

混合表达与生成式重建渲染

高林

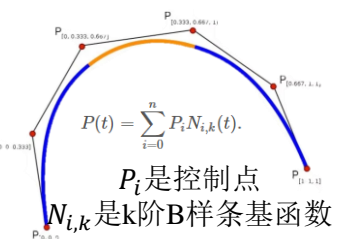
中国科学院计算技术研究所

2025.04.13

几何表示的基础作用



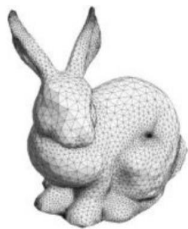
几何表示的发展



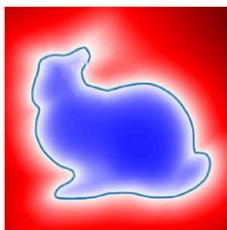
样条曲线



点云



三角网格



隐式场



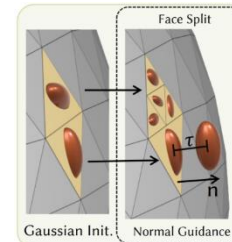
神经辐射场



高斯泼溅



CLAY



高斯网

1975

1990

2005

2020

2023

2024

2020

2021

2022

2023

2024

NeRF
最佳论文提名
ECCV 2020



GIRAFFE
最佳论文
CVPR 2021



Mip-NeRF
最佳论文提名
ICCV 2021



InstaNGP
最佳论文
SIGGRAPH 2022



Ref-NeRF
Best Student Paper
Honorable Mention
CVPR 2022



Gaussian Splatting
最佳论文
SIGGRAPH 2023



Tri-MipRF
最佳论文提名
ICCV 2023



Mip-Splatting
最佳论文
CVPR 2024

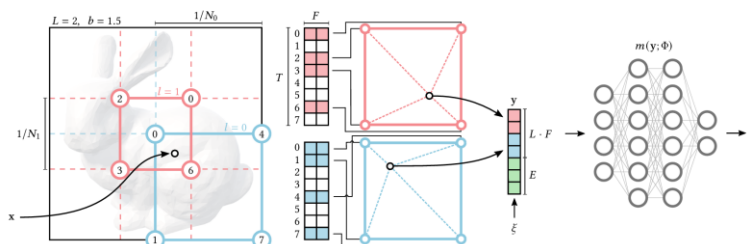


VEGS
3DGS街景
ECCV 2024



混合表达与生成式重建

Instant-NGP - SIGGRAPH' 22 最佳论文



3D GS - SIGGRAPH' 23 最佳论文



CLAY - SIGGRAPH' 24 最佳论文提名

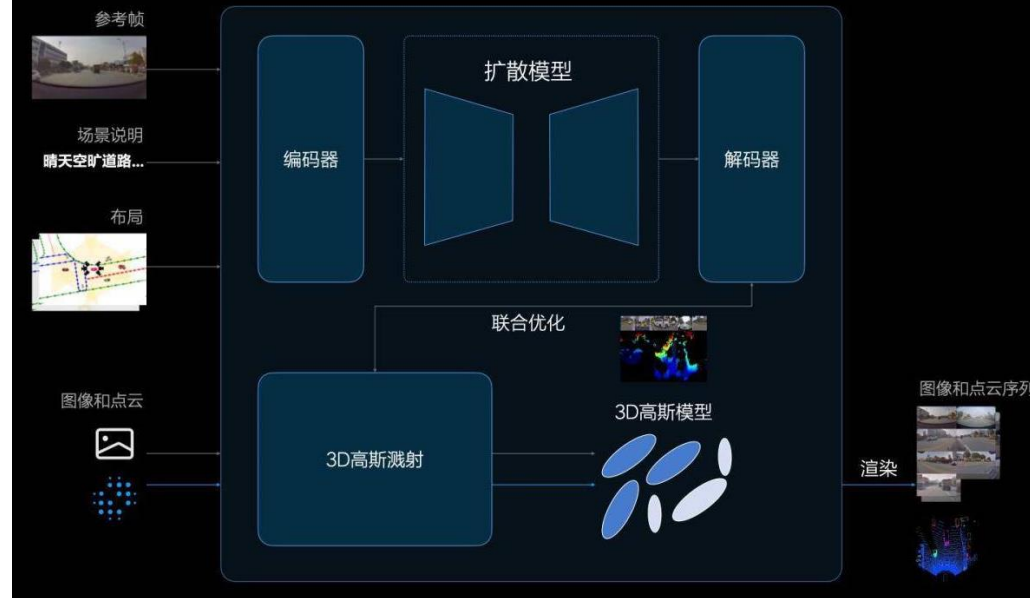


重建是带约束的生成

VEGS - ECCV' 24

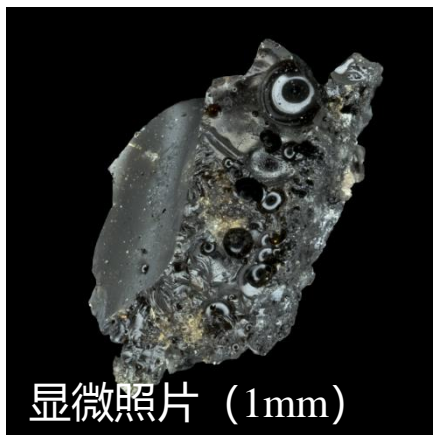


理想汽车的自动驾驶系统考试方案：重建+生成的世界模型



生成式重建是未来的趋势

基于单张显微照片的三维月壤颗粒重建



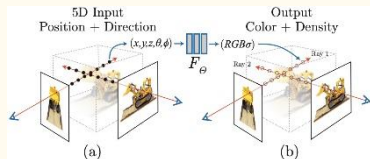
通过单张显微照片
三维重建



生成式渲染



光线追踪渲染



神经辐射场
ECCV2020
最佳论文提名



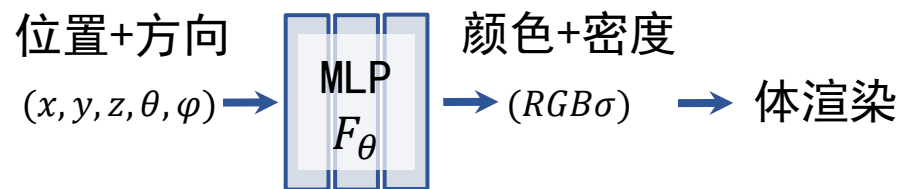
视频生成模型
Sora, OpenSora
CogVideo, ...

