DecoupledGaussian: Object-Scene Decoupling for Physics-Based Interaction

Presenter: Miaowei Wang

Motivation: Simulation from Reality

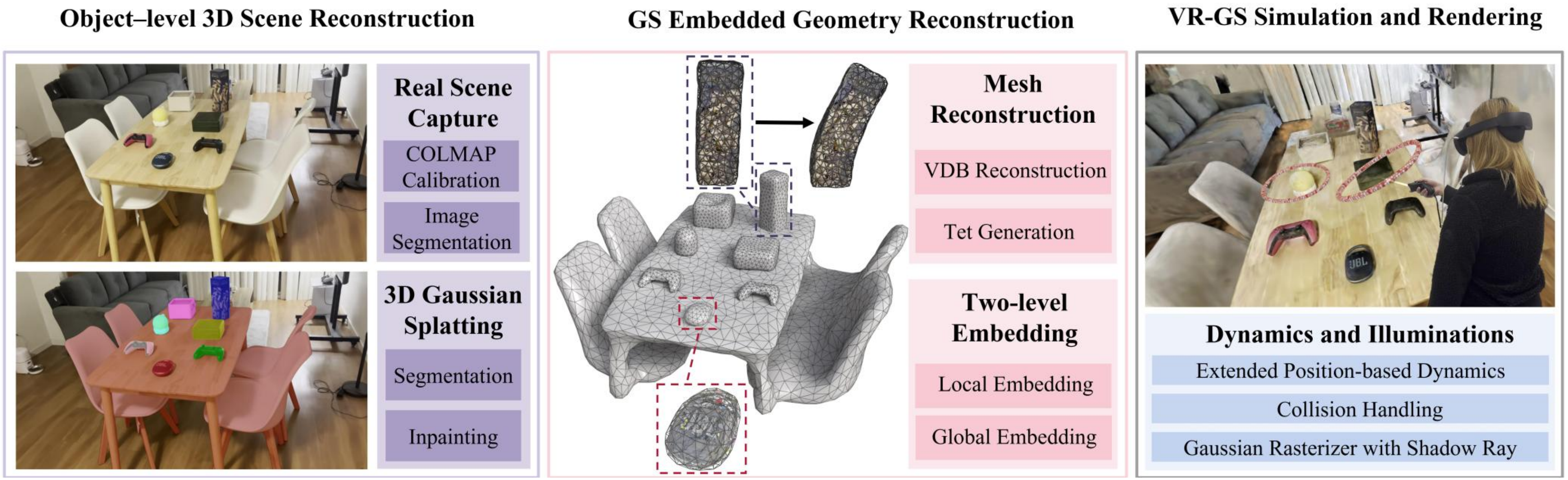
PhysGaussian: start the trend of simulation captured from real scenes

- Introduce GS Kinematics: rotate covariance matrix, spherical harmonic coefficients
- Insert Gaussian Splatting (GS) into the Material Point Method(MPM) simulation engine.

[Image Source: PhysGaussian CVPR2024 Highlight]

Motivation: Simulation from Reality

VR-GS deforms semantically segmented objects by SAM through XPBD simulation engine



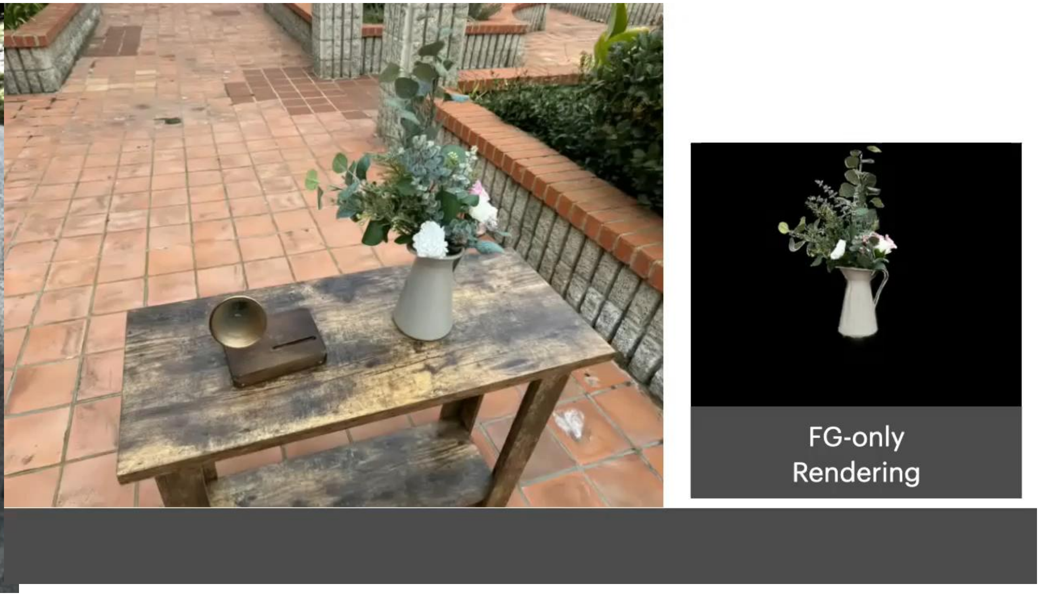
[Image Source: VR-GS Siggraph 2024 Conference]

Limitation: Object Depart from Surface

Current methods fail to simulate object motion departing from the contact surface.

PhysGaussian

VR-GS



FeatureSplatting (ECCV2024, Language-driven)

The core issue is the failure to recover occluded geometry and texture.

