RoboDuet

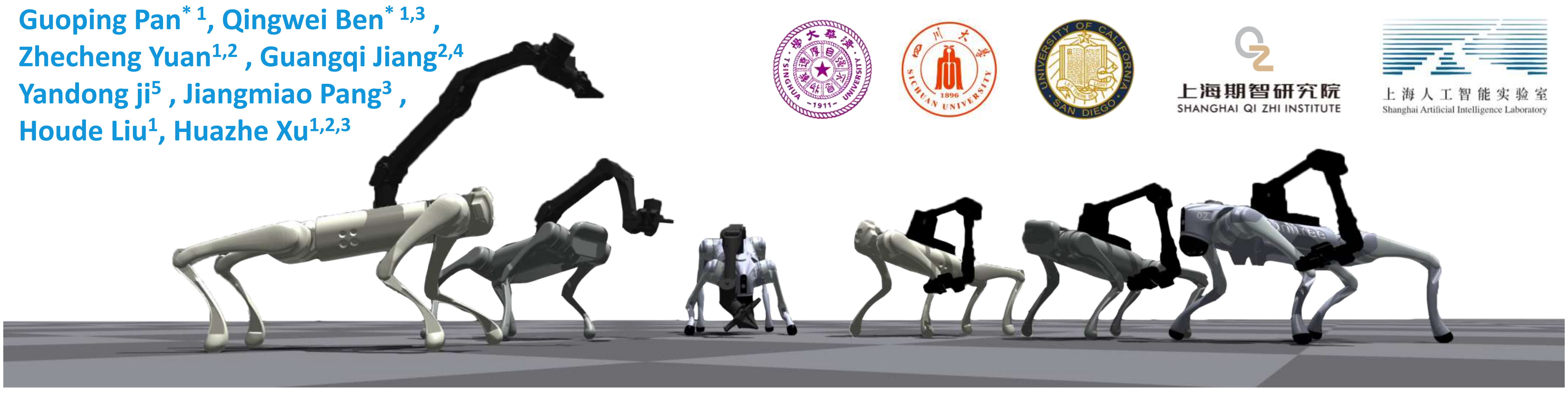
A Framework Affording Mobile-Manipulation and Cross-Embodiment

Guoping Pan* [1], Qingwei Ben* [1,3], Zhecheng Yuan[1,2], Guangqi Jiang[2,4], Yandong ji[5], Jiangmiao Pang[3], Houde Liu[1], Huazhe Xu[1,2,3]

上海期智研究院 SHANGHAI QI ZHI INSTITUTE

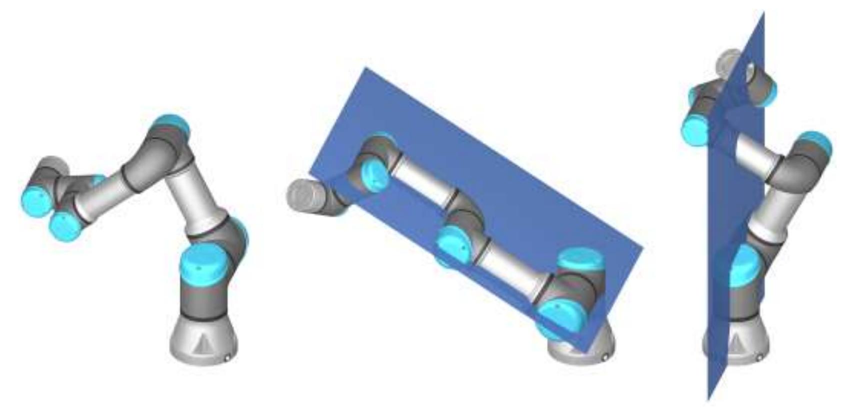
上海人工智能实验室 Shanghai Artificial Intelligence Laboratory

ICRA2024 YOKOHAMA | JAPAN



Challenges

- Classic control strategies for dexterous arm manipulation are susceptible to singularity issues of inverse kinematics (IK) solution.
- Whole Body Control (WBC) offers expanded workspace capabilities but often demands substantial expertise.
- Learning-based controllers often require complete retraining when confronted with different robots.



