Any2Any: Efficient Cross-Embodiment Transfer for Humanoid Whole-Body Tracking

Anonymous Authors

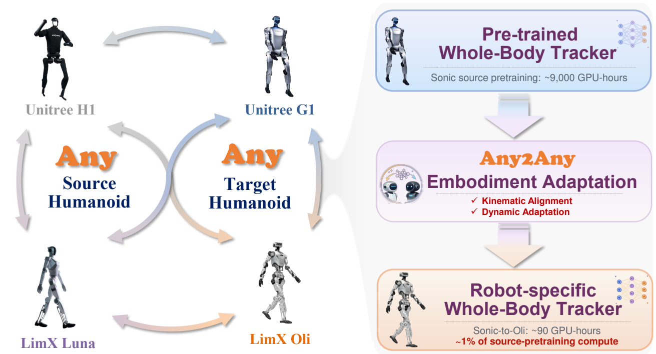

Fig. 1. Overview of ANY2ANY. A pretrained whole-body tracker is transferred between humanoid embodiments through kinematic alignment and localized dynamics adaptation.

***Abstract*— Whole-body tracking (WBT) models have become a key foundation for humanoid robots, enabling them to imitate diverse motions with high fidelity. Training such models from scratch requires large-scale data and computation, making rapid deployment on new humanoid platforms costly. This raises a natural question: Can pretrained WBT models transfer across embodiments with minimal adaptation? To answer this question, we propose ANY2ANY (Fig. 1), a paradigm that efficiently transfers an existing robot-specific WBT policy to a new humanoid embodiment with only a small amount of data and compute. Any2Any first performs kinematic alignment between source and target humanoids, aligning their input and output spaces so that the pretrained source policy can be meaningfully reused on the target embodiment. Any2Any then performs dynamics adaptation, inserting lightweight parameter-efficient fine-tuning (PEFT) components into only the most dynamics-sensitive modules to preserve the source policy's behavioral priors while correcting for the target robot. Extensive experiments on multiple humanoid platforms and pretrained backbones show that Any2Any substantially accelerates convergence and reduces training cost compared with training from scratch, while achieving competitive or superior tracking performance. Notably, Any2Any successfully transfers Sonic models pretrained on Unitree G1 to LimX Oli and LimX Luna, among others, with Sonic-to-Oli adaptation using approximately 90 GPU-hours, about 1% of Sonic's reported 9,000 GPU-hours of source pretraining. These results suggest that pretrained robot-specific WBT policies can be efficiently reused across embodiments, providing a path toward a humanoid *System 0*: a reusable low-level whole-body controller. Code, results, and videos are available on our project page: `https://any2any.top/`.**

I. INTRODUCTION

Behavior foundation models (BFMs) [1], [2] aim to learn reusable motor skills and behavioral priors from large-scale behavior data for downstream adaptation [3], [4]. For humanoids, whole-body tracking (WBT) [5], [6] reproduces diverse full-body reference motions while coordinating legs, torso, arms, and head to maintain balance, and can serve as a motor prior for teleoperation [7], [8], motion imitation [9], [10], and loco-manipulation.

TWIST [8] and GMT [9] extend WBT motion coverage through reinforcement learning, behavior cloning, teacher-student training, adaptive sampling, and mixture-of-experts designs. SONIC [10] treats tracking as a scalable foundation task, growing policy capacity from 1.2M to 42M parameters with over 100M motion frames and 9k GPU hours. HoloMotion [11] trains a whole-body foundation model on extensive motion data; OmniXtreme [12] extends tracking to high-dynamic, extreme motions while retaining fidelity as libraries grow. These advances require large-scale motion data, massively parallel simulation, and substantial compute. Beyond training cost, embodiment dependence [13], [14] remains a fundamental limitation of current large-scale WBT policies. Such policies are tightly coupled to the source robot's morphology and actuator configuration. Consequently, even between similar humanoids, structural and dynamic differences make direct deployment unreliable, while full-policy fine-tuning tends to overwrite the source behavioral prior. Existing cross-embodiment methods mainly address this issue by training universal controllers over many embodiments with morphology randomization or multi-embodiment data [14]–[16], or by scaling robot foundation models with embodiment-aware representations, unified action spaces, post-training, or expert routing [17]–[19]. However, these approaches typically require large-scale multi-embodiment datasets, broad morphology randomization, or expensive large-scale pretraining from scratch, and they mostly address locomotion or manipulation rather than high-fidelity whole-body tracking. In contrast, we study a complementary post-training problem: given a pretrained robot-specific WBT policy, can we efficiently adapt it to another humanoid with limited data and compute?

Humanoid control commonly separates kinematic reference planning from dynamics-aware tracking [20], [21]. Cross-embodiment transfer must likewise resolve kinematic interface incompatibility and adapt to dynamics differences that remain after alignment. This division also appears in two-stream WBT architectures [8], [10]: reference-motion encoders emphasize transferable kinematic representations, whereas action decoders depend more on embodiment-specific dynamics. We thus hypothesize that motor priors such as balance and inter-limb coordination may transfer across embodiments, while embodiment-sensitive modules require adaptation. We therefore insert lightweight PEFT

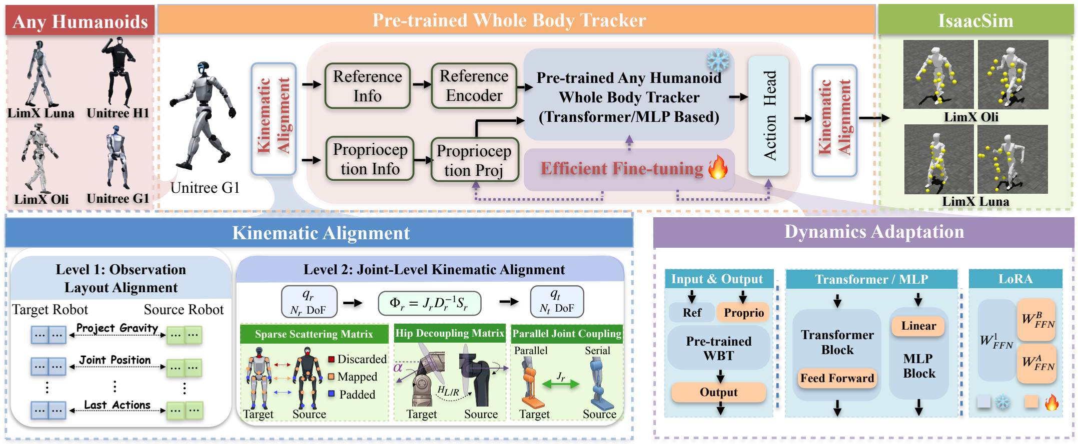

Fig. 2. ANY2ANY: kinematic alignment reconciles target observations and source-policy actions; localized dynamics adaptation specializes a frozen pretrained WBT backbone.

