Dynamic Scene Representations for Volumetric Video

Xiaowei Zhou

浙江大学
ZHEJIANG UNIVERSITY

Spatial Video in Vision Pro

CHINA3DV 2025

基于Gaussian Splatting的高质量4D空间视频

孙佳明，徐震，王一凡，弓胜虎，周晓巍

浙江大学

随着Gaussian Splatting系列技术的逐渐成熟，4D空间视频的清晰度、渲染效率与文件体积等性能参数都得到了长足的进展，高质量捕捉并重建时间与空间信息的能力已经从科幻电影中走向现实。在本次演示中，观众可以在VR头显里面身临其境地体验到多段武打、舞蹈等精彩表演片段，亲身感受4D空间视频这个未来媒体格式的魅力。


亿点点不一样
双节棍
372.2万
8441
08:11
我们这次还使用了一个新的技术 叫作4D高斯

Volumetric video

2D video

Single-view capture

Coding

Code

Decoding

Fixed viewpoint

3D volumetric video

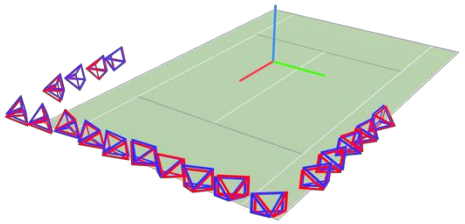
Multi-view capture

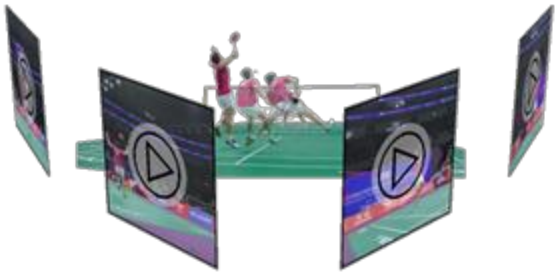
Reconstruction

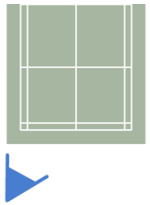
3D scene representation

Rendering

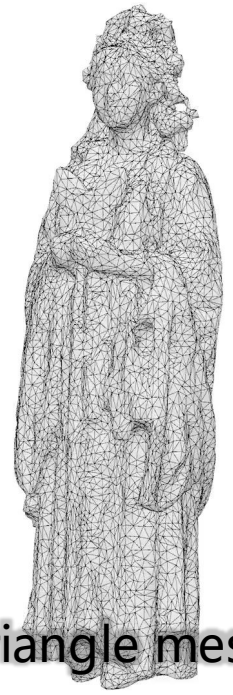
Free viewpoint







3D scene representations for volumetric video



Triangle mesh

