OMNIINTERACT: TOWARDS UNIFIED WHOLE-BODY INTERACTION CONTROL FOR DUAL HUMANOIDS

Anonymous authors
Paper under double-blind review

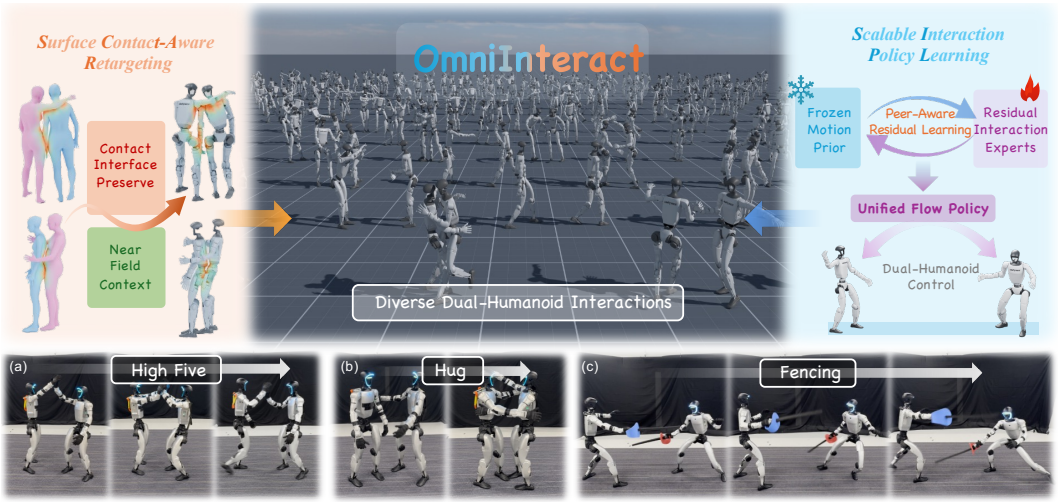


Figure 1: **Top center: OmniInteract enables diverse dual-humanoid interactions from human demonstrations. Left:** SCAR (Surface Contact-Aware Retargeting) preserves fine-grained surface interaction geometry during retargeting. **Right:** Specialist-to-unified residual learning consolidates diverse interaction skills into a unified flow policy. **Bottom (Real, a–c):** Deployment on dual Unitree G1 robots performing High-Five, Hug, and Fencing whole-body interactions.

ABSTRACT

Physical interaction is a key capability for humanoid collaboration, requiring coordinated whole-body motion and precise contact. However, learning diverse dual-humanoid interactions faces two major challenges: (1) human-robot morphological differences disrupt contact interfaces and spatial relationships during retargeting, and (2) jointly learning motor skills and partner coordination makes it difficult to retain control fidelity as interaction diversity grows. In this paper, we propose **OmniInteract** to address these challenges through surface-aware modeling and scalable residual policy learning. First, we introduce **Surface Contact-Aware Retargeting**, which leverages dense human-robot surface correspondences to jointly retarget both humanoids, preserving explicit interaction contact interfaces and their near-field geometric context. Second, we propose a **Scalable Flow-based Interaction Policy** for partner-conditioned residual generation. It factorizes dual-humanoid control into a reusable motor foundation and interaction-specific adaptations: partner-aware experts learn corrections over a frozen whole-body motion prior, while conditional flow matching unifies them in a shared residual action space. This enables scalable skill consolidation while retaining pretrained motor capabilities. Extensive experiments demonstrate that OmniInteract improves retargeting fidelity and dual-humanoid coordination while better retaining interaction performance as skill diversity grows. Real-world deployment further demonstrates successful contact establishment, stable whole-body coordination, and recovery from external perturbations.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

1 INTRODUCTION

Advances in whole-body control and motion imitation have equipped humanoid robots with increasingly versatile motor capabilities (Ben et al., 2025; Fu et al., 2024; He et al., 2024). However, Most existing systems, however, focus on single-humanoid skills (Chen et al., 2025; Jiang et al., 2026; He et al., 2025; Li et al., 2026b), whereas many real-world tasks require collaboration and physical interaction between agents (Shaoul et al., 2025; Doshi et al., 2026). Human-human interaction demonstrations offer a rich source of such behaviors. Transferring them to dual humanoids requires preserving the intended relationships between the bodies across embodiments and maintaining coordinated motion and contact under coupled dynamics.

The first challenge is ***preserving surface interaction geometry during retargeting.*** Morphological differences in body proportions and surface shape can disrupt contact interfaces despite plausible individual poses (Reda et al., 2023; Yin et al., 2025). Although recent interaction-aware approaches have begun to explicitly model relative geometries between bodies, sparse body representations remain insufficient for capturing fine-grained surface-level interactions (Chen et al., 2026). Explicit contact relationships constrain where two bodies should interact, but they do not fully determine how surrounding limbs and body regions should be arranged (Liu et al., 2025b). Conversely, preserving neighborhood-level geometric structures may maintain overall spatial consistency while hiding local contact deviations. Therefore, High-fidelity retargeting requires preserving both where the bodies meet and how nearby regions are arranged.

The second challenge is ***retaining coordination fidelity as interaction diversity grows.*** Independent motion trackers lack the partner feedback needed to correct deviations under coupled dynamics. Training from scratch instead requires learning individual tracking, whole-body balance, and partner coordination together. More importantly, different interactions exhibit substantially different contact patterns and coordination requirements. As interaction diversity increases, jointly training multiple interaction behaviors within a unified policy can introduce severe gradient interference, making it increasingly difficult to maintain high-fidelity interaction performance.

Motivated by this analysis, we propose **OmniInteract**, a framework for learning diverse dual-humanoid interactions from human demonstrations through surface-aware modeling and scalable residual policy learning. To overcome the retargeting bottleneck, we introduce **Surface Contact-Aware Retargeting (SCAR)**, which leverages dense human–robot surface correspondences to construct a contact-aware surface interaction graph between humanoid. It preserves interaction geometry through complementary surface constraints: explicit bilateral contact edges maintain the contact interface, while near-field Laplacian relationships preserve the surrounding spatial structures. SCAR jointly optimizes both humanoids under kinematic constraints, generating synchronized interaction motion references for downstream control.

To overcome the interaction control bottleneck, we propose a **Scalable Flow-based Interaction Policy** that consolidates partner-dependent corrections from task-specific dual-humanoid experts into a unified generative residual controller through flow matching. This design decouples stable single-humanoid motion priors from interaction control learning, leveraging high-capacity generative policies to expand interaction representation capability while avoiding severe interference multi-interaction behaviors reinforcement learning. fidelity and dual-humanoid coordination, together with stronger generalization to unseen motions and better performance retention as interaction diversity grows. Deployment on two Unitree G1 robots further demonstrates successful contact establishment, stable whole-body coordination, and recovery from external pushes.

Our contributions are summarized as follows:

- We present **OmniInteract**, a unified framework for learning dual-humanoid interactions from human demonstrations, integrating interaction-preserving retargeting with scalable coordination.
- We introduce a **surface contact-aware retargeting (SCAR)** method that leverages dense human–robot surface correspondences to preserve bilateral contact interfaces and surrounding interaction geometry for generating synchronized dual-humanoid references.
- We propose a **scalable flow-based interaction policy** that distills partner-aware corrections into a unified residual action space, enabling diverse coordination control while retaining a stable whole-body motion prior.

2 Related Works

Motion Retargeting for Humanoid. Retargeting human motions to humanoid robots is a fundamental step for transforming large-scale human demonstrations into robot-executable references (Park et al., 2025; Luo et al., 2023). Early approaches typically rely on manually designed correspondences between human and robot joints or body parts, followed by optimization under kinematic constraints (Araujo et al., 2025; Qin et al., 2022). Recent studies have improved retargeting quality by incorporating whole-body geometric structures, interaction constraints, and physical feasibility (Yang et al., 2025; Pan et al., 2025). However, existing representations often rely on sparse body configurations or predefined correspondences, which is insufficient to capture fine-grained surface interactions between multiple humanoids (Chen et al., 2026; Liu et al., 2025b). Recent dense correspondence-based approaches represent humanoid bodies through surface-level geometric correspondence (Cao et al., 2026), enabling more scalable motion transfer across embodiments. Our work builds upon this direction and extends surface-based retargeting toward dual-humanoid interaction by explicitly modeling bilateral contact interfaces and surrounding interaction geometry.

Whole-body Control for Humanoid Robots. Learning robust whole-body control policies for humanoid robots has recently advanced through motion imitation (Ji et al., 2024; Xie et al., 2026; Han et al., 2025), reinforcement learning (Li et al., 2026a; Peng et al., 2018; Zhang et al., 2025), and generative policy learning (Li et al., 2026c; Liu et al., 2025a; Shao et al., 2025). Prior works have demonstrated impressive capabilities in single-humanoid locomotion, manipulation (Gu et al., 2026; He et al., 2026), and dynamic motion tracking (Wang et al., 2026). Extending these abilities to dual-humanoid interaction introduces additional challenges, as the policy must simultaneously maintain individual motion stability and adapt to partner-dependent coordination requirements. Existing dual-humanoid control approaches typically learn interaction policies directly through reinforcement learning or task-specific optimization, which can become difficult to scale as interaction diversity increases (Liu et al., 2025b; Chen et al., 2026). In contrast, our approach leverages pretrained single-humanoid motion priors and learns interaction-specific residual adaptations, followed by specialist-to-unified policy consolidation, enabling scalable learning of diverse dual-humanoid interactions while preserving stable individual motion capabilities.

3 OmniInteract

Our goal is to learn whole-body humanoid–humanoid interactions from human–human demonstrations. Unlike independent motion imitation, this task requires preserving inter-body geometry across embodiments during retargeting and sustaining coordinated behavior under coupled dynamics during control. Therefore, we present **OmniInteract**, a unified framework combines **surface contact-aware retargeting (SCAR)** with **scalable flow-based interaction policy training** to address these challenges. In this section, we detail (i) how SCAR jointly retargets dual humanoids while preserving contact surface and surrounding spatial geometry, and (ii) how our flow-based interaction policy training specializes partner-aware residual corrections and consolidates them into a unified interaction policy.

3.1 Surface Contact-Aware Retargeting

Challenge of Interaction Retargeting. Conventional motion retargeting methods typically rely on sparse skeletal or keypoint constraints, which preserve coarse body motion but provide insufficient supervision for the surface contacts that define close human interactions. For example, in an embrace, aligning the wrists and torso does not ensure that the palms reach the intended regions of the partner's back. The core challenge lies in synthesizing motions that simultaneously preserve precise inter-body surface contacts and their surrounding local geometry.

Surface Correspondence and Interaction Coordinates. To model fine-grained inter-body surface contacts, we adopt Dense Point Cloud Correspondence (Cao et al., 2026) to establish surface-level associations across embodiments. For humanoid $a \in \{1, 2\}$, let $h^a_{t,i}$ denote the i-th source surface point at time t, and $x^a_i(q^a_t)$ its corresponding robot surface point under configuration q^a_t. The correspondence is established in the canonical pose and remains fixed throughout the motion,while

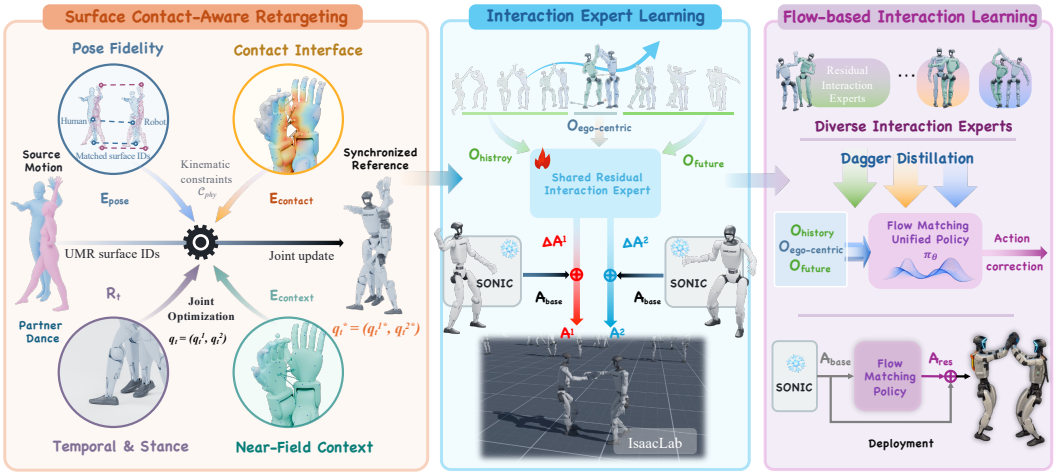


Figure 2: **Overview of the OMNIINTERACT pipeline. Left:** Starting from human-human demonstrations, SCAR jointly retargets the two humanoids while preserving individual motion, surface contacts, and near-field interaction context, yielding synchronized inetraction motions. **Center:** A frozen SONIC tracker provides the base actions, while task-specific residual experts learn partner-aware interaction corrections. **Right:** These diverse interaction experts are then distilled into a unified flow matching policy through DAgger. During deployment, the unified policy predicts residual actions on top of the frozen SONIC for coordinated dual-humanoid control.

each robot point is updated through forward kinematics according to its local body binding. See Appendix A.1 for preprocessing and surface-binding details.

To represent how a surface point is arranged relative to the opposite body, we utilize Laplacian coordinates (Alexa, 2003; Zhou et al., 2005) to define cross-body interaction coordinates. Specifically, For a surface point p_i (human or robot), let $\mathcal{N}_t(i)$ denote its local neighbors on the opposite body. The interaction surface coordinates can be formulated as:

$$\mathcal{L}_t(p_i) = p_i - \sum_{j \in \mathcal{N}_t(i)} \omega_{t,ij} p_j. \tag{1}$$

Here, $\omega_{t,ij}$ denotes the normalized weight of neighbor j, computed from the source geometry. $\mathcal{L}_t(p_i)$ encodes the local spatial relation between a surface point and the opposite body. Minimizing the deformation energy can therefore preserves local inter-body geometry across embodiments.

