Real-time self-contact sensitive finger and full-body animation of avatars with different morphologies and proportions

Mathias DELAHAYE
mathias.delahaye@epfl.ch
EPFL
Lausanne, Vaud, Switzerland

Bruno HERBELIN
bruno.herbelin@epfl.ch
EPFL
Geneva, Geneva, Switzerland

Ronan BOULIC
ronan.boulic@epfl.ch
EPFL
Lausanne, Vaud, Switzerland

ABSTRACT

We use our hands and fingers to interact with the world and it is crucial in VR (Virtual Reality) to enable users to control their avatars with realistic finger movements. However, when the user's body and avatar proportions differ, applying the user's motion directly to the avatar can lead to self-contact mismatches. Here, we present a real-time animation retargeting method sensitive to the self-contact concern. We evaluated this approach and found it to be preferred over traditional avatar control based on the direct captured movement.

ACM Reference Format:

Mathias DELAHAYE, Bruno HERBELIN, and Ronan BOULIC. 2023. Real-time self-contact sensitive finger and full-body animation of avatars with different morphologies and proportions. In *Proceedings of M2023 ACM SIGGRAPH Conference on Motion, Interaction and Games (ACM MIG23).* ACM, New York, NY, USA, 2 pages. https://doi.org/XXXXXXX.XXXXXXX

Figure 1: Decomposition of a target point's position p_j into a surface contact point and vectors (left) that are re-applied on the target avatar (middle) as p'_j. The final result of the attraction process is illustrated on the right.

1 INTRODUCTION

Our bodies allow us to express ourselves through various poses and gestures. In VR, when users wear a HMD (Head-Mounted Display), their physical body vanishes, making it necessary to introduce an avatar for virtual interactions.

In a multi-user VR setting, failing to accurately translate a user's original movements onto their avatar can distort the meaning of their gestures, especially regarding poses involving self-contacts. Simply applying raw joint angles to the avatar may result in unrealistic animations (e.g., animating an avatar with a large belly for a thin user). Therefore, it is crucial to retarget the user's motions to fit the destination avatar.

Despite the importance of self-contact congruence, many existing approaches for retargeting user motion to avatars are either offline or primarily focused on interactions with objects, as discussed in the following section. Here, we introduce a retargeting pipeline that specifically addresses self-contact consistency, even at the finger level, and we evaluate its ability to convey motion semantics through subjective assessments.

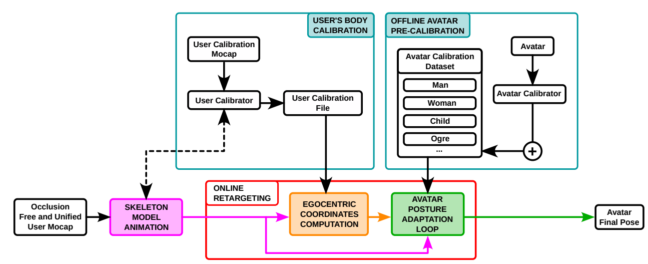

Figure 2: System overview: the upper stages on the schema represent the user and the avatar calibration. Finally, the online stage computes in real-time the instantaneous pose to be applied to the avatar.

2 RELATED WORKS

Animation retargeting is covered in the literature through a large panel of techniques [3]. In particular, in [1] or [4], the authors used an internal representation of the limb structure to adapt a limb pose onto another character easily, but the focus was put on the interaction with the environment and not for self-contacts. In [2], the authors combined the former methods to provide a body-independent retargeting animation pipeline that can handle self-contact congruency in real-time. However, those approaches do not address finger-level interactions, which are crucial in interacting with the virtual world.

ACM MIG23, Nov. 15–17, 2023, Rennes, FR

ACM ISBN 978-1-4503-XXXX-X/18/06…$15.00
https://doi.org/XXXXXXX.XXXXXXX

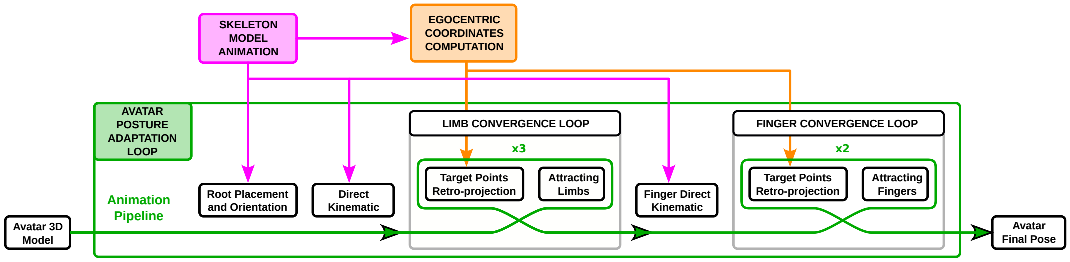

Figure 3: The posture adaptation pipeline resets the placement of the avatar and pre-orient limbs using direct forward kinematic. Limbs are then progressively attracted toward their retro-projected target points (p'_j).

