Object-Centric Dexterous Manipulation from Human Motion Data

Anonymous Author(s)
Affiliation
Address
email

Abstract: Manipulating objects to achieve desired goal states is a basic but important skill for dexterous manipulation. Human hand motions demonstrate proficient manipulation capability, providing valuable data for training robots with multi-finger hands. Despite this potential, substantial challenges arise due to the embodiment gap between human and robot hands. In this work, we introduce a hierarchical policy learning framework that uses human hand motion data for training object-centric dexterous robot manipulation. At the core of our method is a high-level trajectory generative model, learned with a large-scale human hand motion capture dataset, to synthesize human-like wrist motions conditioned on the desired object goal states. Guided by the generated wrist motions, deep reinforcement learning is further used to train a low-level finger controller that is grounded in the robot's embodiment to physically interact with the object to achieve the goal. Through extensive evaluation across 10 household objects, our approach not only demonstrates superior performance but also showcases generalization capability to novel object geometries and goal states. Furthermore, we transfer the learned policies from simulation to a real-world bimanual dexterous robot system, further demonstrating its applicability in real-world scenarios. Project website: `https://sites.google.com/view/obj-dex`.

Keywords: Dexterous Manipulation, RL, Learning from Human

1 Introduction

Developing bimanual multi-fingered robotic systems capable of handling complex manipulation tasks with human-level dexterity has been a longstanding goal in robotics research. Regardless of how the goals are specified, a common element across these definitions is an object-centric perspective focusing on the state of the objects being manipulated. As such, the goal of our work is to train a policy for a bimanual dexterous robot to manipulate the objects according to the task goal defined as a sequence of object pose trajectories.

Prior works primarily utilize deep reinforcement learning (RL) to learn object-centric dexterous manipulation skills [1–3]. Training RL policy that controls both robot arms and two multi-finger hands is possible in theory, but presents substantial challenges in practice due to the high degree of freedom of the robot action space. Imitation learning (IL) can potentially tackle this challenge by leveraging the guidance from human motion data to assist policy learning. However, another challenge arises due to the morphological differences between human and robotic hands, often referred to as the "embodiment gap".

One critical observation is that human finger motions are not consistently useful across various manipulation tasks due to the embodiment gap. Based on this observation, we propose a hierarchical policy learning framework consisting of a high-level planner for the wrist and a low-level controller for the hand. The high-level planner is a generative-based policy, trained by imitation learning with human wrist movements, to generate robot arm actions conditioned on a desired trajectory of the object's movements. Based on the generated arm motions, the low-level controller outputs fine-grained finger actions learned through RL exploration rather than imitation of human data. Our experiments demonstrate that

