

# CS233, CME251: Geometric and Topological Data Analysis

Leonidas Guibas  
Computer Science Department  
Stanford University

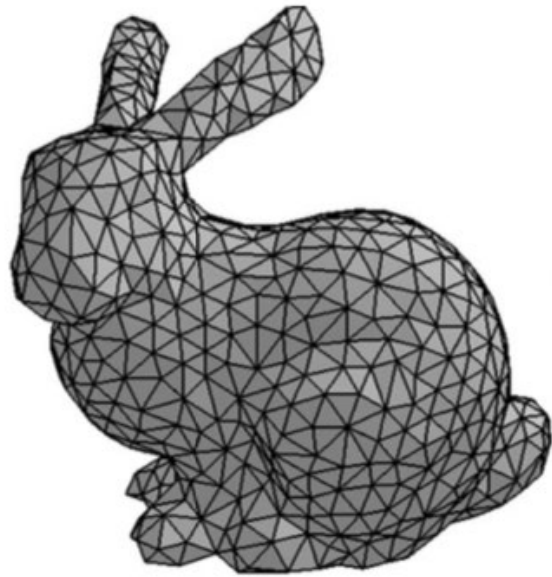


Lecture 18  
10 June 2020



# Last Time: Deep Learning on Graphs and Meshes

# CNNs on Irregular, Non-Euclidean Domains



3D shape graph



General graphs

# Different Formulations of Non-Euclidean CNNs



Spectral domain



Spatial domain

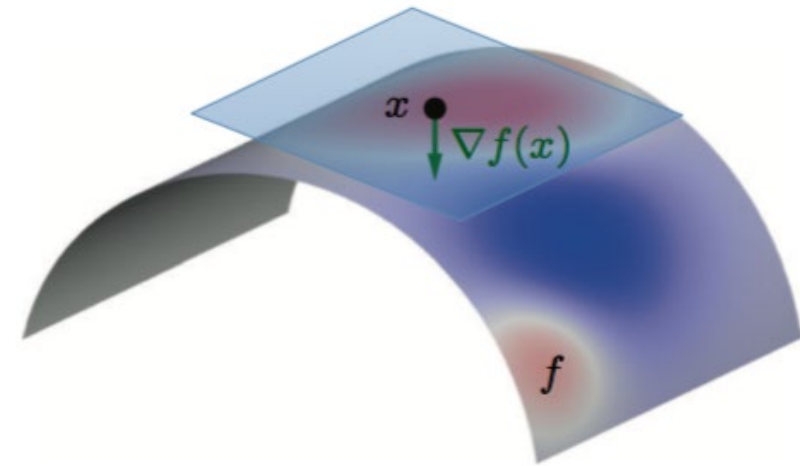
# Spectral Methods: Laplacian is Key



- Intrinsic gradient operator

$$\nabla f : L^2(\mathcal{X}) \rightarrow L^2(T\mathcal{X})$$

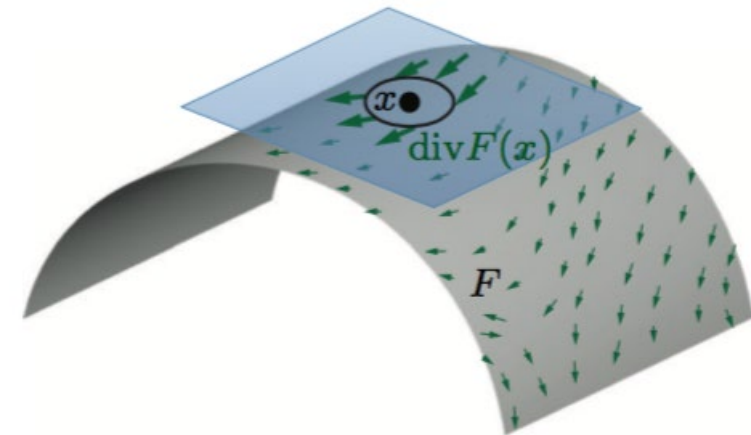
“direction of steepest change of  $f$ ”



- Intrinsic divergence operator

$$\text{div} : L^2(T\mathcal{X}) \rightarrow L^2(\mathcal{X})$$

“net flow of field  $F$  at  $x$ ”

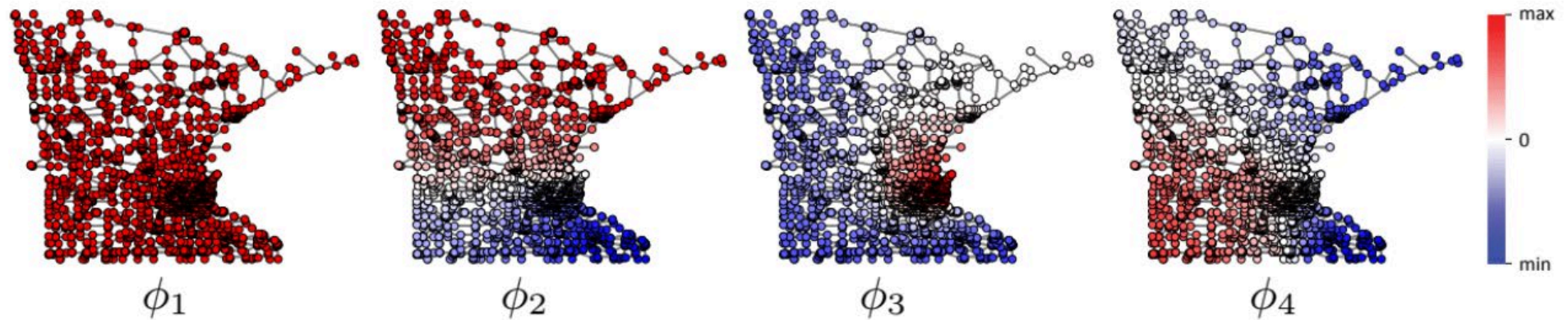


# Going Non-Euclidean: Laplacian is Key

Eigendecomposition of a graph Laplacian

$$\Delta = \Phi \Lambda \Phi^\top$$

where  $\Phi = (\phi_1, \dots, \phi_n)$  are **orthogonal eigenvectors** ( $\Phi^\top \Phi = \mathbf{I}$ ) and  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  the corresponding **non-negative eigenvalues**



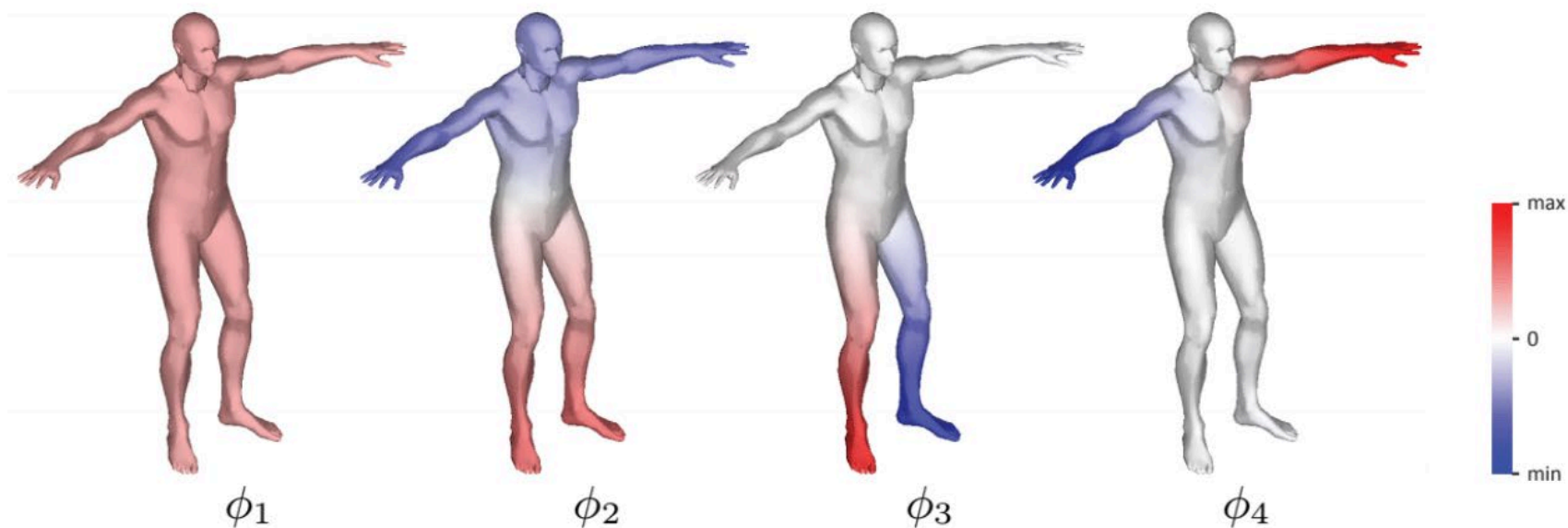
First eigenfunctions of a graph Laplacian

# Going Non-Euclidean: Laplacian is Key

Eigendecomposition of a graph Laplacian

$$\Delta = \Phi \Lambda \Phi^\top$$

where  $\Phi = (\phi_1, \dots, \phi_n)$  are **orthogonal eigenvectors** ( $\Phi^\top \Phi = \mathbf{I}$ ) and  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$  the corresponding **non-negative eigenvalues**



First eigenfunctions of a manifold Laplacian

# Convolution: Euclidean space

Given two functions  $f, g : [-\pi, \pi] \rightarrow \mathbb{R}$  their **convolution** is a function

$$(f \star g)(x) = \int_{-\pi}^{\pi} f(x')g(x - x')dx'$$

- **Shift-invariance:**  $f(x - x_0) \star g(x) = (f \star g)(x - x_0)$
- **Convolution theorem:** Fourier transform diagonalizes the convolution operator  $\Rightarrow$  convolution can be computed in the Fourier domain as

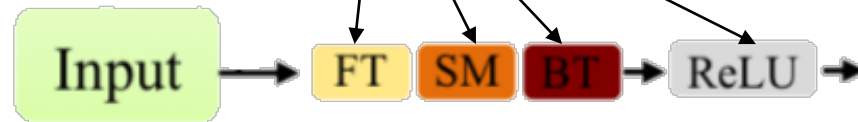
$$\widehat{(f \star g)} = \hat{f} \cdot \hat{g}$$

# Spectral CNNs

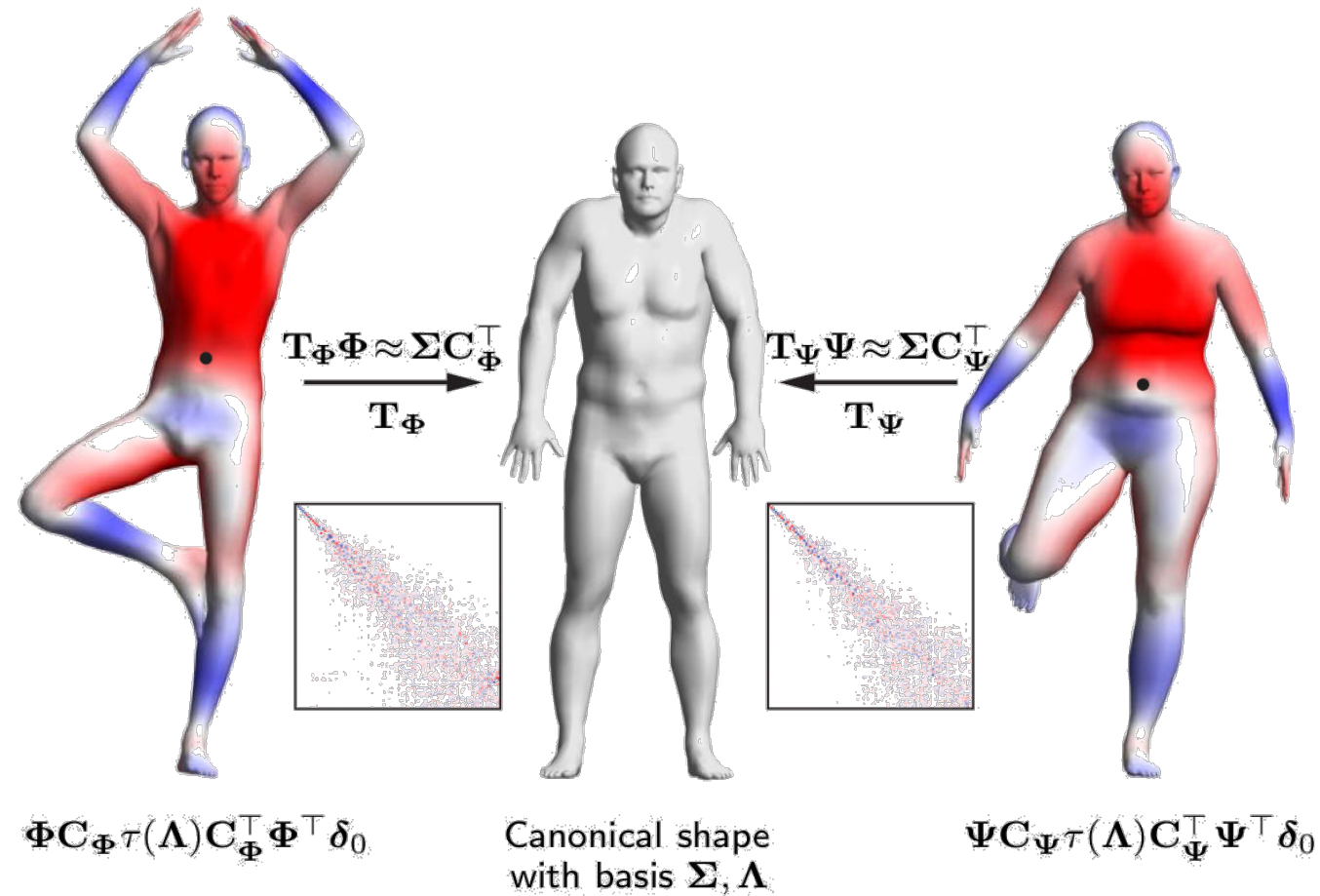
Convolutional layer expressed in the **spectral domain**

$$\mathbf{g}_l = \xi \left( \sum_{l'=1}^p \Phi \hat{\mathbf{W}}_{l,l'} \Phi^\top \mathbf{f}_{l'} \right) \quad \begin{array}{l} l = 1, \dots, q \\ l' = 1, \dots, p \end{array}$$

where  $\hat{\mathbf{W}}_{l,l} = n \times n$  diagonal matrix of filter coefficients

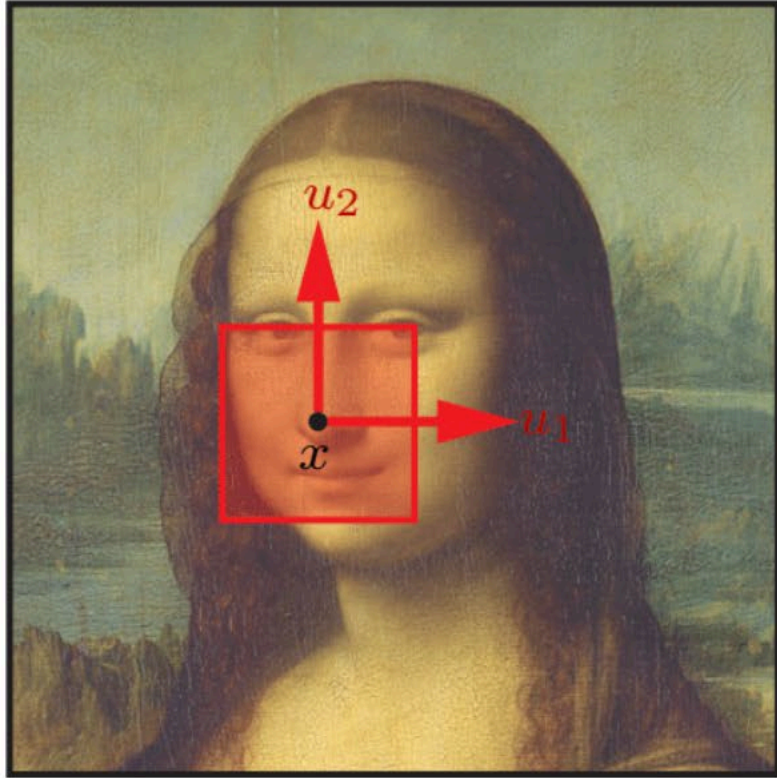


# Spectral Basis Synchronization

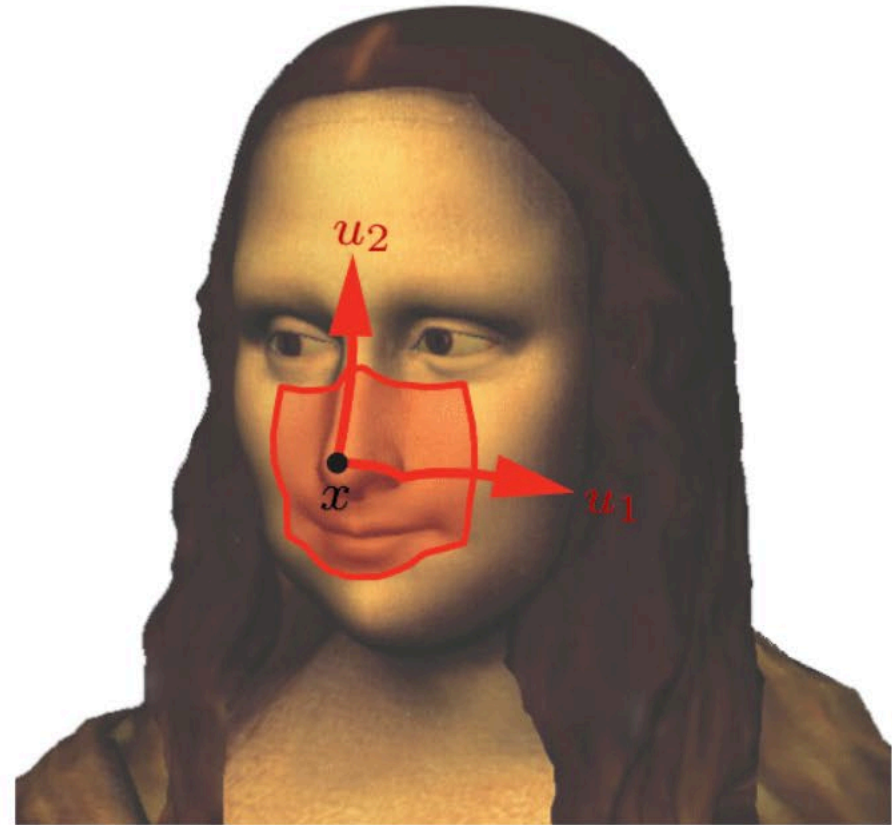


Apply spectral filter  $\tau(\lambda)$  in **synchronized bases**  $\Phi C_\Phi$  and  $\Psi C_\Psi$   
 $\Rightarrow$  **similar results!**

# Primal Methods: Patch Operators



Image

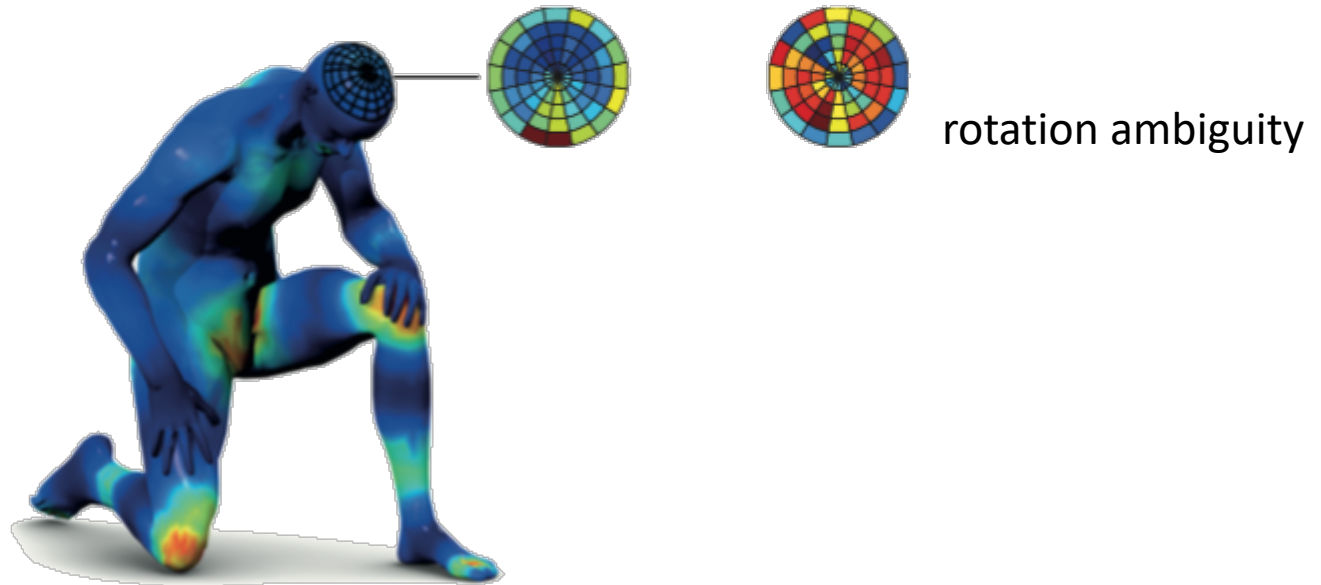


Manifold

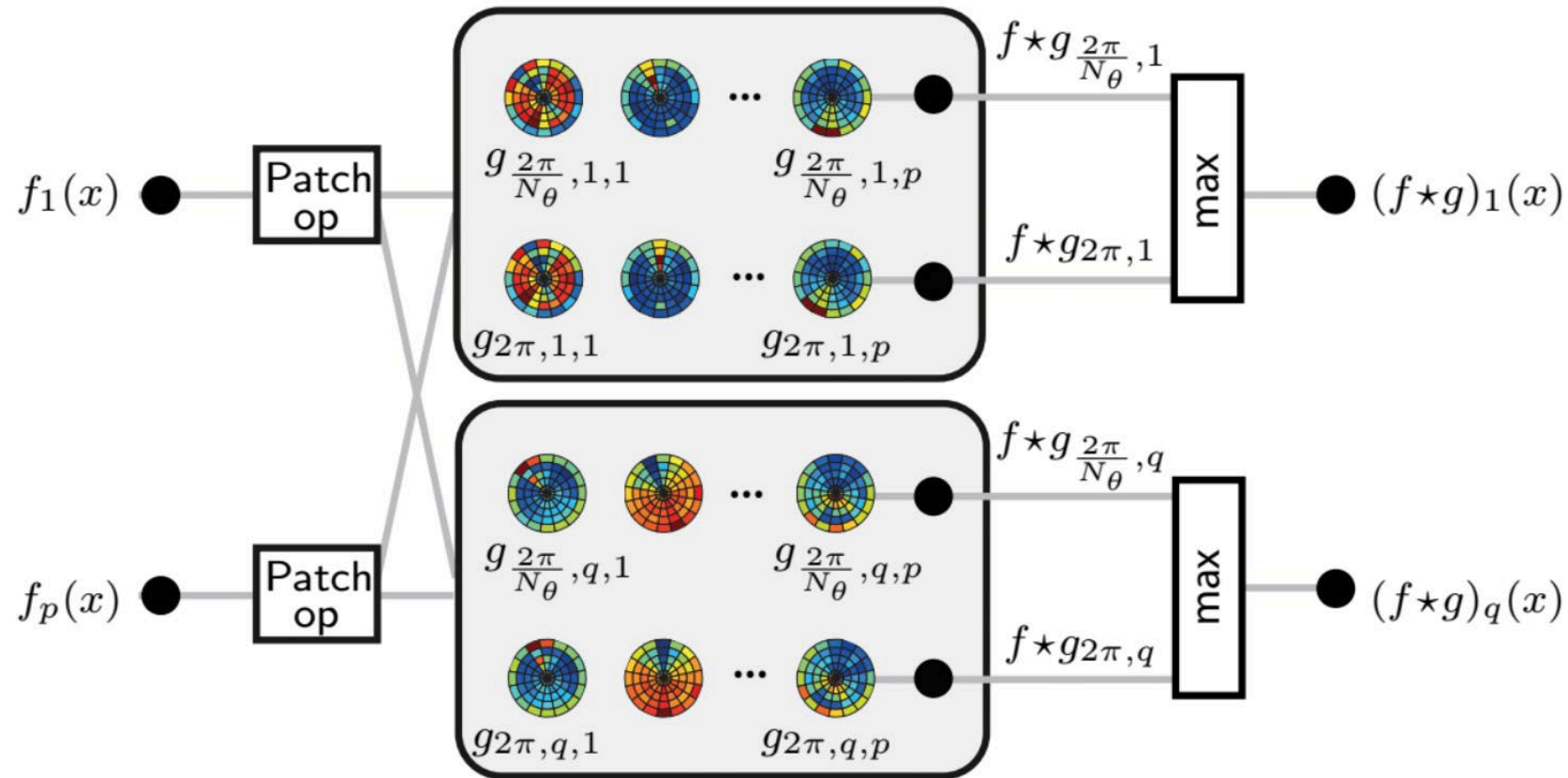
# Geodesic CNNs

- **Geodesic convolution** = apply filter  $a$  to patches extracted from  $f \in L^2(X)$  in local geodesic polar coordinates

$$(f \star a)(x) = \sum_{\theta, r} \underbrace{(D(x)f)(r, \theta)}_{\text{patch}} \underbrace{a(\theta, r)}_{\text{filter}}$$



# Handling Rotation Ambiguity via Pooling



Conv. layer  $(f_l \star g)_{\Delta\theta, l}(x) = \xi \left( \sum_{\ell=1}^p (f_\ell \star g_{\Delta\theta, l, \ell})(x) \right)$

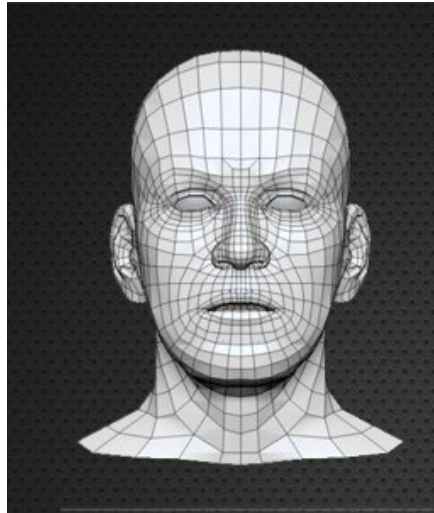
Angular max pooling  $(f \star g)_l(x) = \max_{\Delta\theta} (f \star g)_{\Delta\theta, l}(x)$

Today: Shape Differences and  
Variability --  
Class Wrap-Up

# Object Shape, Material and Appearance Differences



# What Exactly is a Shape Difference?



**vs.**



**vs.**



Making 3D shape differences first class citizens

# Search Engines Based on Differences?

The collage features three main components:

- Technical Diagram:** A line drawing of a shoe's sole and upper with labels: FOXING, VAMP, SPIKE, COLLAR, LINING, and ACHIL PROT.
- YouTube Video:** A video player showing a person holding a white and blue Nike Air Max Torch 4 shoe. The video title is "Nike - Women's Air Max Torch 4 SKU#7938363".
- Product Review Page:** A screenshot of a product page for "Nike Air Max Torch - Womens" (Model: 1135638) with a 4.8 star rating based on 5 reviews. It includes a star rating bar, a "The Good and The Bad" section, and several customer reviews with their dates and locations.

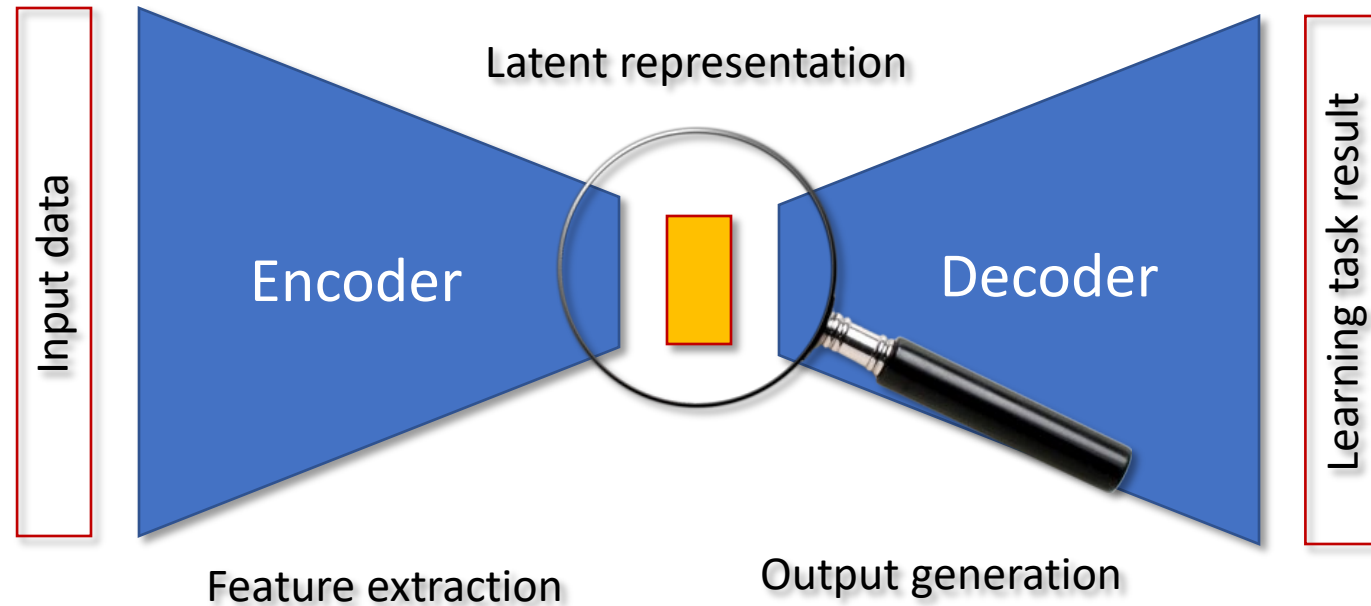
**Star Rating Legend:**

Star Rating	Count
5 star	4
4 star	1
3 star	0
2 star	0
1 star	0

**Customer Reviews for Nike Air Max Torch 4:**

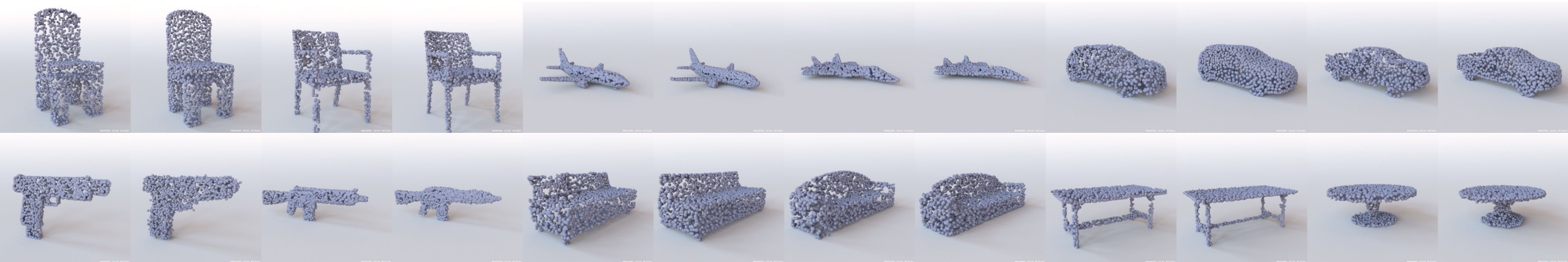
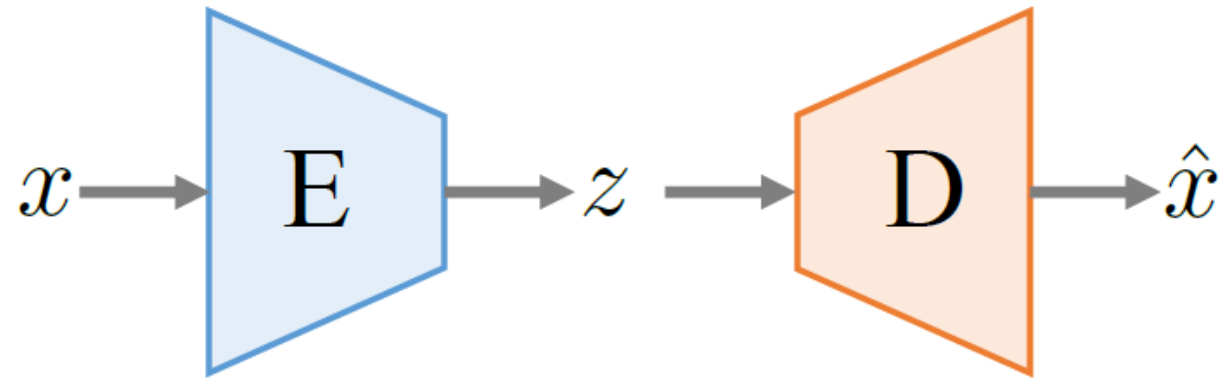
- ATG 7 2009:** "I love these shoes!!" (5 stars)  
"I wear these shoes to school and run in them during gym class and when I work out for soccer. They are also totally cute. -And these are the lowest priced shoes I've ever bought but still better than all the nike shoes I have owned."  
PROS: Breathable, Comfortable, Flexible, Great traction, Lightweight, Love them, Stable  
CONS: None
- OCT 1 2007:** "Good Shoes!" (4 stars)  
"These shoes are very comfortable and perfect for running! Plus the pink and grey colors are really cute too!"  
PROS: Good value
- Review 2 for Nike Air Max Torch 3 Women's Running Shoe:** (5 stars)  
"Extremely comfortable shoes once I bought a 1/2 size bigger."  
"I run 8-15 miles a week. I run primarily to Stay in shape"
- Review 3 for Nike Air Max Torch 3 Women's Running Shoe:** (3 stars)  
"I really like the shoe but unfortunately, I had to send it back because the company had sent me the wrong size shoe. I ordered a size 10 and received a size 11 instead which was waaaaay too big for my foot. I am just waiting to see if I am going to receive a refund or the correct size 10."

