



生成式稀疏视角三维场景重建

段岳圻



清华大学电子工程系

2025.04

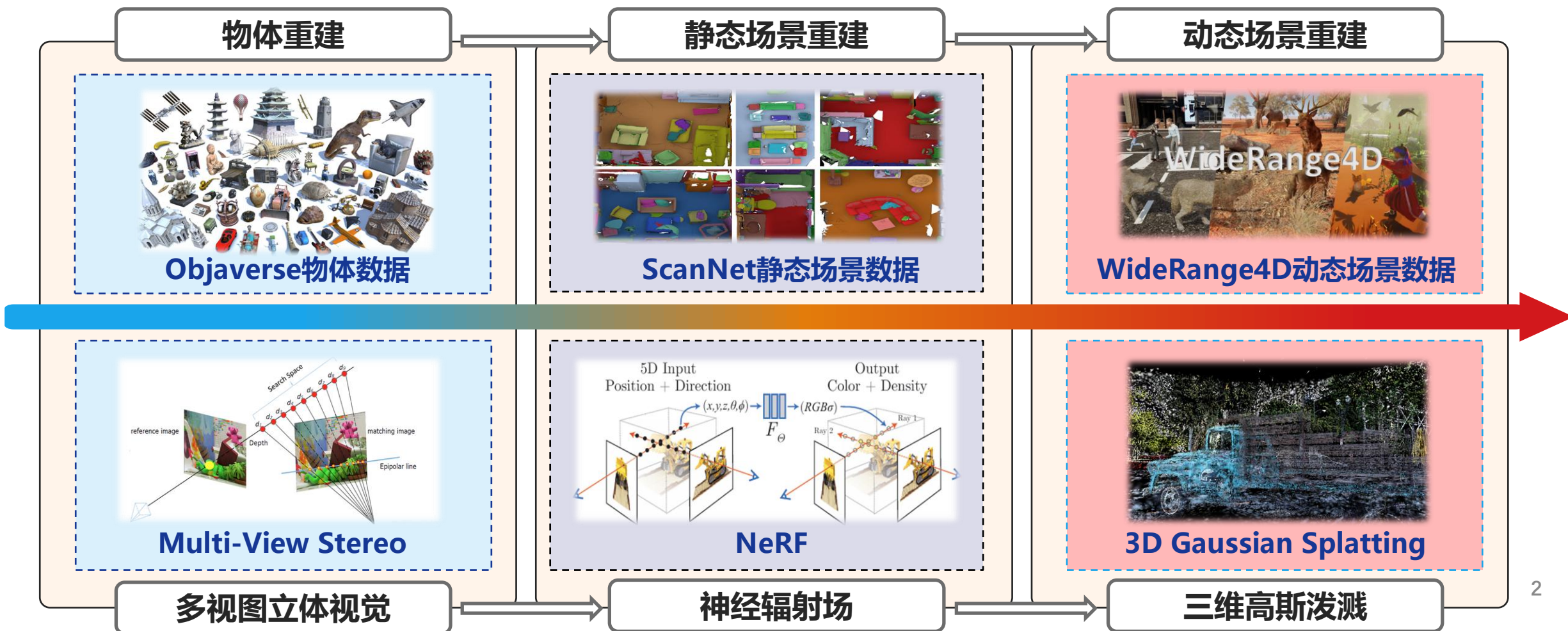
duanyueqi@tsinghua.edu.cn

Homepage: duanyueqi.github.io

多视角三维场景重建



三维场景重建是计算机视觉与图形学的基础问题，旨在从多视角观测精准恢复三维结构及实现新视角渲染



□ 现有基于NeRF和3DGS的三维场景重建范式

■ 优化模型: SparseNeRF, DNGaussian,

■ 优化时间较长, 体验差, 极少输入视角 (0.1~1) 下效果差

在复杂场景中稀疏视角信息观测不完备, “所见即所得” 的重建范式难以取得理想效果





ReconX: Reconstruct Any Scene from Sparse Views with Video Diffusion Model

