DualAvatar: Robust Gaussian Avatar with Dual Representation

Jinsong Zhang[1,2], I-Chao Shen[2], Jotaro Sakamiya[2], Yu-Kun Lai[3], Takeo Igarashi[2*], Kun Li[1*]

[1]Tianjin University, [2]The University of Tokyo, [3]Cardiff University, [*]Corresponding author

Goal

Given monocular images of a person rotating in an A-pose, our goal is to reconstruct a robust GS avatar that has fewer artifacts under unseen poses.

Related Work

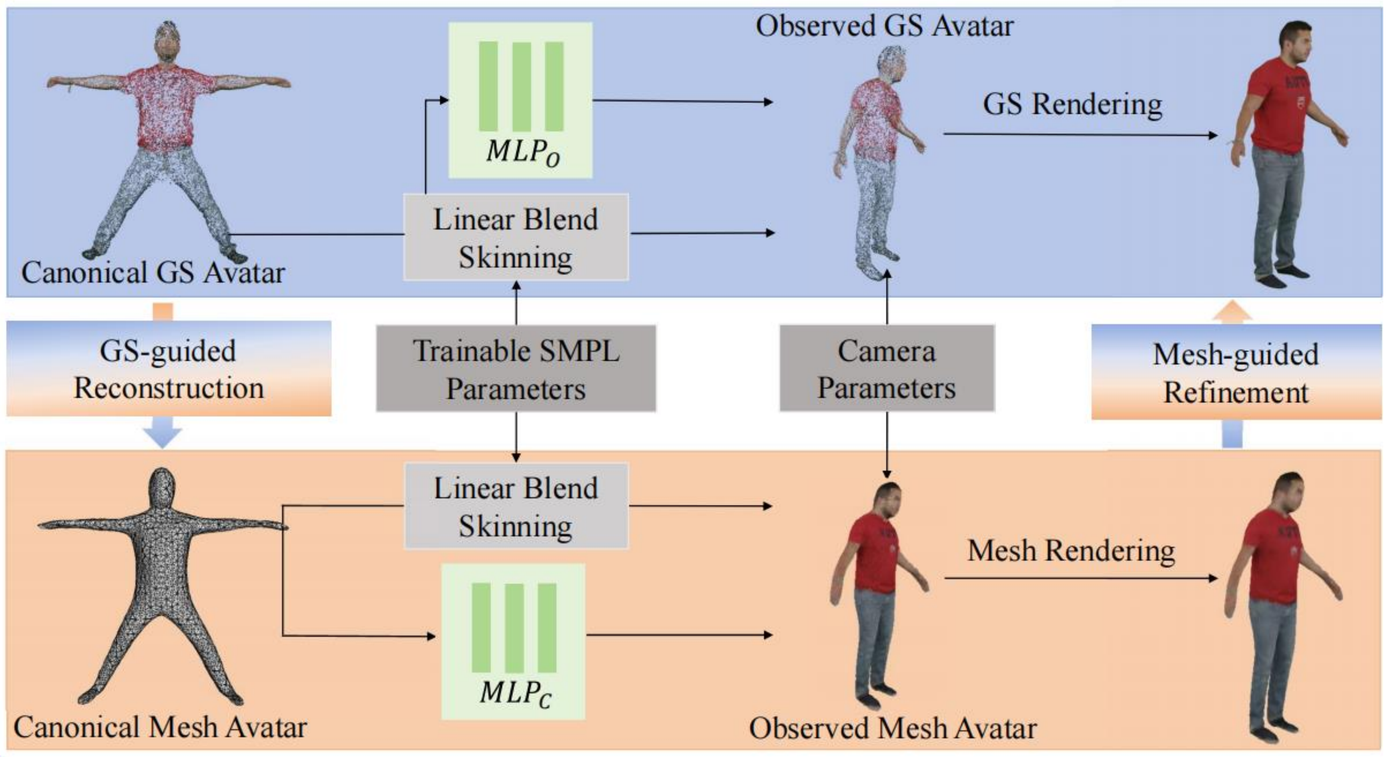
GART (Gaussian Articulated Template Model) [1] models deformable geometry and appearance using dynamic 3D Gaussians [2]. They achieve superior performance on avatar reconstruction, but their method generates unsatisfactory rendering results for unseen poses and unseen regions due to limited observations.

Mesh-based avatars offer greater robustness in handling animation and unseen areas, their rendering quality remains suboptimal.

