



Aesthetic Camera Viewpoint Suggestion with 3D Aesthetic Field

Sheyang Tang¹, Armin Shafiee Sarvestani¹, Jialu Xu¹, Xiaoyu Xu²

Zhou Wang¹

¹University of Waterloo, ²City University of Hong Kong



UNIVERSITY OF
WATERLOO



香港城市大學
City University of Hong Kong


CVPR
JUNE 3-7, 2026

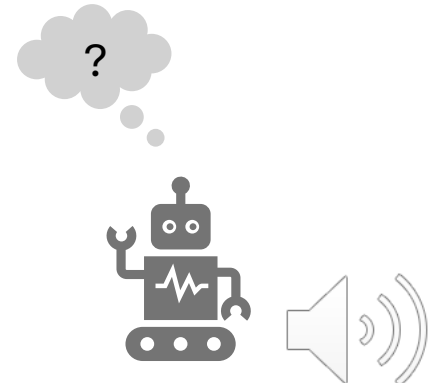
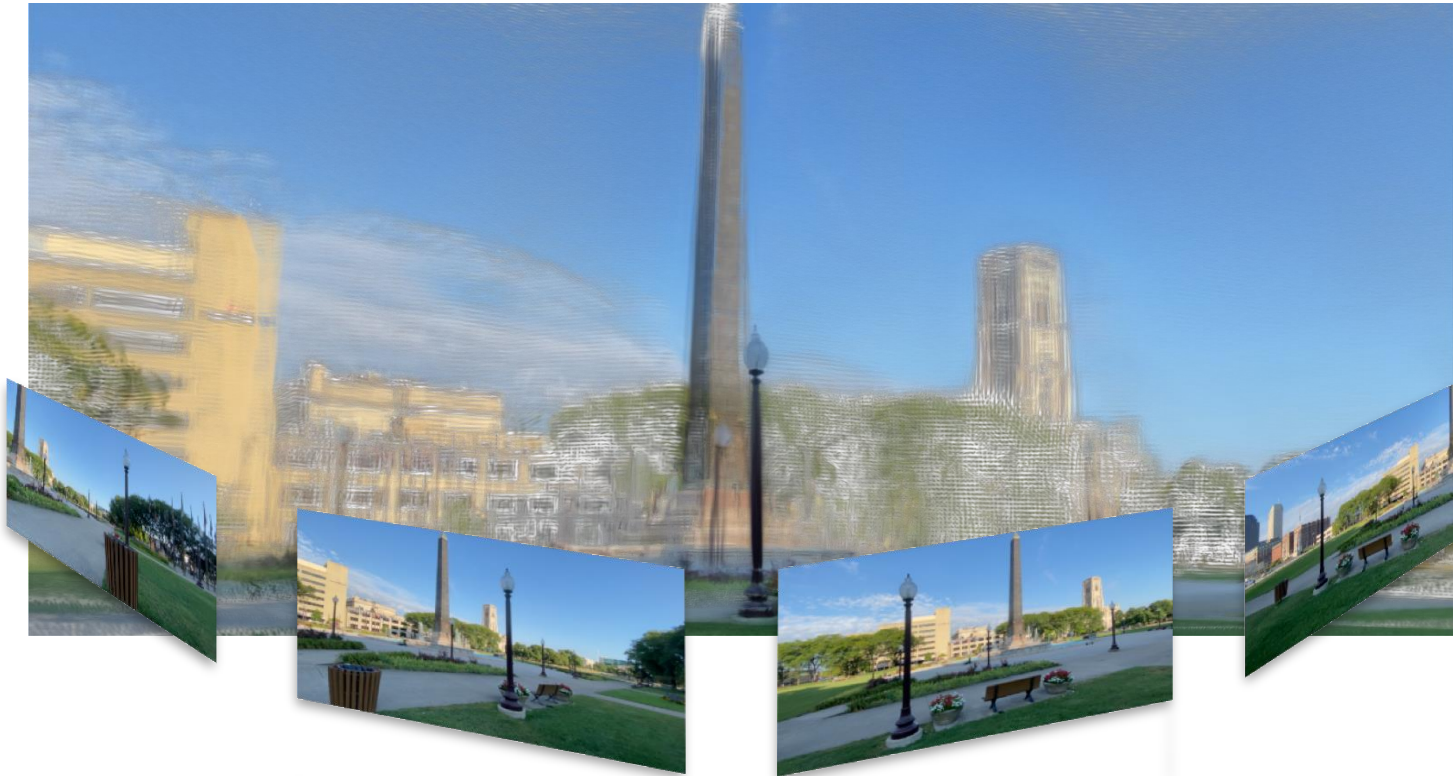