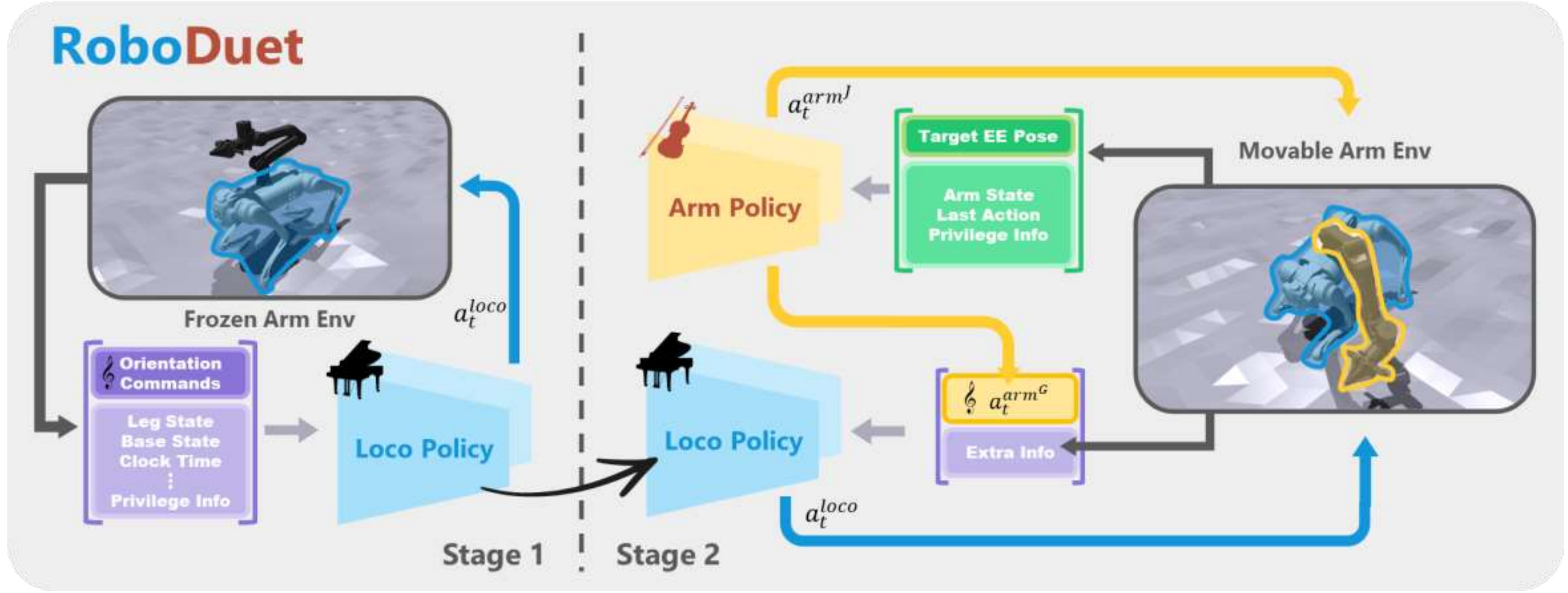
singularities of IK solution

WBC can expand the workspace

Manipulation with six degrees of freedom (6DoF) is essential for executing many daily task.

Inspired by the exquisite, harmonious, and layered duet of a violinist and pianist, we introduce RoboDuet.

Framework



Cooperative Policy | Two-stage Traing | Cross-Embodiment

The framework consists of two stages and two policy. Stage 1 focuses on obtaining the robust locomotion capability with the arm joints fixed and treated as payload. Stage 2 aims to coordinate locomotion and manipulation to achieve whole-body large-range mobile manipulation, when the arm policy is activated to control arm joints and the orientation of the quadruped robot.

Loco Policy

$$r_t^{loco} = r_t^{follow} + r_t^{gait} + r_t^{reg}$$

Raibert Heuristic reward

$$r_t^{c_f^{cmd}} = \sum_{\text{foot}} \left[1 - C_{\text{foot}}^{\text{cmd}}\left(\boldsymbol{\theta}^{cmd}, t\right)\right] \exp\left\{-\left|\mathbf{f}^{\text{foot}}\right|^2 / \sigma_{cf}\right\}$$

$$r_t^{c_v^{cmd}} = \sum_{\text{foot}} \left[C_{\text{foot}}^{\text{cmd}}\left(\boldsymbol{\theta}^{cmd}, t\right)\right] \exp\left\{-\left|\mathbf{v}_{xy}^{\text{foot}}\right|^2 / \sigma_{cv}\right\}$$

$$r_t^{s^{cmd}} = \left(\mathbf{p}_{x,y,\text{ foot}}^{f} - \mathbf{p}_{x,y,\text{ foot}}^{f,\text{ cmd}}\left(\boldsymbol{s}^{cmd}\right)\right)^2$$

$$r_t^{h^{cmd}} = \sum_{\text{foot}} \left(\boldsymbol{h}_{z,\text{ foot}}^{f} - \boldsymbol{h}_z^{f,cmd}\right)^2 C_{\text{foot}}^{\text{cmd}}\left(\boldsymbol{\theta}^{cmd}, t\right)$$

