Geometric Deep Learning

— From the Perspective of 3D Shape Synthesis

三维内容生成

3D content generation refers to the process of creating 3D objects, scenes, and animations for use in various media, such as films, video games, and virtual reality experiences.

几何形状生成

Let's focus on geometric shapes first.

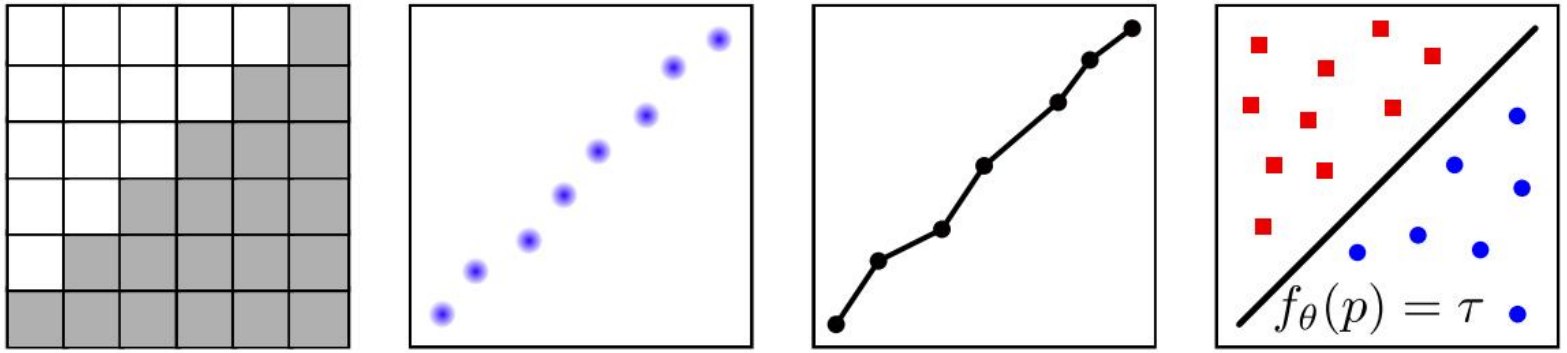
How to synthesize an object with complex geometric shape?

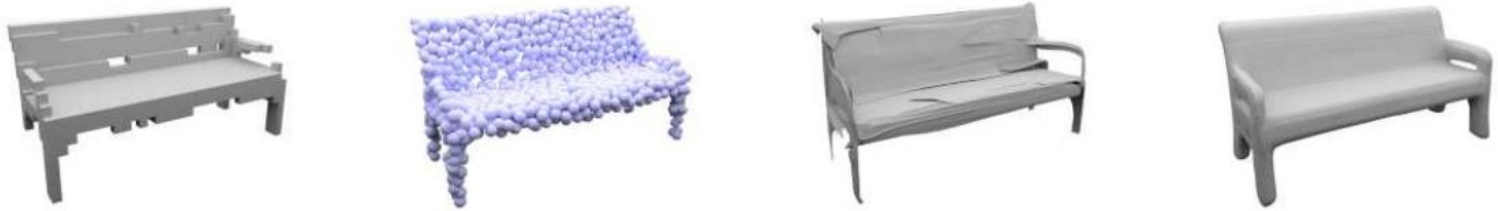
Data-Driven Modeling Methods

几何形状生成

Let's focus on geometric shapes first.

How to synthesize an object with complex geometric shape?



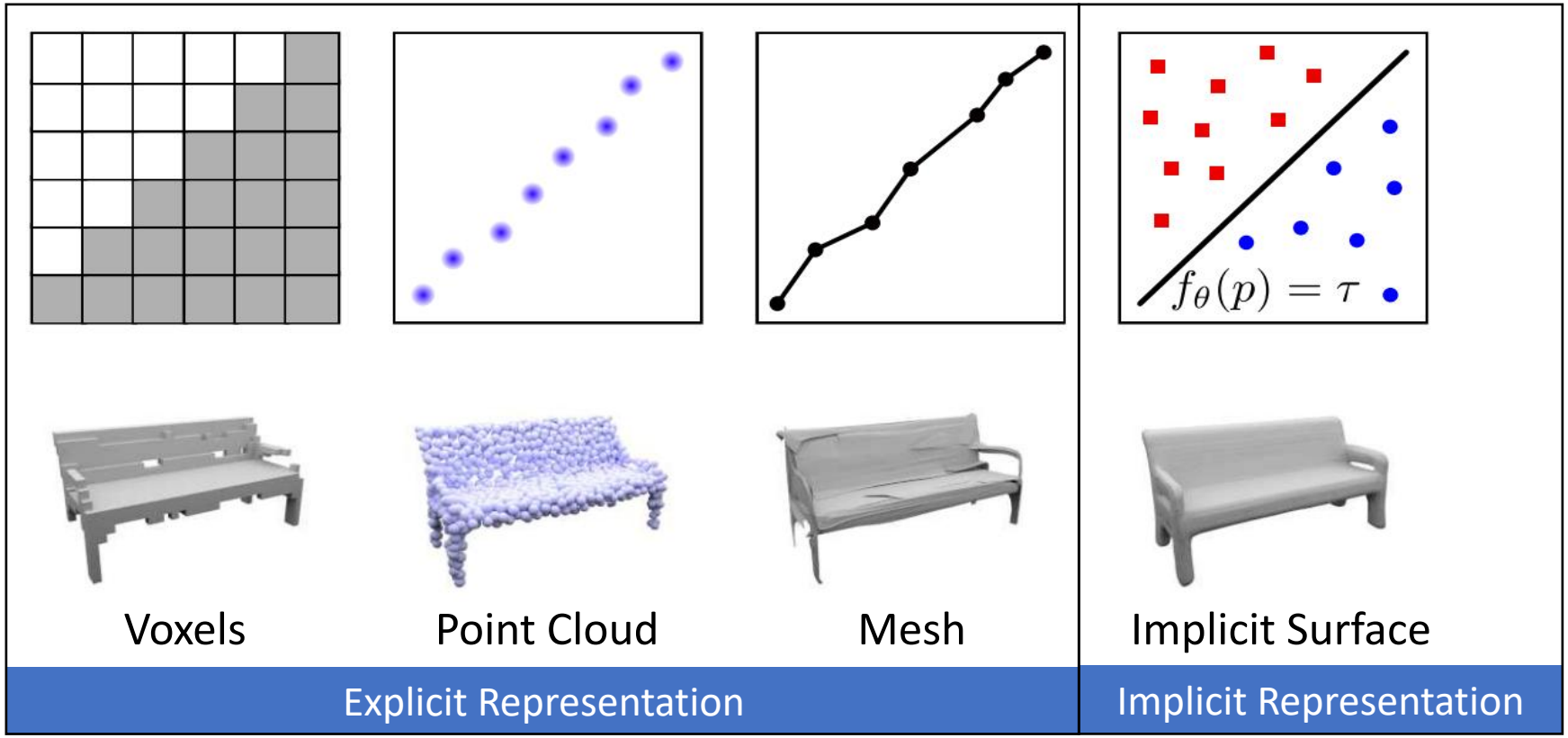


Voxels Point Cloud Mesh Implicit Surface

几何形状生成

Let's focus on geometric shapes first.