# Latent Spaces in ML, Supervised or Not



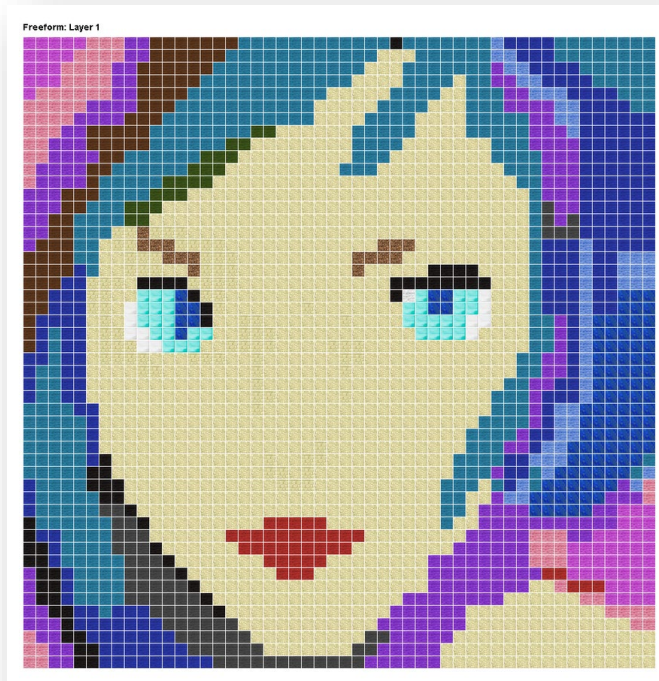
A latent code acts as a low-d proxy for input data w.r.t. a learning task

# Point Cloud Auto-Encoders

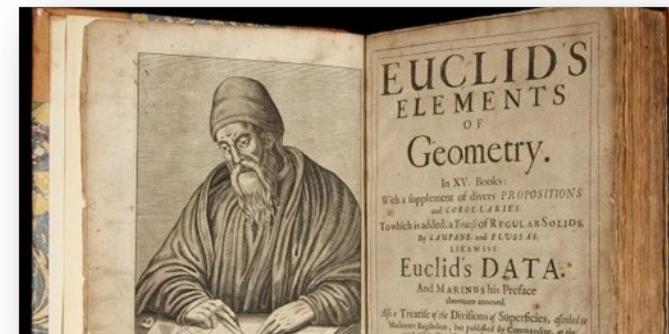
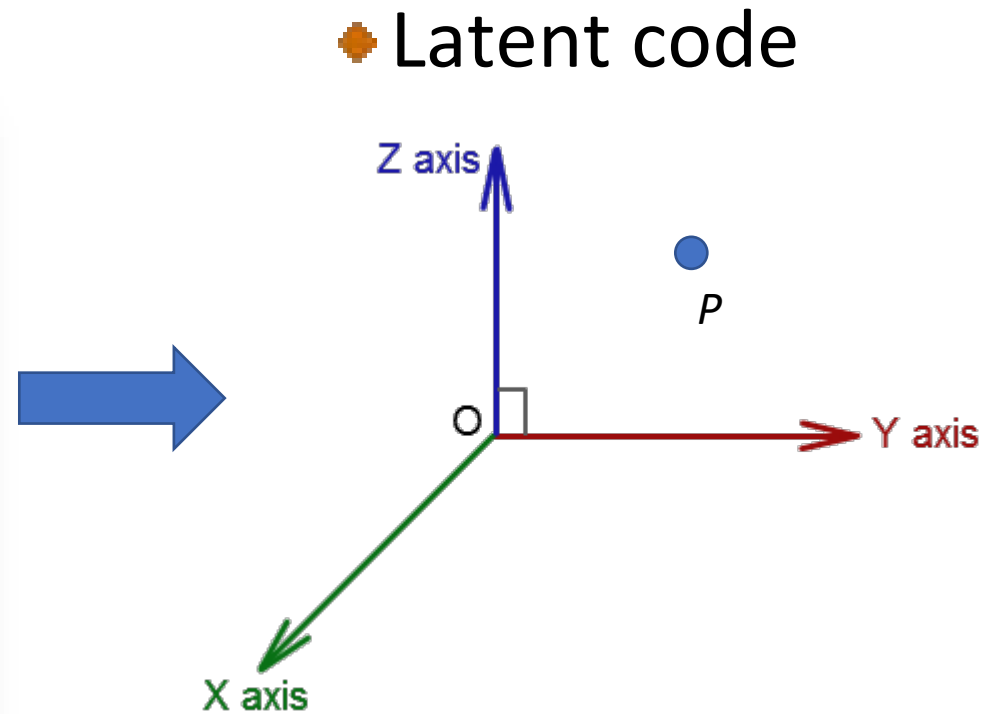


# What Exactly is a Latent Space Representation?

◆ Input

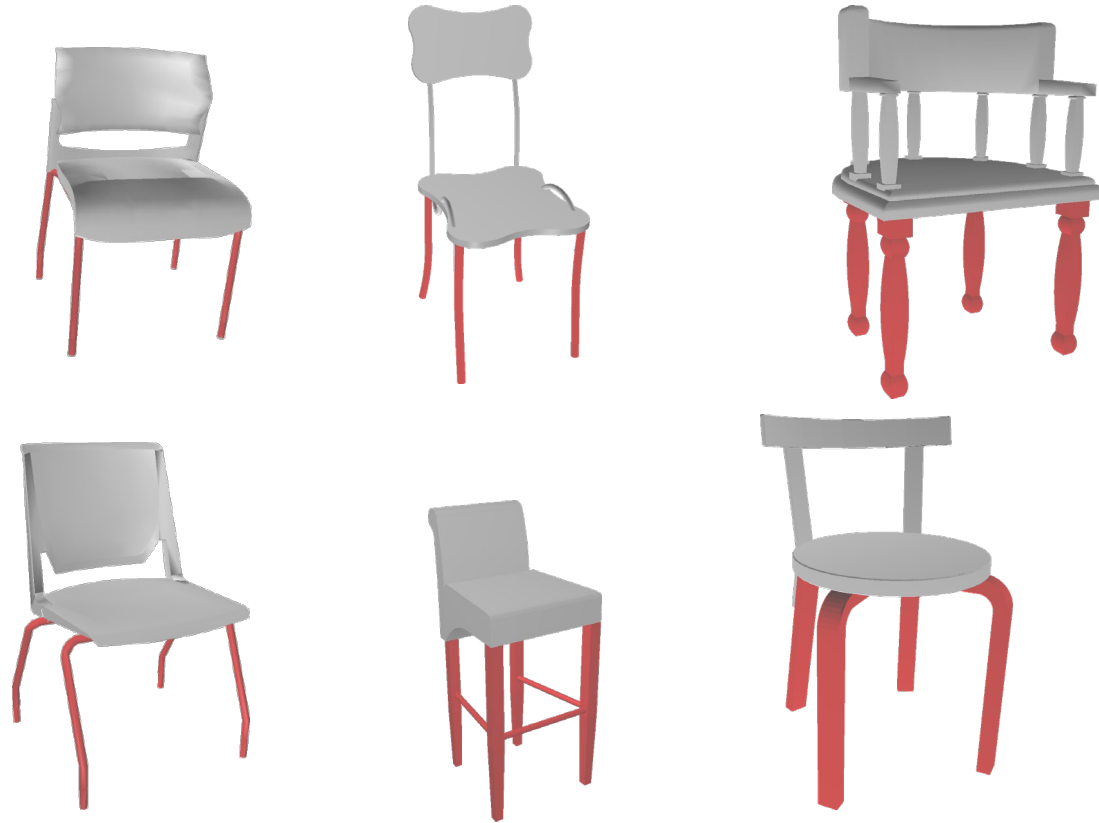


◆ Latent code

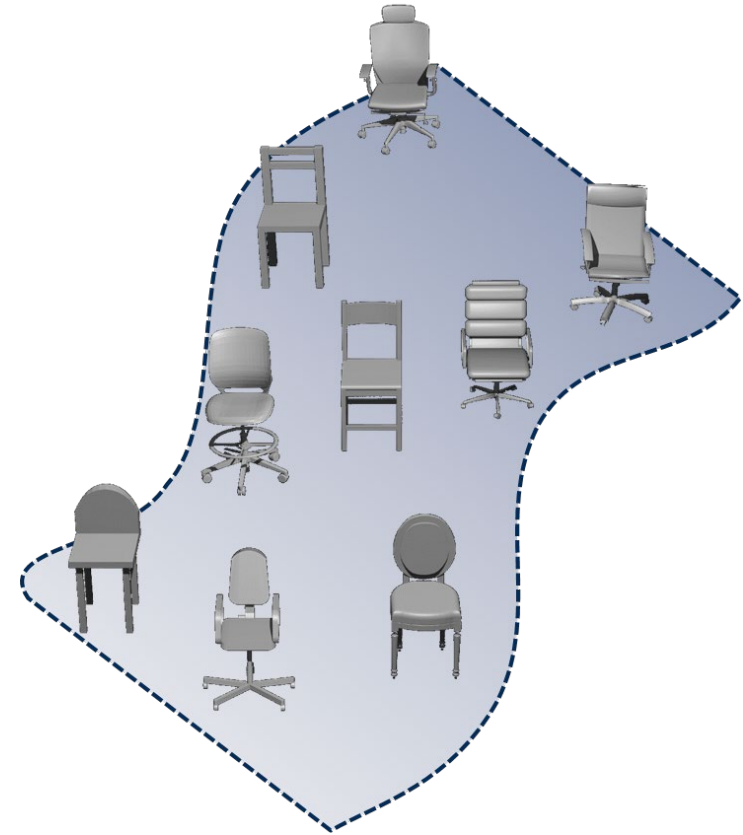


Is a Shape Difference Just a  
Vector in a Latent Space?

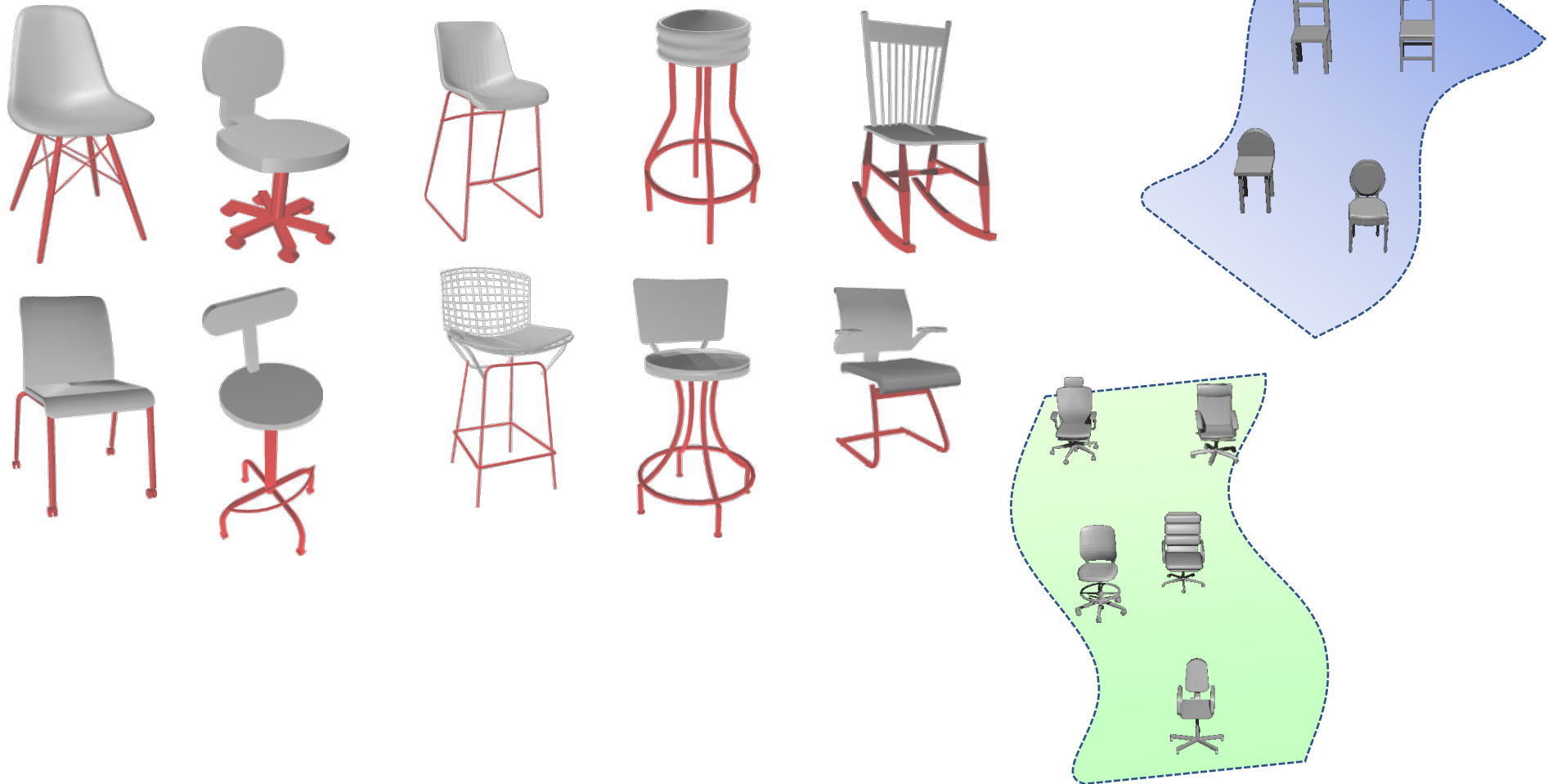
# Continuous Shape Variability



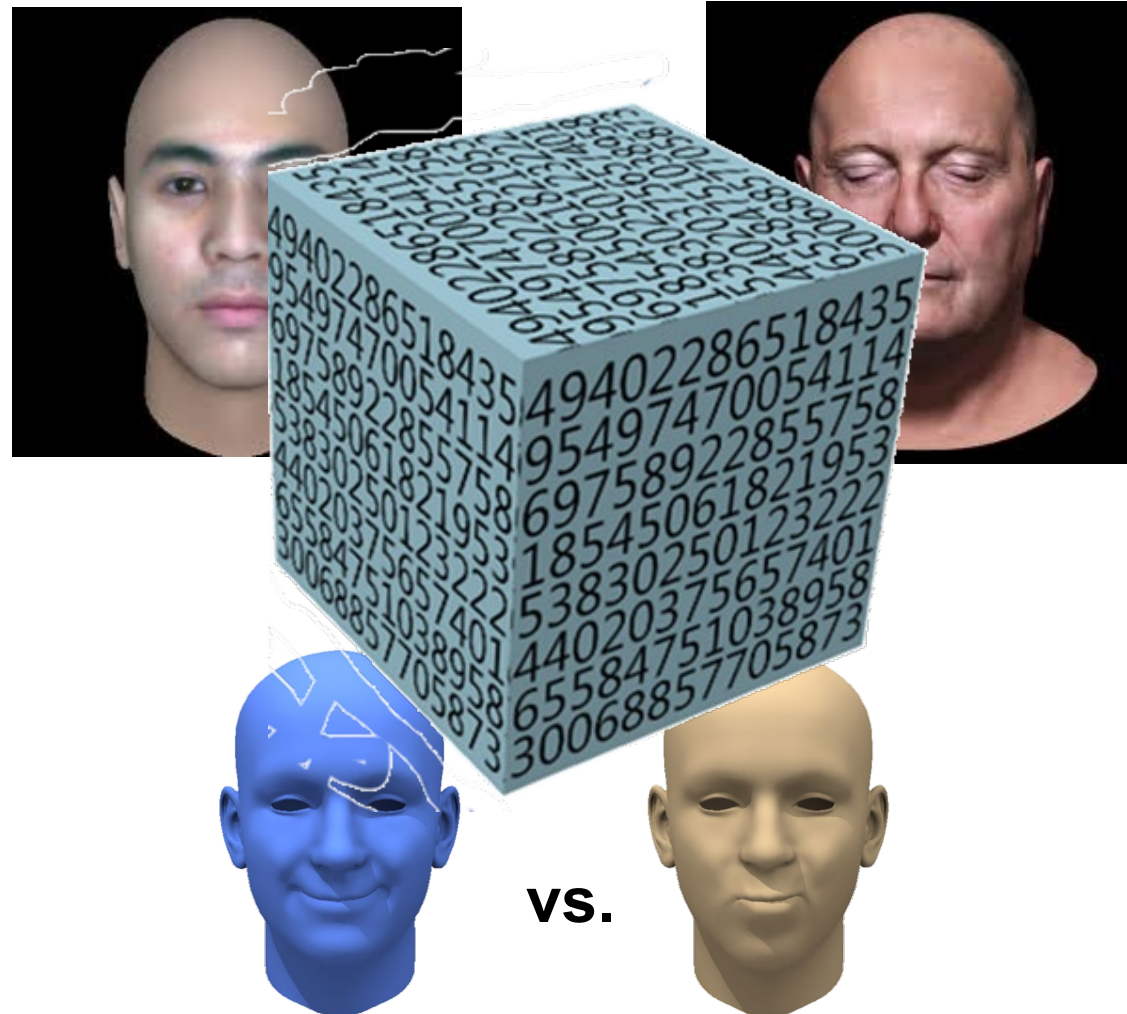
A chair manifold?



# Combinatorial or Discrete Variability

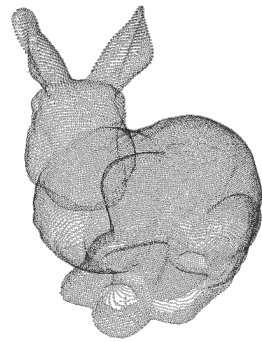


# What Exactly is a Shape Difference?

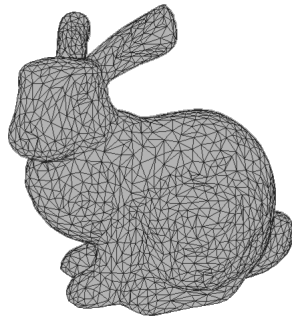


Where and how are the shapes different?

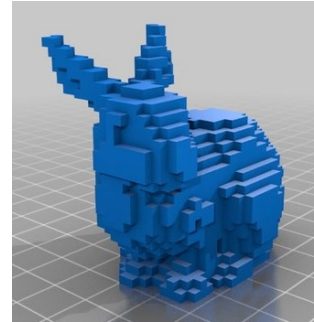
# A Challenge: Multiple 3D Representations



Point Cloud



Mesh



Volumetric



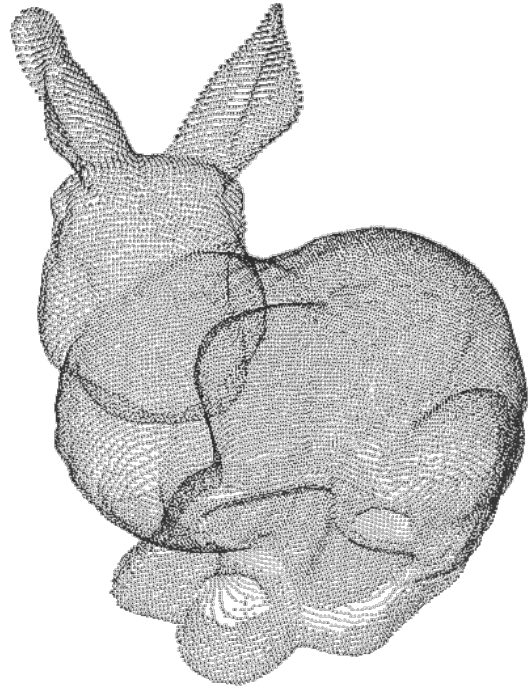
Projected View  
RGB(D)

...

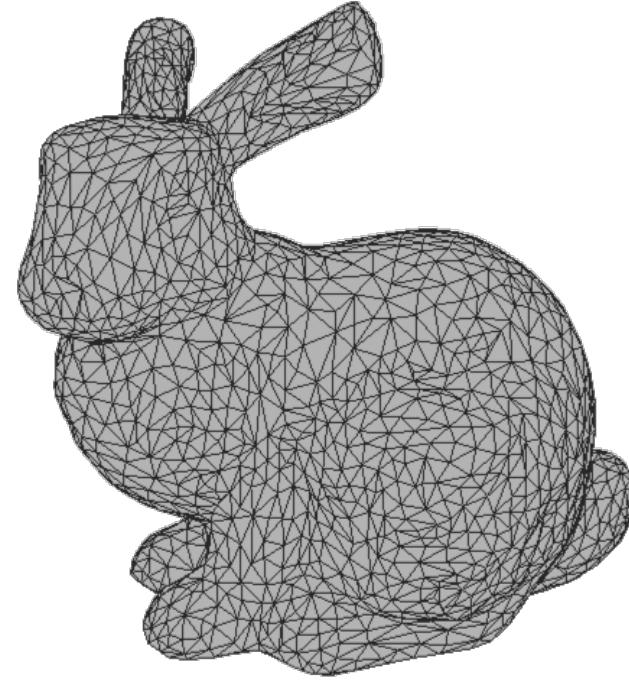


Irregular representations such as point clouds or meshes are a challenge for machine learning algorithms

# Underlying Shape Surface Discretizations



Point cloud



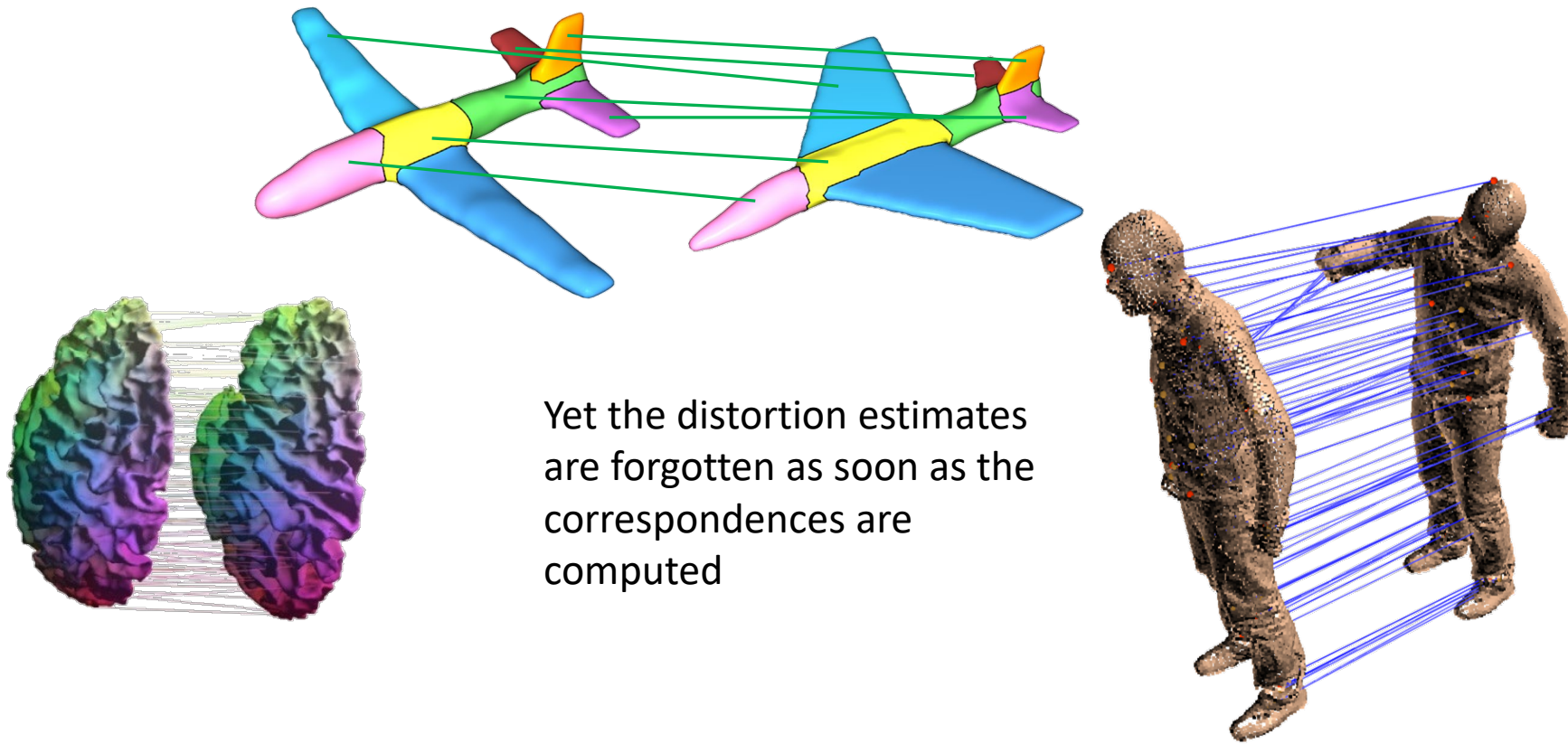
Mesh

Must distinguish differences of the representations from differences of the shapes

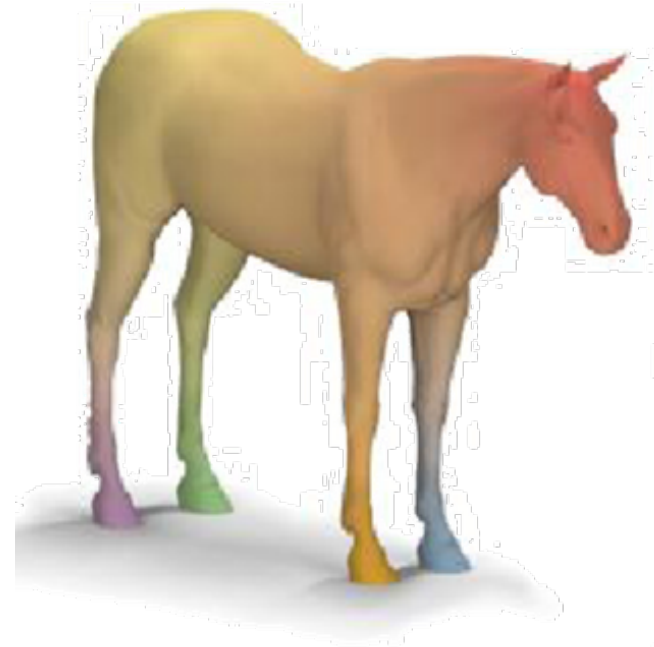
# Continuous Shape Differences Under a Map

# Surface Maps and Distortions

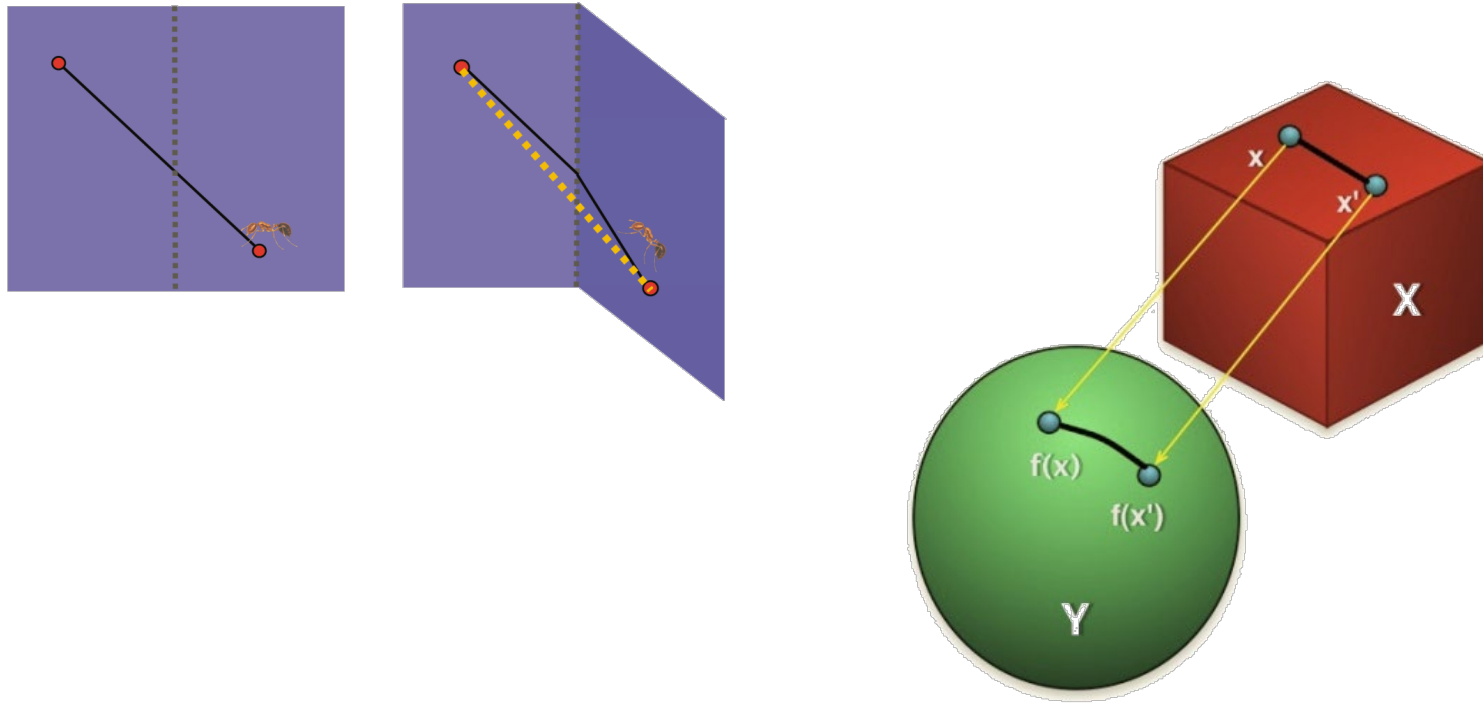
- Shape correspondences
- Often computed by minimizing some measure of distortion



# Subtlety 1: Correspondences at Multiple Scales



# Subtlety 2: Intrinsic or Extrinsic Distances

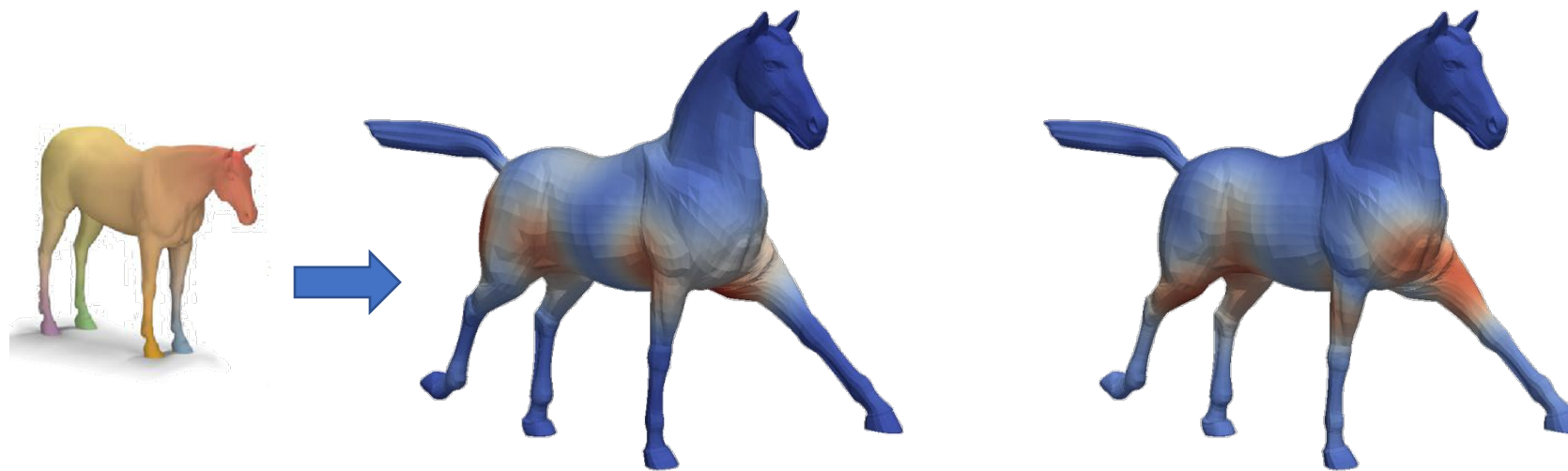
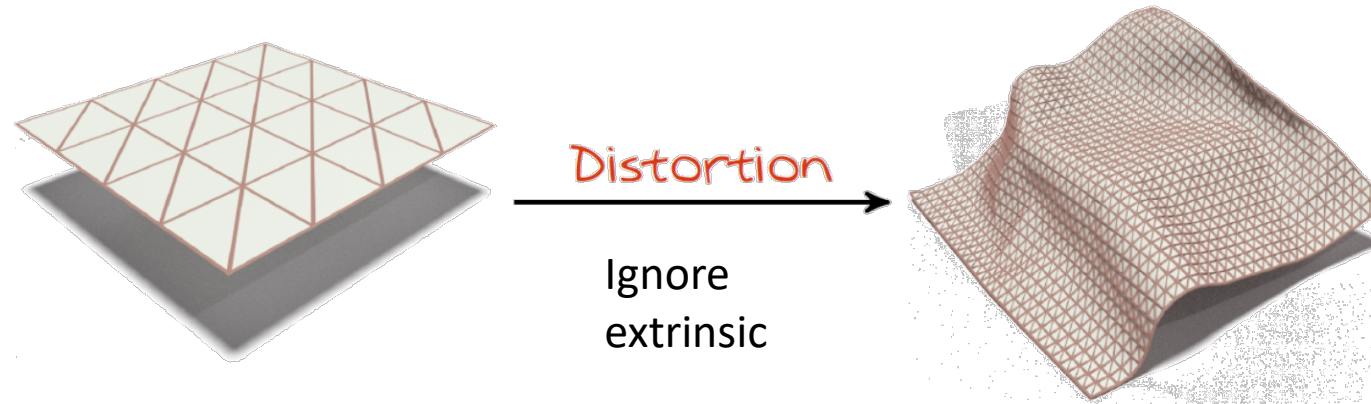


Measure on the surface or in the ambient space?

# Shape Differences from Intrinsic Distortions

[R. Rustomov, M. Ovsjanikov, O. Azercot, M. Ben-Chen, F. Chazal, L. Guibas; Siggraph '13]

# Intrinsic Changes to a Metric



Intrinsic distortions:

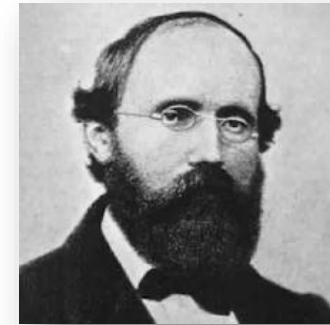
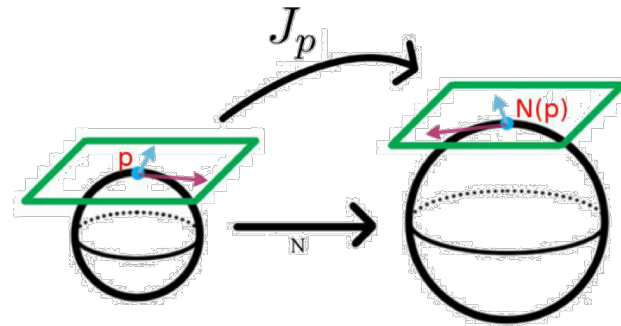
Area distortion

Conformal distortion

Length ...

# Classical Approach to Relating Intrinsic Metrics

To measure distortions induced by a map, we track how inner products of **vectors** change after transporting



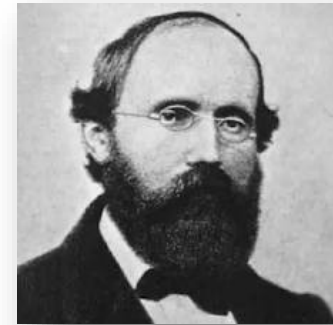
Riemann

## Challenges:

- point-wise information only, hard to aggregate
- noisy

# A Functional View of Distortions

To measure distortions induced by a map, track how inner products of **vectors** change after transporting.