Fangfu Liu, Wenqiang Sun, Hanyang Wang, Yikai Wang, Haowen Sun,

Junliang Ye, Jun Zhang, **Yueqi Duan***

Website: <https://liuff19.github.io/ReconX>

Github: <https://github.com/liuff19/ReconX>



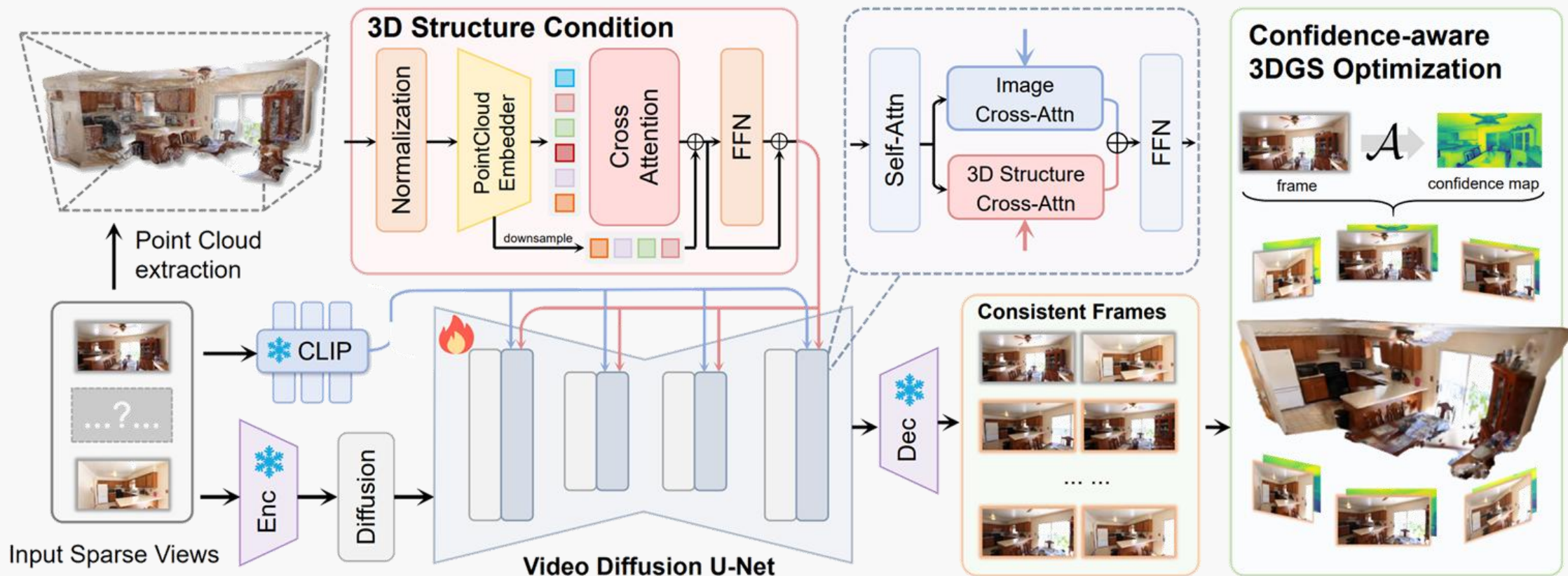
将稀疏视角重建任务建模为三维一致的图生视频任务

- “看得准的地方重建，看不准的地方生成”
- 充分释放视频生成模型的先验创造更多可见视角信息



实现统一生成和重建的新范式

- 三维一致视频生成：三维几何结构嵌入隐空间引导
