Iterative Closed-Loop Motion Synthesis for Scaling the Capabilities of Humanoid Control

Weisheng Xu[1*], Qiwei Wu[1*], Jiaxi Zhang[1], Jing Tan[1], Yangfan Li[1], Yuetong Fang[1], Jiaqi Xiong[2], Kai Wu[1], Rong Ou[1], Renjing Xu[1†]

[1]Hong Kong University of Science and Technology (Guangzhou)
[2]University of Oxford

{wxu421, qwu720}@connect.hkust-gz.edu.cn, renjingxu@hkust-gz.edu.cn

Abstract

Physics-based humanoid control relies on training with motion datasets that have diverse data distributions. However, the fixed difficulty distribution of datasets limits the performance ceiling of the trained control policies. Additionally, the method of acquiring high-quality data through professional motion capture systems is constrained by costs, making it difficult to achieve large-scale scalability. To address these issues, we propose a closed-loop automated motion data generation and iterative framework. It can generate high-quality motion data with rich action semantics, including martial arts, dance, combat, sports, gymnastics, and more. Furthermore, our framework enables difficulty iteration of policies and data through physical metrics and objective evaluations, allowing the trained tracker to break through its original difficulty limits. On the PHC single-primitive tracker, using only approximately 1/10 of the AMASS dataset size, the average failure rate on the test set (2201 clips) is reduced by 45% compared to the baseline. Finally, we conduct comprehensive ablation and comparative experiments to highlight the rationality and advantages of our framework.

1. Introduction

Physics-based humanoid control supports high dynamic locomotion, character animation, and interactive VR. The standard pipeline imitates MoCap with RL. DeepMimic [22] showed complex skills are learnable; AMP [23] improved realism; ASE [24], DReCon [1], and PHC [18] advanced reuse and generalization. Progress remains limited by costly expert MoCap and datasets that skew toward low difficulty or narrow domains.

Existing corpora reflect this trade-off. AMASS covers general low dynamic motion, while AIST++ targets professional dance [12]. Humanoid-X [21] and HuBE [19] scale via video mining or cross-embodiment aggregation, but still favor low difficulty or limited scope. Controllers trained on fixed low dynamic distributions often fail on acrobatic motions due to missing high dynamic features. The scarcity of professional, high dynamic data further constrains progress.

We introduce CLAIMS (as shown in Fig. 1), a closed-loop automated framework that co-evolves motion data synthesis and controllers. Each iteration generates harder professional motions with motion diffusion model (MDM) from language prompts, trains the controller, analyzes failures with a multimodal agent, and refines data to expand capability.

Our contributions are: (1) A semantic definition and normalization for action data, yielding a scalable dataset with difficulty tiers across diverse professional skills. (2) A competitive iterative procedure that alternates data generation and controller optimization in a game-like setup, enabling the policy to surpass its difficulty ceiling and master high difficulty motions.

After iterative optimization, the dataset exhibits clear difficulty stratification and professional characteristics while preserving motion–text alignment. Only the controller requires training compute; other components are training-free and low cost, making the framework controller agnostic. As shown in Table 1, with fewer than 400 training sequences, our tracker surpasses baselines on AIST++, Motion-X/Kungfu [34], EMDB [10], and Video-Convert [8] on average. Performance improves with more iterations, and the framework consistently boosts diverse trackers.

*Equal contribution.
†Corresponding author.
[0]Project page: https://wesleyxu224.github.io/CLAIMS/

