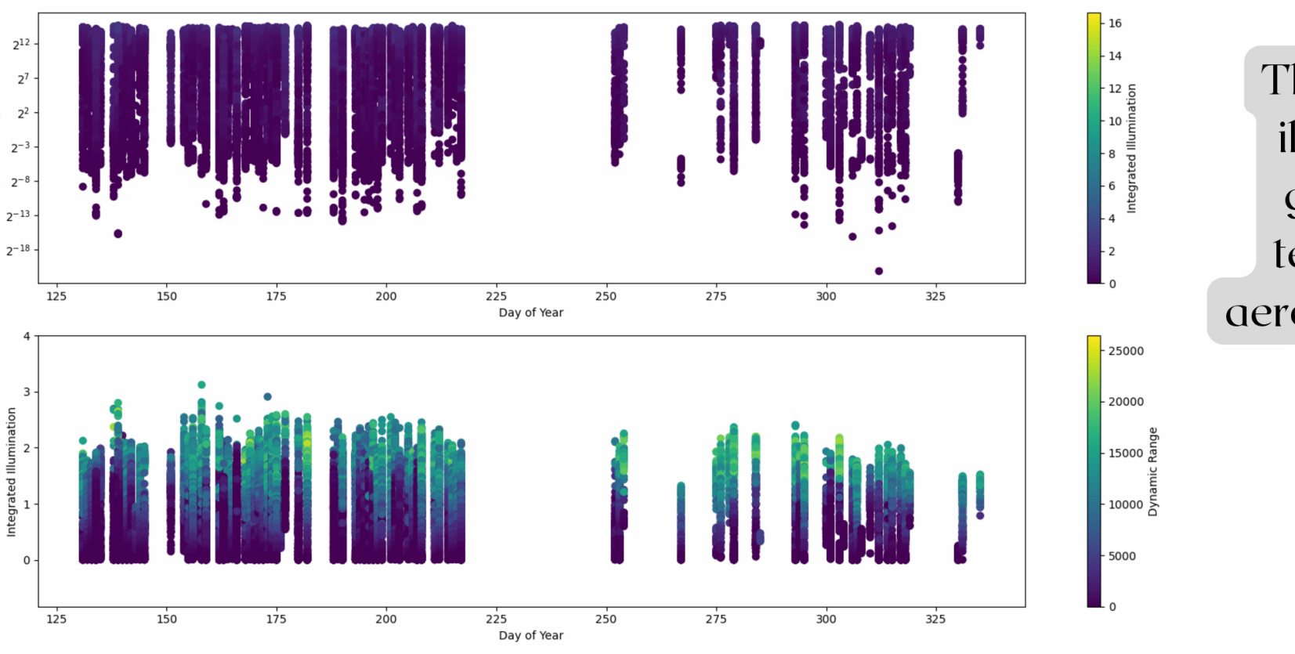
Can We Model Skies?

