

# 基于多视点图像的三维重建

刘利刚

中国科学技术大学

# 计算机图形学 (CG) 的两大应用方向

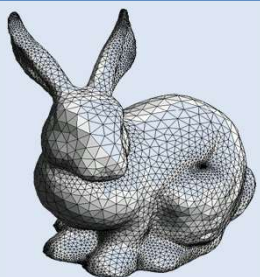
几何表达

仿真计算

输出结果

应用方向

3D数字模型



三角网格  
神经辐射场

物理仿真和动画



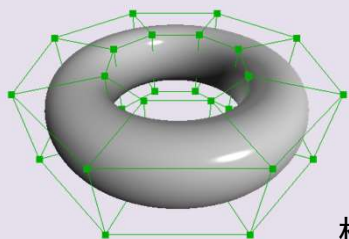
高效率

真实感渲染：视频



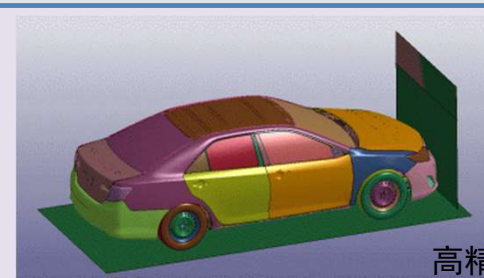
影视  
娱乐  
游戏  
元宇宙  
虚拟现实

计算机辅助设计 (CAD)



样条曲面

计算机辅助分析 (CAE)



高精度

制造 (CAM)：实物



工业4.0  
智能制造  
数字孪生

# 什么是“渲染”？

渲染（广义）：计算生成从某个视点观察场景的图像



如果场景是存在的，最简单的方式：拍照（直接采集）

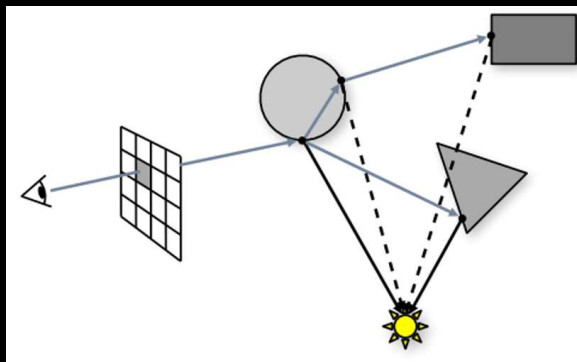
# CG渲染：图像生成的第一性原理

基于几何的场景**表达**



面表达：网格曲面、  
纹理、材质、光源

基于物理的光照**计算**



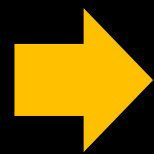
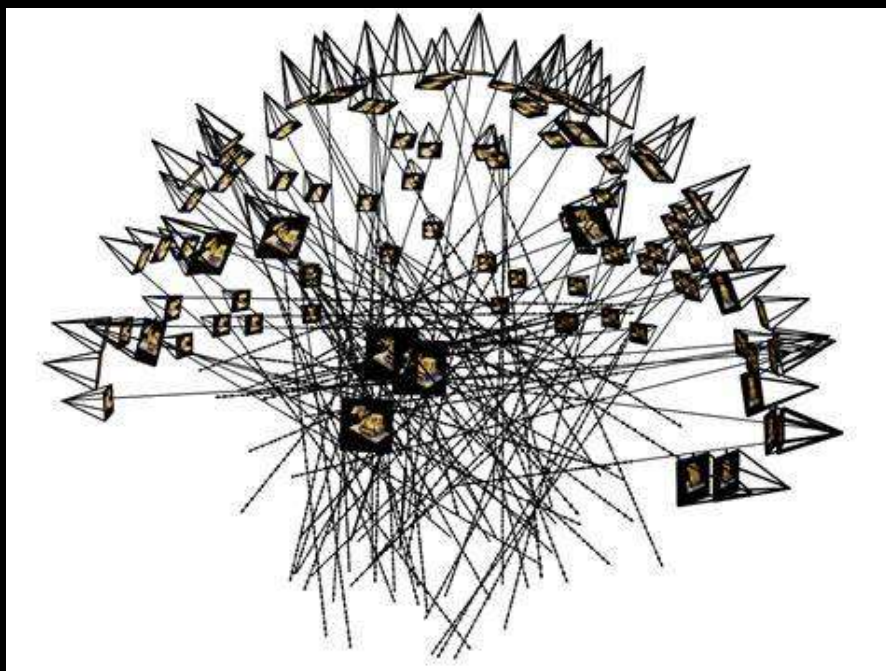
光线追踪算法  
求解渲染方程

$$L_o(p, \omega_o) = L_e(p, \omega_o) + \int_{\Omega^+} L_i(p, \omega_i) f_r(p, \omega_i, \omega_o) (n \cdot \omega_i) d\omega_i$$



两个关键点

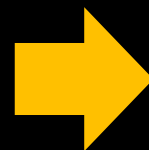
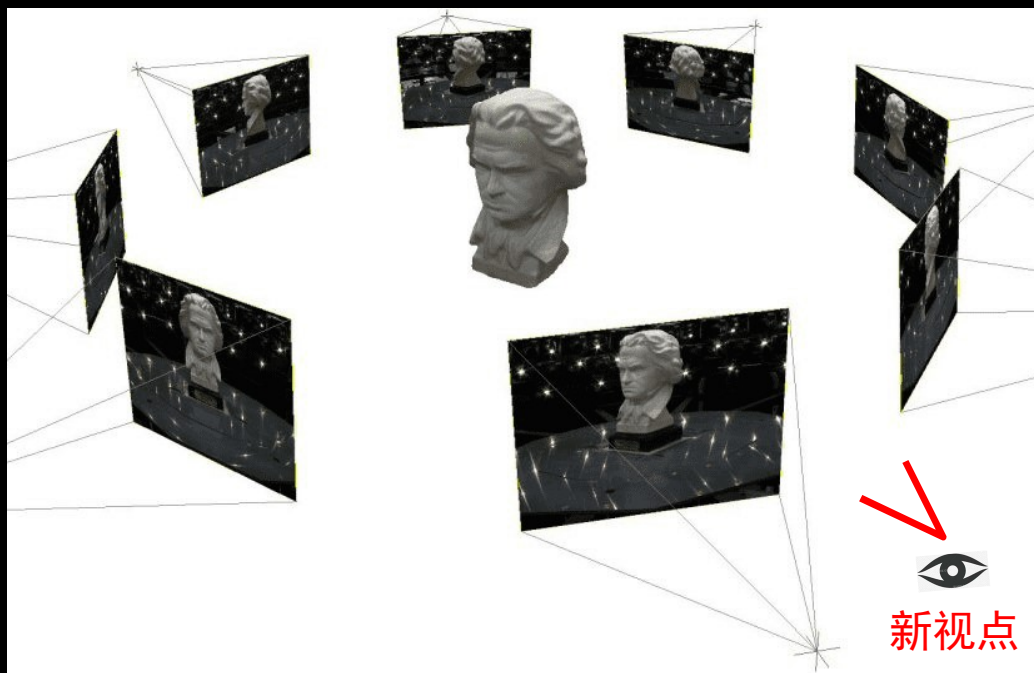
# 反问题：从**观察**获得**未知对象**的认知、处理



未知对象的  
认知及处理

例子：测量、3D重建、CT、地质勘探…

问题：由已知场景多视点图像 → 新视点图像

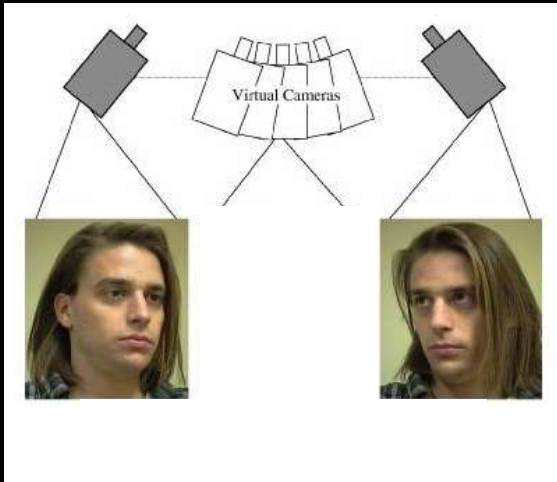


新视点图像  
(合成/渲染)

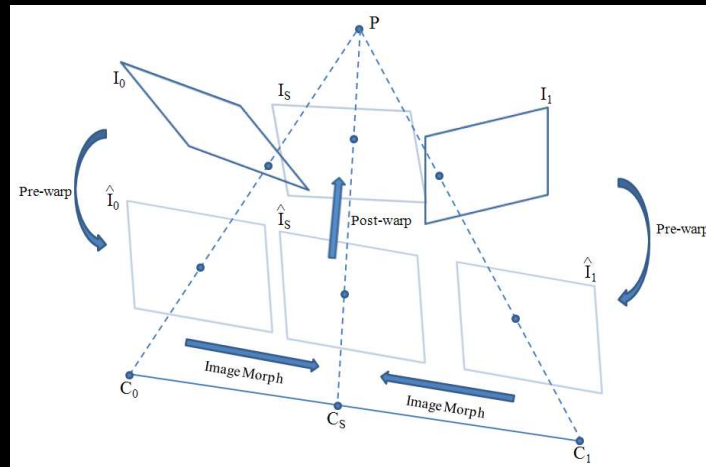
Novel view synthesis/  
Novel view rendering