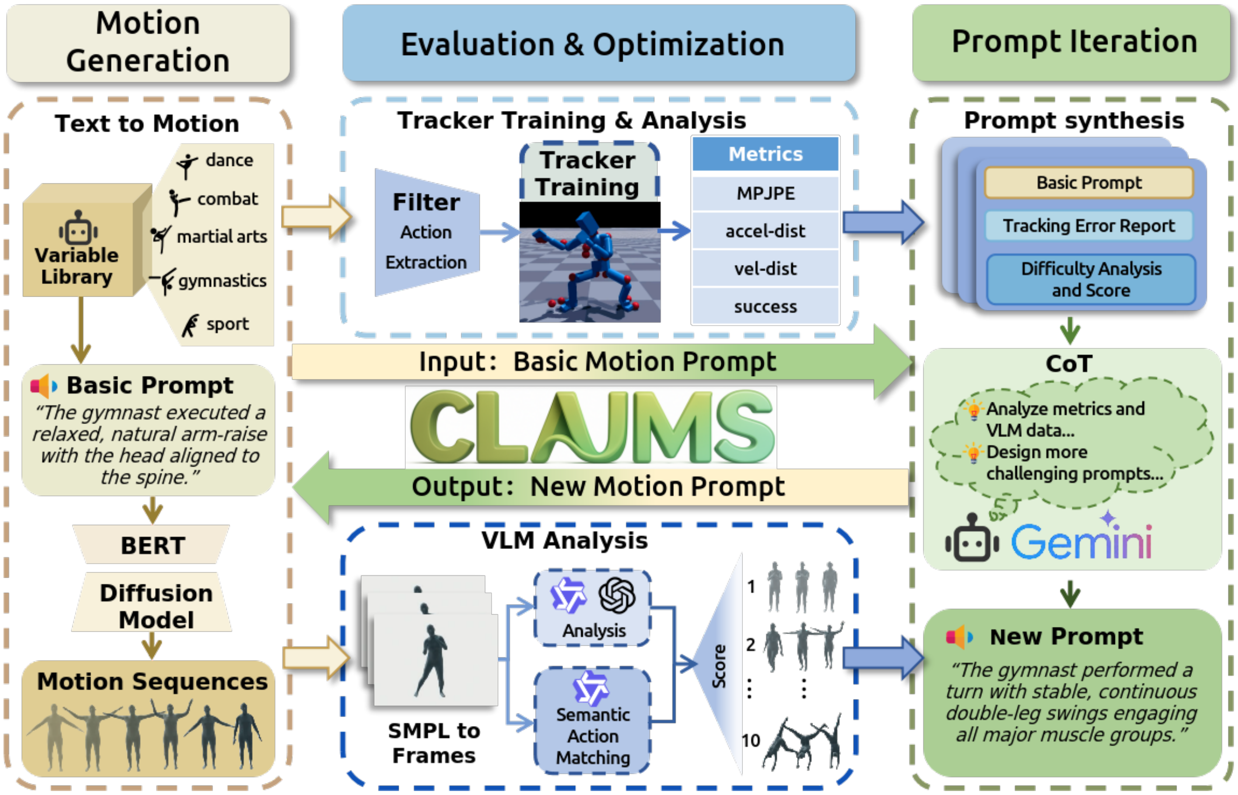


Figure 1. Overview of the CLAIMS pipeline: a closed-loop system that refines prompts from a 5-domain library (martial arts, dance, combat, sports, gymnastics), synthesizes motions with MDM, filters them by physics and VLM checks, trains humanoid trackers with RL, and uses multimodal feedback to generate progressively harder tasks.

2. RELATED WORK

Humanoid Motion Datasets. AMASS [20] unifies motion under SMPL [16] but over 90% of its sequences are low dynamic daily activities, and controllers trained on it generalize poorly to expert actions, as shown by PHC [18]. Domain-specific sets such as AIST++ [12] and EMDB [10] raise dynamism or annotation fidelity but remain narrow in coverage, scale, and diversity. PHUMA [11] scales video-based data with physics-constrained retargeting to improve plausibility, yet focuses on locomotion. To reduce collection cost at scale, Humanoid-X [21] and HuBE [19] mine internet and cross-embodiment motion, but lack reliable semantics and difficulty stratification. Recent million-scale corpora aim to couple scale with semantics and evaluation, for example Go to Zero [6] with automated high precision labels and a zero-shot benchmark, MotionLib [32] with hierarchical text and Motionbook encoding for semantic alignment, and Motion-X++ [34] with expanded multimodal full-body coverage. Despite these advances, most datasets assume a static difficulty distribution and do not provide controller-aware grading or a dynamic expansion process that co-evolves with policy capability. Models for text or multimodal motion generation. MDM [27] established a text-conditioned diffusion baseline; MoFusion [4] improves denoising and sample quality; MotionGPT and MotionGPT-2 [7, 31] unify generation and understanding with a language model formulation; T2M-X [15] extends to partially annotated full body motion. For real-time controllability and deployment, Being-M0.5 [2] leverages the HuMo100M scale corpus and motion tokenization to enable unified vision, language, and motion control. However, these models largely inherit static, low dynamic training distributions and insufficiently cover professional actions; even with scaling and zero-shot evaluation [6, 32], they lack difficulty adaptive mechanisms tied to controller mastery.

Contrast to our work. We build a scalable dataset across five professional domains—martial arts, dance, combat, sports, and gymnastics—with semantic tags and explicit difficulty tiers. Our closed-loop framework expands the high difficulty distribution in response to the controller's current competence, overcoming fixed difficulty corpora and enabling stable gains in professional highly dynamic scenarios.

Physics-Based Humanoid Control. Physics-based humanoid control has advanced in tracking yet still faces two bottlenecks: reliance on fixed training distributions and weak generalization across trackers. Early imitation and tracking methods such as DeepMimic and Skeleton2Humanoid [13, 22] improved skill coverage and plausibility. Adversarial and skill reuse frameworks including AMP, ASE, and PHC [18, 23, 24] further enhanced realism and coverage, and Neural Categorical Priors [35] reduced behavioral imbalance. Despite this, controllers inherit static data and struggle on unseen high dynamic skills; universal controllers such as UHC and MaskedMimic [17, 26] often fail on gymnastics tumbling. FARM [8] targets explosive motions with frame-level augmentation and a residual MoE and introduces the HDHM benchmark, yet a gap in high dynamic capability and evaluation remains.

To broaden capability, diffusion and RL have been integrated in a closed loop for multi-task control [14, 28, 30], and language or multimodality has been used for richer task specification [5, 9]. These approaches still rely on static corpora and bespoke pipelines and lack difficulty adaptation to controller mastery. At a broader level, zero-shot whole body control [29] leverages behavioral foundation models but depends on large offline pretraining and does not deliver universally adaptive training for diverse single primitive trackers. We propose a general training framework for single primitive trackers that couples controller optimization with adaptive difficulty escalation. It raises data difficulty in response to controller competence, requires no specialized high dynamic datasets or pipeline changes, and yields consistent gains on professional high difficulty motion benchmarks.

Iterative Data Generation. Data scarcity remains the main bottleneck across domains, with pipelines still limited in domain specificity, scale, and physical realism. A growing trend is adaptive closed loop curation that couples data selection and generation so models and datasets improve together.

In reasoning and language modeling, SAI-DPO and SEAL [25, 36] adapt data to the model during training by selecting hard examples, self-generating fine-tuning data, and self-editing instructions. In robotics, RoboTwin 2.0 [3] uses a multimodal language model pipeline to auto generate tasks and objects and improves sim-to-real transfer, but its loop relies mainly on success rates and image metrics and leaves evaluation under specified. In character control, PARC [33] pairs a motion generator with a physics-based tracker in an iterative generate-correct-augment-retrain loop, yet adopts a single evaluation criterion and does not generalize well across scenarios.

