フィジカルAI研究の潮流：基礎から融合領域まで

平塚謙良（京都大学 工学部 B4）

目次

1. イントロダクション
2. フィジカルAIの事例紹介
3. ロボット基盤モデル（VLA）
4. 今後の展開とまとめ

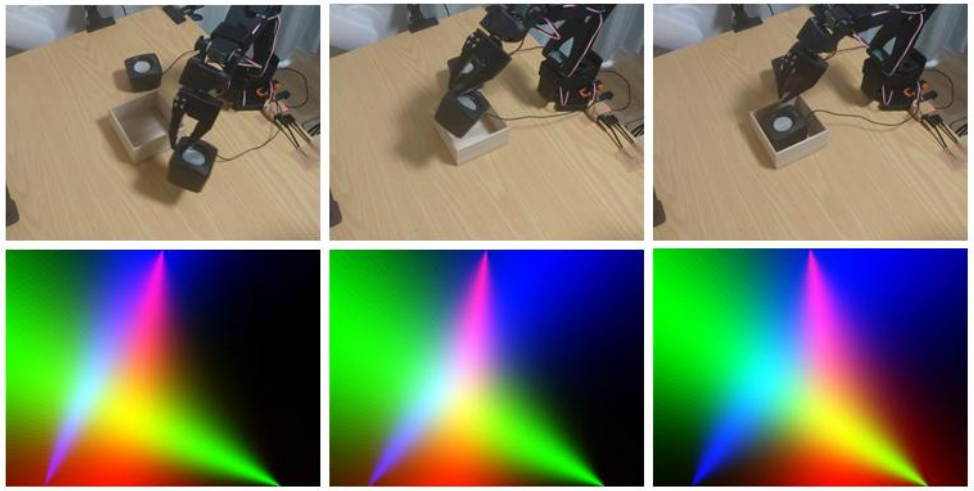
目次

自己紹介

氏名：平塚謙良（Kaneyoshi Hiratsuka）

研究領域：

- 強化学習
- 模倣学習
- ロボット聴覚



音環境情報に基づくピックアンドプレースの模倣学習

インターン等：

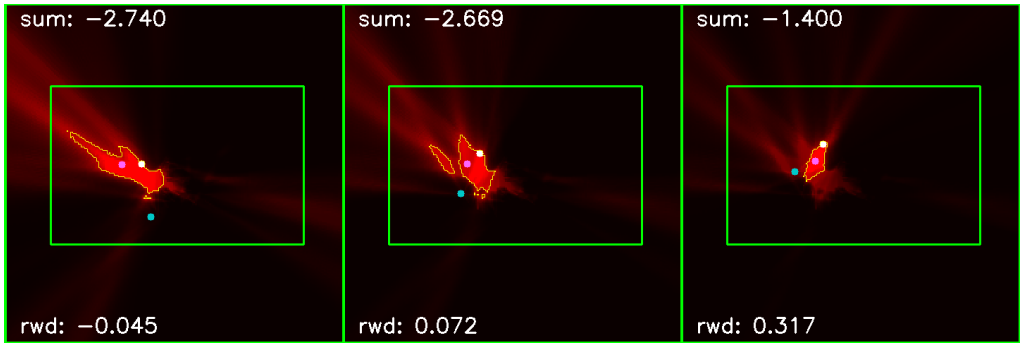
- 松尾研究所

→世界モデルやVLAの研究開発

- AIRoAコンペ

→VLAと強化学習

- Sony SSUP



世界モデルベースの深層強化学習による音源追跡の検討

