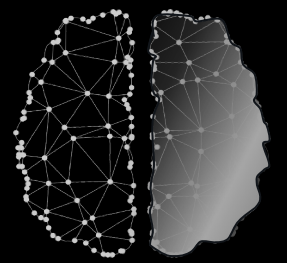
What do Generative Image Models Know?

Anand Bhattad

Research Assistant Professor

Toyota Technological Institute at Chicago



StyLitGAN: Image-based Relighting via Latent Control

CVPR 2024

Anand Bhattad

James Soole

David Forsyth

https://anandbhattad.github.io/stylitgan/

StyLitGAN: Single Image Diverse Relighting

Videos show scene geometry and albedo doesn’t change. Only thing that changes is lighting.

StyLitGAN: Relighting via Latent Control

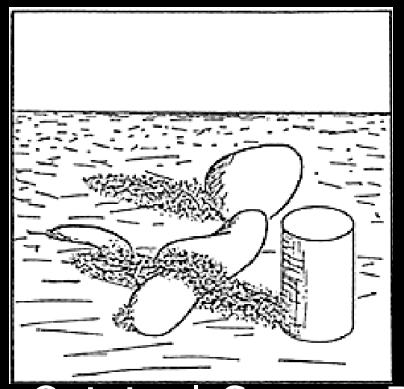
Same direction means similar lighting for all scenes

How Generative Models are doing this?

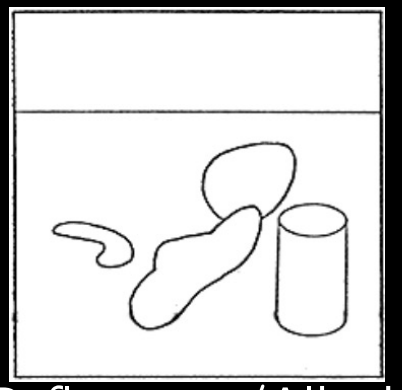
Abstract Representations

“Intrinsic Images”

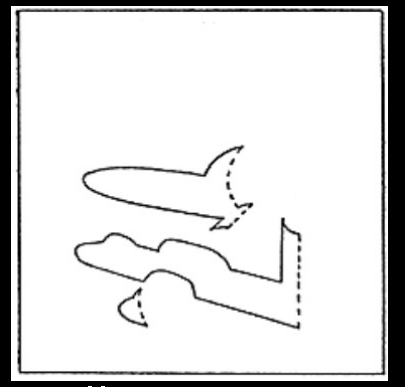
Barrow and Tenenbaum 1978



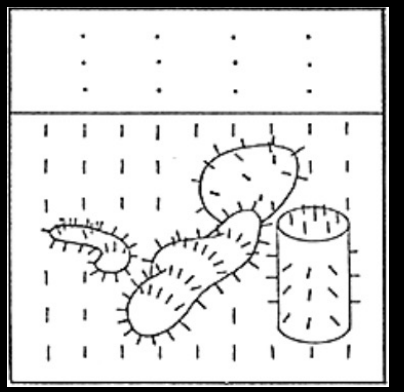
Original Scene



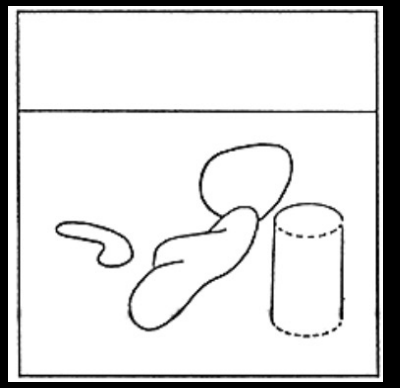
Reflectance/Albedo



Illumination



Orientation



Distance

StyleGAN knows Normal, Depth, Albedo, and More

NeurIPS 2023

Anand Bhattad

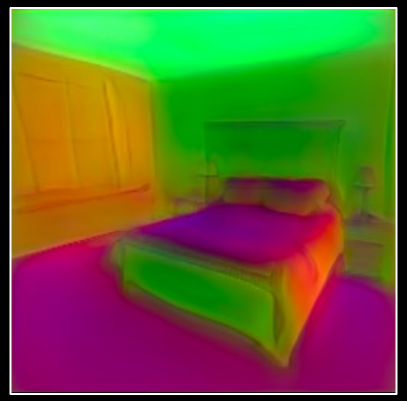
Daniel McKee

Derek Hoiem

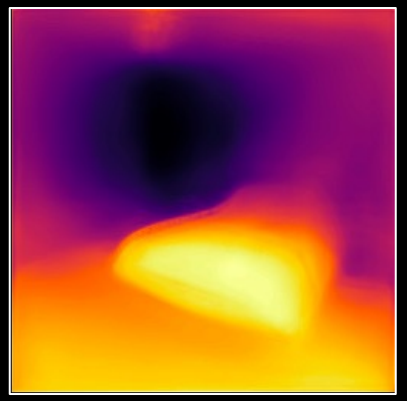
David Forsyth

StyleGAN knows “Intrinsic Images”

