



中山大學
SUN YAT-SEN UNIVERSITY



3D Gaussian Splatting for One-shot Hand Avatars

Xuan Huang^{1*}, Hanhui Li^{1*}, Wanquan Liu¹, Xiaodan Liang¹, Yiqiang Yan², Yuhao Cheng²,
Chengqiang Gao^{1†}

¹Shenzhen Campus of Sun Yat-sen University, Shenzhen, China

²Lenovo Research

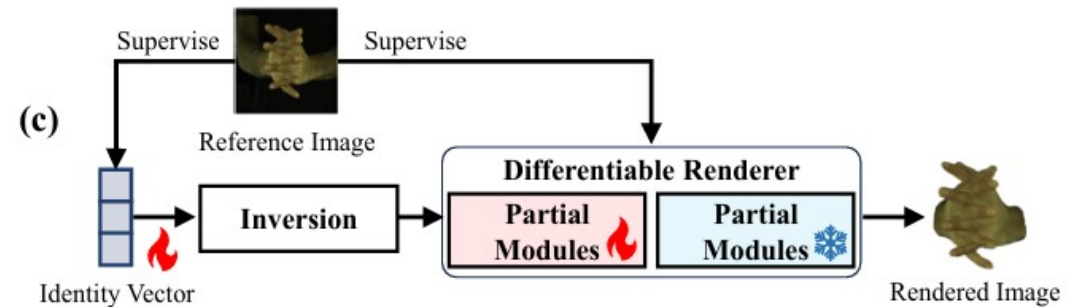
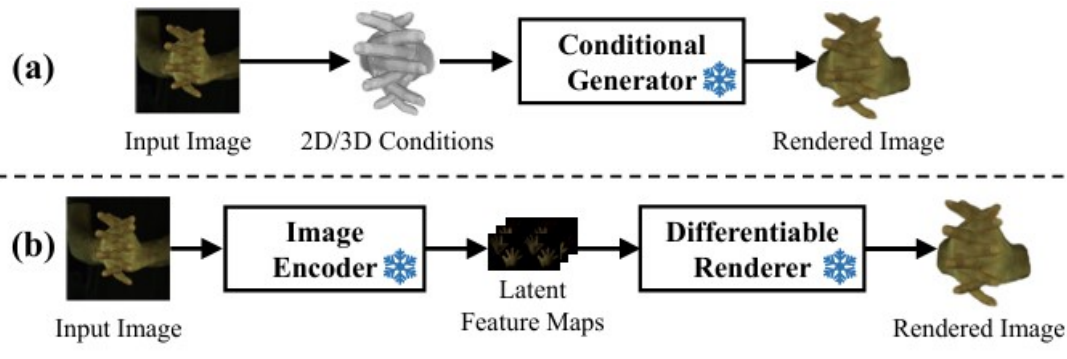
Introduction



- We present a novel interaction-aware Gaussian splatting framework that creates **animatable interacting hand avatars** from a single image.
- These high-fidelity avatars support **various applications**, such as editing, animation, combination, duplication, re-scaling, and text-to-avatar conversion.

