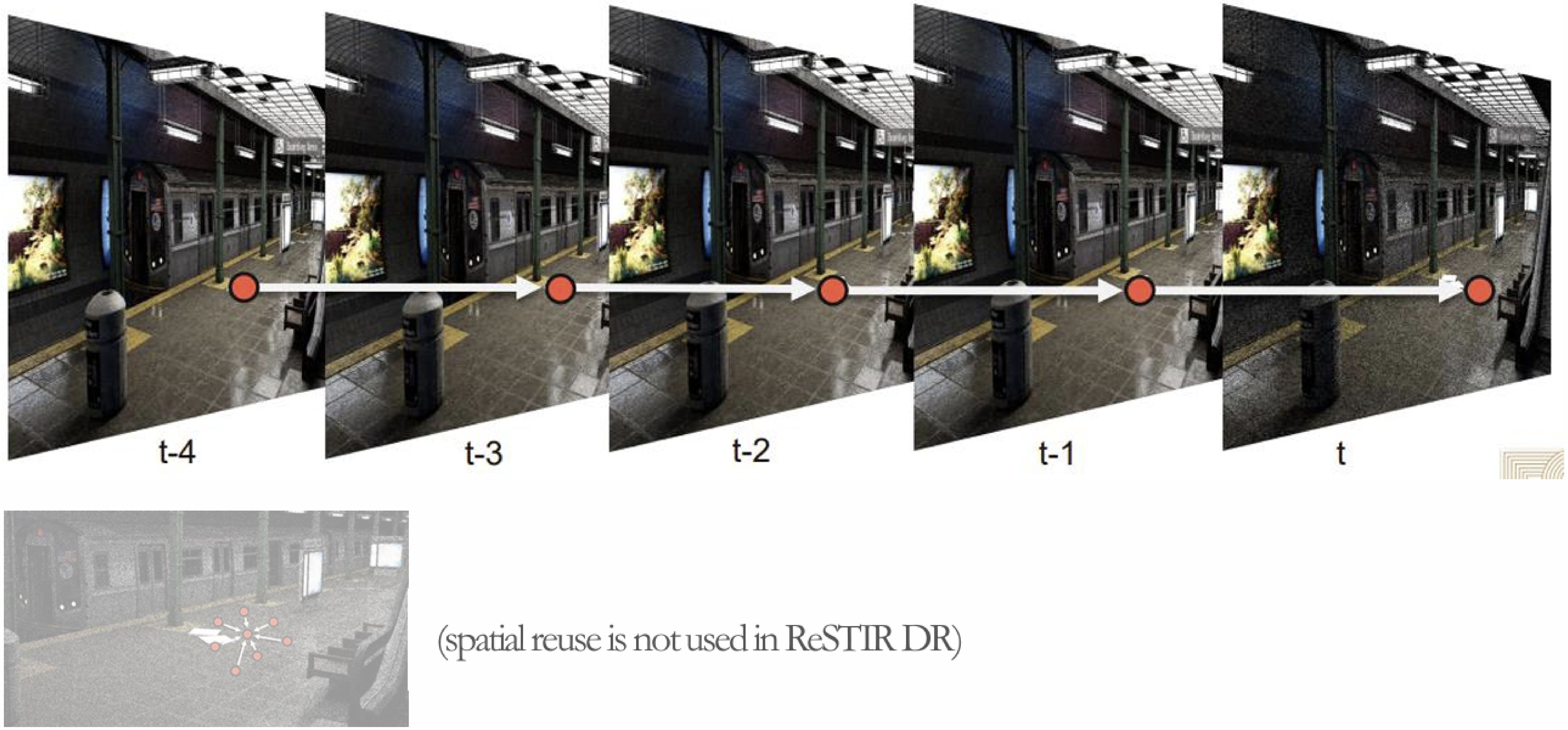
Previous talk: Team 1

Parameter-space ReSTIR for Differentiable and Inverse Rendering

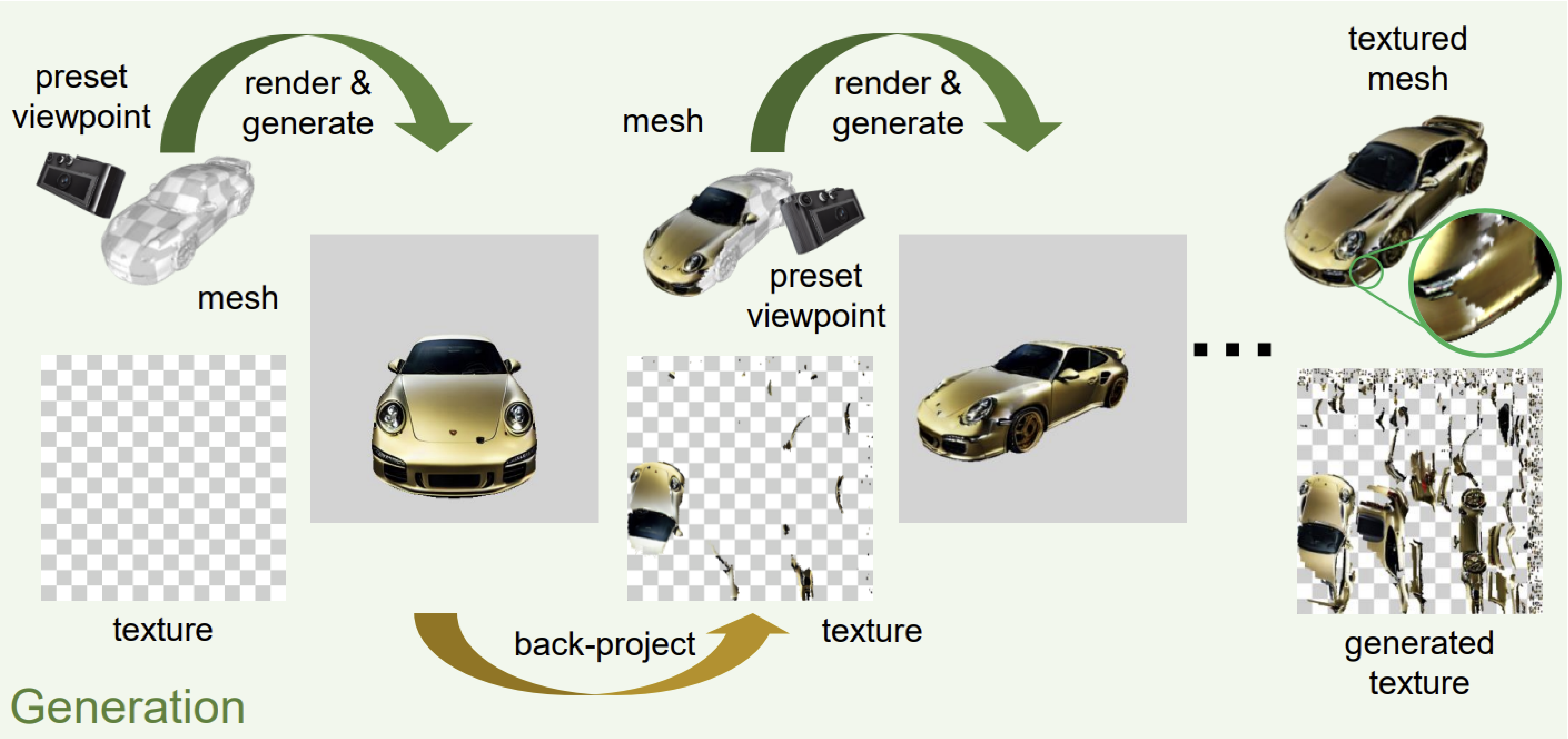


(spatial reuse is not used in ReSTIR DR)

A Gentle Introduction to ReSTIR: Path Reuse in Real-time, Wyman et al. *ACM SIGGRAPH 2023 Courses*

Previous talk: Team 4

Text2Tex: Text-driven Texture Synthesis via Diffusion Models



ECCV 2024

LGM: Large Multi-View Gaussian Model for High -Resolution 3D Content Creation

Team 3

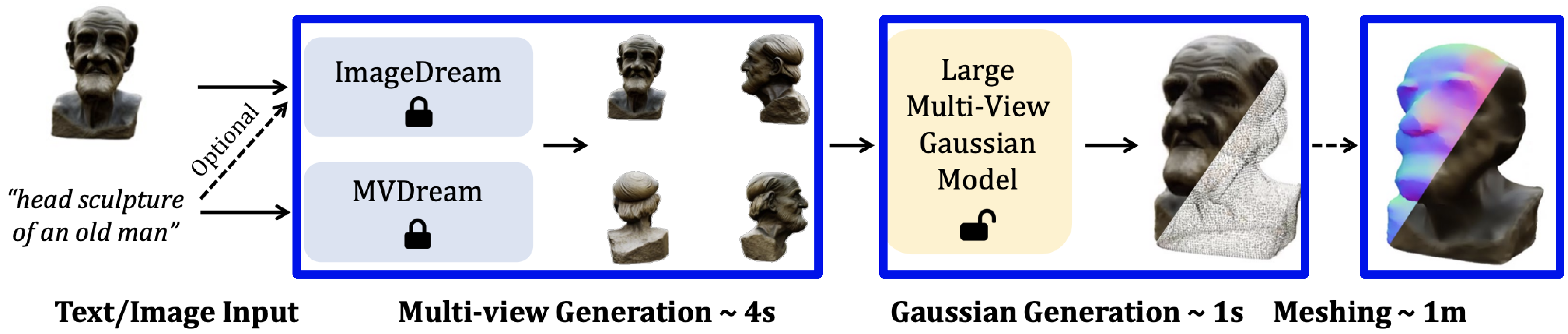
Woo Won Jung, Shin Rim Soo

Outline

1. Overview of LGM
2. Multi view Generation
3. Gaussian Generation
4. Mesh Extraction
5. Result & limitation

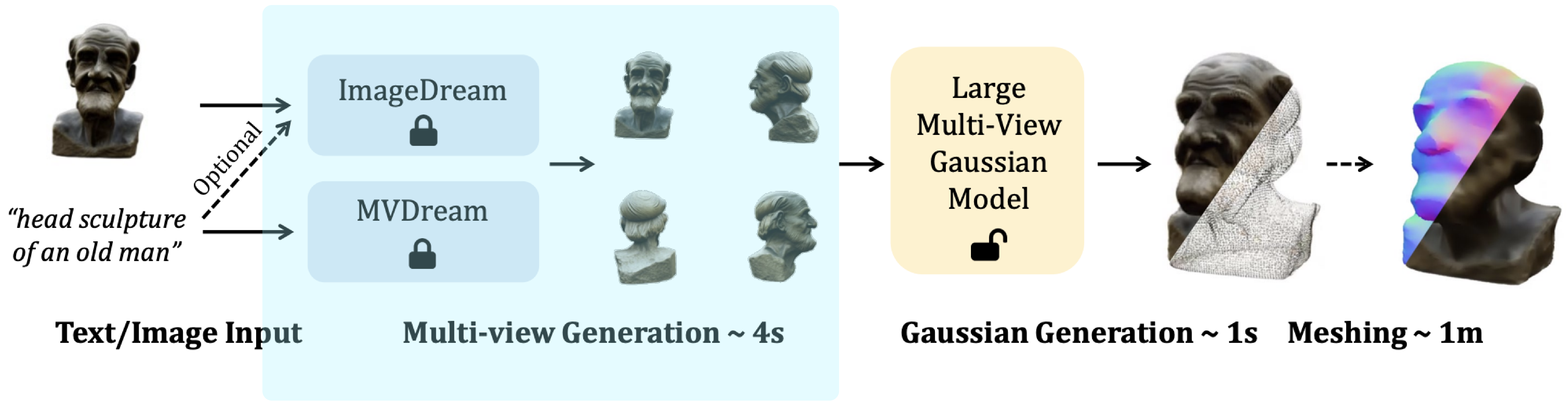
Motivation

LGM: Over-view


"head sculpture of an old man"
Optional
ImageDream
MVDream
Large Multi-View Gaussian Model
Text/Image Input
Multi-view Generation ~ 4s
Gaussian Generation ~ 1s
Meshing ~ 1m

Multi-view Generation	Creates images of multiple views
Gaussian Generation	Create Gaussian from multiple view pixels
Mesh Extraction	Convert 3D Gaussian into polygons

LGM: Multi-view Generation


"head sculpture of an old man"
Optional
ImageDream
MVDream
Large Multi-View Gaussian Model
Text/Image Input
Multi-view Generation ~ 4s
Gaussian Generation ~ 1s
Meshing ~ 1m