3 RETARGETING ANIMATION PIPELINE

Our solution firstly relies on an initial calibration of both the user's and the virtual character's approximate body shapes, which is crucial to anticipate when self-contacts are about to occur (Figure 2). A set of 28 target points $(p_j)_{j \in \mathbb{N}}$ are attached to key skeleton's structure points (knees, elbows, hands, feet, fingertips) and are used to express a body part's location in the body's surface referential through a sum of contributions (Figure 1 left) for each frame. Those coordinates are called egocentric coordinates [2] and are then reapplied onto the avatar's body (initialized with the captured pose) to determine the avatar's body part corresponding location (Figure 1 middle). Once the avatar's target position is determined, an iterative process progressively attracts the limbs, followed by the fingers, towards their expected targeted positions (Figure 3) to produce the final avatar pose (Figure 1 right).

4 SUBJECTIVE EVALUATION

We compared two approaches, our animation method vs. direct kinematics, through a subjective evaluation using the third PV (Person Viewpoint). This evaluation consisted of showing three videos to naive observers: one showing a short video recorded source movement (involving self-contacts, ground interaction, and semantic poses) and two synthesized videos with a similar viewpoint (randomly placed and unlabeled) using each approach, and then asking the participants to evaluate both animations.

5 ANALYSIS

To verify the presence of an effect linked to the animation method factor, we performed a pair-wised comparison between both samples drawn from the retargeted approach and the direct kinematic one. Finally, a test on the direction of the difference was performed in post-hoc using the one-sided version of the t-test, and the effect sizes were measured using Cohen's D. In total, 20 participants (15 women), aged between 20 and 30 years old (average: 21.7, std: 2.43), participated in this study. Measured scores are plotted in Figure 4.

6 DISCUSSION

Prior to performing the evaluation, we observed that our method can produce instability/jitter in the animation and that some complex situations might still induce some interpenetrations when the crude mesh resolution is not high enough to approximate correctly the body surface. Furthermore, in our approach, only the vertical height of the feet is constrained, making it subject to the common footskate phenomenon. Finally, the involved setup is currently heavy and could be simplified to make it more user-friendly.

The subjective evaluation showed that, except for the ground interaction, our retargeting approach was significantly preferred over the direct kinematic animation pipeline, with an overall good appreciation for both methods. In particular, the score difference was more pronounced when the avatar's shape differed more from the performer's one, hence highlighting the importance of retargeting the user's motion. The participant feedback emphasized the role of smoothness in the animation, which was not sufficiently addressed in our approach, hence providing an opportunity for improvement. Ultimately, this user evaluation sets the stage for conducting an immersive first PV user evaluation in the future.

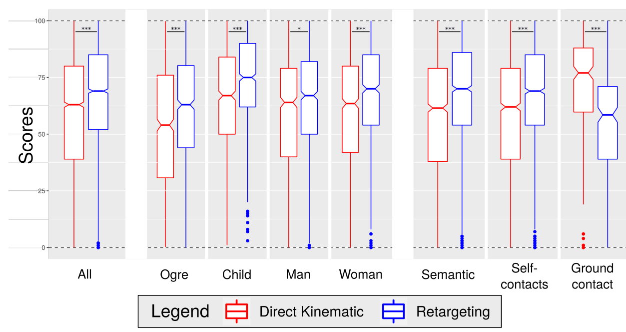

Figure 4: Normalized boxplots of the participants' evaluation scores.

REFERENCES

[1] Richard Kulpa, Franck Multon, and Bruno Arnaldi. 2005. Morphology-independent representation of motions for interactive human-like animation. In *Eurographics*.
[2] Eray Molla, Henrique Galvan Debarba, and Ronan Boulic. 2017. Egocentric Mapping of Body Surface Constraints. *IEEE Transactions on Visualization and Computer Graphics* (2017), 1–1. https://doi.org/10.1109/TVCG.2017.2708083
[3] Lucas Mourot, Ludovic Hoyet, François Le Clerc, François Schnitzler, and Pierre Hellier. 2022. A Survey on Deep Learning for Skeleton-Based Human Animation. In *Computer Graphics Forum*. Wiley Online Library, 122–157.
[4] Franck Multon, Richard Kulpa, Ludovic Hoyet, and Taku Komura. 2009. Interactive animation of virtual humans based on motion capture data. *Computer Animation and Virtual Worlds* 20, 5-6 (2009), 491–500.

Received 14 July 2023