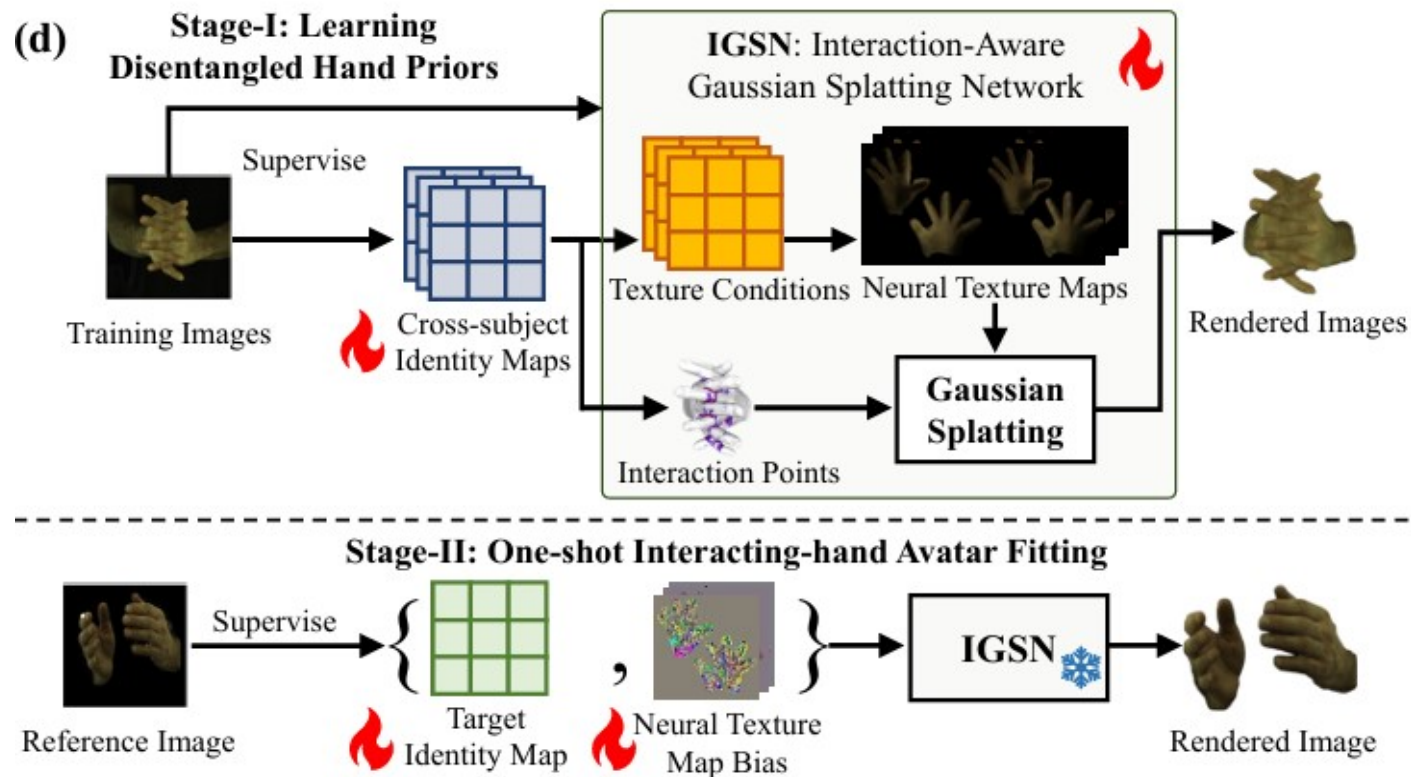
Motivation

- Recent advancements in 3D reconstruction have significantly improved hand avatar creation and related applications. However, creating avatars for **interacting hands** from a single image remains challenging.
- Existing GS-based methods designed for single subjects often yield unsatisfactory results due to limited input views, various hand poses, and occlusions.