# 基于图像的渲染

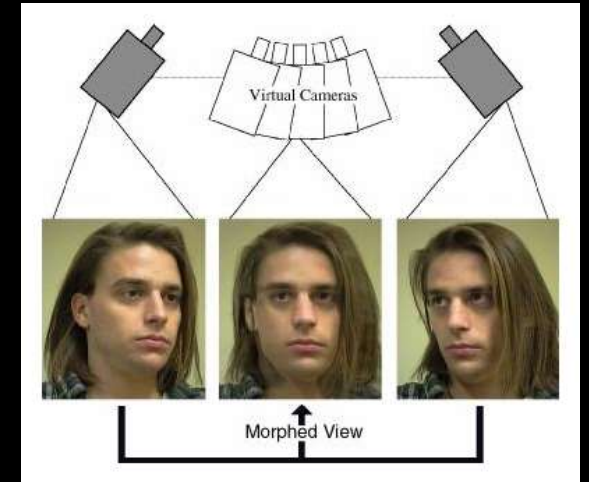
Image based rendering (空间视频) : View interpolation, view morphing, MPI...



Multi-view images



Warping + blending



无显式的场景表达

# 渲染：由“场景表达”到“图像像素”的函数

$X$  : 场景表达

- 确定性表达
  - 显式表达 或 隐式表达
  - 面表达 或 体表达
  - ...

$F$  : 渲染计算

- 确定性计算
  - 面渲染 (光线追踪)
  - 体渲染
  - ...

$Y$  : 图像或像素

$$Y = F(X)$$

两个关键点

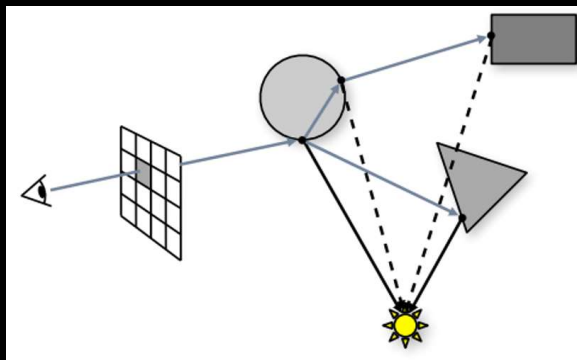
# 渲染-1: 场景表达到图像的函数

X: 场景表达



面表达: 几何 (网格)、  
纹理、材质、光源

F: 渲染计算



光线追踪算法

$$L_o(p, \omega_o) = L_e(p, \omega_o) + \int_{\Omega^+} L_i(p, \omega_i) f_r(p, \omega_i, \omega_o) (n \cdot \omega_i) d\omega_i$$

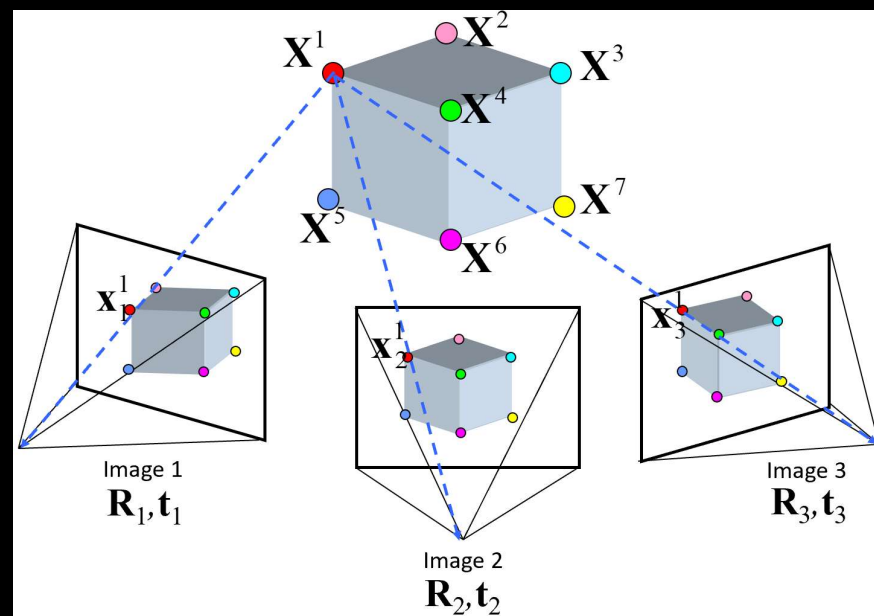
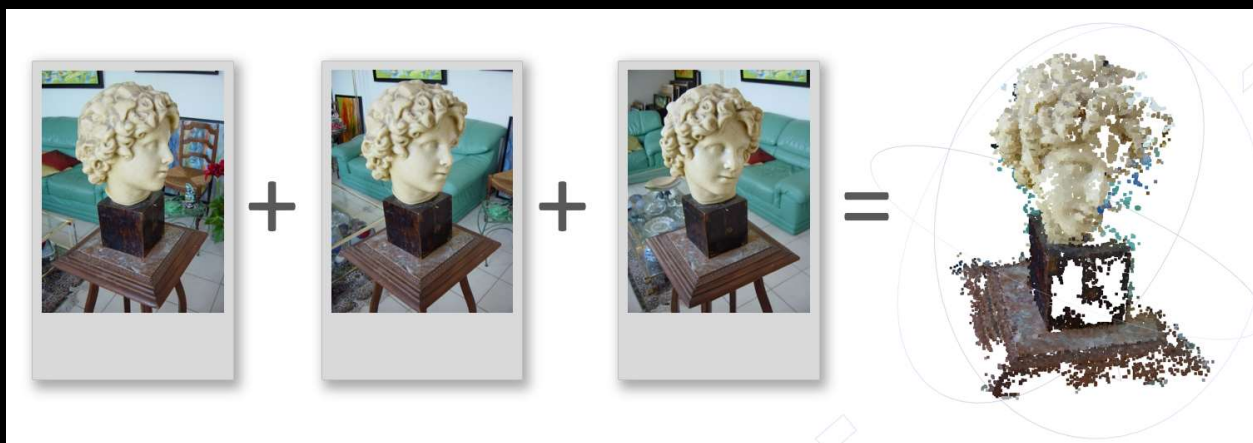
Y: 图像或像素



方式1: 面表达 + 面渲染  
像素颜色 = 采样的折线形光路上所有点颜色的加权求和 (积分)

# 由多视点图像重建3D物体（面表达）

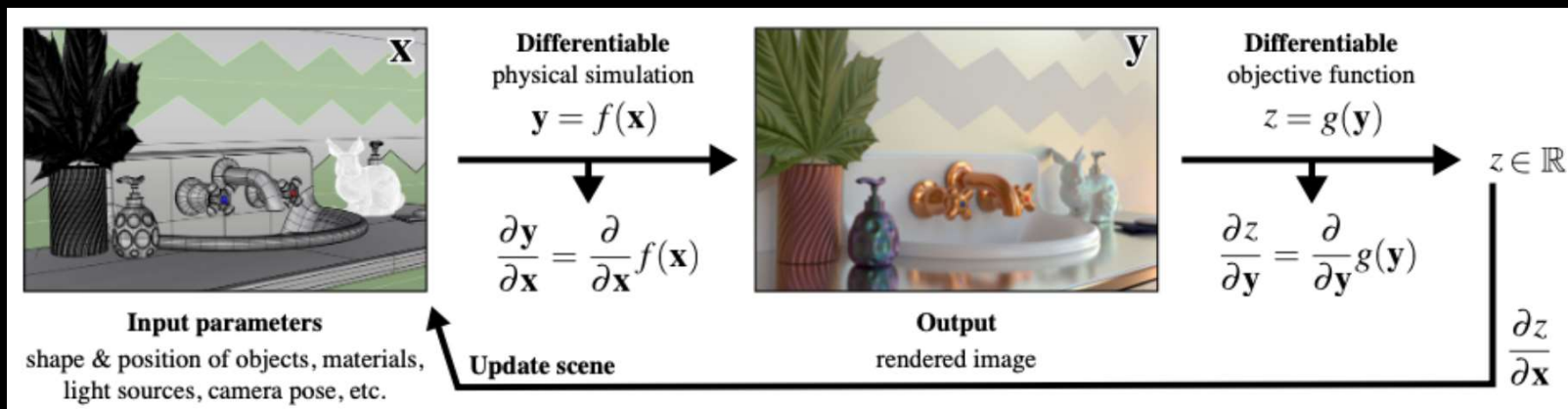
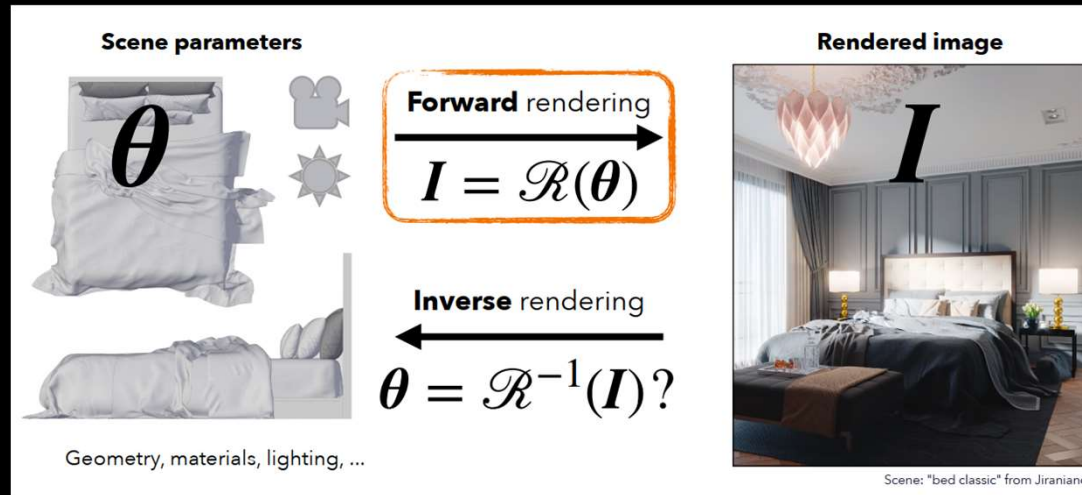
- Multi-view geometry
  - Feature correspondences
  - Multi-view Stereo (MVS)
  - Structure from motion (SfM)
- 问题
  - 特征点对应，多步求解（非端到端），优化技巧



