





# Appendix of Physics-Based Motion Tracking of Contact-Rich Interacting Characters

Xiaotang Zhang<sup>1</sup>, Ziyi Chang<sup>1</sup>, Qianhui Men<sup>2</sup>, and Hubert P. H. Shum<sup>†1</sup>

<sup>1</sup>Durham University, United Kingdom  
{xiaotang.zhang, ziyi.chang, hubert.shum}@durham.ac.uk

<sup>2</sup>University of Bristol, United Kingdom  
qianhui.men@bristol.ac.uk

## 1. Implementation Details

### 1.1. Hyperparameters

The hyperparameters for training are available in Table 1.

Parameter	Value
Number of Environments	20,000
Rollout Memory Size	16
Learning Epochs Per Update	8
Mini-batch Size	2
Learning Rate	$5 \times 10^{-5}$
GAE Discount Factor	0.99
GAE Lambda Coefficient	0.95
Policy Loss Ratio Clipping Coefficient	0.2

**Table 1:** Hyperparameters for training.

### 1.2. Reward Function

The state rewards are calculated using exponential functions to encourage close tracking of the reference motion:

$$r_t^{\text{track}} = w^p r_t^p + w^q r_t^q + w^v r_t^v + w^a r_t^a + w^{jp} r_t^{jp} + w^{jv} r_t^{jv}. \quad (1)$$

For position and velocity features,  $r^{\{\cdot\}}$  is calculated as mean squared error and normalized as:

$$\exp\left(-\frac{\lambda^{\{\cdot\}}}{B} \sum_{b=1}^B \|s_t^{\{\cdot\}} - s_t^{\{\cdot\}}\|^2\right), \quad (2)$$

where  $B$  denotes the amount of rigid body of the simulated character,  $s_t^{\{\cdot\}}$  is the ground truth state,  $\lambda^{\{\cdot\}}$  is a scaling factor to control the decay rate of the exponential reward terms.

For quaternion feature,  $r^q$  is calculated as rotational difference (denoted as  $\ominus$ ) and normalized as:

$$\exp\left(-\frac{\lambda^q}{B} \sum_{b=1}^B (s_t^q \ominus s_t^q)\right). \quad (3)$$

The energy reward is calculated as:

$$r_t^{\text{energy}} = \exp\left(-\frac{\lambda^{\text{energy}}}{J} \sum_{j=1}^J |s_{t,j}^\tau s_{t,j}^{jv}|\right), \quad (4)$$

where  $J$  denotes the degrees of freedom of the simulated character,  $s^\tau$  is the applied joint torque and  $s^{jv}$  is the joint velocity. The energy penalty reward trade off the tracking performance as it suppresses the joint torque at each degree of freedom.

Detailed values for weights and scaling factors can be found in Table 2 and 3.

Parameter	Value
$w^p$ Rigid Body Position Reward Weight	0.4
$w^q$ Rigid Body Quaternion Reward Weight	0.2
$w^v$ Rigid Body Linear Velocity Reward Weight	0.1
$w^a$ Rigid Body Angular Velocity Reward Weight	0.1
$w^{jp}$ Joint Position Reward Weight	0.1
$w^{jv}$ Joint Velocity Reward Weight	0.1

**Table 2:** Reward weights for different state features.

Parameter	Value
$\lambda^p$ Rigid Body Position Reward Scaling Factor	-162
$\lambda^q$ Rigid Body Quaternion Reward Scaling Factor	-201
$\lambda^v$ Rigid Body Linear Velocity Reward Scaling Factor	-3.17
$\lambda^a$ Rigid Body Angular Velocity Reward Scaling Factor	-0.116
$\lambda^{jp}$ Joint Position Reward Scaling Factor	-2
$\lambda^{jv}$ Joint Rotation Reward Scaling Factor	-0.02
$\lambda^{\text{energy}}$ Energy Reward Scaling Factor	0.03

**Table 3:** Empirically tuned scaling factors for the exponential reward terms.