Our task: Decoupling

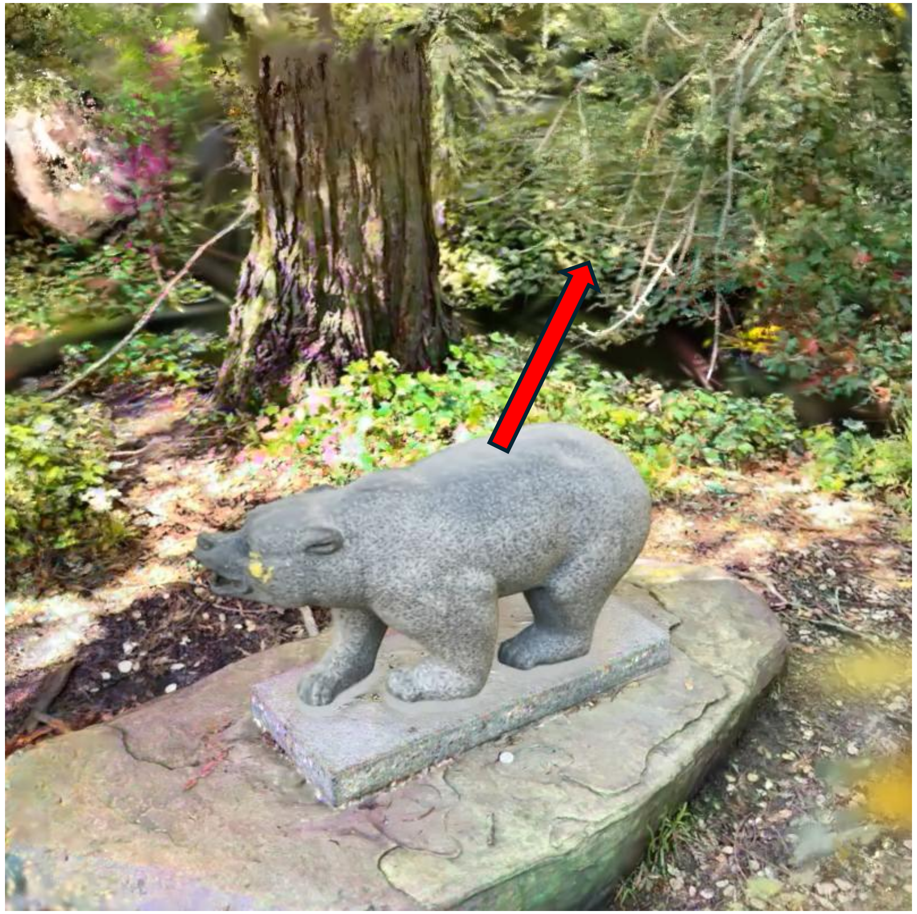
Input Video

Decoupling

Scene

Object

Support Various Interactive Simulation



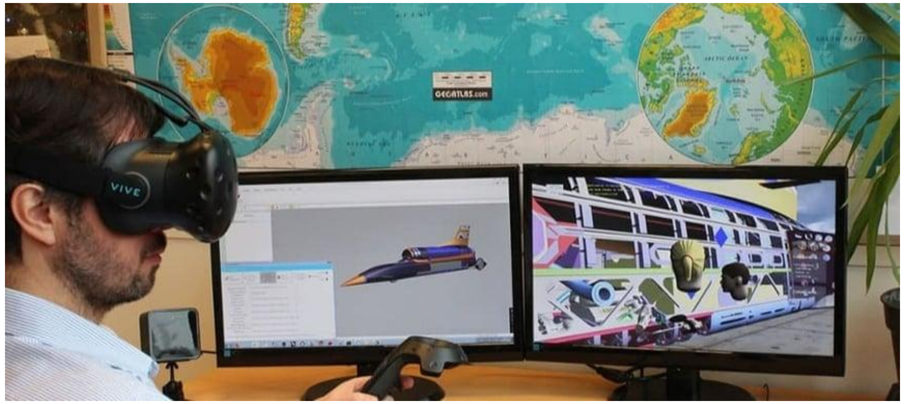
Scene Collisions

Object Melting

DecoupledGaussion: Meaning and Application

Autonomous driving Simulation

Entertainment industry



VR/AR Interaction

Robotic Learning Interaction

Our pipeline

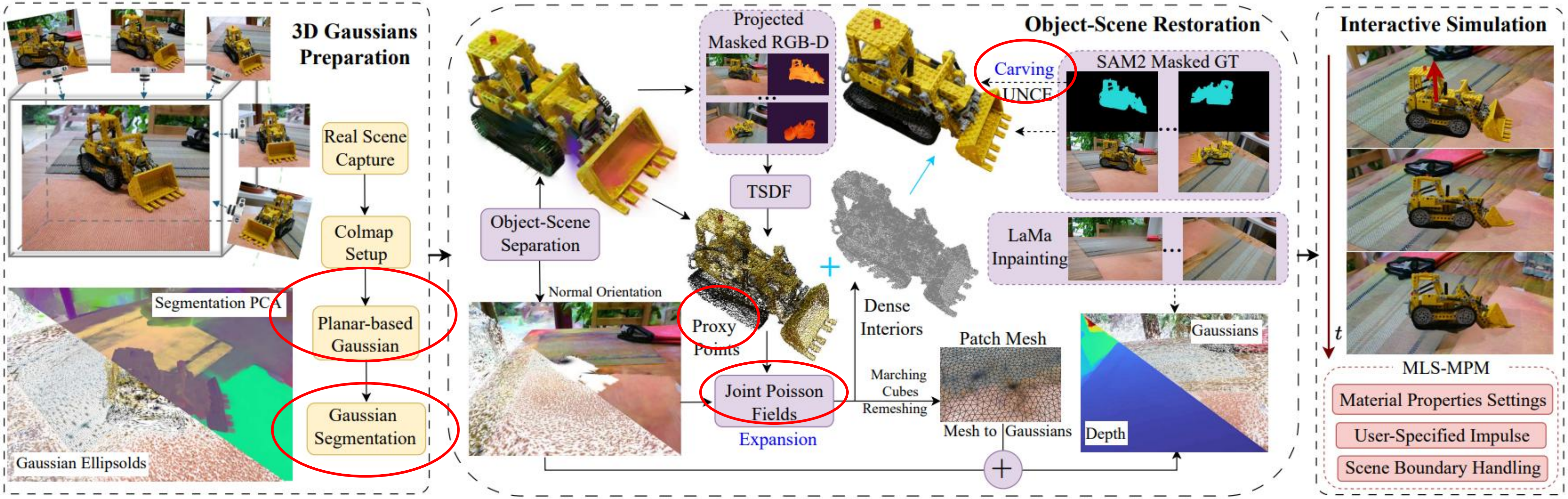


Figure 3. **System Overview**. DecoupledGaussian is an interactive simulation system that enables objects to detach from their initial contact surfaces after applying our proposed restoration pipeline, driven by user-specified impulses (red arrow on the right).

3D Gaussian Preparation: PGSR

The goal of using PGSR:
1) estimate correct geometry from GS
2) remove the large floaters in the scene

Planar-based Gaussian Splatting (PGSR):

-- Flatting the ellipsoid

-- Unbiased depth estimation

-- Single/Multi-view geometry regularization

-- Exposure compensation

Opacity is set to1, with×0.4 scaling
for better Gaussian kernel visualization.

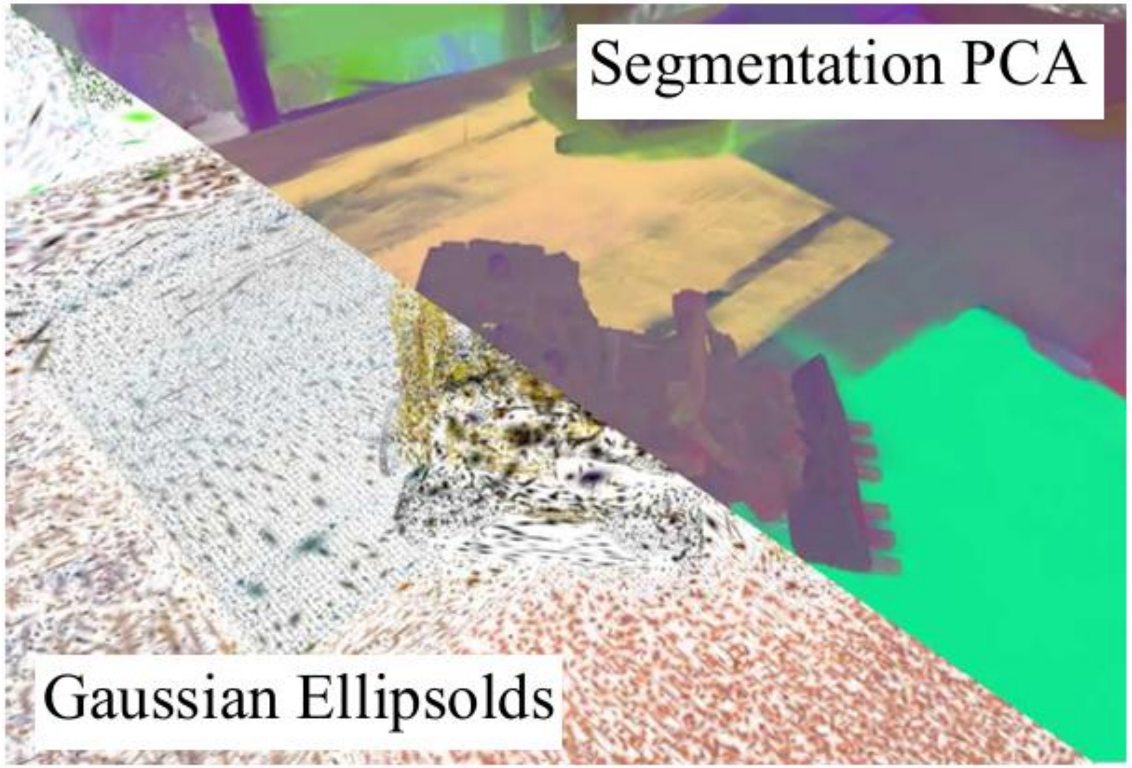
[PGSR: Planar-based Gaussian Splatting for Efficient and High-Fidelity Surface Reconstruction, TVCG2024]

