PEKING UNIVERSITY

SIGGRAPH 2025
Vancouver+ 10-14 August

基于八叉树的多尺度自回归模型

OCTGPT：OCTREE-BASED MULTISCALE AUTOREGRESSIVE MODELS FOR 3D SHAPE GENERATION





SI-TONG WEI, RUI-HUAN WANG, CHUAN-ZHI ZHOU, BAOQUAN CHEN, PENG-SHUAI WANG

北京大学


ChatGPT
image generation
follows
function
This entire poster was generated by ChatGPT image generation.


Tube Shorts

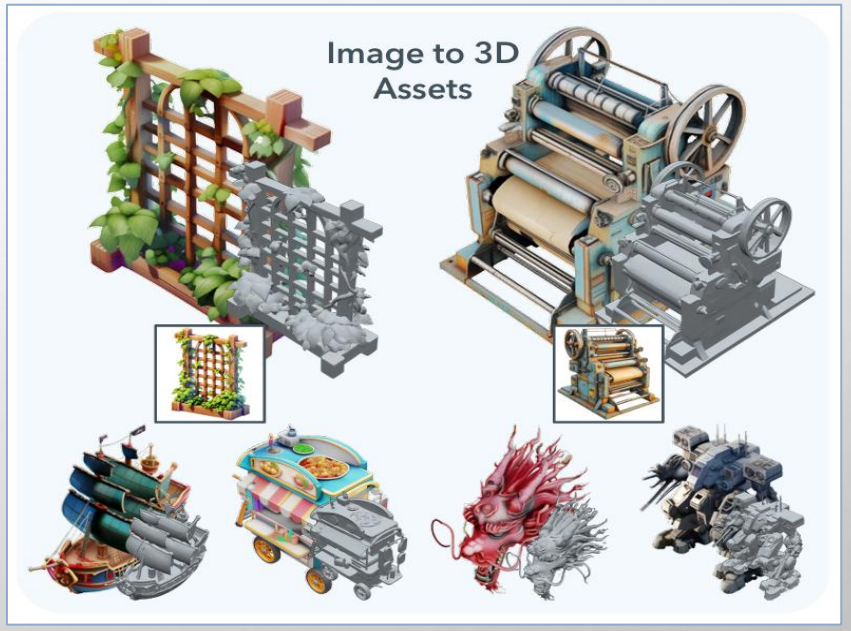
GPT-4o： 原生多模态

3D AIGC

目前主流的 3D AIGC 依赖于扩散模型

LAS-DIFFUSION[ZHENG ET AL. 2023], 3DSHAPE2VECSET [ZHANG ET AL. 2023], CLAY[ZHANG ET AL. 2024], XCUBE[REN ET AL. 2024], TRELLIS[XIANG ET AL. 2024] …