StyleGAN Image



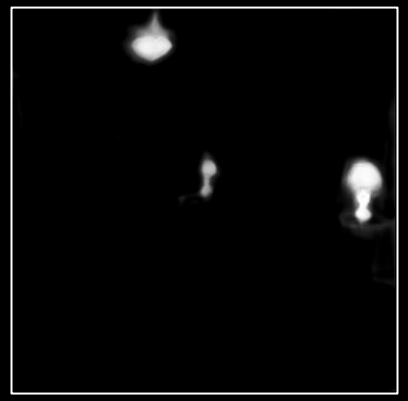
StyleGAN Normals



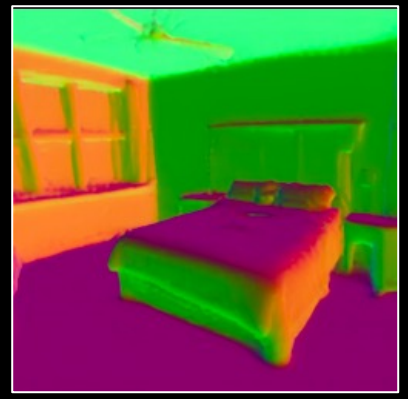
StyleGAN Depth

StyleGAN Albedo

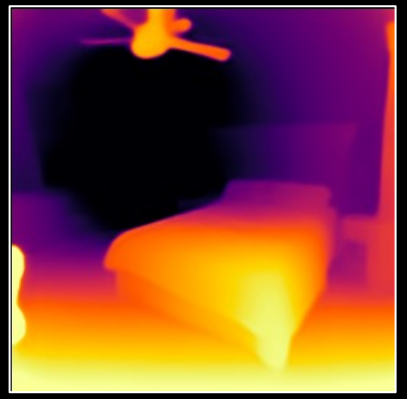
StyleGAN Shading



StyleGAN Segment



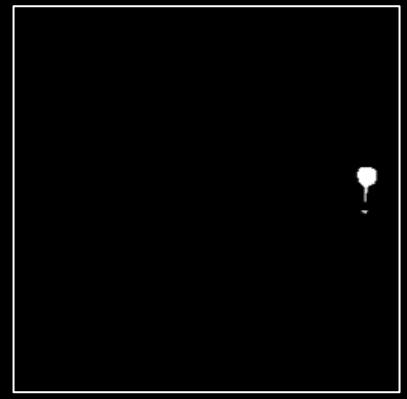
Normals
[Kar et al. ’22]



Depth
[Bhat et al. ’23]

Albedo
[Forsyth and Rock ’22]

Shading
[Forsyth and Rock ’22]



Segment
[Fang et al. ’23]

Is it only StyleGAN that knows Intrinsic Images?

Nope! All Generative Models seem to know about Intrinsic Images

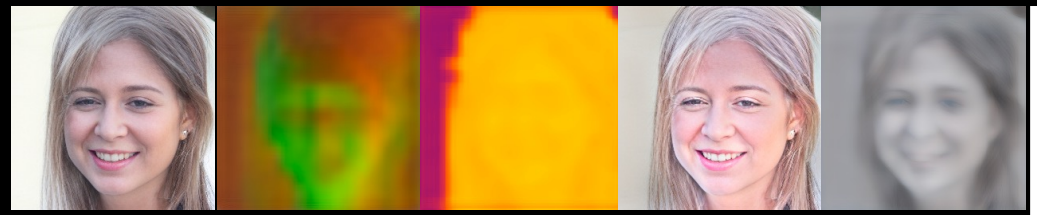


Image Normals Depth Albedo Shading

(a) VQGAN FFHQ (Autoregressive)

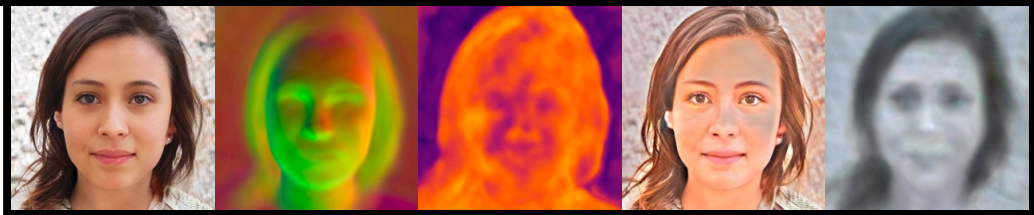


Image Normals Depth Albedo Shading

(b) StyleGAN-XL FFHQ (GAN)