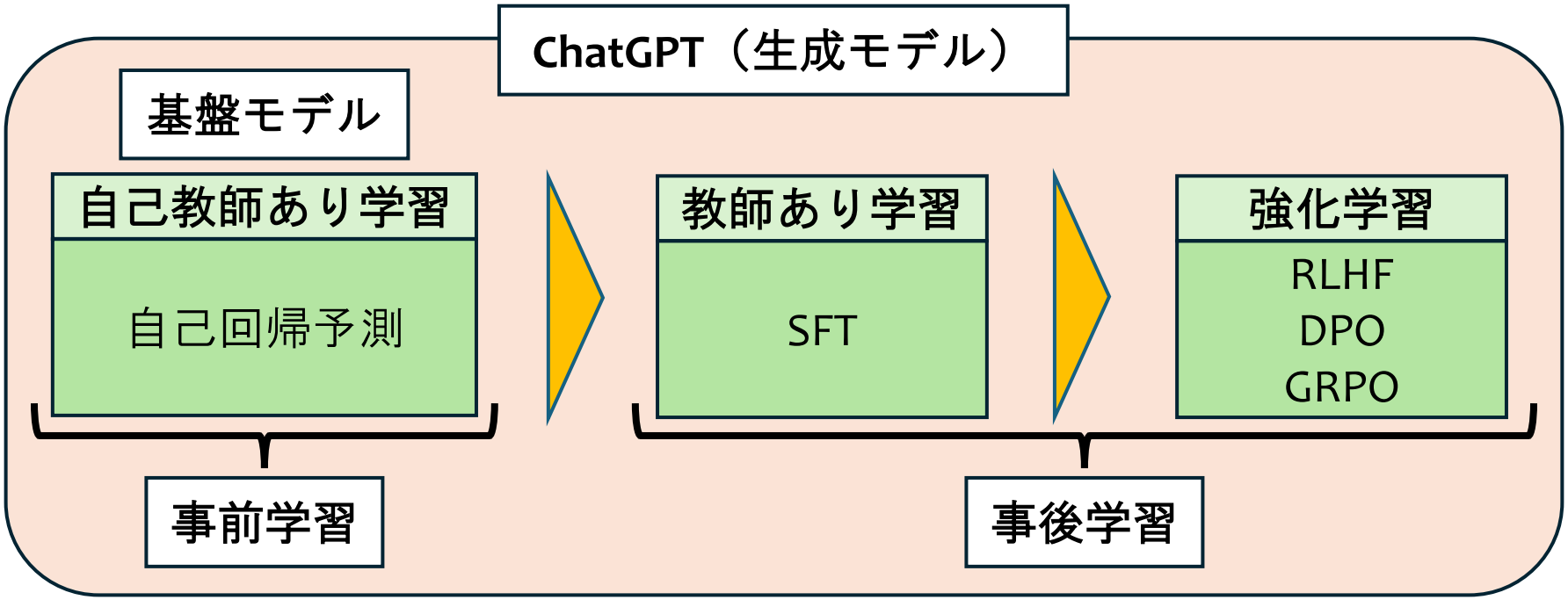
ChatGPTにみるAIの主要概念

- ChatGPTはテキストを生成する**生成モデル・生成AI**の一種
- 事前学習：

大規模データを使って**自己教師あり学習**し，幅広く知識を獲得 (**基盤モデル**)

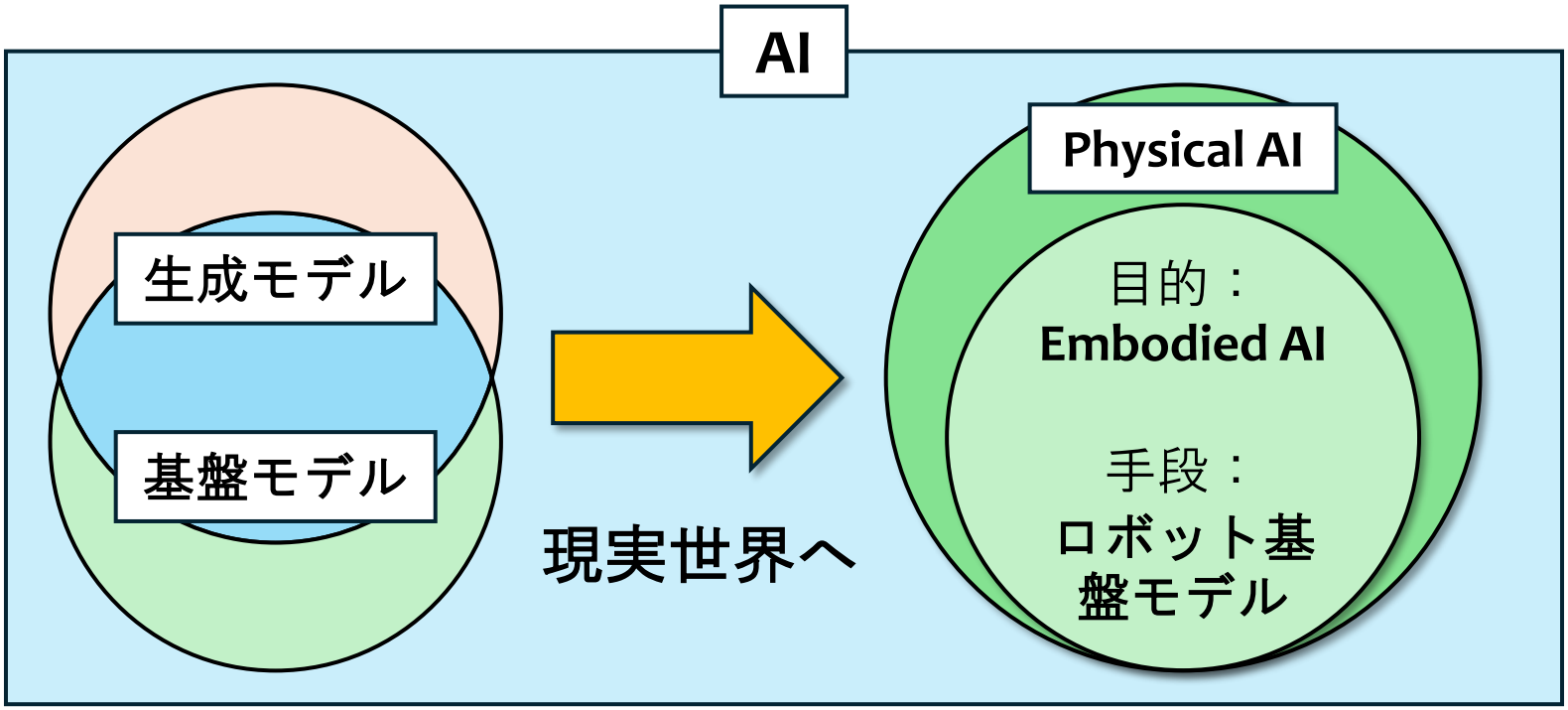
- 事後学習：

基盤モデルを**教師あり学習**や**強化学習**で使いやすく調整



概念の整理

- **生成モデル / AI**：新しいコンテンツを生成するAI
- **基盤モデル**：大規模なデータで事前学習され，特定機能を持つモデルの基盤となる
- **Embodied AI**：身体性を持って環境に作用できるAI. その手段の１つとしてのロボット基盤モデル.
- **Physical AI**：**現実**世界で知的・自律的に振る舞うシステム全般



