Artificial Intelligence for 3D Asset Model Reconstruction from 2D Images

Authors : Jeeraboon Sukparungsee, Pataranan Sethpakdee
Advisor : Atthapon Lorphan, Art-ong Mukngoen

Thailand - Japan Student ICT Fair 2024 · SIF 2024

CHITRALADA SCHOOL

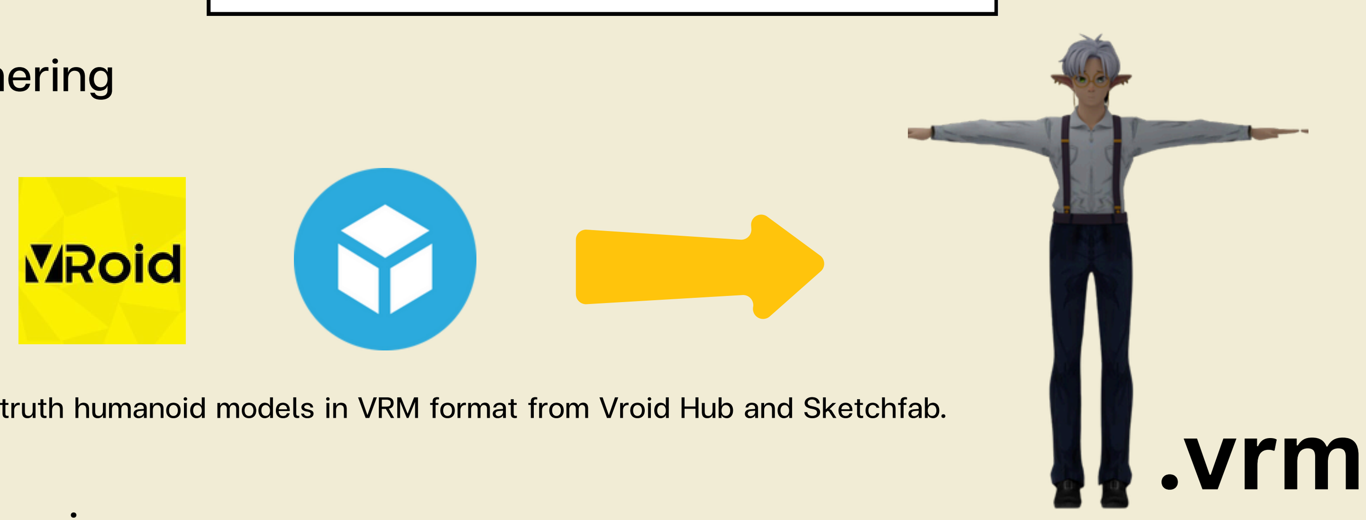
Problem

3D asset models are becoming increasingly popular in various fields such as the gaming and animation industries. However, creating a 3D model from scratch is a complex process that requires significant time and patience, leading to a growing demand for skilled 3D artists.

Currently, Artificial Intelligence (AI) is widely used to assist in creating 2D images. However, the process of generating 3D asset models remains more complex and labor-intensive. This project aims to explore how AI can be leveraged to generate 3D asset models from 2D images by utilizing 3D Point Cloud Reconstruction. This technique involves the use of a 3D Point Cloud, which is a set of points represented by their coordinates (x, y, z) in 3D space [1],[2].

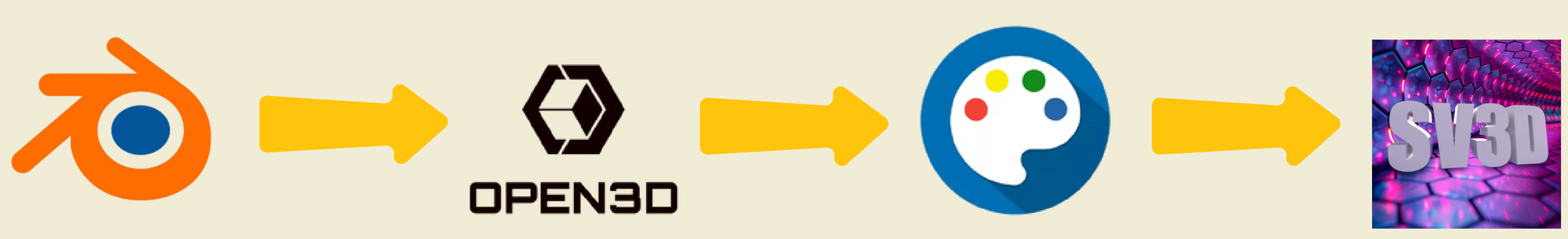
Framework

1. Data Gathering



Gather 3D ground-truth humanoid models in VRM format from Vroid Hub and Sketchfab.