DENVER
COLORADO



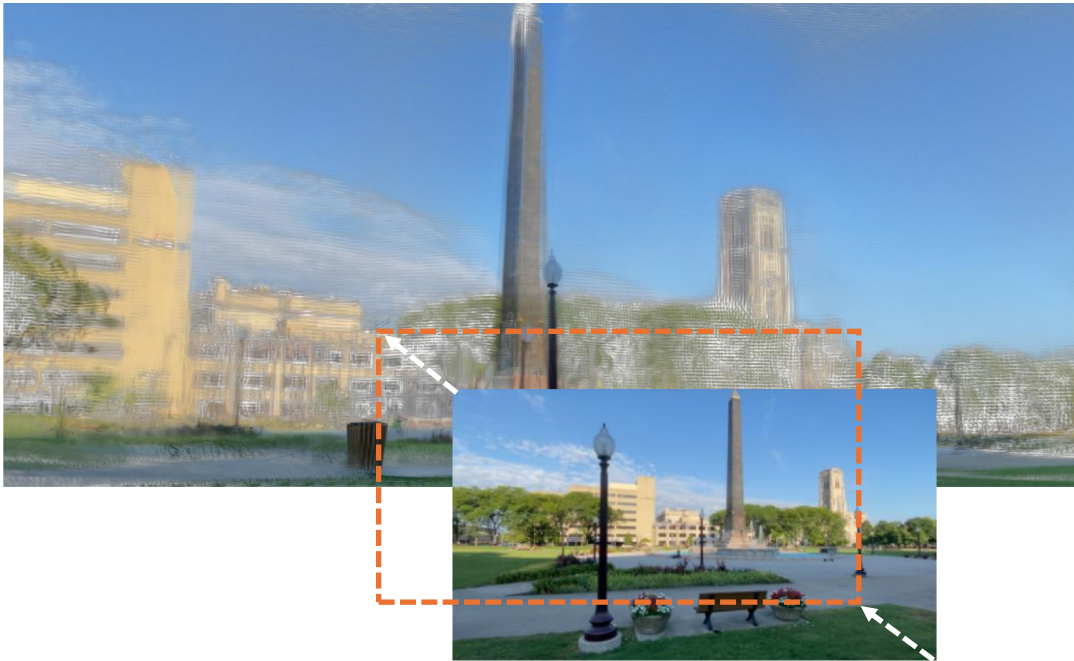
Task

-  How to find the camera pose that gives the best *framing and composition* of a *3D scene*, given only *few observations*?

