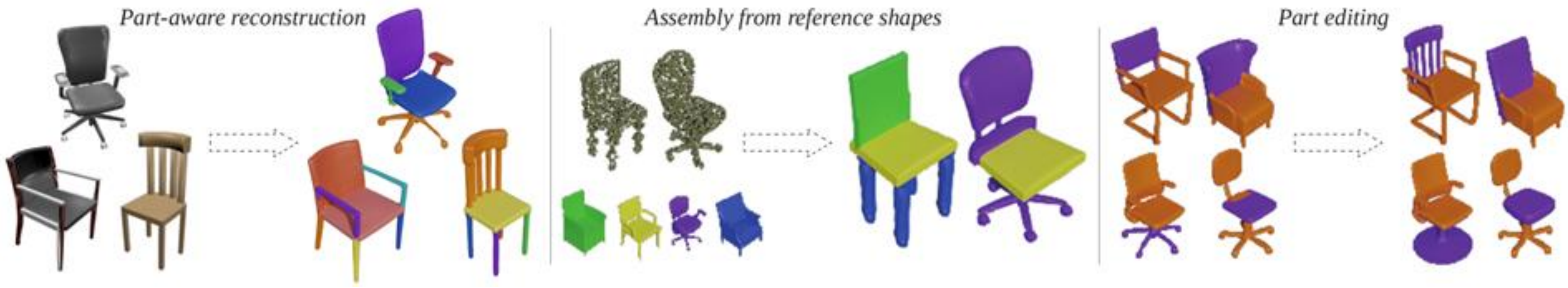
ANISE: Assembly-based Neural Implicit Surface rEconstruction

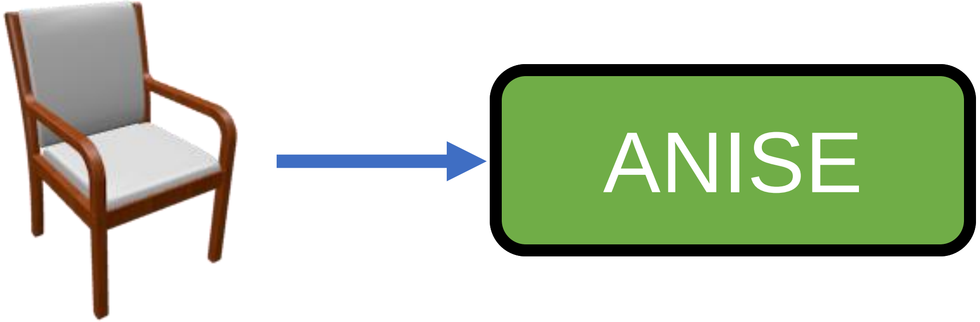


Dmitry Petrov[1†] **Matheus Gadelha[2]** **Radomír Měch[2]** **Evangelos Kalogerakis[1]**

[1] University of Massachusetts Amherst *[2] Adobe Research*

† Part of the work done as Adobe Research internship

Goal: reconstruct shapes as an assembly of neural implicit parts



Input RGB Image

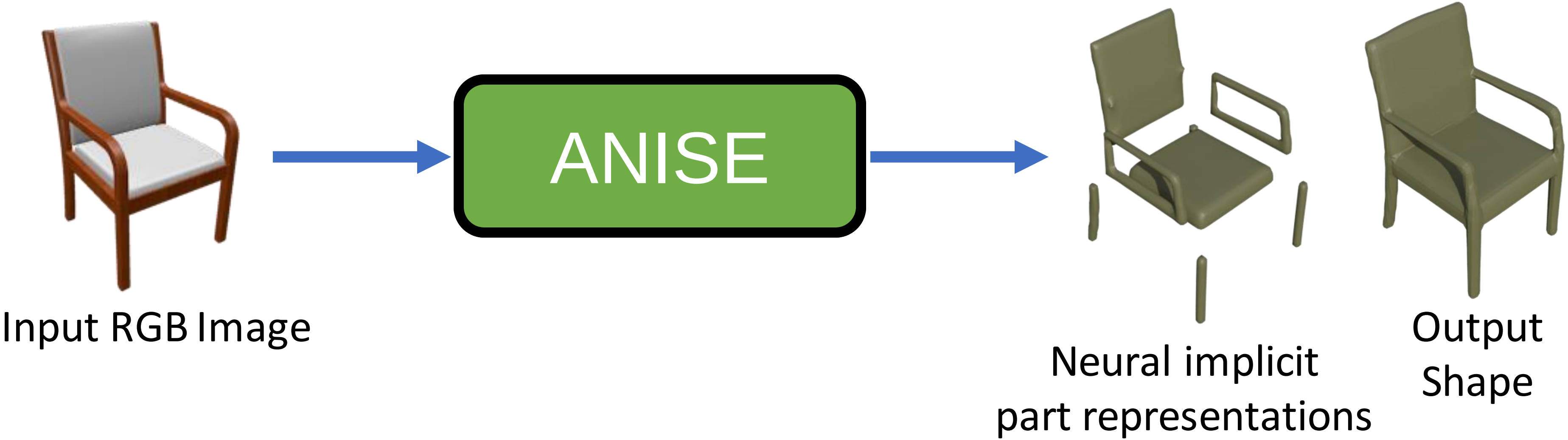
Goal: reconstruct shapes as an assembly of neural implicit parts

Input RGB Image

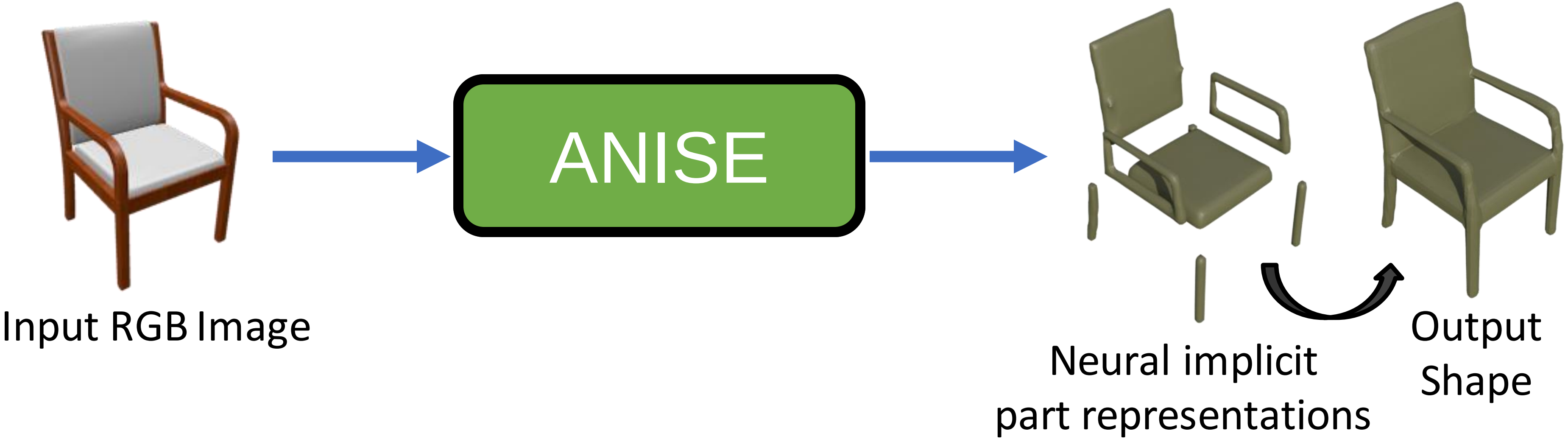
ANISE

Output Shape

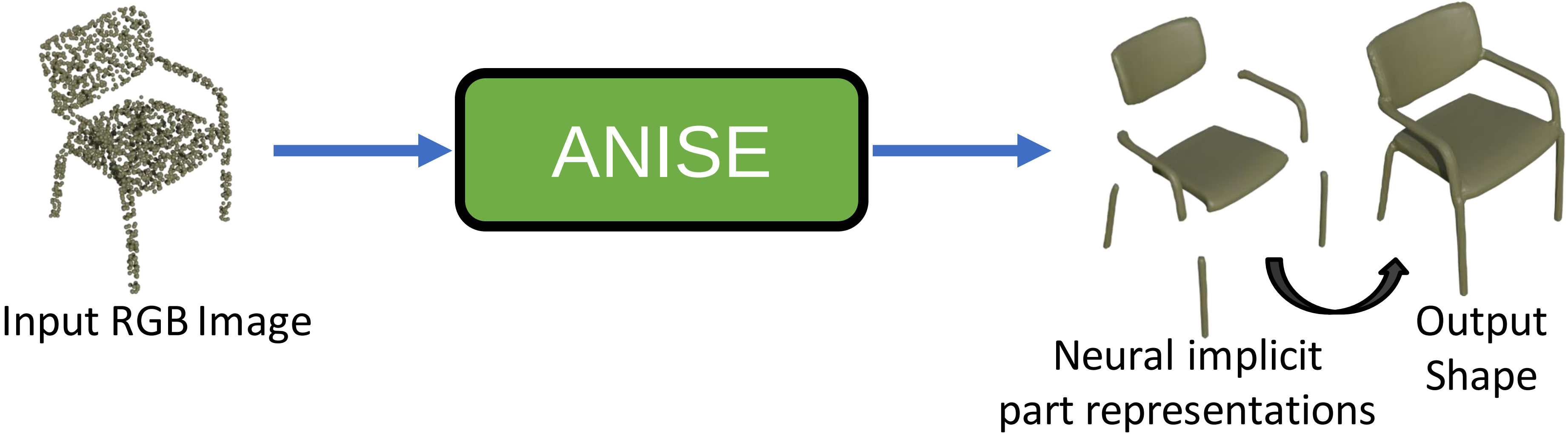
Goal: reconstruct shapes as an assembly of neural implicit parts



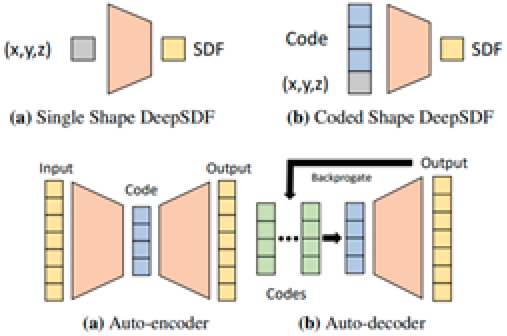
Goal: reconstruct shapes as an assembly of neural implicit parts



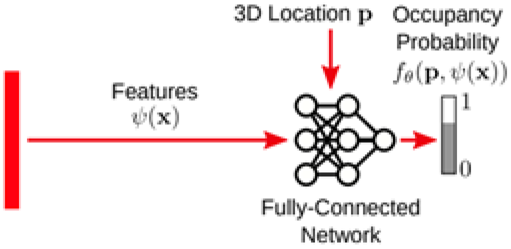
Goal: reconstruct shapes as an assembly of neural implicit parts



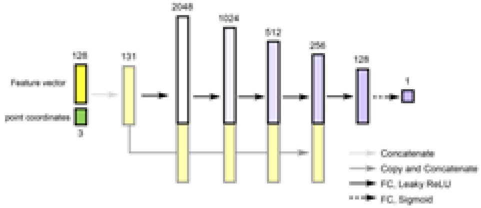
Prior work: neural implicits for 3D shape reconstruction



DeepSDF [Park et. al 2019]

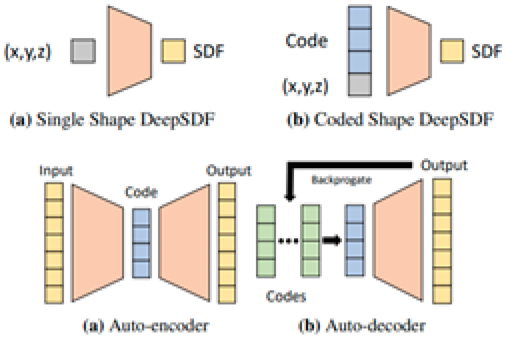


OccupancyNet [Mescheder et. al 2019]

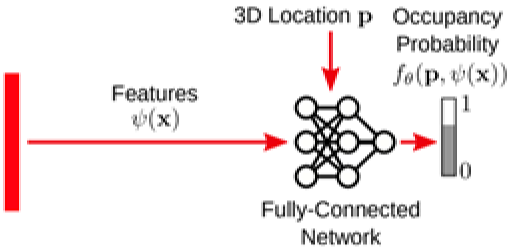


IM-NET [Chen et. al 2019]

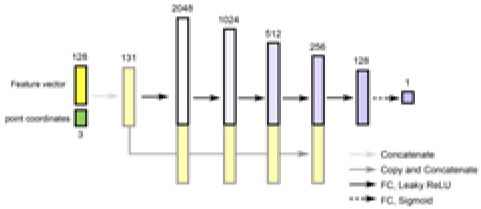
Prior work: neural implicits for 3D shape reconstruction



DeepSDF [Park et. al 2019]

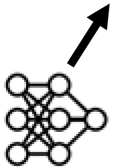


OccupancyNet [Mescheder et. al 2019]



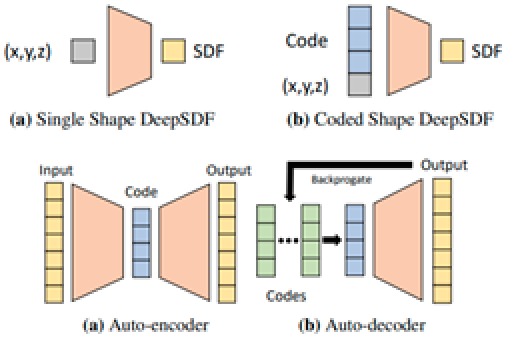
IM-NET [Chen et. al 2019]

$$f(x, y, z, \boldsymbol{\psi}) = scalar$$

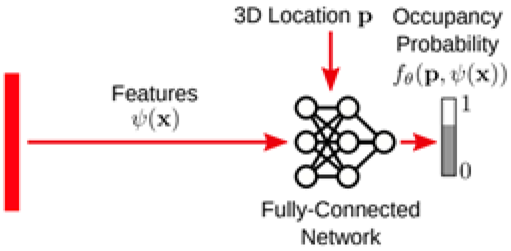


neural network

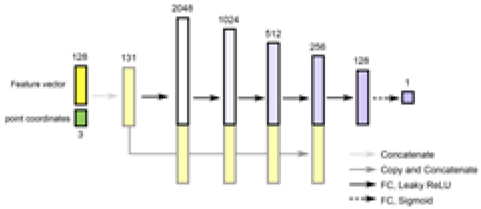
Prior work: neural implicits for 3D shape reconstruction



DeepSDF [Park et. al 2019]



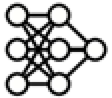
OccupancyNet [Mescheder et. al 2019]



IM-NET [Chen et. al 2019]

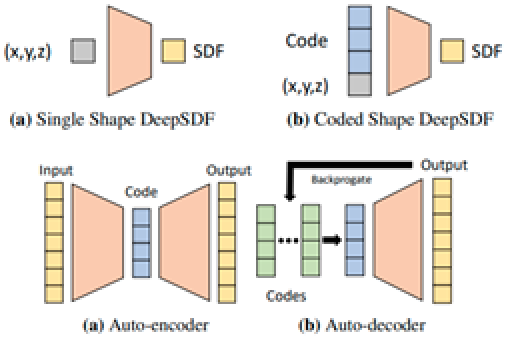
$$f(x, y, z, \boldsymbol{\psi}) = scalar$$

$$\underbrace{x, y, z}_{\text{3D point}}$$

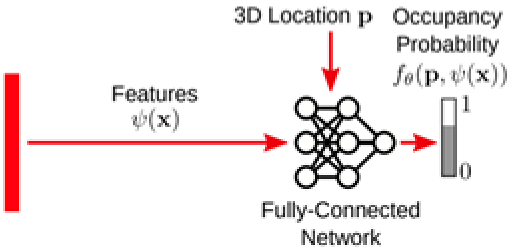


neural network

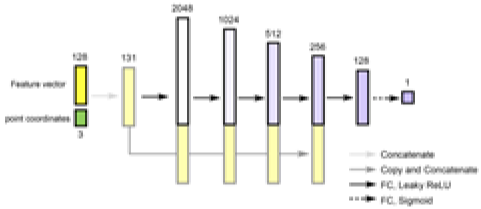
Prior work: neural implicits for 3D shape reconstruction



DeepSDF [Park et. al 2019]



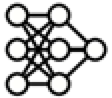
OccupancyNet [Mescheder et. al 2019]



IM-NET [Chen et. al 2019]

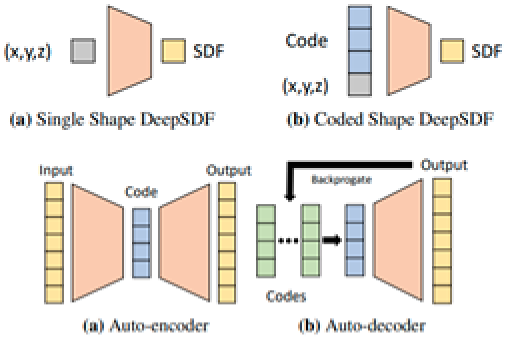
$$f(\underbrace{x, y, z}_{\text{3D point}}, \psi) = scalar$$

input encoding

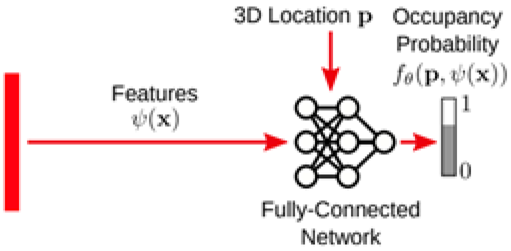


neural network

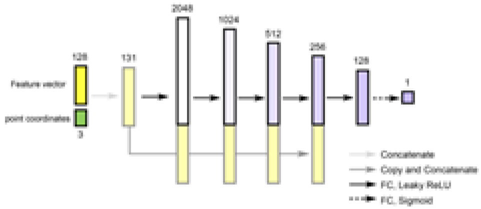
Prior work: neural implicits for 3D shape reconstruction



