

动态场景重建和预测

Dynamic Scene Reconstruction & Prediction

汇报人：李亚城

时间：2025/6/13



Contents

结构大纲

01 Gaussian Prediction

02 研究背景

03 方法框架

04 实验分析

05 讨论与总结





01

Gaussian Prediction



GaussianPrediction: Dynamic 3D Gaussian Prediction for Motion Extrapolation and Free View Synthesis

Boming Zhao*
bmzhao@zju.edu.cn
Zhejiang University
Hangzhou, China

Lin Zeng
22251265@zju.edu.cn
Zhejiang University
Hangzhou, China

Yinda Zhang
yindaz@google.com
Google Inc.
Mountain View, USA

Yuan Li*
yuan_li@zju.edu.cn
Zhejiang University
Hangzhou, China

Yujun Shen
shenyujun0302@gmail.com
Ant Group
Hangzhou, China

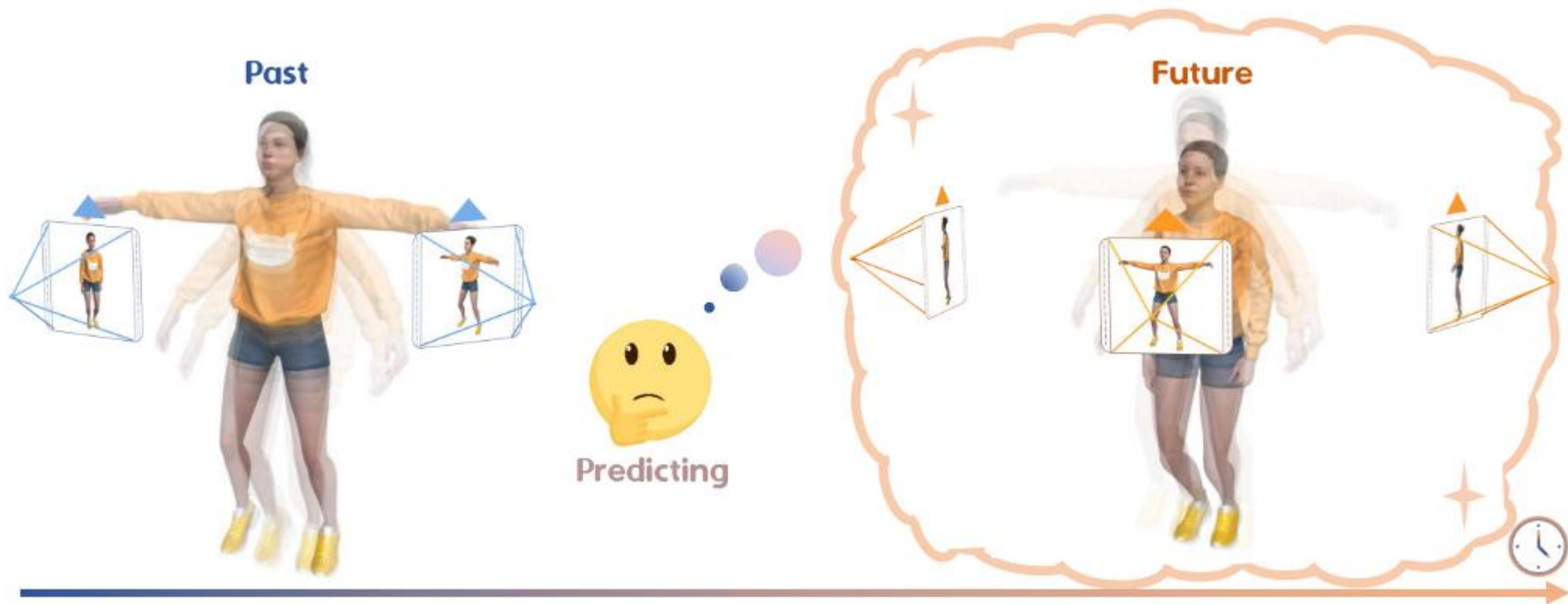
Hujun Bao
bao@cad.zju.edu.cn
Zhejiang University
Hangzhou, China

Ziyu Sun
sunzy2121@mails.jlu.edu.cn
Jilin University
Changchun, China

Rui Ma
ruim@jlu.edu.cn
Jilin University
Changchun, China

Zhaopeng Cui[†]
zhpcui@gmail.com
Zhejiang University
Hangzhou, China

GaussianPrediction: Dynamic 3D Gaussian Prediction for Motion Extrapolation and Free View Synthesis





02

研究背景



“ 如何预测动态环境中的未来情景？ ”

