Implicit Models for Scene Representations

Yiming Huang, Yiqiao Lin, Zichen Tang

Advised by Prof.Qifeng Chen

THE DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
計算機科學及工程學系

Overview

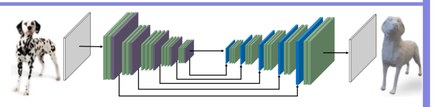
Background

There is a growing demand for complex and realistic 3D models in industries such as movie, gaming, virtual reality and e-commerce. However, creating 3D models from scratch is difficult and time-consuming, and therefore expensive.

To solve this problem, implicit modeling has been used to generate 3D models directly from 2D images using neural networks, which is also known as 3D reconstruction.

Figure 1. *Left*: DC League of Super-Pets is a latest animated film, whose story basically revolves around two dogs. *Right*: Metal dogs is a video game with a fighting dog as the main character.

Objectives



Since most of the work on 3D reconstruction has been about static objects and human faces or bodies, we decided to explore 3D reconstruction of dogs. Our project aims to:

- Build a geometry 3D reconstruction module.
- Build an appearance 3D reconstruction module to enrich the model details.
- Provide a user-friendly interface for demonstration and modification. (e.g. Changing poses)

Methodology

Datasets

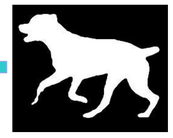
- StanfordExtra: 12k labelled images of dogs
- Metadog: 9x100f dog videos with labels.

2D Image/Video Input

Input

Geometry Reconstruction

Joint and pose



Contour

Features extracted by a pretrained stack hourglass network

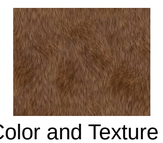
Compute loss and update with gradients

Rendering

3D Geometry

GT keypoints

Appearance Reconstruction



Color and Texture

3D Output model

Figure 2: Procedure of the whole 3D reconstruction process

Progress

Work completed

- Trained pose-branch reconstruction module
- Trained shape-branch reconstruction module
- Improved results with a video input refinement module

Figure 3: Current result using image input.

Work in progress

- Designing a pose modification module
- Optimizing geometry network with new loss functions
- Constructing the appearance reconstruction module

Conclusion

Our current system can predict accurate estimates of 3D dogs shape and pose from images while also producing consistent 3D geometry reconstruction. Our generated dog models assist the development on softwares such as Unity and Unreal.