3D Gaussian Preparation: Segmentation

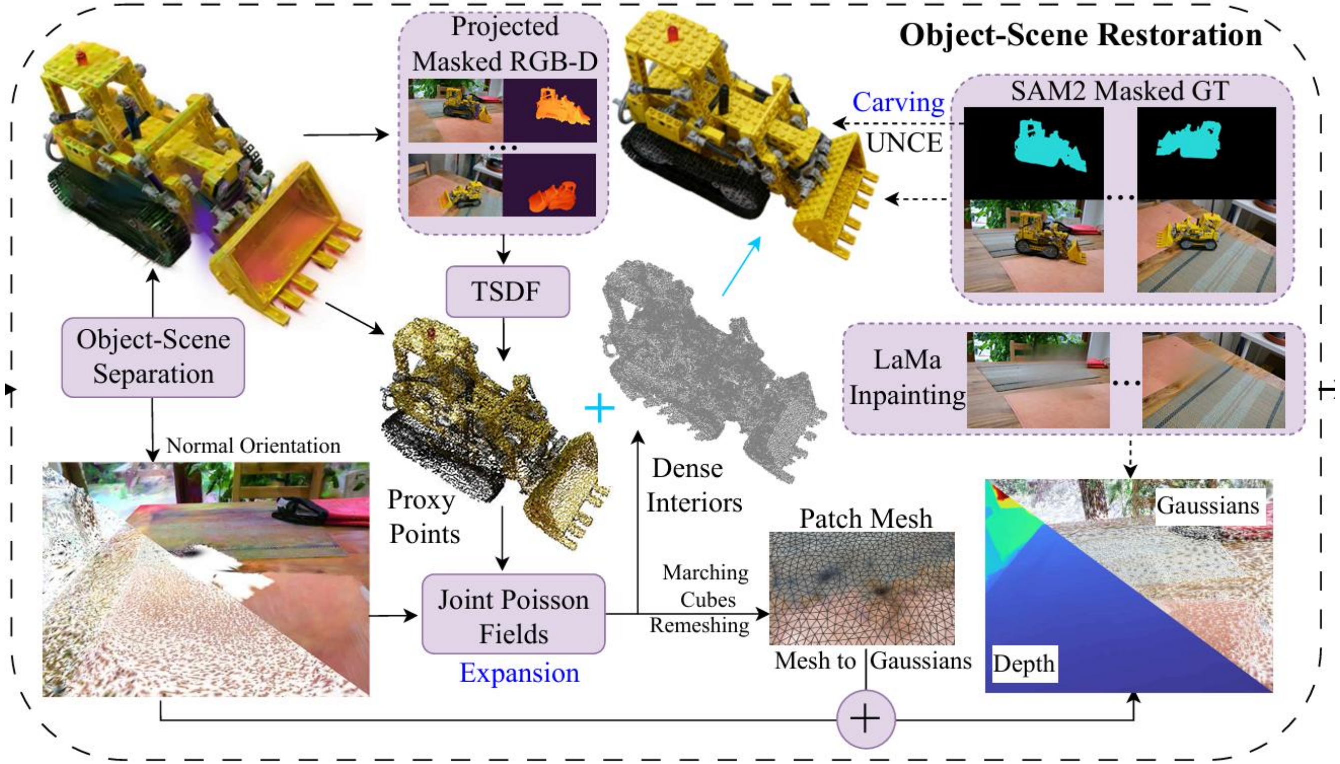
1. Select 2Dsegmentation labels from SAM2
2. Manually selecting classes in the first frame
3. Each kernel is assigned semantic features

3. A gating network to predict segmentation

4. Local feature smoothing.



Object-Scene Restoration



1. User-specified click position to get object (affinity)
2. Contact-surface use KNN to remove nearby artifacts
3. Restore Object and Contact scene

-- Joint Poisson Fields

-- Proxy Points

-- Unilateral Negative Cross Entropy

-- Gaussian Restoration

2D Inpainting Tools

Masked Image LaMa Photoroom

Figure 2. Inpainting tools (LaMa [79]; PhotoRoom [1]) introduce artifacts and inconsistent textures across frames.

(1) inpainted regions often fail to blend seam lessly with surrounding geometry, creating artifacts

(2) texture inconsistencies across frames due to the lack of robust video inpainting tools.

Our approach: prioritizing geometry restoration, leveraging *intrinsic GS geometry priors* to ensure a coherent surface even when texture inpainting is imperfect

Joint Poisson Fields

Geometry Prior: assuming both Object and Scene are *smooth, closed* shapes

We resort to Poisson Surface Reconstruction, equivalent to the winding number field

Joint Poisson Fields STEPS

1. Solving Screen Poisson Indicator for object and scene. Points&Normals
2. Coordinates transformation
3. Solving conflicted regions, by curvature and prioritizing surface
4. Extract ***dense interior points*** for objects while marching cubes to get cropped ***mesh patch*** for surface

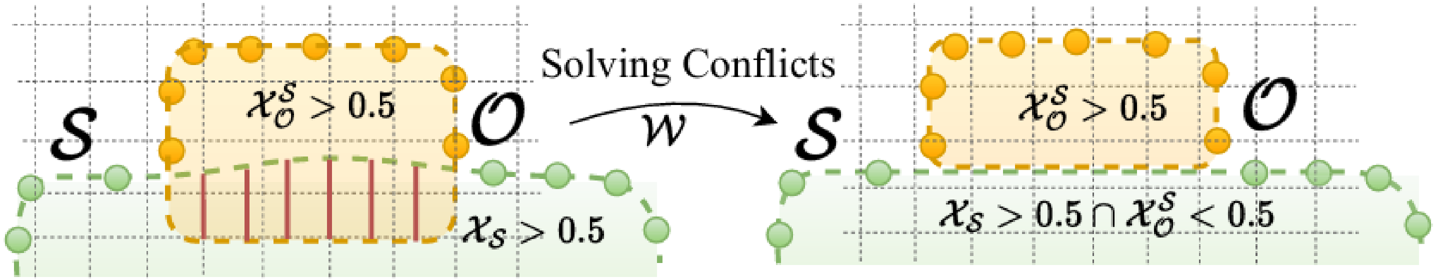
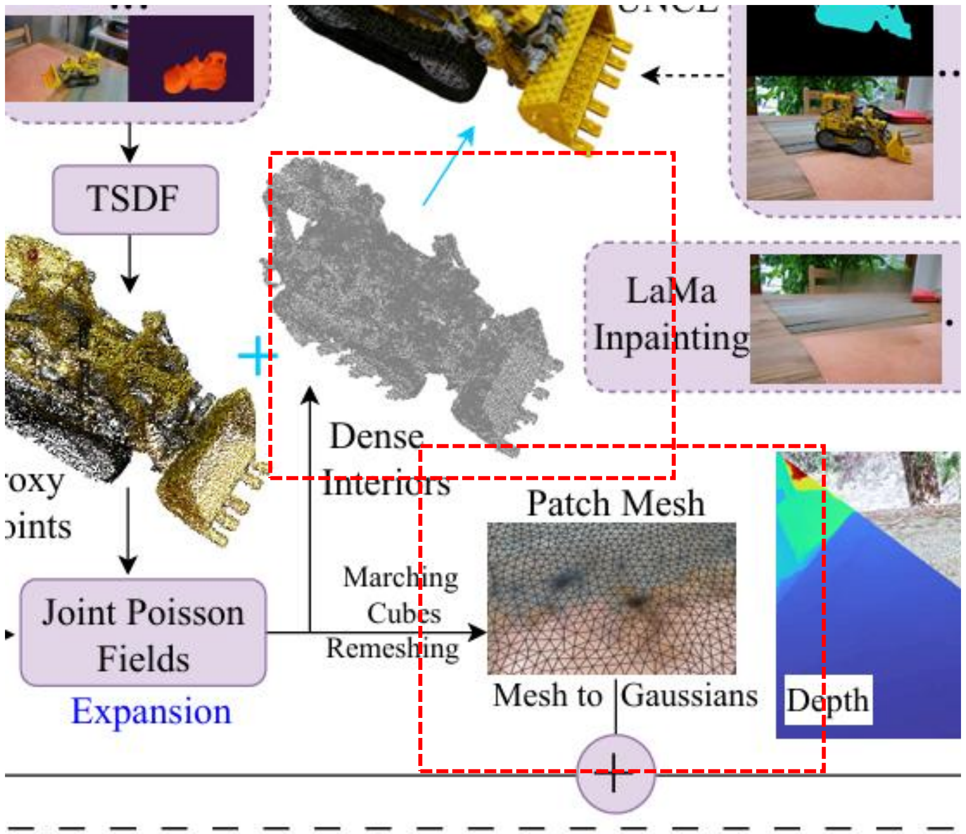


