

Compensating Motion-Induced Errors in Smartphone-Based VR Avatar Reconstruction

Supplementary Material

Friedemann Runte*
TU Dortmund University
Dortmund, Germany

Ulrich Schwanecke
RheinMain University of Applied Sciences
Wiesbaden, Germany

Timo Menzel*[†]
TU Dortmund University
Dortmund, Germany

Mario Botsch[†]
TU Dortmund University
Dortmund, Germany

CCS Concepts

• **Computing methodologies** → **Mesh geometry models**.

Keywords

Virtual Humans, 3D Reconstruction, Avatars, Template Fitting

ACM Reference Format:

Friedemann Runte, Timo Menzel, Ulrich Schwanecke, and Mario Botsch. 2025. Compensating Motion-Induced Errors in Smartphone-Based VR Avatar Reconstruction Supplementary Material. In *31st ACM Symposium on Virtual Reality Software and Technology (VRST '25), November 12–14, 2025, Montreal, QC, Canada*. ACM, New York, NY, USA, 17 pages. <https://doi.org/10.1145/3756884.3765995>

In Figures 1 – 11, we show results from iHuman [3], RMAvatar [4], Avatars for the Masses (A4M) [2], and our method in training poses. This means the subject is rendered as it was captured. Subjects 2, 8, and 11 asked to remain anonymous.

Figures 12 – 22 show results from iHuman, RMAvatar, A4M, and our method in novel poses, that were not seen during capture/training.

Figures 23 and 24 show a female subject with loose, curly hair and a long skirt and a male subject with long hair. These are examples of worst-case conditions for the reconstruction with our approach. Loose hair in windy conditions is difficult to reconstruct, as seen in Figure 25 (see also the corresponding avatar in Figure 42). While the point cloud captures the face perfectly, hair is almost not present. In contrast to Gaussian-based approaches, the training image reconstruction appears considerably worse. In novel pose, iHuman and RMAvatar lose all detail and become blurry, while A4M and ours keep a consistent quality. Animation without information about the clothing itself leads to artifacts as the template is modeled wearing minimal clothing.

In Figures 26 – 44 we show more generated avatars from four directions, including previous subjects and additional ones.

*Both authors contributed equally to this research.

[†] Also with Lamarr Institute for Machine Learning and Artificial Intelligence.

References

- [1] Agisoft. 2025. Agisoft Metashape. <https://www.agisoft.com> [visited on 2025-06-05].
- [2] Timo Menzel, Erik Wolf, Stephan Wenninger, Niklas Spinczyk, Lena Holderrieth, Carolin Wienrich, Ulrich Schwanecke, Marc Erich Latoschick, and Mario Botsch. 2025. Avatars for the Masses: Smartphone-Based Reconstruction of Humans for Virtual Reality. *Frontiers in Virtual Reality* 6 (2025), 1583474. doi:10.3389/frvir.2025.1583474
- [3] Pramish Paudel, Anubhav Khanal, Danda Pani Paudel, Jyoti Tandukar, and Ajad Chhatkuli. 2025. iHuman: Instant Animatable Digital Humans From Monocular Videos. In *Computer Vision – ECCV 2024*. 304–323. doi:10.1007/978-3-031-73226-3_18
- [4] Sen Peng, Weixing Xie, Zilong Wang, Xiaohu Guo, Zhonggui Chen, Baorong Yang, and Xiao Dong. 2025. RMAvatar: Photorealistic human avatar reconstruction from monocular video based on rectified mesh-embedded Gaussians. *Graphical Models* 139 (2025), 101266. doi:10.1016/j.gmod.2025.101266



This work is licensed under a Creative Commons Attribution 4.0 International License. *VRST '25, Montreal, QC, Canada*