- 帧间高频信息保持：置信度引导的优化方式



□ 仅输入稀疏视角 (少至两张), 释放视频扩散模型的**生成能力**解决**稀疏重建**问题



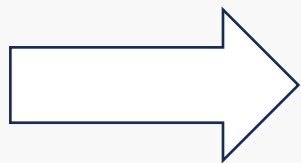
输入：两张稀疏视角

输出：三维一致场景渲染

□ 仅输入稀疏视角 (少至两张), 释放视频扩散模型的**生成能力**解决**稀疏重建**问题



输入：三张稀疏视角

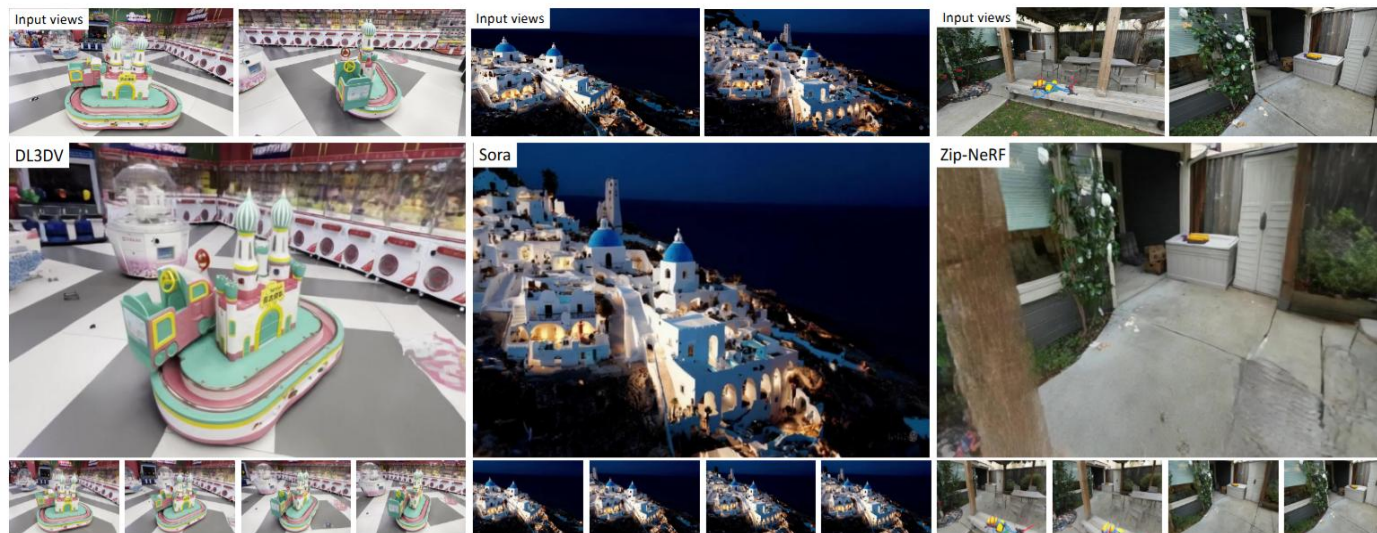


输出：三维一致场景渲染

VideoScene: Distilling Video Diffusion Model to Generate 3D Scenes in One Step

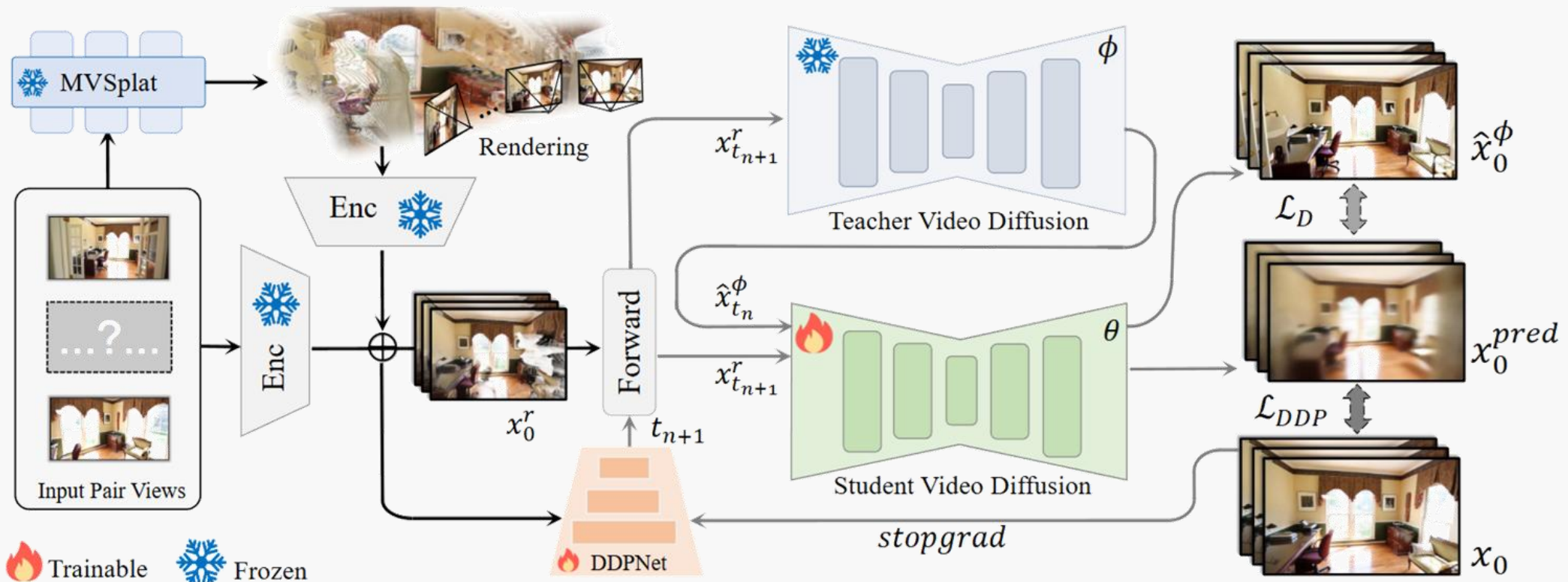
(Highlight)