Optimization Objective. Given the surface correspondences and interaction coordinates defined above, SCAR jointly retargets the two humanoids by balancing individual motion fidelity with inter-body interaction preservation. Let $q_t = (q_t^1, q_t^2)$ denote the root poses and joint configurations of the two robots at time t. We optimize the paired configuration through the following objective:

$$q_t^\star = \arg\min_{q_t \in \mathcal{C}_{phy}} \mathcal{E}_{pose}(q_t) + \lambda_c \mathcal{E}_{contact}(q_t) + \lambda_n \mathcal{E}_{context}(q_t) + \mathcal{R}_t(q_t). \tag{2}$$

Here, $\mathcal{E}_{pose}$ preserves individual motion by matching corresponding surface positions and orientations, $\mathcal{E}_{contact}$ preserves explicit contact interfaces, and $\mathcal{E}_{context}$ maintains the surrounding near-field geometry. $\mathcal{R}_t$ regularizes temporal motion and stance-foot drift, and $\mathcal{C}_{phy}$ defines the feasible set under joint, velocity, floor, and collision constraints. The two humanoids are optimized jointly in a coupled configuration space. We solve Eq. (2) sequentially for each timestep using a Sequential Quadratic Programming (SQP) approach, warm-started with the optimal solution from the previous frame $q_t^\star$. Solver details, constraint handling, and hyperparameters are provided in Appendix A.2.

Contact-Interface Objective. To explicitly preserve physical contact at the surface level, we construct a sampled contact interface from the dense source correspondences. At each frame, we perform bidirectional nearest-neighbor search between the two source surfaces and collect pairs whose distances fall within a contact threshold. Temporal hysteresis is then applied to stabilize persistent

contacts, followed by region-balanced subsampling to obtain a budgeted contact set $\mathcal{C}_t$. For each $(i,j) \in \mathcal{C}_t$, we define the demonstrated and retargeted surface offsets as $d^h_{t,ij} = h^1_{t,i} - h^2_{t,j}$ and $d^r_{t,ij}(q_t) = x^1_i(q^1_t) - x^2_j(q^2_t)$, respectively. We preserve the contact interface by minimizing

$$\mathcal{E}_{contact}(q_t) = \sum_{(i,j)\in\mathcal{C}_t} \alpha_{t,ij} \left\| d^r_{t,ij}(q_t) - d^h_{t,ij} \right\|_2^2, \tag{3}$$

where $\alpha_{t,ij}$ is the normalized weights from the source surface. This objective preserves both the distance and direction of each contact interface by matching the relative contact-offset vectors.

Near-Field Context Objective. To maintain the arrangement of nearby limbs and surfaces, we construct a compact set of near-field context relations around the contact interface. For source surface points that were not selected as contact endpoints, we first identify near-field candidates according to their distance to the opposite body. For each humanoid $a \in \{1,2\}$, we retain the closest candidates up to a fixed number to form the context-center set $\mathcal{S}^a_t$. For each center $i \in \mathcal{S}^a_t$, we find its nearest point on the opposite body and augment it with valid one-ring surface neighbors within a local distance threshold to form $\mathcal{N}_t(i)$. Using the interaction coordinates defined in Eq. (1), we minimize

$$\mathcal{E}_{context}(q_t) = \sum_{a\in\{1,2\}} \sum_{i\in\mathcal{S}^a_t} \beta^a_{t,i} \left\| \mathcal{L}_t(x^a_i(q^a_t)) - \mathcal{L}_t\big(h^a_{t,i}\big) \right\|_2^2. \tag{4}$$

where $\beta^a_{t,i}$ is the normalized proximity weight of context center i. This objective aims to preserve the local organization of near-field surface structure while accommodating morphological differences. Contact selection, neighborhood construction, and weight definitions are detailed in Appendix A.2.

3.2 Scalable Flow-Based Interaction Policy Training

Problem Formulation and Challenge of Interactive Control. We formulate the interaction control as a Multi-Agent Markov Decision Process (MA-MDP): $\mathcal{G} = (\mathcal{S},\ \mathcal{O}, \mathcal{A}, \mathcal{P}, \mathcal{R}, \gamma)$, where $\mathcal{S}$ is the global state space, $\mathcal{O} = \prod_a \mathcal{O}^a$ and $\mathcal{A} = \prod_a \mathcal{A}^a$ are the joint observation and action spaces of humanoid $a \in \{1,2\}$, while the transition function $\mathcal{P} : \mathcal{S} \times \mathcal{A} \to \mathcal{S}$ specifies the environment dynamics. The reward function $\mathcal{R} : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ provides a shared interaction scalar signal, and γ is the discount factor. The objective is to learn a joint policy π that maximizes $\mathbb{E}_\pi[\sum_{\tau=t}^{\infty} \gamma^\tau r_\tau]$.

Given the interaction motions $\mathcal{M}^a$ ($a \in \{1,2\}$) retargeted by SCAR, our goal is to achieve coordinated execution under coupled dynamics. However, most existing strong trackers are designed for single-humanoid and cannot effectively handle partner-dependent deviations induced by physical interaction. Training a dual-humanoid controller from scratch, in contrast, requires simultaneously learning motion tracking, balance, and inter-body coordination. This challenge becomes more pronounced as increasingly diverse interaction patterns are incorporated into a single policy.

Residual Interaction Expert Learning. To avoid relearning dual-humanoid low-level motor control from scratch, we build upon *SONIC* (Luo et al., 2026), a strong general whole-body tracker, and preserve its pretrained tracking and balance capabilities. We freeze the SONIC policy π_{base} and learn only the interaction-specific correction through a residual expert π^k_{res} for each interaction task k. For humanoid a and its partner $\bar{a}$, the final corrected action is formulated as:

$$\mathcal{A}^a = \underbrace{\pi_{base}(\mathcal{O}^a, \mathcal{M}^a)}_{\text{frozen action } \mathcal{A}^a_{base}} + \underbrace{\alpha_{res} \tanh\big(\pi^k_{res}(\mathcal{O}^a_{his}, \mathcal{O}^a_{ego}, \mathcal{O}^a_{fur})\big)}_{\text{residual interaction action } \Delta\mathcal{A}^{\bar{a}\to a}}. \tag{5}$$

Here, $\mathcal{A}^a_{base}$ denotes the base action produced by the frozen SONIC tracker, while $\Delta\mathcal{A}^{\bar{a}\to a}$ represents the partner-dependent residual correction for humanoid a. α_{res} bounds the residual magnitude. Here, $\mathcal{O}^a_{his}$ encodes the recent history of self and partner states, previous actions, and relative-geometry errors; $\mathcal{O}^a_{ego}$ represents the partner's state and relative pose in the ego-centric frame; and $\mathcal{O}^a_{fur}$ provides short-horizon motion references for both humanoids and their desired relative geometry. Both are encoded by temporal 1D CNNs, and the resulting features are fused through an MLP decoder to predict the residual action. For each interaction task, the residual expert is shared by both humanoids and trained with Multi-Agent Proximal Policy Optimization (MAPPO) under centralized training with decentralized execution (Littman, 1994; Gao et al., 2024), using a shared reward composed of motion tracking, interaction-geometry alignment, contact-timing consistency,

and residual regularization. This design decouples interaction coordination from low-level tracking and balance, allowing reinforcement learning to focus on partner-dependent corrections rather than relearning individual control. Detailed observations, network architecture and reward definitions are provided in Appendix B.1.

Flow-Matching Interaction Policy Learning. To consolidate task-specific interaction skills into a unified policy, we distill the residual experts into a flow-based interaction policy using DAgger (Ross et al., 2011). We first roll out the current flow-based policy $\pi_\theta(\Delta\mathcal{A} \mid \hat{\mathcal{O}})$ on top of the frozen Sonic base action in dual humanoids simulation and collecting a trajectory of visited states given interaction motions dataset retargeted by SCAR, where $\hat{\mathcal{O}}$ contains the motion histories and future reference features of both humanoids. For each visited state, we obtain the residual interaction expert action $\Delta\mathcal{A}_{\text{expert}}$ by querying the corresponding interaction expert policy. Since the Sonic tracker remains frozen and shared across all experts, all supervision lies in a common residual action space, allowing the unified policy to focus on interaction coordination without relearning individual motion. The flow-based model then learn to recover the interaction expert residual action $\Delta\mathcal{A}_{\text{expert}}$ from noised action by optimizing:

$$\mathcal{L}_{fm}(\theta) = \mathbb{E}_{t,\epsilon,\Delta\mathcal{A}_{\text{expert}}}\left[\left\| v_\theta(\Delta\mathcal{A}_t, t, \hat{\mathcal{O}}) - (\boldsymbol{\epsilon} - \Delta\mathcal{A}_{\text{expert}})\right\|_2^2\right], \tag{6}$$

where $\Delta\mathcal{A}_t = (1-t)\Delta\mathcal{A}_{\text{expert}} + t\boldsymbol{\epsilon}$ is the noisy residual action interpolated between the expert residual $\Delta\mathcal{A}_{\text{expert}}$ and Gaussian noise $\boldsymbol{\epsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ along the flow timestep $t \in [0, 1]$ (Lipman et al., 2023). This objective learns a conditional vector field $v_\theta(\Delta\mathcal{A}_t, t, \hat{\mathcal{O}})$ to predict the target velocity $\boldsymbol{\epsilon} - \Delta\mathcal{A}_{\text{expert}}$, learning the denoising direction from expert residual actions to random noise along the flow trajectory. With the learned vector field, the residual interaction action can be generated from Gaussian noise by integrating v_θ from $t = 1$ to $t = 0$ using the forward Euler rule:

$$\Delta\mathcal{A}_{t-\frac{1}{N}} = \Delta\mathcal{A}_t - \frac{1}{N} v_\theta(\Delta\mathcal{A}_t, t, \hat{\mathcal{O}}), \tag{7}$$

where N denotes the number of integration steps controlling the sampling accuracy. By iteratively rolling out trajectories and supervising them with interaction expert residuals through DAgger, we learn a unified flow-based policy that maps interaction conditions to partner-aware residual corrections. During deployment, the generated residual action is added to the frozen Sonic base action for dual-humanoid control. Detailed deployment procedures and settings are provided in the Appendix B.2.

4 Experiments

We present extensive experiments in simulation and on physical robots to evaluate the interaction retargeting fidelity, control scalability, and real-world executability of our proposed **OmniInteract** framework. Our experiments are organized around the following key questions:

- **Q1: Interaction-Preserving Retargeting.** Can **SCAR** generate feasible trajectories that better preserve fine-grained inter-body geometry than prior retargeting methods? (Sec. 4.2)
- **Q2: Residual Interaction Learning.** Do residual experts built on frozen motion priors improve dual-humanoid coordination over independent or scratch-trained controllers? (Sec. 4.3)
- **Q3: Flow-Based Interaction Scaling.** Does scaling a flow-based generative policy improve multi-interaction control more effectively than conventional MLP controllers? (Sec. 4.4)
- **Q4: Real-World Robustness.** Can **OmniInteract** transfer to physical humanoids under perturbations while maintaining accurate contacts and stable whole-body coordination? (Sec. 4.5)

4.1 Experiments Settings

Datasets. We evaluate OmniInteract on two human–human interaction datasets, **InterHuman** (Liang et al., 2024) and **Inter-X** (Xu et al., 2024). We use the full motion sets from both datasets for interaction retargeting evaluation. For control, we select 30 diverse interaction motions from each dataset to form the control benchmark to assess performance across physical dynamics. Moreover, we further construct a unseen motion set to evaluate generalization beyond the curated interaction motion library. More dataset details are provided in Appendix C.1.

Table 1: **Quantitative Results of Retargeting.** Metrics include physical feasibility (MPD), interaction fidelity (CIE, NGE, C-F1), and utility (DSR). Best results are shown in **bold**.

Dataset	Method	Feasibility	Interaction Fidelity			Utility
		$\text{MPD}_{(\text{cm})} \downarrow$	$\text{CIE}_{(\%)} \downarrow$	$\text{NGE}_{(\%)} \downarrow$	$\text{C-F1}_{(\%)} \uparrow$	$\text{DSR}_{(\%)} \uparrow$
InterHuman	PHC	2.59	10.07	10.89	47.32	40.6
	GMR	4.76	13.94	13.03	31.98	36.4
	OmniRetarget	3.54	6.71	6.91	58.69	58.7
	UMR	4.04	7.83	7.73	51.14	56.9
	SCAR (Ours)	**0.93**	**4.96**	**6.49**	**65.18**	**75.4**
Inter-X	PHC	3.15	10.94	11.12	33.90	31.2
	GMR	5.10	9.64	9.36	43.07	34.6
	OmniRetarget	5.02	8.09	8.32	56.21	48.7
	UMR	4.98	7.19	7.04	50.03	55.1
	SCAR (Ours)	**1.20**	**3.98**	**4.99**	**64.08**	**72.1**

Baselines. We compare against representative baselines for both interaction retargeting and control. For retargeting, we evaluate PHC (Luo et al., 2023), GMR (Araujo et al., 2025), OmniRetarget (Yang et al., 2025), and UMR (Cao et al., 2026) using the same source motions and skeletal-scaling preprocessing. For control, we consider two groups. *Single-humanoid trackers* include TWIST2 (Ze et al., 2025), BeyondMimic (Liao et al., 2026), and SONIC (Luo et al., 2026), which independently execute the paired reference motions for dual humanoids. *Peer-aware variants* include **From-scratch RL**, which directly learns full-body control without frozon motion priors, and **Specialist→Unified MLP**, which retains the residual specialists but replaces flow matching with deterministic regression. These comparisons isolate the advantages of residual interaction learning and flow-based policy consolidation. Implementation for all baselines are detailed in Appendix C.2.

Metrics. We evaluate retargeting quality and control performance using task-specific metrics. For retargeting, we report **MPD**(cm) for physical plausibility, **CIE**(%), **NGE**(%), and **C-F1**(%) for interaction fidelity, and **DSR**(%) for downstream learnability. For control, we report E_{mpjpe}(mm) for motion tracking error, $E_{inter\text{-}geo}$(%) for interaction preservation, and **Succ**(%) for overall interaction execution success. For conciseness, detailed metric definitions are provided in Appendix C.3.

4.2 Fidelity: Interaction-Preserving Retargeting (Q1)

Quantitative Analysis. Table 1 reveals three clear trends. First, sparse single-body fitting is insufficient for interaction preservation: GMR exhibits consistently large CIE and NGE, indicating that accurate keypoint tracking alone cannot recover fine-grained cross-body geometry. Second, surface-aware baselines improve interaction fidelity but still struggle to balance geometry and feasibility when the two humanoids are optimized independently. OmniRetarget suffers from noticeable penetration and degraded near-field consistency, while UMR benefits from dense surface correspondences but still lacks explicit pairwise interaction constraints. Third, SCAR consistently resolves this trade-off, achieving the best MPD, CIE, NGE, and C-F1 on both datasets. On Inter-X, it reduces CIE and NGE by 44.6% and 29.1% over the second-best methods. Importantly, this improvement also translates to downstream learnability, with DSR reaching 75.4% and 72.1% on InterHuman and

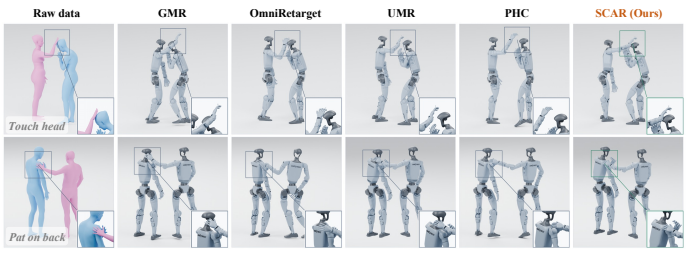


Figure 3: Qualitative visualization comparison of interaction retargeting. (**Top**: Touch Head, **Bottom**: Pat on Back)

Table 2: **Comparison of controllers on dual-humanoid interaction benchmarks.** InterHuman† and Inter-X† denote the motion subsets selected from each dataset for control evaluation.

