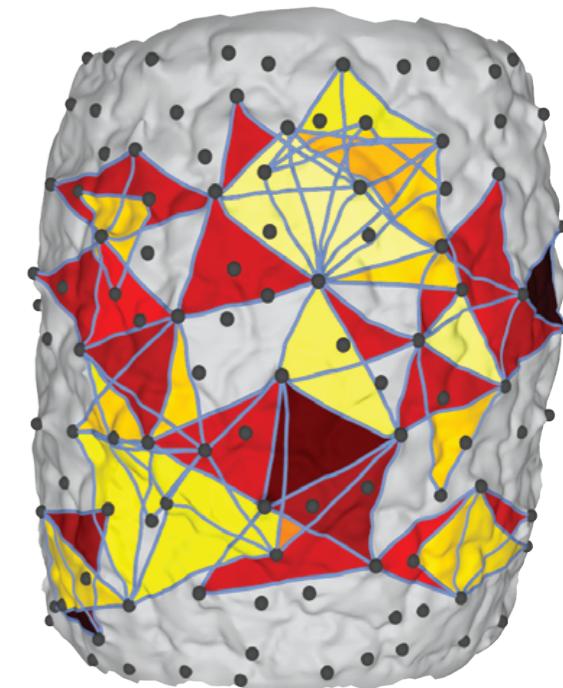
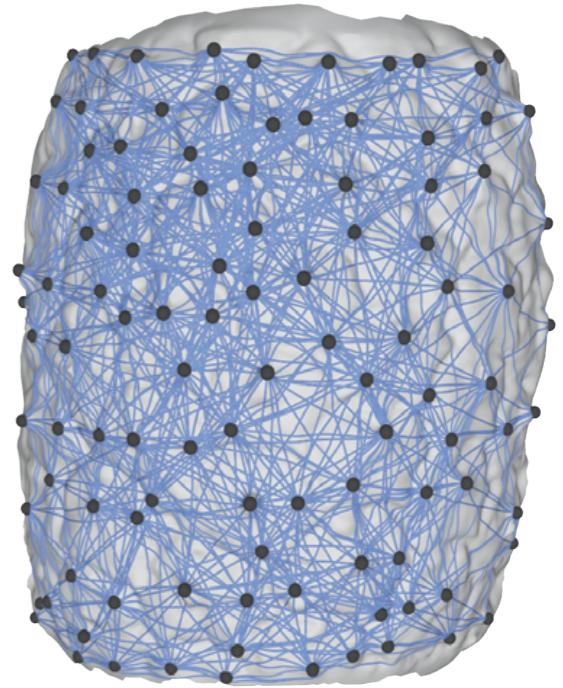
- candidate curves:
 - geodesic curves connecting samples



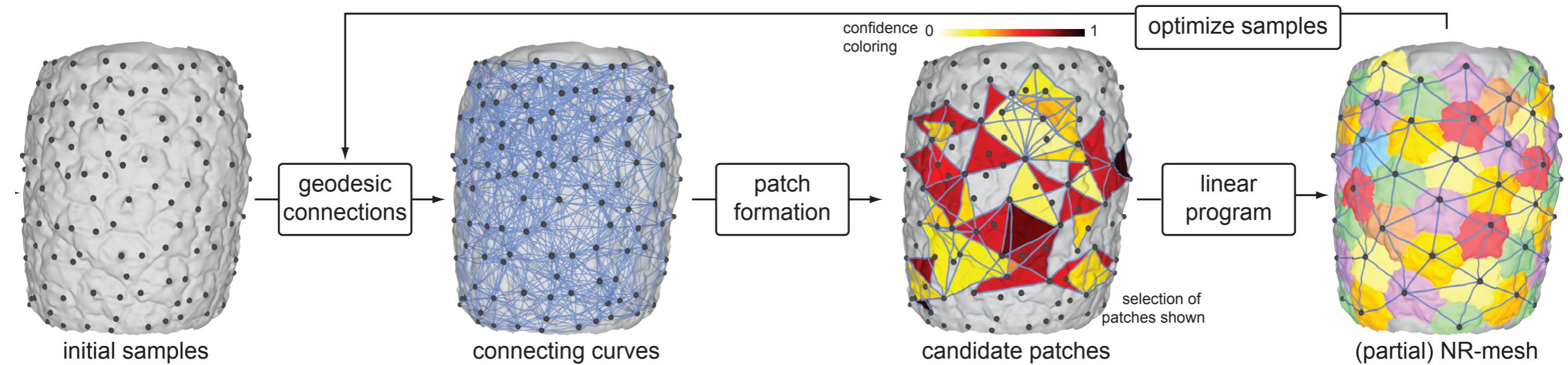
Candidate Curves/Patches



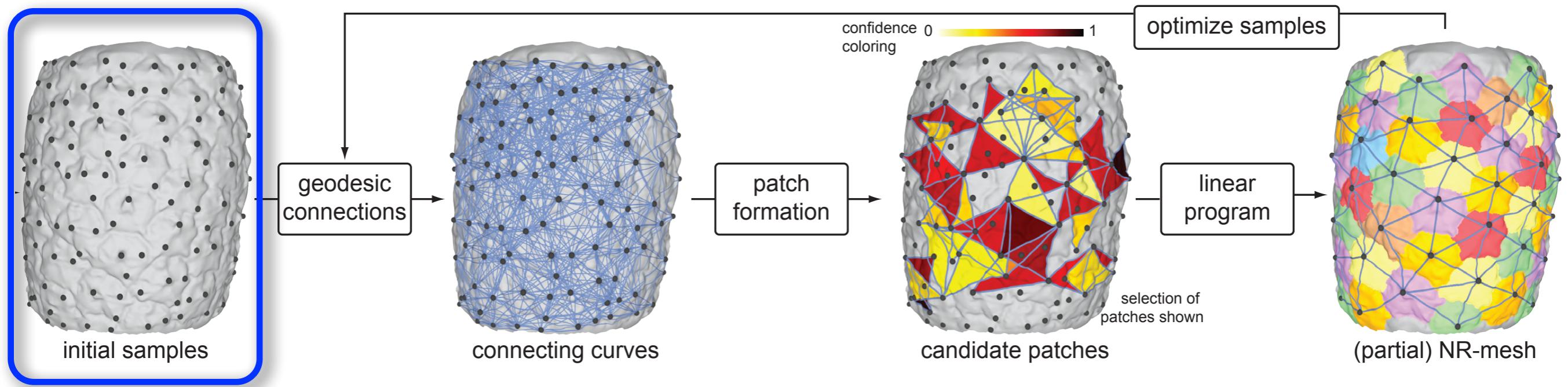
- candidate curves:
 - geodesic curves connecting samples
- candidate patches:
 - self-intersection free loops of 3/4 curves



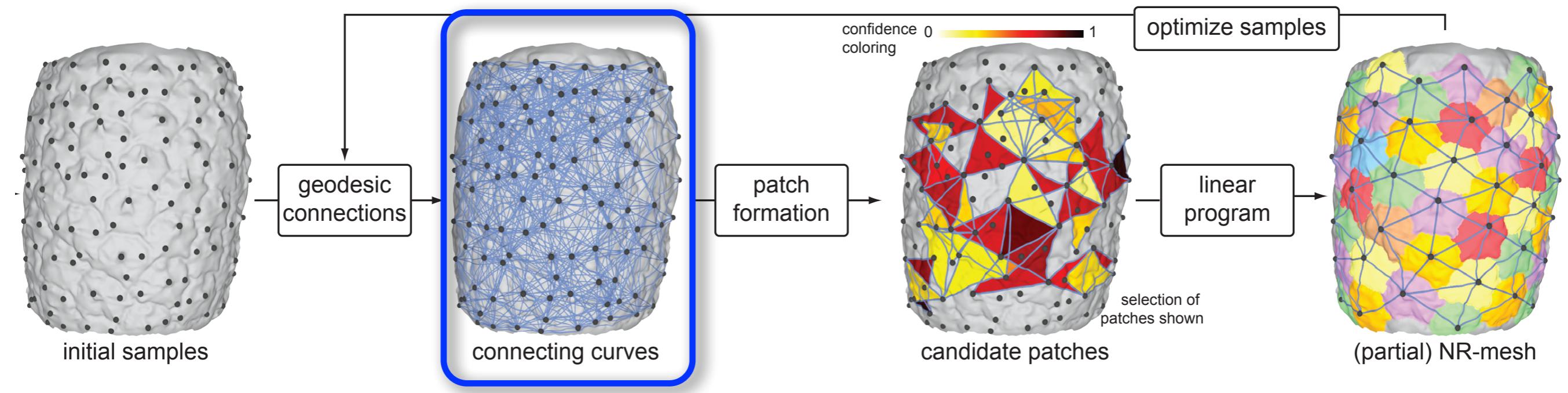
Basic Approach



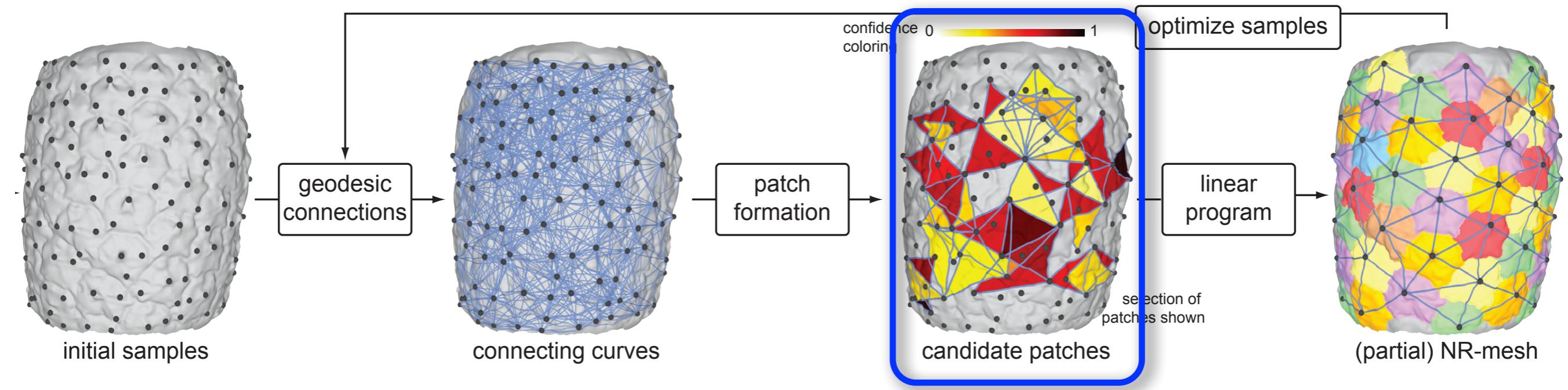
Basic Approach



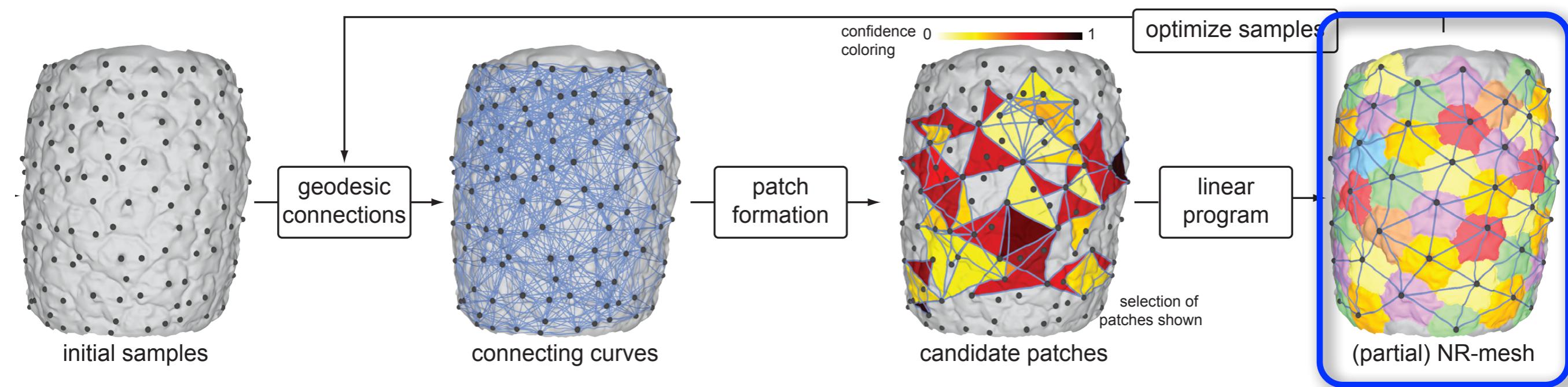
Basic Approach



Basic Approach



Basic Approach



Regularity Energy



Regularity Energy



$z_p = 1$, if patch p is selected.

Regularity Energy



$z_p = 1$, if patch p is selected.

face regularity

$$\sum_p w_p z_p$$

Regularity Energy



$z_p = 1$, if patch p is selected.

face regularity

$$\sum_p w_p z_p$$

$x_v = 1$, if vertex v is selected.

Regularity Energy



$z_p = 1$, if patch p is selected.

face regularity

$$\sum_p w_p z_p$$

$x_v = 1$, if vertex v is selected.

vertex regularity

$$\sum_v w_v \left| \sum_{p \in N(v)} z_p - \eta x_v \right|$$

What makes a Valid Patch Set?

Curve constraint:

$$d_{\mathcal{M}}(c) := |\mathcal{A}_L(c, \mathcal{M}_P)| - |\mathcal{A}_R(c, \mathcal{M}_P)| = 0$$

Uniqueness constraint:

$$|\mathcal{M}_P(f_i)| = 1$$

Integer Programming Formulation



$$\min_{z_P, x_v \in \{0,1\}}$$

Regularity Energy

s.t.

patch collection
results in a manifold mesh

Integer Programming Formulation



$$\min_{z_P, x_v \in \{0,1\}} - \sum_{P \in \mathcal{P}} w_P z_P + \lambda_v \sum_{v \in \mathcal{S}} w_v \left| \sum_{P \in \mathcal{A}_P(v)} z_P - \eta x_v \right|$$

s.t.

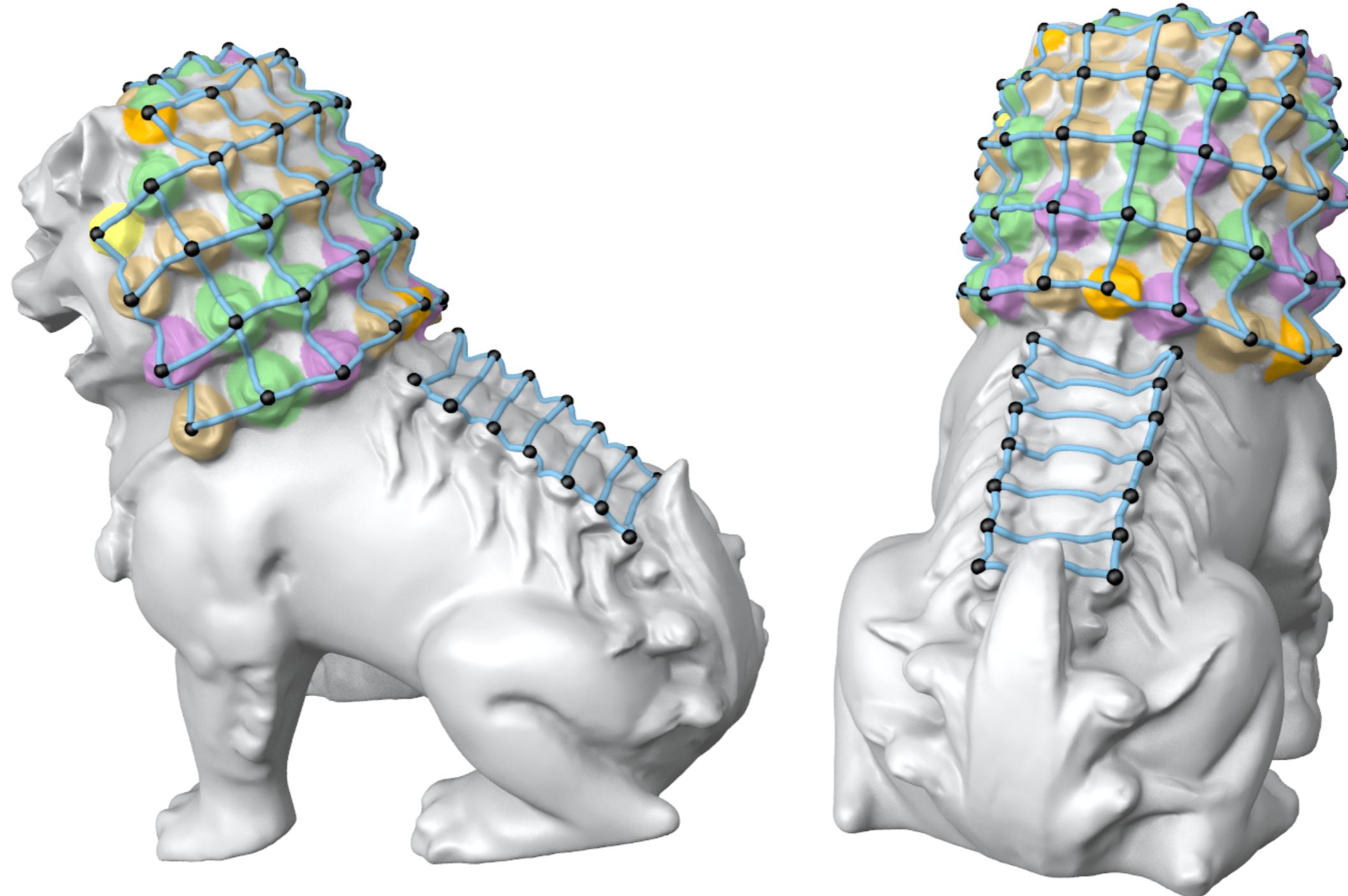
patch collection
results in a manifold mesh

Integer Programming Formulation



$$\begin{aligned}
 \min_{z_P, x_v \in \{0,1\}} \quad & - \sum_{P \in \mathcal{P}} w_P z_P + \lambda_v \sum_{v \in \mathcal{S}} w_v \left| \sum_{P \in \mathcal{A}_P(v)} z_P - \eta x_v \right| \\
 \text{s.t.} \quad & z_P \leq x_v, \quad \forall v \in \mathcal{P}, \quad \forall P \in \mathcal{P}, \\
 & \sum_{P \in \mathcal{A}_L(c)} z_P = \sum_{P \in \mathcal{A}_R(c)} z_P, \quad c \in \mathcal{C}, \\
 & \sum_{P \in \mathcal{C}(f)} z_P = 1.
 \end{aligned}$$

Asian Lion



model collections

Structure among Relations



Structure among Relations



aRb and $bRc \Rightarrow aRc$

Structure among Relations



$$aRb \quad \text{and} \quad bRc \Rightarrow aRc$$

$$T_{ab} \cdot T_{bc} \cdot T_{ca} = I$$

Structure among Relations



$$aRb \quad \text{and} \quad bRc \Rightarrow aRc$$

$$T_{ab} \cdot T_{bc} \cdot T_{ca} = I$$

$$\begin{bmatrix} a_{00} & a_{10} & a_{20} & a_{30} & a_{40} \\ a_{01} & a_{11} & a_{21} & a_{31} & a_{41} \\ a_{02} & a_{12} & a_{22} & a_{32} & a_{42} \\ a_{03} & a_{13} & a_{23} & a_{33} & a_{43} \\ a_{04} & a_{14} & a_{24} & a_{34} & a_{44} \end{bmatrix}$$

Structure among Relations



$$aRb \quad \text{and} \quad bRc \Rightarrow aRc$$

$$T_{ab} \cdot T_{bc} \cdot T_{ca} = I$$

$$\begin{bmatrix} a_{00} & a_{10} & a_{20} & a_{30} & a_{40} \\ \cdot & a_{11} & a_{21} & a_{31} & a_{41} \\ \cdot & \cdot & a_{22} & a_{32} & a_{42} \\ \cdot & \cdot & \cdot & a_{33} & a_{43} \\ \cdot & \cdot & \cdot & \cdot & a_{44} \end{bmatrix}$$

Structure among Relations



$$aRb \quad \text{and} \quad bRc \Rightarrow aRc$$

$$T_{ab} \cdot T_{bc} \cdot T_{ca} = I$$

$$\begin{bmatrix} a_{00} & a_{10} & ? & a_{30} & ? \\ . & a_{11} & a_{21} & ? & a_{41} \\ . & . & a_{22} & a_{32} & a_{42} \\ . & . & . & a_{33} & a_{43} \\ . & . & . & . & a_{44} \end{bmatrix}$$

Structure among Relations



$$aRb \quad \text{and} \quad bRc \Rightarrow aRc$$

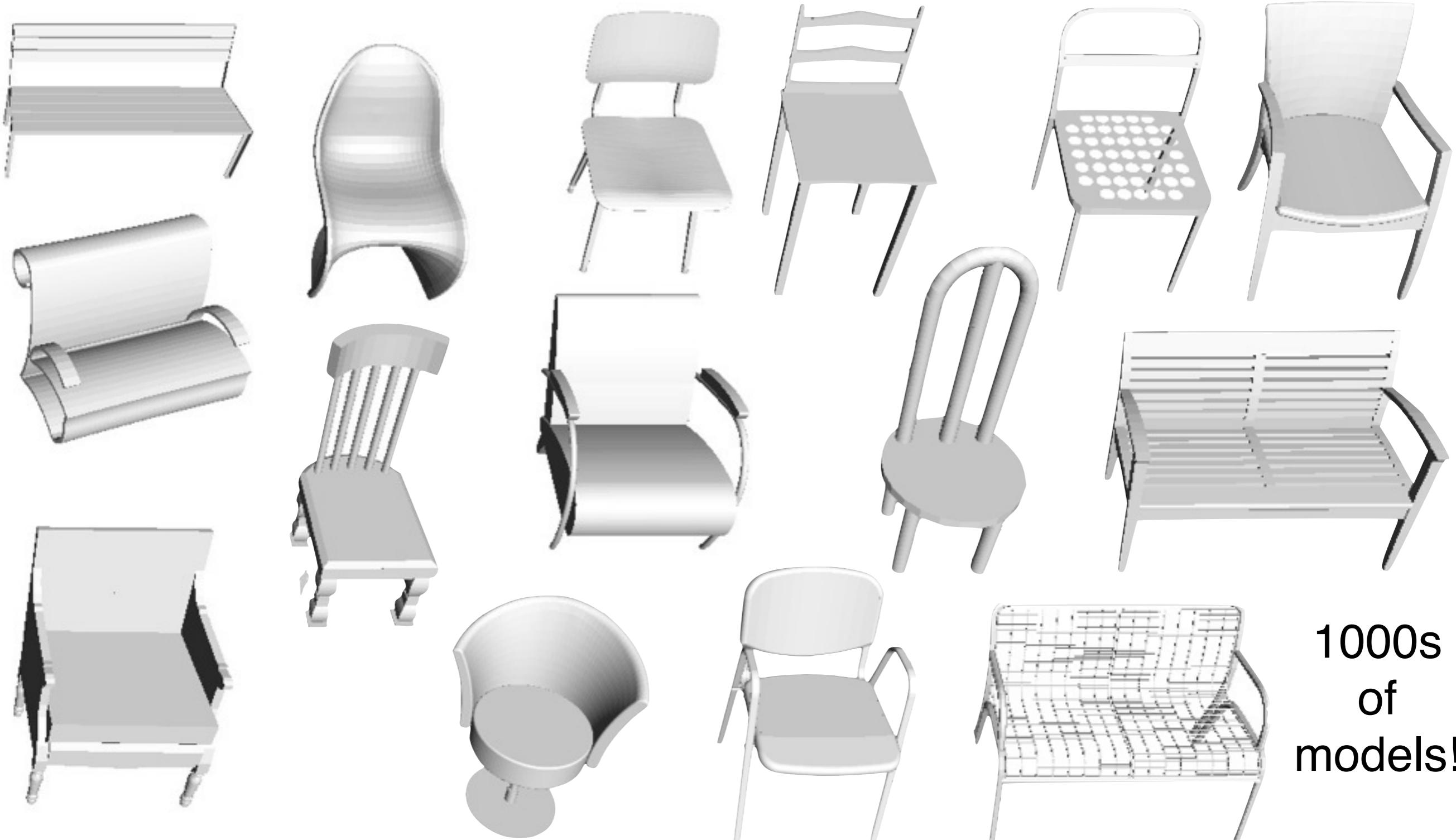
$$T_{ab} \cdot T_{bc} \cdot T_{ca} = I$$

exploit ‘structure’
for

compactness (redundancy)
robustness (constraints)

a_{00}	a_{10}	?	a_{30}	?
.	a_{11}	a_{21}	?	a_{41}
.	.	a_{22}	a_{32}	a_{42}
.	.	.	a_{33}	a_{43}
.	.	.	.	a_{44}

Unorganized Model Collections

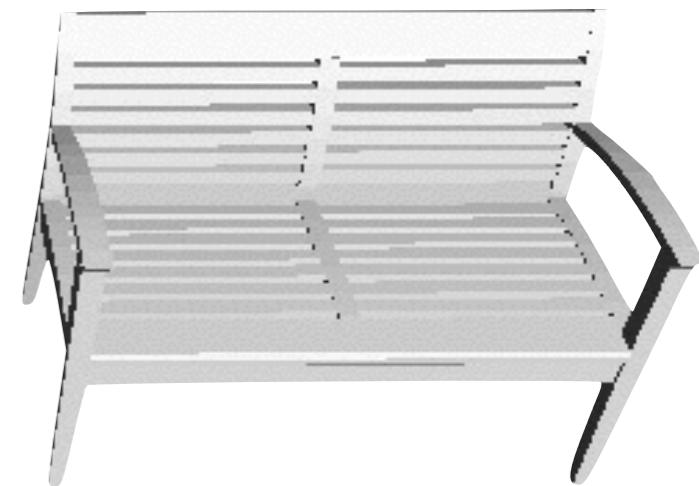
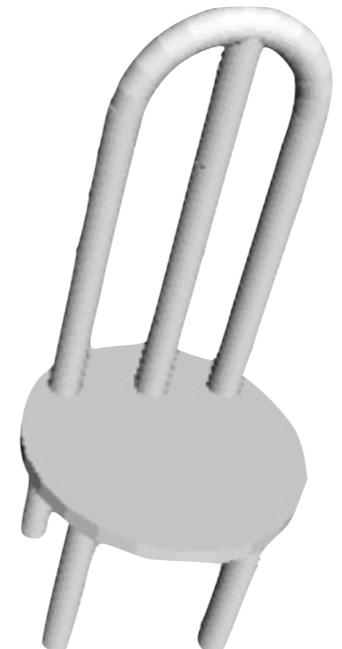


1000s
of
models!

Why Joint Analysis?