Multi-view Generation

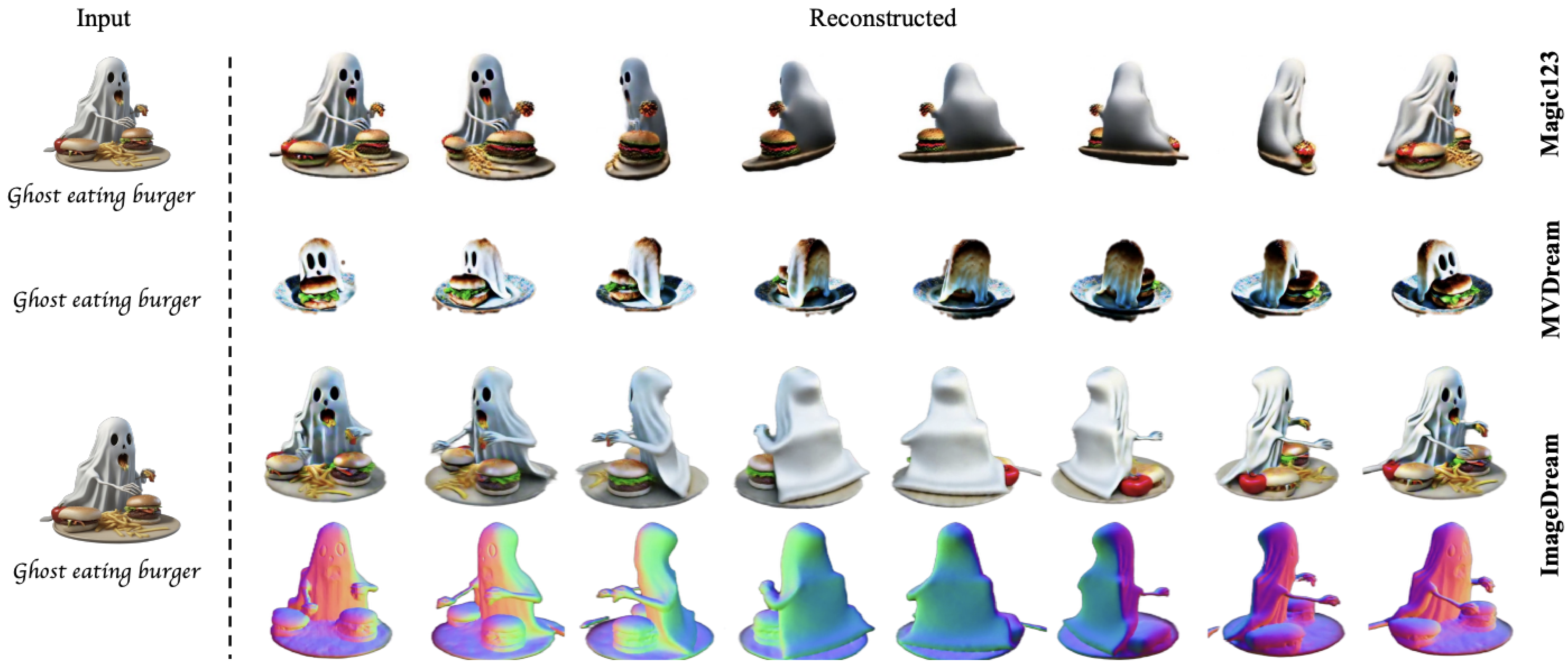
Creates images of multiple views

Gaussian Generation

Create Gaussian from multiple view pixels

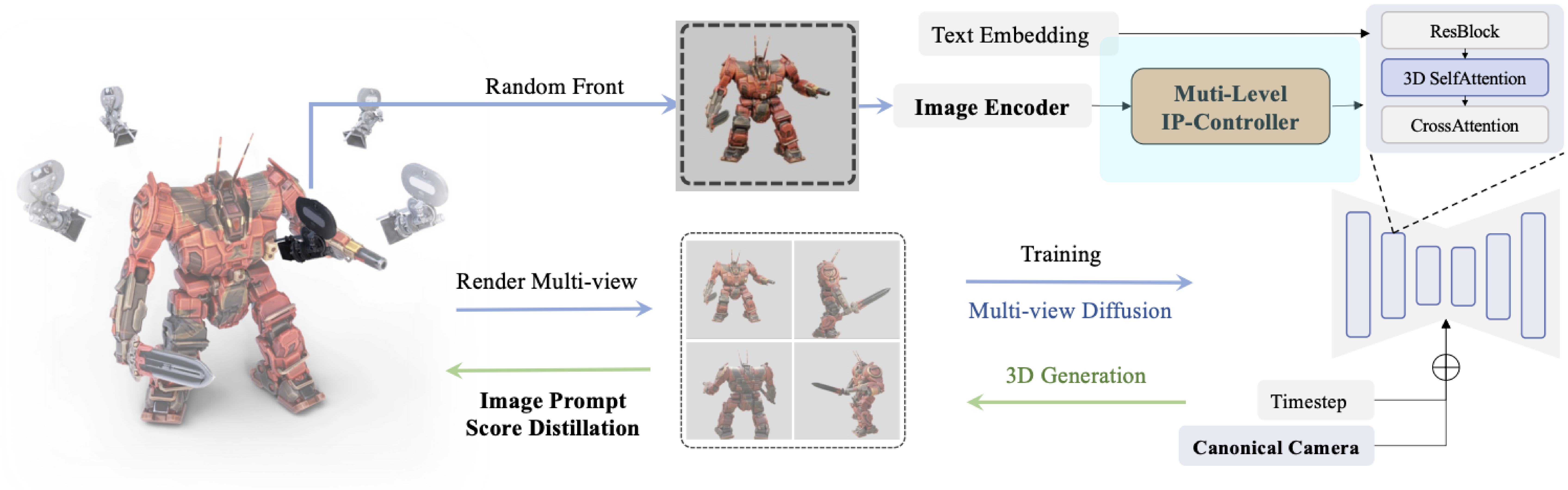
Prerequisites: Image Dream

: framework that generates high quality 3D model from any viewpoint given a single image



Prerequisites: Image Dream

: framework that generates high quality 3D model from any viewpoint given a single image

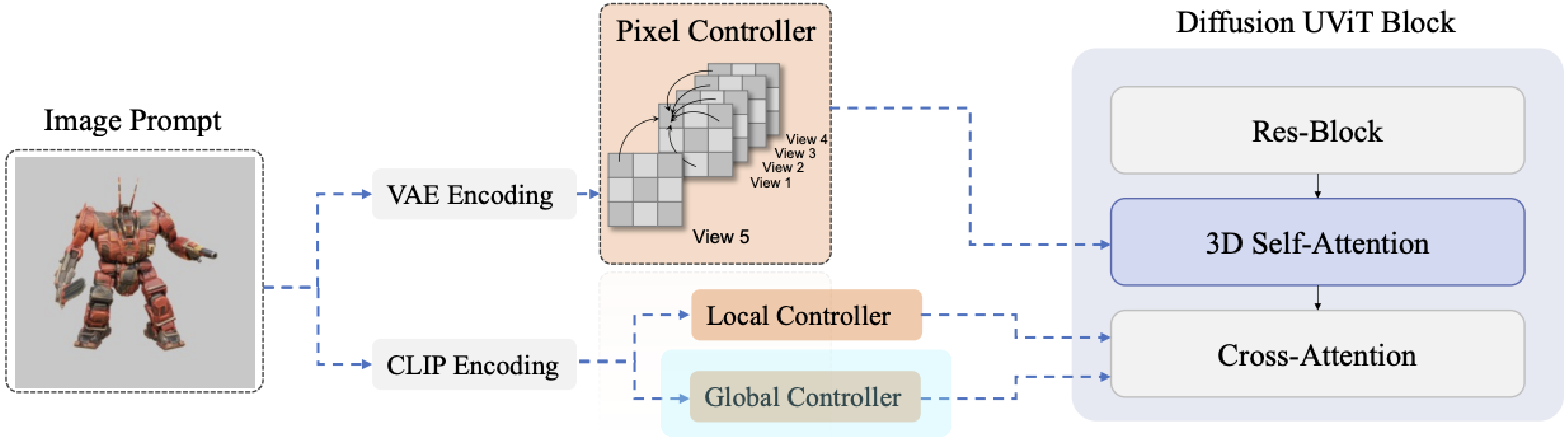


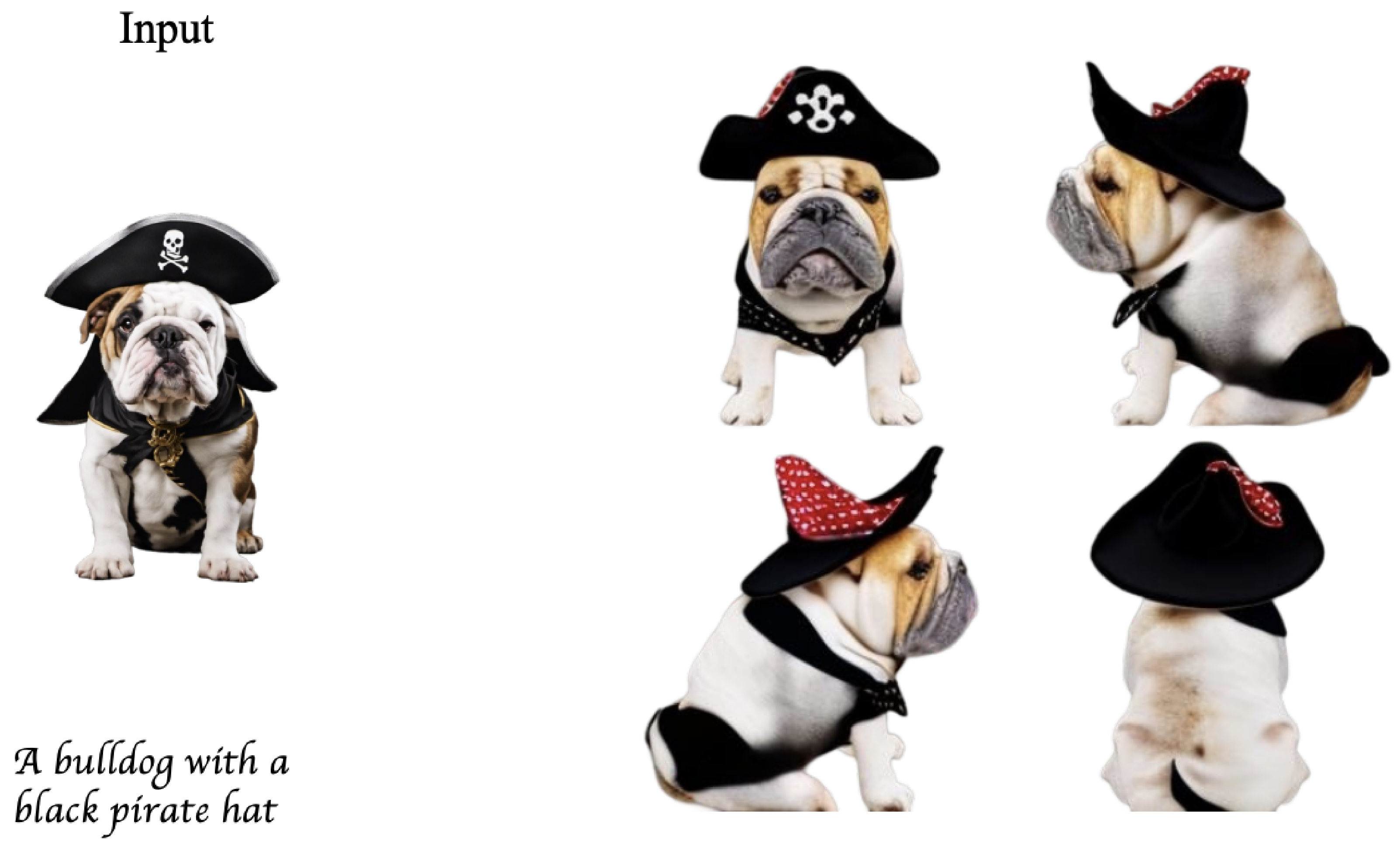
→ Training of the diffusion network

→ Training of the NeRF model

Prerequisites: Image Dream

: framework that generates high quality 3D model from any viewpoint given a single image

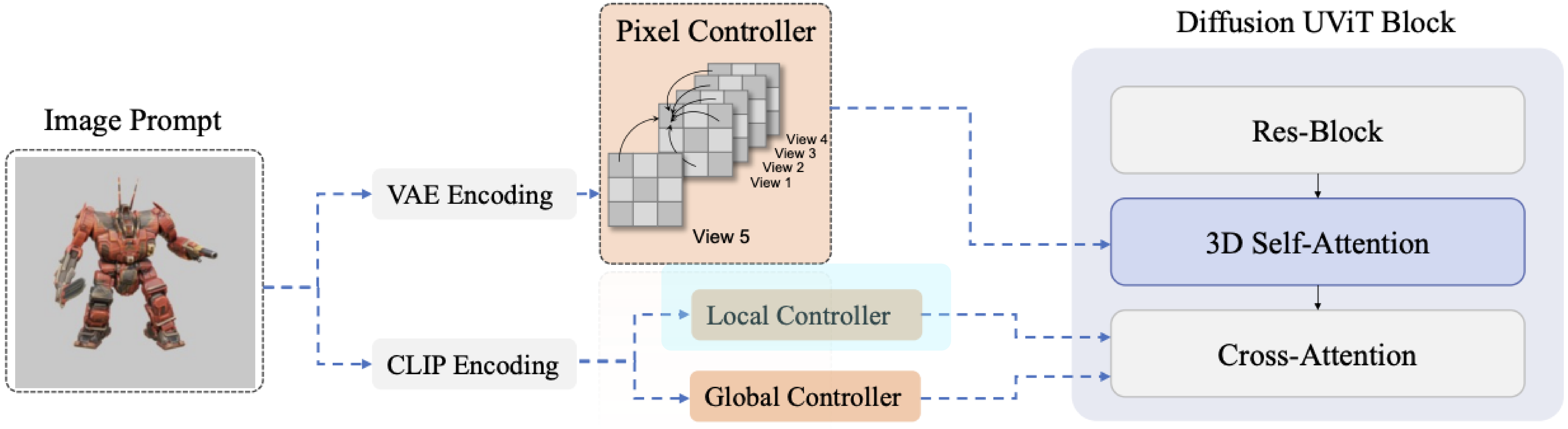


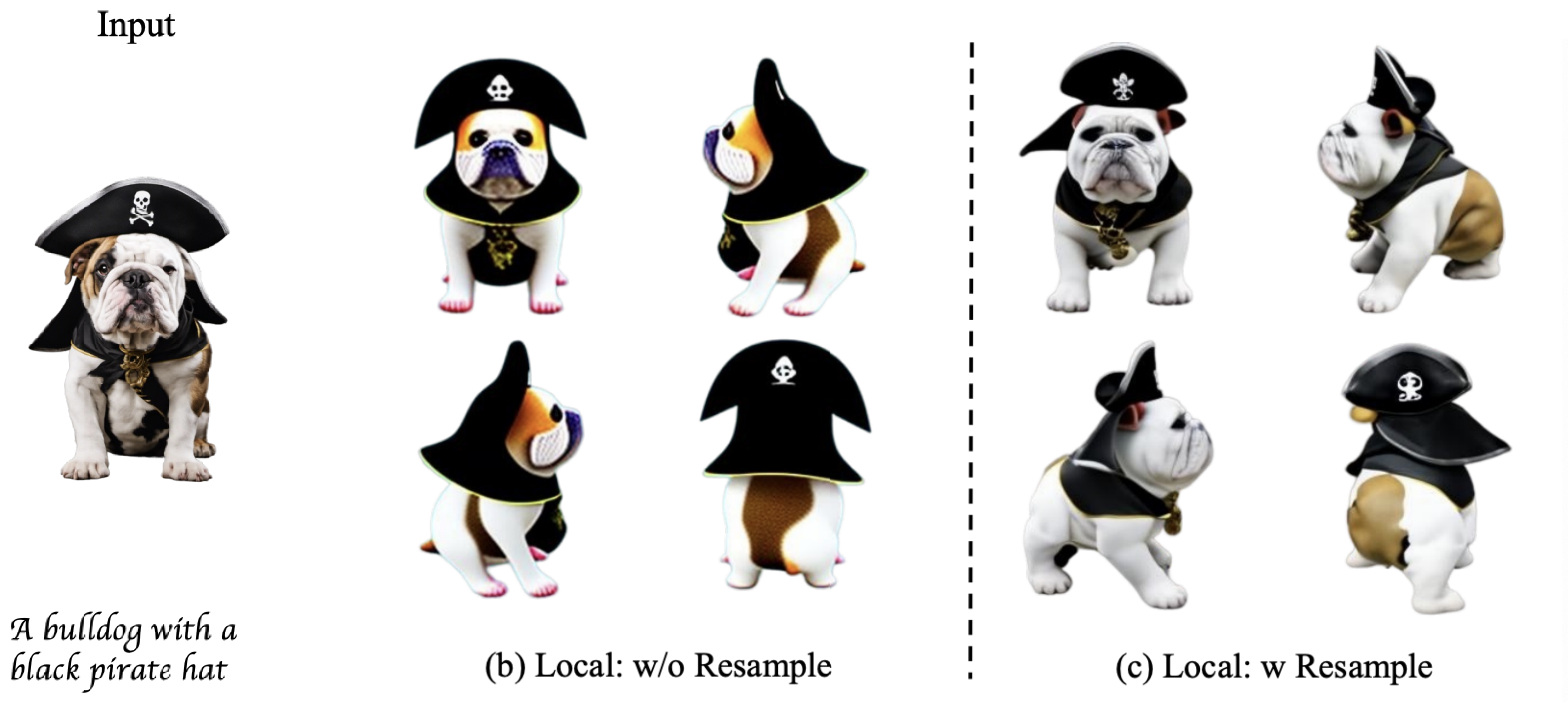


(a) Global

Prerequisites: Image Dream

: framework that generates high quality 3D model from any viewpoint given a single image