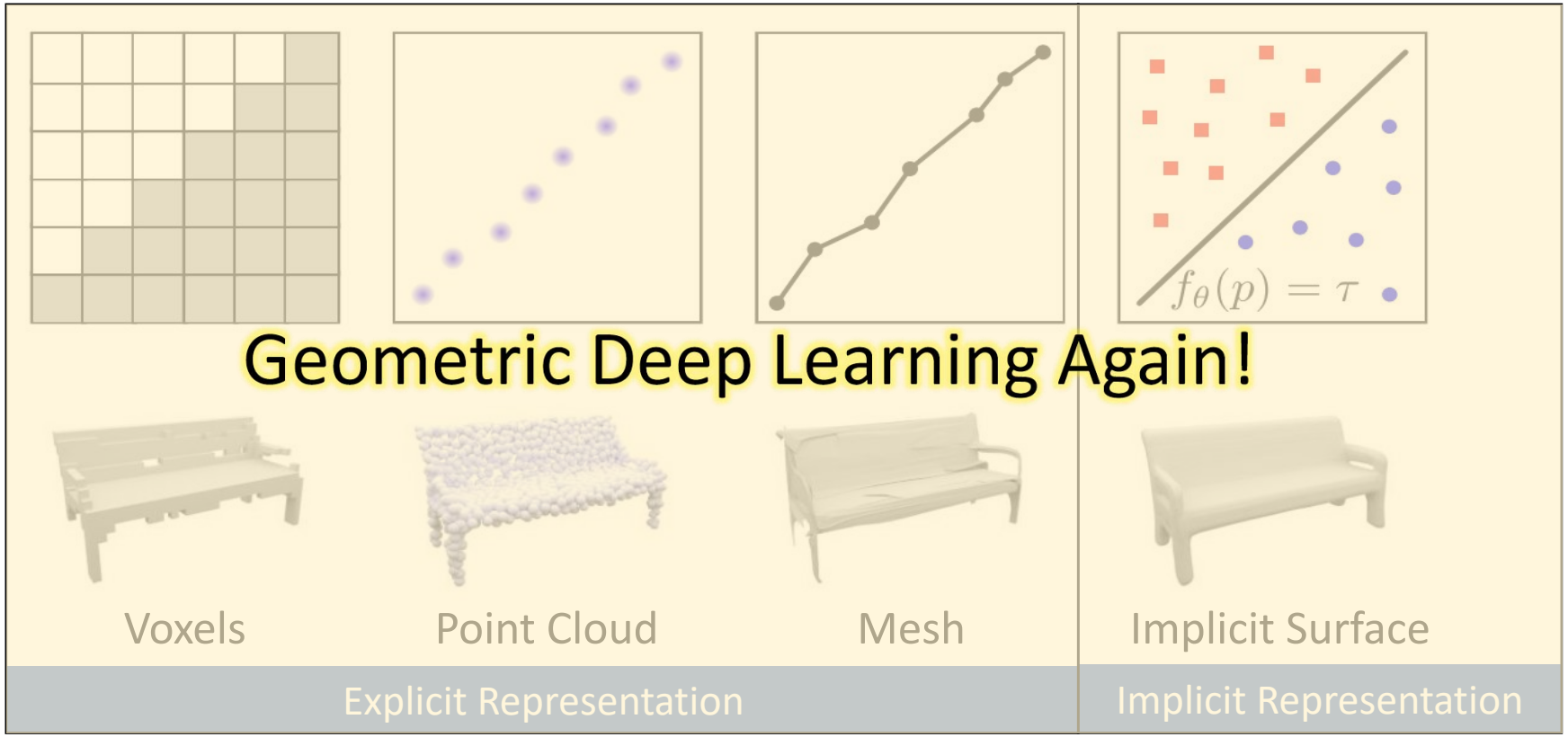
How to synthesize an object with complex geometric shape?



几何形状生成

Let's focus on geometric shapes first.

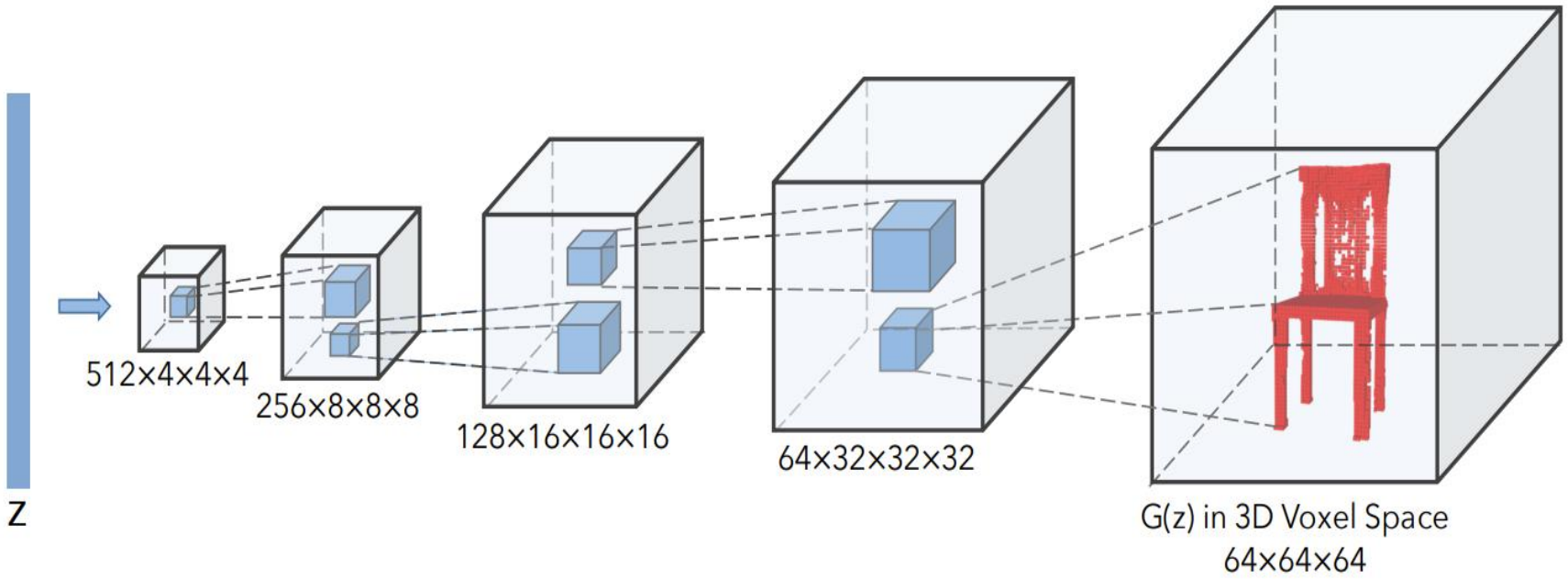
How to synthesize an object with complex geometric shape?