场景重建

从灵活选择的视点渲染场景图像，但它没有纳入预测未来环境状态所需的时间预测元素。

视频预测

从同一视点观察到的结果，从特定视点预测场景的未来动态，但在可视化任意视点方面存在不足，限制了其在理解近期未来方面的有效性。

场景重建

3D GaussianSplatting 技术凭借其快速重建和高效高质量渲染，成为动态重建场景的研究热点，方法包括显式扩展时变高斯特征和利用隐式变形场。此外，一些研究将显式三维高斯表示与隐式神经表示（如 **MLPs** 或 HexPlanes）结合，实现了高质量的新视图合成。该方法不仅具备交互式渲染帧率，还支持灵活编辑，对象插入和场景修改。

动态预测

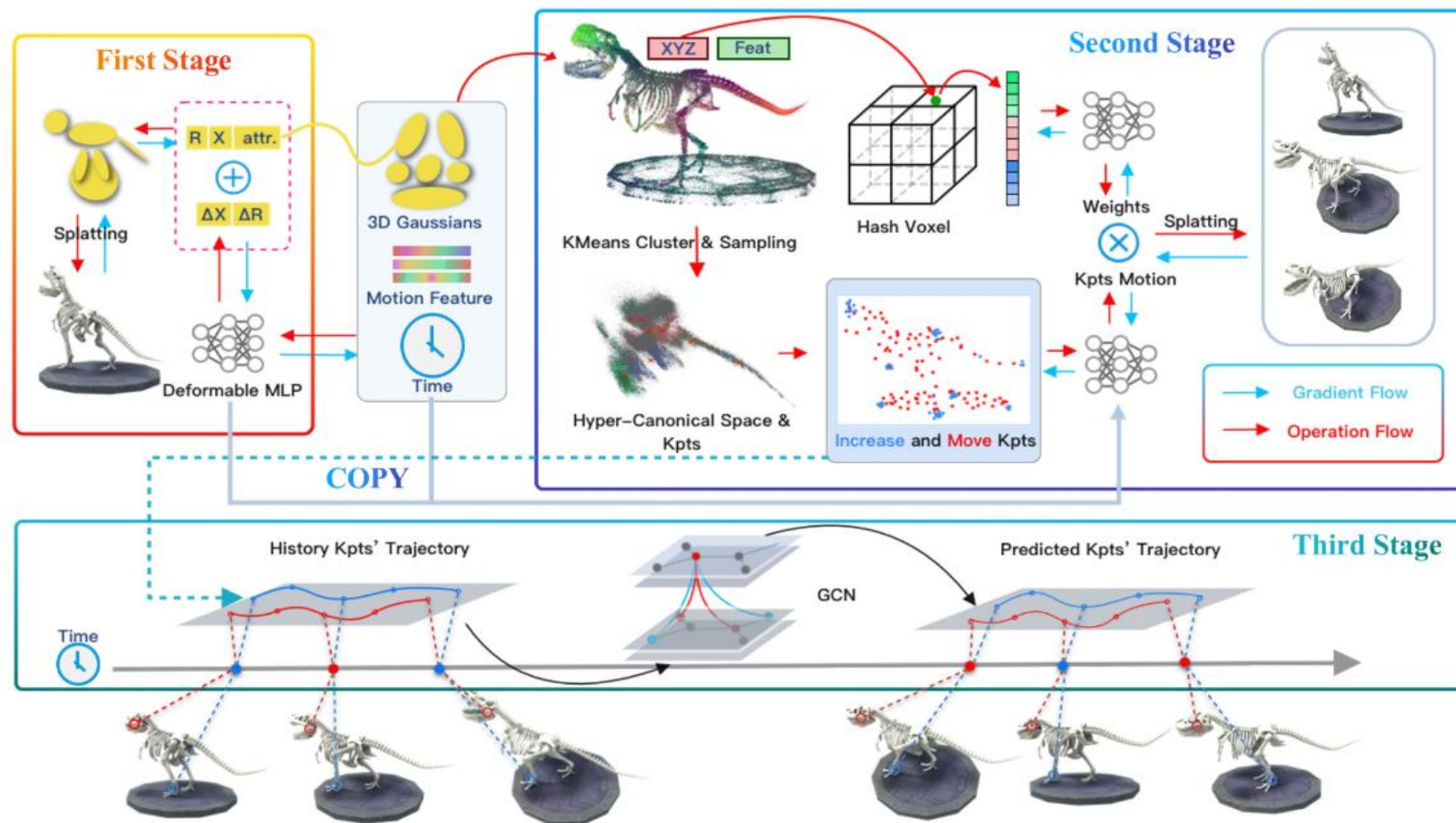
现有方法主要包括三类：一是3D-CNN、LSTM或Transformer结合时空信息进行连续帧预测；二是基于VAE的随机潜在变量方法生成多样化未来帧；三是GAN和扩散模型通过历史数据生成逼真视频。此外，4D数据如人体骨骼运动的预测采用RNN、VAE、GAN和扩散模型效果良好，而场景级动态点云预测则依赖**GCN**，具备优异的泛化能力和快速预测性能。