Riemann

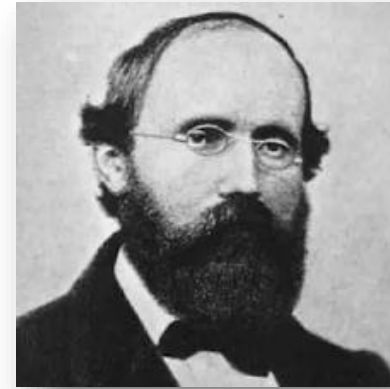
To measure distortions induced by a map, track how inner products of **functions** change after transporting.

# The Art of Measurement

- A metric is defined by a **functional** inner product

$$h^M(f, g) = \int_M f(x)g(x)d\mu(x)$$

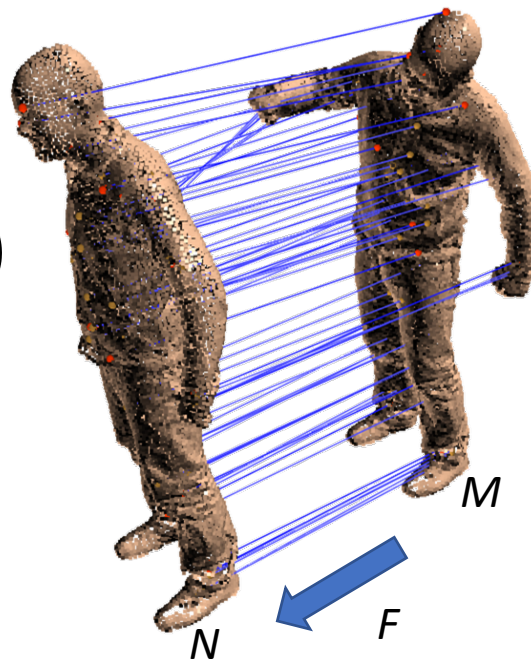
- So we can compare  $M$  and  $N$  by comparing



Riemann

$$h^N(F(f), F(g))$$

The functional map  $F$  transports these functions to  $N$ , where we repeat this measurement with the inner product  $h^N(F(f), F(g))$



$$h^M(f, g)$$



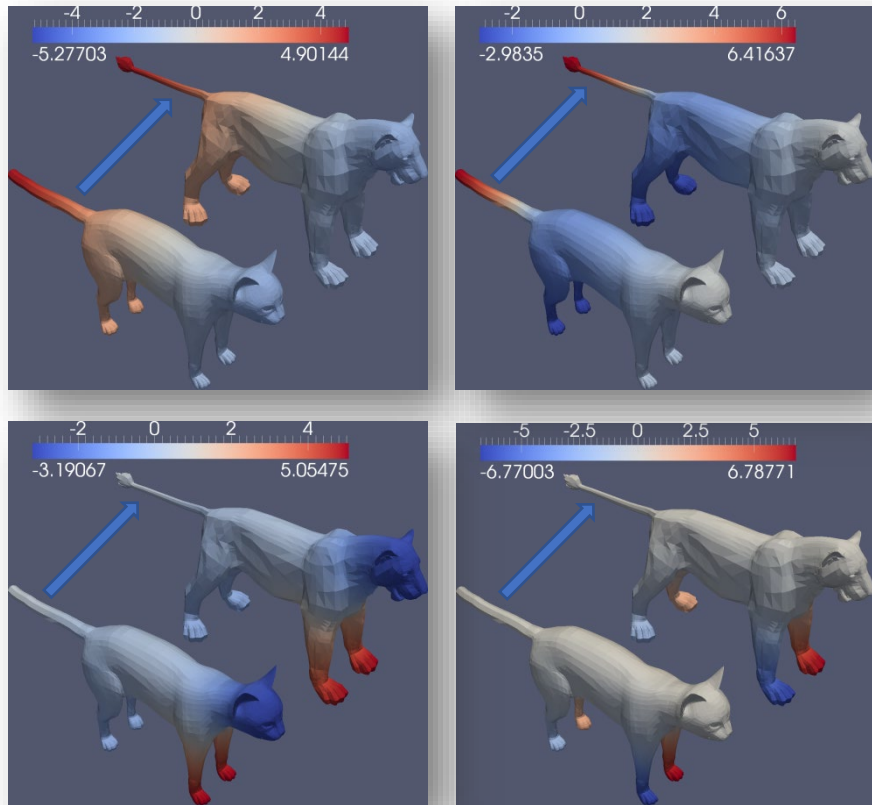
# Inner Products of Functions

$$\begin{aligned}\langle f, g \rangle &= \left\langle \sum_i f_i h_i(x), \sum_j g_j h_j(x) \right\rangle \\ &= \sum_{ij} f_i g_j \langle h_i(x), h_j(x) \rangle \\ &= f^\top A g\end{aligned}$$

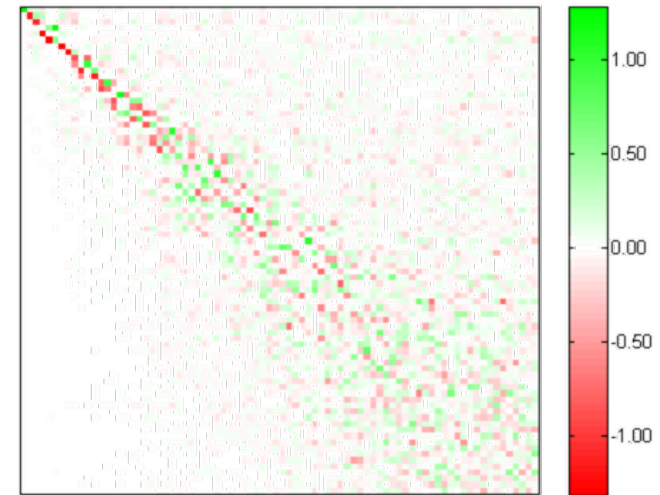
**Area weights matrix**

# Starting from a Functional Map $F$

from cat to lion



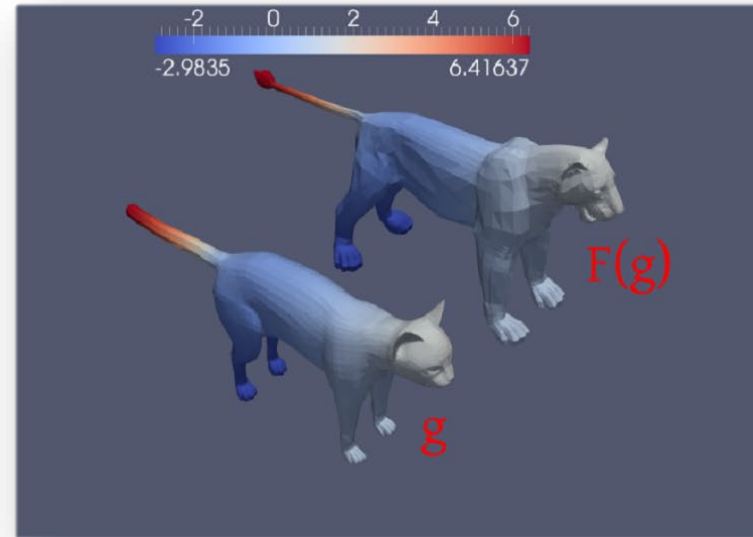
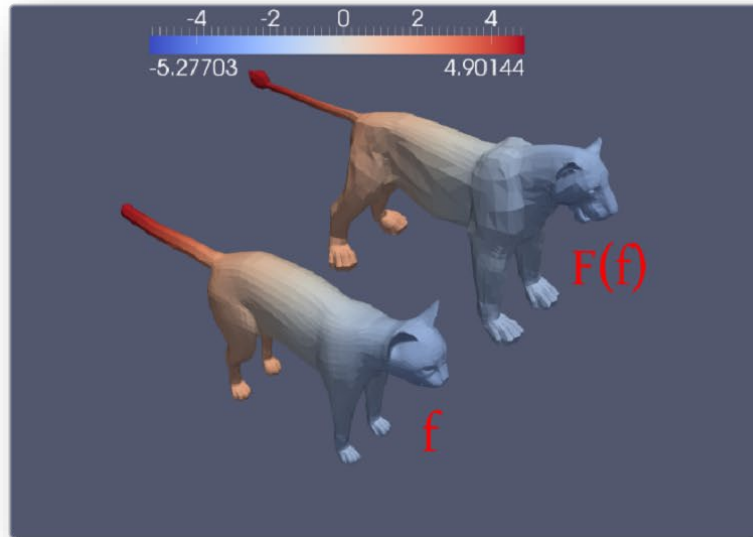
Functions on cat are transferred to lion using  $F$



$F$  is a linear operator (matrix)

$$F : L^2(\text{cat}) \rightarrow L^2(\text{lion})$$

# Measurement Discrepancies

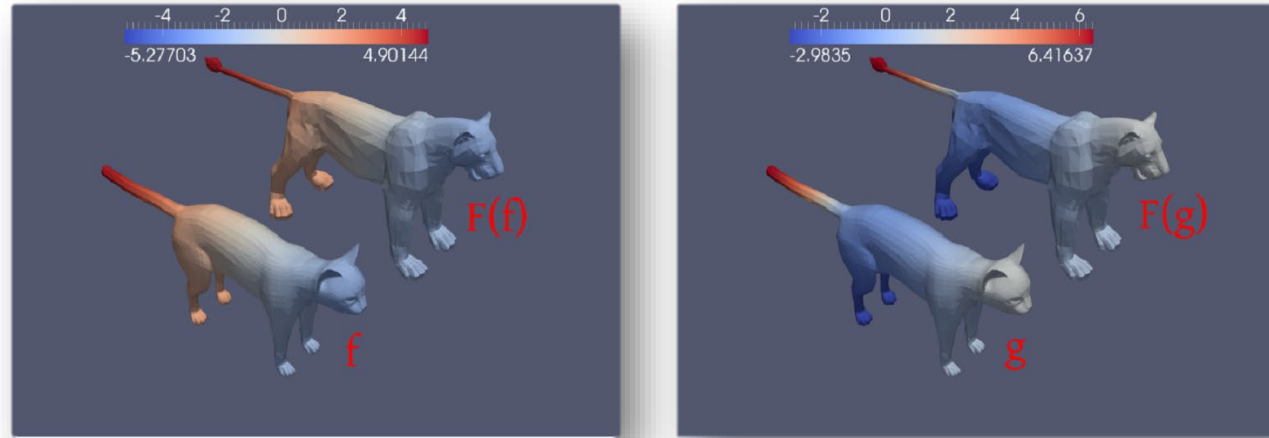


$$\int_{lion} F(f)F(g) d\mu_l \neq \int_{cat} fg d\mu_c$$

after

before

# Measurement Discrepancies



$$\int_{lion} F(f)F(g) d\mu_l \neq \int_{cat} fg d\mu_c$$

after before

**Both** can be considered as inner products on the cat

# The Universal Compensator

Comptes Rendus Hebdomadaires des  
Séances de l'Académie des Sciences de Paris

## Riesz Representation Theorem

There exists a **linear** operator

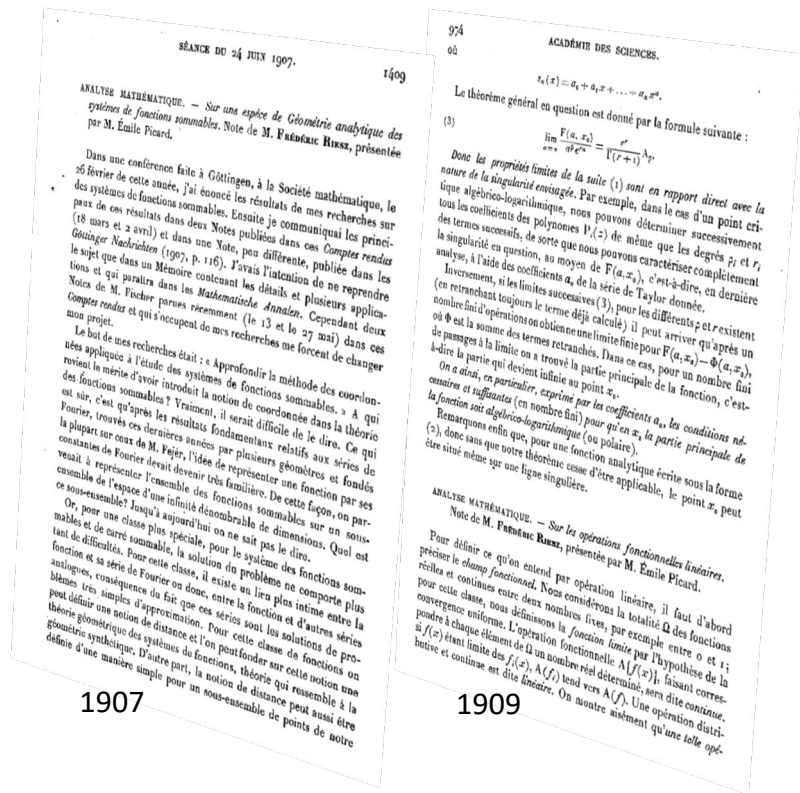
$$V : L^2(\text{cat}) \rightarrow L^2(\text{cat})$$

such that

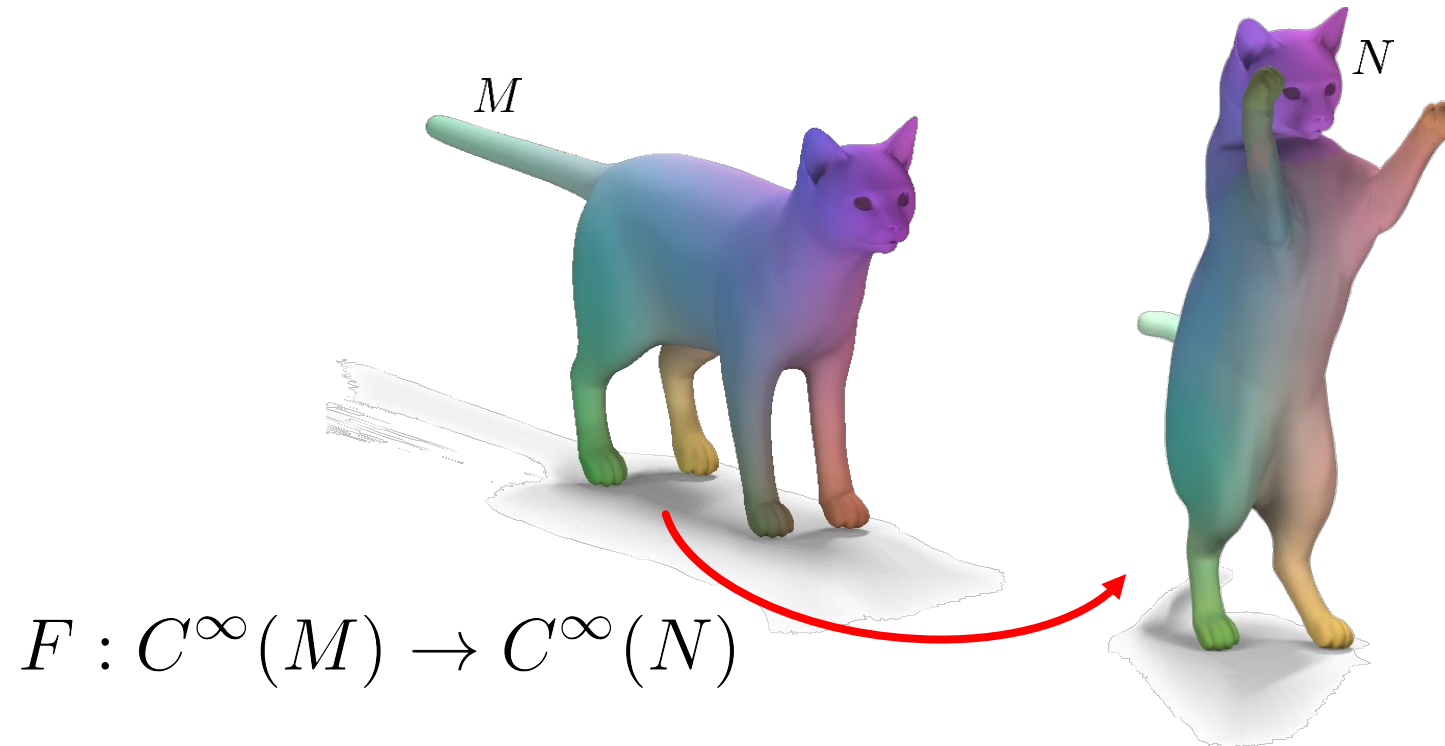
$$\langle f, g \rangle_{\text{after}} = \langle f, V(g) \rangle_{\text{before}}$$



Frigyes Riesz



# Riesz Representation Theorem



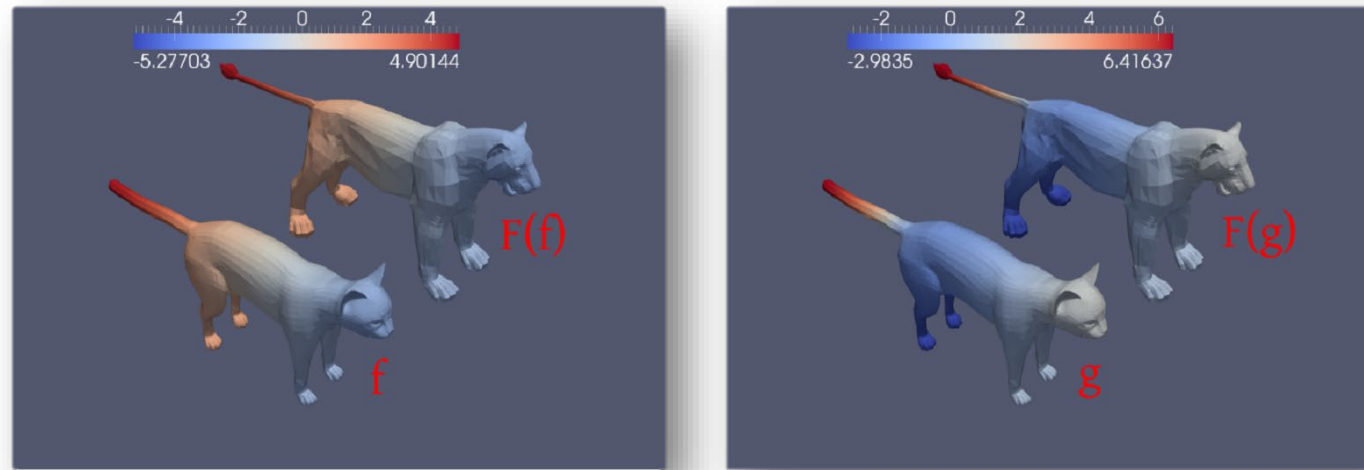
$$\exists V \text{ s.t.} \\ \langle F(f), F(g) \rangle^N = \langle f, V(g) \rangle^M \quad \forall f, g$$

# Sanity Check

$$\begin{aligned}\langle f, g \rangle &\approx f^\top A g \\ \langle F(f), F(g) \rangle^N &\approx [F(f)]^\top A_N [F(g)] \\ &= f^\top \cdot F^\top A_N F \cdot g \\ &= f^\top \cdot (A_M A_M^{-1}) F^\top A_N F \cdot g \\ &= f^\top \cdot A_M (A_M^{-1} F^\top A_N F \cdot g) \\ &\approx \langle f, (A_M^{-1} F^\top A_N F) g \rangle\end{aligned}$$

# Area-Based Shape Difference:

$$V = A_M^{-1} F^T A_N F$$



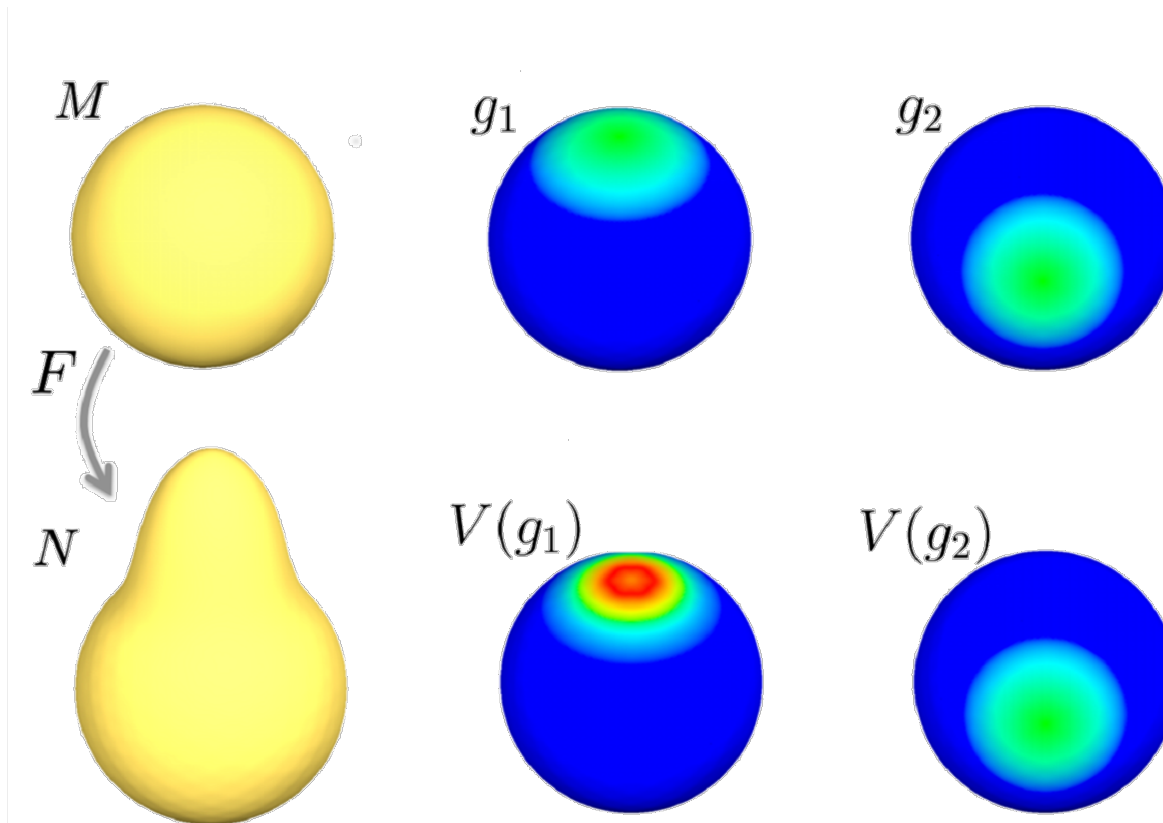
$$\int_{lion} F(f)F(g) \neq \int_{cat} fg$$



$$\int_{lion} F(f)F(g) = \int_{cat} fV(g)$$

$V$  maps functions on the cat to functions on the cat -- is a self-map of the domain

# A Small Example of $V$



Note that  $V$  maps functions on  $M$  to functions on  $N$

$$\int_N F(f)F(g) = \int_M fV(g)$$

# Conformal Shape Difference $R$

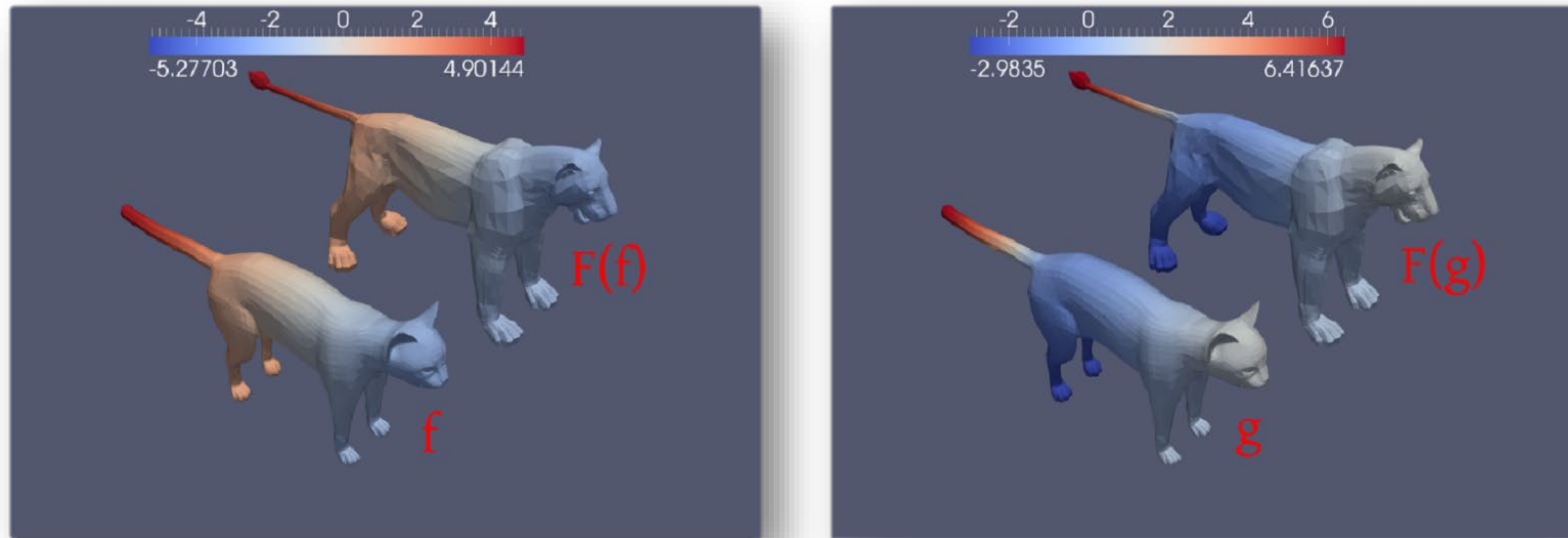
Consider a different inner-product of functions ...

get information about **conformal** distortion

$$\int_N \nabla F(f) \nabla F(g) = \int_M \nabla f \nabla R(g)$$

The choice of inner product should be driven by the application at hand.

# Conformal Shape Difference $R$

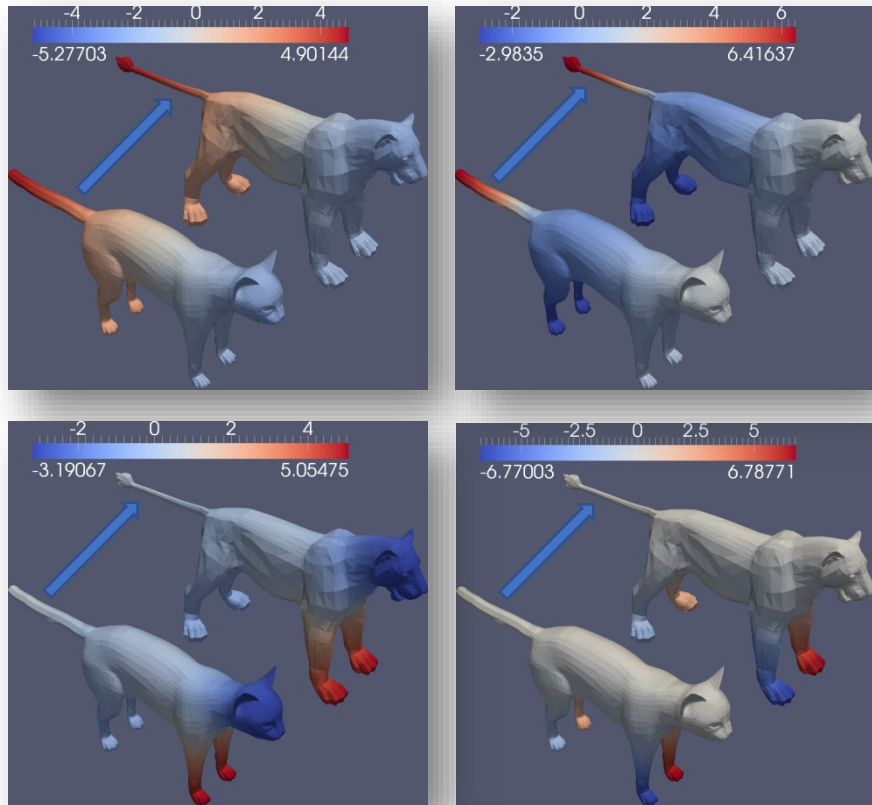


$$\int_{lion} \nabla F(f) \cdot \nabla F(g) \neq \int_{cat} \nabla f \cdot \nabla g \quad \forall f, g$$

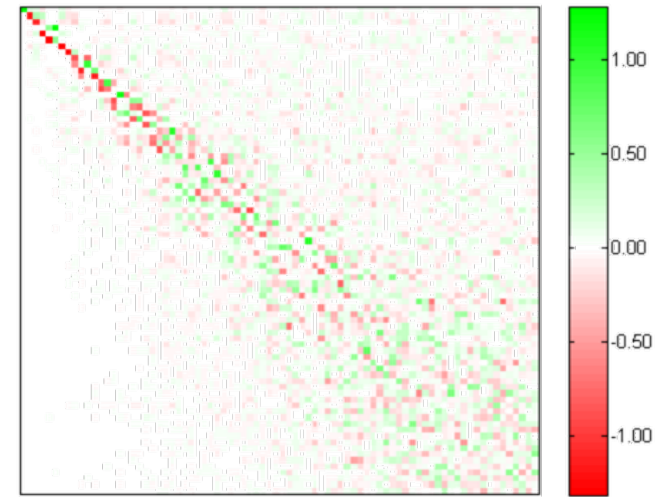
$$\int_{lion} \nabla F(f) \cdot \nabla F(g) \stackrel{\downarrow}{=} \int_{cat} \nabla f \cdot \nabla R(g)$$

# Input: Functional Map $F$

from cat to lion



Functions on cat are transferred to lion using  $F$

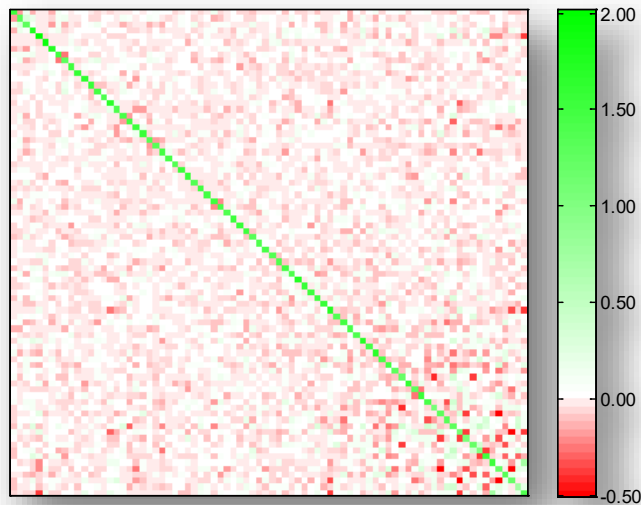


$F$  is a linear operator (matrix)

$$F : L^2(\text{cat}) \rightarrow L^2(\text{lion})$$

# Shape Difference Operators

**V** – area-based shape difference

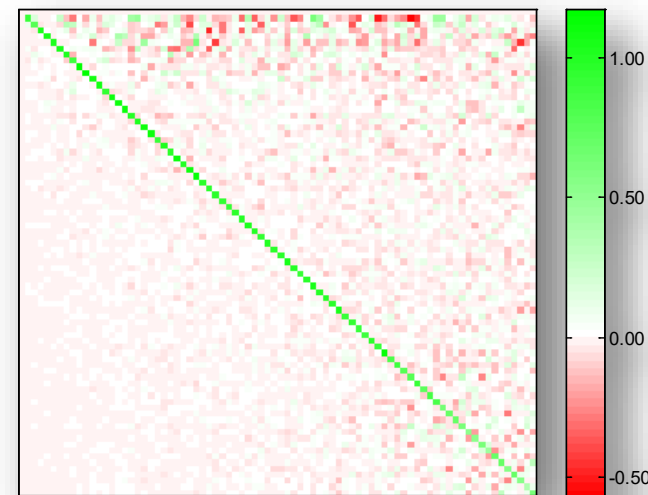


**linear operator (matrix)**

$$V : L^2(\text{cat}) \rightarrow L^2(\text{cat})$$

$$\int_N F(f)F(g) = \int_M fV(g)$$

**R** – conformal shape difference

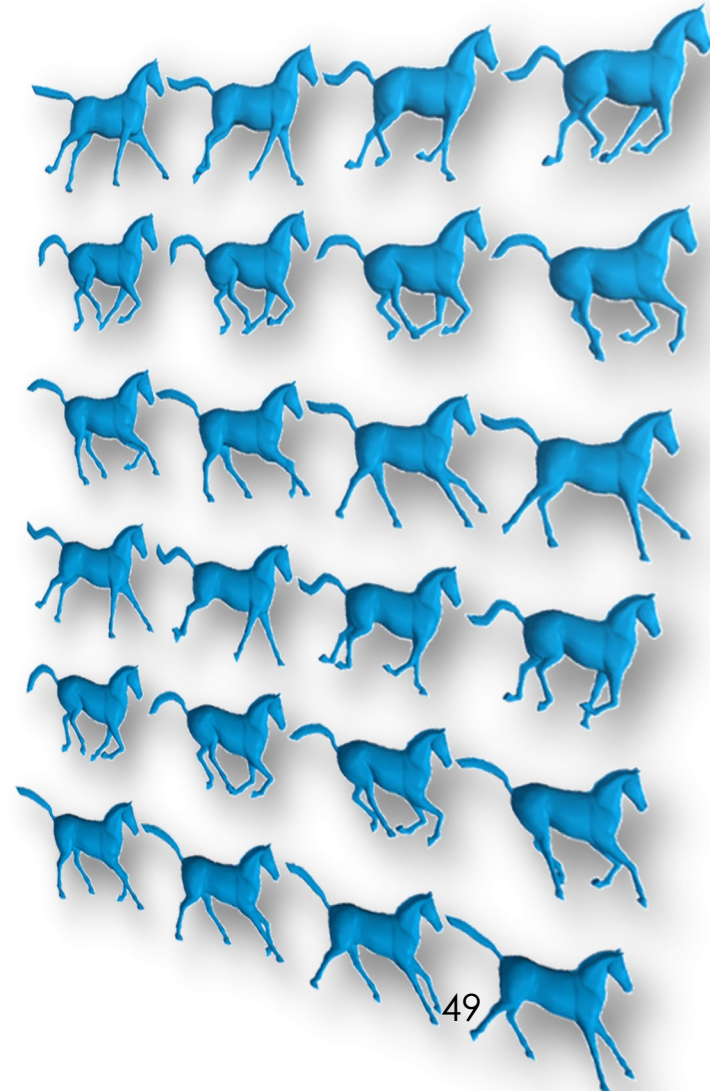
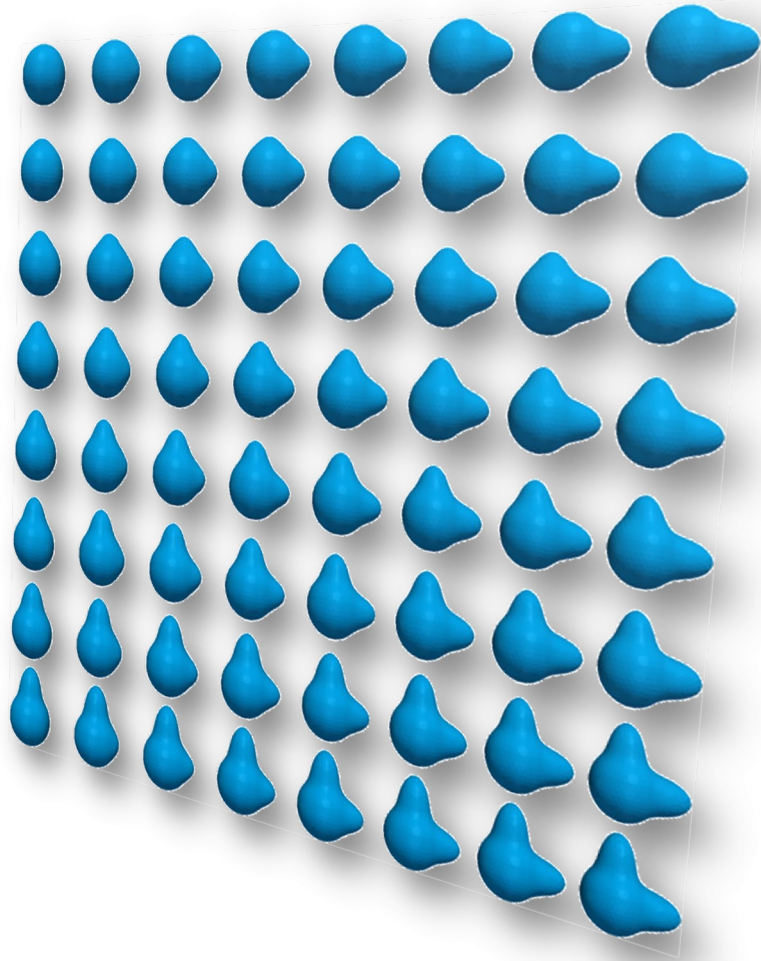


