

УДК 004.896

**ИССЛЕДОВАНИЕ МОДЕЛЕЙ НА БАЗЕ ТРАНСФОРМЕРОВ ДЛЯ УПРАВЛЕНИЯ  
РОБОТИЗИРОВАННЫМИ МАНИПУЛЯТОРАМИ И РАССМОТРЕНИЕ  
ВОЗМОЖНЫХ МОДИФИКАЦИЙ ДЛЯ УЧЕТА ДИНАМИКИ В КОНТАКТНЫХ  
ЗАДАЧАХ**

**Титов Д.С.** (Университет ИТМО)

**Научный руководитель – д.т.н., профессор Колюбин С.А.**  
(Университет ИТМО)

**Введение.** Современные достижения в области искусственного интеллекта и робототехники позволяют разрабатывать все более сложные системы управления для роботизированных манипуляторов. Настоящие исследования фокусируются на использовании трансформеров для формирования токенов действий или генерации функций вознаграждений для обучения на основе естественного языка, а также моделях имитационного обучения для осуществления для передачи навыков. Однако эти модели не имеют продвинутого механизма контроля усилий при контактных задачах. Мы можем разместить силомоментные датчики на манипуляторе, но собрать данные действий человека для обучения, которые бы включали нужные нам усилия проблематично, т.к. это потребует разработки сложного механизма считывания контактных сил от рук человека.

**Основная часть.** Мы можем применить следующий подход: получаем приближенные решения задач с помощью моделей, не учитывающих динамику взаимодействия с объектом; используем AI агента EUREKA[1] и среду Isaac Gym для получения ряда частных решений, учитывающих силы взаимодействия конечного эффектора с объектом; модифицируем изначальную модель и обучаем её на расширенном наборе данных.

Рассмотрим подход в рамках моделей семейства RT[2]. Сначала мы получим решение требуемых задач с помощью моделей обученные без учёта сил и моментов. В результате мы получим 7-мерные векторы действий. Далее мы используем эти данные для RL обучения тех же задач в среде Isaac Gym. Воспользуемся агентом EUREKA для формирования функций вознаграждения нацеленных на получение оптимальных значений сил и моментов для выполнения задачи. Таким образом мы получим решение конкретной задачи по кинематике приближенное к изначальному решению, но уже учитывающее силы и моменты взаимодействия с объектами.

Произведя данные манипуляции для ряда различных задач и манипуляторов, требующих точного контроля силы, мы получим набор решений, из которых можно извлечь данные для обучения новых модифицированных моделей, разработанных на базе старых архитектур.

Далее мы можем произвести обучение используя грубое выровненное пространство действий и наблюдений, предложенное при создании X-embodiment[2]. Выберем один канонический вид камеры из каждого набора данных в качестве входного изображения, изменим его размер до общего разрешения и преобразуется исходный набор действий в 13-мерное действие конечного эффектора (7 изначальных параметров и 6 дополнительных, учитывающих силы и моменты).

Таким образом модифицированные модели RT также будут принимать визуальный ввод и инструкцию на естественном языке, описывающую задачу, и выводить тонизированное действие, но действие уже будет токенизировано в ячейки, равномерно распределенные по каждому из пятнадцати измерений, а не восьми: одно измерение для завершения эпизода, семь изначальных измерений для перемещения энд эффектора и управления захватом, семь расширенных измерений для контроля сил взаимодействия энд эффектора с окружающей средой и силы сжатия захвата.

Также данный подход мы можем применить для модификации модели имитационного обучения MimicPlay[3], применяемой для передачи навыков. Данная модель обучается на данных игры человека и демонстрации роботов, управляемых человеком. По данным данным

человеческой игры модель создаёт скрытые планы для предсказания траектории человеческой руки на основе целевого изображения. По данным демонстрации робота, низкоуровневый контроллер генерирует действия робота из скрытых планов, этот момент мы можем использовать для модификации модели.

Низкоуровневая стратегия обучается с помощью алгоритма управляемого клонирования поведения. Мы можем собирать данные при помощи телеоперации используя манипулятор с силомоментными датчиками, либо, как и в случае моделей RT использовать AI агента EUREKA для формирования набора данных для обучения, включающий силы и моменты. Т.к. модель имеет тенденцию к кластеризации в латентном пространстве, используя данные по требуемым силам и языковые данные, используемые для формирования наград, мы получим ассоциацию действий с определёнными силами и языковым контекстом.

**Выводы.** Предложенный подход позволяет получить через симуляцию данные оптимального динамического взаимодействия конечного эффектора с объектами для дальнейшего обучения. Данный подход позволяет быстро собрать данные при наличии необходимых мощностей для вычислений, однако ограничен возможностями используемой среды. Для решения проблемы динамики в контактных задачах предложена концепция модифицированного вектора действий, который контролирует конечный эффектор (позиция  $x$ ,  $y$ ,  $z$ , крен, тангаж, рыскание, открытие захвата, компоненты силы по осям  $x$ ,  $y$ ,  $z$  и три момента сил, а также сила сжатия захвата).

#### **Список использованных источников:**

1. Ma Y.J., Liang W., Wang G., Huang D.-A., Bastani O., Jayaraman D., Zhu Y., Fan L., Anandkumar A. Eureka: Human-Level Reward Design via Coding Large Language Models // arXiv:2310.12931 [cs.RO]. – 2023. – 19 октября. DOI: 10.48550/arXiv.2310.12931.

2. Open X-Embodiment Collaboration, Padalkar A., Pooley A., Mandlekar A., Jain A., Tung A., Bewley A., Herzog A., Irpan A., Khazatsky A., Rai A., Singh A., Garg A., Brohan A., Raffin A., Wahid A., Burgess-Limerick B., Kim B., Schölkopf B., Ichter B., Lu C., Xu C., Finn C., Xu C., Chi C., Huang C., Chan C., Pan C., Fu C., Devin C., Driess D., Pathak D., Shah D., Büchler D., Kalashnikov D., Sadigh D., Johns E., Ceola F., Xia F., Stulp F., Zhou G., Sukhatme G.S., Salhotra G., Yan G., Schiavi G., Kahn G., Su H., Fang H.-S., Shi H., Ben Amor H., Christensen H.I., Furuta H., Walke H., Fang H., Mordatch I., Radosavovic I., Leal I., Liang J., Abou-Chakra J., Kim J., Peters J., Schneider J., Hsu J., Bohg J., Bingham J., Wu J., Wu J., Luo J., Gu J., Tan J., Oh J., Malik J., Booher J., Tompson J., Yang J., Lim J.J., Silvério J., Han J., Rao K., Pertsch K., Hausman K., Go K., Gopalakrishnan K., Goldberg K., Byrne K., Oslund K., Kawaharazuka K., Zhang K., Rana K., Srinivasan K., Chen L.Y., Pinto L., Fei-Fei L., Tan L., Ott L., Lee L., Tomizuka M., Spero M., Du M., Ahn M. et al. "Open X-Embodiment: Robotic Learning Datasets and RT-X Models" // arXiv:2310.08864 [cs.RO]. – 2023. – 13 октября. DOI: 10.48550/arXiv.2310.08864.

3. Wang C., Fan L., Sun J., Zhang R., Fei-Fei L., Xu D., Zhu Y., Anandkumar A. "MimicPlay: Long-Horizon Imitation Learning by Watching Human Play" // arXiv:2302.12422 [cs.RO], 24 февраля 2023, последняя ревизия 13 октября 2023. Презентация на 7-й Конференции по Обучению Роботов (CoRL 2023). DOI: 10.48550/arXiv.2302.12422.