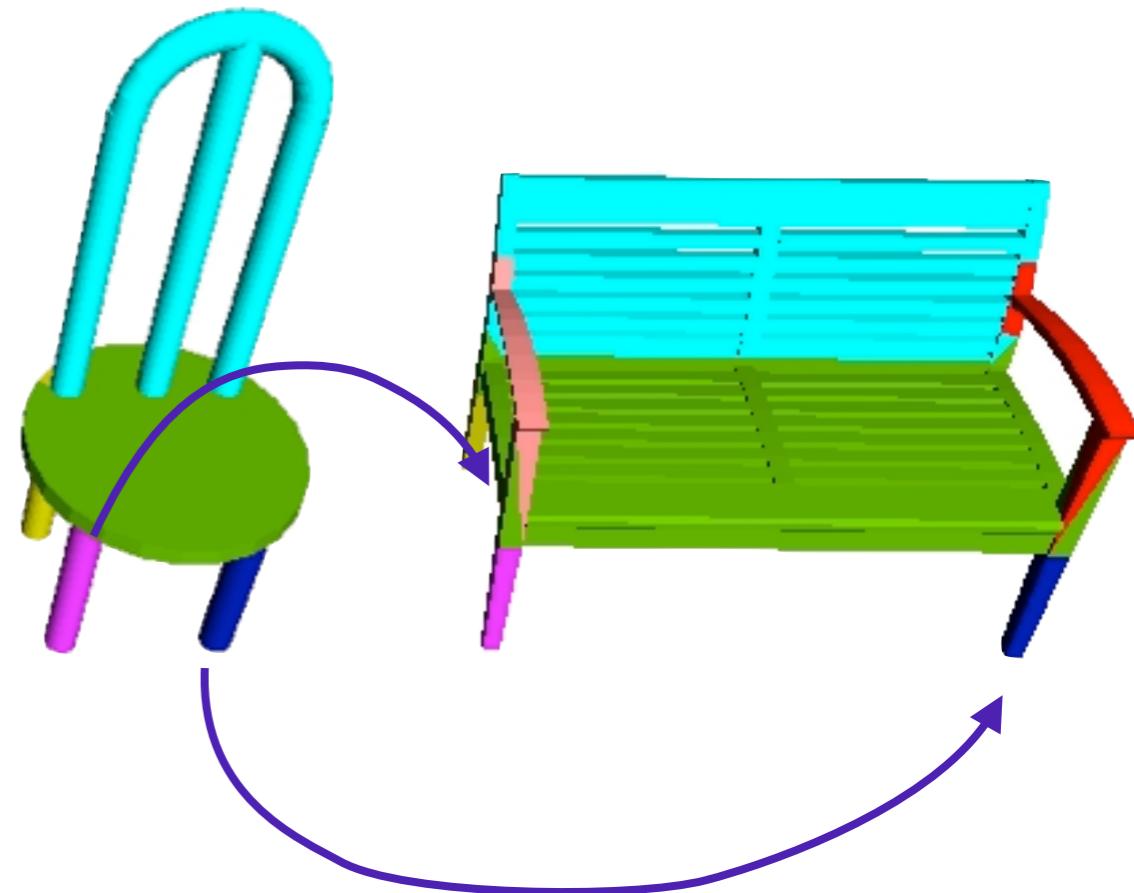
- The problems are inter-related:
 - Correspondences
 - Consistent Segmentations
 - Shape Variations



Why Joint Analysis?

- The problems are inter-related:
 - Correspondences
 - Consistent Segmentations
 - Shape Variations

Smaller search space

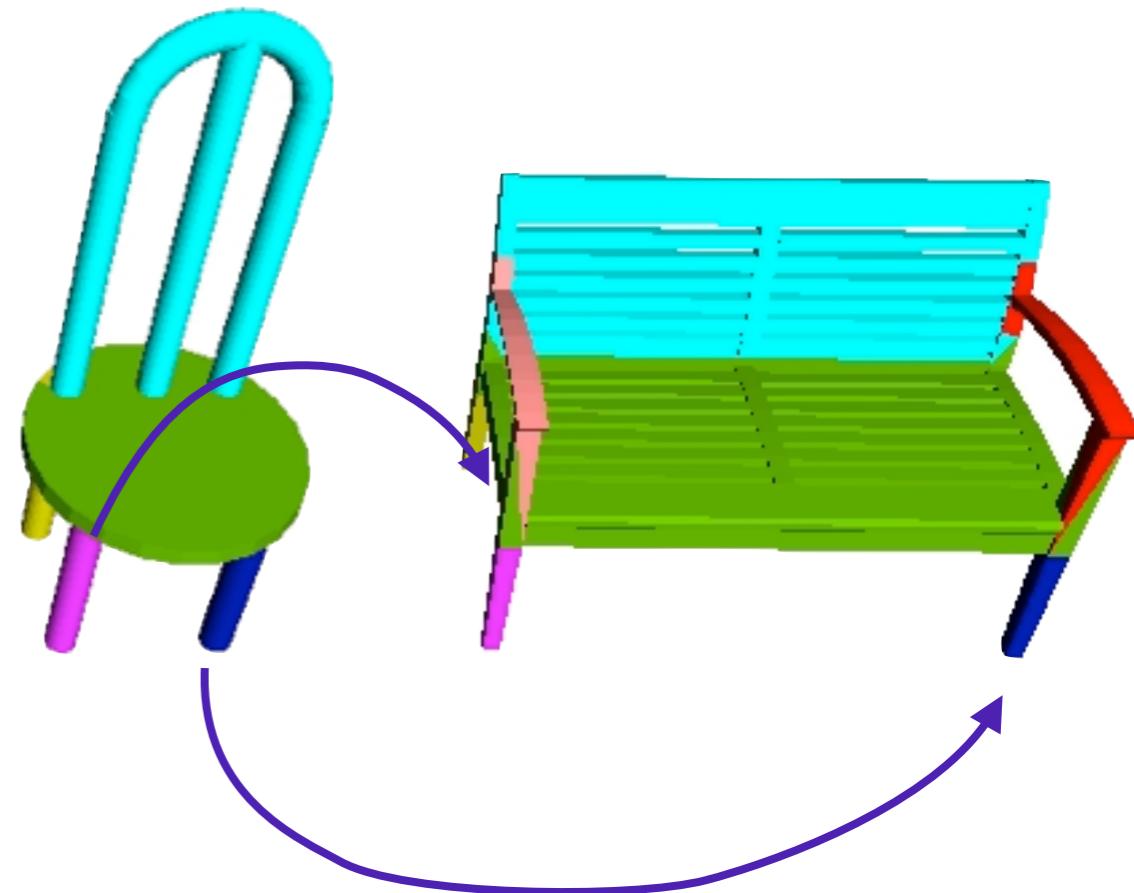


Why Joint Analysis?

- The problems are inter-related:
 - Correspondences
 - Consistent Segmentations
 - Shape Variations

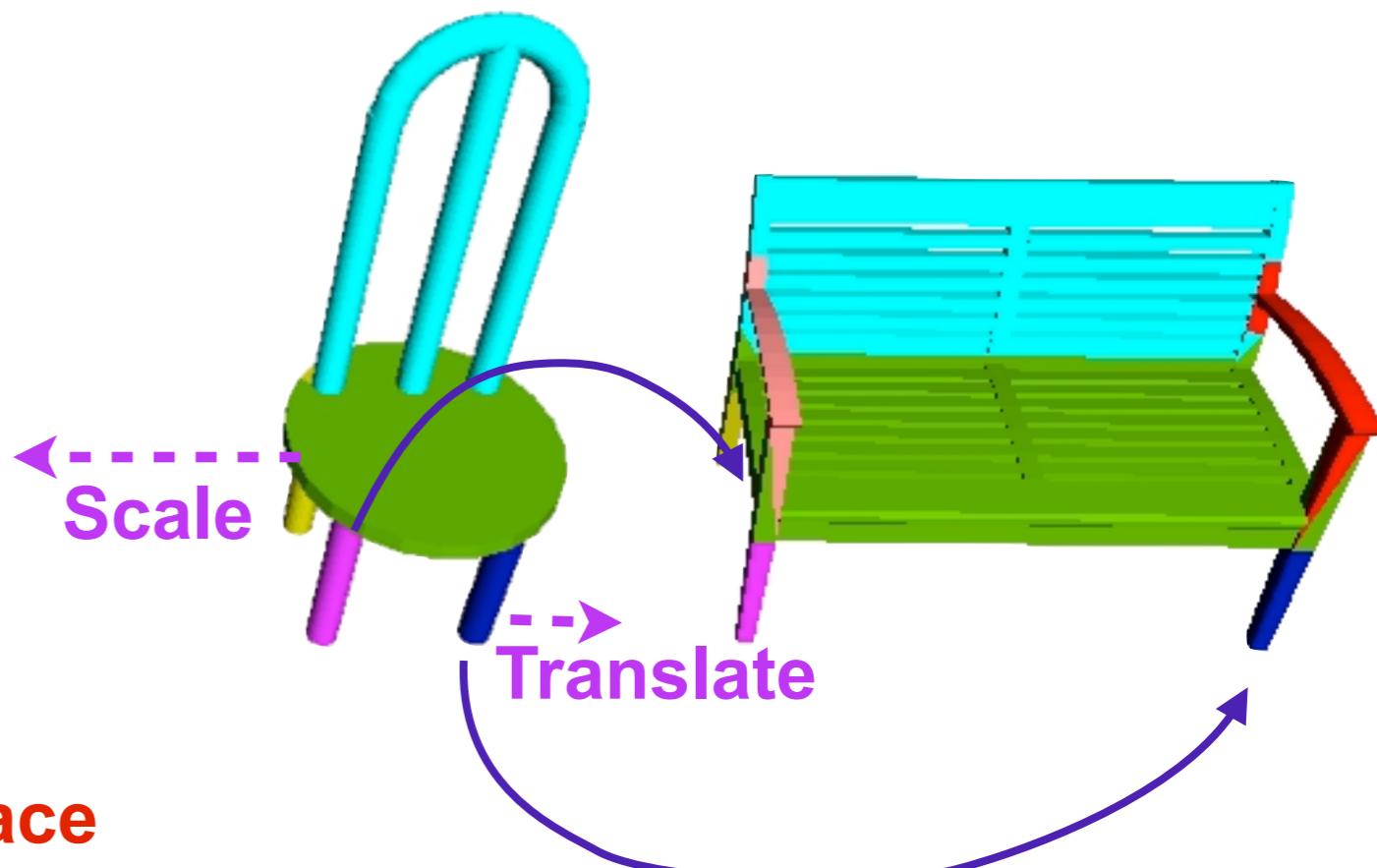


Smaller search space



Why Joint Analysis?

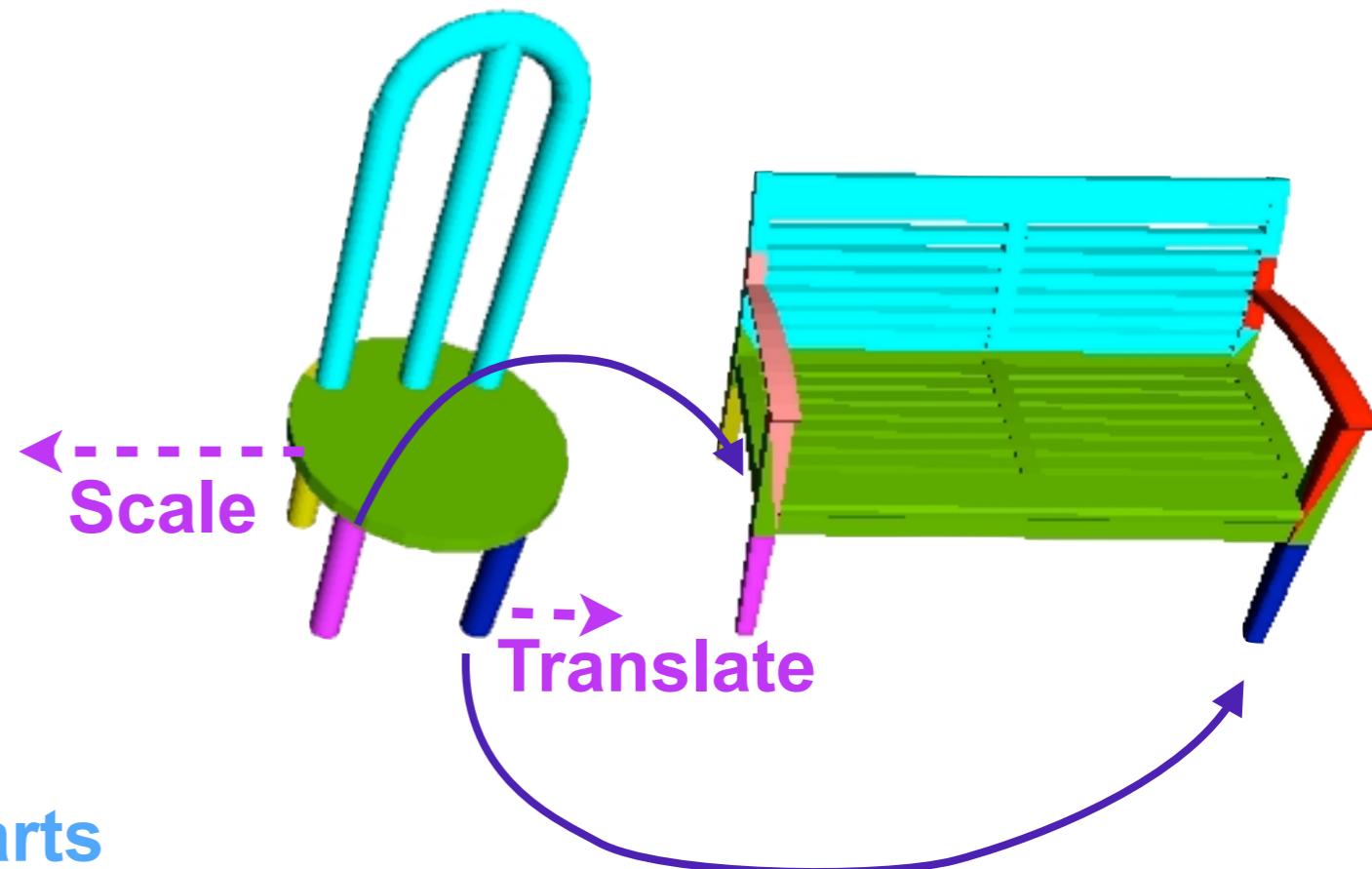
- The problems are inter-related:
 - Correspondences
 - Consistent Segmentations
 - Shape Variations



**Smaller search space
More accurate search space**

Why Joint Analysis?

- The problems are inter-related:
 - Correspondences
 - Consistent Segmentations
 - Shape Variations