We address these gaps with two components: (1) a professional semantic taxonomy for motion that enables principled categorization and difficulty stratification; and (2) a fused evaluation and selection loop that combines controller performance metrics with vision language model feedback under a multi-dimensional protocol. Guided by the taxonomy, the loop escalates difficulty adaptively, pushes generated motions away from pretraining biases, and provides a clear basis for sustained improvement.

Figure 2. Difficulty-aware variable library across five domains and four compositional axes.

3. Methodology

3.1. Defining High-Difficulty Motion Data

We formalize motion professionalism and difficulty for high-skill controller training. Guided by expert requirements and capture risk, we curate five domains (Fig. 2): sports, dance, combat, gymnastics, and martial arts, emphasizing high-dynamic motions. Difficulty is defined along four axes: base action (atomic skill); combo action (composition logic); detail (technical cues like limb placement); and speed and rhythm (temporal structure). For example, a dance prompt combines a grand allegro (base), a saut de basque chain (combo), and a triple pirouette (detail) at a steady tempo (speed). These templates constrain generation and guide dataset optimization, ensuring principled difficulty escalation aligned with professional characteristics.

3.2. Dataset Generation

We synthesize training data with a low cost, pretrained text conditioned motion diffusion model, MDM [27], using a 50 step sampler and a DistilBERT text encoder. Although MDM is trained on HumanML3D, its latent space enables compositional mixing of action primitives, yielding behaviors that remain distributionally consistent with the source

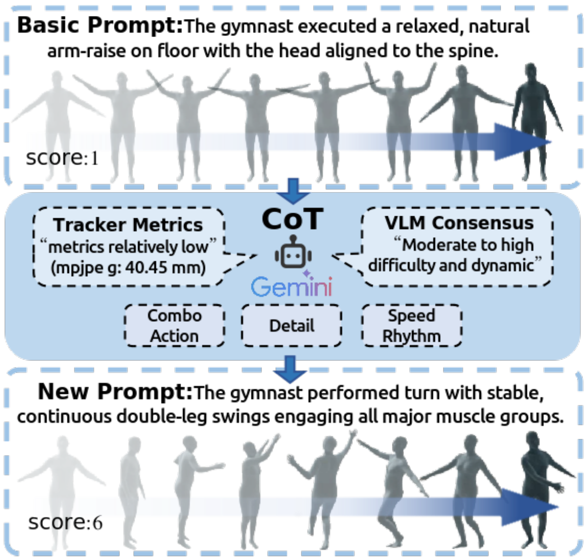


Figure 3. Prompt-to-prompt data generation.

corpus while introducing novel combinations absent from the original data. Sampling follows standard conditional diffusion with fixed weights. To elicit complex professional behaviors without modifying the generator, we use templated action prompts derived from our semantic taxonomy and automatically instantiate them with an auxiliary LLM. The template-based design provides more stable, domain faithful control over motion composition and difficulty than free form LLM descriptions.

In each generation round, we sample text prompts from a difficulty-aware variable bank that spans five expert domains: martial arts, dance, combat, gymnastics, and sports. DistilBERT encodes each prompt into a conditioning embedding for MDM [27], which synthesizes the corresponding motion trajectory. We apply validity checks such as root joint height bounds to remove non-physical artifacts including floating, sinking, and interpenetration, and use a vision and language model (VLM) to assess semantic alignment between prompt and motion. Segments that pass both filters are added to the synthetic training set. The pipeline repeats over multiple rounds with gradually increased difficulty, allowing the corpus to evolve toward high dynamic professional behaviors while keeping computational cost low and preserving the pretrained generator.

3.3. Training and evaluation of Controller

Our single primitive tracker uses reinforcement learning imitation under a single policy with dense rewards on pose, joint velocities, end effectors, and contact events, plus sample filtering, to achieve high coverage tracking and stable optimization. The same RL setup remains compatible with diverse data and feature pipelines such as PHC [18] and MaskedMimic [26], and we keep their original hyperparameters and training settings. The physics-based controller is trained only on our synthesized motions and is the only compute intensive component.

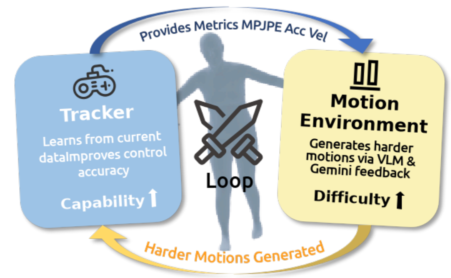


Figure 4. Competitive iteration between the controller and the dataset

After convergence, we evaluate with mpjpe-g, the average joint position error in the world frame; mpjpe-l, the joint position error in the root relative frame; vel-dist, the mean per joint linear velocity discrepancy that reflects smoothness; and accel-dist, the mean per joint acceleration discrepancy that exposes high frequency jitter. These metrics delineate the controller's capability frontier and feed back into the next iteration to guide difficulty escalation in the data generation loop.

3.4. Competitive iteration between the controller and the dataset

We employ a competitive iterative curriculum (Fig. 4) to raise the controller performance ceiling. After each training cycle, if objective metrics exceed predefined thresholds, we treat the current distribution as mastered and escalate data difficulty. The controller then trains on harder samples, which advances its capability. Progress is driven by co-evolution of controller performance and data difficulty, and is measured by a joint evaluation that combines objective physical tracking metrics with subjective visual judgments, producing a self-reinforcing curriculum. For closed loop feedback, we render each training motion as a concatenated SMPL sequence and evaluate it with two vision and language models, GPT-4o and Qwen-VL-MAX. For each clip, the VLMs return a subjective difficulty score and descriptors covering action sequence, technical complexity, intensity, balance, continuity, plus a rationale, while the controller reports objective physical metrics on the same motion. We concatenate these signals into a semantic observation vector (obs) that captures both physical execution and visual perception.