DeepSDF [Park et. al 2019]

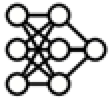


OccupancyNet [Mescheder et. al 2019]



IM-NET [Chen et. al 2019]

$$f(\underbrace{x, y, z}_{\text{3D point}}, \boldsymbol{\psi}) = scalar$$

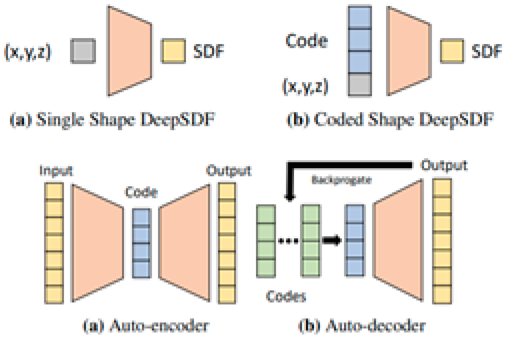


neural network

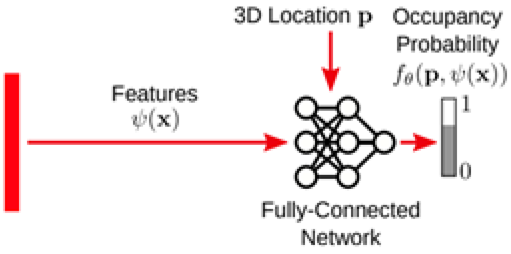
input encoding

e.g., SDF, UDF,
occupancy

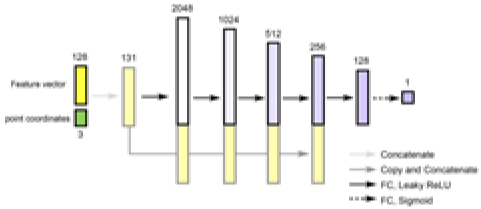
Prior work: neural implicits for 3D shape reconstruction



DeepSDF [Park et. al 2019]

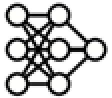


OccupancyNet [Mescheder et. al 2019]



IM-NET [Chen et. al 2019]

$$f(x, y, z, \boldsymbol{\psi}) = scalar$$

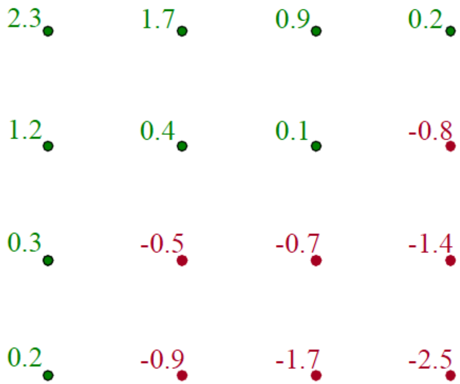


neural network

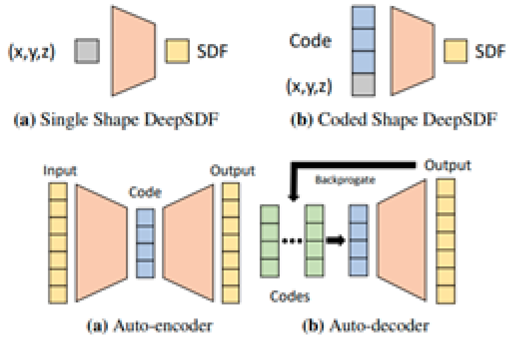
3D point

input encoding

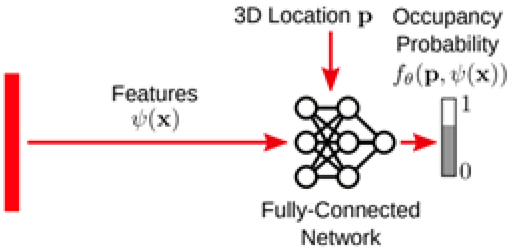
e.g., SDF, UDF, occupancy



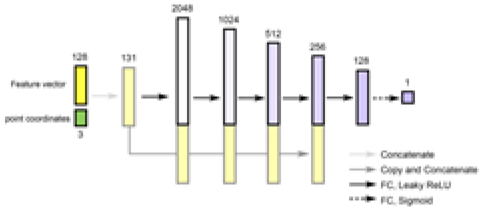
Prior work: neural implicits for 3D shape reconstruction



DeepSDF [Park et. al 2019]

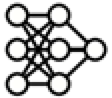


OccupancyNet [Mescheder et. al 2019]



IM-NET [Chen et. al 2019]

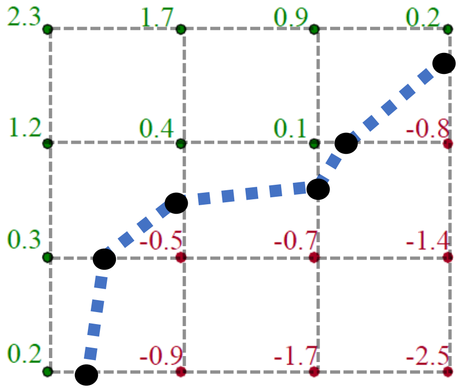
$$f(\underbrace{x,y,z}_{\text{3D point}},\boldsymbol{\psi}) = scalar$$



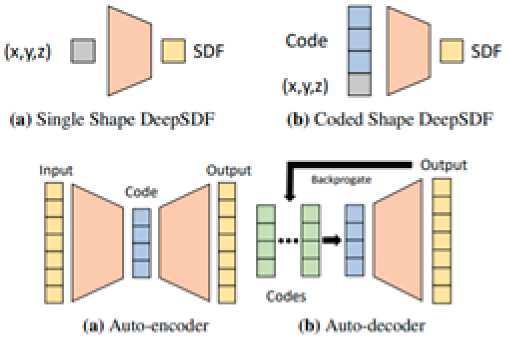
neural network

input encoding

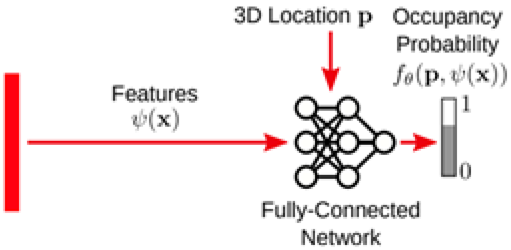
e.g., SDF, UDF, occupancy



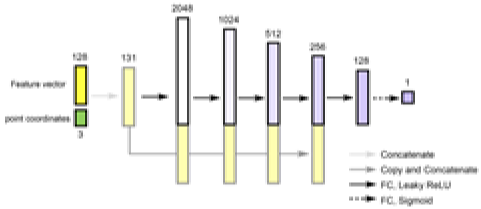
Prior work: neural implicits for 3D shape reconstruction



DeepSDF [Park et. al 2019]

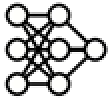


OccupancyNet [Mescheder et. al 2019]



IM-NET [Chen et. al 2019]

$$f(x, y, z, \boldsymbol{\psi}) = scalar$$

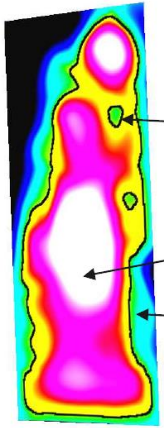


neural network

3D point

input encoding

e.g., SDF, UDF, occupancy



Implicit Surface:

$f(x, y, z, \boldsymbol{\psi}) = 0$

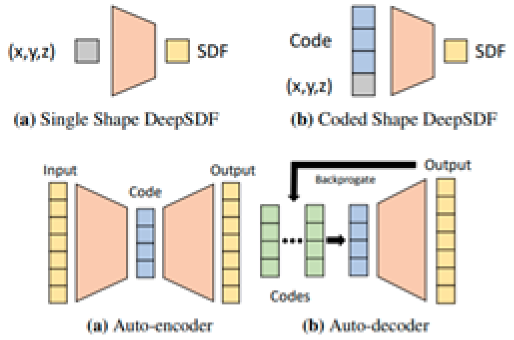
Inside:

$f(x, y, z, \boldsymbol{\psi}) < 0$

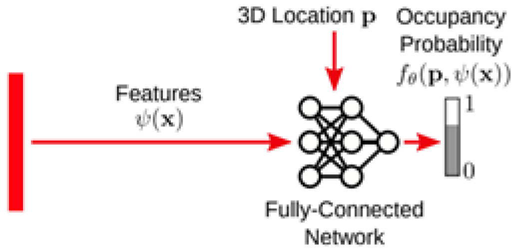
Outside:

$f(x, y, z, \boldsymbol{\psi}) > 0$

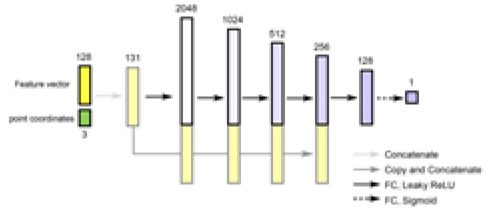
Prior work: neural implicits for 3D shape reconstruction



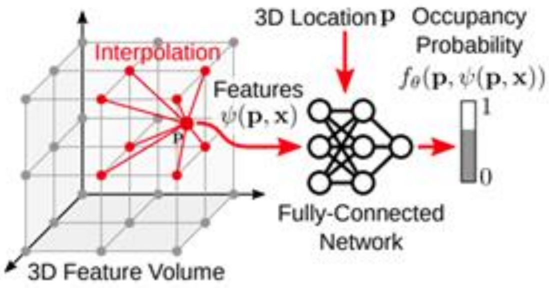
DeepSDF [Park et. al 2019]



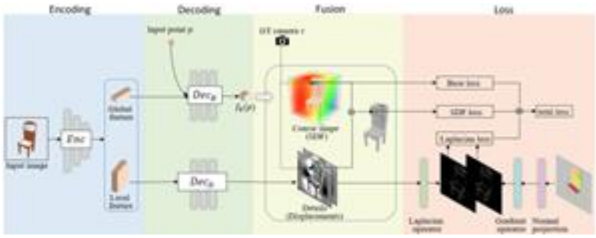
OccupancyNet [Mescheder et. al 2019]



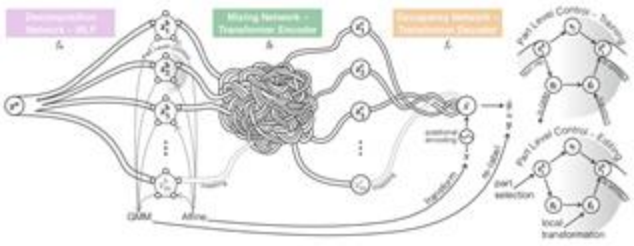
IM-NET [Chen et. al 2019]



ConvOccNet [Mescheder et. al 2020]

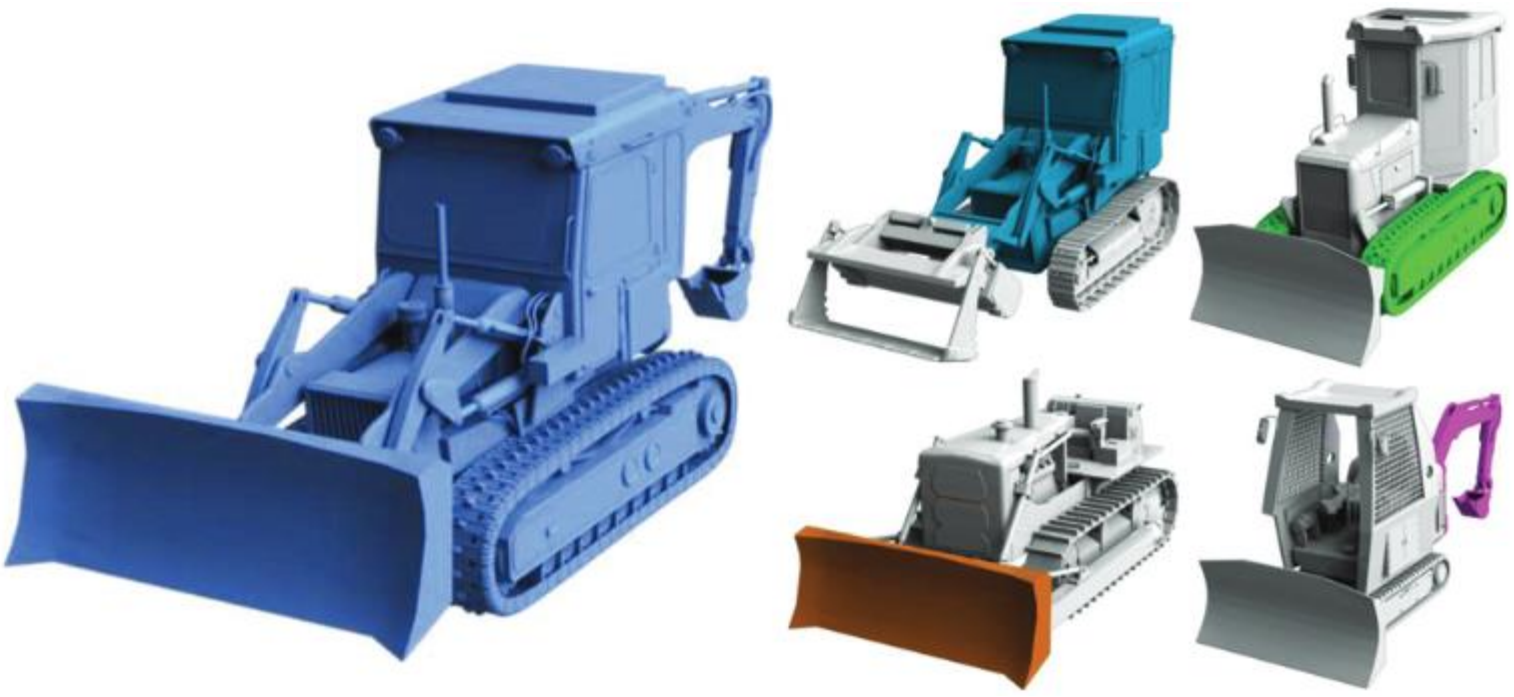


D2IM-Net [Li et. al 2021]



SPAGHETTI [Hertz et al. 2022]

Prior work: assembly-based modeling

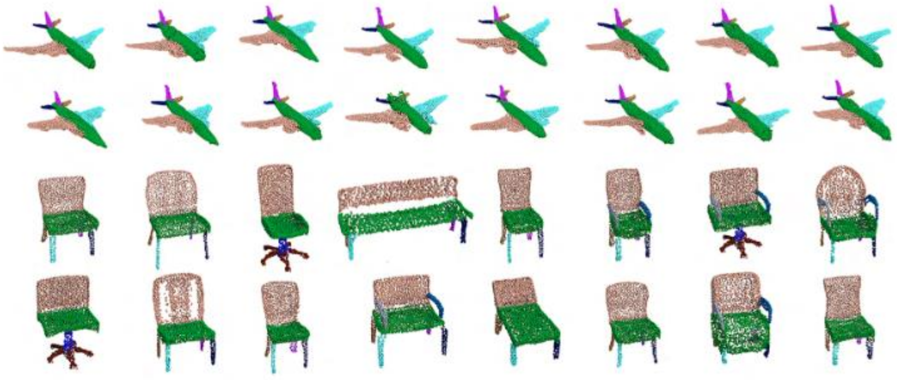


A probabilistic model of Component-based Shape Synthesis [Kalogerakis et al. 2012]

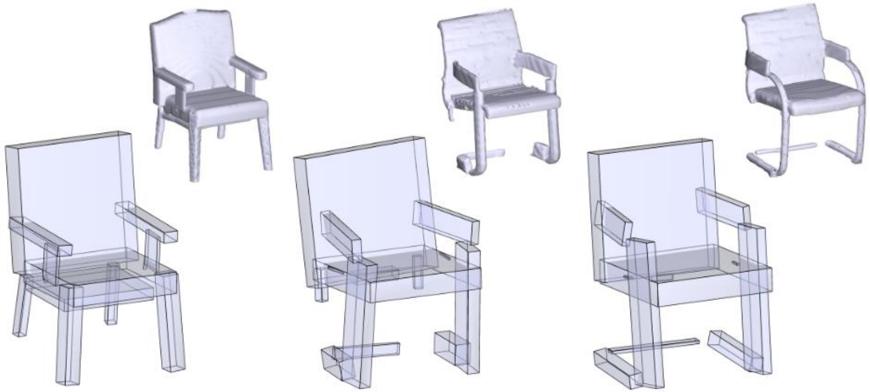
Prior work: assembly-based modeling

Probabilistic Reasoning for Assembly-Based 3D Modeling [Chaudhuri et al. 2011]

Prior work: supervised composite shape modeling



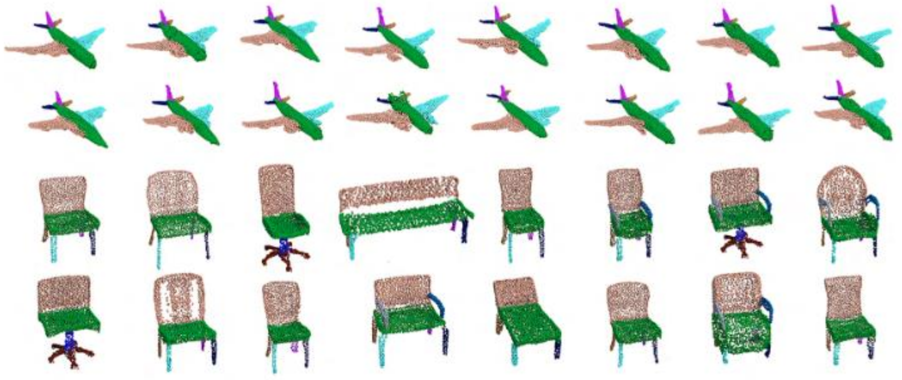
Beta Shape Machine [Huang et al. 2015]



