Tencent 腾讯

NeurIPS

Mesh-RFT: Enhancing Mesh Generation via Fine-grained Reinforcement Fine-Tuning

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

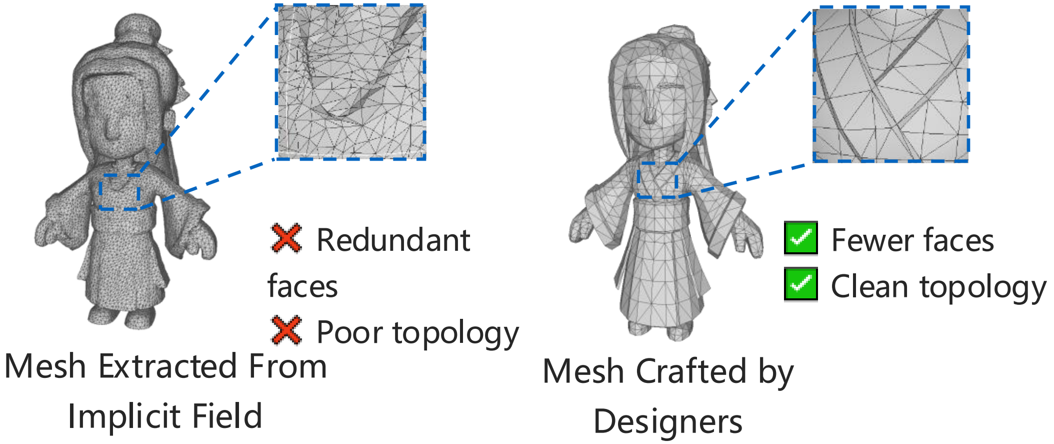
Jian Liu, HKUST&Hunyuan3D

腾讯混元

Native Mesh Generation: Taking 3D Generation from "Viewable" to "Usable"

Current Key Challenges in 3D Generation: The Inability to Produce Artist-like Meshes

While current 3D models achieve geometric precision, mesh topology quality lags artist standards, precluding direct use (e.g., gaming). Issues include:



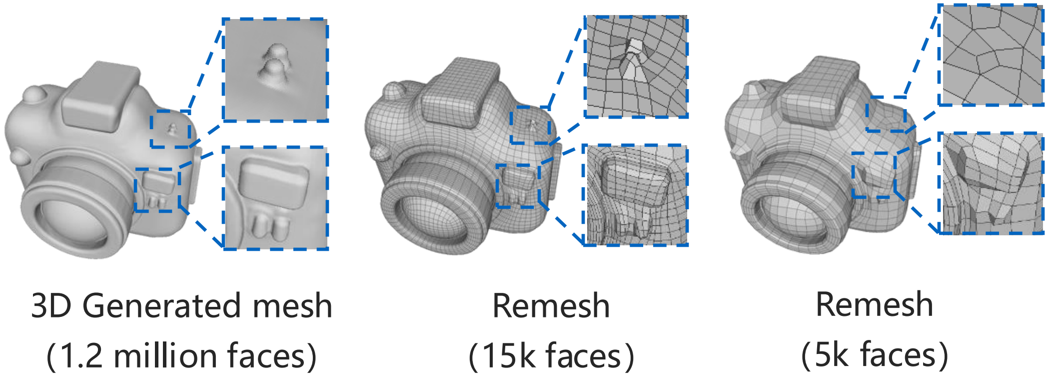
1) **Redundant Faces:** extracted meshes typically contain excessive polygons, which are unsuitable for real-time applications (e.g., games).
2) **Poor Topology:** Chaotic edge flow in extracted meshes blocks animation-ready workflows (e.g., rigging/skinning).

Casual Analysis:

Hunyuan2.0 employs Implicit Field and marching cube for mesh extraction in 3D generation, which only models geometric shapes without learning mesh topology.

Mainstream Solutions and Existing Issues:

Current methods use remesh algorithms for post-processing, improving mesh topology partially. But for low-poly counts in professional pipelines, they cause detail loss and fail on different objects.