Hanyang Wang, Fangfu Liu, Jiawei Chi, Yueqi Duan*



利用三维结构先验蒸馏视频生成的三维一致性

- 不需要视频动态性，仅需要三维一致性，充分蒸馏
- 输入稀疏视角，一步去噪 \rightarrow 生成视频时间从5min降至5s以内



□ 利用三维结构先验蒸馏视频生成的三维一致性

- 实验证明我们一步去噪的三维一致性充分优于现有视频模型
50step去噪的结果





Unique3D: High-Quality and Efficient 3D Mesh Generation from a Single Image

Kailu Wu, Fangfu Liu, Zhihan Cai, Runjie Yan, Hanyang Wang, Yating Hu,

Yueqi Duan*, Kaisheng Ma*

Website: <https://wukailu.github.io/Unique3D/>

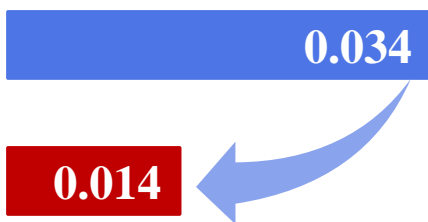
Github: <https://github.com/AiuniAI/Unique3D>

☆ Star 3.3k

- 实现**单图30秒**生成2K分辨率高清三维物体，比之前Adobe提出的方法精度误差**降低到原来的1/2**，真实感**提升近50%**
- 五个月内GitHub星标超过3100星，被开源社区选为最佳三维生成模型，连续一周成为Hacker News 榜首

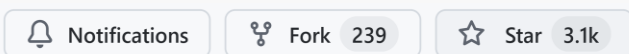
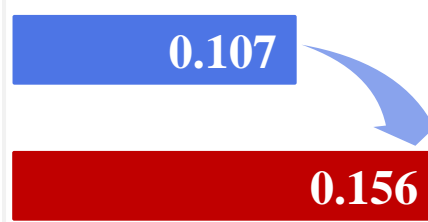
精度误差

Adobe
方法
我们
方法



真实感

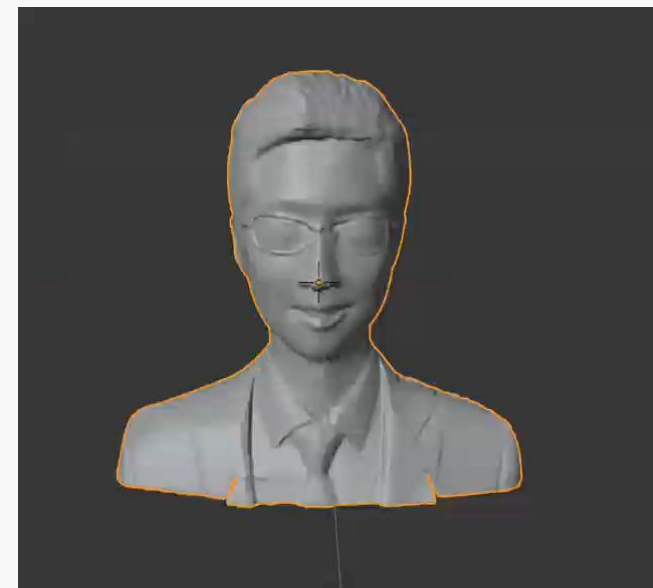
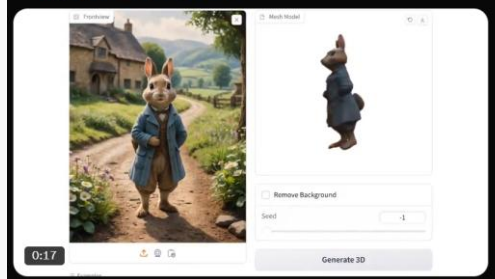
Adobe
方法
我们
方法



Rishub @itsrishub · 6月9日
This is the best model for generating 3D models ~ Unique3D.

