CVPR #1

ConTrack: Constrained Hand Motion Tracking with Adaptive Trade-off Control

Anonymous CVPR submission

Paper ID 1

Abstract

Human demonstrations provide strong priors for robot manipulation, yet transferring them to real robots is non-trivial because of the kinematic gap. In dexterous manipulation, even simulation tracking remains difficult for long-horizon, contact-rich sequences: a reference-tracking policy must keep objects on their target trajectories while preserving demonstrated joint motion and contact timing. Existing approaches often rely on hand-crafted reward tuning, which requires per-sequence tuning and can break under limited interaction budgets. We introduce ConTrack, a reinforcement learning framework that scales with tracking data. ConTrack treats object tracking as a constraint and allocates the remaining control authority to motion fidelity, allowing the task–style trade-off to adapt online through a dual-variable update. ConTrack also stabilizes long-horizon learning with an adaptive mid-trajectory reset library that reuses policy-reachable simulator states. Qualitative and quantitative results in simulation tracking and on a real robot show that ConTrack improves success and object pose accuracy over prior methods while preserving joint and contact fidelity.

1. Introduction

Human hand-object recordings are becoming a practical source of supervision for robot dexterity. They contain long sequences, natural contact schedules, and object use patterns that are difficult to script by hand [3, 4, 6, 7, 9, 22]. The hard part is execution. Many pipelines can recover high-level intent or coarse motion cues from human data, but contact-rich manipulation still needs an additional stage that calibrates the interaction to the target robot. This stage matters because human data introduces a systematic mismatch between the demonstrated hand motion and what a robot hand can physically realize. Morphology and actuation differences appear as a kinematic gap when one attempts to execute human motion directly.

One intuitive response is retargeting. Optimization and neural retargeting align human motion to robot kinematics by minimizing geometric errors [8, 18, 26]. Purely geometric objectives tend to be fragile once contact schedules matter. Recent systems add physical feasibility, object-centric goals, or learned tracking controllers [10, 12, 14, 15], but they still leave an open question: when the reference is not jointly realizable, how should a system choose the trade-off between task success and motion style without per-sequence manual tuning?

Figure 1. ConTrack tracks long-horizon, contact-rich dexterous hand motion across tool use, articulated object interaction, and in-hand rotation. Adaptive task-style control balances object tracking, hand motion, and contact fidelity.

Inspired by tracking-centric control in physics simulation [13, 16, 17, 24], we propose ConTrack, a reference-tracking reinforcement learning framework for contact-rich hand-object interaction. Each input clip specifies retargeted robot joint references, object pose targets, and link-level contact annotations. ConTrack casts object tracking as a constraint rather than another reward term, following the Lagrangian view used in CMDPs and constrained policy optimization [1, 2, 5, 19, 21, 23, 25]. The policy can deviate from the reference when required by physics while remaining anchored to demonstrated motion and contact cues.

Long horizons create a second problem. If training always starts from the first frame, early failures dominate the rollout distribution and later contact phases receive too few updates. Resetting directly to the reference mid-trajectory can also be physically inconsistent because contacts and object configurations are coupled through dynamics. ConTrack instead stores simulator states visited by the current policy and resamples difficult frames from this policy-reachable library. Together, the adaptive task-style controller and reset library transform physically infeasible reference motion into executable robot trajectories.

2. Method

Problem setup. For a reference clip of length T, the simulator state x_t contains robot joints, object poses, and velocities. The reference provides joint targets $q_t^{\mathrm{ref}} \in \mathbb{R}^D$, object poses for O objects, and binary contact targets $c_{t,o,\ell}^{\mathrm{ref}}$ for hand link ℓ. The policy observes the simulator state and time index, then outputs a residual action $a_t \in \mathbb{R}^D$:

$$q_t^{\mathrm{tar}} = q_t^{\mathrm{ref}} + a_t. \tag{1}$$

This keeps exploration close to the retargeted motion while still allowing corrections. We separate the reward into object tracking r_{g}, style fidelity r_{s}, and a smoothness penalty r_{p}. The corresponding discounted returns are $J_{\mathrm{g}}(\pi)$ and $J_{\mathrm{s}}(\pi)$.