Method	InterHuman† + Inter-X†			Unseen Motions		
	$E_{mpjpe}\downarrow$	$E_{inter\text{-}geo}\downarrow$	Succ(%)↑	$E_{mpjpe}\downarrow$	$E_{inter\text{-}geo}\downarrow$	Succ(%)↑
Single-Agent Trackers						
TWIST2	86.56	11.44	41.36	113.35	14.72	28.58
BeyondMimic	64.08	8.91	69.27	–	–	–
SONIC	65.31	9.56	63.32	66.31	9.51	62.65
Peer-Aware Controllers						
From-Scratch RL	59.93	8.54	68.49	68.71	10.83	63.78
Specialist→Unified MLP	46.23	7.29	80.68	63.52	9.76	68.20
OmniInteract	**41.67**	**6.42**	**84.32**	**61.67**	**8.92**	**72.89**

Inter-X. These results highlight that faithful interaction retargeting requires modeling geometry between the two bodies, rather than fitting each body independently.

Qualitative Analysis. Fig. 3 compares the retargeted interactions with the source motions, with close-up views highlighting surface alignment. In *Touch Head* (top, InterHuman), baselines reproduce the overall gesture but exhibit displaced or separated hand–head contact, whereas SCAR better preserves the intended contact location. In *Pat on Back* (bottom, Inter-X), SCAR more faithfully retains the hand–back interface and the relative arm–torso arrangement. These examples highlight the importance of preserving contact interface and near-field context.

4.3 Effectiveness: Partner-Aware Residual Control (Q2)

Simulation Evaluation. To evaluate the effectiveness of partner-aware residual learning in leveraging single-humanoid motion priors for coordinated dual-humanoid interaction, we compare OmniInteract with independent trackers and peer-aware controllers on the selected InterHuman and Inter-X subset. Table 2 reveals a clear gap between individual motion tracking and interactive coordination. Although single-agent trackers achieve reasonable pose imitation, they fail to maintain interaction consistency under coupled dynamics. In particular, SONIC attains comparable tracking accuracy but only achieves 63.32% success, indicating that independently following paired references is insufficient to preserve partner-relative configurations. Moreover, From-scratch RL reaches only 68.49% success, as it must jointly discover individual tracking, balance, and interaction coordination. Building interaction capability on top of the frozen SONIC prior improves this limitation. In contrast, Specialist→Unified MLP achieves 80.68% success, showing the benefit of learning interaction-specific corrections in a residual space. OmniInteract further improves success to 84.32% while achieving the lowest interaction-geometry error (6.42%) and E_{mpjpe} (41.67 mm). These results demonstrate that preserving pretrained motor priors while learning partner-dependent residual corrections provides a more effective route to scalable dual-humanoid coordination.

4.4 Scalability: Unified Control Cross Diverse Interactions (Q3)

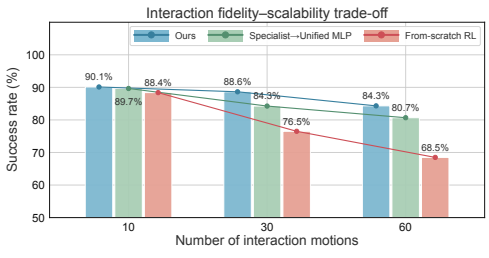


Figure 4: Fidelity–scalability trade-off.

Table 3: Real-world evaluation on two Unitree G1 robots across three interaction tasks.

Task	Attempts	Method	Succ. (%)↑
High-Five	14	SONIC	28.6
		Ours	**85.7**
Hug	22	SONIC	22.7
		Ours	**72.7**
Shoulder	18	SONIC	11.1
		Ours	**77.8**

To evaluate the generalization and scalability of unified interaction control, we first test on unseen motions (randomly sampled from retargeted InterHuman & Inter-X). As summarized in Table 2, OmniInteract outperforms all applicable baselines across the all metrics, achieving 72.89% success and demonstrating that its coordination capabilities transfer beyond the training references. To characterize the fidelity–scalability trade-off in multi-interaction control, we further scale the interaction diversity by training on an expanding set of motions drawn from our control benchmark. As shown in Fig. 4, the methods perform similarly at small scale, but From-scratch RL deteriorates more rapidly as motion diversity increases. At 60 motions, OmniInteract retains 84.3% success, compared with 80.7% for Specialist→Unified MLP and 68.5% for From-scratch RL. These results support flow-based specialist consolidation as an effective way to accommodate diverse interactions while limiting the loss of execution fidelity.

4.5 Robustness: Real-World Dual-Humanoid Interaction (Q4)

We deploy OmniInteract on two Unitree G1 humanoids and evaluate its real-world performance on contact-rich interactions and under external force perturbations. Table 3 reports success rates based on contact establishment at specific keyframes across three tasks: High-Five, Hug, and Shoulder Hold. OmniInteract consistently outperforms the single-humanoid tracker baseline SONIC by over 50%, with the largest margin on Shoulder Hold (77.8% vs. 11.1%). SONIC is particularly vulnerable to accumulated relative drift in this long-horizon interaction, while OmniInteract continuously corrects partner-dependent misalignment through its residual corrections policy. These results demonstrate that partner-aware corrections transfer effectively to physical robots, while independent tracking often fails to establish the intended contact positions.



Figure 5: Real-world deployment of contact interaction tasks and robustness to disturbances.

Robustness in Realistic Deployment. To evaluate robustness under external disturbances, we apply pushes during interaction execution. As shown in Fig. 5(a), the humanoids adapt their poses and relative configuration to recover the intended arm-grabbing interaction after perturbation. The Shoulder Hold and Blocking & Pushing examples in Fig. 5(b–c) further demonstrate stable coordination under sustained physical coupling. These results highlight the closed-loop correction capability of the partner-aware residual policy, which compensates for interaction-induced deviations and restores task-relevant contacts online. (Additional examples of real-world humanoid interactions are provided in the *supplementary video* and Appendix D.2.)

5 Conclusion

In this paper, we present **OmniInteract** to address the challenges of preserving interaction geometry across embodiments and scaling coordinated control across diverse dual-humanoid interactions. We introduce SCAR, which jointly retargets both humanoids through contact-interface and near-field context modeling, producing synchronized references that preserve fine-grained surface relationships. We further propose a scalable flow-based interaction policy that factorizes control into a frozen motor prior and partner-dependent residuals and consolidates expert corrections into a unified generative controller through flow matching, while retaining pretrained motor capabilities. Extensive experiments demonstrate improved retargeting fidelity, higher coordination success, and better performance retention as interaction diversity grows, with physical deployment validating contact establishment and recovery from external perturbations. By connecting surface interaction modeling with scalable residual control, OmniInteract provides an effective route from human demonstrations to coordinated physical interactions between humanoids.

AI USE STATEMENT

We used generative AI tools to assist with refining research ideas and experimental visualization. We additionally used generative AI tools for literature summary, manuscript polishing, LaTeX formatting. All AI-assisted content was critically reviewed and revised by the authors, and literature references were manually verified. We take full responsibility for the final content of this work, including any text, or artifacts produced with the assistance of generative AI.

ETHICS STATEMENT

This work studies interaction control between humanoid robots and does not involve the collection of new human-subject data. Our experiments use existing human motion datasets and robot–robot interactions in simulated or controlled experimental environments. As physical interaction policies may pose safety risks when deployed on real robotic systems, appropriate safety measures and human supervision should be employed in real-world applications.

REPRODUCIBILITY STATEMENT

We provide the implementation of **OmniInteract** in the supplementary material. Appendix A details the retargeting formulation, optimization procedure, and hyperparameters. Appendix B specifies the policy architectures, training configurations, and real-world deployment setup. Appendix C documents the datasets, motion selection, baseline implementations, evaluation protocols, and metric definitions. Appendix D provides additional ablation studies, sensitivity analysis, and qualitative results. Together, these materials support faithful reproduction of the proposed framework and its evaluation.

REFERENCES

Marc Alexa. Differential coordinates for local mesh morphing and deformation. *The Visual Computer*, 19(2):105–114, 2003.

Joao Pedro Araujo, Yanjie Ze, Pei Xu, Jiajun Wu, and C Karen Liu. Retargeting matters: General motion retargeting for humanoid motion tracking. *arXiv preprint arXiv:2510.02252*, 2025.

Qingwei Ben, Feiyu Jia, Jia Zeng, Junting Dong, Dahua Lin, and Jiangmiao Pang. HOMIE: Humanoid loco-manipulation with isomorphic exoskeleton cockpit. In *RSS 2025 Workshop on Whole-body Control and Bimanual Manipulation: Applications in Humanoids and Beyond*, 2025. URL https://openreview.net/forum?id=WSZKGX2Ty7.

Hanyang Cao, Yuetong Fang, Taesoo Kwon, Runyi Yu, Ji Ma, Jing Tan, Yangchen Zhou, Baoze Du, Yi Gu, Yukang Gao, et al. Unified motion retargeting for humanoids with learned point cloud correspondence. *arXiv preprint arXiv:2609.02134*, 2026.

Hongjin Chen, Wei Zhang, Pengfei Li, Shihao Ma, Ke Ma, Yujie Jin, Zijun Xu, Xiaohui Wang, Yupeng Zheng, Zining Wang, et al. Rhythm: Learning interactive whole-body control for dual humanoids. *arXiv preprint arXiv:2603.02856*, 2026.

Zixuan Chen, Mazeyu Ji, Xuxin Cheng, Xuanbin Peng, Xue Bin Peng, and Xiaolong Wang. Gmt: General motion tracking for humanoid whole-body control. *arXiv preprint arXiv:2506.14770*, 2025.

Ria Doshi, Tian Gao, Annie Chen, Chelsea Finn, and Jeannette Bohg. Chorus: Decentralized multi-embodiment collaboration with one vla policy. *arXiv preprint arXiv:2606.12352*, 2026.

Zipeng Fu, Qingqing Zhao, Qi Wu, Gordon Wetzstein, and Chelsea Finn. Humanplus: Humanoid shadowing and imitation from humans. In *8th Annual Conference on Robot Learning*, 2024. URL https://openreview.net/forum?id=WnSl42M9Z4.

Jiawei Gao, Ziqin Wang, Zeqi Xiao, Jingbo Wang, Tai Wang, Jinkun Cao, Xiaolin Hu, Si Liu, Jifeng Dai, and Jiangmiao Pang. Coohoi: Learning cooperative human-object interaction with manipulated object dynamics. *Advances in Neural Information Processing Systems*, 37:79741–79763, 2024.

Zhaoyuan Gu, Junheng Li, Wenlan Shen, Wenhao Yu, Zhaoming Xie, Stephen McCrory, Xianyi Cheng, Abdulaziz Shamsah, Robert Griffin, C Karen Liu, et al. Humanoid locomotion and manipulation: Current progress and challenges in control, planning, and learning. *IEEE/ASME Transactions on Mechatronics*, 31(2):2300–2330, 2026.

Jinrui Han, Weiji Xie, Jiakun Zheng, Jiyuan Shi, Weinan Zhang, Ting Xiao, and Chenjia Bai. Kungfubot2: Learning versatile motion skills for humanoid whole-body control. *arXiv preprint arXiv:2509.16638*, 2025.

Dongjiao He, Wei Xu, Nan Chen, Fanze Kong, Chongjian Yuan, and Fu Zhang. Point-lio: robust high-bandwidth light detection and ranging inertial odometry. *Advanced Intelligent Systems*, 5 (7):2200459, 2023.

Tairan He, Zhengyi Luo, Xialin He, Wenli Xiao, Chong Zhang, Weinan Zhang, Kris M. Kitani, Changliu Liu, and Guanya Shi. Omnih2o: Universal and dexterous human-to-humanoid whole-body teleoperation and learning. In *8th Annual Conference on Robot Learning*, 2024. URL https://openreview.net/forum?id=oL1WEZQal8.

Tairan He, Jiawei Gao, Wenli Xiao, Yuanhang Zhang, Zi Wang, Jiashun Wang, Zhengyi Luo, Guanqi He, Nikhil Sobanbab, Chaoyi Pan, et al. Asap: Aligning simulation and real-world physics for learning agile humanoid whole-body skills. *arXiv preprint arXiv:2502.01143*, 2025.

Tairan He, Zi Wang, Haoru Xue, Qingwei Ben, Zhengyi Luo, Wenli Xiao, Ye Yuan, Xingye Da, Fernando Castañeda, Shankar Sastry, et al. Viral: Visual sim-to-real at scale for humanoid loco-manipulation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 13430–13441, 2026.

Yih Huang, Eric Fleury, and Philip K McKinley. Lcm: A multicast core management protocol for link-state routing networks. In *ICC'98. 1998 IEEE International Conference on Communications. Conference Record. Affiliated with SUPERCOMM'98 (Cat. No. 98CH36220)*, volume 2, pp. 1197–1201. IEEE, 1998.

Mazeyu Ji, Xuanbin Peng, Fangchen Liu, Jialong Li, Ge Yang, Xuxin Cheng, and Xiaolong Wang. Exbody2: Advanced expressive humanoid whole-body control. *arXiv preprint arXiv:2412.13196*, 2024.

Haoran Jiang, Jin Chen, Qingwen Bu, Li Chen, Modi Shi, Yanjie Zhang, Delong Li, Chuanzhe Suo, Hongyang Li, et al. Wholebodyvla: Towards unified latent vla for whole-body loco-manipulation control. In *International Conference on Learning Representations*, volume 2026, pp. 157438–157461, 2026.

Peng Li, Zihan Zhuang, Yangfan Gao, Yi Dong, Sixian Li, Changhao Jiang, Shihan Dou, Zhiheng Xi, Enyu Zhou, Jixuan Huang, et al. From-w1: Towards general humanoid whole-body control with language instructions. *arXiv preprint arXiv:2601.12799*, 2026a.