The model almost perfectly fills in the missing parts of an image to turn your photo into a 3D model. Unfortunately, the best doesn't mean perfect, the developers can still improve it.

HF: huggingface.co/spaces/Wuvin/U...





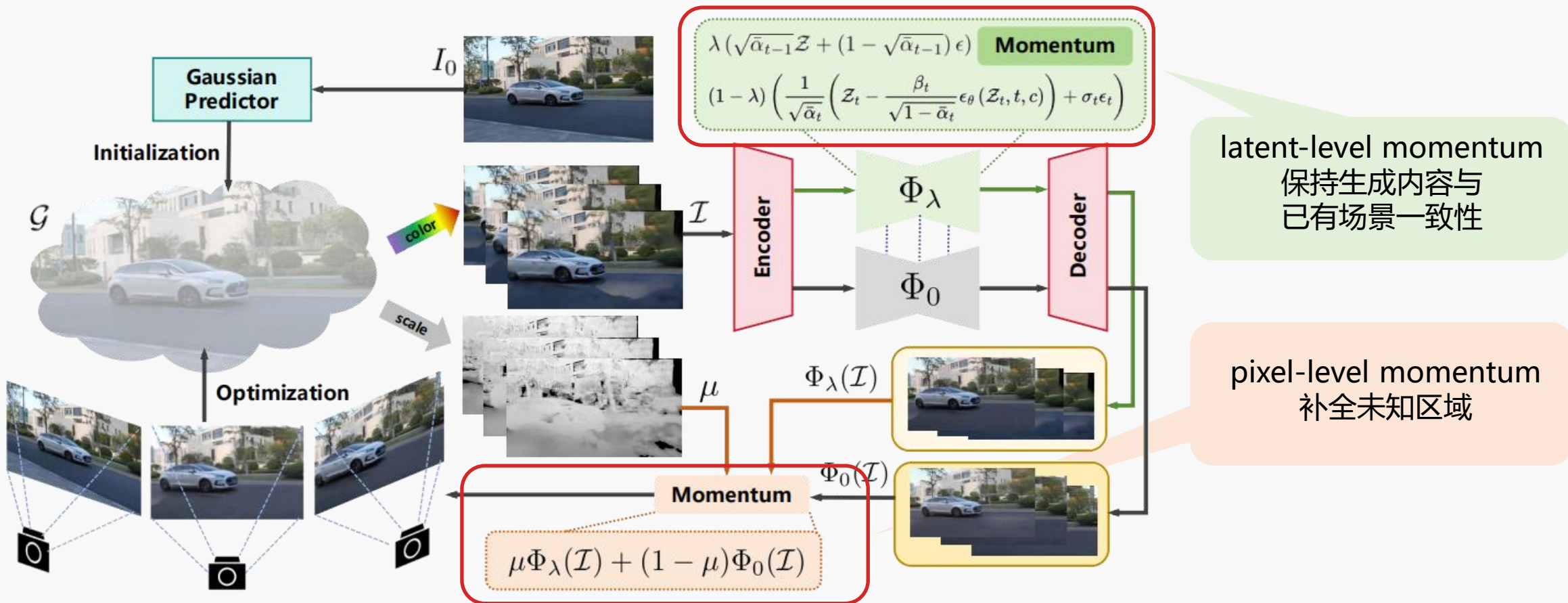
Scene Splatter: Momentum 3D Scene Generation from Single Image with Video Diffusion Model

Shengjun Zhang, Jinzhao Li, Xin Fei, Hao Liu, Yueqi Duan*

Github: <https://shengjun-zhang.github.io/SceneSplatter/>

视频生成模型驱动的单图三维生成

- 将已知场景信息作为**动量**引入生成框架，保持生成内容一致性

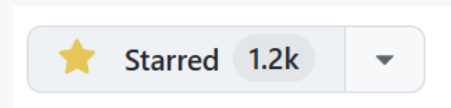


DimensionX: Create Any 3D and 4D Scenes from a Single Image with Controllable Video Diffusion

Wenqiang Sun, Shuo Chen, Fangfu Liu, Zilong Chen, **Yueqi Duan**, Jun Zhang, Yikai Wang