components [22], [23], such as LoRA [24] and adapters [25], only into these modules, freezing the remaining parameters to preserve the source prior. We ask how to resolve kinematic mismatch and which localized module updates enable target adaptation without destroying the source prior. We propose ANY2ANY, a parameter-efficient cross-embodiment WBT post-training framework combining *kinematic alignment* and *dynamics adaptation*. Our contributions are threefold:

1) We demonstrate that humanoid WBT policies pre-trained on a single source robot can be effectively transferred to other humanoid embodiments through principled handling of the kinematic and dynamics mismatches. To the best of our knowledge, this constitutes the first systematic study of cross-embodiment transfer for humanoid WBT.
2) We propose ANY2ANY, a cross-embodiment post-training framework for humanoid WBT that decomposes transfer into two stages: *kinematic alignment*, which unifies source and target policy interfaces, and *dynamics adaptation*, which adapts only the dynamics-sensitive modules while freezing the rest of the pre-trained backbone.
3) We validate Any2Any on six source-target transfer pairs spanning two pretrained WBT policy families and four humanoid platforms, including Sonic-to-Oli adaptation at approximately 1% of Sonic's original source-pretraining compute, and deploy the adapted policies on real hardware across multiple downstream tasks. We further demonstrate aligned eight-robot WBT pre-training and adaptation to Oli, excluded from that pretraining set.

II. Related Work

a) Humanoid Whole-Body Tracking.: Whole-body tracking has developed from reference-motion imitation [26], [27] to real-humanoid teleoperation, general tracking, and long-horizon control [8], [9], [28]. Recent systems extend motion coverage, balance, and dynamic capabilities [12], [29], [30], while Sonic scales policy capacity and motion data toward a reusable motor interface [10]. These policies nevertheless remain tied to robot-specific observations, actions, and dynamics. ANY2ANY asks how an existing specialist's motion prior can be reused on another humanoid, preserving the value of source pretraining while adapting to the target's structural and dynamical differences.

b) Cross-Embodiment Robot Learning.: Multi-embodiment controllers share control knowledge through morphology randomization [15], topology-aware architectures [31], unified interfaces [16], or residual adaptation [14]. Robot foundation models further exploit multi-robot datasets [32], [33], embodiment-aware architectures [17], and unified representations with expert routing [18], [19]. Such approaches typically build generalists using broad robot coverage and substantial pretraining. Our primary setting adapts an already-trained, robot-specific WBT policy to a new humanoid under a limited target-side budget. Section V further evaluates alignment during multi-robot pretraining and subsequent adaptation to an excluded embodiment.

c) Parameter-Efficient Fine-tuning for Robotics.: Adapters, prefix tuning, and LoRA specialize frozen backbones through small trainable components [23]–[25], with growing applications to robotic foundation models [34], [35]. For closed-loop WBT, however, parameter updates alter actions and thereby the subsequent contact, balance, and observation distributions. This makes both adaptation mechanism and placement consequential. We compare PEFT mechanisms and injection scopes to identify which components require target-specific correction and which pretrained motion representations can remain frozen during cross-embodiment transfer.

III. Method

A. Transfer Objective and Policy Backbones

With aligned interfaces and frozen source weights $\theta_{\mathcal{S}}$, ANY2ANY learns $\Delta\theta_{\mathcal{T}}$, with $|\Delta\theta_{\mathcal{T}}| \ll |\theta_{\mathcal{S}}|$. The policy $\pi_{\mathcal{T}} = \pi_{\theta_{\mathcal{S}} \oplus \Delta\theta_{\mathcal{T}}}$ uses $\oplus$ to insert trainable components.

Proximal Policy Optimization (PPO) maximizes

$$\max_{\Delta\theta_{\mathcal{T}}} \mathbb{E}_{\pi_{\mathcal{T}}}\left[\sum_{t=0}^{T_{\mathrm{ep}}} \gamma^t r(s_t, a_t)\right], \tag{1}$$

where r is the tracking reward, γ the discount factor, and T_{ep} the episode horizon. The actor predicts joint-position offsets from reference motion and proprioceptive history: base angular velocity, projected gravity, joint positions and velocities, and previous actions. The critic estimates state values using additional base velocity, contact forces, and body mass distribution, available only during training.

We use three backbones. Our MLP and Transformer policies are pretrained on Oli and collectively termed **Oli-WBT**.

Oli-WBT (MLP). For proprioceptive/reference dimensions d_p, d_r, the MLP flattens $H+1$ timesteps into $(H+1)(d_p+d_r)$ inputs for L layers with GELU activations.

Oli-WBT (Transformer). Each modality is embedded in d dimensions; N causal Transformer decoder blocks model temporal dependencies, and the current-timestep representation predicts the action. Both Oli-WBT critics use the corresponding backbone with privileged inputs.

Sonic. Pretrained on G1, Sonic [10] combines a Robot Motion Encoder and finite scalar quantization (FSQ) bottleneck for motion encoding with a dynamics decoder for action generation.

B. Kinematic Alignment

Alignment has two levels (Fig. 2): observation-layout alignment reorders base states, reference features, proprioception, and action history into the source convention; joint-level alignment handles joint correspondence, inclined hips, and parallel mechanisms, returning actions in target coordinates.

a) Joint correspondence and coordinate conversion: For a target with N_r joints and a source with T joints, a partial injection $\pi_r : \{0,\dots,N_r-1\} \rightharpoonup \{0,\dots,T-1\}$ defines a sparse matrix $S_r \in \{0,1\}^{T\times N_r}$ through $(S_r)_{ij} = \mathbf{1}[\pi_r(j)=i]$. Correspondence follows joint semantics: (i) shared joints use one-to-one mappings with unit coefficients; (ii) source-only joints receive zero-padded inputs, and their predicted actions are discarded; (iii) target-only joints are excluded from source inputs and held at their default positions with zero action offsets. The source network's dimensions remain unchanged. These rules define joint correspondence; inclined axes and closed-chain coupling additionally require the coordinate conversions below.

