Learning Multiple Peg-in-hole Assembly Skills Using Multimodal Representations and Impedance Control

Wenkai Chen[1*], Chao Zeng[1], Fuchun Sun[2] and Jianwei Zhang[1]

***Abstract*— Currently, existing RL methods are difficult to apply to multiple peg-in-hole issues due to more complicated geometric and physical constraints. In addition, previously limited solutions for multiple peg-in-hole assembly are hard to transfer into real industrial scenarios flexibly. To effectively address these issues, this work designs a novel and more challenging multiple peg-in-hole assembly setup. Based on it, multiple modalities including vision, proprioception, and force/torque are learned as compact representations to account for the complexity and uncertainties of the environment. Furthermore, RL is used in the simulation to train the policy and the learned policy is transferred to the real world without extra exploration. Domain randomization and impedance control are embedded into the policy to narrow the gap between simulation and reality. Finally, the evaluation results demonstrate that the proposed solution can achieve successful multiple peg-in-hole assembly with the ability to generalize over different object shapes in the real world.**

I. Introduction

To prompt the high-quality development of the industry, intelligent robots have become indispensable in realizing many manufacturing processes [1]. Taking the assembly task as an example, the global intelligent assembly market is expected to grow by 30% over the next four years [2]. The most obvious characteristic of an assembly task is that it involves mechanical interaction and fits between two or more objects, such as clearance fits, transition fits, and interference fits. Therefore, in order to achieve a high-precision assembly, research in multiple dimensions should be considered, such as the redundancy and clearance of the robot's own mechanical precision, pose uncertainties between peg and hole objects, and the complex physical models involved in each assembly scene, consisting of geometry, contact force, and kinematics [3]–[6]. Especially for the single peg-in-hole assembly, many studies are conducted to achieve promising results [5], [7]–[11]. However, there exist few studies on multiple peg-in-hole manipulation because of a more complicated geometric and physical interaction model [12]. Despite this, the experimental setup of previous multiple peg-in-hole has many flaws, like the peg is fixed on the end-effector, the 6-DOF pose of the holes object stays constant, the shape of the holes object and pegs object is immutable, and lacking the visual feedback. Practically, their setup with these limitations is not in line with the actual multiple peg-in-hole assembly scenes [13]–[15].

In this paper, we design a new multiple peg-in-hole assembly setup to solve the flaws mentioned above from previous work and maximize the transferability which means a successful policy learned from our setup can be easier to be deployed in real manufacturing scenarios. Based on this more challenging task, we also propose a multimodal learning architecture using reinforcement learning, where features of multiple modalities are compacted into latent representations at a high level via a tokenization-based model. It enables robotic agents to leverage the complementary nature of these sensing modalities for policy learning. A simulation environment with the setup is constructed to train the policy with a soft actor-critic (SAC) algorithm [16]. In addition, domain randomization is used in simulation to narrow the gap between the simulation and the real experimental setup. Furthermore, impedance control is designed and embedded into the proposed architecture, which helps the policy deal with our physical contact-rich task. Finally, the proposed assembly task is evaluated both in the simulation and real robot experiments, demonstrating that the proposed multiple modality-driven impedance-based policy trained with domain randomization achieves successful dynamic assembly. The primary contributions and novelties of this paper are:

1) We define a novel and more challenging experimental setup for multiple peg-in-hole assembly task, which is easier to apply to the real application scenario than previous work [13]–[15].

2) A tokenization method based on the transformer architecture is proposed to extract features from robot proprioception and force/torque signals and the extracted features are further fused with visual representations into a compact multimodal representation.

3) With domain randomization and impedance control, the policy for dynamic assembly can be learned successfully in simulation and then transferred to reality without extra exploration.

4) Experimental results show the trained policy could achieve generalization to tasks with different peg shapes under object uncertainties.