**linear operator (matrix)**

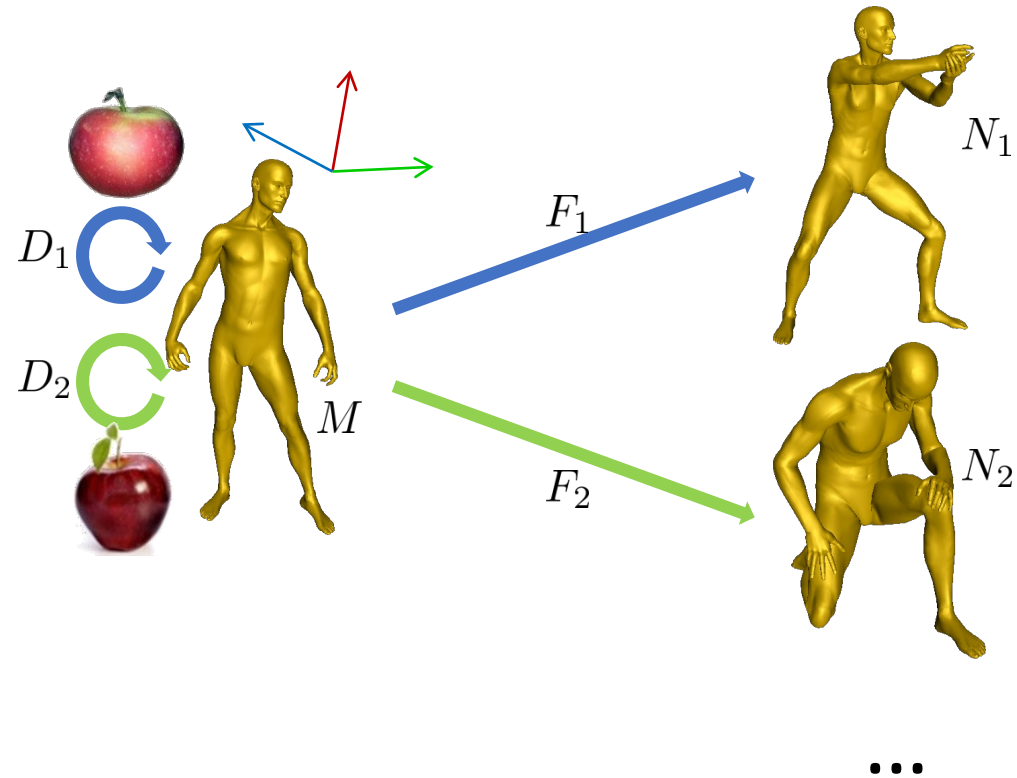
$$R : L^2(\text{cat}) \rightarrow L^2(\text{cat})$$

$$\int_N \nabla F(f)\nabla F(g) = \int_M \nabla f\nabla R(g)$$

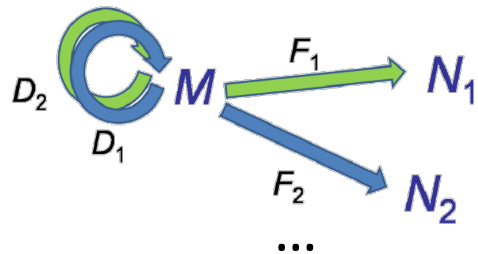
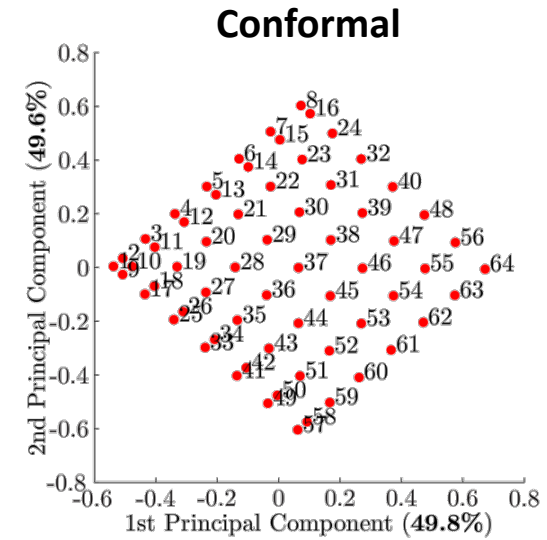
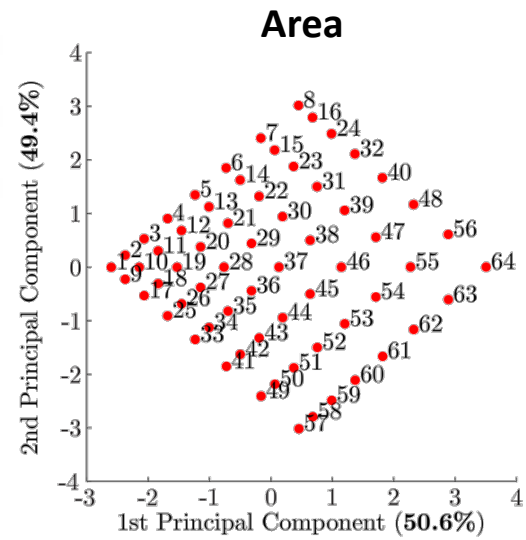
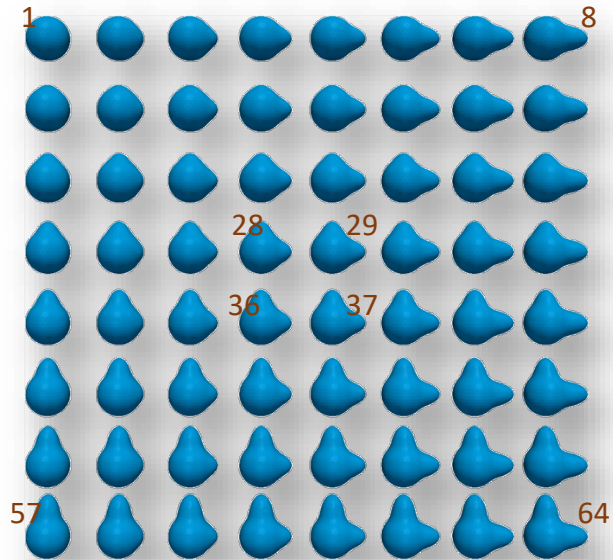
# Shape Differences in Collections



# Comparing Differences

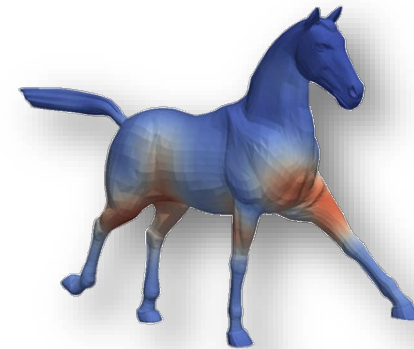
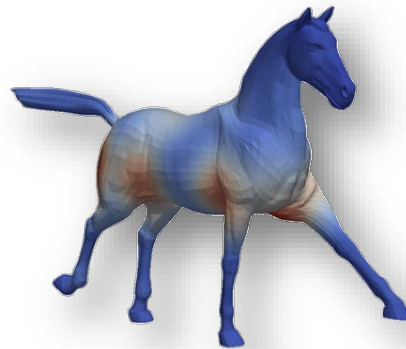
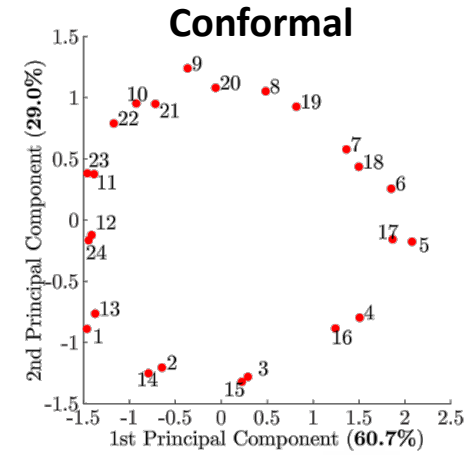
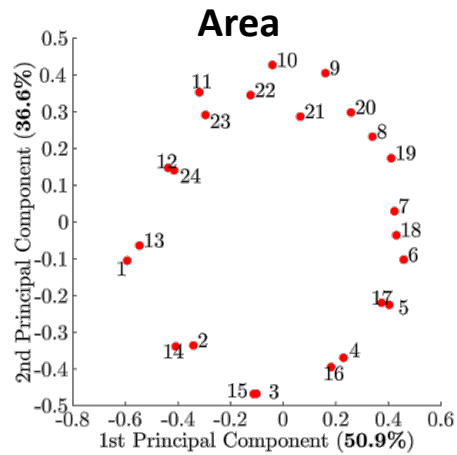
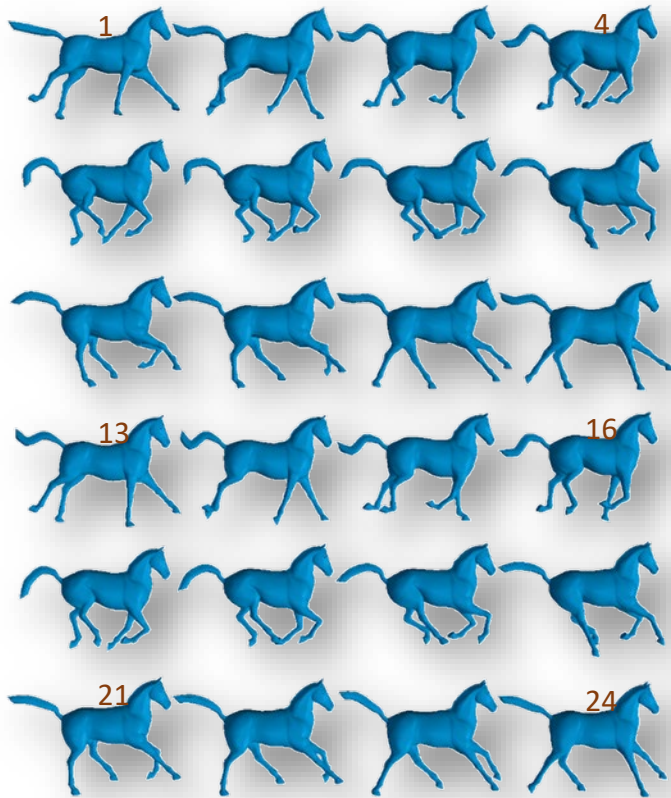


# Intrinsic Shape Space



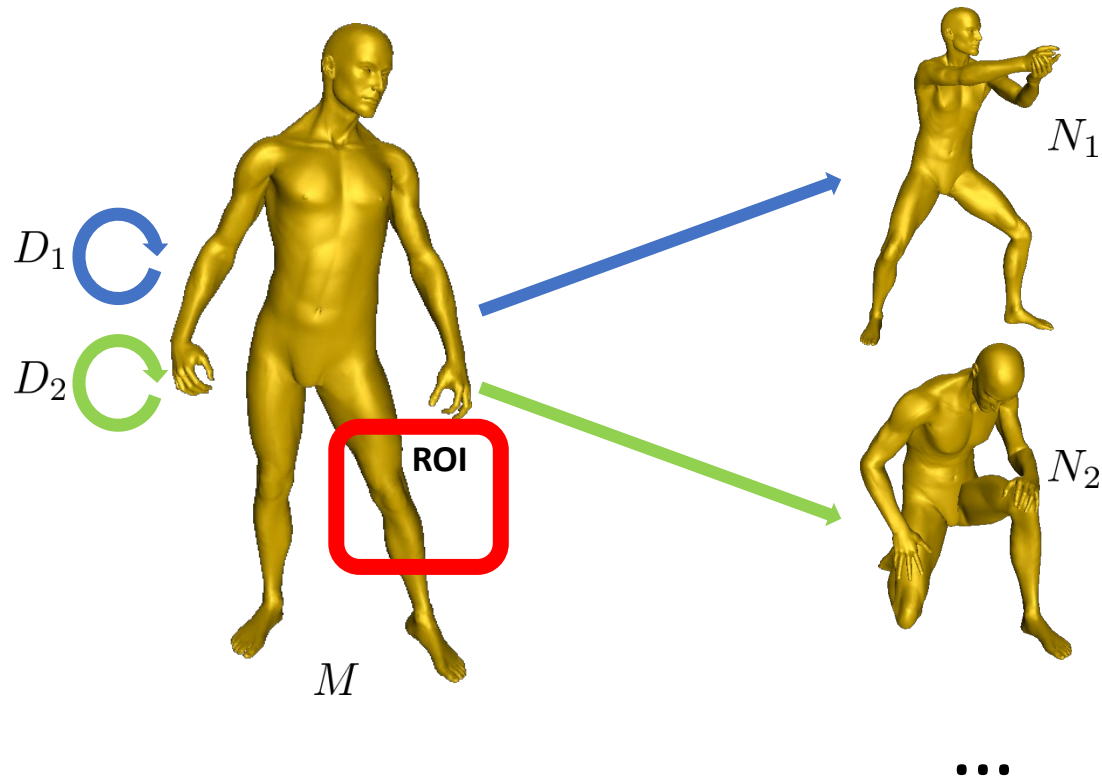
Deformation localization

# Intrinsic Shape Space



Deformation localization

# Localized Comparisons

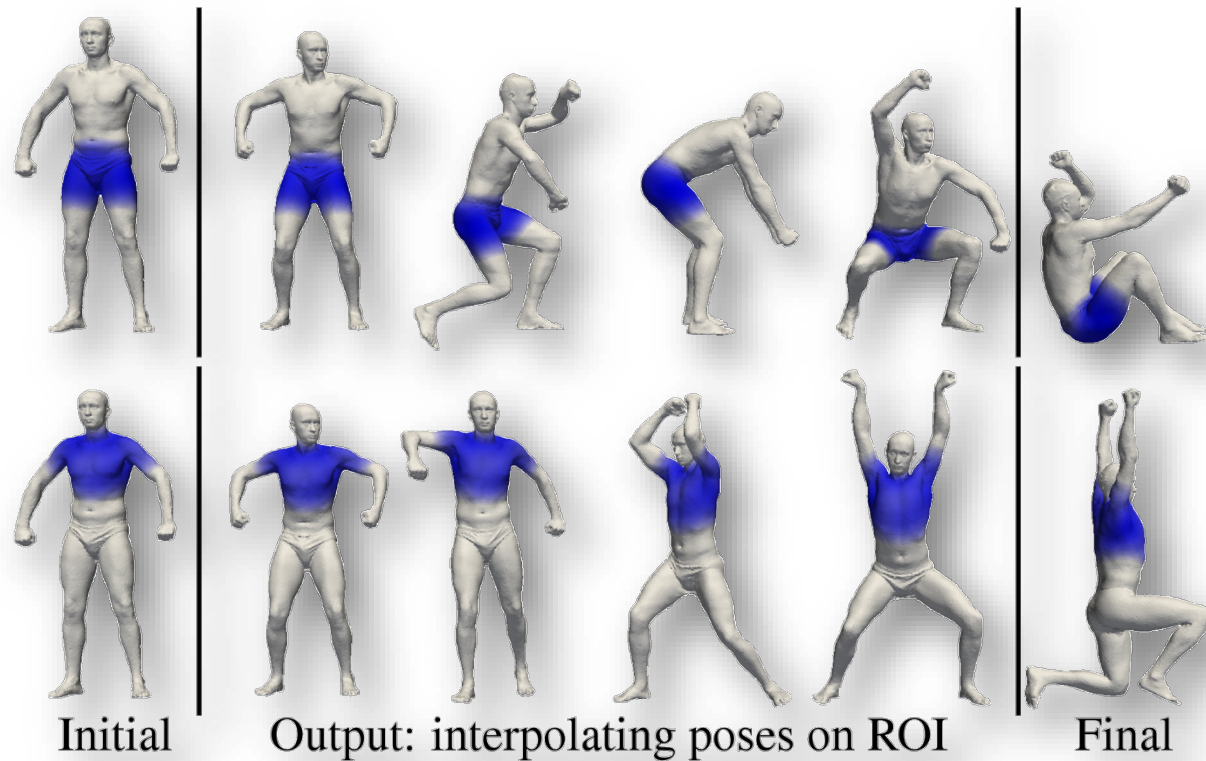


$$\rho : M \rightarrow \mathbb{R}$$

supported in ROI

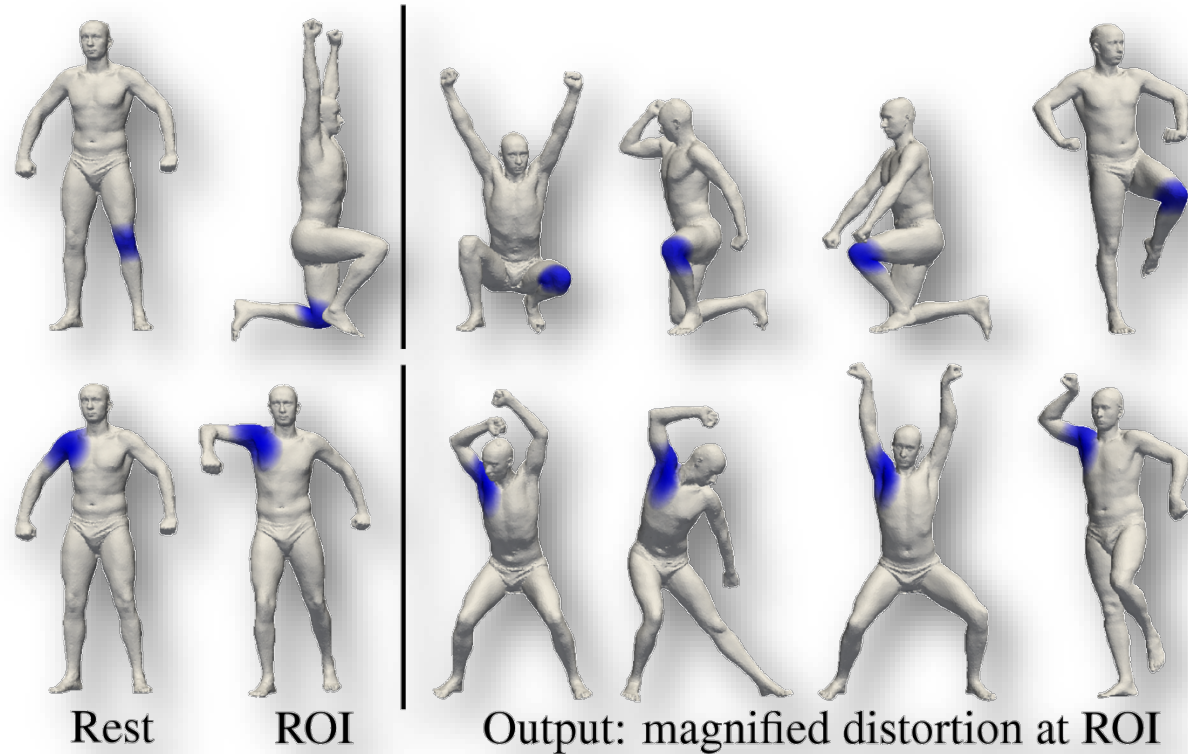
$$D_1\rho \text{ to } D_2\rho$$

# Interpolation Between Poses Along ROI

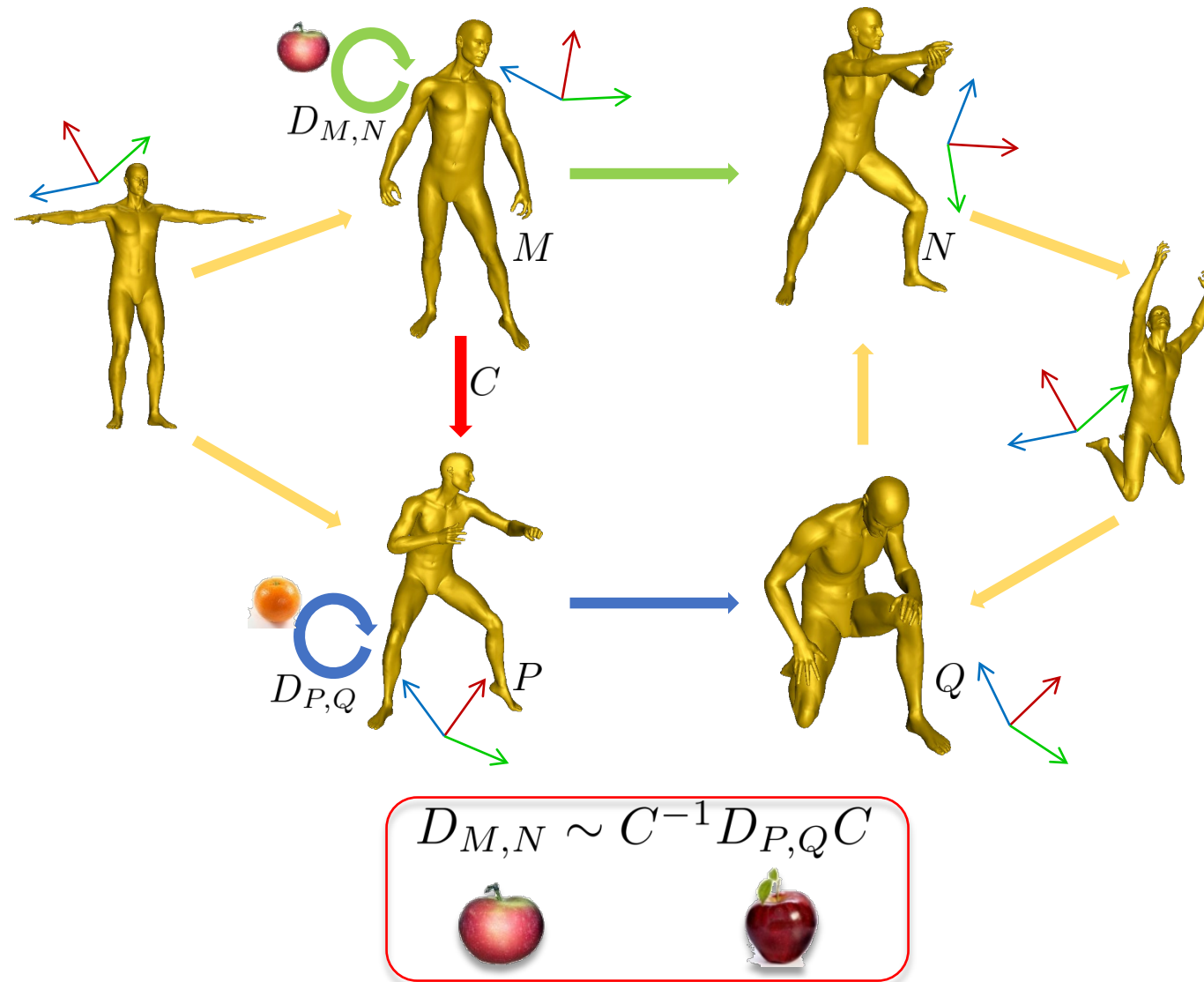


SCAPE Human Shapes

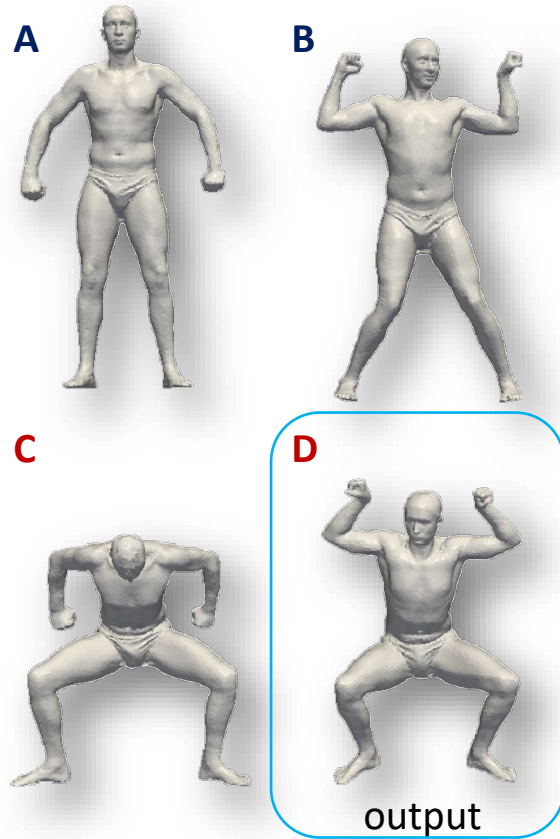
# Exaggeration of Difference in Roi



# Comparing Differences I



# Analogies: D relates to C as B relates to A

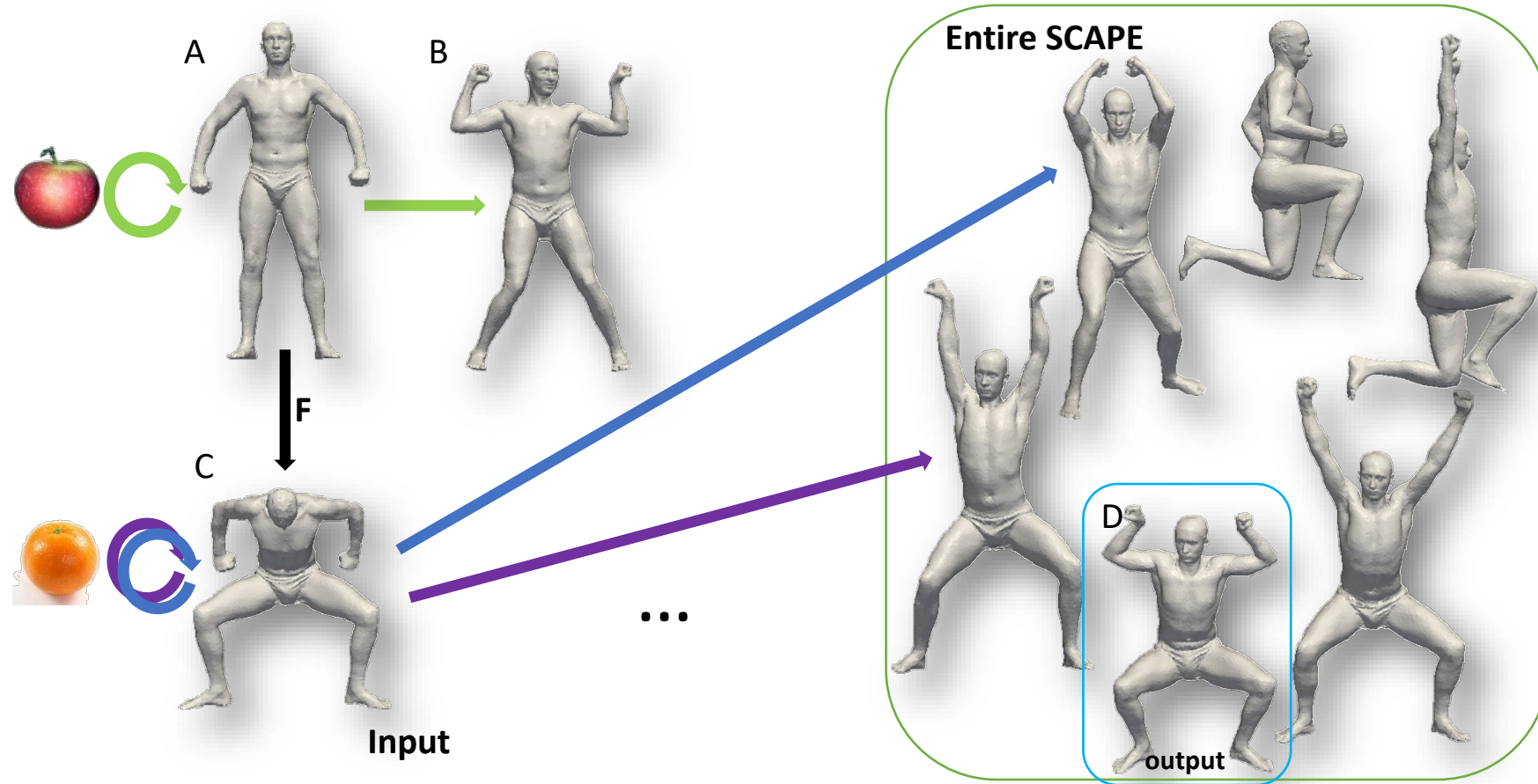


$$D = C + \underbrace{(B - A)}_{\text{hands raised up}}$$

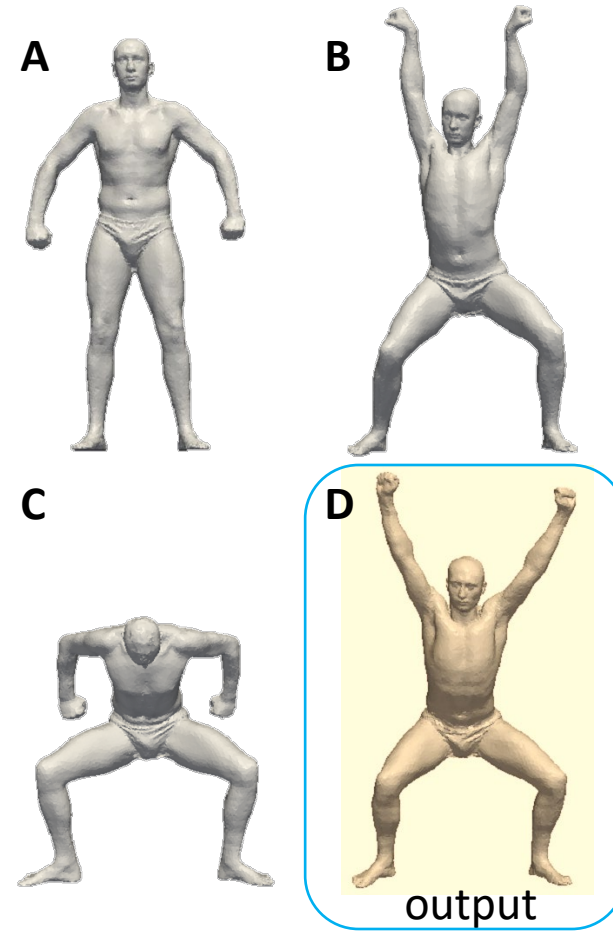
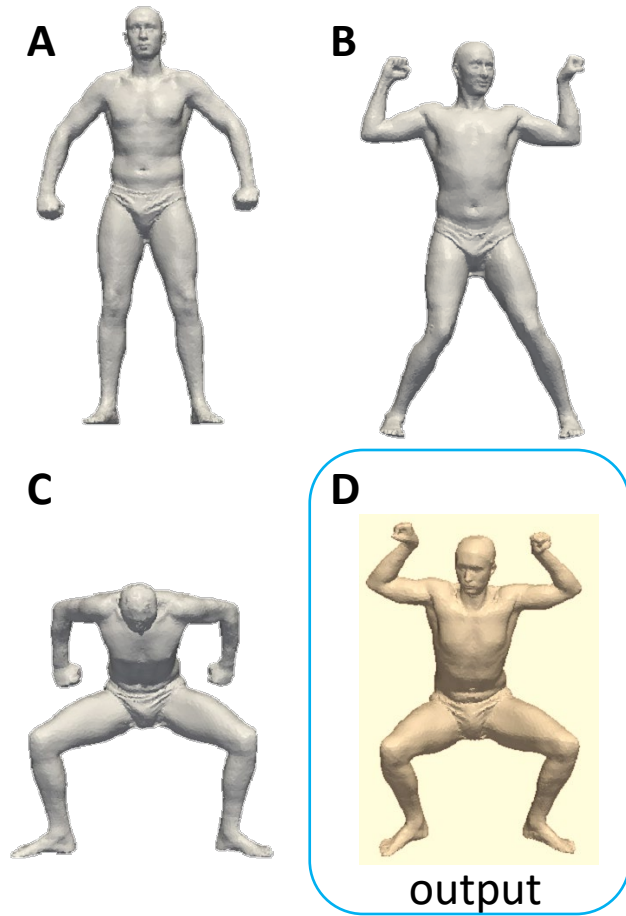
or

$$D = C B A^{-1}$$

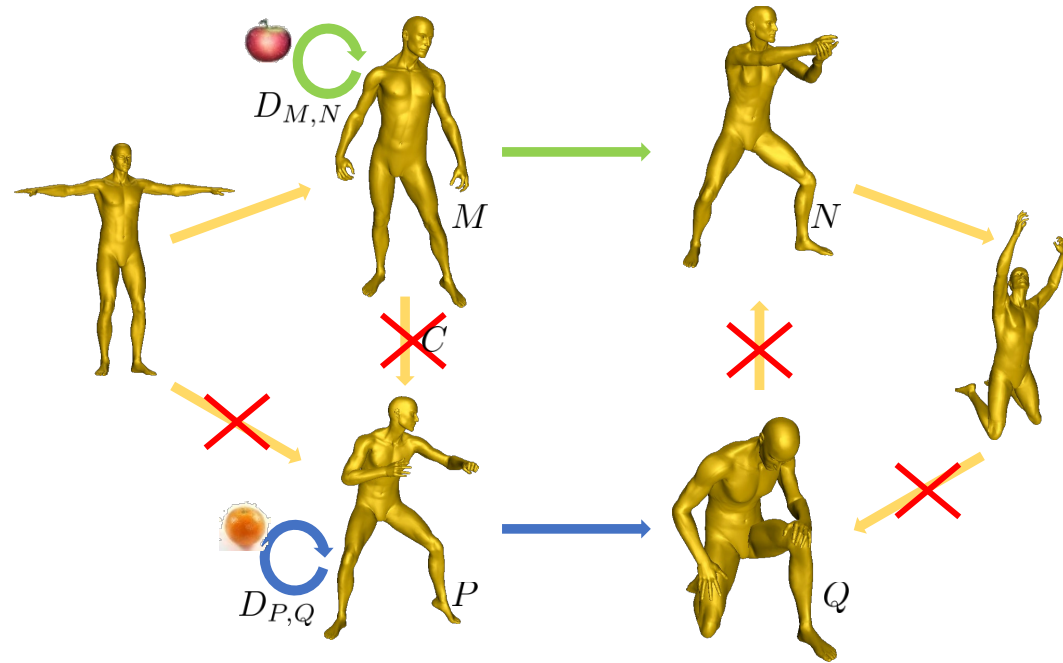
# Analogies: D relates to C as B relates to A