Adaptive task-style mixing. ConTrack optimizes style subject to a normalized task constraint. This is not a full safe-RL guarantee; it is a lightweight CMDP-style controller for the task-style conflict in reference tracking:

$$\max_{\pi} \; J_{\mathrm{s}}(\pi) \quad \text{s.t.} \quad J_{\mathrm{g}}(\pi) \geq \alpha J_{\mathrm{g}}^{\max}, \tag{2}$$

where $\alpha \in (0, 1]$ is a target task ratio and $J_{\mathrm{g}}^{\max}$ is the running maximum of a moving task-return estimate for the clip. The normalization makes the update less sensitive to the raw scale of different objects and clips. We implement Eq. (2) with a scalar dual state λ, similar in spirit to primal-dual constrained RL updates [1, 2, 23]:

$$\lambda \leftarrow \lambda + \eta \left(\alpha - \hat{J}_{\mathrm{g}} / J_{\mathrm{g}}^{\max} \right), \qquad w_{\mathrm{task}} = \sigma(\lambda). \tag{3}$$

PPO then uses a mixed advantage

$$A_{\mathrm{mix}} = w_{\mathrm{task}} A_{\mathrm{g}} + (1 - w_{\mathrm{task}}) A_{\mathrm{s}} + A_{\mathrm{p}}. \tag{4}$$

When the policy falls below the task ratio, w_{task} increases and the learner pushes harder on object tracking. Once the object motion is reliable, the update releases weight back to style.

Policy-reachable reset library. ConTrack maintains a cached simulator state $B[k]$ for each reference frame k. After an episode that starts at t_s and ends at t_e, every visited frame $k \in [t_s, t_e]$ receives a continuation length $\ell_k = t_e - k + 1$. We maintain an exponential moving average $\bar{\ell}_k$ and refresh $B[k]$ with the visited simulator state when the rollout exceeds the previous best continuation at frame k. Sampling then favors frames with low survival ratio,

$$u_k = \bar{\ell}_k / (T - k), \qquad p(k) \propto \exp(-u_k / \tau). \tag{5}$$

Method	Prog.↑	Pos.↓	Rot.↓	Fing.↓	F1↑	Pt.↓
ConTrack	.899	.026	.272	.163	.784	.018
ManipTrans [10]	.743	.012	.207	.277	.620	.030
DexMachina [14]	.246	.038	.348	.147	.708	.024
SPIDER [15]	.444	.201	1.104	.157	.191	.036

Table 1. Main comparison averaged over clips. Position and contact point errors are in meters; rotation and finger errors are in radians.

Variant	Prog.↑	Pos.↓	Rot.↓	Fing.↓	F1↑	Pt.↓
ConTrack	.899	.026	.272	.163	.784	.018
Fixed task	.764	.026	.250	.157	.679	.022
Fixed 1:1	.868	.029	.297	.165	.701	.023
Start reset	.700	.039	.370	.168	.739	.017
Uniform reset	.727	.033	.298	.152	.714	.021
w/o contact	.861	.032	.320	.168	.699	.020

Table 2. Selected ablations under the same training budget. Adaptive mixing improves progress while preserving contact fidelity, and the reset library improves long-horizon coverage.

Early in training, resets concentrate near short feasible suffixes. As the policy learns, the reachable set moves backward through the clip and training spends more rollouts near the remaining difficult contact transitions.

Contact priors. Object pose alone does not specify how a hand should touch the object. We therefore use dataset contacts as a style prior: the reward encourages agreement with reference contact events and, when both reference and simulated contacts are active, penalizes the object-local distance between the realized and target contact points. This term is not a hard constraint; it biases the remaining freedom after object tracking toward human-like contact timing.

3. Experiments

We evaluate one policy per clip on three benchmark tiers. GRAB contains bimanual rigid-object interaction [22]; ARCTIC adds articulated objects and multi-object contact [6]; DexterHand focuses on continuous in-hand rotation [11]. Learning-based methods are trained under the same 5000-update PPO budget with PPO [20], and all methods use the same evaluation metrics. Evaluation rollouts start from the first reference frame, so progress measures end-to-end tracking. We report rollout progress, object pose error, finger tracking error, contact F1, and contact point error.

Table 1 shows the main result. ConTrack reaches 0.899 average progress with strong contact F1. ManipTrans attains the lowest object pose error but drifts in finger motion and contacts. DexMachina preserves finger motion and contact fidelity more closely but makes limited progress under the same budget. SPIDER stays close in kinematics but struggles to maintain object tracking once contact dynamics dominate. The hardest in-hand rotation clip remains challenging within the fixed budget.