Motivation

The mesh-based avatars [3] may have low-quality appearance, but they are robust to invisible regions and are capable of adapting to novel poses effectively. By leveraging the mesh avatar, we can refine the unseen regions and poses of the GS avatar.

Approach



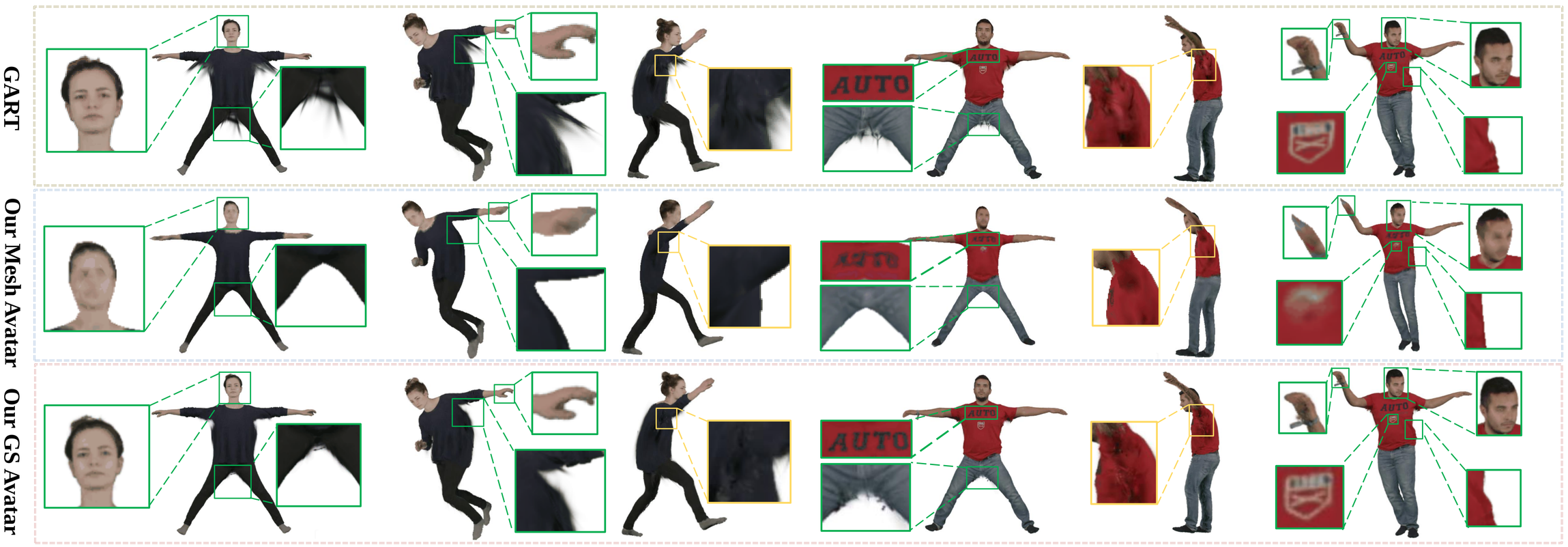
GS-guided Reconstruction: Chamfer distance between GS avatar and mesh avatar in canonical space to provide geometric information for mesh avatar

Mesh-guided Refinement: $\mathcal{L}_1$ regularization between rendered images of GS avatar and mesh avatar with specific poses to inpaint invisible regions

Other loss functions: $\mathcal{L}_1$ term and D-SSIM term between rendered images and the target images

Results

- Dataset: PeopleSnapshot [4](20 images for training)
- Our method can generate high-quality avatar with robust animation
- Our method can inpaint visible regions with the guidance of mesh avatar, and do not decrease the performance of GS avatar



Reference

[1] Lei, Jiahui, et al. "GART: Gaussian articulated template models." in *CVPR*. 2024.
[2] Kerbl, Bernhard, et al. "3D Gaussian Splatting for Real-Time Radiance Field Rendering." *ToG*. 42.4 (2023): 139-1.
[3] Shen, Tianchang, et al. " Deep Marching Tetrahedra: a hybrid representation for high-resolution 3D shape synthesis. ” in *NeurIPS, 2021*.
[4] Alldieck, Thiemo, et al. "Video based reconstruction of 3D people models." in *CVPR*. 2018.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (62122058 and 62171317), and the Science Fund for Distinguished Young Scholars of Tianjin (No. 22JCJQJC00040), and JSPS Grant-in-Aid JP23K16921, Japan..