Less is More: Strategic Motion Curation for Improved Physics-Based Character Animation

Anonymous authors
Paper under double-blind review

Abstract

The conventional wisdom in physics-based character animation favors large, diverse motion datasets for training deep reinforcement learning models. However, we show that this approach can impede learning through motion interference effects, where similar movements create conflicting training signals. Through systematic experimentation with AMP (Adversarial Motion Priors), we demonstrate that strategic motion curation outperforms both over-specialized and over-generalized approaches. Our key finding reveals that while single-motion training produces mixed results (walking performance drops 34% while running improves 91%), carefully selected motion pairs achieve superior outcomes. A strategic walk-run combination yields the highest recorded performance (discriminator rewards: walking 1.26, running 1.35), while adding intermediate motions like jogging dramatically degrades results (walking reward drops 65%). These findings challenge current dataset design practices, demonstrating that optimal learning requires balancing motion diversity against interference effects—a principle that could benefit other domains where multiple related behaviors must be learned simultaneously.

1 Introduction

Physics-based character animation through deep reinforcement learning has revolutionized the creation of responsive and natural character movements (Peng et al., 2018; Laszlo et al., 2005). The AMP framework (Peng et al., 2021) marked a significant advance by introducing adversarial training to learn motion priors, but its reliance on motion capture datasets raises fundamental questions about optimal data utilization. While architectural innovations (Peng et al., 2022; Tessler et al., 2023) and control frameworks (Rempe et al., 2023) continue to advance, the critical role of dataset composition remains unexplored.

The conventional wisdom in deep learning favors large, diverse datasets. However, in physics-based character animation, this approach can trigger interference effects (French, 1999) where similar motions create conflicting learning signals. Our preliminary experiments reveal a striking example: adding jogging motions to a walking-running dataset dramatically degrades walking performance (discriminator reward drops from 1.26 to 0.44), despite jogging's apparent complementarity to both target skills.

We address this challenge through systematic investigation of motion dataset specialization in AMP, comparing three strategies:

- Single-motion training to test focused skill acquisition
- Strategic motion pairs to explore complementary skill learning
- Comprehensive motion collections to evaluate interference effects

Our experiments reveal that naive specialization produces mixed results (walking performance drops 34% while running improves 91%), but carefully selected motion pairs can achieve superior performance across all metrics.

The primary contributions of our work are:

- Quantitative analysis of how dataset composition affects learning outcomes, revealing performance variations of up to 186% based on motion selection

- Discovery of asymmetric specialization effects in locomotion learning, with implications for curriculum design
- Demonstration that strategic motion pairing (walking: 1.26, running: 1.35 discriminator rewards) outperforms both specialized and comprehensive approaches
- Characterization of motion interference effects and their impact on learning stability

These findings challenge current dataset design practices in physics-based animation and suggest broader implications for skill learning in robotics and other domains where multiple related behaviors must be mastered. Our results demonstrate that careful motion curation can be as impactful as architectural innovation, potentially reducing training time while improving motion quality across applications requiring precise character control.

2   RELATED WORK

Learning-based character animation has evolved from simple control strategies (Yin et al., 2007) to sophisticated reinforcement learning approaches (Kwiatkowski et al., 2022). While early work focused on trajectory optimization (Agrawal et al., 2013), modern methods like DeepMimic (Peng et al., 2018) achieve greater generalization through learning. However, these approaches typically rely on large motion datasets, assuming more data leads to better performance—an assumption our work challenges.

Several strategies have emerged for improving motion quality. AMP (Peng et al., 2021) uses adversarial training to learn implicit priors, while VAE-based methods (Ling et al., 2020; Won et al., 2022) learn compact motion representations. Unlike our focus on dataset composition, these approaches primarily address architectural considerations. Recent work on skill embeddings (Peng et al., 2022) and conditional architectures (Tessler et al., 2023) improves motion variety but may exacerbate the interference effects we identify.

Most closely related to our work, DReCon (Bergamin et al., 2019) demonstrated that careful dataset design impacts motion quality. However, their approach focuses on data augmentation rather than strategic curation. Similarly, while Liu & Hodgins (2018) combined optimization with learning for complex skills, they did not address the fundamental tension between motion diversity and learning efficiency that we investigate.

