Muscles in Time - Rebuttal

Dear reviewers and chairs, we are thankful and appreciate the generally positive feedback on our work. We address the key questions and suggestions from **Bfvp** , **pwPd** , **6vHW** , and **kyb6** in this global rebuttal, detailed responses will further be provided directly to the respective reviewers.

Improved Documentation **Bfvp** rightfully pointed out lacking documentation of the MusclesInTime repository. We added a detailed README and Jupyter notebooks for visualizing data samples including motion and intend to migrate this information to a readthedocs.io space.

SMPL-H We acknowledge **Bfvp** 's concerns about SMPL-H motion quality, particularly for hands and feet. The AMASS dataset largely lacks hand articulation, so our muscle simulation omits related muscles, as shown in Figure 2. Unlike [9], which used VIBE for pose estimation, MinT's motions are based on high quality motion capture. This difference is evident in the animations linked below as well as in visualizations provided by the BABEL dataset: https://human-movement.is.tue.mpg.de/explore/. Detailed marker visualization and used mesh vertex indices will shortly be added to the repository.

Simulation quality and limitations All reviewers acknowledged our efforts to list limitations which are unavoidable when using simulated data. **Bfvp** highlighted lines 119-122 where we mention that we follow the community in using pre-validated models, avoiding major changes. In line 128 we mention that we make use of Opensim's scaling mechanism, we means that we do scale the mass properties and body segment length to match an individuals anatomical properties. We will change the paper text to clarify this. We will also highlight the limitation of not adapting other model parameters such as muscle forces, raising awareness in future users of our dataset. The usage of pre-validated models is an important aspect of our simulation pipeline, specifically to address the lack of EMG ground truth recordings for verification in AMASS, significant changes to muscle geometry and properties would require an extensive per-individual re-validation using such recordings. **kyb6** inquires about quantitative evaluation of data quality. The optimization loss used in OpenSimAD which we use for muscle activation includes terms for position tracking, velocity, and acceleration errors as well as terms for activation minimization and reserve actuator minimization. While we do not make use of individual thresholds for these properties, we discard any results for which the optimization problem does not converge within the tolerance recommended by the the authors of OpenSimAD.

Additional results and qualitative examples **6vHW** and **kyb6** suggested evaluation on further real world datasets. While Camargo *et al*. [4] and MiA [9] from Table 1 were the only viable options, Camargo *et al*. do not provide upper body motion capture points. MiA provides only eight EMG measurement points, four of which align with our model. We tried mapping the hamstring measurement to the *biceps femoris long head* and the quadriceps to the *rectus femoris* muscle. Our main problem is that MiA motions are stationary, hovering above the ground and not aligning well with locomotion induced by the feet, making our model generalize badly due to its training on very precise AMASS motions. Fine-tuning on MiA was necessary, as presented in Table A5 in the supplementary. We plan to develop advanced domain adaptation methods to address these issues in future work. **kyb6** suggested additional baseline experiments. The T2M-GPT architecture was chosen for its effectiveness in generating useful intermediate representations for cross-modality translation tasks in human motion, making it a stronger baseline than arbitrary architectures. To address the request of **kyb6** , we are currently preparing additional baseline results which will be provided during the discussion period. To address **kyb6** ' suggestion for showing predictions on MinT as well as for the model in Table A5 combined with the corresponding motion, we prepared two sets of videos containing animated movements with muscle activations and their respective predictions. These videos are available at https://s.kit.edu/mint-vis for MinT and at https://s.kit.edu/mia-vis for MiA. The jupyter notebooks used for generation can be found in the MinT repository.

Submitted to the 38th Conference on Neural Information Processing Systems (NeurIPS 2024) Track on Datasets and Benchmarks. Do not distribute.