Yitang Li, Zhengyi Luo, Tonghe Zhang, Cunxi Dai, Anssi Kanervisto, Andrea Tirinzoni, Haoyang Weng, Kris Kitani, Mateusz Guzek, Ahmed Touati, et al. Bfm-zero: A promptable behavioral foundation model for humanoid control using unsupervised reinforcement learning. In *International Conference on Learning Representations*, volume 2026, pp. 79697–79725, 2026b.

Zhe Li, Yangyang Wei, Boan Zhu, Yibo Peng, Tao Huang, Pengwei Wang, Zhongyuan Wang, Cheng Chi, Chang Xu, and Shanghang Zhang. From language to locomotion: Retargeting-free humanoid control via motion latent guidance. In *International Conference on Learning Representations*, volume 2026, pp. 129292–129319, 2026c.

Han Liang, Wenqian Zhang, Wenxuan Li, Jingyi Yu, and Lan Xu. Intergen: Diffusion-based multi-human motion generation under complex interactions. *International Journal of Computer Vision*, 132(9):3463–3483, 2024.

Qiayuan Liao, Takara E Truong, Xiaoyu Huang, Yuman Gao, Guy Tevet, Koushil Sreenath, and C Karen Liu. Beyondmimic: From motion tracking to versatile humanoid control via guided diffusion. *Science Robotics*, 11(117):eadx8924, 2026.

Yaron Lipman, Ricky T. Q. Chen, Heli Ben-Hamu, Maximilian Nickel, and Matthew Le. Flow matching for generative modeling. In *International Conference on Learning Representations*, 2023.

Michael L Littman. Markov games as a framework for multi-agent reinforcement learning. In *Machine learning proceedings 1994*, pp. 157–163. Elsevier, 1994.

Zhirui Liu, Kaiyang Ji, Ke Yang, Yahao Fan, Jingyi Yu, Ye Shi, and Jingya Wang. Commanding humanoid by free-form language: A large language action model with unified motion vocabulary. *arXiv preprint arXiv:2511.22963*, 2025a.

Zuhong Liu, Junhao Ge, Minhao Xiong, Jiahao Gu, Bowei Tang, Wei Jing, and Siheng Chen. It takes two: Learning interactive whole-body control between humanoid robots. *arXiv preprint arXiv:2510.10206*, 2025b.

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*, pp. 10895–10904, 2023.

Zhengyi Luo, Ye Yuan, Tingwu Wang, Chenran Li, Fernando Castañeda, Sirui Chen, Zi-Ang Cao, Jiefeng Li, David Minor, Qingwei Ben, et al. Sonic: Supersizing motion tracking for natural humanoid whole-body control. *Science Robotics*, 11(117):eaed4592, 2026.

Chaoyi Pan, Changhao Wang, Haozhi Qi, Zixi Liu, Homanga Bharadhwaj, Akash Sharma, Tingfan Wu, Guanya Shi, Jitendra Malik, and Francois Hogan. Spider: Scalable physics-informed dexterous retargeting. *arXiv preprint arXiv:2511.09484*, 2025.

Sungjae Park, Homanga Bharadhwaj, and Shubham Tulsiani. Demodiffusion: One-shot human imitation using pre-trained diffusion policy. *arXiv preprint arXiv:2506.20668*, 2025.

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.

Yuzhe Qin, Hao Su, and Xiaolong Wang. From one hand to multiple hands: Imitation learning for dexterous manipulation from single-camera teleoperation. *IEEE Robotics and Automation Letters*, 7(4):10873–10881, 2022.

Daniele Reda, Jungdam Won, Yuting Ye, Michiel Van De Panne, and Alexander Winkler. Physics-based motion retargeting from sparse inputs. *Proceedings of the ACM on Computer Graphics and Interactive Techniques*, 6(3):1–19, 2023.

Stéphane Ross, Geoffrey Gordon, and Drew Bagnell. A reduction of imitation learning and structured prediction to no-regret online learning. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pp. 627–635. JMLR Workshop and Conference Proceedings, 2011.

Aleksandr Segal, Dirk Haehnel, Sebastian Thrun, et al. Generalized-icp. In *Robotics: science and systems*, volume 2, pp. 435. Seattle, WA, 2009.

Yiyang Shao, Xiaoyu Huang, Bike Zhang, Qiayuan Liao, Yuman Gao, Yufeng Chi, Zhongyu Li, Sophia Shao, and Koushil Sreenath. Langwbc: Language-directed humanoid whole-body control via end-to-end learning. *arXiv preprint arXiv:2504.21738*, 2025.

Yorai Shaoul, Zhe Chen, Mohamed Naveed Gul Mohamed, Federico Pecora, Maxim Likhachev, and Jiaoyang Li. Collaborative multi-robot non-prehensile manipulation via flow-matching co-generation. *arXiv preprint arXiv:2511.10874*, 2025.

Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco: A physics engine for model-based control. In *2012 IEEE/RSJ international conference on intelligent robots and systems*, pp. 5026–5033. IEEE, 2012.

Yunshen Wang, Shaohang Zhu, Peiyuan Zhi, Yuhan Li, Jiaxin Li, Yong-Lu Li, Yuchen Xiao, Xingxing Wang, Baoxiong Jia, and Siyuan Huang. Omnixtreme: Breaking the generality barrier in high-dynamic humanoid control. *arXiv preprint arXiv:2602.23843*, 2026.

Weiji Xie, Jinrui Han, Jiakun Zheng, Huanyu Li, Xinzhe Liu, Jiyuan Shi, Weinan Zhang, Chenjia Bai, and Xuelong Li. Kungfubot: Physics-based humanoid whole-body control for learning highly-dynamic skills. *Advances in Neural Information Processing Systems*, 38:62406–62433, 2026.

Liang Xu, Xintao Lv, Yichao Yan, Xin Jin, Shuwen Wu, Congsheng Xu, Yifan Liu, Yizhou Zhou, Fengyun Rao, Xingdong Sheng, et al. Inter-x: Towards versatile human-human interaction analysis. In *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 22260–22271. IEEE, 2024.

Lujie Yang, Xiaoyu Huang, Zhen Wu, Angjoo Kanazawa, Pieter Abbeel, Carmelo Sferrazza, C Karen Liu, Rocky Duan, and Guanya Shi. Omniretarget: Interaction-preserving data generation for humanoid whole-body loco-manipulation and scene interaction. *arXiv preprint arXiv:2509.26633*, 2025.

Zhao-Heng Yin, Changhao Wang, Luis Pineda, Francois Hogan, Krishna Bodduluri, Akash Sharma, Patrick Lancaster, Ishita Prasad, Mrinal Kalakrishnan, Jitendra Malik, et al. Dexteritygen: Foundation controller for unprecedented dexterity. *arXiv preprint arXiv:2502.04307*, 2025.

Yanjie Ze, Siheng Zhao, Weizhuo Wang, Angjoo Kanazawa, Rocky Duan, Pieter Abbeel, Guanya Shi, Jiajun Wu, and C Karen Liu. Twist2: Scalable, portable, and holistic humanoid data collection system. *arXiv preprint arXiv:2511.02832*, 2025.

Yuanhang Zhang, Yifu Yuan, Prajwal Gurunath, Ishita Gupta, Shayegan Omidshafiei, Ali-akbar Agha-mohammadi, Marcell Vazquez-Chanlatte, Liam Pedersen, Tairan He, and Guanya Shi. Falcon: Learning force-adaptive humanoid loco-manipulation. *arXiv preprint arXiv:2505.06776*, 2025.

Kun Zhou, Jin Huang, John Snyder, Xinguo Liu, Hujun Bao, Baining Guo, and Heung-Yeung Shum. Large mesh deformation using the volumetric graph laplacian. In *ACM SIGGRAPH 2005 Papers*, pp. 496–503. 2005.

APPENDIX

Overview. This appendix provides implementation details and additional experimental results for **OmniInteract**. It is organized into four sections (A–D):

- **Section A: Details of Surface Contact-Aware Retargeting.** We describe motion standardization, surface correspondence, joint optimization objectives and constraints, and solver settings, and explain the complementary roles of contact-interface and near-field context representations.
- **Section B: Details of Scalable Flow-Based Interaction Policy.** We detail residual interaction expert learning and flow-based policy consolidation, including observation spaces, network architectures, rewards, training settings, and real-world deployment.
- **Section C: Experimental Details.** We introduce the interaction motion benchmark and unseen-motion evaluation set, describe baseline implementations, and specify the evaluation protocol and metrics for retargeting and interaction control.
- **Section D: Additional Experiments.** We present SCAR component ablations and qualitative retargeting comparisons, followed by residual-scale sensitivity analysis and qualitative interaction control results in simulation and the real world.

A DETAILS OF SURFACE CONTACT-AWARE RETARGETING

This section provides the implementation details of surface contact-aware retargeting (SCAR), following the formulation and notation in the main text. For humanoid $a \in \{1, 2\}$, $h^a_{t,i}$ denotes a source surface point and $x^a_i(q^a_t)$ its corresponding robot surface point. The paired configuration is $q_t = (q^1_t, q^2_t)$, and $\bar{a} = 3 - a$ denotes the opposite humanoid. We first describe motion standardization and surface binding, then detail the objectives and constraints used by the joint optimization, and finally explain the complementary roles of the two interaction representations.

A.1 MOTION STANDARDIZATION AND SURFACE CORRESPONDENCE

Standardization of motion sources. The InterHuman and Inter-X datasets provide synchronized, SMPL-X-compatible two-person motions at 30 Hz. Both actors are expressed in a common z-up world frame, preserving their relative translations and headings. To adapt the source pair to the target embodiment, we apply the same scale to both actors:

$$s = \frac{H_{robot}}{(H^1 + H^2)/2}, \tag{8}$$

where H^a is the canonical mesh height of source actor a and H_{robot} is the robot height used by the correspondence pipeline. The source points $h^a_{t,i}$ used throughout SCAR refer to the surface samples *after* this common-world standardization and scaling. We do not independently recenter or rescale the two actors, since such operations would change their demonstrated inter-body relationships. The current source decoding configuration neutralizes finger articulation; the objectives operate on the resulting body and hand surfaces.

Canonical correspondence and surface binding. We use the Dense Point Cloud Correspondence (Cao et al., 2026) adopted in the main text, together with UMR's surface-binding implementation, to associate source and robot surface samples in the canonical pose. A source sample is bound to a template triangle with fixed barycentric coordinates and follows that triangle through the motion. Its corresponding robot sample is projected onto robot geometry and bound in the geometry's local frame. Consequently,

$$x^a_i(q^a_t) = R^a_{b(i)}(q^a_t)\bar{x}^a_i + t^a_{b(i)}(q^a_t), \tag{9}$$

where $b(i)$ identifies the rigid geometry carrying the sample, and $\bar{x}^a_i$ is its fixed local position. The correspondence indices remain unchanged over time; only the posed surface locations change. This allows interaction relationships detected on the human surfaces to be evaluated at the corresponding robot surfaces through forward kinematics.

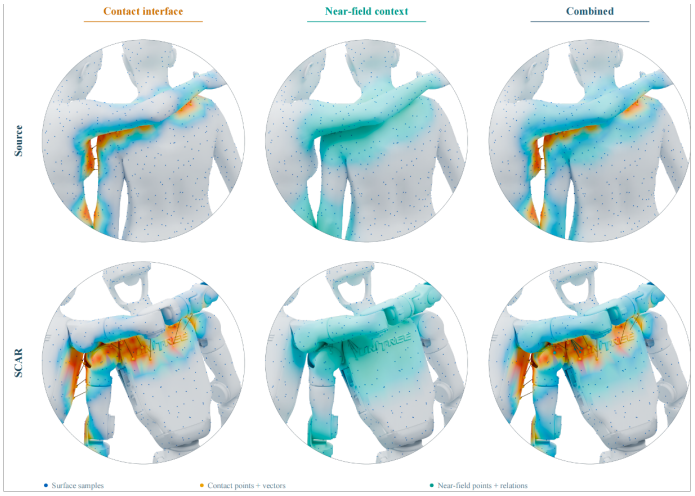


Figure 6: Visualization of contact interface and near-field context.

Surface positions and orientations. The correspondence caches used in our evaluation contain 4,096 surface slots per actor, where a slot is an indexed human–robot surface association. For individual motion tracking, we select a fixed subset $\mathcal{I}^a_{pose}$ with at most 12 slots per nonempty body region, sampled without replacement. These indices are reused throughout the motion. Contact and context construction instead query the complete surface pools before applying their own budgets. Thus, the pose-sampling limit does not restrict the number or location of candidate interaction points. Surface orientations are represented by normals: the target robot normals $\widehat{n}^a_{t,i}$ are obtained by calibrating canonical robot normals and transporting them with the source surface deformation, following UMR, rather than directly copying absolute human normals.

A.2 Optimization Details and Hyperparameters

Joint retargeting formulation. Following the main text, SCAR approximately solves the frame-wise problem

$$q_t^\star = \arg\min_{q_t \in \mathcal{C}_{phy}} \mathcal{E}_{pose}(q_t) + \lambda_c \mathcal{E}_{contact}(q_t) + \lambda_n \mathcal{E}_{context}(q_t) + \mathcal{R}_t(q_t). \tag{10}$$

Both robots' root poses and joint configurations are optimization variables. The contact and context terms couple the two configurations, while $\mathcal{C}_{phy}$ accounts for joint, velocity, floor, and collision constraints. Below, we specify each term and the numerical realization of these constraints.

Individual motion preservation. The pose objective matches the selected surface positions and normal orientations:

$$\begin{aligned}\mathcal{E}_{pose}(q_t) = \sum_{a\in\{1,2\}} \sum_{i\in\mathcal{I}^a_{pose}} \Big(& w_i^{a,pos} \| x_i^a(q_t^a) - h_{t,i}^a \|_2^2 \\ & + w_i^{a,nor} \| n_i^a(q_t^a) - \widehat{n}_{t,i}^a \|_2^2 \Big).\end{aligned} \tag{11}$$

Here $w_i^{a,pos}$ and $w_i^{a,nor}$ are region-dependent weights inherited from UMR. They are distinct from the neighborhood weights $\omega_{t,ij}$ in the interaction coordinates.

Contact-Interface Objective. We construct $\mathcal{C}_t$ from the union of directed nearest-neighbor candidates in both directions, $1 \to 2$ and $2 \to 1$, together with the retained pairs from the previous frame. Recall the demonstrated contact offset $d^h_{t,ij} = h^1_{t,i} - h^2_{t,j}$, and let $\ell_{t,ij} = \|d^h_{t,ij}\|_2$ be its length. A new candidate is eligible if $\ell_{t,ij} \le \tau_{in}$; a previously retained pair remains eligible while $\ell_{t,ij} \le \tau_{out}$, with $\tau_{in} < \tau_{out}$. This hysteresis reduces contact switching near the threshold. Eligible pairs are grouped by their two body regions, ordered by distance within each group, and selected in round-robin order across groups until the contact budget is reached.