Arm Policy

$$r_t^{arm} = r_t^{manip} + r_t^{reg}$$

Manipulation Reward

$$r_t^{manip} = e^{-w \cdot \Delta lpy} \cdot e^{-\Delta\alpha\beta\gamma}$$

$$\Delta lpy = \Delta l + \Delta p + \Delta y$$

$$\Delta\alpha\beta\gamma = \Delta\alpha + \Delta\beta + \Delta\gamma$$

Result

Mobile Manipulation. To assess the effectiveness of two-stage training and cooperative policy, we compared four different training setups. Performance was evaluated by assigning random commands to robots and considering a command successful if the end-effector's tracking resulted in D ≤ 0.03m and θ ≤ π/18. The workspace was calculated as the area of the convex hull formed by the points from successful commands. We selected five checkpoints every 400 iterations up to 45,000 iterations, and performance metrics were derived from averaging 60,000 simulations.

TABLE II: Metrics from various training methods (scaled by 10−2). The initial three categories measure mean errors in robot velocity and end-effector position/orientation. The fourth assesses the robot's survival rate against external forces, and the fifth evaluates the robot workspace. "Still" tests occur with velx and ωz at zero, while "Move" tests are within command range limits.

Metrics		Still				Move			
		Baseline	Two-Stage	Cooperated	**RoboDuet**	Baseline	Two-Stage	Cooperated	**RoboDuet**
velocity tracking ↓	v_x (m/s)	0.95±0.34	0.81±0.19	0.66±0.03	**0.49±0.07**	12.49±2.48	10.39±0.35	12.13±0.99	**10.16±1.05**
	ω_z (rad/s)	0.83±0.50	0.42±0.12	0.47±0.03	**0.35±0.05**	52.12±0.45	52.97±0.38	**51.29±0.22**	51.53±0.67
position tracking↓	l (m)	4.53±0.21	4.77±0.26	**2.97±0.24**	3.01±0.32	4.61±0.31	4.96±0.23	**2.99±0.34**	3.02±0.28
	p (rad)	22.96±1.99	20.75±2.99	17.87±1.12	**17.65±2.05**	21.12±0.46	19.60±3.49	18.15±0.91	**17.66±2.01**
	y (rad)	24.87±3.11	23.27±0.79	18.57±1.11	**17.71±0.72**	23.12±2.83	22.64±1.12	18.69±1.43	**17.58±0.39**
	D (m)	13.36±0.25	12.17±0.93	10.43±0.54	**10.21±0.83**	12.31±0.43	11.91±1.29	10.44±0.36	**10.16±0.71**
orientation tracking↓	α (rad)	51.37±5.30	51.71±2.13	**44.31±2.99**	45.11±1.39	48.65±4.97	49.15±0.98	**43.81±3.02**	44.71±0.82
	β (rad)	90.68±10.24	89.91±8.83	79.48±3.56	**75.80±2.65**	90.83±10.26	88.92±6.45	78.09±3.05	**75.70±2.53**
	γ (rad)	83.94±11.32	78.43±3.36	66.21±1.87	**64.78±2.25**	82.88±10.07	76.98±4.50	66.43±1.71	**65.24±2.22**
	θ (rad)	92.50±4.85	93.69±3.99	**82.32±1.86**	84.08±3.13	92.39±5.64	94.62±4.33	**81.99±1.39**	83.88±3.23
survival rate (-) ↑		97.21±2.52	97.90±3.09	98.49±0.47	**98.99±0.52**	98.49±1.65	98.83±1.93	99.81±0.14	**99.92±0.03**
workspace (m^3) ↑		59.94±9.06	52.87±0.43	73.03±3.41	**75.63±2.78**	73.43±5.18	73.29±0.32	82.44±5.61	**85.39±0.43**

