NeurIPS 2024

GSGAN: Adversarial Learning for Hierarchical Generation of 3D Gaussian Splats

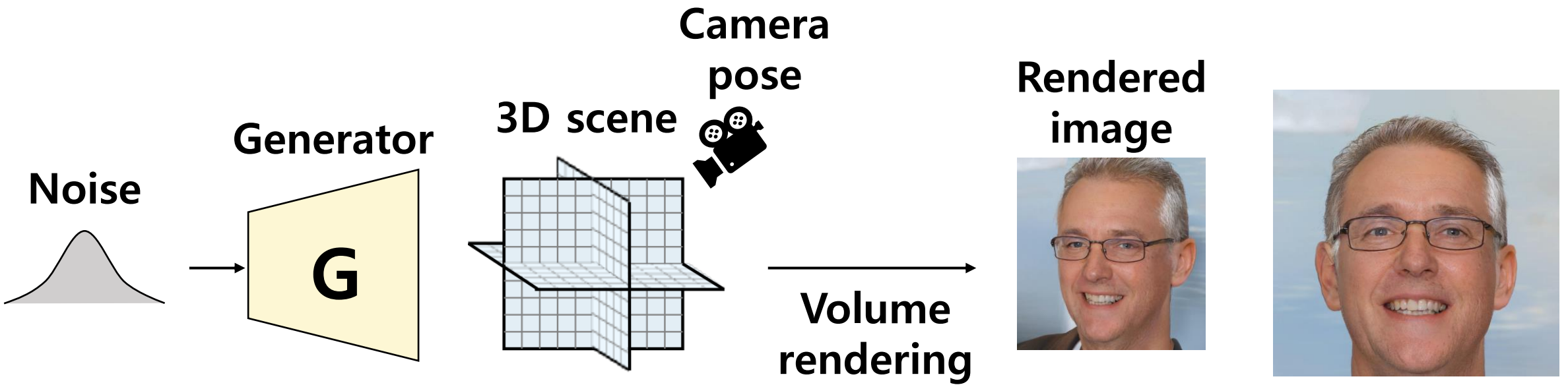
Sangeek Hyun, Jae-Pil Heo

Visual Computing Lab

Sungkyunkwan University

3D Generative Adversarial Networks

- 3D Generative Adversarial Networks (3D GANs) aims to train the generator that synthesizes 3D scene



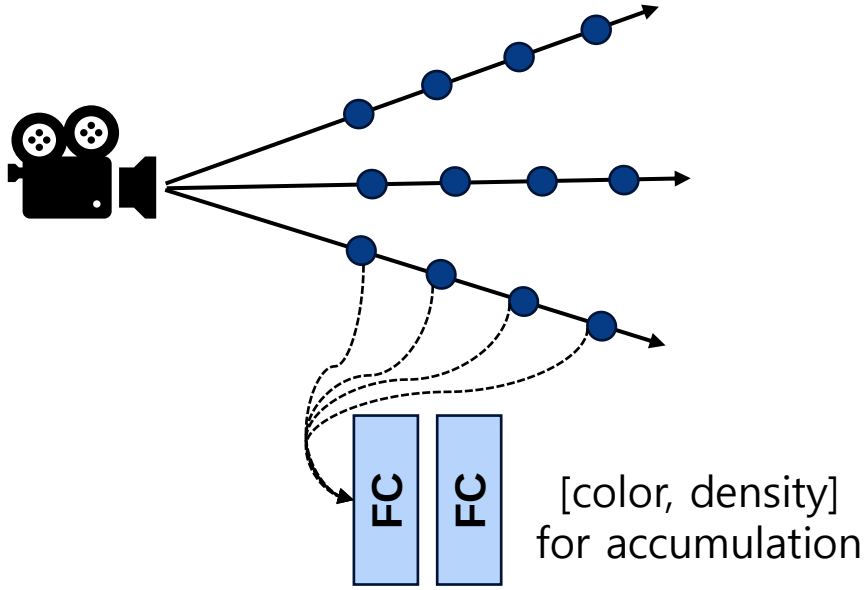
Motivation

- Previous 3D GANs utilize **ray marching** for volume rendering which is **time-consuming**
- Leverage rasterization-based Gaussian Splatting as 3D representation for **fast rendering**

Ray marching (NeRF)

