

Supplementary Material - Social EgoMesh Estimation

Luca Scofano* Alessio Sampieri* Edoardo De Matteis*
Indro Spinelli Fabio Galasso

{scofano, sampieri}@diag.uniroma1.it {dematteis, spinelli, galasso}@di.uniroma1.it

Sapienza University of Rome, Italy

This supplemental material provides extra qualitative poses and offers a more thorough explanation of our social interaction ablation studies (Section 4.3 of our main paper).

Interpersonal distance between Wearer and Interactee

Table 3 in our main paper illustrates that the Mean Per Joint Position Error (MPJPE) is predominantly influenced by the distance separating the wearer and the interactee, indicating the error in the reconstructed pose. Figure 1 illustrates a series of poses with the persons positioned a meter apart, while Figure 2 presents a contrasting scenario where the distance between the poses is extended to more than two meters. As indicated in Table 3, the orientation error remains consistent, which is also observable in the figures. However, it is noteworthy that the MPJPE is higher when the distance increases, as evidenced by the higher error in Figure 2 given by the legs and the arms. Nevertheless, it is consistent with how a human would reasonably approach another by extending their arm. On the contrary, Figure 1 generated pose (blue) is realistic and reflects real-life human behavior. While distance reliably serves as a proxy for social interaction, excessive closeness may yield the opposite outcome, as depicted in Figure 3, where the wearer mirrors the behavior of the interactee.

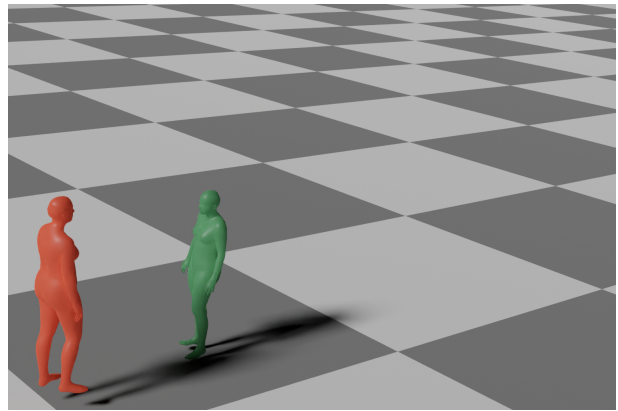
Gaze between wearer and interactee Table 4 in our primary paper demonstrates how an individual’s gaze affects all metrics collectively. This effect is also apparent qualitatively in Figures 5 and 4. In Figure 5, where two individuals are looking at each other (with a field of view smaller than 30 degrees), the reconstructed pose appears realistic. Conversely, in Figure 4, when the two are not looking at each other, the wearer’s pose mimics the interactee’s. In Figure 6, it is evident that the predicted pose (blue) closely resembles the ground truth pose (green). Nevertheless, when out of sight, the orientation error increases.

Moreover, in Figure 7 we demonstrate how the pose reconstruction of the wearer can be affected negatively by the interaction with the interactee. While our main emphasis has been on favorable effects, we now focus on an adverse scenario. The seated posture of the interactee results in the generated wearer also adopting an unnatural bent posture, resembling the other person. The orientation of the seated person also makes the actor slanted, these effects are undesirable within our context.

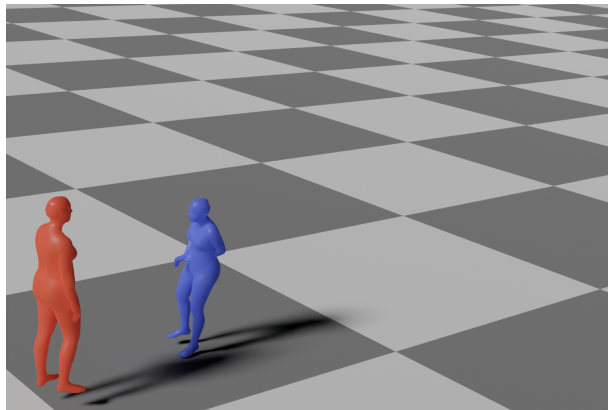
* Authors contributed equally.



(a) Single frame of the wearer's egocentric view.