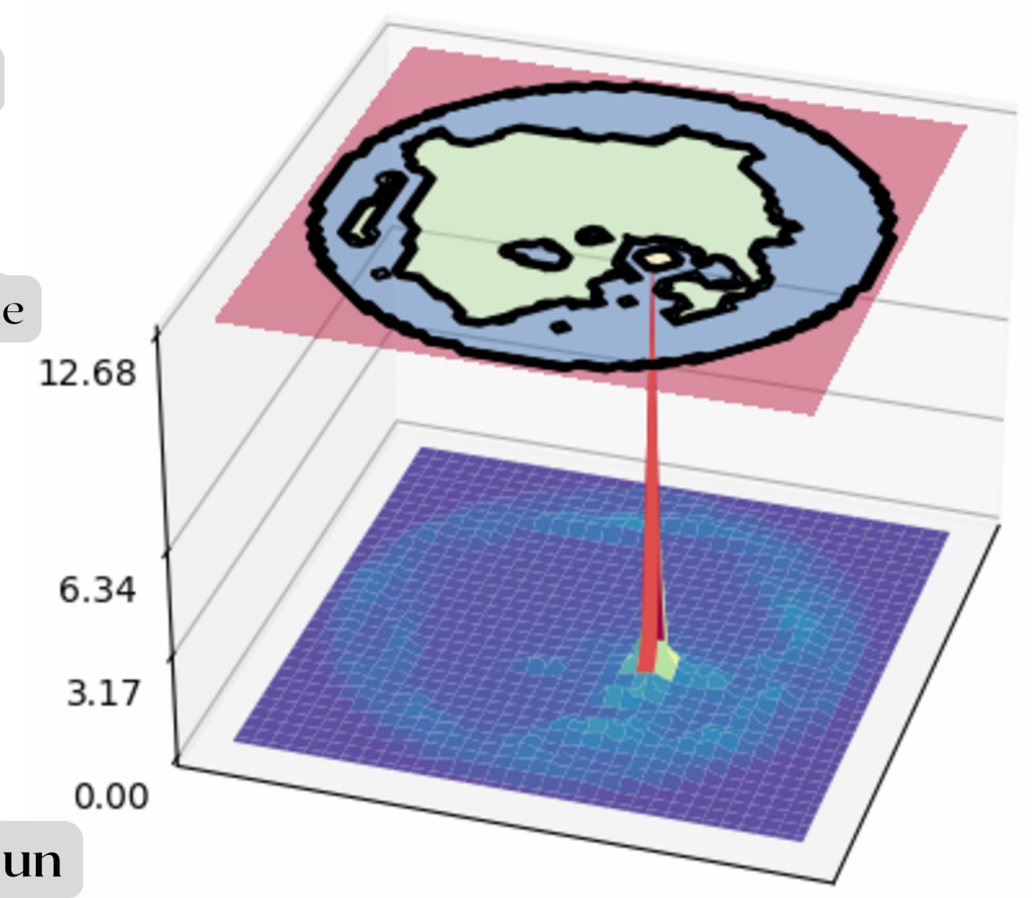
FOR IMAGE BASED LIGHTING

Ian J. Maquignaz, Laval University

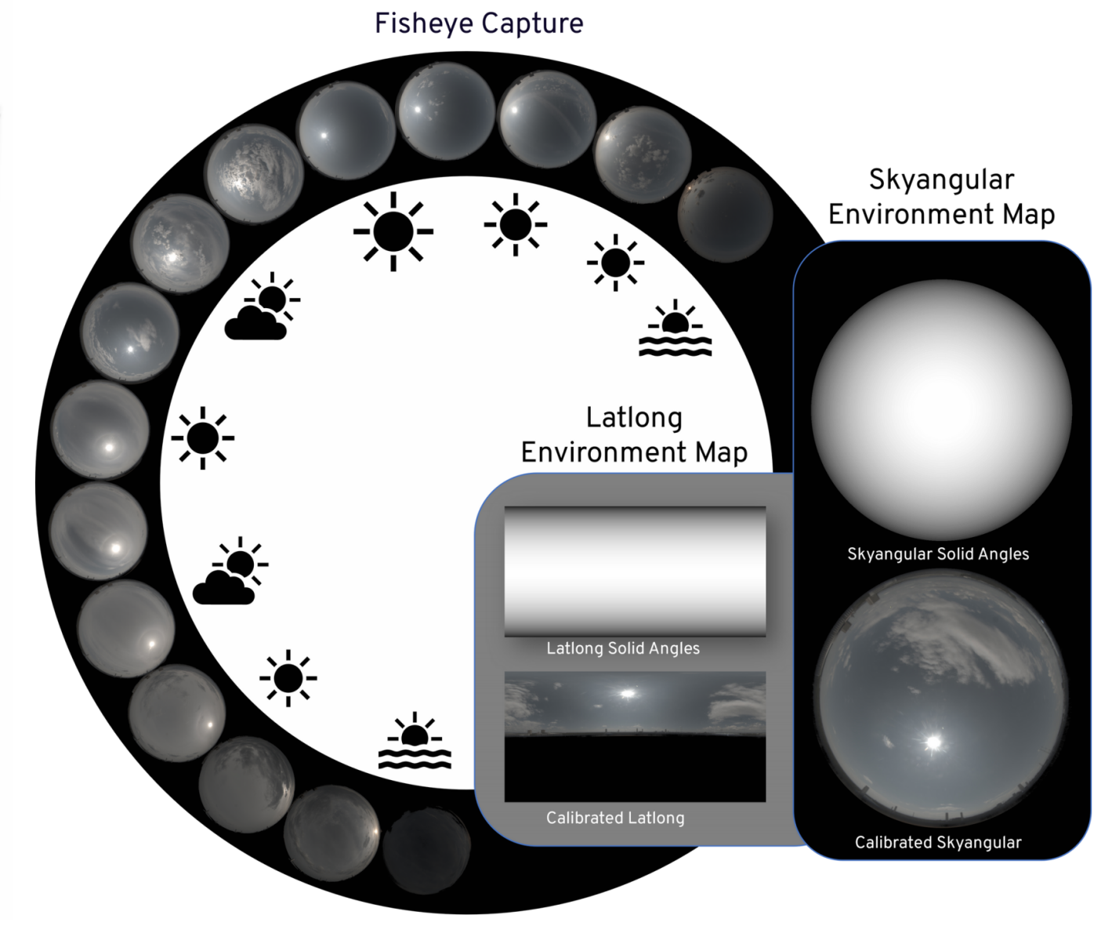
Skydome Dataset

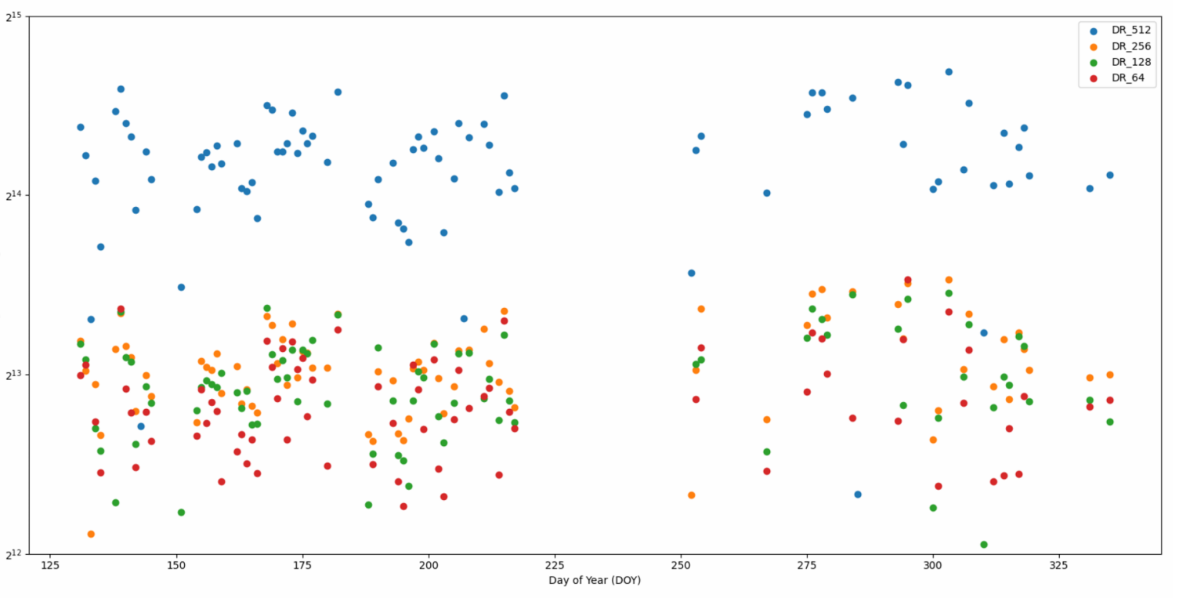


The dynamic range and illumination fluctuates greatly, depending on temperature, pressure, aerosols and other airborne particles.



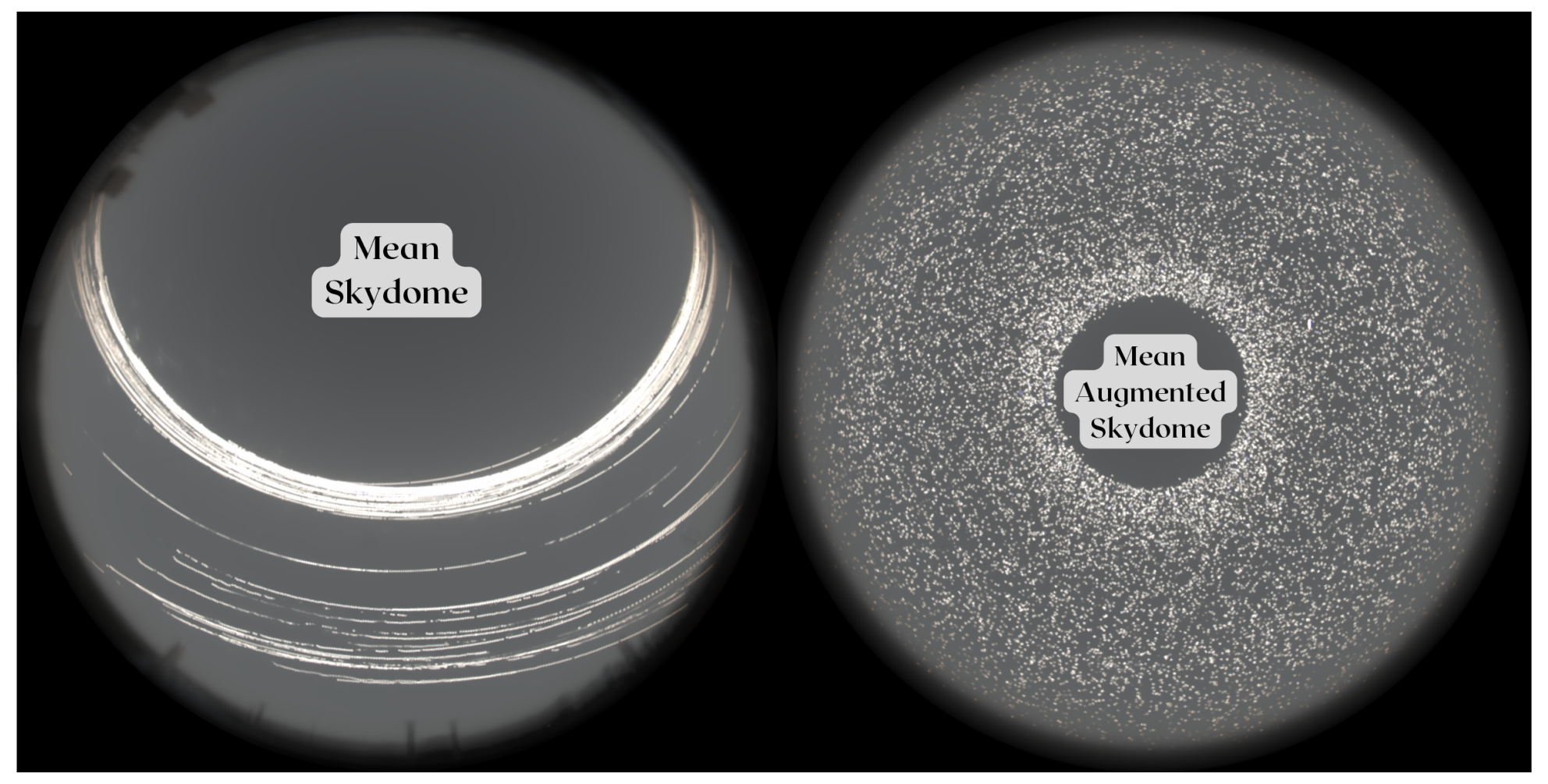
Dynamic range is primarily required for sun pixels, representing less than 5 of skydome pixels, but 60% of illumination.