03

方法框架





高斯正则空间构建

通过变形建模捕获动态场景的外观和几何形状

同心运动蒸馏

提取关键点简化高斯节点数量

超规范空间特征聚类

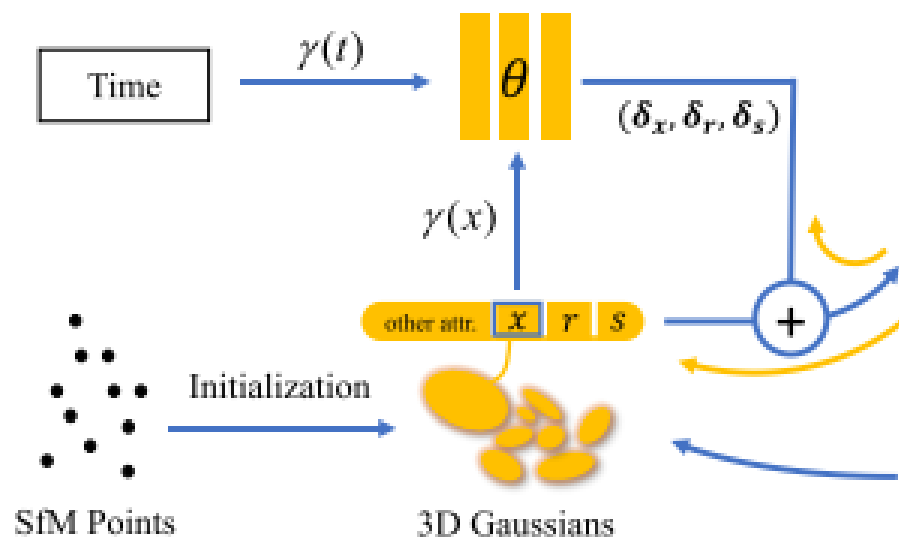
特征聚类选择关键点提高外观和运动表征能力

GCN 未来运动预测

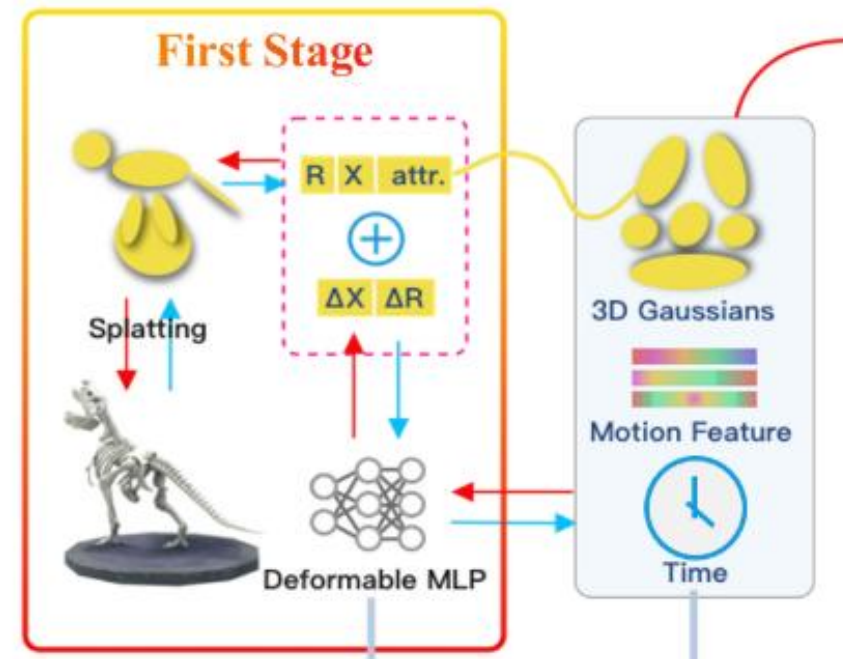
图卷积网络预测3D关键点的未来运动

动态场景变形驱动

基于预测的关键点运动生成未来场景



$$(\delta x, \delta r, \delta s) = \mathcal{F}_{\theta}(\gamma(\text{sg}(\mathbf{x})), \gamma(t))$$



$$\Delta \mu^t, \Delta q^t = D(\gamma(\mu), m, \gamma(t))$$

与Deformable 3D GS^[1]相比，GaussianPrediction对每个高斯函数设置了一个运动特征 m 来编码运动信息

■ 针对Deformable MLP容易陷入局部最优问题：

➤ 在输入 μ 时引入一种退火噪声指数 ε （施加微小扰动）

➤
$$\varepsilon(i) = \mathcal{N}(0, 1) \cdot N_s \cdot (1 - \min(1, \frac{i}{10000}))$$