II. Methodology

An overview of our proposed architecture for multiple peg-in-hole assembly is depicted in Fig. 1. In this section, we focus on utilizing the multiple modalities in the simulated robotic environment to learn a robust policy and then transfer it to the challenging assembly task. First, we propose a

[1]Technical Aspects of Multimodal Systems (TAMS), Department of Informatics, Universität Hamburg

[2]Beijing National Research Center for Information Science and Technology (BNRist), State Key Lab on Intelligent Technology and Systems, Department of Computer Science and Technology, Tsinghua University

*Corresponding author to provide e-mail: wchen@informatik.uni-hamburg.de

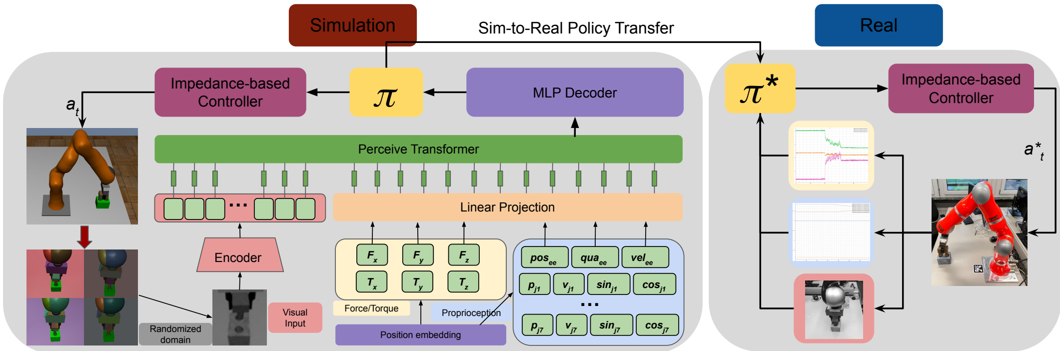

Fig. 1. Overview of the proposed architecture. The left part represents policy training in the simulation environment, where domain randomization is used to sample the interaction of the assembly task like color, lighting, camera, and robotic dynamics (left). During the training process, multiple modalities including visual image, proprioception, and force/torque signals are all tokenized and fused into a perceived transformer module. Each predicted policy π is embedded into an impedance controller to execute torque commands to control the manipulator. Finally, the trained policy π^* is transferred to the real world without additional exploration (right).

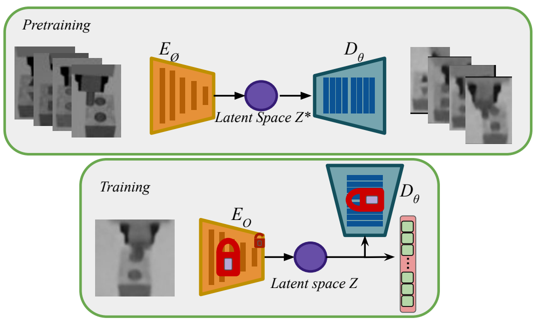

Fig. 2. Overview of the visual representation module. 1) pretraining: We first pre-train the feature encoder and decoder. After that, the latent space Z^* to recognize different object shapes is obtained. 2) training: the decoder and most encoder layers are frozen, and we train the encoder during the policy learning process.

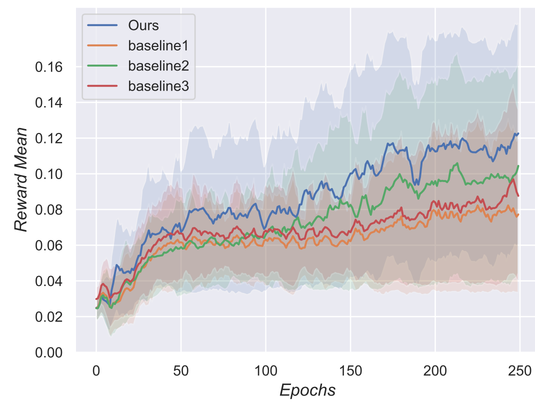

Fig. 3. Ablation study for the proposed architecture considering the effect of the unfrozen encoder and transformer module.