The hip-convention matrix $D_r \in \mathbb{R}^{T\times T}$ is the identity except for its left/right hip blocks,

$$H_{\mathrm{L/R}} = \begin{pmatrix} \cos\alpha & 0 \\ \mp\sin\alpha & 1 \end{pmatrix}, \tag{2}$$

where α is the hip-axis inclination difference between the target and source robots; the signs account for mirrored left/right axes.

b) Parallel-mechanism conversion: For joints with parallel actuation, such as the ankles and waist of some humanoids, we use a calibrated linear approximation between equivalent serial-joint and motor coordinates. For the two-motor pitch–roll pairs considered here, serial coordinates (p,ρ) and motor coordinates (A,B) are related by

$$A = c_p p + s c_r \rho, \qquad B = c_p p - s c_r \rho. \tag{3}$$

The nonzero coefficients c_p, c_r are obtained by least-squares fitting to joint scans for each mechanism; $s \in \{-1,+1\}$ accounts for mirrored roll conventions. The motor-to-serial block for each coupled pair b is

$$J_b = \frac{1}{2}\begin{pmatrix} c_p^{-1} & c_p^{-1} \\ s c_r^{-1} & -s c_r^{-1} \end{pmatrix}. \tag{4}$$

The full fitted matrix $J_r \in \mathbb{R}^{T\times T}$ places each J_b at the corresponding paired indices in the source layout, with identity on all remaining coordinates. Its inverse uses J_b^{-1} at the same indices, recovering the serial-to-motor relation in Eq. 3. Interfaces already exposing serial coordinates retain identity blocks. The complete interface maps are

$$\Phi_r = J_r D_r^{-1} S_r, \qquad \Phi_r^{+} = S_r^{\top} D_r J_r^{-1}. \tag{5}$$

Joint-structured observations, reference features, and action history use $\widetilde{\mathbf{q}}_r = \Phi_r \mathbf{q}_r$; source-space actions return through $\mathbf{a}_r = \Phi_r^{+}\widetilde{\mathbf{a}}$. For invertible D_r, J_r, $\Phi_r^{+}\Phi_r = S_r^{\top} S_r$ masks unmatched joints, equaling I_{N_r} only when all target joints are matched. Alignment unifies policy interfaces without exact kinematic equivalence or recovering missing DoF; subsequent adaptation can compensate for residual alignment errors and dynamics differences.

C. Dynamics Adaptation

After alignment, dynamics still differ. For the same reference motion in aligned joint coordinates q, the target-minus-source rigid-body force residual is

$$\Delta\tau(q,\dot q,\ddot q) = \Delta M(q)\ddot q + \Delta C(q,\dot q)\dot q + \Delta G(q), \tag{6}$$

Here $\Delta M = M_{\mathcal{T}} - M_{\mathcal{S}}$, with analogous definitions for ΔC and ΔG. The terms M, C, and G denote inertia, Coriolis/centrifugal coupling, and gravity loading; actuation, friction, and contact add residuals. Per-link inertial differences scale with link count rather than backbone size, motivating a compact correction for related humanoids. This does not prove low rank in weight space; LoRA learns through closed-loop tracking returns, without torque-residual supervision.

For an adapted linear layer, LoRA [24] learns

$$W' = W + BA, \quad A \in \mathbb{R}^{k\times d_{\mathrm{in}}},\ B \in \mathbb{R}^{d_{\mathrm{out}}\times k}. \tag{7}$$

Training only A, B adds $k(d_{\mathrm{in}} + d_{\mathrm{out}})$ parameters per layer while sharing the frozen W; the total depends on rank and injection scope. Oli-WBT adapts the actor backbone, proprioception and action-output projections, and critic backbone, retaining the reference projection. Sonic adapts its dynamics decoder and critic while freezing the Robot Motion Encoder and FSQ bottleneck.

Proprioceptive processing interprets target feedback and action history; the action decoder generates commands for target dynamics. The frozen reference projection or motion encoder reuses motion features, while critic adaptation supports target-specific value estimation. Our choice of LoRA is supported by the cost–performance comparison in Table IV, while the injection-scope ablation in Fig. 7 supports the selected configuration for the Oli-WBT Transformer.

D. Adaptation Procedure

Target PPO rollouts use aligned observation layouts and source-coordinate joint-structured reference, proprioceptive, and history features. The actor predicts source-space joint-position offsets, which the reverse map converts into target commands. Target transitions and rewards update only the inserted actor and critic LoRA factors; source weights and alignment remain fixed. Deployment uses the same actor and mappings, without privileged observations or the critic.

IV. Experiments

A. Experimental Setup

a) Robots and source policies: The four humanoids span 19–31 body DoF and serial/parallel mechanisms (Table I); G1 excludes hand DoF, and dimensions are nominal manufacturer values. Oli/Luna hip inclinations are $25°/30°$; G1/H1 hip axes are orthogonal. Oli's low-level interface resolves parallel ankle/waist mappings into serial WBT coordinates, whereas Luna requires the parallel-mechanism correction. Luna's fitted (c_p, c_r) are $(1.0028, 0.4885)$ for both ankles ($s = +1/-1$ for left/right) and $(-1.0027, 0.8808)$ for the waist ($s = +1$).

TABLE I

Benchmark robot specifications [36]–[39].

Robot	DoF	WBT model	Height	Mass
Oli	31	Serial	1.65 m	55 kg
Luna	27	Parallel	1.60 m	54 kg
G1	29	Serial	1.32 m	35 kg
H1	19	Serial	1.80 m	47 kg

Oli-WBT uses MLP/Transformer policies pretrained on Oli with over 500 motion hours; Sonic [10] is pretrained on G1. We transfer Oli-WBT to Luna/G1/H1 and Sonic to Oli/Luna/H1, using four A100 GPUs for about 22.5 hours on average (90 GPU-hours) per target adaptation. LoRA ranks are $k = 32$ for Oli-WBT and $k = 64$ for Sonic.

b) Data, training, and baselines: Our evaluation addresses two complementary questions: whether reusing a pretrained WBT policy improves cross-embodiment learning relative to training from scratch, and which adaptation design offers a favorable balance of tracking performance and training cost. We address the first through six source–target transfers comparing Any2Any with Scratch, and the second through adaptation-mechanism and injection-scope ablations on Oli-WBT-to-Luna.