Rendering

Gaussian primitives

Rendering

Surface representation

Hard to optimize with rendering loss,
resulting in limited rendering quality

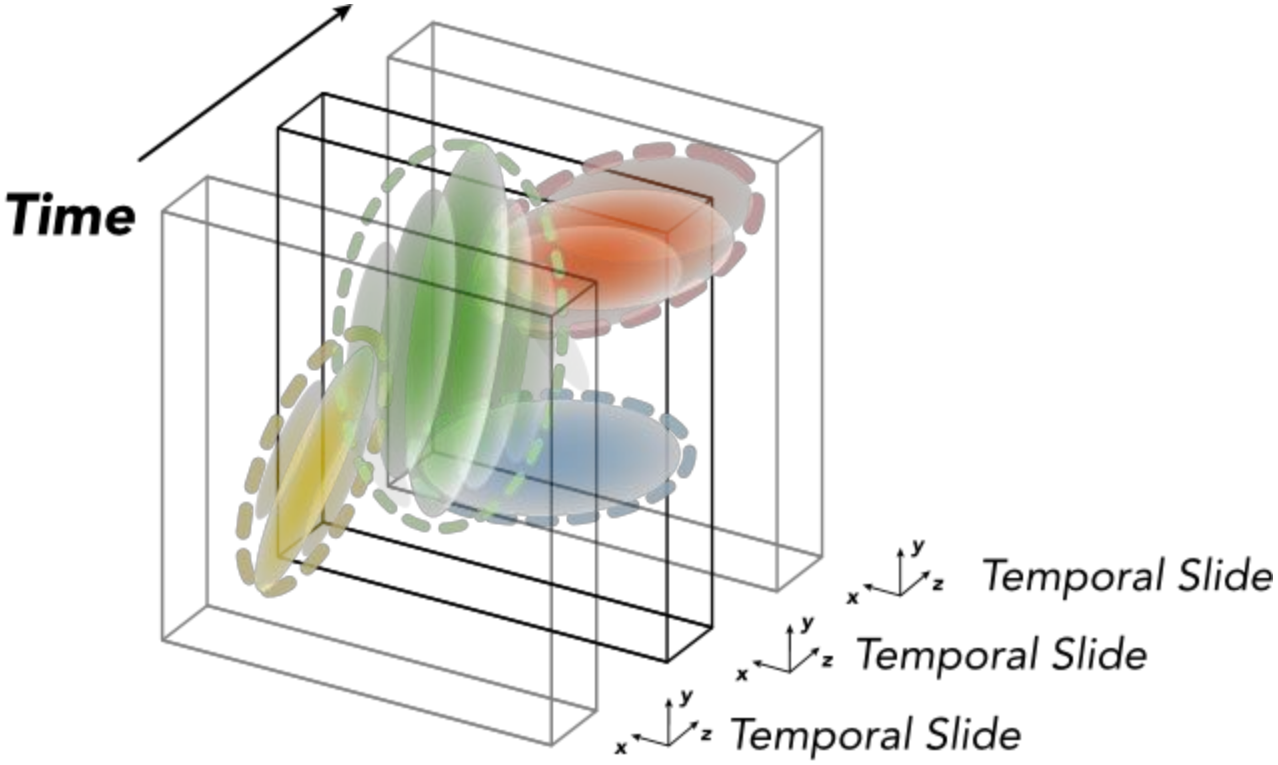
Volume representation

Easy to optimize with rendering loss,
resulting in good rendering quality

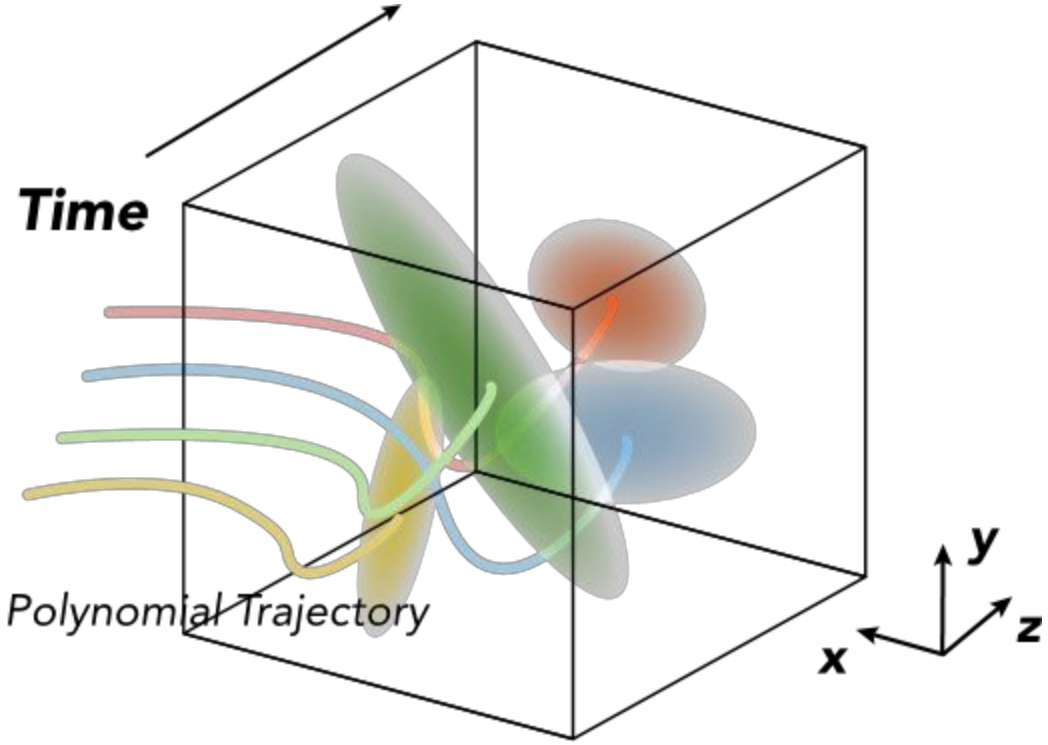
4K4D: Real-Time 4D View Synthesis at 4K Resolution

CVPR 2024

Recent work extending 3DGS to 4D



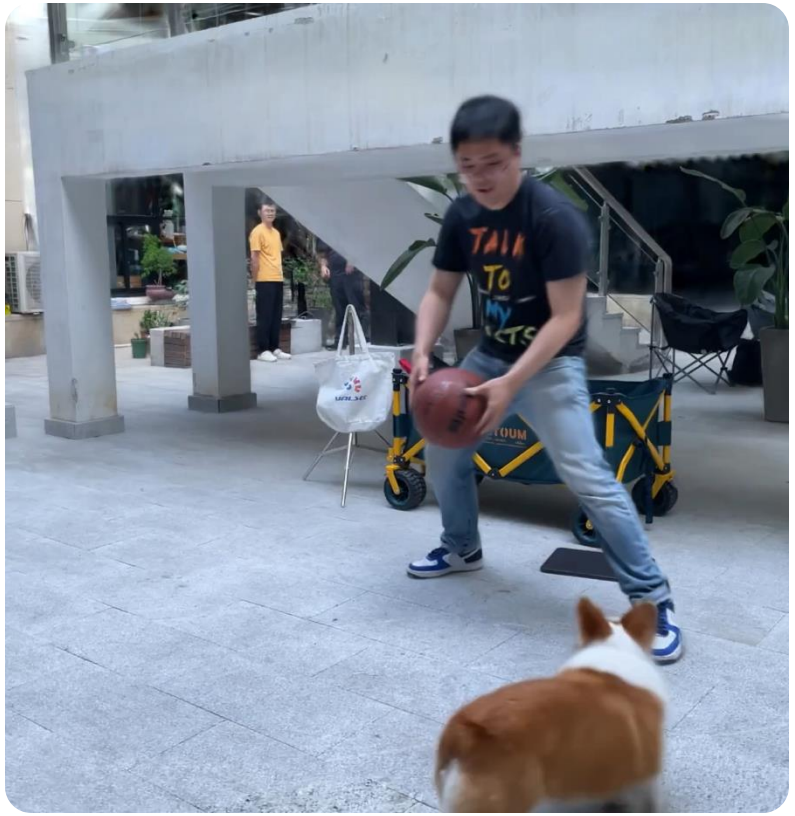
4DGS



Trajectory + 3DGS

- 4D-Rotor Gaussian Splatting: Towards Efficient Novel View Synthesis for Dynamic Scenes, SIGGRAPH 2024
- Real-time Photorealistic Dynamic Scene Representation and Rendering with 4D Gaussian Splatting, ICLR 2024
- Spacetime Gaussian Feature Splatting for Real-Time Dynamic View Synthesis, CVPR 2024
- Shape of Motion: 4D Reconstruction from a Single Video