pretraining-training approach to learn the latent representation from the visual frames where domain randomization is used to collect a robust assembly dataset, shown in Fig. 2. With the pretraining of the encoder-decoder architecture, a prior of the visual image for each standard peg and hole shape can be learned. To speed up the whole policy learning, we further freeze all encoder feature layers except the output layer. Though in the pretraining process, we need multiple images of the different hole and peg shapes. Our final pipeline only needs a single image captured by the camera at inference time during RL training. Second, a self-attention-based transformer architecture is applied to learn the dependencies between robot proprioception states, force-torque signals in the Cartesian space, and the extracted visual embedding. Next, the predicted actions are mapped to the impedance space after decoding the compact representation, and the SAC-based RL algorithm is trained to achieve action control. Finally, the trained policy is transferred to different real assembly tasks directly and adapted to new task situations.

The encoded feature vector from perceive transformer is input into an MLP decoder with three hidden layers to predict the gripper and end-effector actions. The action space in our task is 7-dimensional, consisting of moving of position, orientation from the end-effector, and the open/close state of the gripper. The action a_t is defined as the difference between the current kinematic state and the desired kinematic state:

$$a_t^{pos} = s_{t+1}^{pos} - s_t^{pos}, a^{pos} \in [\Delta x, \Delta y, \Delta z] \quad (1)$$

$$a_t^{ori} = s_{t+1}^{ori} * inv(s_t^{ori}), a^{ori} \in [\Delta\alpha, \Delta\beta, \Delta\gamma] \quad (2)$$

$$a_t^{gri} = s_{t+1}^{gri} - s_t^{gri}, a^{ori} \in [0, 1] \quad (3)$$

where t is the timestamp of the current action, and s^{pos}, s^{ori}, and s^{grpi} represent the scalar value for position, orientation, and gripper.

III. Results

First, to evaluate the proposed architecture, ablation studies about the unfrozen encoder module from visual representation and the perceived transformer module from multimodal tokenization are analyzed. Fig. 3 shows that the proposed architecture could achieve better performance, where baseline 1 means without both modules, baseline 2 means without the perceive transformer, and baseline 3 means without the unfrozen encoder. However, the success rate of the proposed approach is not so high because we find the peg grasped by the gripper cannot be reliably fixed at

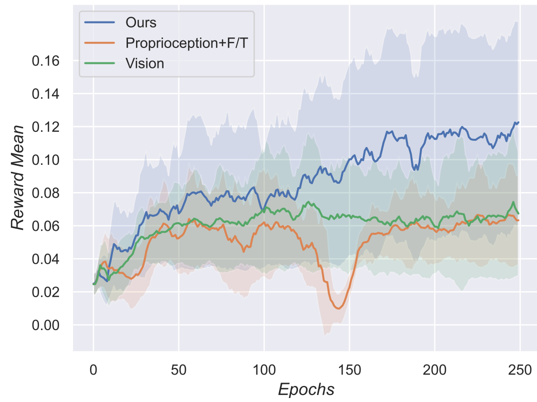

Fig. 4. Comparison of multimodal and various single-modal training.

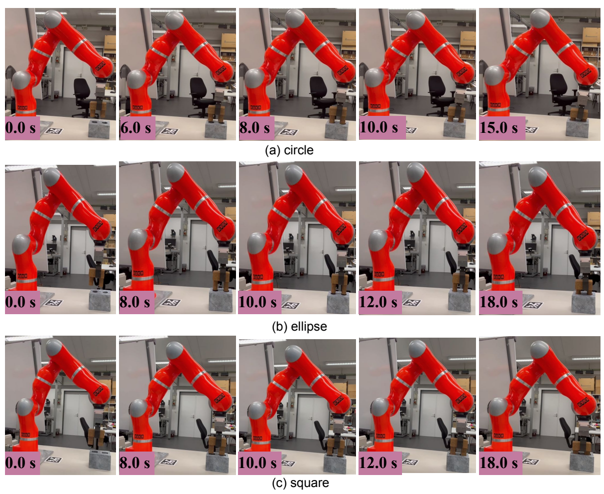

Fig. 5. Example of the proposed solution on multiple peg-in-hole assembly for different trained object shapes (a) circle (b) ellipse (c) square.