神经渲染

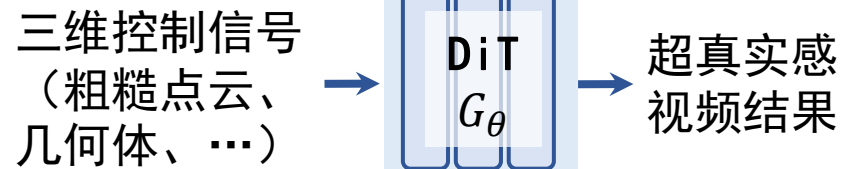
神经渲染优势

超真实感、多样性、动态性

神经辐射场 NeRF

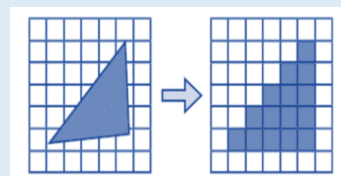


DiT生成模型

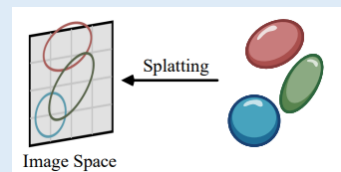


逐像素
渲染

生成式
渲染



传统图形
像素化渲染



高斯泼溅
SIGGRAPH 2023
最佳论文

光栅化渲染

现有视频生成模型缺陷

Runway Gen3



软体模拟不真实

CogVideoX I2V



刚体碰撞不真实

Sora



流体运动不真实

Owl-1



长视频生成突变

基于线稿的视频生成和编辑

难点

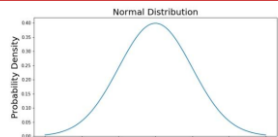
文本输入

矛盾

精确控制视频生成过程

目标：基于线稿精确控制视频生成和编辑过程。

“一只猫咪在镜头前移动”



视频生成模型



几何随机、动作随机



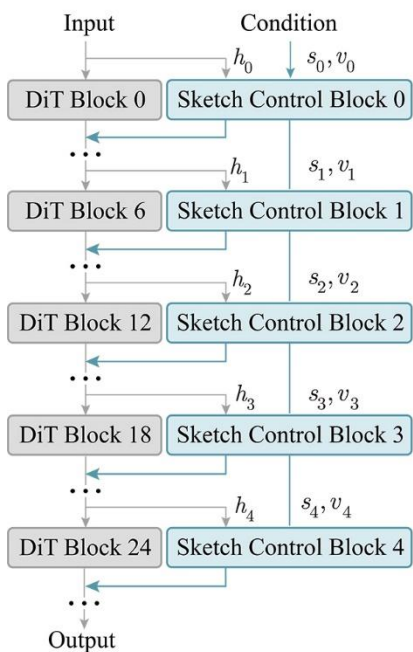
“一只带有浅色礼帽的小狗”



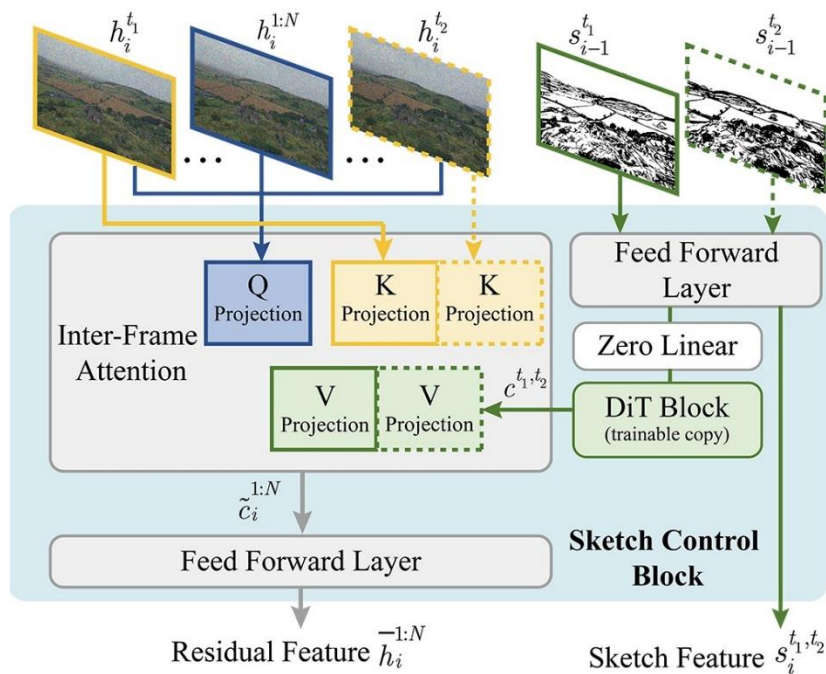
无法精确二次编辑

基于线稿的视频生成和编辑

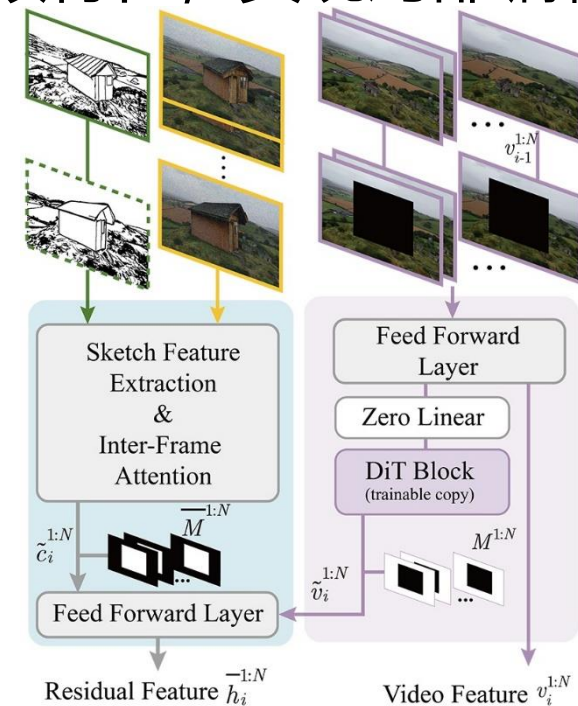
- 设计跳跃连接条件控制网络，实现有效的几何控制
- 稀疏线稿传播：时序注意力机制，根据帧间相似度，传播线稿控制特征，残差叠加
- 视频融合机制：根据编辑区域的mask，融合线稿与原始视频特征，实现局部编辑