なぜ今Physical AIなのか

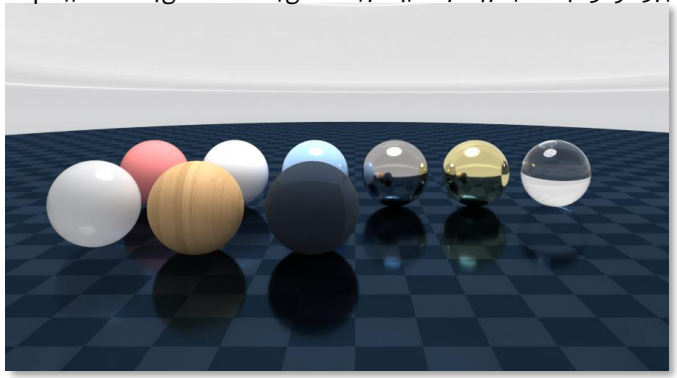
技術的成熟

アーキテクチャや学習法の進化：
- 大規模，高精度のモデルの登場

ハードウェアの進化：
- 高性能なGPU
- 安価なセンサーやアクチュエータ

シミュレータの進化：
- 高速に動作
- 物理的に忠実

https://www.4gamer.net/games/784/G078478/20240319057/



https://genesis-world.readthedocs.io/en/latest/user_guide/getting_started/visualization.html

社会的要求

労働力不足
社会インフラの老朽化

https://www.figure.ai/news/helix-loads-the-dishwasher

目次

事例１：多数の対象の制御

人流制御

対象: 予測不能な群衆

アプローチ: 間接的誘導

駅や空港，スタジアムのセンサー情報を基に，AIが混雑を予測

→空いているルートへ誘導

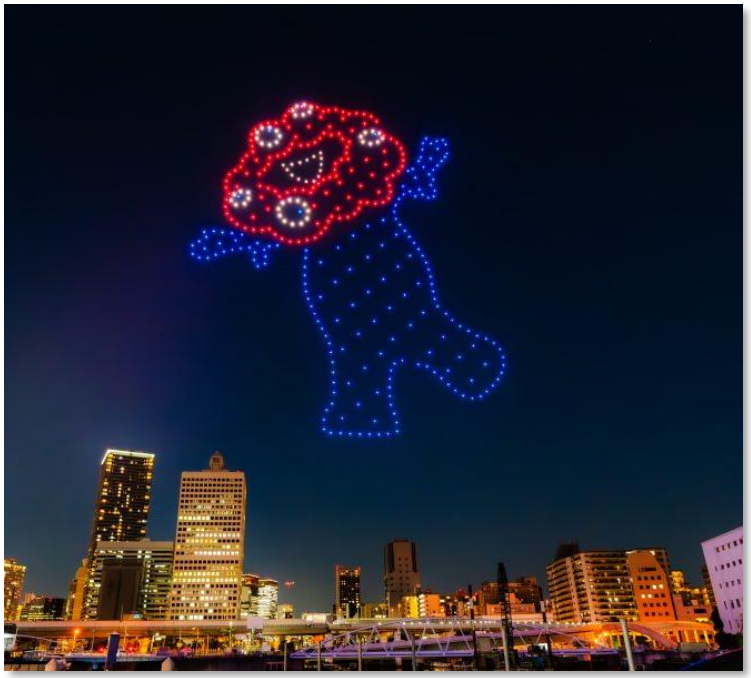
https://www.mlit.go.jp/common/001355040.pdf