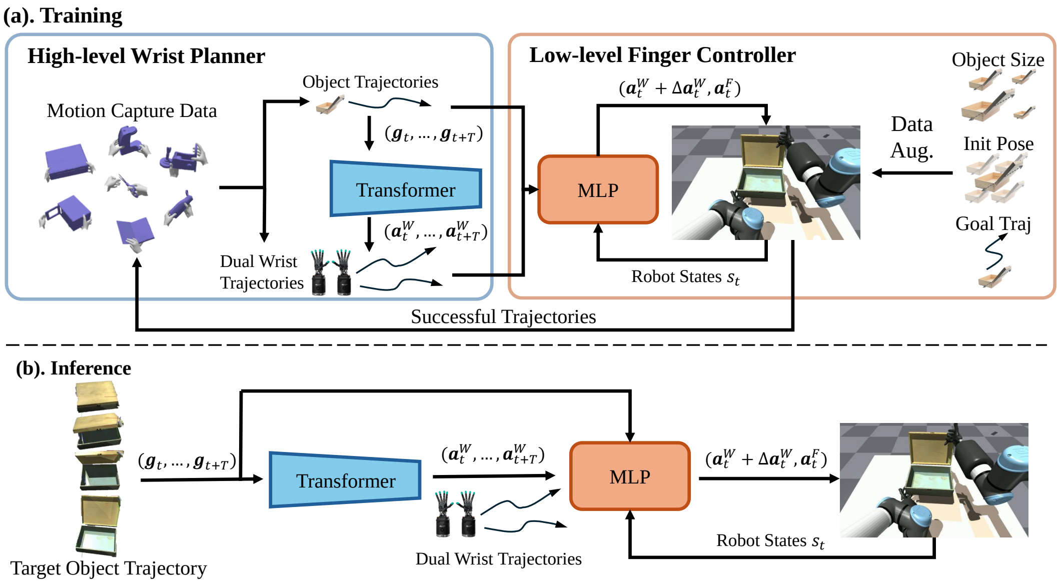

Figure 1: Overview of our framework. (A) Training: We train a generation model to synthesize dual hand trajectory and then use the RL to train a low-level robot controller. (B) Inference: Given a single object goal trajectory, our framework generates dual hand reference trajectory and guides the low-level controller to accomplish the task.

the learned policy exhibits generalization to novel object geometries and unseen motion trajectories. In addition, we successfully transfer our policy from simulation environments to a real-world bimanual dexterous robot, further validating its practical applicability in real-world manipulation tasks.

2 Related Works

2.1 Dexterous Manipulation

Dexterous manipulation is a long-standing research topic in robotics [4–7]. Traditional methods rely on analytical dynamic models for trajectory optimization [4, 5, 7–10], which fall short in complex tasks due to the simplification of contact dynamics. Recently, deep reinforcement learning (RL) has showcased promising results in training dexterous manipulation skills such as in-hand object reorientation [11–16, 16–21], bimanual manipulation [1, 2, 22], sequential manipulation [23–25], and human-like activities [26]. Despite the progress, successfully training a dexterous RL policy often requires extensive reward engineering and system design, which limits its practicality in some scenarios. Besides RL, imitation learning (IL) is also widely used for training dexterous policies [27, 28]. By performing supervised-learning with human teleoperation data [29–33], prior works show impressive results in dexterous grasping [34, 35] and general manipulation tasks [36–43].

2.2 Learning from Human Motion

Recently, learning from human motion data has started to receive more attention because it allows scaling up data collection without robot hardware. Prior works leverage human data [44–49], motion capture data [50–55] to extract valuable motion hints for manipulation [44–46, 48, 52]. For dexterous manipulation, [31, 36, 48, 51, 56, 57] showcase the potential of using analytical methods (e.g., inverse kinematics) to retarget human hand motion to robot hardware. However, due to the embodiment gap between human and robot hands, position-based retargeting methods do not guarantee the replication of task success. In contrast, our approach uses human data as guidance for RL training, which learns the motion retargeting conditioned on the robot's embodiment. Notably, [27, 58–62] share the same idea of utilizing human data as guidance or reward for reinforcement learning.

3 Task Formulation

The goal of an object-centric manipulation task is to let the robot physically interact with the object to achieve the desired motion trajectory. We define the motion trajectory as the sequence of the object's $SE(3)$ transformation $G=(\boldsymbol{g}_1,\boldsymbol{g}_2,...,\boldsymbol{g}_T)$, where each time step $\boldsymbol{g}_i=(\boldsymbol{g}_i^R,\boldsymbol{g}_i^T,g_i^J)$ consists a 3D rotation $\boldsymbol{g}_i^R$, a 3D translation $\boldsymbol{g}_i^T$, and the joint angle g_i^J. g_i^J can be omitted if the object is a single rigid body.

We then formulate an object-centric manipulation task as a Markov Decision Process (MDP) $\mathcal{M}=(\boldsymbol{S},\boldsymbol{A},\pi,\mathcal{T},R,\gamma,\rho,G)$, where $\boldsymbol{S}$ is the state space, $\boldsymbol{A}$ is the action space, π is the agent's policy, $\mathcal{T}(\boldsymbol{s}_{t+1}|\boldsymbol{s}_t,\boldsymbol{a}_t)$ is the transition distribution, R is the reward function, γ is the discount factor, and ρ is the initial state distribution. The policy π conditions on the reference object state trajectory G and the current state $\boldsymbol{s}_t$, and generates robot action distributions $\boldsymbol{a}_t$ to maximize the likelihood between the future object states $(\boldsymbol{s}_{t+1},\boldsymbol{s}_{t+2},...,\boldsymbol{s}_{t+T})$ and the reference trajectory G.