	Point 1	Point 2	Point 3
Image 1	$x_1^1 = \mathbf{K}[\mathbf{R}_1   \mathbf{t}_1] \mathbf{X}^1$	$x_1^2 = \mathbf{K}[\mathbf{R}_1   \mathbf{t}_1] \mathbf{X}^2$	
Image 2	$x_2^1 = \mathbf{K}[\mathbf{R}_2   \mathbf{t}_2] \mathbf{X}^1$	$x_2^2 = \mathbf{K}[\mathbf{R}_2   \mathbf{t}_2] \mathbf{X}^2$	$x_2^3 = \mathbf{K}[\mathbf{R}_2   \mathbf{t}_2] \mathbf{X}^3$
Image 3	$x_3^1 = \mathbf{K}[\mathbf{R}_3   \mathbf{t}_3] \mathbf{X}^1$		$x_3^3 = \mathbf{K}[\mathbf{R}_3   \mathbf{t}_3] \mathbf{X}^3$

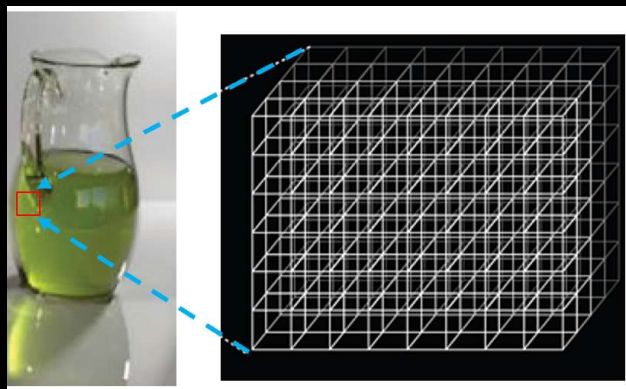
# 由多视点图像重建3D物体（面表达）

- 可微渲染（Differential rendering）



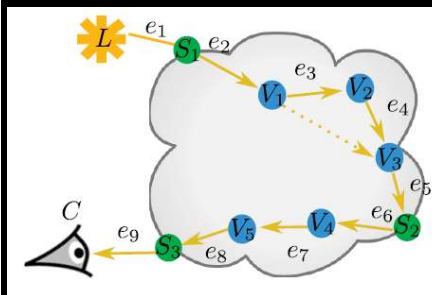
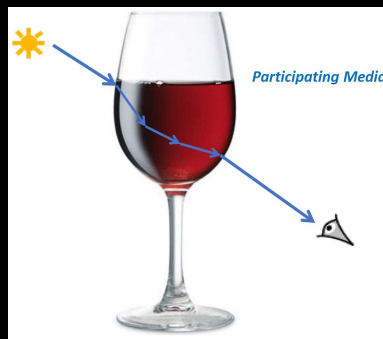
# 渲染-2: 场景表达到图像的函数

$X$ : 场景表达



体表达 (体素): 颜色、材质  
(透明度、反射率、折射率)

$F$ : 渲染计算



体渲染  
求解体渲染方程

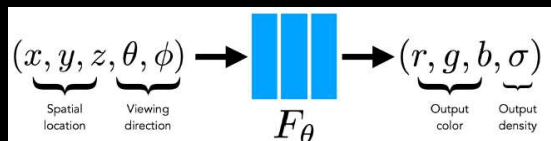
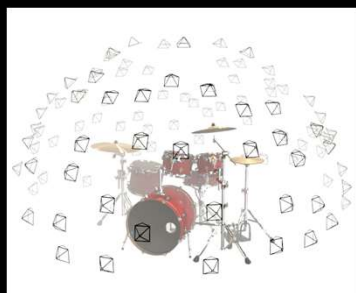
$Y$ : 图像或像素



方式2: 体表达 + 体渲染  
像素颜色 = 采样的折线形光路上所有点颜色的加权求和 (积分)

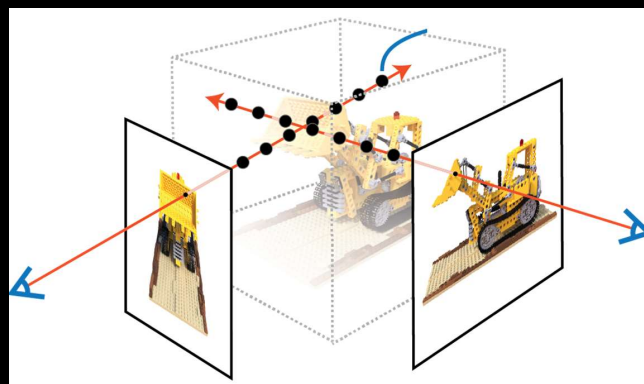
# 渲染-3：场景表达到图像的函数

$X$ ：场景表达



体表达（体素）：颜色、  
简化的材质（透明度）

$F$ ：渲染计算



简化的体渲染

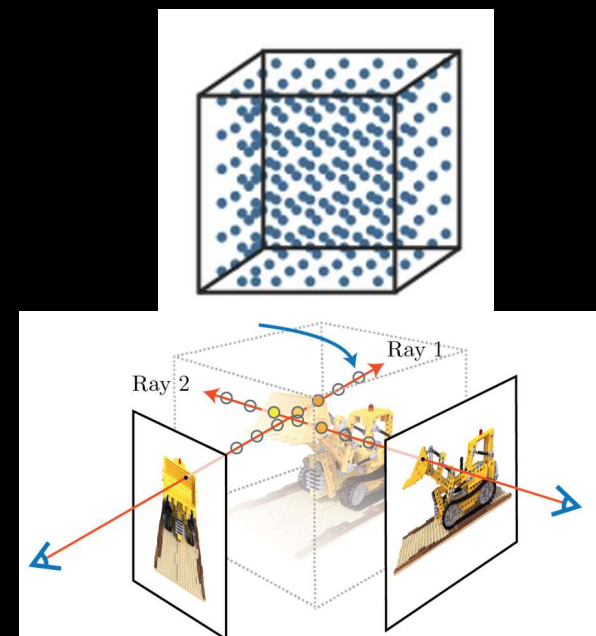
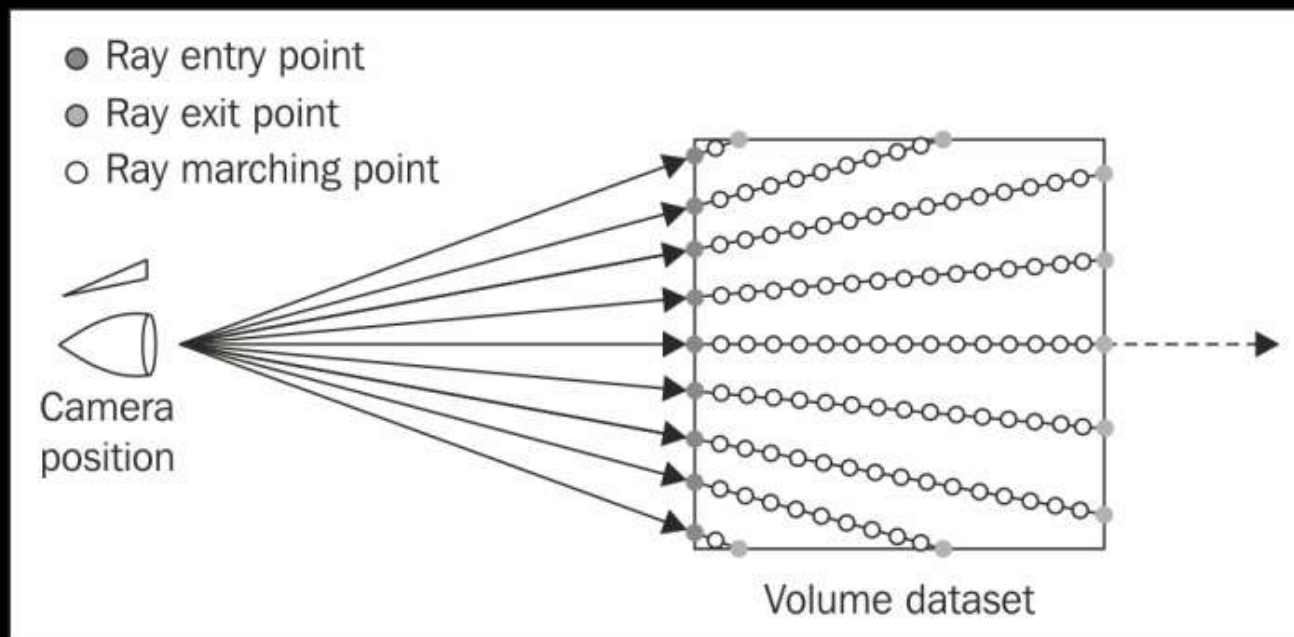
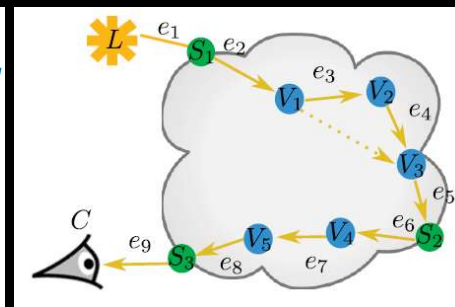
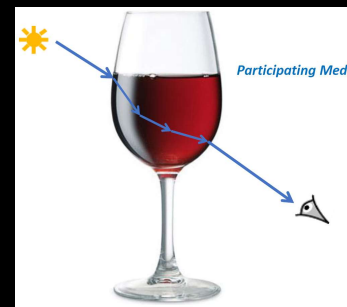
$Y$ ：图像或像素