Challenge 1: Fast and complex motion



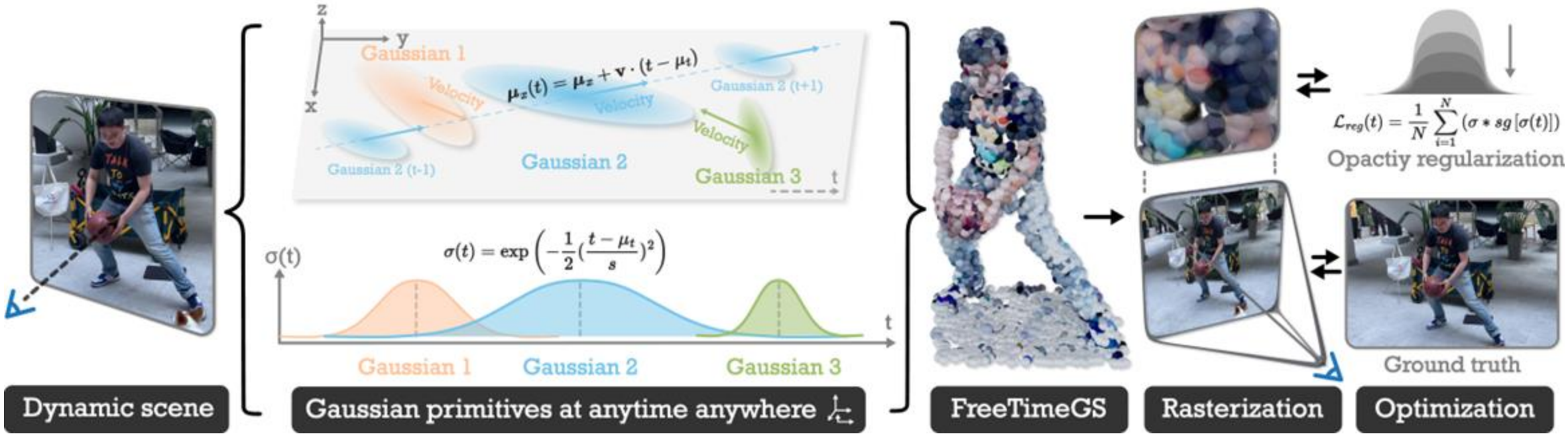
Hard to optimize

4DGS-Based Motion
Entangled Spatial and temporal dimension

Trajectory-Based Motion
Over-parameterization of trajectory

FreeTimeGS

1) Disentangle geometry and motion
2) Piece-wise linear motion



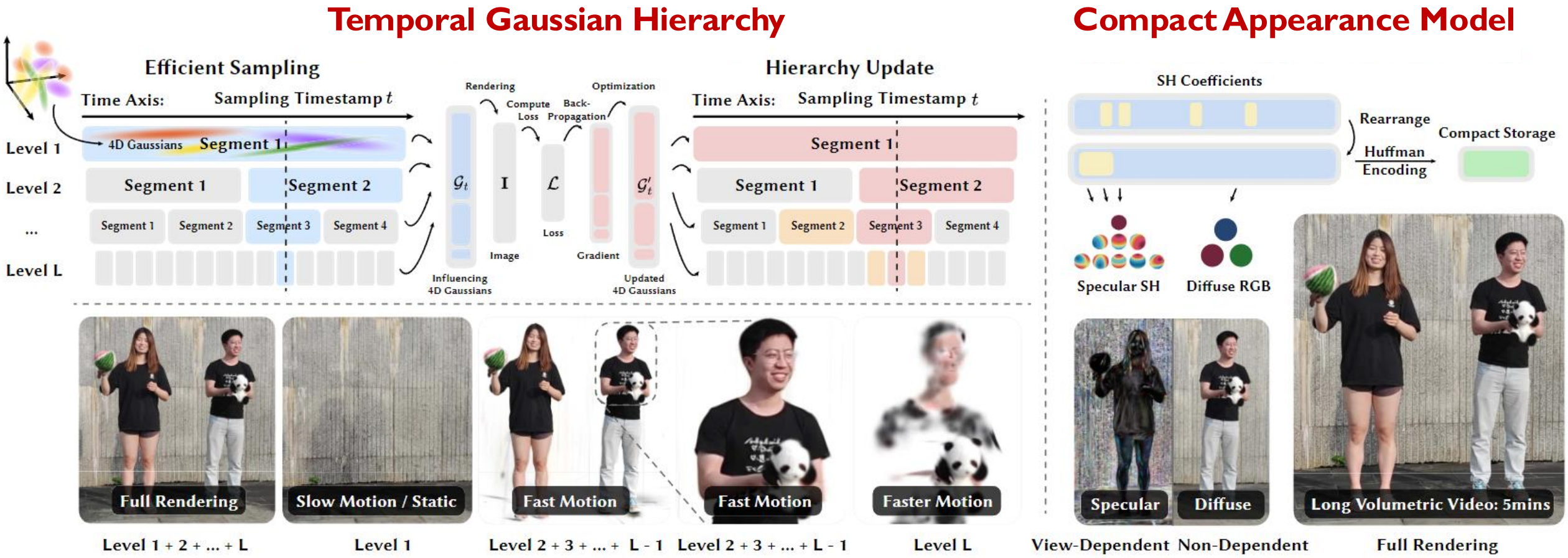
FreeTimeGS: Free Gaussian Primitives at Anytime Anywhere for Dynamic Scene Reconstruction, CVPR 2025

Challenge 2: Representing long volumetric video

Challenge :
- Large training cost and storage usage

Goal:
- Long videos
- SOTA quality!
- Real-time Rendering!