The ablations in Table 2 isolate the two training mechanisms. Fixed task weighting improves object terms locally but loses contact fidelity and progress. A fixed 1:1 mixture is closer, yet remains below the adaptive controller. Resetting only from the start wastes updates on already learned prefixes, while uniform mid-clip resets include states that are not aligned with the current policy. Removing contact reward reduces both contact F1 and progress, suggesting that contact priors help the policy enter the right local dynamics rather than merely improve a style score.

4. Conclusion

ConTrack is a constrained reference-tracking reinforcement learning framework for long-horizon, contact-rich hand-object interaction in physics simulation. It separates task success from style fidelity, controls their trade-off online with a scalar dual controller, and trains long clips from policy-reachable simulator states. Across bimanual interaction and in-hand manipulation benchmarks under a fixed training budget, ConTrack improves progress and contact fidelity while retaining accurate object tracking. The formulation still has limits: the online normalization does not strictly guarantee constraint satisfaction, contact priors rely on usable annotations, the hardest in-hand clips remain difficult under limited interaction budgets, and the hardware study currently executes streamed joint commands without perception feedback.

References

[1] Joshua Achiam, David Held, Aviv Tamar, and Pieter Abbeel. Constrained policy optimization. In *International Conference on Machine Learning*, 2017. 1, 2

[2] Eitan Altman. *Constrained Markov Decision Processes*. Chapman and Hall/CRC, 1999. 1, 2

[3] Samarth Brahmbhatt, Chengcheng Tang, Christopher D. Twigg, Charles C. Kemp, and James Hays. ContactPose: A dataset of grasps with object contact and hand pose. In *European Conference on Computer Vision*, 2020. 1

[4] Yu-Wei Chao, Wei Yang, Yu Xiang, Pavlo Molchanov, Ankur Handa, Jonathan Tremblay, Yashraj S. Narang, Karl Van Wyk, Umar Iqbal, Stan Birchfield, Jan Kautz, and Dieter Fox. DexYCB: A benchmark for capturing hand grasping of objects. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2021. 1

[5] Yinlam Chow, Ofir Nachum, Edgar Duenez-Guzman, and Mohammad Ghavamzadeh. A lyapunov-based approach to safe reinforcement learning. In *Advances in Neural Information Processing Systems*, 2018. 1

[6] Zicong Fan, Omid Taheri, Dimitrios Tzionas, Muhammed Kocabas, Manuel Kaufmann, Michael J. Black, and Otmar Hilliges. ARCTIC: A dataset for dexterous bimanual hand-object manipulation. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023. 1, 2

[7] Rao Fu, Deqing Zhang, Anqi Jiang, Wen Fu, Adam Funk, Daniel Ritchie, and Srinath Sridhar. GigaHands: A massive annotated dataset of bimanual hand activities. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2025. 1

[8] Ankur Handa, Karl Van Wyk, Wei Yang, Jacky Liang, Yu-Wei Chao, Qian Wan, Stan Birchfield, Nathan Ratliff, and Dieter Fox. DexPilot: Vision-based teleoperation of dexterous robotic hand-arm system. In *IEEE International Conference on Robotics and Automation*, 2020. 1

[9] Nan Jiang, Tianxing Liu, Zhaowen Cao, Jialiang Cui, Yixin Chen, He Wang, Yixin Zhu, and Siyuan Huang. Full-body articulated human-object interaction. In *IEEE/CVF International Conference on Computer Vision*, 2023. 1

[10] Kuan Li, Peiyuan Li, Tianxing Liu, Yuyang Li, and Siyuan Huang. ManipTrans: Efficient dexterous bimanual manipulation transfer via residual learning. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2025. 1, 2

[11] Yutong Liang, Sirui Xu, Yifeng Zhang, Bowen Zhan, Hao Zhang, and Libin Liu. DexterCap: An affordable and automated system for capturing dexterous hand-object manipulation, 2026. 2

[12] Xueyi Liu, Jieru Adalibieke, Qiayuan Han, Yuzhe Qin, and Li Yi. DexTrack: Towards generalizable neural tracking control for dexterous manipulation from human references. In *International Conference on Learning Representations*, 2025. 1

[13] Zhengyi Luo, Ye Yuan, Tianyu Wang, Cheng Li, Siwei Chen, Fernando Castaneda, Zhaowen Cao, Jiaman Li, David Minor, and Qiyang Ben. Sonic: Supersizing motion tracking for natural humanoid whole-body control, 2025. 1