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2118-2/2025/11

<https://doi.org/10.1145/3756884.3765995>

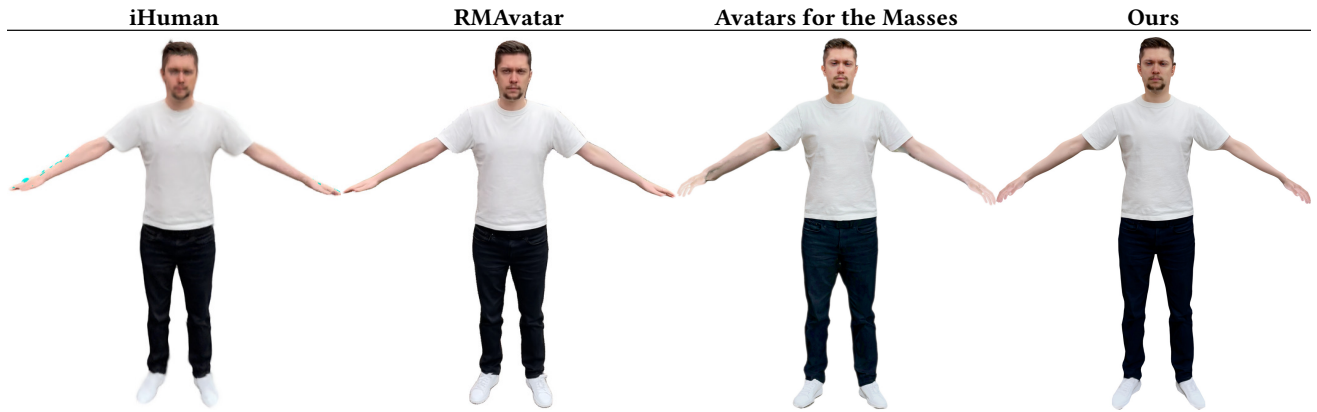


Figure 1: Subject 1 in training pose

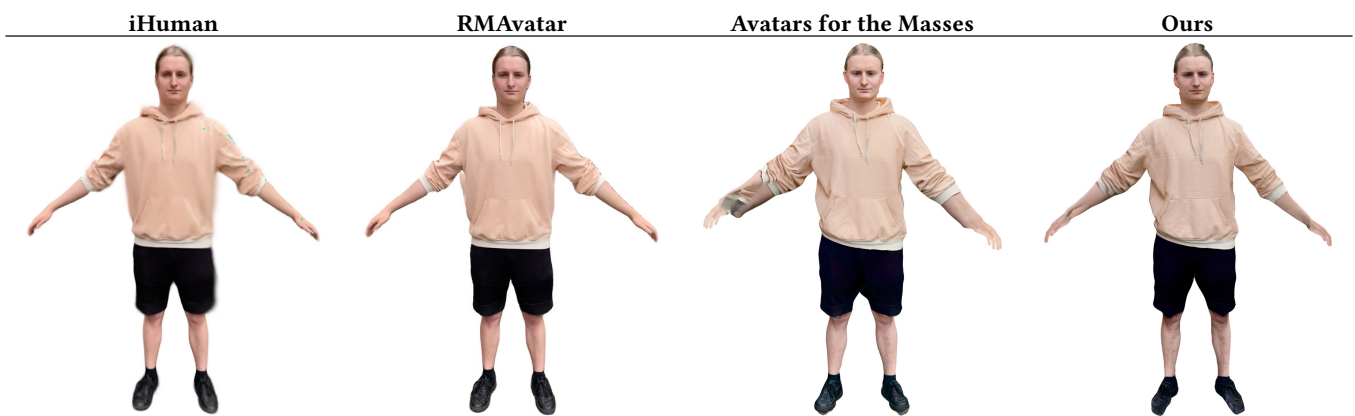


Figure 2: Subject 2 in training pose

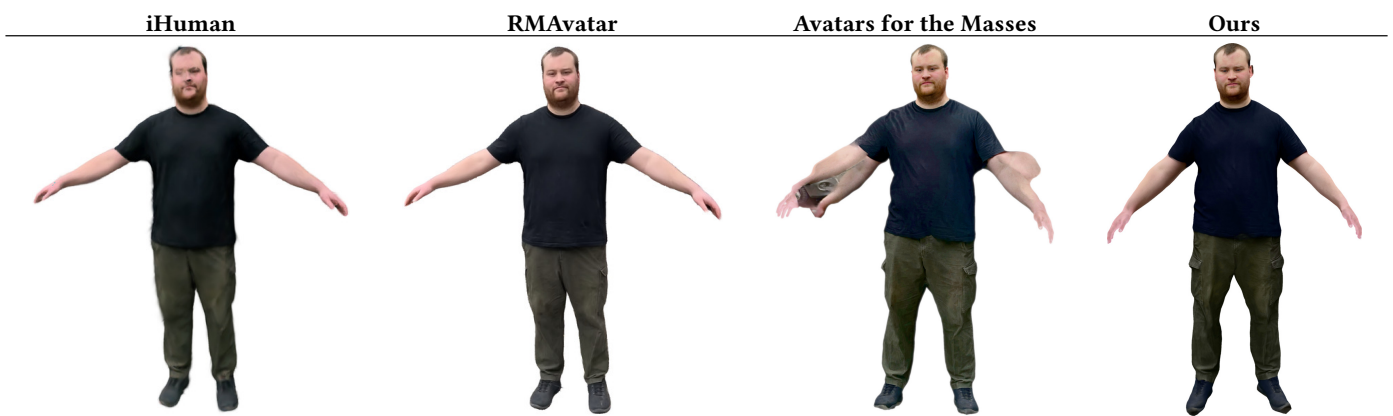


Figure 3: Subject 3 in training pose