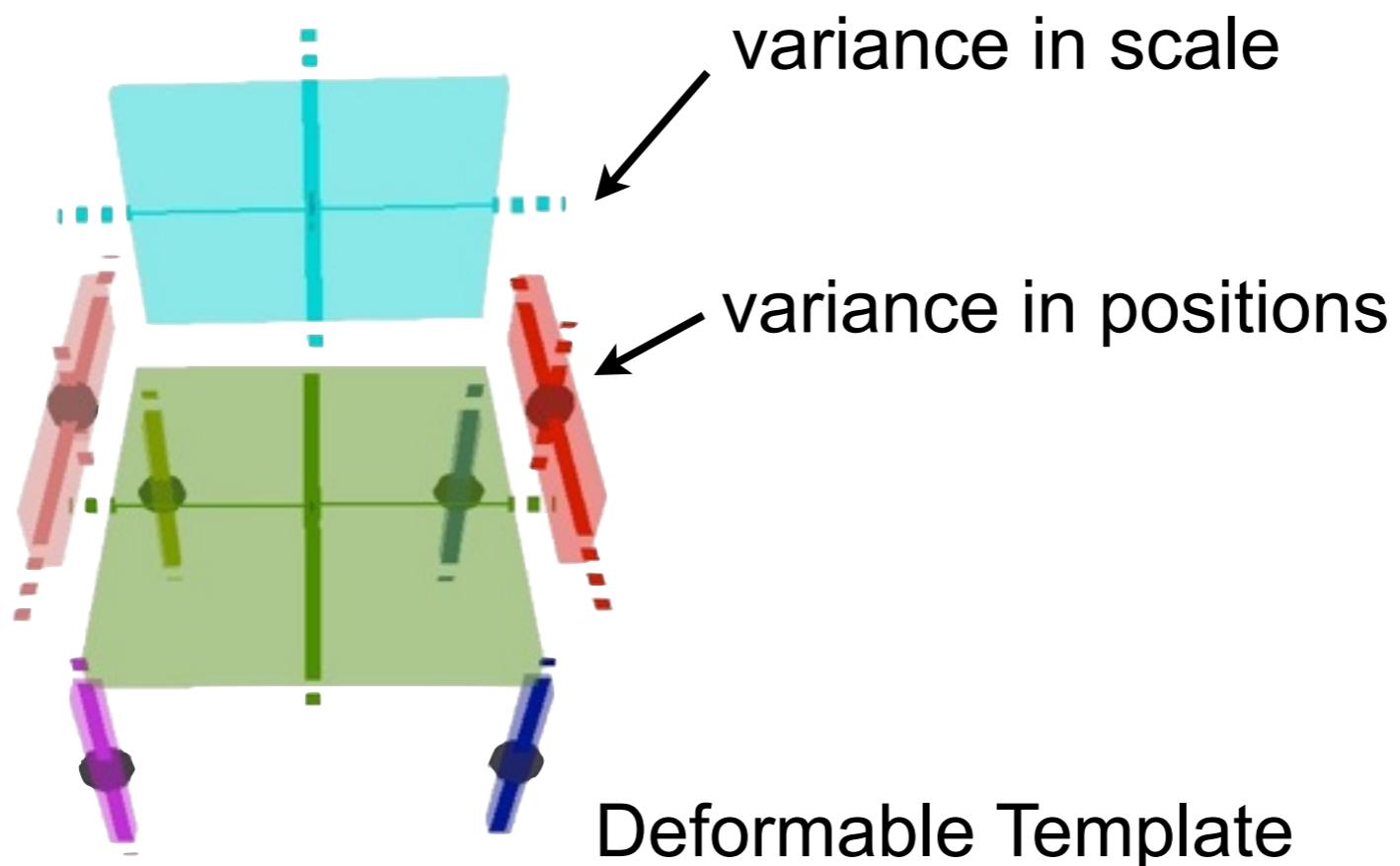


Smaller search space

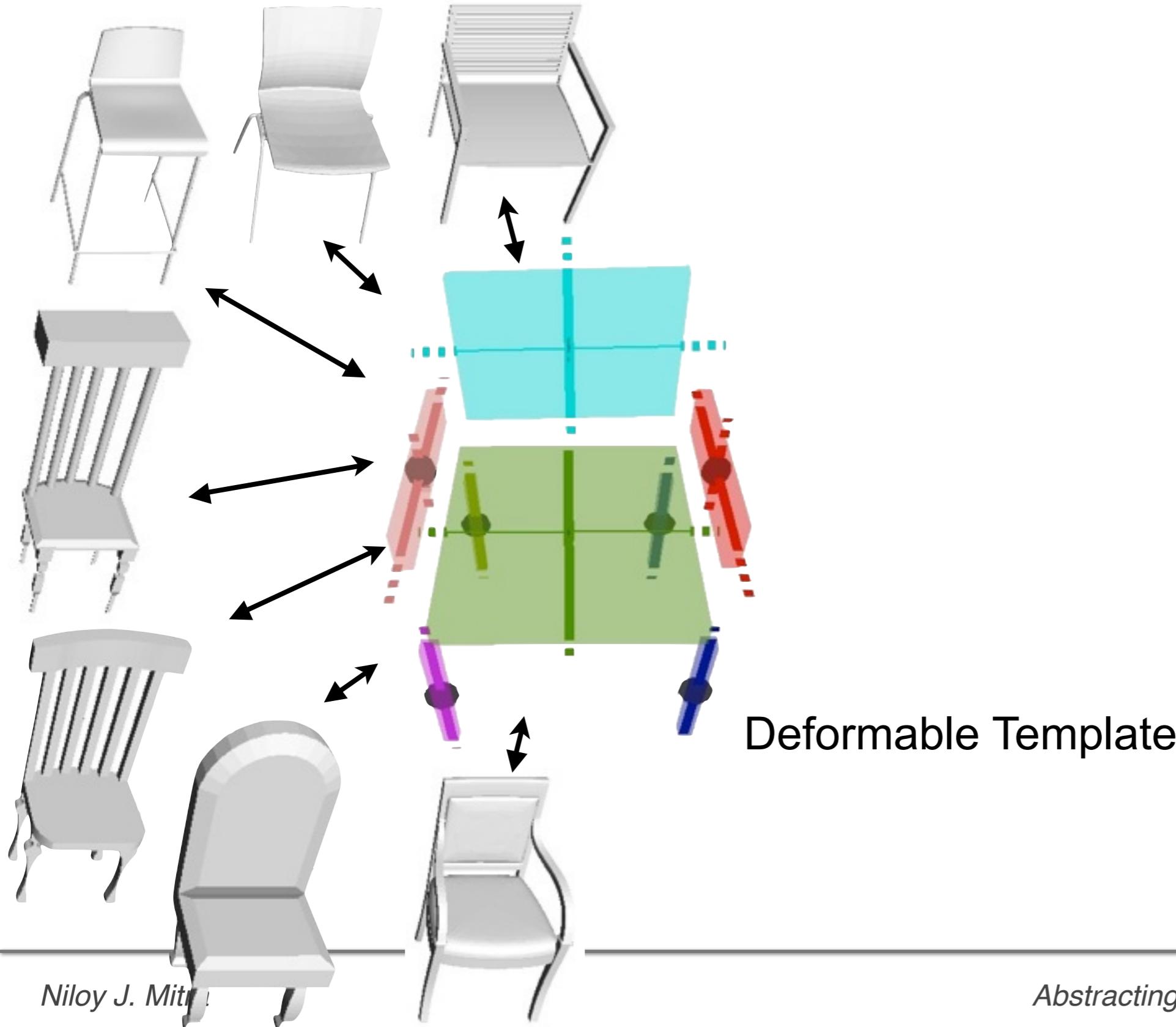
More accurate search space

Man-made object vary w.r.t. parts

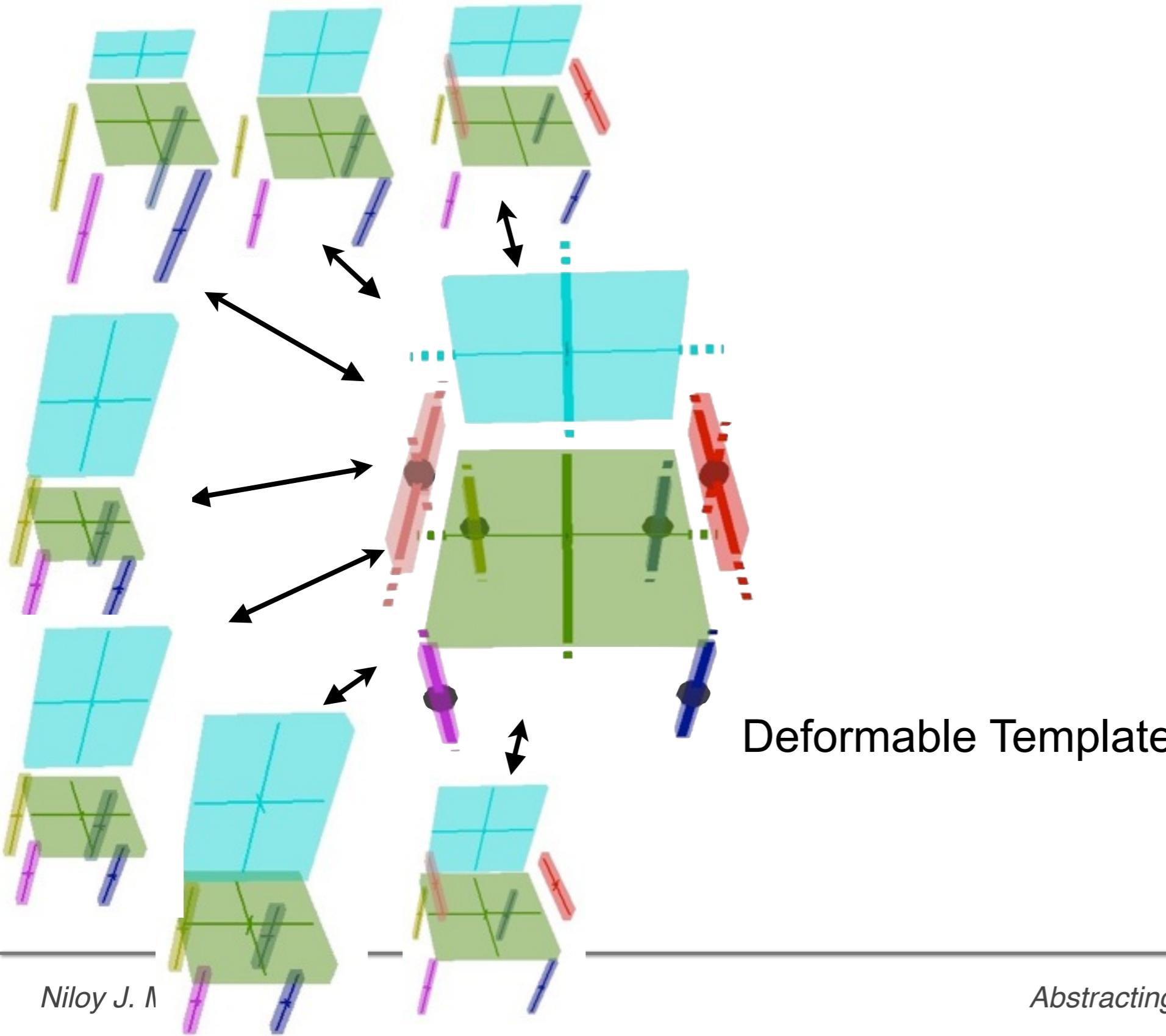
Our Approach



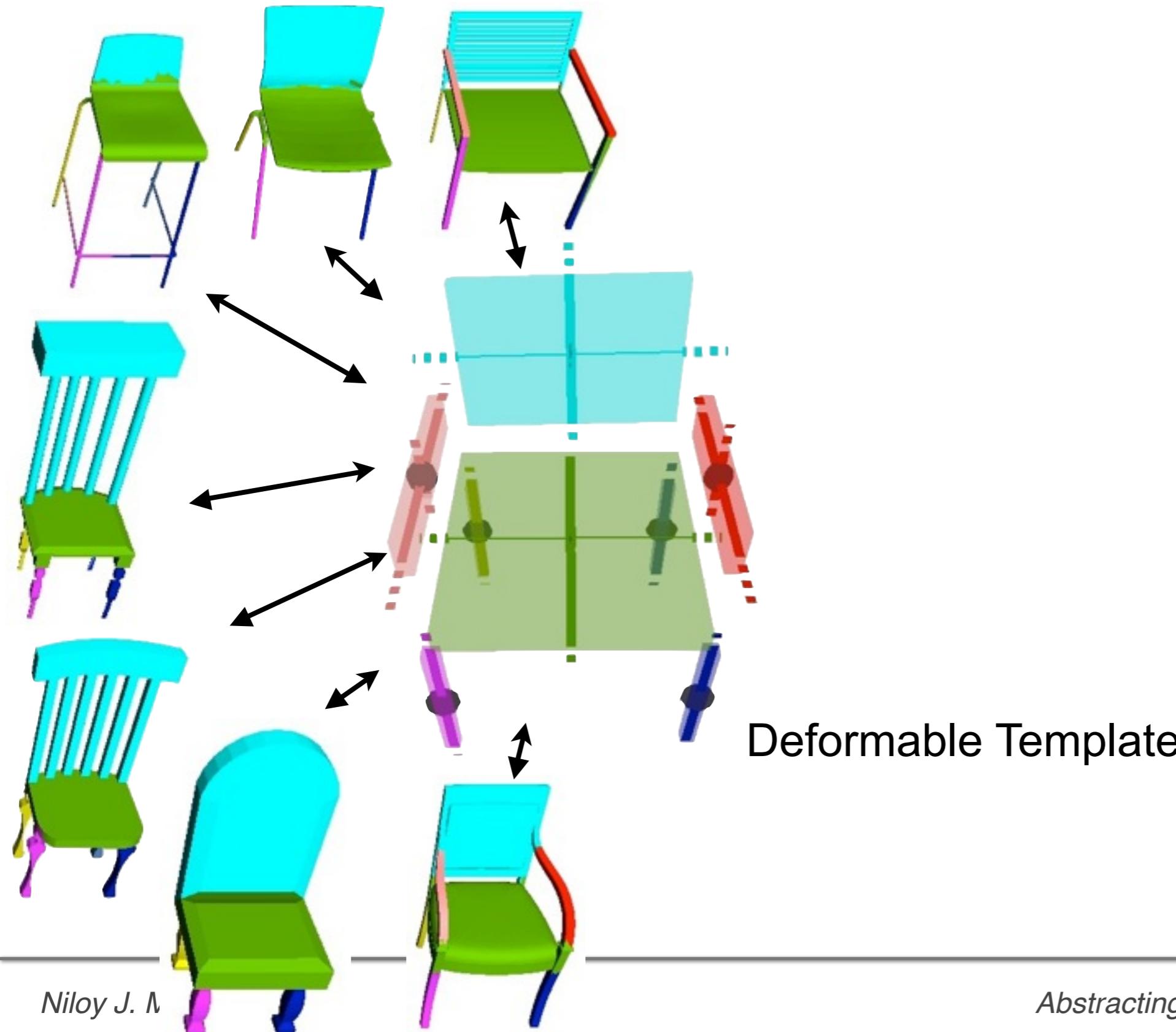
Our Approach



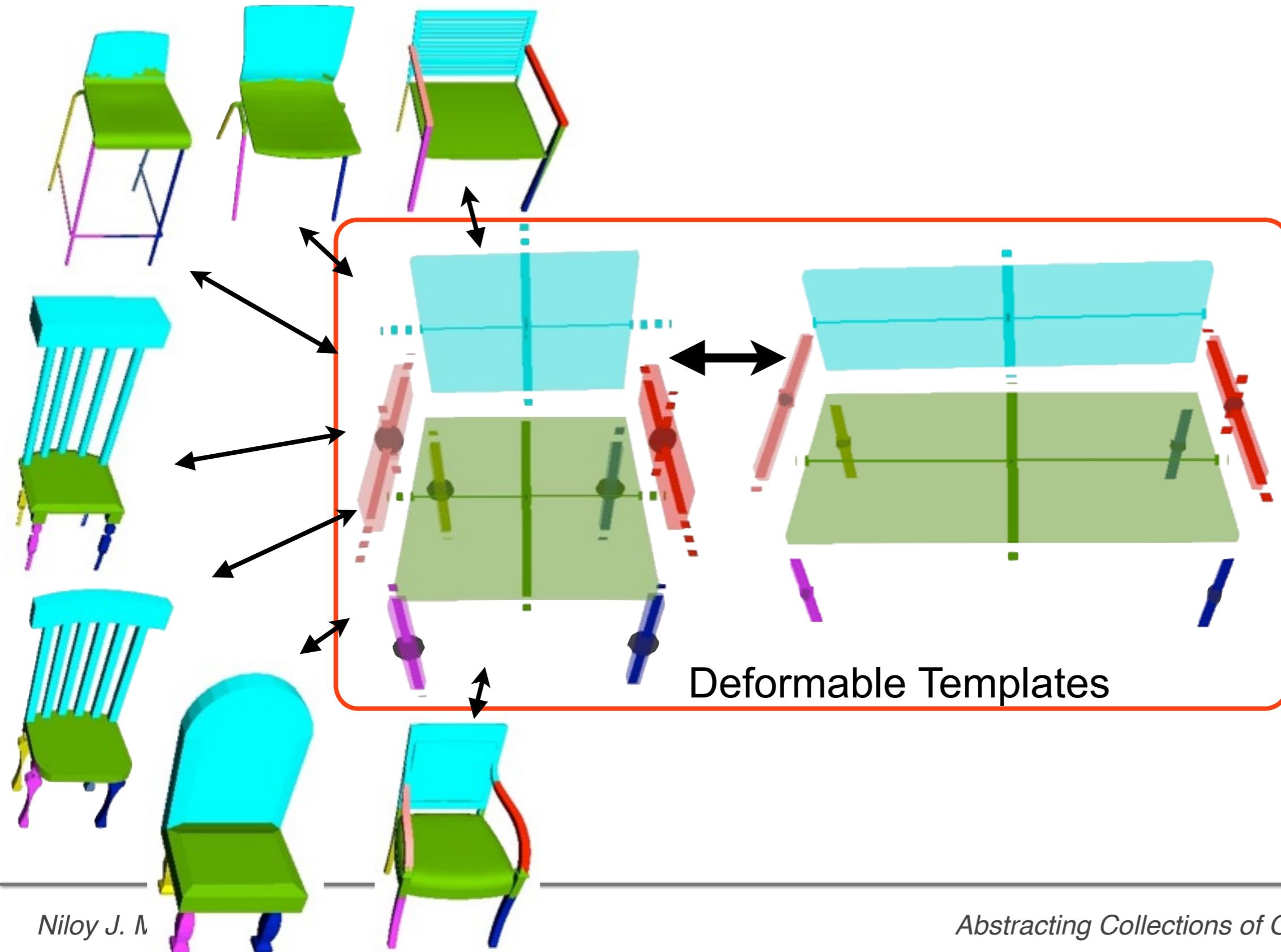
Our Approach



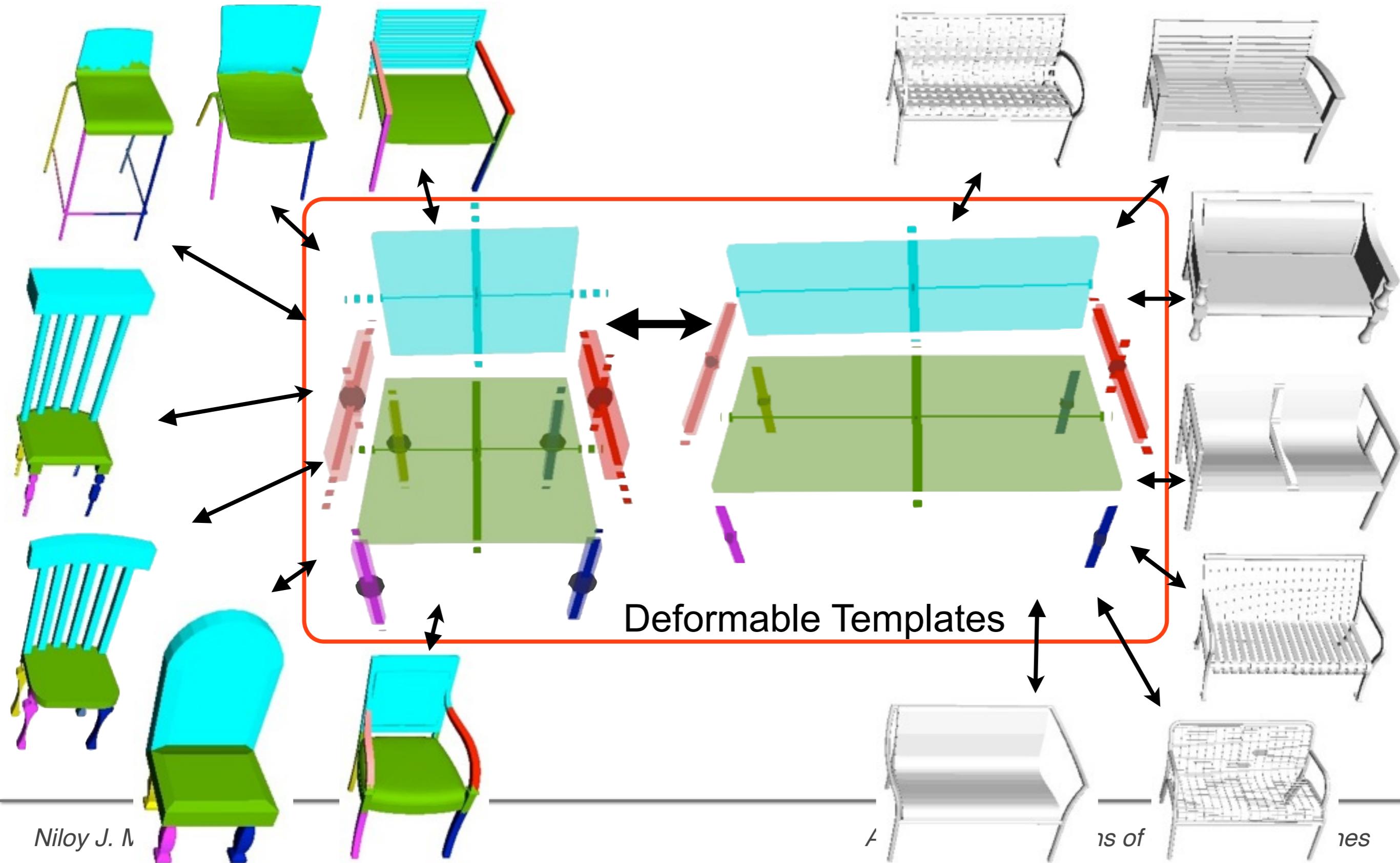
Our Approach



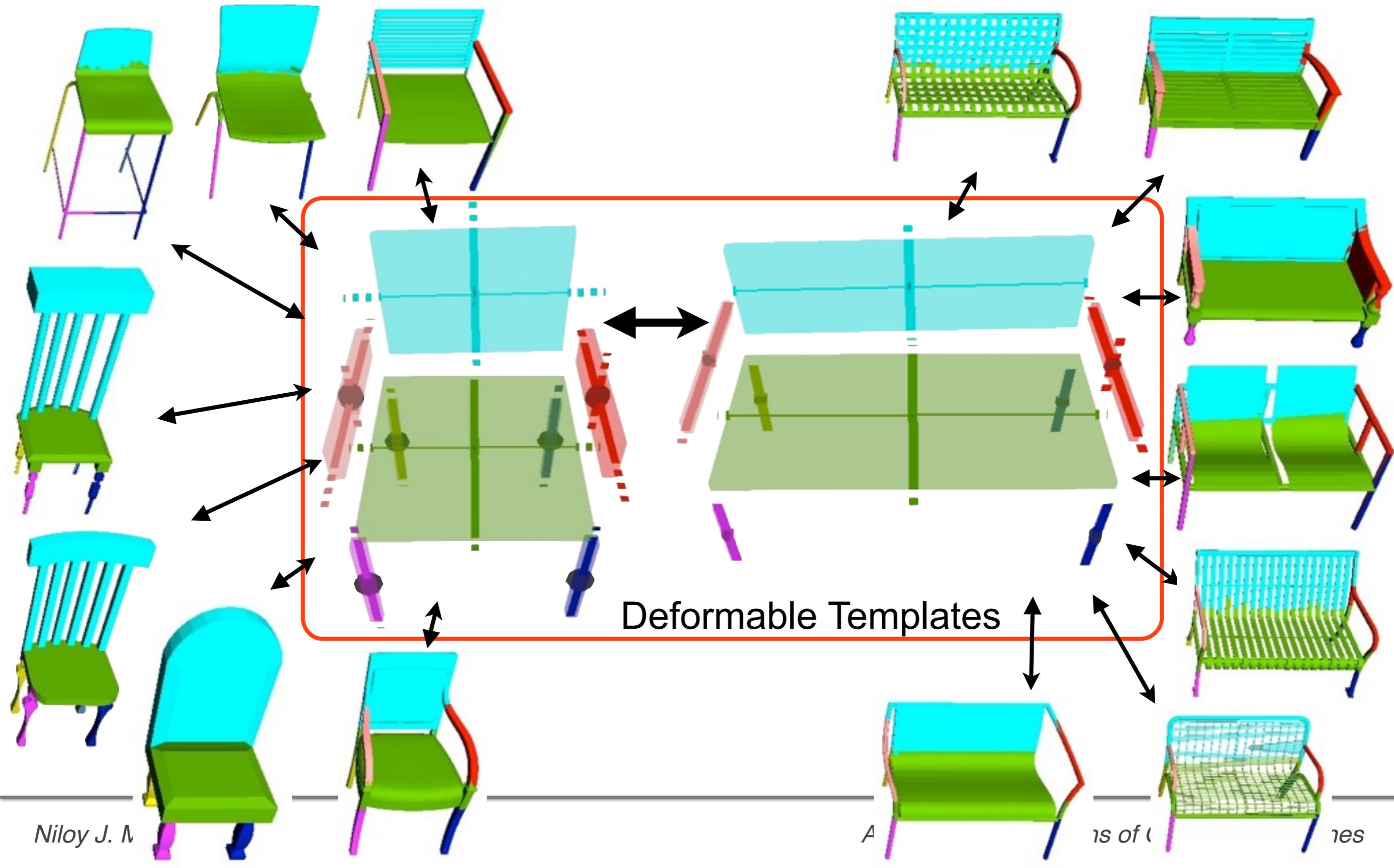
Our Approach



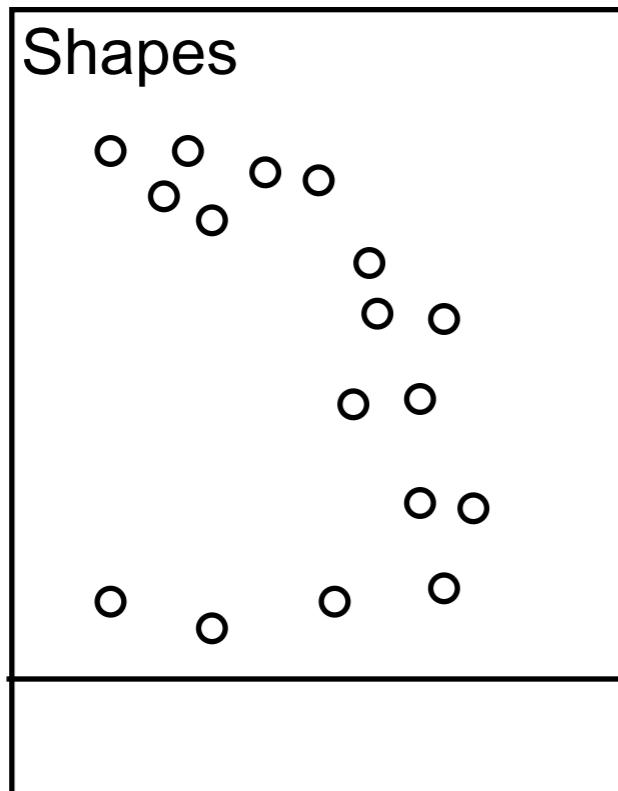
Our Approach



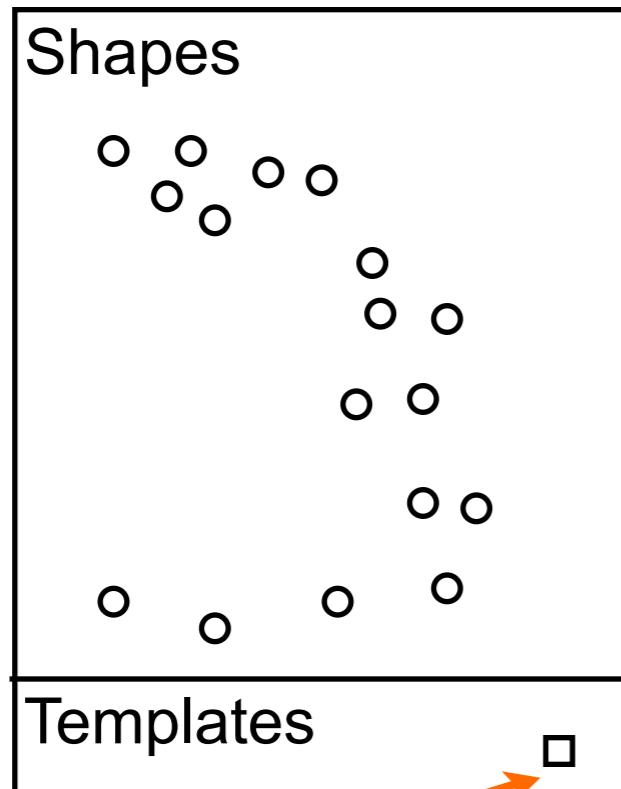
Our Approach



Algorithm Overview

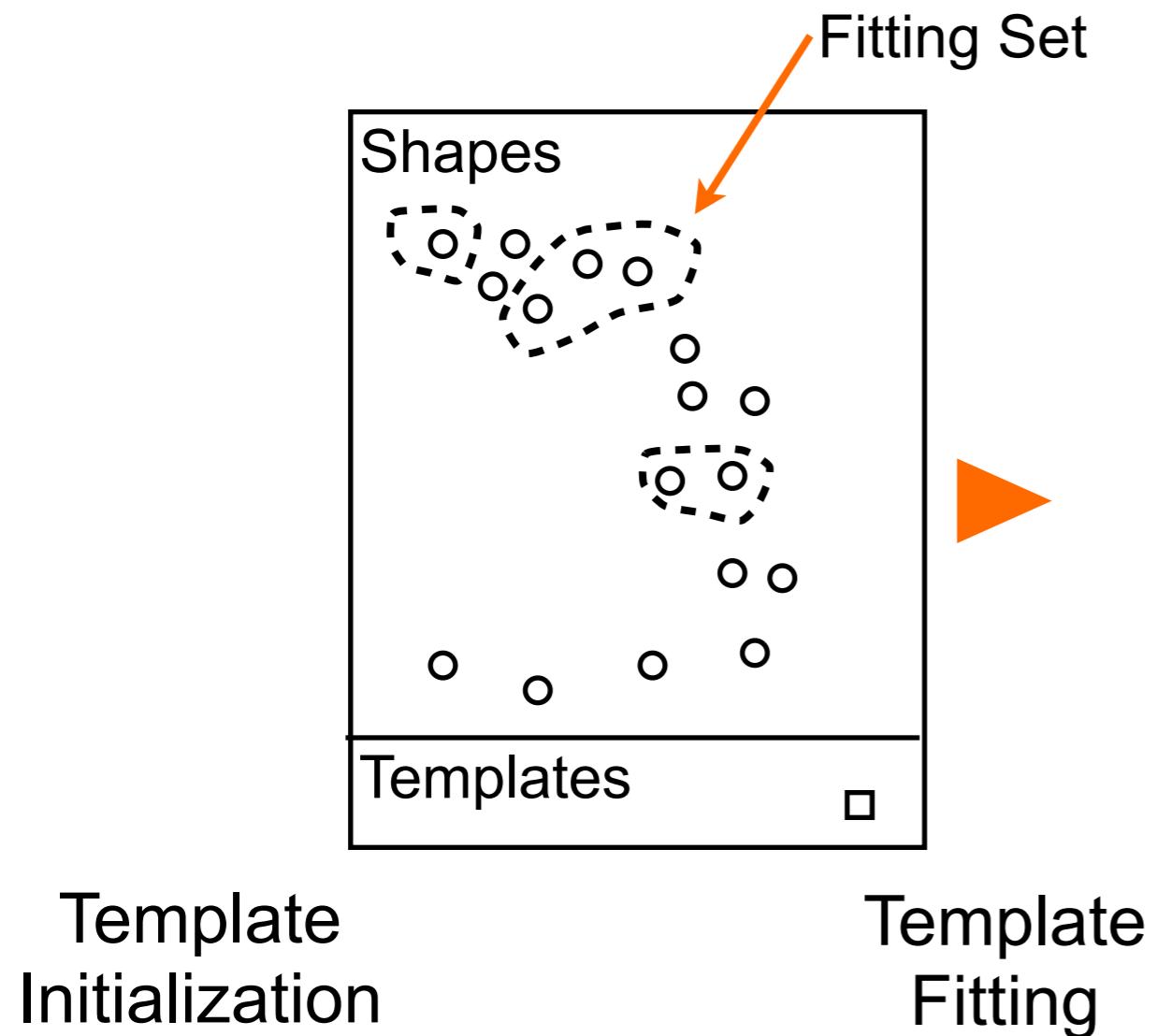


Algorithm Overview

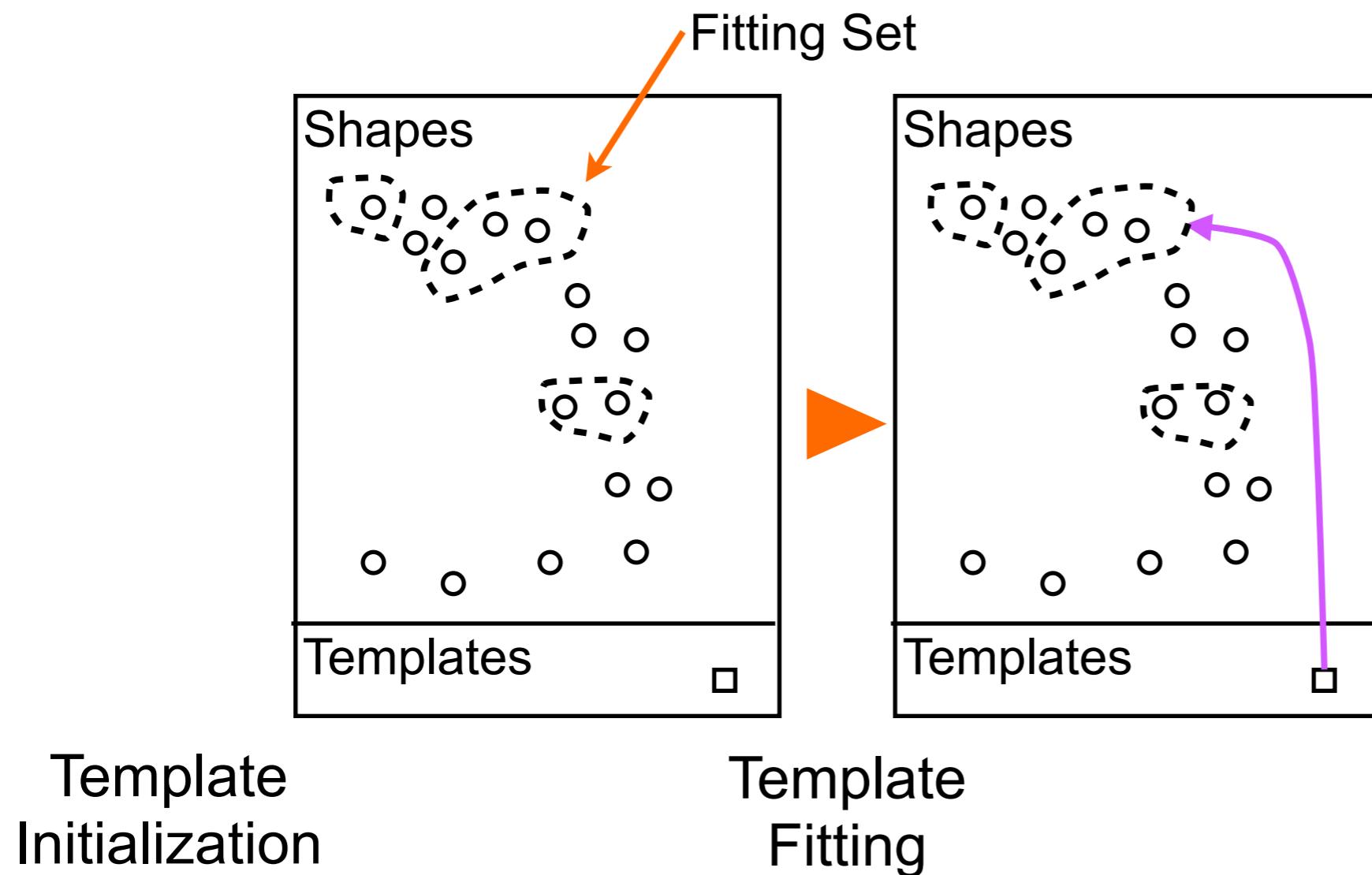


Template
Initialization

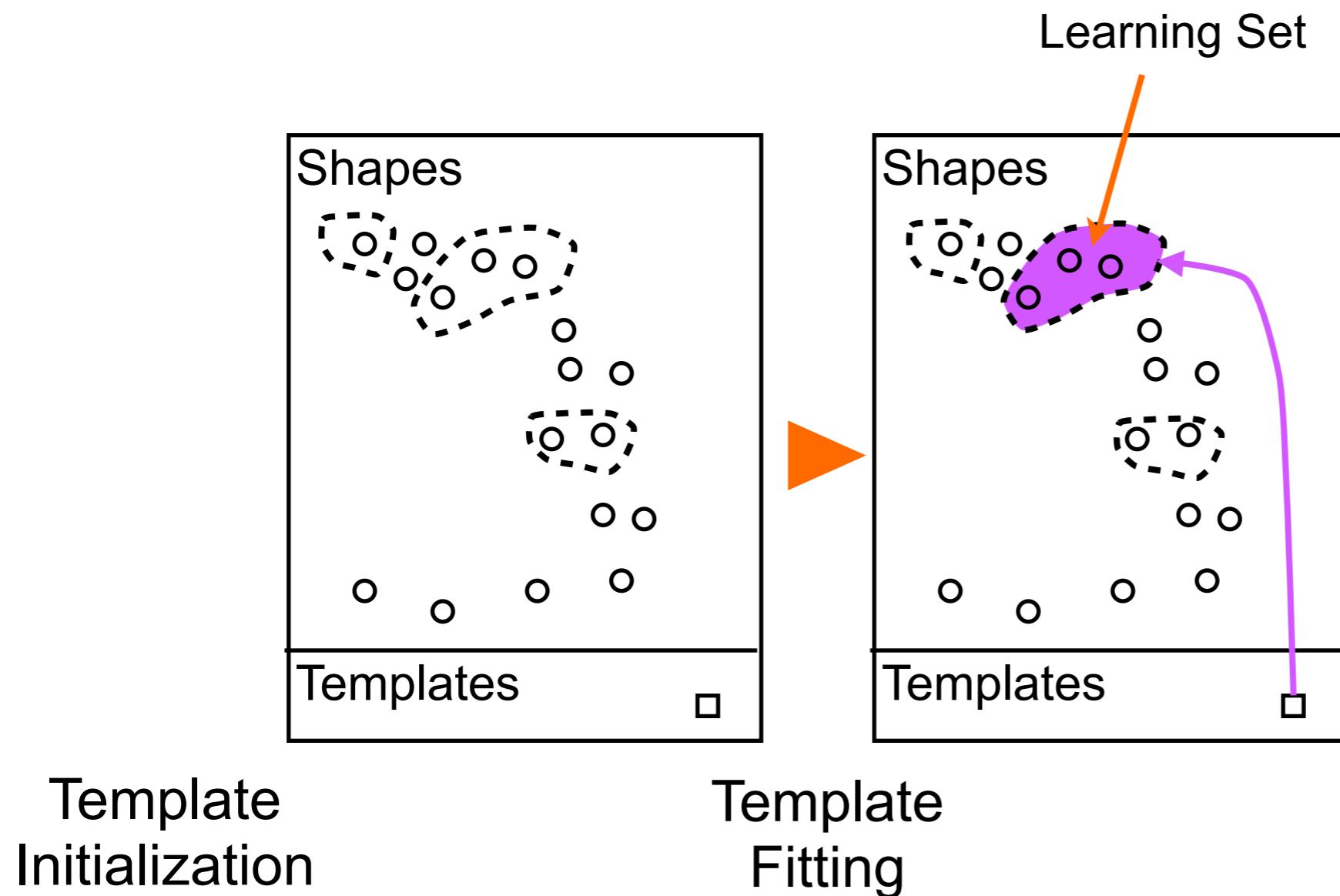
Algorithm Overview



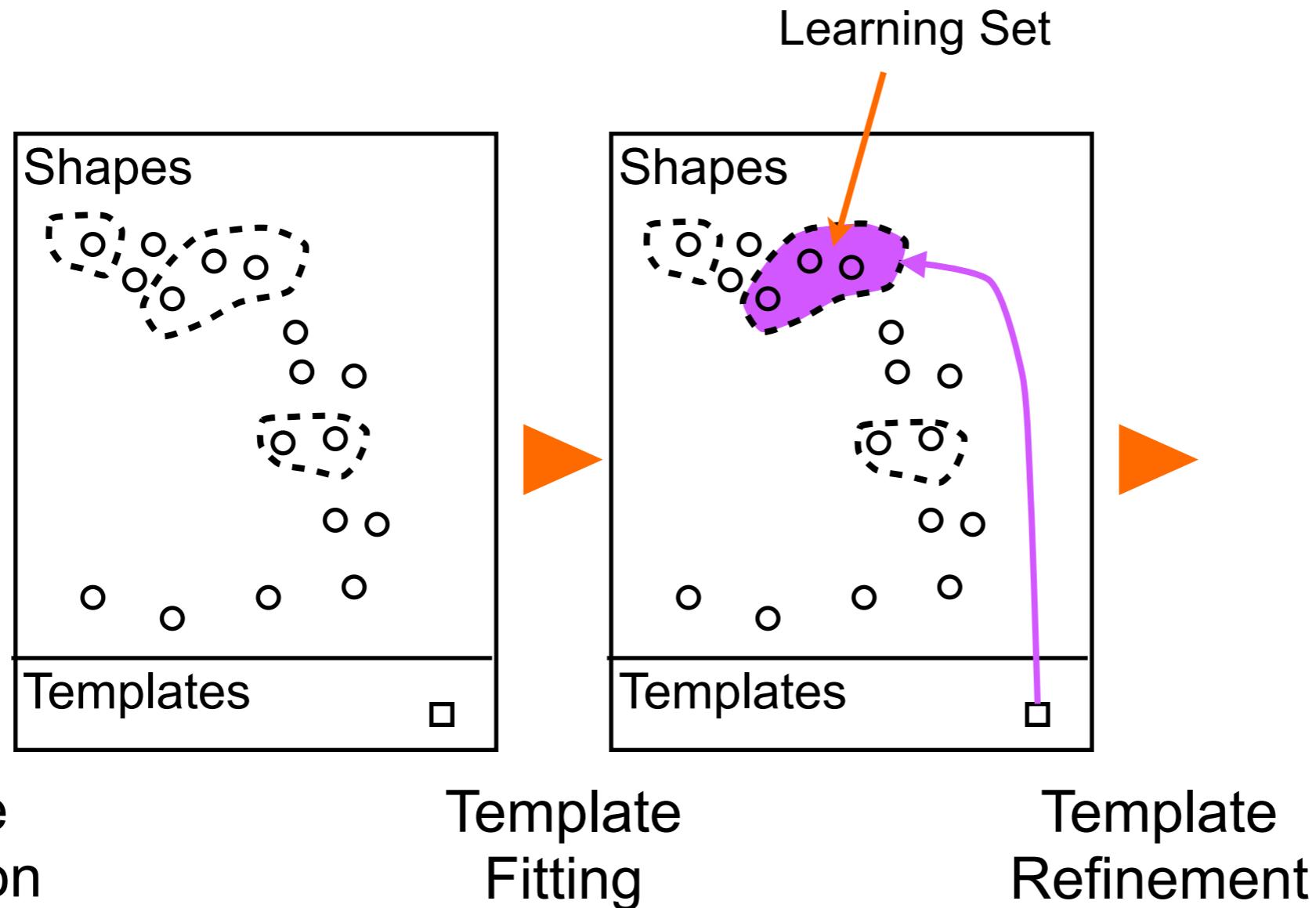
Algorithm Overview



