SSII30周年記念技術マップ-ロボティクス基盤モデル

元田智大（産総研）、牧原 昂志（大阪大学）

把持モデル
把持位置や成功可否の推定

GQCNN [RSS'17]
GG-CNN [RSS'18]
GraspNet [CVPR'20]
Contact-GraspNet [ICRA'21]

把持データセット
画像や言語と把持のラベルをペアにしたもの

DexNet [RSS'17]
Jacquard [IROS'18]
Cornell [IJRR'14]
Supersizing Self-supervision [ICRA'16]
GraspNet-1Billion [CVPR'20]
OCID-grasp [ICRA'21]

4DOF　3次元位置と手首周りの回転
6DOF　3次元位置と姿勢

データ収集ツール
ロボットの視覚、触覚など多様なセンサーから得られるデータを収集する。近年、オープンソースかつ低コストになり、扱いやすくなった*

AnyTeleop [RSS'23]
GELLO [arXiv'23]
ALOHA [RSS'23]
Mobile ALOHA [arXiv'24]
ALOHA Stationary [arXiv'24]
UMI [RSS'24]

Leader-follower
同じ構造のロボットを直接操作する

* 入手可能なツールにより小規模研究機関もデータ拡張に参加可能に。

Language
物体の名前や位置，Instructionを付与している。

OCID-VLG [CoRL'23]
Grasp-Anything [ICRA'24]

行動データセット
デモンストレーションから取得したロボット軌道データセット

MIME [CoRL'18]
Grasping in-the-wild [RA-L'20]
BridgeData v1 [arXiv'21]
RT-1 [RSS'23]
FurnitureBench [RSS'23]
RoboSet [arXiv'23]
DROID [arXiv'24]
RH20T [ICRA'24]
OXE [ICRA'24]
BridgeData v2 [CoRL'23]
RoboVQA [ICRA'24]
AutoRT [ICRAWS'24]

Augmentation
Sim-to-Realやスケールアップが目的。

ArmFarm [IJRR'18]
QT-Opt [CoRL'18]
RoboNet [CoRL'19]
MT-Opt [arXiv'21]
BC-Z [PMLR'22]

Community
データセットを大規模化するコミュニティを形成。

Domain Randomization [IROS'17]
RL-CycleGAN [CVPR'20]
RetineGAN [ICRA'21]
ROSIE [RSS'23]
OPTIMUS [CoRL'23]

行動モデル
視覚情報（画像や映像）と言語情報（文章や音声）を組み合わせて理解し、特定の行動やタスクを実行するためのAIの一種。
特に「行動」を含む場合にロボット基盤モデルを呼ばれる。

Language-based
テキスト入力を条件に「行動」を出力する形式のモデルを指す。視覚情報と共に行動を出力するモデルはVLAと呼ばれ、Googleを始めとしたBigTech企業によって2022年以降トレンドである。

Code-based
高性能なLLM（大規模言語モデル）は、コンピュータ言語を記述可能であることが知られている。つまり、事前に固有の行動（スキル）のライブラリを設定すると、ある要求に対して適切なアルゴリズムを生成できる。

Goal-based
目標状態に向かうように学習された行動モデルを指す。主に事前に決められた環境の状況やロボットの終端姿勢に向かって行動する。

Pretraining
画像言語分野と同様にモデルの汎化性能を上げるための学習手法。
データコストが大きいロボットタスクには有効になる。

ACT [RSS'23]
Diffusion Policy [RSS'23]
Instruct2Act [arXiv'23]
RT-1 [RSS'23]
RoboAgent [arXiv'23]
RT-2 [CoRL'23]
RT-1-X RT-2-X [ICRA'24]
SARA-RT [ICRA'24]
RT-H [arXiv'24]
F3RM [CoRL'23]
DAA [arXiv'24]
CLIPort [CoRL'21]
SayCan [arXiv'22]
PaLM-E [arXiv'23]
Inner Mobologue [CoRL'23]
VoxPoser [CoRL'23]
KnowNo [CoRL'23]
RFM-1 ['24]***
Octo [ICRAWS'24]**
ChatGPT for Robotics [arXiv'23]
ProgPrompt [ICRA'23]
Code as Policy [ICRA'23]
Gato [TMLR'23]
VIMA [ICML'23]
UniPi [NeurIPS'23]
RoboCat [TMLR'24]
RT-Trajectory [ICLR'24]
RT-Sketch [arXiv'24]
MVP [CoRL'22]
R3M [CoRL'22]
RPT [CoRL'23]
GR-1 [ICLR'24]

** オープンソースで提供されたロボット基盤モデルの代表例。

*** 物流倉庫のデータによる、初の商用ロボティクス基盤モデル。

2014　2020　2021　2022　2023　2024