4 Method

In this section, we introduce our framework for object-centric manipulation. The overview of the framework is shown in Figure 1. Our framework consists of three parts: high-level planner (Section 4.1), low-level controller (Section 4.2). The data augmentation loop and the details of our sim-to-real policy transfer are introduced in Appendix.

4.1 High-Level Planner

We train a Transformer-based generative model π^H that takes object category ID c, and the desired object motion trajectory $G=(\boldsymbol{g}_t,\boldsymbol{g}_{t+1},...,\boldsymbol{g}_{t+T})$ as inputs and outputs a sequence of 6-DoF wrist actions $(\boldsymbol{a}_t^W,\boldsymbol{a}_{t+1}^W,...,\boldsymbol{a}_{t+T}^W)$, where each action $\boldsymbol{a}_i^W=(\boldsymbol{p}_i^l,\boldsymbol{p}_i^r)$ consists the 6-DoF pose of the left hand $\boldsymbol{p}_i^l$ and right hand $\boldsymbol{p}_i^r$ in SE(3). In our experiments, we use $T=10$.

4.2 Low-Level Controller

We use Proximal Policy Optimization (PPO) [63] to train π^L. The policy π^L takes the current observation $\boldsymbol{s}_i$, the desired object motion trajectory $G=(\boldsymbol{g}_t,\boldsymbol{g}_{t+1},...,\boldsymbol{g}_{t+T})$, and a sequence of 6-DoF wrist actions $(\boldsymbol{a}_t^W,\boldsymbol{a}_{t+1}^W,...,\boldsymbol{a}_{t+T}^W)$ generated by high-level planner as inputs, and outputs the finger joint action $\boldsymbol{a}_t^F$. Here the observation $\boldsymbol{s}_t$ contains the object pose and robot proprioception. The reward function is defined as $r_t=\exp^{-(\lambda_1*\|\boldsymbol{g}_t^R-\hat{\boldsymbol{g}}_t^R\|_2+\lambda_2*\|\boldsymbol{g}_t^T-\hat{\boldsymbol{g}}_t^T\|_2+\lambda_3*\|g_t^J-\hat{g}_t^J\|_2)}$, aiming to minimize the distance between object's movements and the desired goal trajectory. π^L also learns to output a residual wrist action $\Delta\boldsymbol{a}_t^W$ within a fixed range. The final robot action is a combination of $(\boldsymbol{a}_t^W+\Delta\boldsymbol{a}_t^W,\boldsymbol{a}_t^F)$. Please refer to Appendix B for more detail about the observation space and the reward function.

5 Experiments

The experiments are designed to answer the following research questions: (1) Can the high-level planner generalize to unseen trajectories and unseen objects? (Sec. 5.1) (2) Does our hierarchical approach help bridge the embodiment gap between human and robot hands? (Sec. 5.2) (3) Can our trained policy generalize to unseen object geometries and goal trajectories? (Sec. 5.3) (4) Can we transfer the policy from simulation to a real-world bimanual dexterous robot system? (Sec. 5.4).

5.1 Performance of the high-level planner

Table 1 shows that Ours performs the best in generating wrist motions, with the lowest cumulative translation and orientation error.

5.2 Effectiveness of learning from human with hierarchical pipeline

Table 3 demonstrates that our hierarchical learning framework outperforms traditional hand pose matching methods (*Finger Joint Mapping*, *Fingertip Mapping*) by 50% in completion rate, indicating that our low-level RL significantly helps in bridging the embodiment gap when learning from human

		MLP	RNN	Ours
Trained	TE	5.4±0.2	5.0±0.1	**4.2**±0.2
	OE	14.6±0.7	12.4±0.8	**9.6**±0.6
Unseen Traj	TE	5.6±0.5	6.2±0.3	**5.0**±0.8
	OE	9.4±1.0	9.2±0.5	**7.2**±0.7
Unseen Object	TE	18.2±1.5	17.7±1.2	**12.4**±1.1
	OE	109.4±5.8	82.2±4.2	**75.5**±4.7