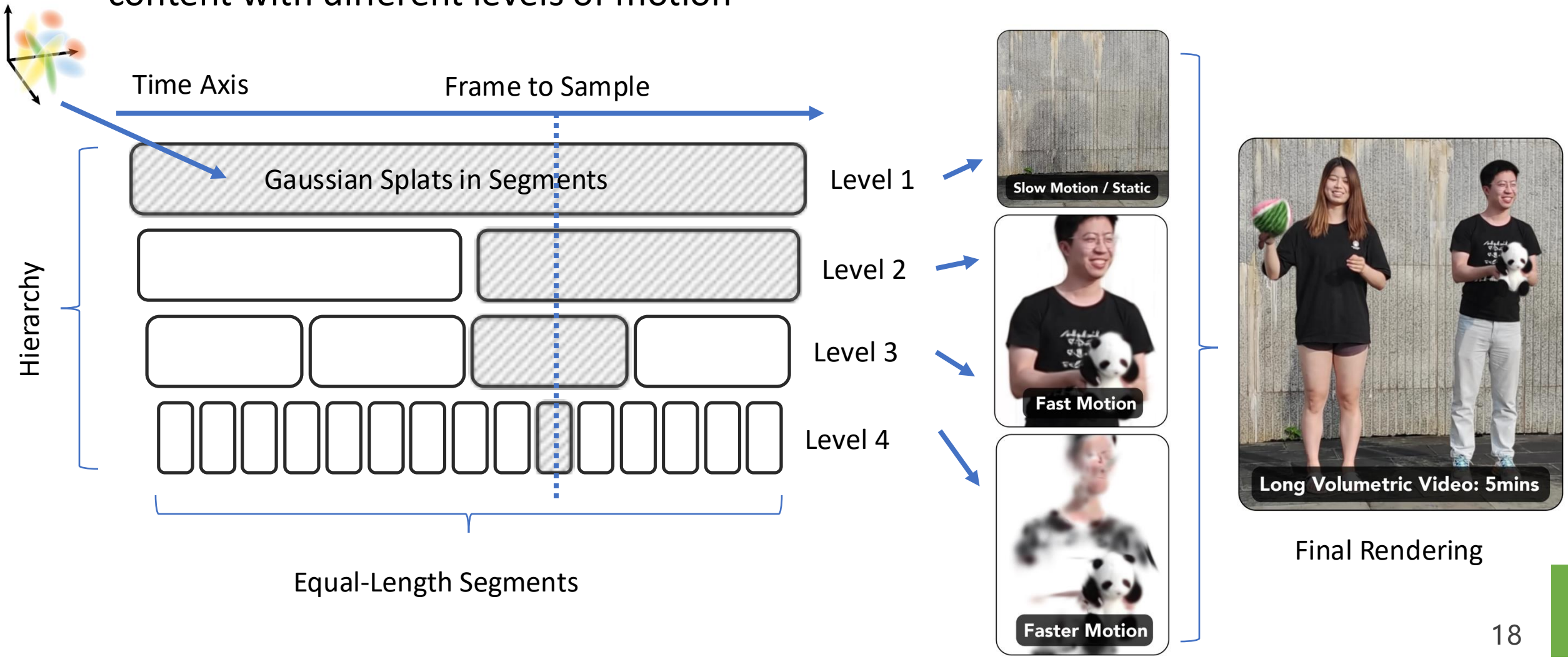
Representing long volumetric video



Representing Long Volumetric Video with Temporal Gaussian Hierarchy, SIGGRAPH Asia 2024 (Journal Track)

Temporal Gaussian Hierarchy

Building a hierarchical structure of Gaussian primitives to represent scene content with different levels of motion



Compact Appearance Model

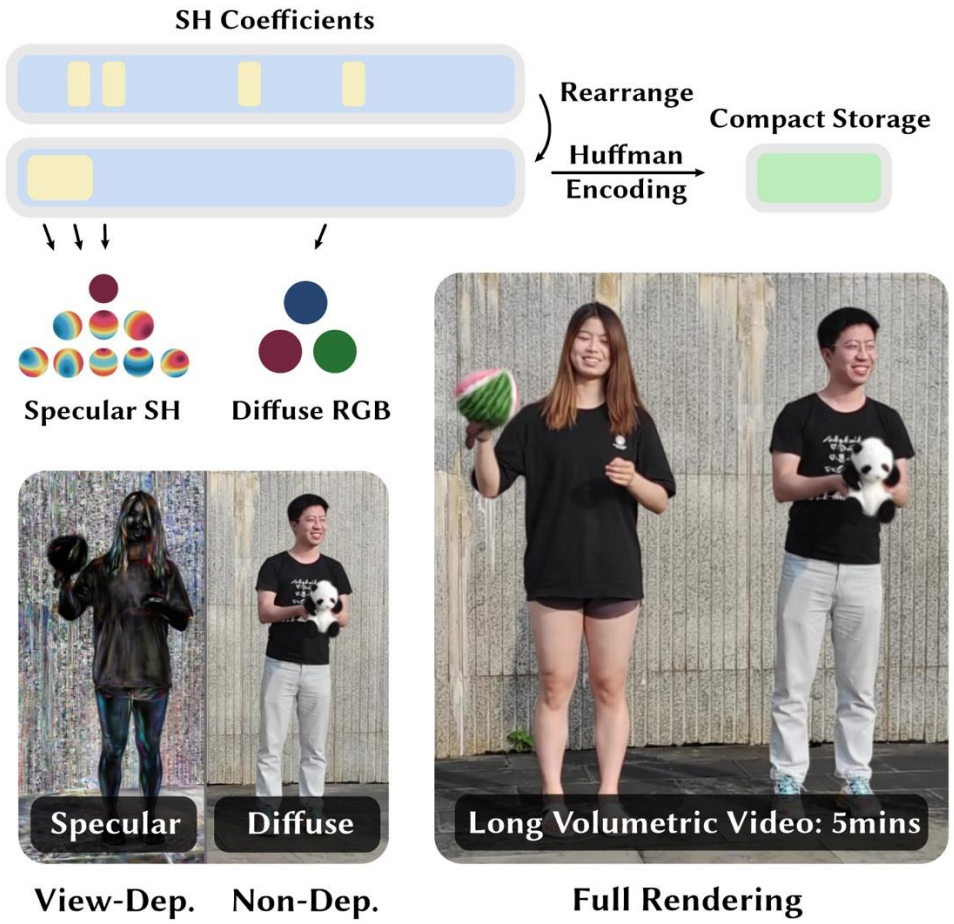
Insight: activating SH based on gradients -> compact model size