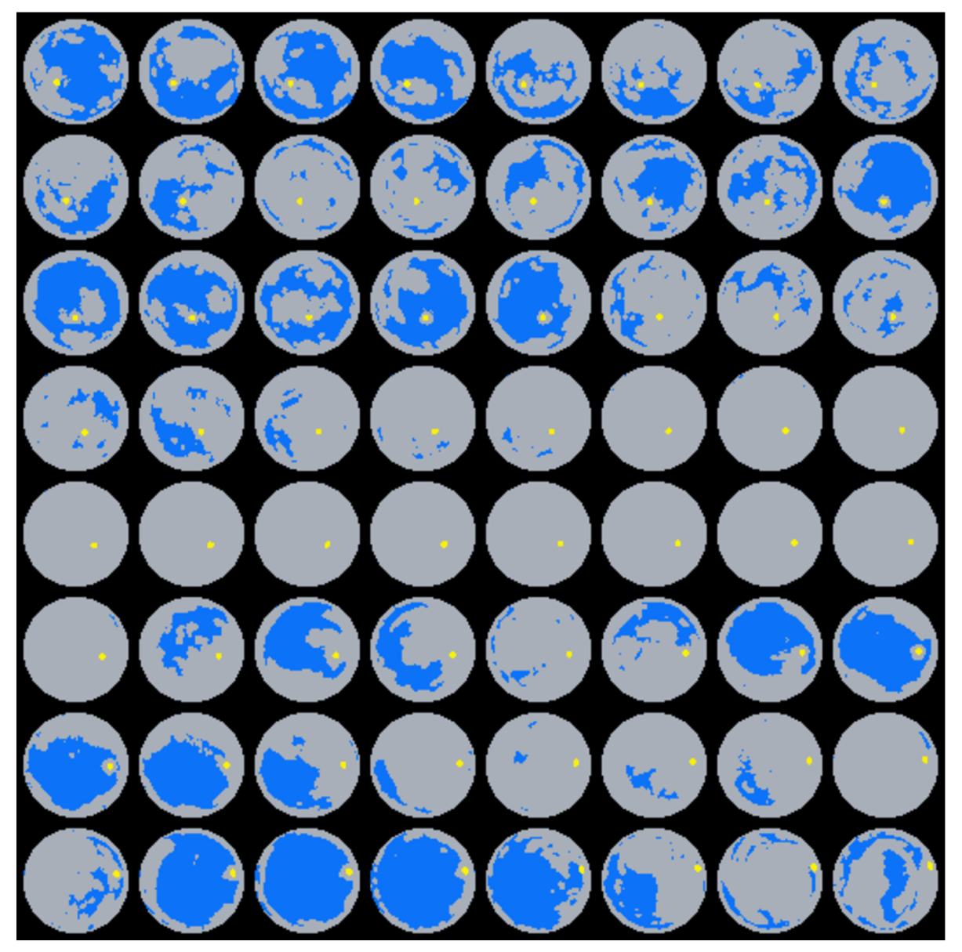


To complicate things, most methods of re-scaling images are more-or-less energy preserving, but do not preserve dynamic range

Augmenting the dataset through rotations and flips around the zenith allows for greater coverage.



Input Modalities



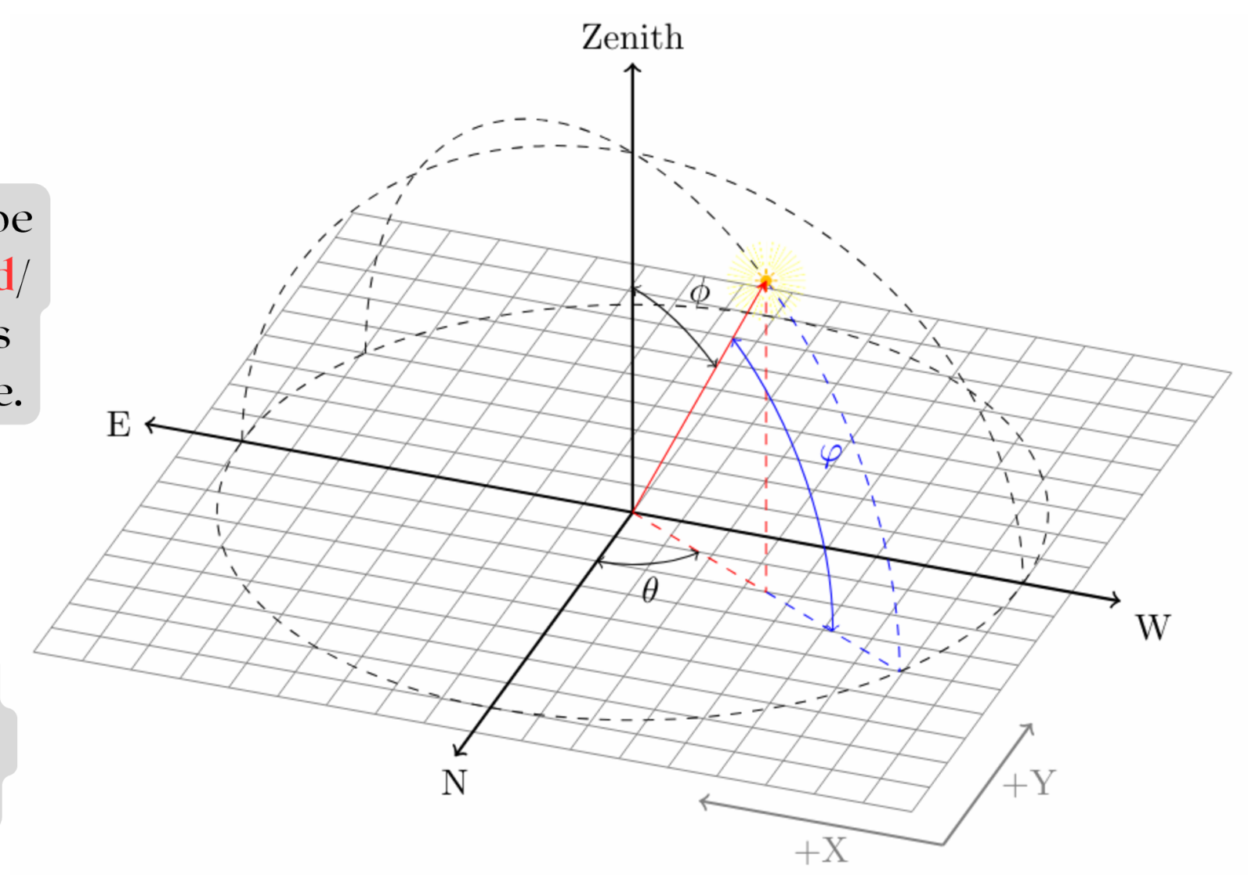
Method #1
"Hand-Drawn" Segmentation Masks

Clouds

Segmentation for cloud formations can be achieved by thresholding the ratio Y=Red/Blue or Y= (Blue-Red)/(Blue+Red). This method is crude, but simple and effective.