When the observation indicates improved mastery, we raise data difficulty. Guided by our professional templates and difficulty attributes, we generate harder samples within the target domains using language prompts and MDM-based compositional synthesis. Each composition step is

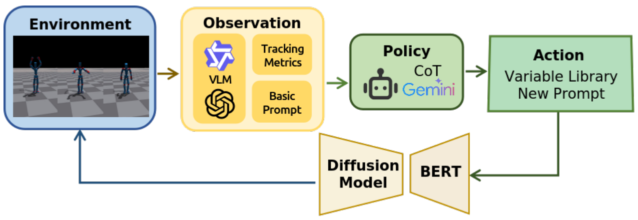


Figure 5. Schematic Diagram of the Automated Iterative Loop

treated as an action that yields a new motion sequence, and the loop is automated end-to-end so data difficulty and controller competence advance together.

3.5. Automation of the iterative process

We instantiate a central LLM policy with a role prompt to drive a competitive closed loop curriculum where dataset difficulty co-evolves with controller performance. Concretely, we use Gemini with chain-of-thought to fuse tracker metrics, subjective visual feedback, and the previous action prompt into a semantic observation, while a VLM ensemble (GPT-4o and Qwen-VL-MAX) supplies prompt–motion alignment signals. The policy outputs the next prompt from a difficulty aware variable bank grounded in our professional templates, and the environment executes motion tracking training on the synthesized clips. Optimization is implicit: the policy aims to improve physical tracking scores while steadily raising annotated difficulty, yielding a self-reinforcing curriculum. The entire loop process is illustrated in Algorithm 1 and Fig. 7. Running the loop for K iterations yields a five domain corpus with K difficulty tiers across martial arts, dance, combat, gymnastics, and sports, together with a physics-based controller that adapts to heterogeneous high difficulty motions. A centralized LLM with chain-of-thought uses multimodal feedback to select informative harder prompts while preserving semantic fidelity, enabling stable performance gains without pipeline specific modifications.

4. Experiments

Our experiment aims to address the following core question: (1) Do our definitions and generation extrapolate beyond MDM's training sets (AMASS, HumanAct12)? (2) Does the framework achieve generalization to unseen high-dynamics and high-difficulty data? (3) Does the feedback loop raise difficulty over time and improve controller stability and performance?

4.1. Experimental Setup

All experiments are performed on a single NVIDIA A6000 GPU. We use the PHC single-primitive training configuration [18] as the base setting for reinforcement learning. Each loop iteration trains a physics-based tracker on the synthetic motion dataset generated from the current prompt batch. Unless otherwise noted, we report metrics on six standard motion test sets: kungfu, emdb, amass, mdm, aist++, and video-converted.

Algorithm 1 LLM-Driven Competitive Dataset–Controller Iteration

Input: variable library $\mathcal{L}$, templates $\mathcal{T}$, generator G, VLM evaluator F_{VLM}, policy LLM π_θ (Gemini CoT).
Initialize dataset $\mathcal{D} \leftarrow \varnothing$, motion sets $\mathcal{M} \leftarrow \varnothing$, tracker π_0^{trk}, action a_0.
for $k = 0$ **to** $K-1$ **do**
  Compute tracking metrics m_k and VLM difficulty/feedback v_k.
  Encode previous action $e_k = \phi(a_k)$ and form observation $o_k = [m_k, v_k, e_k]$.
  Generate new action prompts:

$$A_k = \{a_k^1, \ldots, a_k^M\} \sim \pi_\theta(o_k, \mathcal{L}, \mathcal{T}).$$

  Initialize $M_k \leftarrow \varnothing$.
  for $a_k^j \in A_k$ **do**
    $q_k^j \leftarrow G(a_k^j)$; **if** fails physics **continue**.
    if VLM alignment is sufficient: $M_k \leftarrow M_k \cup \{(q_k^j, a_k^j)\}$.
  end for
  $\mathcal{D} \leftarrow \mathcal{D} \cup M_k$; $\pi_{k+1}^{\text{trk}} \leftarrow \text{TrainTracker}(\mathcal{D})$.
  Store motion set: $\mathcal{M} \leftarrow \mathcal{M} \cup \{M_k\}$.
  Update summary prompt a_{k+1}.
end for
Return: best tracker π_*^{trk} and motion sets $\mathcal{M}$.

4.2. Motion professionalism verification

Motion dataset features. We build training data by combinatorially instantiating a variable prompt library, synthesizing motions with MDM-step50-BERT, rendering poses via SMPL, and filtering with a VLM; Fig. 7a illustrates representative prompt categories and shows semantic alignment between action terms and rendered clips(Fig. 8). To assess domain specificity, we compare (i) professionally curated martial-arts motions, (ii) motions synthesized from our expert martial arts prompts, and (iii) motions from Gemini-generated random prompts. Using per-frame mean global joint rotation features and t-SNE visualization, expert prompt syntheses substantially overlap the professional martial-arts manifold, while random-prompt motions lie far away. This indicates our prompt design encodes salient domain priors and mitigates distributional bias in MDM prompt-conditioned synthesis.