Learning Explicit and Implicit Representations

三维体素生成

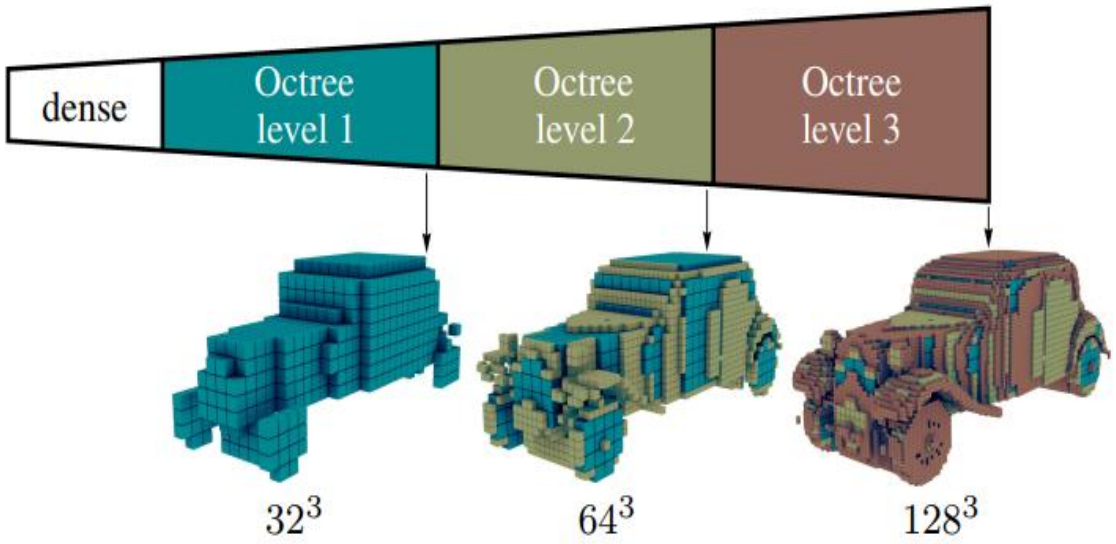
- First attempt: 3D CNN + GAN



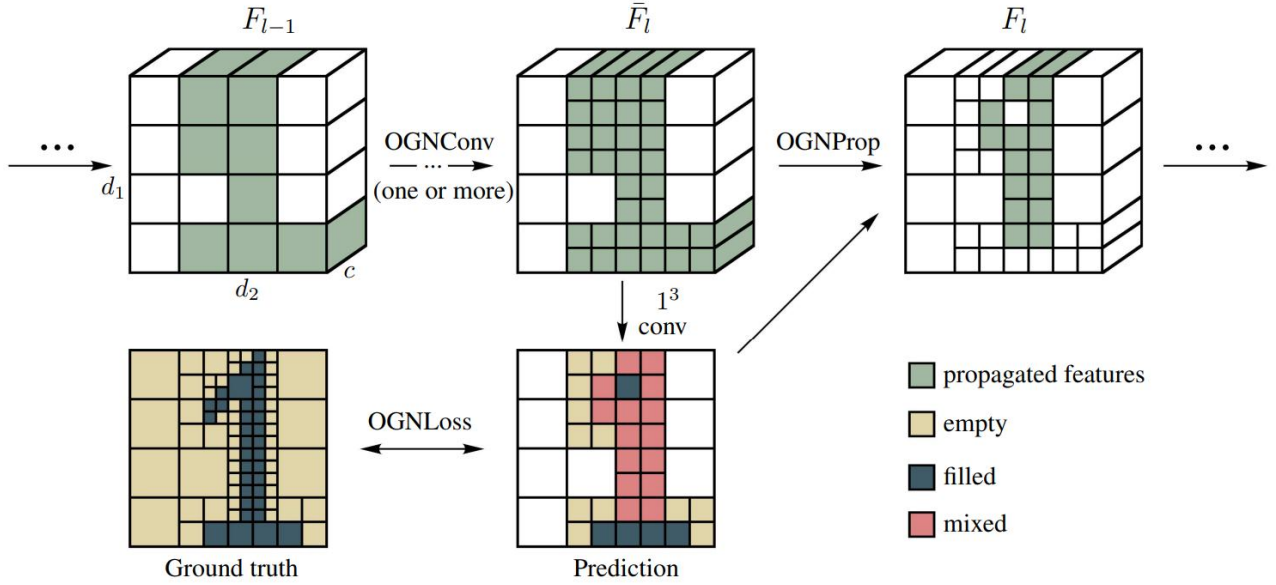
[Wu et al. 2016]

三维体素生成

- First attempt: 3D CNN + GAN
- How to save memory and time cost: octree...



The volumetric output as an octree

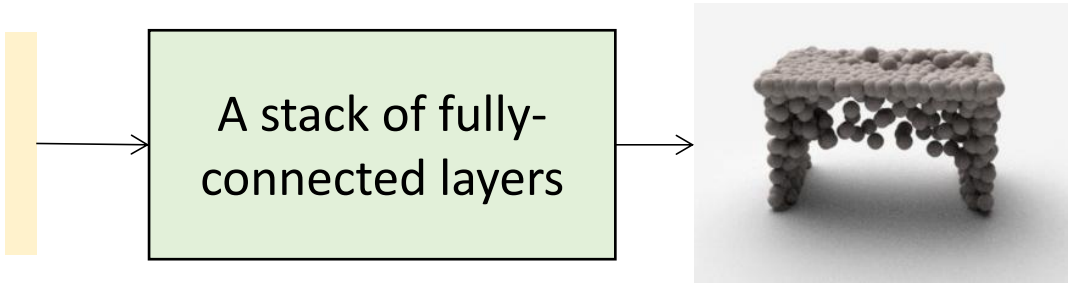


The “empty” cells are no longer needed and thus not propagated

Octree Generating Networks: Efficient Convolutional Architectures for High-resolution 3D Outputs

三维点云生成

- First attempt: MLP networks

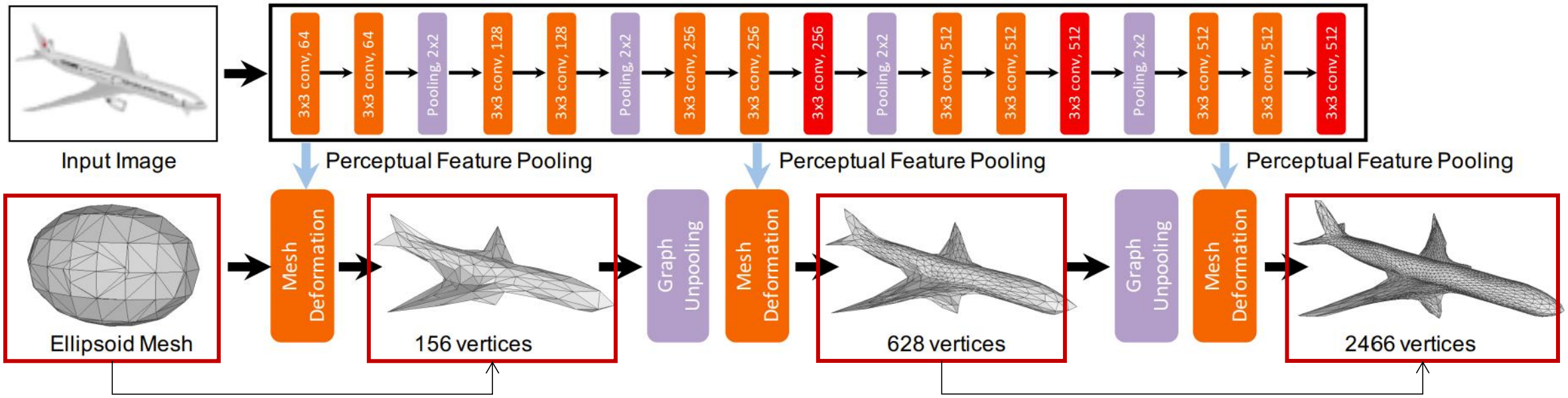


Point clouds generated from a hourglass architecture

三维网格生成

Challenge: varying topology

Solution 1: every shape is deformed with a fixed topology



Network predicts a deformation of an ellipsoid.

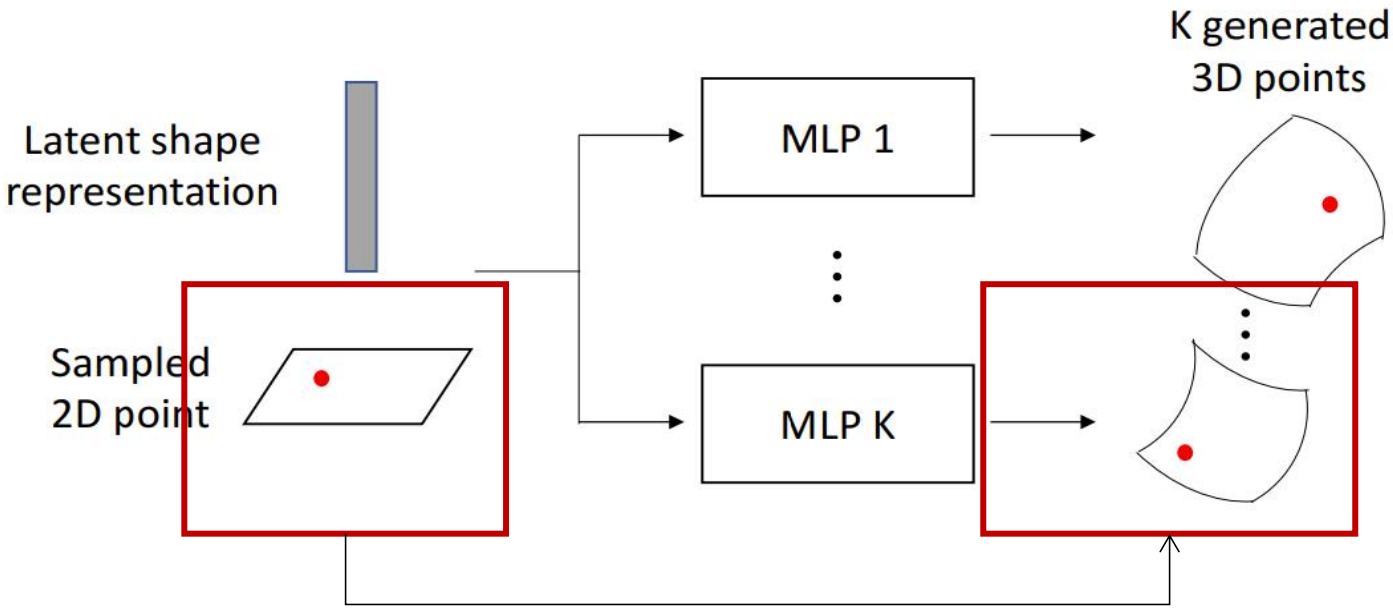
Unpooling provides the upsampling of meshes.

Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images

三维网格生成

Challenge: varying topology

Solution 1: every shape is deformed with a fixed topology



Network predicts a deformation of a plane.

The patches are stitched to a 3D shape.