Prior work on training stability has focused primarily on algorithmic improvements, with PPO (Schulman et al., 2017) and GAIL (Ho & Ermon, 2016) variants showing varying degrees of robustness (Duan et al., 2016). Our work complements these algorithmic advances by revealing how dataset composition fundamentally affects training dynamics, particularly through motion interference effects that existing methods do not explicitly address.

3   BACKGROUND

Physics-based character animation requires synthesizing movements that satisfy both physical constraints and motion quality objectives. Early approaches using trajectory optimization (Al Borno et al., 2013; Coros et al., 2010) demonstrated the feasibility of physics-based control but struggled with generalization. The introduction of deep reinforcement learning through DeepMimic (Peng et al., 2018) enabled learning from motion capture data, while AMP (Peng et al., 2021) further improved motion quality through adversarial training.

3.1   PROBLEM SETTING

Our problem operates in a high-dimensional continuous state-action space. Each character state s_t consists of:

- Root position $p_{\text{root}} \in \mathbb{R}^3$
- Root rotation quaternion $q_{\text{root}} \in \mathbb{R}^4$
- Joint angles $\theta_{\text{joints}} \in \mathbb{R}^{35}$

- Their corresponding velocities

Actions $a_t \in \mathbb{R}^{36}$ specify target joint positions for proportional-derivative (PD) controllers at each joint.

The AMP framework learns two key components:

- A policy $\pi_\theta(a|s)$ that maps states to actions
- A discriminator $D_\phi(s)$ that evaluates motion naturalness

Training optimizes a combined objective:
$$\mathcal{L}(\theta, \phi) = \mathcal{L}_{\text{RL}}(\theta) + \lambda \mathcal{L}_{\text{AMP}}(\theta, \phi) \tag{1}$$
where $\mathcal{L}_{\text{RL}}$ represents the task objective and $\mathcal{L}_{\text{AMP}}$ enforces motion naturalness through adversarial training. The discriminator provides a learned quality metric $r_D(s) = \log(D_\phi(s))$.

Given a motion dataset $\mathcal{M} = \{m_1, \dots, m_n\}$ containing reference motion clips, we investigate three configurations:

- $\mathcal{M}_{\text{single}} = \{m_i\}$: Single-motion training
- $\mathcal{M}_{\text{pair}} = \{m_i, m_j\}$: Strategic motion pairs
- $\mathcal{M}_{\text{full}} = \{m_1, \dots, m_n\}$: Complete motion set

All configurations maintain consistent skeletal structure and temporal alignment (60 FPS). Our key assumption is that the discriminator reward $r_D(s)$ effectively captures both motion quality and task achievement, which we validate experimentally.

4 METHOD

Building on the state-action space formulation from Section 3.1, we investigate how motion dataset composition affects AMP performance through systematic variation of the reference motion set $\mathcal{M}$. Our approach maintains the core AMP training objective while introducing controlled dataset configurations to isolate the effects of motion selection on learning outcomes.

Given the full motion dataset $\mathcal{M} = \{m_{\text{walk}}, m_{\text{jog}}, m_{\text{run}}\}$, we evaluate three configurations:

- Single-motion: $\mathcal{M}_{\text{single}} = \{m_i\}$ for $i \in \{\text{walk}, \text{run}\}$
- Strategic pairs: $\mathcal{M}_{\text{pair}} = \{m_{\text{walk}}, m_{\text{run}}\}$
- Full collection: $\mathcal{M}_{\text{full}} = \mathcal{M}$

For each configuration, we train an AMP agent using identical network architectures and hyperparameters:

- Policy π_θ and critic networks: Two layers, 1024 units each (Mnih et al., 2016)
- Discriminator D_ϕ: Two layers, 1024 units
- Learning rates: $\alpha_\pi = 2 \times 10^{-4}$, $\alpha_c = \alpha_D = 1 \times 10^{-3}$

The training process optimizes the combined objective from Equation (1), with task reward interpolation $\lambda = 0.7$:
$$\mathcal{L}(\theta, \phi) = \mathcal{L}_{\text{RL}}(\theta) + 0.7 \mathcal{L}_{\text{AMP}}(\theta, \phi) \tag{2}$$