跳跃连接结构

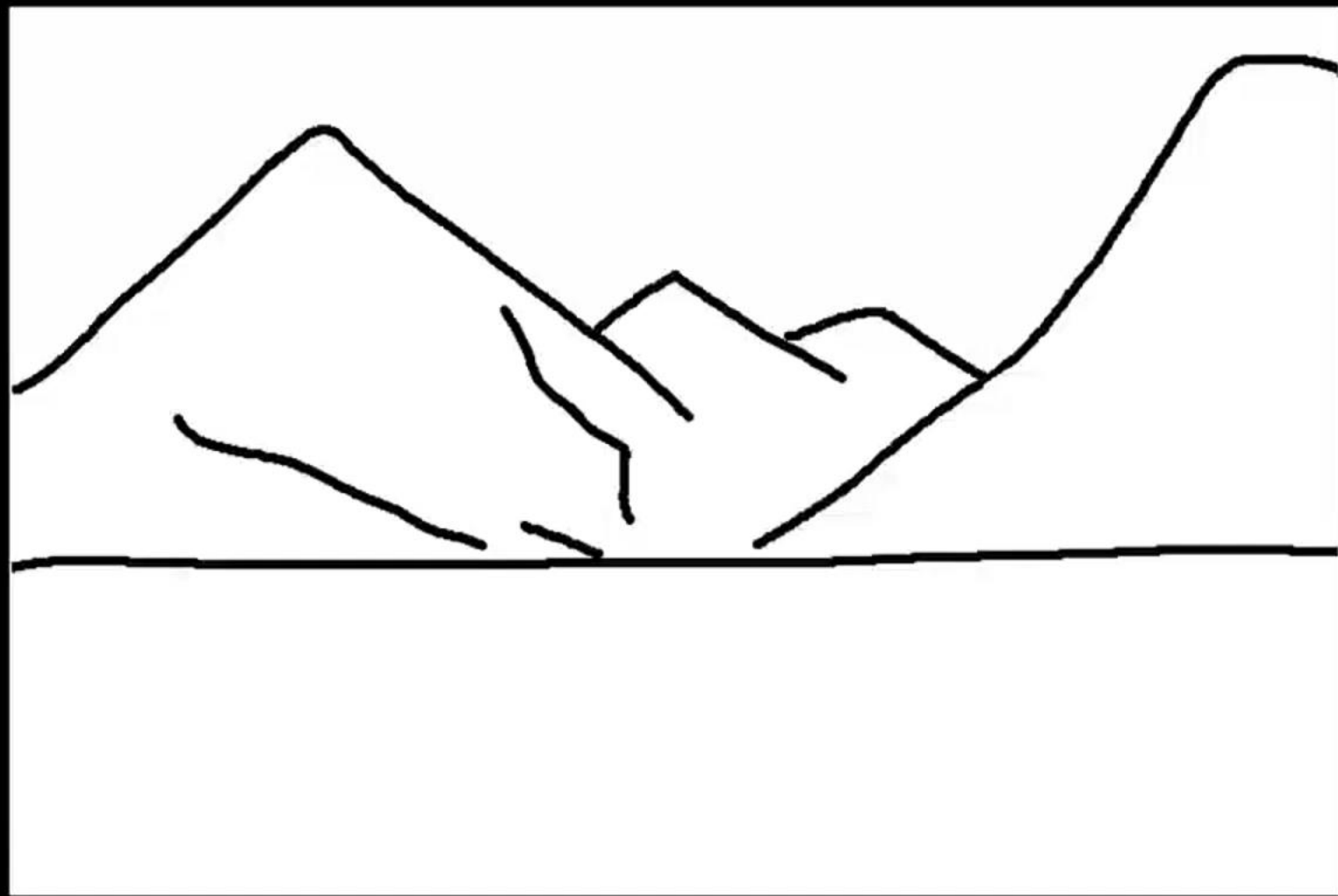


稀疏线稿传播



视频融合机制

视频生成



视频编辑

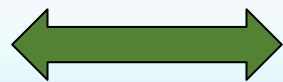


基于线稿的三维场景视频编辑

难点

二维稀疏线稿输入

矛盾



三维交互及新视角生成

目标：基于线稿实现大视角变换的场景视频编辑。



无法处理几何编辑

视频编辑方法[1]



无法处理大视角变换

视频编辑方法[2]



存在混合瑕疵

三维引导生成方法[3]

[1] AnyV2V: A Tuning-Free Framework For Any Video-to-Video Editing Tasks.

[2] I2VEdit: First-Frame-Guided Video Editing via Image-to-Video Diffusion Models.

[3] ViewCrafter: Taming Video Diffusion Models for High-fidelity Novel View Synthesis.

基于线稿的三维场景视频编辑

- 构建三维点云几何代理：预测编辑后的深度图，与原始深度图对齐，三维反投影。
- 视频生成模型渲染：根据编辑后的几何，获取三维Mask并进行渲染。以编辑后的三维点云作为控制条件，引导视频生成。