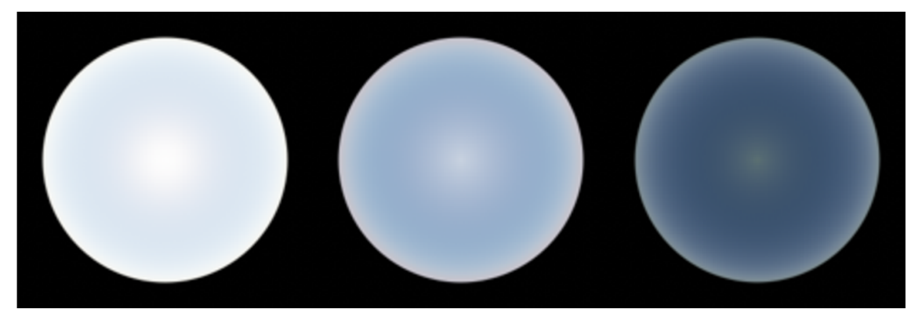
Sun

Solar position can be infered as the brightest spot or calculated via emphemeris tables. The extraterrestrial solar solar disk is 0.5 angular degrees but, with atmospheric attenuation, is visually closer to 5 angular degrees.



Method #2
Clear Sky Augmentation

Many recent models such as DeepClouds, SkyGAN, and LM-GAN elect to use parametric sky-models as priory. This enables users to control the generation of a clear-sky (sometimes inclusive of a solar disk) via minimal set of mostly-intuitive parameters, which can then be augmented by a DNN to include atmospheric formations.

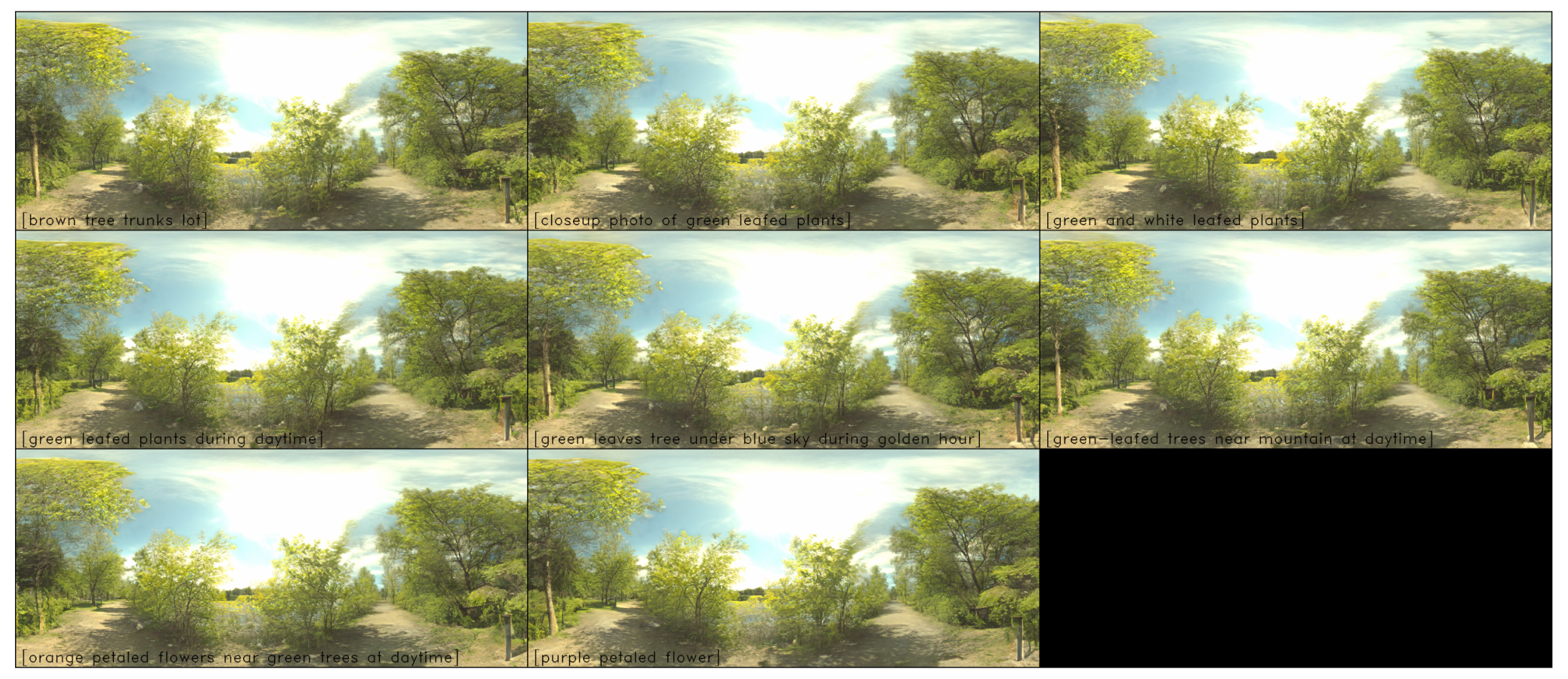


Pretham, Nishita and Hošek-Wilkie Clear Sky Models

Though effective, the viability of these methods is questionable. No proposed augmentation model has been demonstrated to retain the clear-sky parameter configuration nor, if applicable, retain the include solar disk.