2. Data Processing



Use Blender's Python API to capture 2D images of the model and convert the model's format from VRM to PLY.

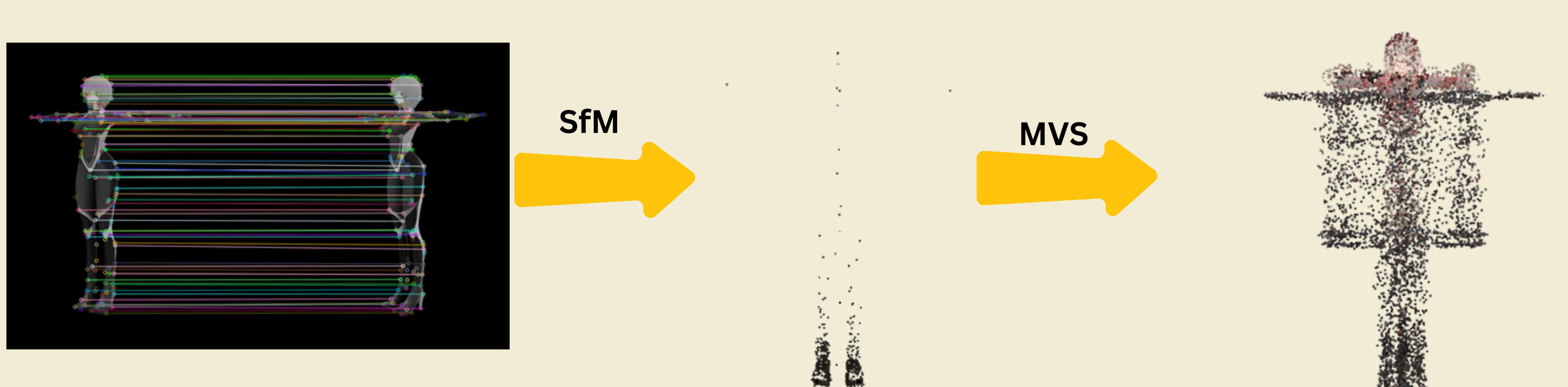
Use Open3D to convert the mesh into a 3D Point Cloud.

Use ControlNet for pose estimation.

Use Stable Video 3D to generate additional views of the image.