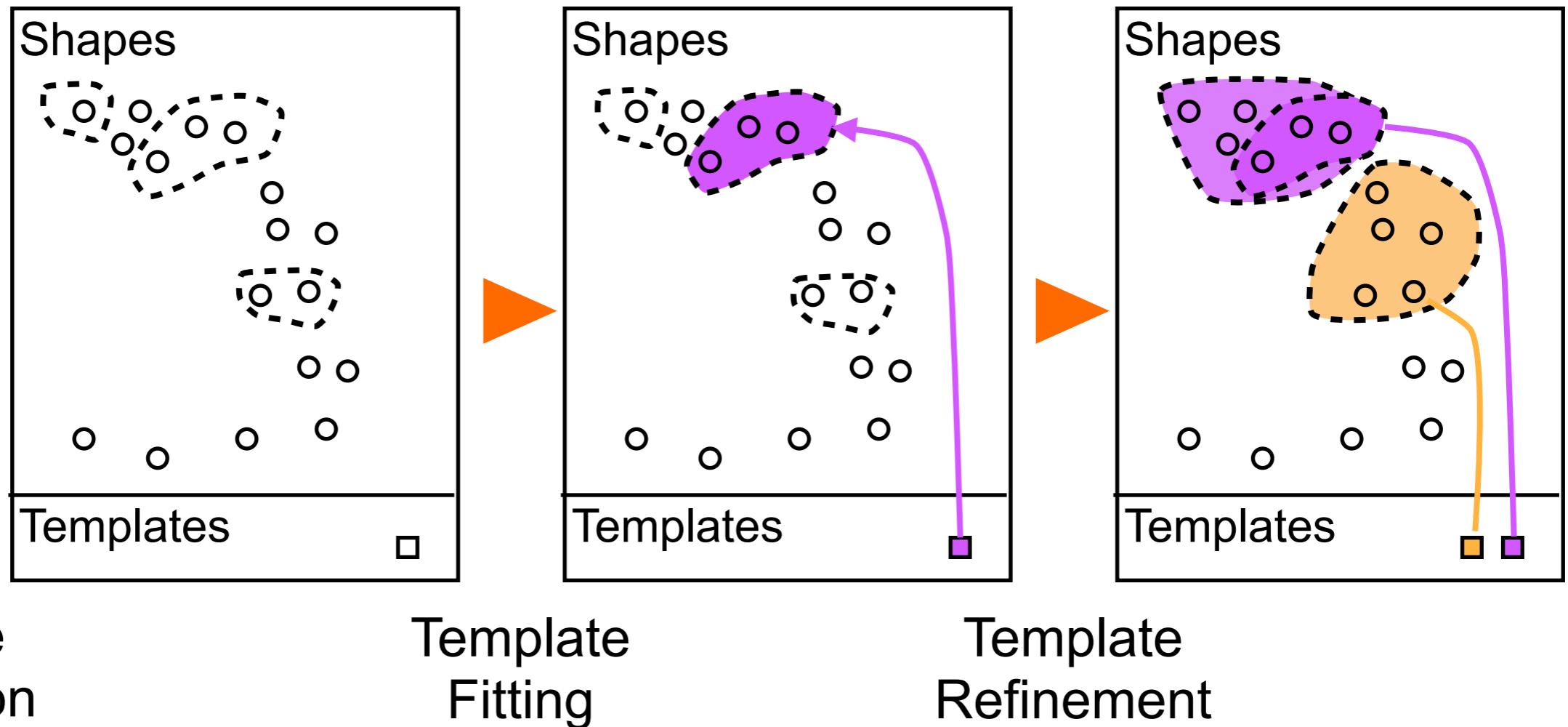
Algorithm Overview



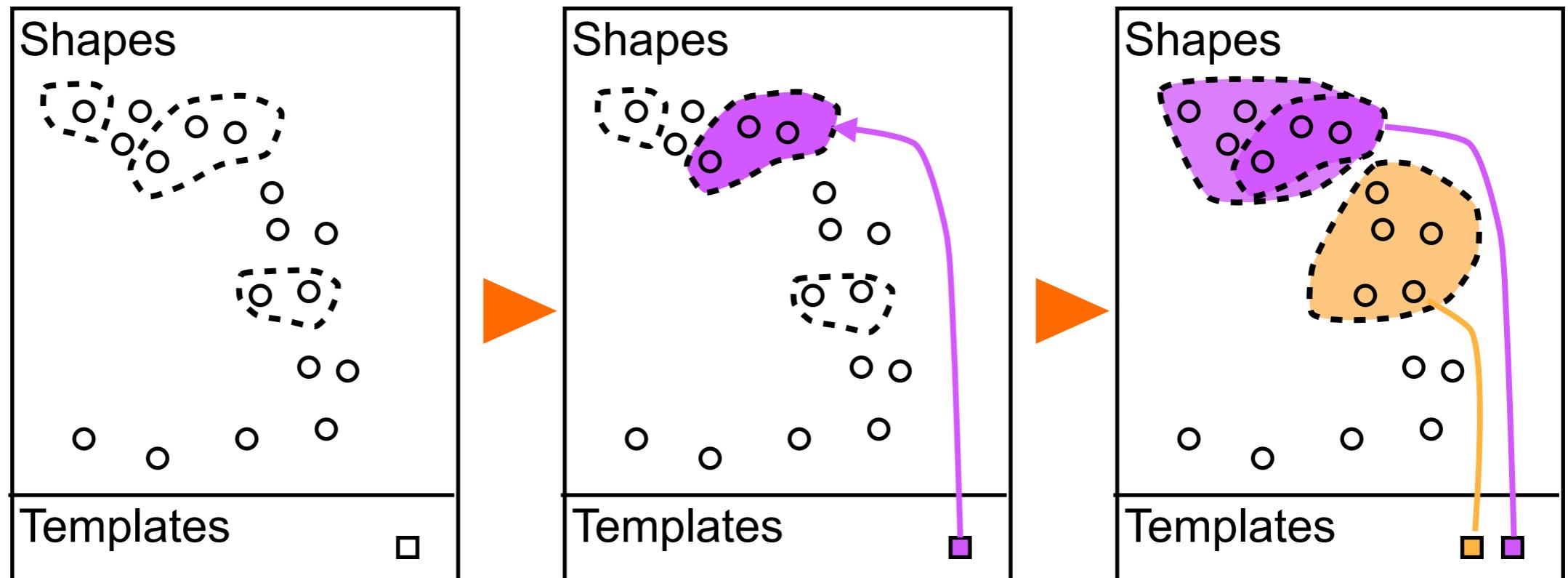
Algorithm Overview



Algorithm Overview



Algorithm Overview



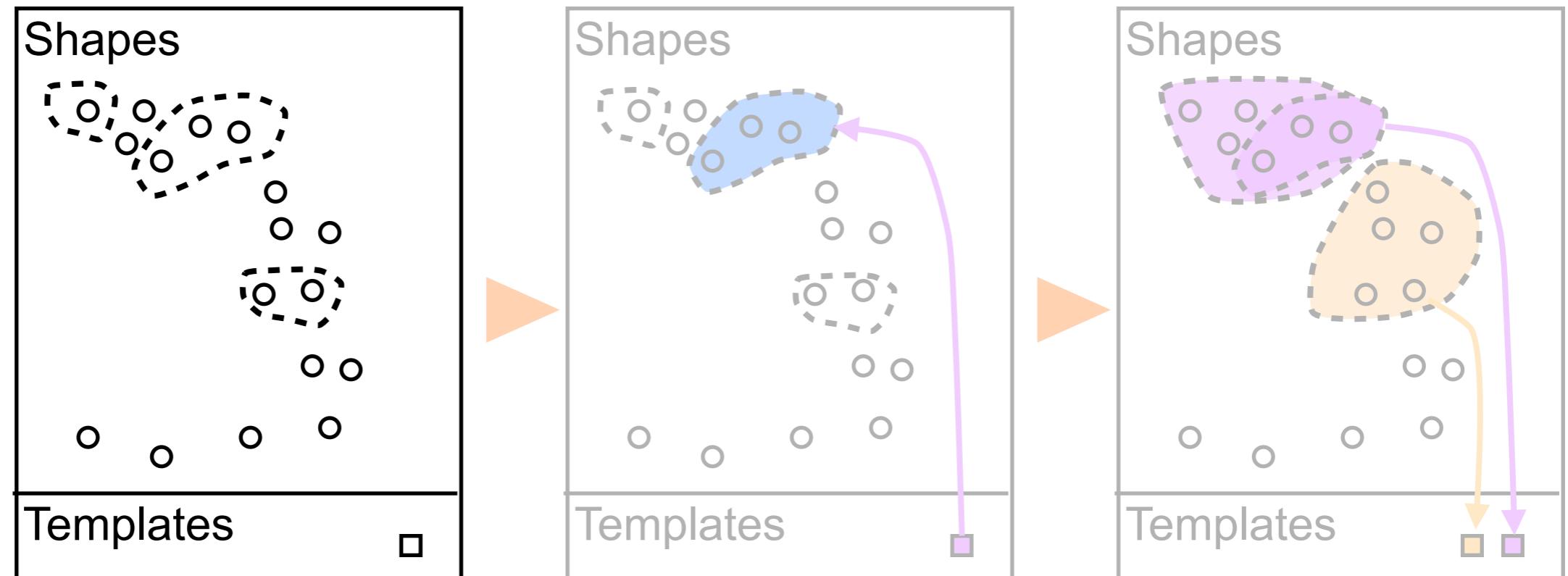
Template
Initialization

Template
Fitting

Template
Refinement

repeat until convergence

Algorithm Overview



Template
Initialization

Template
Fitting

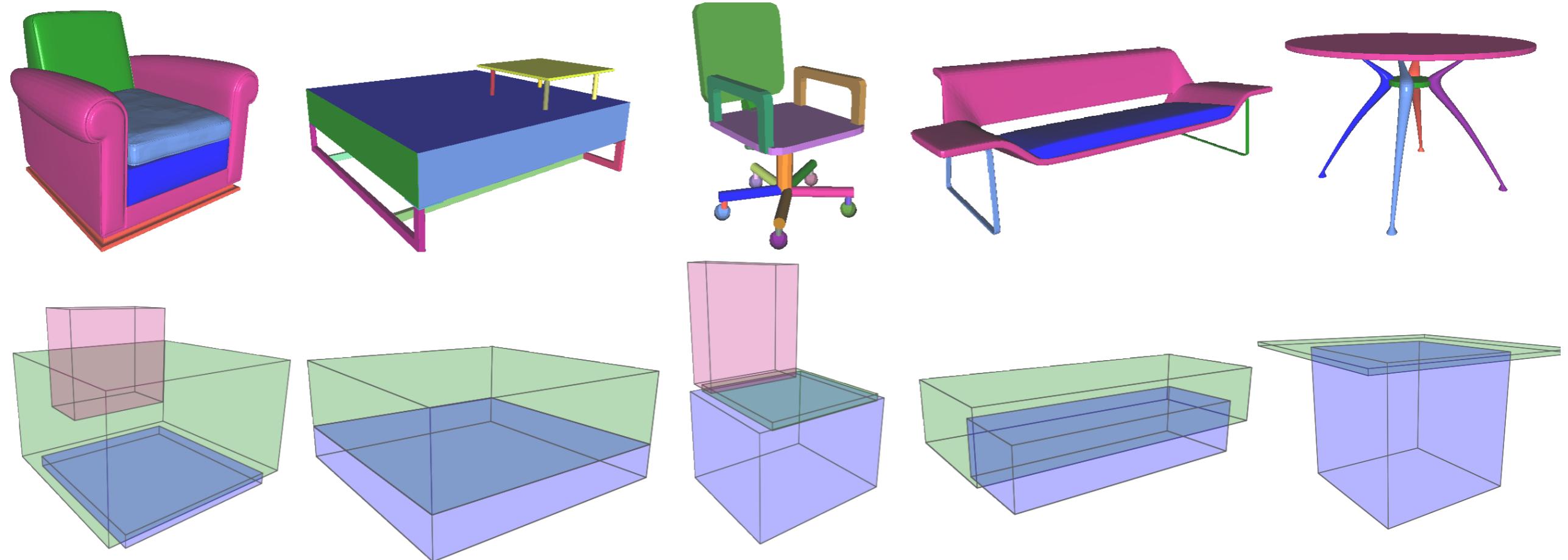
Template
Refinement

repeat until convergence

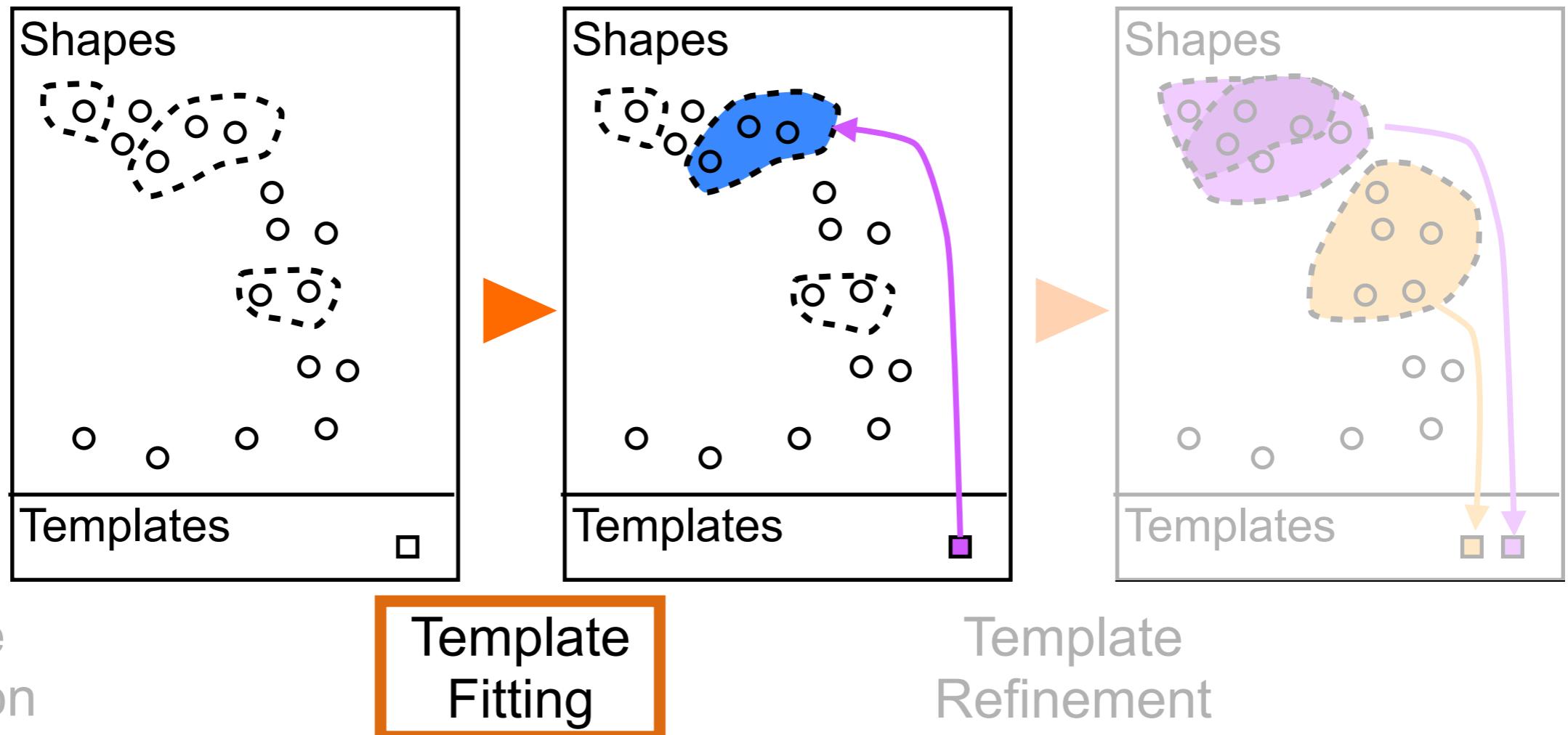
Template Initialization



Template Initialization



Algorithm Overview



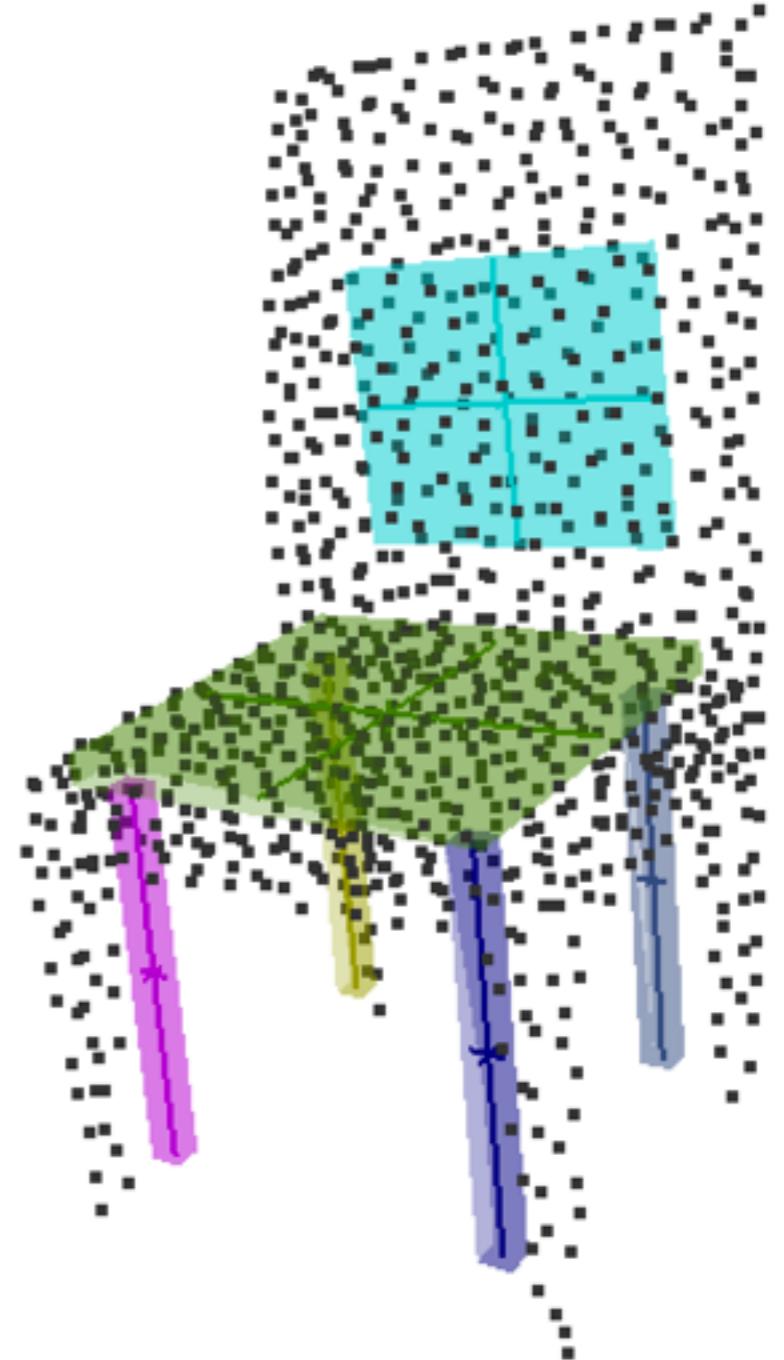
Template
Initialization

Template
Fitting

Template
Refinement

Fitting Parameters

- Rigid alignment
- Per part deformations
 - Existence
 - Centroid position
 - Anisotropic scale
- Labeling of points in the shape
- Shape \leftrightarrow template mapping