Figure 4. Joint Poisson Fields $\mathcal{W}$ first reconstruct $\mathcal{O}$ and $\mathcal{S}$ independently, then resolve conflicts (red area) by defining a boundary that separates them into distinct, non-intersecting entities.

[Adaptive multigrid solvers]

Joint Poisson Fields



Dense Points: helpful for MPM simulation

Patch Mesh: fix the broken geometry of the Surface

How about the input to Joint Poisson Fields:

- Contact Surface
  -- Gaussian Centers as points
  -- Minimum scaling as normal (90-degree disambiguate)

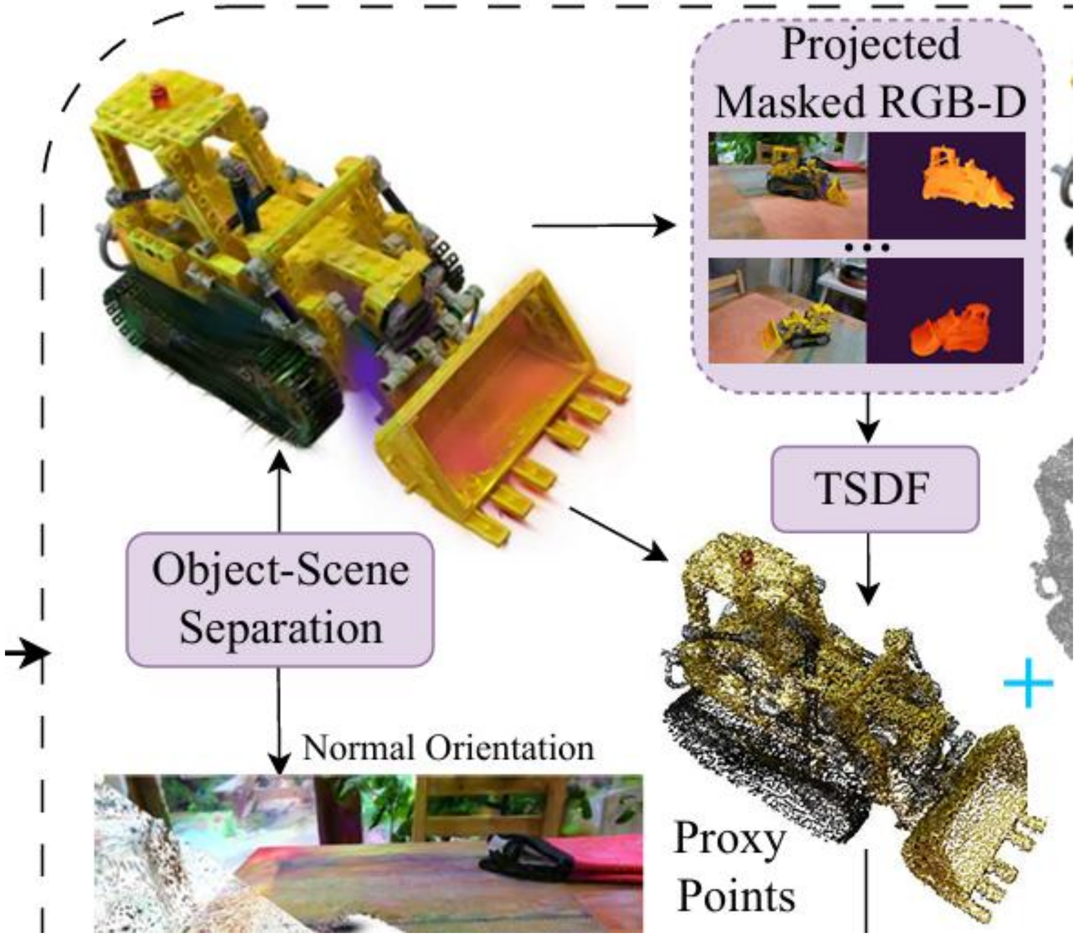
- Target Object
  -- Geometric complexity
  -- Use Proxy Points instead of Gaussian Centers

Proxy Points: Enhance Geometry Estimations

To get the proxy points, we use TSDF but different from others:

1. Projected depth **mask** to boost fusion speed
   Obtain the projected mask by setting zero/one opacity

2. Segment the final proxy points based on raw Gaussian kernel with nearest neighbor search



Proxy Points: Quick qualitative ablation

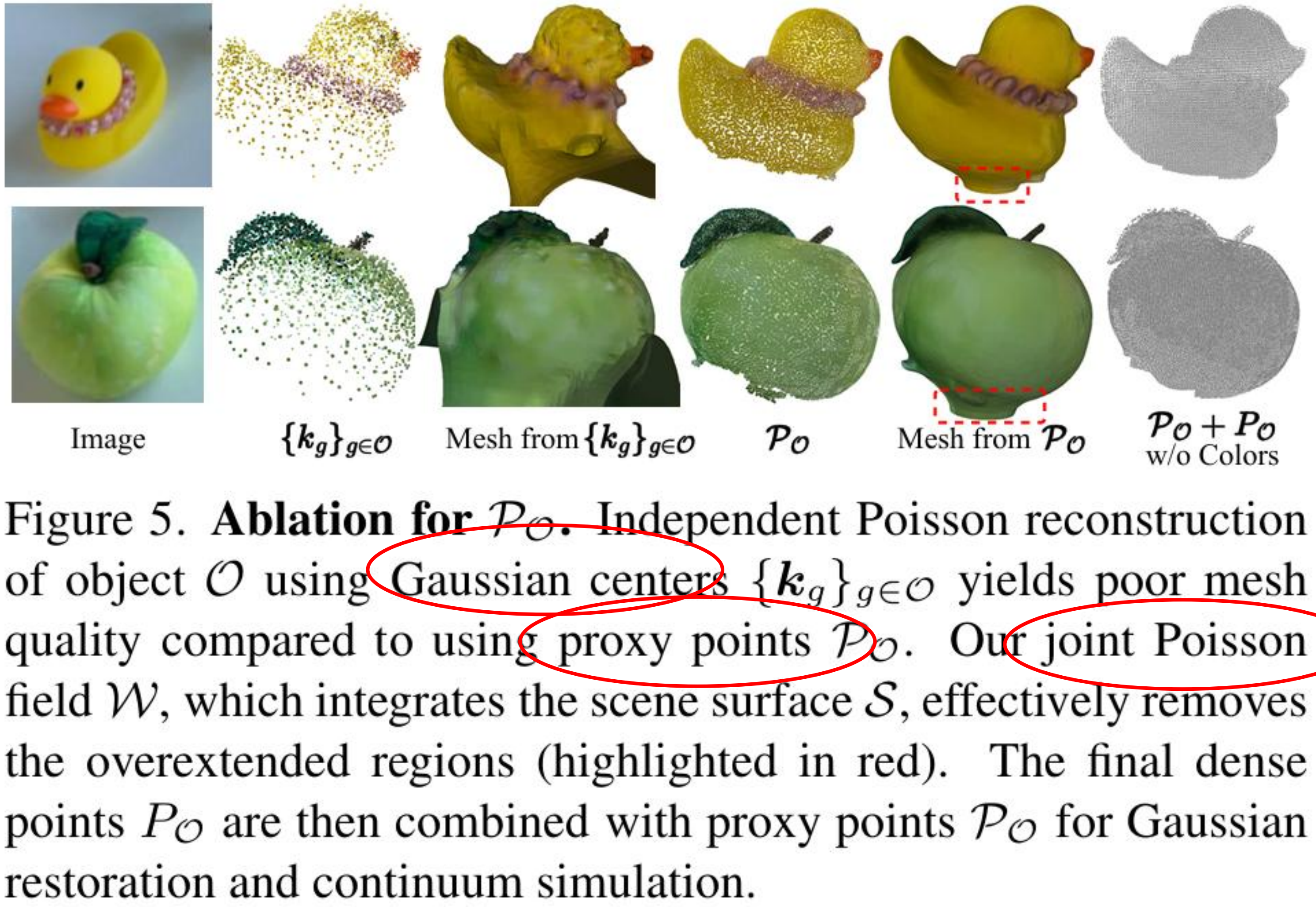


Figure 5. **Ablation for $\mathcal{P}_{\mathcal{O}}$.** Independent Poisson reconstruction of object $\mathcal{O}$ using Gaussian centers $\{\boldsymbol{k}_g\}_{g\in\mathcal{O}}$ yields poor mesh quality compared to using proxy points $\mathcal{P}_{\mathcal{O}}$. Our joint Poisson field $\mathcal{W}$, which integrates the scene surface $\mathcal{S}$, effectively removes the overextended regions (highlighted in red). The final dense points $P_{\mathcal{O}}$ are then combined with proxy points $\mathcal{P}_{\mathcal{O}}$ for Gaussian restoration and continuum simulation.

Unilateral Negative Cross Entropy

- Geometry Expansion from Poisson:

  -- Over-smooth close surface

  -- Influence by noises or outliers

  -- Introducing particles beyond observable viewpoints

  -- Points Converted to GS, we will get like:

Opacity is set to one for TRUCK to highlight artifacts

Unilateral Negative Cross Entropy

Multi-view carving:

- Silhouette consistency between rendered and GT
- Discrepancy between blended opacity and 2D GT object mask

$$\mathrm{UNCE}(p) = -(1 - \underset{\substack{\uparrow\\ \text{2D GT object mask}}}{M_{\mathcal{O}}^{\mathrm{GT}}(p)}) \log(1 - \underset{\substack{\uparrow\\ \text{Blended opacity}}}{\mathbb{1}_{\mathcal{O}}(p)}).$$

- Every 100 iterations, we clean Gaussians with low opacity

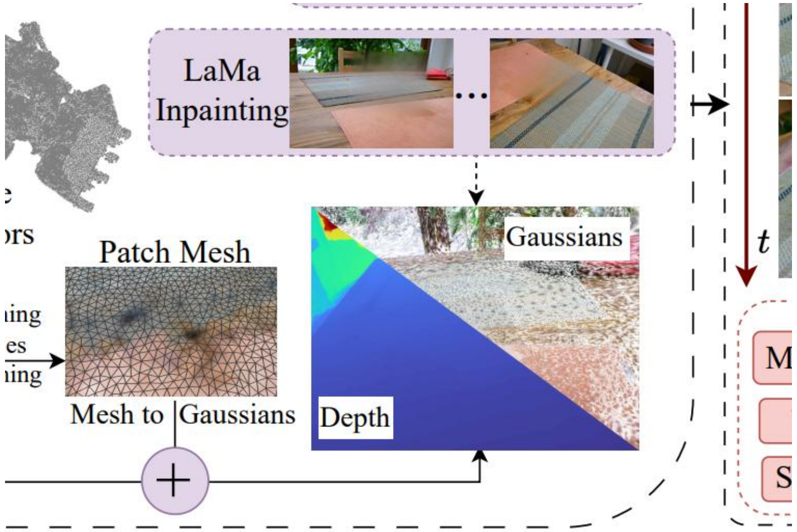
w/o UNCE

Ours

Gaussian Restoration

- Object, Points to Isometry Gaussian, during finetune:
  -- UNCE multi-carve geometry artifacts
  -- Fine-tune textures based on GT-masked train images.

- Contact Surface: fixed by patch mesh
  -- From mesh to new 3D flattened Gaussian
  -- 2D Tools to adjust the texture properties



Interactive Simulation

- User-specified force as an impulse
- Manually specified material properties
- Setting boundary conditions to bounce back
- Estimate Normals of the Ground plane by RANSAC to set up gravity.

……

Evaluation

Object-Scene Interaction Simulation

Object Restoration

Scene Restoration

Qualitative Evaluation: Object Restoration

Segmented Gaussians

GIC [NIPS 2024]

PhysGaussian [CVPR 2024]

Ours

GIC and PhysGaussian don't account for incomplete object surfaces during simulation, unlike our approach.

Qualitative Evaluation: Scene Restoration


GScream [ECCV 2024]


VR-GS [SIGGRAPH 2024]


Ours

since we don't rely on 2D inpainting tools for geometry restoration, we avoid the catastrophic errors commonly seen in Gscream or VR-GS.

Object-Scene Interactive Simulation

VR-GS (Scene) +
PhysGaussian (Object)

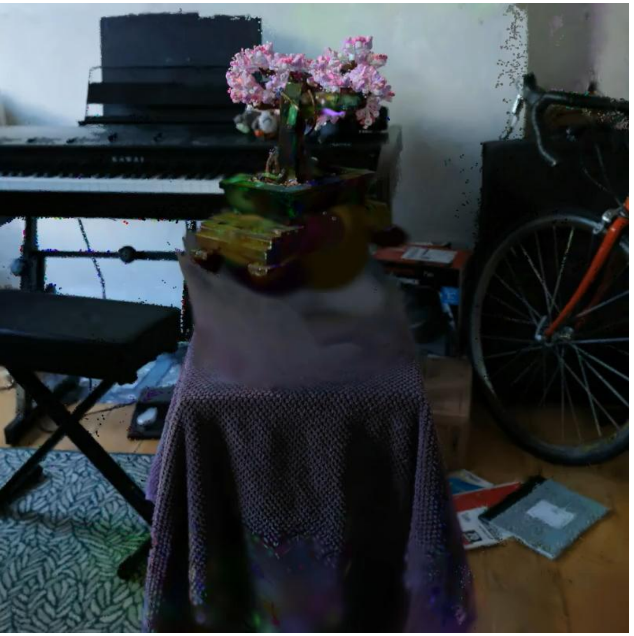
Ours (Scene) +
PhysGaussian (Object)

Ours (Scene) +
GIC (Object)

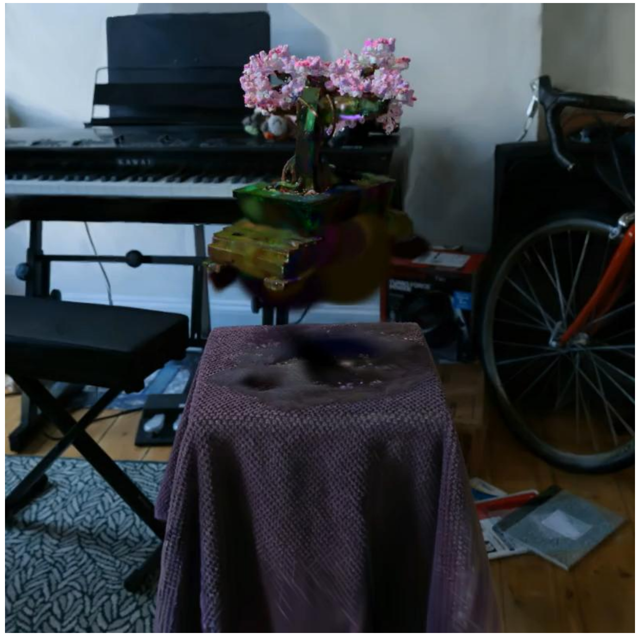
Ours (Scene) +
Ours (Object)

Correct geometry recovery of both the scene and object, as in our method, is crucial for the simulation.