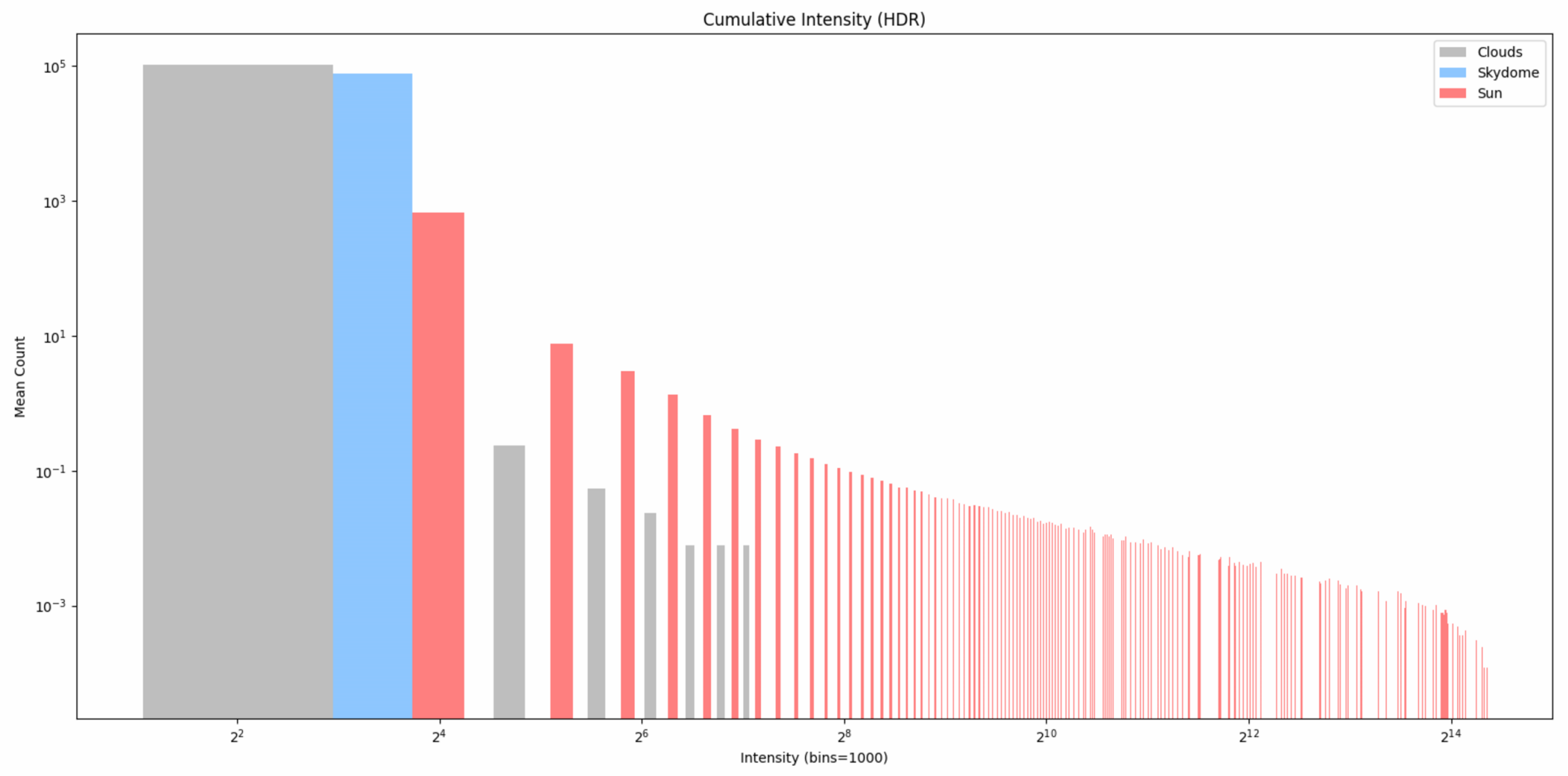
Method #3
Inference from Textual Promps

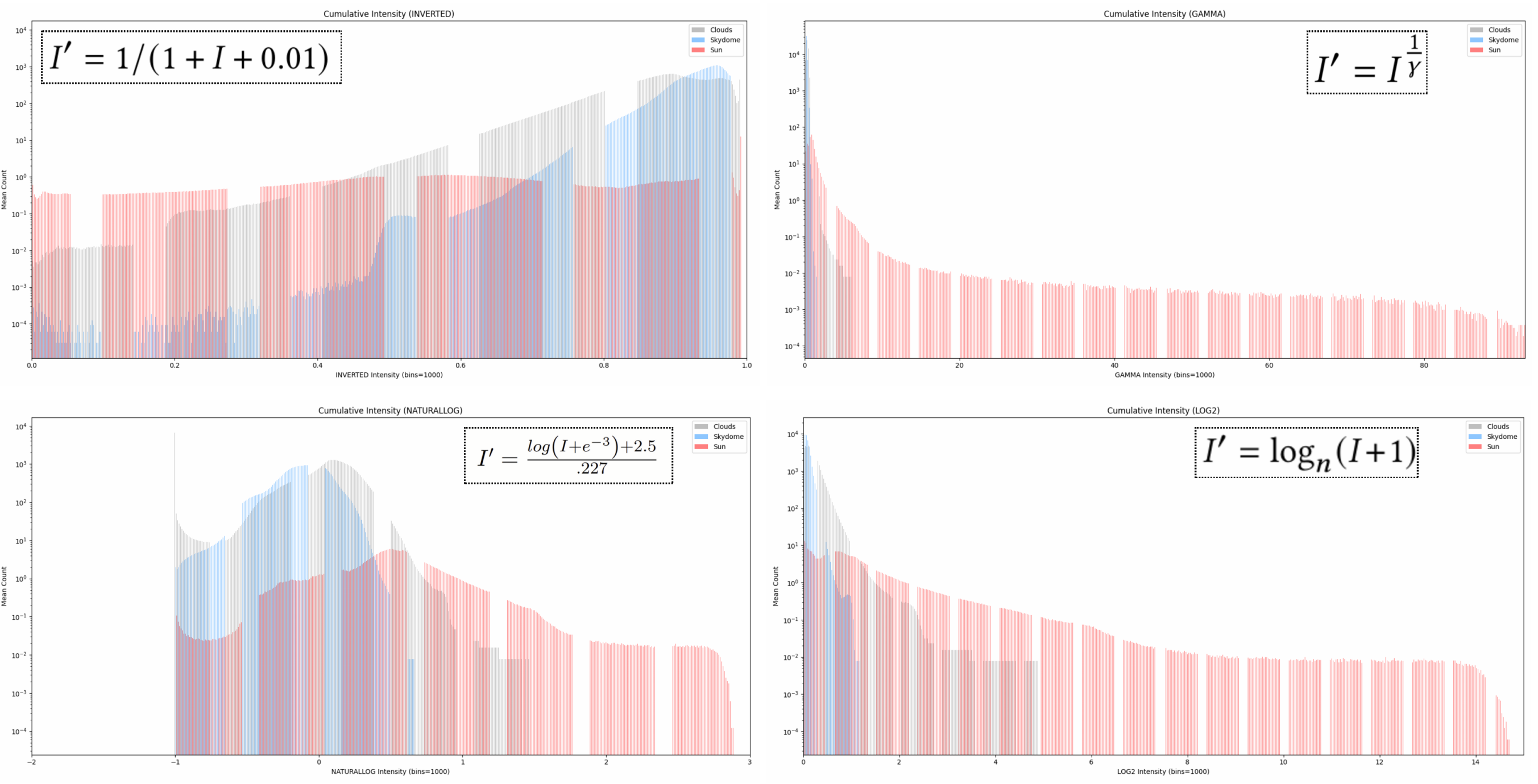
With the emergence of CLIP, models such as Text2Light championed textual prompts for IBL. As shown, this input modality is grossly imprecise, for the language required to accurately describe a skydome is non-colloquial.



The importance of HDR

Coherence of illumination can be key in distinguishing a real- vs. virtual-scene, in particular when 'natural' outdoor illumination is involved. To accurately represent outdoor illumination, High Dynamic Range Imagery (HDRI) with 22-fstops is required to faithfully capture the highlights and shadows of an average real-world outdoor scene.