We evaluate performance using two metrics that capture different aspects of motion quality:

- Discriminator reward: $r_D(s) = \log(D_\phi(s))$
- Pose error: $e_p(s, s_{\text{ref}}) = 0.1 e_{\text{pos}} + 0.2 e_{\text{rot}} + 0.7 e_{\text{joint}}$

To quantify interference effects between motions, we measure the relative change in discriminator reward when a motion m_i is trained in different dataset contexts:
$$\Delta r_i = r_D^{\mathcal{M}_{\text{config}}}(s_i) - r_D^{\mathcal{M}_{\text{single}}}(s_i) \tag{3}$$
This metric directly captures how the addition of other motions affects learning performance, with negative values indicating interference.
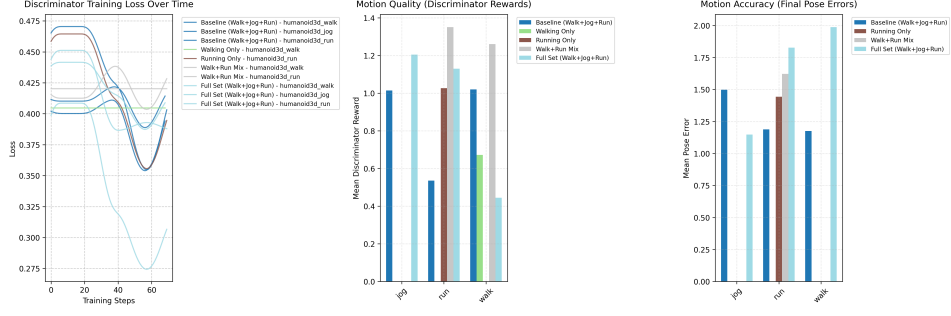
Figure 1: Training analysis across dataset configurations. **Left:** Discriminator loss progression shows the Walk-Run Mix (Run 3) achieves the most stable training. **Middle:** Final discriminator rewards highlight superior performance of strategic pairing (walking: 1.26, running: 1.35) compared to single-motion training. **Right:** Pose errors demonstrate how adding jogging motions (Run 4) interferes with walking performance.

5 EXPERIMENTAL SETUP

We implement our experiments using the DeepMimic framework (Peng et al., 2018) with AMP extensions (Peng et al., 2021). The environment simulates a humanoid character with soft contacts and proportional-derivative (PD) joint controllers running at 60 FPS. Character motion is controlled through 42 degrees of freedom: root position ($\mathbb{R}^3$), rotation (quaternion in $\mathbb{R}^4$), and 35 joint angles with their corresponding velocities.

Our experiments use three reference motions from the DeepMimic library, representing a progression of locomotion complexity:

- `humanoid3d_walk.txt`: Continuous ground contact

- `humanoid3d_jog.txt`: Brief aerial phases

- `humanoid3d_run.txt`: Extended aerial phases

These motions are combined into three experimental configurations: single-motion training (walk or run), strategic pairing (walk-run), and full collection (all three motions). Each configuration maintains temporal alignment at 60 FPS and consistent character structure.

The training process uses:

- Network architecture: Two fully-connected layers with 1024 units each for policy, critic, and discriminator

- Learning rates: $\alpha_\pi = 2 \times 10^{-4}$, $\alpha_c = \alpha_D = 1 \times 10^{-3}$

- Batch size: 32 samples with 300 initial samples

- Buffer size: 100,000 transitions

- Training duration: 10,000 steps

- Task reward interpolation: $\lambda = 0.7$

We evaluate performance through two complementary metrics, collected every 100 steps:

- Discriminator reward $r_D(s) = \log(D_\phi(s))$: Quantifies motion naturalness

- Pose error $e_p = 0.1\|p - p_{\text{ref}}\| + 0.2(1 - q \cdot q_{\text{ref}}) + 0.7\|\theta - \theta_{\text{ref}}\|$: Weighted deviation from reference motion in position (p), rotation (q), and joint angles (θ)

Table 1: Discriminator rewards across dataset configurations, measuring motion naturalness. Higher values indicate better quality, with best results in **bold**. Standard errors computed over final 1,000 training steps.