TABLE I

THE COMPARISON OF FINAL SUCCESS RATE AND ASSEMBLY TIME

Object shape	Circle	Ellipse	Square	Triangle*	Triangle
Time (s)	15.0	18.0	18.0	24.2	17.6
Success rate % avg (std)	90.9 (9.09)	77.3 (13.6)	81.8 (9.09)	54.5 (9.09)	81.8 (9.09)

Note: the untrained object shape is represented with *.

the same position in the MuJoCo simulator especially when the pegs collide with the holes object, causing some failing cases during pose adjustment and search depth periods.

Furthermore, we discuss the effect of different modalities on our challenged assembly task. Our task introduces flexible grippers and a movable pose of the holes object, so vision is necessary to obtain the relative position between pegs and holes. Based on this, we use pure visual features and pure state information inside the robot to compare the performance of multi-modal fusion. As shown in Fig. 4, the training curves demonstrate that the fusion of multiple modalities can significantly improve the performance of this task.

Finally, by randomly adjusting the position and orientation of the holes object in a limited range, we run 22 evaluations for each object in the real experiment. Fig. 5 shows the robotic multiple peg-in-hole assembly processes for different trained object shapes. seen from Table. I, the trained policy from the circle, ellipse, and square could be generalized to the new triangle shape. The mean assembly time of each object shape is also computed to justify the relationship between task difficulty and assembly performance. Finally, after adapting to the simulation environment again, the test on the triangle object shape achieves a significant improvement of over 25% in the success rate and a reduction of over 6.5s in the assembly time, which demonstrates that our model has a good generalization ability over uncertainties and object shapes. As seen in Table.II, we further retrain the policy without adding domain randomization for all kinds of object shapes. We found that the success rate for circle, ellipse, square, and triangle shapes are 13.6%, 9.09%, 0, and 0, respectively. And the final trained model cannot generalize to a new object shape. Furthermore, we retrain the policy without adding impedance control and introduce basic position control to execute the output action. It can be seen that the success rates for all object shapes are significantly lower than our method, demonstrating the embedding of impedance skills is necessary to improve the assembly performance.

TABLE II

THE COMPARISON OF FINAL SUCCESS RATE CONSIDERING THE EFFECT OF DOMAIN RANDOMIZATION AND IMPEDANCE CONTROL

Object shape	Circle	Ellipse	Square	Triangle*	Triangle
w/o domain randomization % avg (std)	13.6 (9.09)	9.09 (9.09)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)
w/o impedance controller % avg (std)	63.6 (13.6)	54.5 (9.09)	50.0 (9.09)	22.7 (9.09)	45.5 (13.6)

Note: the untrained object shape is represented with *.

IV. CONCLUSION AND DISCUSSION

We present a solution for the multiple peg-in-hole assembly task using a multimodal representation by transferring the trained policy in simulation to the real world without extra exploration. A special visual representation module and tokenization-based transformer module are separately proposed to compact the feature as the backbone of reinforcement learning. Furthermore, the policy learning process also incorporates domain randomization and an impedance controller, which speeds up the transferring process and narrows the gap between simulation and reality. Experimental results on a real robot show that our solution could achieve a high success rate for a more challenging multiple peg-in-hole assembly setup, and the generalization ability is also validated by different object shapes.

This work sets the number of pegs and holes as two and the experimental objects consist of four kinds of shapes. In the future, we need to continue studying smart assembly involved in more complicated object interactions in terms of object number, object size, and object shape. We believe that further research about robotic multiple peg-in-hole assembly based on reinforcement learning can significantly improve the efficiency of related manufacturing processes.