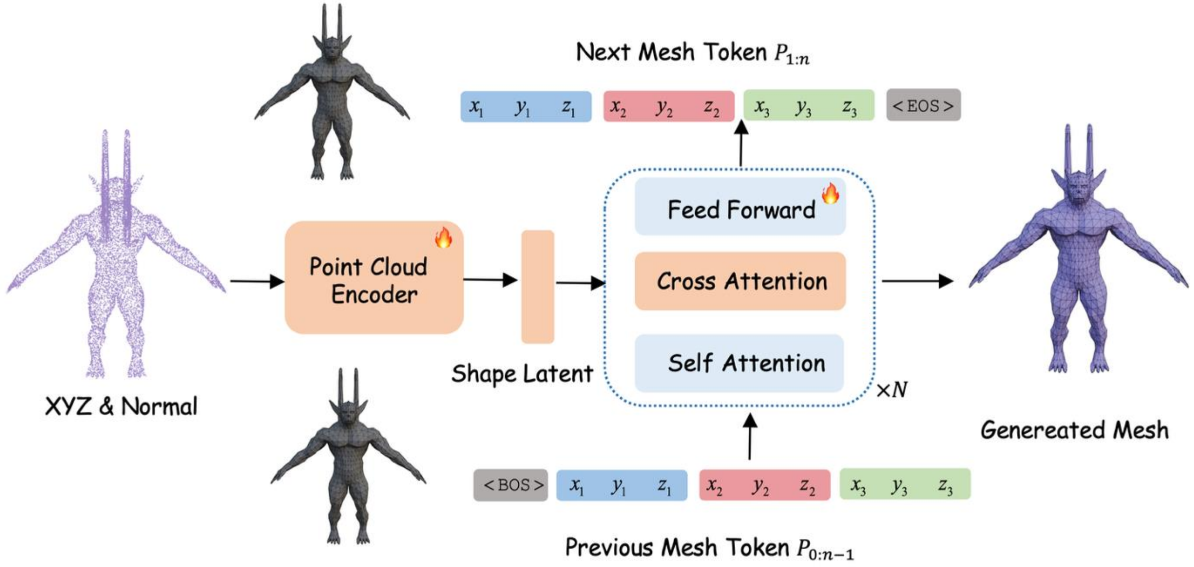


Native Mesh Generation: Artistic-Mesh Aligned 3D Generation

Native Mesh Generation via Autoregression — A new paradigm in 3D generation

- Core Idea: Explicitly model vertices and faces, directly learn edge topology from high-quality meshes.

- Formulate mesh generation as next-mesh-token prediction:
  1) Tokenize mesh (triangle soups) into a 1D coordinate sequence.
  2) Model the mesh sequence with GPT-style Transformer.
  3) Decode the sequence back to the final mesh.



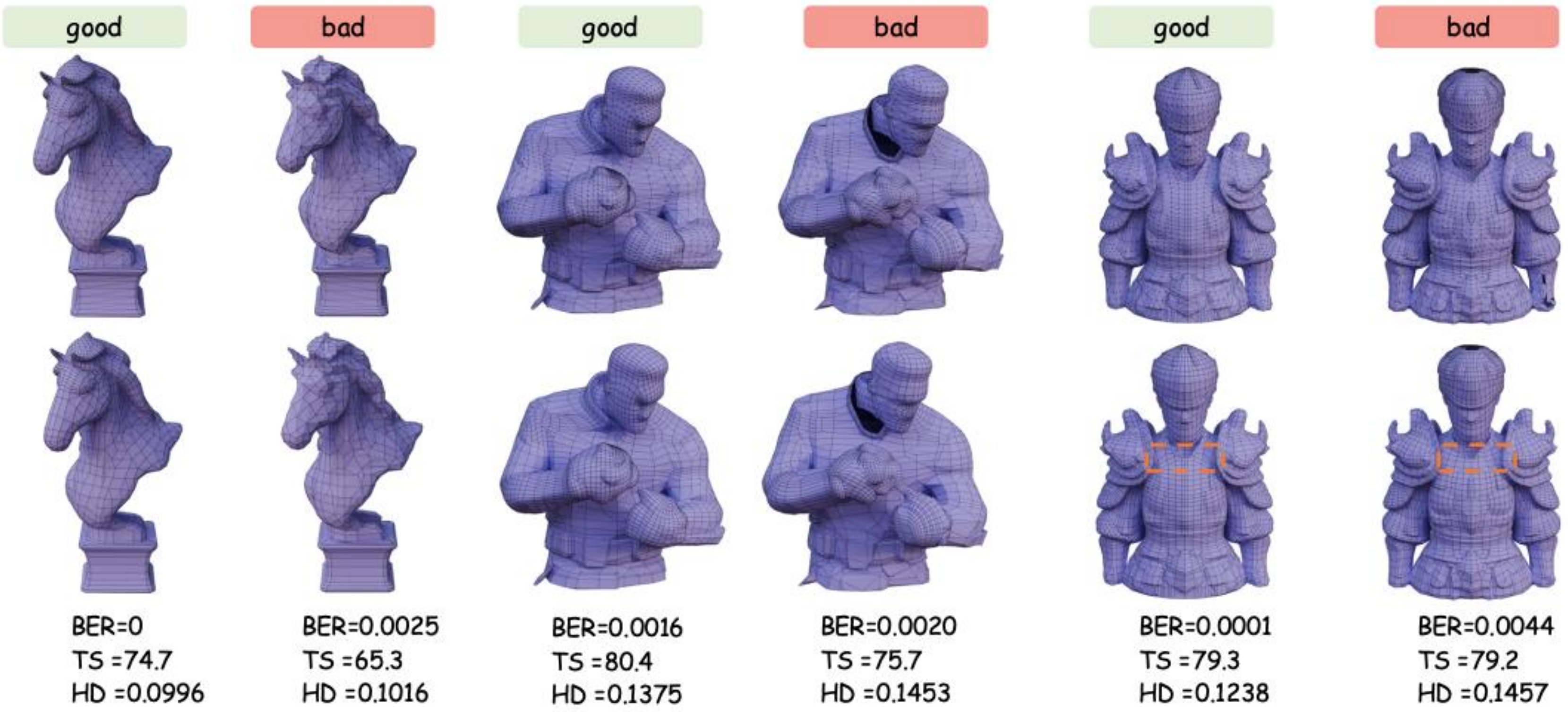
Autoregressive Mesh Generation FrameWork

