RGB to Intrinsic Channel Conversion Project

Deepanshu
210311

October 22, 2024

1 Introduction

This project focuses on converting RGB images into intrinsic channels such as albedo, normal, roughness, metallicity, and irradiance. The aim is to enhance image processing applications, including rendering and material recognition.

2 Dataset Description

The dataset consists of pairs of RGB images and their corresponding intrinsic channels. The RGB images are sourced from public datasets and are complemented by synthetically generated intrinsic channels. The dataset is structured as follows:

- **RGB Images:** Stored in the directory `data/rgb_images/`, containing various scenes.
- **Intrinsic Channels:** Stored in `data/intrinsic_channels/`, with each file corresponding to an RGB image.

3 Implementation Details

We implemented two models: 1. RGB to Intrinsic Channels Model (RGBtoXModel): This model takes an RGB image as input and predicts its intrinsic channels. 2. Intrinsic Channels to RGB Model (XtoRGB-Model): This model predicts the RGB image from the intrinsic channels.

Both models are built using PyTorch, utilizing convolutional neural networks (CNNs). The models were trained using Mean Squared Error (MSE) loss and Adam optimizer.

3.1 Model Architecture

- **RGBtoXModel:**
  - Encoder: Consists of a convolutional layer followed by ReLU activation and max pooling.
  - Decoder: A transposed convolutional layer to output five intrinsic channels.
- **XtoRGBModel:**
  - Encoder: Similar to RGBtoXModel but takes intrinsic channels as input.
  - Decoder: Outputs the RGB image from the encoded intrinsic representation.