GRASS [Li et al. 2017]

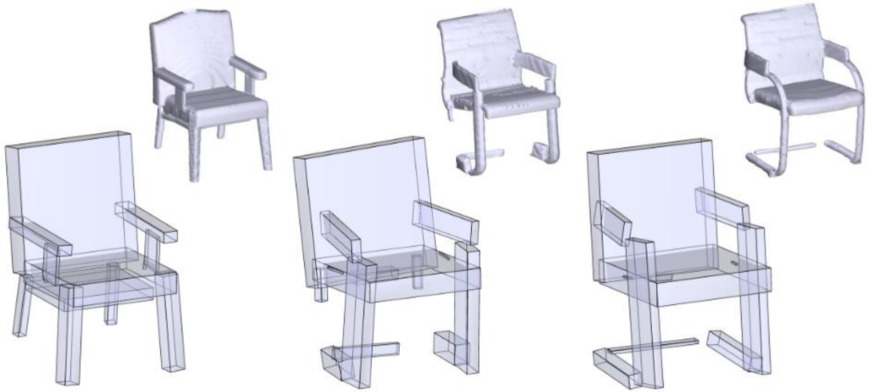
Composite Shape Modeling via Latent Space Factorization [Dubrovina et al. 2019]

Prior work: supervised composite shape modeling



Beta Shape Machine [Huang et al. 2015]

Points



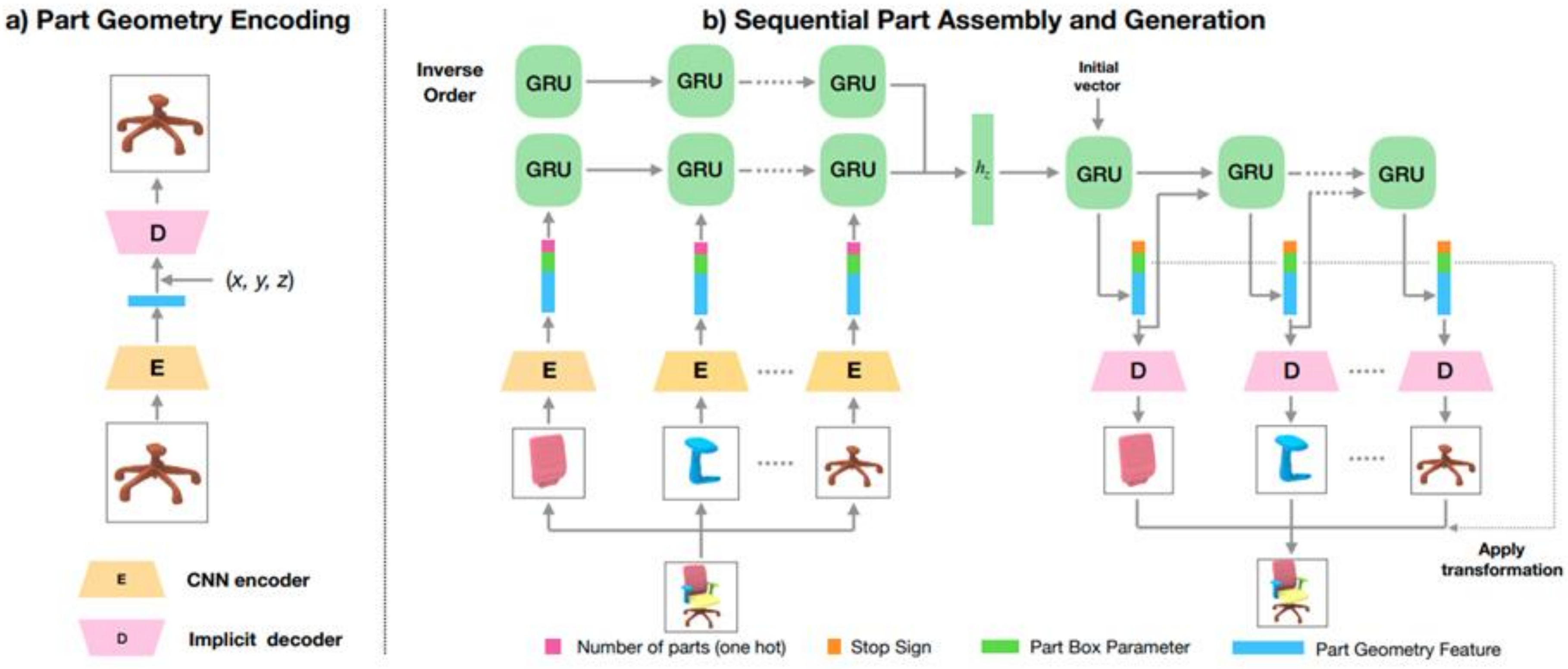
GRASS [Li et al. 2017]

Boxes/Voxels

Composite Shape Modeling via Latent Space Factorization [Dubrovina et al. 2019]

Voxels

Prior work: supervised composite shape modeling



PQ-NET [Wu et al. 2020]

ANISE: Contributions

- Assembly is treated as a **set of part implicits** -- **order does not matter**

ANISE: Contributions

- Assembly is treated as a **set of part implicits** -- **order does not matter**
- **End-to-end training**: supervision after assembling parts into full shape

ANISE: Contributions

- Assembly is treated as a **set of part implicits** -- **order does not matter**
- **End-to-end training**: supervision after assembling parts into full shape
- **Coarse-to-fine generation** – predict structure, then part geometry

ANISE: Contributions

- Assembly is treated as a **set of part implicits -- order does not matter**
- **End-to-end training**: supervision after assembling parts into full shape
- **Coarse-to-fine generation** – predict structure, then part geometry