Examples

Fewer faces、Clean topology、Part structure

Native Mesh Generation: Current Challenge

Challenge : Inconsistent Pre-trained Outputs

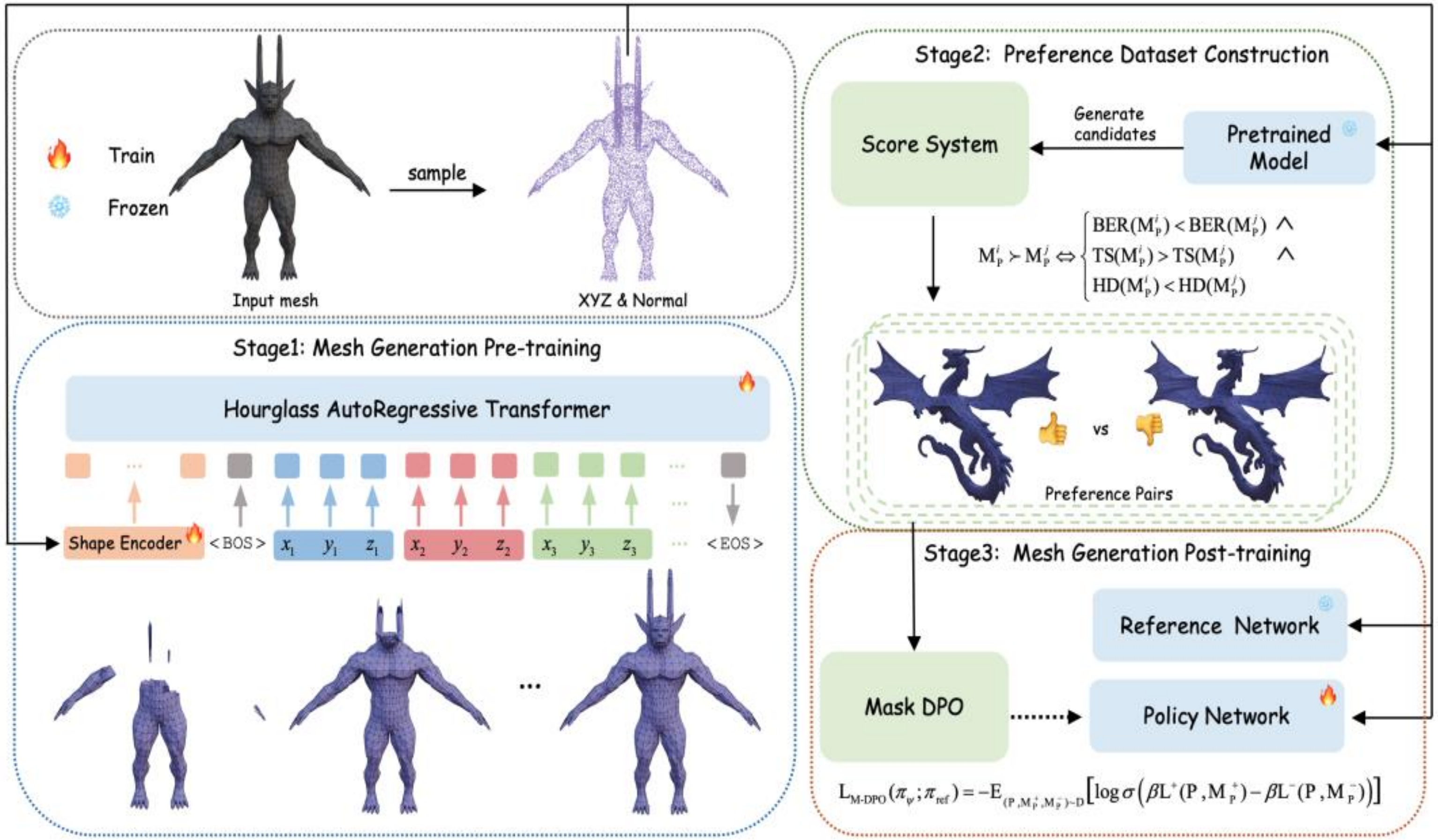


Observation: Pre-trained autoregressive models produce meshes of mixed quality.
Problem: Well-structured regions ("good") often coexist with severe topological defects ("bad"). A global reward signal would damage the "good" parts while trying to fix the "bad".

Q: How can we selectively refine only the flawed regions?

Our Method: Mesh-RFT

Pioneering the "Pre-train & Fine-tune" Paradigm for Mesh Generation



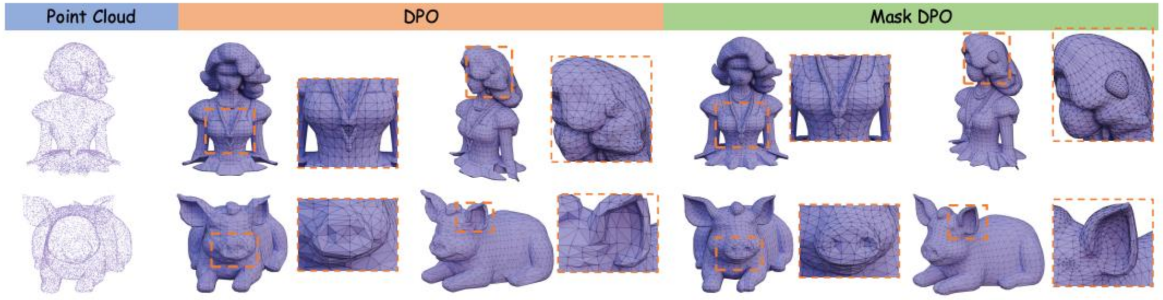
Topology-Aware Scoring System: We propose two objective metrics, Boundary Edge Ratio (BER) and Topology Score (TS), to automatically evaluate mesh quality at both object and face levels, eliminating the need for manual annotation.

Masked Direct Preference Optimization: A novel fine-grained RL strategy that uses quality-aware masks to focus the optimization on low-quality regions of the mesh, enabling localized refinement while preserving global structure.