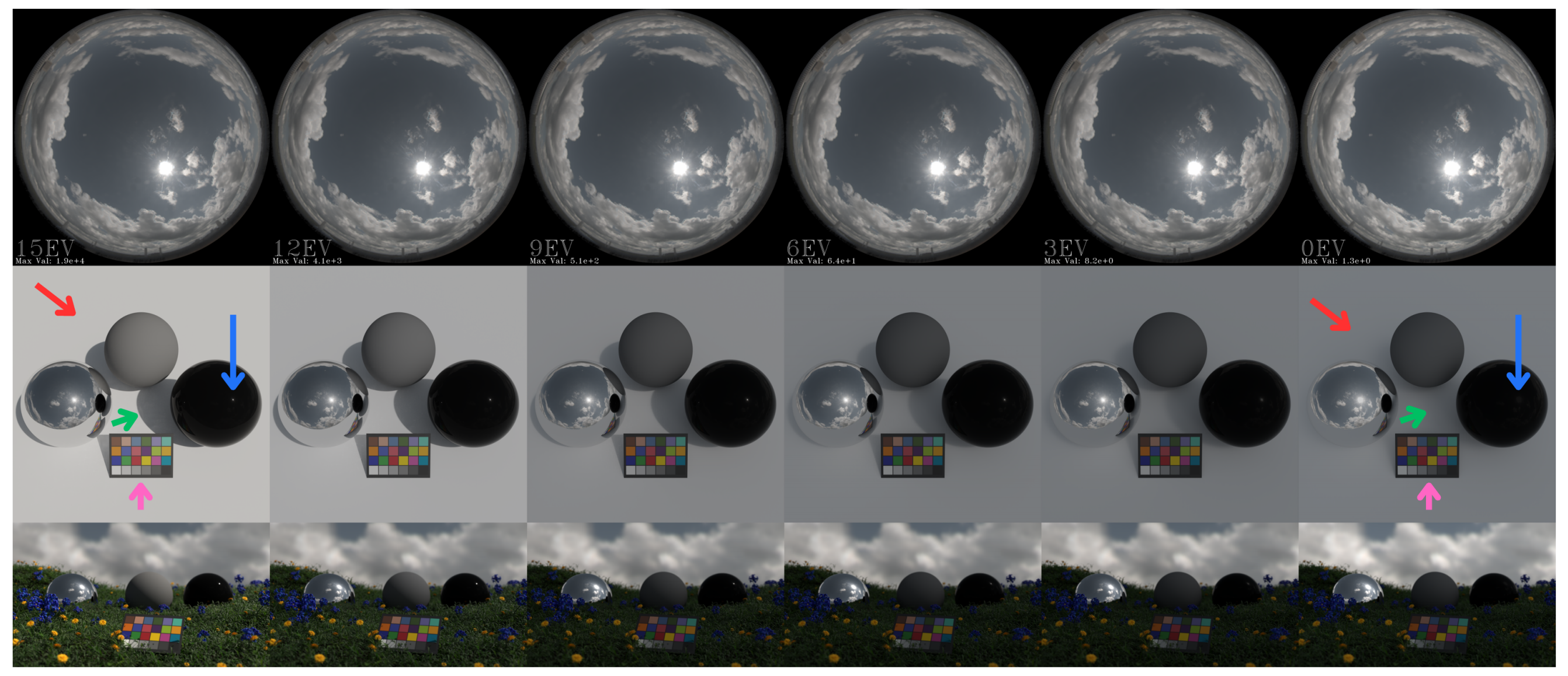


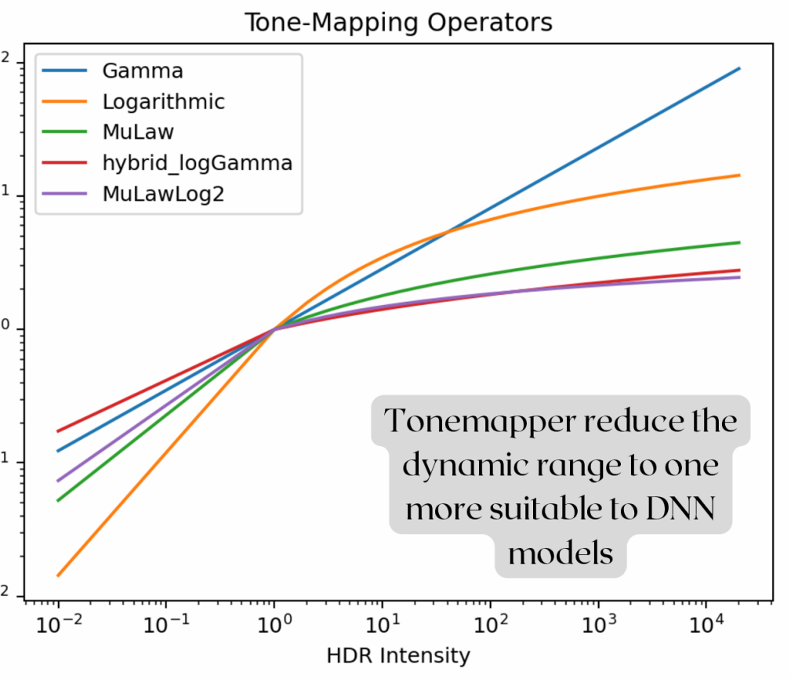


Skydomes which fail to preserve Dynamic Range are **indistinguishable** when viewed as **Low Dynamic Range** (LDR) images and through most common metrics.

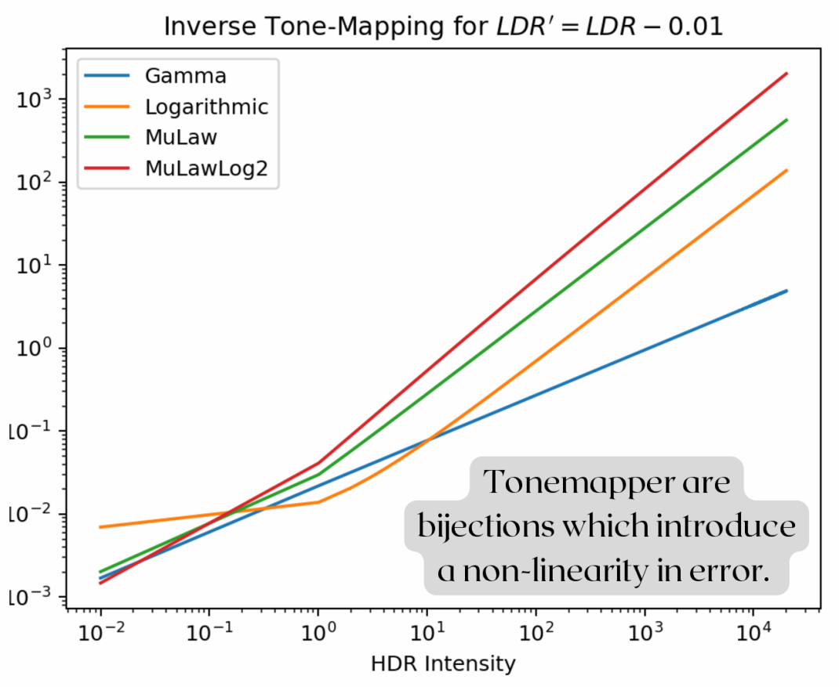
When rendered, flaws the skydome capture- or generation-process emerge, including changing illumination, surface albedos, shadows, and light transmission by materials.

