

Improving Rendering and Texture Reconstruction in 3D Face Models

Valentina Tang

University of California, Santa Cruz

Computer Science and Engineering

Jun 11, 2024

Abstract

This project aims to enhance the quality of rendering and address the issue of incomplete texture data in 3D face models. Utilizing advanced techniques and tools, including Blender, DeepBump, Principal Component Analysis (PCA), and Masked Autoencoders (MAE), the project seeks to improve visual fidelity and efficiently reconstruct missing texture data. The findings indicate that while PCA is effective, MAE offers superior performance, particularly when dealing with significant data loss.

1 Introduction

The goal of this project is to improve rendering quality and address incomplete texture data in 3D face models. High-quality rendering and complete textures are essential for applications in gaming, animation, virtual reality, and various design fields. Incomplete texture maps, often resulting from having only a single image from a specific angle, lead to poor rendering quality. Therefore, this project focuses on developing methods to enhance visual fidelity and reconstruct missing texture data efficiently.

2 Related Work

Rendering and texture reconstruction have been extensively studied in computer graphics and computer vision. Traditional methods for improving rendering include using various texture maps and lighting simulations. Recent advances in machine learning have introduced techniques such as PCA and autoencoders for texture reconstruction, showing promising results in handling missing data and improving texture quality.

3 Methodology

3.1 Rendering Enhancement

Blender and DeepBump were utilized to improve rendering quality and efficiency. The key components involved in this process included roughness maps, subsurface scattering, normal maps, lighting conditions, and an automation script. Roughness maps define the roughness of a surface, enhancing the material's appearance. Subsurface scattering simulates the way light penetrates and diffuses through translucent materials such as skin, adding realism. Normal maps simulate surface detail without adding extra geometry, enhancing the visual complexity of the model. High Dynamic Range Imaging (HDRI) was used to create realistic lighting and reflections, simulating real-world conditions. An automation script was developed to streamline the rendering process, ensuring consistency and efficiency.



Figure 1: example normal map generated from Deep Bump



Figure 2: example roughness map generated from Deep Bump



Figure 3: example subsurface map generated from Deep Bump

3.2 Texture Reconstruction

However, in many cases, only a single image of the 3D face model is available, leading to missing texture data for other sides. This creates a need for effective texture reconstruction methods. Principal Component Analysis (PCA) was one such method used. PCA transforms original data into orthogonal components, capturing the most variance. The basic PCA approach was modified to exclude masked points (representing missing data) during transformation, focusing on available data for accurate reconstruction. Instead of the original matrix in trained PCA, we deleted corresponding columns and rows. In this way the

masked pixels are not taken into account for loss. The PCA training pipeline involved using the FFHQ-UV dataset with a training set of 1000 samples and 1000 components.

Masked Autoencoders (MAE) were also employed. MAE is an asymmetric neural network architecture designed for self-supervised learning. The encoder processes visible data, while the decoder reconstructs masked (missing) data. MAE was finetuned on the ImageNet-1K dataset. The training pipeline for MAE included data augmentation techniques such as random resize crop, random augment, label smoothing, mixup and cutmix, and drop path. Training involved 2000 samples from the FFHQ-UV dataset using the AdamW optimizer with a cosine decaying learning rate. The masked ratio was set to 75%, and training was conducted over 50 epochs with the mean squared error (MSE) of reconstructed patches as the loss function.

4 Results

The use of Blender and DeepBump significantly improved rendering quality. Roughness maps, subsurface scattering, normal maps, and realistic lighting conditions contributed to more visually appealing and realistic 3D face models. The automation script ensured consistent and efficient rendering.

The comparison of PCA and MAE methods revealed that PCA effectively handles texture reconstruction by leveraging principal components, but its performance declines with increased missing data. MAE consistently outperforms PCA, especially with significant data loss. The quality of reconstruction with MAE is highly dependent on the amount of available unmasked data. Visual comparisons were made using random masking at various masking ratios (25%, 50%, 75%, 90%) as well as center masking, outer masking, grid masking, and left half masking scenarios. For more results please refer to appendix. As we can see, PCA’s reconstructed images has some color difference. This is due to the deleted columns and rows in the weight matrix. While the MAE focus more on the reconstruction of the masked patches, which they only compute loss on patches that needs reconstruction. Therefore the reconstructed images has weird looking cubes for parts that don’t need reconstruction.