ドローンショー

対象: 制御可能なドローン群

アプローチ: 直接的制御

中央集権的なシステムが計算した飛行経路を各機が正確に追従



https://www.expo2025.or.jp/news/news-20240515-02/

事例２：電力グリッド制御

課題: 再生可能エネルギーの増加

→電力の需給バランスが不安定になり，大規模停電のリスク

解決策：スマートグリッド

発電所，家庭，工場などの
データを収集

AIで分析

- 気象・過去データ→**需要予測**
- **需給バランスの最適化**
  - 発電・蓄電制御

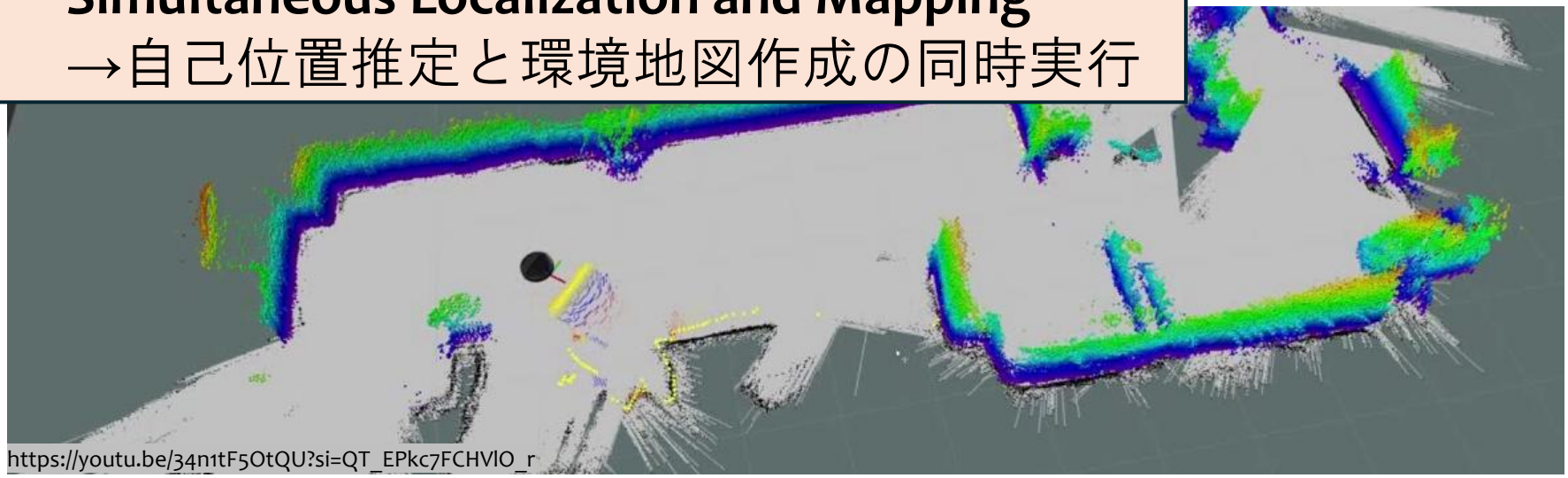
http://www.emhousing.net/taiyoko.html

事例３： 移動ロボットとSLAM

SLAMとは？

Simultaneous Localization and Mapping

→自己位置推定と環境地図作成の同時実行



https://youtu.be/34n1tF5OtQU?si=QT_EPkc7FCHVlO_r

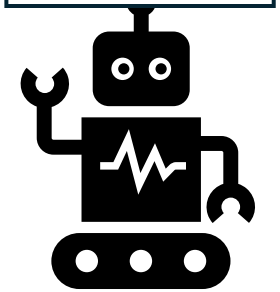
カメラやレーザーセンサーで周囲の環境の「目印」を捉え、自身が移動した時の見え方の変化から、自己位置と環境の地図を同時に推定

https://www.rentio.jp/matome/2019/01/roomba-compare/

応用先:

- **掃除ロボット:** 部屋の地図を作成し効率的に清掃
- **自動運転:** GPSがなくても正確な位置を維持
- **AR/VR:** 仮想オブジェクトを現実空間に配置