# Shape Analogies



# Comparing Differences III

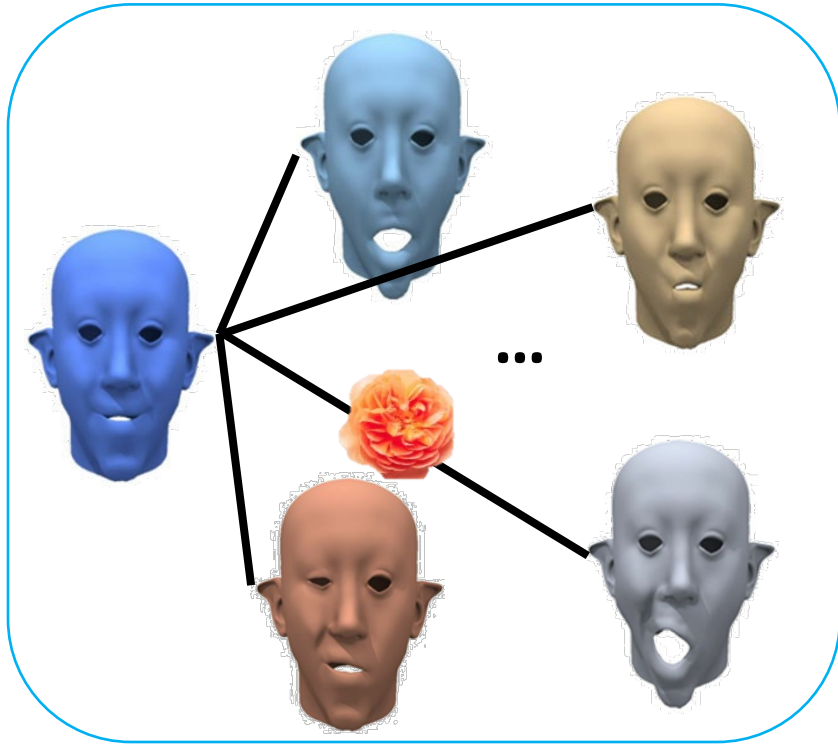


$$D_{M,N} \sim C^{-1} D_{P,Q} C$$

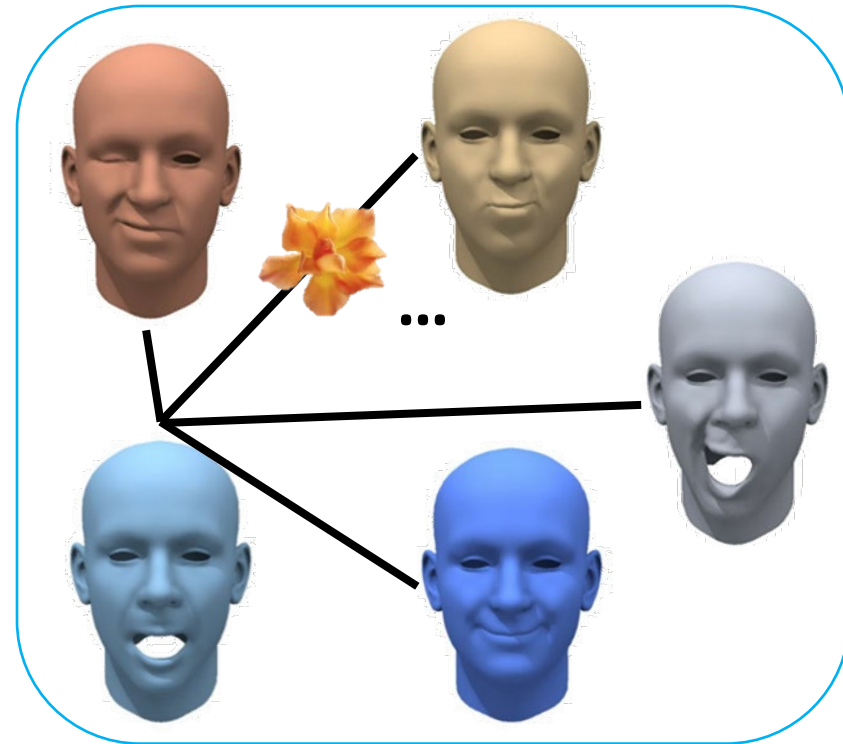
$$\text{Spec}(D_{M,N}) \sim \text{Spec}(D_{P,Q})$$



# Aligning Disconnected Collections

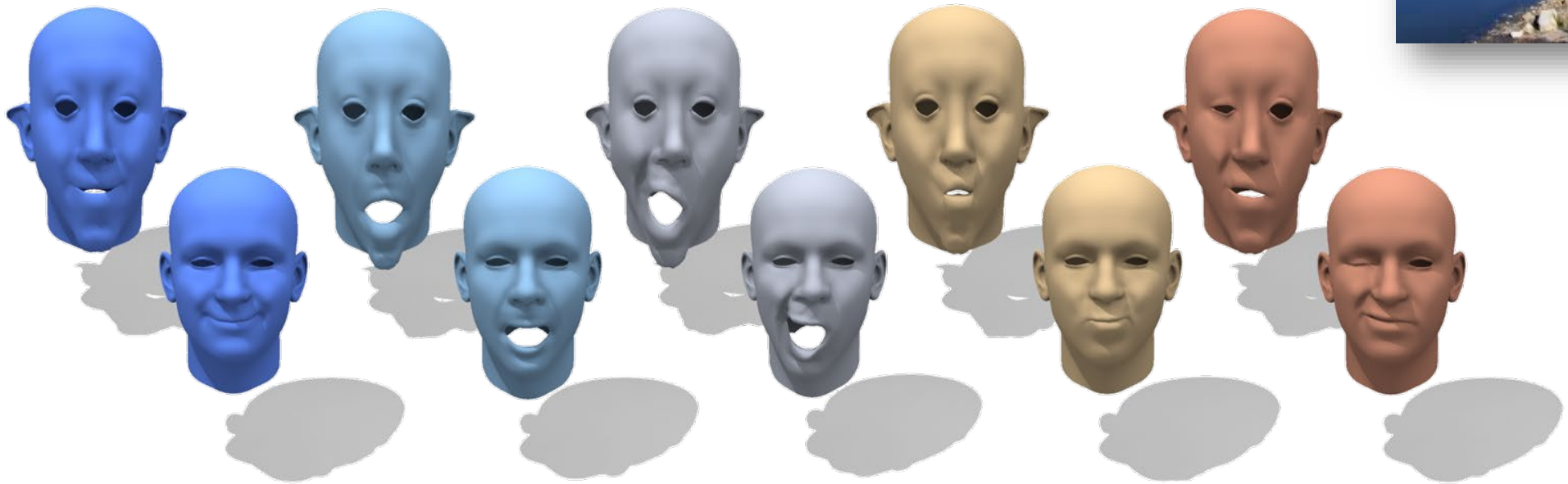


Complete graph



Complete graph

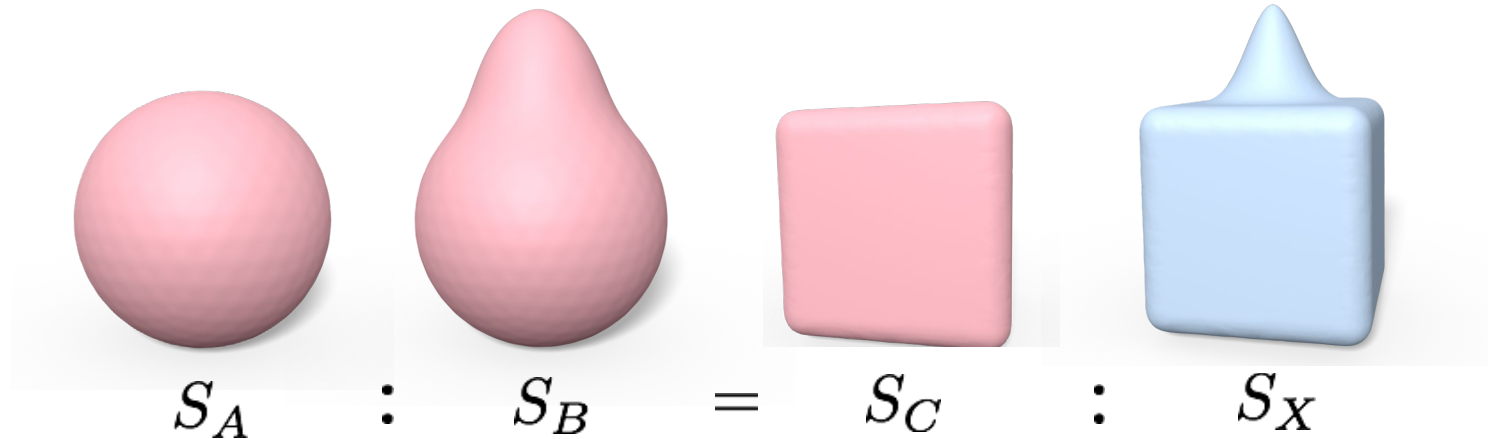
# Aligning, Without “Crossing the River”



Comparing the differences is sometimes easier than comparing the originals

# Shape from Differences

[E. Corman, J. Solomon, M. Ben-Chen, L. Guibas, M. Ovsjanikov; ACM ToG '17]  
[R. Huang, P. Achlioptas, M.J. Rakotosaona, M. Ovsjanikov, L. Guibas; ICCV '19]



# Shape Reconstruction from Differences (Full Basis)

Given the area weights, can solve for triangle areas.

[Slide ack: J. Solomon]

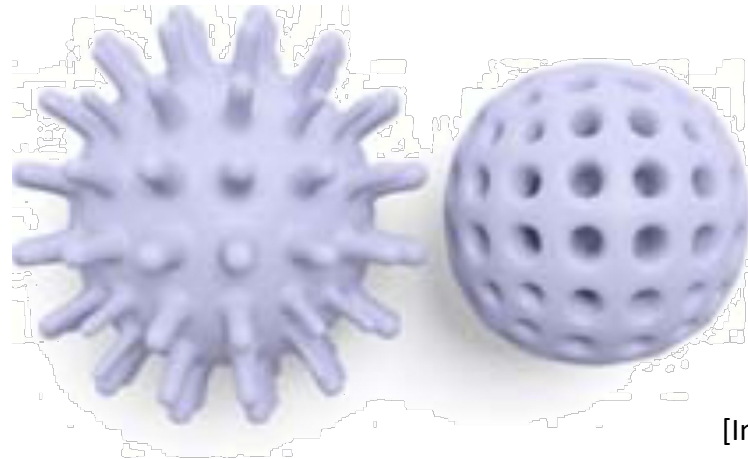
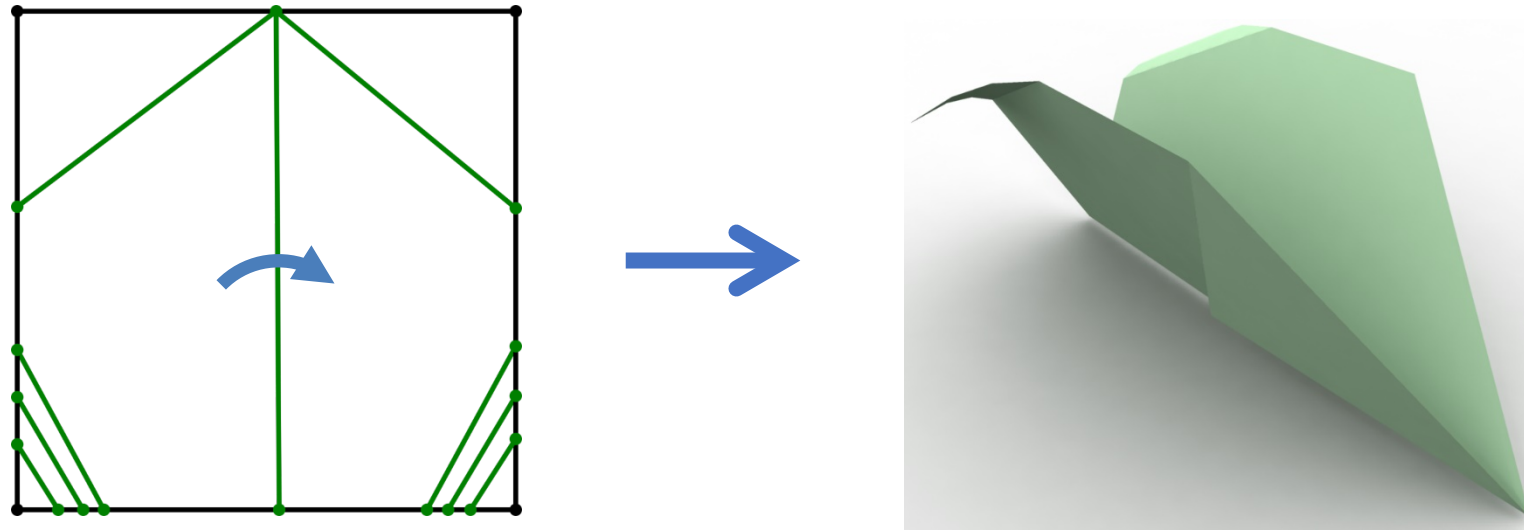
Area-based shape difference  $\rightarrow$  Area weights  $\rightarrow$  Triangle areas

Given triangle areas and conformal inner products, can solve for squared edge lengths.

Area-based shape difference  $\rightarrow$  Area weights  $\rightarrow$  Triangle areas  $\rightarrow$  Squared edge lengths

# How to Encode Extrinsic Information?

[Image: J. Solomon]



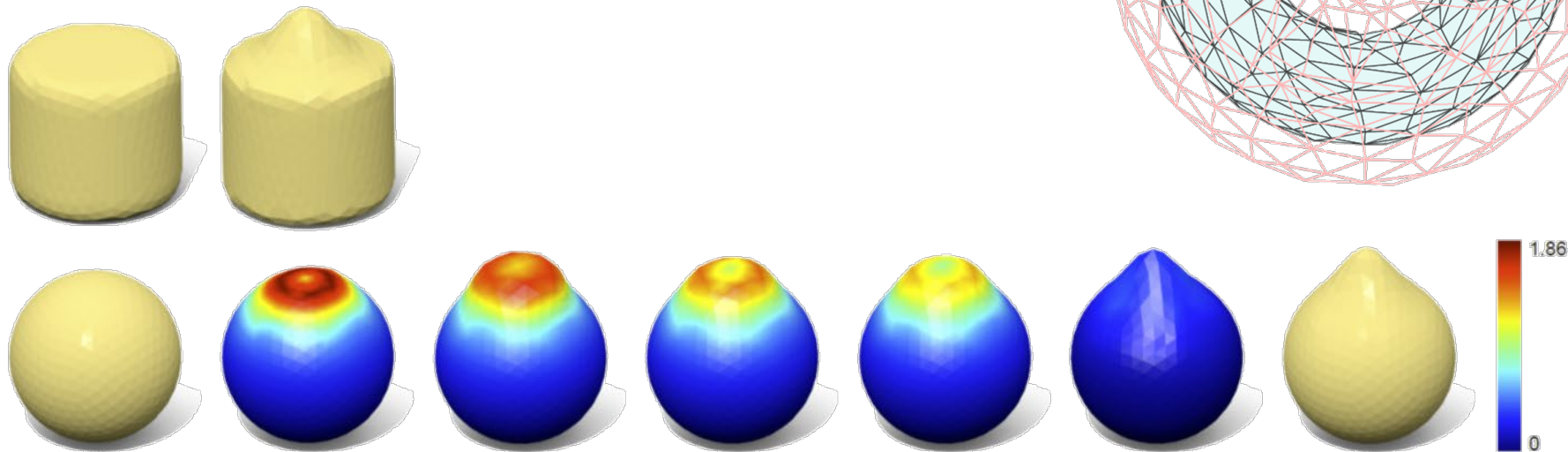
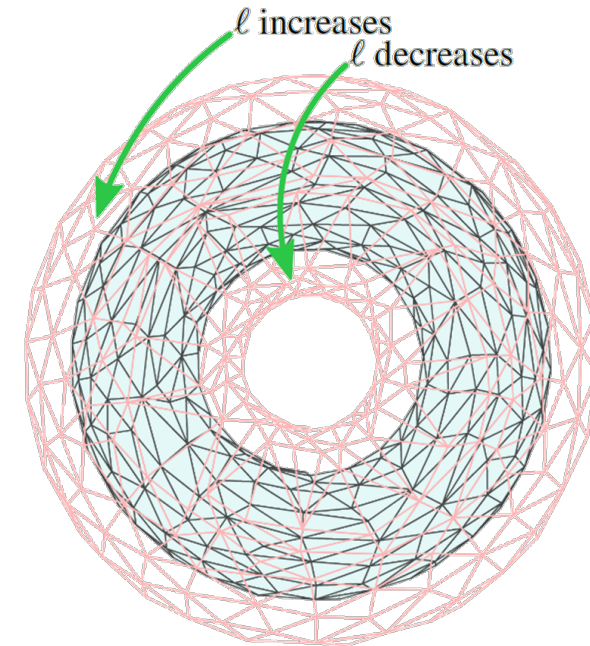
[Image: K. Crane]

# Extrinsic Shape Differences, Version 1

By adding intrinsic differences of an offset surface, we capture extrinsic distortions of the original surface!

Full recovery is provably possible

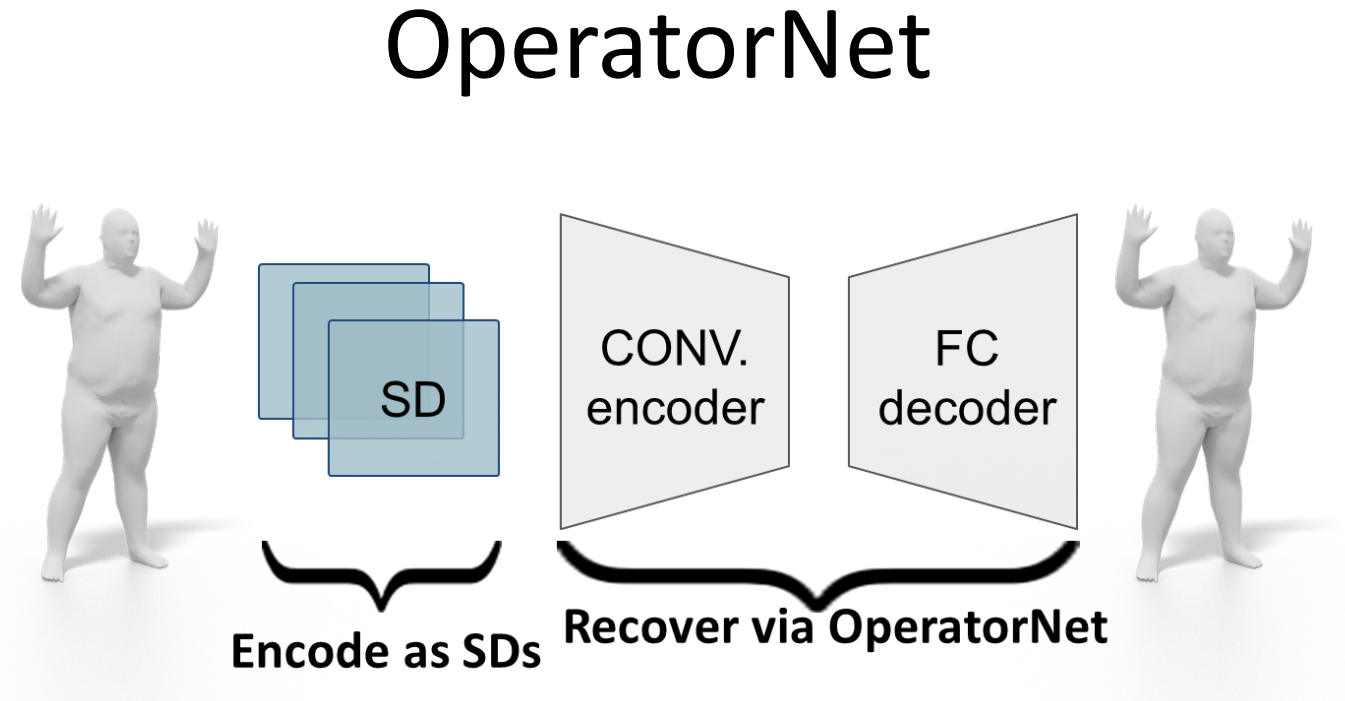
In practice, challenging optimization problem, especially when the functional basis has been truncated



# Extrinsic Shape Differences, Version 2

Decode 3D shapes via deep nets directly from shape difference operators

- **Advantages:**
  - **Compact encodings** (small matrices of size)
  - **Natural algebraic** manipulation
  - **Invariant** to rigid transformation
  - Adapted to **convolutional** neural networks
- Applications: shape interpolation, style transfer, up-sampling



# Intrinsic and Extrinsic

Let us assume that the shapes are in vertex to vertex correspondence. Then we define (1) the intrinsic area-based and conformal differences as before, and (2) an extrinsic shape difference, as follows:

$$\begin{aligned}V &= A_{\text{source}}^{-1} F^T A_{\text{target}} F \\R &= (L_{\text{source}} A_{\text{source}})^{-1} F^T L_{\text{target}} A_{\text{target}} F \\E &= (M_{\text{source}} A_{\text{source}})^{-1} F^T M_{\text{target}} A_{\text{target}} F\end{aligned}$$

Here  $F$  is the truncated basis functional map,  $A$  is the area-weights mass matrix, and  $L$  is the standard Laplacian, and  $M$  is an “extrinsic” Laplacian.

$$M_{i,j} = -\|v_i - v_j\|^2 \text{ if } i \neq j, \text{ or } \sum_{k,k \neq i} M_{i,k} \text{ if } i = j.$$

# Example



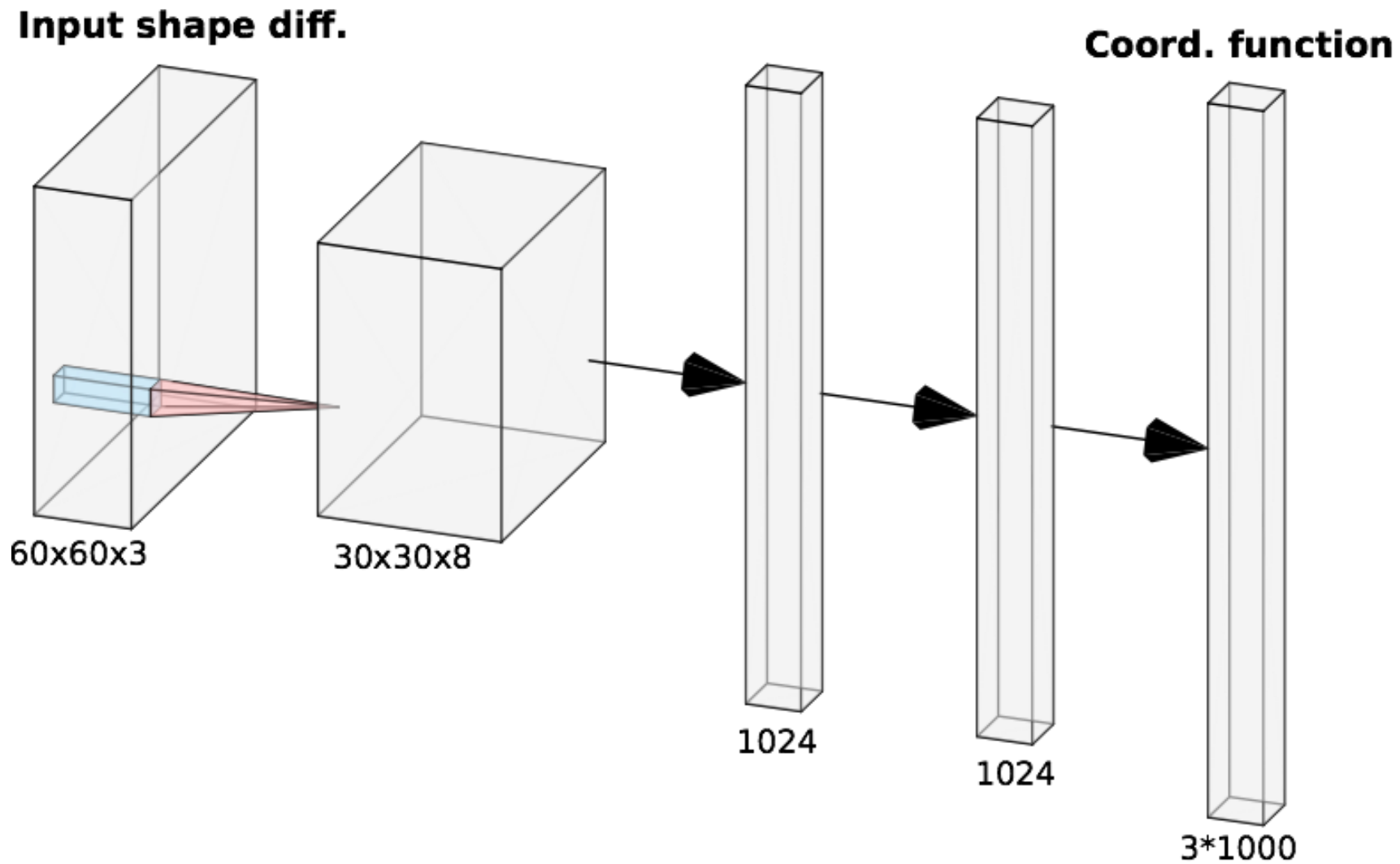
Pose oblivious -- intrinsic



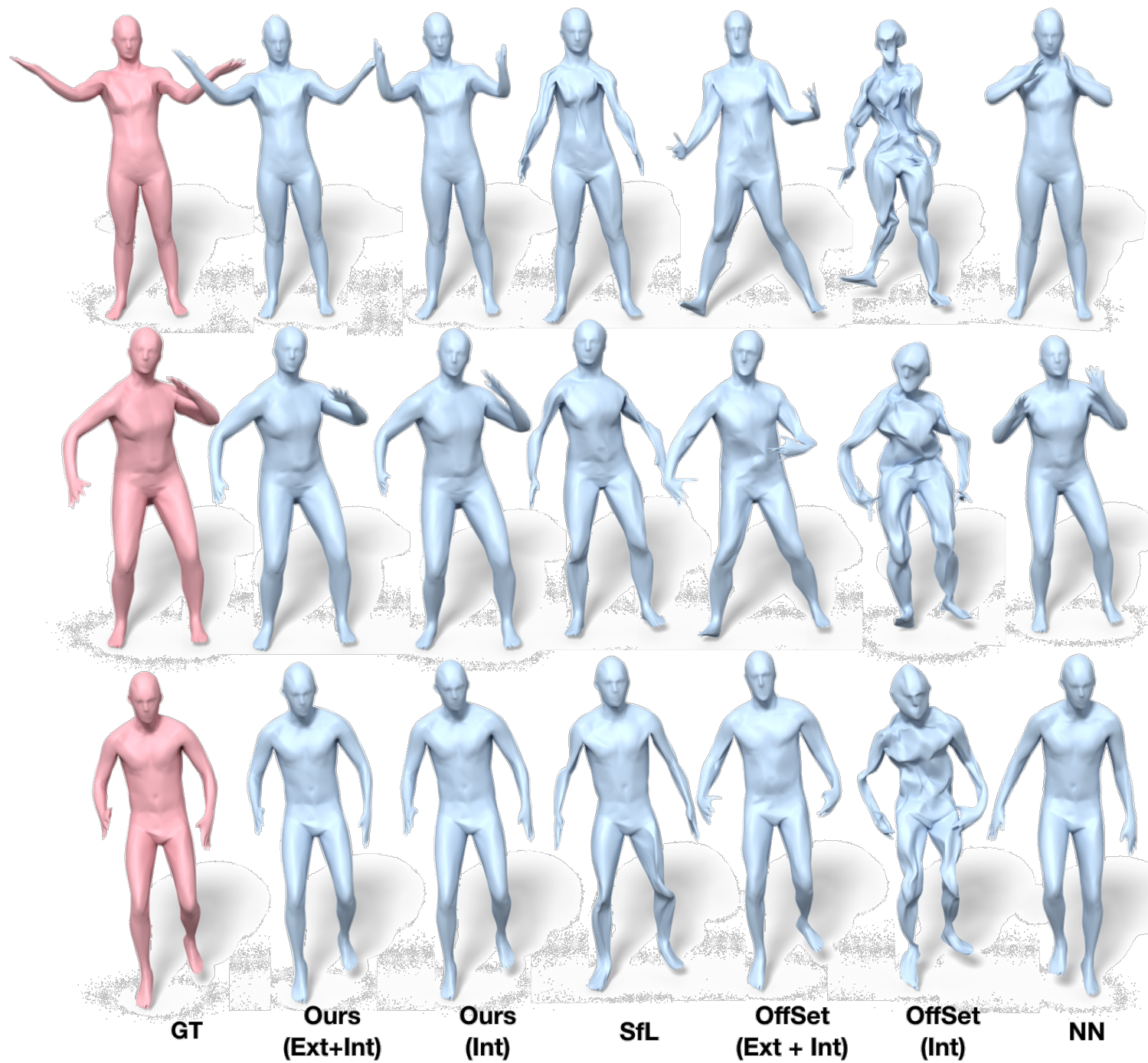
Pose dependent - extrinsic



# OperatorNet Reconstruction



# Reconstruction Comparisons

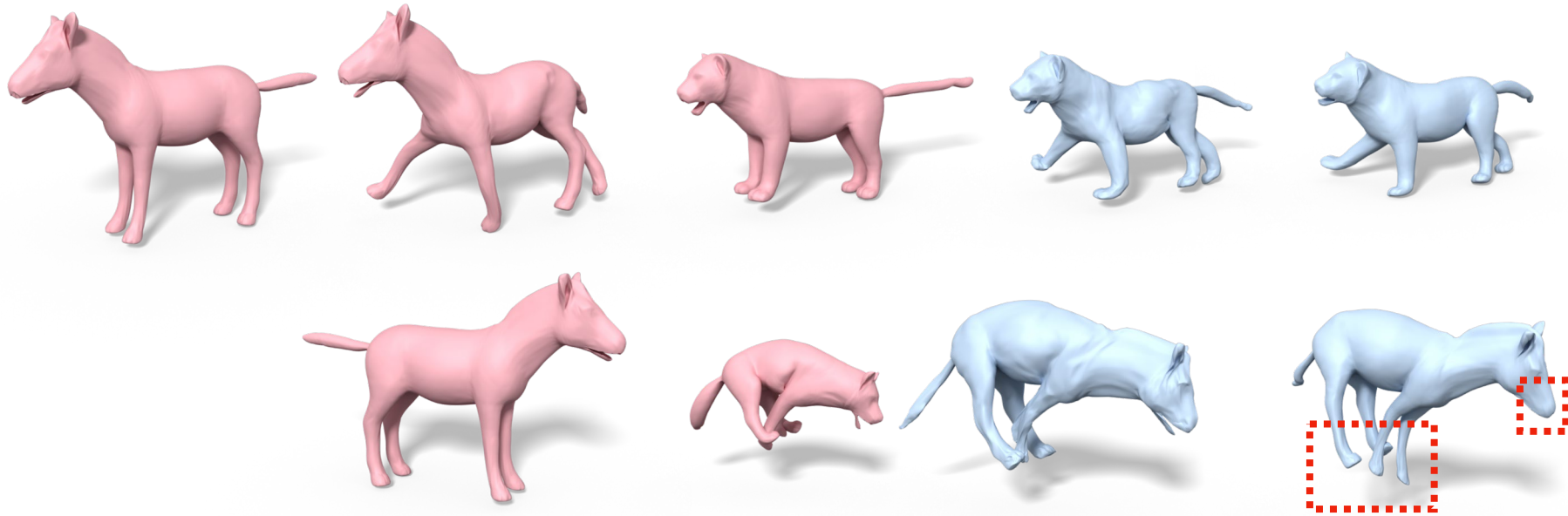


# More Examples



# Shape Analogies

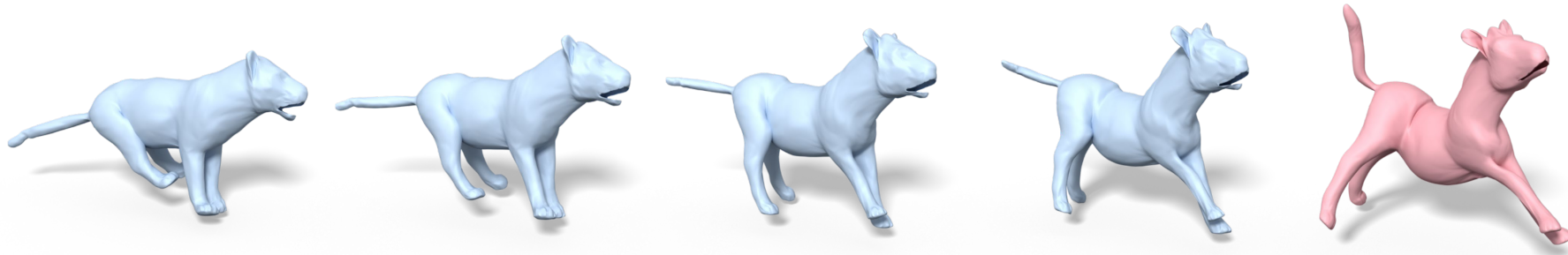
$$S_A : S_B = S_C : S_X$$



**OperatorNet**

**PointNet**

# Interpolation



Interpolation between a dog and a horse

# Shape Interpolation

**PointNet**



**NN**

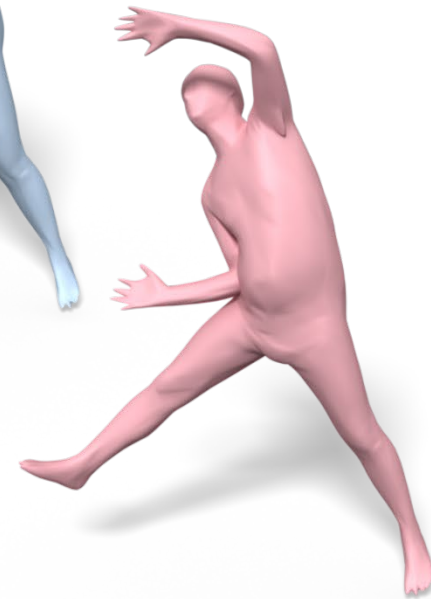
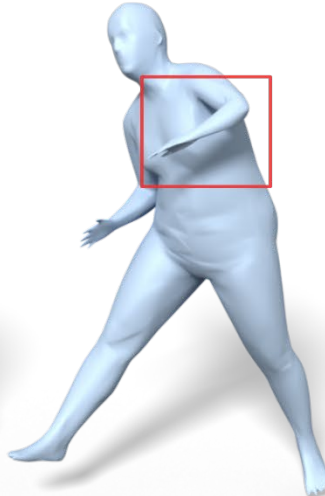