CLAY



Trellis

A 3D model of a Pokémon character

3D model of a chandelier hanging from a chain

3D model of a dragon with wings

自回归模型用于 3D AIGC 的挑战

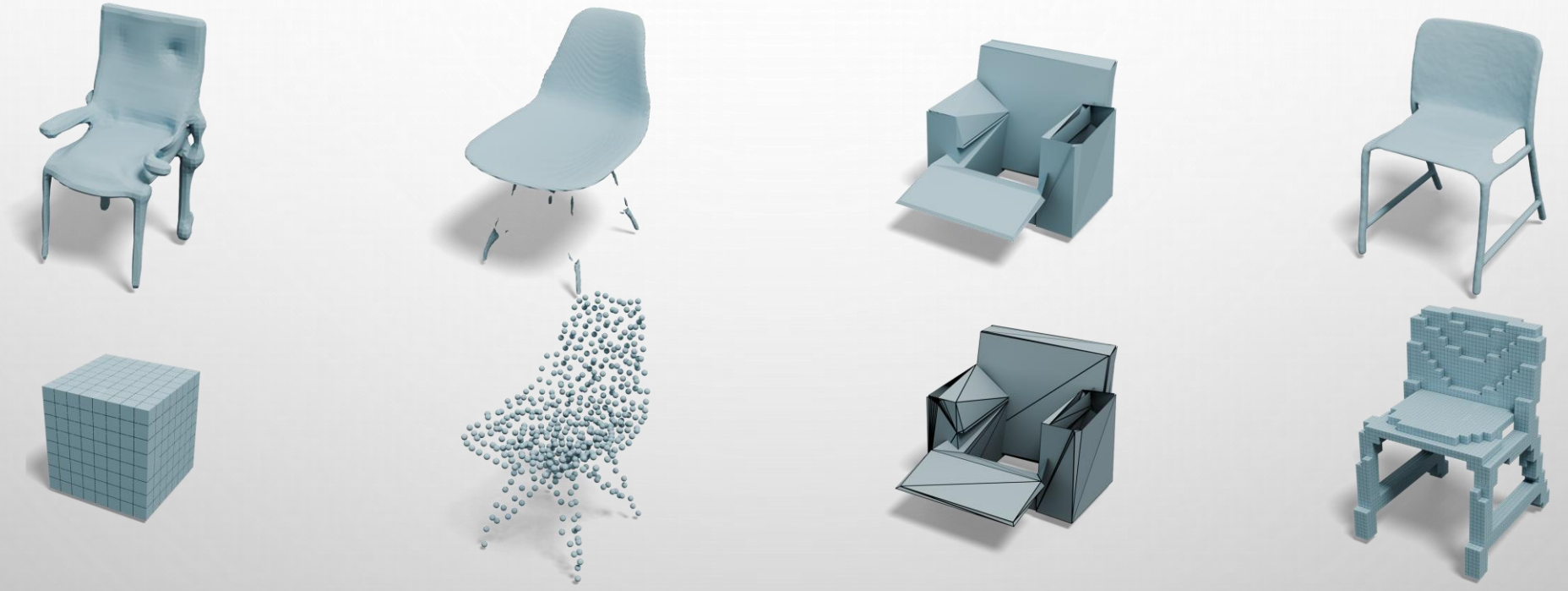
现有的序列化忽略了物体的层次结构和局部性，导致收敛变慢

RASTERIZATION ORDER: AUTOSDF[MITTAL ET AL. 2022], ARGUS3D[QIAN ET AL. 2024]

SERIALIZED MESHES: MESHGPT[SIDDIQUI ET AL. 2024], LLAMA-MESH[WANG ET AL. 2024]