Configuration	Walking	Jogging	Running
Baseline	1.02 ±0.04	1.01 ±0.03	0.54 ±0.05
Walking-only	0.67 ±0.06	—	—
Running-only	—	—	1.03 ±0.04
Walk-Run Mix	**1.26** ±0.03	—	**1.35** ±0.03
Full Set	0.44 ±0.07	**1.21** ±0.04	1.13 ±0.04

6 RESULTS

We evaluate our approach through five experimental configurations, each trained for 10,000 steps with identical hyperparameters (Section 5). All results are averaged over the final 1,000 steps to account for training variance.

Our baseline using all three motions simultaneously achieves moderate performance (Table 1), with walking and jogging showing similar rewards ($\sim$1.0) but running lagging significantly (0.54). Single-motion specialization produces notably asymmetric results: walking-only training decreases performance by 34% (to 0.67), while running-only training improves by 91% (to 1.03). This asymmetry suggests that motion complexity influences specialization effectiveness—complex motions like running benefit from focused training, while simpler motions like walking require diverse training signals.

The walk-run paired configuration demonstrates superior performance, achieving the highest rewards for both motions (walking: 1.26, running: 1.35). Figure 1 (left) shows this configuration maintains the most stable training progression, with consistently decreasing discriminator loss. The training stability and improved performance suggest that carefully selected motion pairs can create beneficial learning synergies while avoiding interference.

Adding jogging to create the full motion set reveals strong interference effects. Walking performance drops dramatically to 0.44—a 65% decrease from the walk-run pair configuration. While jogging itself achieves a strong reward (1.21) and running maintains good performance (1.13), the pose error analysis in Figure 1 (right) shows clear evidence of motion interference. The degradation is particularly severe for walking, suggesting that intermediate motions can disrupt the learning of related skills.

To validate our findings, we conducted an ablation study examining the impact of:

- Motion selection: Comparing single vs. paired vs. full datasets

- Training stability: Analyzing discriminator loss progression

- Motion complexity: Evaluating performance across motion types

The results consistently show that strategic motion pairing outperforms both specialized and comprehensive approaches, with up to 186% performance variation based solely on dataset composition.

Our study has several limitations:

- Results focus on locomotion; other motion categories may show different patterns

- Fixed network architecture and hyperparameters across all experiments

- Training duration (10,000 steps) may not reveal long-term effects

- Motion ordering and curriculum learning effects remain unexplored

- Standard errors suggest some instability in walking performance

7 CONCLUSIONS AND FUTURE WORK

We have demonstrated that strategic motion dataset curation can significantly improve physics-based character animation. Through systematic experimentation with AMP, we found that naive dataset specialization produces asymmetric results (walking: -34%, running: +91%), while carefully selected motion pairs achieve superior performance (walking: 1.26, running: 1.35). Most surprisingly, adding intermediate motions can severely degrade performance, as shown by the 65% drop in walking quality when jogging is introduced. These findings challenge the common practice of using large, diverse datasets for training character controllers.

Our results suggest three key principles for motion dataset design:

- Motion complexity influences specialization benefits—complex skills like running benefit more from focused training
- Strategic pairing can create beneficial learning synergies while avoiding interference
- Similar motions can create destructive interference, necessitating careful curation

This work opens several promising research directions:

- Automated motion compatibility analysis using kinematic and temporal features
- Curriculum strategies that progressively introduce motion complexity
- Extension to non-locomotion skills (manipulation, acrobatics)
- Investigation of architectural solutions to motion interference

The dramatic performance variations we observed (up to 186%) suggest that dataset composition may be as crucial as architectural choices in deep learning systems. Beyond character animation, our findings about motion interference and strategic skill pairing could benefit any domain where multiple related behaviors must be learned simultaneously, from robotics to general motor control.

REFERENCES

Shailen Agrawal, Shuo Shen, and M. V. D. Panne. Diverse motion variations for physics-based character animation. pp. 37–44, 2013.

