

レプリカントの真実

レプリカントの真実 - フィジカルAIの新聞

スモレンスク （ロシア）
予報 · 2026-11-03 · MSK

朝 · 09:00
2.1°
曇り
風: 北 1.81 m/s · 38%

昼 · 12:00
3.2°
雨
風: 北 1.46 m/s · 33%

夕方 · 18:00
3.3°
曇り
風: 西 1.9 m/s · 0%

一日: 1.8–4.1 °C
降水確率: 38%

雨の降るあいだ、私たちにできるのは、人間であることを学ぶことだけ。

号 № 6 · 2026-11-03 rev. 1 ·

天気の出典: Open-Meteo · 取得日時: 2026-11-03 07:25 MSK

ロボット研究：データ、接触、安全性

五つの研究から、動作表現、動画、接触、安全確認がロボットの性能をどう変えるかを検討します。

ニュース

WorldBridge、動作ラベルのない動画で新規課題の成功率55.90%

だれかが物をたなに並べる動画を見たとします。でも、手をどの方向へ何センチ動かしたかは書かれていません。ロボットにとっても、映像だけから作業を覚えるのはむずかしいことです。ジュン・ツェンさんたちの研究チームは、練習したことのない作業でも、動画が助けになるかを調べました。大切なのは、見た場面と実際の手の動きをつなぐことです。

研究では、動画から次にどんな場面になるかを予想するしくみと、その場面に近づくための動きを考えるしくみを組み合わせました。これを動画からの学習と考えることができます。ただし、人の動きをそのままねするわけではありません。人とロボットでは手の形や動かせる場所がちがうため、ロボットに合った命令を別に決める必要があります。

新しい作業を行うコンピューター上の試験では、最初の成功率は25.07%でした。その作業の動画を追加すると55.90%になりました。さらに、動きの命令が付いたお手本を作業ごとに一つ加えると70.70%になりました。成功率とは、決められた回数のうち、うまくできた割合のことです。この数字は研究で使った試験の結果です。

WorldBridgeは、映像の続きを予測する部分と動作を作る部分を分け、構造化された注意機構でつなぎます。両者はフローマッチングで共同学習され、短期と長期の視覚記憶も使います。研究では、新しい課題の動画だけを追加する条件と、動作ラベル付きのお手本を一つ加える条件を分けました。後者では成功率が70.70%になりますが、学習に使える情報量が異なります。同じ条件の数字として扱うと、動画そのものの効果を誤って評価します。

だからといって、どんなロボットでも動画を一度見れば料理や家具の組み立てができるわけではありません。動画に写っていない力のかけ方や、初めて見る道具への対応には、まだ調べることがあります。今回わかったのは、動画だけでも新しい作業の練習に役立つ場合があるということです。

RoboTwin 2.0の全データ条件で報告された96.23%は、新規課題の比較とは別の試験設定です。異なるロボットの手やカメラ、実機への移行で同じ改善が得られるとは確認されていません。再現には動画の種類、動作ラベルの数、試験課題の分離を明記する必要があります。

出典

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

281種類の制御方針を実機のコネクター挿入で比較

通信ケーブルの小さな先を差しこみ口に入れるとき、向きが少しずれるだけで入りません。ロボットの五本指の手にも、これはむずかしい作業です。アマー・アッバスさんたちの研究チームは、手を動かすプログラムを比べました。実物のロボットを使い、およそ1700回の試験をしました。画面の中だけの実験ではありません。

研究者は七つの種類に分けた281通りのプログラムを調べました。カメラの情報をどう使うか、指の動きをどんな数字で表すかを変えました。同じ作業でも、情報のまとめ方によって結果が変わります。たとえば、物の位置を手から見た数字で表すのか、別の場所を基準にするのかで、プログラムの使いやすさがちがいます。

実際に差しこむ前に計算できる67種類の点数も調べました。そのうち、動きのなめらかさやまっすぐさに関係する五つだけが、実物での成功と強く結びつきました。それも一つの種類のプログラムの中だけでした。コンピューター