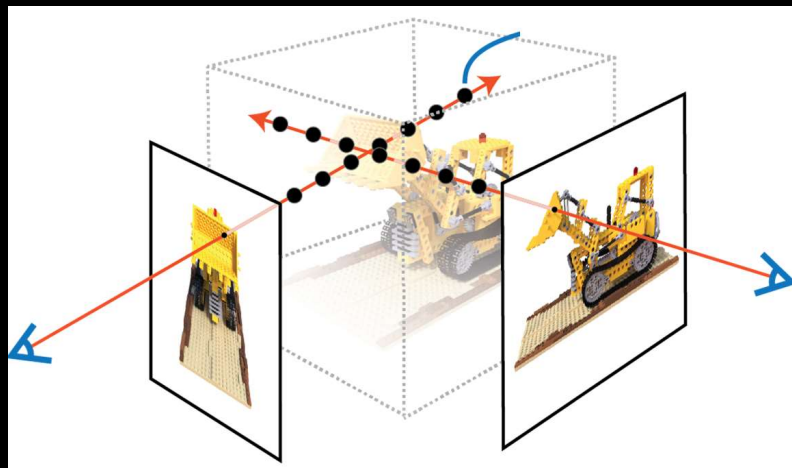
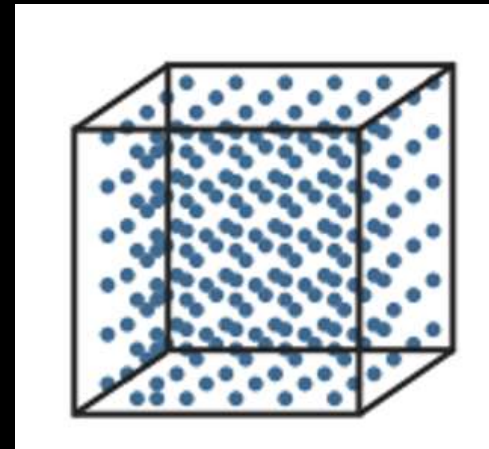
方式3：体表达 + 直线体渲染  
像素颜色 = 直线光路上所有点颜色的加权求和（积分）

# 直线体渲染：直线上的体素颜色的加权求和

- 基于物理的体渲染：光路复杂，难以可微求解
- 简化！！
  - 将光路简化：直线光路
  - 将材质简化：透明度（无反射、散射、折射等）



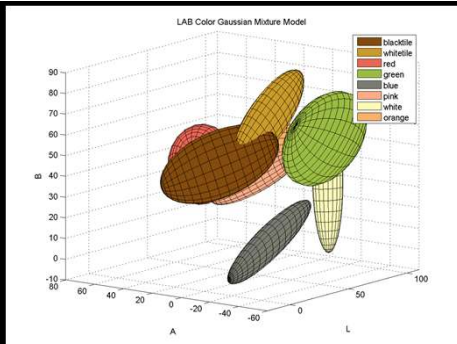
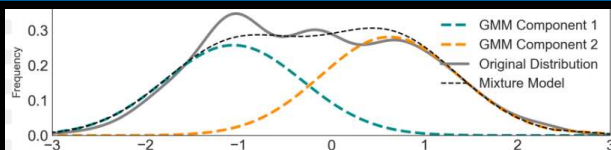
# 反问题：由多视点图像求体表达



$$\min_{\Omega} \sum_i \| \text{render}^{(i)}(F_{\Omega}) - I_{\text{gt}}^{(i)} \|^2$$

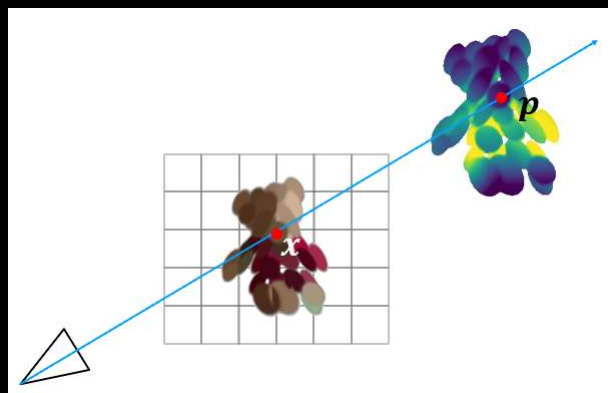
# 渲染-4: 场景表达到图像的函数

$X$ : 场景表达



体表达 (3DGS):  
高斯型的RBF函数逼近

$F$ : 渲染计算



简化的体渲染

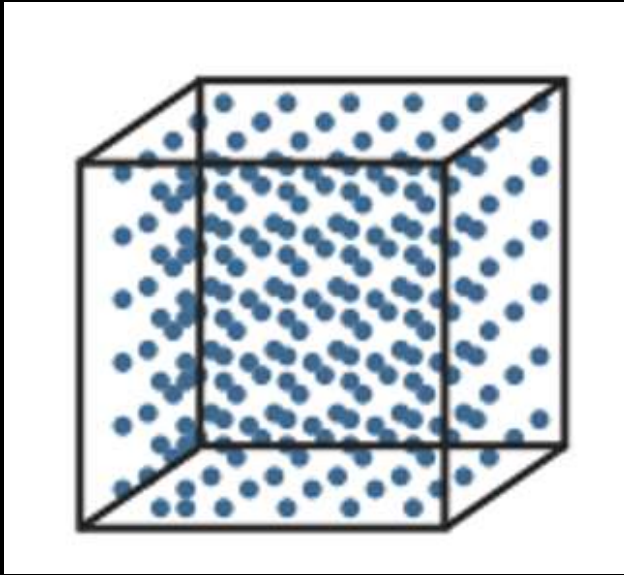
$Y$ : 图像或像素



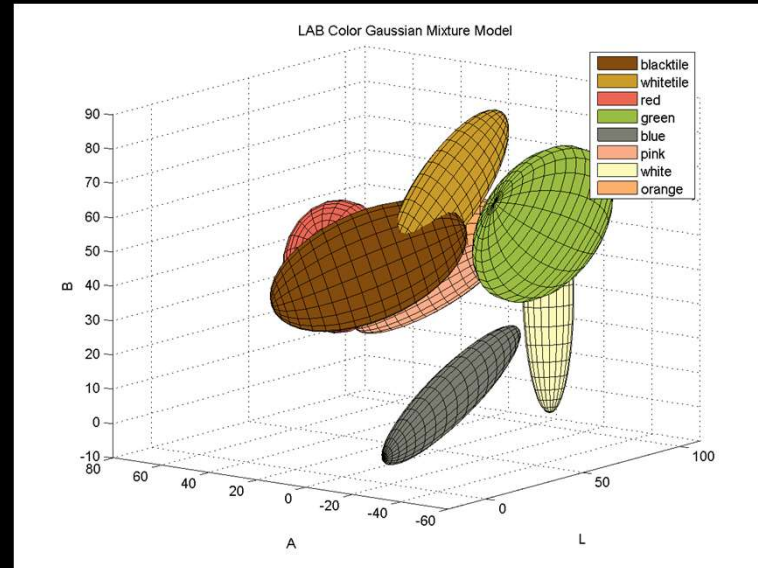
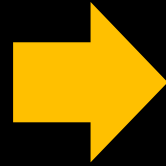
方式4: 3DGS体表达 + 直线体渲染