自回归模型用于 3D AIGC 的挑战

TOKEN 长度较短 （<1K），无法捕获精细细节



AutoSDF
512 tokens

3DILG
512 tokens

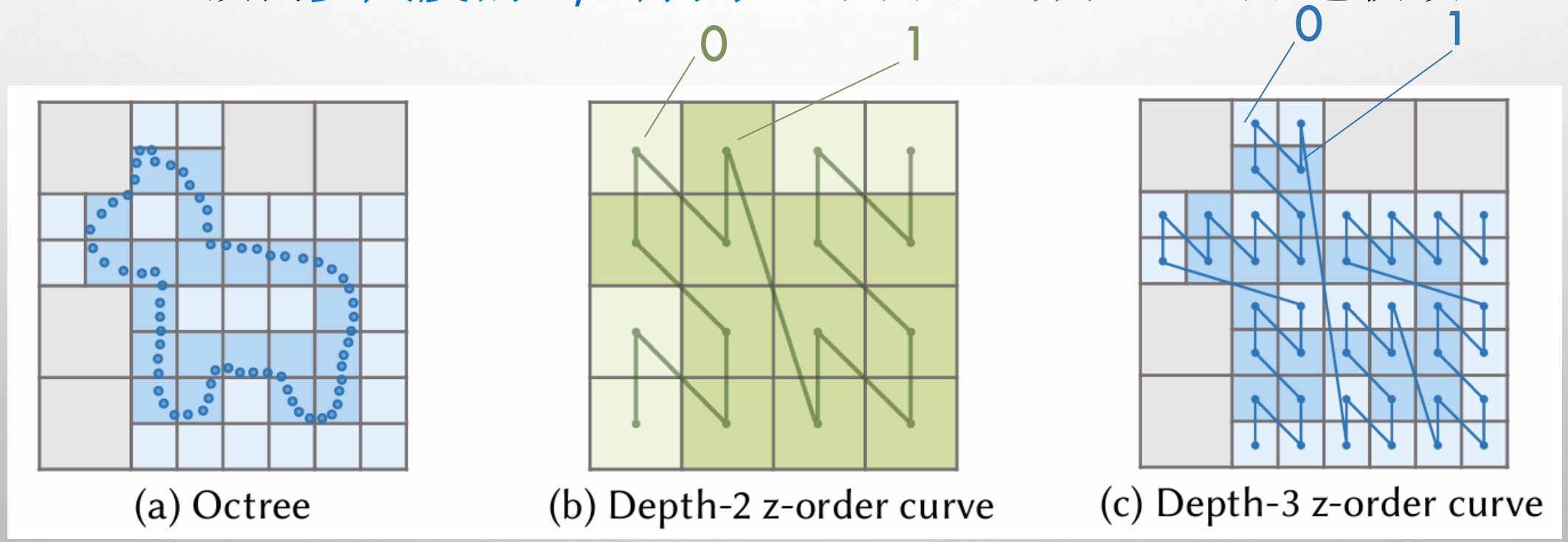
MeshGPT
187 tokens

Ours
49408 tokens

OCTGPT 的核心思想

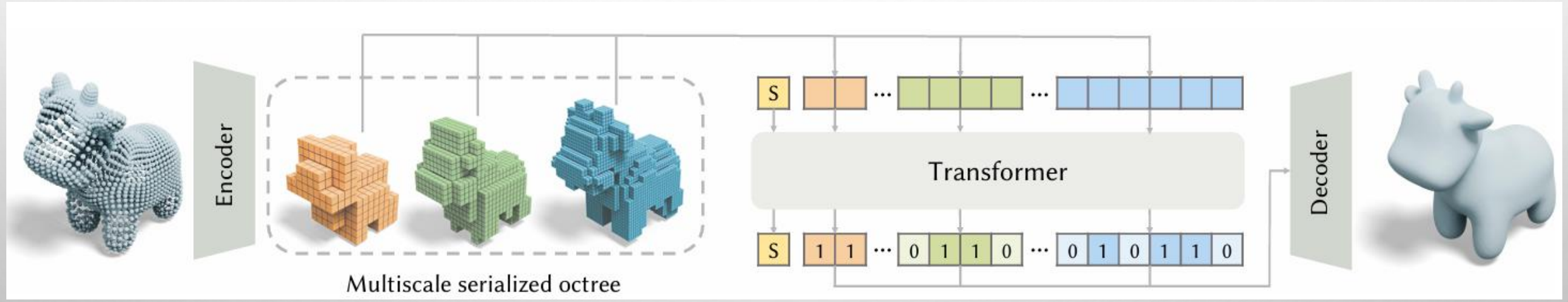
八叉树的Z-ORDER提供了多尺度的序列化，保持了空间局部性

GPT预测**多尺度的 0/1 序列**，~~而不是坐标位置~~，加速收敛



OCTGPT：基于八叉树的 VQVAE 和 TRANSFORMER

使用OctFormer [Wang SIGGRAPH 2023] 加速训练 **69** 倍
多个Token并行预测，加速推理 **13** 倍



4 Nvidia 4090 GPUs | 3天完成训练





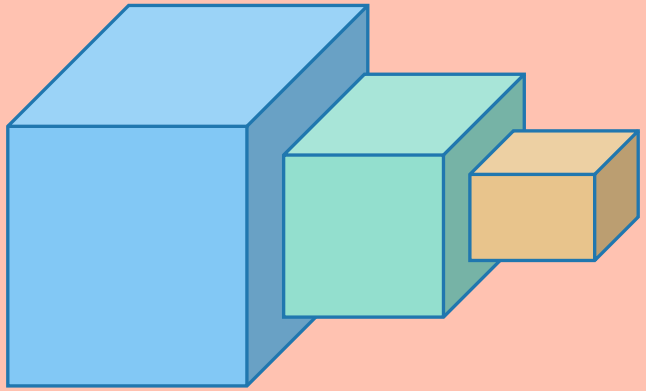







稀疏性

Multiscale Tokenization



多层级

序列性

Transformers