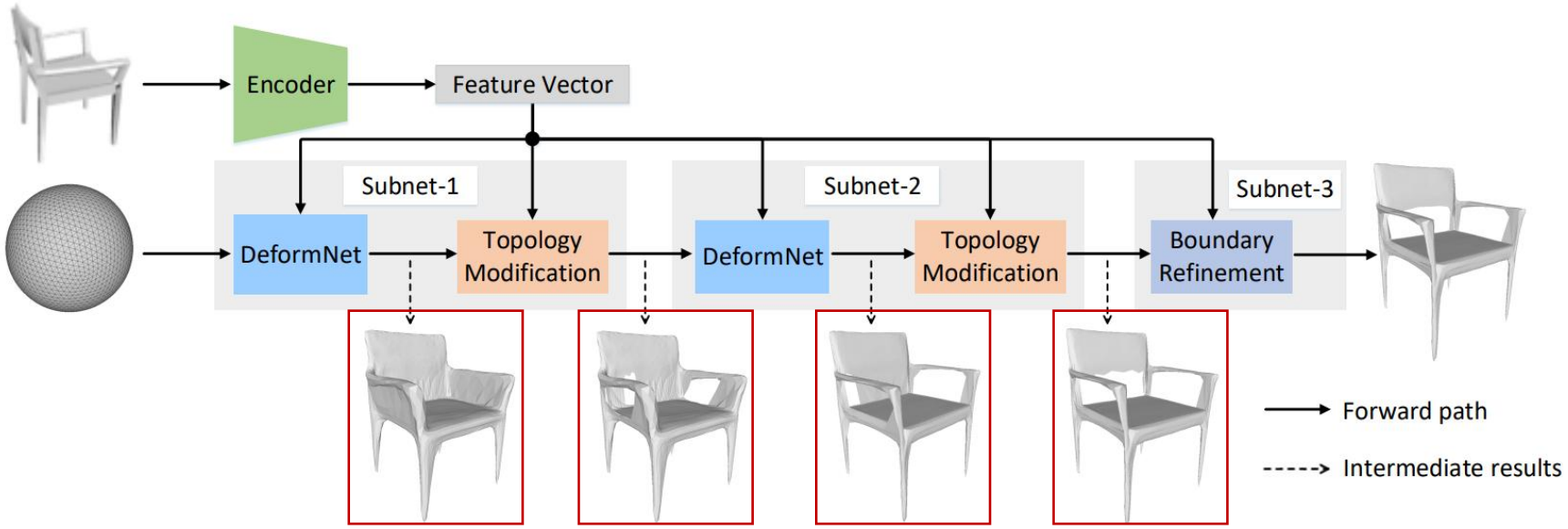
AtlasNet: A Papier-Mache Approach to Learning 3D Surface Generation

三维网格生成

Challenge: varying topology

Solution 1: every shape is deformed with a fixed topology

Solution 2: iteratively update the topology



Deep Mesh Reconstruction from Single RGB Images via Topology Modification Networks

三维网格生成

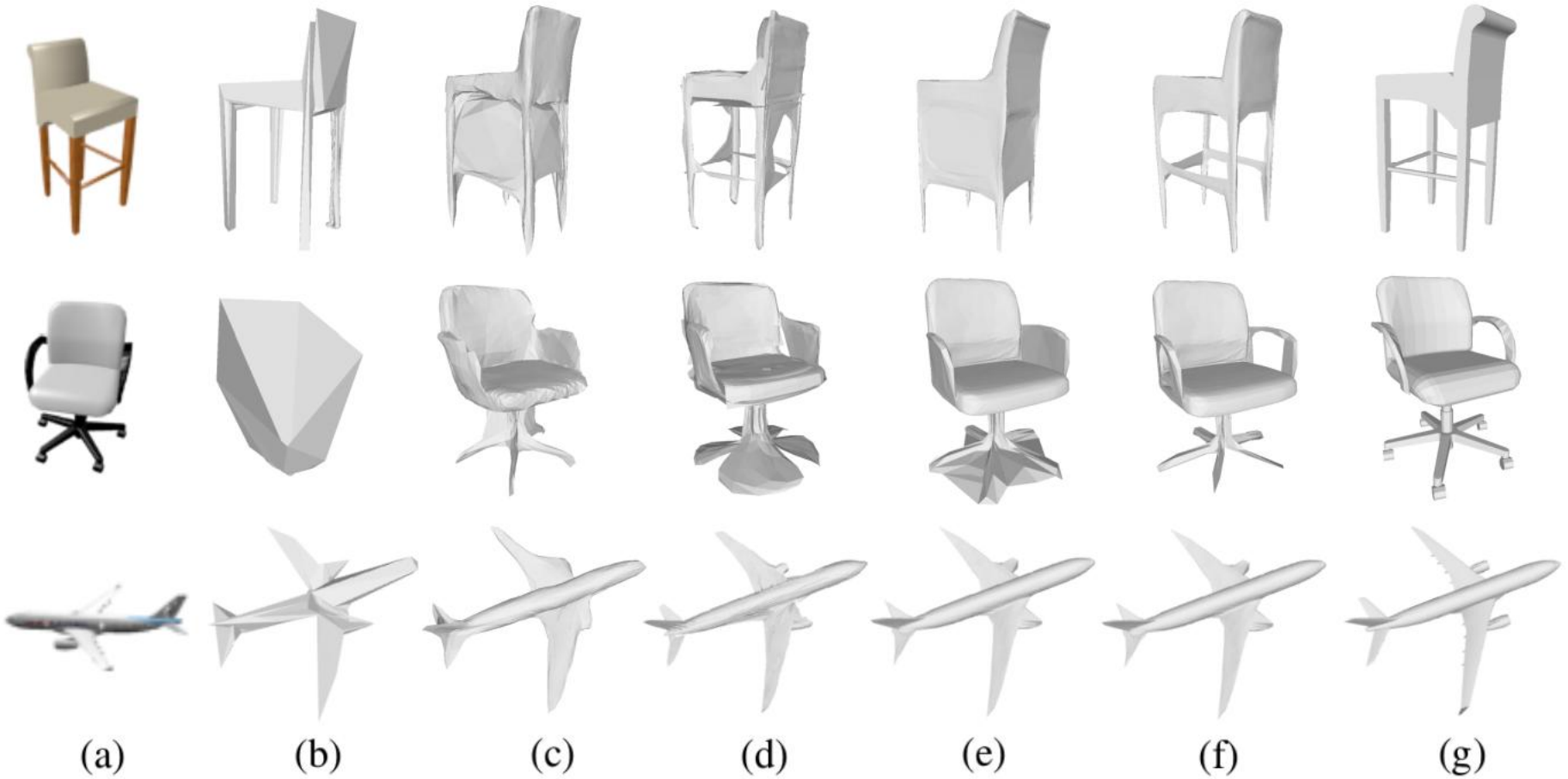
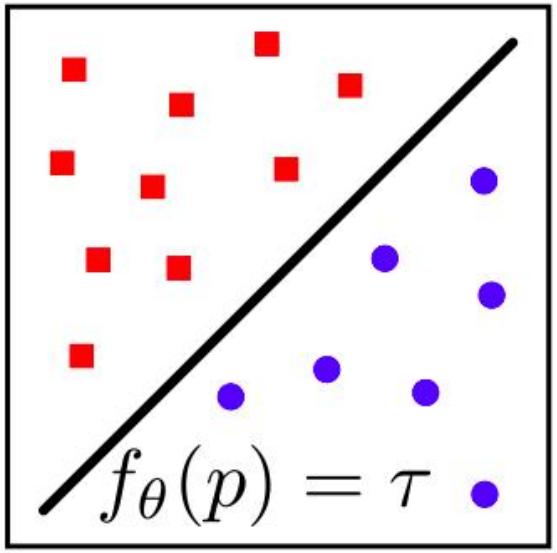


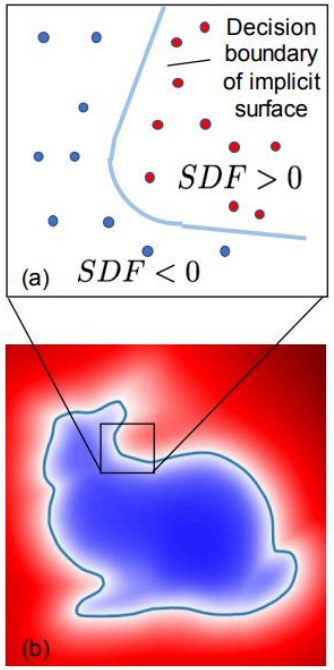
Figure 4. Qualitative results. (a) Input image; (b) N3MR; (c) Pixel2Mesh; (d) AtlasNet-25; (e) Baseline; (f) Ours; (g) Ground truth.