ANISE: neural modules

Input

ANISE: neural modules

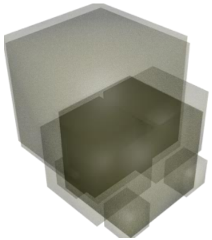
Input

Structure prediction

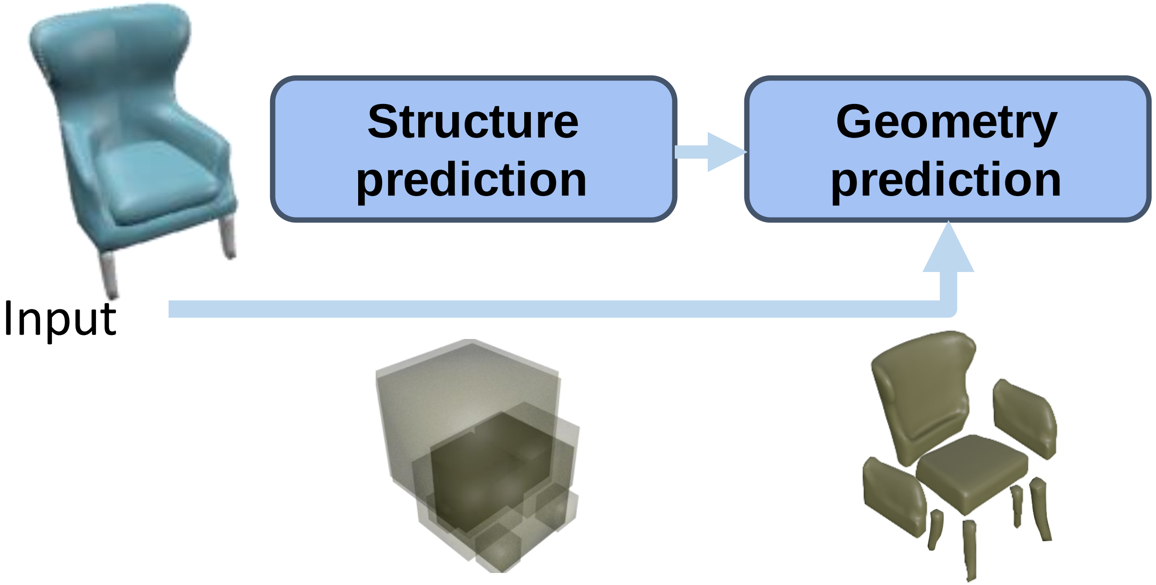
ANISE: neural modules

Input

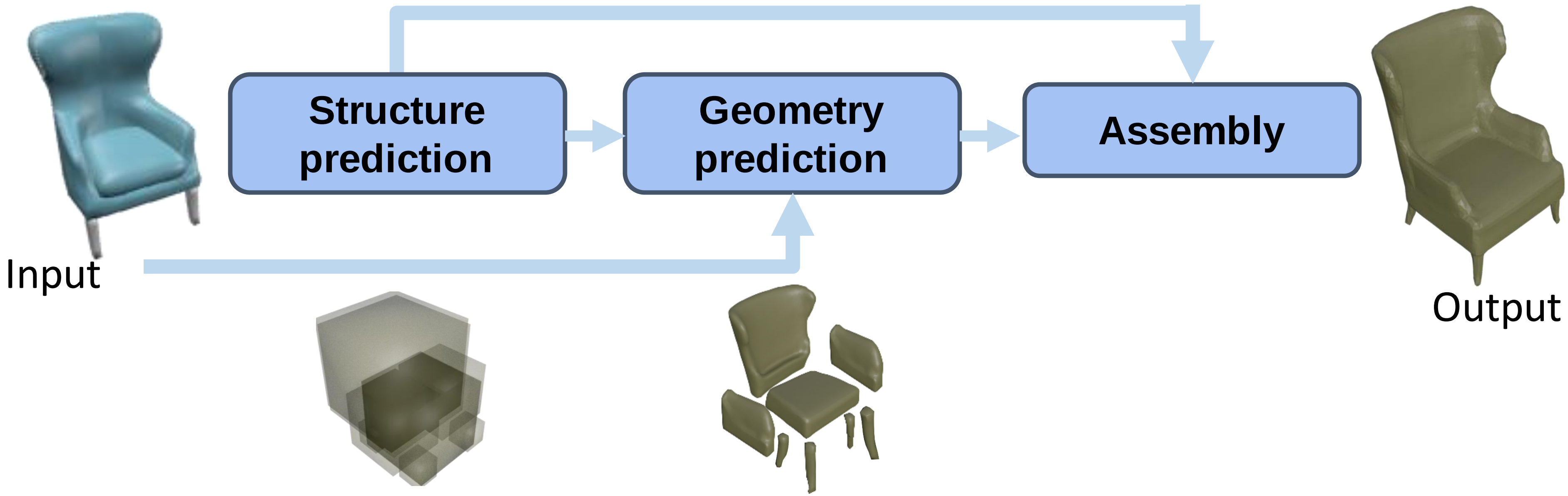
Structure prediction



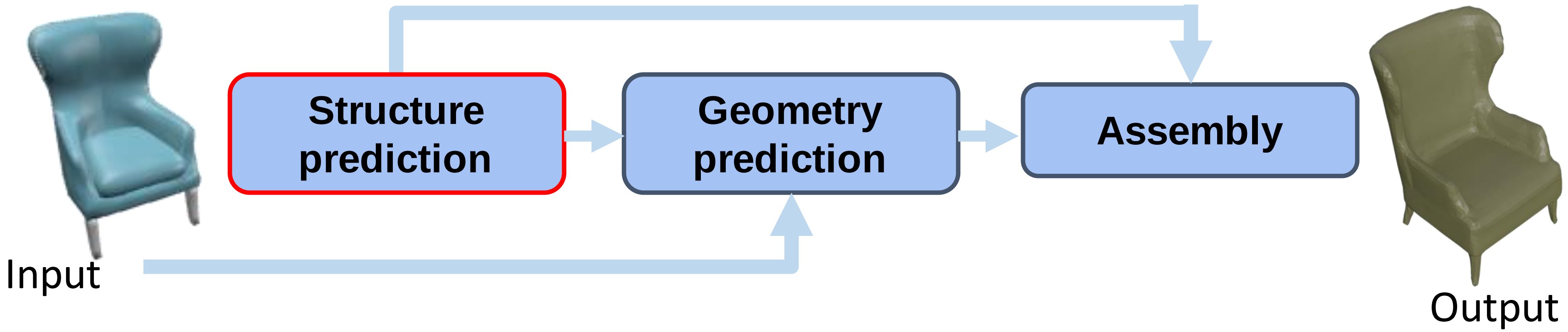
ANISE: neural modules



ANISE: neural modules



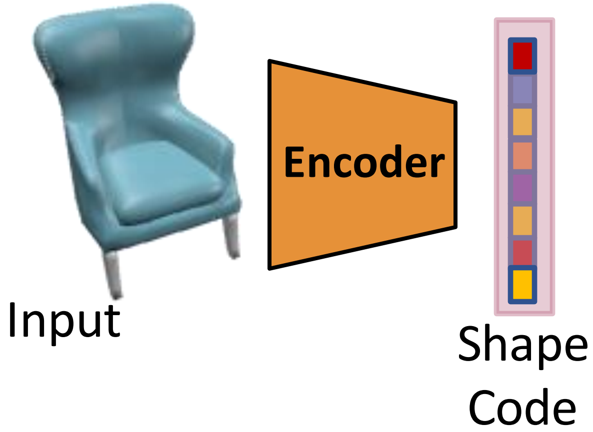
ANISE: neural modules



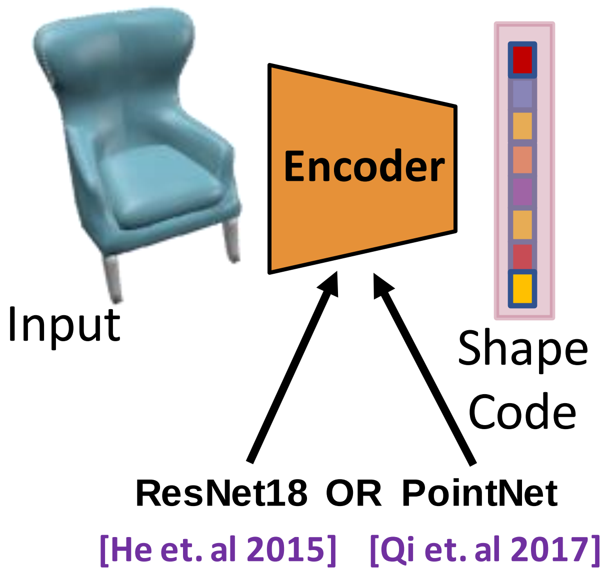
Structure prediction module

Input

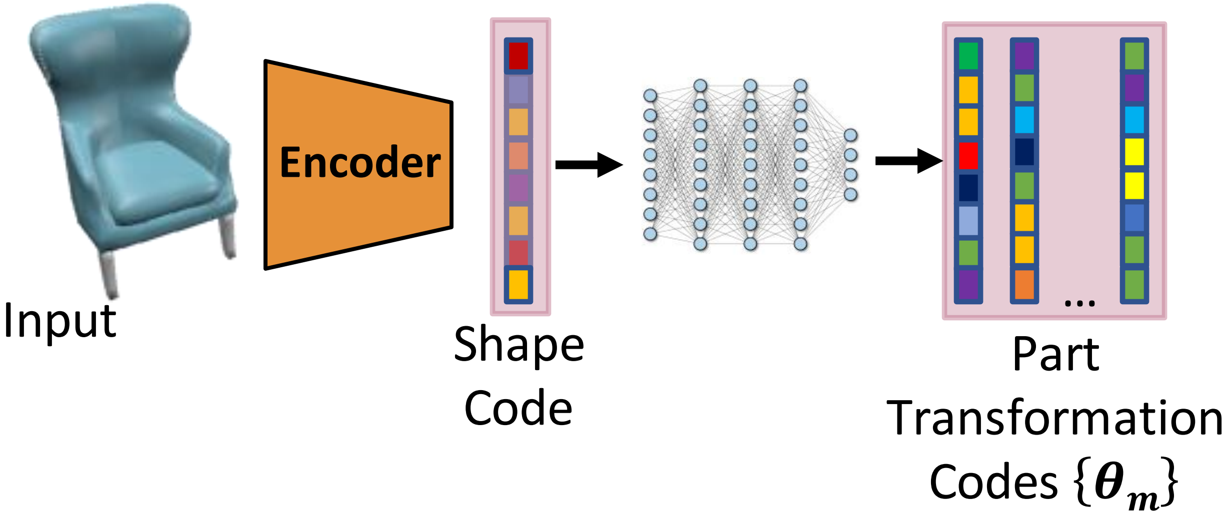
Structure prediction module



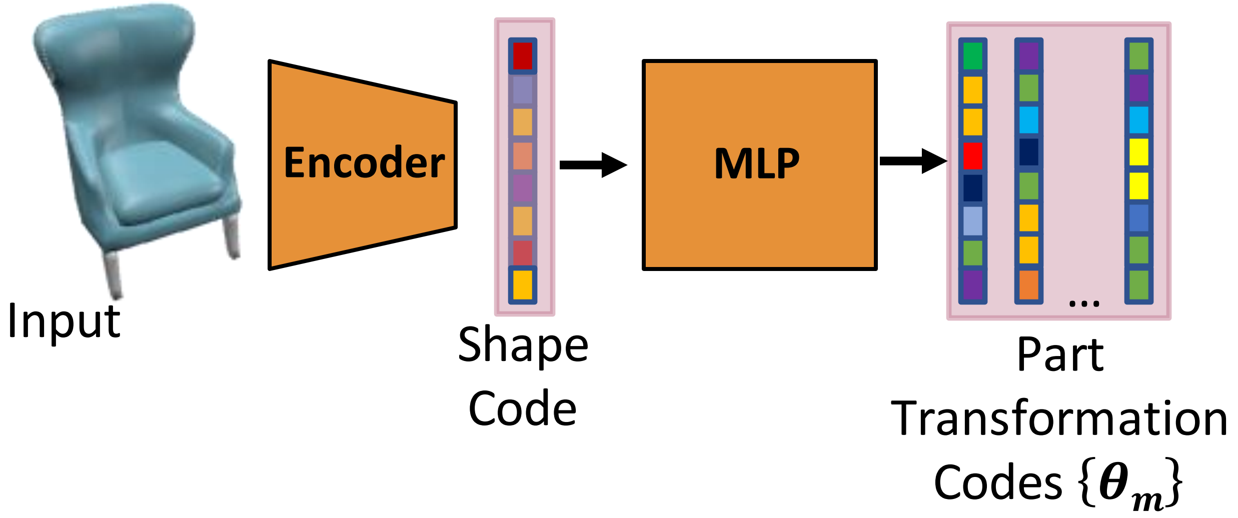
Structure prediction module



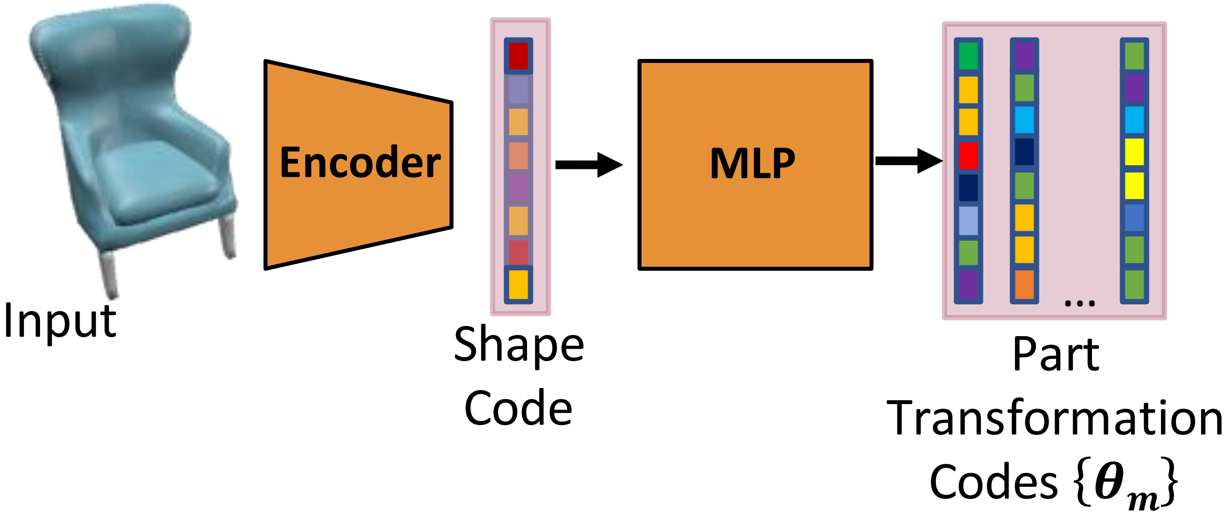
Structure prediction module



Structure prediction module

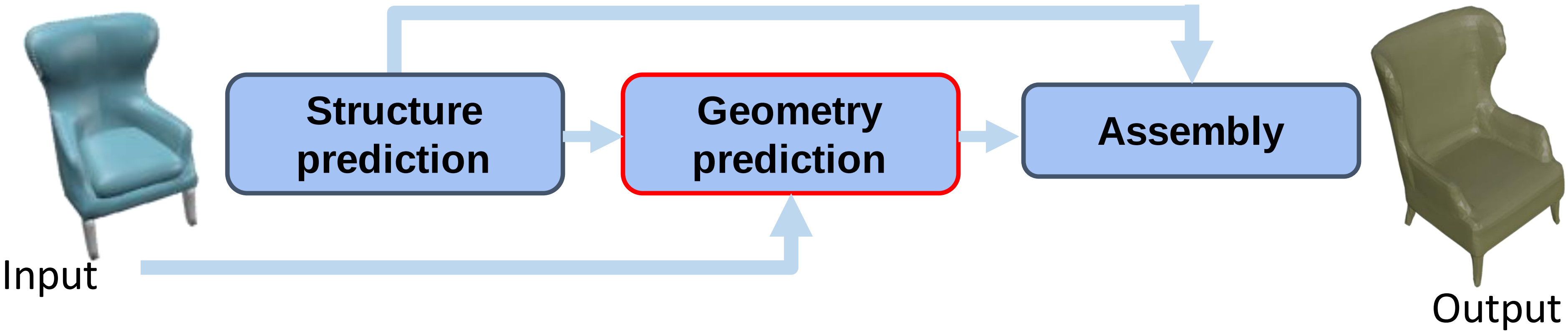


Structure prediction module

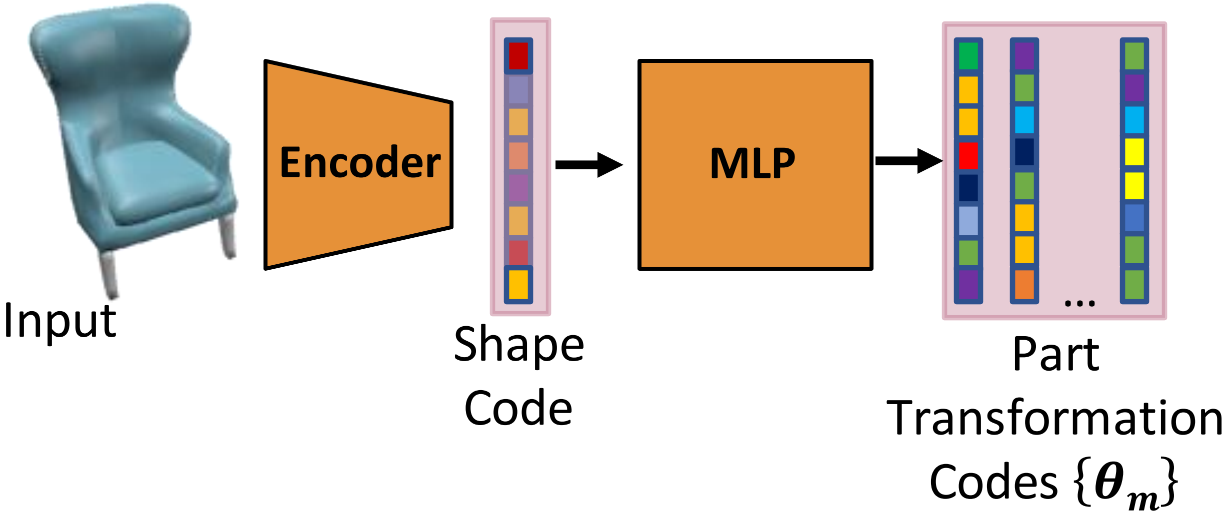