The normalized source weights in the main-text contact objective are

$$\widetilde{\alpha}_{t,ij} = \mathrm{clip}\left(\frac{\tau_{out} - \ell_{t,ij}}{\tau_{out} - \tau_{in}},\, 0.1,\, 1\right), \qquad \alpha_{t,ij} = \frac{\widetilde{\alpha}_{t,ij}}{\sum_{(u,v)\in\mathcal{C}_t} \widetilde{\alpha}_{t,uv}}. \tag{12}$$

Using $d^r_{t,ij}(q_t) = x^1_i(q^1_t) - x^2_j(q^2_t)$, we minimize

$$\mathcal{E}_{contact}(q_t) = \sum_{(i,j)\in\mathcal{C}_t} \alpha_{t,ij} \| d^r_{t,ij}(q_t) - d^h_{t,ij} \|_2^2. \tag{13}$$

Each contact pair contributes an explicit relative-vector residual, penalizing both separation and tangential misalignment without averaging errors across the interface. The target is the demonstrated offset, not a forced zero vector, so small source clearances are retained. Both endpoints are movable: if $J^a_{t,i}$ denotes the surface-point Jacobian with respect to robot a's tangent coordinates, the contact residual has Jacobian $[J^1_{t,i}, -J^2_{t,j}]$. Thus, contact preservation can be achieved through coordinated changes to both robots. It remains a soft geometric objective, not a contact-force or exact zero-gap constraint.

Near-Field Context Objective. For each source point not retained as a contact endpoint, define its distance to the opposite surface as $\ell^a_{t,i} = \min_j \|h^a_{t,i} - h^{\bar{a}}_{t,j}\|_2$. Points within radius r_n are eligible context centers. We jointly rank candidates from both actors by $\ell^a_{t,i}$ and retain them under a shared budget. $\mathcal{S}^a_t$ denotes the retained centers belonging to actor a; hence $|\mathcal{S}^1_t| + |\mathcal{S}^2_t| \le B_n$.

For each center $i \in \mathcal{S}^a_t$, we find the nearest point on actor $\bar{a}$ and its one-ring neighbors in the cached correspondence topology. Among these points, the closest at most K lying within r_n of the center form $\mathcal{N}_t(i)$. Actor identity is implicit in this notation: all $j \in \mathcal{N}_t(i)$ belong to $\bar{a}$. The normalized source-geometric weights are

$$\omega_{t,ij} = \frac{\exp\!\big(-\|h^a_{t,i} - h^{\bar{a}}_{t,j}\|_2^2/(2\sigma_n^2)\big)}{\sum_{u\in\mathcal{N}_t(i)} \exp\!\big(-\|h^a_{t,i} - h^{\bar{a}}_{t,u}\|_2^2/(2\sigma_n^2)\big)}. \tag{14}$$

As in the main text, the interaction coordinate of a human or robot point p_i is

$$\mathcal{L}_t(p_i) = p_i - \sum_{j\in\mathcal{N}_t(i)} \omega_{t,ij} p_j, \qquad \sum_{j\in\mathcal{N}_t(i)} \omega_{t,ij} = 1. \tag{15}$$

This coordinate measures the displacement from a weighted neighborhood centroid on the opposite body to the center point. For robot coordinates, the center is evaluated at q^a_t and all neighbors at $q^{\bar{a}}_t$; therefore $\mathcal{L}_t(x^a_i(q^a_t))$ depends on the paired configuration despite being indexed by its center. The neighborhoods and weights are constructed from the source geometry and held fixed throughout the current frame's optimization. Contact endpoints are excluded only as context *centers*; they can still serve as neighbors.

The center-proximity weights are normalized across both humanoids:

$$\widetilde{\beta}^a_{t,i} = \exp\!\left(-\frac{(\ell^a_{t,i})^2}{2r_n^2}\right), \qquad \beta^a_{t,i} = \frac{\widetilde{\beta}^a_{t,i}}{\sum_{b\in\{1,2\}} \sum_{u\in\mathcal{S}^b_t} \widetilde{\beta}^b_{t,u}}. \tag{16}$$

The context objective is then

$$\mathcal{E}_{context}(q_t) = \sum_{a\in\{1,2\}} \sum_{i\in\mathcal{S}^a_t} \beta^a_{t,i} \left\| \mathcal{L}_t(x^a_i(q^a_t)) - \mathcal{L}_t\big(h^a_{t,i}\big) \right\|_2^2. \tag{17}$$

This term preserves the surrounding surface arrangement through local cross-body coordinates rather than requiring every nearby pair to match independently. If $\mathcal{C}_t$ or both context-center sets are empty, the corresponding interaction term is set to zero without evaluating its weight normalization. Fig. 6 shows the visualization of contact interface and near-field context.

Temporal and stance-foot regularization. We instantiate the main-text regularizer as

$$\mathcal{R}_t(q_t) = \lambda_{temp} \left\| \frac{q_t \ominus q^\star_{t-1}}{\Delta t} \right\|_2^2 + \lambda_{foot} \mathcal{E}_{foot}(q_t), \tag{18}$$

where $\ominus$ denotes a tangent-space configuration difference. The first term penalizes generalized velocity magnitude and is omitted at the first frame. A source foot is marked as supporting when its foot joint lies less than 8 cm above that actor's lowest foot-joint height over the clip and moves slower than 0.25 m/s. At each support onset, up to eight robot foot-surface points are recorded from the frame initialization as fixed anchors $\bar{x}_{f,i}^{foot}$. For the active supporting feet $f \in \mathcal{B}_t$, with nonempty anchor-index sets $\mathcal{I}_f^{foot}$, we use

$$\mathcal{E}_{foot}(q_t) = \sum_{f \in \mathcal{B}_t} \frac{1}{|\mathcal{I}_f^{foot}|} \sum_{i \in \mathcal{I}_f^{foot}} \|x_{f,i}(q_t) - \bar{x}_{f,i}^{foot}\|_2^2. \tag{19}$$

This softly discourages three-dimensional stance-foot drift; it does not impose a hard no-slip constraint.

Constraints and SQP subproblems. We implement SQP using damped Gauss–Newton quadratic subproblems. Let $e_t(q_t)$ stack the weighted residuals of Eq. (10), and let J_t be their tangent-space linearization. At each inner iteration, we solve

$$\min_{\Delta q_t,\, \xi \geq 0} \frac{1}{2}\|e_t(q_t) + J_t \Delta q_t\|_2^2 + \frac{\mu}{2}\|\Delta q_t\|_2^2 + \frac{\lambda_{slack}}{2}\|\xi\|_2^2, \tag{20}$$

subject to the following conditions, which numerically realize $\mathcal{C}_{phy}$.

- **Joint and velocity limits.** The actuated joint coordinates are bounded by the robot model's joint ranges. Componentwise tangent bounds additionally enforce $-v_{\max}\Delta t \leq (q_t \ominus q_{t-1}^{\star}) + \Delta q_t \leq v_{\max}\Delta t$ in the local approximation, using separate limits for root translation, root rotation, and actuated joints. Velocity bounds are omitted at the first frame.
- **Floor and collision constraints.** For each active geometric row, let $\phi_\ell(q_t)$ denote a sampled surface height or a signed collision-geometry distance, with margin m_ℓ. We impose $\phi_\ell(q_t) + J_{\phi_\ell}\Delta q_t \geq m_\ell - \xi_\ell$. The nonnegative slack ξ_ℓ carries a large penalty, allowing recovery from initially violated configurations. Collision checks include self- and inter-robot pairs, excluding same-actor body pairs within three kinematic-tree edges.
- **Trust region and tolerance envelope.** Each update satisfies $\|\Delta q_t\|_\infty \leq \delta$. Protected geometric rows also obey the hard, slack-free bound $\phi_\ell(q_t) + J_{\phi_\ell}\Delta q_t \geq m_\ell - \varepsilon_{env}$, where $\varepsilon_{env} = 0.9\varepsilon$ and ε is the acceptance tolerance. All active geometric rows are protected when the current iterate passes the feasibility check; otherwise, only rows already inside this tighter envelope are protected.

Initialization and nonlinear acceptance. The first frame uses source-derived root positions and headings with a fixed bent-arm initialization. Subsequent frames are warm-started from the previous solution $q_{t-1}^{\star}$. After each QP solve, we evaluate $q_t^{+} = q_t \oplus \zeta\Delta q_t$ with $\zeta \in \{1, 1/2, \ldots, 1/64\}$, where $\oplus$ integrates the floating-base rotations on their configuration manifold. When the current iterate is feasible, a candidate must stay within the acceptance tolerance and not increase the tracking objective in Eq. Eq. (10). When it is infeasible, the rule prioritizes decreasing violation, permitting objective descent when violation is unchanged within numerical tolerance. If no step is accepted, the current iterate is retained and the frame solve terminates. Final validity flags are saved alongside the trajectories; invalid frames are not silently removed.

Implementation and hyperparameters. We use MuJoCo (Todorov et al., 2012) for forward kinematics, surface Jacobians, and collision queries, and UMR's (Cao et al., 2026) Clarabel interface to solve the quadratic subproblems. The paired 29-DoF G1 configuration has 72 stored coordinates, including root quaternions, and a 70-dimensional tangent update. The settings in Tables 4 and 5 are shared by the evaluated InterHuman and Inter-X configurations. Contact weights $\alpha_{t,ij}$ sum to one over $\mathcal{C}_t$; neighbor weights $\omega_{t,ij}$ sum to one within each neighborhood; center weights $\beta_{t,i}^{a}$ sum to one across the retained centers of both actors. The pose objective instead sums its region-weighted residuals.

The geometric constraints use floor and collision margins of -2 and -3 mm, respectively, and a 5 mm acceptance tolerance relative to these margins. Their effective nonlinear bounds are therefore $z \geq -7$ mm and $d_{signed} \geq -8$ mm; the hard envelope leaves a further 0.5 mm buffer. These are

Table 4: SCAR interaction and optimization settings. Context centers share one budget across the two humanoids.

Parameter	Value
Contact thresholds τ_{in}/τ_{out}	$0.02/0.03$ m
Contact-pair budget	128 per frame
Context-center budget B_n	128 total across both actors
Context radius r_n / Gaussian width σ_n	$0.15/0.06$ m
Maximum neighbors K	6 per center
Pose slots per region / sampling seeds	12 / $0, 1$ for actors $1, 2$
λ_c/λ_n	$2000/100$
$\lambda_{temp}/\lambda_{foot}$	$0.02/50$
Damping μ / slack penalty λ_{slack}	$0.01/10^4$
Step bound δ	0.12 m or 0.12 rad by coordinate
Joint / root angular speed limits	$8/4$ rad/s
Root linear speed limit	2 m/s
Floor / collision activation distance	$0.04/0.04$ m
Maximum active floor samples	128 per actor
First-frame / later-frame SQP iterations	$30/8$
Backtracking scales ζ	2^{-k}, $k = 0, \ldots, 6$

Table 5: Region-dependent pose weights inherited from UMR. Left and right regions share weights. `Arm` and `UpperArm` denote distinct regions in the 19-region configuration.

Region	Position w^{pos}	Normal w^{nor}
Torso	20	5
Hips	20	0
Head	1	1
Hand, Foot	10	1
Arm, ForeArm, UpLeg, Leg	10	5
Shoulder	0	0.5
UpperArm	0	0.1

numerical tolerances, not exact non-penetration guarantees. The current implementation compares the numerical maximum of geometric margin violations and componentwise speed excess against the same tolerance. Its validity flag is consequently an implementation-level diagnostic rather than a physical-executability certificate. SCAR generates coupled kinematic references; sustaining their execution under dynamics is the role of the interaction-policy stage of OmniInteract.

A.3 Surface Interaction Graph and Complementary Roles

SCAR's interaction representation separates *where the surfaces should interact* from *how the surrounding surfaces should be arranged*. Contact-interface vectors connect individual pairs in $\mathcal{C}_t$ and can be displayed in orange–gold. They directly visualize $d^h_{t,ij}$ and $d^r_{t,ij}(q_t)$ on the source and robot surfaces. Near-field relations can be displayed in teal: neighborhood spokes identify the points in $\mathcal{N}_t(i)$, and the arrow from their weighted centroid to the center depicts $\mathcal{L}_t(p_i)$. The spokes contribute to one weighted interaction coordinate; they are not separate pairwise objectives. Surface heatmaps highlight proximity to the displayed graph endpoints and should not be interpreted as forces or optimization residuals.

In an embrace, for example, contact pairs specify the palm–back and arm–torso interfaces, while context centers describe the arrangement of nearby limbs and torso regions. Matching the contact pairs alone leaves this surrounding arrangement underconstrained. Conversely, a neighborhood's weighted coordinate may remain similar even when individual contact offsets deviate. Combining $\mathcal{E}_{contact}$ and $\mathcal{E}_{context}$ addresses these two complementary requirements within the same paired configuration, rather than independently retargeting two motions and subsequently repairing their interaction. The resulting trajectories provide synchronized robot-space references for OmniInteract's interaction-policy training. Algorithm 1 summarizes the retargeting procedure.