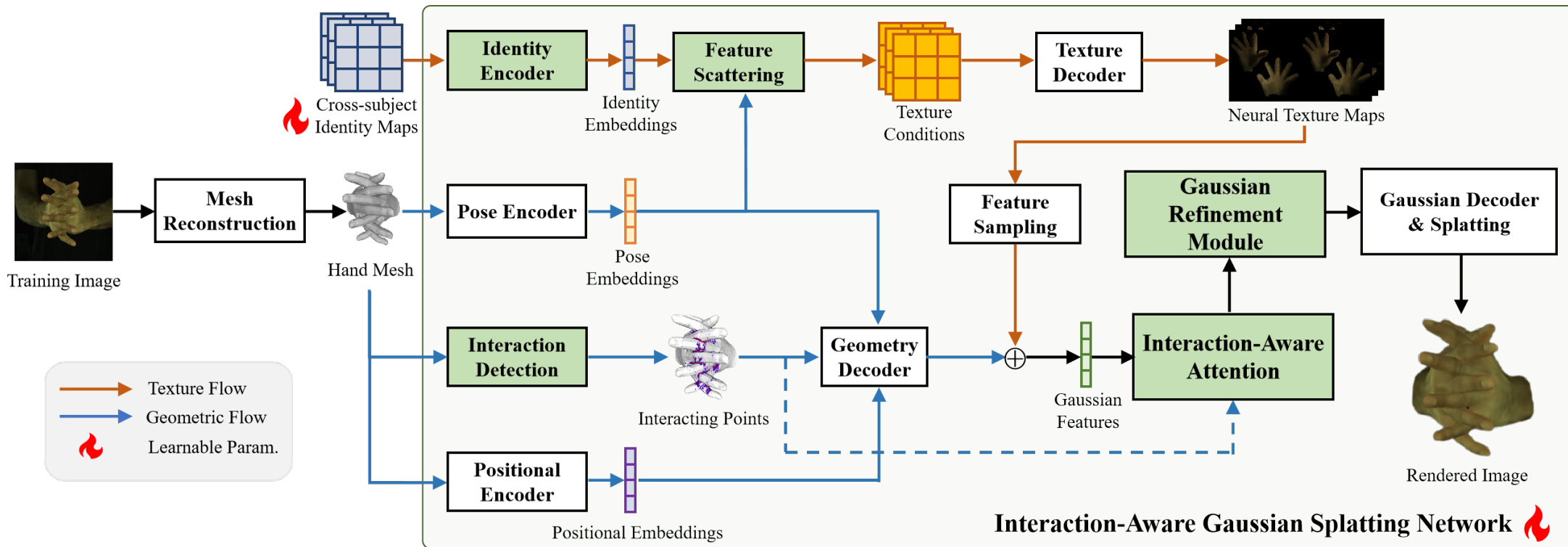
Motivation

- To address these challenges, we introduce a novel two-stage interaction-aware GS framework that exploits cross-subject hand priors and refines 3D Gaussians in interacting areas.