- Cast the rays on every pixel in image plane
- Sampling points on ray and accumulate them

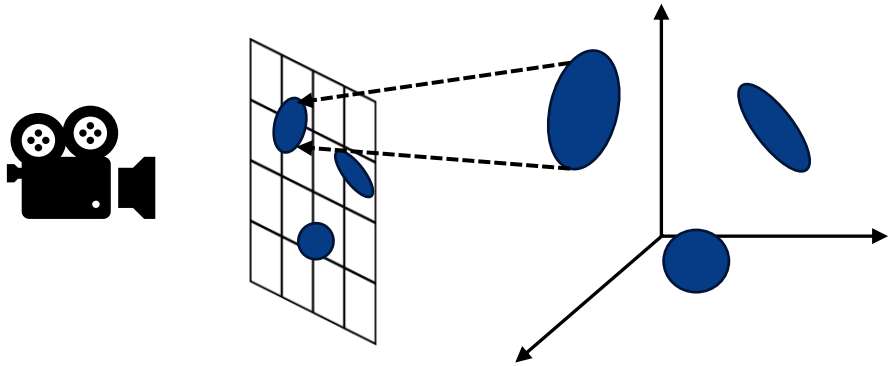
- Slow rendering speed



Rasterization (3D Gaussian Splatting)

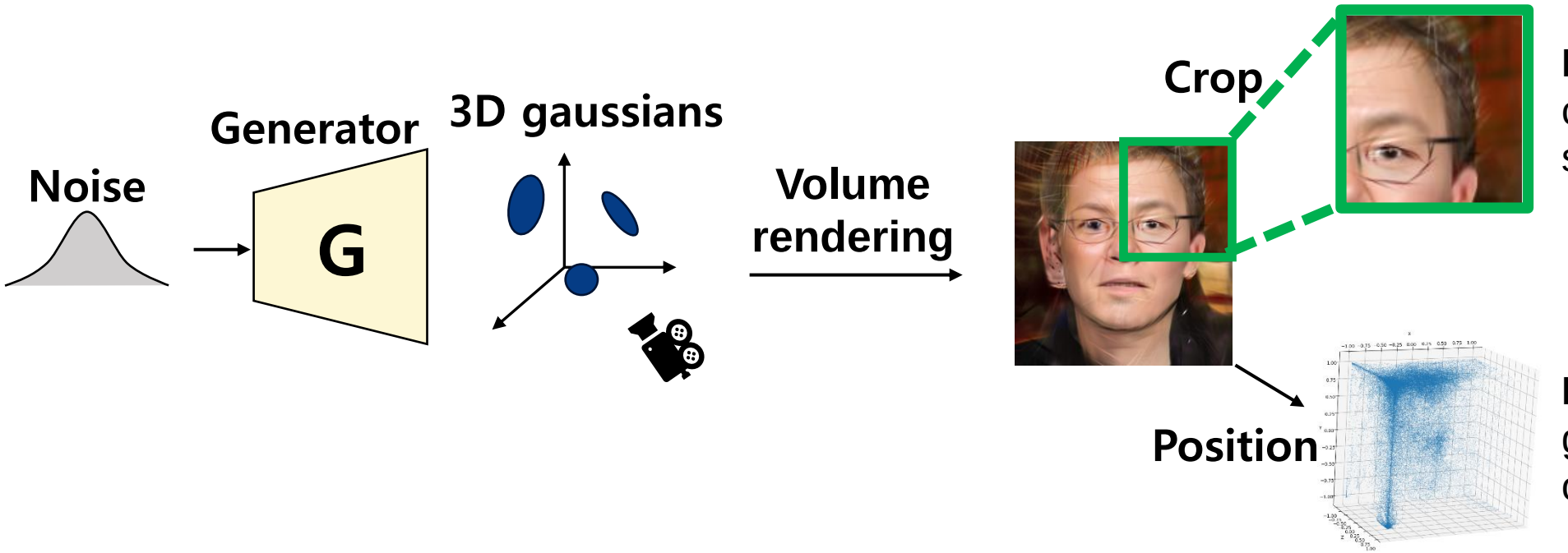
- Project geometric shapes on image plane
- Sorting and alpha compositing them