Website: <https://chenshuo20.github.io/DimensionX/>

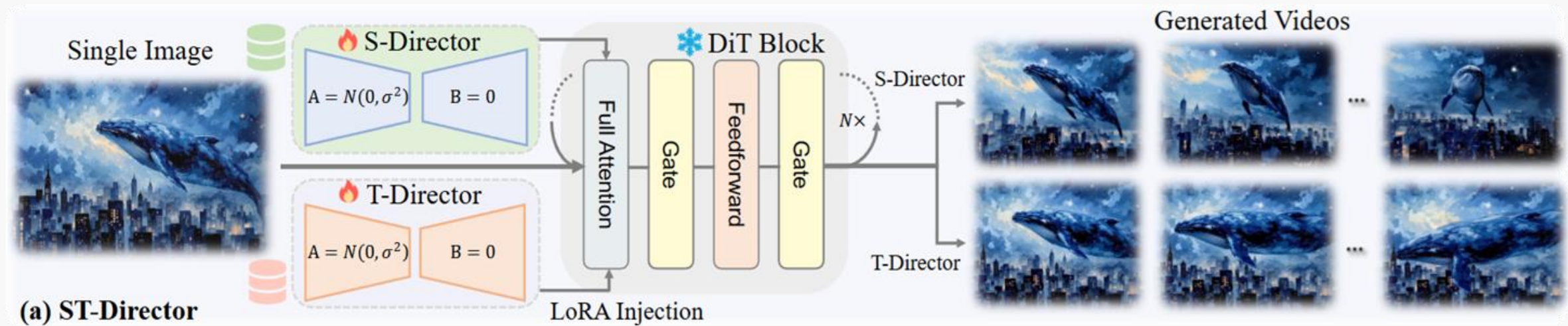
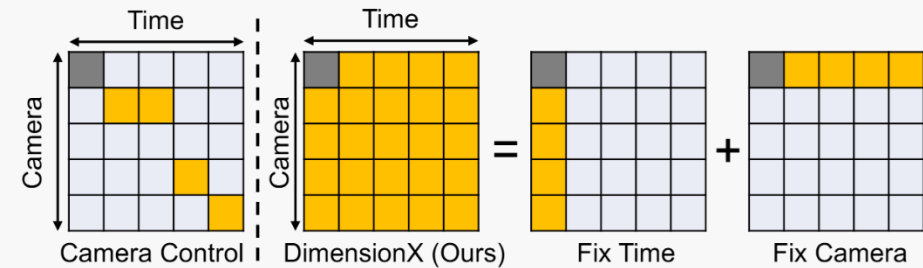
Github: <https://chenshuo20.github.io/DimensionX/>