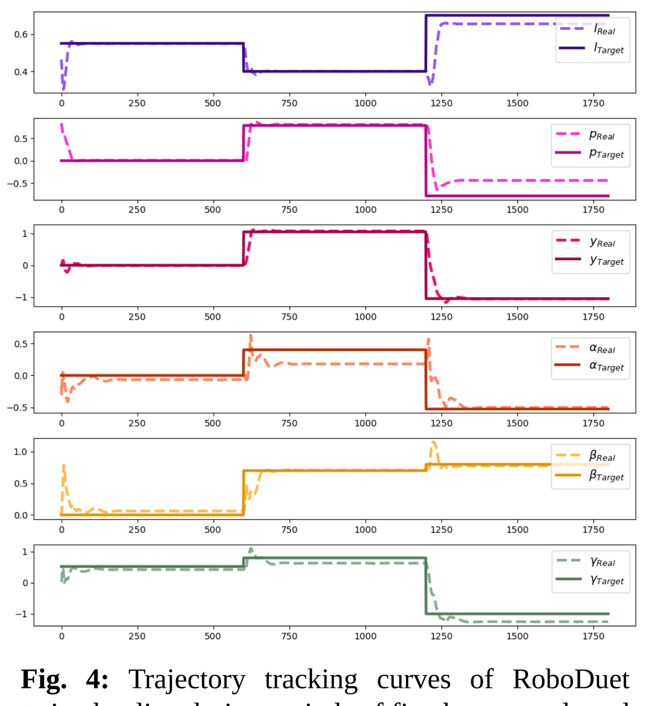


Fig. 4: Trajectory tracking curves of RoboDuet trained policy during periods of fixed commands and sudden changes.

In summary, RoboDuet adeptly integrates cooperative policy and two-stage training, markedly enhancing control performance in tracking accuracy and gait stability, which proves that both ingredients are indispensable.

In Fig.4, we present the variation in tracking error over time for a policy of RoboDuet, demonstrating its performance during periods of fixed commands and sudden changes. It can be observed that the policy converges quickly after the command change, achieving stable velocity and end-effector pose. This indicates that the average error in the final third of the timesteps is reasonable and also demonstrates the stability and robustness of the policy we have trained.

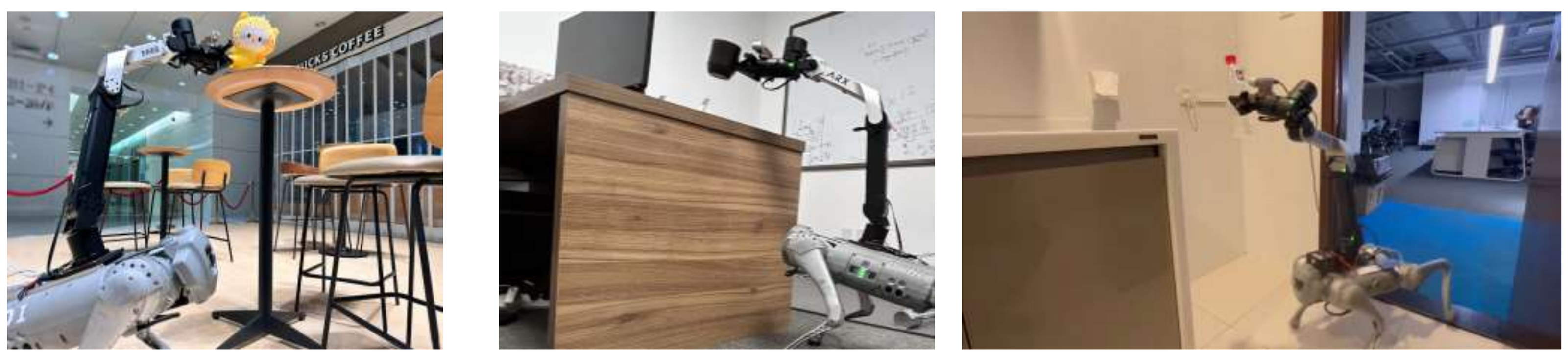
Cross-Embodiment. To exhibit the cross-embodiment capability of RoboDuet, we select two additional quadruped robots (Unitree A1 and Unitree Go2). We place Arx5 on these legged robots and train the robot system with the same training method as stage 1.

As the results shown in Table III, even without additional training for the whole-body control capabilities of the new robot systems, the new combined robots can maintain excellent velocity tracking and pose tracking ability, which not only saves the training costs associated with introducing new equipment, but also greatly demonstrate the superior zero-shot cross-embodiment capability of RoboDuet.