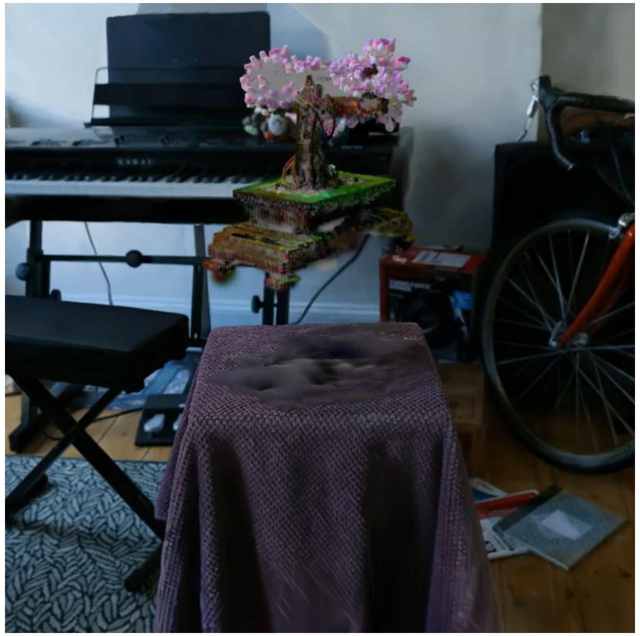
Object-Scene Interactive Simulation



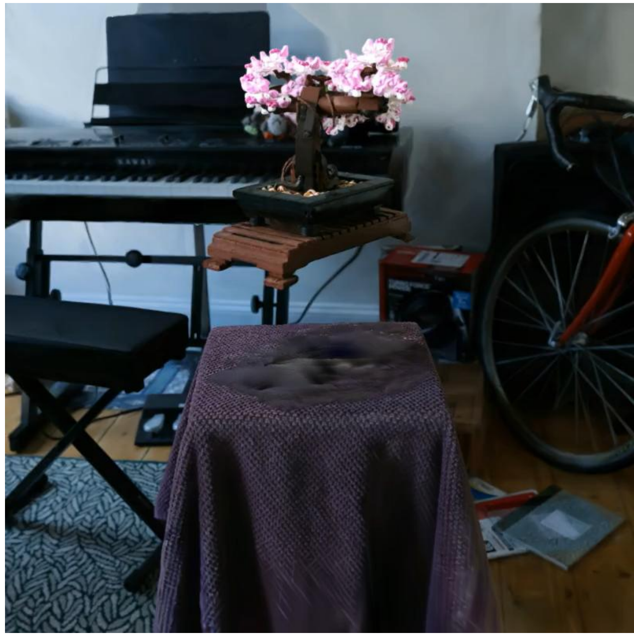
VR-GS (Scene) +
PhysGaussian (Object)



Ours (Scene) +
PhysGaussian (Object)



Ours (Scene) +
GIC (Object)



Ours (Scene) +
Ours (Object)

VR-GS fails to restore the surface of the contact scene, causing the bonsai to get stuck.

Object-Scene Interactive Simulation

VR-GS (Scene) +
PhysGaussian (Object)

Ours (Scene) +
PhysGaussian (Object)

Ours (Scene) +
GIC (Object)

Ours (Scene) +
Ours (Object)

Also, take note of the table in VR-GS. While it appears correct, it still traps the object.

Object-Scene Interactive Simulation

VR-GS (Scene) +
PhysGaussian (Object)

Ours (Scene) +
PhysGaussian (Object)

Ours (Scene) +
GIC (Object)

Ours (Scene) +
Ours (Object)

Of course, our method can also be applied to multiple objects.

Object Restoration

Segmented Gaussians

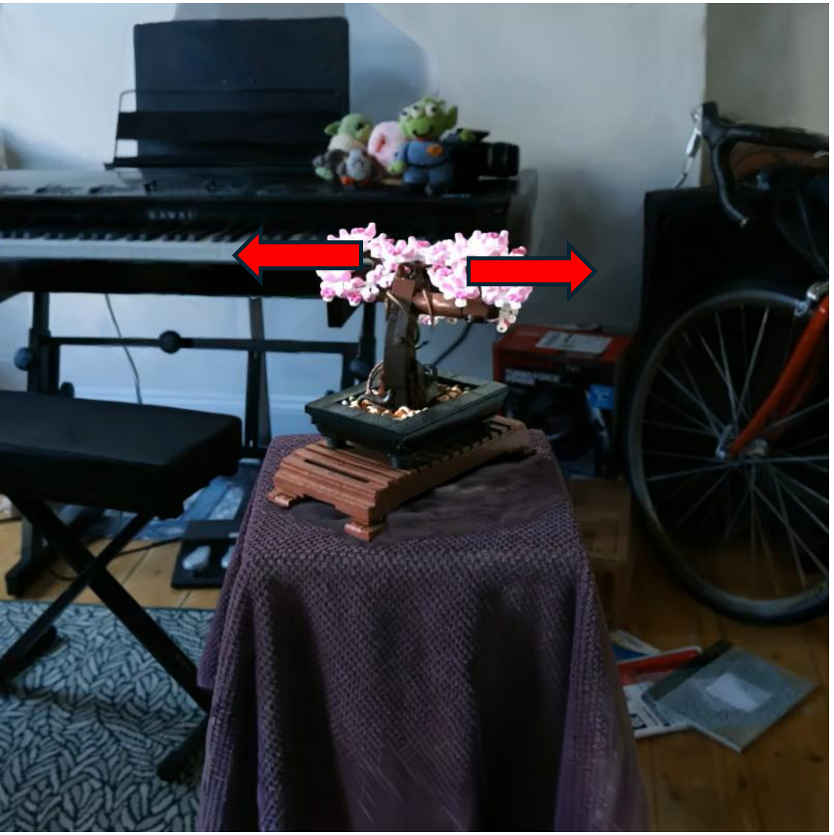
GIC [NIPS 2024]

PhysGaussian [CVPR 2024]