Algorithm 1 Surface Contact-Aware Retargeting (SCAR)

Require: Synchronized source motions, robot models, canonical surface correspondences
Ensure: Paired configurations $\{q_t^\star\}_{t=1}^T$ and validity diagnostics

Phase 1: Shared surface preparation
Standardize both actors; apply the shared scale s
Construct $h_{t,i}^a$, bindings for $x_i^a(q_t^a)$, and normal targets
Fix $\mathcal{I}_{pose}^a$; identify source foot-support intervals
Initialize q_{init} and $\mathcal{C}_0 \leftarrow \varnothing$
Phase 2: Joint frame-wise SQP
for $t = 1, \ldots, T$ **do**
    Initialize $q_t \leftarrow q_{t-1}^\star$; use q_{init} if $t = 1$
    Construct $\mathcal{C}_t$ with bidirectional search, hysteresis, and region balancing
    Compute contact weights $\alpha_{t,ij}$
    Jointly retain $\mathcal{S}_t^1, \mathcal{S}_t^2$ with $|\mathcal{S}_t^1| + |\mathcal{S}_t^2| \le B_n$
    Build $\mathcal{N}_t(i)$, $\omega_{t,ij}$, $\beta_{t,i}^a$, and source coordinates $\mathcal{L}_t(h_{t,i}^a)$
    Update stance anchors; keep the source-defined interaction targets fixed
    for each SQP iteration **do**
        Evaluate the paired residuals, Jacobians, and active constraints
        Solve Eq. (20) with bounds and the hard tolerance envelope
        if the solve fails or $\|\Delta q_t\|_\infty < 10^{-6}$ **then**
            Terminate the current frame solve
        end if
        Backtrack over $q_t \oplus \zeta \Delta q_t$ using nonlinear acceptance checks
        if a candidate is accepted **then**
            Update q_t to the accepted candidate
        else
            Retain q_t and terminate the current frame solve
        end if
    end for
    Store $q_t^\star \leftarrow q_t$ and its validity diagnostics
end for
return Paired robot trajectories and diagnostics

B Details of Scalable Flow-Based Interaction Policy

B.1 Residual Interaction Expert Learning

Partner-aware residual control. SCAR provides synchronized robot-space kinematic references for the two humanoids. We learn to execute these references under simulated dynamics by augmenting the frozen SONIC tracker π_{base} (Luo et al., 2026) with a task-specific residual expert π_{res}^k. For each interaction task k, the expert is shared by both humanoids and evaluated using their respective observations. Following Eq. (5), the final action $\mathcal{A}^a$ combines the frozen tracking action $\mathcal{A}_{base}^a$ with the partner-dependent correction $\Delta\mathcal{A}^{\bar{a}\to a}$, where $a \in \{1, 2\}$ and $\bar{a} = 3 - a$.

The base and residual actions use consistent joint ordering and normalization. We set $\alpha_{res} = 0.3$ in normalized action coordinates and zero-initialize the residual output layer. Consequently, the initial deterministic controller reproduces the frozen tracker, while subsequent optimization learns bounded interaction-specific corrections. Joint targets are updated at 50 Hz and tracked by implicit PD actuators in a 200 Hz physics simulation. Below, n denotes a control step.

Observation space. The observation groups $\mathcal{O}_{his}^a$, $\mathcal{O}_{ego}^a$, and $\mathcal{O}_{fur}^a$ retain the definitions in the main text. They are implemented as two temporal streams together with a current-frame input to the action decoder.

- **History, $\mathcal{O}_{his}^a \in \mathbb{R}^{20\times 239}$.** The history covers control steps $n-19, \ldots, n$. Each entry contains ego tracking features (64), ego proprioception and the previous complete action (93), partner-state features (64), and reference and measured relative torso poses (18). Ego tracking comprises reference joint positions and velocities (29 each) and torso orientation error (6). Proprioception comprises projected gravity (3), base angular velocity (3), centered joint positions and velocities (29 each), and the previous action (29).

Table 6: Network architecture and MAPPO hyperparameters for residual interaction expert learning.

Hyperparameter	Value
Architecture	
History / future window length	20 / 20
Input projection, history / future	$239 \to 60$ / $93 \to 60$
Temporal CNN channels	$60 \to 40 \to 20$
CNN kernel sizes / strides	$(6, 4)$ / $(2, 2)$
Temporal pooling / output projection	Average / $20 \to 64$
Residual decoder input / output	370 / 29
Actor and critic hidden dimensions	$[512, 256, 128]$
Hidden activation	RELU
Residual output initialization / scale	Zero / 0.3
MAPPO training	
Rollout steps per environment	24
Learning epochs / mini-batches	5 / 4
Optimizer / initial learning rate	Adam / 10^{-4}
Learning-rate schedule / target KL	Adaptive / 0.02
PPO clip parameter	0.1
Entropy / value-loss coefficient	0.001 / 1.0
Discount factor γ / GAE parameter λ	0.99 / 0.95
Clipped value loss	Enabled
Maximum gradient norm	0.5
Initial action standard deviation	0.1

- **Ego-centric partner information, $\mathcal{O}^a_{ego} \in \mathbb{R}^{82}$.** This block contains the partner's reference and measured joint positions (29 each), torso orientation error (6), and reference and measured relative torso poses (9 each). Each relative pose consists of translation (3) and rotation (6), expressed in the corresponding ego torso frame. This block is included in the latest history entry rather than processed by an additional temporal encoder.
- **Future references, $\mathcal{O}^a_{fur} \in \mathbb{R}^{20\times 93}$.** The next 20 reference frames provide ego joint positions and velocities (29 each), future torso orientation relative to the current torso (6), and partner joint positions (29). Desired relative-pose information is supplied separately through $\mathcal{O}^a_{ego}$ and its history.

Network architecture and optimization. Two independent temporal 1D CNNs encode the normalized history and future streams into 64-dimensional embeddings. A learned three-dimensional projection of the history embedding is appended to that embedding, yielding 67 dimensions; this projection has no separate auxiliary supervision. The residual decoder concatenates the history embedding (67), future embedding (64), and latest history entry (239), including $\mathcal{O}^a_{ego}$, into a 370-dimensional input.

We train the shared expert with MAPPO under centralized training and decentralized execution. The centralized critic receives both humanoids' tracking states and temporal inputs. Each actor evaluation uses its own observations, communicated partner information, and paired motion references. During training, a diagonal Gaussian is centered on the concatenated corrected actions of both humanoids, giving a 58-dimensional joint action distribution. Thus, α_{res} bounds the deterministic residual correction, not the sampled exploration noise. PPO uses the joint action likelihood ratio and updates the residual expert and critic while keeping SONIC frozen. The core settings are summarized in Table 6.

Interaction-aware rewards. We train the residual experts with a shared reward that balances interaction geometry, contact timing, individual tracking, and control regularization:

$$r_n = \Delta t_{ctrl} \left(r_{inter} + \sum_{a=1}^{2} r^a_{track} + r_{reg} \right), \qquad \Delta t_{ctrl} = 0.02\,\mathrm{s}. \tag{21}$$

Each component is the weighted sum of the corresponding block in Table 7, including the negative regularization weights. Tracking rewards are evaluated separately for each humanoid; interaction rewards and penalties are evaluated once for the pair. The action-rate penalty acts on the complete applied action $\mathcal{A}_n^a$, rather than the residual alone.

Superscripts sim and ref denote simulated quantities and targets from the synchronized retargeted robot trajectories. We use $\mathbf{p}_{a,b}$ and $Q_{a,b}$ for body position and unit quaternion, with the torso as the anchor anc. The relative quantities are $\Delta\mathbf{p} = \mathbf{p}_{1,anc} - \mathbf{p}_{2,anc}$ in world coordinates and $\Delta Q = Q_{1,anc}^{-1} \otimes Q_{2,anc}$, where $\otimes$ denotes quaternion multiplication and d_Q measures the relative rotation angle between the represented orientations. A tilde denotes a reference aligned to the simulated torso's horizontal position using the yaw component of the reference-to-simulated torso rotation difference, while preserving reference torso height. Velocity errors are evaluated in world coordinates. The quantities $\mathbf{v}$ and $\boldsymbol{\omega}$ denote linear and angular velocities, and the superscript xy denotes horizontal components.

The tracked set $\mathcal{B}_{track}$ contains the pelvis, torso, and bilateral hip-roll, knee, ankle-roll, shoulder-roll, elbow, and wrist-yaw links. The topology set $\mathcal{P}_{topo}$ contains the five corresponding cross-humanoid pairs at the left wrist, right wrist, torso, left ankle, and right ankle. For each pair, $\ell_{uv} = \|\mathbf{p}_{1,u} - \mathbf{p}_{2,v}\|_2$ and $\rho_{uv} = 1/\max(\ell_{uv}^{ref}, 0.1)$.

The set $\mathcal{P}_{cont}$ contains the two wrist pairs. Their binary proximity labels c_{uv}^{sim} and c_{uv}^{ref} use a distance threshold of $0.18\,\mathrm{m}$. For the two feet $\mathcal{F}$, reference support $s_{a,f}^{ref}$ is determined by ankle height below $0.06\,\mathrm{m}$, and simulated support $s_{a,f}^{sim}$ by contact-force magnitude above $1\,\mathrm{N}$ in any of the three stored sensor samples. The contact reward jointly exponentiates the separately averaged wrist-proximity and foot-support errors.

Finally, $q_{a,j}^{act}$ denotes an actuated joint position, $q_{a,j}^{min}$ and $q_{a,j}^{max}$ are soft joint limits, and $[x]_+ = \max(x, 0)$. The set $\mathcal{U}$ contains monitored bodies other than wrist and ankle links; $\mathbf{F}_{a,u,h}^{sim}$ denotes the measured net contact force at history index h, and $\boldsymbol{\tau}_a$ denotes joint torques. This fixed contact penalty is a soft regularizer, not a reference-conditioned exclusion rule. These body-level execution rewards complement SCAR's kinematic surface objectives without reusing its surface graph or contact thresholds.

Domain randomization. We randomize contact properties, torso center of mass, and default joint offsets independently for the two humanoids. At reset, both humanoids are initialized from a common reference frame and receive the same sampled root-pose and root-velocity perturbations, together with independent joint-position offsets. We additionally randomize the delay of measured partner states, using the same sampled delay in both communication directions while leaving reference features undelayed. These randomizations expose the residual expert to physical parameter variation, initialization errors, and stale partner information. Table 8 lists the ranges. Continuous parameters are sampled uniformly within their intervals, and communication delay is sampled uniformly from the listed discrete values.

B.2 Flow-Matching Interaction Policy Learning

Shared policy and residual supervision. We consolidate the task-specific interaction experts into a shared flow policy while retaining the frozen SONIC tracker. The two humanoids share student parameters θ, but each policy evaluation produces only the ego humanoid's 29-dimensional residual. For task k and humanoid a, the supervision target is the deterministic correction from Eq. (5):

$$\Delta\mathcal{A}_{expert}^{a} = \Delta\mathcal{A}^{\bar{a}\to a} \in \mathbb{R}^{29}, \qquad a \in \{1, 2\}, \quad \bar{a} = 3 - a. \tag{22}$$

Each target includes α_{res} but excludes exploration noise. Consistent joint ordering, action normalization, and the shared frozen motor prior provide a common residual supervision space across tasks. Samples from both humanoids jointly update the same policy, while supervision and prediction remain per-humanoid.

Ego-conditioned velocity network. Each policy evaluation organizes the paired observation windows from the ego humanoid's perspective:

$$\hat{\mathcal{O}}^{a} = \left(\mathcal{O}_{his}^{a}, \mathcal{O}_{fur}^{a}, \mathcal{O}_{his}^{\bar{a}}, \mathcal{O}_{fur}^{\bar{a}}\right). \tag{23}$$

Table 7: Reward terms and weights for residual interaction expert learning.

Category	Reward term	Equation	Weight
Interaction	Relative anchor position	$\exp\left(-\frac{\|\Delta\mathbf{p}^{sim}-\Delta\mathbf{p}^{ref}\|_2^2}{0.5^2}\right)$	1.5
	Relative anchor orientation	$\exp\left(-\frac{d_Q(\Delta Q^{sim},\Delta Q^{ref})^2}{0.75^2}\right)$	0.5
	Interaction topology	$\exp\left(-\frac{\sum_{(u,v)\in\mathcal{P}_{topo}}\rho_{uv}(\ell_{uv}^{sim}-\ell_{uv}^{ref})^2}{0.35^2\,\lvert\mathcal{P}_{topo}\rvert}\right)$	1.5
	Contact/support timing	$\exp\left(-\frac{\sum_{(u,v)\in\mathcal{P}_{cont}}\lvert c_{uv}^{sim}-c_{uv}^{ref}\rvert}{\lvert\mathcal{P}_{cont}\rvert}-\frac{\sum_{a=1}^{2}\sum_{f\in\mathcal{F}}\lvert s_{a,f}^{sim}-s_{a,f}^{ref}\rvert}{2\lvert\mathcal{F}\rvert}\right)$	1.0
Motion tracking	Anchor position	$\exp\left(-\frac{\|\mathbf{p}_{a,anc}^{sim}-\mathbf{p}_{a,anc}^{ref}\|_2^2}{0.3^2}\right)$	0.25
	Anchor orientation	$\exp\left(-\frac{d_Q(Q_{a,anc}^{sim},Q_{a,anc}^{ref})^2}{0.4^2}\right)$	0.25
	Aligned body position	$\exp\left(-\frac{\sum_{b\in\mathcal{B}_{track}}\|\mathbf{p}_{a,b}^{sim}-\tilde{\mathbf{p}}_{a,b}^{ref}\|_2^2}{0.3^2\,\lvert\mathcal{B}_{track}\rvert}\right)$	0.5
	Aligned body orientation	$\exp\left(-\frac{\sum_{b\in\mathcal{B}_{track}}d_Q(Q_{a,b}^{sim},\tilde{Q}_{a,b}^{ref})^2}{0.4^2\,\lvert\mathcal{B}_{track}\rvert}\right)$	0.5
	Body linear velocity	$\exp\left(-\frac{\sum_{b\in\mathcal{B}_{track}}\|\mathbf{v}_{a,b}^{sim}-\mathbf{v}_{a,b}^{ref}\|_2^2}{1.0^2\,\lvert\mathcal{B}_{track}\rvert}\right)$	0.5
	Body angular velocity	$\exp\left(-\frac{\sum_{b\in\mathcal{B}_{track}}\|\boldsymbol{\omega}_{a,b}^{sim}-\boldsymbol{\omega}_{a,b}^{ref}\|_2^2}{3.14^2\,\lvert\mathcal{B}_{track}\rvert}\right)$	0.5
Regularization	Action rate	$\sum_{a=1}^{2}\|\mathcal{A}_n^a-\mathcal{A}_{n-1}^a\|_2^2$	-0.3
	Foot slip	$\sum_{a=1}^{2}\sum_{f\in\mathcal{F}}s_{a,f}^{sim}\|\mathbf{v}_{a,f}^{xy,sim}\|_2^2$	-0.5
	Joint limits	$\sum_{a=1}^{2}\sum_j\left([q_{a,j}^{act}-q_{a,j}^{max}]_+ + [q_{a,j}^{min}-q_{a,j}^{act}]_+\right)$	-10.0
	Undesired contacts	$\sum_{a=1}^{2}\sum_{u\in\mathcal{U}}\mathbf{1}\left[\max_{h\in\{0,1,2\}}\|\mathbf{F}_{a,u,h}^{sim}\|_2>1\,\mathrm{N}\right]$	-0.1
	Torque	$\sum_{a=1}^{2}\|\boldsymbol{\tau}_a\|_2^2$	-10^{-4}

Table 8: Domain randomization for residual interaction expert learning.