Generation

Input Image



Camera Trajectory

Editing



Input Video

可重光照的高效三维高斯人头生成

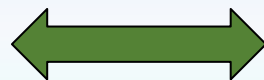
难点

材质和光照二义性



可重光照三维人头高质量生成

结合



利用三维高斯泼溅进行高效生成与渲染

目标：材质与光照解耦的三维人头高效生成与渲染

自然图像数据学习Phong模型，
主要处理白光



FaceLit CVPR 2023
35FPS / 0.5FPS

自然图像数据学习阴影三平面
解耦，仅处理白光



NeRFFaceLighting TOG 2023
37FPS / 0.5FPS

OLAT数据学习阴影三平面解耦，
可处理彩色光

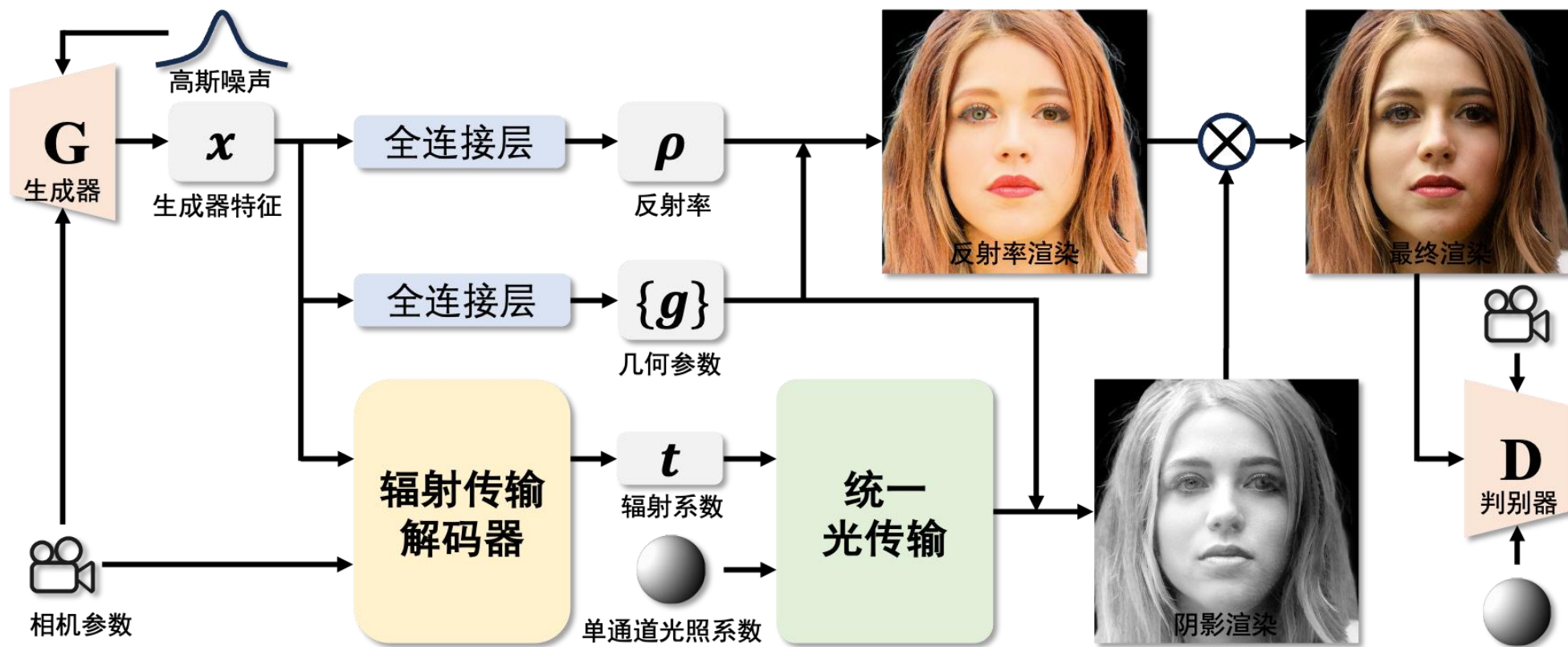


Holo-Relighting CVPR 2024

可重光照的高效三维高斯人头生成

- 仅从自然图像数据学习光照解耦，可以处理彩色光渲染，渲染速度与原生三维高斯泼溅保持一致

统一辐射传输，线性光照模型，消除光照和反照率二义性



可重光照的高效三维高斯人头生成

240+ FPS