Ours

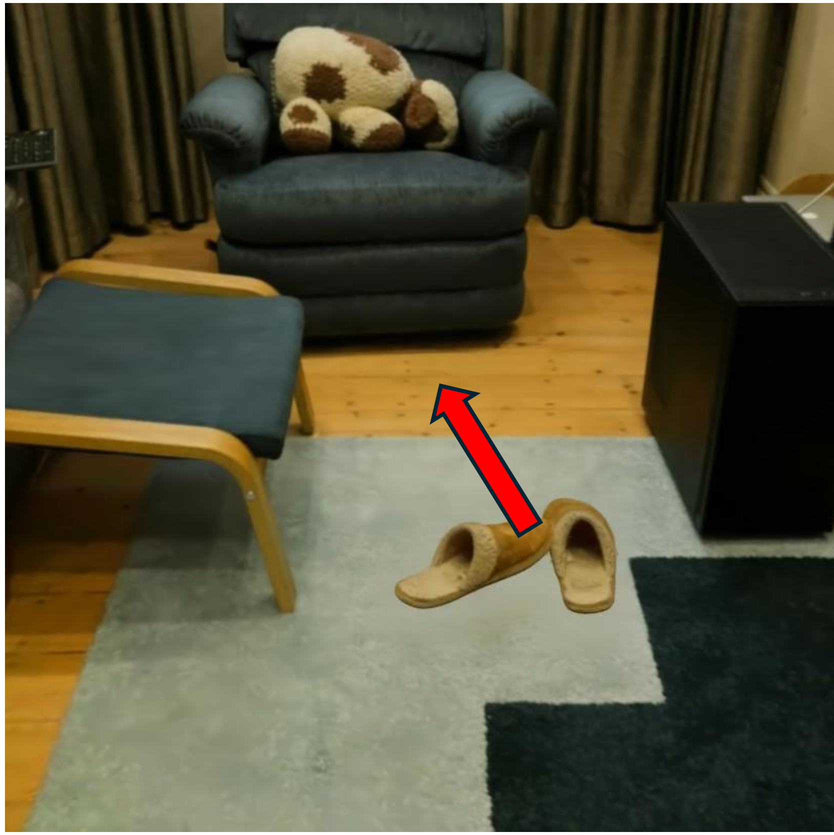
Our method recovers hidden object regions better than others, with slight issues in consistency and high-frequency textures.

More Cases (Ours)



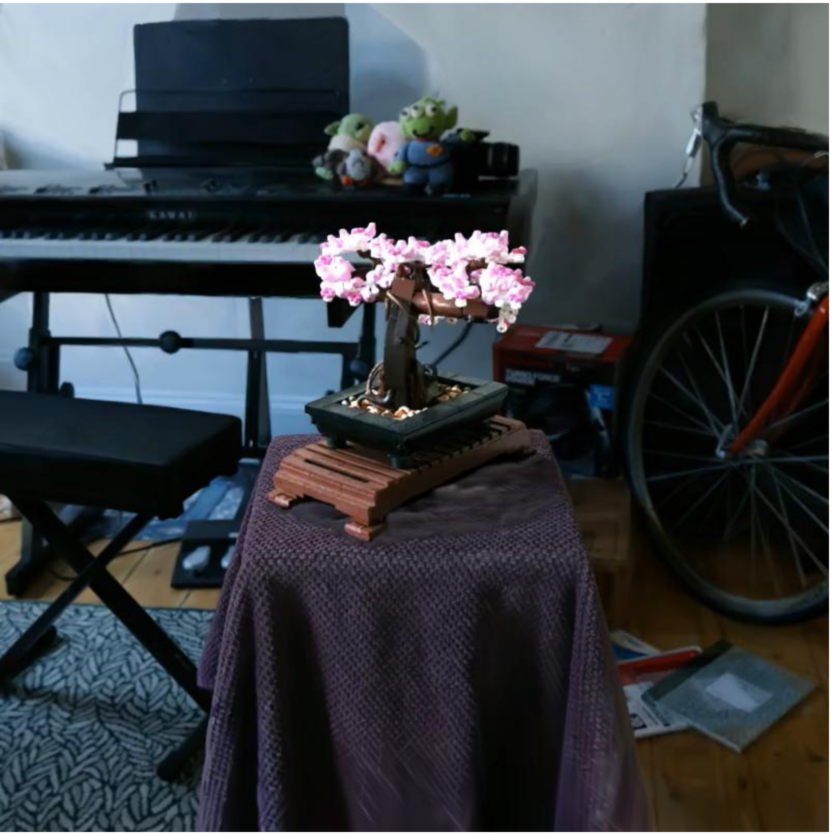
Bonsai

Kitchen



Room

More Cases (Ours)



Bonsai

Kitchen

Room

Interactions Across Diverse Scenes

Input Video

Truck in Bicycle

Quantitative Evaluation

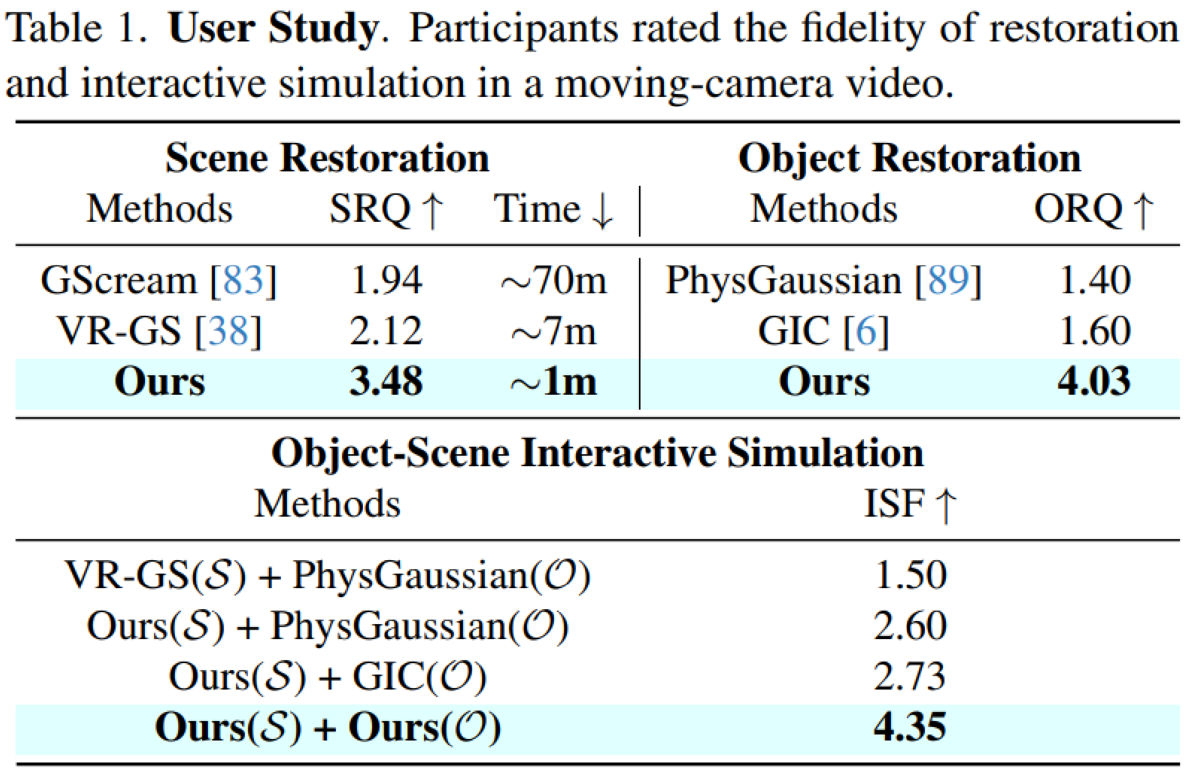


Table 1. **User Study**. Participants rated the fidelity of restoration and interactive simulation in a moving-camera video.