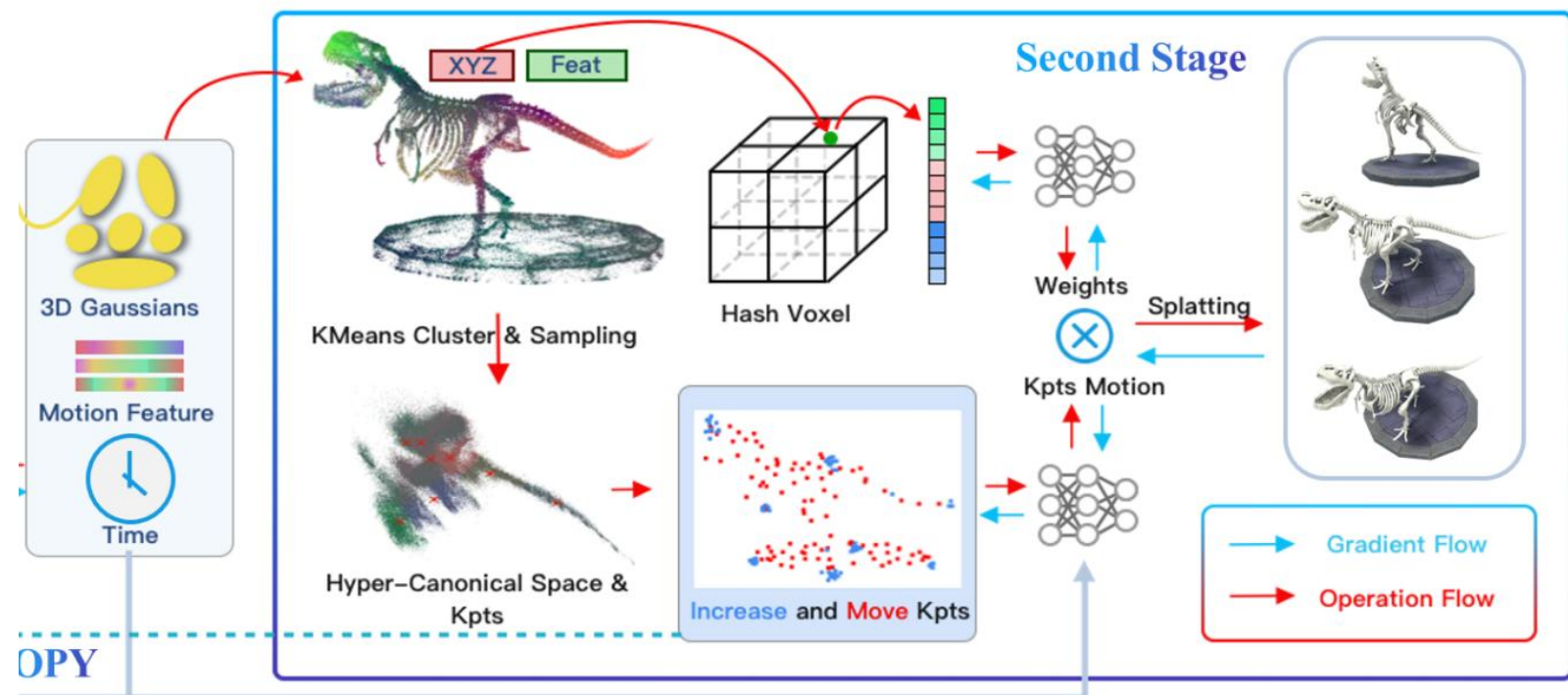
■ 针对动态运动的一些不可逆现象（表面出现或者消失）：

➤ 一个直观的想法是，允许3D高斯模型保持相同的属性——但在某个时间点之前可渲染，之后它们就失去渲染能力。

➤ 引入生命周期，用计算出的生命周期乘以高斯函数的不透明度

$$\psi(G_i, t) = \frac{1}{1 + e^{(-10\Delta_o(G_i, t))}},$$
$$\Delta_o = D_o(\gamma(\mu_i), m_i, \gamma(t)).$$

步骤二：基于关键点的同心运动蒸馏



初始化Kpts: 在超规范空间中使用聚类技术将3D-GS函数进行聚类, 使得具备运动相似性和空间接近性的高斯函数由相同的关键点驱动

自适应增加Kpts: 复杂运动3D-GS附近添加关键点, 选择较大梯度的GS函数进行100倍的均匀下采样, 以生成新添加的关键点位置

时间无关的权重学习: 学习 W_i^T 和 W_i^Q , 认为正则空间中的 G_i 的运动受 N_{near} 关键点的影响 (在超规范空间中离每个高斯点最近的关键点——同时考虑空间接近性 μ 和运动相似性 m 。)

关键点运动: $T_i^t, Q_i^t = D(\gamma(\mu_i^k), m_i^k, \gamma(t))$

高斯属性回推: $P_t^{key} = \{G_i^t : (\mu_i + \Delta\mu_i^t), (q_i \otimes \Delta q_i^t), s_i, c_i, \sigma_i, m_i\},$

$$\Delta\mu_i^t = \sum_{k \in K} (w_{i \leftarrow k}^T \cdot T_k^t), \Delta q_i^t = \sum_{k \in K} (w_{i \leftarrow k}^Q \cdot Q_k^t),$$

