HAAG Weekly Report (Simplified) – Nikita Angarski – 3D Modeling

Week 1-2

Time-Log

- What did you do this week?
    - Started work on exploring Dr. Hirose's approach, forking the repo and going through the code, have not made any edits/commits yet, so far just setting up a coding environment and getting it working.
    - Decided on a dataset to use, CAESAR works well, because it is the same subject as Hirose's (faces, human bodies)
    - Decided on what the experiment is going to attempt: testing Hirose's code for robustness and scalability with large point representation data sets.
    - Edited the submission doc to reflect these steps.
- What are you going to do next week
    - Get Hirose's C code working on current setup
    - Transform CAESAR dataset into point set/cloud representation

-

- Blockers, things you want to flag, problems, etc.
    - One rather large hurdle I currently foresee for myself is gaining an understanding of how point by point representations are transformed into matrices as in the paper, I anticipate I will need some help on this.

Abstracts:

3DCoMPaT200: Language Grounded Large-Scale 3D Vision Dataset for Compositional Recognition

https://neurips.cc/virtual/2024/poster/97745

Abstract:

Understanding objects in 3D at the part level is essential for humans and robots to navigate and interact with the environment. Current datasets for part-level 3D object understanding encompass a limited range of categories. For instance, the ShapeNetPart and PartNet datasets only include 16, and 24 object categories respectively. The 3DCoMPaT dataset, specifically designed for compositional understanding of parts and materials, contains only 42 object categories. To foster richer and fine-grained part-level 3D understanding, we introduce 3DCoMPaT200, a large-scale dataset tailored for compositional understanding of object parts and materials, with 200 object categories with 5 times larger object vocabulary compared to 3DCoMPaT and 4 times larger part categories. Concretely, 3DCoMPaT200 significantly expands upon 3DCoMPaT, featuring 1,031 fine-grained part categories and 293 distinct material classes for compositional application to 3D object parts. Additionally, to address the complexities of compositional 3D modeling, we propose a novel task of Compositional Part Shape Retrieval using ULIP to provide a strong 3D foundational model for 3D Compositional Understanding. This method evaluates the model shape retrieval performance given one, three, or six parts described in text format. These results show that the model's performance improves with an increasing number of style compositions, highlighting the critical role of the compositional dataset. Such results underscore the dataset's effectiveness in enhancing models' capability to understand complex 3D shapes from a compositional perspective. Code and Data can be found here: https://github.com/3DCoMPaT200/3DCoMPaT200/

Summary:

The paper introduces a new dataset called 3DCoMPaT200, which significantly expands the range of 3D object categories and part classifications for better understanding of objects at a detailed level. It also presents a new task to improve how models retrieve and understand 3D shapes based on descriptions of their parts, showing that more detailed data leads to better performance.

What did you do and prove it

This week was mainly about 1) making a testable hypothesis with the information we have, and editing the submission doc to reflect more detailed write up of what we aim to do, and making sure it fits scope. 2) beginning to extrapolate what testable data is to look like with respect to existing code, and 3) exploring methods to transform existing data into a testable format (point representation of 3D images). As I am new to computer imaging, this and last week were both very focused on learning the ins and outs of data representation in the formats we want to explore. I have been able to find both a dataset to get started on testing, and a Python library that will help transform the data into point cloud representations:

https://humanshape.org/CAESAR/

https://pypi.org/project/open3d/

Depending on the scope of the project that we decide on, the Python library does a good job of allowing users to select the specificity of the models and add points to test for scalability, add spurious points to test for robustness, and test for accuracy. The CAESAR dataset's general breadth will allow us to test in-vs-out of sample, which is a nice extension of Dr. Hirose's work.

Unfortunately nothing here yet, but a link to a fork of Dr. Hirose's work is included.

https://github.com/Nikitos1865/dld