Deep Mesh Reconstruction from Single RGB Images via Topology Modification Networks

隐式三维表示



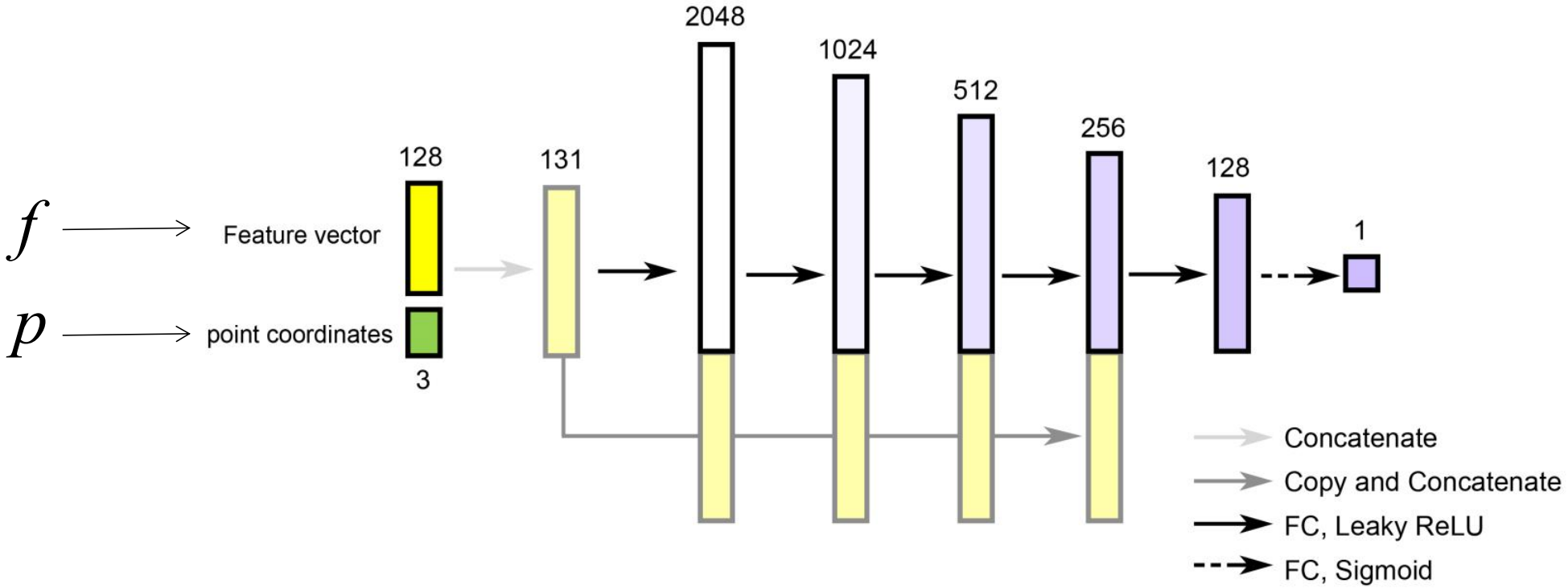
Occupancy Field



Signed Distance Field

神经隐式表示

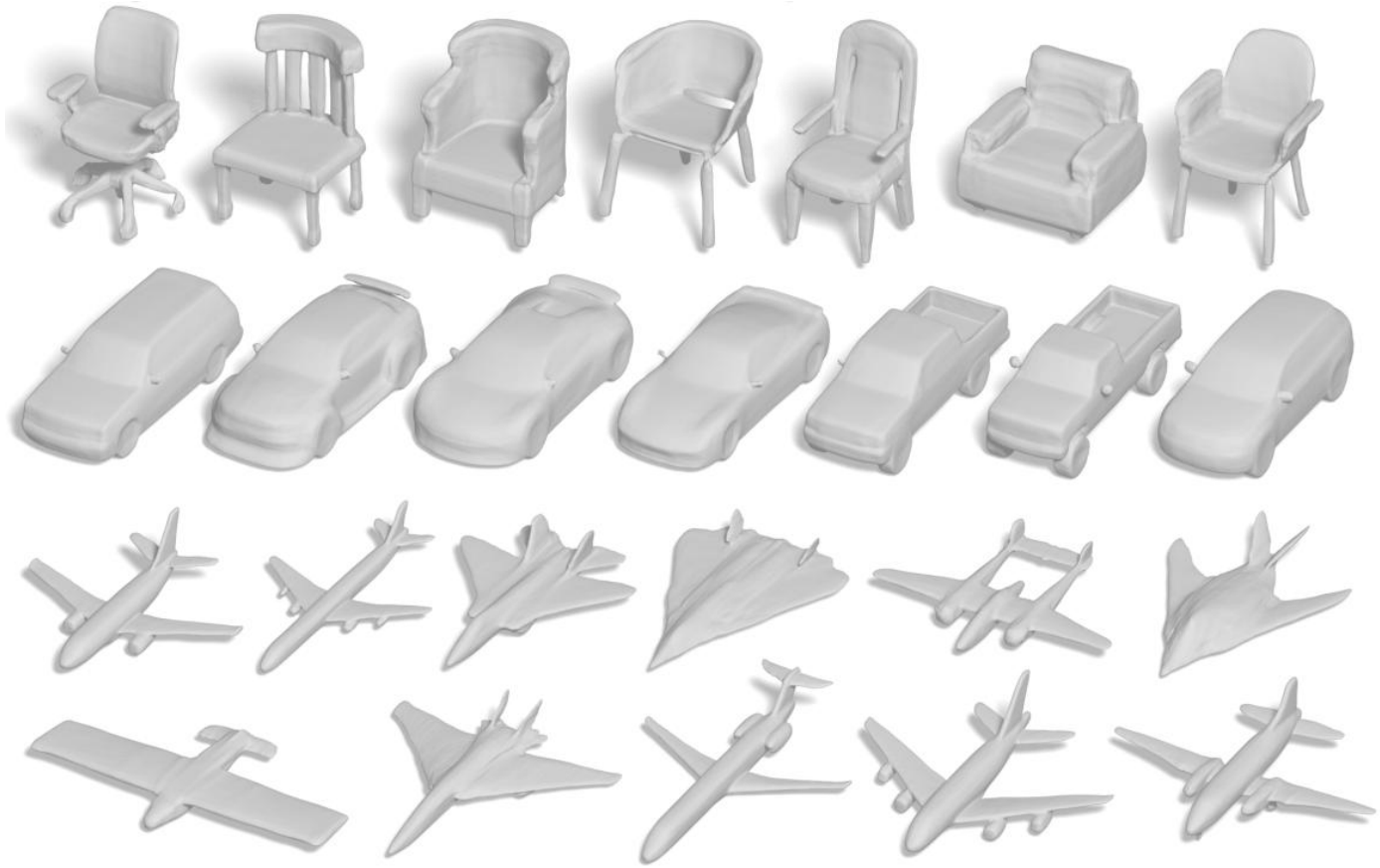
- Neural Implicit Representation：$\sigma = g_\theta(f, p)$



Learning Implicit Fields for Generative Shape Modeling

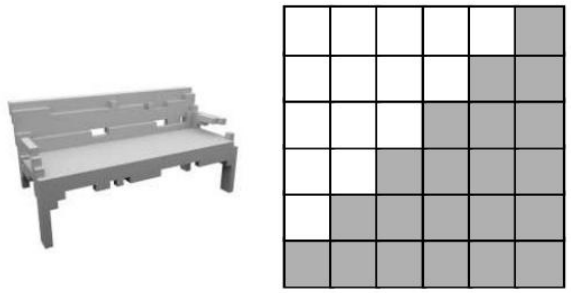
神经隐式表示

- Neural Implicit Representation：$\sigma = g_\theta(f, p)$



Learning Implicit Fields for Generative Shape Modeling

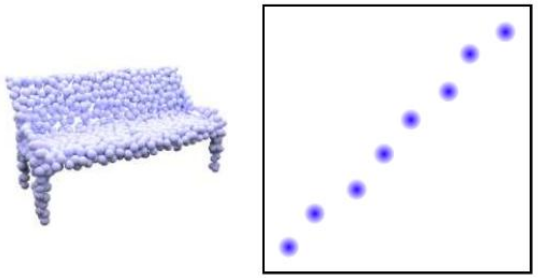
各方法优缺点



Voxels

Pros:
- Regular data suitable for convolutions

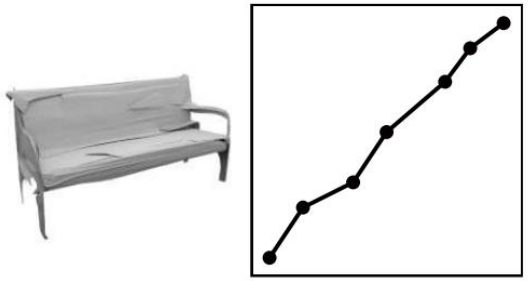
Cons:
- Memory consumption issues



Point Cloud

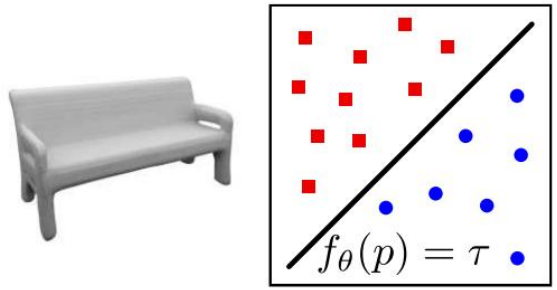
Pros:
- Simple and intuitive representation

