Generating 3D Shoe models from single/multi view 2D images
COMP 6381 - Project Proposal

Nasir Khalid
ID 40192389

1 Motivation

The inverse rendering problem or 3D shape extraction from 2D images is one with a wide variety of applications in commercial environments. A niche such application is in the world of sneakers and shoe design. Brands are looking for ways to allow users to test shoes in AR/VR environments as well as explore designs in interactive ways that build meaningful experiences.

This is only possible by creating 3D models of the shoes, which proves to be an expensive task because of man hours and software costs involved. The goal of this project is therefore to simplify and speed up the creation of 3D shoe models

2 Project Description

The idea is to train and end-to-end model which takes single (or multiple) shoe images and it can output a 3D model of the shoe

2.1 Data Requirements

Training for this model requires images of shoes from multiple angles. Links to images of sneakers from 35 viewpoints can be pulled from "The Sneaker Database" through a python script. From the image links and another script all the images can be downloaded.

2.2 Training

The training can be broken down in to the steps below

1. Create an encoder model that takes the shoe image from one/mutliple views and outputs a latent embedding
2. Create a decoder model that takes this latent embedding and outputs a list of vertex deformations
3. Apply the vertex deformations from 2. to an initialized shoe model
4. Use a differetiable renderer to get multi view images that can be used as a loss for target views
5. Backpropogate image losses and mesh regularizer losses to train model

The figure below shows this visually

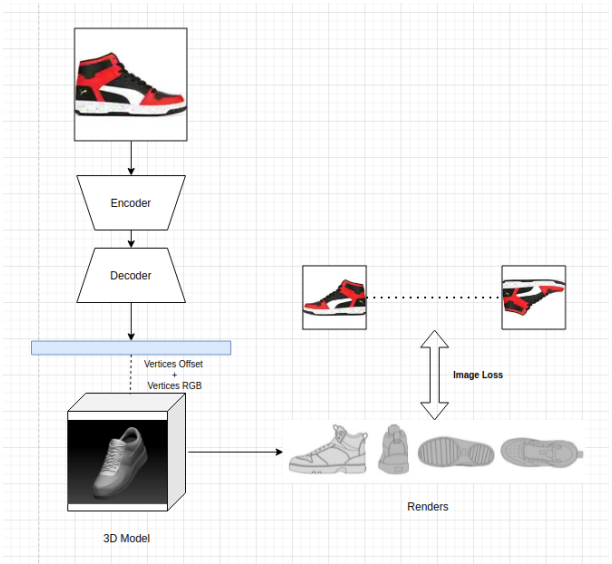

Figure 1: Model Diagram

3 Objectives

Below is a list of fixed objectives that the project is broken down in to:

1. Develop script to get multiview shoe images from API
2. Set up model and training loop
3. Adjust weights and test what provides best results
4. Extract models and texture from renderer
5. Test with hand drawn designs and images outside of training distribution