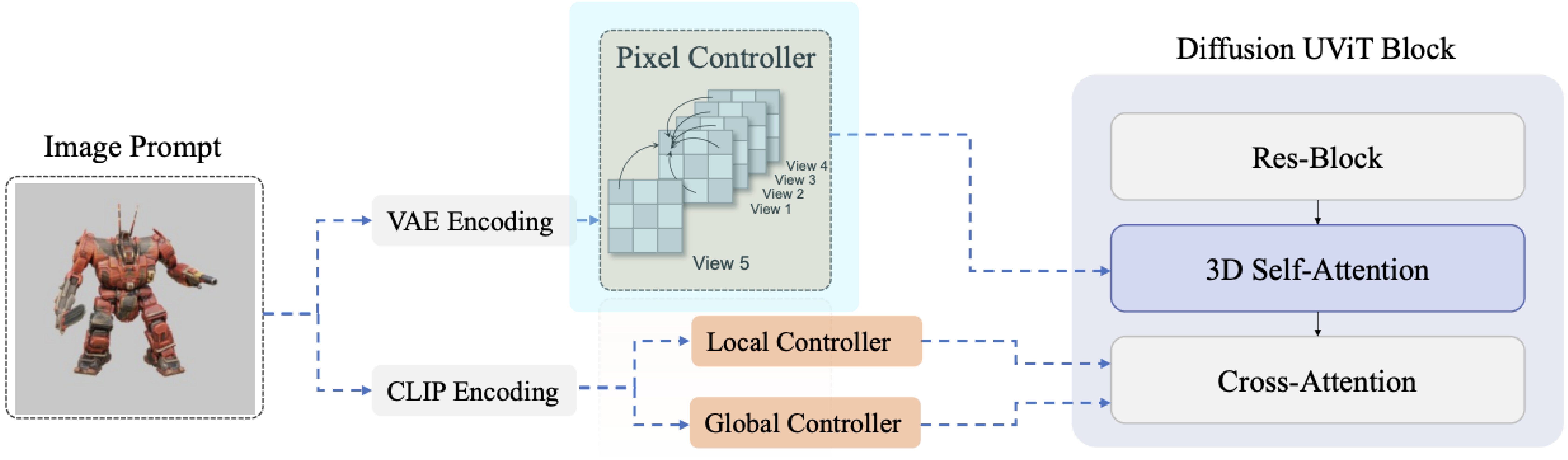


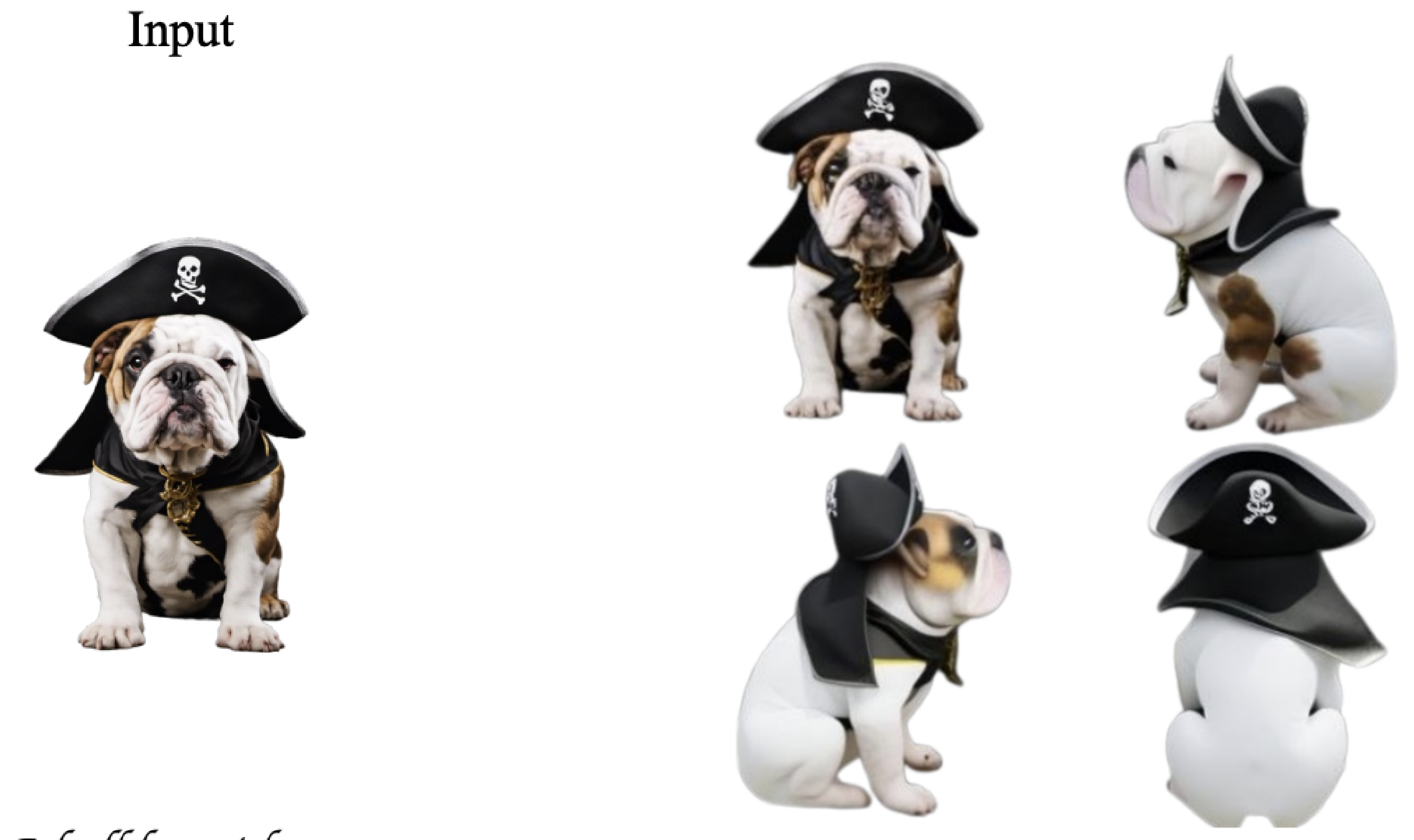
(b) Local: w/o Resample

(c) Local: w Resample

Prerequisites: Image Dream

: framework that generates high quality 3D model from any viewpoint given a single image