像素颜色 = 直线光路上所有点颜色的加权求和 (光追), 即GS函数投影的加权求和 (光栅)

# 使用3D Gaussians函数逼近3D场函数



3D场函数



3D Gaussians (RBF)

# 面表达 vs. 体表达

表达	渲染计算	重建方法	重建难度	重建速度	渲染速度	渲染效果	解耦性	优势	劣势
面表达 (几何、材质、光照)	光追	MVS 端到端 不可微 (光路复杂)	难 分多步, 每步都 有损失 不可微	离线 不实时	慢	真实感弱	场景解耦为: 几何、材质、光照	面向编辑, 可控性强	制作难
体表达 (NeRF/ 3DGS)	光栅	端到端 可微 (光路简单-直线)	易 端到端, 极小图 像损失	离线 不实时	快	真实感强	只是记录了 辐射场: 同时耦合了 几何、材质、光照 (烘培)	面向渲染, 渲染快	编辑难

# 渲染：由“场景表达”到“图像像素”的函数

$X$  : 场景表达

- 确定性表达
  - 显式表达 或 隐式表达
  - 面表达 或 体表达
  - ...
- 模糊性表达
  - 文本
  - 图像
  - 草图
  - ...

$F$  : 渲染计算

- 确定性计算
  - 面渲染 (光线追踪)
  - 体渲染
  - ...
- 概率性计算 (生成模型)
  - GAN/VAE
  - Diffusion
  - Auto regression
  - ...

$Y$  : 图像或像素

$$Y = F(X)$$

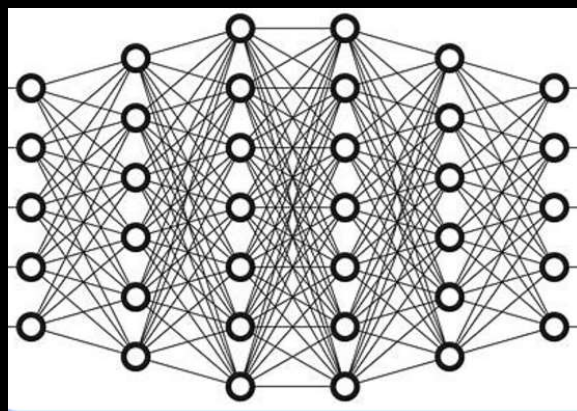
两个关键点

# 渲染-5：场景表达到图像的函数

$X$  : 场景表达

$F$  : 渲染计算

$Y$  : 图像



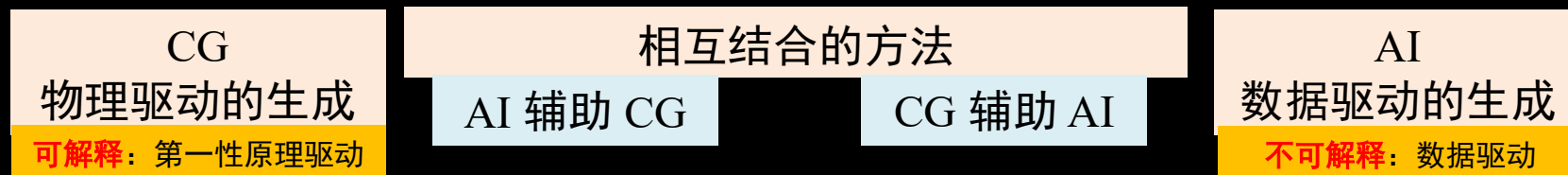
多模态输入：  
文本、图像、草图、...

生成模型

方式5：数据驱动的方法

图像 = 由输入根据一定概率由生成模型得到（具有不确定性）

# 生成图像/视频的不同方法

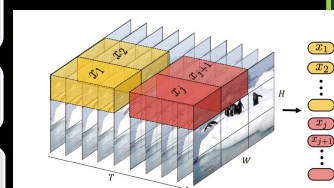


- 几何
- 纹理
- 材质
- 光照



**混合表达**  
NeRF, 3DGS, MLP...

- Tokens
- 隐空间特征
- 多模态



CG

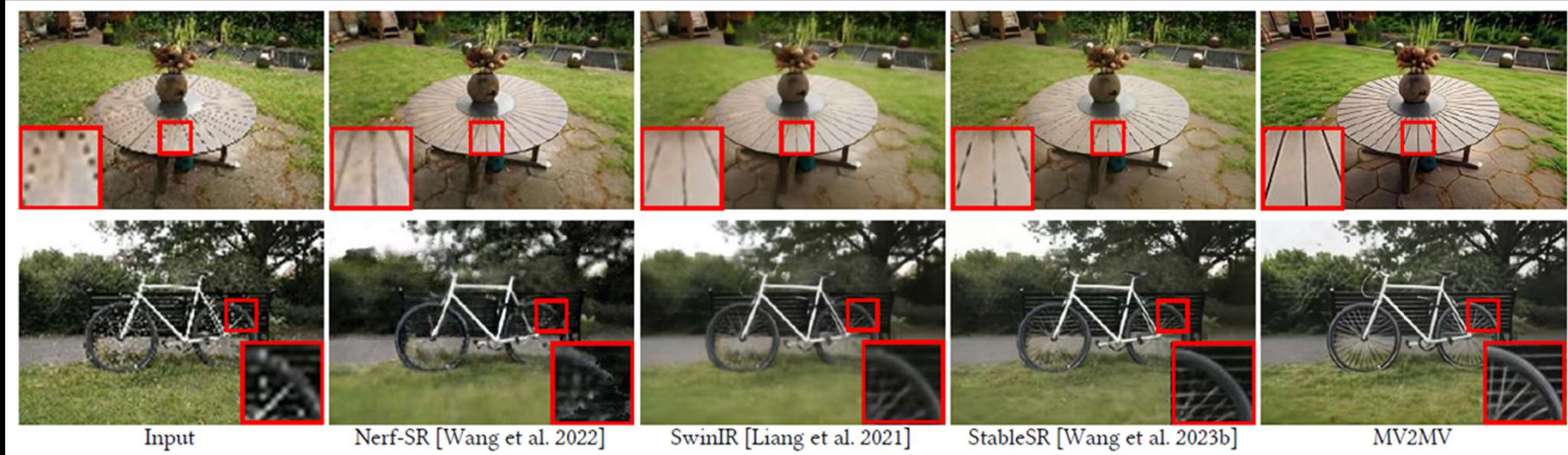


Sora

我们近期在3D重建/生成的最新工作

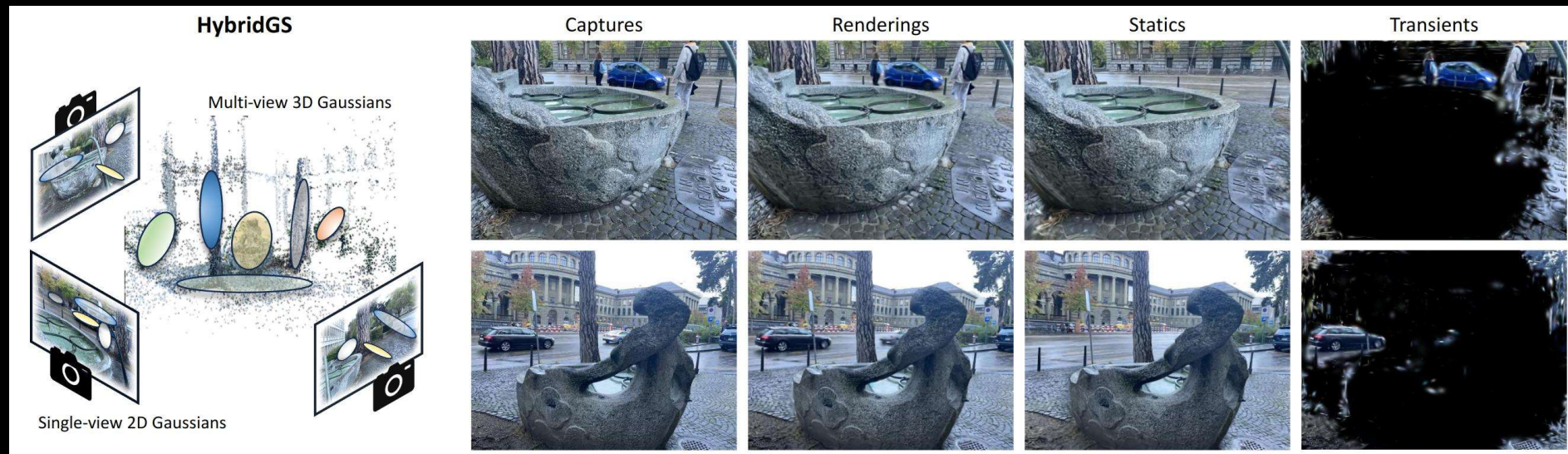
# MV2MV: Multi-View Image Translation via View-Consistent Diffusion Models