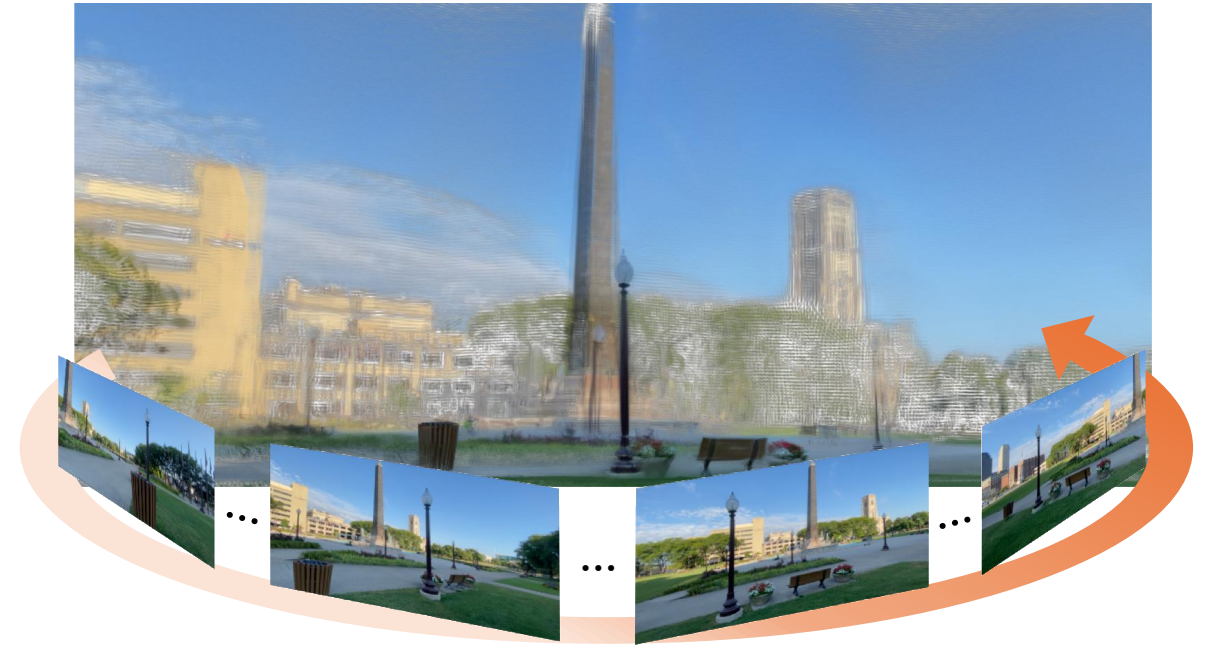


Solution

- **✗** Existing: either single view adjustment or full 3D exploration.