Motion dataset and MDM relationship. We run a t-SNE analysis on clip-level pose embeddings to test whether our

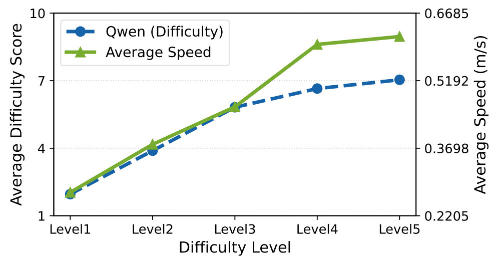
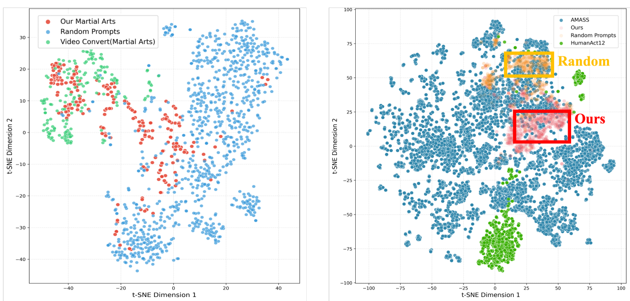


Figure 6. Loop-wise Difficulty and Speed Trends of Qwen Evaluations



(a) (b)

Figure 7. (a) The t-SNE graph of martial arts datasets comparison; (b) The t-SNE graph of all datasets(AMASS/Ours/Random Prompts/HumanAct).

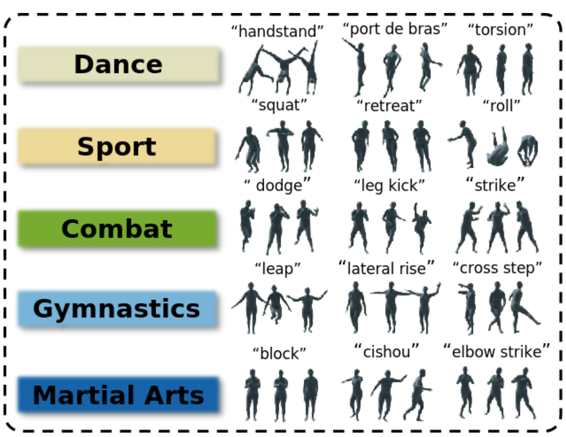


Figure 8. Visualization of 5 categories of professional movements

prompt-conditioned synthesis departs from MDM's training distribution. The corpus includes HumanML3D training data (AMASS, HumanAct12), motions synthesized by MDM from Gemini-generated random prompts, and motions synthesized by MDM from our expert prompts. Fig. 7b shows our expert-prompt samples lie largely outside the MDM training manifold, while random-prompt samples cluster within it—indicating MDM can produce out-of-distribution motions and that expert prompts effectively

Table 1. Success rate (%) of different pipelines across test sets (totaling 2201 clips). The test suite includes: Motion-X/Kungfu (663), EMDB (45), AIST++ (1320), and Video-Convert (173). We report per-set results and the average (Avg) calculated by clip.

Method	Kungfu	EMDB	AIST++	VC	Avg
AMASS	47.1	53.3	67.6	31.2	58.3
L0	37.8	31.1	68.8	33.3	55.9
L1	47.7	33.3	75.3	38.7	64.0
L2	51.7	51.1	73.3	41.6	63.8
L3	59.1	64.4	82.1	50.9	72.4
L4	55.8	55.6	84.0	45.1	71.8
L5	**60.6**	60.0	85.2	54.3	74.8
L6	60.3	**64.4**	**88.1**	**58.9**	**76.9**†

† Compared to the AMASS baseline, our L6 model reduces the average failure rate by 45% (from 41.7% to 23.1%).

elicit novel content beyond the original dataset coverage.

4.3. Generalization research

OOD dataset testing. We evaluate portability by instantiating our iterative framework with the PHC single-primitive controller trained on AMASS and comparing to a non-iterative baseline under matched protocols. The iterative pipeline synthesizes motions with MDM-step50-BERT, applies root-height and VLM-based filtering to form loop0, trains PHC, and collects tracking metrics. Using SMPL renderings and VLMs for difficulty/attribute signals, we construct an observation passed to a Gemini CoT module to propose harder, controller-targeted prompts; repeating synthesis and filtering yields loop1 and subsequent loops. On four high-difficulty third-party benchmarks (Tab. 1), loop0 lags the baseline, loop1 already surpasses it, and later loops continue improving—showing the feedback-driven, prompt-conditioned curriculum is plug-and-play with the PHC single-primitive paradigm and outperforms one-shot training on large, generic corpora.

Compatibility of different trackers. We test cross-paradigm effectiveness by instantiating our iterative framework with the DeepMimic-based MaskedMimic controller, comparing to a baseline MaskedMimic (FC) trained on AMASS in IsaacGym. Our method uses loop0 set, trains the controller, collects PHC-style inference-time metrics, and uses Gemini to generate harder loop1 prompts for data synthesis and continued training. On third-party, high-difficulty suites (Tab. 2 and Fig.9), loop0 matches the baseline, while loop1 delivers consistent, significant improvements—showing that feedback-driven, prompt-conditioned data-and-policy refinement transfers to distinct tracking-by-imitation frameworks.