Table 1: Results for the high-level planner

	Fingertip Mapping	Vanilla RL	Ours
Box	13.1±1.8	20.4±1.8	**69.8**±6.6
Micro.	60.6±2.7	56.1±9.8	**100**±0.0
Laptop	9.2±0.5	8.6±1.4	**76.7**±4.1
Coffee.	8.2±0.6	9.2±3.1	**74.8**±3.8
Mixer	8.3±2.1	10.8±3.2	**82.8**±2.1
Notebook	4.5±0.1	4.5±0.3	**64.3**±8.4

Table 2: Results for the real-world experiments

	Fingertip Mapping	Finger Joint Mapping	Vanilla RL	Ours (w. FR)	Ours (w.o. DAL)	Ours
Box	14.6±0.3	8.9±0.2	23.5±3.5	56.2±7.4	100±0.0	**100**±0.0
Coffee Maker	9.3±0.6	9.0±0.5	10.7±2.7	78.6±1.6	71.5±2.6	**86.1**±5.5
Espresso Machine	22.2±0.4	7.0±0.8	14.3±1.5	70.7±3.5	75.4±4.3	**81.1**±8.6
Ketchup	14.8±0.7	9.5±0.2	4.9±2.7	15.2±1.7	21.8±7.2	**41.2**±13.3
Microwave	38.7±0.2	27.5±0.6	43.5±2.4	61.2±5.3	100±0.0	**100**±0.0
Mixer	21.7±0.9	10.7±0.8	42.1±1.4	42.2±4.0	44.2±6.4	**57.6**±4.9
Notebook	10.1±0.5	5.9±0.4	10.6±2.9	31.1±4.1	38.1±4.8	**38.7**±3.3
Scissors	4.2±0.5	4.1±0.6	4.4±0.6	20.7±2.0	35.9±4.0	**41.4**±14.9
Laptop	9.9±0.4	8.8±1.1	33.0±2.1	42.5±5.4	100±0.0	**100**±0.0

Table 3: Results for the experiments of using one policy per object.

	Fingertip Mapping	Finger Joint Mapping	Vanilla RL	Ours (w. FR)	Ours (w.o. DAL)	Ours
Single Obj - Trained Traj	10.4±2.8	7.1±4.8	17.1±9.6	30.8±12.1	59.6±14.4	**83.8**±9.1
Single Obj - Unseen Traj	4.8±0.3	5.5±0.8	18.8±5.7	19.6±10.1	42.5±7.5	**57.1**±10.2
Multi Obj - Trained Obj	6.9±1.7	5.3±1.2	17.7±10.1	15.5±6.2	35.2±2.8	**47.6**±4.2
Multi Obj - Unseen Obj	3.2±0.6	2.9±0.2	8.1±4.2	8.2±5.3	18.6±3.7	**36.4**±5.0

Table 4: Results for the generalization experiments.

data. Moreover, *Ours* surpasses *Vanilla RL* by 47.3% on average, underscoring the challenge of training arm and hand actions together with RL, and emphasizing the advantage of our high-level planner for guiding the RL in high-dimensional action space.

5.3 Generalization to unseen scenarios

In Table 4, our algorithm surpassing the results of *Vanilla RL* on *Single Obj - Unseen Traj* and *Multi Obj - Unseen Obj* by more than 28%. This indicates that our hierarchical structure substantially improves generalization capabilities across unseen trajectories and unseen object geometries.

5.4 Transfer from simulation to real-world

In Table 2 real-world experiments, our approach has more than a 50% completion rate improvements compared to prior methods, which barely achieve any success ($<$ 20% completion rate) on several objects. This result showcases the ability of our approach on tackling real-world bimaual dexterous manipulation tasks.