事例４：強化学習（RL）



Agent

試行錯誤

行動

状態，報酬

環境

マルコフ性
最適方策の存在を保障

現在の**状態**と**行動**のみに依存して，次の時刻の**状態**が決まる

メリット

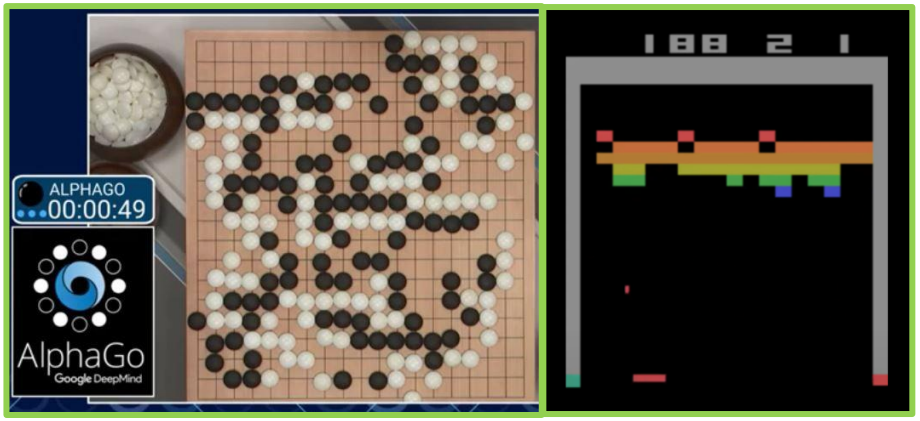
- 教師データが不要
- 人間を超えるような効率的な行動を発見できる

デメリット

- サンプル効率が悪い
- 報酬設計が難しい
- 学習が不安定

Sim2Real

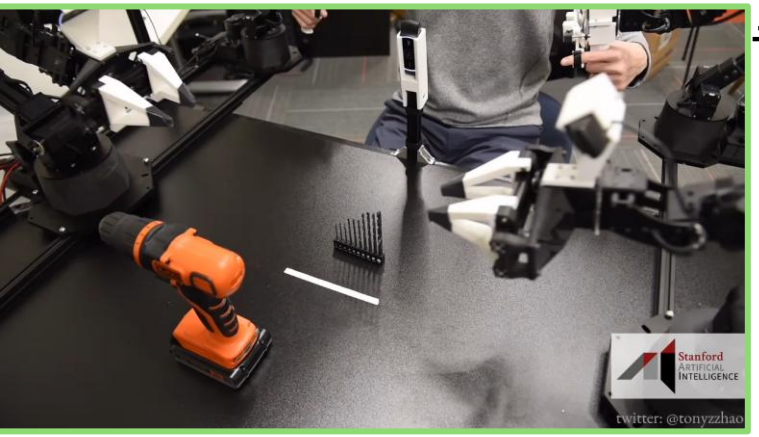
ゲーム領域での成功

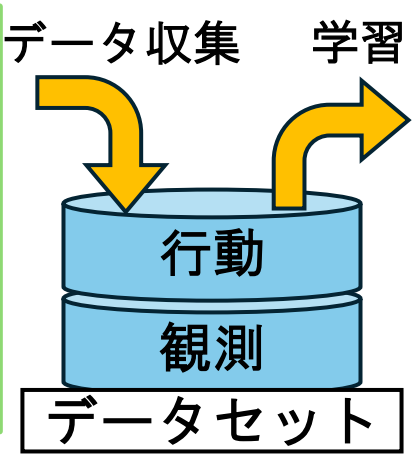


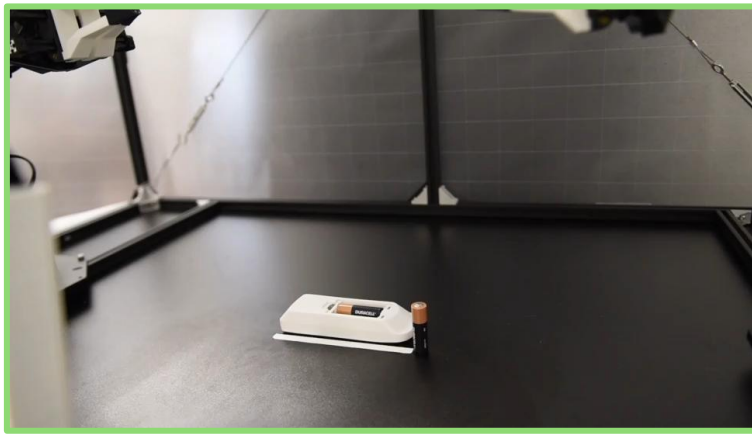
実世界応用へ

http://dimensionless.in/reinforcement-learning-super-mario-alphago/ https://github.com/pierclgr/Atari-DQN https://youtu.be/29xLWhqME2Q?si=iysKy9P9wbBZbUaA

事例５：模倣学習 (IL)







https://aloha-unleashed.github.io/

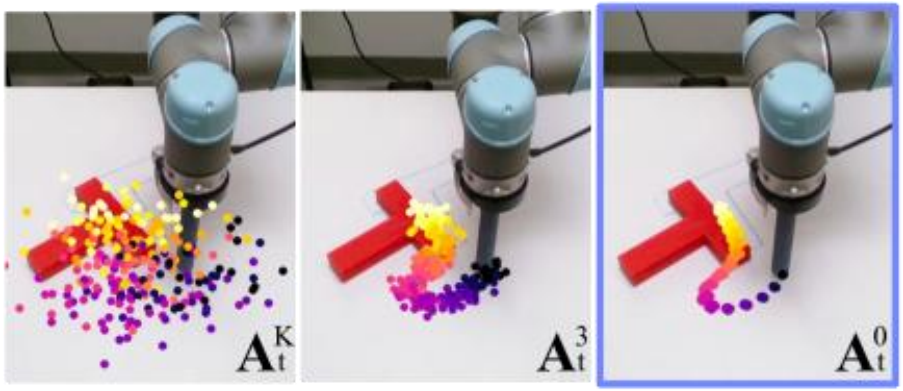
メリット

- データを集めれば学習可能
- 評価関数や報酬関数の設計が不要

デメリット

- 外乱に対して脆弱
  →**OOD問題**が顕著

未知の状況でモデルが適切に動作しないこと



[Cheng+ 2023]

Diffusion Policy：

画像生成で成功を収めている拡散モデルを模倣学習に応用.

現在の観測を条件づけて，ノイズから行動を生成.

