

СИНТЕТИЧЕСКИЙ НАБОР ДАННЫХ И БАЗОВЫЕ СТРАТЕГИИ СЕНСОМОТОРНЫХ НАВЫКОВ РОБОТИЗИРОВАННЫХ СБОРОЧНЫХ ОПЕРАЦИЙ

Друк А.Д. (университет ИТМО)

Научный руководитель – доктор технических наук, профессор Колюбин С.А.
(университет ИТМО)

Введение. Активное развитие методов глубокого обучения тесно связано с возрастающей доступностью объёмных данных. Область управления роботизированными манипуляторами — не исключение: для неё уже созданы специализированные симуляционные среды и крупные датасеты, охватывающие различные задачи [1]. Однако существующие публичные датасеты преимущественно сосредоточены на базовых операциях, таких как перемещение объектов, и не учитывают сложность контактных взаимодействий, например, сборку или точную вставку деталей. Настоящее исследование направлено на заполнение этого пробела путём создания специализированного датасета, включающего базовые навыки контактного взаимодействия, предварительно обученные в симуляторе.

Основная часть. Задачи роботизированной сборки остаются актуальными на протяжении десятилетий, несмотря на их кажущуюся простоту. Методы, основанные на данных, например, обучение с подкреплением (Reinforcement Learning, RL), обладают потенциалом для решения подобных задач в различных условиях благодаря способности агента адаптироваться и обучаться. Однако применение RL к задачам с длительным горизонтом планирования сопряжено с трудностями. Во-первых, требуется разработка узкоспециализированных функций вознаграждения, которые сложно формализовать. Во-вторых, проблема отложенного вознаграждения усложняет обучение: для получения положительного отклика алгоритм должен выполнить серию полезных действий, что затрудняет оптимизацию стратегии.

Для преодоления этих ограничений разрабатываются иерархические и композиционные методы, где обучение с подкреплением используется на нижних уровнях для решения примитивных задач. Например, в работе [2] RL применяется для генерации низкоуровневых навыков (поворот, перемещение, перехват объекта в антропоморфной руке), которые затем интегрируются в высокоуровневый контроллер. Успех таких подходов подчёркивает важность создания качественных датасетов, охватывающих широкий спектр базовых операций. В данной работе рассматривается обучение RL стратегий базовым навыкам вставки деталей различной формы в отверстие в симуляции и формирование набора решений, для обучения новых модифицированных моделей.

Экспериментальная часть включает обучение RL-политики на основе алгоритма PPO (Proximal Policy Optimization) задаче типа peg-in-hole. Для повышения устойчивости модели к реальным условиям применяется доменная рандомизация параметров (масса, размер деталей, коэффициент трения), а также вводятся помехи: смещение начальной позиции, вариации освещения и визуальные шумы. Функция вознаграждения объединяет три компонента, отражающих требования к сборочным операциям: точность совмещения осей детали и отверстия, прогресс вставки по глубине, минимизация боковых сил для предотвращения заклинивания. Наблюдения включают данные с камер, показатели силы и момента, а также состояние манипулятора.

Выводы. Исследование подтверждает потенциал симуляционных сред и обучения с подкреплением для генерации датасетов, ориентированных на контактные задачи. Итоговый датасет содержит временные ряды наблюдений, действий и меток успешности, аннотированные по фазам операции (например, захват, совмещение, вставка). Его структура обеспечивает совместимость с различными алгоритмами машинного обучения, что открывает

возможности для обучения высокоуровневых моделей, способных обобщать навыки на новые условия. Дальнейшие исследования могут быть направлены на расширение датасета за счёт дополнительных типов контактных задач и интеграции мультимодальных данных.

Список использованных источников:

1. Fang, H., Fang, H., Tang, Z., Liu, J., Wang, J., Zhu, H., & Lu, C. (2023). RH20T: A Comprehensive Robotic Dataset for Learning Diverse Skills in One-Shot. *2024 IEEE International Conference on Robotics and Automation (ICRA)*, 653-660.
2. Yin, Z., Wang, C., Pineda, L., Hogan, F., Bodduluri, K., Sharma, A., Lancaster, P., Prasad, I., Kalakrishnan, M., Malik, J., Lambeta, M., Wu, T., Abbeel, P., & Mukadam, M. (2025). DexterityGen: Foundation Controller for Unprecedented Dexterity. *arXiv preprint arXiv:2502.04307*.