Physical AI — путь в прекрасное далёко.



ПРАВДА РЕПЛИКАНТА

ГАЗЕТА О ФИЗИЧЕСКОМ ИСКУССТВЕННОМ ИНТЕЛЛЕКТЕ

ИТИКАВА (Япония)

Прогноз · 2026-11-03 · JST

УТРО · 09:00	15°	Облачно Ветер: С 1.63 м/с · 0%
ДЕНЬ · 12:00	18.1°	Облачно Ветер: С 2.04 м/с · 0%
ВЕЧЕР · 18:00	17.4°	Облачно Ветер: В 0.81 м/с · 0%

За сутки: 12.9–19.5 °C
Вероятность осадков: 4%

Всё, что мы можем, — учиться быть людьми, пока идёт дождь.

ВЫПУСК № 6 · 2026-11-03 rev. 1 ·

Погода: источник: Open-Meteo · Получено: 2026-11-03 13:25 JST

Робототехника: данные, контакт и безопасность

Пять исследований проверяют, как представление действий, видео, контакт и проверка риска меняют результат управления.

Новости

WorldBridge: видео новых задач без разметки повышает успех до 55,90%

Представь, что тебе показали видео, как человек раскладывает вещи по полкам, но не рассказали, какие именно движения нужно повторить. Роботу тоже трудно учиться только по картинке. Группа исследователей во главе с Цзюнем Цэнем изучила, может ли видео помочь роботу справляться с заданиями, которых не было среди обычных тренировочных примеров.

Для этого учёные использовали обучение по видео: система пытается представить, как будет выглядеть следующий момент действия. Другая часть системы подбирает движения, которые могли бы привести к такой картине. Это не означает, что робот просто копирует человека: у него другие руки, размеры и возможности, а движение приходится рассчитывать отдельно.

В виртуальном испытании с новыми заданиями первоначальная система выполняла их успешно в 25,07% случаев. Когда ей добавили видео этих заданий без записанных команд, результат поднялся до 55,90%. А после одного примера с командами для каждого нового задания он достиг 70,70%. Эти числа относятся к определённому компьютерному набору испытаний, а не ко всем роботам.

WorldBridge объединяет два экспертных компонента: предсказатель визуальной последовательности и генератор действий. Их связывают структурированное внимание и совместное согласование потоков, а память поддерживает два временных масштаба.

Авторы отдельно измеряют пользу видео неизвестных задач и одного дополнительного примера с размеченными действиями. Последний поднимает успех до 70,70%, но требует иной объём супервизии, поэтому сравнение с режимом без меток должно явно фиксировать бюджет данных.

Исследование показывает, почему видео может быть полезным учебным материалом. Но пока нельзя утверждать, что любой робот научится по одному ролику собирать шкаф или готовить обед дома. В статье проверяли конкретные задания и условия.

Указанные 96,23% для полной конфигурации RoboTwin 2.0 относятся к другому протоколу, а не к испытанию на новых задачах. Переход на другой захват, камеру или реальный манипулятор в этих цифрах не проверен. Воспроизводимый эксперимент обязан разделять видео, метки действий и тестовые задания.

Источники

WorldBridge: Bridging Video and Action for Task Generalization in World Action Models · 2026-10-29

281 политика в реальных испытаниях вставки разъёма

Вставить маленький сетевой разъём в нужное гнездо труднее, чем кажется. Нужно правильно повернуть его, подвести ровно и не промахнуться. Исследователи под руководством Аммара Аббаса сравнили способы обучения руки робота с пятью пальцами. Они провели около 1700 попыток на настоящем оборудовании, а не только на экране.

Всего учёные испытали 281 вариант управляющей программы из семи групп. Они меняли, какую информацию программа получает от камер и как

записывает движение пальцев. Это называется представлением данных: одну и ту же задачу можно описать разными числами и картинками. Иногда способ описания влиял на результат сильнее, чем обычная настройка обучения.

Исследователи также проверили 67 показателей, которые можно вычислить без настоящей попытки вставить разъём. Только пять показателей, связанных с плавностью и прямотой движения, хорошо совпадали с успехом, и то лишь внутри одной группы программ. Поэтому красивое движение на экране ещё не гарантирует, что пальцы попадут в гнездо.

Для методов на облаках точек обрезка нужной области, система координат относительно основания робота и нормализация признаков дали по 70–100 процентных пунктов улучшения в отдельных контролируемых сравнениях. Для модели, связывающей зрение, текст и действие, переход от относительных команд к абсолютным дал до 70 пунктов. Это результаты разных экспериментов; их нельзя суммировать или переносить на любую архитектуру. Авторы проверяли семь семейств политик, что позволяет увидеть различия между способами представления данных.