[Siggraph Asia 2024]



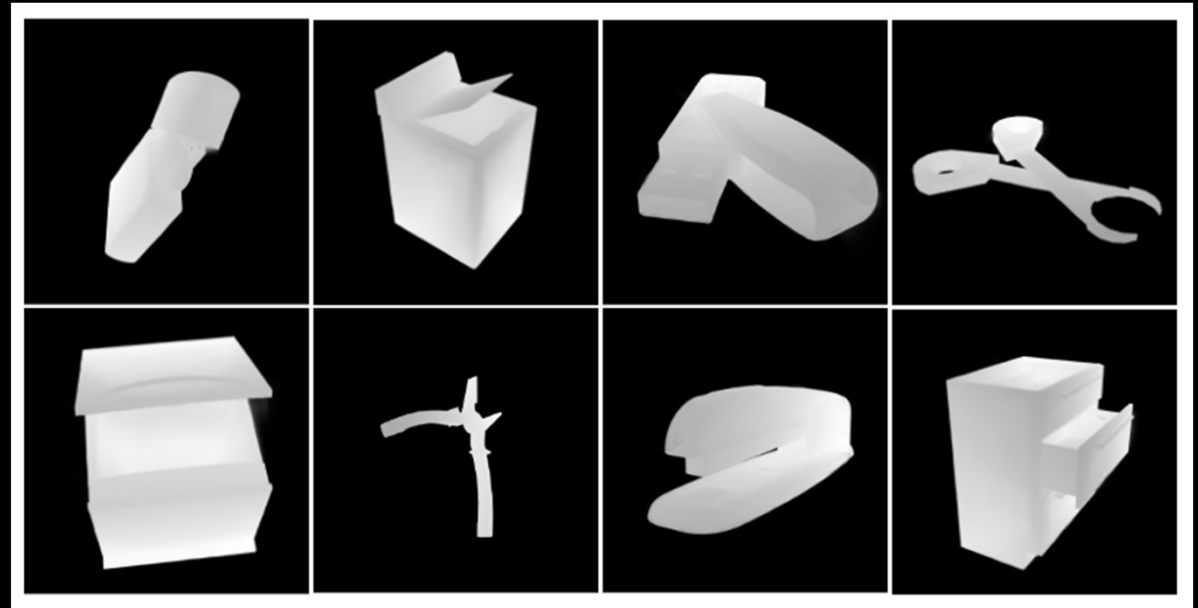
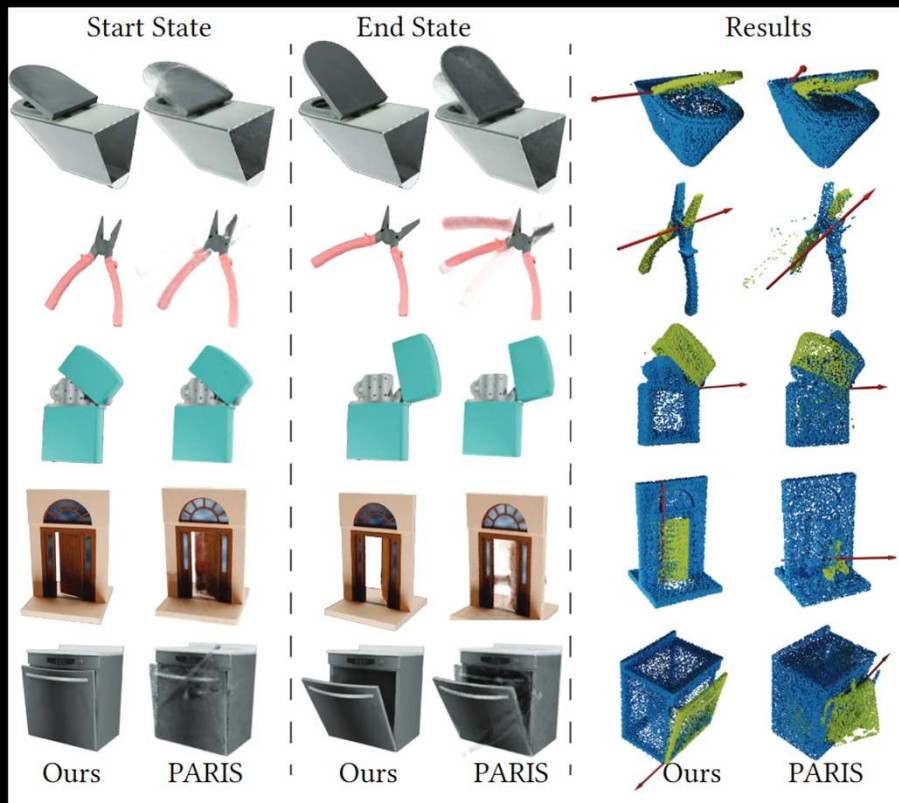
# HybridGS: Decoupling Transients and Statics with 2D and 3D Gaussian Splatting

[CVPR 2025]



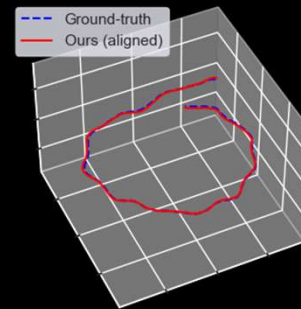
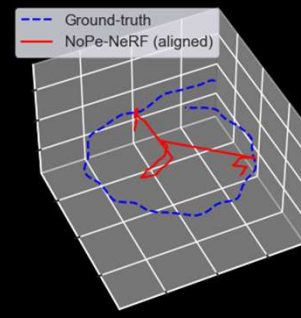
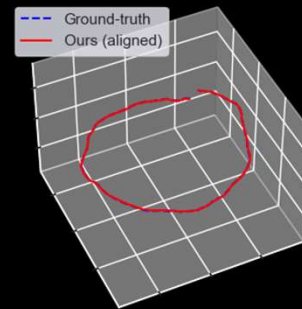
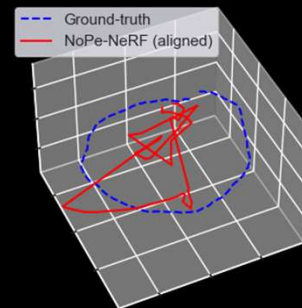
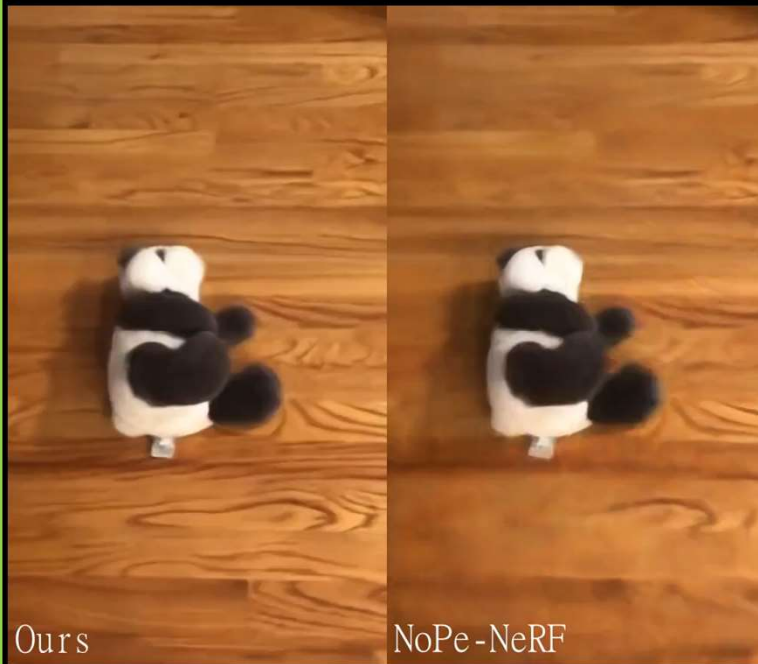
# ArticulatedGS: Self-supervised Digital Twin Modeling of Articulated Objects using 3D Gaussian Splatting

[CVPR 2025]



