

AMP: Adversarial Motion Priors for Stylized Physics-Based Character Control

摘要

合成真实自然的物理角色动画一直是计算机动画领域的一大挑战。现有的基于运动跟踪与奖励函数工程的方法可以生成高质量且逼真的物理动画，然而它们需要人为精细地设计奖励函数，并且需要规划器来选择作为运动跟踪的片段。所复现方法提出利用基于生成式对抗模仿学习的方法，避免了对模仿奖励函数的手工设计，并且可以很好的与任务奖励函数很好地结合而无需复杂的规划器。在各种复杂的模拟角色和一系列控制任务中，展现出了所复现方法的有效性与生成逼真动画的能力。本复现的改进重点在于提出使用模仿奖励与任务奖励相乘的设计方法，避免了单一奖励函数主导使得生成结果偏向于模仿动画或完成任务上；此外设计了基于物理动画数据集与动力学数据集相互转换的程序。

关键词：角色动画；强化学习；对抗模仿学习

1 引言

计算机动画旨在借助计算机辅助创建生动、自然的角色动画，广泛应用于电影、游戏等行业。近年来，随着深度学习的兴起，数据驱动的方法在计算机动画领域快速发展。其中，基于动力学的动画（kinematics-based motion）[1–7]通过在每一时刻描述角色当前的状态（如关节位置、关节旋转角）来表示动画，然而由于缺乏真实的物理环境约束，该方法会生成脚步滑动（foot sliding）、自我穿透（self-penetration）等不自然且不符合物理规律的行为；基于物理仿真的动画通过在物理环境中对角色进行仿真，在每一时刻通过控制器来控制角色每个躯干的力矩或比例-微分控制（Proportional-Derivative Target, PD Target），从而使其在物理环境的约束下完成相应的任务，例如模仿人类的行为[8–11]。然而，基于物理仿真的方法通常受限于角色高自由度的复杂性，使得学习一个鲁棒的控制器是困难的。为此，基于模仿学习的强化学习方法被提出用于物理动画生成上。其中的 AMP 算法[12]通过引入对抗模仿学习（Generative Adversarial Imitation Learning, GAIL）避免了复杂的奖励函数工程（Reward Engineering），并且可以生成真实自然的物理角色动画。

本文的复现工作为 2021 年发表于 SIGGRAPH 会议的文章：AMP: Adversarial Motion Priors for Stylized Physics-Based Character Control。在接下来的章节中，首先会先回顾基于模仿学习的强化学习在物理角色动画生成领域的相关工作，然后阐述所复现的 AMP 算法的具体工作，其次介绍本文的复现内容，最后给出实验结果并作出简单的总结。

2 相关工作

基于模仿学习的强化学习方法在物理动画生成上有着广泛的应用，根据算法类型的不同可以大致分类为如下两种方法：基于模型的方法（model-based methods）和无模型的方法（model-free methods）；其中的无模型方法又可以类为基于奖励函数工程法与生成式对抗强化学习法。下面就这些算法进行简单的介绍。

2.1 奖励函数工程法

基于奖励函数工程的物理角色动画生成方法 [9, 13–18] 通过认为设计一个奖励函数，使模仿的物理角色与模仿对象越接近时得到的奖励越大、而模仿约不显示甚至脱离模仿目标时奖励减小。这种方法训练收敛速度快、生成的物理动画质量高，然而它需要人为地设计物理角色的每个关节与模仿角色之间的相似程度，使得该方法需要复杂的超参数设计并且通用性差；此外该类方法只能严格地模仿所给的目标，而不能脱离目标生成新的动画；最后该方法需要设计复杂的计划器（Planner）来给定模仿的目标，设计复杂。

2.2 生成式对抗强化学习法

基于生成式对抗强化学习的物理角色动画生成方法 [11, 12, 19–21] 通过引入对抗学习的概念，通过判别器给出的评分作为奖励函数，避免了复杂的奖励函数设计，从而具有很大的通用性。此外，该方法也可以很容易地将任务奖励与模仿奖励结合，使角色在模仿人类动作的同时也能完成相应的任务。此类方法目前是基于物理模仿学习这一领域的研究热点，因此本文复现最早提出此类方法的 AMP 算法，旨在帮助今后的学习研究。

2.3 基于模型的方法

基于模型的物理角色动画生成方法 [22–24] 通过对环境的建模或使用世界模型（world model）的近似，并根据其状态转移概率来直接优化相应的策略函数，而避免了随机搜索与策略梯度下降这种收敛慢的方法。然而这种方法对环境有很强的假设，使得其在更复杂的转移条件下（例如环境中存在复杂的物体交互）效果较差。

