Relightable and Animatable Neural Avatar from Sparse-View Video

Zhen Xu Sida Peng Chen Geng Linzhan Mou
Zihan Yan Jiaming Sun Hujun Bao Xiaowei Zhou

Project Page:
https://zju3dv.github.io/relightable_avatar

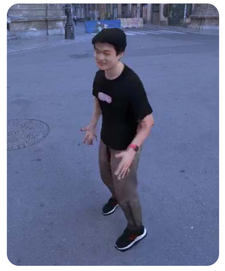
Contact:
zhenx@zju.edu.cn

(highlight)
CVPR
SEATTLE, WA
JUNE 17-21, 2024

Input & Output

Input: **sparse-view video**

Condition:
novel pose & nove lighting



Output:
animatable & relightable avatar

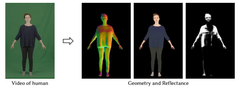
Our method takes only a sparse-view (or monocular) video as input and reconstructs a relightable and animatable neural avatar.

Motivation

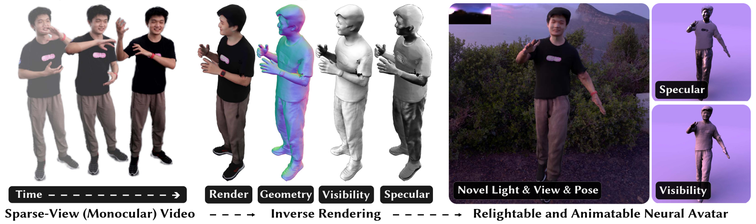
NeuralBody:
non-relightable non-animatable

AnimatableNeRF / AnimatableSDF:
non-relightable / non-relightable animatable / animatable

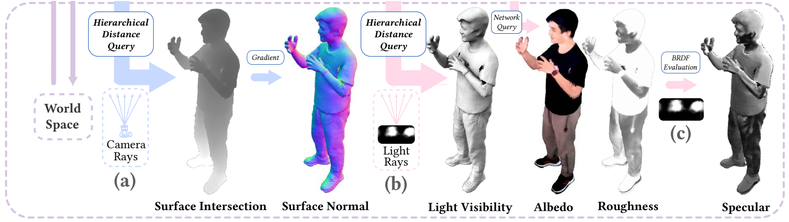
Relighting4D:
relightable non-animatable



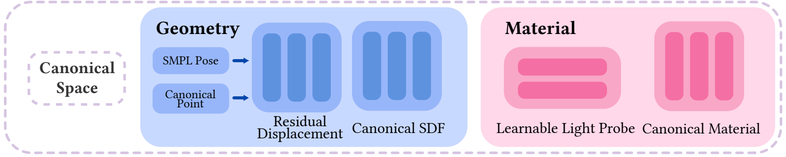
Ours (RelightableAvatar): **relightable & animatable neural human avatar**



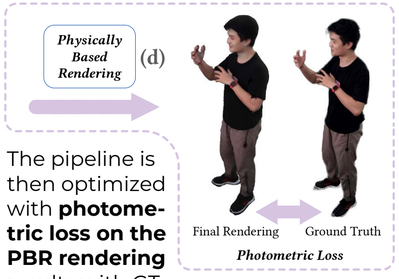
Our Method



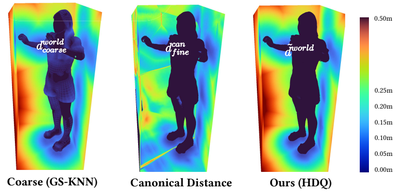
Given world space camera rays, we perform **sphere tracing on the hierarchically queried distances** to find surface intersections.

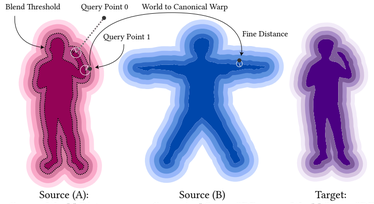


Material properties and surface normals are **queried on the canonical correspondences** and warped to world space using the SMPL prior.



The pipeline is then optimized with **photometric loss on the PBR rendering** results with GT.



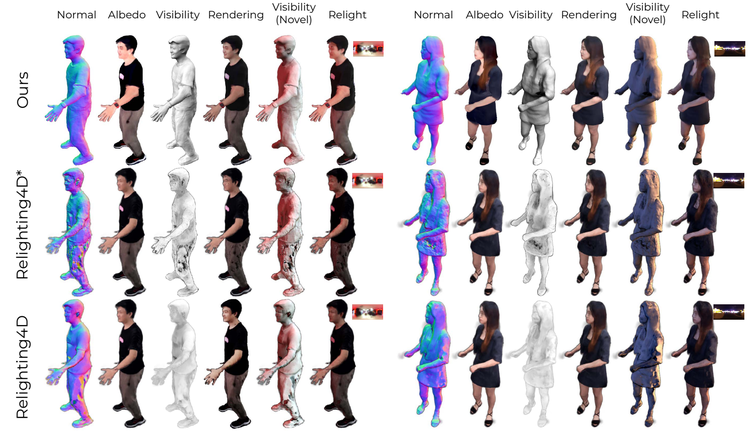


$$d^{world}_{coarse} = \frac{\sum_{k=0}^{K} d_k}{K} \qquad d^{can}_{fine} = S(\boldsymbol{x}'') \qquad \tilde{d}^{world} =$$

$$\begin{cases} d^{world}_{coarse} & \text{, if } d^{world}_{coarse} > \tilde{T}_d \\ d^{can}_{fine}(1 - \frac{d^{can}_{fine}}{T_d}) + d^{world}_{coarse}\frac{d^{can}_{fine}}{T_d} & \text{, otherwise} \end{cases}$$

For point 0, it falls out of the cut-off threshold. For point 1, it is within the range of local SDF values. Thus we only use the coarse distance for point 0 and **blend the coarse and fine distances** for point 1.

Comparison Results



Animation & Relighting