AMASS [40] adaptation motions are retargeted by General Motion Retargeting [41]. Held-out evaluation uses 500 OMOMO manipulation-motion clips [42] and LAFAN locomotion clips [43], excluding near-ground, crawling, and corrupted retargeting. Ablations compare full fine-tuning with/without alignment and aligned LoRA, Adapter, and Prefix-Tuning. PPO in Isaac Lab retains source interfaces, rewards, hyperparameters, motion sampling, and domain randomization. Comparisons share architectures, environments, retargeting, target budgets, and iteration counts; the compute sweep varies GPU count at fixed iterations.

c) Metrics: Training curves report rewards; MuJoCo evaluation reports clip completion (%) and tracking errors (Table II). For both policy families, mean per-joint position error (MPJPE, mm) averages Euclidean distances over all corresponding robot/reference bodies in their respective root frames. Base position is the world-frame XYZ Euclidean distance between roots (cm). Base orientation is the full rotation angle $2\arccos(|u^\top v|)$ for unit root quaternions u, v, converted to degrees. A clip fails when absolute root or any of the four end-effector height errors exceed 0.25 or 0.30 m, respectively, or squared root rotation-angle error exceeds $1\,\mathrm{rad}^2$.

B. Cross-Embodiment Transfer

a) Sonic transfers: Any2Any converges faster with higher final reward across Oli/Luna/H1 (Fig. 3, left), outperforming Scratch in success, MPJPE, and base-orientation accuracy in all six held-out conditions (Table II(a)). On Oli/LAFAN, success rises from 83.1% to 93.1% and MPJPE drops 28.1% (56.5 to 40.6 mm). Base position improves except on H1.

b) Oli-WBT transfers: Table II(b) reports Oli-WBT Transformer results. Any2Any also accelerates Oli-WBT convergence with comparable or higher final reward (Fig. 3, right). It reduces base-position/orientation errors in all six conditions and MPJPE in five (Table II(b)). On Luna/OMOMO, these base errors fall by 12.5%/24.2%, respectively. Small completion-rate and Luna/LAFAN MPJPE trade-offs accompany these broad gains, supporting faster adaptation and better overall tracking. Oli's actuation limits restrict its WBT policy's reliable motion repertoire, potentially contributing to larger tracking errors on more dynamic LAFAN motions after transfer.

C. Real-World Deployment

The original Sonic policy on physical G1 and adapted Any2Any on physical Oli both complete the same 15 motions (Fig. 4). Their real-world MPJPEs are close: 30.06 and 32.96 mm, respectively (Table III), indicating comparable positional tracking accuracy on these motions.

D. Data and Compute Efficiency

At fixed iterations, Any2Any outperforms Scratch in held-out success and MPJPE at all tested AMASS fractions (25/50/100%; Fig. 5, left), although individual metrics are not monotonic with data size.

We compare held-out quality at matched target-side budgets of 20/60/90 GPU-hours (GPU count × wall-clock

TABLE II

CROSS-EMBODIMENT TRACKING RESULTS.

(a) Sonic (G1)					**OMOMO**
Target	Policy	Success ↑ (%)	MPJPE ↓ (mm)	Base pos. ↓ (cm)	Base ori. ↓ (°)
Oli	Scratch	96.6	52.8	30.51	7.0
	Any2Any	**99.6**	**43.2**	**26.53**	**5.9**
Luna	Scratch	74.4	116.4	43.15	14.5
	Any2Any	**80.8**	**104.0**	**39.95**	**10.7**
H1	Scratch	67.2	59.0	**36.82**	7.9
	Any2Any	**77.6**	**55.2**	39.63	**7.4**
					LAFAN
Oli	Scratch	83.1	56.5	69.67	7.3
	Any2Any	**93.1**	**40.6**	**55.17**	**5.6**
Luna	Scratch	88.1	96.5	77.02	11.8
	Any2Any	**91.9**	**80.4**	**62.87**	**8.5**
H1	Scratch	72.0	56.3	**81.48**	7.6
	Any2Any	**85.4**	**48.9**	83.22	**6.6**
(b) Oli-WBT (Oli)					**OMOMO**
Target	Policy	Success ↑ (%)	MPJPE ↓ (mm)	Base pos. ↓ (cm)	Base ori. ↓ (°)
Luna	Scratch	**99.78**	59.68	48.68	21.11
	Any2Any	98.66	**54.12**	**42.61**	**16.00**
G1	Scratch	**99.60**	66.99	47.27	27.57
	Any2Any	99.40	**60.92**	**40.80**	**22.29**
H1	Scratch	90.08	68.19	53.76	27.38
	Any2Any	**94.62**	**64.98**	**50.17**	**26.85**
					LAFAN
Luna	Scratch	98.83	**66.11**	85.79	27.44
	Any2Any	**98.83**	67.14	**78.19**	**24.42**
G1	Scratch	80.77	84.67	100.59	35.81
	Any2Any	**83.44**	**74.26**	**98.67**	**34.86**
H1	Scratch	88.22	80.97	102.49	35.16
	Any2Any	**90.46**	**78.03**	**99.34**	**34.80**

hours), using one/two/four A100 GPUs for 8,000 iterations (Fig. 5, right). ANY2ANY achieves higher success and lower MPJPE than Scratch within each budget. With fixed environments per GPU, more GPUs yield more transitions despite equal iterations. The 90/9,000 GPU-hour ratio compares target adaptation with Sonic source pretraining [10]; it does not measure equal-performance speedup over Scratch.

E. Ablation Studies

For Oli-WBT-to-Luna, we ablate alignment and adaptation with MLP and Transformer backbones, and injection locations with Transformer.

a) Alignment and adaptation mechanism: Alignment improves full fine-tuning (Fig. 6, left). Table IV compares Any2Any@Full FT and Any2Any@LoRA, two implementations under identical alignment. LoRA reduces trainable parameters, PPO iteration time (collection plus learning), and MuJoCo errors on the same 15 real-world motions (Sec. IV-C). We adopt LoRA to lower training cost while limiting changes to pretrained dynamics priors. Adapter is slower with lower rewards; Prefix-Tuning is less stable (Fig. 6, right).