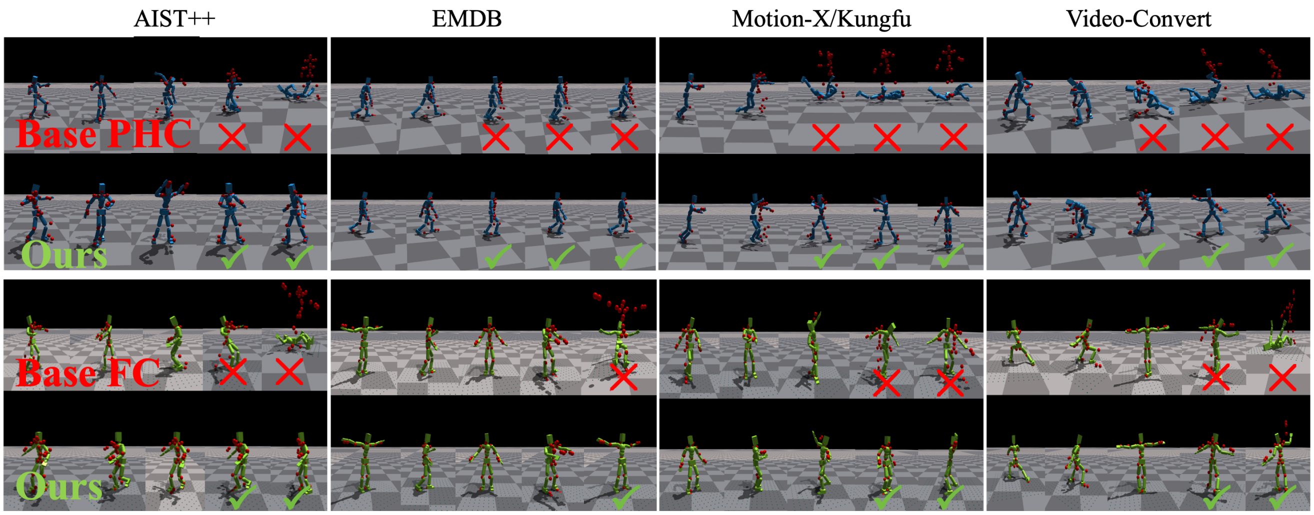


Figure 9. Qualitative tracking performance between Ours and Baselines of PHC and Maskedmimic.

Table 2. Maskedmimic FC*: Success Rate on each test set."*" means early period.

Dataset	AMASS	loop0	loop1
Motion-X/Kungfu	57.2	54.0	**65.8**
EMDB	53.3	48.9	**71.1**
AIST++	68.9	75.3	**83.9**
Video-Convert	41.6	47.4	**62.4**

4.4. Ablation Study

VLM scoring reliability verification. To ensure an objective verification, we construct five tiers of compositional test prompts from predefined variable libraries (Levels 1–5) and instantiate 200 motions per tier (N=1,000). For each motion, we compute the mean velocity as a coarse physical difficulty proxy and render frame sequences. A vision–language model, Qwen, blinded to the text prompts and given only the rendered frames, assigns a difficulty score on a 1–10 scale. As shown in Fig. 6, both the prompt-defined difficulty level and the physical metric exhibit a clear monotonic trend: average velocity increases with tier, and Qwen's difficulty ratings rise accordingly. This alignment across prompt design, physics-based indicators, and VLM-based assessments supports the credibility of the automated VLM scoring.

Observation role ablation. We ablate the observation vector in our iterative feedback pipeline on the PHC single-primitive controller, fixing the Gemini-CoT difficulty objective and comparing: (1) full (physics metrics + VLM), (2) w/o VLM, (3) w/o physics metrics, and (4) w/o both. Starting from the same loop0, each variant is trained through 'loop3' and evaluated by average success on four third-party, high-difficulty benchmarks (Tab. 3). Results are consistent: no observations $<$ no physics $<$ no VLM $<$ full. This indicates the scheduler becomes stochastic without observations; physics-based tracking metrics are the most diagnostic for targeted challenge design, while VLM feedback provides complementary, subjective difficulty cues. Both signals help, with physics being critical and VLM offering additional gains.

Variable template library role verification. To assess the contribution of the variable library, we conduct an ablation in which it is removed and Gemini-CoT is tasked with proposing harder, expert-oriented prompts directly. The 'loop0' model and training protocol remain unchanged; we then iterate to 'loop3'. As summarized in Tab. 3, this variant underperforms the configuration with the variable library across all third-party benchmarks. These results indicate that the variable library provides a structured prior that stabilizes prompt generation, yielding more targeted and diagnostically challenging motions. By improving the specificity and controllability of the synthesized data, the library enables Gemini-CoT to more reliably escalate task difficulty and thus push the controller's performance frontier.

Feedback iteration role verification. To isolate the effect of feedback-driven iteration from mere data scale, we construct a size-matched, non-iterative baseline. Concretely, we randomly sample from the variable library a training set whose cardinality equals the cumulative data

Table 3. Ablation study on success rate across different test sets and loops.

Method	Motion-X/Kungfu			EMDB			AIST++			Video-Convert		
	L1	L2	L3	L1	L2	L3	L1	L2	L3	L1	L2	L3
full (var)	47.7	51.7	59.1	33.3	51.1	**64.4**	75.3	73.3	**82.1**	38.7	41.6	50.9
w/o var	53.5	59.1	**59.9**	53.3	57.8	53.3	75.6	80.2	76.5	50.3	48.0	46.8
w/o VLMs	46.3	54.8	58.4	42.2	62.2	60.0	73.8	80.0	81.9	43.4	45.1	**57.2**
w/o metrics	48.0	50.7	54.6	31.1	31.1	35.6	78.3	77.5	78.9	38.2	28.9	40.5
w/o VLMs+metrics	46.5	54.3	54.3	35.6	46.7	46.7	71.0	76.1	77.4	38.2	46.8	50.9

Table 4. Metrics(Success Rate) on each dataset for 1400 clips testset Loop0 and our Loop6.

Dataset	**Loop0(1400 clips) SR**	**Loop6 SR**
Kungfu(663)	58.7	**60.3**
EMDB(45)	**73.3**	64.4
AIST++(1320)	85.3	**88.1**
Video-Convert(173)	55.5	**59.0**

used across loop0 to loop6, and train a single model to convergence under the same optimization protocol. As reported in Tab. 4, despite identical data budgets, our feedback-iterative regimen (‘loop0‘→‘loop6‘) consistently outperforms the one-shot baseline across multiple large-scale test suites. This evidences that the gains stem from iterative feedback and curriculum effects rather than simply from increasing the number of training motions.