Single view adjustment
Pure 2D; No 3D understanding

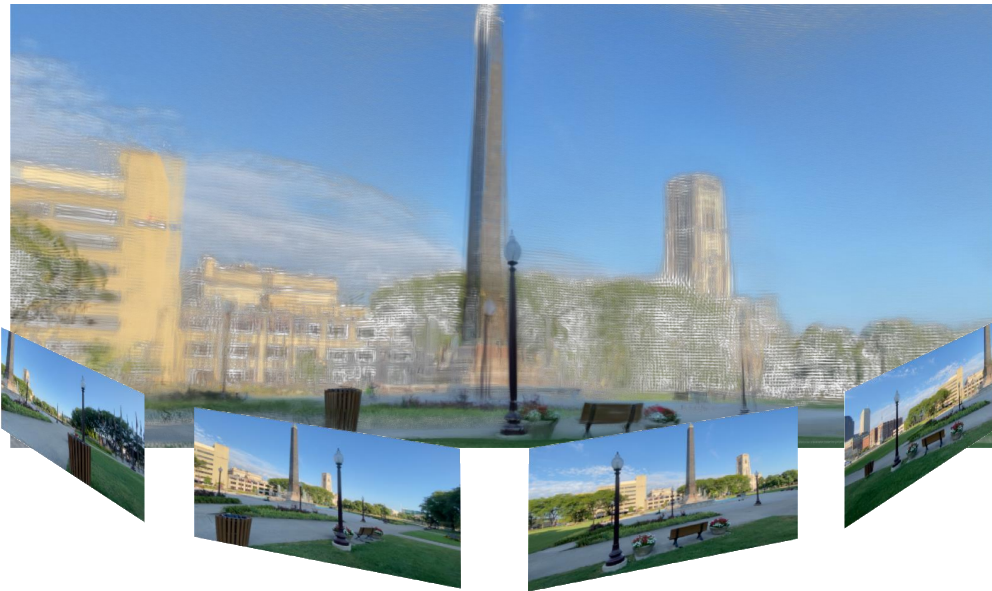


3D exploration
RL search; Needs dense captures



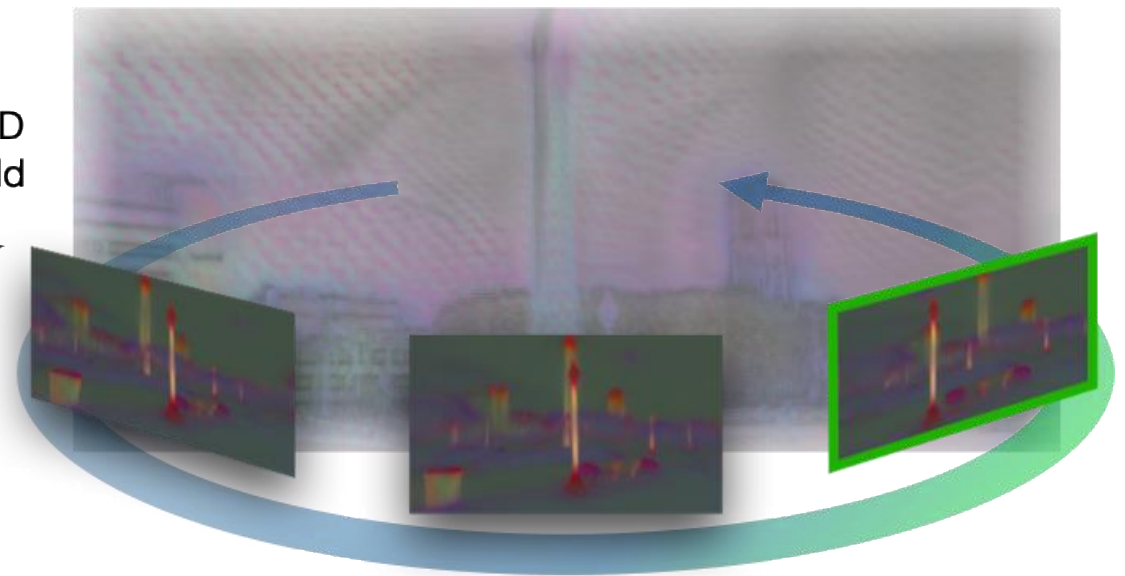
Solution

-  Our contribution: A **3D aesthetic field** for fast viewpoint search from very few input views.