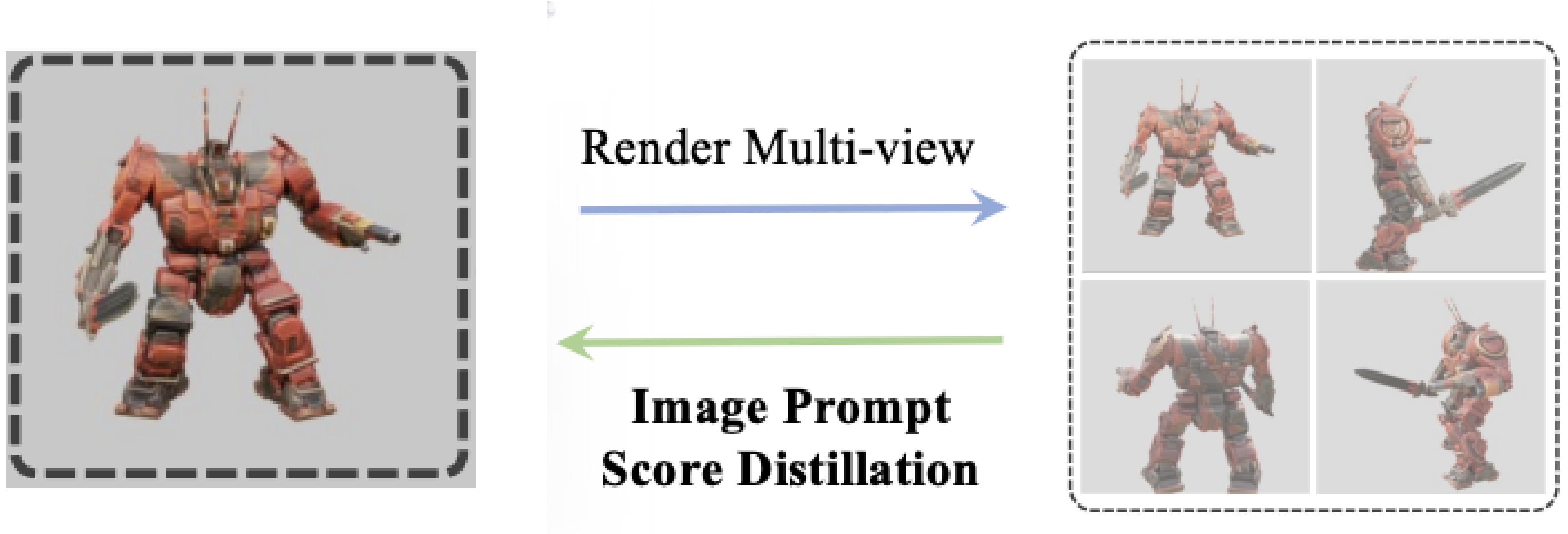




A bulldog with a black pirate hat

(d) +Pixel

LGM: Multi-view Generation



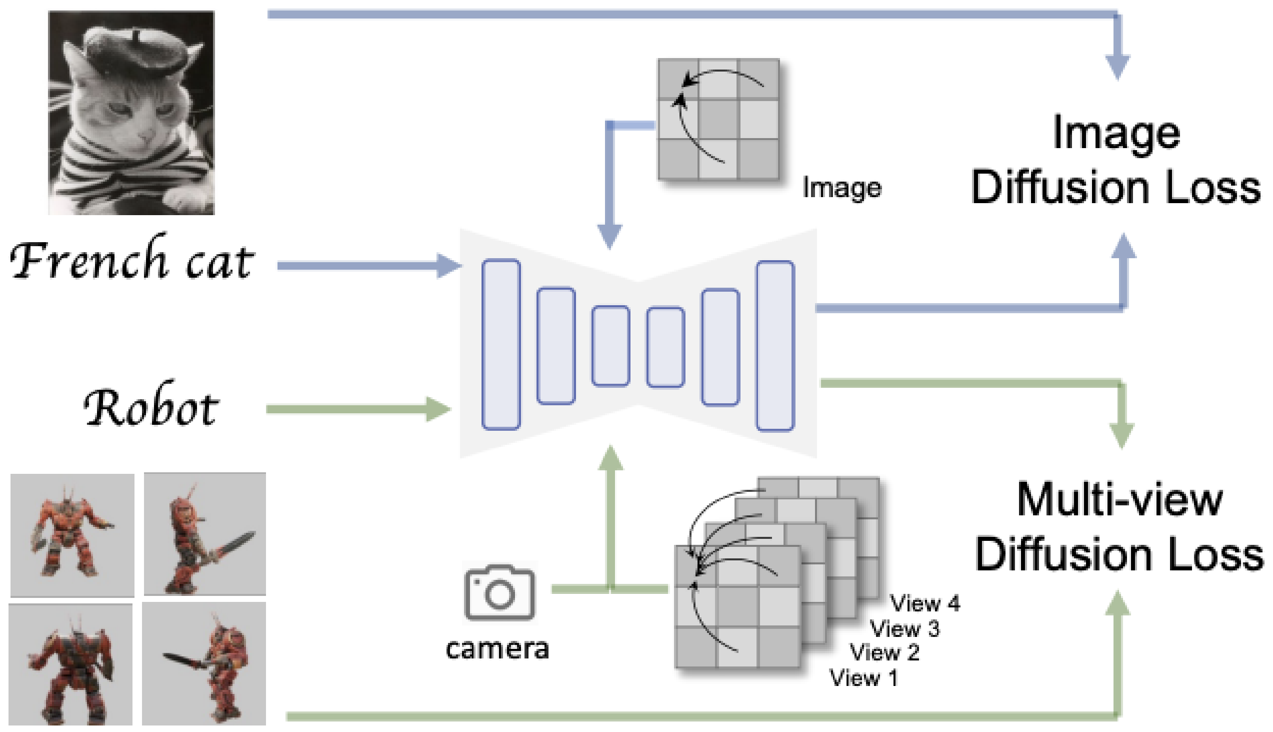


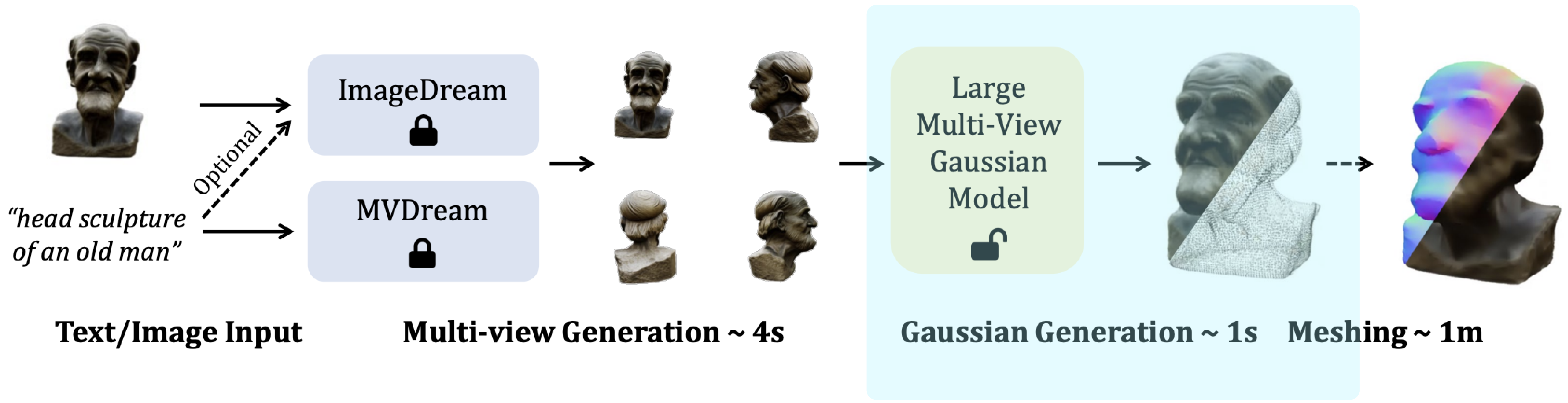
Image Dream

Takes photo as an input to create images of multiple view

MVDream

Takes text as an input to create images of multiple view

LGM: Gaussian Generation



Multi-view Generation

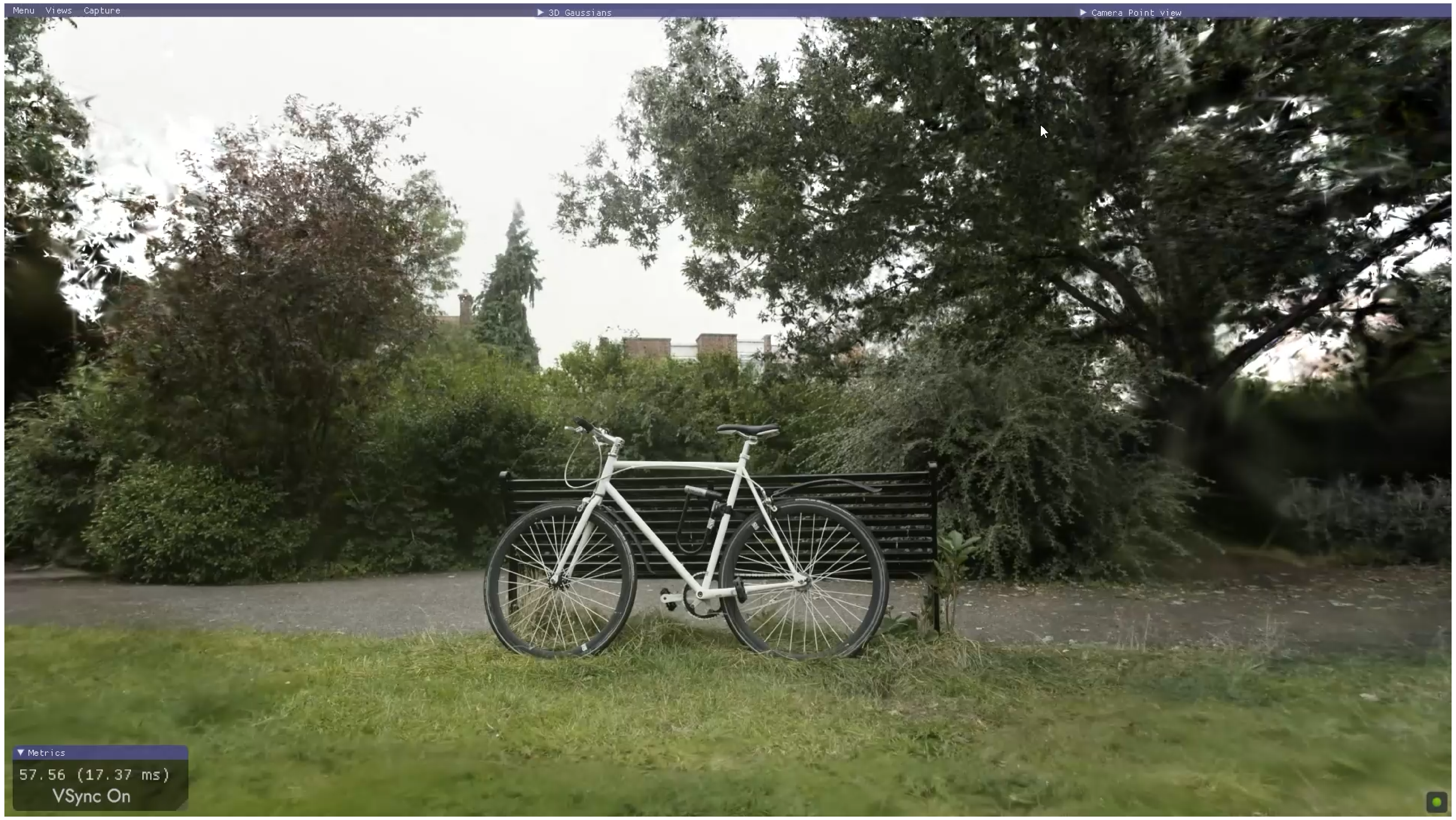
Creates images of multiple views

Gaussian Generation

Create Gaussian from multiple view pixels

Prerequisites: 3D Gaussian Splatting

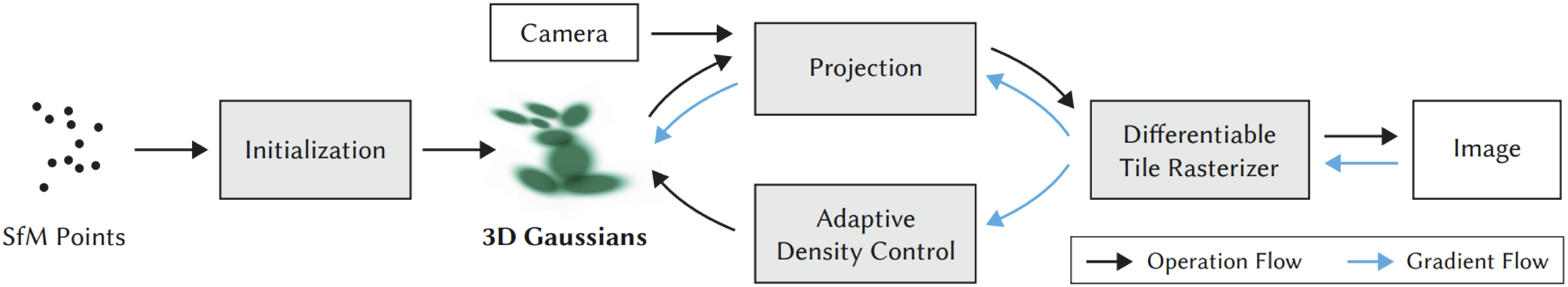
: technique used in the field of real-time radiance field rendering, which represents scenes with 3D Gaussians



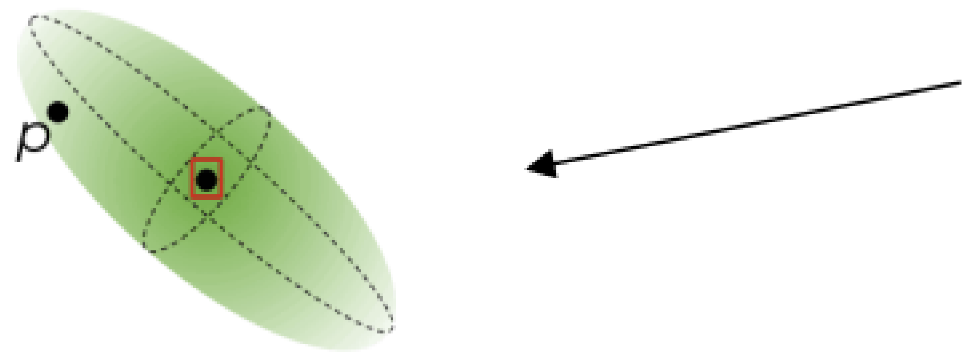
57.56 (17.37 ms)
VSync On

Prerequisites: 3D Gaussian Splatting

: (images from different views, camera position) -> (3D Gaussians)



$$f_i(p) = \sigma(\alpha_i)\exp(-\frac{1}{2}(p-\boxed{\mu_i})\Sigma_i^{-1}(p-\boxed{\mu_i}))$$

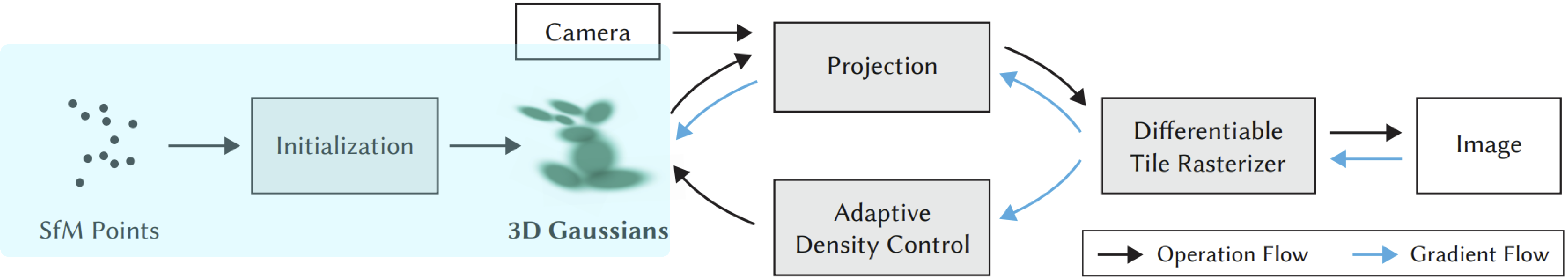


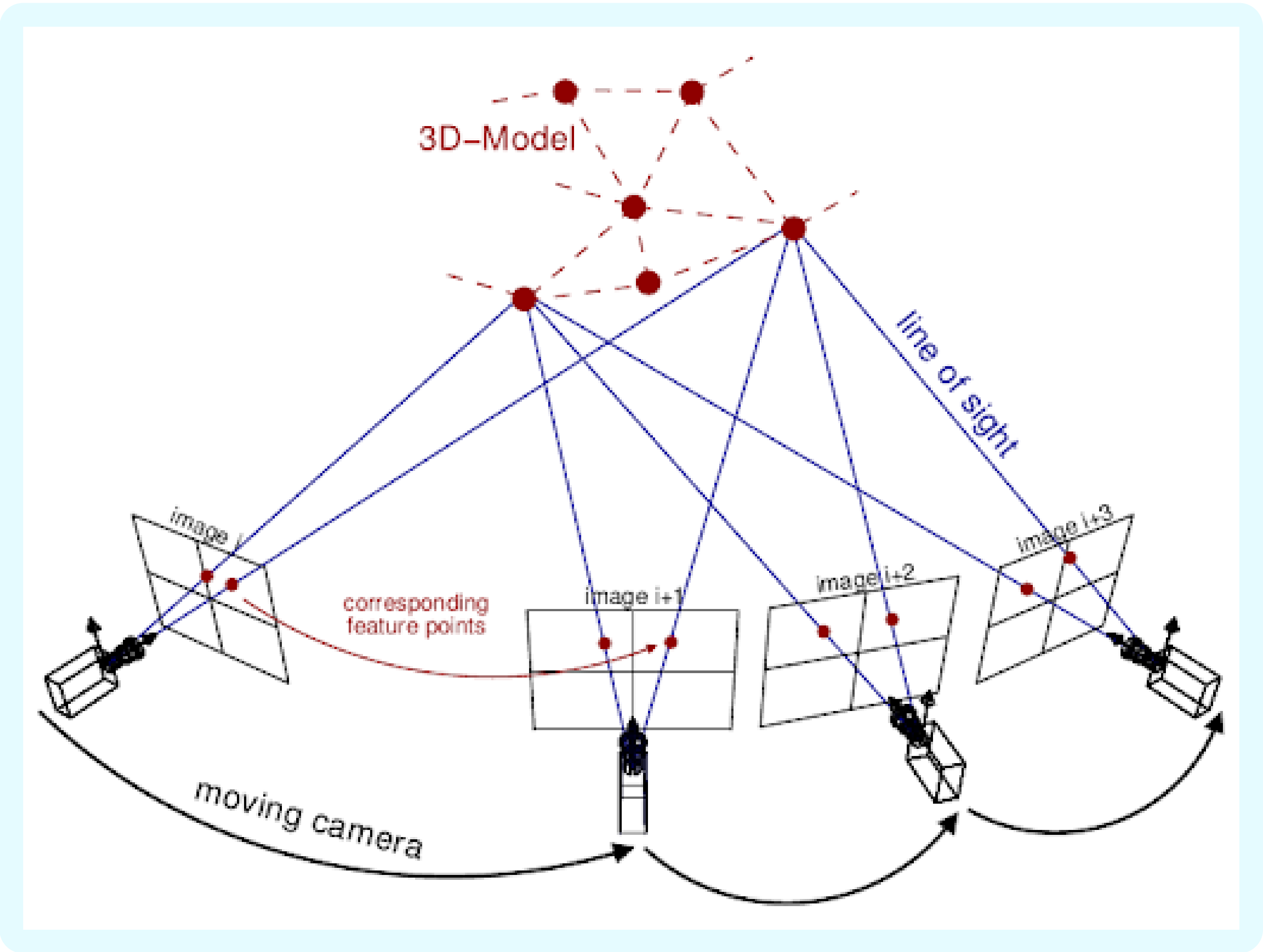
Each 3D Gaussian is parametrized by:

- **Mean μ:** 3D position (x, y, z)
- **Covariance $\Sigma = RSS^TR^T$**; (Scale S, Rotation R)
- **Opacity**: **$\sigma(\alpha)$**
- **Color parameters:** spherical harmonics (SH) coefficients.

Prerequisites: 3D Gaussian Splatting

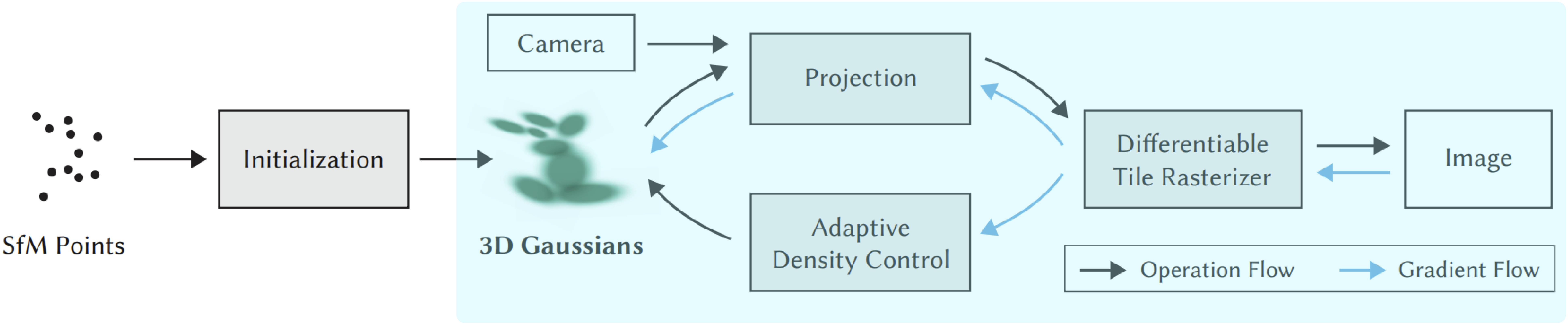
: (images from different views, camera position) -> (3D Gaussians)

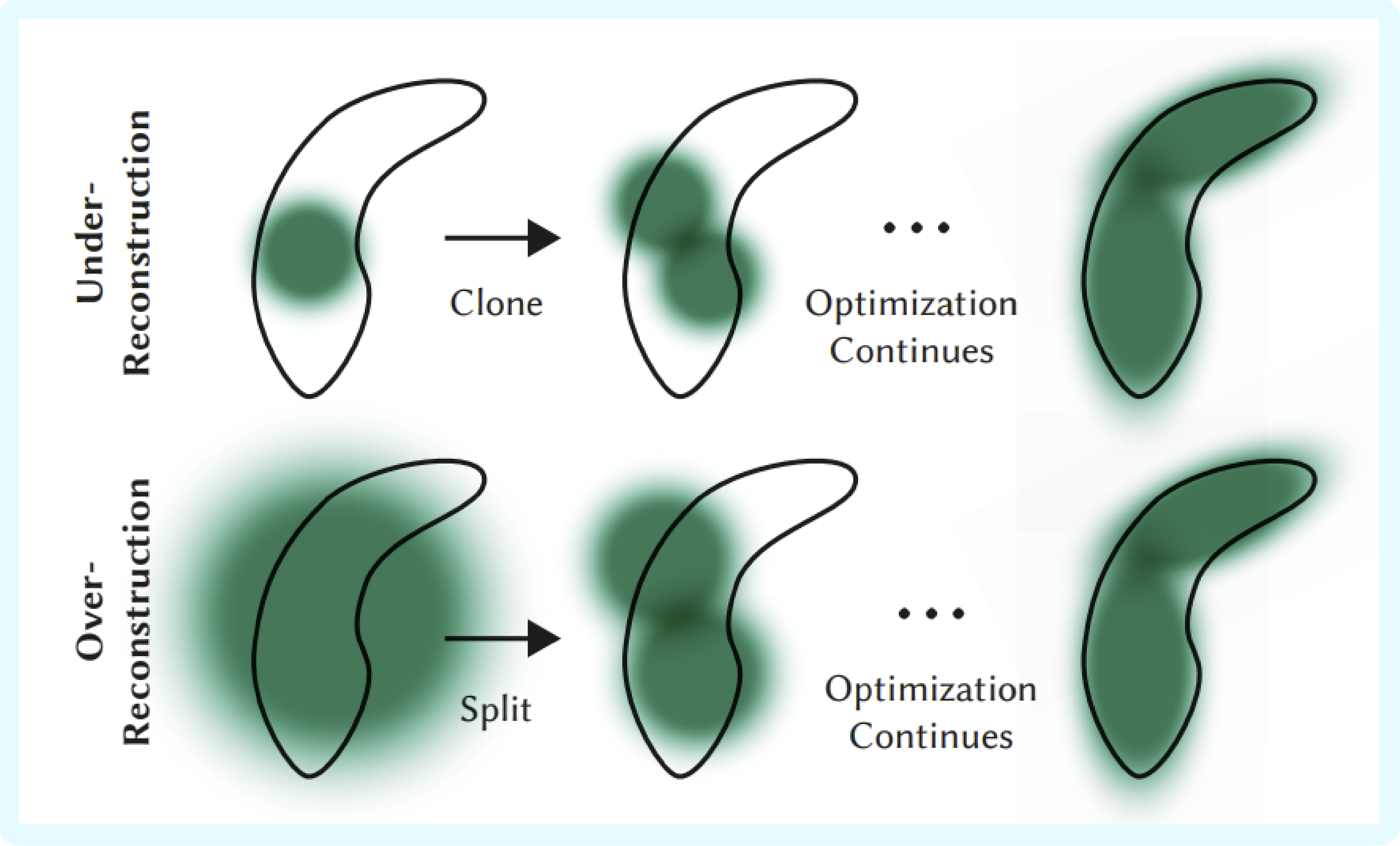




Prerequisites: 3D Gaussian Splatting

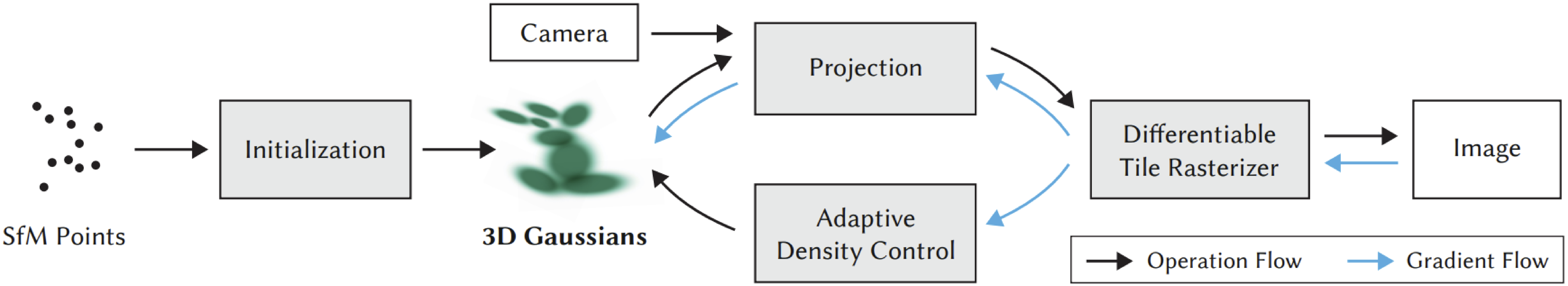
: (images from different views, camera position) -> (3D Gaussians)