Figure 4: 0.25 mask ratio image reconstructed by PCA



Figure 5: 0.5 mask ratio image reconstructed by PCA



Figure 6: 0.75 mask ratio image reconstructed by PCA

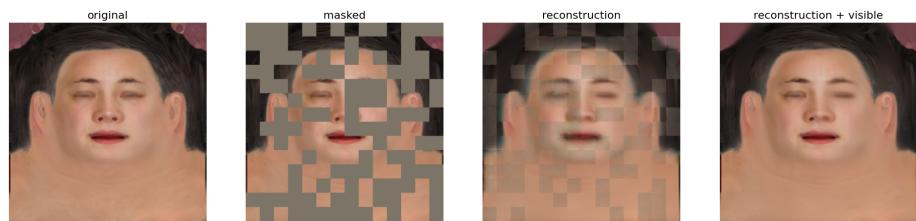


Figure 7: 0.5 mask ratio random masked image reconstructed by MAE

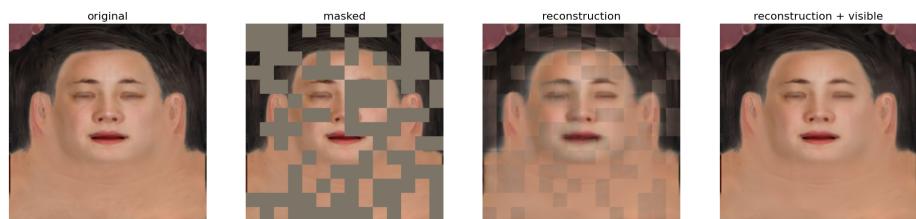


Figure 8: 0.5 mask ratio image reconstructed by pre-trained MAE

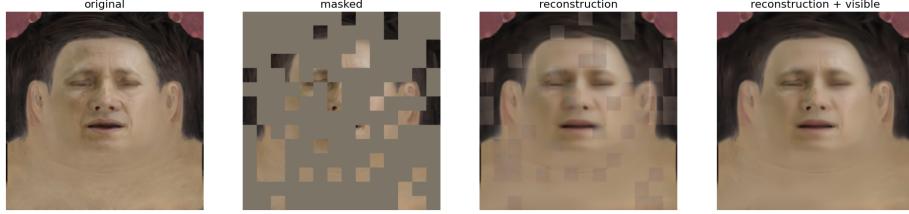


Figure 9: 0.75 mask ratio image reconstructed by pre-trained MAE



Figure 10: 0.9 mask ratio image reconstructed by pre-trained MAE

5 Discussion

The project demonstrates that while PCA is effective for texture reconstruction, its performance is limited when a large portion of data is missing. MAE, on the other hand, provides superior performance in such scenarios, making it a more robust choice for handling incomplete texture data. The effectiveness of MAE is closely tied to the amount of information available in the unmasked data, highlighting the importance of data quality and availability.

6 Conclusion

Both PCA and MAE can be used for texture reconstruction. PCA effectively handles texture reconstruction but struggles with large amounts of missing data. MAE consistently outperforms PCA, particularly with significant data loss, and is more robust in handling large gaps in texture data. The effectiveness of MAE is closely tied to the amount of information available in the unmasked data.

The weakness of this method is that both PCA and MAE generate low resolution images (224x224), however what we want is textures that can be put on 3D face models with high fidelity, which should be high resolutions.

Future work could explore combining PCA and MAE to leverage the strengths of both approaches and it's very possible to reconstruct with several images instead of just one image. Reconstruction of high resolution images is also one

of the directions. Additionally, exploring advanced neural network architectures and enhancing training data quality could further improve reconstruction results.

7 References

He, K., Chen, X., Xie, S., Li, Y., Dollar, P., & Girshick, R. (2021). Masked Autoencoders Are Scalable Vision Learners. arXiv preprint arXiv:2111.06377.

