Gaussian Head Avatar: Ultra High-fidelity Head Avatar via Dynamic Gaussians

Supplementary Material

1. Training Details

During the geometry-guided initialization stage, we use an Adam optimizer, and set the learning rate to 1×10^{-3} for all the networks and 1×10^{-4} for the neutral 3D landmarks $\boldsymbol{P}_0$. Then We train the model for 10000 iterations with a batch size of 4. During the Gaussian model training stage, we also use an Adam optimizer, and set the learning rate to 1×10^{-4} for the two color MLPs, the two deformation MLPs and the two attribute MLPs, 1×10^{-5} for the neutral positions $\boldsymbol{X}_0$ and the point-wise feature vectors $\boldsymbol{F}_0$, 1×10^{-5} for the neutral rotation $\boldsymbol{Q}_0$, 3×10^{-5} for the neutral scale $\boldsymbol{S}_0$, 1×10^{-4} for the neutral opacity $\boldsymbol{Q}_0$ and 1×10^{-4} for the super resolution network $\boldsymbol{\Psi}$. Finally, we train the Gaussian model for 600000 iterations with a batch size of 1 until fully convergence.

2. Failure Case

For non-face areas, our method inputs the head pose as the condition to control the deformation, which is not able to model the complex dynamic deformation of long hair, resulting in blurred rendering results as shown in Fig. 1. On the other hand, the reconstructed head avatar cannot make expressions other than those in the training set. Therefore, when the actor's expression is too exaggerated, our method will output relatively less exaggerated results as shown in Fig. 2.

Figure 1. Failure case: our method can not reconstruct dynamic long hair.

Figure 2. Failure case: our method produce relatively less exaggerated results.