Cons:
- Lack of topology information



Pros:
- Complex shapes

Cons:
- Poor quality of surfaces



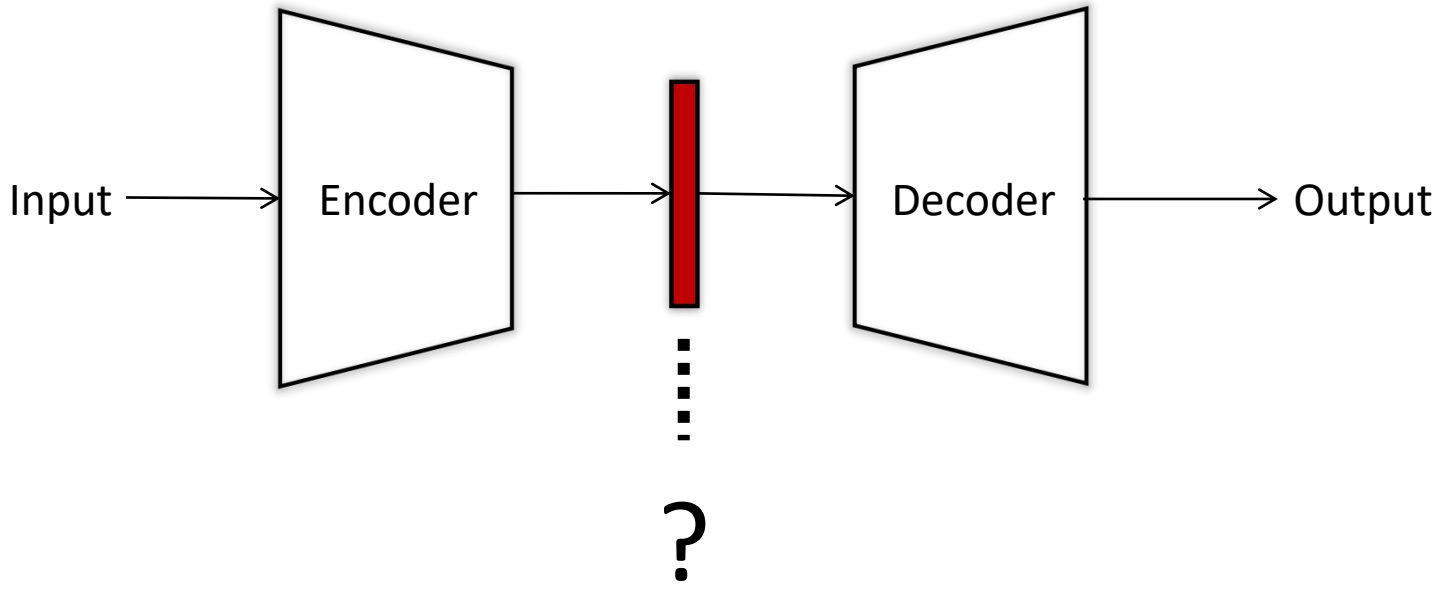
Pros:
- High quality of surfaces

Cons:
- Computation cost

Learning different levels of 3D priors

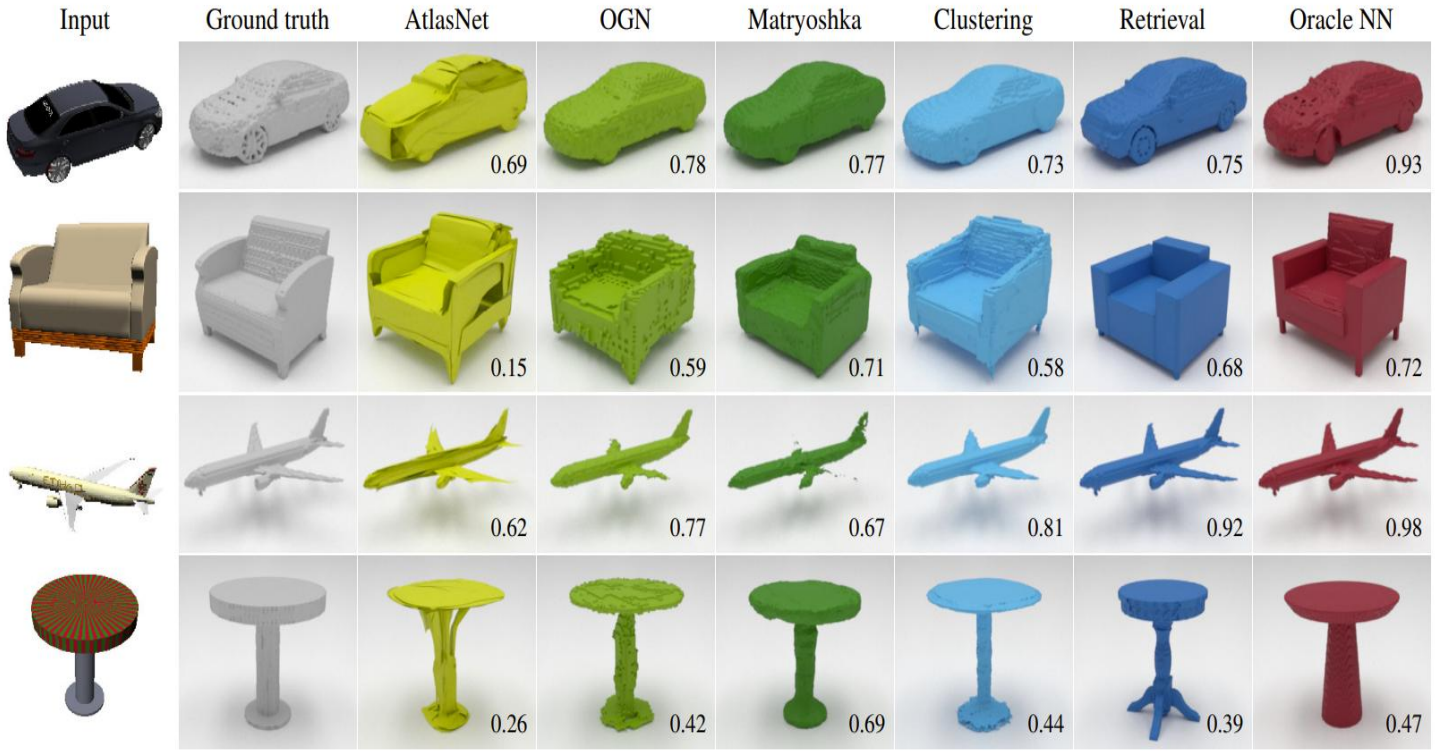
三维先验知识

In Bayesian statistics, a "prior" represents the beliefs we have before observing some data.



全局形状先验

Global prior: a shape is encoded as a feature sampled from the distribution.

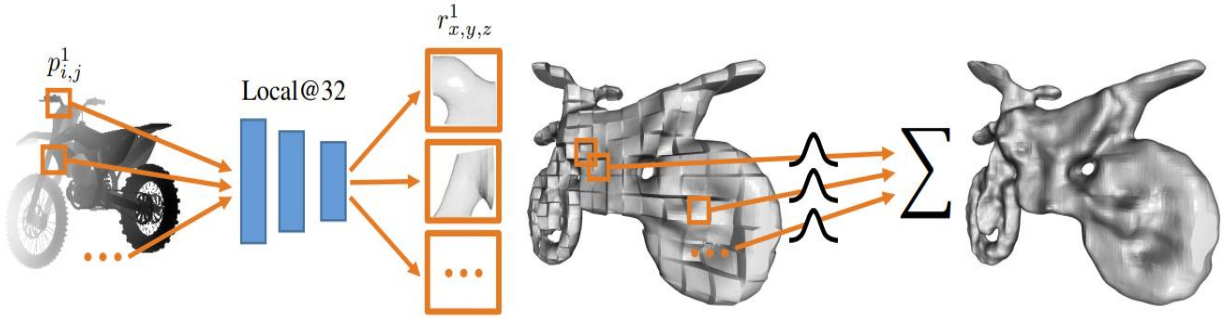


- The network usually learns several categories.
- Good quality, but poor generalization.