Mesh-RFT：Experiment Results

Superior Geometric Fidelity & Topological Coherence

Data Type	Artist Meshes					Dense Meshes				
Metrics	CD ↓	HD ↓	TS ↑	BER ↓	US ↑	CD ↓	HD ↓	TS ↑	BER ↓	US ↑
MeshAnythingv2 [16]	0.2143	0.4197	68.3	0.0749	9%	0.2265	0.4760	72.0	0.0913	8%
BPT [19]	0.1275	0.2735	72.7	0.0280	20%	0.1615	0.3347	73.7	0.0113	18%
DeepMesh* [23]	0.1331	0.2866	74.9	0.0296	22%	0.1760	0.3570	75.8	0.0044	20%
Ours	**0.0973**	**0.1826**	**77.5**	**0.0182**	**45%**	**0.1286**	**0.2411**	**79.4**	**0.0015**	**40%**



Method	**CD ↓**	**HD ↓**	**TS ↑**	**BER ↓**	**US ↑**
Pretrain	0.1588	0.3196	76.5	0.0033	30%
N-DPO	0.1455	0.2919	75.7	0.0028	32%
S-DPO	0.1348	0.2625	77.9	0.0023	35%
M-DPO	**0.1286**	**0.2411**	**79.4**	**0.0015**	**40%**

Effectiveness of Mask-DPO

Mesh-RFT：Demo

Tencent腾讯

Thanks for your listening!

香港科技大學
THE HONG KONG UNIVERSITY OF SCIENCE AND TECHNOLOGY

腾讯混元