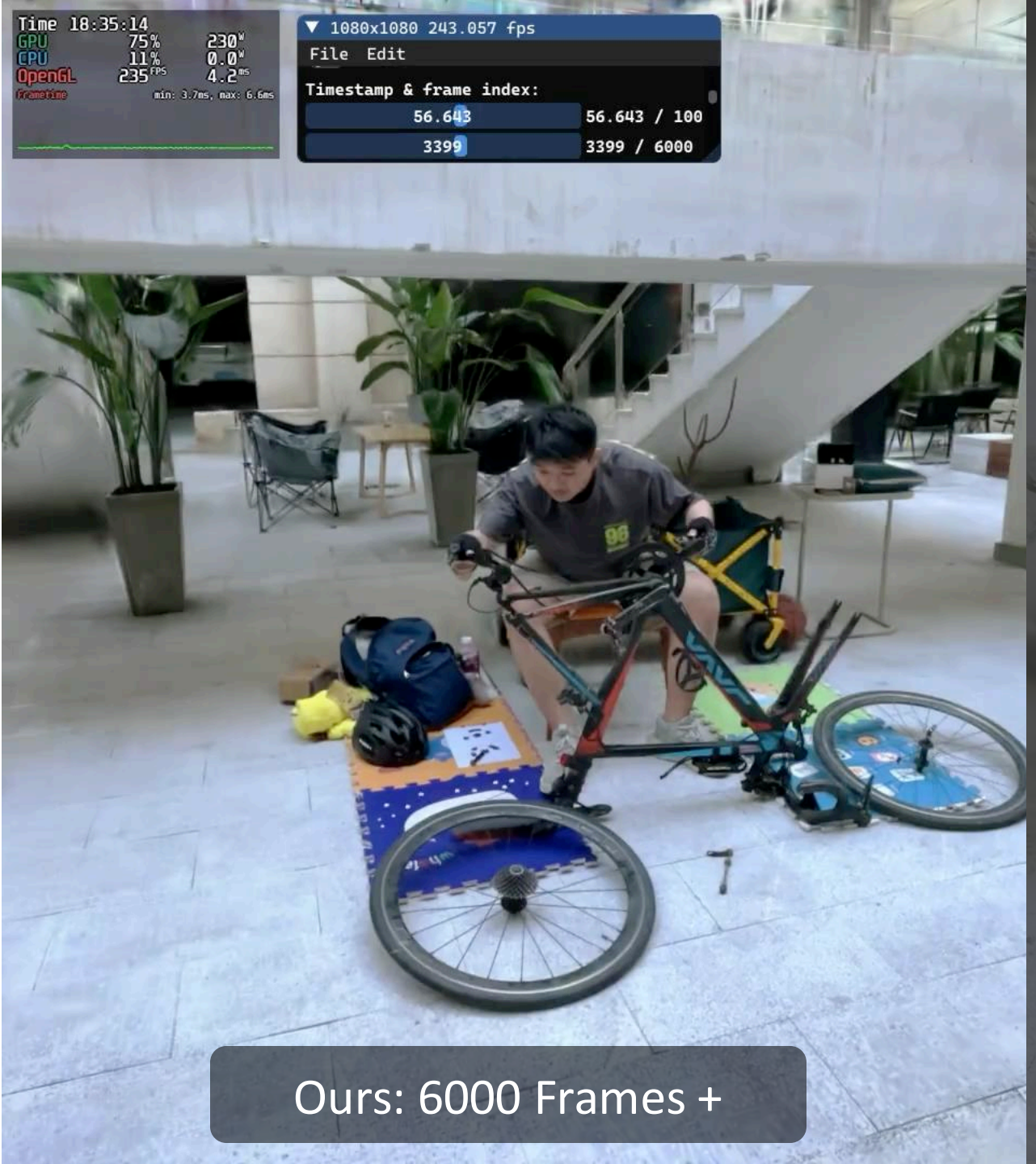
Spherical Harmonics

- Degree m
- Param count: $3x(m+1)^2$

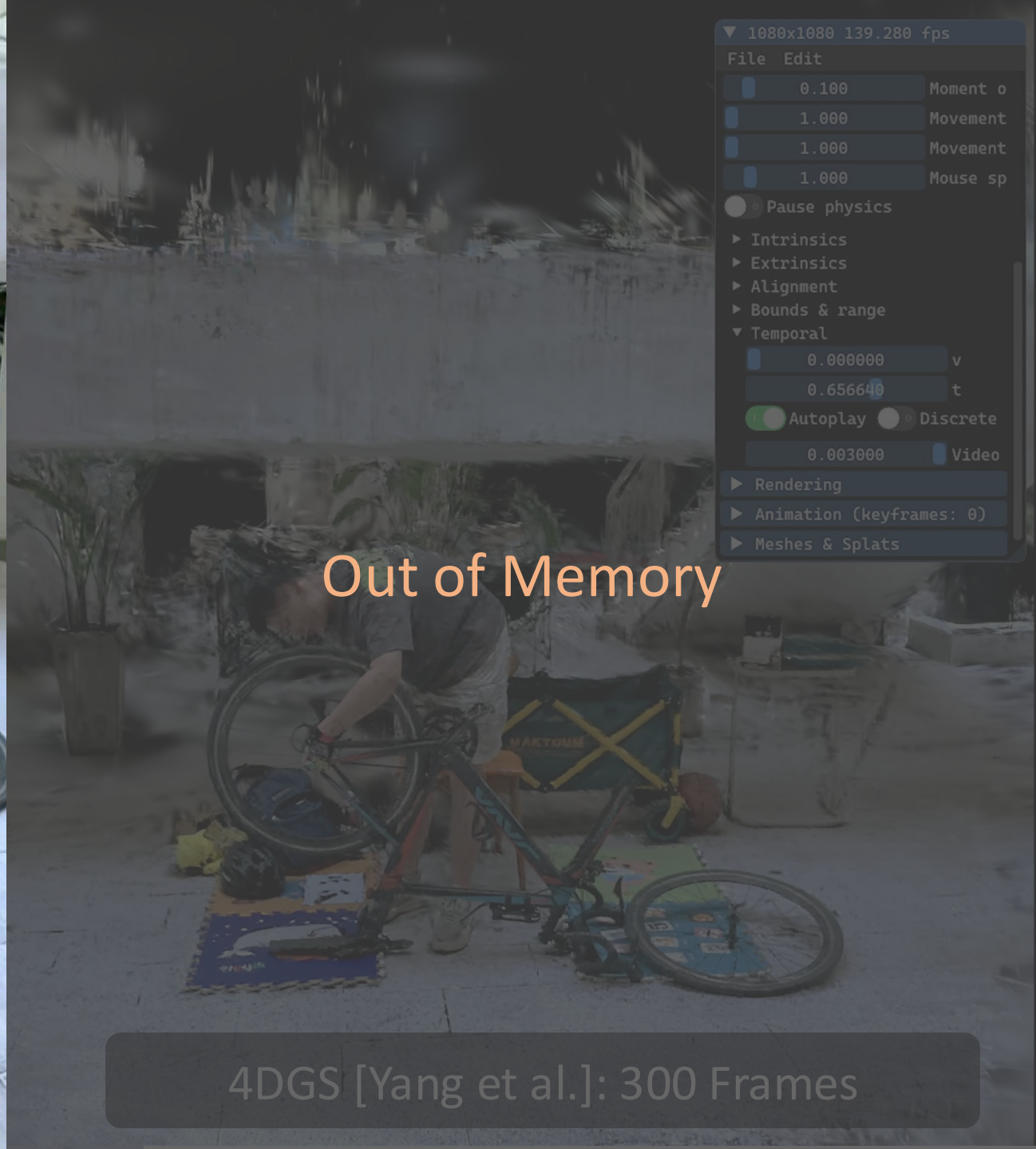
Optimization

- Compute gradient as if all SH params are enabled
- Only SH entries with gradient larger than threshold will be updated
- Other values are kept at 0





Ours: 6000 Frames +



Out of Memory

4DGS [Yang et al.]: 300 Frames


1920x1080 442.993 fps
File Edit
Player control
Play
0.000050
0.000
Playing time
Temporal Forest
fast_gauss
Render mode
linear
Scalar color
color
Visual mode (?)
Point mode


1920x1080 392.674 fps
File Edit
50.000
Video FPS
0
0 / 1000
0.000
0.000 / 20.0
Temporal Forest
fast_gauss
Render mode
linear
Scalar color
color
Visual mode (?)
Point mode


1920x1080 276.956 fps
File Edit
Autoplay
Discrete time
30.000
Video FPS
0
0 / 1600
0.000
0.000 / 53
Temporal Forest
fast_gauss
linear
color
Visual mode (?)
Point mode


1920x1080 277.940 fps
File Edit
Running on NVIDIA GeForce RTX 4090: 0
FPS : 277.940 FPS (?)
torch VRAM: 1885.25 MB (?)
frame time: 3.598 ms (?)
Record timing
Save timing
Unsync timing
Enable VSync
0.103825
Camera
Rendering
Temporal Forest Rendering
fast_gauss
Filter margintal_t
Filter frustum
0.050000
1.000000
1.000000
SIGGRAPH ASIA 2024 TOKYO 東京