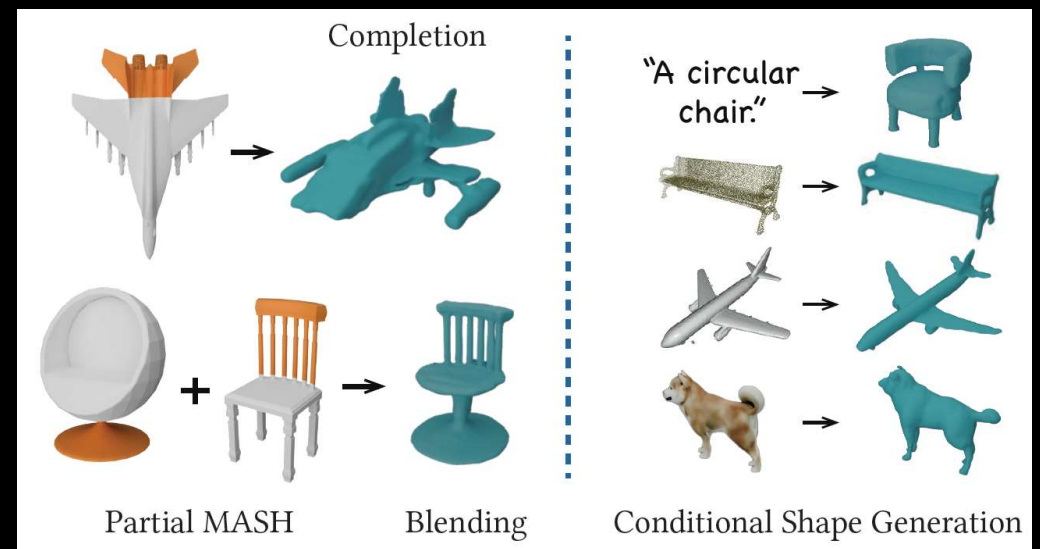
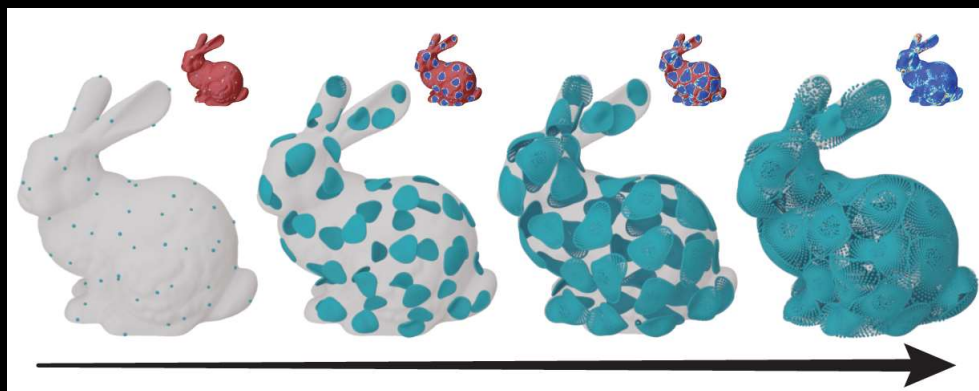
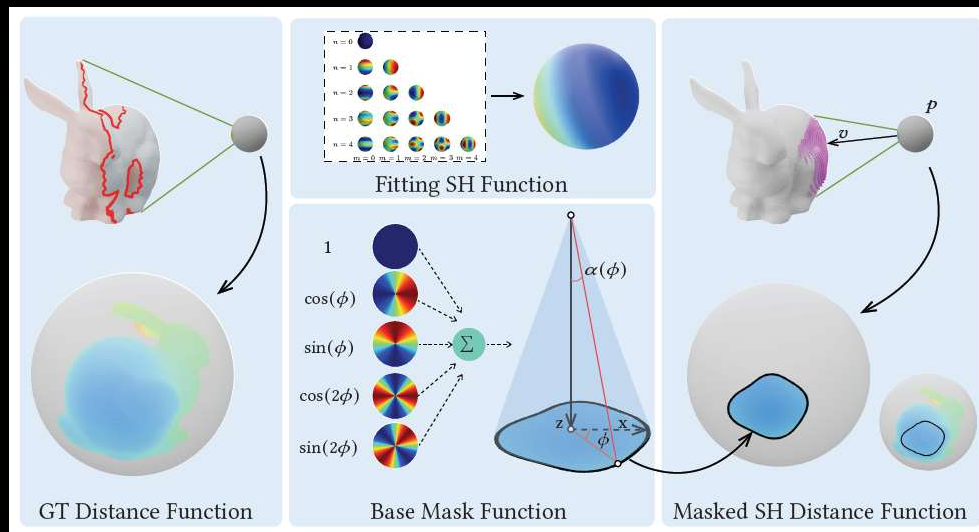
# NoPe-NeRF++: Local-to-Global Optimization of NeRF with No Pose Prior

[Eurographics 2025]



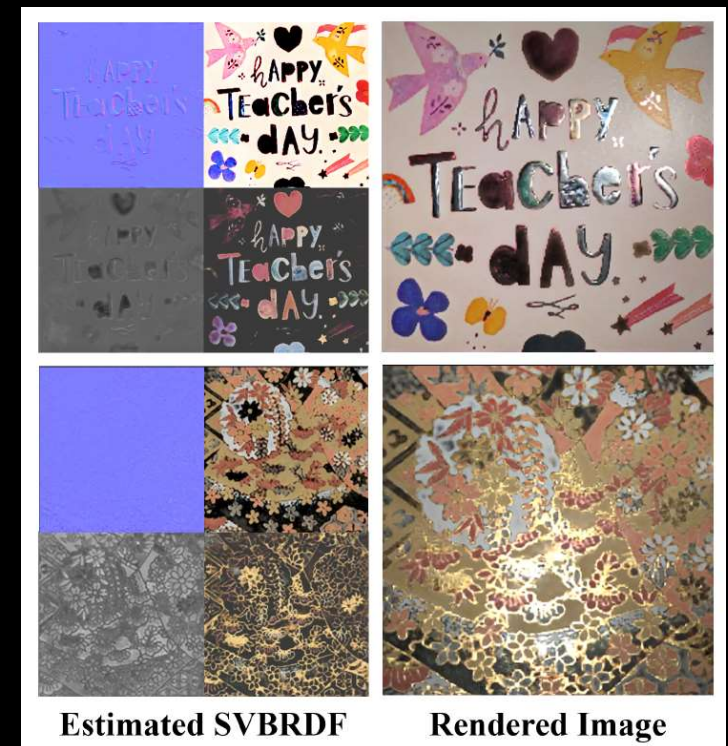
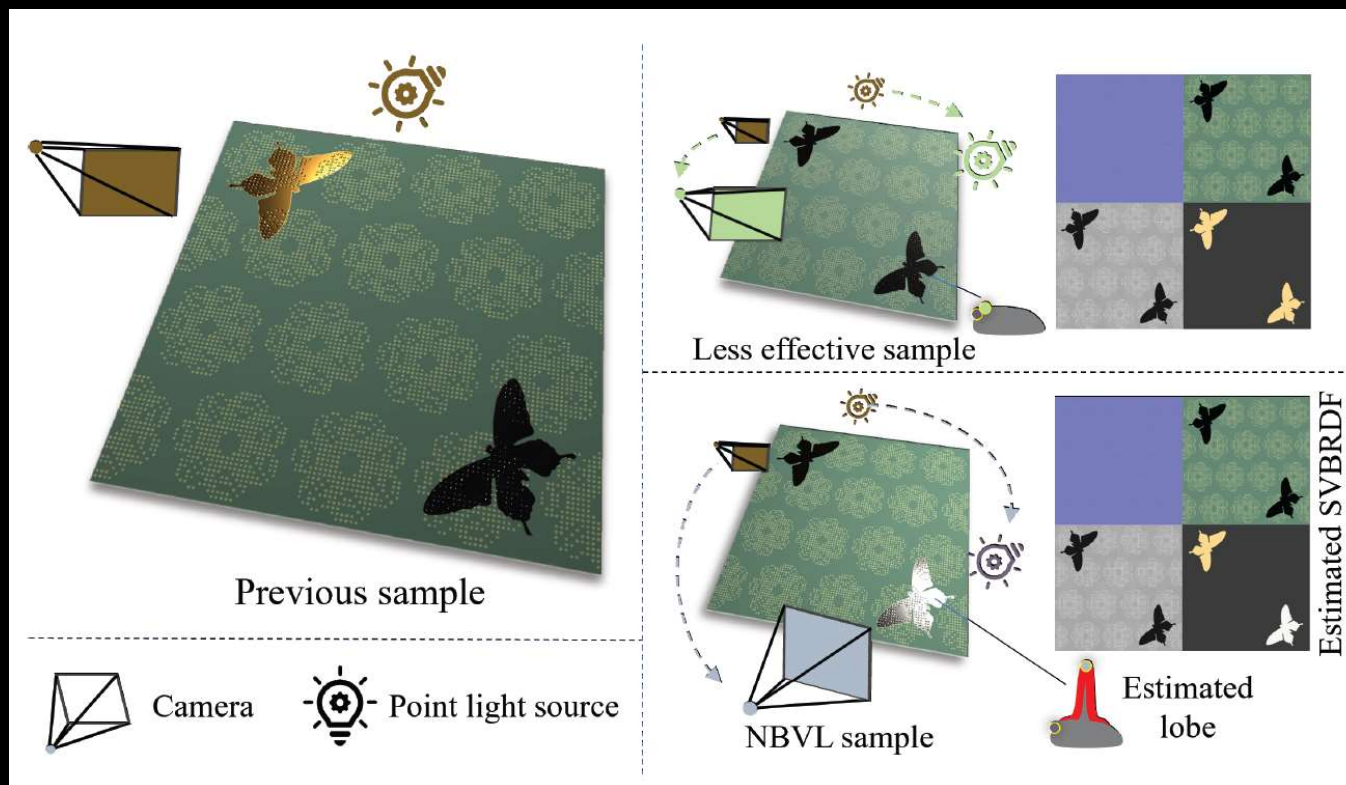
# MASH: Masked Anchored SpHERical Distances for 3D Shape Representation and Generation