事例６：世界モデル

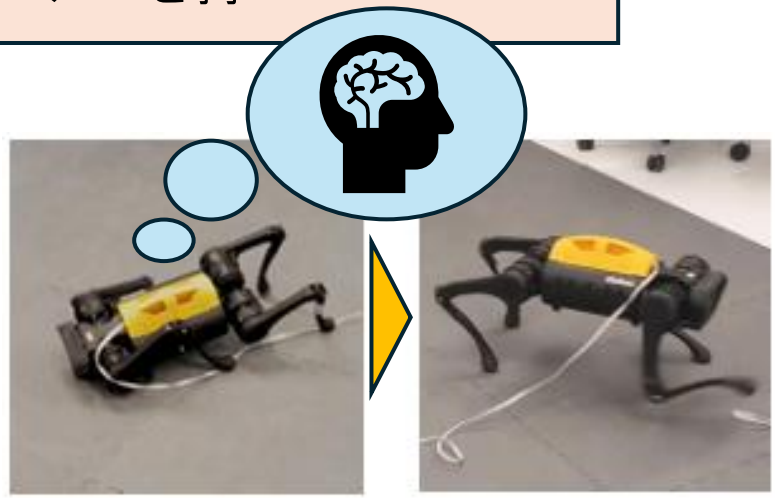
世界モデルとは？

コンセプト：環境がどう変化するかを予測するシミュレータを
AI内部に構築
共通項：内部に状態遷移モデルを持つ

強化学習 × 世界モデル

- **DayDreamer**

世界モデルを用いて，学習したり行動を計画することで，通常よりもはるかに短い１時間程度で歩行動作を学習



動画生成モデル × 世界モデル

- **Nvidia Cosmos Autoregressive**

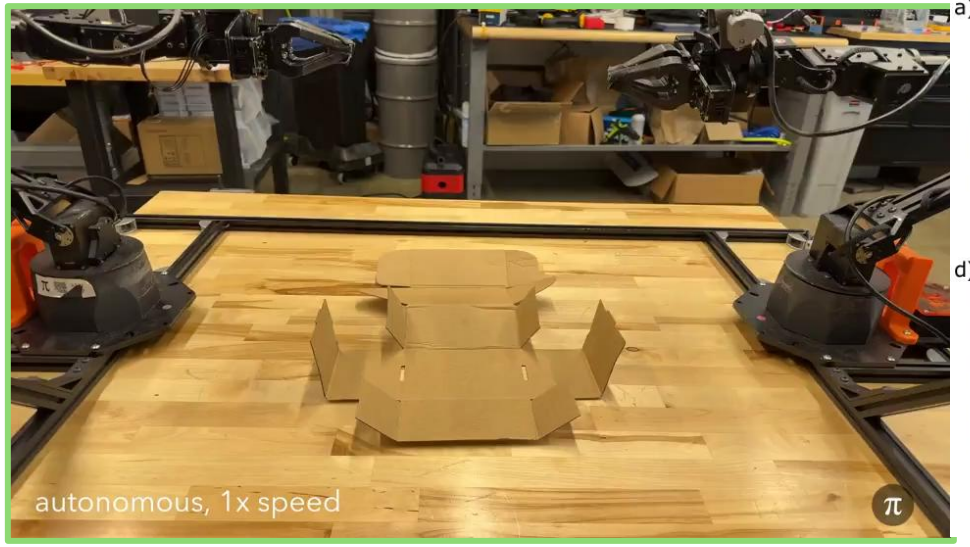
状態遷移モデルを内部に持つことで，より物理法則に忠実な動画生成を可能に

事例７：ロボット基盤モデル

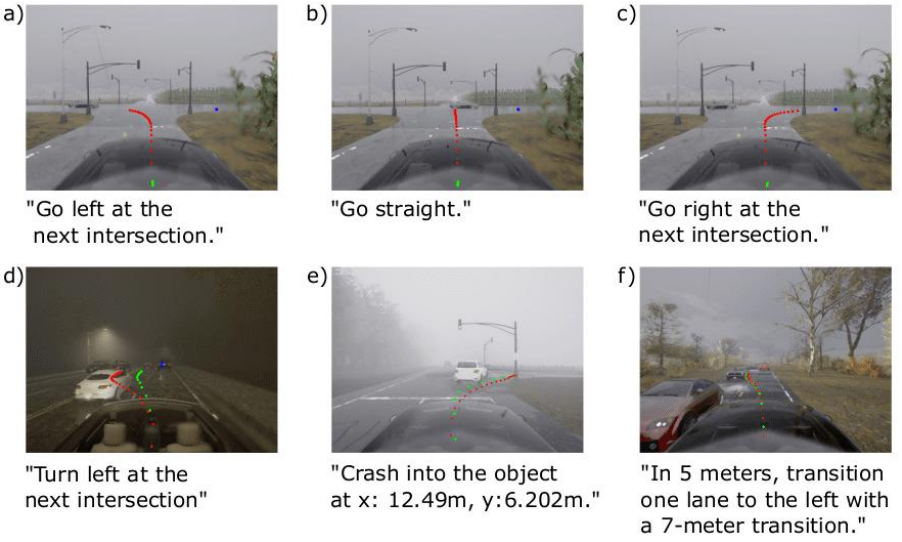
Vision-Language-Action (VLA) モデル：

- 視覚，言語，行動を単一のEnd-to-Endモデルに統一
- Webスケールで事前学習されたVLMの知識をロボット制御に活用

→汎化性能，マルチタスク性能



ロボットアームの制御（π0）



自動運転への応用
（SimLingo [Katrin+ 2025]）

https://www.physicalintelligence.company/blog/pio

目次

VLA：RT-2 [DeepMind 2022]