[14] Zhao Mandi, Yifeng Hou, Dieter Fox, Yashraj Narang, Ajay Mandlekar, and Shuran Song. DexMachina: Functional retargeting for bimanual dexterous manipulation, 2025. 1, 2

[15] Chuer Pan, Chen Wang, Haozhi Qi, Zeyu Liu, Homanga Bharadhwaj, Archit Sharma, Tony Wu, Guanya Shi, Jitendra Malik, and Francois Hogan. SPIDER: Scalable physics-informed dexterous retargeting, 2025. 1, 2

[16] Xue Bin Peng, Pieter Abbeel, Sergey Levine, and Michiel van de Panne. DeepMimic: Example-guided deep reinforcement learning of physics-based character skills, 2018. 1

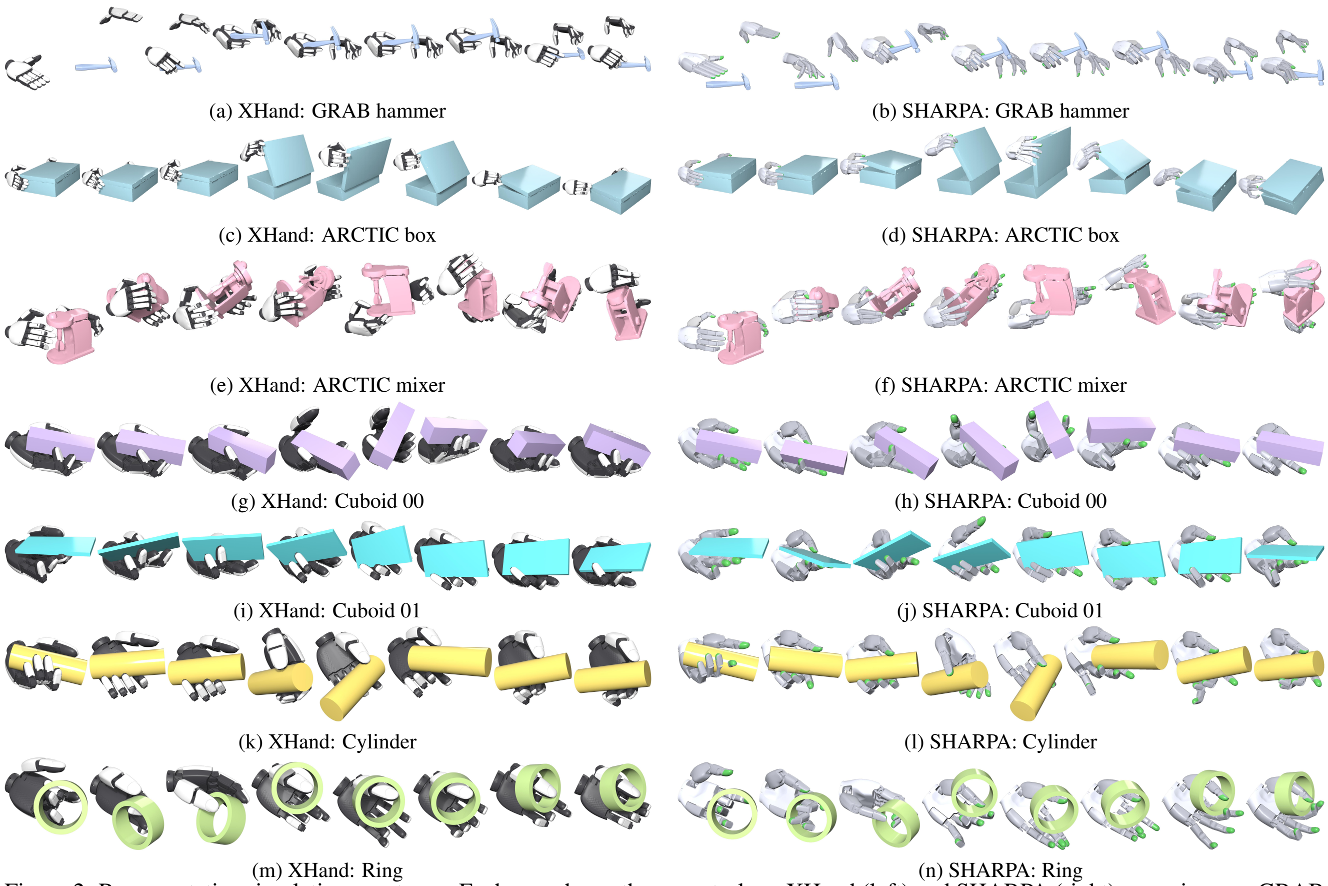
[17] Xue Bin Peng, Yunrong Guo, Laura Halper, Sergey Levine, and Sanja Fidler. ASE: Large-scale reusable adversarial skill embeddings for physically simulated characters. *ACM Transactions on Graphics*, 41(4), 2022. 1