步骤三：基于GCN的运动预测

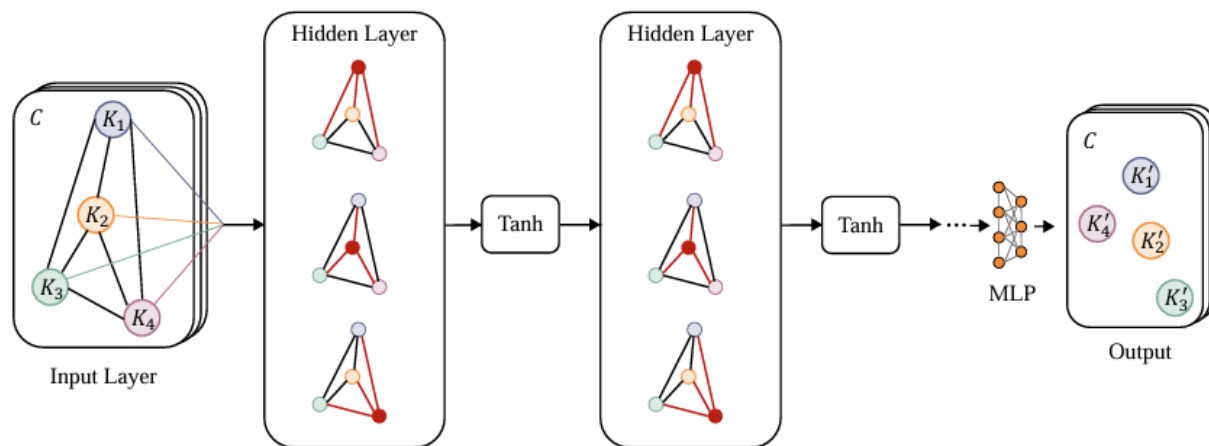
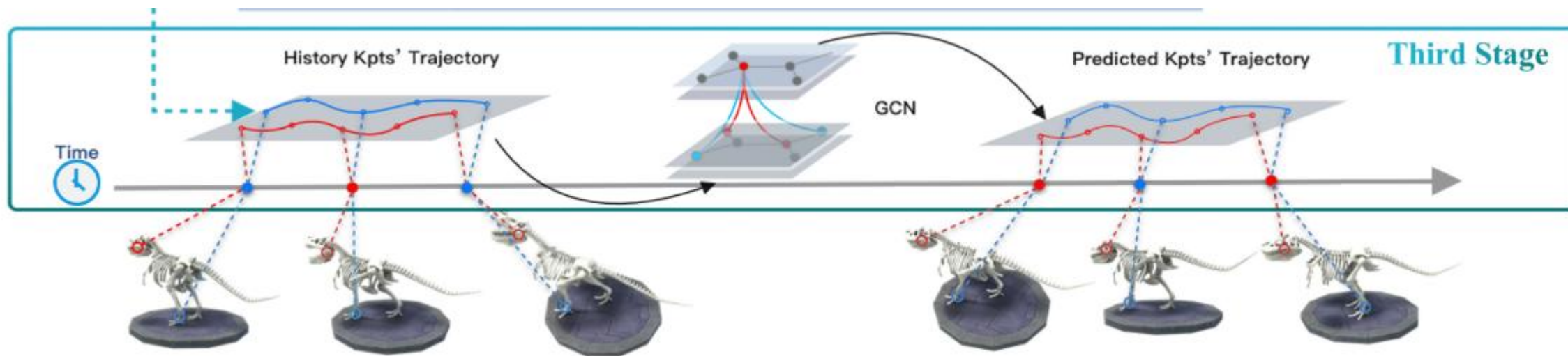


Figure A: GCN network architecture.

利用图卷积网络（GCN）的功能对给定场景中关键点的动态运动模式进行建模和预测，使用高斯属性计算公式计算每个3D高斯的运动

- 使用两个单独的GCN网络分别预测关键点的3D位置和四元数表示的旋转：
 - 将整个场景的 N 关键点建模为一个全连接图
 - 图卷积层 g 从上一层中获取一组权重 $W_g \in \mathbb{R}^{F \times F}$ 和特征 $X_g \in \mathbb{R}^{N \times F}$ ，并输出下一层的特征 $X^{g+1} = \sigma(AX^gW^g)$
 - 模型以10帧关键点的三维位置和旋转作为输入，预测关键点的后续信息（滑动窗口策略）