RT-2の革新性

1. ロボットのアームの角度変化などを，数値ではなくテキストトークンとして表現

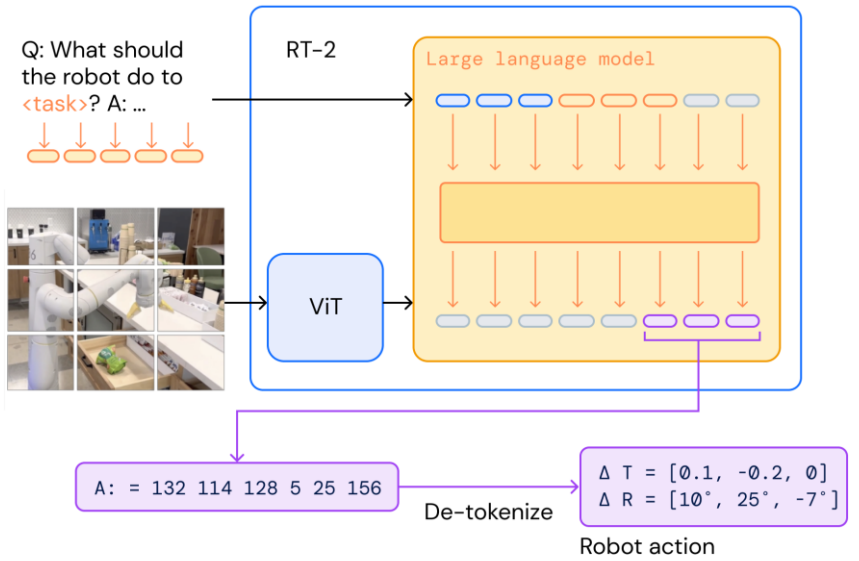
→VLMの構造を変更することなく，ロボットの行動を学習可能に

2. Webスケールデータで事前学習されたVLMをバックボーンとして利用

→高度な意味理解：常識に基づく推論が可能に

「疲れている人向けの飲み物は？」
→エナジードリンクを運ぶ

→思考の連鎖(Chain of Thought)：行動計画を立てて行動可能に



https://robotics-transformer2.github.io/

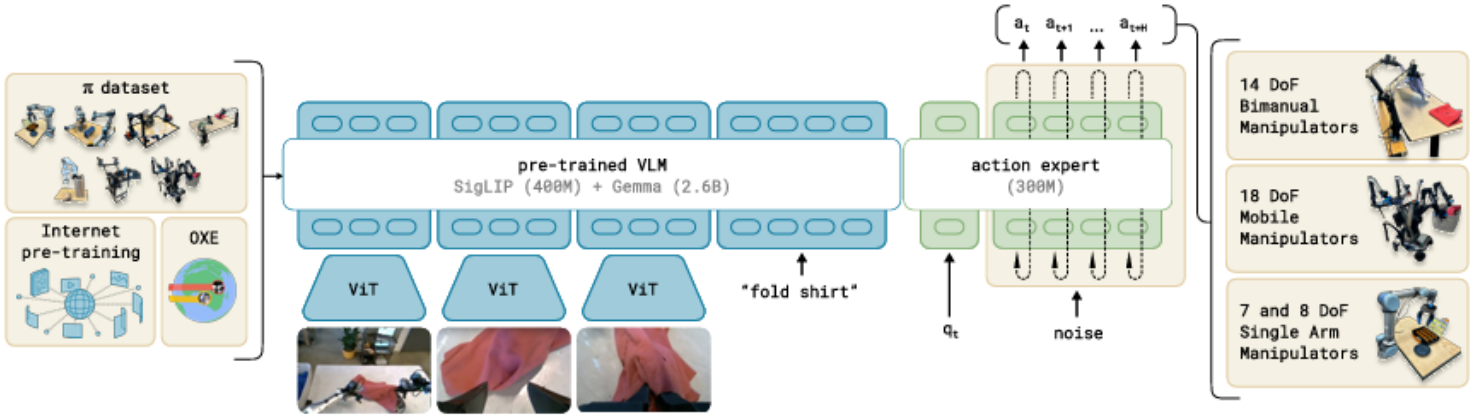
マルチタスク性能や汎化性能が飛躍的に向上し，ロボット基盤モデルへの道筋を示した

VLA：π0 & π0.5 [Physical Intelligence 2024, 2025]

π0 の革新性：より滑らかな動きへ

1. **Action Chunking**
   1回の推論で複数ステップ分の行動を出力
   →時間的に一貫した行動を高い動作周期で出力

2. **Flow Matching**
   行動を離散トークンではなく、**滑らかな連続軌道**として生成



全体的なアーキテクチャの概要 [Kevin+ 2024]

π0.5の進化：

- **階層的制御:** 高レベルな指示を**サブタスクに分解**して順に実行

例：「部屋を掃除して」→「机の上のごみを捨てる」＋・・・

オープンロボットデータセット

VLAの学習には大量のロボットデータセットが必要だが，収集コストが高く，研究機関や企業単体で用意するのは困難

⟹ **オープンロボットデータセットが重要**

Open X-Embodiment (OXE)
- 世界21機関が協力
- 22種類のロボット
- 100万軌道を超えるデータを集約

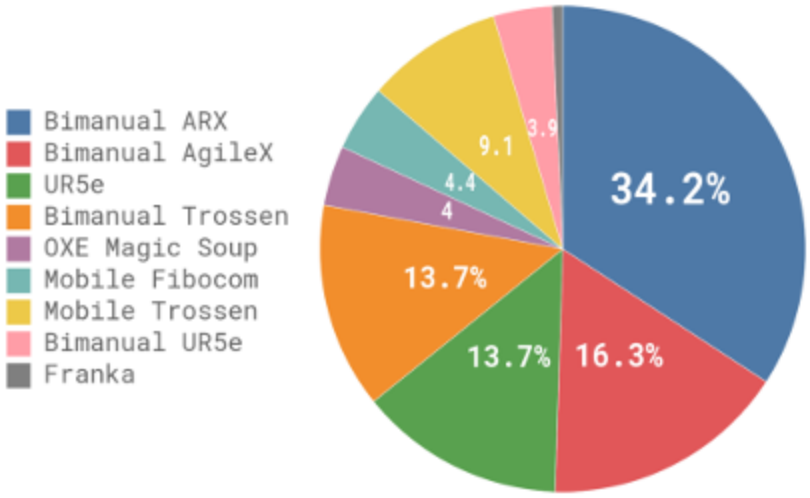
- RoboMIND
- RLDS
- AgiBot-World

- Bridge
- DROID
- LIBERO

日本の取り組み
- AIST-Bimanual Manipulation (産総研)
- AIロボット協会 (AIRoA)

⟹ 複数の異なるロボットで集められたデータ

データの多様性がロボット基盤モデルの実現には重要



$\pi 0$は学習にさまざまなロボットデータを利用
[Kevin+ 2024]

目次

LLM技術の恩恵とVLAの制約

LLMの技術革新

新アーキテクチャ:

- Mixture of Experts (MoE)
  →入力に応じて利用する内部モデルを切り替える
- State Space Models (SSM)
  →Transformerより高速な推論