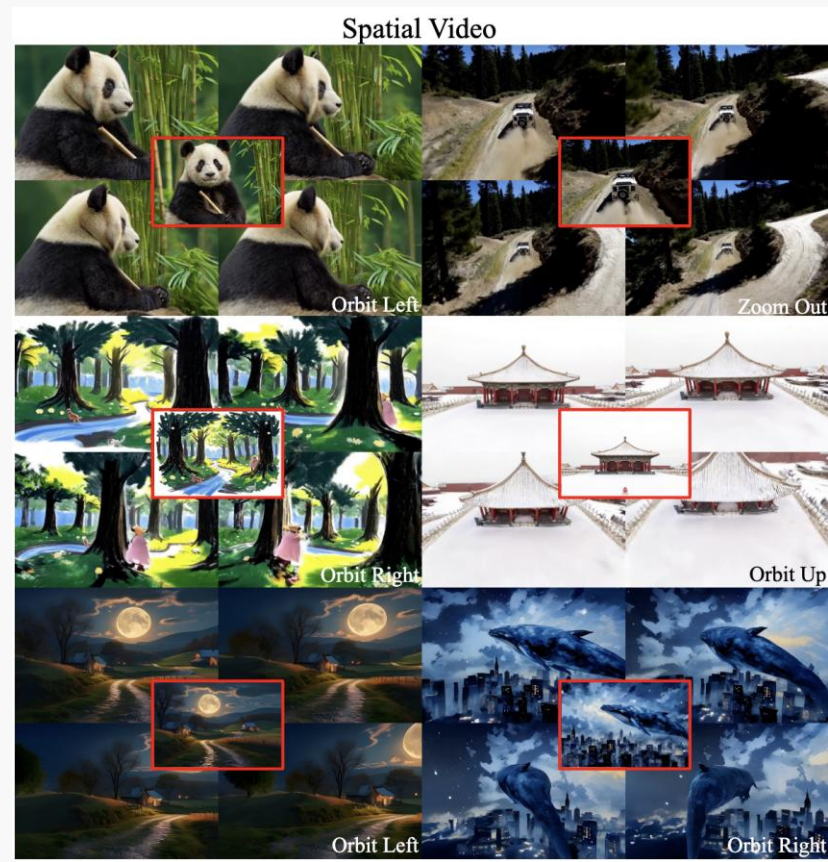
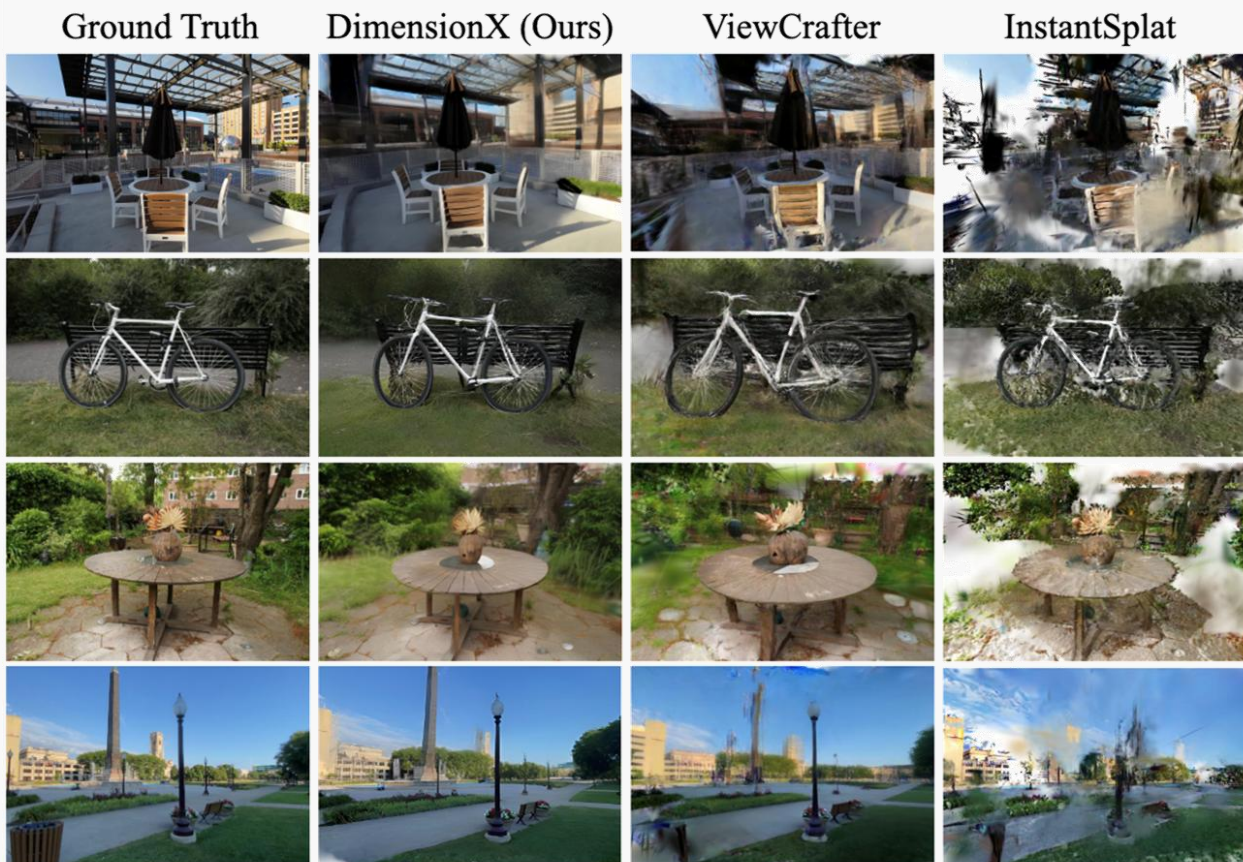
□ 实现全时空360三维场景生成

□ 将视频的时空维度进行分解

- 沿时间维度切割 → 360环绕三维场景
- 沿相机维度切割 → 多视角可控视频生成



- 稀疏视角下效果明显好于ViewCrafter等最新基线方法
- **高泛化性**：单图可以生成卡通漫画三维场景



□ 效果：单图生成三维场景 + 可控视频生成