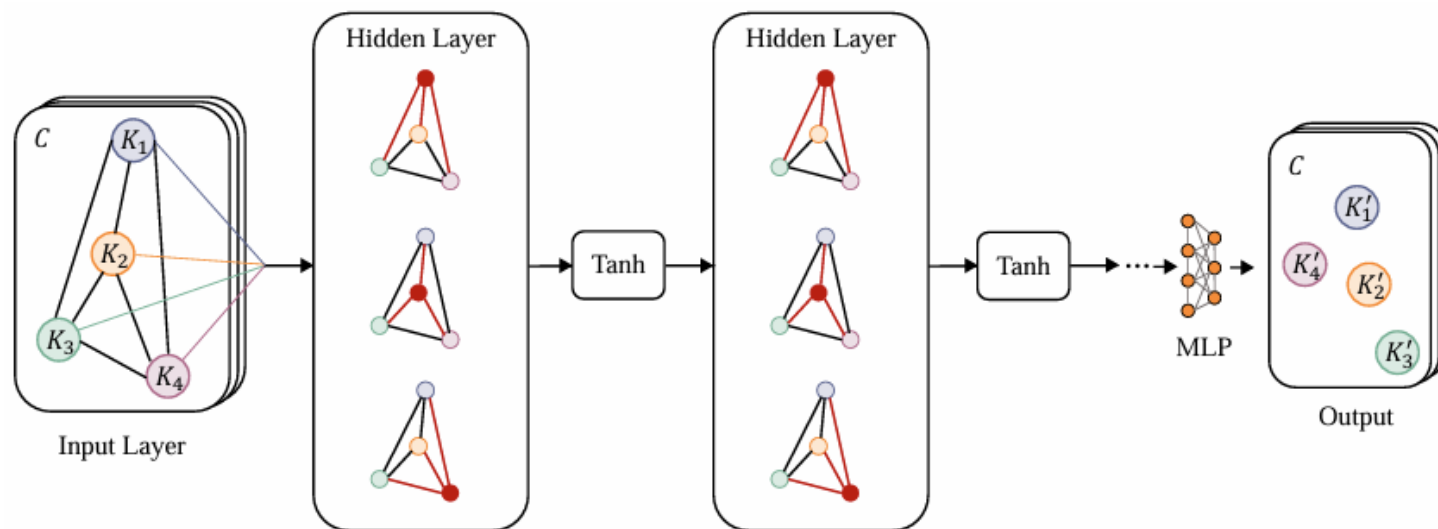


Figure A: GCN network architecture.



04

实验分析





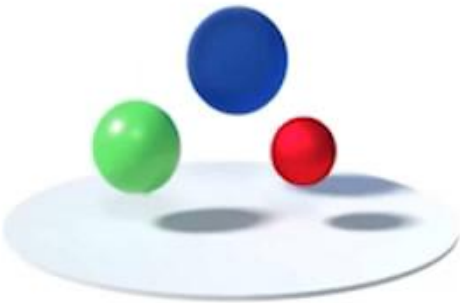
结果表明，本文的方法优于现有的基于NeRF或基于高斯的方法。图中显示的渲染结果可以看出，本文的方法比其他方法实现了更高的质量，重建了更多的动态场景细节。

cks	Bouncingballs			Hellwarrior		
	LPIPS(↓)	PSNR(↑)	SSIM(↑)	LPIPS(↓)	PSNR(↑)	SSIM(↑)
	.0360	35.00	.9835	.0391	39.20	.9763
	.0210	37.69	.9919	.0150	38.52	.9754
	.0126	41.01	.9953	.0093	41.54	.9873
	.0099	41.57	.9954	.0086	41.73	.9874
Hook	Hook			Average		
	LPIPS(↓)	PSNR(↑)	SSIM(↑)	LPIPS(↓)	PSNR(↑)	SSIM(↑)
	.0208	33.34	.9711	.0458	35.18	.9786
	.0164	33.83	.9728	.0338	36.92	.9836
	.0063	37.42	.9867	.0144	40.43	.9918
	.0057	37.44	.9868	.0137	40.58	.9919

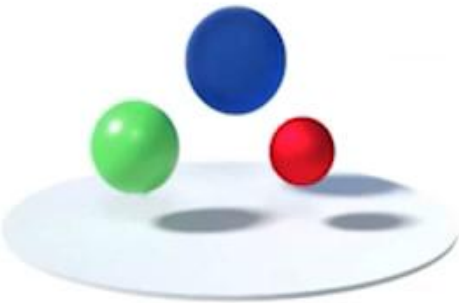


Prediction Result on Synthetic Dataset

Bouncing-Balls



Ours



TiNeuVox-B
[Fang et al. 2022]



Deform-GS
[Yang et al. 2024]