Max number=10

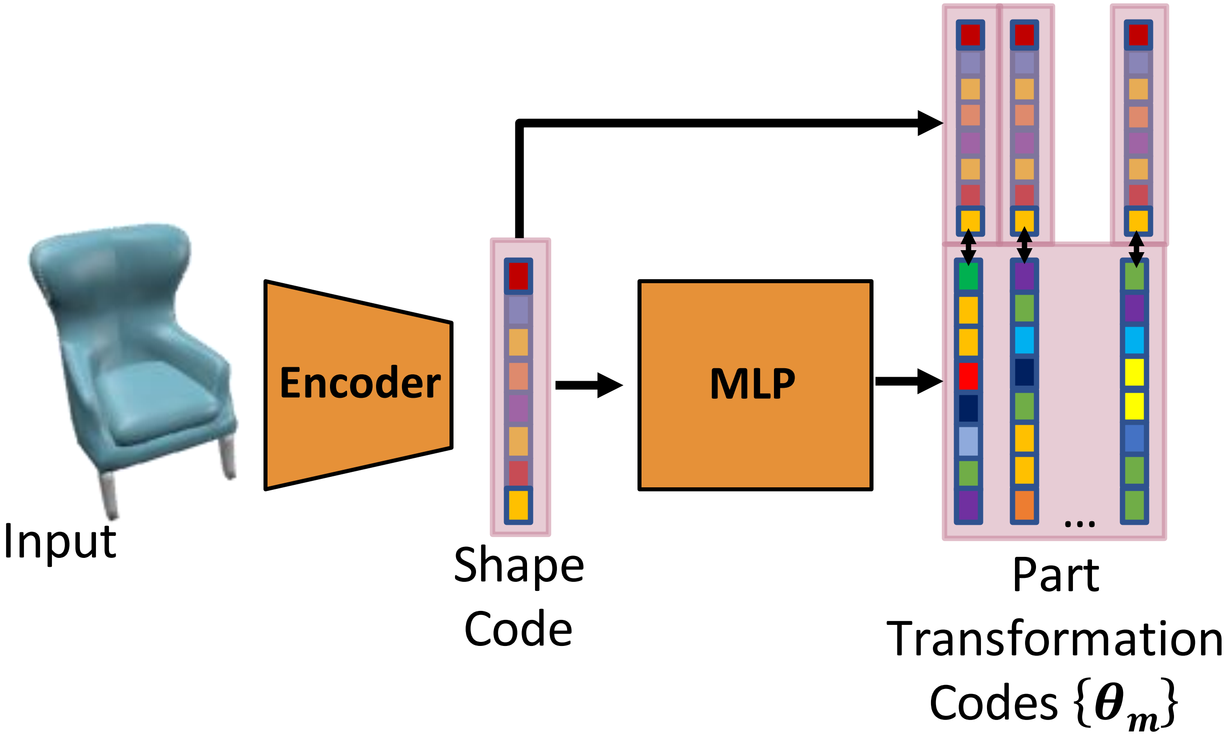
ANISE: neural modules



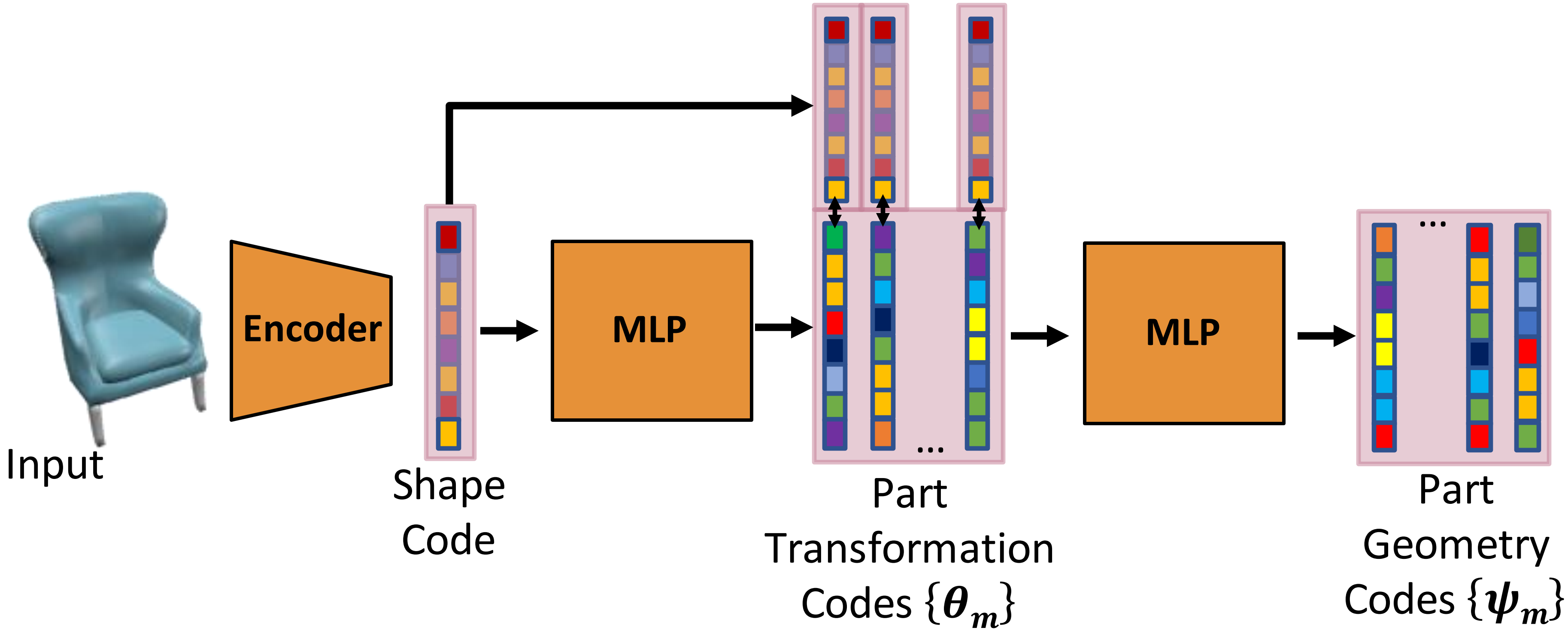
Geometry prediction module



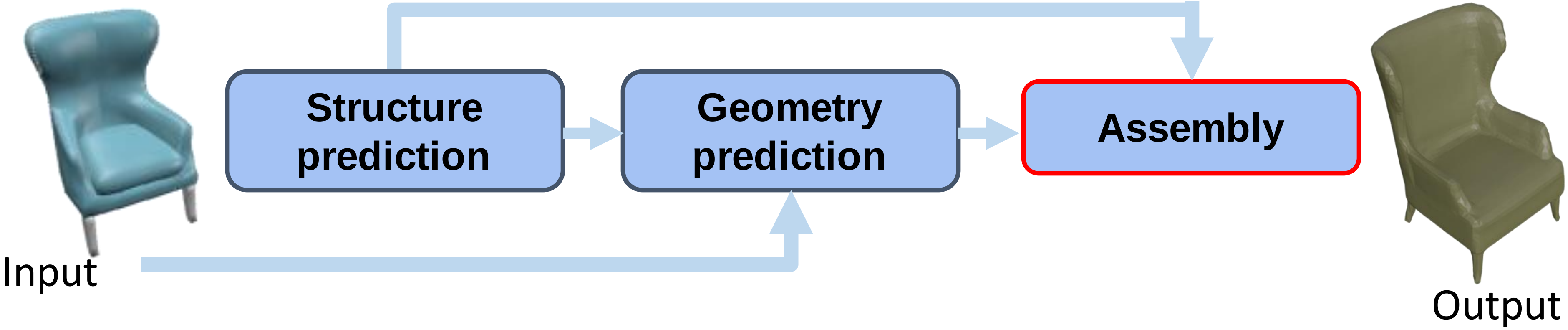
Geometry prediction module



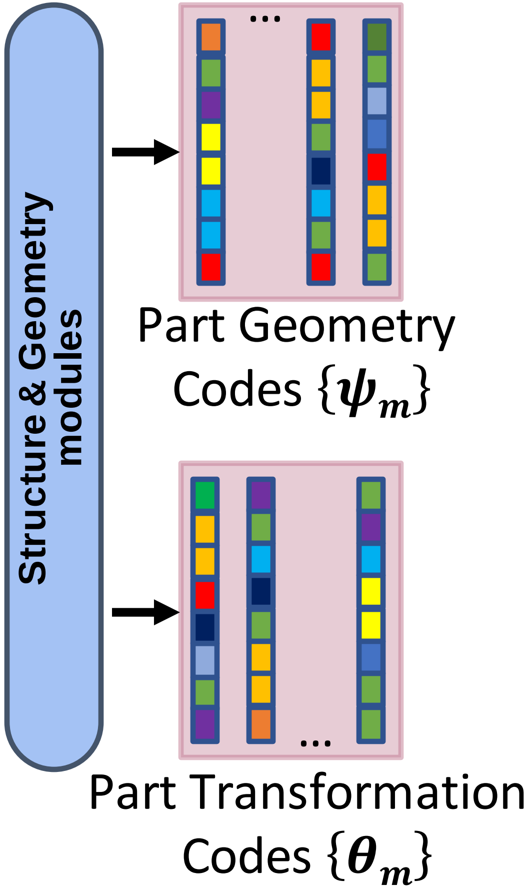
Geometry prediction module



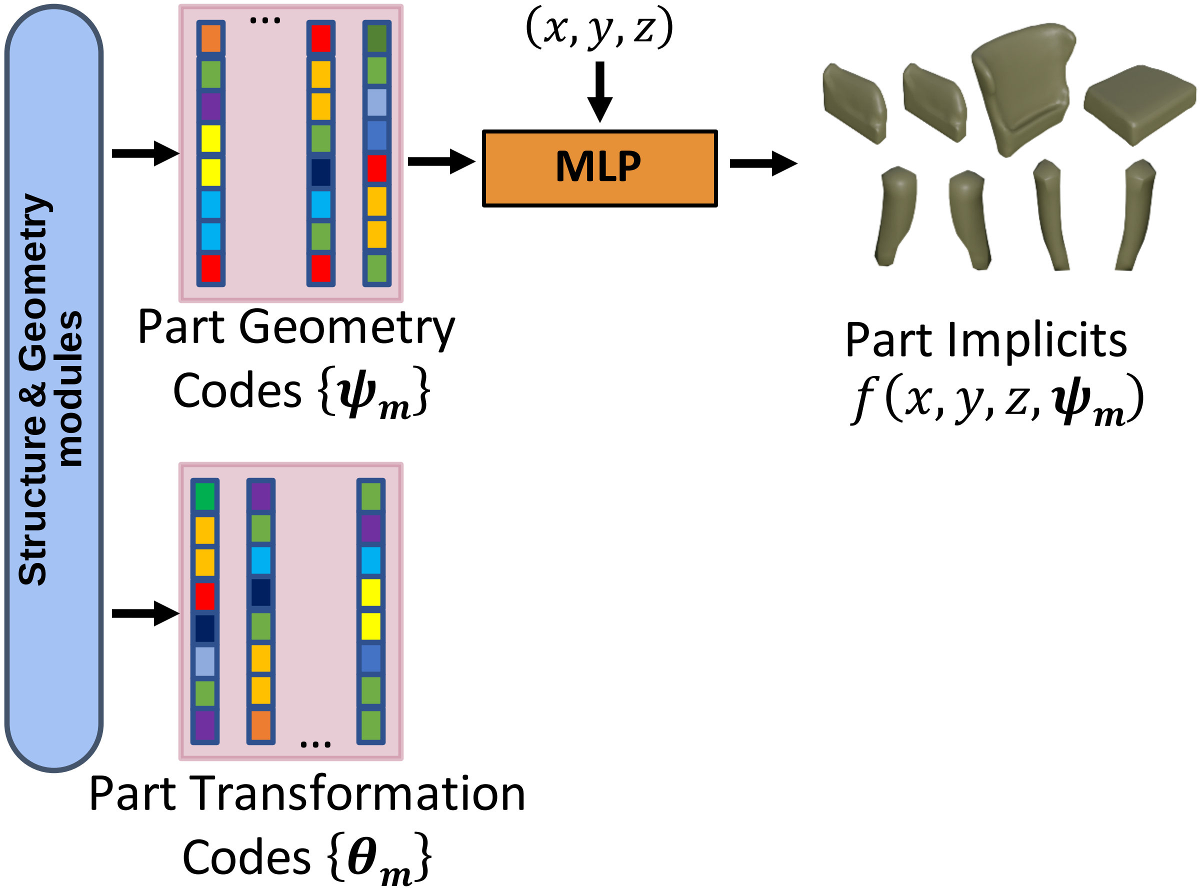
ANISE: neural modules



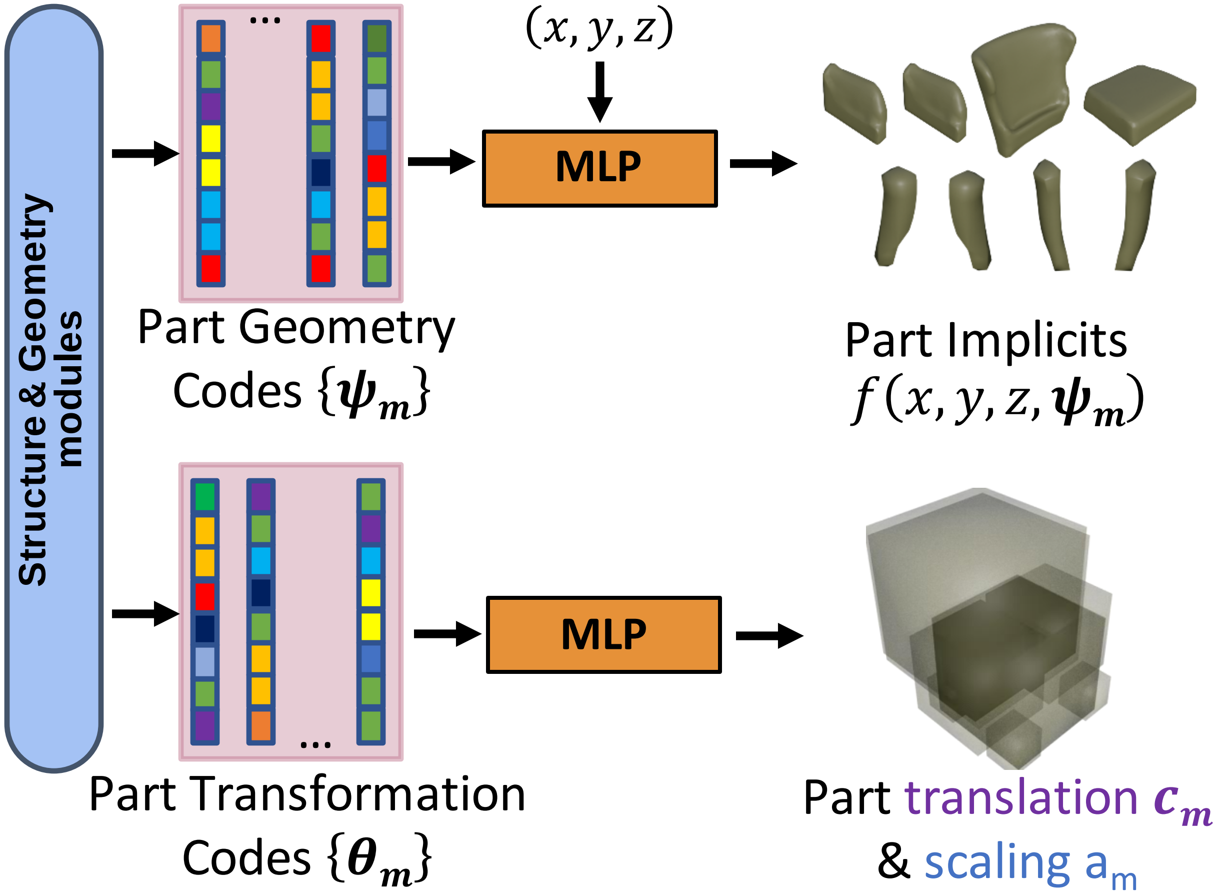
Assembly module



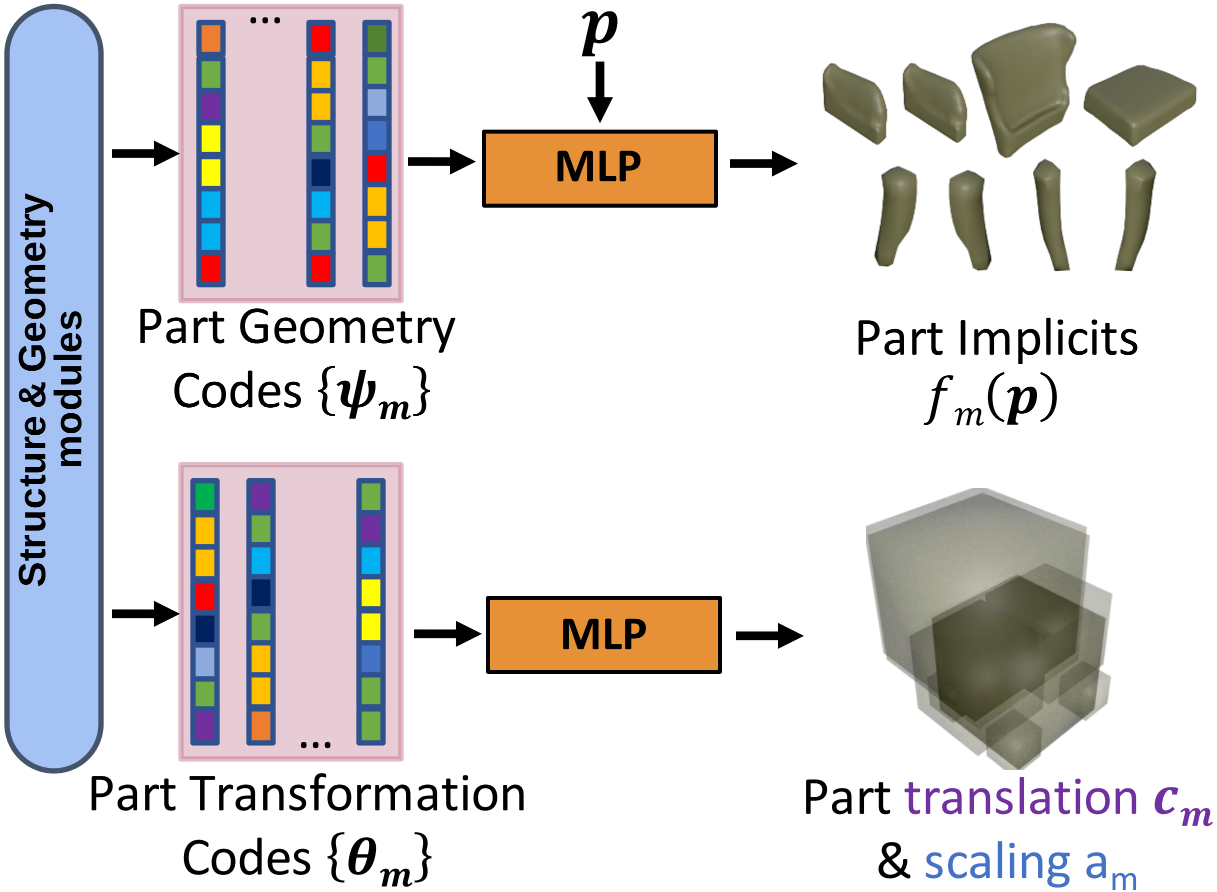
Assembly module



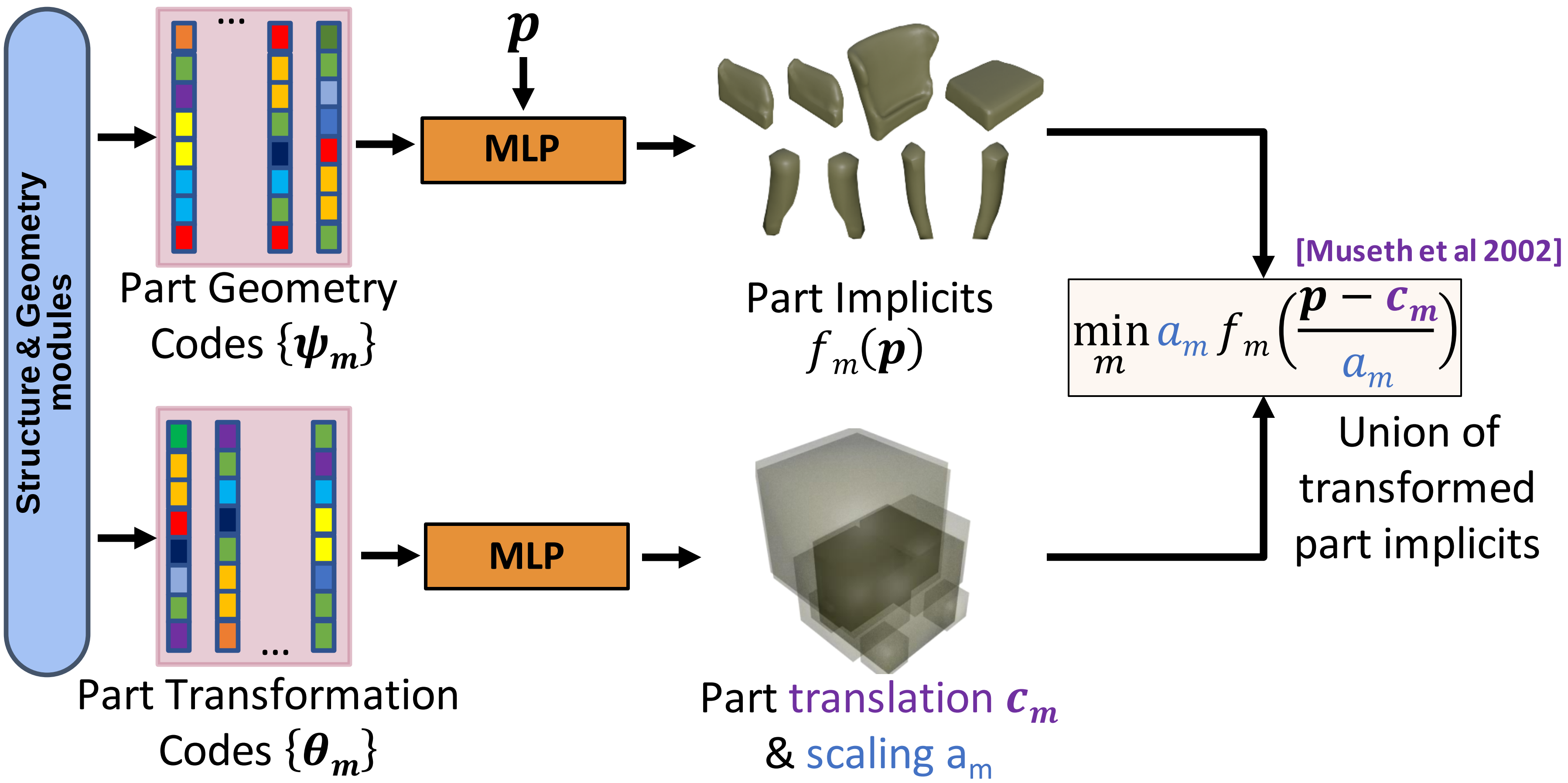
Assembly module



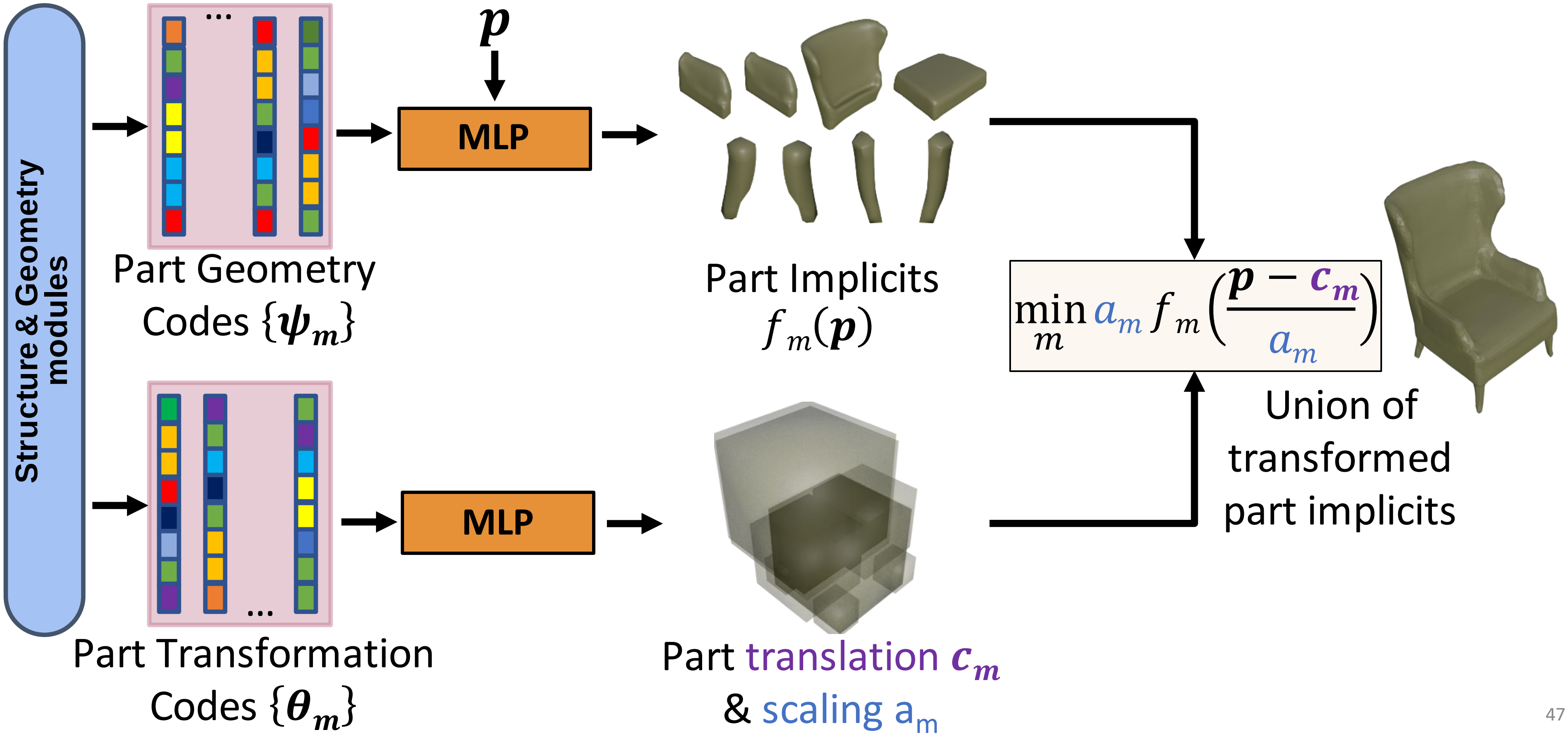
Assembly module



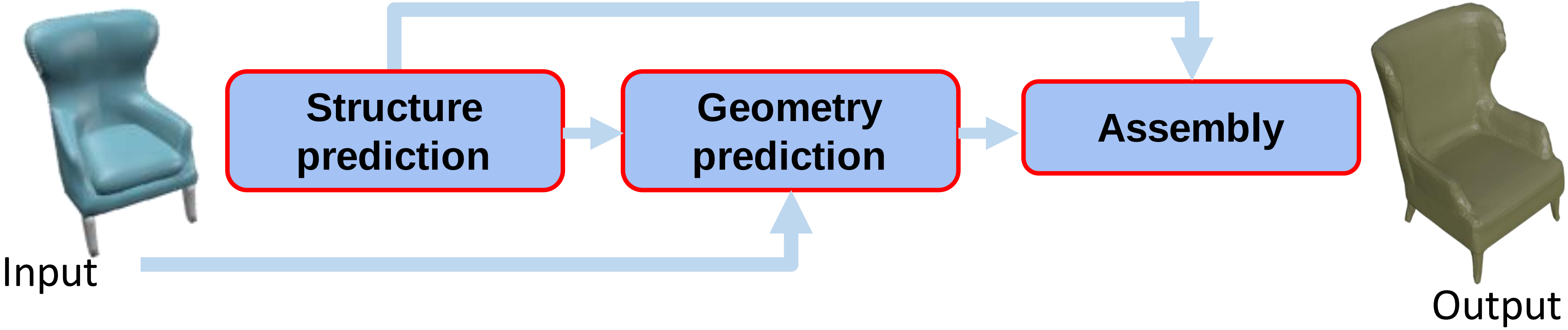
Assembly module



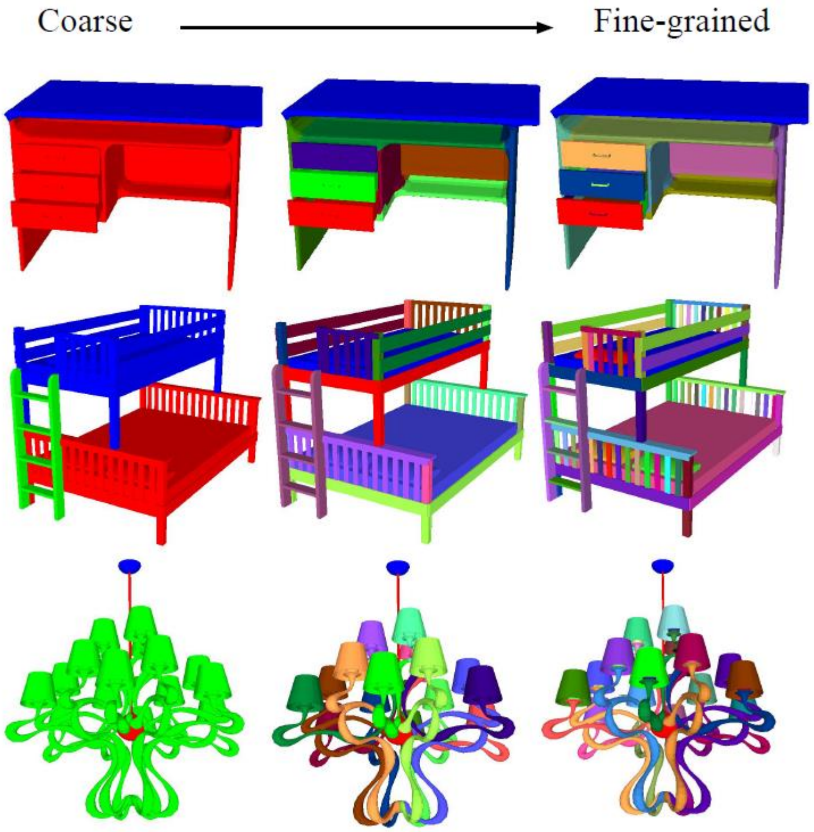
Assembly module



ANISE: training



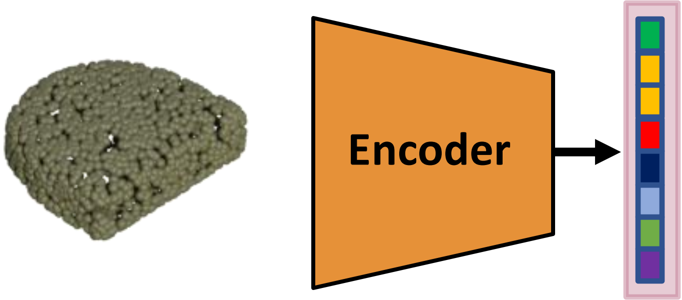
Need dataset of segmented parts => PartNet