3 本文方法

3.1 本文方法概述

如图 1 所示为本文复现方法的模型。给定一个动画数据集，方法要求在模仿该数据集的人类行为的同时，完成另一指定的任务。对于当前环境的观察 S_t ，策略函数 Policy 会输出一个相应的动作 a_t ，并且得到相应的奖励 r_t ，然后使用策略梯度和 GAE 的方法来训练策略函数：

$$J(\pi) = \mathbb{E}_{p(g)} \mathbb{E}_{p(\tau|\pi, g)} \left(\sum_{t=0}^{T-1} \gamma^t r_t \right) \quad (1)$$

其中， $p(\tau|\pi, g) = p(S_0) \prod_{t=0}^{T-1} p(s_{t+1}|s_t, a_t) \pi(a_t|s_t, g)$ 表示了轨迹 $\tau = (s_t, a_t, r_t)_{t=0}^{T-1}, s_T$ 在策略 π

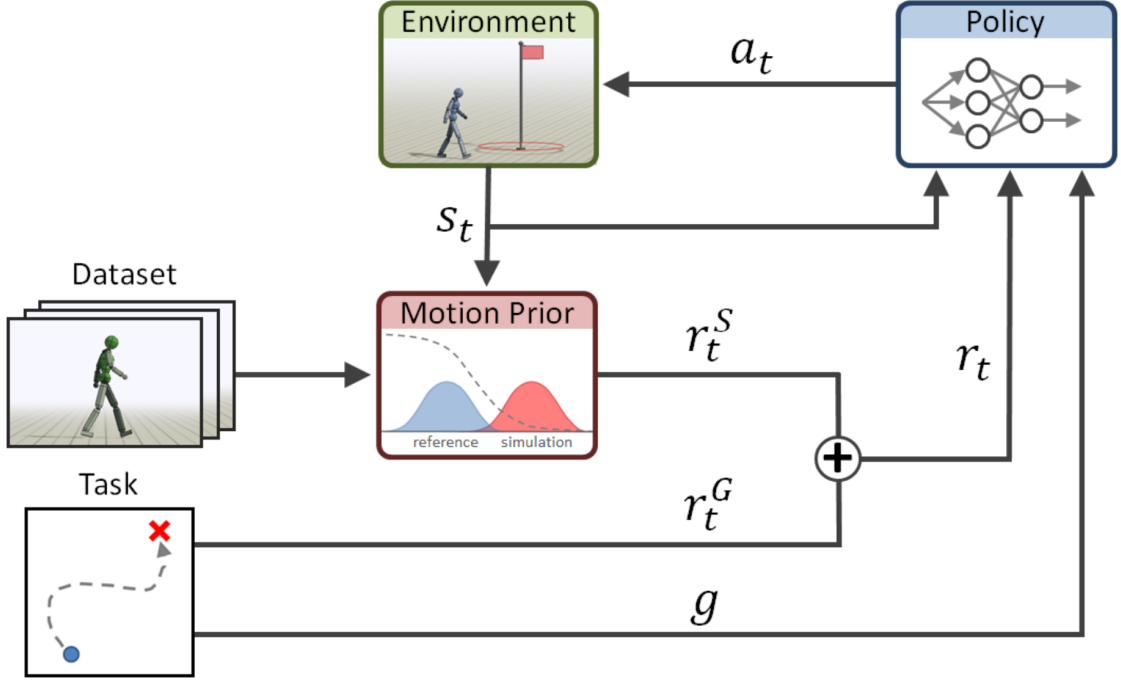


图 1. 方法示意图

和目标 g 下的似然值, $p(s_0)$ 是初始的状态分布, $p(s_{t+1}|s_t, a_t)$ 是环境的动态转移方程。 T 表示轨迹的时间视野, $\gamma \in [0, 1)$ 是折扣因子。在奖励函数中, 模仿的奖励函数为 r_t^S , 由判别器 Discriminator 给出; 任务的奖励函数为 r_t^G , 总奖励函数为

$$r_t = r_t^S + r_t^G \quad (2)$$

对抗模仿学习的奖励函数根据判别器的输出值设置如下:

$$r_t^S = -\log(1 - D(s_t, a_t)) \quad (3)$$

而判别器的优化目标为:

$$\arg_D \min -\mathbb{E}_{d^M(s,a)}[\log(D(s,a))] - \mathbb{E}_{d^\pi(s,a)}[\log(1 - D(s,a))] \quad (4)$$