Dataset difficulty variation verification Using the pretrained composition controller PHC+, we run inference on each loop's motion set and observe a monotonic drop in tracking success with increasing loop index (Tab. 5), indicating a feedback-driven curriculum that yields systematically harder motions. Velocity analyses—global root-velocity distributions (Fig. 10) and per-clip mean inter-frame speed on length-matched 180-frame sequences—show that AMASS is predominantly low-dynamics, whereas our synthesized data exhibits progressively higher velocities, broader tails, higher-frequency variations, and higher peak speeds. These trends confirm that later-loop datasets are intrinsically higher-dynamic and higher-difficulty, aligning with the observed PHC+ performance drop.

4.5. Limitations

CLAIMS faces two constraints. First, synthesis quality is bounded by the generative model's capacity on extreme high-dynamics; however, our modular architecture enables seamless integration of future generators. Second, the manually curated variable library lacks objective calibration and comprehensive coverage, necessitating future work on automated, multi-modal libraries with domain knowledge graphs to scale capabilities.

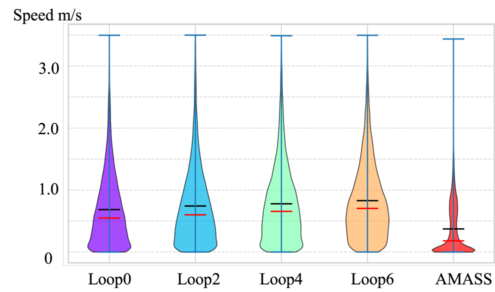


Figure 10. The velocity distribution of AMASS and ours.

Table 5. PHC+ tracking performance on our datasets.A lower success rate indicates that the motions are more challenging for third-party trackers.

Data Loop	SR	g-MPJPE	Acc	Vel
L0	75.3	49.78	5.97	8.54
L1	65.8	53.84	6.66	9.34
L2	65.2	61.29	7.65	10.86
L3	61.2	57.24	7.03	9.95
L4	59.0	57.70	7.08	9.99
L5	52.7	59.10	7.49	10.65
L6	53.6	59.61	7.94	10.97

5. Conclusion

This work targets the core bottlenecks in training data and generalization for physics-based humanoid control. We propose a low-cost, adaptive, closed-loop pipeline for motion data generation and iterative controller optimization, offering a new path to high-dynamics professional motion control. The proposed framework cuts data generation cost, boosts controller generalization, and establishes a data–controller co-evolution paradigm for advancing physics-based humanoid control toward high dynamics and diverse scenarios.

Acknowledgements

This work was supported by the Red Bird MPhil Program at The Hong Kong University of Science and Technology (Guangzhou).

References

[1] 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. 1

[2] Bin Cao, Sipeng Zheng, Ye Wang, Lujie Xia, Qianshan Wei, Qin Jin, Jing Liu, and Zongqing Lu. Being-m0. 5: A real-time controllable vision-language-motion model. *arXiv preprint arXiv:2508.07863*, 2025. 2

[3] Tianxing Chen, Zanxin Chen, Baijun Chen, Zijian Cai, Yibin Liu, Zixuan Li, Qiwei Liang, Xianliang Lin, Yiheng Ge, Zhenyu Gu, et al. Robotwin 2.0: A scalable data generator and benchmark with strong domain randomization for robust bimanual robotic manipulation. *arXiv preprint arXiv:2506.18088*, 2025. 3

[4] Rishabh Dabral, Muhammad Hamza Mughal, Vladislav Golyanik, and Christian Theobalt. Mofusion: A framework for denoising-diffusion-based motion synthesis. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9760–9770, 2023. 2

[5] Pengxiang Ding, Jianfei Ma, Xinyang Tong, Binghong Zou, Xinxin Luo, Yiguo Fan, Ting Wang, Hongchao Lu, Panzhong Mo, Jinxin Liu, et al. Humanoid-vla: Towards universal humanoid control with visual integration. *arXiv preprint arXiv:2502.14795*, 2025. 3

[6] Ke Fan, Shunlin Lu, Minyue Dai, Runyi Yu, Lixing Xiao, Zhiyang Dou, Junting Dong, Lizhuang Ma, and Jingbo Wang. Go to zero: Towards zero-shot motion generation with million-scale data. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 13336–13348, 2025. 2

[7] Biao Jiang, Xin Chen, Wen Liu, Jingyi Yu, Gang Yu, and Tao Chen. Motiongpt: Human motion as a foreign language. *Advances in Neural Information Processing Systems*, 36:20067–20079, 2023. 2

[8] Tan Jing, Shiting Chen, Yangfan Li, Weisheng Xu, and Renjing Xu. Farm: Frame-accelerated augmentation and residual mixture-of-experts for physics-based high-dynamic humanoid control. *arXiv preprint arXiv:2508.19926*, 2025. 1, 3

[9] 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. 3

[10] Manuel Kaufmann, Jie Song, Chen Guo, Kaiyue Shen, Tianjian Jiang, Chengcheng Tang, Juan José Zárate, and Otmar Hilliges. Emdb: The electromagnetic database of global 3d human pose and shape in the wild. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 14632–14643, 2023. 1, 2

[11] Kyungmin Lee, Sibeen Kim, Minho Park, Hyunseung Kim, Dongyoon Hwang, Hojoon Lee, and Jaegul Choo. Phuma: Physically-grounded humanoid locomotion dataset. *arXiv preprint arXiv:2510.26236*, 2025. 2

[12] Ruilong Li, Shan Yang, David A Ross, and Angjoo Kanazawa. Ai choreographer: Music conditioned 3d dance generation with aist++. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 13401–13412, 2021. 1, 2

