Abdullah Saafan
Supervision by Prof. Mohammed Salem

On Image to 3D Model Construction for E-Commerce Applications

German University in Cairo

Motivation

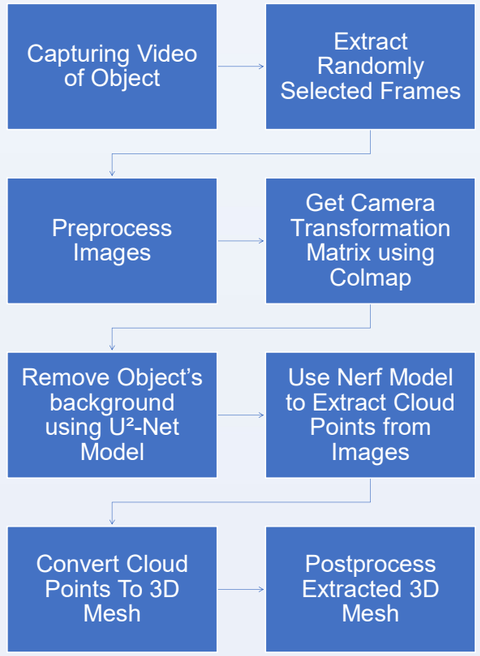
Many business owners now seek an online shopping experience for their customers where they can have the same engagement with their products online as on-site in physical stores. Also, many customers don't prefer online shopping as they cannot imagine exactly how the product will look like on-hand from only seeing images. This gives online shopping a huge disadvantage and discourage business owners from selling online as their products might be misperceived by the targeted customers due to the lack of proper visualization, in addition to the large number of returned purchased items from customers in case they changed their minds when they saw the product in person however clear and descriptive the product images were. Showing a 3D Models for each product on the Seller's website will drastically impact the accuracy of how the user view the product as well as it can provide an interactive engagement with the product or customizability that images doesn't provide (changing colors or sizes, adding extra features or accessories) in an instant and allow customer to create limitless versions of the same product. However, hiring 3D Artists for creating an accurate good-looking 3D Models for each product is an extremely expensive, time consuming and cumbersome way to create these 3D Models.

Problem Statement

An alternative efficient solution would be to use deep learning to extract these 3D Models from 2D images of the product. The Goal of this Thesis will be to create new techniques using deep learning, computer vision and computer graphics to obtain the 3D Model's parameters and properties from one or more image and then deform a 3D base mesh from an existing Assets library of the same image object's category to match the object in the image structurally.

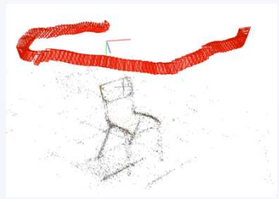
Methodology

1. Capture 360 Degrees Video Footage of the to-be-reconstructed Object
2. Extract Randomly Selected Frames from the Captured Video
3. Preprocess Image to enhance Transformation Matrix Extraction and background Removal
4. Use Colmap to Extract Camera Transformation Matrix from all frames
5. Remove Object's Background so that It doesn't get reconstructed together with the Object
6. Use Nerf Model to 3D-Reconstruct the Object by extracting Cloud Points from Images
7. Use MeshLab to Convert Cloud Points to 3D Mesh and postprocess the extracted Mesh



Example Pipeline

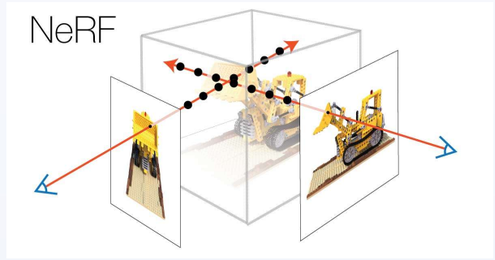
Input Multiple Images For Camera Pose Estimation



Background Removal

Use NeRF to Extract 3D Volumetric Grid

Overview of Neural Radiance Fields



A neural radiance field (NeRF) is a fully connected neural network capable of generating novel views of complex 3D scenes from a limited set of 2D images. It has been trained to reproduce input views of a scene using a rendering loss. It works by interpolating between input images representing a scene to render one complete scene. NeRF is a powerful tool for creating images from synthetic data. Using volume rendering, a NeRF network is trained to map directly from viewing direction and spatial location (5D input) to opacity and colour (4D output). NeRF is a computationally intensive algorithm that can take hours or days to process complex scenes. New algorithms, on the other hand, are now available that significantly improve performance.

RESEARCH POSTER PRESENTATION TEMPLATE © 2019
www.PosterPresentations.com