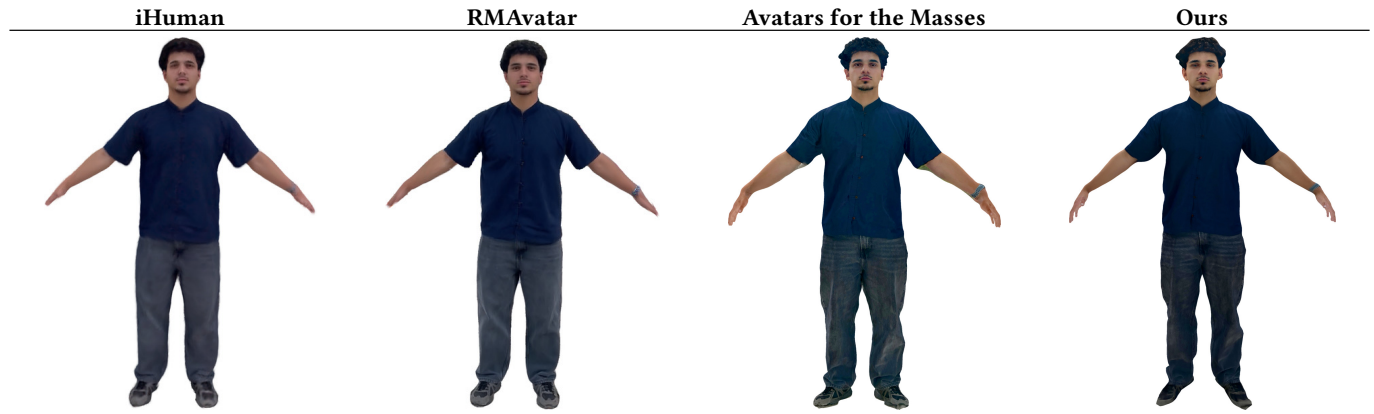


Figure 4: Subject 4 in training pose

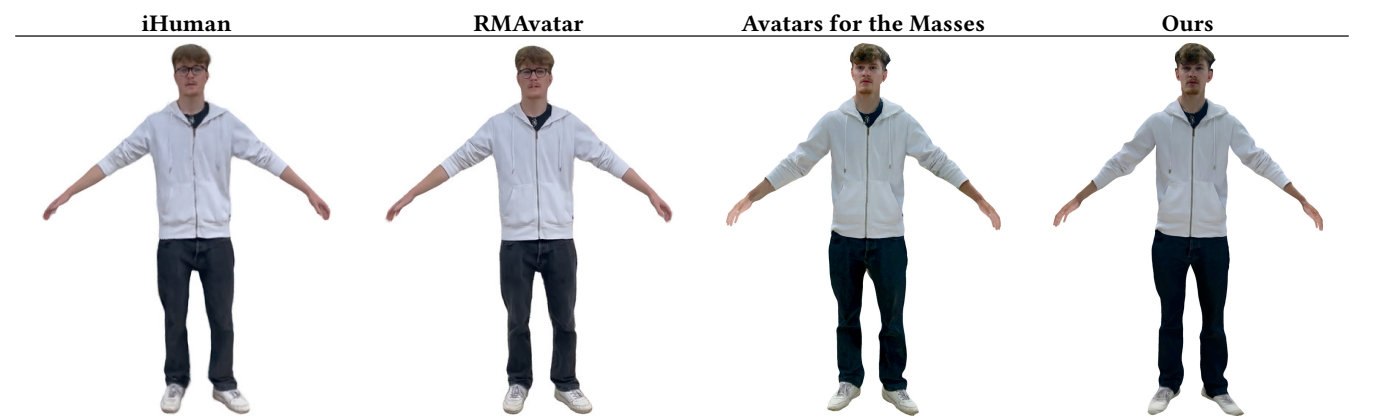


Figure 5: Subject 5 in training pose

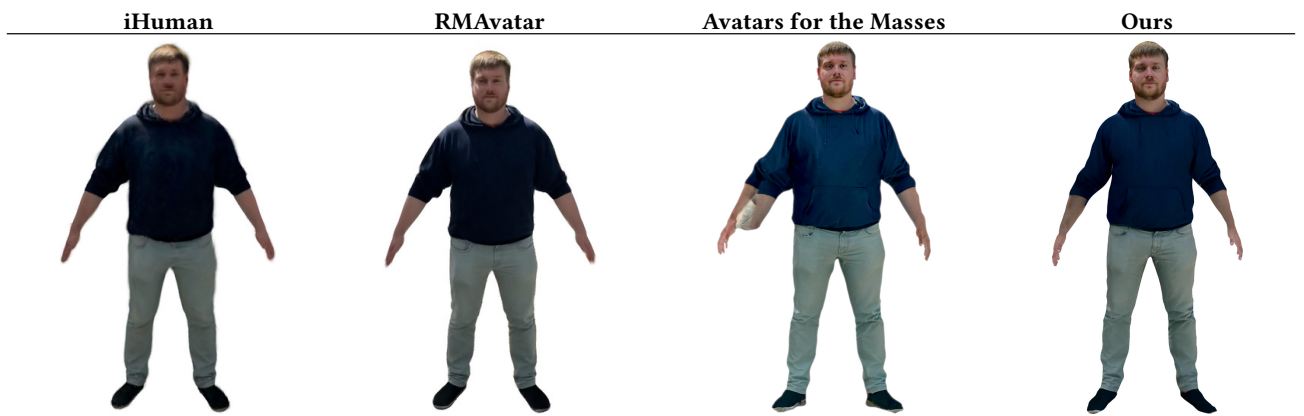


Figure 6: Subject 6 in training pose

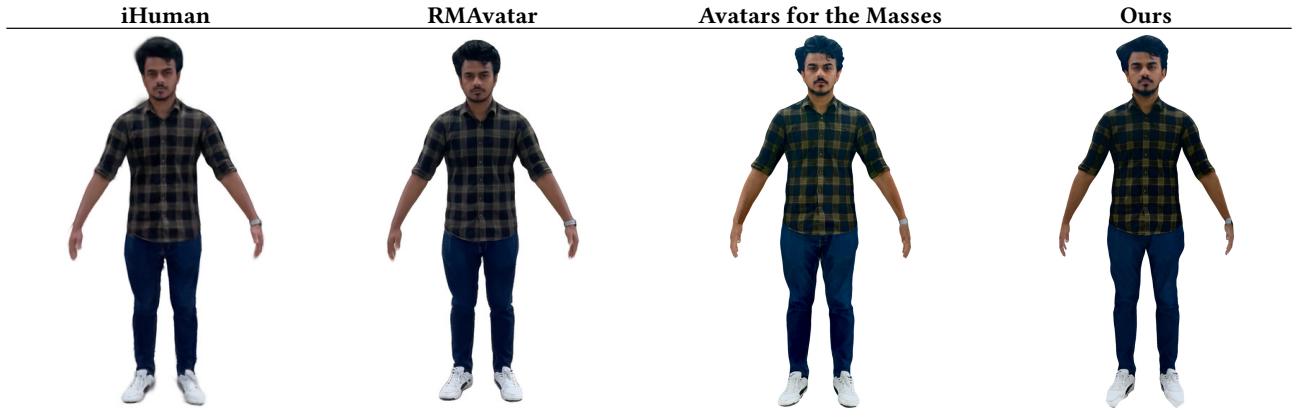


Figure 7: Subject 7 in training pose