(b) Ground Truth pose of the wearer (*green*) and the interactee (*red*)

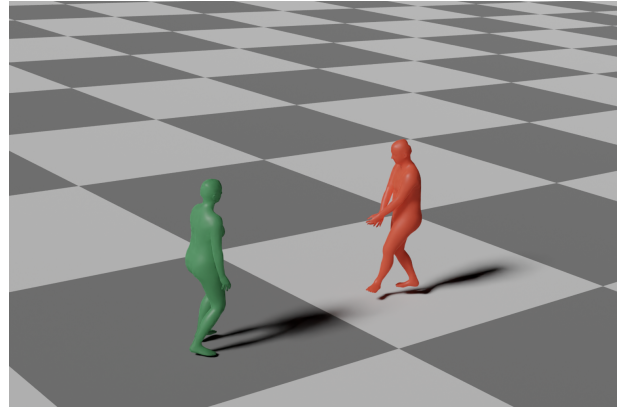


(c) Predicted pose of the wearer (*blue*) and Ground Truth of the interactee(*red*).

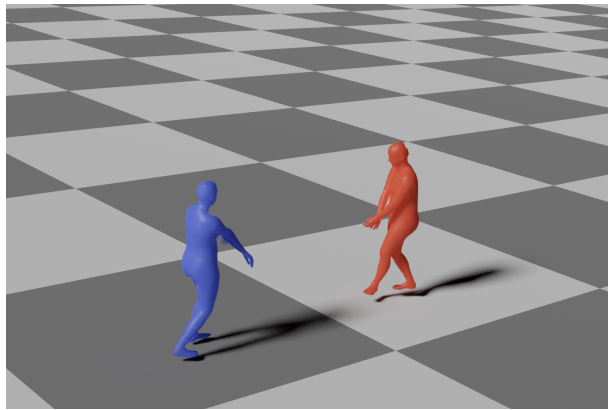
Figure 1. SEE-ME in a scenario with two standing people facing each other.



(a) Single frame of the wearer's egocentric view.



(b) Ground Truth pose (*green*) of the wearer and the interactee (*red*).

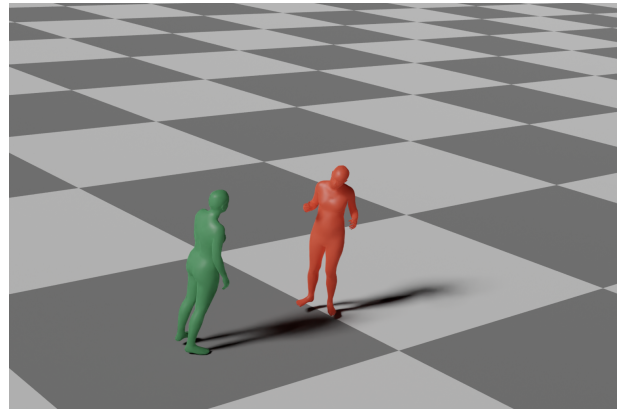


(c) Predicted pose of the wearer (*blue*) and Ground Truth of the interactee(*red*).

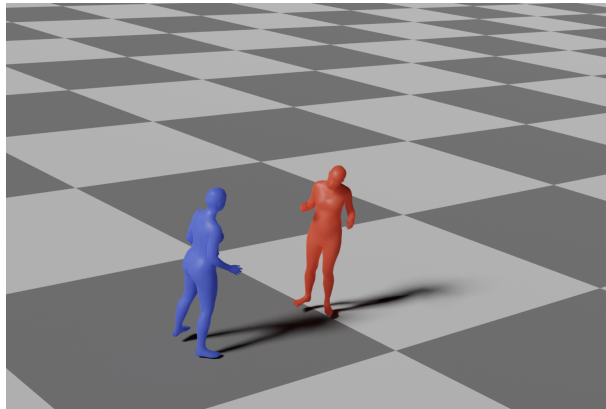
Figure 2. The interactee (red) is coming towards the wearer.



(a) Single frame of the wearer's egocentric view.



(b) Ground Truth pose (*green*) of the wearer and the interactee (*red*).