8 Appendix



Figure 11: grid mask image reconstructed by PCA



Figure 12: grid mask image reconstructed by PCA



Figure 13: center masked image reconstructed by PCA



Figure 14: center masked image reconstructed by PCA



Figure 15: left masked image reconstructed by PCA



Figure 16: left masked image reconstructed by PCA



Figure 17: outer masked image reconstructed by PCA



Figure 18: center masked image reconstructed by PCA



Figure 19: 0.25 mask ratio random masked image reconstructed by PCA



Figure 20: 0.25 mask ratio random masked image reconstructed by PCA



Figure 21: 0.5 mask ratio random masked image reconstructed by PCA



Figure 22: 0.5 mask ratio random masked image reconstructed by PCA



Figure 23: 0.75 mask ratio random masked image reconstructed by PCA



Figure 24: 0.75 mask ratio random masked image reconstructed by PCA

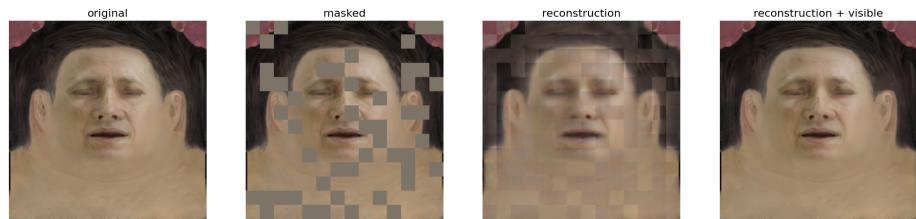


Figure 25: 0.25 mask ratio random masked image reconstructed by MAE

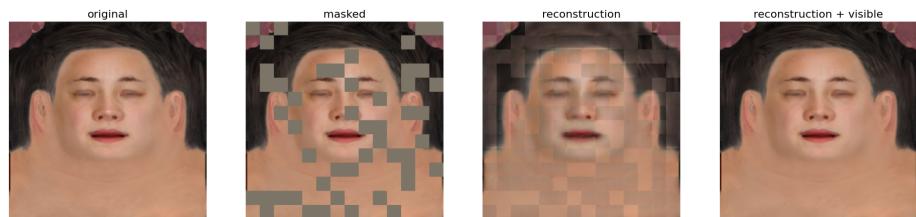


Figure 26: 0.25 mask ratio random masked image reconstructed by MAE

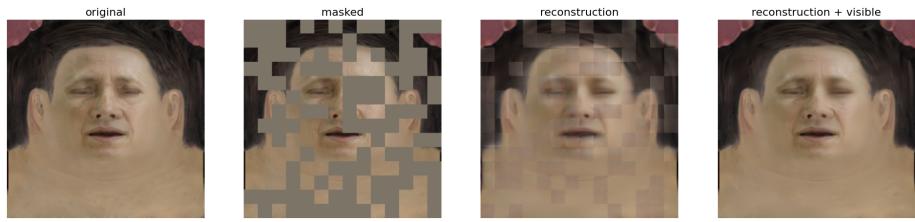


Figure 27: 0.5 mask ratio random masked image reconstructed by MAE

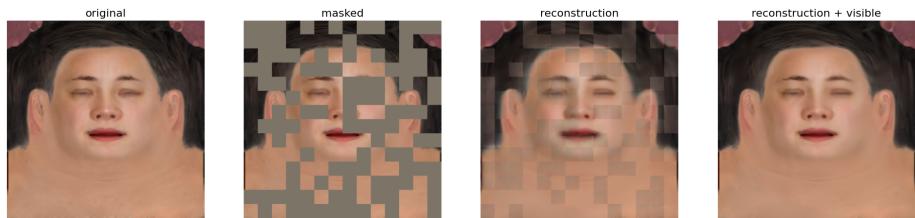


Figure 28: 0.5 mask ratio random masked image reconstructed by MAE



Figure 29: 0.75 mask ratio random masked image reconstructed by MAE

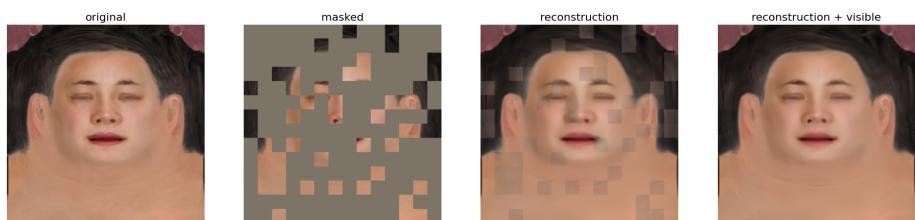


Figure 30: 0.75 mask ratio random masked image reconstructed by MAE

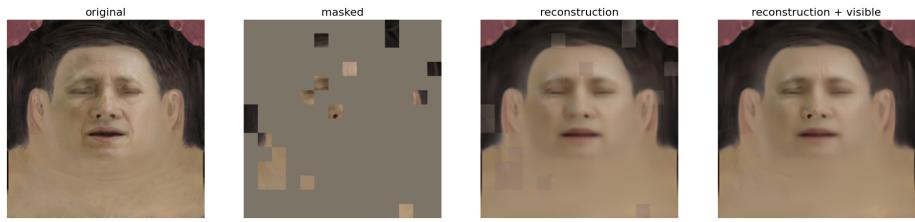


Figure 31: 0.9 mask ratio random masked image reconstructed by MAE

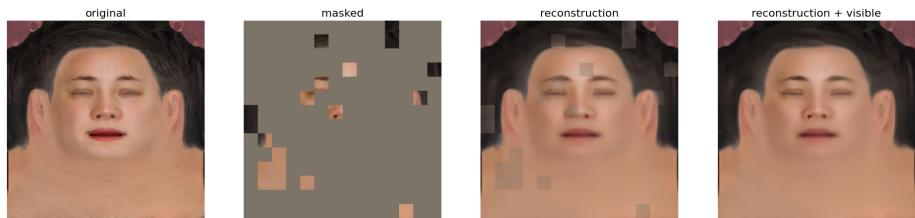


Figure 32: 0.9 mask ratio random masked image reconstructed by MAE