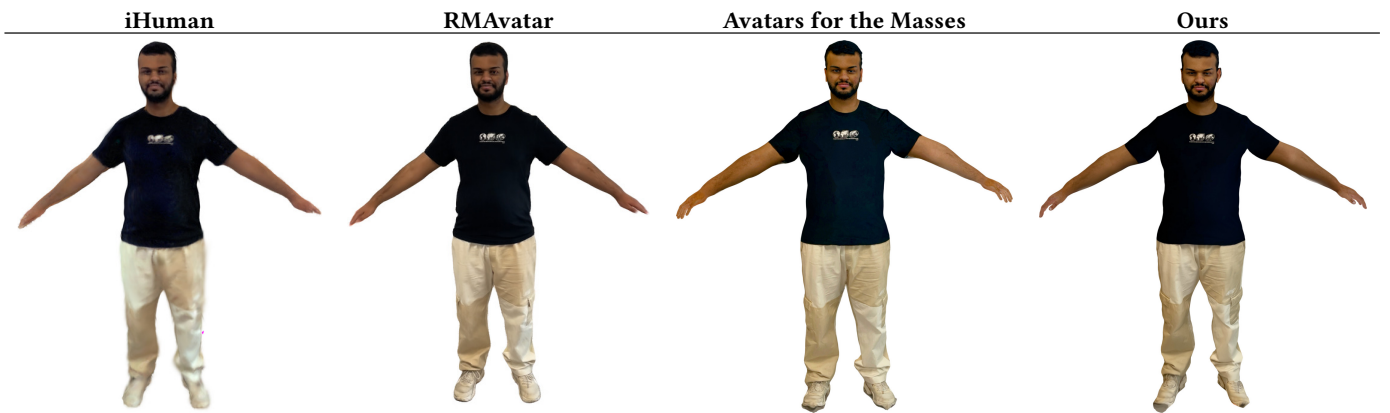


Figure 8: Subject 8 in training pose

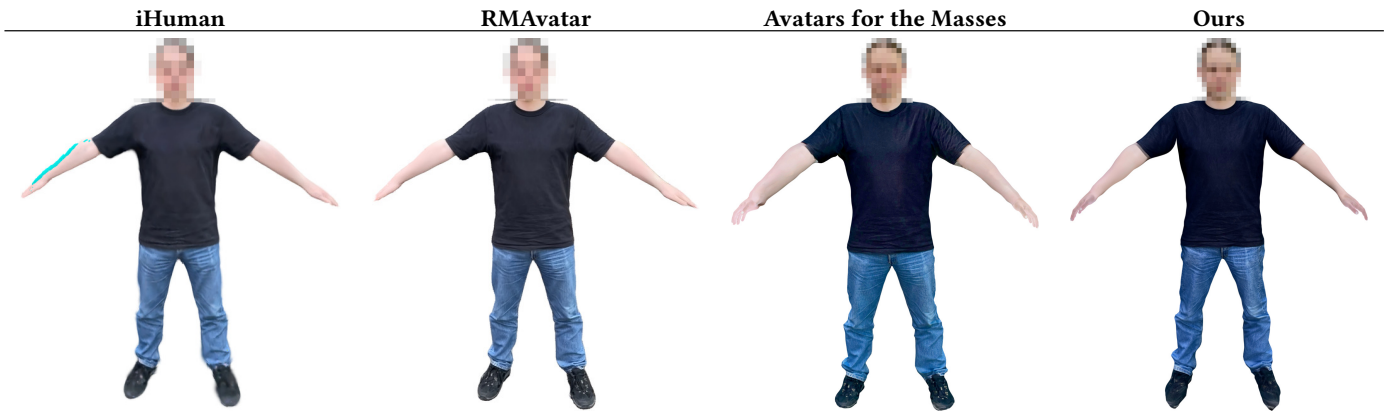


Figure 9: Subject 9 in training pose

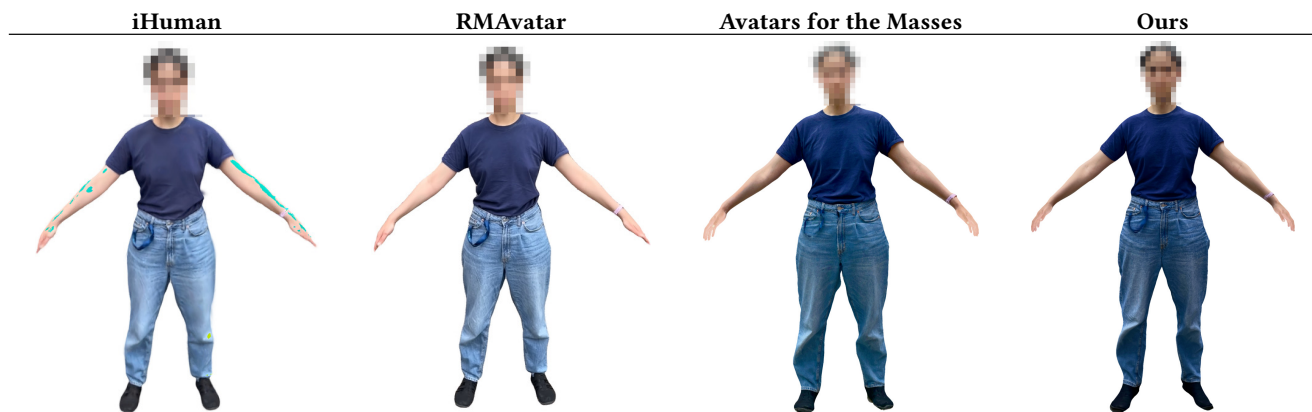


Figure 10: Subject 10 in training pose

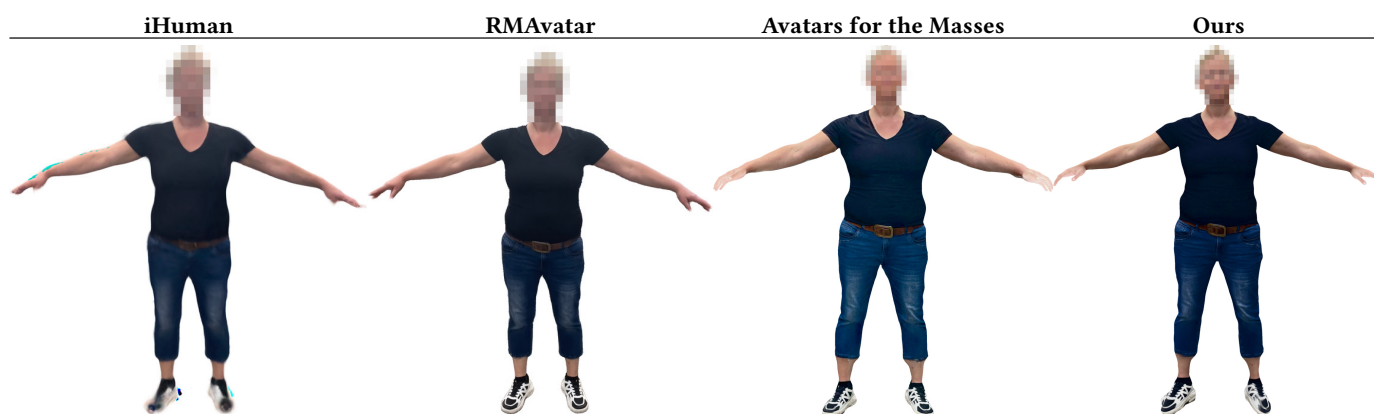


Figure 11: Subject 11 in training pose