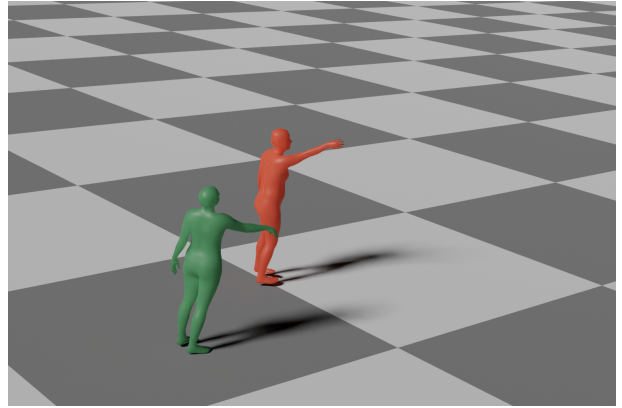


(c) Predicted pose of the wearer (*blue*) and Ground Truth of the interactee (*red*).

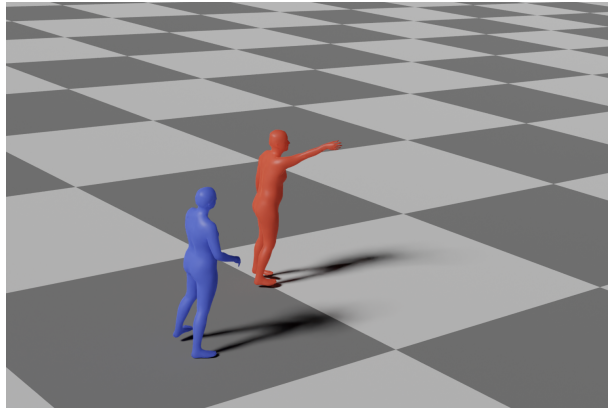
Figure 3. Closer interaction between two agents (less than a meter), we can notice how SEE-ME's wearer (*blue*) mirrors the interactee's hand gestures.



(a) Single frame of the wearer's egocentric view.



(b) Ground Truth pose (*green*) of the wearer and the interactee (*red*).

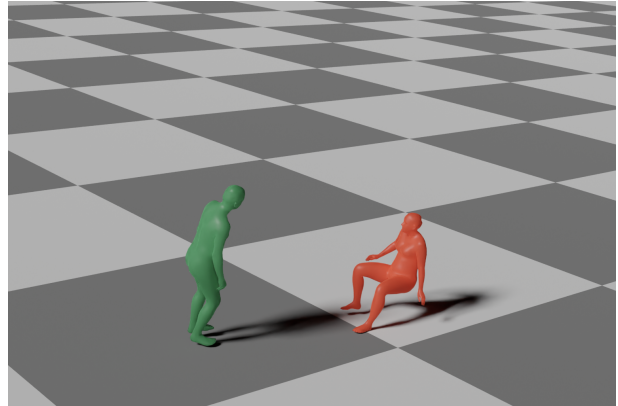


(c) Predicted pose of the wearer (*blue*) and Ground Truth of the interactee (*red*).

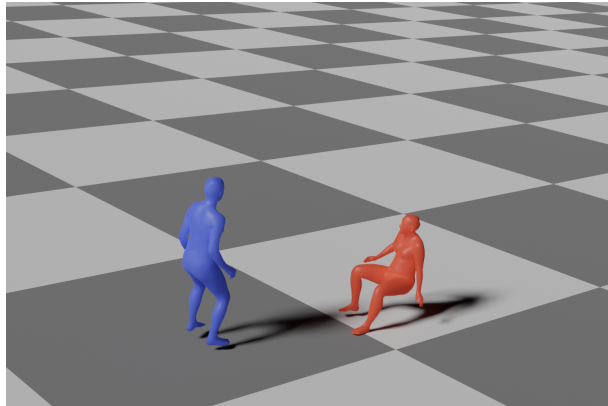
Figure 4. Scenario in which the interactee (red) and the wearer are not facing each other.



(a) Single frame of the wearer's egocentric view.



(b) Ground Truth pose (*green*) of the wearer and the interactee (*red*).