Fitting Energy



$$E = E_{\text{data}} + \gamma E_{\text{deform}} + \beta E_{\text{smooth}}$$

Optimization Details



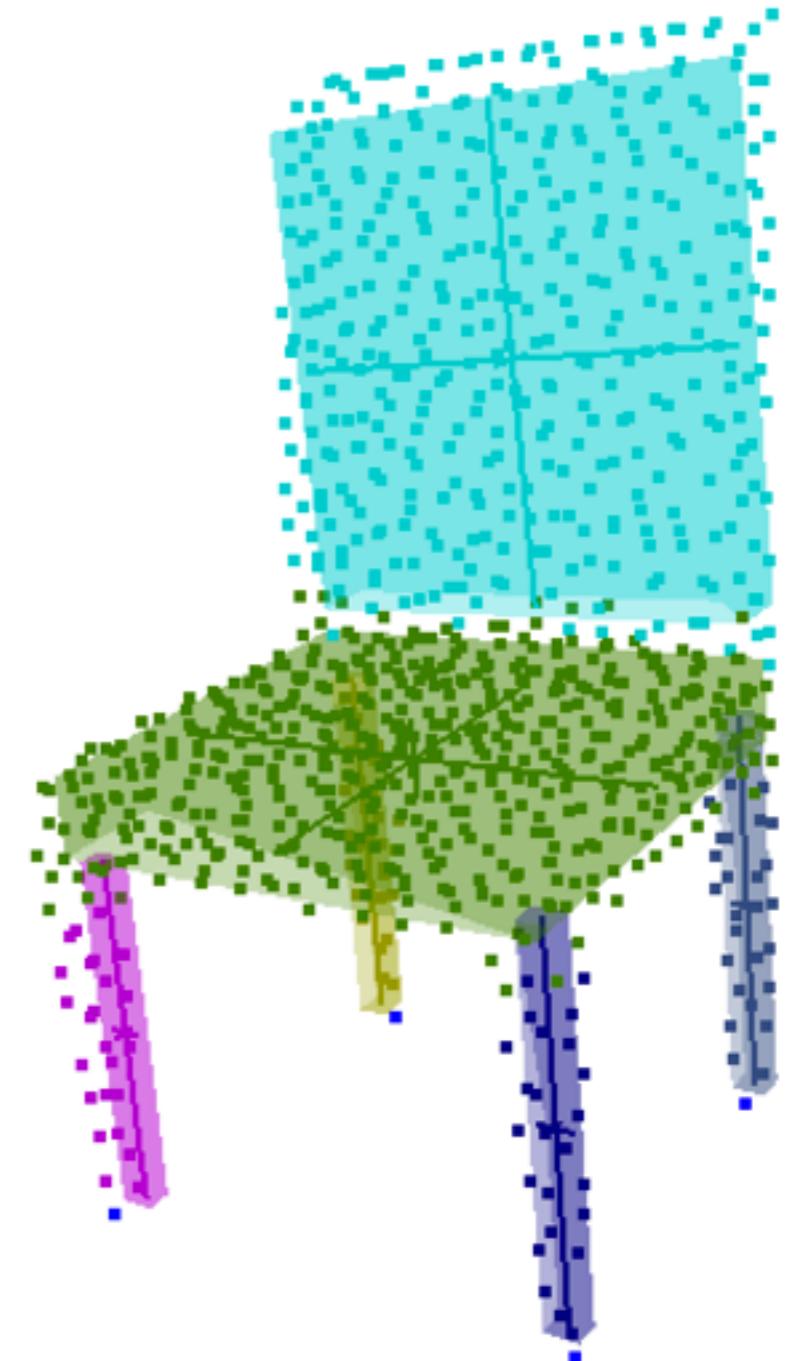
- Alternate steps until shape segmentation converges:
 - Segmentation
 - Correspondences
 - Deformation

$$E = \underline{E_{\text{data}}} + \gamma \underline{E_{\text{deform}}} + \beta E_{\text{smooth}}$$

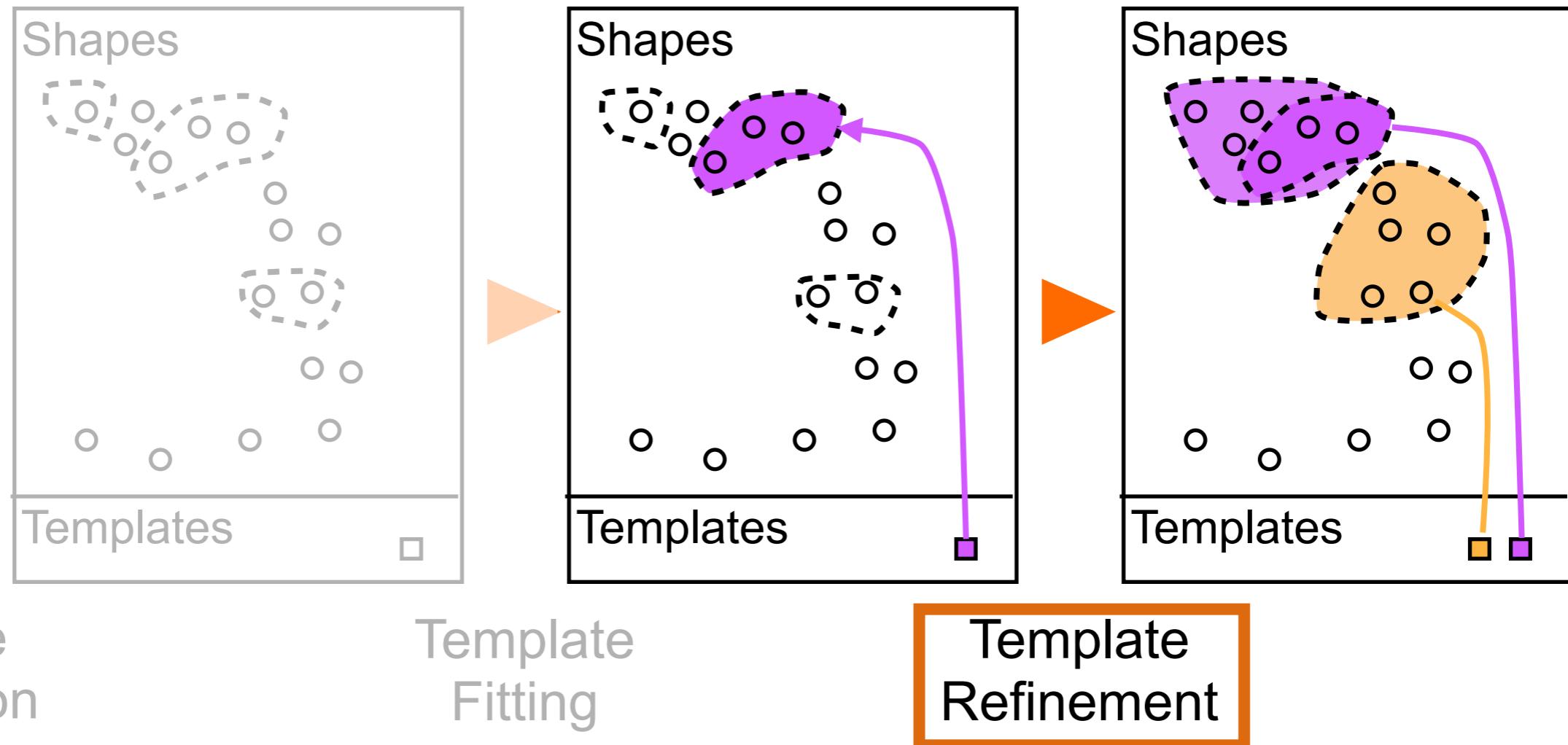
Method: Solve for critical points.

position: $\frac{\partial(E_{\text{data}} + E_{\text{deform}})}{\partial b_p} = 0$

scale: $\frac{\partial(E_{\text{data}} + E_{\text{deform}})}{\partial b_s} = 0$



Algorithm Overview

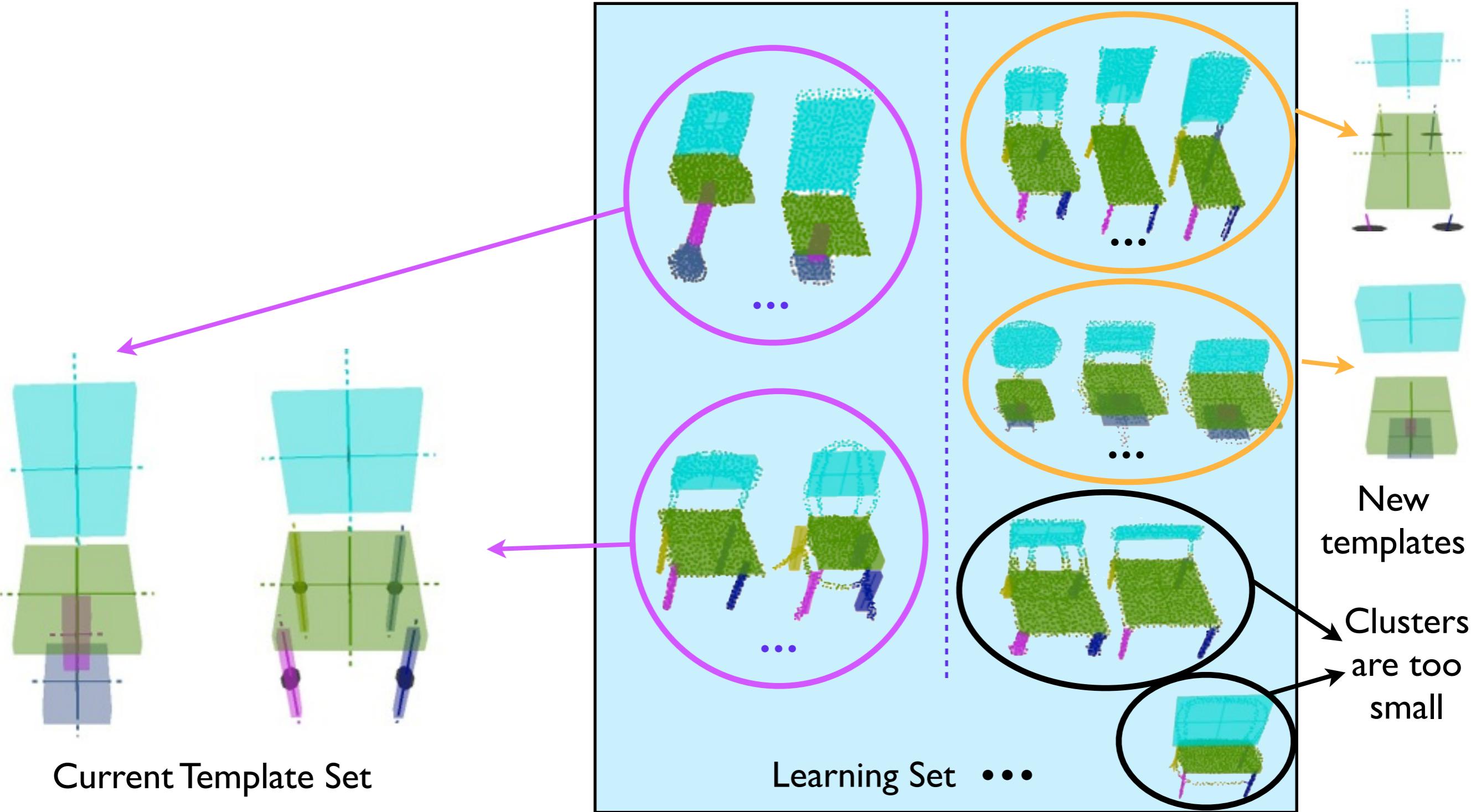


Template
Initialization

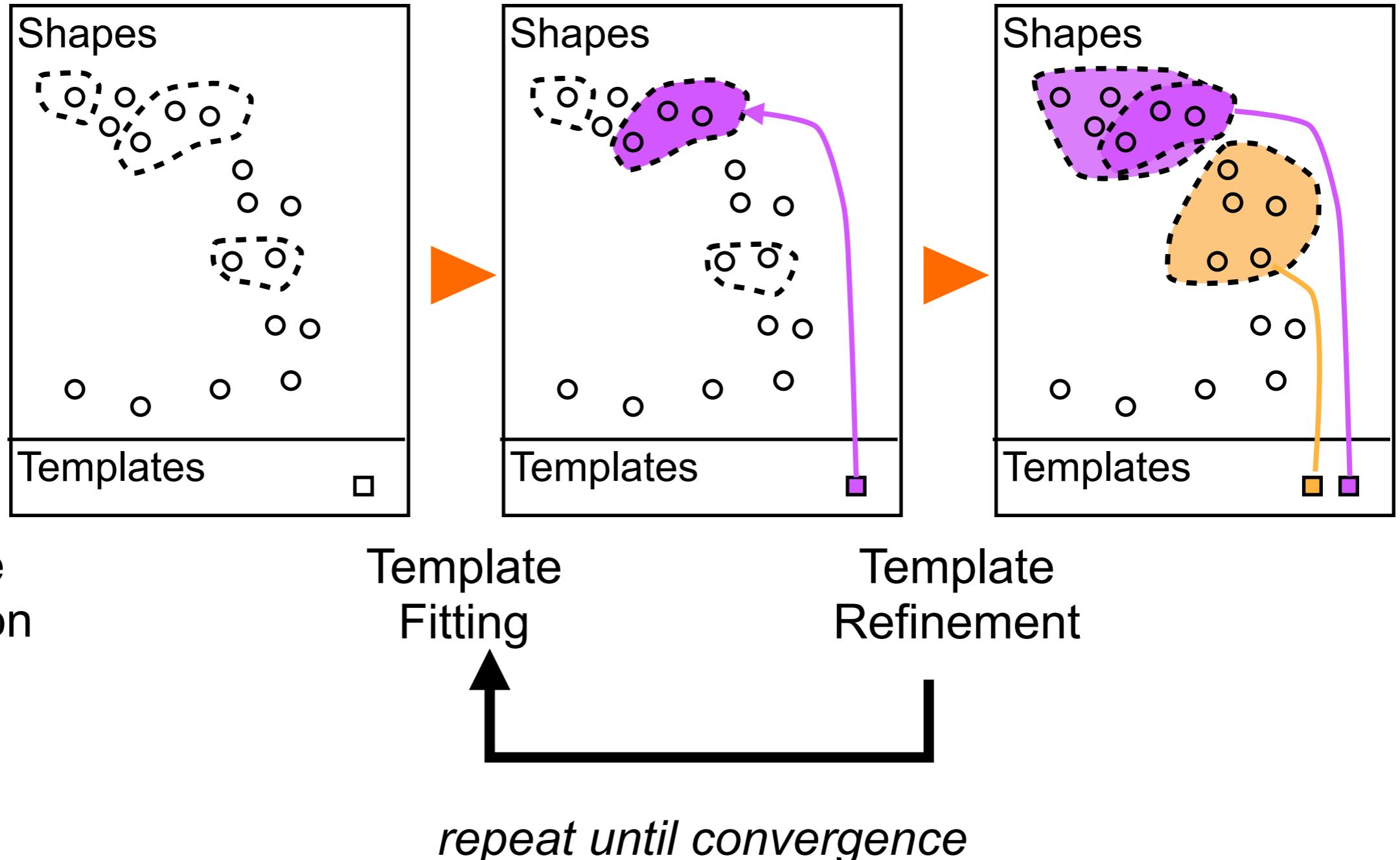
Template
Fitting

Template
Refinement

Template Refinement



Algorithm Overview

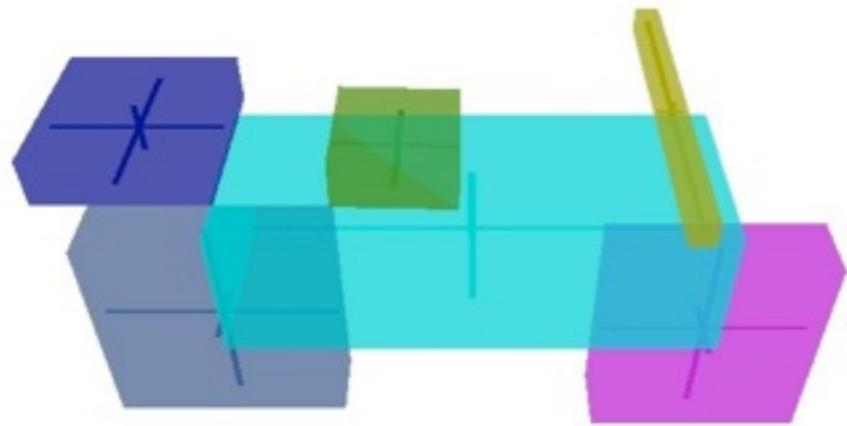


Results

Bikes Dataset



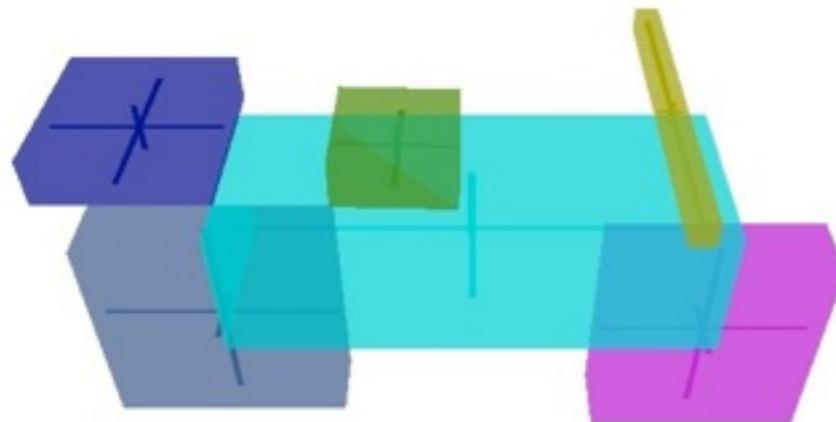
Initial Template:



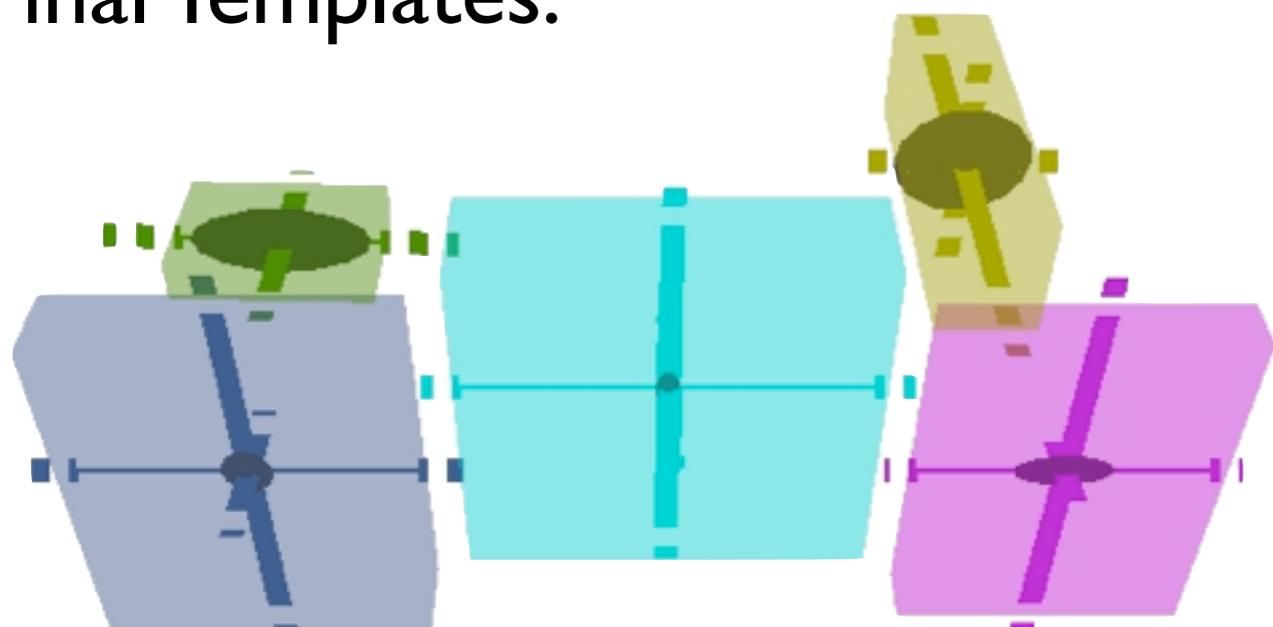
Bikes Dataset



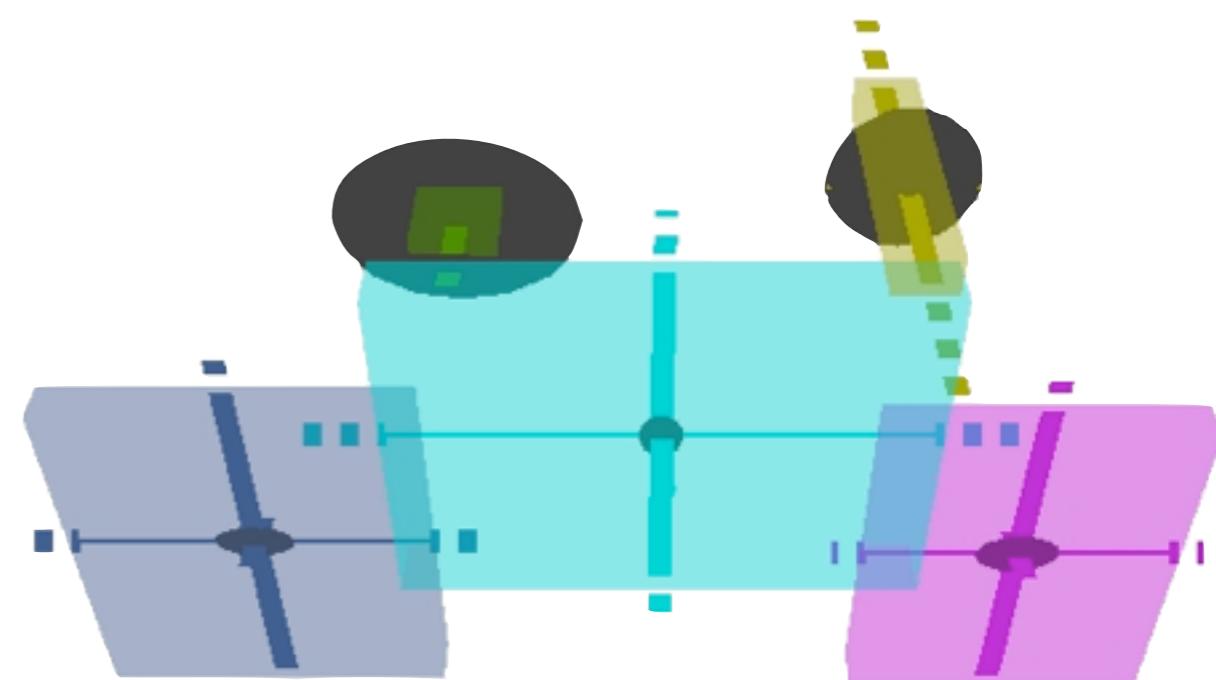
Initial Template:



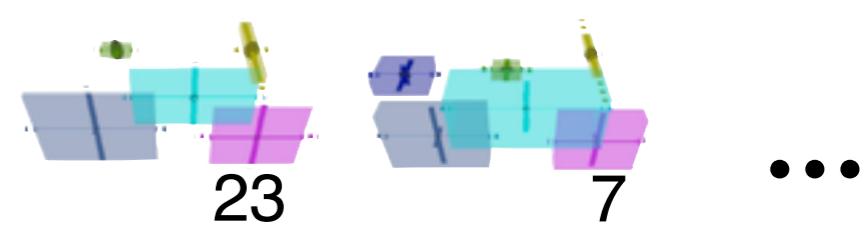
Final Templates:



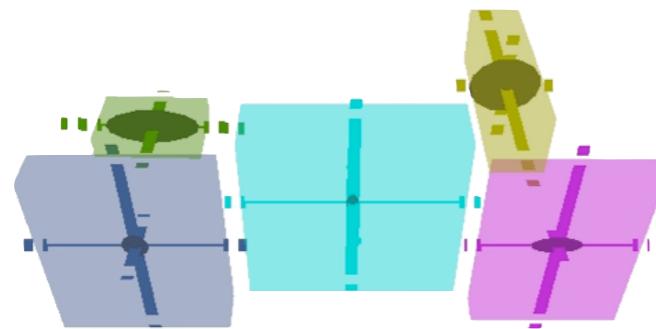
378



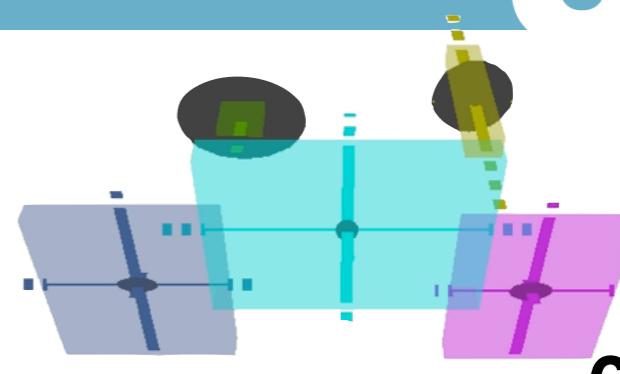
63



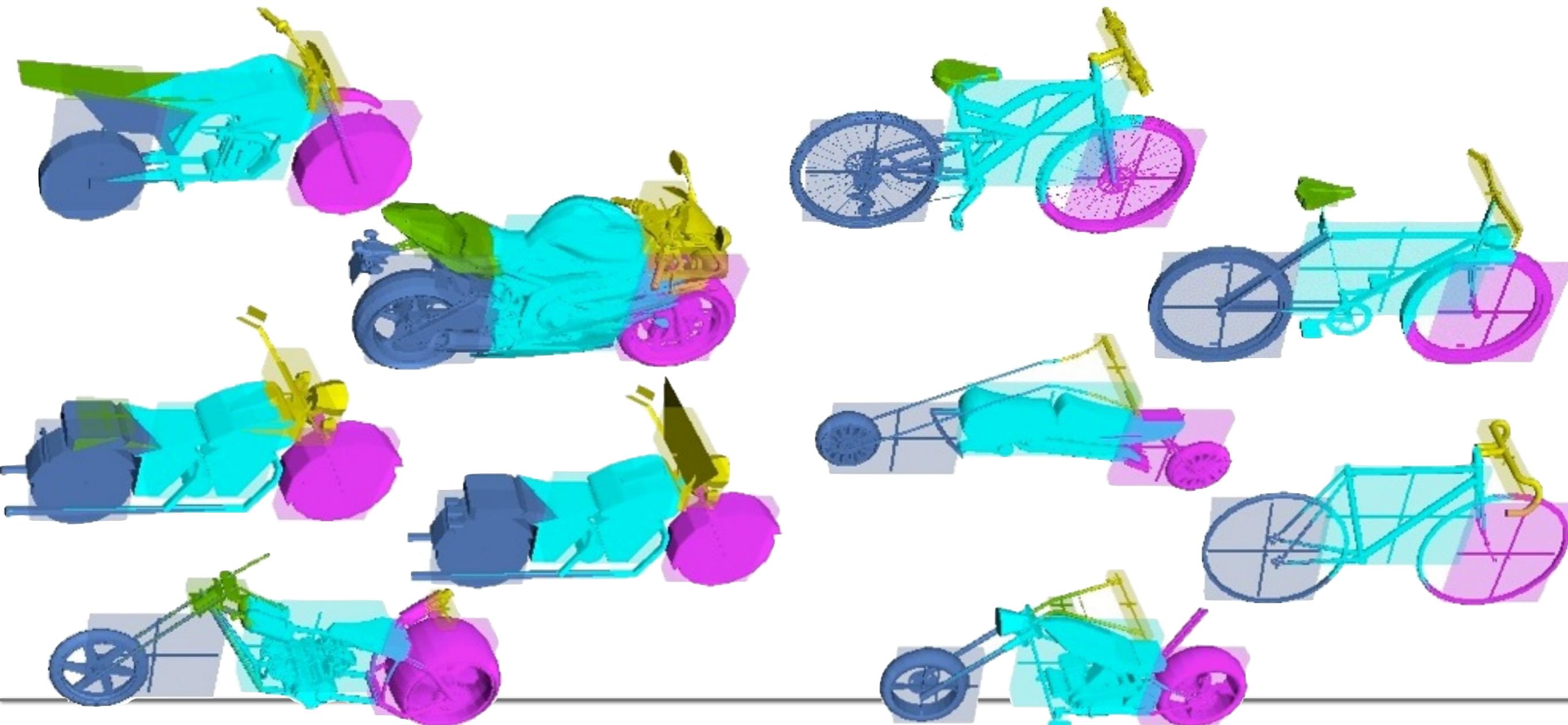
Bikes Dataset



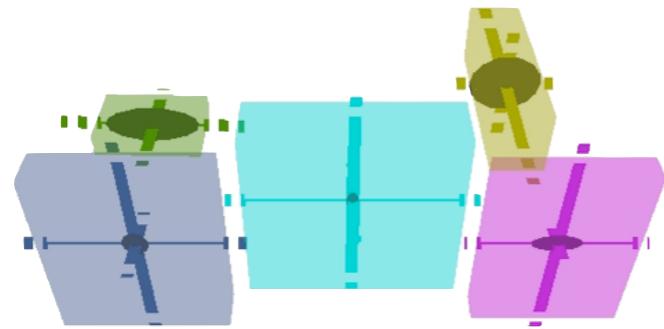
378



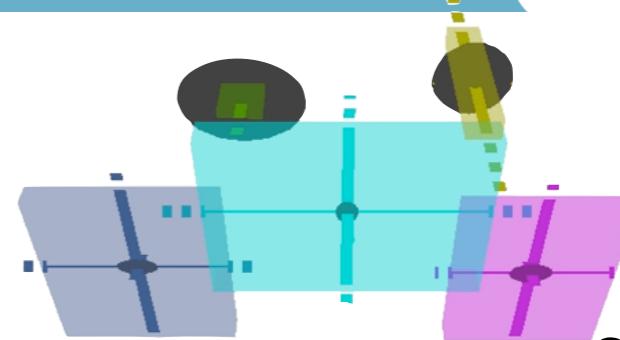
63



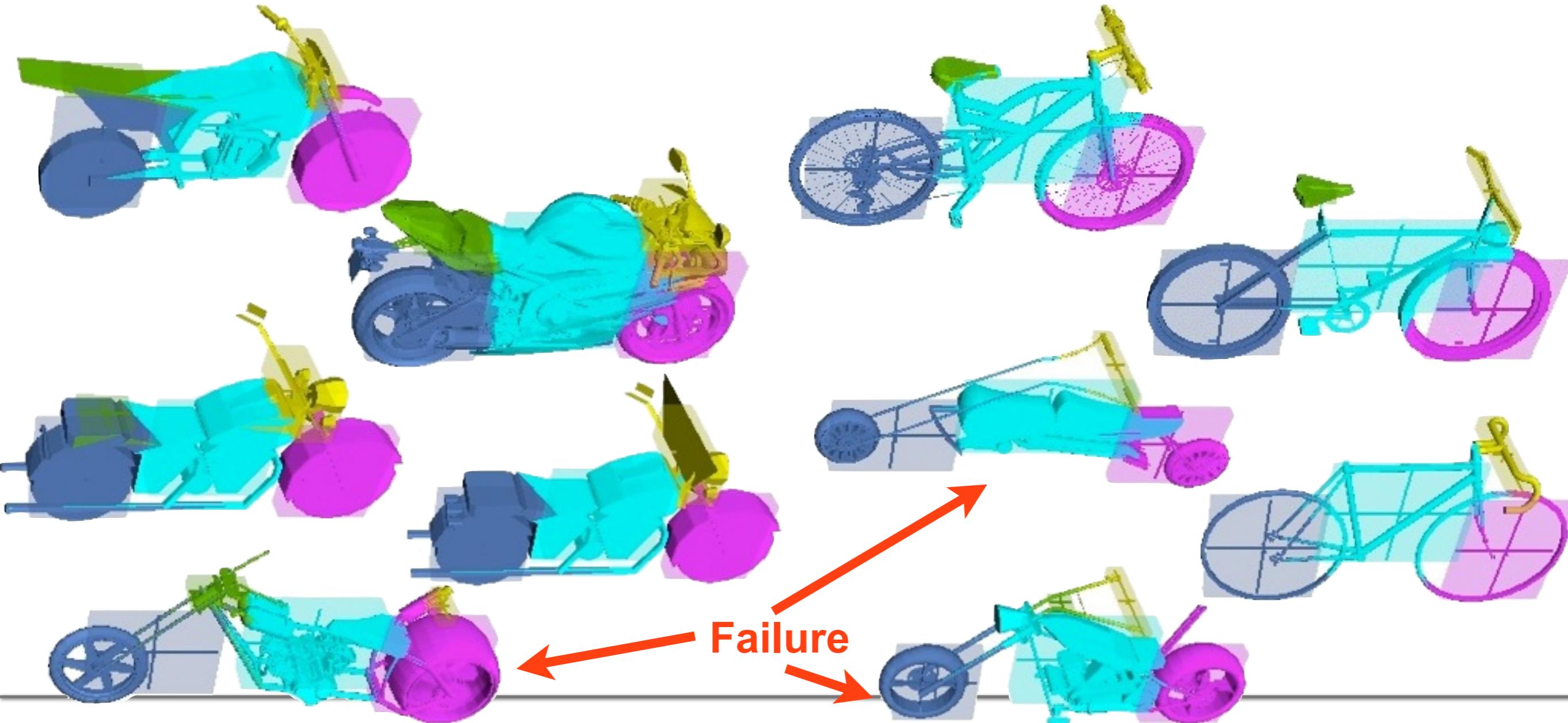
Limitation



378



63



Shape Variations: 400 chairs



400 chairs

Shape Variations: 400 chairs



400 chairs



Rigid

Shape Variations: 400 chairs



400 chairs



Rigid

Deformed

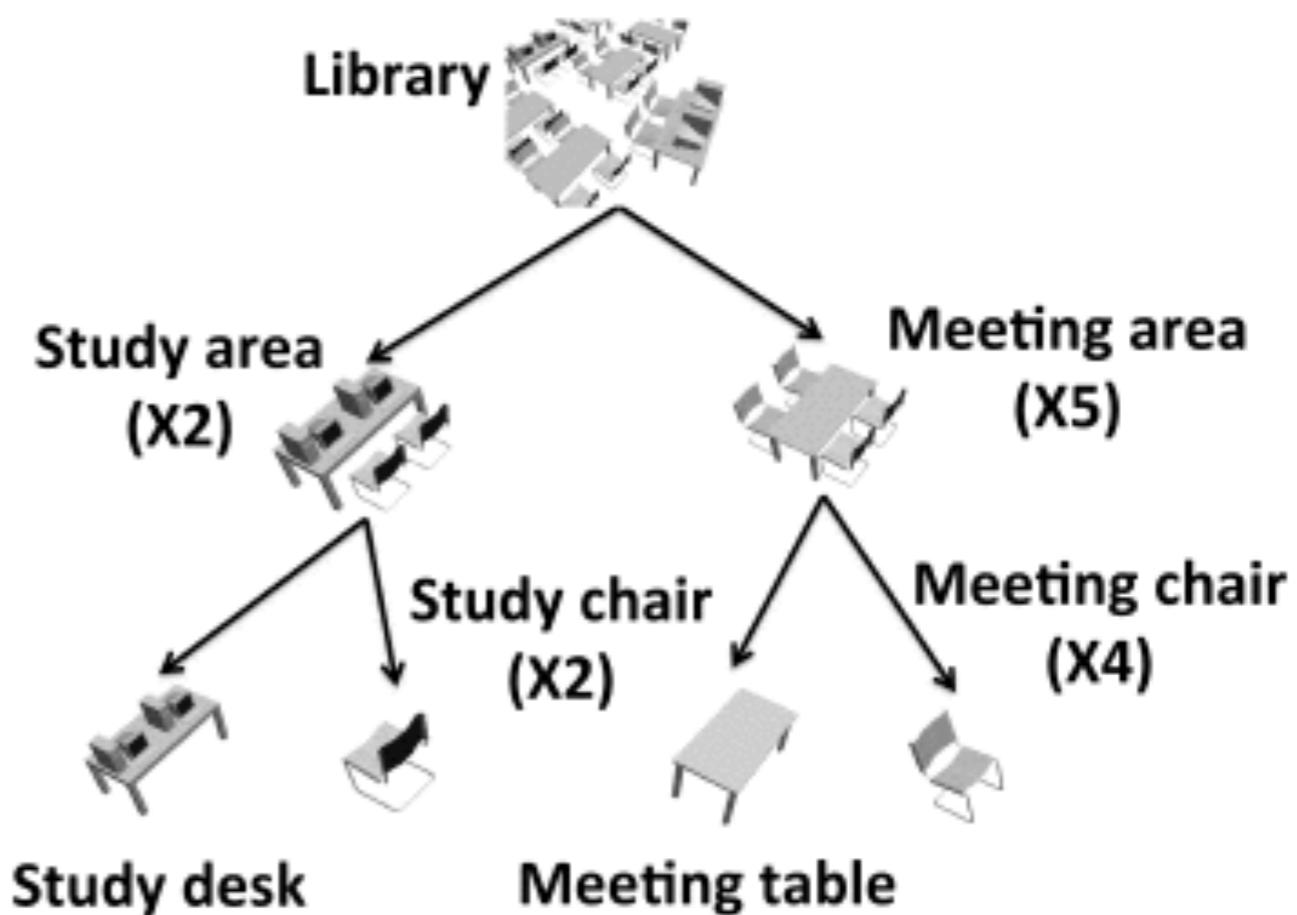
scene collections

How about Scenes?



scene as multi-levels of context (hierarchy)

How about Scenes?

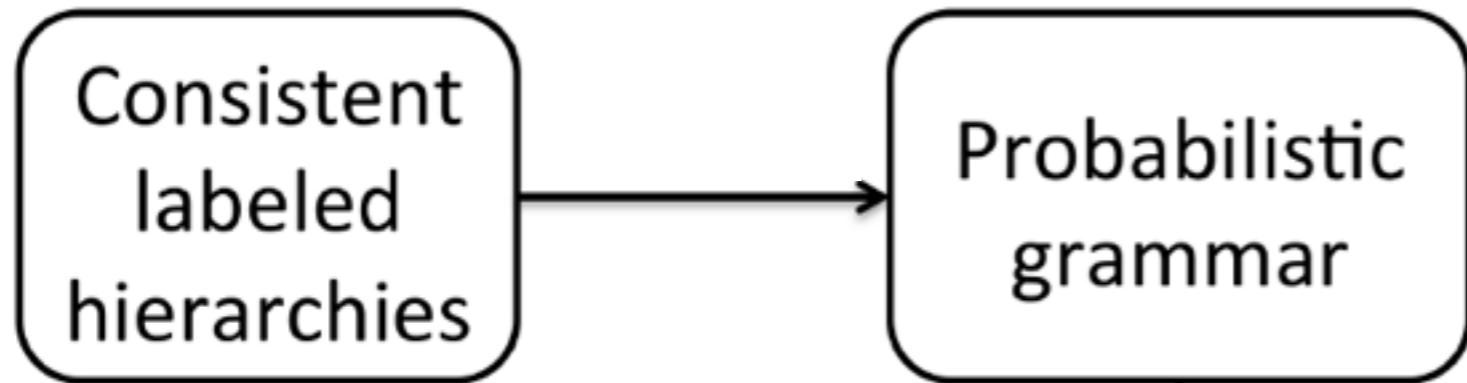


scene as multi-levels of context (hierarchy)

Approach

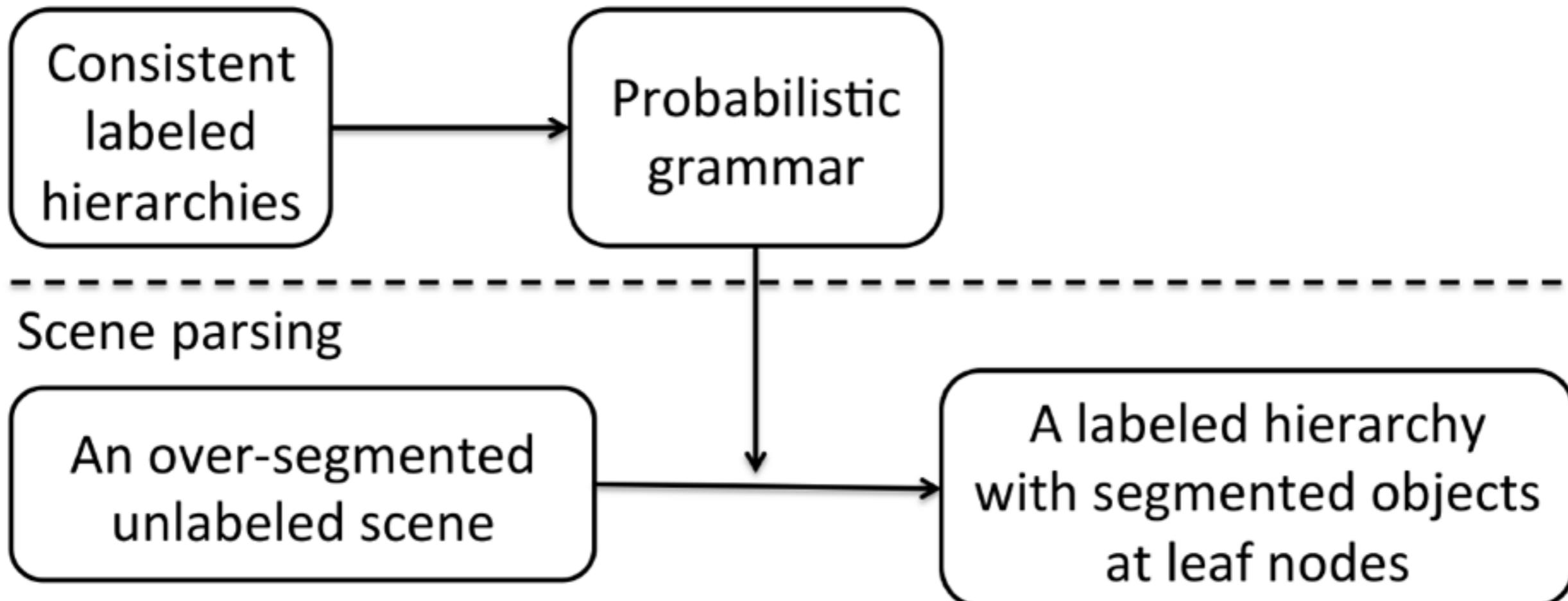


Grammar learning



Approach

Grammar learning



Formulation

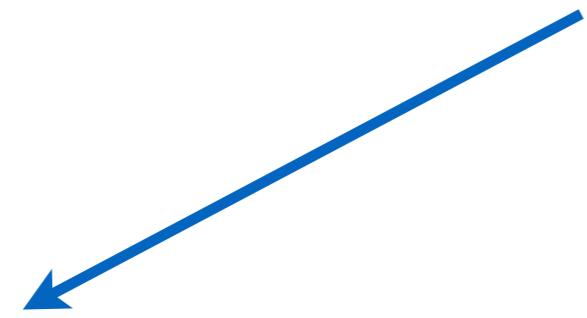


$$P(H|S, \mathcal{G}) \propto P(H|\mathcal{G}) \cdot P(S|H, \mathcal{G})$$

Formulation



$$P(H|S, \mathcal{G}) \propto P(H|\mathcal{G}) \cdot P(S|H, \mathcal{G})$$

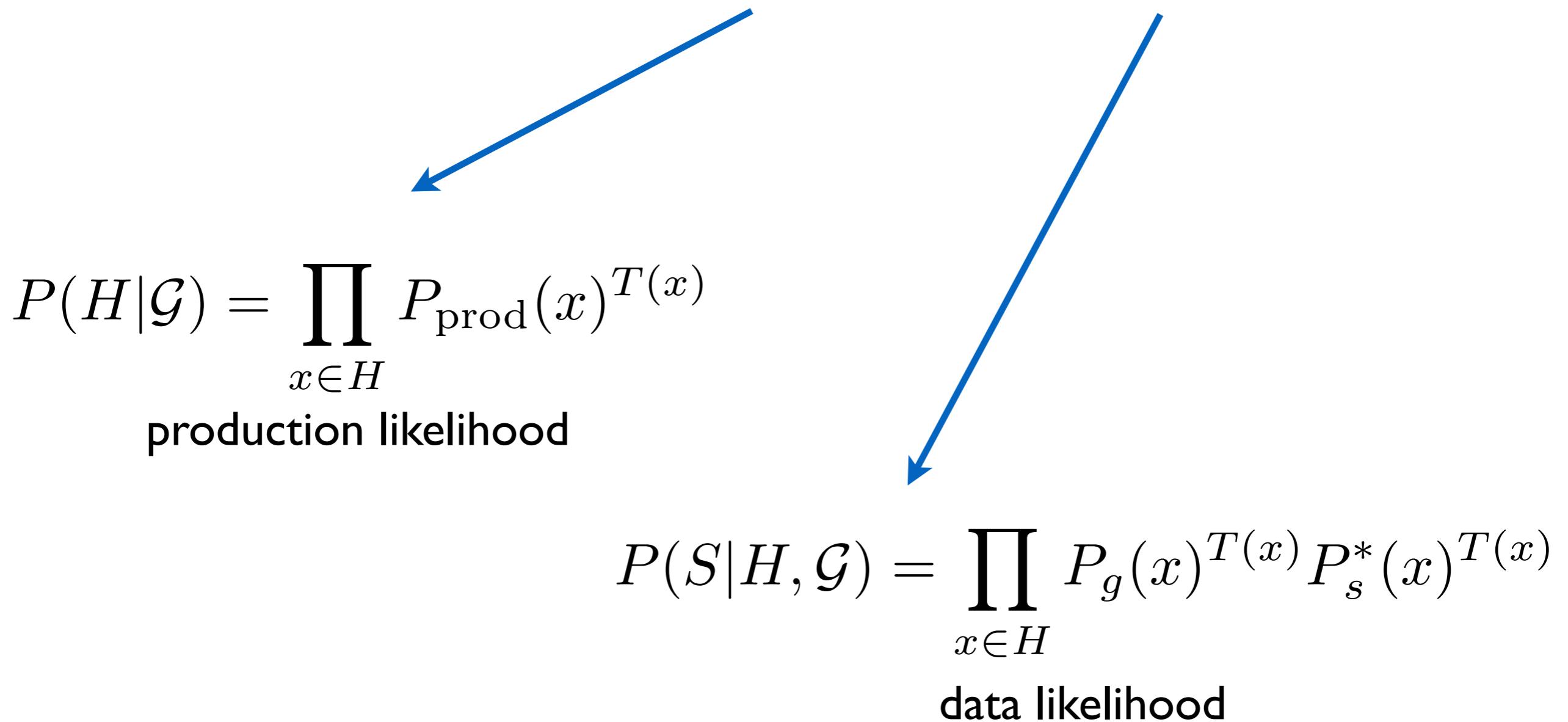


$$P(H|\mathcal{G}) = \prod_{x \in H} P_{\text{prod}}(x)^{T(x)}$$

production likelihood

Formulation

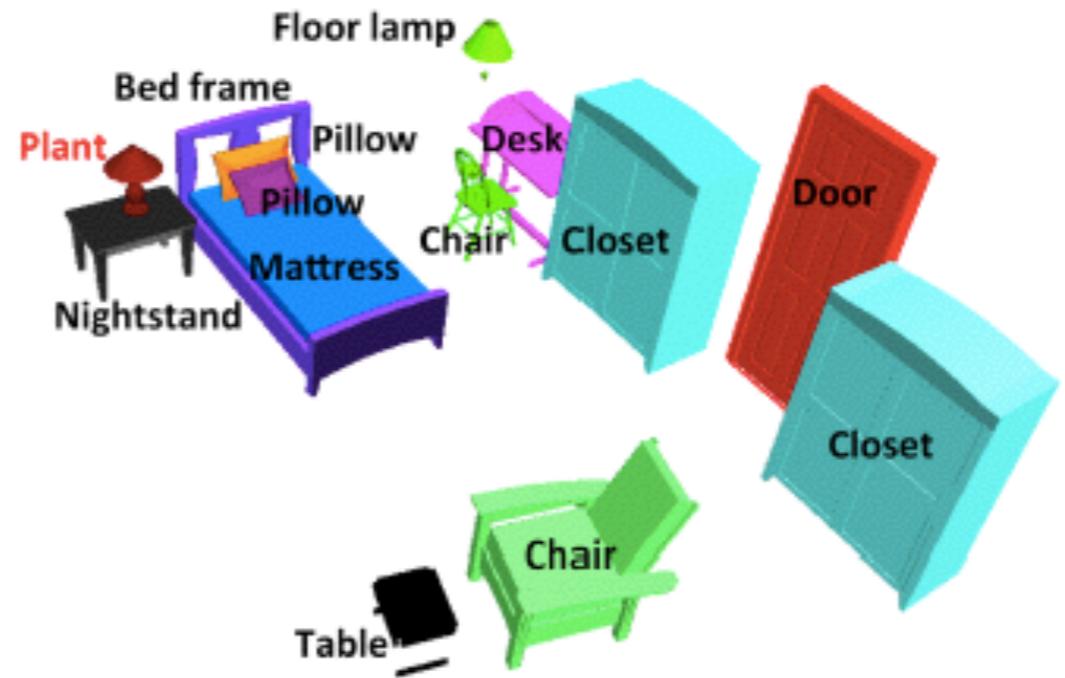
$$P(H|S, \mathcal{G}) \propto P(H|\mathcal{G}) \cdot P(S|H, \mathcal{G})$$



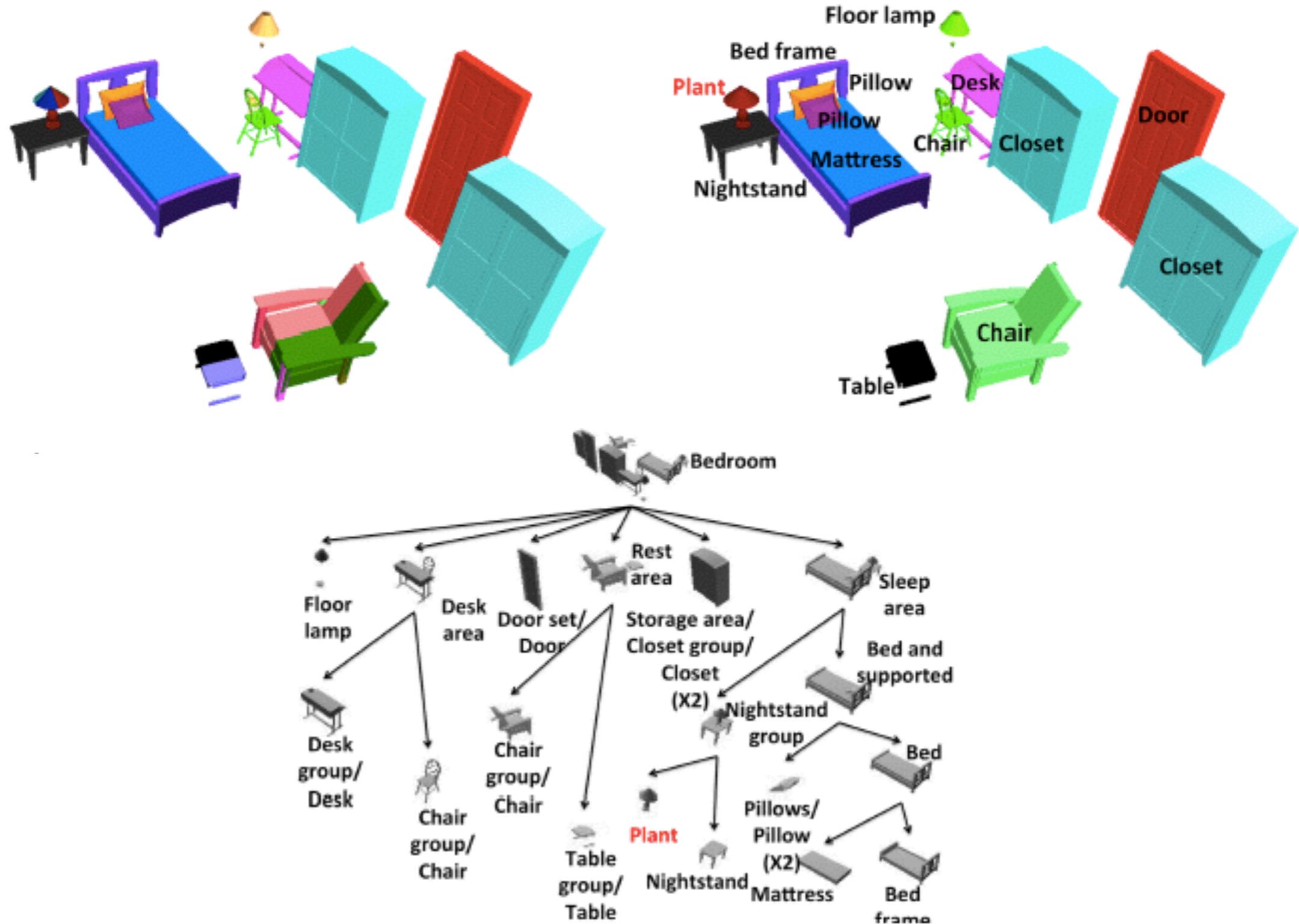
Result



Result



Result



Overview



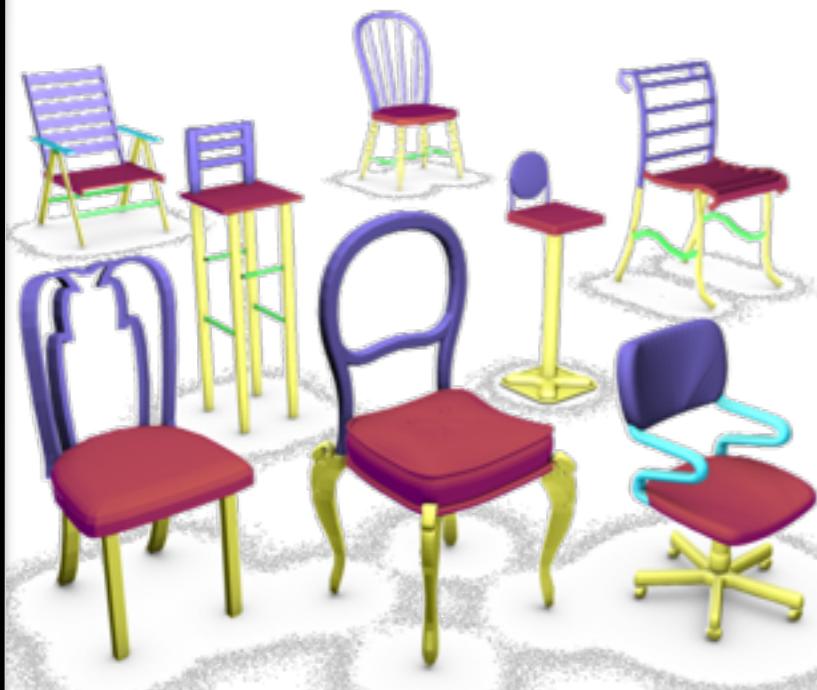
- Motivation
- Factored representations
 - in individual models
 - in model collections
 - in scene collections
- Applications
 - data exploration
 - ‘lifting’ to 3D

exploring smartly

Pdf-s for Exploration



Input



[Siggraph 2014]