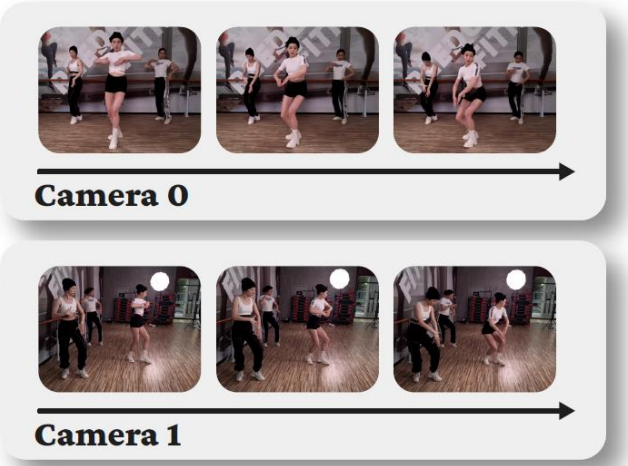
EasyVolcap

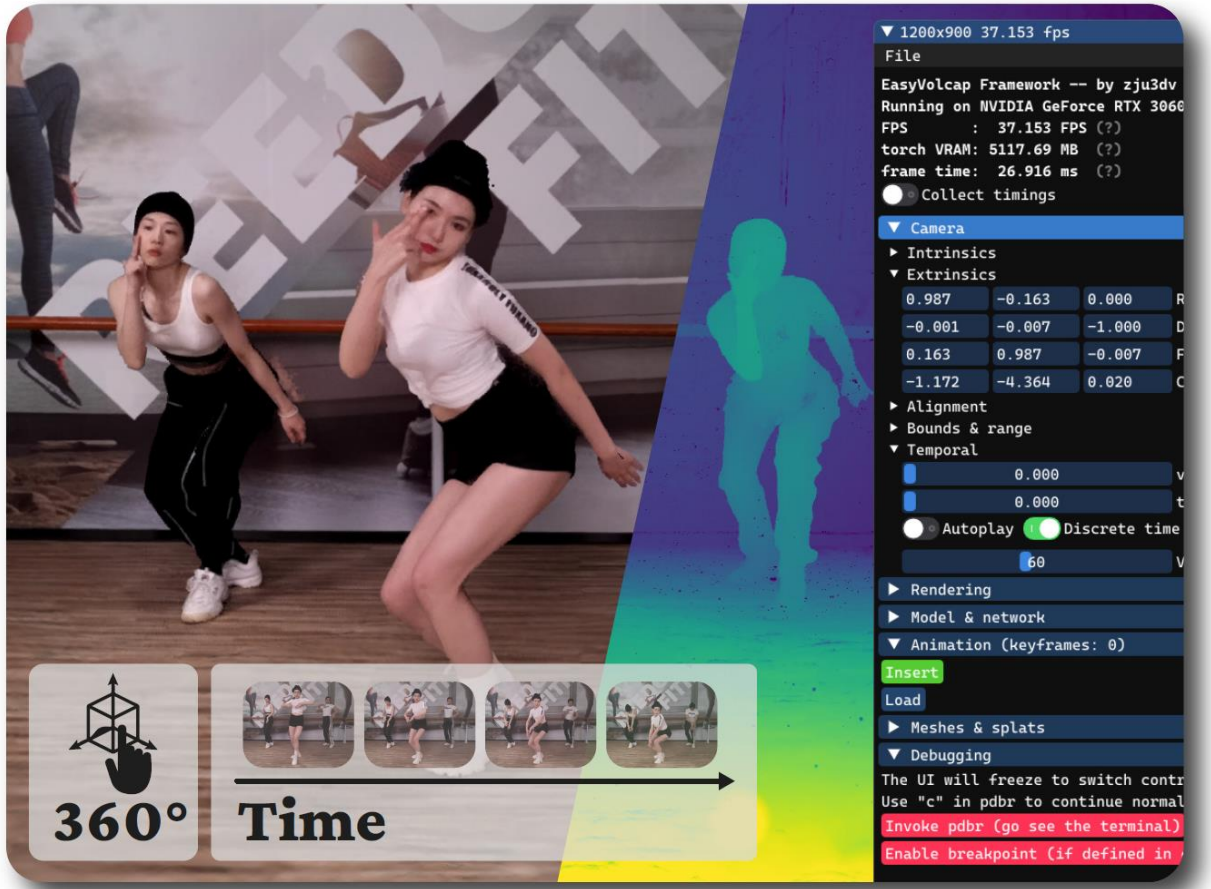
An open-source PyTorch library
to accelerate volumetric video research