We express Dynamic Range as $EV = log_2(|I|_{max} - |I|_{min})$, where $|I|$ is grayscale intensity. Tone-mapping operators are commonplace to compress Dynamic Range (DR) to a visible (or latent) color-space more favourable for DNN training. We investigate a range of operators (τ), including logarithmic (Log_n) $I' = \log_n(I+1)$, Power-Law (γ) $I' = I^{\frac{1}{\gamma}}$, μ-law $I' = \frac{\log_e(1+\mu I)}{\log_e(1.0+\mu)}$, and combinations thereof as shown





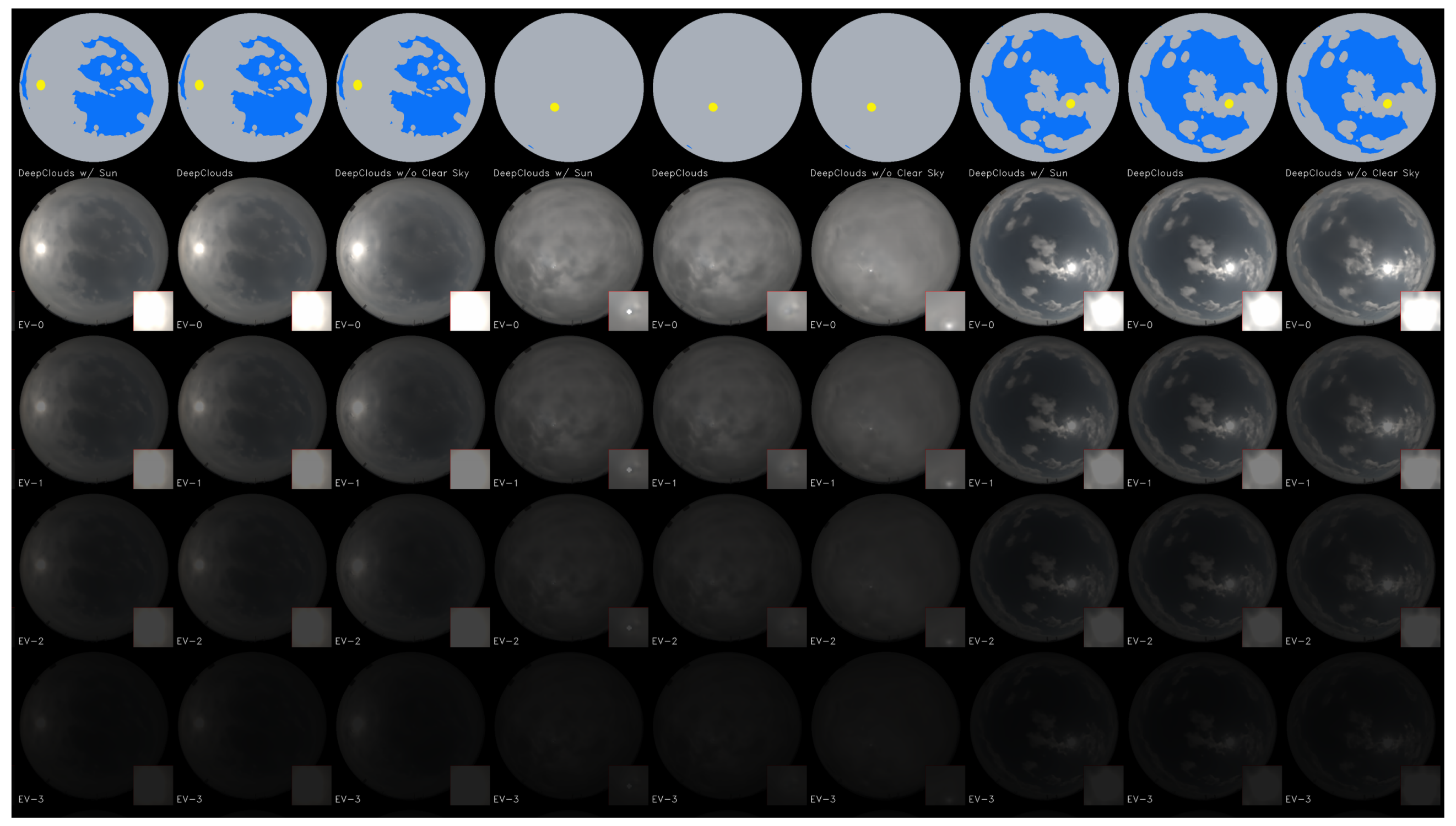
Tonemapper reduce the dynamic range to one more suitable to DNN models



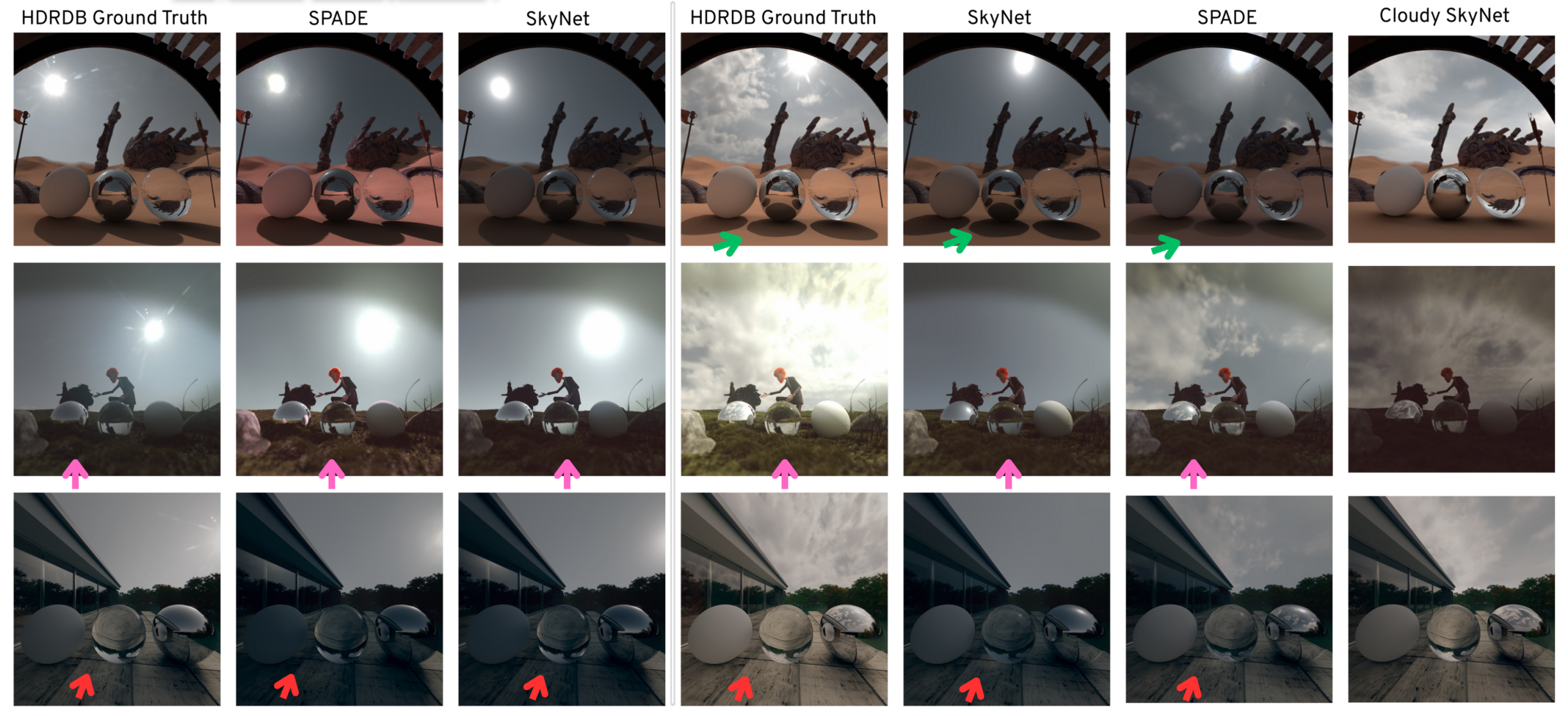
Tonemapper are bijections which introduce a non-linearity in error.