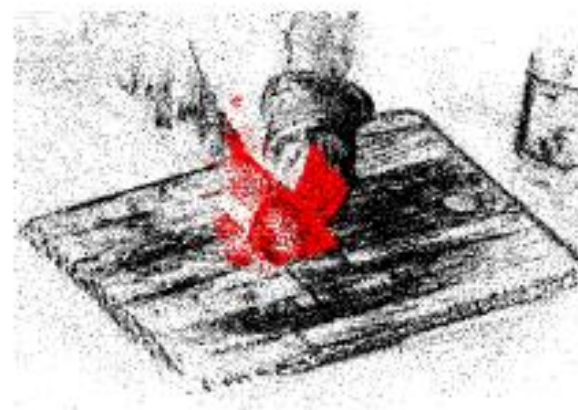
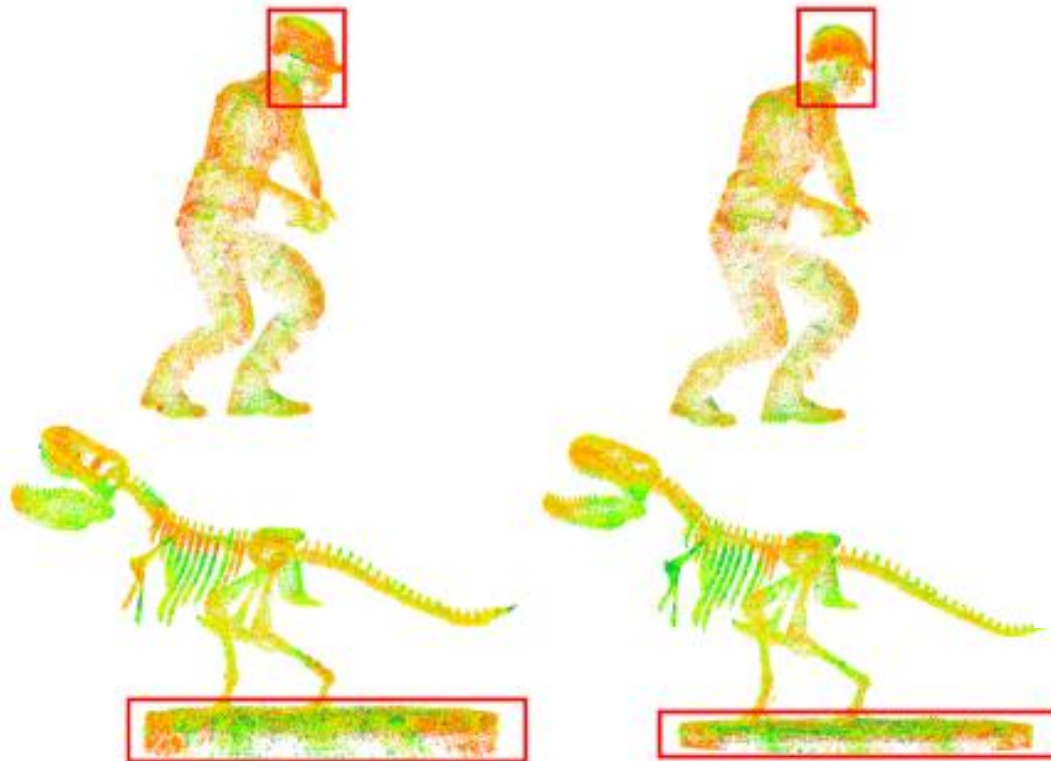


4D-GS
[Wu et al. 2024]

Train Sequence

W/O Annealing Noise

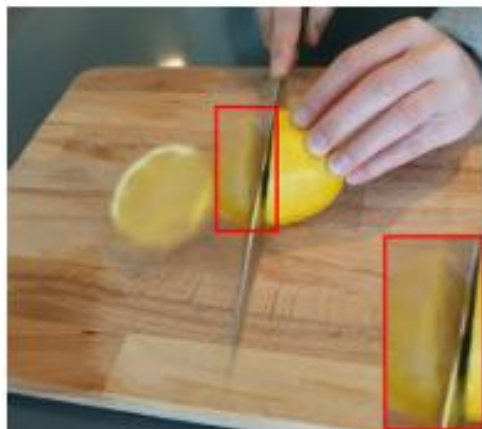
Full Model



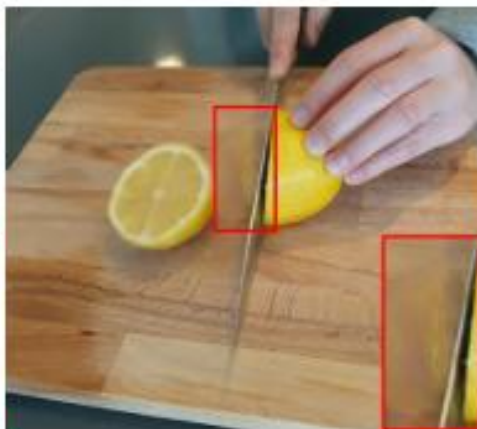
(a) Searching nearest key points in 3D space.