Very few scene captures

Learning a 3D
Aesthetic Field

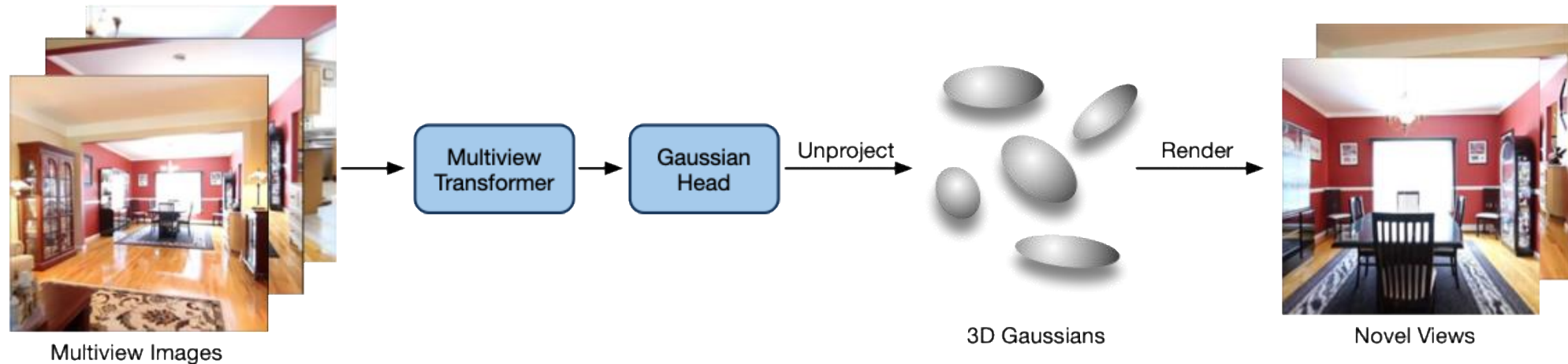


3D Aesthetic Field and Viewpoint Search



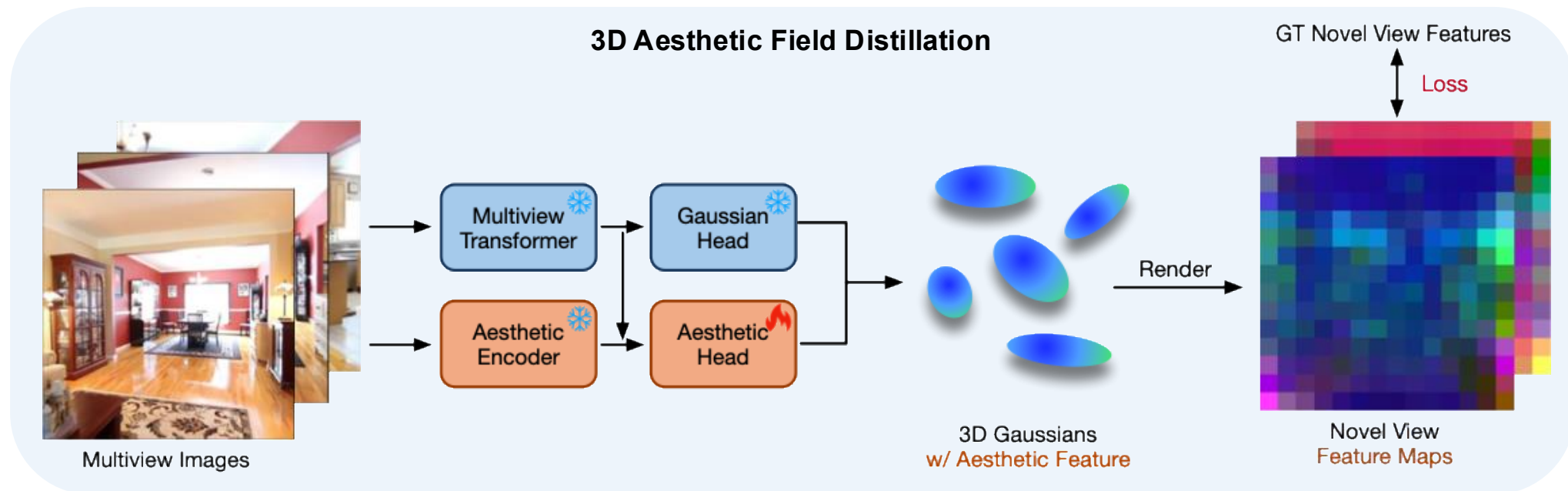