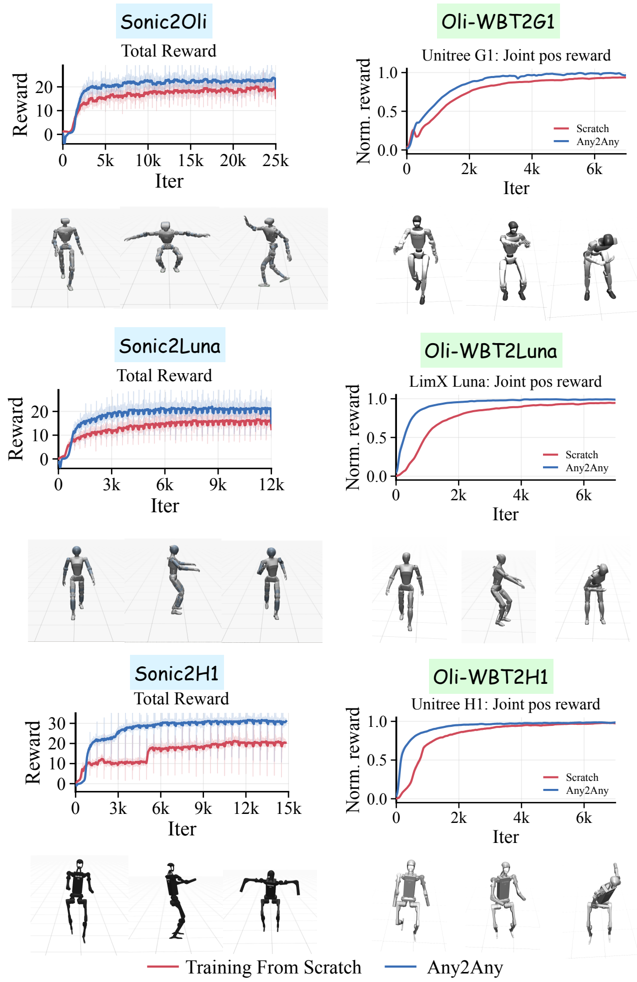

Fig. 3. Cross-embodiment transfer with Sonic (left) and Oli-WBT Transformer (right). Training curves compare total reward for Sonic and normalized joint-position tracking reward for Oli-WBT against Scratch; snapshots show simulated rollouts on each target.

TABLE IV

ALIGNED ANY2ANY VARIANTS ON OLI-WBT-TO-LUNA.

	Training cost and reward				MuJoCo evaluation		
Method	Trainable (%) ↓	Iteration time (s) ↓	Joint ↑ reward	Total ↑ reward	MPJPE (mm) ↓	Base pos. (cm) ↓	Base ori. (°) ↓
Any2Any@ Full FT	100.00	8.38	0.43	19.04	36.76	41.05	18.67
Any2Any@ LoRA	**5.26**	**7.60**	**0.54**	**23.86**	**33.44**	**31.66**	**16.53**

b) Injection locations: Table V defines all nine scopes in Fig. 7. Backbone denotes the main actor/critic network. Both actor and critic have input and output projections. Actor inputs are split into Ref. (reference) and Prop. (proprioception); the critic input is labeled Input. Output produces actions for the actor and state values for the critic. Checks mark adapted modules; unmarked modules are frozen. S7 gives the best observed trade-off by adapting proprioception, action output, and actor/critic backbones while retaining

Fig. 4. Real-world deployment of the adapted Sonic-to-Oli policy after 90 GPU-hours of AMASS adaptation. Snapshots show squatting, forward and backward walking, turning, bowing, running, manipulation, and dancing.

TABLE III

SIMULATION VS. REAL WORLD ON SELECTED MOTIONS.

Setting		Success ↑ (%)	MPJPE ↓ (mm)	Base ori. ↓ (°)
Any2Any (Oli)	Sim.	100.00	26.41	4.04
	Real	100.00	32.96	4.56
Sonic (G1)	Sim.	100.00	28.91	2.66
	Real	100.00	30.06	2.79

the reference projection. Critic adaptation supports target value estimation. This result does not establish a universally optimal scope, and frozen weights alone do not guarantee behavioral preservation.

V. DISCUSSION AND CONCLUSION

ANY2ANY combines kinematic alignment and localized low-rank adaptation, accelerating six transfers across four humanoids and improving most held-out tracking metrics, with condition-specific trade-offs and physical Oli deployment. Adaptation costs approximately 90 GPU-hours, about 1% of Sonic's original source-pretraining compute; the reported Oli-WBT-to-Luna comparison updates 5.26% of parameters.

To our knowledge, this is the first evidence that large-scale WBT priors need not remain bound to their training embodiment. One explanation is a shared human-centric motion structure: the data, reference behaviors, and robot designs derive from human movement and anatomy. Balance, inter-limb coordination, timing, and contact-aware tracking may remain reusable, while adaptation addresses dynamics differences and residual alignment errors. This suggests reusing a WBT backbone for motion understanding and coordination, with a kinematic mapping and small dynamics adapter per target.

Looking forward, two complementary directions follow. *First*, a community *Reference Humanoid* could provide a standardized, human-like transfer platform and common benchmark for whole-body motion, loco-manipulation, and dexterous manipulation.

Second, pretraining can cover diverse robot configurations: ANY2ANY's kinematic alignment (K.A.) enables backbone sharing, with separate adaptation modules for different embodiments. Both reference-embodiment and multi-robot foundation models can then support WBT on new robots through limited target-specific motion data, K.A., and lightweight LoRA fine-tuning.

Our Sonic/Oli-WBT transfers are close to the first setting. To explore the second, we train a Sonic-based WBT policy across eight robots using LAFAN motions [43] (Fig. 8(a)). First, K.A. improves normalized reward and mean episode length over unaligned pretraining at equal PPO iterations (Fig. 8(b)), showing its value for constructing shared WBT foundation models. Second, ANY2ANY reuses this prior to adapt to Oli, excluded from pretraining, achieving higher reward and longer episodes than Scratch at equal iterations, with a clearer episode-length advantage (Fig. 8(c)). These results support efficient specialization of multi-robot priors and provide initial evidence toward a humanoid *System 0*: a reusable low-level whole-body controller.

REFERENCES

[1] M. Pirotta, A. Tirinzoni, A. Touati, A. Lazaric, and Y. Ollivier, "Fast imitation via behavior foundation models," in *International Conference on Learning Representations*, vol. 2024, 2024, pp. 12 685–12 724.

[2] M. Yuan, T. Yu, W. Ge, X. Yao, D. Li, H. Wang, J. Chen, B. Li, W. Zhang, W. Zeng *et al.*, "A survey of behavior foundation model: Next-generation whole-body control system of humanoid robots," *IEEE transactions on pattern analysis and machine intelligence*, 2025.

[3] E. Cetin, A. Touati, and Y. Ollivier, "Finer behavioral foundation models via auto-regressive features and advantage weighting," *arXiv preprint arXiv:2412.04368*, 2024.