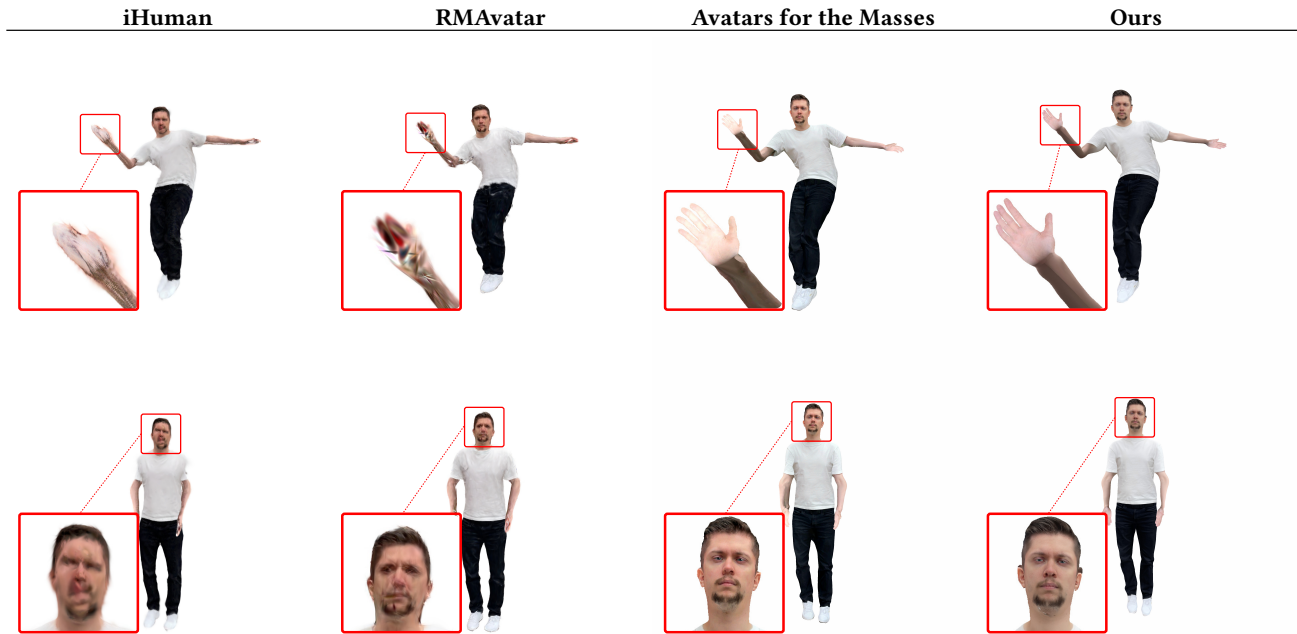


Figure 12: Subject 1 in novel pose

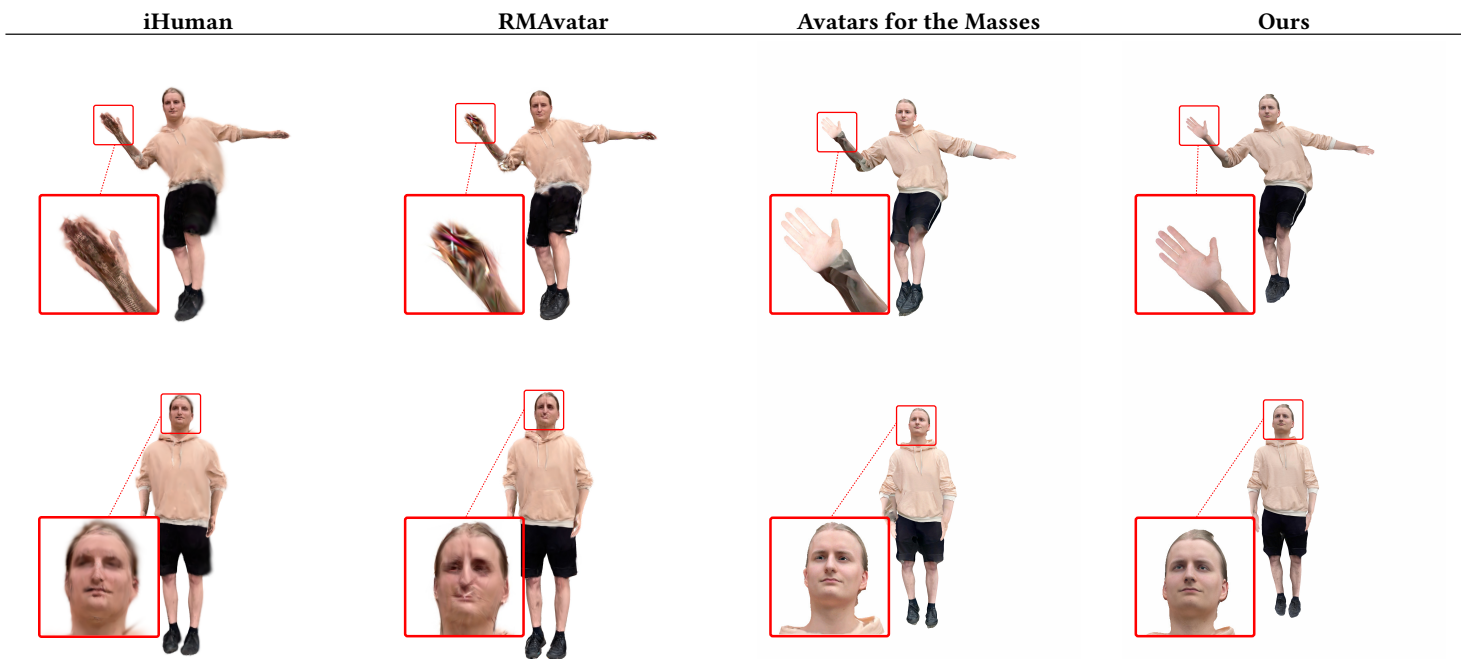


Figure 13: Subject 2 in novel pose

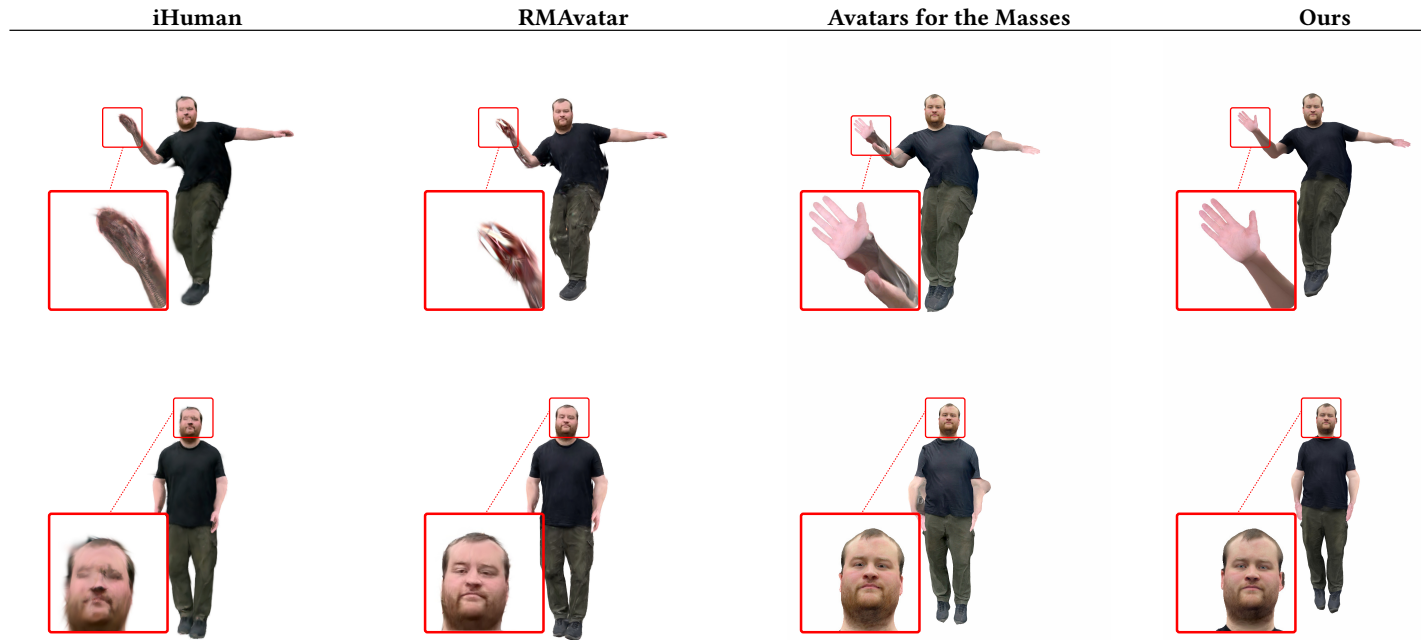


Figure 14: Subject 3 in novel pose

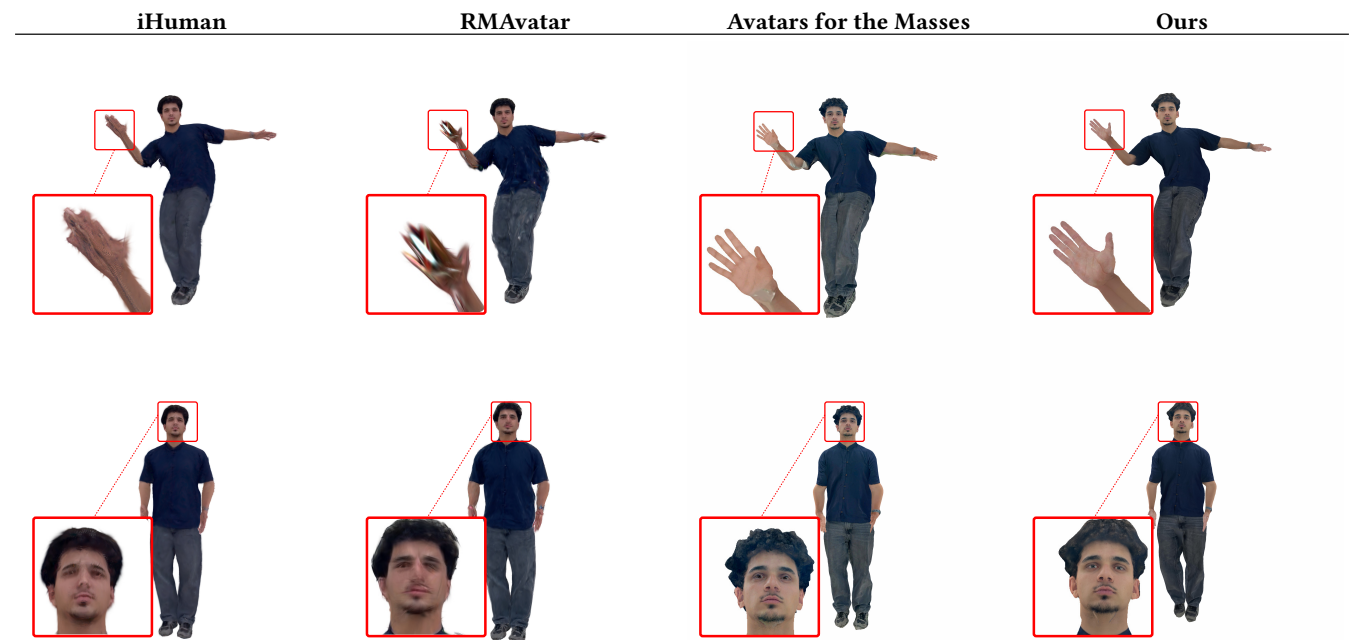


Figure 15: Subject 4 in novel pose

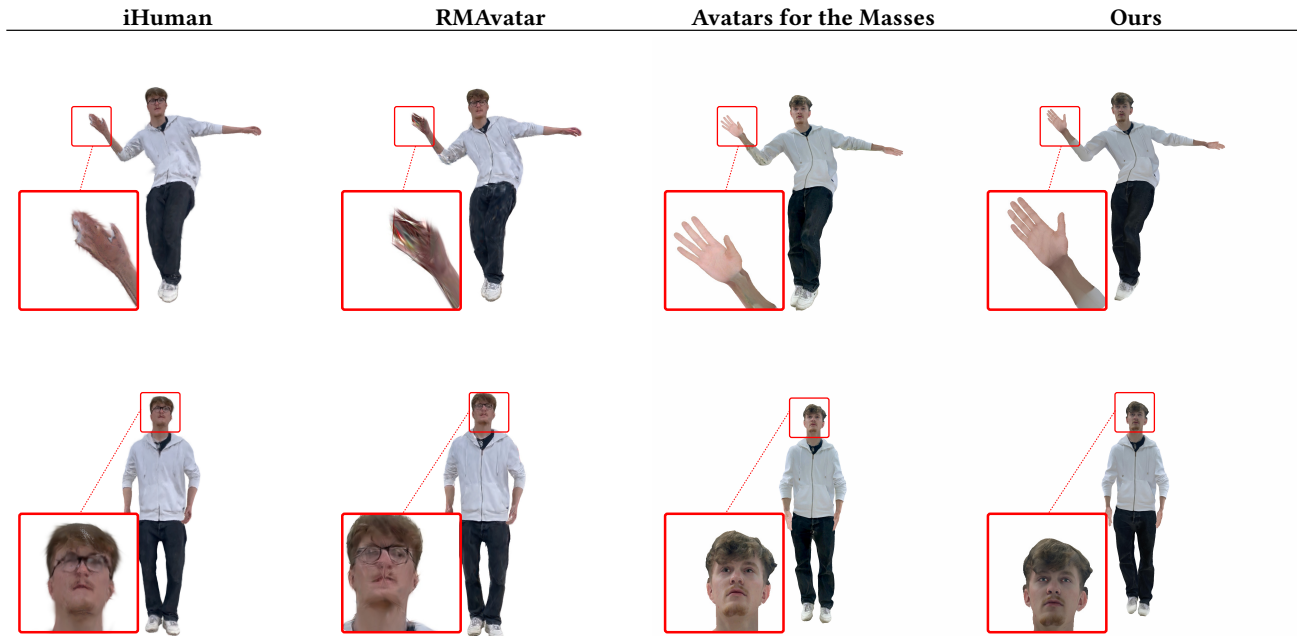