Our method

- We base our framework on feedforward **3D Gaussian Splatting**



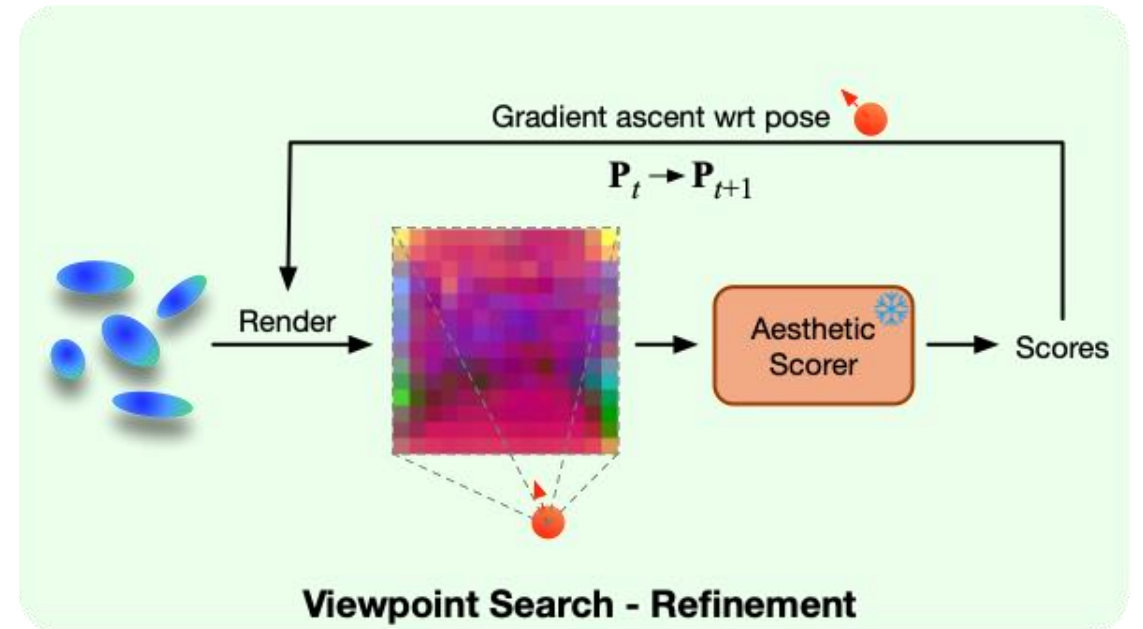
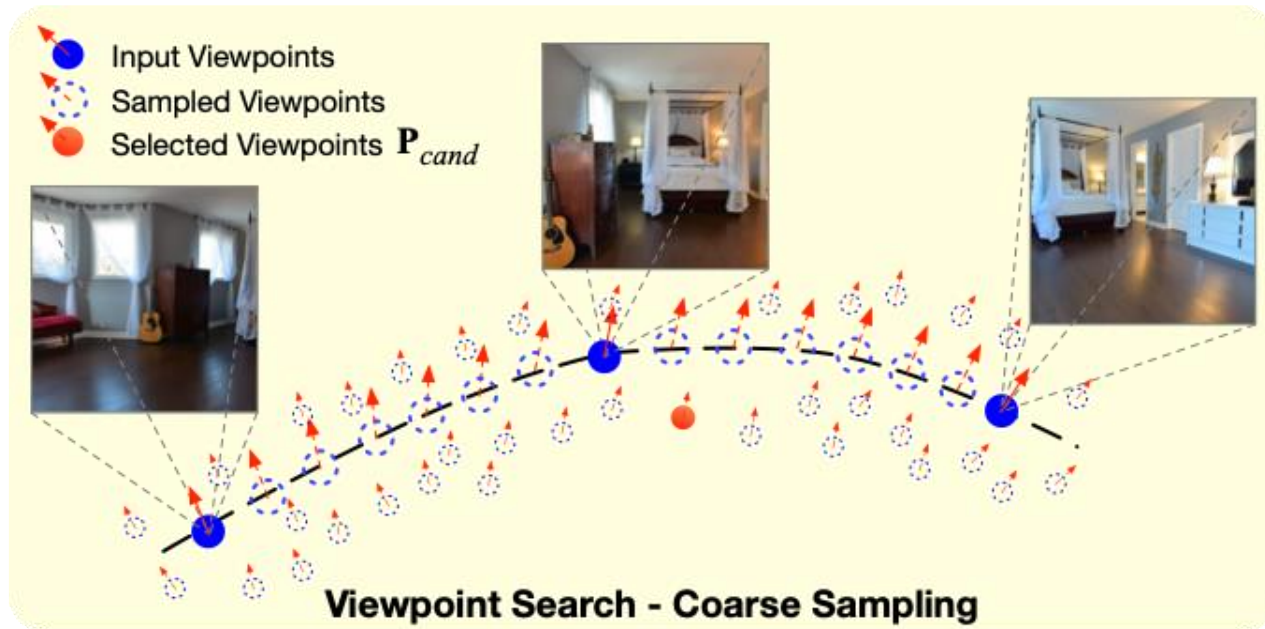
Our method