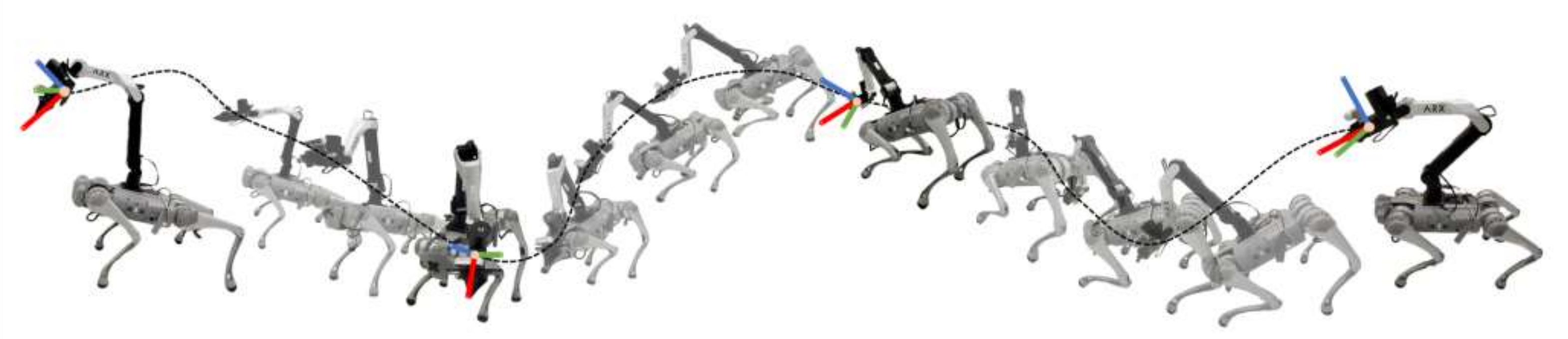
TABLE III: Workspace of Arx5 with different legged robots (×10−2). RoboDuet achieves large-range workspace for different quadruped embodiments.

Arx+X	Workspace (m^3)	
	Still	Move
Go1	75.63±2.78	85.39±0.43
Go2	75.74±2.78	77.04±3.87
A1	74.51±3.22	80.32±2.82

Experiments



Pick and Place. Regarding the task of pick and place, we evaluate the robot's ability to execute mobile manipulation tasks in various scenarios.



Trajectory Following. To observe the performance of the real robot when faced with sudden, discrete command changes, we randomly selected several commands to remain constant for specific periods. We then switched directly between these commands as observation without any transition.

The robot will rapidly reach the posture required by the command after a brief adjustment while ensuring a relatively high stability.

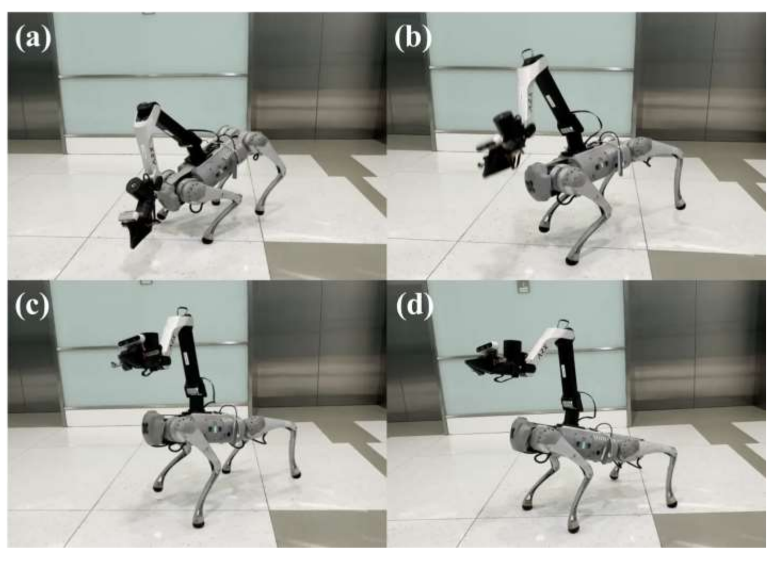


Fig. 6: A moment of the transition between two different discrete commands.

Cross-Embodiment. Despite the lack of physical access to a variety of quadruped robots and robotic arms, we conducted tests on six robotic system combinations in a simulation environment, involving the Unitree Go1, Go2, A1 models, and two types of robotic arms, the Arx5 and WidowX 250s.