Project Proposal

SWE3002_41

2024 .10. 24.

1. Team name and team members

Team Name : 비전이동 (Team5)

Team members :

박진호(2018312612), 김주현(2019314455) , 최준영(2019314833), 박재영(2019310338), 김동현(2019312104), 이호준(2020312401)

2. Project background including related work

Crafting 3D assets from single-view images can facilitate a broad range of applications, eg, virtual reality, industrial design, gaming and animation.

Related works :

zero123++

LRM (Large Reconstruction Model)

3. Goal

Towards High-Quality Image-to-3D Generation

The project aims to achieve **High-Quality Image-to-3D Generation** with a specific focus on improving the visual realism, and efficiency of generating 3D models from single-view image.

4. Plan

These are our tentative plan

1. Investigating current SOTA models of 3D generation / 3D reconstruction (~11/01)

-> Since we have time constraints and limited GPU resources, we will focus on improving existing models rather than proposing entirely new methodologies or training models from scratch.

-> we are planning it by adapting super-resolution, or just fine-tuning. we will discuss about more.

2. Datasets Preparation (~11/01)

3. Model Inference and Improvement (~12/06)

-> Perform inference with the investigated models, identifying areas for improvement based on their performance and results. In addition, we will evaluate our improvements via baseline models.

Also, we are planning sketch to 3D.

If time allows, we will extend the project to incorporate sketch-to-3D generation with controlnet.

5. Datasets

We change our datasets with these :

CO3D datasets

https://ai.meta.com/datasets/co3d-dataset/

Modelnet40 datasets

https://www.kaggle.com/datasets/balraj98/modelnet40-princeton-3d-object-dataset

6. Evaluation methods and baselines for comparison.

Metrics :

User-based visual assessment, IOU score, Chamfer Distance (TBD), LPIPS (TBD)

Since this project is about vision tasks, we think that user-based visual assessment is more important than metric-based evaluation. The quality of 3D models will be ultimately determined by the visual experience of the users.

Baselines :

CRM

InstantMesh

... (TBD)

7. Github Link

https://github.com/2024-2-AIP-Team5/Image_to_3D