Machine Vision Project Proposal

Geometry and Physics-Guided Image Relighting via Intrinsic Decomposition and Latent Control

Satrajit Ghosh
Rutgers University

October 2025

1 Introduction

Realistic image relighting, the process of altering scene illumination from a limited set of images, is a central problem in computer vision with applications in augmented and virtual reality, visual effects, robotics, and modern game engines. Achieving physically convincing relighting requires modeling how light interacts with geometry and material properties. Recent advances such as StyLitGAN [1] demonstrate that generative adversarial networks can synthesize diverse lighting effects through latent-space manipulation without paired data. While visually impressive, these methods lack explicit geometric and photometric reasoning, limiting their physical interpretability and fine-grained control.

This project aims to extend StyLitGAN by incorporating geometric reconstruction and intrinsic image decomposition to achieve physically grounded relighting. The goal is to evaluate whether integrating surface normals, albedo maps, and photometric constraints can enhance realism and controllability.

2 Related Work

This project is closely related to recent developments in image-based relighting and generative modeling. The most relevant work is **StyLitGAN** [1], which performs relighting by manipulating latent directions within a pretrained StyleGAN model. The method estimates intrinsic components such as albedo and shading and then adjusts lighting by traversing specific latent vectors that influence illumination without paired training data. During the process of reviewing this paper for the course paper review, the proposer identified an opportunity to combine the strengths of this latent-space approach with physically grounded methods from classical computer vision. This realization during the review process led directly to the idea for the proposed project, which aims to explore how explicit geometric and photometric reasoning can enhance neural relighting systems.

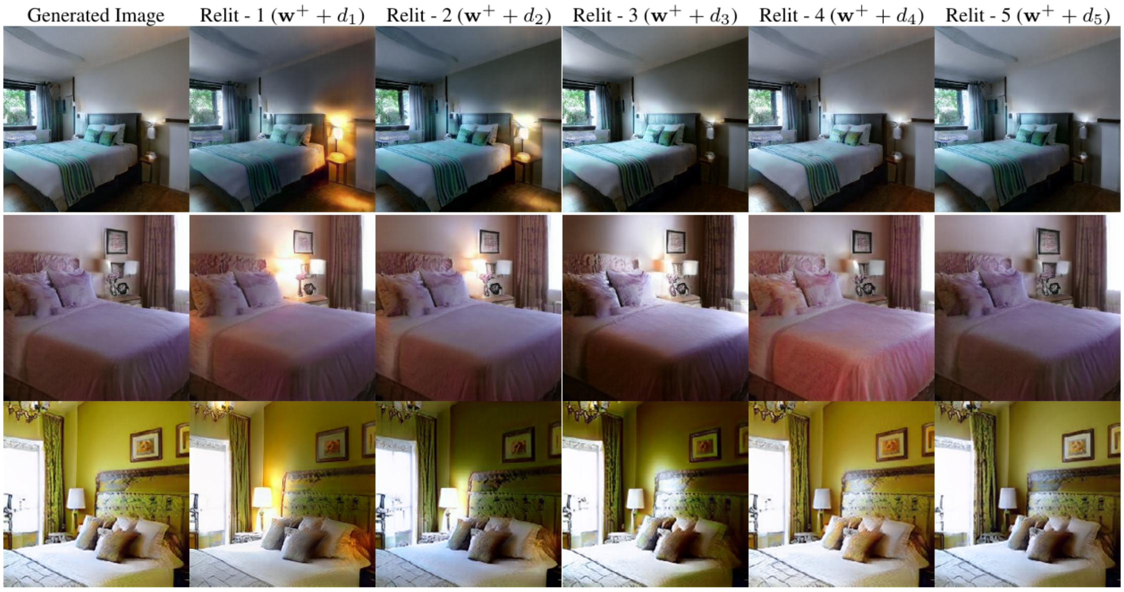

Figure 1: Relighting results generated by **StyLitGAN** [1]. Each column shows the same indoor scene rendered under different latent lighting directions ($w^+ + d_i$). The figure represents one of the best current examples of data-driven relighting, showing how latent-space manipulation can produce diverse illumination effects without explicit geometry or physical modeling. While highly photorealistic, such approaches remain limited in physical interpretability and fine-grained lighting control.

3 Proposed Plan

This project focuses on developing a geometry- and physics-guided relighting framework that extends the ideas introduced in StyLitGAN(Bhattad et al., CVPR 2024). The aim is to integrate geometric reconstruction and intrinsic image decomposition with latent-space relighting to achieve physically consistent and controllable illumination. The approach is structured into four main stages.

Stage 1: Geometry and Normal Recovery

I plan to use COLMAP to perform multi-view stereo reconstruction from a small set of input images, producing a dense point cloud, camera poses, and surface normals. This step provides the necessary 3D structure for the scene, allowing the recovery of per-pixel depth and surface orientation maps. These geometric quantities are crucial because they determine how light interacts with each surface under different illumination directions. By projecting the reconstructed normals back into image space, I can generate accurate normal maps that serve as physically grounded conditioning inputs for the relighting model, bridging the gap between 3D geometry and 2D generative synthesis.

Stage 2: Intrinsic Image Decomposition

I want to separate surface reflectance from illumination by decomposing each image into albedo and shading components through a convex optimization approach. This process should yield physically interpretable maps that describe how color and brightness vary due to material and lighting rather than texture. The resulting albedo and shading fields will complement the recovered geometry, enabling more accurate simulation of how light propagates across surfaces.

Stage 3: Physics-Guided Relighting

Using the depth, normals, and intrinsic components, I aim to simulate new lighting directions to generate physically plausible shading patterns. These maps, along with the albedo estimates, will be used to condition a pretrained StyLitGAN model for controlled relighting. In this configuration, the GAN will no longer rely solely on latent manipulations to infer lighting but will instead be guided by explicit geometric and photometric cues derived from the reconstruction process. I also plan to apply a photometric consistency constraint to ensure that the generated results respect physically meaningful illumination behavior.

Stage 4: Evaluation and Comparison

I plan to compare the results of my system with two baselines: the original StyLitGAN and a purely optimization-based reconstruction. Evaluation will use PSNR, SSIM, and LPIPS for quantitative comparison, along with qualitative assessments of lighting realism and controllability. Testing will begin with simple synthetic objects and progress toward real-world indoor scenes.

Objectives and Expected Outcomes

- To develop an automated geometry- and physics-guided relighting framework that integrates geometric reconstruction, intrinsic image decomposition, and latent-space manipulation to produce physically consistent illumination from standard image inputs.
- To design a geometry-aware conditioning module that uses surface normals and albedo maps to guide a pretrained StyleGAN model, improving lighting control, photometric accuracy, and overall realism.
- To implement a physics-informed loss function based on photometric consistency, ensuring that generated images align with physically valid shading derived from 3D surface geometry.
- To systematically evaluate and compare physics-based, data-driven, and hybrid approaches using quantitative metrics (PSNR, SSIM, LPIPS) and qualitative assessments of lighting realism and interpretability.
- To demonstrate that explicit geometric priors can enhance relighting generalization across different scenes and illumination conditions, providing insights into the role of 3D information in generative relighting.
- To identify strategies suitable for integration into my Master's thesis OpenGL-based game engine, contributing to real-time, physically consistent rendering and future research on intelligent lighting control.

References

[1] A. Bhattad, J. Soole, and D. A. Forsyth, "StyLitGAN: Image-Based Relighting via Latent Control," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2024, pp. 4231–4240.