Env. Map Rendering



Env. Map
& Shading

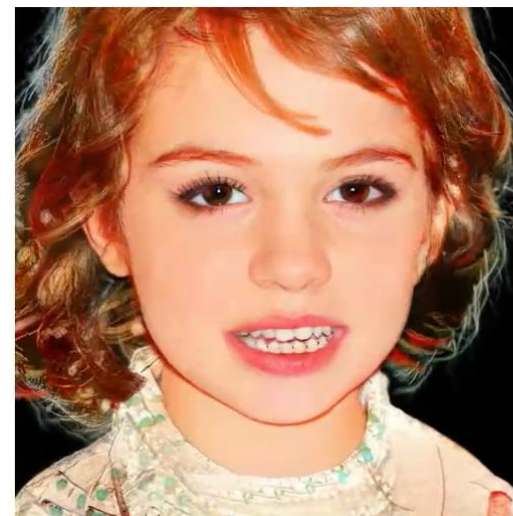
Albedo

Relighting

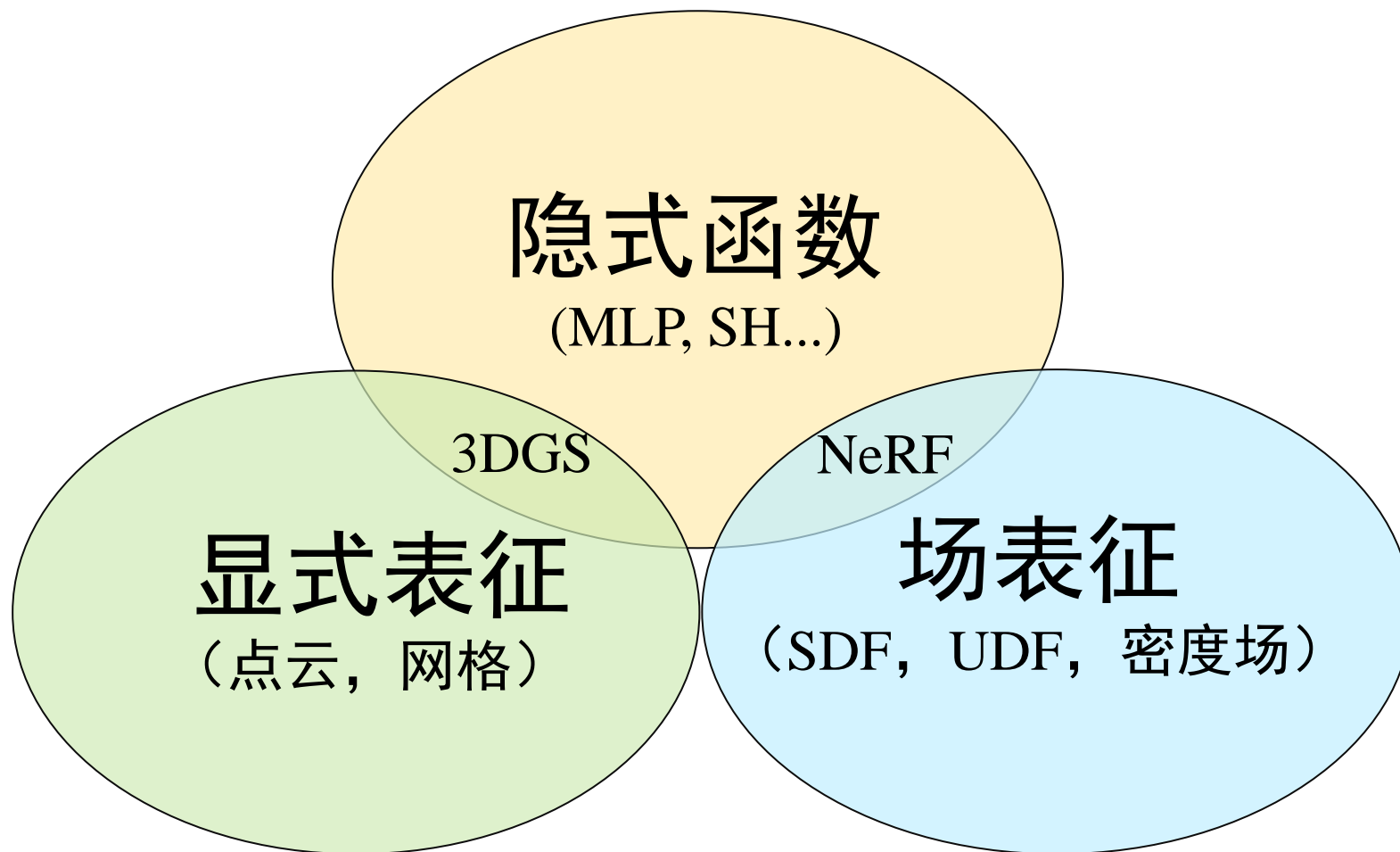


Albedo

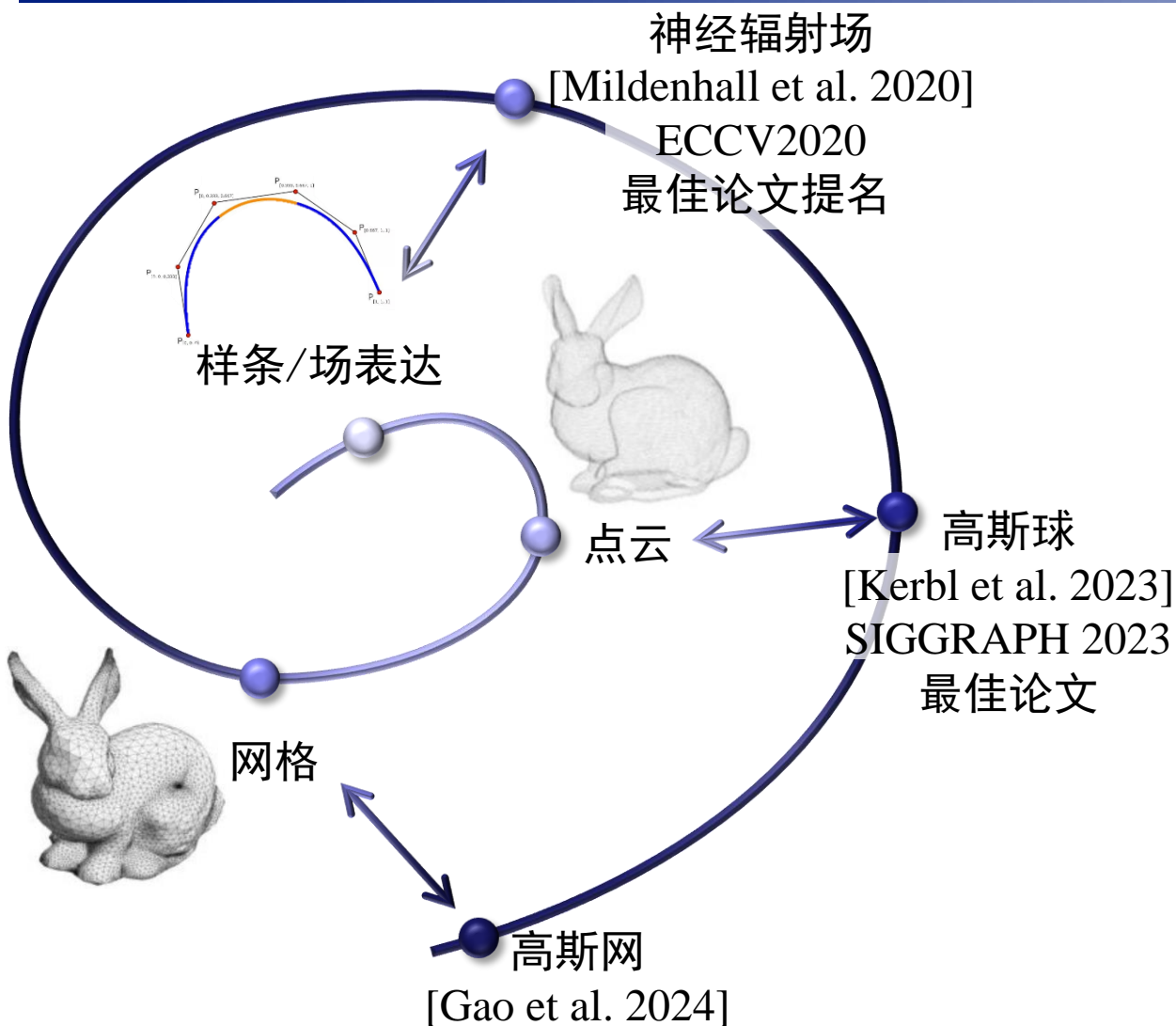
Relighting



混合表征在推动几何表征的发展



几何表示进展



R E A L

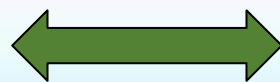
	Photo-Realistic 照片级真实	Efficient 效率	Scalable Large-Scale Deformation 大尺度形变	Low-Cost 低成本
神经辐射场	✓	✗	✗	✓
高斯球	✓	✓	✗	✓
高斯网	✓	✓	✓	✓

高斯泼溅几何、纹理和光照解耦编辑

难点

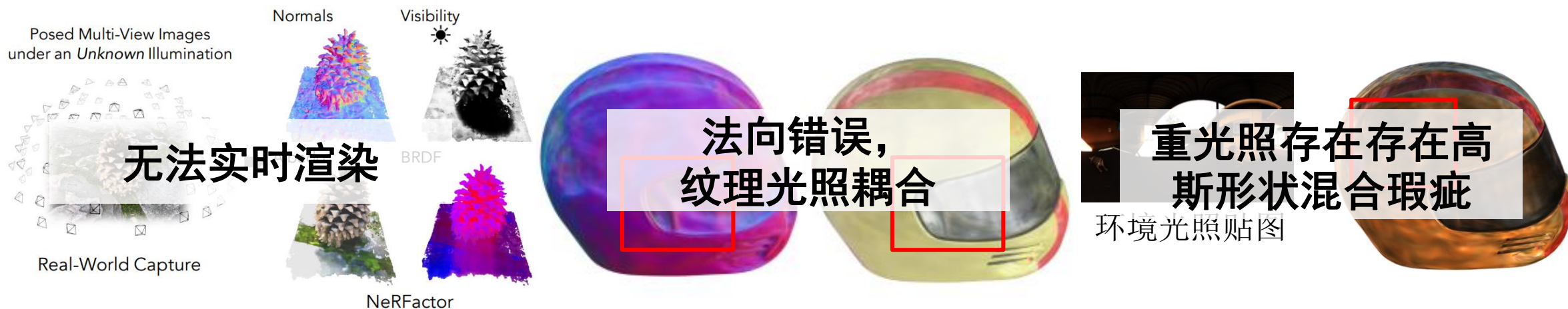
离散几何表示

矛盾



精确的几何、纹理和光照解耦

目标：从多张图像解耦几何、纹理和光照，并支持三者的独立编辑。



NeRF编辑方法[1]

3DGS解耦编辑方法[2]