[4] Y. Li, Z. Luo, T. Zhang, C. Dai, A. Kanervisto, A. Tirinzoni, H. Weng, K. Kitani, M. Guzek, A. Touati *et al.*, "Bfm-zero: A promptable behavioral foundation model for humanoid control using unsupervised reinforcement learning," *arXiv preprint arXiv:2511.04131*, 2025.

[5] Z. Gu, J. Li, W. Shen, W. Yu, Z. Xie, S. McCrory, X. Cheng, A. Shamsah, R. Griffin, C. K. Liu *et al.*, "Humanoid locomotion and manipulation: Current progress and challenges in control, planning, and learning," *IEEE/ASME Transactions on Mechatronics*, vol. 31, no. 2, pp. 2300–2330, 2026.

[6] X. Cheng, Y. Ji, J. Chen, R. Yang, G. Yang, and X. Wang, "Expressive whole-body control for humanoid robots," *arXiv preprint arXiv:2402.16796*, 2024.

[7] T. He, Z. Luo, X. He, W. Xiao, C. Zhang, W. Zhang, K. Kitani, C. Liu, and G. Shi, "Omnih2o: Universal and dexterous human-to-humanoid whole-body teleoperation and learning," *arXiv preprint arXiv:2406.08858*, 2024.

[8] Y. Ze, Z. Chen, J. P. Araújo, Z.-a. Cao, X. B. Peng, J. Wu, and C. K. Liu, "Twist: Teleoperated whole-body imitation system," *arXiv preprint arXiv:2505.02833*, 2025.

[9] Z. Chen, M. Ji, X. Cheng, X. Peng, X. B. Peng, and X. Wang, "Gmt: General motion tracking for humanoid whole-body control," *arXiv preprint arXiv:2506.14770*, 2025.

[10] Z. Luo, Y. Yuan, T. Wang, C. Li, S. Chen, F. Castaneda, Z.-A. Cao, J. Li, D. Minor, Q. Ben *et al.*, "Sonic: Supersizing motion tracking for natural humanoid whole-body control," *arXiv preprint arXiv:2511.07820v1*, 2025.

[11] M. Chen, K. Wang, B. Zhang, X. Ma, Z. Yang, Y. Ren, Q. Huang, Z. Zhu, Y. Wang, and Z. Su, "Holomotion-1 technical report," *arXiv preprint arXiv:2605.15336*, 2026.

[12] Y. Wang, S. Zhu, P. Zhi, Y. Li, J. Li, Y.-L. Li, Y. Xiao, X. Wang, B. Jia, and S. Huang, "Omnixtreme: Breaking the generality barrier in high-dynamic humanoid control," *arXiv preprint arXiv:2602.23843*, 2026.

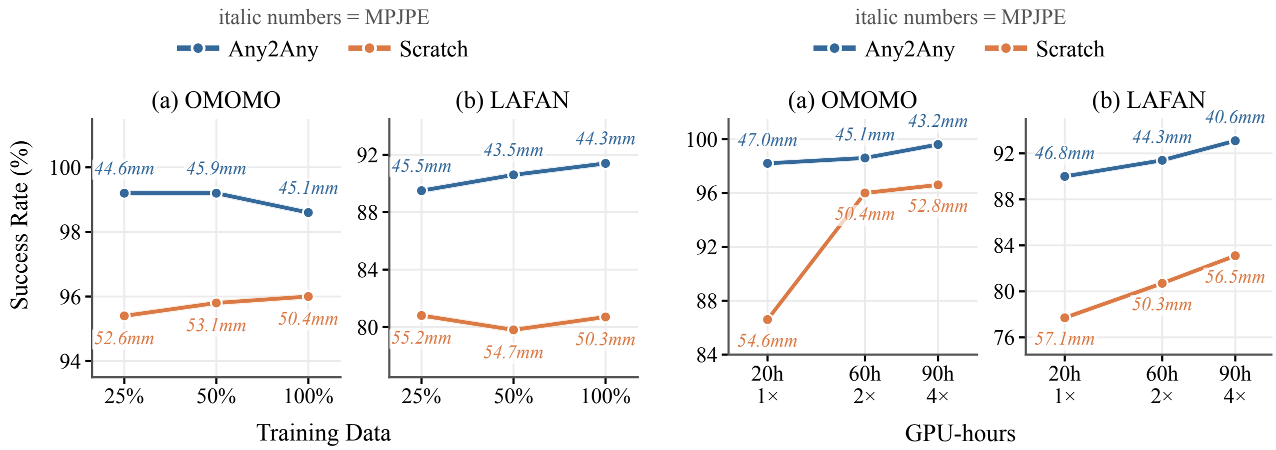

Fig. 5. Sonic-to-Oli efficiency. Left: held-out success and annotated MPJPE at 25%, 50%, and 100% of the adaptation data. Right: the same metrics under one-, two-, and four-GPU settings, labeled 20, 60, and 90 GPU-hours in the figure.