Figure 16: Subject 5 in novel pose

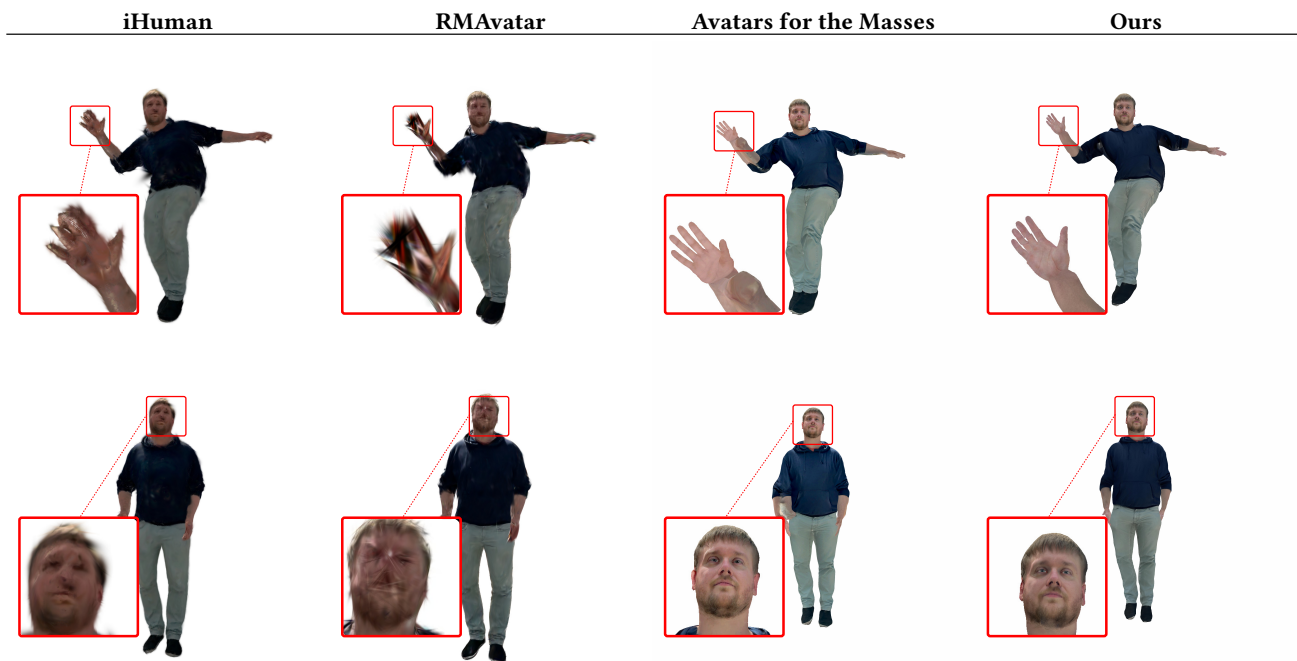


Figure 17: Subject 6 in novel pose

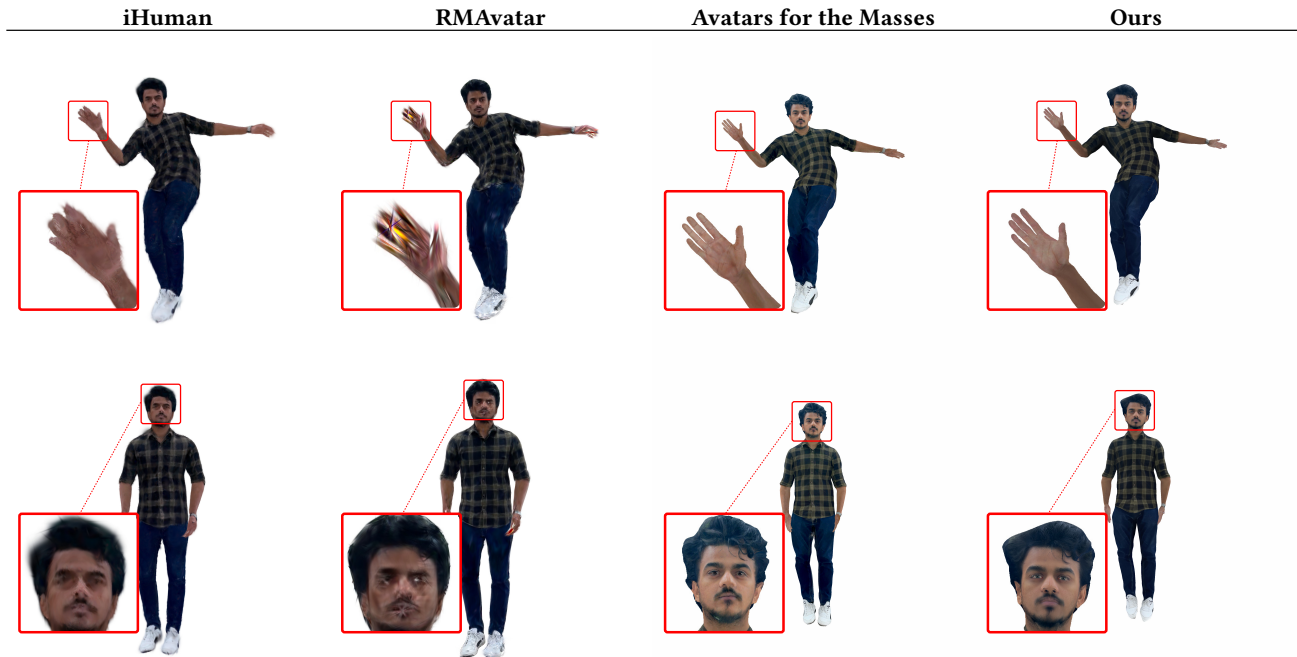


Figure 18: Subject 7 in novel pose

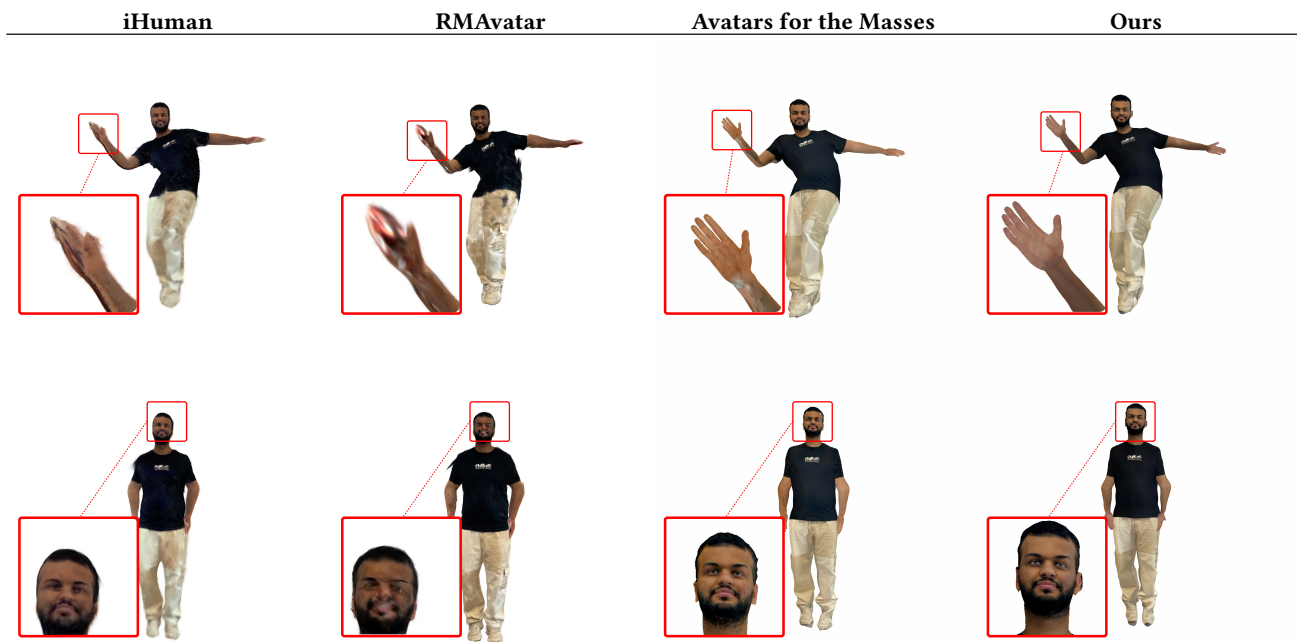


Figure 19: Subject 8 in novel pose

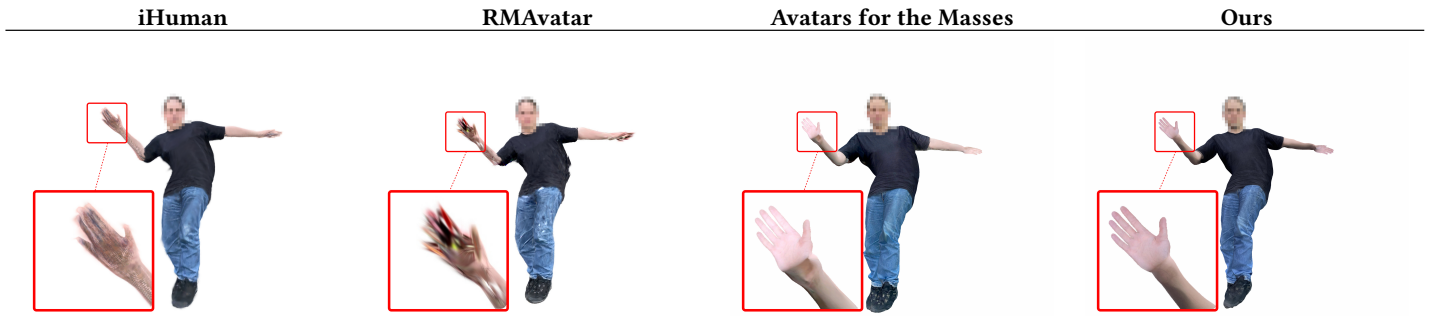


Figure 20: Subject 9 in novel pose