- **Distill** aesthetic knowledge from pretrained image aesthetic models into **3D Gaussians**



Our method

- Two-stage **viewpoint search** pipeline:



Experiments & Results:

- We train our model on multi-view datasets.
- No existing benchmarks.
- We adapt existing multi-view datasets for evaluation:
 - Novel view **aesthetic quality assessment**.
 - **Aesthetic viewpoint suggestion**: use 2D aesthetic models as judges.



Experiments & Results:

- **Novel view aesthetic quality assessment** w/ sparse input views.
- Baseline: render **RGB images**, then evaluate quality.
- Ours: render **aesthetic features**, then decode to quality.

Methods	#Views	RE10k [320]		DL3DV [127]	
		PLCC	SRCC	PLCC	SRCC
Baseline	2	0.657	0.628	0.326	0.307
Ours		0.780	0.740	0.509	0.477
Baseline	4	0.657	0.633	0.513	0.481
Ours		0.796	0.758	0.722	0.682
Baseline	6	0.745	0.701	0.580	0.553
Ours		0.836	0.794	0.753	0.719

(a) Image resolution 256×256 .

Methods	#Views	RE10k [320]		DL3DV [127]	
		PLCC	SRCC	PLCC	SRCC
Baseline	2	0.634	0.611	0.299	0.286
Ours		0.733	0.704	0.479	0.445
Baseline	4	0.596	0.578	0.552	0.535
Ours		0.753	0.722	0.700	0.668
Baseline	6	0.711	0.656	0.646	0.625
Ours		0.821	0.764	0.737	0.707

(b) Image resolution 256×448 .



Experiments & Results:

- **Aesthetic viewpoint suggestion w/ sparse input views.**

<i>RE10k</i> [320]						
Methods	2 Input Views		4 Input Views		6 Input Views	
	VEN \uparrow	SAMPNet \uparrow	VEN \uparrow	SAMPNet \uparrow	VEN \uparrow	SAMPNet \uparrow
Baseline	1.48	2.29	1.79	2.26	2.01	2.29
In-plane Shift*	1.52	2.31	1.81	2.36	2.10	2.41
Rotation*	1.78	2.38	1.95	2.42	2.13	2.45
UNIC [132] \dagger	1.15	2.17	1.61	2.33	1.82	2.35
Uchida <i>et al.</i> [231] \dagger	1.58	2.32	1.89	2.37	2.13	2.42
Ours	1.89	2.40	2.03	2.45	2.20	2.49
<i>DL3DV</i> [127]						
Baseline	2.08	2.23	2.31	2.26	2.47	2.30
In-plane Shift*	2.16	2.28	2.45	2.29	2.59	2.33
Rotation*	2.52	2.30	2.67	2.32	2.85	2.37
UNIC [132] \dagger	2.07	2.21	2.35	2.27	2.48	2.30
Uchida <i>et al.</i> [231] \dagger	2.34	2.25	2.68	2.32	2.81	2.36
Ours	2.56	2.33	2.76	2.36	2.91	2.41