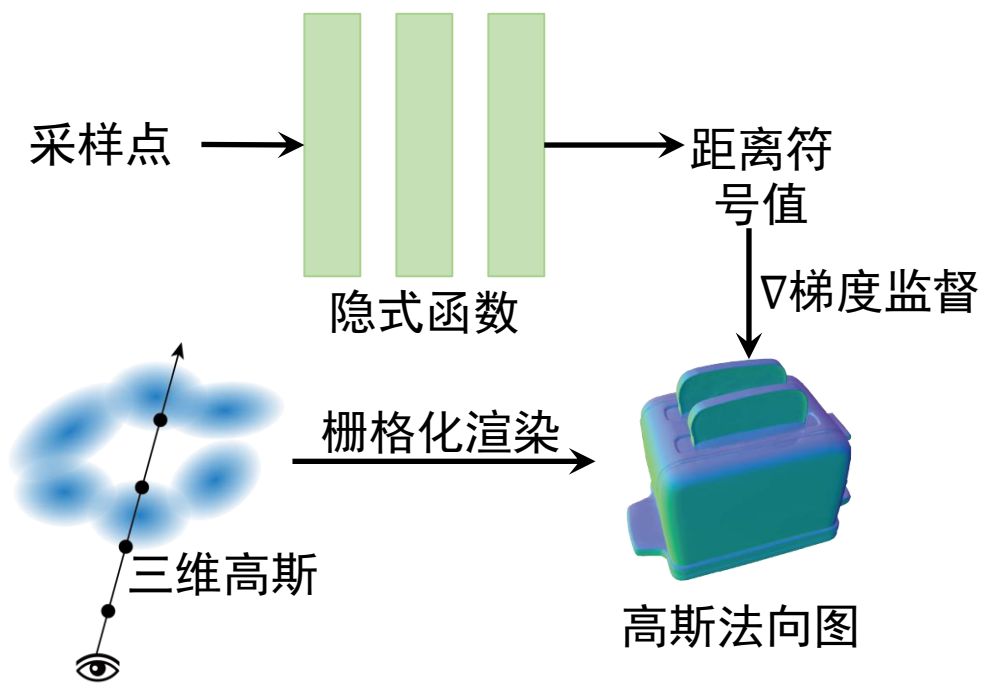
3DGS解耦编辑方法[3]

[1] NeRFactor: Neural Factorization of Shape and Reflectance Under an Unknown Illumination

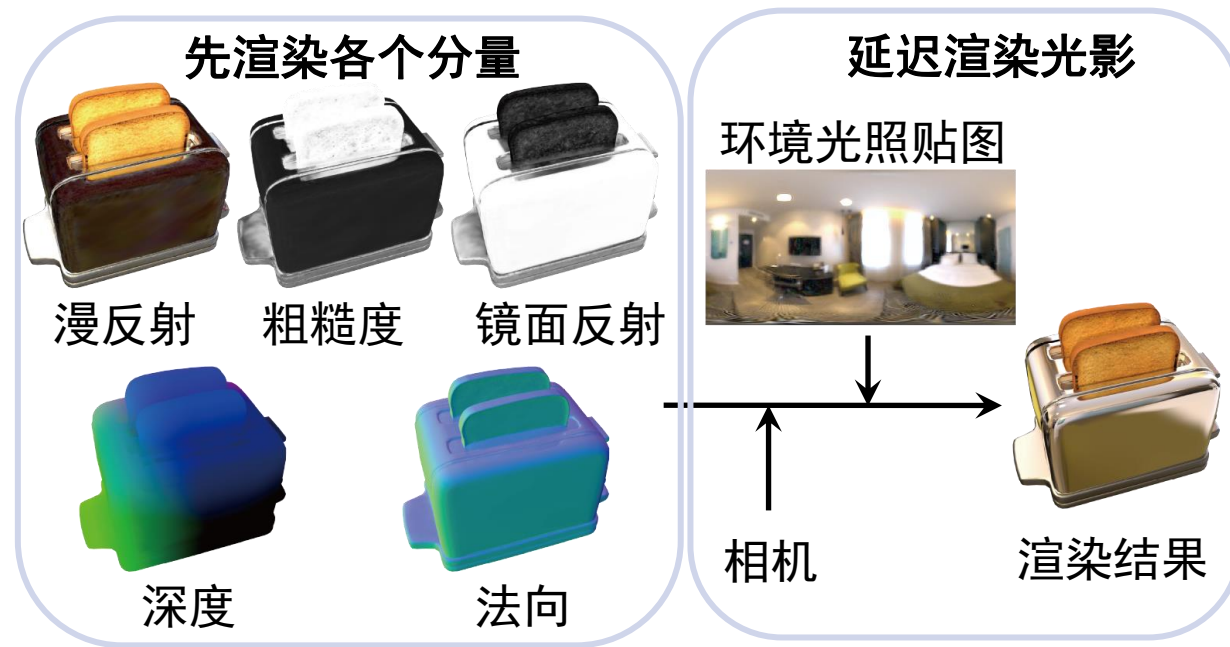
[2] Relightable 3D Gaussian: Real-time Point Cloud Relighting with BRDF Decomposition and Ray Tracing.

[3] GaussianShader: 3D Gaussian Splatting with Shading Functions for Reflective Surfaces.

- 法向场蒸馏：将隐式函数的法向蒸馏到高斯上以得到更好的几何。
- 延迟渲染：首先渲染各个几何、纹理属性图像，之后在像素级别进行光影计算，避免高斯混合带来的瑕疵。