**OperatorNet**

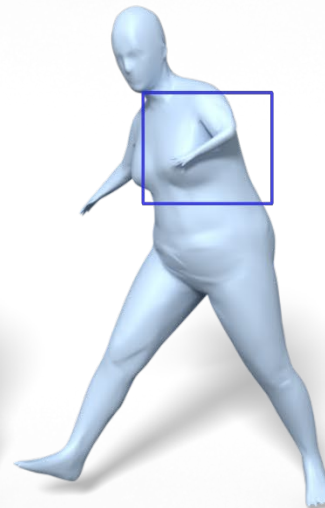


# Interpolation Detail

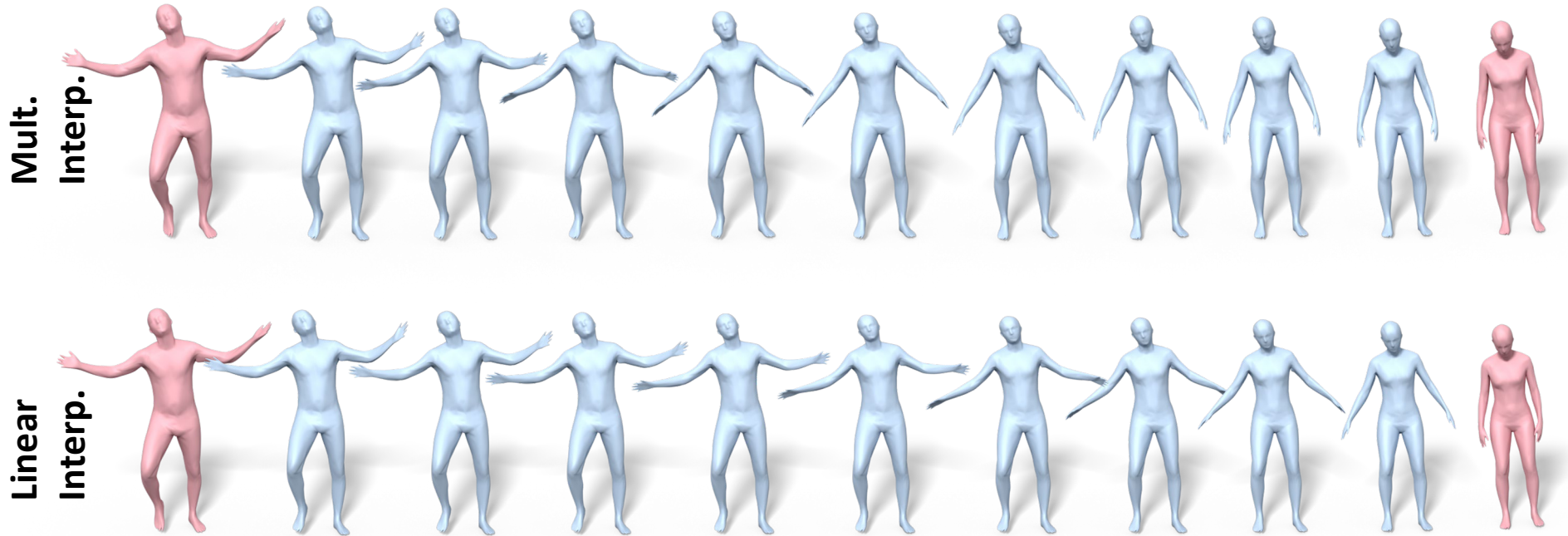
**OperatorNet**



**Pointnet**

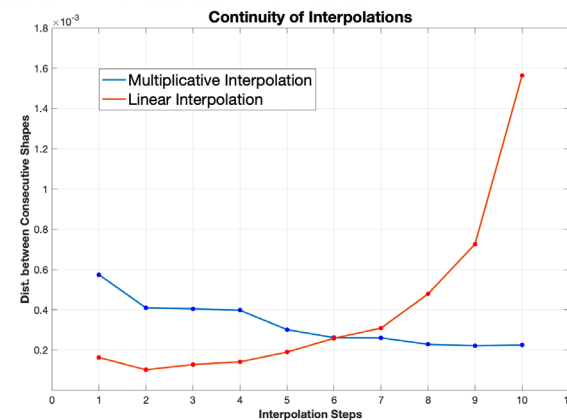


# Additive or Multiplicative?



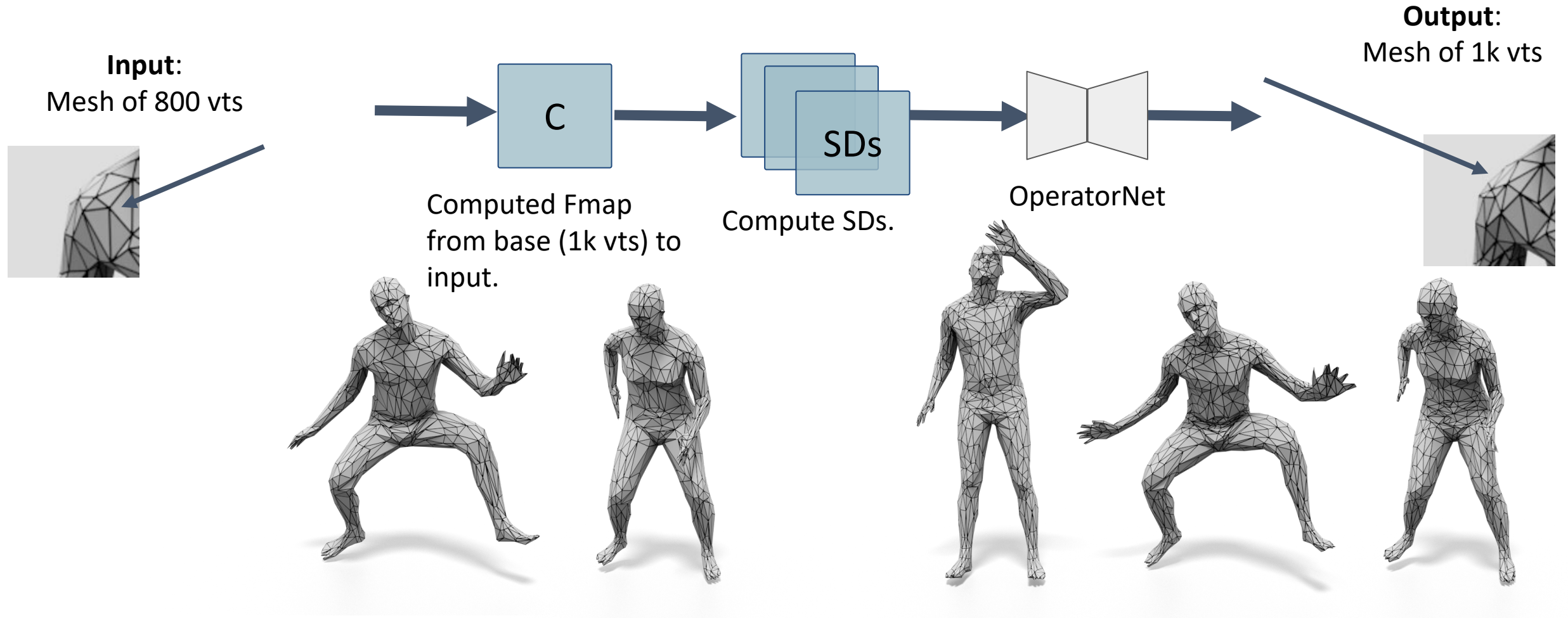
$$C \leftarrow A^t B^{1-t}$$

$$C \leftarrow (1-t)A + tB$$



# Different Triangulations

## Recovery of shapes in different triangulations



# Shape Differences in Language

[P. Achlioptas, J. Fan, R. Hawkins, N. Goodman, L. Guibas; ICCV '19]

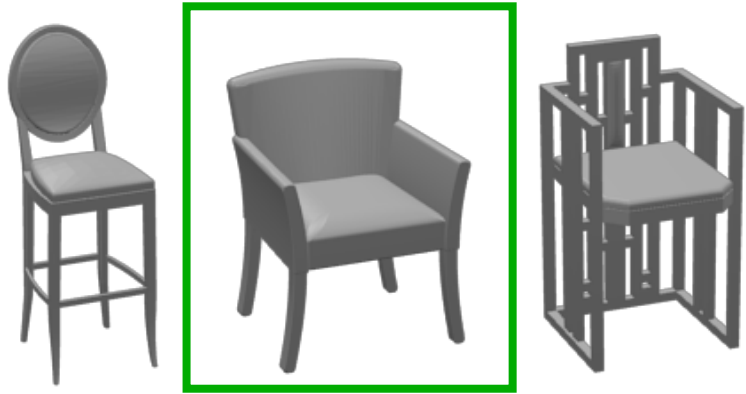
# Differences in Geometry Expressed in Language



Target Object

“Gaps in the back”

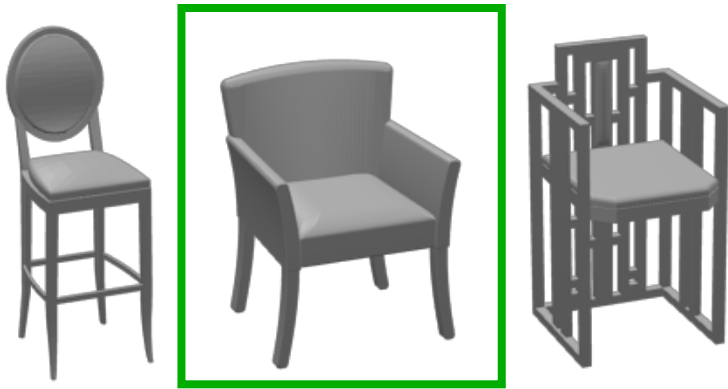
# A Reference Language Game



'looks like a sofa'



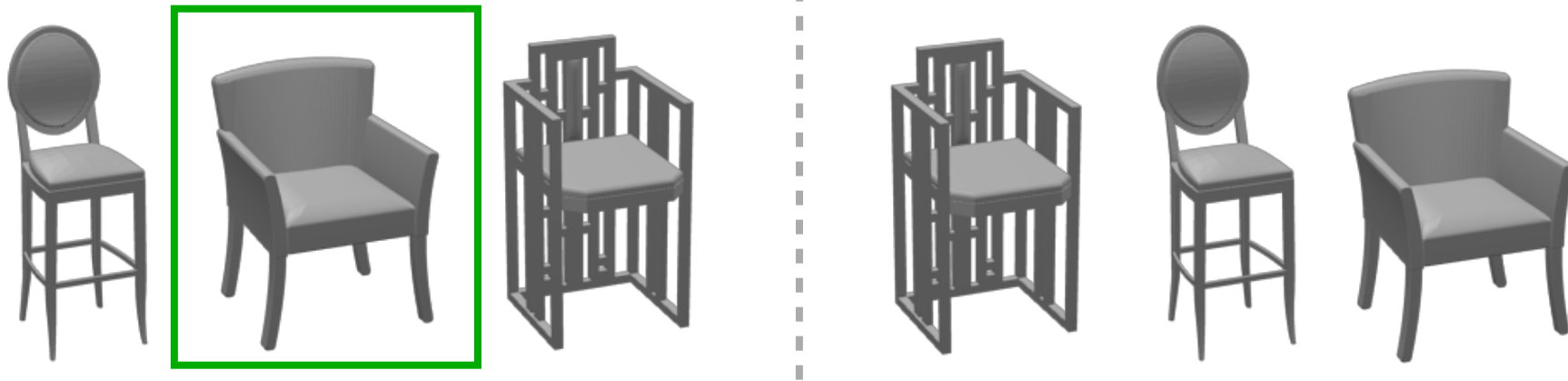
# A Reference Language Game



nailed it



# Considerations



scrambled order  
same view-point  
no texture

no location  
no orientation  
yes geometry

# Communication Context

impossible

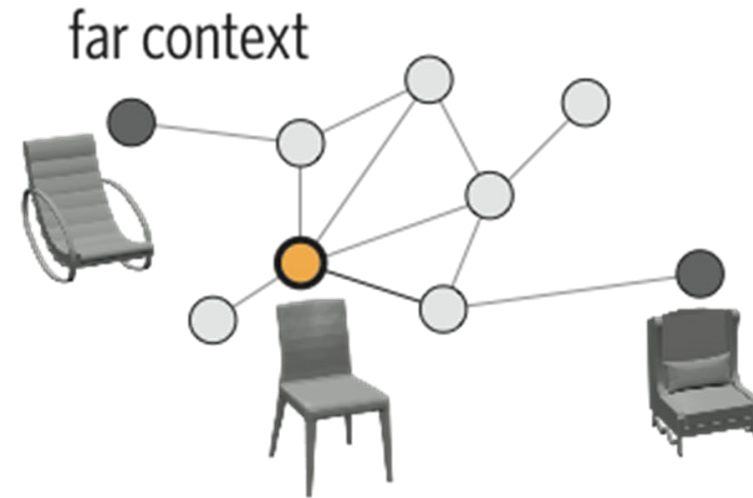
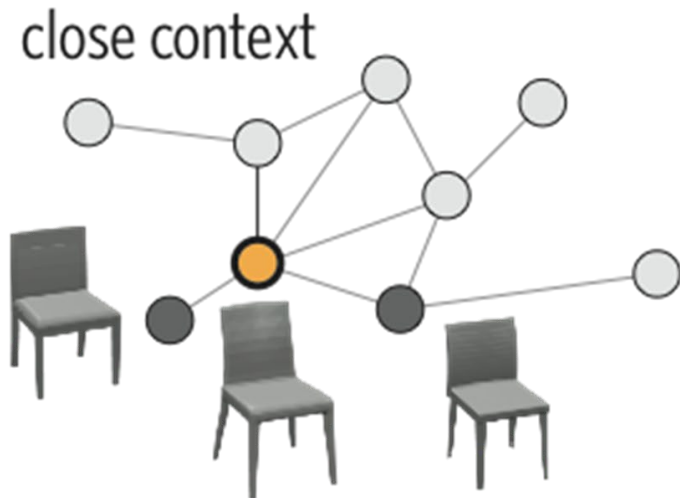
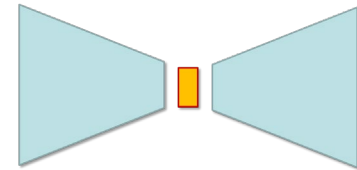


trivial



# Shape Similarity

- Point cloud auto-encoder latent space
- Popular chairs via *max-degree*



# “Chairs in Context” Corpus / Data Set

- 4,054 distinct contexts covering 4,511 chairs
- 78,789 utterances by 2,124 unique AMT participants
  
- “close” vs. “far” human accuracy: 94.2% vs. 97.2%
- 85% of utterances contain a part-related words
- “close” utterances rich in adjectives, comparatives

# Utterance Examples

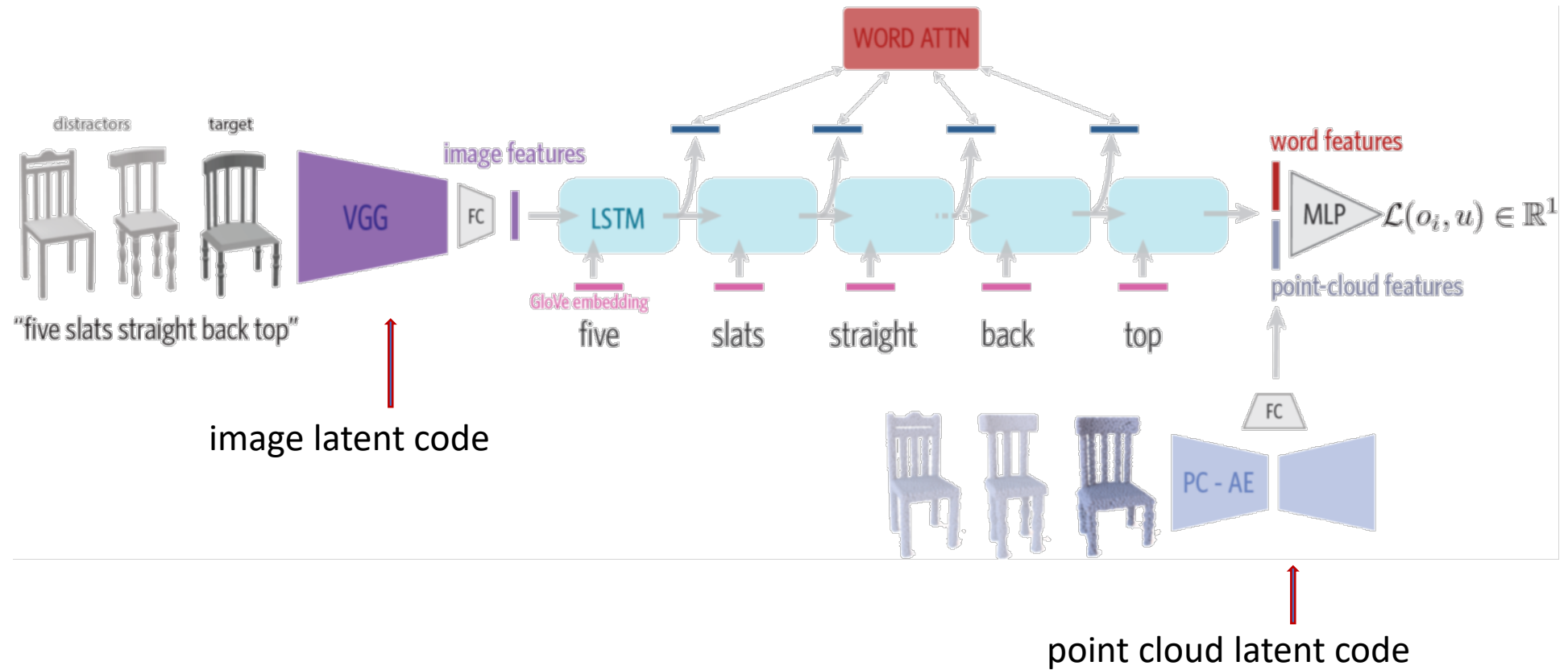


“it has wheels”

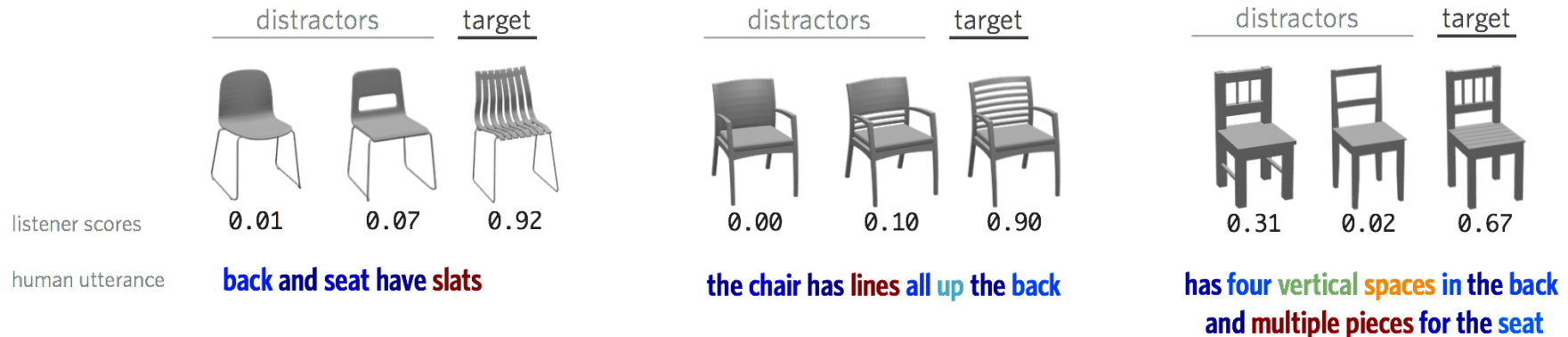
“has vertical lines on the back”

“rectangle back with straight legs”

# Neural Listener w. Attention Model

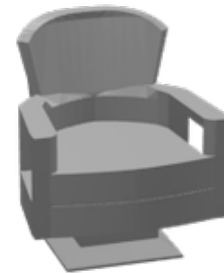


# Attentive Neural Listening



# Performance

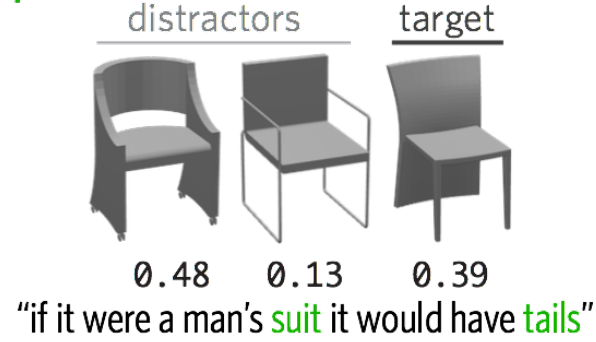
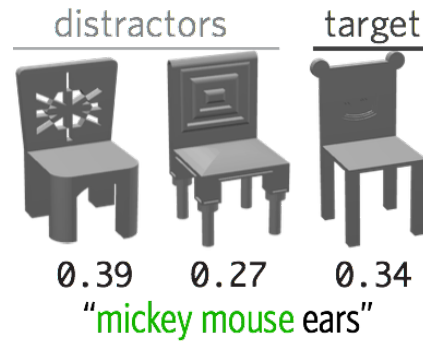
Architecture	Subpopulations			
	Overall	Close	Far	Sup-Comp
Together-with-Context	$75.9 \pm 0.5\%$	$67.4 \pm 1.0\%$	$83.8 \pm 0.6\%$	$74.4 \pm 1.3\%$
Separate-with-Context	$79.4 \pm 0.8\%$	<b><math>70.1 \pm 1.3\%</math></b>	<b><math>88.1 \pm 0.6\%</math></b>	$75.2 \pm 2.1\%$
Separate (proposed)	<b><math>79.6 \pm 0.8\%</math></b>	$69.9 \pm 1.3\%$	<b><math>88.1 \pm 0.4\%</math></b>	<b><math>76.0 \pm 1.6\%</math></b>



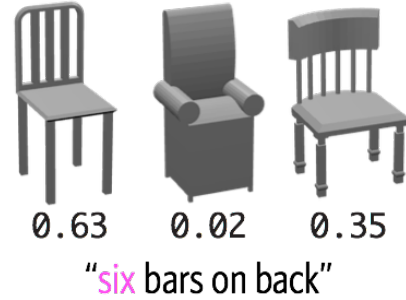
“tallest”

# Listener Failures

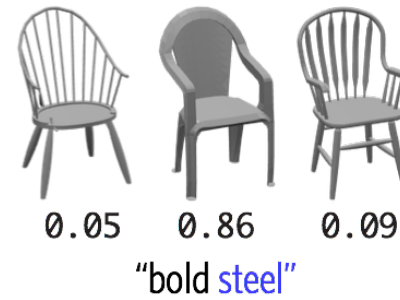
## Metaphors



## Counting



## Material



## Ambiguous

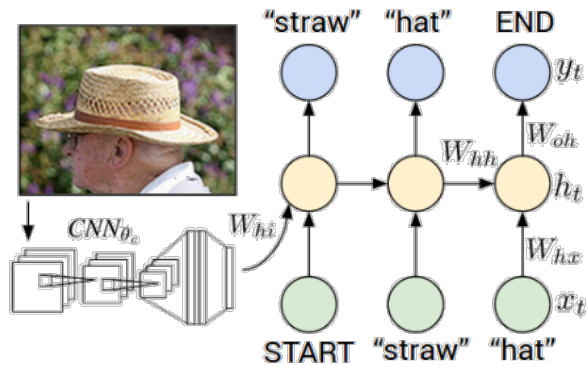


## Negation

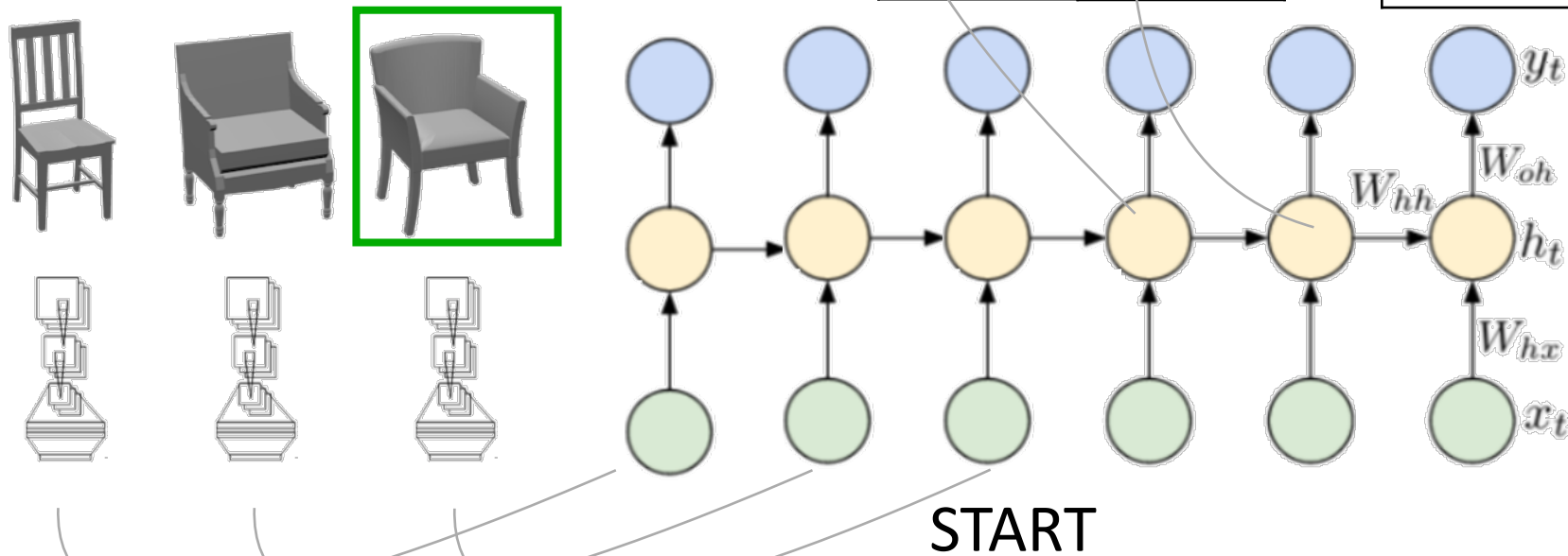


# Neural Speaker

show-and-tell w. teacher forcing



$(p_{u_1})$	$(p_{u_2})$	$(p_{u...})$
curved	curved	curved
the	the	the
...	chair	...
shortest	...	END



# Speaker Examples

## image-based speakers



**pragmatic speaker**

**square arms**

literal speaker

with the tall-est back and seat



**knobby legs**

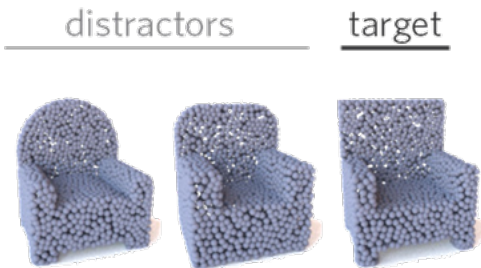
the one with the thick-est legs



**no arm rests**

the one with high-est back

## point-cloud based speakers

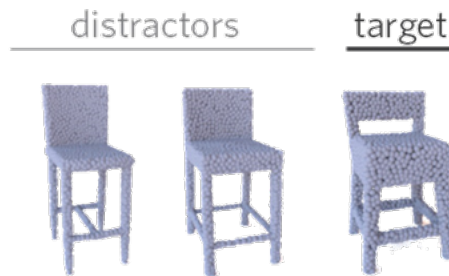


**pragmatic speaker**

**most square back**

literal speaker

thin-est seat



**thick-est legs**

square rack at bottom of chair



**tall-est back**

has arms

# Speaker from Images



# Listener Examples: Shape-Based Product Retrieval

Novel  
Chair  
Collection

*curved seat*



*curved seat, hole on back*



*rectangular hole on back*



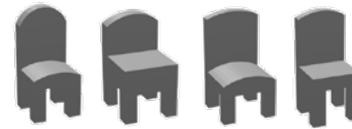
*rectangular hole on back,  
connected legs*



*curved back top*



*curved back top, fat legs*



*thin legs*

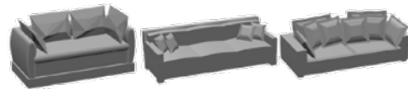


*thin legs, no arms*



Out-of-Train  
Shape  
Collections

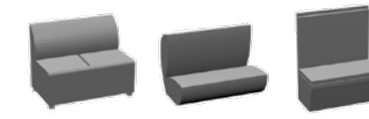
*has pillows*



*three seater*



*no armrests*



*circular*



*skinny legs*



*no legs*



*antique, old looking*



*circular base*



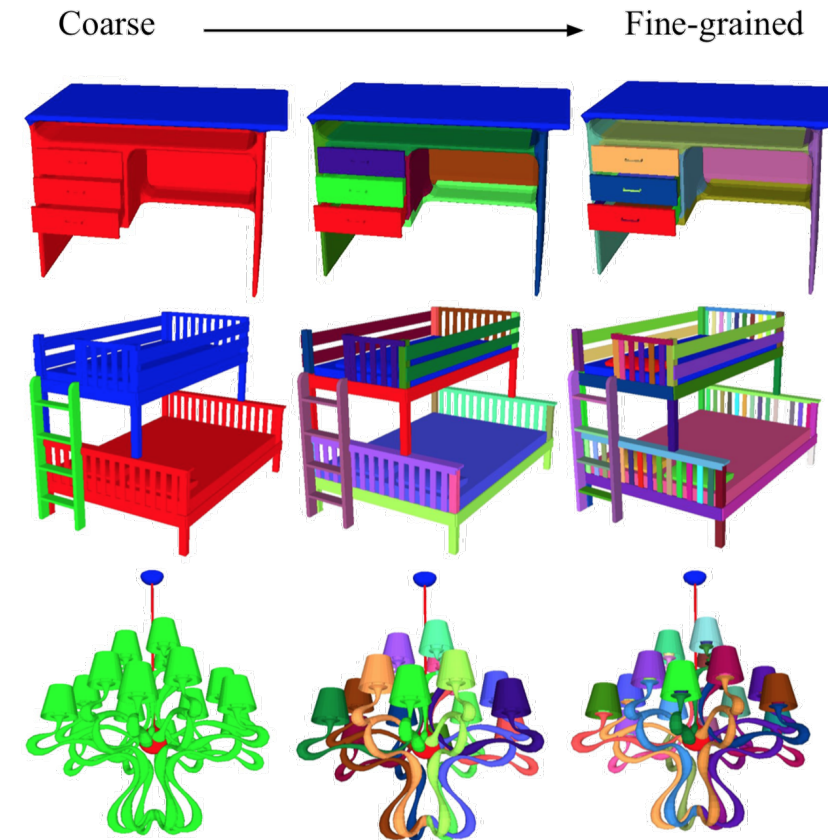
(bottom rows includes *out-of-training* classes)

# Compositional Shape Structure: Shape Parts

[K. Mo, S. Zhu, A. Chang, L. Yi, S. Tripathi, L. Guibas, H. Su; CVPR '19]

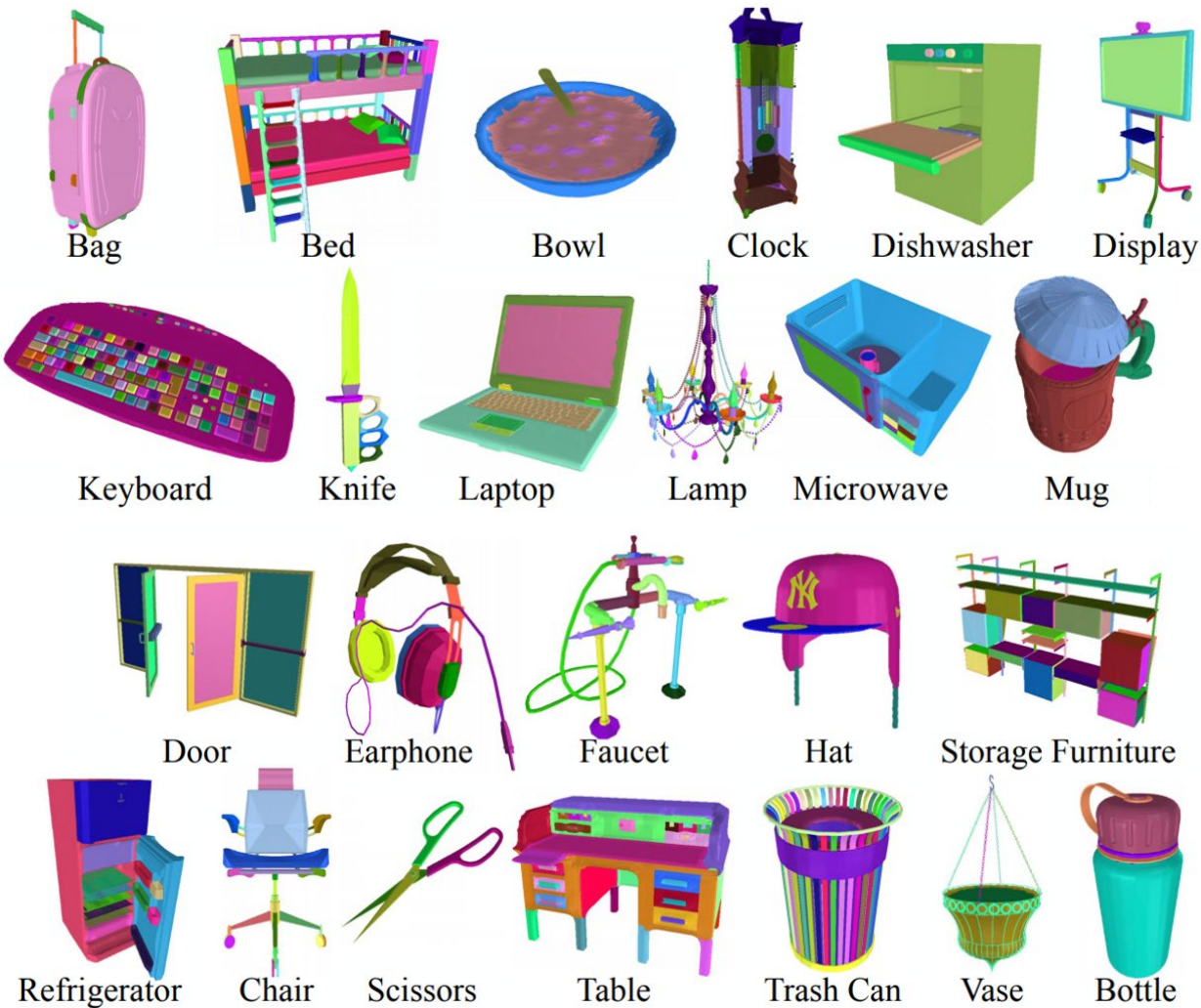
# PartNet: Fine-Grained Object Part Annotation