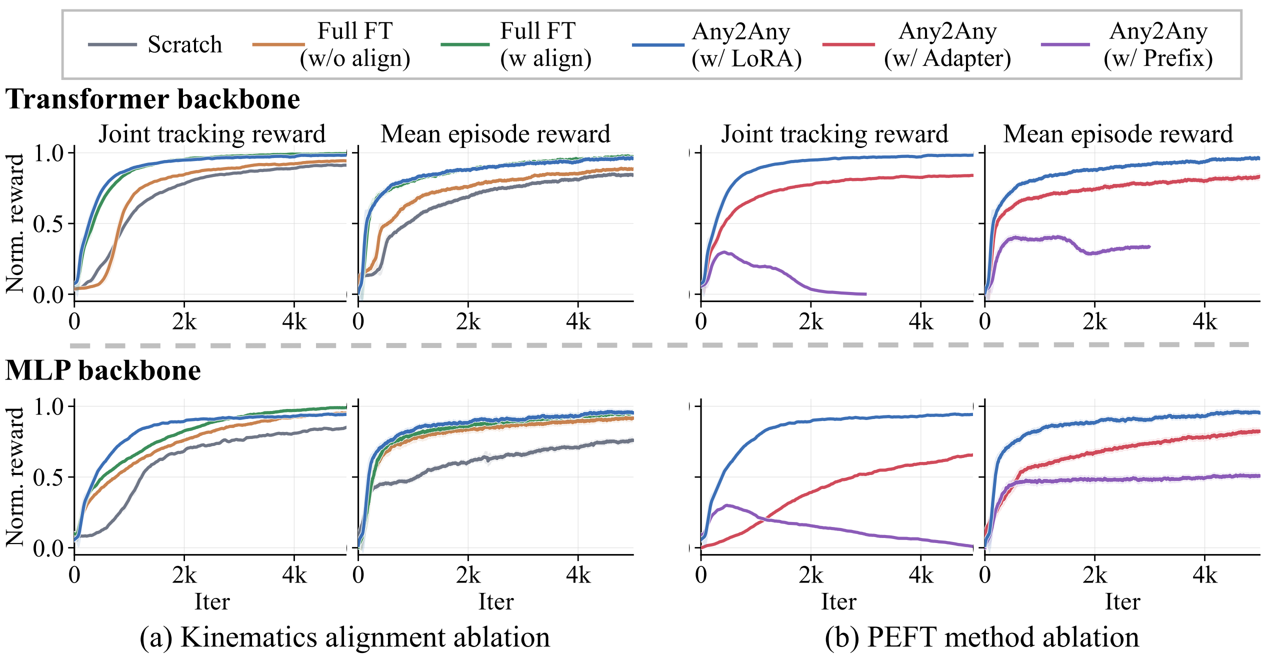

Fig. 6. Oli-WBT-to-Luna ablations with Transformer and MLP backbones. Left: kinematic alignment and adaptation. Right: LoRA, Adapter, and Prefix-Tuning under alignment.

[13] A. Gupta, L. Fan, S. Ganguli, and L. Fei-Fei, "Metamorph: Learning universal controllers with transformers," *arXiv preprint arXiv:2203.11931*, 2022.

[14] S. Yang, Z. Fu, Z. Cao, J. Guo, P. Wensing, W. Zhang, and H. Chen, "Multi-loco: Unifying multi-embodiment legged locomotion via reinforcement learning augmented diffusion," *arXiv preprint arXiv:2506.11470*, 2025.

[15] Y. Xue, Y. Lin, W. Dong, Y. Tang, J. Wang, J. Pang, M. Zhou, M. Liu, and W. Zhang, "Scalable and general whole-body control for cross-humanoid locomotion," *arXiv preprint arXiv:2602.05791*, 2026.

[16] M. Liu, D. Pathak, and A. Agarwal, "Locoformer: Generalist locomotion via long-context adaptation," in *Proceedings of The 9th Conference on Robot Learning*, 2025.

[17] J. Bjorck, F. Castañeda, N. Cherniadev, X. Da, R. Ding, L. Fan, Y. Fang, D. Fox, F. Hu, S. Huang *et al.*, "Gr00t n1: An open foundation model for generalist humanoid robots," *arXiv preprint arXiv:2503.14734*, 2025.

[18] H. Luo, Y. Wang, W. Zhang, S. Zheng, Z. Xi, C. Xu, H. Xu, H. Yuan, C. Zhang, Y. Wang *et al.*, "Being-h0.5: Scaling human-centric robot learning for cross-embodiment generalization," *arXiv preprint arXiv:2601.12993*, 2026.

[19] S. Bai, M. Li, X. Lv, J. Wang, X. Wang, F. Liao, C. Hou, L. Gu, W. Zhou, K. Wu *et al.*, "Hex: Humanoid-aligned experts for cross-embodiment whole-body manipulation," *arXiv preprint arXiv:2604.07993*, 2026.

[20] D. Kim, J. Lee, J. Ahn, O. Campbell, H. Hwang, and L. Sentis, "Computationally-robust and efficient prioritized whole-body controller with contact constraints," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 1–8.

[21] M. Chignoli, D. Kim, E. Stanger-Jones, and S. Kim, "The mit humanoid robot: Design, motion planning, and control for acrobatic behaviors," in *2020 IEEE-RAS 20th International Conference on Humanoid Robots (Humanoids)*. IEEE, 2021, pp. 1–8.

[22] N. Ding, Y. Qin, G. Yang, F. Wei, Z. Yang, Y. Su, S. Hu, Y. Chen, C.-M. Chan, W. Chen *et al.*, "Parameter-efficient fine-tuning of large-scale pre-trained language models," *Nature Machine Intelligence*, vol. 5, no. 3, pp. 220–235, 2023. [Online]. Available:

TABLE V

COMPLETE LORA INJECTION SCOPES.

	Actor				Critic		
	Backbone	Ref.	Prop.	Output	Backbone	Input	Output
S1	✓						
S2	✓				✓		
S3	✓	✓			✓		
S4	✓		✓		✓		
S5	✓			✓	✓		
S6	✓				✓	✓	
S7	✓		✓	✓	✓		
S8	✓	✓	✓	✓			
S9	✓	✓	✓	✓	✓	✓	✓

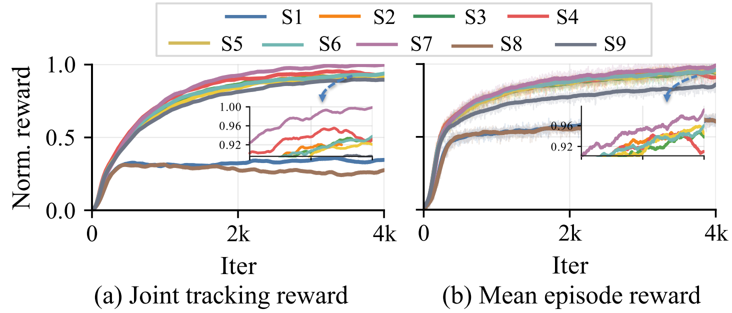

Fig. 7. LoRA injection-scope ablation for Oli-WBT Transformer transfer to Luna. All nine configurations are defined in Table V; S7 is the selected configuration.