Результаты относятся к определённой руке, разъёму и способам испытаний. Они не говорят, что одна программа лучше для любых предметов. Зато теперь яснее, почему перед использованием робота нужно дать ему поработать с настоящими вещами.

Из 67 автономно вычисляемых метрик только пять надёжно коррелировали с успехом, и только внутри одной семьи моделей. Корреляции от −0,70 до −0,80 для плавности и прямоты предсказанной траектории не означают, что эти показатели заменят испытание на оборудовании. Особенно уязвима попытка выбирать модель исключительно по потере на проверочной выборке.

Источники

From Teleoperation to Long-Horizon Dexterity: Evaluating Visuomotor Skills and Offline Proxies for Dexterous Manipulation · 2026-10-29

Контактные поверхности вместо копирования углов пальцев

Когда берёшь стакан, важно не просто согнуть пальцы. Нужно коснуться его в подходящих местах, иначе он выскользнет. Исследователи во главе с Чжисюанем Сюем решили учить роботизированные руки по тому, как меняются места прикосновения к предмету. Для этого они использовали изображения с камер у кончиков пальцев и описание поверхности контакта.

Один этап их метода записывает, где пальцы человека касались вещи. Другой помогает роботу выбрать свои места прикосновения и движения кисти. Это важно потому, что человеческая ладонь и рука робота могут сильно отличаться. Слепое повторение углов человеческих суставов не обязательно позволит роботу удержать тот же предмет.

Метод сравнили с другими вариантами на шести сочетаниях заданий и роботизированных рук. Авторы сообщили, что средний успех на заключительном этапе вырос как минимум на 50 процентных пунктов. Процентный пункт означает разницу между долями успеха: например, рост с 20% до 70% равен 50 пунктам.

Архитектура разделяет интерфейс демонстрации и интерфейс обучения. Первый совмещает общую геометрию контакта с цветными изображениями от камер у кончиков пальцев и при переносе сначала согласует контакт, а затем позу. Второй использует планировщик относительных траекторий контактных поверхностей и низкоуровневое управление кистью и пальцами. Авторы опубликовали программное обеспечение и аппаратные чертежи, что облегчает независимые эксперименты с другими кистями.

В опытах участвовали определённые модели рук и набор заданий. Это не доказательство, что новый способ уже умеет обращаться с любой посудой или инструментом. Но внимание к самому прикосновению помогает объяснить, почему похожее движение пальцев иногда даёт совсем разный результат.

Результат «не менее 50 пунктов» относится к заключительному этапу шести пар рука–задача, а не к любой манипуляции целиком. Изменение материала предмета, ошибок калибровки и начального подвода руки требует отдельной проверки. Метод сохраняет прежде всего характер соприкосновения, а не визуальное сходство человеческих и роботизированных поз.

Источники

Contact Imitation Interface: Transferring Dexterous Skills by Showing How Contact Unfolds · 2026-10-29

TRACE снижает столкновения при слежении за человеком

Представь робота, который должен идти за определённым человеком в людном коридоре. Если перед ним пройдёт кто-то в похожей одежде, робот может ошибиться. Группа исследователей во главе с Минъяном Люем проверила, как улучшение слежения за человеком влияет на безопасность движения.

Когда исследователи обучили программу не терять человека из виду, она стала успешнее выполнять задание. Но в некоторых испытаниях выросло число столкновений. Поэтому учёные добавили отдельную проверку опасности. Она смотрит на признаки того, где находится нужный человек и насколько уверенно робот его отличает от других.

Если риск слишком высок, программа уменьшает шаг вперёд, который робот собирался сделать. При этом она не меняет его поворот. В испытаниях более 98% шагов остались такими же, как раньше. А число столкновений в сложной проверке снизилось на 3,96 процентного пункта по сравнению с вариантом без дополнительной защиты.

Авторы дообучили только головную часть планировщика компактной модели зрения, языка и действий, сохранив основной блок фиксированным. Дополнительный оценщик риска использует 84 признака, включая кандидатов на цель, разницу уверенности, оценку расстояния и направления и команду робота. При превышении порога он уменьшает поступательное смещение следующей точки пути, не меняя поворот. Таким образом, вмешательство привязано к состоянию цели, а не является постоянным ограничением скорости.

Такой результат не означает, что робот теперь безопасен в любом торговом центре. Исследователи сравнивали программы в определённых условиях и не объявляли их полностью защищёнными от всех ошибок. Особенно интересно, что просто постоянное замедление не дало такого же эффекта.

В одновременном повторном сравнении снижение столкновений составило 3,77 пункта против 0,29 пункта в контрольных парах. Метод не даёт формальной гарантии безопасности за пределами испытанных сценариев. Постоянное замедление без анализа положения цели не воспроизвело эффект — это важная граница между фильтром и простым регулятором скорости.