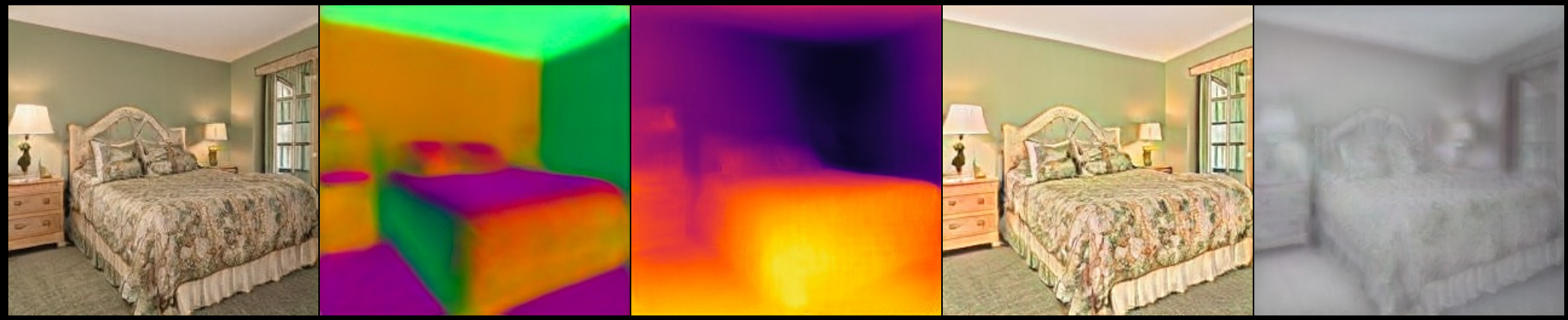


Image Normals Depth Albedo Shading

(c) StyleGAN-v2 LSUN Bedroom (GAN)

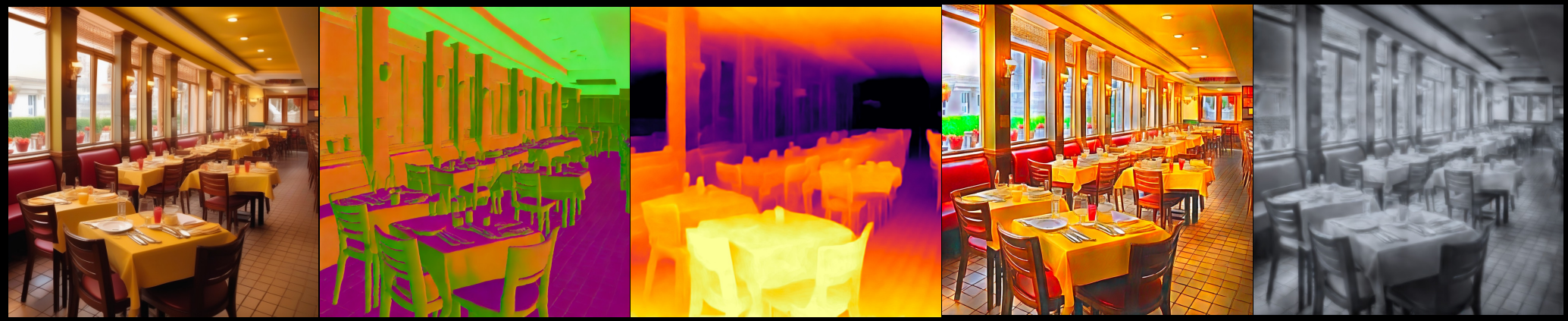


Image Normals Depth Albedo Shading

(d) Stable Diffusion 2.1 (Diffusion)

Generative Models seems to “understand” 3D world

Is there anything that they get it wrong?

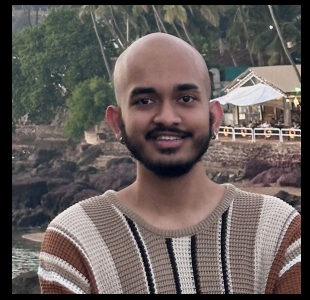
Shadows Don't Lie and Lines Can't Bend!