6 Conclusion

In this work, we present a hierarchical policy learning framework that effectively utilizes human hand motion data to train object-centric dexterous robot manipulation. Our approach demonstrated superior performance across various household objects and showcased generalization capabilities to novel object geometries and goal trajectories. Moreover, the successful transfer of the learned policies from simulation to a real-world bimanual dexterous robot system underscores the practical applicability of our method in real-world scenarios.

References

[1] B. Huang, Y. Chen, T. Wang, Y. Qin, Y. Yang, N. Atanasov, and X. Wang. Dynamic handover: Throw and catch with bimanual hands. *arXiv preprint arXiv:2309.05655*, 2023.

[2] Y. Chen, T. Wu, S. Wang, X. Feng, J. Jiang, Z. Lu, S. McAleer, H. Dong, S.-C. Zhu, and Y. Yang. Towards human-level bimanual dexterous manipulation with reinforcement learning. *Advances in Neural Information Processing Systems*, 35:5150–5163, 2022.

[3] Y. Zhang, A. Clegg, S. Ha, G. Turk, and Y. Ye. Learning to transfer in-hand manipulations using a greedy shape curriculum. In *Computer Graphics Forum*, volume 42, pages 25–36. Wiley Online Library, 2023.

[4] Y. Bai and C. K. Liu. Dexterous manipulation using both palm and fingers. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1560–1565. IEEE, 2014.

[5] V. Kumar, Y. Tassa, T. Erez, and E. Todorov. Real-time behaviour synthesis for dynamic hand-manipulation. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6808–6815. IEEE, 2014.

[6] M. T. Mason and J. K. Salisbury Jr. Robot hands and the mechanics of manipulation. 1985.

[7] I. Mordatch, Z. Popović, and E. Todorov. Contact-invariant optimization for hand manipulation. In *Proceedings of the ACM SIGGRAPH/Eurographics symposium on computer animation*, pages 137–144, 2012.

[8] S. Chen, J. Bohg, and C. K. Liu. Springgrasp: An optimization pipeline for robust and compliant dexterous pre-grasp synthesis. *arXiv preprint arXiv:2404.13532*, 2024.

[9] T. Pang, H. T. Suh, L. Yang, and R. Tedrake. Global planning for contact-rich manipulation via local smoothing of quasi-dynamic contact models. *IEEE Transactions on robotics*, 2023.

[10] X. Cheng, S. Patil, Z. Temel, O. Kroemer, and M. T. Mason. Enhancing dexterity in robotic manipulation via hierarchical contact exploration. *IEEE Robotics and Automation Letters*, 9 (1):390–397, 2023.

[11] T. Chen, J. Xu, and P. Agrawal. A system for general in-hand object re-orientation. *Conference on Robot Learning*, 2021.

[12] T. Chen, M. Tippur, S. Wu, V. Kumar, E. Adelson, and P. Agrawal. Visual dexterity: In-hand dexterous manipulation from depth. *arXiv preprint arXiv:2211.11744*, 2022.

[13] Z.-H. Yin, B. Huang, Y. Qin, Q. Chen, and X. Wang. Rotating without seeing: Towards in-hand dexterity through touch. *arXiv preprint arXiv:2303.10880*, 2023.

[14] H. Qi, B. Yi, S. Suresh, M. Lambeta, Y. Ma, R. Calandra, and J. Malik. General in-hand object rotation with vision and touch. In *Conference on Robot Learning*, pages 2549–2564. PMLR, 2023.

[15] H. Qi, A. Kumar, R. Calandra, Y. Ma, and J. Malik. In-hand object rotation via rapid motor adaptation. In *Conference on Robot Learning*, pages 1722–1732. PMLR, 2023.

[16] S. Dasari, A. Gupta, and V. Kumar. Learning dexterous manipulation from exemplar object trajectories and pre-grasps. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 3889–3896. IEEE, 2023.

[17] OpenAI, I. Akkaya, M. Andrychowicz, M. Chociej, M. Litwin, B. McGrew, A. Petron, A. Paino, M. Plappert, G. Powell, R. Ribas, J. Schneider, N. Tezak, J. Tworek, P. Welinder, L. Weng, Q. Yuan, W. Zaremba, and L. Zhang. Solving rubik's cube with a robot hand. *CoRR*, abs/1910.07113, 2019.