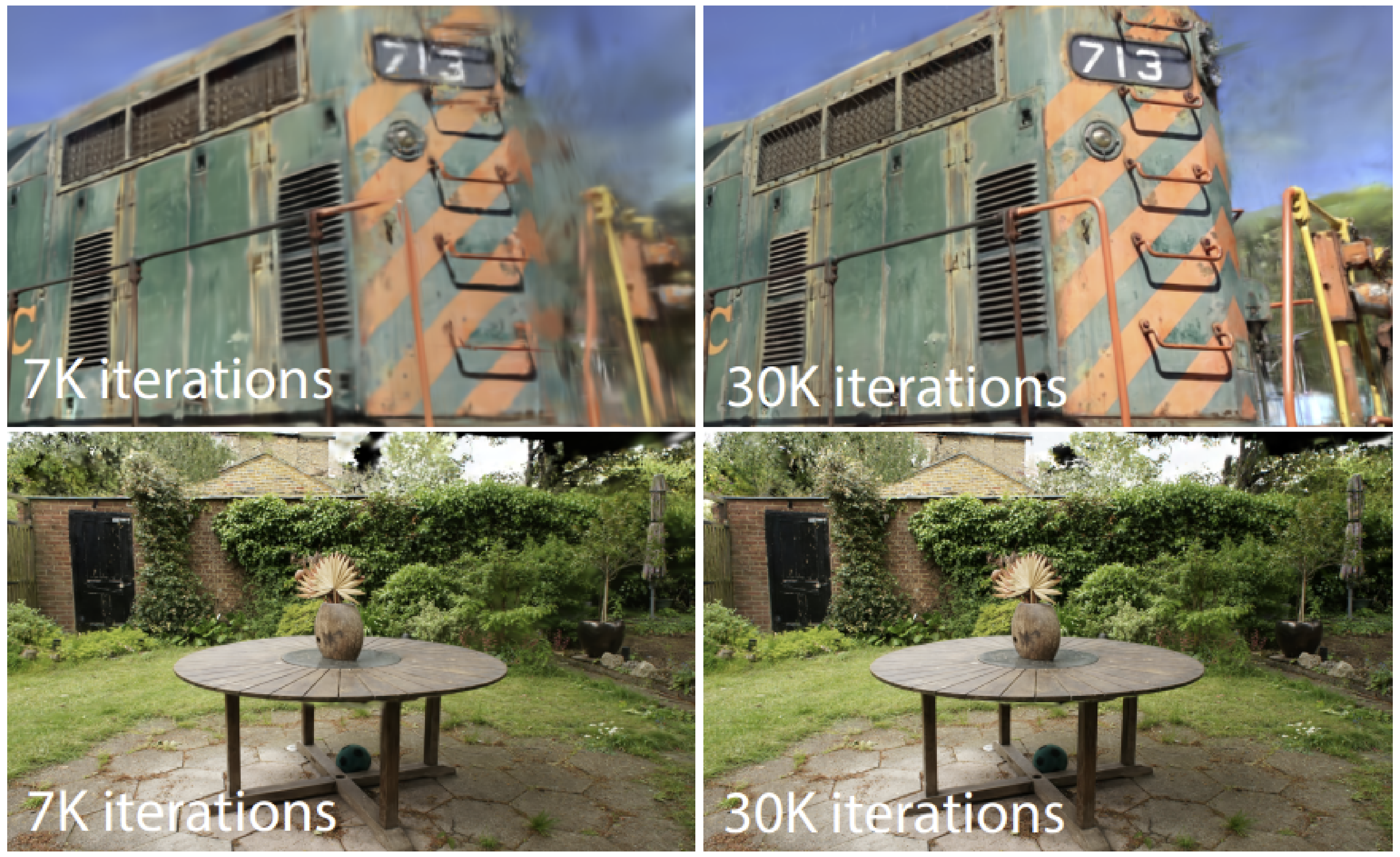




Prerequisites: 3D Gaussian Splatting

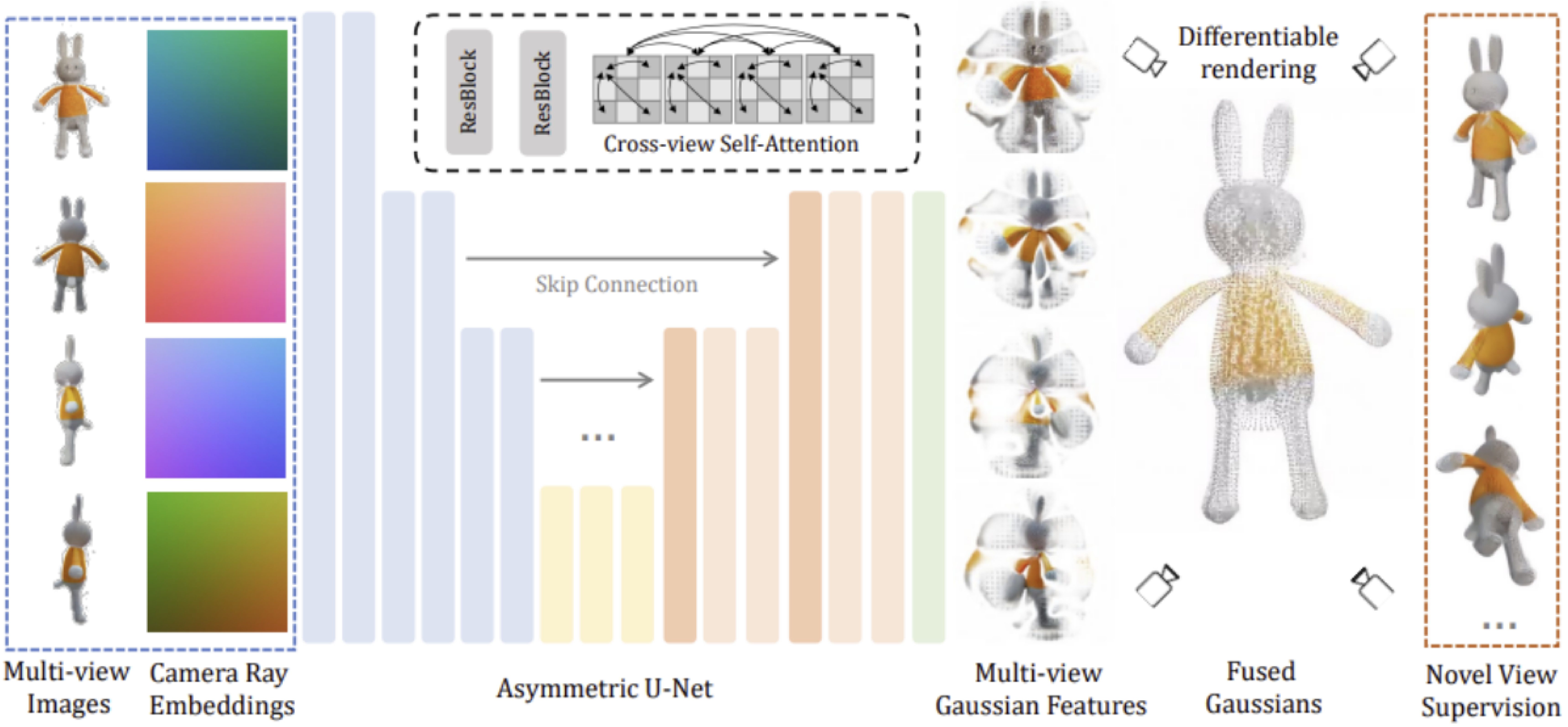
: (images from different views, camera position) -> (3D Gaussians)





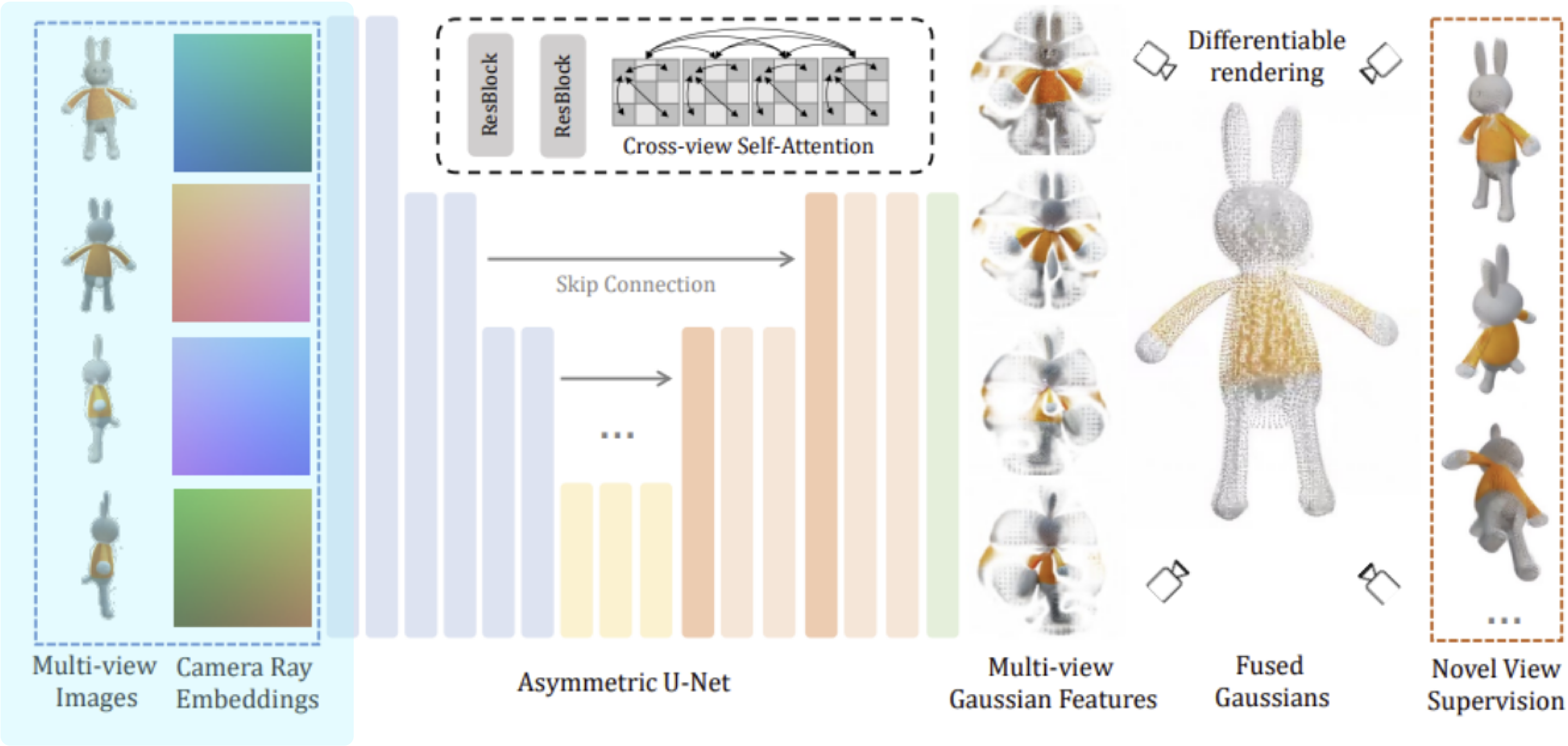
LGM: Gaussian Generation

: (images from different views, camera position) -> (3D Gaussians)



LGM: Gaussian Generation

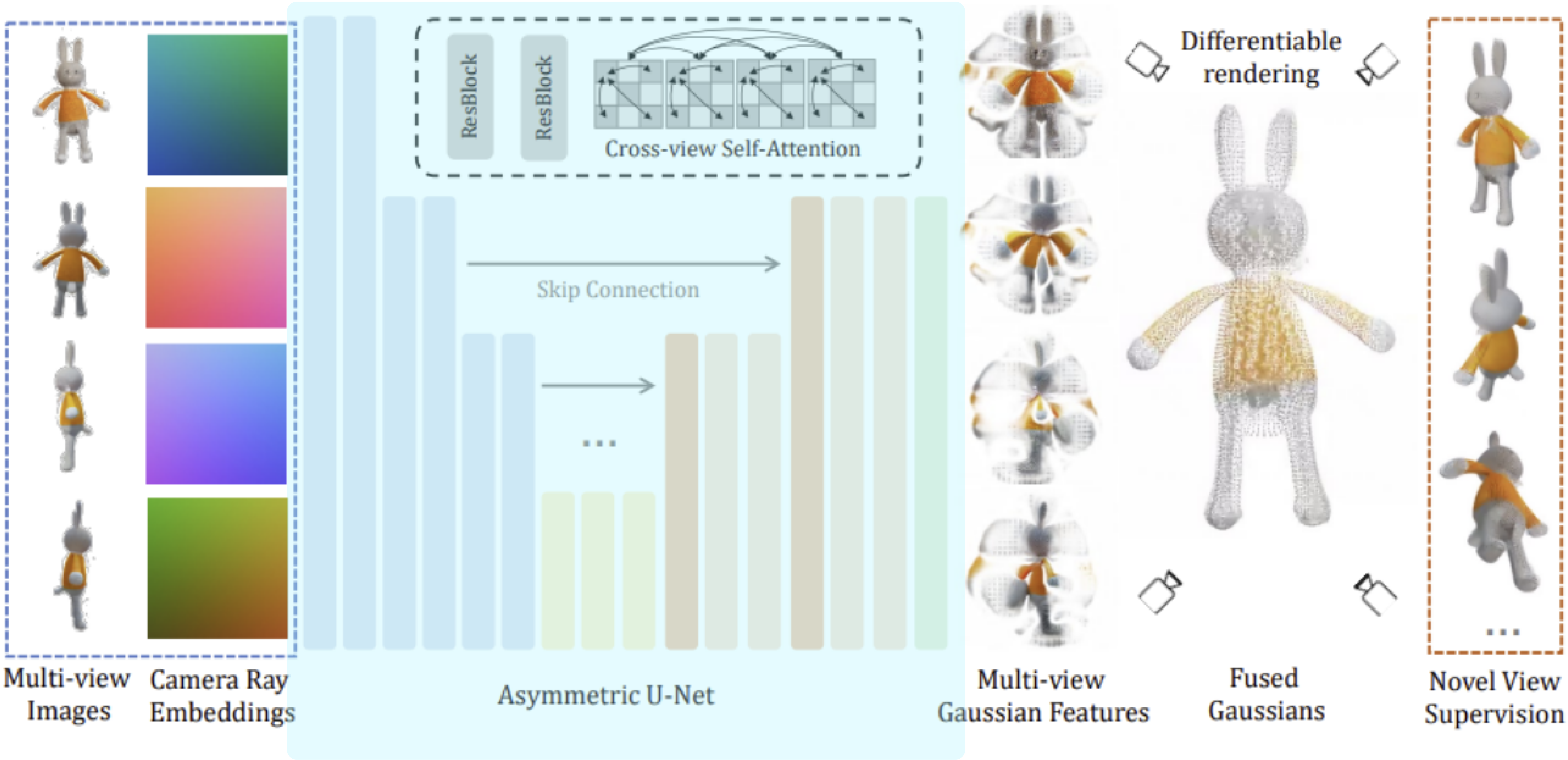
: (images from different views, camera position) -> (3D Gaussians)