- Dataset
  - Fine-grained Parts
  - Hierarchical Segmentation
  - Instance-level Segmentation
  - Consistent Semantics
- Statistics
  - 24 Object Categories
  - 26,671 Different Shapes
  - 573,585 Different Parts
  - Avg 18 Part/shape, Max 230



PartNet Dataset

# Based on Curated Part Hierarchies



# Latent Representations for Shape Structure and Structural Differences

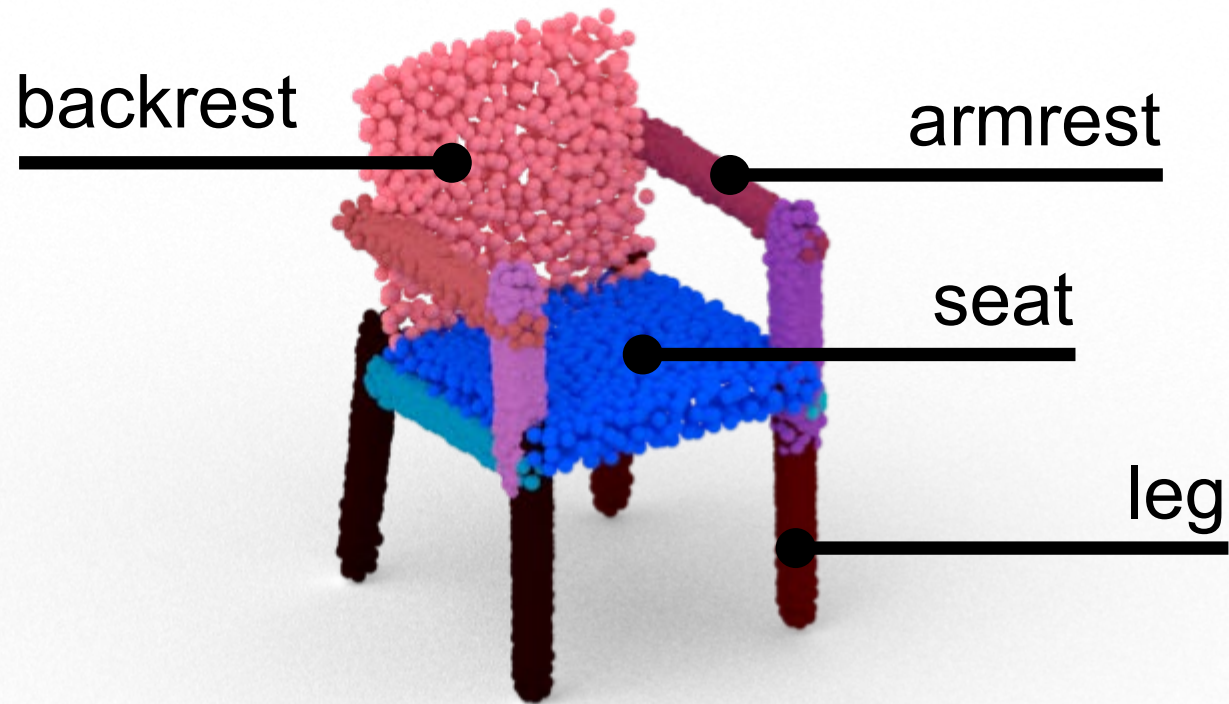
[K. Mo, P. Guerrero, L. Yi, H. Su, P. Wonka, N. Mitra, L. Guibas; Siggraph Asia '19]

[K. Mo, P. Guerrero, L. Yi, H. Su, P. Wonka, N. Mitra, L. Guibas; '20]

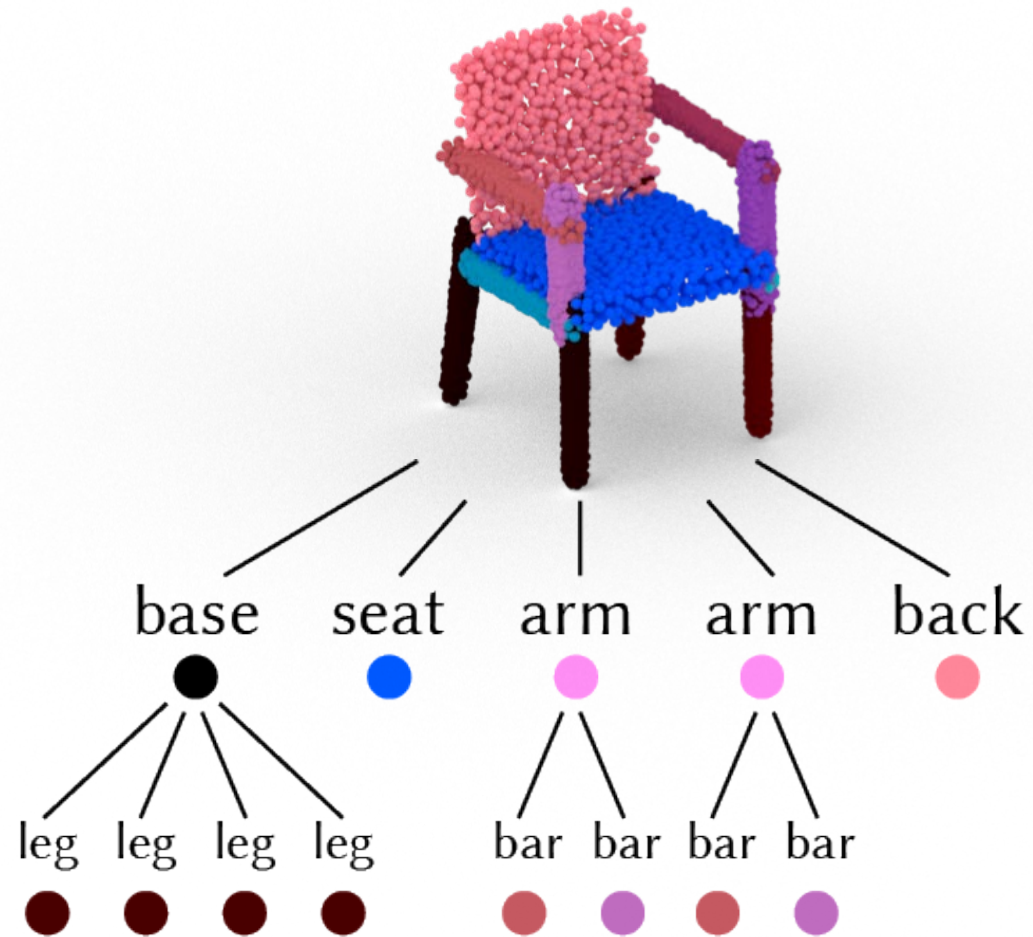
# Geometry and Structure



# Geometry and Structure



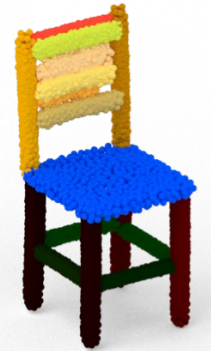
# Structure: Part Hierarchy



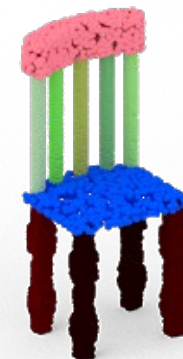
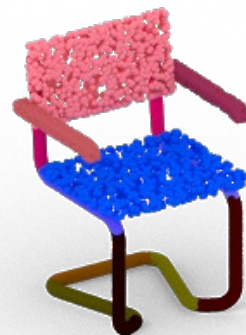
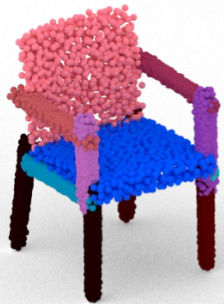
# Goal: A Smooth, Explorable Shape Space



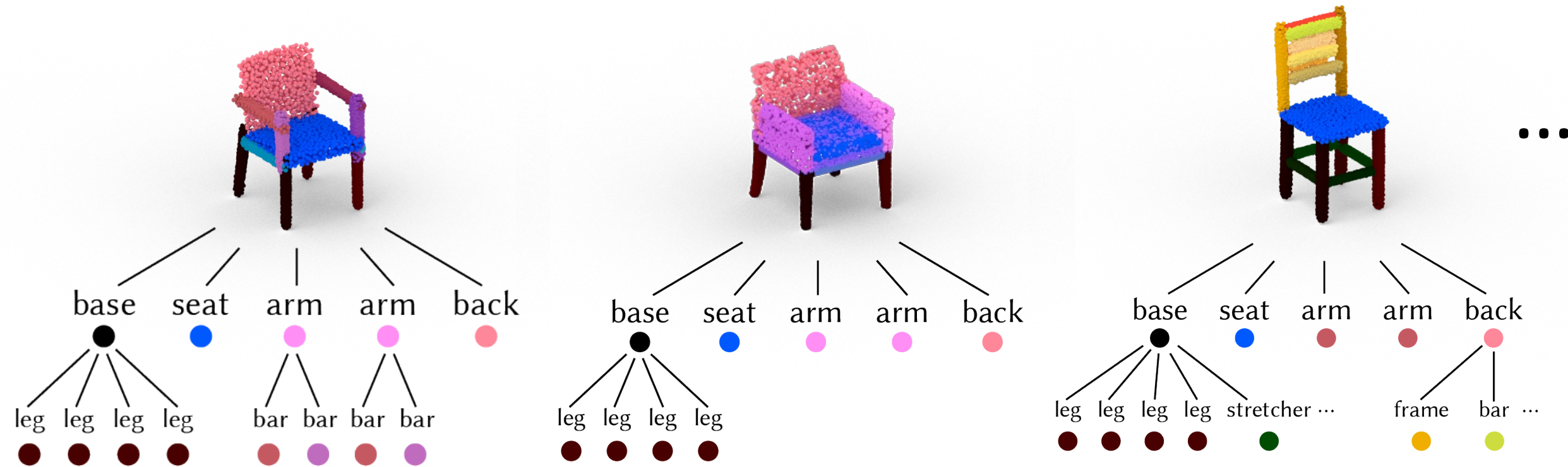
A smooth, common shape space allows for interpolation, generation, exploration, ...



... of both **geometry** and **structure**



# Structural Consistency

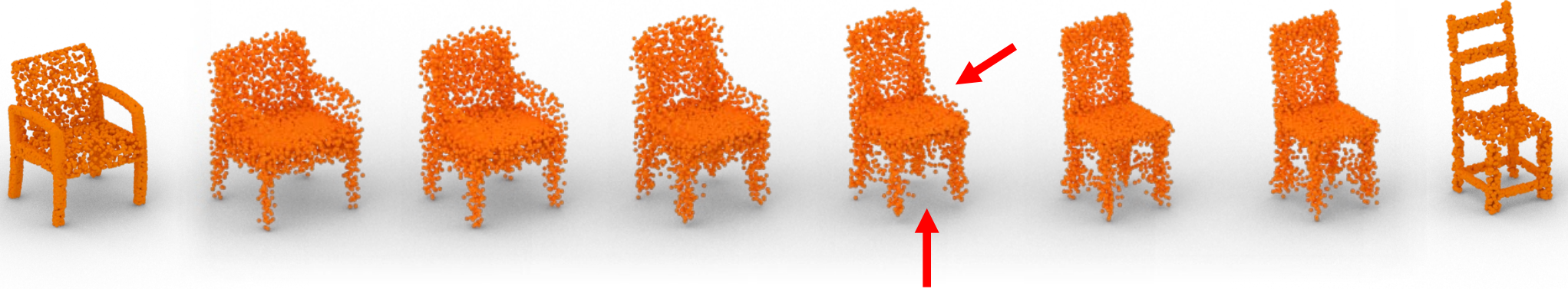


# With vs. Without Structure

source

target

without  
structure



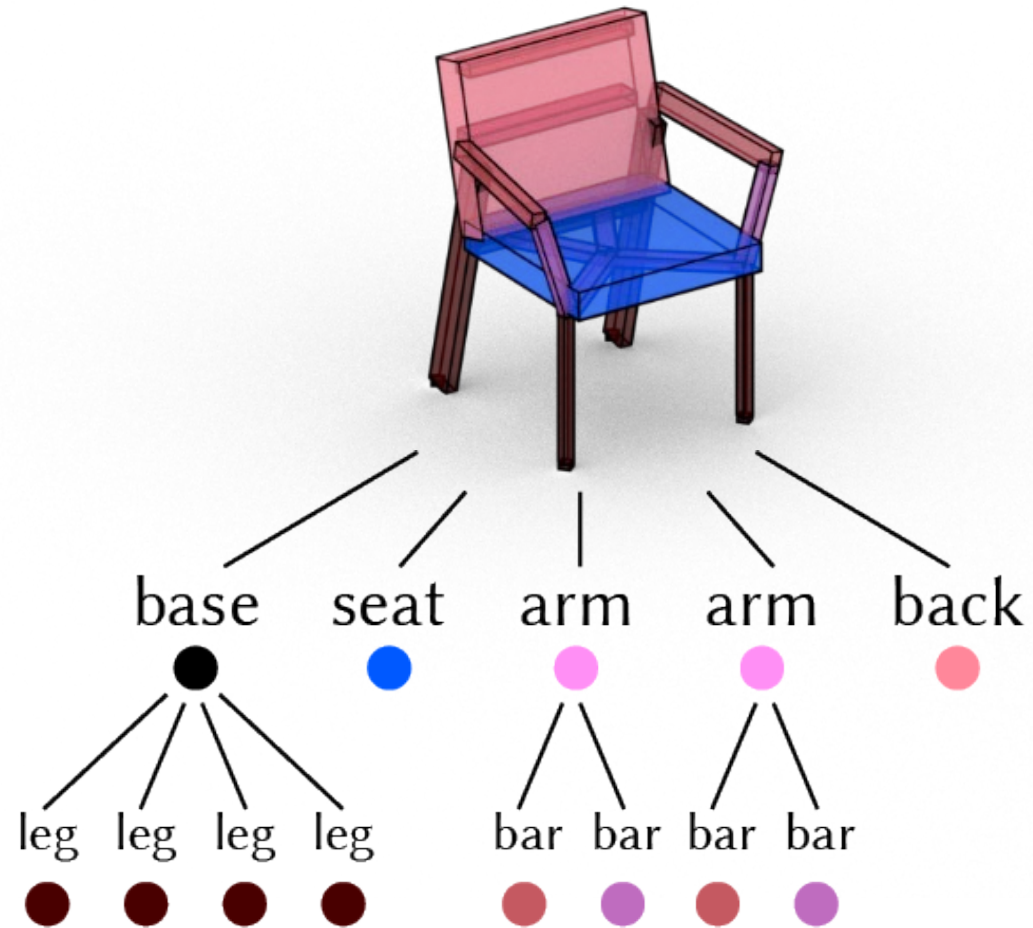
with  
structure



# Object Representation: Part Geometry



# Object Representation: Part Structure



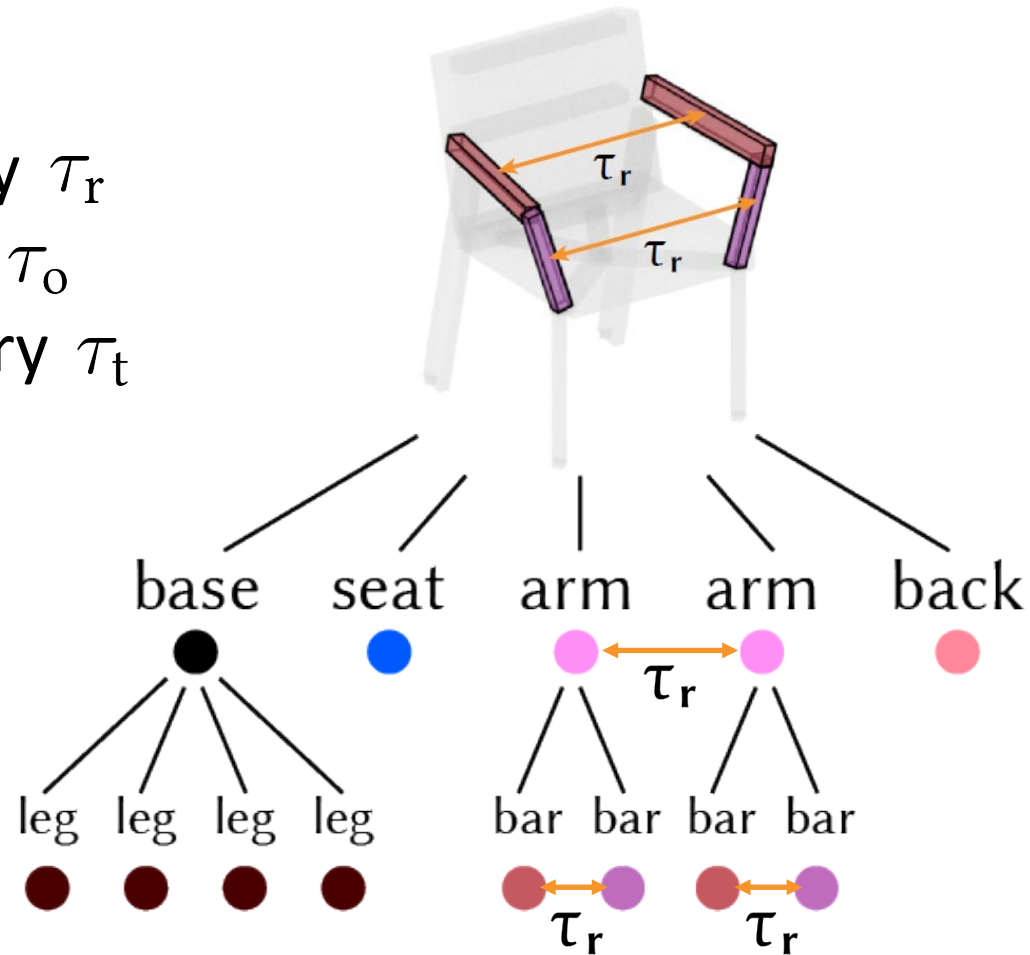
# Object Representation: Sibling Relationships