[18] M. Yang, C. Lu, A. Church, Y. Lin, C. Ford, H. Li, E. Psomopoulou, D. A. Barton, and N. F. Lepora. Anyrotate: Gravity-invariant in-hand object rotation with sim-to-real touch. *arXiv preprint arXiv:2405.07391*, 2024.

[19] J. Pitz, L. Röstel, L. Sievers, and B. Bäuml. Dextrous tactile in-hand manipulation using a modular reinforcement learning architecture. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1852–1858. IEEE, 2023.

[20] A. Handa, A. Allshire, V. Makoviychuk, A. Petrenko, R. Singh, J. Liu, D. Makoviichuk, K. Van Wyk, A. Zhurkevich, B. Sundaralingam, et al. Dextreme: Transfer of agile in-hand manipulation from simulation to reality. *arXiv preprint arXiv:2210.13702*, 2022.

[21] G. Khandate, S. Shang, E. T. Chang, T. L. Saidi, Y. Liu, S. M. Dennis, J. Adams, and M. Ciocarlie. Sampling-based exploration for reinforcement learning of dexterous manipulation. *arXiv preprint arXiv:2303.03486*, 2023.

[22] T. Lin, Z.-H. Yin, H. Qi, P. Abbeel, and J. Malik. Twisting lids off with two hands. *arXiv preprint arXiv:2403.02338*, 2024.

[23] K. Xu, Z. Hu, R. Doshi, A. Rovinsky, V. Kumar, A. Gupta, and S. Levine. Dexterous manipulation from images: Autonomous real-world rl via substep guidance. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5938–5945. IEEE, 2023.

[24] Y. Chen, C. Wang, L. Fei-Fei, and C. K. Liu. Sequential dexterity: Chaining dexterous policies for long-horizon manipulation. *arXiv preprint arXiv:2309.00987*, 2023.

[25] A. Gupta, J. Yu, T. Z. Zhao, V. Kumar, A. Rovinsky, K. Xu, T. Devlin, and S. Levine. Reset-free reinforcement learning via multi-task learning: Learning dexterous manipulation behaviors without human intervention. In *2021 IEEE International Conference on Robotics and Automation (ICRA)*, pages 6664–6671. IEEE, 2021.

[26] K. Zakka, L. Smith, N. Gileadi, T. Howell, X. B. Peng, S. Singh, Y. Tassa, P. Florence, A. Zeng, and P. Abbeel. Robopianist: A benchmark for high-dimensional robot control. *arXiv preprint arXiv:2304.04150*, 2023.

[27] A. Rajeswaran, V. Kumar, A. Gupta, G. Vezzani, J. Schulman, E. Todorov, and S. Levine. Learning complex dexterous manipulation with deep reinforcement learning and demonstrations. *arXiv preprint arXiv:1709.10087*, 2017.

[28] I. Radosavovic, X. Wang, L. Pinto, and J. Malik. State-only imitation learning for dexterous manipulation. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7865–7871. IEEE, 2021.

[29] S. P. Arunachalam, I. Güzey, S. Chintala, and L. Pinto. Holo-dex: Teaching dexterity with immersive mixed reality. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pages 5962–5969. IEEE, 2023.

[30] I. Guzey, B. Evans, S. Chintala, and L. Pinto. Dexterity from touch: Self-supervised pre-training of tactile representations with robotic play. *arXiv preprint arXiv:2303.12076*, 2023.

[31] A. Handa, K. Van Wyk, W. Yang, J. Liang, Y.-W. Chao, Q. Wan, S. Birchfield, N. Ratliff, and D. Fox. Dexpilot: Vision-based teleoperation of dexterous robotic hand-arm system. In *2020 IEEE International Conference on Robotics and Automation (ICRA)*, pages 9164–9170. IEEE, 2020.

[32] A. Sivakumar, K. Shaw, and D. Pathak. Robotic telekinesis: Learning a robotic hand imitator by watching humans on youtube. *arXiv preprint arXiv:2202.10448*, 2022.