Источники

TRACE: Target-State Risk-Aligned Conditional Execution for Embodied Tracking · 2026-10-30

Hourglass переносит навигацию между разными роботами

В узком коридоре маленький робот на колёсах и большой робот с другой системой поворота не могут ехать совершенно одинаково. Команда «поверни направо» приведёт к разным движениям. Группа Чжидуна Вана предложила учить одну программу планировать путь для нескольких видов роботов, не забывая об их различиях.

В основе метода лежит модель мира: программа пытается предсказать, как изменится картинка вокруг робота после движения. Но одной картинки недостаточно. Учёные добавили описание того, как именно конкретный робот разгоняется и поворачивает. Поэтому программа может выбирать путь с учётом возможностей машины, а не только свободного места впереди.

В компьютерных испытаниях участвовали семь роботов пяти типов движения. Общая программа выполняла задания успешнее некоторых сравниваемых способов на 3,8–10,6 процентного пункта. Затем исследователи три минуты управляли настоящим роботом вручную, чтобы узнать особенности его движения, и проверили путь в незнакомом помещении.

Авторы связывают общую модель мира с дифференцируемой упрощённой моделью движения конкретного робота через изменение положения камеры. Общий предсказатель обучен на данных нескольких платформ, а параметры кинематики идентифицируют отдельно. При мысленном развёртывании будущего команда сначала проходит через модель робота, затем через предсказатель сцены. Это позволяет учитывать ограничения поворота и скорости без обучения отдельной политики для каждого типа машины.

Проверка настоящего робота важна, но она не означает, что программа уже подходит для любого транспорта или лестниц. В работе отдельно говорится о движении по ровной поверхности и определённых типах машин.

В реальном опыте параметры новой платформы определяли после трёх минут ручного движения, затем проверяли навигацию в незнакомых помещениях. Но центральное предположение о ровной поверхности не охватывает лестницы, крутые уклоны и контактное преодоление препятствий. Числовое превосходство над базовыми методами получено в симуляции, а реальный опыт подтверждает более узкий сценарий переноса.

Источники

Cross-Embodiment Motion Planning by Coupling Body-Agnostic World Model with Differentiable Ego-Dynamics · 2026-10-30

Анализ

Представление данных меняет результат

Качество зависит не только от обучения модели, но и от системы координат, данных о контакте и исполнительных ограничений.

Числа из разных протоколов нельзя напрямую сравнивать без проверки аппаратуры и доступных данных.

Источники

WorldBridge: Bridging Video and Action for Task Generalization in World Action Models · 2026-10-29

From Teleoperation to Long-Horizon Dexterity: Evaluating Visuomotor Skills and Offline Proxies for Dexterous Manipulation · 2026-10-29

Contact Imitation Interface: Transferring Dexterous Skills by Showing How Contact Unfolds · 2026-10-29

TRACE: Target-State Risk-Aligned Conditional Execution for Embodied Tracking · 2026-10-30

Cross-Embodiment Motion Planning by Coupling Body-Agnostic World Model with Differentiable Ego-Dynamics · 2026-10-30

Словарь выпуска

Ключевые понятия этого выпуска.

01 **WorldBridge**

Видео и команды робота связывающая модель для новых задач.

02 **RoboTwin 2.0**

Виртуальный набор задач манипулирования для сравнения методов.

03 **представление действий**

Формат записи команд и координат, влияющий на обучение и исполнение.

04 **облако точек**

Трёхмерные точки, описывающие видимую поверхность предметов.

05 **Contact Imitation Interface**

Метод переноса навыков по геометрии контакта пальцев.

06 **траектория контакта**

Последовательность положений участка соприкосновения пальца и предмета.

07 **TRACE**

Фильтр риска для корректировки движения при слежении за человеком.

08 **оценка риска**

Расчёт вероятности нежелательного события по признакам ситуации.

09 **Hourglass**

Архитектура общей навигации с отдельной динамикой каждой платформы.

10 **модель мира**

Предсказатель изменений наблюдаемой сцены после действия.

11 **разъём RJ45**

Сетевой соединитель, использованный в эксперименте с точной вставкой пятипальцевой рукой.

Источники

WorldBridge: Bridging Video and Action for Task Generalization in World Action Models · 2026-10-29

From Teleoperation to Long-Horizon Dexterity: Evaluating Visuomotor Skills and Offline Proxies for Dexterous Manipulation · 2026-10-29

Contact Imitation Interface: Transferring Dexterous Skills by Showing How Contact Unfolds · 2026-10-29

TRACE: Target-State Risk-Aligned Conditional Execution for Embodied Tracking · 2026-10-30

Cross-Embodiment Motion Planning by Coupling Body-Agnostic World Model with Differentiable Ego-Dynamics · 2026-10-30