Reflectional Symmetry  $\tau_r$

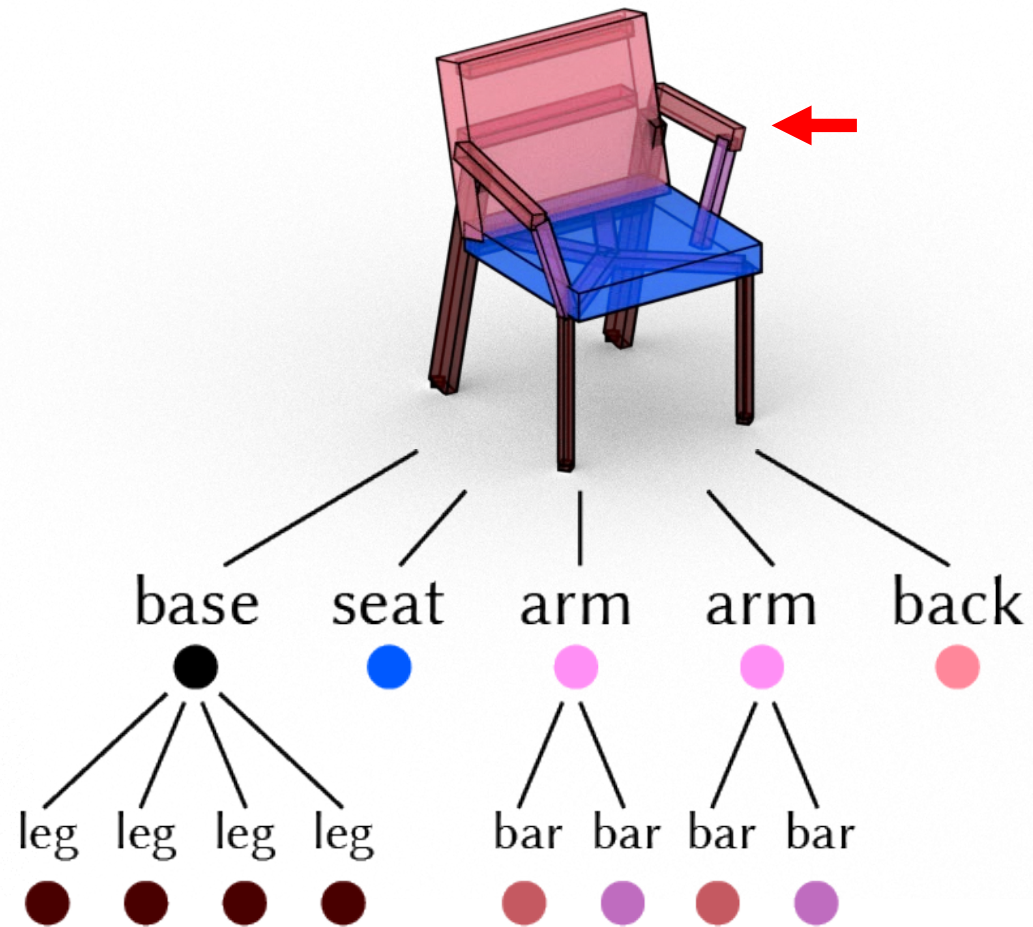
Rotational Symmetry  $\tau_o$

Translational Symmetry  $\tau_t$

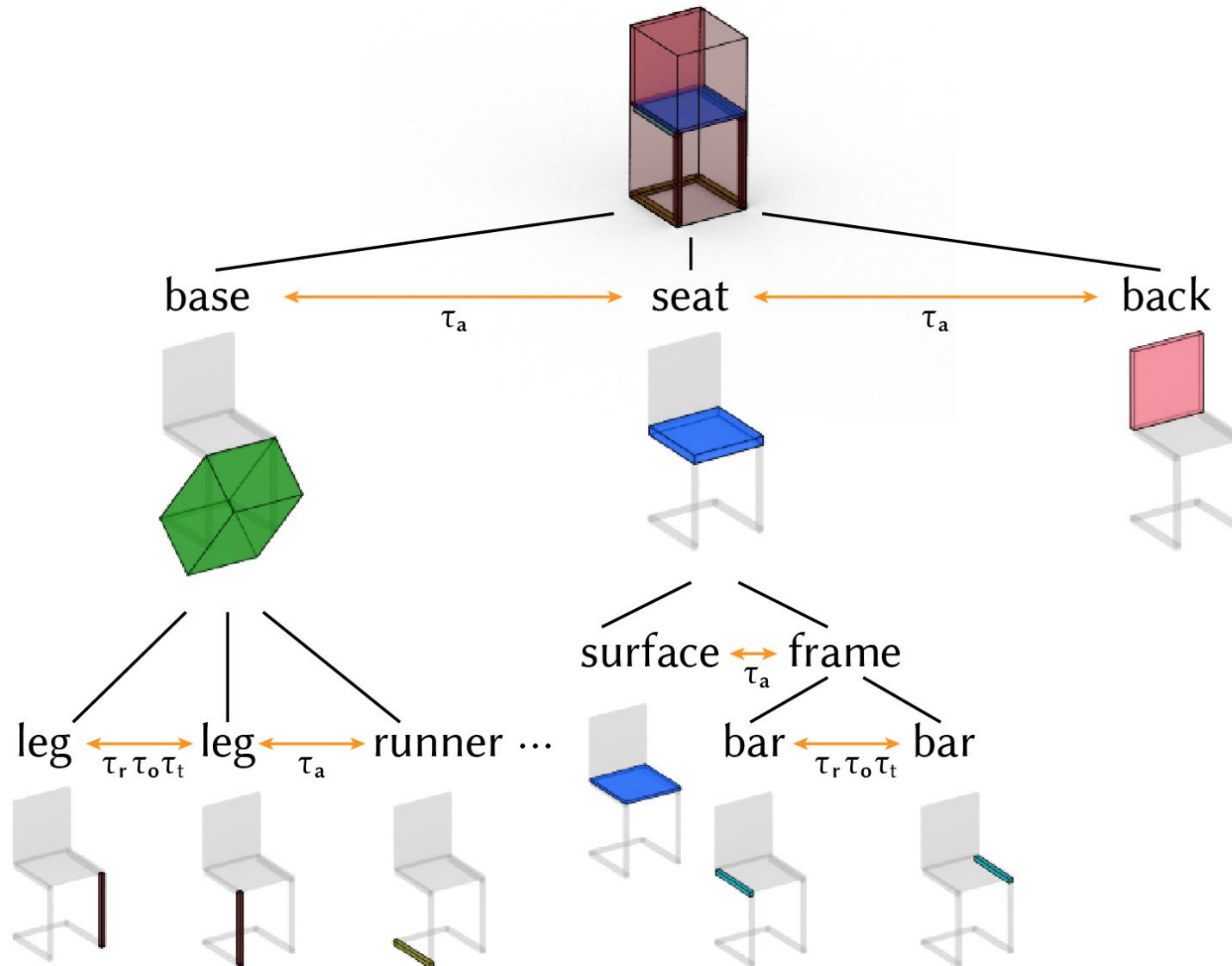
Adjacency  $\tau_a$



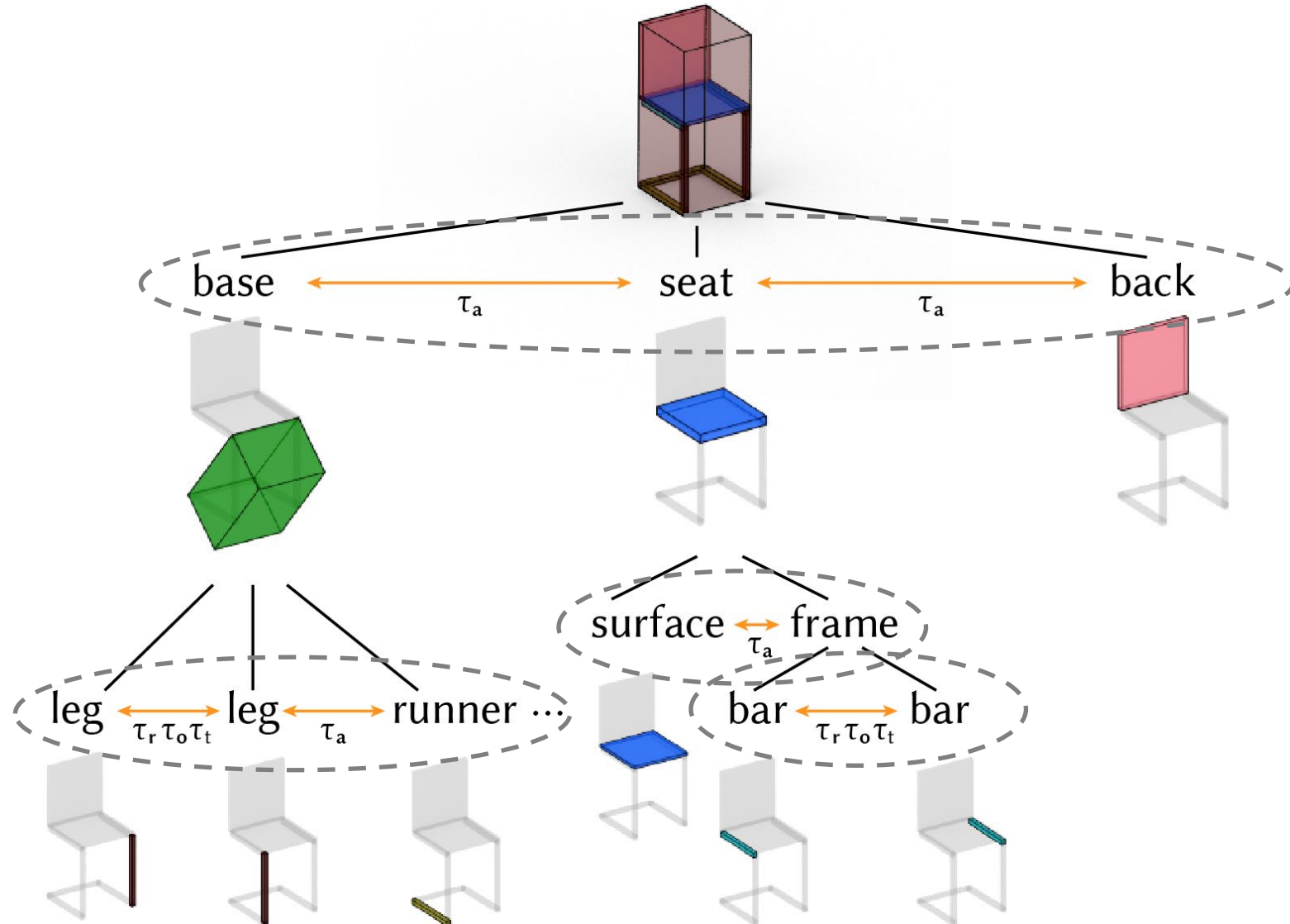
# Object Representation: Part Structure



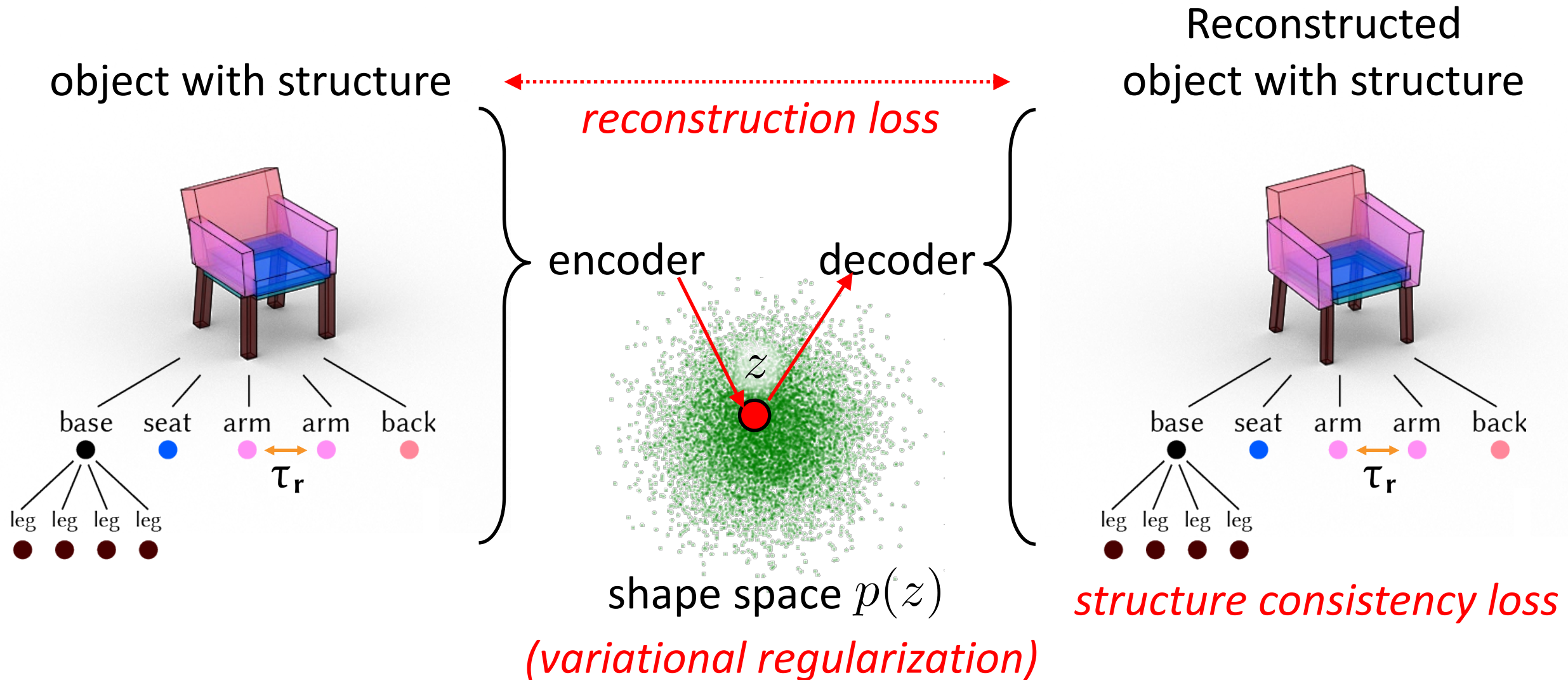
# Object Representation: Example



# A Hierarchy of Graphs

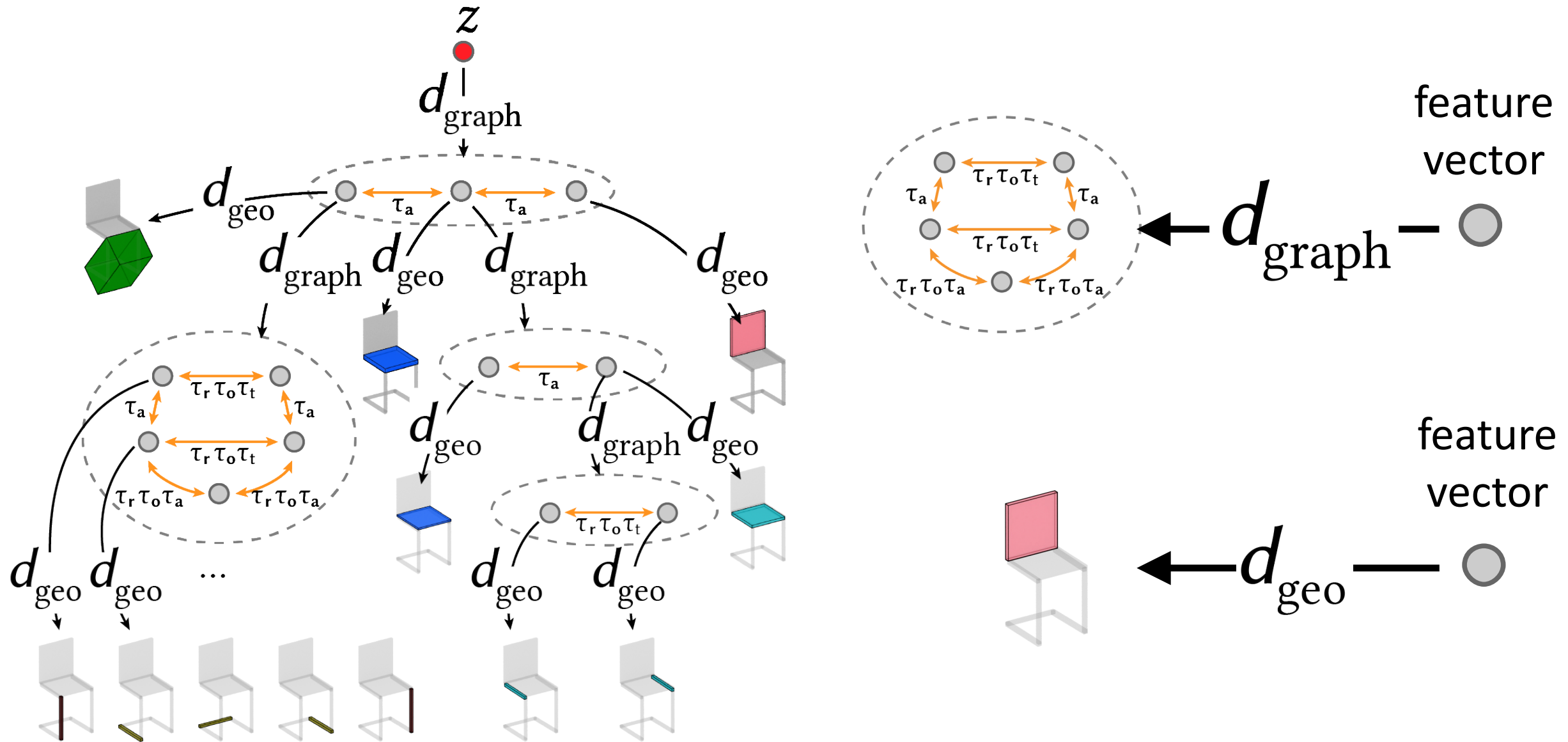


# Architecture Overview: VAE Training

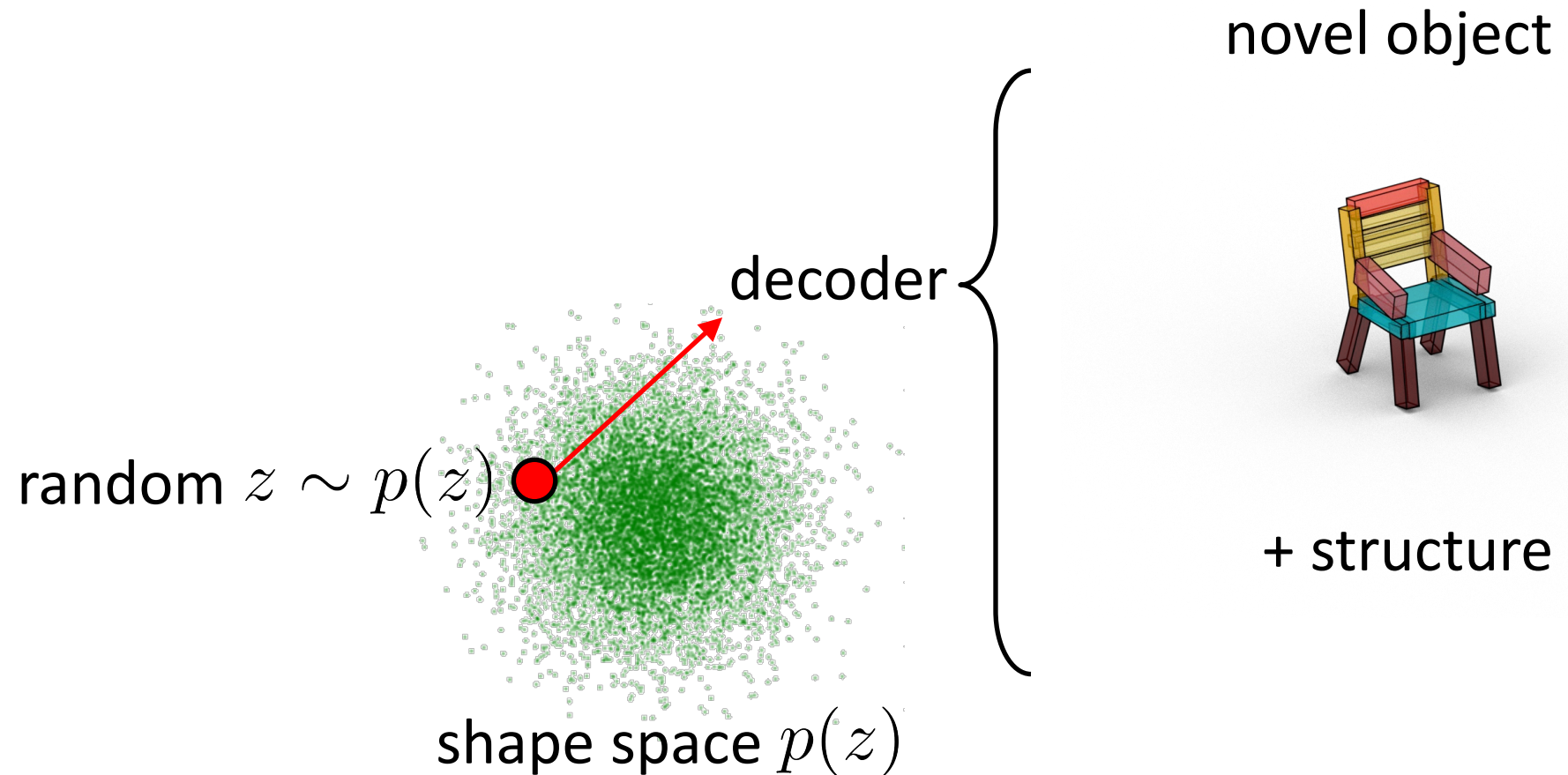




# Hierarchical Graph Decoder



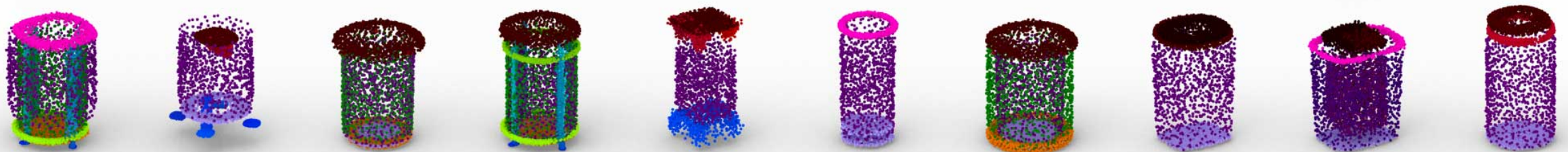
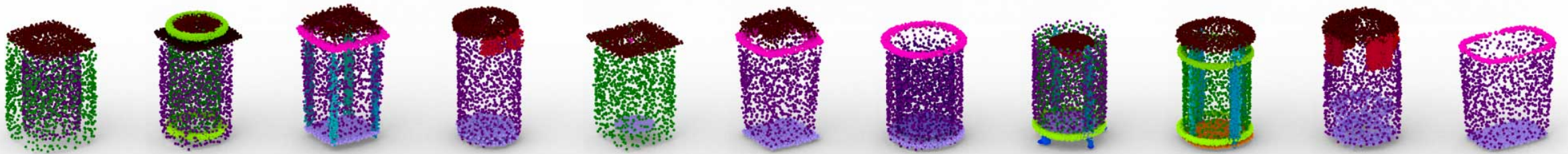
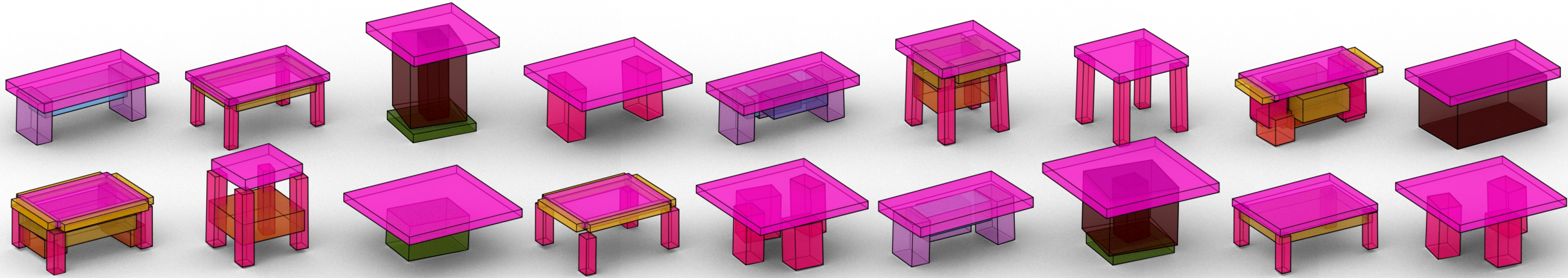
# Application 1: Generation



# Generation



# Generation



# Novelty

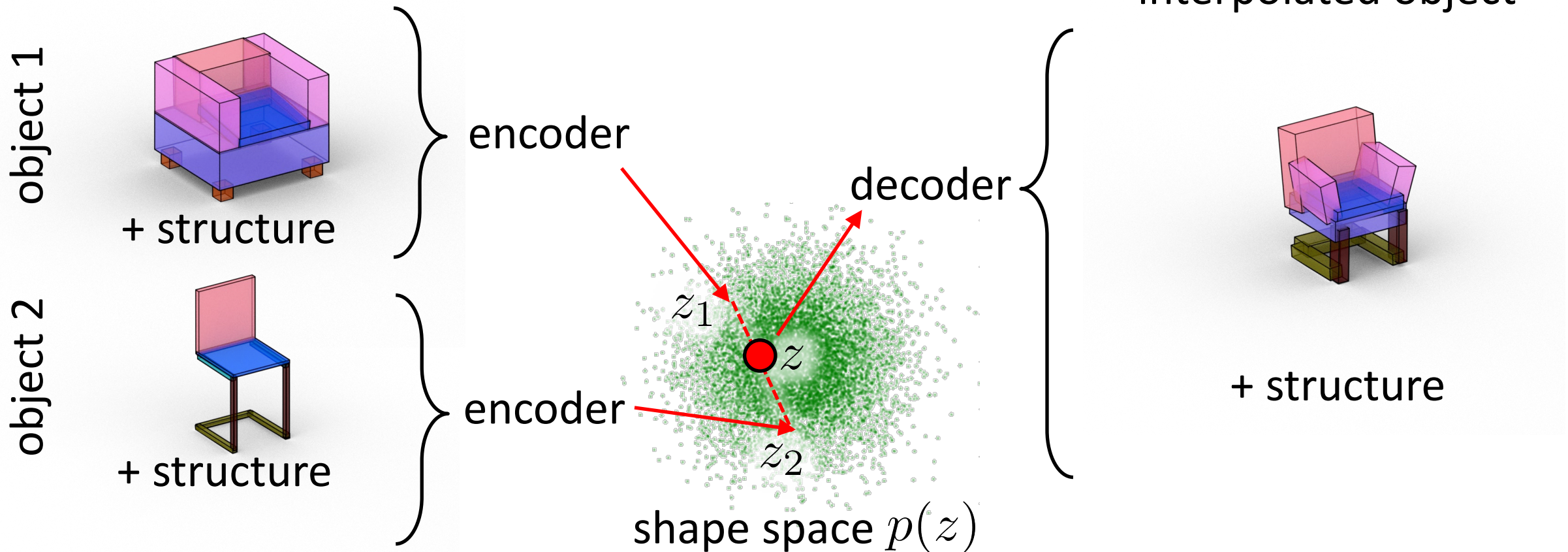
generated



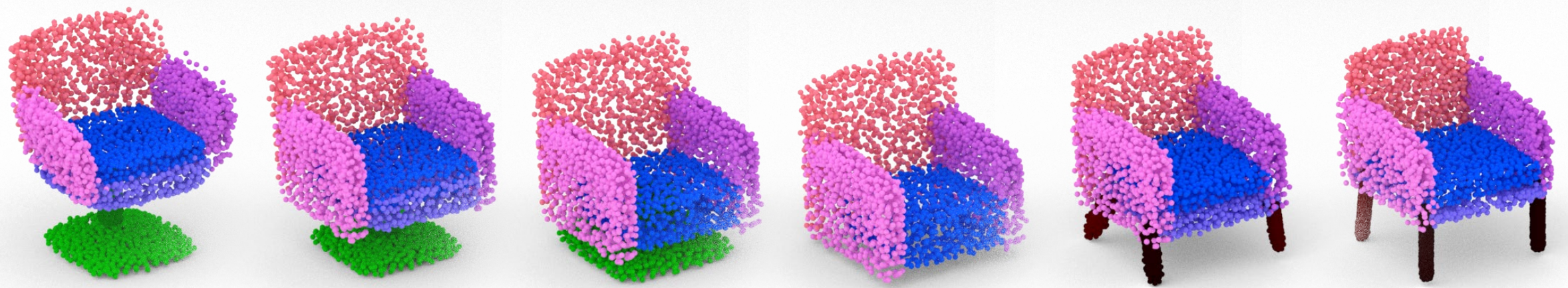
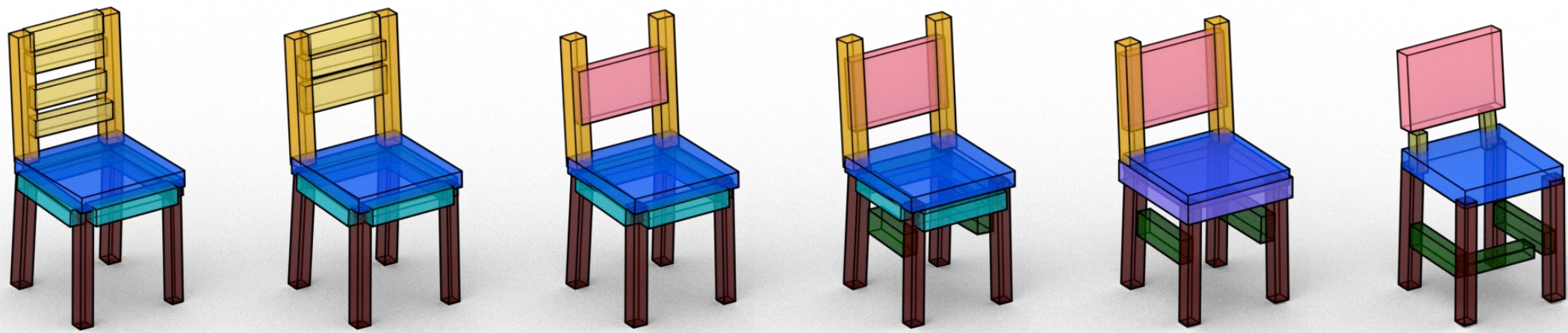
closest training samples



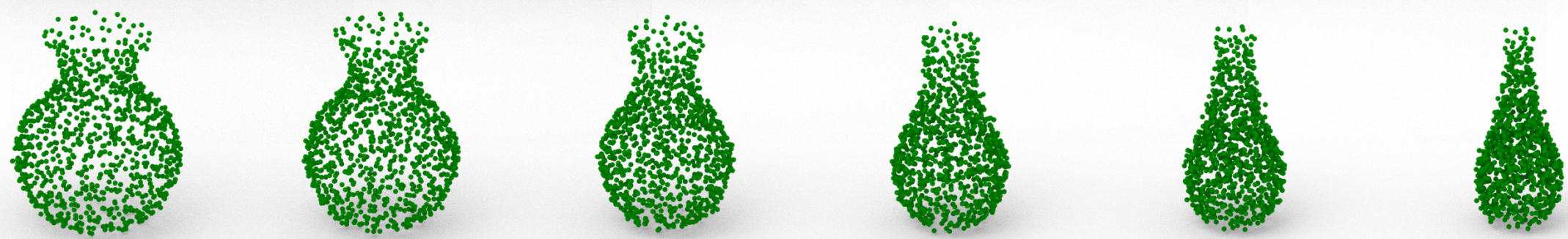
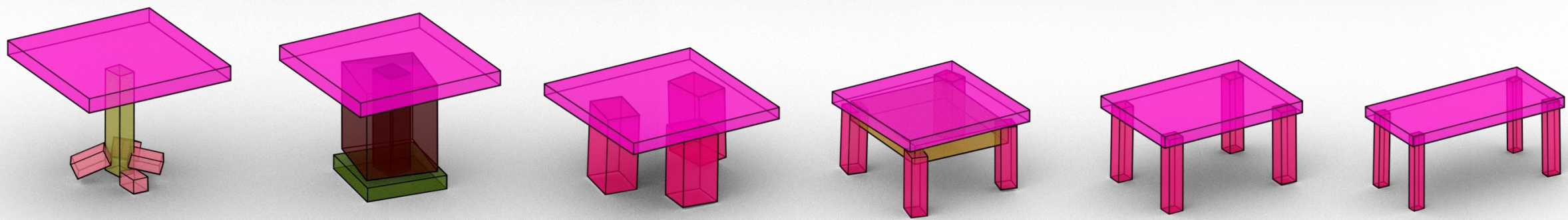
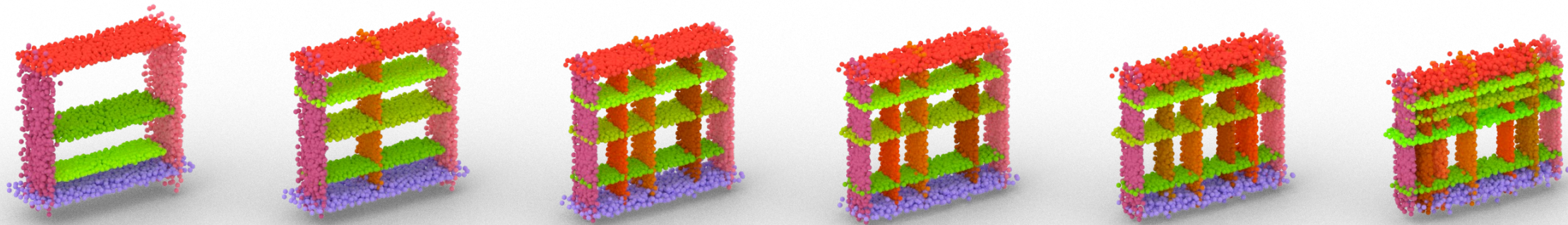
# Application 2: Interpolation



# Interpolation



# Interpolation



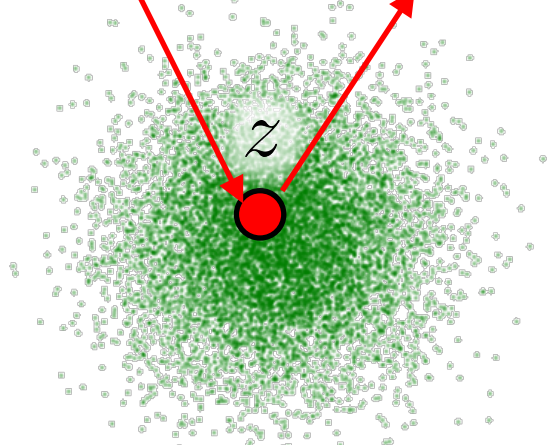
# Application 3: Scan Abstraction

partial scan



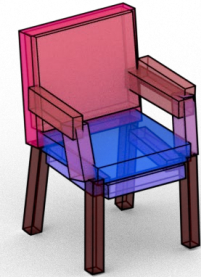
point cloud  
encoder

decoder



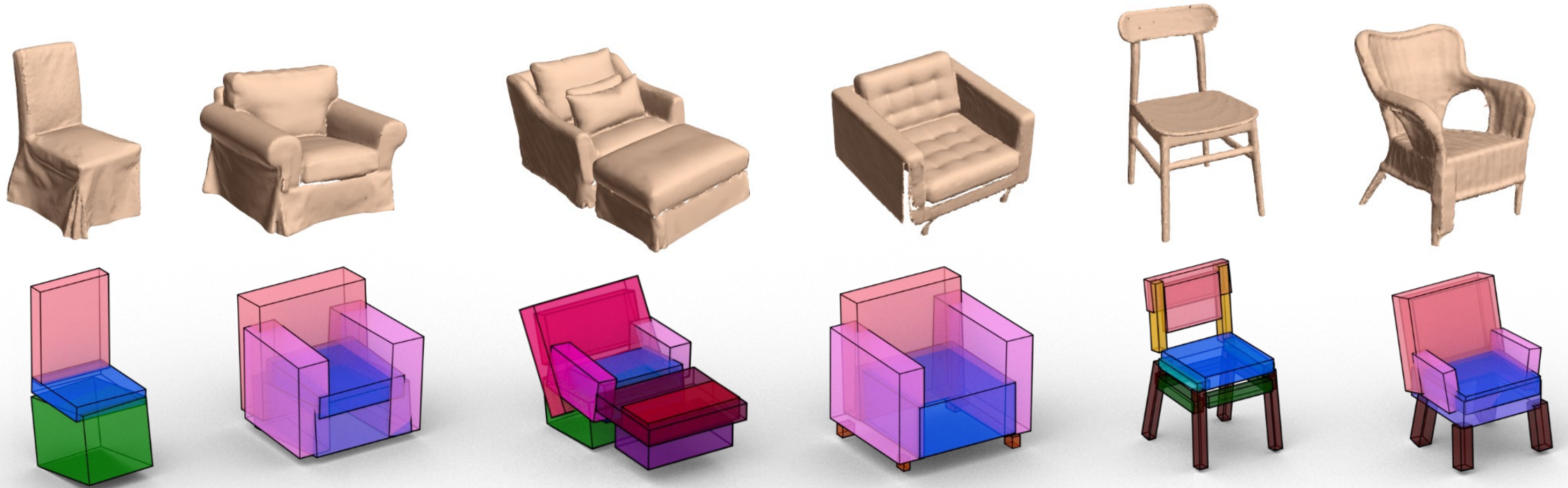
shape space  $p(z)$

reconstructed object

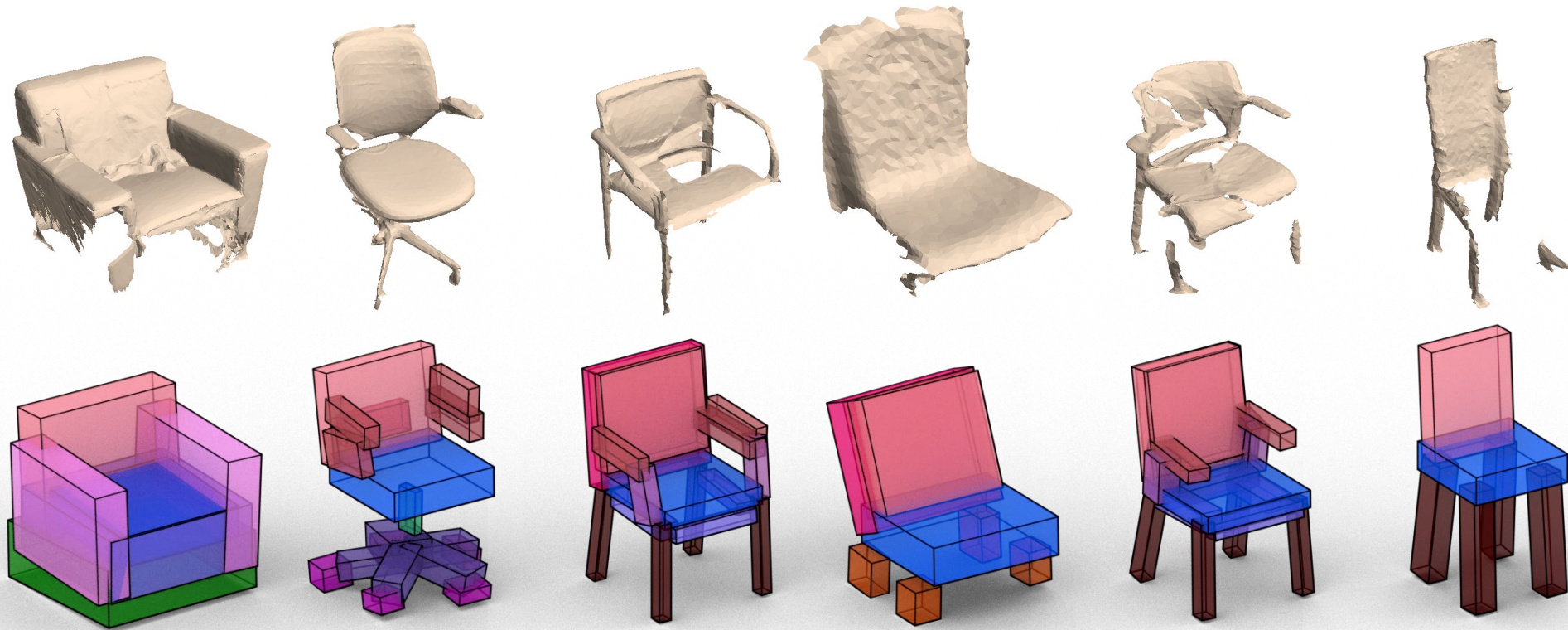


+ structure

# Abstraction of Full Scans

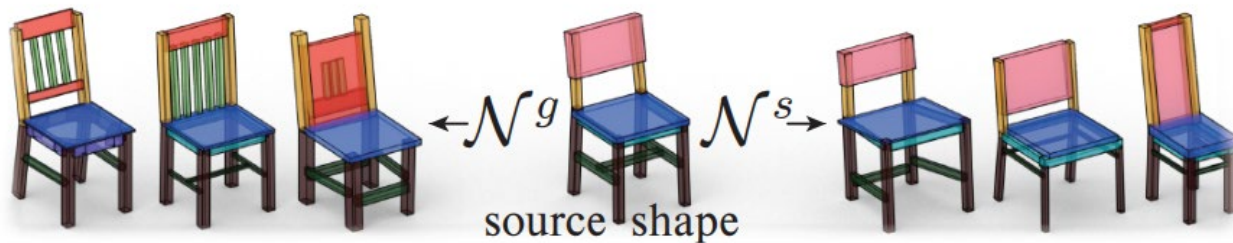


# Abstraction of Partial Scans



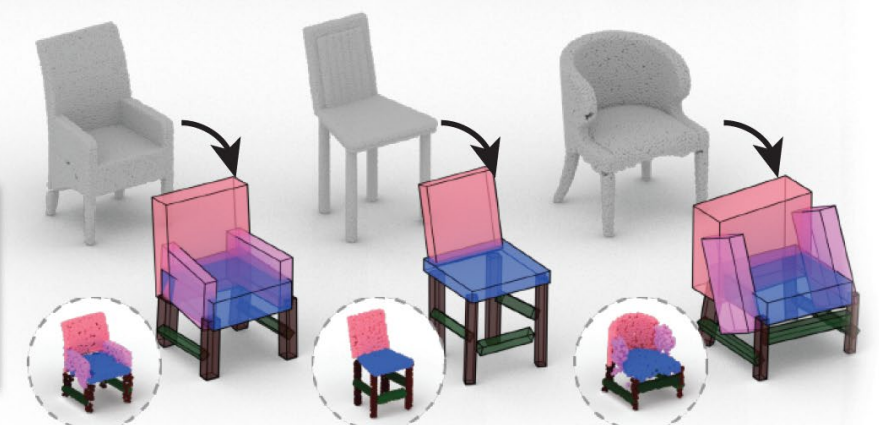
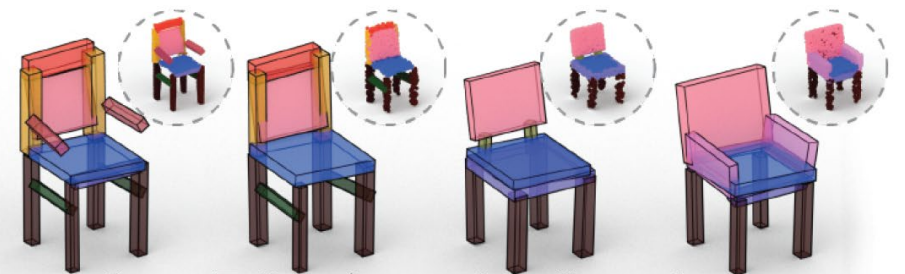
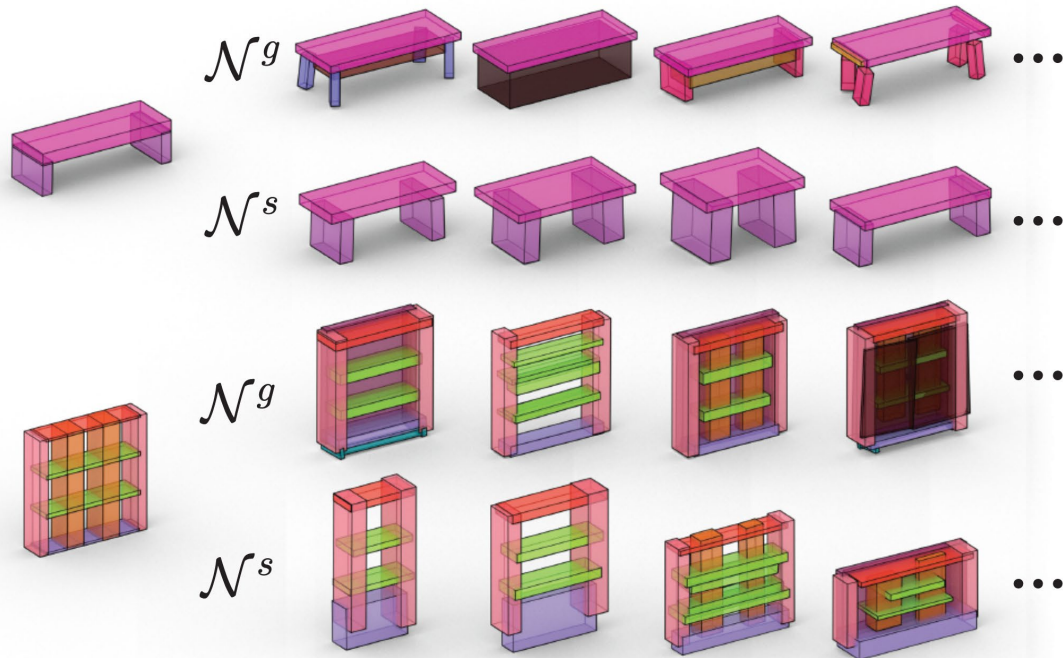
# Learning Shape Variations: Geometric and Structural

## Two Types of Shape Neighborhoods



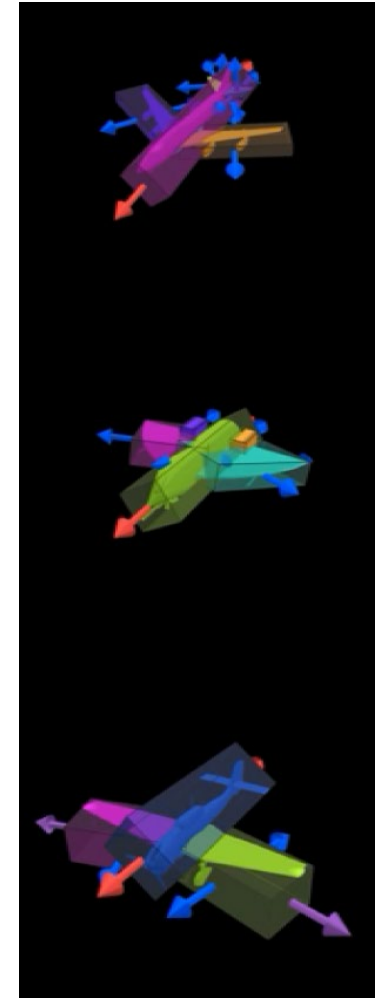
source shape

generated modifications



# Conclusion: What s a Shape Difference?

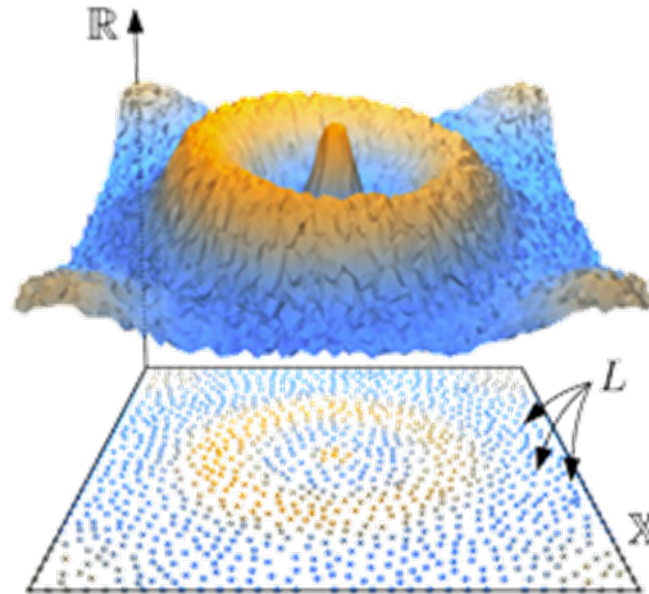
- The notion of shape differences under a map can be given a formal meaning useful in shape collection analysis.
- Differences can be used for shape interpolation, analogies (differences of differences), or reconstruction.
- Geometric shape differences can be mapped to/from language, using deep neural networks.
- Generating proper variations of a shape is an important tool for understanding its semantics
- Shape parts and compositional structure is an essential aspect of both shape generation as well as of understanding the function of shapes.



# The Course

# Data Analysis

- (Geometric and Topological Data) (Analysis)
- (Geometric and Topological) (Data Analysis)



# Topics Covered

- Visual data sets
- PCA
- CCA
- Spectral graph methods
- MDS; NLDR
- Intro to computational topology
- Homology and persistent homology
- Mapper
- 3D geometry representations
- Shape descriptors (shape context, spin images, HKS, WKS)
- Rigid alignments (ICP, RANSAC, geometric hashing)
- Non-rigid alignments (isometric, conformal)
- Shape correspondences
- Volumetric and Multiview CNNs
- Deep learning on point clouds
- Graph and mesh CNNs
- Functional Maps
- Shape differences
- Map networks and cycle consistency



# Topics Not Covered

- Factor analysis
- Independent components analysis
- Nearest neighbor search
- Locality sensitive hashing
- Clustering
- Topic modeling (LDA, etc)
- Tensor decompositions
- Dictionary learning
- Sparse recovery / compressive sensing
- Mixture models
- Non-negative matrix factorization
- Matrix completion
- Zig-zag persistence
- Reeb graphs
- Random graphs and network models

Please send e-mail about topics you'd like to see covered in future CS233 offerings

# The End

If you would like to pursue projects or research related to the topics of this class, please get in touch:

[guibas@cs.stanford.edu](mailto:guibas@cs.stanford.edu)