3. Feature Extraction & Matching

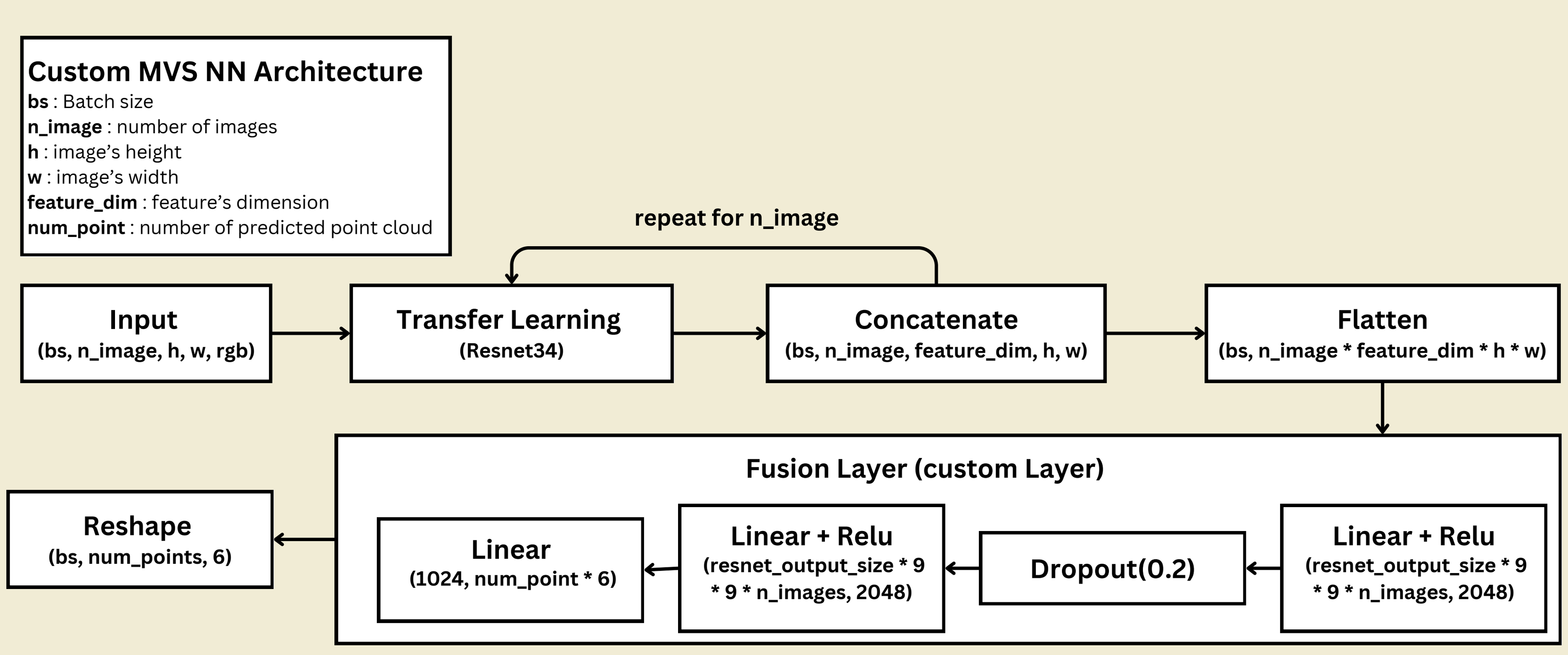
4.Structure from Motion (SfM) and Multi-View Stereo (MVS)



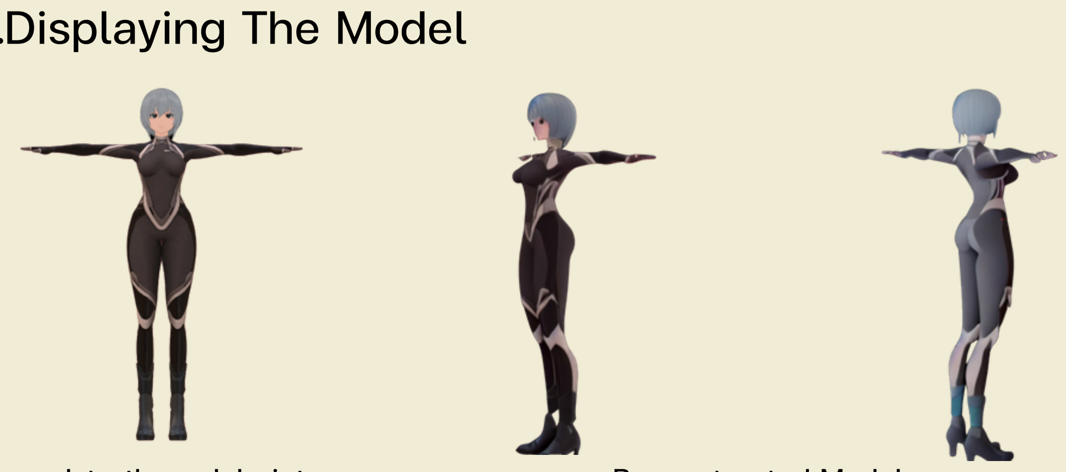
21 Matches of Feature In different orientation