[33] Y. Qin, W. Yang, B. Huang, K. Van Wyk, H. Su, X. Wang, Y.-W. Chao, and D. Fox. Anyteleop: A general vision-based dexterous robot arm-hand teleoperation system. *arXiv preprint arXiv:2307.04577*, 2023.

[34] P. Mandikal and K. Grauman. Learning dexterous grasping with object-centric visual affordances. In *2021 IEEE international conference on robotics and automation (ICRA)*, pages 6169–6176. IEEE, 2021.

[35] Z. Q. Chen, K. Van Wyk, Y.-W. Chao, W. Yang, A. Mousavian, A. Gupta, and D. Fox. Learning robust real-world dexterous grasping policies via implicit shape augmentation. *arXiv preprint arXiv:2210.13638*, 2022.

[36] Y. Qin, Y.-H. Wu, S. Liu, H. Jiang, R. Yang, Y. Fu, and X. Wang. Dexmv: Imitation learning for dexterous manipulation from human videos. In *European Conference on Computer Vision*, pages 570–587. Springer, 2022.

[37] Z. J. Cui, Y. Wang, N. M. M. Shafiullah, and L. Pinto. From play to policy: Conditional behavior generation from uncurated robot data. *arXiv e-prints*, pages arXiv–2210, 2022.

[38] S. Haldar, J. Pari, A. Rai, and L. Pinto. Teach a robot to fish: Versatile imitation from one minute of demonstrations. *arXiv preprint arXiv:2303.01497*, 2023.

[39] Y. Qin, H. Su, and X. 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.

[40] S. P. Arunachalam, S. Silwal, B. Evans, and L. Pinto. Dexterous imitation made easy: A learning-based framework for efficient dexterous manipulation. *arXiv preprint arXiv:2203.13251*, 2022.

[41] I. Guzey, Y. Dai, B. Evans, S. Chintala, and L. Pinto. See to touch: Learning tactile dexterity through visual incentives. *arXiv preprint arXiv:2309.12300*, 2023.

[42] T. Lin, Y. Zhang, Q. Li, H. Qi, B. Yi, S. Levine, and J. Malik. Learning visuotactile skills with two multifingered hands. *arXiv preprint arXiv:2404.16823*, 2024.

[43] Y. Ze, G. Zhang, K. Zhang, C. Hu, M. Wang, and H. Xu. 3d diffusion policy. *arXiv preprint arXiv:2403.03954*, 2024.

[44] S. Bahl, A. Gupta, and D. Pathak. Human-to-robot imitation in the wild. *arXiv preprint arXiv:2207.09450*, 2022.

[45] P. Mandikal and K. Grauman. Dexvip: Learning dexterous grasping with human hand pose priors from video. In *Conference on Robot Learning*, pages 651–661. PMLR, 2022.

[46] S. Bahl, R. Mendonca, L. Chen, U. Jain, and D. Pathak. Affordances from human videos as a versatile representation for robotics. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 13778–13790, 2023.

[47] H. Bharadhwaj, A. Gupta, V. Kumar, and S. Tulsiani. Towards generalizable zero-shot manipulation via translating human interaction plans. *arXiv preprint arXiv:2312.00775*, 2023.

[48] K. Shaw, S. Bahl, and D. Pathak. Videodex: Learning dexterity from internet videos. In *Conference on Robot Learning*, pages 654–665. PMLR, 2023.

[49] A. Bahety, P. Mandikal, B. Abbatematteo, and R. Martín-Martín. Screwmimic: Bimanual imitation from human videos with screw space projection. *arXiv preprint arXiv:2405.03666*, 2024.

[50] H. Zhang, S. Christen, Z. Fan, L. Zheng, J. Hwangbo, J. Song, and O. Hilliges. Artigrasp: Physically plausible synthesis of bi-manual dexterous grasping and articulation. *arXiv preprint arXiv:2309.03891*, 2023.