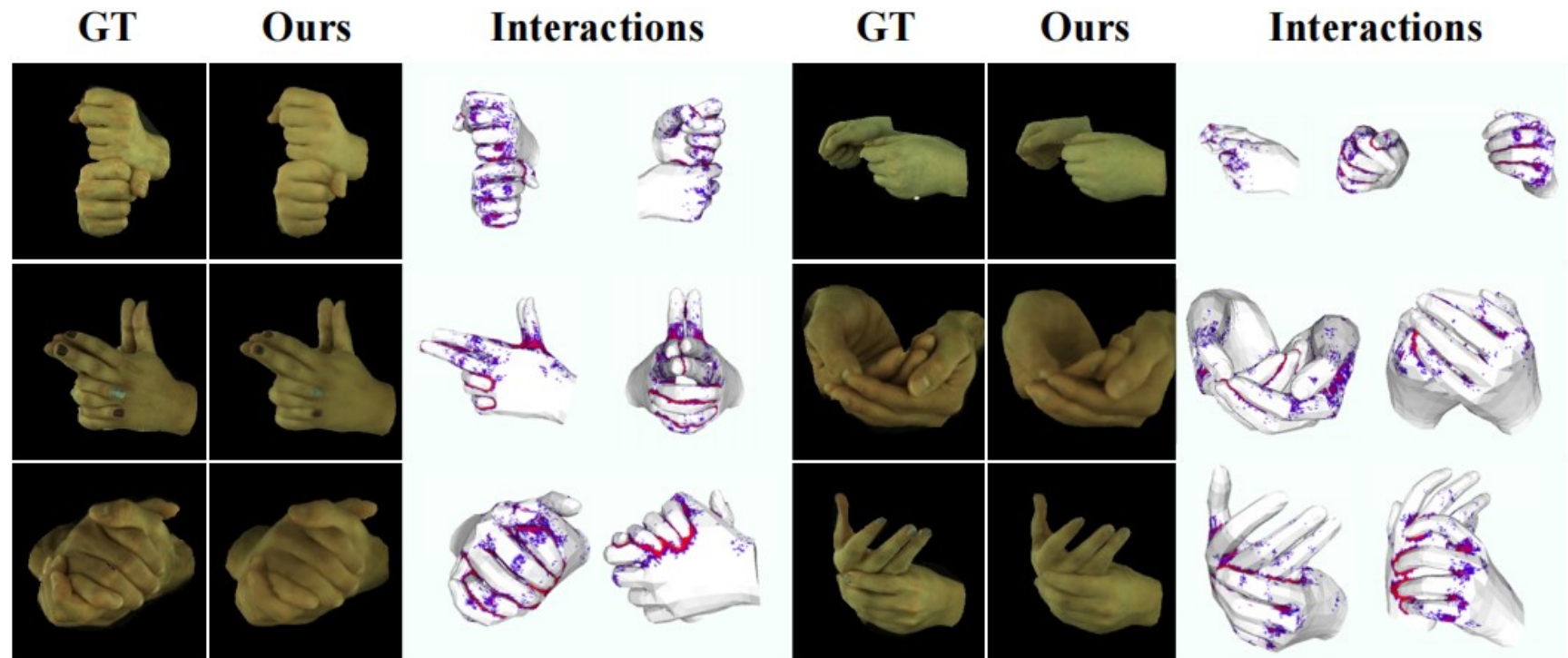


Methodology

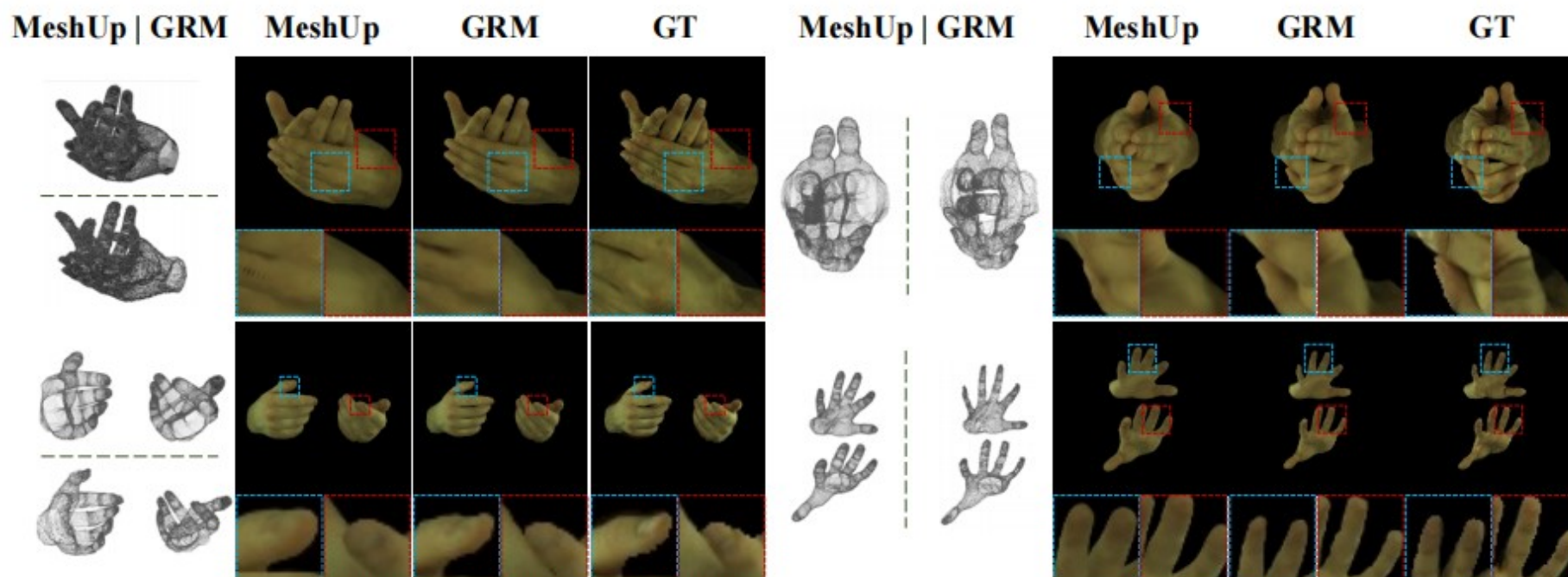
- The architecture of the proposed interaction-aware Gaussian splatting network, of which the core components are the disentangled hand representation, the interaction detection module, the interaction-aware attention module, and the Gaussian refinement module labeled in green.