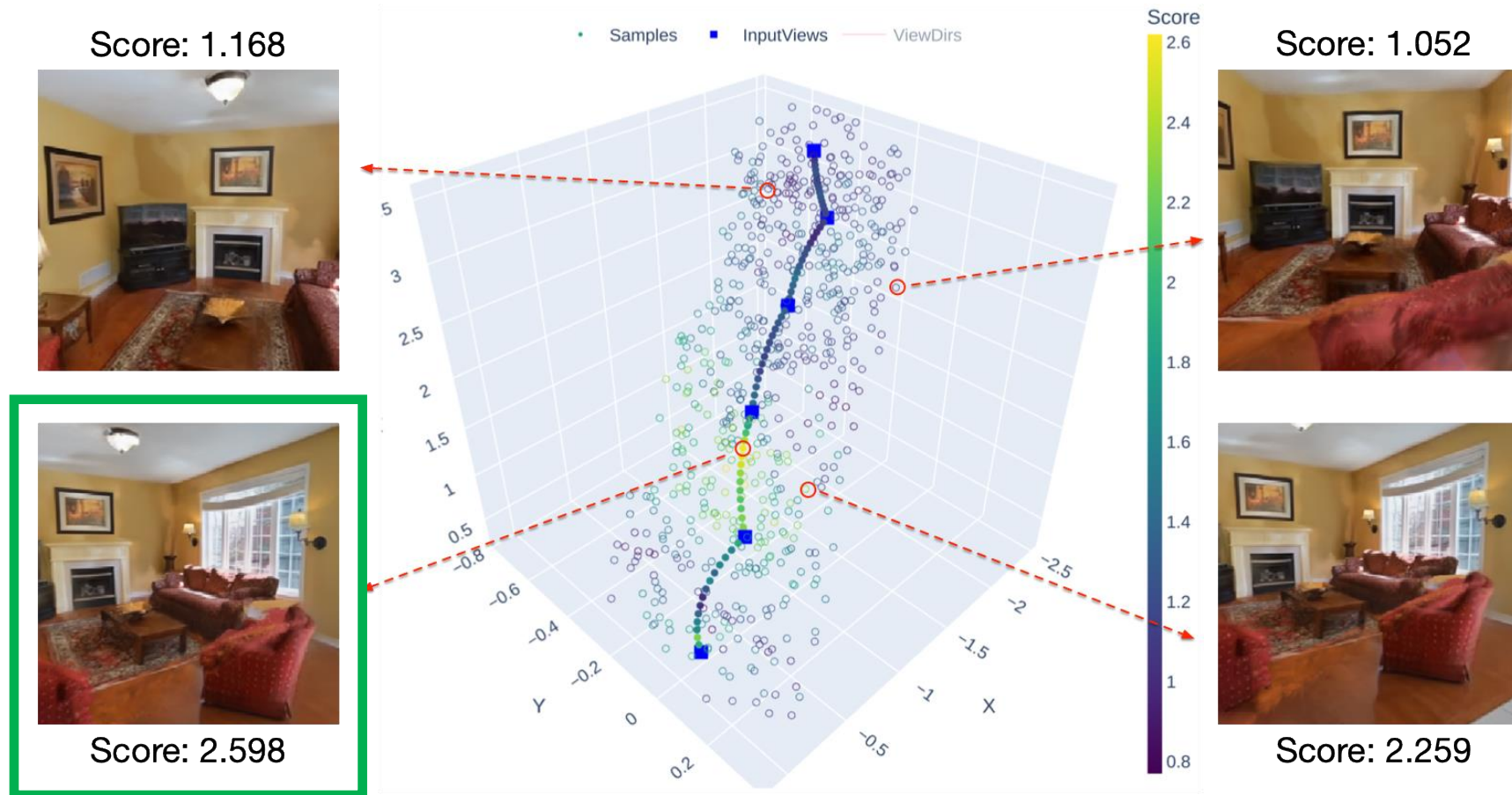
Experiments & Results:

- Our model finds well-composed views that align with human aesthetic preferences.



Experiments & Results:

- Visualization of viewpoint search results in 3D.



Thanks!

Check out our webpage for more results, poster, and implementations.

sheyangtang.com/aesthetic-field/



UNIVERSITY OF
WATERLOO



香港城市大學
City University of Hong Kong

CVPR
JUNE 3-7, 2026



DENVER
COLORADO