+ Fast rendering speed



Problem of 3D GANs with Gaussian Splats

- ✓ Naïve generator architecture fails to synthesize visually plausible 3D scenes
  - Elongated scale of gaussians; the shape of gaussians is elongated
  - Improper position of gaussians; the location of most gaussians resides outside of scene
- ! Introduce regularizations for resolving above problems

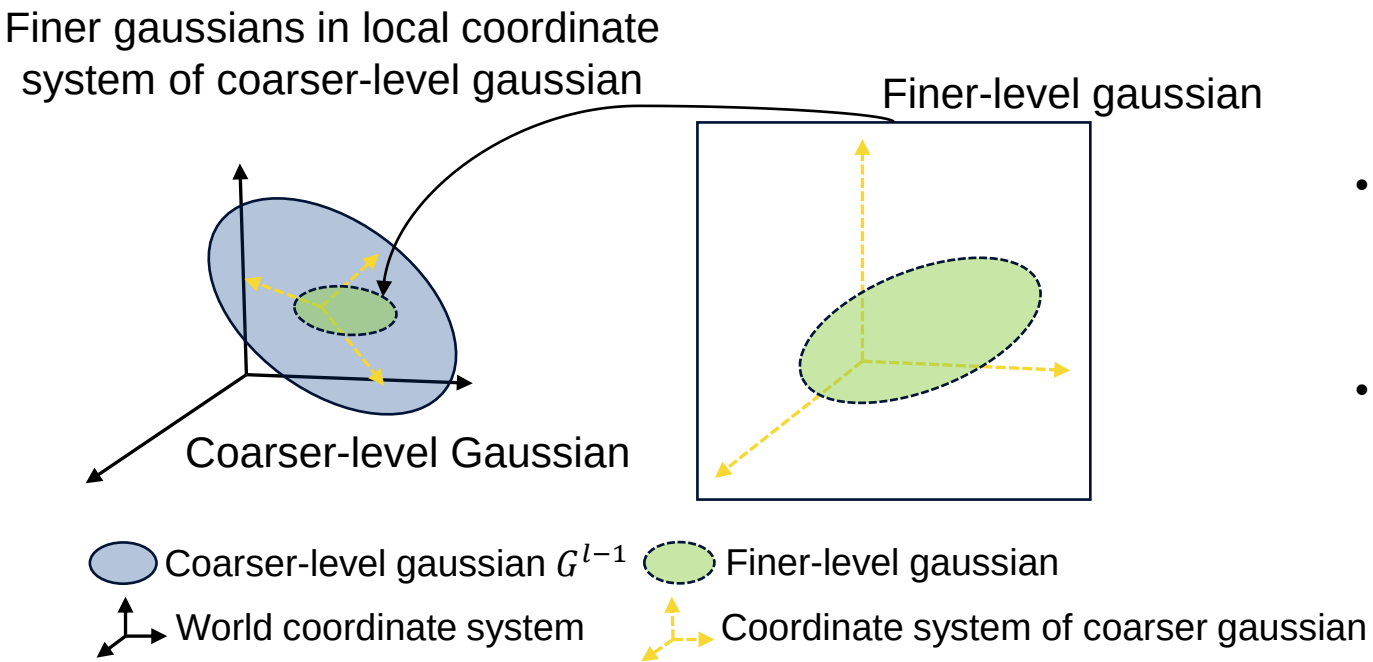


Elongated artifact caused by large scale of gaussians

Improper position; gaussians are located outside of scene bound

Hierarchical 3D Gaussian representation

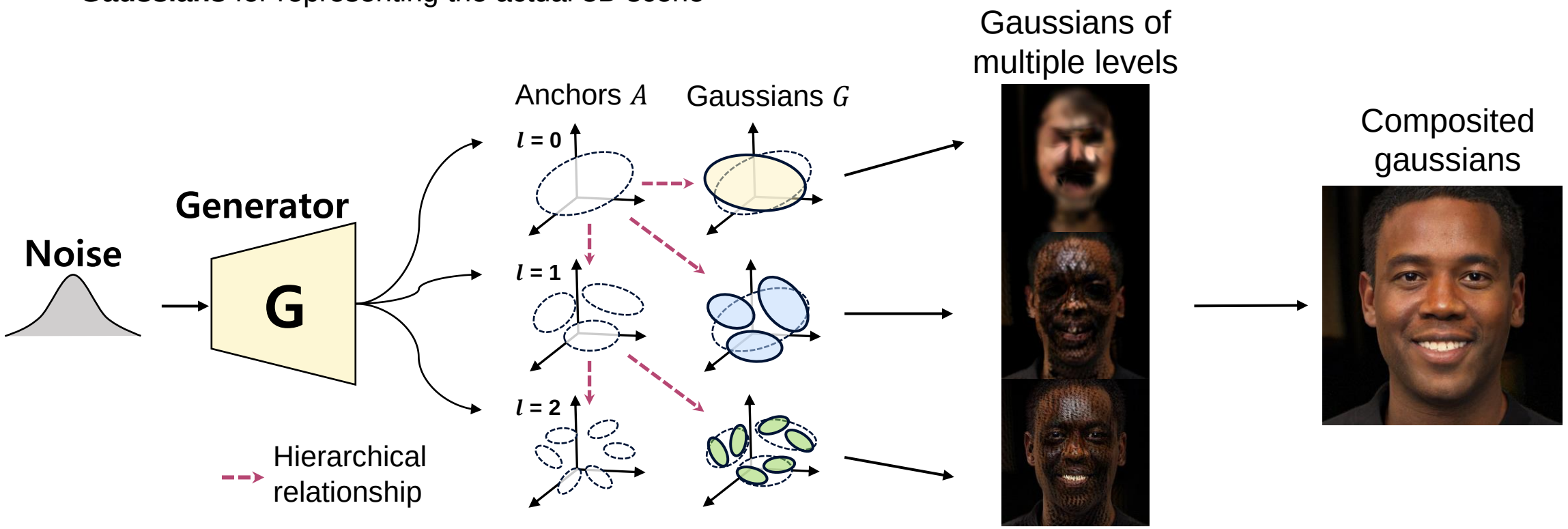
- ✓ Propose the hierarchical gaussian representation such that
  - Position of finer level should be located nearby its coarser counterpart