9-channel feature map: color(RGB: 3), ray direction(3), ray origin(3)

LGM: Gaussian Generation

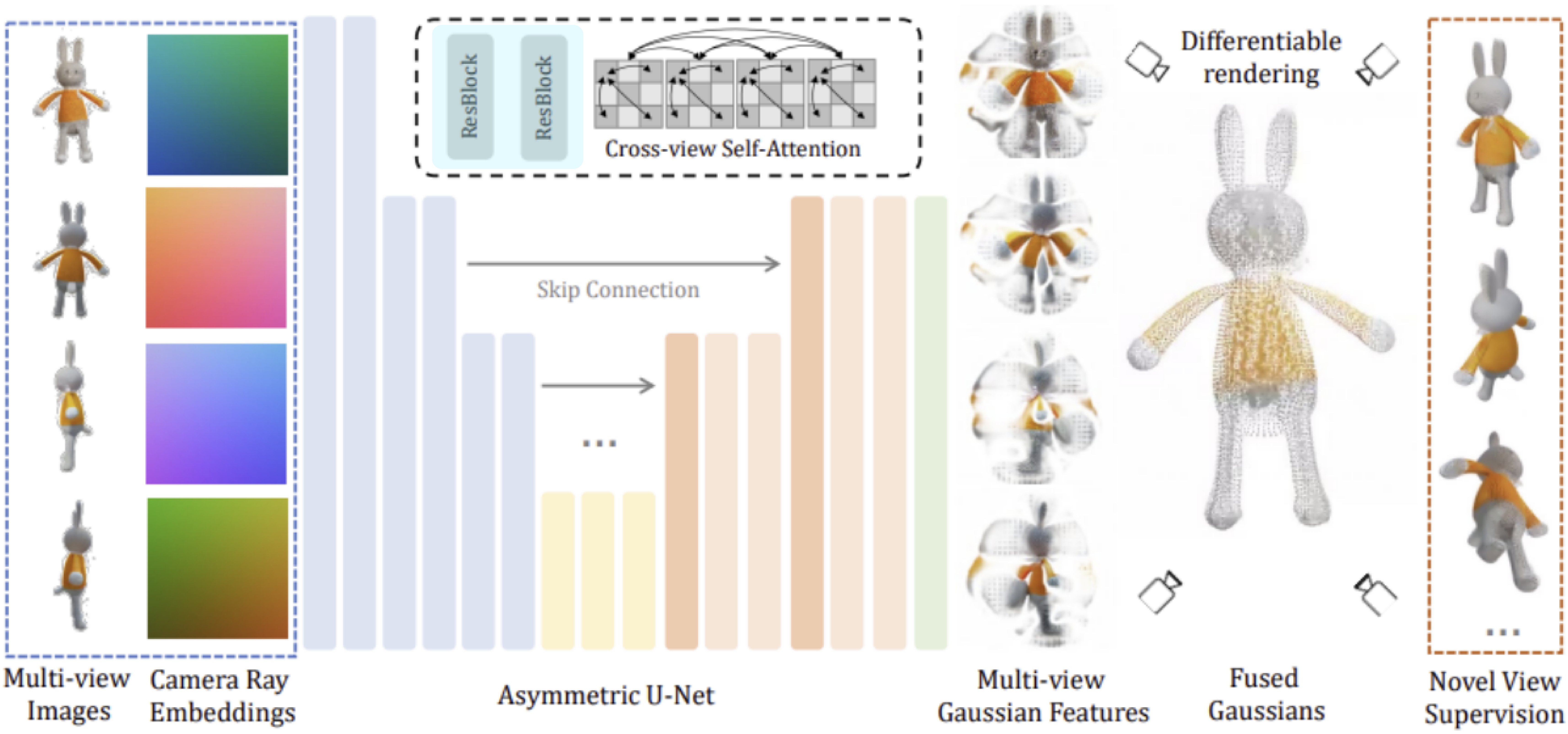
: (images from different views, camera position) -> (3D Gaussians)



Asymmetric U-Net: residual block, cross-view self-attention

LGM: Gaussian Generation

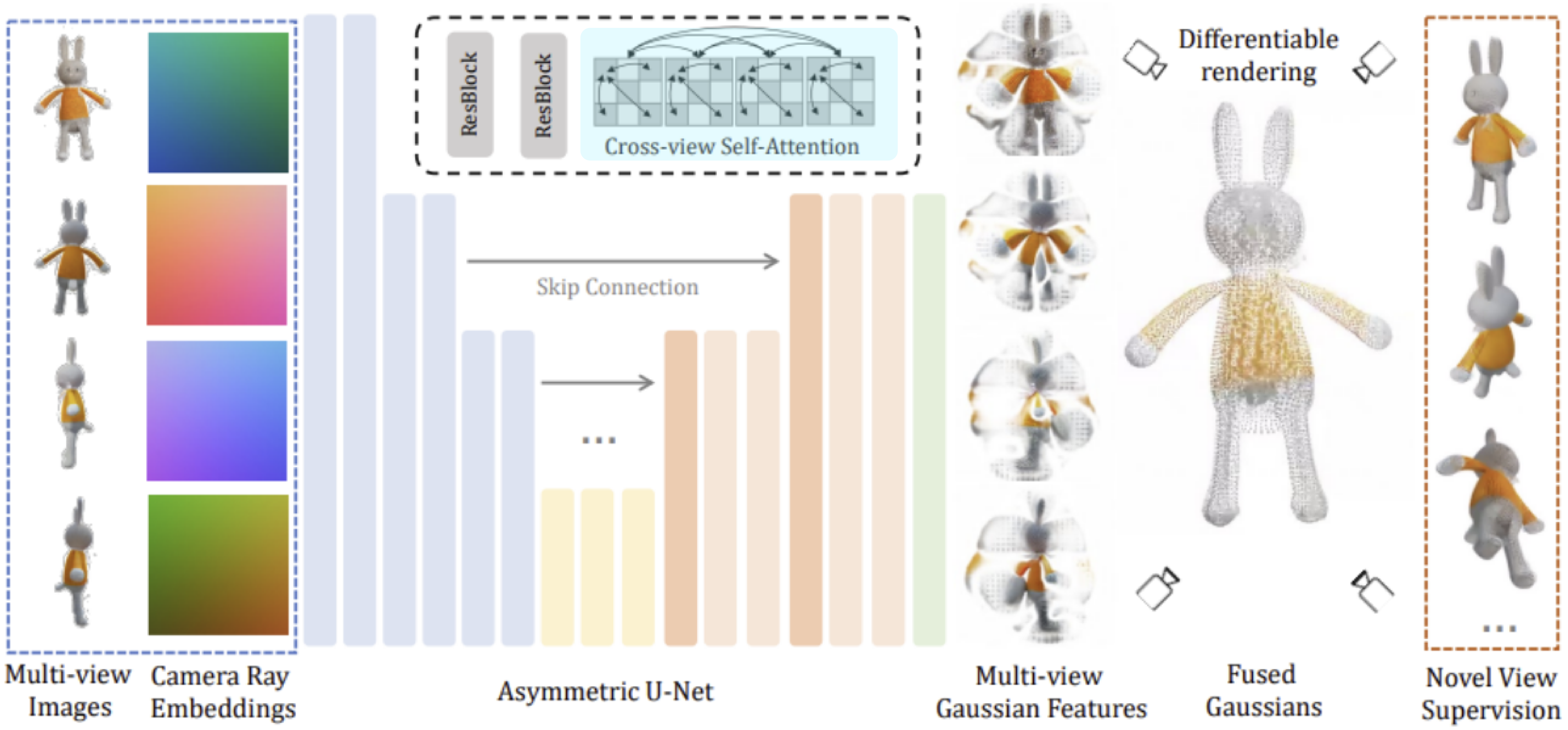
: (images from different views, camera position) -> (3D Gaussians)



Asymmetric U-Net: residual block, cross-view self-attention

LGM: Gaussian Generation

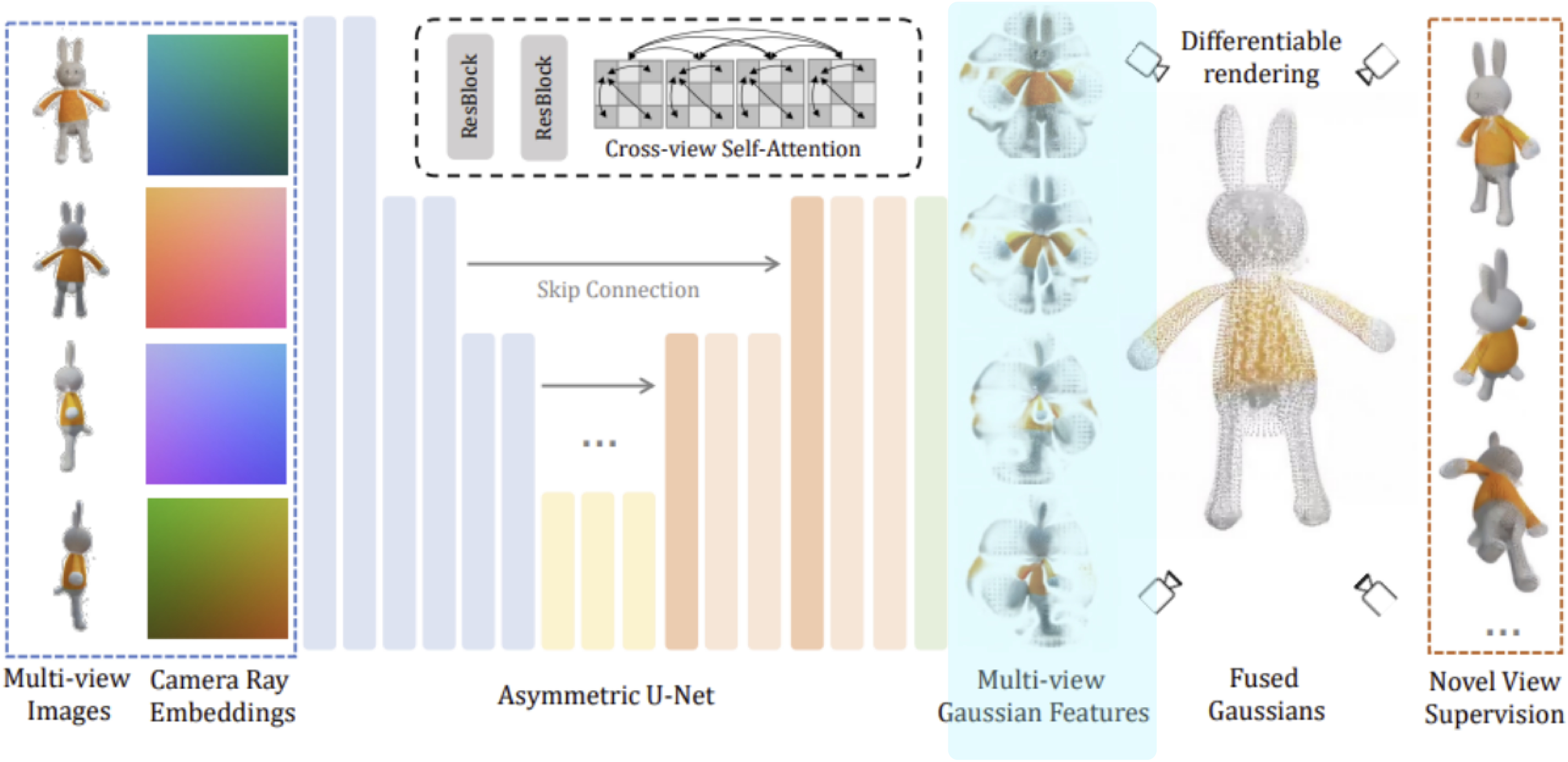
: (images from different views, camera position) -> (3D Gaussians)



Asymmetric U-Net: residual block, cross-view self-attention

LGM: Gaussian Generation

: (images from different views, camera position) -> (3D Gaussians)

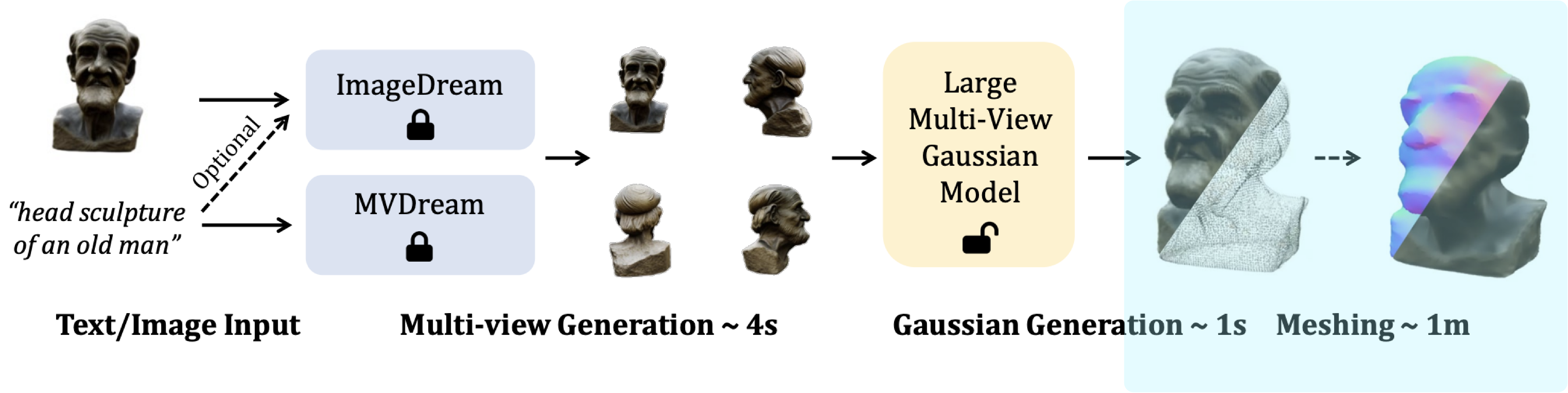


Each feature map -> 14-channel Gaussian features
* 14-channel: point cloud position (3), opacity (1), scale matrix (3), rotation matrix (4), color (3)

LGM: Gaussian Generation

: (images from different views, camera position) -> (3D Gaussians)