[Mo et al. 2019]

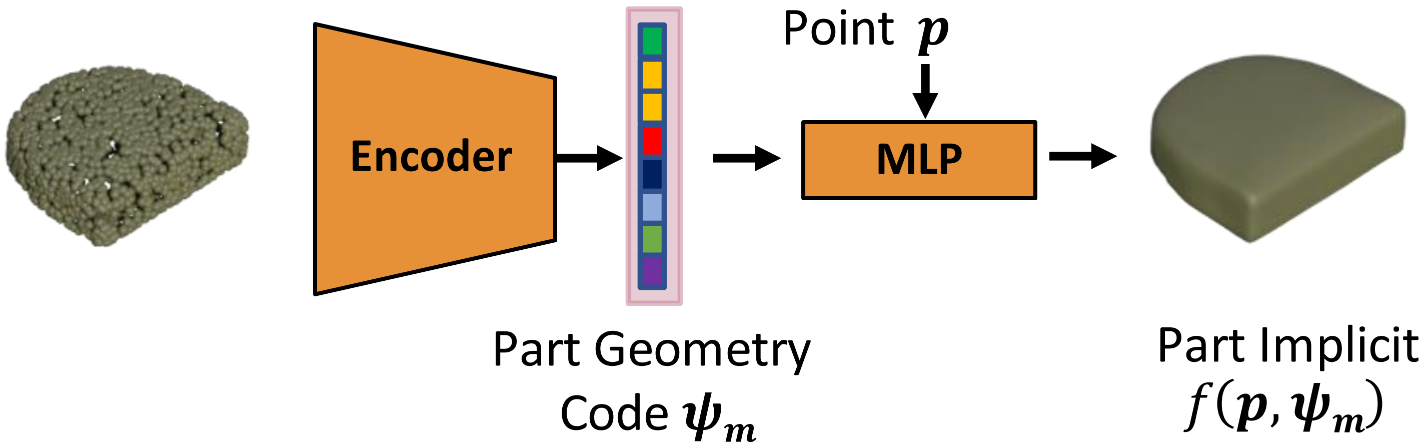
Pre-training stage: learning part geometry codes

Pre-training stage: learning part geometry codes

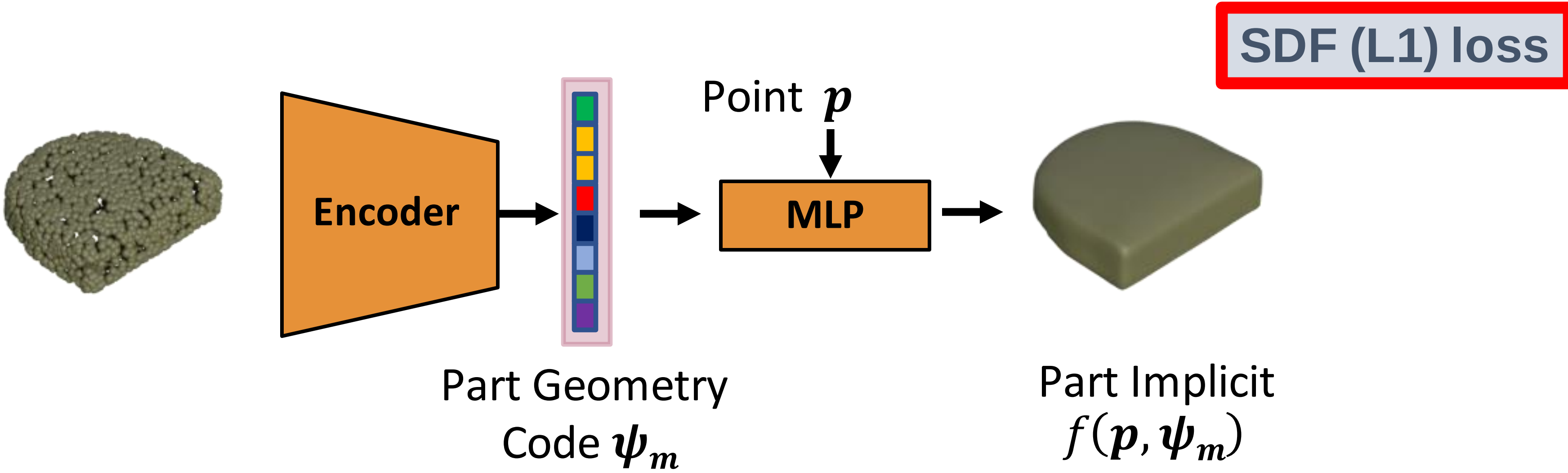


Part Geometry Code $\boldsymbol{\psi}_{\boldsymbol{m}}$

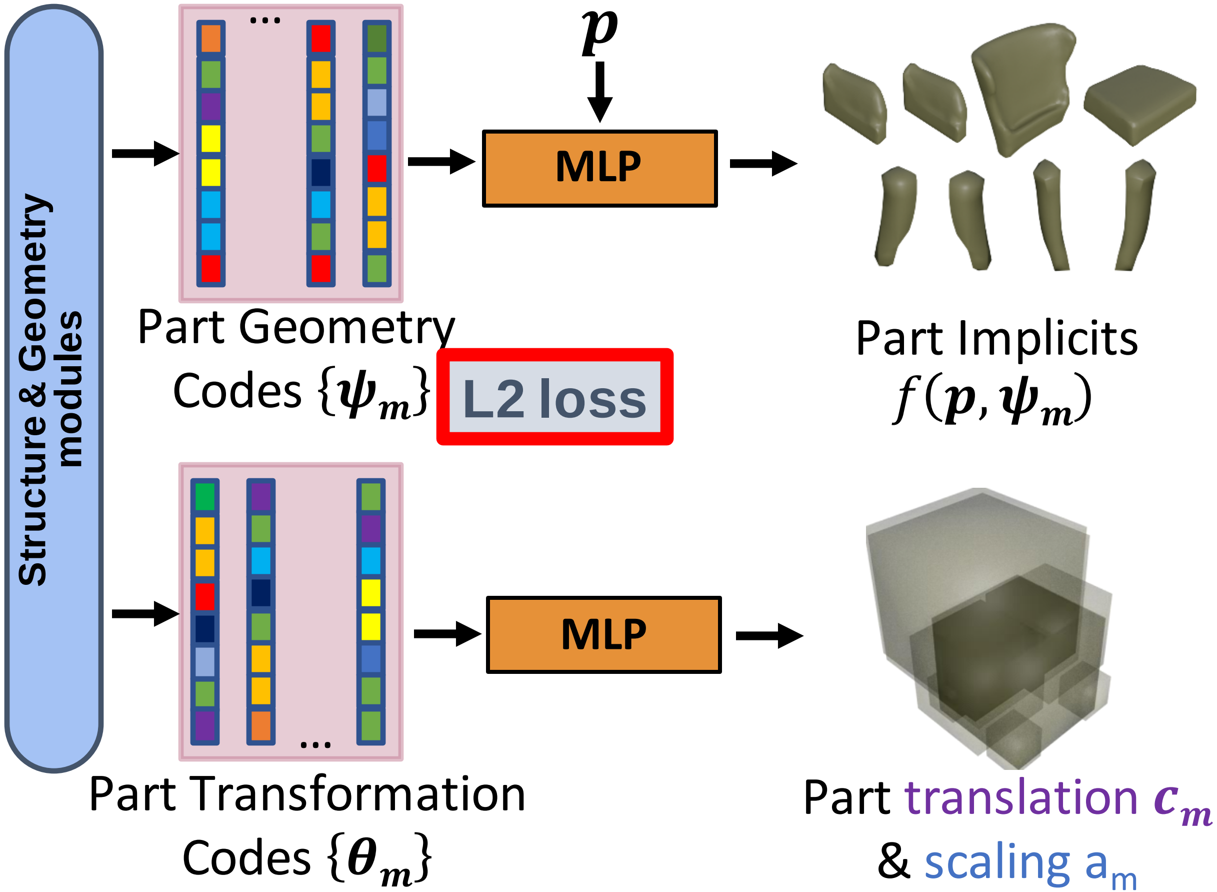
Pre-training stage: learning part geometry codes



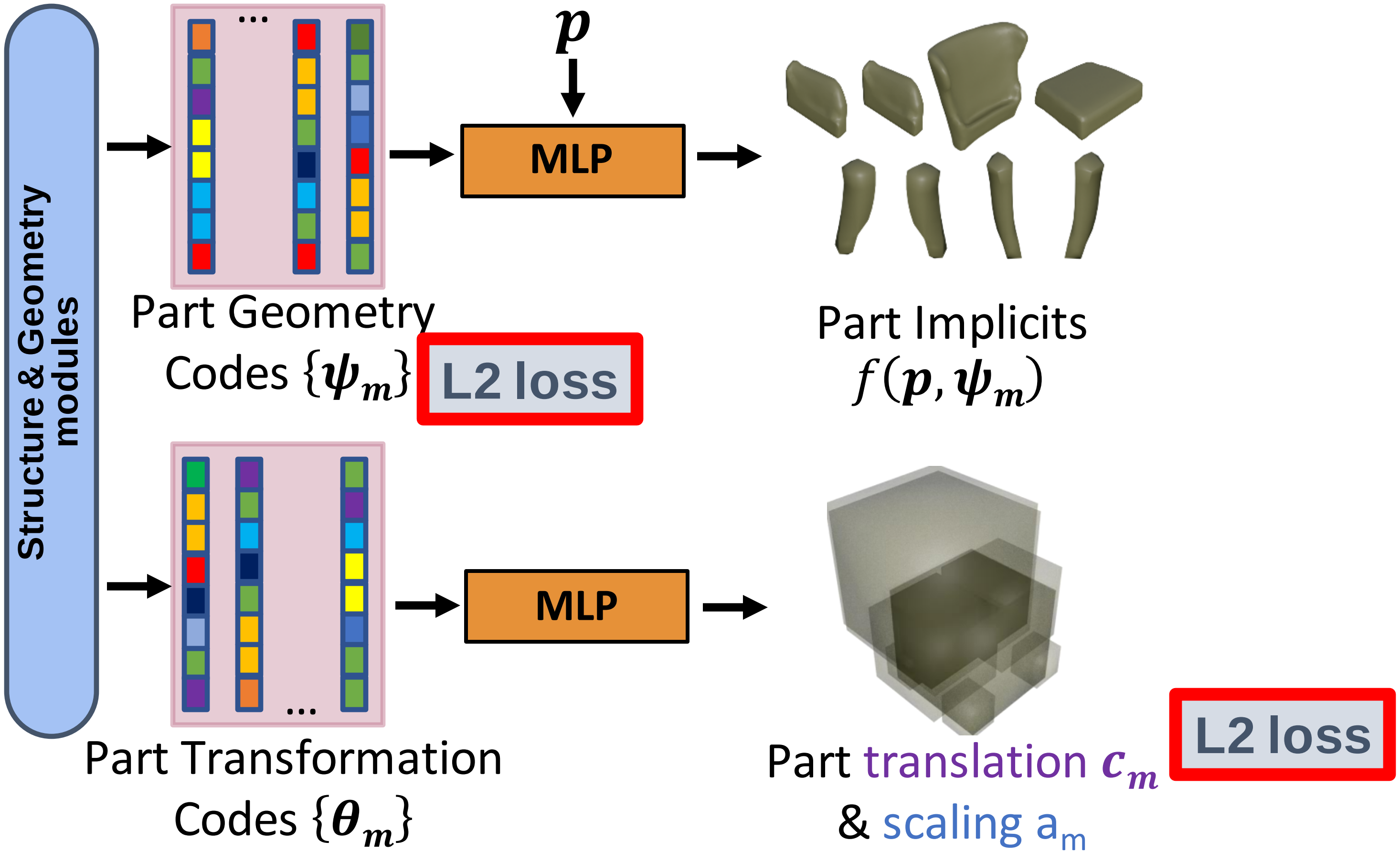
Pre-training stage: learning part geometry codes



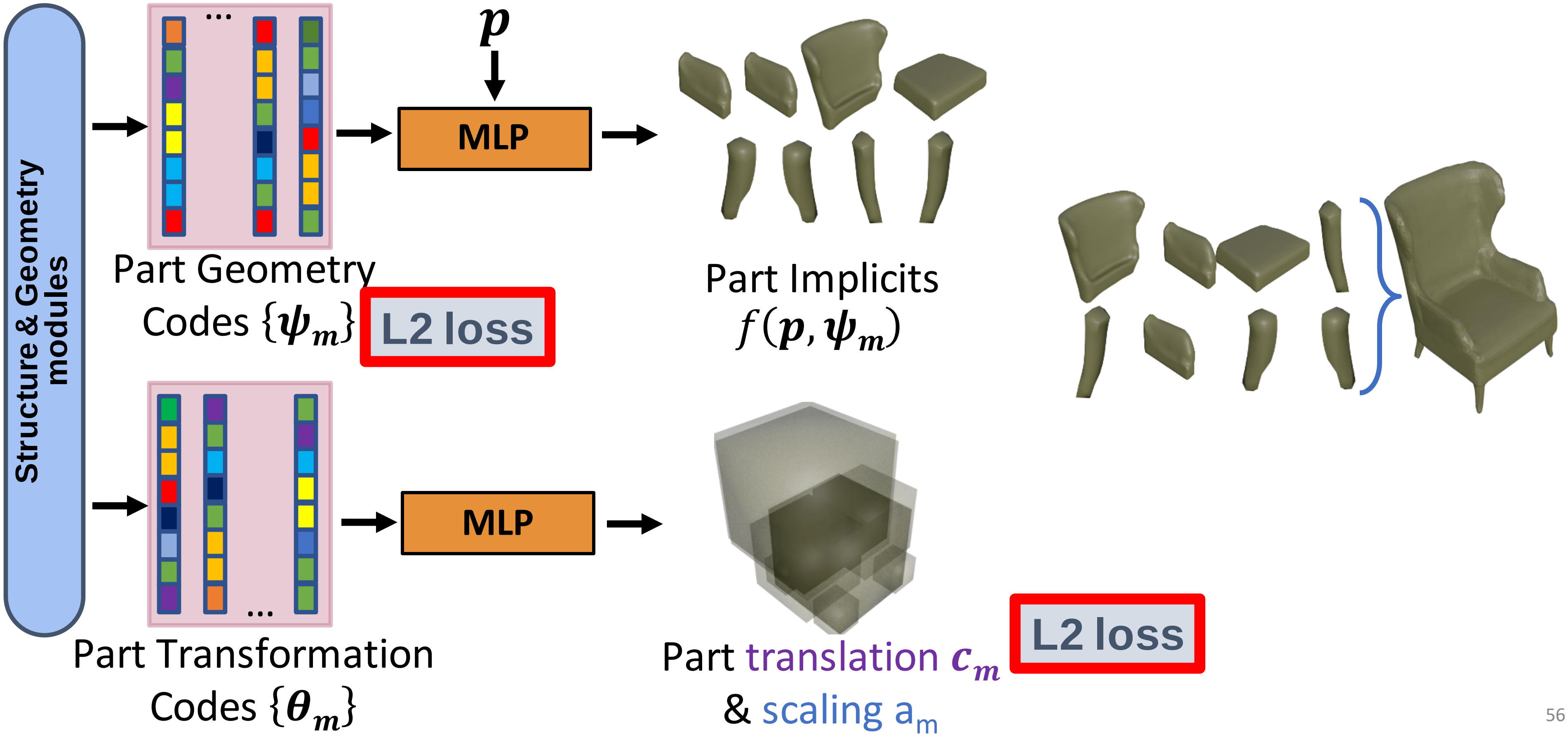
More pre-training: geometry & structure modules



