Human Motion Capture and Avatar Creation for XR Scenarios using Sparse Observations

Zhuo Su 苏卓

ByteDance 字节跳动

1. Background

- XR Applications

Motion Capture | **3D Gaming**

Enhancing gaming, social interactions, and fitness for immersive experiences


PICO Motion Tracker

Avatar Creation | **3D communication**

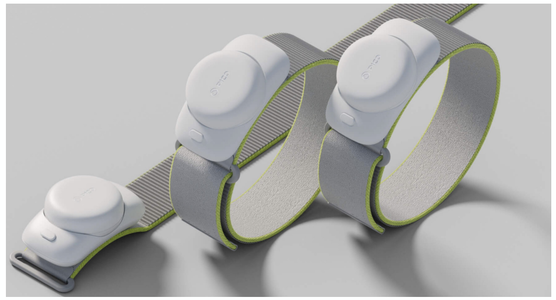
Connecting people through lifelike 3D reconstruction for enhanced interaction

Apple Vision Pro Persona

Meta Codec Avatar

- Sparse Observations

HMD and hand controllers



3DoF of motion trackers

Tracking signals

Ego-centric view images

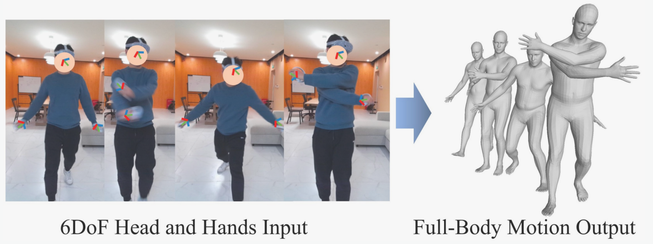
Front view image

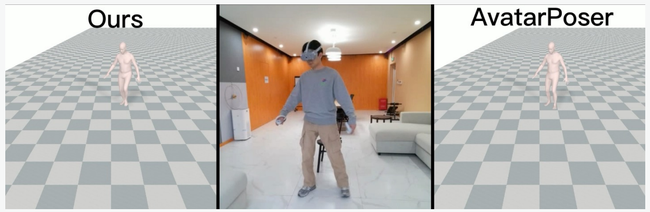
...

2. Motion Capture using Sparse Sensors

AvatarJLM

ICCV'23

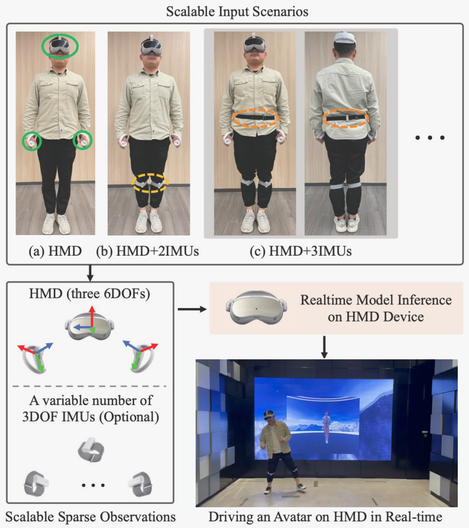




6DoF of HMD and hand controllers

HMD-Poser

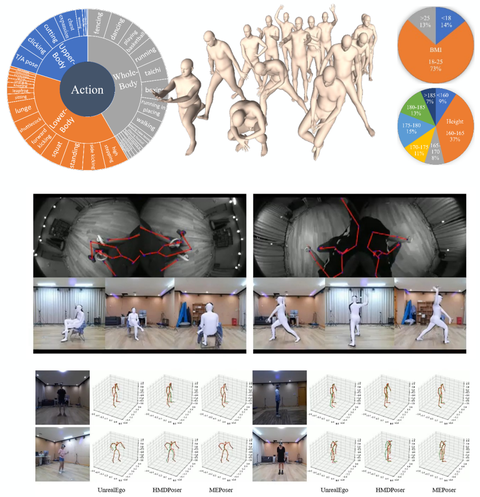
CVPR'24



+ scalable IMUs on legs or pelvis

EMHI (MEPoser)