  - Scale of finer level should be smaller than its coarser counterpart
- ✓ Generator models the scene in a coarse-to-fine manner by this regularization



- The position of gaussian would reside nearby the coarser counterparts
  → **prevent the divergence of position**
- The scale of gaussian is strictly reduced as the level increase
  → **prevent the elongated artifacts**

GSGAN: Generator with hierarchical Gaussians

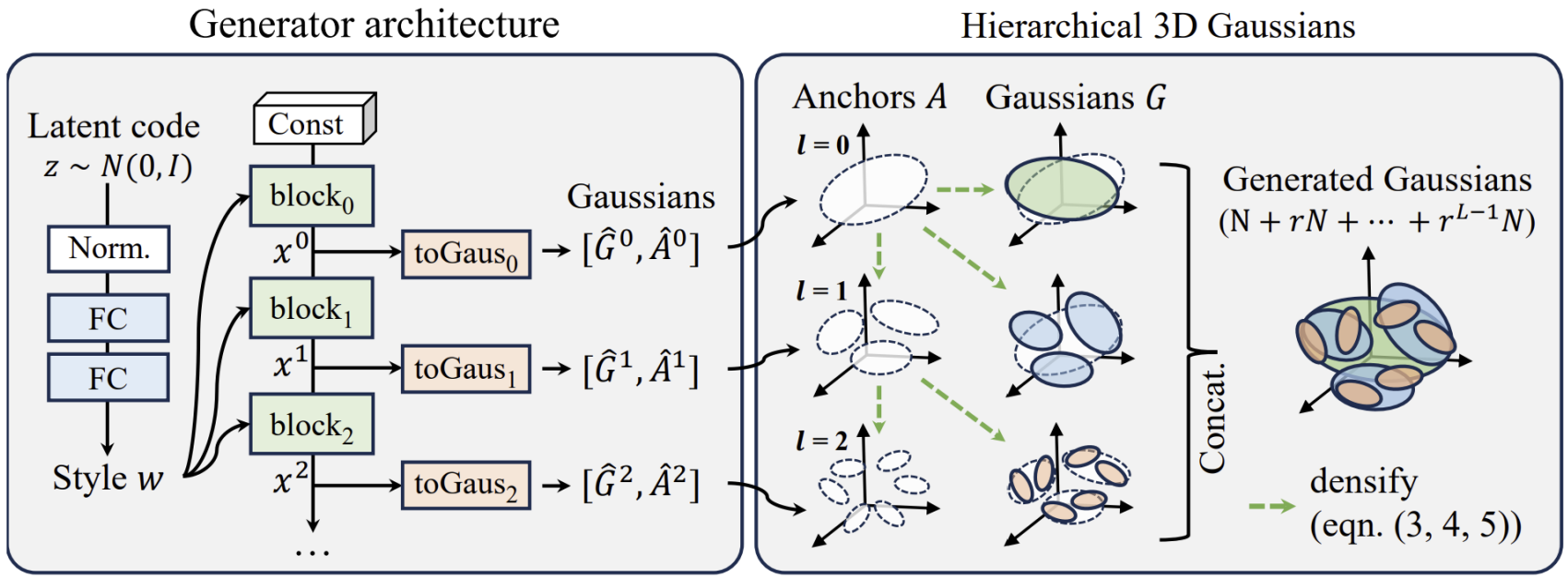
- ✓ Generator synthesizes the hierarchical gaussians and composite them
- ✓ Hierarchical gaussians consist of two types of gaussians
  - **Anchor** for guiding the position and scale of gaussians to be rendered
  - **Gaussians** for representing the actual 3D scene