Pose + Sparse 3D Pointcloud

Dense 3D Pointcloud



5.Displaying The Model

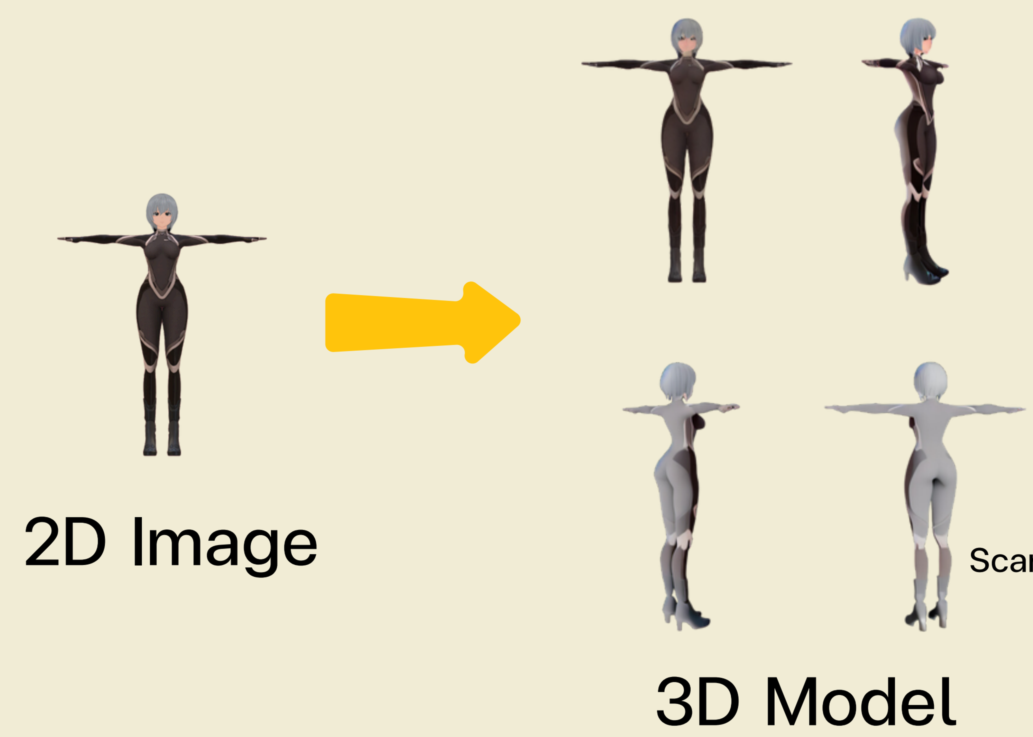


Ground-truth model picture

Reconstructed Model

Use reprojection loss analysis to compare the similarity between the 2D image and the 3D model from the same angle.

Results



Scan to download 3D model

The trained model can successfully reconstruct 3D model from 2D image.

Reprojection Loss Analysis

Image	Reprojection Loss
1	2.55
2	4.34
3	3.65
4	4.32
5	3.15
Average	**3.60**

The above table indicates that the results are based on a reprojection loss analysis, comparing the similarity between the 2D image and the 3D model from the same angle. The analysis shows that the average error rate is 3.6.

Conclusion

This project successfully demonstrates a method for generating 3D asset models from 2D images by utilizing techniques such as 3D Point Cloud Reconstruction, Multi-View Stereo, and Iterative Closest Point alignment. By combining tools like Blender's Python API, Open3D, Deep Point, and Deep Glue, we were able to reconstruct 3D models from 2D images. The reprojection error analysis, which compared the similarity between the 2D images and the reconstructed 3D models, revealed an average error rate of 3.6. This result highlights the potential of AI-driven approaches to simplify and accelerate 3D model creation, offering significant benefits to industries such as gaming, animation, and media. The proposed framework not only reduces the complexity of manual 3D modeling but also opens the door to further exploration of AI in automating 3D asset generation.

Acknowledgements

We would like to express our gratitude to our advisors, Mr. Atthaphon Lorphan and Mr. Art-ong Mukngoen for their invaluable advice. Our gratitude also goes to Assistant Professor Yuvadee Nakapadungrat, Ph.D. and Associate Professor Arunee Intasorn for their advice on improving this report and experiment. This project would not have been possible without their substantial help.

References

[1]Choe, J., Joung, B., Rameau, F., Park, J., & Kweon, I. S. (2022). Deep point cloud reconstruction. arXiv.
[2] Kingkarn, C. (n.d.). 2D to 3D construction - Paper: Dynamic Graph CNN for learning on point clouds.