References

[1] J. Arents and M. Greitans, “Smart industrial robot control trends, challenges and opportunities within manufacturing,” *Applied Sciences*, vol. 12, no. 2, p. 937, 2022.

[2] Í. Elguea-Aguinaco, A. Serrano-Muñoz, D. Chrysostomou, I. Inziarte-Hidalgo, S. Bøgh, and N. Arana-Arexolaleiba, “A review on reinforcement learning for contact-rich robotic manipulation tasks,” *Robotics and Computer-Integrated Manufacturing*, vol. 81, p. 102517, 2023.

[3] K. Valavanis and K. Stellakis, “A general organizer model for robotic assemblies and intelligent robotic systems,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 21, no. 2, pp. 302–317, 1991.

[4] C.-H. Wu and M. G. Kim, “Modeling of part-mating strategies for automating assembly operations for robots,” *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 24, no. 7, pp. 1065–1074, 1994.

[5] T. Inoue, G. De Magistris, A. Munawar, T. Yokoya, and R. Tachibana, “Deep reinforcement learning for high precision assembly tasks,” in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 819–825.

[6] H. Chen, G. Zhang, H. Zhang, and T. A. Fuhlbrigge, “Integrated robotic system for high precision assembly in a semi-structured environment,” *Assembly Automation*, 2007.

[7] Y. Chebotar, A. Handa, V. Makoviychuk, M. Macklin, J. Issac, N. Ratliff, and D. Fox, “Closing the sim-to-real loop: Adapting simulation randomization with real world experience,” in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 2019, pp. 8973–8979.

[8] C. C. Beltran-Hernandez, D. Petit, I. G. Ramirez-Alpizar, and K. Harada, “Variable compliance control for robotic peg-in-hole assembly: A deep-reinforcement-learning approach,” *Applied Sciences*, vol. 10, no. 19, p. 6923, 2020.

[9] A. A. Apolinarska, M. Pacher, H. Li, N. Cote, R. Pastrana, F. Gramazio, and M. Kohler, “Robotic assembly of timber joints using reinforcement learning,” *Automation in Construction*, vol. 125, p. 103569, 2021.

[10] M. Hebecker, J. Lambrecht, and M. Schmitz, “Towards real-world force-sensitive robotic assembly through deep reinforcement learning in simulations,” in *2021 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM)*. IEEE, 2021, pp. 1045–1051.

[11] F. Li, Q. Jiang, S. Zhang, M. Wei, and R. Song, “Robot skill acquisition in assembly process using deep reinforcement learning,” *Neurocomputing*, vol. 345, pp. 92–102, 2019.

[12] K. Sathirakul and R. H. Sturges, “Jamming conditions for multiple peg-in-hole assemblies,” *Robotica*, vol. 16, no. 3, pp. 329–345, 1998.

[13] Z. Hou, H. Dong, K. Zhang, Q. Gao, K. Chen, and J. Xu, “Knowledge-driven deep deterministic policy gradient for robotic multiple peg-in-hole assembly tasks,” in *2018 IEEE International Conference on Robotics and Biomimetics (ROBIO)*. IEEE, 2018, pp. 256–261.

[14] J. Xu, Z. Hou, W. Wang, B. Xu, K. Zhang, and K. Chen, “Feedback deep deterministic policy gradient with fuzzy reward for robotic multiple peg-in-hole assembly tasks,” *IEEE Transactions on Industrial Informatics*, vol. 15, no. 3, pp. 1658–1667, 2018.

[15] X. Li, J. Xiao, W. Zhao, H. Liu, and G. Wang, “Multiple peg-in-hole compliant assembly based on a learning-accelerated deep deterministic policy gradient strategy,” *Industrial Robot: the international journal of robotics research and application*, 2021.

[16] T. Haarnoja, A. Zhou, P. Abbeel, and S. Levine, “Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor,” in *International conference on machine learning*. PMLR, 2018, pp. 1861–1870.