- Visualization of intra- and inter-hand interactions detected by the proposed method. Areas with interactions from sparse to dense are labeled in colors from blue to red, respectively.

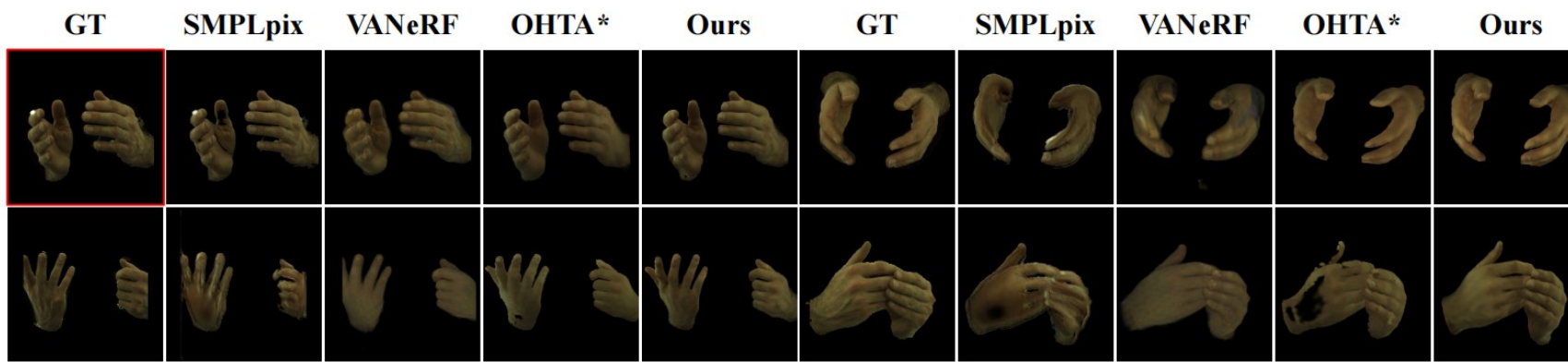


- Qualitative comparisons between Gaussians by mesh upsampling (denoted as MeshUp) and the proposed GRM.



- Project page: <https://github.com/XuanHuang0/GuassianHand>

| Method | Novel View Synthesis | | | Novel Pose Synthesis | | |
|--------------|----------------------|-----------------|--------------------|----------------------|-----------------|--------------------|
| | PSNR \uparrow | SSIM \uparrow | LPIPS \downarrow | PSNR \uparrow | SSIM \uparrow | LPIPS \downarrow |
| KeypointNeRF | 23.55 | 0.804 | 0.326 | - | - | - |
| SMPLpix | 24.50 | 0.868 | 0.170 | 24.26 | 0.854 | 0.173 |
| VANeRF | 25.38 | 0.848 | 0.226 | 24.42 | 0.822 | 0.250 |
| OHTA* | 25.31 | 0.851 | 0.184 | 25.93 | 0.880 | 0.156 |
| Ours | 26.14 | 0.869 | 0.161 | 26.56 | 0.890 | 0.133 |



- One shot synthesis comparison with state-of-the-art methods on Interhand2.6M.



- We propose a novel two-stage interaction-aware GS framework to create **animatable avatars for interacting hands from single-image inputs**. Our method generates high-fidelity rendering results and supports various applications.
- We **disentangle the 3D presentation of hands** into learning-based features that can be generalized well to different subjects and identity maps that are individually optimized for each subject. This disentanglement provides us with flexible and reliable priors for poses, shapes, and textures.
- We introduce an **interaction-aware attention module**, which identifies intra- and inter-hand interactions and further exploits interaction context to improve rendering quality.
- We devise a **Gaussian refinement module** that adaptively adjusts the number and positions of 3D Gaussian, which results in rendered images of higher quality under various hand poses and shapes.



THANKS

Q & A