[Siggraph 2025]



# Appearance-aware Multi-view SVBRDF Reconstruction via Deep Reinforcement Learning

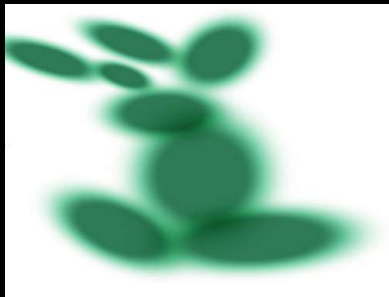
[Siggraph 2025]



# PDE-based 3DGS Reconstruction

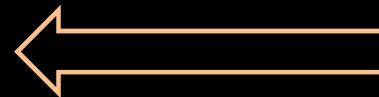
[Ongoing]

$$\left( s_i \frac{\partial L}{\partial \mu_i} \right) \sim \frac{\partial L}{\partial c_i} \sim \frac{\partial L}{\partial o_i} \sim \frac{\partial L}{\partial s_i} \geq \frac{\partial L}{\partial \theta(r_i)} \quad \Rightarrow \quad \left| \frac{\partial L}{\partial \mu_i} \right| \gg \left| \frac{\partial L}{\partial c_i} \right| \sim \left| \frac{\partial L}{\partial o_i} \right| \sim \left| \frac{\partial L}{\partial s_i} \right|$$

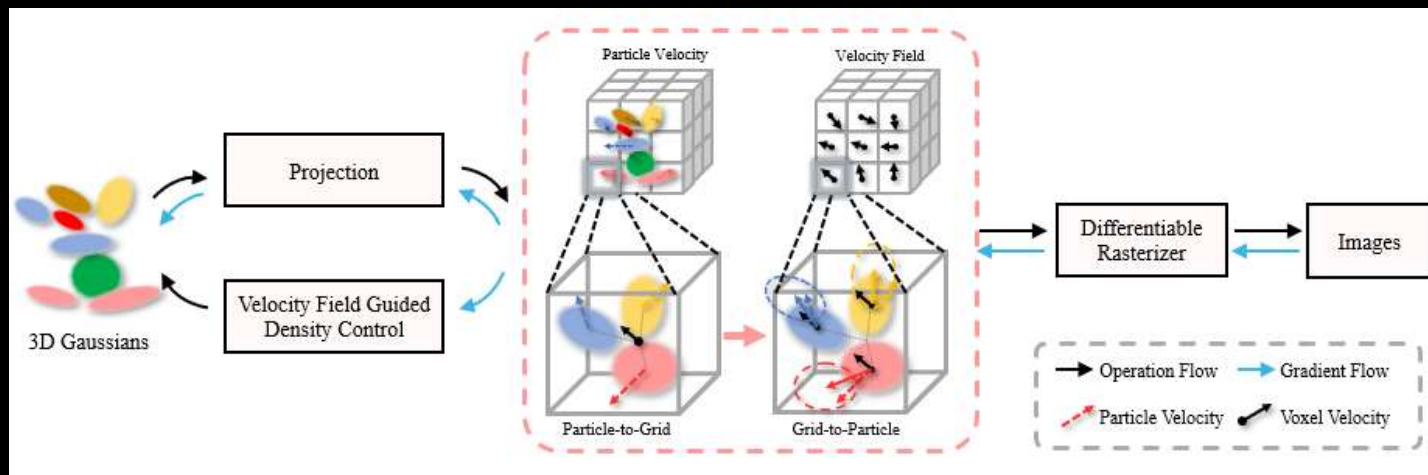


- $\mu$  : Position
- $o$  : Opacity
- $c$  : Color
- $s$  : Scaling
- $r$  : Rotation

Gradient Flow



Loss Function  $L$



# 3DGS的实时渲染-超分/插帧

[Ongoing]

计算步骤	3DGS Original	3DGS-Graphics	Sort-free GS	Mesh Rasterization
Preprocess	高斯球投影; 高斯包围盒(Tile);	高斯球投影	高斯球投影	点的投影
Sorting	Tile-based Sorting	Global Sorting	✗	✗
硬件化	✗	✓	✓	✓
着色方式	像素 $\xrightarrow{\text{Retrieve}}$ 颜色	颜色 $\xrightarrow{\text{Splat}}$ 像素	颜色 $\xrightarrow{\text{Splat}}$ 像素	颜色 $\xrightarrow{\text{Splat}}$ 像素
维护信息	$C = \sum_{i=1}^N \alpha_i T_i c_i, \quad \text{with } T_i = \prod_{j=1}^{i-1} (1 - \alpha_j).$ <ul style="list-style-type: none"> <li>透射率(乘法&amp;指数运算)</li> <li>颜色(求和)</li> <li>维护每个像素</li> </ul>	$C = \sum_{i=1}^N \alpha_i T_i c_i, \quad \text{with } T_i = \prod_{j=1}^{i-1} (1 - \alpha_j).$ <ul style="list-style-type: none"> <li>透射率(乘法&amp;指数运算)</li> <li>颜色(求和)</li> <li>维护Gaussian影响范围</li> </ul>	$C = \frac{c_B w_B + \sum_{i=1}^N c_i \alpha_i w(d_i)}{w_B + \sum_{i=1}^N \alpha_i w(d_i)},$ <ul style="list-style-type: none"> <li>总权重(乘法&amp;指数运算)</li> <li>颜色(求和)</li> <li>维护Gaussian影响范围</li> </ul>	$L = L_a + L_d + L_s$ <ul style="list-style-type: none"> <li>颜色(仅替换值)</li> <li>深度缓冲(仅替换值)</li> <li>维护三角面影响范围</li> </ul>
耗时(1920 × 1080)	143.0 ms	42.2 ms	34.1 ms	/

# 3DGS的重建、渲染及编码

- 重建计算：离线方式为主，可不断加速
  - 基函数的不断改进：Gaussian, GES, DoG, 小波, 有限元, ...
  - 更新策略：cloning, splitting, ...
- 渲染计算：优势在实时计算
  - 光栅化：可重用光栅渲染硬件管线,
  - Blending：参考透明渲染(OIT)的一些策略, 减少排序、blending计算
    - 白盒3DGS有更有效策略
- 压缩：3DGS的压缩及标准已有一些工作
  - 离线编码、实时解码硬化

# 一些思考：3DGS的发展方向

- 函数表达（函数空间）
  - 基函数：Gaussian, GES, DoG, 小波, 有限元, ...
  - 表达能力：低频、高频
- 渲染计算
  - 投影(splatting)要快, 渲染实时
  - 在某些应用, 渲染不需要实时, 有更大自由度
- 三步走
  - 保留渲染方法, 探索更好的基函数
  - 探索新的渲染方法, Gaussian基函数
  - 探索新的渲染方法, 探索更好的基函数



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

# “数字人建模、生成与渲染技术” 专栏

## ➤ 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

黄干芄,刘骊,付晓东,刘利军,彭玮

# GAMES: Graphics And Mixed Environment Symposium

## 图形学与混合现实在线平台

非常多的学习资料！！



- 主页: <http://games-cn.org>
- 宗旨: 图形学及相关领域交流的华人在线社区
- 隶属: 中国计算机学会计算机辅助设计与图形学专委会
  - 线上活动运营负责人: 刘利刚 (2017.6.-2018.12.), 周晓巍 (2019.1.-2021.9.), 胡瑞珍 (2021.10.-2024.7)
  - 新的接力棒: 韩晓光 (2024.8.-), ...

所有资料 (视频/PPT) 云端保存, 总观看 660+ 万人次

- 在线直播活动:
  - 每周四晚8:00-9:30的在线报告 (340+期)
  - 专题: CAD、虚拟现实、几何、绘制、模拟、视觉、可视化
  - 系列在线课程
- 在线交流微信群: 27个群 (12000+人)

加入微信群的方法: 在微信中搜索微信号, 加gameswebinar为好友。  
然后回复 “GAMES”即可获取群聊邀请。



# GAMES在线课程

● GAMES 基础课程 (1\*\*)  
● GAMES 专题课程 (3\*\*)

● GAMES 高级课程 (2\*\*)  
● GAMES 开发课程 (4\*\*)



闫令琪

101

现代计算机图形学入门

2020.2



胡鸣渊

201

高级物理引擎  
实战指南

2020.6



刘利刚

102

几何建模与处理

2020.10



闫令琪

202

高质量实时渲染

2021.3



黄其兴

203

三维视觉与理解

2021.7



王华民

103

基于物理的  
计算机动画入门

2021.11



王希

104

现代游戏引擎

2022.3



徐岗

302

等几何分析

2023.4



霍宇驰、袁亚振、高希  
峰胡义伟、高涛

106

现代图形绘制流水线  
原理与实践

2023.4



何小伟、蔡勇

401

泛动引擎(PeriDyno)  
物理仿真编程与实践

2023.3



刘利斌

105

计算机角色动画基础

2022.10



刘利刚、陈仁杰  
傅孝明、方清

301

曲面参数化

2022.10



孙启霖、彭神帆

204

计算成像

2022.7

谢谢！