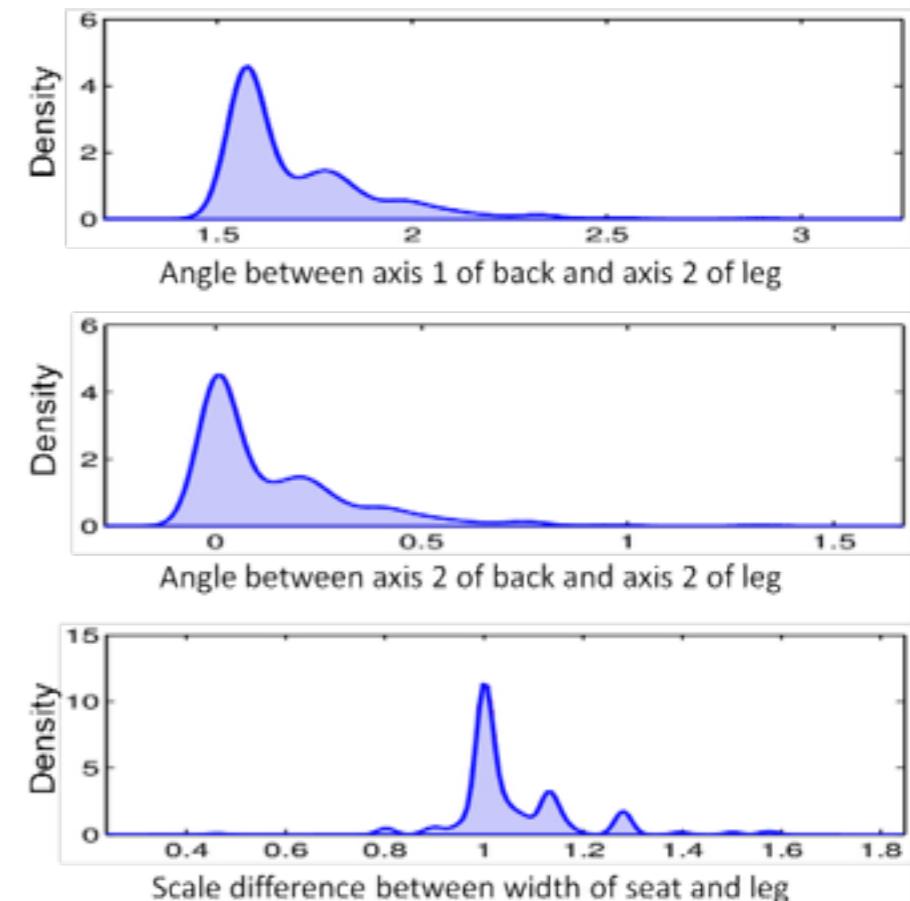
Pdf-s for Exploration



Input

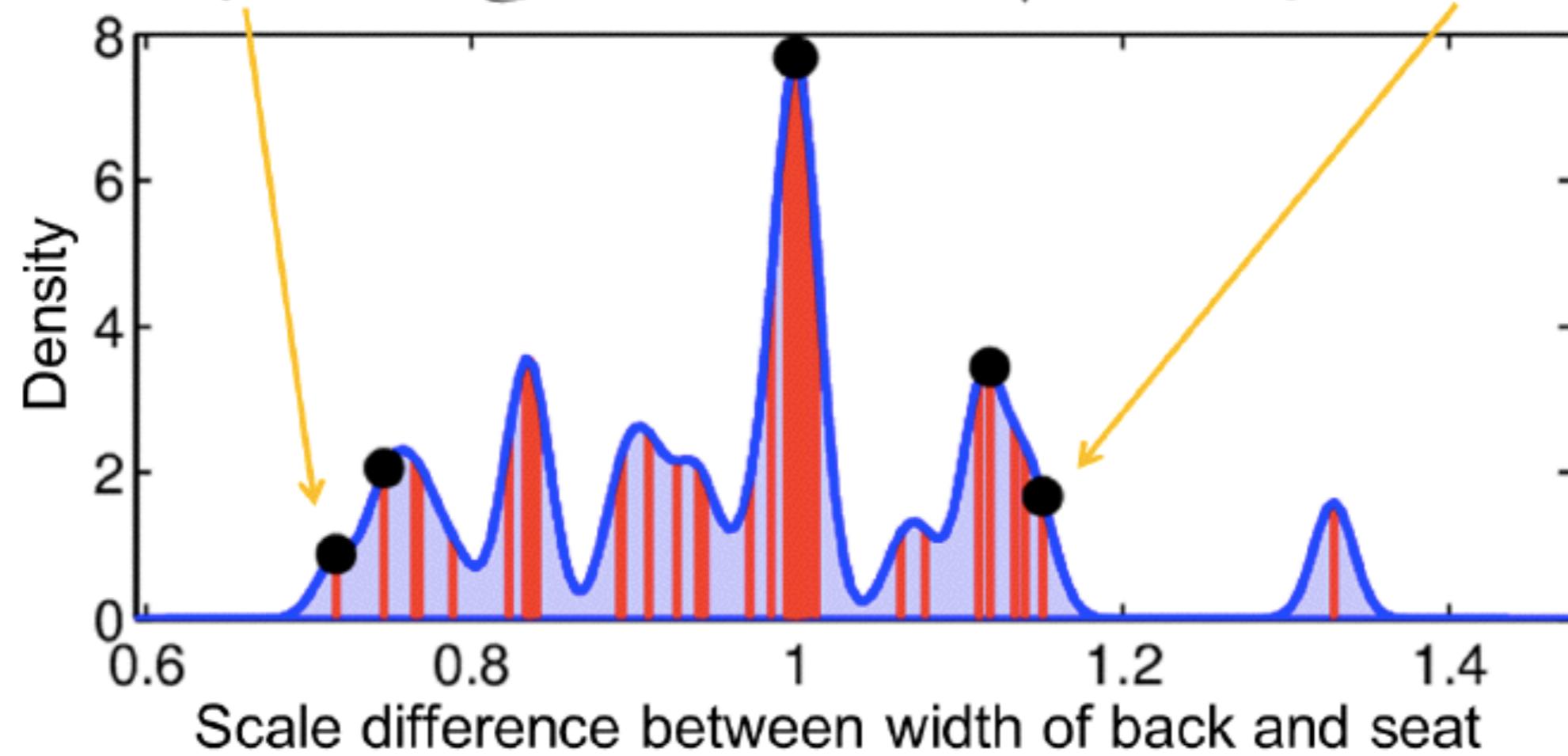


Meta-representation

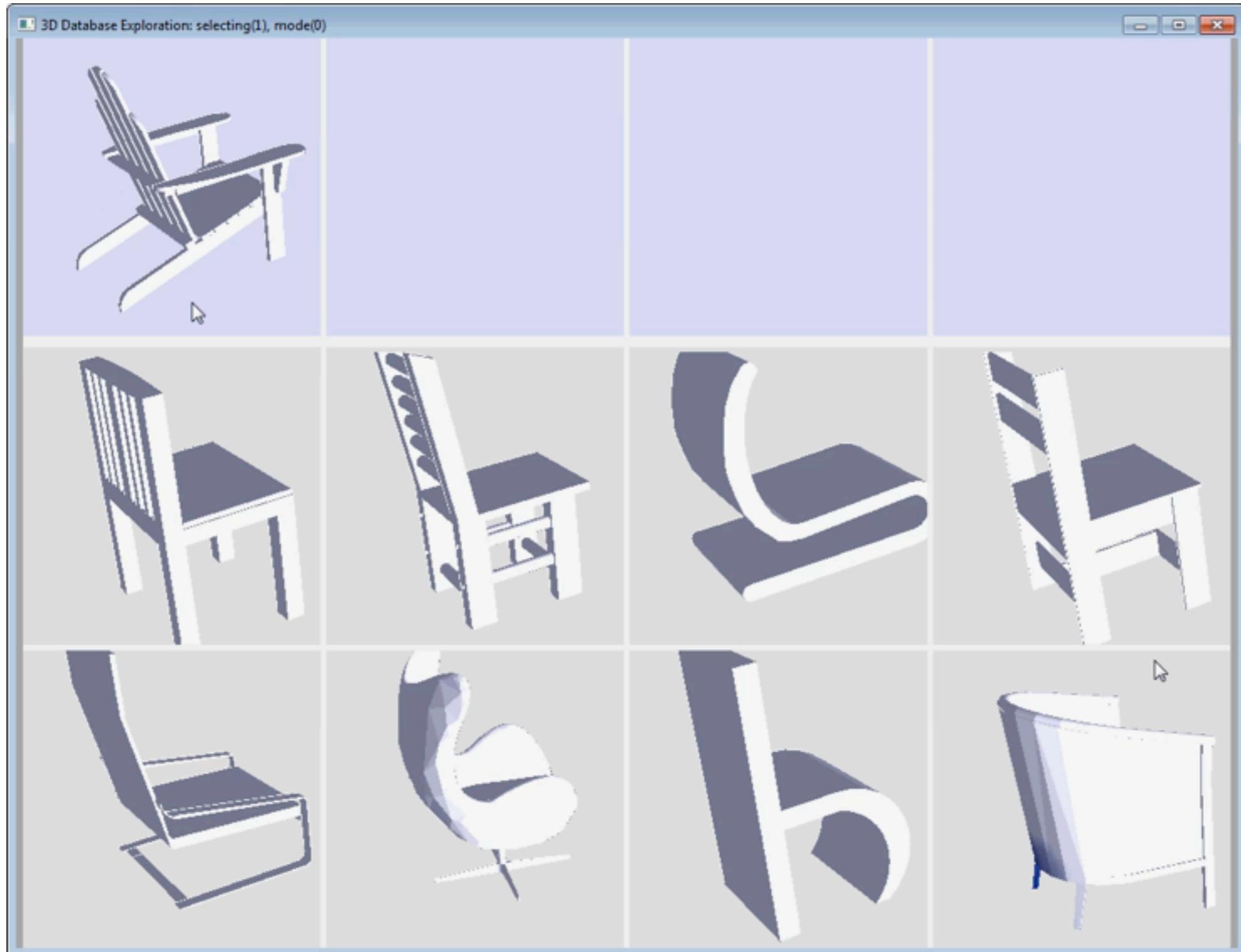


[Siggraph 2014]

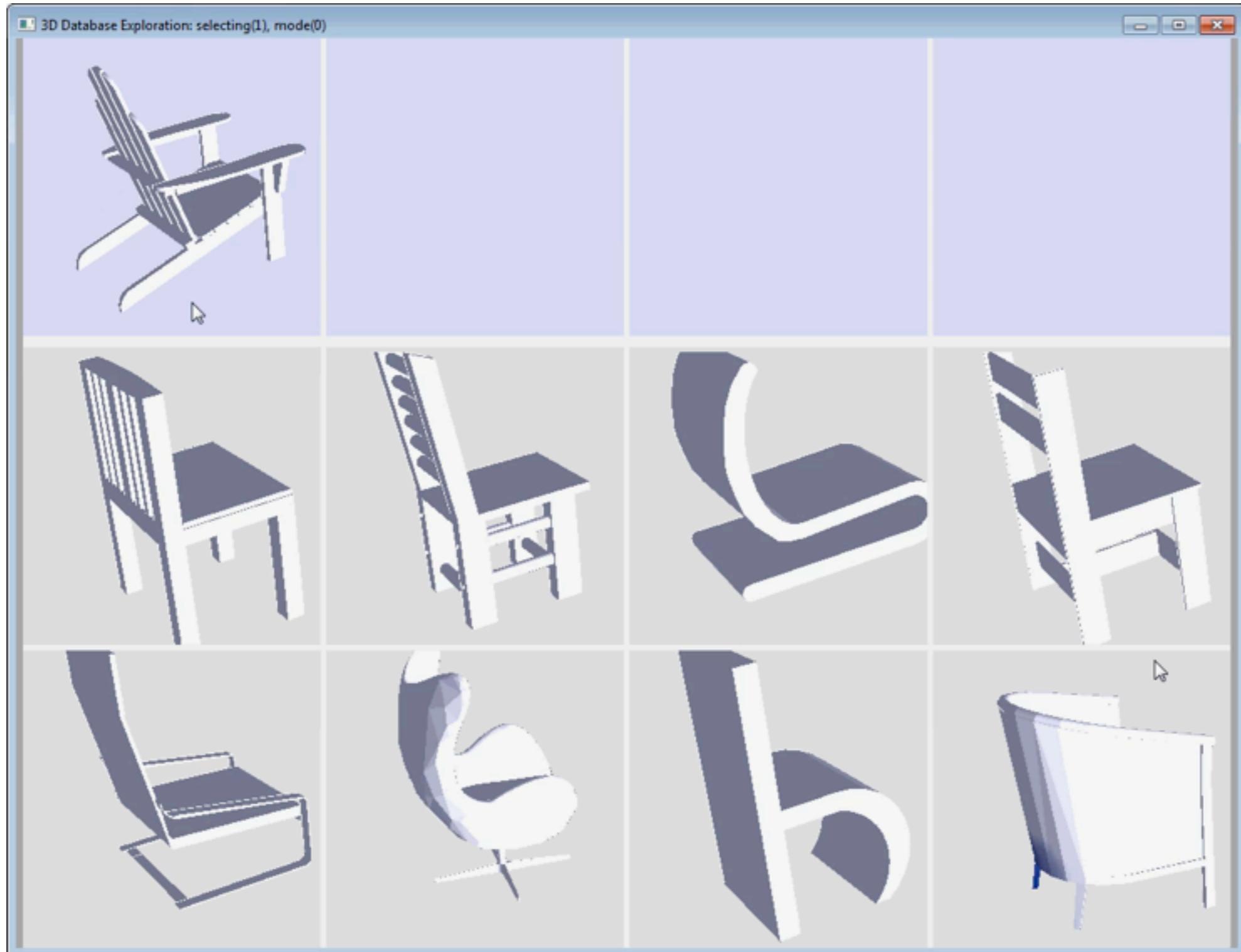
Using the Meta-representation



Exploration



Exploration



lifting to 3D

Image2Depth



[Siggraph 2014]

Image2Depth



Init. Depth Estimation

Depth Refinement

[Siggraph 2014]

Image2Depth



Init. Depth Estimation

Pose Est.

2D Matching

Depth Refinement

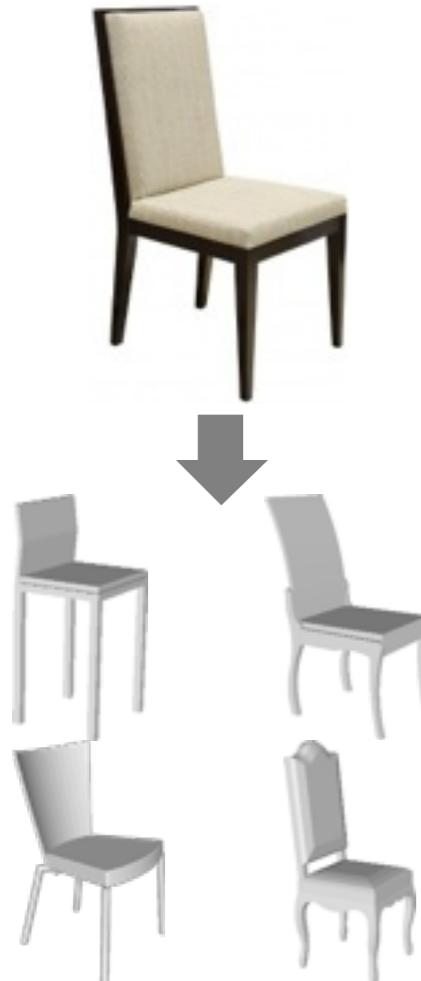
[Siggraph 2014]

Image2Depth



Init. Depth Estimation

Pose Est.



2D Matching

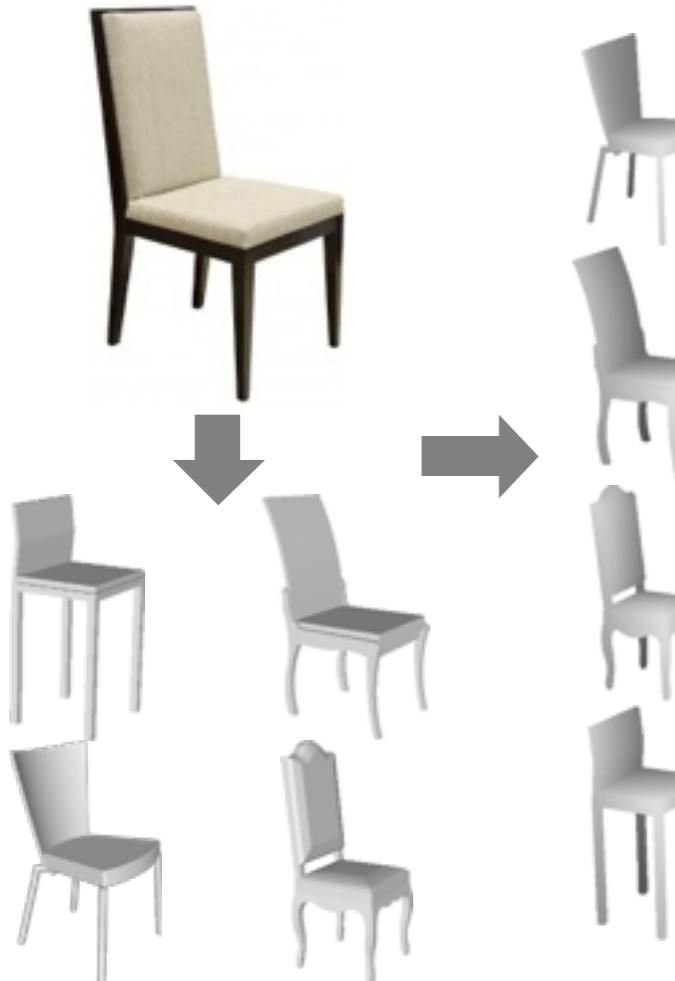
Depth Refinement

[Siggraph 2014]

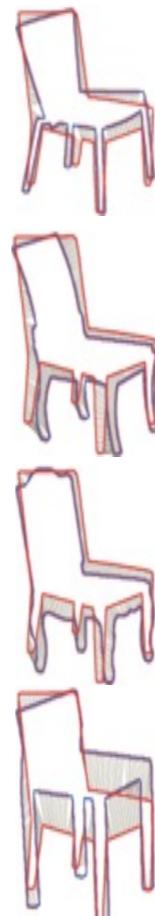
Image2Depth

Init. Depth Estimation

Pose Est.



2D Matching



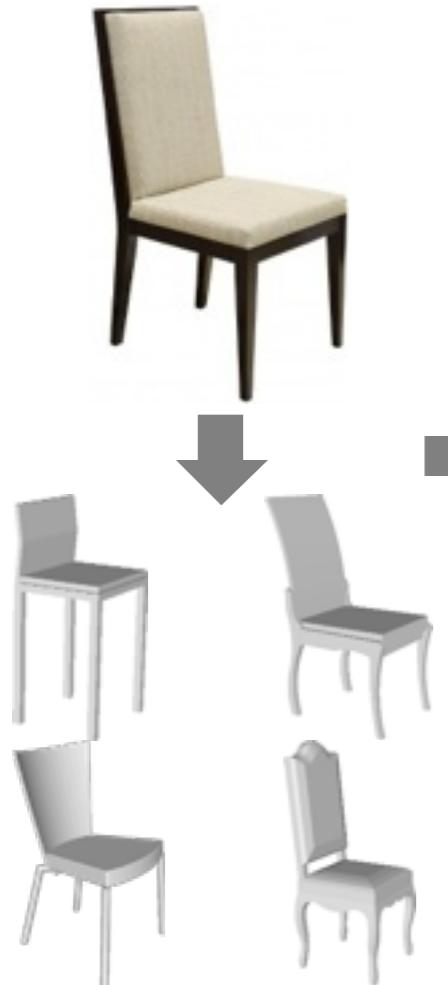
Depth Refinement

[Siggraph 2014]

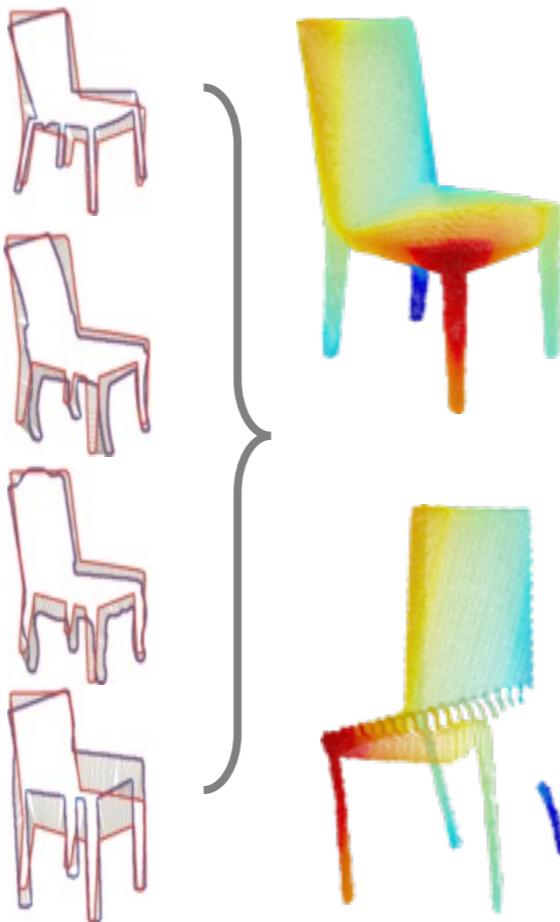
Image2Depth

Init. Depth Estimation

Pose Est.



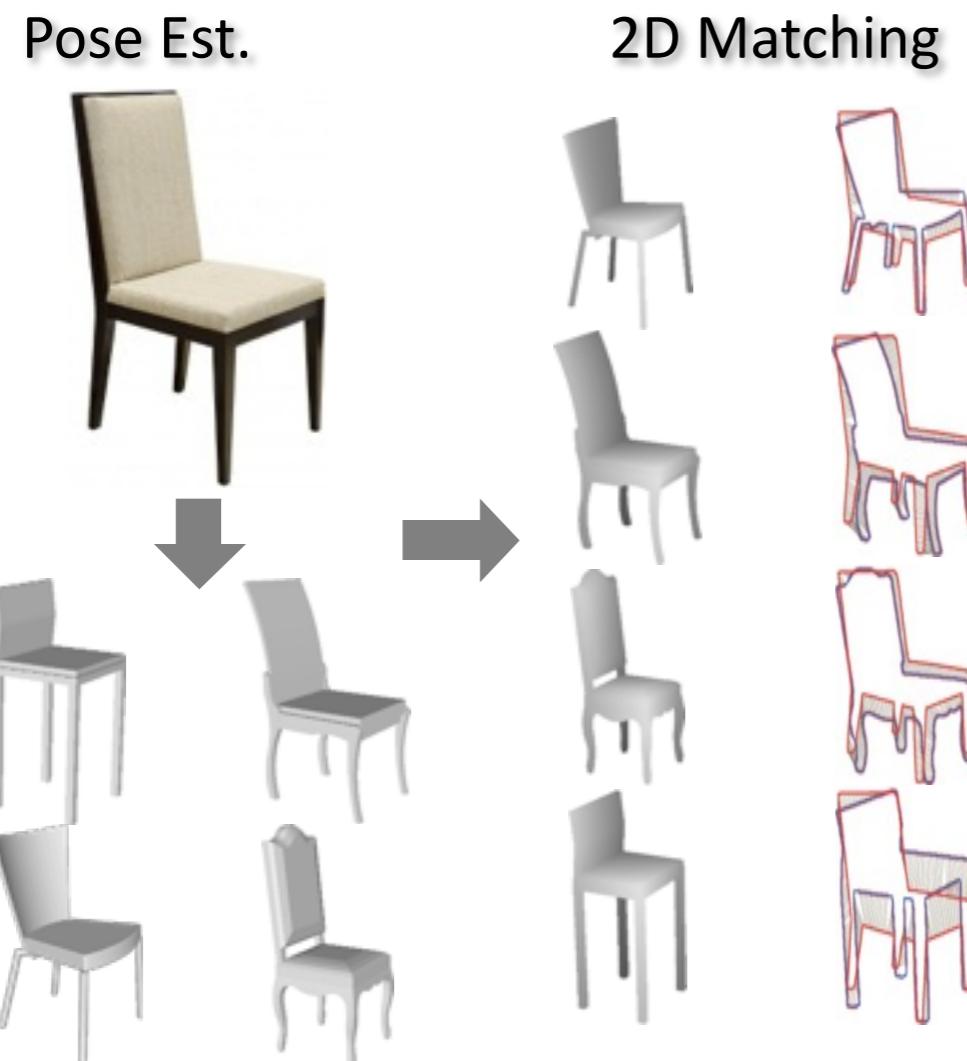
2D Matching



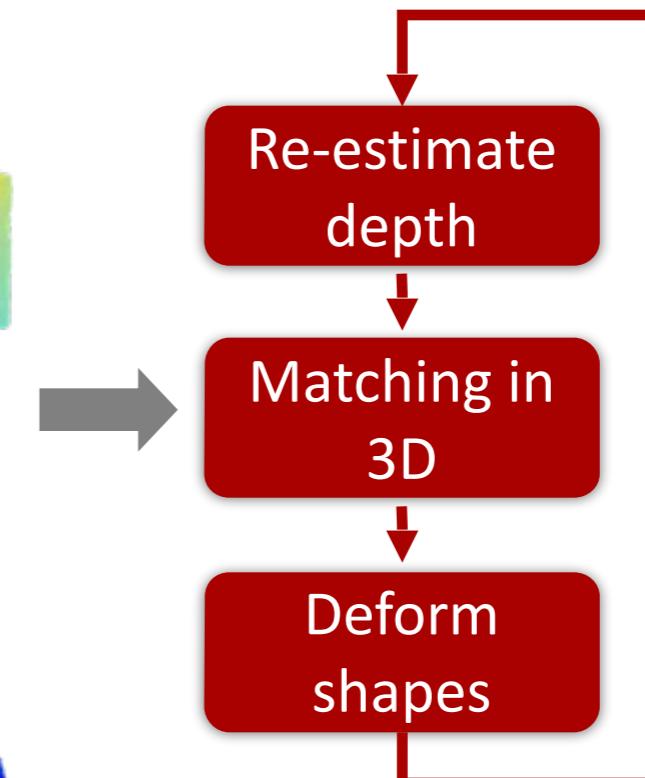
Depth Refinement

Image2Depth

Init. Depth Estimation



Depth Refinement



Init. Depth Est.