Results

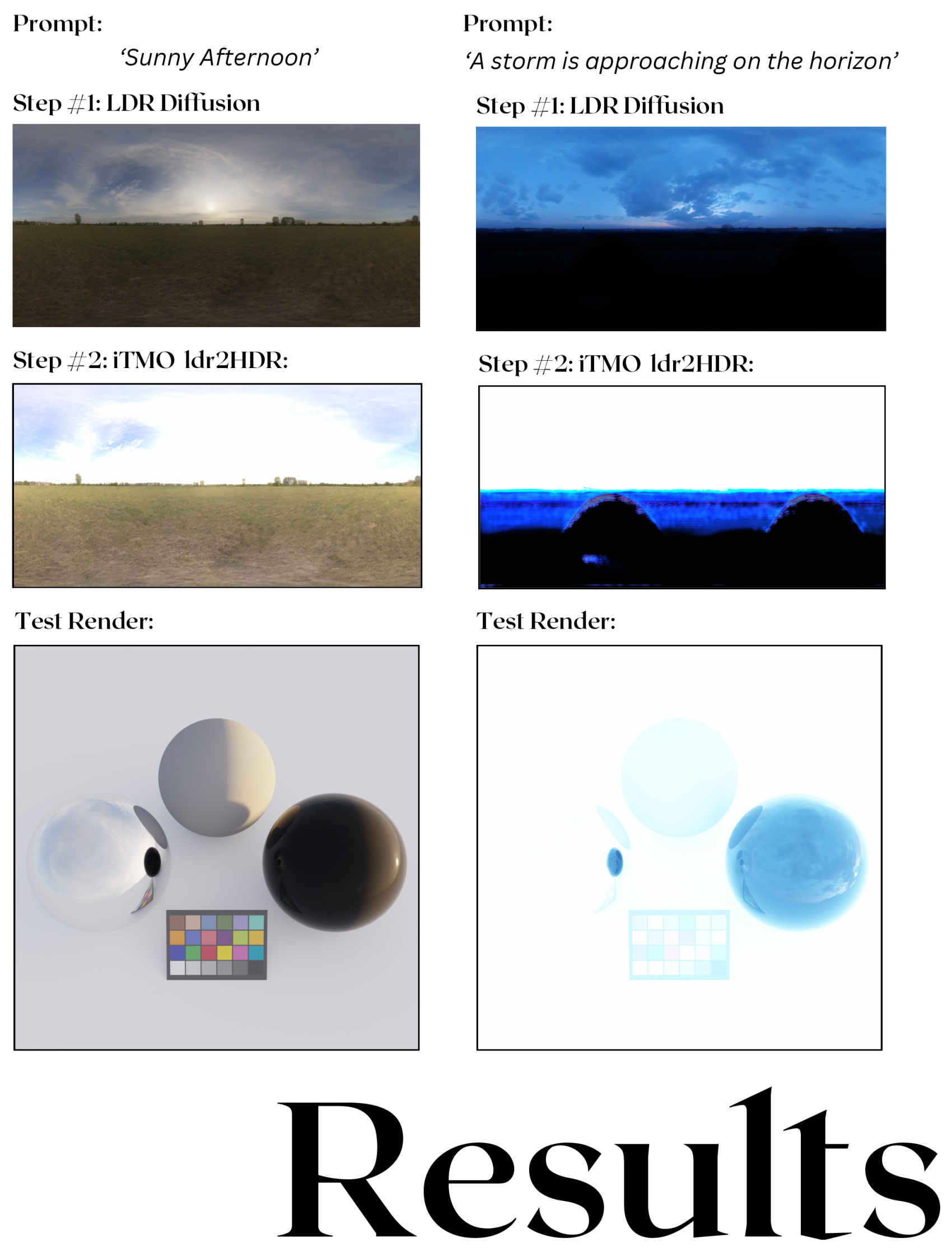
DeepClouds Augments Hošek-Wilkie Clear Skies per "hand-drawn" cloud mask



SPADE
SkyNet (FixUpUnet) Clear and cloudy skies generated from "hand-drawn" cloud masks



Text2Light Skydomes generated via user-provided textual prompts



Prompt:	Prompt:
'Sunny Afternoon'	*'A storm is approaching on the horizon'*
Step #1: LDR Diffusion	Step #1: LDR Diffusion
Step #2: iTMO ldr2HDR:	Step #2: iTMO ldr2HDR:
Test Render:	Test Render:

Quantitative Comparison of Sky Model HDRI. Ground Truth CLIP-IQA is 0.36, and Text2Light Ground Truth CLIP-IQA is 0.38

		$L_1\downarrow$	DR I_f/I_r	$\oiint_I I_f/I_r$	LPIPS$\downarrow$	CLIP-IQA$\uparrow$	UIQI$\uparrow$
Text2Light	LDR	0.16	-	0.71	0.44	0.39	0.026
	HDR	1.39	2.42	3.55	0.42	0.52	0.022

	LDR $L_1\downarrow$	HDR $L_1\downarrow$	LPIPS $\downarrow$	CLIP-IQA $\uparrow$	DR I_f/I_r	$\oiint_I I_f/I_r$
SkyNet	0.05	0.07	0.16	0.50	0.39	0.55
DeepClouds	0.01	0.07	0.21	0.54	0.36	0.59
SkyGAN	NA	NA	NA	0.36	0.96	0.71
Text2Light	0.16	1.39	0.44	0.39	2.42	3.55

	LDR $L_1\downarrow$	LDR $L_2\downarrow$	HDR $L_1\downarrow$	HDR $L_2\downarrow$	LPIPS $\downarrow$	CLIP-IQA $\uparrow$	DR I_f/I_r	$\oiint_I I_f/I_r$
DeepClouds	0.013	0.0008	0.07	412	0.21	0.54	0.36	0.59
DeepClouds w/ Sun	0.019	0.006	inf	inf	0.21	0.53	0.36	$3e^{31}$
DeepClouds w/o Clear Sky	0.016	0.0012	0.07	417	0.21	0.54	0.36	0.62

	Tomemapper	LDR $L_1\downarrow$	LDR $L_2\downarrow$	HDR $L_1\downarrow$	HDR $L_2\downarrow$	LPIPS $\downarrow$	CLIP-IQA $\uparrow$	DR I_f/I_r	$\oiint_I I_f/I_r$
SkyNet (FixUpUnet)	γ	0.057	0.031	-	-	0.25	0.36	0.19	0.48
	Log_2	**0.047**	**0.009**	0.087	428.48	0.28	0.36	**0.27**	**0.58**
	μ-lawLog_2	0.058	0.010	0.078	**422.76**	**0.23**	0.36	**0.27**	**0.58**
	μ-lawLog_2 Global LDR L_1	0.055	**0.009**	**0.077**	423.15	**0.23**	0.36	0.14	0.52
	μ-lawLog_2 Global HDR L_1	0.056	**0.009**	**0.077**	422.96	**0.23**	0.36	0.19	0.54