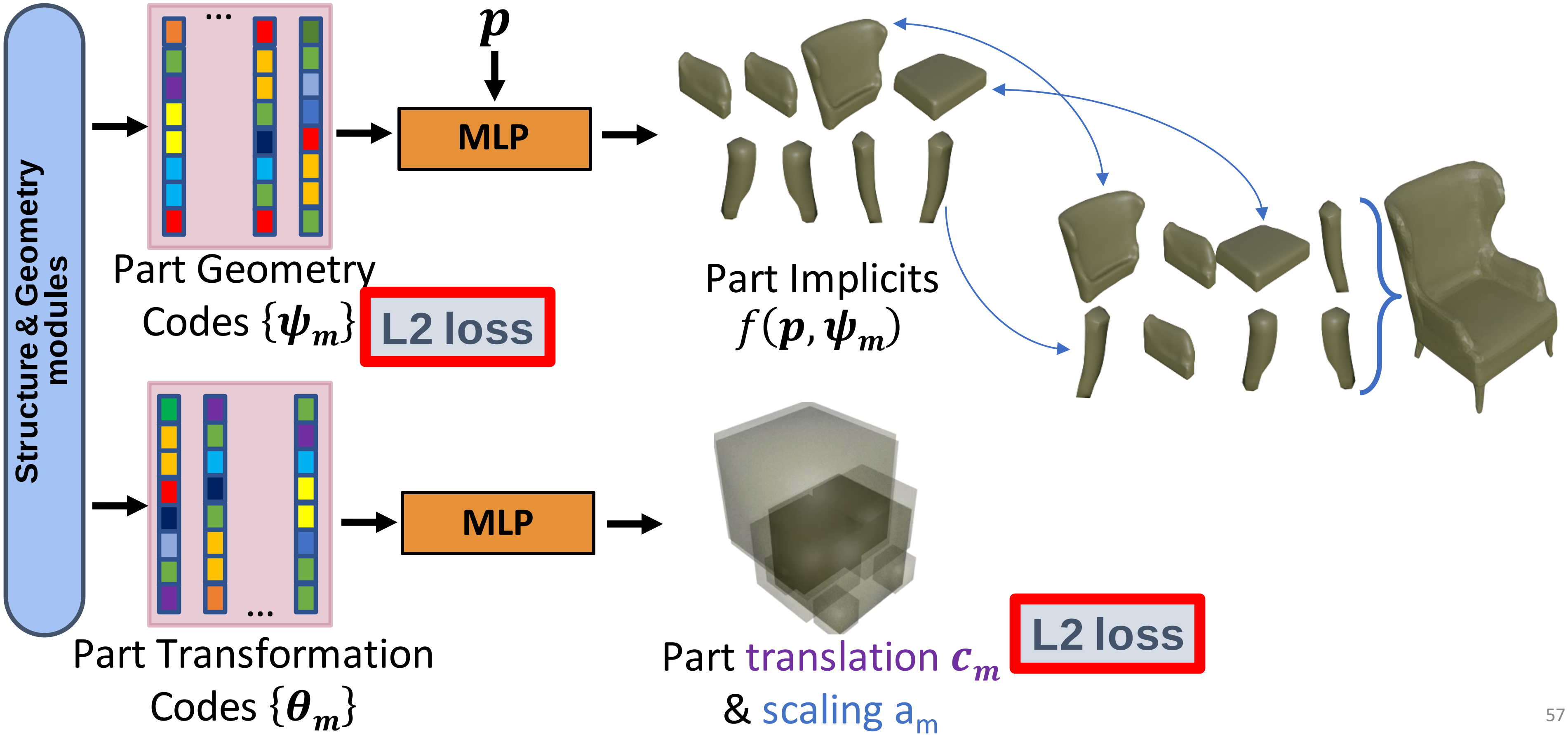
More pre-training: geometry & structure modules



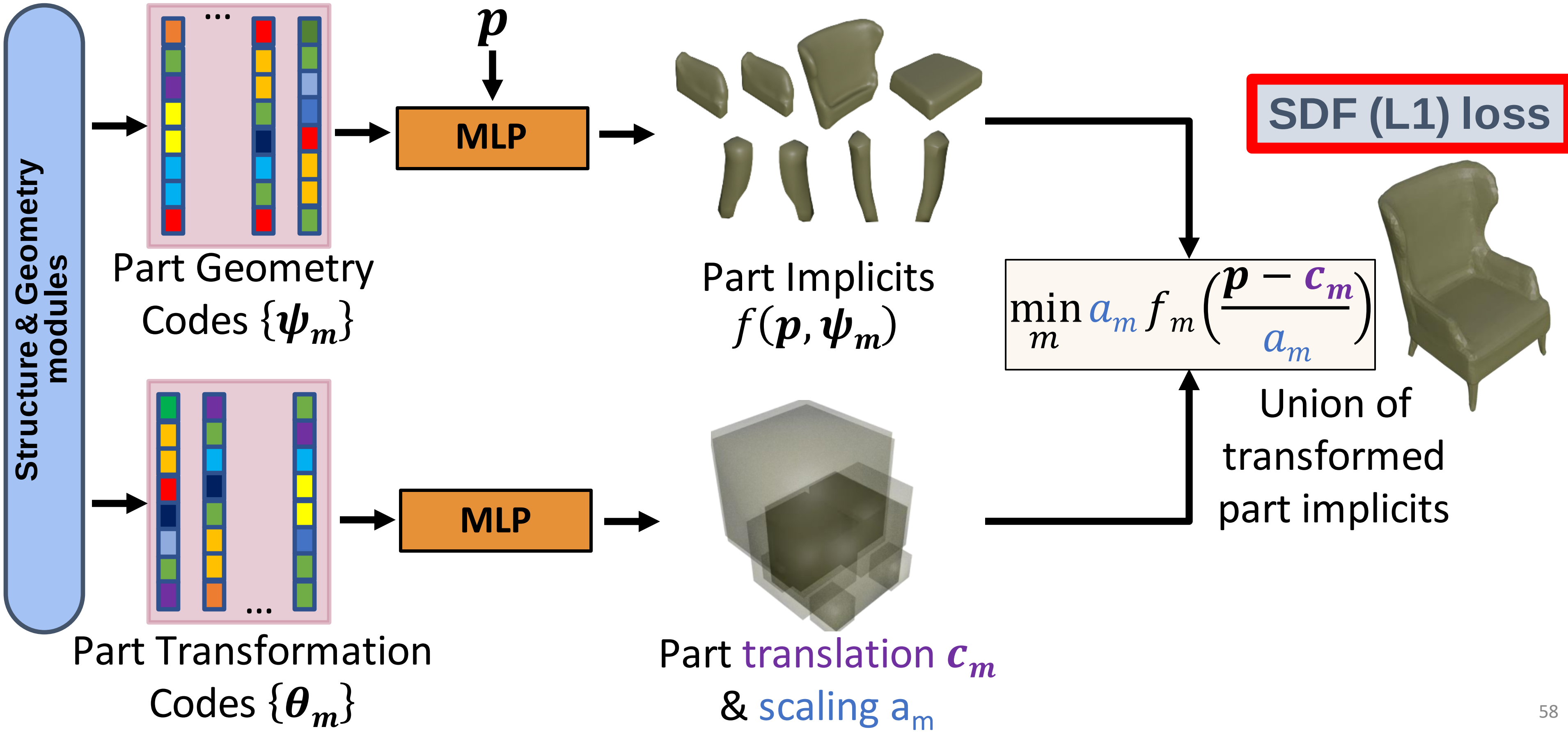
More pre-training: geometry & structure modules



More pre-training: geometry & structure modules



Final training stage



Results

PartNet training / testing

ANISE vs other part-aware methods

Input point cloud

Ground-truth

ANISE vs other part-aware methods

Input point cloud

PQ-NET
[Wu et al. 2020]

Ground-truth

ANISE vs other part-aware methods

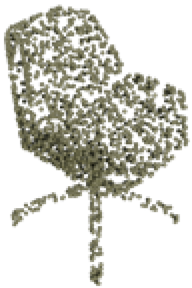
Input point cloud | **PQ-NET [Wu et al. 2020]** | **JLRD [Uy et al. 2021]** | **Ground-truth**

ANISE vs other part-aware methods

ANISE vs other part-aware methods

Input point cloud

Ground-truth

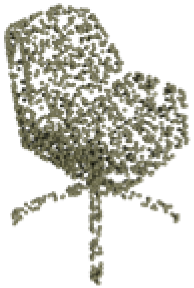


ANISE vs other part-aware methods

Input point cloud

PQ-NET [Wu et al. 2020]

Ground-truth



ANISE vs other part-aware methods




Input point cloud

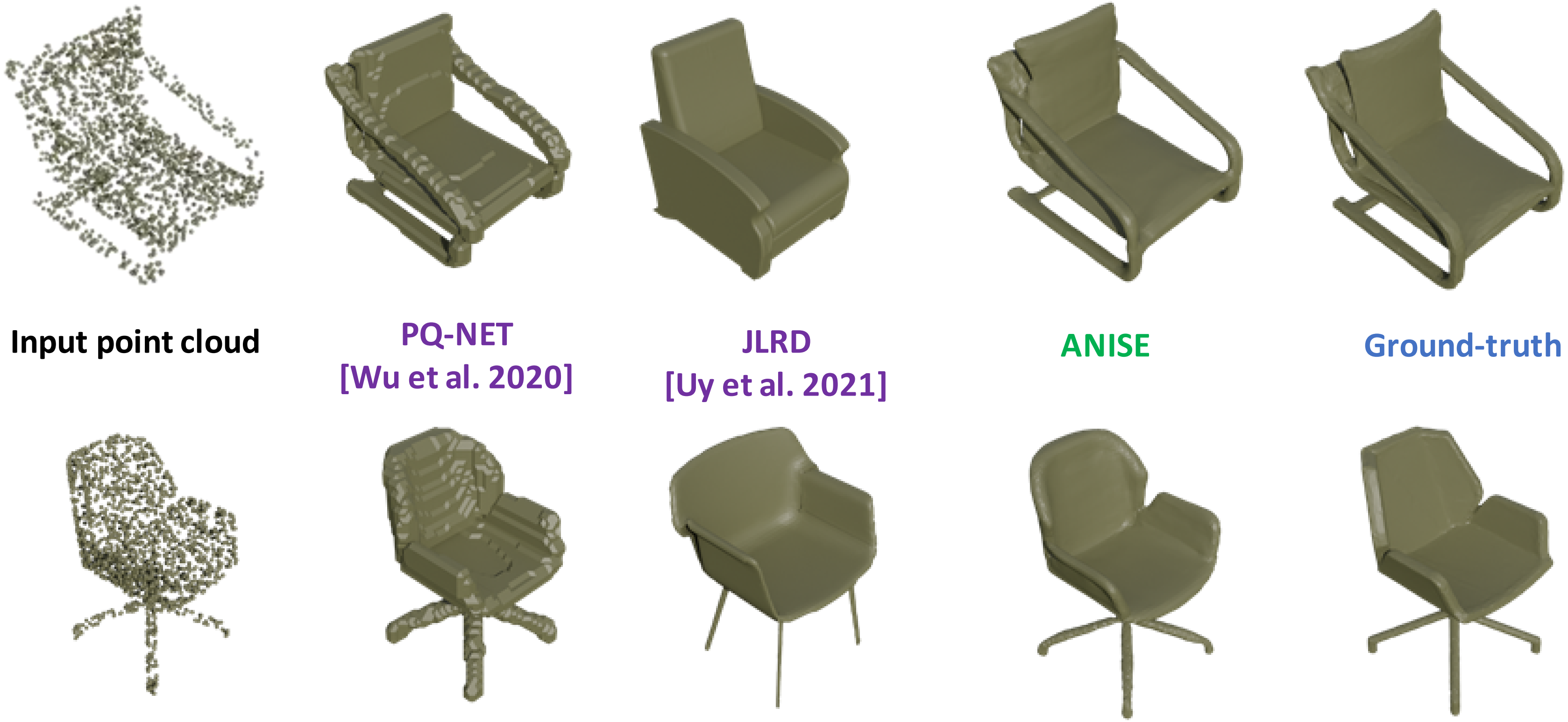
PQ-NET [Wu et al. 2020]

JLRD [Uy et al. 2021]

Ground-truth




ANISE vs other part-aware methods



ANISE vs other part-aware methods

Input RGB image

Ground-truth

ANISE vs other part-aware methods

Input RGB image

PQ-NET
[Wu et al. 2020]

Ground-truth

ANISE vs other part-aware methods




Input RGB image

PQ-NET [Wu et al. 2020]

JLRD [Uy et al. 2021]

Ground-truth




ANISE vs other part-aware methods

Input RGB image

PQ-NET [Wu et al. 2020]

JLRD [Uy et al. 2021]

ANISE

Ground-truth

ANISE vs other part-aware methods

Single-view reconstruction

	Chair			Table		
Method	**IOU (↑)**	**CD (↓)**	**F1 (↑)**	**IOU (↑)**	**CD (↓)**	**F1 (↑)**
JLRD						
PQ-Net						
ANISE						

ANISE vs other part-aware methods

Single-view reconstruction

	Chair			Table		
Method	**IOU (↑)**	**CD (↓)**	**F1 (↑)**	**IOU (↑)**	**CD (↓)**	**F1 (↑)**
JLRD	31.1			29.4		
PQ-Net	47.0			29.7		
ANISE	**56.7**			**57.4**		

ANISE vs other part-aware methods

Single-view reconstruction

	Chair			Table		
Method	**IOU (↑)**	**CD (↓)**	**F1 (↑)**	**IOU (↑)**	**CD (↓)**	**F1 (↑)**
JLRD	31.1	5.09		29.4	4.92	
PQ-Net	47.0	4.27		29.7	18.41	
ANISE	**56.7**	**2.99**		**57.4**	**2.54**	

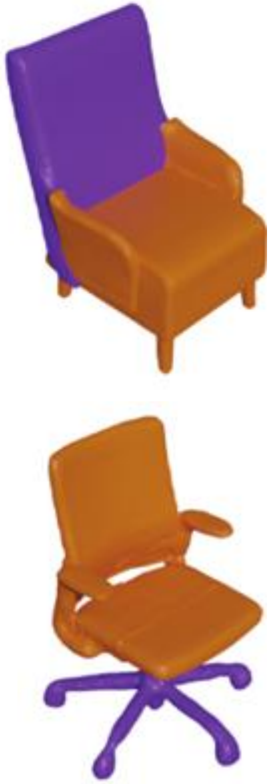
ANISE vs other part-aware methods

Single-view reconstruction

	Chair			Table		
Method	**IOU (↑)**	**CD (↓)**	**F1 (↑)**	**IOU (↑)**	**CD (↓)**	**F1 (↑)**
JLRD	31.1	5.09	48.1	29.4	4.92	58.0
PQ-Net	47.0	4.27	50.2	29.7	18.41	32.1
ANISE	**56.7**	**2.99**	**67.0**	**57.4**	**2.54**	**77.8**

Application: part editing

Rec. shapes

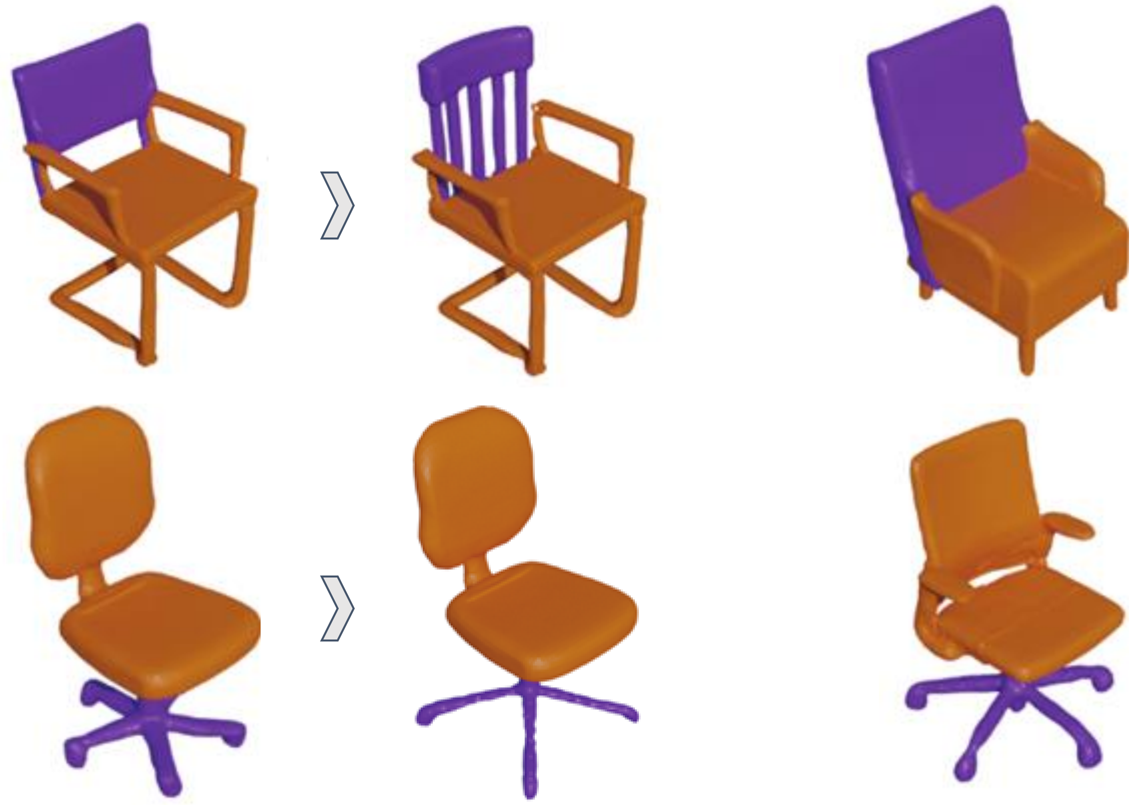


Rec. shapes

Application: part editing

Rec. shapes **Edited Shapes** **Rec. shapes**

Application: part editing



Rec. shapes **Edited Shapes** **Rec. shapes**

Application: part editing

Rec. shapes **Edited Shapes** **Rec. shapes** **Edited Shapes**

Application: part editing

Rec. shapes **Edited Shapes** **Rec. shapes** **Edited Shapes**

Application: Part-constrained shape assembly

Input Collection of Shapes

Application: Part-constrained shape assembly

Input point cloud

Input Collection of Shapes

?