でよい点が出ても、実際の指が正しい場所に入るとはかぎらないのです。

点群を使う方式では、必要な領域だけを切り出すこと、ロボットの土台を基準に座標を表すこと、特徴量を正規化することが、それぞれ特定の比較で70から100ポイントの成功率改善につながりました。画像と指示から動作を出す方式では、相対移動より絶対位置を予測することで最大70ポイント改善しました。これらは条件の異なる試験結果なので、数値を足し合わせたり全方式へ一般化したりできません。七種類の制御方針を比べた点も重要です。

この研究で試したのは、決まった手と小さな差しこみ部品です。どんな物でも同じ方法が一番よいと決まったわけではありません。研究者は、画面での評価だけにたよらず、実物を使った試験が必要だと考えています。

実機を動かさずに計算できる67種類の指標のうち、成功と安定して相関したのは五つだけで、しかも一つの方式の中に限られました。予測した動きのなめらかさや直線性が有用でも、異なるモデルを同じ点数で選べるとは限りません。検証用データの損失だけで実機性能を決めるのは危険です。

出典

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

接触面を合わせる模倣学習、指の角度の単純なコピーから転換

コップを持つときは、指を曲げるだけでは足りません。ふれる場所が悪いと、コップがすべってしまいます。ジーシュエン・シューさんたちの研究チームは、物のどこに指がふれたかに注目しました。指先の近くにある小さなカメラの映像などを使い、人の手の動きをロボットに教える方法を考えました。

人の手とロボットの手は、指の長さも曲がり方もちがいます。だから、人の指の角度だけをそのままねしても、同じ物を持てるとはかぎりません。研究者は、まず人の指が物にふれた場所を記録しました。次に、ロボット自身の指と手首をどう動かせば、似た場所にふれられるかを計算しました。

六つの手と作業の組み合わせで、ほかの方法と比べました。研究チームによると、作業の最後の段階での平均の成功率は、少なくとも50ポイント高くなりました。たとえば20%から70%になると、50ポイントの増加です。ここでの数字は、研究で決めた作業の結果であり、どんな物でも持てるという意味ではありません。

この方法は、人の実演を記録する部分とロボットが学ぶ部分を分けています。前者では共通の接触形状と指先のカラー画像を使い、手の姿勢を移す前に接触面を合わせます。後者では接触面の相対的な軌道を上位で予測し、下位の制御部分が手首と指を動かします。ソフトウェアと機器の設計情報も公開されています。人の指の角度をそのまま再現するのではなく、物との接触関係を保存する考え方です。

まだ、いろいろな形の道具や、ぬれた物を持つ場合まで確かめたわけではありません。それでも、指がどこにふれるかを考えることは、見た目が同じ動きでも成功が変わる理由を知る手がかりになります。

50ポイント以上という数字は、六つの手と作業で測った最終段階の平均成功率です。作業全体や異なる材質でも同じ効果が出るとは限りません。初期の接近動作、指先カメラの校正、接触のずれは別々に調べる必要があります。

出典

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

TRACE、人の追跡中の衝突を3.96ポイント低減

人が多い通路で、一人の人についていくロボットを考えてみましょう。似た服の人が前を通ったら、だれを追えばよいか分からなくなるかもしれません。ミンヤン・リューさんたちの研究チームは、人を見失わないようにする学習と、安全な動きがいつも同じ方向に進むのかを調べました。

人を追い続けるように学習すると、追いかける作業は上手になりました。しかし、一部の試験では、ぶつかる回数も増えてしまいました。そこで研究者は、動く直前に危険を調べるしくみを加えました。目当ての人がどこにいるか、ほかの人とまちがえていないかなど、いくつもの手がかりを調べます。

危険が大きいと判断したときだけ、次に前へ進む長さを短くします。向きを変える命令はそのままです。試験では、98%をこえる動きが元のままでした。それでも、まぎらわしい人がいる試験で、ぶつかる割合は3.96ポイント下がりました。いつでもゆっくり動かすだけでは、同じ改善は出ませんでした。