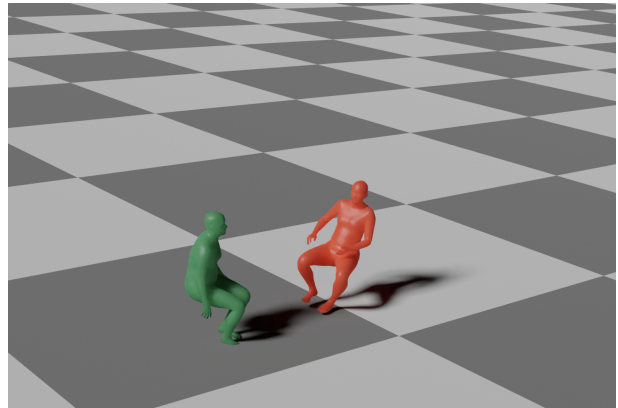


(c) Predicted pose of the wearer (*blue*) and Ground Truth of the interactee (*red*).

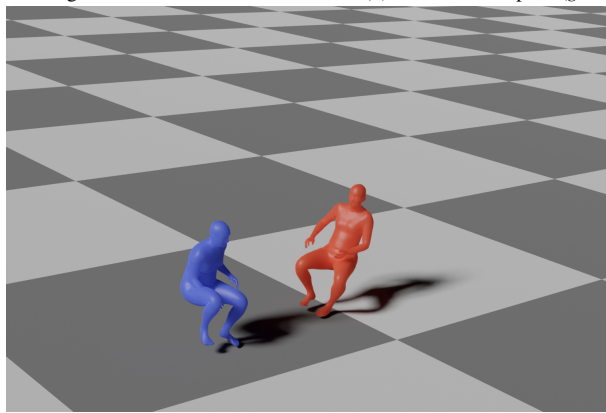
Figure 5. The interactee (*red*) and the wearer are not looking at each other.



(a) Single frame of the wearer's egocentric view.

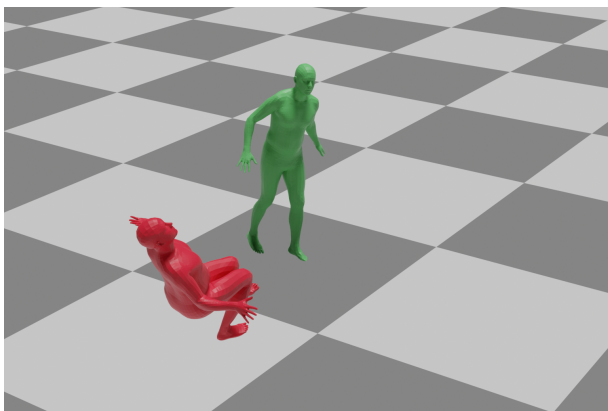


(b) Ground Truth pose (*green*) of the wearer and the interactee (*red*).

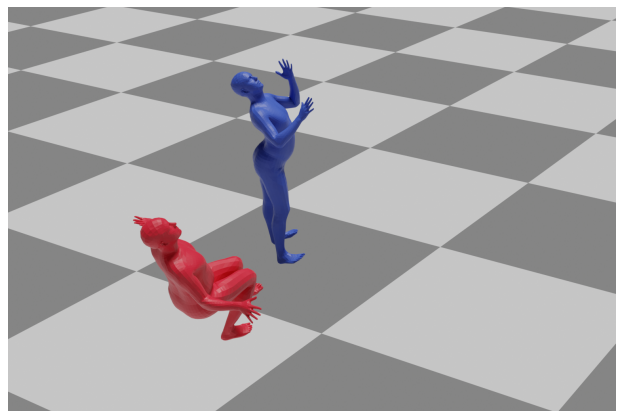


(c) Predicted pose of the wearer (*blue*) and Ground Truth of the interactee (*red*).

Figure 6. The interactee (*red*) is sitting. SEE-ME predicts the wearer by considering the interactee's line of sight.



(a) Seated interactee (*red*) and standing ground truth wearer (*green*).



(b) Seated interactee (*red*) and standing wearer predicted by SEE-ME (*blue*), showing an unnatural posture.

Figure 7. The influence with the interactee can have a negative influence in some scenarios. The seated position of the interactee can condition the wearer into having a contrived posture.