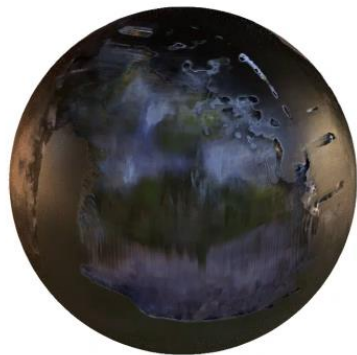
法相场蒸馏



延迟渲染



Input



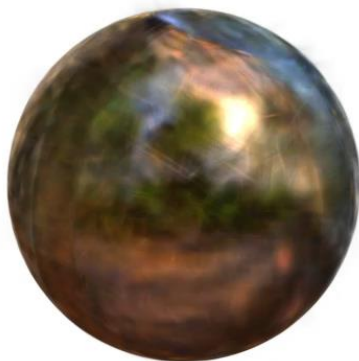
TensoIR



NDR



Ours



RelightableGaussian



GaussianShader



NeRO



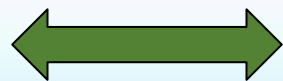
GT

高斯泼溅几何表征的挑战

难点

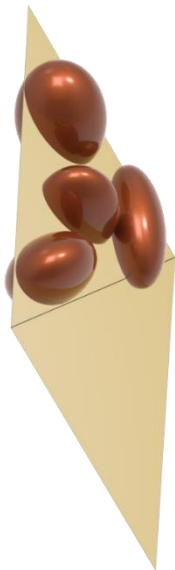
3DGS离散，无拓扑结构

矛盾

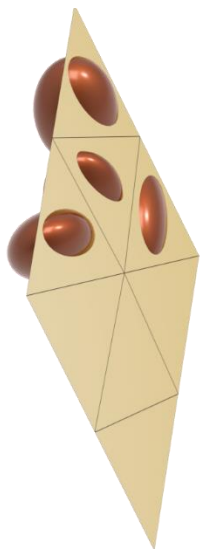


实时大尺度变形

目标：对3DGS进行实时高质量大尺度变形



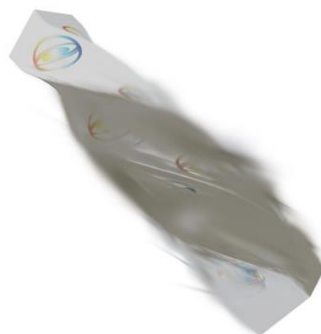
一般基于显式代理的方法
多个高斯对一个面片



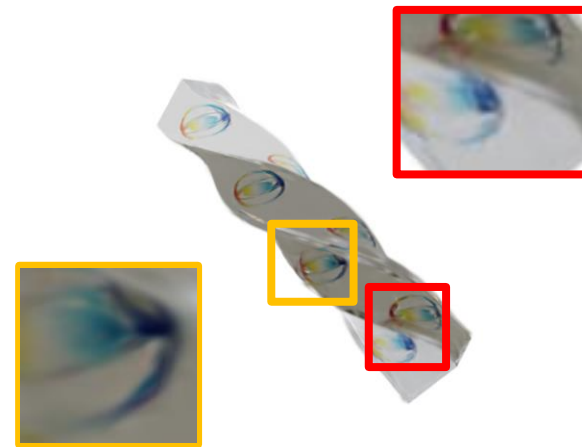
高斯网
一个高斯对一个面片



基于NeRF的方法
受限于NeRF渲染速度，
难以实时变形



基于3DGS的方法1
高斯表示直接与网格建立映射，
可能存在错误映射

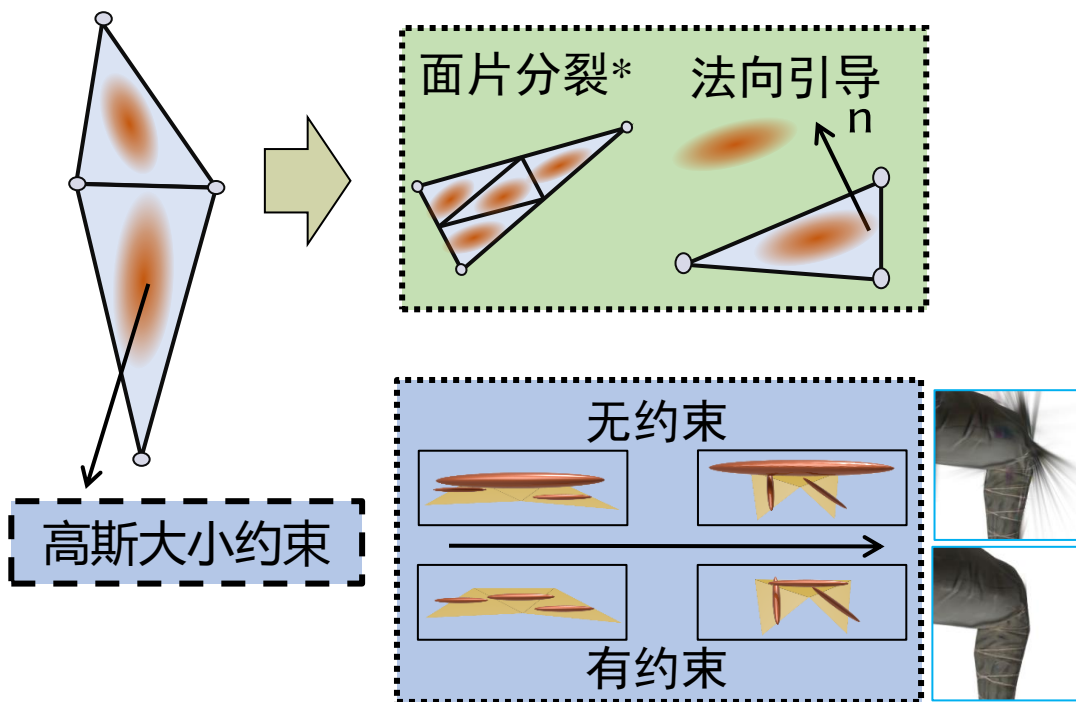


基于3DGS的方法2
缺少对高斯优化和分裂的约束，变
形产生瑕疵

■ 基于网格的高斯优化:

- ◆ 面片分裂: 高斯的渲染引导网格面分割以进行自适应细化, 并且网格面分割指导高斯球的分裂
- ◆ 法向导引: 沿面片法向导引优化高斯位置

高斯初始化



$$\mu = (w_a \mathbf{v}_a + w_b \mathbf{v}_b + w_c \mathbf{v}_c) + \tau R \mathbf{n},$$

重心坐标: $[w_a, w_b, w_c]$

沿法向导引: τR , R 是当前面片外接圆半径, \mathbf{n} 是面片单位法向导

$$L_r = \sum_{g \in G} \max(\max(s_i) - \gamma R_i, 0)$$

高斯大小约束:

γ : 超参

R_i : 面片外接圆半径

s_i : 高斯核的scale

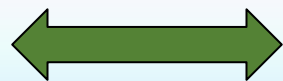


嵌套透明物体重建

难点

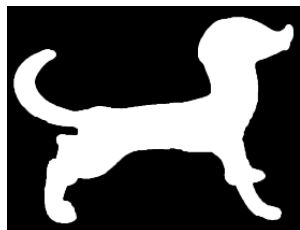
折射和全反射二义性

矛盾



高质量重建结果

目标：从多目图像重建带透明材质物体的几何模型



需要提供mask和光照情况



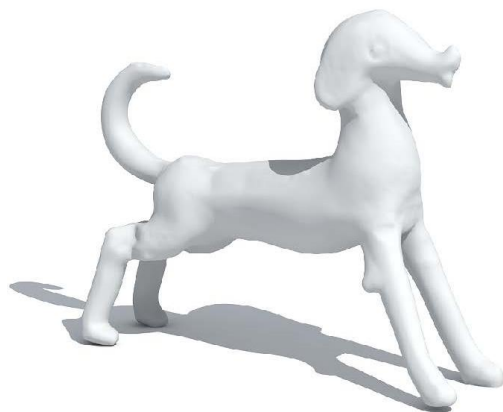
需要特殊的背景和捕捉设备



需要物体在完美的平面上



基于图像处理的方法 CVPR 2020



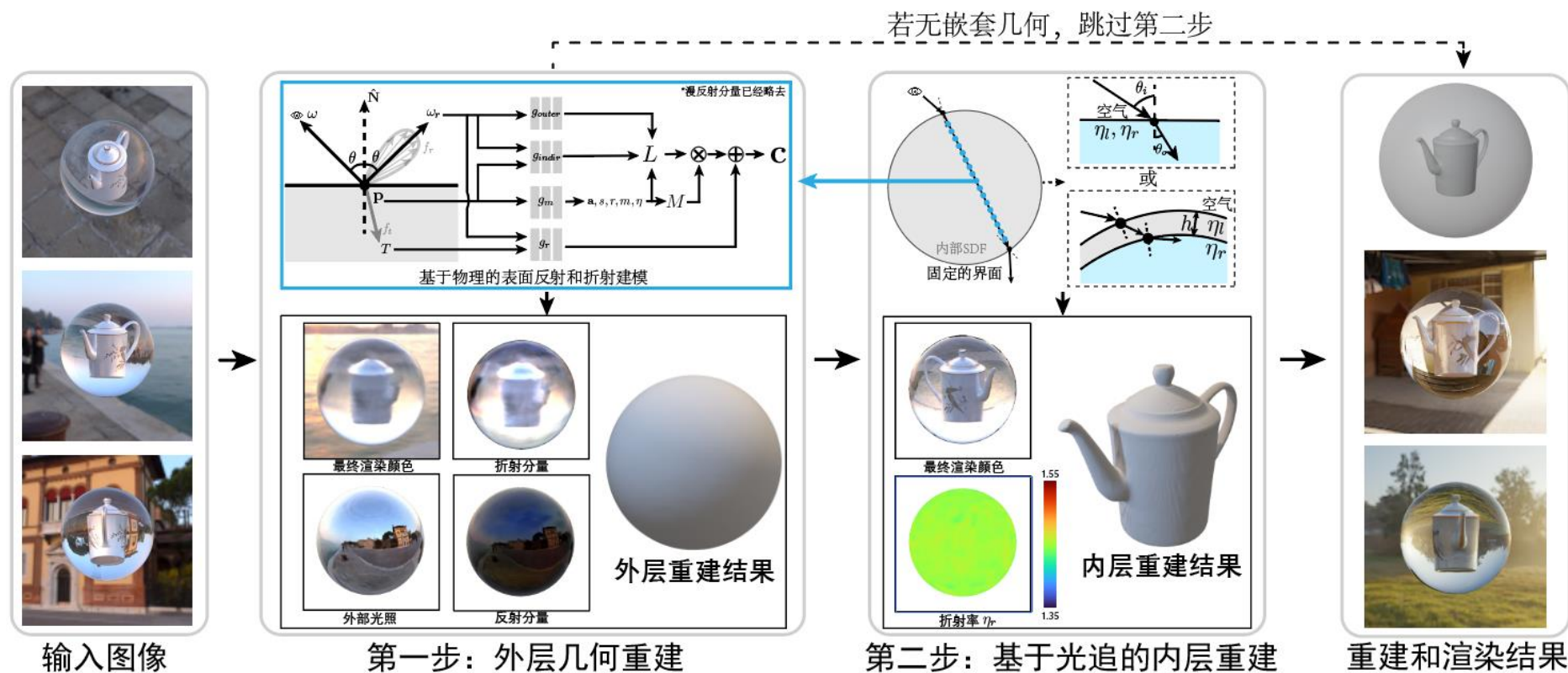
基于NeRF的方法1 ICCV 2023



基于NeRF的方法2 SIGGRAPH Asia 2023 28

- 不需要任何的额外输入和拍摄条件，仅需要图片和对应的相机位姿。

两阶段重建，引入光追，减少二义性



Scene: BallStatue

*Scene Captured **in the Wild** Using a Mobile Phone

■ 结果展示

Scene: GlassIce

玻璃杯，吸管和冰块

Scene: BallStatue

*Scene Captured **in the Wild** Using a Mobile Phone

雕像

Scene: Bottle

*Scene Captured **in the Wild** Using a Mobile Phone

装水的矿泉水瓶

团队合作老师和同学



杨洁



刘锋林



吴桐



孙嘉木



吕恒磊



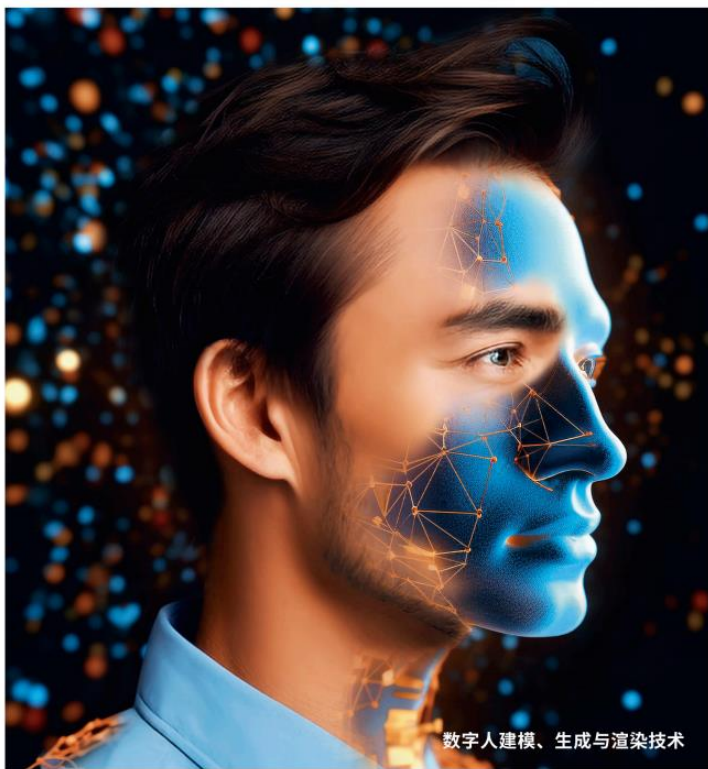
张博韬

“数字人建模、生成与渲染技术” 专栏



中国图象
图形学报

2024
09
VOL.29
ISSN1006-8961
CN11-3758 / TB



数字人建模、生成与渲染技术

2024年9月出版

6篇综述+2篇算法论文

➤ 专题编委

刘利刚 中国科学技术大学 (特邀主编)

高 林 中科院计算所

郭裕兰 中山大学

李 策 兰州理工大学

吕 琳 山东大学

王贝贝 南京大学

张举勇 中国科学技术大学

➤ 发表论文

1. 三维人脸成像及重建技术综述

刘菲,张堃博,杨青,周树波,王云龙,孙哲南

2. 数字室内三维场景构建综述

岳亮,谈皓,黄俊凯,张少魁

3. 多模态数字人建模、合成与驱动综述

高玄,刘东宇,张举勇

扫码阅读

专栏论文



4. 数字人脸渲染与外观恢复方法综述

郝琮晖,杜悠扬,王璐,王贝贝

5. 多模态信息引导的三维数字人运动生成综述

赵宝全,付一愉,苏卓,王若梅,吕辰雷,罗笑南

6. 三维穿衣人体重建综述——

从传统方法到高保真模型

陈鸿鹄,陶云帆,张举勇

7. 室内稀疏全景图的神经辐射场重建

肖强,陈铭林,张晔,黄小红

8. 单视角三维人体重建的着装特征学习

黄千芄,刘骊,付晓东,刘利军,彭玮

谢谢!