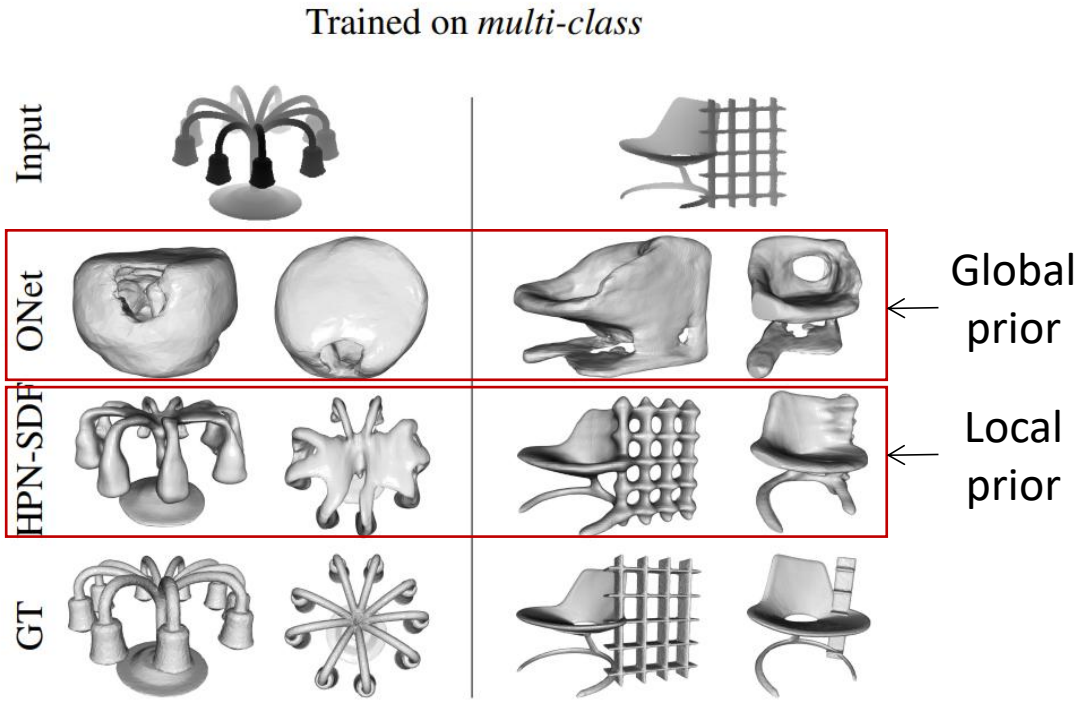
Why? Data? Network?

局部形状先验

Local prior: a shape is decomposed as lots of patches. Each patch is encoded as a feature sampled from the learned distribution.



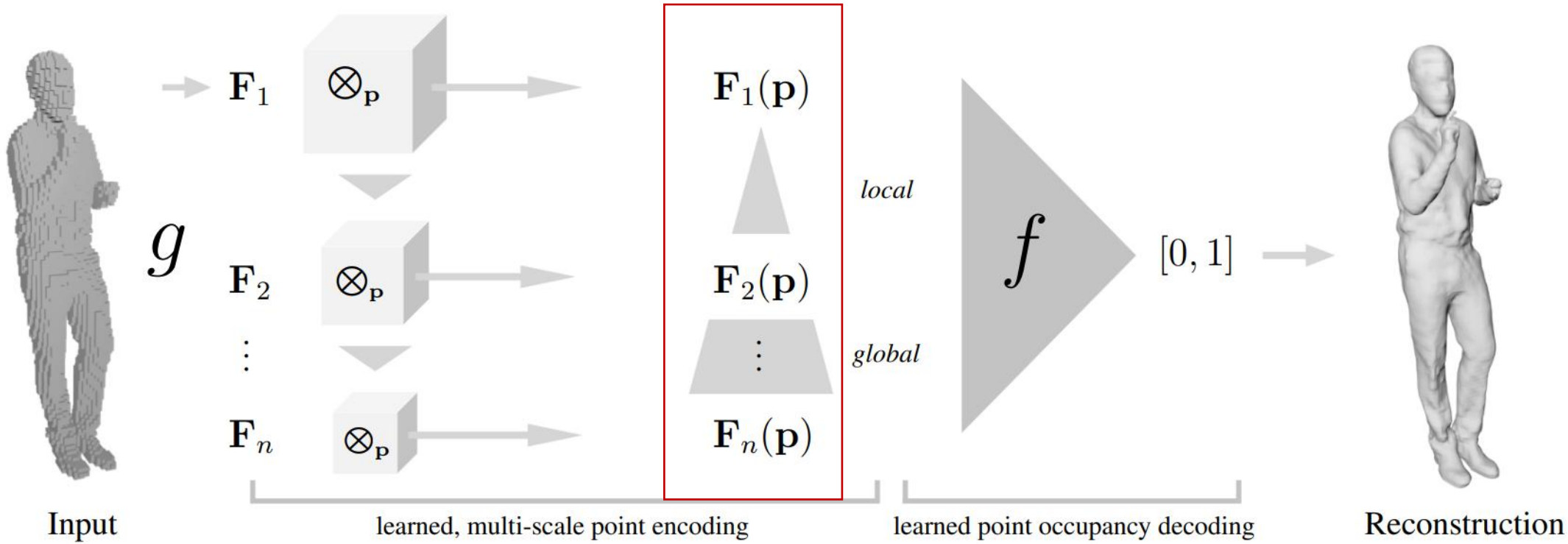
- Poor quality
- but better generalization ability



Fostering Generalization in Single-view 3D Reconstruction by Learning a Hierarchy of Local and Global Shape Priors