Category	Parameter	Range / setting
Dynamics and contact	Static friction	$[0.3, 1.6]$
	Dynamic friction	$[0.3, 1.2]$
	Restitution	$[0, 0.5]$
	Torso CoM offset	$x:\ \pm0.025$ m; $y, z:\ \pm0.05$ m
	Default joint offset	±0.01 rad
Initial state	Root position offset	$x, y:\ \pm0.05$ m; $z:\ \pm0.01$ m
	Root orientation offset	Roll/pitch: ±0.1 rad; yaw: ±0.2 rad
	Root linear velocity offset	$x, y:\ \pm0.5$ m/s; $z:\ \pm0.2$ m/s
	Root angular velocity offset	Roll/pitch: ±0.52 rad/s; yaw: ±0.78 rad/s
	Joint position offset	±0.1 rad, clipped to soft joint limits
Communication	Peer measurement delay	$\{0, 1, 2, 3\}$ control steps (0–60 ms)

The observation groups follow Appendix B.1. The current partner-aware block $\mathcal{O}_{ego}^a$ is already contained in the history and is not introduced as an additional token stream. Separate history and future preprocessing MLPs, with weights shared across roles, convert the four 20-entry windows into 80 tokens. Their representations retain temporal, stream, and ego/partner identities.

A Transformer encodes these tokens, followed by global average pooling and a linear projection to a conditioning embedding. This embedding is concatenated with embeddings of the ego humanoid's noisy residual and flow time. An MLP then predicts the conditional velocity $v_\theta(\Delta\mathcal{A}_t^a, t, \hat{\mathcal{O}}^a) \in \mathbb{R}^{29}$. The same network is evaluated for the other humanoid after exchanging the ego and partner roles. Thus, paired interaction context conditions each individual correction without requiring a joint action-output head. Table 9 summarizes the architecture.

Student-state supervision. We alternate paired student rollouts and expert labeling using DAgger-style supervision (Ross et al., 2011). At each visited interaction state, the shared policy is evaluated for both humanoids, and each generated residual is added to the corresponding frozen SONIC action. Before the simulator advances, task k selects the expert that labels each humanoid using the same pre-action observations, histories, and reference phase. Expert parameters and normalization statistics remain fixed.

Each visited paired state therefore provides two training examples, $(\hat{\mathcal{O}}^a, \Delta\mathcal{A}^a_{expert})$, one for each humanoid. Both examples contribute to the same student update. The labels specify corrections at student-visited states rather than actions from separately executed expert trajectories. Task identity routes the teacher but is not an input to the student.

Conditional flow matching. For each humanoid, we sample $\boldsymbol{\epsilon}^a \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{29})$ and construct

$$\Delta\mathcal{A}_t^a = (1-t)\Delta\mathcal{A}^a_{expert} + t\boldsymbol{\epsilon}^a, \qquad t \in [0,1]. \tag{24}$$

Flow time t is sampled from a Beta distribution and is distinct from the physical control-step index n. The endpoint $t = 0$ corresponds to the expert residual and $t = 1$ to Gaussian noise. Applying Eq. (6) to both humanoids yields the shared-parameter objective

$$\mathcal{L}_{fm}(\theta) = \mathbb{E}\left[\frac{1}{2}\sum_{a=1}^{2}\left\| v_\theta(\Delta\mathcal{A}_t^a, t, \hat{\mathcal{O}}^a) - \left(\boldsymbol{\epsilon}^a - \Delta\mathcal{A}^a_{expert}\right)\right\|_2^2\right]. \tag{25}$$

The expectation covers labeled student states, flow times, and noise. This objective jointly updates θ using both humanoids' data, while each velocity prediction remains 29-dimensional. Only the flow policy is optimized; SONIC and the expert bank remain frozen. The updated shared policy then collects the next round of paired interaction states. At deployment, each humanoid runs a copy of the same trained policy on top of its own SONIC tracker. For humanoid a, we hold $\hat{\mathcal{O}}^a$ fixed, initialize $\Delta\mathcal{A}_1^a$ with Gaussian noise, and integrate from $t = 1$ to $t = 0$ using N explicit Euler steps:

$$\Delta\mathcal{A}^a_{t_{j+1}} = \Delta\mathcal{A}^a_{t_j} - \frac{1}{N} v_\theta\left(\Delta\mathcal{A}^a_{t_j}, t_j, \hat{\mathcal{O}}^a\right), \qquad t_j = 1 - \frac{j}{N}, \tag{26}$$

where $j = 0, \ldots, N-1$. The negative step traverses the learned expert-to-noise velocity field in reverse. Each policy evaluation generates only the ego humanoid's residual, which is applied as

$$\mathcal{A}^a = \mathcal{A}^a_{base} + \Delta\mathcal{A}^a_0. \tag{27}$$

No additional factor of α_{res} is applied because the supervision targets already include this scale. Independent deployment refers to separate policy evaluations and action application, while partner information remains available through the interaction-conditioning interface. The shared flow policy therefore replaces the task-specific expert bank while preserving the additive control interface used during expert learning.

Real-world deployment. We deploy OmniInteract on two Unitree G1 humanoids (approximately 1.3 m tall, 29 DoF), with each robot independently executing the frozen SONIC tracker and the shared residual flow policy onboard. To construct partner-aware observations without privileged simulator states, we employ a coarse-to-fine localization system combining LiDAR and IMU measurements. Point-LIO (He et al., 2023) provides LiDAR–inertial odometry at 10 Hz. For global initialization, GeoTransformer registers the current scan against a pre-built point-cloud map, establishing a common spatial frame for both robots. During execution, a C++ relocalization node refines the pose using Generalized Iterative Closest Point (GICP) (Segal et al., 2009), and a Kalman filter fuses the odometry with global pose corrections to obtain smooth state estimates and mitigate accumulated drift. The robots exchange their estimated global poses and peer-state observations through

Table 9: Architecture and learning specification of the shared per-humanoid residual flow policy.

Component	Specification
	Architecture
History preprocessing MLP	$239 \to 256 \to 256$; ReLU
Future preprocessing MLP	$93 \to 256 \to 256$; ReLU
Preprocessing weight sharing	Across roles; separate history/future weights
Token sequence per policy evaluation	2 roles $\times 2$ streams $\times 20$ entries
Token width / count	256 / 80
Token identity information	Temporal position, stream, and ego/partner role
Condition encoder	Transformer; global average pooling
Condition projection	Linear; 1024-dimensional output
Noisy residual embedding	Linear $29 \to 1024$
Flow-time embedding	Sinusoidal; 1024 dimensions
Feature fusion	Concatenation; 3072 dimensions
Velocity decoder hidden / output	$[2048, 2048, 2048]$ / 29
Policy parameter sharing	One parameter set for both humanoids
	Learning and sampling
Training data	Both humanoids' samples from paired rollouts
Teacher target	Deterministic scaled 29-D residual per humanoid
Task-ID conditioning	Teacher routing only; excluded from student input
Trainable / frozen components	Shared flow network / SONIC and residual experts
Noise distribution	$\mathcal{N}(\mathbf{0}, \mathbf{I}_{29})$ per humanoid
Flow-time sampling	Beta distribution on $[0, 1]$
Velocity target	$\boldsymbol{\epsilon}^a - \Delta\mathcal{A}^a_{expert}$
Loss aggregation	Mean of both humanoids' flow-matching losses
Sampling direction / solver	$t : 1 \to 0$ / N explicit Euler steps
Deployment	Shared weights; separate per-humanoid inference
Action composition	Each residual added to its own SONIC action

Lightweight Communications and Marshalling (LCM) (Huang et al., 1998). Received partner information is transformed into the ego-centric frame and incorporated into the observation histories, while the paired SCAR references advance with a common motion phase. The high-level policy are performed onboard. We utilize ONNX Runtime to execute the trained policies, achieving an inference latency of less than 8 ms, ensuring a stable 50 Hz control loop. Each robot constructs its conditioning observations from the estimated states, evaluates its frozen tracker, and integrates the flow policy to obtain its own 29-dimensional residual. This correction is added to the base action following Eq. (27), and converted into desired joint positions. The low-level interface publishes motor targets and proportional–derivative gains at 500 Hz for joint-level tracking.

C Experimental Details

C.1 Interactoin Motion Benchmark

Interaction Motion Library For retargeting, we evaluate on the full motion sets from the InterHuman (Liang et al., 2024) and Inter-X (Xu et al., 2024) datasets. For control, we select a diverse subset of interaction motions from both datasets, as summarized in Table 10.

Unseen Motion Set. In addition to interaction motions drawn from the training distributions, we evaluate generalization on an unseen motion set. This set is constructed by uniformly sampling 100 motion clips (10 seconds each) from each of the retargeted InterHuman and Inter-X datasets, yielding 200 clips in total, while explicitly excluding all motions used for residual-expert training or unified policy distillation. The unseen set covers diverse interaction patterns and contact configurations, and is used to assess generalization beyond the curated interaction motion library.

Table 10: Summary of interaction motions in the control benchmark.

Motion ID	Motion Source	Duration	Category	Motion Description
InterHuman				
1	InterHuman-2312	8.0 s	Dance	Synchronized dance with shoulder and waist contact.
2	InterHuman-2853	58.4 s	Push	Repeated right-hand pushing, blocking.
3	InterHuman-6436	29.6 s	Dance	Two people dance together.
4	InterHuman-6494	23.9 s	Embrace	Two people dance while embracing.
5	InterHuman-6774	32.3 s	Dance	Face-to-face ballroom dance.
6	InterHuman-6791	25.5 s	Dance	Collaborative dance after approaching with raised arms.
7	InterHuman-2206	22.1 s	Embrace	Close embrace, swaying, walking, and a guided turn.
8	InterHuman-2239	57.0 s	Dance	Dancing followed by hugging and spinning.
9	InterHuman-2670	77.2 s	Spin	Hand-clasped spinning followed by an embrace.
10	InterHuman-2861	60.4 s	TaiChi	Partner tai chi with continuous hand contact.
11	InterHuman-2916	53.9 s	Push	Alternating left- and right-hand pushing and defending.
12	InterHuman-2924	85.7 s	Combat	Repeated attacks, defense, and counter-pushing.
13	InterHuman-2932	83.7 s	Struggle	Bilateral hand-contact struggle with pushing and defense.
14	InterHuman-3857	44.8 s	Sparring	Martial-arts sparring between two people.
15	InterHuman-6391	20.3 s	Sway	Handholding with repeated side-to-side swaying.
16	InterHuman-6809	23.4 s	Dance	Contact dance with circular spinning.
17	InterHuman-6893	71.8 s	Embrace	Approaching, embracing, and swaying together.
18	InterHuman-2115	10.6 s	Stretch	Alternating two-handed grasp and forward arm stretch.
19	InterHuman-2116	12.1 s	Spin	Handholding, raised arms, and turning.
20	InterHuman-2251	19.9 s	Embrace	Embracing while moving in a circle.
21	InterHuman-2736	10.2 s	Conversation	Laughing and conversing face to face.
22	InterHuman-2851	63.5 s	TaiChi	Partner tai chi practice.
23	InterHuman-319	10.7 s	Grasp	Approaching quickly, grasping a hand, and bowing.
24	InterHuman-327	12.0 s	Handshake	Approaching, shaking hands, and stepping backward.
25	InterHuman-7023	16.5 s	Wrestling	Argument followed by wrestling.
26	InterHuman-7207	14.2 s	Grapple	Argument, shoving, and arm grappling.
27	InterHuman-7335	10.0 s	Care	Bandaging a partner's hand followed by a hug.
28	InterHuman-7619	10.0 s	Snatch	Snatching an object followed by a physical struggle.
29	InterHuman-1557	61.7 s	Boxing	Boxing with repeated exchanges of punches.
30	InterHuman-785	9.6 s	Fencing	Two-person fencing practice.
Inter-X				
31	Inter-X-G013T005A021R003	23.3 s	Dance	Dancing and spinning while holding both hands.
32	Inter-X-G014T004A021R005	21.4 s	Dance	Face-to-face partner dance with arm and hand contact.
33	Inter-X-G017T000A000R002	20.9 s	Hug	A hesitant embrace followed by reciprocation.
34	Inter-X-G017T004A021R007	25.5 s	Dance	One partner guides the other's arms while dancing.
35	Inter-X-G031T005A021R005	23.5 s	Dance	Partner swaying with shoulder and waist contact.
36	Inter-X-G053T000A001R011	27.7 s	Handshake	Face-to-face two-handed handshake.
37	Inter-X-G055T003A016R008	27.0 s	High-five	Playful sequence of stepping on each other's feet.
38	Inter-X-G057T000A000R004	20.4 s	Hug	Side embrace around the partner's waist.
39	Inter-X-G004T008A021R005	20.1 s	Dance	Face-to-face dance with side jumps and kicks.
40	Inter-X-G007T001A006R002	21.7 s	Posing	Coordinated hand gestures and paired posing.
41	Inter-X-G024T005A028R019	20.3 s	Wrestling	Standing hand wrestling with crossed wrists.
42	Inter-X-G028T004A021R005	57.4 s	Dance	Close partner dance with waist and shoulder contact.
43	Inter-X-G028T004A021R010	25.3 s	Dance	Partner dance with handholding and raised arms.
44	Inter-X-G029T003A016R009	59.5 s	High-five	Mutual foot-stepping followed by a high-five.
45	Inter-X-G031T005A021R006	27.3 s	Dance	Partner dance with waist contact and handholding.
46	Inter-X-G034T003A012R012	20.2 s	Pointing	Pointing at the partner and then behind them.
47	Inter-X-G034T007A036R008	21.9 s	Covering	Reaching out to cover the partner's mouth.
48	Inter-X-G052T004A022R011	22.2 s	Linking	Walking side by side with linked arms.
49	Inter-X-G053T005A021R007	37.6 s	Dance	Handholding and circular partner dance.
50	Inter-X-G055T001A008R019	23.7 s	Pull	Grasping and pulling the partner closer.
51	Inter-X-G034T004A017R013	25.2 s	Chase	Taking turns chasing and evading each other.
52	Inter-X-G053T000A003R022	12.1 s	Grab	Grabbing the partner's hand and pulling them closer.
53	Inter-X-G053T000A004R027	15.2 s	Hit	Striking the partner's shoulder with a fist.
54	Inter-X-G009T006A033R009	42.6 s	Imitation	Imitating gestures, steps, and leg movements.
55	Inter-X-G016T001A005R001	13.6 s	Kick	Kicking toward a partner's leg while they dodge.
56	Inter-X-G028T002A014R026	12.4 s	Collision	Bumping the partner's shoulder with one's own.
57	Inter-X-G052T006A030R000	15.3 s	Patting	Approaching from the side and patting a cheek.
58	Inter-X-G016T004A020R001	20.7 s	Game	Playing rock-paper-scissors face to face.
59	Inter-X-G017T002A010R002	12.5 s	Slap	Exchanging slaps to the sides of the face.
60	Inter-X-G053T003A018R017	14.9 s	Whisper	Approaching and whispering near a partner's ear.