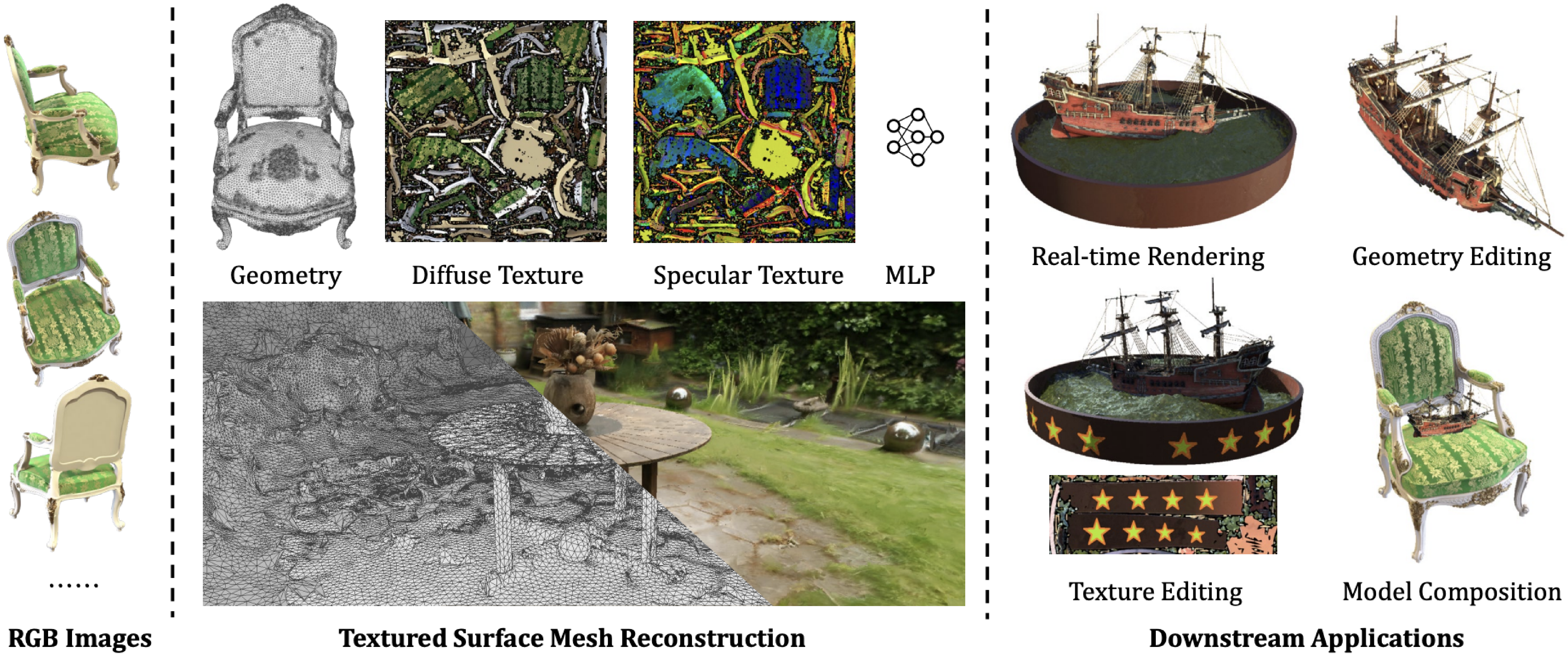
LGM: Mesh Extraction


"head sculpture of an old man"
Optional
ImageDream
MVDream
Large Multi-View Gaussian Model
Text/Image Input
Multi-view Generation ~ 4s
Gaussian Generation ~ 1s
Meshing ~ 1m

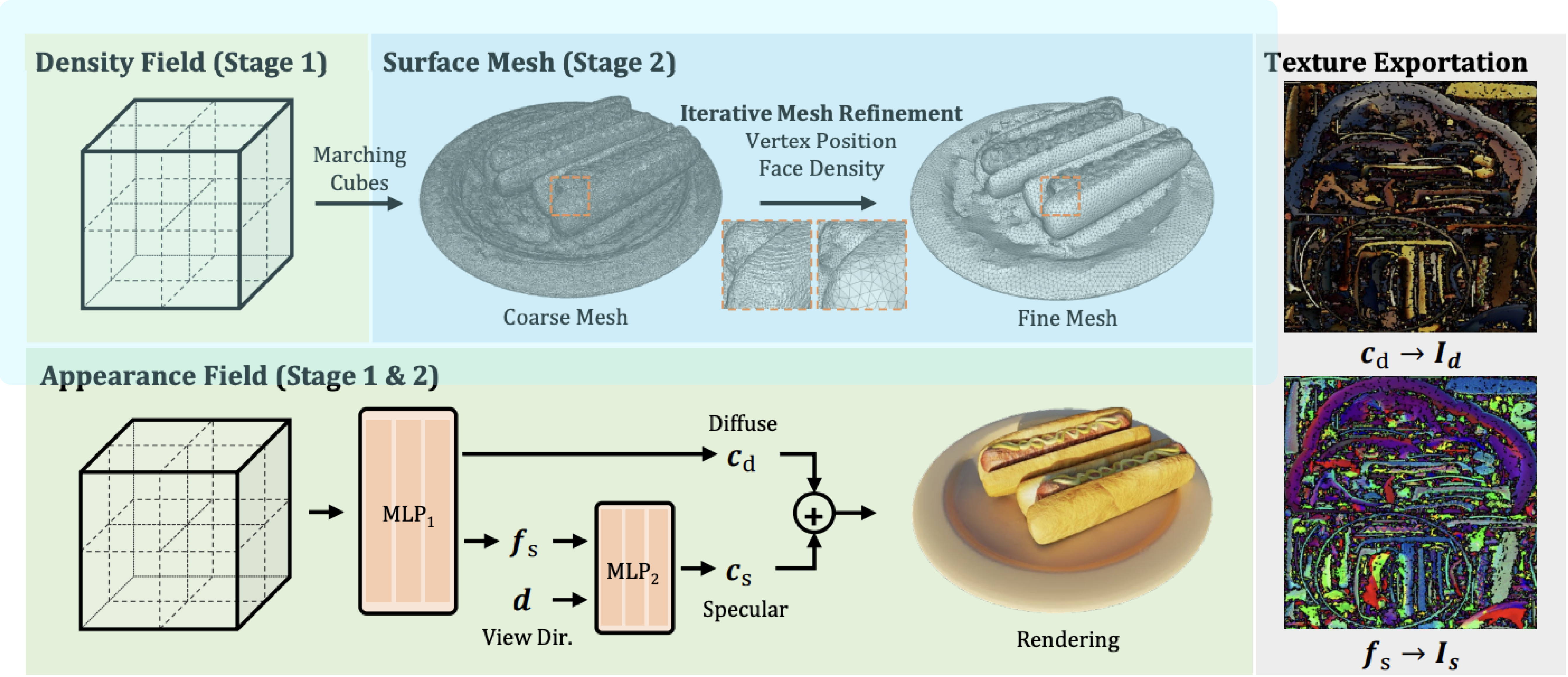
Mesh extraction: Process of converting 3D gaussian into mesh

Prerequisites: NeRF2Mesh

: framework to reconstruct textured surface meshes from multi-view RGB images



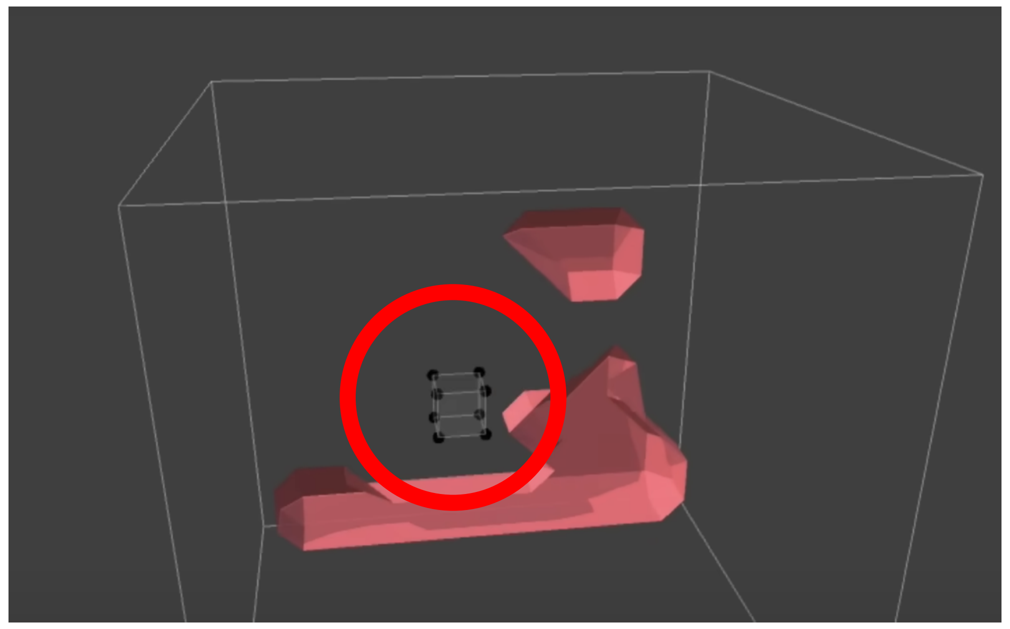
Prerequisites: NeRF2Mesh

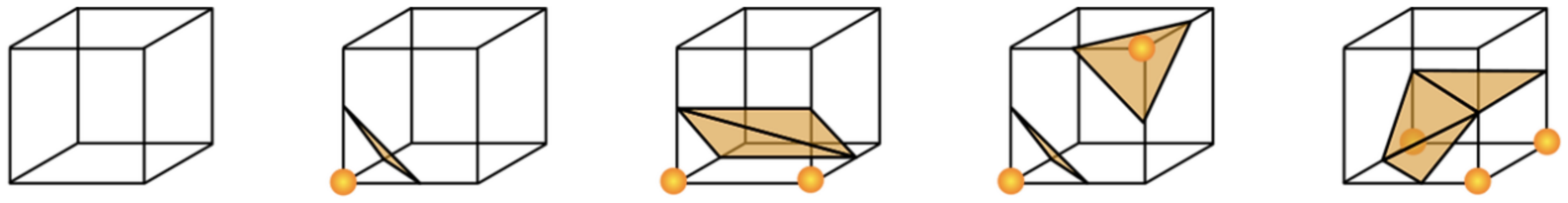


1. Geometry

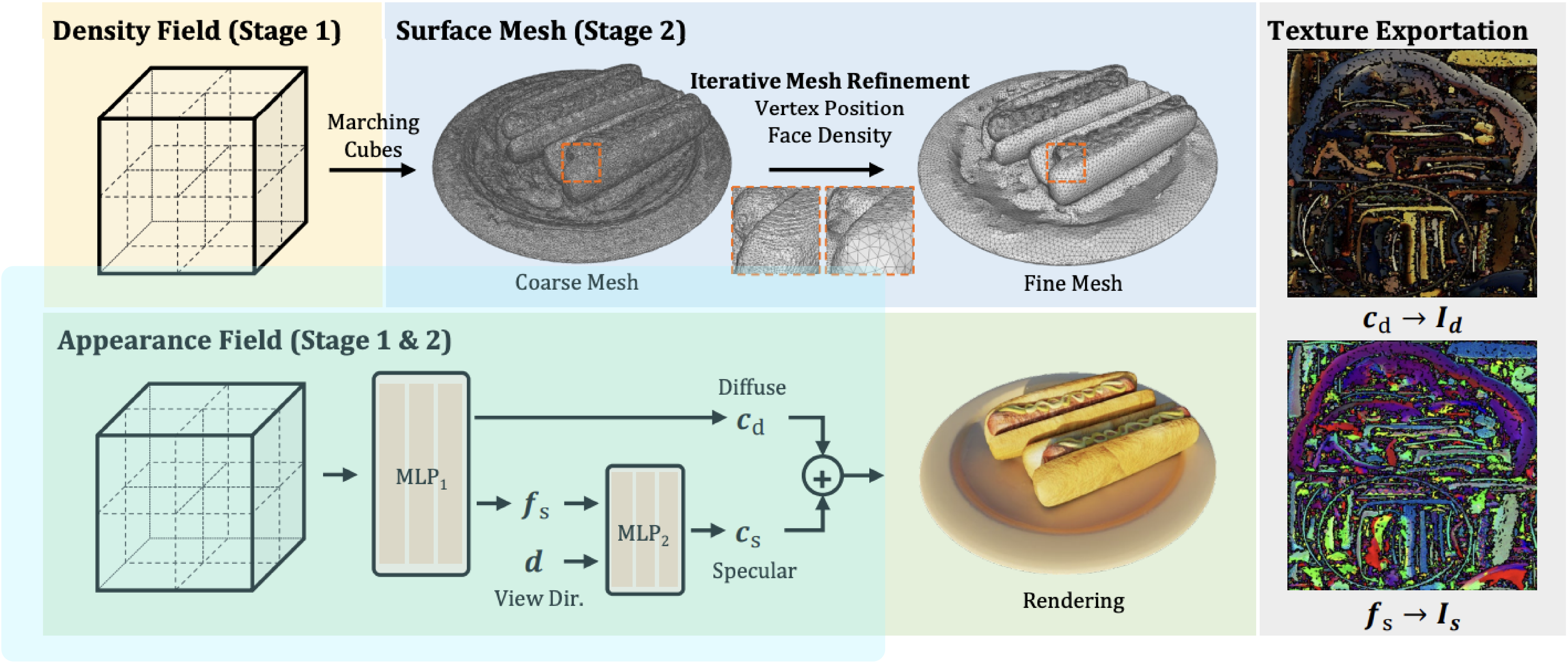
→ Initially learned from density grid, used Marching Cube to obtain coarse mesh. Using Iterative mesh refinement to Coarse Mesh obtain Fine Mesh.