多层次特征融合

Multi-scale Feature Fusion

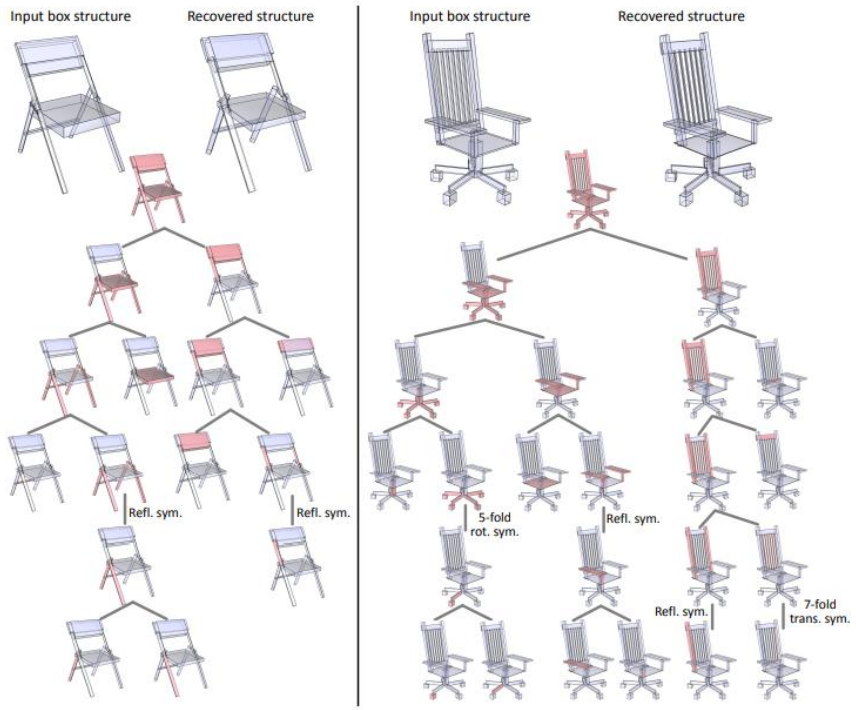


Implicit Functions in Feature Space for 3D Shape Reconstruction and Completion

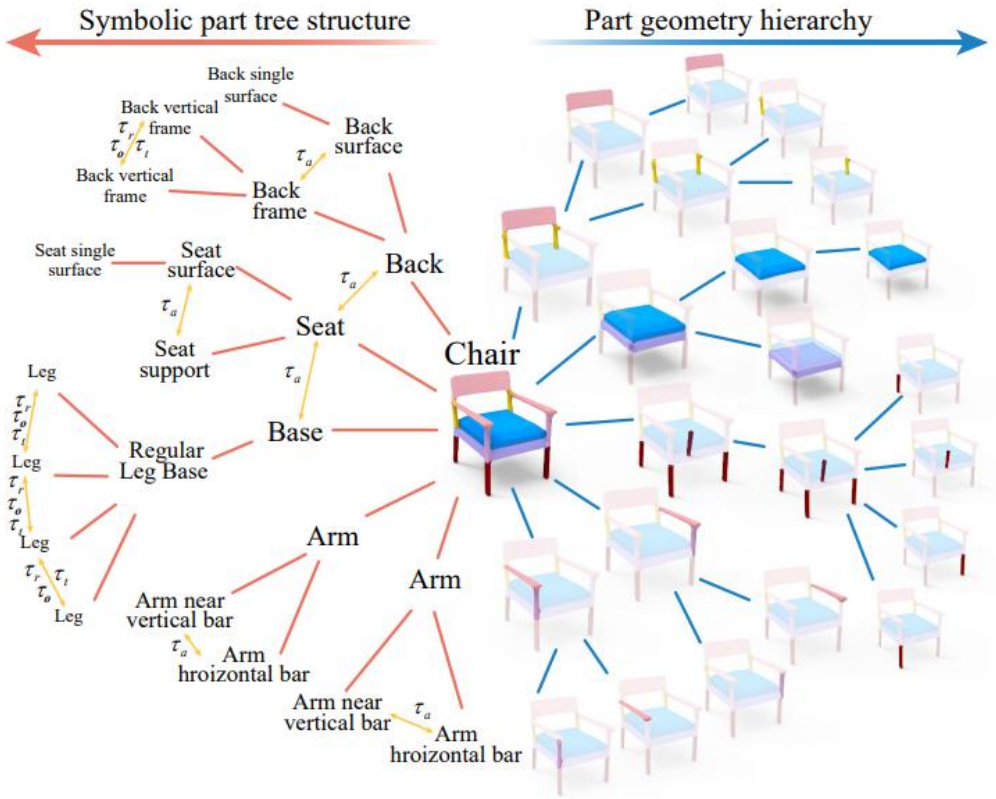
Learning other representations more than just a surface

结构化形状生成

Structured Shape Synthesis: Each object is decomposed as a set of parts and their assembly.



GRASS, 2017



DSG-Net, 2022

结构化形状生成

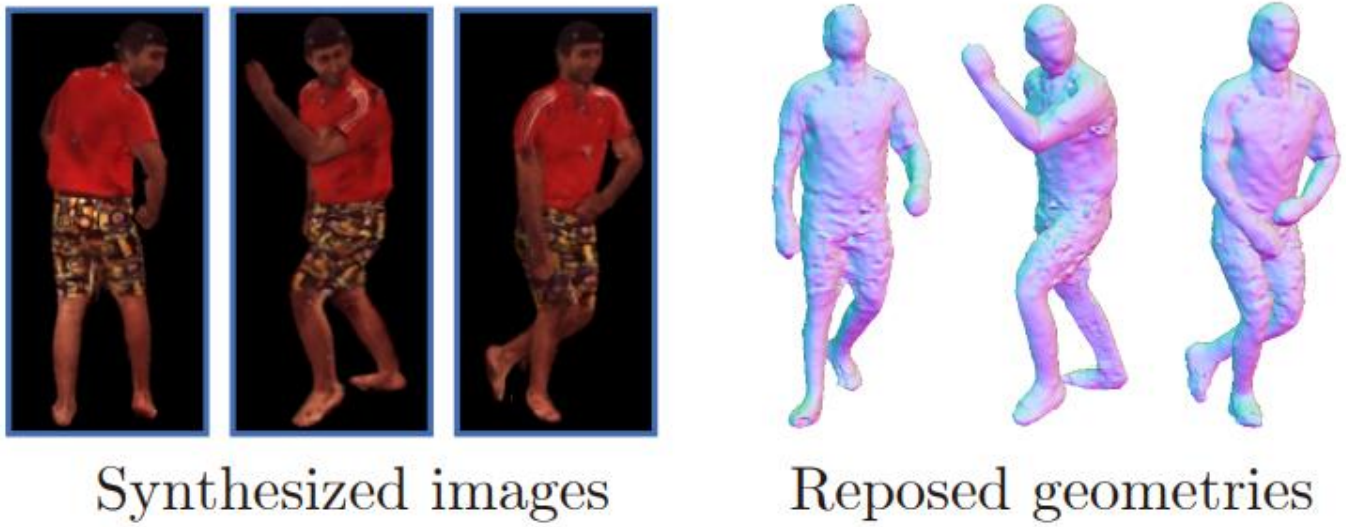
Structured Shape Synthesis: Each object is decomposed as a set of parts and their assembly.

Part-Mobility Dataset

Easy to represent the assembly and the part mobility

结构化形状生成

Structured Shape Synthesis: Each object is decomposed as a set of parts and their assembly.



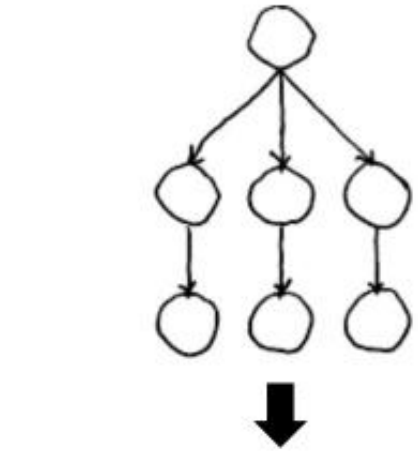
Animatable Human Reconstruction

Easy to represent the assembly and the part mobility

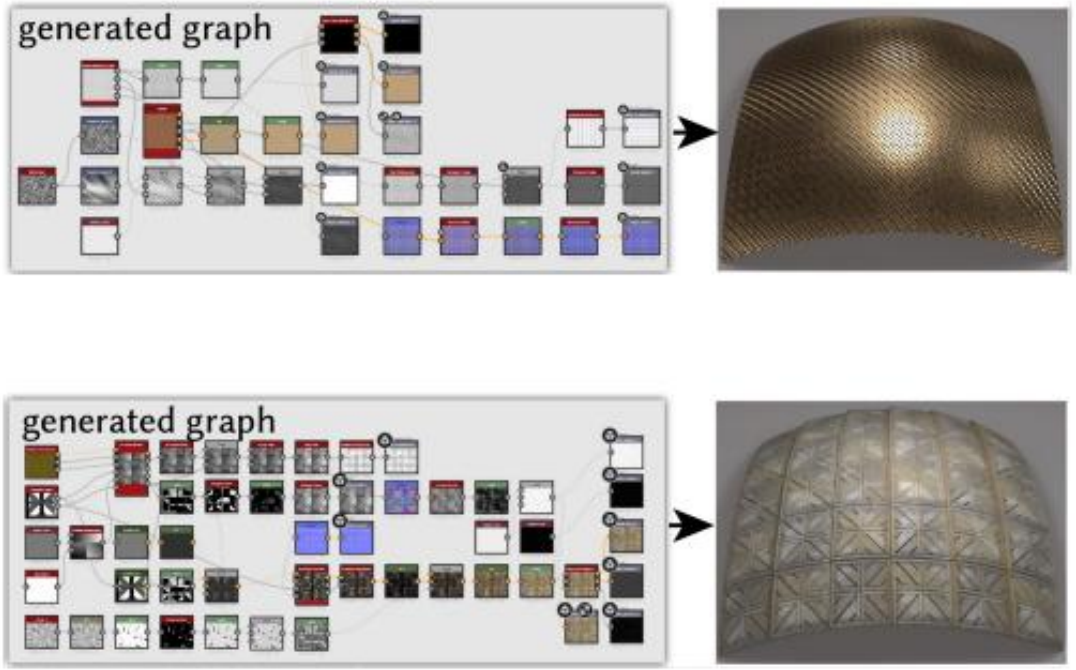
Animatable Neural Radiance Fields for Modeling Dynamic Human Bodies

神经符号表示

Neurosymbolic models produce visual data via a combination of symbolic programs and machine learning.



```
circle(4,10)
for(i<3)
 circle(-3*i+7,5)
 circle(-3*i+7,1)
 line(-3*i+7,4,-3*i+7,2,arrow)
 line(4,9,-3*i+7,6,arrow)
```



Neurosymbolic Models for Computer Graphics

Summary

Summary

- To synthesize 3D shapes, the first problem is to determine the proper data representation and shape prior.
- Targeting at different application purposes, sometimes we need to design our own data representation/structure.
- No one general method that rule all the applications!

Thank you