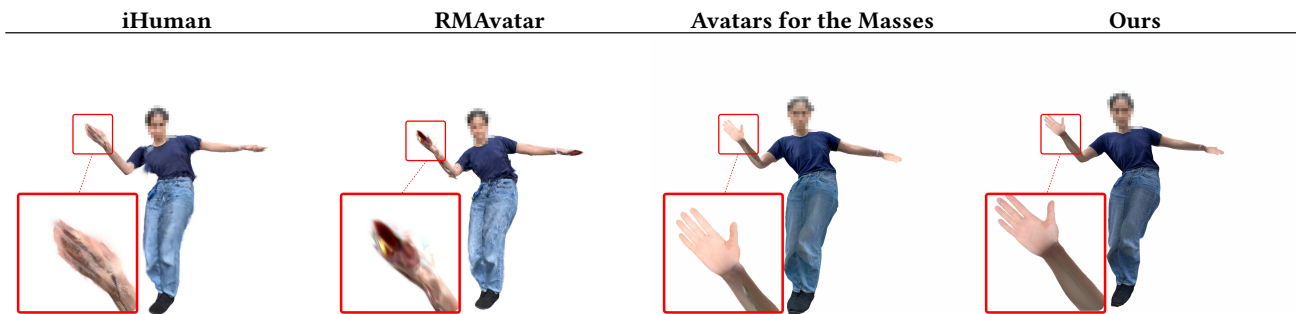


Figure 21: Subject 10 in novel pose

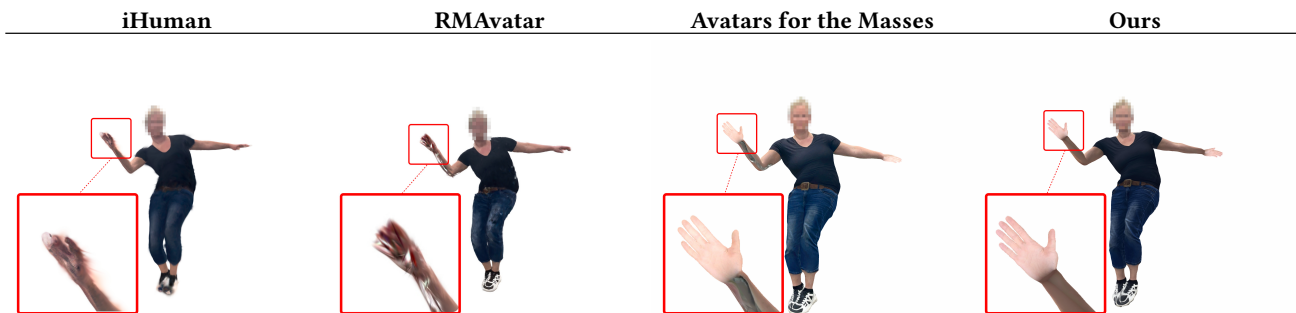


Figure 22: Subject 11 in novel pose

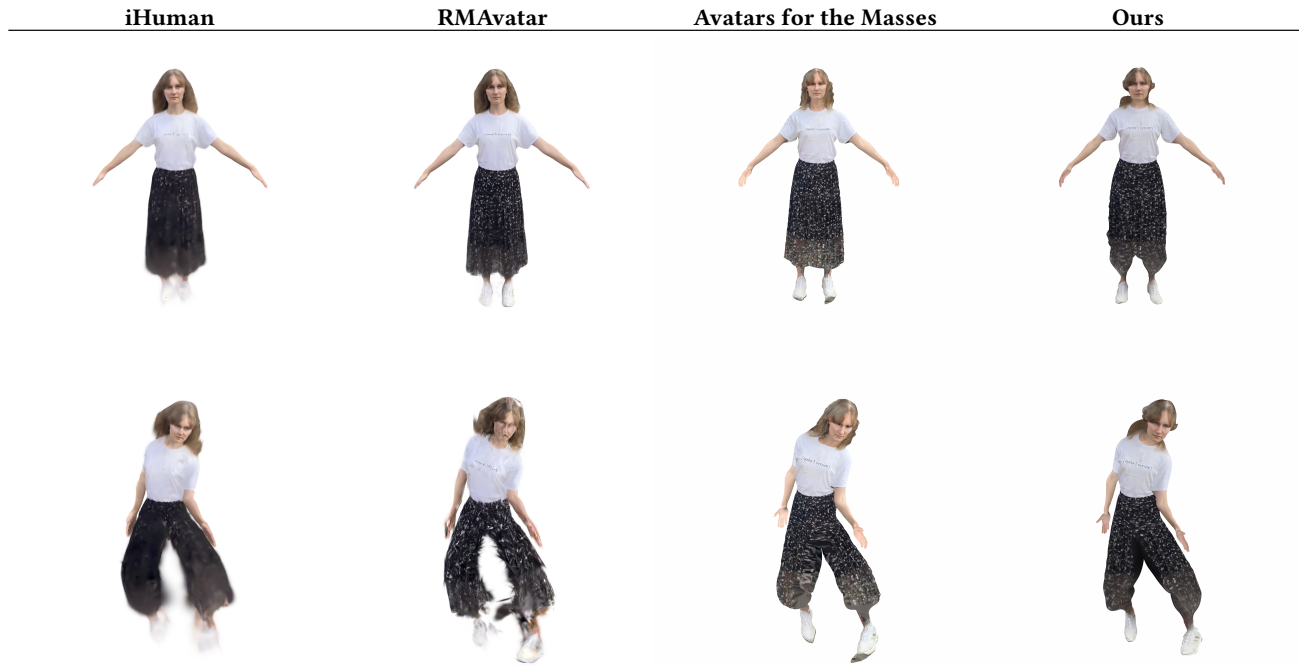


Figure 23: Comparison of female subject with loose, curly hair and skirt. First in train-pose, then in novel-pose animation.