2. Appearance Decomposition





Prerequisites: NeRF2Mesh

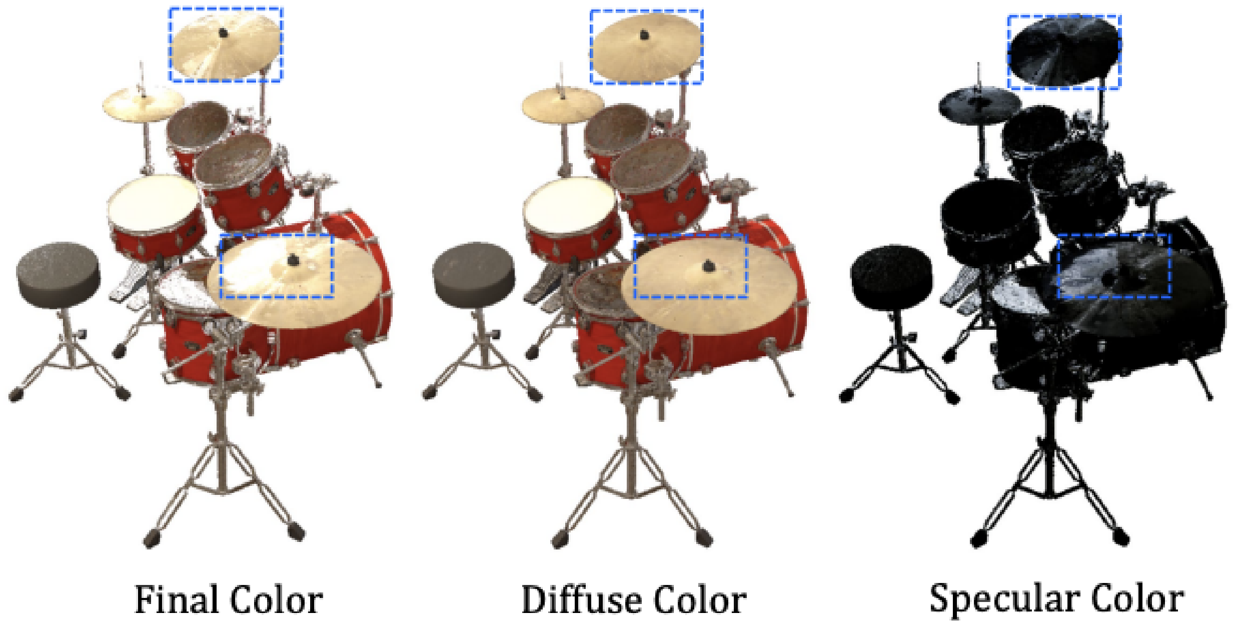


1. Geometry

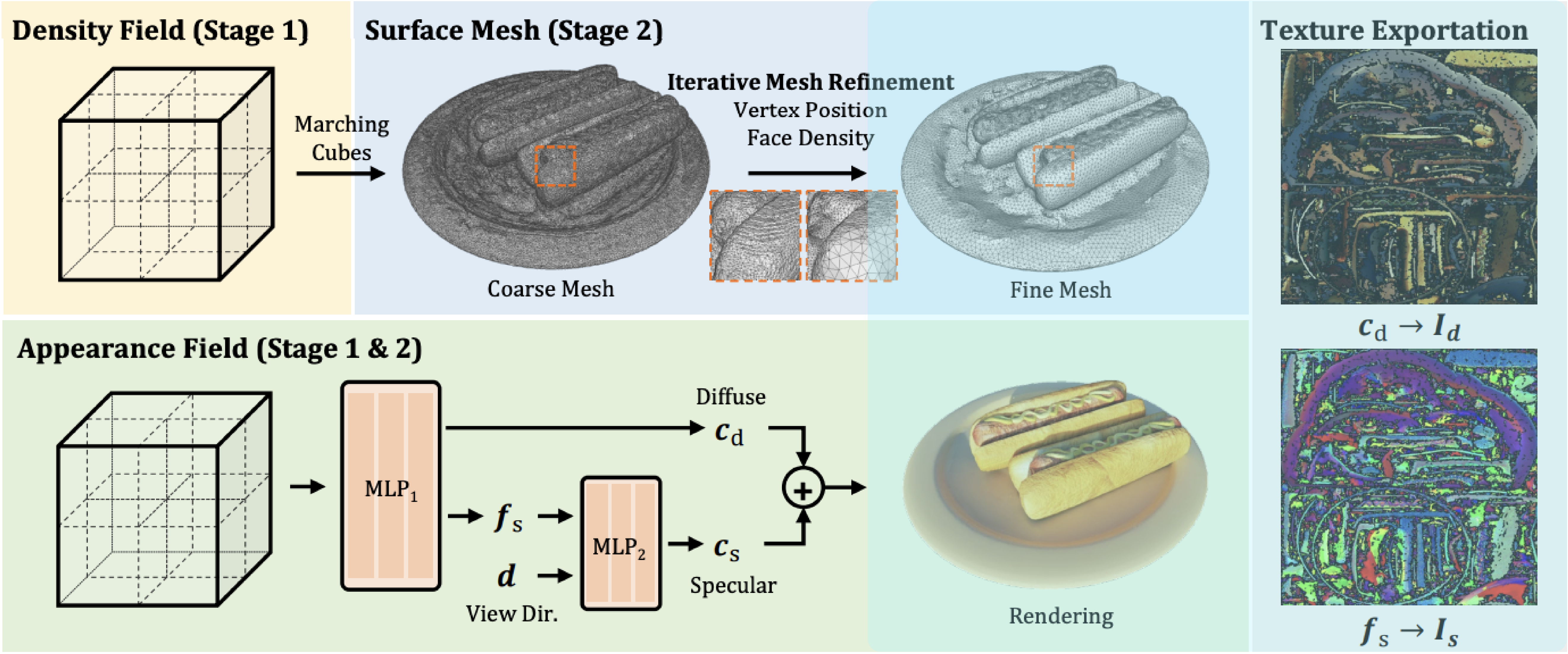
→ Initially learned from density grid, Marching Cube to obtain coarse mesh

2. Appearance Decomposition

→ 5D implicit function operates under no assumption of illumination or material properties



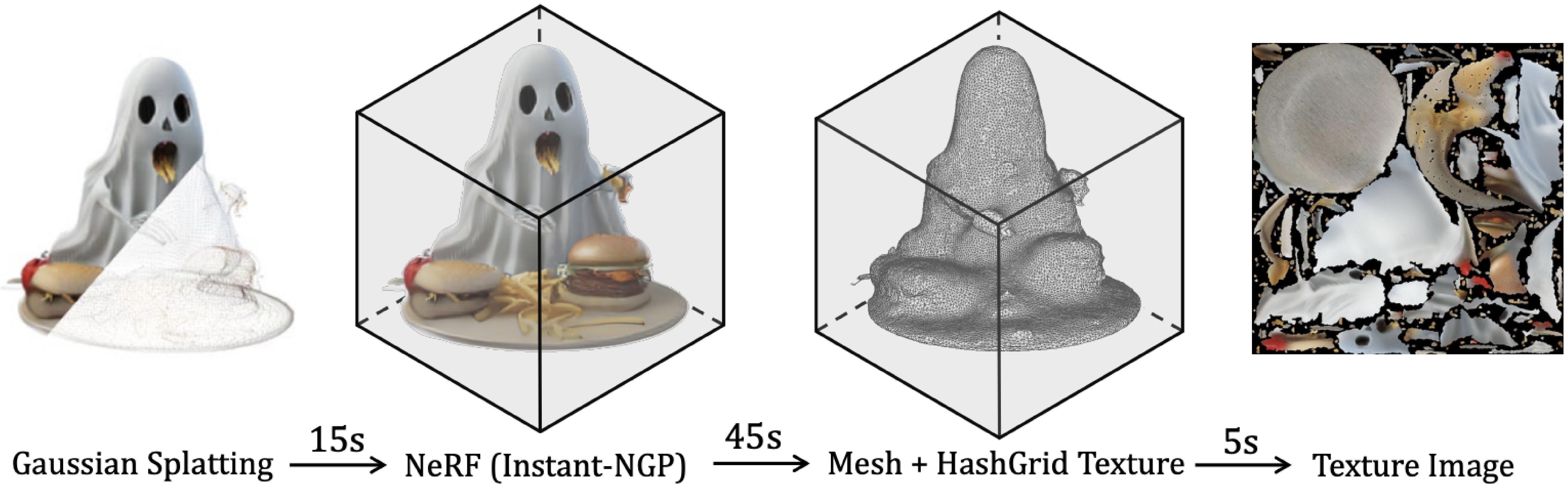
Prerequisites: NeRF2Mesh



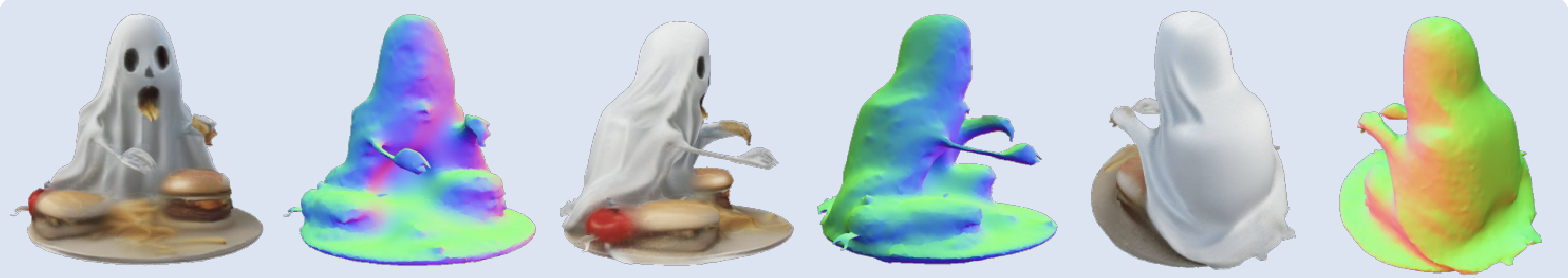
Mesh Exportation:

- Unwrap the UV coordinates of Fine Mesh
- Bake the surface's diffuse color and specular features into two separate images

LGM: Mesh Extraction

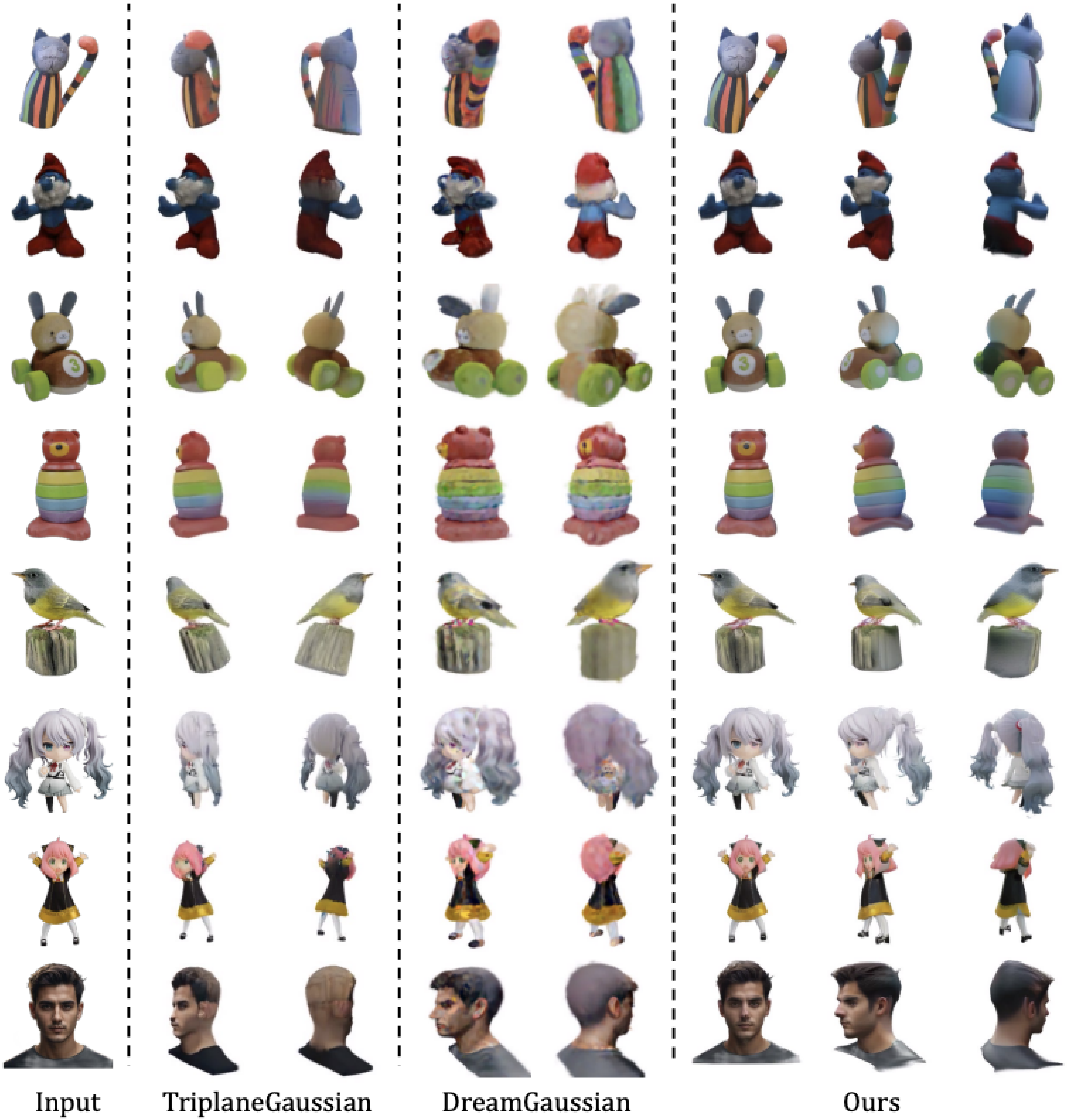


Gaussian Splatting $\xrightarrow{15s}$ NeRF (Instant-NGP) $\xrightarrow{45s}$ Mesh + HashGrid Texture $\xrightarrow{5s}$ Texture Image

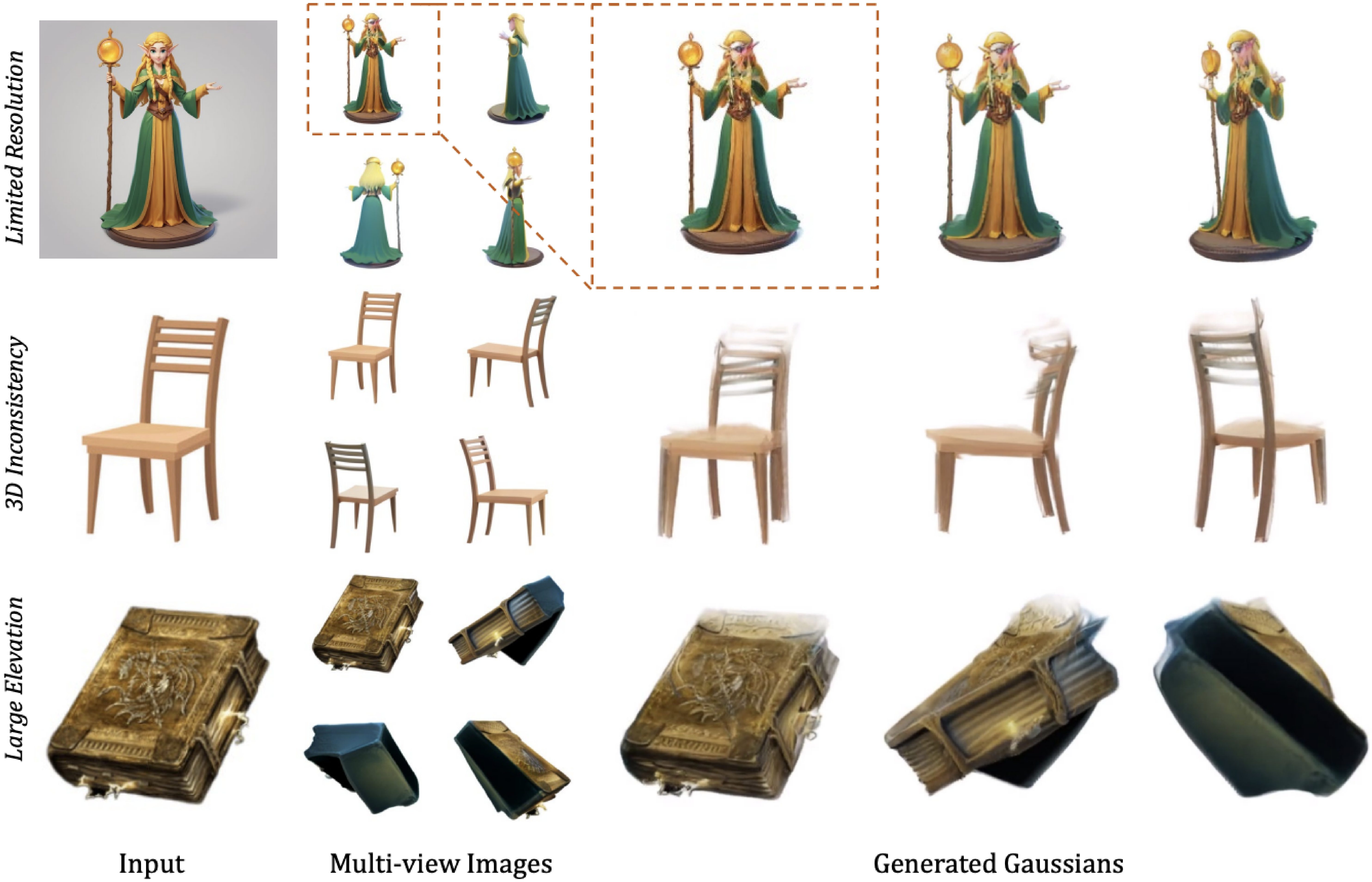


Final Extracted Mesh

Result



Limitation


Limited Resolution
3D Inconsistency
Large Elevation
Input
Multi-view Images
Generated Gaussians

Quiz