其中, $d^M(s,a)$ 和 $d^\pi(s,a)$ 分别表示数据集和测量的状态-动作分布。

4 复现细节

4.1 与已有开源代码对比

这篇文章已有开源代码, 由于文章主要贡献是设计了简洁的模仿奖励函数使其与任务奖励函数结合, 本工作的改进点在于将这个奖励由加法奖励改为乘法奖励, 以避免一个奖励占据主导从而使得生成结果偏离模仿对象或目标任务的现象, 将(2)改进如下:

$$r_t = r_t^S * r_t^G \quad (5)$$

此外，本文还编写了将物理角色动画与动力学动画互相转换的代码，旨在方便后续的科研项目中利用动力学动画作为训练数据集。

4.2 实验环境搭建

使用 Isaac Gym 作为物理仿真平台、rl-games 环境作为强化学习平台、Linux 软件作为操作系统。并在 RTX 4090 GPU 硬件上训练上述网络，大约需要 30 分钟的时间收敛。

4.3 复现结果展示

如下图所示为本文复现的结果展示，更多细节请参阅视频附录材料。

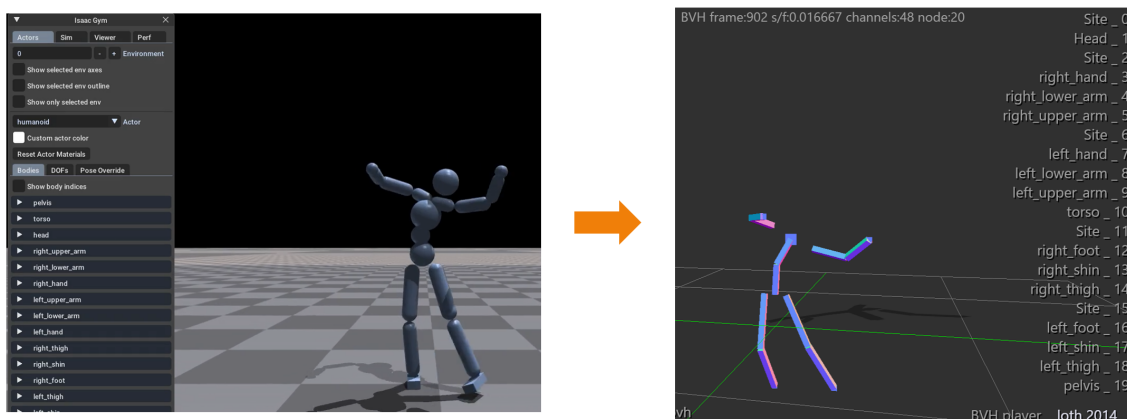


图 2. 基于物理动画与动力学方法相互转换

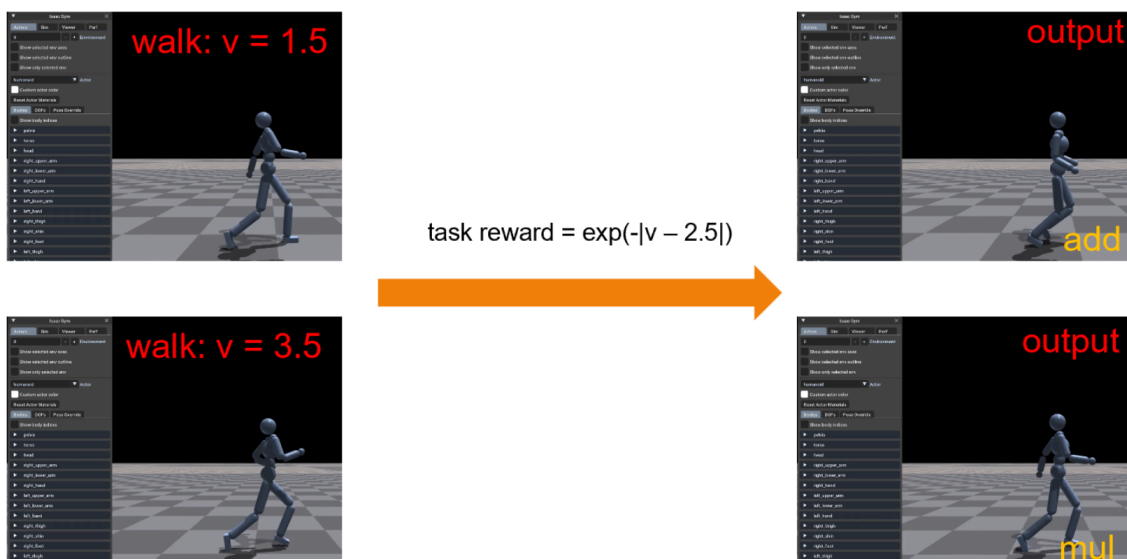


图 3. 改进后的总奖励函数生成的动画更自然

角色类型	数据集名称	动画长度 (秒)	动画片段数量	主题数量
Humanoid	Cartwheel	13.6	3	1
	Jump	28.6	10	4
	Locomotion	434.1	56	8
	Run	204.4	47	3
	Run+Leap+Roll	22.1	10	7
	Stealthy	136.5	3	1
	Walk	229.6	9	5
	Walk+Punch	247.8	15	9
T-Rex	Zombie	18.3	1	1
	Locomotion	10.5	5	1