Generative Models Don't Know Projective Geometry...for now

CVPR 2024

Ayush Sarkar*

Hanlin Mai*


Amitabh Mahapatra*

Lana Lazebnik

David Forsyth

Anand Bhattad

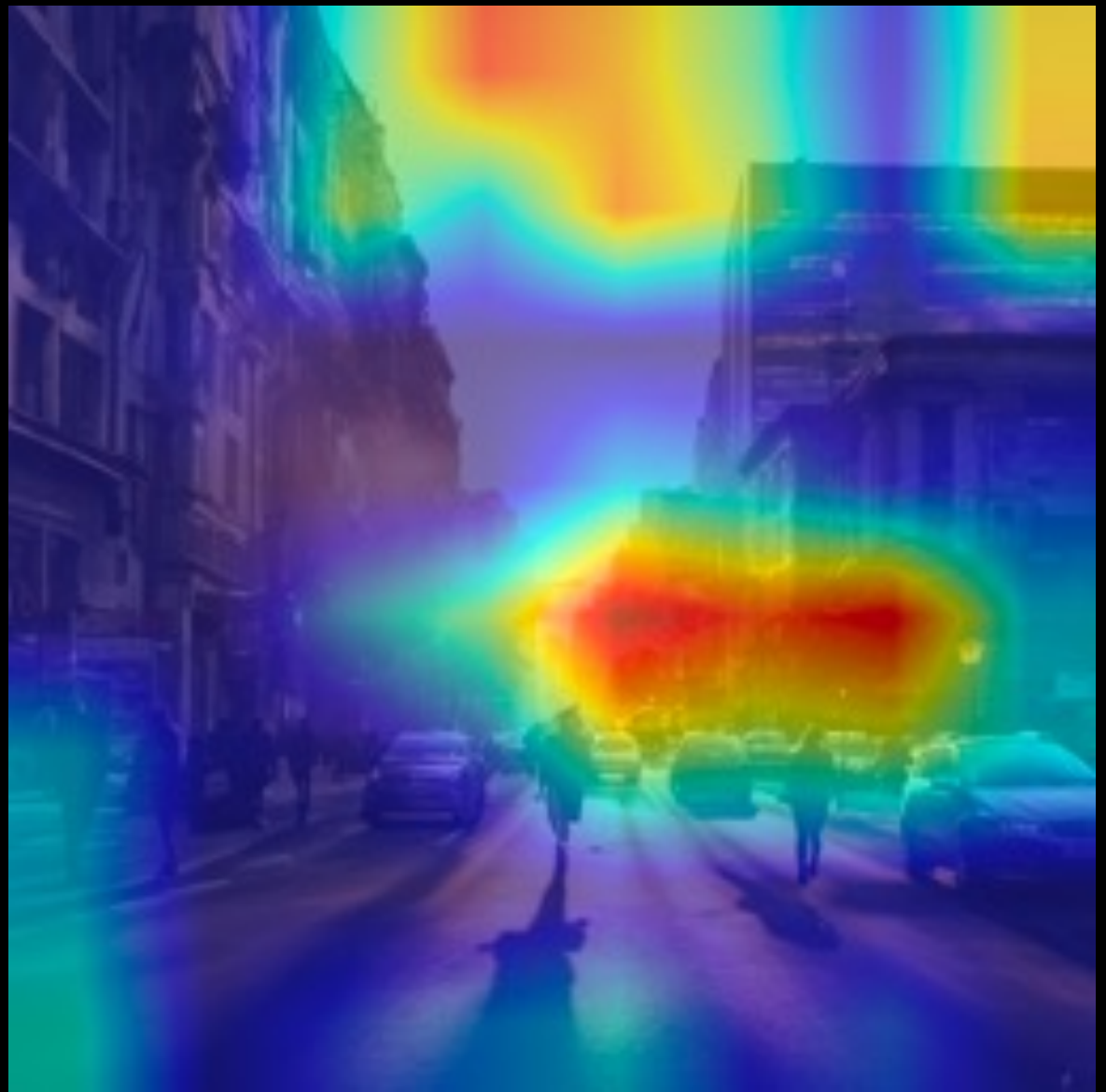
Realistic Generated Images ...

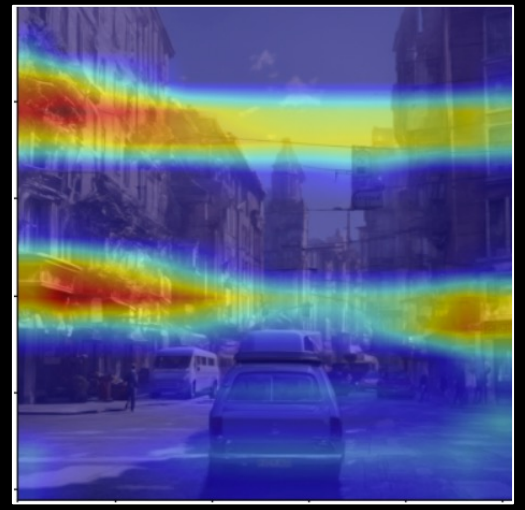
… but there are Shadow Errors

… but there are Perspective Errors

Detected Shadow Errors

Detected Perspective Errors



What do Generative Image Models know?

- a great deal of unexplored knowledge about scenes
- seem to “understand” 3D Scene
- but makes interesting mistakes