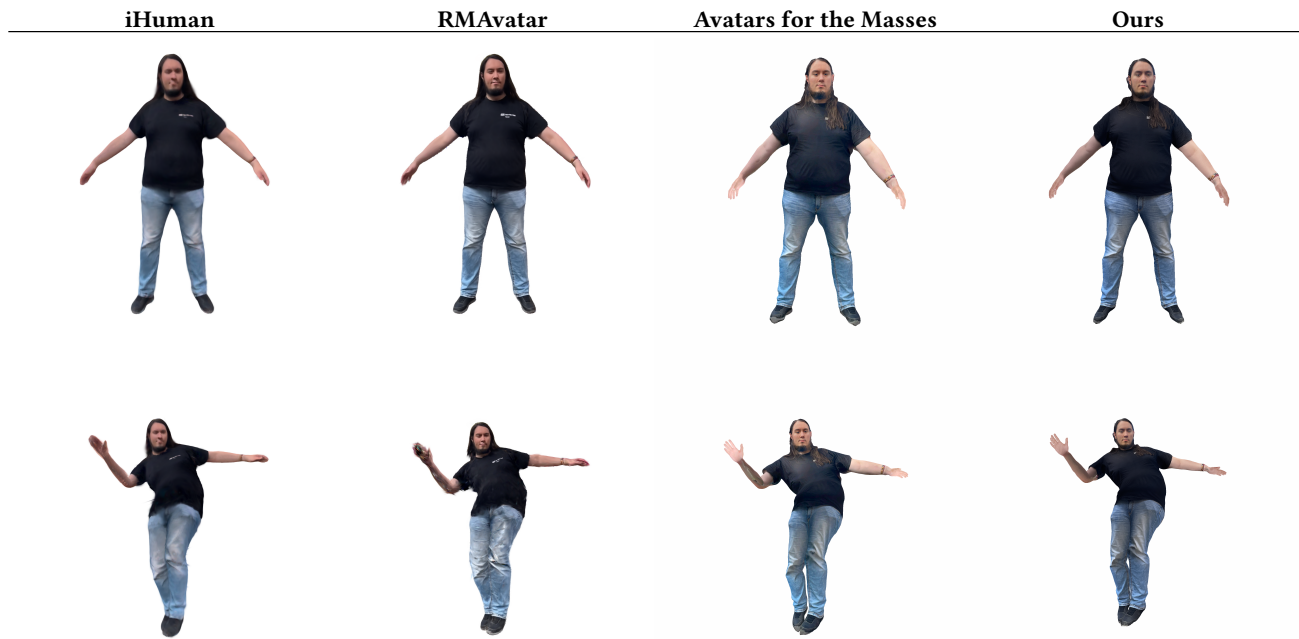


Figure 24: Comparison of male subject with long hair in training-pose and novel-poses.



Figure 25: The point cloud generated by Agisoft Metashape [1] for a female subject with loose hair. Due to windy conditions during the scan, the photogrammetry was unable to match hair features across the image sequence, resulting in only partial reconstructions.



Figure 26: Subject 1



Figure 27: Subject 2



Figure 28: Subject 3

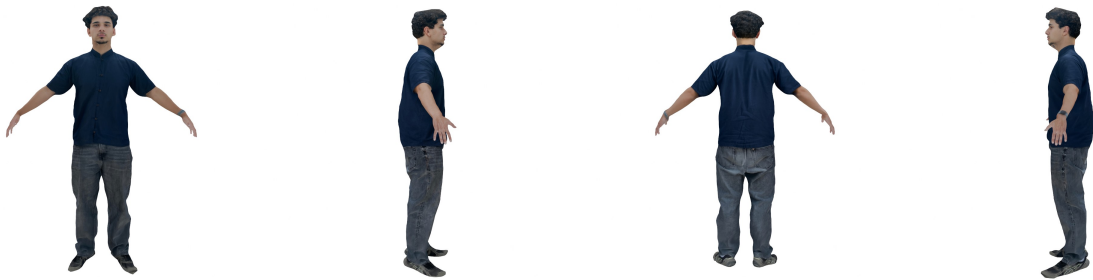


Figure 29: Subject 4



Figure 30: Subject 5



Figure 31: Subject 6



Figure 32: Subject 7

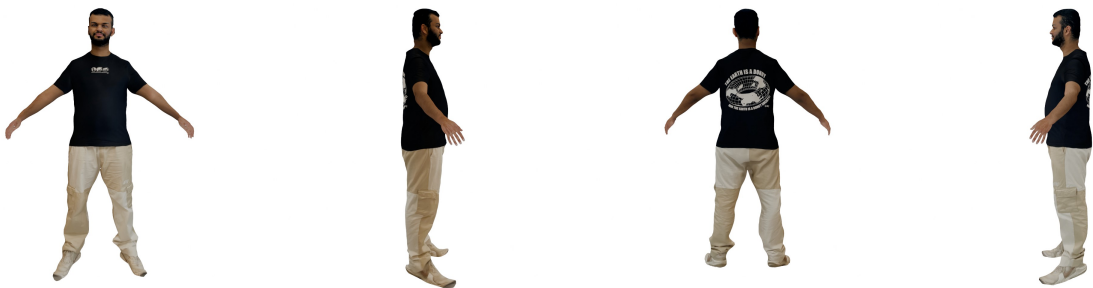


Figure 33: Subject 8



Figure 34: Subject 9

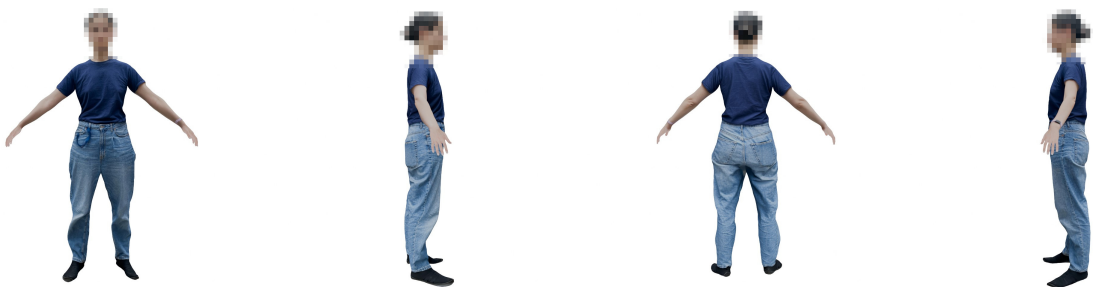


Figure 35: Subject 10



Figure 36: Subject 11



Figure 37: Subject 12



Figure 38: Subject 13



Figure 39: Subject 14



Figure 40: Subject 15



Figure 41: Subject 16



Figure 42: Additional Avatar: Female subject with loose hair and long skirt.



Figure 43: Additional Avatar: Female subject with skirt.

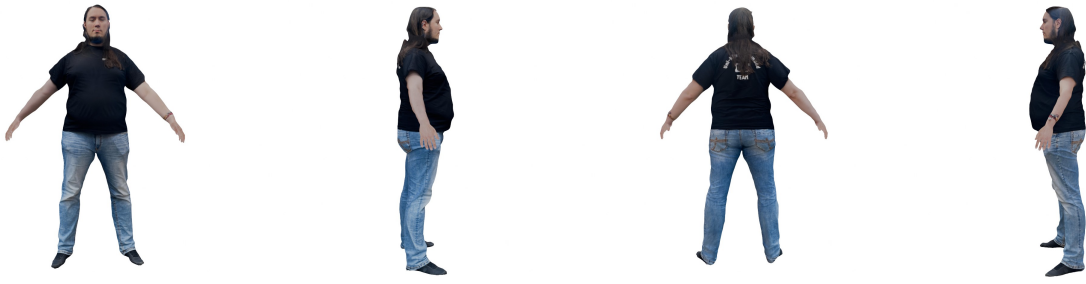


Figure 44: Additional Avatar: Male subject with long hair.