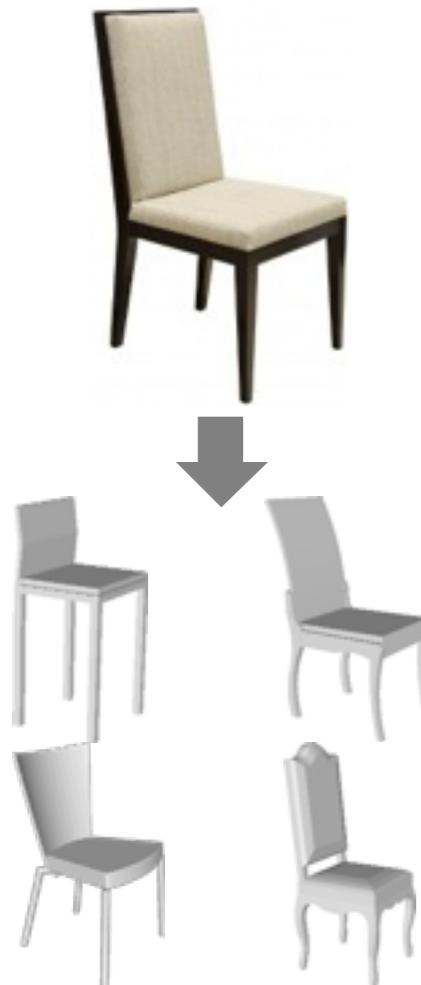
[Siggraph 2014]

Image2Depth



Init. Depth Estimation

Pose Est.



2D Matching



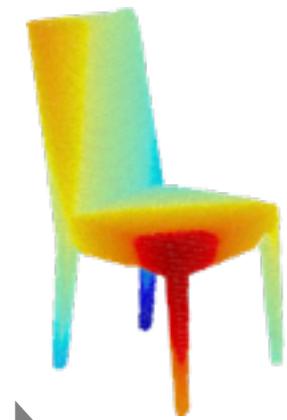
Init. Depth Est.

Depth Refinement

Re-estimate depth

Matching in 3D

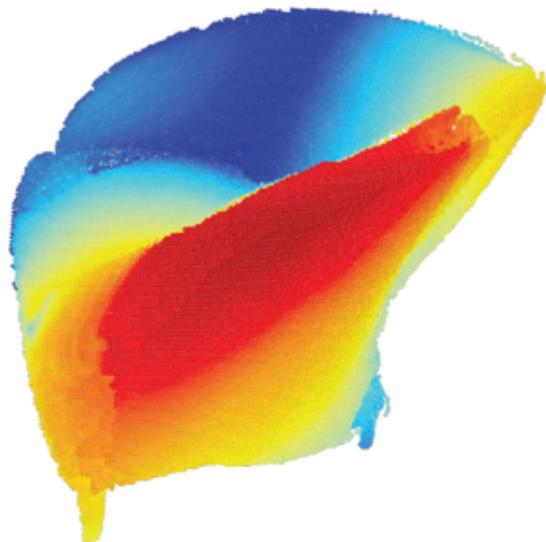
Deform shapes



Final Depth Est.

[Siggraph 2014]

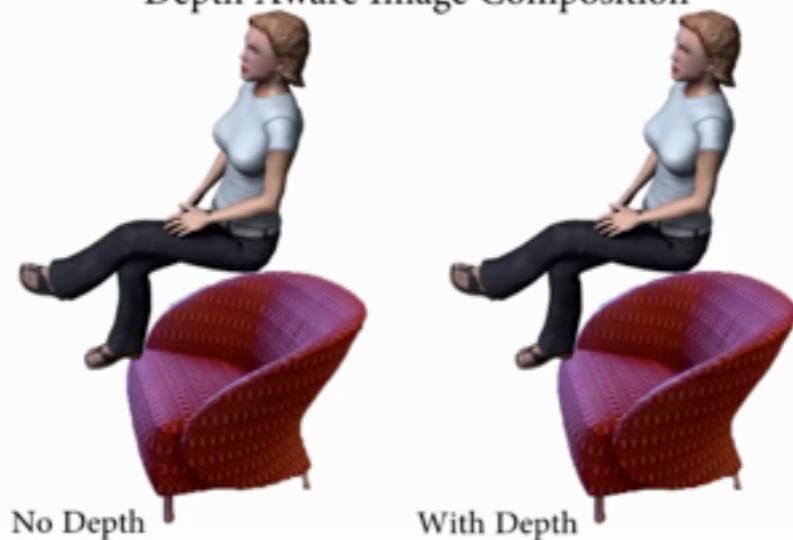
High-quality Point Cloud Reconstruction



Depth Aware Image Composition



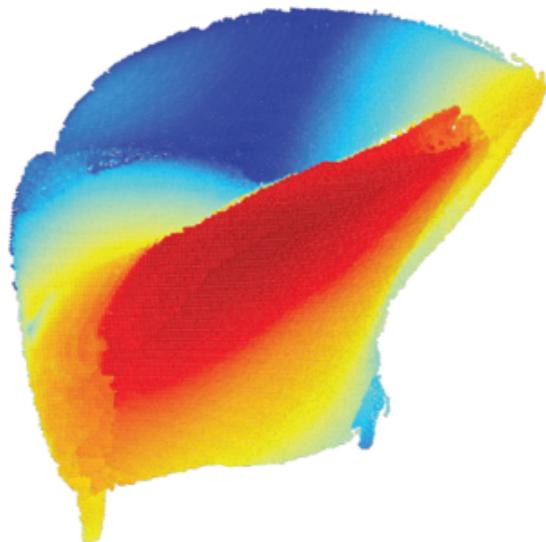
Depth Aware Image Composition



Relighting



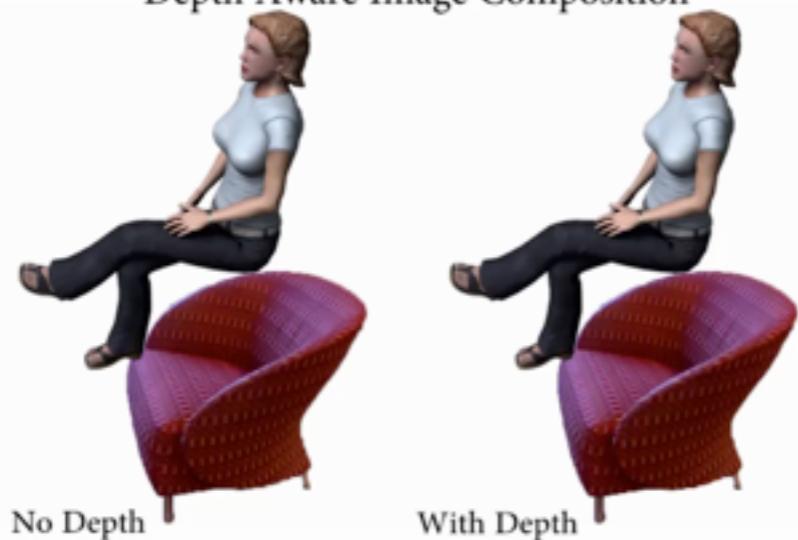
High-quality Point Cloud Reconstruction



Depth Aware Image Composition



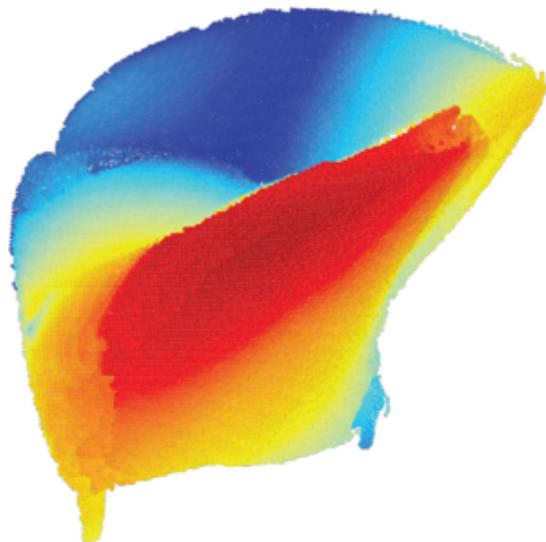
Depth Aware Image Composition



Relighting



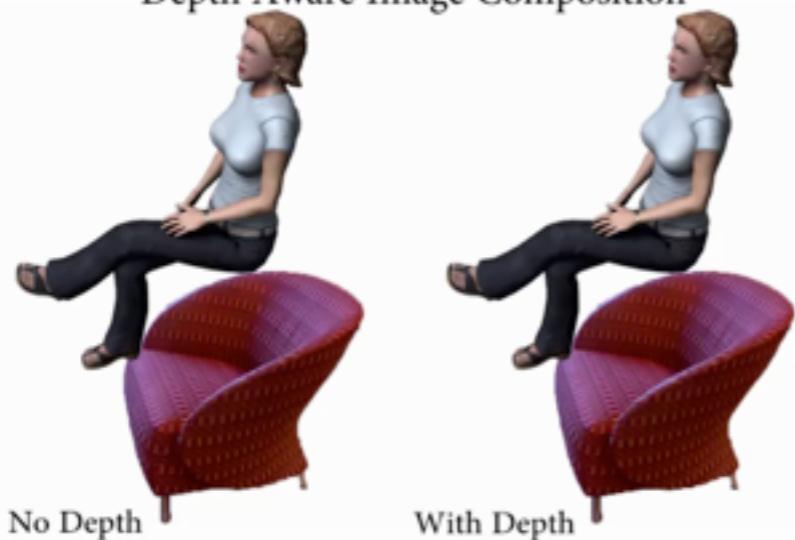
High-quality Point Cloud Reconstruction



Depth Aware Image Composition



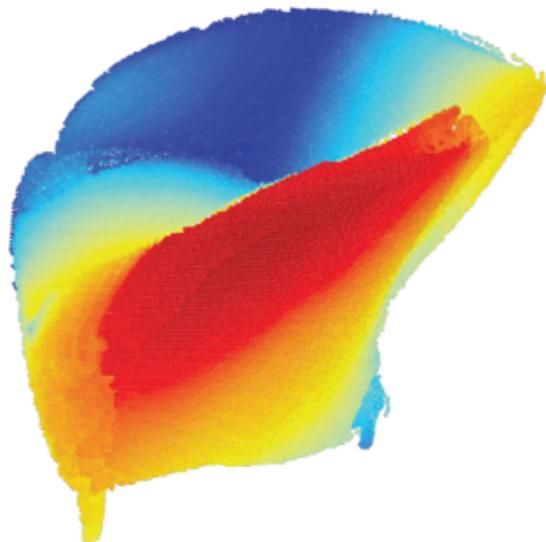
Depth Aware Image Composition



Relighting



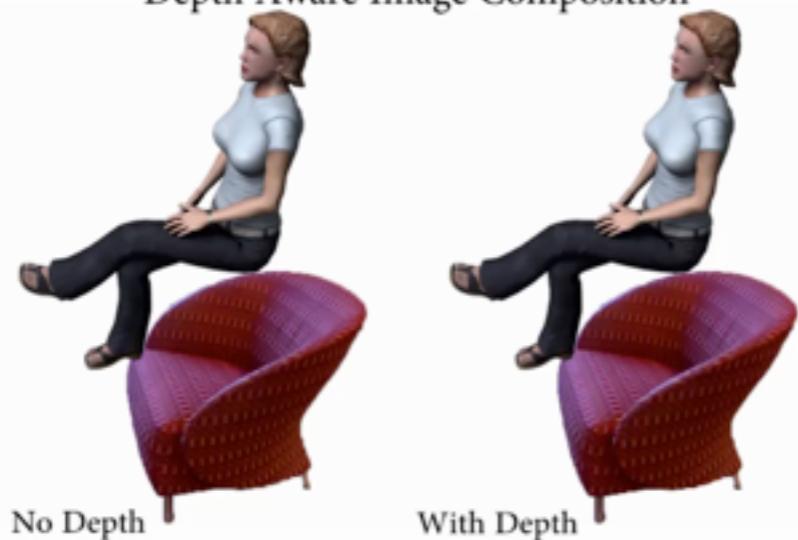
High-quality Point Cloud Reconstruction



Depth Aware Image Composition



Depth Aware Image Composition



Relighting



Summary



Summary



- Relations > individuals (parts)

Summary



- Relations > individuals (parts)
- Exploit *structure* of constraints and correspondence encoded as distributions

Summary

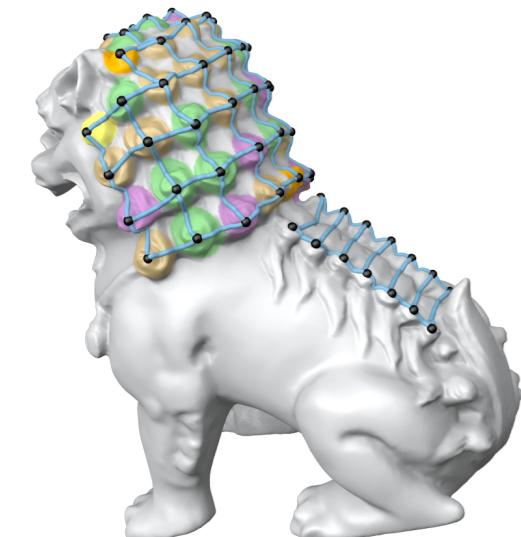


- Relations > individuals (parts)
- Exploit *structure* of constraints and correspondence encoded as distributions
- Exploit *correlation* in data

Summary

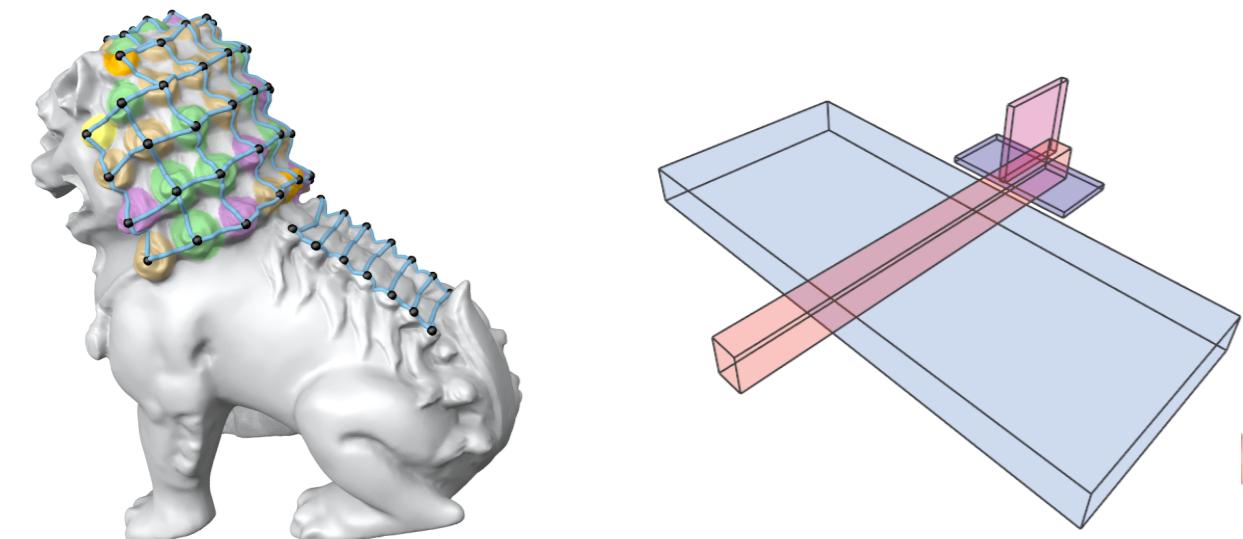


- Relations > individuals (parts)
- Exploit *structure* of constraints and correspondence encoded as distributions
- Exploit *correlation* in data



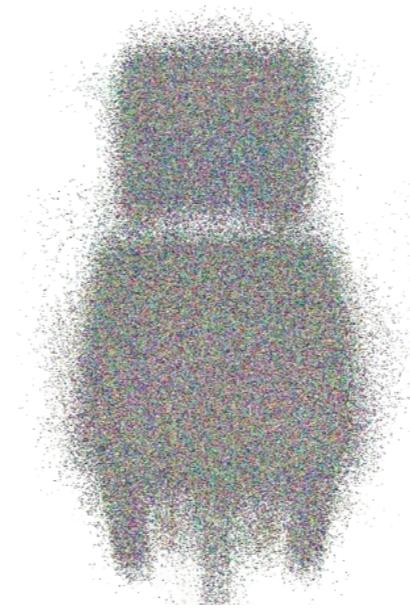
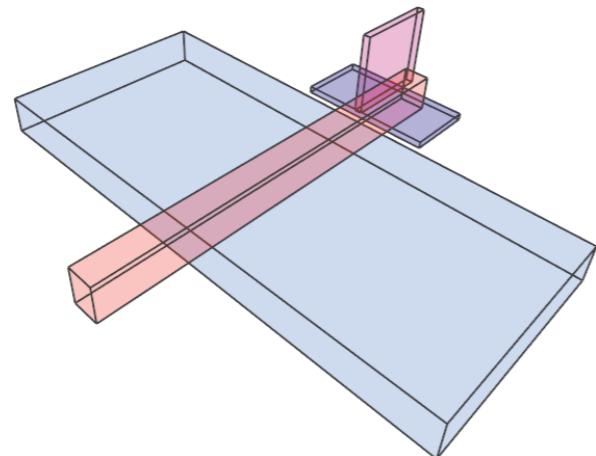
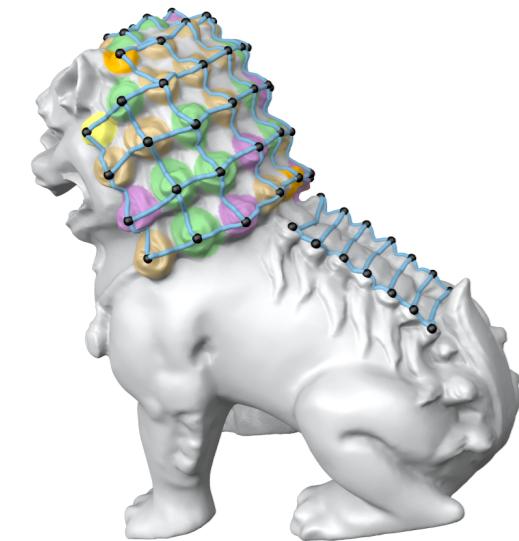
Summary

- Relations > individuals (parts)
- Exploit *structure* of constraints and correspondence encoded as distributions
- Exploit *correlation* in data



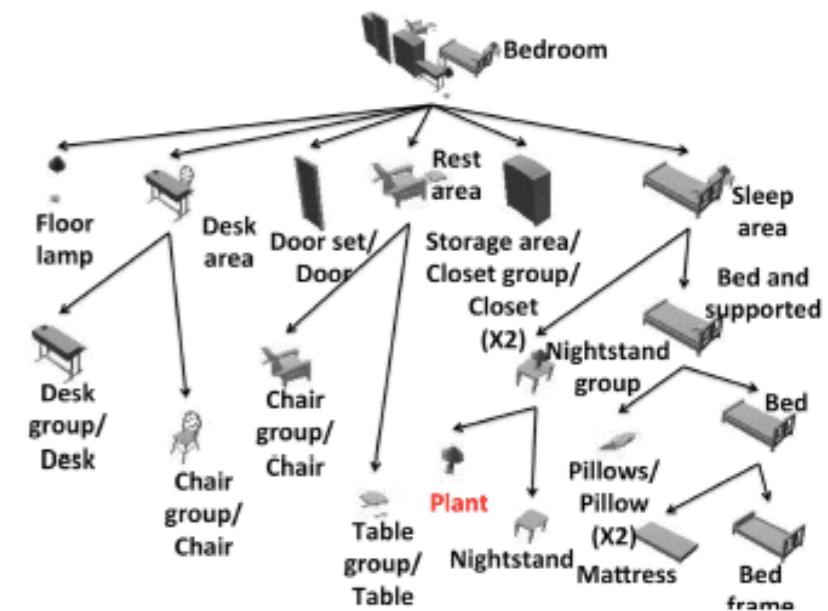
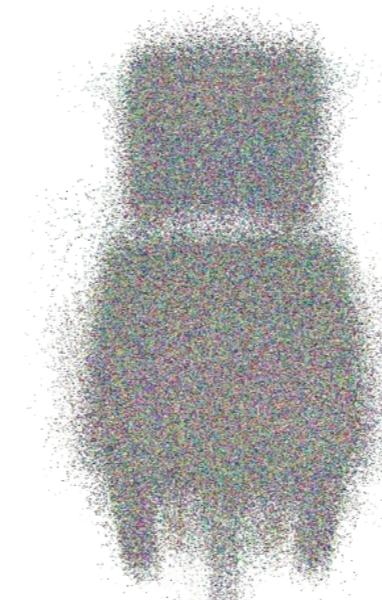
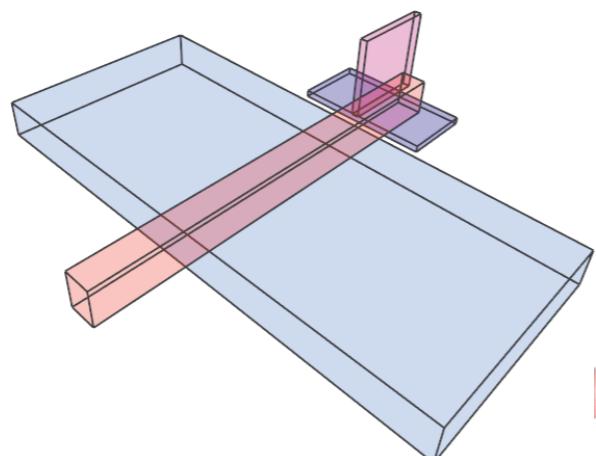
Summary

- Relations > individuals (parts)
- Exploit **structure** of constraints and correspondence encoded as distributions
- Exploit **correlation** in data

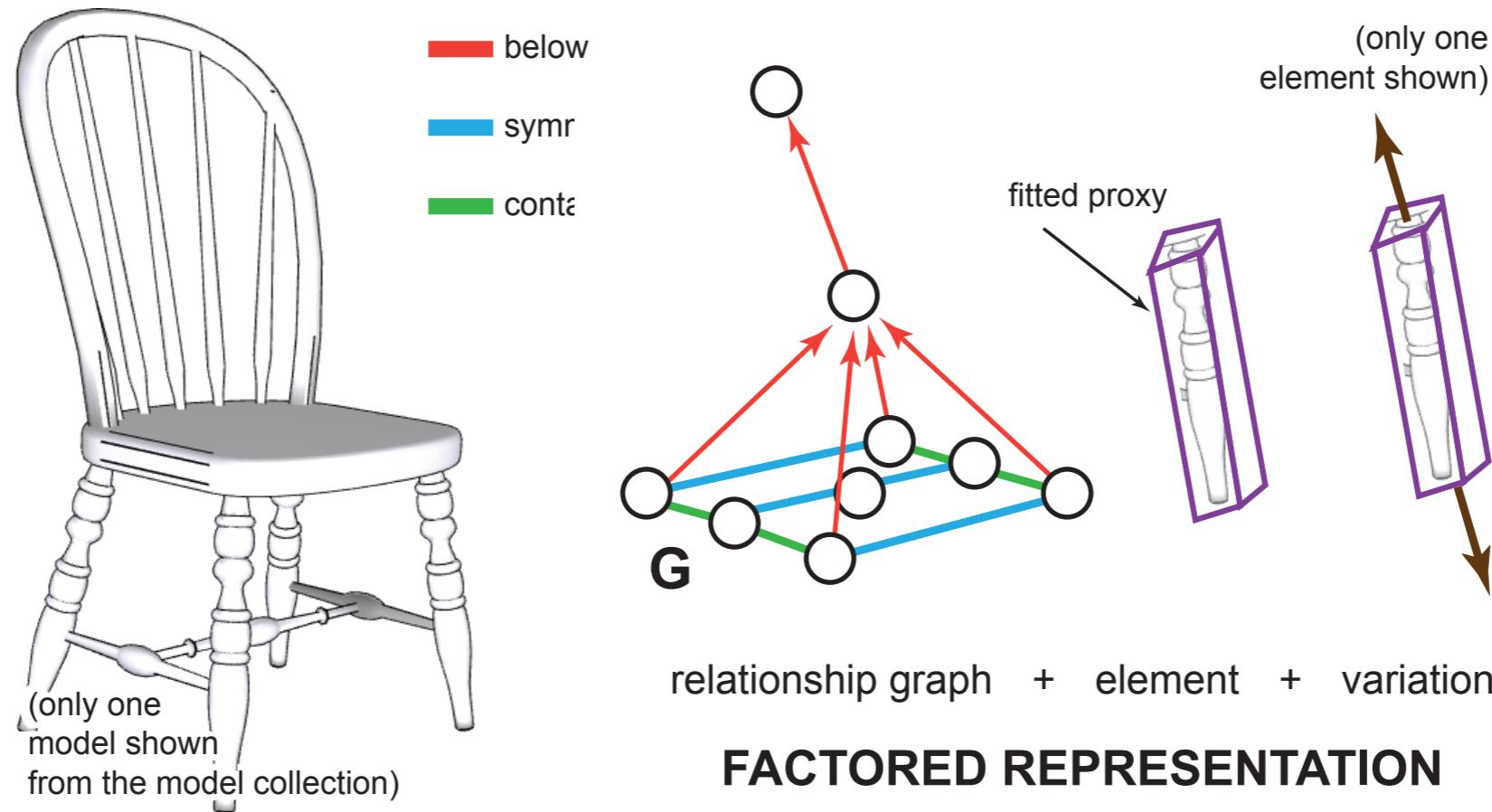


Summary

- Relations > individuals (parts)
- Exploit *structure* of constraints and correspondence encoded as distributions
- Exploit *correlation* in data



Thank You



<http://vecg.cs.ucl.ac.uk/Projects/SmartGeometry/>

(code + data available)