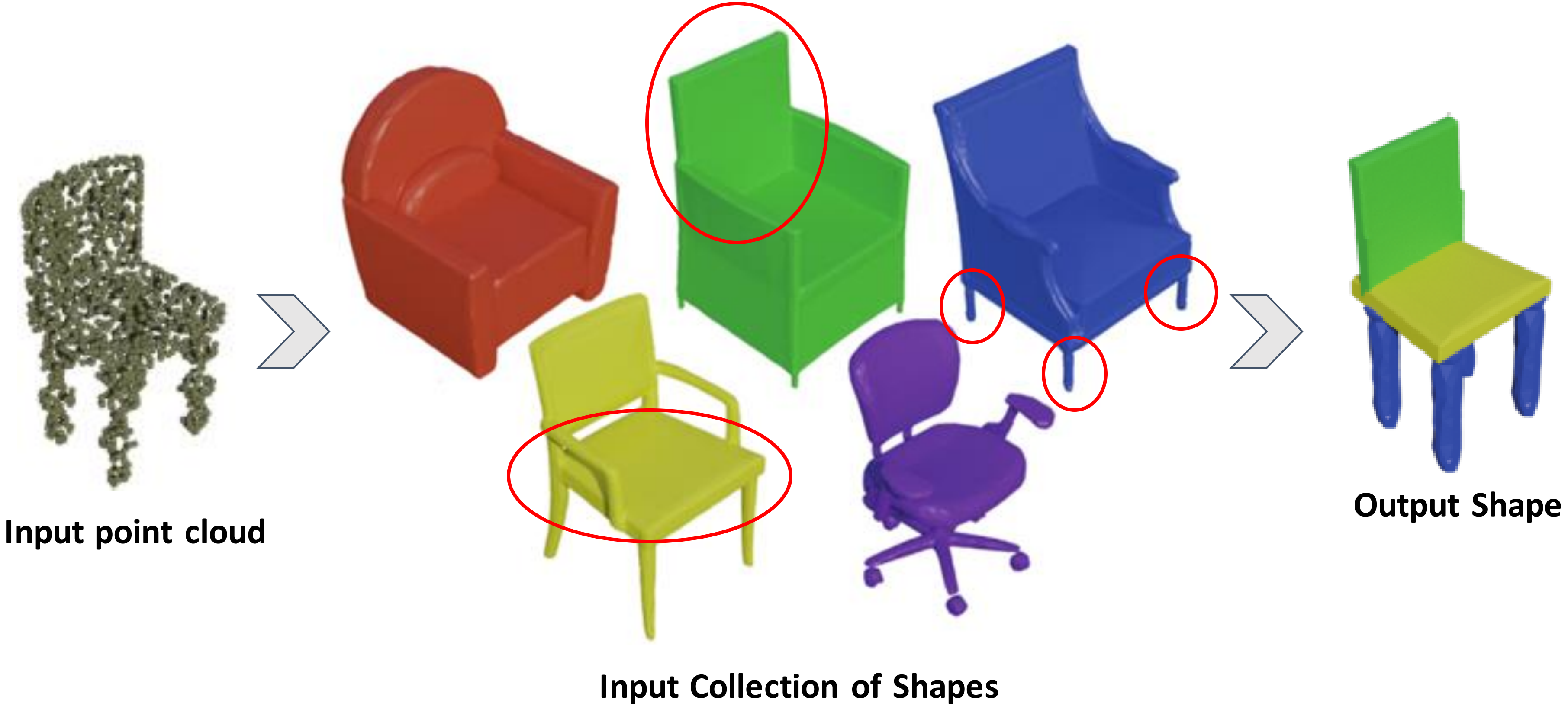
Application: Part-constrained shape assembly

Input point cloud

Input Collection of Shapes

?

Application: Part-constrained shape assembly



Application: Part-constrained shape assembly

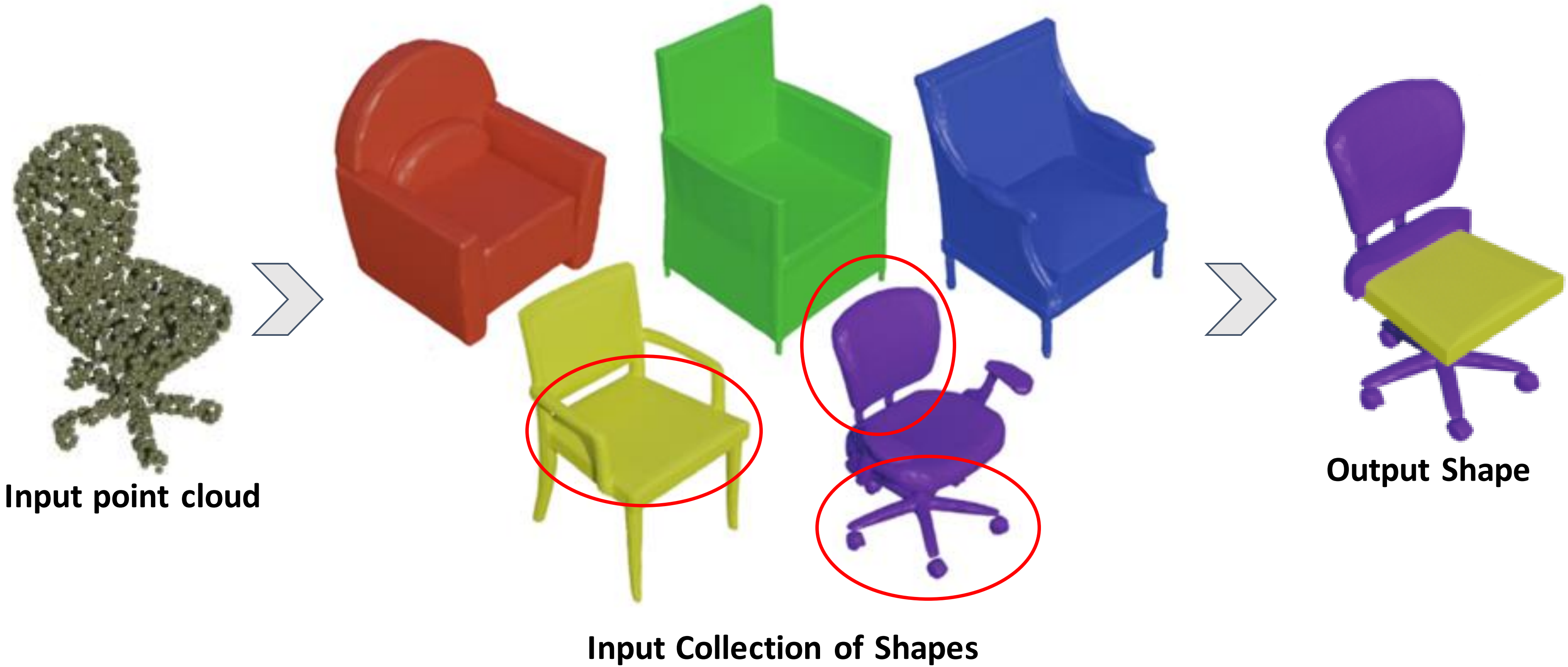
Input Collection of Shapes

Application: Part-constrained shape assembly

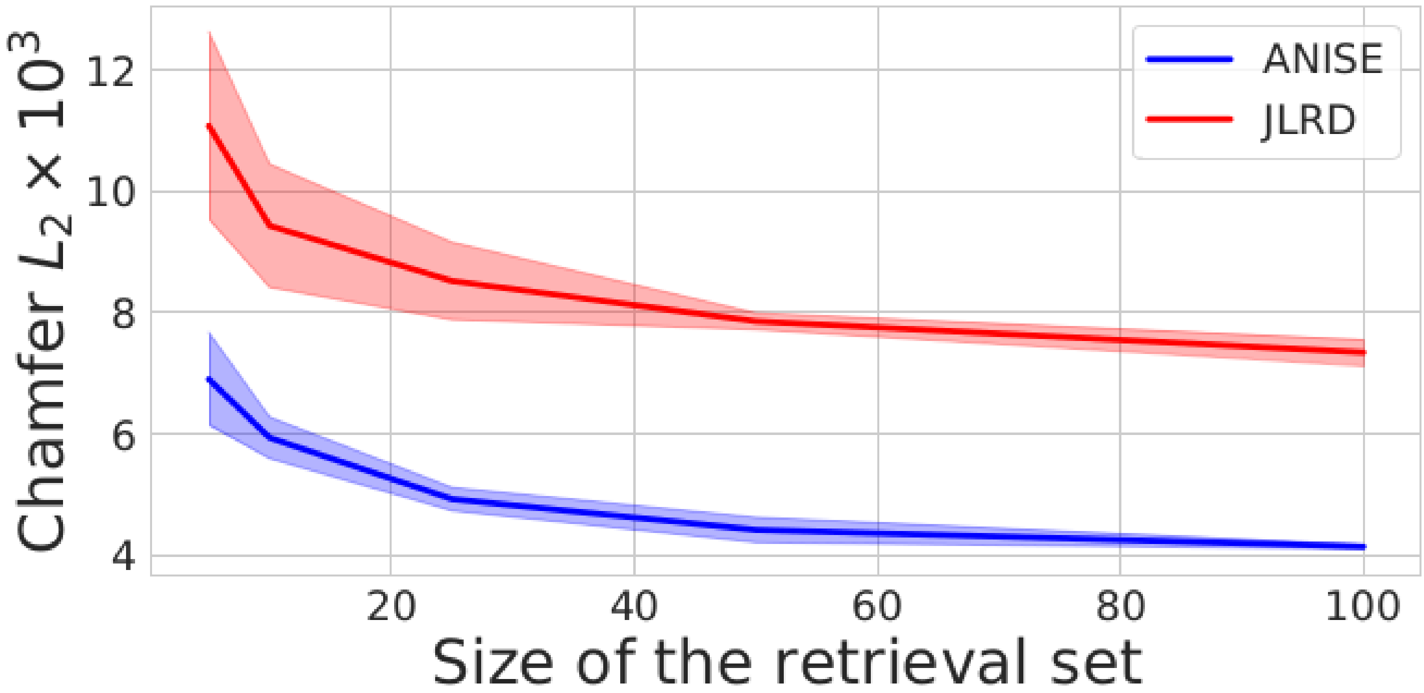
Input point cloud

Input Collection of Shapes

Application: Part-constrained shape assembly



Application: Part-constrained shape assembly

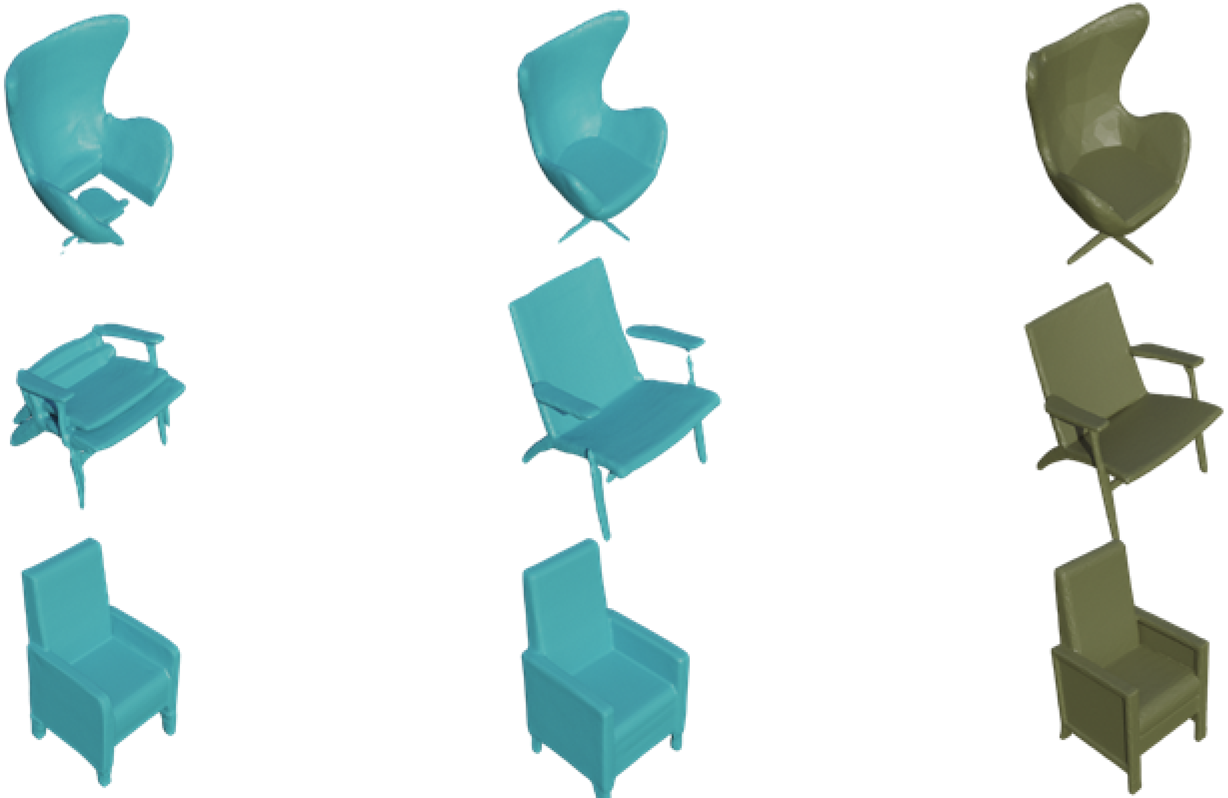


Ablation study: importance of full-shape supervision

w/o full shape supevision

Ground truth

Ablation study: importance of full-shape supervision



w/o full shape supevision

w/ full shape supevision

Ground truth

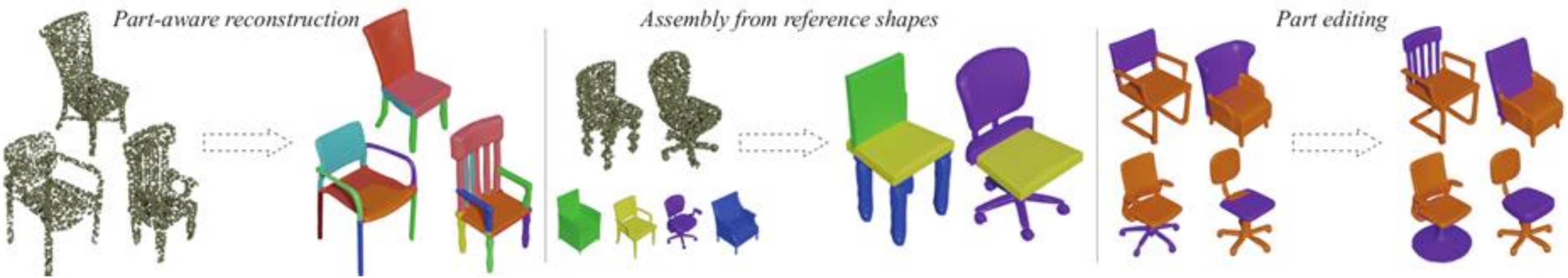
Ablation study

Full-shape supervision	CD (↓)
X	2.44
✓	**1.69**

Ablation study

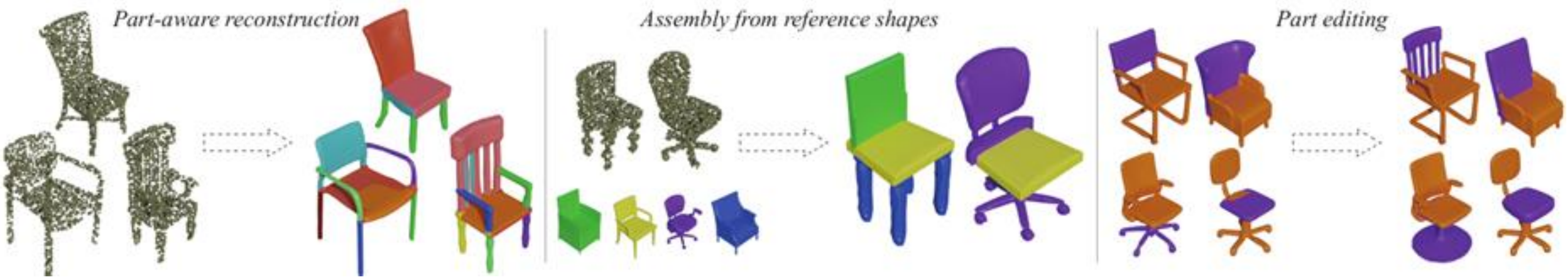
Geometry conditioned on structure	Full-shape supervision	CD (↓)
X	X	2.81
✓	X	2.44
✓	✓	**1.69**

Summary



- Neural architecture for **part-aware implicit surface reconstruction**
- **SOTA performance** compared to ptrior methods
- Enables **part-based editing** and **part assembly from reference shapes**

Limitations



- Needs part supervision & part-segmented datasets
- Transformations are limited to translation and uniform scaling -- no rotations
- Coarse-to-fine synthesis, yet no multiple levels of part hierarchies

Thank you!

Acknowledgements:

Adobe

Our project web page:

https://lodurality.github.io/ANISE