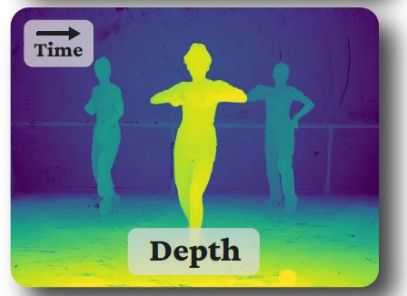
https://github.com/zju3dv/EasyVolcap

EasyVolcap Codebase

From Dataset to Training to Output



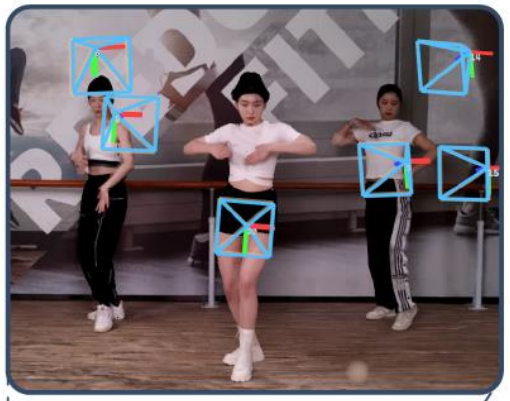




Input: Multi-View Video ⟶ **Unified & Extensible Framework** ⟶ **Output**

EasyVolcap Codebase

Modular Volumetric Video Algorithm Pipeline



Captured Videos

ray_o, ray_d
ray_t

Sampling Strategy

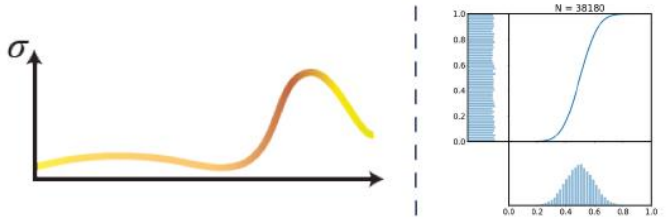
x,y,z,t,θ,φ

Space-Time Embedding & Decoding

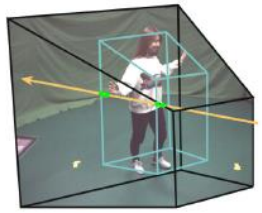
Appearance
Geometry

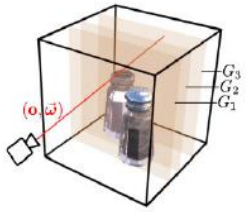
Volume Rendering

Sampling Strategy

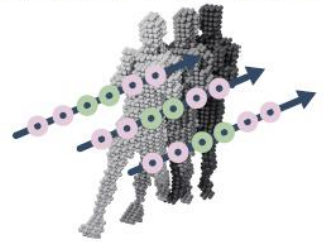


Ray Importance Sampling



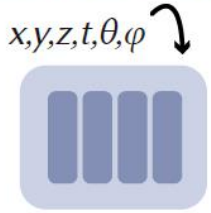


MVS/MLP Depth Guided Sampling

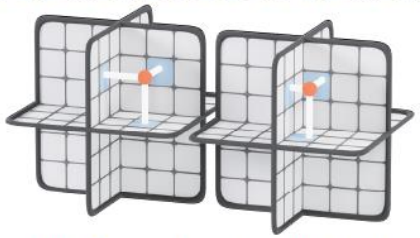


Explicit Geometry

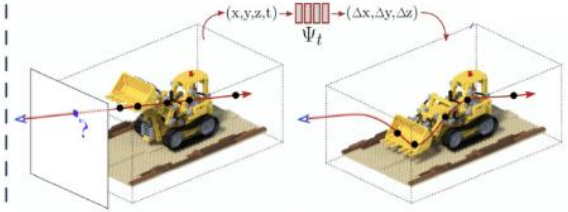
Space-Time Embedding & Decoding



MLP Networks



4D Coordinates & Grids



Deformation Fields

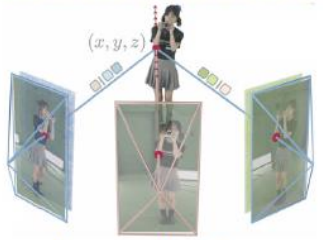
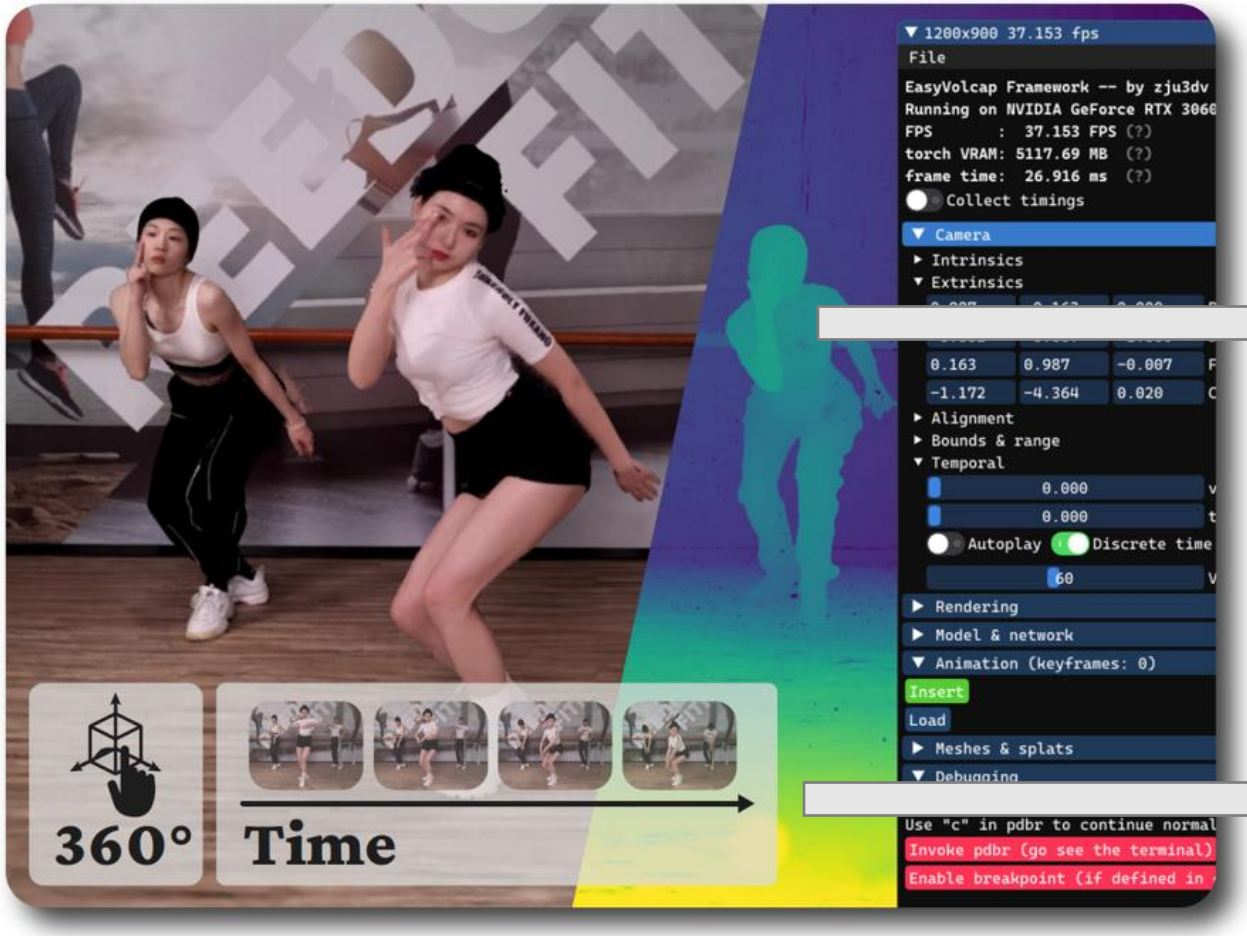


Image Features

EasyVolcap Codebase
High-performance Interactive Viewer



- Configurable IMGUI python components
- High-performance PyTorch-to-screen copy
- Cross-platform support by GLFW (Win & Linux)
- Camera path editor for animation control

Unsolved problems

- Reconstruction from sparse/single views
- Real-time reconstruction
- Relighting and editing
- Compression and streaming
- Generation

Thank you

https://github.com/zju3dv