表 1. 所用动画数据集

5 实验结果

5.1 数据集

如表1所示为所复现方法使用的数据集规模（该方法采用的数据集未开源），展示了不同角色类型的动画数据集，包括 Humanoid 和 T-Rex。每个角色类型有特定的动画集，涵盖了不同长度、片段数量和主题数量的动画，提供了关于各个动画特征的详细信息。

5.2 结果展示

在上述数据集上进行训练，并渲染展示结果如图4和图5所示。

6 总结与展望

本文复现了”AMP: Adversarial Motion Priors for Stylized Physics-Based Character Control”的研究工作。该研究的主要目标是解决计算机动画领域中生成真实自然的基于物理角色动画的挑战。通过引入基于生成式对抗模仿学习的 AMP 算法，避免了对复杂奖励函数的手工设计，并成功将模仿奖励与任务奖励相结合。在复现工作中，本文对总奖励函数进行了改进，采用乘法代替加法以避免奖励函数主导的问题。此外，还编写了动画数据集与动力学数据集相互转换的代码，为后续研究提供了方便。

参考文献

- [1] Guy Tevet, Sigal Raab, Brian Gordon, Yonatan Shafir, Daniel Cohen-Or, and Amit H Bermano. Human motion diffusion model. *arXiv preprint arXiv:2209.14916*, 2022.
- [2] Sebastian Starke, Ian Mason, and Taku Komura. Deepphase: Periodic autoencoders for learning motion phase manifolds. *ACM Transactions on Graphics (TOG)*, 41(4):1–13, 2022.