[13] Yunhao Li, Zhenbo Yu, Yucheng Zhu, Bingbing Ni, Guangtao Zhai, and Wei Shen. Skeleton2humanoid: Animating simulated characters for physically-plausible motion inbetweening. In *Proceedings of the 30th ACM international conference on multimedia*, pages 1493–1502, 2022. 3

[14] Qiayuan Liao, Takara E Truong, Xiaoyu Huang, Guy Tevet, Koushil Sreenath, and C Karen Liu. Beyondmimic: From motion tracking to versatile humanoid control via guided diffusion. *arXiv e-prints*, pages arXiv–2508, 2025. 3

[15] Mingdian Liu, Yilin Liu, Gurunandan Krishnan, Karl S Bayer, and Bing Zhou. T2m-x: Learning expressive text-to-motion generation from partially annotated data. *arXiv preprint arXiv:2409.13251*, 2024. 2

[16] Matthew Loper, Naureen Mahmood, Javier Romero, Gerard Pons-Moll, and Michael J. Black. SMPL: A skinned multi-person linear model. *ACM Trans. Graphics (Proc. SIGGRAPH Asia)*, 34(6):248:1–248:16, 2015. 2

[17] Zhengyi Luo, Ye Yuan, and Kris M Kitani. From universal humanoid control to automatic physically valid character creation. *arXiv preprint arXiv:2206.09286*, 2022. 3

[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. 1, 2, 3, 4, 5

[19] Shipeng Lyu, Fangyuan Wang, Weiwei Lin, Luhao Zhu, David Navarro-Alarcon, and Guodong Guo. Hube: Cross-embodiment human-like behavior execution for humanoid robots. *arXiv preprint arXiv:2508.19002*, 2025. 1, 2

[20] Naureen Mahmood, Nima Ghorbani, Nikolaus F. Troje, Gerard Pons-Moll, and Michael J. Black. AMASS: Archive of motion capture as surface shapes. In *International Conference on Computer Vision, ICCV*, pages 5442–5451, 2019. 2

[21] Jiageng Mao, Siheng Zhao, Siqi Song, Tianheng Shi, Junjie Ye, Mingtong Zhang, Haoran Geng, Jitendra Malik, Vitor Guizilini, and Yue Wang. Learning from massive human videos for universal humanoid pose control. *arXiv preprint arXiv:2412.14172*, 2024. 1, 2

[22] 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. 1, 3

[23] 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 (ToG)*, 40(4):1–20, 2021. 1, 3

[24] 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. 1, 3

[25] Jun Rao, Xuebo Liu, Hexuan Deng, Zepeng Lin, Zixiong Yu, Jiansheng Wei, Xiaojun Meng, and Min Zhang. Dynamic sampling that adapts: Iterative dpo for self-aware mathematical reasoning. *arXiv preprint arXiv:2505.16176*, 2025. 3

[26] Chen Tessler, Yunrong Guo, Ofir Nabati, Gal Chechik, and Xue Bin Peng. Maskedmimic: Unified physics-based character control through masked motion inpainting. *ACM Transactions on Graphics (TOG)*, 43(6):1–21, 2024. 3, 4

[27] Guy Tevet, Sigal Raab, Brian Gordon, Yoni Shafir, Daniel Cohen-or, and Amit Haim Bermano. Human motion diffusion model. In *The Eleventh International Conference on Learning Representations*, 2023. 2, 3, 4

[28] Guy Tevet, Sigal Raab, Setareh Cohan, Daniele Reda, Zhengyi Luo, Xue Bin Peng, Amit H Bermano, and Michiel van de Panne. Closd: Closing the loop between simulation and diffusion for multi-task character control. *arXiv preprint arXiv:2410.03441*, 2024. 3

[29] Andrea Tirinzoni, Ahmed Touati, Jesse Farebrother, Mateusz Guzek, Anssi Kanervisto, Yingchen Xu, Alessandro Lazaric, and Matteo Pirotta. Zero-shot whole-body humanoid control via behavioral foundation models. *arXiv preprint arXiv:2504.11054*, 2025. 3

[30] Takara Everest Truong, Michael Piseno, Zhaoming Xie, and Karen Liu. Pdp: Physics-based character animation via diffusion policy. In *SIGGRAPH Asia 2024 Conference Papers*, pages 1–10, 2024. 3

[31] Yuan Wang, Di Huang, Yaqi Zhang, Wanli Ouyang, Jile Jiao, Xuetao Feng, Yan Zhou, Pengfei Wan, Shixiang Tang, and Dan Xu. Motiongpt-2: A general-purpose motion-language model for motion generation and understanding. *arXiv preprint arXiv:2410.21747*, 2024. 2

[32] Ye Wang, Sipeng Zheng, Bin Cao, Qianshan Wei, Weishuai Zeng, Qin Jin, and Zongqing Lu. Scaling large motion models with million-level human motions. *arXiv preprint arXiv:2410.03311*, 2024. 2

[33] Michael Xu, Yi Shi, KangKang Yin, and Xue Bin Peng. Parc: Physics-based augmentation with reinforcement learning for character controllers. In *Proceedings of the Special Interest Group on Computer Graphics and Interactive Techniques Conference Conference Papers*, page 1–11. ACM, 2025. 3

[34] Yuhong Zhang, Jing Lin, Ailing Zeng, Guanlin Wu, Shunlin Lu, Yurong Fu, Yuanhao Cai, Ruimao Zhang, Haoqian Wang, and Lei Zhang. Motion-x++: A large-scale multimodal 3d whole-body human motion dataset. *arXiv preprint arXiv:2501.05098*, 2025. 1, 2

[35] 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. 3

[36] Adam Zweiger, Jyothish Pari, Han Guo, Ekin Akyürek, Yoon Kim, and Pulkit Agrawal. Self-adapting language models. *arXiv preprint arXiv:2506.10943*, 2025. 3