研究チームは、画像と言葉から動作を作る小型モデルの計画部分だけを追加学習し、基礎部分を固定しました。別の危険評価器には、目標候補、人物の見分けやすさ、距離と方向の推定、予定する命令など84種類の手がかりを入れます。危険がしきい値をこえると、次の目標位置までの前進距離だけを短くし、回転は変えません。常に減速させるのではなく、危険が高い瞬間に限って介入する設計です。

これは、どんな駅や店でも安全に人を追えると証明したわけではありません。研究で決めた場面と条件での結果です。だれかを見つけるのが上手になることと、ぶつからずに動くことは、別々に確かめる必要があります。

同時に行った再試験でも衝突率は3.77ポイント下がり、対照の差は0.29ポイントでした。ただし、あらゆる環境で衝突しないという保証ではありません。目標人物の状態を考えずに一律に遅くするだけでは同じ効果が出なかった点が

、単純な速度制限との違いです。

出典

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

Hourglass、七台・五種類の移動方式に共通の経路計画

小さな車輪のロボットと、大きくて曲がり方がちがうロボットを、同じせまい通路で走らせるとします。「右に曲がる」という命令だけでは、二台が同じように進むとはかぎりません。ジードン・ワンさんたちの研究チームは、形や動き方がちがうロボットに、一つの道案内のしくみを使えないか考えました。

研究者は、動いた後に周りの景色がどう変わるかを予想するモデルを使いました。これを世界モデルと呼びます。でも、景色の予想だけでは十分ではありません。それぞれのロボットが、どれくらい曲がれるか、どう進むかという情報も組み合わせました。道が空いていても、ロボットがそこへ曲がれなければ通れないからです。

コンピューターの試験では、五つの動き方に分かれる七台のロボットを使いました。新しい方法は、比べた一部の方法より、成功する割合が3.8から10.6ポイント高くなりました。さらに研究者は、実物のロボットを三分間だけ人が動かして、その動き方を調べました。その後、初めての室内で道を進めるか試しました。

共通部分をつなぐのは、機体が動くことで生じるカメラ視点の変化です。世界モデルは異なる機体のデータから学び、それぞれの機体には物理パラメーターを調べた簡略な運動モデルを用意します。経路を予測するとき、命令をまず機体の運動モデルに通し、その結果を世界モデルへ渡します。これにより、曲がり方や加速の違いを考えながら、一つの制御方針を使えます。機体ごとに別々の方針を一から学ぶ必要を減らす設計です。

ただし、階段やどんな乗り物にも使えると分かったわけではありません。研究は主に平らな地面での移動を考えています。ロボットの体のちがいを無視せず、同じ道案内のしくみを分け合えるかが研究の中心です。

実機では三分間の手動走行で機体の特徴を調べ、初めての室内で移動を試しました。ただし平らな地面という前提があり、階段や急な斜面での利用は確認されていません。数値上の改善は仮想環境での比較であり、実機試験とは分けて評価する必要があります。

出典

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

分析

データの表現が結果を変える

性能は学習だけでなく、座標系、接触情報、動作の制約にも左右されます。

異なる試験の数字は、機器とデータの条件をそろえずに直接比較できません。

出典

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

本号の用語集

この号のキーワードを解説します。

01 **WorldBridge**
動画予測とロボットの動作を結ぶモデル。

02 **RoboTwin 2.0**
ロボット操作を仮想環境で比較する試験課題集。

03 **動作表現**
位置や動きの命令を数値として表す方式。

04 **点群**
物の表面などを三次元の点の集まりで表したデータ。

05 **Contact Imitation Interface**
指先の接触形状を使って技能を移す方法。

06 **接触軌道**
指と物がふれる部分の位置の変化。

07 **TRACE**
人を追うロボットの動作を危険に応じて調整するフィルター。

08 **危険評価**
周囲の情報から危険の大きさを計算すること。

09 **Hourglass**
共通の経路計画と機体ごとの運動モデルを組み合わせる方式。

10 **世界モデル**

行動の後に周囲の様子がどう変わるか予測するモデル。

11 **RJ45コネクター**

五本指のロボットによる差しこみ試験で使った通信ケーブルの接続部品。

出典

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