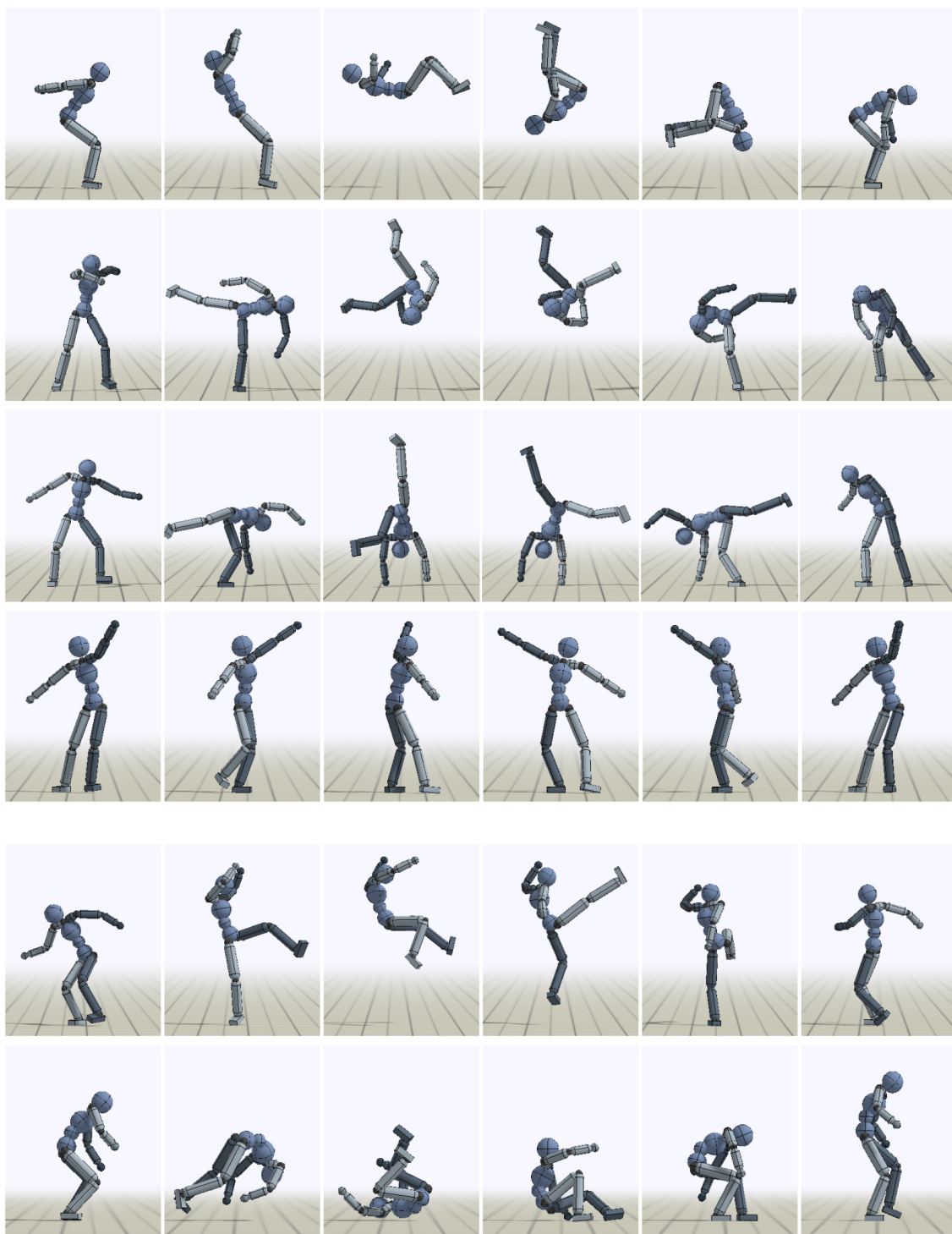


图 4. 在 Humanoid 模型上的结果展示

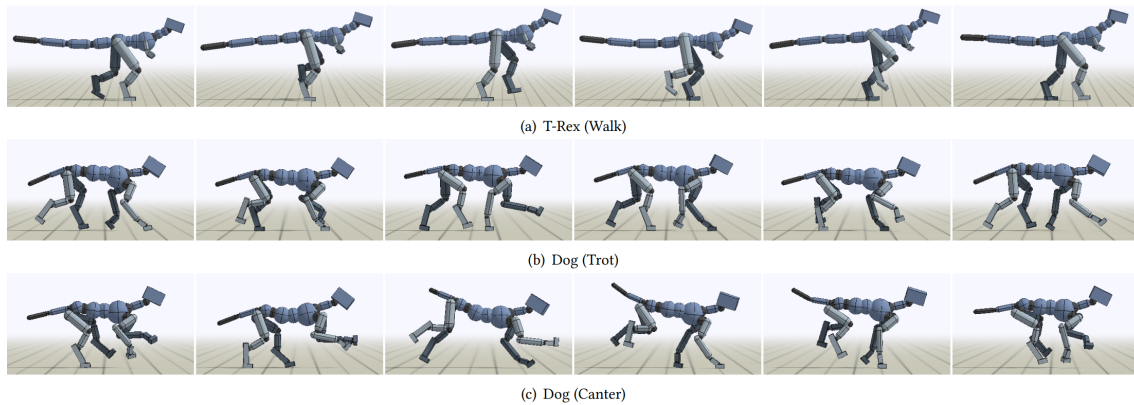


图 5. 在 T-Rex 模型上的结果展示

- [3] Jia Qin, Youyi Zheng, and Kun Zhou. Motion in-betweening via two-stage transformers. *ACM Transactions on Graphics (TOG)*, 41(6):1–16, 2022.
- [4] Zhongfei Qing, Zhongang Cai, Zhitao Yang, and Lei Yang. Story-to-motion: Synthesizing infinite and controllable character animation from long text. In *SIGGRAPH Asia 2023 Technical Communications*, pages 1–4. 2023.
- [5] Jiaman Li, Jiajun Wu, and C Karen Liu. Object motion guided human motion synthesis. *ACM Transactions on Graphics (TOG)*, 42(6):1–11, 2023.
- [6] Yifeng Jiang, Jungdam Won, Yuting Ye, and C Karen Liu. Drop: Dynamics responses from human motion prior and projective dynamics. In *SIGGRAPH Asia 2023 Conference Papers*, pages 1–11, 2023.
- [7] Hongyu Tao, Shuaiying Hou, Changqing Zou, Hujun Bao, and Weiwei Xu. Neural motion graph. In *SIGGRAPH Asia 2023 Conference Papers*, pages 1–11, 2023.
- [8] Qingxu Zhu, He Zhang, Mengting Lan, and Lei Han. Neural categorical priors for physics-based character control. *ACM Transactions on Graphics (TOG)*, 42(6):1–16, 2023.
- [9] Daniele Reda, Jungdam Won, Yuting Ye, Michiel van de Panne, and Alexander Winkler. Physics-based motion retargeting from sparse inputs. *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, 6(3):1–19, 2023.
- [10] Zhiyang Dou, Xuelin Chen, Qingnan Fan, Taku Komura, and Wenping Wang. C · ase: Learning conditional adversarial skill embeddings for physics-based characters. In *SIGGRAPH Asia 2023 Conference Papers*, pages 1–11, 2023.
- [11] Mohamed Hassan, Yunrong Guo, Tingwu Wang, Michael Black, Sanja Fidler, and Xue Bin Peng. Synthesizing physical character-scene interactions. *arXiv preprint arXiv:2302.00883*, 2023.

- [12] Eric Vollenweider, Marko Bjelonic, Victor Klemm, Nikita Rudin, Joonho Lee, and Marco Hutter. Advanced skills through multiple adversarial motion priors in reinforcement learning. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5120–5126. IEEE, 2023.
- [13] Xue Bin Peng, Pieter Abbeel, Sergey Levine, and Michiel Van de Panne. Deepmimic: Example-guided deep reinforcement learning of physics-based character skills. *ACM Transactions On Graphics (TOG)*, 37(4):1–14, 2018.