Scene Restoration			**Object Restoration**	
Methods	SRQ ↑	Time ↓	Methods	ORQ ↑
GScream [83]	1.94	∼70m	PhysGaussian [89]	1.40
VR-GS [38]	2.12	∼7m	GIC [6]	1.60
Ours	**3.48**	**∼1m**	**Ours**	**4.03**

Object-Scene Interactive Simulation	
Methods	ISF ↑
VR-GS($\mathcal{S}$) + PhysGaussian($\mathcal{O}$)	1.50
Ours($\mathcal{S}$) + PhysGaussian($\mathcal{O}$)	2.60
Ours($\mathcal{S}$) + GIC($\mathcal{O}$)	2.73
Ours($\mathcal{S}$) + Ours($\mathcal{O}$)	**4.35**

SQR: Scene Restoration Quality

OQR: Object Restoration Quality

ISF: Interactive Simulation Fidelity

Quantitative Evaluation

Composite complete real scene and real object

Placed by Rigid simulation PyBullet

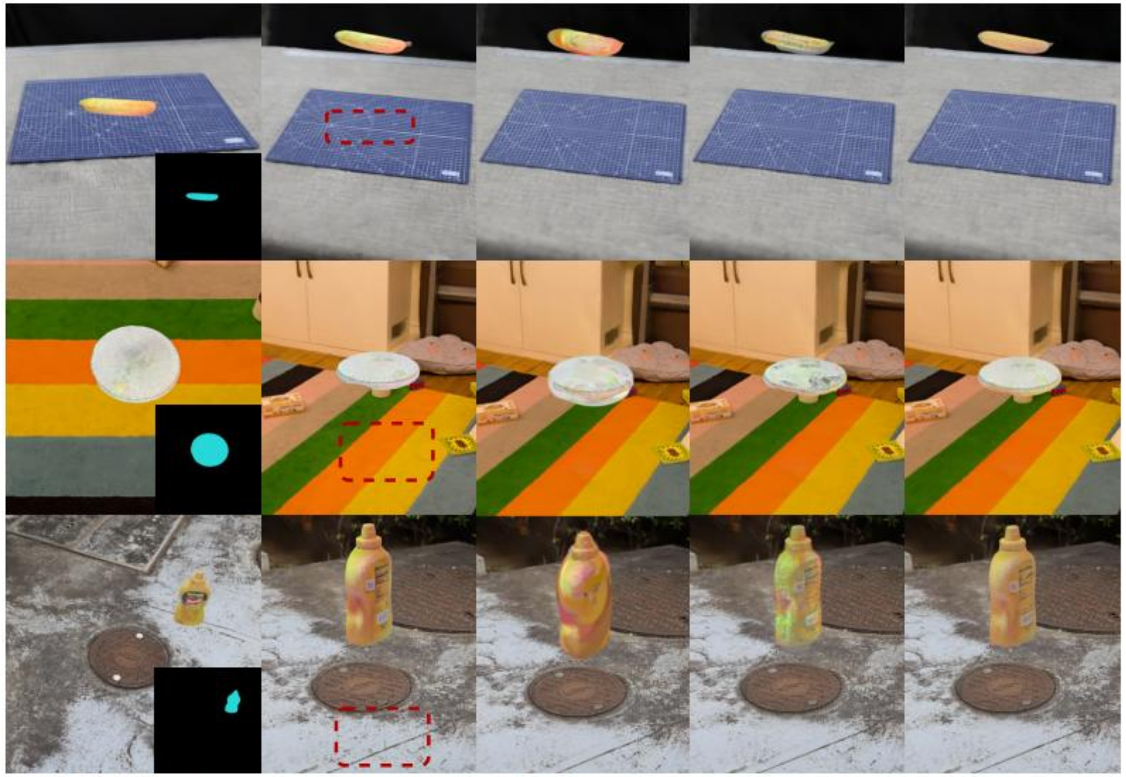
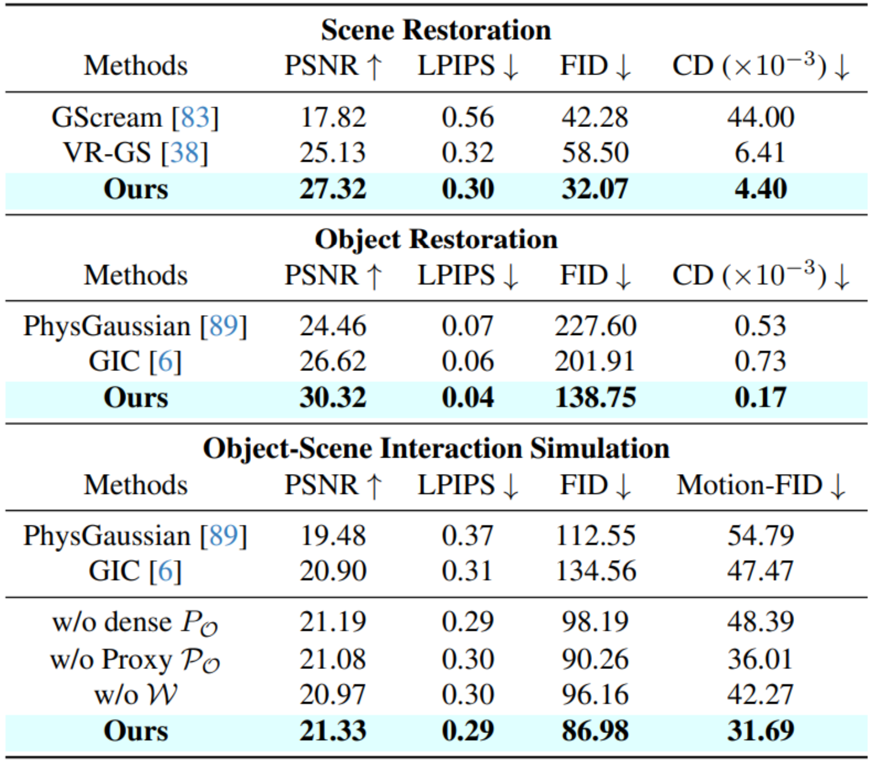


Figure 8. **Benchmark Comparisons.** A test viewpoint visualizes comparisons using the restored scene from our method, with inpainting regions marked by a red rectangle in the Ground Truth.

Table 2. **Quantitative Comparisons & Ablations.** We create a decoupling benchmark with comprehensive metrics comparing baselines and ablations to validate design choices.



	Scene Restoration			
Methods	PSNR ↑	LPIPS ↓	FID ↓	CD ($\times 10^{-3}$) ↓
GScream [83]	17.82	0.56	42.28	44.00
VR-GS [38]	25.13	0.32	58.50	6.41
Ours	**27.32**	**0.30**	**32.07**	**4.40**
	Object Restoration			
Methods	PSNR ↑	LPIPS ↓	FID ↓	CD ($\times 10^{-3}$) ↓
PhysGaussian [89]	24.46	0.07	227.60	0.53
GIC [6]	26.62	0.06	201.91	0.73
Ours	**30.32**	**0.04**	**138.75**	**0.17**
	Object-Scene Interaction Simulation			
Methods	PSNR ↑	LPIPS ↓	FID ↓	Motion-FID ↓
PhysGaussian [89]	19.48	0.37	112.55	54.79
GIC [6]	20.90	0.31	134.56	47.47
w/o dense $P_{\mathcal{O}}$	21.19	0.29	98.19	48.39
w/o Proxy $\mathcal{P}_{\mathcal{O}}$	21.08	0.30	90.26	36.01
w/o $\mathcal{W}$	20.97	0.30	96.16	42.27
Ours	**21.33**	**0.29**	**86.98**	**31.69**

Geometry accuracy using CD
Motion accuracy use Motion-FID

Ablations

Dense Points prevent collapse under gravity, while Joint Point Fields eliminate intersection regions.

Ours (Scene) +
Ground Truth (Object)

Ours (Scene) +
Ours (Object)

Ours (Scene) +
w/o Dense Points

Ours (Scene) +
w/o Joint Poisson Fields

Ablations

Note: Opacity is set to one for TRUCK to highlight artifacts

Ours (Scene) +
Ours (Object)

Ours (Scene) +
w/o UNCE (Object)

Take away

- Restore the object, and contact scene to simulate.
- Not rely on 2D impainting models to repair broken geometry.

Scan to our project website

Thank you!