C.2 Baseline Implementation Details

To validate our contributions, we benchmark our framework against two sets of baselines: retargeting methods (answering Q1) and control policy variants (answering Q2 & Q3).

Retargeting Baselines. For fair comparison, all baselines use the same source motions and identical skeletal-scaling preprocessing.

- **GMR (General Motion Retargeting)** (Araujo et al., 2025): A Cartesian-space optimization-based retargeting method that independently fits each humanoid to the source motion by minimizing positional and rotational errors of selected **body keypoints**, such as the pelvis and end-effectors. Since the two agents are optimized separately, GMR does not explicitly model inter-agent interaction constraints.
- **OmniRetarget** (Yang et al., 2025): A topology-aware retargeting method that uses an **Interaction Mesh** (I-Mesh) to preserve geometric structure, originally developed for human–object interaction. In our dual-humanoid setting, we apply OmniRetarget independently to each agent. As a result, it preserves per-agent body geometry but does not explicitly constrain the relative spatial relationships between the two humanoids.
- **PHC** (Luo et al., 2023): A **sparse correspondence-based** humanoid retargeting pipeline that fits predefined human joints to robot links. We apply it independently to each agent, preserving individual motion but without explicit inter-agent geometric constraints.
- **UMR (Unified Motion Retargeting)** (Cao et al., 2026):A dense surface correspondence-based retargeting method that aligns human and robot exterior **point clouds** through constrained optimization. We apply it independently to each agent, providing fine-grained body-surface alignment but without explicit pairwise constraints between the two humanoids.

Control Policy Baselines. We evaluate our scalable flow-based interaction policy training framework by comparing against representative single-agent baselines and ablated variants. All variants are trained using the same PPO hyperparameters.

- **TWIST2** (Ze et al., 2025): An RL-based whole-body motion tracker conditioned on reference motion and robot proprioception. We initialize from the official checkpoint and adapt the tracker on our interaction-motion subsets, applying it independently to each humanoid with its own reference motion. The policy receives neither partner-state information nor explicit inter-humanoid interaction cues.
- **BeyondMimic** (Liao et al., 2026): An RL-based motion-tracking framework that trains a dedicated policy for each reference motion, targeting high-dynamic whole-body tracking. In our dual-humanoid setting, we train one tracker per motion and execute the two corresponding policies independently for the two robots.
- **SONIC** (Luo et al., 2026): A large-scale universal whole-body motion tracker that uses a single unified policy to track diverse reference motions, including unseen motions. We independently run the same tracker for both humanoids with their respective references, without explicit peer-state or interaction-aware constraints.
- **From scratch RL (w/o Residual Learning)**: Evaluates the contribution of residual interaction learning built upon a pretrained single-agent motion prior. *Implementation*: The frozen base tracker and residual decomposition are removed, and a peer-aware policy is trained from scratch with PPO to directly predict the full control actions for both humanoids. The observation space and interaction rewards remain unchanged.
- **Specialist→Unified MLP (w/o Flow Matching)**: Evaluates the contribution of flow-based policy consolidation when integrating interaction specialists into a unified controller. *Implementation*: The residual specialists and DAgger data-collection procedure remain unchanged, while the conditional flow-matching policy is replaced with a deterministic MLP that directly regresses the paired residual actions from the same conditioning inputs and expert targets.

C.3 Evaluation Protocol and Metric Details

This subsection details the evaluation protocol and metrics used throughout the paper. We use task-specific metrics to evaluate retargeting quality and dynamic control performance of our framework, respectively.

1) Evaluation Protocol

Motion segmentation and episode length. For control policy evaluation, to ensure a consistent evaluation horizon, we divide all reference motions into 10-second clips, corresponding to 500 control steps and matching the episode length used during training. Motions shorter than 10 seconds are evaluated in their original duration. Each motion clip is evaluated over 10 independent rollouts, and metrics are averaged across these trials.

2) Metrics Details

Retargeting Evaluation Metrics. We evaluate retargeting quality from three complementary aspects: physical feasibility, interaction fidelity, and downstream learnability.

- **Maximum Penetration Depth (MPD).** Measures physical plausibility by reporting the maximum inter-humanoid penetration depth over each motion sequence. Penetration is computed using the MuJoCo (Todorov et al., 2012) collision proxy geometries. We first compute MPD for each sequence and then average it over the evaluation set. Lower values indicate fewer severe inter-body collisions.
- **Contact Interaction Error (CIE).** Measures fine-grained contact fidelity using cross-body surface pairs whose distance in the source motion is within the contact threshold $\tau_c = 2\,\mathrm{cm}$. For each pair, we compute the L_2 deviation between the source and retargeted 3D relative vectors and normalize it by the canonical humanoid body scale $s_{\mathrm{body}} = 1.32\,\mathrm{m}$. CIE is reported as a percentage, with lower values indicating better preservation of contact distance and direction.
- **Near-field Geometry Error (NGE).** Measures preservation of the surrounding interaction geometry using cross-body surface pairs whose source distances lie between $\tau_c = 2\,\mathrm{cm}$ and $\tau_n = 15\,\mathrm{cm}$. Pairs whose two endpoints are both contact points are excluded. The L_2 deviation between the source and retargeted relative vectors is normalized by s_{body} and reported as a percentage. Lower NGE indicates better preservation of near-field spatial structure.
- **Contact F1 (C-F1).** Measures region-level contact preservation. A body-region pair is considered in contact if at least one cross-body surface pair lies within $\tau_c = 2\,\mathrm{cm}$. We compare the source and retargeted contact labels and report the micro-averaged F1 score over all frames as a percentage.
- **Downstream Success Rate (DSR).** Serves as a proxy for downstream physical learnability. We define DSR as the percentage of RL rollouts that successfully execute the retargeted reference while preserving the interaction structure, requiring $\mathrm{CIE} < 20\%$ and $\mathrm{NGE} < 20\%$ relative to the scaled reference.

Control Policy Evaluation Metrics. We evaluate dual-humanoid control in terms of motion tracking, interaction preservation, and overall execution success with respect to the retargeted reference trajectories.

- **Root-relative Mean Per-Joint Position Error (E_{mpjpe}).** Measures individual motion-tracking fidelity as the mean per-joint position error after removing global root translation. We compute E_{mpjpe} over both humanoids and average it over the rollout. Lower values indicate more accurate whole-body tracking independent of global translational drift.
- **Interaction Geometry Error ($E_{inter\text{-}geo}$).** Measures how faithfully the executed motion preserves the contact interfaces and surrounding interaction geometry of the retargeted reference. We reuse the CIE and NGE definitions above, replacing the source–retargeted comparison with a reference–execution comparison. Contact and near-field surface pairs are determined exclusively from the retargeted reference using the same thresholds and exclusion rules. For each component, relative-vector deviations are averaged over all corresponding frame–pair instances in the rollout, normalized by s_{body} and reported as a percentage.
- **Success Rate (Succ.).** Measures overall interaction execution success. A rollout is considered successful when $E_{mpjpe} < 10\,\mathrm{cm}$, $E_{inter\text{-}geo} < 20\%$, and neither humanoid falls or terminates early. We report the percentage of successful rollouts over all evaluation trials.

D ADDITIONAL EXPERIMENTS

D.1 ABLATION STUDY & ADDITIONAL VISUALIZATION ON SCAR

Table 11: **SCAR component ablation on Inter-X and InterHuman.** All metrics are reported in %; arrows indicate the preferred direction, and the best results are in bold.

Variant	Inter-X			InterHuman		
	CIE ↓	NGE ↓	DSR ↑	CIE ↓	NGE ↓	DSR ↑
w/o CIO	6.78	5.43	64.2	7.45	7.21	68.8
w/o NCO	4.21	6.63	65.8	5.23	8.21	69.2
SCAR (full)	**3.98**	**4.99**	**72.1**	**4.96**	**6.49**	**75.4**



Figure 7: Qualitative comparison of interaction retargeting on Inter-X and InterHuman. (Top: High-five, Middle: Push, Bottom: Holding Hands.) Columns show the raw motion and results from GMR, OmniRetarget, UMR, PHC, and SCAR (ours). Zoomed-in views highlight contact details.

Component ablation. As shown in Table 11, full SCAR achieves the best results across all metrics on both datasets. Removing **Contact-Interface Objective (CIO)** primarily increases CIE, from 3.98% to 6.78% on Inter-X and from 4.96% to 7.45% on InterHuman. Removing **Near-Field Context Objective (NCO)** instead leads to a larger increase in NGE, from 4.99% to 6.63% and from 6.49% to 8.21%, respectively. Both ablations also reduce DSR, suggesting that CIO and NCO play complementary roles in preserving interaction quality and supporting successful motion execution.

Qualitative comparison. Figure 7 provides additional comparisons on high-five, push, and holding-hands interactions. The enlarged views highlight differences in the relative placement of interacting body parts. Compared with the baselines, SCAR more closely preserves the contact relationships in the source motions, particularly hand-to-body alignment in pushing and hand-to-hand alignment in holding hands. These examples qualitatively support SCAR's effectiveness in preserving interaction-specific spatial relationships during retargeting.

D.2 Sensitivity Analysis and Qualitative Results of OmniInteract

We further examine the sensitivity of OmniInteract to the residual action scale and provide qualitative examples of interaction retargeting in simulation and humanoid interactions in the real world.

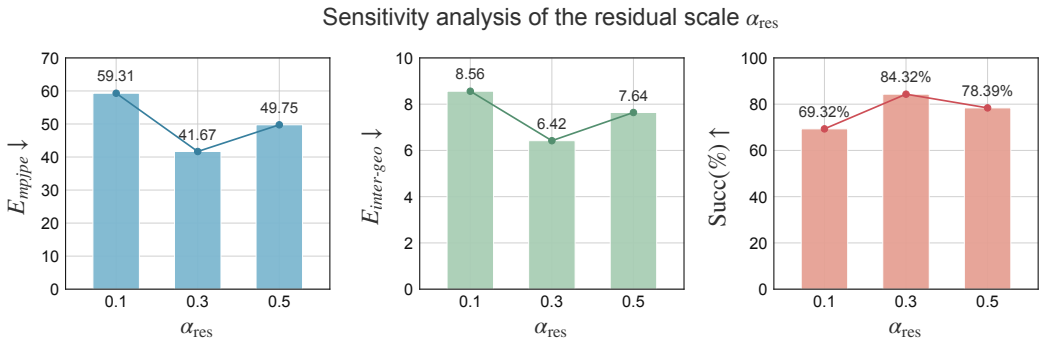


Figure 8: Sensitivity to the residual action scale α_{res}. A scale of 0.3 achieves the lowest tracking and interaction geometry errors and the highest success rate.

Figure 8 evaluates the residual action scale $\alpha_{\text{res}} \in \{0.1, 0.3, 0.5\}$. Among the tested values, $\alpha_{\text{res}} = 0.3$ achieves the best performance across all three metrics, with $E_{mpjpe} = 41.67$, $E_{inter\text{-}geo} = 6.42$, and a success rate of 84.32%. Both smaller and larger scales increase the errors and reduce the success rate, highlighting the importance of appropriately scaling the residual correction.



Figure 9: Qualitative results of OmniInteract for interaction control in Isaac simulation. From top to bottom: dancing, fighting, and walking with an arm around a partner's shoulder.

Figure 9 visualizes the results of OmniInteract for interaction control in Isaac simulation, including dancing, fighting, and walking with an arm around a partner's shoulder. These examples complement the quantitative evaluation by illustrating the robot poses and relative spatial configurations associated with different interaction patterns.

Figure 10 presents additional real-world examples of boxing, dancing, and meeting. Each row contains five keyframes ordered chronologically from left to right, showing how the two humanoid

robots' poses and relative positions evolve during execution. Together, these examples illustrate the range of coordinated and close-range interactions considered in our evaluation.

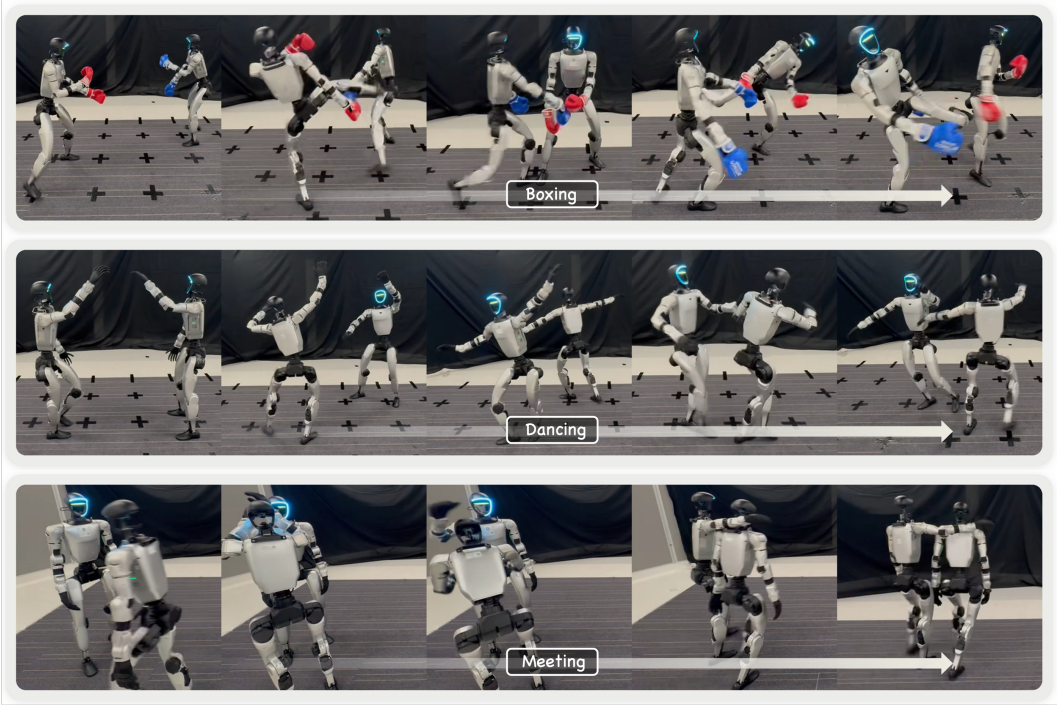


Figure 10: Qualitative visualization of real-world humanoid interactions. (Top: Boxing, Middle: Dancing, Bottom: Meeting.) Each row presents five keyframes arranged chronologically from left to right, illustrating coordinated movements and close-range interactions between two humanoid robots.)