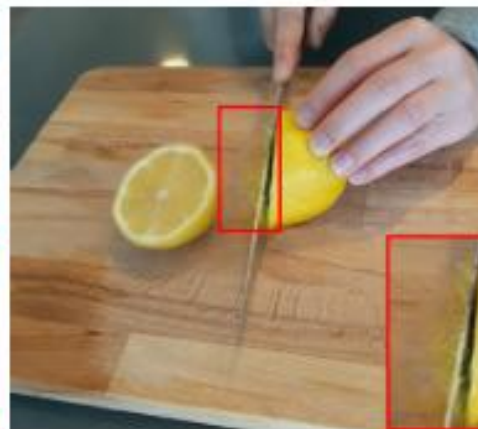
(b) Searching nearest key points in Hyper-Canonical space.



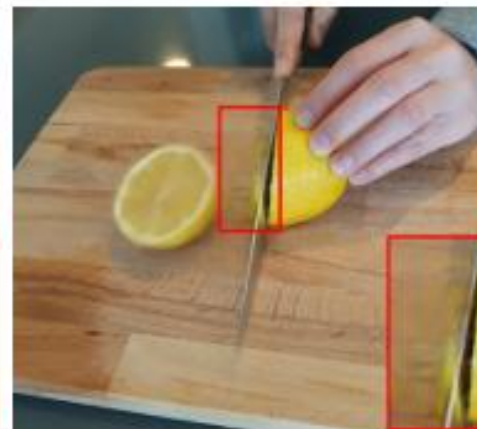
(a) w/o Hyper Space Init.



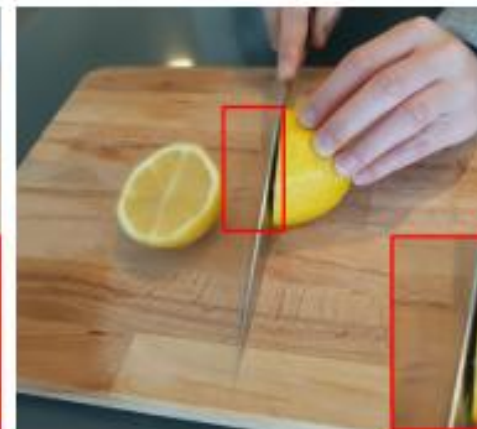
(b) w/o Adap Increasing.



(c) w/o Hyper Space K-NN.



(d) w/o Lifecycle.



(e) Full Model.



05

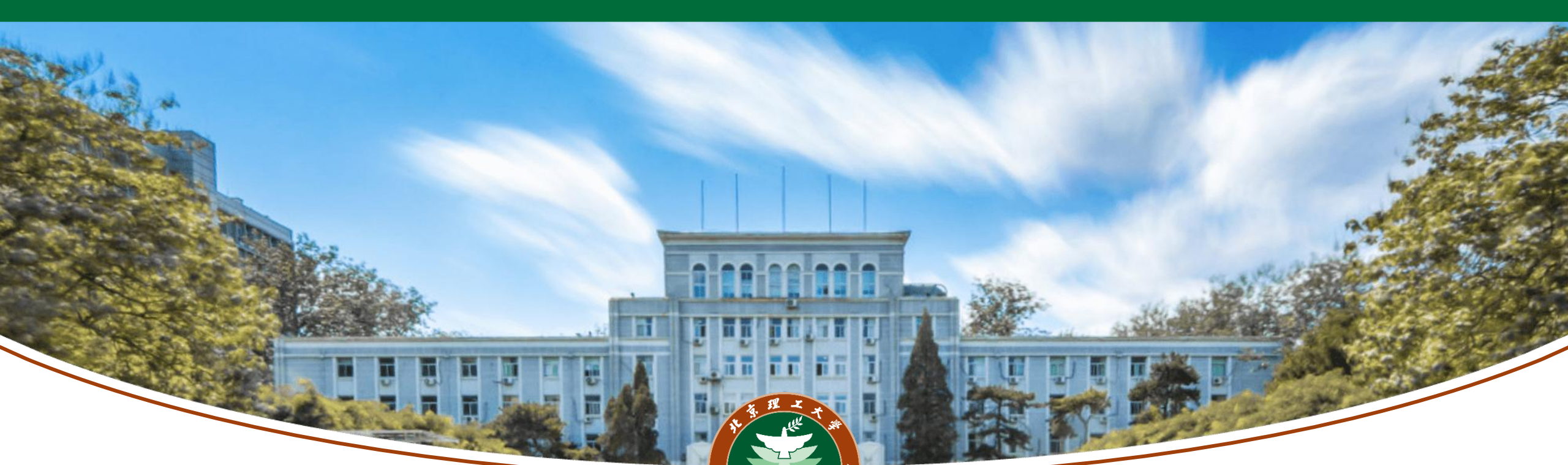
讨论与总结





■ GaussianPrediction:

- 使用deformable MLP记录高斯属性随时间的变化（引入了额外的时间表征-顺势提出超空间的概念）
- 提出关键点驱动运动框架，减少模型预测需要处理的信息数量，同时维持一定的几何特征
- 利用GCN来学习关键点的运动模式，实现未来场景（短时间）预测



请批评指正

Thanks for Your Attention