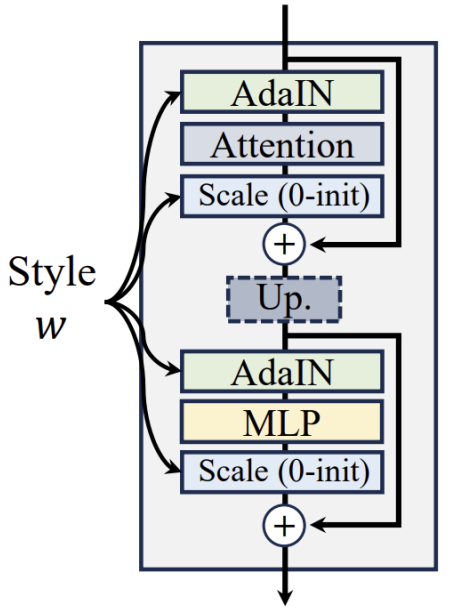
The Proposed Generator architecture

✓ Transformer-based gaussian generator

- Synthesizing gaussians from coarsest to finest level
- StyleGAN-like generator block conditioned by style code w
- Layer-scale for stabilizing the generator in a early stage of training



(a) Illustration of generator architecture with hierarchical 3D Gaussians



(b) `block` architecture

Experimental Result

✓ Quantitative comparison in terms of FID and rendering time

– Ours achieves comparable quality of rendered image (FID) while reporting much faster rendering speed (x100)

Methods	3D consistency	FFHQ 256×256	FFHQ 512×512	AFHQ-Cat 256×256	AFHQ-Cat 512×512	Rendering time (ms) 256×256	Rendering time (ms) 512×512
EG3D [12]		4.80	4.70	3.41	2.72	-	15.5*
GRAM [13]	✓	13.8	-	13.4	-	-	-
GMPI [49]	✓	11.4	8.29	-	7.67	-	-
EpiGRAF [15]	✓	9.71	9.92	6.93	-	-	-
Voxgraf [24]	✓	9.60	-	9.60	-	-	-
GRAM-HD [50]	✓	10.4	-	-	7.67	173.0	197.9
Mimic3D [25]	✓	**5.14**	**5.37**	<u>4.14</u>	<u>4.29</u>	106.8	402.1
Ours	✓	<u>6.59</u>	<u>5.60</u>	**3.43**	**3.79**	**2.7**	**3.0**

Reported score is FID except rendering time

Experimental Result

✓ Generated examples of 3D scenes

FFHQ dataset

AFHQ-Cat dataset

Experimental Result

✓ Visualization of gaussians for each level

FFHQ dataset

AFHQ-Cat dataset

Experimental Result

✓ w-space interpolation & Novel view synthesize from a real-world image by w+ inversion

W-space interpolation

Novel view synthesis
(Top – input / Bottom – synthesized result)

SUNG KYUN KWAN UNIVERSITY(SKKU)

Thank you for watching!

- **Project page: https://hse1032.github.io/gsgan**