[51] Z. Q. Chen, K. Van Wyk, Y.-W. Chao, W. Yang, A. Mousavian, A. Gupta, and D. Fox. Dextransfer: Real world multi-fingered dexterous grasping with minimal human demonstrations. *arXiv preprint arXiv:2209.14284*, 2022.

[52] C. Wang, L. Fan, J. Sun, R. Zhang, L. Fei-Fei, D. Xu, Y. Zhu, and A. Anandkumar. Mimicplay: Long-horizon imitation learning by watching human play. *arXiv preprint arXiv:2302.12422*, 2023.

[53] G. Garcia-Hernando, S. Yuan, S. Baek, and T.-K. Kim. First-person hand action benchmark with rgb-d videos and 3d hand pose annotations. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 409–419, 2018.

[54] Y. Hasson, G. Varol, D. Tzionas, I. Kalevatykh, M. J. Black, I. Laptev, and C. Schmid. Learning joint reconstruction of hands and manipulated objects. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11807–11816, 2019.

[55] O. Taheri, N. Ghorbani, M. J. Black, and D. Tzionas. Grab: A dataset of whole-body human grasping of objects. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part IV 16*, pages 581–600. Springer, 2020.

[56] C. Wang, H. Shi, W. Wang, R. Zhang, L. Fei-Fei, and C. K. Liu. Dexcap: Scalable and portable mocap data collection system for dexterous manipulation. *arXiv preprint arXiv:2403.07788*, 2024.

[57] Y. Ye and C. K. Liu. Synthesis of detailed hand manipulations using contact sampling. *ACM Transactions on Graphics (ToG)*, 31(4):1–10, 2012.

[58] J. Ho and S. Ermon. Generative adversarial imitation learning. *Advances in neural information processing systems*, 29, 2016.

[59] X. B. Peng, P. Abbeel, S. Levine, and M. Van de Panne. Deepmimic: Example-guided deep reinforcement learning of physics-based character skills. *ACM Transactions On Graphics (TOG)*, 37(4):1–14, 2018.

[60] X. B. Peng, Z. Ma, P. Abbeel, S. Levine, and A. Kanazawa. Amp: Adversarial motion priors for stylized physics-based character control. *ACM Transactions on Graphics (ToG)*, 40(4):1–20, 2021.

[61] X. B. Peng, Y. Guo, L. Halper, S. Levine, and S. Fidler. Ase: Large-scale reusable adversarial skill embeddings for physically simulated characters. *ACM Transactions On Graphics (TOG)*, 41(4):1–17, 2022.

[62] Y. Wang, J. Lin, A. Zeng, Z. Luo, J. Zhang, and L. Zhang. Physhoi: Physics-based imitation of dynamic human-object interaction. *arXiv preprint arXiv:2312.04393*, 2023.

[63] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov. Proximal policy optimization algorithms. *arXiv preprint arXiv:1707.06347*, 2017.

[64] T. Yu, D. Quillen, Z. He, R. Julian, K. Hausman, C. Finn, and S. Levine. Meta-world: A benchmark and evaluation for multi-task and meta reinforcement learning. In *Conference on robot learning*, pages 1094–1100. PMLR, 2020.

[65] H. J. Charlesworth and G. Montana. Solving challenging dexterous manipulation tasks with trajectory optimisation and reinforcement learning. In *International Conference on Machine Learning*, pages 1496–1506. PMLR, 2021.

[66] B. Wen, W. Yang, J. Kautz, and S. Birchfield. Foundationpose: Unified 6d pose estimation and tracking of novel objects. *arXiv preprint arXiv:2312.08344*, 2023.

[67] S. Ross, G. Gordon, and D. 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*, pages 627–635. JMLR Workshop and Conference Proceedings, 2011.

[68] J. Li, A. Clegg, R. Mottaghi, J. Wu, X. Puig, and C. K. Liu. Controllable human-object interaction synthesis. *arXiv preprint arXiv:2312.03913*, 2023.

[69] C. Pezzato, C. Salmi, M. Spahn, E. Trevisan, J. Alonso-Mora, and C. H. Corbato. Sampling-based model predictive control leveraging parallelizable physics simulations, 2023.