新学習手法:

- GRPO
  →強化学習により推論能力を強化

マルチモーダル化：

- 画像出力，音声入出力

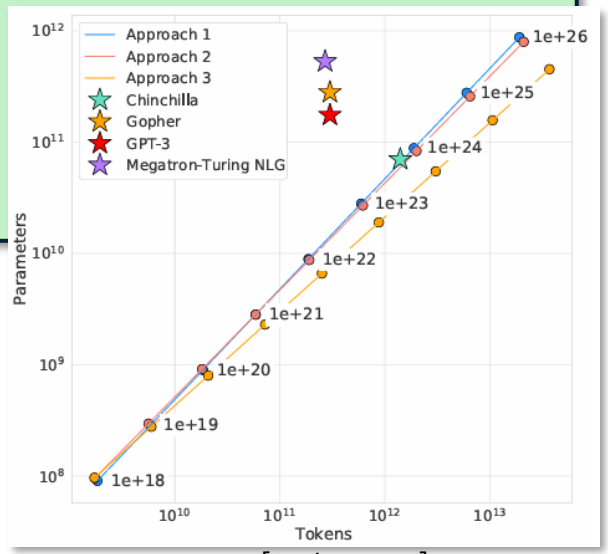
VLA特有の制約

リアルタイム性:

- ロボット制御には低遅延の推論が必須
  →LLMのようにonサーバの推論は難しい

データ不足:

- Chinchilla則
  →性能最適化にはモデルサイズとデータ量のバランスが重要



[Jordan+ 2022]

今後の展開

- モデルは大規模化しない
- VLAの強化学習→VLA-RL, DSRL
- VLAのマルチモーダル化

今後の展開：マルチモーダル化

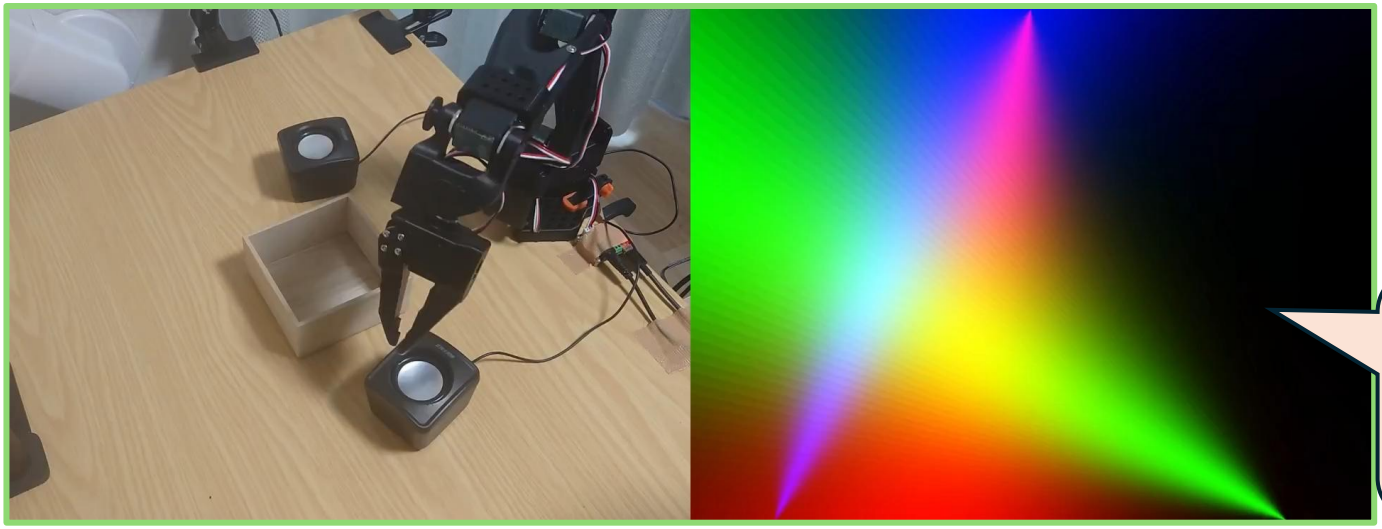
現状：視覚＋言語が中心

次のステップ：より豊かな感覚情報を統合

- **画像出力 (未来予測): WorldVLA** [Jun+ 2025]
  世界モデルとVLAを組み合わせて学習し，
  物理法則を理解した行動を生成
- **触覚入力:** 物理的な接触が重要なタスク
  例：柔らかいものを掴む
- **聴覚入力:** 音が手がかりになるタスク

Prediction

Action



マイクロフォンアレイの信号を可視化

聴覚情報を活用し、ピックアンドプレースの成功率向上を目指す

まとめ

本日の総括:

- **技術の成熟**と**社会の需要**が交差し，Physical AIがより重要に
- さまざまな技術により，AIが現実世界で動作可能に
- ロボット基盤モデルはLLM技術などの活用により急速に発展

日本の勝ち筋：

- LLM開発は莫大な計算資源を持つ巨大IT企業の戦い
- VLA開発は計算資源より高品質なデータが重要

→現場を持つ企業がデータ競争で優位に

	LLM	VLA
プレイヤー	巨大IT企業	現場を持つ企業
主戦場	計算資源	実世界データ

Physical AI，特にロボット基盤モデルは日本にとって大きな可能性を秘めている

質疑応答

ご清聴ありがとうございました