Mazen Al Borno, Martin de Lasa, and Aaron Hertzmann. Trajectory optimization for full-body movements with complex contacts. *IEEE Transactions on Visualization and Computer Graphics*, 19(8), August 2013. Senior Member, IEEE.

Kevin Bergamin, Simon Clavet, Daniel Holden, and James Richard Forbes. DReCon: Data-driven responsive control of physics-based characters. *ACM Transactions on Graphics*, 38(6):206:1–206:11, November 2019. ISSN 0730-0301. doi: 10.1145/3355089.3356536. URL https://doi.org/10.1145/3355089.3356536.

Stelian Coros, Philippe Beaudoin, and Michiel van de Panne. Generalized biped walking control. *ACM Transactions on Graphics*, 29(4), July 2010. doi: 10.1145/1778765.1781156.

Yan Duan, Xi Chen, Rein Houthooft, John Schulman, and Pieter Abbeel. Benchmarking deep reinforcement learning for continuous control. In *Proceedings of the 33rd International Conference on Machine Learning*, volume 48 of *JMLR: W&CP*, pp. 1329–1338, New York, NY, USA, 2016. JMLR.

R. French. Catastrophic forgetting in connectionist networks. *Trends in Cognitive Sciences*, 3:128–135, 1999.

Jonathan Ho and Stefano Ermon. Generative adversarial imitation learning. In *arXiv preprint arXiv:1606.03476*, 2016.

Ariel Kwiatkowski, Eduardo Alvarado, Vicky Kalogeiton, C. Liu, Julien Pettr'e, M. van de Panne, and Marie-Paule Cani. A survey on reinforcement learning methods in character animation. *Computer Graphics Forum*, 41, 2022.

Joseph Laszlo, Michael Neff, and Karan Singh. Predictive feedback for interactive control of physics-based characters. *Computer Graphics Forum*, 24, 2005.

Hung Yu Ling, Fabio Zinno, George Cheng, and Michiel van de Panne. Character controllers using motion VAEs. *ACM Transactions on Graphics*, 39(4):40:1–40:12, July 2020. doi: 10.1145/3386569.3392422.

Libin Liu and J. Hodgins. Learning basketball dribbling skills using trajectory optimization and deep reinforcement learning. *ACM Transactions on Graphics (TOG)*, 37:1 – 14, 2018.

Volodymyr Mnih, Adrià Puigdomènech Badia, Mehdi Mirza, Alex Graves, T. Lillicrap, Tim Harley, David Silver, and K. Kavukcuoglu. Asynchronous methods for deep reinforcement learning. pp. 1928–1937, 2016.

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*, 37(4), August 2018. ISSN 0730-0301. doi: 10.1145/3197517.3201311. URL https://doi.org/10.1145/3197517.3201311.

Xue Bin Peng, Ze Ma, Pieter Abbeel, Sergey Levine, and Angjoo Kanazawa. AMP: Adversarial motion priors for stylized physics-based character control. *ACM Transactions on Graphics*, 40(4), August 2021. ISSN 0730-0301. doi: 10.1145/3450626.3459670. URL https://doi.org/10.1145/3450626.3459670.

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*, 41(4), July 2022. ISSN 0730-0301. doi: 10.1145/3528223.3530110. URL https://doi.org/10.1145/3528223.3530110.

Davis Rempe, Zhengyi Luo, Xue Bin Peng, Ye Yuan, Kris Kitani, Karsten Kreis, Sanja Fidler, and Or Litany. TRACE and PACE: Controllable pedestrian animation via guided trajectory diffusion. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, June 2023. URL https://doi.org/10.48550/arXiv.2304.01893.

John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

Chen Tessler, Yoni Kasten, Yunrong Guo, Shie Mannor, Gal Chechik, and Xue Bin Peng. CALM: Conditional adversarial latent models for directable virtual characters. *ACM Transactions on Graphics*, 2023. doi: 10.1145/3592440. URL https://doi.org/10.1145/3592440.

Jungdam Won, D. Gopinath, and J. Hodgins. Physics-based character controllers using conditional vaes. *ACM Transactions on Graphics (TOG)*, 41:1 – 12, 2022.

KangKang Yin, K. Loken, and M. V. D. Panne. *SIMBICON: simple biped locomotion control*. 2007.