- [14] Xue Bin Peng, Angjoo Kanazawa, Jitendra Malik, Pieter Abbeel, and Sergey Levine. Sfv: Reinforcement learning of physical skills from videos. *ACM Transactions On Graphics (TOG)*, 37(6):1–14, 2018.
- [15] Kevin Bergamin, Simon Clavet, Daniel Holden, and James Richard Forbes. Drecon: data-driven responsive control of physics-based characters. *ACM Transactions On Graphics (TOG)*, 38(6):1–11, 2019.
- [16] Nazmul Karim, Mamshad Nayeem Rizve, Nazanin Rahnavard, Ajmal Mian, and Mubarak Shah. Unicon: Combating label noise through uniform selection and contrastive learning. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 9676–9686, 2022.
- [17] Alexander Winkler, Jungdam Won, and Yuting Ye. Questsim: Human motion tracking from sparse sensors with simulated avatars. In *SIGGRAPH Asia 2022 Conference Papers*, pages 1–8, 2022.
- [18] Zhengyi Luo, Jinkun Cao, Kris Kitani, Weipeng Xu, et al. Perpetual humanoid control for real-time simulated avatars. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 10895–10904, 2023.
- [19] Xue Bin Peng, Yunrong Guo, Lina Halper, Sergey Levine, and Sanja Fidler. Ase: Large-scale reusable adversarial skill embeddings for physically simulated characters. *ACM Transactions On Graphics (TOG)*, 41(4):1–17, 2022.
- [20] Jordan Juravsky, Yunrong Guo, Sanja Fidler, and Xue Bin Peng. Padl: Language-directed physics-based character control. In *SIGGRAPH Asia 2022 Conference Papers*, pages 1–9, 2022.
- [21] Chen Tessler, Yoni Kasten, Yunrong Guo, Shie Mannor, Gal Chechik, and Xue Bin Peng. Calm: Conditional adversarial latent models for directable virtual characters. In *ACM SIGGRAPH 2023 Conference Proceedings*, pages 1–9, 2023.
- [22] Levi Fussell, Kevin Bergamin, and Daniel Holden. Supertrack: Motion tracking for physically simulated characters using supervised learning. *ACM Transactions on Graphics (TOG)*, 40(6):1–13, 2021.

- [23] Jungdam Won, Deepak Gopinath, and Jessica Hodgins. Physics-based character controllers using conditional vaes. *ACM Transactions on Graphics (TOG)*, 41(4):1–12, 2022.
- [24] Heyuan Yao, Zhenhua Song, Baoquan Chen, and Libin Liu. Controlvae: Model-based learning of generative controllers for physics-based characters. *ACM Transactions on Graphics (TOG)*, 41(6):1–16, 2022.