[18] Yuzhe Qin, Wei Yang, Binghao Huang, Karl Van Wyk, Hao Su, Xiaolong Wang, Yu-Wei Chao, and Dieter Fox. AnyTeleop: A general vision-based dexterous robot arm-hand teleoperation system, 2023. 1

[19] Alex Ray, Joshua Achiam, and Dario Amodei. Benchmarking safe exploration in deep reinforcement learning. In *NeurIPS Workshop on Deep Reinforcement Learning*, 2019. 1

[20] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017. 2

[21] Adam Stooke, Joshua Achiam, and Pieter Abbeel. Responsive safety in reinforcement learning by pid lagrangian methods. In *International Conference on Machine Learning*, 2020. 1



(a) XHand: GRAB hammer
(b) SHARPA: GRAB hammer
(c) XHand: ARCTIC box
(d) SHARPA: ARCTIC box
(e) XHand: ARCTIC mixer
(f) SHARPA: ARCTIC mixer
(g) XHand: Cuboid 00
(h) SHARPA: Cuboid 00
(i) XHand: Cuboid 01
(j) SHARPA: Cuboid 01
(k) XHand: Cylinder
(l) SHARPA: Cylinder
(m) XHand: Ring
(n) SHARPA: Ring

Figure 2. Representative simulation montages. Each row shows the same task on XHand (left) and SHARPA (right), covering one GRAB clip, two ARCTIC clips, and four DexterHand clips.

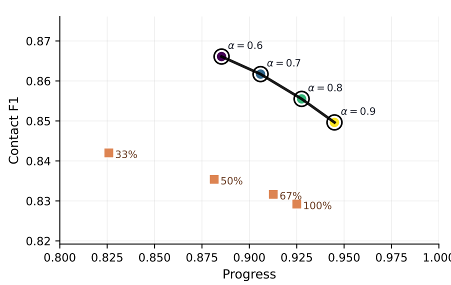


Figure 3. Sweeping the target task ratio traces a task-style Pareto Frontier. Fixed reward mixtures fall inside the frontier under the tested budget.

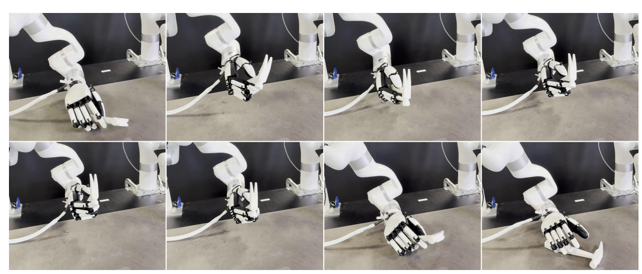


Figure 4. Real-world feasibility on a bimanual setup. We stream policy-predicted joint references from simulation to a real-time controller.

[22] Omid Taheri, Nima Ghorbani, Michael J. Black, and Dimitrios Tzionas. GRAB: A dataset of whole-body human grasping of objects. In *European Conference on Computer Vision*, 2020. 1, 2

[23] Chen Tessler, Daniel J. Mankowitz, and Shie Mannor. Reward constrained policy optimization. In *International Conference on Learning Representations*, 2019. 1, 2

[24] Chen Tessler, Yunrong Guo, Ofir Nabati, Gal Chechik, and Xue Bin Peng. MaskedMimic: Unified physics-based character control through masked motion inpainting. *ACM Transactions on Graphics*, 2024. 1

[25] Kai Wen, Cheng Li, Jun He, and Marco Hutter. Constrained style learning from imperfect demonstrations under task optimality, 2025. 1

[26] Zhi-Hao Yin, Chen Wang, Luis Pineda, Kaushik Bodduluri, Tony Wu, Pieter Abbeel, and Mustafa Mukadam. Geometric retargeting: A principled, ultrafast neural hand retargeting algorithm. In *IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2025. 1