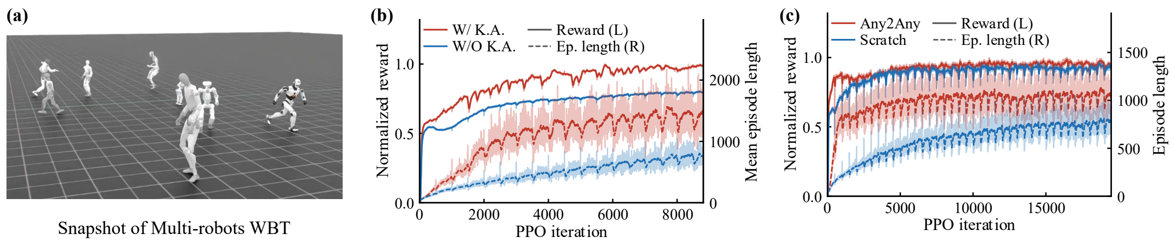

Fig. 8. Multi-robot WBT and transfer. (a) Simulated multi-robot tracking. (b) Eight-robot pretraining on LAFAN with and without kinematic alignment (K.A.). (c) Any2Any adaptation to Oli, excluded from pretraining, versus Scratch. In (b,c), solid curves show normalized reward (left axes); dashed curves show mean episode length (right axes). Comparisons use equal PPO iteration counts within each experiment.

https://www.nature.com/articles/s42256-023-00626-4

[23] X. L. Li and P. Liang, "Prefix-tuning: Optimizing continuous prompts for generation," in *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics*, 2021, pp. 4582–4597. [Online]. Available: https://aclanthology.org/2021.acl-long.353/

[24] E. J. Hu, Y. Shen, P. Wallis, Z. Allen-Zhu, Y. Li, S. Wang, L. Wang, and W. Chen, "Lora: Low-rank adaptation of large language models," in *International Conference on Learning Representations*, 2022. [Online]. Available: https://openreview.net/forum?id=nZeVKeeFYf9

[25] N. Houlsby, A. Giurgiu, S. Jastrzebski, B. Morrone, Q. de Laroussilhe, A. Gesmundo, M. Attariyan, and S. Gelly, "Parameter-efficient transfer learning for nlp," in *International Conference on Machine Learning*, 2019, pp. 2790–2799. [Online]. Available: https://arxiv.org/abs/1902.00751

[26] X. B. Peng, P. Abbeel, S. Levine, and M. van de Panne, "Deepmimic: Example-guided deep reinforcement learning of physics-based character skills," *ACM Transactions on Graphics*, vol. 37, no. 4, pp. 1–14, 2018. [Online]. Available: https://arxiv.org/abs/1804.02717

[27] Z. Luo, J. Cao, A. Winkler, K. Kitani, and W. Xu, "Perpetual humanoid control for real-time simulated avatars," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023. [Online]. Available: https://openaccess.thecvf.com/content/ICCV2023/html/Luo_Perpetual_Humanoid_Control_for_Real-time_Simulated_Avatars_ICCV_2023_paper.html

[28] Y. Li, Y. Lin, J. Cui, T. Liu, W. Liang, Y. Zhu, and S. Huang, "Clone: Closed-loop whole-body humanoid teleoperation for long-horizon tasks," in *Proceedings of The 9th Conference on Robot Learning*, 2025. [Online]. Available: https://arxiv.org/abs/2506.08931

[29] Y. Pan, R. Qiao, L. Chen, K. Chitta, L. Pan, H. Mai, Q. Bu, H. Zhao, C. Zheng, P. Luo *et al.*, "Agility meets stability: Versatile humanoid control with heterogeneous data," *arXiv preprint arXiv:2511.17373*, 2025.

[30] Z. Sun, B.-S. Huang, Y. Peng, X. Li, J. Ma, Y. Sun, Z. Li, H. Jiang, B. Gao, Z. Bing *et al.*, "Mosaic: Bridging the sim-to-real gap in generalist humanoid motion tracking and teleoperation with rapid residual adaptation," *arXiv preprint arXiv:2602.08594*, 2026.

[31] Y. Lin, M. Liu, Y. Xue, M. Zhou, Y. Yu, J. Pang, and W. Zhang, "H-zero: Cross-humanoid locomotion pretraining enables few-shot novel embodiment transfer," *arXiv preprint arXiv:2512.00971*, 2025. [Online]. Available: https://arxiv.org/abs/2512.00971

[32] Open X-Embodiment Collaboration, A. O'Neill, A. Rehman, A. Gupta, A. Maddukuri, A. Gupta, A. Padalkar, A. Lee, A. Pooley, A. Gupta *et al.*, "Open x-embodiment: Robotic learning datasets and rt-x models," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, 2024, pp. 6892–6903. [Online]. Available: https://arxiv.org/abs/2310.08864

[33] Octo Model Team, D. Ghosh, H. Walke, K. Pertsch, K. Black, O. Mees, S. Dasari, J. Hejna, T. Kreiman, C. Xu, J. Luo, Y. L. Tan, L. Y. Chen, P. Sanketi, Q. Vuong, T. Xiao, D. Sadigh, C. Finn, and S. Levine, "Octo: An open-source generalist robot policy," in *Proceedings of Robotics: Science and Systems (RSS)*, 2024. [Online]. Available: https://arxiv.org/abs/2405.12213

[34] M. J. Kim, C. Finn, and P. Liang, "Fine-tuning vision-language-action models: Optimizing speed and success," 2025.

[35] Y. Wang, P. Ding, L. Li, C. Cui, Z. Ge, X. Tong, W. Song, H. Zhao, W. Zhao, P. Hou, S. Huang, Y. Tang, W. Wang, R. Zhang, J. Liu, and D. Wang, "Vla-adapter: An effective paradigm for tiny-scale vision-language-action model," 2025.

[36] LimX Dynamics, "LimX Oli: Full-Size General-Purpose Humanoid Robot," https://www.limxdynamics.com/en/products/oli, 2025, accessed: 2026-05-22.

[37] ——, "LimX Luna Humanoid Robot," https://x.com/LimX_Dynamics, 2026, official product page not yet publicly available at the time of access; accessed: 2026-05-22.

[38] Unitree Robotics, "Unitree G1 Humanoid Robot," https://www.unitree.com/g1, 2024, accessed: 2026-05-22.

[39] ——, "Unitree H1 Universal Humanoid Robot," https://www.unitree.com/h1, 2023, accessed: 2026-05-22.

[40] N. Mahmood, N. Ghorbani, N. F. Troje, G. Pons-Moll, and M. J. Black, "Amass: Archive of motion capture as surface shapes," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 5442–5451.

[41] J. P. Araujo, Y. Ze, P. Xu, J. Wu, and C. K. Liu, "Retargeting matters: General motion retargeting for humanoid motion tracking," *arXiv preprint arXiv:2510.02252*, 2025.

[42] J. Li, J. Wu, and C. K. Liu, "Object motion guided human motion synthesis," in *ACM Transactions on Graphics (SIGGRAPH Asia)*, 2023.

[43] F. G. Harvey, M. Yurick, D. Nowrouzezahrai, and C. Pal, "Robust motion in-betweening," *ACM Transactions on Graphics (SIGGRAPH)*, vol. 39, no. 4, 2020.