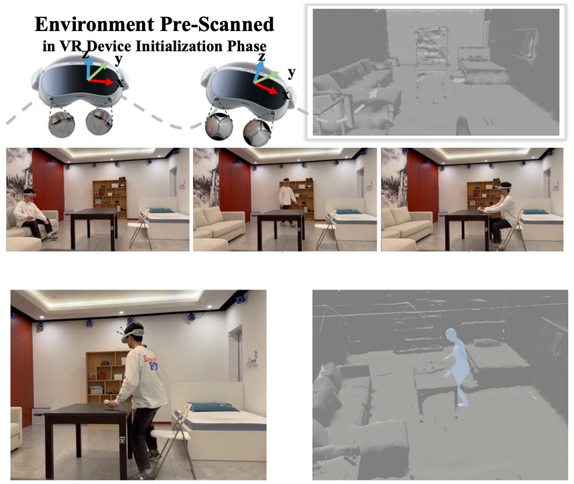
AAAI'25



+ egocentric cameras of HMD

EnvPoser

CVPR'25



+ 3D environment from VR device

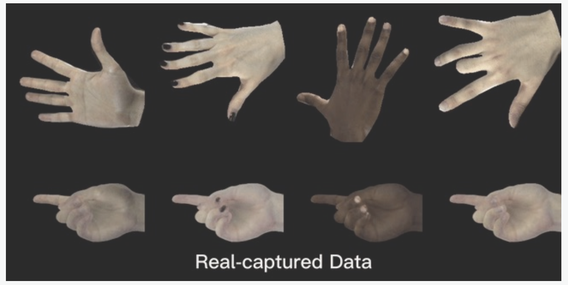
* Explore these research papers at https://suzhuo.github.io/

ByteDance字节跳动

3. Avatar Creation using One/Few Image(s)

OHTA

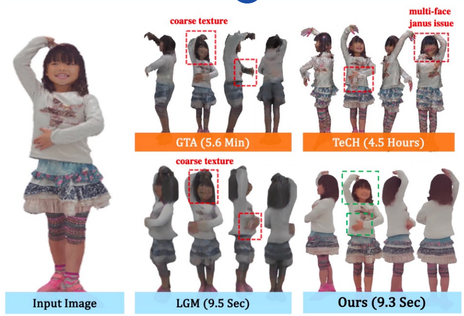
CVPR'24

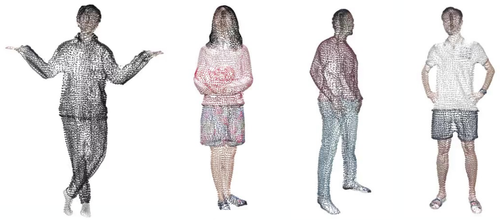


One-shot hand avatar creation

HumanSplat

NeurIPS'24

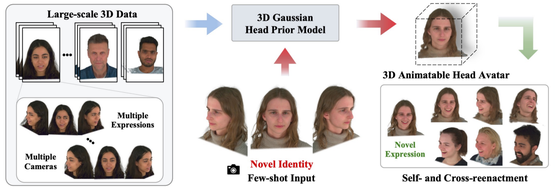


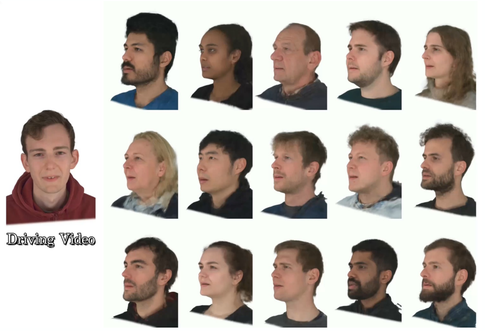


Single-Image full-body Gaussian Splatting

HeadGap

3DV'25

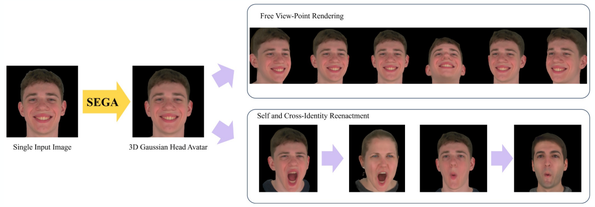




Few-shot head avatar creation

SEGA

arXiv'25



One-shot head avatar creation

* Explore these research papers at https://suzhuo.github.io/

4. Conclusion & Future Work

- Conclusion

Over the past year, my research focused on addressing the challenges of sparse observations in XR scenarios, yielding significant achievements:

- **Motion Capture:** Enabled high-accuracy motion capture under sparse sensor configurations
- **Avatar Creation:** Achieved high-fidelity reconstruction of hands, faces, and full-body avatars from single or few images

These advancements enable immersive XR applications and drive real-world adoption.

- Future Work

- **Real-World Robustness:** Expand datasets, optimize generalizable algorithm, and enhance real-time performance in real-world scenarios
- **Integration with Video Foundation Models:** Further Leverage large generative models to enhance motion realism and avatar fidelity
- **Seamless Motion-Avatar Integration:** Tightly integrate motion capture and avatar creation to enable real-time, full-body XR experiences

- Acknowledgments

Thanks to my collaborators and interns (Xiaozheng Zheng , Panwang Pan, Chen Guo, Zhen Fan, Peng Dai, Songpengcheng Xia, ...).

THANKS